Algebraic & Geometric Topology

Volume 18 (2018)

Issue 1 (pages 1–633)

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

MANAGING EDITOR

Colin Rourke
agt@msp.warwick.ac.uk
University of Warwick

BOARD OF EDITORS

Matthew Ando	University of Illinois mando@math.uiuc.edu	Tsuyoshi Kobayashi	Nara Women's University tsuyoshi09@gmail.com
Julie Bergner	University of Virginia jeb2md@eservices.virginia.edu	Christine Lescop	Université Joseph Fourier lescop@ujf-grenoble.fr
Joan Birman	Columbia University jb@math.columbia.edu	Robert Lipshitz	University of Oregon lipshitz@uoregon.edu
Steven Boyer	Université du Québec à Montréal cohf@math.rochester.edu	Dan Margalit	Georgia Institute of Technology margalit@math.gatech.edu
Fred Cohen	University of Rochester cohf@math.rochester.edu	Norihiko Minami	Nagoya Institute of Technology nori@nitech.ac.jp
Alexander Dranishnikov	University of Florida dranish@math.ufl.edu	Andrés Navas Flores	Universidad de Santiago de Chile andres.navas@usach.cl
Tobias Ekholm	Uppsala University, Sweden tobias.ekholm@math.uu.se	Robert Oliver	Université Paris-Nord bobol@math.univ-paris13.fr
Mario Eudave-Muñoz	Univ. Nacional Autónoma de México mario@matem.unam.mx	Luis Paris	Université de Bourgogne lparis@u-bourgogne.fr
David Futer	Temple University dfuter@temple.edu	Jérôme Scherer	École Polytech. Féd. de Lausanne jerome.scherer@epfl.ch
Soren Galatius	Stanford University galatius@math.stanford.edu	Peter Scott	University of Michigan pscott@umich.edu
John Greenlees	University of Sheffield j.greenlees@sheffield.ac.uk	Zoltán Szabó	Princeton University szabo@math.princeton.edu
J. Elisenda Grigsby	Boston College grigsbyj@bc.edu	Ulrike Tillmann	Oxford University tillmann@maths.ox.ac.uk
Ian Hambleton	McMaster University ian@math.mcmaster.ca	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		

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

The subscription price for 2018 is US $465/year for the electronic version, and $730/year (+ $60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. 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) at Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published 6 times per year and continuously online. Periodical rate postage paid at Berkeley, CA 94704, 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

http://msp.org/

© 2018 Mathematical Sciences Publishers

msp

Algebraic & Geometric Topology 18 (2018) 1–13

Width of a satellite knot and its companion

QILONG GUO
ZHENKUN LI

In this paper, we give a proof of a conjecture which says that $w(K) \geq n^2 w(J)$, where $w(\cdot)$ is the width of a knot, K is a satellite knot with J as its companion, and n is the winding number of the pattern. We also show that equality holds if K is a satellite knot with braid pattern.

57M25, 57M27

1 Introduction

Width is an important invariant of knots which was introduced by Gabai [2]. It gives rise to the notion of thin position (of knots), which is essentially used in Gabai's proof of property R [2] and Gordon and Luecke's proof of the knot complement conjecture [3], among others. We can view width as a kind of refinement of bridge number. It is an interesting question how those knot invariants behave under the operations of connected sum and taking satellites. For the bridge number, we know that $b(K_1 \# K_2) = b(K_1) + b(K_2) - 1$ and $b(K) \geq nb(J)$, given that K is a satellite knot with companion J and n is the wrapping number; see Schubert [6] and Schultens [7]. In the case of width, it is known that

$$w(K_1) + w(K_2) - 2 \geq w(K_1 \# K_2) \geq \max\{w(K_1), w(K_2)\},$$

and both bounds are tight. See Blair and Tomova [1], Rieck and Sedgwick [4] and Scharlemann and Schultens [5]. For width under taking satellites, it is conjectured that $w(K) \geq n^2 w(J)$, where n is the wrapping number, similar to the case of the bridge number. There is also a weak version conjecturing that $w(K) \geq n^2 w(J)$, where n is the winding number instead of the wrapping number. Zupan [8; 9] proves that $w(K) \geq 8n^2$, where n is the winding number, and $w(K) = q^2 w(J)$, where K is a (p, q)–cable knot with companion J and q acts as the winding number. Both of his results give partial positive answers to the weak version. In this paper, we give a complete positive answer to the weak version involving the winding number.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.1

Theorem 1.1 *Let K be a satellite knot with companion J, and suppose the winding number of the pattern is n. Then*

$$w(K) \geqslant n^2 w(J).$$

In Section 2, we introduce some basic concepts and construct a graph associated to the neighborhood of the companion; in Section 3, we prove that there is a simple loop in this graph, and such a loop is unique; in Section 4, we associate each knot with a word in Z_2, the free monoid of rank 2, and then use it to help calculate the width.

2 Preliminaries

First we introduce some basic definitions.

Definition 2.1 Suppose $\widehat{V}$ is a standard solid torus in S^3, and $\widehat{k}$ is a knot in $\operatorname{int}(\widehat{V})$ such that $\widehat{k}$ is not contained in any 3–ball $B \subset \widehat{V}$. Let $j \subset S^3$ be a nontrivial knot and let $V = N(j)$ be the closure of a tubular neighborhood of j in S^3. Let $f\colon \widehat{V} \to S^3$ be an embedding such that $f(\widehat{V}) = V$, and let $k = f(\widehat{k})$. Then k is called a *satellite knot* with *companion* j and *pattern* $\widehat{k}$. The *winding number* (of the pattern) is defined to be the algebraic intersection number (up to a sign) of the pattern with a meridian disk. Furthermore, if K (or $J, \widehat{K}$) is the knot type represented by k (or $j, \widehat{k}$), we could say that K is a satellite knot with companion J and pattern $\widehat{K}$ without ambiguity.

Let $\pi\colon \mathbb{R}^4 \to \mathbb{R}$ be the projection $(x_1, x_2, x_3, x_4) \mapsto x_4$, and regard S^3 as the unit sphere in $\mathbb{R}^4$. In the rest of this paper, we always assume that $h = \pi|_{S^3}$. Then h is a Morse function on S^3 with exactly two critical points. These two critical points are $h^{-1}(1)$ and $h^{-1}(-1)$, and we call them infinite (critical) points. For each $r \in (-1, 1)$, $h^{-1}(r)$ is obviously a 2–sphere, which is called a level sphere.

Definition 2.2 Let K be a knot type and let $\mathcal{K}$ be the set of all knots $k \in K$ such that

- k does not contain the two infinite points,
- $h|_k$ is Morse, and
- the critical points of $h|_k$ are in distinct levels.

For each $k \in \mathcal{K}$, suppose all the critical values of $h|_k$ are $c_1 < c_2 < \cdots < c_m$. Choose regular values $r_1, r_2, \ldots, r_{m-1}$ such that $c_i < r_i < c_{i+1}$ for $i = 1, 2, \ldots, m-1$, and

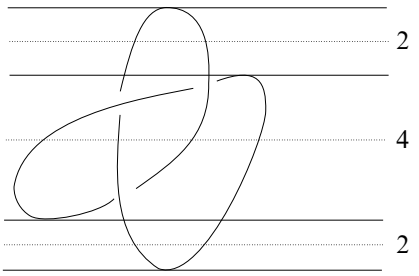


Figure 1: The width of the trefoil is 8.

let $\omega_i(k) = |k \cap h^{-1}(r_i)|$. Define

$$w(k) = \sum_{i=1}^{m-1} \omega_i(k) \quad \text{and} \quad w(K) = \min_{k \in \mathcal{K}} w(k).$$

We call $w(K)$ the *width* of the knot type K. See Figure 1 for the width of trefoil.

In this paper, we focus on "nice" solid tori defined as follows.

Definition 2.3 Let V be a solid torus in S^3. We say that V is *nice* if

- V does not contain the two infinite points,
- $h|_{\partial V}$ is also a Morse function, and
- all critical points of $h|_{\partial V}$ are in distinct levels.

If a solid torus contains a infinity point in its interior, then the first condition can be achieved by "digging them out"; ie pick an arc connecting the infinity point and a point on the boundary ∂V (disjoint from k, j), and remove a tubular neighborhood of the arc. Then we can modify the new boundary to satisfy the other two conditions and get a nice solid torus.

Definition 2.4 Let V be a nice solid torus in S^3. We construct a graph as follows. Let $c_1 < c_2 < \cdots < c_m$ be all critical values of $h|_{\partial V}$, and let $M = V - \left(\bigcup_{i=1}^{m} h^{-1}(c_i)\right)$. Then vertices of the graph correspond to connected components of M and edges correspond to components of $h^{-1}(c_i) \cap V$ for $i = 1, 2, \ldots, m$, which are not points. We require that two vertices v_1 and v_2 are connected by an edge if and only if the two corresponding components of M are separated by a component of $h^{-1}(c_i) \cap V$ which corresponds to the edge. We call this graph the *Reeb graph* and denote it by $\Gamma(V)$.

We need some results from [5]. In general, critical points of $h|_{\partial V}$ are classified as maximal, minimal and saddle points. Following [5], maximal (or minimal) points can be further divided into external and internal maximal (or minimal) points; saddle points can be divided into nested and unnested saddle points. We don't want to introduce the detailed definition here as they are not important for the use of this paper. Now the connectivity graph studied in [5] can be defined as follows.

Definition 2.5 Suppose $c_1', \dots, c_l'$ are all critical values of $h|_{\partial V}$ corresponding to all external maximal, external minimal and unnested saddle points, and let $M' = V - \left(\bigcup_{i=1}^{l} h^{-1}(c_i')\right)$. If we carry out the construction in Definition 2.4 using M' and c_i', then the graph we get is called the *connectivity graph*, and is denoted by $\Gamma_C(V)$.

We can see from the definition that M is obtained from M' by cutting off those critical levels containing internal maximal, internal minimal and nested saddle points. Then a component of M' is either unchanged or is cut off into a few components. For graphs, this corresponds to the fact that a vertex of $\Gamma_C(V)$ is either unchanged or replaced by some other graph which can be easily seen to be connected. Hence we have the following lemma:

Lemma 2.6 *Let V be a nice solid torus. If the connectivity graph $\Gamma_C(V)$ contains a loop, then Reeb graph $\Gamma(V)$ also has a loop.*

In this paper, we also need following results.

Lemma 2.7 [5] *For any knots K_1, K_2, we have $w(K_1 \# K_2) \geq \max\{w(K_1), w(K_2)\}$.*

Lemma 2.8 [5] *Let $V \subset S^3$ be a nice solid torus with boundary T. Then there is an new embedding $i\colon V \to S^3$ satisfying the following properties:*

- *$H = \overline{S^3 - i(V)}$ is a solid torus.*
- *The connectivity graph is a tree if and only if there is a meridian disk D of H such that $i^{-1}(\partial D) \subset T$ is horizontal in T under $h|_T$ (ie $i^{-1}(\partial D) \subset (h|_T)^{-1}(r)$ for some r, where r is a regular value of $h|_T$).*

Remark. Lemma 2.8 is a special case of Proposition 2.3 in [5].

We will also use the word "vertex" to refer to its corresponding connected component of M. It is not hard to see that each vertex has a product structure $P \times (c_i, c_{i+1})$,

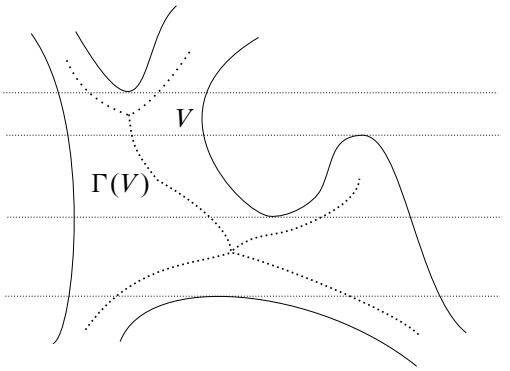


Figure 2: Embedding of $\Gamma(V)$ into V

where P is a horizontal planar surface (what we call a horizontal piece in Definition 3.5), and c_i, c_{i+1} are the two critical values that bound the vertex from below and above. We could embed $\Gamma(V)$ into $\mathrm{int}(V)$ as follows:

(1) Pick one point in the interior of each component of M.

(2) For two adjacent vertices, connecting the two points in the two vertices by a monotone decreasing arc in the interior of V.

Then $\Gamma(V)$ can be thought of as a 1–dimensional complex in $\mathring{V}$ as in Figure 2. The solid curves are the boundary of the solid torus V, and the interior of V is bounded by them. The dashed curves indicate the embedding of $\Gamma(V)$ into $\mathring{V}$.

3 $\Gamma(V)$ contains a unique simple loop

In this section, we always assume that $V \subset S^3$ is a nice solid torus and $\Gamma(V)$ is the graph constructed as in Definition 2.4. We need some preliminary results before proving that there is a unique loop in $\Gamma(V)$.

Lemma 3.1 *Suppose V is knotted in S^3, $T = \partial V$, and r is a regular value of $h|_T$ such that one component of $(h|_T)^{-1}(r)$ is an essential curve on T. Then at least one component X of $h^{-1}(r) \cap V$ is a surface with boundary such that exactly one boundary component δ is essential on T. Furthermore, δ is a meridian on T.*

Proof Since r is a regular value of $h|_T$, we have that $(h|_T)^{-1}(r)$ is a disjoint union of some simple closed curves. Note that $h^{-1}(r)$ is a sphere in S^3, so each component α of $(h|_T)^{-1}(r)$, which is essential in T, bounds two disks on $h^{-1}(r)$. We say α is *innermost* if α bounds a disk in $h^{-1}(r)$ which does not contain any other essential

curve of $(h|_T)^{-1}(r)$. By the innermost arguments, we can find a component P of $h^{-1}(r)-T$ such that only one component of ∂P is essential in T.

Claim $P \subset V$.

Proof of claim Suppose P is not contained in V. Then $P \cap \mathring{V} = \varnothing$. Let δ be the component of ∂P which is essential in T. Then any other component of P bounds a disk in T; hence δ bounds a singular disk in $\overline{S^3 - V}$. By Dehn's lemma, δ bounds an embedded disk in $\overline{S^3 - V}$, and thus T is compressible in $\overline{S^3 - V}$, which contradicts to the assumption that V is knotted. □

By similar argument as above, we can see that δ bounds an embedded disk in V; hence δ is a meridian of T. □

Corollary 3.2 *Suppose V is knotted and r is a regular value of $h|_T$. Suppose δ is a component of $(h|_T)^{-1}(r)$ which is essential in T. Then δ is a meridian of T.*

Corollary 3.3 *The graph $\Gamma(V)$ is not a tree if V is a knotted solid torus.*

Proof Suppose, on the contrary, that $\Gamma(V)$ is a tree. By Lemma 2.6, the connectivity graph defined in [5] for V is also a tree. Let $i\colon V \to S^3$ be the embedding as in Lemma 2.8. Then there is a meridian disk D of solid torus $H = \overline{S^3 - i(V)}$ such that $i^{-1}(\partial D) \subset (h|_T)^{-1}(r)$ for some r. Obviously, $i^{-1}(\partial D)$ is an essential curve in T, so by Corollary 3.2, $i^{-1}(\partial D)$ is a meridian of V; hence ∂D bounds a disk in $i(V)$, which contradicts that $S^3 = i(V) \cup H$. □

Definition 3.4 Let l be a simple loop of Γ. A vertex of l which is locally minimal (maximal) under h is called a *minimal* (*maximal*) *vertex*. We say that a vertex is a *critical vertex* if it is either minimal or maximal. A vertex which is neither minimal nor maximal is called a *vertical vertex*.

Definition 3.5 Let r be a regular value of $h|_T$. We call each component of $h^{-1}(r) \cap V$ a *horizontal piece*.

Let v be a vertex in l, regard v as a component of M (see the discussion at the end of Section 2) and let P be a horizontal piece in v. We want to describe the intersection of l with P. If v is a maximal vertex, then the two adjacent vertices in l are both below v.

Thus $l \cap v$ is an arc with one maximal point with respect to the height function h, and P intersects l in either zero, one or two points. Note that the intersection at one point is not transversal because if we move the piece slightly above or below, then the intersection would be zero or two points. A similar result holds for minimal vertices. If v is vertical, then $l \cap v$ is a monotonic arc, so every horizontal piece intersects l exactly once.

Another observation is that the two vertices adjacent to a critical vertex must both be vertical, so in any simple loop, vertical vertices must exist.

Lemma 3.6 *Let l be a simple loop in $\Gamma(V)$. Then, as a simple closed curve in V, we have that l (with any orientation) represents a generator of $H_1(V)$.*

Proof Let v be a vertical vertex of l, and pick a horizontal piece of v. Then P is a properly embedded surface in V and intersects l transversally once, so the algebraic intersection number of l and P is just ± 1; hence l must be a generator of $H_1(V)$. □

Proposition 3.7 *There is a unique simple loop in $\Gamma(V)$.*

Proof The existence of a loop follows from Corollary 3.3. To show the uniqueness, suppose, on the contrary, there are two different simple loops l_1, l_2. If l_1 and l_2 do not have the same vertical vertices, then there is a vertical vertex v with respect to one loop but not the other, say, with respect to l_1 but not l_2. Then pick a generic horizontal piece P in v and calculate the intersection number of l_2 and P. Since the geometric intersection number is even, the algebraic intersection number is also even and cannot be ± 1. But the algebraic intersection number of l_1 and P is ± 1. This contradicts the fact that both l_1 and l_2 are generators of $H_1(V)$.

Finally, observe that if the two simple loops have same vertical vertices, then they must be the same loop. We conclude that the simple loop must be unique. □

4 The inequalities

To calculate width, we use a technique coming from Zupan; see Section 5 in [9]. Let Z_2 be the free monoid generated by $\{a, b\}$, and let $\varphi\colon Z_2 \to \mathbb{Z}$ be a homomorphism such that $\varphi(a) = 2$, $\varphi(b) = -2$ and $\varphi(\alpha\beta) = \varphi(\alpha) + \varphi(\beta)$. For a word $x = \alpha_1\alpha_2\cdots\alpha_m \in Z_2$, where each α_j is a or b, write $x_i = \alpha_1\alpha_2\cdots\alpha_i \in Z_2$, and define

$$w(x) = \sum_{i=1}^{m} \varphi(x_i).$$

For each knot $k \in \mathcal{K}$ (see Definition 2.2), associate a word $x = x(k) \in Z_2$ to it as follows: suppose all the critical points of k, from the lowest to highest, are $p_1, p_2, \dots, p_m$; then define

$$x = x(k) = \alpha_1 \alpha_2 \cdots \alpha_m,$$

where $\alpha_i = a$ if p_i is a local minimal critical point and $\alpha_i = b$ if p_i is a local maximal critical point. Let $w_i(k)$ be defined as in Definition 2.2. It is not hard to see that $w_i(k) = \varphi(x_i)$ and

$$w(k) = \sum_{i=1}^{m} w_i(k) = \sum_{i=1}^{m} \varphi(x_i) = w(x).$$

Lemma 4.1 *Suppose $x = \alpha_1 \alpha_2 \cdots \alpha_m \in Z_2$ is a word.*

(i) *Suppose $\varphi(x_i) \geqslant 0$ for $i = 1, 2, \dots, m$. Let x' be a word obtained by deleting two letters α_i, α_j in F, where $i < j$, $\alpha_i = a$ and $\alpha_j = b$. Then $w(x) \geqslant w(x')$.*

(ii) *Suppose x' is obtained from x by exchanging two letters α_i, α_{i+1}, where $\alpha_i = a$ or $\alpha_{i+1} = b$. Then $w(x) \geqslant w(x')$.*

The proof is straightforward. We call the operation on words in (i) (or in (ii)) of this lemma the type I (or II) operation. The next lemma is useful when estimating $w(k)$:

Lemma 4.2 *Suppose n is a fixed positive integer, and $x = x(\widetilde{k})$ is a word, associated with a knot $\widetilde{k}$, which has the form $\omega_1 \alpha_1^{s_1} \omega_2 \alpha_2^{s_2} \cdots \omega_m \alpha_m^{s_m} \omega_{m+1}$, where $\omega_i = \beta_{i1} \beta_{i2} \cdots \beta_{it_i}$ is a word for $i = 1, 2, \dots, m+1$. Assume that each $s_i \geqslant n$.*

Furthermore, suppose that $x = \alpha_1 \cdots \alpha_m$ is the word associated to another knot $\widehat{l}$, and

(1) $\varphi(\omega_1 \alpha_1^{s_1} \cdots \omega_i \alpha_i^{s_i}) \geqslant n\varphi(\alpha_1 \cdots \alpha_i)$ *if* $\alpha_i = a$,

(2) $\varphi(\omega_1 \alpha_1^{s_1} \cdots \omega_{i-1} \alpha_{i-1}^{s_{i-1}} \omega_i) \geqslant n\varphi(\alpha_1 \cdots \alpha_{i-1})$ *if* $\alpha_i = b$.

Then we have

$$w(\widetilde{k}) \geqslant n^2 w(\alpha_1 \alpha_2 \cdots \alpha_m) = n^2 w(\widehat{l}).$$

Proof Suppose $1 \leqslant i \leqslant m$. If $\alpha_i = a$, then for $0 \leqslant j \leqslant n-1$, we have

$$\varphi(\omega_1 \alpha_1^{s_1} \cdots \omega_i \alpha_i^{s_i - j}) \geqslant n\varphi(\alpha_1 \cdots \alpha_i) - 2j;$$

if $\alpha_i = b$, then for $1 \leqslant j \leqslant n$, we have

$$\begin{aligned} \varphi(\omega_1 \alpha_1^{s_1} \cdots \omega_{i-1} \alpha_{i-1}^{s_{i-1}} \omega_i \alpha_i^{j}) &\geqslant n\varphi(\alpha_1 \cdots \alpha_{i-1}) - 2j \\ &= n(\varphi(\alpha_1 \cdots \alpha_i) + 2) - 2j \\ &= n\varphi(\alpha_1 \cdots \alpha_i) + 2n - 2j. \end{aligned}$$

Since the word comes from a knot, $\varphi(x_l) \geq 0$ for any l. Hence we have

$$\begin{aligned} w(\tilde{k}) &\geq \sum_{\alpha_i=a} \sum_{j=0}^{n-1} \varphi(\omega_1 \alpha_1^{s_1} \cdots \omega_i \alpha_i^{s_i-j}) + \sum_{\alpha_i=b} \sum_{j=1}^{n} \varphi(\omega_1 \alpha_1^{s_1} \cdots \omega_{i-1} \alpha_{i-1}^{s_{i-1}} \omega_i \alpha_i^{j}) \\ &\geq \sum_{i=1}^{m} n^2 \varphi(\alpha_1, \ldots, \alpha_i) + \sum_{\alpha_i=a} \sum_{j=0}^{n-1} (-2j) + \sum_{\alpha_i=b} \sum_{j=1}^{n} 2n - 2j \\ &= n^2 w(\alpha_1, \ldots, \alpha_n) - \sum_{\alpha_i=a} n(n-1) + \sum_{\alpha_i=b} n(n-1) \\ &= n^2 w(\alpha_1, \ldots, \alpha_n) \\ &= n^2 w(\hat{l}). \end{aligned}$$

□

Now suppose k is a satellite knot with companion j, and let V be a closed regular neighborhood of j that contains k. Without loss of generality, we can assume that V is a nice solid torus. Let l be the unique loop in $\Gamma(V)$ as in Proposition 3.7; then l can also be viewed as a knot in $\mathring{V} \subset S^3$. Denote by J (or K, L) the knot type of j (or k, l).

Lemma 4.3 $w(L) \geq w(J)$.

Proof Picking a horizontal piece P as in Lemma 3.1, and capping off all inessential boundaries of P near the boundary T, we get a meridian disk D of V such that $D \cap l = P \cap l$. By Section 3, l represents a generator of $H_1(V)$, so the algebraic intersection number of l and P is ± 1. Note also that l intersects P (and hence D) at most two points; see the discussion below Definition 3.5. So l must intersect the meridian disk D (transversally) only once and hence can be viewed as a composition of j and possibly another knot l'. Then by Lemma 2.7, $w(L) \geq w(J)$. □

Lemma 4.4 *Let P be a horizontal piece in a vertical vertex of l. Then the geometric intersection number of k and P is no less than the winding number of the pattern k.*

Proof Since the winding number for the pattern of k is n, and l represents a generator of $H_1(V) \cong \mathbb{Z}$, we have $[k] = \pm n[l] \in H_1(V)$. Since P is a vertical vertex, the algebraic intersection number (up to sign) of ± 1 and l is 1. Consequently, P and k must have algebraic intersection number (up to sign) $\pm n$, hence the geometric intersection number is at least n. □

Now we isotope l into an equivalent knot $\hat{l}$ and change k into another knot $\tilde{k}$ as follows.

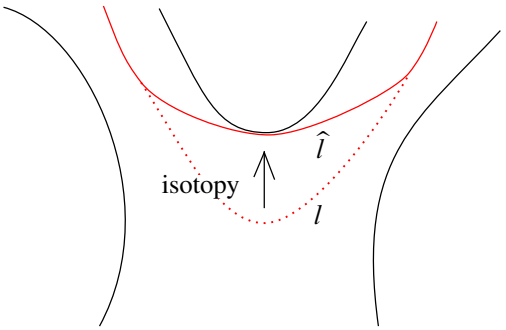


Figure 3: Isotopy between l and $\widehat{l}$

Suppose all the critical points of l are $q_1, q_2, \ldots, q_m$, from the lowest to the highest. Each q_j corresponds to a critical vertex of $\Gamma(V)$ and hence corresponds to a component of $M = V - \left(\bigcup_{i=1}^{n} h^{-1}(c_i)\right)$ (see Definition 2.4), denoted by C_j. When q_j is a local minimal point of l, suppose C_j is bounded from above by $h^{-1}(c_{i_j})$. Since v_j has a product structure, we can move q_j up to a point $\widehat{q}_j$ so that $\widehat{q}_j$ is a critical point of $h|_T$ and $h(\widehat{q}_j) = c_{i_j}$. Furthermore, we can assume that no more critical points of l are created. Do similar operations on local maximal points of l and after all such operations, l becomes a new knot $\widehat{l}$. Obviously, $\widehat{l}$ and l are equivalent knots, and all the critical points of $\widehat{l}$ are $\widehat{q}_1, \widehat{q}_2, \ldots, \widehat{q}_m$. See Figure 3 for the isotopy near a local minimal point of l.

Suppose $\widehat{q}_j$ is a local minimal point, then pick a regular value r_j slightly larger than c_{i_j} so that no other critical point of T, $\widehat{l}$, k lies between two level spheres $h^{-1}(r_j)$ and $h^{-1}(c_{i_j})$. There are two horizontal pieces P_j and Q_j on $h^{-1}(r_j)$ which respectively belong to the two vertical vertices adjacent to the vertex C_j. The pieces P_j and Q_j cut k into arcs, and each arc of $k - P_j - Q_j$ that lies in C_j is disjoint from any other C_t for $t \neq j$ since l is the unique simple loop. If β is such an arc intersecting C_j, we can create a new arc γ so that β and γ have the same end points, γ has exactly one inner critical point which is local minimal and γ is contained in $h^{-1}(c_{i_j}, r_j)$. Then replace β by γ and do this repeatedly until no arc of $k - P_j - Q_j$ intersects C_j. Then we finish the operation for a particular local minimal point of $\widehat{l}$. See Figure 4. Do similar replacements for local maximal points of $\widehat{l}$. After such replacement for all m critical points of $\widehat{l}$, we have that k becomes a new knot $\widetilde{k}$. The knots k and $\widetilde{k}$ may have different knot type, but it does not matter. We only need the following inequality.

Lemma 4.5 $w(k) \geqslant w(\widetilde{k})$.

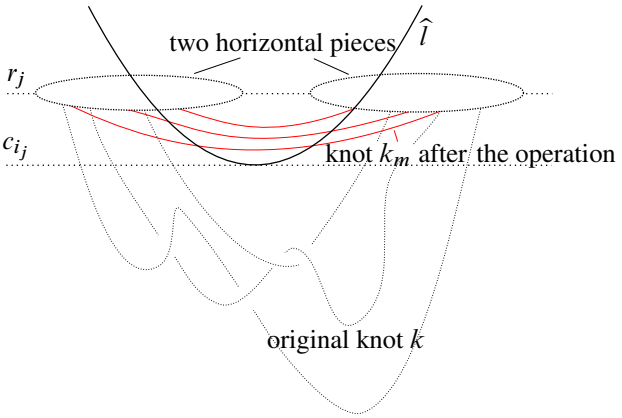


Figure 4: Operation on k

Proof We study how k becomes $\widetilde{k}$. Let $\widehat{q}_j$ be a local minimal point of $\widehat{l}$, and let P_j, Q_j be as above. Also let β be an arc in $k-P_j-Q_j$ which has end points in $P_i \cup Q_i$ and has interior below them. Then the operation of creating γ and replacing β can be done by two steps. the first step is to cancel pairs of maximal and minimal points of β to make β to have a unique critical point which is minimal. Since the interior of β is below its two end points, we can always pair a maximal point with another minimal point which is lower. This corresponds to the type I operation on words and, by Lemma 4.1, will not increase width. The condition that $\varphi(y_i) > 0$ in Lemma 4.1 holds because, after cancelling each pair of points, k still remains a knot. The second step is to lift the unique minimal point of β above the level c_{i_j}. This corresponds to type II operation on words and will not increase width. Similar arguments apply to local maximal points of $\widehat{l}$. □

Lemma 4.6 *$w(\widetilde{k}) \geq n^2 w(\widehat{l})$, where n is the winding number.*

Proof Now we need to estimate $w(\widetilde{k})$. The difficulty is that we do not know every critical point of $\widetilde{k}$ but only the ones near a critical point of $\widehat{l}$. For a local minimal point $\widehat{q}_j$ of $\widehat{l}$, let r_i be a regular value slightly larger than c_{i_j} as in the discussion above. Then $|h^{-1}(r_j) \cap \widehat{l}| = \omega_j(\widehat{l})$, so there are exactly $\omega_j(\widehat{l})$ horizontal pieces of V on $h^{-1}(r_j)$ which intersect $\widehat{l}$. By Lemma 4.4, we have

$$|h^{-1}(r_j) \cap \widetilde{k}| \geq n\omega_j(\widehat{l}).$$

A similar argument applies when $\widehat{q}_i$ is local maximal point. The difference is that we should pick a regular level r_i slightly lower than c_{i_j}, and hence

$$|h^{-1}(r_j) \cap \widetilde{k}| \geq n(\omega_{j-1}(\widehat{l})).$$

Note that ω_{j-1} appears on the right because we pick a regular level slightly lower than the critical level c_{i_j}. Compare Definition 2.2.

Suppose the word for $\widehat{l}$ is $\alpha_1\alpha_2\cdots\alpha_m$. Then the word for $\widetilde{k}$ can be written as $\omega_1\alpha_1^{s_1}\omega_2\alpha_2^{s_2}\cdots\omega_m\alpha_m^{s_m}\omega_{m+1}$, where ω_i is an arbitrary word for $i=1,2,\dots,m+1$ and $s_i \geqslant n$ for all $i=1,2,\dots,m$. The argument above shows that the words for $\widetilde{k}$ and $\widehat{l}$ satisfy the conditions of Lemma 4.2, and hence we have

$$w(\widetilde{k}) \geqslant n^2 w(\alpha_1\alpha_2\cdots\alpha_m) = n^2 w(\widehat{l}). \qquad \square$$

Theorem 4.7 *Let K be a satellite knot with companion J and winding number n. Then*

$$w(K) \geqslant n^2 w(J).$$

Proof Choose $k \in \mathcal{K}$ which realizes the width of its knot type. Without loss of generality, we can assume that V is a closed regular neighborhood of j which is nice and contains k. Furthermore, we could assume that all critical points of ∂V and k are in distinct levels. Construct $\Gamma(V)$ and pick the unique loop l by Proposition 3.7. Combining Lemmas 4.3, 4.5 and 4.6, we have

$$w(K) = w(k) \geqslant n^2 w(\widetilde{k}) \geqslant n^2 w(\widehat{l}) \geqslant n^2 w(L) \geqslant n^2 w(J). \qquad \square$$

Corollary 4.8 *Let K be a satellite knot with knotted companion J, and suppose the winding number of the pattern is n. If K has a braid pattern, then*

$$w(K) = n^2 w(J).$$

Proof Suppose j realizes the width of its knot type and has the associated word $x = \alpha_1\alpha_2\cdots\alpha_t$. Since k has a braid pattern, we can embed k so that the word associated to it is $y = \alpha_1^n\alpha_2^n\cdots\alpha_m^n$. By direct calculation, we have

$$w(K) \leqslant w(k) = w(y) = n^2 w(x) = n^2 w(j) = n^2 w(J).$$

Together with Theorem 4.7, we conclude that

$$w(K) = n^2 w(J). \qquad \square$$

Acknowledgements Guo is supported by NSFC (No. 11601519) and Science Foundation of China University of Petroleum, Beijing (No. 2462015YJRC034 and No. 2462015YQ0604).

References

[1] **R Blair**, **M Tomova**, *Width is not additive*, Geom. Topol. 17 (2013) 93–156 MR

[2] **D Gabai**, *Foliations and the topology of* 3*–manifolds, III*, J. Differential Geom. 26 (1987) 479–536 MR

[3] **C M Gordon**, **J Luecke**, *Knots are determined by their complements*, J. Amer. Math. Soc. 2 (1989) 371–415 MR

[4] **Y Rieck**, **E Sedgwick**, *Thin position for a connected sum of small knots*, Algebr. Geom. Topol. 2 (2002) 297–309 MR

[5] **M Scharlemann**, **J Schultens**, 3*–manifolds with planar presentations and the width of satellite knots*, Trans. Amer. Math. Soc. 358 (2006) 3781–3805 MR

[6] **H Schubert**, *Über eine numerische Knoteninvariante*, Math. Z. 61 (1954) 245–288 MR

[7] **J Schultens**, *Additivity of bridge numbers of knots*, Math. Proc. Cambridge Philos. Soc. 135 (2003) 539–544 MR

[8] **A Zupan**, *Properties of knots preserved by cabling*, Comm. Anal. Geom. 19 (2011) 541–562 MR

[9] **A Zupan**, *A lower bound on the width of satellite knots*, Topology Proc. 40 (2012) 179–188 MR

College of Science, China University of Petroleum-Beijing
Beijing, China

Department of Mathematics, Massachusetts Institute of Technology
Cambridge, MA, United States

guoqilong1984@hotmail.com, zhenkun@mit.edu

Received: 25 November 2014 Revised: 15 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 15–68

The closed-open string map for S^1–invariant Lagrangians

DMITRY TONKONOG

Given a monotone Lagrangian submanifold invariant under a loop of Hamiltonian diffeomorphisms, we compute a piece of the closed-open string map into the Hochschild cohomology of the Lagrangian which captures the homology class of the loop's orbit.

Our applications include split-generation and nonformality results for real Lagrangians in projective spaces and other toric varieties; a particularly basic example is that the equatorial circle on the 2–sphere carries a nonformal Fukaya A_∞ algebra in characteristic 2.

53D37, 53D40, 57R58; 53D45

1 Introduction

1.1 Overview of main results

Let X be a compact monotone symplectic manifold, $L \subset X$ a monotone Lagrangian submanifold, and $\mathbb{K}$ a field. We assume that L satisfies the usual conditions making its Floer theory well defined over $\mathbb{K}$, namely, L has Maslov index at least 2 and is oriented and spin if $\operatorname{char}\mathbb{K} \neq 2$. In this case, one can define a unital algebra over $\mathbb{K}$, the Floer cohomology $\mathrm{HF}^*(L,L)$, which is invariant under Hamiltonian isotopies of L. A larger amount of information about L is captured by the Fukaya A_∞ algebra of L, and given this A_∞ algebra, one can build another associative unital algebra called the Hochschild cohomology $\mathrm{HH}^*(L,L)$. There is the so-called (full) closed-open string map

$$\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(L,L),$$

which is a map of unital algebras, where $\mathrm{QH}^*(X)$ is the (small) quantum cohomology of X. This map is of major importance in symplectic topology, particularly in light of Abouzaid's split-generation criterion (see Abouzaid [1], Sheridan [45] and Ritter and Smith [38]), one of whose versions in the case $\operatorname{char}\mathbb{K} = 2$ says the following: if the closed-open map is *injective*, then L split-generates the w–summand $\mathcal{F}uk(X)_w$ of the Fukaya category, where $w = w(L) \in \mathbb{K}$ is the so-called obstruction number of L.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.15

(When $\operatorname{char}\mathbb{K} \neq 2$, the hypothesis can be weakened to say that $\mathcal{CO}^*$ is injective on a relevant eigensummand of $\mathrm{QH}^*(X)$; we will recall this later.)

Split-generation of the Fukaya category $\mathcal{F}uk(X)_w$ by a Lagrangian submanifold L is an algebraic phenomenon which has important geometric implications. For example, in this case, L must have nonempty intersection with any other monotone Lagrangian submanifold L' which is a nontrivial object in $\mathcal{F}uk(X)_w$, namely such that $\mathrm{HF}^*(L', L') \neq 0$ and $w(L') = w$. Another application, though not discussed here, is that split-generation results are used in proofs of homological mirror symmetry.

The present paper contributes new calculations of the closed-open map, motivated by the split-generation criterion and the general lack of explicit calculations known so far. (The closed-open map is defined by counting certain pseudoholomorphic disks with boundary on L, which makes it extremely hard to compute in general.)

There is a simplification of the full closed-open map, called the "zeroth-order" closed-open map, which is a unital algebra map

$$\mathcal{CO}^0\colon \mathrm{QH}^*(X) \to \mathrm{HF}^*(L, L).$$

It is the composition of $\mathcal{CO}^*$ with the canonical projection $\mathrm{HH}^*(L, L) \to \mathrm{HF}^*(L, L)$, and if $\mathcal{CO}^0$ is injective, so is $\mathcal{CO}^*$ (but not vice versa). Although $\mathcal{CO}^0$ generally carries less information than $\mathcal{CO}^*$, it is sometimes easier to compute. For example, we compute $\mathcal{CO}^0$ when L is the real locus of a complex toric Fano variety X; see Theorem 1.13. This map turns out to be noninjective in many cases, eg for $\mathbb{R}P^{2n+1} \subset \mathbb{C}P^{2n+1}$ over a characteristic-2 field. The aim of the present paper is to study the higher-order terms of the full closed-open map $\mathcal{CO}^*$, and to find examples when $\mathcal{CO}^*$ is injective but $\mathcal{CO}^0$ is not.

Specifically, let us consider the following setting: a loop γ of Hamiltonian symplectomorphisms preserves a Lagrangian L setwise. Let $S(\gamma) \in \mathrm{QH}^*(X)$ be the Seidel element of γ; then from Charette and Cornea [11], one can see that

$$\mathcal{CO}^0(S(\gamma)) = 1_L,$$

where 1_L denotes the unit in $\mathrm{HF}^*(L, L)$. Our main result, Theorem 1.7, is a tool for distinguishing $\mathcal{CO}^*(S(\gamma))$ from the Hochschild cohomology unit in $\mathrm{HH}^*(L, L)$; this way, it captures a nontrivial piece of the full closed-open map $\mathcal{CO}^*$ not seen by $\mathcal{CO}^0$. We apply Theorem 1.7 to show that $\mathcal{CO}^*$ is injective for some real Lagrangians in toric manifolds and also for monotone toric fibres which correspond to (non-Morse) A_2–type critical points of the Landau–Ginzburg superpotential.

After this paper had appeared, Evans and Lekili [19] proved split-generation for all orientable real toric Lagrangians, and all monotone toric fibres in zero characteristic, by completely different methods. They make use of the fact that these are homogeneous Lagrangians (ie they are orbits of Hamiltonian group actions), while we only use the fact these Lagrangians are invariant under certain Hamiltonian loops.

We will now mention our examples regarding real Lagrangians and postpone all discussion of monotone toric fibres, along with an introductory part, to Section 4.

Proposition 1.1 *Let $\mathbb{K}$ be a field of characteristic 2 and $\mathbb{R}P^n$ the standard real Lagrangian in $\mathbb{C}P^n$. Then $\mathcal{CO}^*\colon \mathrm{QH}^*(\mathbb{C}P^n) \to \mathrm{HH}^*(\mathbb{R}P^n, \mathbb{R}P^n)$ is injective for all n. In contrast, $\mathcal{CO}^0\colon \mathrm{QH}^*(\mathbb{C}P^n) \to \mathrm{HF}^*(\mathbb{R}P^n, \mathbb{R}P^n)$ is injective if and only if n is even.*

Corollary 1.2 *Over a field of characteristic 2, $\mathbb{R}P^n$ split-generates $\mathcal{F}uk(\mathbb{C}P^n)_0$.*

As hinted above, this corollary leads to a result on nondisplaceability of $\mathbb{R}P^n$ from other monotone Lagrangians which are Floer-theoretically nontrivial. This has been known due to Biran and Cornea [8, Corollary 8.1.2] and Entov and Polterovich [18]. Very recently, Konstantinov [29] showed that the Chiang Lagrangian in $\mathbb{C}P^3$ admits a higher-rank local system making it Floer-theoretically nontrivial over a characteristic-2 field; then he concludes via Corollary 1.2 that the Chiang Lagrangian is nondisplaceable from $\mathbb{R}P^3$. It is possible that for Lagrangians with higher-rank local systems, a generalisation of [8] and [18] can be invoked instead of Corollary 1.2, but we have not checked this.

We can extract another interesting consequence about projective spaces from our main computation of the closed-open map.

Proposition 1.3 *The Fukaya A_∞ algebra of the Lagrangian $\mathbb{R}P^{4n+1} \subset \mathbb{C}P^{4n+1}$ is not formal over a characteristic-2 field for any $n \geq 0$.*

Here formality means an existence of a quasi-isomorphism with the associative algebra $\mathrm{HF}^*(\mathbb{R}P^{4n+1}, \mathbb{R}P^{4n+1}) \cong \mathbb{K}[u]/(u^{4n+2} - 1)$ considered as an A_∞ algebra with trivial higher-order structure maps. In particular, the Fukaya A_∞ algebra of the equator $S^1 \subset S^2$ is not formal in characteristic 2; we devote a separate discussion to this fact in Section 3, where we explicitly exhibit a nontrivial Massey product which provides an alternative proof of the nonformality. Below is another example of split-generation which we can prove using the same methods.

Proposition 1.4 *Let $\mathbb{K}$ be a field of characteristic 2, $X = \mathrm{Bl}_{\mathbb{C}P^1} \mathbb{C}P^9$ the blowup of $\mathbb{C}P^9$ along a complex line which intersects $\mathbb{R}P^9$ in a circle, and $L \subset X$ the blowup of $\mathbb{R}P^9$ along that circle. Then $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(L,L)$ is injective, although $\mathcal{CO}^0\colon \mathrm{QH}^*(X) \to \mathrm{HF}^*(L,L)$ is not. Consequently, L split-generates $\mathcal{F}uk(X)_0$.*

(The manifold $\mathrm{Bl}_{\mathbb{C}P^1} \mathbb{C}P^9$ is the first instance among $\mathrm{Bl}_{\mathbb{C}P^k} \mathbb{C}P^n$ for which L is monotone of Maslov index at least 2 and such that $\mathcal{CO}^0$ is not injective; the last requirement makes the use of our general results essential in this example.) In general, it is known that the real Lagrangian in a toric Fano variety is not displaceable from the monotone toric fibre; this was proved by Alston and Amorim [5]. Proposition 1.4 implies a much stronger nondisplaceability result like the one which has been known for $\mathbb{R}P^n \subset \mathbb{C}P^n$.

Corollary 1.5 *Let $\mathbb{K}$ and $L \subset X$ be as in Proposition 1.4, and $L' \subset X$ any other monotone Lagrangian, perhaps equipped with a local system $\pi_1(L) \to \mathbb{K}^\times$, with minimal Maslov number at least 2 and such that $\mathrm{HF}^*(L',L') \neq 0$. If $w(L') \neq 0$, we also assume the technical Hypothesis 3.9, which is expected to hold following Ganatra* [22]. *Then $L \cap L' \neq \varnothing$.*

Here $\mathrm{HF}^*(L',L')$ denotes the Floer cohomology of L' with respect to the local system ρ, so a better notation would be $\mathrm{HF}^*((L',\rho),(L',\rho))$. For brevity, we decided to omit ρ from our notation of Floer and Hochschild cohomologies throughout the article when it is clear that a Lagrangian is equipped with such a local system. The point of allowing local systems in Corollary 1.5 is to introduce more freedom in achieving the nonvanishing of $\mathrm{HF}^*(L',L')$.

Note that Corollary 1.5 does not require that the obstruction number of L' matches the one of L, namely zero. If $w(L') \neq 0$, we can pass to $X \times X$ noticing that $w(L' \times L') = 2w(L') = 0$ and similarly $w(L \times L) = 0$, so we have well-defined Floer theory between the two product Lagrangians. This trick was observed by Abreu and Macarini [3] and has also been used in Alston and Amorim [5]. So it suffices to show that $L \times L$ split-generates $\mathcal{F}uk(X \times X)_0$; this follows from Proposition 1.4 by the general expectation that the condition of the Abouzaid's split-generation criterion is "preserved" under Künneth isomorphisms. As we explain later, this general expectation is contingent upon a certain commutative diagram which we formulate as Hypothesis 3.9, and which is largely substantiated by Ganatra [22]; see also Abouzaid and Smith [2] and Amorim [6].

As in the case of $\mathbb{R}P^n$, we also prove a nonformality statement.

Proposition 1.6 *The Fukaya A_∞ algebra of the Lagrangian $\mathrm{Bl}_{\mathbb{R}P^1}\mathbb{R}P^9 \subset \mathrm{Bl}_{\mathbb{C}P^1}\mathbb{C}P^9$ from Proposition 1.4 is not formal over a characteristic-2 field.*

Although we cannot prove that $\mathcal{CO}^*$ is injective for the real locus of an arbitrary toric Fano variety, we are able to do this in a slightly wider range of examples, which we postpone to Section 3. We will prove Proposition 1.1 and Corollary 1.2 at the end of the introduction, and the remaining statements from above will be proved in Section 3. Now we state the main theorem; the new pieces of notation are explained right after the statement.

Theorem 1.7 *Let X be a compact monotone symplectic manifold and $L \subset X$ a monotone Lagrangian submanifold of Maslov index at least 2, possibly equipped with a local system $\rho\colon H_1(L) \to \mathbb{K}^\times$. If $\operatorname{char}\mathbb{K} \neq 2$, assume L is oriented and spin.*

Let $\gamma = \{\gamma_t\}_{t\in S^1}$ be a loop of Hamiltonian symplectomorphisms of X, and denote by $S(\gamma) \in \mathrm{QH}^*(X)$ the corresponding Seidel element. Suppose the loop γ preserves L setwise, that is, $\gamma_t(L) = L$. Denote by $l \in H_1(L)$ the homology class of an orbit $\{\gamma_t(q)\}_{t\in S^1}$, $q \in L$. Finally, assume $\mathrm{HF}^*(L,L) \neq 0$.

(a) *Then $\mathcal{CO}^0(S(\gamma)) = (-1)^{\epsilon(l)}\cdot\rho(l)\cdot 1_L$, where $1_L \in \mathrm{HF}^*(L,L)$ is the unit.*

(b) *Suppose there exists no $a \in \mathrm{HF}^*(L,L)$ such that*

$$(*)\quad \mu^2(a,\Phi(y)) + \mu^2(\Phi(y),a) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle y,l\rangle\cdot 1_L \quad \text{for each } y \in H^1(L).$$

Then $\mathcal{CO}^*(S(\gamma)) \in \mathrm{HH}^*(L,L)$ is linearly independent from the Hochschild cohomology unit.

(c) *More generally, suppose $Q \in \mathrm{QH}^*(X)$ and there exists no $a \in \mathrm{HF}^*(L,L)$ such that*

$$(**)\quad \mu^2(a,\Phi(y)) + \mu^2(\Phi(y),a) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle y,l\rangle\cdot\mathcal{CO}^0(Q) \quad \text{for each } y \in H^1(L).$$

Then $\mathcal{CO}^*(S(\gamma) * Q)$ and $\mathcal{CO}^*(Q)$ are linearly independent in the Hochschild cohomology $\mathrm{HH}^*(L,L)$.

Here μ^2 is the product on $\mathrm{HF}^*(L,L)$, $\langle -,-\rangle$ is the pairing $H^1(L)\otimes H_1(L) \to \mathbb{K}$, and $S(\gamma) * Q$ is the quantum product of the two elements. Next,

$$\Phi\colon H^1(L) \to \mathrm{HF}^*(L,L)$$

is the PSS map of Albers [4], which is canonical and well defined if $\mathrm{HF}^*(L,L) \neq 0$. Its well-definedness in a setting closer to ours was studied by eg Biran and Cornea [9],

and later we discuss it in more detail. Note that Φ is not necessarily injective, although in our applications, where $\mathrm{HF}^*(L,L) \cong H^*(L)$, it will be. Finally, in the theorem, we have allowed L to carry an arbitrary local system, which modifies the Fukaya A_∞ structure of L by counting the same punctured holomorphic disks as in the case without a local system with coefficients which are the values of ρ on the boundary loops of such disks. The algebras $\mathrm{HF}^*(L,L)$ and $\mathrm{HH}^*(L,L)$ get modified accordingly, although their dependence on ρ is not reflected by our notation, as mentioned earlier. We allow nontrivial local systems in view of our application to toric fibres and will only need the trivial local system $\rho \equiv 1$ for applications to real Lagrangians.

To complete the statement of Theorem 1.7, we need to explain the appearance of the sign $(-1)^{\epsilon(l)} = \pm 1$. By a spin Lagrangian, we always mean a Lagrangian with a fixed spin structure (rather than admitting one). We have two natural trivialisations of TL over the loop $\gamma_t(q) \subset L$: the one induced from a fixed basis of T_qL by the Hamiltonian loop γ and the one determined by the spin structure on L. We put $\epsilon(l)$ to be 0 if the two trivialisations agree and 1 otherwise.

Outline of proof We mentioned earlier that part (a) of Theorem 1.7 is an easy consequence of the paper by Charette and Cornea [11]. The proof of parts (b) and (c) also starts by using a result from that paper, and then the main step is an explicit computation of

$$\mathcal{CO}^1(S(\gamma))|_{\mathrm{CF}^1(L,L)} \colon \mathrm{CF}^1(L,L) \to \mathrm{CF}^0(L,L)$$

on the cochain level, which turns out to be dual to taking the γ–orbit of a point up to the factor $(-1)^{\epsilon(l)}\rho(l)$: this is Proposition 2.8. The final step is to check whether the computed nontrivial piece of the Hochschild cocycle $\mathcal{CO}^*(S(\gamma))$ survives to cohomology; this is controlled by equations $(*)$ and $(**)$.

Remark 1.8 In our examples, we will never encounter a nontrivial sign $(-1)^{\epsilon(l)}$: for real Lagrangians, we shall be working over characteristic 2, and for toric fibres with the standard spin structure, this sign is easily seen to be $+1$. The examples when the sign $(-1)^{\epsilon(l)}$ is negative have been found by J Smith [47, Remark 5.3.2]; they occur for $\mathrm{PSU}(N-1)$–homogeneous Lagrangians. We are grateful to him for pointing out the presence of this sign in general, which was missed in the previous versions of the paper.

1.2 The split-generation criterion

We will now briefly discuss the split-generation criterion in more detail, particularly because we wish to pay attention to both cases: $\mathrm{char}\,\mathbb{K} = 2$ and $\mathrm{char}\,\mathbb{K} \neq 2$. We continue

to denote by $L \subset X$ a monotone Lagrangian submanifold with minimal Maslov number at least 2, which is oriented and spin if $\operatorname{char}\mathbb{K} \neq 2$. If $\operatorname{char}\mathbb{K} = 2$, we allow L to be nonorientable. Consider the quantum multiplication by the first Chern class as an endomorphism of quantum cohomology, $-*c_1(X)\colon \mathrm{QH}^*(X) \to \mathrm{QH}^*(X)$. If $\mathbb{K}$ is algebraically closed, we have an algebra decomposition $\mathrm{QH}^*(X) = \bigoplus_w \mathrm{QH}^*(X)_w$, where $\mathrm{QH}^*(X)_w$ is the generalised w–eigenspace of $-*c_1(X)$ for $w \in \mathbb{K}$.

Recall that $w(L) \in \mathbb{K}$ denotes the obstruction number of L, ie the count of Maslov index 2 disks with boundary on L. By an observation of Auroux, Kontsevich and Seidel, $\mathcal{CO}^0(2c_1) = 2w(L)\cdot 1_L$, which in $\operatorname{char}\mathbb{K} \neq 2$ implies that $\mathcal{CO}^0(c_1) = w(L)\cdot 1_L$; see eg [45]. Now suppose that $\operatorname{char}\mathbb{K} = 2$ and $c_1(X)$ lies in the image of $H^2(X, L; \mathbb{K}) \to H^2(X; \mathbb{K})$, which is true if L is orientable (because the Maslov class goes to twice the Chern class under $H^2(X, L; \mathbb{Z}) \to H^2(X; \mathbb{Z})$, and the Maslov class of an orientable manifold is integrally divisible by two). In this case, the same argument shows again that $\mathcal{CO}^0(c_1) = w(L)\cdot 1_L$. This way one deduces the following lemma, which is well known but usually stated only for $\operatorname{char}\mathbb{K} \neq 2$.

Lemma 1.9 *For $\mathbb{K}$ of any characteristic, if L is orientable, then $\mathcal{CO}^0\colon \mathrm{QH}^*(X) \to \mathrm{HF}^*(L, L)$ vanishes on all summands except maybe $\mathrm{QH}^*(X)_{w(L)}$.* □

(If $w(L)$ is not an eigenvalue of $-*c_1(X)$, then $\mathcal{CO}^0$ vanishes altogether, and it follows that $\mathrm{HF}^*(L, L) = 0$. Recall that L is required to be monotone.) The same vanishing statement is expected to hold for the full map $\mathcal{CO}^*$. Keeping this vanishing in mind, we see that the "naive" version of the split-generation criterion stated in the introduction, that $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(L, L)$ is injective, can only be useful when $\operatorname{char}\mathbb{K} = 2$ and L is nonorientable. In other cases, it must be replaced by a more practical criterion which does not ignore the eigenvalue decomposition; we will now state both versions of the criterion. Let $\mathit{Fuk}(X)_w$ denote the Fukaya category whose objects are monotone Lagrangians in X with minimal Maslov number at least 2, oriented and spin if $\operatorname{char}\mathbb{K} \neq 2$, and whose obstruction number equals $w \in \mathbb{K}$.

Theorem 1.10 *Let $L_1, \dots, L_n \subset X$ be Lagrangians which are objects of $\mathit{Fuk}(X)_w$, and $\mathcal{G} \subset \mathit{Fuk}(X)_w$ the full subcategory generated by $L_1, \dots, L_n$. Then $\mathcal{G}$ split-generates $\mathit{Fuk}(X)_w$ if either of the two following statements hold:*

(a) *$\operatorname{char}\mathbb{K} \neq 2$ and $\mathcal{CO}^*|_{\mathrm{QH}^*(X)_w}\colon \mathrm{QH}^*(X)_w \to \mathrm{HH}^*(\mathcal{G})$ is injective.*

(b) *$\mathbb{K}$ is arbitrary and $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(\mathcal{G})$ is injective.* □

In the monotone case, this theorem is due to Ritter and Smith [38] and Sheridan [45]. It is more common to only state part (a), but it is easy to check the same proof works for part (b) as well. (In part (a), we could also allow $\operatorname{char}\mathbb{K}=2$, if L is orientable.) Theorem 1.10 is most easily applied when $\mathrm{QH}^*(X)_w$ is 1–dimensional: because $\mathcal{CO}^*$ is unital, it automatically becomes injective. We apply this theorem in more complicated cases. Before we proceed, let us mention one easy corollary of split-generation. We say that $L_1,\dots,L_n$ split-generate the Fukaya category when $\mathcal{G}$ does.

Lemma 1.11 *If Lagrangians $L_1,\dots,L_n\subset X$ split-generate $\mathcal{F}uk(X)_w$, and $L\subset X$ is another Lagrangian that is an object of $\mathcal{F}uk(X)_w$ with $\mathrm{HF}^*(L,L)\neq 0$, then L has nonempty intersection and nonzero Floer cohomology with one of the L_i.* □

1.3 $\mathcal{CO}^0$ for real toric Lagrangians

In this subsection, we state a theorem that computes $\mathcal{CO}^0$ for real Lagrangians in toric manifolds. Using this, it is easy to identify the cases when $\mathcal{CO}^0$ is injective (and the split-generation follows immediately), and the cases when $\mathcal{CO}^0$ is not injective, and therefore, a further study of $\mathcal{CO}^*$ is required to establish the split-generation. Our subsequent goal is to apply the main result, Theorem 1.7, to some examples of the latter type.

Let X be a (smooth, compact) toric Fano variety with minimal Chern number at least 2, ie $\langle c_1(X), H_2(X;\mathbb{Z})\rangle = N\mathbb{Z}$ with $N\geq 2$. As a toric manifold, X has a canonical antiholomorphic involution $\tau\colon X\to X$. Its fixed locus is the so-called real Lagrangian $L\subset X$ which is smooth (see Duistermaat [16, page 419]), monotone, and whose minimal Maslov number equals the minimal Chern number of X (see Haug [23]). When speaking of such real Lagrangians, we will always be working over a field $\mathbb{K}$ of characteristic 2. In particular, there is the Frobenius map

$$\mathcal{F}\colon \mathrm{QH}^*(X)\to \mathrm{QH}^{2*}(X),\quad \mathcal{F}(x)=x^2.$$

Because $\operatorname{char}\mathbb{K}=2$, we have that $\mathcal{F}$ is a map of unital algebras. We have reflected in our notation that $\mathcal{F}$ multiplies the $\mathbb{Z}/2N$–grading by two. A classical theorem of Duistermaat [16] constructs, again in $\operatorname{char}\mathbb{K}=2$, the isomorphisms $H^i(L)\cong H^{2i}(X)$. We can package these isomorphisms into a single isomorphism of unital algebras,

$$\mathcal{D}\colon H^{2*}(X)\xrightarrow{\cong} H^*(L).$$

Let us now recall a recent theorem of Haug [23].

Theorem 1.12 *If* $\operatorname{char}\mathbb{K} = 2$, *then* $\mathrm{HF}^*(L,L) \cong H^*(L)$ *as vector spaces. Using the identification coming from a specific perfect Morse function from* [23], *and also identifying* $\mathrm{QH}^*(X) \cong H^*(X)$, *the same map*

$$\mathcal{D}\colon \mathrm{QH}^{2*}(X) \xrightarrow{\cong} \mathrm{HF}^*(L,L)$$

is again an isomorphism of unital algebras. □

It turns out that it is possible to completely compute $\mathcal{CO}^0$ for real toric Lagrangians. This rather quickly follows by combining the works of Charette and Cornea [11], Hyvrier [25], and McDuff and Tolman [32]; we explain this theorem in Section 3.

Theorem 1.13 *The diagram below commutes:*

$$\begin{array}{ccc} \mathrm{QH}^*(X) & \xrightarrow{\mathcal{F}} & \mathrm{QH}^{2*}(X) \\ & {\scriptstyle \mathcal{CO}^0} \searrow & \Big\downarrow {\scriptstyle \mathcal{D}} \cong \\ & & \mathrm{HF}^*(L,L) \end{array}$$

In particular, $\mathcal{CO}^0$ is injective if and only if $\mathcal{F}$ is injective.

1.4 Split-generation for the real projective space

We conclude the introduction by proving Proposition 1.1 and Corollary 1.2. The crucial idea is that when n is odd, the kernel of $\mathcal{CO}^0\colon \mathrm{QH}^*(\mathbb{C}P^n) \to \mathrm{HF}^*(\mathbb{R}P^n, \mathbb{R}P^n)$ is the ideal generated by the Seidel element of a nontrivial Hamiltonian loop preserving $\mathbb{R}P^n$; this allows us to apply Theorem 1.7 and get new information about $\mathcal{CO}^*$. Recall that $\mathrm{QH}^*(X) \cong \mathbb{K}[x]/(x^{n+1}-1)$ and $w(\mathbb{R}P^n) = 0$ because the minimal Maslov number of $\mathbb{R}P^n$ equals $n+1$ (when $n=1$, we still have $w(S^1) = 0$ for $S^1 \subset S^2$).

Proof of Proposition 1.1 If n is even, the Frobenius map on $\mathrm{QH}^*(\mathbb{C}P^n)$ is injective, so by Theorem 1.13, $\mathcal{CO}^0\colon \mathrm{QH}^*(\mathbb{C}P^n) \to \mathrm{HF}^*(\mathbb{R}P^n, \mathbb{R}P^n)$ is injective, and hence so is $\mathcal{CO}^*$.

Now suppose n is odd, and let $n = 2p-1$. Given $\operatorname{char}\mathbb{K} = 2$, we have $\mathrm{QH}^*(\mathbb{C}P^n) \cong \mathbb{K}[x]/(x^p+1)^2$, so $\ker\mathcal{F} = \ker\mathcal{CO}^0$ is the ideal generated by x^p+1. Consider the Hamiltonian loop γ on $\mathbb{C}P^n$ which in homogeneous coordinates $(z_1 : \cdots : z_{2p})$ is the

rotation $\left(\begin{smallmatrix}\cos t & \sin t\\ -\sin t & \cos t\end{smallmatrix}\right)$ for $t \in [0,\pi]$, applied simultaneously to the pairs $(z_1,z_2),\dots,(z_{2p-1},z_{2p})$. Note that t runs to π, not 2π. This loop is Hamiltonian isotopic to the loop

$$(z_0 : \cdots : z_{2p-1}) \mapsto (e^{2it}z_0 : z_1 : \cdots : e^{2it}z_{2p-1} : z_{2p}), \quad t \in [0,\pi],$$

so $S(\gamma) = x^p$; see [32]. The loop γ clearly preserves the real Lagrangian $\mathbb{R}P^n \subset \mathbb{C}P^n$, and its orbit l is a generator of $H_1(\mathbb{R}P^n) \cong \mathbb{K}$. Taking $y \in H^1(\mathbb{R}P^n)$ to be the generator, we get $\langle y, l\rangle = 1$, and the right-hand side of equation $(**)$ from Theorem 1.7 equals $\mathcal{CO}^0(Q)$. On the other hand, the product on $\mathrm{HF}^*(\mathbb{R}P^n,\mathbb{R}P^n)$ is commutative by Theorem 1.12, so the left-hand side of $(**)$ necessarily vanishes. We conclude that the hypothesis of Theorem 1.7(c) is satisfied for any $Q \notin \ker \mathcal{CO}^0$.

Let us prove that $\mathcal{CO}^*(P) \neq 0$ for each nonzero $P \in \mathrm{QH}^*(\mathbb{C}P^n)$. If $\mathcal{CO}^0(P) \neq 0$, we are done, so it suffices to suppose that $\mathcal{CO}^0(P) = 0$. This means that

$$P = (x^p + 1) * Q = (S(\gamma) + 1) * Q$$

for some $Q \in \mathrm{QH}^*(\mathbb{C}P^n)$. Note that if $Q \in \ker \mathcal{CO}^0 = \ker \mathcal{F}$, then $P \in (\ker \mathcal{F})^2 = \{0\}$. So if $P \neq 0$, then $\mathcal{CO}^0(Q) \neq 0$, and thus $\mathcal{CO}^*(P) \neq 0$ by Theorem 1.7(c) and the observation earlier in this proof. □

Remark 1.14 When n is even, $c_1(\mathbb{C}P^n)$ is invertible in $\mathrm{QH}^*(\mathbb{C}P^n)$, so the 0–eigenspace $\mathrm{QH}^*(\mathbb{C}P^n)_0$ is trivial; but L is nonorientable, so this does not contradict Lemma 1.9. On the other hand, when n is odd, L is orientable, but $c_1(\mathbb{C}P^n)$ vanishes in $\mathrm{char}\,\mathbb{K} = 2$, so the whole $\mathrm{QH}^*(\mathbb{C}P^n)$ is its 0–eigenspace; this is also consistent with Lemma 1.9.

Proof of Corollary 1.2 This follows from Proposition 1.4 and Theorem 1.10(b). □

The same trick of finding a real Hamiltonian loop whose Seidel element generates $\ker \mathcal{CO}^0$ works for some other toric manifolds which have "extra symmetry" in addition to the toric action, like a Hamiltonian action of $\mathrm{SU}(2)^{\dim_{\mathbb{C}} X/2}$ which was essentially used above. As already mentioned, we will provide more explicit examples in Section 3.

Acknowledgements The author is most grateful to his supervisor Ivan Smith for many useful comments as well as constant care and enthusiasm. Yankı Lekili and Jack Smith have provided valuable feedback and pointed out two inaccuracies in the previous versions of this paper. The paper has also benefited from discussions with Mohammed Abouzaid, Lino Amorim, François Charette, Georgios Dimitroglou Rizell, Jonny Evans,

Alexander Ritter, Paul Seidel, Nick Sheridan and Renato Vianna. We are grateful to the referee for helpful suggestions on improving the exposition.

The author was funded by the Cambridge Commonwealth, European and International Trust, and acknowledges travel funds from King's College, Cambridge. The work was partially supported by RFBR grant 15-01-06302.

2 Proof of Theorem 1.7

Let X be a monotone symplectic manifold and $w \in \mathbb{K}$. We recall that the objects in the monotone Fukaya category $\mathcal{F}uk(X)_w$ are monotone Lagrangian submanifolds $L \subset X$ with minimal Maslov number at least 2, oriented and spin if $\operatorname{char}\mathbb{K} \neq 2$, equipped with local systems $\rho\colon \pi_1(L) \to \mathbb{K}^\times$, whose count of Maslov index 2 disks (weighted using ρ) equals w. We will use the definition of the Fukaya category based on achieving transversality by explicit Hamiltonian perturbations of the pseudoholomorphic equation. This setup was developed by Seidel [40] for exact manifolds and carries over to monotone ones; see [38; 45]. There is a notion of bounding cochains from [20], generalising the notion of a local system, and all results are expected carry over to them as well.

2.1 A theorem of Charette and Cornea

Suppose $\gamma = \{\gamma_t\}_{t\in S^1}$ is a loop of Hamiltonian symplectomorphisms on X. As explained by Seidel in [40, Section 10c], the loop γ gives rise to a natural transformation $\gamma^\sharp$ from the identity functor on $\mathcal{F}uk(X)_w$ to itself. Any such natural transformation is a cocycle of the Hochschild cochain complex $\mathrm{CC}^*(\mathcal{F}uk(X)_w)$ [40, Section 1d]. Denote the corresponding Hochschild cohomology class by

$$[\gamma^\sharp] \in \mathrm{HH}^*(\mathcal{F}uk(X)_w).$$

We denote, as earlier, the closed-open map by $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(\mathcal{F}uk(X)_w)$ and the Seidel element by $S(\gamma) \in \mathrm{QH}^*(X)$. The following theorem was proved by Charette and Cornea [11].

Theorem 2.1 *If we take for $\mathcal{F}uk(X)_w$ the Fukaya category of Lagrangians with trivial local systems only, then $\mathcal{CO}^*(S(\gamma)) = [\gamma^\sharp]$.* □

Let us now restrict to a single Lagrangian L which is preserved by the Hamiltonian loop γ and denote by $l \in H_1(L)$ the homology class of an orbit of γ on L. Let $\mathrm{CC}^*(L,L)$ denote the Hochschild cochain complex of the A_∞ algebra $\mathrm{CF}^*(L,L)$,

and let $\mathrm{HH}^*(L,L)$ be its Hochschild cohomology. (The definition of Hochschild cohomology will be reviewed later in this section.) We will now need to recall the proof of Theorem 2.1 for several reasons: first, we wish to see how Theorem 2.1 gets modified in the presence of a local system on L; second, we shall see the appearance of a sign "hidden" in $\gamma^\sharp$; third and most importantly, we will recall the definition of the moduli spaces computing $\gamma^\sharp$ in the process. Eventually, for later use, we need a form of Theorem 2.1 expressed by formula (2-2) below, which takes the local system and the sign into account.

Pick some Floer datum $\{H_s, J_s\}_{s\in[0,1]}$ and perturbation data defining an A_∞ structure on Floer's complex $\mathrm{CF}^*(L,L)$ [40]. Recall that the maps

$$\mathcal{CO}^k(S(\gamma))\colon \mathrm{CF}^*(L,L)^{\otimes k} \to \mathrm{CF}^*(L,L)$$

count 0–dimensional moduli space of disks satisfying a perturbed pseudoholomorphic equation (with appropriately chosen perturbation data) with $k+1$ boundary punctures (k inputs and one output) and one interior marked point. These disks satisfy the Lagrangian boundary condition L, and their interior marked point is constrained to a cycle dual to $S(\gamma)$; see Figure 1, top left (in this figure, we abbreviate the datum $\{H_s, J_s\}$ simply to H). A disk u is counted with coefficient $\pm\rho(\partial u)$, where the sign $\pm$ comes from the orientation on the moduli space, and $\rho(\partial u)\in\mathbb{K}^\times$ is the monodromy of the local system. The collection of maps $\mathcal{CO}^*(S(\gamma)) := \{\mathcal{CO}^k\}_{k\geq 0}$ is a Hochschild cochain in $\mathrm{CC}^*(L,L)$ if all perturbation data are chosen consistently with gluing.

The argument of Charette and Cornea starts by passing to a more convenient definition of the closed-open map in which $\mathcal{CO}^k$ counts holomorphic disks with $k+1$ boundary punctures and one interior puncture (instead of a marked point). We can view the neighbourhood of the interior puncture as a semi-infinite cylinder; then the pseudoholomorphic equation restricts on this semi-infinite cylinder to a Hamiltonian Floer equation with some Floer datum $\{F_t, J_t\}_{t\in S^1}$. We input the PSS image of $S(\gamma)$ to the interior puncture (see Figure 1, top middle) given as a linear combination of some Hamiltonian orbits z (in the figure, we abbreviate the datum $\{F_t, J_t\}$ simply to F_t).

The PSS image of $S(\gamma)$ counts configurations shown in the upper part of Figure 1, top middle, consisting of disks with one output puncture (say, asymptotic to an orbit y), and a cylinder counting continuation maps from $(\gamma_t^{-1})^* y$ seen as an orbit of Floer's complex with datum pulled back by the loop γ_t^{-1} [39, Lemmas 2.3 and 4.1] to another orbit z of the original Floer's complex with datum $\{F_t, J_t\}$. Let us glue the z–orbits

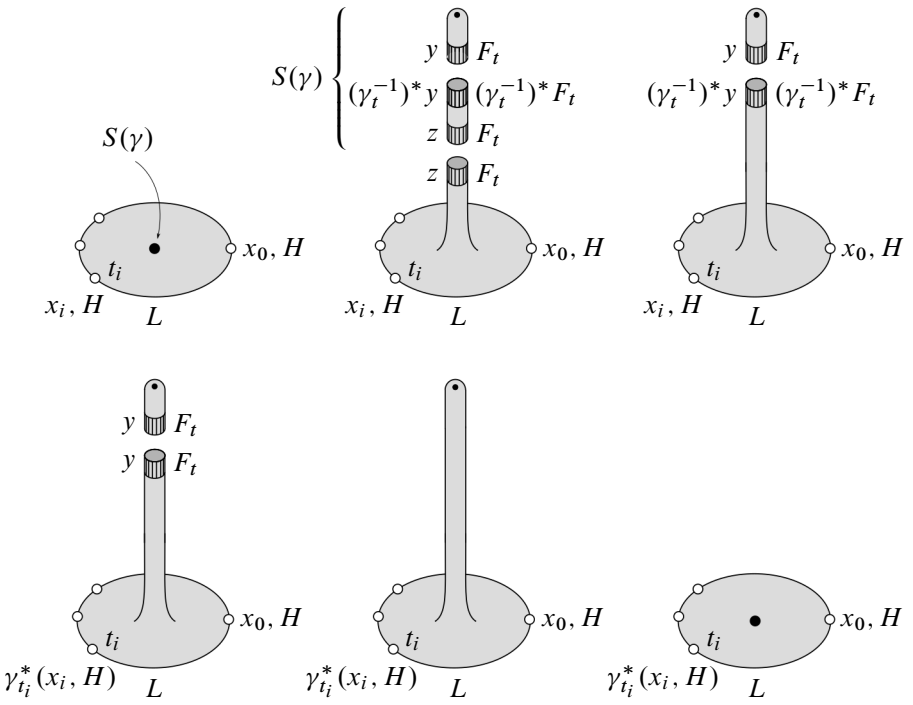


Figure 1: A computation of $\mathcal{CO}^*(S(\gamma))$ by Charette and Cornea

together, passing to Figure 1, top right, and then substitute each lower punctured pseudoholomorphic disk u in Figure 1, top right, by $\widetilde{u}$ defined as

$$\widetilde{u}(re^{2\pi it}) = \gamma_t^{-1} \circ u(re^{2\pi it}), \tag{2-1}$$

assuming that the interior puncture is located at $0 \in \mathbb{C}$ and the output puncture at $1 \in \mathbb{C}$. Let us look at the effect of this substitution.

First, $[\partial u] = [\partial \widetilde{u}] + l \in H_1(L)$, so the count of configurations in Figure 1, top right (before substitution) is equal to the count of configurations in Figure 1, bottom left (after substitution) multiplied by $\rho(l)$.

Second, $\widetilde{u}$ satisfies the same boundary condition L because $\gamma_t L = L$, but the perturbation data defining the pseudoholomorphic equation get pulled back accordingly. In particular, the Lagrangian Floer datum $\{H_s, J_s\}_{s\in[0,1]}$ and the asymptotic chord at a strip-like end corresponding to the boundary puncture at $t_i \in S^1$ get pulled back by γ_{t_i}.

Third, (2-1) gives an abstract bijection $u \mapsto \widetilde{u}$ between the respective zero-dimensional moduli spaces, but we should discuss how this bijection behaves with respect to the

signs attached to u and $\widetilde{u}$ by the orientations on the moduli spaces. Assume for simplicity that the Hamiltonian perturbation is small enough so that we can canonically deform ∂u and $\partial\widetilde{u}$ to loops inside L. Take the trivialisation of $TL|_{\partial u}$ defined by the spin structure on L and push it forward by γ to a trivialisation of $TL|_{\partial\widetilde{u}}$. Obviously, (2-1) preserves the signs computed using these trivialisations; we remark that this trivialisation of $TL|_{\partial\widetilde{u}}$ is used in the general definition of $\gamma^\sharp$. In our specific case, $\widetilde{u}$ is again a curve with boundary on L (rather than with a moving Lagrangian boundary condition), and we wish to consider a different orientation scheme for $\widetilde{u}$, namely the usual scheme for orienting moduli spaces of curves with boundary on L using the given spin structure. We shall be using this orientation scheme from now on, and we observe that it uses the trivialisation of $TL|_{\partial\widetilde{u}}$ coming from the spin structure, which may be different from the pushforward trivialisation of $TL|_{\partial\widetilde{u}}$ mentioned before. We denote the sign difference between the two orientations for $\widetilde{u}$ by $(-1)^\epsilon = \pm 1$; this sign equals $+1$ if and only if the two trivialisations of $TL|_{\partial\widetilde{u}}$ from above are homotopic. (Ultimately, we are going to use the fact that for our choice of orientation scheme, the moduli space of constant unconstrained disks is positively oriented; this may not be true for the γ pushforward orientation scheme.)

Remark 2.2 In general, the number ϵ does not necessarily equal $\epsilon(l)$ from the introduction. This equality holds when $\partial\widetilde{u}$ is contractible, as easily seen from the definitions.

Fourth, near the interior puncture, $\widetilde{u}$ satisfies the Hamiltonian Floer equation with original datum $\{F_t, J_t\}$ at the interior puncture, and is asymptotic orbit y. So we can glue the y–orbits, passing to Figure 1, bottom middle, and Figure 1, bottom right, is another drawing of the same domain we got after gluing, namely the disk with $k+1$ boundary punctures and one interior unconstrained marked point fixed at $0 \in \mathbb{C}$. Let us explain the presence of the marked point $0 \in \mathbb{C}$: it is carried over from a marked point on the upper disk in Figure 1, top middle, where the interior marked point serves to stabilise the domain; such a marked point is present in the definition of the Seidel element.

Summing up,

$$\mathcal{CO}^k(S(\gamma))(x_1 \otimes \cdots \otimes x_k) = (-1)^\epsilon \cdot \rho(l) \cdot \sum \#\mathcal{M}^\gamma(x_1, \ldots, x_k; x_0) \cdot x_0, \tag{2-2}$$

where $\mathcal{M}^\gamma(x_1, \ldots, x_k; x_0)$ is the 0–dimensional moduli space of disks shown in Figure 1, bottom right, which satisfy the inhomogeneous pseudoholomorphic equation defined by domain- and modulus-dependent perturbation data in the sense of [40] such that:

- The disks carry the unconstrained interior marked point fixed at $t = 0$, the output boundary puncture fixed at $t_0 = 1$, and k free input boundary punctures at $t_i \in S^1$, $i = 1, \dots, k$.
- On a strip-like end corresponding to a boundary puncture $t_i \in S^1$, perturbation data restrict to the Floer datum which is the γ_{t_i}–pullback of the original Floer datum $\{H_s, J_s\}_{s\in[0,1]}$, and the asymptotic chord for this strip must be the γ_{t_i}–pullback of the asymptotic chord x_i of the original Floer datum.
- The data must be consistent with gluing strip-like ends at $t_i \in S^1$ to strip-like ends of punctured pseudoholomorphic disks carrying the γ_{t_i}–pullbacks of the perturbation data defining the A_∞ structure on $\mathrm{CF}^*(L, L)$.
- We use the standard orientation scheme for curves with boundary on L to orient the $\mathcal{M}^\gamma$, and the sign $(-1)^\epsilon$ was explained above.

The counts $\#\mathcal{M}^\gamma$ are signed and weighted by ρ; the third condition guarantees that $\mathcal{CO}^*(S(\gamma))$ is a Hochschild cocycle. Formula (2-2) coincides with the formula from [40, Section 10c] defining the natural transformation $[\gamma^\sharp]$ up to $(-1)^\epsilon \rho(l)$, and we have clarified this difference.

Remark 2.3 The fixed interior marked point at $t = 0$ and the fixed boundary marked point at $t_0 = 1$ make sure our disks have no automorphisms, so the values $t_i \in S^1$ of the other boundary punctures are uniquely defined.

Before proceeding, note that we are already able to compute $\mathcal{CO}^0(S(\gamma))$.

Corollary 2.4 *If $\{\gamma_t\}_{t\in S^1}$ is a Hamiltonian loop such that $\gamma_t(L) = L$, then on the chain level,*

$$\mathcal{CO}^0(S(\gamma)) = (-1)^{\epsilon(l)} \cdot \rho(l) \cdot 1_L \in \mathrm{HF}^0(L, L).$$

(Here 1_L is a chain-level representative of the cohomology unit; see (2-3) below.)

Proof When $k = 0$, the moduli space in formula (2-2) is exactly the moduli space defining the cohomological unit in $\mathrm{CF}^*(L, L)$; see eg [45, Section 2.4]. The equality $\epsilon = \epsilon(l)$ holds by Remark 2.2: for a small Hamiltonian, the curves computing the unit are close to being constant and therefore have contractible boundary. □

Proof of Theorem 1.7(a) This is the homology-level version of Corollary 2.4. □

2.2 The PSS maps in degree one

Our goal will be to compute a "topological piece" of $\mathcal{CO}^1(S(\gamma))$. This subsection introduces some background required for the computation: in particular, we recall that there is a canonical map $\Phi\colon H^1(L) \to \mathrm{HF}^*(L,L)$ which was used in the statement of Theorem 1.7. This is the Lagrangian PSS map of Albers [4], and the fact it is canonical was discussed, for instance, by Biran and Cornea [9, Proposition 4.5.1(ii)] in the context of Lagrangian quantum cohomology.

First, recall that once the Floer datum is fixed, the complex $\mathrm{CF}^*(L,L)$ acquires the Morse $\mathbb{Z}$–grading. This grading is not preserved by the Floer differential or the A_∞ structure maps, but is still very useful. Assume that the Hamiltonian perturbation, as part of the Floer datum, is chosen to have a unique minimum x_0 on L, which means that $\mathrm{CF}^0(L,L)$ is one-dimensional and generated by x_0. We denote by

$$1_L \in \mathrm{CF}^0(L,L) \tag{2-3}$$

the chain-level cohomological unit defined in [45, Section 2.4], which is proportional to x_0. Now pick a metric and a Morse–Smale function f on L with a single minimum; together they define the Morse complex which we denote by $C^*(L)$. Consider the "Maslov index 0" versions of the PSS maps, which are linear (but *not* chain) maps

$$\Psi\colon \mathrm{CF}^*(L,L) \to C^*(L), \quad \Phi\colon C^*(L) \to \mathrm{CF}^*(L,L). \tag{2-4}$$

These maps are defined as in the paper of Albers [4], with the difference that Φ and Ψ count configurations with Maslov index 0 disks only. For example, the map Ψ counts configurations consisting of a Maslov index 0 pseudoholomorphic disk with boundary on L and one input boundary puncture, followed by a semi-infinite gradient trajectory of f which outputs an element of $C^*(L)$. Similarly, Φ counts configurations in which a semi-infinite gradient trajectory is followed by a Maslov index 0 disk with an output boundary puncture. The maps Ψ and Φ preserve $\mathbb{Z}$–gradings on the two complexes.

Let $d_0\colon \mathrm{CF}^*(L,L) \to \mathrm{CF}^{*+1}(L,L)$ be the "Morse" part of the Floer differential counting the contribution of Maslov index 0 disks; see Oh [34]. Let $d_{\mathrm{Morse}}\colon C^*(L) \to C^{*+1}(L)$ be the usual Morse differential. The lemma below is a version of [4, Theorem 4.11].

Lemma 2.5 *Φ and Ψ are chain maps with respect to d_0 and d_{Morse}, and are cohomology inverses of each other.* □

Lemma 2.6 *Suppose $\mathrm{HF}^*(L,L) \neq 0$. If $y \in C^1(L)$ is a Morse cocycle (resp. coboundary) then $\Phi(y)$ is a Floer cocycle (resp. coboundary).*

Proof This follows from the fact that the image and the kernel of d_0 and the full Floer differential d coincide on $\mathrm{CF}^1(L,L)$ if $\mathrm{HF}^*(L,L)\neq 0$ by Oh's decomposition of the Floer differential [34]. □

Consequently, if $\mathrm{HF}^*(L,L)\neq 0$, we get a map

$$\Phi\colon H^1(L)\to \mathrm{HF}^*(L,L).$$

For Ψ, we have a weaker lemma using [34] (this lemma is not true for coboundaries instead of cocycles).

Lemma 2.7 *If $y\in \mathrm{CF}^*(L,L)$ is a Floer cocycle, then $\Psi(y)\in C^*(L)$ is a Morse cocycle.* □

By Lemmas 2.6 and 2.7, given $\mathrm{HF}^*(L,L)\neq 0$, we have the following induced maps which we denote by the same symbols Ψ and Φ, abusing notation:

$$\Psi\colon \mathrm{CF}^*(L,L)\to H^*(L),\quad \Phi\colon H^*(L)\to \mathrm{HF}^*(L,L). \tag{2-5}$$

In particular, Theorem 1.7 in Section 1 refers to this cohomological version of the map Φ. We remind that Ψ does not necessarily descend to a map from $\mathrm{HF}^*(L,L)$.

2.3 Computing the topological part of $\mathcal{CO}^1(S(\gamma))$

We continue to use the above conventions and definitions, namely we use the $\mathbb{Z}$–grading on $\mathrm{CF}^*(L,L)$, the maps Φ and Ψ, and the choice of a Hamiltonian perturbation on L with a unique minimum x_0. From now on, we assume $\mathrm{HF}^*(L,L)\neq 0$. Recall that $\mathcal{CO}^*(S(\gamma))$ is determined via formula (2-2) by the moduli spaces $\mathcal{M}^\gamma(x_1,\dots,x_k;x_0)$. The connected components of $\mathcal{M}^\gamma(x_1,\dots,x_k;x_0)$ corresponding to disks of Maslov index μ have dimension

$$|x_0|+k+\mu-\sum_{i=1}^{k}|x_i|,$$

where $|x_i|$ are the $\mathbb{Z}$–gradings of the $x_i\in \mathrm{CF}^*(L,L)$. Consequently,

$$\mathcal{CO}^k(S(\gamma))\colon \mathrm{CF}^*(L,L)^{\otimes k}\to \mathrm{CF}^*(L,L)$$

is a sum of maps of degrees

$$-k-mN_L,\quad m\geq 0,$$

where N_L is the minimal Maslov number of L. In particular, the restriction of $\mathcal{CO}^1(S(\gamma))$ to $\mathrm{CF}^1(L,L)$ is of pure degree -1; that is, its image lands in $\mathrm{CF}^0(L,L)$:

$$\mathcal{CO}^1(S(\gamma))|_{\mathrm{CF}^1(L,L)}\colon \mathrm{CF}^1(L,L)\to \mathrm{CF}^0(L,L).$$

Moreover, this map is determined by the moduli space consisting of Maslov index 0 disks only and can be computed in purely topological terms. This is the main technical computation which we now perform; recall that $l\in H_1(L)$ is the homology class of an orbit of γ in L, and the sign $(-1)^{\epsilon(l)}$ was defined in Section 1.

Proposition 2.8 *Suppose* $\mathrm{HF}^*(L,L)\neq 0$. *If* $x\in \mathrm{CF}^1(L,L)$ *is a Floer cocycle, then on the chain level,*

$$\mathcal{CO}^1(S(\gamma))(x) = (-1)^{\epsilon(l)}\cdot\rho(l)\cdot\langle\Psi(x),l\rangle\cdot 1_L.$$

(Here $\Psi(x)\in H^1(L)$ is from (2-5), $\langle -,-\rangle$ denotes the pairing $H^1(L)\otimes H_1(L)\to\mathbb{K}$, we consider $\mathcal{CO}^1$ on the chain level, and 1_L is a chain-level cohomology unit as in (2-3).)

Proof All disks with boundary on L we consider in this proof are assumed to have Maslov index 0. We identify the domains of all disks that appear in the proof with the unit disk in $\mathbb{C}$, and their boundaries are identified with the unit circle $S^1\subset\mathbb{C}$. In the subsequent figures, punctured marked points will be drawn by circles filled white, and unpunctured marked points by circles filled black. According to formula (2-2), for a generator $x\in\mathrm{CF}^1(L,L)$ we have

$$\mathcal{CO}^1(S(\gamma))(x) = (-1)^{\epsilon}\cdot\rho(l)\cdot\#\mathcal{M}^{\gamma}(x;x_0),$$

where $\mathcal{M}^{\gamma}(x;x_0)$ consists of (perturbed pseudoholomorphic) Maslov index 0 disks whose domains are shown in Figure 2, left.

Step 1 (perturbation data producing bubbles with unpunctured points) Recall that the domains appearing in the moduli space $\mathcal{M}^{\gamma}(x;x_0)$ are disks with the interior marked point 0 and boundary punctures 1 and t, where $t\in S^1\setminus\{1\}$. For further use, we will choose perturbation data defining $\mathcal{M}^{\gamma}(x;x_0)$ whose bubbling behaviour as $t\to 1$ differs from the standard one. Usually, the perturbation data would be chosen so as to be compatible, as $t\to 1$, with the gluing G_1 shown in Figure 2, where the bubble meets the principal disk along a puncture, meaning that near this puncture it satisfies a Floer equation and shares an asymptotic Hamiltonian chord with the corresponding puncture

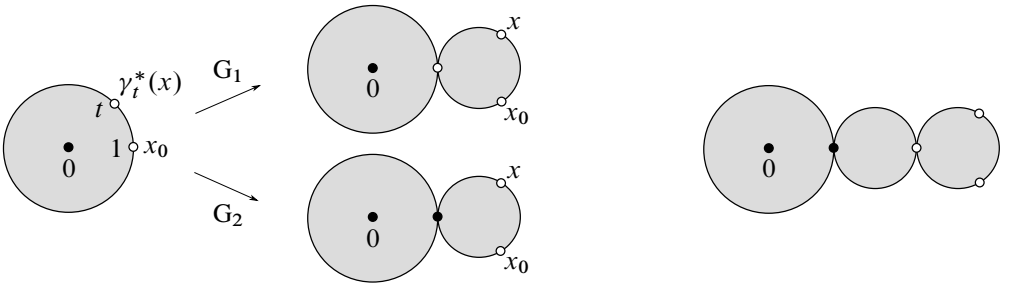


Figure 2: Two types of gluings for $\mathcal{M}^\gamma(x;x_0)$ and a way to interpolate between the glued perturbation data

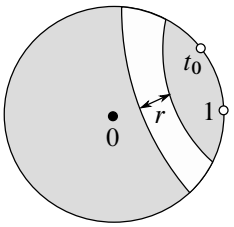


Figure 3: Collision of two boundary marked points seen as stretching the unshaded strip with parameter $r \to +\infty$

of the principal disk. On the other hand, we will use perturbation data consistent with gluing G_2 shown in Figure 2, where the bubble is attached to the principal disk by an unpunctured marked point. Near the unpunctured marked point, the disks satisfy a holomorphic equation with no Hamiltonian term.

Let us explain how to define both types of data more explicitly. The domain in Figure 2, left, with free parameter t close to 1, is biholomorphic to the domain shown in Figure 3 whose boundary marked points are fixed at 1 and some $t_0 \in S^1$, upon which a stretching procedure along the unshaded strip is performed. This stretching procedure changes the complex structure on the disk by identifying the strip with $[0,1]\times[0,1]$, removing it, and gluing back the longer strip $[0,1]\times[0,r]$. The parameter $r \in [1,+\infty)$ is free and replaces the free parameter t, so that tending $r \to +\infty$ replaces the collision of two marked points $t \to 1$.

In order to get perturbation data which are consistent with the usual bubbling, as shown by G_1 in Figure 2, one requires the perturbed pseudoholomorphic equation to coincide, on the strip $[0,1]\times[0,r]$, with the usual Floer equation defining the Floer differential, which uses a Hamiltonian perturbation translation-invariant in the direction of $[0,r]$. In order to get perturbation data producing the bubbling pattern shown by G_2 in Figure 2,

we simply put an unperturbed pseudoholomorphic equation on the strip $[0,1]\times[0,r]$ without using a Hamiltonian perturbation at all.

Both ways of defining perturbation data are subject to appropriate gluing and compactness theorems, which precisely say that as we tend $r\to\infty$, the solutions bubble in one of the two corresponding ways shown in Figure 2. The standard choice is used, for example, to prove that $[\gamma^\sharp]$ (obtained from the counts of various $\mathcal{M}^\gamma$) is a Hochschild cocycle in the Fukaya A_∞ algebra of L defined using Seidel's setup with Hamiltonian perturbations. The other choice will be more convenient for our computations. Note that the two different types of perturbation data give the same count $\#\mathcal{M}^\gamma(x;x_0)$: this is proved by interpolating between them using the two-parametric space of perturbation data obtained from gluing together the disks in Figure 2, right, with different length parameters. Recall that all disks in $\mathcal{M}^\gamma(x;x_0)$ have Maslov index 0 as the Morse index $|x|=1$; therefore, no unnecessary bubbling occurs. (Since we do not want to compute the moduli spaces $\mathcal{M}^\gamma$ other than the $\mathcal{M}^\gamma(x;x_0)$ for $|x|=1$, we do not have to worry about extending our unusual type of perturbation data to the other moduli spaces.)

In addition, we will assume that the Hamiltonian perturbation vanishes over the principal disk in Figure 2, bottom, making this disk J–holomorphic and hence constant, because the disk has Maslov index 0. Such configurations can be made consistent with gluing: for this, one just needs to make the Hamiltonian perturbation vanish over the shaded subdomain in Figure 3, for all t close to 1. Note that regularity can be achieved by perturbing the pseudoholomorphic equation over the subdomain to the right of the unshaded strip in Figure 3.

Step 2 (a one-dimensional cobordism from $\mathcal{M}^\gamma(x;x_0)$) In what follows, we will use (and assume familiarity with) the theory of holomorphic pearly trees developed by Sheridan in his Morse–Bott definition of the Fukaya category [45]; see also the earlier work of Cornea and Lalonde [15]. Sheridan performs the analysis based on extending Seidel's setup of Fukaya categories from [40]. Although [45] considers exact Lagrangians instead of monotone ones, all the analysis works equally well in the nonexact case if we only consider disks of Maslov index 0, because here unpunctured disk bubbles cannot occur just like in the exact case. Techniques for dealing with holomorphic pearly trees (or "clusters") with disks of arbitrary Maslov index have appeared in [15; 10], but we will not actually need to appeal to them.

We will now define a family of domains depending on two parameters $s\in[0,2\pi]$ and $l\in[-1,+\infty]$. When $s\notin\{0,2\pi\}$, the domains are shown in Figure 4, where $t=e^{is}$;

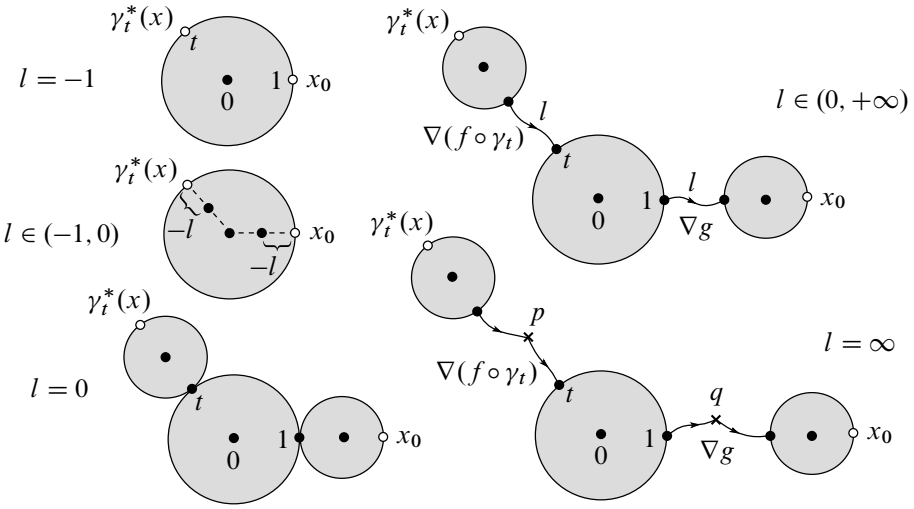


Figure 4: The domains for $s\in(0,2\pi)$ and $l\in[-1,\infty]$, where $t=e^{is}$

we discuss the case $s\in\{0,2\pi\}$ later. When $l=-1$, the domain is the disk from the definition of $\mathcal{M}^{\gamma}(x;x_0)$. When $l\in(-1,0)$, the domain is the same disk (called principal) with two additional interior marked points whose position is determined by the parameter l: the first point lies on the line segment $[0,t]$, the second one lies on the line segment $[0,1]$, and both points have distance $1+l$ from 0. When $l=0$, the domain consists of the principal disk with marked points $0,1,t$, and two bubble disks attached to the principal disk at points 1 and t. The first bubble disk has marked points $0,1$ and a boundary puncture at -1; the second one has marked points at $0,-1$ and a boundary puncture at 1. When $0<l<\infty$, the domain contains the same three disks, now disjoint from each other, plus two line segments of length l connecting the bubble disks to the principal one along the boundary marked points at which the disks used to be attached to each other. (So far, the length is just a formal parameter associated with the domain, but soon it will become the length of the flowline corresponding to the segment.) When $l=\infty$, we replace each line segment by two rays $[0,+\infty)\sqcup(-\infty,0]$.

When $s=0$ or $s=2\pi$, the domains obtain extra bubbles as those discussed above in the definition of $\mathcal{M}^{\gamma}(x;x_0)$, which correspond to the parameter $t=e^{is}\in S^1$ approaching $1\in S^1$ from the two sides. These domains are shown in Figure 5: as l goes from -1 to 0, the two interior marked points move along the punctured paths. Observe that these points are crossing the node between the two disks at some intermediate value of l; this does not cause any difficulty with the definitions because these marked points are only used to represent varying perturbation data consistent with the types of bubbling

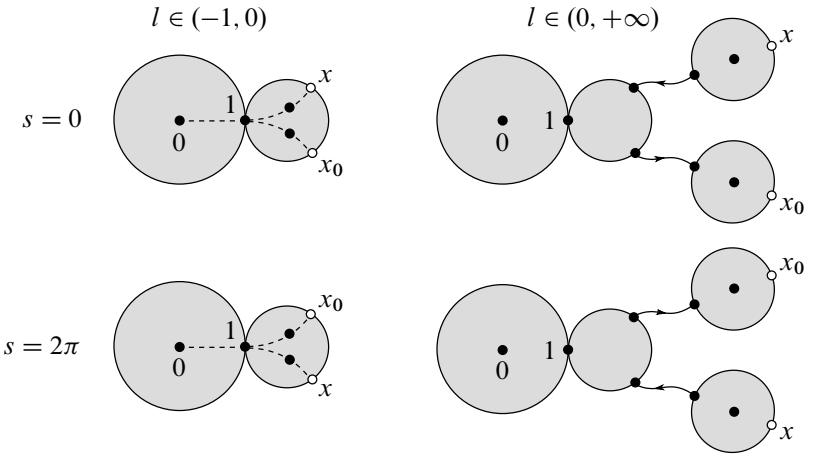


Figure 5: The domains for $s \in \{0, 2\pi\}$, $l \in [-1, \infty]$

we prescribe in the figures. We will soon mention what these varying data are in terms of stretching certain strips inside a fixed disk. When $l > 0$, the length of the paths equals l. When $l = \infty$, one introduces broken lines $[0, +\infty) \sqcup (-\infty, 0]$ as above.

Having specified the domains, we briefly explain how to equip the disks with suitable perturbed pseudoholomorphic equations and line segments with suitable gradient equations to get a moduli space of solutions. When $l = -1$, we choose the equations defining $\mathcal{M}^\gamma(x; x_0)$ as discussed above in Step 1; in particular, this is consistent with bubbling at the unpunctured point as $s \to 0$ or $s \to 2\pi$. When $-1 < l < 0$, we choose the equations with the same properties as for $\mathcal{M}^\gamma(x; x_0)$, which are additionally consistent with bubbling at unpunctured points as $l \to 0$. When $l \geq 0$, we choose the equation on the disk with an input puncture to be the γ_t–pullback of the one appearing in the definition of the PSS map Ψ, and the equation on the disk with an output puncture to be exactly the equation from the PSS map Φ. Finally we fix two generic Morse–Smale functions $f, g\colon L \to \mathbb{R}$. On the line segments and rays, the equation is the gradient equation for g or the γ_t–pullback of the gradient equation for f, as shown in Figure 4.

In general, we can arrange the equations on all disks to have a Hamiltonian perturbation. For the disks with punctures, we must do so anyway; however, for the disks without punctures (with unpunctured marked points only), we can choose the equation to be J–holomorphic without a Hamiltonian term, provided that we check that our moduli space can be made regular with this restricted choice. In order to carry out the computations below, we choose the zero Hamiltonian perturbation on all disks without boundary punctures, namely,

- the central disk in Figure 4, bottom left, top right and bottom right,
- the left disks in all domains of Figure 5, and the central disks in Figure 5, top right and bottom right.

We will now specify one more property of the equations that we choose. When $s = 0$ and $l < 0$, we require the perturbation data on the twice-punctured disk in Figure 5, top left, to be obtained by π–rotation from the perturbation data on the similar disk for $s = 2\pi$ (and the same parameter l) if we identify the two punctures with points 1 and -1 of the unit disk. When $s = 0$ and $l \geq 0$, we make a similar symmetric choice.

Finally, we specify the asymptotic conditions at the punctures. If x is a generator of $\mathrm{CF}^1(L, L)$, we specify that the input puncture in Figure 4 must be asymptotic to the γ_t–pullback of x (as usual, if x is a linear combination of generators, we take the disjoint union of the relevant moduli spaces). The output puncture must be asymptotic to the unique generator $x_0 \in \mathrm{CF}^0(L, L)$. When $l = \infty$, the first pair of rays in Figure 4, bottom right, must be asymptotic to a point p such that $\gamma_t^{-1}(p) \in C^1(L)$ (that is, p is an index-1 critical point of $f \circ \gamma_t$) and the second pair of rays must be asymptotic to $q \in C^0(L)$; we assume q is the unique minimum of g. The interior marked points on the disks are unconstrained.

Above, we have specified a 2–dimensional space of domains and the equations over them. This gives us a moduli space of solutions ("pearly trees") which is 1–dimensional by our choice of indices. We remind the reader that a formal definition of this moduli space falls into the setup of moduli spaces of pearly trajectories given by Sheridan [44]. For our purposes, its description given above will suffice.

The boundary of our 1–dimensional moduli space consists of

- solutions whose domains have parameter $l = -1$ or $l = \infty$,
- solutions whose domains have parameter $s = 0$ or $s = 2\pi$.

We claim that solutions of the second type cancel pairwise. Indeed, recall that the disks without boundary punctures in Figure 5 are constant, and the perturbation data on the punctured disks for $s = 0, 2\pi$ are chosen in a way to provide the same solutions after a π–rotation on each disk. Let us describe more explicitly what happens when $l < 0$ as the case when $l \geq 0$ is clear enough from Figure 5, top right and bottom right. We can represent the domains shown in Figure 4, top left, with free $l < 0$ and free small $s > 0$, where $t = e^{is}$, by a disk with fixed boundary punctures, stretched with length parameters $-1/l$ and $1/s$ along the three strips shown in Figure 6, left. The stretching procedure

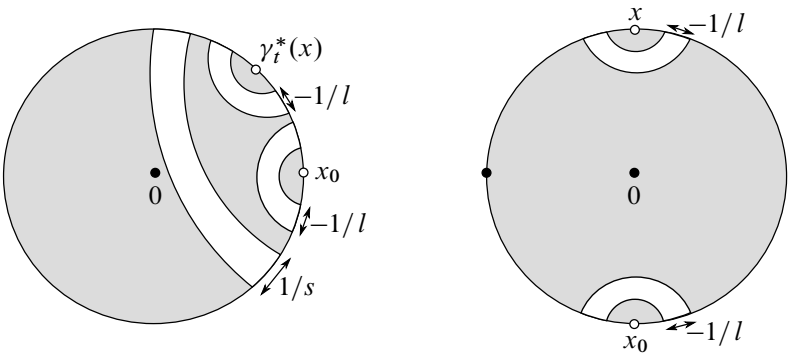


Figure 6: The domains (left) for $l < 0$, $s > 0$, seen as a fixed disk with three stretched strips, and (right) for $l < 0$, $s = 0$, when the principal disk is constant.

was described earlier, and our choice of perturbation data says that the stretched strips, and the subdomain to the left of the $1/s$–strip, carry an unperturbed pseudoholomorphic equation. So for $s = 0$, we get the disks shown in Figure 6, right, with the unpunctured boundary marked point attached to a constant disk, which means this boundary marked point is unconstrained. (As usual, the domain is considered up to complex automorphisms, so the unconstrained point does not prevent us from having rigid solutions.) This way, Figures 5, top left, and 6, right, are drawings of the same configuration for any $l < 0$. If we rotate the disk in Figure 6, right, by π, we get precisely the disk with perturbation data we would have gotten for $s = 2\pi$, except that the boundary marked point (with the attached constant disk) is on the different side of the boundary.

We have shown that the solutions for $s = 0$ and $s = 2\pi$ are in a natural bijection; we claim that this bijection reverses the signs associated to those solutions as parts of the moduli space. This is ultimately related to the fact that the constant disk is being glued to those solutions at the opposite boundary components. One can adopt a proof (which we will not provide in detail here) from the following classical example where an analogous sign issue has been treated. Suppose L_0 and L_1 are monotone Lagrangian submanifolds with obstruction numbers $m_0(L_i) \in \mathbb{Z}$, $i = 0, 1$. Then the Floer differential d on $\mathrm{CF}^*(L_0, L_1)$ satisfies $d^2 = m_0(L_0) - m_0(L_1)$. This relation arises from Maslov index 2 disks bubbling off the two sides of a 1–dimensional moduli space of Floer strips [33]; see also [45, (2.3.9)] and [46, Figure 2]. Here, indeed, the gluing of the side bubbles contributes with the opposite signs and results in the term $m_0(L_0) - m_0(L_1)$. The reference for the signs in this relation is [20]; see specifically Remark 3.2.21(1), Formula (3.3.4) and Chapter 8 from that book. Although in our case

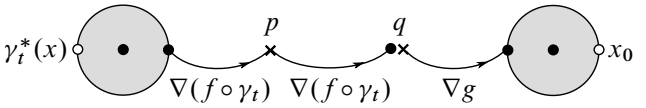


Figure 7: The domains when $l=+\infty$ and the principal disk together with a flowline are constant. Here p is an index-1 critical point of $f\circ\gamma_t$, and q is the minimum of f.

we would be gluing a constant disk rather than Maslov index 2 disk, and the nonconstant curve in Figure 6, right, satisfies a different equation than the standard Floer one (eg our equation is not $\mathbb{R}$–invariant), the required orientation analysis is essentially the same.

The outcome of the cancellation discussed above is that the count of configurations in Figure 4, top left, ie $\#\mathcal{M}^\gamma(x;x_0)$, equals the count of configurations in Figure 4, bottom right, and it remains to compute the latter.

Step 3 (a Morse-theoretic computation) Let us look at Figure 4, bottom right. Recall that $q\in L$ is the minimum of g, so the semi-infinite flowline of ∇g flowing into q must be constant. Second, we have arranged the principal (central) disk to be constant, as well. So the configurations in Figure 4, bottom right, reduce to those shown in Figure 7.

The free parameter $t=e^{is}\in S^1\setminus\{1\}$ is "unseen" by the domain after the principal disk became ghost (ie constant), but the equations still depend on it. First, consider the left disk and the left flowline in Figure 7, forgetting the rest of the configuration. Those disk and flowline satisfy the γ_t–pullback of the equation defining the PSS map Ψ, so for each t, the linear combination of points p appearing as limits of such configurations equals $\gamma_t(\Psi(x))$, where $\Psi(x)\in C^1(L)$ is the PSS image which is a linear combination of index-1 critical points of f, so that $\gamma_t(\Psi(x))$ is a combination of critical points of $f\circ\gamma_t$.

Let us now add back the middle flowline, still forgetting the right flowline and the right disk, and count the resulting configurations. The middle flowline is a semi-infinite flowline of $\nabla(f\circ\gamma_t)$ ending at the point q; note that q is *not* a critical point of f. Suppose for the moment that we allow the right end of the middle flowline to be free (not constrained to q) and denote the moduli space of such configurations by P. Then we can consider the evaluation map at the right end of the flowline, ev: $P\to L$. The image of ev, as a chain, is a linear combination of unstable manifolds, with respect to the function $f\circ\gamma_t$, associated with the linear combination of the critical points p which we have previously computed. Recall that this linear combination of points p equals $\gamma_t(\Psi(x))$. Consequently, if we denote by $C_{\Psi(x)}\subset L$ the disjoint union of

(oriented, codimension-1) unstable manifolds of the Morse cochain $\Psi(x) \in C^1(L)$ with respect to f, then

$$P = (S^1 \setminus \{1\}) \times C_{\Psi(x)}, \quad \mathrm{ev}(t, z) = \gamma_t(z).$$

Those configurations which evaluate at $q \in L$ are the intersection points $C_{\Psi(x)} \cap l$, where $l = \{\gamma_t(q)\}_{t \in S^1}$ is the orbit of q. By perturbing γ_t and f, the intersections can be easily made transverse, and we get

$$\#(P \times_{\mathrm{ev}} \{q\}) = [C_{\Psi(x)}] \cdot [l] = \langle \Psi(x), l \rangle.$$

Recall this is the count of the part of configurations in Figure 7 which end up at q. Finally, the count of the rightmost flowlines (emerging from q) plus the right disks in Figure 7 equals $1_L \in \mathrm{CF}^0(L, L)$. Indeed, the g–unstable manifold of the minimum q is the whole manifold L (minus a codimension-2 subset), so the count is the same as the count of the rightmost disks only, and the latter produces 1_L by definition.

Putting everything together and noting that $\epsilon = \epsilon(l)$ by Remark 2.2 (since the total boundary in Figure 7 is contractible for a small Hamiltonian perturbation), we get the statement of Proposition 2.8. One last thing is to argue that the moduli spaces we have been using are regular.

According to [45], the regularity of moduli spaces of pearly trajectories consisting of pseudoholomorphic disks and flowlines is equivalent to the regularity of the separate disks and flowlines not constrained to satisfy the incidence conditions, plus the transversality of the evaluation maps which account for the incidence conditions.

The nonconstant disks in the proof carry the pseudoholomorphic equation with a Hamiltonian perturbation which makes them regular. The constant disks are known to be regular on their own; and it is easy to see that for generic f and g, the flowlines are transverse to the evaluation maps for all appearing configurations. □

2.4 Checking nontriviality in Hochschild cohomology

In this subsection, we prove Theorem 1.7(b)–(c) (recall that part (a) was proved earlier; see Corollary 2.4). We have computed in Proposition 2.8 the map $\mathcal{CO}^1(S(\gamma))|_{\mathrm{CF}^1(L,L)}$, and it remains to see when the result survives to something nontrivial on the level of Hochschild cohomology, and thus distinguishes $\mathcal{CO}^*(S(\gamma)) \in \mathrm{HH}^*(L, L)$ from the unit in $\mathrm{HH}^*(L, L)$.

First, let us quickly recall the definition of Hochschild cohomology. Let A be an A_∞ algebra, and assume it is $\mathbb{Z}/2$ graded if $\operatorname{char}\mathbb{K} \neq 2$. When $\operatorname{char}\mathbb{K} = 2$, we consider A as an ungraded algebra. The space of Hochschild cochains is, by definition,

$$\mathrm{CC}^*(A,A) = \prod_{k\geq 0} \operatorname{Hom}(A^{\otimes k}, A).$$

If A is $\mathbb{Z}/2$–graded, then $\mathrm{CC}^*(A,A)$ is $\mathbb{Z}/2$–graded:

$$\mathrm{CC}^r(A,A) = \prod_{k\geq 0} \operatorname{Hom}(A^{\otimes k}, A[r-k]).$$

If $h = \{h^k\}_{k\geq 0} \in \mathrm{CC}^*(A)$ with $h^k\colon A^{\otimes k} \to A$, then the Hochschild differential of h is the sequence of maps

$$\begin{aligned}(\partial h)^k(a_k,\dots,a_1) &= \\ &\sum_{i+j\leq k} (-1)^{(r+1)(|a_1|+\cdots+|a_i|+i)} \mu^{k+1-i}(a_k,\dots,a_{i+j+1}, h^j(a_{i+j},\dots,a_{i+1}), a_i,\dots,a_1) \\ &\quad + \sum_{i+j\leq k} (-1)^{r+1+|a_1|+\cdots+|a_i|+i} \\ &\qquad\qquad \cdot h^{k+1-i}(a_k,\dots,a_{i+j+1}, \mu^j(a_{i+j},\dots,a_{i+1}), a_i,\dots,a_1).\end{aligned}$$

Here r is the $\mathbb{Z}/2$–degree of h. (When $k=0$, the agreement is that $\operatorname{Hom}(A^{\otimes 0}, A) = A$, so h^0 is an element of A.) If $\operatorname{char}\mathbb{K} = 2$, we do not need the gradings as the signs do not matter.

Let us return to the A_∞ algebra $\mathrm{CF}^*(L,L)$. We continue to use the Morse $\mathbb{Z}$–grading on the vector space $\mathrm{CF}^*(L,L)$ keeping in mind this grading is not respected by the A_∞ structure. If L is oriented, the reduced $\mathbb{Z}/2$–grading is preserved by the A_∞ structure, so $\mathrm{CF}^*(L,L)$ is a $\mathbb{Z}/2$–graded A_∞ algebra. If L is not oriented, we must suppose $\operatorname{char}\mathbb{K} = 2$.

Proof of Theorem 1.7(b) We continue to work with $\mathcal{CO}^*$ on the chain level. Because the homological closed-map is unital [45, Lemma 2.3], the Hochschild cohomology unit is realised by the cochain $1_{\mathrm{HH}} := \mathcal{CO}^*(1) \in \mathrm{CC}^*(L,L)$, where 1 is the unit in $\mathrm{QH}^*(X)$. The A_∞ category $\mathrm{CF}^*(L,L)$ need not be strictly unital, so the maps $(1_{\mathrm{HH}})^k = \mathcal{CO}^k(1)$ need not vanish for $k>0$. However, because the identity Hamiltonian loop preserves L and has homologically trivial orbits on it, Proposition 2.8 applies to $1 = S(\mathrm{Id}) \in \mathrm{QH}^*(X)$ and says that $(1_{\mathrm{HH}})^1(x) = 0$ for any Floer cocycle $x \in \mathrm{CF}^1(L,L)$.

Suppose $\mathcal{CO}^*(S(\gamma)) + \alpha\cdot 1_{\mathrm{HH}}$ is the coboundary of an element $h \in \mathrm{CC}^*(L,L)$ for some $\alpha \in \mathbb{K}$. By comparing $(\partial h)^0$ with $\mathcal{CO}^0(S(\gamma)) + \alpha\cdot(1_{\mathrm{HH}})^0$ (see Corollary 2.4),

we get

$$\mu^1(h^0) = (-1)^{\epsilon(l)}\rho(l)\cdot 1_L + \alpha\cdot 1_L.$$

Here μ^1 is the Floer differential and 1_L is a chain-level cohomology unit; see (2-3). The assumption $\mathrm{HF}^*(L,L)\neq 0$ implies that the Floer cohomology unit 1_L cannot be killed by the Floer differential. Therefore, we cannot solve the above equation unless $\alpha = -(-1)^{\epsilon(l)}\rho(l)$ and $\mu^1(h^0)=0$. Next, by comparing $(\partial h)^1$ with $\mathcal{CO}^1(S(\gamma))+\alpha\cdot(1_{\mathrm{HH}})^1$ (see Proposition 2.8) for any Floer cocycle $x\in \mathrm{CF}^1(L,L)$, we get

$$(-1)^{|h^0|+1}\mu^2(x,h^0)+(-1)^{(|h^0|+1)(|x|+1)}\mu^2(h^0,x)+\mu^1(h^1(x))+(-1)^{|h^1|+1}h^1(\mu^1(x)) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle\Psi(x),l\rangle\cdot 1_L.$$

Here we are using the version of Ψ as in (2-5). Because x is a Floer cocycle, the last summand of the left-hand side vanishes. If $\operatorname{char}\mathbb{K}=2$, let $a:=h^0\in\mathrm{CF}^*(L,L)$. If $\operatorname{char}\mathbb{K}\neq 2$, let $a\in\mathrm{CF}^{\mathrm{odd}}(L,L)$ be the odd degree part of h^0. By computing the signs in the above equality, we get for any Floer cocycle $x\in\mathrm{CF}^1(L,L)$,

$$\mu^2(x,a)+\mu^2(a,x)+\mu^1(h^1(x)) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle\Psi(x),l\rangle\cdot 1_L.$$

Recall that $\mu^1(h^0)=0$, so $\mu^1(a)=0$ as well, and we get the following equality for Floer cohomology classes $[x],[a]\in\mathrm{HF}^*(L,L)$ and $\Psi(x)\in H^1(L)$:

$$\mu^2([x],[a])+\mu^2([a],[x]) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle\Psi(x),l\rangle\cdot 1_L\in\mathrm{HF}^*(L,L).$$

Now put $x=\Phi(y)$, where $y\in C^1(L)$ is a Morse cochain and Φ is the chain-level map from (2-4). The above equality means that for all $[y]\in H^1(L)$,

$$\mu^2(\Phi([y]),[a])+\mu^2([a],\Phi([y])) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle[y],l\rangle\cdot 1_L\in\mathrm{HF}^*(L,L).$$

This time, we have used the homology-level version of Φ from (2-5). The above equality is exactly prohibited by the hypothesis of Theorem 1.7(b), so Theorem 1.7(b) is proved. □

Proof of Theorem 1.7(c) Note that, on the homology level,

$$\mathcal{CO}^0(S(\gamma)*Q)=\mathcal{CO}^0(S(\gamma))\cdot\mathcal{CO}^0(Q)=1_L\cdot\mathcal{CO}^0(Q)=\mathcal{CO}^0(Q)$$

(here the dot denotes the μ^2 product), so the only possible linear relation between $\mathcal{CO}^*(S(\gamma)*Q)$ and $\mathcal{CO}^*(Q)$ is the following, where 1 is the unit in $\mathrm{QH}^*(X)$:

$$\mathcal{CO}^*((S(\gamma)-1)*Q)=0.$$

We have $\mathcal{CO}^*((S(\gamma)-1)*Q) = \mathcal{CO}^*(S(\gamma)-1) \star \mathcal{CO}^*(Q)$, where the symbol $\star$ denotes the Yoneda product in Hochschild cohomology.

Let us now return to working with $\mathcal{CO}^*$ on the chain level. Recall that if $\phi = \{\phi^k\}_{k\geq 0}$, $\psi = \{\psi^k\}_{k\geq 0} \in \mathrm{CC}^*(L,L)$ are Hochschild cochains, the $k=1$ part of their Yoneda product by definition equals

$$(\phi \star \psi)^1(x) = \pm\mu^2(\phi^1(x), \psi^0) \pm \mu^2(\phi^0, \psi^1(x)).$$

There is an explicit formula for the signs which we do not need. Let us apply this formula to $\mathcal{CO}^*(S(\gamma)-1)$ and $\mathcal{CO}^*(Q)$. We know that $(\mathcal{CO}^*(S(\gamma)-1))^0 = 0$ by Corollary 2.4, and $(\mathcal{CO}^*(S(\gamma)-1))^1(x) = \mathcal{CO}^1(S(\gamma))(x)$ is given by Proposition 2.8 for any Floer cocycle $x \in \mathrm{CF}^1(L,L)$. Consequently, we get

$$\mathcal{CO}^1((S(\gamma)-1)*Q)(x) = (-1)^{\epsilon(l)}\rho(l)\cdot\langle\Psi(x), l\rangle\cdot\mathcal{CO}^0(Q).$$

From this point, the rest of the proof follows the one of Theorem 1.7(b). □

3 The closed-open map for real toric Lagrangians

In this section, after a short proof of Theorem 1.13, we look for further examples of real toric Lagrangians where Theorem 1.7 can be effectively applied. We also discover that Proposition 2.8, after additional work, allows us to show that the Fukaya A_∞ algebra of some of the considered Lagrangians is not formal. In particular, we prove the results about real toric Lagrangians stated in Section 1 (except for Proposition 1.1 and Corollary 1.2, which have been proved therein). We work with a coefficient field $\mathbb{K}$ of characteristic 2 throughout this section.

3.1 A proof of Theorem 1.13

Let X be a compact, smooth toric Fano variety, and $D \subset X$ be a toric divisor corresponding to one of the facets of the polytope defining X. There is a Hamiltonian circle action γ on X associated with D, which comes from the toric action by choosing a Hamiltonian which achieves maximum on D. A theorem of McDuff and Tolman [32] says the following.

Theorem 3.1 *$S(\gamma) = D^*$, where $D^* \in \mathrm{QH}^*(X)$ is the Poincaré dual of D.* □

The loop γ never preserves the real Lagrangian $L \subset X$, but if we parametrise $\gamma = \{\gamma_t\}_{t\in[0,1]}$, then $\gamma_{1/2}(L) = L$; see [23]. Consequently, $\alpha = \{\gamma_t(L)\}_{t\in[0,1/2]}$ is a

loop of Lagrangian submanifolds, and moreover, we have $\alpha^2 = \{\gamma_t(L)\}_{t\in[0,1]}$ in the space of Lagrangian loops. There is an associated Lagrangian Seidel element $S_L(\alpha) \in \mathrm{HF}^*(L,L)$, which counts pseudoholomorphic disks with rotating boundary condition α, and a single boundary puncture which evaluates to an element of $\mathrm{HF}^*(L,L)$. A theorem of Hyvrier [25, Theorem 1.13], based on the disk doubling trick, computes $S_L(\alpha)$.

Theorem 3.2 *We have $S_L(\alpha) = [L\cap D]^*$, where $L\cap D$ is the clean intersection that has codimension 1 in L, and $[L\cap D]^* \in H^1(L) \subset \mathrm{HF}^*(L,L)$ is its dual class.* □

The inclusion $H^1(L)\subset \mathrm{HF}^*(L,L)$ is the PSS map Φ from Section 2, which is injective because $\mathrm{HF}^*(L,L)\cong H^*(L)$ by Theorem 1.12.

Proof of Theorem 1.13 It suffices to prove that $\mathcal{CO}^0(D^*) = \mathcal{D}(\mathcal{F}(D^*))$, where $D\subset X$ is a toric divisor as above and $D^*\in \mathrm{QH}^*(X)$ is its dual class, because such D^* generate $\mathrm{QH}^*(X)$ as an algebra [32]. Let γ be the Hamiltonian loop corresponding to D as above, and α be the Lagrangian loop as above, such that $\alpha^2 = \{\gamma_t(L)\}_{t\in[0,1]}$. It follows from Theorem 2.1 that

$$\mathcal{CO}^0(S(\gamma)) = S_L(\alpha^2),$$

and the latter can be rewritten as $\mathcal{F}(S_L(\alpha))$, where $\mathcal{F}$ is the Frobenius map on $\mathrm{HF}^*(L,L)$. By Theorem 3.1, $S(\gamma)=D^*$, and by Theorem 3.2, $S_L(\alpha)=[L\cap D]^*$. Finally, if we look at Haug's construction [23] of the Duistermaat isomorphism $\mathcal{D}$, we will see that $[L\cap D]^* = \mathcal{D}(D^*)$. Putting everything together, we get

$$\mathcal{CO}^0(D^*) = \mathcal{F}(\mathcal{D}(D^*)).$$

Because $\mathcal{D}$ is a ring map, it commutes with the Frobenius maps on $\mathrm{HF}^*(L,L)$ and $\mathrm{QH}^*(X)$, and the theorem follows. □

3.2 Split-generation for toric varieties with Picard rank 2

It is known that the unique toric variety with Picard number 1 is the projective space. By a theorem of Kleinschmidt [28] (see also [14]), every n–dimensional toric Fano variety whose Picard group has rank 2 (ie whose fan has $n+2$ generators) is isomorphic to the projectivisation of a sum of line bundles over $\mathbb{C}P^{n-k}$

$$X(a_1,\dots,a_k) := \mathbb{P}_{\mathbb{C}P^{n-k}}(\mathcal{O}\oplus\mathcal{O}(a_1)\oplus\cdots\oplus\mathcal{O}(a_k)) \quad \text{with } a_i\geq 0,\ \sum_{i=1}^{k} a_i \leq n-k-1. \tag{3-1}$$

(The imposed conditions on the a_i are equivalent to X being toric Fano.) The $n+2$ vectors in $\mathbb{Z}^n$ generating the fan of $X(a_1,\dots,a_k)$ are the columns of the matrix

$$\begin{pmatrix} & & -1 & a_1 \\ & & \vdots & \vdots \\ & & -1 & a_k \\ & I_{n\times n} & 0 & -1 \\ & & \vdots & \vdots \\ & & 0 & -1 \end{pmatrix} \tag{3-2}$$

The minimal Chern number of $X(a_1,\dots,a_k)$ equals $\gcd\big(k+1, n-k+1-\sum a_i\big)$; see [36]. Some of these varieties provide further examples where, using Theorems 1.13 and 1.7, we can prove the injectivity of $\mathcal{CO}^*$ and deduce split-generation.

Theorem 3.3 *Let $X := X(a_1,\dots,a_k)$ be as above, $L \subset X$ the real Lagrangian, and $\mathbb{K}$ a field of characteristic 2. Suppose all a_i are odd and $\gcd\big(k+1, n-k+1-\sum a_i\big) \geq 2$.*

(a) *If $n-k+1$ is odd, then $\mathcal{CO}^0\colon \mathrm{QH}^*(X) \to \mathrm{HF}^*(L,L)$ is injective.*

(b) *If $n-k+1$ is even, k is even and the numbers a_i come in equal pairs, then $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(L,L)$ is injective while $\mathcal{CO}^0$ is not.*

In both cases, L split-generates $\mathcal{F}uk(X)_0$.

Proof Let $x, y \in H^2(X)$ be the generators corresponding to the last two columns of the matrix (3-2). They generate $\mathrm{QH}^*(X)$ as an algebra and satisfy the following relations when $\operatorname{char}\mathbb{K} = 2$:

$$x(x+y)^k = 1, \quad y^{n-k+1}(x+y)^{-\sum a_i} = 1.$$

(For brevity, we no longer use the symbol $*$ to denote the quantum product.) If $n-k+1$ is odd, one can show that the Frobenius endomorphism $\mathcal{F}$ on $\mathrm{QH}^*(X)$ is an isomorphism, so $\mathcal{CO}^0$ is injective by Theorem 1.13. It follows that $\mathcal{CO}^*$ is also injective, and split-generation follows from Theorem 1.10. Part (a) is proved.

In the rest of the proof we work with case (b), so we let $n-k+1 = 2r$, $k = 2q$, and $\sum a_i = 2p$. The rewritten relations in $\mathrm{QH}^*(X)$ are

$$x(x+y)^{2q} = 1, \quad y^{2r}(x+y)^{-2p} = 1. \tag{3-3}$$

Lemma 3.4 *For the ring $\mathrm{QH}^*(X)$ as in* (3-3)*, the kernel of $\mathcal{F}$ is the ideal generated by $y^r(x+y)^{-p}+1$.*

Proof Equations (3-3) are equivalent to

$$x^{-p} = y^{2rq}, \quad y^{4rq+2r} + y^{4rp+2p} + 1 = 0,$$

where the second equation is rewritten from the second equation in (3-3) using the substitution $x^{-p} = y^{2rq}$. This means if we let

$$R(y) = y^{2rq+r} + y^{2rq+p} + 1,$$

then $R(y) = y^r(x+y)^{-p} + 1$. Let $g = \gcd(2rq, p)$ and let $\alpha, \beta \in \mathbb{Z}$ be such that

$$-2rq \cdot \alpha + p \cdot \beta = g.$$

Consider the map $\phi\colon \mathbb{K}[u] \to \mathrm{QH}^*(X)$ given by $u \mapsto x^\alpha y^\beta$; this map is onto because

$$\phi(u^{p/g}) = y, \quad \phi(u^{2qr/g}) = x^{-1} \tag{3-4}$$

using the given relations (note that the powers p/g, $2qr/g$ are integral). Further, $\ker\phi$ is obviously the ideal generated by $V(u)^2$, where $V(u) := R(u^{p/g})$, and we conclude that ϕ provides an isomorphism

$$\phi\colon \mathbb{K}[u]/V(u)^2 \xrightarrow{\cong} \mathrm{QH}^*(X), \quad V(u) = u^{(p/g)(2rq+r)} + u^{(p/g)(2rq+p)} + 1. \tag{3-5}$$

It is clear that $V(u)$ generates the kernel of the Frobenius map on $\mathbb{K}[u]/V(u)^2$. Because $V(u)$ corresponds to $y^r(x+y)^{-p}+1$ under ϕ, Lemma 3.4 follows. □

We continue the proof of Theorem 3.3(b). It turns out that, similarly to the case of $\mathbb{R}P^n \subset \mathbb{C}P^n$ studied in the introduction, the generator of $\ker\mathcal{F}$ from Lemma 3.4 equals $S(\gamma)+1$ for a real Hamiltonian loop γ on X which preserves L setwise and has homologically nontrivial orbits on it. To construct γ, we will need the additional assumption that the a_i come in equal pairs, so we assume the sequence $(a_i)_{i=1}^{2q}$ is $(a_1, a_1, \dots, a_q, a_q)$.

Recall that X, being a toric manifold, is a quotient of $\mathbb{C}^{2r+2q+1}$ minus some linear subspaces determined by the fan, by an action of $(\mathbb{C}^*)^2$. Using the common notation, this action is given by $z \mapsto t_1^{v_1} t_2^{v_2} z$, where $z \in \mathbb{C}^{2r+2q+1}$ and v_1, v_2 are the vectors in $\mathbb{Z}^{2r+2q+1}$ given by the following two rows:

(3-6)

$2q$ entries					$2r-1$ entries			2 entries	
a_1	a_1	$\cdots$	a_q	a_q	-1	$\cdots$	-1	0	-1
-1	-1	$\cdots$	-1	-1	0	$\cdots$	0	-1	0

Let $(z_1, \dots, z_{2r+2q+1})$ be the coordinates on $\mathbb{C}^{2r+2q+1}$. The action of $(\mathbb{C}^*)^2$ on $\mathbb{C}^{2r+2q+1}$ commutes with the action of

$$G = \mathrm{SU}(2)^q \times \mathrm{SU}(2r),$$

where the $\mathrm{SU}(2)$ factors act respectively on $(z_1, z_2), \dots, (z_{2q-1}, z_{2q})$, and $\mathrm{SU}(2r)$ acts on $(z_{2q+1}, \dots, z_{2q+2r-1}, z_{2q+2r+1})$; note we have omitted z_{2q+2r}. (If we view X as a projective bundle over $\mathbb{C}P^{2r-1}$ as in (3-1), the coordinates on which $\mathrm{SU}(2r)$ acts are the homogeneous coordinates on the base.) Denote the real form of G by $G^{\mathbb{R}} = \mathrm{SO}(2)^q \times \mathrm{SO}(2r)$. Because all a_i are odd, the action of $(-1, +1) \in (\mathbb{C}^*)^2$ coincides with the action of $-I \in G$. Consequently, the action of G descends to a Hamiltonian action of $G/\pm I$ on X. Its real form $G^{\mathbb{R}}/\pm I$ preserves the real Lagrangian $L \subset X$, and we let γ be the S^1–subgroup of $G^{\mathbb{R}}/\pm I$ defined as follows. This subgroup lifts to the path from I to $-I$ in $G^{\mathbb{R}}$ which is the image of the rotation $\left(\begin{smallmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{smallmatrix}\right) \in \mathrm{SO}(2)$, $t \in [0, \pi]$, under the diagonal inclusions

$$\mathrm{SO}(2) \subset \mathrm{SO}(2)^q \times \mathrm{SO}(2)^r \subset \mathrm{SO}(2)^q \times \mathrm{SO}(2r) = G^{\mathbb{R}}.$$

Recall that we are assuming $\operatorname{char} \mathbb{K} = 2$.

Lemma 3.5 *The homology class of γ–orbits on L is nonzero in $H_1(L; \mathbb{K})$.*

Proof Indeed, L is a real projective bundle over $\mathbb{R}P^{2r-1}$, and the orbits project to the nontrivial cycle on the base, provided $\operatorname{char} \mathbb{K} = 2$. □

Lemma 3.6 *We have $S(\gamma) + 1 = y^r (x + y)^{-p} + 1$ (which is the generator of $\ker \mathcal{F}$ from Lemma 3.4).*

Proof Inside the complex group $G/\pm I$, the loop γ is homotopic to the loop γ' lifting to the path from I to $-I$ in G which is the image of the path $\left(\begin{smallmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{smallmatrix}\right) \in \mathrm{SU}(2)$, $t \in [0, \pi]$, under the diagonal inclusions

$$\mathrm{SU}(2) \subset \mathrm{SU}(2)^q \times \mathrm{SU}(2)^r \subset \mathrm{SU}(2)^q \times \mathrm{SU}(2r) = G.$$

By using the action of $\mathbb{C}^* \subset (\mathbb{C}^*)^2$ corresponding to the first vector in (3-6), we see that γ' descends to the same Hamiltonian loop in X as the loop γ'' in G which acts on $\mathbb{C}^{2r+2q+1}$, for $t \in [0, 2\pi]$, as follows:

$$\begin{aligned}(z_1, \dots, z_{2r+2q+1}) \mapsto \\ (e^{it(a_1+1)/2} z_1, e^{it(a_1-1)/2} z_2, \dots, e^{it(a_q+1)/2} z_{2q-1}, e^{it(a_q-1)/2} z_{2q}, \\ z_{2q+1}, e^{-it} z_{2q+2}, \dots, e^{-it} z_{2r+2q-2}, z_{2r+2q-1}, z_{2r+2q}, e^{-it} z_{2r+2q+1}).\end{aligned}$$

Note that here t runs through $[0, 2\pi]$, hence the $\frac{1}{2}$–factors. Because all a_i are odd, γ'' is now a closed loop in G, not only in $G/\pm I$. So by [32], its Seidel element $S(\gamma'') \in \mathrm{QH}^*(X)$ can be computed as the quantum product of powers of the divisors corresponding to the coordinates on $\mathbb{C}^{2r+2q+1}$, where the powers are the multiplicities of rotations. Given $\mathrm{char}\,\mathbb{K} = 2$, and recalling that $S(\gamma'') = S(\gamma') = S(\gamma)$, we get

$$\begin{aligned} S(\gamma) &= (x+y)^{(a_1+1)/2}(x+y)^{(a_1-1)/2}\cdots(x+y)^{(a_q+1)/2}(x+y)^{(a_q-1)/2}y^{-1}\cdots y^{-1} \\ &= (x+y)^p y^{-r}. \end{aligned}$$

This element squares to 1 by (3-3) (in agreement with the fact γ has order 2 in $\pi_1(G/\pm I) \cong \mathbb{Z}/2$), so it also equals $y^r(x+y)^{-p}$, which proves Lemma 3.6. □

We conclude the proof of Theorem 3.3(b). By Lemmas 3.4 and 3.6, $\ker \mathcal{F}$ is the ideal generated by $S(\gamma)+1$. Suppose $P \in \mathrm{QH}^*(X)$ such that $\mathcal{CO}^*(P) = 0 \in \mathrm{HH}^*(L, L)$. Then $\mathcal{CO}^0(P) = 0$, so $P \in \ker \mathcal{F}$ by Theorem 1.13. Consequently, $P = (S(\gamma)+1) * Q$, and if $P \neq 0$, then $Q \notin \ker \mathcal{F}$ (because otherwise we would get $P \in (\ker \mathcal{F})^2 = \{0\}$). Apply Theorem 1.7(b) to the product $(S(\gamma)+1) * Q$; the left-hand side of $(**)$ vanishes because μ^2 is commutative on $\mathrm{HF}^*(L, L)$ [23], and the right-hand side is nontrivial for some y by Lemma 3.5 and because $\mathcal{CO}^0(Q) \neq 0$. It follows that $\mathcal{CO}^*(P) \neq 0$. We have shown that $\mathcal{CO}^*$ is injective, and split-generation follows from Theorem 1.10. Note that $w(L) = 0$ holds for all real Lagrangians, as Maslov index 2 disks come in pairs because of the action of the antiholomorphic involution; see [23]. □

The following corollary in particular implies Proposition 1.4 from the introduction.

Corollary 3.7 *Let $X = \mathrm{Bl}_{\mathbb{C}P^{2q-1}} \mathbb{C}P^{2r+2q-1}$, and let $L \subset X$ be the real Lagrangian (diffeomorphic to $\mathrm{Bl}_{\mathbb{R}P^{2q-1}} \mathbb{R}P^{2r+2q-1}$). Assume $\gcd(2q+1, 2r-2q) \geq 2$ and that either r or q are odd. Then $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HF}^*(L, L)$ is injective, although $\mathcal{CO}^0$ is not. Consequently, L split-generates $\mathcal{F}uk(X)_0$.*

Proof of Proposition 1.4 Take X as in (3-1) with $a_1 = \cdots = a_k = 1$. Then $X = \mathrm{Bl}_{\mathbb{C}P^{k-1}} \mathbb{C}P^n$; see eg [17, Proposition 11.14]. The additional hypotheses of the current corollary make sure X satisfies all conditions of Theorem 3.3(b), which together with the split-generation criterion (Theorem 1.10(b)) implies the corollary. □

In order to deduce nondisplaceability results between the real Lagrangian L and other Lagrangians with arbitrary obstruction numbers, we need the following lemma.

Lemma 3.8 *Suppose* $\operatorname{char}\mathbb{K} = 2$, L *is an object of* $\mathcal{F}uk(X)_w$ *and* $\mathcal{CO}^*\colon \mathrm{QH}^*(X) \to \mathrm{HH}^*(L,L)$ *is injective. Assuming Hypothesis 3.9 below,* $L \times L$ *split-generates* $\mathcal{F}uk(X \times X)_0$.

Note that by Lemma 1.9, the condition of Lemma 3.8 can only hold if $\mathrm{QH}^*(X) = \mathrm{QH}^*(X)_w$ or L is nonorientable.

Proof First, observe that $w(L \times L) = 2w(L) = 0$. By [45], the injectivity of $\mathcal{CO}^*$ is equivalent to the fact that the open-closed map $\mathcal{OC}^*\colon \mathrm{HH}_*(L,L) \to \mathrm{QH}^*(X)$ hits the unit $1 \in \mathrm{QH}^*(X)$.

Hypothesis 3.9 *There is a commutative diagram*

$$\begin{array}{ccc} \mathrm{HH}_*(L,L) \otimes \mathrm{HH}_*(L,L) & \longrightarrow & \mathrm{HH}_*^{\mathrm{split}}(L \times L, L \times L) \\ \downarrow{\scriptstyle \mathcal{OC}^* \otimes \mathcal{OC}^*} & & \downarrow{\scriptstyle \mathcal{OC}^*_{\mathrm{prod}}} \\ \mathrm{HF}^*(X) \otimes \mathrm{HF}^*(X) & \xrightarrow{=} & \mathrm{HF}^*(X \times X) \end{array}$$

where $\mathcal{OC}^*_{\mathrm{prod}}$ *is the open-closed map on the product, and the* HF^* *are Hamiltonian Floer cohomologies, isomorphic to the quantum cohomologies of the corresponding spaces.*

Here $\mathrm{HH}_*^{\mathrm{split}}(L \times L, L \times L)$ indicates that the A_∞ structure on $L \times L$ is computed using a split Hamiltonian perturbation and a product almost complex structure; such a choice can be made regular. We expect the hypothesis to hold following Ganatra [22, Remark 11.1], who stated it on chain level, in the setup of the wrapped Fukaya category of an exact manifold. A slight complication is that the A_∞ algebra of $L \times L$ appearing in the top right corner of the diagram had to be equipped with so-called one-sided homotopy units; their presence is denoted by a tilde in [22, Remark 11.1]. This does not affect the diagram on the homology level [22, Proposition 10.10], but we have not checked how this subtlety carries over to the monotone setup; therefore we leave Hypothesis 3.9 as a conjecture.

Given Hypothesis 3.9, if $\mathcal{OC}^*$ hits the unit, then $\mathcal{OC}^* \otimes \mathcal{OC}^*$ and $\mathcal{OC}^*_{\mathrm{prod}}$ also do. The latter fact implies that $\mathcal{CO}^*$ is injective on the product, and split-generation follows from Theorem 1.10(b). □

Corollary 1.5 from the introduction is a particular case of the following.

Corollary 3.10 *Let $\mathbb{K}$ be a field of characteristic 2 and $L \subset X$ be as in Theorem 3.3(a) or (b), or as in Corollary 3.7. Suppose $L' \subset X$ is another monotone Lagrangian, perhaps equipped with a local system $\pi_1(L) \to \mathbb{K}^\times$, with minimal Maslov number at least 2 and such that $\mathrm{HF}^*(L', L') \neq 0$. If $w(L') \neq 0$, assume Hypothesis 3.9. Then $L \cap L' \neq \varnothing$.*

Proof If $w(L') = 0$, this follows from the fact that L split-generates $\mathcal{F}uk(X)_0$ and from Lemma 1.11. If $w(L') \neq 0$, we have that $w(L' \times L') = 2w(L') = 0$, so $L' \times L'$ is an object of $\mathcal{F}uk(X \times X)_0$ which is split-generated by $L \times L$ by Lemma 3.8. Then $(L \times L) \cap (L' \times L') \neq 0$ by Lemma 1.11, and so $L \cap L' \neq 0$. □

3.3 An application to nonformality

Recall that if $\mathcal{A} \to \mathcal{A}'$ is a quasi-isomorphism of A_∞ categories, it induces an isomorphism $\mathrm{HH}^*(\mathcal{A}) \to \mathrm{HH}^*(\mathcal{A}')$; see eg Seidel [42, (1.14)]. We will need an explicit chain-level formula for this isomorphism, which can be obtained by combining Seidel's argument with Ganatra's functoriality formulas [22, Section 2.9], and this requires a short account. We are assuming the reader is familiar with the basic language of A_∞ categories from eg [40; 45; 22], so that we can skip some basic definitions and present the other ones rather informally. For simplicity, we are working with $\mathrm{char}\,\mathbb{K} = 2$ so we won't have to worry about signs, and we restrict to A_∞ algebras rather than categories.

Recall that if $\mathcal{A}$ is an A_∞ algebra, its Hochschild cohomology $\mathrm{HH}^*(\mathcal{A})$ can be seen as Hochschild cohomology $\mathrm{HH}^*(\mathcal{A}, \mathcal{A})$ of $\mathcal{A}$ as an $\mathcal{A}$–$\mathcal{A}$ bimodule. If $F\colon \mathcal{A} \to \mathcal{A}'$ is a quasi-isomorphism between A_∞ algebras, it induces quasi-isomorphisms

$$\mathrm{CC}^*(\mathcal{A}, \mathcal{A}) \xrightarrow{F_*} \mathrm{CC}^*(\mathcal{A}, F^*\mathcal{A}') \xleftarrow{F^*} \mathrm{CC}^*(\mathcal{A}', \mathcal{A}'), \tag{3-7}$$

which proves that $\mathrm{HH}^*(\mathcal{A}, \mathcal{A}) \cong \mathrm{HH}^*(\mathcal{A}', \mathcal{A}')$. Chain-level formulas for the two intermediate quasi-isomorphisms, which we will now recall, were written down eg by Ganatra [22, Section 2.9] (in the context of Hochschild homology, but these are easily adjusted to cohomology).

If $\mathcal{B}$, $\mathcal{B}'$ are two $\mathcal{A}$–$\mathcal{A}$ bimodules, a morphism $G\colon \mathcal{B} \to \mathcal{B}'$ is a sequence of maps $G^k\colon \mathcal{A}^{\otimes i} \otimes \mathcal{B} \otimes \mathcal{A}^{\otimes j} \to \mathcal{B}'$, $i + j + 1 = k$, satisfying a sequence of relations which we informally write down as

$$\textstyle\sum_\star G^\star(\mathrm{Id}^{\otimes\star} \otimes \mu^\star_{\mathcal{A}\text{ or }\mathcal{B}} \otimes \mathrm{Id}^{\otimes\star}) = \sum_\star \mu^\star_{\mathcal{B}'}(\mathrm{Id}^{\otimes\star} \otimes G^\star \otimes \mathrm{Id}^{\otimes\star}).$$

Here $\star$ are positive integers which are mutually independent but are such that the total number of inputs on both sides of the equation is the same, the sum is over all such

possibilities, and the structure map on the left is $\mu^\star_{\mathcal{A}}$ or $\mu^\star_{\mathcal{B}}$ depending on whether one of its arguments is in $\mathcal{B}$. In its full form, the above relation should be written as follows:

$$\begin{aligned}
&\sum G^{i_1+i_4-i_3+1}(a_1,\dots,a_{i_1},\mu_{\mathcal{B}}^{i_3-i_1+1}(a_{i_1+1},\dots,a_{i_2},b,a_{i_2+1},\dots,a_{i_3}),a_{i_3+1},\dots,a_{i_4})\\
&+\sum G^{i_1+i_4-i_2+2}(a_1,\dots,a_{i_1},\mu_{\mathcal{A}}^{i_2-i_1}(a_{i_1+1},\dots,a_{i_2}),a_{i_2+1},\dots,a_{i_3},b,a_{i_3+1},\dots,a_{i_4})\\
&+\sum G^{i_2+i_4-i_3+2}(a_1,\dots,a_{i_1},b,a_{i_1+1},\dots,a_{i_2},\mu_{\mathcal{A}}^{i_3-i_2}(a_{i_2+1},\dots,a_{i_3}),a_{i_3+1},\dots,a_{i_4})\\
&=\sum \mu_{\mathcal{B}'}^{i_1+i_4-i_3+1}(a_1,\dots,a_{i_1},G^{i_3-i_1+1}(a_{i_1+1},\dots,a_{i_2},b,a_{i_2+1},\dots,a_{i_3}),a_{i_3+1},\dots,a_{i_4})
\end{aligned}$$

for $a_i \in \mathcal{A}$ and $b \in \mathcal{B}$, where the sums are all over $i_1+i_2+i_3+i_4 = l$. We will keep the informal style of notation, in which the inputs are omitted and the valencies are replaced by $\star$. The induced map $G_*\colon \mathrm{CC}^*(\mathcal{A},\mathcal{B}) \to \mathrm{CC}^*(\mathcal{A},\mathcal{B}')$ is defined by

$$(G_*(h))^\star = \textstyle\sum_\star G^\star(\mathrm{Id}^{\otimes\star} \otimes h^\star \otimes \mathrm{Id}^{\otimes\star}), \tag{3-8}$$

where $h^\star\colon \mathcal{A}^{\otimes\star} \to \mathcal{B}$ and $(G_*(h))^\star\colon \mathcal{A}^{\otimes\star} \to \mathcal{B}'$. If G is a quasi-isomorphism, so is G_*.

If $\mathcal{A}$, $\mathcal{A}'$ are two A_∞ algebras, a morphism $F\colon \mathcal{A} \to \mathcal{A}'$ is a sequence of maps $F^\star\colon \mathcal{A}^{\otimes\star} \to \mathcal{A}'$ such that

$$\textstyle\sum_\star \mu^\star_{\mathcal{A}'}(F^\star \otimes \cdots \otimes F^\star) = \sum_\star F^\star(\mathrm{Id}^{\otimes\star} \otimes \mu^\star_{\mathcal{A}} \otimes \mathrm{Id}^{\otimes\star}).$$

Next, if $\mathcal{B}$ is an $\mathcal{A}'-\mathcal{A}'$ bimodule, its two-sided pullback $F^*\mathcal{B}$ is an $\mathcal{A}$–$\mathcal{A}$ bimodule based on the same vector space $\mathcal{B}$, whose structure maps [22, Section 2.8] are

$$\mu^\star_{F^*\mathcal{B}} = \textstyle\sum_\star \mu^\star_{\mathcal{B}}(F^\star \otimes \cdots \otimes F^\star \otimes \mathrm{Id}_{\mathcal{B}} \otimes F^\star \otimes \cdots \otimes F^\star). \tag{3-9}$$

There is also a morphism $F^*\colon \mathrm{CC}^*(\mathcal{A}',\mathcal{B}) \to \mathrm{CC}^*(\mathcal{A},F^*\mathcal{B})$ defined by

$$(F^*(h))^\star = \textstyle\sum_\star h^\star(F^\star \otimes \cdots \otimes F^\star), \tag{3-10}$$

where $h^\star\colon (\mathcal{A}')^{\otimes\star} \to \mathcal{B}$ and $(F^*(h))^\star\colon \mathcal{A}^{\otimes\star} \to \mathcal{B}$. The total number of inputs here can be zero, and $F^*(h)^0 = h^0$. If F is a quasi-isomorphism, so is F^*.

If, again, $F\colon \mathcal{A} \to \mathcal{A}'$ is a morphism of A_∞ algebras, let $F^*\mathcal{A}'$ be the $\mathcal{A}$–$\mathcal{A}$ bimodule which is the pullback of $\mathcal{A}'$ seen as an $\mathcal{A}'-\mathcal{A}'$ bimodule.

Lemma 3.11 *The same sequence of maps $F^\star\colon \mathcal{A}^{\otimes\star} \to \mathcal{A}'$ provides a morphism of $\mathcal{A}$–$\mathcal{A}$ bimodules $\mathcal{A} \to F^*\mathcal{A}'$, also denoted by f.*

Proof We must check $\sum_\star F^\star(\mathrm{Id}^{\otimes\star}\otimes\mu^\star_{\mathcal{A}}\otimes\mathrm{Id}^{\otimes\star}) = \sum_\star \mu^\star_{F^*\mathcal{A}}(\mathrm{Id}^{\otimes\star}\otimes F^\star\otimes\mathrm{Id}^{\otimes\star})$. If we apply formula (3-9) to rewrite the right-hand sum, the unique Id–factor in (3-9), which in our case is $\mathrm{Id}_{\mathcal{A}'}$, gets applied to the $F^\star$–factor. So our right-hand sum equals $\sum_\star \mu^\star_{\mathcal{A}'}(F^\star\otimes\cdots\otimes F^\star\otimes\cdots\otimes F^\star)$, which is exactly the condition that F is a morphism of A_∞ algebras $\mathcal{A}\to\mathcal{A}'$. □

This lemma explains the precise meaning of (3-7): if $F\colon \mathcal{A}\to\mathcal{A}'$ is a morphism of A_∞ algebras, then the first map F_* from (3-7) is the pushforward of F considered as a morphism of modules $\mathcal{A}\to F^*\mathcal{A}'$ given by formula (3-8). The second map in (3-7) is the pullback as in (3-10).

Next, recall that a general property of quasi-isomorphisms between A_∞ algebras (bimodules, etc.) is that they have quasi-inverses [40, Chapter 1]. If $\mathcal{F}\colon\mathcal{A}\to\mathcal{A}'$ is a quasi-isomorphism, then the $\mathcal{A}$–$\mathcal{A}$ bimodule morphism F from Lemma 3.11 is also a quasi-isomorphism; hence there is an $\mathcal{A}$–$\mathcal{A}$ bimodule quasi-isomorphism $G\colon F^*\mathcal{A}'\to\mathcal{A}$ which is a quasi-inverse of F, so we have quasi-isomorphisms

$$\mathrm{CC}^*(\mathcal{A},\mathcal{A}) \xleftarrow{G_*} \mathrm{CC}^*(\mathcal{A},F^*\mathcal{A}') \xleftarrow{F^*} \mathrm{CC}^*(\mathcal{A}',\mathcal{A}'). \tag{3-11}$$

Their composition acts on Hochschild cochains by

$$(G_*F^*(h))^\star = \textstyle\sum_\star G^\star(\mathrm{Id}^{\otimes\star}\otimes h^\star(F^\star\otimes\cdots\otimes F^\star)\otimes\mathrm{Id}^{\otimes\star}),$$

where $h^\star\colon(\mathcal{A}')^{\otimes\star}\to\mathcal{A}'$ and $(G_*F^*(h))^\star\colon\mathcal{A}^{\otimes\star}\to\mathcal{A}$. In particular, $(G_*F^*(h))^0 = G^1(h^0)$, and if $h^0=0\in\mathcal{A}'$, then

$$(G_*F^*(h))^1(u) = G^1(h^1(F^1(u))),\quad u\in\mathcal{A}'. \tag{3-12}$$

Note that $G^1\colon\mathcal{A}'\to\mathcal{A}$ and $F^1\colon\mathcal{A}\to\mathcal{A}'$ are chain maps with respect to $\mu^1_{\mathcal{A}}$ and $\mu^1_{\mathcal{A}'}$, and they are cohomology inverses of each other.

Assume $L\subset X$ is a Lagrangian preserved by a Hamiltonian loop γ which together satisfy the conditions of Theorem 1.7 (those conditions which are common to all parts of the theorem). Assume the A_∞ algebra $\mathrm{CF}^*(L,L)$ is formal, ie there is an A_∞ quasi-isomorphism $F\colon\mathrm{HF}^*(L,L)\to\mathrm{CF}^*(L,L)$. Let

$$h := (G_*F^*)(\mathcal{CO}^*(S(\gamma))-1)\in\mathrm{CC}^*(\mathrm{HF}^*(L,L),\mathrm{HF}^*(L,L)), \tag{3-13}$$

where G is a quasi-inverse of F, and G_* and F^* are as in (3-11). So h is a Hochschild cochain for the associative algebra $\mathrm{HF}^*(L,L)$. Then by Corollary 2.4, $h^0=0$, Proposition 2.8 and (3-12), we have

$$h^1(x) = \rho(l)\cdot\langle\Psi(F^1(x)),l\rangle\cdot G^1(1_L).$$

(We have dropped the extra sign, working in characteristic two.) Let us additionally assume that L is *wide* [9, Definition 1.2.1], ie there is a vector space isomorphism between $H^*(L)$ and $\mathrm{HF}^*(L,L)$, and that L admits a perfect Morse function. These conditions enable us to identify $\mathrm{CF}^*(L,L) \cong \mathrm{HF}^*(L,L)$ as vector spaces. Because G^1 is cohomologically unital, $G^1(1_L) = 1_L \in \mathrm{HF}^*(L,L)$, so

$$h^1(x) = \rho(l) \cdot \langle \Psi(F^1(x)), l \rangle \cdot 1_L \in \mathrm{HF}^*(L,L). \tag{3-14}$$

Under our identifications, Ψ becomes an isomorphism between the vector spaces below, and F^1 can be considered as an algebra isomorphism from $\mathrm{HF}^*(L,L)$ to itself:

$$\mathrm{HF}^*(L,L) \xrightarrow[\text{algebra iso.}]{F^1} \mathrm{HF}^*(L,L) \xrightarrow[\text{linear iso.}]{\Psi} H^*(L). \tag{3-15}$$

We now turn the discussion to Hochschild cohomology of monic algebras. Let $f(u) \in \mathbb{K}[u]$ be a polynomial and $A := \mathbb{K}[u]/(f)$ the quotient algebra; it is called a monic algebra. This is an algebra in the ordinary associative sense; we can consider it as an A_∞ algebra by equipping it with trivial higher-structure maps. The Hochschild cohomology algebra $\mathrm{HH}^*(A)$ was computed by Holm [24]. Recall that Hochschild cohomology of ungraded associative algebras is $\mathbb{Z}$–graded (unlike Hochschild cohomology of non-$\mathbb{Z}$–graded A_∞ algebras): cochains $A^{\otimes k} \to A$ are said to have degree k, and the differential has degree 1. By [24, Proposition 2.2],

$$\mathrm{HH}^k(A) = \begin{cases} A & \text{if } k = 0, \\ \mathrm{Ann}_A(f') & \text{if } k > 0 \text{ is odd}, \\ A/(f') & \text{if } k > 0 \text{ is even}. \end{cases}$$

Although we already know the answer, let us compute $\mathrm{HH}^1(A)$ explicitly, as this will be helpful later.

Lemma 3.12 *For a Hochschild cocycle $h\colon A \to A$ in $\mathrm{CC}^1(A)$, we must have*

$$h(u^m) = amu^{m-1} \tag{3-16}$$

for some fixed $a \in A$. Note that $a = h(u)$.

Proof Let us compute the Hochschild differential $\partial h\colon A \otimes A \to A$ on the two elements u and u^{m-1} for some $m \in \mathbb{N}$. Because ∂h vanishes by assumption, we get

$$0 = (\partial h)(u, u^{m-1}) = uh(u^{m-1}) + u^{m-1}h(u) + h(u^m).$$

The desired formula follows by induction on m. □

So any cocycle $h \in \mathrm{CC}^1(A)$ is completely determined by a single element $a = h(u) \in A$, and h must meet an additional condition that $h(f(u)) = h(0) = 0$, which is equivalent to $a \in \mathrm{Ann}_A(f')$. As the differential $\mathrm{CC}^0(A) \to \mathrm{CC}^1(A)$ vanishes, we get an isomorphism $\mathrm{HH}^1(A) \to \mathrm{Ann}_A(f')$, $h \mapsto h(u)$.

We will further assume that $\mathrm{char}\,\mathbb{K} = 2$ and $f' = 0$. The latter condition means that f is a sum of even powers of u. Denote by

$$\psi\colon \mathrm{HH}^1(A) \to A$$

the isomorphism $\phi(h) = h(u)$ from above. Note that if $s(u) \in A$ is an arbitrary element given by a polynomial with derivative $s'(u)$, then by (3-16), we get

$$\psi(h) = s'(u) \cdot h(s(u)). \tag{3-17}$$

For $k > 1$, we also have isomorphisms $\psi\colon \mathrm{HH}^k(A) \to A$, all of which we denote by the same letter by abusing notation; we will not need an explicit formula for these isomorphisms when $k > 1$.

Moreover, [24, Lemma 5.1] computes the Yoneda product on $\mathrm{HH}^*(A)$. In particular, given $h_1, h_2 \in \mathrm{HH}^1(A)$, their Yoneda product $h_1 \star h_2$ is determined by

$$\psi(h_1 \star h_2) = \psi(h_1) \cdot \psi(h_2) \sum_{j \text{ odd}} f_{2j} u^{2j-2} \in A, \tag{3-18}$$

where $f = \sum_j f_j u^j$, $f_j \in \mathbb{K}$.

The two strands of discussion can be combined in the following theorem.

Theorem 3.13 *Let $\mathbb{K}$ be a field of characteristic 2 and $L \subset X$ a Lagrangian preserved by a Hamiltonian loop γ which together satisfy the conditions of Theorem 1.7 (those conditions which are common to all parts of the theorem). Assume there is an algebra isomorphism $\mathrm{HF}^*(L,L) \cong \mathbb{K}[u]/(f)$, where $f(u) = \sum_{j \geq 0} f_j u^j$ is a polynomial, and also that L is wide and admits a perfect Morse function, so that we can identify the vector spaces $\mathrm{HF}^*(L,L) \cong \mathrm{CF}^*(L,L)$, and $\Psi\colon \mathrm{HF}^*(L,L) \to H^*(L)$ becomes an isomorphism of vector spaces. Further, assume that*

- *$f' = 0$ and $\sum_{j \text{ odd}} f_{2j} u^{2j-2}$ is invertible in $\mathbb{K}[u]/(f)$,*
- *$\langle \Psi(r(u)), l \rangle = 1$ for an element $r(u) \in \mathbb{K}[u]/(f) \cong \mathrm{HF}^*(L,L)$ which generates $\mathrm{HF}^*(L,L)$ as an algebra,*
- *$S(\gamma)^2 = 1 \in \mathrm{QH}^*(X)$.*

Then the Fukaya A_∞ algebra of L is not formal over $\mathbb{K}$.

Proof Supposing $\mathrm{CF}^*(L,L)$ is formal, let h be as in (3-13) and F^1 as in (3-15). Then there exists $s(u) \in \mathrm{HF}^*(L,L)$ (we view this element as a polynomial in $\mathbb{K}[u]/(f)$) such that $F^1(s(u)) = r(u)$. Then by (3-14), $h^1(s(u)) = \rho(l)\cdot 1 \in \mathrm{HF}^*(L,L)$, so by (3-17),

$$\psi(h^1) = \rho(l)\cdot s'(u) \in \mathrm{HF}^*(L,L).$$

Further, note that $h \star h = 0$ because $(S(\gamma)+1)^2 = 0$, so (3-18) yields

$$\rho(l)^2 \cdot (s'(u))^2 \sum_{j \text{ odd}} f_{2j} u^{2j-2} = 0 \in \mathrm{HF}^*(L,L).$$

By hypothesis, this implies $(s'(u))^2 = 0$, so $s'(u) \in \ker \mathcal{F}$, where $\mathcal{F}\colon \mathbb{K}[u]/(f) \to \mathbb{K}[u]/(f)$ is the Frobenius endomorphism. In general, over $\operatorname{char}\mathbb{K} = 2$, it is always true that $s'(u)$ is a sum of even powers of u, so $s'(u)$ is a square of another polynomial: $s'(u) = (t(u))^2$. Then $t(u)^2 \in \ker\mathcal{F}$, which implies $t(u) \in \ker\mathcal{F}$ because $\ker\mathcal{F}$, being an ideal in $\mathbb{K}[u]/(f)$, is necessarily prime. Consequently, $s'(u) = 0$. So $s(u)$ is a sum of even powers of u, hence the subalgebra generated by $s(u)$ lies in the subalgebra of $\mathbb{K}[u]/(f)$ generated by u^2, which is smaller than the whole $\mathbb{K}[u]/(f)$: for example, it does not contain the element u. (Recall that f is also a sum of even powers of u.) On the other hand, we know that F^1 is an algebra isomorphism, $F^1(s(u)) = r(u)$ and $r(u)$ generates the whole $\mathrm{HF}^*(L,L)$ by hypothesis. So $s(u)$ should also generate $\mathrm{HF}^*(L,L)$, which is a contradiction. □

Proof of Proposition 1.3 Take the real loop γ preserving $\mathbb{R}P^{4n+1}$ defined in the proof of Proposition 1.1, and let $L = \mathbb{R}P^{4n+1}$ and $X = \mathbb{C}P^{4n+1}$. Recall that if $x \in H^2(X)$ is the generator, then $\mathrm{QH}^*(X) \cong \mathbb{K}[x]/(x^{4n+2}+1)$ and $S(\gamma) = x^{2n+1}$, so $S(\gamma)^2 = 1$. Also recall that $l \in H_1(L) \cong \mathbb{K}$ is nonzero. By Theorem 1.12, $\mathrm{HF}^*(L,L) \cong \mathbb{K}[u]/(u^{4n+2}+1)$, where $u \in \mathrm{CF}^1(L,L) \cong \mathbb{K}$, and we have $\langle \Psi(u), l\rangle = 1$. Now apply Theorem 3.13 taking $r(u) = u$ to conclude the proof. □

Proposition 3.14 *Let X be $\mathrm{Bl}_{\mathbb{C}P^{2q-1}} \mathbb{C}P^{2r+2q-1}$ and $L \subset X$ the real Lagrangian (diffeomorphic to $\mathrm{Bl}_{\mathbb{R}P^{2q-1}} \mathbb{R}P^{2r+2q-1}$). Assume $\gcd(2q+1, 2r-2q) \geq 2$ and either r or q is odd. Then the A_∞ algebra of L is not formal over a characteristic-2 field.*

Proof We recall that all real Lagrangians are wide by Theorem 1.12 and admit a perfect Morse function by [23]. The fact that $\gcd(2q+1, 2r-2q) \geq 2$ means we are in the situation of Theorem 3.3(b) with $k = 2q = 2p$ and $a_1 = \cdots = a_k = 1$. We have already seen by (3-5) that $\mathrm{HF}^*(L,L) \cong \mathbb{K}[u]/(f)$ with $f' = 0$, and it is easy to check

that $\sum_{j \text{ odd}} f_{2j} u^{2j-2}$ is invertible provided that either r or q is odd (otherwise this element would vanish). Moreover, via (3-4) and Haug's isomorphism (Theorem 1.12), $u^{p/g}$ corresponds to the generator of $\mathrm{CF}^1(L,L) \cong \mathbb{K}^2$ such that $\langle \Psi(u), l \rangle = 1$. Now apply Theorem 3.13 taking $r(u) = u^{p/g}$. □

3.4 Nonformality of the equator on the sphere

Proposition 1.3 says in particular that the A_∞ algebra of an equatorial circle on S^2 is not formal over $\operatorname{char} \mathbb{K} = 2$. This is an especially simple case which can be verified by hand, and it is worth discussing it in more detail. Let $L_1 \subset S^2$ be a fixed equator and $L_2, L_3, \dots$ a sequence of its small Hamiltonian perturbations; assume $|L_i \cap L_j| = 2$ for each i, j. Then $\mathrm{CF}^0(L_i, L_j) \cong \mathbb{K}$ is generated by an element which we denote by 1 (this is the cohomological unit), and $\mathrm{CF}^1(L_i, L_j) \cong \mathbb{K}$ is generated by an element which we denote by u (we use the same letter for all i, j). Of the two intersection points $L_i \cap L_j$, the point u is the one at which $T_u L_j$ is obtained from $T_u L_i$ by a small positive rotation with respect to the ω–induced orientation on S^2. Consider the A_∞ structure maps between the consecutive Lagrangians:

$$\mu^k \colon \mathrm{CF}^*(L_k, L_{k+1}) \otimes \cdots \otimes \mathrm{CF}^*(L_1, L_2) \to \mathrm{CF}^*(L_1, L_{k+1}) \tag{3-19}$$

given by counting immersed polygons as in [40]. These give a model for the A_∞ algebra of L, because all the L_i differ by small perturbations, and we can canonically identify the spaces $\mathrm{CF}^*(L_i, L_{i+1})$ with each other. The A_∞ maps will depend on the particular arrangement of the L_i, although up to quasi-isomorphism they give the same A_∞ algebra.

Remark 3.15 The fact the A_∞ algebra of L defined using the count of polygons is quasi-isomorphic to the one defined using Hamiltonian perturbations seems not to have been written down in detail, but it is widely accepted. An approach is sketched in [41, Remark 7.2] and also performed in [44] in a slightly different setup.

Let us compute some of the A_∞ structure maps using a specific choice of the L_i. Fix a Hamiltonian H whose flow is the rotation of $S^2 \subset \mathbb{R}^3$ around an axis which is not orthogonal to the plane intersecting S^2 along the equator L_1. Let $L_2, L_3, \dots$ be obtained from L_1 by applying that rotation by small but consecutively increasing angles, ie L_i are time-t_i pushoffs of L_1 under the flow of H for $0 = t_1 < t_2 < t_3 < \cdots$. The first four resulting circles L_i are represented in Figure 8, left. The pairwise intersections of the L_i are contained in two opposite patches of the sphere; those

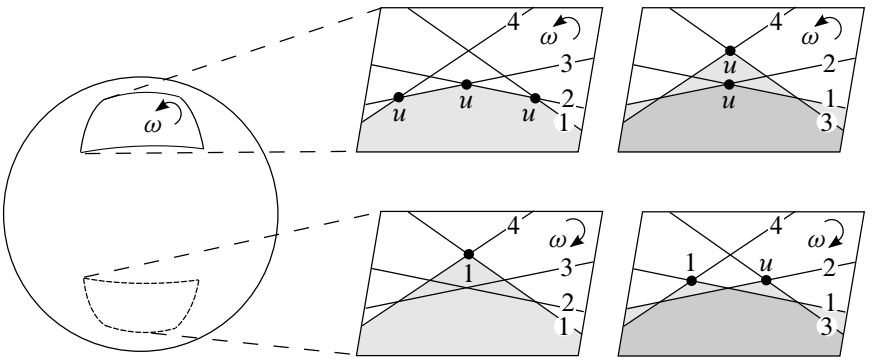


Figure 8: Two different configurations (left) and (right) consisting of four small Hamiltonian pushoffs $L_1, \dots, L_4$ (marked by numbers) of an equatorial circle on S^2. The image of the disk contributing to $\mu^3(u,u,u) = 1$ is shaded.

patches are shown in the top and bottom of Figure 8, left, together with the L_i on them, which are depicted by straight lines. Both patches are drawn as if we look at them from the same point "above" the sphere, so that the positive rotation (with respect to the orientation on S^2) is counterclockwise on the upper patch and clockwise on the lower patch. For this particular choice of perturbations, and for each $i < j$, all degree-one points $u \in \mathrm{CF}^*(L_i, L_j)$ are located on the upper patch, and all points $1 \in \mathrm{CF}^*(L_i, L_j)$ are on the lower patch.

We claim that in this model, we get

$$\mu^3(u,u,u) = 1, \quad \mu^3(u,u,1) = 0, \quad \mu^3(u,1,u) = 0, \quad \mu^3(1,u,u) = 0.$$

For grading reasons, $\mu^k(u, \dots, u)$ is a multiple of 1 and is determined by counting Maslov index 2 disks. There is a unique such disk; for $k = 3$, it is shown in grey shade in Figure 8, left, on the two patches; away from the patches, this disc is just a strip between L_1 and L_4. Also for grading reasons, the only other products which can possibly be nontrivial are $\mu^k(u, \dots, u, 1, u, \dots, u) \in \{0, u\}$, where exactly one input is 1. It possible to check that these vanish for our configuration of the circles L_i, at least when $k = 3$. Now note that

$$\mu^2(1+u, 1+u) = 0, \quad \mu^3(1+u, 1+u, 1+u) = 1.$$

The latter equality exhibits a nontrivial Massey product seen as a well-defined element of

$$\mathbb{K}[u]/(1+u) \cong \mathbb{K}.$$

An explanation of where the Massey products generally belong is found in Remark 1.2 of [40], which explains the quotient by $1+u$ above. The presence of a nontrivial Massey product is invariant under quasi-isomorphisms. To see this, recall that the analogous fact for dg algebras is easy, and any A_∞ algebra is quasi-isomorphic to a dg algebra. Moreover, the Massey products for the A_∞ and dg models satisfy a simple relation [31, Theorem 3.1 and Corollary A.5]; in particular, if triple Massey products of an A_∞ algebra are nontrivial, they remain nontrivial for its dg model. This gives us an alternative proof of the fact that the A_∞ algebra of the equator on S^2 is not formal.

For any other arrangement of the L_i, we will necessarily have $\mu^3(1+u,1+u,1+u)=1$ modulo $1+u$ because of invariance of Massey products, meaning that

$$\mu^3(1+u,1+u,1+u)\in\{1,u\}.$$

For example, another possible configuration of $L_1,\dots,L_4$ is shown in Figure 8, right; it is simply obtained from the earlier configuration by changing the ordering of the L_i. In this new model, the maps μ^k from (3-19) are now

$$\mu^3(u,u,u)=1,\quad \mu^3(u,u,1)=0,\quad \mu^3(u,1,u)=u,\quad \mu^3(1,u,u)=u.$$

The unique disk contributing to $\mu^3(u,u,u)$ is shown in Figure 8, right, by grey shade. It is an immersed disk, and the domain over which it self-overlaps has darker shade. Note that in this model, the degree-one generators $u\in \mathrm{CF}^1(L_1,L_2)$, $\mathrm{CF}^1(L_3,L_4)$, $\mathrm{CF}^1(L_1,L_4)$ correspond to the intersection points on the upper patch, and the degree-one generator $u\in\mathrm{CF}^1(L_2,L_3)$ corresponds to the intersection point on the lower patch. We see that we again get $\mu^3(1+u,1+u,1+u)=u$.

The existence of the Massey product above crucially required $\operatorname{char}\mathbb{K}=2$, because otherwise we would not get $\mu^2(1+u,1+u)=0$, which is necessary to speak of the triple Massey product of $1+u$ with itself. If $\operatorname{char}\mathbb{K}\neq 2$, then

$$\mathrm{HF}^*(L_1,L_1)\cong\mathbb{K}[u]/(u^2-1)\cong\mathbb{K}[u]/(u-1)\oplus\mathbb{K}[u]/(u+1)$$

is a direct sum of fields, whose Hochschild cohomology as an ordinary algebra vanishes [27] in degree 2, in contrast to the case where $\operatorname{char}\mathbb{K}=2$. So any A_∞ algebra over $\mathbb{K}[u]/(u-1)\oplus\mathbb{K}[u]/(u+1)$ is formal by [26] or [43, Section 3]; in particular, the A_∞ algebra of the equator on S^2 is formal. For example, the product $\mu^3(1+u,1+u,1+u)$ can be made to vanish after a formal diffeomorphism. Because of the nontrivial Massey product in characteristic 2, such a formal diffeomorphism, say over $\mathbb{Q}$, will necessarily involve division by 2, and cannot be realised by any geometric choice of the pushoffs L_i.

In comparison, the topological A_∞ algebra of the circle is formal over a field of any characteristic. Indeed, the topological A_∞ algebra is $\mathbb{Z}$–graded, so if we make this algebra to be based on the cohomology ring $H^*(S^1) \cong \mathbb{K}[x]/x^2$ where $|x| = 1$, the only possibly nontrivial products will be $\mu^k(x, \dots, x, 1, x, \dots, x)$ for grading reasons. On the other hand, every A_∞ algebra is quasi-isomorphic to a minimal, strictly unital one over a field of any characteristic [40, Lemma 2.1; 30, Theorem 3.1.1]. In a minimal strictly unital model, those products vanish by definition when $k \geq 3$.

4 The closed-open map for monotone toric fibres

4.1 The mechanism of Theorem 1.7 for toric fibres

Let X be an n–dimensional compact toric Fano variety, and $T \subset X$ the unique monotone toric fibre. Evans and Lekili [19] proved (after this paper had appeared as a preprint) that if $\operatorname{char} \mathbb{K} = 0$, the Fukaya category $\mathcal{F}uk(X)_w$ is split-generated by several copies of T, equipped with the local systems corresponding to the critical points of the Landau–Ginzburg superpotential with critical value $w \in \mathbb{K}$. We shall recall the formula for the superpotential of a toric manifold in the next subsection; the common general references are [13; 7; 21].

Prior to [19], the split-generation by toric fibres had been proved only in the case when the superpotential is Morse; see Ritter [37]. (For Ritter, proving split-generation requires considerable effort, even in the Morse case, if W has several critical points with the same critical value. However, the difficulty is mainly related to the fact that he allows some noncompact toric varieties, where the injectivity of $\mathcal{CO}^*$ is no longer a criterion for split-generation and one must look at $\mathcal{OC}^*$ instead. If we work with compact manifolds, checking that $\mathcal{CO}^*$ is injective for an arbitrary Morse potential is easy; see Corollary 4.4.) An example of a toric Fano variety with non-Morse superpotential over $\mathbb{C}$ has been obtained by Ostrover and Tyomkin [35], and one can check that the superpotential in their case has an A_3 singularity.

To complete the literature overview, we should mention the work in progress by Abouzaid, Fukaya, Oh, Ohta and Ono that will prove the split-generation result for toric manifolds that are not necessarily Fano.

Because the toric fibre T is invariant under all the Hamiltonian loops coming from the torus action, it is an obvious example where Theorem 1.7 can be put to the test. It turns out that it does allow one to prove split-generation away from the Morse case, though

not too far from it: the superpotential is required to have at worst A_2 singularities, and the extra condition that $\operatorname{char}\mathbb{K} \neq 2, 3$ is required; see Corollary 4.6.

Although our result is much weaker than the general one from [19], we find it interesting in our approach that the ability to solve equation $(*)$ from Theorem 1.7 depends on whether W is Morse or not. Equip T with a local system ρ which corresponds to a critical point of W; then (T, ρ) is wide, and we can identify the vector spaces $\mathrm{HF}^*(T, \rho) \cong H^*(T)$ via the PSS map Φ. For convenience, let us rewrite equation $(*)$ from Theorem 1.7:

$$(*') \qquad \mu^2(a, y) + \mu^2(y, a) = \rho(l) \cdot \langle y, l \rangle \cdot 1_T \quad \text{for each } y \in H^1(T).$$

We are using the standard spin structure on T, and hence the sign $(-1)^{\epsilon(l)}$ is positive; see Remark 1.8. Recall that Theorem 1.7(b) can be applied if there exists no $a \in \mathrm{HF}^*(T, \rho)$ making $(*')$ hold. The Floer cohomology algebra of (T, ρ) is a Clifford algebra determined by the Hessian of W at the point ρ, so the left-hand side of $(*')$ is equal to $\operatorname{Hess}_\rho W(a, y) \cdot 1_T$, at least when $a \in H^1(T)$; we are using informal notation for the moment. Therefore, finding an element a solving $(*')$ reduces to finding an a such that

$$\operatorname{Hess}_\rho W(a, -) = \text{const} \cdot \langle -, l \rangle. \tag{4-1}$$

The ability to find such an a depends on how degenerate $\operatorname{Hess}_\rho W$ is. If ρ is a Morse point of W, such an a can always be found, so Theorem 1.7(b) does not apply. However, the Morse case can actually be covered by Theorem 1.7(a) as we explain below. On the other hand, when $\operatorname{Hess}_\rho W$ has kernel, we will have some elements $l \in H_1(T)$ for which (4-1) has no solution a. If we consider the S^1–action whose orbit is such an element l, Theorem 1.7(b) can be applied to the Seidel element of this S^1–action to get some new information on $\mathcal{CO}^*$ which is not seen by $\mathcal{CO}^0$. This information turns out to be sufficient only when the superpotential has A_2 singularities, however, there is a possible way of improvement which we speculate upon in the end of this section.

4.2 The results

Recall [12; 13; 21; 20] that the Landau–Ginzburg superpotential of X is a Laurent polynomial $W\colon (\mathbb{K}^\times)^n \to \mathbb{K}$ given by

$$W(x_1, \dots, x_n) = \sum_e \prod_{j=1}^n x_j^{e^j},$$

where the first sum is over the outer normals $e \in \mathbb{Z}^n$ to the facets of the polyhedron defining X, and $e^j \in \mathbb{Z}$ are their coordinates. (Sometimes, the superpotential is written

down with a Novikov parameter, but we can ignore it because we will only be working with the monotone torus T.) We identify $(\mathbb{K}^\times)^n$ with the space of all local systems $H_1(T;\mathbb{Z}) \to \mathbb{K}^\times$. For $\rho \in (\mathbb{K}^\times)^n$, we write (T,ρ) for the torus equipped with this local system. Also, we will abbreviate

$$\mathrm{HF}^*(T,\rho) := \mathrm{HF}^*((T,\rho),(T,\rho)),$$

and the same for Hochschild cohomology. It is known (see [35, Proposition 3.3] for example) that

$$\mathrm{QH}^*(X) \cong \mathbb{K}[x_1^{\pm 1},\dots,x_n^{\pm 1}]/\operatorname{Jac}(W) = \mathcal{O}(Z), \tag{4-2}$$

where the Jacobian ideal $\operatorname{Jac}(W)$ is generated by $(\partial W/\partial x_1,\dots,\partial W/\partial x_n)$, and Z is the subscheme of $\operatorname{Spec}\mathbb{K}[x_1^{\pm 1},\dots,x_n^{\pm 1}]$ defined by the ideal sheaf $\operatorname{Jac}(W)$. Then Z is a 0–dimensional scheme supported at the critical points of W,

$$\{\rho_1,\dots,\rho_q\} = \operatorname{Crit} W, \quad \rho_i \in (\mathbb{K}^\times)^n.$$

The obstruction number of the torus is given by

$$w(T,\rho) = W(\rho) \in \mathbb{K}.$$

Under the isomorphism (4-2), the quantum product is the usual product on $\mathcal{O}(Z)$, and the first Chern class of X is given by the function W itself. The generalised eigenspace decomposition with respect to $-*c_1(X)$ is simply the decomposition into the local rings at the points ρ_i:

$$\mathbb{K}[x_1^{\pm 1},\dots,x_n^{\pm 1}]/\operatorname{Jac}(W) \cong \bigoplus_{\rho_i \in \operatorname{Crit} W} \mathcal{O}_{\rho_i}(Z),$$

the eigenvalue of the ρ_i–summand being the critical value $W(\rho_i)$. From Lemma 1.9, we see that $\mathrm{HF}^*(T,\rho) = 0$ if $\rho \notin \operatorname{Crit} W$. On the other hand, it is known that (T,ρ_i) is wide for $\rho_i \in \operatorname{Crit} W$; ie $\mathrm{HF}^*(T,\rho_i)$ is isomorphic as a vector space to $H^*(T)$.

Lemma 4.1 *Under the isomorphism* (4-2)*, the map $\mathcal{CO}^0\colon \mathrm{QH}^*(X) \to \mathrm{HF}^*(T,\rho_i)$ is given by*

$$\mathcal{CO}^0(f) = f(\rho_i)\cdot 1_T.$$

Here $f(x_1,\dots,x_n) \in \mathrm{QH}^*(X)$, $f(\rho_i) \in \mathbb{K}$ is the value of the function at $\rho_i \in \operatorname{Crit} W$, and $1_T \in \mathrm{HF}^*(T,\rho_i)$ is the unit.

Proof Because $\mathcal{CO}^0$ is a map of algebras, it suffices to prove the lemma when $f = x_k$ is a linear function for $1 \le k \le n$. By [32], $f = S(\gamma)$ for a Hamiltonian loop γ coming from the Hamiltonian torus action such that the value of the local system ρ_i on an orbit of γ equals the kth coordinate ρ_i^k, which is the same as the value $f(\rho_i)$. So $\mathcal{CO}^0(f) = f(\rho_i) \cdot 1_T$ by Theorem 1.7(a). □

Corollary 4.2 *For $\rho_i \ne \rho_j \in \operatorname{Crit} W$, the map $\mathcal{CO}^*|_{\mathcal{O}_{\rho_i}(Z)} \to \operatorname{HH}^*(T, \rho_j)$ vanishes.*

Remark 4.3 If $W(\rho_i) \ne W(\rho_j)$, Corollary 4.2 follows from Lemma 1.9. When $W(\rho_i) = W(\rho_j)$, the statement is implicit in [37, Proof of Theorem 6.17] where it is shown that, dually,

$$\mathcal{OC}\colon \operatorname{HH}_*(T, \rho_j) \to \operatorname{QH}^*(X)$$

hits at most one summand of the form $\mathcal{O}_{\rho_i}$, and we know by Lemma 4.1 that this summand must actually be $\mathcal{O}_{\rho_j}$. The proof in [37] is very different and relies on the variation of the symplectic form.

Proof Let $f \in \mathbb{K}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$ be such that $f(\rho_i) \ne 0$ and $f(\rho_j) = 0$. Then as an element of $\mathcal{O}_{\rho_i}(Z)$, f is invertible. If the corollary does not hold, $\mathcal{CO}^*(f)$ is also invertible. On the other hand, $\mathcal{CO}^0(f) = 0 \in \operatorname{HF}^*(T, \rho_j)$ by Lemma 4.1. The map $\operatorname{HH}^*(T, \rho_j) \to \operatorname{HF}^*(T, \rho_j)$, which takes a Hochschild cochain to its zeroth-order term, is a map of unital algebras by the formula for the Yoneda product and because the Hochschild cohomology unit is represented by a cochain whose zeroth-order term is the Floer cohomology unit (this follows, for example, from the unitality of $\mathcal{CO}^*$). We have determined that f lies in the kernel of $\operatorname{HH}^*(T, \rho_j) \to \operatorname{HF}^*(T, \rho_j)$, but that contradicts the fact that f is invertible. This implies the corollary. □

For $w \in \mathbb{K}$, let

$$\operatorname{Crit}_w W = \{\rho \in \operatorname{Crit} W : W(\rho) = w\}$$

denote the set of all critical points of W with the same critical value w. We will sometimes denote the restrictions of $\mathcal{CO}^0$ and $\mathcal{CO}^*$ to subalgebras of $\operatorname{QH}^*(X)$ by the same symbol when it is otherwise clear that we are considering a restriction.

Corollary 4.4 *If $\operatorname{char} \mathbb{K} \ne 2$, the map*

$$\mathcal{CO}^0\colon \operatorname{QH}^*(X)_w \to \bigoplus_{\rho_i \in \operatorname{Crit}_w W} \operatorname{HF}^*(T, \rho_i)$$

is injective if and only if all points of $\operatorname{Crit}_w W$ are Morse.

Proof By Corollary 4.2, $\mathcal{CO}^0$ is injective if and only if its restrictions $\mathcal{CO}^0\colon \mathcal{O}_{\rho_i}(Z) \to \mathrm{HF}^*(T,\rho_i)$ are injective for each ρ_i. The map $\mathcal{O}_{\rho_i}(Z) \to \mathbb{K}$ which takes $f \in \mathcal{O}_{\rho_i}(Z)$ to its value $f(\rho_i)$ is injective if and only if $\mathcal{O}_{\rho_i}(Z)$ is a field, which is equivalent to the fact that ρ_i is a Morse point of W when $\operatorname{char}\mathbb{K} \neq 2$. Now apply Lemma 4.1. □

Proposition 4.5 *Suppose that* $\operatorname{char}\mathbb{K} \neq 2,3$ *and that* W *has an* A_2 *singularity at a point* ρ*; then* $\mathcal{CO}^*\colon \mathcal{O}_\rho(Z) \to \mathrm{HH}^*(T,\rho)$ *is injective.*

Proof After an integral linear change of coordinates, we may assume that the Hessian of W at ρ is the diagonal matrix $\operatorname{Hess}_\rho W = \operatorname{diag}(1,\dots,1,0)$. We claim that $\mathcal{O}_\rho(Z)$ is generated, as a vector space, by the two elements 1 and x_n, where the linear function x_n corresponds to the kernel of $\operatorname{Hess}_\rho W$. Indeed, after a further nonlinear change of coordinates with the identity linear part, we can bring W to the canonical form

$$W(\widetilde{x}_1,\dots,\widetilde{x}_n) = W(\rho) + \sum_{i=1}^{n-1}(\widetilde{x}_s - \rho^i)^2 \ + (\widetilde{x}_n - \rho^n)^3.$$

Here $\rho^i \in \mathbb{K}$ are the coordinates of ρ. Then

$$\operatorname{Jac}(W) = ((\widetilde{x}_1 - \rho^1),\dots,(\widetilde{x}_{n-1} - \rho^{n-1}),(\widetilde{x}_n - \rho^n)^2),$$

so $\mathcal{O}_\rho(Z)$ is generated, as a vector space, by 1 and $\widetilde{x}_n$. Because x_n, as a function of $\widetilde{x}_1,\dots,\widetilde{x}_n$, equals $\widetilde{x}_n$ plus terms of order at least 2, it is easy to see that the elements 1 and x_n also generate the vector space $\mathcal{O}_\rho(Z)$.

Let us identify $\mathrm{HF}^*(T,\rho)$ with $H^*(T)$ via the PSS map Φ. Recall that, in general, $\mathrm{HF}^*(T,\rho)$ is the algebra generated by $y_1,\dots,y_n \in H^1(T)$ with relations

$$y_p y_q + y_q y_p = \partial^2_{x_p x_q} W(\rho).$$

In particular, in our case, we get $y_p y_n + y_n y_p = 0$ for any $1 \le p \le n$, so $y_n \in \mathrm{HF}^1(T,\rho)$ anticommutes with any element of $\mathrm{HF}^*(T,\rho)$ of odd degree. Consequently, the left-hand side of equation $(*)$ from Theorem 1.7 vanishes if we put $y = y_n$ and allow a to be of arbitrary odd degree.

Returning to our generator $x_n \in \mathcal{O}_\rho(Z)$, we have $x_n = S(\gamma)$ for a Hamiltonian S^1–action (coming from the toric action) such that the element $y_n \in \mathrm{HF}^1(T,\rho)$ is dual to the orbit $l \in H_1(T)$ of γ, so $\langle y_n, l\rangle = 1$. Thus if we put $y = y_n$, the right-hand side of equation $(*)$ from Theorem 1.7 becomes $\rho^n \cdot 1_T \neq 0$. Hence $(*)$ has no solution, and Theorem 1.7(b) says that $\mathcal{CO}^*(x_n)$ and $1_{\mathrm{HH}} = \mathcal{CO}^*(1)$ are linearly independent. □

Combining the above discussion with the split-generation criterion, we get the following:

Corollary 4.6 *Suppose that* $\operatorname{char}\mathbb{K} \neq 2,3$ *and that each critical point* $\rho_i \in \operatorname{Crit}_w W$ *is either Morse or an* A_2 *singularity. Then the copies of the monotone toric fibre with local systems* $\{(T,\rho_i)\}_{\rho_i \in \operatorname{Crit}_w W}$ *split-generate* $\mathcal{F}uk(X)_w$. □

4.3 A way of extending Theorem 1.7

It is in fact not surprising that Theorem 1.7 turned out to be efficient only for A_2 singularities. The main result on which Theorem 1.7 is based upon is Proposition 2.8, which computes the *linear* part $\mathcal{CO}^1$ of the closed-open map, while the only non-Morse singularity whose local Jacobian is generated as a vector space by constant and *linear* functions is the A_2 singularity (for which the Jacobian is generated by 1 and x_n as above). One could extend the computation in Proposition 2.8 to all orders of $\mathcal{CO}^*$ when applied to products of 1–cochains on L; we conjecture that the following holds.

Conjecture 4.7 *The restriction*

$$\mathcal{CO}^k(S(\gamma))|_{\mathrm{CF}^1(L,L)^{\otimes k}} \colon \mathrm{CF}^1(L,L)^{\otimes k} \to \mathrm{CF}^0(L,L)$$

equals

$$(-1)^{\epsilon(l)}\rho(l)\cdot(l^*)^{\otimes k}\cdot 1_L \tag{4-3}$$

on symmetrised tensor products of Floer 1–cocycles. Here $l^* \colon \mathrm{CF}^1(L,L) \to \mathbb{K}$ *is given by* $l^*(x) = \langle \Psi(x), l\rangle$, *and* $l \in H_1(L)$ *is the orbit of* γ.

Remark 4.8 As in Proposition 2.8, part of the statement is that the image of this restriction necessarily lands in $\mathrm{CF}^0(L,L)$; this follows for degree reasons. Although the proof of the above formula should be analogous to Proposition 2.8, one new issue arises which we have not checked in detail. Consider the moduli spaces $\mathcal{M}^\gamma(x_1,\dots,x_k;x_0)$ from Section 2 and the pearly moduli spaces analogous to Figure 7 but with more inputs. The new issue is a different type of domain degenerations coming from the collision of input points: eg several punctured inputs for a curve in $\mathcal{M}^\gamma(x_1,\dots,x_k;x_0)$ may collide and create a bubble. To prove (4-3), one would need to argue that these collisions cancel out when the input string is symmetrised.

Formula (4-3) is a chain level computation, and whether it survives to something nontrivial in Hochschild cohomology will be governed by equations generalising

equation $(*)$ from Theorem 1.7; those equations will be determined by the A_∞ structure maps on L up to order $k+1$. When L is the monotone toric fibre, the A_∞ structure maps have been related to higher-order partial derivatives of W by Cho [12], and intuitively, the more degenerate the superpotential is, the more nontrivial information from (4-3) survives to Hochschild cohomology. Consequently, these observations are a possible starting point for proving split-generation results for toric Fano varieties with other degenerate superpotentials. However, further development of this discussion seems both complicated and not particularly demanded, given the general results of [19].

References

[1] **M Abouzaid**, *A geometric criterion for generating the Fukaya category*, Publ. Math. Inst. Hautes Études Sci. 112 (2010) 191–240 MR

[2] **M Abouzaid**, **I Smith**, *Homological mirror symmetry for the 4–torus*, Duke Math. J. 152 (2010) 373–440 MR

[3] **M Abreu**, **L Macarini**, *Remarks on Lagrangian intersections in toric manifolds*, Trans. Amer. Math. Soc. 365 (2013) 3851–3875 MR

[4] **P Albers**, *A Lagrangian Piunikhin–Salamon–Schwarz morphism and two comparison homomorphisms in Floer homology*, Int. Math. Res. Not. 2008 (2008) art. id. rnm134, 1–56 MR

[5] **G Alston**, **L Amorim**, *Floer cohomology of torus fibers and real Lagrangians in Fano toric manifolds*, Int. Math. Res. Not. 2012 (2012) 2751–2793 MR

[6] **L Amorim**, *The Künneth theorem for the Fukaya algebra of a product of Lagrangians*, Internat. J. Math. 28 (2017) art. id. 1750026, 1–38 MR

[7] **D Auroux**, *Mirror symmetry and T–duality in the complement of an anticanonical divisor*, J. Gökova Geom. Topol. 1 (2007) 51–91 MR

[8] **P Biran**, **O Cornea**, *A Lagrangian quantum homology*, from "New perspectives and challenges in symplectic field theory" (M Abreu, F Lalonde, L Polterovich, editors), CRM Proc. Lecture Notes 49, Amer. Math. Soc., Providence, RI (2009) 1–44 MR

[9] **P Biran**, **O Cornea**, *Rigidity and uniruling for Lagrangian submanifolds*, Geom. Topol. 13 (2009) 2881–2989 MR

[10] **F Charest**, *Source spaces and perturbations for cluster complexes*, Thesis, Université de Montréal (2012) MR Available at `https://search.proquest.com/docview/1352774099`

[11] **F Charette**, **O Cornea**, *Categorification of Seidel's representation*, Israel J. Math. 211 (2016) 67–104 MR

[12] **C-H Cho**, *Products of Floer cohomology of torus fibers in toric Fano manifolds*, Comm. Math. Phys. 260 (2005) 613–640 MR

[13] **C-H Cho**, **Y-G Oh**, *Floer cohomology and disc instantons of Lagrangian torus fibers in Fano toric manifolds*, Asian J. Math. 10 (2006) 773–814 MR

[14] **S Choi**, **M Masuda**, **D Y Suh**, *Quasitoric manifolds over a product of simplices*, Osaka J. Math. 47 (2010) 109–129 MR

[15] **O Cornea**, **F Lalonde**, *Cluster homology: an overview of the construction and results*, Electron. Res. Announc. Amer. Math. Soc. 12 (2006) 1–12 MR

[16] **J J Duistermaat**, *Convexity and tightness for restrictions of Hamiltonian functions to fixed point sets of an antisymplectic involution*, Trans. Amer. Math. Soc. 275 (1983) 417–429 MR

[17] **D Eisenbud**, **J Harris**, 3264 *and all that—a second course in algebraic geometry*, Cambridge Univ. Press (2016) MR

[18] **M Entov**, **L Polterovich**, *Rigid subsets of symplectic manifolds*, Compos. Math. 145 (2009) 773–826 MR

[19] **J D Evans**, **Y Lekili**, *Generating the Fukaya categories of Hamiltonian G–manifolds*, preprint (2015) arXiv

[20] **K Fukaya**, **Y-G Oh**, **H Ohta**, **K Ono**, *Lagrangian intersection Floer theory: anomaly and obstruction, I*, AMS/IP Studies in Advanced Mathematics 46, Amer. Math. Soc., Providence, RI (2009) MR

[21] **K Fukaya**, **Y-G Oh**, **H Ohta**, **K Ono**, *Lagrangian Floer theory on compact toric manifolds: survey*, preprint (2010) arXiv

[22] **S Ganatra**, *Symplectic cohomology and duality for the wrapped Fukaya category*, preprint (2013) arXiv

[23] **L Haug**, *On the quantum homology of real Lagrangians in Fano toric manifolds*, Int. Math. Res. Not. 2013 (2013) 3171–3220 MR

[24] **T Holm**, *Hochschild cohomology rings of algebras $k[X]/(f)$*, Beiträge Algebra Geom. 41 (2000) 291–301 MR

[25] **C Hyvrier**, *Lagrangian circle actions*, Algebr. Geom. Topol. 16 (2016) 1309–1342 MR

[26] **T V Kadeishvili**, *The structure of the $A(\infty)$–algebra, and the Hochschild and Harrison cohomologies*, Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR 91 (1988) 19–27 MR

[27] **C Kassel**, *A Künneth formula for the cyclic cohomology of $\mathbf{Z}/2$–graded algebras*, Math. Ann. 275 (1986) 683–699 MR

[28] **P Kleinschmidt**, *A classification of toric varieties with few generators*, Aequationes Math. 35 (1988) 254–266 MR

[29] **M Konstantinov**, *Higher rank local systems in Lagrangian Floer theory*, preprint (2017) arXiv

[30] **K Lefèvre-Hasegawa**, *Sur les A_∞ catégories*, thesis, Université Paris 7 (2003) Available at `http://www.theses.fr/en/2003PA077063`

[31] **D-M Lu**, **J H Palmieri**, **Q-S Wu**, **J J Zhang**, *A–infinity structure on Ext-algebras*, J. Pure Appl. Algebra 213 (2009) 2017–2037 MR

[32] **D McDuff**, **S Tolman**, *Topological properties of Hamiltonian circle actions*, Int. Math. Res. Pap. (2006) art. id. 72826, 1–77 MR

[33] **Y-G Oh**, *Addendum to: "Floer cohomology of Lagrangian intersections and pseudo-holomorphic disks, I"*, Comm. Pure Appl. Math. 48 (1995) 1299–1302 MR

[34] **Y-G Oh**, *Relative Floer and quantum cohomology and the symplectic topology of Lagrangian submanifolds*, from "Contact and symplectic geometry" (C B Thomas, editor), Publ. Newton Inst. 8, Cambridge Univ. Press (1996) 201–267 MR

[35] **Y Ostrover**, **I Tyomkin**, *On the quantum homology algebra of toric Fano manifolds*, Selecta Math. 15 (2009) 121–149 MR

[36] **Z Qin**, **Y Ruan**, *Quantum cohomology of projective bundles over $\mathbf{P}^n$*, Trans. Amer. Math. Soc. 350 (1998) 3615–3638 MR

[37] **A F Ritter**, *Circle actions, quantum cohomology, and the Fukaya category of Fano toric varieties*, Geom. Topol. 20 (2016) 1941–2052 MR

[38] **A F Ritter**, **I Smith**, *The monotone wrapped Fukaya category and the open-closed string map*, Selecta Math. 23 (2017) 533–642 MR

[39] **P Seidel**, *π_1 of symplectic automorphism groups and invertibles in quantum homology rings*, Geom. Funct. Anal. 7 (1997) 1046–1095 MR

[40] **P Seidel**, *Fukaya categories and Picard–Lefschetz theory*, Eur. Math. Soc., Zürich (2008) MR

[41] **P Seidel**, *Homological mirror symmetry for the genus two curve*, J. Algebraic Geom. 20 (2011) 727–769 MR

[42] **P Seidel**, *Abstract analogues of flux as symplectic invariants*, Mém. Soc. Math. Fr. 137, Soc. Math. Fr., Paris (2014) MR

[43] **P Seidel**, *Homological mirror symmetry for the quartic surface*, Mem. Amer. Math. Soc. 1116, Amer. Math. Soc., Providence, RI (2015) MR

[44] **N Sheridan**, *On the homological mirror symmetry conjecture for pairs of pants*, J. Differential Geom. 89 (2011) 271–367 MR

[45] **N Sheridan**, *On the Fukaya category of a Fano hypersurface in projective space*, Publ. Math. Inst. Hautes Études Sci. 124 (2016) 165–317 MR

[46] **I Smith**, *A symplectic prolegomenon*, Bull. Amer. Math. Soc. 52 (2015) 415–464 MR

[47] **J Smith**, *Discrete and continuous symmetries in monotone Floer theory*, preprint (2017) arXiv

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge
Cambridge, United Kingdom

dt385@cam.ac.uk

Received: 8 April 2015 Revised: 19 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 69–109

Heegaard Floer homology and knots determined by their complements

FYODOR GAINULLIN

We investigate the question of when different surgeries on a knot can produce identical manifolds. We show that given a knot in a homology sphere, unless the knot is quite special, there is a bound on the number of slopes that can produce a fixed manifold that depends only on this fixed manifold and the homology sphere the knot is in. By finding a different bound on the number of slopes, we show that non-null-homologous knots in certain homology $\mathbb{R}P^3$ are determined by their complements. We also prove the surgery characterisation of the unknot for null-homologous knots in L–spaces. This leads to showing that all knots in some lens spaces are determined by their complements. Finally, we establish that knots of genus greater than 1 in the Brieskorn sphere $\Sigma(2,3,7)$ are also determined by their complements.

57M25; 57M27

1 Introduction

Dehn surgery is an important and widely used technique for constructing 3–manifolds, yet many natural questions about it are still unanswered. For example, given a knot K in a manifold Y, how many different surgeries on K can produce a fixed manifold Z? Conjecturally, in generic circumstances, the answer is 1. More precisely, we have:

Conjecture 1.1 (cosmetic surgery conjecture; see Gordon [7, Conjecture 6.1], Kirby [12, Problem 1.81(A)] and Ni and Wu [20, Conjecture 1.1]) *Let K be a knot in a closed, connected, orientable 3–manifold Y such that the exterior of K is irreducible and not homeomorphic to the solid torus. Suppose there are two different slopes r_1 and r_2 such that there is an orientation-preserving homeomorphism between $Y_{r_1}(K)$ and $Y_{r_2}(K)$. Then the slopes r_1 and r_2 are equivalent.*

We call two slopes equivalent if there is a homeomorphism of the knot exterior taking one to the other. If there are two distinct surgeries on K (with inequivalent slopes) that produce the same oriented manifolds, then we call such surgeries purely cosmetic.

Another very natural question about knots in 3–manifolds is whether knots are determined by their complements. In other words, given two distinct knots $K_1, K_2 \subset Y$, can

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.69

there exist an orientation-preserving homeomorphism between $Y \setminus K_1$ and $Y \setminus K_2$? We remark that by "distinct" here we mean that there is no orientation-preserving homeomorphism of Y taking K_1 to K_2.

Given a knot $K_1 \subset Y$, we say that K_1 is *determined by its complement* if there is no other knot $K_2 \subset Y$ such that there is an orientation-preserving homeomorphism between $Y \setminus K_1$ and $Y \setminus K_2$. We say that K_1 is *strongly* determined by its complement if the condition of the previous sentence holds without the insistence that the homeomorphism be orientation-preserving.

By Edwards [5], the question for complements is equivalent to the analogous question for exteriors. It is not difficult to see then that the question of whether a knot is determined by its complement can be reformulated in terms of Dehn surgery as follows. A knot $K \subset Y$ is determined by its complement if and only if the following condition holds: if a surgery of some slope r on K gives Y, then r is equivalent to the meridian of K.

Knots in S^3 are determined by their complements (but not strongly, as there exist chiral knots); see Gordon and Luecke [8]. Apart from some obvious ones, no examples of knots that are not determined by their complements have been exhibited. Thus the following conjecture seems natural:

Conjecture 1.2 (knot complement conjecture; see Gordon [7, Conjecture 6.2], Kirby [12, Problem 1.81(D)], Boyer [1, Conjecture 6.2]) *Let K be a knot in a closed, connected, orientable 3–manifold Y such that the exterior of K is irreducible and not homeomorphic to the solid torus. Suppose there is a nontrivial slope r such that there is an orientation-preserving homeomorphism between $Y_r(K)$ and Y. Then r is equivalent to the meridian of K.*

We remark that dealing with equivalent slopes may seem complicated, but this issue does not arise at all if we can show that for no slope r (other than the meridian) we have $Y_r(K) \cong Y$. This not only shows that K is determined by its complement, but also that there is a unique slope which is equivalent to the meridian (ie the meridian itself). In all statements for which we will be able to show that a knot is determined by its complement, this will be the case.

Now a knot $K \subset Y$ is strongly determined by its complement if and only if it is determined by its complement and the following condition holds. If $Y_r(K)$ is homeomorphic to $-Y$ by an orientation-preserving homeomorphism, then there is an

orientation-reversing homeomorphism of the exterior of K that takes the meridian to r. For example, achiral knots in S^3 are strongly determined by their complements.

It is not difficult to see that Conjectures 1.1 and 1.2 are, in fact, equivalent. However, they are not equivalent if we concentrate on a given manifold. In other words, if Y in the conjectures is fixed, then they are genuinely different (Conjecture 1.1 implies Conjecture 1.2).

Conjecture 1.1 is wide open. In contrast to Conjecture 1.2, it is not even proven for knots in S^3. However, in [20], Ni and Wu (generalising some results of Ozsváth and Szabó [31]) used Heegaard Floer homology with great success to address the cosmetic surgery conjecture in S^3 (or other L–space homology spheres). They have been able to show that

- many manifolds (including all Seifert fibred spaces) cannot be results of purely cosmetic surgery;
- there are at most 2 slopes on a knot that can yield the same (oriented) manifold by surgery and they are negatives of each other;
- if p/q is a purely cosmetic surgery slope, then $q^2 \equiv -1 \pmod{p}$;
- knots that admit purely cosmetic surgery satisfy certain conditions on their knot Floer homology.

In fact, it can be shown that any knot that admits a Seifert fibred surgery satisfies the cosmetic surgery conjecture (see Gainullin [6]).

Boyer and Lines ruled out cosmetic surgeries on many knots in homology spheres. They showed [2, Proposition 5.1] that knots with $\Delta_K''(1) \neq 0$ satisfy the cosmetic surgery conjecture. Here Δ_K is the Alexander polynomial of K normalised so as to be symmetric and satisfy $\Delta_K(1) = 1$.

One might try to approach Conjectures 1.1 and 1.2 by trying to at least find some bound on the number of slopes on a knot that can yield the same manifold. (The cosmetic surgery conjecture states that this bound is 1.) In fact, it follows from the work of Cooper and Lackenby [3] that, given two manifolds Y and Z, there are only finitely many slopes α such that there exists a hyperbolic knot $K \subset Y$ with $Y_\alpha(K) = Z$.

Our first result says that given a knot K in any homology sphere Y, if the Heegaard Floer homology of a 3–manifold Z satisfies a certain property, then there is a bound on the number of slopes that can produce Z that only depends on the first homology of Z. More precisely, we have:

Theorem 4.3 *Let K be a knot in a homology sphere Y. Let Z be a rational homology sphere with $|H_1(Z)| = p$ such that p does not divide $\chi(HF_{\mathrm{red}}(Z))$. Suppose that there exist q_1 and q_2 such that*

$$Z = Y_{p/q_1}(K) = Y_{p/q_2}(K).$$

Then there is no multiple of p between q_1 and q_2. In particular, there are at most $\phi(p)$ surgeries on K that give Z.

Here ϕ denotes the Euler totient function.

In particular, if Z is a homology $\mathbb{R}P^3$ (ie $|H_1(Z)| = 2$) whose order of reduced Floer homology is odd, then it can be obtained by at most one surgery on any fixed knot in any homology sphere. A standard homological argument then gives:

Corollary 7.1 *Let Z be a closed, connected, oriented manifold with $|H_1(Z)| = 2$. Suppose that $\dim(HF_{\mathrm{red}}(Z))$ is odd. Then non-null-homologous knots in Z are determined by their complements.*

Spaces that satisfy the conditions of Theorem 4.3 exist, some of which we exhibit in:

Corollary 4.4 *Let $Z^1_m = S^2((3,-1),(2,1),(6m-2,-m))$ for odd $m \geq 3$ and Z^2_n be the result of $2/n$–surgery on the figure-eight knot for any odd n. If K is a knot in a homology sphere that gives one of Z^1_m or Z^2_n by surgery of some slope, then this surgery slope is unique.*

Results of Ni and Wu show that Heegaard Floer homology is a relatively good invariant when restricted to the set of manifolds obtained by surgery on a fixed knot in S^3 in the sense that at most finitely many (ie 2) of them can have the same Heegaard Floer homology (examples, when 2 different surgeries have the same Heegaard Floer homology, do occur — see Ozsváth and Szabó [31, Section 9]).

Let K be a knot in an arbitrary homology sphere. We show that, unless K is very special, the set of Heegaard Floer homologies of spaces obtained by surgery on K still contains at most finitely many repetitions. In the statement below, subscripts e and o stand for even and odd (respectively) parts of homology groups. The rest of the notation will be explained in the next section.

For a homology sphere Y and a rational homology sphere Z, define

$$N(Y,Z) = 2|H_1(Z)|\dim(HF_{\mathrm{red}}(Y)) + \dim(HF_{\mathrm{red}}(Z)).$$

Theorem 5.4 *Let Y be a non-L–space homology sphere, Z be a rational homology sphere and $K \subset Y$ be a knot and suppose there are coprime integers p and q such that $Z = Y_{p/q}(K)$.*

If $|q| > N(Y, Z)$, then

- $V_0(K) = 0$;
- $\Delta_K \equiv 1$;
- $\dim(A_k^{\mathrm{red}}(K)_e) = \dim(HF_{\mathrm{red}}(Y)_e)$ *for all* k;
- $\dim(A_k^{\mathrm{red}}(K)_o) = \dim(HF_{\mathrm{red}}(Y)_o)$ *for all* k.

Note that if Y is an L–space homology sphere, K a knot in it and $Z = Y_{p/q}(K)$, then by Gainullin [6, Theorem 7] we have $|q| \leq p + \dim(HF_{\mathrm{red}}(Z))$.[1] Therefore we lose nothing by assuming that Y is not an L–space.

Theorem 5.4 shows that for a knot to have a surgery not satisfying the bound $|q| \leq N(Y, Z)$, the knot has to be quite "special", ie satisfy all four of the bullet points listed in the statement of the theorem. In the next proposition, we show that even for such "special" knots that have genus > 1 we can provide an alternative upper bound on q.

For a rational homology sphere Z, define

$$\widehat{D}(Z) = \max\{\deg(z) - d(Z, \mathfrak{t}) \mid z \in HF_{\mathrm{red}}(Z, \mathfrak{t}) \text{ is homogeneous, } \mathfrak{t} \in \mathrm{Spin}^c(Z)\}$$

and

$$\check{D}(Z) = \min\{\deg(z) - d(Z, \mathfrak{t}) \mid z \in HF_{\mathrm{red}}(Z, \mathfrak{t}) \text{ is homogeneous, } \mathfrak{t} \in \mathrm{Spin}^c(Z)\},$$

where by $\deg(z)$ we understand the absolute grading of z.

Proposition 6.3 *Let $K \subset Y$ be a knot such that $Z = Y_{p/q}(K)$ for $p, q > 0$ and $q > N(Y, Z)$. Suppose the genus of K is larger than 1. Then*

$$\lfloor q/p \rfloor \leq \tfrac{1}{2}(\widehat{D}(Z) - \check{D}(Y)).$$

Thus if infinitely many spaces obtained by surgery on a knot K in a homology sphere have the same Heegaard Floer homology then K has genus 1 and trivial Alexander polynomial.

In [13], Kronheimer, Mrowka, Ozsváth and Szabó prove the following "surgery characterisation of the unknot":

[1] In [6] the theorem is stated for $Y = S^3$, but the same proof is valid for arbitrary L–space homology spheres.

Theorem 1.3 [13, Theorem 1.1] *Let U denote the unknot in S^3 and let K be any knot. If there is an orientation-preserving diffeomorphism $S^3_r(K) \cong S^3_r(U)$ for some rational number r, then $K = U$.*

Clearly, this theorem provides an alternative proof of the fact that knots in S^3 are determined by their complements. The proof has been adapted to the setting of Heegaard Floer homology in Ozsváth and Szabó [29] (see also Manion [15]). Crucial for the proof is the fact that S^3 is an L–space, ie it has "the smallest possible" Heegaard Floer homology.

The Brieskorn space $\Sigma(2,3,7)$ has perhaps the simplest possible Heegaard Floer homology a non-L–space can have. We use this to prove:

Theorem 7.2 *Knots of genus larger than 1 in the Brieskorn sphere $\Sigma(2,3,7)$ are determined by their complements. Moreover, if $K \subset \Sigma(2,3,7)$ is a counterexample to Conjecture 1.2 then the surgery slope is integral, $\widehat{HFK}(\Sigma(2,3,7), K, 1)$ has dimension 2 and its generators lie in different $\mathbb{Z}_2$–gradings.*

Nonfibred knots of genus larger than 1 in $\Sigma(2,3,7)$ are strongly determined by their complements.

Returning to L–spaces, we show that Theorem 1.3 admits a generalisation as follows.

Theorem 8.2 *Let Y be an L–space and $K \subset Y$ a null-homologous knot. Suppose that*

$$HF^+(Y_{p/q}(K)) \cong HF^+(Y \# L(p,q)).$$

Then K is the unknot.

In particular, null-homologous knots in L–spaces are determined by their complements.

In the statement above, by $L(p,q)$ we understand the result of p/q surgery on the unknot in S^3.

Shortly after the first version of this preprint was published, Ravelomanana [35] published his proof of the fact that knots in L–space integral homology spheres are determined by their complements, thus reproving a part of our Theorem 8.2. In a later version of his paper, he was also able to show that knots in $\Sigma(2,3,5)$ are (in our terminology) strongly determined by their complements.

Lens spaces are L–spaces (indeed, they are the reason for the name), so it follows that null-homologous knots in lens spaces are determined by their complements. In fact, Mauricio [16] has proven the above theorem for integral slopes, so, coupled with the cyclic surgery theorem of Culler, Gordon, Luecke and Shalen [4], Mauricio's result implies that null-homologous nontorus knots in lens spaces are determined by their complements (though Mauricio does not phrase it in this way).

Some homological arguments allow us to prove a stronger result for lens spaces. First, we need to fix some notation. Note that we will be making some arbitrary choices. Suppose K is a knot in a lens space $L = L(p,q)$ and view L as a union of two Heegaard solid tori V and W. Isotope K into W and fix thus obtained isotopy class of K in W. Now that K is viewed as a fixed knot in W, it has a well-defined winding number w in W (ie the algebraic intersection number of K with a meridional disc of W — it does not make sense if we allow K to leave W). Embed W into S^3 in a standard way. This endows both K and W with a preferred longitude (note, again, that this only makes sense after we fix the embedding, which is chosen arbitrarily). We use the thus-obtained longitude of K to identify slopes with rational numbers.

Note that even though w was fixed rather arbitrarily, it features in the corollary below. This is because its remainder modulo p is well-defined and the slope n in the statement also depends upon the choices above.

Then we have the following result:

Corollary 8.3 *If p is square-free, then all knots in $L = L(p,q)$ are determined by their complements.*

More precisely, let K be a knot whose exterior is not a solid torus and such that a nontrivial surgery on it gives L. Then the exterior of K is not Seifert fibred, $p \mid w^2$ and the surgery slope, n, is an integer that satisfies (with some choice of sign)

$$n = -q\frac{w^2}{p} \pm 1.$$

Moreover, there is at most one such slope (ie we can choose either $+$ or $-$ but not both in the equation above).

Acknowledgements I would like to thank András Juhász and Tye Lidman for valuable comments about the earlier versions of this paper. I am very grateful to Tom Hockenhull and Marco Marengon for persistently suggesting I should write some of the results

in this paper up. Special thanks go to Tye Lidman and Duncan McCoy who found a mistake in the previous version of this paper. I would also like to thank my supervisor Dorothy Buck for her continued support over the course of my PhD studies.

Finally, I am very grateful to the anonymous referee. They have been very careful in reading the drafts of this paper and their detailed comments and suggestions improved the paper very substantially.

2 Review of the mapping cone formula

In this section, we review the mapping cone formula of [31, Theorem 1.1]. We use notation largely similar to that of Ni and Wu [20].

Given a knot K in a homology sphere Y we can associate to it a doubly pointed Heegaard diagram as in [24]. We define a complex $C = CFK^\infty(Y, K)$ generated (over an arbitrary field $\mathbb{F}$) by elements of the form $[\boldsymbol{x}, i, j]$, where $\boldsymbol{x}$ is an "intersection point" of the Heegaard diagram (as defined in [24]) and $(i, j) \in \mathbb{Z} \times \mathbb{Z}$. Generators of C are not all triples $[\boldsymbol{x}, i, j]$, but only those that satisfy a certain condition. The differential on C does not increase either i or j, so C is doubly filtered by the pair $(i, j) \in \mathbb{Z} \times \mathbb{Z}$. The doubly filtered chain homotopy type of this complex is a knot invariant [24, Theorem 3.1].

By [33, Lemma 4.5], the complex C is homotopy equivalent (as a filtered complex) to a complex for which all filtration-preserving differentials are trivial. In other words, at each filtration level, we replace the group, viewed as a chain complex with the filtration-preserving differential, by its homology. From now on we work with this *reduced* complex.

The complex C is invariant under the shift by the vector $(-1, -1)$. Thus there is an action of a formal variable U on C, which is simply the translation by the vector $(-1, -1)$. In other words, the group at the filtration level (i, j) is the same as the one at the filtration level $(i - 1, j - 1)$ and U is the identity map from the first one to the second. Of course, U is a chain map. In C the map U is invertible (but note that it will not be in various subcomplexes and quotients), so C is an $\mathbb{F}[U, U^{-1}]$–module.

This means that as an $\mathbb{F}[U, U^{-1}]$–module C is generated by the elements with the first filtration level $i = 0$. In the reduced complex, the group at filtration level $(0, j)$ is denoted by $\widehat{HFK}(Y, K, j)$ and is known as the knot Floer homology of K at Alexander grading j.

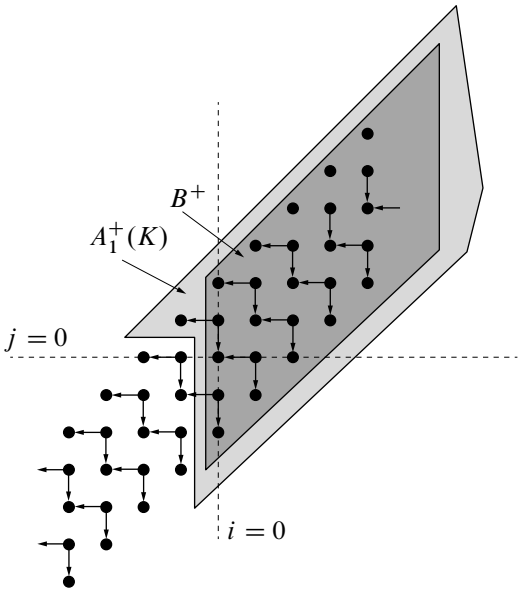


Figure 1: Schematic representation of (a part of) the complex C (for some genus-2 knot). Dots represent groups at various filtration levels and arrows stand for components of the differential. The part shaded in grey (including the dark grey part over it) is the complex $A_1^+(K)$. The part shaded dark grey represents B^+.

The complex C possesses an absolute $\mathbb{Q}$–grading and a relative $\mathbb{Z}$–grading, ie differences of absolute $\mathbb{Q}$–gradings of elements of C are integers. In fact, C is the complex used to compute the (∞–flavour of the) Heegaard Floer homology of Y; the knot provides an additional filtration for it. By grading the Heegaard Floer homology of Y (as in [22]), we obtain the grading on C. The map U decreases this grading by 2.

Using the filtration on C we can define the quotients (see Figure 1)

$$A_k^+(K) = C\{i \geq 0 \text{ or } j \geq k\}, \quad k \in \mathbb{Z},$$

and

$$B^+ = C\{i \geq 0\} \cong CF^+(Y).$$

We also define two chain maps, $v_k, h_k \colon A_k^+(K) \to B^+$. The first one is just the projection (ie it sends to zero all generators with $i < 0$ and acts as the identity map for everything else). The second one is the composition of three maps: firstly we project to $C\{j \geq k\}$, then we multiply by U^k (this shifts everything by the vector $(-k, -k)$) and finally we apply a chain homotopy equivalence that identifies $C\{j \geq 0\}$ with $C\{i \geq 0\}$.

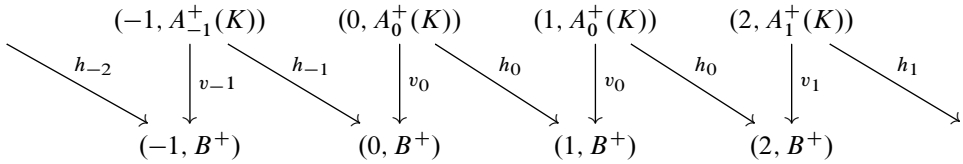


Figure 2: Schematic representation of the portion of the map $D^+_{i,p/q}$ for $i = 0$ and $p/q = \frac{2}{3}$

Such a chain homotopy equivalence exists because the two complexes both represent $CF^+(Y)$ and, by general theory [24], there is a chain homotopy equivalence between them, induced by the moves between the Heegaard diagrams. We usually do not know the explicit form of this chain homotopy equivalence.

Knot Floer homology detects the knot genus. It does so in the following way [18, Theorem 3.1]:

Theorem 2.1 (Ni) *Let Y be a homology sphere and $K \subset Y$ a knot.*
Then $g(K) = \max\{j \in \mathbb{Z} \mid \widehat{HFK}(Y, K, j) \neq 0\}$.

From this (together with symmetries of C) we can see that the maps v_k (resp. h_k) are isomorphisms if $k \geq g$ (resp. $k \leq -g$). For example, Figure 1 represents some knot of genus 2.

We define chain complexes

$$\mathcal{A}^+_{i,p/q}(K) = \bigoplus_{n\in\mathbb{Z}}(n, A^+_{\lfloor (i+pn)/q\rfloor}(K)), \quad \mathcal{B}^+ = \bigoplus_{n\in\mathbb{Z}}(n, B^+).$$

The first entry in the brackets here is simply a label used to distinguish different copies of the same group. There is a chain map $D^+_{i,p/q}$ from $\mathcal{A}^+_{i,p/q}(K)$ to $\mathcal{B}^+$ defined by taking sums of all maps v_k and h_k with appropriate domains and requiring that the map v_k goes to the group with the same label n and h_k increases the label by 1. Explicitly

$$D^+_{i,p/q}\big(\{(k, a_k)\}_{k\in\mathbb{Z}}\big) = \{(k, b_k)\}_{k\in\mathbb{Z}},$$

where $b_k = v^+_{\lfloor (i+pk)/q\rfloor}(a_k) + h^+_{\lfloor (i+p(k-1))/q\rfloor}(a_{k-1})$ — see Figure 2.
Each of $A^+_k(K)$ and B^+ inherits a relative $\mathbb{Z}$–grading from the one on C. Let $\mathbb{X}^+_{i,p/q}$ denote the mapping cone of $D^+_{i,p/q}$. We fix a relative $\mathbb{Z}$–grading on the whole of it by requiring that the maps v_k and h_k (and so $D^+_{i,p/q}$) decrease it by 1. The following is proven in [31]:

Theorem 2.2 (Ozsváth and Szabó) *There is a relatively graded isomorphism of $\mathbb{F}[U]$–modules*

$$H_*(\mathbb{X}^+_{i,p/q}) \cong HF^+(Y_{p/q}(K), i).$$

The index i in $HF^+(Y_{p/q}(K), i)$ stands for a Spin^c structure. The numbering of Spin^c structures we refer to is defined in [31], but we do not need precise details of how to obtain this numbering for our purposes.

We can also determine the absolute grading on the mapping cone. The group $\mathcal{B}^+$ is independent of the knot. Now if we insist that the absolute grading on the mapping cone for the unknot should coincide with the grading of HF^+ of the surgery on it (ie $d(L(p,q),i) + d(Y)$), this fixes the grading on $\mathcal{B}^+$. We then use this grading to fix the grading on $\mathbb{X}^+_{i,p/q}$ for arbitrary knots — this grading then is the correct grading, ie it coincides with the one HF^+ should have.

It seems quite complicated to understand the homology of the mapping cone of $D^+_{i,p/q}$ by direct inspection. Thus we pass to homology of the objects we introduced above. Specifically, let $\boldsymbol{A}^+_k(K) = H_*(A^+_k(K))$, $\boldsymbol{B}^+ = H_*(B^+)$, $\mathbb{A}^+_{i,p/q}(K) = H_*(\mathcal{A}^+_{i,p/q}(K))$, $\mathbb{B}^+ = H_*(\mathcal{B}^+)$ and let $\boldsymbol{v}_k$, $\boldsymbol{h}_k$ and $\boldsymbol{D}^+_{i,p/q}$ denote the maps induced by v_k, h_k and $D^+_{i,p/q}$, respectively, in homology.

When we talk about $\mathbb{A}^+_{i,p/q}(K)$ as an absolutely graded group, we mean the grading that it inherits from the absolute grading of the mapping cone that we described above.

Recall that the short exact sequence

$$0 \to \mathcal{B}^+ \xrightarrow{i} \mathbb{X}^+_{i,p/q} \xrightarrow{j} \mathcal{A}^+_{i,p/q}(K) \to 0$$

induces the exact triangle

(1)
$$\begin{array}{ccc} \mathbb{A}^+_{i,p/q}(K) & \xrightarrow{\boldsymbol{D}^+_{i,p/q}} & \mathbb{B}^+ \\ & \overset{j_*}{\nwarrow} & \downarrow{\scriptstyle i_*} \\ & & H_*(\mathbb{X}^+_{i,p/q}) \cong HF^+(Y_{p/q}(K), i) \end{array}$$

All maps in these sequences are U–equivariant.

Define $\mathcal{T}^+_d$ to be the graded $\mathbb{F}[U]$–module $\mathbb{F}[U, U^{-1}]/U \cdot \mathbb{F}[U]$ with (the equivalence class of) 1 having grading d and multiplication by U decreasing the grading by 2. Similarly, let $\mathcal{T}(N)$ be the submodule of $\mathcal{T}^+_d$ generated by $\{1, U^{-1}, \dots, U^{-(N-1)}\}$.

We omit the subscript d if the absolute grading does not exist or is not relevant. However, even without the absolute grading, these groups are still relatively $\mathbb{Z}$–graded (by requiring that U decreases the grading by 2).

If Z is a rational homology sphere, then $HF^+(Z,\mathfrak{s}) = \mathcal{T}^+_d \oplus HF_{\mathrm{red}}(Z,\mathfrak{s})$, where $d = d(Z,\mathfrak{s})$ is the d–invariant (or the correction term) of Z in the Spin^c structure $\mathfrak{s}$ and $HF_{\mathrm{red}}(Z,\mathfrak{s})$ is the reduced Floer homology of Z in the same Spin^c structure. The reduced Floer homology of Z is the sum of reduced Floer homologies in all Spin^c structures, which we denote by

$$HF_{\mathrm{red}}(Z) = \bigoplus_{\mathfrak{s}\in \mathrm{Spin}^c(Z)} HF_{\mathrm{red}}(Z,\mathfrak{s}).$$

For each $\mathfrak{s} \in \mathrm{Spin}^c(Z)$, $HF_{\mathrm{red}}(Z,\mathfrak{s})$ is a finitely generated $\mathbb{F}[U]$–module in the kernel of a large enough power of U, thus it has the form $\bigoplus_{i=1}^m \mathcal{T}(n_i)$ for some n_i.

We have that $\boldsymbol{A}^+_k(K) \cong \boldsymbol{A}^T_k(K) \oplus \boldsymbol{A}^{\mathrm{red}}_k(K)$ and $\boldsymbol{B}^+ = \boldsymbol{B}^T \oplus \boldsymbol{B}^{\mathrm{red}}$, where $\boldsymbol{A}^T_k(K) \cong \mathcal{T}^+ \cong \boldsymbol{B}^T$ and $\boldsymbol{A}^{\mathrm{red}}_k(K)$ and $\boldsymbol{B}^{\mathrm{red}}$ are finitely generated-$\mathbb{F}[U]$ modules in the kernel of a large enough power of U. Define

$$\mathbb{A}^T_{i,p/q}(K) = \bigoplus_{n\in\mathbb{Z}} (n, \boldsymbol{A}^T_{\lfloor (i+pn)/q\rfloor}(K)), \quad \mathbb{A}^{\mathrm{red}}_{i,p/q}(K) = \bigoplus_{n\in\mathbb{Z}} (n, \boldsymbol{A}^{\mathrm{red}}_{\lfloor (i+pn)/q\rfloor}(K)),$$
$$\mathbb{B}^T = \bigoplus_{n\in\mathbb{Z}} (n, \boldsymbol{B}^T), \qquad \mathbb{B}^{\mathrm{red}} = \bigoplus_{n\in\mathbb{Z}} (n, \boldsymbol{B}^{\mathrm{red}}).$$

We decompose the maps in a similar manner. Let $\boldsymbol{D}^+_{i,p/q} = \boldsymbol{D}^T_{i,p/q} \oplus \boldsymbol{D}^{\mathrm{red}}_{i,p/q}$, where the first map is the restriction of $\boldsymbol{D}^+_{i,p/q}$ to $\mathbb{A}^T_{i,p/q}(K)$ and the second one is the restriction to $\mathbb{A}^{\mathrm{red}}_{i,p/q}(K)$. Let $\boldsymbol{v}^T_k$ and $\boldsymbol{h}^T_k$ be the restrictions of $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$, respectively, to $\boldsymbol{A}^T_k(K)$. Then $\boldsymbol{D}^T_{i,p/q}$ is defined using $\boldsymbol{v}^T_k$ and $\boldsymbol{h}^T_k$ in the same way as $\boldsymbol{D}^+_{i,p/q}$ is defined using $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$.

Notice that the images of $\boldsymbol{v}^T_k$ and $\boldsymbol{h}^T_k$ are contained in $\boldsymbol{B}^T$ — this is because they are $\mathbb{F}[U]$–module maps. In fact, since these maps are homogeneous and are isomorphisms for large enough gradings, they are multiplications by some powers U^{V_k} and U^{H_k} for $\boldsymbol{v}^T_k$ and $\boldsymbol{h}^T_k$, respectively.

The following are some useful properties of V_k and H_k. The proofs are completely analogous to the case of knots in S^3; see [20]:

- $V_k = H_{-k}$ for any $k \in \mathbb{Z}$.
- $V_k \geq V_{k+1}$ and $H_k \leq H_{k+1}$ for any $k \in \mathbb{Z}$.

- $V_k \to +\infty$ as $k \to -\infty$ and $H_k \to +\infty$ as $k \to +\infty$.
- $V_k = 0$ for $k \geq g(K)$ and $H_k = 0$ for $k \leq -g(K)$.

In other words, V_k form a nonincreasing unbounded sequence of nonnegative numbers, which become zero at $g(K)$ or earlier, and $H_k = V_{-k}$.

3 Correction terms

The next lemma essentially shows that when $p, q > 0$, the map $D^+_{i,p/q}$ becomes an isomorphism "at the ends", so in the mapping cone formula we only need to consider a finite central part.

Lemma 3.1 *Fix a number $G \geq g(K)$. Let $p, q > 0$. Let $\mathbb{B}^+_{G^+}$ and $\mathbb{B}^+_{G^-}$ be the subgroups of $\mathbb{B}^+$ consisting of all $(n, \boldsymbol{B}^+)$ with n satisfying*

$$\left\lfloor \frac{i + pn}{q} \right\rfloor \geq G$$

and

$$\left\lfloor \frac{i + p(n-1)}{q} \right\rfloor \leq -G,$$

respectively.

Let $\mathbb{A}^+_{i,p/q}(K)_{G^+}$ and $\mathbb{A}^+_{i,p/q}(K)_{G^-}$ be the subgroups of $\mathbb{A}^+_{i,p/q}(K)$ consisting of all $(n, \boldsymbol{A}^+_k(K))$ with n satisfying

$$\left\lfloor \frac{i + pn}{q} \right\rfloor \geq G$$

and

$$\left\lfloor \frac{i + pn}{q} \right\rfloor \leq -G,$$

respectively.

Then $\boldsymbol{D}^+_{i,p/q}$ maps $\mathbb{A}^+_{i,p/q}(K)_{G^\pm}$ isomorphically onto $\mathbb{B}^+_{G^\pm}$.

Proof The cases $n \geq 0$ and $n \leq 0$ are similar, so we will only consider $n \geq 0$.

First we want to show that the image of $\boldsymbol{D}^+_{i,p/q}$ is all of $\mathbb{B}^+_{G^+}$. Suppose $\xi \in (n, \boldsymbol{B}^+)$, $n \geq 0$ and $\lfloor (i + pn)/q \rfloor \geq G \geq g(K)$.

Note that $\boldsymbol{v}_k$ is an isomorphism for $k \geq g(K)$. Moreover, if we identify each $\boldsymbol{A}^+_k(K)$ with $\boldsymbol{B}^+$ via $\boldsymbol{v}_k$ for $k \geq g(K)$, then any fixed element of $\boldsymbol{A}^+_k(K)$ is in the kernel of $\boldsymbol{h}_k$ for big enough k and $\boldsymbol{h}_k$ decreases the grading by any amount we want if k is big enough.

Let $\eta_0 = \boldsymbol{v}^{-1}_{\lfloor (i+pn)/q \rfloor}(\xi) \in (n, \boldsymbol{A}^+_{\lfloor (i+pn)/q \rfloor})$. Define η_m inductively by

$$\eta_m = \boldsymbol{v}^{-1}_{\lfloor (i+(n+m)p)/q \rfloor}(\boldsymbol{h}_{\lfloor (i+(n+m-1)p)/q \rfloor}(\eta_{m-1})).$$

By properties of $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$ described above, for big enough m we have $\eta_m = 0$. This shows that $\xi = \boldsymbol{D}^+_{i,p/q}\left(\sum_k \eta_k\right)$ is in the image of $\boldsymbol{D}^+_{i,p/q}$.

Now suppose $\eta \in \mathbb{A}^+_{i,p/q}(K)_{G^+}$ is in the kernel of the restriction of $\boldsymbol{D}^+_{i,p/q}$ to $\mathbb{A}^+_{i,p/q}(K)_{G^+}$. Then the leftmost component must be in the kernel of the corresponding map $\boldsymbol{v}_k$ — a contradiction. □

In the previous section, we mentioned that it is possible to put an absolute grading on $\mathbb{B}^+$ that does not depend on the knot we consider. In the next lemma, we explicitly describe this grading. Note that since the relative grading is already fixed, it is enough to put an absolute grading on any homogeneous element of $\mathbb{B}^+$.

Lemma 3.2 *Let Y be a homology sphere. Consider the mapping cone for the* Spin^c *structure i. The grading of 1 in $(0, \boldsymbol{B}^T)$ is $d(Y) + d(L(p,q), i) - 1$.*

Proof As a result of p/q–surgery on the unknot, we get $Y \# L(p,q)$, whose correction terms are $d(Y \# L(p,q), i) = d(Y) + d(L(p,q), i)$.

By Lemma 3.1 applied to the unknot (which has genus 0), the map $\boldsymbol{D}^+_{i,p/q}$ is surjective for the unknot. Thus $HF^+(Y \# L(p,q)) \cong \ker(\boldsymbol{D}^+_{i,p/q})$.

Just as in [20, Proof of Proposition 1.6], there is a tower in the kernel of $\boldsymbol{D}^T_{i,p/q}$ and the element U^{-n} in this tower has U^{-n} as a component in $(0, \boldsymbol{A}^T_0(K))$. This shows that 1 in $(0, \boldsymbol{A}^T_0(K))$ has grading $d(Y \# L(p,q), i) = d(Y) + d(L(p,q), i)$, so, since $V_0 = 0$, 1 in $(0, \boldsymbol{B}^T)$ has grading $d(Y) + d(L(p,q), i) - 1$. □

Let Y be a homology sphere. Recall that its Heegaard Floer homology possesses an absolute $\mathbb{Z}_2$–grading, defined to be 0 on the tower part and be the reduction mod 2 of the relative $\mathbb{Z}$–grading.[2] Decompose the Heegaard Floer homology of Y in the following way: $HF^+(Y) \cong \mathcal{T}^+ \bigoplus_{i=1}^{l} \mathcal{T}(n_i^+) \oplus \bigoplus_{i=1}^{m} \mathcal{T}(n_i^-)$, where $\mathcal{T}(n_i^+)$ (resp. $\mathcal{T}(n_i^-)$) lie in even (resp. odd) $\mathbb{Z}_2$–grading. The following proposition may be seen as a generalisation of [20, Proposition 1.6].

[2] In other words, every element of the tower has grading 0 and the grading of an element is 1 if and only if it has odd relative $\mathbb{Z}$–grading with some element of the tower.

Proposition 3.3 *With notation as above, suppose $Z = Y_{p/q}(K)$ for $p > 0$ and $q > 0$. Then*

$$d(Y) + d(L(p,q),i) - 2\max\{V_{\lfloor i/q\rfloor}, H_{\lfloor (i-p)/q\rfloor}\} - 2\max_j\{n_j^-\} \le d(Z,i) \tag{2}$$

and

$$d(Z,i) \le d(Y) + d(L(p,q),i) - 2\max\{V_{\lfloor i/q\rfloor}, H_{\lfloor (i-p)/q\rfloor}\}. \tag{3}$$

Proof Since the grading of $\mathbb{B}^+$ is independent of the knot, by Lemma 3.2 we have that 1 in $(0, \boldsymbol{B}^T)$ has grading $d(Y) + d(L(p,q),i) - 1$. As usual, the proof subdivides into two cases, depending on whether $V_{\lfloor i/q\rfloor} = \max\{V_{\lfloor i/q\rfloor}, H_{\lfloor (i-p)/q\rfloor}\}$ or otherwise. The two cases are analogous, so we only consider the case $V_{\lfloor i/q\rfloor} = \max\{V_{\lfloor i/q\rfloor}, H_{\lfloor (i-p)/q\rfloor}\}$.

Then, as in [20, Proof of Proposition 1.6] (or see [6, Lemmas 12–13]) we can show that there is a tower in the kernel of $\boldsymbol{D}^T_{i,p/q}$ such that U^{-n} in this tower has U^{-n} as the component in $(0, \boldsymbol{A}^T_0(K))$. Suppose that 1 in this tower has grading d. Then also 1 in $(0, \boldsymbol{A}^T_0(K))$ has grading d and thus $U^{-V_{\lfloor i/q\rfloor}} \in (0, \boldsymbol{A}^T_0(K))$ has grading $d + 2V_{\lfloor i/q\rfloor}$.

On the other hand, $U^{-V_{\lfloor i/q\rfloor}} \in (0, \boldsymbol{A}^T_0(K))$ is mapped to $1 \in (0, \boldsymbol{B}^T)$ by $\boldsymbol{v}^+_{\lfloor i/q\rfloor}$, which has grading -1. So

$$d + 2V_{\lfloor i/q\rfloor} - 1 = d(Y) + d(L(p,q),i) - 1,$$

from which it follows that $d = d(Y) + d(L(p,q),i) - 2V_{\lfloor i/q\rfloor}$.

By the exact triangle (1), everything in the kernel of $\boldsymbol{D}^+_{i,p/q}$ must be in the image of j_*. So, in particular, the tower we identified in the kernel of $\boldsymbol{D}^T_{i,p/q}$ (and thus $\boldsymbol{D}^+_{i,p/q}$) must be in the image of j_*. At high enough gradings only the elements of the tower in $HF^+(Y_{p/q}(K),i)$ may hit the elements of the tower in the kernel of $\boldsymbol{D}^+_{i,p/q}$. Since the maps in the triangle are U–equivariant, the tower in $HF^+(Y_{p/q}(K),i)$ must be mapped onto the tower in the kernel of $\boldsymbol{D}^+_{i,p/q}$. It follows that

$$d(Y) + d(L(p,q),i) - 2V_{\lfloor i/q\rfloor} = d \ge d(Z,i).$$

This argument also shows that the map j_* has submodule $\mathcal{T}\big(\frac{1}{2}(d - d(Z,i))\big)$ in its kernel and moreover this submodule lies in $\mathbb{Z}_2$–grading 0.

Thus there has to be a submodule $\mathcal{T}(N)$ in $\mathbb{B}^+$ with $N \ge \frac{1}{2}(d - d(Z,i))$ that is not in the image of $\boldsymbol{D}^+_{i,p/q}$. Moreover, it must have odd $\mathbb{Z}_2$–grading in $\mathbb{B}^+$.

However, the odd part of $\mathbb{B}^+$ is in the kernel of $U^{\max_j\{n_j^-\}}$, so

$$\max_j\{n_j^-\} \geq \tfrac{1}{2}(d - d(Z,i))$$

and therefore

$$d(Z,i) \geq d - 2\max_j\{n_j^-\} = d(Y) + d(L(p,q),i) - 2V_{\lfloor i/q\rfloor} - 2\max_j\{n_j^-\}.$$

This completes the proof in the case $V_{\lfloor i/q\rfloor} \geq H_{\lfloor (i-p)/q\rfloor}$. The other case is completely analogous. □

The following straightforward corollary may be of interest.

Corollary 3.4 *Let Y be a positively oriented Seifert fibred homology sphere and $K \subset Y$ a knot. Suppose $Z = Y_{p/q}(K)$, where $p/q > 0$. Then*

$$d(Z,i) = d(Y) + d(L(p,q),i) - 2\max\{V_{\lfloor i/q\rfloor}, H_{\lfloor (i-p)/q\rfloor}\}.$$

Proof Positively oriented Seifert fibred homology spheres have $n_i^- = 0$ for all i, by [23, Corollary 1.4]. □

4 Surgery producing spaces with $p \nmid \chi(HF_{\mathrm{red}})$

In this section, we want to prove Theorem 4.3. We use the Casson–Walker invariant, normalised as in [20]. Our normalisation for the Alexander polynomial of a null-homologous knot in a rational homology also differs from that used in some other sources (in particular [38]). Specifically, we require that the Alexander polynomial Δ_K of a null-homologous knot K in a rational homology sphere Y satisfies $\Delta_K(t) = \Delta_K(t^{-1})$ and $\Delta_K''(1) = |H_1(Y)|$.

To a rational homology sphere Y, the Casson–Walker invariant assigns a rational number $\lambda(Y)$. Two key properties we will need are as follows.

For a null-homologous knot K in a rational homology sphere Y we have (see [38, Proposition 6.2] and note we are using slightly different normalisations)

$$\lambda(Y_{p/q}(K)) = \lambda(Y) + \lambda(L(p,q)) + \frac{q}{2p|H_1(Y)|}\Delta_K''(1). \tag{4}$$

The following formula appears in [37, Theorem 3.3]:

$$|H_1(Y)|\lambda(Y) = \chi(HF_{\mathrm{red}}(Y)) - \frac{1}{2}\sum_{\mathfrak{s}\in \mathrm{Spin}^c(Y)} d(Y,\mathfrak{s}). \tag{5}$$

If we apply this formula to a lens space $L(p,q)$ we get

$$\sum_{\mathfrak{s}\in \mathrm{Spin}^c(L(p,q))} d(L(p,q),\mathfrak{s}) = -2p\lambda(L(p,q)). \tag{6}$$

Another invariant we will briefly use is the Casson–Gordon invariant, τ, which satisfies the following surgery formula: Suppose W is an integral homology sphere and K a knot in it. Then

$$\tau(Y_{p/q}(K)) = \tau(L(p,q)) - \sigma(K,p), \tag{7}$$

where $\sigma(K,p)$ is a number depending only on K and p.

Finally, both Casson–Walker and Casson–Gordon invariants of lens spaces can be expressed in terms of Dedekind sums. For our purposes, it is enough to know that a Dedekind sum assigns to a pair of coprime numbers (p,q) a number $s(q,p)$. We have

$$\lambda(L(p,q)) = -\tfrac{1}{2}s(q,p) \tag{8}$$

and

$$\tau(L(p,q)) = -4ps(q,p). \tag{9}$$

Proposition 4.1 *Let Y be a homology sphere, $K\subset Y$ a knot and suppose there is a rational homology sphere Z with*

$$Z = Y_{p/q_1}(K) = Y_{-p/q_2}(K),$$

where $p, q_1, q_2 > 0$.

Then

$$\chi(HF_{\mathrm{red}}(Z)) = p\chi(HF_{\mathrm{red}}(Y)).$$

Proof By combining equations (7) and (9) we get $s(q_1,p) = s(-q_2,p)$. From this and (8) we get $\lambda(L(p,q_1)) = \lambda(L(p,-q_2))$.

Now (4) implies that

$$\frac{q_1}{2p|H_1(Z)|}\Delta_K''(1) = \frac{-q_2}{2p|H_1(Z)|}\Delta_K''(1),$$

from which it follows that $\Delta_K''(1) = 0$.

Formula (3) gives $d(Z,i)\le d(Y)+d(L(p,q_1),i)$. If we can get Z from Y by $-p/q_2-$surgery, then by reversing orientations we see that we can get $-Z$ from $-Y$ by p/q_2-surgery. Using formula (3) then gives $d(-Z,i)\le d(-Y)+d(L(p,q_2),i)$, which yields $-d(Z,i)\le -d(Y)-d(L(p,-q_2),i)$ and thus $d(Z,i)\ge d(Y)+d(L(p,-q_2),i)$.

Summing over all Spin^c structures yields

$$\sum_{\mathfrak{s}\in\mathrm{Spin}^c(Z)} d(Z,\mathfrak{s}) \le pd(Y) + \sum_{\mathfrak{s}\in\mathrm{Spin}^c(L(p,q_1))} d(L(p,q_1),\mathfrak{s}) \tag{10}$$

and

$$\sum_{\mathfrak{s}\in\mathrm{Spin}^c(Z)} d(Z,\mathfrak{s}) \ge pd(Y) + \sum_{\mathfrak{s}\in\mathrm{Spin}^c(L(p,-q_2))} d(L(p,-q_2),\mathfrak{s}). \tag{11}$$

By (6), the two last terms (sums of d–invariants for lens spaces) in the two equations above are equal.

The inequalities (10) and (11) now imply

$$\sum_{\mathfrak{s}\in\mathrm{Spin}^c(Z)} d(Z,\mathfrak{s}) \;=\; pd(Y) - 2p\lambda(L(p,q_1)) \;=\; pd(Y) - 2p\lambda(L(p,-q_2)). \tag{12}$$

Now using (5) we get

$$\chi(HF_{\mathrm{red}}(Z)) - \frac{1}{2}\sum_{\mathfrak{s}\in\mathrm{Spin}^c(Z)} d(Z,\mathfrak{s}) = p\lambda(Z).$$

Using (4) and the fact that $\Delta_K''(1) = 0$, we obtain

$$p\lambda(Z) = p\lambda(Y) + p\lambda(L(p,q_1)).$$

Applying (5) to Y gives us

$$\begin{aligned} p\lambda(Z) &= p\chi(HF_{\mathrm{red}}(Y)) - \tfrac{1}{2}pd(Y) + p\lambda(L(p,q_1)) \\ &= p\chi(HF_{\mathrm{red}}(Y)) - \tfrac{1}{2}\big(pd(Y) - 2p\lambda(L(p,q_1))\big). \end{aligned}$$

Substituting from (12) and combining the equalities we arrive at

$$\chi(HF_{\mathrm{red}}(Z)) - \frac{1}{2}\sum_{\mathfrak{s}\in\mathrm{Spin}^c(Z)} d(Z,\mathfrak{s}) = p\chi(HF_{\mathrm{red}}(Y)) - \frac{1}{2}\sum_{\mathfrak{s}\in\mathrm{Spin}^c(Z)} d(Z,\mathfrak{s}),$$

from which the conclusion of the proposition follows. □

Since this proposition holds for arbitrary homology spheres, we have relative freedom to "change coordinates", ie to see a surgery on a knot in some homology sphere as a surgery on its dual in another homology sphere. This is the essence of what is going on in the next lemma.

Lemma 4.2 *Let K be a knot in a homology sphere Y and suppose for some rational homology sphere Z that*

$$Z = Y_{p/q_1}(K) = Y_{p/q_2}(K).$$

Suppose further that there exists $k \in \mathbb{Z}$ such that $q_1 < pk < q_2$. Then

$$p \,|\, \chi(HF_{\mathrm{red}}(Z)).$$

Proof Consider a homology sphere Y_1 given by

$$Y_1 = Y_{1/k}(K).$$

Let K' be the surgery dual of K in Y_1. Denote by μ the meridian of K' and by m and l the meridian and the (preferred) longitude, respectively, of K. Longitudes of K and K' coincide. We view the curves μ, m and l as slopes on the boundary of $Y \setminus \mathrm{nb}(K) = Y_1 \setminus \mathrm{nb}(K')$.

We have $\mu = m + kl$. So $pm + q_1 l = p\mu + (q_1 - pk)l$ and $pm + q_2 l = p\mu + (q_2 - pk)l$. Since $q_1 - pk < 0 < q_2 - pk$, this shows that Z can be obtained by both positive and negative surgery on K' in Y_1. Then, by Proposition 4.1,

$$\chi(HF_{\mathrm{red}}(Z)) = p\,\chi(HF_{\mathrm{red}}(Y_1)) \Rightarrow p \,|\, \chi(HF_{\mathrm{red}}(Z)). \qquad \square$$

We are now in position to prove:

Theorem 4.3 *Let K be a knot in a homology sphere Y. Let Z be a rational homology sphere with $|H_1(Z)| = p$ such that p does not divide $\chi(HF_{\mathrm{red}}(Z))$. Suppose that there exist q_1 and q_2 such that*

$$Z = Y_{p/q_1}(K) = Y_{p/q_2}(K).$$

Then there is no multiple of p between q_1 and q_2. In particular, there are at most $\phi(p)$ surgeries on K that give Z.

Proof If $p/q_1, \ldots, p/q_N$ are distinct slopes that give Z by surgery on K, then by Lemma 4.2 there is $k \in \mathbb{Z}$ such that $pk < q_i < p(k+1)$ for all i (clearly the case $p = 1$ is vacuous). Since the q_i are coprime to p, the conclusion follows. $\square$

If there are spaces that satisfy the condition of Theorem 4.3 and have order of homology 2 then for any knot in any homology sphere there can be at most one slope that

gives such a space by surgery (since $\phi(2) = 1$). Note also that $\dim(HF_{\mathrm{red}}(Z)) \equiv \chi(HF_{\mathrm{red}}(Z)) \pmod 2$, so the condition is then equivalent to $\dim(HF_{\mathrm{red}}(Z))$ being odd. Such spaces do exist and the next corollary demonstrates some.

Corollary 4.4 *Let $Z^1_m = S^2((3,-1),(2,1),(6m-2,-m))$ for odd $m \geq 3$ and Z^2_n be the result of $2/n$–surgery on the figure-eight knot for any odd n. If K is a knot in a homology sphere that gives one of Z^1_m or Z^2_n by surgery of some slope, then this surgery slope is unique.*

Proof Note that Z^1_m is the result of $2/m$–surgery on the right-handed trefoil. It is enough to show that the order of $HF_{\mathrm{red}}(Z^1_m)$ or $HF_{\mathrm{red}}(Z^2_n)$ is odd.

The trefoil has $V_0 = 1$ and $V_1 = 0$ and all $\boldsymbol{A}^{\mathrm{red}}_k(K)$ trivial (since it is a genus 1 L–space knot). The dimension of its reduced Floer homology can be found using the formula of [21, Corollary 3.6] (or see [6, Proposition 16] for the same formula with notation as in this paper). In this case, the dimension is $m-2$ for $m \geq 3$, which is clearly odd.

Since the figure-eight knot is alternating, we can calculate its knot Floer homology from the Alexander polynomial and the signature (see [26, Theorem 6.1]). This (after some calculations) shows that for the figure-eight $\boldsymbol{A}^+_0(K) \cong \mathcal{T}^+_d \oplus \mathcal{T}_{d-1}(1)$ for some d. We also have $V_0 = 0$ and $\boldsymbol{A}^{\mathrm{red}}_k(K) = 0$ for $k \neq 0$. Thus the dimension of the reduced Floer homology of surgery is equal to n by [20, Proposition 5.3]. □

5 Bound on q for knots that are not too exceptional

In this section, we prove Theorem 5.4. First, we deal with knots for which the sequence $\{V_k\}_{k\geq 0}$ is not identically zero.

Lemma 5.1 *Let K be a knot in a homology sphere Y with $V_0 > 0$. Suppose $Z = Y_{p/q}(K)$ for $p \neq 0$. Then*

$$|q| \leq |p| + \frac{\dim(HF_{\mathrm{red}}(Z))}{V_0}.$$

Proof Let $|q| \geq |p|$. It follows from [6, Lemma 13] for positive surgeries and [6, Lemma 17] for negative surgeries that $HF_{\mathrm{red}}(Z,i)$ contains a submodule isomorphic to $\mathcal{T}(V_0)^{\oplus n_i}$, where $n_i = \#\{j \in \mathbb{Z} \mid 0 \leq j < |q|,\ j \equiv i \pmod{|p|}\} - 1$.

So $\dim(HF_{\mathrm{red}}(Z,i)) \geq n_i V_0$. Therefore,

$$\dim(HF_{\mathrm{red}}(Z)) \geq V_0 \sum_{i=0}^{|p|-1} n_i = V_0(|q|-|p|).$$

The desired inequality now follows upon rearranging the terms. □

The mapping cone complex is sometimes unnecessarily large for our purposes. By using some elementary linear algebra contained in the next lemma, we want to be able to pass to a smaller complex when necessary.

Lemma 5.2 *Let T_1, T_2, R_1 and R_2 be graded vector spaces and let $f\colon T_1 \to T_2$, $g\colon R_1 \to R_2$ and $h\colon R_1 \to T_2$ be graded linear maps. Suppose that f is surjective. Then the homology of the complex*

$$0 \to T_1 \oplus R_1 \xrightarrow{D} T_2 \oplus R_2 \to 0,$$

where D is given by $D = \left(\begin{smallmatrix} f & h \\ 0 & g \end{smallmatrix}\right)$, is isomorphic (as a graded vector space) to the direct sum of the kernel of the map f and the homology of the complex

$$0 \to R_1 \xrightarrow{g} R_2 \to 0.$$

Proof It is clear that the cokernels of the maps D and g are isomorphic as graded vector spaces — indeed they are generated by the same elements not in the image of g.

We need to show that the kernels agree. Since f is surjective, there is a graded map $f^*\colon T_2 \to T_1$ such that $f \circ f^* = \mathrm{id}_{T_2}$. Consider the map

$$\theta\colon \ker(D) \to \ker(f) \oplus \ker(g)$$

given by $\theta((t,r)) = (t - f^*(h(r)), r)$. It is easy to see that this map is a graded isomorphism. □

Let $\widetilde{\boldsymbol{v}}_k$ be the restriction of $\boldsymbol{v}_k$ to $\boldsymbol{A}_k^{\mathrm{red}}(K)$ followed by the projection to $\boldsymbol{B}^{\mathrm{red}}$. Define $\widetilde{\boldsymbol{h}}_k$ similarly using $\boldsymbol{h}_k$. Define also $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ to be the restriction of $\boldsymbol{D}^+_{i,p/q}$ to $\mathbb{A}^{\mathrm{red}}_{i,p/q}(K)$ followed by the projection to $\mathbb{B}^{\mathrm{red}}$. The map $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ is a sum of various maps $\widetilde{\boldsymbol{v}}_k$ and $\widetilde{\boldsymbol{h}}_k$.

In terms of the notation in Lemma 5.2 we have $D = \boldsymbol{D}^+_{i,p/q}$, $f = \boldsymbol{D}^T_{i,p/q}$, $g = \widetilde{\boldsymbol{D}}^+_{i,p/q}$, $T_1 = \mathbb{A}^T_{i,p/q}(K)$, $T_2 = \mathbb{B}^T$, $R_1 = \mathbb{A}^{\mathrm{red}}_{i,p/q}(K)$ and $R_2 = \mathbb{B}^{\mathrm{red}}$.

Lemma 5.3 *Let Y be a homology sphere and $K \subset Y$ a knot with $V_0 = 0$. Define $Z = Y_{p/q}(K)$. Then* $\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q}))$, $\dim(\operatorname{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q})) < \infty$,

$$\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q})) + \dim(\operatorname{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \leq \dim(HF_{\mathrm{red}}(Z,i)) + \dim(HF_{\mathrm{red}}(Y))$$

and

$$\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \leq \dim(HF_{\mathrm{red}}(Z,i)).$$

Proof By the proof of Proposition 3.3 (specifically inequality (2)) the tower in $HF^+(Z,i)$ is isomorphic to a direct sum of two pieces (one of which may be trivial). One piece is the kernel of $\boldsymbol{D}^T_{i,p/q}$ (the whole kernel, because $V_0 = 0$ — see eg [6, Lemma 13]). Another piece (which may be trivial) is a subspace of the cokernel of $\boldsymbol{D}^+_{i,p/q}$ (isomorphic to the cokernel of $\widetilde{\boldsymbol{D}}^+_{i,p/q}$) of dimension at most $\dim(HF_{\mathrm{red}}(Y))$.

Now, by Lemma 5.2, the whole resulting homology is isomorphic (in a graded way) to the kernel of $\boldsymbol{D}^T_{i,p/q}$ and the homology of $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ (which correspond to f and g in Lemma 5.2, respectively). The paragraph above implies that the whole tower part is covered by the kernel of $\boldsymbol{D}^T_{i,p/q}$ and perhaps a piece of dimension at most $\dim(HF_{\mathrm{red}}(Y))$. Thus the homology of $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ is finite-dimensional,[3] so $\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q}))$, $\dim(\operatorname{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q})) < \infty$ and since the dimension of the homology of $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ equals $\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q})) + \dim(\operatorname{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q}))$ we have

$$\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q})) + \dim(\operatorname{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \leq \dim(HF_{\mathrm{red}}(Z,i)) + \dim(HF_{\mathrm{red}}(Y))$$

For the second inequality note that the reduced part of $HF^+(Z,i)$ consists of the part in the kernel and the part in the cokernel. If we forget about the part in the cokernel altogether, we can see that the kernel contributes to the dimension exactly $\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q}))$. This verifies the second inequality. □

For an absolutely $\mathbb{Z}_2$–graded abelian group H, let H_e denote the subgroup of elements of grading 0 and H_o denote the subgroup of elements of grading 1.

Recall that we defined

$$N(Y,Z) = 2|H_1(Z)|\dim(HF_{\mathrm{red}}(Y)) + \dim(HF_{\mathrm{red}}(Z)).$$

We are now ready to prove the main theorem of this section.

[3] The structure of the graded vector spaces here allows us to talk about the dimensions: once we identified what covers the elements of high enough grading in the tower, the rest has to happen in a finite-dimensional vector space.

Theorem 5.4 *Let Y be a non-L–space homology sphere, Z be a rational homology sphere and $K \subset Y$ be a knot, and suppose there are coprime integers p and q such that $Z = Y_{p/q}(K)$.*

If $|q| > N(Y, Z)$, then

- $V_0(K) = 0$;
- $\Delta_K \equiv 1$;
- $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) = \dim(HF_{\mathrm{red}}(Y)_e)$ *for all* k;
- $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_o) = \dim(HF_{\mathrm{red}}(Y)_o)$ *for all* k.

Proof Suppose $|q| > N(Y, Z)$. We know from Lemma 5.1 that $V_0(K) = 0$. Since changing the orientation of a manifold does not change the dimension of its reduced Floer homology [25, Proposition 2.5], we can assume that $p > 0$ and $q > 0$.

Claim $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) \geq \dim(\boldsymbol{B}_e^{\mathrm{red}})$ *and* $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_o) \geq \dim(\boldsymbol{B}_o^{\mathrm{red}})$.

Proof The even and odd cases are completely analogous so we only prove

$$\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) \geq \dim(\boldsymbol{B}_e^{\mathrm{red}}).$$

Suppose for contradiction that $\dim(\boldsymbol{B}_e^{\mathrm{red}}) - \dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) \geq 1$.

Let B_i be the sum of all $(n, \boldsymbol{B}_e^{\mathrm{red}})$ with n satisfying

$$\left\lfloor \frac{i + pn}{q} \right\rfloor = \left\lfloor \frac{i + p(n-1)}{q} \right\rfloor = k.$$

Let A_i be the sum of all $(n, \boldsymbol{A}_k^{\mathrm{red}}(K)_e)$ with n satisfying $\lfloor (i + pn)/q \rfloor = k$. Then $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ maps A_i into B_i, so $\dim(\mathrm{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \geq \dim(B_i) - \dim(A_i)$.

Define $N_i = \{\, j \mid j \equiv i \pmod p,\ \lfloor j/q \rfloor = k\}$. Then $\dim(A_i) = N_i \dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e)$ and $\dim(B_i) = (N_i - 1)\dim(\boldsymbol{B}_e^{\mathrm{red}})$.

By Lemma 5.3 we have

$$\begin{aligned}\dim(HF_{\mathrm{red}}(Z, i)) + \dim(HF_{\mathrm{red}}(Y)) &\geq \dim(\mathrm{coker}(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \\ &\geq N_i(\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) - \dim(\boldsymbol{B}_e^{\mathrm{red}})) - \dim(\boldsymbol{B}_e^{\mathrm{red}}) \\ &\geq N_i - \dim(HF_{\mathrm{red}}(Y)).\end{aligned}$$

Noting that $\sum_{i=0}^{p-1} N_i = q$ and summing over all Spin^c structures, we get

$$\dim(HF_{\mathrm{red}}(Z)) + p\dim(HF_{\mathrm{red}}(Y)) \geq q - p\dim(HF_{\mathrm{red}}(Y)),$$

which contradicts the assumption made on q. ◁

Claim $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) \le \dim(\boldsymbol{B}_e^{\mathrm{red}})$ *and* $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_o) \le \dim(\boldsymbol{B}_o^{\mathrm{red}})$.

Proof Again, the two cases are analogous, so we only show that $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) \le \dim(\boldsymbol{B}_e^{\mathrm{red}})$. Suppose for a contradiction that $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) - \dim(\boldsymbol{B}_e^{\mathrm{red}}) \ge 1$.

Let $\widehat{A}_i$ be the sum of all $(n, \boldsymbol{A}_k^{\mathrm{red}}(K)_e)$ with n satisfying

$$\left\lfloor \frac{i+pn}{q} \right\rfloor = \left\lfloor \frac{i+p(n+1)}{q} \right\rfloor = k.$$

Let $\widehat{B}_i$ be the sum of all $((n, \boldsymbol{B}_e^{\mathrm{red}})$ with n satisfying $\lfloor (i+pn)/q \rfloor = k$.

Clearly $\widetilde{\boldsymbol{D}}^+_{i,p/q}$ maps $\widehat{A}_i$ into $\widehat{B}_i$, so $\dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \ge \dim(\widehat{A}_i) - \dim(\widehat{B}_i))$.

We have $\dim(\widehat{B}_i) = N_i \dim(\boldsymbol{B}_e^{\mathrm{red}})$ and $\dim(\widehat{A}_i) = (N_i - 1)\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e)$.

Hence, by Lemma 5.3, we have

$$\begin{aligned}\dim(HF_{\mathrm{red}}(Z,i)) &\ge \dim(\ker(\widetilde{\boldsymbol{D}}^+_{i,p/q})) \\ &\ge (N_i - 1)(\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) - \dim(\boldsymbol{B}_e^{\mathrm{red}})) - \dim(\boldsymbol{B}_e^{\mathrm{red}}) \\ &\ge N_i - 1 - \dim(HF_{\mathrm{red}}(Y)).\end{aligned}$$

Summing over all Spin^c structures, we get

$$\dim(HF_{\mathrm{red}}(Z)) \ge q - p - p\dim(HF_{\mathrm{red}}(Y)),$$

which is again a contradiction to the assumed inequality for q. ◁

Combining the results in the two claims, we see that the assumption that q violates the bound in the statement of the lemma implies that for all k we have $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) = \dim(\boldsymbol{B}_e^{\mathrm{red}})$ and $\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_o) = \dim(\boldsymbol{B}_o^{\mathrm{red}})$. Thus $\chi(\boldsymbol{A}_k^{\mathrm{red}}(K)) = \chi(\boldsymbol{B}^{\mathrm{red}})$ for all k.

Let K be a knot and $\Delta_K(T) = a_0 + \sum_i a_i(T^i + T^{-i})$ be its symmetrised Alexander polynomial, with normalisation convention $\Delta_K(1) = 1$. Define its *torsion coefficients* $t_i(K)$ for $i \ge 0$ by

$$t_i(K) = \sum_{j \ge 1} j a_{i+j}.$$

We now want to show that $t_k(K) = \chi(\boldsymbol{A}_k^{\mathrm{red}}(K)) - \chi(\boldsymbol{B}^{\mathrm{red}})$. This will imply that all torsion coefficients of K are 0 and thus its Alexander polynomial is trivial.

Define $\Delta_k = C\{j \ge k \text{ and } i < 0\}$. Note that $\chi(\Delta_k) = t_k(K)$. We have an exact sequence

$$0 \to \Delta_k \xrightarrow{i} \boldsymbol{A}_k^+(K) \xrightarrow{v_k} \boldsymbol{B}^+ \to 0,$$

which leads to an exact triangle

$$\begin{array}{ccc} H_*(\Delta_k) & \xrightarrow{i_*} & \boldsymbol{A}_k^+(K) \\ & \nwarrow & \downarrow \boldsymbol{v}_k \\ & & \boldsymbol{B}^+ \end{array} \tag{13}$$

Since $V_0 = 0$, the map $\boldsymbol{v}_k^T$ maps $\boldsymbol{A}_k^T(K)$ isomorphically onto $\boldsymbol{B}^T$. So, up to graded isomorphism, we also have an exact triangle

$$\begin{array}{ccc} H_*(\Delta_k) & \xrightarrow{i_*} & \boldsymbol{A}_k^{\mathrm{red}}(K) \\ & \nwarrow & \downarrow \widetilde{\boldsymbol{v}}_k \\ & & \boldsymbol{B}^{\mathrm{red}} \end{array} \tag{14}$$

It follows that $t_k(K) = \chi(\Delta_k) = \chi(\boldsymbol{A}_k^{\mathrm{red}}(K)) - \chi(\boldsymbol{B}^{\mathrm{red}}) = 0$. □

6 A bound on q for exceptional knots of genus larger than 1

In this section we want to show that for knots that do not satisfy the bound of Theorem 5.4 and have genus larger than 1, the integer q is still bounded by a quantity depending only on the pair of manifolds connected by surgery. We first prove the following lemma, a version of which for S^3 was proven in [10, Lemma 2.5].

Lemma 6.1 *Let $\{V_k\}_{k\in\mathbb{Z}}$ and $\{H_k\}_{k\in\mathbb{Z}}$ be numbers associated with a knot K in a homology sphere, as defined in Section 2. Then*

$$H_k - V_k = k$$

for all $k \in \mathbb{Z}$.

Proof According to [30, Theorem 2.3], the modules $\boldsymbol{A}_k^+(K)$ can be identified with HF^+ of N–surgeries on K (in a certain Spin^c structure), where N is a sufficiently large integer. Moreover, after this identification, the maps $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$ coincide with the maps into $HF^+(Y)$ induced by the (turned-around) surgery cobordism.

More specifically, the maps $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$ can be thought of as the maps corresponding to the Spin^c structures $\mathfrak{v}_k$ and $\mathfrak{h}_k$, respectively, where

$$\langle c_1(\mathfrak{v}_k), [\widehat{F}]\rangle + N = 2k$$

and

$$\langle c_1(\mathfrak{h}_k), [\widehat{F}]\rangle - N = 2k.$$

Here $[\widehat{F}]$ is the homology class of the surface obtained by capping off a Seifert surface F of K with the core of the 2–handle.

From this we can deduce that $c_1(\mathfrak{v}_k)^2 = -\frac{1}{N}(2k-N)^2$ and $c_1(\mathfrak{h}_k)^2 = -\frac{1}{N}(2k+N)^2$ (see [15, Proposition 2.69] for a nice exposition of this calculation). The difference in the grading shifts of the two maps identified with $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$ is given by $2(H_k - V_k)$. On the other hand, we can deduce from [28, Theorem 7.1] that the difference in the grading shifts is also given by

$$\tfrac{1}{4}(c_1(\mathfrak{v}_k)^2 - c_1(\mathfrak{h}_k)^2) = 2k.$$

Comparing the two expressions, we get the desired result. □

For two homogeneous elements u and v in the mapping cone complex, denote their relative $\mathbb{Z}$–grading by $\deg(u,v)$. For a homogeneous element w of Heegaard Floer homology of some rational homology sphere, denote by $\deg(w)$ its absolute $\mathbb{Q}$–grading. Recall that the modules $\boldsymbol{A}_k^+(K)$ and $\boldsymbol{B}^+$ decompose as the sum of the "tower" $\mathcal{T}^+$ and the reduced part. For a homogeneous element c in either one of $\boldsymbol{A}_k^+(K)$ or $\boldsymbol{B}^+$ denote by $\widetilde{d}(c)$ its relative grading with the 1 in the tower part (ie $\widetilde{d}(c) = \deg(c,1)$).

As already mentioned, if Y is an L–space homology sphere, there is a bound on q similar to that of Theorem 5.4 that holds for all knots. Thus, as before, we will assume throughout this section that Y is a non-L–space homology sphere.

Lemma 6.2 *Let $K \subset Y$ be a knot and suppose $Z = Y_{p/q}(K)$, where $p, q > 0$. Let $N(Y,Z)$ be defined as in the statement of Theorem 5.4 and suppose $q > N(Y,Z)$. Then, for every homogeneous $z \in \boldsymbol{A}_k^{\mathrm{red}}(K)$,*

$$\widetilde{d}(z) \geq \min\{\widetilde{d}(c) \mid c \in \boldsymbol{B}^{\mathrm{red}} \text{ is homogeneous}\}.$$

Proof Suppose there exists k and $z \in \boldsymbol{A}_k^{\mathrm{red}}(K)$ with

$$\widetilde{d}(z) < \min\{\widetilde{d}(c) \mid c \in \boldsymbol{B}^{\mathrm{red}} \text{ is homogeneous}\}.$$

Both $\boldsymbol{v}_k$ and $\boldsymbol{h}_k$ do not increase $\widetilde{d}$, so z is in the kernel of both $\widetilde{\boldsymbol{v}}_k$ and $\widetilde{\boldsymbol{h}}_k$, hence also in the kernel of $\widetilde{\boldsymbol{D}}^+_{i,p/q}$. This holds for every copy of $\boldsymbol{A}_k^{\mathrm{red}}(K)$ in the mapping cone complexes for all Spin^c structures, so, summing contributions from all Spin^c structures and using Lemma 5.3, we deduce

$$q \leq \dim(HF_{\mathrm{red}}(Z)) < N(Y,Z).$$

This is a contradiction. □

Recall that for a rational homology sphere Z we defined

$$\hat{D}(Z) = \max\{\deg(z) - d(Z,\mathfrak{t}) \mid z \in HF_{\mathrm{red}}(Z,\mathfrak{t}) \text{ is homogeneous, } \mathfrak{t} \in \mathrm{Spin}^c(Z)\}$$

and

$$\check{D}(Z) = \min\{\deg(z) - d(Z,\mathfrak{t}) \mid z \in HF_{\mathrm{red}}(Z,\mathfrak{t}) \text{ is homogeneous, } \mathfrak{t} \in \mathrm{Spin}^c(Z)\},$$

where $\deg(z)$ is the absolute grading of z.

We can now formulate a bound on q that holds for knots that have genus > 1 and are not covered by Theorem 5.4.

Proposition 6.3 *Let $K \subset Y$ be a knot such that $Z = Y_{p/q}(K)$ for $p, q > 0$ and $q > N(Y,Z)$. Suppose the genus of K is larger than 1. Then*

$$\lfloor q/p \rfloor \leq \tfrac{1}{2}(\hat{D}(Z) - \check{D}(Y)).$$

Proof By Theorem 2.1 and the exact triangle (13), the map $\boldsymbol{v}^+_{g-1}$ must not be an isomorphism (where g is the genus of K). Since $V_0 = 0$, the map $\boldsymbol{v}^T_{g-1}$ is an isomorphism, so the map $\widetilde{\boldsymbol{v}}_{g-1}$ must not be an isomorphism. Since the spaces $\boldsymbol{A}^{\mathrm{red}}_{g-1}(K)$ and $\boldsymbol{B}^{\mathrm{red}}$ have the same dimension, the map $\widetilde{\boldsymbol{v}}_{g-1}$ must have some kernel. Suppose $z \in \ker(\widetilde{\boldsymbol{v}}_{g-1})$. By adding an element of $\boldsymbol{A}^T_{g-1}(K)$ if necessary, we can assume that $z \in \ker(\boldsymbol{v}^+_{g-1})$.

Let $N = \max\{n \mid \lfloor pn/q \rfloor = g-1\}$. Then $(N,z) \in (N, \boldsymbol{A}^+_{g-1}(K))$ is in the kernel of $\boldsymbol{v}^+_{g-1}$. By Lemma 3.1, we can assume that (N,z) is in the kernel of $\boldsymbol{D}^+_{0,p/q}$. Denoting as usual by 1 the generators of the kernel of U in the tower modules, we have

$$\deg((N,1),(0,1)) = 2\sum_{i=1}^{g-1} n_i H_i,$$

where $n_i \geq \lfloor q/p \rfloor$. Since $V_0 = 0$, by Lemma 6.1 we have $H_i = i$ for $i \geq 0$. So

$$\deg((N,1),(0,1)) \geq g(g-1)\lfloor q/p \rfloor.$$

By the proof of Proposition 3.3, $\mathrm{gr}(0,1) \geq d(Z,0)$. In homology, (N,z) represents some element of $HF_{\mathrm{red}}(Z,0)$. Suppose its absolute grading there is G.

Then $G - d(Z,0) \geq \deg((N,z),(0,1)) = \widetilde{d}(z) + \deg((N,1),(0,1))$. By Lemma 6.2,

$$\widetilde{d}(z) \geq \check{D}(Y).$$

Combining the various inequalities, we obtain

$$\lfloor q/p \rfloor \leq \frac{G - d(Z,0) - \check{D}(Y)}{g(g-1)} \leq \frac{\hat{D}(Z) - \check{D}(Y)}{2}. \qquad \square$$

We can combine Theorem 5.4 and Proposition 6.3 into the following:

Corollary 6.4 *Let Y be a homology sphere and $K \subset Y$ a knot. Suppose $Y_{p/q}(K) = Z$ for $p \neq 0$. There exists a constant $C(Y,Z)$ that depends only on the Heegaard Floer homology of Y and Z such that if $q > C(Y,Z)$ then*

- *the genus of K is 1;*
- *K has trivial Alexander polynomial;*
- *$V_0(K) = 0$;*
- *$\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_e) = \dim(HF_{\mathrm{red}}(Y)_e)$ for all k;*
- *$\dim(\boldsymbol{A}_k^{\mathrm{red}}(K)_o) = \dim(HF_{\mathrm{red}}(Y)_o)$ for all k.*

7 Some knots determined by their complements

Results of Section 4 can be applied to show that, in certain homology $\mathbb{R}P^3$, non-null-homologous knots are determined by their oriented complements.

Corollary 7.1 *Let Z be a closed, connected, oriented manifold with $|H_1(Z)| = 2$. Suppose that $\dim(HF_{\mathrm{red}}(Z))$ is odd. Then non-null-homologous knots in Z are determined by their complements.*

Proof By Theorem 4.3, if a knot in one of these spaces has a homology sphere surgery, then it is determined by its complement. Thus it will be enough to show that non-null-homologous knots in such spaces have homology sphere surgeries.

Let Z be a space as in the statement. Denote by S a solid torus regular neighbourhood of K and denote the exterior of K by Z_0. Let T be the boundary of S. Consider the exact sequence for the pair (Z,S),

$$0 \to H_2(Z,S) \to H_1(S) \to H_1(Z) \to H_1(Z,S) \to 0.$$

By considering this sequence with coefficients in $\mathbb{Q}$, $H_2(Z,S)$ is a direct sum of one copy of $\mathbb{Z}$ with a torsion group. Then excision and Poincaré–Lefschetz duality show that $H_2(Z,S) \cong H_2(Z_0,T) \cong H^1(Z_0)$. The latter group is free abelian, thus

$H_2(Z_0, T) \cong \mathbb{Z}$. Similarly, $H_1(Z, S) \cong H^2(Z_0)$, which is equal (by the universal coefficients theorem) to the torsion part of $H_1(Z_0)$; call it G. We have $H_1(Z_0) \cong \mathbb{Z} \oplus G$. Now the sequence above becomes

$$0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}_2 \to G \to 0.$$

Since K is not null-homologous (but obviously $2K$ is), there is a rational Seifert surface for K that winds twice longitudinally. This implies that the map between two copies of $\mathbb{Z}$ in the sequence above is multiplication by 2. It follows that $G = 0$.

Now consider the exact sequence of the pair (T, Z_0),

$$\cdots \to H_2(Z_0, T) \to H_1(T) \to H_1(Z_0) \to 0.$$

This sequence shows that $H_1(Z_0)$ is generated by the images of the meridian and the longitude. This means that if we perform a surgery with a slope given by the generator of $H_1(Z_0)$, we get a homology sphere. □

The Brieskorn sphere $\Sigma(2,3,7)$ has perhaps the simplest Heegaard Floer homology a non-L–space can possibly have — the rank of its reduced Floer homology is 1. This makes it possible to show that "most" knots in $\Sigma(2,3,7)$ are determined by their complements. The first part of the proof shows that a surgery from $\Sigma(2,3,7)$ to $\Sigma(2,3,7)$ must be integral. The thinking behind the proof is similar to that of Section 5, but we get a better bound due to the fact that $1 < 2$ and linear maps into 1–dimensional spaces are either trivial or surjective.

Theorem 7.2 *Knots of genus larger than 1 in the Brieskorn sphere $\Sigma(2,3,7)$ are determined by their complements. Moreover, if $K \subset \Sigma(2,3,7)$ is a counterexample to Conjecture 1.2, then the surgery slope is integral, $\widehat{HFK}(\Sigma(2,3,7), K, 1)$ has dimension 2 and its generators lie in different $\mathbb{Z}_2$–gradings.*

Nonfibred knots of genus larger than 1 in $\Sigma(2,3,7)$ are strongly determined by their complements.

Proof The Heegaard Floer homology of $-\Sigma(2,3,7)$ has been computed in [22, Section 8.1]. Alternatively, we can calculate it using the program HFNem2 by Çağrı Karakurt.[4] It is clear that proving what we want for knots in $-\Sigma(2,3,7)$ is equivalent to proving it for knots in $\Sigma(2,3,7)$.

[4] At the time of writing available for download at https://www.ma.utexas.edu/users/karakurt/.

The calculation shows that $-\Sigma(2,3,7)$ has reduced Floer homology of rank 1, situated in the absolute grading 0 which is equal to the d–invariant of $-\Sigma(2,3,7)$. By [25, Proposition 2.5; 22, Proposition 4.2], reduced Floer homology of $\Sigma(2,3,7)$ has rank 1 (and odd absolute $\mathbb{Z}_2$–grading) and d–invariant 0.

Suppose a knot $K \subset \Sigma(2,3,7)$ gives $\Sigma(2,3,7)$ by $1/q$–surgery. Since the analogous statement can be proven for $-\Sigma(2,3,7)$, we can assume that $q > 0$.

Let $Y = \Sigma(2,3,7)$. Suppose for a contradiction that $q > 1$. Note that by Proposition 3.3 we must have $V_0 = 0$. Consider $(qg-1, \boldsymbol{A}^+_{g-1})$, where $g = g(K)$. This is the "rightmost" group for which the corresponding map $\boldsymbol{v}$ is not an isomorphism.

Suppose $\boldsymbol{A}^{\mathrm{red}}_{g-1} = 0$. Since $q \geq 2$, the group $(qg-1, \boldsymbol{B}^{\mathrm{red}})$ is not in the image of $\boldsymbol{D}^+_{i,p/q}$. Thus it gives rise to a generator of $HF_{\mathrm{red}}(Y)$. However, this element is in the even $\mathbb{Z}_2$–grading. This gives a contradiction.

Now assume $\dim(\boldsymbol{A}^{\mathrm{red}}_{g-1}) \geq 1$. Since the map $\boldsymbol{v}_{g-1}$ cannot be an isomorphism, the map $\boldsymbol{v}^{\mathrm{red}}_{g-1}$ must have kernel. According to Lemma 3.1, this element of the kernel gives rise to an element in $HF_{\mathrm{red}}(Y)$. However, applying the same argument "on the other end", ie to $\boldsymbol{h}_{-(g-1)}$, we see that we must have $\dim(HF_{\mathrm{red}}(Y)) \geq 2$, which gives a contradiction.

All in all, we have $q = 1$, irrespective of the genus. Now assume $g > 1$. According to [25, Theorem 9.1], there is an exact triangle of $\mathbb{F}[U]$–modules

$$\begin{array}{ccc} HF^+(Y) & \xrightarrow{f} & HF^+(Y_0(K)) \\ & \overset{h}{\nwarrow} & \downarrow g \\ & & HF^+(Y_1(K)) \end{array} \tag{15}$$

Much as in the proof of [24, Corollary 4.5], we have[5]

$$HF^+(Y_0(K), g-1) \cong HF^+(Y_0(K), -(g-1)) \cong \widehat{HFK}(Y, K, g),$$

so these groups are nontrivial by Theorem 2.1.

By [25, Theorem 10.4], $HF^+(Y_0(K), 0) \cong \mathcal{T}_{d_{-1/2}(Y_0(K))} \oplus \mathcal{T}_{d_{+1/2}(Y_0(K))} \oplus R$, where $R = HF_{\mathrm{red}}(Y_0(K), 0)$ is a finite-dimensional vector space in the kernel of some power of U and $d_{\pm 1/2}(Y_0(K)) \in \mathbb{Z} + \frac{1}{2}$ is an analogue of the correction term.

[5]The labellings of the Spin^c structures are obtained by taking the half of the pairing of their Chern classes with the generator of $H^2(Y_0(K))$ obtained by capping a Seifert surface.

By [22, Lemma 3.1], the component of f mapping into $HF^+(Y_0(K),0)$ has grading $-\frac{1}{2}$ and the restriction of g to $HF^+(Y_0(K),0)$ also has grading $-\frac{1}{2}$.

Since in high enough gradings the group $HF^+(Y_0(K),0)$ consists of two tower modules (in different relative gradings) and only one of them can be in the image of the map f, it follows that the other one has to map onto the tower in $HF^+(Y_1(K))$. It follows that the restriction of h to the tower part of $HF^+(Y_1(K))$ is zero. Since we assume $g>1$ and so $HF^+(Y_0(K))$ contains two nontrivial vector spaces $HF^+(Y_0(K),\pm(g-1))$, the map h has to be trivial. If h was not trivial, it would surject onto the reduced part of $HF^+(Y)$ and so there would be nothing left to map onto $HF^+(Y_0(K),\pm(g-1))$. Thus we can assume $h=0$. Then the maps f and g map a tower module isomorphically onto another one, so comparing the gradings we see that $d_{\pm 1/2}(Y_0(K))=\pm\frac{1}{2}$.

Moreover, the dimension of $HF_{\mathrm{red}}(Y_0(K))$ has to be 2, thus

$$HF^+(Y_0(K),g-1)\cong HF^+(Y_0(K),-(g-1))\cong \widehat{HFK}(Y,K,g)\cong \mathbb{F}_2,$$
$$HF^+(Y_0(K),k)=0$$

for $k\notin\{0,\pm(g-1)\}$ and $R=0$.

By [25, Proposition 10.14 and Theorem 10.17], we have

$$t_k(K)=0$$

for $0\le k<g$ and $t_{g-1}(K)=\pm 1$.

Notice that

$$\Delta_K''(1)=2t_0(K)+4\sum_{i=1}^{g-1}t_i(K)=\pm 4\neq 0,$$

which contradicts (4).

If $g=q=1$, then, by the reasoning of Lemma 3.1, $\boldsymbol{A}_0^{\mathrm{red}}\cong HF_{\mathrm{red}}(Y)=\mathbb{F}_2$. Since $\boldsymbol{v}_0^{\mathrm{red}}$ cannot be an isomorphism, it must be zero. It follows that the dimension of $\widehat{HFK}(Y,K,1)$ is 2. Since $\Delta_K''(1)=0$ forces the Alexander polynomial to be trivial, the two generators have to be in different $\mathbb{Z}_2$–gradings.

Now suppose $K\subset Y$ produces $-Y$ by $1/q$–surgery for $q>0$. Moreover, let $g(K)>1$. Just as before we cannot have $\dim(\boldsymbol{A}_{g-1}^{\mathrm{red}})\ge 1$, thus $\boldsymbol{A}_{g-1}^{\mathrm{red}}=0$. Since the d–invariants of Y and $-Y$ coincide, we have $V_0=0$, thus the map $\boldsymbol{v}_{g-1}^{\mathrm{red}}$ has no kernel and has cokernel of dimension 1. It follows that $\dim\big(\widehat{HFK}(Y,K,g(K))\big)=1$, so, by [17, Theorem 1.1] or [11, Theorem 9.11], K is fibred. □

8 Null-homologous knots in L–spaces

The aim of this section is to show that null-homologous knots in L–spaces are determined by their complements. Moreover, we will also show that all knots in lens spaces of the form $L(p,q)$ with p square-free are determined by their complements.

First of all, we need to verify that the mapping cone formula for rational surgeries as proved in [31] for knots in integer homology spheres also applies to null-homologous knots in rational homology spheres with only minor modifications.

Let Y be a rational homology sphere and K a null-homologous knot in it. Let Y_0 be the exterior of K. To understand relative Spin^c structures, we first need to calculate the first homology.

Lemma 8.1 *We have $H_1(Y_0) \cong \mathbb{Z} \oplus H_1(Y)$, where the first factor is the group generated by the meridian of K.*

Proof Let S be a (closed) regular neighbourhood of K and $T = \partial Y_0$ its boundary. Consider the long exact sequence for the pair,

$$0 \to H_2(Y,S) \to H_1(S) \to H_1(Y) \to_1 (Y,S) \to 0.$$

By excision, $H_*(Y,S) \cong H_*(Y_0,T)$. Since K is null-homologous, the boundary of its Seifert surface generates $H_1(S)$. Thus the map $H_2(Y,S) \to H_1(S)$ is an isomorphism.

It follows that $H_1(Y) \cong H_1(Y_0,T)$. By Poincaré–Lefschetz duality, $H_1(Y_0,T) \cong H^2(Y_0)$ and, by the universal coefficients theorem, the torsion part of $H_1(Y_0)$ is equal to the torsion part of $H^2(Y_0)$, which, by the previous sentence, equals $H_1(Y)$.

Now consider the Mayer–Vietoris sequence

$$0 \to H_1(T) \to H_1(S) \oplus H_1(Y_0) \to H_1(Y) \to 0.$$

The same sequence with rational coefficients shows that $H_1(Y_0)$ has rank 1, so, combining with the argument above, $H_1(Y_0) \cong \mathbb{Z} \oplus H_1(Y)$. Moreover, the map $H_1(T) \to H_1(S)$ takes the longitude to a generator and the meridian to 0 and also the longitude is trivial in $H_1(Y_0)$. It follows that we have an exact sequence

$$0 \to \langle m \rangle = \mathbb{Z} \xrightarrow{f} H_1(Y_0) \cong \mathbb{Z} \oplus H_1(Y) \to H_1(Y) \to 0.$$

Since the map f is an injection, the subgroup $0 \oplus H_1(Y)$ is disjoint from the image of f. Hence it is mapped injectively into $H_1(Y)$. However, since this group is finite,

the restriction of the second map from the right to $H_1(Y)$ is an isomorphism. It follows that the $\mathbb{Z}$ factor of $H_1(Y_0)$ (perhaps after changing the splitting) is equal to the kernel of this map and thus is generated by the meridian. □

Since $H_1(Y_0)$ acts freely and transitively on the set of relative Spin^c structures on Y_0, we can identify (noncanonically) this group with the set of relative Spin^c structures. In other words, we label relative Spin^c structures on Y_0 by a pair (n, h) with $n \in \mathbb{Z}$ and $h \in H_1(Y)$. Moreover, by adding a multiple of the meridian we can ensure that

$$\langle c_1((n,h)), [F]\rangle = 2n,$$

where F is a Seifert surface for K.

Recall from [31, Section 7] that doing p/q surgery on K in Y is equivalent to doing integral surgery with slope $a = \lfloor p/q \rfloor$ on the knot $\widehat{K} = K \# O_{q/r}$ in $\widehat{Y} = Y \#(-L(q,r))$. Here $p = aq + r$ and $O_{q/r}$ is a core of one of the Heegaard solid tori in $-L(q,r)$, thought of as the image of one component of the Hopf link after the $-q/r$–surgery on the other component.

Denote by $\widehat{Y}_0$ and L_0 the exteriors of $\widehat{K}_0$ and $O_{q/r}$, respectively. Notice that $\widehat{Y}_0$ is obtained by glueing Y_0 and L_0 along an annulus A whose core maps to meridians of K and $O_{q/r}$. Consider the associated Mayer–Vietoris sequence

$$\cdots \to H_1(A) \to H_1(Y_0) \oplus H_1(L_0) \to H_1(\widehat{Y}_0) \to 0.$$

We can rewrite it as

$$\cdots \to \mathbb{Z} \to \mathbb{Z} \oplus H_1(Y) \oplus \mathbb{Z} \to H_1(\widehat{Y}_0) \to 0.$$

Moreover, the generator of $H_1(A)$ is mapped to meridians of both knots. Thus the second map from the right is given by $1 \mapsto (1, 0, -q)$. It follows that we can write $H_1(\widehat{Y}_0) \cong \mathbb{Z} \oplus H_1(Y)$. Moreover, the meridian of K maps to q times the generator of $\mathbb{Z}$ and the meridian of $O_{q/r}$ maps to the generator of $\mathbb{Z}$. As in the case of homology spheres, the push-off of K with respect to the framing a is mapped to p times the generator of the $\mathbb{Z}$ summand.

Now, just as in [31, Proof of Theorem 1.1], we can assemble the mapping cone, whose homology will coincide with the Heegaard Floer homology of p/q–surgery on K. The only difference with the case of homology spheres is that instead of indices i (that represented relative Spin^c structures) we have to use pairs (i, h), but when considering every Spin^c structure on the resulting space separately, h stays the same.

We use the same notation as in [31], replacing relative Spin^c structures with the labelling we defined after the proof of Lemma 8.1, ie (n,h). Note that $B^+_{(n,h)}$ only depends on h (up to a shift in the filtration) and in homology gives $HF^+(Y,h)$.[6] Consequently, we denote this group simply by B^+_h.

For each $h \in H_1(Y)$ and $0 \le i < p$, consider the set of groups $(s, A^+_{(\lfloor (i+ps)/q \rfloor, h)})$ for $s \in \mathbb{Z}$. Combine them into

$$\mathcal{A}^+_{(i,h)} = \bigoplus_{s\in\mathbb{Z}} (s, A^+_{(\lfloor (i+ps)/q \rfloor, h)}).$$

Similarly, define

$$\mathcal{B}_{(i,h)} = \bigoplus_{s\in\mathbb{Z}} (s, B^+_h).$$

Define $D^+_{(i,h),p/q} \colon \mathcal{A}^+_{(i,h)} \to \mathcal{B}_{(i,h)}$ componentwise by

$$D^+_{(i,h),p/q}(s,a_s) = (s, v^+_{(\lfloor (i+ps)/q \rfloor, h)}(a_s)) + (s+1, h^+_{(\lfloor (i+ps)/q \rfloor, h)}(a_s)).$$

Denote by $\mathbb{X}^+_{(i,h),p/q}$ the mapping cone of $D^+_{(i,h),p/q}$.Then the Heegaard Floer homology of $Y_{p/q}(K)$ in a certain Spin^c structure is given by the homology of $\mathbb{X}^+_{(i,h),p/q}$. We denote this Spin^c structure on $Y_{p/q}(K)$ by (i,h). In other words, we have

$$H_*(\mathbb{X}^+_{(i,h),p/q}) \cong HF^+(Y_{p/q}(K),(i,h)).$$

Reusing the notation from above, let $\boldsymbol{A}^+_{(n,h)}$, $\boldsymbol{B}^+_h$, $\mathbb{A}^+_{(i,h)}$ and $\mathbb{B}_{(i,h)}$ be the homologies of $A^+_{(n,h)}$, B^+_h, $\mathcal{A}^+_{(i,h)}$ and $\mathcal{B}_{(i,h)}$, respectively. Let $\boldsymbol{v}^+_{(n,h)}$, $\boldsymbol{h}^+_{(n,h)}$ and $\boldsymbol{D}^+_{(i,h),p/q}$ be the maps induced by $v^+_{(n,h)}$, $h^+_{(n,h)}$ and $D^+_{(i,h),p/q}$, respectively, in homology.

As before, we denote by $\boldsymbol{D}^T_{(i,h),p/q}$, $\boldsymbol{v}^T_{(n,h)}$ and $\boldsymbol{h}^T_{(n,h)}$ the restrictions of $\boldsymbol{D}^+_{(i,h),p/q}$, $\boldsymbol{v}^+_{(n,h)}$ and $\boldsymbol{h}^+_{(n,h)}$ to the tower parts (in the case of $\boldsymbol{D}^T_{(i,h),p/q}$, the restriction to the sum of the tower parts).

The maps $\boldsymbol{v}^T_{(n,h)}$ and $\boldsymbol{h}^T_{(n,h)}$ are multiplications by powers of U. Denote these powers by $V_{(n,h)}$ and $H_{(n,h)}$, respectively. For each $h \in H_1(Y)$ we have:

- $V_{(n,h)} \ge V_{(n+1,h)}$ and $H_{(n,h)} \le H_{(n+1,h)}$.
- There is $N \in \mathbb{N}$ such that $V_{(n,h)} = 0$ for all $n \ge N$ and $H_{(n,h)} = 0$ for all $n \le -N$.
- $V_{(n,h)} \to +\infty$ as $n \to -\infty$ and $H_{(n,h)} \to +\infty$ as $n \to +\infty$.

[6] The statement should be read as saying that there is an identification between $H_1(Y)$ and $\mathrm{Spin}^c(Y)$ such that this correspondence holds.

By [6, Lemma 12], the map $\boldsymbol{D}^T_{(i,h),p/q}$ is surjective (when $p/q > 0$), so

$$HF^+(Y_{p/q}(K),(i,h)) \cong \ker(\boldsymbol{D}^+_{(i,h),p/q}).$$

If one of $\boldsymbol{A}^+_{(n,h)}$ is not a tower, ie contains some reduced part (which we denote by $\boldsymbol{A}^{\mathrm{red}}_{(n,h)}$), then every element of $\boldsymbol{A}^{\mathrm{red}}_{(n,h)}$ will be a component of some element of the kernel of $\boldsymbol{D}^+_{(i,h),p/q}$. However, such an element will not be in the image of large enough power of U. It follows that $HF^+(Y_{p/q}(K),(i,h))$ will have some reduced Floer homology. Thus if $Y_{p/q}(K)$ is an L–space, then $\boldsymbol{A}^+_{(n,h)} \cong \mathcal{T}^+$ for all n and h.

Write $\widehat{A}_{(n,h)} = \ker(U\colon A^+_{(n,h)} \to A^+_{(n,h)})$ and denote its homology by $\widehat{\boldsymbol{A}}_{(n,h)}$. Since $\boldsymbol{A}^+_{(n,h)} \cong \mathcal{T}^+$ for all n and h, we have $\widehat{\boldsymbol{A}}_{(n,h)} \cong \mathbb{F}$ for all n and h.

8.1 Alexander polynomial

Just as in the case of homology spheres, given a knot $K \subset Y$ in a rational homology sphere, one can define its Alexander module to be the first homology of the covering space $\widehat{Y}$ of $Y \setminus K$ with deck transformation group $\mathbb{Z}$. The two differences are as follows: firstly, to define $\widehat{Y}$ instead of the abelianisation map we use $\phi\colon \pi_1(Y \setminus K) \to \mathbb{Z}$ gotten by composing abelianisation with the projection onto $\mathbb{Z}$ (so the subgroup defining $\widehat{Y}$ is the preimage of the torsion subgroup of $H_1(Y)$ under the abelianisation map). Secondly, for a more convenient definition of the Alexander polynomial later we use the ring $\mathbb{Q}[t,t^{-1}]$ instead of $\mathbb{Z}[t,t^{-1}]$.

With these changes, the method for obtaining the presentation matrix for $H_1(\widehat{Y})$ as a $\mathbb{Q}[t,t^{-1}]$–module using Fox calculus works in the same way as for knots in homology spheres — see [14, Chapter 11].

Now the Alexander polynomial Δ_K is defined to be a specific generator of the ideal generated by the maximal size minors of the presentation matrix of the Alexander module. The specific generator is fixed by the requirement that $\Delta_K(t) = \Delta_K(t^{-1})$ and $\Delta_K(1) = |H_1(Y)|$.

Suppose we have a genus g doubly pointed Heegaard diagram for $K \subset Y$. To get a Heegaard diagram of the knot exterior we add one more α–curve, α_{g+1}. As described in [33, Section 3], this leads to a presentation of $\pi_1(Y)$ in which there is one generator for each α–curve and one relator for each β–curve. Denote the generators by $\{a_i\}_{i=1}^{g+1}$ and relators (words in a_i) by $\{w_j\}_{j=1}^{g}$. Denote the free differential with respect to a_i by d_{a_i} (this time with respect to the map ϕ, not abelianisation).

Define

$$\widehat{HFK}(Y,K,n)=\bigoplus_{h\in H_1(Y)}\widehat{HFK}(Y,K,(n,h)).$$

Then, as in [33, Section 3], we can see that

$$\chi(\widehat{HFK}(Y,K))=\sum_{i,j}(-1)^i t^j \dim \widehat{HFK}_i(Y,K,j)=\det(d_{a_i}w_j)_{1\le i,j\le g}.$$

Following [34, Proposition 3.1], we see that, in fact,

$$\chi(\widehat{HFK}(Y,K))=\Delta_K(t).$$

For every $h\in H_1(Y)$ define[7]

$$\Delta_{K,h}(t)=\sum_{i,j}(-1)^i t^j \dim \widehat{HFK}_i(Y,K,(j,h)).$$

We have $\Delta_K(t)=\sum_h \Delta_{K,h}(t)$ and $\Delta_{K,h}(1)=1$.

Let Y be an L–space and K a null-homologous knot in it with L–space surgery. Recall that $\widehat{A}_{(n,h)}\cong\mathbb{F}$ for all n and h. Using the same algebraic manipulations as in [27, Section 3] we deduce that, for each fixed h, $\widehat{HFK}(Y,K,(n,h))$ has dimension 0 or 1, successive copies of $\mathbb{F}$ are concentrated in different $\mathbb{Z}_2$ gradings and the first (and the last) copies of $\mathbb{F}$ are concentrated in grading 0.

It follows that, for each h, $\Delta''_{K,h}(1)\ge 0$ and equality is only possible if $\Delta_{K,h}(t)=1$ or t.

8.2 Knots determined by their complements

We are now ready to prove the surgery characterisation of the unknot for null-homologous knots in rational homology L–spaces.

Theorem 8.2 *Let Y be an L–space and $K\subset Y$ a null-homologous knot. Suppose that*

$$HF^+(Y_{p/q}(K))\cong HF^+(Y\# L(p,q)).$$

Then K is the unknot.

In particular, null-homologous knots in L–spaces are determined by their complements.

[7] This definition works only up to an affine identification, since we should really have $h\in \mathrm{Spin}^c(Y)$.

Proof The Casson–Walker invariant is additive under connected sums and, by (5), determined by Heegaard Floer homology, thus we have

$$\lambda(Y)+\lambda(L(p,q))=\lambda(Y_{p/q}(K))=\lambda(Y)+\lambda(L(p,q))+\frac{q}{2p|H_1(Y)|}\Delta_K''(1).$$

It follows that $\Delta_K''(1)=0$. Thus for each h we must have $\Delta_{K,h}(t)=1$ or t. However, by symmetry if there is a multiple of t in $\Delta_K(t)$ there must also be a multiple of t^{-1}. Hence $\Delta_{K,h}(t)=1$ for all h. Note that it also means that

$$\widehat{HFK}(Y,K,(n,h))=0$$

for all $n\neq 0$. Hence, by [19, Theorem 2.2], $g(K)=0$, ie K is the unknot. □

A straightforward homological argument provides more restrictions on knots not being determined by their complements in lens spaces. In particular, all knots in lens spaces $L(p,q)$ with p square-free satisfy Conjecture 1.2. For notation used in the statement below, see Section 1 or the proof. By a nontrivial surgery we mean a surgery with a slope that is not the meridian, so even if a slope is equivalent to the meridian, surgery with this slope is still nontrivial.

It is clear that cores of Heegaard solid tori of $L(p,q)$ admit nontrivial surgery which give back $L(p,q)$. Hence, in the corollary below we assume that K is not a core of one of the Heegaard solid tori.

Corollary 8.3 *If p is square-free, then all knots in $L=L(p,q)$ are determined by their complements.*

More precisely, let K be a knot whose exterior is not a solid torus and such that a nontrivial surgery on it gives L. Then the exterior of K is not Seifert fibred, $p\,|\,w^2$ and the surgery slope, n, is an integer that satisfies the following (with some choice of sign):

$$n=-q\frac{w^2}{p}\pm 1.$$

Moreover, there is at most one such slope (ie we can choose either $+$ or $-$ but not both in the equation above).

Proof By Theorem 8.2 we only need to consider non-null-homologous knots. Let $L=L(p,q)$ be a lens space and K a non-null-homologous knot in it.

Suppose the exterior of K is not Seifert fibred. Then, by the cyclic surgery theorem [4], the slope has to be integral.

We can isotope K into one of the Heegaard solid tori W of L. Then we can get L by first performing an integral surgery on K in W and then glueing the other solid torus from the outside so that its meridian becomes the (p,q)–curve.

Let μ be the meridian of W, fix a curve λ in ∂W with $\lambda \cdot \mu = 1$ and embed W into S^3 in the standard way with respect to μ and λ. This endows K with a well-defined longitude l. Let m be the meridian of K. Suppose K has winding number w in W. Then, in $H_1(W \setminus K)$ (which is generated by m and λ), $l = w\lambda$ and $\mu = wm$. Let n be the surgery slope with respect to these coordinates. Then surgery on K introduces a relation $nm + w\lambda = 0$. The other Heegaard solid torus introduces a relation $-qwm + p\lambda$. All in all, the first homology of L has presentation matrix

$$\begin{pmatrix} n & w \\ -qw & p \end{pmatrix}.$$

The order of the first homology of L is the absolute value of the determinant of the relation matrix. Thus we must have $\pm p = np + qw^2$, and thus $n = -qw^2/p \pm 1$. However, q is coprime with p, so $p \mid w^2$. By the cyclic surgery theorem, the distance between slopes that give lens spaces is at most 1, so $-qw^2/p + 1$ and $-qw^2/p - 1$ cannot both produce a lens space.

If p is square-free, $p \mid w^2$ implies that $p \mid w$, so K is null-homologous in L — a contradiction.

Now suppose the exterior of K is Seifert fibred. As stated, Conjecture 1.2 has been proven for knots with Seifert fibred exteriors in [36, Theorem 1].

However, to demonstrate that there are no nontrivial slopes equivalent to the meridian we will provide a different proof.

By [36, Lemma 2], we can assume that K is a fibre in some fibration of L.

Fibrations of lens spaces come in two families (see eg [9, Theorem 2.3]). All lens spaces can be fibred over a sphere with at most two exceptional fibres. There are also some lens spaces that can be fibred over the projective plane with one exceptional fibre of invariant $(n,1)$.

In the first (and most common) case, to avoid the exterior being a solid torus, K must be an ordinary fibre in a fibration with two exceptional fibres. Suppose the invariants of these fibres are (p_1, x_1) and (p_2, x_2). Since Seifert fibred spaces with three exceptional fibres are not lens spaces, surgery on K must introduce a fibre with invariant $(1,n)$. If

so, the order of homology changes from $|p_1x_2 + p_2x_2|$ to $|p_1x_2 + p_2x_2 + np_1p_2|$. Equality is only possible if $p_1 = p_2 = 2$, ie L is obtained by filling the Seifert fibred space over a disc with two exceptional fibres with the same invariant $(2, 1)$.

Different fillings of this space that produce lens spaces can be indexed by integers $m \in \mathbb{Z}$ and they produce $L(4m, 2m + 1)$. If $L(4m, 2m + 1) = L(4n, 2n + 1)$ for $m \neq n$ (the equality sign here means "there exists an orientation-preserving homeomorphism"), then $m = -n$, ie the two lens spaces are $L(4m, 2m \pm 1)$. If they are to be homeomorphic by an orientation-preserving homeomorphism, then we must have $(2m + 1)(2m - 1) \equiv 1 \pmod{4m}$ (see [32] and references therein), which is not true (even though this spaces *are* homeomorphic by an orientation-reversing homeomorphism).

This deals with the case when K is a fibre of a fibration of L over a sphere. L may also have a fibration with one exceptional fibre over the projective plane, in which case the invariant of the exceptional fibre is $(n, 1)$. In this case, $L = L(4n, 2n + 1)$. If K is the exceptional fibre, then the only nontrivial surgery which still gives a lens space yields $L(4m, 2m + 1)$ for $n \neq m$. This case has been dealt with above. Similarly, if K is an ordinary fibre then surgeries on it are indexed by an integer k and their effect is to change the invariant of the exceptional fibre to $(n, nk + 1)$. It follows that we must have $nk + 1 = -1$, so surgery again gives $L(4n, 2n - 1)$. □

References

[1] **S Boyer**, *Dehn surgery on knots*, from "Handbook of geometric topology" (R J Daverman, R B Sher, editors), North-Holland, Amsterdam (2002) 165–218 MR

[2] **S Boyer**, **D Lines**, *Surgery formulae for Casson's invariant and extensions to homology lens spaces*, J. Reine Angew. Math. 405 (1990) 181–220 MR

[3] **D Cooper**, **M Lackenby**, *Dehn surgery and negatively curved 3–manifolds*, J. Differential Geom. 50 (1998) 591–624 MR

[4] **M Culler**, **C M Gordon**, **J Luecke**, **P B Shalen**, *Dehn surgery on knots*, Ann. of Math. 125 (1987) 237–300 MR

[5] **C H Edwards, Jr**, *Concentricity in 3–manifolds*, Trans. Amer. Math. Soc. 113 (1964) 406–423 MR

[6] **F Gainullin**, *The mapping cone formula in Heegaard Floer homology and Dehn surgery on knots in S^3*, Algebr. Geom. Topol. 17 (2017) 1917–1951 MR

[7] **C M Gordon**, *Dehn surgery on knots*, from "Proceedings of the International Congress of Mathematicians, I" (I Satake, editor), Math. Soc. Japan, Tokyo (1991) 631–642 MR

[8] **C M Gordon**, **J Luecke**, *Knots are determined by their complements*, J. Amer. Math. Soc. 2 (1989) 371–415 MR

[9] **A E Hatcher**, *Notes on basic* 3*–manifold topology*, book project (2007) Available at `http://www.math.cornell.edu/~hatcher/3M/3Mdownloads.html`

[10] **J Hom**, **T Lidman**, **N Zufelt**, *Reducible surgeries and Heegaard Floer homology*, Math. Res. Lett. 22 (2015) 763–788 MR

[11] **A Juhász**, *Floer homology and surface decompositions*, Geom. Topol. 12 (2008) 299–350 MR

[12] **R C Kirby, editor**, *Problems in low-dimensional topology*, from "Geometric topology, II" (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc., Providence, RI (1997) 35–473 MR

[13] **P Kronheimer**, **T Mrowka**, **P Ozsváth**, **Z Szabó**, *Monopoles and lens space surgeries*, Ann. of Math. 165 (2007) 457–546 MR

[14] **W B R Lickorish**, *An introduction to knot theory*, Graduate Texts in Mathematics 175, Springer (1997) MR

[15] **A Manion**, *Heegaard Floer homology and knots*, Part III essay, University of Cambridge (2010) Available at `http://tinyurl.com/ManionIII`

[16] **M Mauricio**, *Distance bounding and Heegaard Floer homology methods in reducible Dehn surgery*, PhD thesis, Imperial College London (2012) Available at `http://hdl.handle.net/10044/1/10226`

[17] **Y Ni**, *Knot Floer homology detects fibred knots*, Invent. Math. 170 (2007) 577–608 MR

[18] **Y Ni**, *Link Floer homology detects the Thurston norm*, Geom. Topol. 13 (2009) 2991–3019 MR

[19] **Y Ni**, **Z Wu**, *Heegaard Floer correction terms and rational genus bounds*, Adv. Math. 267 (2014) 360–380 MR

[20] **Y Ni**, **Z Wu**, *Cosmetic surgeries on knots in* S^3, J. Reine Angew. Math. 706 (2015) 1–17 MR

[21] **Y Ni**, **X Zhang**, *Characterizing slopes for torus knots*, Algebr. Geom. Topol. 14 (2014) 1249–1274 MR

[22] **P Ozsváth**, **Z Szabó**, *Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary*, Adv. Math. 173 (2003) 179–261 MR

[23] **P Ozsváth**, **Z Szabó**, *On the Floer homology of plumbed three-manifolds*, Geom. Topol. 7 (2003) 185–224 MR

[24] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and knot invariants*, Adv. Math. 186 (2004) 58–116 MR

[25] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and three-manifold invariants: properties and applications*, Ann. of Math. 159 (2004) 1159–1245 MR

[26] **P Ozsváth**, **Z Szabó**, *Knot Floer homology, genus bounds, and mutation*, Topology Appl. 141 (2004) 59–85 MR

[27] **P Ozsváth**, **Z Szabó**, *On knot Floer homology and lens space surgeries*, Topology 44 (2005) 1281–1300 MR

[28] **P Ozsváth**, **Z Szabó**, *Holomorphic triangles and invariants for smooth four-manifolds*, Adv. Math. 202 (2006) 326–400 MR

[29] **P Ozsváth**, **Z Szabó**, *Lectures on Heegaard Floer homology*, from "Floer homology, gauge theory, and low-dimensional topology" (D A Ellwood, P S Ozsváth, A I Stipsicz, Z Szabó, editors), Clay Math. Proc. 5, Amer. Math. Soc., Providence, RI (2006) 29–70 MR

[30] **P S Ozsváth**, **Z Szabó**, *Knot Floer homology and integer surgeries*, Algebr. Geom. Topol. 8 (2008) 101–153 MR

[31] **P S Ozsváth**, **Z Szabó**, *Knot Floer homology and rational surgeries*, Algebr. Geom. Topol. 11 (2011) 1–68 MR

[32] **J H Przytycki**, **A Yasukhara**, *Symmetry of links and classification of lens spaces*, Geom. Dedicata 98 (2003) 57–61 MR

[33] **J A Rasmussen**, *Floer homology and knot complements*, PhD thesis, Harvard University (2003) MR Available at `https://search.proquest.com/docview/305332635`

[34] **J Rasmussen**, *Lens space surgeries and L–space homology spheres*, preprint (2007) arXiv

[35] **H C Ravelomanana**, *Knot complement problem for L–space $\mathbb{Z}HS^3$*, preprint (2015) arXiv

[36] **Y W Rong**, *Some knots not determined by their complements*, from "Quantum topology" (L H Kauffman, R A Baadhio, editors), Ser. Knots Everything 3, World Sci., River Edge, NJ (1993) 339–353 MR

[37] **R Rustamov**, *Surgery formula for the renormalized Euler characteristic of Heegaard Floer homology*, preprint (2004) arXiv

[38] **K Walker**, *An extension of Casson's invariant*, Annals of Mathematics Studies 126, Princeton University Press (1992) MR

Department of Mathematics, Imperial College London
London, United Kingdom

`fyodor.gainullin@gmail.com`

Received: 13 October 2015 Revised: 7 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 111–152

Classification of tight contact structures on small Seifert fibered L–spaces

IRENA MATKOVIČ

We identify tight contact structures on small Seifert fibered L–spaces as exactly the structures having nonvanishing contact invariant, and classify them by their induced Spin^c structures. The result (in the new case of $M(-1;r_1,r_2,r_3)$) is based on the translation between convex surface theory and the tightness criterion of Lisca and Stipsicz.

57R17

1 Introduction

By small Seifert fibered 3–manifold we mean a Seifert fibration over the sphere S^2 with three singular fibers, standardly given as $M(e_0;r_1,r_2,r_3)$, where $e_0 \in \mathbb{Z}$ and $r_i \in \mathbb{Q} \cap (0,1)$ with $r_1 \geq r_2 \geq r_3$. For a surgery presentation of this manifold, see the bottom diagram of Figure 1.

L–spaces (by definition, Heegaard Floer homology lens spaces) among Seifert fibered manifolds are geometrically characterized by absence of transverse contact structures — see Lisca and Stipsicz [11, Theorem 1.1] — which is essential for our classification. The restriction can be simply described in terms of the Seifert constants: L–spaces are all manifolds with $e_0 \geq 0$ and with $e_0 \leq -3$, while for $e_0 = -1, -2$ some explicit numerical inequalities (see Section 4.4) are imposed on the triple (r_1,r_2,r_3).

Problem to classify tight contact structures up to contact isotopy is usually asked for prime atoroidal manifolds; the first because tight contact structures respect connected sum decomposition of 3–manifolds, the second because an embedded essential torus is a known source of infinitely many nonisotopic tight structures. Small Seifert fibered manifolds, beside hyperbolic ones, share these properties. On the other hand, the existence question for Seifert manifolds has been completely answered by Lisca and Stipsicz [13]: only the ones which belong to a one-parameter family of $(2n-1)$–surgeries on the torus knot $T_{2,2n+1}$ — equivalently, which are orientation-preserving diffeomorphic to $M\left(-1;\frac{1}{2},\frac{n}{2n+1},\frac{1}{2n+3}\right)$ for some $n \in \mathbb{N}$ — do not admit any tight

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.111

structure. Classification then arises from the comparison of bounds: the lower bound is obtained constructively by contact surgery complemented with the use of invariants, and for the upper bound the convex surface theory is applied.

The main invariant in the classification of tight Seifert fibered manifolds is the maximal twisting number (that is, the difference between the contact framing and the fibration framing, maximized in the smooth isotopy class of a regular fiber) — applied in convex surface theory, it allows one to give upper bounds on the number of tight structures. By the results of Wu [20] all tight contact structures when $e_0 \leq -2$ have negative maximal twisting, while for $e_0 \geq 0$ they are all zero-twisting; in the work of Ghiggini [3] the negative maximal twisting is further related to the existence of transverse contact structures. This, in the case of L–spaces, results in a simple division: maximal twisting is equal to zero when $e_0 \geq -1$, and has value -1 when $e_0 \leq -2$. The fixed maximal twisting of a regular fiber in all the cases gives some unique contact structure on the complement of singular fibers relative to boundary, pushing the classification into tubular neighborhoods of the three singular fibers.

The classification whenever $e_0 \neq -1$ is then finished by Legendrian surgery construction — the diagrams are simply given by Legendrianization of the standard presentation of a Seifert manifold; this has been done by Wu [20] for $e_0 \neq -2, -1, 0$, by Ghiggini, Lisca and Stipsicz [4] for $e_0 \geq 0$, and by Ghiggini [3] for L–spaces with $e_0 = -2$. In particular, all these tight structures are Stein fillable, and are classified by the first Chern class of their fillings (according to Lisca and Matić [10]), equivalently by their Spin^c structure, or closer to the present context by their contact Ozsváth–Szabó invariants (according to Plamenevskaya [17]).

The remaining case of $M(-1; r_1, r_2, r_3)$ has been partly addressed already by Ghiggini, Lisca and Stipsicz [5]. Here, the constructive side needs to be attacked differently because of the existence of nonfillable tight structures. Invoking that all tight structures are zero-twisting [3], Lisca and Stipsicz gave a uniform description of all possible tight structures by certain surgery diagrams.

Proposition 1.1 [11, Proposition 6.1] *Each tight contact structure with maximal twisting equal to zero on the small Seifert fibered space $M(-1; r_1, r_2, r_3)$ is given by one of the surgery presentations of Figure 1, top left.* □

This reduces the classification problem to the recognition of tightness and isotopies between the finite collection of structures, listed by the associated Thurston–Bennequin and rotation numbers.

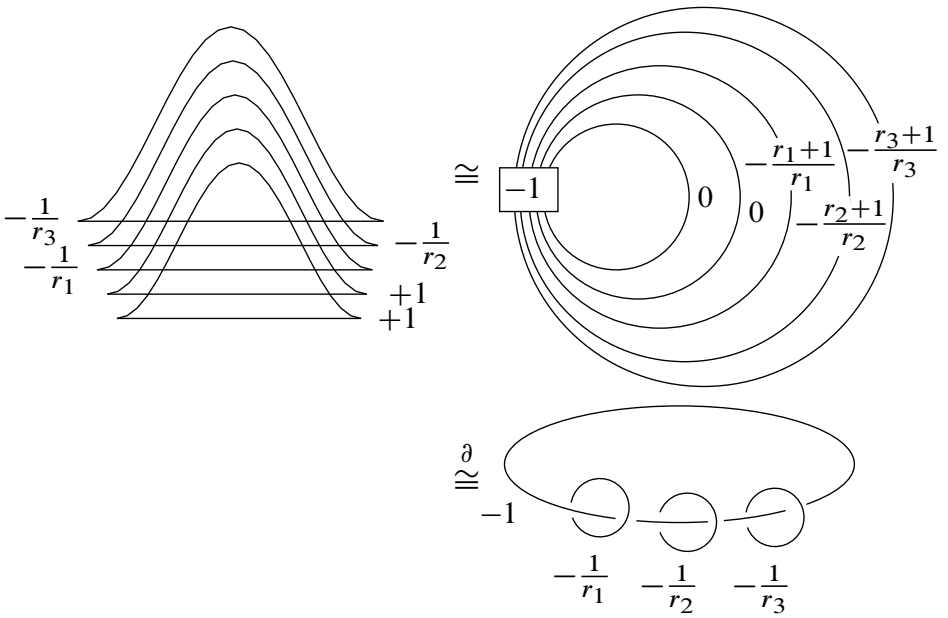


Figure 1: Contact structures on $M(-1;r_1,r_2,r_3)$, followed by the smoothened surgery diagram of the underlying 3–manifold and its standard presentation; when referring to them as 4–manifolds, we assume inverse slam-dunks to be done.

The underlying topological question, classification of oriented 2–plane fields $\xi \in \Xi$ up to homotopy, is given by their induced Spin^c structure $\boldsymbol{t}_\xi$ together with the 3–dimensional invariant $d_3(\xi)$. (Recall that $\pi_0(\Xi)$ can be identified with homotopy classes of maps $[M, S^2]$, which can, through Pontryagin–Thom construction, be given by framed links in M; here a link up to oriented cobordism represents the class in $H_1(M;\mathbb{Z})$, equivalently $\boldsymbol{t}_\xi$, while the framing corresponds to the Hopf invariant as a 3–dimensional obstruction for homotopies between plane fields.)

To detect tightness, we basically use the Ozsváth–Szabó contact invariant [16], implicitly expecting all tight structures to have nonvanishing one. But we address it indirectly, by showing the sufficient condition of Lisca and Stipsicz.

Theorem 1.2 [11, Theorem 1.2] *If for a contact structure ξ of Figure 1 on Seifert fibered L–space $M = M(-1;r_1,r_2,r_3)$ the equality $d_3(\xi) = d(M,\boldsymbol{t}_\xi)$ holds, then its contact invariant $c(M,\xi) \in \widehat{\mathrm{HF}}(-M,\boldsymbol{t}_\xi)$ does not vanish.* □

Then, to confirm the overtwistedness of all undetected structures and to obtain the isotopies between tight ones, as always, convex surface theory is applied.

The observations accumulate in the confirmation of Conjecture 4.7 of Stipsicz [18].

Theorem 1.3 *Let M be a small Seifert fibered L–space of the form $M(-1; r_1, r_2, r_3)$. Then a contact structure ξ on M is tight if and only if it is given by a contact surgery presentation of Figure 1 and its 3–dimensional invariant $d_3(\xi)$ is equal to the d–invariant $d(M, \boldsymbol{t}_\xi)$. Moreover, two tight structures ξ_1 and ξ_2 on M are contact isotopic if and only if their induced* Spin^c *structures $\boldsymbol{t}_{\xi_1}$ and $\boldsymbol{t}_{\xi_2}$ are isomorphic.*

Expressed in terms of the Ozsváth–Szabó contact invariant, all tight structures on small Seifert fibered L–spaces satisfy the following:

Corollary 1.4 *Let ξ be a contact structure on a small Seifert fibered L–space $M = M(e_0; r_1, r_2, r_3)$. Then ξ is tight if and only if its contact invariant $c(\xi) \in \widehat{\mathrm{HF}}(-M, \boldsymbol{t}_\xi)$ is nonzero. Moreover, two tight structures ξ_1 and ξ_2 are isotopic if and only if their contact invariants $c(\xi_1)$ and $c(\xi_2)$ coincide, if and only if their induced* Spin^c *structures $\boldsymbol{t}_{\xi_1}$ and $\boldsymbol{t}_{\xi_2}$ are isomorphic.*

Proof If there are less than three singular fibers, the manifold considered is a lens space. Here as well as when $e_0 \neq -1$, all tight structures are Stein fillable according to previous results: by Eliashberg [2] for the 3–sphere, by Honda [7] for the lens spaces, and due to Wu [20], Ghiggini, Lisca and Stipsicz [4] and Ghiggini [3] in the small Seifert fibered cases. By that and Theorem 1.3 above, a tight structure on any considered M has nonvanishing contact invariant.

The fillable structures are classified by the contact invariant due to Plamenevskaya [17]. In fact, for L–spaces the nontrivial contact invariant of ξ is the unique generator of $\widehat{\mathrm{HF}}(-M, \boldsymbol{t}_\xi)$, hence ξ is the only tight representative of its induced Spin^c structure. Its 3–dimensional invariant $d_3(\xi)$ is specified as the absolute grading of the contact invariant, which equals $d(M, \boldsymbol{t}_\xi)$. □

Our result reduces the classification problem to a well-understood computation of invariants. Although our method does not result in the number of tight structures on a given small Seifert manifold, the problem is translated to a completely combinatorial (so not geometric) count. Indeed, in any special case the number can be easily determined by, say, a computer calculation (as here both d_3 and d are computable, and the Spin^c structure can be given as an element of H_1). What is more, since there is a surgery presentation of considered contact manifolds, we have a very explicit description of tight structures.

Remark 1.5 In contrast to the cases with $e_0 \neq -1$, not all tight structures on $M(-1; r_1, r_2, r_3)$ are fillable. Whenever $r_1 + r_2 < 1$, in the language of Lecuona and Lisca [9] for manifolds of special type, the existence of Stein filling is even topologically obstructed. And as all contact structures of the form given by Figure 1 are known to be supported by open books with planar pages Lisca and Stipsicz [11], the theorem of Wendl [19] implies they are not fillable at all. Most manifolds with $r_1 + r_2 \geq 1$ admit Stein fillable as well as nonfillable tight structures, as specified in Matkovič [14].

Overview

In Section 2 we explain the structure of our proof of Theorem 1.3, and review main concepts behind it. Then in Section 3 we illustrate the suggested approach by reproving the classification on $M_p := M\left(-1; \frac{1}{2}, \frac{1}{2}, \frac{1}{p}\right)$; see Ghiggini, Lisca and Stipsicz [5]. Technical details are given in the last two sections. In Section 4 we establish paths of characteristic covectors separating the presentations into classes with the same contact invariant. Finally, in Section 5, with the help of convex surface theory, the presentations of the same class are realized to be contact isotopic, more, the failure of the tightness criterion is related to overtwistedness.

Acknowledgement I am indebted to András Stipsicz for his insightful mentoring.

2 Outline of the proof

Following the classification scheme given in the Introduction, we need a construction, a method to detect tightness (Section 2.1), and finally a proof that it is complete (that is, a way to recognize overtwistedness and isotopies between possibly different presentations of the same contact structure; Section 2.2).

To construct tight structures on L–spaces of the form $M(-1; r_1, r_2, r_3)$, by Proposition 1.1 the contact surgery presentations of Figure 1 suffice. This gives a finite collection of contact structures, on which we need to run the following two-step analysis.

2.1 Detect tightness

In order to detect tightness we examine the equality between the 3–dimensional homotopy invariant d_3 of the contact structure and the d–invariant of the induced

Spin^c structure (according to Theorem 1.2). The advantage of this condition over the Ozsváth–Szabó contact invariant is that these two invariants are easily computable. They can be described in terms of characteristic covectors on plumbings bounded by $\pm M$, which brings them into the same picture (as presented below).

We exploit two 4–manifolds, naturally arising as fillings of our M, one given by the smoothened surgery diagram (Figure 1, top right) — call it X — and another given by standard smooth presentation of Seifert fibration (Figure 1, bottom) — say $W = W_\Gamma$ as we can think of it as a simple plumbing along the graph Γ. They are related by $X \# \mathbb{C}P^2 \cong W \# 2\mathbb{C}P^2$. Further, the plumbing description of $-M$ will play the central role in what follows. Call it $W_{\Gamma'}$, where Γ' stands for the plumbing graph dual to Γ.

Now, the 3–dimensional invariant of the contact bundle ξ can be directly read off from the surgery presentation $d_3(\xi) = \frac{1}{4}(c^2(X,J) - 3\sigma(X) - 2b_2(X)) + \#(+1\text{–surgeries})$ [1, Corollary 3.6], where $c(X,J)$ stands for the characteristic element, determined by the ξ–induced almost complex structure J of $X \backslash \{\text{a point in each } +1\text{–handle}\}$. While the d–invariant corresponds to the reversely oriented $-M$ (which bounds a negative-definite plumbing) together with an induced Spin^c structure $\boldsymbol{t}_\xi$. It is realized by the characteristic 2–cohomology class, which gives a Spin^c cobordism from S^3 to $(-M, \boldsymbol{t}_\xi)$ whose associated map in Heegaard Floer homology decreases the absolute grading the least. These are recognized by full paths [15, Section 3.1]. To establish some terminology let us recollect:

Full path Assume that Γ' is a negative-definite plumbing with at most one bad vertex. A t–tuple $(K_1, \dots, K_t)$ of characteristic covectors on $W_{\Gamma'}$ forms a full path if its elements are connected by the following 2 PD steps: for some vertex v with $\langle K_i, v\rangle = -v \cdot v$, the vector K_{i+1} is given by $K_{i+1} = K_i + 2\,\mathrm{PD}(v)$. (In particular, all its elements induce the same Spin^c structure on the boundary $\partial W_{\Gamma'} = -M$ as they differ only by twice generators of $H^2(W_{\Gamma'}, -M; \mathbb{Z})$. Further, their (common) degree can be computed by the formula $\frac{1}{4}(K_i^2 + |\Gamma'|)$.) The path either ends at some characteristic vector K which exceeds the bounds $v \cdot v \le \langle K_i, v\rangle \le -v \cdot v$ at some $v \in \Gamma'$ — we will say that it drops out — or the path reaches the proper ends in the initial vector K_1 satisfying $v \cdot v + 2 \le \langle K_1, v\rangle \le -v \cdot v$ for all $v \in \Gamma'$ and the terminal vector K_t satisfying $v \cdot v \le \langle K_t, v\rangle \le -v \cdot v - 2$ for all $v \in \Gamma'$ — such an ending full path according to [15] determines a nontrivial element of $\widehat{\mathrm{HF}}(-M)$.

Embedding into blown-up $\mathbb{C}P^2$ According to [13, Lemma 4.2] we can embed M as a hypersurface in a closed oriented 4–manifold R such that $R \backslash \nu(M) = W_\Gamma \cup W_{\Gamma'}$.

The configuration of both intersection graphs Γ and Γ' is obtained by blowing up the initial lines $l_1, l_2, l_3 \subset \mathbb{C}P^2$ with $l_1 \cap l_2 \cap l_3 = \{p\}$, and $l \subset \mathbb{C}P^2$ with $p \notin l$ (see Figure 2).

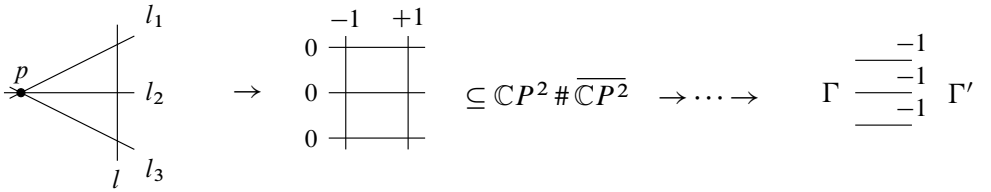


Figure 2: Construction of the 4–manifold $R = \mathbb{C}P^2 \# n\overline{\mathbb{C}P^2}$; at the end, the two plumbings are glued together along the exceptional spheres from the last blow-up of each singular fiber, and (not shown) all regular fibers.

Denote standard generators of $H_2(R;\mathbb{Z})$ as follows: h ($h^2 = 1$) for the initial $\mathbb{C}P^1 \subset \mathbb{C}P^2$ and e_i ($e_i^2 = -1$) for exceptional curves. Then the above description of the embedding $W_\Gamma \cup W_{\Gamma'} \hookrightarrow R$ gives:

- $\{z = \text{center of } \Gamma\} \mapsto e_1$;
- $\{z' = \text{center of } \Gamma'\} \mapsto h - e_2 - e_3 - e_4$;
- $\{x_i = \text{first vertex of the leg } L_i \subset \Gamma\} \mapsto h - e_1 - e_{i+1} - \sum e_j$ for $i = 1, 2, 3$;
- $\{v \text{ vertex}, v \neq z, z', x_i\} \mapsto e_j - \sum e_k$, for example

$$\{x_i' = \text{first vertex of the dual leg } L_i' \subset \Gamma'\} \mapsto e_{i+1} - \sum e_k \quad \text{for } i = 1, 2, 3.$$

We will refer to Γ as the manifold side and Γ' as the dual side. Throughout, we will follow the convention that primed notation belongs to the dual graph: apart from the special vertices denoted above, let v_j^i be the jth vertex of L_i and $v_j^{i'}$ the jth vertex of L_i'.

Tightness criterion We have d_3 given by a characteristic covector on W_Γ, d given by a characteristic covector on $W_{\Gamma'}$ and, what is more, we can glue these two plumbings along a rational homology sphere M, giving a blown-up $\mathbb{C}P^2$ (called R). Having a characteristic covector c on R which agrees with the ones providing d_3 and d on W_Γ and $W_{\Gamma'}$, respectively, the equality $d_3(\xi) = d(M; \boldsymbol{t}_\xi)$ can be rewritten as $c^2 = \sigma(R)$. This can be understood as a geometrization of Theorem 1.2.

Theorem 2.1 [13, Theorem 3.3] *Let $M = M(-1; r_1, r_2, r_3)$ be an L–space. Then a sufficient condition for a contact structure ξ on M to be tight (even for $c(M, \xi) \neq 0$) is in the existence of a characteristic cohomology class $c \in H^2(R;\mathbb{Z})$ such that:*

(i) $d_3(\xi) = \frac{1}{4}((c|_{W_\Gamma})^2 - 3\sigma(W_\Gamma) - 2b_2(W_\Gamma)) + 1$, *that is,* $c|_{W_\Gamma}$ *corresponds to* $c(X, J)$.

(ii) $c|_{W_{\Gamma'}}$ *belongs to an ending full path on* $W_{\Gamma'}$, *that is, gives rise to a class of* $\widehat{\mathrm{HF}}(-M, \boldsymbol{t}_\xi)$.

(iii) $c^2 = \sigma(R)$. □

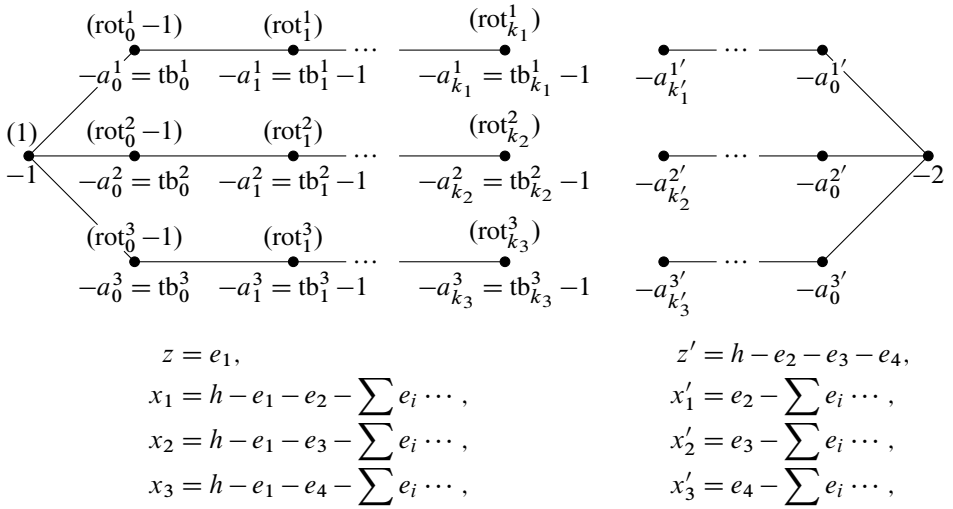


Figure 3: Plumbing graph Γ (left) and its dual Γ' (right) with denoted self-intersections and evaluations of characteristic covector c, $(\cdot) = \langle c, v\rangle$, on the manifold side; the central and the first vertices on legs are given in generating classes of $H_2(R; \mathbb{Z})$.

For a given surgery presentation, the conditions read as follows (see Figure 3):

(i) c on W_Γ is determined by $c(X, J)$, on generators of $H_2(X)$ evaluated as rotation numbers (central blow-up decreases all the (neighboring) values by 1).

(ii) This is to be checked given the constraints from (i).

(iii) We can give c as $\mathrm{PD}(c) = \alpha h + \sum \alpha_i e_i$, where $\alpha, \alpha_i \in \{\pm 1\}$, by construction fulfilling the equality $c^2 = \sigma(R)$.

Technically speaking, given a tuple $(c_v)_{v\in\Gamma}$ of c–evaluations $(c_v) = \langle c, v\rangle$ as in (i), we list all possible ± 1 distributions in the expression $\mathrm{PD}(c)$ of (iii), and calculate corresponding values on Γ'. We know that these different Γ'–evaluations are, for each given surgery presentation, only different representatives of the same connected

component of full paths. We continue these paths towards their ends, so that we connect to them all the characteristic covectors which can be obtained by allowed 2 PD steps. Taking all (c_v)–tuples, this results in the separation of the characteristic lattice into components; denote them by $\mathcal{P}_\xi$ according to the contact structure ξ to which they belong. By abuse of notation, we call $\mathcal{P}_\xi$ a full path; we say that $\mathcal{P}_\xi$ drops out or properly ends if any (hence, every) path in the component has this property.

In this language, Theorem 1.3 takes the following working form:

Theorem 2.2 *The contact structure ξ on $M(-1; r_1, r_2, r_3)$ given by surgery diagram is tight if and only if its full path $\mathcal{P}_\xi$ properly ends in the initial and terminal vector. Two such contact structures ξ_1 and ξ_2 are isotopic if and only if their paths $\mathcal{P}_{\xi_1}$ and $\mathcal{P}_{\xi_2}$ meet (hence, coincide).*

2.2 Prove overtwistedness and describe contact isotopies

Finally, to close our classification we need that the zero elements (drop-outs) correspond to overtwistedness, and for the second part of Theorem 1.3 (Theorem 2.2) that elements giving the same $\widehat{\mathrm{HF}}(M)$–generator (sharing the same path) are actually contact isotopic.

Here, convex surface theory comes in. We need to translate contact surgeries back into convex decomposition. Natural convex decomposition of the manifold M separates the three singular tori from the rest of the manifold. Then the coefficients in the continued fraction expansions of the three surgeries, along with the chosen stabilizations, determine basic slice decompositions of the three tori.

Contact surgery Contact surgery in addition to ordinary surgery prescribes for the contact structure to be preserved in the complement of a tubular neighborhood of the core link, while the extension to glued-up tori needs to be tight. The extended contact structure is determined by the boundary slope [7, Theorem 2.3], given by the surgery coefficient. Contact surgery diagrams [1] encode a basic slice decomposition of the glued-up solid torus. The slope uniquely determines continued fraction blocks. Concretely, writing out the continued fraction expansion of the surgery coefficient

$$-\frac{1}{r_i} = -a_0^i - \cfrac{1}{\ddots - \cfrac{1}{-a_{k_i}^i}} = [a_0^i, \dots, a_{k_i}^i]$$

(we use the convention of [5] with $a_j^i \geq 2$), continued fraction blocks are toric annuli in the layering of the solid torus with boundary slope $[a_{k_i}^i, \dots, a_0^i]$, cut out by pairs of tori of slopes $[a_{k_i}^i, \dots, a_{j+1}^i - 1]$ and $[a_{k_i}^i, \dots, a_j^i - 1]$ (the outermost being $[a_{k_i}^i, \dots, a_0^i]$,

and the innermost -1). In a surgery diagram, they are represented in a chain of pushed-off knots with appropriate integral surgery coefficients (after turning into ± 1 surgeries, captured by Thurston–Bennequin invariants). The remaining ambiguity in the signs of the basic slices within each continued fraction block is reflected in the choice of stabilizations of the corresponding Legendrian knot (equivalently their differences as their total number is determined by the surgery coefficient), so given by its rotation number. In the translation, positive and negative basic slices in the decomposition of a continued fraction block correspond to positive and negative stabilizations (down- and up-cusps) of corresponding Legendrian knot. The loss of basic slice ordering in the transition is explained by the shuffling property [7, Lemma 4.14] of basic slices within a single block.

What we need is to relate steps in the full path with appropriate state traversals, and drop-outs to nontight basic slice configurations.

In other words, we have set up two ways to describe rotation numbers. Provided the surgery coefficient is fixed, they can be equally given by either the number of negative basic slices (up-cusps) or the number of negative signs on the generators forming the corresponding part of the dual leg. Then the nice thing — to be shown — is that the full path connections reflect the known behavior of basic slices.

3 First example

We illustrate our strategy on small Seifert fibered L–spaces $M_p := M\left(-1; \frac{1}{2}, \frac{1}{2}, \frac{1}{p}\right)$. The classification on these manifolds was first obtained by Ghiggini, Lisca and Stipsicz in [5]; wherever applicable, we use their notation. First we describe tight structures on M_p using Theorem 1.3, then we prove Theorem 1.3 in this special case.

Claim 3.1 *The manifold M_p admits exactly three tight contact structures up to isotopy.*

The finite collection of contact structures, given by Figure 1, can be encoded in the following table of invariants:

surgery coefficient	tb	$\mathrm{rot}(\lvert\mathrm{rot}\rvert \leq -\mathrm{tb}-1)$
$+1$	-1	0
$+1$	-1	0
-1	-2	$\mathrm{rot}_1 \in \{-1, 1\}$
-1	-2	$\mathrm{rot}_2 \in \{-1, 1\}$
-1	$-p$	$\mathrm{rot}_3 \in \{-p+1, -p+3, \ldots, p-1\}$

As an application of the theorem, the tightness and isotopies can be recognized solely from the induced Spin^c structures and the two invariants. In our case these are as follows:

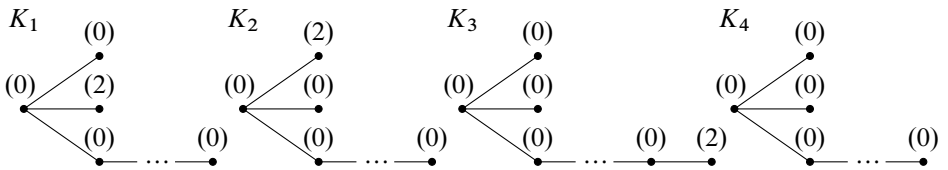
$$\begin{aligned} d_3(\xi) &= \tfrac{1}{4}(c^2(X,J) - 3\sigma(X) - 2b_2(X)) + q \\ &= \tfrac{1}{4}((0,0,\mathrm{rot}_1,\mathrm{rot}_2,\mathrm{rot}_3)\,Q_X^{-1}(0,0,\mathrm{rot}_1,\mathrm{rot}_2,\mathrm{rot}_3)^T - 3\cdot(-1) - 2\cdot 5) + 2. \end{aligned}$$

So, for mixed $(\mathrm{rot}_1,\mathrm{rot}_2) = (\pm 1, \mp 1)$, the d_3 is always zero, as for $(\mathrm{rot}_1,\mathrm{rot}_2) = (\pm 1,\pm 1)$ it runs through the values $\{\frac{1}{4}(2-p),\dots,\frac{1}{4}(-2+3p)\}$ by the step ± 1 as rot_3 increases.

There are exactly four Spin^c structures for each p (as $|H_1(M_p;\mathbb{Z})| = 4$):

$$\begin{aligned} H_1(-M_p;\mathbb{Z}) &= \left\langle \mu,\mu_a,\mu_b,\mu_c \;\middle|\; \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & p & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mu \\ \mu_a \\ \mu_b \\ \mu_c \end{pmatrix} = 0 \right\rangle \\ &= \begin{cases} \langle \mu_b; 4\mu_b = 0\rangle \cong \mathbb{Z}_4 & \text{for } p \text{ odd}, \\ \langle \mu_b,\mu_c; 2\mu_b = 2\mu_c = 0\rangle \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2 & \text{for } p \text{ even}. \end{cases} \end{aligned}$$

They can be given by the set $\{\boldsymbol{t}_1 = \boldsymbol{t}_4 + \mu_b, \boldsymbol{t}_2 = \boldsymbol{t}_4 + \mu_c, \boldsymbol{t}_3 = \boldsymbol{t}_4 + \mu_a, \boldsymbol{t}_4\}$. And corresponding four characteristic 2–cohomology classes, realizing $d(-M_p,\boldsymbol{t}_i)$, are on the generators of $H_2(W_{\Gamma'})$ given by



Therefore,

$$\begin{aligned} d(-M_p,\boldsymbol{t}_i) &= \max\{\tfrac{1}{4}(c_1(\boldsymbol{s})^2 + |\Gamma'|) : \boldsymbol{s} \in \mathrm{Spin}^c(W_{\Gamma'}),\ \boldsymbol{s}|_{-M_p} = \boldsymbol{t}_i\} \\ &= \begin{cases} 0 & \text{if } i = 1,2, \\ \frac{1}{4}(p-2) & \text{if } i = 3, \\ \frac{1}{4}(p+2) & \text{if } i = 4. \end{cases} \end{aligned}$$

Applying Theorem 1.3, the above computations already give that for distinct $\mathrm{rot}_1,\mathrm{rot}_2$ all structures are tight, and belong to two different isotopy classes, while for equal rot_1 and rot_2 the only tight triples are $(\pm 1,\pm 1,\mp(p-1))$, and they are isotopic to each other. This proves Claim 3.1.

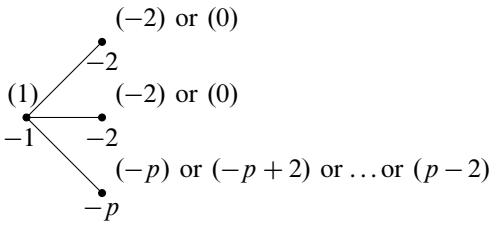
Claim 3.2 *Theorem 1.3 holds for M_p.*

We show this following the two-step analysis described in Section 2.

3.1 Detect tightness

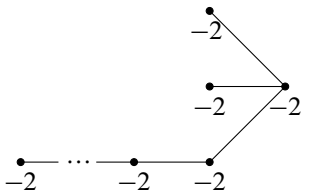
The condition we use to recognize tight structures among all (M_p, ξ) presented by surgery diagrams of Figure 1 is an existence of the characteristic covector c as in Theorem 2.1.

We give c as $\mathrm{PD}(c) = \alpha h + \sum \alpha_i e_i$, where $\alpha, \alpha_i \in \{\pm 1\}$, and such that $(c_i) = \langle c, x_i \rangle = \mathrm{rot}_i - 1$. Concretely, the c–evaluations on Γ belong to one of the following:



$$z = e_1,$$
$$x_1 = h - e_1 - e_2 - e_5,$$
$$x_2 = h - e_1 - e_3 - e_6,$$
$$x_3 = h - e_1 - e_4 - \sum_{7}^{p+5} e_i.$$

Then, for each such (α, α_i) we compute $c|_{\Gamma'}$, and check how its full path ends:



$$z' = h - e_2 - e_3 - e_4,$$
$$x'_1 = e_2 - e_5,$$
$$x'_2 = e_3 - e_6,$$
$$x'_3 = e_4 - e_7,\ e_7 - e_8,\ \ldots,\ e_{p+4} - e_{p+5}.$$

In Tables 1 and 2 we tabulate all possible (α, α_i) for each given triple (c_1, c_2, c_3). We will make explicit how some $c|_{\Gamma'}$ drop out, and connect the others to the right initial and terminal vector. Also, we will emphasize the appearance of the same characteristic covectors $c|_{\Gamma'}$ in some pairs of c–triples.

First observe that (on the level of paths) the order of signs on generators of each leg is unimportant, as they can be shuffled using $\pm 2\,\mathrm{PD}(v')$–steps for $\langle c, v' \rangle = \pm 2$. Then there are essentially only two different sign-vectors (α, α_i) for a chosen c–triple, differing in the sign of h. The two are connected by $\pm 2\,\mathrm{PD}(z')$, applied when $\langle c, z' \rangle = \pm 2$. Notice that all these different sign configurations belong to the same surgery presentation.

In the light of the previous paragraph, we record only the number of positive and negative signs on exceptional generators of each leg. Write $\{m+, n-\}_i$ when there are m

c_3	$(-2,-2,c_3)$		$(0,0,c_3)$	
	$(h-)$	$(h+)$	$(h-)$	$(h+)$
	$\{1+,1-\}_1$	$\{0+,2-\}_1$	$\{2+,0-\}_1$	$\{1+,1-\}_1$
	$\{1+,1-\}_2$	$\{0+,2-\}_2$	$\{2+,0-\}_2$	$\{1+,1-\}_2$
$p-2$	$\{p+,0-\}_3$	$\{(p-1)+,1-\}_3$	$\{p+,0-\}_3$	$\{(p-1)+,1-\}_3$
$p-4$	$\{(p-1)+,1-\}_3$	$\{(p-2)+,2-\}_3$	$\{(p-1)+,1-\}_3$	$\{(p-2)+,2-\}_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$-p$	$\{1+,(p-1)-\}_3$	$\{0+,p-\}_3$	$\{1+,(p-1)-\}_3$	$\{0+,p-\}_3$

Table 1

positive and n negative generators of L_i (counted without h and e_1); not to be confused with vectors of signs which record exact sign configuration on corresponding generators. In addition, let $(h+)_{(c_1,c_2,c_3)}$ and $(h-)_{(c_1,c_2,c_3)}$ denote any of the sign configurations which belongs to (c_1,c_2,c_3) and has positive and negative sign, respectively, on h. We look separately at the cases with the same, and later with distinct (c_1,c_2).

For $(-2,-2,c_3)$ with $c_3\in\{p-4,\dots,-p\}$, there exists a configuration $(h,e_2,e_3,e_4)=(-,-,-,-)$ which drops out: $\langle c,z'\rangle=\langle -h-e_2-e_3-e_4,h-e_2-e_3-e_4\rangle=-4$. Similarly, $(0,0,c_3)$ with $c_3\in\{p-2,\dots,-p+2\}$ drops out at $(h,e_2,e_3,e_4)=(+,+,+,+)$. Therefore, the paths possibly end only for the triples $(-2,-2,p-2)$ and $(0,0,-p)$.

Furthermore, we observe that $(-2,-2,p-2)$ and $(0,0,-p)$ belong to the same full path because the configurations $(h-)_{(-2,-2,p-2)}$ and $(h+)_{(0,0,-p)}$ give the same characteristic vector (zeroes on the third leg, and (h,e_4): $(-,+)\leftrightarrow(+,-)$ with the same evaluation on $z'=h-e_2-e_3-e_4$). This proves also that their (common) path indeed ends, namely at K_3 — given by $(h,e_4,e_7,\dots,e_{p+4},e_{p+5})=(-,-,-,\dots,-,+)$ for $(0,0,-p)$ — on the initial side and at $-K_3$ — as $(h,e_4,e_7,\dots,e_{p+4},e_{p+5})=(+,+,+,\dots,+,-)$ for $(-2,-2,p-2)$ — on the terminal.

Sign configurations adapted to any c–triple with distinct c_1 and c_2 build a connected part of (one of the two) full paths. Indeed, let us see how these parts patch together into a path. For $k\in\{1,\dots,p-1\}$, we have

$$(h+)_{(-2,0,p-2k)}\underset{c|_{\Gamma'}}{=}(h-)_{(0,-2,p-2k-2)}\underset{c|_{\Gamma}}{\equiv}(h+)_{(0,-2,p-2k-2)}\underset{c|_{\Gamma'}}{=}(h-)_{(-2,0,p-2k-4)},$$

c_3	$(-2,0,c_3)$		$(0,-2,c_3)$	
	$(h-)$	$(h+)$	$(h-)$	$(h+)$
	$\{1+,1-\}_1$	$\{0+,2-\}_1$	$\{2+,0-\}_1$	$\{1+,1-\}_1$
	$\{2+,0-\}_2$	$\{1+,1-\}_2$	$\{1+,1-\}_2$	$\{0+,2-\}_2$
$p-2$	$\{p+,0-\}_3$	$\{(p-1)+,1-\}_3$	$\{p+,0-\}_3$	$\{(p-1)+,1-\}_3$
$p-4$	$\{(p-1)+,1-\}_3$	$\{(p-2)+,2-\}_3$	$\{(p-1)+,1-\}_3$	$\{(p-2)+,2-\}_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$-p$	$\{1+,(p-1)-\}_3$	$\{0+,p-\}_3$	$\{1+,(p-1)-\}_3$	$\{0+,p-\}_3$

Table 2

where the first and the last equality denote the same characteristic covector on Γ', while the middle equivalence means (different sign distributions of) the same presentation. This separates all characteristic vectors arising from presentations with mixed (c_1,c_2) into two full paths. One starting at K_1 — as $-h-e_2+e_5+e_3+e_6+e_4+e_7+\cdots+e_{p+5}$ for $(-2,0,p-2)$ — and ending at $-K_2$ — as $+h-e_2-e_5+e_3-e_6-e_4-e_7-\cdots-e_{p+5}$ for $(-2,0,-p)$ — or $-K_1$ — as $+h+e_2-e_5-e_3-e_6-e_4-e_7-\cdots-e_{p+5}$ for $(0,-2,-p)$. The other starting at K_2 — as $-h+e_2+e_5-e_3+e_6+e_4+e_7+\cdots+e_{p+5}$ for $(0,-2,p-2)$ — and ending at $-K_1$ — as $+h+e_2-e_5-e_3-e_6-e_4-e_7-\cdots-e_{p+5}$ for $(0,-2,-p)$ — or $-K_2$ — as $+h-e_2-e_5+e_3-e_6-e_4-e_7-\cdots-e_{p+5}$ for $(-2,0,-p)$. The two terminal possibilities depend on the parity of p (odd or even).

In conclusion, translated back into rotation numbers we have obtained the following paths of tight structures, each sharing the same invariants:

- $(-1,-1,p-1)$ and $(1,1,-p+1)$ with $\mathrm{Spin}^c = \boldsymbol{t}_4+\mu_a$ and $d_3 = \frac{2-p}{4}$.
- $(-1,1,p-1)$ and $(1,-1,p-3)$ and $(-1,1,p-5)$ and … with $\mathrm{Spin}^c = \boldsymbol{t}_4+\mu_b$ and $d_3 = 0$.
- $(1,-1,p-1)$ and $(-1,1,p-3)$ and $(1,-1,p-5)$ and … with $\mathrm{Spin}^c = \boldsymbol{t}_4+\mu_c$ and $d_3 = 0$.

3.2 Prove overtwistedness and describe contact isotopies

In our (simplest possible) cases with boundary slopes $\frac{1}{k}$ for $k \in \mathbb{Z}$, there is a single continued fraction block for each special fiber. Contact surgery presents a direct translation between positive and negative stabilizations (down- and up-cusps) of core Legendrian unknots and positive and negative basic slices in the decomposition of a continued fraction block with slopes -1 and $-k$. The generators forming the

corresponding leg (and by that, the dual vertices) in the plumbings above can be thought of as another way of layering solid torus into k slices.

We need contact topological interpretation for the steps in full paths.

First, the unimportance of sign permutations in the legs coincide with the shuffling of basic slices within a single continued fraction block [7, Lemma 4.14]. Moreover, Section 6 of [5] provides sufficient isotopy moves between contact structures presented by different surgery diagrams. Let us spell this out. Since the moves in [5] are given by the matrices of signs whose coefficients are q^i_j, the number of positive basic slices in the jth continued fraction block of the ith leg — in our case only (q^1_0, q^2_0, q^3_0) — we rewrite the previously obtained paths of tight structures in this language, changing rotation numbers to the q^i_0:

- $(0,0,p-1)$ and $(1,1,0)$.
- $(0,1,p-1)$ and $(1,0,p-2)$ and $(0,1,p-3)$ and ....
- $(1,0,p-1)$ and $(0,1,p-2)$ and $(1,0,p-3)$ and ....

Now we notice that the conditions which caused a full path to drop out, and so prevented our tightness criterion from working, exactly agree with the condition for which overtwistedness can be proved. And, finally, there are contact isotopies between pairs of surgery presentations which share the same path. Let us compare.

Proposition 3.3 [5, Propositions 6.3, 6.1 and 6.4] *Let a contact structure on M_p be given by (q^1_0, q^2_0, q^3_0) as above. Then the triples $(1,1,q^3_0)$ with $q^3_0 \neq 0$ and $(0,0,q^3_0)$ with $q^3_0 \neq p-1$ present overtwisted structures. Between other presentations, there are the following contact isotopies:*

$$(1,0,q^3_0) \simeq \begin{cases} (0,1,q^3_0+1) & \text{when } q^3_0 < p-1, \\ (0,1,q^3_0-1) & \text{when } q^3_0 > 0, \end{cases} \quad \text{and} \quad (1,1,0) \simeq (0,0,p-1). \ \square$$

Problems in general

The examples shown above are special in several ways. In general, it can happen that the full path associated to some presentation $(c_v)_{v\in\Gamma}$ drops out, although all characteristic covectors computed from (α,α_i)–configurations which restrict to $(c_v)_{v\in\Gamma}$ satisfy the bounds $v\cdot v \leq \langle c, v\rangle \leq -v\cdot v$ for all $v \in \Gamma'$. Also, not all the steps in a full path need to be presentable, that is, arising from some tuple of rotation numbers. (For examples of such paths, look at the two "applications" in [13].) That said, we need to

find out how the (subsequent) presentations of the same path are related when neither of their characteristic covectors on Γ' coincides (Corollary 4.14). Finally, we need new conditions for overtwistedness (Proposition 5.1) and isotopies (Proposition 5.2), which will explain such behavior of full paths.

4 Characteristic covectors, tightness, and full paths

In Section 2.1, we have associated characteristic covectors on Γ' to any given surgery presentation. Here we investigate their full paths. Namely, how these paths end, and which presentations share the same path. In order to do so, we first observe that certain 2 PD–steps do not change the presentation (Section 4.1). Then we explore the only remaining central step — concretely, we explain it on the level of homology generators (Section 4.2). In the following subsections, we are then concerned with the associated change in $c|_\Gamma$, whether this new $c|_\Gamma$ comes from some presentation and when it leads to the end of the path (Section 4.3). Moreover, we describe (in Section 4.4) the first presentable $c|_\Gamma$ (or the end of path) following any possible starting point.

Notation 4.1 We describe a characteristic 2–cohomology class $c \in H^2(R;\mathbb{Z})$ as $\mathrm{PD}(c) = \alpha h + \sum \alpha_i e_i$, where $\alpha, \alpha_i \in \{\pm 1\}$. In the following, vectors of signs correspond to parts of the coefficient-vector (α, α_i), covering generators of (usually) a single Γ– or Γ'–vertex.

We often call a single vertex by its self-intersection. When a vertex is written out in generating classes, these are called starting, middle and last, according to the position; explicitly, if $v = e_s - \sum_{j=s+1}^{l} e_j$, then e_s is starting, e_l is last, and all others are middle. On legs, the starting generator of a vertex and the last generator of the previous vertex coincide.

Presentability will be assigned to dual vectors and it means that the corresponding manifold side arises from a contact presentation, that is, the manifold-side evaluations can be expressed by rotation numbers as in Figure 3.

Let our starting point be a characteristic vector which comes from a contact presentation, and which satisfies $v \cdot v \le \langle c, v\rangle \le -v \cdot v$ (otherwise we have already dropped out). We will follow the path only in one direction — towards the initial vector. Recall that the corresponding step is given by $-2\,\mathrm{PD}(v)$ for some v with $\langle c, v\rangle = v \cdot v$, and the vector we aim at satisfies $v \cdot v + 2 \le \langle c, v\rangle \le -v \cdot v$. Everything could be verbatim repeated with opposite signs in the direction of the terminal vector.

4.1 Steps on legs

First, we observe that steps taken for $v \neq z'$ never change the presentation considered, nor does the path drop out at any of these vertices. (To the remaining case $v = z'$ we dedicate Section 4.2.)

Lemma 4.2 *Characteristic vectors c and $c - 2\,\mathrm{PD}(v)$ for $v \neq z'$ with $\langle c, v\rangle = v \cdot v$ always belong to the same surgery presentation.*

Proof As these vertices ($v \in \Gamma'$ and $v \neq z'$) are described by $v = e_i - \sum e_j$, the evaluation of the characteristic covector c reaches the self-intersection when presenting generators all admit the same sign as in the vertex. So $-2\,\mathrm{PD}(v)$ changes their signs from $(+-\cdots-)$ to $(-+\cdots+)$. But this change has no effect on the evaluation of c on any of the Γ–vertices.

Indeed, from the way how the exceptional classes are chosen we see that each e_j starts some new vertex, either one on the manifold side or one on the dual side. So the starting and the last generator of v are nonstarting on the manifold side, while all its middle generators are starting (and last) generators of manifold vertices. Hence, the restriction of c to the generators of v evaluates trivially on Γ, that is, $\langle c|_v, \Gamma\rangle = 0$, and is therefore independent of sign.

Since these (manifold-side) evaluations directly correspond to rotation numbers, with neither of these moves do we switch between presentations. □

Lemma 4.3 *All drop-outs occur in the center $z' = h - e_2 - e_3 - e_4$ of the dual star.*

Proof We notice that all the vertices in legs of Γ' are formed by exactly as many generators (the e_j) as the value of their self-intersections. Hence, there is no way to drop out at any of them. So the only possible drop-out happens at z' when the signs of generators h and e_2, e_3 and e_4 are all the same — $(++++)$ or $(----)$ — and $\langle c, z'\rangle = \pm 4$. □

In sum, we may assume the initial condition $v \cdot v + 2 \leq \langle c, v\rangle \leq -v \cdot v$ is violated only at the central vertex z' — such a vector can be easily reached by finishing all possible $-2\,\mathrm{PD}$–steps on legs, which either sweep out the problem or transfer it to the center. (As each $-2\,\mathrm{PD}$–step pushes the problem to the neighboring vertices, we are successively completing the steps, as long as we do not run into a vertex v which despite of the -2–change does not evaluate as $v \cdot v$, or we reach the end of the leg.) In particular, neither noncentral vertex is of the form $(+-\cdots-)$.

4.2 Central step

After the above reduction, covector c either drops out at z', presents the initial vector or it reaches self-intersection at z'. For the latter, the generators forming $z' = h - e_2 - e_3 - e_4$ take values either $(+---)$ or $(-+--)$, up to reordering the legs. The -2 PD–step taken next changes exactly these generators by twice $(-+++)$. In the first case we stay in the same presentation, as the step only switches the signs in the pairs (h, e_i) for $i = 2, 3, 4$, preserving the evaluation on all the influenced manifold vertices, x_1, x_2 and x_3. In the second case, we can (on the level of generators) instead of simply adding $-2z'$ to the given description of $\mathrm{PD}(c)$, first change the sign configuration, without changing the dual $c|_{\Gamma'}$ and with controlled (seen later) change on the manifold side $c|_{\Gamma}$, and then do the -2 PD–step as above, not influencing the manifold side.

Algorithm 4.4 (central step or turn) Whenever we arrive, after possibly renumbering the legs, at c with $(h, e_2, e_3, e_4) = (-+--)$ and $\langle c, v\rangle \neq v \cdot v$ for all $v \neq z'$, the next step in the full path is given by the characteristic covector $\bar{c}$ as follows. Denoting vertices of L'_1 by $\{v'_0, \dots, v'_{k'_1}\}$ and their generators as $v'_i = e^i_1 - \sum_{j=2}^{l_i} e^i_j$ with $e^i_{l_i} = e^{i+1}_1$ and $e^0_1 = e_2$, define $\mathrm{PD}(\bar{c}) = \mathrm{PD}(c) + 2h - 2e_2$ and modify it as follows:

for $i \in \{0, \dots, k'_1\}$ if $\langle \bar{c}, e^i_1\rangle \neq \langle c, e^i_1\rangle$:
  for $j \in \{2, \dots, l_i\}$ if $\langle c, e^i_j\rangle = +1$:
    $\mathrm{PD}(\bar{c}) = \mathrm{PD}(c) - 2e^i_j$
  endfor
if $\langle \bar{c}, e^i_1\rangle = \langle c, e^i_1\rangle$:
  stop

Then add $-2z'$ to the so-obtained sign configuration $\mathrm{PD}(\bar{c})$.

To prove well-definedness, we need that this reformulation always exists (the inner loop in our algorithm always stops; see Lemma 4.5) and that uniqueness, ensured by always taking the first positive generator (chosen ordering of the inner loop), can be explained by the independence of order, at least as far as contact presentations are concerned (Lemma 4.6).

Lemma 4.5 *Every characteristic vector $c_{\Gamma'}$ with $\mathrm{PD}(c|_{z'}) = (-+--)$ can be achieved by another distribution of signs, with positive sign on h; it is associated to a different manifold vector* (possibly nonpresentable).

Proof Starting at the center z', the two distributions are given by $(-+--)$ and $(+---)$. The switch of the h–sign with the opposite sign of e_2 does not impose any change into the second and the third dual leg. For the first leg, the appropriate adaptation of signs, which results in the same dual evaluation, exists because of the exclusion of any $(+-\cdots-)$–configurations (that is, the assumption $\langle c, v\rangle \neq v \cdot v$ for any $v \neq z'$). □

Lemma 4.6 *If a sign on one middle generator of a dual vertex is changed, all of them need to be changed (independent of order) before we get back into presentable. A turn of the last generator can result in a presentable vector only when all prior middle generators are negative.*

Proof For a covector to be presentable, all dual vertices have to have same-signed middle generators, because these generators on the manifold side are forming a chain of -2s, zero being their only possible rotation number.

For the second claim, suppose on the contrary the middle signs on some v' are positive. Changing the sign of its last generator (from positive to negative) forces a switch of all the signs in the following chain (if any) of dual -2s (to preserve dual evaluations). Then, this influences the evaluation on the next non-(-2) dual vertex w', which can be corrected by changing one of its later generators from positive to negative. If the middle generators of w' are already negative or if we get them all negative by the current turn, we have found (independent of further changes) a manifold-side vertex which starts at positive (second last) generator in v' and has all further signs negative. If by the change of one middle generator not all of them are negative, the vector is nonpresentable by the first part. If all generators of w' are positive, and we turn the last one, we need to repeat the same argument with w' in place of v'. It remains to check whether we could get presentable result by correcting only starting and last generators of all following (necessarily, fully positive) dual vertices. But if not before, the process ends in nonpresentable, giving $(+-\cdots-)$ on the last manifold vertex. □

To sum up, the central turns are the only significant steps in following possible changes on manifold vectors, and by that, in presentations. We may assume that after each central turn also all -2 PD–steps on legs are finished.

4.3 On turning sequences and presentability

To begin, notice how to recognize the ends of a full path.

Lemma 4.7 *If after a central step, the covector c on the starting dual vertices evaluates as their self-intersection, that is, $\langle c, x_i' \rangle = x_i' \cdot x_i'$, on*

- *at most one leg, we have arrived at the initial end;*
- *two legs, the full path continues;*
- *all three legs, this causes a drop-out.*

Proof The maximal starting dual evaluations tell us on how many legs we need further -2 PD–steps. The evaluation $\langle c, z' \rangle$ on z' right after a central turn is $+2$. If further turns are needed for one leg only, we do not reach a -2 central evaluation again and the corresponding vector is initial; with two we get back to $\langle c, z' \rangle = -2$ and we continue with another central turn; three gives a drop-out in $(h, e_2, e_3, e_4) = (----)$. □

Corollary 4.8 *A presentation ξ whose $\mathcal{P}_\xi$ properly ends necessarily admits a leg, starting in a fully positive vertex.* (If the presentation corresponds to the initial vector, there are two fully positive starting vertices.)

Proof For PD(c) take a sign configuration which evaluates on manifold vertices according to the rotation numbers of ξ, which takes minus sign on h, and for which $\langle c, v' \rangle \neq v' \cdot v'$ for all $v' \in \Gamma' \setminus \{x_i' : i = 1, 2, 3\}$. (This is the stage right after a central turn.) As in Lemma 4.7 above, there is a leg, say L_1, for which $\langle c, x_1' \rangle \neq x_1' \cdot x_1'$. We prove that on this leg $\langle c, x_1 \rangle = a_0^1 - 2$ holds, that is, the generators of x_1 (apart from h and e_1) are positive.

Write out x_1 as $h - e_1 - e_2 - e_5 - \cdots - e_J$. The signs on the generators up to e_{J-1} are positive as otherwise we would have shuffled the negative sign to e_5 by -2 PD–steps on consecutive dual vertices of square -2 (resulting in $\langle c, x_1' \rangle = -2$ for $x_1' = e_2 - e_5$). The positivity of e_J follows from presentability via the following claim.

Claim *A presentable covector on neither dual vertex takes the form $(+ - \cdots - +)$.*

Proof A proof of this fact is basically the same as the second part of the proof of Lemma 4.6. Suppose on the contrary, there is such a dual vertex; it is not the last vertex of the dual leg, because it would give the last vertex on the manifold side with self-intersection -2 and $+2$ c–evaluation. But then, every non-(-2) dual vertex further on the dual leg needs to have again negative middle signs (otherwise we have found a manifold vertex, starting in the negative sign of the previous non-(-2) with all following generators positive) and positive last one (because of $(+ - \cdots -)$ exclusion). After all, we end in the impossible last dual vertex. □

Since also $(+-\cdots-)$–configuration on any dual vertex, except x_i', is excluded by $\langle c, v'\rangle \neq v'\cdot v'$, and since middle generators of any dual vertex are same-signed, we get that e_J is positive. It is a middle generator of a dual vertex starting in positive e_{J-1}. □

The leg with fully positive starting vertex is the one which in the reordering of Algorithm 4.4 takes the role of L_1. When we wish to emphasize the leg according to which the central step is done, we refer to it as a turn of L_i.

Since the evaluation of characteristic covector on L_i–vertices changes only by turns of L_i, we may separately study their influence.

Lemma 4.9 *Let c be a presentable noninitial characteristic covector. Assume that it evaluates on the vertices of some leg $L = (-a_0, -a_1, \dots, -a_j, -a_{j+1}, \dots, -a_k)$ as*

$$\langle c, L\rangle = (a_0-2, a_1-2, \dots, a_j-2, a_{j+1}-2-2n_{j+1}, \dots, a_k-2-2n_k),$$

where $k \geq j$, $n_{j+1}, \dots, n_k \geq 0$ and $n_{j+1} > 0$.

The path runs into the next possibly presentable covector $\bar{c}$, only after

$$1+1+(a_1-1)+(a_2-1)(a_1-1)+\cdots+(a_{j-1}-1)\cdots(a_1-1) \quad \textit{turns of } L,$$

in

$$\langle \bar{c}, L\rangle = (-a_0, -a_1+2, \dots, -a_j+2, a_{j+1}-2n_{j+1}, \dots, a_k-2-2n_k).$$

Proof To be illustrative, we explicitly write out all the generators involved in the first few turns. Below are the two sides, L_1 and L_1', in homology generators; the $*$–symbol stands for truncation only:

$$\begin{aligned}
L_1\colon\quad & x_1 = h - e_1 - e_2 - e_5 - \cdots - e_{J-1} - e_J \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad e_J - e_{J+1} \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad e_{J+1} - e_{J+2} \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \ddots \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad e_{K-1} - e_K - e_{K+1} - *, \\
L_1'\colon\quad & x_1' = e_2 - e_5 \\
& \qquad\qquad\quad \ddots \\
& \qquad\qquad\quad e_{J-2} - e_{J-1} \\
& \qquad\qquad\qquad\quad e_{J-1} - e_J - e_{J+1} - \cdots - e_K \\
& \qquad\qquad\qquad\qquad\quad *.
\end{aligned}$$

In this notation, the starting part of L_1 and the evaluation of c on it take values

$$L_1 = (-J+3, \underbrace{-2, \dots, -2}_{K-J-1}, -T, -S, *) \quad \text{and} \quad \langle c, L_1 \rangle = (J-5, 0, \dots, 0, M, N, *).$$

By the first turn, according to the algorithm, we change generators up to e_J —it does not influence further dual vertices, but a new vector can be presentable only when all the middle generators $e_J, \dots, e_{K-1}$ are same-signed. Therefore, in order to (possibly) reach a presentable vector again we have to repeat turning of this particular leg $K-J$ times. The resulting manifold vector is of the form $(-J+3, 0 \dots, 0, M+2, N, *)$; its presentability depends on the $(+2)$–changed manifold vertex $e_{K-1} - e_K - e_{K+1} - *$.

In terms of generators, we have reached another presentation exactly when the e_K- sign is negative. The positive e_K–sign, on the other hand, requires another turn, but this forces some further changes to preserve the dual. Namely, we need to change signs on generators of the following chain of -2s, and one (without loss of generality, first) middle generator afterwards. The resulting vector is not necessarily presentable, provided the starting point was; it depends on presentability of the vertex starting in the (last changed) middle generator ($+2$ rotation change). But if it is, the new presentation is $(-J+3, 0, \dots, 0, -T+2, N+2, *)$; for this, we need to turn this leg $K-J+1$ times.

Continuing in the same manner, we trace similar behavior at all levels. Concretely: We are successively turning fully positive vertices, which influences the evaluation on the following manifold vertex by $+2$. If the result is presentable, we have finished. Otherwise, the following vertex was also fully positive, at the moment its evaluation is minus self-intersection, and it will have turned under the influence of another turn of the previous vertex. For that we need to bring the previous vertex back to maximal rotation, using (again) influence of the previous vertices on the leg. But notice that each vertex is influenced only by turns of the vertex just before it. Therefore, to come from maximal rotation through minus self-intersection to minimal rotation on some vertex v_{k+1}, we need to influence it by two turns of its immediately previous vertex v_k. This in turn is obtained by $a_k - 1$ turns of its previous vertex v_{k-1}, first to get from minus self-intersection to minimal rotation, and then by the step of $+2$ to maximal rotation. This explains the number of steps and finishes the proof. □

Obviously, the leg (its vertices with self-intersections) together with the sign configuration (in presentable, rotation numbers) determine when the leg is turned. In particular, it specifies the gaps between the subsequent turnings of the same leg, when some other leg needs to be turned in order for the path to continue. Actually, the reverse also holds.

Lemma 4.10 *A form of a leg together with a distribution of signs on its generators is completely described by the sequence of its turns.*

Proof As before, we separately state (and argue for) the first step.

Claim *Between two subsequent turnings of the same leg L, there are always either a_0-2 or a_0-1 turnings of other legs.*

Proof Remember that all generators (except possibly the last) of the starting manifold vertex on the turning leg are positive. Since by each turn of other legs we change starting evaluation by $+2$ (through the change of h–sign from negative to positive), the gap is determined by the number of generators of the starting vertex. Its variation by one is due to whether the dual vertex following -2s is also fully negative after the L–turn. □

That said, given a turning sequence, we get a_0 out of the size of gap between subsequent turnings. If the gap is always the same (a_0-1), this means that v_0 is the only vertex on L, and its self-intersection is $-a_0$.

Call this number of central turns between two turns of L, the 0th period of L. Let us define higher periods for a turning sequence:

1st period Number of times the 0th period is a_0-2 before it turns to a_0-1.

2nd period Number of times the 1st period is a'_0-1 before it turns to a'_0-2.

k^{th} period Number of times the $(k-1)^{\text{st}}$ period is $a'_{k-2}-1$ before it turns to $a'_{k-2}-2$.

The numbers should not be read from the first time round. As suggested by notation, they correspond to the self-intersections of dual vertices, hence they determine L', and from that L. The initial distribution of signs can be now recognized from the values of periods before the first change. □

4.4 Restrictions on the whole structure

We look at all possible (presentable) entries. For each we continue its path as long as it reaches another presentation, or otherwise it ends, either by a drop-out or a (nonpresentable) initial vector. Throughout we assume that the (normalized) Seifert constants are ordered $r_1 \geq r_2 \geq r_3$. In order to reduce the possibilities we invoke the L–space condition.

L–space condition Recall the numerical condition for M to be an L–space: there are no coprime integers m and a such that $1/r_1 > m/a$, $1/r_2 > m/(m-a)$ and $1/r_3 > m$; we say that the coefficients (r_1, r_2, r_3) are not realizable. As a direct consequence of this condition we observe that:

(i) $r_1 \geq \frac{1}{2}$; equivalently, one leg starts with -2 (otherwise the realizability condition is satisfied for coprime $m = 2, a = 1$).

(ii) If
$$L_1 = (\underbrace{-2, \dots, -2}_{k}, *), \quad \text{so} \quad \frac{1}{r_1} > \frac{k+2}{k+1} = -(\underbrace{[2, \dots, 2]}_{k+1}),$$
then $1/r_2 \leq k+2$; equivalently, $x_2 \cdot x_2 \geq -k-2$.

Dual configurations In the following arguments, there will frequently appear a pair of (truncated) legs which are dual to each other, that is, describing a lens space and its dual. Recall that the coefficients of the two are related as follows (here, $-2^{\times b_\iota}$ means a chain of b_ι–many -2s):
$$L_i = (-b_1 - 2, -2^{\times b_2}, -b_3 - 3, \dots, -b_m - 2),$$
$$L_j = (-2^{\times b_1}, -b_2 - 3, -2^{\times b_3}, \dots, -2^{\times b_m}).$$
The inverses of the continued fractions they describe add up to -1.

So, our starting point is a presentable characteristic covector, which does not present an initial end, or a drop-out. Thus, exactly one leg of the corresponding presentation starts in a fully positive vertex. We separate the cases: in Proposition 4.11 we gather presentations for which either v_0^1 or v_0^2 is stabilized fully positively, and in Proposition 4.13 we cover presentations for which v_0^3 is fully positively stabilized.

Proposition 4.11 *Let c be a presentable noninitial characteristic covector, associated to a presentation with fully positive starting vertex on L_i, either L_1 or L_2. This means that it evaluates on the vertices of*
$$L_i = (-a_0^i, -a_1^i, \dots, -a_j^i, -a_{j+1}^i, \dots, -a_{k_i}^i)$$
as in Lemma 4.9:
$$\langle c, L_i \rangle = (a_0^i - 2, a_1^i - 2, \dots, a_j^i - 2, a_{j+1}^i - 2 - 2n_{j+1}^i, \dots, a_{k_i}^i - 2 - 2n_{k_i}^i)$$
for some $0 \leq j \leq k_i$, $n_{j+1}^i, \dots, n_{k_i}^i \geq 0$ and $n_{j+1}^i > 0$.

Denote coefficients on other two legs by a^{λ}_{κ}, and let c evaluate as $\langle c, v^3_0\rangle = a^3_0 - 2 - 2n^3_0$ on the first vertex of the third leg L_3, and

$$\langle c, L_l\rangle = (-a^l_0, -a^l_1+2, \dots, -a^l_k+2, a^l_{k+1}-2-2n^l_{k+1}, \dots, a^l_{k_l}-2-2n^l_{k_l})$$

on the leg L_l for $l \neq i, 3$ for some $-1 \le k \le k_l$ and $n^l_{k+1} \le a^l_{k+1} - 3$.

Furthermore, define m and N^l_{m+1} as follows:

$$m := \max\{\kappa \le k : \text{denominator of } [a^l_0, \dots, a^l_\kappa, N] \le n^3_0 \text{ for some } N\},$$

$$N^l_{m+1} := \begin{cases} n^l_{k+1}+1 & \text{if the denominator of } [a^l_0, \dots, a^l_k, n^l_{k+1}+1]^{-1} \le n^3_0, \\ N & \text{otherwise.} \end{cases}$$

where $N \in [1, a^l_{m+1})$ is such that the denominator of $[a^l_0, \dots, a^l_m, N]^{-1} \le n^3_0$ and the denominator of $[a^l_0, \dots, a^l_m, N+1]^{-1} > n^3_0$. Then the full path of c behaves as follows:

(1) *If $-[a^i_0, \dots, a^i_{j-1}]^{-1} - [a^l_0, \dots, a^l_m, N^l_{m+1}]^{-1} < 1$, the full path drops out.*

(2) *If $-[a^i_0, \dots, a^i_{j-1}]^{-1} - [a^l_0, \dots, a^l_m, N^l_{m+1}]^{-1} = 1$ and $m = k$ and $N^l_{m+1} = n^l_{k+1} + 1$, we reach a new presentation $\bar{c}$ which on the three legs takes the following values:*

$$\langle \bar{c}, L_i\rangle = (-a^i_0, -a^i_1+2, \dots, -a^i_j+2, a^i_{j+1}-2n^i_{j+1}, \dots, a^i_{k_i}-2-2n^i_{k_i}),$$

$$\langle \bar{c}, L_l\rangle = (a^l_0-2, a^l_1-2, \dots, a^l_{k+1}-2, a^l_{k+2}-2n^l_{k+2}, \dots, a^l_{k_l}-2-2n^l_{k_l}),$$

$$\langle \bar{c}, v^3_0\rangle = a^3_0 - 2 - 2n^3_0 + 2D,$$

where D stands for the denominator of $[a^l_0, \dots, a^l_k, n^l_{k+1}+1]^{-1}$ and the evaluations on the rest of L_3 remain the same as for c.

(3) *Otherwise, the path continues in nonpresentable and reaches the nonpresentable initial end.*

Remark 4.12 Rewrite the coefficients up to a^i_j in the b–notation used for dual configurations above, so for appropriate $b_\iota \ge 0$ (notice $b_1 > 0$ on L_1):

$$L_i = (-2^{\times b_1}, -b_2-3, -2^{\times b_3}, \dots, -2^{\times b_J}, -a^i_{j+1}, \dots, -a^i_{k_i}).$$

Also for L_l, truncated as continued fractions in the proposition, take

$$(a^l_0, \dots, a^l_m, N^l_{m+1}) = (-b'_1-2, -2^{\times b'_2}, -b'_3-2, \dots, -b'_J-2).$$

Then the conditions, given in the proposition in terms of continued fraction sums, can be restated as:

- The two continued fractions add up to 1 when $b'_k = b_k$ for all k.
- The sum is greater than 1 when, for $K = \min\{k : b'_k \neq b_k\}$, $b'_K < b_K$ if K is odd and $b'_K > b_K$ if K is even.
- The sum is smaller than 1 when, for $K = \min\{k : b'_k \neq b_k\}$, $b'_K > b_K$ if K is odd and $b'_K < b_K$ if K is even.

Proof of Proposition 4.11 We need to observe how specific behavior of the path restricts possible forms of a covector, and by that, of a presentation.

To meet another presentation, recall that we need to swap the signs of all generators forming the fully positive vertices $v^i_0, \dots, v^i_j$ (Lemma 4.9). To achieve this, we need certain number of L_i–turns, which are arranged in the turning sequence, uniquely determined by the form of L_i. So, any turn of other two legs should appear at exactly specified nonturning stages of L_i.

Now notice that any turn of L_3 (before finishing the specified sequence) would immediately end the path (in nonpresentable). Indeed, it has to appear after k (or $k+1$) turns of L_1 (which is the 1st period of L_1). So after it L_2 needs to be turned (its starting coefficient being $a^2_0 \leq k+2$, hence its 0th period being at most $k+1$) and also L_1 needs to be turned (having 0th period 0 or 1). But this already means we have arrived at the initial end; see Lemma 4.7.

Therefore, the turning sequence of L_i (up to its $(j+1)^{\text{st}}$ vertex) exactly specifies the turning sequence of L_l (needed to reach presentability on L_i again), its turns being in nonturning points of L_i, and vice versa. It is exactly the turning sequence of the dual leg with all evaluations fully negative. Rewritten in terms of relations between continued fractions, the two legs are of dual forms if and only if the corresponding continued fractions add up to one. In the dual (negative) leg for the last entry the significant information is the number of negative signs, as turning sequence depends only on whether we have reached maximal evaluation.

Taken together, we have obtained:

(1) If not all gaps in the turning sequence of L_i are filled by turns of L_l, and at the same time, the sequence is not quit by the turn of L_3 before or at the time when first such nonfilled gap appears, then the full path drops out.

(2) If the two continued fractions add up to exactly one, this means the turning sequences of corresponding legs exactly fit together, and we reach another

presentation — if and only if the evaluation on the starting vertex of the third leg is negative enough, not to quit the sequence of turnings interchanging between L_i and L_l. That is, there have to be more negative generators than there are turns of L_i and L_l, which equals the denominator of the corresponding continued fractions. The form of the new presentation is determined by Lemma 4.9.

(3) Otherwise, either we hit into some turning point of L_l before reaching the next gap in the sequence of L_i, or we reach a turning point of L_3 at or before the time when the L_i–gap is not filled by the L_l–turn for the first time. As observed above, in these cases, the full path properly ends with the initial vector, but it is necessarily nonpresentable because we have not yet reached the first possibly presentable stage as specified in Lemma 4.9. □

Proposition 4.13 *Throughout the path, there can be at most two noninitial characteristic covectors for which the starting vertex of L_3 is fully positive, that is $\langle c, v_0^3\rangle = a_0^3 - 2$.*

If $\langle c, v_1^3\rangle \neq a_1^3 - 2$ or $L_3 = (v_0^3)$, the turn of L_3 is presentable, the two presentations differ in $\langle \bar{c}, v_0^3\rangle = -a_0^3$ and $\langle \bar{c}, v_1^3\rangle = \langle c, v_1^3\rangle + 2$, and, for $l = 1, 2$, also $\langle \bar{c}, v_0^l\rangle = \langle c, v_0^l\rangle + 2$.

If $\langle c, v_1^3\rangle = a_1^3 - 2$, and c is not terminal, the turn of L_3 necessarily makes the continuation of the path nonpresentable, and ends it in a nonpresentable initial end.

If $\langle c, v_1^3\rangle = a_1^3 - 2$, and c is terminal, let us write out the c–evaluations at the terminal end:

$$\langle c, L_1\rangle = (-a_0^1, -a_1^1 + 2, \dots, -a_j^1 + 2, a_{j+1}^1 - 2 - 2n_{j+1}^1, \dots, a_{k_1}^1 - 2 - 2n_{k_1}^1),$$
$$\langle c, L_2\rangle = (-a_0^2, -a_1^2 + 2, \dots, -a_k^2 + 2, a_{k+1}^2 - 2 - 2n_{k+1}^2, \dots, a_{k_2}^2 - 2 - 2n_{k_2}^2),$$

for some $0 \le j \le k_1$ and $0 \le k \le k_2$. Then:

(1) *If for maximal $J \le j$ and $K \le k$ such that*

$$-[a_0^1, \dots, a_J^1, 3]^{-1} - [a_0^2, \dots, a_K^2, 2]^{-1} = 1,$$

the denominator of the two fractions is smaller than a_0^3, then the full path drops out.

(2) *If $\langle c, v_2^3\rangle \neq a_2^3 - 2$, and there exist $J \le j$ and $K \le k$ such that*

$$-[a_0^1, \dots, a_J^1, 3]^{-1} - [a_0^2, \dots, a_K^2, 2]^{-1} = 1,$$
$$n_{J+1}^1 \ge 2 \quad \text{or} \quad L^1 = (v_0^1, \dots, v_J^1),$$
$$n_{K+1}^2 \ge 1 \quad \text{or} \quad L^2 = (v_0^2, \dots, v_K^2),$$

and the denominator of the two fractions equals a_0^3, we reach a new presentation $\bar{c}$ which on the three legs takes the values

$$\langle \bar{c}, L_1\rangle = (a_0^1-2, a_1^2-2, \dots, a_J^1-2, \langle c, v_{J+1}^1\rangle + 4, \langle c, v_\iota^1\rangle_{\iota=J+2}^{k_1}),$$
$$\langle \bar{c}, L_2\rangle = (a_0^2-2, a_1^2-2, \dots, a_K^2-2, \langle c, v_{K+1}^2\rangle + 2, \langle c, v_\kappa^2\rangle_{\kappa=K+2}^{k_2}),$$
$$\langle \bar{c}, L_3\rangle = (-a_0^3, -a_1^3, \langle c, v_2^3\rangle + 2, \langle c, v_\mu^3\rangle_{\mu=3}^{k_3}).$$

This $\bar{c}$ presents the initial end of the full path.

(3) *Otherwise, the path continues in nonpresentable and reaches the nonpresentable initial end.*

Proof As observed in the proof of Proposition 4.11, any time when the path runs into a presentation with fully positive v_0^3 (not at its terminal end), it reaches the initial vector, either before or after a turn of L_3. Therefore, if the characteristic vector before the L_3–turn is noninitial and presentable, the only other presentation which can appear as we continue the path can occur straight after this turn. The resulting vector is presentable if and only if $\langle c, v_1^3\rangle \neq a_1^3-2$ (when it exists). The relation between the two presentations is as always read from Lemma 4.9 (the simplest possible — one-turn — case).

The only remaining option is to have a presentable terminal end with $\langle c, v_0^3\rangle = a_0^3-2$. In that case the turn of L_3 does not end the path, and if this turn is not presentable itself, we need to look for any possible following presentation. Since presentability on L_3 can be recovered only by a turn of L_3, and since according to the above this turn ends the path, we might meet such a presentation (only) at the initial end. This in particular means that turns of L_1 and L_2 in between the two turns of L_3 should begin and end with a vector which is presentable on these two legs. The first turn after the L_3–turn and the last turn before another L_3–turn are done according to L_1.

As before, we inductively determine that, in order for the turning sequences of L_1 and L_2 to fit together (being interchangeably turned until the second L_3–turn), the two legs need to have fully negative starting vertices, forming almost dual vectors — "almost" in the sense that there is no "last pair", that is, the two vectors as given in the paragraph on dual configurations end with $-b_m-3$ and $-2^{\times b_m}$ instead of $-b_m-2$ and $-2^{\times b_m}$. In other words, they are dual when enlarged by 3 and 2, respectively.

The three possibilities are now given as before:

(1) We have not reached the turning point of L_3 yet, but the sequence of turnings of L_1 and L_2 cannot continue.

(2) The turn of L_3 appears exactly in the moment when neither L_1 nor L_2 can be turned, and the sign configuration on them is presentable. Additionally, we need that $\langle c, v_2^3\rangle \neq a_2^3 - 2$ to reach presentability of L_3 as well.

(3) Otherwise, either the first two legs hit a common turning point, the third leg finishes the sequence early, not having enough negative generators, or (simply) the terminal vector obtained as in (2) is not presentable because $\langle c, v_2^3\rangle = a_2^3 - 2$ (hence, after the L_3–turn, $\langle \bar{c}, v_2^3\rangle = a_2^3$). In all the cases, the path stops in a nonpresentable initial end. □

Above we described how the successive presentations in the path are related to each other and indicate what property causes a drop-out. Any given presentation can now be either walked through these stages to the proper ends of the full path or it drops out. Joining results (taking into account also their analogues obtained by following the path in the terminal direction) we obtain the following picture. Here, the presentations are given as evaluations of characteristic covectors on generators of $H_2(W_\Gamma)$, written as triples of vectors c^i whose entries are $c^i_j = \langle c, v^i_j\rangle$. Vectors are truncated — we write out only the relevant part and hide the rest into $*$.

Corollary 4.14 (full path components) *If a given presentation ξ does not admit both a fully positive and a fully negative starting vertex, its full path drops out at ξ. Moreover, a full path drops out when it runs into a presentation given by either of the following characteristic covectors $c|_\Gamma$, independently of how the three vectors continue in the hidden $*$–part.*

For some $(i,l) \in \{(1,2),(2,1)\}$,

$$c|_\Gamma = \left(\begin{array}{c|c|c} a^i_0 - 2 & -a^l_0 & \\ a^i_1 - 2 & -a^l_1 + 2 & \\ \vdots & \vdots & \\ a^i_j - 2 & a^l_{k+1} - 2 - 2n^l_{k+1} & a^3_0 - 2 - 2n^3_0 \\ a^i_{j+1} - 2 - 2n^i_{j+1} & a^l_{k+2} - 2 - 2n^l_{k+2} & * \\ * & * & \end{array}\right)$$

for which $-[a^i_0,\dots,a^i_{j-1}]^{-1} - [a^l_0,\dots,a^l_m, N^l_{m+1}]^{-1} < 1$ holds for N^l_{m+1} defined as in Proposition 4.11, or

$$c|_\Gamma = \left(\begin{array}{c|c|c} -a_0^1 & -a_0^2 & \\ -a_1^1+2 & -a_1^2+2 & a_0^3-2 \\ \vdots & \vdots & a_1^3-2 \\ -a_j^1+2 & -a_k^2+2 & * \\ a_{j+1}^1-2-2n_{j+1}^1 & a_{k+1}^2-2-2n_{k+1}^2 & \\ * & * & \end{array}\right)$$

such that for maximal $J \le j$, $K \le k$ with $-[a_0^1,\dots,a_J^1,3]^{-1}-[a_0^2,\dots,a_K^2,2]^{-1}=1$, the denominator of the two fractions is smaller than a_0^3.

In the terminal direction, symmetrically, a drop-out occurs at presentations with oppositely stabilized surgery link (that is, surgery diagrams given by the same but reversely oriented link).

Any two presentations ξ_1 and ξ_2 whose associated characteristic vectors meet at the same path $\mathcal{P}_{\xi_1}=\mathcal{P}_{\xi_2}$ are related by the sequence of rotation number changes, each taking one of the following forms. The pairs are presented in the form of $c|_\Gamma$ and they have to be identical on all further generators, hidden in $*$.

Either for $(i,l)\in\{(1,2),(2,1)\}$,

$$\left(\begin{array}{c|c|c} a_0^i-2 & -a_0^l & \\ a_1^i-2 & -a_1^l+2 & \\ \vdots & \vdots & a_0^3-2-2n_0^3 \\ a_j^i-2 & a_{k+1}^l-2-2n_{k+1}^l & * \\ a_{j+1}^i-2-2n_{j+1}^i & a_{k+2}^l-2-2n_{k+2}^l & \\ * & * & \end{array}\right)$$

$$\simeq \left(\begin{array}{c|c|c} -a_0^i & a_0^l-2 & \\ -a_1^i+2 & a_1^l-2 & \\ \vdots & \vdots & a_0^3-2-2(n_0^3-D) \\ -a_j^i+2 & a_{k+1}^l-2 & * \\ a_{j+1}^i-2n_{j+1}^i & a_{k+2}^l-2n_{k+2}^l & \\ * & * & \end{array}\right),$$

where D is the denominator of $[a_0^l,\dots,a_k^l,n_{k+1}^l+1]^{-1}$, and k and n_{k+1}^l satisfy

$$1=-[a_0^i,\dots,a_{j-1}^i]^{-1}-[a_0^l,\dots,a_k^l,n_{k+1}^l+1]^{-1},$$

or

$$\left(\begin{array}{c|c|c} a_0^1-2-2n_0^1 & a_0^2-2-2n_0^2 & a_0^3-2 \\ * & * & a_1^3-2-2n_1^3 \\ & & * \end{array}\right) \simeq \left(\begin{array}{c|c|c} a_0^1-2n_0^1 & a_0^2-2n_0^2 & -a_0^3 \\ * & * & a_1^3-2n_1^3 \\ & & * \end{array}\right),$$

or

$$\left(\begin{array}{c|c|c} -a_0^1 & -a_0^2 & \\ -a_1^1+2 & -a_1^2+2 & \\ \vdots & \vdots & a_0^3-2 \\ -a_J^1+2 & -a_K^2+2 & a_1^3-2 \\ \vdots & \vdots & a_2^3-2-2n_2^3 \\ -a_j^1+2 & -a_k^2+2 & * \\ a_{j+1}^1-2-2n_{j+1}^1 & a_{k+1}^2-2-2n_{k+1}^2 & \\ * & * & \end{array}\right)$$

$$\simeq \left(\begin{array}{c|c|c} a_0^1-2 & a_0^2-2 & -a_0^3 \\ \vdots & \vdots & -a_1^3+2 \\ a_J^1-2 & a_K^2-2 & a_2^3-2n_2^3 \\ a_{J+1}^1+2-2n_{J+1}^1 & a_{K+1}^2-2n_{K+1}^2 & * \\ * & * & \end{array}\right),$$

where, for $J \le j$, $K \le k$,

$$-[a_0^1,\dots,a_J^1,3]^{-1}-[a_0^2,\dots,a_K^2,2]^{-1}=1,$$
$$n_{J+1}^1 \ge 2 \quad \text{or} \quad L^1=(v_0^1,\dots,v_J^1),$$
$$n_{K+1}^2 \ge 1 \quad \text{or} \quad L^2=(v_0^2,\dots,v_K^2),$$

and the denominator of the two fractions equals a_0^3. □

5 Convex surface theory, overtwistedness and isotopies

To prove that the isotopic classification of tight structures is contained in the full paths of their dual covectors, we need to observe that presentations sharing the same path are indeed isotopic, and relate drop-outs to overtwistedness. This section covers part of the proof outlined in Section 2.2.

To begin with, remember two simple properties of full paths, which have a direct convex-theoretic interpretation. The first is the shuffling property of basic slices within a single continued fraction block [7, Section 4.4.5], which can be, in the Heegaard Floer

interpretation, recovered by 2 PD steps on the consecutive dual vertices of square -2. The second is a necessary condition for tightness, that the presentation contains both a leg starting in a fully positive vertex and a leg starting in a fully negative vertex. In full paths, a fully positive starting vertex is required by Corollary 4.8, and a fully negative one by its analogue when following the path in the terminal direction. In convex surface theory, other presentations can be seen to fail the conditions of the gluing lemma [7, Theorem 4.25], as in [5, Proposition 6.3], but can be also understood as a special case of overtwistedness, proved below.

Let us now state the result as predicted from the Heegaard Floer picture, as in Corollary 4.14. We encode contact presentations into "matrices of negative signs", that is, triples of vectors q^i, possibly of different length, whose coefficients are q^i_j, the number of negative basic slices in the jth continued fraction block of the ith singular fiber. The three vectors in the propositions are truncated, so that we write out only the relevant part (on which overtwistedness is decided, or which behaves nontrivially under isotopy moves) and hide the rest into $*$. Analogously, we can define "matrices of positive signs". Notice that the ones counting negative slices directly correspond to the relations obtained in the initial direction of the full path in Section 4. With positive slices they correspond to symmetric relations in the terminal direction. To describe isotopy moves it is enough to give only pairs of matrices of negative (or only positive) signs, while to encode conditions for overtwistedness, the two are different.

Proposition 5.1 (overtwistedness conditions) *Let a contact presentation be described by either of the following matrices of signs, negative or positive:*

(O1) *For some* $(i,l) \in \{(1,2),(2,1)\}$,

$$(q^i \mid q^l \mid q^3) = \left(\begin{array}{c|c|c} 0 & a^l_0 - 1 & \\ 0 & a^l_1 - 2 & \\ \vdots & \vdots & \\ 0 & a^l_k - 2 & n^3_0 \\ n^i_{j+1} & n^l_{k+1} & * \\ * & n^l_{k+2} & \\ & * & \end{array}\right)$$

for which $-[a^i_0,\dots,a^i_{j-1}]^{-1} - [a^l_0,\dots,a^l_m,N^l_{m+1}]^{-1} < 1$ *holds for* N^l_{m+1} *defined as in Proposition 4.11.*

(O2) *Or*

$$(q^1 \mid q^2 \mid q^3) = \left(\begin{array}{c|c|c} a_0^1-1 & a_0^2-1 & \\ a_1^1-2 & a_1^2-2 & 0 \\ \vdots & \vdots & 0 \\ a_j^1-2 & a_k^2-2 & * \\ n_{j+1}^1 & n_{k+1}^2 & \\ * & * & \end{array}\right)$$

such that for maximal $J \le j$ and $K \le k$ with

$$-[a_0^1,\dots,a_J^1,3]^{-1} - [a_0^2,\dots,a_K^2,2]^{-1} = 1,$$

the denominator of the two fractions is smaller than a_0^3.

Then, independently of the basic slice decompositions of further continued fraction blocks (the $*$–part of vectors)*, the corresponding contact structure is overtwisted.*

Proposition 5.2 (isotopy conditions) *The following pairs of matrices give isotopic contact structures, provided all coefficients are in the range $n_j^i \in [0, a_j^i - 2]$ and $n_0^i \in [0, a_0^i - 1]$, and the further basic slice decompositions* (the $*$–parts) *are the same.*

(I1) *On $(q^i \mid q^l \mid q^3)$, for $(i,l) \in \{(1,2),(2,1)\}$,*

$$\left(\begin{array}{c|c|c} 0 & a_0^l-1 & \\ 0 & a_1^l-2 & \\ \vdots & \vdots & \\ 0 & a_k^l-2 & n_0^3 \\ n_{j+1}^i & n_{k+1}^l & * \\ * & n_{k+2}^l & \\ & * & \end{array}\right) \simeq \left(\begin{array}{c|c|c} a_0^i-1 & 0 & \\ a_1^i-2 & 0 & \\ \vdots & \vdots & \\ a_j^i-2 & 0 & n_0^3-D \\ n_{j+1}^i-1 & n_{k+2}^l-1 & * \\ * & * & \end{array}\right),$$

where D is the denominator of $[a_0^l,\dots,a_k^l,n_{k+1}^l+1]^{-1}$, and k and n_{k+1}^l satisfy

$$1 = -[a_0^i,\dots,a_{j-1}^i]^{-1} - [a_0^l,\dots,a_k^l,n_{k+1}^l+1]^{-1}.$$

(I2) *On $(q^1 \mid q^2 \mid q^3)$,*

$$\left(\begin{array}{c|c|c} n_0^1 & n_0^2 & 0 \\ * & * & n_1^3 \\ & & * \end{array}\right) \simeq \left(\begin{array}{c|c|c} n_0^1-1 & n_0^2-1 & a_0^3-1 \\ * & * & n_1^3-1 \\ & & * \end{array}\right).$$

(I3) *On* $(q^1 \mid q^2 \mid q^3)$,

$$\left(\begin{array}{c|c|c} a_0^1-1 & a_0^2-1 & \\ a_1^1-2 & a_1^2-2 & \\ \vdots & \vdots & 0 \\ a_J^1-2 & a_K^2-2 & 0 \\ \vdots & \vdots & n_2^3 \\ a_j^1-2 & a_k^2-2 & * \\ n_{j+1}^1 & n_{k+1}^2 & \\ * & * & \end{array}\right) \simeq \left(\begin{array}{c|c|c} 0 & 0 & \\ \vdots & \vdots & a_0^3-1 \\ 0 & 0 & a_1^3-2 \\ n_{J+1}^1-2 & n_{K+1}^2-1 & n_2^3-1 \\ * & * & * \end{array}\right),$$

where for $J \le j, K \le k$,

$$\begin{gathered} -[a_0^1,\dots,a_J^1,3]^{-1}-[a_0^2,\dots,a_K^2,2]^{-1}=1, \\ n_{J+1}^1 \ge 2 \quad \text{or} \quad L^1=(v_0^1,\dots,v_J^1), \\ n_{K+1}^2 \ge 1 \quad \text{or} \quad L^2=(v_0^2,\dots,v_K^2), \end{gathered}$$

and the denominator of the two fractions equals a_0^3.

The proof of both propositions is postponed till the end of the section, after a note on contact topological foundations of isotopies and some general computation of slopes.

5.1 State traversals and contact isotopies

Convex decomposition of the Seifert fibrations we are working with consists of the neighborhoods of the three singular fibers F_i and the background circle bundle over the pair of pants. To ensure the product structure in the complement of the F_i we use nonnormalized coefficients $M(0; r_1-1, r_2, r_3)$. The results here rely on Honda's classification of tight structures on separated pieces, namely solid tori [7] and circle bundles over the pair of pants Σ [8, Section 5.1]. In the context of isotopies, we are mostly concerned with the changes of boundary slopes obtained by thickening tubular neighborhoods of singular fibers F_i. The corresponding picture is the 3–punctured sphere with given boundary slopes.

Concretely, we recognize correspondences between presentations relying on the following lemma, due to Ghiggini and Schönenberger.

Lemma 5.3 [6, Lemma 4.13] *Let* Σ *be a pair of pants and* ξ *a tight contact structure on* $\Sigma \times S^1$ *whose boundary* $-\partial(\Sigma \times S^1) = T_1 \cup T_2 \cup T_3$ *consists of tori in standard*

form with $\#\Gamma_{T_i} = 2$ *for* $i = 1, 2, 3$, *and slopes* $s(T_1) = -p_1/q$, $s(T_2) = -p_2/q$ *and* $s(T_3) = \infty$. *Suppose that there exists a pair of pants* $\Sigma' \subset \Sigma$ *such that* $\Sigma \times S^1$ *decomposes as* $\Sigma \times S^1 = \Sigma' \times S^1 \cup C_1 \cup C_2$, *where* $C_i = T_i \times I$, *with* $\xi|_{C_i}$ *minimally twisting, and where* $\xi|_{\Sigma' \times S^1}$ *is a tight contact structure with infinite boundary slopes such that the section* $\Sigma' \times \{\theta\}$ *for some* $\theta \in S^1$ *is convex with dividing set consisting of arcs, each connecting two different boundary components.*

If $s(T_2) = -p_2/q < 0$ *and both* $\xi|_{C_i}$ *decompose into basic slices of the same sign, there exists a convex annulus* A *bounded by the Legendrian rulings of* T_1 *and* T_2, *and without boundary parallel dividing curves.* □

In our case, the decomposition $\Sigma \times S^1 = \Sigma' \times S^1 \cup C_1 \cup C_2$ always exists as we are dealing with the zero-twisting structures; both thickened tori are minimally twisting [8, Lemma 5.1]. The condition on the background structure — no boundary parallel dividing curves on the convex section — follows (as in [5, Lemma 5.4]) from the fact that ξ is appropriate on $\Sigma \times S^1$ (such that there is no embedded $T^2 \times I$ with T^2 isotopic to a boundary component and I–twisting at least π). And the latter is satisfied for any $\Sigma \times S^1$, cut as a background out of tight small Seifert manifold [20, Lemma 2.4]. With the addition of [3, Section 3], Lemma 5.3 can be reformulated in the sense of [5, Lemmas 5.7–5.8].

Lemma 5.4 *Let* Σ *be a pair of pants and let* ξ *be an appropriate contact structure on* $\Sigma \times S^1$ *with convex boundary* $-\partial(\Sigma \times S^1) = T_1 \cup T_2 \cup T_3$ *with* $\#\Gamma_{T_i} = 2$ *for* $i = 1, 2, 3$, *and boundary slopes* $s(T_1) = -p_1/q$, $s(T_2) = -p_2/q$ *and* $s(T_3) = \infty$.

(L1) *If there exists a collar neighborhood* $C_3 \subset \Sigma \times S^1$ *of* T_3, *which is minimally twisting with boundary slopes* ∞ *and* $(p_1 + p_2 - 1)/q$, *whose basic slices are all same-signed, and for which* $\xi|_{(\Sigma \times S^1) - C_3}$ *coincides with the unique tight structure with boundary slopes* $-p_1/q$, $-p_2/q$ *and* $(p_1 + p_2 - 1)/q$, *and maximal twisting* $-q$, *then signs of basic slices in the decomposition of* C_1 *and* C_2 *are all opposite to* C_3*–signs.*

(L2) *And conversely, if* C_1 *and* C_2 *decompose into same-signed basic slices, then there exists* C_3 *composed of opposite-signed slices, with boundary slopes* ∞ *and* $(p_1 + p_2 - 1)/q$, *and such that its complement is a unique tight structure as above.*

Proof (following [5]) For (L1), uniqueness of a tight structure with given properties is stated in [3, Proposition 3.3]. The fact that the signs in the decomposition of collars

are opposite can be read from the relative Euler class evaluation on vertical annuli $A_i \subset C_i$, which have boundaries in vertical Legendrian divides on ∞–side and in Legendrian rulings on the other boundary. If we complete these annuli with annuli through the pair of pants up to T_3 for A_1 and A_2, and up to T_1 and T_2 for A_3, we get two pairs of homologically equivalent, but oppositely oriented, annuli. As the Euler class evaluation on all the extended parts is zero (the first having boundary in Legendrian divides, the second living inside $-q$–maximal twisting), the evaluation on A_1 and A_2 is opposite to that of A_3. Therefore, the collars C_1 and C_2 decompose into basic slices all of the same sign, opposite to the signs in C_3.

For (L2), we take the unique tight $\Sigma'' \times S^1$ (as described in (L1)) and attach to it a thickened torus with slopes $(p_1 + p_2 - 1)/q$ and ∞, and slices signed oppositely to the ones in C_1. Then according to (L1) the signs on collars in decomposition $\Sigma' \times S^1 \cup C_1 \cup C_2$ are again opposite. And we have built up a contact structure, isotopic to the original in all three pieces ($\Sigma' \times S^1$ has the same dividing set on the pair of pants, while C_1 and C_2 have the same Euler class evaluations). □

5.2 Slicing and continued fractions

We give a short reflection on the slopes of glued-up torus and its slicing. Denote by V_i the standard convex neighborhood of F_i with boundary $-\partial(M \setminus V_i)$ trivialized by $\binom{1}{0}$, the horizontal direction of $\Sigma \times 1$, and $\binom{0}{1}$, the direction of fiber, and from the other side ∂V_i by the meridian and some longitude, the last being chosen so that translation $A_i : \partial V_i \to -\partial(M \setminus V_i)$ is given by

$$A_i = \begin{pmatrix} \alpha_i & \alpha_i' \\ -\beta_i & -\beta_i' \end{pmatrix},$$

where $\beta_i/\alpha_i = r_i$ (or $r_1 - 1$ for the first leg); in terms of the continued fraction expansion we have $-\alpha_i/\beta_i = [a_0^i, \dots, a_{k_i}^i]$, $-\alpha_i/\alpha_i' = [a_{k_i}^i, \dots, a_0^i]$ and $-\beta_i/\beta_i' = [a_{k_i}^i, \dots, a_1^i]$. Now, the ∞–slope of a thickened neighborhood U_i of a singular fiber corresponds to $[a_{k_i}^i, \dots, a_0^i]$ in the torus basis, and the slopes of the factorization can be obtained in order (from outside in) by decreasing the last entry of this fraction.

We will be interested in slopes of tori which peel off certain sequences of basic slices from thickened neighborhoods U_i, and their expression in the background basis. Notice the following general behavior:

Lemma 5.5 *The slope of a torus which peels off $\sum_0^{j-1}(a_l^i - 2) + m$ outer basic slices from U_i, as seen from $-\partial(M \setminus V_i)$, is independent of inner continued fraction blocks*

in the decomposition of U_i, that is, of vertices $a^i_j,\dots,a^i_{k_i}$, farther down the legs. It equals $[a^i_0,\dots,a^i_{j-1},m]^{-1}$, or $[a^i_0,\dots,a^i_{j-1},m]^{-1}+1$ in the case $i=1$.

Proof The slope of interest is in the torus basis expressed as $[a^i_{k_i},\dots,a^i_j-m]$.

Recall the matrix form of a negative continued fraction

$$[a^i_{k_i},\dots,a^i_0]\leftrightarrow\left[\begin{pmatrix}-a^i_{k_i}&1\\-1&0\end{pmatrix}^{-1}\cdots\begin{pmatrix}-a^i_0&1\\-1&0\end{pmatrix}^{-1}\right]^2=\left[\begin{pmatrix}-\beta_i'&-\alpha_i'\\\beta_i&\alpha_i\end{pmatrix}\right]^2$$

and notice that it is exactly the inverse of our identification $A\colon \partial V_i\to-\partial(M\setminus V_i)$.

Hence, we get the desired slope in the second column of

$$\begin{pmatrix}-a^i_0&1\\-1&0\end{pmatrix}\cdots\underbrace{\begin{pmatrix}-a^i_{j+1}&1\\-1&0\end{pmatrix}\cdots\begin{pmatrix}-a^i_{k_i}&1\\-1&0\end{pmatrix}\begin{pmatrix}-a^i_{k_i}&1\\-1&0\end{pmatrix}^{-1}\cdots\begin{pmatrix}-a^i_{j+1}&1\\-1&0\end{pmatrix}^{-1}}_{I}\begin{pmatrix}-a^i_j+m&1\\-1&0\end{pmatrix}^{-1}$$
$$=\begin{pmatrix}-a^i_0&1\\-1&0\end{pmatrix}\cdots\begin{pmatrix}-a^i_j&1\\-1&0\end{pmatrix}\begin{pmatrix}0&-1\\1&-a^i_j+m\end{pmatrix}$$
$$=\begin{pmatrix}-a^i_0&1\\-1&0\end{pmatrix}\cdots\begin{pmatrix}-a^i_{j-1}&1\\-1&0\end{pmatrix}\begin{pmatrix}1&m\\0&1\end{pmatrix}.$$

Indeed, this is independent of a^i_ι for $\iota\ge j$.

Now, comparing the second columns,

$$\left[\begin{pmatrix}-a^i_0&1\\-1&0\end{pmatrix}\cdots\begin{pmatrix}-a^i_{j-1}&1\\-1&0\end{pmatrix}\begin{pmatrix}1&m\\0&1\end{pmatrix}\right]^2=:\begin{pmatrix}A\\B\end{pmatrix}$$
$$\leftrightarrow\begin{pmatrix}-B\\-A\end{pmatrix}=\left[\begin{pmatrix}0&-1\\1&-a^i_0\end{pmatrix}\cdots\begin{pmatrix}0&-1\\1&-a^i_{j-1}\end{pmatrix}\begin{pmatrix}0&-1\\1&-m\end{pmatrix}\right]^2,$$

we express the slope as $[a^i_0,\dots,a^i_{j-1},m]^{-1}$, or $[a^i_0,\dots,a^i_{j-1},m]^{-1}+1$ for $i=1$. □

This independence of inner layers allows us to compute the background-basis slope of any sequence of slices (from outside in) on a truncated leg. In the opposite direction, if the slope of peeled-off slices is $[a^i_0,\dots,a^i_{j-1},m]^{-1}$, or $[a^i_0,\dots,a^i_{j-1},m]^{-1}+1$ if $i=1$, this in the torus basis corresponds to $[a^i_k,\dots,a^i_j-m]$ when $m\le a^i_j-1$. When $m=a^i_j$, we get $[a^i_k,\dots,a^i_{j+1},0]$, undefined as a continued fraction, but in terms of the chain of surgeries, the 0–framed meridian cancels a_{j+1}, which results in $[a^i_k,\dots,a^i_{j+2}]$.

5.3 Proofs

The proofs of Propositions 5.1 and 5.2 are stated for matrices of negative signs, but they can be verbatim repeated for positive ones. Without loss of generality, basic slices within each continued fraction block are shuffled so that the negative slices are outer.

Proof of Proposition 5.1 The guiding principle is as follows. Look at the two singular tori U_μ and U_ν whose outermost slices are negative. If we can peel such sequences of negative basic slices from U_μ and U_ν that their inner boundary tori T_μ and T_ν have slopes with the same denominator, say $-p_\mu/q$ and $-p_\nu/q$, we can use Lemma 5.3 to find a torus parallel to ∂U_λ of slope $(p_\mu + p_\nu - 1)/q$, call it T. Whenever this slope is not greater than the critical slope of the singular fiber, $\mathrm{Crit}(F_\lambda)$ (that is, the slope of meridian of the glued-up torus in the background basis), there exists a torus of critical slope between ∂U_λ and T, which proves overtwistedness (see also [12, Section 2]). Furthermore, if the slope of T is such that the thickened torus between ∂U_λ and T (whose basic slices are all positive by (L2)) forms a basic slice together with some of the slices in the original decomposition of U_λ, any negative basic slice in this glued-together basic slice implies overtwistedness by the gluing lemma. Below we analyze the slopes in each of the cases separately.

(O1) Consider first the structures of the first kind with $(i,l) = (1,2)$. Around F_2 there are only negative slices in the first $k+1$ continued fraction blocks and n^2_{k+1} of them in the block corresponding to the vertex v^2_{k+1} (here we shuffle them to be its outer). Around F_3, on the other hand, we have n^3_0 negative slices (again, shuffled so that they are the outer). Thus, peeling off from U_2 basic slices up to the $(N^2_{m+1})^{\text{th}}$ slice of the $-a^2_{m+1}$–block, we obtain the slope $[a^2_0,\dots,a^2_m,N^2_{m+1}]^{-1}$ (in the background basis), which can be joined by cutting the annulus to the torus with slope $-\frac{1}{D}$ around F_3 (peeling off D slices from U_3, where D is the denominator of $[a^2_0,\dots,a^2_m,N^2_{m+1}]^{-1}$, which is at most n^3_0 by assumption). That way, we have found a torus T parallel to T_1 of slope $-[a^2_0,\dots,a^2_m,N^2_{m+1}]^{-1}$.

Now, observe that the critical slope of F_1 satisfies

$$[a^1_0,\dots,a^1_{j-1}-1]^{-1}+1 \le \mathrm{Crit}(F_1) = [a^1_0,\dots,a^1_{k_1}]^{-1}+1 \le [a^1_0,\dots,a^1_{j-1}]^{-1}+1.$$

Our assumed condition gives $-[a^2_0,\dots,a^2_m,N^2_{m+1}]^{-1} < [a^1_0,\dots,a^1_{j-1}]^{-1}+1$. If also a bit more holds true, namely $-[a^2_0,\dots,a^2_m,N^2_{m+1}]^{-1} \le \mathrm{Crit}(F_1)$, the torus T embraces the critical one. Otherwise, we have $-[a^1_0,\dots,a^1_{k_1}]^{-1} - [a^2_0,\dots,a^2_m,N^2_{m+1}]^{-1} \ge 1$ and we can truncate both fractions so that the truncations add up to exactly 1 [9,

Lemma 3.2]. So $-[a_0^1,\dots,a_J^1]^{-1}-[a_0^2,\dots,b_M^2]^{-1}=1$ for some $J\in\{j,\dots,k_1\}$ and $M\in\{0,\dots,m+1\}$, and $b_M^2=a_M^2$ for $M\le m$ or $b_M^2=N_{m+1}^2$ for $M=m+1$. Peeling off from U_2 only slices of the first $M+1$ outer blocks and corresponding (as many as the denominator of $[a_0^2,\dots,b_M^2]^{-1}$, which is certainly less than or equal to $D\le n_0^3$) slices in U_3, the slope of T is $[a_0^1,\dots,a_J^1]^{-1}+1$ in the background basis. By the text under Lemma 5.5 this equals $[a_{k_1}^1,\dots,a_{J+2}^1]$ in torus basis, and T bounds a basic slice with a torus T_1 of slope $[a_{k_1}^1,\dots,a_{J+2}^1-1]$. For tightness, the conditions of the gluing lemma require for all the subslices of a glued-together basic slice to be positive, but this is not satisfied as the toric annulus bounded by T_1 and T contains the $(j+1)^{\text{st}}$ continued fraction block ($J\ge j$) and $n_{j+1}^1>0$ by assumption.

For $(i,l)=(2,1)$, the arguments are the same, but here the induced slope of T (built from peeling-off tori in U_1 and U_3) equals $[a_0^1,\dots,a_m^1,N_{m+1}^1]^{-1}+1$, while the critical slope is given by

$$[a_0^2,\dots,a_{j-1}^2-1]^{-1}\le \mathrm{Crit}(F_2)=[a_0^2,\dots,a_{k_2}^2]^{-1}\le[a_0^2,\dots,a_{j-1}^2]^{-1}.$$

(O2) Structures of the second kind admit negative basic slices in the outer layers of U_1 and U_2. The background-basis slopes $[a_0^1,\dots,a_J^1,3]^{-1}+1$ on T_1 around F_1 and $[a_0^2,\dots,a_K^2,2]^{-1}$ on T_2 around F_2 — which add up to zero — are reached by peeling off the corresponding sequences of (negative) basic slices when v_{J+1}^1 and v_{K+1}^2 exist, and by decreasing the twisting number of the Legendrian singular fibers F_1 or F_2 by stabilizing when $L_1=(v_0^1,\dots,v_J^1)$ or $L_2=(v_0^2,\dots,v_K^2)$. Joining the two tori T_1 and T_2 by an annulus interpolating between the rulings, and edge-rounding, we obtain a torus T around F_3 of slope $-\frac{1}{D}$, where D is the denominator of $[a_0^1,\dots,a_J^1,3]^{-1}$. By assumption, the denominator D is not greater than a_0^3-1, hence the obtained slope $-\frac{1}{D}$ is smaller than or equal to $-1/(a_0^3-1)<\mathrm{Crit}(F_3)$, that is, T embraces the critical torus. □

Proof of Proposition 5.2 Recognition of isotopies, in all cases, follows the same steps. First, we apply (L2) to get an additional thickened torus C around the singular torus U_λ with positive outermost slices — traverse outer layers, whose basic slices are all negative, from U_μ and U_ν. This new collar together with some, say n, continued fraction blocks around F_λ join into a thicker basic slice, separated into positive basic subslices. The isotopy is now given by reversing all these signs. In case we have used all continued fraction blocks of U_λ, it is interpreted in destabilization followed by opposite stabilization of a core knot. Otherwise, C together with $n+1$ outermost continued fraction blocks in U_λ builds a continued fraction block. Its signs can be

shuffled, resulting in the $+2$–change in its innermost $(n+1)^{\text{st}}$ block (one negative slice replaced by positive) and turn of sign on all basic slices that form C and the first n continued fraction blocks (from positive to negative). The basic slices around the other two fibers, F_μ and F_ν, are then adapted according to (L1) — the peeled off ones change their signs from negative to positive, others remain untouched. The relevant slopes for the three isotopy moves are analyzed below.

(I1) Since the equality $-[a^i_0,\dots,a^i_{j-1}]^{-1}-[a^l_0,\dots,a^l_k,n^l_{k+1}+1]^{-1}=1$ holds and n^3_0 is at least as much as the denominator of the two fractions, we can peel off from U_l and U_3 as many negative slices that the slope of the torus T we get around F_i via (L2) equals $[a^i_{k_i},\dots,a^i_{j+1}]$ in the torus basis. The torus T bounds a basic slice with the torus of slope $[a^i_{k_i},\dots,a^i_{j+1}-1]$, which cuts off the positive outermost slices from U_i. The torus of slope $[a^i_{k_1},\dots,a^i_{j+2}]$ then gives a continued fraction block with the torus T.

(I2) These are essentially isotopies from [5, Proposition 6.4]. Peeling off a single (negative outermost) basic slice from U_1 and U_2, we obtain a torus of slope 0, expressed in the background basis, around F_3. It corresponds to the slope $[a^3_{k_3},\dots,a^3_1]$, and forms a glued-together basic slice with the outermost continued fraction block (with inner slope $[a^3_{k_3},\dots,a^3_1-1]$), and hence, a continued fraction block with $[a^3_{k_3},\dots,a^3_2]$.

(I3) In the proof of Proposition 5.1, for structures of the second kind (O2), we have obtained that the slope of the torus T built via (L2) from the two tori of slopes $[a^1_0,\dots,a^1_J,3]^{-1}+1$ around F_1 and $[a^2_0,\dots,a^2_K,2]^{-1}$ around F_2, which sum up to zero, equals $-\frac{1}{D}$ for D the denominator of $[a^1_0,\dots,a^1_J,3]^{-1}$. By assumption, D equals a^3_0; moreover, $-1/a^3_0$ is in the torus basis expressed as $[a^3_{k_3},\dots,a^3_2]$. Thus, the torus T bounds a basic slice with the torus of slope $[a^3_{k_3},\dots,a^3_2-1]$ and a continued fraction block with the torus of slope $[a^3_{k_3},\dots,a^3_3]$ in the slicing of U_3. □

Proof of Theorem 1.3 In Section 2 we have reduced Theorem 1.3 to Theorem 2.2.

Suppose we are given a contact surgery presentation ξ as in Figure 1, whose full path $\mathcal{P}_\xi$ properly ends. By Theorem 2.1 such a presentation describes a tight structure. Furthermore, all presentations which share this same path induce the same Spin^c structure (defining the same class in $\widehat{\mathrm{HF}}(-M)$ [15]) and have the same 3–dimensional invariant (which is, through the equality $d_3(\xi)=d(M,\boldsymbol{t}_\xi)$ of Theorem 1.2, determined by $\boldsymbol{t}_\xi$). In Corollary 4.14 we identify how presentations in the same path are related to each other, and in Proposition 5.2 we realize all these relations by contact isotopies.

Any further isotopies are, of course, excluded by the fact that different paths present nonhomotopic bundles.

On the other hand, if the path $\mathcal{P}_\xi$ drops out (fails the tightness criterion), the corresponding structure ξ admits one of the features recognized in Corollary 4.14 or it can be walked through presentations, related by the above isotopy moves, to some presentation which admits such a feature. Finally, for these structures either the existence of a torus with critical slope or the gluing lemma argument proves their overtwistedness, in Proposition 5.1.

This finishes the proof of Theorem 2.2, originally stated as Theorem 1.3. □

References

[1] **F Ding**, **H Geiges**, **A I Stipsicz**, *Surgery diagrams for contact 3–manifolds*, Turkish J. Math. 28 (2004) 41–74 MR

[2] **Y Eliashberg**, *Contact 3–manifolds twenty years since J Martinet's work*, Ann. Inst. Fourier (Grenoble) 42 (1992) 165–192 MR

[3] **P Ghiggini**, *On tight contact structures with negative maximal twisting number on small Seifert manifolds*, Algebr. Geom. Topol. 8 (2008) 381–396 MR

[4] **P Ghiggini**, **P Lisca**, **A I Stipsicz**, *Classification of tight contact structures on small Seifert 3–manifolds with $e_0 \geq 0$*, Proc. Amer. Math. Soc. 134 (2006) 909–916 MR

[5] **P Ghiggini**, **P Lisca**, **A I Stipsicz**, *Tight contact structures on some small Seifert fibered 3–manifolds*, Amer. J. Math. 129 (2007) 1403–1447 MR

[6] **P Ghiggini**, **S Schönenberger**, *On the classification of tight contact structures*, from "Topology and geometry of manifolds" (G Matić, C McCrory, editors), Proc. Sympos. Pure Math. 71, Amer. Math. Soc., Providence, RI (2003) 121–151 MR

[7] **K Honda**, *On the classification of tight contact structures, I*, Geom. Topol. 4 (2000) 309–368 MR

[8] **K Honda**, *On the classification of tight contact structures, II*, J. Differential Geom. 55 (2000) 83–143 MR

[9] **A G Lecuona**, **P Lisca**, *Stein fillable Seifert fibered 3–manifolds*, Algebr. Geom. Topol. 11 (2011) 625–642 MR

[10] **P Lisca**, **G Matić**, *Tight contact structures and Seiberg–Witten invariants*, Invent. Math. 129 (1997) 509–525 MR

[11] **P Lisca**, **A I Stipsicz**, *Ozsváth–Szabó invariants and tight contact 3–manifolds, III*, J. Symplectic Geom. 5 (2007) 357–384 MR

[12] **P Lisca**, **A I Stipsicz**, *Ozsváth–Szabó invariants and tight contact three-manifolds, II*, J. Differential Geom. 75 (2007) 109–141 MR

[13] **P Lisca**, **A I Stipsicz**, *On the existence of tight contact structures on Seifert fibered 3–manifolds*, Duke Math. J. 148 (2009) 175–209 MR

[14] **I Matkovič**, *Fillability of small Seifert fibered spaces*, preprint (2016) arXiv

[15] **P Ozsváth**, **Z Szabó**, *On the Floer homology of plumbed three-manifolds*, Geom. Topol. 7 (2003) 185–224 MR

[16] **P Ozsváth**, **Z Szabó**, *Heegaard Floer homology and contact structures*, Duke Math. J. 129 (2005) 39–61 MR

[17] **O Plamenevskaya**, *Contact structures with distinct Heegaard Floer invariants*, Math. Res. Lett. 11 (2004) 547–561 MR

[18] **A I Stipsicz**, *Ozsváth–Szabó invariants and 3–dimensional contact topology*, from "Proceedings of the International Congress of Mathematicians, II" (R Bhatia, A Pal, G Rangarajan, V Srinivas, M Vanninathan, editors), Hindustan Book Agency, New Delhi (2010) 1159–1178 MR

[19] **C Wendl**, *Strongly fillable contact manifolds and J –holomorphic foliations*, Duke Math. J. 151 (2010) 337–384 MR

[20] **H Wu**, *Legendrian vertical circles in small Seifert spaces*, Commun. Contemp. Math. 8 (2006) 219–246 MR

Department of Mathematics, Central European University
Budapest, Hungary

matkovic_irena@phd.ceu.edu

Received: 29 January 2016 Revised: 30 March 2017

msp

Algebraic & Geometric Topology 18 (2018) 153–186

An infinite family of links with critical bridge spheres

DANIEL RODMAN

A closed orientable splitting surface in an oriented 3–manifold is a topologically minimal surface of index n if its associated disk complex is $(n-2)$–connected but not $(n-1)$–connected. A critical surface is a topologically minimal surface of index 2. In this paper, we use an equivalent combinatorial definition of critical surfaces to construct the first known critical bridge spheres for nontrivial links.

57M25

1 Introduction

In the 1960s Haken developed a framework for studying manifolds that contain incompressible surfaces. He demonstrated that for such manifolds, one can reduce the proofs of many theorems to induction arguments by using "hierarchies", consecutively cutting a manifold into pieces along incompressible surfaces until a collection of balls is obtained. This powerful approach clearly demonstrated the utility of incompressible surfaces in the study of low-dimensional topology.

In 1987 Casson and Gordon [4] introduced the idea of strongly irreducible surfaces. Unlike incompressible surfaces, which have no compressing disks, strongly irreducible surfaces have potentially many compressing disks to both sides, but any two compressing disks on opposite sides necessarily intersect. Somewhat surprisingly, it turns out that many theorems that are easy to prove for manifolds with incompressible surfaces also hold true for manifolds with strongly irreducible surfaces, although the proofs can be somewhat more involved. Casson and Gordon's watershed result in the theory of strongly irreducible surfaces is that if a closed, connected, orientable 3–manifold's minimal-genus Heegaard splitting is not strongly irreducible, then the manifold must contain an essential surface [4].

In 2002, Bachman [1] introduced the notion of a critical surface, which he followed in 2009 [2] with the introduction of the more general concept of a topologically minimal surface. A closed orientable splitting surface in an oriented 3–manifold M

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.153

is a topologically minimal surface of index n if its associated disk complex is $(n-2)$–connected but not $(n-1)$–connected. This latter definition provides the framework into which incompressible, strongly irreducible, and critical surfaces all fit: they are topologically minimal surfaces of indices 0, 1 and 2 respectively. Just like the incompressible and strongly irreducible surfaces before them, critical surfaces and topologically minimal surfaces in general have been used to prove long-standing conjectures that had remained unresolved for many years. For example, Bachman [3] used these surfaces to prove the Gordon conjecture[1] and to provide counterexamples to the stabilization conjecture.[2]

Most of the work done thus far regarding topologically minimal surfaces deals specifically with surfaces which are Heegaard splittings, two-sided surfaces that split a manifold into two compression bodies. Less is known about topologically minimal bridge surfaces, which are a natural type of surface to consider when the 3–manifold is a link complement. Lee has shown in [8] that all bridge spheres for the unknot with any number of bridges in S^3 are topologically minimal, and his results provide upper bounds for the indices of such bridge spheres. In particular, he concludes that the bridge sphere for an unknot in a 3–bridge position has index exactly 2 (ie it is critical).

In this paper we provide the first known examples of critical bridge spheres for nontrivial links. Our construction is inspired by recent work of Johnson and Moriah [6] in which they construct links with bridge surfaces of arbitrarily high distance. The central result of this paper is the following:

Theorem 6.9 *There is an infinite family of nontrivial links with critical bridge spheres.*

In Section 2 we go over some of the foundational topological definitions upon which this paper relies, including an equivalent, combinatorial definition of a critical surface. Then in Section 2 we describe what it means for a link to be in a plat position. Following that, in Section 3 we embed a link L in S^3 in a plat position with bridge sphere F and discuss some of the specific details of the embedding as well as build some of the tools (certain arcs, loops, disks and projection maps) which we will use throughout the rest of the paper. Sections 2 and 3 should be considered setup for the rest of the paper, and this is essentially the same setup as in Johnson and Moriah's paper [6]. In particular, we make use of Johnson and Moriah's plat links and projection maps. In Section 4 we

[1] Gordon conjecture: if the Heegaard splitting $U \cup_H V$ is a connected sum of two Heegaard splittings, and H is stabilized, then one of its summands is stabilized.

[2] Stabilization conjecture: given any pair of Heegaard splittings, stabilizing the higher-genus splitting once results in a stabilization of the other splitting.

develop a link-diagrammatic way to visualize boundary loops of compressing disks for F. Theorem 6.9 is proved in Sections 5 and 6.

2 Definitions

Suppose Σ is a compact orientable surface embedded in a compact orientable 3–manifold M, and D is a disk embedded in M (not necessarily properly). We say that D is a *compressing disk* for Σ if $D \cap \Sigma = \partial D$ and ∂D neither bounds a disk in Σ nor is parallel to a boundary component of Σ.

A compact, orientable surface Σ embedded in a connected 3–manifold M is a *splitting surface* if $M \backslash \Sigma$ has two components. We associate a simplicial complex Γ called the *disk complex* to (M, Σ) in the following way: Vertices of Γ are isotopy classes of compressing disks for Σ. A set of $m+1$ vertices will be filled in with an m–simplex if and only if the corresponding isotopy classes of compressing disks have pairwise disjoint representatives. For $n \geq 0$, Σ is called *topologically minimal of index n* if Γ is $(n-2)$–connected but not $(n-1)$–connected. (Note that every space is (-2)–connected, and a space is (-1)–connected if and only if it is nonempty.) Σ is called a *critical surface* if it is a topologically minimal surface of index 2. In [2], Bachman gives an alternative, combinatorial definition for critical surfaces and proves the two definitions are equivalent. This second definition, given below, is the one we will use in this paper.

A *critical surface* is a splitting surface $\Sigma \subset M$ whose isotopy classes of compressing disks can be partitioned into $\mathcal{C}_1 \sqcup \mathcal{C}_2$ in a way that satisfies the following two conditions:

(1) Whenever $[C] \in \mathcal{C}_1$ and $[D] \in \mathcal{C}_2$ for C and D on opposite sides of Σ, $\partial C \cap \partial D$ must be nonempty.

(2) For $i = 1, 2$, there exists a pair of disjoint compressing disks, one on either side of Σ, where each disk belongs to an isotopy class in $\mathcal{C}_i$.

Any link $L \subset S^3$ can be isotoped so that all of its maxima lie above all of its minima (with respect to the standard height function on S^3). After such an isotopy, L is said to be in *bridge position*. Let $\eta(L)$ be an open regular neighborhood of L, and let F be a level sphere in S^3 which separates all of the maxima of L from all of the minima. Then the surface $F' = F \backslash \eta(L) \subset M = S^3 \backslash \eta(L)$ (a sphere with a finite number of open disks removed) is called a *bridge sphere* for L. See Figure 2.

In our context, studying the bridge sphere $F' \subset S^3 \backslash \eta(L)$ is equivalent to studying the "sphere $F \subset S^3$ with marked points", where the *marked points* are the points of $L \cap F$.

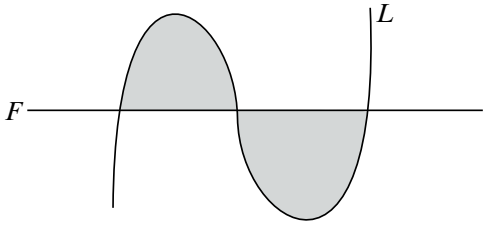


Figure 1: If a bridge disk on one side of F intersects a bridge disk below F in a single point of the link, and are disjoint otherwise, then the link is perturbed.

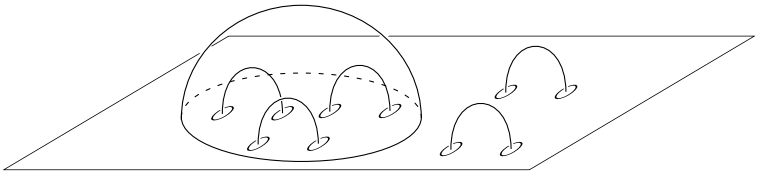


Figure 2: A bridge sphere and a compressing disk

A disk in S^3 is considered a compressing disk for the sphere F with marked points if and only if it is a compressing disk for F' in M. This implies that a compressing disk for F cannot intersect L, and two compressing disks for F are considered to be in the same isotopy class if and only if they are isotopic in $S^3/\eta(L)$.

A *bridge arc* is a component of $L\backslash F$, and each of these bridge arcs has exactly one critical point, so each is parallel into F. An isotopy of a bridge arc into F sweeps out a disk, called a *bridge disk*. (Bridge disks are not unique, even up to isotopy.)

A link is *perturbed* if there is a bridge disk above F and another bridge disk below F which intersect in a point of L, and which are otherwise disjoint. See Figure 1. We will say the link is *perturbed at* a bridge arc α if α corresponds to either of these two bridge disks. If a link L is perturbed, then there is an isotopy of L through the bridge disks which reduces the number of critical points of L by one maximum and one minimum.

A *cap* for F is a compressing disk C such that ∂C bounds a disk in F that contains exactly two marked points. The existence of caps is guaranteed by the existence of bridge disks. See Figure 3.

The notion of boundary-compression along boundary-compressing disks will also be useful. Suppose Σ is a surface with boundary properly embedded in a 3–manifold M.

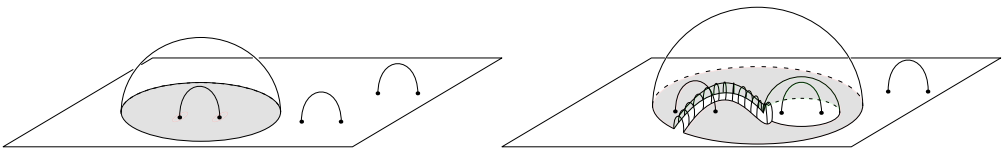


Figure 3: The two disks depicted are both caps for the leftmost bridge arc. In each figure, the disk in F with two marked points is shaded.

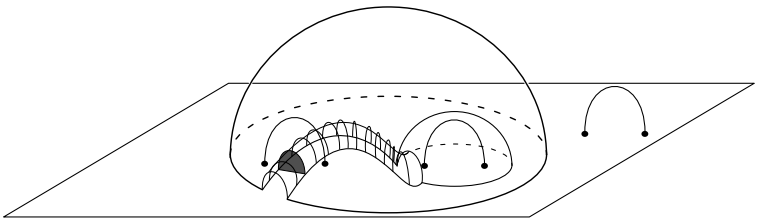


Figure 4: A boundary-compressing disk is shaded dark gray.

A *boundary-compressing disk* (See Figure 4 for an example) is defined to be an embedded disk D with the following properties:

- $D \cap \partial M$ is an arc β.
- $D \cap \Sigma$ is an arc $\widetilde{\beta}$.[3]
- $\partial D = \beta \cup \widetilde{\beta}$.
- $\beta \cap \widetilde{\beta} = \partial\beta = \partial\widetilde{\beta}$.

Next we define a half-twist. Let U be a twice punctured (marked) disk. The half-twist can be simply understood to be the homeomorphism $U \to U$ depicted in Figure 5. To be technical, the *half-twist* H of U is the homeomorphism $H\colon U \to U$ defined as follows: View U as a union $A \cup_\phi D$ of the closed unit disk $D \subset \mathbb{C}$ with an outer annulus A whose intersection is an embedded circle containing the two marked points. See Figure 5. Give D the coordinates $re^{i\theta}$, and identify A with $S^1 \times I$, with the coordinates (θ, t) (here A is glued to D by the map ϕ taking $(\theta, 0)$ in one component of ∂A to $e^{i\theta}$ in ∂D). Now define

$$H(p) = \begin{cases} (\theta + \pi t, t) & \text{for } p = (\theta, t) \in A, \\ re^{i(\theta+\pi)} & \text{for } p = re^{i\theta} \in D. \end{cases}$$

[3]Our definition here is looser than the standard definition since we do not require $\widetilde{\beta}$ to be an *essential* arc in Σ. We drop this condition because we will want to perform boundary-compressions on disks, which have no essential arcs.

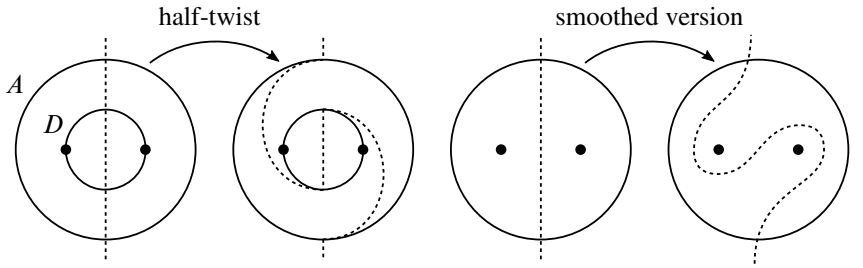


Figure 5: Half-twist

For any punctured (marked) surface S with more than one puncture (marked point), if $l \subset S$ is a loop which bounds an embedded twice-punctured (marked) disk $U_l \subset S$, we can identify U_l with U and define the *half-twist* H_l *about* l to be the homeomorphism $H_l\colon S \to S$ given by

$$H_l(p) = \begin{cases} H(p) & \text{for } p \in U, \\ p & \text{for } p \notin U. \end{cases}$$

Now we will define plat position for a link. We start by describing plat position for a braid. Fix integers $h \geq 2$ and $b \geq 3$. We begin by constructing a set of vertical cylinders in $\mathbb{R}^3$ which will provide a frame for our braid. See Figure 6. In the xy–plane, define $c^1, c^2, \ldots, c^{2b-1}$ to be the circles of radius $\frac{1}{2}$ such that c^j has center $\left(j + \frac{1}{2}, 0, 0\right)$. Now define Cyl^j to be the cylinder obtained by crossing c^j with the z–interval $[1, h]$.

Next we define the *twist regions*, subcylinders of $\{\mathrm{Cyl}^j\}$ depicted in Figure 6. Let $i \in \{2, 3, \ldots, h\}$. For odd i, let j range from 1 to b; for even i, let j range from 1 to $b-1$. Then define

$$\mathrm{Tw}_i^j = \begin{cases} c^{2j-1} \times \left[i - \frac{3}{4}, i - \frac{1}{4}\right] & \text{for } i \text{ odd}, \\ c^{2j} \times \left[i - \frac{3}{4}, i - \frac{1}{4}\right] & \text{for } i \text{ even}. \end{cases}$$

In this paper, superscripts will usually denote an object's horizontal position (ie in the x–direction), and subscripts will usually denote an object's vertical position (ie in the z–direction). This is reminiscent of typical matrix notation: an entry a_i^j of a matrix is in the ith position vertically and the jth position horizontally. Here, however, unlike in a matrix, an array of objects is enumerated from bottom to top instead of from top to bottom.

Now we define a (nonclosed) braid B to be in (h, b)*–plat position* if it satisfies the following conditions:

(1) B is a $2b$–strand braid embedded in $\bigcup \mathrm{Cyl}^j$ whose bottom endpoints' coordinates are $(j, 0, 1)$ for $j = 1, 2, \ldots, 2b$ and whose top endpoints' coordinates are $(j, 0, h)$ for $j = 1, 2, \ldots, 2b$.

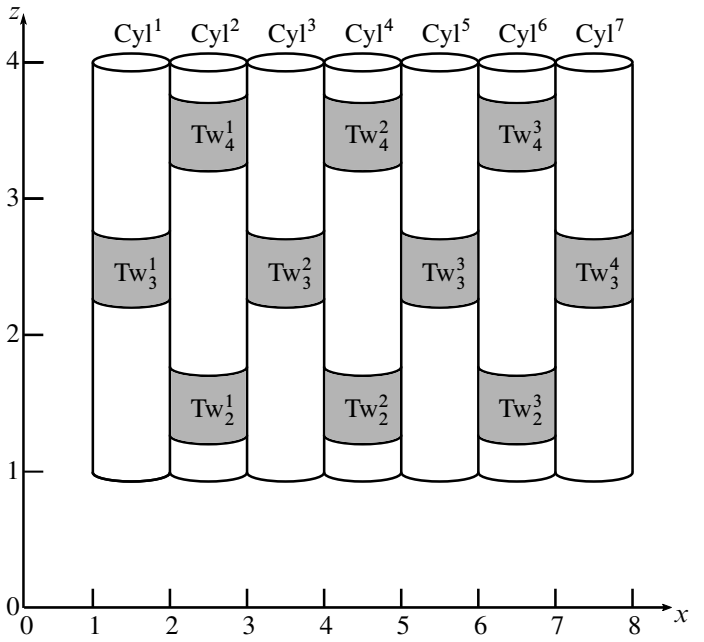


Figure 6: The cylinders which form a frame for plat links (here $h = b = 4$)

(2) Arcs of B are partitioned into two types: vertical arcs which lie in the intersection of $\bigcup \mathrm{Cyl}^j$ with the xz–plane, and twisting arcs which are not contained in the xz–plane but are properly embedded in the twist regions.

(3) Every twisting arc is strictly increasing in a twist region, and as it ascends, it moves either strictly clockwise around the cylinder or strictly counterclockwise.

These conditions guarantee that each twist region contains exactly two arcs of B. Observe that the endpoints of every twisting arc must lie in the xz–plane. Since twisting arcs move strictly clockwise or counterclockwise, they intersect the xz–plane minimally. That is, there is no isotopy of a twisting arc in a twist region, relative to its endpoints, which would decrease the intersection of that twisting arc and the xz–plane. Looking down at the link from above, if the twisting arc moves counterclockwise as it ascends, we define t to be -1 plus the number of times the twisting arc intersects the xz–plane. Similarly, if the twisting arc moves clockwise as it ascends, we define $-t$ to be -1 plus the number of times the twisting arc intersects the xz–plane. In other words, to each twisting arc we assign an integer t such that the twisting arc travels around the twist region through an angle of $t\pi$. Notice that the other twisting arc in the same twist region must twist through an equal angle $t\pi$ (or else the two arcs would

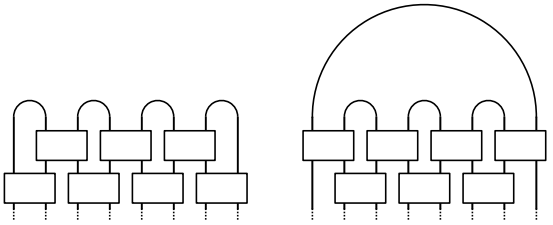


Figure 7: Closing a braid in (h,b)–plat position, for (left) h even, and (right) h odd. In this paper, we only consider $h=4$.

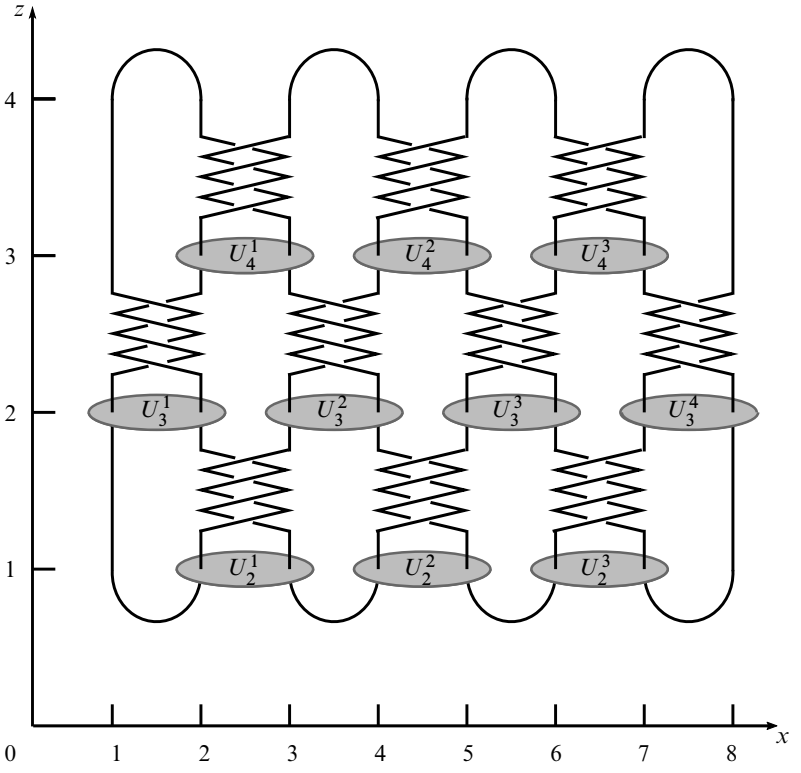


Figure 8: A 4–twisted $(4,4)$–plat link with $t_1^j = t_3^j = 4$ for all j and $t_2^j = -4$ for all j. Also pictured are the U–disks defined in Section 3.

intersect). Thus we can define the number t to be the *twist number* corresponding to that twist region. We will use t_i^j to denote the twist number for Tw_i^j.

There is a standard way to create a link from a braid in plat position. See Figure 7. Along the bottom of B, we have a row of $2b$ endpoints at a height of $z=1$. For each odd j between 1 and $2b$, connect the endpoints $(j,0,1)$ and $(j+1,0,1)$ with a lower semicircle in the xz–plane. The way we connect the top endpoints of B depends

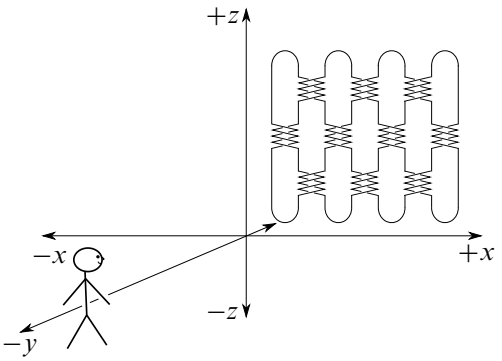


Figure 9: Point of view

on the parity of h. Notice that if h is even, then the highest row of twist regions consists of $b-1$ twist regions. Along the top of B we have a row of $2b$ endpoints at a height of $z=h$. For each odd j between 1 and $2b$, connect the endpoints $(j,0,h)$ and $(j+1,0,h)$ with an upper semicircle in the xz–plane. If h is odd, there are b twist regions at the top. For each even j between 2 and $2b-2$, connect the endpoints $(j,0,h)$ and $(j+1,0,h)$ with an upper semicircle in the xz–plane. Then connect the points $(j,0,1)$ and $(j,0,2b)$ with a larger upper semicircle in the xy–plane. See Figure 7. Thus we obtain a link from B.

Any link constructed in this way from a braid in (h,b)–plat position will be said to be *a link in (h,b)–plat position*. Observe that such a link is in a bridge position with bridge number b. For $n\in\mathbb{N}$, a link in (h,b)–plat position is called *n–twisted* if $|t_i^j|\geq n$ for all i,j. See Figure 8 for an example of a 4–twisted link in $(4,4)$–plat position.

3 Setting

Let $f\colon S^3\to[-\infty,\infty]$ be a height function on S^3, and let Λ be a strictly increasing arc in S^3 whose endpoints are the two critical points of S^3. Then $S^3\backslash\Lambda$ is homeomorphic to $\mathbb{R}^3$. Let $\phi\colon\mathbb{R}^3\to S^3\backslash\Lambda$ be a homeomorphism that respects the height function f on S^3 and the standard height function on $\mathbb{R}^3$. That is, if P_i is the level plane in $\mathbb{R}^3$ of height i, and F_i is the level sphere in S^3 of height i, then $\phi(P_i)$ is F_i minus the point $F_i\cap\Lambda$. Throughout the rest of the paper, we adopt the point of view of someone standing on the $-y$ side of the xz–plane. This gives meaning to words like "left", "right", "up", "down", "horizontal" and "vertical". See Figure 9.

The link L Let L be a 2–twisted link in $(4,4)$–plat position with bridge sphere $F\simeq F_1=f^{-1}(1)\subset S^3$. Suppose L has twist numbers $\{t_i^j\}$ such that for all j, the

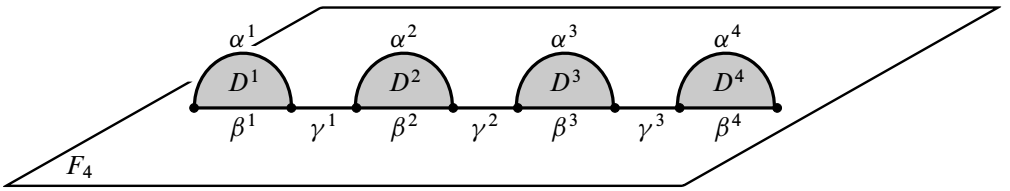


Figure 10: The bridge disks D^k above F

Figure 11: The bridge arcs below F. Of these, we care especially about the rightmost bridge arc $\alpha^{4'}$, the straight line segment β' between its endpoints, and the bridge disk $D^{4'}$ cobounded by $\alpha^{4'}$ and β'.

twist numbers t_2^j, t_4^j are positive, and t_3^j is negative. We will work with this link L for the rest of the paper.

Consider the diagram $D(L)$ for L obtained by projection to the xz–plane. Our choice of signs for the twist numbers $\{t_i^j\}$ makes $D(L)$ an alternating diagram, and we know that L is a split link if and only if $D(L)$ is a split diagram (see Theorem 4.2 of [9]). Since $D(L)$ is clearly a nonsplit diagram, it follows that L is a nonsplit link.

The bridge disks Refer to Figure 10. Above F_4 in the xz–plane lie the four upper bridge arcs. We will name these arcs consecutively from left to right α^1, α^2, α^3, α^4. Vertical projection of the bridge arcs into F gives us four straight line segments at level 4, which we will name consecutively from left to right β^1, β^2, β^3, β^4. We will refer to these as *β–arcs*. Observe that α^j and β^j form a loop. Let D^j be the disk in the xz–plane bounded by this loop. Notice each D^j is a bridge disk. Define γ^j to be the straight line segment connecting the points $(2j, 0, 4)$ and $(2j+1, 0, 4)$. We will refer to these arcs as *γ–arcs*.

Below F_1 in the xz–plane lie the four lower bridge arcs, which we will call consecutively from left to right $\alpha^{1'}$, $\alpha^{2'}$, $\alpha^{3'}$, $\alpha^{4'}$. We will also define β' to be the straight line segment between the endpoints of $\alpha^{4'}$. See Figure 11. The union $\beta' \cup \alpha^{4'}$ is a loop bounding a bridge disk in the xz–plane, which we will call $D^{4'}$.

***l*–loops and *U*–disks** For $i = 1, 3$, define l_i^j to be the circle of radius $\frac{3}{4}$ in the horizontal plane $z = i - 1$ centered at the point $\left(2j + \frac{1}{2}, 0, i - 1\right)$. For $i = 2$, define l_2^j to be the circle of radius $\frac{3}{4}$ in the horizontal plane $z = 2$ centered at the point

$\left(2j-\frac{1}{2},0,i-1\right)$. Then each l_i^j bounds a twice-punctured disk which we call U_i^j. Each disk U_i^j lies below a corresponding twist region Tw_i^j. See Figure 8.

σ–projection It will be convenient to be able to talk about simple vertical projections. If P_i denotes the horizontal plane in $\mathbb{R}^3$ of height i, let $\sigma_i\colon \mathbb{R}^3 \to P_i$ be the vertical projection map defined by $(x,y,z)\mapsto(x,y,i)$.

τ–projection We will also need a more subtle type of projection which respects L. We closely follow Johnson and Moriah [6]. Note that L intersects each F_i in the same number of points and these points vary continuously as i varies from 1 to h. This can be thought of as an isotopy of these $2b$ marked points in F which extends to an ambient isotopy of F_h. To be precise, there is a projection map $\tau_i\colon F\times[1,h]\to F_i$ for each $i\in[1,h]$ that sends each arc component of the plat braid to a point $(j,0,i)$ for some $j=1,2,\dots,2b$ and defines a homeomorphism $F_{i'}\to F_i$ for each $i'\in[1,h]$. These homeomorphisms are canonical up to isotopy fixing the points $L\cap F_i$, and the induced homeomorphism $F_i\to F_i$ is the identity. Further, for each $i=2,3,4$, the homeomorphism induced by $\tau_i\colon F_{i-1}\to F_i$ is a composition of half-twists about the l–loops, which can be expressed as the product $\prod_j H_{l_i^j}^{t_i^j}$. (Note that since the U–disks at any given level are pairwise disjoint, the corresponding half-twists all commute with each other.) Therefore the homeomorphism $g\colon F_0\to F_4$ induced by τ_4 can be expressed as

$$(*)\qquad g=\Big(\prod_j H_{l_4^j}^{t_4^j}\Big)\Big(\prod_j H_{l_3^j}^{t_3^j}\Big)\Big(\prod_j H_{l_2^j}^{t_2^j}\Big).$$

Disk partition Our goal is to show that F is critical, so we need to exhibit a partition $\mathcal{C}_1\sqcup\mathcal{C}_2$ of the isotopy classes of the compressing disks for F. (Recall the definition of "critical" from the beginning of Section 2.) Above F, let B be the frontier of a regular neighborhood of D^1. Below F, let B' be the frontier of a regular neighborhood of $D^{4'}$. B and B' are depicted in Figure 12. Let $[B]$, $[B']\in\mathcal{C}_2$, and let all of the other isotopy classes of compressing disks be in $\mathcal{C}_1$. It will be convenient to refer to a compressing disk C as *red* if $[C]\in\mathcal{C}_1$ and *blue* if $[C]\in\mathcal{C}_2$.

4 The labyrinth

We are interested in the image of $\partial B'$ under g, the homeomorphism induced by τ_4. As τ_4 takes $\partial B'$ up from level 1 to level 4, $\partial B'$ undergoes a series of twists, following the strands of the link. As explained in Section 3, g can be expressed as a product of

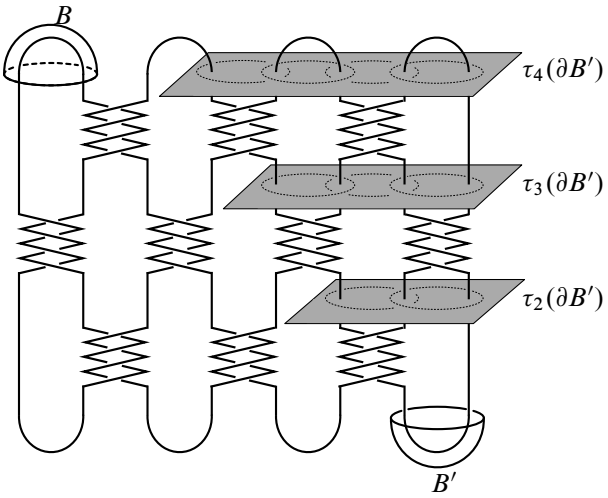


Figure 12: The two blue compressing disks, B and B'

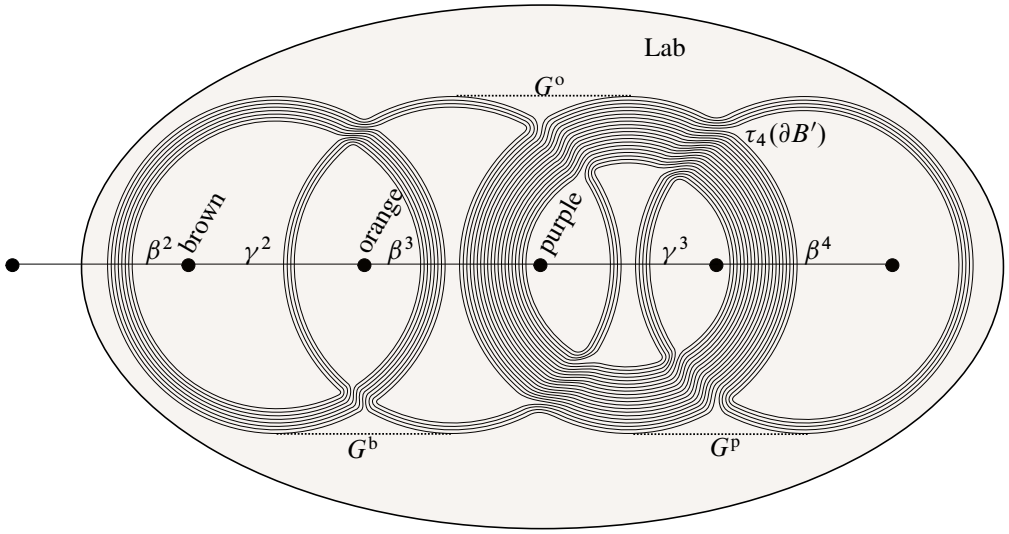


Figure 13: The thin spiraling line is $\tau_4(\partial B')$ in the case when $t_1^j = t_3^j = 2$ and $t_2^j = -2$ for all j. Also depicted is the disk Lab which contains $\tau_4(\partial B')$, and the three gates, G^{b}, G^{o}, and G^{p}, as well as the three colored points. In general, the five marked points in Lab may be permuted differently when different twist numbers are chosen.

half-twists around the l–loops via the equation $(*)$. In Figure 13 we show exactly what $\partial B'$ looks like under g in a case where $|t_i^j| = 2$ for all j.

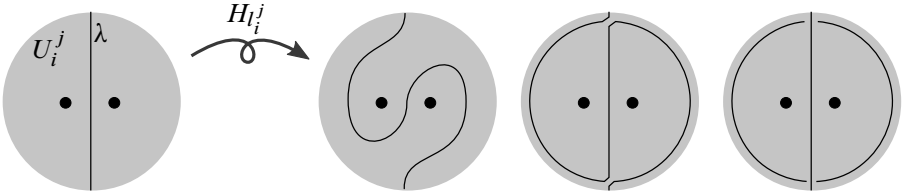


Figure 14: How to view a half-twist about an l–loop as a link diagram

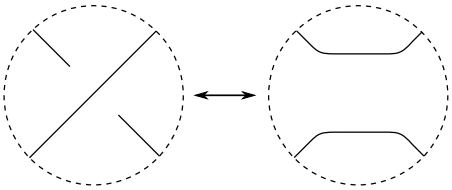


Figure 15: Rule for smoothing a crossing

Of course, we want to understand what $\tau_4(\partial B')$ will look like in general for any set of twist numbers $\{t_i^j\}$. To do so, we will devise a convenient pictorial notation for loops in F which undergo half-twists about l–loops. Suppose λ is a loop in F with one arc of intersection with U_i^j which passes between the two marked points, as in the first picture in Figure 14. The second and third pictures in Figure 14 are homotopic pictures of $H_{l_i^j}(\lambda)$. Observe that the third picture is reminiscent of knot diagram smoothing. In fact, using the smoothing rule in Figure 15, we can "unsmooth" $H_{l_i^j}(\lambda)$ into the fourth picture, which consists of a link diagram on F with two components: (1) the original λ and (2) l_i^j. This suggests a convenient algorithm to draw any such loop λ after a (positive) half-twist about an l–loop: We start by drawing λ. Next we draw the l–loop so that it passes *under* λ. Last we smooth the crossings using the rule in Figure 15. If we want to perform a negative half-twist instead, the procedure is the same, except that we draw the l–loop crossing *over* λ.

We can generalize this process to the situation where there are n strands of λ passing through U_i^j. In Figure 16, the top row is an example in which $n = 3$, and the bottom row depicts the general case. As in Figure 14, the second and third pictures of each row in Figure 16 are homotopic pictures of $H_{l_i^j}(\lambda)$, and we can "unsmooth" $H_{l_i^j}(\lambda)$ into the fourth picture, which consists of a link diagram on F. This time, the link diagram contains $n+1$ components: λ and n parallel copies of l_i^j. Thus our generalized procedure for drawing a half-twist of λ about l_i^j is the following: First determine the number n of times that λ passes between the two punctures of U_i^j, then draw n parallel, disjoint copies of l_i^j so that at each crossing, l_i^j passes *under* λ for a positive

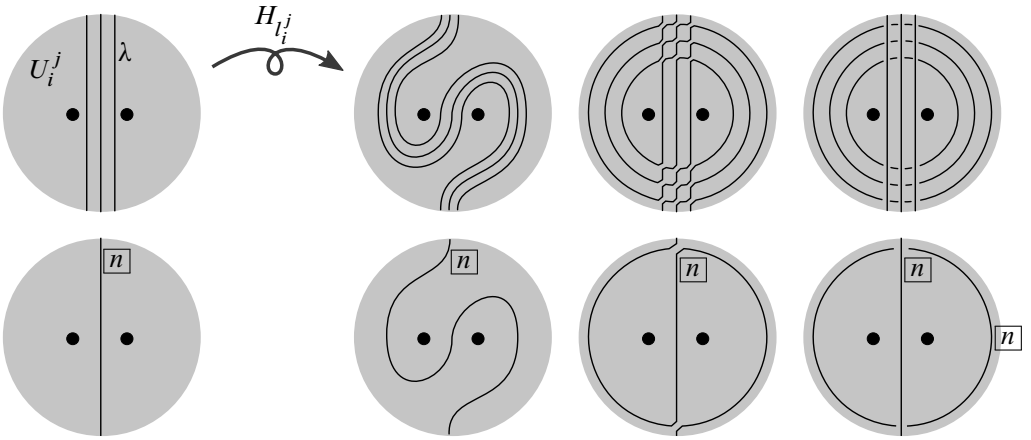


Figure 16: How to view a half-twist about an l–loop as a link diagram. In the top row of pictures, we show specifically what the situation looks like when $n = 3$. The bottom row shows that the whole process can be done with an arbitrary number n of strands through U_i^j. In all the following figures that depict half-twists, when we label the number of parallel strands present, as in the bottom row here, we will box the labels with squares.

half-twist, or *over* λ for a negative half-twist, and then smooth the crossings using the rule from Figure 15.

Note that to perform $|t_i^j|$ consecutive half-twists around l_i^j, we simply perform this process $|t_i^j|$ times, taking care that every time we add new parallel copies of l_i^j, they are nested inside the ones previously drawn. Thus if λ is half-twisted around l_i^j a total of $|t_i^j|$ times, and λ passes between the punctures of U_i^j a total of n times, then to build the link diagram depicting $H_{l_i^j}^{t_i^j}(\lambda)$, we will add a total of $n|t_i^j|$ parallel copies of l_i^j (all passing under λ if $t_i^j > 0$ or over λ if $t_i^j < 0$). Thus after any sequence of half-twists about l–loops, we obtain a link diagram representing λ. Instead of drawing all of the parallel copies of each l–loop, we will draw only one of each and label it with the number $\boxed{n}$ of strands it represents.

In the resulting link diagram, we will have crossings involving two strands which are labeled with different numbers. Figure 17 depicts an example in which we show how to recover what λ actually looks like near such a crossing, where a strand marked with a $\boxed{7}$ crosses a strand marked with a $\boxed{3}$. Figure 18 shows the general case.

Our purpose in developing this link-diagrammatic representation of a twisted loop is to understand what $\partial B'$ looks like after performing (in order) half-twists about all of the l–loops for any set of twist numbers $\{t_i^j\}$. With this method in hand, this is a

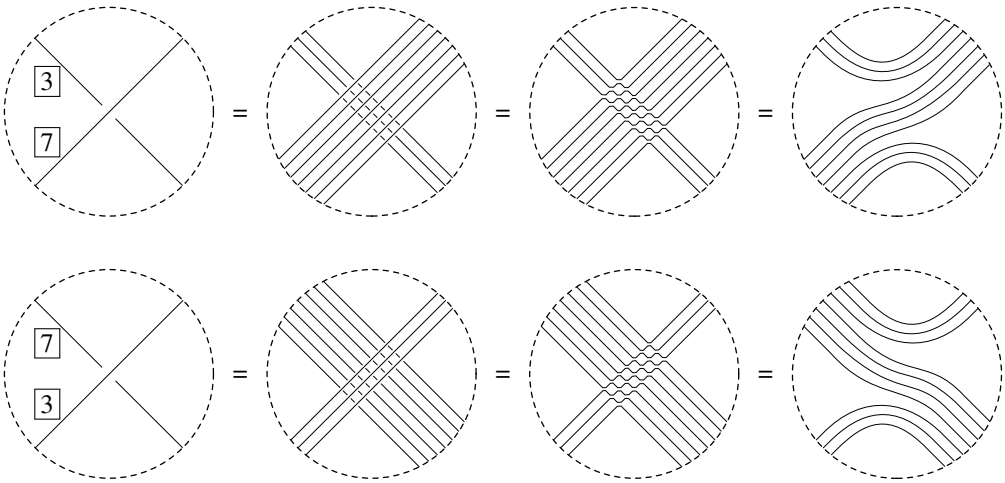


Figure 17: After twisting a loop λ about l–loops, we obtain a link diagram on F. Here is an example of a crossing in such a diagram: In this crossing, one strand represents three parallel strands, and the other represents seven parallel strands. Consider the top row of pictures, where the over-strand has a larger number. The first picture is a neighborhood of the crossing of the link diagram we have obtained. The second picture explicitly shows all of the parallel strands at the crossing. The third picture shows the result after smoothing all 21 crossings. The fourth picture is a smoothed version of the third. All four pictures represent the same thing. The bottom row is similar, but this time the under-strand has the larger number.

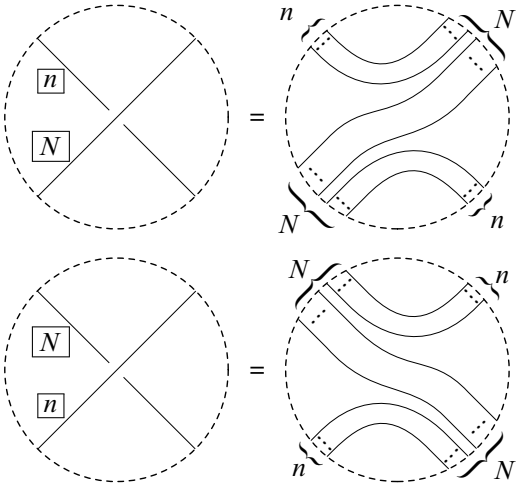


Figure 18: How to interpret a crossing in the link diagram of N strands with n strands, where $N > n$

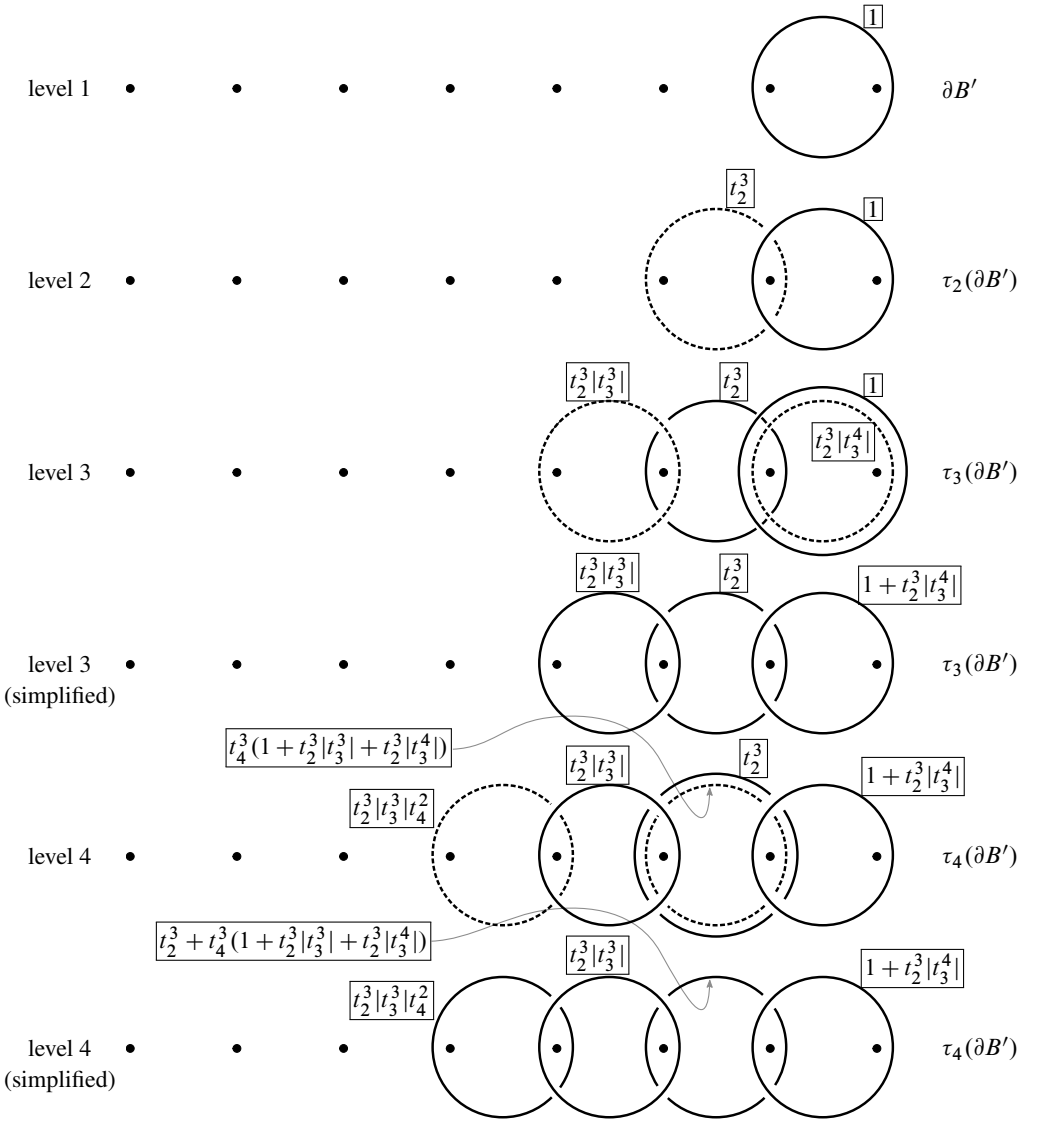


Figure 19: The components of the link diagram that are newly added as a result of having just moved up a level are highlighted by being drawn with dotted lines. The reader may find it helpful to compare this figure to Figure 12, which shows the link diagram's location with respect to L.

straightforward task. Refer to Figure 19 as we describe this process. We start with the simple circle $\partial B'$ in F at level 1, and then push it up to level 2 (ie we apply τ_2). The boundary $\partial B'$ had to pass through the twist region Tw_2^3, so this means it underwent t_2^3

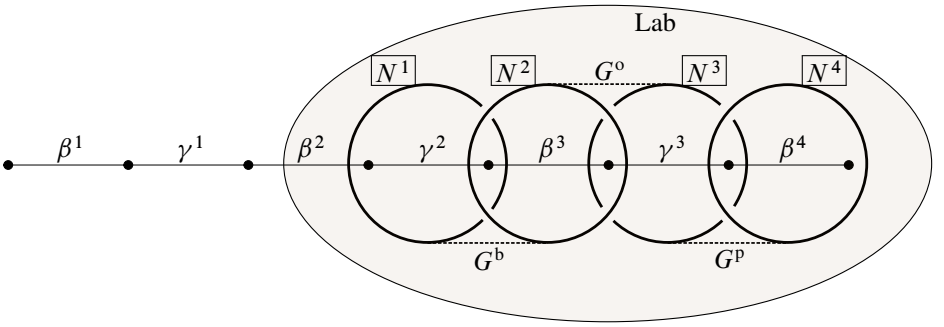


Figure 20: The link diagram representing $\tau_4(\partial B')$ (with values of the N^i and inequalities between them as given below)

half-twists about l_2^3. Before twisting (ie at level 1), $\partial B'$ passed one time (ie $n = 1$) between the punctures of U_2^3; therefore what $\partial B'$ looks like at level 2 is a link diagram consisting of the original circle $\partial B'$, plus $1 \times t_2^3$ parallel copies of l_i^j. Since we specified in Section 3 that t_2^j is positive for all j, each copy of l_i^j will be drawn to pass under $\partial B'$. Now the original $\partial B'$ and the t_2^3 copies of l_i^j together with the crossing information make a link diagram which represents the image of $\partial B'$ under τ_2. Since the value of t_2^3 is arbitrary, we cannot smooth the crossings to see exactly what $\tau_2(\partial B')$ looks like. But that is not a problem; we can still push this loop up from level 2 to level 3. As we do, we add copies of l_3^3 and l_3^4 to the link diagram. This time, since $t_3^j < 0$ for all j, these l–loops will all be drawn with overcrossings, which gives us a link diagram representing $\tau_3(\partial B')$. Finally we push this up to level 4, adding copies of l_4^2 and l_4^3 (with undercrossings) to the link diagram. In this way we see that $\tau_4(\partial B')$ can be represented by the diagram of the four-component unlink in Figure 20.

Each circle in the diagram in Figure 20 is marked with a number N^i indicating how many parallel copies of that circle are present, where

$$\begin{aligned} N^1 &= t_2^3|t_3^3|t_4^2, \\ N^2 &= t_2^3|t_3^3|, \\ N^3 &= t_2^3 + t_4^3(1 + t_2^3|t_3^3| + t_2^3|t_3^4|), \\ N^4 &= 1 + t_2^3|t_3^4|. \end{aligned}$$

Recall that $t_2^j, t_4^j \geq 2$ and $t_3^j \leq -2$ for all j. No matter what values the twist numbers $\{t_i^j\}$ take on, provided they follow this rule, it can be shown that $N^i > 0$ for each i, and these three inequalities hold true:

$$N^1 > N^2, \quad N^2 < N^3, \quad N^3 > N^4.$$

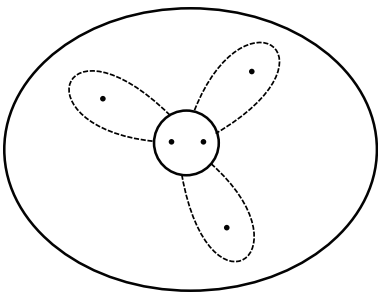


Figure 21: Proposition 4.2 asserts that Lab is isotopic to this picture. The boundary $\partial B'$ and the three gates cut Lab into five components: three once-punctured disks, an annulus, and a twice-punctured disk.

This means that at each of the six crossings, the strand labeled with the higher number passes under the strand labeled with the lower number, so all six crossings look like the bottom picture in Figure 18.

Proposition 4.1 *The position of $\partial B'$ described by the link diagram in Figure 20 is minimal with respect to $\beta^2 \cup \beta^3 \cup \beta^4$.*

Proof The bigon criterion in [5] tells us that as long as $\partial B'$ cobounds no bigons with $\beta^2 \cup \beta^3 \cup \beta^4$, then they are in minimal position. Consider $\partial B'$ as being cut into component strands by $\beta^2 \cup \beta^3 \cup \beta^4$. To bound a bigon, one of these strands would have to have both endpoints on a single β–arc. However, each strand has endpoints on distinct β–arcs. □

Let Lab be the disk in F_4 which contains $\tau_4(\partial B')$ and whose complement is a regular neighborhood of $\beta^1 \cup \gamma^1$. (Lab is depicted in Figures 13 and 20, and also, Lab is isotopic in F to the top gray rectangle in Figure 12.) We will refer to Lab as the *labyrinth*.

Define $G^{\mathrm{b}}, G^{\mathrm{o}}$, and G^{p} to be the dotted arcs depicted in Figures 13 and 20. We will call these the *brown, orange*, and *purple gates* of the labyrinth, respectively. Each gate is an arc in Lab with endpoints on $\partial B'$.

Proposition 4.2 *$\partial B'$ and the three gates cut* Lab *into five components: three once-punctured disks, an annulus, and a twice-punctured disk.*

Proof Since each crossing in our link diagram in Figure 20 looks like the bottom picture of Figure 18, it is apparent that there is an annulus component, which we will

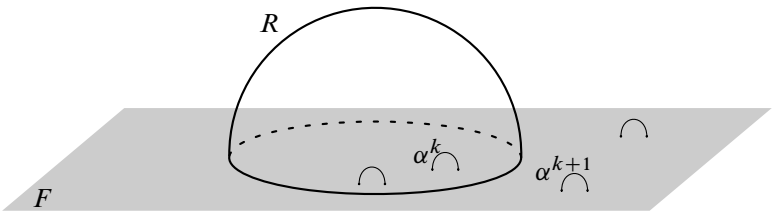


Figure 22: In Case 1, α^2, α^3 and α^4 appear on both sides of R.

call A. A has ∂(Lab) as one boundary component, and the other component of ∂A alternates between the three gates and three subarcs of $\partial B'$. Observe that A is not punctured. The boundary $\partial B'$ cuts off the twice-punctured disk U_2^3 from Lab by definition, and clearly none of the gates are in this disk. Lab $\setminus U_2^3$ is a thrice-punctured annulus in which the three gates are properly embedded but not nested: if any subset of the gates were nested, then at least one of them would not be a boundary component of A. None of the gates can be parallel to $\partial B'$ because that would either force A to be punctured or force the other gates to be nested. Thus the only possible configuration is (isotopic to) Figure 21. □

We will refer to the three once-punctured disks in Proposition 4.2 as the *brown*, *orange*, and *purple punctured disks*, according to the color of the corresponding gate, and we will call the marked point contained therein a *brown*, *orange*, or *purple marked point*. In the brown disk, there is a unique arc (up to isotopy) connecting the brown point to G^{b}, which we will call the *brown escape route*. We define the *orange* and *purple escape routes* similarly. (The three escape routes are depicted in the second picture of Figure 32.) We can think of the colored points as escaping from a maze whose walls are ($\partial B'$), and the escape routes are the paths they take to get to the exits (the gates).

5 Red disks enter the labyrinth

Refer to Figure 10.

Lemma 5.1 *If R is a red disk above F, then R must intersect γ^2 or γ^3.*

Proof F cuts S^3 into two 3–balls: M_+ above F and M_- below F. R divides M_+ into two 3–balls which we will call M_+^1 and M_+^2 and refer to as the two sides of R.

Case 1 (α^2, α^3 and α^4 are not all on the same side of R) Then for $k = 2$ or $k = 3$, α^k and α^{k+1} are on opposite sides of R. Since γ^k connects an endpoint of α^k and an endpoint of α^{k+1}, γ^k must intersect R.

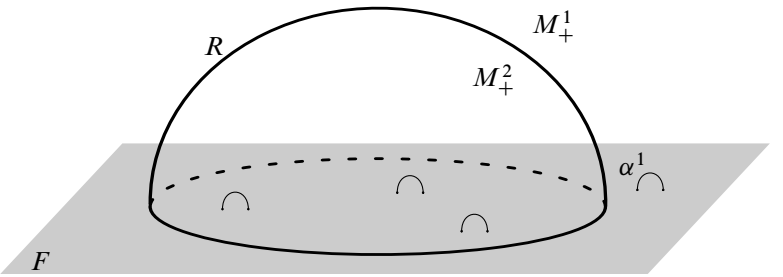


Figure 23: In Case 2, $\alpha^1 \subset M_+^1$ and $\alpha^2, \alpha^3, \alpha^4 \subset M_+^2$.

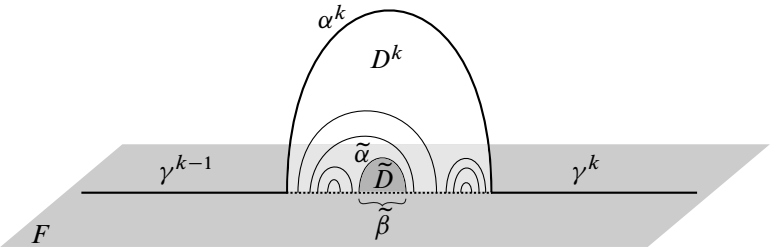


Figure 24: In Subcase 2.2, $R \cap D^k$ is nonempty. Possible arcs of intersection are depicted on D^k.

Case 2 (all of the bridge arcs α^2, α^3 and α^4 are on the same side of R) Without loss of generality, say they are in M_+^2. See Figure 23. Notice that α^1 cannot also be in M_+^2 because that would imply ∂R is null-homotopic in F, contradicting the fact that R is a compressing disk. Therefore, in Case 2, R separates α^1 from the other bridge arcs, which implies R is a cap for α^1. Let $\mathfrak{D} = D^2 \cup D^3 \cup D^4$, and isotope R to minimize $\#|R \cap \mathfrak{D}|$.

Subcase 2.1 (R is disjoint from $\mathfrak{D}$) Assume (for contradiction) that R is also disjoint from γ^2 and γ^3. Then we have a single straight arc $\xi = \beta^2 \cup \gamma^2 \cup \beta^3 \cup \gamma^3 \cup \beta^4$ disjoint from ∂R because $\beta^j \subset D^j \subset \mathfrak{D}$. So ∂R cuts F into two disks, one of which contains ξ, and the other of which contains $\partial\beta^1$. But of course, there is only one such loop in F, which is ∂B. Thus $R \simeq B$, so the red disk R is blue, a contradiction. We conclude that in this subcase, ∂R must intersect γ^2 or γ^3.

Subcase 2.2 (R is not disjoint from $\mathfrak{D}$) For some $D^k \subset \mathfrak{D}$, $R \cap D^k$ is nonempty. $R \cap D^k$ cannot contain loop intersections because $\#|R \cap \mathfrak{D}|$ is minimal, and $S^3 \backslash \eta(L)$ is irreducible. $R \cap D^k$ cannot contain an arc with either endpoint on α^k because that would imply that R intersects L. Thus the components of intersection in D^k must all be arcs with both endpoints on β^k. See Figure 24.

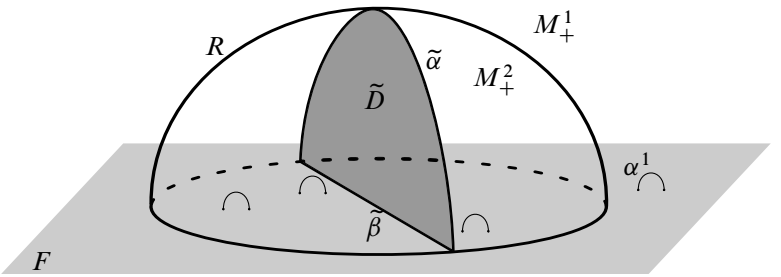


Figure 25: In Subcase 2.2, $R \cap D^k$ is nonempty. Here we see R and the outermost disk $\widetilde{D}$.

Take an arc of intersection outermost on D^k and call it $\widetilde{\alpha}$. Let $\widetilde{\beta}$ be the subarc of β^k that shares endpoints with $\widetilde{\alpha}$. Then let $\widetilde{D}$ be the subdisk of D^k cobounded by $\widetilde{\alpha}$ and $\widetilde{\beta}$. See Figures 24 and 25. $\widetilde{D} \subset M_+^j$ for either $j = 1$ or $j = 2$. Whichever is the case, $\widetilde{D}$ cuts M_+^j into two balls which we call the sides of $\widetilde{D}$. Recall that we assume $\alpha^2, \alpha^3, \alpha^4 \subset M_+^2$, so therefore $\alpha^1 \subset M_+^1$. $\widetilde{D}$ cannot be in M_+^1 with α^1. If it were, then since α^1 cannot intersect R or $\widetilde{D}$, α^1 would have to be completely contained on one side or the other of $\widetilde{D}$. Then the other side of $\widetilde{D}$ would be an empty 3–ball through which we can isotope R, removing at least one component of intersection with D^k, which is a contradiction since $\#|R \cap \mathfrak{D}|$ is already minimal. Thus $\widetilde{D}$ is in M_+^2 with α^2, α^3 and α^4.

If α^2, α^3 and α^4 were on the same side of $\widetilde{D}$, that would mean that the other side of $\widetilde{D}$ was an empty 3–ball through which we could isotope R, removing the arc of intersection $\widetilde{\alpha} \subset R \cap \mathfrak{D}$, contradicting minimality. Thus it must be the case that there are some α–arcs in either side of $\widetilde{D}$.

Then for $k = 2$ or $k = 3$, α^k and α^{k+1} are on opposite sides of $\widetilde{D}$. An endpoint of α^k and an endpoint of α^{k+1} are connected by the arc γ^k in F. None of the γ–arcs have interiors that intersect $\mathfrak{D}$, so γ^k must pass through ∂R in order to connect α^k to α^{k+1}. Thus we have proved that in Subcase 2.2, $\partial R \cap (\gamma^2 \cup \gamma^3) \neq \varnothing$.

Now we have proved both subcases, concluding Case 2 and finishing the proof of Lemma 5.1. □

Corollary 5.2 *Every red disk intersects at least one of $\tau_4(\partial B')$, G^b, G^o and G^p.*

Proof Recall that by Proposition 4.2, $\text{Lab} \setminus \left(\tau_4(\partial B') \cup G^{\text{b}} \cup G^{\text{o}} \cup G^{\text{p}}\right)$ has five components: three (colored) once-punctured disks, one twice-punctured disk (U_2^3), and

an annulus. See Figure 21. Let R be a red compressing disk above F. The boundary of a compressing disk cannot bound a once-punctured disk in F. Thus ∂R cannot be contained in one of the three colored once-punctured disks. If ∂R lies in the twice-punctured disk, then ∂R must be isotopic to $\tau_4(\partial B')$. But then after the isotopy, $R \cup B'$ would be a splitting sphere for L, a contradiction since L is nonsplit. We see in Figures 13 and 20 that γ^2 and γ^3 are disjoint from the annulus component of Lab, so if ∂R is contained in the annulus, ∂R would fail to intersect γ^2 or γ^3, contradicting Lemma 5.1. We conclude ∂R cannot lie in a single component of $\text{Lab} \setminus \big(\tau_4(\partial B') \cup G^{\mathrm{b}} \cup G^{\mathrm{o}} \cup G^{\mathrm{p}}\big)$; therefore it must intersect $\tau_4(\partial B') \cup G^{\mathrm{b}} \cup G^{\mathrm{o}} \cup G^{\mathrm{p}}$. □

6 All red disks above F intersect all blue disks below

Recall the disks and arcs defined in Section 3 and pictured in Figures 10 and 11. Let $\mathfrak{D} = D^2 \cup D^3 \cup D^4$ as in Section 5. Let $\mathcal{R}$ be the set of red compressing disks above F which are disjoint from β'. Throughout Section 6, whenever $\mathcal{R}$ is nonempty, assume that R is a disk in $\mathcal{R}$ such that $|R \cap \mathfrak{D}| \le |R' \cap \mathfrak{D}|$ for all $R' \in \mathcal{R}$, and assume R is in minimal position with respect to $\mathfrak{D}$ and with respect to the gates. β' cuts β^i into multiple subarcs, which we call *lanes*.

Lemma 6.1 *Assume $\mathcal{R}$ is nonempty. If two points of $R \cap \beta^i$ lie in the same lane, then they cannot be the endpoints of a common arc of $R \cap D^i$.*

Proof Suppose there exists an arc of $R \cap D^i$ whose endpoints lie in the same lane. We may assume the arc is outermost on D^i. It cuts off a small subdisk $\widetilde{D}^i \subset D^i$ which does not intersect β'. Boundary-compress R along $\widetilde{D}^i$. The result is two disks, R^1, R^2 whose boundaries lie in F, and which are disjoint from β'. Suppose R^1 is trivial. Then it represents an isotopy through which we can move R to decrease its intersection with the bridge disks. But that implies R was not in minimal position with respect to the bridge disks, a contradiction. Thus R^1 is not trivial, and similarly, neither is R^2.

This means both R^1 and R^2 are compressing disks. By construction, neither of them intersects β', and $|R^i \cap \mathfrak{D}| < |R \cap \mathfrak{D}|$ for both i. Thus $R^i \notin \mathcal{R}$ by our choice of R as a minimal representative from $\mathcal{R}$. Then by the definition of $\mathcal{R}$, both R^i must be blue. There is only one blue disk above F, which is B, so R^1 and R^2 must be parallel copies of B. This implies R is a band sum of parallel disks, but any band sum of parallel disks is trivial, so R is a trivial disk, which is a contradiction. □

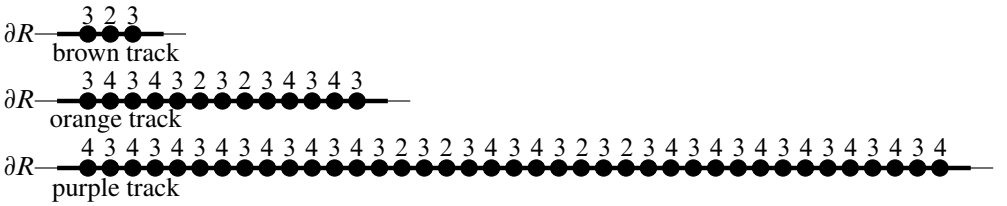


Figure 26: Here we depict the brown, orange, and purple tracks in ∂R with their respective sequences of numbered points for the example depicted in Figure 13 (in which $t_1^j = t_3^j = 2$ and $t_2^j = -2$ for all j). We will not prove this figure, nor does the paper depend on it. The reader can check it by taking a pencil (and perhaps a magnifying glass) to Figure 13 and tracing out a path that winds in a gate, around a marked point inside the labyrinth, and then back out.

By definition, $R \in \mathcal{R}$ implies $R \cap \beta' = \varnothing$, so R can also be made disjoint from $\partial B'$ since $\partial B'$ is the boundary of a regular neighborhood of β'. It follows from Corollary 5.2 that R must intersect at least one gate. Figure 21 makes clear something not at all obvious in Figures 13 and 20, which is that if an arc of R "enters" the labyrinth through a particular gate, it must subsequently "exit" the labyrinth from the same gate. We will call the components of the intersection of ∂R and the brown disk *brown tracks*, and we will define *orange tracks* and *purple tracks* similarly. Observe that these tracks are pairwise disjoint, and each track has endpoints on the gate of the corresponding color.

We label the points of $\partial R \cap D^2$, $\partial R \cap D^3$, and $\partial R \cap D^4$ *2–points*, *3–points*, and *4–points*, respectively, and we will collectively refer to these as *numbered points*. Likewise, we will label the arcs of $R \cap D^2$, $R \cap D^3$, and $R \cap D^4$ *2–arcs,* *3–arcs*, and *4–arcs*, respectively, and we will collectively call them *numbered arcs*. Observe that the endpoints of a j–arc are j–points. The next result follows from these definitions.

Lemma 6.2 *If $\mathcal{R}$ is nonempty, at least one endpoint of each 2–arc must lie in a track, and all 3– and 4–points lie in tracks.*

Proof The leftmost lane of β^2 is not completely contained in Lab, but all the other lanes of β^2, as well as all the lanes of β^3 and β^4, are contained in the union of the three colored disks and U_2^3. The boundary ∂R is disjoint from U_2^3, so it can only intersect these lanes inside the colored disks. Since every arc component of ∂R inside a colored disk is, by definition, a track, we conclude that aside from the leftmost lane of β^2, ∂R only intersects the lanes of β^2, β^3 and β^4 in tracks. Thus all 3– and 4–points lie in tracks. If both endpoints of a 2–arc lie outside of the tracks, then those endpoints must lie in the leftmost lane of β^2, contradicting Lemma 6.1. Therefore at least one endpoint of each 2–arc must lie in a track. □

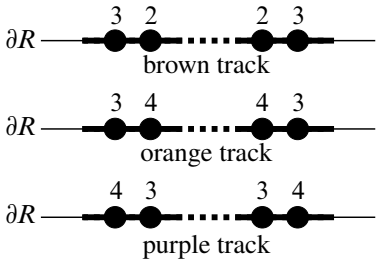


Figure 27: Each colored track contains a sequence of numbered points. No matter the twist numbers $\{t_i^j\}$, we can say what the outermost and second-outermost points will be for each type of track.

Corresponding to each track color is a particular sequence of numbered points. Each of these three sequences is symmetric in the sense that the outermost numbered points in the track have the same number, the second-outermost numbered points match each other, the third-outermost numbered points match, etc. Figure 26 depicts the sequence of numbered points along each colored track in the case depicted in Figure 13.

Observe Figure 28, which depicts the possibilities for which directions tracks can go after they intersect the β–arcs. We see that the outermost numbered points of a brown track are 3–points because the first and last β–arc that the track will intersect is β^3. After intersecting β^3, a brown track will loop around and intersect β^2, so the second-outermost numbered points of a brown track are 2–points. Similarly, the outermost numbered points of an orange track are 3–points, and the second-outermost numbered points of an orange track are 4–points. Finally, the outermost numbered points of a purple track are 4–points, and the second-outermost numbered points of a purple track are 3–points. This information is summarized in Figure 27. Note that this information about the outermost and second-outermost numbered points does not depend on the set $\{t_i^j\}$ of twist numbers.

Lemma 6.3 *Assume $\mathcal{R}$ is nonempty. No numbered arc can have endpoints which lie on the same track.*

Proof Since ∂R must intersect at least one of the gates, it follows that there exists at least one track in ∂R. Refer again to Figure 28: In the top picture, we see that after intersecting β^3, a track will always either intersect β^2 or β^4 or leave the labyrinth. In the bottom picture, we see that after intersecting β^2 or β^4, a track will always either intersect β^3 or leave the labyrinth. This means that in each sequence of numbered points in a track, every other numbered point is a 3–point.

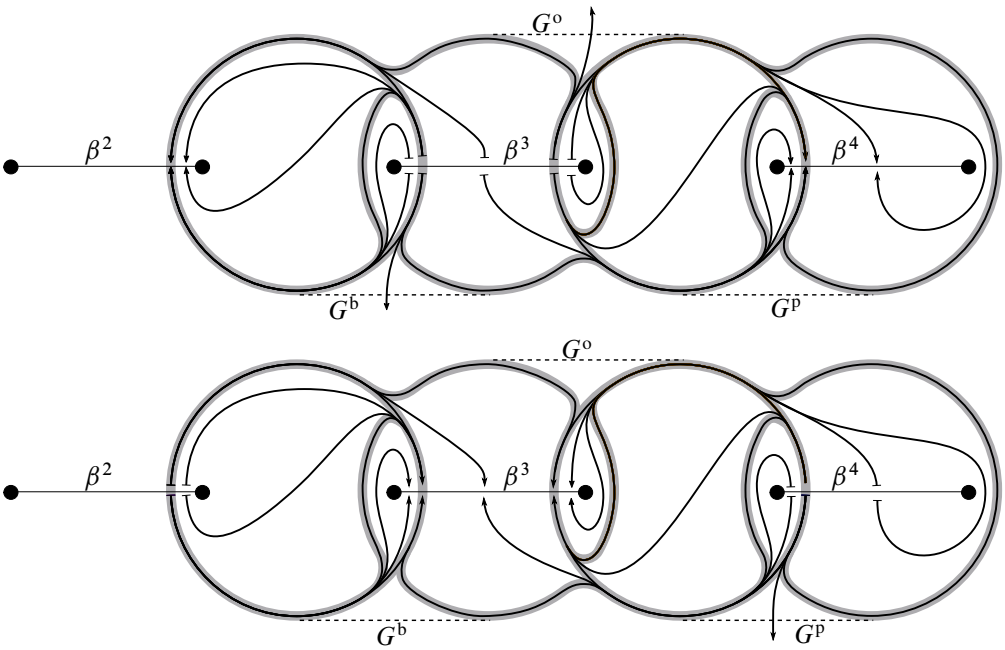


Figure 28: If λ is a track, then λ contains a sequence of numbered points. The top picture shows that each 3–point on λ either has a 2–point or a 4–point to both sides, or it has a 2–point or a 4–point to one side and exits through a gate to the other side. The bottom picture shows that each 2–point on λ has a 3–point to either side, and that each 4–point either has a 3–point to either side, or it has a 3–point to one side and exits through a gate to the other side.

Suppose two numbered points from the same track are connected in R by a numbered arc λ. Since the two endpoints of λ must have the same number, and since every other numbered point in the track is a 3–point, there must be an odd number of numbered points between the endpoints of λ. But this leads to a contradiction: numbered arcs never intersect each other, so it is impossible to pair up the numbered points between the endpoints of λ with disjoint numbered arcs. □

Lemma 6.4 *If $\mathcal{R}$ is nonempty, ∂R contains tracks of all three colors.*

Proof Consider a numbered arc λ which is outermost in R. Being outermost guarantees that the endpoints of λ are adjacent same-numbered points on ∂R. Since numbered points in each track alternate between 3–points and 2– or 4–points, there are only three ways to have adjacent same-numbered points on ∂R:

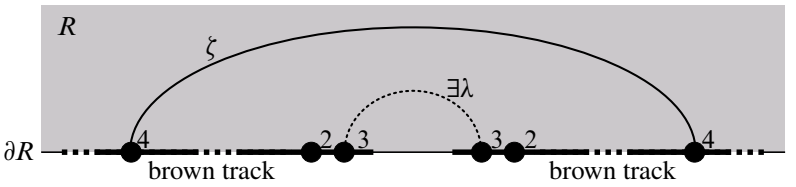


Figure 29: Case 1: The outermost 4–arc ζ connects 4–points in adjacent brown tracks.

(1) The endpoints of λ are consecutive 2–points which do not lie in tracks, which would contradict Lemma 6.2.

(2) The endpoints of λ are a 2–point off of a track, and an outermost numbered point on a track, which is impossible since no outermost point on any track is a 2–point.

(3) The endpoints of λ are outermost same-numbered points of adjacent tracks.

Suppose ∂R contains no brown tracks. Since the outermost numbered points of orange tracks are 3–points, and the outermost numbered points of purple tracks are 4–points, λ must have endpoints which are adjacent outermost numbered points of two adjacent same-colored tracks, and we reach a contradiction by Lemma 6.1 since these two endpoints lie in the same lane. Therefore ∂R must contain at least one brown track. By a similar argument, ∂R contains at least one orange track.

Suppose ∂R contains no purple tracks. Consider a numbered 4–arc ζ of $R \cap D^4$ which is outermost in R (ie outermost among the set of 4–arcs). The arc ζ cuts ∂R into two components. Since ζ is outermost among the 4–arcs, one of these components contains no 4–points. Call this piece the *primary* piece. There are three possibilities: ζ connects two 4–points in two brown tracks, it connects two 4–points in two orange tracks, or it connects a 4–point of a brown track and a 4–point of an adjacent orange track. We will derive a contradiction in each case.

Case 1 (ζ connects two 4–points in two brown tracks)

For this to be possible, the set $\{t_i^j\}$ of twist numbers must be such that brown tracks contain 4–points, which is only true for some sets of twist numbers. In contrast, orange and purple tracks necessarily contain 4–points, no matter the twist numbers. This means that since the primary piece of ∂R cannot contain any 4–points, it cannot contain any complete track of any color. Then the 4–points which are the endpoints of ζ must lie in adjacent brown tracks. See Figure 29. Consider all of the numbered points in the

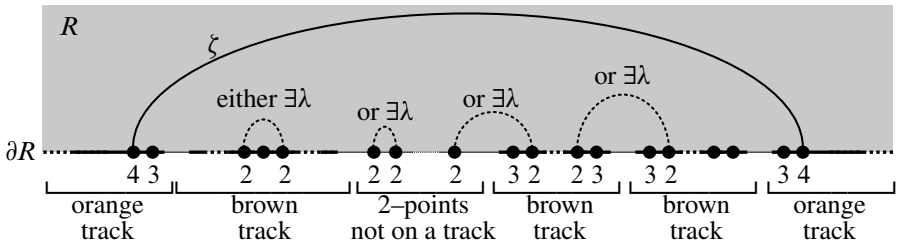


Figure 30: Case 2: The outermost 4–arc ζ connects 4–points in adjacent orange tracks. If brown tracks do not contain 4–points, then there may be any number of brown tracks between these two orange tracks (and there may be any number of 2–points between each pair of tracks).

primary piece of ∂R. They must be paired with numbered arcs, and there must be an outermost such arc λ, which connects same-numbered points adjacent on ∂R. If there are no 2–points between the brown tracks in the primary piece of ∂R, then the only pair of adjacent same-numbered points in the primary part of ∂R are an outermost 3–point of one brown track, and an outermost 3–point of the other brown track. But these points are the same lane of β^3, which contradicts Lemma 6.1. If there is exactly one 2–point between the brown tracks in the primary piece of ∂R, then there does not exist a pair of adjacent same-numbered points in the primary part of ∂R to be the endpoints of λ. If there is more than one 2–point between the brown tracks in the primary piece of ∂R, then the only pair(s) of adjacent same-numbered points in the primary part of ∂R are pairs of these 2–points, but they cannot be the endpoints of λ by Lemma 6.2. We conclude Case 1 is not possible.

Case 2 (ζ connects two 4–points in two orange tracks)

In this case, brown tracks may or may not contain 4–points, depending on the twist numbers; if they do not, then the primary piece of ∂R may contain any number of brown tracks. If there are zero brown tracks in the primary piece of ∂R, then the proof is similar to Case 1. Suppose the primary piece of ∂R contains at least one brown track. Then since the second-outermost numbered points of brown tracks are 2–points, there are at least two 2–points in the primary piece of ∂R. Since numbered arcs cannot intersect each other, they cannot intersect ζ in particular, so all of the 2–points in the primary piece must be paired with each other by a set S of 2–arcs. Consider an outermost 2–arc λ in S (ie outermost in R with respect to the other 2–arcs of S). The endpoints of λ are a pair of 2–points with no other 2–points between them. There are four ways this can happen (depicted in Figure 30), though each one leads to a contradiction:

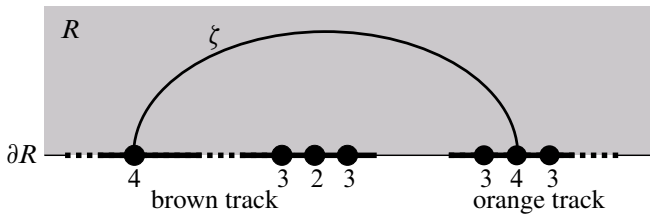


Figure 31: Case 3: Adjacent brown and orange tracks with 4–points connected by the outermost 4–arc ζ.

(1) The pair of 2–points lies in a single brown track, contradicting Lemma 6.3.

(2) Both 2–points are off of tracks, contradicting Lemma 6.1.

(3) The pair consists of a second-outermost numbered point in a brown track and a 2–point not contained in a track. But then λ would separate a single 3–point from all the other numbered points of ∂R, so the numbered arc corresponding to that 3–point would have to intersect λ.

(4) The pair consists of second-outermost numbered points in adjacent brown tracks, points which lie in the same lane, again contradicting Lemma 6.1.

Therefore Case 2 is impossible.

Case 3 (ζ connects a 4–point of a brown track and a 4–point of an adjacent orange track)

See Figure 31. Again, the numbered points in the primary piece of ∂R must all be connected in pairs. Recall that in every track the numbered points alternate between 3–points and 2– or 4–points, so in the intersection of the primary piece and the brown track, there are at least two 3–points, one on either side of the second-outermost 2–point. In the orange track, the second-outermost numbered point is a 4–point, so it must be an endpoint of ζ, so there is exactly one other numbered point in the intersection of the primary piece and the orange track: an outermost 3–point. Since the primary piece of the brown track contains strictly more than one 3–point, and the primary piece of the orange track contains exactly one 3–point, it is impossible to pair up the 3–points in such a way that no pair lies in the same track, so Case 3 contradicts Lemma 6.3 and is therefore impossible. We conclude ∂R must contain purple tracks, finishing the proof of Lemma 6.4. □

Lemma 6.5 *Assume $\mathcal{R}$ is nonempty. Let $b \subset \beta^i$ be a lane whose endpoints are interior points of β'. Then a track must intersect b.*

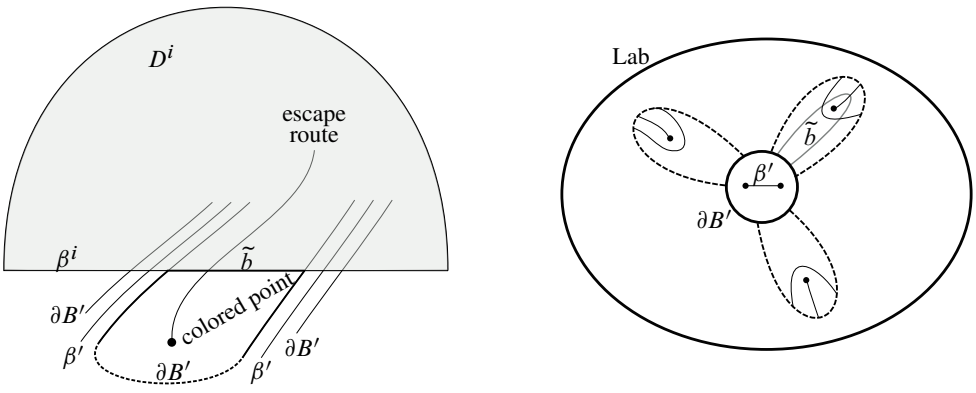


Figure 32: On the left, $\widetilde{b}$ is the bold horizontal segment of β^i. On the right, $\widetilde{b}$ is the gray arc. Also depicted on the right are the three escape routes, and the three corresponding tracks.

Proof Recall that $\partial B'$ and the three gates cut Lab into five components, one of which is the annulus A, and observe that no lane of β^2, β^3 or β^4 intersects A except for the leftmost lane of β^2. By definition, b cannot be that leftmost lane, so b is properly contained in the union of U_2^3 and the three colored disks. Let $\widetilde{b} = b \backslash (U_2^{3\circ})$ as in the first picture of Figure 32; $\widetilde{b}$ is an arc properly contained in a colored disk, with endpoints on ∂U_2^3. $\widetilde{b}$ cannot cobound a bigon with $\partial B'$ because that would contradict Proposition 4.1, so $\widetilde{b}$ must be an arc cutting out a punctured disk from a colored disk, as in the second picture of Figure 32. Suppose the colored disk in question is brown. By definition, the brown escape route does not intersect $\partial B'$, so the only way for it to connect the brown point with the brown gate is to intersect $\widetilde{b}$. Thus if $\widetilde{b}$ is contained in the brown disk, it must intersect the brown escape route. Every brown track is a frontier in the brown disk of the brown escape route, so every brown track must also intersect $\widetilde{b}$. Since the color was arbitrary, the lemma is proved. □

Proposition 6.6 *If L is not perturbed at $\alpha^{4'}$, then $\mathcal{R}$ is empty.*

Proof Assume $\mathcal{R}$ is nonempty. Consider a numbered arc λ which for some i is outermost in D^i. Let p and p' be the endpoints of λ, and let $\widetilde{\beta}^i$ be the segment of β^i between p and p', as in Figure 33. The points p and p' cannot lie in a single lane by Lemma 6.1. If the two lanes containing p and p' have at least one lane between them (ie if the lanes are not adjacent), then by Lemma 6.5, there must be at least one numbered point between p and p'. However, λ is outermost in D^i, so that cannot

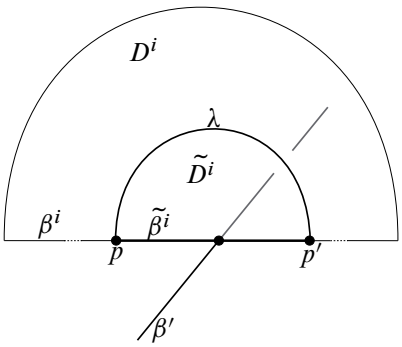


Figure 33: λ straddles exactly one strand of β'

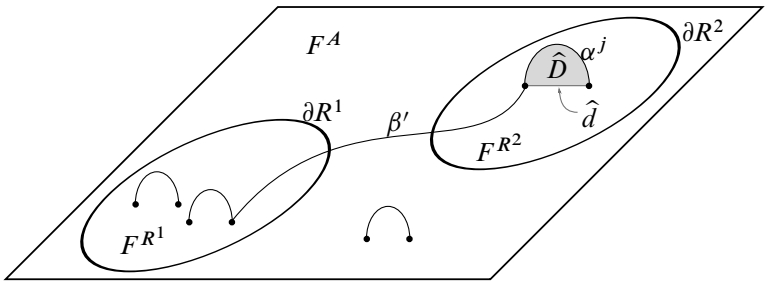


Figure 34: In Case 1, the bridge arcs are all disjoint from R^1 and R^2. One of the disks F^{R^1} or F^{R^2} must contain exactly one pair of partners, the endpoints of a bridge arc α^j. Then L is perturbed at $\alpha^{4'}$.

happen. Therefore the lanes containing p and p' must be adjacent, so $\widetilde{\beta}^i$ contains exactly one point of β'.

The arc λ cuts off a small disk $\widetilde{D}^i$ from D^i. We boundary-compress R along $\widetilde{D}^i$, which results in two new disks, R^1 and R^2, properly embedded in M_+. Since each intersects β' exactly once, ∂R^1 and ∂R^2 cut F into an annulus and two disks, and each of the two disks contains an endpoint of β'. Label these three regions F^{R^1}, F^{R^2} and F^A. The disk R^i cannot be trivial since F is punctured to either side of ∂R^i. Therefore both R^i are compressing disks. Notice that R^1 and R^2 cannot be parallel because the band sum dual to the boundary compression just performed would recover R, but any band sum of parallel disks is a trivial disk.

Define a pair of marked points on F to be *partners* if they are endpoints of a common bridge arc above F. Let the endpoints of β' be q^1 and q^2.

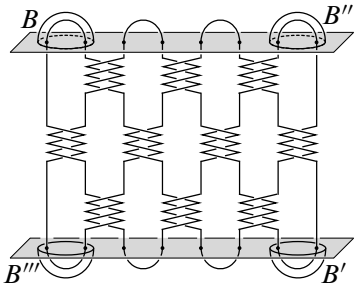


Figure 35: A disjoint pair of blue disks (B and B') and a disjoint pair of red disks (B'' and B''') which fulfill condition (2) of the definition of a critical surface

Case 1 For all four pairs of partners, it is the case that both points are contained in the same region: F^{R^1}, F^{R^2} or F^A. Since an endpoint of β' lies in each of the disks, a pair of partners lies in each of the disks. One pair of partners must be in F^A; otherwise, R^1 and R^2 would be parallel compressing disks. This accounts for three of the four pairs of marked points. The fourth pair may be in any of the three regions F^{R^1}, F^{R^2} or F^A. This means that F^{R^1} or F^{R^2} must contain exactly one pair of partners. See Figure 34. Suppose these partners are the endpoints of α^j for some j. Then there is some bridge disk $\widehat{D}$ for α^j (which is not necessarily isotopic to D^j) which is disjoint from R^i. If $\widehat{d} = F \cap \widehat{D}$, then $\widehat{d}$ and β' can be made disjoint except for their shared endpoint. Then $\widehat{d} \cup \beta'$ is an embedded arc. This means that $\widehat{D}$ and $D^{4'}$ are bridge disks above and below F which intersect in a single point of the link, so L is perturbed at $\alpha^{4'}$. This contradicts the hypothesis of the proposition, so Case 1 is not possible.

Case 2 There is a pair of partners which lie in different regions F^{R^1}, F^{R^2} and F^A. But then the bridge arc connecting them must intersect R^1 or R^2, a contradiction since compressing disks are disjoint from the link. Having arrived at a contradiction, we conclude R cannot exist, so $\mathcal{R}$ is empty, and the proposition is proved. □

Define $\mathcal{R}'$ to be the set of red compressing disks below F disjoint from B. The next result follows directly from Proposition 6.6 and rotational symmetry of L.

Corollary 6.7 *If L is not perturbed at α^1, then $\mathcal{R}'$ is empty.*

Theorem 6.8 *Let L be a 2–twisted link in $(4,4)$–plat position with bridge sphere $F \simeq F_4 = f^{-1}(4) \subset S^3$. Suppose L has twist numbers $\{t_i^j\}$ such that for all j, the twist numbers t_2^j, t_4^j are positive and t_3^j is negative. If L is perturbed at neither α^1 nor $\alpha^{4'}$, then F is a critical bridge sphere.*

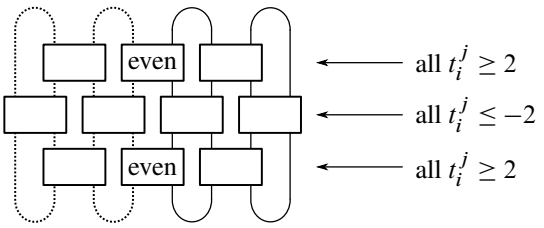


Figure 36: Links L in the family $\mathcal{L}$. Since both t_2^2 and t_4^2 are even, L will have at least two components, L^1 (drawn with dotted arcs) and L^2 (drawn with solid arcs).

Proof By Proposition 6.6 and Corollary 6.7 we conclude that $\mathcal{R}$ and $\mathcal{R}'$ are empty. In other words, all red disks above F intersect all blue disks below F, and vice versa, which fulfills condition (1) of the definition of a critical surface. Observe that the link diagram representing $\partial B'$ (in Section 4) is disjoint from ∂B, so B and B' are disjoint. This gives us a pair of disjoint blue disks, one above F and one below. Consider B'' and B''' depicted in Figure 35. (Explicitly, these are the frontiers of regular neighborhoods of D^4 and $D^{1'}$ above and below F, respectively.) By a symmetric argument, these two compressing disks are a pair of disjoint red disks, one above F and one below. These two pairs of disjoint disks fulfill the condition (2) of the definition of a critical surface. □

We finally come to the proof of Theorem 6.9.

Theorem 6.9 *There is an infinite family of nontrivial links with critical bridge spheres.*

Proof All we must do to prove Theorem 6.9 is to demonstrate that there exists an infinite family of links which satisfy the hypothesis of Theorem 6.8. Let $\mathcal{L}$ be the set of 2–twisted links in $(4,4)$–plat position with twist numbers $\{t_i^j\}$, such that all twist numbers in the top and bottom rows are positive, and all twist numbers in the middle row are negative, and such that t_2^2 and t_4^2 are both even. The parity requirement guarantees that if $L \in \mathcal{L}$, then L is a union of two different links, say L^1 and L^2. See Figure 36. Consider L^1; it is a link in a 2–bridge position (with respect to F), so it can have at most two components.

Case 1 (L^1 has one component, ie it is a knot) Let $D(L^1)$ be the diagram for L^1 obtained by projection to the xz–plane. This diagram is alternating (for the same reason $D(L)$ was shown to be alternating in Section 3). Observe that $D(L^1)$ is a reduced diagram. (In other words, as a graph, it has no cut vertex.) One of the

famous Tait conjectures states that a reduced alternating diagram for a knot realizes the knot's minimal crossing number [9]. (This was proved in 1987 by Kauffman [7] and Murasugi [10].) Therefore as $D(L^1)$ has at least eight crossings, L^1 is not the unknot, so it must be a 2–bridge knot. Since L^1 is in bridge position, it is by definition not perturbed at any of its bridge arcs, one of which is α^1. We conclude that L^1 is not perturbed at α^1. Since L^1 and L^2 are not connected to each other, α^1 cannot share an endpoint with either of the two bridge arcs below F contained in L^2. Therefore we conclude that L is not perturbed at α^1.

Case 2 (L^1 has two components) Then each component is individually an unknot in bridge position. Unknots are not perturbed, so in particular, L^1 is not perturbed at α^1. From there, it follows as in Case 1 that L is not perturbed at α^1 either.

By a similar argument, L is not perturbed at $\alpha^{4'}$ either. Therefore L satisfies Theorem 6.8 and its bridge sphere F is critical. This proves $\mathcal{L}$ is an infinite family of links satisfying Theorem 6.9. □

References

[1] **D Bachman**, *Critical Heegaard surfaces*, Trans. Amer. Math. Soc. 354 (2002) 4015–4042 MR

[2] **D Bachman**, *Topological index theory for surfaces in* 3*–manifolds*, Geom. Topol. 14 (2010) 585–609 MR

[3] **D Bachman**, *Stabilizing and destabilizing Heegaard splittings of sufficiently complicated* 3*–manifolds*, Math. Ann. 355 (2013) 697–728 MR

[4] **A J Casson**, **C M Gordon**, *Reducing Heegaard splittings*, Topology Appl. 27 (1987) 275–283 MR

[5] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) MR

[6] **J Johnson**, **Y Moriah**, *Bridge distance and plat projections*, Algebr. Geom. Topol. 16 (2016) 3361–3384 MR

[7] **L H Kauffman**, *State models and the Jones polynomial*, Topology 26 (1987) 395–407 MR

[8] **J H Lee**, *Bridge spheres for the unknot are topologically minimal*, Pacific J. Math. 282 (2016) 437–443 MR

[9] **W B R Lickorish**, *An introduction to knot theory*, Graduate Texts in Mathematics 175, Springer (1997) MR

[10] **K Murasugi**, *Jones polynomials and classical conjectures in knot theory*, Topology 26 (1987) 187–194 MR

Mathematics Department, Taylor University
Upland, IN, United States

daniel_rodman@taylor.edu

Received: 17 February 2016 Revised: 6 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 187–219

A second cohomology class of the symplectomorphism group with the discrete topology

RYOJI KASAGAWA

A second cohomology class of the group cohomology of the symplectomorphism group is defined for a symplectic manifold with first Chern class proportional to the class of symplectic form and with trivial first real cohomology. Some properties of it are studied. In particular, it is characterized in terms of cohomology classes of the universal symplectic fiber bundle over the classifying space of the symplectomorphism group with the discrete topology.

57R17, 57R50; 55R40, 58H10

1 Introduction

A two cocycle on the symplectomorphism group of an exact symplectic manifold with trivial first real cohomology was defined and studied by Ismagilov, Losik and Michor [12]. Gal and Kędra [8] carried investigation of it further. They provided another definition of the two cocycle and obtained vanishing and nonvanishing results of its cohomology class. They also gave some applications of it to foliated symplectic bundles and Hamiltonian actions. This two cocycle is constructed by using a primitive of the symplectic form, the existence of which is the starting point of the construction.

In this paper we consider a symplectic manifold (M,ω) of dimension $2n$ with first Chern class $c_1(M)$ proportional to the class $[\omega]$ of the symplectic form ω, that is, we assume the equality $c_1(M)+\kappa[\omega]=0$ in $H^2(M;\mathbb{R})$ for some $\kappa\in\mathbb{R}$. This is our assumption instead of the exactness of the symplectic form. We define a second cohomology class $\sigma\in H^2(G,\mathbb{R})$ in the group cohomology of the symplectomorphism group $G=\mathrm{Symp}(M,\omega)$ of (M,ω) with values in $\mathbb{R}$ in a similar way to another definition by Gal and Kędra, mentioned above. In order to define σ, we need the symplectic twistor space of (M,ω), which is a fiber bundle $p\colon \mathcal{J}\to M$ over M with fiber the Siegel upper half space; see Albuquerque and Rawnsley [3]. In particular we use the fact that the standard action of G on M induces that as $U(n)$–bundle isomorphisms on the pullback bundle p^*TM over $\mathcal{J}$ of the tangent bundle TM of M by p, which has a canonical Hermitian structure. The class σ is defined in terms

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.187

of differential forms on $\mathcal{J}$ using a unitary connection of p^*TM and other ingredients. We can show vanishing results for the restrictions of it to certain subgroups.

The group cohomology $H^*(G,\mathbb{R})$ is identified with $H^*(BG^\delta;\mathbb{R})$, the singular cohomology of the classifying space BG^δ of G^δ, which is G with the discrete topology. In order to give the class $\|\sigma\|$ in $H^2(BG^\delta;\mathbb{R})$ corresponding to σ, we need to handle differential forms on the fiber bundle $EG^\delta\times_{G^\delta}\mathcal{J}$ with fiber $\mathcal{J}$ associated with the universal principal bundle $EG^\delta\to BG^\delta$. Therefore we give the bundles as simplicial manifolds and we do calculus of differential forms on them; see Dupont [6; 7]. The class $\|\sigma\|$ is characterized so that its pullback by the projection of the associated (M,ω)–bundle $EG^\delta\times_{G^\delta}M\to BG^\delta$ is the sum of the first Chern class of the tangent bundle along the fibers of it and κ times the transverse symplectic class of it. This is a similar result to that for the transverse symplectic class of a foliated surface bundle with area-preserving holonomy in Kotschick and Morita [19], and is used to show the nontriviality of σ for the example in this paper. Let BG be the classifying space of G with the C^∞ topology. The identity map of G induces a continuous map $BG^\delta\to BG$, hence a homomorphism $\iota^*\colon H^*(BG;\mathbb{R})\to H^*(BG^\delta;\mathbb{R})$. If M is closed, a constant multiple of the cohomology class $\|\sigma\|$ is shown to be the image ι^*c by ι^* of a certain characteristic class $c\in H^2(BG;\mathbb{R})$ of (M,ω)–bundles obtained by fiber integration; see Januszkiewicz and Kędra [13]. Thus $\|\sigma\|$, and hence σ, is essentially a known characteristic class in the closed case. Therefore similar results for ι^*c to those for σ hold.

Finally we give a nontrivial example of σ. We consider the moduli space $\mathcal{R}_g$ of $\mathrm{SU}(2)$–representations of the fundamental group of a compact oriented surface of genus g with one boundary whose holonomy along the boundary is $-I$. It is a closed, simply connected symplectic manifold satisfying our assumptions. Using some facts on mapping class groups of surfaces, we can show the nontriviality of the class σ of the symplectomorphism group of the symplectic manifold $\mathcal{R}_g$ for $g\geqq 4$. This result implies that the class ι^*c is also nontrivial in this example, as stated above, which tells us that a characteristic class obtained by fiber integration is generally nontrivial on discrete symplectomorphism groups, in particular of closed simply connected symplectic manifolds. On the other hand, Reznikov defined cohomology classes of the symplectomorphism group of a general symplectic manifold and showed a similar result to this example in Reznikov [23]. One may be curious about the connection between Reznikov's second cohomology class and σ, but we leave it as a question.

This paper is organized as follows. In Section 2, we define the second cohomology class σ and state main results. In Section 3, we show the well-definedness of σ and vanishing results for restriction of σ to certain subgroups of G. In Section 4, we recall some properties of simplicial manifolds and introduce simplicial manifolds $N\overline{\mathcal{J}}$ and others related with G. In Sections 5 and 6, we define a simplicial 2–form P on $N\overline{\mathcal{J}}$

and show a relation between it and σ. In Section 7, we define a second cohomology class $\|P\|$ of the fat realization of $N\overline{\mathcal{J}}$ corresponding to P and prove other main results. In Section 8, we give a nontrivial example of σ.

2 Definitions and main results

Let (M,ω) be a connected, symplectic manifold of dimension $2n$ and let $G = \mathrm{Symp}(M,\omega)$ the symplectomorphism group of it. Let $\mathcal{P} \to M$ be the principal symplectic frame bundle for (M,ω). Let $\mathcal{J}$ be the quotient space $\mathcal{P}/U(n)$; then the projection $p\colon \mathcal{J} \to M = \mathcal{P}/\mathrm{Sp}(2n,\mathbb{R})$ is a fiber bundle over M with fiber $\mathrm{Sp}(2n,\mathbb{R})/U(n)$. It is a smooth manifold of dimension $2n + n(n+1)$. The fiber $\mathrm{Sp}(2n,\mathbb{R})/U(n)$ is identified with the Siegel upper half space and is contractible. It is also identified with the space of positive compatible complex structures on the standard symplectic vector space $\mathbb{R}^n$. Therefore we can consider $\mathcal{J}$ as the fiber bundle over M whose fiber $\mathcal{J}_x$ at $x \in M$ is the space of positive compatible complex structures on the symplectic vector space (T_xM,ω_x) [3; 21; 25]. Since the pullback bundle p^*TM over $\mathcal{J}$ of the tangent bundle TM by the projection p is also given as $p^*TM = \mathcal{P}\times_{U(n)}\mathbb{C}^n$, it is obvious it is a Hermitian vector bundle. It is easily checked that its Hermitian form on the fiber $p^*TM|_{j_x} = T_xM$ at $j_x \in \mathcal{J}$ with $p(j_x) = x$ is given by $h_{j_x}(\cdot,\cdot) = \omega_x(\cdot, j_x\cdot) - \sqrt{-1}\omega_x(\cdot,\cdot)$. The standard action of G on M induces those on $\mathcal{P}$, $\mathcal{J}$ and p^*TM. Since the construction of the bundles is natural, the projections $p^*TM \to \mathcal{J} \to M$ are G–equivariant. Moreover G acts on p^*TM by $U(n)$–bundle isomorphisms. Therefore G acts also on the space of $U(n)$–connections of p^*TM. Let A be a $U(n)$–connection of p^*TM; then we have the pushforward connection of A by $\varphi \in G$, which is denoted by φ_*A. We remark that the actions of $\varphi \in G$ on M and on $\mathcal{J}$ are denoted by $\varphi(x)$ for $x \in M$ and by $\varphi(j)$ for $j \in \mathcal{J}$, respectively, without distinction. In particular its induced action by pullback on the space $\Omega^*(\mathcal{J})$ of forms on $\mathcal{J}$ at $\alpha \in \Omega^*(\mathcal{J})$ is denoted by $\varphi^*\alpha$. But we mainly use the left action given by the pullback $\varphi_*\alpha = (\varphi^{-1})^*\alpha$ of the inverse.

Let $c_1(M)$ and $[\omega]$ be the first Chern class of (M,ω) and the cohomology class of the symplectic form ω respectively in the second cohomology group $H^2(M)$ of M. Here and hereafter singular cohomology groups and de Rham cohomology groups are identified and the coefficients of them are $\mathbb{R}$ if not mentioned explicitly.

We assume that the equality $c_1(M) + \kappa[\omega] = 0$ holds in $H^2(M)$ for some $\kappa \in \mathbb{R}$. Pulling back by the projection p, we have $c_1(p^*TM) + \kappa[p^*\omega] = 0$ in $H^2(\mathcal{J})$. Thus there exists a 1–form $\mu \in \Omega^1(\mathcal{J})$ on $\mathcal{J}$ satisfying

$$c_1(F_A) + \kappa p^*\omega + d\mu = 0. \tag{2-1}$$

Here F_A is the curvature form of A and c_1 is the $\mathrm{Ad}_{U(n)}$–invariant linear form on the Lie algebra $u(n)$ corresponding to the first Chern class. Put

$$f(\varphi) = c_1(\varphi_* A - A) + \varphi_* \mu - \mu$$

for each $\varphi \in G$. It is easy to see that $f(\varphi)$ belongs to the space $Z^1(\mathcal{J})$ of closed 1–forms on $\mathcal{J}$. Thus we have a map $f\colon G \to Z^1(\mathcal{J})$.

Lemma 2.1 *The map f is a crossed homomorphism and its cohomology class $[f] \in H^1(G, Z^1(\mathcal{J}))$ in the group cohomology of G with coefficients in $Z^1(\mathcal{J})$ is independent of the choice of A and μ.*

This lemma is explained in the beginning of Section 3. But its detailed proof is not given there because it was proved in a general setting in [15].

If the first cohomology group of M is nontrivial, the class $[f]$ descents to a class in $H^1(G, H^1(M))$ by considering the projection $Z^1(\mathcal{J}) \to H^1(\mathcal{J}) \cong H^1(M)$. There are some works about it in [15] and related ones in [4; 16; 18; 19].

From now on we assume in addition $H^1(M) = \{0\}$. Under the assumptions, considering the de Rham complex of $\mathcal{J}$, we have a short exact sequence $\{0\} \to \mathbb{R} \hookrightarrow \Omega^0(\mathcal{J}) \xrightarrow{d} B^1(\mathcal{J}) = Z^1(\mathcal{J}) \to \{0\}$ of G–modules, where $B^1(\mathcal{J})$ denotes the space of exact 1–forms on $\mathcal{J}$. It induces a long exact sequence of cohomology groups of G. In particular we have the connecting homomorphism $\delta\colon H^1(G, Z^1(\mathcal{J})) \to H^2(G, \mathbb{R})$ in it [5]. The following proposition is proved in Section 3.

Proposition 2.2 *Let (M, ω) be a connected symplectic manifold of dimension $2n$ with $H^1(M) = \{0\}$. Assume $c_1(M) + \kappa[\omega] = 0$ in $H^2(M)$ for some $\kappa \in \mathbb{R}$. Then $\sigma := \delta([f]) \in H^2(G, \mathbb{R})$ is uniquely defined, that is, it is independent of the construction.*

By the definition of the connecting homomorphism δ, a 2–cocycle representing the class σ can be expressed in terms of f. Using this expression, we can show vanishing results of restrictions of σ to certain subgroups of G. For any $j \in \mathcal{J}$ and any positive compatible almost complex structure J on (M, ω), let G_j and G_J be the subgroups of G consisting of all elements which preserve j and J, respectively. The following theorem is proved in Section 3.

Theorem 2.3 *Let K be any subgroup of G_j for some j or of G_J for some J. If $K \subset G_J$, we assume in addition that M is closed. Then the restriction of the class σ to K vanishes, that is, the equality $\iota^*\sigma = 0$ in $H^2(K, \mathbb{R})$ holds, where $\iota\colon K \hookrightarrow G$ is the inclusion.*

Here we remark that the cohomology group of K in this theorem is that in group theory. In this paper, cohomology groups of groups denote those in group theory like $H^*(G,\mathbb{R})$ for G, as mentioned in the introduction.

Let G^δ be G with the discrete topology and $EG^\delta \to BG^\delta$ the universal principal G^δ–bundle over the classifying space BG^δ of G^δ. Let $\pi\colon E^\delta M = EG^\delta \times_{G^\delta} M \to BG^\delta$ be the universal (M,ω)–bundle over BG^δ and $E^\delta TM \to E^\delta M$ its tangent bundle along the fibers. Since $E^\delta TM$ is a symplectic vector bundle of rank $2n$, we have the first Chern class $c_1(E^\delta TM) \in H^2(E^\delta M)$. The symplectic form ω on M defines a cohomology class $\|\omega\| \in H^2(E^\delta M)$, which is precisely given in Section 7. Since $H^2(BG^\delta)$ is naturally identified with the group cohomology $H^2(G,\mathbb{R})$, we have a homomorphism $H^2(G,\mathbb{R}) = H^2(BG^\delta) \xrightarrow{\pi^*} H^2(E^\delta M)$, which is also denoted by π^*. The following theorem is proved in Section 7.

Theorem 2.4 *Under the setting of Proposition 2.2, the homomorphism*

$$\pi^*\colon H^2(G,\mathbb{R}) \to H^2(E^\delta M)$$

is injective and the equality $\pi^*\sigma = c_1(E^\delta TM) + \kappa\|\omega\|$ *holds in* $H^2(E^\delta M)$.

Let G_b be the subgroup of G consisting of all elements which preserve a given basepoint $b \in M$ and $D\colon G_b \to \mathrm{Sp}(2n,\mathbb{R})$ a homomorphism obtained from the differentials $d\varphi_b\colon T_bM \to T_bM$ at b for all $\varphi \in G_b$. Here we take a symplectic basis for (T_bM,ω_b), so we can consider the symplectic group $\mathrm{Sp}(2n,\mathbb{R})$ as the target of D. Let $c_1 \in H^2(\mathrm{Sp}(2n,\mathbb{R}),\mathbb{R})$ be the class in the group cohomology of $\mathrm{Sp}(2n,\mathbb{R})$ corresponding to the first Chern class of flat symplectic vector bundles. Then we have $D^*c_1 \in H^2(G_b;\mathbb{R})$. It is independent of the choice of the symplectic basis for (T_bM,ω_b). Let $\iota_b\colon G_b \hookrightarrow G$ be the inclusion. The following corollary is proved in Section 7 and is used to show that σ in the example below is nontrivial.

Corollary 2.5 *The equality* $\iota_b^*\sigma = D^*c_1$ *holds in* $H^2(G_b,\mathbb{R})$.

Next we consider the case that M is closed. Let $EM \to BG$ be the universal (M,ω)–bundle over the classifying space BG of G with the C^∞ topology and $ETM \to EM$ the tangent bundle along the fibers of EM. Since ETM is a symplectic vector bundle, we have the first Chern class $c_1(ETM) \in H^2(EM)$ and its $(n+1)^{\text{st}}$ power $c_1(ETM)^{n+1} \in H^{2(n+1)}(EM)$. Since M is closed, the integration along the fibers induces a homomorphism $\pi_!\colon H^{2(n+1)}(EM) \to H^2(BG)$, which is also called the Gysin homomorphism. Put $c = \pi_!(c_1(ETM)^{n+1}) \in H^2(BG)$. It is one of the characteristic classes of (M,ω)–bundles with structure group G. Since the identity map of G induces a continuous homomorphism from G^δ to G with the C^∞ topology,

and hence a continuous map $\iota\colon BG^\delta \to BG$, we have its induced homomorphism $\iota^*\colon H^2(BG) \to H^2(BG^\delta) = H^2(G,\mathbb{R})$. The following theorem is also proved in Section 7.

Theorem 2.6 *If, in the setting of Proposition 2.2, M is in addition supposed to be closed, then the equality $\iota^* c = (n+1)c_1(M)^n([M])\sigma$ holds in $H^2(G,\mathbb{R})$, where $c_1(M)^n([M])$ is the evaluation of $c_1(M)^n \in H^{2n}(M)$ on the fundamental class $[M]$ of M.*

For a subgroup $K \subset G$, let $\iota_K^*\colon H^2(BG) \xrightarrow{\iota^*} H^2(G,\mathbb{R}) \to H^2(K,\mathbb{R})$ be the composite map of ι^* with the induced homomorphism by the inclusion $K \hookrightarrow G$. For a closed manifold M, the following corollary is just Theorem 2.3 in terms of $c \in H^2(BG)$ using Theorem 2.6.

Corollary 2.7 *Under the assumptions of Theorem 2.6, the equality $\iota_K^* c = 0$ holds in $H^2(K,\mathbb{R})$ for any subgroup K of G_j or G_J in Theorem 2.3.*

Finally we give a nontrivial example of σ. Let Σ_g be a closed oriented surface of genus $g \geqq 2$ and $\Sigma_{g,1} \subset \Sigma_g$ the complement of an open disk in Σ_g. The moduli space $\mathcal{R}_g$ of $\mathrm{SU}(2)$–representations of $\pi_1(\Sigma_{g,1})$ with holonomy $-I$ along the boundary $\partial\Sigma_{g,1}$ is given by

$$\mathcal{R}_g = \{\phi \in \mathrm{Hom}(\pi_1(\Sigma_{g,1}), \mathrm{SU}(2)) \mid \phi(\partial\Sigma_{g,1}) = -I\}/\mathrm{Ad}\,\mathrm{SU}(2),$$

where $\partial\Sigma_{g,1}$ also presents the homotopy class of itself. It is known to be a smooth, closed, 1–connected symplectic manifold of dimension $6g-6$ with symplectic form ω, whose cohomology class $[\omega]$ is twice the generator of $H^2(\mathcal{R}_g;\mathbb{Z}) \cong \mathbb{Z}$ (see [24] for properties of $\mathcal{R}_g$). Thus the assumptions of Proposition 2.2 are satisfied for $\mathcal{R}_g$. In particular we have a relation $c_1(\mathcal{R}_g) + \kappa[\omega] = 0$ for some $0 \neq \kappa \in \mathbb{R}$. Therefore we have our cohomology class $\sigma \in H^2(\mathrm{Symp}(\mathcal{R}_g,\omega),\mathbb{R})$.

Let $\Gamma_{g,1}$ be the mapping class group of $\Sigma_{g,1}$, that is, the group of the isotopy classes of diffeomorphisms of $\Sigma_{g,1}$ whose restrictions to $\partial\Sigma_{g,1}$ are the identity. The action of $\Gamma_{g,1}$ on $\pi_1(\Sigma_{g,1})$ induces a homomorphism $\rho\colon \Gamma_{g,1} \to \mathrm{Symp}(\mathcal{R}_g,\omega)$. Considering extension of the diffeomorphisms by the identity on $\Sigma_g \smallsetminus \Sigma_{g,1}$, the group $\Gamma_{g,1}$ acts on $H^1(\Sigma_g;\mathbb{Z})$, hence we have a homomorphism $\tau\colon \Gamma_{g,1} \to \mathrm{Aut}(H^1(\Sigma_g;\mathbb{Z}),\cdot\cup\cdot) \cong \mathrm{Sp}(2g,\mathbb{Z}) \subset \mathrm{Sp}(2g,\mathbb{R})$ and the cohomology class $\tau^* c_1 \in H^2(\Gamma_{g,1},\mathbb{R})$, which is known to be nontrivial for $g \geqq 3$ (see [17; 22]). Here $c_1 \in H^2(\mathrm{Sp}(2g,\mathbb{R}),\mathbb{R})$ is the first Chern class appearing before. The following theorem is proved in Section 8.

Theorem 2.8 *Let $g \geqq 4$, then the equality $\rho^*\sigma = -3\tau^* c_1$ holds in $H^2(\Gamma_{g,1},\mathbb{R})$. Hence σ and $\iota^* c \in H^2(\mathrm{Symp}(\mathcal{R}_g,\omega),\mathbb{R})$ are nontrivial.*

This theorem implies that there is a foliated symplectic bundle over a surface with closed simply connected fiber $\mathcal{R}_g$ and with nontrivial smooth characteristic class given by fiber integration.

3 Proofs of Proposition 2.2 and Theorem 2.3

In this section we prove Proposition 2.2 and Theorem 2.3. We return to the situation above Lemma 2.1.

Lemma 3.1 (1) *The 1–form $f(\varphi)$ on $\mathcal{J}$ is closed for each $\varphi \in G$.*

(2) *The map $f\colon G \to Z^1(\mathcal{J})$, $\varphi \mapsto f(\varphi)$, is a crossed homomorphism.*

By definition, f is a crossed homomorphism if and only if it satisfies $\delta f(\varphi, \psi) := \varphi_*\{f(\psi)\} - f(\varphi\psi) + f(\varphi) = 0$ for any $\varphi, \psi \in G$, that is, it is a 1–cocycle in group cohomology theory. Another choice of connection B and $\nu \in \Omega^1(\mathcal{J})$ instead of A and μ gives a crossed homomorphism

$$g\colon G \to \Omega^1(\mathcal{J}), \quad g(\varphi) = c_1(\varphi_* B - B) + \varphi_*\nu - \nu.$$

Put $\alpha = \nu - \mu + c_1(B - A) \in \Omega^1(\mathcal{J})$.

Lemma 3.2 (1) *α is a closed 1–form on $\mathcal{J}$*

(2) *$g(\varphi) - f(\varphi) = \varphi_*\alpha - \alpha =: \delta\alpha(\varphi)$ for each $\varphi \in G$.*

Lemmas 3.1 and 3.2 are proved by direct computation using the following formulas. Their detailed proofs are omitted because they were proved in a general setting in [15]. Let A, B and C be $U(n)$–connections on p^*TM; then we have

(1) $c_1(A - B) + c_1(B - C) = c_1(A - C)$,

(2) $dc_1(B - A) = c_1(F_B) - c_1(F_A)$, and

(3) $c_1(F_{\varphi_* A}) = \varphi_* c_1(F_A)$ for each $\varphi \in G$.

Here we note that we consider $c_1(A - B)$, $c_1(F_A)$ and others as forms on $\mathcal{J}$, not on the total space $\mathcal{P}$ of the $U(n)$–principal bundle $\mathcal{P} \to \mathcal{P}/U(n) = \mathcal{J}$ associated with p^*TM.

Proof of Lemma 2.1 and Proposition 2.2 Since a crossed homomorphism is nothing but a 1–cocycle, we have a cohomology class $[f] \in H^1(G, Z^1(\mathcal{J}))$ by Lemma 3.1. Lemma 3.2 shows that the difference between 1–cocycles f and g is a coboundary. This implies that the class $[f]$ is independent of the construction, giving Lemma 2.1. Proposition 2.2 is a direct consequence of the uniqueness of the class $[f]$. □

The second cohomology class σ defined in Proposition 2.2 is explicitly given as follows. Since the target of f is $B^1(\mathcal{J})$, there exists a map

$$h\colon G \to \Omega^0(\mathcal{J}) \tag{3-1}$$

satisfying $dh(\varphi) = f(\varphi)$ for each $\varphi \in G$. Let

$$s(\varphi, \psi) := \delta h(\varphi, \psi) := \varphi_* h(\psi) - h(\varphi\psi) + h(\varphi)$$

for each $\varphi, \psi \in G$; then we have a 2–cocycle $s\colon G^2 \to \mathbb{R}$ and its cohomology class is $\sigma = [s] \in H^2(G, \mathbb{R})$. Here we note that $\mathbb{R}$ is identified with the constant functions in $\Omega^0(\mathcal{J})$.

Proof of Theorem 2.3 We have only to prove the theorem for $K = G_j$ and $K = G_J$.

In the case of $K = G_j$, we can take a map h in (3-1) satisfying the additional condition $h(\varphi)(j) = 0$ for all $\varphi \in G$. Such an h is uniquely determined. For the 2–cocycle $s = \delta h\colon G^2 \to \mathbb{R}$ and any $\varphi, \psi \in G_j$, we have

$$s(\varphi, \psi) = s(\varphi, \psi)(j) = \varphi_*\{h(\psi)\}(j) = h(\psi)(\varphi^{-1}(j)) = h(\psi)(j) = 0,$$

where the first equality is the identification of a constant function with its value at $j \in \mathcal{J}$. Thus we have $\iota^*\sigma = 0$ in $H^2(G_j, \mathbb{R})$.

In the case of $K = G_J$, we can take a map h in (3-1) satisfying the condition $\int_M J^*\{h(\varphi)\}\omega^n = 0$ for all $\varphi \in G$, where J is considered as a map $J\colon M \to \mathcal{J}$. Such an h is also unique. For the 2–cocycle $s = \delta h$, we have

$$\begin{aligned} s(\varphi, \psi) \int_M \omega^n &= \int_M J^*\{s(\varphi, \psi)\}\omega^n = \int_M J^*\{\varphi_* h(\psi)\}\omega^n \\ &= \int_M \varphi_*\{J^* h(\psi)\omega^n\} = \int_M J^*\{h(\psi)\}\omega^n = 0 \end{aligned}$$

for any $\varphi, \psi \in G_J$, which implies $\iota^*\sigma = 0$ in $H^2(G_J, \mathbb{R})$. □

4 Simplicial manifolds

In this section we recall simplicial manifolds, simplicial forms and so on, for which we refer mainly to [6; 7]. We also introduce some simplicial manifolds needed in this paper.

A simplicial set X is a sequence $X = \{X_q\}_{q=0,1,\dots}$ of sets together with face operators $\varepsilon_i\colon X_q \to X_{q-1}$ for $i = 0, \dots, q$ and degeneracy ones $\eta_i\colon X_q \to X_{q+1}$ for $i = 0, \dots, q$

which satisfy the identities

$$\begin{aligned}
\varepsilon_i\varepsilon_j &= \varepsilon_{j-1}\varepsilon_i && \text{if } i<j,\\
\eta_i\eta_j &= \eta_{j+1}\eta_i && \text{if } i\leqq j,\\
\varepsilon_i\eta_j &= \begin{cases} \eta_{j-1}\varepsilon_i & \text{if } i<j,\\ \mathrm{id} & \text{if } i=j,\ j+1,\\ \eta_j\varepsilon_{i-1} & \text{if } i>j+1. \end{cases}
\end{aligned}$$

Moreover, if all X_q are topological spaces and all face and degeneracy operators are continuous maps, then X is called a simplicial space. Furthermore, if all X_q are smooth manifolds and all face and degeneracy operators are smooth maps, then X is called a simplicial manifold.

A simplicial n–form P on a simplicial manifold $X=\{X_q\}$ is a sequence $P=\{P_q\}_{q=0,1,\dots}$ of n–forms P_q on $\Delta^q\times X_q$ such that

$$(\varepsilon^i\times\mathrm{id})^*P_q=(\mathrm{id}\times\varepsilon_i)^*P_{q-1} \quad \text{on } \Delta^{q-1}\times X_q \text{ for } i=0,\dots,q \text{ and } q=1,2,\dots,$$

where $\Delta^q\subset\mathbb{R}^{q+1}$ is the standard q–simplex given by

$$\Delta^q=\Big\{t=(t_0,\dots,t_q)\in\mathbb{R}^{q+1}\;\Big|\; t_i\geqq 0 \text{ for } i=0,\dots,q,\ \sum_{i=0}^{q}t_i=1\Big\} \tag{4-1}$$

and $\varepsilon^i\colon\Delta^{q-1}\to\Delta^q$ is the ith face map given by

$$\varepsilon^i(t_0,\dots,t_{q-1})=(t_0,\dots,t_{i-1},0,t_i,\dots,t_{q-1})$$

for $q=0,1,\dots$.

Let $\mathbb{A}^n(X)$ be the set of simplicial n–forms on X. The exterior differentials on the $\Delta^q\times X_q$ define a differential $d\colon\mathbb{A}^n(X)\to\mathbb{A}^{n+1}(X)$ and the exterior multiplication of usual forms on them defines a multiplication $\wedge\colon\mathbb{A}^{n_1}(X)\otimes\mathbb{A}^{n_2}(X)\to\mathbb{A}^{n_1+n_2}(X)$ satisfying obvious identities. Let $\mathbb{A}^{k,l}(X)$ be the set of simplicial $k+l$–forms P whose restriction $P|_{\Delta^q\times X_q}:=P_q$ is locally of the form

$$\sum a_{i_1\dots i_k,j_1\dots j_l}\,dt_{i_1}\wedge\cdots\wedge dt_{i_k}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l},$$

where $t=(t_0,\dots,t_q)$ are the coordinates of Δ^q in (4-1), (x_j) are local ones of X_q and the $a_{i_1\dots i_k,j_1\dots j_l}$ are smooth functions on $\Delta^q\times X_q$. Let $d'\colon\mathbb{A}^{k,l}(X)\to\mathbb{A}^{k+1,l}(X)$ be the exterior differential with respect to t and let $d''\colon\mathbb{A}^{k,l}(X)\to\mathbb{A}^{k,l+1}(X)$ be the exterior differential with respect to (x_j) times $(-1)^k$. Then we have a double complex $(\mathbb{A}^{*,*}(X),d',d'')$ and its total complex $(\mathbb{A}^*(X),d)$ with $\mathbb{A}^n(X)=\bigoplus_{k+l=n}\mathbb{A}^{k,l}(X)$ and $d=d'+d''$.

We have another double complex $(\mathcal{A}^{*,*}(X), \delta', d'')$. Let $\mathcal{A}^{k,l}(X) = \Omega^l(X_k)$ be the set of l–forms on X_k. The differential $\delta'\colon \mathcal{A}^{k,l}(X) = \Omega^l(X_k) \to \mathcal{A}^{k+1,l}(X) = \Omega^l(X_{k+1})$ is given by $\delta' = \sum_{i=0}^{k+1}(-1)^i \varepsilon_i^*$ and the differential $d''\colon \mathcal{A}^{k,l}(X) = \Omega^l(X_k) \to \mathcal{A}^{k,l+1}(X) = \Omega^{l+1}(X_k)$ is given by $(-1)^k$ times the exterior differential on $\Omega^l(X_k)$. Put $\mathcal{A}^n(X) = \bigoplus_{k+l=n} \mathcal{A}^{k,l}(X)$ and $\delta = \delta' + d''$, then we have the total complex $(\mathcal{A}^*(X), \delta)$ of $(\mathcal{A}^{*,*}(X), \delta', d'')$.

Now we have two double complexes, $(\mathbb{A}^{*,*}(X), d', d'')$ and $(\mathcal{A}^{*,*}(X), \delta', d'')$. Restricting elements of $\mathbb{A}^{k,l}(X)$ to $\Delta^k \times X_k$ and integrating them over Δ^k, we obtain a map $I_{\Delta^k}\colon \mathbb{A}^{k,l}(X) \to \mathcal{A}^{k,l}(X)$ for each k. The collection $I_\Delta = \{I_{\Delta^k}\}$ clearly defines a map of double complexes. Moreover, it induces a natural isomorphism $I_\Delta = I_{\Delta,X}\colon H^*(\mathbb{A}^*(X)) \cong H^*(\mathcal{A}^*(X))$ on homology groups of their total complexes, where we use the same symbol I_Δ for the induced map.

We need the fat realization $\|X\|$ and the geometric one $|X|$, of a simplicial space $X = \{X_q\}$. They are the topological spaces given by

$$\|X\| = \coprod_{q \geqq 0} \Delta^q \times X_q / \sim$$

with the identifications

$$(\varepsilon^i t, x) \sim (t, \varepsilon_i x) \quad \text{for } t \in \Delta^{q-1},\ x \in X_q,\ i = 0, \dots, q,\ q = 1, 2, \dots, \tag{4-2}$$

and by

$$|X| = \coprod_{q \geqq 0} \Delta^q \times X_q / \sim, \sim'$$

with the identifications (4-2) and

$$(\eta^i t, x) \sim' (t, \eta_i x) \quad \text{for } t \in \Delta^{q+1},\ x \in X_q,\ i = 0, \dots, q,\ q = 0, 1, \dots.$$

In particular we have a natural map $\|X\| \to |X|$.

A simplicial map $f\colon X \to Y$ of simplicial spaces is a collection $f = \{f_q\colon X_q \to Y_q\}$ of continuous maps that commutes with the face and degeneracy operators. It induces a continuous map $\|f\|\colon \|X\| \to \|Y\|$ of fat realizations. Moreover, if X and Y are simplicial manifolds and all the f_q are differential maps, f is called a differential simplicial map. In this case, it induces homomorphisms $f^*\colon H^*(\mathcal{A}^*(Y)) \to H^*(\mathcal{A}^*(X))$ and $f^*\colon H^*(\mathbb{A}^*(Y)) \to H^*(\mathbb{A}^*(X))$ of cohomologies.

If X is a simplicial manifold, we have an isomorphism

$$\mathcal{I} = \mathcal{I}_X\colon H^*(\mathbb{A}^*(X)) \to H^*(\|X\|), \tag{4-3}$$

which is the composition of isomorphisms

$$H^*(\mathbb{A}^*(X)) \xrightarrow[\cong]{I_\Delta} H^*(\mathcal{A}^*(X)) \xrightarrow[\cong]{I_{\mathrm{dR}}} H^*(\|X\|),$$

where $I_{\mathrm{dR}} = I_{\mathrm{dR},X}$ is the isomorphism in the de Rham theorem in this version. These isomorphisms are natural, that is, the diagram

$$\begin{array}{ccccc} \mathcal{I}\colon H^*(\mathbb{A}^*(X)) & \xrightarrow{I_\Delta} & H^*(\mathcal{A}^*(X)) & \xrightarrow{I_{\mathrm{dR}}} & H^*(\|X\|) \\ {\scriptstyle f^*}\uparrow & & {\scriptstyle f^*}\uparrow & & {\scriptstyle \|f\|^*}\uparrow \\ \mathcal{I}\colon H^*(\mathbb{A}^*(Y)) & \xrightarrow{I_\Delta} & H^*(\mathcal{A}^*(Y)) & \xrightarrow{I_{\mathrm{dR}}} & H^*(\|Y\|) \end{array} \tag{4-4}$$

commutes for any differential simplicial map $f\colon X \to Y$ of simplicial manifolds. The isomorphism $\mathcal{I}$ is multiplicative with respect to the wedge product on the source and the cup product on the target.

We need a spectral sequence of the double complex $(\mathcal{A}^{k,l}(X), \delta', d'')$. Let $F^n\mathcal{A}^*(X) = \bigoplus_{p \geqq n} \mathcal{A}^{p,*}(X)$ be a subcomplex of $(\mathcal{A}^*(X), \delta)$ for each $n \geqq 0$. Then we have a filtration $\mathcal{A}^*(X) = F^0\mathcal{A}^*(X) \supseteq \cdots \supseteq F^n\mathcal{A}^*(X) \supseteq \cdots$. The spectral sequence obtained from it has $E_1^{p,q} = H^{p+q}(F^p\mathcal{A}^*(X)/F^{p+1}\mathcal{A}^*(X), \delta)$ and the differential d_1 is identified as

$$[d_1\colon E_1^{p,q} \to E_1^{p+1,q}] = [\delta'\colon H^q(\mathcal{A}^{p,*}(X), d'') \to H^q(\mathcal{A}^{p+1,*}(X), d'')]. \tag{4-5}$$

Put $F^nH^*(\mathcal{A}^*(X)) = \mathrm{Im}[H^*(F^n\mathcal{A}^*(X)) \to H^*(\mathcal{A}^*(X))]$; then the E_∞–term is given by $E_\infty^{p,q} \cong F^pH^{p+q}(\mathcal{A}^*(X))/F^{p+1}H^{p+q}(\mathcal{A}^*(X))$. Therefore we have short exact sequences $\{0\} \to F^{p+1}H^{p+q}(\mathcal{A}^*(X)) \to F^pH^{p+q}(\mathcal{A}^*(X)) \to E_\infty^{p,q} \to \{0\}$.

For any set Y, we have the simplicial set

$$NY = \{NY_q = Y\}_{q=0,1,\ldots} \tag{4-6}$$

with all face and degeneracy operators equal to the identity. If Y is a smooth manifold, we have a simplicial manifold NY.

Lemma 4.1 *For the spectral sequence for the double complex $(\mathcal{A}^{*,*}(NY), \delta', d'')$, the following hold:*

(1) $E_r^{p,q} = 0$ *for* $p \neq 0$ *and* $r = 2, 3, \ldots, \infty$.

(2) $H^q(\mathcal{A}^*(NY)) \cong E_\infty^{p,q} = E_2^{0,q} \cong H^q(Y)$.

Proof Since $\mathcal{A}^{p,q}(NY) = \Omega^q(Y)$, we have $E_1^{p,q} = H^q(\Omega^*(Y), (-1)^p d) = H^q(Y)$ and $d_1 = \delta' = \sum_{i=0}^{p+1}(-1)^i\varepsilon_i^*$ on $E_1^{p,q}$ under the identification (4-5). Since $\varepsilon_i = \mathrm{id}$ for all i, we get $d_1 = \mathrm{id}$ if p is odd and $d_1 = 0$ if p is even. Thus we have

$$(E_1^{*,q}, d_1)\colon 0 \to E_1^{0,q} \xrightarrow{0} E_1^{1,q} \xrightarrow{\mathrm{id}} E_1^{2,q} \xrightarrow{0} \cdots$$

for all q, hence $E_2^{0,q} = E_1^{0,q} = H^q(Y)$ and $E_2^{p,q} = \{0\}$ for $p \neq 0$. These imply the lemma. □

The identity map $NY_q = Y \xrightarrow{\mathrm{id}} Y$ induces a natural map $\|NY\| \to |NY| = Y$ and vertical maps in the commutative diagram

$$\begin{array}{ccccc} H^*(\Omega^*(Y)) & = & H^*(\mathcal{A}^{0,*}(NY)) & \xrightarrow{I_{\mathrm{dR}}} & H^*(Y) \\ \downarrow & & \downarrow \cong & & \downarrow \\ \mathcal{I}\colon H^*(\mathbb{A}^*(NY)) & \xrightarrow{I_\Delta} & H^*(\mathcal{A}^*(NY)) & \xrightarrow{I_{\mathrm{dR}}} & H^*(\|NY\|) \end{array} \tag{4-7}$$

Since all the horizontal maps are isomorphisms and the middle vertical map gives the isomorphism (2) in Lemma 4.1, the other vertical ones are also isomorphisms.

Next we define the simplicial manifolds needed in this paper. Recall that we have the sequence $p^*TM \to \mathcal{J} \to M$ of fiber bundles defined in Section 2. In particular, the first is a $U(n)$–vector bundle. In the rest of this section, we write simply $G = G^\delta$, where G^δ is the symplectomorphism group $G = \mathrm{Symp}(M, \omega)$ of (M, ω) with the discrete topology.

Put $N\overline{\mathcal{J}}_q = \mathcal{J} \times G^q$. Then we have the simplicial manifold $N\overline{\mathcal{J}} = \{N\overline{\mathcal{J}}_q\}_{q=0,1,\dots}$ equipped with face operators

$$\begin{gathered} \varepsilon_i\colon N\overline{\mathcal{J}}_q \to N\overline{\mathcal{J}}_{q-1}, \\ (J, \varphi_1, \dots, \varphi_q) \mapsto \begin{cases} (\varphi_1^{-1}J, \varphi_2, \dots, \varphi_q) & \text{if } i = 0, \\ (J, \varphi_1, \dots, \varphi_i\varphi_{i+1}, \dots, \varphi_q) & \text{if } 0 < i < q, \\ (J, \varphi_1, \dots, \varphi_{q-1}) & \text{if } i = q, \end{cases} \end{gathered} \tag{4-8}$$

for $q = 1, 2, \dots$ and degeneracy ones

$$\begin{gathered} \eta_i\colon N\overline{\mathcal{J}}_q \to N\overline{\mathcal{J}}_{q+1}, \\ (J, \varphi_1, \dots, \varphi_q) \mapsto (J, \varphi_1, \dots, \varphi_i, \mathrm{id}, \varphi_{i+1}, \dots, \varphi_q), \end{gathered} \qquad \text{for } i = 0, 1, \dots, q \tag{4-9}$$

for $q = 0, 1, \dots$.

For simplicity, we put $T = p^*TM$. Let $N\overline{T}_q = T \times G^q$; then the collection $N\overline{T} = \{N\overline{T}_q\}_{q=0,1,\dots}$ with ε_i and η_i given by the same expression (4-8) and (4-9) with $v \in T$ instead of $J \in \mathcal{J}$ is a simplicial manifold.

Similarly we have the simplicial manifold $N\overline{M} = \{N\overline{M}_q\}_{q=0,1,\dots}$ with $N\overline{M}_q = M \times G^q$.

Let $NG_q = G^q$; then we have the simplicial space $NG = \{NG_q\}_{q=0,1,\dots}$ equipped with face operators

$$\varepsilon_i\colon NG_q \to NG_{q-1}, \quad (\varphi_1,\dots,\varphi_q) \mapsto \begin{cases} (\varphi_2,\dots,\varphi_q) & \text{if } i=0,\\ (\varphi_1,\dots,\varphi_i\varphi_{i+1},\dots,\varphi_q) & \text{if } 0<i<q,\\ (\varphi_1,\dots,\varphi_{q-1}) & \text{if } i=q,\end{cases}$$

for $q = 1, 2, \dots$ and degeneracy ones

$$\begin{aligned} &\eta_i\colon NG_q \to NG_{q+1},\\ &(\varphi_1,\dots,\varphi_q) \mapsto (\varphi_1,\dots,\varphi_i,\mathrm{id},\varphi_{i+1},\dots,\varphi_q), \end{aligned} \qquad \text{for } i = 0,1,\dots,q$$

for $q = 0, 1, \dots$.

We note that if we consider all NG_q as 0–dimensional manifolds, namely, disjoint unions of points, we can also consider NG as a simplicial manifold. In this case the face and degeneracy operators are also given by the same formulas (4-8) and (4-9) with one point $\{*\}$ with trivial G action instead of $\mathcal{J}$.

Since the actions of G on the fiber bundles $T \to \mathcal{J} \to M \to \{*\}$ commute with the projections of them, we have a sequence of simplicial maps

$$N\overline{T} \to N\overline{\mathcal{J}} \to N\overline{M} \to NG.$$

The first map is a simplicial $U(n)$–vector bundle, and the second and the last are simplicial fiber bundles with fibers $\mathcal{J}_x$ for $x \in M$ and M, respectively.

These simplicial manifolds can be also obtained as follows. Let $N\overline{G}_q = G^{q+1}$; then we have the simplicial space $N\overline{G} = \{N\overline{G}_q\}$ with face and degeneracy operators

$$\begin{aligned} \varepsilon_i(\varphi_0,\dots,\varphi_q) &= (\varphi_0,\dots,\varphi_{i-1},\varphi_{i+1},\dots,\varphi_q),\\ \eta_i(\varphi_0,\dots,\varphi_q) &= (\varphi_0,\dots,\varphi_i,\varphi_i,\dots,\varphi_q). \end{aligned}$$

It is a simplicial principal G–bundle over NG with G–action $(\varphi_0,\dots,\varphi_q)\cdot\varphi = (\varphi_0\varphi,\dots,\varphi_q\varphi)$ for $((\varphi_0,\dots,\varphi_q),\varphi) \in N\overline{G}_q \times G$. The simplicial manifolds $N\overline{F}$ for $F = T, \mathcal{J}, M$ can be obtained as $N\overline{F} = \{N\overline{G}_q \times_G F\}$, that is, these simplicial manifolds are associated bundles with the simplicial principal G–bundle $N\overline{G}$ over NG. Here we remark that if we take appropriate identifications of $N\overline{G}_q \times_G F$ with $N\overline{F}_q = F \times G^q$, we have the expressions (4-8) and (4-9) of face and degeneracy operators of $N\overline{F}$. If we take other identifications, we of course obtain other expressions of face and degeneracy operators.

The fat realization $\|NG\|$ is known to be an explicit construction of the classifying space BG of G. The singular cohomology $H^*(\|NG\|)$ with coefficients in $\mathbb{R}$ is identified with the group cohomology $H^*(G,\mathbb{R})$. It is easy to see it if we use the de Rham

isomorphism $I_{\mathrm{dR}}\colon H^*(\mathcal{A}^*(NG)) \cong H^*(\|NG\|)$. In fact, since $\mathcal{A}^{p,q}(NG) = \{0\}$ for $q > 0$ and $d'' = 0$, we have $\mathcal{A}^p(NG) = \mathcal{A}^{p,0}(NG) = \Omega^0(NG_p) = \mathrm{Map}(G^p, \mathbb{R})$ and $\delta = \delta'$. This implies

$$\begin{aligned}
&(\delta a)(\varphi_1, \dots, \varphi_{p+1}) \\
&= \left(\sum_{i=0}^{p+1} (-1)^i \varepsilon_i^* a \right)(\varphi_1, \dots, \varphi_{p+1}) \\
&= a(\varphi_2, \dots, \varphi_{p+1}) + \sum_{i=1}^{p} (-1)^i a(\varphi_1, \dots, \varphi_i \varphi_{i+1}, \dots, \varphi_{p+1}) + (-1)^{p+1} a(\varphi_1, \dots, \varphi_p)
\end{aligned}$$

for $a \in \mathrm{Map}(G^p, \mathbb{R})$ and $(\varphi_1, \dots, \varphi_{p+1}) \in G^{p+1}$. Thus the complex $(\mathcal{A}^*(NG), \delta)$ is nothing but the cochain complex $C^*(G, \mathbb{R})$ of G with coefficients in $\mathbb{R}$ in group cohomology theory [5]. Usually the cochain complex obtained from the CW complex decomposition of $\|NG\|$ in the obvious way is identified with the cochain complex of G, but we identify $C^*(G, \mathbb{R})$ and its cohomology group $H^*(G, \mathbb{R})$ with $(\mathcal{A}^*(NG), \delta)$ and $H^*(\mathcal{A}^*(NG))$, respectively, in this paper.

5 Simplicial 2–forms

In this section we construct a simplicial 2–form P on the simplicial manifold $N\overline{\mathcal{J}}$, which corresponds to the cohomology class σ defined in Proposition 2.2. Let (M, ω) be a connected, symplectic manifold of dimension $2n$, and $N\overline{T}$ and $N\overline{\mathcal{J}}$ the simplicial manifolds defined in Section 4. We assume that $c_1(M) + \kappa[\omega] = 0$ for some $\kappa \in \mathbb{R}$ and $H^1(M) = \{0\}$.

Since $N\overline{T} \to N\overline{\mathcal{J}}$ is a simplicial $U(n)$–vector bundle, we can consider the first Chern class $c_1(N\overline{T}) \in H^2(\mathbb{A}^*(N\overline{\mathcal{J}}))$ of it (see [6; 7] for characteristic classes of simplicial vector bundles). We define some simplicial 2–forms on $N\overline{\mathcal{J}}$. Let A be the $U(n)$–connection on $T = p^*TM$ and μ the 1–form on $\mathcal{J}$ chosen in Section 2, which satisfies the equality (2-1). Then we have the crossed homomorphism $f\colon G \to Z^1(\mathcal{J})$ in Lemma 2.1 or 3.1, the map $h\colon G \to \Omega^0(\mathcal{J})$ (3-1) and the 2–cocycle $s\colon G^2 \to \mathbb{R} \subset \Omega^0(\mathcal{J})$. Let $F \in Z^1(N\overline{\mathcal{J}}_1)$ be the closed 1–form on $N\overline{\mathcal{J}}_1$ corresponding to f, namely it is defined by $F(J, \varphi) = f(\varphi)(J)$ for $(J, \varphi) \in N\overline{\mathcal{J}}_1 = \mathcal{J} \times G$. Here we remark that the letter F is used also as F_B to express the curvature of a connection B, but we don't mind it since we can easily distinguish them. Similarly let $H \in \Omega^0(N\overline{\mathcal{J}}_1)$ and $S \in \Omega^0(N\overline{\mathcal{J}}_2)$ be the functions on $N\overline{\mathcal{J}}_1$ and $N\overline{\mathcal{J}}_2$ corresponding to h and s, respectively. Then we have $dH = F$, where d is the exterior derivative on $N\overline{\mathcal{J}}_1$ as an ordinary manifold, and $dS = 0$ since the restriction of S to each component $\mathcal{J} \times \{(\varphi_1, \varphi_2)\} \subset N\overline{\mathcal{J}}_2$ is constant.

Next we consider induced homomorphisms $\Omega^0(N\overline{\mathcal{J}}_1) \xrightarrow{\varepsilon_i^*} \Omega^0(N\overline{\mathcal{J}}_2) \xrightarrow{\varepsilon_j^*} \Omega^0(N\overline{\mathcal{J}}_3)$ by the face operators ε_i of $N\overline{\mathcal{J}}$.

Lemma 5.1 *The following equalities hold:*

(1) $$(\varepsilon_i^* H)(J,\varphi_1,\varphi_2) = \begin{cases} \varphi_{1*}\{h(\varphi_2)\}(J) & \text{if } i=0,\\ h(\varphi_1\varphi_2)(J) & \text{if } i=1,\\ h(\varphi_1)(J) & \text{if } i=2,\end{cases}$$

for $\varepsilon_i\colon N\overline{\mathcal{J}}_2 \to N\overline{\mathcal{J}}_1$ and $(J,\varphi_1,\varphi_2) \in N\overline{\mathcal{J}}_2$,

(2) $\varepsilon_0^* H - \varepsilon_1^* H + \varepsilon_2^* H = S$ *in* $\Omega^0(N\overline{\mathcal{J}}_2)$, *and*

(3) $\varepsilon_0^* S - \varepsilon_1^* S + \varepsilon_2^* S - \varepsilon_3^* S = 0$ *in* $\Omega^0(N\overline{\mathcal{J}}_3)$.

Proof (1) For $i=0$, we have

$$\begin{aligned}(\varepsilon_0^* H)(J,\varphi_1,\varphi_2) &= H(\varepsilon_0(J,\varphi_1,\varphi_2)) = H(\varphi_1^{-1}J,\varphi_2)\\ &= h(\varphi_2)(\varphi_1^{-1}J) = \varphi_{1*}\{h(\varphi_2)\}(J).\end{aligned}$$

The rest are similarly obtained: (2) is the definition of s in terms of S and H; (3) means $(\delta')^2 = 0$, so it is clear or is easily checked by a direct computation using properties of the ε_i. □

For each $0 \leqq m \leqq q$, let $\tau_{i_0,i_1,\dots,i_m} = \tau^q_{i_0,i_1,\dots,i_m}\colon N\overline{\mathcal{J}}_q \to N\overline{\mathcal{J}}_m$ be the composite map defined by $\tau_{i_0,i_1,\dots,i_m} = \varepsilon_0 \dots \widehat{\varepsilon}_{i_0} \cdots \widehat{\varepsilon}_{i_m} \cdots \varepsilon_q$ for $0 \leqq i_0 < i_1 < \cdots < i_m \leqq q$, where $\widehat{\varepsilon}_{i_j}$ means that ε_{i_j} is removed and $\tau^q_{0,1,\dots,q} = \mathrm{id}$. We need the following lemma.

Lemma 5.2 *The following equalities hold for H, S and $\tau_{i,\dots} = \tau^q_{i,\dots}$:*

(1) $\tau_{i_0,i_1,\dots,i_m}\varepsilon_j = \tau_{i_0,\dots,i_{k-1},i_k+1,\dots,i_m+1}$ *for* $i_{k-1} < j \leqq i_k$ *and* $0 \leqq m \leqq q$,

(2) $\tau^*_{1,j+1}H - \tau^*_{0,j+1}H + \tau^*_{0,1}H - \tau^*_{0,1,j+1}S = 0$ *for* $1 \leqq j \leqq q-1$ *and* $q \geqq 2$, *and*

(3) $\tau^*_{0,1,l+1}S - \tau^*_{0,1,k+1}S + \tau^*_{1,k+1,l+1}S - \tau^*_{0,k+1,l+1}S = 0$ *for* $1 \leqq k < l \leqq q-1$ *and* $q \geqq 3$.

Proof Since $\varepsilon_i\varepsilon_j = \varepsilon_{j-1}\varepsilon_i$ for $i<j$, we have $\varepsilon_k\varepsilon_l = \varepsilon_l\varepsilon_{k+1}$ for $k \geqq l$. Using them, we obtain the equality (1). Similarly we can show that the left-hand sides of the equalities (2) and (3) are equal to $\tau^*_{0,1,j+1}(\varepsilon_0^*H - \varepsilon_1^*H + \varepsilon_2^*H - S)$ and $\tau^*_{0,1,k+1,l+1}(\varepsilon_2^*S - \varepsilon_3^*S + \varepsilon_0^*S - \varepsilon_1^*S)$, respectively. We get (2) and (3) by Lemma 5.1. □

Let $\overline{\omega}_q$ be the 2–form on $\Delta^q \times N\overline{\mathcal{J}}_q$ which is the pullback of the symplectic form ω on M by the composition $\Delta^q \times N\overline{\mathcal{J}}_q = \Delta^q \times \mathcal{J} \times G^q \to \mathcal{J} \xrightarrow{p} M$ of projections for each q. Moreover, we define 1–forms Q_q and R_q on $\Delta^q \times N\overline{\mathcal{J}}_q$ for each q by

$$Q_q = \sum_{j=1}^{q} dt_j \cdot \tau_{0,j}^* H + \sum_{1 \leqq k < l \leqq q} (t_k dt_l - t_l dt_k) \cdot \tau_{0,k,l}^* S,$$
$$R_q = \sum_{l=0}^{q} t_l \tau_l^* \mu,$$

where $\tau_{j,\dots} = \tau^q_{j,\dots}$, $\mu \in \Omega^1(N\overline{\mathcal{J}}_0) = \Omega^1(\mathcal{J})$ and empty sums mean zero. We remark that here and hereafter functions and forms on Δ^q and on $N\overline{\mathcal{J}}_q$ can be considered as those on $\Delta^q \times N\overline{\mathcal{J}}_q$ by pulling them back by the projections $\Delta^q \times N\overline{\mathcal{J}}_q \to \Delta^q, N\overline{\mathcal{J}}_q$.

Recall that $\tau_l \colon N\overline{\mathcal{J}}_q \to N\overline{\mathcal{J}}_0, (t_0, \dots, t_q) \in \Delta^q$ and that A is a $U(n)$–connection of the vector bundle $T = N\overline{T}_0$ over $\mathcal{J} = N\overline{\mathcal{J}}_0$. Thus the sum $B_q = \sum_{l=0}^{q} t_l \widetilde{\tau_l^* A}$ is a $U(n)$–connection of $\Delta^q \times N\overline{T}_q$, where $\widetilde{\tau_l^* A}$ is the product connection of $\tau_l^* A$ with the trivial one in the direction of Δ^q for each l. We have its first Chern form $c_1(F_{B_q}) \in \Omega^2(\Delta^q \times N\overline{\mathcal{J}}_q)$. Put

$$P_q = c_1(F_{B_q}) + \kappa \overline{\omega}_q + d(Q_q + R_q); \tag{5-1}$$

then we have $P_q \in \Omega^2(\Delta^q \times N\overline{\mathcal{J}}_q)$.

Lemma 5.3 (1) *$\overline{\omega} = \{\overline{\omega}_q\}$ and $c_1 = \{c_{1,q} = c_1(F_{B_q})\}$ are closed simplicial 2–forms on $N\overline{\mathcal{J}} = \{N\overline{\mathcal{J}}_q\}$.*

(2) *$Q = \{Q_q\}$ and $R = \{R_q\}$ are simplicial 1–forms on $N\overline{\mathcal{J}}$.*

(3) *$P = \{P_q\}$ is a closed simplicial 2–form on $N\overline{\mathcal{J}}$. Its cohomology class satisfies $[P] = [c_1 + \kappa\overline{\omega}] \in H^2(\mathbb{A}^*(N\overline{\mathcal{J}}))$ and is independent of the choice of the ingredients, in particular of A.*

Proof (1) It is clear for $\overline{\omega}$. For c_1, since $\tau_l \varepsilon_i = \tau_l$ for $l \leqq i-1$ and $\tau_l \varepsilon_i = \tau_{l+1}$ for $l \geqq i$ by Lemma 5.2, we have

$$\begin{aligned}(\mathrm{id} \times \varepsilon_i)^* B_{q-1} &= \sum_{l=0}^{q-1} t_l \widetilde{(\tau_l \varepsilon_i)^* A} = \sum_{l=0}^{i-1} t_l \widetilde{\tau_l^* A} + \sum_{l=i}^{q-1} t_l \widetilde{\tau_{l+1}^* A} \\ &= \sum_{l=0}^{i-1} t_l \widetilde{\tau_l^* A} + \sum_{l=i+1}^{q} t_{l-1} \widetilde{\tau_l^* A} = (\varepsilon^i \times \mathrm{id})^* \left(\sum_{l=0}^{q} t_l \widetilde{\tau_l^* A} \right) = (\varepsilon^i \times \mathrm{id})^* B_q\end{aligned}$$

as connections on $\Delta^{q-1}\times N\overline{T}_q$, hence we obtain $(\varepsilon^i\times\mathrm{id})^*c_{1,q}=(\mathrm{id}\times\varepsilon_i)^*c_{1,q-1}$ for each q. Thus c_1 is a simplicial 2–form on $N\overline{\mathcal{J}}$. Its closedness is obvious.

(2) On $\Delta^{q-1}\times N\overline{\mathcal{J}}_q$, for $i=1,\dots,q$ we have

$$\begin{aligned}
&(\mathrm{id}\times\varepsilon_i)^*Q_{q-1}\\
&=\sum_{j=1}^{q-1}dt_j\cdot(\tau_{0,j}\varepsilon_i)^*H+\sum_{1\leqq k<l\leqq q-1}(t_k\,dt_l-t_l\,dt_k)\cdot(\tau_{0,k,l}\varepsilon_i)^*S\\
&=\sum_{j=1}^{i-1}dt_j\cdot\tau_{0,j}^*H+\sum_{j=i}^{q-1}dt_j\cdot\tau_{0,j+1}^*H+\sum_{1\leqq i\leqq k<l\leqq q-1}(t_k\,dt_l-t_l\,dt_k)\cdot\tau_{0,k+1,l+1}^*S\\
&\quad+\sum_{1\leqq k<i\leqq l\leqq q-1}(t_k\,dt_l-t_l\,dt_k)\cdot\tau_{0,k,l+1}^*S+\sum_{1\leqq k<l<i\leqq q}(t_k\,dt_l-t_l\,dt_k)\cdot\tau_{0,k,l}^*S.
\end{aligned}$$

On the other hand, for $i=1,\dots,q$ we have

$$\begin{aligned}
&(\varepsilon^i\times\mathrm{id})^*Q_q\\
&=\sum_{j=1}^{i-1}dt_j\cdot\tau_{0,j}^*H+\sum_{j=i+1}^{q}dt_{j-1}\cdot\tau_{0,j}^*H\\
&\qquad+\Bigg(\sum_{1\leqq i<k<l\leqq q}+\sum_{1\leqq k<i<l\leqq q}+\sum_{1\leqq k<l<i\leqq q}\Bigg)(\varepsilon^i)^*(t_k\,dt_l-t_l\,dt_k)\cdot\tau_{0,k,l}^*S,
\end{aligned}$$

where we used $(\varepsilon^i)^*(t_k\,dt_l-t_l\,dt_k)=0$ for $i=k,l,$, so

$$\begin{aligned}
&(\varepsilon^i\times\mathrm{id})^*Q_q\\
&=\sum_{j=1}^{i-1}dt_j\cdot\tau_{0,j}^*H+\sum_{j=i}^{q-1}dt_j\cdot\tau_{0,j+1}^*H+\sum_{1\leqq i<k<l\leqq q}(t_{k-1}\,dt_{l-1}-t_{l-1}\,dt_{k-1})\cdot\tau_{0,k,l}^*S\\
&\quad+\sum_{1\leqq k<i<l\leqq q}(t_k\,dt_{l-1}-t_{l-1}\,dt_k)\cdot\tau_{0,k,l}^*S+\sum_{1\leqq k<l<i\leqq q}(t_k\,dt_l-t_l\,dt_k)\cdot\tau_{0,k,l}^*S,
\end{aligned}$$

which is equal to the result of the computation above if we change the variables k and l appropriately. Thus we have $(\varepsilon^i\times\mathrm{id})^*Q_q=(\mathrm{id}\times\varepsilon_i)^*Q_{q-1}$ for $i=1,\dots,q$.

For $i=0$, eliminating t_0 by use of the equality $\sum_{j=0}^{q-1}t_j=1$ on Δ^{q-1} and a similar computation as above implies

$$
\begin{aligned}
&(\varepsilon^0\times\mathrm{id})^*Q_q-(\mathrm{id}\times\varepsilon_0)^*Q_{q-1}\\
&=\sum_{j=1}^{q-1}dt_j\cdot(-\tau_{1,j+1}^*H+\tau_{0,j+1}^*H-\tau_{0,1}^*H+\tau_{0,1,j+1}^*S)\\
&\qquad+\sum_{1\leqq k<l\leqq q-1}(t_k\,dt_l-t_l\,dk_k)\\
&\qquad\qquad\cdot(-\tau_{0,1,l+1}^*S+\tau_{0,1,k+1}^*S-\tau_{1,k+1,l+1}^*S+\tau_{0,k+1,l+1}^*S)\\
&=0.
\end{aligned}
$$

Here the last equality is obtained by Lemma 5.2. Thus we have $(\varepsilon^i\times\mathrm{id})^*Q_q=(\mathrm{id}\times\varepsilon_i)^*Q_{q-1}$ on $\Delta^{q-1}\times N\overline{\mathcal{J}}_q$ for $0\leqq i\leqq q$ and $q=1,2,\dots$, hence $Q=\{Q_q\}$ is a simplicial 1–form on $N\overline{\mathcal{J}}$. R is shown to be a simplicial 1–form in a similar way as c_1 is, so its proof is omitted.

(3) This is clear from (1) and (2) of this lemma, the definition (5-1) of P_q, and properties of the first Chern form. □

Let $a_l=\tau_l^*A-\tau_0^*A$ for $l=0,1,\dots,q$. Since the τ_l^*A are connections of $N\overline{T}_q$, the a_l are $\mathrm{End}_{\mathbb{C}}N\overline{T}_q$–valued 1–forms on $N\overline{\mathcal{J}}_q$, hence the $c_1(a_l)$ are 1–forms on $N\overline{\mathcal{J}}_q$. They are considered as 1–forms on $\Delta^q\times N\overline{\mathcal{J}}_q$, as mentioned before.

Lemma 5.4 (1) $c_1(F_{B_q})=\sum_{l=0}^q t_l c_1(F_{\tau_l^*A})+\sum_{l=1}^q dt_l\wedge c_1(a_l)$.

(2) $\tau_{0,l}^*F=c_1(a_l)+\tau_l^*\mu-\tau_0^*\mu$ *for* $0\leqq l\leqq q$, *where* F *is the closed* 1*–form on* $N\overline{\mathcal{J}}_1$ *defined at the beginning of this section.*

Proof (1) On $\Delta^q\times N\overline{\mathcal{J}}_q$, since $B_q=\widetilde{\tau_0^*A}+\sum_{l=1}^q t_l a_l$, we have

$$F_{B_q}=F_{\widetilde{\tau_0^*A}}+d_{\widetilde{\tau_0^*A}}\Big(\sum_{l=1}^q t_l a_l\Big)+\Big(\sum_{l=1}^q t_l a_l\Big)\wedge\Big(\sum_{l=1}^q t_l a_l\Big),$$

hence

$$c_1(F_{B_q})=c_1\Big(F_{\widetilde{\tau_0^*A}}+d_{\widetilde{\tau_0^*A}}\Big(\sum_{l=1}^q t_l a_l\Big)\Big),$$

where $d_{\widetilde{\tau_0^*A}}$ is the covariant exterior derivative with respect to $\widetilde{\tau_0^*A}$. Since

$$d_{\widetilde{\tau_0^*A}}\Big(\sum_{l=1}^q t_l a_l\Big)=\sum_{l=1}^q(dt_l\wedge a_l+t_l d_{\tau_0^*A}a_l),$$

we obtain

$$c_1(F_{B_q})=c_1\Big(F_{\tau_0^*A}+\sum_{l=1}^q t_l d_{\tau_0^*A}a_l\Big)+\sum_{l=1}^q dt_l\wedge c_1(a_l)$$

$$= \sum_{l=0}^{q} t_l c_1(F_{\tau_0^* A} + d_{\tau_0^* A} a_l) + \sum_{l=1}^{q} dt_l \wedge c_1(a_l)$$
$$= \sum_{l=0}^{q} t_l c_1(F_{\tau_l^* A}) + \sum_{l=1}^{q} dt_l \wedge c_1(a_l),$$

where we used $\sum_{l=0}^{q} t_l = 1$, $a_0 = 0$, and $F_{\widetilde{\tau_0^* A}} = F_{\tau_0^* A}$ as forms on $\Delta^q \times N\overline{\mathcal{J}}_q$.

(2) Since $F \in \Omega^1(N\overline{\mathcal{J}}_1)$ and $\tau_{0,l}\colon N\overline{\mathcal{J}}_q \to N\overline{\mathcal{J}}_1$, we have $\tau_{0,l}^* F \in \Omega^1(N\overline{\mathcal{J}}_q)$. Put $\Phi_l := \Phi_l(\varphi_1, \dots, \varphi_q) := \varphi_1 \varphi_2 \cdots \varphi_l$ for $(J, \varphi_1, \dots, \varphi_q) \in N\overline{\mathcal{J}}_q = \mathcal{J} \times G^q$. Then we have $\tau_{0,l}(J, \varphi_1, \dots, \varphi_q) = \varepsilon_1 \cdots \widehat{\varepsilon}_l \cdots \varepsilon_q (J, \varphi_1, \dots, \varphi_q) = (J, \Phi_l)$ and

$$\tau_{0,l}^* F = f(\Phi_l) = c_1((\Phi_l)_* A - A) + (\Phi_l)_* \mu - \mu = c_1(a_l) + \tau_l^* \mu - \tau_0^* \mu$$

on $\mathcal{J} = \mathcal{J} \times \{(\varphi_1, \dots, \varphi_q)\} \subset N\overline{\mathcal{J}}_q$, where $\tau_l = \tau_l^q\colon N\overline{\mathcal{J}}_q \to N\overline{\mathcal{J}}_0 = \mathcal{J}$ is given by $\tau_l(J, \varphi_1, \dots, \varphi_q) = \Phi_l^{-1} J$ for $1 \leqq l \leqq q$ and $\tau_0(J, \varphi_1, \dots, \varphi_q) = J$. □

Lemma 5.5 $P_q = 2 \sum_{1 \leqq k < l \leqq q} dt_k \wedge dt_l \cdot \tau_{0,k,l}^* S$.

Proof On $\Delta^q \times N\overline{\mathcal{J}}_q$, we have $dR_q = \sum_{l=0}^{q} dt_l \wedge \tau_l^* \mu - \sum_{l=0}^{q} t_l c_1(F_{\tau_l^* A}) - \kappa \overline{\omega}_q$, where we used the equality (2-1) and $\sum_{l=0}^{q} t_q = 1$ on Δ^q. Using this, Lemma 5.4 and $\sum_{l=0}^{q} dt_l = 0$, we have

$$c_1(F_{B_q}) + \kappa \overline{\omega}_q + dR_q = \sum_{l=1}^{q} dt_l \wedge c_1(a_l) + \sum_{l=0}^{q} dt_l \wedge \tau_l^* \mu$$
$$= \sum_{l=1}^{q} dt_l \wedge \{c_1(a_l) + \tau_l^* \mu - \tau_0^* \mu\} = \sum_{l=1}^{q} dt_l \wedge \tau_{0,l}^* F.$$

Since $dH = F$ and $dS = 0$, we have

$$dQ_q = -\sum_{j=1}^{q} dt_j \wedge \tau_{0,j}^* F + 2 \sum_{1 \leqq k < l \leqq q} dt_k \wedge dt_l \cdot \tau_{0,k,l}^* S.$$

The definition (5-1) of P_q and the computations above imply this lemma. □

Corollary 5.6 *$P \in \mathbb{A}^{2,0}(N\overline{\mathcal{J}})$ and $I_\Delta(P) = S \in \mathcal{A}^{2,0}(N\overline{\mathcal{J}})$.*

Proof By Lemma 5.5, the former is clear. We then have the latter:

$$I_\Delta(P) = \int_{\Delta^2} P_2 = \tau_{0,1,2}^* S = S \in \Omega^0(N\overline{\mathcal{J}}_2) = \mathcal{A}^{2,0}(N\overline{\mathcal{J}})$$

by the definition of I_Δ and since $\int_{\Delta^2} dt_1 \wedge dt_2 = \frac{1}{2}$ and $\tau_{0,1,2} = \mathrm{id}$ on $N\overline{\mathcal{J}}_2$. □

6 Spectral sequence of $\mathcal{A}^{*,*}(N\overline{\mathcal{J}})$

In this section we consider the double complex $(\mathcal{A}^{*,*}(X), \delta', d'')$ in Section 4 with $X = N\overline{\mathcal{J}}$. Its spectral sequence has $E_1^{p,q} = H^q(\mathcal{A}^{p,*}(N\overline{\mathcal{J}}), d'')$ and

$$d_1 = \delta' \colon H^q(\mathcal{A}^{p,*}(N\overline{\mathcal{J}}), d'') \to H^q(\mathcal{A}^{p+1,*}(N\overline{\mathcal{J}}), d'').$$

Since $N\overline{\mathcal{J}}_p = \mathcal{J} \times G^p$ and the topology of G is the discrete one, we have the identification

$$\mathcal{A}^{p,*}(N\overline{\mathcal{J}}) = \Omega^*(N\overline{\mathcal{J}}_p) \cong \mathrm{Map}(G^p, \Omega^*(\mathcal{J})) = \mathrm{Map}(NG_p, \Omega^*(\mathcal{J})),$$

where Map denotes the set of all maps. Under this identification, the differential d'' on $\mathcal{A}^{p,*}(N\overline{\mathcal{J}})$ corresponds to $(-1)^p$ times the exterior differential on the target $\Omega^*(\mathcal{J})$ of the maps $NG_p \to \Omega^*(\mathcal{J})$. Thus we get $E_1^{p,q} = H^q(\mathcal{A}^{p,*}(N\overline{\mathcal{J}}), d'') \cong \mathrm{Map}(NG_p, H^q(\mathcal{J}))$ and the differential

$$d_1 = \delta' \colon \mathrm{Map}(NG_p, H^q(\mathcal{J})) \to \mathrm{Map}(NG_{p+1}, H^q(\mathcal{J}))$$

given by

(6-1)
$$\begin{aligned}(\delta' a)(\varphi_1, \dots, \varphi_{p+1}) &= \left(\sum_{i=0}^{p+1} (-1)^i \varepsilon_i^* a \right)(\varphi_1, \dots, \varphi_{p+1}) \\ &= \varphi_{1*} a(\varphi_2, \dots, \varphi_{p+1}) + \sum_{i=1}^{p} (-1)^i a(\varphi_1, \dots, \varphi_i \varphi_{i+1}, \dots, \varphi_{p+1}) \\ &\qquad + (-1)^{p+1} a(\varphi_1, \dots, \varphi_p)\end{aligned}$$

for $a \in \mathrm{Map}(NG_p, H^q(\mathcal{J}))$ and $(\varphi_1, \dots, \varphi_{p+1}) \in NG_{p+1} = G^{p+1}$, where the ε_i are given by (4-8).

On the other hand, we have cohomology groups of G in group theory [5]. In particular, we consider the group cohomology $H^*(G, H^q(\mathcal{J}))$ with values in the G–module $H^q(\mathcal{J})$ whose G–action is given by $(\varphi, a) \mapsto \varphi_* a := (\varphi^{-1})^* a$ for any $(\varphi, a) \in G \times H^q(\mathcal{J})$. Its cochain group is given by $C^p(G, H^q(\mathcal{J})) = \mathrm{Map}(G^p, H^q(\mathcal{J}))$ and its coboundary operator $\delta \colon C^p(G, H^q(\mathcal{J})) \to C^{p+1}(G, H^q(\mathcal{J}))$ by the same expression (6-1). Thus the differential $d_1 = \delta'$ of the spectral sequence is identified with the coboundary operator δ of the cochain complex $C^*(G, H^q(\mathcal{J}))$ and the E_2–term is given by $E_2^{p,q} = H^p(G, H^q(\mathcal{J}))$.

Let $\pi \colon N\overline{\mathcal{J}} \to N\overline{M} \to NG$ be the composition of simplicial bundle projections.

Lemma 6.1 *Assume $H^1(M)=\{0\}$, then there exists an exact sequence*

$$\{0\}\to H^2(G,\mathbb{R})\xrightarrow{\pi^*} H^2(\mathcal{A}^*(N\overline{\mathcal{J}}))\to H^2(\mathcal{J})^G, \tag{6-2}$$

where the middle map π^* is the induced homomorphism by π under the identification $H^2(G,\mathbb{R})=H^2(\mathcal{A}^*(NG))$, and $H^2(\mathcal{J})^G$ is the invariant subspace of $H^2(\mathcal{J})$ under the action of G.

Proof We consider the spectral sequence above. The assumption $H^1(\mathcal{J})\cong H^1(M)=\{0\}$ implies $E_2^{p,1}=H^p(G,H^1(\mathcal{J}))=\{0\}$ for all p. Hence we have $H^2(\mathcal{A}^*(N\overline{\mathcal{J}}))\cong E_\infty^{0,2}\oplus\{0\}\oplus E_\infty^{2,0}$, more precisely a short exact sequence

$$\{0\}\to E_\infty^{2,0}\to H^2(\mathcal{A}^*(N\overline{\mathcal{J}}))\to E_\infty^{0,2}\to\{0\}. \tag{6-3}$$

Moreover we can easily check that $E_\infty^{0,2}=E_4^{0,2}\subset E_2^{0,2}=H^0(G,H^2(\mathcal{J}))=H^2(\mathcal{J})^G$ and $E_\infty^{2,0}=E_2^{2,0}\cong H^2(G,\mathbb{R})$. Thus we obtain the exact sequence (6-2). The middle map is given as follows. By considering NG as a simplicial manifold as mentioned before, we have $H^2(\mathcal{A}^*(NG))\cong\{0\}\oplus\{0\}\oplus E_\infty^{2,0}(NG)$ from the spectral sequence $\{E_r^{p,q}(NG)\}$ for the double complex $(\mathcal{A}^{*,*}(NG),\delta',d'')$. The induced map $\pi^*\colon H^2(\mathcal{A}^*(NG))\to H^2(\mathcal{A}^*(N\overline{\mathcal{J}}))$ preserves the decompositions by the $E_\infty-$terms of them because π is a simplicial map. This and our convention $H^2(G,\mathbb{R})=H^2(\mathcal{A}^*(NG))$ imply the middle map is given as in the statement of this lemma. □

We return to the situation in Section 5, namely, we assume $c_1(M)+\kappa[\omega]=0$ and $H^1(M)=\{0\}$. Thus we have the cohomology class $[P]$ appearing in Lemma 5.3. By Corollary 5.6 and the proof of Lemma 6.1, $I_\Delta([P])$ belongs to the image of π^*.

Lemma 6.2 *The equality $\pi^*\sigma=I_\Delta([P])$ holds in $H^2(\mathcal{A}^*(N\overline{\mathcal{J}}))$. Moreover, if $[P]\neq 0\in H^2(\mathbb{A}^*(N\overline{\mathcal{J}}))$, then $\sigma\neq 0\in H^2(G,\mathbb{R})$.*

Proof The image of $s\in\mathrm{Map}(G^2,\mathbb{R})$ under the map $\mathrm{Map}(G^2,\mathbb{R})=\mathcal{A}^2(NG)\xrightarrow{\pi^*}\mathcal{A}^2(N\overline{\mathcal{J}})$ is nothing but $S\in\mathcal{A}^2(N\overline{\mathcal{J}})$. Thus we get $\pi^*\sigma=[S]$ in $H^2(\mathcal{A}^*(N\overline{\mathcal{J}}))$. Corollary 5.6 implies the required equality.

Since $I_\Delta\colon H^*(\mathbb{A}^*(N\overline{\mathcal{J}}))\to H^*(\mathcal{A}^*(N\overline{\mathcal{J}}))$ is an isomorphism and π^* is injective on the second cohomology by Lemma 6.1, we have the rest. □

7 Cohomology classes

In this section we define the cohomology class in $H^2(\|N\overline{\mathcal{J}}\|)$ corresponding to $[P]$ and prove Theorem 2.4, Corollary 2.5 and Theorem 2.6.

The fat realization $\|N\overline{T}\| \to \|N\overline{\mathcal{J}}\|$ of the simplicial $U(n)$–vector bundle $N\overline{T} \to N\overline{\mathcal{J}}$ is an ordinary $U(n)$–vector bundle. Thus we have the first Chern class $c_1(\|N\overline{T}\|) \in H^2(\|N\overline{\mathcal{J}}\|)$. On the other hand we defined closed simplicial 2–forms $\overline{\omega}$ and P on $N\overline{\mathcal{J}}$ in Lemma 5.3, hence we have cohomology classes $[\overline{\omega}]$ and $[P]$ in $H^2(\mathbb{A}^*(N\overline{\mathcal{J}}))$. Put $\|\overline{\omega}\| = \mathcal{I}([\overline{\omega}])$ and $\|P\| = \mathcal{I}([P])$ in $H^2(\|N\overline{\mathcal{J}}\|)$, where $\mathcal{I} = \mathcal{I}_{N\overline{\mathcal{J}}}$ is the isomorphism (4-3) with $X = N\overline{\mathcal{J}}$. Since the map $\mathcal{I}$ is natural and Chern–Weil theory holds for simplicial vector bundles (see [6; 7]), Lemma 5.3(3) implies

$$\|P\| = c_1(\|N\overline{T}\|) + \kappa\|\overline{\omega}\| \in H^2(\|N\overline{\mathcal{J}}\|). \tag{7-1}$$

Consider the Leray–Serre spectral sequence $\{E_r^{p,q}\}$ for the bundle $\|\pi\|\colon \|N\overline{\mathcal{J}}\| \to \|NG\|$ with fiber $\mathcal{J}$ which is the fat realization of $\pi\colon N\overline{\mathcal{J}} \to NG$. The E_2–term is given by $E_2^{p,q} = H^p(\|NG\|, \mathcal{H}^q)$, where $\mathcal{H}^q$ is the local system associated with this bundle, the coefficients of which are in $H^q(\mathcal{J})$. In the same way as the proof of Lemma 6.1, we have a short exact sequence

$$\{0\} \to E_\infty^{2,0} \to H^2(\|N\overline{\mathcal{J}}\|) \to E_\infty^{0,2} \to \{0\}, \tag{7-2}$$

$E_\infty^{2,0} = E_2^{2,0} = H^2(\|NG\|)$ and $E_\infty^{0,2} = E_4^{0,2} \subset E_2^{0,2} = H^0(\|NG\|, \mathcal{H}^2) = H^2(\mathcal{J})^G$. We then obtain the exact sequence

$$\{0\} \to H^2(\|NG\|) \xrightarrow{\|\pi\|^*} H^2(\|N\overline{\mathcal{J}}\|) \to H^2(\mathcal{J})^G, \tag{7-3}$$

where the last map is the induced homomorphism by the inclusion $\mathcal{J} \hookrightarrow \|N\overline{\mathcal{J}}\|$ of the fiber at the 0–cell $\Delta^0 \times NG_0$.

Let $\|\sigma\| \in H^2(\|NG\|)$ be the image of $\sigma \in H^2(G, \mathbb{R})$ under the map

$$H^2(G, \mathbb{R}) = H^2(\mathcal{A}^*(NG)) \xrightarrow{I_{\mathrm{dR}}} H^2(\|NG\|).$$

Lemma 7.1 *Let $\{E_r^{p,q}\}$ be the Leray–Serre spectral sequence for the bundle*

$$\|\pi\|\colon \|N\overline{\mathcal{J}}\| \to \|NG\|,$$

then the following hold in the exact sequence (7-2) *or* (7-3)*:*

(1) *$\|\overline{\omega}\| \in E_\infty^{0,2}$; precisely, the image of $\|\overline{\omega}\|$ in $E_\infty^{0,2}$ is nonzero if ω is not exact;*

(2) *$\|P\| \in E_\infty^{2,0}$; precisely, $\|P\|$ belongs to the image of $E_\infty^{2,0}$ in $H^2(\|N\overline{\mathcal{J}}\|)$; and*

(3) *$\|\pi\|^*(\|\sigma\|) = \|P\|$.*

Proof Let $N\mathcal{J}$ be the simplicial manifold (4-6) with $Y = \mathcal{J}$. The injections $N\mathcal{J}_q = \mathcal{J} \to \mathcal{J} \times G^q = N\overline{\mathcal{J}}_q$, $J \mapsto (J, \mathrm{id}, \dots, \mathrm{id})$, induce a simplicial map $\iota\colon N\mathcal{J} \to N\overline{\mathcal{J}}$.

Thus we have simplicial maps $N\mathcal{J} \xrightarrow{\iota} N\overline{\mathcal{J}} \xrightarrow{\pi} NG$. By the naturality of the de Rham isomorphism I_{dR}, the diagram

$$\begin{array}{ccccc} H^2(\mathcal{A}^*(NG)) & \xrightarrow{\pi^*} & H^2(\mathcal{A}^*(N\overline{\mathcal{J}})) & \xrightarrow{\iota^*} & H^2(\mathcal{A}^*(N\mathcal{J})) \\ \downarrow{\scriptstyle I_{\mathrm{dR}}} & & \downarrow{\scriptstyle I_{\mathrm{dR}}} & & \downarrow{\scriptstyle I_{\mathrm{dR}}} \\ H^2(\|NG\|) & \xrightarrow{\|\pi\|^*} & H^2(\|N\overline{\mathcal{J}}\|) & \xrightarrow{\|\iota\|^*} & H^2(\|N\mathcal{J}\|) \end{array} \tag{7-4}$$

commutes. Moreover we have $H^*(\mathcal{A}^*(N\mathcal{J})) \cong H^*(\mathcal{J}) \cong H^*(\|N\mathcal{J}\|)$ by Lemma 4.1 and the diagram (4-7). The commutative diagram (7-4) implies that the exact sequences (6-2) and (7-3), and hence (6-3) and (7-2), are isomorphic through I_{dR}.

Since $\overline{\omega} \in \mathbb{A}^{0,2}(N\overline{\mathcal{J}})$, we have $I_\Delta(\overline{\omega}) = p^*\omega$ in $\mathcal{A}^{0,2}(N\overline{\mathcal{J}}) = \Omega^2(\mathcal{J})$, hence $[I_\Delta(\overline{\omega})] = [p^*\omega] \neq 0$ in $E_\infty^{0,2} \subset H^2(\mathcal{J})^G$ of (6-3). Since $P \in \mathbb{A}^{2,0}(N\overline{\mathcal{J}})$ by Corollary 5.6, we have $I_\Delta(P) \in \mathcal{A}^{2,0}(N\overline{\mathcal{J}})$, hence $[I_\Delta(P)] \in E_\infty^{2,0}$ in (6-3). Thus we have $\|\overline{\omega}\| = I_{\mathrm{dR}}([I_\Delta(\overline{\omega})]) \in E_\infty^{0,2}$ and $\|P\| = I_{\mathrm{dR}}\big([I_\Delta(P)]\big) \in E_\infty^{2,0}$ in (7-2). This completes the proof of (1) and (2). (3) is clear from Lemma 6.2 and the commutativity of the first square of the diagram (7-4). □

Recall that we defined the sequence of fiber bundles $N\overline{T} \to N\overline{\mathcal{J}} \xrightarrow{\overline{\beta}} N\overline{M} \to NG$ of simplicial manifolds in Section 4, where the second simplicial map is named $\overline{\beta}$. Taking the fat realization of it, we have a sequence of fiber bundles

$$\|N\overline{T}\| \xrightarrow{\alpha} \|N\overline{\mathcal{J}}\| \xrightarrow{\beta=\|\overline{\beta}\|} \|N\overline{M}\| \xrightarrow{\gamma} \|NG\|,$$

where the projections are (re)named as written. We note that $\gamma \circ \beta = \|\pi\|$. Since the fibers of β are contractible, there exists a section $j\colon \|N\overline{M}\| \to \|N\overline{\mathcal{J}}\|$ of it. In the same way as the definition of $\overline{\omega}$, the symplectic form ω on M defines a simplicial closed 2–form $\omega' \in \mathbb{A}^2(N\overline{M})$. It satisfies $\overline{\omega} = \overline{\beta}^*\omega'$ in $\mathbb{A}^2(N\overline{\mathcal{J}})$. Put $\|\omega\| = \mathcal{I}([\omega']) \in H^2(\|N\overline{M}\|)$; then we have $j^*\|\overline{\omega}\| = \|\omega\|$.

Let $\zeta \to \|N\overline{M}\|$ be the tangent bundle along the fibers of the M–bundle $\gamma\colon \|N\overline{M}\| \to \|NG\|$. Since ζ is a symplectic vector bundle, we have the first Chern class $c_1(\zeta) \in H^2(\|N\overline{M}\|)$, which agrees with $j^*c_1(\|N\overline{T}\|)$.

Corollary 7.2 *Let $\{E_r^{p,q}\}$ be the Leray–Serre spectral sequence for the bundle*

$$\gamma\colon \|N\overline{M}\| \to \|NG\|;$$

then the following hold:

(1) $\|\omega\| \in E_\infty^{0,2}$,

(2) $\gamma^*\|\sigma\| \in E_\infty^{2,0}$, *and*

(3) $\gamma^*\|\sigma\| = c_1(\zeta) + \kappa\|\omega\|$ *in* $H^2(\|N\overline{M}\|)$.

Proof The spectral sequences $\{E_r^{p,q}\}$ in Lemma 7.1 and in this corollary are isomorphic to each other for $r \geqq 1$ under the induced homomorphisms by fiber-preserving maps $\beta\colon \|N\overline{\mathcal{J}}\| \rightleftarrows \|N\overline{M}\| : j$ over the identity of $\|NG\|$ since the fibers of β are contractible and j is a section of β. Considering the pullback by j, we obtain the results from Lemma 7.1 and the equality (7-1) because of $\|\pi\| \circ j = \gamma$. □

Proof of Theorem 2.4 By the definition of $\|\sigma\|$, Corollary 7.2(3) is nothing but the equality of Theorem 2.4, because $\gamma\colon \|N\overline{M}\| \to \|NG\|$ is a realization of $\pi\colon E^\delta M \to BG^\delta$ in Theorem 2.4. The injectivity of π^*, namely that of γ^*, is obtained from that of $\|\pi\|^*$ in the exact sequence (7-3) and the equality $\|\pi\| \circ j = \gamma$ since the section j induces an isomorphism on cohomology. □

Take a basepoint $b \in M$ and fix it. Let $G_b = \{\varphi \in G \mid \varphi(b) = b\}$ be a subgroup of G. Replacing G with G_b in the definitions of NG, $N\overline{M}$, $N\overline{\mathcal{J}}$ and $N\overline{T}$, and considering the inclusion $G_b \hookrightarrow G$, we have a sequence of simplicial fiber bundles $N\overline{T}_b \to N\overline{\mathcal{J}}_b \to N\overline{M}_b \to NG_b$ and a commutative diagram of simplicial maps

$$\begin{array}{ccccccc}
N\overline{T}_b & \longrightarrow & N\overline{\mathcal{J}}_b & \longrightarrow & N\overline{M}_b & \underset{\leftarrow s}{\longrightarrow} & NG_b \\
\downarrow & & \downarrow & & \downarrow{\scriptstyle \bar{\mu}} & & \downarrow \\
N\overline{T} & \longrightarrow & N\overline{\mathcal{J}} & \xrightarrow{\bar{\beta}} & N\overline{M} & \longrightarrow & NG
\end{array}$$

Here maps needed later are named as in the diagram and $s = \{s_q\}\colon NG_b \to N\overline{M}_b$ is the simplicial map given by

$$s_q\colon NG_{b,q} = G_b^q \to M \times G_b^q = N\overline{M}_{b,q}, \quad s_q(\varphi_1, \dots, \varphi_q) = (b, \varphi_1, \dots, \varphi_q),$$

which is a section of the simplicial fiber bundle $N\overline{M}_b \to NG_b$.

Taking the fat realization of the right square of the diagram above, we have the following commutative one:

$$\begin{array}{ccc}
\|N\overline{M}_b\| & \underset{\leftarrow \|s\|}{\xrightarrow{\gamma_b}} & \|NG_b\| \\
\downarrow{\scriptstyle \mu = \|\bar{\mu}\|} & & \downarrow{\scriptstyle \iota} \\
\|N\overline{M}\| & \xrightarrow{\gamma} & \|NG\|
\end{array} \tag{7-5}$$

Here the horizontal arrows are projections of fiber bundles and the vertical ones are inclusions, some of which are (re)named as in the diagram. The simplicial map $s\colon NG_b \to N\overline{M}_b$ induces a continuous section $\|s\|\colon \|NG_b\| \to \|N\overline{M}_b\|$ of γ_b and it satisfies $\gamma \circ \mu \circ \|s\| = \iota$.

Proof of Corollary 2.5 Corollary 2.5 is equivalent to that the equality $\iota^*(\|\sigma\|) = c_1((\mu \circ \|s\|)^*\zeta)$ holds in $H^2(\|NG_b\|)$. Thus we show this.

The equality $(\overline{\mu} \circ s)^*\omega' = 0$ in $\mathbb{A}^2(NG_b)$ is obvious. Considering the commutative diagram

$$\begin{array}{ccccc} H^*(\mathbb{A}^*(N\overline{M})) & \xrightarrow{\overline{\mu}^*} & H^*(\mathbb{A}^*(N\overline{M}_b)) & \xrightarrow{s^*} & H^*(\mathbb{A}^*(NG_b)) \\ \mathcal{I}\downarrow & & \mathcal{I}\downarrow & & \mathcal{I}\downarrow \\ H^*(\|N\overline{M}\|) & \xrightarrow{\mu^*=\|\overline{\mu}\|^*} & H^*(\|N\overline{M}_b\|) & \xrightarrow{\|s\|^*} & H^*(\|NG_b\|) \end{array}$$

obtained from the naturality (4-4) of $\mathcal{I}$ for simplicial maps, we have

$$\|s\|^*\mu^*(\|\omega\|) = \|s\|^*\mu^*\mathcal{I}([\omega']) = \mathcal{I}((\overline{\mu} \circ s)^*[\omega']) = 0.$$

This implies

$$\iota^*(\|\sigma\|) = \|s\|^*\mu^*\gamma^*(\|\sigma\|) = \|s\|^*\mu^*(c_1(\zeta) + \kappa\|\omega\|) = c_1\big((\mu \circ \|s\|)^*\zeta\big),$$

where we used the commutative diagram (7-5) and (3) of Corollary 7.2. □

Remark 7.3 Let $[\mathfrak{G}]$ be the second cohomology class of G for an exact symplectic manifold with trivial first real cohomology group defined by Ismagilov, Losik and Michor [12] mentioned in the beginning of the introduction. Here the symbol $[\mathfrak{G}]$ is that of Gal and Kędra in [8]. We remark that $[\mathfrak{G}]$ is generally different from our cohomology class σ even if the first Chern class of the exact symplectic manifold is trivial. In this case, we can take any $\kappa \in \mathbb{R}$. We apply Corollary 2.5 to the standard symplectic vector space $\mathbb{R}^{2n}$ with origin b. The inclusion $\iota\colon \mathrm{Sp}(2n,\mathbb{R}) \hookrightarrow G = \mathrm{Symp}(\mathbb{R}^{2n})$ factors through $\iota_b\colon G_b \hookrightarrow G$. Since the restriction of D to $\mathrm{Sp}(2n,\mathbb{R})$ is the identity map under the obvious identification, we have $\iota^*\sigma = c_1$ in $H^2(\mathrm{Sp}(2n,\mathbb{R}),\mathbb{R})$ by Corollary 2.5. On the other hand we have $\iota^*([\mathfrak{G}]) = 0$ since the restriction of $[\mathfrak{G}]$ to G_b vanishes by Example 3.9(1) in [8]. Thus σ and $[\mathfrak{G}]$ are different in this example because of the term of the first Chern class.

Proof of Theorem 2.6 Let $\{E_r^{p,q}\}$ be the Leray–Serre spectral sequence for the bundle $\gamma\colon \|N\overline{M}\| \to \|NG\|$. We have $E_2^{p,q} = H^p(\|NG\|, \mathcal{H}^q)$, where $\mathcal{H}^q$ is the local system associated with this bundle, the coefficients of which are in $H^q(M)$. Since M is a $2n$–dimensional manifold, we have $E_2^{p,q} = \{0\}$ for $q > 2n$, hence

$$H^{2+2n}(\|N\overline{M}\|) = \{0\} \oplus \cdots \oplus \{0\} \oplus E_\infty^{2,2n} \oplus E_\infty^{3,2n-1} \oplus \cdots \oplus E_\infty^{2+2n,0}.$$

Thus we get the Gysin homomorphism

$$\pi_!\colon H^{2+2n}(\|N\overline{M}\|) \to E_\infty^{2,2n} \subset E_2^{2,2n} = H^2(\|NG\|, \mathcal{H}^{2n}) \cong H^2(\|NG\|),$$

where we used the fact that the local system $\mathcal{H}^{2n} = H^{2n}(M) \cong \mathbb{R}$ is trivial.

On the other hand, we have $H^2(\|N\overline{M}\|) = E_\infty^{0,2} \oplus E_\infty^{2,0} = E_4^{0,2} \oplus E_2^{2,0}$ and $c_1(\zeta) = \gamma^*\|\sigma\| - \kappa\|\omega\|$ with $\gamma^*\|\sigma\| \in E_\infty^{2,0}$ and $\|\omega\| \in E_\infty^{0,2}$ by Corollary 7.2. Then we have

$$\begin{aligned} c_1(\zeta)^{n+1} &= \sum_{k=0}^{n+1} \binom{n+1}{k} (\gamma^*\|\sigma\|)^k (-\kappa\|\omega\|)^{n+1-k} \\ &= (-\kappa)^{n+1}\|\omega\|^{n+1} + (n+1)(-\kappa)^n(\gamma^*\|\sigma\|)\|\omega\|^n + \cdots, \end{aligned}$$

where the multiplication is the cup product of $H^*(\|N\overline{M}\|)$.

Here we recall the cup product in the Leray–Serre spectral sequence [20]. The cup products $H^p(\|NG\|, \mathcal{H}^q) \otimes H^{p'}(\|NG\|, \mathcal{H}^{q'}) \to H^{p+p'}(\|NG\|, \mathcal{H}^{q+q'})$ in cohomologies with local coefficients induce products

$$\vee\colon E_r^{p,q} \otimes E_r^{p',q'} \to E_r^{p+p',q+q'}$$

for $r = 2, 3, \ldots, \infty$. Let ab be the cup product of $a \in E_\infty^{p,q}$ and $b \in E_\infty^{p',q'}$ in $H^*(\|N\overline{M}\|)$, then we have $ab = (-1)^{qp'} a \vee b$.

Taking this fact into consideration, we have $(\gamma^*\|\sigma\|)^k\|\omega\|^{n+1-k} \in E_\infty^{2k,2(n-k+1)}$ and $\|\omega\|^{n+1} = 0$. Therefore we obtain

$$\pi_!(c_1(\zeta)^{n+1}) = (n+1)(-\kappa)^n[\omega]^n([M])\|\sigma\| = (n+1)c_1(M)^n([M])\|\sigma\|$$

in $H^2(\|NG\|)$, where we remark that $\gamma^*\|\sigma\| \in E_\infty^{2,0}$ is $\|\sigma\| \in H^2(\|NG\|)$ under the identification $E_\infty^{2,0} = E_2^{2,0} = H^2(\|NG\|)$. The sequence of fiber bundles $E^\delta TM \to E^\delta M \to BG^\delta$ is isomorphic to the pullback of that of $ETM \to EM \to BG$ by the map $\iota\colon BG^\delta \to BG$ induced from the identity map of G appearing above Theorem 2.6. Thus we have $\iota^* c = \pi_!(c_1(E^\delta TM)^{n+1})$ by the naturality of the Gysin homomorphism. Since $\zeta \to \|N\overline{M}\|$ is a realization of $E^\delta TM \to E^\delta M$, we obtain Theorem 2.6 under appropriate identification. □

Example 7.4 We give an application of Corollary 2.7. We refer to [14; 1; 2] for ingredients in this example. Let $\omega_{\mathbb{C}P^i}$ be the Fubini–Study form on $\mathbb{C}P^i$ normalized so that $\frac{1}{2\pi}\int_{\mathbb{C}P^1} \omega_{\mathbb{C}P^i} = 1$ for a line $\mathbb{C}P^1 \subset \mathbb{C}P^i$ for $i = 1, 2$. Let

$$M = \{([z_0, z_1, z_2], [w_1, w_2]) \in \mathbb{C}P^2 \times \mathbb{C}P^1 \mid z_1 w_2^k = z_2 w_1^k\}$$

be the kth Hirzebruch surface for a nonnegative integer k. For any positive numbers r and s, let ω be the Kähler form on M which is the restriction to M of the Kähler form $s\pi_1^*\omega_{\mathbb{C}P^2} + r\pi_2^*\omega_{\mathbb{C}P^1}$ on $\mathbb{C}P^2 \times \mathbb{C}P^1$, where π_1 and π_2 are the projections from $\mathbb{C}P^2 \times \mathbb{C}P^1$ to the first factor $\mathbb{C}P^2$ and the second $\mathbb{C}P^1$, respectively. The effective action of $\mathbb{T} = \{(\alpha, \beta) \in \mathbb{C}^2 : |\alpha| = |\beta| = 1\}$ on M given by

$$(\alpha, \beta) \cdot ([z_0, z_1, z_2], [w_1, w_2]) = ([\alpha z_0, z_1, \beta^k z_2], [w_1, \beta w_2])$$

is a Kähler one. It is also a Hamiltonian action and its moment map is given by

$$\Phi([z_0,z_1,z_2],[w_1,w_2])=\left(\frac{-s|z_0|^2}{|z_0|^2+|z_1|^2+|z_2|^2},\frac{-sk|z_2|^2}{|z_0|^2+|z_1|^2+|z_2|^2}+\frac{r|w_1|^2}{|w_1|^2+|w_2|^2}\right).$$

Thus the Kähler manifold M with Kähler form ω endowed with this toric action is a Kähler toric variety. Its Delzant polytope $\mathrm{Im}(\Phi)$ is checked to be the quadrilateral with vertices $(s,0)$, (s,r), $(0,-sk)$ and $(0,r)$. After changing the basis of $\mathbb{T}$ and adding a constant vector to the moment map Φ, it becomes the quadrilateral with vertices $(0,0)$, $(r+ks,0)$, (r,s) and $(0,s)$ [14].

Let $E\mathbb{T}\to B\mathbb{T}$ be the universal principal $\mathbb{T}$–bundle over the classifying space $B\mathbb{T}$ of $\mathbb{T}$ with the usual topology and $E_{\mathbb{T}}M=E\mathbb{T}\times_{\mathbb{T}}M\to B\mathbb{T}$ its associate M–bundle. Then we have cohomology classes in $H^*(B\mathbb{T})$ obtained from Chern classes of the tangent bundle $E_{\mathbb{T}}TM$ along the fibers of $E_{\mathbb{T}}M$ by the fiber integration $\pi_!\colon H^*(E_{\mathbb{T}}M)\to H^{*-4}(B\mathbb{T})$. In particular we consider the fiber integral $\pi_!(c_1(E_{\mathbb{T}}TM)^3)\in H^2(B\mathbb{T})$ of the third power of the first Chern class of $E_{\mathbb{T}}TM$.

Let $r=1+\lambda-\frac{1}{2}(k-1)>0$ and $\lambda>-1$ if k is odd, and $r=1+\lambda-\frac{1}{2}k>0$ and $\lambda\geqq 0$ if k is even, and $s=1$. Then the Delzant polytope of M is the quadrilateral with vertices $(0,0)$, $\big(2+\lambda+\frac{1}{2}(k-1),0\big)$, $\big(1+\lambda-\frac{1}{2}(k-1),1\big)$, $(0,1)$ if k is odd, and with ones $(0,0)$, $\big(1+\lambda+\frac{1}{2}k,0\big)$, $\big(1+\lambda-\frac{1}{2}k,1\big)$, $(0,1)$ if k is even. If $1+\lambda>\big[\frac{1}{2}k\big]$, then Januszkiewicz and Kędra computed the fiber integrals of characteristic classes of $E_{\mathbb{T}}TM$ in terms of k and λ in Proposition 4.11 in [13], which holds more generally for symplectic toric varieties whose Delzant polytopes are given as above. Using a formula in the proposition, we obtain $\pi_!(c_1(E_{\mathbb{T}}TM)^3)=-2k(kx-2y)\in H^2(B\mathbb{T})=\mathbb{R}[x,y]$, where x and y are the cohomology classes used in the proposition.

For $k\geqq 2$, ω is not proportional to the first Chern class $c_1(M)$ of M because $c_1(M)$ takes both positive and nonpositive values on the classes of holomorphic spheres in M. Thus we take $k=1$. In this case M is also known as the one point blow-up of $\mathbb{C}P^2$. Let E be the exceptional divisor and F the inverse image of π_2 at any point of $\mathbb{C}P^1$. Since E and F are $\mathbb{C}P^1$ and their self-intersection numbers are -1 and 0, respectively, the first Chern class of M takes values 1 on E and 2 on F. On the other hand, we have $\int_E\omega=2\pi r=2\pi(1+\lambda)$ and $\int_F\omega=2\pi s=2\pi$. Therefore it must be $\lambda=-\frac{1}{2}$, which satisfies the condition of Januszkiewicz and Kędra's proposition, for $[\omega]$ to be proportional to $c_1(M)$.

We take $k=1$, $\lambda=-\frac{1}{2}$ and $s=1$, hence $r=\frac{1}{2}$, as explained above. The cohomology class $c=\pi_!(c_1(ETM)^3)\in H^2(BG)$ is mapped to nonzero the element $j^*c=-2(x-2y)\in H^2(B\mathbb{T})$ by the homomorphism j^* induced from the injection $j\colon\mathbb{T}\hookrightarrow G$. Let J be the complex structure of M. Since the action of $\mathbb{T}$ is Kähler,

the image of j is a subgroup of $G_J \subset G$. Let $\mathbb{T}^\delta$ be $\mathbb{T}$ with the discrete topology. By Corollary 2.7, the image of j^*c by the homomorphism $H^2(B\mathbb{T}) \to H^2(B\mathbb{T}^\delta)$ induced from the continuous map $\mathbb{T}^\delta \to \mathbb{T}$ is zero.

8 An example

In this section we give a nontrivial example of the class σ. Let Σ_g be a closed oriented surface of genus $g \geqq 2$. Let $\Sigma_{g,1}$ be a subsurface of Σ_g which is the complement of a small open disk in Σ_g. Take a basepoint $* \in \partial\Sigma_{g,1} \subset \Sigma_g$. Put $\widetilde{\mathcal{R}}_g = \{\phi \in \mathrm{Hom}(\pi_1(\Sigma_{g,1}), \mathrm{SU}(2)) \mid \phi(\partial\Sigma_{g,1}) = -I\}$, where $\partial\Sigma_{g,1} \cong S^1$ also presents the homotopy class of itself in $\pi_1(\Sigma_{g,1}) := \pi_1(\Sigma_{g,1}, *)$ and $I \in \mathrm{SU}(2)$ is the identity matrix. We consider the moduli space $\mathcal{R}_g = \widetilde{\mathcal{R}}_g/\mathrm{Ad}\,\mathrm{SU}(2)$ of $\mathrm{SU}(2)$–representations of $\pi_1(\Sigma_{g,1})$ with holonomy $-I$ along the boundary $\partial\Sigma_{g,1}$. It is known as a closed, 1–connected, symplectic manifold of real dimension $6g-6$. Its first Chern class $c_1(\mathcal{R}_g)$ is given by twice the generator of $H^2(\mathcal{R}_g;\mathbb{Z}) \cong \mathbb{Z}$ [24]. Let ω be the symplectic form of $\mathcal{R}_g$; then we have $c_1(\mathcal{R}_g) + \kappa[\omega] = 0$ in $H^2(\mathcal{R}_g)$ for some $0 \neq \kappa \in \mathbb{R}$. Therefore we have the second cohomology class $\sigma \in H^2(\mathrm{Symp}(\mathcal{R}_g, \omega), \mathbb{R})$ in Proposition 2.2.

We show the nontriviality of the class σ. Let $\Sigma_{g-1,1} \subset \Sigma_{g,1}$ be the subsurface of $\Sigma_{g,1}$ the complement of which is diffeomorphic to a torus with two open disks deleted as depicted in Figure 1. Let $\Gamma_{g,1}$ be the mapping class group of $\Sigma_{g,1}$, that is, it is the group of the isotopy classes of diffeomorphisms of $\Sigma_{g,1}$ which preserve the boundary pointwise, and $\Gamma_{g-1,1}$ that of $\Sigma_{g-1,1}$. $\Gamma_{g-1,1}$ is considered as a subgroup of $\Gamma_{g,1}$ if we extend diffeomorphisms on $\Sigma_{g-1,1}$ by the identity on $\Sigma_{g,1} \smallsetminus \Sigma_{g-1,1}$. $\Gamma_{g,1}$ acts symplectically on $\mathcal{R}_g$ by $\Gamma_{g,1} \times \mathcal{R}_g \to \mathcal{R}_g$, $(\varphi, [\phi]) \mapsto [\varphi_*\phi]$, where $\varphi_*\phi = (\varphi^{-1})^*\phi$ is the pullback of ϕ by the inverse of φ. This action induces a homomorphism $\rho\colon \Gamma_{g,1} \to \mathrm{Symp}(\mathcal{R}_g, \omega)$.

Let a_i and b_i for $1 \leqq i \leqq g$ be the elements of $\pi_1(\Sigma_{g,1})$ given as depicted in Figure 1. Since $\pi_1(\Sigma_{g,1})$ is the free group generated by them, the homomorphism $\phi_0\colon \pi_1(\Sigma_{g,1}) \to \mathrm{SU}(2)$ given by

$$\phi_0(a_i) = \phi_0(b_i) = I \quad \text{for } 1 \leqq i \leqq g-1,$$
$$\phi_0(a_g) = \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix},$$
$$\phi_0(b_g) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

is well-defined.

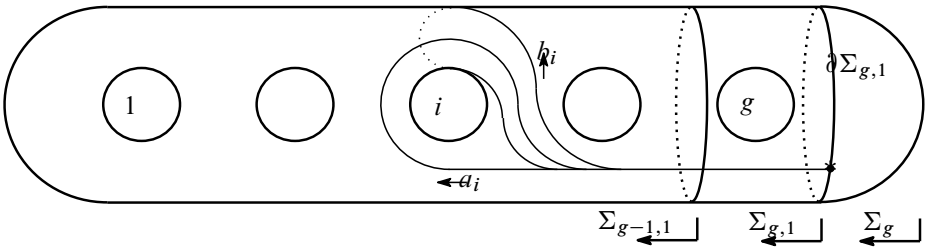


Figure 1: Surfaces and generators

Lemma 8.1 (1) *The representation ϕ_0 belongs to $\widetilde{\mathcal{R}}_g$, hence it defines a point $[\phi_0] \in \mathcal{R}_g$.*

(2) *The restriction of the action of $\Gamma_{g,1}$ on $\mathcal{R}_g$ to $\Gamma_{g-1,1}$ fixes $[\phi_0]$.*

Proof (1) We have only to show the equality $\phi_0\big(\prod_{i=1}^{g}[a_i, b_i]\big) = -I$. It is easily checked.

(2) For any $\varphi \in \Gamma_{g-1,1}$, the elements $\varphi_*(a_i)$ and $\varphi_*(b_i)$ for $1 \leqq i \leqq g-1$ belong to the subgroup S of $\pi_1(\Sigma_{g,1})$ generated by $\{a_i, b_i \mid 1 \leqq i \leqq g-1\}$, and $\varphi_*(a_g) = a_g$ and $\varphi_*(b_g) = b_g$. Since $\phi_0(S) = \{I\}$, we have $\varphi_*\phi_0 = \phi_0$. □

By Lemma 8.1, the differential of the action of $\Gamma_{g-1,1}$ at $[\phi_0]$ induces a homomorphism $D\colon \Gamma_{g-1,1} \to \mathrm{Aut}(T_{[\phi_0]}\mathcal{R}_g, \omega_{[\phi_0]})$. For any $\phi \in \widetilde{\mathcal{R}}_g$, the tangent space $T_{[\phi]}\mathcal{R}_g$ and the symplectic form $\omega_{[\phi]}$ on it are given as follows. Let Z^1_ϕ be the vector space of Ad_ϕ–crossed homomorphisms $u\colon \pi_1(\Sigma_{g,1}) \to \mathrm{su}(2)$ satisfying $u\big(\prod_{i=1}^{g}[a_i, b_i]\big) = 0$ and B^1_ϕ that of principal ones, namely, each $u \in B^1_\phi$ is given by $u(c) = \mathrm{Ad}_{\phi(c)}\, w - w$ for $c \in \pi_1(\Sigma_{g,1})$ for some $w \in \mathrm{su}(2)$. Then we have $Z^1_\phi \supset B^1_\phi$. On the other hand, ϕ descends to an $\mathrm{SO}(3)$–representation $\bar{\phi}\colon \pi_1(\Sigma_g) := \pi_1(\Sigma_g, *) \to \mathrm{SO}(3) = \mathrm{SU}(2)/\{\pm I\}$. Let $m_{\bar\phi}$ be the $\pi_1(\Sigma_g)$–module $\mathrm{so}(3)$ with respect to the adjoint representation $\mathrm{Ad}_{\bar\phi}$ of $\bar\phi$. Under a natural identification of $\mathrm{so}(3)$ with $\mathrm{su}(2)$, the spaces $Z^1_{\bar\phi}$ of $\mathrm{Ad}_{\bar\phi}$–crossed homomorphisms from $\pi_1(\Sigma_g)$ to $\mathrm{so}(3)$ and $B^1_{\bar\phi}$ of principal ones are identified with Z^1_ϕ and B^1_ϕ, respectively, in an obvious way. Thus $T_{[\phi]}\mathcal{R}_g = Z^1_\phi/B^1_\phi \cong Z^1_{\bar\phi}/B^1_{\bar\phi} = H^1(\pi_1(\Sigma_g), m_{\bar\phi})$. Under this isomorphism, the symplectic form $\omega_{[\phi]}$ is given by $\omega_{[\phi]}(x, y) = \langle B_*(x \cup y), [z]\rangle$ for $x, y \in H^1(\pi_1(\Sigma_g), m_{\bar\phi})$. Here $z = \sum_{i=1}^{g}\{(R_{i-1}, a_i) - (R_i b_i, a_i) + (R_{i-1}a_i, b_i) - (R_i, b_i)\}$ with $R_i = \prod_{j=1}^{i}[a_j, b_j]$ for $0 \leqq i \leqq g$ is a 2–cycle representing the generator $[z] \in H_2(\pi(\Sigma_g), \mathbb{Z})$,

$$B_*(\cdot \cup \cdot)\colon H^1(\pi_1(\Sigma_g), m_{\bar\phi}) \otimes H^1(\pi_1(\Sigma_g), m_{\bar\phi}) \to H^2(\pi_1(\Sigma_g), \mathbb{R})$$

is the cup product together with coefficient pairing B, which is the Killing form, and $\langle \cdot, \cdot \rangle$ denotes the natural pairing of cohomology and homology classes [9].

Next we consider the case of $\phi = \phi_0$. Let $\mathbb{T} = \{u \in Z^1_{\phi_0} \mid u(a_g) = u(b_g) = 0\}$ be a subspace of $Z^1_{\phi_0}$. It is isomorphic to $\mathrm{su}(2)^{2(g-1)}$. We can easily check that

$$B^1_{\phi_0} = \left\{ \phi \in Z^1_{\phi_0} \;\middle|\; \phi(a_i) = \phi(b_i) = 0 \text{ for } 1 \leqq i \leqq g-1, \right.$$
$$\left. \phi(a_g) = \begin{pmatrix} 0 & -\overline{\alpha} \\ \alpha & 0 \end{pmatrix},\ \phi(b_g) = \begin{pmatrix} \sqrt{-1}b & \frac{1}{2}(\alpha - \overline{\alpha}) \\ \frac{1}{2}(\alpha - \overline{\alpha}) & -\sqrt{-1}b \end{pmatrix},\ \alpha \in \mathbb{C},\ b \in \mathbb{R} \right\}$$

and $Z^1_{\phi_0} = \mathbb{T} \oplus B^1_{\phi_0}$. Thus the projection $Z^1_{\phi_0} \to Z^1_{\phi_0}/B^1_{\phi_0}$ induces an isomorphism $\mathbb{T} \to T_{[\phi_0]}\mathcal{R}_g$.

Let H be the vector space

$$H^1(\Sigma_{g-1,1}, \partial\Sigma_{g-1,1}; \mathbb{R})$$

and Ω_H the symplectic form on H given by the cup product and the evaluation on the fundamental class $[\Sigma_{g-1,1}, \partial\Sigma_{g-1,1}]$. Let (H^3, Ω) be the direct sum of three copies of the symplectic vector space (H, Ω_H). Since $-\Omega$ is also a symplectic form on H^3, we get a symplectic vector space $(H^3, -\Omega)$. Take an orthonormal basis $\{e_1, e_2, e_3\}$ of $\mathrm{su}(2)$ with respect to minus the Killing form $-B$. We can consider H as a subspace of $\mathrm{Hom}(\pi_1(\Sigma_{g,1}, \mathbb{R}), \mathbb{R})$ by the injection

$$H \cong H^1(\Sigma_{g,1}, \Sigma_{g,1} \smallsetminus \Sigma_{g-1,1}; \mathbb{R}) \hookrightarrow H^1(\Sigma_{g,1}; \mathbb{R}) = \mathrm{Hom}(\pi_1(\Sigma_{g,1}), \mathbb{R}).$$

Then the map $H^3 \to \mathbb{T}$ given by $w \mapsto \sum_{i=1}^3 w_i e_i$ for $w = (w_1, w_2, w_3) \in H^3$ is well-defined homomorphism by the definition of ϕ_0. Let $\xi\colon H^3 \to T_{[\phi_0]}\mathcal{R}_g$ be the composition of it with the isomorphism $\mathbb{T} \cong T_{[\phi_0]}\mathcal{R}_g$ above. Since the standard action of $\Gamma_{g-1,1}$ on H is symplectic, its diagonal action on $(H^3, -\Omega)$ is also symplectic.

Lemma 8.2 *The map $\xi\colon (H^3, -\Omega) \to (T_{[\phi_0]}\mathcal{R}_g, \omega_{[\phi_0]})$ is an isomorphism of symplectic vector spaces. Moreover, the restricted action of $\Gamma_{g-1,1}$ on $T_{[\phi_0]}\mathcal{R}_g$ is isomorphic to the diagonal action of it on H^3 under the isomorphism ξ, that is, the following diagram commutes:*

$$\begin{array}{ccc} \Gamma_{g-1,1} \times H^3 & \longrightarrow & H^3 \\ \downarrow{\scriptstyle \mathrm{id}\times\xi} & & \downarrow{\scriptstyle \xi} \\ \Gamma_{g-1,1} \times T_{[\phi_0]}\mathcal{R}_g & \longrightarrow & T_{[\phi_0]}\mathcal{R}_g \end{array}$$

where the horizontal arrows are given by the actions.

Proof Since the map $H^3 \to \mathbb{T}$ above is clearly an isomorphism, so is ξ. For any $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3) \in H^3$, we have

$$\xi^*\omega_{\phi_0}(x,y) = \omega_{\phi_0}(\xi(x),\xi(y)) = \left\langle B_*\left(\left[\sum_i x_i e_i\right] \cup \left[\sum_j y_j e_j\right]\right), [z]\right\rangle$$
$$= \sum_{i,j} \langle x_i \cup y_j, [z]\rangle B(e_i,e_j) = -\sum_i \langle x_i \cup y_i, [z]\rangle = -\Omega(x,y),$$

where we note that $B(e_i,e_j) = -\delta_{ij}$ and that we can compute these equalities like those in trivial module since $\phi_0(S) = \{I\}$ and $x(c) = y(c) = 0$ for $c = a_g, b_g$. Thus we have a symplectic isomorphism ξ.

For the latter part, we have only to check $\xi(\varphi_* w) = \varphi_* \xi(w)$ for all $\varphi \in \Gamma_{g-1,1}$ and $w \in H^3$, where $\varphi_* w = (\varphi_* w_1, \varphi_* w_2, \varphi_* w_3)$. Since we can compute them like with trivial coefficients for the same reason as above, we can easily check the equality. □

Proof of Theorem 2.8 Let $\tau\colon \Gamma_{g,1} \to \mathrm{Sp}(2g,\mathbb{R})$ be the homomorphism in Theorem 2.8 for each $g \geqq 1$. Since $\mathrm{Aut}(H^1(\Sigma_g;\mathbb{Z}), \cdot \cup \cdot) = \mathrm{Aut}(H^1(\Sigma_g;\mathbb{Z}), -(\cdot \cup \cdot))$ as subsets of the automorphisms of $H^1(\Sigma_g;\mathbb{Z})$, we also have a homomorphism $\tau^-\colon \Gamma_{g,1} \to \mathrm{Aut}(H^1(\Sigma_g;\mathbb{Z}), -(\cdot \cup \cdot)) \cong \mathrm{Sp}(2g,\mathbb{Z}) \subset \mathrm{Sp}(2g,\mathbb{R})$ for each $g \geqq 1$. We simply use the symbols τ and τ^- for different g. On the other hand we denote $c_1 \in H^2(\mathrm{Sp}(2g,\mathbb{R}),\mathbb{R})$ by $c_{1,g}$ for each g. Then we have $(\tau^-)^* c_{1,g} = -\tau^* c_{1,g}$.

Applying Corollary 2.5 to the situation here, we have $\iota_b^*\sigma = D^* c_{1,3(g-1)}$ in $H^2(G_b,\mathbb{R})$, where $G = \mathrm{Symp}(\mathcal{R}_g,\omega)$ and $b = [\phi_0]$. By Lemma 8.2, we have a commutative diagram

$$\begin{array}{ccccccc}
\mathrm{Sp}(2(g-1),\mathbb{R})^3 & \xleftarrow{\Delta} & \Gamma_{g-1,1} & \xrightarrow{\iota_1} & \Gamma_{g,1} & \xrightarrow{\tau} & \mathrm{Sp}(2g,\mathbb{R}) \\
\downarrow{\scriptstyle \iota_2} & & \downarrow{\scriptstyle \rho|_{\Gamma_{g-1,1}}} & & \downarrow{\scriptstyle \rho} & & \\
\mathrm{Sp}(6(g-1),\mathbb{R}) & \xleftarrow{D} & G_b & \xrightarrow{\iota_b} & G & &
\end{array}$$

where $\Delta = \tau^- \oplus \tau^- \oplus \tau^-$, and ι_1 and ι_2 are inclusions. Thus we have

$$\iota_1^*\rho^*\sigma = (\rho|_{\Gamma_{g-1,1}})^* \iota_b^* \sigma = (\rho|_{\Gamma_{g-1,1}})^* D^* c_{1,3(g-1)} = \Delta^* \iota_2^* c_{1,3(g-1)}$$

in $H^2(\Gamma_{g-1,1},\mathbb{R})$, and

$$\Delta^*\iota_2^* c_{1,3(g-1)} = 3(\tau^-)^* c_{1,g-1} = -3\tau^* c_{1,g-1} = -3\iota_1^*\tau^* c_{1,g}$$

by properties of the first Chern class. Thus we obtain $\iota_1^*\rho^*\sigma = -3\iota_1^*\tau^* c_{1,g}$. Since $\iota_1^*\colon H^2(\Gamma_{g,1},\mathbb{R}) \to H^2(\Gamma_{g-1,1},\mathbb{R})$ is injective for $g \geqq 4$ [10; 11; 17], we have $\rho^*\sigma = -3\tau^* c_{1,g}$. The nontriviality of the class $\tau^* c_{1,g} \in H^2(\Gamma_{g,1},\mathbb{R})$ for $g \geqq 3$ is a well-known fact [22]. This completes the proof of Theorem 2.8. □

Acknowledgments The author would like to thank the referee for helpful comments for improving the manuscript and in particular for suggesting Example 7.4.

References

[1] **M Abreu**, **G Granja**, **N Kitchloo**, *Compatible complex structures on symplectic rational ruled surfaces*, Duke Math. J. 148 (2009) 539–600 MR

[2] **M Abreu**, **D McDuff**, *Topology of symplectomorphism groups of rational ruled surfaces*, J. Amer. Math. Soc. 13 (2000) 971–1009 MR

[3] **R Albuquerque**, **J Rawnsley**, *Twistor theory of symplectic manifolds*, J. Geom. Phys. 56 (2006) 214–246 MR

[4] **A Banyaga**, *The structure of classical diffeomorphism groups*, Mathematics and its Applications 400, Kluwer, Dordrecht (1997) MR

[5] **K S Brown**, *Cohomology of groups*, Graduate Texts in Mathematics 87, Springer (1982) MR

[6] **J L Dupont**, *Simplicial de Rham cohomology and characteristic classes of flat bundles*, Topology 15 (1976) 233–245 MR

[7] **J L Dupont**, *Curvature and characteristic classes*, Lecture Notes in Mathematics 640, Springer (1978) MR

[8] **S Gal**, **J Kędra**, *A two-cocycle on the group of symplectic diffeomorphisms*, Math. Z. 271 (2012) 693–706 MR

[9] **W M Goldman**, *The symplectic nature of fundamental groups of surfaces*, Adv. in Math. 54 (1984) 200–225 MR

[10] **J Harer**, *The second homology group of the mapping class group of an orientable surface*, Invent. Math. 72 (1983) 221–239 MR

[11] **J L Harer**, *Stability of the homology of the mapping class groups of orientable surfaces*, Ann. of Math. 121 (1985) 215–249 MR

[12] **R S Ismagilov**, **M Losik**, **P W Michor**, *A 2–cocycle on a symplectomorphism group*, Mosc. Math. J. 6 (2006) 307–315 MR

[13] **T Januszkiewicz**, **J Kędra**, *Characteristic classes of smooth fibrations*, preprint (2002) arXiv

[14] **Y Karshon**, *Periodic Hamiltonian flows on four-dimensional manifolds*, Mem. Amer. Math. Soc. 672, Amer. Math. Soc., Providence, RI (1999) MR

[15] **R Kasagawa**, *On crossed homomorphisms on symplectic mapping class groups*, Trans. Amer. Math. Soc. 360 (2008) 4815–4839 MR

[16] **J Kędra**, **D Kotschick**, **S Morita**, *Crossed flux homomorphisms and vanishing theorems for flux groups*, Geom. Funct. Anal. 16 (2006) 1246–1273 MR

[17] **M Korkmaz**, **A I Stipsicz**, *The second homology groups of mapping class groups of oriented surfaces*, Math. Proc. Cambridge Philos. Soc. 134 (2003) 479–489 MR

[18] **D Kotschick**, **S Morita**, *Signatures of foliated surface bundles and the symplectomorphism groups of surfaces*, Topology 44 (2005) 131–149 MR

[19] **D Kotschick**, **S Morita**, *Characteristic classes of foliated surface bundles with area-preserving holonomy*, J. Differential Geom. 75 (2007) 273–302 MR

[20] **J McCleary**, *A user's guide to spectral sequences*, 2nd edition, Cambridge Studies in Advanced Mathematics 58, Cambridge Univ. Press (2001) MR

[21] **D McDuff**, **D Salamon**, *Introduction to symplectic topology*, Clarendon, New York (1995) MR

[22] **S Morita**, *Casson invariant, signature defect of framed manifolds and the secondary characteristic classes of surface bundles*, J. Differential Geom. 47 (1997) 560–599 MR

[23] **A Reznikov**, *Continuous cohomology of the group of volume-preserving and symplectic diffeomorphisms, measurable transfer and higher asymptotic cycles*, Selecta Math. 5 (1999) 181–198 MR

[24] **I Smith**, *Floer cohomology and pencils of quadrics*, Invent. Math. 189 (2012) 149–250 MR

[25] **I Vaisman**, *Symplectic geometry and secondary characteristic classes*, Progress in Mathematics 72, Birkhäuser, Boston (1987) MR

Department of Mathematics, College of Science and Technology, Nihon University
Surugadai, Chiyoda-ku, Japan

kasagawa@math.cst.nihon-u.ac.jp

Received: 26 May 2016 Revised: 22 February 2017

msp

Algebraic & Geometric Topology 18 (2018) 221–245

On the stability of asymptotic property C for products and some group extensions

GREGORY COPELAND BELL
ANDRZEJ NAGÓRKO

We show that Dranishnikov's asymptotic property C is preserved by direct products and the free product of discrete metric spaces. In particular, if G and H are groups with asymptotic property C, then both $G \times H$ and $G * H$ have asymptotic property C. We also prove that a group G has asymptotic property C if $1 \to K \to G \to H \to 1$ is exact, asdim $K < \infty$ and H has asymptotic property C. The groups are assumed to have left-invariant proper metrics and need not be finitely generated. These results settle questions of Dydak and Virk (2016), of Bell and Moran (2015) and an open problem in topology.

54F45; 20F69

1 Introduction

Asymptotic property C is a coarse invariant of metric spaces introduced by A N Dranishnikov in 2000 [11]. Recently it received a lot of attention; see Dranishnikov and Zarichnyi [13], Dydak [14], Dydak and Virk [15], Bell and Moran [7], Beckhardt [1], Beckhardt and Goldfarb [2] and Yamauchi [30].

Asymptotic property C is a coarse analog of topological property C; see Haver [21]. Topological property C fails to be preserved by direct products in very striking ways: There is an example of a separable metrizable space X with property C whose product with the irrationals fails to have property C (under the continuum hypothesis); see Pol [24]. There is also a separable complete metric space X with property C whose square $X \times X$ fails to have property C; see Pol and Pol [25]. The question whether asymptotic property C is preserved by direct products is widely known and was published by Pearl [23]. We answer it in the affirmative in the following theorem.

Theorem 1.1 *Let X and Y be metric spaces. If X and Y have asymptotic property C, then $X \times Y$ has asymptotic property C.*

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.221

Theorem 1.1 is an analog of a theorem of D M Rohm [27, Theorem 3], which states that topological property C is preserved by a direct product if one of the factors is compact. Asymptotic property C has a compactness-like property built into the definition, which requires that the sequence of constructed families is finite.

The technique used to prove Theorem 1.1 can also be used to prove a useful fibering lemma:

Fibering lemma *If $f\colon X \to Y$ is a uniformly expansive map of metric spaces, Y has asymptotic property C, and coarse fibers of f have uniform asymptotic property C, then X has asymptotic property C.*

As an immediate consequence of this lemma, we can provide an affirmative answer to a question of J Dydak and Ž Virk [15] regarding preservation of asymptotic property C via maps with finite asymptotic dimension.

The fibering lemma is effective in settling various questions concerning asymptotic property C of countable groups or finitely generated groups. Consider the short exact sequences

$$1 \to G \to G \times H \to H \to 1$$

and

$$1 \to [\![G, H]\!] \to G * H \to G \times H \to 1$$

with groups G and H endowed with left-invariant proper metrics. We show that the surjective maps $G \times H \to H$ and $G * H \to G \times H$ satisfy the assumptions of the fibering lemma if G and H have asymptotic property C. In the latter sequence $[\![G, H]\!]$ is the commutator subgroup of $G * H$, which is quasi-isometric to a tree — see Serre [28] — hence it is of finite asymptotic dimension. Therefore the last exact sequence is a special case of the group extension

$$1 \to K \to G \to H \to 1,$$

which (as we show) satisfies assumptions of the fibering lemma if K has finite asymptotic dimension and H has asymptotic property C. These observations are summarized in Theorem 1.2.

Theorem 1.2 *Let G and H be groups with proper left-invariant metrics.*

(1) *If G and H have asymptotic property C, then $G \times H$ has asymptotic property C.*

(2) *If G and H have asymptotic property C, then $G * H$ has asymptotic property C.*

(3) *If $1 \to K \to G \to H \to 1$ is an exact sequence, H has asymptotic property C and K has finite asymptotic dimension, then G has asymptotic property C.*

Theorem 1.2(2) settles the question of whether asymptotic property C is preserved by free products, which was posed by Bell and Moran [7]. We do not consider the more general question of free products with an amalgamated subgroup.

The result about free products of groups can be extended to free products of metric spaces.

Theorem 1.3 *Let X and Y be metric spaces that are discrete in the metric sense* (see Definition 6.7)*. If X and Y have asymptotic property C, then $X * Y$ has asymptotic property C.*

The proof differs from the proof given in the case of groups; it relies on the following decomposition lemma and is of independent interest.

Decomposition lemma *Let X be a metric space. Assume that there exists k such that for each (infinite) sequence $R_1, R_2, \dots$ of real numbers there exists a finite sequence $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n$ of families of subsets of X such that*

(1) *$\bigcup_i \mathcal{U}_i$ covers X,*

(2) *$\mathcal{U}_i$ is R_i–disjoint,*

(3) *$\mathcal{U}_i$ has asymptotic dimension bounded by k uniformly.*

Then X has asymptotic property C.

Acknowledgements The authors would like to thank the anonymous referee for many helpful suggestions and improvements. The authors have also been made aware that Theorem 1.1 was shown independently by Davila [10]. This research was supported by the NCN (Narodowe Centrum Nauki) grant no. 2011/01/D/ST1/04144.

2 Preliminaries

Let U and V be subsets of a metric space X. Let $R > 0$ be a real number. The subsets U and V are said to be *R–disjoint* if $\operatorname{dist}(u, u') > R$ whenever $u \in U$ and $u' \in U'$.

A family of subsets $\mathcal{U}$ of a metric space X will be said to be *R–disjoint* if U and U' are R–disjoint whenever U and U' are distinct elements of $\mathcal{U}$. We define the *mesh* of a collection $\mathcal{U}$ of subsets of X by $\operatorname{mesh}(\mathcal{U}) = \sup\{\operatorname{diam}(U) : U \in \mathcal{U}\}$. A family of subsets $\mathcal{U}$ of a metric space X will be said to be *uniformly bounded* if $\operatorname{mesh}(\mathcal{U}) < \infty$. We will occasionally describe a family as being *B–bounded* when $\operatorname{mesh}(\mathcal{U}) \leq B$ for some $B > 0$.

Asymptotic dimension was defined by Gromov [18] as an asymptotic analog of covering dimension.

Definition 2.1 A metric space X is said to have *asymptotic dimension at most n*, $\operatorname{asdim} X \leq n$, if for every $R > 0$ there are $n + 1$ uniformly bounded families of R–disjoint sets $\mathcal{U}_0, \mathcal{U}_1, \ldots, \mathcal{U}_n$ such that $\bigcup_i \mathcal{U}_i$ covers X.

Occasionally we will need more control over the disjointness and mesh of covers of a collection of sets. We say that a family $\{X_\alpha\}_\alpha$ of subsets of a metric space X has asymptotic dimension no more than n *uniformly* if for every R there is a B such that each X_α can be covered by at most $n + 1$ R–disjoint families of B–bounded sets.

Dranishnikov defined asymptotic property C [11] as an asymptotic analog of Haver's topological property C [21].

Definition 2.2 A metric space X has *asymptotic property C* if for each (infinite) sequence $R_1, R_2, \ldots$ of real numbers there exists a finite sequence $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_n$ of families of subsets of X such that

(1) $\bigcup_i \mathcal{U}_i$ covers X,

(2) $\mathcal{U}_i$ is R_i–disjoint,

(3) $\mathcal{U}_i$ is uniformly bounded.

There is no loss of generality in assuming that the numbers R_i comprise a nondecreasing sequence. It is clear that spaces with finite asymptotic dimension have asymptotic property C; it is also true that metric spaces with asymptotic property C have Yu's property A [11, Theorem 7.11].

It is easy to see that the union of a finite collection of sets with asymptotic property C has asymptotic property C. In fact, this is true for certain types of infinite unions, too (see [7, Theorem 4.2]).

Proposition 2.3 *Let X and Y have asymptotic property C. Then $X \cup Y$ has asymptotic property C.*

Finally, we mention that the concepts of asymptotic dimension and asymptotic property C (among others) belong to the realm of so-called coarse geometry. It is possible to define these concepts in the context of the coarse category without reference to a metric (see [8; 17; 26]); however, we do not pursue this course further in this work.

Definition 2.4 Let $f\colon X \to Y$ be a map of metric spaces. We say that f is

(i) *effectively proper* if there is some proper (see Definition 2.5), nondecreasing $\rho_1\colon [0,\infty) \to [0,\infty)$ such that $\rho_1(\mathrm{dist}_X(x,x')) \le \mathrm{dist}_Y(f(x),f(x'))$;

(ii) *uniformly expansive* if there is a nondecreasing $\rho_2\colon [0,\infty) \to [0,\infty)$ such that $\mathrm{dist}_Y(f(x),f(x')) \le \rho_2(\mathrm{dist}_X(x,x'))$;

(iii) *coarsely uniform embedding* if it is both effectively proper and uniformly expansive;

(iv) *coarsely surjective* if there is some R such that for each $y \in Y$, there is some $x \in X$ for which $\mathrm{dist}_Y(y, f(x)) < R$;

(v) a *coarse equivalence* if it is a coarsely uniform embedding that is coarsely surjective.

It is straightforward to verify that asymptotic dimension and asymptotic property C are invariants of coarse equivalence.

Definition 2.5 A metric space (X, dist) is said to be *proper* if closed balls are compact. We also say that the metric itself is proper in this case. A map $f\colon X \to Y$ between topological spaces is called *proper* if inverse images of compact sets are compact.

A metric d on a group Γ will be called *left-invariant* if $\mathrm{dist}(gh, gh') = \mathrm{dist}(h,h')$ for all g, h and h' in Γ.

Let Γ be a finitely generated group. Fix a finite generating set S that is symmetric in the sense that $s \in S$ implies $s^{-1} \in S$. The pair (Γ, S) carries a natural proper left-invariant metric called the word metric. This metric is given by first defining a norm on Γ as follows: $\|g\|_S = 0$ if and only if g is the identity element; otherwise $\|g\|_S$ is the length of the shortest word in the alphabet S that presents the element g. The (left-invariant) *word metric* is given by $\mathrm{dist}_S(g,h) = \|g^{-1}h\|_S$. It is easy to see that any two word metrics corresponding to different (finite) generating sets give rise to

coarsely equivalent spaces. In the case of a general countable group Γ, which may not be finitely generated, Dranishnikov and Smith showed that up to coarse equivalence there is a unique left-invariant proper metric on Γ [12, Proposition 1.1]. Moreover, each such metric arises from a weight function as follows.

Definition 2.6 Let Γ be a countable group with (possibly infinite) generating set S. Give S the discrete topology. A *weight function* on S is a map $w\colon S \to [0,\infty)$ that is positive, proper and has the property that $w(s) = w(s^{-1})$.

Theorem 2.7 [12, Proposition 1.3] *Every weight function w on a countable group defines a proper norm* (which induces a proper left-invariant metric). *The norm is given by*

$$\|\gamma\|_w = \inf\left\{ \sum_{i=1}^{n} w(s_i) \;\middle|\; \gamma = s_1 \cdots s_n,\ s_i \in S \right\}$$

and the corresponding metric is $\operatorname{dist}_w(g,h) = \|g^{-1}h\|_w$.

Definition 2.8 An *action* of a group Γ on a topological space X is a homomorphism $\varphi\colon \Gamma \to \operatorname{Homeo}(X)$. We denote $\varphi(\gamma)(x)$ by $\gamma.x$. The *orbit* of $x_0 \in X$ is the equivalence class of all $\{x \in X : x = g.x_0 \text{ for some } g \in \Gamma\}$. The action is *transitive* if there is only one orbit. For $x \in X$, the *stabilizer* of x is denoted by Γ_x and is the subgroup of Γ defined by $\Gamma_x = \{\gamma \in \Gamma : \gamma.x = x\}$. Let Γ be a group acting on a metric space X. The action of Γ on X is said to be an *action by isometries* (or an *isometric action*) if $\varphi(\Gamma) \subset \operatorname{Isom}(X)$. For $R > 0$ define the *R–stabilizer of $x_0 \in X$* by $W_R(x_0) = \{\gamma \in \Gamma : \operatorname{dist}_X(\gamma.x_0, x_0)\} \le R$.

3 Direct products preserve asymptotic property C

Several authors have studied the so-called permanence properties — that is, the extent to which properties are preserved by unions, products, etc — of coarse invariants such as finite asymptotic dimension [4], finite decomposition complexity [19; 15], or property A [3]. The permanence properties of asymptotic property C have proved to be slightly more elusive than others, with incremental special cases being proved for unions [7] and certain types of group extensions and products [1; 2]. Indeed, although it had been conjectured for some time, only recently was a group with infinite asymptotic dimension and asymptotic property C shown to exist [30].

Let $(X, \operatorname{dist}_X)$ and $(Y, \operatorname{dist}_Y)$ be metric spaces. We can define a metric on the product $X \times Y$ by $\operatorname{dist}_{X\times Y}(x \times y, x' \times y') = \sqrt{(\operatorname{dist}_X(x,x'))^2 + (\operatorname{dist}_Y(y,y'))^2}$.

R_{21}	R_{27}	R_{34}	R_{42}	R_{51}	R_{61}	$R_{1,6}$	$R_{2,6}$	$R_{3,6}$	$R_{4,6}$	$R_{5,6}$	$R_{6,6}$
R_{15}	R_{20}	R_{26}	R_{33}	R_{41}	R_{50}	$R_{1,5}$	$R_{2,5}$	$R_{3,5}$	$R_{4,5}$	$R_{5,5}$	$R_{6,5}$
R_{10}	R_{14}	R_{19}	R_{25}	R_{32}	R_{40}	$R_{1,4}$	$R_{2,4}$	$R_{3,4}$	$R_{4,4}$	$R_{5,4}$	$R_{6,4}$
R_{6}	R_{9}	R_{13}	R_{18}	R_{24}	R_{31}	$R_{1,3}$	$R_{2,3}$	$R_{3,3}$	$R_{4,3}$	$R_{5,3}$	$R_{6,3}$
R_{3}	R_{5}	R_{8}	R_{12}	R_{17}	R_{23}	$R_{1,2}$	$R_{2,2}$	$R_{3,2}$	$R_{4,2}$	$R_{5,2}$	$R_{6,2}$
R_{1}	R_{2}	R_{4}	R_{7}	R_{11}	R_{16}	$R_{1,1}$	$R_{2,1}$	$R_{3,1}$	$R_{4,1}$	$R_{5,1}$	$R_{6,1}$

Figure 1: We arrange the given sequence as indicated and assign new labels corresponding to the row and column in which the number now appears.

The main goal of this section is to answer the following question from the Lviv Topological Seminar [23, Question 1.3, page 656]: Let X and Y be two metric spaces with asymptotic property C. Does $X \times Y$ have asymptotic property C?

Theorem 3.1 *Let X and Y be metric spaces with asymptotic property C. Then $X \times Y$ has asymptotic property C.*

The idea for the proof was inspired by the proof that a product of a compact space with (topological) property C and a space with property C has property C [27, Theorem 3].

Proof Let $R_1 \leq R_2 \leq \cdots$ be a given nondecreasing sequence of positive numbers. Rearrange the sequence into subsequences $R_{i,1} \leq R_{i,2} \leq \cdots$ using the rearrangement shown in Figure 1.

Fix $i = 1$, so that we are considering the column $R_{1,j}$ with $j = 1, 2, \ldots$. Let $\mathcal{U}_{1,1}, \mathcal{U}_{1,2}, \ldots, \mathcal{U}_{1,n_1}$ be uniformly bounded families of subsets of X such that $\mathcal{U}_{1,j}$ is $R_{1,j}$–disjoint and the union $\bigcup_j \mathcal{U}_{1,j}$ covers X. Similarly, for each $i = 2, 3, \ldots$, we construct finite collections of $R_{i,j}$–disjoint families $\mathcal{U}_{i,j}$ for $j = 1, 2, \ldots, n_i$ of uniformly bounded subsets of X whose union covers X. This construction yields a sequence of positive numbers R_{i,n_i} with the property that $R_{i,n_i} \geq R_{i,j}$ for all $j \leq n_i$. Using this sequence we obtain covers $\mathcal{V}_i$ for $i = 1, 2, \ldots, m$ of Y; see Figure 2.

Let $\mathcal{W}_{i,j} = \{U \times V \colon U \in \mathcal{U}_{i,j},\, V \in \mathcal{V}_i\}$.

Given $(x, y) \in X \times Y$ there is some i_0 such that $\mathcal{V}_{i_0}$ contains an element V containing y, since the union of the $\mathcal{V}_i$ covers Y. Then there is some j_0 such that $\mathcal{U}_{i_0,j_0}$ contains a set U containing x, since each $\bigcup_j \mathcal{U}_{i,j}$ is a cover of X for each i. Thus, $U \times V$ is some $W \in \mathcal{W}_{i_0,j_0}$ that contains (x, y). Thus, the collection $\bigcup_{i,j} \mathcal{W}_{i,j}$ covers $X \times Y$.

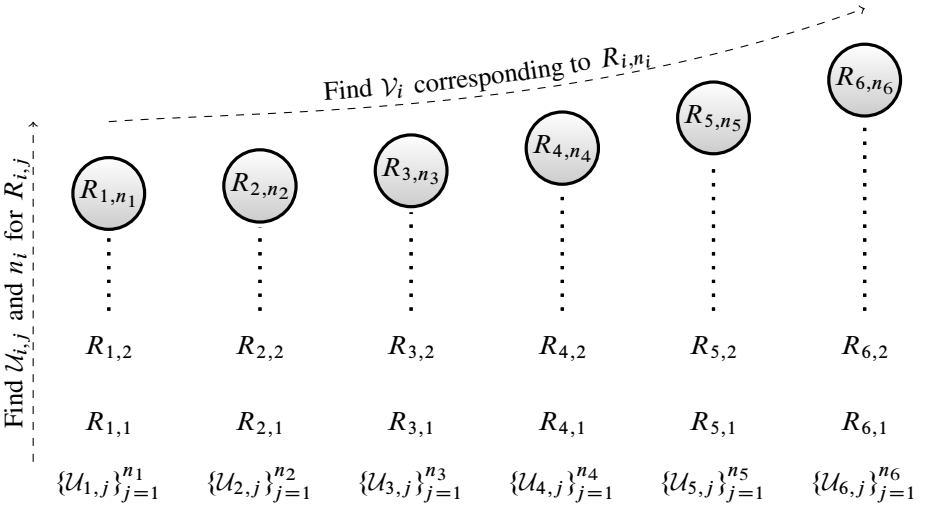


Figure 2: We find covers $\{\mathcal{U}_{i,j}\}_{j=1}^{n_i}$ of X for each column $R_{i,1}, R_{i,2}, \ldots$, and then construct a cover $\{\mathcal{V}_i\}_{i=1}^{m}$ of Y corresponding to R_{i,n_i}.

The family $\mathcal{W}_{i,j}$ is uniformly bounded; indeed, if $W \in \mathcal{W}_{i,j}$ then $\operatorname{diam}(W)^2 \leq \operatorname{mesh}(\mathcal{U}_{i,j})^2 + \operatorname{mesh}(\mathcal{V}_i)^2$.

Next, we check that each family $\mathcal{W}_{i,j}$ is $R_{i,j}$–disjoint; we suppose that distinct elements W and W' are in some fixed $\mathcal{W}_{i,j}$. Write $W = U \times V$ and $W' = U' \times V'$. Take arbitrary elements $(u, v) \in W$ and $(u', v') \in W'$. If $U = U'$ then $\operatorname{dist}((u, v), (u', v')) \geq \operatorname{dist}_Y(v, v') > R_{i,n_i}$. If $V = V'$, then $\operatorname{dist}((u, v), (u', v')) \geq \operatorname{dist}_X(u, u') > R_{i,j}$. Otherwise, $\operatorname{dist}((u, v), (u', v')) \geq \min\{\operatorname{dist}_X(u, u'), \operatorname{dist}_Y(v, v')\} > \min\{R_{i,n_i}, R_{i,j}\} = R_{i,j}$, since $R_{i,j} \leq R_{i,n_i}$ for each i and each $j \leq n_i$. Thus, $\mathcal{W}_{i,j}$ is $R_{i,j}$–disjoint.

Finally, we rearrange the $\mathcal{W}_{i,j}$ back into a single sequence as in Figure 1 to obtain $\mathcal{W}_{k_1}, \mathcal{W}_{k_2}, \ldots, \mathcal{W}_{k_N}$ (where some indices may have been omitted). Notice that this produces a finite sequence $\{k_1, k_2, \ldots, k_N\}$ with the property that $k_i \geq i$ for each i. Thus, $R_{k_i} \geq R_i$ for each i. Thus, the collection $\mathcal{W}_{k_i}$ is R_{k_i}–disjoint and therefore it is R_i–disjoint. Therefore, we have a finite collection $\mathcal{W}_{k_1}, \mathcal{W}_{k_2}, \ldots, \mathcal{W}_{k_N}$ of uniformly bounded R_i–disjoint families of subsets of $X \times Y$ with the property that $\bigcup \mathcal{W}_{k_i}$ covers $X \times Y$. □

4 The fibering lemma

In this section we prove a Hurewicz-type theorem for asymptotic property C that is motivated by the technique of the proof of Theorem 3.1. The classical Hurewicz

theorem on mappings that lower dimension goes back to 1927 and can be stated as follows:

Hurewicz theorem [16, Theorem 1.12.4] *If $f\colon X \to Y$ is a closed mapping of separable metric spaces and there is an integer $k \geq 0$ such that $\dim f^{-1}(y) \leq k$ for every $y \in Y$, then $\dim X \leq \dim Y + k$.*

This theorem was first translated to asymptotic dimension for finitely generated groups in [5]. It was proved for countable groups in [12] and translated to Assouad–Nagata dimension in [9]. A version of the theorem for metric spaces appears in [15]. In a more general context, this theorem belongs to the class of fibering theorems (see for example [19]) in which a property is preserved by a mapping when the range and fibers have this property.

4.1 Fibers with uniform asymptotic property C

Definition 4.1 Let $f\colon X \to Y$ be a map of metric spaces. Let M be a real number. We say that $A \subset X$ is a *coarse fiber at scale M* if $\operatorname{diam} f(A) < M$.

Definition 4.2 Let $f\colon X \to Y$ be a map of metric spaces. We say that *coarse fibers of f have uniform asymptotic property C* if for each infinite sequence $R_1, R_2, R_3, \ldots$ of real numbers there exists an integer k such that, for each real number M, there exists a real number B such that, for every $A \subset X$ with $\operatorname{diam} f(A) < M$, there exists a sequence $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_k$ of families of subsets of A that satisfies the following conditions:

(1) $\bigcup_{i=1}^{k} \mathcal{U}_i$ covers A.

(2) $\mathcal{U}_i$ is R_i–disjoint for each i.

(3) $\operatorname{mesh} \mathcal{U}_i \leq B$ for each i.

Comment 4.3 Note the order of quantifiers in Definition 4.2: the number k of families depends only on the sequence $R_1, R_2, \ldots$; the bound B on meshes of the families depends on the sequence $R_1, R_2, \ldots$, the number k and the scale M of fibers.

If $f\colon X \times Y \to Y$ is the projection map from the product of metric spaces $X \times Y$, then it is straightforward to show that coarse fibers of f have uniform asymptotic property C if X has asymptotic property C.

Fibering lemma *If $f\colon X \to Y$ is a uniformly expansive map of metric spaces, Y has asymptotic property C and coarse fibers of f have uniform asymptotic property C, then X has asymptotic property C.*

It is known that topological property C is preserved by preimages if the map is closed, fibers have topological property C and the codomain is paracompact; see [20].

Proof Let $\rho\colon [0,\infty) \to [0,\infty)$ be strictly increasing and such that

$$\operatorname{dist}_Y(f(x), f(x')) \leq \rho(\operatorname{dist}_X(x, x')).$$

Let $R_1, R_2, R_3, \ldots$ be an infinite sequence of real numbers. Without any loss of generality we may assume that $R_1, R_2, R_3, \ldots$ is increasing.

Rearrange the sequence $R_1 \leq R_2 \leq R_3 \leq \cdots$ into subsequences $R_{i,1} \leq R_{i,2} \leq R_{i,3} \leq \cdots$ using the rearrangement shown in Figure 1. For each i, let n_i be the number of families obtained by applying Definition 4.2 to the map f and the sequence $R_{i,1} \leq R_{i,2} \leq R_{i,3} \leq \cdots$. Consider the sequence

$$\rho(R_{1,n_1}),\ \rho(R_{2,n_2}),\ \rho(R_{3,n_3}),\ \ldots$$

Using the assumption that Y has asymptotic property C, we obtain a sequence $\mathcal{V}_1$, $\mathcal{V}_2, \ldots, \mathcal{V}_m$ of families of subsets of Y such that:

(1) $\bigcup_{i=1}^{m} \mathcal{V}_i$ covers Y.

(2) $\mathcal{V}_i$ is $\rho(R_{i,n_i})$–disjoint for each i.

(3) $\mathcal{V}_i$ is uniformly bounded for each i.

Let $M_i = \operatorname{mesh} \mathcal{V}_i$. Apply Definition 4.2 to the map f, the sequence $R_{i,1} \leq R_{i,2} \leq \cdots \leq R_{i,n_i}$ and the scale M_i to obtain a bound B_i.

Let $V \in \mathcal{V}_i$. The set $f^{-1}(V)$ is a coarse fiber of f at scale M_i. By Definition 4.2 applied to the sequence $R_{i,1} \leq R_{i,2} \leq \cdots \leq R_{i,n_i}$ and the coarse fiber $f^{-1}(V)$ at scale M_i, there exists a sequence $\mathcal{U}_1^V, \mathcal{U}_1^V, \ldots, \mathcal{U}_{n_i}^V$ of families of subsets of $f^{-1}(V)$ such that:

(1) $\bigcup_{j=1}^{n_i} \mathcal{U}_j^V$ covers $f^{-1}(V)$.

(2) $\mathcal{U}_j^V$ is $R_{i,j}$–disjoint for each j.

(3) $\operatorname{mesh} \mathcal{U}_j^V \leq B_i$ for each j.

Observe that if $V, W \in \mathcal{V}_i$ with $v \in V$, $w \in W$ and $V \neq W$, then

$$\operatorname{dist}_X(f^{-1}(v), f^{-1}(w)) > R_{i,n_i}.$$

Since $R_{i,n_i} \geq R_{i,j}$ for $j \leq n_i$, the family

$$\mathcal{U}_{i,j} = \bigcup_{V \in \mathcal{V}_i} \mathcal{U}_j^V$$

is $R_{i,j}$–disjoint and $\operatorname{mesh} \mathcal{U}_{i,j} < B_i$. Hence we obtain a collection of uniformly bounded families $\{\mathcal{U}_{i,j} : i \leq m,\ j \leq n_i\}$ that are $R_{i,j}$–disjoint and whose union covers X. Rearranging these families back into the sequence $R_1, R_2, \ldots$ (possibly leaving some places empty) shows that X has asymptotic property C. □

By the comments following Comment 4.3 we see that the product theorem (Theorem 3.1) can be deduced from the fibering lemma.

4.2 Fibers with uniformly bounded asymptotic dimension

We turn our attention to the preservation of coarse properties discussed in [15], from which we recall the following definition.

Definition 4.4 Given a function $f\colon X \to Y$ of metric spaces, its *asymptotic dimension* $\operatorname{asdim}(f)$ is the supremum of $\operatorname{asdim}(A)$ for $A \subset X$ such that $\operatorname{asdim}(f(A)) = 0$.

Lemma 4.5 *Let $f\colon X \to Y$ be a uniformly expansive map of metric spaces. The following are equivalent:*

(1) *For every $M > 0$ and every $R > 0$, there is a $B > 0$ such that if $U \subset X$ and $\operatorname{diam}(f(U)) < M$, then U can be covered by $n+1$ R–disjoint families of B–bounded subsets of X.*

(2) $\operatorname{asdim} f \leq n$.

Proof (1) $\Longrightarrow$ (2) Since f is uniformly expansive, there is an increasing control function $\rho\colon [0,\infty) \to [0,\infty)$ such that $\operatorname{dist}_Y(f(x), f(x')) \leq \rho(\operatorname{dist}_X(x,x'))$. Let $R > 0$ be given and take an arbitrary $A \subset X$ with the property that $\operatorname{asdim} f(A) = 0$. We may cover $f(A)$ with uniformly bounded sets U_i that are $\rho^{-1}(R)$–disjoint. Let M be an upper bound for $\operatorname{diam} U_i$ and take $B > 0$ as above. Then each $f^{-1}(U_i)$ can be covered by the union of $n+1$ families $\mathcal{U}_{i,j}$ of R–disjoint, uniformly B–bounded subsets of $f^{-1}(U_i)$. By the choice of the U_i, we see that if $U_i \neq U_j$ then $\operatorname{dist}(f^{-1}(u), f^{-1}(v)) > R$ whenever $u \in U_i$ and $v \in U_j$. Thus, by taking the collection $\mathcal{V}_j = \mathcal{U}_{i,j}$ for $j = 0, 1, \ldots, n$ we obtain a collection of $n+1$ R–disjoint families of

uniformly bounded sets whose union covers A. Thus, asdim $A \leq n$ and, since A was arbitrary, we conclude that asdim $f \leq n$.

(2) $\Longrightarrow$ (1) Suppose that asdim $f \leq n$. Fix $R > 0$ and $M > 0$. For each $U \subset X$ with $\operatorname{diam}(f(U)) < M$, let $B(U)$ denote the smallest integer B for which U can be expressed as a union of B–bounded, R–disjoint families $\mathcal{U}_0, \ldots, \mathcal{U}_n$.

If $\sup\{B(U) : \operatorname{diam}(f(U)) < M\} < \infty$ then there is nothing to prove. Thus, we assume that we can find a sequence of subsets $U_1, U_2, \ldots$ of X with $\operatorname{diam}(f(U_i)) < M$ for which $B(U_1), B(U_2), \ldots$ is unbounded. We may take the sequence of such $B(U_i)$ to be monotonically increasing. If the set $\{f(U_1), f(U_2), \ldots\}$ is bounded, then $\operatorname{asdim}\big(\bigcup_i f(U_i)\big) = 0$ and therefore the sequence of the $B(U_i)$ is bounded from above. Otherwise, the sequence $f(U_i)$ is unbounded. Thus, we take $i_1 = 1$. Because the $f(U_i)$ are bounded, yet the sequence $\{f(U_i)\}$ is unbounded, for each $k > 1$ we can find an increasing sequence of indices i_k such that $\operatorname{dist}\big(\bigcup_{j=1}^{k-1} f(U_{i_j}), f(U_k)\big) > 2^k$. (Here the distance between subsets is defined by $\operatorname{dist}(A, B) = \inf\{\operatorname{dist}(a, b) : a \in A, b \in B\}$.) Thus, $\bigcup_{k=1}^{\infty} f(U_{i_j})$ has asymptotic dimension 0, which contradicts the fact that the $B(U_i)$ increase monotonically to infinity since asdim $f \leq n$. □

Dydak and Virk have the following theorem regarding such maps and property C:

Theorem 4.6 [15, Theorem 6.3] *Suppose that $f\colon X \to Y$ is coarsely surjective and* $\operatorname{asdim}(f) = 0$. *Then, if Y has asymptotic property C, X has asymptotic property C.*

This naturally leads them to the following question, which we are able to settle in the affirmative [15, Question 6.4]: Suppose $f\colon X \to Y$ is uniformly expansive, coarsely surjective and of finite asymptotic dimension $\operatorname{asdim}(f)$. If Y has asymptotic property C, does X have asymptotic property C?

Corollary 4.7 *Suppose that $f\colon X \to Y$ is a uniformly expansive, coarsely surjective map of finite asymptotic dimension* $\operatorname{asdim}(f)$ *and that Y has asymptotic property C. Then X has asymptotic property C.*

Proof Suppose that $\operatorname{asdim}(f) = n$. Then, given a sequence $R_1 \leq R_2 \leq \cdots$, we take $k = n + 1$ and apply Lemma 4.5 to R_{n+1} to find that for any $M > 0$ there is a B such that whenever $\operatorname{diam} f(A) < M$, we can cover A by $n + 1$ R_{n+1}–disjoint families $\mathcal{U}_1, \ldots, \mathcal{U}_{n+1}$ that are each B–bounded. Since the sequence R_i is nondecreasing, we see that this shows that the coarse fibers of f have uniform asymptotic property C. Thus, by the fibering lemma, we conclude that X has asymptotic property C. □

5 Fibers of group homomorphisms

In the assumptions of the fibering lemma we impose the asymptotic property C condition *uniformly* on *coarse* fibers. The following example shows that this is indeed necessary. In the present section we show that, for group homomorphisms, in some cases the uniformity of the fiber condition is implicit.

Example 5.1 Let $\{0,1\}^n$ be the set of vertices of an n–dimensional cube endowed with the ℓ_1–metric. The disjoint union $\bigsqcup_{n=1}^{\infty}\{0,1\}^n$ can be metrized in such a way that it is a locally finite metric space that fails to have property A [22] and hence fails to have asymptotic property C as well [11, Theorem 7.11]. Consider the map $\bigsqcup_{n=1}^{\infty}\{0,1\}^n \to \mathbb{Z}$ that collapses $\{0,1\}^n$ into a single integer n^2. It is a uniformly expanding map onto a zero-dimensional space. Its coarse fibers are all bounded (hence they are zero-dimensional and have asymptotic property C).

5.1 Fibers vs coarse fibers

Throughout this section, G and H denote countable groups endowed with proper left-invariant metrics.

Lemma 5.2 *Let G be a countable group with a proper left-invariant metric acting by isometries on a metric space X. Let $x_0 \in X$ and define $\varphi\colon G \to X$ by $\varphi(g) = g.x_0$. Fix n and suppose that, for every $R > 0$, the R–stabilizer* (see Definition 2.8) *satisfies* $\operatorname{asdim} W_R(x_0) \le n$. *Then coarse fibers of φ have uniform asymptotic property C.*

We remark that the R–stabilizer is a coarse fiber in the sense of Definition 4.1.

Proof We may assume that the action is transitive. Suppose that a sequence $R_1 \le R_2 \le \cdots$ is given and take $k = n+1$. Now, if $M > 0$ is given, we may use the fact that $\operatorname{asdim} W_M(x_0) \le k-1$ to find families of R_k–disjoint sets $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_k$ such that $\bigcup_i \mathcal{U}_i$ covers $W_M(x_0)$ and $\operatorname{mesh}(\mathcal{U}_i) \le B$ for some B.

Let $A \subset G$ have the property that $\operatorname{diam}(\varphi(A)) \le M$. Then take any $g_A \in G$ and observe that $g_A(x_0) \in \varphi(A)$. Thus, $\varphi(A) \subset B_M(g_A.x_0)$ and so $g_A^{-1}A \subset W_M(x_0)$. Since the metric on G is left-invariant, the families $\mathcal{V}_i$ given by $\{g_A^{-1}U : U \in \mathcal{U}_i\}$ have the property that they are R_k–disjoint and B–bounded. Moreover, their union covers A, as required. □

We actually proved a stronger statement: for each R the coarse fibers at scale R have asymptotic dimension at most n uniformly.

Lemma 5.3 *Let $\varphi\colon G \to X$ describe the transitive action by isometries of a countable discrete group G on a proper discrete metric space X. If the stabilizer G_x has finite asymptotic dimension for some $x \in X$, then coarse fibers of the map $\varphi\colon G \to X$ describing the action have uniform asymptotic property C.*

Proof Let $R > 0$ be given and consider $B_R(x) \subset X$. Since this set is finite, it follows that $W_R(x)$ is a finite union of sets that are isometric to G_x. Hence, this union has the same asymptotic dimension as G_x and the result follows from Lemma 5.2. □

Lemma 5.4 *If the kernel of a group homomorphism $f\colon G \to H$ has finite asymptotic dimension, then coarse fibers of f have uniform asymptotic property C.*

We do not know whether the weaker assumption that the kernel of f has asymptotic property C is sufficient; see Question 7.2.

Proof Define an action of G on H by $g.h \mapsto f(g)h$. Notice that this action is isometric; ie $\operatorname{dist}(g.h, g.h') = \operatorname{dist}(h, h')$ for all $g \in G$. Also, the stabilizer of this action at e is $\ker f$. Since H is assumed to be a countable group with a proper metric, as a metric space it has bounded geometry, and the result follows from Lemma 5.3. □

Lemma 5.5 *If the kernel of a projection $p\colon G \times H \to G$ has asymptotic property C, then coarse fibers of p have uniform asymptotic property C.*

Proof We observe that the kernel of p is $\{e\} \times H$, which is isometric to H. Let $R_1 \le R_2 \le \cdots$ be a given sequence. Take families $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_k$ of subsets of H such that $\mathcal{U}_i$ is R_i–disjoint, $\operatorname{mesh}\mathcal{U}_i$ is finite for each i, and $\bigcup_i \mathcal{U}_i$ covers H. Given $M > 0$, we take $B = M + \max_i\{\operatorname{mesh}\mathcal{U}_i\}$. If $A \subset G \times H$ has the property that $\operatorname{diam}(p(A)) \le M$, then there is some $g \in G$ such that $p(A) \subset B^G(g, M)$, where the superscript indicates that the M–ball is taken in the group G. Consider the collection $\mathcal{V}_i$ of subsets of $G \times H$ given by $\{B^G(g, M) \times U : U \in \mathcal{U}_i\}$. Observe that this collection covers A, $\operatorname{mesh}\mathcal{V}_i \le \operatorname{mesh}\mathcal{U}_i + M \le B$, and each $\mathcal{V}_i$ is R_i–disjoint. □

We would like to remark that we did not require the bounded geometry condition on G or H in the previous proof. Indeed, the same technique can be used to show that if point fibers of a projection $p\colon X \times Y \to X$ have asymptotic property C, then coarse fibers of p have uniform asymptotic property C.

In [1] it is shown that a finitely generated group acting on a space with asymptotic property C will have asymptotic property C provided the R–stabilizers have finite

asymptotic dimension. We can use our fibering theorem to recover this result, which we state for arbitrary countable groups with proper left-invariant metrics.

Theorem 5.6 *Let G be a countable group with a proper left-invariant metric acting by isometries on a metric space X. Let $x_0 \in X$ and define $\phi\colon G \to X$ by $\phi(g) = g.x_0$. If, for every $R > 0$, $\operatorname{asdim} W_R(x_0) \leq n$ and, X has asymptotic property C, then G has asymptotic property C.*

Proof This will follow from the fibering lemma and Lemma 5.2 provided we can show that ϕ is uniformly expansive.

To this end, by [12] we know that the metric on G is a weighted word metric. Therefore, we assume that $S = \{s_i\}$ is a generating set for G with corresponding weights $w(s_i)$. By their definition and the fact that the metric is proper, the weights have the following property. For each fixed $N \in \mathbb{N}$, there are only finitely many (finite) sequences of weights whose sums are bounded from above by N; that is, the collection of all sequences $(w(s_{i_1}), w(s_{i_2}), \dots)$ for which $\sum w(s_{i_j}) \leq N$ is a finite collection. Thus, it makes sense to define a function $\rho\colon \mathbb{N} \to \mathbb{N}$ by $\rho(N) = \max\{\sum \operatorname{dist}(s_{i_j}.x_0, x_0)\}$, where the maximum is taken over all sequences of s_i for which $\sum w(s_{i_j}) \leq N$.

Then, by applying the triangle inequality and the fact that the metric is left-invariant, we find $\operatorname{dist}_X(g'.x_0, g.x_0) = \operatorname{dist}_X(g^{-1}g'.x_0, x_0) \leq \sum_{j=1}^{k} \operatorname{dist}_X(s_{i_j}.x_0, x_0)$, where $g^{-1}g' = s_{i_1}s_{i_2}\cdots s_{i_k}$. But now

$$\sum_{j=1}^{k} \operatorname{dist}_X(s_{i_j}.x_0, x_0) \leq \rho(\|g^{-1}g'\|) = \rho(\operatorname{dist}_G(g', g))$$

and so we conclude that $\operatorname{dist}_X(\phi(g'), \phi(g)) \leq \rho(\operatorname{dist}_G(g', g))$. □

Theorem 5.7 *Let $\phi\colon G \to H$ be a surjective homomorphism of finitely generated groups; suppose that H has property C. If $\operatorname{asdim} \ker \phi < \infty$ then G has asymptotic property C.*

Proof Let S be a generating set for H with corresponding weights W. For each $s \in S$, let $\tilde{s} \in \phi^{-1}(s)$ be some element of G. Extend the set $\{\tilde{s} : s \in S\}$ to a generating set T for G with the property that the weight of each $\tilde{s}$ is the same as the weight of the corresponding s. This yields a left-invariant proper metric on G which is coarsely equivalent to any other choice of left-invariant proper metric on G, so it suffices to show that G has asymptotic property C in this metric. In this metric, ϕ is coarsely expansive and so the result follows from the fibering lemma and Lemma 5.4. □

Corollary 5.8 *If G and H have asymptotic property C, then $G * H$ has asymptotic property C.*

Proof We consider the exact sequence

$$1 \to [\![G, H]\!] \to G * H \to G \times H \to 1.$$

The commutator group $[\![G, H]\!]$ is free [28] and so as a subspace of $G * H$ it is quasi-isometric to a tree (albeit one in which edges can have arbitrarily long length). Thus, its asymptotic dimension is 1. By Theorem 3.1, $G \times H$ has asymptotic property C. We apply Theorem 5.7 to complete the proof. □

6 Free products preserve asymptotic property C

In the present section we use the decomposition lemma to prove that the free product of metric spaces preserves asymptotic property C; however, the decomposition lemma has generalizations that make it interesting in its own right.

6.1 The decomposition lemma

Lemma 6.1 *Let X be a metric space. Assume that there exists k such that for each (infinite) sequence $R_1, R_2, \dots$ of real numbers there exists a finite sequence $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n$ of families of subsets of X such that*

(1) *$\bigcup_i \mathcal{U}_i$ covers X,*

(2) *$\mathcal{U}_i$ is R_i–disjoint,*

(3) *$\mathcal{U}_i$ has asymptotic dimension bounded by $k-1$ uniformly.*

Then X has asymptotic property C.

Proof Let $R_1 \le R_2 \le \cdots$ be a given sequence of real numbers. For $i \in \{1, 2, \dots, n\}$, take $\mathcal{U}_i$ to be an R_{ik}–disjoint collection of subsets of X that has asymptotic dimension less than k uniformly such that $\bigcup_i \mathcal{U}_i$ covers X.

Fix some i and suppose that $U \in \mathcal{U}_i$. Using the uniformity of the asymptotic dimension of the $\mathcal{U}_i$, take some B and take families $\mathcal{V}_j^U$ for $j = 1, \dots, k$ that comprise an R_{ik}–disjoint collection of subsets of U whose diameter is bounded by B such that the union covers U. Let $\mathcal{V}_{(i-1)k+j}$ denote the collection $\{V \in \mathcal{V}_j^U : U \in \mathcal{U}_i\}$. Since the family $\mathcal{U}_i$ is itself R_{ik}–disjoint, it is clear that the collection $\mathcal{V}_{(i-1)k+j}$ for $j = 1, \dots, k$

will be an R_{ik}–disjoint, B–bounded cover of $\bigcup_{U\in\mathcal{U}_i} U$. Since we assume the given sequence to be nondecreasing, we have found a uniformly bounded cover of X by nk families $\mathcal{V}_1,\dots,\mathcal{V}_{nk}$ of subsets of X with the property that $\mathcal{V}_t$ is R_t–bounded for each $t=1,\dots,nk$. Thus, X has property C, as required. □

Corollary 6.2 [29, Theorem III.8] *If X has asymptotic property C and Y has finite asymptotic dimension, then $X\times Y$ has asymptotic property C.*

6.2 Free product of metric spaces

Definition 6.3 Let (X,x_0) be a metric space with a fixed basepoint x_0. The *free product* $*X$ is the metric space whose elements are words in the alphabet $X\setminus\{x_0\}$ along with the trivial word ϵ.

We denote $X\setminus\{x_0\}$ by X^*. We identify elements of X^* with the set of single-letter (nontrivial) words in $*X$.

There is a natural notion of *concatenation* of words, which we usually denote by juxtaposition. When we want to emphasize the concatenation of two elements u and v in $*X$, we write $u\cdot v$. If $u\in *X$ and $A\subset *X$, we write $u\cdot A$ for the set of all elements of $*X$ that are obtained by concatenating an element of A on the left of the element $u\in *X$. The notation X^m is used to denote those elements of $*X$ that are obtained by concatenating exactly m elements of X^*. We define $X^0=\{\epsilon\}$ and use the notation $X^{\le m}$ for $\bigcup_{i=0}^m X^i$. Finally, if A and B are subsets of $*X$, then we define their *concatenation* $A\cdot B$ to be all elements of the form $a\cdot b$ with $a\in A$ and $b\in B$.

We define the *norm* of the trivial word to be 0 and the *norm* of a letter $x\in X^*$ to be $\|x\|=\operatorname{dist}(x_0,x)$. The *norm* of a composite word is the sum of norms of its letters: $\|x_1x_2\cdots x_k\|=\sum_{i=1}^k\|x_i\|$. The *distance* between two elements of $*X$ is determined by eliminating their common prefix: if $w=u\cdot x\cdot v$ and $w'=u\cdot x'\cdot v'$ (with $u,v,v'\in *X$ and $x,x'\in X^*$ with $x\neq x'$), then $\operatorname{dist}(w,w')=\operatorname{dist}(x,x')+\|v\|+\|v'\|$. The distance from w to the trivial word ϵ is equal to the norm $\|w\|$ of w.

Comment 6.4 For simplicity of proofs we defined the *unary* free product of a *single* metric space X. For applications, we use the fact that if X and Y are metric spaces, then $X*Y$ embeds isometrically into $*(X\vee Y)$, where $X\vee Y$ denotes the wedge of X and Y and $X*Y$ denotes the free product of X and Y, defined as in [4].

Definition 6.5 A subset $A\subset *X$ is said to be *flat* if there exists some $x\in *X$ such that $A\subset x\cdot(X^*)$.

Definition 6.6 Let $R > 0$. The *R–cone* of a set $A \subset {}^*X$ is defined to be

$$\operatorname{con}_R A = A \cdot ({}^*\overline{B}(x_0, R)) \subset {}^*X,$$

where $\overline{B}(x_0, R)$ denotes the closed ball with center x_0 and radius R.

Definition 6.7 We will say that the metric space X is *discrete in the metric sense* if $\inf\{\operatorname{dist}(x, y) : x, y \in X, x \neq y\} > 0$.

Lemma 6.8 *Let X be a metric space that is discrete in the metric sense. If $\{A_\alpha\}$ is a family of flat uniformly bounded subsets of ${}^*X$, then, for each M, the family $\{\operatorname{con}_M A_\alpha\}$ has asymptotic dimension bounded by 1 uniformly.*

Proof Let $E = \inf\{\operatorname{dist}(x, y) : x, y \in X, x \neq y\}$. By the assumption, $E > 0$. Let $D = \sup_\alpha\{\operatorname{diam}(A_\alpha)\}$. By the assumption, $D < \infty$.

For each α we define a simplicial tree T_α. The set of vertices of T_α consists of elements of $\operatorname{con}_M A_\alpha$ and of an additional element r_α, which we call the root vertex. The root vertex r_α is connected by edges to all vertices from the base A_α of the cone $\operatorname{con}_M A_\alpha$. Two vertices p and t of the cone $\operatorname{con}_M A_\alpha$ are connected by an edge if and only if there exists $x \in \overline{B}(x_0, M)$ with $x \neq x_0$ such that $p = t \cdot x$ or $t = p \cdot x$. The metric $\operatorname{dist}_{T_\alpha}$ on T_α is the geodesic metric with edge lengths equal to 1.

Consider the map $f_\alpha\colon \operatorname{con}_M A_\alpha \to T_\alpha$ that maps elements of $\operatorname{con}_M A$ to their corresponding vertices of T_α. We will show that f_α satisfies the inequalities

$$\frac{\operatorname{dist}(x, y)}{M} - \frac{D}{M} \leq \operatorname{dist}_{T_\alpha}(f_\alpha(x), f_\alpha(y)) \leq \frac{\operatorname{dist}(x, y)}{E} + 3$$

for each $x, y \in \operatorname{con}_M A_\alpha$.

As a first case we consider $x, y \in \operatorname{con}_M A_\alpha$ such that $x = a x_1 x_2 \cdots x_k$ and $y = b y_1 y_2 \cdots y_l$, where $a, b \in A_\alpha$ with $a \neq b$ and $x_i, y_i \in \overline{B}(x_0, R)^*$ with $k, l \geq 0$. We have

$$\operatorname{dist}(x, y) = \left(\sum_{i=1}^{k} \|x_i\|\right) + \operatorname{dist}(a, b) + \left(\sum_{i=1}^{l} \|y_i\|\right)$$

and

$$\operatorname{dist}_{T_\alpha}(f_\alpha(x), f_\alpha(y)) = 2 + k + l.$$

We have

$$E \leq \operatorname{dist}(a, b) \leq D, \quad E \leq \|x_i\| \leq M, \quad E \leq \|y_i\| \leq M.$$

Therefore,

$$(k+l+1)E \leq \operatorname{dist}(x,y) \leq (k+l)M + D.$$

As a second case consider $x, y \in \operatorname{con}_M A_\alpha$ such that $x = c\cdot x_1 x_2 \cdots x_k$, $y = c\cdot y_1 y_2 \cdot y_l$, $c \in \operatorname{con}_M A_\alpha$, $x_i, y_i \in \overline{B}(x_0, R)^*$ and $x_1 \neq y_1$. We have

$$\operatorname{dist}(x,y) = \left(\sum_{i=2}^{k} \|x_i\|\right) + \operatorname{dist}(x_1, y_1) + \left(\sum_{i=2}^{l} \|y_i\|\right)$$

and

$$\operatorname{dist}_{T_\alpha}(f_\alpha(x), f_\alpha(y)) = k + l.$$

We have

$$E \leq \operatorname{dist}(x_1, y_1) \leq D, \quad E \leq \|x_i\| \leq M, \quad E \leq \|y_i\| \leq M.$$

Therefore,

$$(k+l-1)E \leq \operatorname{dist}(x,y) \leq (k+l)M.$$

Combining the inequalities from both cases we have

$$\operatorname{dist}_{T_\alpha}(f_\alpha(x), f_\alpha(y)) \leq k + l + 2 \leq \frac{\operatorname{dist}(x,y)}{E} + 3$$

and

$$\frac{\operatorname{dist}(x,y)}{M} - \frac{D}{M} \leq k + l \leq \operatorname{dist}_{T_\alpha}(f_\alpha(x), f_\alpha(y)).$$

Hence, the inequalities that we sought:

$$\frac{\operatorname{dist}(x,y)}{M} - \frac{D}{M} \leq \operatorname{dist}_{T_\alpha}(f_\alpha(x), f_\alpha(y)) \leq \frac{\operatorname{dist}(x,y)}{E} + 3.$$

Therefore, each f_α is a quasi-isometry with constants that do not depend on α.

In the example following Theorem 19 of [6], it is shown that for a given $r > 0$ one can take a cover of any simplicial tree by $3r$–bounded, r–disjoint families of subsets. It follows that any family of simplicial trees has asymptotic dimension bounded by 1 uniformly. Since the family $\{\operatorname{con}_M A_\alpha\}$ is uniformly quasi-isometric to a family of simplicial trees, it has asymptotic dimension bounded by 1 uniformly. □

Definition 6.9 We say that a set A is *R–connected* if for every $x, y \in A$ there exists a finite sequence $x_0 = x, x_1, x_2, \ldots, x_k = y$ such that $\operatorname{dist}(x_i, x_{i+1}) \leq R$ for each $i = 1, 2, \ldots, k-1$. An *R–connected component* of A (or simply an *R–component*) is a maximal R–connected subset of A. We recall that in Section 2 we described the sets A and B as being *R–disjoint* if, for each $x \in A$ and $y \in B$, we have $\operatorname{dist}(x,y) > R$.

Definition 6.10 The *order* of an element x of $*X$ is the nonnegative integer m such that $x \in X^m$. We write $\operatorname{ord}(x) = m$ in this case.

Lemma 6.11 *Let X be a discrete metric space and let $R > 0$. Let $A \subset *X$ be such that $A \subset *X \cdot (A \setminus \overline{B}(x_0, R))$. Assume that R–connected components of A are uniformly bounded. Then, for each M, R–connected components of $\operatorname{con}_M A$ have asymptotic dimension bounded by 1 uniformly.*

Proof As a first step we will show that

$$C \text{ is an } R\text{–connected component of } \operatorname{con}_M A$$
$$\iff \quad C \cap X^{\leq n} \text{ is an } R\text{–connected component of } \operatorname{con}_M A \cap X^{\leq n} \text{ for each } n.$$

The implication $\Leftarrow$ is trivial.

We will say that two words w_1, w_2 in $*X$ are *adjacent* if

(1) $w_1 = w_2 \cdot x$ for some $x \in X$, or

(2) $w_2 = w_1 \cdot x$ for some $x \in X$, or

(3) $w_1 = w \cdot x_1$ and $w_2 = w \cdot x_2$ for some $w \in *X$ and $x_1, x_2 \in X$.

In other words, w_1 and w_2 are adjacent if they differ by their last letter or if one is equal to the other with a single letter appended at the end.

Let $x = p \cdot x_1 x_2 \cdot x_k$ and $y = p \cdot y_1 y_2 \cdot y_l$ with $x, y \in \operatorname{con}_M A$ and $\operatorname{dist}(x, y) \leq R$, where p is the common prefix of x and y. We have

$$\operatorname{dist}(x, y) = \left(\sum_{i=2}^{k} \|x_i\| \right) + \operatorname{dist}(x_1, y_1) + \left(\sum_{i=2}^{l} \|y_i\| \right) \leq R.$$

Therefore $x_i, y_i \in \overline{B}(x_0, R)$ for $i \geq 2$, so $x_i, y_i \notin A$ for $i \geq 2$. Hence, $p \cdot x_1 x_2 \cdots x_i \in \operatorname{con}_M A$ and $p \cdot y_1 y_2 \cdots y_i \in \operatorname{con}_M A$ for each i. Therefore, $x = p \cdot x_1 x_2 \cdots x_k$, $p \cdot x_1 x_2 \cdots x_{k-1}, \ldots, p \cdot x_1, p \cdot y_1, p \cdot y_1 y_2, \ldots, p \cdot y_1 y_2 \cdots y_l = y$ is a sequence of R–close adjacent elements of $\operatorname{con}_M A$.

Therefore, if C is an R–connected component of $\operatorname{con}_M A$, then for each $x, y \in C$ there exists a sequence $x = x_0, x_1, \ldots, x_k = y$ of adjacent R–close points of C.

Observe that if $x_1, x_2, \ldots, x_k$ is a sequence of adjacent points with $\operatorname{ord}(x_1) = \operatorname{ord}(x_k)$ and $\operatorname{ord}(x_i) > \operatorname{ord}(x_1)$ for all i with $2 \leq i \leq k-1$, then $x_1 = x_k$. This implies that if $x, y \in \operatorname{con}_M A$ are in the same R–connected component and $x, y \in X^{\leq n}$, then x and y are in the same R–connected component of $C \cap X^{\leq n}$. The implication $\Rightarrow$ is proven.

Let D be a bound on the diameters of the R–connected components of A. Let C be an R–connected component of $\mathrm{con}_M A$. Let $C_k = C \cap X^{\leq k}$. We have proven that each C_k is R–connected. Let k_0 be the smallest integer such that $C_{k_0} \neq \varnothing$. The diameter of C_{k_0} is bounded by $2M + 2R + 2D$. We will show by induction on k that

$$C_k \subset \mathrm{con}_{M+R+D}\, C_{k_0}.$$

Consider $x \in (C_{k+1} \setminus C_k)$. Since $C \subset \mathrm{con}_M A$, we have $x \in A$ or $x \in \mathrm{con}_M C_k$.

Let $x \in A$. Then x has to be in an R–connected component of A that is R–close to $\mathrm{con}_M\, C_k$. Therefore $x \in \mathrm{con}_{M+R+D}\, C_k$.

Let $x \in \mathrm{con}_M\, C_k$. Then $\mathrm{con}_M\, C_k \subset \mathrm{con}_{M+R+D}\, C_k$. Hence, $x \in \mathrm{con}_{M+R+D}\, C_k \subset \mathrm{con}_{M+R+D}\, \mathrm{con}_{M+R+D}\, C_{k_0} = \mathrm{con}_{M+R+D}\, C_{k_0}$.

We proved by induction that if C is an R–connected component of $\mathrm{con}_M A$, then $C \subset \mathrm{con}_{M+R+D}\, C_{k_0}$. Observe that C_{k_0} is flat and has uniformly bounded diameter (by $2M + 2R + 2D$). Therefore, by Lemma 6.8, the collection of R–connected components of $\mathrm{con}_M A$ has asymptotic dimension bounded by 1 uniformly. □

Theorem 6.12 *Let X be a discrete metric space with a fixed basepoint x_0. If X has asymptotic property C, then $*X$ has asymptotic property C.*

Proof Let $R_1 \leq R_2 \leq \cdots$ be a nondecreasing sequence of real numbers. Let $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_n$ be a finite sequence of R_i–disjoint uniformly bounded families of subsets of X such that $\bigcup \mathcal{U}_i$ covers X. For $i = 1, 2, \ldots, n$, let

$$\mathcal{V}_i = \{x \cdot (U \setminus \overline{B}(x_0, R_{n+1})) : U \in \mathcal{U}_i,\ x \in *X\}$$

and

$$\mathcal{V}_{n+1} = \{\{x_0\}\}.$$

We will prove that

(1) $\bigcup_{i=1}^{n+1} \big(\mathrm{con}_{R_{n+1}} \bigcup \mathcal{V}_i\big) = *X$.

(2) $\mathcal{V}_i$ is R_i–disjoint, uniformly bounded and its elements are flat for each i.

To prove (1), consider an element $x = x_1 x_2 \cdots x_k \in *X$. If $\|x_i\| \leq R_{n+1}$ for each i, then $x \in \mathrm{con}_{R_{n+1}}\{x_0\}$. Otherwise, let m be the largest integer such that $\|x_m\| > R_{n+1}$. Since $\bigcup \mathcal{U}_i$ covers X, there exists an l and a $U \in \mathcal{U}_l$ such that $x_m \in U$. We have $x \in x_1 x_2 \cdots x_{m-1} \cdot (U \setminus \overline{B}(x_0, R_{n+1})) \cdot *\overline{B}(x_0, R_{n+1})$; hence, $x \in \mathrm{con}_{R_{n+1}} \bigcup \mathcal{V}_l$.

To prove (2), assume that i is given, with $1 \leq i \leq n$. Observe that elements of $\mathcal{V}_i$ are flat and uniformly bounded directly from the definition. So, it remains only to show that $\mathcal{V}_i$ is R_i–disjoint. Suppose that V_1 and V_2 are distinct elements of $\mathcal{V}_i$, say $V_1 = x_1 \cdot U_1$ and $V_2 = x_2 \cdot U_2$. We consider two cases: (a) If $x_1 = x_2$, then the result follows from the fact that the family $\mathcal{U}_i$ is R_i–disjoint. (b) If $V_1 \cup V_2$ is not flat, then we take arbitrary elements $v_1 = x_1 u_1 \in V_1 = x_1 \cdot U_1$ and $v_2 = x_2 u_2 \in V_2 = x_2 \cdot U_2$. Then $\operatorname{dist}(v_1, v_2) \geq \min\{\|u_1\|, \|u_2\|\} > R_{n+1} \geq R_i$, by the definition of $\mathcal{V}_i$ and the definition of the metric in $*X$.

By Lemma 6.11, the R_i–connected components of $\operatorname{con}_{R_{n+1}+1} \bigcup \mathcal{V}_i$ have asymptotic dimension bounded by 1 uniformly and are R_i–disjoint. By the decomposition lemma, $*X$ has asymptotic property C. □

Corollary 6.13 *Let X and Y be metric spaces with fixed basepoints. If X and Y have asymptotic property C, then $X * Y$ has asymptotic property C.*

Proof By the union theorem (Proposition 2.3), $X \vee Y$ has asymptotic property C. By Theorem 6.12, $(X \vee Y) * (X \vee Y)$ has asymptotic property C. Since $X * Y$ embeds isometrically into $(X \vee Y) * (X \vee Y)$ and asymptotic property C is inherited by subspaces, we are done. □

7 Open questions

Although we were able to resolve several questions related to asymptotic property C, we would like to mention the following problems, which remain open.

We were able to show that the free product of two metric spaces (or groups) with asymptotic property C has asymptotic property C, but this invites the obvious question for amalgamated products [7, Questions 4.3]: Is asymptotic property C preserved by amalgamated free products? Generalizing this a bit more, one could consider the following question:

Let Γ be a finite graph. To each vertex v of Γ we assign a group G_v. We let $\mathfrak{G}\Gamma$ denote the graph product of these groups; ie the group generated by the G_v along with the additional relation that two elements of the groups G_v and G_w commute precisely when an edge of Γ connects v and w. If all the G_v have asymptotic property C, does $\mathfrak{G}\Gamma$ have asymptotic property C?

These questions and the question of whether a semidirect product preserves asymptotic property C could be addressed via a true fibering theorem. We state this in two forms, both as a statement for metric spaces and a statement in terms of exact sequences of groups.

Question 7.1 Let $f\colon X \to Y$ be a uniformly expansive map between metric spaces. Assume that Y has asymptotic property C and that $f^{-1}(A)$ has asymptotic property C for every bounded subset $A \subset Y$. Does X have to have asymptotic property C?

Question 7.2 Let $f\colon G \to H$ be a group homomorphism such that H and $\ker f$ have asymptotic property C. Does G have to have asymptotic property C?

One could also ask several questions related to the construction of Yamauchi [30]. In particular, one could ask [30, Question 3.2]: Does every countable direct union of groups with finite asymptotic dimension have asymptotic property C? More generally, is asymptotic property C preserved by countable direct unions?

Dydak and Virk showed that the notions of countable asymptotic dimension and straight finite decomposition complexity are equivalent [15]. It is easy to see that asymptotic property C implies either of these two notions, but the exact nature of the relationship between these notions is unknown for groups. We believe that the asymptotic property C can be generalized to a property we propose to call ω–APC. This generalization enables us to see that our fibering lemma and decomposition lemma are both special cases of a more general framework of which straight finite decomposition complexity and asymptotic property C are the first nontrivial special cases.

References

[1] **S Beckhardt**, *Groups acting on metric spaces with asymptotic property C*, preprint (2015) arXiv

[2] **S Beckhardt**, **B Goldfarb**, *Extension properties of asymptotic property C and finite decomposition complexity*, preprint (2016) arXiv

[3] **G C Bell**, *Property A for groups acting on metric spaces*, Topology Appl. 130 (2003) 239–251 MR

[4] **G Bell**, **A Dranishnikov**, *On asymptotic dimension of groups*, Algebr. Geom. Topol. 1 (2001) 57–71 MR

[5] **G C Bell**, **A N Dranishnikov**, *A Hurewicz-type theorem for asymptotic dimension and applications to geometric group theory*, Trans. Amer. Math. Soc. 358 (2006) 4749–4764 MR

[6] **G Bell**, **A Dranishnikov**, *Asymptotic dimension*, Topology Appl. 155 (2008) 1265–1296 MR

[7] **G Bell**, **D Moran**, *On constructions preserving the asymptotic topology of metric spaces*, North Carolina J. Math. Stat. 1 (2015) 46–57

[8] **G Bell**, **D Moran**, **A Nagórko**, *Coarse property C and decomposition complexity*, Topology Appl. 227 (2017) 30–50 MR

[9] **N Brodskiy**, **J Dydak**, **M Levin**, **A Mitra**, *A Hurewicz theorem for the Assouad–Nagata dimension*, J. Lond. Math. Soc. 77 (2008) 741–756 MR

[10] **T Davila**, *On asymptotic property C*, preprint (2016) arXiv

[11] **A N Dranishnikov**, *Asymptotic topology*, Uspekhi Mat. Nauk 55 (2000) 71–116 MR In Russian; translated in Russian Math. Surveys 55 (2000) 1085–1129

[12] **A Dranishnikov**, **J Smith**, *Asymptotic dimension of discrete groups*, Fund. Math. 189 (2006) 27–34 MR

[13] **A Dranishnikov**, **M Zarichnyi**, *Asymptotic dimension, decomposition complexity, and Haver's property C*, Topology Appl. 169 (2014) 99–107 MR

[14] **J Dydak**, *Coarse amenability and discreteness*, J. Aust. Math. Soc. 100 (2016) 65–77 MR

[15] **J Dydak**, **Ž Virk**, *Preserving coarse properties*, Rev. Mat. Complut. 29 (2016) 191–206 MR

[16] **R Engelking**, *Theory of dimensions finite and infinite*, Sigma Series in Pure Mathematics 10, Heldermann, Lemgo (1995) MR

[17] **B Grave**, *Asymptotic dimension of coarse spaces*, New York J. Math. 12 (2006) 249–256 MR

[18] **M Gromov**, *Asymptotic invariants of infinite groups*, from "Geometric group theory, II" (G A Niblo, M A Roller, editors), London Math. Soc. Lecture Note Ser. 182, Cambridge Univ. Press (1993) 1–295 MR

[19] **E Guentner**, *Permanence in coarse geometry*, from "Recent progress in general topology, III" (K P Hart, J van Mill, P Simon, editors), Atlantis Press, Paris (2014) 507–533 MR

[20] **Y Hattori**, **K Yamada**, *Closed pre-images of C–spaces*, Math. Japon. 34 (1989) 555–561 MR

[21] **W E Haver**, *A covering property for metric spaces*, from "Topology conference" (R F Dickman, Jr, P Fletcher, editors), Lecture Notes in Math. 375, Springer (1974) 108–113 MR

[22] **P W Nowak**, *Coarsely embeddable metric spaces without Property A*, J. Funct. Anal. 252 (2007) 126–136 MR

[23] **E Pearl** (editor), *Open problems in topology, II*, Elsevier, Amsterdam (2007) MR

[24] **E Pol**, *A weakly infinite-dimensional space whose product with the irrationals is strongly infinite-dimensional*, Proc. Amer. Math. Soc. 98 (1986) 349–352 MR

[25] **E Pol**, **R Pol**, *A metric space with the Haver property whose square fails this property*, Proc. Amer. Math. Soc. 137 (2009) 745–750 MR

[26] **J Roe**, *Lectures on coarse geometry*, University Lecture Series 31, Amer. Math. Soc., Providence, RI (2003) MR

[27] **D M Rohm**, *Products of infinite-dimensional spaces*, Proc. Amer. Math. Soc. 108 (1990) 1019–1023 MR

[28] **J-P Serre**, *Trees*, Springer (2003) MR

[29] **L Sher**, *Asymptotic dimension and asymptotic property C*, Master's thesis, The University of North Carolina at Greensboro (2011) Available at `http://tinyurl.com/SherUNC`

[30] **T Yamauchi**, *Asymptotic property C of the countable direct sum of the integers*, Topology Appl. 184 (2015) 50–53 MR

Department of Mathematics and Statistics, The University of North Carolina
Greensboro, NC, United States

Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw
Warszawa, Poland

`gcbell@uncg.edu, amn@mimuw.edu.pl`

Received: 22 July 2016 Revised: 29 June 2017

msp

Algebraic & Geometric Topology 18 (2018) 247–312

Higher cohomology operations and R–completion

DAVID BLANC
DEBASIS SEN

Let R be either $\mathbb{F}_p$ or a field of characteristic 0. For each R–good topological space $\boldsymbol{Y}$, we define a collection of higher cohomology operations which, together with the cohomology algebra $H^*(\boldsymbol{Y}; R)$, suffice to determine $\boldsymbol{Y}$ up to R–completion. We also provide a similar collection of higher cohomology operations which determine when two maps $f_0, f_1\colon \boldsymbol{Z} \to \boldsymbol{Y}$ between R–good spaces (inducing the same algebraic homomorphism $H^*(\boldsymbol{Y}; R) \to H^*(\boldsymbol{Z}; R)$) are R–equivalent.

55P60; 55N99, 55P15, 55P20

Introduction

We describe complete sets of invariants for the R–homotopy types of R–good topological spaces and maps between them, consisting of systems of higher R–cohomology operations (where R is either $\mathbb{F}_p$ or a field of characteristic 0).

Higher homotopy or cohomology operations (see Toda [37], Spanier [35; 36], Adams [1] and Maunder [30]) should be thought of as inductively defined systems of obstructions to rectifying homotopy-commutative diagrams. In particular, an *n^{th} order cohomology operation* is attached to a diagram $F\colon I \to \mathrm{ho}\,\mathsf{Top}$ indexed by a *lattice* in the sense of Blanc and Markl [15, Section 2] — that is, a directed category I of length $n+1$ such that, for all but the initial object i_0 in I, $F(i)$ is a product of R–module Eilenberg–Mac Lane spaces (an R–GEM).

0.1 Example The simplest example is a Toda bracket [3], for a diagram of the form

(0.2)
$$\begin{array}{ccccccc} & & * & & & & \\ & \overset{f}{\nearrow} & & \overset{g}{\searrow} & & \overset{h}{\nearrow} & \\ \boldsymbol{Y} & \longrightarrow & \boldsymbol{W}_0 & \longrightarrow & \boldsymbol{W}_1 & \longrightarrow & \boldsymbol{W}_2, \\ & & & \underset{*}{\searrow} & & \underset{*}{\nearrow} & \end{array}$$

with $\boldsymbol{W}_0$, $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ R–GEMs, and each adjacent composition nullhomotopic (see Adams [1] and Harper [26]).

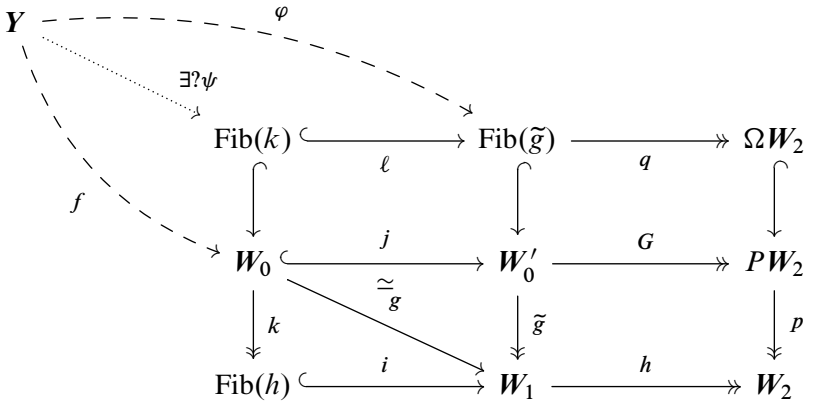
Published: 19 January 2018 DOI: 10.2140/agt.2018.18.247

This defines a secondary cohomology operation in the sense of Adams [1], since f represents a set of cohomology classes in $H^*(\boldsymbol{Y}; R)$, on which the set of primary R–cohomology operations represented by g vanish. The fact that $h \circ g \sim *$ indicates a relation among primary operations.

Our general strategy for rectification of any $F\colon I \to \mathrm{ho}\,\mathsf{Top}$ is to inductively rectify, and then make fibrant, longer and longer final segments of the given diagram. In our example, we first make $\boldsymbol{W}_2$ fibrant and change h into a fibration (so the subdiagram $\boldsymbol{W}_1 \xrightarrow{h} \boldsymbol{W}_2$ is fibrant). We then change g up to homotopy so that $h \circ g = *$, using Blanc, Johnson and Turner [11, Lemma 5.11]. Factoring g as $\boldsymbol{W}_0 \xrightarrow{k'} \mathrm{Fib}(h) \xhookrightarrow{i} \boldsymbol{W}_1$, and then changing k' into a fibration k makes $\boldsymbol{W}_0 \xrightarrow{g} \boldsymbol{W}_1 \xrightarrow{h} \boldsymbol{W}_2$ fibrant. To simplify notation we denote $i \circ k$ simply by $g\colon \boldsymbol{W}_0 \to \boldsymbol{W}_1$.

We think of the following solid diagram of vertical and horizontal fibration sequences as the *template* for our Toda bracket (depending only on $\boldsymbol{W}_0 \xrightarrow{g} \boldsymbol{W}_1 \xrightarrow{h} \boldsymbol{W}_2$):

(0.3)



The nullhomotopy G exists since $h \circ \widetilde{g} \circ j = g \circ h = *$ and j is a weak equivalence.

Now, any map $f\colon \boldsymbol{Y} \to \boldsymbol{W}_0$ with $g \circ f \sim *$ factors up to homotopy through a map $\varphi\colon \boldsymbol{Y} \to \mathrm{Fib}(\widetilde{g})$. To rectify (0.2), φ and f should induce a map $\psi\colon \boldsymbol{Y} \to \mathrm{Fib}(k) = \mathrm{Fib}(q)$. The obstruction to doing so — namely the homotopy class of the composite $q \circ \varphi\colon \boldsymbol{Y} \to \Omega \boldsymbol{W}_2$ — is called the *value* of the Toda bracket $\langle f, g, h\rangle$ (for the given choices of k and φ).

0.4 The basic construction Our object in this paper is to associate to each R–good space $\boldsymbol{Y}$ a sequence $\langle\langle \boldsymbol{Y} \rangle\rangle = (\langle\langle \boldsymbol{Y} \rangle\rangle_n)_{n=2}^{\infty}$ of higher cohomology operations which serve as a complete set of invariants for the R–homotopy type of $\boldsymbol{Y}$, constructed roughly as follows:

(a) We start with $H^*(\boldsymbol{Y};R)$ as a Θ_R–*algebra* — ie a graded R–algebra with an action of the primary R–cohomology operations (such as Steenrod squares). We then choose a *CW resolution* of this Θ_R–algebra, given by an inductively defined simplicial Θ_R–algebra $V_\bullet$, with the nth stage obtained from the $(n-1)$–truncation by attaching a free Θ_R–algebra $\overline{V}_n$ along a map $\overline{\partial}^n_0\colon \overline{V}_n \to V_{n-1}$ (see Section 1.11 below). Our goal is to realize $V_\bullet \to H^*(\boldsymbol{Y};R)$ by a coaugmented cosimplicial space $\boldsymbol{Y} \to \boldsymbol{W}^\bullet = \lim_n \boldsymbol{W}^\bullet_{[n]}$, for which $\boldsymbol{Y} \to \operatorname{Tot}\boldsymbol{W}^\bullet$ is the R–completion map of Bousfield and Kan [20].

(b) To carry out this program, we use a double induction: if $\boldsymbol{W}^\bullet_{[n-1]}$ realizes $V_\bullet$ through simplicial dimension $n-1$, we can choose a map $\overline{d}^0_{n-1}\colon \boldsymbol{W}^{n-1}_{[n-1]} \to \overline{\boldsymbol{W}}^n$ realizing $\overline{\partial}^n_0\colon \overline{V}_n \to V_{n-1}$ (with $\overline{\boldsymbol{W}}^n$ an R–GEM). Then $\overline{d}^0_{n-1}\circ d^0$ is nullhomotopic by a nullhomotopy $F^{n-2}\colon \boldsymbol{W}^{n-2}_{[n-1]} \to P\overline{\boldsymbol{W}}^n$, with $F^{n-2}\circ d^0$ factoring through $a^{n-3}\colon \boldsymbol{W}^{n-3}_{[n-1]} \to \Omega\overline{\boldsymbol{W}}^n$, as in Example 0.1 above. One can in fact choose F^{n-2} so that a^{n-3} is nullhomotopic.

(c) Step (b) can be repeated, and the kth obstruction a^{n-k} again turns out to be nullhomotopic. We end up with an nth order cohomology operation $\langle\langle \boldsymbol{Y}\rangle\rangle_n$ with value $a^{-1}\colon \boldsymbol{Y} \to \Omega^{n-1}\overline{\boldsymbol{W}}^n$. The fact that this too vanishes allows us to extend $\boldsymbol{W}^\bullet_{[n-1]}$ to a cosimplicial space $\boldsymbol{W}^\bullet_{[n]}$ (which will be n–coskeletal, up to homotopy). We call the system $\mathcal{W} = (\boldsymbol{W}^\bullet_{[n]})_{n\in\mathbb{N}}$ a *sequential realization* of $V_\bullet$ for the space $\boldsymbol{Y}$, and think of it as a template for a sequence $\langle\langle \boldsymbol{Y}\rangle\rangle = (\langle\langle \boldsymbol{Y}\rangle\rangle_n)_{n=2}^\infty$ of higher cohomology operations.

(d) Finally, given a space $\boldsymbol{Z}$ with $H^*(\boldsymbol{Y};R) \cong H^*(\boldsymbol{Z};R)$, the augmentation of Θ_R–algebras $\varepsilon\colon V_0 \to H^*(\boldsymbol{Z};R)$ can be realized by a map $\boldsymbol{\varepsilon}_{[0]}\colon \boldsymbol{Z} \to \boldsymbol{W}^0_{[0]}$. If we can extend this to a coaugmentation $\boldsymbol{\varepsilon}\colon \boldsymbol{Z} \to \boldsymbol{W}^\bullet := \lim \boldsymbol{W}^\bullet_{[n]}$, we would obtain an R–equivalence between $\boldsymbol{Z}$ and $\boldsymbol{Y}$. The obstruction to extending the $(n-1)^{\text{st}}$ approximation $\boldsymbol{\varepsilon}_{[n-1]}\colon \boldsymbol{Z} \to \boldsymbol{W}^\bullet_{[n]}$ for $\boldsymbol{\varepsilon}$ to $\boldsymbol{\varepsilon}_{[n]}$ is a map $a^{-1}\colon \boldsymbol{Z} \to \Omega^{n-1}\overline{\boldsymbol{W}}^n$ as above — the value associated to $\boldsymbol{Z}$ for the nth order cohomology operation $\langle\langle \boldsymbol{Y}\rangle\rangle_n$.

0.5 Main results In order to apply the machinery described above, we need the following important technical result:

Theorem A *Any CW resolution $V_\bullet$ of the Θ_R–algebra $H^*(\boldsymbol{Y};R)$ can be realized by a coaugmented cosimplicial space $\boldsymbol{Y} \to \boldsymbol{W}^\bullet$ with each $\boldsymbol{W}^n$ an R–GEM, obtained as the limit of a sequential realization as above.*

See Theorem 2.33 below. This allows us to produce various templates for the system of higher cohomology operations $\langle\langle \boldsymbol{Y}\rangle\rangle$, based on the algebraic resolution of our choice.

The sequence of higher cohomology operations presented here is dual to the higher homotopy operations of Blanc [6] and Blanc, Johnson and Turner [13], which correspond to the André–Quillen cohomology obstructions for distinguishing between different realizations $\pi_*\boldsymbol{Y}$ of a Π–algebra (see Blanc, Dywer and Goerss [10]). Such André–Quillen classes appear also in the dual context of distinguishing between different realizations of a given abstract Θ_R–algebra Γ (see Blanc [8] and Biedermann, Raptis and Stelzer [5]). However, the higher cohomology operations of Blanc [8] correspond to André–Quillen *cocycles* (for a specific algebraic resolution $V_\bullet$ of $H^*(\boldsymbol{Y};R)$). We therefore would like to collect together the various higher order cohomology operations corresponding to a given André–Quillen cohomology class. We do this by means of suitable (split) weak equivalences, called *comparison maps*, between sequential realizations for various algebraic resolutions, and show:

Theorem B *Any two sequential realizations of two CW resolutions $V_\bullet$ for the same space $\boldsymbol{Y}$ are connected by a zigzag of comparison maps.*

See Theorem 3.20 below.

Our two main results may then be summarized as follows:

Theorem C *For R either $\mathbb{F}_p$ or a field of characteristic 0, let $\boldsymbol{Y}$ and $\boldsymbol{Z}$ be R–good spaces and $\vartheta\colon H^*(\boldsymbol{Y};R)\to H^*(\boldsymbol{Z};R)$ an isomorphism of Θ_R–algebras. Then ϑ is realizable by a zigzag of R–equivalences between $\boldsymbol{Y}$ and $\boldsymbol{Z}$ if and only if the system of higher operations associated to this initial data vanishes.*

See Theorem 4.18 below.

By extending the ideas sketched in Section 0.4, one can use any sequential realization for $\boldsymbol{Y}$ to define a system of higher cohomology operations associated to any two maps $f_0, f_1\colon \boldsymbol{Z}\to\boldsymbol{Y}$ which induce the same map in cohomology; although the construction is more complicated, these operations still take values in the groups $[\boldsymbol{Z}, \Omega^{n-1}\overline{\boldsymbol{W}}^n]$, and we have:

Theorem D *For R either $\mathbb{F}_p$ or a field of characteristic 0, let $f_0, f_1\colon \boldsymbol{Z}\to\boldsymbol{Y}$ be two maps between R–good spaces which induce the same morphism of Θ_R–algebras $H^*(\boldsymbol{Y};R)\to H^*(\boldsymbol{Z};R)$. Then f_0 is R–equivalent to f_1 if and only if the associated system of higher operations vanishes.*

See Theorem 5.29 below.

0.6 Organization Section 1 provides some background material on (co)simplicial constructions and on sketches and their algebras. In Section 2 we define and study sequential realizations of algebraic resolutions. In Section 3 we show how any two such sequential realizations may be connected by a zigzag of comparison maps. In Section 4 we construct the higher cohomology operations used to distinguish between spaces, including a detailed rational example in Section 4.24, while in Section 5 we define the analogous invariants for maps.

The appendix reviews the notions of enriched sketches and their mapping algebras, which are used to generalize and prove Theorem A.

0.7 Acknowledgements We would like to thank the referee for his or her careful and pertinent comments.

The research of Blanc was supported by Israel Science Foundation grants 74/11 and 770/16, and that of Sen by INSPIRE grant No. IFA MA-12.

1 Background

In this section we present some background material on (co)simplicial theory and algebraic theories that will be used throughout the paper.

1.1 Definition Let $\mathbf{\Delta}$ denote the category of finite ordered sets and order-preserving maps (see May [31, Section I.2]), and $\mathbf{\Delta}_+$ the subcategory with the same objects, but only monic maps. A *cosimplicial object* $G^\bullet$ in a category $\mathcal{C}$ is a functor $\mathbf{\Delta} \to \mathcal{C}$, and a *restricted* cosimplicial object is a functor $\mathbf{\Delta}_+ \to \mathcal{C}$. More concretely, we write G^n for the value of $G^\bullet$ at the ordered set $[\boldsymbol{n}] = \{0 < 1 < \cdots < n\}$. The maps in the diagram $G^\bullet$ are generated by the *coface* maps $d^i = d^i_n \colon G^n \to G^{n+1}$ for $0 \le i \le n+1$, as well as *codegeneracy* maps $s^j = s^j_n \colon G^n \to G^{n-1}$ for $0 \le j < n$ in the nonrestricted case, satisfying the usual cosimplicial identities. Dually, a *simplicial object* $G_\bullet$ in $\mathcal{C}$ is a functor $\mathbf{\Delta}^{\mathrm{op}} \to \mathcal{C}$. The category of cosimplicial objects over $\mathcal{C}$ will be denoted by $\mathcal{C}^{\mathbf{\Delta}}$, and the category of simplicial objects over $\mathcal{C}$ will be denoted by $\mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}}$. However, the category of simplicial sets is denoted simply by $\mathcal{S} = s\mathsf{Set}$, and that of pointed simplicial sets by $\mathcal{S}_* = s\mathsf{Set}_*$. By a *space* we always mean a pointed simplicial set.

There are natural embeddings $c(-)^\bullet \colon \mathcal{C} \to \mathcal{C}^{\mathbf{\Delta}}$ and $c(-)_\bullet \colon \mathcal{C} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}}$, defined by letting $c(A)^\bullet$ denote the constant cosimplicial object which is A in every cosimplicial dimension, and similarly for $c(A)_\bullet$.

1.2 Latching and matching objects For a cosimplicial object $G^\bullet \in \mathcal{C}^{\mathbf{\Delta}}$ in a complete category $\mathcal{C}$, the nth *matching object* for $G^\bullet$ is defined to be

$$M^n G^\bullet := \lim_{\phi:[n]\to[k]} G^k, \tag{1.3}$$

where ϕ ranges over the nonidentity surjective maps $[n]\to[k]$ in $\mathbf{\Delta}$. There is a natural map $\zeta^n\colon G^n \to M^n G^\bullet$ induced by the structure maps of the limit, and any iterated codegeneracy map $s^I = \phi_*\colon G^n \to G^k$ factors as

$$s^I = \mathrm{proj}_\phi \circ \zeta^n, \tag{1.4}$$

where $\mathrm{proj}_\phi\colon M^n G^\bullet \to G^k$ is the structure map for the copy of G^k indexed by ϕ (see Bousfield and Kan [20, Section X.4.5]).

Note that the inclusion $\mathbf{\Delta}_+ \hookrightarrow \mathbf{\Delta}$ induces a forgetful functor $\mathcal{U}\colon \mathcal{C}^{\mathbf{\Delta}} \to \mathcal{C}^{\mathbf{\Delta}_+}$, and its right adjoint $\mathcal{F}\colon \mathcal{C}^{\mathbf{\Delta}_+} \to \mathcal{C}^{\mathbf{\Delta}}$ is given by $(\mathcal{F}G^\bullet)^n = G^n \times M^n G^\bullet$, with codegeneracies given by (1.4) and coface maps by the cosimplicial identities.

The nth *latching object* for $G^\bullet \in \mathcal{C}^{\mathbf{\Delta}}$ is

$$L^n G^\bullet := \operatorname*{colim}_{\theta:[k]\to[n]} G^k,$$

where the maps θ are now nonidentity injective maps, with $\sigma^n\colon L^n G^\bullet \to G^n$ defined by the structure maps.

These two constructions have analogues for a simplicial object $G_\bullet$ over a (co)complete category $\mathcal{C}$: the *latching object*

$$L_n G_\bullet := \operatorname*{colim}_{\theta:[k]\to[n]} G_k,$$

and the *matching object*

$$M_n G_\bullet := \lim_{\phi:[n]\to[k]} G_k,$$

equipped with the obvious canonical maps.

1.5 Definition Let $\mathcal{C}$ be a pointed category. If it is complete, the nth *Moore chain* object of $G_\bullet \in \mathcal{C}^{\mathbf{\Delta}}$ is

$$C_n G_\bullet := \bigcap_{i=1}^{n} \mathrm{Ker}\{d_i\colon G_n \to G_{n-1}\}, \tag{1.6}$$

with differential $\partial_n^{G_\bullet} = \partial_n := (d_0)|_{C_n G_\bullet}\colon C_n G_\bullet \to C_{n-1} G_\bullet$. The nth *Moore cycle* object is $Z_n G_\bullet := \mathrm{Ker}(\partial_n^{G_\bullet})$.

Dually, if $\mathcal{C}$ is cocomplete, the nth *Moore cochain* object of $G^\bullet \in \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}}$ is

$$C^n G^\bullet := \operatorname{Coker}\Big(\coprod_{i=1}^{n-1} G^n \xrightarrow{\perp_i d^i} G^n\Big) \tag{1.7}$$

with differential

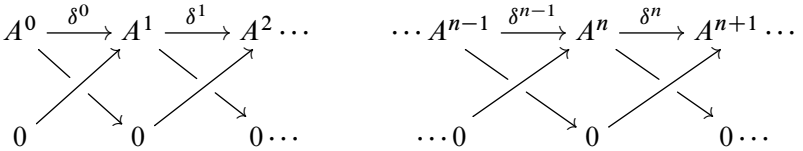
$$\delta^{n-1}\colon C^{n-1}G^\bullet \to C^n G^\bullet$$

induced by d^0_{n-1}, and structure map $v^n\colon G^n \to C^n G^\bullet$.

We denote the cofiber of δ^{n-1} by $Z^n G^\bullet$, with structure map $w^n\colon C^n G^\bullet \to Z^n G^\bullet$.

1.8 Cochain complexes in $\mathcal{C}$ In general, a *cochain complex* in $\mathcal{C}$ is a commuting diagram A^* of the form

(1.9)



(so $\delta^n \circ \delta^{n-1} = 0$ for all n).

We let $\mathsf{Ch}^{\mathcal{C}}$ denote the category of nonnegatively graded cochain complexes over $\mathcal{C}$, and by $\mathsf{Ch}^{\mathcal{C}}_{\leq n}$ the category of *n–truncated* cochain complexes A^* in $\mathcal{C}$ (for which $A^i = 0$ unless $0 \leq i \leq n$).

The category $\mathsf{Ch}_{\mathcal{C}}$ of nonnegatively graded chain complexes over $\mathcal{C}$ is defined analogously.

The Moore cochain functor $C^*\colon \mathcal{C}^{\mathbf{\Delta}+} \to \mathsf{Ch}^{\mathcal{C}}$ has a right adjoint (and left inverse) $\mathcal{E}\colon \mathsf{Ch}^{\mathcal{C}} \to \mathcal{C}^{\mathbf{\Delta}+}$ with $(\mathcal{E}\boldsymbol{A}_*)^n = A^n$, $d^0_n = \delta^n$, and $d^i_n = 0$ for $i \geq 1$. This holds also for $\mathsf{Ch}^{\mathcal{C}}_{\leq n}$ if we truncate $\mathcal{C}^{\mathbf{\Delta}+}$, too.

When $\mathcal{C}$ is a model category, we have several possible model category structures on $\mathcal{C}^{\mathbf{\Delta}}$, $\mathcal{C}^{\mathbf{\Delta}+}$, $\mathsf{Ch}^{\mathcal{C}}$, and $\mathsf{Ch}^{\mathcal{C}}_{\leq n}$ (see eg Hirschhorn [28, Section 15.3], Bousfield and Kan [19] and Chachólski and Scherer [21, Section 12]).

In particular, $\mathsf{Ch}^{\mathcal{C}}_{\leq n}$ has two different Reedy model category structures, depending on how we choose the degrees in (1.9); in both cases, the weak equivalences are defined levelwise. In the *right Reedy* model structure, fibrations are also defined levelwise, and an n–cochain complexes A^* is cofibrant if for each $k \leq n$ the natural map $Z^k A^* := \operatorname{Cof}(\delta^{k-1}) \to A^{k+1}$ is a cofibration (with $A^{-1} := *$). In the *left Reedy* model structure, cofibrations are defined levelwise, and an n–cochain complex A^*

is fibrant if for each $k \le n$ the natural map $A^k \to \mathrm{Ker}(\delta^{k+1})$ is a fibration (with $A^{-1} := *$). Evidently, right Reedy cofibrancy implies left Reedy cofibrancy, and left Reedy fibrancy implies right Reedy fibrancy.

Note that the Moore cochains functor $C^*\colon \mathcal{S}_*^{\mathbf{\Delta}} \to \mathsf{Ch}^{\mathcal{S}_*}$ preserves cofibrancy and weak equivalences among cofibrant objects in the (right) Reedy model structures (see Bousfield and Kan [20, Proposition X.6.3]).

By analogy with the usual fiber/cone construction for (co)chain complexes we have:

1.10 Fibers and cones (i) For any map $f\colon A^\bullet \to B^\bullet$ in $\mathcal{C}^{\mathbf{\Delta}+}$ we define the restricted cosimplicial object $C^\bullet = \mathrm{Fib}(f)$ by setting $C^n := A^n \times B^{n-1}$, with coface maps

$$\begin{array}{ccccc} C^{n+1} & = & A^{n+1} & \times & B^n \\ {\scriptstyle d^0}\uparrow & & {\scriptstyle d^0}\uparrow & \nearrow_{f^n} & \\ C^n & = & A^n & \times & B^{n-1} \end{array} \quad \text{and} \quad \begin{array}{ccccc} C^{n+1} & = & A^{n+1} & \times & B^n \\ {\scriptstyle d^i}\uparrow & & {\scriptstyle d^i}\uparrow & & {\scriptstyle d^{i-1}}\uparrow \\ C^n & = & A^n & \times & B^{n-1} \end{array} \quad \text{for } i \ge 1$$

and a natural projection $\ell\colon C^\bullet \to A^\bullet$. For $d^j \circ d^0\colon C^{n-1} \to C^{n+1}$ we verify that

$$\begin{aligned} \mathrm{pr}_{B^n} \circ d^j \circ d^0 &= d^{j-1}_{B^{n-1}} \circ f^{n-1} \circ \mathrm{pr}_{A^{n-1}} \\ &= f^n \circ d^{j-1}_{A^{n-1}} \circ \mathrm{pr}_{A^{n-1}} = \mathrm{pr}_{B_n} \circ d^0 \circ d^{j-1} \end{aligned}$$

for all $j > 0$, while clearly $\mathrm{pr}_{B^n}(d^j \circ d^i) = \mathrm{pr}_{B^n}(d^i \circ d^{j-1})$ for $1 \le i < j$.

(ii) Similarly, for a map $f\colon A_\bullet \to B_\bullet$ in $\mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+}$ we define the restricted simplicial object $C_\bullet = \mathrm{Cone}(f)$ by setting $C_n := A_{n-1} \amalg B_n$, with

$$d_i^{C_n} := \begin{cases} f_{n-1} \perp d_0^{B_n} & \text{if } i = 0, \\ d_{i-1}^{A_{n-1}} \perp d_i^{B_n} & \text{if } i \ge 1, \end{cases}$$

and a natural inclusion $m\colon B_\bullet \hookrightarrow C_\bullet$.

1.11 Simplicial CW objects A simplicial object $G_\bullet \in \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}}$ over a pointed category $\mathcal{C}$ is called a *CW object* if it is equipped with a *CW basis* $(\overline{G}_n)_{n\in\mathbb{N}}$ in $\mathcal{C}$ such that $G_n = \overline{G}_n \amalg L_n G_\bullet$, and $d_i|_{\overline{G}_n} = 0$ for $1 \le i \le n$. In this case, $\overline{\partial}_0^{G_n} := d_0|_{\overline{G}_n}\colon \overline{G}_n \to G_{n-1}$ is called the *attaching map* for $\overline{G}_n$. By the simplicial identities $\overline{\partial}_0^{G_n}$ factors as

$$\overline{\partial}_0^{G_n}\colon \overline{G}_n \to Z_{n-1}G_\bullet \subset G_{n-1}. \tag{1.12}$$

Note that we have an explicit formula

$$L_n G_\bullet := \coprod_{0\le k<n} \ \coprod_{0\le i_1<\cdots<i_{n-k}\le n-1} \overline{G}_k, \tag{1.13}$$

where the iterated degeneracy map $s_{i_{n-k}} \dots s_{i_2} s_{i_1}$, restricted to the basis object $\overline{G}_k$, is the inclusion into the copy of $\overline{G}_k$ indexed by $(i_1, \dots, i_{n-k})$.

1.14 Remark Given $\overline{G} \in \mathcal{C}$, define $\overline{G} \otimes_* S^n$ in $\mathsf{Ch}_{\mathcal{C}}$ be the chain complex with $\overline{G}$ in dimension n and $*$ elsewhere. A CW object $G_\bullet$ over $\mathcal{C}$ with CW basis $(\overline{G}_n)_{n=0}^{\infty}$ is the colimit of an inductively constructed sequence of skeleta $\mathrm{sk}_0 G_\bullet \hookrightarrow \mathrm{sk}_1 G_\bullet \hookrightarrow \cdots$, in the usual sense (see Goerss and Jardine [25, Section VII.1]), starting with $\mathrm{sk}_0 G_\bullet := c(\overline{G}_0)_\bullet$. To do so, note that the attaching map $\overline{\partial}_0^{G_n} \colon \overline{G}_n \to Z_{n-1} G_\bullet$ defines a chain map

$$\phi \colon \overline{G}_n \otimes_* S^{n-1} \to C_* \,\mathrm{sk}_{n-1} G_\bullet,$$

which has an adjoint $\widetilde{\phi} \colon \mathcal{E}'(\overline{G}_n \otimes_* S^{n-1}) \to \mathcal{U}' \,\mathrm{sk}_{n-1} G_\bullet$, where $\mathcal{E}' \colon \mathsf{Ch}_{\mathcal{C}} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+}$ is left adjoint to the Moore chain functor $C_* \colon \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+} \to \mathsf{Ch}_{\mathcal{C}}$ and $\mathcal{U}' \colon \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+}$ is the forgetful functor (compare Section 1.8).

Note that $\mathcal{U}'$ has a left adjoint $\mathcal{F}' \colon \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}}$ given by $(\mathcal{F}' G_\bullet)_n = G_n \amalg L_n G_\bullet$, (compare Section 1.2). If $\vartheta \colon \mathcal{F}'\mathcal{U}' \to \mathrm{Id}$ is the counit for the adjunction, we have $\mathrm{sk}_n G_\bullet$ as the pushout in

$$\begin{array}{ccc} \mathcal{F}'\mathcal{U}' \,\mathrm{sk}_{n-1} G_\bullet & \xrightarrow{\vartheta} & \mathrm{sk}_{n-1} G_\bullet \\ {\scriptstyle \mathcal{F}' m}\downarrow & \boxed{\mathrm{PO}} & \downarrow \\ \mathcal{F}' \,\mathrm{Cone}(\widetilde{\phi}) & \longrightarrow & \mathrm{sk}_n G_\bullet \end{array} \tag{1.15}$$

for m as in Section 1.10(ii).

1.16 Cosimplicial CW objects A cosimplicial CW object $G^\bullet \in \mathcal{C}^{\mathbf{\Delta}}$ with CW basis $(\overline{G}^n)_{n \in \mathbb{N}}$ may be defined analogously as the limit of a tower of coskeleta

$$\cdots \to \mathrm{csk}^2 G^\bullet \to \mathrm{csk}^1 G^\bullet \to \mathrm{csk}^0 G^\bullet$$

(see [loc. cit.]), starting with $\mathrm{csk}^0 G^\bullet := c(\overline{G}^0)^\bullet$, by thinking of its attaching maps as a cochain map $\varphi \colon C^* \,\mathrm{csk}^{n-1} G^\bullet \to \overline{G}^n \otimes^* S^{n-1}$ (where $\overline{G}^n \otimes^* S^{n-1}$ is the cochain complex with $\overline{G}^n$ in dimension $n-1$ and zero elsewhere).

The map φ has an adjoint $\widehat{\varphi} \colon \mathcal{U} \,\mathrm{csk}^{n-1} G^\bullet \to \mathcal{E}(\overline{G}^n \otimes^* S^{n-1})$, and we have $\mathrm{csk}^n G^\bullet$ as the pullback in

$$\begin{array}{ccc} \mathrm{csk}^n G^\bullet & \longrightarrow & \mathcal{F}\,\mathrm{Fib}(\widehat{\varphi}) \\ \downarrow & \boxed{\mathrm{PB}} & \downarrow {\scriptstyle \mathcal{F}\ell} \\ \mathrm{csk}^{n-1} G^\bullet & \xrightarrow{\theta} & \mathcal{F}\mathcal{U} \,\mathrm{csk}^{n-1} G^\bullet \end{array} \tag{1.17}$$

with θ the unit for $\mathcal{F}\mathcal{U}$ and ℓ as in Section 1.10(i), using the notation of Section 1.2.

1.18 List of functors For the reader's convenience we list the main functors we have defined for simplicial and cosimplicial objects in a category $\mathcal{C}$:

(a) The Moore cochain complex functor $C^*\colon \mathcal{C}^{\mathbf{\Delta}+} \to \mathsf{Ch}^{\mathcal{C}}$ and its right adjoint (and left inverse) $\mathcal{E}\colon \mathsf{Ch}^{\mathcal{C}} \to \mathcal{C}^{\mathbf{\Delta}+}$.

(b) The Moore chain complex functor $C_*\colon \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+} \to \mathsf{Ch}_{\mathcal{C}}$ and its left adjoint (and right inverse) $\mathcal{E}'\colon \mathsf{Ch}_{\mathcal{C}} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+}$.

(c) The forgetful functor $\mathcal{U}\colon \mathcal{C}^{\mathbf{\Delta}} \to \mathcal{C}^{\mathbf{\Delta}+}$ and its right adjoint $\mathcal{F}\colon \mathcal{C}^{\mathbf{\Delta}+} \to \mathcal{C}^{\mathbf{\Delta}}$ (adding codegeneracies).

(d) The forgetful functor $\mathcal{U}'\colon \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+}$ and its left adjoint $\mathcal{F}'\colon \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}_+} \to \mathcal{C}^{\mathbf{\Delta}^{\mathrm{op}}}$ (adding degeneracies).

When there is no danger of confusion, we denote $C^*(\mathcal{U}(\boldsymbol{W}^\bullet))$ simply by $C^*\boldsymbol{W}^\bullet$ and $C_*(\mathcal{U}'(G_\bullet))$ by $C_*G_\bullet$.

1.19 Definition A *simplicial* category $\mathcal{C}$ (in the sense of Quillen) is one in which, for each (finite) $K \in \mathcal{S}$ and $\boldsymbol{X} \in \mathcal{C}$, we have objects $\boldsymbol{X} \otimes K$ and $\boldsymbol{X}^K$ in $\mathcal{C}$ equipped with appropriate adjunction-like isomorphisms. In particular, such categories are simplicially enriched. A *simplicial model category* is a simplicial category with a model category structure satisfying axiom SM7 of Quillen [32, Sections II.1–II.2]). The basic examples are $\mathcal{S}$ and $\mathcal{S}_*$.

1.20 Assumption From now on $\mathcal{C}$ will be a pointed simplicial model category in which all objects are cofibrant — so in particular it is left proper.

The main example we shall be concerned with is $\mathcal{C} = \mathcal{S}_*$, so we shall sometimes refer to the objects of $\mathcal{C}$ — denoted by boldface letters $\boldsymbol{X}$, $\boldsymbol{Y}$, and so on — as "spaces".

1.21 $\mathcal{G}$–resolution model structure Let $\mathcal{G}$ be a class of homotopy group objects in a model category $\mathcal{C}$ as above, closed under loops. A map $i\colon \boldsymbol{A} \to \boldsymbol{B}$ in $\mathrm{ho}\,\mathcal{C}$ is called *$\mathcal{G}$–monic* if $i^*\colon [\boldsymbol{B}, \boldsymbol{G}] \to [\boldsymbol{A}, \boldsymbol{G}]$ is onto for each $\boldsymbol{G} \in \mathcal{G}$. An object $\boldsymbol{Y}$ in $\mathcal{C}$ is called *$\mathcal{G}$–injective* if $i^*\colon [\boldsymbol{B}, \boldsymbol{Y}] \to [\boldsymbol{A}, \boldsymbol{Y}]$ is onto for each $\mathcal{G}$–monic map $i\colon \boldsymbol{A} \to \boldsymbol{B}$ in $\mathrm{ho}\,\mathcal{C}$. A fibration in $\mathcal{C}$ is called *$\mathcal{G}$–injective* if it has the right lifting property for the $\mathcal{G}$–monic cofibrations in $\mathcal{C}$.

The homotopy category $\mathrm{ho}\,\mathcal{C}$ is said to have *enough $\mathcal{G}$–injectives* if each object is the source of a $\mathcal{G}$–monic map to a $\mathcal{G}$–injective target. In this case, $\mathcal{G}$ is called a class of *injective models* in $\mathrm{ho}\,\mathcal{C}$.

Recall that a homomorphism in the category $s\mathsf{Gp}$ of simplicial groups is a weak equivalence or fibration when its underlying map in $\mathcal{S}_*$ is such.

A map $f\colon \boldsymbol{W}^\bullet \to \boldsymbol{Y}^\bullet$ in $\mathcal{C}^{\boldsymbol{\Delta}}$ is called a $\mathcal{G}$–*equivalence* if $f^*\colon [\boldsymbol{Y}^\bullet, \boldsymbol{G}] \to [\boldsymbol{W}^\bullet, \boldsymbol{G}]$ is a weak equivalence in $s\mathsf{Gp}$ for each $\boldsymbol{G} \in \mathcal{G}$. Bousfield [19, Theorem 3.3] showed that if $\mathcal{G}$ is a class of injective models in $\mathrm{ho}(\mathcal{C})$, then $\mathcal{C}^{\boldsymbol{\Delta}}$ has a left proper pointed simplicial model category structure with such maps as weak equivalences.

1.22 $\mathcal{G}$–completion Given a class $\mathcal{G}$ of injective models in $\mathcal{C}$ as above, a $\mathcal{G}$–*resolution* of an object $\boldsymbol{Y} \in \mathcal{C}$ is a $\mathcal{G}$–fibrant $\boldsymbol{W}^\bullet$ equipped with a $\mathcal{G}$–trivial cofibration $c(\boldsymbol{Y})^\bullet \hookrightarrow \boldsymbol{W}^\bullet$ (see Definition 1.1). In this case $\widehat{L}_{\mathcal{G}}\boldsymbol{Y} := \operatorname{Tot}\boldsymbol{W}^\bullet$ is called the $\mathcal{G}$–*completion* of Y, where $\operatorname{Tot}\boldsymbol{W}^\bullet \in \mathcal{C}$ is constructed as in [19, Section 2.8]. Moreover, a map $f\colon \boldsymbol{Y} \to \boldsymbol{Z}$ in $\mathcal{C}$ is a $\mathcal{G}$–equivalence if and only if $\widehat{L}_{\mathcal{G}} f\colon \widehat{L}_{\mathcal{G}}\boldsymbol{Y} \to \widehat{L}_{\mathcal{G}}\boldsymbol{Z}$ is a weak equivalence in $\mathcal{C}$. An object $\boldsymbol{Y} \in \mathcal{C}$ is called $\mathcal{G}$–*complete* if $\boldsymbol{Y} \to \widehat{L}_{\mathcal{G}}\boldsymbol{Y}$ is a weak equivalence, and it is $\mathcal{G}$–*good* if $\boldsymbol{Y} \to \widehat{L}_{\mathcal{G}}\boldsymbol{Y}$ is a $\mathcal{G}$–equivalence — so $\widehat{L}_{\mathcal{G}}\boldsymbol{Y}$ is $\mathcal{G}$–complete (see Bousfield [19, Section 8]). We say that two maps $f_0, f_1\colon \boldsymbol{Z} \to \boldsymbol{Y}$ between $\mathcal{G}$–good objects in $\mathcal{C}$ are $\mathcal{G}$–*equivalent* if $\widehat{L}_{\mathcal{G}} f_0 \sim \widehat{L}_{\mathcal{G}} f_1\colon \widehat{L}_{\mathcal{G}}\boldsymbol{Z} \to \widehat{L}_{\mathcal{G}}\boldsymbol{Y}$ are homotopic (for a suitable fibrant and cofibrant model of the $\mathcal{G}$–completion).

A cosimplicial object $\boldsymbol{W}^\bullet \in \mathcal{C}^{\boldsymbol{\Delta}}$ is called *weakly $\mathcal{G}$–fibrant* if it is Reedy fibrant (see Hirschhorn [28, Section 15.3]), and every $\boldsymbol{W}^n$ is in $\mathcal{G}$ for $n \geq 0$. A *weak $\mathcal{G}$–resolution* of an object $\boldsymbol{Y} \in \mathcal{C}$ is a weakly $\mathcal{G}$–fibrant $\boldsymbol{W}^\bullet$ which is $\mathcal{G}$–equivalent to $c(\boldsymbol{Y})^\bullet$. In this case there is a natural weak equivalence $\widehat{L}_{\mathcal{G}}\boldsymbol{Y} \to \operatorname{Tot}\boldsymbol{W}^\bullet$, by [19, Theorem 6.5].

1.23 Example The example we have in mind is the class $\mathcal{G} = \mathcal{G}_R$ of all R–GEMs in $\mathcal{C} = \mathcal{S}_*$, for some ring R. In this case the $\mathcal{G}$–completion is the Bousfield–Kan R–completion [20].

1.24 Sketches and their algebras An (FP–)*sketch* in the sense of Ehresmann [22] (cf Lawvere [29] is a small category Θ with a specified set $\mathcal{P}$ of (finite) products. This generalizes Lawvere's notion of a *theory*, which requires that $\mathrm{Obj}(\Theta) = \mathbb{N}$. A Θ–*algebra* (or Θ–model) is a functor $\Gamma\colon \Theta \to \mathsf{Set}_*$ preserving the products in $\mathcal{P}$, with natural transformations as model morphisms (see Borceux [18, Section 4.1]). We write $\Gamma\{\boldsymbol{B}\}$ for the value of Γ at $\boldsymbol{B} \in \Theta$. The category of Θ–algebras is denoted by Θ–Alg. Examples of categories of such models include varieties of universal algebras, such as groups, rings, and modules over rings.

1.25 Remark In homotopy theory such theories often arise by choosing a collection $\mathcal{A}$ of objects in a model category $\mathcal{C}$ (as in Assumption 1.20), and letting $\Theta = \Theta_{\mathcal{A}}$ denote

the full subcategory of the homotopy category $\mathrm{ho}\,\mathcal{C}$ consisting of objects of $\mathcal{C}$ which are representable as *finite-type* products of objects in $\mathcal{A}$ — that is, products of the form

$$\prod_{\boldsymbol{A}\in\mathcal{A}}\ \prod_{i\in I_{\boldsymbol{A}}} \boldsymbol{A}, \tag{1.26}$$

where each indexing set $I_{\boldsymbol{A}}$ is finite. When each $\boldsymbol{A}\in\mathcal{A}$ is a homotopy group object in $\mathcal{C}$, Θ–algebras have a natural underlying group structure; in such cases we call the sketch Θ *algebraic*.

For every $\boldsymbol{Y}\in\mathcal{C}$ we then have a *realizable* Θ–algebra $\Gamma=H^*_\Theta\boldsymbol{Y}$, with $\Gamma\{\boldsymbol{B}\}=[\boldsymbol{Y},\boldsymbol{B}]_{\mathcal{C}}$ for every $\boldsymbol{B}\in\Theta$. A Θ–algebra is called *free* if is isomorphic to $H^*_\Theta\boldsymbol{B}$ for some $\boldsymbol{B}\in\Theta$.

Note, however, that even though the forgetful functor $U\colon \Theta\text{–}\mathsf{Alg}\to \mathsf{Set}^{\mathcal{A}}$ to $\mathcal{A}$–graded sets has a left adjoint $F\colon \mathsf{Set}^{\mathcal{A}}\to\Theta\text{–}\mathsf{Alg}$, not all Θ–algebras in its image are free by our definition: because F preserves colimits, any Θ–algebra in the image of F is a coproduct of *monogenic* free Θ–algebras (those realizable by on object of $\mathcal{A}$), and conversely. For our purposes, however (see Theorem 2.33 below), it is necessary that any free Θ–algebra be realizable in $\mathcal{C}$.

1.27 Example For any commutative ring R, let $\mathcal{A}:=\{\boldsymbol{K}(R,n)\}_{n=1}^{\infty}$ in $\mathcal{C}=\mathcal{S}_*$; we then have an FP–sketch Θ_R in $\mathrm{ho}\,\mathcal{S}_*$ whose objects are "finite-type" R–GEMs of the form $\prod_{n=1}^{\infty}\boldsymbol{K}(V_n,n)$, where $V_n=R^{k_n}$ is a free R–module of dimension $k_n<\infty$. Since each $\boldsymbol{B}\in\Theta_R$ is an R–module object, all Θ_R–algebras take values in R–modules.

Note that the realizable Θ_R–algebra $\Gamma=H^*_{\Theta_R}\boldsymbol{Y}$ has $\Gamma\{\boldsymbol{K}(R,n)\}=H^n(\boldsymbol{Y};R)$, so we denote it simply by $H^*(\boldsymbol{Y};R)$; it is the R–cohomology algebra of $\boldsymbol{Y}$, equipped with the action of the primary R–cohomology operations. When $R=\mathbb{F}_p$, Θ_R is the homotopy category of such $\mathbb{F}_p$–GEMs, and a Θ_R–algebra is an unstable algebra over the mod p Steenrod algebra, as in Schwartz [34, Section 1.4]. When $R=\mathbb{Q}$, a Θ_R–algebra is just a graded commutative $\mathbb{Q}$–algebra.

More generally, for any limit cardinal λ, let $R\text{–}\mathsf{Mod}^{\lambda}$ denote the set of isomorphism types of free R–modules of dimension $<\lambda$, and let

$$\mathcal{A}^{\lambda}:=\{\boldsymbol{K}(V,n)\mid V\in R\text{–}\mathsf{Mod}^{\lambda},\ 1\le n\le\infty\}.$$

The FP–sketch Θ_R^{λ} then consists of the R–GEMs which are finite-type products of the form $\prod_{n=1}^{\infty}\boldsymbol{K}(V_n,n)$ with V_n a free R–module of dimension $<\lambda$.

Note that since finite products in R–Mod are also coproducts, $\prod_{I_n} V_i$ is itself in R–Mod^λ, since I_n is finite and λ is infinite. However, we must be able to distinguish the generating set $\mathcal{A}$ from the resulting sketch Θ_R, in order to define what we mean by "finite-type" products or coproducts.

We note the following two standard facts about an algebraic sketch Θ:

1.28 Lemma *If Γ is an Θ–algebra and $\boldsymbol{B} \in \Theta$, there is a natural isomorphism* $\mathrm{Hom}_{\Theta\text{–}\mathsf{Alg}}(H^*_\Theta \boldsymbol{B}, \Gamma) \cong \Gamma\{\boldsymbol{B}\}$.

1.29 Proposition [16, Section 6] *The category $s\Theta$–Alg of simplicial Θ–algebras has a model category structure, in which the weak equivalences and fibrations are defined on the underlying (graded) simplicial groups.*

1.30 Definition For any algebraic sketch Θ, a *CW resolution* of a Θ–algebra Γ is a cofibrant replacement $\varepsilon\colon V_\bullet \xrightarrow{\simeq} c(\Gamma)_\bullet$ equipped with a CW basis $(\overline{V}_n)_{n\in\mathbb{N}}$ as in Section 1.11, with each $\overline{V}_n$ a free Θ–algebra.

1.31 Remark In fact, any CW object $V_\bullet$ for which each $\overline{V}_n$ is a free Θ–algebra and each attaching map $\overline{\partial}_0^{V_n}$ for $n \geq 0$ surjects onto $Z_{n-1}V_\bullet$ is a CW resolution. Here we set $Z_{-1}V_\bullet := \Gamma$ and $\overline{\partial}_0^{V_0} := \varepsilon$, so that

$$\varepsilon \circ \overline{\partial}_0^{V_1} = 0. \tag{1.32}$$

Thus one can easily construct (nonfunctorial) CW resolutions of any Θ–algebra Γ.

Moreover, by dualizing [8, Proposition 3.12] (see also [7, Proposition 12]) we can choose a CW basis for any free simplicial resolution in Θ_R–Alg when R is a field. However, we shall not make use of this fact.

2 Realizing simplicial Θ–algebra resolutions

Let Θ be an algebraic sketch obtained as in Remark 1.25 from a set $\mathcal{A}$ of homotopy group objects in a model category $\mathcal{C}$. In this section we show how a CW resolution $V_\bullet$ of a realizable Θ–algebra Γ can itself be realized over $\mathcal{C}$.

Any algebraic resolution $V_\bullet \to \Gamma$ is clearly realizable by a cosimplicial object $\widetilde{\boldsymbol{W}}^\bullet \in c(\mathrm{ho}\,\mathcal{C})$ (for which the cosimplicial identities hold up to homotopy), but it is not clear a priori that this can be rectified to a strict cosimplicial object $\boldsymbol{W}^\bullet$ in $\mathcal{C}$. This is in fact

possible in the two cases described in Theorem A.11 in the appendix. However, for the purposes of this paper, we need not only the existence of this realization $\boldsymbol{W}^\bullet$, but also the particular form it takes, described as follows:

2.1 Definition Let Θ be a sketch in $\mathrm{ho}\,\mathcal{C}$ as in Remark 1.25, and let $V_\bullet$ be a simplicial CW resolution of the Θ–algebra $H^*_\Theta \boldsymbol{Y}$ for some $\boldsymbol{Y} \in \mathcal{C}$, with CW basis $(\overline{V}_n)_{n=0}^\infty$. A *sequential realization*

$$\mathcal{W} = \langle \boldsymbol{W}^\bullet_{[n]}, \widetilde{\boldsymbol{W}}^\bullet_{[n]} \rangle_{n\in\mathbb{N}}$$

of $V_\bullet$ for $\boldsymbol{Y}$ consists of a tower of Reedy fibrant and cofibrant $\boldsymbol{Y}$–coaugmented cosimplicial objects

$$\cdots \to \boldsymbol{W}^\bullet_{[n]} \xrightarrow{\pi_{[n]}} \boldsymbol{W}^\bullet_{[n-1]} \xrightarrow{\pi_{[n-1]}} \boldsymbol{W}^\bullet_{[n-2]} \to \cdots \to \boldsymbol{W}^\bullet_{[0]} \tag{2.2}$$

in $\mathcal{C}^{\boldsymbol{\Delta}}$, with each $\pi_{[n]}$ a Reedy fibration. We call $\boldsymbol{W}^\bullet_{[n]}$ the nth *stage* of $\mathcal{W}$, and set $\boldsymbol{W}^\bullet := \mathrm{holim}_n \boldsymbol{W}^\bullet_{[n]}$ to be the limit of the tower (2.2). This tower must satisfy the following requirements for each $n \geq 0$:

(a) The coaugmentation $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Y} \to \boldsymbol{W}^\bullet_{[n]}$ realizes $V_\bullet \to H^*_\Theta \boldsymbol{Y}$ through simplicial dimension n — that is, we have a natural isomorphism

$$H^*_\Theta \boldsymbol{W}^k_{[n]}\{\boldsymbol{B}\} \xrightarrow{\cong} V_k\{\boldsymbol{B}\} \quad \text{for all } \boldsymbol{B} \in \Theta \text{ and } -1 \leq k \leq n, \tag{2.3}$$

using the notation of Example 1.27.

(b) The coaugmentation $\boldsymbol{\varepsilon}_{[n-1]}\colon \boldsymbol{Y} \to \boldsymbol{W}^\bullet_{[n-1]}$ lifts along the $\pi_{[n]}\colon \boldsymbol{W}^\bullet_{[n]} \to \boldsymbol{W}^\bullet_{[n-1]}$ to $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Y} \to \boldsymbol{W}^\bullet_{[n]}$.

(c) Each $\boldsymbol{W}^\bullet_{[n]}$ is obtained from $\boldsymbol{W}^\bullet_{[n-1]}$ as follows: assume given a fibrant realization $\overline{\boldsymbol{W}}^n \in \mathcal{C}$ of $\overline{V}_n$ and let $\boldsymbol{D}^* \in \mathsf{Ch}^{\mathcal{C}}$ be a left Reedy fibrant replacement of $\overline{\boldsymbol{W}}^n \otimes^* S^{n-1}$. Assume that we can realize the nth attaching map $\overline{\partial}^{V_n}_0\colon \overline{V}_n \to V_{n-1}$ of $V_\bullet$ by a cochain map $F\colon C^*\boldsymbol{W}^\bullet_{[n-1]} \to \boldsymbol{D}^*$ in the category of coaugmented cochain complexes (that is, defined in dimensions $-1 \leq k \leq n-1$). Note that $(\overline{\boldsymbol{W}}^n \otimes^* S^{n-1})^{-1} = 0$.

Let $\widetilde{F}\colon \mathcal{U}\boldsymbol{W}^\bullet_{[n-1]} \to \mathcal{E}\boldsymbol{D}^*$ the corresponding map of restricted cosimplicial objects as in Section 1.16, and set $\widetilde{\boldsymbol{W}}^\bullet_{[n]} := \mathrm{Fib}(\widetilde{F})$ as in Section 1.10, take a pullback

$$\begin{array}{ccc} \widehat{\boldsymbol{W}}^\bullet_{[n]} & \longrightarrow & \mathcal{F}\widetilde{\boldsymbol{W}}^\bullet_{[n]} \\ \big\downarrow & \boxed{\text{PB}} & \big\downarrow{\scriptstyle \mathcal{F}\ell} \\ \boldsymbol{W}^\bullet_{[n-1]} & \xrightarrow{\theta} & \mathcal{F}\mathcal{U}\boldsymbol{W}^\bullet_{[n-1]} \end{array} \tag{2.4}$$

as in (1.17), with $W^{\bullet}_{[n]}$ a Reedy fibrant and cofibrant replacement (see Hirschhorn [28, Section 15.3]) for $\widehat{W}^n$.

We start the process with $W^{\bullet}_{[0]} := c(\overline{W}^0)^{\bullet}$.

(d) For reasons which will appear later, we sometimes require the chain map $F\colon C^*W^{\bullet}_{[n-1]} \to D^*$ in step (c) to be a levelwise (ie left Reedy) cofibration. If this is true for each $n \geq 1$, we say that the sequential realization $\mathcal{W}$ is *cofibrant*.

2.5 Understanding the passage from $W^{\bullet}_{[n-1]}$ to $W^{\bullet}_{[n]}$ We now give an explicit description of the construction of $W^{\bullet}_{[n]}$ from $W^{\bullet}_{[n-1]}$, also introducing some auxiliary notation:

(i) What does a left Reedy fibrant replacement for $\overline{W}^n \otimes^* S^{n-1}$ look like? In order for the coaugmented cochain complex D^* in $\mathsf{Ch}^{\mathcal{C}}$ to be left Reedy fibrant, the structure map $D^{k-1} \to \mathrm{Ker}(\delta^k_D)$ (into the "k–cocycles" — not the same as Z^kD^* of Definition 1.5) must be a fibration for each k. In addition, since $(\overline{W}^n \otimes^* S^{n-1})^k = 0$ for $k \neq n-1$, D^k must be contractible.

It is natural for such a D^* to be described by a downward induction, starting with $D^n = 0$ and $D^{n-1} = \overline{W}^n$. Thus $\mathrm{Ker}(\delta^{n-1}_D) = \overline{W}^n$, too, so D^{n-2} must be a contractible object equipped with a fibration to $\overline{W}^n$ — ie it is a path object for $\overline{W}^n$, in the sense of Quillen [32, Section I.2]. We could choose it to be the standard path object $P\overline{W}^n$ of Equation (A.3), but this may not be a good choice if we also want $F^k\colon C^kW^{\bullet}_{[n-1]} \to D^k$ to be a cofibration for each k, as in Definition 2.1(d). Thus we let $\delta^{n-2}_D\colon D^{n-2} \to D^{n-1}$ be some fibration $\bar{p}^0\colon \overline{PW^n} \to \overline{W}^n$ with $\overline{PW^n}$ contractible.

At the $(n-3)^{\mathrm{rd}}$ stage we see that $\mathrm{Ker}(\delta^{n-2}_D)$ is the fiber $\overline{\Omega^1W^n}$ of $\bar{p}^0$, and again $\delta^{n-3}_D\colon D^{n-3} \to D^{n-2}$ must be a fibration $\bar{p}^1\colon \overline{P\Omega^1W^n} \to \overline{\Omega^1W^n}$ (with $\overline{P\Omega^1W^n}$ contractible), followed by the inclusion $\bar{\imath}^1\colon \overline{\Omega^1W^n} \hookrightarrow \overline{P\Omega^0W^n}$.

Proceeding by induction, for each $-1 \leq j \leq n-1$ we obtain a fibration sequence

$$\overline{\Omega^{j+1}W^n} \xhookrightarrow{\bar{\imath}^{j+1}} \overline{P\Omega^jW^n} \xrightarrow{\bar{p}^j} \overline{\Omega^jW^n} \tag{2.6}$$

in $\mathcal{C}$, called the *j^{th} modified path-loop fibration*, with $\overline{P\Omega^jW^n}$ contractible. By convention we set $\overline{W}^n := \overline{\Omega^0W^n} = \overline{P\Omega^{-1}W^n}$, with the identity map as

$$\bar{\imath}^0\colon \overline{\Omega^0W^n} \xrightarrow{=} \overline{P\Omega^{-1}W^n}. \tag{2.7}$$

Thus $D^k = \overline{P\Omega^{n-k-2}W^n}$ for each $-1 \leq k \leq n-1$, and the differential $\delta^k_D\colon D^k \to D^{k+1}$ is the composite

$$\overline{P\Omega^{n-k-2}W^n} \xrightarrow{\bar{p}^{n-k-2}} \overline{\Omega^{n-k-2}W^n} \xrightarrow{\bar{\imath}^{n-k-2}} \overline{P\Omega^{n-k-3}W^n}. \tag{2.8}$$

For later use we may (and shall) assume (see Step VIII in proof of Theorem A.11) that for each $0 \le j \le n-1$, (2.6) fits into a commutative diagram

$$\begin{array}{ccccc} \overline{\Omega^{j+1}W^n} & \xhookrightarrow{\bar{\iota}^{j+1}} & \overline{P\Omega^j W^n} & \xrightarrow{\bar{p}^j} & \overline{\Omega^j W^n} \\ \simeq\downarrow{\scriptstyle \sigma^{j+1}} & & \simeq\downarrow{\scriptstyle P\tau^j} & & \downarrow= \\ \Omega\overline{\Omega^j W^n} & \xhookrightarrow{\iota} & P\overline{\Omega^j W^n} & \xrightarrow{p} & \overline{\Omega^j W^n} \end{array} \tag{2.9}$$

in which the bottom sequence is the usual path-loop fibration for $\overline{\Omega^j W^n}$, and the vertical maps are all trivial fibrations.

(ii) The map $F\colon C^*W^\bullet_{[n-1]} \to D^*$ and its fiber As in Definition 2.1(c), the nth attaching map for $V_\bullet$ is to be realized by a cochain map $F\colon C^*W^\bullet_{[n-1]} \to D^*$ in the category of coaugmented cochain complexes given by maps $F^k\colon C^kW^\bullet_{[n-1]} \to D^k = \overline{P\Omega^{n-k-2}W^n}$ for each $-1 \le k \le n-1$, so by Section 1.10 the restricted cosimplicial object $\widetilde{W}^\bullet_{[n]} = \mathrm{Fib}(\widetilde{F})$ is given by

$$\widetilde{W}^k_{[n]} := W^k_{[n-1]} \times \overline{P\Omega^{n-k-1}W^n} \tag{2.10}$$

in dimension $0 \le k \le n$, while by (2.7) we have

$$\widetilde{W}^n_{[n]} := W^n_{[n-1]} \times \overline{W}^n. \tag{2.11}$$

We denote the two structure maps for the product (2.10) by

$$\psi^k_{[n]}\colon \widetilde{W}^k_{[n]} \to W^k_{[n-1]}, \quad q^k_{[n]}\colon \widetilde{W}^k_{[n]} \to \overline{P\Omega^{n-k-1}W^n}, \tag{2.12}$$

respectively. By Section 1.10 we see that the coface maps $\widetilde{d}^i_k\colon \widetilde{W}^k_{[n]} \to \widetilde{W}^{k+1}_{[n]}$ are determined by

$$F^k \circ v^k \circ \psi^k_{[n]} = q^{k+1}_{[n]} \circ \widetilde{d}^0_k, \tag{2.13}$$

where $v^k\colon W^k_{[n-1]} \to C^kW^\bullet_{[n-1]}$ is the structure map for (1.7). For $i = 1$ we have

$$\delta^{k-1}_{\boldsymbol{D}} \circ q^k_{[n]} = \bar{\iota}^{n-k-1} \circ \bar{p}^{n-k-1} \circ q^k_{[n]} = q^{k+1}_{[n]} \circ \widetilde{d}^1_k, \tag{2.14}$$

while for $i \ge 2$ we have

$$q^{k+1}_{[n]} \circ \widetilde{d}^i_k = 0. \tag{2.15}$$

When $k = n-1$ we have $\bar{p}^0 \circ q^{n-1}_{[n]} = q^n_{[n]} \circ \widetilde{d}^1_{n-1}$, in accordance with (2.7).

We note also that if the given $W^{\bullet}_{[n-1]}$ is equipped with a coaugmentation $\varepsilon_{[n-1]}\colon Y \to W^{\bullet}_{[n-1]}$, we also have a coaugmented version $\varepsilon_{[n]}\colon Y \to \widetilde{W}^{\bullet}_{[n]}$ for the restricted cosimplicial object $\widetilde{W}^{\bullet}_{[n]}$, which is defined in the same way by setting $\widetilde{W}^{-1}_{[n-1]} := Y$.

(iii) Making $\widetilde{W}^{\bullet}_{[n]}$ into a full cosimplicial object $W^{\bullet}_{[n]}$ Given the restricted cosimplicial object $\widetilde{W}^{\bullet}_{[n]}$ obtained in step (ii), we first define a full cosimplicial object $\widehat{W}^{\bullet}_{[n]}$ by setting

$$\overline{G}^k := \begin{cases} \overline{P\Omega^{n-k-1}W^n} & \text{if } k \le n, \\ * & \text{if } k > n \end{cases} \tag{2.16}$$

(using (2.7)). We then let

$$\begin{aligned} \widehat{W}^r_{[n]} &:= \widetilde{W}^r_{[n]} \times \prod_{0<k<r} \prod_{0\le i_1<\cdots<i_k\le r-1} \overline{G}^{r-k} \\ &= W^r_{[n-1]} \times \prod_{0\le k\le r} \prod_{0\le i_1<\cdots<i_k\le r} \overline{G}^{r-k} \end{aligned} \tag{2.17}$$

be the construction denoted by $\mathrm{csk}^n G^{\bullet}$ in (1.17). See Section 1.2 and (2.10) (and compare (1.13) in the dual case).

The codegeneracy map $s^t\colon \widehat{W}^{r+1}_{[n]} \to \widehat{W}^r_{[n]}$ is defined into the factor $\overline{G}^{r-k}$ of $\widehat{W}^r_{[n]}$ indexed by the k–tuple $I = (i_1, \dots, i_k)$ by projecting $\widehat{W}^{r+1}_{[n]}$ onto the factor $\overline{G}^{r-k}$ indexed by the unique $(k+1)$–tuple $J = (j_1, \dots, j_{k+1})$ satisfying the cosimplicial identity $s^I \circ s^t = s^J$. The coface maps of $\widehat{W}^{\bullet}_{[n]}$ are determined by those of $\widetilde{W}^{\bullet}_{[n]}$ and the cosimplicial identities, and we have a natural map of restricted cosimplicial objects $g\colon \mathcal{U}\widehat{W}^{\bullet}_{[n]} \to \widetilde{W}^{\bullet}_{[n]}$, which is a dimensionwise trivial fibration.

2.18 Remark $\widehat{W}^{\bullet}_{[n]}$ is obviously n–coskeletal. Moreover, it is Reedy fibrant, since the natural map

$$\hat{\zeta}^r\colon \widehat{W}^r_{[n]} \to M^r \widehat{W}^{\bullet}_{[n]}$$

(cf Section 1.2) is just the product of $\zeta^r\colon W^r_{[n-1]} \to M^r W^{\bullet}_{[n-1]}$ (which is a fibration, since $W^{\bullet}_{[n-1]}$ is Reedy fibrant) with the projection onto the appropriate factors in (2.17) (which is a fibration since all objects $\overline{G}^k$ are fibrant). Moreover, the composites of

$$\widehat{W}^k_{[n]} \xrightarrow{g^k} \twoheadrightarrow \widetilde{W}^k_{[n]} \xrightarrow{\psi^k_{[n]}} \twoheadrightarrow W^k_{[n-1]}$$

for each k fit together to define a map of cosimplicial objects ${}'\pi_{[n]}\colon \widehat{W}^{\bullet}_{[n]} \twoheadrightarrow W^{\bullet}_{[n-1]}$, which is a Reedy fibration (for the same reason).

Finally, we let $h_{[n]}\colon W^{\bullet}_{[n]} \to \widehat{W}^{\bullet}_{[n]}$ be a (functorial) Reedy cofibrant replacement (see Hirschhorn [28, Section 15], and compare Bousfield and Kan [20, Section X.4.2]), so

$h_{[n]}$ is a trivial Reedy fibration, and we set $\pi_{[n]}\colon \boldsymbol{W}^{\bullet}_{[n]} \to \boldsymbol{W}^{\bullet}_{[n-1]}$ to be the composite ${}'\pi_{[n]} \circ h_{[n]}$ — again a Reedy fibration. Thus the full cosimplicial object $\boldsymbol{W}^{\bullet}_{[n]}$ is Reedy fibrant and cofibrant, and dimensionwise weakly equivalent to $\widehat{\boldsymbol{W}}^{\bullet}_{[n]}$. Since the latter is n–coskeletal, we see that $\boldsymbol{W}^{\bullet}_{[n]}$ is n–coskeletal "up to homotopy".

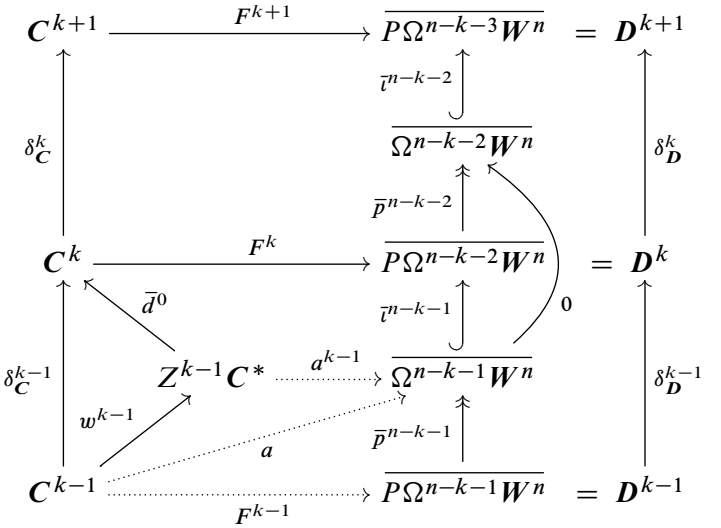
2.19 Lemma *Let $(\boldsymbol{C}^*, \delta_{\boldsymbol{C}})$ be a cochain complex and $(\boldsymbol{D}^*, \delta_{\boldsymbol{D}})$ a left Reedy fibrant $(n-1)$–truncated cochain complex as in Section 2.5(i) in a pointed model category $\mathcal{C}$, and let $Z^j\boldsymbol{C}^* := \mathrm{Coker}(\delta^{j-1}_{\boldsymbol{C}})$ with structure map $w^j\colon \boldsymbol{C}^j \to Z^j\boldsymbol{C}^*$ as in Definition 1.5. Any cochain map $F^j\colon \boldsymbol{C}^j \to \boldsymbol{D}^j$ defined for $k \le j < n$ induces a unique map $a^{k-1}\colon Z^{k-1}\boldsymbol{C}^* \to \overline{\Omega^{n-k-1}\boldsymbol{W}^n}$ with*

$$\bar{\imath}^{n-k-1} \circ a^{k-1} \circ w^{k-1} = F^k \circ \delta^{k-1}_{\boldsymbol{C}}, \tag{2.20}$$

in the notation of (2.6) *and* (2.8).

Proof By assumption, we have the following solid commuting diagram:

(2.21)



Then

$$\bar{\imath}^{n-k-2} \circ \bar{p}^{n-k-2} \circ F^k \circ \delta^{k-1}_{\boldsymbol{C}} = \delta^k_{\boldsymbol{D}} \circ F^k \circ \delta^{k-1}_{\boldsymbol{C}} = F^{k+1} \circ \delta^k_{\boldsymbol{C}} \circ \delta^{k-1}_{\boldsymbol{C}} = 0.$$

Since $\bar{\imath}^{n-k-2}$ is a monomorphism, in fact $\bar{p}^{n-k-2} \circ F^k \circ \delta^{k-1}_{\boldsymbol{C}} = 0$, Therefore, since (2.6) is a fibration sequence, $F^k \circ \delta^{k-1}_{\boldsymbol{C}}$ factors through $a\colon \boldsymbol{C}^{k-1} \to \overline{\Omega^{n-k-1}\boldsymbol{W}^n}$ as indicated in (2.21). Moreover, since $\bar{\imath}^{n-k-1} \circ a \circ \delta^{k-2}_{\boldsymbol{C}} = F^k \circ \delta^{k-1}_{\boldsymbol{C}} \circ \delta^{k-2}_{\boldsymbol{C}} = 0$ and $\bar{\imath}^{n-k-1}$ is a monomorphism, too, the map a factors through $a^{k-1}\colon Z^{k-1}\boldsymbol{C}^* \to \overline{\Omega^{n-k-1}\boldsymbol{W}^n}$ as in (2.21), satisfying (2.20). □

We next note the following technical fact about Moore chain objects:

2.22 Lemma *Let $\boldsymbol{W}^\bullet \in \mathcal{C}^{\boldsymbol{\Delta}}$ be a Reedy cofibrant cosimplicial object over a model category $\mathcal{C}$ as in Assumption 1.20, and $\boldsymbol{B}$ a homotopy group object in $\mathcal{C}$. Then for any Moore chain $\alpha \in C_n[\boldsymbol{W}^\bullet, \boldsymbol{B}]$ for the simplicial group $[\boldsymbol{W}^\bullet, \boldsymbol{B}]$:*

(a) *α can be realized by a map $a\colon \boldsymbol{W}^n \to \boldsymbol{B}$ with $a \circ d^i_{n-1} = 0$ for all $1 \le i \le n$, and thus induces a map $\overline{a}\colon C^n\boldsymbol{W}^\bullet \to \boldsymbol{B}$ with $\overline{a} \circ v^n = a$.*

(b) *If α is a Moore cycle, we can choose a nullhomotopy $H\colon \boldsymbol{W}^{n-1} \to P\boldsymbol{B} \subseteq \boldsymbol{B}^{[0,1]}$ for $a \circ d^0_{n-1}$ such that $H \circ d^j_{n-2} = 0$ for $1 \le j \le n-1$, and thus induces a map $\overline{H}\colon C^n\boldsymbol{W}^\bullet \to P\boldsymbol{B}$ with $\overline{H} \circ v^n = H$.*

Proof Since $\boldsymbol{W}^\bullet$ is Reedy cofibrant, the simplicial space $\boldsymbol{U}_\bullet = \mathrm{map}_*(\boldsymbol{W}^\bullet, \boldsymbol{B}) \in s\mathcal{S}_*$ is Reedy fibrant, so we have an isomorphism

$$\iota_\star\colon \pi_i C_n \boldsymbol{U}_\bullet \to C_n \pi_i \boldsymbol{U}_\bullet \quad \text{for all } i \ge 0 \tag{2.23}$$

(see [20, Proposition X.6.3]). Thus we can represent $\alpha \in C_n\pi_0\boldsymbol{U}_\bullet$ by a map $a \in C_n\boldsymbol{U}_\bullet$, which implies (i).

If α is a cycle, then $\partial_n(\alpha) = [a \circ d^0_{n-1}]$ vanishes in $\pi_0 C_{n-1}\boldsymbol{U}_\bullet$, so we have a nullhomotopy H for $a \circ d^0_{n-1}$ in

$$PC_{n-1}\,\mathrm{map}_*(\boldsymbol{W}^\bullet, \boldsymbol{B}) = C_{n-1}\,\mathrm{map}_*(\boldsymbol{W}^\bullet, P\boldsymbol{B}) \subseteq \mathrm{map}_*(\boldsymbol{W}^{n-1}, P\boldsymbol{B}),$$

which implies (ii). □

From the description in Section 2.5 we can actually deduce:

2.24 Proposition *Any $\boldsymbol{W}^\bullet_{[n]}$ obtained from $\boldsymbol{W}^\bullet_{[n-1]}$ as in Definition 2.1(c) will satisfy Definition 2.1(a)–(b), and the limit $\boldsymbol{W}^\bullet$ of (2.2) is a Reedy fibrant cosimplicial object over $\mathcal{C}$ which realizes the given algebraic resolution $V_\bullet \to H^*_\Theta \boldsymbol{Y}$.*

Proof In the setting of Definition 2.1, with $\boldsymbol{C}^* = C^*\boldsymbol{W}^\bullet_{[n-1]}$ for Reedy cofibrant $\boldsymbol{W}^\bullet_{[n-1]}$, we can use Lemma 2.22(a) to represent the attaching map $\overline{\partial}^n_0\colon \overline{V}_n \to Z_{n-1}V_\bullet$ for the CW resolution $V_\bullet$ by a map $F^{n-1}\colon C^{n-1}\boldsymbol{W}^\bullet_{[n-1]} \to \overline{\boldsymbol{W}}^n$ (see Step IV in the proof of Theorem A.11 below). From (2.16) and (2.17), we then see that $\boldsymbol{W}^\bullet_{[n]}$ realizes $V_\bullet$ through simplicial dimension n.

Moreover, $\boldsymbol{W}^\bullet$ is as stated because the maps $\pi_{[n]}$ restrict to trivial fibrations

$$\pi^k_{[n]}\colon \boldsymbol{W}^k_{[n]} \to \boldsymbol{W}^k_{[n-1]}$$

for each $0 \le k < n$, so $\boldsymbol{W}^k = \mathrm{holim}_n \boldsymbol{W}^k_{[n]}$ in $\mathcal{C}$. □

2.25 Higher cohomology operations When F^{k-1} also exists that makes (2.21) commute, we have

$$a^{k-1} \circ w^{k-1} = \overline{p}^{n-k-1} \circ F^{k-1}, \tag{2.26}$$

so F^{k-1} is a nullhomotopy for $a^{k-1} \circ w^{k-1}$ in the sense of Quillen [32, Section I.2]. Note that Lemma 2.19 also makes sense for $k = n-1$, where $F^n = 0$.

If we choose F^{n-1} as in the proof of Proposition 2.24, the map $F^{n-1} \circ \delta^{n-2}$, which induces $a^{n-2} \circ w^{n-2}$, is nullhomotopic, with nullhomotopy F^{n-2}. We can then think of $a^{n-3} \circ w^{n-3}\colon C^{n-2} \to \overline{\Omega^1 \boldsymbol{W}^n}$ as the value of the secondary cohomology operation corresponding to the diagram

$$\begin{tikzcd} \boldsymbol{C}^{n-3} \arrow[r, "\delta^{n-3}"] \arrow[rr, bend left, "0"] & \boldsymbol{C}^{n-2} \arrow[r, "\delta^{n-2}"] \arrow[rr, bend right, "0"'] & \boldsymbol{C}^{n-1} \arrow[r, "F^{n-1}"] & \overline{\boldsymbol{W}}^n \end{tikzcd} \tag{2.27}$$

as in (0.2). Only $\overline{\boldsymbol{W}}^n$ is an R–GEM, but this suffices to let us think of each value $a^{n-3} \circ w^{n-3}$ of this Toda bracket in $[\boldsymbol{C}^{n-3}, \Omega\overline{\boldsymbol{W}}^n]$ as a collection of cohomology classes for $\boldsymbol{C}^{n-3}$. This nevertheless qualifies as a higher cohomology operation as described in the introduction, if we use a truncation of $\boldsymbol{\Delta}_+$ as our indexing category I.

By what we say above, if $a^{n-3} \circ w^{n-3} \sim *$ — that is, the secondary operation corresponding to (2.27) vanishes — then the choice of a nullhomotopy F^{n-3} yields a value $a^{n-4} \circ w^{n-4}\colon \boldsymbol{C}^{n-4} \to \overline{\Omega^2 \boldsymbol{W}^n}$ for the corresponding third order operation, and so on. This observation is the key to what we are doing in this paper.

2.28 Definition Let Θ be an algebraic sketch, as in Remark 1.25, associated to $\mathcal{A} \subseteq \operatorname{Obj} \mathcal{C}$, so that by definition any $\boldsymbol{B} \in \Theta$ is of the form $\boldsymbol{B} := \prod_{\boldsymbol{A} \in \mathcal{A}} \prod_{i \in I_{\boldsymbol{A}}} \boldsymbol{A}$ with each $I_{\boldsymbol{A}}$ a finite indexing set (see (1.26)).

We then say that Θ is *allowable* if the natural map

$$\coprod_{\boldsymbol{A} \in \mathcal{A}} \coprod_{i \in I_{\boldsymbol{A}}} H^*_\Theta \boldsymbol{A} \to H^*_\Theta \boldsymbol{B} \tag{2.29}$$

is an isomorphism for any such $\boldsymbol{B} \in \Theta$.

2.30 Remark If we write $I := \coprod_{\boldsymbol{A} \in \mathcal{A}} I_{\boldsymbol{A}}$ and denote the copy of $\boldsymbol{A}$ indexed by $i \in I_{\boldsymbol{A}}$ by $\boldsymbol{B}_i$, we have $\boldsymbol{B} = \prod_{i \in I} \boldsymbol{B}_i$. For any Θ–algebra Γ we then have, by

Lemma 1.28 and (2.29),

$$\textbf{(2.31)}\qquad \prod_{i\in I}\Gamma\{\boldsymbol{B}_i\} = \prod_{i\in I}\mathrm{Hom}_{\Theta\text{–}\mathsf{Alg}}(H^*_\Theta \boldsymbol{B}_i, \Gamma)$$
$$= \mathrm{Hom}_{\Theta\text{–}\mathsf{Alg}}\Big(\coprod_{i\in I} H^*_\Theta \boldsymbol{B}_i, \Gamma\Big) = \mathrm{Hom}_{\Theta\text{–}\mathsf{Alg}}(H^*_\Theta \boldsymbol{B}, \Gamma).$$

2.32 Lemma *If R is $\mathbb{F}_p$ or a field of characteristic 0 and λ is any limit cardinal, the FP–sketch Θ^λ_R of Example 1.27 is allowable.*

Proof Every $\boldsymbol{B}\in\Theta^\lambda_R$ has the form $\boldsymbol{B}=\prod_{n=1}^\infty\prod_{I_n}\boldsymbol{K}(V_i,n)$ for $V_i\in R\text{–}\mathsf{Mod}^\lambda$ and finite indexing sets I_n. Then $H^*(\boldsymbol{B};R)=\coprod_{n=1}^\infty\coprod_{I_n}H^*(\boldsymbol{K}(V_i,n);R)$ (see Blanc and Sen [17, Lemma 4.17]). □

We are now in a position to state our first important technical result:

2.33 Theorem *Let Θ be an allowable algebraic sketch in $\mathcal{C}$ and let $V_\bullet$ be any CW resolution of the realizable Θ–algebra $H^*_\Theta\boldsymbol{Y}$. Then there is a cofibrant sequential realization $\mathcal{W}=\langle \boldsymbol{W}^\bullet_{[n]}, \widetilde{\boldsymbol{W}}^\bullet_{[n]}\rangle_{n\in\mathbb{N}}$ of $V_\bullet$ for $\boldsymbol{Y}$, with each $\boldsymbol{W}^k_{[n]}$ in Θ.*

We defer the proof to the appendix, where we actually prove a more general result (which is needed elsewhere).

2.34 Remark If we want to use the allowability of Θ^λ_R in Lemma 2.32 for Theorem 2.33, the choice of the cardinal λ may depend on the size of graded R–vector space $H^*(\boldsymbol{Y};R)$ (see Blanc and Sen [17, Section 3]). However, in the most commonly encountered case, $H^*(\boldsymbol{Y};R)$ will be of finite type, and we may choose the CW resolution $V_\bullet$ to be finite type in each simplicial dimension, too. In this case we can make do with the original $\Theta_R=\Theta^\omega_R$ of Example 1.27.

3 Comparing cosimplicial resolutions

Cosimplicial resolutions of the type constructed in Section 2 play a central role in our theory of higher cohomology operations, but they depend on many particular choices. In this section we shall show how any two such cosimplicial objects are related by a zigzag of maps of a particularly simple form.

Although many of the results hold more generally, from now on we restrict attention to the algebraic sketch Θ_R for R either $\mathbb{F}_p$ or a field of characteristic 0 (see Example 1.27),

with $\mathcal{C} = \mathbb{S}_*$. This allows us to assume for convenience that all the objects in each stage of our sequential realizations are simplicial R–modules (though the maps between them need not be strict simplicial homomorphisms).

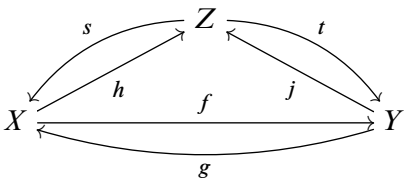
We shall also assume from here on that all spaces are connected. The modifications needed for the nonconnected case should be clear.

We first note the following general facts about model categories:

3.1 Lemma *Let X and Y be two weakly equivalent fibrant and cofibrant objects in a simplicial model category $\mathcal{C}$.*

(a) *There is a diagram of weak equivalences*

(3.2)

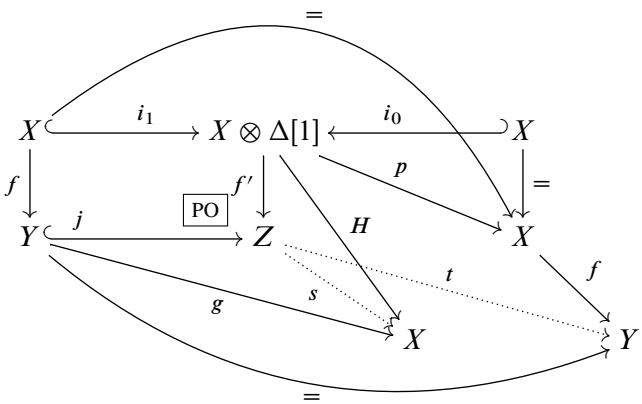


where $h := f' \circ i_0$, $s \circ j = g$, $t \circ h = f$, $s \circ h = \mathrm{Id}_X$, $t \circ j = \mathrm{Id}_Y$ and the notation f', i_0 is as in the proof.

(b) *There are maps $X \amalg Y \xrightarrow{F} Z \xrightarrow{G} X \times Y$, with F a cofibration which is a trivial cofibration on each summand, G a fibration which is a trivial fibration onto each factor, and the induced maps $X \to X$ and $Y \to Y$ are identities.*

Proof (a) By Quillen [32, Section I.1]) we have homotopy equivalences $f\colon X \to Y$ and $g\colon Y \to X$ with a homotopy $H\colon g \circ f \sim \mathrm{Id}_X$ fitting into a commutative diagram

(3.3)



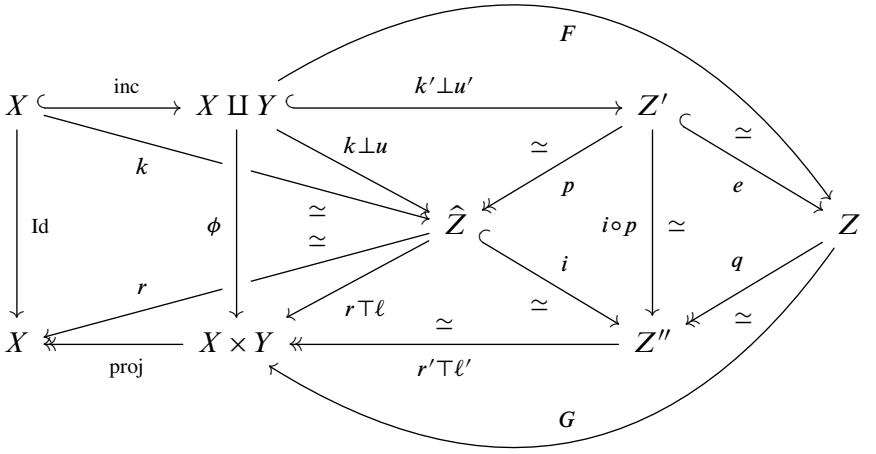
with all maps weak equivalences, where Z is the pushout, the maps i_0 and i_1 are induced by the inclusions $\Delta[0] \hookrightarrow \Delta[1]$ and p is induced by $\Delta[1] \twoheadrightarrow \Delta[0]$.

(b) Choose a weak equivalence $f\colon X \to Y$, and factor it as $X \xrightarrow{k} \widehat{Z} \xrightarrow{\ell} Y$, with k a trivial cofibration and ℓ a trivial fibration. By the LLP and fibrancy of X we have a retraction $r\colon \widehat{Z} \to X$ for k, and by RLP and cofibrancy of Y we have a section $u\colon Y \to \widehat{Z}$ for ℓ, both weak equivalences. Set

$$\phi := (\mathrm{Id}_X \top (r \circ u)) \perp (f \top \mathrm{Id}_Y)\colon X \amalg Y \to X \times Y.$$

Factor $k \perp u\colon X \amalg Y \to \widehat{Z}$ as $X \amalg Y \xrightarrow{k' \perp u'} Z' \xrightarrow{p} \widehat{Z}$ (a cofibration followed by a trivial fibration), and $r \top \ell\colon \widehat{Z} \to X \times Y$ as $\widehat{Z} \xrightarrow{i} Z'' \xrightarrow{r' \top \ell'} X \times Y$ (a trivial cofibration followed by a fibration). Finally, factor $i \circ p\colon Z' \xrightarrow{\simeq} Z''$ as $Z' \xrightarrow{e} Z \xrightarrow{q} Z'$ (a trivial cofibration followed by a trivial fibration):

(3.4)



Then $F := e \circ (k' \perp u')$ is a cofibration, $G := (r' \top \ell') \circ q$ is a fibration, and the claim follows by tracking the weak equivalences in (3.4) (and similarly for Y). □

3.5 Definition Given two CW resolutions $\varepsilon\colon V_\bullet \to \Gamma$ and ${}'\varepsilon\colon {}'V_\bullet \to \Gamma$ of a Θ_R–algebra Γ, with CW bases $(\overline{V}_n)_{n\in\mathbb{N}}$ and $({}'\overline{V}_n)_{n\in\mathbb{N}}$, an *algebraic comparison map* $\Psi\colon V_\bullet \to {}'V_\bullet$ is a system

$$\Psi = \langle \varphi, \rho, (\overline{\varphi}_n, \overline{\rho}_n)_{n\in\mathbb{N}} \rangle, \tag{3.6}$$

where $\varphi\colon V_\bullet \to {}'V_\bullet$ is a split monic weak equivalence of simplicial Θ_R–algebras with retraction $\rho\colon {}'V_\bullet \to V_\bullet$ (with ${}'\varepsilon \circ \overline{\varphi}_0 = \varepsilon$), induced by inclusions of coproduct summands $\overline{\varphi}_n\colon \overline{V}_n \hookrightarrow {}'\overline{V}_n$ with retractions $\overline{\rho}_n$ for each $n \geq 0$.

3.7 Lemma *Any two CW resolutions $\varepsilon^{(0)}\colon V^{(0)}_\bullet \to \Gamma$ and $\varepsilon^{(1)}\colon V^{(1)}_\bullet \to \Gamma$ of the same Θ_R–algebra Γ have a common "algebraic h–cobordism" CW resolution $\varepsilon\colon V_\bullet \to \Gamma$, with algebraic comparison maps $\Psi^{(i)}\colon V^{(i)}_\bullet \to V_\bullet$ for $i = 0, 1$.*

Proof Let $(\overline{V}_n^{(i)})_{n\in\mathbb{N}}$ be CW bases for $V_\bullet^{(i)}$ for $i=0,1$. Since $X=V_\bullet^{(0)}$ and $Y=V_\bullet^{(1)}$ are fibrant and cofibrant in $s\Theta_R$–Alg (cf Proposition 1.29), they have homotopy equivalences $f\colon X\to Y$ and $g\colon Y\to X$ as in Lemma 3.1(a). We make explicit the construction of the lemma by producing a CW basis $(\overline{V}_n)_{n\in\mathbb{N}}$ for $Z=V_\bullet$ in (3.3), together with inclusions of coproduct summands $\overline{\varphi}^n_{(i)}\colon \overline{V}_n^{(i)}\hookrightarrow \overline{V}_n$ for each $n\geq 0$ as in Definition 3.5.

If we write $e_0, e_1\in\Delta[1]_0$ and $\sigma\in\Delta[1]_1$ for the nondegenerate simplices, we have $V_n={}'\overline{V}_n \amalg L_nV_\bullet$, with

$$\begin{aligned}\textbf{(3.8)}\quad {}'\overline{V}_n := \overline{V}_n^{(0)}\otimes(\boldsymbol{s}^{n-1}e_0)\amalg\coprod_{k=0}^{n-1}[\overline{V}_n^{(0)}\otimes(\boldsymbol{s}^{n-1}_{\widehat{k}}\sigma)]\amalg[\overline{V}_n^{(1)}\otimes(\boldsymbol{s}^{n-1}e_1)]\\ \amalg\coprod_{k=0}^{n-1}[s_k\overline{V}_{n-1}^{(0)}\otimes(\boldsymbol{s}^{n-1}_{\widehat{k}}\sigma)].\end{aligned}$$

Here $\boldsymbol{s}^n$ is the iterated degeneracy map $s_n\dots s_0$ and $\boldsymbol{s}^n_{\widehat{k}}:=s_n\dots s_{k+1}\widehat{s}_k s_{k-1}\dots s_0$.

The face maps are calculated as usual on each factor of $a\otimes b$, except that

$$\textbf{(3.9)}\qquad d_n(s_{n-1}u\otimes(\boldsymbol{s}^{n-2}\sigma))=f_{n-1}u\otimes(\boldsymbol{s}^{n-2}e_1)\in[\overline{V}_{n-1}^{(1)}\otimes(\boldsymbol{s}^{n-2}e_1)]$$

for $k=n-1$ in the second line of (3.8), by (3.3), where $f\colon V_\bullet^{(0)}\to V_\bullet^{(1)}$ is the chosen homotopy equivalence.

Note that ${}'\overline{V}_n$ is not a CW basis object for $V_\bullet$, since the summands in the second line of (3.8) always have at least two nonvanishing face maps.

However, for any $v\in{}'\overline{V}_n$ we can define $v^{[0]}:=v$ and $v^{[k+1]}:=v^{[k]}-s_{n-k-1}d_{n-k}v^{[k]}$ by induction on k, and find that $d_iv^{[k]}=0$ for $n-k<i\leq n$, so $v^{[n]}$ is a Moore chain. Note that for $v\in U$ in summands U on the first line of (3.8) we have simply $v=v^{[n]}$.

Explicitly, we replace each generator $v=s_ku\otimes(\boldsymbol{s}^{n-1}_{\widehat{k}}\sigma)$ of a summand in the second line of (3.8) for ${}'\overline{V}_n$ by

$$\textbf{(3.10)}\quad v^{[n]}=\begin{cases}\sum_{i=0}^{k}(-1)^i[s_{k-i}u]\otimes[(\boldsymbol{s}^{n-1}_{\widehat{k}}\sigma)-(\boldsymbol{s}^{n-1}_{k\widehat{+1}}\sigma)] & \text{if } k<n-1,\\ \sum_{i=1}^{n}(-1)^i[s_{n-i}f_{n-1}u\otimes(\boldsymbol{s}^{n-2}e_1)-s_{n-i}u\otimes(\boldsymbol{s}^{n-2}\sigma)] & \text{if } k=n-1.\end{cases}$$

Thus $v^{[n]}$ always has the form $v+\sum_{i=1}^k u_i$, with the elements u_i all degenerate.

The Θ_R–algebra retraction $\rho'\colon V_n\to{}'\overline{V}_n$ onto the summand ${}'\overline{V}_n$ therefore takes $v^{[n]}$ to v. This implies that if $\{v_i\}_{i\in I}$ are generators of ${}'\overline{V}_n$, the new elements $\{v_i^{[n]}\}_{i\in I}$

still generate a free sub-Θ_R–algebra of V_n, because any relation of the form

$$\psi(v_1^{[n]},\dots,v_k^{[n]})=0$$

in V_n, where ψ is some primary R–cohomology operation, implies that also

$$0=\rho'(\psi(v_1^{[n]},\dots,v_k^{[n]}))=\psi(\rho'(v_1^{[n]}),\dots,\rho'(v_k^{[n]}))=\psi(v_1,\dots,v_k),$$

which can hold only if $\psi\equiv 0$ since the elements v_i are generators of a free Θ_R–algebra.

Therefore, if we write $\overline{V}_n$ for the sub-Θ_R–algebra of V_n generated by all elements $v^{[n]}$ as v varies over a set of generators for (each summand of) ${}'\overline{V}_n$, we still have a coproduct of free Θ_R–algebras $V_n=\overline{V}_n\amalg L_nV_\bullet$, where $\overline{V}_n$ now serves as an nth CW basis object for $V_\bullet$.

Moreover, we still have inclusions of coproduct summands $\overline{\varphi}^n_{(i)}\colon \overline{V}^{(i)}_n\hookrightarrow\overline{V}_n$ inducing split trivial cofibrations of simplicial Θ_R–algebras $\varphi^{(i)}\colon V^{(i)}_\bullet\to V_\bullet$ for $i=0,1$. We can use these to further write $\overline{V}_n:=\overline{U}_n\amalg\overline{V}^{(0)}_n\amalg\overline{V}^{(1)}_n$. □

3.11 Definition Given an algebraic comparison map $\Psi=\langle\varphi,\rho,(\overline{\varphi}_n,\overline{\rho}_n)_{n\in\mathbb{N}}\rangle$ between two CW resolutions $\varepsilon\colon V_\bullet\to\Gamma$ and ${}'\varepsilon\colon {}'V_\bullet\to\Gamma$ of a realizable Θ_R–algebra Γ (see Definition 3.5), and sequential realizations $\mathcal{W}$ and ${}'\mathcal{W}$ of $V_\bullet$ and ${}'V_\bullet$, respectively, a *comparison map* $\Phi\colon\mathcal{W}\to{}'\mathcal{W}$ *over* Ψ is a system

$$\Phi=\langle e_{[n]},r_{[n]},(\overline{Pe}^k_n)_{k=0}^{n-1},(\bar{e}^k_n)_{k=0}^{n-1},(\overline{Pr}^k_n)_{k=0}^{n-1},(\bar{r}^k_n)_{k=0}^{n-1}\rangle_{n\in\mathbb{N}} \tag{3.12}$$

consisting of:

(i) Split fibrations of the modified path-loop fibrations of (2.6) fitting into a diagram

$$\begin{array}{ccccc}
\overline{\Omega^{k+1}{}'W^n} & \xhookrightarrow{{}'\bar{\imath}^{k+1}} & \overline{P\Omega^k{}'W^n} & \xrightarrow{{}'\bar{p}^k} & \overline{\Omega^k{}'W^n} \\
{\scriptstyle\bar{r}^{k+1}_n}\uparrow\downarrow{\scriptstyle\bar{e}^{k+1}_n} & & {\scriptstyle\overline{Pr}^k_n}\uparrow\downarrow{\scriptstyle\overline{Pe}^k_n} & & {\scriptstyle\bar{r}^k_n}\uparrow\downarrow{\scriptstyle\bar{e}^k_n} \\
\overline{\Omega^{k+1}W^n} & \xhookrightarrow{\bar{\imath}^{k+1}} & \overline{P\Omega^kW^n} & \xrightarrow{\bar{p}^k} & \overline{\Omega^kW^n}
\end{array} \tag{3.13}$$

for each $0\le k<n$, in which both upward and downward squares commute, as well as

$$\overline{Pe}^k_n\circ\overline{Pr}^k_n=\mathrm{Id}\quad\text{and}\quad\bar{e}^k_n\circ\bar{r}^k_n=\mathrm{Id}\quad\text{for all }0\le k<n. \tag{3.14}$$

We require that for all $0\le k<n$, the maps $\bar{e}^k_n$ realize $\Omega^k\overline{\varphi}_n$ and the maps $\bar{r}^k_n$ realize $\Omega^k\overline{\rho}_n$.

(ii) A cosimplicial map $e_{[n]}\colon {}'\boldsymbol{W}^{\bullet}_{[n]} \to \boldsymbol{W}^{\bullet}_{[n]}$ realizing $\varphi\colon V_\bullet \to {}'V_\bullet$ through simplicial dimension n, with section $r_{[n]}\colon \boldsymbol{W}^{\bullet}_{[n]} \to {}'\boldsymbol{W}^{\bullet}_{[n]}$ realizing ρ, such that for each $0 \le k < n$, both squares in the following diagram commute:

$$\begin{array}{ccc} C^k({}'\boldsymbol{W}^{\bullet}_{[n-1]}) & \xrightarrow{C^k(e^k_{[n-1]})} & C^k(\boldsymbol{W}^{\bullet}_{[n-1]}) \\ {\scriptstyle {}'F^k}\downarrow & & \downarrow{\scriptstyle F^k} \\ \overline{P\Omega^{n-k-2}{}'\boldsymbol{W}^n} & \xrightarrow[\overline{P}\bar{e}^{n-k-2}_n]{} & \overline{P\Omega^{n-k-2}\boldsymbol{W}^n} \end{array} \qquad \begin{array}{ccc} C^k({}'\boldsymbol{W}^{\bullet}_{[n-1]}) & \xleftarrow{C^k(r^k_{[n-1]})} & C^k(\boldsymbol{W}^{\bullet}_{[n-1]}) \\ {\scriptstyle {}'F^k}\downarrow & & \downarrow{\scriptstyle F^k} \\ \overline{P\Omega^{n-k-2}{}'\boldsymbol{W}^n} & \xleftarrow[\overline{P}\bar{r}^{n-k-2}_n]{} & \overline{P\Omega^{n-k-2}\boldsymbol{W}^n} \end{array} \tag{3.15}$$

When each map $\bar{e}^k_n\colon \overline{\Omega^k{}'\boldsymbol{W}^n} \twoheadrightarrow \overline{\Omega^k\boldsymbol{W}^n}$ and $e^k_{[n]}\colon {}'\boldsymbol{W}^k_{[n]} \twoheadrightarrow \boldsymbol{W}^k_{[n]}$ is a trivial fibration (and thus each map $\bar{r}^k_n\colon \overline{\Omega^k\boldsymbol{W}^n} \hookrightarrow \overline{\Omega^k{}'\boldsymbol{W}^n}$ and $r^k_{[n]}\colon {}'\boldsymbol{W}^k_{[n]} \hookrightarrow \boldsymbol{W}^k_{[n]}$ is a trivial cofibration in $\mathcal{S}_*$), we say that Φ is a *trivial* comparison map.

If we only have

$$\Phi = \big\langle e_{[n]}, r_{[n]}, (\overline{P}e^k_n)^{n-1}_{k=0}, (\bar{e}^k_n)^{n-1}_{k=0}, (\overline{P}r^k_n)^{n-1}_{k=0}, (\bar{r}^k_n)^{n-1}_{k=0}\big\rangle^N_{n=0} \tag{3.16}$$

as above, we say that $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$ is an *N–stage comparison map* over Ψ.

3.17 Remark If we let $\bar{\jmath}^k_{[n]}\colon \overline{\Omega^k\boldsymbol{X}^n} \hookrightarrow \overline{\Omega^k{}'\boldsymbol{W}^n}$ denote the inclusion of the fiber of $\bar{e}^k_n\colon \overline{\Omega^k{}'\boldsymbol{W}^n} \twoheadrightarrow \overline{\Omega^k\boldsymbol{W}^n}$, we see that the splitting $\bar{r}^k_n\colon \overline{\Omega^k\boldsymbol{W}^n} \hookrightarrow \overline{\Omega^k{}'\boldsymbol{W}^n}$ induces a retraction $\bar{s}^k_{[n]}\colon \overline{\Omega^k{}'\boldsymbol{W}^n} \twoheadrightarrow \overline{\Omega^k\boldsymbol{X}^n}$ for $\bar{\jmath}^k_{[n]}$, defined $x \mapsto x - \bar{r}^k_n\bar{e}^k_n(x)$, and thus a map

$$\overline{\Omega^k{}'\boldsymbol{W}^n} \xrightarrow[\simeq]{\bar{s}^k_{[n]} \top \bar{e}^k_n} \overline{\Omega^k\boldsymbol{X}^n} \times \overline{\Omega^k\boldsymbol{W}^n}, \tag{3.18}$$

which is a weak equivalence for each $0 \le k < n$ (using the abelian group structure on all spaces). As we shall see, in many cases we can assume (3.18) is actually an equality.

3.19 Definition A *zigzag of comparison maps* between two sequential realizations $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ of a realizable $\Theta_{\boldsymbol{R}}$–algebra Γ is a (possibly infinite) sequence of cospans of comparison maps starting from $\mathcal{W}^{(0)}$ and ending at $\mathcal{W}^{(1)}$, which is *locally finite* in the sense that for each $n \ge 0$, only finitely many of the comparison maps in the zigzag between the nth stages $(\boldsymbol{W}^{\bullet}_{[n]})^{(0)}$ and $(\boldsymbol{W}^{\bullet}_{[n]})^{(1)}$ are not the identity map.

We say that two abstract sequential realizations $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ (of arbitrary spaces $\boldsymbol{Y}^{(0)}$ and $\boldsymbol{Y}^{(1)}$) are *weakly equivalent* if they are related by a zigzag of comparison maps.

Similarly, if only the nth stages $(\boldsymbol{W}^{\bullet}_{[n]})^{(0)}$ and $(\boldsymbol{W}^{\bullet}_{[n]})^{(1)}$ of two such sequential realizations $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ are related by a zigzag of comparison maps, we say that $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ are *n–equivalent*, or that $(\boldsymbol{W}^{\bullet}_{[n]})^{(0)}$ and $(\boldsymbol{W}^{\bullet}_{[n]})^{(1)}$ are *weakly equivalent*.

3.20 Theorem *Any two cofibrant sequential realizations $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ of $\boldsymbol{Y}$ are weakly equivalent.*

Proof We prove the theorem in two main steps:

(i) Different algebraic resolutions We first show that, given an algebraic comparison map $\Psi\colon V_\bullet \to {}'V_\bullet$ for $\boldsymbol{Y}$ and a cofibrant sequential realization $\mathcal{W}$ of $V_\bullet$, there is a cofibrant sequential realization ${}'\mathcal{W}$ of ${}'V_\bullet$ with a comparison map $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$ over Ψ, constructed (with the maps $e_{[n]}\colon {}'\boldsymbol{W}^{\bullet}_{[n]} \to \boldsymbol{W}^{\bullet}_{[n]}$ and sections $r_{[n]}\colon \boldsymbol{W}^{\bullet}_{[n]} \to {}'\boldsymbol{W}^{\bullet}_{[n]}$) by induction on $n \geq 0$:

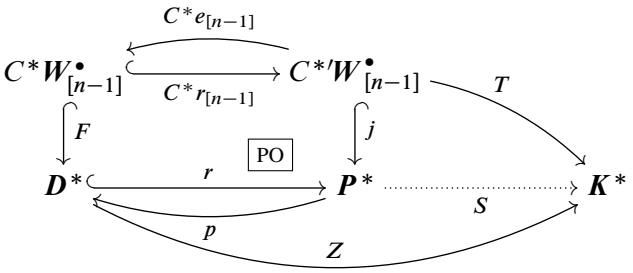
At the nth stage, we may assume by Lemma 3.1 that $e_{[n-1]}$ is a fibration and $r_{[n-1]}$ is a cofibration in the resolution model category $\mathsf{Ch}^{\mathcal{C}}_{\leq n}$, so in particular $e^{j}_{[n-1]}$ is a fibration and $r^{j}_{[n-1]}$ a cofibration for $0 \leq j \leq n-1$ (see Bousfield [19, Section 3.2]).

Since $\overline{V}_n$ is a coproduct summand in ${}'\overline{V}_n = \overline{V}_n \amalg \overline{U}_n$, the map $\overline{\varphi}_n\colon \overline{V}_n \hookrightarrow {}'\overline{V}_n$ is simply the inclusion, while $\overline{\rho}_n\colon {}'\overline{V}_n \hookrightarrow \overline{V}_n$ has the form $\mathrm{Id} \perp \zeta$. If we realize $\overline{V}_n$ by $\overline{\boldsymbol{W}}^n$ and $\overline{U}_n$ by $\overline{\boldsymbol{X}}^n$ then ${}'\overline{V}_n$ is realized by ${}'\widehat{\boldsymbol{W}}^n := \overline{\boldsymbol{X}}^n \times \overline{\boldsymbol{W}}^n$. By Definition 2.1, the nth stage of $\mathcal{W}$ is determined by the choice of left Reedy fibrant replacement $\boldsymbol{D}^*$ of $\overline{\boldsymbol{W}}^n \otimes^* S^{n-1}$, equipped with a left Reedy cofibration $F\colon C^*\boldsymbol{W}^{\bullet}_{[n-1]} \to \boldsymbol{D}^*$ realizing the given attaching map $\overline{\partial}^n_0\colon \overline{V}_n \to C_{n-1}V_\bullet$.

If $\boldsymbol{K}^*$ is similarly a left Reedy fibrant replacement for $\overline{\boldsymbol{X}}^n \otimes^* S^{n-1}$, the attaching map ${}'\overline{\partial}^n_0\colon {}'\overline{V}_n \to C_{n-1}{}'V_\bullet$ has the form $\overline{\partial}^n_0 \perp \tau$, and we may realize $\tau\colon \overline{U}_n \to C_{n-1}{}'V_\bullet$ by $T\colon C^*{}'\boldsymbol{W}^{\bullet}_{[n-1]} \to \boldsymbol{K}^*$ (not a cofibration) and $\zeta\colon \overline{U}_n \to \overline{V}_n$ by $Z\colon \boldsymbol{D}^* \to \boldsymbol{K}^*$.

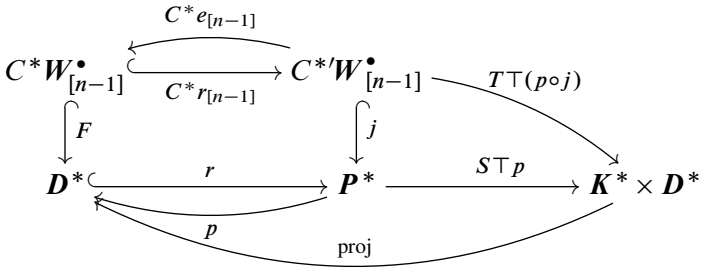
Consider the following diagram in the left Reedy model category of n–truncated cochain complexes over $\mathcal{C}$, in which $\boldsymbol{P}^*$ is the pushout of the upper left square and the map p with section r is induced by $C^*r_{[n-1]}$:

(3.21)



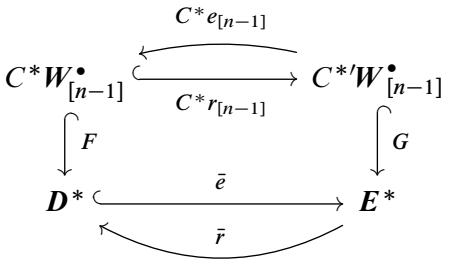
Since, by Definition 3.5, $C_{n-1}\rho\circ{}'\overline{\partial}^n_0=\overline{\partial}^n_0\circ\overline{\rho}_n$, also $C_{n-1}\rho\circ\tau=\overline{\partial}^n_0\circ\overline{\rho}_n|_{\overline{U}_n}=\overline{\partial}^n_0\circ\zeta$, so the outer square in (3.21) commutes up to homotopy. Since F is a cofibration, we may change Z up to homotopy to make it commute on the nose by Blanc, Johnson and Turner [11, Lemma 5.11]. The maps Z and T then induce S as indicated. This allows us to extend (3.21) to a commuting diagram

(3.22)



We now factor $S\top p$ as a cofibration $G'\colon \boldsymbol{P}^*\hookrightarrow\boldsymbol{E}^*$ followed by a trivial fibration $t\colon \boldsymbol{E}^*\twoheadrightarrow\boldsymbol{K}^*\times\boldsymbol{D}^*$ (in the left Reedy model structure on truncated cochain complexes). If we set $G\colon C^*{}'\boldsymbol{W}^\bullet_{[n-1]}\to\boldsymbol{E}^*$ equal to $G'\circ j$, $\bar{r}\colon \boldsymbol{E}^*\to\boldsymbol{D}^*$ equal to $\mathrm{proj}\circ t$, and $\bar{e}\colon \boldsymbol{D}^*\to\boldsymbol{E}^*$ equal to the cofibration $G'\circ e$, we see that $\boldsymbol{E}^*$ is a left Reedy fibrant replacement for ${}'\widehat{\boldsymbol{W}}^n\otimes^* S^{n-1}$ (since $\boldsymbol{K}^*\times\boldsymbol{D}^*$ is a product of fibrant objects), G is a left Reedy cofibration, and they fit into a diagram

(3.23)



in which both the left and right squares commute, and $\bar{r}\circ\bar{e}=\mathrm{Id}$.

Applying the functorial procedure of Definition 2.1(c) to the two vertical arrows in (3.23) again, we obtain n–stage comparison map $\Phi\colon\mathcal{W}\to{}'\mathcal{W}$ extending the given $(n-1)$–stage comparison map.

(ii) Reducing to the case of one algebraic resolution Assume $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ are associated respectively to the two CW resolutions $V^{(0)}_\bullet$ and $V^{(1)}_\bullet$ of the Θ_R–algebra $\Gamma=H^*(\boldsymbol{Y};R)$, with CW bases $(\overline{V}^{(i)}_n)_{n\in\mathbb{N}}$ for $i=0,1$.

By Lemma 3.7, there is a third CW resolution ${}'\varepsilon\colon {}'V_\bullet\to\Gamma$, with CW basis $({}'\overline{V}_n)_{n\in\mathbb{N}}$, equipped with algebraic comparison maps $\Psi^{(i)}\colon V^{(i)}_\bullet\to{}'V_\bullet$ for $i=0,1$. By step (i),

there are then two sequential realizations ${}'\mathcal{W}^{(i)}$ of ${}'V_\bullet \to \Gamma$, for $i = 0, 1$, each equipped with a comparison map $\Psi^{(i)}\colon \mathcal{W}^{(i)} \to {}'\mathcal{W}^{(i)}$ over $\Psi^{(i)}$. Thus we are reduced to dealing with the case where the two (cofibrant) sequential realizations $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$ (ie the ${}'\mathcal{W}^{(i)}$ just constructed) are of the same CW resolution $V_\bullet \to \Gamma$, with CW basis $(\overline{V}_n)_{n\in\mathbb{N}}$. We construct a zigzag of comparison maps between them, by induction on $n \geq 0$:

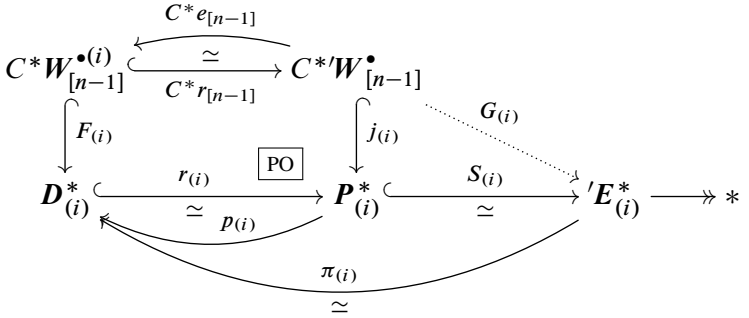
We assume by induction the existence of a cospan of $(n-1)$–stage trivial comparison maps $\Phi^{(i)}\colon \mathcal{W}^{(i)} \to \mathcal{W}$ for $i = 0, 1$ over $\mathrm{Id}_{V_\bullet}$. For $n = 0$ this is simply

$$\boldsymbol{W}^{\bullet(0)}_{[-1]} = c(\boldsymbol{Y})^\bullet = \boldsymbol{W}^{\bullet(1)}_{[-1]}.$$

By Definition 2.1(c), the nth stage for $\mathcal{W}^{(i)}$ is determined by the choice of left Reedy fibrant replacements $\boldsymbol{D}^*_{(i)}$ of $\overline{\boldsymbol{W}}^n \otimes^* S^{n-1}$ (where $\overline{\boldsymbol{W}}^n$ is some realization of the nth algebraic CW basis object $\overline{V}_n$), together with left Reedy cofibrations $F^{(i)}\colon C^*\boldsymbol{W}^{\bullet(i)}_{[n-1]} \to \boldsymbol{D}^*_{(i)}$ for $i = 0, 1$ realizing the given attaching map $\overline{\partial}^n_0\colon \overline{V}_n \to C_{n-1}V_\bullet$.

For $i = 0, 1$, consider the following diagram in the left Reedy model category $\mathsf{Ch}^{\mathcal{C}}_{\leq n}$, in which $\boldsymbol{P}^*$ is again the pushout of the upper left square

(3.24)

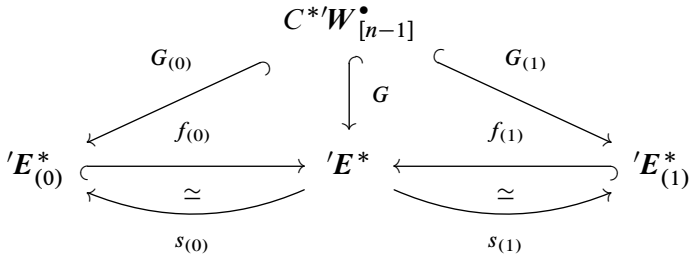


The retraction $p_{(i)}$ for the trivial cofibration $r_{(i)}$ is induced by the retraction $C^*e^{(i)}_{[n-1]}$ for the trivial cofibration $C^*r^{(i)}_{[n-1]}$, so $p_{(i)}$ is a weak equivalence. Factor $p_{(i)}$ as a trivial cofibration $k_{(i)}\colon \boldsymbol{P}^*_{(i)} \to {}'\boldsymbol{E}^*_{(i)}$ followed by a fibration $\pi_{(i)}\colon {}'\boldsymbol{E}^*_{(i)} \to \boldsymbol{D}^*_{(i)}$ (also a weak equivalence), so ${}'\boldsymbol{E}^*_{(i)}$ is in particular a fibrant replacement for $\boldsymbol{P}^*_{(i)}$ since ${}'\boldsymbol{E}^*_{(i)}$ is fibrant.

Now set $G_{(i)} := S_{(i)} \circ j_{(i)}\colon C^*{}'\boldsymbol{W}^\bullet_{[n-1]} \hookrightarrow {}'\boldsymbol{E}^*_{(i)}$ (a cofibration). Because the maps $\pi_i \circ G_{(i)} = F_{(i)} \circ C^*e_{[n-1]}$ realize the same algebraic attaching map $\phi\colon \overline{V}_n \otimes_* S^{n-1} \to C^*V_\bullet$ for $i = 0, 1$, they are weakly equivalent in the arrow category of $\mathsf{Ch}^{\mathcal{C}}_{\leq n-1}$. Thus $G^{(0)}$ and $G^{(1)}$ are weakly equivalent fibrant and cofibrant objects in the under category $C^*{}'\boldsymbol{W}^\bullet_{[n-1]}\backslash\mathsf{Ch}^{\mathcal{C}}_{\leq n-1}$ with its standard model category structure (see Hirschhorn [28, Theorem 7.6.5(a)]). We can therefore apply Lemma 3.1(b) to obtain an intermediate

object G fitting into the following diagram, in which all four triangles commute, and $s_{(i)} \circ f_{(i)} = \mathrm{Id}$ for $i = 0, 1$:

(3.25)



Again applying the functors of Definition 2.1(c) to all three downward arrows of (3.25) yields a new n–stage sequential realization ${}'\mathcal{W}$ (corresponding to $G\colon C^*\boldsymbol{W}^{\bullet}_{[n-1]} \hookrightarrow {}'\boldsymbol{E}^*$) with two new n–stage trivial comparison maps ${}'\Phi_{(i)}\colon {}'\mathcal{W}^{(i)} \to {}'\mathcal{W}$ for $i = 0, 1$.

The two composites

$$\mathcal{W}^{(0)} \xrightarrow{\Phi^{(0)}} {}'\mathcal{W}^{(0)} \xrightarrow{{}'\Phi_{(0)}} {}'\mathcal{W} \xleftarrow{{}'\Phi_{(1)}} {}'\mathcal{W}^{(1)} \xleftarrow{\Phi^{(1)}} \mathcal{W}^{(1)} \tag{3.26}$$

then yield the required cospan of n–stage comparison maps. □

4 Higher cohomology operations

The notion of secondary and higher cohomology operations has a long history in homotopy theory, going back to the 1950s, but there is no completely satisfactory general theory of such operations. Here we follow the point of view taken in Blanc and Markl [15] and Blanc, Johnson and Turner [12], where they are subsumed under the notion of general pointed higher homotopy operations.

We want to think of a cofibrant sequential realization $\mathcal{W}$ for a space $\boldsymbol{Y}$ as providing a template for an infinite sequence of operations of order n for $n = 2, 3, \dots$, potentially acting on any space $\boldsymbol{Z}$ with $H^*(\boldsymbol{Z}; R)$ (abstractly) isomorphic to $H^*(\boldsymbol{Y}; R)$. The operation of order n is defined only when specific choices have been made inductively for all lower order operations in such a way that they all vanish.

The sequence of such choices is called a "strand" of the higher cohomology operation associated to the given sequential realization $\mathcal{W}$, and the corresponding "system of higher cohomology operations" will be an equivalence class of strands under comparison maps.

4.1 Definition Let $\Gamma = H^*(\boldsymbol{Y}; R)$ be the Θ_R–algebra associated to a space $\boldsymbol{Y}$, and $\varepsilon\colon V_\bullet \to \Gamma$ a CW resolution with CW basis $(\overline{V}_n)_{n\in\mathbb{N}}$. Assume we are also given *initial data* consisting of a sequential realization $\mathcal{W} = \langle \boldsymbol{W}^\bullet_{[n]}, \widetilde{\boldsymbol{W}}^\bullet_{[n]}\rangle_{n\in\mathbb{N}}$ of $V_\bullet$, and an isomorphism of Θ_R–algebras $\vartheta\colon \Gamma \to H^*(\boldsymbol{Z}; R)$ for some space $\boldsymbol{Z}$.

An *n–strand* for $0 \le n \le \infty$ $\mathcal{S}_{[n]} = (\boldsymbol{\varepsilon}_{[0]}, \boldsymbol{\varepsilon}_{[1]}, \dots, \boldsymbol{\varepsilon}_{[n]})$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ consists of a compatible collection of coaugmentations $\boldsymbol{\varepsilon}_{[k]}\colon \boldsymbol{Z} \to \boldsymbol{W}^\bullet_{[k]}$ for $k = 0, \dots, n$ realizing $\vartheta \circ \varepsilon\colon V_\bullet \to H^*(\boldsymbol{Z}; R)$ through simplicial dimension n. Compatibility means that $\boldsymbol{\varepsilon}_{[k-1]} = \pi_{[k]} \circ \boldsymbol{\varepsilon}_{[k]}$ (cf Definition 2.1(c)). In particular, an *∞–strand* for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ is an infinite sequence $\mathcal{S}_{[\infty]} := (\boldsymbol{\varepsilon}_{[0]}, \dots, \boldsymbol{\varepsilon}_{[n]}, \dots)$ of such compatible coaugmentations.

Given an $(n-1)$–strand $\mathcal{S}_{[n-1]}$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$, consider the composite ξ of

$$\boldsymbol{Z} \xrightarrow{\boldsymbol{\varepsilon}_{[n-1]}} \boldsymbol{W}^0_{[n-1]} \xrightarrow{F^0_{[n-1]}} \overline{P\Omega^{n-2}\boldsymbol{W}^n} \xrightarrow{\bar{p}^{n-2}} \overline{\Omega^{n-2}\boldsymbol{W}^n} \xrightarrow{\bar{\iota}^{n-2}} \overline{P\Omega^{n-3}\boldsymbol{W}^n},$$

which by (2.14) represents the component

$$q^2_{[n-1]} \circ d^1 \circ d^0 \circ \boldsymbol{\varepsilon}_{[n-1]}$$

of the iterated coface map from $\boldsymbol{Z}$ into $\overline{P\Omega^{n-3}\boldsymbol{W}^n}$. As in Lemma 2.19, since

$$d^1 \circ d^0 \circ \boldsymbol{\varepsilon}_{[n-1]} = d^2 \circ d^1 \circ \boldsymbol{\varepsilon}_{[n-1]}$$

and $q^2_{[n-1]} \circ d^2 = 0$ by (2.15), we see that ξ is the zero map. Since $\bar{\iota}^{n-2}$ is monic, this means that the composite

$$\bar{p}^{n-2} \circ F^0_{[n-1]} \circ \boldsymbol{\varepsilon}_{[n-1]}$$

is already zero, so $F^0_{[n-1]} \circ \boldsymbol{\varepsilon}_{[n-1]}$ factors through the fiber $\overline{\Omega^{n-1}\boldsymbol{W}^n}$ of $\bar{p}^{n-2}$. We denote the resulting map by

$$a^{-1}_{[n-1]}\colon \boldsymbol{Z} \to \overline{\Omega^{n-1}\boldsymbol{W}^n},$$

with

$$\bar{\iota}^{n-1} \circ a^{-1}_{[n-1]} = F^0_{[n-1]} \circ \boldsymbol{\varepsilon}_{[n-1]}, \tag{4.2}$$

as in (2.20).

Note that ϑ induces an isomorphism $[\boldsymbol{Z}, \overline{\Omega^{n-1}\boldsymbol{W}^n}] \cong \Gamma\{\overline{\Omega^{n-1}\boldsymbol{W}^n}\}$ so the homotopy class $[a^{-1}_{[n-1]}]$ may be identified with an element

$$\mathrm{Val}(\mathcal{S}_{[n-1]}) \in \Gamma\{\overline{\Omega^{n-1}\boldsymbol{W}^n}\} \cong \Gamma\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\}, \tag{4.3}$$

called the *value* of the $(n-1)$–strand $\mathcal{S}_{[n-1]}$.

4.4 Remark We do not need the full sequential realization $\mathcal{W}$ to define $\mathrm{Val}(\mathcal{S}_{[n]})$, but only the restricted cosimplicial set $\widetilde{\boldsymbol{W}}^{\bullet}_{[n]}$ of its $(n-1)^{\text{st}}$ stage $\boldsymbol{W}^{\bullet}_{[n-1]}$. Thus a 0–*strand* is completely determined by a choice of a realization $\boldsymbol{\varepsilon}\colon \boldsymbol{Z} \to \boldsymbol{W}^0_{[0]}$ of $\vartheta \circ \varepsilon\colon V_0 \to H^*(\boldsymbol{Z}; R)$. Such an $\boldsymbol{\varepsilon}$ always exists, and is unique up to homotopy.

Note also that Definition 4.1 can be stated purely in the language of $\boldsymbol{\Theta}_R$–mapping algebras — see Baues and Blanc [2] and the appendix below.

4.5 Lemma *Given an $(n-1)$–strand $\mathcal{S}_{[n-1]} = (\boldsymbol{\varepsilon}_{[0]}, \boldsymbol{\varepsilon}_{[1]}, \dots, \boldsymbol{\varepsilon}_{[n-1]})$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$, the coaugmentation $\boldsymbol{\varepsilon}_{[n-1]}\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}_{[n-1]}$ extends to a coaugmentation $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}_{[n]}$ if and only if $\mathrm{Val}(\mathcal{S}_{[n-1]}) = 0$ in $\Gamma\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\}$.*

Proof If the value is zero, we can choose a nullhomotopy $F^{-1}_{[n]}\colon \boldsymbol{Z} \to \overline{P\Omega^{n-1}\boldsymbol{W}^n}$ for $a^{-1}_{[n-1]}$, (in the sense of Quillen [32, Section I.2]), with $\bar{p}^{n-1} \circ F^{-1}_{[n]} = a^{-1}_{[n-1]}$. This extends $\boldsymbol{\varepsilon}_{[n-1]}$ to a coaugmentation $\widetilde{\boldsymbol{\varepsilon}}_{[n]}\colon \boldsymbol{Z} \to \widetilde{\boldsymbol{W}}^0_{[n]}$ as in the proof of Theorem A.11, which also shows how to extend $\widetilde{\boldsymbol{\varepsilon}}_{[n]}$ to a coaugmentation for $\widetilde{\boldsymbol{W}}^{\bullet}_{[n]}$, and thus (after making it Reedy cofibrant) for $\boldsymbol{W}^{\bullet}_{[n]}$.

Conversely, since $\boldsymbol{W}^0_{[n]} = \widetilde{\boldsymbol{W}}^0_{[n]} = \boldsymbol{W}^0_{[n-1]} \times \overline{P\Omega^{n-1}\boldsymbol{W}^n}$ by (2.10) and step (iii) of Section 2.5, given an extension $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Z} \to \boldsymbol{W}^0_{[n]}$ of $\boldsymbol{\varepsilon}_{[n-1]}$, we can compose it with the projection $q^0_{[n]}\colon \widetilde{\boldsymbol{W}}^0_{[n]} \to \overline{P\Omega^{n-1}\boldsymbol{W}^n}$ to obtain a nullhomotopy $F^{-1}_{[n]}$ for $a^{-1}_{[n-1]}$. □

4.6 Correspondence of strands Note that if $\Psi\colon V_\bullet \to {}'V_\bullet$ is an algebraic comparison map between two CW resolutions for a Θ_R–algebra Γ, as in (3.6), the mutually inverse weak equivalences $\varphi\colon V_\bullet \to {}'V_\bullet$ and $\rho\colon {}'V_\bullet \to V_\bullet$ induce mutually inverse isomorphisms of Θ_R–algebras $\varphi_\#\colon \pi_0 V_\bullet \cong \Gamma \cong \pi_0 {}'V_\bullet :\rho_\#$.

On the other hand, if $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$ is an n–stage comparison map between two sequential realizations over Ψ, as in (3.16), then $(\bar{r}^{n-1})_*\colon \Gamma\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\} \hookrightarrow \Gamma\{\Omega^{n-1}\overline{{}'\boldsymbol{W}}^n\}$ is just a split inclusion, with retraction $(\bar{e}^{n-1})_*\colon \Gamma\{\Omega^{n-1}\overline{{}'\boldsymbol{W}}^n\} \twoheadrightarrow \Gamma\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\}$.

Let $\vartheta\colon \Gamma \cong H^*(\boldsymbol{Z}; R)$ be an isomorphism of Θ_R–algebras, and let $\mathcal{S}_{[n]}$ and ${}'\mathcal{S}_{[n]}$ be two n–strands for the initial data $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ and $({}'\mathcal{W}, \boldsymbol{Z}, \vartheta)$, respectively. If $\Phi = \big\langle e_{[n]}, r_{[n]}, (\overline{Pe}^k_n)^{n-1}_{k=0}, (\bar{e}^k_n)^{n-1}_{k=0}, (\overline{Pr}^k_n)^{n-1}_{k=0}, (\bar{r}^k_n)^{n-1}_{k=0}\big\rangle^N_{n=0}$ is an n–stage comparison map as above, we write ${}'\mathcal{S}_{[n]} = r^{\#}(\mathcal{S}_{[n]})$ if ${}'\boldsymbol{\varepsilon}_{[k]} = r^0_{[k]} \circ \boldsymbol{\varepsilon}_{[k]}$ for each $0 \le k \le n$, and $\mathcal{S}_{[n]} = e^{\#}({}'\mathcal{S}_{[n]})$ if $\boldsymbol{\varepsilon}_{[k]} = e^0_{[k]} \circ {}'\boldsymbol{\varepsilon}_{[k]}$ for each $0 \le k \le n$.

From (3.13), (3.15), and (4.2) we see that

$$\textbf{(4.7)} \qquad \mathrm{Val}(r^{\#}(\mathcal{S}_{[n]})) = \bar{r}^n_*(\mathrm{Val}(\mathcal{S}_{[n]})) \quad \text{and} \quad \mathrm{Val}(e^{\#}({}'\mathcal{S}_{[n]})) = \bar{e}^n_*(\mathrm{Val}({}'\mathcal{S}_{[n]})).$$

Therefore, given an n–stage comparison map $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$ as above, an n–strand $\mathcal{S}_{[n]}$ for $\mathcal{W}$ and an n–strand ${}'\mathcal{S}_{[n]}$ for ${}'\mathcal{W}$, we see that:

4.8 (a) $\mathrm{Val}(\mathcal{S}_{[n]}) = 0$ if and only if $\mathrm{Val}(r^{\#}(\mathcal{S}_{[n]})) = 0$.

(b) If $\mathrm{Val}({}'\mathcal{S}_{[n]}) = 0$ then $\mathrm{Val}(e^{\#}({}'\mathcal{S}_{[n]})) = 0$, but not necessarily conversely.

This explains the need for the following:

4.9 Definition Given spaces $\boldsymbol{Y}$ and $\boldsymbol{Z}$ with $\vartheta\colon H^*(\boldsymbol{Y}; R) \xrightarrow{\cong} H^*(\boldsymbol{Z}; R)$, we define two equivalence relations $\sim$ and $\approx$ on n–strands for $\boldsymbol{Z}$ (with respect to various sequential realizations):

The *weak equivalence* relation of strands $\sim$ is generated by $e^{\#}({}'\mathcal{S}_{[n]}) \sim {}'\mathcal{S}_{[n]}$ for any n–strand ${}'\mathcal{S}_{[n]}$ for $({}'\mathcal{W}, \boldsymbol{Z}, \vartheta)$ and any comparison map $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$. We denote the corresponding equivalence class by $[\mathcal{S}_{[n]}]$.

The *strong equivalence* relation of strands $\approx$ is generated by the relation $e^{\#}({}'\mathcal{S}_{[n]}) \approx {}'\mathcal{S}_{[n]}$ for any n–strand ${}'\mathcal{S}_{[n]}$ for $({}'\mathcal{W}, \boldsymbol{Z}, \vartheta)$ and any comparison map $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$ satisfying

$$(\bar{s}^n_{[n+1]})_{\#}(\mathrm{Val}({}'\mathcal{S}_{[n]})) = 0, \tag{4.10}$$

in the notation of Remark 3.17. We denote strong equivalence classes by $[\![\mathcal{S}_{[n]}]\!]$.

4.11 Remark Clearly $\mathcal{S}_{[n]} \approx {}'\mathcal{S}_{[n]}$ implies that $\mathcal{S}_{[n]} \sim {}'\mathcal{S}_{[n]}$, and both notions coincide if the comparison map $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$ is trivial — that is, if in the underlying algebraic comparison map Ψ, $V_\bullet$ and ${}'V_\bullet$ have isomorphic CW bases.

Also $\mathcal{S}_{[n]} \approx r^{\#}(\mathcal{S}_{[n]})$ (and thus $\mathcal{S}_{[n]} \sim r^{\#}(\mathcal{S}_{[n]})$) for *any* comparison map $\Phi\colon \mathcal{W} \to {}'\mathcal{W}$, since $e^{\#}r^{\#}(\mathcal{S}_{[n]}) = \mathcal{S}_{[n]}$ and $(\bar{s}^n_{[n+1]})_{\#}\big(\mathrm{Val}(r^{\#}(\mathcal{S}_{[n]}))\big) = (\bar{s}^n_{[n+1]})_{\#}(\bar{r}^{\,n})_*(\mathrm{Val}(\mathcal{S}_{[n]})) = 0$ by (4.7) and (3.18).

4.12 Lemma *When $\mathcal{S}_{[n]} \approx {}'\mathcal{S}_{[n]}$, $\mathrm{Val}(\mathcal{S}_{[n]}) = 0$ if and only if $\mathrm{Val}({}'\mathcal{S}_{[n]}) = 0$.*

Proof This follows from (4.7), since by (4.10) we see that $\mathrm{Val}({}'\mathcal{S}_{[n]}) \in \Gamma\{\Omega^{n\prime}\overline{\boldsymbol{W}}^{n+1}\}$ is uniquely determined by its image $\bar{e}^n_*(\mathrm{Val}({}'\mathcal{S}_{[n]}))$ under the projection in $\Gamma\{\Omega^n\overline{\boldsymbol{W}}^{n+1}\}$. □

4.13 Definition Given a space $\boldsymbol{Y}$ with $\Gamma := H^*(\boldsymbol{Y}; R)$, we want to think of each sequential realization $\mathcal{W}$ for $\boldsymbol{Y}$ as a template for a countable sequence $\langle\langle \boldsymbol{Y} \rangle\rangle = (\langle\langle \boldsymbol{Y} \rangle\rangle_n)_{n=2}^{\infty}$ of higher operations, where for each $n \geq 2$, we define the *universal*

n^{th} order cohomology operation $\langle\langle \boldsymbol{Y}\rangle\rangle_n$ associated to the space $\boldsymbol{Y}$ to be the function that assigns to every $(n-1)$–strand $\mathcal{S}_{[n-1]}$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ the class

$$\langle\langle \boldsymbol{Y}\rangle\rangle_n(\mathcal{S}_{[n-1]}) := \mathrm{Val}(\mathcal{S}_{[n-1]}) \in \Gamma\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\}.$$

We say that the operation $\langle\langle \boldsymbol{Y}\rangle\rangle_n$ *vanishes for* $\boldsymbol{Z}$ if there is a cofibrant sequential realization $\mathcal{W}$ and an $(n-1)$–strand $\mathcal{S}_{[n-1]}$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ such that $\langle\langle \boldsymbol{Y}\rangle\rangle_n(\mathcal{S}_{[n-1]}) = 0$. By Lemma 4.12, this notion of the vanishing depends only the strong equivalence classes $[\![\mathcal{S}_{[n-1]}]\!]$ of the strand.

4.14 Definition Given spaces $\boldsymbol{Y}$ and $\boldsymbol{Z}$ with $\vartheta\colon H^*(\boldsymbol{Y}; R) \xrightarrow{\cong} H^*(\boldsymbol{Z}; R)$ and a sequential realization $\mathcal{W}$ of $\boldsymbol{Y}$, we say that $\langle\langle \boldsymbol{Y}\rangle\rangle$ *vanishes coherently for* $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ if there is an ∞–strand $\mathcal{S}_{[\infty]}$ for $\mathcal{W}$ as in Definition 4.1; that is, for each $n \geq 2$, we have a given $(n-1)$–strand $\mathcal{S}_{[n-1]}$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ such that $\mathrm{Val}(\mathcal{S}_{[n-1]}) = 0$, which extends to the next n–strand $\mathcal{S}_{[n]}$ using Lemma 4.5.

4.15 Example For any sequential realization $\mathcal{W}$ of a space $\boldsymbol{Y}$, the sequence $\langle\langle \boldsymbol{Y}\rangle\rangle$ vanishes coherently for $(\mathcal{W}, \boldsymbol{Y}, \mathrm{Id}_\Gamma)$, since then we have a given coaugmentation $\boldsymbol{\varepsilon}\colon \boldsymbol{Y} \to \boldsymbol{W}^\bullet$, which we can then project to each $\boldsymbol{W}^\bullet_{[n]}$ (see Proposition 2.24) to yield $\boldsymbol{\varepsilon}_{[n]}$ showing that $\mathrm{Val}(\mathcal{S}_{[n-1]}) = 0$ for the corresponding $(n-1)$–strand $\mathcal{S}_{[n-1]}$, by Lemma 4.5.

4.16 Remark A priori, each individual strand $\mathcal{S}_{[n-1]}$ has a different template for $\langle\langle \boldsymbol{Y}\rangle\rangle$ — namely, the restricted cosimplicial set $\widetilde{\boldsymbol{W}}^\bullet_{[n]}$ of the $(n-1)^{\mathrm{st}}$ stage $\boldsymbol{W}^\bullet_{[n-1]}$ of $\mathcal{W}$. However, the following result, which follows from Theorem 3.20, shows that we can in fact use any one cofibrant sequential realization to calculate $\langle\langle \boldsymbol{Y}\rangle\rangle_n$:

4.17 Key Lemma *Given $\boldsymbol{Y}$ and $\vartheta\colon H^*(\boldsymbol{Y}; R) \xrightarrow{\cong} H^*(\boldsymbol{Z}; R)$ as above, $\langle\langle \boldsymbol{Y}\rangle\rangle_n$ vanishes for $\boldsymbol{Z}$ if and only if for* every *n–stage cofibrant sequential realization $\mathcal{W}$ of $\boldsymbol{Y}$, there is an $(n-1)$–strand $\mathcal{S}_{[n-1]}$ for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ such that* $\mathrm{Val}(\mathcal{S}_{[n-1]}) = 0$.

Proof By definition, $\langle\langle \boldsymbol{Y}\rangle\rangle_n$ vanishes for $\boldsymbol{Z}$ if there is *some* cofibrant n–stage sequential realization $'\mathcal{W}$ of $\boldsymbol{Y}$ and an $(n-1)$–strand $\mathcal{S}'_{[n-1]}$ for $('\mathcal{W}, \boldsymbol{Z}, \vartheta)$ such that $\mathrm{Val}(\mathcal{S}'_{[n-1]}) = 0$. By Theorem 3.20 we know that there is a finite zigzag of cospans of comparison maps connecting $'\mathcal{W}$ to $\mathcal{W}$, say

$$'\Phi_{(1)}\colon \mathcal{W}^{(0)} = {}'\mathcal{W} \to \mathcal{W}^{(1)}, \quad {}'\Phi_{(2)}\colon \mathcal{W}^{(2)} \to \mathcal{W}^{(1)}, \quad {}'\Phi_{(3)}\colon \mathcal{W}^{(2)} \to \mathcal{W}^{(3)},$$

and so on until ${}'\Phi_{(N)}\colon \mathcal{W}^{(N-1)} \to \mathcal{W}^{(N)} = \mathcal{W}$. If ${}'\Phi_{(1)} = \langle e_{[k]}, r_{[k]}, \dots\rangle_{k=0}^{n}$ as in (3.16), we set $\mathcal{S}^{(1)}_{[n-1]} := r^{\#}(\mathcal{S}'_{[n-1]})$ (an $(n-1)$–strand for $\mathcal{W}^{(1)}$), and see from (4.7) that $\mathrm{Val}(\mathcal{S}^{(1)}_{[n-1]}) = 0$. Similarly, if ${}'\Phi_{(2)} = \langle {}'e_{[k]}, {}'r_{[k]}, \dots\rangle_{k=0}^{n}$ we set $\mathcal{S}^{(2)}_{[n-1]} := {}'e^{\#}(\mathcal{S}^{(1)}_{[n-1]})$ (an $(n-1)$–strand for $\mathcal{W}^{(2)}$), and again see from (4.7) that $\mathrm{Val}(\mathcal{S}^{(2)}_{[n-1]}) = 0$. Continuing in this way we finally obtain an $(n-1)$–strand $\mathcal{S}_{[n-1]} = \mathcal{S}^{(N)}_{[n-1]}$ for $\mathcal{W}^{(N)} = \mathcal{W}$ with $\mathrm{Val}(\mathcal{S}_{[n-1]}) = 0$, as required.

The converse follows from the fact that $H^*(\boldsymbol{Y}; R)$ has at least one CW resolution by Remark 1.31, and thus there is at least one cofibrant sequential realization for $\boldsymbol{Y}$ by Theorem 2.33. □

Clearly, the $(n-1)$–strands $\mathcal{S}'_{[n-1]}$ and $\mathcal{S}_{[n-1]}$ are weakly equivalent. However, they are not necessarily strongly equivalent, since there is no reason for (4.10) to hold for the even-numbered comparison maps above ${}'\Phi_{(2)}$, ${}'\Phi_{(4)}$, and so on.

4.18 Theorem *For R either $\mathbb{F}_p$ or a field of characteristic 0, let $\boldsymbol{Y}$ and $\boldsymbol{Z}$ be R–good spaces with isomorphic Θ_R–algebras. Then the following are equivalent:*

(a) *The system of higher cohomology operations $\langle\langle \boldsymbol{Y}\rangle\rangle$ vanishes coherently for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ for* some *cofibrant sequential realization $\mathcal{W}$ of $\boldsymbol{Y}$ and some ϑ.*

(b) *$\langle\langle \boldsymbol{Y}\rangle\rangle$ vanishes coherently for* every *cofibrant sequential realization of $\boldsymbol{Y}$.*

(c) *$\boldsymbol{Y}$ and $\boldsymbol{Z}$ are R–equivalent.*

Proof (a) $\Longleftrightarrow$ (b) This is by Key Lemma 4.17.

(a) $\Longrightarrow$ (c) Assume that $\langle\langle \boldsymbol{Y}\rangle\rangle$ vanishes coherently — that is, there is an ∞–strand $\mathcal{S}_{[\infty]}$ for some $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ (where $\mathcal{W}$ need not be cofibrant), and thus coaugmentations $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}_{[n]}$ for all $n \geq 0$. These fit together to define a coaugmentation $\boldsymbol{\varepsilon}\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}$ for $\boldsymbol{W}^{\bullet} := \operatorname{holim} \boldsymbol{W}^{\bullet}_{[n]}$, which induces an isomorphism

$$H^*(\operatorname{Tot} \boldsymbol{W}^{\bullet}; R) \to H^*(\boldsymbol{Z}; R).$$

Since $\boldsymbol{Y}$ is R–good, it is R–equivalent to the total space $\operatorname{Tot} \boldsymbol{W}^{\bullet} \simeq \widehat{L}_{\mathcal{G}} \boldsymbol{Y}$ (see Section 1.22), and thus the map $f\colon \boldsymbol{Z} \to \operatorname{Tot} \boldsymbol{W}^{\bullet}$ induced by the coaugmentation $\boldsymbol{\varepsilon}$ realizes ϑ, so $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are related by a cospan of R–equivalences.

(c) $\Longrightarrow$ (a) Conversely, if $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are R–equivalent, we have a zigzag of R–equivalences from $\boldsymbol{Y}$ to $\boldsymbol{Z}$ inducing an isomorphism of Θ_R–algebras $\vartheta\colon H^*(\boldsymbol{Y}; R) \to H^*(\boldsymbol{Z}; R)$, so it suffices to consider the following two special cases:

(i) Given a R–equivalence $f\colon \boldsymbol{Z} \to \boldsymbol{Y}$, and some (not necessarily cofibrant) sequential realization $\mathcal{W}$ for $\boldsymbol{Y}$, by precomposing the coaugmentations $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Y} \to \boldsymbol{W}^{\bullet}_{[n]}$ with f we obtain coaugmentations $\boldsymbol{\varepsilon}_{[n]} \circ f\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}_{[n]}$, still realizing $V_{\bullet} \to \Gamma$, since $f^{\#}\colon \Gamma \to H^*(\boldsymbol{Z}; R)$ is an isomorphism. This yields an ∞–strand for $(\mathcal{W}, \boldsymbol{Z}, \vartheta)$ by Lemma 4.5.

(ii) On the other hand, given a R–equivalence $g\colon \boldsymbol{Y} \to \boldsymbol{Z}$ and any cofibrant sequential realization ${}'\mathcal{W}$ for $\boldsymbol{Z}$, by precomposing the coaugmentations $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Z} \to {}'\boldsymbol{W}^{\bullet}_{[n]}$ with g as in (a) we obtain coaugmentations $\boldsymbol{\varepsilon}_{[n]} \circ g\colon \boldsymbol{Y} \to {}'\boldsymbol{W}^{\bullet}_{[n]}$ realizing $V_{\bullet} \to \Gamma$, and thus making ${}'\mathcal{W}$ itself with the new coaugmentations into a cofibrant sequential realization ${}''\mathcal{W}$ for $\boldsymbol{Y}$. The coaugmentations $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Z} \to {}'\boldsymbol{W}^{\bullet}_{[n]}$ form an ∞–strand $\mathcal{S}_{[\infty]}$ for ${}''\mathcal{W}$, showing that $\langle\langle \boldsymbol{Y} \rangle\rangle$ vanishes coherently for $({}''\mathcal{W}, \boldsymbol{Z}, \vartheta)$.

This completes the proof. □

4.19 Corollary *If $\boldsymbol{Y}$ and $\boldsymbol{Y}'$ are R–equivalent R–good spaces, any cofibrant sequential realization $\mathcal{W} = \langle \boldsymbol{W}^{\bullet}_{[n]}, \widetilde{\boldsymbol{W}}^{\bullet}_{[n]} \rangle_{n\in\mathbb{N}}$ for $\boldsymbol{Y}$ is also a (cofibrant) sequential realization for $\boldsymbol{Y}'$.*

Proof The R–equivalence implies that there is an isomorphism $\vartheta\colon H^*(\boldsymbol{Y}; R) \cong H^*(\boldsymbol{Y}'; R)$, so by the theorem there is an ∞–strand $\mathcal{S}_{[\infty]}$ for $(\mathcal{W}, \boldsymbol{Y}', \vartheta)$, and thus a coaugmentation $\boldsymbol{\varepsilon}'\colon \boldsymbol{Y}' \to \boldsymbol{W}^{\bullet}$. □

4.20 Low-dimensional cases As noted in Remark 4.4, given a simplicial set $\boldsymbol{Z}$ equipped with an isomorphism of Θ_R–algebras $\vartheta\colon \Gamma \to H^*(\boldsymbol{Z}; R)$ and a sequential realization $\mathcal{W}$ for Γ as above, we can always define a coaugmentation $\boldsymbol{\varepsilon}_{[0]}\colon \boldsymbol{Z} \to \boldsymbol{W}^0_{[0]}$ realizing $\phi = \vartheta \circ \varepsilon\colon V_0 \to H^*(\boldsymbol{Z}; R)$, which is unique up to homotopy, as in the proof of Lemma 4.5. Moreover, if $\overline{d}^0_0\colon \overline{\boldsymbol{W}}^0 \to \overline{\boldsymbol{W}}^1$ realizes the first attaching map $\overline{\partial}^1_0\colon \overline{V}_1 \to V_0 = \overline{V}_0$, then $a^{-1}_{[0]} := \overline{d}^0_0 \circ \boldsymbol{\varepsilon}_{[0]}$ is nullhomotopic, since it realizes $\varepsilon \circ \overline{\partial}^1_0$ (see (A.12)). Thus we can always choose a nullhomotopy $F^{-1}_{[0]}$ for $a^{-1}_{[0]}$, and use it do define $\boldsymbol{\varepsilon}_{[1]}\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}_{[1]}$, as in the proof of the lemma. Note however that while the map $\boldsymbol{\varepsilon}_{[0]}$ is unique up to homotopy, the map $\boldsymbol{\varepsilon}_{[1]}$ depends on our choice of $F^{-1}_{[0]}$.

This explains why our definition of nth order cohomology operations only makes sense for $n \geq 2$.

Thus the first case of interest is $n = 2$. Since $\boldsymbol{\Theta}_R$ consists of abelian group objects, we can replace our 2–truncated restricted cosimplicial diagram $\boldsymbol{W}^{\bullet}_{[1]} \to \overline{\boldsymbol{W}}^2$ by

$$\boldsymbol{Z} \xrightarrow{\boldsymbol{\varepsilon}_{[1]}} \boldsymbol{W}^0_{[1]} \xrightarrow{d^0 - d^1} \boldsymbol{W}^1_{[1]} \xrightarrow{\overline{d}^0_1} \overline{\boldsymbol{W}}^2, \tag{4.21}$$

where the fact that $d^0 \circ \boldsymbol{\varepsilon}_{[1]} = d^1 \circ \boldsymbol{\varepsilon}_{[1]}$ means that the first composite is nullhomotopic, while the fact that $\overline{d}^0_1 \circ d^0$ is nullhomotopic and $\overline{d}^0_1 \circ d^1 = 0$ means that the second composite is also nullhomotopic.

In particular, the first map in (4.21) represents (a collection of) R–cohomology classes α, while the remaining two represent R–cohomology operations ξ and ζ, with $\xi(\alpha) = 0$ and $\zeta \circ \xi = 0$. Thus our universal secondary operation is just a (collection of) secondary cohomology operations in the sense of Adams (see [1]), taking values in $[\boldsymbol{Z}, \Omega\overline{\boldsymbol{W}}^2] = H^{*-1}(\boldsymbol{Z}; R)$.

4.22 Remark Our higher cohomology operations are modeled on Adams' (stable) secondary cohomology operations (see also Harper [26]). However, the more delicate questions involving the prerequisites for an $(n+1)^{\text{st}}$ order operation to be defined, and the dependence on various choices made, are hidden here in the two components of the n–strand $\mathcal{S}_{[n]}$, consisting of:

(a) The nth stage $\boldsymbol{W}^{\bullet}_{[n]}$ in the sequential approximation $\mathcal{W}$ encodes a preliminary choice of nullhomotopies for that part of the diagram consisting only of spaces in $\boldsymbol{\Theta}_R$ (the representing spaces for cohomology).

(b) The data associated to the specific simplicial set $\boldsymbol{Z}$ consists of the coaugmentation $\boldsymbol{\varepsilon}_{[n]}\colon \boldsymbol{Z} \to \boldsymbol{W}^{\bullet}_{[n]}$, which itself is determined by:

  i. The coherent system of earlier choices made, encoded in the $(n-1)$–strand $\mathcal{S}_{[n-1]} = (\boldsymbol{\varepsilon}_{[0]}, \boldsymbol{\varepsilon}_{[1]}, \dots, \boldsymbol{\varepsilon}_{[n-1]})$ (essentially, the single map $\boldsymbol{\varepsilon}_{[n-1]}$).

  ii. The single choice of the nullhomotopy $F^{-1}_{[n]}$ for $a^{-1}_{[n]}$ (ie the value of the previously defined n^{th} order operation, which must necessarily vanish in order to proceed to the $(n+1)^{\text{st}}$ step).

4.23 Models for rational homotopy theory When working with $R = \mathbb{Q}$ it is convenient to use some of the known models for rational homotopy theory (see Quillen [33]). In particular, finite-type, simply connected rational spaces $\boldsymbol{Y} \in \mathcal{S}_{\mathbb{Q}}$ can be modelled in the category CDGA of differential graded commutative $\mathbb{Q}$–algebras (CDGAs), using a suitable Sullivan model $(A^*, d) \in \mathsf{CDGA}$ for $\boldsymbol{Y}$ (see Félix and Halperin [24, Section 12]).

The equivalence of homotopy categories $\mathrm{ho}\,\mathcal{S}_{\mathbb{Q}} \to \mathrm{ho}\,\mathsf{CDGA}$ is contravariant and takes products to coproducts and path or loop spaces to cone or suspension objects. Thus if we try to apply Theorem 2.33 to the model $(A^*, d) \in \mathsf{CDGA}$ directly, rather than

to $\boldsymbol{Y} \in \mathcal{S}_{\mathbb{Q}}$, we will end up with a *simplicial* CDGA $W_\bullet$, obtained as the homotopy colimit of sequential simplicial realization (see Blanc, Johnson and Turner [14] for full details of the simplicial version). We could in fact replace this simplicial CDGA by a bigraded CDGA (see Félix [23], and compare Blanc [9]).

In particular, we have an adjunction $\Lambda\colon \mathsf{Ch}^*_{\mathbb{Q}} \rightleftharpoons \mathsf{CDGA} : U$ between cochain complexes and CDGAs, where U is the forgetful functor and $\Lambda(V^*, d)$ the free graded commutative algebra on the graded vector space V^*, with $d^i\colon V^i \to V^{i+1}$ extended as a derivation. Note that each $\Lambda(V^*, d)$ is a Sullivan algebra, and thus a cofibrant CDGA (see Hess [27, Section 1]).

The functor Λ yields formal minimal models for each $\mathbb{Q}$–GEM, as well as their cylinders, cones and suspensions. For example, if V^* is a graded $\mathbb{Q}$–vector space which is degreewise finite-dimensional, then $\Lambda(V^*, 0)$ is a minimal model for $\prod_{i=0}^{\infty} \boldsymbol{K}(V^i, i)$. Similarly, if $A^* := \Lambda(V^*, d)$ is a Sullivan model for some space $\boldsymbol{Y}$, and

$$i\colon (V^*, d) \hookrightarrow C(V^*, d)$$

is the inclusion into the cone (see Weibel [38, Section 1.5]), then $i_*\colon \Lambda(V^*, d) \to \Lambda(C(V^*, d))$, the corresponding cone inclusion in CDGA, is a CDGA model for the path fibration $p\colon P\boldsymbol{Y} \to \boldsymbol{Y}$ (see Félix and Halperin [24, Section 14]).

4.24 A rational example Even though the above discussion was stated in terms of the sketch Θ_R in $\mathrm{ho}\,\mathcal{C}$ (for $\mathcal{C} = \mathcal{S}_*$), when $R = \mathbb{Q}$, as we just pointed out, we can also apply it (mutatis mutandis) to the corresponding CDGA models.

For example, let $\boldsymbol{Y}$ be the simply connected, $\mathbb{Q}$–local, finite-type space represented by the free CDGA (A^*, d) with:

i. $A^n = \mathbb{Q}\langle x, y, z\rangle$ with $dx = dy = dz = 0$.

ii. $A^{2n-1} = \mathbb{Q}\langle u, v\rangle$ with $du = xy$ and $dv = xz$.

iii. $A^{3n-2} = \mathbb{Q}\langle q, r, s, t\rangle$ with $dq = xu$, $dr = xv$, $ds = yu$ and $dt = zv$.

iv. A^i for $i > 3n$ chosen so that $H^i(A^*) = 0$ for $i \geq 3n$.

Here $n > 1$ is odd.

Thus (A^*, d) has rational cohomology $\Gamma = H^*(\boldsymbol{Y}; Q)$ (as a $\Theta_{\mathbb{Q}}$–algebra — that is, a graded $\mathbb{Q}$–algebra) with:

i. $\Gamma^n = \mathbb{Q}\langle [x], [y], [z]\rangle$.

ii. $\Gamma^{2n} = \mathbb{Q}\langle [y]\cdot[z]\rangle$.

iii. $\Gamma^{3n-1} = \mathbb{Q}\langle\omega\rangle$, where ω is represented in A^* by $zu + yv$.

iv. $\Gamma^i = 0$ for $i \neq n, 2n, 3n-1$.

Note that the (formal) rational space $\boldsymbol{Z} := (\boldsymbol{S}^n \vee (\boldsymbol{S}^n \times \boldsymbol{S}^n) \vee \boldsymbol{S}^{3n-1})_{\mathbb{Q}}$ also has $H^*(\boldsymbol{Z}; Q) \cong \Gamma$ (as $\Theta_{\mathbb{Q}}$–algebras). It is represented by the Sullivan model (B^*, d) with:

i. $B^n = \mathbb{Q}\langle x, y, z\rangle$ with $dx = dy = dz = 0$.

ii. $B^{2n-1} = \mathbb{Q}\langle u, v\rangle$ with $du = xy$ and $dv = xz$.

iii. $B^{3n-2} = \mathbb{Q}\langle p, q, r, s, t\rangle$ with $dp = zu + yv$, $dq = xu$, $dr = xv$, $ds = yu$, and $dt = zv$.

iv. $B^{3n-1} = \mathbb{Q}\langle w\rangle$ with $dw = 0$, where ω is represented in B^* by w.

v. Again, B^i for $i > 3n$ chosen so that $H^i(B^*) = 0$ for $i \geq 3n$.

Let us denote by $\mathcal{F}_{\mathbb{Q}}\langle x_{n_1}, \dots, x_{n_k}\rangle$ the free $\Theta_{\mathbb{Q}}$–algebra generated by elements x_{n_i} in degree n_i for $i = 1, \dots, k$ — so that $\mathcal{F}_{\mathbb{Q}}\langle x_{n_1}, \dots, x_{n_k}\rangle \cong H^*\big(\prod_{i=1}^k \boldsymbol{K}(\mathbb{Q}, n_i); \mathbb{Q}\big)$ (cf Section 4.23).

We may choose a (minimal) CW resolution of $\Theta_{\mathbb{Q}}$–algebras $V_\bullet \to \Gamma$ with CW basis $(\overline{V}_n)_{n\in\mathbb{N}}$ as follows:

(a) $\overline{V}_0 = \mathcal{F}_{\mathbb{Q}}\langle x_n, y_n, z_n, w_{3n-1}\rangle$, with the obvious augmentation $\varepsilon\colon \overline{V}_0 \to \Gamma$.

(b) $\overline{V}_1 = \mathcal{F}_{\mathbb{Q}}\langle u_{2n}, v_{2n}\rangle \amalg \overline{U}_1$, where $\overline{U}_1$ is a free $\Theta_{\mathbb{Q}}$–algebra with generators in degrees $> 3n$. The attaching map $\overline{\partial}_0\colon \overline{V}_1 \to V_0 = \overline{V}_0$ is defined by $u_{2n} \mapsto x_n y_n$ and $v_{2n} \mapsto x_n z_n$.

(c) $\overline{V}_2 = \mathcal{F}_{\mathbb{Q}}\langle p_{3n}, q_{3n}, r_{3n}, s_{3n}, t_{3n}\rangle \amalg \overline{U}_2$, where $\overline{U}_2$ is again a free $\Theta_{\mathbb{Q}}$–algebra with generators in degrees $> 3n$. The attaching map $\overline{\partial}_0\colon \overline{V}_2 \to V_1 = \overline{V}_1 \amalg s_0\overline{V}_0$ is defined by

(4.25)
$$\begin{gathered} p_{3n} \mapsto (s_0 z_n)u_{2n} + (s_0 y_n)v_{2n}, \quad q_{3n} \mapsto (s_0 x_n)u_{2n}, \\ r_{3n} \mapsto (s_0 x_n)v_{2n}, \quad s_{3n} \mapsto (s_0 y_n)u_{2n}, \quad t_{3n} \mapsto (s_0 z_n)v_{2n}. \end{gathered}$$

(d) For $k \geq 3$ the basis $\Theta_{\mathbb{Q}}$–algebra $\overline{V}_k$ has generators in degrees $> 3n$.

Denote by $\Lambda[\boldsymbol{x}_n] = \Lambda(M^*, 0)$ the formal free CDGA model for $\boldsymbol{K}(\mathbb{Q}, n)$ (where M^* is the graded vector space concentrated in degree n with basis $\{\boldsymbol{x}_n\}$; see Section 4.23).

We can realize $V_\bullet \to \Gamma$ through (co)simplicial dimension 2 and degree $3n$ by an augmented simplicial CDGA $W_\bullet^{[2]} \to B^*$ with CW basis $(\overline{W}_n)_{n\in\mathbb{N}}$ constructed as follows:

Step A First, we construct the simplicial analogue of $W_\bullet^{[1]}$ through simplicial dimension 1:

(a) We let

$$\overline{W}_0 = \Lambda[\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{z}_n, \boldsymbol{w}_{3n-1}] := \Lambda[\boldsymbol{x}_n] \amalg \Lambda[\boldsymbol{y}_n] \amalg \Lambda[\boldsymbol{z}_n] \amalg \Lambda[\boldsymbol{w}_{3n-1}]$$

with augmentation $\boldsymbol{\varepsilon}\colon \overline{W}_0 \to A^*$ defined by

$$\boldsymbol{x}_n \mapsto x, \quad \boldsymbol{y}_n \mapsto y, \quad \boldsymbol{z}_n \mapsto z, \quad \boldsymbol{w}_{3n-1} \mapsto zu + yv.$$

(b) We let

$$\widetilde{W}_1^{[1]} = \overline{W}_1 := \Lambda[\boldsymbol{u}_{2n}, \boldsymbol{v}_{2n}],$$

where the untruncated version of $\overline{W}_1$ has an additional free CDGA coproduct summand D_2 with generators in degrees $> 3n$. The attaching map $\overline{\partial}_0\colon \overline{W}_1 \to W_0^{[2]}$ is defined by $\boldsymbol{u}_{2n} \mapsto \boldsymbol{x}_n\boldsymbol{y}_n$, $\boldsymbol{v}_{2n} \mapsto \boldsymbol{x}_n\boldsymbol{z}_n$, and all other generators sent to 0.

(c) Dually to Step III in the proof of Theorem 2.33, we must add a coproduct summand $C\overline{W}_1$ to obtain $\widetilde{W}_0^{[1]} := \overline{W}_0 \amalg C\overline{W}_1$, where the cone on the formal CDGA $\overline{W}_1 = \Lambda[\boldsymbol{u}_{2n}, \boldsymbol{v}_{2n}]$, which models the path space $P\overline{\boldsymbol{W}}^1$, is the CDGA

$$C\overline{W}_1 = \Lambda(\mathbb{Q}(\boldsymbol{u}'_{2n}, \boldsymbol{v}'_{2n}, \overline{\boldsymbol{u}}_{2n-1}, \overline{\boldsymbol{v}}_{2n-1}), d)$$

with differential

$$d(\overline{\boldsymbol{u}}_{2n-1}) = -\boldsymbol{u}'_{2n} \quad \text{and} \quad d(\overline{\boldsymbol{v}}_{2n-1}) = -\boldsymbol{v}'_{2n}$$

(see Section 4.23). We will denote this simply by $\Lambda(\boldsymbol{u}'_{2n}, \boldsymbol{v}'_{2n}, \overline{\boldsymbol{u}}_{2n-1}, \overline{\boldsymbol{v}}_{2n-1})$.

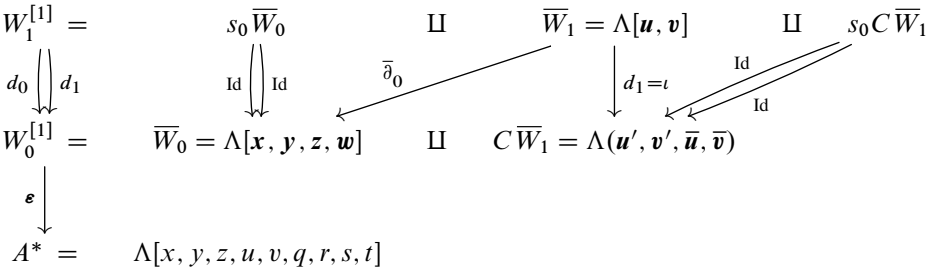
The augmentation $\boldsymbol{\varepsilon}\colon C\overline{W}_1 \to A^*$ is given by

$$\boldsymbol{u}'_{2n} \mapsto xy, \quad \boldsymbol{v}'_{2n} \mapsto xz, \quad \overline{\boldsymbol{u}}_{2n-1} \mapsto u, \quad \overline{\boldsymbol{v}}_{2n-1} \mapsto v.$$

(d) Finally, we must add the degeneracies to $\widetilde{W}_\bullet^{[1]}$ (as in step (iii) of Section 2.5), that is, we add two coproduct summands

$$s_0\overline{W}_0 = \Lambda[s_0\boldsymbol{x}_n, s_0\boldsymbol{y}_n, s_0\boldsymbol{z}_n, s_0\boldsymbol{w}_{3n-1}],$$
$$s_0C\overline{W}_1 = \Lambda(s_0\boldsymbol{u}'_{2n}, s_0\boldsymbol{v}'_{2n}, s_0\overline{\boldsymbol{u}}_{2n-1}, s_0\overline{\boldsymbol{v}}_{2n-1}),$$

to obtain the 1–truncation of the augmented simplicial CDGA $W_\bullet^{[1]}$ (dual to (A.12)) in degrees $\le 3n$ given by



Step B To extend $W_\bullet^{[1]}$ to $W_\bullet^{[2]}$, we proceed as follows:

(a) First, we set

$$\widetilde{W}_2^{[2]} = \overline{W}_2 := \Lambda[\boldsymbol{p}_{3n}, \boldsymbol{q}_{3n}, \boldsymbol{r}_{3n}, \boldsymbol{s}_{3n}, \boldsymbol{t}_{3n}]$$

(where again we omit generators in degrees $> 3n$).

As a first approximation, we would like to use (4.25) to determine $\overline{\partial}_0^{\overline{W}_2}\colon \overline{W}_2 \to W_1^{[2]}$. However, although this will guarantee that $d_0 \circ \overline{\partial}_0^{\overline{W}_2} = 0$, we would then not have $d_1 \circ \overline{\partial}_0^{\overline{W}_2} = 0$ (see Step VII in the proof of Theorem 2.33).

We therefore define $\overline{\partial}_0^{\overline{W}_2}$ by

$$\begin{aligned}
&\boldsymbol{p}_{3n} \mapsto (s_0 z_n)\boldsymbol{u}_{2n} + (s_0 y_n)\boldsymbol{v}_{2n} - s_0(z_n \boldsymbol{u}'_{2n}) - s_0(y_n \boldsymbol{v}'_{2n}),\\
&\boldsymbol{q}_{3n} \mapsto (s_0 x_n)\boldsymbol{u}_{2n} - s_0(x_n \boldsymbol{u}'_{2n}), \quad \boldsymbol{r}_{3n} \mapsto (s_0 x_n)\boldsymbol{v}_{2n} - s_0(x_n \boldsymbol{v}'_{2n}),\\
&\boldsymbol{s}_{3n} \mapsto (s_0 y_n)\boldsymbol{u}_{2n} - s_0(y_n \boldsymbol{u}'_{2n}), \quad \boldsymbol{t}_{3n} \mapsto (s_0 z_n)\boldsymbol{v}_{2n} - s_0(z_n \boldsymbol{v}'_{2n}),
\end{aligned} \tag{4.26}$$

with $d_1\colon \overline{W}_2 \to \widetilde{W}_1^{[2]}$ the inclusion into the new cone summand

$$C\overline{W}_2 = \Lambda(\boldsymbol{p}'_{3n}, \boldsymbol{q}'_{3n}, \boldsymbol{r}'_{3n}, \boldsymbol{s}'_{3n}, \boldsymbol{t}'_{3n}, \overline{\boldsymbol{p}}_{3n-1}, \overline{\boldsymbol{q}}_{3n-1}, \overline{\boldsymbol{r}}_{3n-1}, \overline{\boldsymbol{s}}_{3n-1}, \overline{\boldsymbol{t}}_{3n-1})$$

in $\widetilde{W}_1^{[2]} = W_1^{[1]} \amalg C\overline{W}_2$ (with $d(\overline{\boldsymbol{p}}_{3n-1}) = \boldsymbol{p}'_{3n}$, and so on).

The map $d_0 = F_1\colon C\overline{W}_2 \to W_0^{[1]}$ is given by

$$\begin{aligned}
&\boldsymbol{p}'_{3n} \mapsto -z_n \boldsymbol{u}'_{2n} - y_n \boldsymbol{v}'_{2n}, \quad && \overline{\boldsymbol{p}}_{3n-1} \mapsto z_n \overline{\boldsymbol{u}}_{2n-1} + y_n \overline{\boldsymbol{v}}_{2n-1} - \boldsymbol{w}_{3n-1},\\
&\boldsymbol{q}'_{3n} \mapsto -x_n \boldsymbol{u}'_{2n}, && \overline{\boldsymbol{q}}_{3n-1} \mapsto x_n \overline{\boldsymbol{u}}_{2n-1},\\
&\boldsymbol{r}'_{3n} \mapsto -x_n \boldsymbol{v}'_{2n}, && \overline{\boldsymbol{r}}_{3n-1} \mapsto x_n \overline{\boldsymbol{v}}_{2n-1},\\
&\boldsymbol{s}'_{3n} \mapsto -y_n \boldsymbol{u}'_{2n}, && \overline{\boldsymbol{s}}_{3n-1} \mapsto y_n \overline{\boldsymbol{u}}_{2n-1},\\
&\boldsymbol{t}'_{3n} \mapsto -z_n \boldsymbol{v}'_{2n}, && \overline{\boldsymbol{t}}_{3n-1} \mapsto z_n \overline{\boldsymbol{v}}_{2n-1},
\end{aligned} \tag{4.27}$$

as in (4.25).

(b) In dimension 0 we have $\widetilde{W}_0^{[2]} = W_0^{[1]} \amalg C\Sigma\overline{W}_2$, where

$$C\Sigma\overline{W}_2 = \Lambda(\overline{\boldsymbol{p}}'_{3n-1}, \overline{\boldsymbol{q}}'_{3n-1}, \overline{\boldsymbol{r}}'_{3n-1}, \overline{\boldsymbol{s}}'_{3n-1}, \overline{\boldsymbol{t}}'_{3n-1}, \overline{\overline{\boldsymbol{p}}}_{3n-2}, \overline{\overline{\boldsymbol{q}}}_{3n-2}, \overline{\overline{\boldsymbol{r}}}_{3n-2}, \overline{\overline{\boldsymbol{s}}}_{3n-2}, \overline{\overline{\boldsymbol{t}}}_{3n-2})$$

(with $d(\overline{\overline{\boldsymbol{p}}}_{3n-2}) = \overline{\boldsymbol{p}}'_{3n-1}$, and so on).

The face map

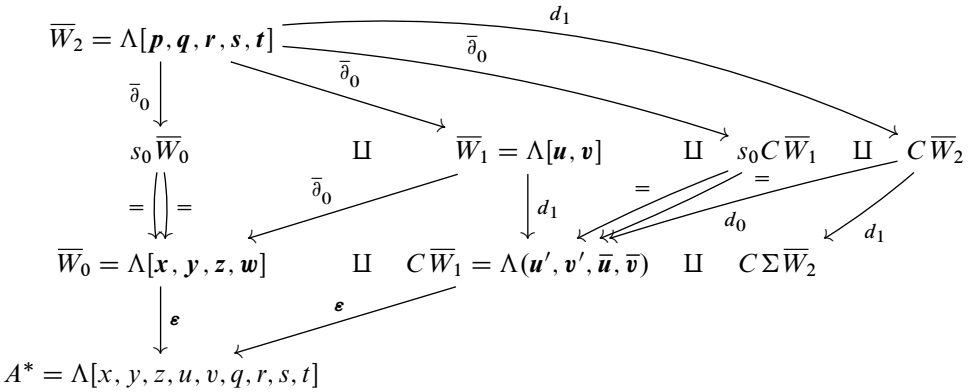
$$d_1\colon \widetilde{W}_1^{[2]} \to \widetilde{W}_0^{[2]}$$

is defined on the new summand $C\overline{W}_2$ to be the quotient $C\overline{W}_2 \twoheadrightarrow \Sigma\overline{W}_2$ followed by the inclusion $\Sigma\overline{W}_2 \hookrightarrow C\Sigma\overline{W}_2$, which is given by $\boldsymbol{p}_{3n} \mapsto 0$, $\overline{\boldsymbol{p}}_{3n-1} \mapsto -\overline{\boldsymbol{p}}'_{3n-1}$, and so on.

The augmentation $\boldsymbol{\varepsilon}\colon C\Sigma\overline{W}_2 \to A^*$ is given by

$$\textbf{(4.28)}\qquad \begin{aligned} \overline{\boldsymbol{p}}'_{3n-1} &\mapsto 0, & \overline{\overline{\boldsymbol{p}}}_{3n-2} &\mapsto 0,\\ \overline{\boldsymbol{q}}'_{3n-1} &\mapsto x_n u_{2n-1}, & \overline{\overline{\boldsymbol{q}}}_{3n-2} &\mapsto -q_{3n-2},\\ \overline{\boldsymbol{r}}'_{3n-1} &\mapsto x_n v_{2n-1}, & \overline{\overline{\boldsymbol{r}}}_{3n-2} &\mapsto -r_{3n-2},\\ \overline{\boldsymbol{s}}'_{3n-1} &\mapsto y_n u_{2n-1}, & \overline{\overline{\boldsymbol{s}}}_{3n-2} &\mapsto -s_{3n-2},\\ \overline{\boldsymbol{t}}'_{3n-1} &\mapsto z_n v_{2n-1}, & \overline{\overline{\boldsymbol{t}}}_{3n-2} &\mapsto -t_{3n-2}. \end{aligned}$$

The 2–truncation of $\widetilde{W}_\bullet^{[2]}$ in degrees $\le 3n$ is given by



Step C When we try to map $W_\bullet^{[2]}$ to the CDGA B^*, we must modify the augmentation $\boldsymbol{\varepsilon}\colon \widetilde{W}_0^{[2]} \to A^*$ as follows:

In order to realize $\varepsilon\colon V_0 \to \Gamma$ we must have $\boldsymbol{\varepsilon}(\boldsymbol{w}_{3n-1}) = w$ and otherwise $\boldsymbol{\varepsilon}$ is the same as in Step A. Therefore,

$$\boldsymbol{\varepsilon}(\overline{\boldsymbol{p}}'_{3n-1}) = \boldsymbol{\varepsilon}(d_1(\overline{\boldsymbol{p}}_{3n-1})) = \boldsymbol{\varepsilon}(\boldsymbol{z}_n\overline{\boldsymbol{u}}_{2n-1} + \boldsymbol{y}_n\overline{\boldsymbol{v}}_{2n-1} - \boldsymbol{w}_{3n-1}) = zu + yv - w,$$

by (4.27). Therefore, we must map $\bar{\bar{\boldsymbol{p}}}_{3n-2}$ to an element a in B^{3n-2} with $d(a) = zu + yv - w$. Since $d(p) = zu + yv$ but w represents a nonzero element in $\Gamma = H^{3n-1}(\boldsymbol{Z}; \mathbb{Q})$, no such a exists. Thus the two rational spaces $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are not weakly equivalent by Corollary 4.19.

Intuitively, the element $\omega \in \Gamma^{3n-1}$ is represented by the Massey product $[zu + yv] = \langle [y], [x], [z] \rangle$ in $H^{3n-1}(\boldsymbol{Y}; \mathbb{Q})$, while ω is not a Massey product in $H^{3n-1}(\boldsymbol{Z}; \mathbb{Q})$ since $\boldsymbol{Z}$ is formal. In our language these facts are represented by the nullhomotopies $\bar{\boldsymbol{p}}_{3n-1} \mapsto \boldsymbol{z}_n \bar{\boldsymbol{u}}_{2n-1} + \boldsymbol{y}_n \bar{\boldsymbol{v}}_{2n-1} - \boldsymbol{w}_{3n-1}$ in $\boldsymbol{Y} \to \boldsymbol{W}^\bullet$, where in the analogous construction for $\boldsymbol{Z}$ we would have had $\bar{\boldsymbol{p}}_{3n-1} \mapsto \boldsymbol{z}_n \bar{\boldsymbol{u}}_{2n-1} + \boldsymbol{y}_n \bar{\boldsymbol{v}}_{2n-1}$.

5 Higher cohomology invariants for maps

The system of higher cohomology operations associated to a Θ_R–algebra $H^*(\boldsymbol{Y}; R)$ described in the previous section may be thought of as a sequence of obstructions to realizing an algebraic isomorphism $\vartheta\colon H^*(\boldsymbol{Y}; R) \xrightarrow{\cong} H^*(\boldsymbol{Z}; R)$ by a map $f\colon \boldsymbol{Z} \to \boldsymbol{Y}$ (necessarily an R–equivalence) — as well as constituting a complete collection of higher invariants for the weak R–homotopy type of spaces.

In this section we address the analogous problem for arbitrary maps: given two maps $f_0, f_1\colon \boldsymbol{Z} \to \boldsymbol{Y}$ which induce the same morphism of Θ_R–algebras $\psi\colon H^*(\boldsymbol{Y}; R) \to H^*(\boldsymbol{Z}; R)$, we define a sequence of higher cohomology operations which vanish coherently if and only if f_0 and f_1 are R–equivalent.

5.1 Obstructions for lifting homotopies We start with the *initial data* $(\mathcal{W}, f_0, f_1)$, consisting of a sequential realization $\mathcal{W}$ for $\boldsymbol{Y}$, and two maps f_0 and f_1 as above. Since f_0 and f_1 induce the same map of Θ_R–algebras, there is a homotopy $H^0_{[0]}\colon \boldsymbol{Z} \otimes I \to \overline{\boldsymbol{W}}^0 = \boldsymbol{W}^0_{[0]}$ between $\boldsymbol{\varepsilon}_{[0]} \circ f_0$ and $\boldsymbol{\varepsilon}_{[0]} \circ f_1$, so

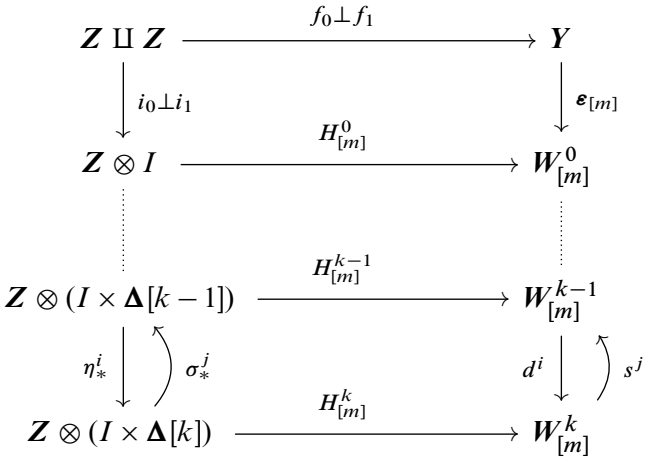
$$\boldsymbol{\varepsilon}_{[0]} \circ (f_0 \perp f_1) = H^0_{[0]} \circ (i_0 \perp i_1). \tag{5.2}$$

We call $\mathcal{H}_{[0]} = (H^0_{[0]})$ a 0–*strand* for $(\mathcal{W}, f_0, f_1)$.

Recall that the standard cosimplicial space $\boldsymbol{\Delta}^\bullet$ is given by the diagram of n–simplices with the standard maps between them, where $\eta^i\colon \boldsymbol{\Delta}[k-1] \hookrightarrow \boldsymbol{\Delta}[k]$ is the inclusion of the ith face, and $\sigma^j\colon \boldsymbol{\Delta}[k] \twoheadrightarrow \boldsymbol{\Delta}[k-1]$ is the jth collapse map (see Bousfield and Kan [20, Section X.2.2]). Applying the simplicial structure operation $- \otimes \boldsymbol{\Delta}^\bullet$ to a fixed space $\boldsymbol{Z}$ yields a cosimplicial space $\boldsymbol{Z} \otimes \boldsymbol{\Delta}^\bullet$.

5.3 Definition An *n–strand* $\mathcal{H}_{[n]} = (H_{[0]}, \dots, H_{[m]}, \dots, H_{[n]})$ for $(\mathcal{W}, f_0, f_1)$ is a compatible sequence of maps of cosimplicial spaces $H_{[m]}\colon \boldsymbol{Z} \otimes I \otimes \boldsymbol{\Delta}^\bullet \to \boldsymbol{W}^\bullet_{[m]}$ for $m = 0, \dots, n$, each determined by the collection of k–homotopies $H^k_{[m]}\colon \boldsymbol{Z} \otimes (I \times \boldsymbol{\Delta}[k]) \to \boldsymbol{W}^k_{[m]}$ for $0 \le k \le m$ such that all the downward and upward squares in the diagram

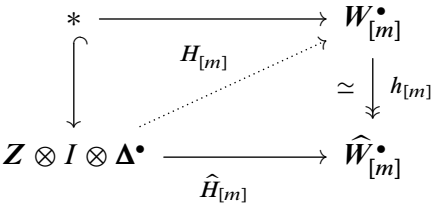
(5.4)



commute for all choices of $0 \le i \le k \le m$ and $0 \le j \le k-1$.

More precisely, the choices of $H^k_{[m]}$ for $0 \le k \le m$ uniquely determine a map of cosimplicial spaces $\widehat{H}_{[m]}\colon \boldsymbol{Z} \otimes I \otimes \boldsymbol{\Delta}^\bullet \to \widehat{\boldsymbol{W}}^\bullet_{[m]}$, since the target is m–coskeletal. We then use the left lifting property for the solid commuting square of cosimplicial spaces

(5.5)



to obtain the required map $H_{[m]}$, unique up to weak equivalence, using the fact that $\boldsymbol{Z} \otimes I \otimes \boldsymbol{\Delta}^\bullet$ is cofibrant and the map $h_{[m]}\colon \boldsymbol{W}^\bullet_{[m]} \to \widehat{\boldsymbol{W}}^\bullet_{[m]}$ is a trivial fibration by Remark 2.18.

We say that an n–strand $\mathcal{H}_{[n]}$ *extends* a given $(n-1)$–strand $\mathcal{H}_{[n-1]}$ if

$$H_{[n-1]} = \pi_{[n]} \circ H_{[n]}\colon \boldsymbol{Z} \otimes (I \times \boldsymbol{\Delta}^\bullet) \to \boldsymbol{W}^\bullet_{[n-1]} \tag{5.6}$$

(cf (2.2)).

An *∞–strand* is a sequence $\mathcal{H}_{[\infty]} := (H_{[n]})_{n=1}^\infty$ satisfying (5.6) for each $n > 0$.

5.7 Remark In order to extend a given $(n-1)$–strand $\mathcal{H}_{[n-1]}$ for a sequential realization $\mathcal{W}$ of $\boldsymbol{Y}$ to an n–strand, we need to produce maps $\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to \overline{P\Omega^{n-k-1}\boldsymbol{W}^n}$ for $0\le k\le n$ satisfying

$$\textbf{(5.8)}\qquad \begin{cases} \widehat{H}^k_{[n]}\circ\eta^0_* = F^{k-1}\circ v^{k-1}\circ H^{k-1}_{[n-1]},\\ \widehat{H}^k_{[n]}\circ\eta^1_* = \delta^{k-2}\circ\widehat{H}^{k-1}_{[n]},\\ \widehat{H}^k_{[n]}\circ\eta^i_* = 0 \quad \text{for } i\ge 2, \end{cases}$$

where $\delta^j=\delta^j_{\boldsymbol{D}}$ is the differential of $\boldsymbol{D}^*$ as in (2.8).

Thus, in the following diagram we are given the solid $(n-1)$–strand for $\boldsymbol{W}^\bullet_{[n-1]}$, which we wish to extend by the dashed maps to the (given) restricted cosimplicial space $\widetilde{\boldsymbol{W}}^\bullet_{[n]}$:

$$\textbf{(5.9)}\qquad \begin{array}{ccccc}
\boldsymbol{Z}\amalg\boldsymbol{Z} & \xrightarrow{(f_0\perp f_1)} & & & \boldsymbol{Y}\\
\downarrow{\scriptstyle(i_0\perp i_1)} & \overset{H^0_{[n-1]}}{\nearrow} & & \overset{F^{-1}}{\searrow} & \downarrow{\scriptstyle\varepsilon}\\
\boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[0]) & \overset{\widehat{H}^0_{[n]}}{\dashrightarrow} & \overline{P\Omega^{n-1}\boldsymbol{W}^n} & \times & \boldsymbol{W}^0_{[n-1]} = \widetilde{\boldsymbol{W}}^0_{[n]}\\
\downarrow{\scriptstyle\eta^0_*}\downarrow{\scriptstyle\eta^1_*} & \overset{H^1_{[n-1]}}{\nearrow} & \downarrow{\scriptstyle\delta^0F^0\circ v^0} & & \downarrow{\scriptstyle d^0}\downarrow{\scriptstyle d^1}\\
\boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[1]) & \overset{\widehat{H}^1_{[n]}}{\dashrightarrow} & \overline{P\Omega^{n-2}\boldsymbol{W}^n} & \times & \boldsymbol{W}^1_{[n-1]} = \widetilde{\boldsymbol{W}}^1_{[n]}\\
\downarrow{\scriptstyle\eta^0_*}\downarrow{\scriptstyle\eta^1_*}\downarrow{\scriptstyle\eta^2_*} & \overset{H^2_{[n-1]}}{\nearrow} & \downarrow{\scriptstyle\delta^0F^1\circ v^1} & & \downarrow{\scriptstyle d^0}\downarrow{\scriptstyle d^1}\downarrow{\scriptstyle d^2}\\
\boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[2]) & \overset{\widehat{H}^2_{[n]}}{\dashrightarrow} & \overline{P\Omega^{n-3}\boldsymbol{W}^n} & \times & \boldsymbol{W}^2_{[n-1]} = \widetilde{\boldsymbol{W}}^2_{[n]}\\
\vdots & \overset{H^{n-1}_{[n-1]}}{\nearrow} & \vdots & & \vdots\\
\boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[n-1]) & \overset{\widehat{H}^{n-1}_{[n]}}{\dashrightarrow} & \overline{P\Omega^0\boldsymbol{W}^n} & \times & \boldsymbol{W}^{n-1}_{[n-1]} = \widetilde{\boldsymbol{W}}^{n-1}_{[n]}\\
\downarrow{\scriptstyle\eta^0_*}\ \cdots\ \downarrow{\scriptstyle\eta^n_*} & & \downarrow{\scriptstyle p} & \overset{\delta^{n-2}}{\searrow} & \\
 & \overset{\widehat{H}^n_{[n]}}{\dashrightarrow} & & \overset{\overline{d}^0_n}{\searrow} & \\
\boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[n]) & & \overline{\boldsymbol{W}}^n & &
\end{array}$$

5.10 Folding polytopes Consider the iterated trivial fibration of (2.9),

$$\textbf{(5.11)}\qquad \xi^j := P\Omega^{j-1}\sigma^1\circ\cdots\circ P\Omega\sigma^{j-1}\circ P\sigma^j\circ P\tau^j\colon \overline{P\Omega^j\boldsymbol{W}^n}\twoheadrightarrow P\Omega^j\overline{\boldsymbol{W}}^n.$$

If we identify the k–simplex $\boldsymbol{\Delta}[k]$ with (a quotient of) the k–cube I^k, each map

$$\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to\overline{P\Omega^{n-k-1}\boldsymbol{W}^n},$$

after postcomposing with $\xi^{n-k-1}\colon \overline{P\Omega^{n-k-1}\boldsymbol{W}^n} \to P\Omega^{n-k-1}\overline{\boldsymbol{W}}^n$, can be identified by adjunction with a pointed map $\widetilde{H}^k\colon (\boldsymbol{Z}\otimes I)\otimes I^n \to \overline{\boldsymbol{W}}^n$ taking certain facets of I^n to the basepoint.

Moreover, the compatibility conditions of (5.8) translate into requirements that the restrictions of the maps $\widetilde{H}^k$ to certain facets of I^n match up in an appropriate way. This information can be encoded by gluing together $n+1$ n–cubes (corresponding to cosimplicial dimensions $0,1,\dots,n$) along their facets to obtain a single n–dimensional cubical complex, as follows:

5.12 Definition The barycentric subdivision, as a triangulation of the standard n–simplex $\boldsymbol{\Delta}[n]$, exhibits it as a PL-cone on its boundary $\partial\boldsymbol{\Delta}[n]$. We may similarly define by induction a triangulation of the standard n–cube $I^n=[0,1]^n$ obtained as the cone on its boundary ∂I^n (more precisely, the join of the barycenter of I^n with the inductively defined triangulation of ∂I^n).

This allows us to define PL-homeomorphisms $\zeta^n\colon I^n\to\boldsymbol{\Delta}[n]$, starting with the obvious isomorphism for $n=1$, taking boundary to boundary, and extending to the interior by applying the join with the respective barycenters.

For each $n\geq 2$, we consider the corner C of ∂I^n consisting of all $(n-1)$–facets $(E_k)_{k=0}^{n-1}$ incident with the fixed vertex $v=(0,\dots,0)$, where

$$E_k=\{(t_1,\dots,t_n)\in I^n : t_{k+1}=0\}.$$

We use the previously defined ζ^{n-1} to identify $E_k\cong I^{n-1}$ with the kth face $\boldsymbol{\Delta}_k[n]$ of $\boldsymbol{\Delta}[n]$. For the complementary corner C' (incident with the vertex $v'=(1,\dots,1)$ opposite v), we use the orthogonal projection from the last vertex of $\boldsymbol{\Delta}[n]$ onto the face $\boldsymbol{\Delta}_n[n]$ opposite it to obtain a subdivision of $\boldsymbol{\Delta}_n[n]$ into n $(n-1)$–dimensional simplices, which we identify with the n $(n-1)$–dimensional facets of $\mathbb{C}'$ using $(\zeta^{n-1})^{-1}$.

Now, for each $n\geq 1$ consider $n+1$ standard n–cubes $I^n_{(0)},\dots,I^n_{(n)}$, where we have a PL-isomorphism ζ^n_k given as follows:

$$I^n_{(k)}\cong I^k\times I^{n-k}\xrightarrow{\zeta^k\times\mathrm{Id}}\boldsymbol{\Delta}[k]\times I^{n-k}.$$

For $k<n$ we think of the first direction of I^{n-k} as the *path* direction, and the remaining $n-k-1$ directions as *loop* directions. This allows us to represent a pointed map $h\colon \widehat{\boldsymbol{Z}}\otimes\boldsymbol{\Delta}[k]\to P\Omega^{n-k-1}\overline{\boldsymbol{W}}$ by a map $\widehat{h}\colon \widehat{\boldsymbol{Z}}\otimes I^n_{(k)}\to\overline{\boldsymbol{W}}$ in a canonical way (sending certain facets of $I^n_{(k)}$ to the basepoint).

For any $1 \le k \le n$ we have two $(k-1)$–faces $\mathbf{\Delta}_0[k]$ and $\mathbf{\Delta}_1[k]$ of $\mathbf{\Delta}[k]$, and the $(n-1)$–dimensional prisms $\mathbf{\Delta}_0[k] \times I^{n-k}$ and $\mathbf{\Delta}_1[k] \times I^{n-k}$ are identified under the map $(\zeta_k^n)^{-1}$ with two $(n-1)$–dimensional facets of $I^n_{(k)}$, which we denote by B_0^k and B_1^k, respectively. By our convention, if $0 \le k < n$ we have another special $(n-1)$–dimensional facet of $I^n_{(k)}$, denoted by

$$C^k := \{(t_1, \dots, t_n) \in I^k \times I^{n-k} \mid t_{k+1} = 0\}$$

(the zero-face in the "path direction").

We now define the nth *folding polytope* $\mathcal{P}^n$ for each $n \ge 2$ by taking the disjoint union of the $n+1$ n–cubes $I^n_{(0)}, \dots, I^n_{(n)}$, and identifying B_1^k with C^{k-1} for each $1 \le k \le n$.

5.13 Lemma *For each $n \ge 2$, the folding polytope $\mathcal{P}^n$ is homeomorphic to an n–ball, with boundary $\partial\mathcal{P}^n$ homeomorphic to an $(n-1)$–sphere.*

5.14 Remark All the faces $(B_1^k)_{k=1}^n$ and $(C^k)_{k=0}^{n-1}$ are now interior to $\mathcal{P}^n$, while the remaining facets of the cubes $I^n_{(k)}$, including $(B_0^k)_{k=1}^n$, constitute the boundary $\partial\mathcal{P}^n$.

5.15 Example The four constituent 3–cubes of $\mathcal{P}^3$ are illustrated in Figure 1, with the dotted arrows indicating glued faces. Note that the two faces B_0^k and B_1^k are adjacent for $2 \le k \le n$, while B_0^1 and B_1^1 are opposite each other (since the same is true of $\mathbf{\Delta}_0[k]$ and $\mathbf{\Delta}_1[k]$ in $\mathbf{\Delta}[k]$). On the other hand, C^k is always adjacent to both B_0^k and B_1^k.

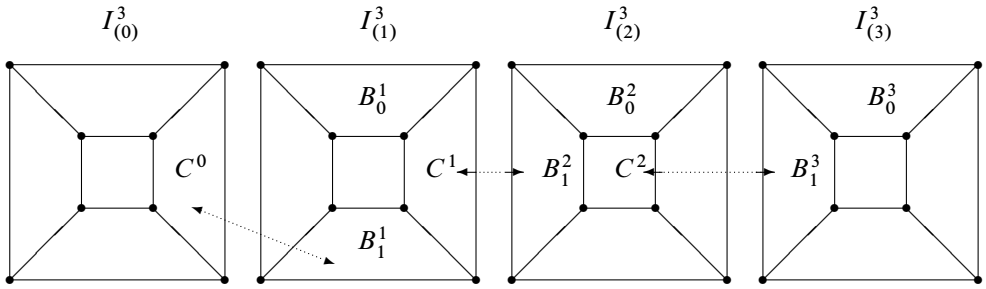


Figure 1: The four 3–cubes of $\mathcal{P}^3$

5.16 Lemma *Given two maps $f_0, f_1 \colon \boldsymbol{Z} \to \boldsymbol{Y}$ which induce the same algebraic homomorphism of Θ_R–algebras $H^*(\boldsymbol{Y}; R) \to H^*(\boldsymbol{Z}; R)$ and a sequential realization*

$\mathcal{W}$ for $\boldsymbol{Y}$, let $H_{[n-1]}\colon \boldsymbol{Z}\otimes I\otimes\boldsymbol{\Delta}^\bullet\to \boldsymbol{W}^\bullet_{[n-1]}$ be an $(n-1)$–strand as in Section 5.1. Then there is a one-to-one correspondence between collections of maps

$$\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to\overline{P\Omega^{n-k-1}\boldsymbol{W}^n}\quad \text{for } 0\le k\le n$$

as in Remark 5.7 (satisfying (5.8)) *and maps $h\colon (\boldsymbol{Z}\otimes I)\otimes\mathcal{P}^n\to\overline{\boldsymbol{W}}^n$ such that*

$$\textbf{(5.17)}\qquad h|_{(\boldsymbol{Z}\otimes I)\otimes B^k_0}=\xi^{n-k-1}\circ F^{k-1}\circ v^{k-1}\circ H^{k-1}_{[n-1]}$$
$$\text{for } 1\le k\le n \text{ and } h|_{(\boldsymbol{Z}\otimes I)\otimes E}=*,$$

where $E:=\partial\mathcal{P}^n\setminus\bigcup_{k=1}^n B^k_0$.

Proof Given a map $\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes I\otimes\boldsymbol{\Delta}[k]\to\overline{P\Omega^{n-k-1}\boldsymbol{W}^n}$, we obtain a map

$$\widetilde{H}^k_{(n)}\colon (\boldsymbol{Z}\otimes I)\otimes I^n\to\overline{\boldsymbol{W}}^n,$$

where we identify I^k with $\boldsymbol{\Delta}[k]$ using ζ^k and take the $(k+1)^{\text{st}}$ coordinate for the path direction and the remaining $n-k-1$ coordinates for the loop directions, as in Definition 5.12.

The first condition in (5.8) says that on B^k_0 (corresponding to the 0–face of $\boldsymbol{\Delta}[k]$), $\widehat{H}^k_{[n]}$ equals $F^{k-1}\circ v^{k-1}\circ H^{k-1}_{[n-1]}$. The second condition there says that on B^k_1 (corresponding to the 1–face of $\boldsymbol{\Delta}[k]$), $\widehat{H}^k_{[n]}$ equals $\delta^{n-k}\circ\widehat{H}^{k-1}_{[n]}$ (where δ^{n-k} is defined in (2.8)), which stated in terms of cubes means that it coincides with $\widehat{H}^{k-1}_{[n-1]}$ restricted to C^{k-1}. Since the coface maps d^i into $P\Omega^{n-k-1}\overline{\boldsymbol{W}}^n$ vanish for $i\ge 2$, and $\widehat{H}^{k-1}_{[n]}$ also vanishes at the other end of the path direction, and at both ends of the loop directions, we obtain the description above.

Conversely, given such a map $\widetilde{H}$, we use its restrictions to the n–cubes $I^0_{(n)},\dots,I^n_{(n)}$ to define the maps $\widetilde{H}^k$, and thus maps ${}'\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to P\Omega^{n-k-1}\overline{\boldsymbol{W}}^n$ satisfying

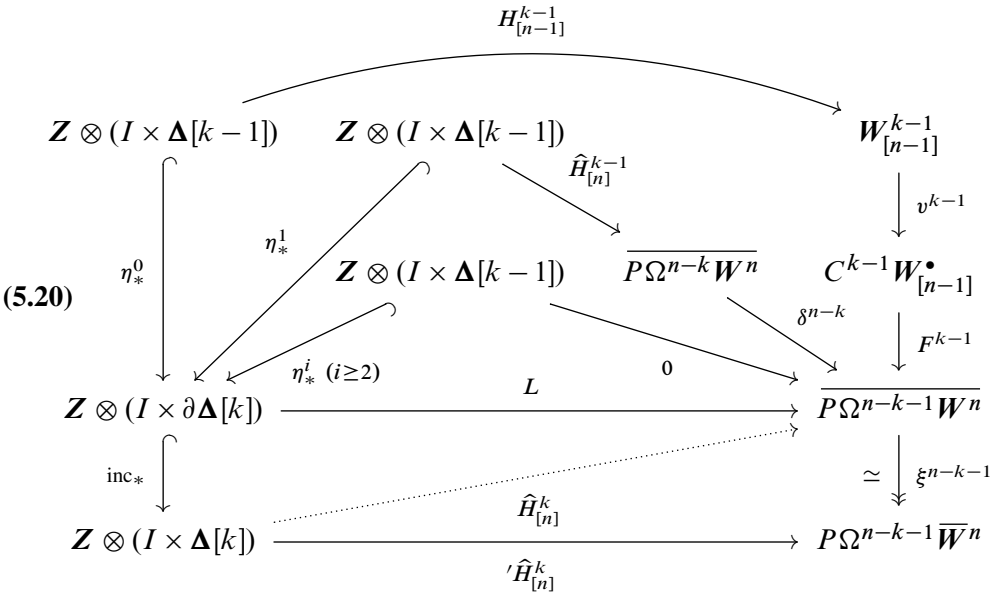
$$\textbf{(5.18)}\qquad \begin{cases} {}'\widehat{H}^k_{[n]}\circ\eta^0_*=\xi^{n-k-1}\circ F^{k-1}\circ v^{k-1}\circ H^{k-1}_{[n-1]},\\ {}'\widehat{H}^k_{[n]}\circ\eta^1_*=\xi^{n-k-1}\circ\delta^{n-k}\circ{}'\widehat{H}^{k-1}_{[n]},\\ {}'\widehat{H}^k_{[n]}\circ\eta^i_*=0\quad \text{for } i\ge 2,\end{cases}$$

for ξ^j as in (5.11). We now show by induction on $0\le k$ that these lift to maps $\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to\overline{P\Omega^{n-k-1}\boldsymbol{W}^n}$ satisfying (5.8), and

$$\textbf{(5.19)}\qquad {}'\widehat{H}^k_{[n]}=\xi^{n-k-1}\circ\widehat{H}^k_{[n]}.$$

Indeed, the inductively defined lift $\widehat{H}^{k-1}_{[n]}$ induces a map $L\colon \boldsymbol{Z}\otimes(I\times\partial\boldsymbol{\Delta}[k])\to\overline{P\Omega^{n-k-1}\boldsymbol{W}^n}$ fitting into the solid commutative diagram

(5.20)



where the upper squares fit together to define L by induction, using (5.8), and the bottom solid square then commutes by (5.18) and (5.19).

Since $\boldsymbol{Z}$ is cofibrant in $\mathcal{C}$ by Assumption 1.20, the map inc_* is a cofibration; see [32, Section II.2]. Moreover ξ^{n-k-1} is a trivial fibration, so we have the lifting $\widehat{H}^k_{[n]}$ by the LLP. The fact that (5.20) commutes implies that (5.8) holds for k, too. To start the induction for $k=0$, we just need the fact that ξ^{n-1} is a trivial fibration and $Z\otimes I$ is cofibrant with $L=0$, since (5.8) is then vacuous. □

5.21 Definition Assume given initial data $(\mathcal{W}, f_0, f_1\colon \boldsymbol{Z}\to\boldsymbol{Y})$ with a corresponding $(n-1)$–strand $\mathcal{H}_{[n-1]}=(H_{[m]})_{m=0}^{n-1}$, as in Section 5.1. We associate to this a map $g\colon(\boldsymbol{Z}\otimes I)\otimes\partial\mathcal{P}^n\to\overline{\boldsymbol{W}}^n$ which sends $(\boldsymbol{Z}\otimes I)\otimes B^k_0$ to $\overline{\boldsymbol{W}}^n$ by $F^k\circ H^k_{[n]}$ for each $1\leq k\leq n$, and all other $(n-1)$–cubes of $\partial\mathcal{P}^n$ to the basepoint. Here we use the convention of (2.7), so $F^n=\bar{d}^0_n$.

Since at most two additional $(n-1)$–facets of $I^n_{(k)}$ are identified with $(n-1)$–facets of $I^n_{(k\pm1)}$, we may think of $P\Omega^{n-k}\overline{\boldsymbol{W}}^n$ as contained in $\mathrm{map}_*(d^0_1 I^n_{(k)},\overline{\boldsymbol{W}}^n)$, so the map induced by $H^{k-1}_{[n]}\circ F^k$ is well-defined. Moreover, these maps are compatible for adjacent values of k by (5.8).

By Lemma 5.13 we can think of g as a map $(\boldsymbol{Z}\otimes I)\otimes\boldsymbol{S}^{n-1}\to\overline{\boldsymbol{W}}^n$, and because all maps are pointed, this actually factors through the half-smash

$$(\boldsymbol{Z}\otimes I)\ltimes\boldsymbol{S}^{n-1}:=((\boldsymbol{Z}\otimes I)\times\boldsymbol{S}^{n-1})/(*\times\boldsymbol{S}^{n-1}),$$

which can be canonically identified with $\Sigma^{n-1}(\boldsymbol{Z}\otimes I)\vee(\boldsymbol{Z}\otimes I)$ (see [4]). Moreover, the map $(\boldsymbol{Z}\otimes I)\to\overline{\boldsymbol{W}}^n$ in question is nullhomotopic for $n\geq 1$, so we may restrict attention to the factor $g'\colon \Sigma^{n-1}(\boldsymbol{Z}\otimes I)\to\overline{\boldsymbol{W}}^n$, and define the *value* of the $(n-1)$–strand $\mathcal{H}_{[n-1]}$ to be the class

$$\textbf{(5.22)}\qquad \mathrm{Val}(\mathcal{H}_{[n-1]}) := [g'] \in [\Sigma^{n-1}(\boldsymbol{Z}\otimes I), \overline{\boldsymbol{W}}^n] \cong [\boldsymbol{Z}, \Omega^{n-1}\overline{\boldsymbol{W}}^n],$$

so it consists of a set of cohomology classes for $\boldsymbol{Z}$.

5.23 Proposition *Under the assumptions of Lemma 5.16,* $\mathrm{Val}(\mathcal{H}_{[n-1]})=0$ *if and only if* $\mathcal{H}_{[n-1]}$ *extends to an* n*–strand* $\mathcal{H}_{[n]}$.

Proof The $(n-1)$–strand $\mathcal{H}_{[n-1]}$ extends to an n–strand $\mathcal{H}_{[n]}$ if and only if we have a collection of maps

$$\widehat{H}^k_{[n]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to\overline{P\Omega^{n-k-1}\boldsymbol{W}^n}$$

for $0\leq k\leq n$ as in Remark 5.7 satisfying (5.8), and by Lemma 5.16 this corresponds to a map $(\boldsymbol{Z}\otimes I)\otimes\mathcal{P}^n\to\overline{\boldsymbol{W}}^n$ whose restriction to $(\boldsymbol{Z}\otimes I)\otimes\partial\mathcal{P}^n$ is the map g determined by $\mathcal{H}_{[n-1]}$ as in Definition 5.21. The map g extends to $(\boldsymbol{Z}\otimes I)\otimes\mathcal{P}^n$ if and only if g' is nullhomotopic. □

5.24 Correspondence of strands for maps Given $f_0, f_1\colon \boldsymbol{Z}\to\boldsymbol{Y}$ with $f_0^* = f_1^*\colon H^*(\boldsymbol{Y};R)\to H^*(\boldsymbol{Z};R)$, an n–stage comparison map $\Phi\colon\mathcal{W}\to{}'\mathcal{W}$ between two sequential realizations for $\boldsymbol{Y}$ as in (3.16), and n–strands $\mathcal{H}_{[n]}$ and ${}'\mathcal{H}_{[n]}$ for $\mathcal{W}$ and ${}'\mathcal{W}$, respectively, we write

$${}'\mathcal{H}_{[n]} = r^{\#}(\mathcal{H}_{[n]})$$

if $(H^k_{[m]})' = r^k_{[m]}\circ H^k_{[m]}\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to{}'\boldsymbol{W}^{\,k}_{[m]}$ and

$$\mathcal{H}_{[n]} = e^{\#}({}'\mathcal{H}_{[n]})$$

if $H^k_{[m]} = e^k_{[m]}\circ(H^k_{[m]})'\colon \boldsymbol{Z}\otimes(I\times\boldsymbol{\Delta}[k])\to\boldsymbol{W}^k_{[m]}$ for each $0\leq k\leq m\leq n$ (compare Section 4.6).

By comparing (3.13) and (3.15) with (5.9) and Definition 5.21, we see that

$$\textbf{(5.25)}\qquad \mathrm{Val}(r^{\#}(\mathcal{H}_{[n]})) = \bar{r}^n_*(\mathrm{Val}(\mathcal{H}_{[n]}))\quad\text{and}\quad \mathrm{Val}(e^{\#}({}'\mathcal{H}_{[n]})) = \bar{e}^n_*(\mathrm{Val}({}'\mathcal{H}_{[n]})),$$

as in (4.7), so

5.26 (a) $\mathrm{Val}(\mathcal{H}_{[n]})=0$ if and only if $\mathrm{Val}(r^{\#}(\mathcal{H}_{[n]}))=0$,

(b) if $\mathrm{Val}({}'\mathcal{H}_{[n]})=0$ then $\mathrm{Val}(e^{\#}({}'\mathcal{H}_{[n]}))=0$, as in (4.8).

We define weak and strong equivalences relations on strands as in Definition 4.9.

5.27 Definition Given two maps $f_0, f_1\colon \boldsymbol{Z} \to \boldsymbol{Y}$ inducing the same homomorphism of Θ_R–algebras $\phi\colon H^*(\boldsymbol{Y}; R) \to H^*(\boldsymbol{Z}; R)$, the associated *universal n^{th} order cohomology operation* $\langle\!\langle f_{(0,1)}\rangle\!\rangle_n$ which assigns to an $(n-1)$–strand $\mathcal{H}_{[n-1]}$ for this data the value

$$\langle\!\langle f_{(0,1)}\rangle\!\rangle_n(\mathcal{H}_{[n-1]}) := \mathrm{Val}(\mathcal{H}_{[n-1]}) \in \Gamma'\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\},$$

where $\Gamma' := H^*(\boldsymbol{Z}; R)$. We say that $\langle\!\langle f_{(0,1)}\rangle\!\rangle_n$ *vanishes* if there is a cofibrant $\mathcal{W}$ with an $(n-1)$–strand $\mathcal{H}_{[n-1]}$ for $(\mathcal{W}, f_0, f_1)$ such that $\langle\!\langle f_{(0,1)}\rangle\!\rangle_n(\mathcal{H}_{[n-1]}) = 0$. Note that this depends only on the strong equivalence class of $\mathcal{H}_{[n-1]}$.

As in Definition 4.14, we then say that the system $\langle\!\langle f_{(0,1)}\rangle\!\rangle = (\langle\!\langle f_{(0,1)}\rangle\!\rangle_n)_{n=2}^{\infty}$ of nth order cohomology operations for (f_0, f_1) *vanishes coherently* for $(\mathcal{W}, f_0, f_1)$ if there is an ∞–strand $\mathcal{H}_{[\infty]}$ for this data — that is, for each $n \geq 1$, we have an n–strand $\mathcal{H}_{[n]}$ for $(\mathcal{W}, f_0, f_1)$ such that $\mathrm{Val}(\mathcal{H}_{[n]} = 0)$, which extends to the $(n+1)$–strand $\mathcal{H}_{[n+1]}$ using Proposition 5.23.

The proof of Key Lemma 4.17 shows also:

5.28 Lemma *Given $f_0, f_1\colon \boldsymbol{Z} \to \boldsymbol{Y}$ as above, $\langle\!\langle f_{(0,1)}\rangle\!\rangle_n$ vanishes if and only if, for* every *n–stage cofibrant sequential realization $\mathcal{W}$ of $\boldsymbol{Y}$, there is an $(n-1)$–strand $\mathcal{H}_{[n-1]}$ for $(\mathcal{W}, f_0, f_1)$ with $\mathrm{Val}(\mathcal{H}_{[n-1]}) = 0$.*

Moreover, if $\langle\!\langle f_{(0,1)}\rangle\!\rangle_n$ vanishes at the $(n-1)$–strand $\mathcal{H}_{[n-1]}$ for $(\mathcal{W}, f_0, f_1)$, then for any other n–stage cofibrant sequential realization ${}'\mathcal{W}$ of $\boldsymbol{Y}$ we can choose the $(n-1)$–strand ${}'\mathcal{H}_{[n-1]}$ for ${}'\mathcal{W}$ to be weakly equivalent to $\mathcal{H}_{[n-1]}$.

In analogy to Theorem 4.18 we therefore have:

5.29 Theorem *For R either $\mathbb{F}_p$ or a field of characteristic 0, let $f_0, f_1\colon \boldsymbol{Z} \to \boldsymbol{Y}$ be two maps between R–good spaces which induce the same map of Θ_R–algebras $H^*(\boldsymbol{Y}; R) \to H^*(\boldsymbol{Z}; R)$. Then the following are equivalent:*

(a) *The system of higher cohomology operations $\langle\!\langle f_{(0,1)}\rangle\!\rangle$ vanishes coherently for* some *cofibrant sequential realization $\mathcal{W}$ of $\boldsymbol{Y}$.*

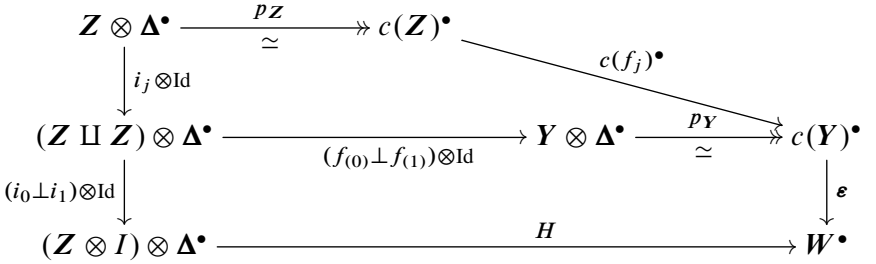
(b) *$\langle\!\langle f_{(0,1)}\rangle\!\rangle$ vanishes coherently for* every *cofibrant sequential realization of $\boldsymbol{Y}$.*

(c) *The maps f_0 and f_1 are R–equivalent* (cf Section 1.22).

Proof (a) $\Longleftrightarrow$ (b) By Lemma 5.28.

(a) $\Longrightarrow$ (c) Note that the projection $p_{\boldsymbol{X}}\colon \boldsymbol{X}\otimes\boldsymbol{\Delta}^\bullet \to c(\boldsymbol{X})^\bullet$ is a trivial Reedy fibration for any $\boldsymbol{X}\in\mathcal{C}$. Since $\boldsymbol{Z}\amalg\boldsymbol{Z}\xrightarrow{i_0\perp i_1}\boldsymbol{Z}\otimes I\xrightarrow{\sigma}\boldsymbol{Z}$ is a cylinder object in $\mathcal{C}$ (see Quillen [32, Section I.1]), the same is true after applying $(-)\otimes\boldsymbol{\Delta}^\bullet$. An ∞–strand $\mathcal{H}_{[\infty]}$ for a sequential realization $\mathcal{W}$ (with associated $\boldsymbol{\varepsilon}\colon \boldsymbol{Y}\to\boldsymbol{W}^\bullet$) defines a map $H\colon (\boldsymbol{Z}\otimes I)\otimes\boldsymbol{\Delta}^\bullet\to\boldsymbol{W}^\bullet$, fitting into a commutative diagram of cosimplicial spaces

(5.30)



for $j=0,1$. Applying Tot yields a cylinder object

$$\mathrm{Tot}((\boldsymbol{Z}\otimes I)\otimes\boldsymbol{\Delta}^\bullet)$$

for $\mathrm{Tot}((\boldsymbol{Z}\amalg\boldsymbol{Z})\otimes\boldsymbol{\Delta}^\bullet)$, and a homotopy $\mathrm{Tot}\,H$ between $\boldsymbol{\varepsilon}_*\circ f_0\circ\mathrm{Tot}(p_{\boldsymbol{Z}})$ and $\boldsymbol{\varepsilon}_*\circ f_1\circ\mathrm{Tot}(p_{\boldsymbol{Z}})$. Since $\mathrm{Tot}(p_{\boldsymbol{Z}})$ is a weak equivalence and $\boldsymbol{\varepsilon}_*\colon \boldsymbol{Y}\to\mathrm{Tot}\,\boldsymbol{W}^\bullet$ is an R–equivalence, we see that f_0 and f_1 are R–equivalent.

(c) $\Longrightarrow$ (b) Let $\mathcal{W}$ be any cofibrant sequential realization for $\boldsymbol{Y}$ with associated $\boldsymbol{\varepsilon}\colon \boldsymbol{Y}\to\boldsymbol{W}^\bullet$. By Definition 2.1, $\boldsymbol{W}^\bullet$ is Reedy fibrant. Thus $\mathrm{Tot}\,\boldsymbol{W}^\bullet$ is an R–complete Kan complex, with $\boldsymbol{\varepsilon}_*\colon \boldsymbol{Y}\to\mathrm{Tot}\,\boldsymbol{W}^\bullet$ the R–completion map, and so the R–equivalent maps $\boldsymbol{\varepsilon}_*\circ f_0$ and $\boldsymbol{\varepsilon}_*\circ f_1$ are actually homotopic (see [20, Lemma I.5.5]). We may therefore choose a homotopy $F\colon \boldsymbol{Z}\otimes I\to\mathrm{Tot}\,\boldsymbol{W}^\bullet$ between them, whose adjoint is the map of cosimplicial spaces $\widetilde{F}\colon \boldsymbol{Z}\otimes I\otimes\boldsymbol{\Delta}^\bullet\to\boldsymbol{W}^\bullet$ (see [20, Section I.3.3]). Composing $\widetilde{F}$ with the structure maps $\boldsymbol{W}^\bullet\to\boldsymbol{W}^\bullet_{[n]}$ for the limit of (2.2) yields a compatible sequence of cosimplicial maps $H_{[n]}\colon \boldsymbol{Z}\otimes I\otimes\boldsymbol{\Delta}^\bullet\to\boldsymbol{W}^\bullet_{[n]}$.

This defines compatible n–strands for $\mathcal{W}$ and all $n\geq 1$, showing that the system $\langle\!\langle f_{(0,1)}\rangle\!\rangle$ of higher-order operations vanishes by Proposition 5.23. □

5.31 Corollary *If $f_0, f_1\colon \boldsymbol{Z}\to\boldsymbol{Y}$ are maps between R–complete Kan complexes inducing the same map of Θ_R–algebras $\psi\colon H^*(\boldsymbol{Y};R)\to H^*(\boldsymbol{Z};R)$, the system of higher operations (f_0,f_1) is a complete set of invariants for the homotopy classes $[f_0]$ and $[f_1]$.*

5.32 A rational example for a map As in Section 4.24, we now consider an example of the obstruction to a map $f\colon \boldsymbol{Z} \to \boldsymbol{Y}$ being rationally trivial when $f^*\colon H^*(\boldsymbol{Y}; Q) \to H^*(\boldsymbol{Z}; Q)$ is the zero map:

Let $\boldsymbol{Z} := \boldsymbol{S}^{2n-1}_{\mathbb{Q}}$ and $\boldsymbol{Y} := (\boldsymbol{S}^n \vee \boldsymbol{S}^n)_{\mathbb{Q}}$ for $n > 1$ odd, with $f := [\iota_n, \iota'_n]\colon \boldsymbol{Z} \to \boldsymbol{Y}$ the Whitehead product map. The free CDGA model for $\boldsymbol{Y}$ is (A^*, d) given in degrees $\leq 2n$ by $A^* = \mathbb{Q}\langle x_n, y_n, u_{2n-1}\rangle$ with $d(u) = xy$, while $\boldsymbol{Z}$ has the formal CDGA model $B^* = \Lambda[\boldsymbol{z}_{2n-1}]$. The CDGA model for f is $\varphi\colon A^* \to B^*$ mapping u to z.

Realizing the obvious minimal free algebraic resolution of $H^*(\boldsymbol{Y}; Q)$, we obtain the 1–truncated augmented simplicial CDGA $\widetilde{W}^{[1]}_\bullet \to A^* \to B^*$ in degrees $\leq 2n$ depicted in Figure 2.

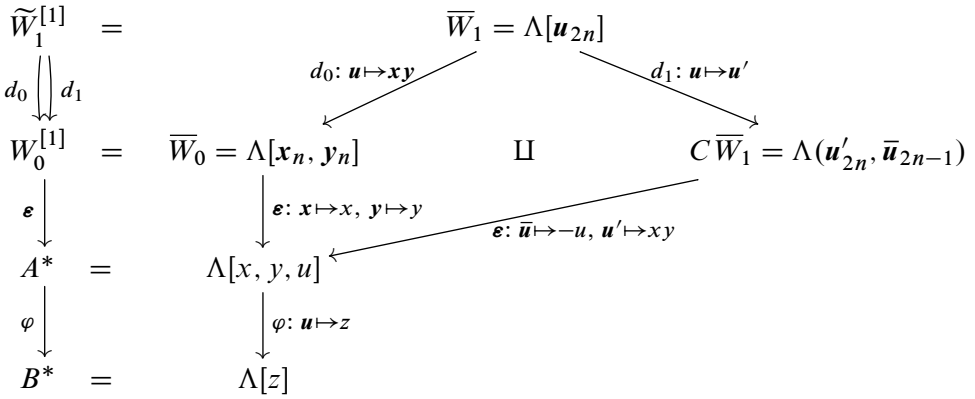


Figure 2: $\widetilde{W}^{[1]}_\bullet$ in degrees $\leq 2n$

As noted above, the original map $f\colon \boldsymbol{Z} \to \boldsymbol{Y}$ is nullhomotopic if and only if we can extend the composite $\widetilde{W}^{[1]}_\bullet \to B^*$ in the diagram in Figure 2 to the cone $C\widetilde{W}^{[1]}_\bullet \to B^*$, which by Theorem 5.29 is equivalent to the vanishing of the associated system of higher operations.

The model for the cofiber of $W^{[1]}_\bullet \hookrightarrow C\overline{W}_1$ (corresponding to the loop space $\Omega\overline{\boldsymbol{W}}^1$) is the formal CDGA $\Sigma\overline{W}_1 := \Lambda(\overline{\boldsymbol{u}}'_{2n-1})$, and its cone (corresponding to the path space $P\Omega\overline{\boldsymbol{W}}^1$) is $C\Sigma\overline{W}_1 := \Lambda(\overline{\boldsymbol{u}}''_{2n-1}, \overline{\overline{\boldsymbol{u}}}_{2n-2})$, with $d(\overline{\overline{\boldsymbol{u}}}_{2n-2}) = \overline{\boldsymbol{u}}''_{2n-1}$.

Thus in order to extend the given map $\varphi \circ \boldsymbol{\varepsilon}\colon W^{[1]}_0 \to B^*$ (sending $\overline{\boldsymbol{u}}_{2n-1}$ to $-z$) to its cone, we need a map $F\colon C\Sigma\overline{W}_1 \to B^*$ sending $\overline{\boldsymbol{u}}''_{2n-1}$ to $-z$. However, since necessarily $F(\overline{\overline{\boldsymbol{u}}}) = 0$, this is impossible — that is, the secondary cohomology operation does not vanish: its value is represented by the map $\psi\colon \Sigma\overline{W}_1 \to B^*$ defined by $\psi(\overline{\boldsymbol{u}}'_{2n-1}) = z$.

Appendix: Proof of Theorem 2.33

In this appendix we state and prove Theorem 2.33 in a more general form needed in Blanc and Sen [17]. For this purpose we recall the notion of a *mapping algebra*, which encodes the extra structure on the mapping spaces $\mathrm{map}_*(\boldsymbol{Y}, \boldsymbol{K}(R,n))$ needed to recover the R–completion of $\boldsymbol{Y}$ from them (see Blanc and Sen [17]).

A.1 Definition An *enriched sketch* $(\boldsymbol{\Theta}, \mathcal{P}, \mathcal{K})$ is a small subcategory $\boldsymbol{\Theta}$ of a simplicial category $\mathcal{C}$ (see Definition 1.19), with $\boldsymbol{\Theta}$ closed under a given set of limits $\mathcal{P}$ and under $(-)^K$ for K in a given subcategory $\mathcal{K}$ of $\mathcal{S}$. We assume that all mapping spaces $\mathrm{map}_{\boldsymbol{\Theta}}(\boldsymbol{B}, \boldsymbol{B}')$ are Kan complexes.

A $\boldsymbol{\Theta}$–*mapping algebra* is a pointed simplicial functor $\mathfrak{X}\colon \boldsymbol{\Theta} \to \mathcal{S}_*$ (written $\mathfrak{X}\colon \boldsymbol{B} \mapsto \mathfrak{X}\{\boldsymbol{B}\}$) which preserves the limits in $\mathcal{P}$ and with $\mathfrak{X}\{(\boldsymbol{B})^K\} = (\mathfrak{X}\{\boldsymbol{B}\})^K$ for any $\boldsymbol{B} \in \boldsymbol{\Theta}$ and $K \in \mathcal{K}$. The category of $\boldsymbol{\Theta}$–mapping algebras will be denoted by $\mathcal{M}ap_{\boldsymbol{\Theta}}$. See Baues and Blanc [2, Section 8] and Blanc and Sen [17, Section 1] for further details.

For any $\boldsymbol{Y} \in \mathcal{C}$ we have a *realizable* $\boldsymbol{\Theta}$–mapping algebra $\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{Y}$ defined for each $\boldsymbol{B} \in \boldsymbol{\Theta}$ by $\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{Y}\{\boldsymbol{B}\} := \mathrm{map}_{\mathcal{C}}(\boldsymbol{Y}, \boldsymbol{B})$.

A.2 Remark Assume that in our enriched sketch $(\boldsymbol{\Theta}, \mathcal{P}, \mathcal{K})$, the category $\mathcal{K}$ includes $\Delta[0] \hookrightarrow \Delta[1]$, and $\mathcal{P}$ includes all finite products and the pullback squares

(A.3)

$$\begin{array}{ccc} P\boldsymbol{B} & \hookrightarrow & \boldsymbol{B}^{\Delta[1]} \\ \simeq\downarrow & \mathrm{PB} & \downarrow{\scriptstyle \mathrm{ev}_0}\simeq \\ * & \hookrightarrow & \boldsymbol{B} \end{array} \qquad \begin{array}{ccc} \Omega\boldsymbol{B} & \xhookrightarrow{\iota_{\boldsymbol{B}}} & P\boldsymbol{B} \\ \downarrow & \mathrm{PB} & \downarrow{\scriptstyle \mathrm{ev}_1} \\ * & \hookrightarrow & \boldsymbol{B} \end{array}$$

for any $\boldsymbol{B} \in \boldsymbol{\Theta}_R$. For each $\boldsymbol{\Theta}$–mapping algebra $\mathfrak{X}$ and objects $\boldsymbol{B}$ and $\{\boldsymbol{B}_i\}_{i=1}^n$ in $\boldsymbol{\Theta}$ we then have natural isomorphisms

(a) $i_\Pi\colon \mathfrak{X}\{\prod_{i=1}^n \boldsymbol{B}_i\} \xrightarrow{\cong} \prod_{i=1}^n \mathfrak{X}\{\boldsymbol{B}_i\}$;

(b) $i_P\colon \mathfrak{X}\{P\boldsymbol{B}\} \xrightarrow{\cong} P\mathfrak{X}\{\boldsymbol{B}\}$;

(c) $i_\Omega\colon \mathfrak{X}\{\Omega\boldsymbol{B}\} \xrightarrow{\cong} \Omega\mathfrak{X}\{\boldsymbol{B}\}$.

We assume that the objects $B \in \boldsymbol{\Theta}$ are fibrant in $\mathcal{C}$. In order to simplify the proof of Theorem A.11 below, we make the following ad hoc assumption: if $f\colon \boldsymbol{W}^\bullet \to \boldsymbol{Y}^\bullet$ is a map of cosimplicial object over $\mathcal{C}$ with each $\boldsymbol{W}^n$ and $\boldsymbol{Y}^n$ in $\boldsymbol{\Theta}$ and $\boldsymbol{W}^\bullet \to \boldsymbol{Z}^\bullet \to \boldsymbol{Y}^\bullet$ is the functorial factorization of f as a (trivial) Reedy cofibration followed by a (trivial) Reedy fibration, then each $\boldsymbol{Z}^n$ is in $\boldsymbol{\Theta}$, too.

This will hold, for instance, when $\mathcal{C} = \mathcal{S}_*$ and $\boldsymbol{\Theta}$ consists of all simplicial R–modules of cardinality $< \lambda$ (for some limit cardinal λ). This is the example for which case (1) of Theorem A.11 is needed in Blanc and Sen [17].

We have the following enriched version of Lemma 1.28:

A.4 Lemma [17, Lemma 1.9] *For any $\boldsymbol{\Theta}$–mapping algebra $\mathfrak{Y}$ and $\boldsymbol{B} \in \boldsymbol{\Theta}$, there is a natural isomorphism* $\mathrm{Hom}_{\mathcal{M}ap_{\boldsymbol{\Theta}}}(\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{B}, \mathfrak{Y}) \xrightarrow{\cong} \mathfrak{Y}\{\boldsymbol{B}\}_0$.

A.5 Definition To any enriched sketch $(\boldsymbol{\Theta}, \mathcal{P}, \mathcal{K})$ we associate a sketch $\pi_0\boldsymbol{\Theta}$ as in Remark 1.25, with the same objects and products as $\boldsymbol{\Theta}$, where $\mathrm{Hom}_{\Theta}(\boldsymbol{B}, \boldsymbol{B}') := \pi_0 \operatorname{map}_{\boldsymbol{\Theta}}(\boldsymbol{B}, \boldsymbol{B}')$. A map of $\boldsymbol{\Theta}$–mapping algebras $f\colon \mathfrak{X} \to \mathfrak{Y}$ is a *weak equivalence* if it induces a weak equivalence $f_{\boldsymbol{B}}\colon \mathfrak{X}\{\mathcal{B}\} \to \mathfrak{Y}\{\mathcal{B}\}$ for any $\mathcal{B} \in \boldsymbol{\Theta}$. This means that f induces an isomorphism $f_{\#}\colon \pi_0\mathfrak{X} \to \pi_0\mathfrak{Y}$ of the corresponding $\pi_0\boldsymbol{\Theta}$–algebras.

A.6 Example For any commutative ring R, let $\boldsymbol{\Theta}_R \subseteq \mathcal{S}_*$ denote the enriched sketch whose objects are finite-type R–GEMs of the form $\prod_{i=1}^{\infty} \boldsymbol{K}(V_i, m_i)$ for $m_i \geq 0$, with V_i a finite-dimensional free R–module where $\mathcal{K}$ is as above and $\mathcal{P}$ includes all such finite-type products. Since each $\boldsymbol{B} \in \boldsymbol{\Theta}_R$ is an R–module object in $\mathcal{S}_*$, the same is true of $\mathfrak{X}\{\boldsymbol{B}\}$, so $\boldsymbol{\Theta}_R$–mapping algebras actually take value in simplicial R–modules. The enriched sketch $\boldsymbol{\Theta}_R^{\lambda}$ is defined analogously (see Example 1.27).

Note that $\pi_0\boldsymbol{\Theta}_R$ includes also 0–dimensional Eilenberg–Mac Lane spaces, while the algebraic sketch Θ_R consists of R–GEMs in dimensions ≥ 1 (to avoid having to deal with the nonreduced cohomology of nonconnected spaces). Thus $\Theta_R \subset \pi_0\boldsymbol{\Theta}_R$, which motivates the following:

A.7 Definition For $\boldsymbol{\Theta}$ an enriched sketch in $\mathcal{C}$ and $\mathfrak{X}$ a $\boldsymbol{\Theta}$–mapping algebra, let $\Theta \subseteq \pi_0\boldsymbol{\Theta}$ be a subalgebraic sketch (still closed under finite products, but not necessarily under loops), and let $V_\bullet \to \pi_0\mathfrak{X}$ be a CW resolution of the corresponding Θ–algebra. A *sequential realization* $\mathcal{W} = \langle \boldsymbol{W}^{\bullet}_{[n]}, \widetilde{\boldsymbol{W}}^{\bullet}_{[n]} \rangle_{n\in\mathbb{N}}$ of $V_\bullet$ for $\mathfrak{X}$ consists of a tower of Reedy fibrant and cofibrant cosimplicial objects as in (2.2), such that:

(a) We have an augmentation $\varepsilon_{[n]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n]} \to \mathfrak{X}$ for the simplicial $\boldsymbol{\Theta}$–mapping algebra $\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{\bullet}_{[n]}$, realizing $V_\bullet \to \pi_0\mathfrak{X}$ through simplicial dimension n — ie we have a natural isomorphism as in (2.3).

(b) The augmentation $\varepsilon_{[n-1]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]} \to \mathfrak{X}$ extends along the $\boldsymbol{\Theta}$–mapping algebra map $\pi^*_{[n]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]} \to \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n]}$ to $\varepsilon_{[n]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n]} \to \mathfrak{X}$.

(c) Each $\boldsymbol{W}^{\bullet}_{[n]}$ is obtained from $\boldsymbol{W}^{\bullet}_{[n-1]}$ as in Definition 2.1(c).

A.8 Remark We do not require the sequential realization $\mathcal{W} = \langle \boldsymbol{W}^{\bullet}_{[n]}, \widetilde{\boldsymbol{W}}^{\bullet}_{[n]} \rangle_{n\in\mathbb{N}}$ for $\mathfrak{X}$ to be cofibrant, as in Definition 2.1(d). Therefore, in the explicit description of Section 2.5 we may take the fibration sequences (2.6) in step (i) to be the standard path-loop fibrations, so

$$\overline{\Omega^j \boldsymbol{W}^n} := \Omega^j \overline{\boldsymbol{W}}^n \tag{A.9}$$

for all $0 \le j \le n-1$. Thus (2.10) becomes simply

$$\widetilde{\boldsymbol{W}}^k_{[n]} = \boldsymbol{W}^k_{[n-1]} \times P\Omega^{n-k-1}\overline{\boldsymbol{W}}^n \tag{A.10}$$

for all $0 \le k \le n$ (with (2.11) still holding for $k = n$).

A.11 Theorem *Let $\boldsymbol{\Theta}$ be an enriched sketch $\boldsymbol{\Theta}$ in a model category $\mathcal{C}$ as in Remark A.2, and $\Theta \subseteq \pi_0\boldsymbol{\Theta}$ an algebraic sketch.*

(1) *If $\mathfrak{X}$ is a $\boldsymbol{\Theta}$–mapping algebra, and $V_\bullet$ a CW resolution of $\Gamma := \pi_0\mathfrak{X}$ with CW basis $(\overline{V}_n)_{n\in\mathbb{N}}$ such that each $\overline{V}_n$ is realizable by an object $\overline{\boldsymbol{W}}^n \in \boldsymbol{\Theta}$, then there is a sequential realization $\mathcal{W} = \langle \boldsymbol{W}^{\bullet}_{[n]}, \widetilde{\boldsymbol{W}}^{\bullet}_{[n]} \rangle_{n\in\mathbb{N}}$ of $V_\bullet$ as in Definition A.7.*

(2) *If Θ is allowable* (see Definition 2.28), *$\boldsymbol{Y} \in \mathcal{C}$, and $V_\bullet$ is any CW resolution of $\Gamma := H^*_\Theta \boldsymbol{Y} = \pi_0\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{Y}$, then $V_\bullet$ has a cofibrant sequential realization $\mathcal{W} = \langle \boldsymbol{W}^{\bullet}_{[n]}, \widetilde{\boldsymbol{W}}^{\bullet}_{[n]} \rangle_{n\in\mathbb{N}}$ of $V_\bullet$ for $\boldsymbol{Y}$.*

Thus, in both cases $V_\bullet$ is realizable by $\boldsymbol{W}^\bullet = \operatorname{holim}_n \boldsymbol{W}^{\bullet}_{[n]}$. In case (2), *if Θ is contained in some class of injective models $\mathcal{G}$ in $\mathcal{C}$, then $\boldsymbol{Y} \to \boldsymbol{W}^\bullet$ is a weak $\mathcal{G}$–resolution* (see Section 1.22).

Proof In case (2), for each $n \ge 0$ we choose an object $\overline{\boldsymbol{W}}^n \in \Theta$ realizing $\overline{V}_n$. In both cases, we then construct a sequential realization $\mathcal{W}$ by a double induction, where in the outer induction $\boldsymbol{W}^{\bullet}_{[n]}$ is obtained from $\boldsymbol{W}^{\bullet}_{[n-1]}$ as in Definition 2.1(c), using an inner descending induction on $0 \le k \le n$:

I Step $n = 0$ of the outer induction We start the induction with $\boldsymbol{W}^{\bullet}_{[0]} := c(\overline{\boldsymbol{W}}^0)^\bullet$ (the constant cosimplicial object), which is weakly $\mathcal{G}$–fibrant. Because $\overline{V}_0$ is a free Θ–algebra, in case (1) the Θ–algebra augmentation $\varepsilon\colon \overline{V}_0 \to \Gamma$ corresponds by Lemma 1.28 to a unique element in $[\overline{\varepsilon}_{[0]}] \in \Gamma\{\overline{\boldsymbol{W}}^0\} = \pi_0\mathfrak{X}\{\overline{\boldsymbol{W}}^0\}$, for which we may choose a representative $\overline{\varepsilon}_{[0]} \in \mathfrak{X}\{\overline{\boldsymbol{W}}^0\}_0$, corresponding to a map of $\boldsymbol{\Theta}$–mapping algebras $\varepsilon_{[0]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\overline{\boldsymbol{W}}^0 \to \mathfrak{X}$ by Lemma A.4. In case (2), we may realize ε by a map $\boldsymbol{\varepsilon}\colon \boldsymbol{Y} \to \overline{\boldsymbol{W}}^0$ by Lemma 1.28 and (2.31), since Θ is allowable.

II Step $n = 1$ of the outer induction We choose a map $C^0\boldsymbol{W}^{\bullet}_{[0]} = \overline{\boldsymbol{W}}^0 \to \overline{\boldsymbol{W}}^1$ realizing the first attaching map $\overline{\partial}^1_0\colon \overline{V}_1 \to V_0 = \overline{V}_0$, with $\boldsymbol{W}^{\bullet}_{[1]}$ given in dimensions

≤ 1 by

(A.12)
$$\begin{array}{ccccccc} \boldsymbol{W}^0_{[1]} & = & \overline{\boldsymbol{W}}^0 & & \times & & P\overline{\boldsymbol{W}}^1 \\ d_0^0 \left(\begin{array}{c} \searrow \\ s^0 \\ \searrow \end{array}\right. \uparrow \left.\begin{array}{c} \downarrow \\ \uparrow \\ \downarrow \end{array}\right) d_0^1 & d_0^0{=}d_0^1{=}\mathrm{Id} & \left(\begin{array}{c} \downarrow \\ \uparrow \\ \downarrow \end{array}\right) & = & \begin{array}{c} \overset{\bar{d}_0^0}{\searrow} \\ \underset{\searrow}{} \end{array} & \begin{array}{c} \overset{d_0^1=p}{\swarrow} \\ \underset{\swarrow}{} \end{array} & = \left(\begin{array}{c} \uparrow \\ \downarrow \\ \downarrow \end{array}\right) d_0^0{=}d_0^1{=}\mathrm{Id} \\ \boldsymbol{W}^1_{[1]} & = & \overline{\boldsymbol{W}}^0 & \times & \overline{\boldsymbol{W}}^1 & \times & P\overline{\boldsymbol{W}}^1 \end{array}$$

In case (1), to define the augmentation $\overline{\varepsilon}_{[1]}$ as a 0–simplex in $\mathfrak{X}\{\boldsymbol{W}^0_{[1]}\}_0$ extending $\overline{\varepsilon}_{[0]} \in \mathfrak{X}\{\boldsymbol{W}^0_{[0]}\} = \mathfrak{X}\{\overline{\boldsymbol{W}}^0\}$, we use the fact that $\mathfrak{X}\{\boldsymbol{W}^0_{[1]}\} = \mathfrak{X}\{\overline{\boldsymbol{W}}^0\} \times P\mathfrak{X}\{\overline{\boldsymbol{W}}^1\}$, by Remark A.2(a)-(b), so we need only to find a 0–simplex H in $P\mathfrak{X}\{\overline{\boldsymbol{W}}^1\}$ — which, by (A.3), is a 1–simplex in $\mathfrak{X}\{\overline{\boldsymbol{W}}^1\}$ with $d_1 H = 0$.

In order to qualify as an augmentation $\mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[1]} \to \mathfrak{X}$ of simplicial $\boldsymbol{\Theta}$–mapping algebras, $\varepsilon_{[1]}$ must satisfy the simplicial identity

(A.13)
$$\varepsilon_{[1]} \circ d_0 = \varepsilon_{[1]} \circ d_1 \colon \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^1_{[1]} \to \mathfrak{X}$$

as maps of $\boldsymbol{\Theta}$–mapping algebras — or equivalently, these must correspond to the same 0–simplex in $\mathfrak{X}\{\boldsymbol{W}^1_{[1]}\} = \mathfrak{X}\{\overline{\boldsymbol{W}}^0\} \times \mathfrak{X}\{\overline{\boldsymbol{W}}^1\} \times P\mathfrak{X}\{\overline{\boldsymbol{W}}^1\}$. In the first and third factors this obviously holds, so we need only consider the two 0–simplices in $\mathfrak{X}\{\overline{\boldsymbol{W}}^1\}$: in other words, since the path fibration p in (A.12) (induced by the inclusion $\Delta[0] \hookrightarrow \Delta[1]$) becomes d_0 in $\mathfrak{X}\{\overline{\boldsymbol{W}}^1\}$, we must choose H so that $d_0 H$ is the 0–simplex $(\bar{d}_0^0)_{\#}\overline{\varepsilon}_{[0]}$. By (1.32), $\varepsilon \circ \bar{\partial}_0^0 = 0$ in Θ–Alg, which implies (by our choices of $\bar{d}_0^0$ and $\overline{\varepsilon}_{[0]}$ representing $\bar{\partial}_0^0$ and ε, respectively) that $(\bar{d}_0^0)_{\#}\overline{\varepsilon}_{[0]}$ is nullhomotopic, so the required H exists.

In case (2), we choose a nullhomotopy for $\bar{d}_0^0 \circ \boldsymbol{\varepsilon}$ to extend $\boldsymbol{\varepsilon}_{[0]}$ to the factor $P\overline{\boldsymbol{W}}^1$, and thus define $\boldsymbol{\varepsilon}_{[1]} \colon \boldsymbol{Y} \to \boldsymbol{W}^0_{[1]}$.

III Step n of the outer induction $n \geq 2$ To construct $\boldsymbol{W}^\bullet_{[n]}$ given $\boldsymbol{W}^\bullet_{[n-1]}$, by Proposition 2.24 it suffices to produce a cochain map $F \colon C^*\boldsymbol{W}^\bullet_{[n-1]} \to \boldsymbol{D}^*$ as in Section 2.5 (where the left Reedy fibrant replacement $\boldsymbol{D}^*$ for $\overline{\boldsymbol{W}}^n \otimes^* S^{n-1}$ is given by (A.9) and (2.8)). We do so by a descending induction on the cosimplicial dimension $0 \leq k \leq n-1$, starting with Step IV below for $k = n-1$, under the following induction hypotheses:

In stage k we assume the existence of $F^j \colon C^j\boldsymbol{W}^\bullet_{[n-1]} \to P\Omega^{n-j-2}\overline{\boldsymbol{W}}^n$ (constituting a cochain map for $k+1 \leq j \leq n-1$) and $a^j \colon Z^j\boldsymbol{W}^\bullet_{[n-1]} \to \Omega^{n-j-2}\overline{\boldsymbol{W}}^n$ for $k \leq j \leq n-1$, with a nullhomotopy

$$\hat{F}^k \colon C^k\boldsymbol{W}^\bullet_{[n-1]} \to P\Omega^{n-k-2}\overline{\boldsymbol{W}}^n$$

such that

$$a^k \circ w^k = p \circ \widehat{F}^k \colon C^k \boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega^{n-k-2}\overline{\boldsymbol{W}}^n, \tag{A.14}$$

as in (2.26). By Lemma 2.19, $\widehat{F}^k$ induces $\widehat{a}^{k-1}\colon Z^{k-1}\boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega^{n-k-1}\overline{\boldsymbol{W}}^n$ with

$$\iota \circ \widehat{a}^{k-1} \circ w^{k-1} = \widehat{F}^k \circ \delta^{k-1}, \tag{A.15}$$

as in (2.20), where $\iota = \overline{\iota}^{n-k-1}\colon \Omega^{n-k-1}\overline{\boldsymbol{W}}^n \hookrightarrow P\Omega^{n-k-2}\overline{\boldsymbol{W}}^n$ is the inclusion. However, we do not assume in our induction hypothesis that $\widehat{a}^{k-1} \circ w^{k-1}$ is nullhomotopic.

IV Step $k = n-1$ of the inner induction To define $F^{n-1}\colon C^{n-1}\boldsymbol{W}^{\bullet}_{[n-1]} \to \overline{\boldsymbol{W}}^n$, note that the simplicial space $\boldsymbol{U}_\bullet := \mathrm{map}_{\mathcal{C}}(\boldsymbol{W}^{\bullet}_{[n-1]}, \overline{\boldsymbol{W}}^n)$ is Reedy fibrant, since $\boldsymbol{W}^{\bullet}_{[n-1]}$ is Reedy cofibrant. Moreover, since $H^*_\Theta \boldsymbol{W}^k_{[n-1]} \cong V_k$ for all $0 \le k < n$ by Definition 2.1(a), the attaching map $\overline{\partial}^n_0\colon \overline{V}_n \to V_{n-1}$ yields a homotopy class

$$\begin{aligned}\alpha \in [\boldsymbol{W}^{n-1}_{[n-1]}, \overline{\boldsymbol{W}}^n] &= \pi_0\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{n-1}_{[n-1]}\{\overline{\boldsymbol{W}}^n\} = V_{n-1}\{\overline{\boldsymbol{W}}^n\} \\ &= \mathrm{Hom}_{\Theta\text{–}\mathsf{Alg}}(H^*_\Theta \overline{\boldsymbol{W}}^n, V_{n-1}) = \mathrm{Hom}_{\Theta\text{–}\mathsf{Alg}}(\overline{V}_n, V_{n-1}),\end{aligned} \tag{A.16}$$

where the next-to-last equality follows from Lemma 1.28, as extended in (2.31).

This α is a Moore chain in $\pi_0\boldsymbol{U}_\bullet$ by Section 1.11, so by Lemma 2.22(a), it can be represented by a map $F^{n-1}\colon C^{n-1}\boldsymbol{W}^{\bullet}_{[n-1]} \to \overline{\boldsymbol{W}}^n$, which induces $a^{n-2}\colon Z^{n-2}\boldsymbol{W}^{\bullet}_{[n-1]} \to \overline{\boldsymbol{W}}^n$ by Lemma 2.19 with $a^{n-2} \circ w^{n-2} = F^{n-1} \circ \delta^{n-2}$. Moreover, by (1.12) α is in fact a Moore $(n-1)$–cycle, so we have a nullhomotopy $\widehat{F}^{n-2}\colon C^{n-2} \to P\overline{\boldsymbol{W}}^n$ for $a^{n-2} \circ w^{n-2}$, as in (A.14).

V Step k of the inner induction $1 < k \le n-2$ Let $\boldsymbol{A} := \Omega^{n-k-2}\overline{\boldsymbol{W}}^n$. By assumption (Step III), we have a nullhomotopy $\widehat{F}^k\colon C^k\boldsymbol{W}^{\bullet}_{[n-1]} \to P\boldsymbol{A}$ for $a^k \circ w^k$, and $\widehat{F}^k \circ \delta^{k-1}$ determines $\widehat{a}^{k-1}\colon Z^{k-1}\boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega\boldsymbol{A}$ satisfying (A.15), which is thus a $(k-1)$–cycle for the Reedy fibrant bisimplicial set $\mathrm{map}_{\mathcal{C}}(\boldsymbol{W}^{\bullet}_{[n-1]}, \Omega\boldsymbol{A})$.

Since $\boldsymbol{W}^{\bullet}_{[n-1]}$ realizes $V_\bullet$ through simplicial dimension $n-1 > k$, by Definition 2.1(a), $\widehat{a}^{k-1} \circ w^{k-1} \circ v^{k-1}\colon \boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega\boldsymbol{A}$ represents a $(k-1)$–cycle $[\widehat{a}^{k-1}]$ for $V_\bullet\{\Omega\boldsymbol{A}\}$, as in (A.16). Because $V_\bullet \to \Gamma$ is a resolution, and thus acyclic, there is a Moore chain γ_k in $C_k V_\bullet\{\Omega\boldsymbol{A}\}$ with $\partial_0^{V_k}(\gamma_k) = [\widehat{a}^{k-1}]$. Since $V_k\{\Omega\boldsymbol{A}\} \cong \pi_0\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^k_{[n-1]}\{\Omega\boldsymbol{A}\} = [\boldsymbol{W}^k_{[n-1]}, \Omega\boldsymbol{A}]$, by Lemma 2.22(a) we can represent γ_k by a map $g^k\colon C^k\boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega\boldsymbol{A}$, while by Lemma 2.22(b) we have a homotopy

$$G\colon g^k \circ \delta^{k-1} \sim \widehat{a}^{k-1} \circ w^k \colon C^{k-1}\boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega\boldsymbol{A}. \tag{A.17}$$

Next, concatenation of homotopies gives an action of $\mathrm{Hom}(\boldsymbol{C}, \Omega\boldsymbol{A})$ on $\mathrm{Hom}(\boldsymbol{C}, P\boldsymbol{A})$ (see [35, Section 1]), which we use to define a new nullhomotopy

$$\textbf{(A.18)} \qquad F^k := \widehat{F}^k \star (\iota \circ g^k)^{-1} \colon C^k \boldsymbol{W}^{\bullet}_{[n-1]} \to P\boldsymbol{A}$$

for $a^k \circ w^k$ (where $\iota \colon \Omega\boldsymbol{A} \hookrightarrow P\boldsymbol{A}$ is the inclusion).

By Lemma 2.19, F^k induces a map $a^{k-1} \colon Z^{k-1}\boldsymbol{W}^{\bullet}_{[n-1]} \to \Omega\boldsymbol{A}$ satisfying

$$\textbf{(A.19)} \qquad \begin{aligned} \iota \circ \widetilde{a} \circ w^{k-1} &= \widetilde{F} \circ \delta^{k-1} = (\widehat{F}^k \star (\iota \circ g^k)^{-1}) \circ \delta^{k-1} \\ &= (\widehat{F}^k \circ \delta^{k-1}) \star (\iota \circ g^k \circ \delta^{k-1})^{-1} \\ &= \iota \circ [(\widehat{a}^{k-1} \circ w^{k-1}) \star (g^k \circ \delta^{k-1})^{-1})] \end{aligned}$$

by (2.20), (A.18), (A.15) and the fact that the H–space structure $\star$ and $(-)^{-1}$ commutes with precomposition of maps into $\Omega\boldsymbol{A}$.

Since ι is a monomorphism, we conclude that

$$\textbf{(A.20)} \qquad \begin{aligned} a^{k-1} \circ w^{k-1} &= (\widehat{a}^{k-1} \circ w^{k-1}) \star (g^k \circ \delta^{k-1})^{-1}) \\ &\sim (\widehat{a}^{k-1} \circ w^{k-1}) \star (\widehat{a}^{k-1} \circ w^{k-1})^{-1} \sim 0 \end{aligned}$$

by (A.17), so a^{k-1} satisfies (A.14) for $k-1$.

In case (2), the last two steps of the downward induction are no different from those for $k \geq 2$ if we set $\widetilde{\boldsymbol{W}}^{-1}_{[n]} = \boldsymbol{W}^{-1}_{[n-1]} := \boldsymbol{Y}$, with $\widetilde{d}^0_{-1} \colon \widetilde{\boldsymbol{W}}^{-1}_{[n]} \to \widetilde{\boldsymbol{W}}^0_{[n]}$ as the coaugmentation $\boldsymbol{\varepsilon}_{[n]}$. However, in case (1) we no longer have an object $\boldsymbol{W}^{k-1}_{[n-1]}$ in $\mathcal{C}$ for $k = 0$, so we must modify our construction somewhat, using the language of $\boldsymbol{\Theta}$–mapping algebras, as follows:

VI Step $k = 1$ of the descending induction By the descending induction hypotheses III for $k = 1$ we have some nullhomotopy $\widehat{F}^1 \colon a^1 \circ w^1 \sim 0$ and $\widehat{a}^0$ with $\iota \circ \widehat{a}^0 \circ w^0 = \widehat{F}^1 \circ d^0_0$ by (A.15). Let $\boldsymbol{B} := \Omega^{n-3}\overline{\boldsymbol{W}}^n$.

We can think of $a^1 \circ w^1 \circ v^1$ as a 0–simplex $\underline{a}^1 \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^1_{[n-1]}\{\boldsymbol{B}\}_0$, of $\widehat{a}^0 \circ w^0 \circ v^0$ as a 0–simplex $\underline{\widehat{a}}^0 \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\Omega\boldsymbol{B}\}_0$, and of $\widehat{F}^1 \circ v^1$ as a 1–simplex $\underline{\widehat{F}}^1 \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^1_{[n-1]}\{\boldsymbol{B}\}_1$, implicitly using the natural isomorphism $i_P \colon \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{P\boldsymbol{B}\} \cong P\mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\boldsymbol{B}\}$ of Remark A.2(b), and the inclusion of $j \colon (PK)_i \subseteq K_{i+1}$ for any Kan complex K and $i \geq 0$ (see [31, Section 23.3]). Thus $\widehat{F}^1 \colon a^1 \circ w^1 \sim 0$ means $d_0 \underline{\widehat{F}}^1 = \underline{a}^1$ and $d_1 \underline{\widehat{F}}^1 = 0$ (simplicial face maps in the mapping space). The fact that the domain of $\widehat{F}^1$ is $C^1 \boldsymbol{W}^{\bullet}_{[n-1]}$ implies that

$$\textbf{(A.21)} \qquad (d^1)^*(\underline{\widehat{F}}^1) = 0 \quad \text{in } \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\boldsymbol{B}\}_1,$$

and (A.15) becomes

$$j\circ\iota\circ i_\Omega(\underline{\widehat{a}}^0)=(d^0)^*(\underline{\widehat{F}}^1)\quad\text{in } \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]}\{\boldsymbol{B}\}_1, \tag{A.22}$$

again using $i_{\boldsymbol{P}}$, the isomorphism i_Ω of Remark A.2(c), the inclusion $\iota\colon \Omega K\hookrightarrow PK$ and $j\colon (PK)_0\subseteq K_1$ as above for $K=\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]}\{\boldsymbol{B}\}$.

Moreover, we have an augmentation of simplicial $\boldsymbol{\Theta}$–mapping algebras

$$\varepsilon_{[n-1]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^\bullet_{[n-1]}\to\mathfrak{X}$$

by Definition 2.1(b) in the outer induction hypothesis. Thus $\varepsilon_{[n-1]}$ satisfies the simplicial identity

$$\varepsilon_{[n-1]}\circ(d^0)^*=\varepsilon_{[n-1]}\circ(d^1)^*. \tag{A.23}$$

In order to apply Lemma 2.19, since $\boldsymbol{Y}$ is not available in cosimplicial dimension -1, we need to verify that $\underline{\widehat{a}}^0$ is a "Moore 0–cycle" in the augmented simplicial $\boldsymbol{\Theta}$–mapping algebra $\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^\bullet_{[n-1]}\to\mathfrak{X}$ — that is,

$$\varepsilon_{[n-1]}(\underline{\widehat{a}}^0)=0\quad\text{in } \mathfrak{X}\{\Omega\boldsymbol{B}\}_0, \tag{A.24}$$

so after postcomposition with the monic maps

$$\mathfrak{X}\{\Omega\boldsymbol{B}\}_0\xrightarrow{\iota_\Omega}\Omega\mathfrak{X}\{\boldsymbol{B}\}_0\xrightarrow{\iota}P\mathfrak{X}\{\boldsymbol{B}\}_0\xrightarrow{j}\mathfrak{X}\{\boldsymbol{B}\}_1$$

it suffices to show

$$j\circ\iota\circ\iota_\Omega\circ\varepsilon_{[n-1]}(\underline{\widehat{a}}^0)=0, \tag{A.25}$$

Note that the following diagram of sets commutes:

$$\begin{array}{ccccc}
 & & \underline{\widehat{F}}^1\in\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^1_{[n-1]}\{\boldsymbol{B}\}_1 & & \\
 & & \downarrow{\scriptstyle (d^0_1)^*} & \searrow^{(d^1_1)^*} & \\
\underline{\widehat{a}}^0\in\mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]}\{\Omega\boldsymbol{B}\}_0 & \xrightarrow{j\circ\iota\circ\iota_\Omega} & \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]}\{\boldsymbol{B}\}_1 & & \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^0_{[n-1]}\{\boldsymbol{B}\}_1 \\
\downarrow{\scriptstyle \varepsilon_{[n-1]}} & & \downarrow{\scriptstyle \varepsilon_{[n-1]}} & \swarrow_{\varepsilon_{[n-1]}} & \\
\mathfrak{X}\{\Omega\boldsymbol{B}\}_0 & \xrightarrow{j\circ\iota\circ\iota_\Omega} & \mathfrak{X}\{\boldsymbol{B}\}_1 & &
\end{array} \tag{A.26}$$

where the left-hand square commutes by naturality of $j\circ\iota\circ\iota_\Omega$, and the right-hand side commutes by (A.23). By (A.22), the simplices $\underline{\widehat{F}}^1$ and $\underline{\widehat{a}}^0$ thus map to the same value in $\mathfrak{X}\{B\}_1$, which is zero by (A.21), proving that (A.25), and thus (A.24), in fact hold.

By the outer induction hypothesis that Definition 2.1(b) holds, the augmented simplicial abelian group

$$V_\bullet\{\Omega \boldsymbol{B}\} \xrightarrow{\varepsilon} \pi_0 \mathfrak{X}\{\Omega \boldsymbol{B}\} = \Gamma\{\Omega^{n-2}\overline{\boldsymbol{W}}^n\} \tag{A.27}$$

is realized by $\boldsymbol{W}^\bullet_{[n-1]}$ through simplicial dimension $n-1$, so that

$$V_k\{\Omega \boldsymbol{B}\} \cong \pi_0 \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^k_{[n-1]}\{\Omega \boldsymbol{B}\} = [\boldsymbol{W}^k_{[n-1]}, \Omega \boldsymbol{B}] \tag{A.28}$$

for $k = 1$ and 0.

Since, by (A.24), $[\underline{\hat{a}}^0]$ is a 0–cycle in (A.27), which is acyclic, there is a Moore 1–chain $\gamma_1 \in V_1\{\Omega \boldsymbol{B}\}$ with $\partial_0^{V_1}(\gamma_1) = [\hat{a}^0]$. As in (A.18) of Step V, this can be used to produce a map $F^1\colon C^1 \boldsymbol{W}^\bullet_{[n-1]} \to P\boldsymbol{B}$ such that

$$F^1 \circ d^1 = 0, \tag{A.29}$$

yielding $a^0\colon \boldsymbol{W}^0_{[n-1]} \to \Omega \boldsymbol{B}$ (by Step III) having a nullhomotopy $\widehat{F}^0\colon a^0 \sim 0$ (again, as in Step V). Moreover, (A.29) and (A.23) together imply that (after replacing $\widehat{F}^1$ by F^1 in (A.26)) one can deduce that

$$\varepsilon_{[n-1]}(\underline{a}^0) = 0 \quad \text{in } \mathfrak{X}\{\Omega \boldsymbol{B}\}, \tag{A.30}$$

as required.

VII Step $k = 0$ of the descending induction For $k = 0$, let $\boldsymbol{C} := \Omega^{n-2}\overline{\boldsymbol{W}}^n$. The nullhomotopy $\widehat{F}^0\colon a^0 \circ w^0 \sim 0$ is represented as in Step VI by a 1–simplex $\underline{\widehat{F}}^0 \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\boldsymbol{C}\}_1$ with $d_0 \underline{\widehat{F}}^0 = \underline{a}^0$, and $d_1 \underline{\widehat{F}}^0 = 0$. Thus for $\varepsilon_{[n-1]}(\underline{\widehat{F}}^0) \in \mathfrak{X}\{\boldsymbol{C}\}_1$ we have $d_0(\varepsilon_{[n-1]}(\underline{\widehat{F}}^0)) = \varepsilon_{[n-1]}(\underline{a}^0) = 0$, by (A.30), as well as $d_1(\varepsilon_{[n-1]}(\underline{\widehat{F}}^0)) = 0$.

Therefore, as in (A.22) we obtain $\underline{\hat{a}} \in \mathfrak{X}\{\Omega \boldsymbol{C}\}_0$ with $j \circ \iota \circ \iota_\Omega(\underline{\hat{a}}) = \varepsilon_{[n-1]}(\underline{\widehat{F}}^0)$.

Note that this argument fails if we do not have the mapping algebra $\mathfrak{X}$ (with $\boldsymbol{\Theta}$ as in Remark A.2), which allows us to obtain an element $[\underline{\hat{a}}] \in \Gamma\{\Omega \boldsymbol{C}\} = \Gamma\{\Omega^{n-1}\overline{\boldsymbol{W}}^n\}$ from the nullhomotopy $\widehat{F}^0\colon \boldsymbol{W}^0_{[n-1]} \to P\boldsymbol{C}$.

Again we use the acyclicity of (A.27) to deduce the existence of a Moore 0–chain $\gamma_0 \in V_0\{\Omega \boldsymbol{C}\}$ with

$$\varepsilon(\gamma_0) = [\underline{\hat{a}}] \in \Gamma\{\Omega \boldsymbol{C}\}. \tag{A.31}$$

We no longer require Lemma 2.22 to deduce that we can represent γ_0 by a map $g^0\colon \boldsymbol{W}^0_{[n-1]} \to \Omega \boldsymbol{C}$, corresponding to a 0–simplex $\underline{g}^0 \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\Omega \boldsymbol{C}\}_0$, and thus

a 1–simplex

$$\textbf{(A.32)} \qquad \underline{H} := (j \circ \iota \circ \iota_\Omega)(\underline{g}^0) \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\boldsymbol{C}\}_1 \quad \text{with } d_0\underline{H} = d_1\underline{H} = 0.$$

Since $\mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\boldsymbol{C}\}$ is a homotopy group object in $\mathcal{S}_*$, with homotopy group structure $\star$ and inverse $(-)^{-1}$ induced from those of $\boldsymbol{C} \in \Theta \subset \mathcal{C}$, we may define a new 1–simplex

$$\textbf{(A.33)} \quad \underline{F}^0 := \widehat{\underline{F}}^0 \star \underline{H}^{-1} \in \mathfrak{M}_{\boldsymbol{\Theta}} \boldsymbol{W}^0_{[n-1]}\{\boldsymbol{C}\}_1 \quad \text{with } d_0\underline{F}^0 = \underline{a}^0 \text{ and } d_1\underline{F}^0 = 0.$$

Since $\varepsilon_{[n-1]}(\underline{a}^0) = 0$, the 1–simplex $\varepsilon_{[n-1]}(\underline{F}^0) \in \mathfrak{X}\{\boldsymbol{C}\}_1$ is in the image of the composite inclusion

$$\mathfrak{X}\{\Omega\boldsymbol{C}\}_0 \xrightarrow{\iota_\Omega} \Omega\mathfrak{X}\{\boldsymbol{C}\}_0 \xrightarrow{\iota} P\mathfrak{X}\{\boldsymbol{C}\}_0 \xrightarrow{j} \mathfrak{X}\{\boldsymbol{C}\}_1.$$

Thus, there is a 0–simplex $\underline{a} \in \mathfrak{X}\{\Omega\boldsymbol{C}\}_0$ with

$$\textbf{(A.34)} \qquad (j \circ \iota \circ \iota_\Omega)(\underline{a}) = \varepsilon_{[n-1]}(\underline{F}^0).$$

Moreover, since $(j \circ \iota \circ \iota_\Omega)$ and $\varepsilon_{[n-1]}$ commute with the homotopy group structure,

$$(j\circ\iota\circ\iota_\Omega)(\underline{a}) = \varepsilon_{[n-1]}(\widehat{\underline{F}}^0)\star(\varepsilon_{[n-1]}(\underline{H})^{-1}) = [(j\circ\iota\circ\iota_\Omega)(\widehat{\underline{a}})]\star[(j\circ\iota\circ\iota_\Omega)(\varepsilon_{[n-1]}\underline{g}^0)]^{-1},$$

by (A.32) and (A.34). Since $j \circ \iota \circ \iota_\Omega$ is monic, this implies that

$$\textbf{(A.35)} \qquad \underline{a} = \widehat{\underline{a}} \star (\varepsilon_{[n-1]}\underline{g}^0)^{-1}.$$

By (A.31) there is a 1–simplex $G \in \mathfrak{X}\{\Omega\boldsymbol{C}\}$ with $d_0G = \widehat{\underline{a}}$ and $d_1G = \varepsilon_{[n-1]}(\underline{g}^0)$ in $\mathfrak{X}\{\Omega\boldsymbol{C}\}_0$. Therefore, $G' := G \star (s_0\varepsilon_{[n-1]}(\underline{g}^0))^{-1} \in \mathfrak{X}\{\Omega\boldsymbol{C}\}_1$ satisfies

$$\textbf{(A.36)} \quad d_0G' = \widehat{\underline{a}} \star (\varepsilon_{[n-1]}(\underline{g}^0))^{-1} \quad \text{and} \quad d_1G' = \varepsilon_{[n-1]}(\underline{g}^0) \star (\varepsilon_{[n-1]}(\underline{g}^0))^{-1}.$$

Since $\Omega\boldsymbol{C}$ is a homotopy group object, there is a 1–simplex $K \in \mathfrak{X}\{\Omega\boldsymbol{C}\}_1$ with

$$\textbf{(A.37)} \qquad d_0K = \varepsilon_{[n-1]}(\underline{g}^0) \star (\varepsilon_{[n-1]}(\underline{g}^0))^{-1} \quad \text{and} \quad d_1K = 0.$$

Since $\mathfrak{X}\{\Omega\boldsymbol{C}\}$ is a Kan complex, we have $\sigma \in \mathfrak{X}\{\Omega\boldsymbol{C}\}_2$ with $d_0\sigma = G'$ and $d_1\sigma = K$, so if we set $\widehat{\underline{F}} := d_2\sigma$, we find

$$\textbf{(A.38)} \qquad d_0\widehat{\underline{F}} = \widehat{\underline{a}} \star (\varepsilon_{[n-1]}(\underline{g}^0))^{-1} \quad \text{and} \quad d_1\widehat{\underline{F}} = 0.$$

We deduce from (A.35) that $\widehat{\underline{F}}$ is a nullhomotopy for $\underline{a}$ in $\mathfrak{X}\{\Omega\boldsymbol{C}\}$, as required in Step III, thus completing the construction of the restricted cosimplicial object $\widetilde{\boldsymbol{W}}^\bullet_{[n]}$.

Since making the cosimplicial object $\widehat{\boldsymbol{W}}^{\bullet}_{[k]}$ as in Section 2.5(iii) Reedy fibrant or cofibrant does not change $\widehat{\boldsymbol{W}}^{0}_{[k]}$, we see by induction from (2.17) that

$$\widetilde{\boldsymbol{W}}^{0}_{[n]} = \boldsymbol{W}^{0}_{[n-1]} \times P\Omega\boldsymbol{C} = \prod_{k=0}^{n} P\Omega^{k-1}\overline{\boldsymbol{W}}^{k} \tag{A.39}$$

(using the convention of (2.7)). Note that all the factors on the right-hand side of (A.39) are in $\boldsymbol{\Theta}$ by Remark A.2, so $\widetilde{\boldsymbol{W}}^{0}_{[n]}$ is, too.

Thus, by Lemma A.4, the augmentation $\widetilde{\varepsilon}_{[n]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\widetilde{\boldsymbol{W}}^{0}_{[n]} \to \mathfrak{X}$ is determined by the choice of a suitable 0–simplex

$$e = (e', e'') \in \mathfrak{X}\{\widetilde{\boldsymbol{W}}^{0}_{[n]}\}_0 = \mathfrak{X}\{\boldsymbol{W}^{0}_{[n-1]} \times P\Omega\boldsymbol{C}\}_0 = \mathfrak{X}\{\boldsymbol{W}^{0}_{[n-1]}\}_0 \times P\mathfrak{X}\{\Omega\boldsymbol{C}\}_0$$

by (A.10), using Remark A.2(a), where the component $e' \in \mathfrak{X}\{\boldsymbol{W}^{0}_{[n-1]}\}$ corresponds to the given $\varepsilon_{[n-1]}\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{0}_{[n-1]} \to \mathfrak{X}$.

On the other hand, as before, e'' corresponds to the nullhomotopy $\underline{\hat{F}} \in \mathfrak{X}\{\Omega\boldsymbol{C}\}_1$, which actually lands in $P\mathfrak{X}\{\Omega\boldsymbol{C}\}_0$.

The cosimplicial identity

$$\widetilde{\varepsilon}_{[n]} \circ (d^0_0)^* = \widetilde{\varepsilon}_{[n]} \circ (d^1_0)^*\colon \mathfrak{M}_{\boldsymbol{\Theta}}\widetilde{\boldsymbol{W}}^{1}_{[n]} \to \mathfrak{X} \tag{A.40}$$

may be verified using Lemma A.4 by representing both sides of the equation by elements in

$$\mathfrak{X}\{\widetilde{\boldsymbol{W}}^{1}_{[n]}\}_0 = \mathfrak{X}\{\boldsymbol{W}^{1}_{[n-1]}\}_0 \times \mathfrak{X}\{P\boldsymbol{C}\}_0,$$

where the components in $\mathfrak{X}\{\boldsymbol{W}^{1}_{[n-1]}\}_0$ agree because $(\boldsymbol{\varepsilon}_{[n-1]})^*\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{\bullet}_{[n-1]} \to \mathfrak{X}$ is an augmented simplicial $\boldsymbol{\Theta}$–mapping algebra, by the outer induction hypothesis. Here we are using the last assumption in Remark A.2 to deduce that $\boldsymbol{W}^{1}_{[n-1]} \in \boldsymbol{\Theta}$.

Thus it remains to identify the corresponding components in $\mathfrak{X}\{P\boldsymbol{C}\}_0$. On the one hand, $q^1_{[n]} \circ d^0_0\colon \widetilde{\boldsymbol{W}}^{0}_{[n-1]} \to P\boldsymbol{C}$ is the map $F^0 \circ \psi^0_{[n]}$, and $\widetilde{\varepsilon}_{[n]} \circ (\psi^0_{[n]})^*\colon \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{0}_{[n-1]} \to \mathfrak{X}$ is the given $\varepsilon_{[n-1]}$ by construction, and since $(F^0)^*\colon \mathfrak{M}_{\boldsymbol{\Theta}}P\boldsymbol{C} \to \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{0}_{[n-1]}$ is represented according to Lemma A.4 by $\underline{F}^0 \in \mathfrak{M}_{\boldsymbol{\Theta}}\boldsymbol{W}^{0}_{[n-1]}\{P\boldsymbol{C}\}$, we see that the composite $\widetilde{\varepsilon}_{[n]} \circ (q^1_{[n]} \circ d^0_0)^*$ is represented in $\mathfrak{X}\{P\boldsymbol{C}\}_0$ by

$$\varepsilon_{[n-1]}(\underline{F}^0) = (j \circ \iota_{\Omega})(\underline{a}),$$

according to (A.34).

On the other hand, $q^1_{[n]} \circ d^1_0 \colon \widetilde{\boldsymbol{W}}^0_{[n-1]} \to P\boldsymbol{C}$ equals $\iota \circ p \circ q^0_{[n]}$, so

$$\widetilde{\varepsilon}_{[n]} \circ (d^1_0)^* = \widetilde{\varepsilon}_{[n]} \circ (q^0_{[n]})^* \circ (\iota \circ p)^*,$$

where $\widetilde{\varepsilon}_{[n]} \circ (q^0_{[n]})^* \colon \mathfrak{M}_{\boldsymbol{\Theta}} P\Omega\boldsymbol{C} \to \mathfrak{X}$ is represented by $\underline{\widehat{F}} \in \mathfrak{X}\{\Omega\boldsymbol{C}\}_1$, so

$$\widetilde{\varepsilon}_{[n]} \circ (q^0_{[n]})^* \circ (\iota \circ p)^*$$

is represented by

$$(j \circ \iota_\Omega)(d_0 \underline{\widehat{F}}) = (j \circ \iota_\Omega)(\underline{a}) \in \mathfrak{X}\{P\boldsymbol{C}\}_0$$

as above.

This completes the nth step of the outer induction in case (1).

VIII Making $\mathcal{W}$ cofibrant in case (2) Note that when we represent the attaching map $\overline{\partial}^n_0 \colon \overline{V}_n \to Z_{n-1}V_\bullet$ for $V_\bullet$ by a map $\phi \colon \overline{V}_n \otimes_* S^{n-1} \to C_*V_\bullet$ as in Remark 1.14, ϕ extends to a map of *augmented* $(n-1)$–truncated chain complexes in Θ–Alg, ending in $C_{-1}V_\bullet := \Gamma$. Moreover, in case (2) the cochain map $F \colon C^*\boldsymbol{W}^\bullet_{[n-1]} \to \boldsymbol{D}^*$ realizing ϕ which we have constructed in the outer induction above also extends to a map of coaugmented $(n-1)$–truncated cochain complexes, with $C^{-1}\boldsymbol{W}^\bullet_{[n-1]} := \boldsymbol{Y}$, δ^{-1} given by $\boldsymbol{\varepsilon}_{[n-1]}$, and F^{-1} given by Lemma 4.5.

In the left Reedy model category structure on the category $\mathsf{Ch}^{\mathcal{C}}_{-1\le *\le n-1}$ of $(n-1)$–truncated coaugmented cochain complexes in $\mathcal{C}$, factor F as a cofibration

$$G \colon C^*\boldsymbol{W}^\bullet_{[n-1]} \hookrightarrow \boldsymbol{E}^*$$

followed by a trivial fibration $T \colon \boldsymbol{E}^* \twoheadrightarrow \boldsymbol{D}^*$. Applying the functorial construction of Definition 2.1(c) to G yields $\boldsymbol{W}^\bullet_{[n]}$ as required for the cofibrant sequential realization $\mathcal{W}$. The vertical trivial fibrations of (2.9) are induced by T. □

References

[1] **J F Adams**, *On the non-existence of elements of Hopf invariant one*, Ann. of Math. 72 (1960) 20–104 MR

[2] **H-J Baues**, **D Blanc**, *Comparing cohomology obstructions*, J. Pure Appl. Algebra 215 (2011) 1420–1439 MR

[3] **H-J Baues**, **D Blanc**, **S Gondhali**, *Higher Toda brackets and Massey products*, J. Homotopy Relat. Struct. 11 (2016) 643–677 MR

[4] **H-J Baues**, **M Jibladze**, *Suspension and loop objects in theories and cohomology*, Georgian Math. J. 8 (2001) 697–712 MR

[5] **G Biedermann**, **G Raptis**, **M Stelzer**, *The realization space of an unstable coalgebra*, preprint (2014) arXiv

[6] **D Blanc**, *Higher homotopy operations and the realizability of homotopy groups*, Proc. London Math. Soc. 70 (1995) 214–240 MR

[7] **D Blanc**, *CW simplicial resolutions of spaces with an application to loop spaces*, Topology Appl. 100 (2000) 151–175 MR

[8] **D Blanc**, *Realizing coalgebras over the Steenrod algebra*, Topology 40 (2001) 993–1016 MR

[9] **D Blanc**, *Homotopy operations and rational homotopy type*, from "Categorical decomposition techniques in algebraic topology" (G Arone, J Hubbuck, R Levi, M Weiss, editors), Progr. Math. 215, Birkhäuser, Basel (2004) 47–75 MR

[10] **D Blanc**, **W G Dwyer**, **P G Goerss**, *The realization space of a Π–algebra: a moduli problem in algebraic topology*, Topology 43 (2004) 857–892 MR

[11] **D Blanc**, **M W Johnson**, **J M Turner**, *On realizing diagrams of Π–algebras*, Algebr. Geom. Topol. 6 (2006) 763–807 MR

[12] **D Blanc**, **M W Johnson**, **J M Turner**, *Higher homotopy operations and cohomology*, J. K–Theory 5 (2010) 167–200 MR

[13] **D Blanc**, **M W Johnson**, **J M Turner**, *Higher homotopy operations and André–Quillen cohomology*, Adv. Math. 230 (2012) 777–817 MR

[14] **D Blanc**, **M W Johnson**, **J M Turner**, *Higher invariants for spaces and maps*, preprint (2015)

[15] **D Blanc**, **M Markl**, *Higher homotopy operations*, Math. Z. 245 (2003) 1–29 MR

[16] **D Blanc**, **G Peschke**, *The fiber of functors between categories of algebras*, J. Pure Appl. Algebra 207 (2006) 687–715 MR

[17] **D Blanc**, **D Sen**, *Mapping spaces and R–completion*, preprint (2013) arXiv

[18] **F Borceux**, *Handbook of categorical algebra, II: Categories and structures*, Encyclopedia of Mathematics and its Applications 51, Cambridge Univ. Press (1994) MR

[19] **A K Bousfield**, *Cosimplicial resolutions and homotopy spectral sequences in model categories*, Geom. Topol. 7 (2003) 1001–1053 MR

[20] **A K Bousfield**, **D M Kan**, *Homotopy limits, completions and localizations*, Lecture Notes in Mathematics 304, Springer (1972) MR

[21] **W Chachólski**, **J Scherer**, *Homotopy theory of diagrams*, Mem. Amer. Math. Soc. 736, Amer. Math. Soc., Providence, RI (2002) MR

[22] **C Ehresmann**, *Esquisses et types des structures algébriques*, Bul. Inst. Politehn. Iaşi 14 (1968) 1–14 MR

[23] **Y Félix**, *Modèles bifiltrés: une plaque tournante en homotopie rationnelle*, Canad. J. Math. 33 (1981) 1448–1458 MR

[24] **Y Félix**, **S Halperin**, **J-C Thomas**, *Rational homotopy theory*, Graduate Texts in Mathematics 205, Springer (2001) MR

[25] **P G Goerss**, **J F Jardine**, *Simplicial homotopy theory*, Progress in Mathematics 174, Birkhäuser, Basel (1999) MR

[26] **J R Harper**, *Secondary cohomology operations*, Graduate Studies in Mathematics 49, Amer. Math. Soc., Providence, RI (2002) MR

[27] **K Hess**, *Rational homotopy theory: a brief introduction*, from "Interactions between homotopy theory and algebra" (L L Avramov, J D Christensen, W G Dwyer, M A Mandell, B E Shipley, editors), Contemp. Math. 436, Amer. Math. Soc., Providence, RI (2007) 175–202 MR

[28] **P S Hirschhorn**, *Model categories and their localizations*, Mathematical Surveys and Monographs 99, Amer. Math. Soc., Providence, RI (2003) MR

[29] **F W Lawvere**, *Functorial semantics of algebraic theories*, Proc. Nat. Acad. Sci. U.S.A. 50 (1963) 869–872 MR

[30] **C R F Maunder**, *Cohomology operations of the Nth kind*, Proc. London Math. Soc. 13 (1963) 125–154 MR

[31] **J P May**, *Simplicial objects in algebraic topology*, Van Nostrand Mathematical Studies 11, Van Nostrand, Princeton, NJ (1967) MR

[32] **D G Quillen**, *Homotopical algebra*, Lecture Notes in Mathematics 43, Springer (1967) MR

[33] **D Quillen**, *Rational homotopy theory*, Ann. of Math. 90 (1969) 205–295 MR

[34] **L Schwartz**, *Unstable modules over the Steenrod algebra and Sullivan's fixed point set conjecture*, Univ. of Chicago Press (1994) MR

[35] **E Spanier**, *Secondary operations on mappings and cohomology*, Ann. of Math. 75 (1962) 260–282 MR

[36] **E Spanier**, *Higher order operations*, Trans. Amer. Math. Soc. 109 (1963) 509–539 MR

[37] **H Toda**, *Composition methods in homotopy groups of spheres*, Annals of Mathematics Studies 49, Princeton Univ. Press (1962) MR

[38] **C A Weibel**, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics 38, Cambridge Univ. Press (1994) MR

Department of Mathematics, University of Haifa
Haifa, Israel
Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur
Kanpur, India
blanc@math.haifa.ac.il, debasis@iitk.ac.in

Received: 5 September 2016 Revised: 4 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 313–332

On high-dimensional representations of knot groups

STEFAN FRIEDL
MICHAEL HEUSENER

Given a hyperbolic knot K and any $n \geq 2$ the abelian representations and the holonomy representation each give rise to an $(n-1)$–dimensional component in the $\mathrm{SL}(n,\mathbb{C})$–character variety. A component of the $\mathrm{SL}(n,\mathbb{C})$–character variety of dimension $\geq n$ is called high-dimensional.

It was proved by D Cooper and D Long that there exist hyperbolic knots with high-dimensional components in the $\mathrm{SL}(2,\mathbb{C})$–character variety. We show that given any nontrivial knot K and sufficiently large n the $\mathrm{SL}(n,\mathbb{C})$–character variety of K admits high-dimensional components.

57M25, 57M27; 57M50

1 Introduction

Given a knot $K \subset S^3$ we denote by $E_K = S^3 \setminus \nu K$ the knot exterior and we write $\pi_K = \pi_1(E_K)$. Furthermore, given a group G and $n \in \mathbb{N}$ we denote by $X(G, \mathrm{SL}(n,\mathbb{C}))$ the $\mathrm{SL}(n,\mathbb{C})$–character variety. We recall the precise definition in Section 2. It is straightforward to see that the abelian representations of a knot group π_K give rise to an $(n-1)$–dimensional subvariety of $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$ consisting solely of characters of abelian representations (see [17, Section 2]).

If K is hyperbolic, then we denote by $\widetilde{\mathrm{Hol}}\colon \pi_K \to \mathrm{SL}(2,\mathbb{C})$ a lift of the holonomy representation. For $n \geq 2$ we denote by $\zeta_n\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(n,\mathbb{C})$ the, up to conjugation, unique rational irreducible representation of $\mathrm{SL}(2,\mathbb{C})$. P Menal-Ferrer and J Porti [20; 19] showed that for any n, the representation $\rho_n := \zeta_n \circ \widetilde{\mathrm{Hol}}$ is a smooth point of the $\mathrm{SL}(n,\mathbb{C})$–representation variety $R(\pi_k, \mathrm{SL}(n,\mathbb{C}))$. Moreover, Menal-Ferrer and Porti proved that the corresponding character χ_{ρ_n} is a smooth point on the character variety $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$; in particular it is contained in a unique component of dimension $n-1$ [20, Theorem 0.4]. Also, the deformations of reducible representations studied in [16] and [2] give rise to $(n-1)$–dimensional components in the character variety $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.313

The above discussion shows that the character variety $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$ of any given knot K contains an $(n-1)$–dimensional subvariety consisting of abelian representations, and if K is hyperbolic, $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$ also contains an $(n-1)$–dimensional subvariety that contains characters of irreducible representations. Furthermore E Falbel and A Guilloux [13] showed that every component of $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$ satisfying a mild technical hypothesis is of dimension at least $n-1$.

This motivates the following definition. Given a knot K we say that a component of $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$ is *high-dimensional* if its dimension is greater than $n-1$. We summarize some known facts about the existence and nonexistence of high-dimensional components of character varieties of knot groups:

- For $n=3$ and K a nonalternating torus knot, the variety $X(\pi_K, \mathrm{SL}(3,\mathbb{C}))$ has 3–dimensional components, whereas for alternating torus knots, $X(\pi_K, \mathrm{SL}(3,\mathbb{C}))$ has only 2–dimensional components. In particular for alternating torus knots, the variety $X(\pi_K, \mathrm{SL}(3,\mathbb{C}))$ does not contain any high-dimensional components. For more details see [21, Theorem 1.1].
- For $n=3$ and $K=4_1$ the variety $X(\pi_K, \mathrm{SL}(3,\mathbb{C}))$ has five 2–dimensional components. Three of the five components contain characters of irreducible representations. There are no high-dimensional components. (See [17, Theorem 1.2].)
- It was proved by D Cooper and D Long [9, Section 8] that for a given n there exists an alternating hyperbolic knot K_n in S^3 such that the $\mathrm{SL}(2,\mathbb{C})$–character variety admits a component of dimension at least n.

The main result of this note is the proof that the $\mathrm{SL}(n,\mathbb{C})$–character variety of every nontrivial knot admits high-dimensional components for n sufficiently large. More precisely, building on Cooper, Long, and A Reid [10, Theorem 1.3] we prove the following theorem.

Theorem 1.1 *Let $K \subset S^3$ be a nontrivial knot. Then for all $N \in \mathbb{N}$ there exists an $n \geq N$ such that the character variety $X(\pi_K, \mathrm{SL}(n,\mathbb{C}))$ contains a high-dimensional component.*

Given a group G we now denote by $X^{\mathrm{irr}}(G, \mathrm{SL}(n,\mathbb{C}))$ the character variety corresponding to irreducible representations. We refer the reader to Section 3 for the precise definition. The following is then a more refined version of Theorem 1.1.

Theorem 1.2 *Let $K \subset S^3$ be a nontrivial knot. Then given any $N \in \mathbb{N}$ there exists a $p \geq N$ such that $X^{\mathrm{irr}}(\pi_K, \mathrm{SL}(p,\mathbb{C}))$ contains a high-dimensional component.*

Remark It is straightforward to generalize the proofs of the above theorems to any knot K in a rational homology sphere Σ such that $\Sigma \setminus \nu K \neq S^1 \times D^2$ and such that $\Sigma \setminus \nu K$ is irreducible. We restrict ourselves in the exposition to knots in S^3, to simplify the notation and to make it easier to compare our results to earlier results.

In the case of the figure-eight knot we obtain in Section 4.2 a refined quantitative result:

Proposition 1.3 *Let $K \subset S^3$ be the figure-eight knot. Then for all $n \in \mathbb{N}$ the representation variety $X(\pi_K, \mathrm{SL}(10n, \mathbb{C}))$ has a component C of dimension at least $4n^2 - 1$. Moreover, C contains characters of irreducible representations.*

Remark For a free group F_r we have $\dim X(F_r, \mathrm{SL}(n, \mathbb{C}))/(n^2-1) = (r-1)$; hence

$$\limsup_{n\to\infty} \frac{\dim X(G, \mathrm{SL}(n, \mathbb{C}))}{n^2 - 1} \leq (r-1)$$

if G is generated by r elements. It follows from Proposition 1.3 that for the figure-eight knot $K = 4_1$ the following inequality holds:

$$\tfrac{1}{25} \leq \limsup_{n\to\infty} \frac{\dim X(\pi_K, \mathrm{SL}(n, \mathbb{C}))}{n^2 - 1} \leq 1.$$

At the end of Section 4.2 we discuss the question of the extent to which the result of Proposition 1.3 can be generalized to other 2–bridge knots.

2 Representation and character varieties

Before proving Theorem 1.1, we recall some definitions and facts. The general reference for representation and character varieties is the book [18] of A Lubotzky and A Magid.

Given two representations $\rho_1\colon G \to \mathrm{GL}(n_1, \mathbb{C})$ and $\rho_2\colon G \to \mathrm{GL}(n_2, \mathbb{C})$ we define the *direct sum* $\rho_1 \oplus \rho_2\colon G \to \mathrm{GL}(n_1 + n_2, \mathbb{C})$ by

$$(\rho_1 \oplus \rho_2)(\gamma) = \left(\begin{array}{c|c} \rho_1(\gamma) & 0 \\ \hline 0 & \rho_2(\gamma) \end{array}\right).$$

We also define the *tensor product* $\rho_1 \otimes \rho_2\colon G \to \mathrm{GL}(n_1 \cdot n_2, \mathbb{C}) = \mathrm{Aut}(\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2})$ by

$$((\rho_1 \otimes \rho_2)(\gamma))(v_1 \otimes v_2) = \rho_1(\gamma)(v_1) \otimes \rho_2(\gamma)(v_2).$$

Definition We call a representation $\rho\colon G \to \mathrm{GL}(n, \mathbb{C})$ *reducible* if there exists a nontrivial, proper subspace $V \subset \mathbb{C}^n$ such that V is $\rho(G)$–stable. The representation ρ is called *irreducible* or *simple* if it is not reducible. A *semisimple* representation is a direct sum of simple representations.

Let $G = \langle g_1, \dots, g_r \rangle$ be a finitely generated group. A $\mathrm{SL}(n,\mathbb{C})$–representation is a homomorphism $\rho\colon G \to \mathrm{SL}(n,\mathbb{C})$. The $\mathrm{SL}(n,\mathbb{C})$–representation variety is

$$R(G,\mathrm{SL}(n,\mathbb{C})) = \mathrm{Hom}(G,\mathrm{SL}(n,\mathbb{C})) \subset \mathrm{SL}(n,\mathbb{C})^r \subset M_n(\mathbb{C})^r \cong \mathbb{C}^{n^2 r}.$$

The representation variety $R(G,\mathrm{SL}(n,\mathbb{C}))$ is an affine algebraic set. It is contained in $\mathrm{SL}(n,\mathbb{C})^r$ via the inclusion $\rho \mapsto (\rho(g_1),\dots,\rho(g_r))$, and it is the set of solutions of a system of polynomial equations in the matrix coefficients.

The group $\mathrm{SL}(n,\mathbb{C})$ acts by conjugation on $R(G,\mathrm{SL}(n,\mathbb{C}))$. More precisely, for $A \in \mathrm{SL}(n,\mathbb{C})$ and $\rho \in R(G,\mathrm{SL}(n,\mathbb{C}))$ we define $(A.\rho)(g) = A\rho(g)A^{-1}$ for all $g \in G$. In what follows we will write $\rho \sim \rho'$ if there exists an $A \in \mathrm{SL}(n,\mathbb{C})$ such that $\rho' = A.\rho$, and we will call ρ and ρ' *equivalent*. For $\rho \in R(G,\mathrm{SL}(n,\mathbb{C}))$ we define its *character* $\chi_\rho\colon G \to \mathbb{C}$ by $\chi_\rho(\gamma) = \mathrm{tr}(\rho(\gamma))$. We have $\rho \sim \rho' \Rightarrow \chi_\rho = \chi_{\rho'}$. Moreover, if ρ and ρ' are semisimple, then $\rho \sim \rho'$ if and only if $\chi_\rho = \chi_{\rho'}$. (See Theorems 1.27 and 1.28 in the book [18] of Lubotzky and Magid.)

The *algebraic quotient* or *GIT quotient* for the action of $\mathrm{SL}(n,\mathbb{C})$ on $R(G,\mathrm{SL}(n,\mathbb{C}))$ is called the *character variety*. This quotient will be denoted by $X(G,\mathrm{SL}(n,\mathbb{C})) = R(G,\mathrm{SL}(n,\mathbb{C}))/\!\!/\,\mathrm{SL}(n,\mathbb{C})$. The character variety is not necessarily an irreducible affine algebraic set. Work of C Procesi [23] implies that there exists a finite number of group elements $\{\gamma_i \mid 1 \le i \le M\} \subset G$ such that the image of $t\colon R(G,\mathrm{SL}(n,\mathbb{C})) \to \mathbb{C}^M$ given by

$$t(\rho) = (\chi_\rho(\gamma_1),\dots,\chi_\rho(\gamma_M))$$

can be identified with the affine algebraic set $X(G,\mathrm{SL}(n,\mathbb{C})) \cong t(R(G,\mathrm{SL}(n,\mathbb{C})))$; see also [18, page 27]. This justifies the name *character variety*. For an introduction to algebraic invariant theory see I Dolgachev's book [12]. For a brief introduction to $\mathrm{SL}(n,\mathbb{C})$–representation and character varieties of groups see [15].

Example 2.1 For a free group F_r of rank r we have $R(F_r,\mathrm{SL}(n,\mathbb{C})) \cong \mathrm{SL}(n,\mathbb{C})^r$ is an irreducible algebraic variety of dimension $r(n^2-1)$, and the dimension of the character variety $X(F_k,\mathrm{SL}(n,\mathbb{C}))$ is $(r-1)(n^2-1)$.

The first homology group of the knot exterior is isomorphic to $\mathbb{Z}$. A canonical surjection $\varphi\colon \pi_K \to \mathbb{Z}$ is given by $\varphi(\gamma) = \mathrm{lk}(\gamma, K)$ where lk denotes the linking number in S^3 (see [7, 3.B]). Hence, every *abelian* representation of a knot group π_K factors through $\varphi\colon \pi_K \to \mathbb{Z}$. Here, we call ρ *abelian* if its image is abelian. Therefore, we obtain for each nonzero complex number $\eta \in \mathbb{C}^*$ an abelian representation

$\eta^{\varphi}\colon \pi_K \to \mathrm{GL}(1,\mathbb{C}) = \mathbb{C}^*$ given by $\gamma \mapsto \eta^{\varphi(\gamma)}$. Notice that a 1–dimensional representation is always irreducible.

Let W be a finite dimensional $\mathbb{C}$–vector space. For every representation $\rho\colon G \to \mathrm{GL}(W)$ the vector space W turns into a left $\mathbb{C}[G]$–module via ρ. This $\mathbb{C}[G]$–module will be denoted by W_ρ or simply W if no confusion can arise. Notice that every finite dimensional $\mathbb{C}$–vector space W which is a left $\mathbb{C}[G]$–module gives a representation $\rho\colon G \to \mathrm{GL}(W)$, and by fixing a basis of W we obtain a matrix representation.

The following lemma follows from Proposition 1.7 in [18] and the discussion therein.

Lemma 2.2 *Any group epimorphism $\alpha\colon G \twoheadrightarrow F$ between finitely generated groups induces a closed embedding $R(F,\mathrm{SL}(n,\mathbb{C})) \hookrightarrow R(G,\mathrm{SL}(n,\mathbb{C}))$ of algebraic varieties, and an injection*

$$X(F,\mathrm{SL}(n,\mathbb{C})) \hookrightarrow X(G,\mathrm{SL}(n,\mathbb{C})).$$

Let $H \leq G$ be a subgroup of finite index. Then the restriction of a representation $\rho\colon G \to \mathrm{SL}(n,\mathbb{C})$ to H will be denoted by $\mathrm{res}^G_H \rho$ or simply by $\rho|_H$ if no confusion can arise. This restriction is compatible with the action by conjugation and it induces a regular map $\nu\colon X(G,\mathrm{SL}(n,\mathbb{C})) \to X(H,\mathrm{SL}(n,\mathbb{C}))$. In what follows we will make use of the following result of A S Rapinchuk which follows directly from [24, Lemma 1].

Lemma 2.3 *If $H \leq G$ is a subgroup of finite index k, then*

$$\nu\colon X(G,\mathrm{SL}(n,\mathbb{C})) \to X(H,\mathrm{SL}(n,\mathbb{C}))$$

has finite fibers.

2.1 The induced representation

Let G be a group and let $H \leq G$ be a subgroup of finite index k. Given a representation $\alpha\colon H \to \mathrm{GL}(m,\mathbb{C})$ we refer to the representation of G that is given by left multiplication by G on

$$\mathbb{C}[G] \otimes_{\mathbb{C}[H]} \mathbb{C}^m$$

as the induced representation. We denote by $e_1,\dots,e_m$ the standard basis of $\mathbb{C}^m$ and we pick representatives $g_1,\dots,g_k$ of G/H. It is straightforward to see that $g_i \otimes e_j$ with $i \in \{1,\dots,k\}$ and $j \in \{1,\dots,m\}$ form a basis for $\mathbb{C}[G]\otimes_{\mathbb{C}[H]}\mathbb{C}^m$ as a complex vector space. Using the ordered basis

$$g_1 \otimes e_1,\ \dots,\ g_1 \otimes e_m,\ \dots,\ g_k \otimes e_1,\ \dots,\ g_k \otimes e_m,$$

the induced representation can be viewed as a representation $\mathrm{ind}^G_H \alpha\colon G \to \mathrm{GL}(mk,\mathbb{C})$.

If $\alpha\colon H \to \mathrm{SL}(m,\mathbb{C})$ is a representation into the special linear group, then for $g \in G$ a priori the determinant of $\mathrm{ind}_H^G \alpha(g)$ is in $\{\pm 1\}$. But it is straightforward to see that if m is even, then $\mathrm{ind}_H^G \alpha$ defines in fact a representation $G \to \mathrm{SL}(mk,\mathbb{C})$.

Lemma 2.4 *Let m be even and $H \leq G$ a subgroup of finite index k. Then the map*

$$\iota\colon R(H,\mathrm{SL}(m,\mathbb{C})) \to R(G,\mathrm{SL}(mk,\mathbb{C}))$$

given by $\iota(\alpha) = \mathrm{ind}_H^G \alpha$ is an injective algebraic map. It depends on the choice of a system of representatives, and it is compatible with the action of $\mathrm{SL}(m,\mathbb{C})$ and $\mathrm{SL}(mk,\mathbb{C})$ respectively.

Moreover, the corresponding regular map (which does not depend on the choice of a system of representatives)

$$\overline{\iota}\colon X(H,\mathrm{SL}(m,\mathbb{C})) \to X(G,\mathrm{SL}(mk,\mathbb{C}))$$

has finite fibers.

Proof A very detailed proof of the first statement can be found in [11, Section 10.A] (see also [18, pages 9–10] and [24]). The second part is [24, Lemma 3]. □

Finally we recall the following theorem of Cooper, Long, and Reid [10, Theorem 1.3] (see also [8, Corollary 6] or [1, page 94]).

Theorem 2.5 *Let N be an irreducible 3–manifold with incompressible nonempty toroidal boundary.* (In our application N will be the exterior of a nontrivial knot in S^3.) *Then $\pi_1(N)$ admits a finite-index subgroup H that admits an epimorphism $\alpha\colon H \to F_2$ onto a free group on two generators.*

We are now in the position to prove our main result:

Proof of Theorem 1.1 Let K be a nontrivial knot. We write $G = \pi_1(E_K)$. By Theorem 2.5 the group G admits a finite-index subgroup H that admits an epimorphism $\alpha\colon H \to F_2$ onto a free group on two generators. It is clear (see Example 2.1) that $R(F_2,\mathrm{SL}(m,\mathbb{C})) \cong \mathrm{SL}(m,\mathbb{C})^2$, and

$$\dim X(F_2,\mathrm{SL}(m,\mathbb{C})) = m^2 - 1.$$

It follows from Lemma 2.2 that the variety $X(H,\mathrm{SL}(m,\mathbb{C}))$ has a component of dimension at least m^2-1. We denote by k the index of H in G, and we will suppose that m is even. Then it follows from Lemma 2.4 that $X(G,\mathrm{SL}(mk,\mathbb{C}))$ contains an irreducible component of dimension at least m^2-1.

Now for all $m > k$ we have $m^2 - 1 > mk - 1$. Therefore, for a given $N \in \mathbb{N}$ we choose an $m \in \mathbb{N}$ that is even and greater than k such that $n := mk \geq N$. The character variety $X(\pi_K, \mathrm{SL}(n, \mathbb{C}))$ contains an irreducible component whose dimension is bigger than $m^2 - 1 > mk - 1 = n - 1$. □

3 Proof of Theorem 1.2

We let $R^{\mathrm{irr}}(G, \mathrm{SL}(n, \mathbb{C})) \subset R(G, \mathrm{SL}(n, \mathbb{C}))$ denote the Zariski-open subset of irreducible representations. The set $R^{\mathrm{irr}}(G, \mathrm{SL}(n, \mathbb{C}))$ is invariant by the $\mathrm{SL}(n, \mathbb{C})$–action, and we will denote by $X^{\mathrm{irr}}(G, \mathrm{SL}(n, \mathbb{C})) \subset X(G, \mathrm{SL}(n, \mathbb{C}))$ its image in the character variety. Notice that $X^{\mathrm{irr}}(G, \mathrm{SL}(n, \mathbb{C}))$ is an orbit space for the action of $\mathrm{SL}(n, \mathbb{C})$ on $R^{\mathrm{irr}}(G, \mathrm{SL}(n, \mathbb{C}))$ (see [22, Chapter 3, Section 3]).

Before we can give the proof of Theorem 1.2 we need to introduce several further definitions. These notations are classic (see [28; 4] for more details).

Let H and K be two subgroups of finite index of G, and let $\alpha\colon H \to \mathrm{GL}(W)$ be a linear representation. Then for all $g \in G$ we obtained the *twisted* representation $\alpha^g\colon gHg^{-1} \to \mathrm{GL}(W)$ given by

$$\alpha^g(x) = \alpha(g^{-1}xg) \quad \text{for } x \in gHg^{-1}. \tag{1}$$

Notice that the twisted representation α^g is irreducible or semisimple if and only if α is irreducible or semisimple respectively.

Now, we choose a set of representatives S of the (K, H) double cosets of G. For $s \in S$, we let $H_s = sHs^{-1} \cap K \leq K$. We obtain a homomorphism $\mathrm{res}^{sHs^{-1}}_{H_s} \alpha^s\colon H_s \to \mathrm{GL}(W)$ by restriction of α^s to H_s. The representation $\mathrm{res}^G_K \mathrm{ind}^G_H \alpha$ is equivalent to the direct sum of twisted representations:

$$\mathrm{res}^G_K \mathrm{ind}^G_H \alpha \cong \bigoplus_{s \in S} \mathrm{ind}^K_{H_s} \mathrm{res}^{sHs^{-1}}_{H_s} \alpha^s. \tag{2}$$

Equation (2) takes a simple form if $H = N = K$ is a normal subgroup of finite index of G. We obtain

$$\mathrm{res}^G_N \mathrm{ind}^G_N \alpha \cong \bigoplus_{s \in S} \alpha^s, \tag{3}$$

where S is a set of representatives of the N cosets of G.

In what follows we will make use of the following lemmas:

Lemma 3.1 *Let G be a group, $H \leq G$ a subgroup of finite index, and $\rho\colon G \to \mathrm{GL}(V)$ a representation. If $\mathrm{res}^G_H \rho\colon H \to \mathrm{GL}(V)$ is semisimple, then $\rho\colon G \to \mathrm{GL}(V)$ is semisimple.*

Proof This is [30, Theorem 1.5]. □

Lemma 3.2 *Let G be a group, $H \leq G$ a subgroup of finite index, and $\alpha\colon H \to \mathrm{GL}(W)$ a representation. If α is irreducible, then $\mathrm{ind}^G_H \alpha$ is semisimple.*

Proof We can choose a normal subgroup $N \trianglelefteq G$ of finite index such that $N \leq H$. More precisely, we can take

$$N = \bigcap_g gHg^{-1}$$

to be the *normal core* of H in G. We choose a set of representatives S of the (N, H) double cosets of G. In this case we obtain that $H_s = sHs^{-1} \cap N = s(H \cap N)s^{-1} = N$, and the double coset NsH is equal to sH since $N \subset H$ is normal. Therefore, (2) gives

$$\mathrm{res}^G_N \mathrm{ind}^G_H \alpha \cong \bigoplus_{s \in S} \mathrm{res}^{sHs^{-1}}_N \alpha.$$

Now, $\mathrm{res}^{sHs^{-1}}_N \alpha\colon N \to \mathrm{GL}(W)$ is a *twist* of $\alpha|_N$ ie for all $g \in N$ we have

$$\mathrm{res}^{sHs^{-1}}_N \alpha(g) = \alpha(s^{-1}gs) = (\alpha|_N)^s(g).$$

By Clifford's theorem [30, Theorem 1.7], we obtain that $\alpha|_N$ is semisimple. We have that $\alpha|_N = \alpha_1 \oplus \cdots \oplus \alpha_k$ is a direct sum of simple representations. Therefore,

$$(\mathrm{ind}^G_H \alpha)|_N \cong \bigoplus_{s \in S} \alpha_1^s \oplus \cdots \oplus \alpha_k^s$$

is the direct sum of irreducible representations. This proves that $(\mathrm{ind}^G_H \alpha)|_N$ is semisimple, and it follows from Lemma 3.1 that $\mathrm{ind}^G_H \alpha$ is semisimple. □

Corollary 3.3 *Let G be a group, and let $N \lhd G$ be a normal subgroup of finite index. If $\alpha\colon N \to \mathrm{SL}(V)$ is irreducible, then $\mathrm{ind}^G_N(\alpha)$ is semisimple.*

Moreover, if $\mathrm{ind}^G_N(\alpha) \cong \rho_1 \oplus \cdots \oplus \rho_l$ is a decomposition of $\mathrm{ind}^G_N(\alpha)$ into irreducible representations $\rho_j\colon G \to \mathrm{SL}(V_j)$, then $\dim V$ divides $\dim V_j$ and hence

$$\dim V \leq \dim V_j \leq \dim(V) \cdot [G : N].$$

Proof The first part follows directly from (3) and Lemma 3.1 since α^s is irreducible for all $s \in G$. Notice that S is now a set of representatives of the cosets G/N. Moreover,

we obtain

$$\bigoplus_{s\in S} \alpha^s \cong \operatorname{res}_N^G \operatorname{ind}_N^G \alpha \cong (\rho_1|_N) \oplus \cdots \oplus (\rho_l|_N).$$

If $\rho_j|_N$ is irreducible, then it must be isomorphic to one of the twisted representations α^s. Otherwise $\rho_j|_N$ is isomorphic to a direct sum of twisted representations α^s, $s \in S$, and hence $\dim V_j$ is a multiple of $\dim V$. □

Lemma 3.4 *Let G be a group and let $H \leq G$ be a finite-index subgroup.*

If there exists a surjective homomorphism $\varphi\colon H \twoheadrightarrow F_2$ onto a free group of rank two, then there exists a normal subgroup $N \trianglelefteq G$ of finite index such that $N \leq H$, and $\varphi(N) \leq F_2$ is a free group of finite rank $r \geq 2$.

Proof Let N be the normal core of H ie $N = \bigcap_{g\in G} gHg^{-1} \trianglelefteq G$. The normal subgroup is a finite-index subgroup of G, and $N \leq H$. Now, $H \geq \varphi^{-1}(\varphi(N)) \geq N$, and therefore $\varphi^{-1}(\varphi(N))$ is also of finite index in H (and hence in G). Hence, $\varphi(N) \trianglelefteq F_2$ is of finite index, and $\varphi(N)$ is a free group of rank $r = [F_2 : \varphi(N)] + 1 \geq 2$. □

We conclude this section with the following theorem.

Theorem 3.5 *Let $K \subset S^3$ be a nontrivial knot. Then there exists $k \in \mathbb{N}$ such that for all even $m \in \mathbb{N}$ there exists $p \in \mathbb{N}$ such that $m \leq p \leq mk$, and*

$$\dim X^{\mathrm{irr}}(\pi_K, \mathrm{SL}(p,\mathbb{C})) \geq \frac{m^2-k}{k}.$$

In particular, for m even with $m > k^2$ there exists $p \in \mathbb{N}$ such that $m \leq p < m\sqrt{m}$, and

$$\dim X^{\mathrm{irr}}(\pi_K, \mathrm{SL}(p,\mathbb{C})) \geq \frac{m^2-k}{k} > km - 1 \geq p - 1.$$

The last sequence of inequalities in Theorem 3.5 implies that $X^{\mathrm{irr}}(\pi_K, \mathrm{SL}(p,\mathbb{C}))$ contains a high-dimensional component. In particular, Theorem 3.5 implies Theorem 1.2 from the introduction.

Proof By Lemma 3.4 there exists a finite-index normal subgroup $N \trianglelefteq \pi_K$ of the knot group π_K, and an epimorphism $\psi\colon N \twoheadrightarrow F_2$. We put $k = [\pi_K : N]$.

For each even $m \in \mathbb{N}$, we obtain a regular map

$$\psi^*\colon X^{\mathrm{irr}}(F_2, \mathrm{SL}(m,\mathbb{C})) \to X(N, \mathrm{SL}(m,\mathbb{C}))$$

and we let $C \subset X(N, \mathrm{SL}(m,\mathbb{C}))$ denote the image of ψ^*. By Chevalley's theorem [3, Corollary 10.2], the set $C \subset X(N, \mathrm{SL}(m,\mathbb{C}))$ is constructible. Again, by Chevalley's theorem the image $D := \overline{\iota}(C)$, where $\overline{\iota}\colon X(N, \mathrm{SL}(m,\mathbb{C})) \to X(\pi_K, \mathrm{SL}(km,\mathbb{C}))$, is

also a constructible set. Notice that $\dim C = \dim D = m^2-1$ since ψ^* is an embedding by Lemma 2.2 and since $\overline{\iota}$ has finite fibers by Lemma 2.3.

If D contains a character of an irreducible representation, then

$$D \cap X^{\mathrm{irr}}(\pi_K, \mathrm{SL}(km, \mathbb{C}))$$

contains a Zariski-open subset of $\bar{D}$ which is of dimension $m^2-1 \geq (m^2-k)/k$. Hence the conclusion of the theorem is satisfied for $p = km$.

If D does not contain irreducible representations, then

$$D \subset X^{\mathrm{red}}(\pi_K, \mathrm{SL}(km, \mathbb{C})).$$

In this case we can choose for a given $\chi \in D$ a semisimple representation ρ such that $\chi_\rho = \chi$. Now, we follow the argument in the proof of Corollary 3.3 and we obtain $\rho \sim \rho_1 \oplus \cdots \oplus \rho_l$, where $\rho_j \colon \pi_K \to \mathrm{GL}(p_j, \mathbb{C})$ and $m \leq p_j < km$. For the l–tuple $(p_1, \ldots, p_l)$ we consider the regular map

$$\Phi_{(p_1,\ldots,p_l)} \colon R(\pi_K, \mathrm{SL}(p_1, \mathbb{C})) \times \cdots \times R(\pi_K, \mathrm{SL}(p_l, \mathbb{C})) \times (\mathbb{C}^*)^{l-1} \to R(\pi_K, \mathrm{SL}(km, \mathbb{C}))$$

given by

$$\Phi_{(p_1,\ldots,p_l)}(\rho_1, \ldots, \rho_l, \lambda_1, \ldots, \lambda_{l-1}) = \bigoplus_{i=1}^{l-1} (\rho_i \otimes \lambda_i^{p_l \varphi}) \oplus (\rho_l \otimes (\lambda_1^{-p_1} \cdots \lambda_{l-1}^{-p_{l-1}})^{\varphi}).$$

The map $\Phi_{(p_1,\ldots,p_l)}$ induces a map $\bar{\Phi} := \bar{\Phi}_{(p_1,\ldots,p_l)}$ between the character varieties

$$\bar{\Phi} \colon X(\pi_K, \mathrm{SL}(p_1, \mathbb{C})) \times \cdots \times X(\pi_K, \mathrm{SL}(p_l, \mathbb{C})) \times (\mathbb{C}^*)^{l-1} \to X^{\mathrm{red}}(\pi_K, \mathrm{SL}(km, \mathbb{C})) \tag{4}$$

(see [18, page 20]).

By Lemma 3.7, the map $\bar{\Phi}_{(p_1,\ldots,p_l)}$ also has finite fibers; denote its image by $D_{(p_1,\ldots,p_l)}$. Again, by Chevalley's theorem, the image $D_{(p_1,\ldots,p_l)} \subset X^{\mathrm{red}}(\pi_K, \mathrm{SL}(km, \mathbb{C}))$ is a constructible set, and

$$\dim D_{(p_1,\ldots,p_l)} = \sum_{j=1}^{l} \dim X(\pi_K, \mathrm{SL}(p_j, \mathbb{C})) + l - 1. \tag{5}$$

By Corollary 3.3, D is covered by finitely many sets of the form $D_{(p_1,\ldots,p_l)}$ (see also [18, Proposition 1.29]). Since $\dim D = m^2-1$, there must be at least one set $D_{(p_1,\ldots,p_l)}$ of dimension at least m^2-1. If we apply (5) to this choice we obtain that

$$\sum_{j=1}^{l} \dim X(\pi_K, \mathrm{SL}(p_j, \mathbb{C})) \geq m^2 - l.$$

In particular there is a j such that the corresponding summand is greater than or equal to $(m^2-l)/l$. Note that from $m \le p_j \le mk$ for $j = 1, \dots, l$ and $p_1 + \cdots + p_l = mk$ it follows that $l \le k$ which in turn implies that $(m^2-l)/l \ge (m^2-k)/k$. Summarizing we see that

$$\dim X^{\mathrm{irr}}(\pi_K, \mathrm{SL}(p_j, \mathbb{C})) \ge \frac{m^2-k}{k}.$$

This concludes the proof of the first statement of the theorem.

The second statement follows from the first statement using some elementary algebraic inequalities. □

In order to prove Lemma 3.7 we will make use of Schur's lemma (see [14, 1.7]).

Lemma 3.6 (Schur's lemma) *If V and W are irreducible representations of G and $f\colon V \to W$ is a $\mathbb{C}[G]$–module homomorphism, then:*

(1) *Either f is an isomorphism or $f = 0$.*

(2) *If $V = W$, then $f = \lambda \cdot \mathrm{id}$ for some $\lambda \in \mathbb{C}$.*

Lemma 3.7 *The map $\overline{\Phi} := \overline{\Phi}_{(p_1,\dots,p_l)}$ from* (4),

$$\overline{\Phi}\colon X(\pi_K, \mathrm{SL}(p_1, \mathbb{C})) \times \cdots \times X(\pi_K, \mathrm{SL}(p_l, \mathbb{C})) \times (\mathbb{C}^*)^{l-1} \to X(\pi_K, \mathrm{SL}(km, \mathbb{C})),$$

has finite fibers.

Proof We show that $\overline{\Phi}_{(p_1,\dots,p_l)}$ is a composition of maps having finite fibers.

Let G be a finitely generated group, and let $n_1 + \cdots + n_k = n$ be a partition of n. Then the map

$$\Psi\colon R(G, \mathrm{GL}(n_1, \mathbb{C})) \times \cdots \times R(G, \mathrm{GL}(n_k, \mathbb{C})) \to R(G, \mathrm{GL}(n, \mathbb{C}))$$

given by $\Psi(\varrho_1, \dots \varrho_k) = \varrho_1 \oplus \cdots \oplus \varrho_k$ induces a regular map

$$\overline{\Psi}\colon X(G, \mathrm{GL}(n_1, \mathbb{C})) \times \cdots \times X(G, \mathrm{GL}(n_k, \mathbb{C})) \to X(G, \mathrm{GL}(n, \mathbb{C}))$$

(see [18, page 20]). Let $\chi \in X(G, \mathrm{GL}(n, \mathbb{C}))$ be in the image of $\overline{\Psi}$. We will use that a character $\chi = \chi_\varrho$ is the character of a semisimple representation ϱ, which is unique up to conjugation [18, Theorem 1.28]. The representation $\varrho \sim \varrho_1 \oplus \cdots \oplus \varrho_l$ decomposes into irreducible representations. Schur's lemma implies that the representations ϱ_j are unique up to equivalence [14, Proposition 1.8]. If $\overline{\Psi}(\chi_1, \dots, \chi_k) = \chi$ then each χ_i is also the character of a semisimple representation α_i, and $\alpha_1 \oplus \cdots \oplus \alpha_k \sim \varrho$. Now, the

irreducible representations which can occur in the decomposition of α_i are equivalent to the representations ϱ_j. This means that the semisimple representations α_i such that $\Psi(\alpha_1,\dots,\alpha_k)$ is equivalent to ϱ, can have among their irreducible components only representations equivalent to $\varrho_1,\dots,\varrho_l$. Hence $\overline{\Psi}^{-1}(\chi)$ is finite.

Next, for $G=\pi_K$ and the surjection $\varphi\colon \pi_K \twoheadrightarrow \mathbb{Z}$ we consider the map

$$\Lambda_p\colon R(\pi_K,\mathrm{SL}(p,\mathbb{C}))\times\mathbb{C}^*\to R(\pi_K,\mathrm{GL}(p,\mathbb{C})),\quad \Lambda_p(\rho,\lambda)=\rho\otimes\lambda^{\varphi}.$$

The cyclic group $C_p=\langle\omega\rangle$ of pth roots of unity acts on $R(\pi_K,\mathrm{SL}(p,\mathbb{C}))\times\mathbb{C}^*$ by $\omega(\rho,\lambda)=(\rho\otimes\omega^{\varphi},\overline{\omega}\lambda)$. This action commutes with the action of $\mathrm{SL}(p,\mathbb{C})$ by conjugation. We obtain an action of C_p on $X(\pi_K,\mathrm{SL}(p,\mathbb{C}))\times\mathbb{C}^*$. The map

$$\overline{\Lambda}_p\colon X(\pi_K,\mathrm{SL}(p,\mathbb{C}))\times\mathbb{C}^*\to X(\pi_K,\mathrm{GL}(p,\mathbb{C}))$$

factors through $(X(\pi_K,\mathrm{SL}(p,\mathbb{C}))\times\mathbb{C}^*)/C_p\cong X(\pi_K,\mathrm{GL}(p,\mathbb{C}))$ (see Lemma 2.2 in [17]). Hence $\overline{\Lambda}_p$ has finite fibers.

Note that the map $\varepsilon\colon(\mathbb{C}^*)^{l-1}\to(\mathbb{C}^*)^l$ given by

$$\varepsilon(\lambda_1,\dots,\lambda_{l-1})=(\lambda_1^{p_l},\dots,\lambda_{l-1}^{p_l},(\lambda_1^{p_1}\cdots\lambda_{l-1}^{p_{l-1}})^{-1})$$

has finite fibers. Finally,

$$\overline{\Phi}_{(p_1,\dots,p_l)}=\overline{\Psi}\circ(\overline{\Lambda}_{p_1}\times\cdots\times\overline{\Lambda}_{p_l})\circ(\mathrm{id}\times\cdots\times\mathrm{id}\times\varepsilon)$$

has finite fibers. □

4 The character variety of the figure-eight knot

The aim of this section is to prove Proposition 1.3. Before studying the character variety of the figure-eight knot we will extend Mackey's irreducibility criterion to infinite groups. Suppose that $H\le G$ is a subgroup, and that W is a representation of H. Mackey's irreducibility criterion gives necessary and sufficient conditions under which $\mathrm{ind}_H^G W$ is irreducible as a $\mathbb{C}[G]$–module.

Let $H\le G$ be a subgroup of finite index, and let $\alpha\colon H\to\mathrm{GL}(W)$ be a representation. For $s\in G$ we obtain the twisted representation $\alpha^s\colon sHs^{-1}\to\mathrm{GL}(W)$ (see (1)). Define

$$H_s:=sHs^{-1}\cap H\quad\text{ and }\quad W_s:=\mathrm{res}^{sHs^{-1}}_{H_s}W_{\alpha^s}.$$

In what follows we call two semisimple representations V and V' of G *disjoint* if $\mathrm{Hom}_{\mathbb{C}[G]}(V,V')=0$. The aim of the next subsection is to prove the following:

Proposition 4.1 (Mackey's irreducibility criterion) *Let $H \leq G$ be a subgroup of finite index. We suppose that $\alpha\colon H \to \mathrm{GL}(W)$ is an irreducible representation. Then $\mathrm{ind}_H^G \alpha$ is irreducible if and only if for all $s \in G - H$ the $\mathbb{C}[H_s]$–modules W_s and $\mathrm{res}_{H_s}^H W$ are disjoint.*

Remark Proposition 4.1 is well known for finite groups (see [28]). We will see that it also holds for infinite groups under the assumption that the subgroup $H \leq G$ has finite index. The main difference is that for infinite groups the induced and *coinduced* modules are in general not isomorphic. We will make use of Lemma 4.2 in place of the classical Frobenius reciprocity theorem.

4.1 Mackey's criterion

We follow Serre's notation and exposition in [28]. First let us recall the adjoint isomorphism. Let Q and R be rings and let ${}_QS$, ${}_RT_Q$, and ${}_RU$ be modules (here the location of the indices indicates whether these are left or right module structures). Then $\mathrm{Hom}_R(T,U)$ is a left Q–module via $qf(v) = f(vq)$ for all $v \in V$, and

$$\mathrm{Hom}_Q(S, \mathrm{Hom}_R(T,U)) \cong \mathrm{Hom}_R(T \otimes_Q S, U) \tag{6}$$

(see [27, Theorem 2.76]).

Lemma 4.2 *Let G be a group and $H \leq G$ a subgroup of finite index. For each left $\mathbb{C}[H]$–module W and a left $\mathbb{C}[G]$–module V we obtain*

$$\mathrm{Hom}_{\mathbb{C}[G]}(V, \mathrm{ind}_H^G W) \cong \mathrm{Hom}_{\mathbb{C}[H]}(\mathrm{res}_H^G V, W), \tag{7}$$

and

$$\mathrm{Hom}_{\mathbb{C}[H]}(W, \mathrm{res}_H^G V) \cong \mathrm{Hom}_{\mathbb{C}[G]}(\mathrm{ind}_H^G W, V). \tag{8}$$

Proof For proving (7) we apply (6) with $Q = \mathbb{C}[G]$, $R = \mathbb{C}[H]$, $S = V$, $T = \mathbb{C}[G]$, and $U = W$:

$$\mathrm{Hom}_{\mathbb{C}[G]}(V, \mathrm{Hom}_{\mathbb{C}[H]}(\mathbb{C}[G], W)) \cong \mathrm{Hom}_{\mathbb{C}[H]}(\mathbb{C}[G] \otimes_{\mathbb{C}[G]} V, W).$$

Since $H \leq G$ is of finite index we obtain that the coinduced module $\mathrm{coind}_H^G(W) := \mathrm{Hom}_{\mathbb{C}[H]}(\mathbb{C}[G], W)$ and the induced module $\mathrm{ind}_H^G(W) = \mathbb{C}[G] \otimes_{\mathbb{C}[H]} W$ are isomorphic as left $\mathbb{C}[G]$–modules (see [4, Lemma III.5.9]). Moreover, $\mathrm{res}_H^G(U)$ and $\mathbb{C}[G] \otimes_{\mathbb{C}[G]} U$ are isomorphic as left $\mathbb{C}[H]$–modules, and (7) follows.

To prove (8) we apply (6) with $Q=\mathbb{C}[H],\ R=\mathbb{C}[G],\ S=W,\ T=\mathbb{C}[G],\ U=V$:

$$\mathrm{Hom}_{\mathbb{C}[H]}(W,\mathrm{Hom}_{\mathbb{C}[G]}(\mathbb{C}[G],V))\cong\mathrm{Hom}_{\mathbb{C}[G]}(\mathbb{C}[G]\otimes_{\mathbb{C}[H]}W,V).$$

The left $\mathbb{C}[H]$–module $\mathrm{Hom}_{\mathbb{C}[G]}(\mathbb{C}[G],V)$ is isomorphic to $\mathrm{res}^G_H(V)$, and hence (8) follows. □

Before proving Mackey's criterion we state the following:

Lemma 4.3 *Let V be a semisimple G–module. Then V is irreducible if and only if $\mathrm{Hom}_{\mathbb{C}[G]}(V,V)\cong\mathbb{C}$.*

Proof Let $V\cong V_1\oplus\cdots\oplus V_k$ be a decomposition of V into irreducible modules. Then

$$\mathrm{Hom}_{\mathbb{C}[G]}(V,V)=\bigoplus_{i,j}\mathrm{Hom}_{\mathbb{C}[G]}(V_i,V_j).$$

Now, the lemma follows from Schur's lemma since $\dim_{\mathbb{C}}\mathrm{Hom}_{\mathbb{C}[G]}(V,V)=1$ if V is irreducible and $\dim_{\mathbb{C}}\mathrm{Hom}_{\mathbb{C}[G]}(V,V)>1$ if V is not irreducible. □

Proof of Proposition 4.1 It follows from Lemma 3.2 that $\mathrm{ind}^G_H\alpha$ is semisimple. Therefore, by Lemma 4.3 it follows that $\mathrm{ind}^G_H\alpha$ is irreducible if and only if

$$\mathrm{Hom}_{\mathbb{C}[G]}(\mathrm{ind}^G_H W,\mathrm{ind}^G_H W)\cong\mathbb{C}.$$

We choose a system S of the (H,H) double cosets of G. Then we obtain that

$$\begin{aligned}\mathrm{Hom}_{\mathbb{C}[G]}(\mathrm{ind}^G_H W,\mathrm{ind}^G_H W)&\cong\mathrm{Hom}_{\mathbb{C}[H]}(W,\mathrm{res}^G_H\,\mathrm{ind}^G_H W)&&\text{(by (8))}\\&\cong\bigoplus_{s\in S}\mathrm{Hom}_{\mathbb{C}[H]}(W,\mathrm{ind}^H_{H_s}W_s)&&\text{(by (2) for } K=H)\\&\cong\bigoplus_{s\in S}\mathrm{Hom}_{\mathbb{C}[H_s]}(\mathrm{res}^H_{H_s}W,W_s)&&\text{(by (7)).}\end{aligned}$$

Now, if $s\in H$, then $H_s=H$ and $W_s\cong W$, and $\mathrm{Hom}_H(W,W)\cong\mathbb{C}$ since W is an irreducible $\mathbb{C}[H]$–module (see Lemma 4.3). Thus $\mathrm{Hom}_{\mathbb{C}[G]}(\mathrm{ind}^G_H W,\mathrm{ind}^G_H W)\cong\mathbb{C}$ if and only if $\mathrm{Hom}_{\mathbb{C}[H_s]}(\mathrm{res}^H_{H_s}W,W_s)=0$ for all $s\in G-H$. □

4.2 The character varieties of the figure-eight knot

Let us consider the figure-eight knot K_{4_1} and its group π_{4_1}:

$$\pi_{4_1}\cong\langle s,t\mid st^{-1}s^{-1}ts=tst^{-1}s^{-1}t\rangle\cong\langle s,a\mid a^{-1}s^{-1}asa^{-1}sas^{-1}a^{-1}\rangle,\tag{9}$$

where $a = ts^{-1}$. Following Reidemeister [26, Section 15], we obtain a representation $\delta\colon \pi_{4_1} \to S_5$ into the symmetric group S_5 given by

$$\delta(s) = \sigma := (1)(2,5)(3,4) \quad \text{and} \quad \delta(a) = \tau := (1,2,3,4,5). \tag{10}$$

The image of δ is a dihedral group. We adopt the convention that permutations act on the right on $\{1,\dots,\alpha\}$, and hence π_{4_1} acts on the right. We put $N = \mathrm{Ker}(\delta)$ and $H = \mathrm{Stab}(1) = \{g \in G \mid 1^{\delta(g)} = 1\}$. We have $N \le H$, $N \trianglelefteq \pi_{4_1}$, $[\pi_{4_1} : H] = 5$, and $[H : N] = 2$. We obtain

$$\pi_{4_1} = H \sqcup Ha \sqcup Ha^2 \sqcup Ha^3 \sqcup Ha^4,$$

and $g \in Ha^i$ if and only if $1^{\delta(g)} = i + 1$ for $0 \le i \le 4$. Notice that $\delta(H) = \langle\sigma\rangle$ is the cyclic group generated by σ.

Now we can use the Reidemeister–Schreier method [31] for finding a presentation for H: $\{1, a, a^2, a^3, a^4\}$ is a Schreier representative system for the right cosets modulo H. Hence generators of H are

$$y_0 = s, \quad y_i = a^i s a^{i-5} \text{ for } i = 1,2,3,4, \quad \text{and} \quad y_5 = a^5. \tag{11}$$

We obtain defining relations r_i, where $i = 0,1,2,3,4$, for H by expressing each

$$r_i = a^i(a^{-1}s^{-1}asa^{-1}sas^{-1}a^{-1})a^{-i}$$

as a word in the y_j:

$$r_0 = y_5^{-1}y_1^{-1}y_2^2y_1^{-1}, \quad r_1 = y_0^{-1}y_1y_3y_2^{-1}, \quad r_2 = y_4^{-1}y_5y_0y_5^{-1}y_4y_3^{-1},$$
$$r_3 = y_3^{-1}y_4y_0y_4^{-1}, \quad r_4 = y_2^{-1}y_3y_1y_5y_0^{-1}y_5^{-1}.$$

It follows that $H/\langle\langle y_0\rangle\rangle \cong \langle y_1,y_2,y_3,y_4,y_5 \mid y_3 = y_5 = 1, y_1 = y_2\rangle \cong F(y_1,y_4)$. Therefore, a surjection $\psi\colon H \twoheadrightarrow F_2 = F(x,y)$ is given by

$$\psi(y_0) = \psi(y_3) = \psi(y_5) = 1, \quad \psi(y_1) = \psi(y_2) = x, \quad \text{and} \quad \psi(y_4) = y. \tag{12}$$

We have $y_0 \in H - N$ since by (10) the permutation $\delta(y_0) = \delta(s) = \sigma$ fixes 1 but $\delta(y_0) = \sigma \ne \mathrm{id}$. Moreover, $y_0 \in \mathrm{Ker}(\psi)$.

We need also generators of N: we have $y_0 \notin N$, $\delta(y_5) = \mathrm{id}$, and $\delta(y_i) = \sigma$ for $0 \le i \le 4$. Hence Reidemeister–Schreier gives that N is generated by

$$y_i y_0^{-1}, \quad y_5, \quad y_0^2, \quad y_0 y_i, \quad y_0 y_5 y_0^{-1} \qquad \text{where } i = 1,2,3,4. \tag{13}$$

Lemma 4.4 *Let $\beta\colon F_2 \to \mathrm{SL}(2m,\mathbb{C})$ be irreducible. Then $\alpha = \beta\circ\psi\colon H \to F_2 \to \mathrm{SL}(2m,\mathbb{C})$ and $\alpha|_N\colon H \to F_2 \to \mathrm{SL}(2m,\mathbb{C})$ are also irreducible.*

Proof If $\beta\colon F_2 \to \mathrm{SL}(2m,\mathbb{C})$ is irreducible, then $\alpha = \beta\circ\psi$ is also irreducible since $\psi\colon H \to F_2$ is surjective (the representations α and β have the same image). Similarly, (13) and (12) give that $\psi|_N\colon N \to F_2$ is also surjective (β and $\alpha|_N$ have the same image). □

We obtain a component of representations $X_0 \subset X(H,\mathrm{SL}(2m,\mathbb{C}))$ with $\dim X_0 \geq 4m^2-1$, and X_0 contains irreducible representations. In order to apply Proposition 4.1 we notice that $\{1,a,a^2\}$ is a representative system for the (H,H) double cosets

$$\pi_{4_1} = H \sqcup HaH \sqcup Ha^2H.$$

More precisely, we have that $HaH = Ha \sqcup Ha^4$ and $Ha^2H = Ha^2 \sqcup Ha^3$ since $1^{\delta(hah')} \in \{2,5\}$ and $1^{\delta(ha^2h')} \in \{3,4\}$ for all $h,h' \in H$ (recall that $\delta(H) = \langle\sigma\rangle$).

We have also

$$H_a = H \cap aHa^{-1} = N = H_{a^2} = H \cap a^2Ha^{-2},$$

since an element in the image of the dihedral representation $\delta\colon \pi_{4_1} \to S_5$ given by (10) which fixes two numbers is the identity, and $N = \mathrm{Ker}(\delta)$.

For the rest of this section we let $G := \pi_{4_1}$ denote the group of the figure-eight knot.

Lemma 4.5 *Let $\beta\colon F(x,y) \to \mathrm{SL}(2m,\mathbb{C})$ be given by $\beta(x) = A$ and $\beta(y) = B$. If β is irreducible, then $\rho = \mathrm{ind}_H^G(\beta\circ\psi)$ is also irreducible.*

Proof We let $\alpha = \beta\circ\psi$ denote the corresponding representation of H. We want to apply Proposition 4.1 to prove that $\rho = \mathrm{ind}_H^G\,\alpha$ is irreducible. For $s \in \{a,a^2\}$ we have

$$\mathrm{res}_{H_s}^H \alpha = \alpha|_N \quad \text{since } H_a = H_{a^2} = N.$$

Moreover, for $s \in \{a,a^2\}$ we obtain

$$W_s = \mathrm{res}_{H_s}^{sHs^{-1}} W_{\alpha^s} = (\alpha|_N)^s.$$

By Lemma 4.4 we obtain that α, $\alpha|_N$, $(\alpha|_N)^a$, and $(\alpha|_N)^{a^2}$ are irreducible. Recall also that by Schur's lemma (Lemma 3.6) two irreducible representations are disjoint if and only if they are not equivalent. Hence by Proposition 4.1 we obtain that $\rho = \mathrm{ind}_H^G\,\alpha$ is irreducible if and only if

$$\alpha|_N \not\sim (\alpha|_N)^a \quad \text{and} \quad \alpha|_N \not\sim (\alpha|_N)^{a^2}.$$

In order to see this, we consider the element $y_0^2 \in N$. We have $\psi(y_0) = 1$ and therefore $\alpha|_N(y_0^2) = I_{2m}$, where $I_{2m} \in \mathrm{SL}(2m, \mathbb{C})$ denotes the identity matrix. Recall that $\beta(x) = A$ and $\beta(y) = B$, where $A, B \in \mathrm{SL}(2m, \mathbb{C})$. We have

$$a^{-1}y_0^2 a = a^{-1}s^2 a = a^{-5}\cdot a^4 s a^{-1}\cdot a s a^{-4}\cdot a^5 = y_5^{-1}y_4 y_1 y_5,$$
$$a^{-2}y_0^2 a^2 = a^{-2}s^2 a^2 = a^{-5}\cdot a^3 s a^{-2}\cdot a^2 s a^{-3}\cdot a^5 = y_5^{-1}y_3 y_2 y_5.$$

This gives $\psi(a^{-1}y_0^2 a) = yx$ and $\psi(a^{-2}y_0^2 a^2) = x$.

Therefore, $(\alpha|_N)^a(y_0^2) = BA$ and $(\alpha|_N)^{a^2}(y_0^2) = A$. Now, if β is irreducible, then $A \neq I_{2m}$, and $AB \neq I_{2m}$. Hence $\alpha|_N \not\sim (\alpha|_N)^a$ and $\alpha|_N \not\sim (\alpha|_N)^{a^2}$. □

Proof of Proposition 1.3 The subgroup $H \leq \pi_{4_1}$ is of index 5, and $\psi\colon H \to F_2$ is a surjective homomorphism onto a free group of rank 2. Now, by Lemma 4.4 and Lemma 4.5, and the same argument as in the proof of Theorem 1.2, we obtain that for all $m \in \mathbb{N}$ the character variety $X(\pi_{4_1}, \mathrm{SL}(10m, \mathbb{C}))$ has a component C of dimension at least $4m^2 - 1$. Finally, Lemma 4.5 implies that C contains characters of irreducible representations. □

More explicitly, if $\beta\colon F(x, y) \to \mathrm{SL}(2m, \mathbb{C})$ is a representation given by $\beta(x) = A$ and $\beta(y) = B$, then, following the construction given in Section 2.1, the induced representation $\rho = \mathrm{ind}_H^G(\beta \circ \psi)\colon \pi_{4_1} \to \mathrm{GL}(10m, \mathbb{C})$ is given by

$$\rho(s) = \begin{pmatrix} I_{2m} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & B \\ 0 & 0 & 0 & I_{2m} & 0 \\ 0 & 0 & A & 0 & 0 \\ 0 & A & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \rho(t) = \begin{pmatrix} 0 & A & 0 & 0 & 0 \\ I_{2m} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & B \\ 0 & 0 & 0 & I_{2m} & 0 \\ 0 & 0 & A & 0 & 0 \end{pmatrix}.$$

Here, s and t are the generators of π_{4_1} from (9), and $I_{2m} \in \mathrm{SL}(2m, \mathbb{C})$ is the identity matrix.

Remark We finish the section with some historical comments. Subgroups similar to H and N from Section 4.2 are more generally defined for two-bridge knots. Let $K = \mathfrak{b}(p, q) \subset S^3$ be a two-bridge knot. A presentation of the knot group $\pi_{p,q}$ is given by the following (see [7, E 12.1]):

$$\pi_{p,q} = \langle s, t \mid l_s s = t l_s \rangle \qquad \text{where } l_s = s^{\epsilon_1} t^{\epsilon_2} \cdots t^{\epsilon_{\alpha-1}} \text{ and } \epsilon_k = (-1)^{[k(\beta/\alpha)]}.$$

K Reidemeister [26, Section 15] defined the representation $\delta\colon \pi_{p,q} \to S_p$ into the

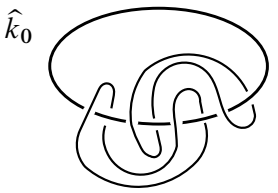


Figure 1: The link $\widehat{K}_{4_1} \subset \widehat{S}^3$

symmetric group S_p by

$$\delta(s) = (1)(2, 2n+1)(3, 2n)\cdots(n+1, n+2) \quad \text{and} \quad \delta(a) = (1, 2, \dots, p), \tag{14}$$

where $p = 2n+1$ and $a = ts^{-1}$. The image of δ is a dihedral group. We put $N = \mathrm{Ker}(\delta)$ and $H = \mathrm{Stab}(1) = \{g \in G \mid 1^{\delta(g)} = 1\}$. We have $N \leq H$, $N \trianglelefteq \pi_{p,q}$, $[\pi_{p,q} : H] = p$, and $[H : N] = 2$.

The irregular covering of E_K corresponding to H has been studied since the beginning of knot theory. Reidemeister calculated a presentation of H. Moreover, he showed that the total space of the corresponding irregular branched covering $(\widehat{S}^3, \widehat{K}) \to (S^3, \mathfrak{b}(p,q))$ is simply connected. He proved also that the branching set $\widehat{K}$ consists of $(n+1)$ unknotted components (see [25; 26]). G Burde [5] proved that $\widehat{S}^3$ is in fact the 3–sphere and he determined the nature of the branching set explicitly in [6]. More recently, G Walsh studied the regular branched covering corresponding to N [29]. She proved that the corresponding branching set is a great circle link in S^3.

For the figure-eight knot $K_{4_1} = \mathfrak{b}(5,3)$, the link $\widehat{K}_{4_1} \subset S^3$ has a particularly simple form (see Figure 1 and [7, Example 14.22]). If we fill in the component $\widehat{k}_0$, then $\widehat{K}_{4_1}$ transforms into the trivial link of two components. Therefore the isomorphism $H/\langle\langle y_0 \rangle\rangle \cong F_2$, where $\langle\langle y_0 \rangle\rangle$ denotes the normal subgroup of H generated by the meridian y_0 of $\widehat{k}_0$, has a geometric origin.

We expect similar methods can be used to prove that for any hyperbolic 2–bridge knot $\mathfrak{b}(p,q)$, where $q \geq 2$, the character variety $X(\pi_{p,q}, \mathrm{SL}(2pm, \mathbb{C}))$ admits an at least $(4m^2-1)$–dimensional component. In [16, Section 4] it was proved that for the trefoil knot $\mathfrak{b}(3,1)$, the variety $X(\pi_{3,1}, \mathrm{SL}(6, \mathbb{C}))$ contains an at least 7–dimensional component. The same construction can be used to prove that for the torus knots $\mathfrak{b}(p,1)$, the variety $X(\pi_{p,1}, \mathrm{SL}(2pk, \mathbb{C}))$, where $k \in \mathbb{N}$, is at least $(2k^2p^2 - 4pk^2 + 1)$–dimensional.

Acknowledgements The authors where supported by the SFB 1085 "Higher Invariants" at the University of Regensburg, funded by the Deutsche Forschungsgemeinschaft

(DFG), and by the Laboratoire de Mathématiques Blaise Pascal UMR 6620 of the Université Clermont Auvergne and CNRS.

We wish to thank the referee for many helpful comments.

References

[1] **M Aschenbrenner**, **S Friedl**, **H Wilton**, *3–manifold groups*, EMS Series of Lectures in Mathematics 20, Eur. Math. Soc., Zürich (2015) MR

[2] **L Ben Abdelghani**, **M Heusener**, *Irreducible representations of knot groups into* $\mathrm{SL}(n,\mathbf{C})$, Publ. Mat. 61 (2017) 363–394 MR

[3] **A Borel**, *Linear algebraic groups*, 2nd edition, Graduate Texts in Mathematics 126, Springer (1991) MR

[4] **K S Brown**, *Cohomology of groups*, corrected 1st edition, Graduate Texts in Mathematics 87, Springer (1994) MR

[5] **G Burde**, *On branched coverings of* S^3, Canad. J. Math. 23 (1971) 84–89 MR

[6] **G Burde**, *Links covering knots with two bridges*, Kobe J. Math. 5 (1988) 209–219 MR

[7] **G Burde**, **H Zieschang**, **M Heusener**, *Knots*, fully revised and extended 3rd edition, De Gruyter Studies in Mathematics 5, De Gruyter, Berlin (2013) MR

[8] **J O Button**, *Strong Tits alternatives for compact* 3*–manifolds with boundary*, J. Pure Appl. Algebra 191 (2004) 89–98 MR

[9] **D Cooper**, **D D Long**, *Remarks on the* A*–polynomial of a knot*, J. Knot Theory Ramifications 5 (1996) 609–628 MR

[10] **D Cooper**, **D D Long**, **A W Reid**, *Essential closed surfaces in bounded* 3*–manifolds*, J. Amer. Math. Soc. 10 (1997) 553–563 MR

[11] **C W Curtis**, **I Reiner**, *Methods of representation theory, I*, Wiley, New York (1981) MR

[12] **I Dolgachev**, *Lectures on invariant theory*, London Math. Soc. Lecture Note Series 296, Cambridge Univ. Press (2003) MR

[13] **E Falbel**, **A Guilloux**, *Dimension of character varieties for* 3*–manifolds*, Proc. Amer. Math. Soc. 145 (2017) 2727–2737 MR

[14] **W Fulton**, **J Harris**, *Representation theory*, Graduate Texts in Mathematics 129, Springer (1991) MR

[15] **M Heusener**, $\mathrm{SL}(n,\mathbb{C})$*–representation spaces of knot groups*, from "Topology, geometry and algebra of low-dimensional manifolds" (T Kitano, editor), Kôkyûroku 1991, RIMS, Kyoto (2016) 1–26

[16] **M Heusener**, **O Medjerab**, *Deformations of reducible representations of knot groups into* $\mathrm{SL}(n,\mathbf{C})$, Math. Slovaca 66 (2016) 1091–1104 MR

[17] **M Heusener**, **V Muñoz**, **J Porti**, *The* $\mathrm{SL}(3,\mathbb{C})$*–character variety of the figure eight knot*, Illinois J. Math. 60 (2016) 55–98 MR

[18] **A Lubotzky**, **A R Magid**, *Varieties of representations of finitely generated groups*, Mem. Amer. Math. Soc. 336, Amer. Math. Soc., Providence, RI (1985) MR

[19] **P Menal-Ferrer**, **J Porti**, *Local coordinates for* $\mathrm{SL}(n,\mathbf{C})$*–character varieties of finite-volume hyperbolic* 3*–manifolds*, Ann. Math. Blaise Pascal 19 (2012) 107–122 MR

[20] **P Menal-Ferrer**, **J Porti**, *Twisted cohomology for hyperbolic three manifolds*, Osaka J. Math. 49 (2012) 741–769 MR

[21] **V Muñoz**, **J Porti**, *Geometry of the* $\mathrm{SL}(3,\mathbb{C})$*–character variety of torus knots*, Algebr. Geom. Topol. 16 (2016) 397–426 MR

[22] **P E Newstead**, *Introduction to moduli problems and orbit spaces*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics 51, Narosa, New Delhi (1978) MR

[23] **C Procesi**, *The invariant theory of* $n \times n$ *matrices*, Advances in Math. 19 (1976) 306–381 MR

[24] **A S Rapinchuk**, *On SS–rigid groups and A Weil's criterion for local rigidity, I*, Manuscripta Math. 97 (1998) 529–543 MR

[25] **K Reidemeister**, *Knoten und Verkettungen*, Math. Z. 29 (1929) 713–729 MR

[26] **K Reidemeister**, *Knotentheorie*, Ergebnisse der Mathematik (1) 1, Springer (1932)

[27] **J J Rotman**, *An introduction to homological algebra*, 2nd edition, Springer (2009) MR

[28] **J-P Serre**, *Linear representations of finite groups*, Graduate Texts in Mathematics 42, Springer (1977) MR

[29] **G S Walsh**, *Great circle links and virtually fibered knots*, Topology 44 (2005) 947–958 MR

[30] **B A F Wehrfritz**, *Infinite linear groups: an account of the group-theoretic properties of infinite groups of matrices*, Ergebnisse der Mathematik (2) 76, Springer (1973) MR

[31] **H Zieschang**, **E Vogt**, **H-D Coldewey**, *Surfaces and planar discontinuous groups*, Lecture Notes in Mathematics 835, Springer (1980) MR

Fakultät für Mathematik, Universität Regensburg
Regensburg, Germany

Laboratoire de Mathématiques Blaise Pascal - UMR 6620 - CNRS
Université Clermont Auvergne, Campus des Cézeaux
Aubière, France

stefan.friedl@mathematik.uni-regensburg.de, michael.heusener@uca.fr

http://math.univ-bpclermont.fr/~heusener/

Received: 14 October 2016 Revised: 4 April 2017

msp

Algebraic & Geometric Topology 18 (2018) 333–385

DAHA and plane curve singularities

IVAN CHEREDNIK
IAN PHILIPP

We suggest a relatively simple and totally geometric conjectural description of uncolored DAHA superpolynomials of arbitrary algebraic knots (conjecturally coinciding with the reduced stable Khovanov–Rozansky polynomials) via the flagged Jacobian factors (new objects) of the corresponding unibranch plane curve singularities. This generalizes the Cherednik–Danilenko conjecture on the Betti numbers of Jacobian factors, the Gorsky combinatorial conjectural interpretation of superpolynomials of torus knots and that by Gorsky and Mazin for their constant term. The paper mainly focuses on nontorus algebraic knots. A connection with the conjecture due to Oblomkov, Rasmussen and Shende is possible, but our approach is different. A motivic version of our conjecture is related to p–adic orbital A–type integrals for anisotropic centralizers.

14H50, 17B45, 20C08, 20F36, 33D52, 57M25; 17B22, 22E50, 22E57, 30F10, 33D80

0 Introduction

We propose a relatively simple and totally computable conjectural *geometric* description of uncolored DAHA superpolynomials of arbitrary algebraic knots in terms of *flagged Jacobian factors* (new objects) of the corresponding unibranch plane curve singularities, presumably coinciding with the corresponding *stable Khovanov–Rozansky polynomials*. This description significantly generalizes (a) Cherednik and Danilenko's conjecture on the Betti numbers of Jacobian factors (any unibranch singularities), (b) Gorsky's conjectural interpretation of superpolynomials of torus knots from [16] etc, and (c) that from Gorsky and Mazin [19; 20] for their a–constant term. Our conjecture is different from the *ORS conjecture* from Gorsky, Oblomkov, Rasmussen and Shende [22] and Oblomkov, Rasmussen and Shende [36] (though some connection is not impossible).

Motivation Algebrogeometric theory of topological invariants of algebraic links has a long history, starting with the well-known algebraic interpretation of the *Alexander polynomials*. This paper provides an algebrogeometric description of stable

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.333

Khovanov–Rozansky polynomials via the DAHA-superpolynomials. See eg Khovanov [26], Khovanov and Rozansky [27; 28], Rasmussen [39] and Webster and Williamson [42]. The geometry of flagged Jacobian factors conjecturally provides the DAHA-superpolynomials from Cherednik [6; 7] and Cherednik and Danilenko [8; 9]. For instance, this explains the positivity of the latter in the uncolored case, conjectured in [8].

For uncolored *torus knots*, this positivity results from the combinatorial construction from Gorsky [16; 17], which conjecturally provides both stable *KhR-polynomials* and those via DAHA (and is closely related to rational DAHA). Our conjecture makes the positivity entirely geometric (generalizing [19; 20]) for torus and arbitrary algebraic knots. We expect important implications in the theory of plane curve singularities, the theory of p–adic *orbital integrals* and *affine Springer fibers*; see Conjecture 2.5(iii) and Section 5.

Algebraic knots Torus-type (quasihomogeneous) plane singularities are very special. Not much is actually known on the Jacobian factors of nontorus plane singularities; the paper by Piontkowski [38] still remains the main source of examples. It was an important development when the DAHA approach from [6; 7] and Gorsky and Neguț [21] was extended from torus knots to arbitrary algebraic knots in [8] and then to any algebraic link in [9]. The *Newton pairs* and the theory of *Puiseux expansion*, the key in the *topological* classification of plane curve singularities, naturally emerge in the DAHA approach.

One of the key advantages of the usage of DAHA is that adding colors is relatively direct (via the Macdonald polynomials), which is well understood for any *iterated torus link* (including all algebraic links). This is well ahead of any other approaches (topology included) for such links. We expect that our present paper can be enhanced by adding colors via (presumably) the curves suggested in Maulik [31]. The case of rectangle Young diagrams is exceptional due to the conjectured positivity of the corresponding *reduced* DAHA superpolynomials for algebraic knots [6; 8]. The switch from the rank-one torsion free modules in the definition of compactified Jacobians to arbitrary ranks is expected here (among other modifications), which is in progress.

The passage to arbitrary algebraic knots and links from torus knots is important because of multiple reasons. The generality is an obvious advantage, but not the only one. All algebraic links (not only torus knots) are necessary to employ the technique of

the *resolved conifold* and similar tools used in [31] to prove the (colored generalization of) the *OS conjecture* concerning the HOMFLYPT polynomials; see Oblomkov and Shende [37]. Also, all algebraic *links* are needed for the theory of Hitchin and affine Springer fibers since *spectral curves* are generally not unibranch. Topologically, the class of iterated links is closed with respect to *cabling*, a major operation in knot theory.

ORS conjecture Let us briefly comment on Conjecture 2 from [36]; see Section 5 for some further discussion. It relates the geometry of *nested Hilbert schemes* of arbitrary (germs of) plane curve singularities to the Khovanov–Rozansky *unreduced* stable polynomials of the corresponding links. The main component of their conjecture versus the OS conjecture is the *weight filtration*. The polynomials there, $\mathscr{P}_{\mathrm{alg}}$, conjecturally coincide with uncolored ones in the q,t–DAHA theory (upon switching to the standard parameters); see [8]. They are connected with the *perverse filtration* on the cohomology of the compactified Jacobians from Maulik and Yun [32] and Migliorini and Shende [34] (see Proposition 4 in [36]).

Our approach is based on *admissible flags* of submodules $\{\mathcal{M}_i\}$ in the normalization ring; here $\dim_{\mathbb{C}}(\mathcal{M}_{i+1}/\mathcal{M}_i)=1$, but they are not full flags, and the admissibility is a very restrictive condition. The absence of (nested) Hilbert schemes is due to the *reduced setting* of our paper (continuing [8]); there are other important deviations from [36]. For instance, we do not need the weight filtration, and our approach is quite computable. There may be a connection with Section 9.1 from [22] (a reduced version of the construction of [36]), but this is unclear. Actually, the weight filtration appears naturally in (5-2), which follows from a modular variant (2-7) of our conjecture, but it is associated with a parameter different from that in [36].

Main results The key is Conjecture 2.5; anything else is about confirmations, examples and connections. It extends Conjecture 2.4(iii) from [8] for Betti numbers of Jacobians factors for unibranch plane curve singularities (the case $q=1$, $a=0$). It was essentially checked in Mellit [33] for torus knots. The Betti numbers for torus knots are due to Lusztig and Smelt [30] (see also [19] and [38]). We focus in this paper on nontorus knots.

The series of the plane curve singularities for *Puiseux exponents* $(4,2u,v)$ for odd u,v and $v>2u>4$ is the simplest of nontorus type; the corresponding links are $\mathrm{Cab}(2u+v,2)T(u,2)$. Here we generalize the formulas from [38] for the dimensions of cells in the corresponding CW-presentation. The most convincing demonstrations of

our main conjecture are the examples where such cells are not all affine. Such examples are well beyond [38] and are actually a new vintage in the theory of compactified Jacobians as well as our flagged generalization.

Some perspectives An extension of the geometric approach to superpolynomials from this paper to *all* root systems is of obvious interest, especially due to connections with p–adic orbital integrals. Cherednik and Elliot [10] hint that such a uniform theory may exist, in spite of the fact that there can be no rank stabilization for the systems EFG. Such a theory can be expected to provide *refined* generalizations of *orbital integrals* from the geometric fundamental lemma; local spectral curves are taken here as plane curve singularities. See the end of the paper.

The case $q = 1$, $a = 0$ is directly related to p–adic orbital integrals of nil-elliptic type A. An immediate corollary of Conjecture 2.5 is that such orbital integrals are *topological* invariants of the corresponding plane curve singularities. This readily follows from [30] for torus knots, but seems beyond any existing approaches for nontorus singularities, especially in the presence of nonaffine cells (see the online supplement). This invariance, the refined orbital integrals, the connections with HOMFLYPT homology and an extension of our paper to arbitrary algebraic links, any colors and all root systems are natural challenges.

1 DAHA superpolynomials

We will provide here the main facts of DAHA theory needed for the definition of the DAHA–Jones polynomials and DAHA superpolynomials. See [6; 7; 5] for details. The construction is totally uniform for any root systems and weights.

1.1 Definition of DAHA

Let $R = \{\alpha\} \subset \mathbb{R}^n$ be a root system of type $A_n, \dots, G_2$ with respect to a euclidean form $(\cdot, \cdot)$ on $\mathbb{R}^n$, W the Weyl group generated by the reflections s_α, and R_+ the set of positive roots corresponding to fixed simple roots $\alpha_1, \dots, \alpha_n$. The form is normalized by the condition $(\alpha, \alpha) = 2$ for *short* roots. The weight lattice is $P = \bigoplus_{i=1}^n \mathbb{Z}\omega_i$, where $\{\omega_i\}$ are fundamental weights. The root lattice is $Q = \bigoplus_{i=1}^n \mathbb{Z}\alpha_i$. Replacing $\mathbb{Z}$ by $\mathbb{Z}_+ = \{\mathbb{Z} \ni m \geq 0\}$, we obtain P_+, Q_+; see eg [3] or [5].

Setting $v_\alpha := \frac{1}{2}(\alpha, \alpha)$, the vectors $\widetilde{\alpha} = [\alpha, v_\alpha j] \in \mathbb{R}^n \times \mathbb{R} \subset \mathbb{R}^{n+1}$ for $\alpha \in R$ and $j \in \mathbb{Z}$ form the *twisted affine root system* $\widetilde{R} \supset R$ ($z \in \mathbb{R}^n$ is identified with $[z, 0]$).

We add $\alpha_0 := [-\vartheta, 1]$ to the simple roots for the *maximal short root* $\vartheta \in R_+$. The corresponding set $\widetilde{R}_+$ of positive roots is $R_+ \cup \{[\alpha, \nu_\alpha j] \mid \alpha \in R,\ j > 0\}$.

The set of the indices of the images of α_0 by all automorphisms of the affine Dynkin diagram will be denoted by O (with $O = \{0\}$ for E_8, F_4, G_2). Let $O' := \{r \in O \mid r \neq 0\}$; then $O' = [1, \dots, n]$ for A_n. The elements ω_r for $r \in O'$ are *minuscule weights*, defined by the inequalities $(\omega_r, \alpha^\vee) \leq 1$ for all $\alpha \in R_+$. We set $\omega_0 = 0$ for the sake of uniformity.

Affine Weyl groups Given $\widetilde{\alpha} = [\alpha, \nu_\alpha j] \in \widetilde{R}$ and $b \in P$, let

$$s_{\widetilde{\alpha}}(\widetilde{z}) = \widetilde{z} - (z, \alpha^\vee)\widetilde{\alpha}, \quad b'(\widetilde{z}) = [z, \zeta - (z, b)] \tag{1-1}$$

for $\widetilde{z} = [z, \zeta] \in \mathbb{R}^{n+1}$. The *affine Weyl group* $\widetilde{W} = \langle s_{\widetilde{\alpha}} \mid \widetilde{\alpha} \in \widetilde{R}_+ \rangle$ is the semidirect product $W \ltimes Q$ of its subgroups $W = \langle s_\alpha \mid \alpha \in R_+ \rangle$ and Q, where α is identified with

$$s_\alpha s_{[\alpha, \nu_\alpha]} = s_{[-\alpha, \nu_\alpha]} s_\alpha \quad \text{for } \alpha \in R \text{ considered in } \langle s_{\widetilde{\alpha}} \rangle.$$

Using the presentation of $\widetilde{W}$ as $W \ltimes Q$, the *extended Weyl group* $\widehat{W}$ can be defined as $W \ltimes P$, where the corresponding action is

$$(wb)([z, \zeta]) = [w(z), \zeta - (z, b)] \quad \text{for } w \in W,\ b \in P. \tag{1-2}$$

It is canonically isomorphic to $\widetilde{W} \ltimes \Pi$ for $\Pi := P/Q$. The latter group consists of $\pi_0 = \mathrm{id}$ and the images π_r of minuscule ω_r in P/Q.

The group Π will be naturally identified with the subgroup of $\widehat{W}$ of the elements of the length zero; the *length* is defined as follows:

$$l(\widehat{w}) = |\lambda(\widehat{w})| \quad \text{for } \lambda(\widehat{w}) := \widetilde{R}_+ \cap \widehat{w}^{-1}(-\widetilde{R}_+).$$

One has $\omega_r = \pi_r u_r$ for $r \in O'$, where u_r is the (unique) element $u \in W$ of minimal length such that $u(\omega_r) \in -P_+$.

Setting $\widehat{w} = \pi_r \widetilde{w} \in \widehat{W}$ for $\pi_r \in \Pi$ and $\widetilde{w} \in \widetilde{W}$, the length $l(\widehat{w})$ coincides with the length of any reduced decomposition of $\widetilde{w}$ in terms of the simple reflections $s_i, 0 \leq i \leq n$ (a standard and important fact).

Let m be the least natural number such that $(P, P) = (1/\mathsf{m})\mathbb{Z}$; then $\mathsf{m} = n + 1$ for A_n. The double affine Hecke algebra, DAHA, depends on the parameters q, t_ν $(\nu \in \{\nu_\alpha\})$; to be exact, it is defined over the ring of polynomials in terms of $q^{\pm 1/\mathsf{m}}$ and $\{t_\nu^{\pm 1/2}\}$.

For $\widetilde{\alpha} = [\alpha, \nu_\alpha j] \in \widetilde{R}$, we set $t_{\widetilde{\alpha}} = t_\alpha = t_{\nu_\alpha}$ and $q_{\widetilde{\alpha}} = q_\alpha = q^{\nu_\alpha}$, and introduce $k_{\widetilde{\alpha}} = k_\alpha = k_{\nu_\alpha}$ from the relation $t_\nu = q^{\nu k_\nu}$. For $i = 1, \dots, n$, let

$$\rho_k := \frac{1}{2}\sum_{\alpha>0} k_\alpha \alpha = k_{\mathrm{sht}}\rho_{\mathrm{sht}} + k_{\mathrm{lng}}\rho_{\mathrm{lng}}, \quad \rho_\nu = \frac{1}{2}\sum_{\nu_\alpha=\nu} \alpha = \sum_{\nu_{\alpha_i}=\nu} \omega_i,$$

where sht and lng are used for short and long roots. We note that the specialization $k_{\mathrm{sht}} = 1 = k_{\mathrm{lng}}$ corresponds to quantum groups and provides the *WRT invariants* in the construction below; see [6].

For pairwise commutative $X_1, \dots, X_n$,

$$X_{\widetilde{b}} := \prod_{i=1}^{n} X_i^{l_i} q^j \quad \text{if } \widetilde{b} = [b, j],\ \widehat{w}(X_{\widetilde{b}}) = X_{\widehat{w}(\widetilde{b})}, \tag{1-3}$$

where

$$b = \sum_{i=1}^{n} l_i \omega_i \in P, \quad j \in (1/\mathsf{m})\mathbb{Z}, \quad \widehat{w} \in \widehat{W}.$$

For instance, $X_0 := X_{\alpha_0} = X_{[-\vartheta,1]} = qX_\vartheta^{-1}$.

Recall that $\omega_r = \pi_r u_r$ for $r \in O'$ (see above). Note that π_r^{-1} is $\pi_{\iota(i)}$, where ι is the standard involution of the *nonaffine* Dynkin diagram, induced by $\alpha_i \mapsto -w_0(\alpha_i)$; it is the reflection of $[1, \dots, n]$ in type A_n. Finally, we set $m_{ij} = 2, 3, 4, 6$ when the number of links between α_i and α_j in the affine Dynkin diagram is $0, 1, 2, 3$.

Definition 1.1 The double affine Hecke algebra $\mathcal{HH}$ is generated by the elements $\{T_i \mid 0 \le i \le n\}$, pairwise commutative $\{X_b \mid b \in P\}$ satisfying (1-3) and the group Π, where the following relations are imposed:

(o) $(T_i - t_i^{1/2})(T_i + t_i^{-1/2}) = 0$ for $0 \le i \le n$;

(i) $T_iT_jT_i\cdots = T_jT_iT_j\cdots$, with m_{ij} factors on each side;

(ii) $\pi_r T_i \pi_r^{-1} = T_j$ if $\pi_r(\alpha_i) = \alpha_j$;

(iii) $T_i X_b = X_b X_{\alpha_i}^{-1} T_i^{-1}$ if $(b, \alpha_i^\vee) = 1$ for $0 \le i \le n$;

(iv) $T_i X_b = X_b T_i$ if $(b, \alpha_i^\vee) = 0$ for $0 \le i \le n$;

(v) $\pi_r X_b \pi_r^{-1} = X_{\pi_r(b)} = X_{u_r^{-1}(b)} q^{(\omega_{\iota(r)}, b)}$ for $r \in O'$.

Given $\widetilde{w} \in \widetilde{W}, r \in O$, the product

$$T_{\pi_r \widetilde{w}} := \pi_r T_{i_l} \cdots T_{i_1}, \quad \text{where } \widetilde{w} = s_{i_l} \cdots s_{i_1} \text{ and } l = l(\widetilde{w}), \tag{1-4}$$

does not depend on the choice of the reduced decomposition. Moreover,

$$T_{\widehat{v}} T_{\widehat{w}} = T_{\widehat{v}\widehat{w}} \quad \text{whenever} \quad l(\widehat{v}\widehat{w}) = l(\widehat{v}) + l(\widehat{w}), \qquad \text{for } \widehat{v}, \widehat{w} \in \widehat{W}. \tag{1-5}$$

In particular, we arrive at the pairwise commutative elements

$$Y_b := \prod_{i=1}^{n} Y_i^{l_i} \qquad \text{if} \quad b = \sum_{i=1}^{n} l_i \omega_i \in P, \quad Y_i := T_{\omega_i}, \quad b \in P. \tag{1-6}$$

1.2 Main features

The following maps can be (uniquely) extended to automorphisms of $\mathcal{H}$, where $q^{1/(2m)}$ must be added to the ring of constants [5, (3.2.10)–(3.2.15)]:

$$\tau_+ \colon \begin{cases} X_b \mapsto X_b, \\ Y_r \mapsto X_r Y_r q^{-(\omega_r, \omega_r)/2}, \\ T_i \mapsto T_i \ (i > 0), \\ T_0 \mapsto q^{-1} X_\vartheta T_0^{-1}, \\ \pi_r \mapsto q^{-(\omega_r, \omega_r)/2} X_r \pi_r \ (r \in O'), \end{cases} \qquad \tau_- \colon \begin{cases} Y_b \mapsto Y_b, \\ X_r \mapsto Y_r X_r q^{(\omega_r, \omega_r)/2}, \\ T_i \mapsto T_i \ (i \geq 0), \\ X_\vartheta \mapsto q T_0 X_\vartheta^{-1} T_{s_\vartheta}^{-1}, \end{cases} \tag{1-7}$$

$$\sigma := \tau_+ \tau_-^{-1} \tau_+ = \tau_-^{-1} \tau_+ \tau_-^{-1}. \tag{1-8}$$

These automorphisms fix t_v, q and their fractional powers.

The span of $\tau_\pm$ is the *projective* $\mathrm{PSL}_2(\mathbb{Z})$ (due to Steinberg), which is isomorphic to the braid group B_3. Let us list the matrices corresponding to the automorphisms above upon the natural projection onto $\mathrm{SL}_2(\mathbb{Z})$, which is upon the specialization $t_v^{1/(2m)}, q^{1/(2m)} \mapsto 1$. The matrix $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ will represent the map

$$X_b \mapsto X_b^\alpha Y_b^\gamma, \quad Y_b \mapsto X_b^\beta Y_b^\delta$$

for $b \in P$. One has $\tau_+ \mapsto \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $\tau_- \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, and $\sigma \mapsto \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.

We note that there are some simplifications with the definition of DAHA and $\tau_\pm$ for A_n and in Theorem 1.2(i), but they are not significant (the theory is very much uniform for any root systems). However, A_n is obviously needed in part (ii) in this theorem.

Following [5], we use the PBW Theorem to express any $H \in \mathcal{H}$ in the form $\sum_{a,w,b} c_{a,w,b} X_a T_w Y_b$ for $w \in W$ and $a, b \in P$ (this presentation is unique). Then

we substitute

$$\{\cdot\}_{\mathrm{ev}}\colon X_a \mapsto X_a(q^{-\rho_k}) = q^{-(\rho_k,a)}, \quad Y_b \mapsto q^{(\rho_k,b)}, \quad T_i \mapsto t_i^{1/2}. \tag{1-9}$$

The functional $\mathcal{HH} \ni H \mapsto \{H\}_{\mathrm{ev}}$, which is called *coinvariant*, acts via the projection $H \mapsto H\Downarrow := H(1)$ of $\mathcal{HH}$ onto the *polynomial representation* $\mathcal{V}$, which is the $\mathcal{HH}$–module induced from the one-dimensional character $T_i(1) = t_i^{-1/2} = Y_i(1)$ for $1 \le i \le n$ and $T_0(1) = t_0^{-1/2}$. Here $t_0 = t_{\mathrm{sht}}$; see [5; 6; 7].

The polynomial representation $\mathcal{V}$ is linearly generated by X_b $(b \in P)$, and the action of T_i $(0 \le i \le n)$ there is given by the *Demazure–Lusztig operators*:

$$T_i = t_i^{1/2} s_i + (t_i^{1/2} - t_i^{-1/2})(X_{\alpha_i} - 1)^{-1}(s_i - 1), \quad 0 \le i \le n. \tag{1-10}$$

The elements X_b become the multiplication operators and π_r $(r \in O')$ act via the general formula $\widehat{w}(X_b) = X_{\widehat{w}(b)}$ for $\widehat{w} \in \widehat{W}$.

Macdonald polynomials The Macdonald polynomials $P_b(X)$, $b \in P_+$ are uniquely defined as follows. For $c \in P$, let c_+ be a unique element such that $c_+ \in W(c) \cap P_+$. Given $b \in P_+$ and assuming that $c \in P$ is such that $b \ne c_+ \in b - Q_+$,

$$P_b - \sum_{a \in W(b)} X_a \in \bigoplus_c \mathbb{Q}(q, t_v) X_c \quad \text{and} \quad \mathrm{CT}\big(P_b X_{c'} \mu(X; q, t)\big) = 0, \tag{1-11}$$

where

$$\mu(X; q, t) := \prod_{\alpha \in R_+} \prod_{j=0}^{\infty} \frac{(1 - X_\alpha q_\alpha^j)(1 - X_\alpha^{-1} q_\alpha^{j+1})}{(1 - X_\alpha t_\alpha q_\alpha^j)(1 - X_\alpha^{-1} t_\alpha q_\alpha^{j+1})}.$$

Here CT is the constant term; μ is considered a Laurent series of X_b with the coefficients expanded in terms of positive powers of q. The coefficients of P_b belong to the field $\mathbb{Q}(q, t_v)$. The following evaluation formula (the Macdonald evaluation conjecture) is important to us:

$$(P_b(q^{-\rho_k})) = q^{-(\rho_k,b)} \prod_{\alpha > 0} \prod_{j=0}^{(\alpha^\vee,b)-1} \left(\frac{1 - q_\alpha^j t_\alpha X_\alpha(q^{\rho_k})}{1 - q_\alpha^j X_\alpha(q^{\rho_k})} \right). \tag{1-12}$$

1.3 Algebraic knots

Torus knots $T(\mathsf{r}, \mathsf{s})$ are defined for any integers $\mathsf{r}, \mathsf{s} > 0$ such that $\gcd(\mathsf{r}, \mathsf{s}) = 1$. One has the symmetry $T(\mathsf{r}, \mathsf{s}) = T(\mathsf{s}, \mathsf{r})$, where we use "$=$" for the ambient isotopy equivalence. Also $T(\mathsf{r}, 1) = \bigcirc$ for the *unknot*, denoted by $\bigcirc$.

Algebraic knots $\mathcal{T}(\vec{r},\vec{s})$ are associated with two sequences of (strictly) positive integers:

$$\vec{r}=\{r_1,\dots,r_\ell\},\quad \vec{s}=\{s_1,\dots,s_\ell\}\quad \text{such that } \gcd(r_i,s_i)=1; \tag{1-13}$$

we will call ℓ the *length* of $\vec{r},\vec{s}$. The pairs $\{r_i,s_i\}$ are *characteristic* or *Newton pairs*.

We will need one more sequence:

$$a_1=s_1,\quad a_i=a_{i-1}r_{i-1}r_i+s_i\ (i=2,\dots,m). \tag{1-14}$$

See eg [12] and [38]. Then

$$\mathcal{T}(\vec{r},\vec{s}):=\mathrm{Cab}(\vec{a},\vec{r})(\bigcirc)=(\mathrm{Cab}(a_\ell,r_\ell)\cdots\mathrm{Cab}(a_2,r_2))(T(a_1,r_1)) \tag{1-15}$$

in terms of the *cabling* defined below. Note that the first iteration (application of Cab) is for $\{a_1,r_1\}$ (not for the last pair!).

Cabling The *cabling* $\mathrm{Cab}(a,b)(K)$ of any oriented knot K in (oriented) S^3 is defined as follows; see eg [35; 12]. We consider a small 2–dimensional torus around K and put there the torus knot $T(a,b)$ in the direction of K, which is $\mathrm{Cab}(a,b)(K)$ (up to ambient isotopy).

This procedure depends on the order of a,b and the orientation of K. We choose the latter in the standard way (compatible with almost all sources, including the Mathematica package "KnotTheory"); the parameter a gives the number of turns around K. This construction also depends on the framing of the cable knots; we take the natural one, associated with the parallel copy of the torus where a given cable knot sits (its parallel copy has zero linking number with this knot).

By construction, $\mathrm{Cab}(a,0)(K)=\bigcirc$ and $\mathrm{Cab}(a,1)(K)=K$ for any knot K and $a\neq 0$. See [8] for further discussion of relations. The pairs $\{r_i,a_i\}$ are sometimes called *topological*; the isotopy equivalence of algebraic knots generally can be seen only at the level of r,a–parameters (not at the level of the Newton or Puiseux pairs).

Newton–Puiseux theory Given a sequence $r_i,s_i>0$ of Newton (characteristic) pairs, the knot $\mathcal{T}(\vec{r},\vec{s})$ is the link of the *germ of the singularity*

$$y=x^{s_1/r_1}(c_1+x^{s_2/(r_1r_2)}(c_2+\cdots+x^{s_\ell/(r_1r_2\cdots r_\ell)}))\quad \text{at } 0, \tag{1-16}$$

which is the intersection of the corresponding plane curve with a small 3–dimensional sphere in $\mathbb{C}^2$ around 0. We will always assume in this paper that this germ is *unibranch*.

The inequality $\mathsf{s}_1 < \mathsf{r}_1$ is commonly imposed here (otherwise x and y can be switched). Formula (1-16) is the celebrated *Newton–Puiseux expansion*; see eg [12]. All algebraic knots can be obtained in such a way.

Jacobian factors One can associate with a unibranch $\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}$ the *Jacobian factor* $J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}})$. Up to a homeomorphism, it can be introduced as the canonical compactification of the *generalized Jacobian* of an integral *rational* planar curve with $\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}$ as its *only* singularity. It has a purely local definition, which we will use below. Its dimension is the *δ–invariant* of the singularity $\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}$, also called the *arithmetic genus*.

Calculating the *Euler number* $e(J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}))$, the topological Euler characteristic of $J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}})$, and the corresponding *Betti numbers* in terms of $\vec{\mathsf{r}},\vec{\mathsf{s}}$ is a challenging problem. For torus knots $T(\mathsf{r},\mathsf{s})$, one has $e(J(\mathcal{C}_{\mathsf{r},\mathsf{s}})) = (1/(\mathsf{r}+\mathsf{s}))\binom{\mathsf{r}+\mathsf{s}}{\mathsf{r}}$ due to [2]. This formula is related to the perfect modules of rational DAHA and the combinatorics of generalized Catalan numbers; see eg [19].

The Euler numbers of $J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}})$ were calculated in [38] (the Main Theorem) for the following triples of *Puiseux characteristic exponents*:

$$(4,2u,v),\ (6,8,v),\ (6,10,v) \quad \text{for odd } u,v>0, \tag{1-17}$$

where $4<2u<v$, $8<v$, and $10<v$, respectively. Here $\delta = \dim J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}})$ is $\frac{1}{2}(\mathsf{r}-1)(\mathsf{s}-1)$ for $T(\mathsf{r},\mathsf{s})$ and $2u+\frac{1}{2}(v-1)-1$ for the series $(4,2u,v)$. Generally, δ equals the cardinality $|\mathbb{N}\setminus\Gamma|$, where Γ is the *valuation semigroup* associated with $\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}$; see [38] and [45]. The Euler numbers of the Jacobian factors can be also calculated via the HOMFLYPT polynomials of the corresponding links (see below) due to [37; 31].

Concerning the Betti numbers for the torus knots and the series $(4,2u,v)$, the odd (co)homology of $J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}})$ vanishes. The formulas for the corresponding even Betti numbers $h^{(2k)} = \mathrm{rk}(H^{2k}(J(\mathcal{C}_{\vec{\mathsf{r}},\vec{\mathsf{s}}})))$ were calculated explicitly for many values of k in [38], where $0\le k\le\delta$. Not much was and is known/expected beyond these two series.

1.4 DAHA–Jones theory

The following results and conjectures are mainly from [8]; see also [7, Theorem 1.2] and [6; 21].

The construction is given directly in terms of the parameters $\{\vec{\mathsf{r}},\vec{\mathsf{s}}\}$, though it actually depends only on the corresponding topological parameters $\{\vec{\mathsf{a}},\vec{\mathsf{r}}\}$. Recall that torus knots $T(\mathsf{r},\mathsf{s})$ are naturally represented by $\gamma_{\mathsf{r},\mathsf{s}}\in \mathrm{PSL}_2(\mathbb{Z})$ with the first column $(\mathsf{r},\mathsf{s})^{\mathrm{tr}}$

(tr denotes transposition), where $\mathsf{r},\mathsf{s}>0$ and we assume that $\gcd(\mathsf{r},\mathsf{s})=1$. Let $\widehat{\gamma}_{\mathsf{r},\mathsf{s}}$ be any pullback of $\gamma_{\mathsf{r},\mathsf{s}}$ to the projective $\mathrm{PSL}_2(\mathbb{Z})$.

For a polynomial F in terms of fractional powers of q and t_v, the *tilde-normalization* $\widetilde{F}$ will be the result of the division of F by the lowest q,t_v–monomial, assuming that it is well defined. We put $q^\bullet t^\bullet$ for a monomial factor (possibly fractional) in terms of q,t_v. See [8] for the following theorem. We will also apply this definition to the superpolynomials, where the lowest q,t–monomial is picked from the constant a–term.

Theorem 1.2 *Let R be a reduced irreducible root system. Recall $H\mapsto H\Downarrow := H(1)$, where the action of $H\in\mathcal{HH}$ in $\mathcal{V}$ is used.*

(i) *Given two strictly positive sequences $\vec{\mathsf{r}},\vec{\mathsf{s}}$ of length ℓ as in* (1-13)*, we lift $(\mathsf{r}_i,\mathsf{s}_i)^{\mathrm{tr}}$ to γ_i and then to $\widehat{\gamma}_i$ (acting in $\mathcal{HH}$) as above. For a weight $b\in P_+$, the* **DAHA–*Jones polynomial*** *is*

$$\mathrm{JD}^R_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(b;q,t)=\mathrm{JD}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(b;q,t) := \big\{\widehat{\gamma}_1\big(\cdots\big(\widehat{\gamma}_{\ell-1}\big((\widehat{\gamma}_\ell(P_b)/P_b(q^{-\rho_k}))\Downarrow\big)\Downarrow\big)\cdots\big)\big\}_{\mathrm{ev}}. \tag{1-18}$$

It does not depend on the particular choice of the lifts γ_i and $\widehat{\gamma}_i\in\mathrm{GL}_2^\wedge(\mathbb{Z})$. The ***tilde-normalization*** *$\widetilde{\mathrm{JD}}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(b;q,t)$ is well defined and is a polynomial in terms of q,t_v with the constant term 1.*

(ii) *Let us switch to the root system A_n for $\mathfrak{sl}_{n+1}$, setting $t=t_{\mathrm{sht}}=q^k$ and considering $P_+\ni b=\sum_{i=1}^n b_i\omega_i$ as (dominant) weights for any A_m (for $\mathfrak{sl}_{m+1}$) with $m\ge n-1$, where we assume that $\omega_n=0$ upon the restriction to A_{n-1}.*

Then given $\mathcal{T}(\vec{\mathsf{r}},\vec{\mathsf{s}})$ as above, there exists a **DAHA-*superpolynomial*** *$\mathcal{H}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(b;q,t,a)$ from $\mathbb{Z}[q,t^{\pm1},a]$ satisfying the relations*

$$\mathcal{H}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(b;q,t,a=-t^{m+1})=\widetilde{\mathrm{JD}}^{A_m}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(b;q,t)\quad \text{for any } m\ge n-1; \tag{1-19}$$

its a–constant term is automatically tilde-normalized. □

Topological connection Let us briefly discuss the conjectural relation of DAHA-superpolynomials to *stable Khovanov–Rozansky polynomials* denoted by $\mathrm{KhR}_{\mathrm{stab}}$; see [27; 28; 26; 39]. We consider only the reduced setting (actually not quite developed topologically).

The passage to the Khovanov–Rozansky polynomials for $\mathfrak{sl}_N$ for sufficiently large N is the substitution $a\mapsto t^N\sqrt{q/t}$. Note the relation to the *Heegaard–Floer homology* for $N=0$. Equivalently, this passage is $a_{\mathrm{st}}\mapsto q_{\mathrm{st}}^N$ in the *standard topological parameters*

(also used in the *ORS conjecture*), which are related to the DAHA parameters as follows:

$$t = q_{\mathrm{st}}^2, \quad q = (q_{\mathrm{st}}t_{\mathrm{st}})^2, \quad a = a_{\mathrm{st}}^2 t_{\mathrm{st}}, \quad q_{\mathrm{st}}^2 = t, \quad t_{\mathrm{st}} = \sqrt{q/t}, \quad a_{\mathrm{st}}^2 = a\sqrt{t/q}. \tag{1-20}$$

For the DAHA-superpolynomials from (1-19),

$$\mathcal{H}_{\vec{r},\vec{s}}(\square; q, t, a)_{\mathrm{st}} = \widetilde{\mathrm{KhR}}_{\mathrm{stab}}(q_{\mathrm{st}}, t_{\mathrm{st}}, a_{\mathrm{st}}) \quad \text{where } \square = \omega_1, \tag{1-21}$$

and $\widetilde{\mathrm{KhR}}_{\mathrm{stab}}$ is *reduced* $\mathrm{KhR}_{\mathrm{stab}}$ divided by the smallest power of a_{st} and then by $q_{\mathrm{st}}^{\bullet}t_{\mathrm{st}}^{\bullet}$ such that $\widetilde{\mathrm{KhR}}_{\mathrm{stab}}(a_{\mathrm{st}} = 0) \in \mathbb{Z}_+[q_{\mathrm{st}}, t_{\mathrm{st}}]$ with the constant term 1. Here $\{\cdot\}_{\mathrm{st}}$ means the switch from the DAHA parameters to the *standard topological parameters*.

Also, the polynomials $\mathrm{KhR}_{\mathrm{stab}}$ are expected to coincide with the (reduced) physics *superpolynomials* based on the *BPS states* [11; 1; 14; 18] and those obtained in terms of *rational* DAHA [22; 21] for torus knots. The latter approach is developed so far only for torus knots and in the uncolored case; there is some progress for symmetric powers of the fundamental representation; see [18]. We will not touch the connections with rational DAHA in this paper. Concerning physics origins, let us mention that using the Macdonald polynomials at roots of unity q (for $t = q^k, k \in \mathbb{Z}_+$) instead of Schur functions in the usual construction of *knot operators* was suggested in [1].

Betti numbers $h^{(i)} = \mathrm{rk}\, H^i(J(\mathcal{C}_{\vec{r},\vec{s}}))$ We are very much motivated by the DAHA approach to these numbers. Technically, we generalize the interpretation of Gorsky's superpolynomials for torus knots at $a = 0$ from [19; 20] and the following conjecture from [8]:

$$\mathcal{H}_{\vec{r},\vec{s}}(\square; q = 1, t, a = 0) = \sum_{i=0}^{2\delta} h^{(i)} t^{i/2} \quad \text{for } \delta = \dim J(\mathcal{C}_{\vec{r},\vec{s}}). \tag{1-22}$$

It implies that $h^{\mathrm{odd}} = 0$ (the van Straten–Warmt conjecture). Relation (1-22) will be generalized below to the whole superpolynomial $\mathcal{H}_{\vec{r},\vec{s}}(\square; q, t, a)$, which is the main result of our paper.

2 Geometric superpolynomials

2.1 Modules of semigroups

Let $\mathcal{R}$ be the complete local ring of the *unibranch* germ of the plane curve singularity, embedded into the normalization ring $\mathcal{O} = \mathbb{C}[\![z]\!]$. The *conductor* of $\mathcal{R}$ is the smallest c

such that $z^{\mathsf{c}}\mathcal{O} \subset \mathcal{R}$; actually it is the ideal $(z^{\mathsf{c}}) \subset \mathcal{O}$, but we will call c the conductor in this paper. We set $\mathcal{K} := \mathbb{C}((z))$.

The corresponding *semigroup* $\Gamma_{\mathcal{R}}$ is formed by the orders of the smallest powers, ie *valuations* $v(x)$ (minimal z–degrees) $x \in \mathcal{R}$. The δ–invariant (the arithmetic genus) is then $\delta_{\mathcal{R}} = |\mathbb{Z}_+ \setminus \Gamma_{\mathcal{R}}|$. We will call $\mathbb{Z}_+ \setminus \Gamma_{\mathcal{R}}$ the *set of gaps* and denote it by $G_{\mathcal{R}}$; thus $\delta = |G_{\mathcal{R}}|$. Also, $\mathsf{c} = 2\delta$.

Compactified Jacobians for projective curves are generally defined as the varieties of coherent torsion free sheaves of rank one and fixed degree up to isomorphisms. The *Jacobian factor* $J_{\mathcal{R}}$, we are going to define, is a local version of the compactified Jacobian. It is (as a set) formed by all finitely generated $\mathcal{R}$–submodules $\mathcal{M} \subset \mathcal{K} = \mathbb{C}((z))$, of (any) prescribed degree, also called *$\mathcal{R}$–lattices*.

We define *$\mathcal{O}$–degree* $\deg_{\mathcal{O}}\mathcal{M}$ with respect to $\mathcal{O}$; it is $\dim_{\mathbb{C}}(\mathcal{O}/\mathcal{M})$ if $\mathcal{M} \subset \mathcal{O}$. For arbitrary submodules $\mathcal{M} \subset \mathcal{K}$,

$$\deg_{\mathcal{O}}(\mathcal{M}) = \dim_{\mathbb{C}}(\mathcal{O}/(\mathcal{O}\cap\mathcal{M})) - \dim_{\mathbb{C}}(\mathcal{M}/(\mathcal{O}\cap\mathcal{M})). \tag{2-1}$$

This definition is a natural counterpart of the degree of a divisor at a given point (here at $z = 0$) in the smooth situation. Actually we will mainly need below $-\deg_{\mathcal{R}}(\mathcal{M}) = \delta - \deg_{\mathcal{O}}(\mathcal{M})$.

The valuations $v(x)$ of the elements $x \in \mathcal{M} \subset \mathcal{K}$ form a $\Gamma_{\mathcal{R}}$–module $\Delta_{\mathcal{M}}$; the modules for semigroups Γ (with 0) are subsets $\Delta \subset \mathbb{Z}_+$ such that $\Gamma + \Delta \subset \Delta$. Unless stated otherwise, we assume that $\mathcal{M} \subset \mathcal{O}$ and that it contains the element $1 + \sum_{i=1}^{\mathsf{c}-1} \lambda_0^i z^i$ (of valuation 0), such an embedding can be achieved by the division by z^m for $m = \min(\Delta_{\mathcal{M}})$. Here the upper limit $\mathsf{c}-1$ is sufficient in the sum due to the definition of the conductor. The notation $\Delta_{\circ}$ is used for such a normalization in [38], we call it the *standard normalization*. See there for these and related definitions and facts.

For a *standard* $\mathcal{M}$, we will use the notation $D_{\mathcal{M}}$ or $D[\mathcal{M}]$ for $G_{\mathcal{R}} \cap \Delta_{\mathcal{M}}$ and call it the *set of added gaps* or simply the *D–set*. The square brackets will be used for the list of its elements. For instance, $D_{\mathcal{R}} = \varnothing = []$ corresponds to the (trivial) Γ–module $\Delta_{\mathcal{R}} = \Gamma$, and $D_{\mathcal{O}} = G_{\mathcal{R}} = [1, \dots, \mathsf{c}-1]$ for $\mathcal{M} = \mathcal{O}$ (recall that $\mathsf{c}-1$ is the last gap in G).

Due to the normalization we impose, one has

$$\deg_{\mathcal{O}}(\mathcal{M}) = \delta - |D_{\mathcal{M}}|, \quad -\deg_{\mathcal{R}}(\mathcal{M}) = |D_{\mathcal{M}}| \qquad \text{for standard } \mathcal{M}. \tag{2-2}$$

Not all Γ–modules Δ can be realized as $\Delta_{\mathcal{M}}$ for nontorus singularities. Recall that torus knots $T(\mathsf{r},\mathsf{s})$ are associated with the rings $\mathcal{R}=\mathbb{C}[\![z^{\mathsf{r}},z^{\mathsf{s}}]\!]$. The simplest example of a nontorus singularity is $\mathcal{R}=\mathbb{C}[\![z^4,z^6+z^7]\!]$ with $\Gamma=\langle 4,6,13\rangle$ and $\mathsf{c}=16$. Then the sets of added gaps $D=[2,15]$ and $D=[2,11,15]$ do not come from any modules $\mathcal{M}$; following [38], we call them *nonadmissible*. All other D are admissible in this case (ie can be obtained as $D_{\mathcal{M}}$).

2.2 $J_{\mathcal{R}}$ as a projective variety

Let $J_{\mathcal{R}}[D]\subset J_{\mathcal{R}}$ be the set of modules $\mathcal{M}$ with $D_{\mathcal{M}}=D$, $J_{\mathcal{R}}[d]=\bigcup_{|D|=d}J_r[D]$ and $\bar{J}_{\mathcal{R}}[d]=\bigcup_{d'\geq d}J_r[d']$ for $d\geq 0$. That is, the latter is the set of all standard submodules of $\mathcal{O}$–degree $\delta-d$ or smaller. The set $J_{\mathcal{R}}[0]=J_{\mathcal{R}}[\varnothing]$ is *the big cell*; it is formed by all invertible modules $\mathcal{M}$ (with one generator and of $\mathcal{O}$–degree δ). Also, $J_{\mathcal{R}}[\delta]=\bar{J}_{\mathcal{R}}[\delta]=\{\mathcal{M}=\mathcal{O}\}$ and $J_{\mathcal{R}}[>\delta]=\varnothing$. Finally, we note that the isomorphisms of a particular submodule $\mathcal{M}$ are those induced by the action of the group of units $\mathcal{R}^*$.

We can give $\bar{J}_{\mathcal{R}}[d]$ the structure of a projective subvariety in $J_{\mathcal{R}}=\bar{J}_{\mathcal{R}}[0]$ following [23]; actually they are projective schemes over $\mathbb{Z}$. Generally, the $\mathcal{O}$–degree of $z^a\mathcal{M}$ for an $\mathcal{R}$–module $\mathcal{M}\subset\mathcal{O}$ and $a\in\mathbb{Z}_+$ is $\deg_{\mathcal{O}}\mathcal{M}+a$. Indeed, $\deg_{\mathcal{O}}(z^a\mathcal{M})=|\mathbb{Z}_+\setminus\nu(z^a\mathcal{M})|=a+|\mathbb{Z}_+\setminus\nu(\mathcal{M})|$.

Given $0\leq d\leq\delta$, let $\mathcal{M}^\circ$ be a standard submodule of $\mathcal{O}$–degree $\delta-d^\circ$ for $d^\circ\geq d$ and $a=d^\circ-d$. Then the submodule $\mathcal{M}=z^a\mathcal{M}^\circ$ is of $\mathcal{O}$–degree $\delta-d^\circ+d^\circ-d=\delta-d$, and one has

$$z^{2\delta}\mathcal{O}\subset\mathcal{M}\subset\mathcal{O}\quad\text{and}\quad\dim_{\mathbb{C}}(\mathcal{O}/\mathcal{M})=\delta-d.$$

Vice versa, $\mathcal{M}^\circ=z^{-a}\mathcal{M}$ are standard for any such $\mathcal{R}$–submodules $\mathcal{M}$, where a is the smallest evaluation $\nu(z)$ among all $z\in\mathcal{M}$. Such $\mathcal{M}$ for $d=0$ are called *δ–normalized modules* in [38]; our standard submodules are called there *0–normalized*.

Now consider the Grassmanian $\mathrm{Gr}(\mathcal{O}/z^{2\delta}\mathcal{O},\delta+d)$. A subspace of dimension $\delta+d$ is an $\mathcal{R}$–module if and only if it is invariant under the action of $\mathcal{R}/z^{2\delta}\mathcal{O}$ by multiplication. Thus $\bar{J}_{\mathcal{R}}[d]$ is the set of fixed points in $\mathrm{Gr}(\mathcal{O}/z^{2\delta}\mathcal{O},\delta+d)$ under the group action of $(\mathcal{R}/z^{2\delta}\mathcal{O})^\star$. To obtain the structure we desire, extend the action of $(\mathcal{R}/z^{2\delta}\mathcal{O})^\star$ to $\bigwedge^{\delta+d}\mathcal{O}/z^{2\delta}\mathcal{O}$ and consider the image under the Plücker embedding. The condition of being a fixed point under the action $(\mathcal{R}/z^{2\delta}\mathcal{O})^\star$ defines a linear subspace of the projective space $\mathbb{P}(\bigwedge^{\delta+d}\mathcal{O}/z^{2\delta}\mathcal{O})$, so $J_{\mathcal{R}}$ is the intersection of the image of the Plücker embedding of $\mathrm{Gr}(\mathcal{O}/z^{2\delta}\mathcal{O},\delta+d)$ and this linear subspace. Finally, we go back to our standard modules using the identification maps above.

There is an alternative *intrinsic* interpretation of the projective structure of $J_{\mathcal{R}}$. Let us consider $\mathcal{R}$–submodules $\widetilde{\mathcal{M}} \subset \mathcal{K} \otimes_{\mathbb{C}} \mathbb{C}[\![u]\!]$, finitely generated over $\mathbb{C}[\![u]\!]$, ie a family of submodules $\mathcal{M}$ (formally) analytically depending on a parameter u. Let $\widetilde{\mathcal{M}}$ contain a principle ideal $\mathbb{C}[\![u,z]\!]z^p$ for some p (arbitrarily large).

We define the *enhanced* (*flat*) *limit* $\widetilde{\mathcal{M}}_0$ of $\widetilde{\mathcal{M}}$ as the linear subspace of $\mathcal{K}$ of *all* linear combinations of the vectors in $\widetilde{\mathcal{M}}$ divided by the smallest possible power of u and then evaluated at $u=0$. This is one of the key definitions in the theory of sheaves and bundles over curves.

Such a limit obviously contains the straightforward specialization $\mathcal{M}_0 := \widetilde{\mathcal{M}}(u=0)$. The space $\widetilde{\mathcal{M}}_0 \subset \mathcal{O}$ has a natural structure of a $\mathcal{R}$–submodule by construction. Therefore it becomes a standard module (an element of $J_{\mathcal{R}}$) upon the division by a proper power z^n $(n \geq 0)$.

As in the definition of the standard normalization, we assume here that $\widetilde{\mathcal{M}}$ contains an element with zero z–valuation; its z–constant term can be any nonzero element (possibly noninvertible) from $\mathbb{C}[\![u]\!]$. By the way, one can check that the enhanced limit will remain the same if the principle ideal $\mathbb{C}[\![u,z]\!]z^{\mathsf{c}}$ is added to $\widetilde{\mathcal{M}}$; ie p above can be assumed to be no greater than the conductor c. This results from a standard theory of flat limits. Without general theory here, this fact can be justified by enlarging $\widetilde{\mathcal{M}}$ with all linear combinations of the vectors in $\widetilde{\mathcal{M}}$ divided by the corresponding minimal power of u and verifying that this procedure will eventually add $\mathbb{C}[\![u,z]\!]z^{\mathsf{c}}$ to $\mathcal{M}$. The limit $\widetilde{\mathcal{M}}_0$ will remain unchanged under such an "enhancement" of $\widetilde{\mathcal{M}}$.

Finally, the *boundary* of any family of modules considered in $J_{\mathcal{R}}$ is the collection of enhanced limits for all one-parametric subfamilies $\widetilde{\mathcal{M}}$. Using one u is sufficient in this approach (another general fact).

Proposition 2.1 *Let $\mathcal{M}'$ be the enhanced limit $\widetilde{\mathcal{M}}_0$ of a u–family $\widetilde{\mathcal{M}}$ of modules invertible over $\mathbb{C}((u))$, where we assume that $0 \in \Delta_{\widetilde{\mathcal{M}}}$ as above. Setting $\widetilde{\mathcal{M}}_0 \subset (z^{d'}) = \mathcal{O}z^{d'}$ for the maximal such d', one has that $d' \geq d$ for the number of added gaps for the standard module $\mathcal{M}$ corresponding to $\mathcal{M}'$ and $d' = d$ for sufficiently general such $\widetilde{\mathcal{M}}$.*

Proof For generic $\widetilde{\mathcal{M}}$, the limit $\widetilde{\mathcal{M}}_0$ is of degree 0 relative to $\mathcal{R} \in \mathcal{O}$; we use that $J_{\mathcal{R}}$ is a closed subvariety in $\mathrm{Gr}(\mathcal{O}/z^{2\delta}\mathcal{O}, \delta)$. Therefore $z^{-d}\mathcal{M}'$ is standard for $d = |D_{\mathcal{M}}|$ and $d' \geq d$ for any (nongeneric) $\widetilde{\mathcal{M}}$. □

Two examples of enhanced limits To illustrate Proposition 2.1, let us consider $\mathcal{R} = \langle z^4, z^6 + z^9, z^7 \rangle$ with $\Gamma = \{0, 4, 6, 7, 8, 10, 11, \dots\}$, $G = [1, 2, 3, 5, 9]$, $\delta = |G| = 5$, and the conductor $\mathsf{c} = 10$. We take the following family of modules invertible over $\mathbb{C}((u))$:

$$\widetilde{\mathcal{M}} = \mathcal{R}[\![u, z]\!](u^4 + u(z + z^3) + u^3 z^2) + \mathbb{C}[\![u, z]\!] z^{10},$$

which is an invertible $\mathcal{R}$–module over $\mathbb{C}((u))$ (but not over $\mathbb{C}[\![u]\!]$ due to u^4). Then its *enhanced* limit $\widetilde{\mathcal{M}}_0$ is a module with the corresponding set of added gaps $D = [2, 3, 5, 9]$, so $d = |D| = 4$. However, it is contained in (z^5), which is smaller than (z^4) guaranteed by the proposition.

Taking here $\widetilde{\mathcal{M}} = \mathcal{R}[\![u, z]\!](u(1 + z^5) + z^9) + \mathbb{C}[\![u, z]\!] z^{10}$, the limit $\widetilde{\mathcal{M}}_0$ has the same $D = [2, 3, 5, 9]$. Now it strictly belongs to (z^4), so this is the case of a general position for such D.

The cardinalities of the D–sets will be associated below with the parameter q. If $J_{\mathcal{R}}[D]$ are all affine spaces, their dimensions give the powers of t. The third parameter a of our construction will be due to the flag-length of the *flagged Jacobian factors* defined as follows.

2.3 Flagged Jacobian factors

Definition 2.2 For a D–set D and the sequence $\vec{g} = \{g_i \mid 1 \le i \le m\}$ in $G \setminus D$ such that $g_i < g_{i+1}$, the *D–flag* is

$$\mathcal{D} = \{D_0 = D, D_1 = D \cup g_1, D_2 = D \cup \{g_1, g_2\}, \dots, D_m = D \cup \{g_1, \dots, g_m\}\},$$

provided that all $\Delta_i = \Gamma \cup D_i$ are standard Γ–modules. Then the corresponding *flagged Jacobian factor* $J^m_{\mathcal{R}}$ is defined as the union of the following varieties of flags of standard submodules $\mathcal{M} \subset \mathcal{O}$:

$$J^m_{\mathcal{R}}[\mathcal{D}] := \{\mathscr{M} = \{\mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_m\}, \text{ where } D_i = D_{\mathcal{M}_i}\} \tag{2-3}$$

for all $0 \le i \le m$; we will sometimes omit m here. If at least one such flag $\mathscr{M}$ exists, we call the corresponding $\mathcal{D}$ and $\{\Delta_i\}$ *admissible*.

Considering D–flags $\mathcal{D}$ of length m, we set $J^m_{\mathcal{R}}[d > \delta] = \varnothing$,

$$J^m_{\mathcal{R}}[d] := \bigcup_{\mathcal{D}, |D_0| = d} J^m_{\mathcal{R}}[\mathcal{D}] \subset J^m_{\mathcal{R}}, \quad \bar{J}^m_{\mathcal{R}}[d] = \bigcup_{d' \ge d} J^m_{\mathcal{R}}[d'], \tag{2-4}$$

where $d = 0, 1, \dots, \delta$. For $m = 1$, the D–flag $\mathcal{D}$ is obtained from D by adding one gap g and we use the notation $J^1_{\mathcal{R}}[D, g]$. When $m = 0$, we use $J_{\mathcal{R}}[D]$, which is from [38].

Let us begin with some general properties of flags $\mathscr{M}$. We note that our usage of Nakayama's lemma in Proposition 2.3 below is actually similar to that in [36], Section 2.1.

For an arbitrary module $\mathcal{M}$, we set $\mathcal{M}^{(i)} = \mathcal{M} \cap z^i\mathcal{O}$, which is obviously an $\mathcal{R}$–module, and $\overline{\mathcal{M}} = \mathcal{M}/\mathfrak{m}\mathcal{M}$ for the maximal ideal $\mathfrak{m} \subset \mathcal{R}$. Accordingly, $\overline{\mathcal{M}}^{(i)}$ is the image of $\mathcal{M}^{(i)}$ in $\overline{\mathcal{M}}$. Obviously, $\dim \mathcal{M}^{(g)}/\mathcal{M}^{(g+1)} = 1$ if and only if $g \in \Delta_{\mathcal{M}}$, and $\dim \overline{\mathcal{M}}^{(g)}/\overline{\mathcal{M}}^{(g+1)} \le 1$.

Proposition 2.3 (i) *For a D–flag*

$$\mathcal{D} = \{D_0 \subset D_1 = D_0 \cup \{g_1\} \subset \cdots \subset D_m = D_0 \cup \{g_1, \dots, g_m\}\}$$

from Definition 2.2, all $D_0 \cup \{g_i\}$ for $1 \le i \le m$ are D–sets. Let $\mathscr{M} = \{\mathcal{M}_i\}$ be a flag of (standard) modules corresponding to a D–flag (admissible). Picking an arbitrary $m_i \in \mathcal{M}_i$ such that $\upsilon(m_i) = g_i$, the space $\mathcal{M}_0 \oplus \mathbb{C}m_i$ is an $\mathcal{R}$–module.

Then for $1 \le i \le m$, we have $\mathcal{M}_i = \mathcal{M}_0 \oplus \mathbb{C}m_1 \oplus \cdots \oplus \mathbb{C}m_i$, $\mathfrak{m}\mathcal{M}_m \subset \mathcal{M}_0$ and $\dim \overline{\mathcal{M}}^{(g_i)}/\overline{\mathcal{M}}^{(g_i+1)} = 1$. The elements m_i modulo $\mathcal{M}_0$ are uniquely determined up to proportionality, ie depend only on the flag $\mathscr{M}$.

(ii) *Vice versa, for the D–flag $\mathcal{D}$ as above, let us assume the existence of a module $\mathcal{M}_{\mathrm{top}}$ such that $D[\mathcal{M}_{\mathrm{top}}] = D_m$ and $\dim \overline{\mathcal{M}}^{(g_i)}_{\mathrm{top}}/\overline{\mathcal{M}}^{(g_i+1)}_{\mathrm{top}} = 1$ for all i. We do not impose the admissibility of the whole $\mathcal{D}$. Then $\mathcal{M}_m = \mathcal{M}_{\mathrm{top}}$ can be extended to a flag $\mathscr{M}$ corresponding to $\mathcal{D}$ (so it is admissible) and all such flags can be described as follows.*

(a) *For any subspace $\overline{\mathcal{M}}_0 \subset \overline{\mathcal{M}}_{\mathrm{top}}$ such that for $1 \le i \le m$,*

$$\dim \overline{\mathcal{M}}_{\mathrm{top}}/\overline{\mathcal{M}}_0 = m, \quad \dim(\overline{\mathcal{M}}_0 + \overline{\mathcal{M}}^{(g_i)}_{\mathrm{top}})/(\overline{\mathcal{M}}_0 + \overline{\mathcal{M}}^{(g_i+1)}_{\mathrm{top}}) = 1, \quad \overline{\mathcal{M}}_0 \not\subset \overline{\mathcal{M}}^{(1)}_{\mathrm{top}},$$

let $\mathfrak{m}\mathcal{M}_{\mathrm{top}} \subset \mathcal{M}_0 \subset \mathcal{M}_{\mathrm{top}}$ be a unique lift of $\overline{\mathcal{M}}_0$ (the Nakayama lemma).

(b) *Then for any pullbacks, denoted by $\bar{m}_i$, of the generators of the latter quotients to $(\overline{\mathcal{M}}_0 + \overline{\mathcal{M}}^{(g_i)}_{\mathrm{top}})/\overline{\mathcal{M}}_0$, we take $\mathscr{M} = \{\mathcal{M}_i\}$, where $\mathcal{M}_i$ is a similar lift of the space $\overline{\mathcal{M}}_i = \overline{\mathcal{M}}_0 \oplus \mathbb{C}\bar{m}_1 \oplus \cdots \oplus \mathbb{C}\bar{m}_i$ to an $\mathcal{R}$–submodule of $\mathcal{M}_{\mathrm{top}}$ containing $\mathfrak{m}\mathcal{M}_{\mathrm{top}}$. These imply that $D[\mathcal{M}_0] = D_0$, $\mathcal{M}_{i=m} = \mathcal{M}_{\mathrm{top}}$.*

(iii) *Continuing* (ii), *all flags* $\mathscr{M}$ *in* (i) *are uniquely described by the following data:* (α) *the space* $\overline{\mathcal{M}}_0 \subset \overline{\mathcal{M}}_{\mathrm{top}}$ *as above,* (β) *the set of elements* $\bar{m}_i \in (\overline{\mathcal{M}}_0 + \overline{\mathcal{M}}_{\mathrm{top}}^{(g_i)})/\overline{\mathcal{M}}_0$ *considered up to proportionality. That is, for a fixed* $\mathcal{M}_{\mathrm{top}}$ *and* $\overline{\mathcal{M}}_0$ *as in* (ii), *such flags of modules are naturally parametrized by unipotent complex matrices of size* $m \times m$, *which is due to the action of the Borel subgroup preserving the full flag* $\{(\overline{\mathcal{M}}_0 + \overline{\mathcal{M}}_{\mathrm{top}}^{(g_i)})/\overline{\mathcal{M}}_0\}$ *in* $\overline{\mathcal{M}}_{\mathrm{top}}/\overline{\mathcal{M}}_0$.

Proof Part (iii) formally follows from (i,ii), including the equivalence of the admissibility of a D–flag $\mathcal{D}$ and the existence of the pair $\{\mathcal{M}_0, \mathcal{M}_m\}$ for D_0, D_m such that $\mathfrak{m}\mathcal{M}_m \subset \mathcal{M}_0 \subset \mathcal{M}_m$. Let us justify (i,ii). Below $\{g_i\}$ means a single element g_i considered as a set.

First of all, $D_0 \cup \{g_i\}$ correspond to certain Γ–modules for any $1 \le i \le m$. Indeed, adding g_i to D_{i-1} cannot add anything new to D_i (since the corresponding Δ_i is assumed a Γ–module). The same holds for $D_0 \cup \{g_i\}$, since extra elements in the minimal Γ–module containing the latter can be only greater than g_i, ie can be only in $D_i \setminus D_{i-1}$

This reasoning is equally applicable to the flags $\mathscr{M}$ for admissible $\mathcal{D}$. We claim that $D_0 \cup \{g_i\}$ is admissible for each g_i. Indeed, adding $\mathcal{R}m_i \subset \mathcal{M}_i$ to $\mathcal{M}_0$ for an element m_i with the valuation $\upsilon(m_i) = g_i$ cannot create any new valuations versus Δ_0 but g_i, since this does not create them when going from $\mathcal{M}_{i-1}$ to $\mathcal{M}_i$. Thus $\mathcal{M}_0 + \mathcal{R}m_i = \mathcal{M}_0 + \mathbb{C}m_i$ is an $\mathcal{R}$–module corresponding to $D_0 \cup \{g_i\}$.

Next, ψm_i must belong to $\mathcal{M}_0$ for any $\psi \in \mathfrak{m}$. We use that for any $N > 0$ there exists an element $m' \in \mathcal{M}_0$ such that $\upsilon(\psi m_i - m') > N$; thus this difference can be assumed in $\mathcal{M}_0$.

Similarly, $\mathcal{M}_i$ is the linear span of $\mathcal{M}_{i-1}$ and m_i for the elements m_i introduced above, since for every $m \in \mathcal{M}_i$, there exists $m' \in \mathcal{M}_{i-1} + \mathcal{R}m_i$ such that $\upsilon(m - m')$ is greater than any given number. Thus $m - m'$ can be assumed in $\mathcal{M}_{i-1}$. We obtain that adding the elements $\{m_i \mid 1 \le i \le j\}$ to $\mathcal{M}_0$ generates $\mathcal{M}_j$. Combining this claim for $j = m$ with $\mathfrak{m}m_i \in \mathcal{M}_0$ checked above, we conclude that $\mathfrak{m}\mathcal{M}_m \subset \mathcal{M}_0$.

Part (ii) uses the same arguments and the Nakayama lemma. We lift $\bar{m}_i$ to arbitrary elements in $\mathcal{M}_{\mathrm{top}}^{(g_i)}$. The corresponding $\mathcal{R}$–modules does not depend on such choices. Note that the condition $\overline{\mathcal{M}}_m = \overline{\mathcal{M}}_{\mathrm{top}}$ provides that all $\mathcal{M}_i$ are standard ($\Delta_i \ni 0$). We need only to check that these modules have the required D–sets, which results from the following lemma by induction with respect to the length m of $\mathscr{M}$.

Lemma 2.4 *For two modules $\mathcal{M}' \subset \mathcal{M}$, where $\mathcal{M}'$ is not necessarily under the standard normalization, let $g \in D[\mathcal{M}]$, $\mathfrak{m}\mathcal{M} \subset \mathcal{M}'$, $\dim(\overline{\mathcal{M}}' + \overline{\mathcal{M}}^{(g)})/(\overline{\mathcal{M}}' + \overline{\mathcal{M}}^{(g+1)}) = 1$ and $\overline{\mathcal{M}} = \overline{\mathcal{M}}' \oplus \mathbb{C}\bar{m}$ for a pullback $\bar{m}$ of the generator of the latter quotient. Then we claim that $\mathcal{M}'$ is a standard module and $D[\mathcal{M}'] = D[\mathcal{M}] \setminus \{g\}$.*

Proof The module $\mathcal{M}$ is linearly generated by $\mathcal{M}'$ and any lift of $\bar{m}$ to $\mathcal{M}$ (the Nakayama lemma). Since $\dim(\overline{\mathcal{M}}' + \overline{\mathcal{M}}^{(g)})/(\overline{\mathcal{M}}' + \overline{\mathcal{M}}^{(g+1)})$ is 1, the gap g cannot be among the valuations of $\mathcal{M}'$ and this is the only missing gap there versus $\mathcal{M}$, which gives the required. □

Part (iii) follows from (ii). The inequalities $\dim \overline{\mathcal{M}}_{\mathrm{top}}^{(g_i)}/\overline{\mathcal{M}}_{\mathrm{top}}^{(g_i+1)} > 0$ are obviously necessary and sufficient to ensure the existence of the required $\overline{\mathcal{M}}_0$. The rest is a combination of (i) and (ii). □

Thus given a D–flag $\mathcal{D}$ as in (i), and $\mathcal{M}_{\mathrm{top}}$ such that

$$D[\mathcal{M}_{\mathrm{top}}] = D_m = D_0 \cup \{g_1, \dots, g_m\},$$

the inequalities $\dim \overline{\mathcal{M}}_{\mathrm{top}}^{(g_i)}/\overline{\mathcal{M}}_{\mathrm{top}}^{(g_i+1)} = 1$ are equivalent to the admissibility of $\mathcal{D}$. We note that in all examples we calculated, the admissibility of a D–flag is equivalent to that of every D_i. We cannot justify this in general.

This proposition is of independent interest and can be used to count the dimensions of $J_{\mathcal{R}}[\mathcal{D}]$ and to check (in some cases) that the latter are biregular to affine spaces.

We note that there is a natural (biregular) action of the algebraic group $J_{\mathcal{R}}[\varnothing] = \{\mathcal{L}\}$ in $J_{\mathcal{R}}$. Without going into details, Theorem 1 from [38] can be extended to flagged Jacobian factors as follows. If the GIT quotient $J_{\mathcal{R}}^m[\mathcal{D}]/J_{\mathcal{R}}[\varnothing]$ is biregular to an affine space and with the stabilizers of points in $J_{\mathcal{R}}^m[\mathcal{D}]$ of the same dimension, then the latter space is biregular to some $\mathbb{A}^N$. This is not always the case (see the online supplement), but it is not impossible that $J_{\mathcal{R}}^m[d]$ can be always paved by affine spaces.

2.4 The main conjecture

The flagged Jacobian factor $J_{\mathcal{R}}^m$ has a natural structure of a quasiprojective variety. Accordingly, $J_{\mathcal{R}}^m[\mathcal{D}]$ and $\overline{J}_{\mathcal{R}}^m[d]$ are its subvarieties; the latter is closed. Proposition 2.3 gives that $J_{\mathcal{R}}^m$ is a certain *subvariety* in a proper *parahoric Springer fibers* in the terminology from Section 2.2.9 from [44]. They are formed by *partial* periodic flags of $\mathcal{R}$–invariant lattices. We do not use them in the present work. By H_i, we will mean singular (relative) homology with the $\mathbb{C}$–coefficients.

Conjecture 2.5 *Let $\mathcal{R}$ be the ring of a unibranch plane curve singularity $\mathcal{C}_{\vec{r},\vec{s}}$ from Section 1.3, and $\mathcal{H}_{\vec{r},\vec{s}}(\square;q,t,a)$ the* DAHA *uncolored superpolynomial from* (1-19). *Recall that $\delta=|G_{\mathcal{R}}|$ is the arithmetic genus of $\mathcal{R}$; we assume that $\mathsf{r}_1>\mathsf{s}_1$.*

(i) *We conjecture that the relative homology $H_{2i+1}(\bar{J}^m_{\mathcal{R}}[d],\bar{J}^m_{\mathcal{R}}[d+1])$ vanishes for all $i,d\geq 0$, and*

$$\mathcal{H}_{\vec{r},\vec{s}}(\square;q,t,a)=\sum_{d,i,m}\mathrm{rk}\big(H_{2i}(\bar{J}^m_{\mathcal{R}}[d],\bar{J}^m_{\mathcal{R}}[d+1])\big)q^{d+m}t^{\delta-i}a^m, \tag{2-5}$$

where $0\leq d,i\leq\delta$, and the range of m is from 0 to $\mathsf{s}_1\mathsf{r}_2\cdots\mathsf{r}_\ell-1$, which is the number of (admissible) D such that $\dim(J_{\mathcal{R}}[D])=\delta-1$. The right-hand side of (2-5) *will be called $\mathcal{H}^{\mathrm{hom}}(q,t,a)$. Also, we conjecture that $J^m_{\mathcal{R}}[d]$ are paved by affine spaces.*

(ii) *If all varieties $J^m_{\mathcal{R}}[\mathcal{D}]$ are affine spaces $\mathbb{A}^i$, then their total number for any fixed i,d gives the corresponding rank $\mathrm{rk}(H_{2i})$ from* (2-5), *and this relation reads*

$$\mathcal{H}_{\vec{r},\vec{s}}(\square;q,t,a)=\sum_{\mathcal{D}}q^{d+m}t^{\delta-\dim(J^m_{\mathcal{R}}[\mathcal{D}])}a^m, \tag{2-6}$$

where the summation is over all admissible D–flags $\mathcal{D}$, $d=|D_0|$. Here the cells $J^m_{\mathcal{R}}[\mathcal{D}]$ contribute to $H_{2i}(\bar{J}^m_{\mathcal{R}}[d])\subset H_{2i}(J^m_{\mathcal{R}})$ for $i=\dim(J^m_{\mathcal{R}}[\mathcal{D}])$, and then to $H_{2i}(\bar{J}^m_{\mathcal{R}}[d],\bar{J}^m_{\mathcal{R}}[d+1])$. Formula (2-6) *can be readily extended to any affine cell decompositions of $J^m_{\mathcal{R}}[d]$ (not only those via $J^m_{\mathcal{R}}[\mathcal{D}]$).*

(iii) *Let $1/t=p^\ell$ for prime p and $\ell\in\mathbb{N}$, let $\mathbb{F}=\mathbb{F}_{1/t}$ be the field with p^ℓ elements, and let $|X(\mathbb{F})|$ be the number of $\mathbb{F}$–points of a scheme X defined over $\mathbb{F}$. One can assume that $\mathcal{R}$ is defined over $\mathbb{Z}$ and consider $J^m_{\mathcal{R}}[d]$ as schemes over $\mathbb{F}$. We conjecture that apart from finitely many p,*

$$\mathcal{H}_{\vec{r},\vec{s}}(\square;q,t,a)=t^\delta\sum_{d,m}|J^m_{\mathcal{R}}[d](\mathbb{F})|q^{d+m}a^m:=\mathcal{H}^{\mathrm{mod}}(q,t,a). \tag{2-7}$$

If $J^m_{\mathcal{R}}[d]$ are (a) *paved by affine spaces over $\mathbb{F}$, and* (b) *(non)admissible $\mathcal{D}$ remain such over this field, then* (2-7) *is equivalent to* (2-5). □

Relative homology Vanishing $H_{2i+1}(\bar{J}^m_{\mathcal{R}}[d],\bar{J}^m_{\mathcal{R}}[d+1])$ generalizes the *van Straten–Warmt conjecture*, which claims that odd Betti numbers of $J_{\mathcal{R}}$ vanish. This assumption and the exact sequences for relative (singular) homology imply that the natural maps $H_{2i+1}(\bar{J}^m_{\mathcal{R}}[d+1])\to H_{2i+1}(\bar{J}^m_{\mathcal{R}}[d])$ are surjective, which readily gives that $H_{2i+1}(\bar{J}^m_{\mathcal{R}}[d])=\{0\}$ for any i,d. Using this, the following sequence is exact:

$$0\to H_{2i}(\bar{J}^m_{\mathcal{R}}[d+1])\to H_{2i}(\bar{J}^m_{\mathcal{R}}[d])\to H_{2i}(\bar{J}^m_{\mathcal{R}}[d],\bar{J}^m_{\mathcal{R}}[d+1])\to 0. \tag{2-8}$$

As an application, (2-5) upon the substitution $q=1$ becomes

$$\mathcal{H}_{\vec{r},\vec{s}}(\square;q=1,t,a)=\sum_{i,m}\operatorname{rk} H_{2i}(J_{\mathcal{R}}^{m})t^{\delta-i}a^{m},$$

where the special case $a=0$ is [8, Conjecture 2.4(iii)]. Note that $H_{2i}(\bar{J}_{\mathcal{R}}^{m}[d],\bar{J}_{\mathcal{R}}^{m}[d+1])$ are potentially connected with compactly supported cohomology $H_c^{2k}(J_{\mathcal{R}}^{m}[d])$ for $k=j$ or $k+j=\dim J_{\mathcal{R}}^{m}[d]$, though Poincaré duality fails for singular varieties. Compare with (5-2).

Motivic approach Relation (2-7) readily follows from part (ii). We think that it can generally hold. In examples, it suffices to take any $p^{\ell}=1/t$ such that Γ remains the same over $\mathbb{F}_{1/t}$, but we do not know how far this goes. Say that (2-7) holds for $\mathcal{R}=\mathbb{F}_2[\![z^4+z^5,z^6]\!]$ under $a=0$, but not for $\mathbb{F}_3$, which changes Γ. However, $\mathcal{R}=\mathbb{F}_3[\![z^4,z^6+z^7]\!]$ is fine, and we note that $\nu(\mathbb{F}_2[\![z^4+z^5,z^6]\!])=\nu(\mathbb{F}_3[\![z^4,z^6+z^7]\!])$. These rings over $\mathbb{C}$ correspond to coinciding DAHA-superpolynomials. See (5-2) for a reformulation of (2-7) in terms of the weight filtration.

When $a=0$ and $q=1$, part (iii) is closely related to Theorems 0.1 and 0.2 from [15] in the case of A_n for affine Springer fibers. The latter are not always paved by affine spaces for other types [25]. The A_n–case is exceptional; the positivity of the orbital integrals and vanishing odd rational homology are widely expected to hold for anisotropic centralizers (ie in the nil-elliptic case). This matches well the conjectured positivity of uncolored DAHA superpolynomials.

Knowing the spectral curve is sufficient here. However, it is far from obvious beyond the torus case (quasihomogeneous plane curve singularities) that only the *topological* type of the singularity matters here. This follows from of our conjecture. The analytic equivalence of plane curve singularities is generally very different from the topological perspective. See [29; 4] and Section 5.3 for the identification of affine Springer fibers of anisotropic type with the compactified Jacobian of rational curves such that $\mathcal{R}$ is the local ring at its *unique* singular point.

In full generality, with the nonzero parameters q and a in (2-7), we count points in $J_{\mathcal{R}}^{m}[d](\mathbb{F})$ with q–weights $q^{|D_m|}$, where $|D|=\delta-\deg_{\mathcal{O}}\mathcal{M}$ for the corresponding *standard* $\mathcal{M}$. This seems a new turn in the theory of affine Springer fibers and related p–adic *orbital integrals*. Possible (expected) adding colors and the multibranch generalization of our construction make this even more interesting. Generally, we think that using the powerful modern theory of *topological* invariants of plane curve

singularities can be expected to impact the theory of orbital integrals (at least in type A) and related part of the fundamental lemma; see Sections 5.3 and 5.4 for further discussion.

The case $\mathcal{R} = \langle z^4, z^6 + z^7 \rangle$ We will stick to *admissible pairs* $\{\mathcal{M} = \mathcal{M}_0 \subset \mathcal{M}_1\}$, where $D_1 = D_0 \cup \{g_1\}$ (corresponding to the coefficient of a^1 in the main conjecture). Recall that $D_i = \Delta_i \cap G$ and $G = \mathbb{Z}_+ \setminus \Gamma$. Using the definition of $J_{\mathcal{R}}[D_0 \subset D_1] = J^1_{\mathcal{R}}[D_0, g_1]$ directly, we obtain the following lemma.

Lemma 2.6 *Provided the admissibility of the Γ–modules D_0 and $D_1 = D_0 \cup \{g_1\}$ for $g_1 \notin \Gamma \cup D_0$, one has*

$$\dim J_{\mathcal{R}}[D_0 \subset D_1] \le \dim J_{\mathcal{R}}[D_1] + \big|\{g \in \Delta_0 \mid g < g_1\}\big|, \tag{2-9}$$

$$\dim J_{\mathcal{R}}[D_0 \subset D_1] = \dim J_{\mathcal{R}}[D_1] + 1 \quad \text{if } g_1 < g \text{ for all } g \in D_0,$$

$$\dim J_{\mathcal{R}}[D_0 \subset D_1] = \dim J_{\mathcal{R}}[D_0] \quad \text{if } \{g \in G \mid g > g_1\} \subset D_0. \tag{2-10}$$

In the case of $\mathcal{R} = \langle z^4, z^6 + z^7 \rangle$, the pairs $\{D_0, g_1\}$ that are not included in the latter two formulas are all governed by the first one where "$\le$" is replaced by "$=$" there. Thus these formulas provide all dimensions of the varieties of admissible pairs $D_0 \subset D_1$. Using this for such $\mathcal{R}$, one obtains the formula

$$\begin{aligned}\sum_{D_0 \subset D_1} q^{|D_0|+1} t^{\delta - \dim(J^m_{\mathcal{R}}[D_0 \subset D_1])} &= q + q^2(1+t) + q^3(1+2t+t^2) + q^4(3t+2t^2+t^3) \\ &\quad + q^5(t+4t^2+2t^3+t^4) + q^6(t^2+4t^3+2t^4+t^5) \\ &\quad + q^7(t^3+3t^4+2t^5+t^6) + q^8(t^5+t^6+t^7),\end{aligned}$$

which matches the coefficient of a^1 of the uncolored DAHA *superpolynomial for $\mathcal{R} = \langle z^4, z^6 + z^7 \rangle$ from* [8].

Proof To justify the first inequality, we begin with any parametric family of modules $\mathcal{M}_0$ corresponding to D_0, assuming that they can be extended to $\mathcal{M}_{g_1}$ by adding certain m_{g_1} with $\upsilon(m_{g_1}) = g_1$ and that they are different (as submodules of $\mathcal{O}$) for different values of parameters. The dimension of the resulting family of modules $\mathcal{M}_1$ is no greater than $\dim J_{\mathcal{R}}[D_1]$. Given $\mathcal{M}_1 \ni m_{g_1}$, its different submodules $\mathcal{M}_0$ can be only obtained by adding the terms from $\mathbb{C}m_{g_1}$ to the generators $m_g \in \mathcal{M}_0$ with $\upsilon(m_g) = g$, where $g \in D_0$ such that $g < g_1$. This gives the required inequality.

If $g_1 < D_0$, then only $m_0 = 1 + z(\cdot)$ can be altered by $\mathbb{C}m_{g_1}$ and there is a one-dimensional family of (pairwise distinct) such submodules $\mathcal{M}_0$ inside fixed $\mathcal{M}_1$, which gives the second formula in (2-9). Here we use that the relations for the coefficients of the generators $m \in \mathcal{M}$ necessary for the equality $D_{\mathcal{M}} = D$ allow such a deformation.

Generally, if there several such m_g (not just $g = 0$) with g before g_1, then it is not true that the increase of dimension for $D[\mathcal{M}_0]$ will coincide with the number of such g. It really occurs with five such admissible pairs $\{D_0, D_1\}$ for $\mathcal{R} = \langle z^4, z^6 + z^{13}\rangle$:

(2-11)

D_0	D_1	$\dim_0$	$\dim_1$
$[2, 9, 15]$	$[2, 9, 11, 15]$	7	6
$[2, 7, 11, 15]$	$[2, 7, 9, 11, 15]$	6	5
$[2, 9, 11, 15]$	$[2, 7, 9, 11, 15]$	6	5
$[2, 3, 7, 9, 11, 15]$	$[2, 3, 5, 7, 9, 11, 15]$	6	5
$[1, 2, 5, 7, 9, 11, 15]$	$[1, 2, 3, 5, 7, 9, 11, 15]$	3	0

where $\dim_i = \dim(J_{\mathcal{R}}[D_i])$. However, in all these cases, $\dim(J_{\mathcal{R}}[D_0 \subset D_1])$ coincides with $\dim_0$, which is (2-9) with strict equality there.

Finally, formula (2-10) holds because adding *any* element $m_{g_1} \in \mathcal{O}$ with such a valuation g_1 to $\mathcal{M}_0$ results in the required $\mathcal{M}_1$ in this case. *All* cases from (2-11) are of the latter type. Thus any admissible pair for $\mathcal{R} = \langle z^4, z^6 + z^7\rangle$ satisfies either (2-9) with the equality there or (2-10). Then one can use the formulas for $\dim J_{\mathcal{R}}^{m=0}[D]$ from [38]. □

3 The family $\mathbb{C}[\![z^4, z^{2u} + z^v]\!]$

3.1 Dimensions of $J_{\mathcal{R}}^m[\mathcal{D}]$

It is shown in [38] that the varieties $J_{\mathcal{R}}[D]$ are isomorphic to affine spaces for the family $\mathcal{R} = \mathbb{C}[\![z^4, z^{2u} + z^v]\!]$, where $(uv, 2) = 1$, $v > 2u$ and their dimensions are computed. We will generalize these claims to admissible D–flags and the corresponding varieties $J_{\mathcal{R}}^m[\mathcal{D}]$.

To write our formula for the dimensions we will need a few definitions. Let

$$\mu_{D,g} := \dim(J_{\mathcal{R}}^1[D, g]) - \dim(J_{\mathcal{R}}[D])$$

be the *dimension change*. Also, we write

$$\gamma_\Delta(\ell) := |[\ell, \infty) \setminus \Delta|$$

for the *gap counting function*. Recall that $D = \Delta \setminus \Gamma$ is called a D–set corresponding to a *standard* Γ–module Δ (ie $0 \in \Delta$).

Theorem 3.1 *Let $\mathcal{R} = \mathbb{C}[\![z^4, z^{2u} + z^v]\!]$, where $(uv, 2) = 1$, $v > 2u$, and $\mathcal{D}$ be an admissible D–flag (with D_i corresponding to admissible Δ_i). Then $J^m_{\mathcal{R}}[\mathcal{D}]$ (with the notation as above) is isomorphic to an affine space $\mathbb{A}^N$ with*

$$N = \dim(J_{\mathcal{R}}[D_0]) + \sum_{i=0}^{m-1} \mu_{D_i, g_{i+1}},$$

where

$$\mu_{D,g} = \begin{cases} \gamma_{\Delta\cup\{g\}}(g) - \gamma_{\Delta\cup\{g\}}(g+4) & \text{for } g \equiv 1 \text{ or } 3 \bmod 4, \\ \gamma_{\Delta\cup\{g\}}(g) - \gamma_{\Delta\cup\{g\}}(g+4) & \\ \qquad - (\gamma_{\Delta\cup\{g\}}(g+n) - \gamma_{\Delta\cup\{g\}}(g+4+n)) & \text{for } g \equiv 2 \bmod 4, \end{cases}$$

and $n = n(D)$ is the smallest odd number in $(\Delta \cup v) \cap [2u, \infty)$. Here $D = D_i$, $g = g_{i+1}$ or D is any admissible D–set such that $D \cup \{g\}$ is admissible.

3.2 Basic definitions

Before we proceed with the proof, we will briefly summarize and adjust the approach taken in [38] to prove that the $J_{\mathcal{R}}[D]$ are affine; our argument relies heavily on his method. Let Δ be a Γ–module and begin by choosing a_0, a_1, a_2 and a_3 such that $a_i = \min\{k \in \Delta \mid k \equiv i \bmod 4\}$. Consider the following elements in $\mathcal{O}$:

$$\begin{aligned} m_0 &= 1 + \sum_{k \in \mathbb{N}\setminus\Delta} \lambda^0_k z^k, & m_1 &= z^{a_1} + \sum_{k \in [a_1,\infty)\setminus\Delta} \lambda^1_{k-a_1} z^k, \\ m_2 &= z^{a_2} + \sum_{k \in [a_2,\infty)\setminus\Delta} \lambda^2_{k-a_2} z^k, & m_3 &= z^{a_3} + \sum_{k \in [a_3,\infty)\setminus\Delta} \lambda^3_{k-a_3} z^k, \end{aligned} \tag{3-1}$$

where the λ–coefficients are treated as variables. The valuation Γ–module of the module $\mathcal{M}$ generated by $\{m_i\}$ will then contain Δ, since any element of Δ has the form $a_i + 4n$ for some $n \in \mathbb{N}$ and because $z^4 \in \mathcal{R}$. Thus $\{a_i\}$ form a *basis* for $\upsilon(\mathcal{M})$ in a natural sense.

An important component of Piontkowski's method is the observation that $\upsilon(\mathcal{M}) = \Delta$ if and only if the relations among the elements m_i do not produce elements with valuation not in Δ. Thus the *syzygies* of the set $\{m_i\}$ as well as the syzygies of the set of their leading terms are of importance. Lemma 11 of [38] uses this basic idea to give an equivalent condition for $\upsilon(\mathcal{M}) = \Delta$; it will be provided. We need the notion of *initial vector*, which is also from [38], to state the aforementioned lemma:

Definition 3.2 For $\vec{r} = (r_0, \dots, r_3) \in \mathcal{R}^4$, let $\sigma = \min\{\upsilon(r_j) + a_j\}$. The *initial vector* $\mathrm{in}(\vec{r})$ is as follows: $\mathrm{in}(\vec{r}) = (\zeta_j)$ with ζ_j equal to the monomial of lowest degree in r_j if $\upsilon(r_j) + a_j = \sigma$ and 0 otherwise.

Lemma 3.3 *Let $\mathcal{M}$ be an $\mathcal{R}$–module generated by $\{m_0, m_1, m_2, m_3\}$, and V an $\mathcal{R}$–submodule in $\bigoplus_{j=0}^{3}\mathcal{R}$ such that the initial vectors $\{\mathrm{in}(\vec{r}) \mid \vec{r} \in V\}$ of V linearly generate the* syzygies *of the set (z^{a_j}) of $\mathbb{C}[\Delta]$. Here $\mathbb{C}[\Delta]$ is the vector space generated by the elements $\{z^k\}$ for $k \in \Delta$. We use $\sigma = \min\{\upsilon(r_j m_j)\} = \min\{\upsilon(r_j) + a_j\}$ from Definition 3.2.*

Then $\upsilon(\mathcal{M}) = \Delta$ if and only if, for each $\vec{r} = (r_j) \in V$, the initial terms in $\sum_{j=0}^{3} r_j m_j$ cancel, ie $\upsilon\big(\sum_{j=0}^{3} r_j m_j\big) > \sigma$, and for every j, there exists $p_j \in \mathcal{R}$ such that $\upsilon(p_j m_j) > \sigma$ and $\sum_{j=0}^{3} r_j m_j = \sum_{j=0}^{3} p_j m_j$. If such p_j exist, then the element $\sum_{j=0}^{3} p_j m_j \in \mathcal{M}$ is called a ***higher-order expression*** *for $\sum_{j=0}^{3} r_j m_j \in \mathcal{M}$, which is generally not uniquely determined by the latter element.* □

We will need the following *reduction procedure* from [38] for series $y \in \mathcal{O}$. Let $y_0 = y$. Then define inductively $y_{i+1} = y_i$ if $i \notin \Delta$, and we also set $s_i = 0$ in this case. If $i \in \Delta$, then find the monomial $c_i z^i$ with power i in y_i and the element $s_i \in \mathcal{R}$ such that $s_i m_{j_i} = c_i z^i + \cdots$ for one of the generators m_{j_i} (it can be nonunique). Then we let $y_{i+1} = y_i - s_i m_{j_i}$. The sequence of elements $y_i \in \mathcal{O}$ converges to an element y_∞, which has the form $\sum_{k \in \mathbb{N}\setminus\Delta} d_k z^k$ for some coefficients d_k.

The key facts from [38] about the reduction procedure are that (a) $y_\infty = 0$ if and only if $y \in \mathcal{M}$, and (b) if $y_\infty = 0$, then $\sum s_i m_{j_i}$ is a higher-order expression for y.

Definition 3.4 For any element $y \in \mathcal{O}$, the element $y^\dagger$ will be the result of the reduction procedure applied to y. One has $(y^\dagger)^\dagger = y^\dagger$.

The reduction procedure depends on the choices above. We can standardize the procedure by always taking the elements s_i involved to be of the form $(z^4)^k$ for some nonnegative integers k, which then makes the reduction procedure unique. To see that picking such s_i is possible, observe that when eliminating $c_i z^i$ for $i \in \Delta$, we can choose a unique generator m_j such that $a_j \equiv i \bmod 4$. We will call such a procedure *standard*, and we will always assume such a standardization in what will follow.

One has $x^\dagger = x + f$ for some $f \in \mathcal{M}$. If for two elements $x, y \in \mathcal{O}$, we know that $x^\dagger = x + f_1$ and $y^\dagger = y + f_2$, then the standard reduced form $(x+y)^\dagger$ of $x + y$ is $x + y + f_1 + f_2$. Generally, if $(x+y)^\dagger = x + y + f$, the uniqueness guarantees $f = f_1 + f_2$, and we obtain that $x^\dagger + y^\dagger = (x+y)^\dagger$ for the standardization we will always impose. Thus $\dagger$ is a $\mathbb{C}$–linear projection.

Piontkowski proves that $J_{\mathcal{R}}[D]$ are affine by using the previous lemma and the reduction procedure. In Lemma 3.3, it is sufficient to consider the degrees of syzygies of (z^{a_j}) less than $\max\{a_j \mid j = 0, \dots, 3\} - 3$, and there are only finitely many of such syzygies. Thus it is sufficient to check that only finitely many linear combinations of the $\{m_i\}$ prescribed by the syzygies have higher-order expressions to ensure that $\Delta = \upsilon(\mathcal{M})$. To obtain the higher-order expressions of these elements the reduction procedure is used and the resulting elements of $\mathcal{O}$ must vanish since these elements are in $\mathcal{M}$. The coefficients of the higher-order expressions are polynomials in terms of the coefficients of $\{m_i\}$, which have the form

$$\lambda_k^j - \lambda_k^i + \text{polynomial in } \lambda_\ell^*, \quad \ell < k.$$

These expressions must vanish due to the properties of the reduction procedure. Since they vanish and are linear in the parameters λ_k^j, λ_k^i, we can express λ_k^j in terms of λ_ℓ^* for $\ell < k$ and λ_k^i.

Applying the same process to λ_ℓ^* such that $\ell < k$, we can eventually show that $J_{\mathcal{R}}[D]$ is a graph of a regular function on an affine space $\mathbb{A}^n$. Since the graph of a regular function is always isomorphic to the domain of the function, we have that $J_{\mathcal{R}}[D] \cong \mathbb{A}^n$. We will use a similar technique to prove Theorem 3.1.

3.3 The cells $J_{\mathcal{R}}^m[\mathcal{D}]$ are affine

Let $\mathcal{D}$ be an admissible D–flag with $\vec{g} = \{g_i \mid 1 \le i \le m\}$. Define $\mathcal{E}$ to be the admissible D–flag of length $m-1$ such that $E_0 = D_0$ and $E_j = D_0 \cup \{g_i \mid 1 \le i \le j\}$ where $1 \le j \le m-1$; ie $\mathcal{E}$ is a truncation of $\mathcal{D}$.

Let $\mathscr{M}$ be an element of $J_{\mathcal{R}}^{m-1}[\mathcal{E}]$, and define

$$h_{g_m} = z^{g_m} + \sum_{k \in [g_m, \infty) \setminus \Delta_m} \lambda_{k-g_m}^{h_{g_m}} z^k.$$

By adjoining the module $\mathcal{M}_m = \mathcal{M}_{m-1} \oplus \mathcal{R} h_{g_m}$ to the end of $\mathscr{M}$, we can extend $\mathscr{M}$ to an element of $J_{\mathcal{R}}^m[\mathcal{D}]$ (of course there are restrictions on the λ coefficients which are addressed below). The other way around, every element of $J_{\mathcal{R}}^m[\mathcal{D}]$ can be obtained by this procedure from flags of length $m-1$.

To ensure that $\upsilon(\mathcal{M}_m) = \Delta_m$ and that $\mathcal{M}_{m-1} \subset \mathcal{M}_m$, we need to use Lemma 3.3 above and Lemma 3.5 below. The hypotheses for these lemmas only concern $\mathcal{M}_{m-1}$ and g_m. For this reason, we can reduce the argument to the case of $m = 1$.

From now on we set $D = D_0$, so it corresponds to an admissible Γ–module $\Delta = \Delta_0$. Accordingly, set $g = g_1$ and $D_1 = D \cup \{g\}$, where D_1 will be assumed admissible. Also, we let $\Delta' = \Delta \cup \{g\}$.

Recall that we let the module generated by the $\{m_i\}$ as in (3-1) be denoted by $\mathcal{M}$. Then we consider adding

$$h := z^g + \sum_{k \in [g,\infty)\setminus\Delta} \lambda^h_{k-g} z^k$$

to the set of generators $\{m_i\}$ and set $\mathcal{M}' = \langle \mathcal{M}, h \rangle$. Note that h will replace m_ℓ where $g \equiv \ell \bmod 4$.

The modules $\mathcal{M}$ and $\mathcal{M}'$ must satisfy $\upsilon(\mathcal{M}) = \Delta$ and $\upsilon(\mathcal{M}') = \Delta'$. By Lemma 3.3, this holds for $\mathcal{M}$ if and only if the following elements have higher-order expressions:

$$\begin{aligned} T^1 &:= (z^{2u} + z^v)m_0 - z^{4\alpha_2} m_2, \\ T^2 &:= z^{4(u-\alpha_2)} m_0 - (z^{2u} + z^v) m_2, \\ T^3 &:= \left(z^{2u+v} + \tfrac{1}{2} z^{2v}\right) m_0 - z^{4\alpha_1} m_j, \end{aligned} \tag{3-2}$$

where α_i is the unique integer such that $a_i = \beta_i(2u+v) + \gamma_i(2u) - 4\alpha_i$ for $\beta_i, \gamma_i \in \{0, 1\}$ and $j \equiv 2u + v \bmod 4$. To obtain the higher-order expressions, the reduction procedure is applied to T^i. As a result of the reduction procedure polynomial relations among the λ variables are obtained (as discussed before the beginning of the proof). For $\mathcal{M}'$, we use a similar approach. When we consider the syzygies (3-2) in the context of $\mathcal{M}'$, we will denote the resulting T by $(T^i)'$. See pages 14–17 of [38] concerning the existence of the higher-order expressions.

Given a module $\mathcal{M}$ with $\upsilon(\mathcal{M}) = \Delta$, not every module $\mathcal{N}$ with $\upsilon(\mathcal{N}) = \Delta'$ is its extension. In order to understand when $\mathcal{M} \subset \mathcal{N}$, let

$$F := m_i - (z^4)h \in \mathcal{N},$$

where $i \equiv g \bmod 4$. Note that $\upsilon(F) > a_i$, which gives a new type of syzygy (not from [38]); there is one such syzygy for each pair $\mathcal{M} \subset \mathcal{N}$.

Let us take i such that $g \equiv i \bmod 4$. Then replacing a_i with g results in a basis for Δ'; let us check this. If there were ℓ such that $\ell \equiv a_i \bmod 4$, $g < \ell < a_i$ and $\ell \notin \Delta$, then the presence of g in Δ' implies that $\ell \in \Delta'$ since Δ' is a Γ–module. This contradicts $\Delta' = \Delta \cup \{g\}$. Thus $g + 4 = a_i$, as required.

The following lemma provides necessary and sufficient conditions for $\mathcal{M} \subset \mathcal{N}$ when $\upsilon(\mathcal{N}) = \Delta'$ and $\upsilon(\mathcal{M}) = \Delta$. It is important to distinguish performing the reduction procedure with respect to $\mathcal{M}$ or $\mathcal{N}$ and we will use $\dagger_1$ to denote reduction with respect to the generators (3-1) of $\mathcal{M}$ and $\dagger_2$ for reduction with respect to (3-1) except with the changes necessary for $\mathcal{N}$.

Lemma 3.5 *Suppose $\mathcal{M}$ and $\mathcal{N}$ are standard $\mathcal{R}$–modules in $\mathcal{O}$ such that $\upsilon(\mathcal{M}) = \Delta$ and $\upsilon(\mathcal{N}) = \Delta \cup \{g\}$. Furthermore, suppose that $\mathcal{N}$ contains the generators m_j of $\mathcal{M}$ from* (3-1) *satisfying $g \not\equiv j \mod 4$. Then $\mathcal{M} \subset \mathcal{N}$ if and only if $F^{\dagger_2} = 0$.*

Proof Since $g \in \upsilon(\mathcal{N})$, we see that $h \in \mathcal{N}$ for some choice of values for the variables λ_k^h and that h is a normalized generator of $\mathcal{N}$. For some i, we have that $g \equiv a_i \mod 4$ (ie g replaces a_i in the basis for $\upsilon(\mathcal{M})$). Observe that $M \subset \mathcal{N}$ if and only if $m_i \in \mathcal{N}$, which will happen if and only if $(m_i)^{\dagger_2} = 0$. The first step in the reduction of m_i is F, and so $(m_i)^{\dagger_2} = 0$ if and only if $F^{\dagger_2} = 0$. □

Let us now return to considering $\mathcal{M} \subset \mathcal{M}'$ with $\upsilon(\mathcal{M}) = \Delta$ and $\upsilon(\mathcal{M}') = \Delta'$. We only need to consider the equations resulting from $F^{\dagger_2}$ since the equations resulting from $\upsilon(\mathcal{M}) = \Delta$ and $\upsilon(\mathcal{M}') = \Delta'$ are already solved for in [38]. We may write $F^{\dagger_2} = \sum_{k=1}^{\infty} \widetilde{c}_k z^{a_i+k}$. Recall that the only powers of z present in $F^{\dagger_2}$ are those greater than a_i that are not in Δ. By Lemma 3.5, we have $F^{\dagger_2} = 0$, which implies $\widetilde{c}_k = 0$. Similar to the analysis of $(T^j)^{\dagger_1}$ from the discussion before the proof, the $\widetilde{c}_k$ are in the form

$$\widetilde{c}_k = \lambda_k^i - \lambda_k^h + (\text{a polynomial in terms of } \lambda_p^\bullet \text{ for } p < k).$$

For a given k, we can then express λ_k^h in terms of $\lambda_p^\bullet$ for $p < k$. This gives that $J^1_{\mathcal{R}}[D, g]$ is an affine space because it is a graph of a regular function on an affine space. Since the $\gamma_{\Delta'}(g+4)$ equations $\widetilde{c}_k = 0$ are solvable, we see that $\mu_{D,g} \leq \gamma_{\Delta'}(g) - \gamma_{\Delta'}(g+4)$. The exact value of $\mu_{D,g}$ depends on the congruence class of g modulo 4. We will now obtain the formulas for the dimensions.

3.4 Calculating dimensions

We are going now to justify the dimension formulas in Theorem 3.1. Recall that our approach extends the formulas and techniques used in [38] to the case of flags of modules.

First assume g is odd. Since $2u+v$ is odd, it is either congruent to 1 or 3 mod 4. If $g \not\equiv 2u+v \bmod 4$, then $T^i = (T^i)'$ for all i, and hence we do not need to impose any further relations on the coefficients. Thus $\mu_{D,g} = \gamma_{\Delta'}(g) - \gamma_{\Delta'}(g+4)$ in this case.

When $g \equiv 2u+v \bmod 4$, we have $(T^3)' = z^{2u+v}g_0 - z^{4\alpha_h}h$ and $T^i = (T^i)'$ for $i \neq 3$. At the bottoms of pages 15 and 16 of [38], it is shown that a higher-order expression exists for T^3 when the smallest odd number $n \in \Delta \cap [2u, \infty)$ is less than or equal to v. When $n > v$ it is also shown that we can use the higher-order expressions of T^1 and T^2 to obtain a higher-order expression for T^3 (without imposing any new relations among the λ–parameters). Hence $\mu_{D,g} = \gamma_{\Delta'}(g) - \gamma_{\Delta'}(g+4)$. This finishes the proof when $g \equiv 1$ or $3 \bmod 4$.

In the last case, $g \equiv 2 \bmod 4$ implies that

$$\begin{aligned} (T^1)' &= (z^{2u} + z^v)m_0 - z^{4\alpha_h}h, \\ (T^2)' &= z^{4(u-\alpha_h)}m_0 - (z^{2u} + z^v)h, \end{aligned} \tag{3-3}$$

and $T^3 = (T^3)'$. Following [38], $\gamma_{\Delta'}(2u)$ equations for the λ–parameters result from the coefficients of $((T^1)')^{\dagger_2}$, and $\gamma_{\Delta'}(g+n)$ distinct equations result from the coefficients of $((T^2)')^{\dagger_2}$. We claim that the $\gamma_{\Delta'}(2u)$ equations from $((T^1)')^{\dagger_2}$ are equal to those from $(T^1)^{\dagger_1}$, and all of the $\gamma_{\Delta'}(a_2+n)$ equations from $(T^2)^{\dagger_1}$ can be obtained from the coefficients of $((T^2)')^{\dagger_2}$. To prove this we introduce the following definition and lemma.

Let P be any power series in z whose coefficients are polynomials in the λ variables. We let $\mathfrak{I}(P)$ be the ideal generated by the coefficients of P in the polynomial ring over the λ variables. We have the following basic result concerning $\mathfrak{I}$ and $\dagger_2$.

Lemma 3.6 *$\mathfrak{I}((rP)^{\dagger_2}) \subset \mathfrak{I}(P^{\dagger_2})$, where r is a polynomial in $\mathcal{R}$.*

Proof Since $\dagger_2$ is linear, it is sufficient to prove the lemma when r is a monomial. The reduction procedure for rP is exactly the same as for P until the first k such that $k \notin \Delta \ni k + v(r)$. Beyond this range, a multiple of the coefficient of z^k in $P^{\dagger_2}$ may be added to the remaining coefficients of $(rP)^{\dagger_2}$. Because of such k, all coefficients of $(rP)^{\dagger_2}$ will be coefficients of $P^{\dagger_2}$ plus some multiples of the previous coefficients of $P^{\dagger_2}$. Thus we have $\mathfrak{I}[(rP)^{\dagger_2}] \subset \mathfrak{I}[P^{\dagger_2}]$. □

On page 14 of [38], we see that the valuations of T^1, T^2, $(T^1)'$ and $(T^2)'$ are greater than $2u$ and therefore greater than g. Let us use this.

Lemma 3.7 *If $\mathcal{M} \subset \mathcal{M}'$, then $(T^1)^{\dagger_1} = (T^1)^{\dagger_2}$ and $(T^2)^{\dagger_1} = (T^2)^{\dagger_2}$.*

Proof Observe that $(T^1)^{\dagger_1} - (T^1)^{\dagger_2} = r_0 m_0 + r_1 m_1 + r_2 F + r_3 m_3$, where the $r_i \in R$. When we apply $\dagger_2$ to the left-hand side, we get

$$((T^1)^{\dagger_1} - (T^1)^{\dagger_2})^{\dagger_2} = (T^1)^{\dagger_1} - (T^1)^{\dagger_2}$$

since $(T^1)^{\dagger_1}$ has valuations greater than g, which implies the left-hand side is an eigenvector of $\dagger_2$. Applying $\dagger_2$ to the right-hand side, we have

$$(r_0 m_0 + r_1 m_1 + r_2 F + r_3 m_3)^{\dagger_2} = (r_2 F)^{\dagger_2}$$

since m_0, m_1, m_3 are in $\mathcal{M}'$. Hence we have

$$(T^1)^{\dagger_1} - (T^1)^{\dagger_2} = (r_2 F)^{\dagger_2}.$$

Note that r_2 may be a power series, but only a truncation of it determines $(r_2 F)^{\dagger_2}$ because terms with valuation greater than the conductor will all eventually be eliminated. Therefore, by Lemma 3.6, we have $\mathfrak{I}((r_2 F)^{\dagger_2}) \subset \mathfrak{I}(F^{\dagger_2})$. Now $\mathcal{M} \subset \mathcal{M}'$, which means $F^{\dagger_2} = 0$ by Lemma 3.5, and so we have $(r_2 F)^{\dagger_2} = 0$. The proof for T^2 is identical because T^2 has valuation greater than $2u$. □

Now we prove that $(T^1)^{\dagger_1} = ((T^1)')^{\dagger_2}$ when $\mathcal{M} \subset \mathcal{M}'$. Notice that $T^1 - (T^1)' = -z^{4\alpha_2} F$, where α_2 and α_h are defined in (3-2) and (3-3), respectively, and we have used that $\alpha_h = \alpha_2 + 1$. Thus

$$(T^1)^{\dagger_2} - ((T^1)')^{\dagger_2} = (-z^{4\alpha_2} F)^{\dagger_2}.$$

By Lemmas 3.5 and 3.6 we have $(-z^{4\alpha_2} F)^{\dagger_2} = 0$ so by the previous Lemma 3.7 we have $(T^1)^{\dagger_1} = ((T^1)')^{\dagger_2}$.

To prove the equations from $(T^2)^{\dagger_1}$ are redundant we first observe that $T^2 - z^4 (T^2)' = -(z^{2u} + z^v) F$. This implies that

$$(T^2)^{\dagger_2} - (z^4 (T^2)')^{\dagger_2} = (-(z^{2u} + z^v) F)^{\dagger_2}.$$

Again, by Lemmas 3.5 and 3.6, we see that $(-(z^{2u} + z^v) F)^{\dagger_2} = 0$. By Lemma 3.7, we have $(T^2)^{\dagger_2} = (T^2)^{\dagger_1}$, which means $(z^4 (T^2)')^{\dagger_2} = (T^2)^{\dagger_1}$. Finally, by Lemma 3.6, $\mathfrak{I}((z^4 (T^2)')^{\dagger_2}) \subset \mathfrak{I}(((T^2)')^{\dagger_2})$, which readily implies that the equations from $(T^2)^{\dagger_1}$ are redundant.

Thus we have shown that the $\gamma_{\Delta'}(2u)+\gamma_{\Delta'}(a_2+n)$ equations from $(T^1)^{\dagger_1}$ and $(T^2)^{\dagger_1}$ are actually redundant. So, as required,

$$\begin{aligned}\mu_{D,g}&=\gamma_{\Delta'}(g)-\gamma_{\Delta'}(a_2)-\gamma_{\Delta'}(2u)-\gamma_{\Delta'}(g+n)+\gamma_{\Delta'}(2u)+\gamma_{\Delta'}(a_2+n)\\&=\gamma_{\Delta'}(g)-\gamma_{\Delta'}(a_2)-(\gamma_{\Delta'}(g+n)-\gamma_{\Delta'}(a_2+n))\\&=\gamma_{\Delta\cup\{g\}}(g)-\gamma_{\Delta\cup\{g\}}(g+4)-(\gamma_{\Delta\cup\{g\}}(g+n)-\gamma_{\Delta\cup\{g\}}(g+4+n)).\end{aligned}$$

4 Numerical support

We provide here dimensions of cells for some basic examples and the corresponding nonadmissible D–flags $\mathcal{D}$. It is important to know whether these dimensions, nonadmissible $\mathcal{D}$ and admissible ones with nonaffine $J_r^m[\mathcal{D}]$ are *topological* properties of the singularity. We found no counterexamples, but formally this can be wrong even if our main conjecture holds. Zariski proved in Chapter IV, Section 3 of [45] that $\Gamma=\langle 4,6,13+2v\rangle$ for $v\geq 0$ uniquely determines the corresponding *analytic* singularity (ie that each equisingularity class is one point), but generally such questions can be (very) involved. Also, these examples seem absolutely necessary for (restarting) the theory of Jacobian factors beyond torus knots, which is of obvious importance for topology and geometry of plane curve singularities and for orbital integrals.

4.1 Two simplest cables

We begin with formula (2-6) for the knots $\mathrm{Cab}(13,2)T(3,2)$ and $\mathrm{Cab}(15,2)T(3,2)$. Here one can use the general Theorem 3.1 or the special Lemma 2.6. The latter can be extended to any m for these two cases (with minor adjustments).

Recall that $\mathcal{R}=\langle z^4,z^6+z^7\rangle$, $\Gamma=\langle 4,6,13\rangle$ and $\mathcal{R}=\langle z^4,z^6+z^9\rangle$, $\Gamma=\langle 4,6,15\rangle$ in these cases. The first DAHA superpolynomial is as follows:

$$\vec{\mathsf{r}}=\{3,2\},\quad \vec{\mathsf{s}}=\{2,1\},\quad \mathcal{T}=\mathrm{Cab}(13,2)T(3,2);$$

$$\begin{aligned}&\mathcal{H}_{\vec{\mathsf{r}},\vec{\mathsf{s}}}(\Box;q,t,a)=\\&1+qt+q^8t^8+q^2(t+t^2)+a^3(q^6+q^7t+q^8t^2)+q^3(t+t^2+t^3)\\&\quad+q^4(2t^2+t^3+t^4)+q^5(2t^3+t^4+t^5)+q^6(2t^4+t^5+t^6)+q^7(t^5+t^6+t^7)\\&\quad+a^2\big(q^3+q^4(1+t)+q^5(1+2t+t^2)+q^6(2t+2t^2+t^3)+q^7(2t^2+2t^3+t^4)\\&\qquad\qquad+q^8(t^3+t^4+t^5)\big)\\&\quad+a\big(q+q^2(1+t)+q^3(1+2t+t^2)+q^4(3t+2t^2+t^3)+q^5(t+4t^2+2t^3+t^4)\\&\qquad+q^6(t^2+4t^3+2t^4+t^5)+q^7(t^3+3t^4+2t^5+t^6)+q^8(t^5+t^6+t^7)\big).\end{aligned}$$

D–sets	dim	D–sets	dim
$\varnothing$	8	1, 3, 5, 7, 9, 11, 15	2
15	7	2, 7, 11, 15	6
11, 15	6	2, 9, 15	7
7, 11, 15	6	2, 9, 11, 15	6
9, 15	7	2, 7, 9, 11, 15	5
9, 11, 15	5	2, 3, 7, 9, 11, 15	4
7, 9, 11, 15	4	2, 5, 9, 11, 15	5
3, 7, 9, 11, 15	4	2, 5, 7, 9, 11, 15	3
5, 9, 11, 15	5	2, 3, 5, 7, 9, 11, 15	1
5, 7, 9, 11, 15	3	1, 2, 5, 7, 9, 11, 15	3
3, 5, 7, 9, 11, 15	2	1, 2, 3, 5, 7, 9, 11, 15	0
1, 5, 7, 9, 11, 15	4		

Table 1: Dimensions for $\Gamma = \langle 4, 6, 13\rangle$, $m = 0$

Here and henceforth, we use [8]. Let us list the necessary information to verify (2-6). For the greatest possible $m = 3$, there are only three admissible D–sets D_0 that can occur in such a (long) flag. Namely, these flags and the dimensions $\dim J^3_{\mathcal{R}}[\mathcal{D}]$ are

$$\begin{aligned}
&D_0 = [9, 11, 15], && \vec{g} = (2, 5, 7), && \dim = 8 && \rightsquigarrow q^6 t^0 a^3,\\
&D_0 = [7, 9, 11, 15], && \vec{g} = (2, 3, 5), && \dim = 7 && \rightsquigarrow q^7 t^1 a^3,\\
&D_0 = [5, 7, 9, 11, 15], && \vec{g} = (1, 2, 3), && \dim = 6 && \rightsquigarrow q^8 t^2 a^3;
\end{aligned} \tag{4-1}$$

we show their contributions to the corresponding superpolynomial.

Tables 1, 2, 3 show D_0, the corresponding $\vec{g}$ and the dimensions of $J^m_{\mathcal{R}}[\mathcal{D}]$ for all admissible flags as $m = 0, 1, 2$.

There are no admissible extensions of degree 4, so we have

$$\begin{aligned}
&\sum_{D} q^{|D|+m} t^{\delta-\dim(J^m_{\mathcal{R}}[D,D'])} a^m = \\
&1 + qt + q^2(t + t^2) + q^3(t + t^2 + t^3) + q^4(2t^2 + t^3 + t^4)\\
&\qquad + q^5(2t^3 + t^4 + t^5) + q^6(2t^4 + t^5 + t^6) + q^7(t^5 + t^6 + t^7) + q^8 t^8\\
&+ a^3\big(q^6 + q^7 t + q^8 t^2\big)\\
&+ a^2\big(q^3 + q^4(1 + t) + q^5(1 + 2t + t^2) + q^6(2t + 2t^2 + t^3)\\
&\qquad + q^7(2t^2 + 2t^3 + t^4) + q^8(t^3 + t^4 + t^5)\big)\\
&+ a\big(q + q^2(1 + t) + q^3(1 + 2t + t^2) + q^4(3t + 2t^2 + t^3) + q^5(t + 4t^2 + 2t^3 + t^4)\\
&\qquad + q^6(t^2 + 4t^3 + 2t^4 + t^5) + q^7(t^3 + 3t^4 + 2t^5 + t^6) + q^8(t^5 + t^6 + t^7)\big),
\end{aligned}$$

which coincides with $\mathcal{H}_{\vec{r},\vec{s}}(\square; q, t, a)$ from Section 3.1 of [8].

D–sets	g	dim	D–sets	g	dim
$\varnothing$	15	8	5, 7, 9, 11, 15	2	4
15	9	8	5, 7, 9, 11, 15	3	3
15	11	7	3, 5, 7, 9, 11, 15	1	3
11, 15	7	7	3, 5, 7, 9, 11, 15	2	2
11, 15	9	6	1, 5, 7, 9, 11, 15	2	5
7, 11, 15	2	7	1, 5, 7, 9, 11, 15	3	4
7, 11, 15	9	6	1, 3, 5, 7, 9, 11, 15	2	2
9, 15	2	8	2, 7, 11, 15	9	6
9, 15	11	7	2, 9, 15	11	7
9, 11, 15	2	7	2, 9, 11, 15	5	7
9, 11, 15	5	6	2, 9, 11, 15	7	6
9, 11, 15	7	5	2, 7, 9, 11, 15	3	6
7, 9, 11, 15	2	6	2, 7, 9, 11, 15	5	5
7, 9, 11, 15	3	5	2, 3, 7, 9, 11, 15	5	4
7, 9, 11, 15	5	4	2, 5, 9, 11, 15	7	5
3, 7, 9, 11, 15	2	5	2, 5, 7, 9, 11, 15	1	4
3, 7, 9, 11, 15	5	4	2, 5, 7, 9, 11, 15	3	3
5, 9, 11, 15	2	6	2, 3, 5, 7, 9, 11, 15	1	1
5, 9, 11, 15	7	5	1, 2, 5, 7, 9, 11, 15	3	3
5, 7, 9, 11, 15	1	5			

Table 2: Dimensions for $\Gamma = \langle 4, 6, 13\rangle$, $m = 1$

For $\mathcal{R} = \mathbb{C}[\![t^4, t^6+t^9]\!]$ corresponding to the $\Gamma = \langle 4, 6, 15\rangle$ and cable $\mathrm{Cab}(15, 2)T(3, 2)$, the situation is very similar. We checked that

$$\sum_{m=0}^{\infty} \sum_{\{D_0=D,\dots,D_m\}} q^{|D|+m} t^{\delta-\dim(J^m_{\mathcal{R}}[D])} a^m = \mathcal{H}_{\{3,2\},\{2,3\}}(\square; q, t, a) =$$

$$\begin{aligned}
&1 + qt + q^9t^9 + q^2(t + t^2) + q^3(t + t^2 + t^3) + a^3\big(q^6 + q^7t + q^8t^2 + q^9t^3\big)\\
&\quad + q^4(2t^2 + t^3 + t^4) + q^5(2t^3 + t^4 + t^5) + q^6(2t^4 + t^5 + t^6)\\
&\qquad\qquad + q^7(2t^5 + t^6 + t^7) + q^8(t^6 + t^7 + t^8)\\
&\quad + a^2\big(q^3 + q^4(1 + t) + q^5(1 + 2t + t^2) + q^6(2t + 2t^2 + t^3)\\
&\qquad + q^7(2t^2 + 2t^3 + t^4) + q^8(2t^3 + 2t^4 + t^5) + q^9(t^4 + t^5 + t^6)\big)\\
&\quad + a\big(q + q^2(1 + t) + q^3(1 + 2t + t^2)\\
&\qquad + q^4(3t + 2t^2 + t^3) + q^5(t + 4t^2 + 2t^3 + t^4) + q^6(t^2 + 4t^3 + 2t^4 + t^5)\\
&\qquad\quad + q^7(t^3 + 4t^4 + 2t^5 + t^6) + q^8(t^4 + 3t^5 + 2t^6 + t^7) + q^9(t^6 + t^7 + t^8)\big).
\end{aligned}$$

D–sets	$\vec{g}$	dim	D–sets	$\vec{g}$	dim
15	9, 11	8	3, 7, 9, 11, 15	2, 5	5
11, 15	7, 9	7	5, 9, 11, 15	2, 7	6
7, 11, 15	2, 9	7	5, 7, 9, 11, 15	1, 2	6
9, 15	2, 11	8	5, 7, 9, 11, 15	1, 3	5
9, 11, 15	2, 5	8	5, 7, 9, 11, 15	2, 3	4
9, 11, 15	2, 7	7	3, 5, 7, 9, 11, 15	1, 2	3
9, 11, 15	5, 7	6	1, 5, 7, 9, 11, 15	2, 3	5
7, 9, 11, 15	2, 3	7	2, 9, 11, 15	5, 7	7
7, 9, 11, 15	2, 5	6	2, 7, 9, 11, 15	3, 5	6
7, 9, 11, 15	3, 5	5	2, 5, 7, 9, 11, 15	1, 3	4

Table 3: Dimensions for $\Gamma = \langle 4, 6, 13\rangle$, $m = 2$

See the same section of [8]. Note that formulas (3.1) and (3.2) there for the DAHA–Betti polynomials are obtained from the superpolynomials as $a = 0, q = 1$.

4.2 The case of $\Gamma = \langle 4, 14, 31\rangle$

We checked the main conjecture in many cases for the series $(4, 2u, v)$ of *Puiseux exponents*. This example is of importance because quite a few features of Theorem 3.1 (and our proof) cannot be seen for the subfamily $(4, 6, v)$. For $\Gamma = \langle 4, 14, 31\rangle$, we list all admissible D–flags $\mathcal{D}$ and the dimensions $\dim J^m_{\mathcal{R}}[\mathcal{D}]$ for $m = 3$ (the greatest possible value). The calculation of dimensions at the maximal m as a matter of fact includes a lot of information about the dimensions for previous lengths m, so this is a good test of our conjecture. However, we restrict ourselves with a^3 here due to practical reasons. The total number of admissible $\mathcal{D}$ is 1071, but there are "only" 85 such D–flags of the top length for $m = 3$. The dimensions are in Tables 4 and 5 (its continuation) in the same format as above. Compare them with the coefficients of a^3 of the corresponding superpolynomial from [8], which we provide:

$$\mathcal{R} = \mathbb{C}[\![z^4, z^{14} + z^{17}]\!], \quad \mathcal{T} = \mathrm{Cab}(31, 2)T(7, 2);$$

$\mathcal{H}_{\{7,2\},\{2,3\}}(\square; q, t, a) =$

$$1+qt+q^2t+q^3t+q^2t^2+q^3t^2+2q^4t^2+q^5t^2+q^6t^2+q^3t^3+q^4t^3+2q^5t^3+2q^6t^3+2q^7t^3+q^8t^3+q^9t^3+q^4t^4+q^5t^4+2q^6t^4+2q^7t^4+3q^8t^4+2q^9t^4+2q^{10}t^4+q^5t^5+q^6t^5+2q^7t^5+2q^8t^5+3q^9t^5+3q^{10}t^5+3q^{11}t^5+q^6t^6+q^7t^6+2q^8t^6+2q^9t^6+3q^{10}t^6+3q^{11}t^6+4q^{12}t^6+q^7t^7+q^8t^7+2q^9t^7+2q^{10}t^7+3q^{11}t^7+3q^{12}t^7+4q^{13}t^7+q^8t^8+q^9t^8+2q^{10}t^8+2q^{11}t^8+3q^{12}t^8+3q^{13}t^8+4q^{14}t^8+q^9t^9+q^{10}t^9+2q^{11}t^9+2q^{12}t^9+3q^{13}t^9+3q^{14}t^9+4q^{15}t^9+$$

$$q^{10}t^{10}+q^{11}t^{10}+2q^{12}t^{10}+2q^{13}t^{10}+3q^{14}t^{10}+3q^{15}t^{10}+3q^{16}t^{10}+q^{11}t^{11}+q^{12}t^{11}+2q^{13}t^{11}+2q^{14}t^{11}+3q^{15}t^{11}+3q^{16}t^{11}+2q^{17}t^{11}+q^{12}t^{12}+q^{13}t^{12}+2q^{14}t^{12}+2q^{15}t^{12}+3q^{16}t^{12}+2q^{17}t^{12}+q^{18}t^{12}+q^{13}t^{13}+q^{14}t^{13}+2q^{15}t^{13}+2q^{16}t^{13}+3q^{17}t^{13}+q^{18}t^{13}+q^{14}t^{14}+q^{15}t^{14}+2q^{16}t^{14}+2q^{17}t^{14}+2q^{18}t^{14}+q^{15}t^{15}+q^{16}t^{15}+2q^{17}t^{15}+2q^{18}t^{15}+q^{19}t^{15}+q^{16}t^{16}+q^{17}t^{16}+2q^{18}t^{16}+q^{19}t^{16}+q^{17}t^{17}+q^{18}t^{17}+2q^{19}t^{17}+q^{18}t^{18}+q^{19}t^{18}+q^{20}t^{18}+q^{19}t^{19}+q^{20}t^{19}+q^{20}t^{20}+q^{21}t^{21}$$

$$+a^3\big(q^6+q^7t+q^8t+q^9t+q^8t^2+q^9t^2+2q^{10}t^2+q^{11}t^2+q^{12}t^2+q^9t^3+q^{10}t^3+2q^{11}t^3+2q^{12}t^3+2q^{13}t^3+q^{10}t^4+q^{11}t^4+2q^{12}t^4+2q^{13}t^4+3q^{14}t^4+q^{11}t^5+q^{12}t^5+2q^{13}t^5+2q^{14}t^5+3q^{15}t^5+q^{12}t^6+q^{13}t^6+2q^{14}t^6+2q^{15}t^6+3q^{16}t^6+q^{13}t^7+q^{14}t^7+2q^{15}t^7+2q^{16}t^7+3q^{17}t^7+q^{14}t^8+q^{15}t^8+2q^{16}t^8+2q^{17}t^8+2q^{18}t^8+q^{15}t^9+q^{16}t^9+2q^{17}t^9+2q^{18}t^9+q^{19}t^9+q^{16}t^{10}+q^{17}t^{10}+2q^{18}t^{10}+q^{19}t^{10}+q^{17}t^{11}+q^{18}t^{11}+2q^{19}t^{11}+q^{18}t^{12}+q^{19}t^{12}+q^{20}t^{12}+q^{19}t^{13}+q^{20}t^{13}+q^{20}t^{14}+q^{21}t^{15}\big)$$

$$+a^2\big(q^3+q^4+q^5+q^4t+2q^5t+3q^6t+2q^7t+q^8t+q^5t^2+2q^6t^2+4q^7t^2+4q^8t^2+4q^9t^2+2q^{10}t^2+q^{11}t^2+q^6t^3+2q^7t^3+4q^8t^3+5q^9t^3+6q^{10}t^3+5q^{11}t^3+3q^{12}t^3+q^7t^4+2q^8t^4+4q^9t^4+5q^{10}t^4+7q^{11}t^4+7q^{12}t^4+5q^{13}t^4+q^8t^5+2q^9t^5+4q^{10}t^5+5q^{11}t^5+7q^{12}t^5+8q^{13}t^5+6q^{14}t^5+q^9t^6+2q^{10}t^6+4q^{11}t^6+5q^{12}t^6+7q^{13}t^6+8q^{14}t^6+6q^{15}t^6+q^{10}t^7+2q^{11}t^7+4q^{12}t^7+5q^{13}t^7+7q^{14}t^7+8q^{15}t^7+6q^{16}t^7+q^{11}t^8+2q^{12}t^8+4q^{13}t^8+5q^{14}t^8+7q^{15}t^8+8q^{16}t^8+5q^{17}t^8+q^{12}t^9+2q^{13}t^9+4q^{14}t^9+5q^{15}t^9+7q^{16}t^9+7q^{17}t^9+3q^{18}t^9+q^{13}t^{10}+2q^{14}t^{10}+4q^{15}t^{10}+5q^{16}t^{10}+7q^{17}t^{10}+5q^{18}t^{10}+q^{19}t^{10}+q^{14}t^{11}+2q^{15}t^{11}+4q^{16}t^{11}+5q^{17}t^{11}+6q^{18}t^{11}+2q^{19}t^{11}+q^{15}t^{12}+2q^{16}t^{12}+4q^{17}t^{12}+5q^{18}t^{12}+4q^{19}t^{12}+q^{16}t^{13}+2q^{17}t^{13}+4q^{18}t^{13}+4q^{19}t^{13}+q^{20}t^{13}+q^{17}t^{14}+2q^{18}t^{14}+4q^{19}t^{14}+2q^{20}t^{14}+q^{18}t^{15}+2q^{19}t^{15}+3q^{20}t^{15}+q^{19}t^{16}+2q^{20}t^{16}+q^{21}t^{16}+q^{20}t^{17}+q^{21}t^{17}+q^{21}t^{18}\big)$$

$$+a\big(q+q^2+q^3+q^2t+2q^3t+3q^4t+2q^5t+q^6t+q^3t^2+2q^4t^2+4q^5t^2+4q^6t^2+4q^7t^2+2q^8t^2+q^9t^2+q^4t^3+2q^5t^3+4q^6t^3+5q^7t^3+6q^8t^3+5q^9t^3+4q^{10}t^3+q^{11}t^3+q^5t^4+2q^6t^4+4q^7t^4+5q^8t^4+7q^9t^4+7q^{10}t^4+7q^{11}t^4+2q^{12}t^4+q^6t^5+2q^7t^5+4q^8t^5+5q^9t^5+7q^{10}t^5+8q^{11}t^5+9q^{12}t^5+3q^{13}t^5+q^7t^6+2q^8t^6+4q^9t^6+5q^{10}t^6+7q^{11}t^6+8q^{12}t^6+10q^{13}t^6+3q^{14}t^6+q^8t^7+2q^9t^7+4q^{10}t^7+5q^{11}t^7+7q^{12}t^7+8q^{13}t^7+10q^{14}t^7+3q^{15}t^7+q^9t^8+2q^{10}t^8+4q^{11}t^8+5q^{12}t^8+7q^{13}t^8+8q^{14}t^8+10q^{15}t^8+3q^{16}t^8+q^{10}t^9+2q^{11}t^9+4q^{12}t^9+5q^{13}t^9+7q^{14}t^9+8q^{15}t^9+9q^{16}t^9+2q^{17}t^9+q^{11}t^{10}+2q^{12}t^{10}+4q^{13}t^{10}+5q^{14}t^{10}+7q^{15}t^{10}+8q^{16}t^{10}+7q^{17}t^{10}+q^{18}t^{10}+q^{12}t^{11}+2q^{13}t^{11}+4q^{14}t^{11}+5q^{15}t^{11}+7q^{16}t^{11}+7q^{17}t^{11}+4q^{18}t^{11}+q^{13}t^{12}+2q^{14}t^{12}+4q^{15}t^{12}+5q^{16}t^{12}+7q^{17}t^{12}+5q^{18}t^{12}+q^{19}t^{12}+q^{14}t^{13}+2q^{15}t^{13}+4q^{16}t^{13}+5q^{17}t^{13}+6q^{18}t^{13}+2q^{19}t^{13}+q^{15}t^{14}+2q^{16}t^{14}+4q^{17}t^{14}+5q^{18}t^{14}+4q^{19}t^{14}+q^{16}t^{15}+2q^{17}t^{15}+4q^{18}t^{15}+4q^{19}t^{15}+q^{20}t^{15}+q^{17}t^{16}+2q^{18}t^{16}+4q^{19}t^{16}+2q^{20}t^{16}+q^{18}t^{17}+2q^{19}t^{17}+3q^{20}t^{17}+q^{19}t^{18}+2q^{20}t^{18}+q^{21}t^{18}+q^{20}t^{19}+q^{21}t^{19}+q^{21}t^{20}\big).$$

4.3 The series $(6, 8, v)$

The series with *Puiseux exponents* $(4, 2u, v)$ above corresponds to the somewhat special links $\mathrm{Cab}(2u+v, 2)T(u, 2)$; torus knots and cables for $(2p+1, 2)$ are known to have some special symmetries.

D–sets	$\vec{g}$	dim
27, 37, 41	10, 23, 33	21
27, 33, 37, 41	10, 23, 29	20
27, 29, 33, 37, 41	10, 23, 25	19
25, 27, 29, 33, 37, 41	10, 21, 23	18
21, 25, 27, 29, 33, 37, 41	10, 17, 23	18
17, 21, 25, 27, 29, 33, 37, 41	10, 13, 23	18
23, 27, 33, 37, 41	10, 19, 29	20
23, 27, 29, 33, 37, 41	10, 19, 25	19
23, 25, 27, 29, 33, 37, 41	10, 19, 21	17
21, 23, 25, 27, 29, 33, 37, 41	10, 17, 19	16
17, 21, 23, 25, 27, 29, 33, 37, 41	10, 13, 19	16
13, 17, 21, 23, 25, 27, 29, 33, 37, 41	9, 10, 19	16
19, 23, 27, 29, 33, 37, 41	10, 15, 25	19
19, 23, 25, 27, 29, 33, 37, 41	10, 15, 21	17
19, 21, 23, 25, 27, 29, 33, 37, 41	10, 15, 17	15
17, 19, 21, 23, 25, 27, 29, 33, 37, 41	10, 13, 15	14
13, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	9, 10, 15	14
9, 13, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	5, 10, 15	15
15, 19, 23, 25, 27, 29, 33, 37, 41	10, 11, 21	17
15, 19, 21, 23, 25, 27, 29, 33, 37, 41	10, 11, 17	15
15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	10, 11, 13	13
13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	9, 10, 11	12
9, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	5, 10, 11	13
5, 9, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 10, 11	14
11, 15, 19, 21, 23, 25, 27, 29, 33, 37, 41	7, 10, 17	15
11, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	7, 10, 13	13
11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	7, 9, 10	11
9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	5, 7, 10	11
5, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 7, 10	12
7, 11, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	3, 10, 13	14
7, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	3, 9, 10	12
7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	3, 5, 10	11
5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 3, 10	11
10, 17, 21, 25, 27, 29, 33, 37, 41	6, 13, 23	18
10, 23, 27, 33, 37, 41	6, 19, 29	20
10, 23, 27, 29, 33, 37, 41	6, 19, 25	19
10, 23, 25, 27, 29, 33, 37, 41	6, 19, 21	18
10, 21, 23, 25, 27, 29, 33, 37, 41	6, 17, 19	17
10, 17, 21, 23, 25, 27, 29, 33, 37, 41	6, 13, 19	17
10, 13, 17, 21, 23, 25, 27, 29, 33, 37, 41	6, 9, 19	16
10, 19, 23, 27, 29, 33, 37, 41	6, 15, 25	19
10, 19, 23, 25, 27, 29, 33, 37, 41	6, 15, 21	18
10, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 15, 17	16

Table 4: Dimensions for $\Gamma = \langle 4, 14, 31 \rangle$, $m = 3$ (I)

D–sets	$\vec{g}$	dim
10, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 13, 15	15
10, 13, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 9, 15	14
9, 10, 13, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	5, 6, 15	14
10, 15, 19, 23, 25, 27, 29, 33, 37, 41	6, 11, 21	17
10, 15, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 11, 17	16
10, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 11, 13	14
10, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 9, 11	12
9, 10, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	5, 6, 11	12
5, 9, 10, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 6, 11	13
10, 11, 15, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 7, 17	15
10, 11, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 7, 13	13
10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	6, 7, 9	10
9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	5, 6, 7	9
5, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 6, 7	10
7, 10, 11, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	3, 6, 13	13
7, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	3, 6, 9	10
7, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	3, 5, 6	8
5, 7, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 3, 6	8
6, 10, 17, 21, 25, 27, 29, 33, 37, 41	2, 13, 23	18
6, 10, 17, 21, 23, 25, 27, 29, 33, 37, 41	2, 13, 19	17
6, 10, 13, 17, 21, 23, 25, 27, 29, 33, 37, 41	2, 9, 19	16
6, 10, 19, 23, 27, 29, 33, 37, 41	2, 15, 25	19
6, 10, 19, 23, 25, 27, 29, 33, 37, 41	2, 15, 21	18
6, 10, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 15, 17	17
6, 10, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 13, 15	16
6, 10, 13, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 9, 15	15
6, 9, 10, 13, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 5, 15	14
6, 10, 15, 19, 23, 25, 27, 29, 33, 37, 41	2, 11, 21	17
6, 10, 15, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 11, 17	16
6, 10, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 11, 13	15
6, 10, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 9, 11	13
6, 9, 10, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 5, 11	12
5, 6, 9, 10, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 2, 11	12
6, 10, 11, 15, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 7, 17	15
6, 10, 11, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 7, 13	14
6, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 7, 9	11
6, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 5, 7	9
5, 6, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 2, 7	9
6, 7, 10, 11, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 3, 13	13
6, 7, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 3, 9	10
6, 7, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	2, 3, 5	7
5, 6, 7, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 37, 41	1, 2, 3	6

Table 5: Dimensions for $\Gamma = \langle 4, 14, 31\rangle$, $m = 3$ (II)

We will now consider the series $(6,8,v)$ for odd $v \geq 9$ from (1-17) with the link $\mathrm{Cab}(25+(v-9),2)T(4,3)$. The corresponding ring, semigroup, δ–invariant, and Euler number of $J_{\mathcal{R}}$ are

$$\mathcal{R} = \mathbb{C}[\![z^6, z^8 + z^v]\!], \qquad \Gamma = \langle 6, 8, 25+(v-9)\rangle,$$
$$\delta = 18 + \tfrac{1}{2}(v-9), \quad e(J_{\mathcal{R}}) = 227 + \tfrac{25}{2}(v-9).$$

Piontkowski [38] provides the Euler number, the total number of Γ–modules, which is $273+\frac{25}{2}(v-9)$, and the number of *nonadmissible* ones. The latter is 46 for any v. The difference $273+\frac{25}{2}(v-9)-46$ is exactly the Euler number, since all admissible cells in the decomposition from [38] are diffeomorphic to $\mathbb{A}^N$ in the case under consideration and the homology can be readily calculated.

For our conjecture, we need to know the set of all Γ–modules Δ, those that are nonadmissible, and the dimensions $\dim(J_{\mathcal{R}}[\mathcal{D}])$. It is not too difficult to find all nonadmissible modules following [38], but they are not provided in his paper. There are "generic" nonadmissible modules and three exceptional ones, which we are going to describe.

We examine the elements

$$T_{ij}^{pq} = \phi_i m_p - \phi_j m_q \in \mathcal{M},$$

where $D = D_{\mathcal{M}}$ is the set of gaps of $\mathcal{M}$, $i<j\in\Gamma$, $p>q\in\Delta_{\mathcal{M}}=v(\mathcal{M})$, $v(\phi_i)=i$, and $v(m_p)=p$, where we use the valuation $v\colon \mathcal{O}\to\mathbb{Z}_+$. Recall that modules $\mathcal{M}$ are submodules in $\mathcal{O}=\mathbb{C}[\![z]\!]$ with an element of valuation 0.

We will assume that the leading z–monomial in ϕ_i, m_k has the coefficient 1. For instance, $m_0 = 1+\sum_{p>0}\lambda_0^p z^p$. The choice of these elements is of course nonunique (higher terms can be added to them).

Proposition 4.1 *All nonadmissible $\mathcal{M}$ for $\mathcal{R}=\mathbb{C}[\![t^6,t^8+t^v]\!]$ can be described as follows. In the differences from Section 4.3, let*

(a) $q=0$, $p=2,4,10$; (b) $q=2$, $p=4$; (c) $p>q\in\{0,2,4\}$.

Then, let us impose the following relations for i, j in Section 4.3:

(4-2) $\qquad i+p+1\notin\Delta_{\mathcal{M}}$ *and* $i+p=j+q$ $\qquad$ *for* $i,j\in\Gamma$.

Here the latter results in the inequalities $v(T_{ij}^{pq})\geq 1+p+1$, which can be only strict due to the former since $T_{ij}^{pq}\in\mathcal{M}$.

The nonadmissibility of $\mathcal{M}$ of type (a) *or* (b) *occurs if and only if there exist no $m_p, m_q\in\mathcal{M}$ in Section 4.3 satisfying* (4-2) *for* ***all possible*** *choices of i, j there. In*

the case of (c), *the absence of* $m_0, m_2, m_4 \in \mathcal{M}$ *must be checked for each of the three choices of* p, q *there altogether and every possible* i, j *satisfying Section 4.3.* □

Let us list the D–sets (the sets of gaps) for all 46 nonadmissible modules $\mathcal{M}$. In Table 6, $(v-9)$ must be added to all gaps in the second half of the first column (clearly visible). The second column contains the smallest $p>0$ in D that ensures the nonadmissibility in the case of (a), the letters b or c stand for the remaining three exceptional cases. The third column contains the first $g=g_{\min}=i+p+1\notin D$ such that its absence in D contradicts the absence of the previous gaps $g'i'+p+1<g_{\min}$ in D, which is for a given pair from (a,b,c). Such a gap is provided only for the pair $(0,2)$ in the case of (c).

The nonadmissibility can occur only if there are at least two possible pairs (i,j) satisfying (4-2). The corresponding conditions are simple equalities for the differences $\lambda^1_p-\lambda^1_q$, where $m_p=z^p+\lambda^1_p z^{p+1}+\cdots$. If the difference $\lambda^1_p-\lambda^1_q$ takes different values for different pairs (i,j), then the module cannot be admissible. In the case of c, such a contradiction can be reached only if *three* such sequences of equalities are considered together, ie for $\lambda^1_0-\lambda^1_2$, $\lambda^1_0-\lambda^1_4$, and $\lambda^1_2-\lambda^1_4$. Generally, higher-order z–expansions can be necessary here, ie $\lambda^{i>1}_p$ may occur; however this is not the case with $(6,8,v)$.

Using this table and the list of all $273+\frac{25}{2}(v-9)$ modules Δ, we checked Conjecture 2.5 for $t=1$ (ie ignoring the dimensions) for quite a few v. Here and in *all* calculations we performed it appeared sufficient to replace the admissibility of $\mathcal{D}$ by the admissibility of all D_i (separately), a potentially weaker condition. Generalizing [38], we verified here that all subvarieties $J^m_{\mathcal{R}}[\mathcal{D}]$ are diffeomorphic to proper $\mathbb{A}^N$, so we are in the situation of (2-6).

Let us provide some details. Recall that the D–flag of length m originated at D is by definition the sequence of D–sets for Γ–modules:

$$\mathcal{D}=\big\{D_0=D,\ D_1=D\cup\{g_1\},\ \dots,\ D_m=D\cup\{g_1,g_2,\dots,g_m\}\big\}, \tag{4-3}$$

where the inequalities $g_1<g_2<\cdots<g_m\in G\setminus D$ are imposed. Let us illustrate numerically the importance of this very ordering in our definition of D–flags.

As it results from Proposition 2.3, each set $D\cup\{g_i\}$ corresponds to a certain Γ–module for any $1\le i\le m$. However, apart from torus knots, these conditions (imposed together) are generally significantly weaker than the conditions we need. The Γ–modules D_i from (4-3) are not always admissible (coming from some $\mathcal{M}$) if all $D\cup\{g_i\}$ are.

D–sets	under $\{\cdot\}+(v-9)$	p	g	D–sets	under $\{\cdot\}+(v-9)$	p	g
10,	35	10	19	2, 10,	21, 23, 27, 29, 35	2	19
10,	27, 35	10	19	4, 10,	19, 21, 27, 29, 35	4	17
10,	29, 35	10	19	4, 10,	19, 23, 27, 29, 35	4	17
2, 10,	27, 35	2	19	4, 10,	21, 23, 27, 29, 35	4	17
4, 10,	29, 35	4	17	10,	15, 21, 23, 27, 29, 35	10	19
10,	23, 29, 35	10	19	2, 4, 10,	17, 23, 27, 29, 35	2	19
10,	27, 29, 35	10	19	2, 4, 10,	19, 21, 27, 29, 35	4	17
2, 10,	27, 29, 35	2	19	2, 4, 10,	19, 23, 27, 29, 35	4	17
4, 10,	23, 29, 35	4	17	2, 4, 10,	21, 23, 27, 29, 35	2	19
4, 10,	27, 29, 35	4	17	2, 10,	15, 21, 23, 27, 29, 35	2	19
10,	21, 27, 29, 35	10	19	2, 10,	17, 21, 23, 27, 29, 35	2	19
10,	23, 27, 29, 35	10	19	4, 10,	15, 21, 23, 27, 29, 35	4	17
2, 4, 10,	27, 29, 35	2	19	4, 10,	19, 21, 23, 27, 29, 35	4	17
2, 10,	21, 27, 29, 35	2	19	2, 4, 10,	15, 21, 23, 27, 29, 35	2	19
2, 10,	23, 27, 29, 35	2	19	2, 4, 10,	17, 19, 23, 27, 29, 35	***b***	21
4, 10,	19, 27, 29, 35	4	17	2, 4, 10,	17, 21, 23, 27, 29, 35	2	19
4, 10,	21, 27, 29, 35	4	17	2, 4, 10,	19, 21, 23, 27, 29, 35	4	17
4, 10,	23, 27, 29, 35	4	17	2, 10,	15, 17, 21, 23, 27, 29, 35	2	19
10,	21, 23, 27, 29, 35	10	19	4, 10,	15, 19, 21, 23, 27, 29, 35	4	17
2, 4, 10,	19, 27, 29, 35	4	17	2, 4, 10,	15, 17, 21, 23, 27, 29, 35	2	19
2, 4, 10,	21, 27, 29, 35	2	19	2, 4, 10,	15, 19, 21, 23, 27, 29, 35	4	17
2, 4, 10,	23, 27, 29, 35	2	19	2, 4, 10,	17, 19, 21, 23, 27, 29, 35	***c***	11
2, 10,	17, 23, 27, 29, 35	2	19	2, 4, 10,	15, 17, 19, 21, 23, 27, 29, 35	***c***	11

Table 6: Nonadmissible modules for $(6, 8, v)$

We set $\epsilon(D) = 1, 0$ correspondingly for admissible and nonadmissible D and $\epsilon(\mathcal{D}) = \prod_{i=1}^{m} \epsilon(D_i)$, where $D_i = D_{\mathcal{M}_i}$. Then $\epsilon(\mathcal{D}) = 1$ is equivalent to the admissibility of $\mathcal{D}$ (in the example under consideration). We expect this implication to hold in general, but cannot prove this at the moment. Thus $\epsilon(\mathcal{D}) = 1$ implies that $\prod_{i=1}^{m} \epsilon(D \cup g_i) = 1$, but the latter condition is not generally insufficient for the former.

Namely, if the admissibility of D–flags were defined as $\prod_{i=1}^{m} \epsilon(D \cup \{g_i\}) = 1$ instead of $\epsilon(\mathcal{D}) = 1$, ie separately for each and every $D \cup \{g_i\}$, then there would be 14 extra (wrong) terms (with multiplicities) in (4-4) below. This clearly demonstrates that our flags are generally more subtle than using "marks" for torus knots in [16] and other related works.

The smallest example is as follows. Using Table 6 with $v = 9$, the set of all g_i for $D = [10, 17, 19, 23, 27, 29, 35]$ such that $D \cup \{g_i\}$ is admissible is $\{2, 4, 11, 21\}$.

However, $D\cup\{2,4\}=[2,4,10,17,19,23,27,29,35]$ is a nonadmissible D–set (it is marked by b in the table).

Finally, we arrive at the following identity:

$$\text{(4-4)}\quad \sum_{m=0}^{\infty}\sum_{\mathcal{D}=\{D_0,\dots,D_m\}}\epsilon(\mathcal{D})q^{|D_0|+m}a^m=$$

$$\begin{aligned}
&1+q+2q^2+3q^3+5q^4+7q^5+10q^6+12q^7+16q^8+19q^9+22q^{10}\\
&\qquad+24q^{11}+25q^{12}+24q^{13}+22q^{14}+17q^{15}+11q^{16}+5q^{17}+q^{18}\\
&+a\big(q+2q^2+4q^3+7q^4+12q^5+18q^6+26q^7+35q^8+46q^9+56q^{10}\\
&\qquad+66q^{11}+72q^{12}+74q^{13}+70q^{14}+59q^{15}+41q^{16}+21q^{17}+5q^{18}\big)\\
&+a^2\big(q^3+2q^4+5q^5+9q^6+16q^7+24q^8+36q^9+48q^{10}+62q^{11}\\
&\qquad+74q^{12}+82q^{13}+83q^{14}+76q^{15}+58q^{16}+34q^{17}+10q^{18}\big)\\
&+a^3\big(q^6+2q^7+5q^8+9q^9+15q^{10}+22q^{11}+31q^{12}+38q^{13}\\
&\qquad+44q^{14}+44q^{15}+38q^{16}+26q^{17}+10q^{18}\big)\\
&+a^4\big(q^{10}+2q^{11}+4q^{12}+6q^{13}+9q^{14}+11q^{15}+11q^{16}+9q^{17}+5q^{18}\big)\\
&+a^5\big(q^{15}+q^{16}+q^{17}+q^{18}\big).
\end{aligned}$$

The latter sum exactly coincides with $\mathcal{H}_{\vec{r},\vec{s}}(\Box;q,t=1,a)$ from formula (4) in Section 3.2 of [8], where

$$\vec{r}=\{4,2\},\quad \vec{s}=\{3,1\},\quad \mathcal{T}=\mathrm{Cab}(25,2)T(4,3).$$

Also see formula (3.4) there for the corresponding DAHA–Betti polynomial, which is the a–constant term of (4-4) upon the substitution $q\mapsto 1/t$ and multiplication by t^{18}:

$$\begin{aligned}
&1+5t+11t^2+17t^3+22t^4+24t^5+25t^6+24t^7+22t^8+19t^9\\
&\qquad+16t^{10}+12t^{11}+10t^{12}+7t^{13}+5t^{14}+3t^{15}+2t^{16}+t^{17}+t^{18}.
\end{aligned}$$

Here we use the superduality.

We also checked that $J^m_{\mathcal{R}}[\mathcal{D}]$ are always affine spaces, found formulas for the dimensions of $J^{m=0,1}_{\mathcal{R}}[D]$ and correspondingly verified (2-6) for the coefficients of superpolynomials of a^0,a^1 in the considered case.

4.4 Beyond "Piontkowski"

Examples are given in [38] where his approach does not work because the corresponding cells $J_r[D]$ are nonaffine spaces. Thus the count of such cells and knowing their dimensions is insufficient for obtaining the Betti numbers and the Euler number of J_r. Some "negative" examples are provided in the table after Theorem 13 in [38]. We found that not always the cells in his table are really nonaffine, but the phenomenon he discovered is of course important. It is unclear whether nonaffine cells and nonadmissible D and $\mathcal{D}$ are topological. This is of obvious interest due to our main conjecture and because of the exciting link to orbital integral in type A.

We mostly consider a relatively simple singularity with Puiseux exponents $(6, 9, 10)$ and the cable $\mathrm{Cab}(19, 3)T(3, 2)$. Its ring is $\mathcal{R} = \mathbb{C}[\![t^6, t^9 + t^{10}]\!]$, the valuation semigroup is $\Gamma = \langle 6, 9, 19\rangle$ and $\delta = |\mathbb{N} \setminus \Gamma| = 21$ in this case. See [8, Section 3.2] for details and the formula for the DAHA superpolynomial; the DAHA–Betti polynomial ($a = 0, q = 1$) is formula (3.3) there, which does provide correct Betti numbers for the corresponding $J_{\mathcal{R}}$.

This example is transitional in a sense; all cells are still affine spaces, but the justification of this fact (straightforward, using computers) becomes more involved. We considered some deformations of parameters of $\mathcal{R}$ and think that Table 7 depends only on the corresponding Γ (ie this table is of a topological nature) but this is not clear.

The simplest example of nonaffine $J_{\mathcal{R}}[D]$ in this case given in [38] is

$$D = [3, 7, 10, 13, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41],$$

but we found that the corresponding $J_{\mathcal{R}}[D]$ is biregular to $\mathbb{A}^{14}$. The only problem with this and two other similar sets D (our program obtained) is that the natural λ–variables Piontkowski uses (as do we) are inconvenient to parametrize $J_{\mathcal{R}}[D]$; a certain (linear) change of variables is necessary. The other two D–sets with a similar behavior (when a straight elimination of the λ–variables is insufficient) are

$$[3, 10, 13, 16, 20, 22, 23, 26, 29, 32, 35, 41],$$
$$[3, 10, 13, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41].$$

All three (and any other cells) are affine spaces in this case. The next example of a nonaffine cell from [38] is for $\mathbb{C}[\![t^6, t^9 + t^{13}]\!]$; we confirm it. See Appendix A in the online supplement.

We note that our program routinely calculates $|J_{\mathcal{R}}[D](\mathbb{F}_3)|$ for "suspicious" D to double check the direct verification of the affineness of the cells (mostly automated).

22, 41	2, 3, 5, 8, 11, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 41	1, 3, 4, 7, 10, 13, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
22, 35, 41	3, 5, 8, 11, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 32, 41	2, 3, 8, 11, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 35, 41	2, 3, 5, 8, 11, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
16, 22, 35, 41	1, 4, 7, 10, 13, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 35, 41	1, 3, 4, 7, 10, 13, 16, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 29, 35, 41	3, 8, 11, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 32, 35, 41	3, 5, 11, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
16, 22, 32, 35, 41	3, 5, 8, 11, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 29, 35, 41	3, 4, 7, 10, 13, 16, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 32, 35, 41	2, 3, 8, 11, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 26, 32, 35, 41	1, 4, 7, 10, 13, 16, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 29, 32, 35, 41	4, 7, 10, 13, 16, 20, 22, 23, 26, 29, 32, 35, 41
13, 16, 22, 32, 35, 41	3, 11, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 13, 16, 22, 32, 35, 41	3, 8, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 26, 32, 35, 41	3, 8, 11, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 29, 32, 35, 41	3, 5, 11, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 22, 23, 29, 32, 35, 41	3, 4, 7, 10, 13, 16, 22, 23, 26, 29, 32, 35, 41
3, 22, 26, 29, 32, 35, 41	4, 7, 10, 13, 16, 22, 23, 26, 29, 32, 35, 41
13, 16, 22, 29, 32, 35, 41	3, 14, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 13, 16, 22, 29, 32, 35, 41	3, 11, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 23, 29, 32, 35, 41	3, 11, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 26, 29, 32, 35, 41	3, 8, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 20, 22, 26, 29, 32, 35, 41	3, 7, 10, 13, 16, 22, 23, 26, 29, 32, 35, 41
3, 22, 23, 26, 29, 32, 35, 41	7, 10, 13, 16, 22, 23, 26, 29, 32, 35, 41
10, 13, 16, 22, 29, 32, 35, 41	3, 16, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 10, 13, 16, 22, 29, 32, 35, 41	3, 14, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 20, 22, 26, 29, 32, 35, 41	3, 14, 16, 20, 22, 23, 26, 29, 32, 35, 41
3, 16, 22, 23, 26, 29, 32, 35, 41	3, 11, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 17, 22, 23, 26, 29, 32, 35, 41	3, 7, 10, 13, 16, 22, 26, 29, 32, 35, 41
3, 20, 22, 23, 26, 29, 32, 35, 41	7, 10, 13, 16, 22, 26, 29, 32, 35, 41
10, 13, 16, 22, 26, 29, 32, 35, 41	3, 17, 20, 22, 23, 26, 29, 32, 35, 41
3, 10, 13, 16, 22, 26, 29, 32, 35, 41	3, 16, 20, 22, 23, 26, 29, 32, 35, 41
3, 14, 20, 22, 23, 26, 29, 32, 35, 41	3, 16, 17, 22, 23, 26, 29, 32, 35, 41

Table 7: Nonadmissible D for $\Gamma = \langle 6, 9, 19\rangle$

These cardinalities must be $3^{\dim}$ for affine cells. The prime 3 is a "place of good reduction" for this $\mathcal{R}$. The reduction is bad modulo $p = 2$ and one needs to switch to

the *topologically* equivalent ring $\mathbb{C}[\![t^6+t^7,t^9]\!]$ before replacing $\mathbb{C}\mapsto\mathbb{F}_2$. We omit general analysis of places of bad reduction in this paper.

We will not discuss much the flags here, but let us mention that all cells are affine for $J_{\mathcal{R}}^{m=1}[\mathcal{D}]$ (their dimensions are all calculated) in the case of $\mathcal{R}=\mathbb{C}[\![t^6,t^9+t^{10}]\!]$. The list of (all) 70 nonadmissible D (ie for $m=0$) will be provided. The total number of modules Δ is $447=377+70$ in this case, and the Euler number is $e(J_{\mathcal{R}})=377$. Accordingly, we checked (numerically) the coincidence from (2-6) in the following two cases: (a) for all a when $t=1$ (for the admissibility of $\mathcal{D}$ understood as the admissibility of all D_i in this flag), and (b) for the coefficients of $\mathcal{H}(\square;q,t,a)$ from [8] of $a^{0,1}$.

The calculation in (a) greatly demonstrates the role of admissibility and the implications of the ordering $g_1<\cdots<g_m$, which are quite nontrivial combinatorially. The corresponding reduction of the superpolynomial for $\mathcal{R}=\mathbb{C}[\![t^6,t^9+t^{10}]\!]$ is

$$\begin{aligned}
&\mathcal{H}(\square;q,t=1,a)=\\
&1+q+2q^2+3q^3+5q^4+7q^5+10q^6+13q^7+17q^8+21q^9+25q^{10}+29q^{11}+33q^{12}\\
&\qquad+36q^{13}+37q^{14}+37q^{15}+34q^{16}+28q^{17}+20q^{18}+12q^{19}+5q^{20}+q^{21}\\
&+a^5\big(q^{15}+q^{16}+2q^{17}+2q^{18}+3q^{19}+2q^{20}+q^{21}\big)\\
&+a^4\big(q^{10}+2q^{11}+4q^{12}+7q^{13}+11q^{14}+15q^{15}+19q^{16}\\
&\qquad+22q^{17}+23q^{18}+21q^{19}+13q^{20}+5q^{21}\big)\\
&+a\big(q+2q^2+4q^3+7q^4+12q^5+18q^6+27q^7+37q^8+50q^9+63q^{10}+78q^{11}+91q^{12}\\
&\quad+105q^{13}+113q^{14}+118q^{15}+114q^{16}+100q^{17}+76q^{18}+48q^{19}+22q^{20}+5q^{21}\big)\\
&+a^3\big(q^6+2q^7+5q^8+9q^9+16q^{10}+24q^{11}+36q^{12}+47q^{13}+61q^{14}+71q^{15}\\
&\qquad+81q^{16}+82q^{17}+76q^{18}+57q^{19}+32q^{20}+10q^{21}\big)\\
&+a^2\big(q^3+2q^4+5q^5+9q^6+16q^7+25q^8+38q^9+53q^{10}+71q^{11}+90q^{12}+109q^{13}\\
&\qquad+126q^{14}+138q^{15}+143q^{16}+134q^{17}+111q^{18}+75q^{19}+38q^{20}+10q^{21}\big).
\end{aligned}$$

5 Some perspectives

The topics below are mostly open projects, but we believe that this section can be of interest; the relation to orbital integrals, affine Springer fibers and the motivic reformulation (5-2) of (2-7) are the key. We will first address the absence of colors and links in our main conjecture and related issues.

5.1 Adding colors

In contrast to [9] and previous paper, we restrict this one to algebraic uncolored knots and only type A is addressed in our conjecture. Let us briefly comment on this. Adding colors is expected via the curves in [31], so algebraic *links* are/seem necessary for this. Apart from rectangle Young diagrams, the coefficients of DAHA-superpolynomials are nonpositive, which is an obvious challenge for the geometric interpretation. This is the same for links (included uncolored ones). The Jacobian factors become *ind-schemes* for algebraic links and we need to follow [25] (divide by certain split tori) to make them *proper*.

One can try to bypass the nonpositivity issues switching to some *Hilbert schemes* instead of our flagged Jacobians, which corresponds to the unreduced topological setting. This can be similar to [36; 13]; see also Conjecture 5.3(ii) from [9].

Superpolynomials beyond A_n Our flagged Jacobian factors are related to the Hitchin and affine Springer fibers. The hope is that the DAHA-superpolynomials can be directly and geometrically determined by spectral curves for types B_n, C_n, D_n via the corresponding flagged Jacobian factors and/or Hilbert schemes. The rank stabilization was conjectured in [6]; the name "hyperpolynomials" is used instead of superpolynomials for non-A types. The spectral curves are generally nonunibranch, so the passage to links is necessary here. A certain confirmation is the following observation.

It is true in all known examples that $\mathcal{H}(b;q,t,u,a=0)$ for the hyperpolynomials from Section 4.2 of [6] depend only on the corresponding knot and the color b. This holds for the hyperpolynomials of type C, D calculated there for $T(3,2), T(7,2), T(4,3)$ and $b=\omega_1, 2\omega_1, \omega_2$. The hyperpolynomials of type D are the specialization of those of type C upon $u=1$ (see (4.7) there; u is the second t for B, C). The same holds for type B under $q^2\mapsto q$, $u\mapsto t$.

Furthermore, the hyperpolynomials $\mathcal{H}^{\mathfrak{ad}}_{\mathtt{r},\mathtt{s}}(q,t,a)$ for the *Deligne–Gross series* (extending the series $E_{6,7,8}$) introduced in Section 4.2 of [10] coincide at $a=0$ with $\mathcal{H}_{\mathtt{r},\mathtt{s}}(\square;q,t,a=0)$ of type A. The existence of $\mathcal{H}^{\mathfrak{ad}}_{\mathtt{r},\mathtt{s}}$ was confirmed only partially in [10] (and only for $T(3,2), T(4,3)$). These $\mathcal{H}^{\mathfrak{ad}}_{\mathtt{r},\mathtt{s}}$ are mysterious from the viewpoint of [6], since there can be no rank stabilization here. We hope that our present paper can shed some light on this.

5.2 Perfect DAHA modules

A challenge is an interpretation of DAHA superpolynomials for arbitrary algebraic knots, similar to that in [17] for torus ones via perfect modules of rational DAHA. See also Theorem 9.5 from [22]. There are some obstacles.

One can use here the classification results from [41; 40; 43] in terms of K–theory or homology (for the rational and trigonometric DAHA) of *Iwahoric Springer fibers*. See paper [41] for general theory, including a comprehensive analysis in type A.

Following Gorsky, let us consider the root system $A_{\mathsf{s}-1}$ for a given torus knot $T(\mathsf{r},\mathsf{s})$ (so we adjust its rank to the knot) and take $e=e_1+\cdots+e_{\mathsf{s}-1}+z^{\mathsf{r}}f_\theta$ in the notations from Section 4.1 from [40]. Then $H_*(\mathcal{Y}_e)$ for the corresponding Iwahoric Springer fiber $\mathcal{Y}_e$ can be supplied with a natural structure of the *perfect* $\mathcal{HH}$–module for $t=q^{-\mathsf{r}/\mathsf{s}}$ of dimension $\mathsf{r}^{\mathsf{s}-1}$, which is a very explicit quotient of the polynomial representation of rational DAHA. This module becomes simpler with q,t (via K–theory), but the grading then will be missing.

The coefficients of a^m in the superpolynomial are associated with isotypic components of this $\mathcal{HH}$–module for the wedge m^{th} powers of the standard $\mathsf{s}-1$ dimensional representation of $W=\boldsymbol{S}_{\mathsf{s}}$. This construction is quite combinatorial; Gorsky relates a^m to m–sets of *marks*, which are corners in Dyck paths below the diagonal in the $\mathsf{s}\times\mathsf{r}$–rectangle. In the absence of marks (when $a=0$), the space of W–coinvariants is the image of the projection of $H_*(\mathcal{Y}_e)$ onto $H_*(\mathcal{X}_e)$ for the affine Springer fiber $\mathcal{X}_e$ (see also Section 5.3 below).

The number of marks corresponds to m in our $J^m_{\mathcal{R}}$. The latter is a special subvariety of the *parahoric Springer fiber* of full m–subflags starting from the top, defined as the space of flags of $\mathcal{R}$–modules $\{\mathcal{M}_i \mid i=0,\dots,m\}$, where $\mathfrak{m}\mathcal{M}_m\subset\mathcal{M}_i$ and $\dim_{\mathbb{C}}(\mathcal{M}_{i+1}/\mathcal{M}_i)=1$; here $\mathfrak{m}\subset\mathcal{R}$ is the maximal ideal. They can be extended to full *periodic* flags via the forgetful map from $\mathcal{Y}_e$ to the parahoric one.

This construction involves the root system $A_{\mathsf{s}-1}$, and the topological $\mathsf{r}\leftrightarrow\mathsf{s}$ symmetry of the superpolynomial generally becomes far from obvious. Our geometric superpolynomials are defined directly in terms of the singularity and are manifestly $\mathsf{r}\leftrightarrow\mathsf{s}$–symmetric.

Furthermore, the (finite-dimensional) spaces $H_*(\mathcal{Y}_e)$ for arbitrary nil-elliptic (anisotropic) e are not generally related to any DAHA-modules unless in the torus case. These spaces are needed for general unibranch plane curve singularities, but their

dimensions and other features are very different from those of the finite-dimensional modules. It is not impossible that they can appear as some "remarkable" subspaces in infinite-dimensional $\mathcal{HH}$–modules, but this is questionable.

5.3 Affine Springer fibers

Conjecture 2.5 leads to the following connection with the orbital integrals from the fundamental lemma (in the geometric setting). We will mainly follow Section 2.4 from [44]. The example from Section 2.4.2 there is for the spectral curve $y^r = x^s$ [30], which can be actually generalized to any (germs of) plane curve singularities $\mathcal{C}$ such that the Jacobian factor $J_{\mathcal{R}}$ at $(0,0)$ has the Piontkowski-type decomposition with affine cells only.

The following discussion will be mostly restricted to type A in the anisotropic case. Note that fractional ideals of $\mathcal{R}$ of *degree zero* are considered in [44] (and in the fundamental lemma) in the definition of the compactified Jacobians.

Let $1/t = p^\ell$ be the cardinality of a finite field $\mathbb{F}$. The choice $t = 1/|\mathbb{F}|$ is standard when connecting the q,t–theory of spherical functions with the p–adic theory (see eg [5]), though using q instead of $1/t$ is equally fine due to the superduality of the DAHA superpolynomials.

Because we stick to the unibranch case, $\gamma \in \mathfrak{g}(\mathbb{F}((x)))$ is assumed *nil-elliptic*: no split tori over a local field $\mathbb{F}((x))$ in the stabilizer G_γ ($\gamma = e$ was used above). The affine Springer fiber $\mathcal{X}_\gamma$ is then formed by the classes of g in the *affine Grassmannian* $G(\mathbb{F}((x)))/G(\mathbb{F}[\![x]\!])$ over the field $\mathbb{F}$ such that $\mathrm{Ad}_g^{-1}(\gamma) \in \mathfrak{g}(\mathbb{F}[\![x]\!])$. Here $\mathrm{Lie}(G) = \mathfrak{g}$; see eg [44]. Yun denotes the corresponding *parahoric* ones by $\mathcal{X}_{P,\gamma}$, which contain our flagged Jacobian factors (though we do not need any $\mathfrak{g}, G$). The Iwahoric Springer fibers are for $P = I$ in his notation.

The affine Springer fiber $\mathcal{X}_\gamma$ can be naturally identified with the compactified Jacobians for rational curves with the local ring $\mathcal{R}$ at its (unique) singularity; see eg [29; 4]. A general construction is from [4]. Let $\mathscr{G}$ be a *factorizable* Lie group schemes over a smooth projective curve E, defined by the conditions $H^1(E, \mathrm{Lie}(\mathscr{G})) = \{0\} = H^0(E, \mathrm{Lie}(\mathscr{G}))$ for the corresponding sheaf of Lie algebras $\mathrm{Lie}(\mathscr{G}) \to E$. For a scheme subtorus $\mathscr{T} \subset \mathscr{G}$ such that its generic fiber is a maximal torus, embeddings $f\colon \mathscr{T} \hookrightarrow \mathscr{G}$ become conjugations over sufficiently general open $U \subset E$: specifically, $f(\xi) = \phi^{-1}\xi\phi$ for $\xi \in H^0(U, \mathscr{T})$ and meromorphic sections $\phi = \phi_U$ of $\mathscr{G}$ over U. Čech cohomology is used.

Assuming that ϕ_i exist for open U_i such that $E = \bigcup_i U_i$, the map

$$\{f\} \ni f \mapsto \{\phi_i \phi_j^{-1} \in H^0(U_i \cap U_j, \mathscr{G})\} \to H^1(E, \mathscr{T})$$

is an isomorphism. Actually, we use this map in the opposite direction, from $H^1(E, \mathscr{T})$ to $\{f\}$. Here $H^1(E, \mathscr{T})$ is the generalized Jacobian of the cover $F \to E$ if $\mathscr{T} = \mathcal{O}_F^* \subset \mathscr{G}$. Note that E can be only $\mathbb{CP}^1$ or an elliptic curve; generalized Prim varieties appear for the latter.

Let F be rational with exactly one singular point that is the whole fiber F_o over some $o \in E = \mathbb{CP}^1$; here $\mathscr{G}_o$ must be G. Then $f(\mathscr{T})$ can be obtained as $g^{-1}\mathscr{T}g$ for *rational sections* g of $\mathscr{G}$ such that g are regular at $E \setminus o$ and $g^{-1}\mathscr{T}g$ are regular at F_o. Thus such g form an open subset in the affine Springer fiber $\mathscr{X}_\gamma$ at o for a sufficiently general section γ of $\mathrm{Lie}(\mathscr{T})$ regular at o. For any fiber F_o, the map $g \mapsto g^{-1}\mathscr{T}g$ goes through the quotient $L_\gamma \setminus \mathscr{X}_\gamma$ for the group L_γ of rational sections of F regular at $F \setminus F_o$; see [44, Theorem 2.9] and [25].

Type A is not necessary here, but we need this in our paper (and anisotropic G_γ). Recall that our *flagged* Jacobian factors deviate from the usual ones. First, we consider only standard $\mathcal{R}$–modules. Second, admissible D–flags $\mathcal{D} = [D_0 \subset D_1 = D_0 \cup \{g_1\} \subset \cdots \subset D_m]$ are subtle; D_i must be D–sets and the ordering $g_1 < g_2 < \cdots < g_m$ is imposed. This ordering and the admissibility are quite nontrivial geometrically.

5.4 Motivic approach

Our construction results in the following generalization of orbital integrals:

$$t^\delta \sum_{\mathcal{D}} q^{m+|D_0|} a^m |J_{\mathcal{R}}^m[\mathcal{D}](\mathbb{F}_{1/t})|$$

in type A (the nil-elliptic case), where the summation is over all admissible D–flags $\mathcal{D}$. Their interpretation as "natural" orbital integrals is one of the main challenges triggered by this work. It seems doable due to an entirely geometric nature of our approach. Let us also mention here potential adding colors to our construction, another challenge for us and the specialists in orbital integrals.

For sufficiently general prime p, there is solid evidence that such sums coincide with the DAHA superpolynomials $\mathcal{H}_{\mathcal{C}}(\square; q, t, a)$ associated with the singularity $\mathcal{C}$, which is stated in (2-7). For $q = 1$ and $a = 0$, we arrive at *orbital integrals* O_γ. They are expected to be $|\mathbb{F}|$–integral and positive, which matches our main conjecture.

It is not impossible that all classes $[J^m_{\mathcal{R}}[d]]$ in $K_0(\mathrm{Sch}/\mathbb{C})$ are sums of classes $[\mathbb{A}^1]^N$ over $\mathbb{Z}_+$, the strongest possible (motivic) assumption. This would match known and conjectured properties of $\mathcal{H}_{\mathcal{C}}$. It is possible that $J^m_{\mathcal{R}}[d]$ are always paved by affine spaces (even when $J^m_{\mathcal{R}}[\mathcal{D}]$ are not all affine); see Appendix A in the online supplement.

If (2-7) holds, then DAHA superpolynomials provide all virtual Hodge numbers of $J^m_{\mathcal{R}}[d]$. This results from part (3) of Theorem 1 from the appendix by N M Katz in [24]. Following it, let $E(X;x,y)=\sum_{r,s}e_{r,s}x^r y^s$ for a separated scheme $X/\mathbb{C}$ of finite type and *virtual Hodge numbers* $e_{r,s}=\sum_i(-1)^i h^{r,s}(\mathrm{gr}_W^{r+s}(H^i_c(X^{\mathrm{an}},\mathbb{C})))$.

Then (2-7) gives that $X=J^m_{\mathcal{R}}[d]$ is *strongly polynomial-count* and the following formula holds:

$$(5\text{-}1)\quad (xy)^{\delta}\mathcal{H}_{\vec{r},\vec{s}}(\square;q,1/(xy),a)=\sum_{d,m}E(J^m_{\mathcal{R}}[d];x,y)q^{d+m}a^m$$
$$=\sum_{d,m,i,r}(-1)^i q^{d+m}a^m(xy)^r\,\mathrm{rk}(\mathrm{gr}^{2r}_W(H^i_c(J^m_{\mathcal{R}}[d],\mathbb{C}))),$$

which directly links the DAHA superpolynomials to the weight filtration in $H^i_c(X^{\mathrm{an}},\mathbb{C})$. We use that $\mathcal{R}$ can be assumed over $\mathbb{Z}$ and then $J^m_{\mathcal{R}}[d]$ are defined over a localization of $\mathbb{Z}$ by finitely many primes.Thus $e_{r,s}=0$ for $r\neq s$ (ie this is a Tate–Hodge structure) and the right-hand side in (2-7) can be replaced by the following $\mathcal{H}^{\mathrm{wt}}(q,t,a)$:

$$(5\text{-}2)\qquad \mathcal{H}_{\vec{r},\vec{s}}(\square;q,t,a)=\sum_{d,m,i,r}(-1)^i q^{d+m}t^{\delta-r}a^m\,\mathrm{rk}(\mathrm{gr}^{2r}_W(H^i_c(J^m_{\mathcal{R}}[d],\mathbb{C}))).$$

Conclusion In spite of some similarity between (5-2) and [22, Theorem 9.5], we do not see how they can be connected. First, $t_{\mathrm{st}}^2=q/t$ is used there for the weight filtration instead of t in (5-2). Second, our construction does not require affine Springer fibers and picking the corresponding *isotypic* components in their homology. Lie groups do not appear in our approach; in a sense, this corresponds to the "endoscopy part" of the fundamental lemma. Third, related to the second, our admissible flags are new and different from those in [22]. Fourth, our approach is fully computational and we calculate well beyond torus knots and the series in [38]; the examples provided in [36] were only for some simple torus knots. In spite of these differences, the ORS conjecture can be still compatible with our one, but this is not clear.

Conceptually, *nested Hilbert schemes* from [36] and similar objects in related areas of geometry/topology and physics result in some *infinite* Poincaré series. Generally the "ultimate" problem is to transform them into polynomials or some *finite* expressions

so that the resulting coefficients are positive integers. The latter positivity is generally much more subtle than the positivity (if any) of the coefficients of the initial series. In this paper, we provide the conjectural geometric interpretation for the DAHA superpolynomials via the *Jacobian factors* instead of Hilbert schemes, which is therefore ultimate in the sense above. There is recent progress with the geometric interpretation of the positivity conjecture for DAHA superpolynomials colored by symmetric or wedge powers, but so far not with arbitrary rectangles.

We note that not many formulas are known for the stable KhR–polynomials (the celebrated Khovanov polynomials for $\mathfrak{sl}_2$ are exceptional). They are mostly for $T(2m+1,2)$ and no formulas are known for iterated nontorus knots. Thus checking our geometric superpolynomials versus the DAHA ones is actually the only way for iterated torus knots; though see [8] for some conditional verifications of our topological conjecture using the reduction to the Khovanov polynomials.

We do not pay any special attention to torus knots in our work. For such knots, we generalize the approach from [19; 20], where $a=0$. Our usage of *standard modules* is the key; this is not fully understood geometrically, but we already have some motivic interpretation of complete geometric superpolynomials (with all three parameters). We note that even for torus knots, our *flagged Jacobian factors* are new. The conditions on the corresponding Γ–modules Δ in their definition are not clear by now from the viewpoint of usual flagged constructions.

We plan to approach arbitrary *colored* algebraic *links* and possibly reach arbitrary root systems our further papers. It is a must if we want to realize the potential of DAHA in full and for connections with p–adic orbital integrals. Also, there is a realistic program of justifying our conjecture (related to [31]). It requires knowing the behavior of our geometric superpolynomials under monoidal-type transformations of the corresponding singularities. All of them must be considered here, not only those for torus knots.

Acknowledgements We thank Evgeny Gorsky, Anatoly Libgober, Davesh Maulik and Andras Szenes for useful discussions. The referee's suggestions were very helpful. We give special thanks to David Kazhdan. Cherednik thanks University of Geneva for the invitation and hospitality. Cherednik was partially supported by NSF grant DMS–1363138.

References

[1] **M Aganagic**, **S Shakirov**, *Refined Chern–Simons theory and knot homology*, from "String-Math 2011" (J Block, J Distler, R Donagi, E Sharpe, editors), Proc. Sympos. Pure Math. 85, Amer. Math. Soc., Providence, RI (2012) 3–31 MR

[2] **A Beauville**, *Counting rational curves on $K3$ surfaces*, Duke Math. J. 97 (1999) 99–108 MR

[3] **N Bourbaki**, *Éléments de mathématique, XXXIV: Groupes et algèbres de Lie, Chapitres 4–6*, Actualités Scientifiques et Industrielles 1337, Hermann, Paris (1968) MR

[4] **I V Cherednik**, *On the group-theoretical interpretation of Baker functions and τ–functions*, Uspekhi Mat. Nauk 38 (1983) 133–134 MR In Russian; translated in Russ. Math. Surv. 6 (1983) 113–114

[5] **I Cherednik**, *Double affine Hecke algebras*, London Math. Soc. Lec. Note Ser. 319, Cambridge Univ. Press (2005) MR

[6] **I Cherednik**, *Jones polynomials of torus knots via DAHA*, Int. Math. Res. Not. 2013 (2013) 5366–5425 MR

[7] **I Cherednik**, *DAHA-Jones polynomials of torus knots*, Selecta Math. 22 (2016) 1013–1053 MR

[8] **I Cherednik**, **I Danilenko**, *DAHA and iterated torus knots*, Algebr. Geom. Topol. 16 (2016) 843–898 MR

[9] **I Cherednik**, **I Danilenko**, *DAHA approach to iterated torus links*, from "Categorification in geometry, topology, and physics" (A Beliakova, A D Lauda, editors), Contemp. Math. 684, Amer. Math. Soc., Providence, RI (2017) 159–267 MR

[10] **V Cherednik**, **R Elliot**, *Refined composite invariants of torus knots via DAHA*, Ann. Fac. Sci. Toulouse Math. 25 (2016) 433–471 MR

[11] **N M Dunfield**, **S Gukov**, **J Rasmussen**, *The superpolynomial for knot homologies*, Experiment. Math. 15 (2006) 129–159 MR

[12] **D Eisenbud**, **W Neumann**, *Three-dimensional link theory and invariants of plane curve singularities*, Annals of Mathematics Studies 110, Princeton Univ. Press (1985) MR

[13] **P Etingof**, **E Gorsky**, **I Losev**, *Representations of rational Cherednik algebras with minimal support and torus knots*, Adv. Math. 277 (2015) 124–180 MR

[14] **H Fuji**, **S Gukov**, **P Sułkowski**, *Super- A–polynomial for knots and BPS states*, Nuclear Phys. B 867 (2013) 506–546 MR

[15] **M Goresky**, **R Kottwitz**, **R MacPherson**, *Purity of equivalued affine Springer fibers*, Represent. Theory 10 (2006) 130–146 MR

[16] **E Gorsky**, *q,t–Catalan numbers and knot homology*, from "Zeta functions in algebra and geometry" (A Campillo, G Cardona, A Melle-Hernández, W Veys, editors), Contemp. Math. 566, Amer. Math. Soc., Providence, RI (2012) 213–232 MR

[17] **E Gorsky**, *Arc spaces and DAHA representations*, Selecta Math. 19 (2013) 125–140 MR

[18] **E Gorsky**, **S Gukov**, **M Stosic**, *Quadruply-graded colored homology of knots*, preprint (2013) arXiv

[19] **E Gorsky**, **M Mazin**, *Compactified Jacobians and q,t–Catalan numbers, I*, J. Combin. Theory Ser. A 120 (2013) 49–63 MR

[20] **E Gorsky**, **M Mazin**, *Compactified Jacobians and q,t–Catalan numbers, II*, J. Algebraic Combin. 39 (2014) 153–186 MR

[21] **E Gorsky**, **A Neguţ**, *Refined knot invariants and Hilbert schemes*, J. Math. Pures Appl. 104 (2015) 403–435 MR

[22] **E Gorsky**, **A Oblomkov**, **J Rasmussen**, **V Shende**, *Torus knots and the rational DAHA*, Duke Math. J. 163 (2014) 2709–2794 MR

[23] **G-M Greuel**, **G Pfister**, *Moduli spaces for torsion free modules on curve singularities, I*, J. Algebraic Geom. 2 (1993) 81–135 MR

[24] **T Hausel**, **F Rodriguez-Villegas**, *Mixed Hodge polynomials of character varieties*, Invent. Math. 174 (2008) 555–624 MR With an appendix by N M Katz

[25] **D Kazhdan**, **G Lusztig**, *Fixed point varieties on affine flag manifolds*, Israel J. Math. 62 (1988) 129–168 MR

[26] **M Khovanov**, *Triply-graded link homology and Hochschild homology of Soergel bimodules*, Internat. J. Math. 18 (2007) 869–885 MR

[27] **M Khovanov**, **L Rozansky**, *Matrix factorizations and link homology*, Fund. Math. 199 (2008) 1–91 MR

[28] **M Khovanov**, **L Rozansky**, *Matrix factorizations and link homology, II*, Geom. Topol. 12 (2008) 1387–1425 MR

[29] **G Laumon**, *Fibres de Springer et jacobiennes compactifiées*, from "Algebraic geometry and number theory" (V Ginzburg, editor), Progr. Math. 253, Birkhäuser, Boston (2006) 515–563 MR

[30] **G Lusztig**, **J M Smelt**, *Fixed point varieties on the space of lattices*, Bull. London Math. Soc. 23 (1991) 213–218 MR

[31] **D Maulik**, *Stable pairs and the HOMFLY polynomial*, Invent. Math. 204 (2016) 787–831 MR

[32] **D Maulik**, **Z Yun**, *Macdonald formula for curves with planar singularities*, J. Reine Angew. Math. 694 (2014) 27–48 MR

[33] **A Mellit**, *Toric braids and (m,n)–parking functions*, preprint (2016) arXiv

[34] **L Migliorini**, **V Shende**, *A support theorem for Hilbert schemes of planar curves*, J. Eur. Math. Soc. 15 (2013) 2353–2367 MR

[35] **H R Morton**, *The coloured Jones function and Alexander polynomial for torus knots*, Math. Proc. Cambridge Philos. Soc. 117 (1995) 129–135 MR

[36] **A Oblomkov**, **J Rasmussen**, **V Shende**, *The Hilbert scheme of a plane curve singularity and the HOMFLY homology of its link*, preprint (2012) arXiv

[37] **A Oblomkov**, **V Shende**, *The Hilbert scheme of a plane curve singularity and the HOMFLY polynomial of its link*, Duke Math. J. 161 (2012) 1277–1303 MR

[38] **J Piontkowski**, *Topology of the compactified Jacobians of singular curves*, Math. Z. 255 (2007) 195–226 MR

[39] **J Rasmussen**, *Some differentials on Khovanov–Rozansky homology*, Geom. Topol. 19 (2015) 3031–3104 MR

[40] **M Varagnolo**, **E Vasserot**, *Finite-dimensional representations of DAHA and affine Springer fibers: the spherical case*, Duke Math. J. 147 (2009) 439–540 MR

[41] **E Vasserot**, *Induced and simple modules of double affine Hecke algebras*, Duke Math. J. 126 (2005) 251–323 MR

[42] **B Webster**, **G Williamson**, *A geometric construction of colored HOMFLYPT homology*, Geom. Topol. 21 (2017) 2557–2600 MR

[43] **Z Yun**, *Global Springer theory*, Adv. Math. 228 (2011) 266–328 MR

[44] **Z Yun**, *Lectures on Springer theories and orbital integrals*, preprint (2016) arXiv

[45] **O Zariski**, *Le problème des modules pour les branches planes*, 2nd edition, Hermann, Paris (1986) MR

Mathematics Department, University of North Carolina
Chapel Hill, NC, United States

chered@email.unc.edu, iphilipp@live.unc.edu

Received: 28 October 2016 Revised: 15 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 387–408

Inertia groups of high-dimensional complex projective spaces

SAMIK BASU
RAMESH KASILINGAM

For a complex projective space the inertia group, the homotopy inertia group and the concordance inertia group are isomorphic. In complex dimension $4n+1$, these groups are related to computations in stable cohomotopy. Using stable homotopy theory, we make explicit computations to show that the inertia group is nontrivial in many cases. In complex dimension 9, we deduce some results on geometric structures on homotopy complex projective spaces and complex hyperbolic manifolds.

57R55, 57R60; 55P25, 55P42

1 Introduction

The study of manifolds in differential topology presents itself through four different classes of equivalence: homotopy equivalence, homeomorphism, PL homeomorphism and diffeomorphism. The classification of manifolds up to these equivalences is a fundamental question in geometry and topology.

One of the first results in this subject is that of Milnor [22] that there exist smooth manifolds which are homeomorphic to S^7 but not diffeomorphic. This raises the possibility of nondiffeomorphic smooth structures for a given a topological manifold. We refer to these as inequivalent smoothings. For example, the 7–sphere has 28 inequivalent smoothings.[1]

A smooth manifold homeomorphic to S^m is known as a smooth homotopy m–sphere.[2] The existence of smooth homotopy m–spheres was studied in the amazing work of Kervaire and Milnor [16]. The set of diffeomorphism classes of smooth homotopy spheres Θ_m for $m \geq 5$ forms a group under the operation of connected sum. It was shown that there exist exotic spheres in a vast majority of dimensions but also that in each dimension there are only finitely many. The proof of the result points to a curious

[1] This actually occurs as a consequence of the results of Kervaire and Milnor [16].

[2] We know that if $m \geq 5$ as an m–manifold homotopy equivalent to S^m is actually homeomorphic to S^m.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.387

connection to the stable homotopy groups of spheres (denoted by π_n^s for $n \geq 0$) whose values are mysterious but are computable using algebraic techniques, especially in low dimensions.

A possible way to change smooth structure on a smooth manifold M^m, without changing its homeomorphism type, is to take its connected sum $M^m \# \Sigma^m$ with a smooth homotopy sphere Σ^m. This induces a group action of Θ_m on the set of smooth structures on the topological manifold M. The collection of smooth homotopy spheres Σ^m which admit a diffeomorphism $M^m \# \Sigma^m \to M^m$ forms a subgroup $I(M)$ of Θ_m, called the inertia group of M^m.

The calculation of $I(M)$ for an arbitrary manifold M has proven to be a hard problem in general but there are results in certain cases. Tamura [27] constructed explicit nontrivial elements in the inertia group for certain 3–sphere bundles over S^4. Examples with nontrivial inertia group are also constructed by Brown and Steer [4]. In every dimension m, Winkelnkemper [30] proved that there exists a manifold M^m such that $I(M^m) = \Theta_m$. Levine [21] constructed certain nontrivial elements in the inertia group for many manifolds, most notably for simply connected nonspin manifolds in dimensions $8n+2$.

There are certain cases when the inertia group is 0. It was proved by Schultz [26] that $I(S^p \times S^q) = 0$ when $p+q \geq 5$. Kawakubo [14] proved that $I(\mathbb{C}P^m) = 0$ for $m \leq 8$. Limitations on the size of the inertia group have been given by Wall [29], Browder [3], Kosiński [19] and Novikov [24]. There is no systematic approach for computing the inertia groups in general, and many problems are open. In this paper we are interested in the problem: *What are the inertia groups of* $\mathbb{C}P^m$ *if* $m \geq 9$*?*

Analogous to the inertia group, for a manifold M one may define the homotopy inertia group $I_h(M)$ and the concordance inertia group $I_c(M)$. $I_h(M)$ (resp. $I_c(M)$) consists of those $\Sigma \in I(M)$ for which the diffeomorphism $M \# \Sigma \cong M$ is homotopic (resp. concordant) to the identity. Kasilingam [12] proved that for a complex projective space all these groups are the same.

The concordance inertia groups may be understood using homotopy theory. In dimensions $8n+2$, there is a $\mathbb{Z}/2$ summand of π^s_{8n+2} generated by μ_{8n+2} that corresponds to an exotic sphere $\Sigma^{8n+2} \in \Theta_{8n+2}$. Farrell and Jones [8] proved that $\mathbb{C}P^{4n+1} \# \Sigma^{8n+2}$ is not concordant to $\mathbb{C}P^{4n+1}$ using certain relations in KO^*, the cohomology theory induced by real K–theory. Hence this also implies that the element Σ^{8n+2} does not lie in the inertia group.

In this paper we try to compute the inertia groups of $\mathbb{C}P^{4n+1}$ using stable cohomotopy. We prove:

Theorem A (a) $I(\mathbb{C}P^9) \cong \mathbb{Z}/2$ *or* $\mathbb{Z}/4$ *as a subgroup of* $\Theta_{18} = \mathbb{Z}/2 \oplus \mathbb{Z}/8$.
(b) $I(\mathbb{C}P^{13}) \supseteq \mathbb{Z}/2$.

(See Theorem 3.11.) This is the first example of a nontrivial inertia group among the inertia groups of complex projective spaces.[3] These provide examples of exotic spheres whose α–invariant is 0 and where connected sum with projective spaces does not change the smooth structure. For the projective space $\mathbb{C}P^9$, we also get examples of exotic spheres with α–invariant 0 which change the smooth structure. The techniques involved are explicit calculations using Spanier–Whitehead duality and the knowledge of the stable homotopy groups of spheres in low dimensions at the prime 2.

We also make computations at odd primes p, making use of the nontriviality of certain elements in the stable homotopy groups proved by Lee [20] and deduce:

Theorem B *There are infinitely many values of* n *for which there exist nontrivial elements in the inertia group of* $\mathbb{C}P^{4n+1}$.

(For a more accurate statement see Theorem 3.11.)

Computations of the inertia groups carry with them a number of geometric applications. Using previously known results on the triviality of the inertia group $I(\mathbb{C}P^m)$ for $m \leq 8$, it is possible to classify, up to diffeomorphism, all closed smooth manifolds homeomorphic to the complex projective n–space for $n = 3$ and 4; see Kasilingam [13]. Kasilingam [12] showed that a way to generate exotic spheres which are not in the inertia group[4] in dimensions $8n + 2$ is by considering exotic spheres with a nontrivial α–invariant.[5] The exotic spheres outside the inertia group for complex projective spaces have consequences for complex hyperbolic manifolds; see Farrell and Jones [8].

In this paper, from the inertia group of $\mathbb{C}P^9$, we deduce examples of three inequivalent smooth structures of $\mathbb{C}P^9$, different from the standard one, such that one admits a metric of nonnegative scalar curvature and the other two do not (see Theorem 4.3). The example relies on the construction of an element outside the inertia group whose

[3]The reader may observe that $\mathbb{C}P^{4n+1}$ is a spin manifold.
[4]These are defined to be Farrell–Jones spheres.
[5]These are defined as Hitchin spheres.

α–invariant is 0. These examples extend the results in [12] to dimension 18. Following this example, we construct examples of closed negatively curved Riemannian 18–manifolds which are homeomorphic but not diffeomorphic to complex hyperbolic manifolds (see Theorem 4.4).

Organization of the paper In Section 2, we introduce some preliminaries on the inertia group $I(\mathbb{C}P^m)$ and prove a result relating this to a computation in stable cohomotopy for dimension $m = 4n + 1$. In Section 3, we make some computations in stable homotopy related to the above question and prove the nontriviality of $I(\mathbb{C}P^{4n+1})$ for the different values of n. Finally, Section 4 contains some geometric applications.

Notation Denote by $O = \operatorname{colim}_n O(n)$, $\mathrm{Top} = \operatorname{colim}_n \mathrm{Top}(n)$ and $G = \operatorname{colim}_n G(n)$ the direct limit of the groups of orthogonal transformations, homeomorphisms and homotopy equivalences, respectively. In this paper all manifolds will be closed, smooth, oriented and connected, and all homeomorphisms and diffeomorphisms are assumed to preserve orientation, unless otherwise stated.

2 Inertia groups of complex projective spaces

In this section we recall some basic facts about inertia groups, specializing to the case of complex projective spaces, and provide the background of the arguments in the rest of the paper. We start by recalling some terminology from [16]:

Definition 2.1 (a) A homotopy m–sphere Σ^m is an oriented, smooth, closed manifold homotopy equivalent to the standard unit sphere S^m in $\mathbb{R}^{m+1}$.

(b) A homotopy m–sphere Σ^m is said to be exotic if it is not diffeomorphic to S^m.

(c) Two homotopy m–spheres Σ_1^m and Σ_2^m are said to be equivalent if there exists an orientation-preserving diffeomorphism $f\colon \Sigma_1^m \to \Sigma_2^m$.

The set of equivalence classes of homotopy m–spheres is denoted by Θ_m. The equivalence class of Σ^m is denoted by $[\Sigma^m]$. When $m \geq 5$, Θ_m forms an abelian group with group operation given by connected sum # and the zero element represented by the equivalence class of S^m. M Kervaire and J Milnor [16] showed that each Θ_m is a finite group.

Definition 2.2 Let M be a topological manifold. Let (N, f) be a pair consisting of a smooth manifold N together with a homeomorphism $f\colon N \to M$. Two such pairs (N_1, f_1) and (N_2, f_2) are concordant provided there exists a diffeomorphism

$g\colon N_1 \to N_2$ such that the composition $f_2 \circ g$ is topologically concordant to f_1, ie there exists a homeomorphism $F\colon N_1 \times [0,1] \to M \times [0,1]$ such that $F|_{N_1\times 0} = f_1$ and $F|_{N_1\times 1} = f_2 \circ g$. The set of all such concordance classes is denoted by $\mathcal{C}(M)$.

Note that there is a homeomorphism $h\colon M^m \# \Sigma^m \to M^m$ for $m \geq 5$ which is the inclusion map outside of homotopy sphere Σ^m and well defined up to topological concordance. We will denote the class of $(M^m \# \Sigma^m, h)$ in $\mathcal{C}(M)$ by $[M^m \# \Sigma^m]$. (Note that $[M^m \# S^m]$ is the class of (M^m, id_{M^m}).)

Definition 2.3 Let M^m be a closed smooth m–dimensional manifold. The inertia group $I(M) \subset \Theta_m$ is defined as the set of $\Sigma \in \Theta_m$ for which there exists a diffeomorphism $\phi\colon M \to M \# \Sigma$.

The homotopy inertia group $I_h(M)$ is defined as the set of all $\Sigma \in I(M)$ such that there exists a diffeomorphism $M \to M \# \Sigma$ which is homotopic to id: $M \to M \# \Sigma$.

The concordance inertia group $I_c(M)$ is defined as the set of all $\Sigma \in I_h(M)$ such that $M \# \Sigma$ is concordant to M.

We recall the following theorem about complex projective spaces:

Theorem 2.4 [12, Theorem 4.2] *For $n \geq 1$, $I_c(\mathbb{C}P^n) = I_h(\mathbb{C}P^n) = I(\mathbb{C}P^n)$.*

Next we recall a reformulation of inertia groups via homotopy theory. Let $f_M\colon M^m \to S^m$ be a degree 1 map. Note that f_M is well defined up to homotopy. Composition with f_M defines a homomorphism

$$f_M^*\colon [S^m, \mathrm{Top}/O] \to [M^m, \mathrm{Top}/O],$$

and in terms of the identifications

$$\Theta_m = [S^m, \mathrm{Top}/O] \quad \text{and} \quad \mathcal{C}(M^m) = [M^m, \mathrm{Top}/O]$$

given by [17, pages 25 and 194], f_M^* becomes $[\Sigma^m] \mapsto [M^m \# \Sigma^m]$. Therefore the concordance inertia group $I_c(M)$ can be identified with $\mathrm{Ker}(f_M^*)$.

Recall that the based homotopy classes $[X, G]$ can be identified with the 0th stable cohomotopy group $\widetilde{\pi}^0(X)$. We also write f_M^* for the induced homomorphism $[S^m, G] = \widetilde{\pi}^0(S^m) = \pi_m^s \to [M^m, G] = \widetilde{\pi}^0(M^m)$ by a degree 1 map $f_M\colon M^m \to S^m$.

Now recollect some facts from smoothing theory [5]. The natural inclusions of H–spaces $O \subset \mathrm{Top} \subset G$ induce H–space maps $\phi\colon G \to G/O$ and $\psi\colon \mathrm{Top}/O \to G/O$

such that

$$\psi_*\colon \Theta_{8n+2} = \pi_{8n+2}(\mathrm{Top}/O) \to \pi_{8n+2}(G/O)$$

is an isomorphism for $n \geq 1$. The homotopy groups of G are the stable homotopy groups of spheres π^s_m; ie $\pi_m(G) = \pi^s_m$ for $m \geq 1$. For $n \geq 1$,

$$\phi_*\colon \pi^s_{8n+2} \to \pi_{8n+2}(G/O)$$

is an isomorphism.

Theorem 2.5 *Let M^{8n+2} be a closed, smooth $8n+2$–manifold homotopy equivalent to $\mathbb{C}P^{4n+1}$ with $n \geq 1$. Then $I_c(M) = \mathrm{Ker}(\Phi)$, where*

$$\Phi\colon \Theta_{8n+2} \xrightarrow{\psi_*} \pi_{8n+2}(G/O) \xrightarrow{\phi_*^{-1}} \widetilde{\pi}^0(S^{8n+2}) \xrightarrow{f_M^*} \widetilde{\pi}^0(M^{8n+2}).$$

The proof of Theorem 2.5 requires two facts we prove below:

Lemma 2.6 *Let M^{2m} be a closed smooth $2m$–manifold homotopy equivalent to $\mathbb{C}P^m$ with $m \geq 1$. Then the homomorphism $\psi_*\colon [M^{2m}, \mathrm{Top}/O] \mapsto [M^{2m}, G/O]$ is monic.*

Proof Consider the Barratt–Puppe sequence for the inclusion $i\colon \mathbb{C}P^{m-1} \hookrightarrow \mathbb{C}P^m$, which induces the exact sequence on taking homotopy classes $[-, \Omega(G/\mathrm{Top})]$

$$\cdots \to [S\mathbb{C}P^{m-1}, \Omega(G/\mathrm{Top})] \to [S^{2m}, \Omega(G/\mathrm{Top})] \xrightarrow{f^*_{\mathbb{C}P^m}} [\mathbb{C}P^m, \Omega(G/\mathrm{Top})] \xrightarrow{i^*} [\mathbb{C}P^{m-1}, \Omega(G/\mathrm{Top})] \to \cdots,$$

and, by identifying

$$[S^{2m}, \Omega(G/\mathrm{Top})] = [S^{2m+1}, G/\mathrm{Top}] = L_{2m+1}(e) = 0,$$

where $L_k(e)$ is the simply connected surgery obstruction group, and

$$[\mathbb{C}P^1, \Omega(G/\mathrm{Top})] = 0,$$

we can prove that $[\mathbb{C}P^m, \Omega(G/\mathrm{Top})] = 0$ for all m. Now consider the long exact sequence associated to the fibration $\mathrm{Top}/O \to G/O \to G/\mathrm{Top}$,

$$\cdots \to [M^{2m}, \Omega(G/\mathrm{Top})] \to [M^{2m}, \mathrm{Top}/O] \xrightarrow{\psi_*} [M^{2m}, G/O] \to [M^{2m}, G/\mathrm{Top}],$$

and using the fact that $[M^{2m}, \Omega(G/\mathrm{Top})] = [\mathbb{C}P^m, \Omega(G/\mathrm{Top})] = 0$, we have that the homomorphism ψ_* is monic. □

Lemma 2.7 *Let M^{2m} be a closed, smooth $2m$–manifold homotopy equivalent to $\mathbb{C}P^m$ with $m \geq 1$. Then the homomorphism $\phi_*\colon [M^{2m}, G] \to [M^{2m}, G/O]$ is monic.*

Proof Since M^{2m} is homotopy equivalent to $\mathbb{C}P^m$, let $g\colon M^{2m} \to \mathbb{C}P^m$ be a homotopy equivalence. The induced map $g^*\colon [\mathbb{C}P^m, -] \to [M^{2m}, -]$ fits into the following commutative diagram:

$$\begin{array}{ccc} [\mathbb{C}P^m, G] & \xrightarrow{\phi_*} & [\mathbb{C}P^m, G/O] \\ \cong\big\downarrow g^* & & \cong\big\downarrow g^* \\ [M^{2m}, G] & \xrightarrow{\phi_*} & [M^{2m}, G/O] \end{array}$$

Brumfiel [6, page 77] has shown that

$$\phi_*\colon [\mathbb{C}P^m, G] \to [\mathbb{C}P^m, G/O]$$

is monic for all $m \geq 1$. This implies that the homomorphism $\phi_*\colon [M^{2m}, G] \to [M^{2m}, G/O]$ is monic. □

Proof of Theorem 2.5 Consider the following commutative diagram:

$$\begin{array}{ccc} [S^{2m}, \mathrm{Top}/O] = \Theta_{2m} & \xrightarrow{f^*_{M^{2m}}} & [M^{2m}, \mathrm{Top}/O] = \mathcal{C}(M^{2m}) \\ \big\downarrow \psi_* & & \big\downarrow \psi_* \\ [S^{2m}, G/O] & \xrightarrow{f^*_{M^{2m}}} & [M^{2m}, G/O] \\ \phi_*\big\uparrow & & \phi_*\big\uparrow \\ [S^{2m}, G] = \pi^s_{2m} & \xrightarrow{f^*_{M^{2m}}} & [M^{2m}, G] \end{array}$$

Recall that the concordance class $[M^{2m} \# \Sigma] \in [M^{2m}, \mathrm{Top}/O]$ of $M^{2m} \# \Sigma$ is $f^*_{M^{2m}}([\Sigma])$ when $m > 2$, and that $[M^{2m}] = [M^{2m} \# S^{2m}]$ is the zero element of this group. Now Lemmas 2.6 and 2.7, used in conjunction with a simple diagram chase for $m = 4n + 1$, show that $I_c(M^{8n+2}) = \mathrm{Ker}(\Phi = f^*_M \circ \phi_*^{-1} \circ \psi_*)$, thus proving Theorem 2.5. □

Identifying the group Θ_{8n+2} with $\widetilde{\pi}^0(S^{8n+2})$ in view of Theorem 2.5 we consider the following question, related to inertia groups:

Problem 2.8 *What is the kernel of $f^*_{\mathbb{C}P^{4n+1}}\colon \widetilde{\pi}^0(S^{8n+2}) \to \widetilde{\pi}^0(\mathbb{C}P^{4n+1})$?*

We explore some cases of this question in the next section.

3 Computations in stable homotopy

In this section we make computations relating to Problem 2.8. We work in the category of spectra. Throughout this section we use the notation X for both the space and the spectrum $\Sigma^\infty X$ and the notation $\{X, Y\}$ for the stable homotopy classes of maps from X to Y.

Recall [7; 11] that there are models of the category of spectra which are a closed symmetric monoidal category with monoidal structure given by the smash product $\wedge$ and the mapping spectrum denoted by $F(-,-)$. The sphere spectrum S^0 is the unit of the monoidal structure. In this category, for a spectrum X one may form the dual spectrum $DX = F(X, S^0)$. This notion appeared earlier for finite spectra as the Spanier–Whitehead dual of X [2].

One notes that if X is a finite cellular spectrum then so is the dual DX. We briefly recall this construction. Let the cellular structure on X be given by $X = \operatorname{colim}_n X^{(n)}$ such that $X^{(n)}$ is obtained from $X^{(n-1)}$ by attaching a cell of dimension a_n. That is, there are cofibre sequences

$$S^{a_n - 1} \to X^{(n-1)} \to X^{(n)}.$$

The dual structure on DX is given by $DX = \operatorname{colim}_n D(X/X^{(n)})$, with cofibre sequences

$$S^{-a_n - 1} \to D(X/X^{(n)}) \to D(X/X^{(n-1)})$$

obtained by dualizing the cofibre

$$S^{a_n} \to X/X^{(n-1)} \to X/X^{(n)} \to S^{a_n + 1}.$$

Thus, in the colimit $DX = \operatorname{colim}_n D(X/X^{(n)})$, $D(X/X^{(n-1)})$ is obtained from $D(X/X^{(n)})$ by attaching a cell of dimension $-a_n$. Therefore, DX is also a finite cellular spectrum with a $-n$–cell for every n–cell of X.

Note that $\widetilde{\pi}^0(X) \cong \{X, S^0\}$. Note also that $\{X, S^0\} = \pi_0 F(X, S^0) = \pi_0 D(X)$. Therefore, for a map $f\colon X \to Y$ of spectra, the map $\widetilde{\pi}^0(f)\colon \widetilde{\pi}^0(Y) \to \widetilde{\pi}^0(X)$ is equivalent to the map $\pi_0(D(f))\colon \pi_0 D(Y) \to \pi_0 D(X)$ induced by the dual map $D(f)\colon D(Y) \to D(X)$.

For Problem 2.8 we wish to compute $\widetilde{\pi}^0(f)$, where $f\colon \mathbb{C}P^{4n+1} \to S^{8n+2}$ is the usual degree 1 map. Our approach is to compute $\pi_0(D(f))$.

3.1 Computations in dimension 18

We begin by noting that $D\mathbb{C}P^9$ has the filtration

$$S^{-18} = D(\mathbb{C}P^9/\mathbb{C}P^8) \to D(\mathbb{C}P^9/\mathbb{C}P^7) \to \cdots \to D(\mathbb{C}P^9/\mathbb{C}P^0) = D(\mathbb{C}P^9) \tag{3-1}$$

and there are cofibre sequences, for $1 \le k \le 8$,

$$S^{-2k-1} \to D(\mathbb{C}P^9/\mathbb{C}P^k) \to D(\mathbb{C}P^9/\mathbb{C}P^{k-1}) \to S^{-2k}.$$

The degree 1 map $\mathbb{C}P^9 \to S^{18}$ dualises to the inclusion of the bottom cell $S^{-18} \to D(\mathbb{C}P^9)$. Thus we are interested in the question: which elements of $\pi_{18}^s = \pi_0 S^{-18}$ map to 0 under the map above? Recall that the homotopy group $\pi_0 S^{-18} = \pi_{18}^s = \mathbb{Z}/2 \oplus \mathbb{Z}/8$. The $\mathbb{Z}/2$ summand is generated by the element μ_{18} [1]. From [8] it follows that this element maps nontrivially into $\pi_0 D\mathbb{C}P^9$. Therefore the question remains about the other summand, $\mathbb{Z}/8$. In this computation we need formulas for the action of the Steenrod operations on the cohomology of $\mathbb{C}P^n$. We recall them below:

Proposition 3.1 *In $H^*(\mathbb{C}P^n; \mathbb{Z}/2)$ we have the formulas*

$$\mathrm{Sq}^2(x^k) = x^{k+1} \iff k \text{ is odd},$$
$$\mathrm{Sq}^4(x^k) = x^{k+2} \iff k \equiv 2 \text{ or } 3 \pmod 4.$$

Recall that the Steenrod operation Sq^2 detects the Hopf map η, which in our notation is h_1, and Sq^4 detects the map ν, which in our notation is h_2 (modulo $2h_2$). We prove:

Proposition 3.2 *The map $\pi_{18}^s = \pi_0 S^{-18} \to \pi_0(D(\mathbb{C}P^9/\mathbb{C}P^3))$ is injective.*

Proof We compute this map using the filtration (3-1). Note that the group π_{18}^s is 2–torsion so it suffices to work in the 2–local category. We use the notation in [25, Appendix 3.3]. In terms of this notation, π_{18}^s is $\mathbb{Z}/2\{h_1 P^2 h_1\} \oplus \mathbb{Z}/8\{h_2 h_4\}$. It helps to note that the element $h_2 h_4$ is indecomposable in the algebra π_*^s of stable homotopy groups (since the element h_4 supports the differential $d_2(h_4) = h_0 h_3^2$ and all the hidden extensions in the range are written in Corollary 4.4.50).

We start proceeding along the sequence (3-1) with the spectrum $D(\mathbb{C}P^9/\mathbb{C}P^7)$. This fits into a cofibre

$$S^{-17} \to S^{-18} \to D(\mathbb{C}P^9/\mathbb{C}P^7).$$

The map $S^{-17} \to S^{-18} \in \pi_1^s = \mathbb{Z}/2\{h_1\}$. Whether it is nontrivial or not is determined by the action of the Steenrod operation Sq^2 on the cone and hence determined by the

action of Sq^2 on $\mathbb{C}P^9/\mathbb{C}P^7$. Note from Proposition 3.1 that $\mathrm{Sq}^2(x^8) = 0$. Therefore the map is trivial and $D(\mathbb{C}P^9/\mathbb{C}P^7) \simeq S^{-18} \vee S^{-16}$. It follows that the map from $\pi^s_{18} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^7)$ is injective.

The next term is $D(\mathbb{C}P^9/\mathbb{C}P^6)$. We have the cofibre sequence

$$S^{-15} \to D(\mathbb{C}P^9/\mathbb{C}P^7) \to D(\mathbb{C}P^9/\mathbb{C}P^6). \tag{3-2}$$

The map $S^{-15} \to D(\mathbb{C}P^9/\mathbb{C}P^7) \simeq S^{-18} \vee S^{-16}$ is an element of $\pi^s_1 \oplus \pi^s_3 = \mathbb{Z}/2\{h_1\} \oplus \mathbb{Z}/8\{h_2\}$. Note that on $\mathbb{C}P^9/\mathbb{C}P^6$ the Steenrod operations satisfy the formulas $\mathrm{Sq}^2(x^7) = x^8$ and $\mathrm{Sq}^4(x^7) = x^9$. Thus the 16–cell in $\mathbb{C}P^9/\mathbb{C}P^6$ attaches onto the 14–cell by h_1 and the 18–cell attaches via h_2 onto the 14–cell (or some other odd multiple which does not change the computations below). Therefore the map $S^{-15} \to S^{-16} \vee S^{-18}$ is given by (h_1, h_2). On π_0, we have the sequence

$$\pi^s_{15} \xrightarrow{(h_1,h_2)} \pi^s_{16} \oplus \pi^s_{18} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^6).$$

Observe that multiplication by h_2 from π^s_{15} to π^s_{18} is 0. This may be read off from [25, Table A.3.1a] and the fact that h_4, h_0h_4 and $h_0^2h_4$ support nontrivial differentials (see also [18, Figure 5.11]). It follows that the map from $\pi^s_{18} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^6)$ is injective.

The next term in the sequence (3-1) is $D(\mathbb{C}P^9/\mathbb{C}P^5)$ and is formed from $D(\mathbb{C}P^9/\mathbb{C}P^6)$ by the cofibre

$$S^{-13} \to D(\mathbb{C}P^9/\mathbb{C}P^6) \to D(\mathbb{C}P^9/\mathbb{C}P^5).$$

The group $\pi^s_{13} = 0$ and so the map from $\pi^s_{18} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^5)$ is injective.

Next we analyse the cofibre

$$S^{-11} \to D(\mathbb{C}P^9/\mathbb{C}P^5) \to D(\mathbb{C}P^9/\mathbb{C}P^4)$$

and compute the image $\pi^s_{11} = \mathbb{Z}/8\{Ph_2\} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^5)$. Note that the methods above imply that $D(\mathbb{C}P^9/\mathbb{C}P^5)$ is the cofibre

$$S^{-13} \vee S^{-15} \xrightarrow{A} S^{-16} \vee S^{-18} \to D(\mathbb{C}P^9/\mathbb{C}P^5), \tag{3-3}$$

where the matrix A is given by

$$\begin{bmatrix} h_2 & h_1 \\ 0 & h_2 \end{bmatrix}.$$

The map $S^{-11} \to D(\mathbb{C}P^9/\mathbb{C}P^5)$ may be described by a map $S^{-11} \to S^{-12} \vee S^{-14}$ (hence in $\pi_1^s \oplus \pi_3^s = \mathbb{Z}/2\{h_1\} \oplus \mathbb{Z}/8\{h_2\}$) together with a choice of null homotopy after composition by A to $S^{-15} \vee S^{-17}$. The choice of null homotopy lies in $\{S^{-11}, S^{-16} \vee S^{-18}\} = \pi_5^s \oplus \pi_7^s = \mathbb{Z}/16\{h_3\}$.

The map $S^{-11} \to S^{-12} \vee S^{-14}$ is also the attaching map of the -10–cell for the spectrum $D(\mathbb{C}P^7/\mathbb{C}P^4)$. Hence one may try to compute its homotopy class via Steenrod operations. From the formulas in Proposition 3.1 we have $\mathrm{Sq}^2(x^5) = x^6$ and $\mathrm{Sq}^4(x^5) = 0$. So the map onto the -12–cell is h_1 and the map onto the -14–cell is some even multiple of h_2. It follows from the formulas in [23, Proposition 5.2] that the second map is 0.

Now take a class $a \in \pi_{11}^s$, so that a is some multiple of Ph_2. To compute its image onto $\pi_{12}^s \oplus \pi_{14}^s$, one has to multiply by the class $(h_1, 0)$. Notice that $h_1 Ph_2 = 0$. Therefore the image must be zero.

Our case of interest is the image in π_{18}^s. This can be computed using Toda brackets. More precisely, the image $\pi_{11}^s \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^5)$ lies in the image of

$$\pi_0(S^{-16} \vee S^{-18}) \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^5),$$

which is computed as the image of the set $\langle \pi_{11}^s, (h_1, 0), A\rangle$. We obtain the image onto $\pi_{18}^s \subset \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^5)$ as the projection of $\langle \pi_{11}^s, (h_1, 0), A\rangle \in \pi_{16}^s \oplus \pi_{18}^s$ onto π_{18}^s. As $h_2 \cdot \pi_{15}^s$ and $\pi_{11}^s \cdot h_3$ are 0 (from [25, Table A.3.1a]), the indeterminacy of the bracket maps to 0 in π_{18}^s. Now, the -14–cell is the only cell in $D(\mathbb{C}P^9/\mathbb{C}P^5)$ which attaches nontrivially down to the -18–cell and the map from S^{-11} is null-homotopic on the -14–cell, so the above bracket projected down to π_{18}^s must contain 0. Hence the map from $\pi_{18}^s \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^4)$ is injective.

The next term in the sequence (3-1) is $D(\mathbb{C}P^9/\mathbb{C}P^3)$. This fits into a cofibre

$$S^{-9} \to D(\mathbb{C}P^9/\mathbb{C}P^4) \to D(\mathbb{C}P^9/\mathbb{C}P^3)$$

and we compute the image $\pi_9^s = \mathbb{Z}/2\{h_2^3 = h_1^2 h_3, h_1 c_0, Ph_1\} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^4)$. Compose the map $S^{-9} \to D(\mathbb{C}P^9/\mathbb{C}P^4) \to S^{-10}$ to the top cell. This can be detected by computing Sq^2. As $\mathrm{Sq}^2(x^4) = 0$, this map is null-homotopic. Therefore the attaching map goes down to $D(\mathbb{C}P^9/\mathbb{C}P^5)$, which we compute using the cofibre (3-3).

The map $S^{-9} \to D(\mathbb{C}P^9/\mathbb{C}P^5)$ is given by a map $S^{-9} \to S^{-12} \vee S^{-14}$ (hence in $\pi_3^s \oplus \pi_5^s = \mathbb{Z}/8\{h_2\}$) and a choice of null homotopy after composition by A to $S^{-15} \vee S^{-17}$. The choice of null homotopy lies in $\{S^{-9}, S^{-16} \vee S^{-18}\} = \pi_7^s \oplus \pi_9^s =$

$\mathbb{Z}/16\{h_3\} \oplus \mathbb{Z}/2\{h_2^3, h_1c_0, Ph_1\}$. We have $\mathrm{Sq}^2(x^4) = 0$ and hence the map is of the form $2kh_2$. From the formulas in [23, Proposition 5.2] it follows that $k = 1$. Now take a class $a \in \pi_9^s$. Note $h_2 \cdot \pi_9^s = 0$. Thus the image in π_{12}^s is 0.

Our interest is the image in π_{18}^s. This can again be computed using Toda brackets as in the previous case. The image onto π_{18}^s must be obtained by map onto the factor S^{-14} which is the only factor which attaches down to the -18–cell. But the attachment of S^{-9} onto this cell is 0 and hence the entire Toda bracket is forced to contain zero. Also the indeterminacy may be computed to be 0. Hence the map from $\pi_{18}^s \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^3)$ is injective. □

The above computation implies that the classes in π_{18}^s survive in the sequence all the way up to $\pi_0 D(\mathbb{C}P^9/\mathbb{C}P^3)$. However, in the next step we obtain a nontrivial kernel. We use some formulas for Toda brackets from [18].

Theorem 3.3 *The class $4h_2h_4 \in \pi_{18}^s$ maps to 0 in $\pi_0(D(\mathbb{C}P^9/\mathbb{C}P^2))$. It follows that the kernel of the map $\pi_{18}^s \to \pi_0(D\mathbb{C}P^9)$ is at least $\mathbb{Z}/2$.*

Proof The second statement follows from the first and the fact that h_2h_4 represents an element of order 8 in π_{18}^s. Thus we need prove only the first. We have the cofibre

$$S^{-7} \to D(\mathbb{C}P^9/\mathbb{C}P^3) \to D(\mathbb{C}P^9/\mathbb{C}P^2).$$

We compute the image $\pi_7^s = \mathbb{Z}/16\{h_3\} \to \pi_0 D(\mathbb{C}P^9/\mathbb{C}P^3)$. Compose the map $S^{-7} \to D(\mathbb{C}P^9/\mathbb{C}P^3) \to S^{-8}$ to the top cell. This composite is h_1 by the formula $\mathrm{Sq}^2(x^3) = x^4$ in Proposition 3.1. The map $h_1\colon \pi_7^s \to \pi_8^s$ has kernel $\mathbb{Z}/8\{2h_3\}$.

First we prove that the map $\pi_7^s \to \pi_0 D(\mathbb{C}P^7/\mathbb{C}P^3)$ has kernel $\mathbb{Z}/8\{2h_3\}$. That is, we show that $2h_3$ maps to 0 in the latter group. Note from the proof of Proposition 3.2 that none of the cells in dimension -8, -10 or -12 attach to the -14–cell. Thus we have $D(\mathbb{C}P^7/\mathbb{C}P^3) \simeq D(\mathbb{C}P^6/\mathbb{C}P^3) \vee S^{-14}$.

Observe that the -8–cell of $D(\mathbb{C}P^7/\mathbb{C}P^3)$ does not attach to the -10–cell. Together with the fact that π_{12}^s and π_{13}^s are 0, we get that $\pi_0 D(\mathbb{C}P^6/\mathbb{C}P^3) \to \pi_0(S^{-8} \vee S^{-10})$ is an isomorphism. Now observe that the composite map $S^{-7} \to S^{-8} \vee S^{-10}$ is h_1 on the first factor and h_2 on the second factor. Hence the kernel of the map $\pi_7^s \to \pi_0 D(\mathbb{C}P^6/\mathbb{C}P^3)$ is $\mathbb{Z}/8\{2h_3\}$.

Observe that Proposition 5.6 of [23] implies that the 14–cell in $\mathbb{C}P^7/\mathbb{C}P^2$ does not attach to the cells in dimension 8, 10 or 12, and attaches onto the 6–cell by the

map $2h_3$. Hence the map $S^{-7} \to D(\mathbb{C}P^6/\mathbb{C}P^3) \vee S^{-14} \to S^{-14}$ is given by $2h_3$. Thus the map $\pi_7^s \to \pi_{14}^s$ is multiplication by $2h_3$, which is 0. Thus we have that the kernel of the map $\pi_7^s \to \pi_0 D(\mathbb{C}P^7/\mathbb{C}P^3)$ is $\mathbb{Z}/8\{2h_3\}$.

Next we extend the above result to the map $\pi_7^s \to \pi_0 D(\mathbb{C}P^8/\mathbb{C}P^3)$. The map $2h_3\colon S^0 \to S^{-7}$ factors in the diagram as below:

$$\begin{array}{ccccccc} S^0 & \xrightarrow{2h_3} & S^{-7} & \longrightarrow & D(\mathbb{C}P^8/\mathbb{C}P^3) & \longrightarrow & S^{-8} \\ & \searrow^{h_3} & \uparrow 2 & & \uparrow & & \\ & & S^{-7} & \xrightarrow{\alpha} & D(\mathbb{C}P^8/\mathbb{C}P^4) & & \end{array}$$

We prove that the composite

$$S^0 \xrightarrow{h_3} S^{-7} \xrightarrow{\alpha} D(\mathbb{C}P^8/\mathbb{C}P^4)$$

is 0. The above argument shows that

$$S^0 \xrightarrow{h_3} S^{-7} \xrightarrow{\alpha} D(\mathbb{C}P^8/\mathbb{C}P^4) \to D(\mathbb{C}P^7/\mathbb{C}P^4) \simeq D(\mathbb{C}P^6/\mathbb{C}P^4) \vee S^{-14}$$

is 0. It remains to compute the map onto

$$\pi_{16}^s = \pi_0 S^{-16},$$

the bottom cell of $D(\mathbb{C}P^8/\mathbb{C}P^4)$. This may be computed via the cofibre

$$S^{-16} \to D(\mathbb{C}P^8/\mathbb{C}P^4) \xrightarrow{q} D(\mathbb{C}P^7/\mathbb{C}P^4) \to S^{-15}$$

as an element of the sum of Toda brackets $\langle h_3, q\circ\alpha, D(\mathbb{C}P^6/\mathbb{C}P^4) \to S^{-15}\rangle$ and $\langle h_3, q\circ\alpha, S^{-14} \to S^{-15}\rangle$.

We use that the attachment of the -6–cell to the -14–cell is $2h_3$, as noted above. Thus the attachment of α onto the -14–cell is given by $4h_3$. Hence the latter bracket equals $\langle h_3, 2h_3, h_1\rangle = 2h_1h_4 = 0$ modulo trivial indeterminacy.

The first bracket above is the 4–fold bracket $\langle h_3, 2h_2, h_1, h_2\rangle$. The indeterminacy of this bracket lies in the three-fold Toda bracket $\langle Ph_2, h_1, h_2\rangle$. Note that this bracket is in the kernel of the map $\pi_{16}^s \to \pi_0 D(\mathbb{C}P^8/\mathbb{C}P^4)$ killed by the attachment of the -10–cell. Modulo the above indeterminacy, the bracket $\langle h_3, 2h_2, h_1, h_2\rangle$ is a multiple of 2 and hence 0. Thus, the kernel of $\pi_7^s \to \pi_0 D(\mathbb{C}P^8/\mathbb{C}P^3)$ is $2h_3$.

The class $2h_3$ maps to 0 under $S^{-7} \to D(\mathbb{C}P^9/\mathbb{C}P^3) \to D(\mathbb{C}P^8/\mathbb{C}P^3)$. Thus it maps to $\pi_0 S^{-18}$. We prove that it maps to the class $4h_2h_4$ under this map. We have the factorization, as above:

$$\begin{array}{ccccccc} S^0 & \xrightarrow{2h_3} & S^{-7} & \longrightarrow & D(\mathbb{C}P^9/\mathbb{C}P^3) & \longrightarrow & S^{-8} \\ & \searrow^{h_3} & \uparrow 2 & & \uparrow & & \\ & & S^{-7} & \xrightarrow{\alpha} & D(\mathbb{C}P^9/\mathbb{C}P^4) & & \end{array}$$

We write $D(\mathbb{C}P^9/\mathbb{C}P^4)$ as the cofibre

$$\Sigma^{-1}D(\mathbb{C}P^6/\mathbb{C}P^4)\vee S^{-15}\to S^{-16}\vee S^{-18}\to D(\mathbb{C}P^9/\mathbb{C}P^4) \\ \to D(\mathbb{C}P^6/\mathbb{C}P^4)\vee S^{-14}\to S^{-15}\vee S^{-17}.$$

Hence the map onto π_{18}^s is a sum of Toda brackets $\langle h_3, \alpha, D(\mathbb{C}P^6/\mathbb{C}P^4)\to S^{-17}\rangle$ and $\langle h_3, \alpha, S^{-14}\to S^{-17}\rangle$. The latter map is the Toda bracket $\langle h_3, 4h_3, h_2\rangle = 2h_2h_4$ modulo trivial indeterminacy.

For the first map, note that the attaching map $D(\mathbb{C}P^6/\mathbb{C}P^4)\to S^{-17}$ restricting to the bottom cell S^{-12} is trivial and on the top cell is h_3 as computed by the Steenrod operation $\mathrm{Sq}^8(x^5)=x^9$. Therefore the first map is $\langle h_3, 2h_2, h_3\rangle$. This bracket is computed in [18, page 251] as $2h_2h_4$.

Therefore the sum of the two maps is $4h_2h_4$. It follows that the element $4h_2h_4$ maps to 0 in $\pi_0 D(\mathbb{C}P^9/\mathbb{C}P^2)$. This completes the proof. □

It follows from the above result that the kernel of $\pi_{18}^s\to\{\mathbb{C}P^9, S^0\}$ is at least $\mathbb{Z}/2$ and is a subgroup of $\mathbb{Z}/8$. We prove that it cannot be $\mathbb{Z}/8$, so that the kernel can be $\mathbb{Z}/2$ or $\mathbb{Z}/4$.

Theorem 3.4 *The class $h_2h_4\in\pi_{18}^s$ maps to a nontrivial class in $\{\mathbb{C}P^9, S^0\}$.*

Proof We use some computations from [28]. Recall that the element h_2h_4 is denoted by ν^*. For such a stable class, the sphere of origin is the first sphere where this class desuspends to. For the class ν^* the sphere of origin is S^{12} and it desuspends to the class $\xi_{12}\colon S^{30}\to S^{12}$. The class ν^* also desuspends to $\nu_{16}^*\colon S^{34}\to S^{16}$.

From Lemma 12.14 of [28] we know that $H(\nu_{16}^*)=\nu_{31} \pmod{2\nu_{31}}$. The latter maps isomorphically to the stable range and is equivalent to (an odd multiple of) the map $h_2\in\pi_3^s$. We have the commutative diagram:

$$\begin{array}{ccccc} [S^{34}, S^{16}] & \xrightarrow{H_\#} & [S^{34}, S^{31}] & \xrightarrow{\cong} & \{S^{18}, S^{15}\} \\ \downarrow & & \downarrow & & \downarrow \\ [\Sigma^{16}\mathbb{C}P^9, S^{16}] & \xrightarrow{H_\#} & [\Sigma^{16}\mathbb{C}P^9, S^{31}] & \xrightarrow{\cong} & \{\mathbb{C}P^9, S^{15}\} \end{array} \tag{3-4}$$

In terms of the Spanier–Whitehead duality the last map is induced by $\pi_{-15}S^{-18} \to \pi_{-15}D\mathbb{C}P^9$ from the inclusion of the bottom cell. Note that the map $D(\mathbb{C}P^9/\mathbb{C}P^6) \to D\mathbb{C}P^9$ is an isomorphism on π_{-15}. From the cofibre sequence (3-2), we have the exact sequence

$$\cdots \to \pi_{-15}S^{-15} \to \pi_{-15}S^{-18} \oplus \pi_{-15}S^{-16} \to \pi_{-15}D(\mathbb{C}P^9/\mathbb{C}P^6) \to \cdots.$$

The left map sends $1 \in \mathbb{Z}$ to $(h_2, h_1) \in \pi_{-15}S^{-18} \oplus \pi_{-15}S^{-16}$. It follows that $2h_2$ maps to 0 in $\pi_{-15}D(\mathbb{C}P^9/\mathbb{C}P^6)$ and the class h_2 maps nontrivially. Hence, in the diagram (3-4) the element ν_{16}^* in the top-left corner maps to a nontrivial element in the bottom-right group $\{\mathbb{C}P^9, S^{15}\}$. Therefore it maps nontrivially in $[\Sigma^{16}\mathbb{C}P^9, S^{16}]$.

Next we show that this implies that the image in $\{\mathbb{C}P^9, S^0\}$ is nontrivial. We have the diagram:

$$\begin{array}{ccc}
[\Sigma^{16}\mathbb{C}P^9/\mathbb{C}P^7, S^{16}] & \longrightarrow & \{\mathbb{C}P^9/\mathbb{C}P^7, S^0\} \\
\downarrow & & \downarrow \\
[\Sigma^{16}\mathbb{C}P^9, S^{16}] & \longrightarrow & \{\mathbb{C}P^9, S^0\} \\
\downarrow & & \downarrow \\
[\Sigma^{16}\mathbb{C}P^7, S^{16}] & \xrightarrow{\cong} & \{\mathbb{C}P^7, S^0\}
\end{array}$$

Note that the vertical arrows are exact. It follows that the kernel of $[\Sigma^{16}\mathbb{C}P^9, S^{16}] \to \{\mathbb{C}P^9, S^0\}$ is isomorphic to the kernel of $[\Sigma^{16}\mathbb{C}P^9/\mathbb{C}P^7, S^{16}] \to \{\mathbb{C}P^9/\mathbb{C}P^7, S^0\}$. We have observed in the proof of Proposition 3.2 that $\mathbb{C}P^9/\mathbb{C}P^7 \simeq S^{16} \vee S^{18}$. Hence, $[\Sigma^{16}\mathbb{C}P^9/\mathbb{C}P^7, S^{16}] \cong [S^{32}, S^{16}] \oplus [S^{34}, S^{16}]$, so that the above kernel is the direct sum

$$\mathrm{Ker}\big([S^{32}, S^{16}] \to \{S^{16}, S^0\}\big) \oplus \mathrm{Ker}\big([S^{34}, S^{16}] \to \{S^{18}, S^0\}\big).$$

The element $\nu_{16}^* \in \pi_{34}S^{16}$ maps to $\nu^* \in \pi_{18}^s$ and so does not lie in the kernel above. It follows that the image of ν_{16}^* maps nontrivially to $\{\mathbb{C}P^9, S^0\}$. Thus the element h_2h_4 is mapped nontrivially in $\{\mathbb{C}P^9, S^0\}$. □

Summarizing the computations in Proposition 3.2 and Theorems 3.3 and 3.4, we have the following result for Problem 2.8:

Corollary 3.5 *The kernel* $\mathrm{Ker}(f^*_{\mathbb{C}P^9}) \subset \pi_{18}^s = \mathbb{Z}/2 \oplus \mathbb{Z}/8$ *is nontrivial but not the entire group* $\mathbb{Z}/8$*. It is either* $\mathbb{Z}/2$ *or* $\mathbb{Z}/4$ *as a subgroup of* $\mathbb{Z}/8$*.*

3.2 Computations in higher dimensions

We follow up the computations in Section 3.1 by demonstrating that there are many examples where the inclusion of the bottom cell in $D(\mathbb{C}P^{8n+2})$ carries a nontrivial kernel in π_0. The methods here are easier, involving computations of Steenrod operations and the existence of certain stable homotopy classes.

The next example after 18 is 26. Recall from [25] that the group $\pi^s_{26} \cong \mathbb{Z}/2\{\mu_{26}\} \oplus \mathbb{Z}/2\{h_2^2 g\} \oplus \mathbb{Z}/3\{\beta_2\}$. We know that the class μ_{26} survives to $\mathbb{C}P^{13}$ from [8]. We have the following result:

Theorem 3.6 *The class $h_2^2 g$ maps to 0 in $\pi_0 D(\mathbb{C}P^{13}/\mathbb{C}P^{11})$. It follows that the kernel of the map $\pi^s_{26} \to \pi_0(D\mathbb{C}P^{13})$ is at least $\mathbb{Z}/2$.*

Proof As in Section 3.1, we have the filtration

$$S^{-26} = D(\mathbb{C}P^{13}/\mathbb{C}P^{12}) \to D(\mathbb{C}P^{13}/\mathbb{C}P^{11}) \to \cdots \to D(\mathbb{C}P^{13}/\mathbb{C}P^{0}) = D(\mathbb{C}P^{13})$$

and the map $S^{-26} \to D(\mathbb{C}P^{13})$ is the inclusion of the bottom cell. We show that $h_2^2 g$ is in the kernel. It suffices to work 2–locally. We have the cofibre

$$S^{-25} \to S^{-26} \to D(\mathbb{C}P^{13}/\mathbb{C}P^{11}).$$

The map $S^{-25} \to S^{-26}$ is an element of $\pi^s_1 = \mathbb{Z}/2\{h_1\}$ and, since $\mathrm{Sq}^2(x^{12}) = 0$ in the cohomology of $\mathbb{C}P^n$, the map is trivial. It follows that $D(\mathbb{C}P^{13}/\mathbb{C}P^{11}) \simeq S^{-26} \vee S^{-24}$. Next we have the cofibre

$$S^{-23} \to D(\mathbb{C}P^{13}/\mathbb{C}P^{11}) \to D(\mathbb{C}P^{13}/\mathbb{C}P^{10}).$$

The map $S^{-23} \to D(\mathbb{C}P^{13}/\mathbb{C}P^{11})$ is given by a pair of maps to S^{-26} and S^{-24}. We note the Steenrod squares $\mathrm{Sq}^2(x^{11}) = x^{12}$ and $\mathrm{Sq}^4(x^{11}) = x^{13}$ in the cohomology of $\mathbb{C}P^n$. Therefore the map is given by (h_1, h_2). On π_0 we have the sequence

$$\pi^s_{23} \xrightarrow{(h_1,h_2)} \pi^s_{24} \oplus \pi^s_{26} \to \pi_0 D(\mathbb{C}P^{13}/\mathbb{C}P^{10}).$$

Now note that $h_2 g$ represents a nontrivial element in π^s_{23} and $h_1 h_2 g = 0$. Thus $h_2^2 g$ is in the image of the left-hand map and hence maps to 0 in $\pi_0 D(\mathbb{C}P^{13}/\mathbb{C}P^{10})$. Therefore, $h_2^2 g$ maps to 0 in $\pi_0 D\mathbb{C}P^{13}$, proving the theorem. □

We have the corresponding result for the Problem 2.8.

Corollary 3.7 *The kernel $\mathrm{Ker}(f^*_{\mathbb{C}P^{13}}) \subset \pi^s_{26} = \mathbb{Z}/2 \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/3$ is nontrivial and contains a summand $\mathbb{Z}/2$.*

Next we show that techniques as in Theorem 3.6 exist in many high dimensions. More precisely, we demonstrate examples in higher dimensions where the map $S^{-8n-2} \to D(\mathbb{C}P^{8n+2})$ has a nontrivial kernel on π_0 using some p–local computations. We use a result from [20]: For $p \geq 7$ the classes $\alpha_1\beta_1^r\gamma_t$ are nontrivial in the stable homotopy groups of S^0 for $2 \leq t \leq p-1$ and $r \leq p-2$ (in dimension $n(t,p,r) = [2(tp^3-t-p^2)+2r(p^2-1-p)-2]$). With these assumptions, $\beta_1^r\gamma_t$ is also nontrivial in dimension $n(t,p,r)-(2p-3)$.

We note:

Proposition 3.8 *In $H^*(\mathbb{C}P^n;\mathbb{Z}/p)$, $P^1(x^k) \neq 0$ if and only if p does not divide k.*

In the next theorem, note that if $(p-1)(t-r)+r \equiv 3 \pmod 4$ then $n(t,p,r) \equiv 2 \pmod 8$.

Theorem 3.9 *Suppose that p is a prime ≥ 7, $2 \leq t \leq p-1$ and $r \leq p-2$. Assume that p does not divide $t+r$. Then the map $\pi^s_{n(t,p,r)} \to \pi_0 D(\mathbb{C}P^{n(t,p,r)/2})$ has nontrivial p–torsion in the kernel.*

Proof Note that the condition p does not divide $t+r$ implies that $n(t,p,r)-2(p-1)$ is not divisible by p. We work p–locally. The first nontrivial element in $\pi^s_* \otimes \mathbb{Z}_{(p)}$ in positive dimension is α_1 in dimension $2p-3$, and this is detected by the Steenrod operation P^1.

Constructing $D(\mathbb{C}P^{n(t,p,r)/2})$ cell by cell as above in the p–local category, the first possible nontrivial attaching map is the map

$$S^{-n(t,p,r)+2p-3} \to D(\mathbb{C}P^{n(t,p,r)/2}/\mathbb{C}P^{(n(t,p,r)-2(p-1))/2}).$$

The assumption that p does not divide $t+r$ implies that, in $\mathbb{C}P^N$ for $N \gg 0$,

$$P^1(x^{(n(t,p,r)-2(p-1))/2}) = x^{n(t,p,r)/2}.$$

Therefore the map $S^{-n(t,p,r)+2p-3} \to D(\mathbb{C}P^{n(t,p,r)/2}/\mathbb{C}P^{(n(t,p,r)-2(p-1))/2})$ is given by α_1 on the bottom cell. Now, in the long exact sequence of homotopy groups

$$\cdots \to \pi^s_{n(t,p,r)-(2p-3)} \to \pi_0 D(\mathbb{C}P^{n(t,p,r)/2}/\mathbb{C}P^{(n(t,p,r)-2(p-1))/2}) \to \pi_0 D(\mathbb{C}P^{n(t,p,r)/2}/\mathbb{C}P^{(n(t,p,r)-2(p-1))/2}) \to \cdots,$$

the nontrivial element $\beta_1^r\gamma_t$ maps to the nontrivial element $\alpha_1\beta_1^r\gamma_t$. It follows that the nontrivial element $\alpha_1\beta_1^r\gamma_t \in \pi^s_{n(t,p,r)}$ goes to 0 in $\pi_0 D(\mathbb{C}P^{n(t,p,r)/2})$. This completes the proof. □

Remark 3.10 The above conditions are easily satisfied. For example, if $t = 3,\ r = 1$ and $p \geq 7$, we have $n(t, p, r) = 6p^3 - 2p - 10$, which is $\equiv 2 \pmod 8$.

The following result is an immediate consequence of stringing together Theorems 2.4 and 2.5, Corollaries 3.5 and 3.7, and Theorem 3.9.

Theorem 3.11 (i) $I(\mathbb{C}P^9) = \mathbb{Z}_2$ or $\mathbb{Z}_4$.

(ii) $I(\mathbb{C}P^{13}) \supseteq \mathbb{Z}/2$.

(iii) *Suppose that p is a prime ≥ 7, $2 \leq t \leq p-1$ and $r \leq p-2$. Assume that p does not divide $t + r$ and $(p-1)(t-r) + r \equiv 3 \pmod 4$. Then $I(\mathbb{C}P^{n(t,p,r)/2}) \supseteq \mathbb{Z}/p$.*

4 Geometric structures of inequivalent smooth structures

In this section we explore some consequences of the computations in the previous section. It is known that inequivalent smoothings need not share certain basic geometric properties with the standard smooth structure. Hitchin [10] has shown that certain homotopy spheres do not admit a Riemannian metric of positive scalar curvature, while the round metric on the standard S^m has positive sectional curvature. In this section, we also show that there exist two smooth structures on $\mathbb{C}P^9$ such that one admits a metric of nonnegative scalar curvature and the other does not.

Let $\mathrm{Ker}(d_{\mathbb{R}})$ denote the kernel of the Adams d–invariant $d_{\mathbb{R}}\colon \pi^s_{8n+2} \to \mathbb{Z}/2$; see [1]. Under the isomorphism $\Theta_{8n+2} \cong \pi^s_{8n+2}$, the Adams d–invariant $d_{\mathbb{R}}\colon \pi^s_{8n+2} \to \mathbb{Z}/2$ may be identified with the α–invariant homomorphism $\alpha\colon \Theta_{8n+2} \to \mathbb{Z}/2$; see [10]. Therefore $\mathrm{Ker}(d_{\mathbb{R}})$ consists of homotopy $8n+2$–spheres which bound spin manifolds.

In [12], we studied the Adams d–invariant and asked the following question to determine the inertia group $I(\mathbb{C}P^{4n+1})$:

Question 4.1 *Let $f\colon \mathbb{C}P^{4n+1} \to S^{8n+2}$ be any degree 1 map with $n \geq 1$.*

Does there exist an element $\eta \in \mathrm{Ker}(d_{\mathbb{R}}) \subset \pi^s_{8n+2} = \Theta_{8n+2}$ such that the following is true:

($\star$) *If any map $h\colon S^{q+8n+2} \to S^q$ represents η, then*

$$h \circ \Sigma^q f\colon \Sigma^q \mathbb{C}P^{4n+1} \to S^q$$

is not null-homotopic.

The following theorem shows that the answer to the above question is yes for $n = 2$, where the Adams d–invariant $d_{\mathbb{R}}\colon \pi^s_{18} = \mathbb{Z}/2\{h_1P^2h_1\} \oplus \mathbb{Z}/8\{h_2h_4\} \to \mathbb{Z}/2$ is such that $\mathrm{Ker}(d_{\mathbb{R}}) = \mathbb{Z}_8\{h_2h_4\}$; see [1]. From Theorem 3.11 we readily deduce:

Proposition 4.2 *$I(\mathbb{C}P^9)$ is properly contained in* $\mathrm{Ker}(d_{\mathbb{R}}) = \mathbb{Z}/8$.

Picking up an element in $\mathrm{Ker}(d_{\mathbb{R}})$ which is not in $I(\mathbb{C}P^9)$, we have a class $\{S^{18}, S^0\}$ which maps nontrivially in $\{\mathbb{C}P^9, S^0\}$. Thus, for every representative $S^{q+18} \to S^q$, the corresponding map $\Sigma^q\mathbb{C}P^9 \to S^q$ is not null-homotopic. This answers Question 4.1 in the case $n = 2$.

The existence of classes outside the inertia group for $\mathbb{C}P^9$ also has the following consequence:

Theorem 4.3 *There exist three homotopy 18–spheres Σ^{18}_1, Σ^{18}_2 and Σ^{18}_3 such that the following is true:*

(i) *The manifolds $\mathbb{C}P^9$, $\mathbb{C}P^9 \# \Sigma^{18}_1$, $\mathbb{C}P^9 \# \Sigma^{18}_2$ and $\mathbb{C}P^9 \# \Sigma^{18}_3$ are pairwise nondiffeomorphic.*

(ii) *The manifolds $\mathbb{C}P^9 \# \Sigma^{18}_2$ and $\mathbb{C}P^9 \# \Sigma^{18}_3$ do not admit a metric of nonnegative scalar curvature but $\mathbb{C}P^9 \# \Sigma^{18}_1$ does.*

Proof We start with the first statement. Let Σ^{18}_1, Σ^{18}_2 and Σ^{18}_3 be homotopy spheres in $\Theta_{18} \cong \pi^s_{18} = \mathbb{Z}_2\{h_1P^2h_1\} \oplus \mathbb{Z}_8\{h_2h_4\}$ represented by the classes h_2h_4, $h_1P^2h_1$ and $h_1P^2h_1 + h_2h_4$, respectively. Note that $\Sigma^{18}_i \notin I(\mathbb{C}P^9)$ by Theorem 3.4, where $i = 1$, 2 or 3, and $\Sigma^{18}_1 \in \mathrm{Ker}(d_{\mathbb{R}})$ but Σ^{18}_2, $\Sigma^{18}_3 \notin \mathrm{Ker}(d_{\mathbb{R}})$. If $\mathbb{C}P^9 \# \Sigma^{18}_i$ is diffeomorphic to $\mathbb{C}P^9 \# \Sigma^{18}_j$, then $\Sigma^{18}_i \# (\Sigma^{18}_j)^{-1} \in I(\mathbb{C}P^9)$. But, $\Sigma^{18}_i \# (\Sigma^{18}_j)^{-1}$ equals Σ^{18}_k or $(\Sigma^{18}_k)^{-1}$, where $k \neq i$, j. This implies that the manifolds $\mathbb{C}P^9$, $\mathbb{C}P^9 \# \Sigma^{18}_1$, $\mathbb{C}P^9 \# \Sigma^{18}_2$ and $\mathbb{C}P^9 \# \Sigma^{18}_3$ are pairwise nondiffeomorphic. This proves (i).

Now we turn to (ii). Since $\Sigma^{18}_i \notin \mathrm{Ker}(d_{\mathbb{R}})$, where $i = 2$ or 3, and $\mathbb{C}P^9$ is a spin manifold equipped with its natural metric (the Fubini study metric) of positive scalar curvature, the α–invariant satisfies $\alpha(\mathbb{C}P^9) = 0$ and $\alpha(\Sigma^{18}_i) \neq 0$; see [10]. Therefore,

$$\alpha(\mathbb{C}P^9 \# \Sigma^{18}_i) = \alpha(\mathbb{C}P^9) + \alpha(\Sigma^{18}_i) \neq 0.$$

We now proceed by contradiction. Suppose there exists a nonnegative scalar curvature Riemannian metric g on $\mathbb{C}P^9 \# \Sigma^{18}_i$. The nonvanishing of the α–invariant and the well-known deformation properties of scalar curvature [15] now imply that the Riemannian

manifold $(\mathbb{C}P^9 \# \Sigma_i^{18}, g)$ must be scalar-flat and $\mathbb{C}P^9 \# \Sigma_i^{18}$ has a nontrivial parallel spinor. Therefore the Riemannian manifold $(\mathbb{C}P^9 \# \Sigma_i^{18}, g)$ has special holonomy [10]. This is a contradiction, since $\mathbb{C}P^9 \# \Sigma_i^{18}$ has generic holonomy. Hence $\mathbb{C}P^9 \# \Sigma_2^{18}$ and $\mathbb{C}P^9 \# \Sigma_3^{18}$ do not admit a metric of nonnegative scalar curvature. Now consider the connected sum $\mathbb{C}P^9 \# \Sigma_1^{18}$ and, by using the fact $\alpha(\Sigma_1^{18}) = 0$, it follows that $\mathbb{C}P^9 \# \Sigma_1^{18}$ admits a metric of positive scalar curvature [9]. This proves (ii). □

Another application is the following result, which is an immediate consequence of Theorem 4.3 and [12, Theorem 1.4]:

Theorem 4.4 *Let Σ_1^{18}, Σ_2^{18} and Σ_3^{18} be the specific homotopy 18–spheres posited in Theorem 4.3. Given a positive real number ϵ, there exists a closed complex hyperbolic manifold M^{18} of complex dimension 9 such that the following is true:*

(i) *The manifolds M^{18}, $M^{18} \# \Sigma_1^{18}$, $M^{18} \# \Sigma_2^{18}$ and $M^{18} \# \Sigma_3^{18}$ are pairwise nondiffeomorphic.*

(ii) *Each of the manifolds $M^{18} \# \Sigma_1^{18}$, $M^{18} \# \Sigma_2^{18}$ and $M^{18} \# \Sigma_3^{18}$ supports a negatively curved Riemannian metric whose sectional curvatures all lie in the closed interval $[-4-\epsilon, -1+\epsilon]$.*

References

[1] **J F Adams**, *On the groups $J(X)$, IV*, Topology 5 (1966) 21–71 MR

[2] **J F Adams**, *Stable homotopy and generalised homology*, University of Chicago Press (1974) MR

[3] **W Browder**, *On the action of $\Theta^n(\partial\pi)$*, from "Differential and combinatorial topology (a symposium in honor of Marston Morse)" (S S Cairns, editor), Princeton Univ. Press (1965) 23–36 MR

[4] **E H Brown, Jr**, **B Steer**, *A note on Stiefel manifolds*, Amer. J. Math. 87 (1965) 215–217 MR

[5] **G Brumfiel**, *Homotopy equivalences of almost smooth manifolds*, Comment. Math. Helv. 46 (1971) 381–407 MR

[6] **G W Brumfiel**, *Homotopy equivalences of almost smooth manifolds*, from "Algebraic topology" (A Liulevicius, editor), Amer. Math. Soc., Providence, RI (1971) 73–79 MR

[7] **A D Elmendorf**, **I Kriz**, **M A Mandell**, **J P May**, *Rings, modules, and algebras in stable homotopy theory*, Mathematical Surveys and Monographs 47, Amer. Math. Soc., Providence, RI (1997) MR

[8] **F T Farrell**, **L E Jones**, *Complex hyperbolic manifolds and exotic smooth structures*, Invent. Math. 117 (1994) 57–74 MR

[9] **M Gromov**, **H B Lawson, Jr**, *The classification of simply connected manifolds of positive scalar curvature*, Ann. of Math. 111 (1980) 423–434 MR

[10] **N Hitchin**, *Harmonic spinors*, Advances in Math. 14 (1974) 1–55 MR

[11] **M Hovey**, **B Shipley**, **J Smith**, *Symmetric spectra*, J. Amer. Math. Soc. 13 (2000) 149–208 MR

[12] **R Kasilingam**, *Farrell–Jones spheres and inertia groups of complex projective spaces*, Forum Math. 27 (2015) 3005–3015 MR

[13] **R Kasilingam**, *Classification of smooth structures on a homotopy complex projective space*, Proc. Indian Acad. Sci. Math. Sci. 126 (2016) 277–281 MR

[14] **K Kawakubo**, *Inertia groups of low dimensional complex projective spaces and some free differentiable actions on spheres, I*, Proc. Japan Acad. 44 (1968) 873–875 MR

[15] **J L Kazdan**, **F W Warner**, *Scalar curvature and conformal deformation of Riemannian structure*, J. Differential Geometry 10 (1975) 113–134 MR

[16] **M A Kervaire**, **J W Milnor**, *Groups of homotopy spheres, I*, Ann. of Math. 77 (1963) 504–537 MR

[17] **R C Kirby**, **L C Siebenmann**, *Foundational essays on topological manifolds, smoothings, and triangulations*, Annals of Mathematics Studies 88, Princeton Univ. Press (1977) MR

[18] **S O Kochman**, *Bordism, stable homotopy and Adams spectral sequences*, Fields Institute Monographs 7, Amer. Math. Soc., Providence, RI (1996) MR

[19] **A Kosiński**, *On the inertia group of π–manifolds*, Amer. J. Math. 89 (1967) 227–248 MR

[20] **C-N Lee**, *Detection of some elements in the stable homotopy groups of spheres*, Math. Z. 222 (1996) 231–245 MR

[21] **J Levine**, *Inertia groups of manifolds and diffeomorphisms of spheres*, Amer. J. Math. 92 (1970) 243–258 MR

[22] **J Milnor**, *On manifolds homeomorphic to the 7–sphere*, Ann. of Math. 64 (1956) 399–405 MR

[23] **R E Mosher**, *Some stable homotopy of complex projective space*, Topology 7 (1968) 179–193 MR

[24] **S P Novikov**, *Homotopically equivalent smooth manifolds, I*, Izv. Akad. Nauk SSSR Ser. Mat. 28 (1964) 365–474 MR In Russian; translated in Amer. Math. Soc. Transl. 48 (1965) 271–396

[25] **D C Ravenel**, *Complex cobordism and stable homotopy groups of spheres*, 2nd edition, AMS Chelsea 347, Amer. Math. Soc., Providence, RI (2004) MR

[26] **R Schultz**, *On the inertia group of a product of spheres*, Trans. Amer. Math. Soc. 156 (1971) 137–153 MR

[27] **I Tamura**, *Sur les sommes connexes de certaines variétés différentiables*, C. R. Acad. Sci. Paris 255 (1962) 3104–3106 MR

[28] **H Toda**, *Composition methods in homotopy groups of spheres*, Annals of Mathematics Studies 49, Princeton Univ. Press (1962) MR

[29] **C T C Wall**, *Classification of* $(n-1)$*–connected* $2n$*–manifolds*, Ann. of Math. 75 (1962) 163–189 MR

[30] **H E Winkelnkemper**, *On the action of* Θ^n*, I*, Trans. Amer. Math. Soc. 206 (1975) 339–346 MR

Department of Mathematical and Computational Science, Indian Association for the Cultivation of Science
Kolkata, India

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute
Bangalore, India

mcssb@iacs.res.in, mathsramesh1984@gmail.com

Received: 13 November 2016 Revised: 10 July 2017

msp

Algebraic & Geometric Topology 18 (2018) 409–439

Nonexistence of boundary maps for some hierarchically hyperbolic spaces

SARAH C MOUSLEY

We provide negative answers to questions posed by Durham, Hagen and Sisto on the existence of boundary maps for some hierarchically hyperbolic spaces, namely maps from right-angled Artin groups to mapping class groups. We also prove results on existence of boundary maps for free subgroups of mapping class groups.

57M07; 20F36, 20F65

1 Introduction

Let Γ be a finite graph with vertex set $V(\Gamma) = \{s_1, \dots, s_k\}$. The right-angled Artin group determined by Γ, denoted by $A(\Gamma)$, is the group with the presentation

$$A(\Gamma) = \langle s_1, \dots, s_k : [s_i, s_j] = 1 \iff s_i s_j \text{ is an edge in } \Gamma \rangle.$$

Let $S = S_{g,n}$ be a connected, oriented surface of genus g with n punctures, and let $\mathrm{Mod}(S)$ denote the mapping class group of S. Clay, Leininger and Mangahas [6] and Koberda [8] construct "nice" embeddings of right-angled Artin groups to mapping class groups. Behrstock, Hagen and Sisto [3; 2] introduced a geometric structure called a hierarchically hyperbolic space (HHS). Important examples of spaces that are HHSs include mapping class groups of surfaces and right-angled Artin groups. Durham, Hagen and Sisto [7] constructed a boundary for hierarchically hyperbolic spaces (see Section 2). In that paper, the authors ask the following question, motivated by a desire to develop a notion of geometrically finite subgroups of mapping class groups:

Question 1.1 *Let $A(\Gamma)$ be a right-angled Artin group embedded in $\mathrm{Mod}(S)$ in the sense of either Clay, Leininger and Mangahas* [6] *or Koberda* [8]. *Does the embedding $A(\Gamma) \to \mathrm{Mod}(S)$ extend continuously to an injective map $\partial A(\Gamma) \to \partial \mathrm{Mod}(S)$?*

We prove that in general the answer to Question 1.1 is no by providing, for each type of embedding, an explicit example where the embedding does not extend continuously.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.409

Theorem 1.2 *There exists a surface S, a right-angled Artin group Γ, a Clay–Leininger–Mangahas embedding $\phi\colon A(\Gamma)\to \mathrm{Mod}(S)$, and a Koberda embedding $\phi'\colon A(\Gamma)\to \mathrm{Mod}(S)$ such that, regardless of the HHS structure on $A(\Gamma)$, neither ϕ nor ϕ' extends continuously to a map $\partial A(\Gamma)\to \partial\mathrm{Mod}(S)$.*

Clay–Leininger–Mangahas (CLM) embeddings are quasi-isometric embeddings (see Theorem 2.5). Thus Theorem 1.2 shows that the sufficient conditions for extendability of maps between hierarchically hyperbolic spaces are different than those for existence of Cannon–Thurston maps for Gromov hyperbolic spaces. Indeed, quasi-isometric embeddings between Gromov hyperbolic spaces always extend continuously to maps between Gromov boundaries.

We also prove the following result, which gives a complete characterization of the Koberda embeddings of free groups sending all generators to powers of Dehn twists that have continuous extensions:

Theorem 1.3 *Let $\{\alpha_1,\dots,\alpha_k\}$ be a collection of pairwise intersecting curves in S and Γ the graph with $V(\Gamma)=\{s_1,\dots,s_k\}$ and no edges. For sufficiently large N, the homomorphism*

$$\phi\colon A(\Gamma)\to \mathrm{Mod}(S) \quad \text{defined by } \phi(s_i)=T_{\alpha_i}^N \text{ for all } i$$

is injective by the work of Koberda [8]. *Moreover, ϕ extends continuously to a map $\partial A(\Gamma)\to\partial\mathrm{Mod}(S)$ if and only if $\{\alpha_1,\dots,\alpha_k\}$ pairwise fill S, where $A(\Gamma)$ is equipped with any HHS structure.*

In fact, we prove something stronger than Theorem 1.3. We prove a nonexistence result (Theorem 5.3) for a class of Koberda embeddings of right-angled Artin groups that are not necessarily free groups. We also prove an existence result (Theorem 6.1) for a class of embeddings of free groups that includes the Koberda embeddings described in Theorem 1.3 as well as a class of CLM embeddings. The following question remains open:

Question 1.4 *Let $A(\Gamma)\to\mathrm{Mod}(S)$ be a CLM embedding of a free group that sends some pair of generators of $A(\Gamma)$ to mapping classes whose full supports together do not fill S. Is it always the case that ϕ does not extend? In other words, does the forward direction of Theorem 1.3 hold for CLM embeddings?* (Theorem 6.1 proves the backwards direction).

Idea behind the nonexistence proofs (Theorems 1.2, 1.3 and 5.3) All of the embeddings $\phi\colon A(\Gamma)\to \mathrm{Mod}(S)$ we present that do not extend share the following key feature: for some pair of noncommuting generators a and b of $A(\Gamma)$, the subsurface Y filled by the full supports of $\phi(a)$ and $\phi(b)$ is a *proper* subsurface of S. For the embeddings we consider, this allows us to produce two sequences in $A(\Gamma)$ that converge to the same point in $\partial A(\Gamma)$, but whose images do not converge to the same point in $\partial\mathrm{Mod}(S)$. We choose the nth term of the first sequence so that the annular projection of its image to some boundary component γ of Y is distance $O(n)$ from a basepoint, while the projection to γ of the image of the second sequence has bounded diameter. We then show that $O(n)$ is fast enough to conclude that every accumulation point in $\partial\mathrm{Mod}(S)$ of the image of the first sequence has a term associated to γ. On the other hand, accumulation points of the image of the second sequence have no such term. Thus the images of the sequences do not converge to the same point in $\partial\mathrm{Mod}(S)$.

The following open question arises naturally from our nonexistence proofs:

Question 1.5 *Let* $\mathrm{Teich}(S)$ *denote the Teichmüller space of a surface* S*, equipped with the Weil–Petersson metric. There is a hierarchically hyperbolic space structure on* $\mathrm{Teich}(S)$*, where the set of domains is all nonannular subsurfaces of* S (see Brock [5]). *Given that we show annular subsurface projections can obstruct extendability, we wonder if an orbit map from* $A(\Gamma)$ *to* $\mathrm{Teich}(S)$ *corresponding to a CLM embedding* $A(\Gamma)\to\mathrm{Mod}(S)$ *extends continuously to a boundary map.* (Note that this is clearly not the case for the Koberda embeddings described in Theorem 2.6, since applying powers of a Dehn twist to a point in $\mathrm{Teich}(S)$ can move it only a bounded amount.)

In Section 2 we will recall relevant definitions and theorems and introduce notation. Section 3 will establish a handful of lemmas that will be used for proving Theorem 1.2. Section 4 is devoted to proving Theorem 1.2 for a CLM embedding, and in Section 5 we prove Theorem 1.2 for a Koberda embedding. Using similar techniques, we then prove that a more general class of Koberda embeddings of right-angled Artin groups do not extend continuously (Theorem 5.3), which will imply one direction of Theorem 1.3. In Section 6 we will prove Theorem 6.1, which will imply the other direction of Theorem 1.3.

Remark We call the embeddings that send generators of our right-angled Artin group to mapping classes that are pseudo-Anosov on subsurfaces CLM embeddings and those that send generators to powers of Dehn twists Koberda embeddings, even though

Koberda [8] proved that both these types of embeddings are injective. We do this primarily to distinguish the two types of embeddings, but also to emphasize that CLM embeddings have nice geometric properties (see Theorem 2.5).

Acknowledgments The author was supported by the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program and by a Research Assistantship through NSF Grant number DMS-1510034. The author would like to thank her PhD advisor Chris Leininger for his numerous insights, which inspired a great deal of this paper, and also for his patience and consistent support. The author would also like to thank Mark Hagen and Matt Durham for helpful conversations and the referees for their suggestions and careful reading of this paper.

2 Background

In this section, we recall some needed definitions and theorems.

Notation Let $f, g\colon X \to \mathbb{R}$ be functions. Given constants $A \geq 1$ and $B \geq 0$, we write $f \succ_{A,B} g$ to mean $f(x) \geq \frac{1}{A} g(x) - B$ for all $x \in X$, and will just write $f \succ g$ when the constants are understood.

2.1 Curves and subsurfaces

Throughout this paper, we let $S = S_{g,n}$ denote a connected, oriented surface of genus g with n punctures. Define the *complexity of* S to be $\xi(S) = 3g - 3 + n$. We will always assume $\xi(S) \geq 1$. Additionally, we fix a complete hyperbolic metric on S. That is, we assume that S is of the form $S = \mathbb{H}^2/\Lambda$, where $\Lambda \subseteq \mathrm{Isom}^+(\mathbb{H}^2)$ and Λ acts properly discontinuously and freely on $\mathbb{H}^2$.

For $i = 1, 2$, let $\widetilde{\gamma}_i$ be a bi-infinite path in $\mathbb{H}^2$ with ends limiting to distinct points x_i and y_i on $\partial\mathbb{H}^2$. We say that $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ *link* if the geodesic connecting x_1 to y_1 intersects the geodesic connecting x_2 to y_2 in the interior of $\mathbb{H}^2$.

By a *curve* in S, we will always mean the geodesic representative in the homotopy class of an essential, simple, closed curve in S. By a *multicurve* in S, we will always mean a collection of pairwise disjoint curves in S. We write $i(\alpha, \beta)$ to denote the geometric intersection number of curves α and β. We say that a pair of curves α and β *fills* S if for every curve γ in S we have $i(\gamma, \alpha) > 0$ or $i(\gamma, \beta) > 0$.

A *nonannular subsurface Y of S* is a component of S after removing a (possibly empty) collection of pairwise disjoint curves on S. Additionally, we require that Y satisfies $\xi(Y) \geq 1$; in particular, we do not consider a pair of pants to be a subsurface. We define ∂Y to be the collection of curves in S that are disjoint from Y and also are contained in the closure of Y, treating Y as a subset of S. When $Y \neq S$, the path metric completion of Y is a surface with boundary, and the image of this boundary under the map induced by the inclusion $Y \subseteq S$ is ∂Y.

An annular subsurface of S is defined as follows. Let α be a curve in S. Choose a component $\widetilde{\alpha}$ of the preimage of α in $\mathbb{H}^2$, and let $h \in \Lambda$ be a primitive isometry with axis $\widetilde{\alpha}$. Define

$$Y = (\overline{\mathbb{H}^2} - \{x, y\})/\langle h \rangle,$$

where x and y are the fixed points of h on $\partial\mathbb{H}^2$. Observe that Y is a compact annulus and $\operatorname{int}(Y) \to S$ is a covering. We say that Y is the *annular subsurface of S with core curve α*. We define ∂Y to be α.

For any subsurface Y of S, we will write $Y \subseteq S$, even though when Y is an annulus, Y is not a subset of S.

Given $f \in \operatorname{Mod}(S)$ and a curve or simple bi-infinite geodesic γ in S, we define $f(\gamma)$ to be the curve or simple bi-infinite geodesic obtained as follows. Consider a component $\widetilde{\gamma}$ of the preimage of γ in $\mathbb{H}^2$. Choose a representative ψ in the isotopy class of f and lift it to a map $\widetilde{\psi}\colon \mathbb{H}^2 \to \mathbb{H}^2$. We define $f(\gamma)$ to be the image in S of the geodesic in $\mathbb{H}^2$ that connects the endpoints of $\widetilde{\psi}(\widetilde{\gamma})$ on $\partial\mathbb{H}^2$. Given $Y \subseteq S$, if Y is nonannular, we let $f(Y)$ denote the nonannular subsurface in its isotopy class. If Y is an annulus with core curve α, we let $f(Y)$ denote the annular subsurface of S with core curve $f(\alpha)$.

2.2 Curve complex

Let Y be a subsurface of S. If Y satisfies $\xi(Y) \geq 1$, the *curve complex of Y*, denoted by $\mathcal{C}(Y)$, is the simplicial complex whose vertices are curves contained in Y, and if $\xi(Y) > 1$, a set of vertices forms a simplex if and only if they are pairwise disjoint. If $\xi(Y) = 1$, then we define the simplices of $\mathcal{C}(Y)$ differently. In the case that Y is a once-punctured torus, a set of vertices forms a simplex if and only if they pairwise intersect exactly once. If Y is a four-times-punctured sphere, a set of vertices forms a simplex if and only if they pairwise intersect exactly twice.

Now let Y be a compact annulus. Consider all embedded arcs in Y that connect one boundary component to the other. We define two arcs to be equivalent if one can be homotoped to the other, fixing the endpoints of the arcs throughout the homotopy. In this case, the *curve complex of* Y is the simplicial complex whose vertices are equivalence classes of arcs, and a set of vertices forms a simplex if and only if for each pair of vertices there exist representative arcs of each whose restrictions to $\operatorname{int}(Y)$ are disjoint. The following simple formula will be useful to us: given inequivalent arcs α and β in $\mathcal{C}(Y)$,

$$d_{\mathcal{C}(Y)}(\alpha, \beta) = |\alpha \cdot \beta| + 1, \tag{1}$$

where $\alpha \cdot \beta$ denotes the algebraic intersection number of α and β.

2.3 Markings and subsurface projection

A *marking* μ *on* S is a maximal collection of pairwise disjoint curves in S, denoted by $\operatorname{base}(\mu)$, together with another collection of associated curves called *transversals*: for each $\beta \in \operatorname{base}(\mu)$ its associated transversal γ_β is a curve that intersects β minimally (ie once or twice) and is disjoint from all other curves in $\operatorname{base}(\mu)$.

Let Y be a subsurface of S and β a multicurve in S. We will now define *the projection of* β *to* Y, which we will denote by $\pi_Y(\beta)$. Suppose Y is not an annulus and β is a single curve. If β is disjoint from Y, define $\pi_Y(\beta) = \varnothing$. If β is contained in Y, define $\pi_Y(\beta) = \beta$. Otherwise, $\beta \cap Y$ is a collection of essential arcs in Y with endpoints on ∂Y. For each such arc γ, take the geodesic representatives of the boundary components of a small regular neighborhood of $\gamma \cup \partial Y$ that are contained in Y. Define $\pi_Y(\beta)$ to be the collection of all such curves over all arcs γ in $\beta \cap Y$. If β is a multicurve, define $\pi_Y(\beta)$ to be the union of the projections to Y of each curve in β.

Now let Y be an annular subsurface with core curve α and $\operatorname{int}(Y) \to S$ the associated covering. Let β be a multicurve or a bi-infinite, simple geodesic in S. Consider the components of the full preimage of β in $\operatorname{int}(Y)$ that are arcs. We will view each such component as having endpoints on the boundary of Y. In this case, we define $\pi_Y(\beta)$ to be the (equivalence classes of) arcs in this collection that have an endpoint on each boundary component of Y. When convenient, we will write $\pi_\alpha(\beta)$ instead of $\pi_Y(\beta)$.

We now describe how to project a marking μ to $Y \subseteq S$. If Y is nonannular or Y is an annulus whose core curve is not contained in $\operatorname{base}(\mu)$, we define $\pi_Y(\mu) = \pi_Y(\operatorname{base}(\mu))$. Otherwise, Y is an annulus with core curve $\alpha \in \operatorname{base}(\mu)$, and we define $\pi_Y(\mu)$ to be $\pi_Y(\gamma_\alpha)$, where γ_α is the transversal associated to α.

Given any subsurface $Y \subseteq S$, we define

$$d_Y(\mu, \mu') = \operatorname{diam}_{\mathcal{C}(Y)}(\pi_Y(\mu) \cup \pi_Y(\mu')),$$

where μ and μ' are markings, collections of curves, or (when Y is an annulus) bi-infinite simple geodesics in S. A useful fact about subsurface projection is the following: for all $f \in \operatorname{Mod}(S)$,

$$d_Y(f(\mu), f(\mu')) = d_{f^{-1}(Y)}(\mu, \mu').$$

In this paper, we utilize the following theorem, which involves subsurface projections:

Theorem 2.1 [11, Lemma 2.3] *For all subsurfaces W of S, given any marking or multicurve μ such that $\pi_W(\mu) \neq \varnothing$, we have that* $\operatorname{diam}_{\mathcal{C}(W)}(\pi_W(\mu)) \leq 2$. *If W is an annulus, then* $\operatorname{diam}_{\mathcal{C}(W)}(\pi_W(\mu)) \leq 1$.

Masur and Minsky [11] define the *marking graph of S*, denoted by $\widetilde{\mathcal{M}}(S)$, to be the graph whose vertices are markings and vertices are adjacent if one can be obtained from the other by an elementary move; see [11] for a complete definition. Giving $\widetilde{\mathcal{M}}(S)$ the path metric $d_{\widetilde{\mathcal{M}}(S)}$ and $\operatorname{Mod}(S)$ a word metric $d_{\operatorname{Mod}(S)}$, there is an action of $\operatorname{Mod}(S)$ on $\widetilde{\mathcal{M}}(S)$ by isometries for which every orbit map is a quasi-isometry. The following theorem gives a relationship between distances in $\widetilde{\mathcal{M}}(S)$ and subsurface projections:

Theorem 2.2 [11, Lemma 3.5] *For any subsurface W of S and any markings μ and μ' on S, we have that $d_W(\mu, \mu') \leq 4 d_{\widetilde{\mathcal{M}}(S)}(\mu, \mu')$.*

We say that distinct subsurfaces X and Y are *disjoint* if $\pi_X(\partial Y) = \varnothing$ and $\pi_Y(\partial X) = \varnothing$. We say that X is a *proper subsurface of Y*, denoted by $X \subsetneq Y$, if $\pi_Y(\partial X) \neq \varnothing$ and $\pi_X(\partial Y) = \varnothing$. We say that X and Y are *overlapping*, denoted by $X \pitchfork Y$, if $\pi_Y(\partial X) \neq \varnothing$ and $\pi_X(\partial Y) \neq \varnothing$. In the case where X and Y are not annuli, these relationships, respectively, are disjointness, proper containment and intersection without containment as subsets of S. We say X and Y *fill* S if for every curve γ in S we have $\pi_X(\gamma) \neq \varnothing$ or $\pi_Y(\gamma) \neq \varnothing$.

The following theorems will be used to prove our results. The first theorem was proved in [1] and later a simpler proof with constructive constants appeared in [9].

Theorem 2.3 (Behrstock inequality [1, Theorem 4.3; 9, Lemma 2.13]) *Let X and Y be overlapping subsurfaces of S and μ a marking on S. Then*

$$d_X(\mu, \partial Y) \geq 10 \implies d_Y(\mu, \partial X) \leq 4.$$

Theorem 2.4 (Bounded Geodesic Image Theorem [11, Theorem 3.1]) *There exists a constant K_0 depending only on S such that the following is true: Let Y and Z be subsurfaces of S with Y a proper subsurface of Z. Let $v_1, \dots, v_n$ be any geodesic segment in $\mathcal{C}(Z)$ satisfying $\pi_Y(v_i) \neq \varnothing$ for all $1 \leq i \leq n$. Then*

$$\operatorname{diam}_{\mathcal{C}(Y)}(\pi_Y(v_1) \cup \cdots \cup \pi_Y(v_n)) \leq K_0.$$

2.4 Partial order on subsurfaces

Let μ and μ' be markings on S and $K \geq 20$. Let $\Omega(K, \mu, \mu')$ denote the collection of subsurfaces Y of S such that $d_Y(\mu, \mu') \geq K$. Behrstock, Kleiner, Minsky and Mosher [4] define the following partial order on $\Omega(K, \mu, \mu')$: given $X, Y \in \Omega(K, \mu, \mu')$ such that $X \pitchfork Y$, define $X \prec Y$ if and only if one of the following equivalent conditions is satisfied:

$$d_X(\mu, \partial Y) \geq 10, \quad d_X(\partial Y, \mu') \leq 4, \quad d_Y(\mu, \partial X) \leq 4 \quad \text{or} \quad d_Y(\partial X, \mu') \geq 10.$$

That these conditions are equivalent is a consequence of Theorem 2.3; see Corollary 3.7 in [6].

2.5 Embedding RAAGs in Mod(S)

If $f \in \operatorname{Mod}(S)$ is such that there exists a representative in the isotopy class of f that pointwise fixes the complement of a nonannular subsurface Y, we say that f is *supported on* Y. Given such an f, we define the *translation length of f on $\mathcal{C}(Y)$* to be

$$\tau_Y(f) = \lim_{n \to \infty} \frac{d_Y(\mu, f^n(\mu))}{n},$$

where μ is any marking on S. If $f \in \operatorname{Mod}(S)$ is a power of a Dehn twist about a curve α, we say that f is *supported on* the annular subsurface Y with core curve α, and define $\tau_Y(f)$ to be the absolute value of the power. In either case, we say that Y *fully supports* f if $\tau_Y(f) > 0$. By the work of Masur and Minsky [10], when Y is nonannular, Y fully supports f if and only if f is pseudo-Anosov on Y.

Clay, Leininger and Mangahas [6] proved the following result, which allows us to find quasi-isometrically embedded right-angled Artin subgroups inside $\operatorname{Mod}(S)$:

Theorem 2.5 [6, Theorem 2.2] *Let Γ be a finite graph with $V(\Gamma) = \{s_1, \dots, s_k\}$, and let $\{X_1, \dots, X_k\}$ be a collection of nonannular subsurfaces of S. Suppose $s_i s_j$ is an edge in Γ if and only if X_i and X_j are disjoint, and $s_i s_j$ is not an edge in Γ*

if and only if $X_i \pitchfork X_j$ or $i = j$. Then there exists a constant $C > 0$ such that the following holds: Let $\{f_1, \dots, f_k\}$ be a set of mapping classes of S such that f_i is pseudo-Anosov on X_i and satisfies $\tau_{X_i}(f_i) \geq C$ for all i. Then the homomorphism

$$\phi\colon A(\Gamma) \to \mathrm{Mod}(S) \quad \text{defined by } \phi(s_i) = f_i \text{ for all } i$$

is a quasi-isometric embedding, implying that ϕ is injective, since $A(\Gamma)$ is torsion-free.

Koberda [8] also has a result which produces right-angled Artin subgroups of $\mathrm{Mod}(S)$. Below we give a special case of Koberda's result that we will use.

Theorem 2.6 [8, Theorem 1.1] *Let $\{\alpha_1, \dots, \alpha_k\}$ be a collection of distinct curves in S. Let Γ be the graph with $V(\Gamma) = \{s_1, \dots, s_k\}$ and with $s_i s_j$ an edge in Γ if and only if $i(\alpha_i, \alpha_j) = 0$. Then, for sufficiently large N, the homomorphism*

$$\phi\colon A(\Gamma) \to \mathrm{Mod}(S) \quad \text{defined by } \phi(s_i) = T_{\alpha_i}^N \text{ for all } i$$

is injective, where T_{α_i} denotes a Dehn twist about α_i.

2.6 Gromov boundary of hyperbolic spaces

A geodesic metric space X is *Gromov hyperbolic* (or just hyperbolic) if there exists a $\delta \geq 0$ such that, given any geodesic triangle in X, each side is contained in the δ–neighborhood of the union of the other two sides. Given a Gromov hyperbolic space (X, d_X) and points $x, y, z \in X$, the *Gromov product* of x and y with respect to z is defined as

$$(x, y)_z = \tfrac{1}{2}(d_X(x, z) + d_X(y, z) - d_X(x, y)).$$

We say that a sequence (x_n) in X *converges at infinity* if $\liminf_{i,j\to\infty}(x_i, x_j)_z = \infty$ for some (any) $z \in X$. We define two such sequences (x_n) and (y_n) to be equivalent if $\liminf_{i,j\to\infty}(x_i, y_j)_z = \infty$ for some (any) $z \in X$. The *Gromov boundary of X* is the collection of all such sequences up to this equivalence, and is denoted by $\partial_G X$ or just ∂X when it is clear from context that we are using the Gromov boundary.

One Gromov hyperbolic space that this paper is concerned with is the curve complex of S, which was proved to be Gromov hyperbolic by Masur and Minsky [11]. We can now state a corollary of Theorem 2.4 that will be useful later.

Corollary 2.7 *Let X and Y be subsurfaces of S with X a proper subsurface of Y. Suppose $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of markings on S such that $\pi_Y(\mu_n) \to \lambda$ for some $\lambda \in \partial\mathcal{C}(Y)$. Then* $\mathrm{diam}_{\mathcal{C}(X)}(\pi_X(\mu_1) \cup \pi_X(\mu_2) \cup \cdots) < \infty$.

Proof For each n, choose $\alpha_n \in \pi_Y(\mu_n)$. Because $\pi_Y(\mu_n) \to \lambda \in \partial\mathcal{C}(Y)$, we can choose L large so that for all $n \geq L$ we have

$$(\alpha_n, \alpha_L)_{\alpha_1} \geq 2 + d_Y(\partial X, \alpha_1), \tag{2}$$

where the Gromov product is computed in $\mathcal{C}(Y)$. Consider $n \geq L$. Let γ_n be a geodesic in $\mathcal{C}(Y)$ with endpoints α_n and α_L. If there exists a vertex v on γ_n with $\pi_X(v) = \varnothing$, then v and ∂X form a multicurve in Y, which implies that

$$\begin{aligned}
&(\alpha_n, \alpha_L)_{\alpha_1} \\
&\quad = \tfrac{1}{2}(d_Y(\alpha_n, \alpha_1) + d_Y(\alpha_L, \alpha_1) - d_Y(\alpha_n, \alpha_L)) \\
&\quad \leq \tfrac{1}{2}\big(d_Y(\alpha_n, v) + d_Y(v, \alpha_1) + d_Y(\alpha_L, v) + d_Y(v, \alpha_1) - (d_Y(\alpha_n, v) + d_Y(v, \alpha_L))\big) \\
&\quad = d_Y(v, \alpha_1) \leq d_Y(v, \partial X) + d_Y(\partial X, \alpha_1) \leq 1 + d_Y(\partial X, \alpha_1).
\end{aligned}$$

But this contradicts (2), so we conclude that $\pi_X(v) \neq \varnothing$ for all v on γ_n. We can now apply Theorems 2.1 and 2.4 to see that, for all $n \geq L$,

$$d_X(\mu_n, \mu_L) \leq \operatorname{diam}_{\mathcal{C}(X)}(\pi_X(\mu_n)) + d_X(\alpha_n, \alpha_L) + \operatorname{diam}_{\mathcal{C}(X)}(\pi_X(\mu_L)) \leq 2 + K_0 + 2,$$

where K_0 is as in Theorem 2.4. Therefore,

$$\begin{aligned}
\operatorname{diam}_{\mathcal{C}(X)}(\pi_X(\mu_1) \cup \pi_X(\mu_2) \cup \cdots) &\leq \operatorname{diam}_{\mathcal{C}(X)}(\pi_X(\mu_1) \cup \cdots \cup \pi_X(\mu_L)) + 2(K_0 + 4) \\
&< \infty. \qquad \square
\end{aligned}$$

2.7 Hierarchically hyperbolic spaces

In [3], Behrstock, Hagen and Sisto define the notion of a hierarchically hyperbolic space. Roughly, a hierarchically hyperbolic space is a quasi-geodesic metric space $\mathcal{X}$, equipped with additional structure which we will call a hierarchically hyperbolic space (HHS) structure. An HHS structure consists of an index set $\mathcal{G}$ and for each $W \in \mathcal{G}$ a Gromov hyperbolic space $\widehat{C}W$ and a projection map $\pi_W \colon \mathcal{X} \to 2^{\widehat{C}W}$. The elements of $\mathcal{G}$ and the projection maps must satisfy a long list of properties. See [3; 2].

The first example of a hierarchically hyperbolic space is $\operatorname{Mod}(S)$, where here $\mathcal{G}$ is the collection of all subsurfaces of S, $\widehat{C}W$ is the curve graph of W for $W \in \mathcal{G}$, and projection π_W is given by composing an orbit map for the action of $\operatorname{Mod}(S)$ on $\widetilde{\mathcal{M}}(S)$ with the subsurface projection map defined in Section 2.3. The works of Masur and Minsky [10; 11], Behrstock [1] and Behrstock, Kleiner, Minsky and Mosher [4] imply that $\operatorname{Mod}(S)$ is a hierarchically hyperbolic space. See [2, Section 11] for details. In fact, the notion of hierarchical hyperbolicity was motivated by a desire to generalize some of the machinery surrounding mapping class groups.

In [3], it is shown that a large class of CAT(0) cube complexes can be equipped with a hierarchically hyperbolic structure, including the universal covers of Salvetti complexes associated to right-angled Artin groups. The CAT(0) cube complex we are primarily concerned with is the Cayley graph X of $A(\Gamma)$ when Γ has no edges (that is, $A(\Gamma)$ is a free group). We equip $A(\Gamma)$ with a hierarchically hyperbolic structure by equipping X with such a structure and then associating $A(\Gamma)$ with X.

2.8 Boundary of hierarchically hyperbolic spaces

In [7], the authors construct a boundary for hierarchically hyperbolic spaces. Here we will describe convergence in this boundary for $\mathrm{Mod}(S)$ and for free groups. With the exception of Theorem 5.3, these will be the only examples we will need.

As a set, the HHS boundary of $\mathrm{Mod}(S)$ is defined as follows:

$$\partial\mathrm{Mod}(S) = \Big\{ \sum_{Y\subseteq S} c_Y \lambda_Y \;\Big|\; c_Y \geq 0 \text{ and } \lambda_Y \in \partial\mathcal{C}(Y) \text{ for all } Y, \sum_{Y\subseteq S} c_Y = 1, \text{ and if } c_{Y'}, c_Y > 0, \text{ then } Y \text{ and } Y' \text{ are disjoint or equal} \Big\}.$$

In [7], the authors define a topology on $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$. In this topology, [7, Definition 2.10] tells us that a sequence of mapping classes $(g_n)_{n\in\mathbb{N}}$ in $\mathrm{Mod}(S)$ converges to a point $\sum_{i=1}^k c_i\lambda_i$ in $\partial\mathrm{Mod}(S)$, where $c_i > 0$ for all i, $\sum_{i=1}^k c_i = 1$ and $\lambda_i \in \partial\mathcal{C}(Y_i)$ for pairwise disjoint subsurfaces $Y_1, \dots, Y_k$, if and only if the following statements hold: for a fixed marking μ on S,

(1) $\lim_{n\to\infty} \pi_{Y_i}(g_n\mu) = \lambda_i$ for each $i = 1, \dots, k$,

(2) $\lim_{n\to\infty} d_{Y_i}(\mu, g_n\mu)/d_{Y_j}(\mu, g_n\mu) = c_i/c_j$ for each $i, j = 1, \dots, k$, and

(3) $\lim_{n\to\infty} d_W(\mu, g_n\mu)/d_{Y_i}(\mu, g_n\mu) = 0$ for every (any) $i = 1, \dots, k$ and every subsurface $W \subseteq S$ that is disjoint from Y_j for all $j = 1, \dots, k$.

Let Γ be a graph with no edges, and let $A(\Gamma)$ be the corresponding free group, equipped with an HHS structure. The HHS boundary of $A(\Gamma)$ will be denoted by $\partial A(\Gamma)$. We do not define $\partial A(\Gamma)$ here because Theorem 4.3 in [7] implies that the identity map $A(\Gamma) \to A(\Gamma)$ extends to a homeomorphism $A(\Gamma)\cup\partial_G A(\Gamma) \to A(\Gamma)\cup\partial A(\Gamma)$. Thus, two sequences in $A(\Gamma)$ converge to the same point in $\partial_G A(\Gamma)$ if and only if they converge to the same point in $\partial A(\Gamma)$. (See Section 2 of [7] for the definition of $\partial A(\Gamma)$.)

Another useful fact on convergence is that $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$ and $A(\Gamma) \cup \partial A(\Gamma)$ are sequentially compact (see Theorem 3.4 of [7]).

To understand Question 1.1 and the statements of our theorems, one last definition is needed.

Definition 2.8 Let $\phi\colon A(\Gamma)\to \mathrm{Mod}(S)$ be an injective homomorphism and let $A(\Gamma)$ and $\mathrm{Mod}(S)$ be equipped with any fixed HHS structures. We say that ϕ *extends continuously to a map* $\partial A(\Gamma)\to \partial\mathrm{Mod}(S)$ if there exists a function $\bar{\phi}\colon A(\Gamma)\cup\partial A(\Gamma)\to \mathrm{Mod}(S)\cup\partial\mathrm{Mod}(S)$ such that

(1) $\bar{\phi}|_{A(\Gamma)}=\phi$,

(2) $\bar{\phi}(\partial A(\Gamma))\subseteq\partial\mathrm{Mod}(S)$, and

(3) $\bar{\phi}$ is continuous at each point in $\partial A(\Gamma)$.

Remark 2.9 To establish that $\phi\colon A(\Gamma)\to\mathrm{Mod}(S)$ extends continuously, it is enough to show that, for all $x\in\partial A(\Gamma)$, given any two sequences (x_n) and (y_n) in $A(\Gamma)$ that converge to x, we have that $(\phi(x_n))$ and $(\phi(y_n))$ converge to the same point in $\partial\mathrm{Mod}(S)$. This follows from a diagonal sequence argument (see the end of the proof of Theorem 5.6 in [7] for details).

3 Lemmas on subsurface projections

The following lemmas are the heart of our proof of Theorem 1.2:

Lemma 3.1 *Suppose X and Y are disjoint subsurfaces of S, and if Y is an annulus, then the core of Y is not contained in ∂X. If μ and μ' are markings and $f\in\mathrm{Mod}(S)$ a mapping class supported on X, then $|d_Y(\mu, f(\mu'))-d_Y(\mu,\mu')|\leq 4$.*

Proof If Y is not an annulus, then $\pi_Y(f(\mu'))=\pi_Y(\mu')$, so the claim clearly holds. Assume then that Y is an annular subsurface of S with core α, and let $\mathrm{int}(Y)\to S$ be the associated covering. If X is not an annulus, define Z to be the component of $S-X$ that contains α. If X is an annulus with core β, let Z be the component of S containing α after removing a small regular neighborhood of β. Let $\widetilde{\alpha}$ be the component of the preimage of α in $\mathrm{int}(Y)$ that is a closed curve. Let $\widetilde{Z}$ be the component of the preimage of Z in $\mathrm{int}(Y)$ that contains $\widetilde{\alpha}$.

Abusing notation, we let f denote a representative in the isotopy class of f that fixes Z pointwise. Consider the lift of f to $\mathrm{int}(Y)$ that fixes a point on $\widetilde{\alpha}$, and thus fixes

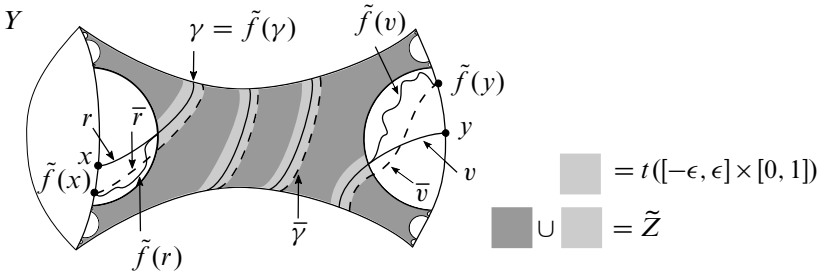


Figure 1: The arc β' is the concatenation of r, γ and v. The concatenation of $\overline{r}, \overline{\gamma}$ and $\overline{v}$ is equivalent to $\tilde{f}(\beta')$, and that representative of $\tilde{f}(\beta')$ intersects β' at most once (drawn is the exactly once case). See the proof of Lemma 3.1.

$\widetilde{Z}$ pointwise. Let $\tilde{f}\colon Y \to Y$ denote the continuous extension of that lift. Consider $\beta' \in \pi_Y(\mu')$. We will show

$$d_{\mathcal{C}(Y)}(\beta', \tilde{f}(\beta')) \leq 2. \tag{3}$$

This will complete the proof because (3), the triangle inequality and Theorem 2.1 imply that

$$\begin{aligned} &|d_Y(\mu, f(\mu')) - d_Y(\mu, \mu')| \\ &\qquad \leq d_Y(\mu', f(\mu')) \\ &\qquad \leq \operatorname{diam}_{\mathcal{C}(Y)}(\pi_Y(\mu')) + d_{C(Y)}(\beta', \tilde{f}(\beta')) + \operatorname{diam}_{\mathcal{C}(Y)}(\pi_Y(f(\mu'))) \\ &\qquad \leq 1 + 2 + 1 = 4. \end{aligned}$$

Inequality (3) holds if β' is contained in $\widetilde{Z}$ because, in that case, $\tilde{f}(\beta') = \beta'$. Thus, we assume β' is not contained in $\widetilde{Z}$. We break β' up into three parts. Let γ be the largest subpath of β' contained in $\widetilde{Z}$. Let x and y denote the endpoints of β' on ∂Y. Removing γ from β' yields rays r and v that limit to x and y, respectively.

We now construct an arc equivalent to $\tilde{f}(\beta')$ that intersects β' at most once. Figure 1 illustrates the construction. Let $t\colon [-\epsilon, \epsilon] \times [0, 1] \to Y$ be a small tubular neighborhood of γ such that $t|_{\{0\}\times[0,1]} = \gamma$ and $t([-\epsilon, \epsilon] \times \{0, 1\}) \subseteq \partial\widetilde{Z}$. Let R and V denote the components of $\operatorname{int}(Y) - \widetilde{Z}$ containing r and v, respectively. Because $\tilde{f}$ fixes $\widetilde{Z}$ pointwise, $\tilde{f}$ restricts to homeomorphisms of both R and V, implying that $\tilde{f}(r)$ and $\tilde{f}(v)$ are contained in R and V, respectively. Consequently, in R there exists a ray $\overline{r}$ based at $t(-\epsilon, 0)$ or $t(\epsilon, 0)$ that limits to $\tilde{f}(x)$ and is disjoint from r. If $\overline{r}$ is based at $t(-\epsilon, 0)$, define $\overline{\gamma} = t|_{\{-\epsilon\}\times[0,1]}$. Otherwise, define $\overline{\gamma} = t|_{\{\epsilon\}\times[0,1]}$. Choose $\overline{v}$ to be an arc in V from $\overline{\gamma}(1)$ to $\tilde{f}(y)$ that intersects v at most once. Observe that the arc

obtained by concatenating $\overline{r}$, $\overline{\gamma}$ and $\overline{v}$ is equivalent to $\tilde{f}(\beta')$ and intersects β' at most once.

Therefore, by (1) we have $d_{\mathcal{C}(Y)}(\tilde{f}(\beta'), \beta') = 1 + |\tilde{f}(\beta') \cdot \beta'| \leq 2$, as desired. □

Lemma 3.2 *Given a homomorphism $\phi\colon A(\Gamma) \to \mathrm{Mod}(S)$ and a marking μ on S, there exists a constant $M \geq 1$ such that the following holds: Let $y_1 \dots y_n \in A(\Gamma)$, where each $y_i \in V(\Gamma)$. Then $d_W(\mu, \phi(y_1 \dots y_n)\mu) \leq Mn$ for all subsurfaces $W \subseteq S$.*

Proof Define $M = 4\max\{d_{\widetilde{\mathcal{M}}(S)}(\mu, \phi(x)\mu) : x \in V(\Gamma)\}$. By the triangle inequality and Theorem 2.2,

$$\begin{aligned} d_W(\mu, \phi(y_1 \dots y_n)\mu) &\leq \sum_{i=1}^{n} d_W(\phi(y_1 \dots y_{i-1})\mu, \phi(y_1 \dots y_i)\mu) \\ &\leq \sum_{i=1}^{n} 4d_{\widetilde{\mathcal{M}}(S)}(\phi(y_1 \dots y_{i-1})\mu, \phi(y_1 \dots y_i)\mu) \\ &= \sum_{i=1}^{n} 4d_{\widetilde{\mathcal{M}}(S)}(\mu, \phi(y_i)\mu) \leq Mn. \end{aligned}$$ □

Lemma 3.3 *Let $\phi\colon A(\Gamma) \to \mathrm{Mod}(S)$ be a homomorphism. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of elements in $A(\Gamma)$ and μ a marking on S. Suppose, for some subsurface $W \subseteq S$, there exist constants $A \geq 1$ and $B \geq 0$ that do not depend on n such that $d_W(\mu, \phi(g_n)\mu) \succ_{A,B} \|g_n\|$, where $\|g_n\|$ denotes the word length of g_n with respect to the standard generating set $V(\Gamma)$ for $A(\Gamma)$. Further suppose that $\lim_{n\to\infty} \|g_n\| = \infty$ and that $(\pi_W(\phi(g_n)\mu))_{n\in\mathbb{N}}$ converges to some point λ_W in $\partial\mathcal{C}(W)$. Then all accumulation points of $(\phi(g_n))_{n\in\mathbb{N}}$ in $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$ are in $\partial\mathrm{Mod}(S)$ and are of the form $\sum_{Y\subseteq S} c_Y\lambda_Y$, where $c_W > 0$.*

Proof After passing to a subsequence, we may assume that $(\phi(g_n))_{n\in\mathbb{N}}$ converges. By assumption, $\lim_{n\to\infty} d_W(\mu, \phi(g_n)\mu) = \infty$. Combine this with Theorem 2.2 to see that $\lim_{n\to\infty} d_{\widetilde{\mathcal{M}}(S)}(\mu, \phi(g_n)\mu) = \infty$. Because $\widetilde{\mathcal{M}}(S)$ is quasi-isometric to $\mathrm{Mod}(S)$ via orbit maps, it follows that $\lim_{n\to\infty} d_{\mathrm{Mod}(S)}(1, \phi(g_n)) = \infty$. Thus, it must be that $\lim_{n\to\infty} \phi(g_n) \in \partial\mathrm{Mod}(S)$.

Suppose $\lim_{n\to\infty} \phi(g_n) = \sum_{Y\subseteq S} c_Y\lambda_Y$ for constants $c_Y \geq 0$ and $\lambda_Y \in \partial\mathcal{C}(Y)$. We will now argue that $c_W > 0$. Let $Z \subseteq S$ be such that $c_Z > 0$. If $W = Z$, we are done. So we assume $W \neq Z$. By definition of the topology on $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$,

we have that $\lim_{n\to\infty} \pi_Z(\phi(g_n)\mu) = \lambda_Z$. If $W \subsetneq Z$, then Corollary 2.7 implies that $\operatorname{diam}_{\mathcal{C}(W)}(\pi_W(\phi(g_1)\mu) \cup \pi_W(\phi(g_2)\mu) \cup \cdots) < \infty$. But this cannot be since $\pi_W(\phi(g_n)\mu) \to \lambda_W \in \partial\mathcal{C}(W)$. Similarly, we cannot have $Z \subsetneq W$ for then Corollary 2.7 implies that $\operatorname{diam}_{\mathcal{C}(Z)}(\pi_Z(\phi(g_1)\mu) \cup \pi_Z(\phi(g_2)\mu) \cup \cdots) < \infty$, contradicting that $\pi_Z(\phi(g_n)\mu) \to \lambda_Z \in \partial\mathcal{C}(Z)$. Now suppose that $Z \pitchfork W$. Then by Theorem 2.3, after passing to a subsequence, we have that

$$d_W(\partial Z, \phi(g_n)\mu) \le 10 \quad \text{for all } n \qquad \text{or} \qquad d_Z(\partial W, \phi(g_n)\mu) \le 10 \quad \text{for all } n.$$

If $d_W(\partial Z, \phi(g_n)\mu) \le 10$ for all n, then, for all n,

$$d_W(\mu, \phi(g_n)\mu) \le d_W(\mu, \partial Z) + d_W(\partial Z, \phi(g_n)\mu) \le d_W(\mu, \partial Z) + 10,$$

contradicting that $\pi_W(\phi(g_n)\mu) \to \lambda_W \in \partial\mathcal{C}(W)$. Similarly, if $d_Z(\partial W, \phi(g_n)\mu) \le 10$ for all n, then $d_Z(\mu, \phi(g_n)\mu)$ is bounded independent of n, contradicting that $\pi_Z(\phi(g_n)\mu) \to \lambda_Z \in \partial\mathcal{C}(Z)$. So it is not the case that $Z \pitchfork W$. Therefore it must be that W and Z are disjoint for all $Z \subseteq S$ with $c_Z > 0$.

Fix $Z \subseteq S$ with $c_Z > 0$. Lemma 3.2 together with the fact that $d_W(\mu, \phi(g_n)\mu) \succ_{A,B} \|g_n\|$ implies that

$$\frac{d_W(\mu, \phi(g_n)\mu)}{d_Z(\mu, \phi(g_n)\mu)} \ge \frac{\frac{1}{A}\|g_n\| - B}{M\|g_n\|}, \tag{4}$$

where $M \ge 1$ is as in Lemma 3.2. Since $\|g_n\| \to \infty$, (4) implies

$$\lim_{n\to\infty} \frac{d_W(\mu, \phi(g_n)\mu)}{d_Z(\mu, \phi(g_n)\mu)} \ge \lim_{n\to\infty} \frac{\frac{1}{A}\|g_n\| - B}{M\|g_n\|} > 0.$$

Therefore, by definition of the topology of $\operatorname{Mod}(S) \cup \partial\operatorname{Mod}(S)$, we have $c_W > 0$, as desired. □

4 Clay–Leininger–Mangahas RAAGs

In this section, we prove the first part of Theorem 1.2. We begin with a description of a CLM embedding $\phi\colon A(\Gamma) \to \operatorname{Mod}(S)$.

Embedding construction Let Γ be the graph with vertex set $V(\Gamma) = \{a, b\}$ and no edges. Let $S = \mathbb{H}^2/\Lambda$, X_a and X_b be the surfaces indicated in Figure 2. For short, let X_{ab} denote $X_a \cup X_b$. Let $\widetilde{S - X_{ab}}$ be a component of the preimage of $S - X_{ab}$ in $\mathbb{H}^2$, and let $\widetilde{\partial X_{ab}}$ be a geodesic in $\mathbb{H}^2$ that is in the boundary of $\widetilde{S - X_{ab}}$.

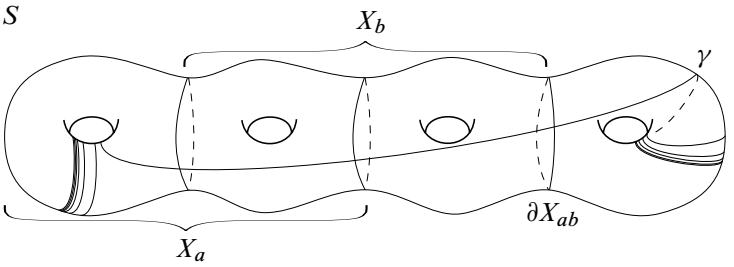


Figure 2: Overlapping subsurfaces X_a and X_b of surface S, curve ∂X_{ab} and bi-infinite simple geodesic γ

Let $\widetilde{\gamma}$ be a geodesic in $\mathbb{H}^2$ that links with $\widetilde{\partial X_{ab}}$ and maps to a simple bi-infinite geodesic γ in S. Further suppose that $\widetilde{\gamma} \cap (\widetilde{S - X_{ab}})$ is an infinite ray and let p be its endpoint on $\partial\mathbb{H}^2$. For example, take γ to be the simple bi-infinite geodesic in S with one end spiraling around a curve essential in $S - X_{ab}$ and the other end spiraling around a curve in X_a, as in Figure 2, and take $\widetilde{\gamma}$ to be an appropriate lift of γ. Choose $f_b \in \mathrm{Mod}(S)$ so that f_b is pseudo-Anosov on X_b. To simplify arguments, we abuse notation and let f_b denote a representative in the isotopy class of f_b that fixes all points outside X_b. This ensures that $\tilde{f}_b$ fixes $\widetilde{S - X_{ab}}$ pointwise, where $\tilde{f}_b\colon \mathbb{H}^2 \to \mathbb{H}^2$ is the lift of f_b fixing some point on $\widetilde{\partial X_{ab}}$. Thus, the extension of $\tilde{f}_b$ to $\partial\mathbb{H}^2$ fixes pointwise p and the endpoints x and y of $\widetilde{\partial X_{ab}}$. Additionally, we choose f_b to have the following properties:

(1) $\tilde{f}_b(\widetilde{\gamma})$ links with $h(\widetilde{\gamma})$, where $h \in \Lambda$ is a primitive isometry with axis $\widetilde{\partial X_{ab}}$; and

(2) $\tau_{X_b}(f_b) \geq C$, where C is as in Theorem 2.5.

We note that a pseudo-Anosov on X_b satisfying (1) can be obtained from any mapping class that is pseudo-Anosov on X_b by postcomposing with some number of Dehn twists (or inverse Dehn twists) about ∂X_{ab}. Finally, a pseudo-Anosov on X_b satisfying (1) and (2) can be obtained from one satisfying (1) by passing to a sufficiently high power.

Let $f_a \in \mathrm{Mod}(S)$ be any mapping class that is pseudo-Anosov on X_a and satisfies $\tau_{X_a}(f_a) \geq C$. Theorem 2.5 says that the homomorphism

$$\phi\colon A(\Gamma) \to \mathrm{Mod}(S) \quad \text{defined by } \phi(a) = f_a \text{ and } \phi(b) = f_b$$

is a quasi-isometric embedding.

Equip $A(\Gamma)$ with any HHS structure. In the remainder of this section, we will prove the following theorem, which proves the first part of Theorem 1.2:

Theorem 4.1 *The sequences $(a^n)_{n\in\mathbb{N}}$ and $(a^n b^n)_{n\in\mathbb{N}}$ converge to the same point in $\partial A(\Gamma)$, but $(\phi(a^n))_{n\in\mathbb{N}}$ and $(\phi(a^n b^n))_{n\in\mathbb{N}}$ do not converge to the same point in $\partial\mathrm{Mod}(S)$.*

We will divide the proof of Theorem 4.1 into two propositions.

Proposition 4.2 *The sequences $(a^n)_{n\in\mathbb{N}}$ and $(a^n b^n)_{n\in\mathbb{N}}$ converge to the same point in $\partial A(\Gamma)$.*

Proof Let X be the Cayley graph of $A(\Gamma)$. By the discussion in Section 2.8, to show that $(a^n)_{n\in\mathbb{N}}$ and $(a^n b^n)_{n\in\mathbb{N}}$ converge to the same point in $\partial A(\Gamma)$, it is enough to show that they converge to the same point in $\partial_G X$. Now the Gromov product satisfies

$$(a^i, a^j b^j)_1 = \min(i, j) \to \infty \quad \text{as } i, j \to \infty.$$

Thus, $\lim_{n\to\infty} a^n = \lim_{n\to\infty} a^n b^n$ in $\partial_G X$, as desired. □

Throughout the rest of this section, μ will denote a fixed marking on S. To continue, we require the following lemma:

Lemma 4.3 *There exist constants $A \geq 1$ and $B \geq 0$ such that for all $n \geq 1$ we have $d_{\partial X_{ab}}(\mu, \phi(a^n b^n)\mu) \succ_{A,B} n$. Consequently, after passing to a subsequence, $(\pi_{\partial X_{ab}}(\phi(a^n b^n)\mu))_{n\in\mathbb{N}}$ converges to a point in $\partial\mathcal{C}(\partial X_{ab})$.*

Proof We begin by establishing the following claim:

Claim 1 *Let $n \geq 1$. Then $\tilde{f}_b^n(\widetilde{\gamma})$ has endpoint p and links with $h^i(\widetilde{\gamma})$ for all $1 \leq i \leq n$.*

Proof of Claim 1 By our choice of $\tilde{f}_b$ and $\widetilde{\gamma}$, we know the claim holds for $n = 1$. Let $n \geq 2$. Inductively, suppose that $\tilde{f}_b^{n-1}(\widetilde{\gamma})$ has endpoint p and links with $h^i(\widetilde{\gamma})$ for all $1 \leq i \leq n-1$. Let I be the interval in $\partial\mathbb{H}^2$ that connects the endpoints of $\widetilde{\partial X_{ab}}$ and does not contain p, oriented from the repelling fixed point of h to the attracting fixed point. We will use interval notation when speaking about connected subsets of I. Now $\tilde{f}_b$ extends continuously to a homeomorphism of $\partial\mathbb{H}^2$, which we will also denote by $\tilde{f}_b$, and because $\tilde{f}_b$ fixes the endpoints of $\widetilde{\partial X_{ab}}$, this extension restricts to a homeomorphism of I. Let z be the endpoint of $\widetilde{\gamma}$ in I, and let $x \in \partial I$ be the attracting fixed point of h. Because $\tilde{f}_b^{n-1}(\widetilde{\gamma})$ links with $h^i(\widetilde{\gamma})$ for all $1 \leq i \leq n-1$ and has endpoint p, we have

$$(\tilde{f}_b^{n-1}(z), x] \subseteq (h^i(z), x] \quad \text{for all } 0 \leq i \leq n-1. \tag{5}$$

Since $\tilde{f}_b(\tilde{\gamma})$ has endpoint p and links with $h(\tilde{\gamma})$, it must be that $\tilde{f}_b(z) \in (hz, x]$. It follows from this, the fact that $\tilde{f}_b$ and h fix x, that $\tilde{f}_b$ and h commute by uniqueness of map lifting, and (5) that, for all $0 \leq i \leq n-1$,

$$\begin{aligned}(6)\qquad \tilde{f}_b^n(z) = \tilde{f}_b^{n-1}(\tilde{f}_b(z)) \in \tilde{f}_b^{n-1}(h(z), x] = h(\tilde{f}_b^{n-1}(z), x] \subseteq h(h^i(z), x] \\ = (h^{i+1}(z), x].\end{aligned}$$

Because $\tilde{f}_b$ fixes p, we have $\tilde{f}_b^n(p) = p$. This combined with (6) implies that $\tilde{f}_b^n(\tilde{\gamma})$ links with $h^{i+1}(\tilde{\gamma})$ for all $0 \leq i \leq n-1$, proving Claim 1. □

By Claim 1, after replacing $\tilde{f}_b^n(\tilde{\gamma})$ with the geodesic connecting its endpoints, the images of $\tilde{f}_b^n(\tilde{\gamma})$ and $\tilde{\gamma}$ in $(\overline{\mathbb{H}}^2 - \{x, y\})/\langle h \rangle$ intersect each other at least n times, and all these intersections have the same sign. Now apply (1) to see that

$$d_{\partial X_{ab}}(\gamma, \phi(b^n)\gamma) \geq n + 1.$$

It follows that

$$\begin{aligned}(7)\quad d_{\partial X_{ab}}(\mu, \phi(b^n)\mu) &\geq d_{\partial X_{ab}}(\gamma, \phi(b^n)\gamma) - d_{\partial X_{ab}}(\mu, \gamma) - d_{\partial X_{ab}}(\phi(b^n)\mu, \phi(b^n)\gamma) \\ &\geq n + 1 - 2d_{\partial X_{ab}}(\mu, \gamma).\end{aligned}$$

Lemma 3.1 says that $|d_{\partial X_{ab}}(\mu, \phi(a^n b^n)\mu) - d_{\partial X_{ab}}(\mu, \phi(b^n)\mu)| \leq 4$. This together with (7) implies that

$$d_{\partial X_{ab}}(\mu, \phi(a^n b^n)\mu)) \succ n.$$

From this and the fact that $\mathcal{C}(\partial X_{ab})$ is quasi-isometric to $\mathbb{R}$, it is immediate that $(\pi_{\partial X_{ab}}(\phi(a^n b^n)\mu))_{n\in\mathbb{N}}$ has a subsequence converging to a point in $\partial\mathcal{C}(\partial X_{ab})$. □

Proposition 4.4 *The sequences $(\phi(a^n))_{n\in\mathbb{N}}$ and $(\phi(a^n b^n))_{n\in\mathbb{N}}$ do not converge to the same point in* $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$.

Proof After passing to a subsequence, we may assume $(\phi(a^n))_{n\in\mathbb{N}}$ and $(\phi(a^n b^n))_{n\in\mathbb{N}}$ converge to points p and q, respectively, in $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$ and, by Lemma 4.3, that $(\pi_{\partial X_{ab}}(\phi(a^n b^n)\mu))_{n\in\mathbb{N}}$ converges to a point in $\partial\mathcal{C}(\partial X_{ab})$. Lemmas 3.3 and 4.3 imply that q is in $\partial\mathrm{Mod}(S)$. Say $q = \sum_{Y\subseteq S} c_Y^q \lambda_Y^q$, where $c_Y^q \geq 0$ and $\lambda_Y^q \in \partial\mathcal{C}(Y)$ for all $Y \subseteq S$. Then Lemmas 3.3 and 4.3 also imply that $c_{\partial X_{ab}}^q > 0$.

Now, if p were in $\mathrm{Mod}(S)$, then we would be done since clearly then $p \neq q$. So we will assume that $p \in \partial\mathrm{Mod}(S)$, and let $p = \sum_{Y\subseteq S} c_Y^p \lambda_Y^p$. Now observe that, by Lemma 3.1 and Theorem 2.1,

$$d_{\partial X_{ab}}(\mu, \phi(a^n)\mu) \leq d_{\partial X_{ab}}(\mu, \mu) + 4 \leq 5.$$

Thus, $(\pi_{\partial X_{ab}}(\phi(a^n)\mu))_{n\in\mathbb{N}}$ does not limit to a point on $\partial\mathcal{C}(\partial X_{ab})$. So, by definition of the topology of $\mathrm{Mod}(S)\cup\partial\mathrm{Mod}(S)$, it must be that $c^p_{\partial X_{ab}}=0$. Since $c^q_{\partial X_{ab}}>0$, we see that $p\neq q$, which completes the proof. □

5 Koberda RAAGs

In this section we complete the proof of Theorem 1.2. Following this, we will discuss how to use similar techniques to prove a large class of Koberda embeddings do not extend.

Let α and β be the pair of intersecting curves on $S=\mathbb{H}^2/\Lambda$ depicted in Figure 3. Let Γ be the graph with $V(\Gamma)=\{a,b\}$ and no edges. For sufficiently large N, Theorem 2.6 says that the homomorphism

$$\phi\colon A(\Gamma)\to\mathrm{Mod}(S)\quad\text{defined by }\phi(a)={T_\alpha}^N\text{ and }\phi(b)={T_\beta}^N$$

is injective, where T_α and T_β denote Dehn twists about α and β, respectively. Throughout this section, we let μ be a fixed marking on S. Equip $A(\Gamma)$ with an HHS structure.

In this section we prove the following theorem, which will complete the proof of Theorem 1.2:

Theorem 5.1 *There exists $g\in A(\Gamma)$ such that the sequences $(a^n)_{n\in\mathbb{N}}$ and $(a^ng^n)_{n\in\mathbb{N}}$ converge to the same point in $\partial A(\Gamma)$, but $(\phi(a^n))_{n\in\mathbb{N}}$ and $(\phi(a^ng^n))_{n\in\mathbb{N}}$ do not converge to the same point in $\partial\mathrm{Mod}(S)$.*

As a step towards proving Theorem 5.1, we prove the following lemma, in which we construct $g\in A(\Gamma)$:

Lemma 5.2 *There exist constants $A\geq 1$ and $B\geq 0$ and a word $g\in A(\Gamma)$ such that for all $n\geq 1$ we have $d_\eta(\mu,\phi(a^ng^n)\mu)\succ_{A,B} n$, where η is the curve shown in Figure 3. Consequently, after passing to a subsequence, $(\pi_\eta(\phi(a^ng^n)\mu))_{n\in\mathbb{N}}$ converges to a point in $\partial\mathcal{C}(\eta)$.*

Proof We will prove that there exist constants c_1, c_2 and c_3 such that $g=b^{c_1}a^{c_2}b^{c_3}$ has the desired properties.

Let A be the annulus in Figure 3. Let $\widetilde{A}$ be a component of the preimage of A in $\mathbb{H}^2$. Let $\widetilde{\beta}$ be a component of the preimage of β such that a segment of $\widetilde{\beta}$ is in the boundary of $\widetilde{A}$, and let $\widetilde{\eta}$ denote the component of the preimage of η in the boundary of $\widetilde{A}$. Let

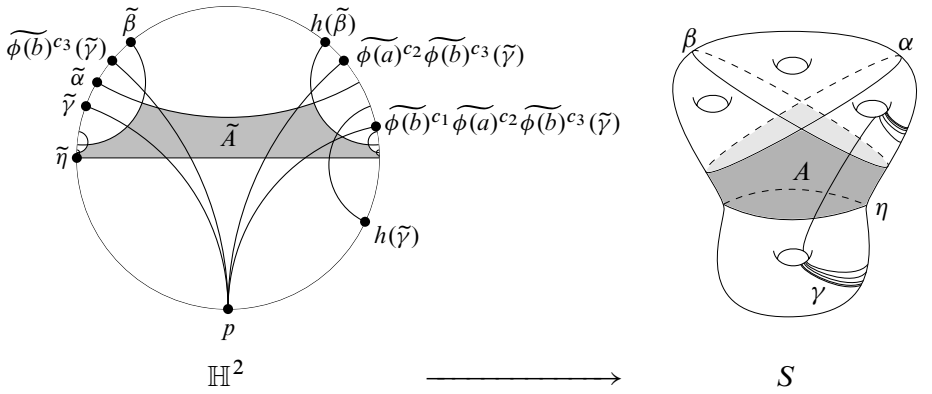


Figure 3: Curves α, β and η bounding an annulus A, and simple bi-infinite geodesic γ on surface S, and the universal cover $\mathbb{H}^2$ of S as in Lemma 5.2

$h \in \Lambda$ be a primitive isometry with axis $\widetilde{\eta}$. Let $\widetilde{\alpha}$ be the component of the preimage of α that links with $\widetilde{\beta}$ and $h(\widetilde{\beta})$ and contains a segment that is in the boundary of $\widetilde{A}$.

Let Y_α be the component of $S-\alpha$ that contains η. To simplify arguments, we let $\phi(a)$ denote a representative in its isotopy class that fixes Y_α pointwise. Let $\widetilde{\phi(a)}\colon \mathbb{H}^2 \to \mathbb{H}^2$ be the lift of $\phi(a)$ that fixes some point on $\widetilde{\alpha}$. Similarly define Y_β to be the component of $S-\beta$ containing η, choose a representative in the isotopy class of $\phi(b)$ that fixes Y_β pointwise, and let $\widetilde{\phi(b)}$ be the lift of $\phi(b)$ that fixes some point on $\widetilde{\beta}$. It then follows that

$$\widetilde{\phi(a)} = \mathbb{1} \quad \text{on } \widetilde{Y}_\alpha \qquad \text{and} \qquad \widetilde{\phi(b)} = \mathbb{1} \quad \text{on } \widetilde{Y}_\beta,$$

where for $i \in \{\alpha, \beta\}$ we let $\widetilde{Y}_i$ denote the component of the preimage of Y_i in $\mathbb{H}^2$ whose boundary contains $\tilde{\imath}$. Observe that $\widetilde{\phi(i)}$ fixes the endpoints of $\widetilde{\eta}$ for $i \in \{a, b\}$.

Choose a geodesic $\widetilde{\gamma}$ in $\mathbb{H}^2$ that links with both $\widetilde{\beta}$ and $\widetilde{\eta}$ and maps to a simple bi-infinite geodesic in S. Further, suppose that $\widetilde{\gamma} \cap \widetilde{Y}_\alpha \cap \widetilde{Y}_\beta$ is an infinite ray, and let p denote its endpoint on $\partial\mathbb{H}^2$. For example, take γ to be the simple bi-infinite geodesic in S with one end spiraling around a curve essential in $Y_\alpha \cap Y_\beta$ and the other end spiraling around a curve essential in $S - Y_\beta$ as in Figure 3, and take $\widetilde{\gamma}$ to be an appropriate component of the preimage of γ. Observe that $\widetilde{\phi(a)}$ and $\widetilde{\phi(b)}$ must fix p.

Now choose $c_3 \in \mathbb{Z}$ so that $\widetilde{\phi(b)}^{c_3}(\widetilde{\gamma})$ links with $\widetilde{\alpha}$. Then choose $c_2 \in \mathbb{Z}$ so that $\widetilde{\phi(a)}^{c_2}\widetilde{\phi(b)}^{c_3}(\widetilde{\gamma})$ links with $h(\widetilde{\beta})$. Finally, choose $c_1 \in \mathbb{Z}$ so that

$$\widetilde{\phi(b)}^{c_1}\widetilde{\phi(a)}^{c_2}\widetilde{\phi(b)}^{c_3}(\widetilde{\gamma})$$

links with $h(\widetilde{\gamma})$. See Figure 3.

To simplify notation, define

$$g = b^{c_1}a^{c_2}b^{c_3} \in A(\Gamma) \quad \text{and} \quad \widetilde{\phi(g)} = \widetilde{\phi(b)}^{c_1}\widetilde{\phi(a)}^{c_2}\widetilde{\phi(b)}^{c_3}.$$

As in Lemma 4.3, we have that $\widetilde{\phi(g)}^n(\widetilde{\gamma})$ has endpoint p and links with $h^i(\widetilde{\gamma})$ for all $1 \le i \le n$, implying that $d_\eta(\gamma, \phi(g^n)\gamma) \ge n+1$. It follows that

$$\begin{aligned} d_\eta(\mu, \phi(g^n)\mu) &\ge d_\eta(\gamma, \phi(g^n)\gamma) - d_\eta(\mu, \gamma) - d_\eta(\phi(g^n)\mu, \phi(g^n)\gamma) \\ &\ge n + 1 - 2d_\eta(\mu, \gamma). \end{aligned} \tag{8}$$

Now Lemma 3.1 says that $|d_\eta(\mu, \phi(a^n g^n)\mu) - d_\eta(\mu, \phi(g^n)\mu)| \le 4$. This together with (8) implies that $d_\eta(\mu, \phi(a^n g^n)\mu) \succ n$. From this and the fact that $\mathcal{C}(\eta)$ is quasi-isometric to $\mathbb{R}$, it is immediate that $(\pi_\eta(\phi(a^n g^n)\mu))_{n\in\mathbb{N}}$ has a subsequence converging to a point in $\partial\mathcal{C}(\eta)$. □

Proof of Theorem 5.1 Let $g \in A(\Gamma)$ be as in Lemma 5.2. By the discussion in Section 2.8, to show that $(a^n)_{n\in\mathbb{N}}$ and $(a^n g^n)_{n\in\mathbb{N}}$ converge to the same point $\partial A(\Gamma)$ it is enough to show that they converge to the same point in $\partial_G X$, where X is the Cayley graph of $A(\Gamma)$. Now the Gromov product satisfies

$$(a^i, a^j g^j)_1 = (a^i, a^j (b^{c_1}a^{c_2}b^{c_3})^j)_1 = \min(i, j) \to \infty \quad \text{as } i, j \to \infty.$$

Therefore, $\lim_{n\to\infty} a^n = \lim_{n\to\infty} a^n g^n$ in $\partial_G X$, as desired.

To finish this proof, we mimic the proof of Proposition 4.4. Replacing b with g, and ∂X_{ab} with η, and Lemma 4.3 with Lemma 5.2, we find that $(\phi(a^n))_{n\in\mathbb{N}}$ and $(\phi(a^n g^n))_{n\in\mathbb{N}}$ do not converge to the same point in $\partial\mathrm{Mod}(S)$. □

Our techniques used to prove Theorem 5.1 can be used to prove a more general statement on nonexistence of boundary maps for right-angled Artin groups that are not necessarily free groups. To prove this more general statement, one needs to understand HHS structures for all right-angled Artin groups. In the following theorem, by a *standard HHS structure on* $A(\Gamma)$, we mean one induced by a factor system generated by a rich family of subgraphs of Γ. We refer the reader to [3], specifically Proposition 8.3 and Remark 13.2, for details and to [7] for a general description of the corresponding HHS boundary. In the proof of the following theorem, we freely use definitions and notations used in [3; 7].

Theorem 5.3 *Let $\{\alpha_1, \dots, \alpha_k\}$ be any collection pairwise distinct of curves in S. Let Γ be the graph with $V(\Gamma) = \{s_1, \dots, s_k\}$ and $s_i s_j$ an edge in Γ if and only if*

$i(\alpha_i, \alpha_j) = 0$. Give $A(\Gamma)$ a standard HHS structure, or if $A(\Gamma)$ is a free group, any HHS structure. If there exist distinct intersecting curves α_i and α_j that do not fill S, then any corresponding Koberda embedding $\phi\colon A(\Gamma) \to \mathrm{Mod}(S)$ does not extend continuously to a map $\partial A(\Gamma) \to \mathrm{Mod}(S)$.

Proof Consider the subgraph Λ of Γ with $V(\Lambda) = \{s_i, s_j\}$. Contained in the Salvetti complex S_Γ associated to Γ there is a subcomplex that is the Salvetti complex associated to $A(\Lambda)$. We let $\widetilde{S}_\Lambda$ denote the lift of this subcomplex to the universal cover $\widetilde{S}_\Gamma$ of S_Γ that contains 1. Let $\mathcal{R}$ be a rich family of induced subgraphs of Γ, and let $\mathcal{F}$ be the corresponding factor system in $\widetilde{S}_\Gamma$. Lemma 8.4 of [3] tells us that

$$\mathcal{F}' = \{F \cap \widetilde{S}_\Lambda : F \in \mathcal{F}\}$$

is a factor system in $\widetilde{S}_\Lambda$. Associating $A(\Gamma)$ and $A(\Lambda)$ with $\widetilde{S}_\Gamma$ and $\widetilde{S}_\Lambda$, respectively, we equip each with the HHS structures corresponding to their respective factor systems. We first argue that the inclusion map $A(\Lambda) \to A(\Gamma)$ extends continuously to a map $\partial A(\Lambda) \to \partial A(\Gamma)$. If ϕ extends continuously to a map $\partial A(\Gamma) \to \partial\mathrm{Mod}(S)$, it will follow that $A(\Lambda) \to \mathrm{Mod}(S)$ extends continuously to a map $\partial A(\Lambda) \to \partial\mathrm{Mod}(S)$; we will show that this is impossible.

First, consider $A(\Lambda) \to A(\Gamma)$. Given $U \in \mathcal{F}'$ such that U is not a 0–cube, define $\pi(U)$ to be the parallelism class of the $\subseteq$–minimal $F \in \mathcal{F}$ such that $U = F \cap \widetilde{S}_\Lambda$. Observe that U and V are nested (respectively orthogonal) if and only if $\pi(U)$ and $\pi(V)$ are nested (respectively orthogonal). This together with Lemma 10.11 of [7] implies that $(A(\Lambda) \to A(\Gamma), \pi)$ is a hieromorphism. Theorem 5.6 of [7] gives a condition guaranteeing that a hieromorphism extends continuously. In our case, if the following claims are true, we can apply Theorem 5.6 to conclude that $A(\Lambda) \to A(\Gamma)$ extends continuously.

Claim 1 *π is injective.*

Proof of Claim 1 Suppose $U, V \in \mathcal{F}'$ and $\pi(U) = \pi(V)$. Then $\pi(U) \sqsubseteq \pi(V)$ and $\pi(V) \sqsubseteq \pi(U)$. Thus, $U \subseteq V$ and $V \subseteq U$, implying $U = V$, as desired. □

Claim 2 *If $[F] \in \overline{\mathcal{F}}$ is not a class of 0–cubes and there exists no $U \in \mathcal{F}'$ satisfying $\pi(U) = [F]$, then $\mathrm{diam}_{\widehat{C}F}(\pi_F(\widetilde{S}_\Lambda))$ is bounded above uniformly for some (any) $F \in [F]$.*

Proof of Claim 2 Let $[F] \in \overline{\mathcal{F}}$ be as in Claim 2. First, suppose there exists $F \in [F]$ such that $F \cap \widetilde{S}_\Lambda \neq \varnothing$. By Lemma 8.5 in [3], we have $\mathfrak{g}_F(\widetilde{S}_\Lambda) \subseteq F \cap \widetilde{S}_\Lambda$. If $F \cap \widetilde{S}_\Lambda$

is a 0–cube, then $\operatorname{diam}_{\widehat{C}F}(\pi_F(\widetilde{S}_\Lambda)) \leq 1$, so the claim holds. Otherwise, there must exist $\overline{F} \in \mathcal{F}$ such that $\overline{F} \subsetneq F$ and $\overline{F} \cap \widetilde{S}_\Lambda = F \cap \widetilde{S}_\Lambda$. It follows that $C\overline{F}$ is coned off in $\widehat{C}F$ and that $\mathfrak{g}_F(\widetilde{S}_\Lambda) \subseteq \overline{F}$. This implies that $\operatorname{diam}_{\widehat{C}F}(\pi_F(\widetilde{S}_\Lambda)) \leq 4$.

Now assume $F \cap \widetilde{S}_\Lambda = \varnothing$ for all $F \in [F]$. An argument like that in the proof of Proposition 8.3 of [3] shows that we can find $g \in A(\Gamma)$, $\Gamma' \in \mathcal{R}$, and $x \in A(\Lambda)$ so that $g\widetilde{S}_{\Gamma'} \in [F]$ and

$$\mathfrak{g}_{g\widetilde{S}_{\Gamma'}}(\widetilde{S}_\Lambda) \subseteq g(\widetilde{S}_{\Gamma' \cap \Lambda \cap \operatorname{Lk} \overline{g}}) \subseteq g(\widetilde{S}_{\Gamma' \cap \operatorname{Lk} \overline{g}}), \tag{9}$$

where $\overline{g} = g^{-1}x$ and $\operatorname{Lk}\overline{g}$ denotes the link of $\overline{g}$. Now if $\Gamma' \cap \Lambda \cap \operatorname{Lk}\overline{g} = \varnothing$, then $\mathfrak{g}_{g\widetilde{S}_{\Gamma'}}(\widetilde{S}_\Lambda) = \{g\}$, implying that $\operatorname{diam}_{\widehat{C}(g\widetilde{S}_{\Gamma'})}(\pi_{g\widetilde{S}_{\Gamma'}}(\widetilde{S}_\Lambda)) \leq 1$. Assume then that $\Gamma' \cap \Lambda \cap \operatorname{Lk}\overline{g} \neq \varnothing$. Then by definition of $\mathcal{R}$ and $\mathcal{F}$, we have that $\Gamma' \cap \operatorname{Lk}\overline{g} \in \mathcal{R}$ and $g(\widetilde{S}_{\Gamma' \cap \operatorname{Lk}\overline{g}}) \in \mathcal{F} - \{0\text{–cubes}\}$. If $g(\widetilde{S}_{\Gamma' \cap \operatorname{Lk}\overline{g}})$ is not a proper subcomplex of $g\widetilde{S}_{\Gamma'}$, then $\Gamma' \subseteq \operatorname{Lk}\overline{g}$, implying that $x\widetilde{S}_{\Gamma'}$ is parallel to $g\widetilde{S}_{\Gamma'}$ (see Lemma 2.4 in [3]). But this cannot be because $(x\widetilde{S}_{\Gamma'}) \cap \widetilde{S}_\Lambda = x(\widetilde{S}_{\Gamma' \cap \Lambda}) \neq \varnothing$ and no factor parallel to $g\widetilde{S}_{\Gamma'}$ intersects $\widetilde{S}_\Lambda$ nontrivially. Therefore, $g(\widetilde{S}_{\Gamma' \cap \operatorname{Lk}\overline{g}})$ must be a proper subcomplex of $g\widetilde{S}_{\Gamma'}$. Thus, $Cg(\widetilde{S}_{\Gamma' \cap \operatorname{Lk}\overline{g}})$ is coned off in $\widehat{C}(g\widetilde{S}_{\Gamma'})$. This together with (9) implies that $\operatorname{diam}_{\widehat{C}(g\widetilde{S}_{\Gamma'})}(\pi_{g\widetilde{S}_{\Gamma'}}(\widetilde{S}_\Lambda)) \leq 4$, completing the proof of Claim 2. □

We now argue that $A(\Lambda) \to \operatorname{Mod}(S)$ does not extend continuously to a map $\partial A(\Lambda) \to \partial\operatorname{Mod}(S)$. Let η denote a geodesic representative of an essential boundary component of a small regular neighborhood of $\alpha_i \cup \alpha_j$. Using the proof techniques of Lemma 5.2, we can construct $g \in A(\Lambda)$ so that $d_\eta(\mu, \phi(s_i^n g^n)\mu)$ grows linearly in n. For later convenience, we construct g so that when written in reduced form, the first letter of g is $s_j^{\pm 1}$. As in Proposition 4.4, we see that the sequences $(\phi(s_i^n))$ and $(\phi(s_i^n g^n))$ do not converge to the same point in $\operatorname{Mod}(S) \cup \partial\operatorname{Mod}(S)$. Now observe that (s_i^n) and $(s_i^n g^n)$ converge to the same point in $\partial_G A(\Lambda)$. Therefore, by the discussion in Section 2.8, (s_i^n) and $(s_i^n g^n)$ converge to the same point in $\partial A(\Lambda)$. We have now established that $A(\Lambda) \to \operatorname{Mod}(S)$ does not extend continuously to a map $\partial A(\Lambda) \to \partial\operatorname{Mod}(S)$. Therefore, $A(\Gamma) \to \operatorname{Mod}(S)$ does not extend continuously when $A(\Gamma)$ is equipped with a standard HHS structure.

Now suppose $A(\Gamma)$ is a free group equipped with any HHS structure. Then, by the discussion in Section 2.8, because (s_i^n) and $(s_i^n g^n)$ converge to the same point in $\partial_G A(\Gamma)$, we have that (s_i^n) and $(s_i^n g^n)$ converge to the same point in $\partial A(\Gamma)$. Because $(\phi(s_i^n))$ and $(\phi(s_i^n g^n))$ do not converge to the same point in $\partial\operatorname{Mod}(S)$, it follows that $A(\Gamma) \to \operatorname{Mod}(S)$ does not extend continuously. □

6 Existence of boundary maps for some free groups

In this section, we show that a class of embeddings of free groups in $\mathrm{Mod}(S)$, which includes a class of Koberda embeddings and a class of CLM embeddings, extend continuously.

Throughout this section, let Γ be the graph with $V(\Gamma) = \{s_1, \dots, s_k\}$ and no edges, and let $A(\Gamma)$ denote the corresponding right-angled Artin group (a rank k free group). Equip $A(\Gamma)$ with an HHS structure. Let $\{X_1, \dots, X_k\}$ be a collection of distinct, pairwise overlapping and pairwise filling subsurfaces of S and $\{f_1, \dots, f_k\}$ a collection of mapping classes such that f_i is fully supported on X_i. Let μ be a fixed marking on S. The main theorem of this section is the following, which implies the remaining direction of Theorem 1.3 in the introduction:

Theorem 6.1 *Let $A(\Gamma)$ be the rank k free group equipped with any HHS structure. Let $\{X_1, \dots, X_k\}$ be a collection of distinct, pairwise overlapping and pairwise filling subsurfaces of S, and $\{f_1, \dots, f_k\}$ a collection of mapping classes such that f_i is fully supported on X_i. There exists a $C > 0$ such that, if $\tau_{X_i}(f_i) \geq C$ for all i, then the homomorphism*

$$\phi\colon A(\Gamma) \to \mathrm{Mod}(S) \quad \text{defined by } \phi(s_i) = f_i \text{ for all } i$$

is a quasi-isometric embedding and extends continuously to a map $\partial A(\Gamma) \to \partial \mathrm{Mod}(S)$.

We emphasize the arguments we will use to establish that ϕ is a quasi-isometric embedding are essentially the same as those used by Clay, Leininger and Mangahas to prove Theorem 2.5. In particular, when the X_i are all nonannular, that ϕ is a quasi-isometric embedding is Theorem 2.5. To prove Theorem 6.1, we require the following proposition:

Proposition 6.2 *There exists $K > 0$ such that the following holds: For each $1 \leq i \leq k$, assume $\tau_{X_i}(f_i) \geq 2K$. Let $\phi\colon A(\Gamma) \to \mathrm{Mod}(S)$ be the homomorphism defined by $\phi(s_i) = f_i$ for all i. Consider $g_1 \dots g_k \in A(\Gamma)$, where for each i we have $g_i = x_i^{e_i}$ for some $x_i \in \{s_1^{\pm 1}, \dots, s_k^{\pm 1}\}$ and $e_i > 0$, and $x_i \neq x_{i+1}$, and $x_1^{e_1} \dots x_k^{e_k}$ is a reduced word. Let Y_i be the subsurface of S that fully supports $\phi(x_i)$. Then:*

(1) *For each $1 \leq i \leq k$, we have $d_{\phi(g_1 \dots g_{i-1})Y_i}(\mu, \phi(g_1 \dots g_k)\mu) \geq Ke_i$.*

(2) *For all $1 \leq i < j \leq k$, we have $\phi(g_1 \dots g_{i-1})Y_i \prec \phi(g_1 \dots g_{j-1})Y_j$, where $\prec$ denotes the partial order on $\Omega(K, \mu, \phi(g_1 \dots g_k)\mu)$.*

(3) *The homomorphism $\phi\colon A(\Gamma) \to \mathrm{Mod}(S)$ is a quasi-isometric embedding.*

Proof Define $K = K_0 + 20 + 2\max\{d_{X_i}(\mu, \partial X_j) : 1 \le i, j \le k \text{ and } i \ne j\}$, where K_0 is maximum of the constants in Theorem 6.12 of [11] and Theorem 2.4. Statements (1) and (2) of this proposition are essentially Theorem 5.2 in [6]. The difference is that Theorem 5.2 does not allow for the homomorphism to send a generator to a power of a Dehn twist. The only obstruction to Theorem 5.2 holding for homomorphisms ϕ of this type is the following: Suppose X_i is the subsurface that fully supports $\phi(s_i)$, and let $\sigma \in A(\Gamma)$ be a nonempty word in letters commuting with s_i, not including s_i. If X_i is nonannular, then $d_{X_i}(\phi(\sigma)\mu', \mu'') = d_{X_i}(\mu', \mu'')$ for any markings μ' and μ''. This not necessarily true if X_i is an annulus. However, this issue does not arise for us because $A(\Gamma)$ a free group implies no such σ exists. Thus, the arguments used to prove Theorem 5.2 in [6] also prove our statements (1) and (2). The proof of our statement (3) is the same as the proof in [6] of Theorem 2.5, using our statement (1) instead of their Theorem 5.2. □

The proof of the next lemma is essentially contained in the proof of Theorem 6.1 in [6]. We include a proof here for completeness.

Lemma 6.3 *Let $\phi\colon A(\Gamma) \to \mathrm{Mod}(S)$, $g_1 \dots g_k \in A(\Gamma)$ and Y_i be as in Proposition 6.2. Let $\mathcal{G}$ be a geodesic in $\mathcal{C}(S)$ with one end in $\pi_S(\mu)$ and the other end in $\pi_S(\phi(g_1 \dots g_k)\mu)$. Then, for each $1 \le i \le k$, there exists a curve γ_i on $\mathcal{G}$ such that $\pi_{\phi(g_1 \dots g_{i-1})Y_i}(\gamma_i) = \varnothing$. If $|i - j| \ge 3$ and γ_i and γ_j are two such curves, then $\gamma_i \ne \gamma_j$.*

Proof Fix $1 \le i \le k$. By way of contradiction, suppose for all curves v on $\mathcal{G}$, we have $\pi_{\phi(g_1 \dots g_{i-1})Y_i}(v) \ne \varnothing$. Then Theorems 2.4 and 2.1 together imply that

$$d_{\phi(g_1 \dots g_{i-1})Y_i}(\mu, \phi(g_1 \dots g_k)\mu) \le 4 + K_0.$$

But Proposition 6.2 says $d_{\phi(g_1 \dots g_{i-1})Y_i}(\mu, \phi(g_1 \dots g_k)\mu) \ge K > K_0 + 4$, a contradiction. Thus, there must exist a curve γ_i on $\mathcal{G}$ such that $\pi_{\phi(g_1 \dots g_{i-1})Y_i}(\gamma_i) = \varnothing$, as desired. Note that this implies that γ_i and $\partial\phi(g_1 \dots g_{i-1})Y_i$ form a multicurve.

Now consider γ_i and γ_j, where $1 \le i < j \le k$ and $|i - j| \ge 3$. We will show that γ_i and γ_j are distinct curves. To the contrary, suppose $\gamma_i = \gamma_j$. Because of the filling assumption on $\{X_1, \dots, X_k\}$, the pair of subsurfaces Y_{i+1} and Y_{i+2} fill S. Thus, $\phi(g_1 \dots g_{i+1})Y_{i+1} = \phi(g_1 \dots g_i)Y_{i+1}$ and $\phi(g_1 \dots g_{i+1})Y_{i+2}$ are also a pair of subsurfaces that fill S. Thus, it must be that $\pi_{\phi(g_1 \dots g_{n-1})Y_n}(\gamma_i) \ne \varnothing$ for some $n \in \{i+1, i+2\}$. In any case, $i < n < j$.

In the remainder of this proof, to simplify notation, for each l we define $\overline{Y}_l = \phi(g_1 \dots g_{l-1})Y_l$. By Proposition 6.2, we have

$$\overline{Y}_i \prec \overline{Y}_n \prec \overline{Y}_j,$$

where $\prec$ is the partial order on $\Omega(K, \mu, \phi(g_1 \dots g_k)\mu)$. In particular, these three subsurfaces are pairwise overlapping. This together with the assumption that $\gamma_i = \gamma_j$ and Theorem 2.1 implies that

$$d_{\overline{Y}_n}(\partial\overline{Y}_i, \partial\overline{Y}_j) \leq d_{\overline{Y}_n}(\partial\overline{Y}_i, \gamma_i) + d_{\overline{Y}_n}(\gamma_j, \partial\overline{Y}_j) \leq 2 + 2 = 4.$$

It follows from this and the definition of $\prec$ that

$$\begin{aligned} d_{\overline{Y}_n}(\mu, \phi(g_1 \dots g_k)\mu) &\leq d_{\overline{Y}_n}(\mu, \partial\overline{Y}_i) + d_{\overline{Y}_n}(\partial\overline{Y}_i, \partial\overline{Y}_j) + d_{\overline{Y}_n}(\partial\overline{Y}_j, \phi(g_1 \dots g_k)\mu) \\ &\leq 4 + 4 + 4 = 12. \end{aligned}$$

But this cannot be, because $d_{\overline{Y}_n}(\mu, \phi(g_1 \dots g_k)\mu) \geq K \geq 20$ by Proposition 6.2. Therefore, γ_i and γ_j are distinct curves. □

We have now developed the tools we will need to prove Theorem 6.1.

Proof of Theorem 6.1 Define $C = 2K$, where K is as in Proposition 6.2 and, for each $1 \leq i \leq k$, assume that $\tau_{X_i}(f_i) \geq C$. By Proposition 6.2, ϕ is a quasi-isometric embedding.

Let X denote the Cayley graph of $A(\Gamma)$. Choose $x \in \partial_G X$. Let γ be the infinite geodesic ray in X based at 1 limiting to x in $\partial_G X$. We think of γ as an infinite word of the form $y_1 y_2 y_3 \dots$, where each $y_i \in \{s_1^{\pm 1}, \dots, s_k^{\pm 1}\}$ and the word $y_1 y_2 \dots y_i$ is a reduced word for all i. By construction, the sequence $(y_1 \dots y_n)$ converges to x in $X \cup \partial_G X$. Let (h_n) be another sequence in $A(\Gamma)$ that converges to x in $X \cup \partial_G X$. We will show that $(\phi(h_n))$ and $(\phi(y_1 \dots y_n))$ converge to the same point in $\partial\mathrm{Mod}(S)$. By the discussion in Section 2.8, this will prove the theorem. We will consider two cases: (1) there does not exist $N \geq 1$ such that $y_i = y_N$ for all $i \geq N$, and (2) such an N exists. In both cases, we will assume each h_n is written in the form $h_n = g_{n,1} \dots g_{n,N(n)}$, where for all i we have $g_{n,i} = x_{n,i}^{e_{n,i}}$ for some $e_{n,i} > 0$ and $x_{n,i} \in \{s_1^{\pm 1}, \dots, s_k^{\pm 1}\}$ satisfying $x_{n,i} \neq x_{n,i+1}$, and $x_{n,1}^{e,1} \dots x_{n,N(n)}^{e_{n,N(n)}}$ is a reduced word.

Case 1 Suppose there does not exist $N \geq 1$ such that $y_i = y_N$ for all $i \geq N$. Then we can think of γ as an infinite word of the form $g_1 g_2 g_3 \dots$, where $g_i = x_i^{e_i}$ for some $e_i > 0$ and $x_i \in \{s_1^{\pm 1}, \dots, s_k^{\pm 1}\}$ satisfying $x_i \neq x_{i+1}$, and $x_1^{e_1} \dots x_i^{e_i}$ is a reduced

word for all i. Define Y_i to be the subsurface that fully supports $\phi(x_i)$. For short, we let $\overline{Y}_i$ denote $\phi(g_1 \dots g_{i-1})Y_i$.

Because (h_n) and $(y_1 \dots y_n)$ converge to the same point in $\partial_G X$ and X is a tree, h_n and $y_1 \dots y_n$ must agree on longer and longer initial segments as $n \to \infty$. In particular, given $L \geq 1$, there exists M such that, for all $n \geq M$, we have $g_{n,1} \dots g_{n,L} = g_1 \dots g_L$. Consider $n \geq M$ and $k \geq e_1 + \cdots + e_L$. Choose a curve $\beta \in \text{base}(\mu)$. Given $\sigma \in A(\Gamma)$, let $\mathcal{G}(\sigma)$ denote some choice of geodesic in $\mathcal{C}(S)$ with endpoints β and $\phi(\sigma)\beta$. By Lemma 6.3, for all $1 \leq i \leq L$ there exist curves γ_i and γ_i' on $\mathcal{G}(y_1 \dots y_k)$ and $\mathcal{G}(h_n)$, respectively, such that $\pi_{\overline{Y}_i}(\gamma_i) = \varnothing$ and $\pi_{\overline{Y}_i}(\gamma_i') = \varnothing$. Observe that

$$d_S(\gamma_i, \partial\overline{Y}_i) \leq 1 \quad \text{and} \quad d_S(\gamma_i', \partial\overline{Y}_i) \leq 1.$$

Choose γ_r to be the curve in $\{\gamma_i : 1 \leq i \leq L\}$ closest to $\phi(y_1 \dots y_k)\beta$. Lemma 6.3 tells us that if $|i - j| \geq 3$, then $\gamma_i \neq \gamma_j$. So necessarily $d_S(\beta, \gamma_r) \geq \frac{1}{3}L$. Thus, the Gromov product, computed in $\mathcal{C}(S)$, is

$$\begin{aligned}
&(\phi(y_1 \dots y_k)\beta, \phi(h_n)\beta)_\beta \\
&\quad = \tfrac{1}{2}[d_S(\beta, \phi(y_1 \dots y_k)\beta) + d_S(\beta, \phi(h_n)\beta) - d_S(\phi(y_1 \dots y_k)\beta, \phi(h_n)\beta)] \\
&\quad \geq \tfrac{1}{2}\big[d_S(\beta, \gamma_r) + d_S(\gamma_r, \phi(y_1 \dots y_k)\beta) + d_S(\beta, \gamma_r') + d_S(\gamma_r', \phi(h_n)\beta) \\
&\qquad\quad - \big(d_S(\phi(y_1 \dots y_k)\beta, \gamma_r) + d_S(\gamma_r, \partial\overline{Y}_r) + d_S(\partial\overline{Y}_r, \gamma_r') + d_S(\gamma_r', \phi(h_n)\beta)\big)\big] \\
&\quad \geq \tfrac{1}{2}[d_S(\beta, \gamma_r) + d_S(\beta, \gamma_r') - 2] \\
&\quad \geq \tfrac{1}{2}\big(\tfrac{1}{3}L - 2\big).
\end{aligned}$$

It follows that

$$\liminf_{k,n\to\infty} (\phi(y_1 \dots y_k)\beta, \phi(h_n)\beta))_\beta = \infty. \tag{10}$$

Because (h_n) is an arbitrary sequence converging to x, we could have taken it to be $(y_1 \dots y_n)$. Thus, (10) tells us two things: (1) $(\phi(y_1 \dots y_n)\mu)$ converges to a point in $\partial\mathcal{C}(S)$, and (2) $(\phi(y_1 \dots y_n)\mu)$ and $(\phi(h_n)\mu)$ converge to the same point in $\partial\mathcal{C}(S)$. By definition of the topology on $\text{Mod}(S) \cup \partial\text{Mod}(S)$, this tells us that $(\phi(y_1 \dots y_n))$ and $(\phi(h_n))$ converge to the same point in $\partial\text{Mod}(S)$.

Case 2 Assume there exists $N \geq 1$ such that $y_i = y_N$ for all $i \geq N$. Corollary 6.2 in [7] tells us that the action of $\text{Mod}(S)$ by left multiplication extends to an action of $\text{Mod}(S)$ on $\text{Mod}(S) \cup \partial\text{Mod}(S)$ by homeomorphisms. Consequently, if we can show that $(\phi((y_1 \dots y_{N-1})^{-1}h_n))_{n\in\mathbb{N}}$ and $(\phi(y_N \dots y_n))_{n\in\mathbb{N}}$ converge to the same

point in $\partial\mathrm{Mod}(S)$, then $(\phi(h_n))_{n\in\mathbb{N}}$ and $(\phi(y_1 \dots y_n))_{n\in\mathbb{N}}$ must converge to the same point in $\partial\mathrm{Mod}(S)$. Furthermore, $((y_1 \dots y_{N-1})^{-1}h_n)_{n\in\mathbb{N}}$ and $(y_N \dots y_n)_{n\in\mathbb{N}}$ converge to the same point in $\partial_G X$. Thus, without loss of generality we assume $N = 1$. By our assumption, $y_1 \dots y_n = y_1^n$ for all n.

Let Y be the subsurface that fully supports $\phi(y_1)$, and let $\partial Y = \{\beta_1, \dots, \beta_l\}$. Then

$$\lim_{n\to\infty} \frac{d_Y(\mu, \phi(y_1^n)\mu)}{n} > 0 \quad \text{and} \quad \pi_Y(\phi(y_1^n)\mu) \to \lambda_Y \quad \text{for some } \lambda_Y \in \partial\mathcal{C}(Y). \tag{11}$$

Further observe that, for all i,

$$\lim_{n\to\infty} \frac{d_{\beta_i}(\mu, \phi(y_1^n)\mu)}{n} \geq 0. \tag{12}$$

If (12) is an equality, let λ_i be any point in $\partial\mathcal{C}(\beta_i)$. Otherwise, define $\lambda_i \in \partial\mathcal{C}(\beta_i)$ to be $\lim_{n\to\infty} \pi_{\beta_i}(\phi(y_1^n)\mu)$. For all subsurfaces W disjoint from Y and not an annulus with core curve in ∂Y, Lemma 3.1 and Theorem 2.1 imply that $d_W(\mu, \phi(y_1^n)\mu) \leq d_W(\mu, \mu) + 4 \leq 6$. Consequently,

$$\lim_{n\to\infty} \phi(y_1^n) = c_Y \lambda_Y + \sum_{i=1}^{l} c_i \lambda_i,$$

where

$$c_Y + \sum_{i=1}^{l} c_i = 1 \quad \text{and} \quad \frac{c_i}{c_Y} = \lim_{n\to\infty} \frac{d_{\beta_i}(\mu, \phi(y_1^n)\mu)}{d_Y(\mu, \phi(y_1^n)\mu)}.$$

Because (h_n) and (y_1^n) converge to the same point in $\partial_G X$, given any $L \geq 1$, for all sufficiently large n we have $x_{n,1} = y_1$ and $e_{n,1} \geq L$. So, by removing finitely many initial terms from (h_n), for convenience we may assume that $g_{n,1} = y_1^{e_{n,1}}$ for all n. Observe that $e_{n,1} \to \infty$ as $n \to \infty$. It is immediate from this and the definition of the topology of $\mathrm{Mod}(S) \cup \partial\mathrm{Mod}(S)$ that $\lim_{n\to\infty} \phi(g_{n,1}) = \lim_{n\to\infty} \phi(y_1^n)$. Thus, to finish the proof, we must show $\lim_{n\to\infty} \phi(g_{n,1}) = \lim_{n\to\infty} \phi(h_n)$. By passing to subsequences, we may assume that either $N(n) = 1$ for all n or $N(n) \geq 2$ for all n. If the former holds, then $h_n = g_{n,1}$, and we are done. Assume then that $N(n) \geq 2$ for all n. To proceed, we require the following claims:

Claim 1 *$d_Y(\phi(g_{n,1})\mu, \phi(h_n)\mu)$ is bounded above, independent of n.*

Claim 2 *Let W be a subsurface that is disjoint from Y. Then $d_W(\phi(g_{n,1})\mu, \phi(h_n)\mu)$ is bounded above, independent of n.*

We postpone the proofs of these claims and for now assume they are true. First, observe that Claim 1 and (11) imply that $\pi_Y(\phi(h_n)\mu) \to \lambda_Y$. If Inequality (12) is strict, then Claim 2 implies that $\pi_{\beta_i}(\phi(h_n)\mu) \to \lambda_i$. Further observe that Claims 1 and 2 imply that, for all W disjoint from Y,

$$\begin{aligned}\lim_{n\to\infty} \frac{d_W(\mu, \phi(g_{n,1})\mu)}{d_Y(\mu, \phi(g_{n,1})\mu)} &= \frac{\lim_{n\to\infty} d_W(\mu, \phi(g_{n,1})\mu)/e_{n,1}}{\lim_{n\to\infty} d_Y(\mu, \phi(g_{n,1})\mu)/e_{n,1}} \\ &= \frac{\lim_{n\to\infty} d_W(\mu, \phi(h_n)\mu)/e_{n,1}}{\lim_{n\to\infty} d_Y(\mu, \phi(h_n)\mu)/e_{n,1}} \\ &= \lim_{n\to\infty} \frac{d_W(\mu, \phi(h_n)\mu)}{d_Y(\mu, \phi(h_n)\mu)}.\end{aligned}$$

It follows that $\lim_{n\to\infty} \phi(g_{n,1}) = \lim_{n\to\infty} \phi(h_n)$, as desired.

To finish the proof, we will now prove Claims 1 and 2. For each n, let Z_n denote the subsurface that fully supports $\phi(x_{n,2})$.

Proof of Claim 1 Fix $n \geq 1$. Because Y fully supports $\phi(x_{n,1})$, by Proposition 6.2, we know $Y \prec \phi(g_{n,1})Z_n$, where $\prec$ denotes the partial order on $\Omega(K, \mu, \phi(h_n)\mu)$. Thus, $d_Y(\partial\phi(g_{n,1})Z_n, \phi(h_n)\mu) \leq 4$. Therefore,

$$\begin{aligned}d_Y(\phi(g_{n,1})\mu, \phi(h_n)\mu) &\leq d_Y(\phi(g_{n,1})\mu, \partial\phi(g_{n,1})Z_n) + d_Y(\partial\phi(g_{n,1})Z_n, \phi(h_n)\mu) \\ &\leq d_Y(\mu, \partial Z_n) + 4.\end{aligned}$$

There are finitely many possibilities for Z_n, so this completes the proof of Claim 1. □

Proof of Claim 2 Fix $n \geq 1$. Because Y and Z_n fill S, and Y and W are disjoint, it must be that $\pi_{Z_n}(\partial W) \neq \varnothing$. There are two cases to consider: (1) $W \pitchfork Z_n$, and (2) $W \subsetneq Z_n$. First, suppose that $W \pitchfork Z_n$. It then follows from Proposition 6.2, Theorem 2.1 and the definition of K that

$$\begin{aligned}&d_{Z_n}(\partial W, \phi(g_{n,2} \dots g_{n,N(n)})\mu) \\ &\qquad \geq d_{Z_n}(\mu, \phi(g_{n,2} \dots g_{n,N(n)})\mu) - d_{Z_n}(\partial Y, \partial W) - d_{Z_n}(\mu, \partial Y) \\ &\qquad \geq K - 2 - \tfrac{1}{2}K \geq 10.\end{aligned}$$

Thus Theorem 2.3 implies that $d_W(\partial Z_n, \phi(g_{n,2} \dots g_{n,N(n)})\mu) \leq 4$. From this and Theorem 2.2 we find that

$$\begin{aligned}(13)\qquad d_W(\phi(g_{n,1})\mu, \phi(h_n)\mu) &= d_W(\mu, \phi(g_{n,2} \dots g_{n,N(n)})\mu) \\ &\leq d_W(\mu, \partial Z_n) + d_W(\partial Z_n, \phi(g_{n,2} \dots g_{n,N(n)})\mu) \\ &\leq 4 \max\{d_{\widetilde{\mathcal{M}}(S)}(\mu, \mu_i) : 1 \leq i \leq k\} + 4,\end{aligned}$$

where μ_i is a fixed choice of marking with $\partial X_i \subseteq \text{base}(\mu_i)$ for each $1 \leq i \leq k$. This provides a uniform bound in the case that $W \pitchfork Z_n$.

Now suppose that $W \subsetneq Z_n$. First, observe that because Z_n fully supports $\phi(x_{n,2})$, the sequence $(\pi_{Z_n}(\phi(x_{n,2})^m \mu))_{m\in\mathbb{N}}$ converges to a point in $\partial\mathcal{C}(Z_n)$. Thus, by Corollary 2.7 there exists a constant M, which depends on W and $x_{n,2}$, such that $d_W(\mu, \phi(g_{n,2})\mu) \leq M$ for all n. Note that there are only finitely many possibilities for $x_{n,2}$, so M can be chosen to be independent of n. This implies that

$$\begin{aligned} d_W(\phi(g_{n,1})\mu, \phi(h_n)\mu) &\leq d_W(\mu, \phi(g_{n,2}\dots g_{n,N(n)})\mu) \\ &\leq d_W(\mu, \phi(g_{n,2})\mu) + d_W(\phi(g_{n,2})\mu, \phi(g_{n,2}\dots g_{n,N(n)})\mu) \\ &\leq M + d_{\phi(g_{n,2})^{-1}W}(\mu, \phi(g_{n,3}\dots g_{n,N(n)})\mu). \end{aligned}$$

Now, if $N(n) = 2$, then we can apply Theorem 2.1 to see that

$$d_{\phi(g_{n,2})^{-1}W}(\mu, \phi(g_{n,3}\dots g_{n,N(n)})\mu) = d_{\phi(g_{n,2})^{-1}W}(\mu, \mu) \leq 2,$$

and Claim 2 is established. Suppose then that $N(n) \geq 3$. Let V_n denote the subsurface that fully supports $\phi(x_{n,3})$. Observe that, because $\tau_{Z_n}(\phi(x_{n,2})) \geq 2K$ and ∂Y and ∂W form a multicurve, we have

$$\begin{aligned} &d_{Z_n}(\partial\phi(g_{n,2})^{-1}W, \partial V_n) \\ &\quad\geq d_{Z_n}(\partial W, \partial\phi(g_{n,2})^{-1}W) - d_{Z_n}(\partial W, \partial Y) - d_{Z_n}(\mu, \partial Y) - d_{Z_n}(\mu, \partial V_n) \\ &\quad\geq 2K - 2 - \tfrac{1}{2}K - \tfrac{1}{2}K > 2. \end{aligned}$$

This together with Theorem 2.2 establishes that $\partial\phi(g_{n,2})^{-1}W$ and ∂V_n do not form a multicurve. Thus, $\phi(g_{n,2})^{-1}W \pitchfork V_n$. So, to bound

$$d_{\phi(g_{n,2})^{-1}W}(\mu, \phi(g_{n,3}\dots g_{n,N(n)})\mu)$$

from above independent of n, we can use the same techniques used above to bound $d_W(\mu, \phi(g_{n,2}\dots g_{n,N(n)})\mu)$ when $W \pitchfork Z_n$. This completes the proof of Claim 2, and thus the proof of Theorem 6.1. □

References

[1] **J A Behrstock**, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. 10 (2006) 1523–1578 MR

[2] **J Behrstock**, **M F Hagen**, **A Sisto**, *Hierarchically hyperbolic spaces, II: Combination theorems and the distance formula*, preprint (2015) arXiv

[3] **J Behrstock**, **M F Hagen**, **A Sisto**, *Hierarchically hyperbolic spaces, I: Curve complexes for cubical groups*, Geom. Topol. 21 (2017) 1731–1804 MR

[4] **J Behrstock**, **B Kleiner**, **Y Minsky**, **L Mosher**, *Geometry and rigidity of mapping class groups*, Geom. Topol. 16 (2012) 781–888 MR

[5] **J F Brock**, *The Weil–Petersson metric and volumes of* 3*–dimensional hyperbolic convex cores*, J. Amer. Math. Soc. 16 (2003) 495–535 MR

[6] **M T Clay**, **C J Leininger**, **J Mangahas**, *The geometry of right-angled Artin subgroups of mapping class groups*, Groups Geom. Dyn. 6 (2012) 249–278 MR

[7] **M G Durham**, **M F Hagen**, **A Sisto**, *Boundaries and automorphisms of hierarchically hyperbolic spaces*, Geom. Topol. 21 (2017) 3659–3758 MR

[8] **T Koberda**, *Right-angled Artin groups and a generalized isomorphism problem for finitely generated subgroups of mapping class groups*, Geom. Funct. Anal. 22 (2012) 1541–1590 MR

[9] **J Mangahas**, *A recipe for short-word pseudo-Anosovs*, Amer. J. Math. 135 (2013) 1087–1116 MR

[10] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, I: Hyperbolicity*, Invent. Math. 138 (1999) 103–149 MR

[11] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, II: Hierarchical structure*, Geom. Funct. Anal. 10 (2000) 902–974 MR

Department of Mathematics, University of Illinois at Urbana-Champaign
Urbana, IL, United States

mousley2@illinois.edu

http://www.math.uiuc.edu/~mousley2/

Received: 17 November 2016 Revised: 17 July 2017

msp

Algebraic & Geometric Topology 18 (2018) 441–492

Finite Dehn surgeries on knots in S^3

YI NI
XINGRU ZHANG

We show that on a hyperbolic knot K in S^3, the distance between any two finite surgery slopes is at most 2, and consequently, there are at most three nontrivial finite surgeries. Moreover, in the case where K admits three nontrivial finite surgeries, K must be the pretzel knot $P(-2,3,7)$. In the case where K admits two noncyclic finite surgeries or two finite surgeries at distance 2, the two surgery slopes must be one of ten or seventeen specific pairs, respectively. For D–type finite surgeries, we improve a finiteness theorem due to Doig by giving an explicit bound on the possible resulting prism manifolds, and also prove that $4m$ and $4m+4$ are characterizing slopes for the torus knot $T(2m+1,2)$ for each $m \geq 1$.

57M25

1 Introduction

For a knot K in S^3, its set of slopes $\{p/q : p,q \in \mathbb{Z},\ (p,q)=1\}$ will be parametrized by the standard meridian/longitude coordinates of K. Recall that the distance between two slopes p_1/q_1, p_2/q_2 of K is the number $|p_1q_2 - p_2q_1|$. A Dehn surgery on a knot $K \subset S^3$ with slope p/q is called a cyclic or finite surgery if the resulting manifold, which we denote by $S^3_K(p/q)$, has cyclic or finite fundamental group, respectively. By Gabai [23], $S^3_K(0)$ has cyclic fundamental group only when K is the trivial knot. It follows that cyclic surgery on nontrivial knots in S^3 is equivalent to finite cyclic surgery. Due to Perelman's resolution of Thurston's geometrization conjecture [47; 48], a connected closed 3–manifold has finite fundamental group if and only if it is a spherical space form. For a spherical space form Y, it has cyclic fundamental group if and only if it is a lens space, and it has noncyclic fundamental group if and only if it has a Seifert fibered structure whose base orbifold is $S^2(a,b,c)$, a 2–sphere with three cone points of orders $a \leq b \leq c$, satisfying $1/a + 1/b + 1/c > 1$; ie (a,b,c) is either $(2,3,3)$, $(2,3,4)$, $(2,3,5)$, or $(2,2,n)$ for some integer $n > 1$. Correspondingly, we say that a spherical space form Y or its fundamental group is C–type, T–type, O–type, I–type, or D–type if Y is a lens space or a Seifert-fibered space with base

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.441

orbifold $S^2(2,3,3)$, $S^2(2,3,4)$, $S^2(2,3,5)$, or $S^2(2,2,n)$, respectively. We shall also refine finite surgeries (slopes) into C–type (often called cyclic), T–type, O–type, I–type and D–type accordingly.

If a nonhyperbolic knot in S^3 admits a nontrivial finite surgery, then the knot is either a torus knot or a cable over a torus knot (see Boyer and Zhang [9]), and finite surgeries on torus knots and on cables over torus knots are classified in Moser [39] and Bleiler and Hodgson [6], respectively. Concerning finite surgeries on hyperbolic knots in S^3, we recall the following:

Known Facts 1.1 Let $K \subset S^3$ be any fixed hyperbolic knot.

(1) Any nontrivial cyclic surgery slope of K must be an integer, K has at most two nontrivial cyclic surgery slopes, and if two, they are consecutive integers; see Culler, Gordon, Luecke and Shalen [16].

(2) Any finite surgery slope of K must be either an integer or a half-integer [9], the distance between any two finite surgery slopes of K is at most 3, and K has at most four nontrivial finite surgery slopes (see Boyer and Zhang [13]); consequently, K has at most one half-integer finite surgery slope.

(3) The distance between a finite surgery slope and a cyclic surgery slope on K is at most 2 [9].

(4) Any D–type finite surgery slope of K must be an integer. There is at most one D–type finite surgery slope on K. If there is a D–type finite surgery on K, then there is at most one nontrivial cyclic surgery on K, and the D–type finite surgery slope and the cyclic surgery slope are consecutive integers [9].

(5) Any O–type finite surgery slope of K must be an integer. If there is an O–type finite surgery on K, then there is at most one nontrivial cyclic surgery on K, and the O–type finite surgery slope and the cyclic surgery slope are consecutive integers [9].

(6) There are at most two T–type finite surgery slopes on K, and if two, one is integral, the other is half-integral, and their distance is 3 [9].

The above results were obtained by using classical techniques, mostly those derived from $\mathrm{PSL}_2(\mathbb{C})$–representations of 3–manifold groups, hyperbolic geometry and geometric and combinatorial topology in dimension 3. Recently, new progress on finite surgeries on knots in S^3 has been made through applications of Floer homology theory. Note

that up to replacing a knot K by its mirror image $-K$, we may and shall assume that any finite surgery slope on a nontrivial knot in S^3 has positive sign. We recall the following:

Known Facts 1.2 (1) If a knot K in S^3 admits a nontrivial finite surgery, then K is a fibered knot; see Ni [40].

(2) If a nontrivial knot K in S^3 admits a nontrivial finite surgery slope p/q, then $p/q \geq 2g(K)-1$, where $g(K)$ is the Seifert genus of K (see Ozsváth and Szabó [46]), and the nonzero coefficients of the Alexander polynomial of K are alternating ± 1 (see Ozsváth and Szabó [44]).

(3) If a knot K in S^3 admits a T–type, O–type or I–type finite surgery with an integer slope, then the surgery slope is one of the finitely many integers listed in Tables 1–3 in Section 2, there is a sample knot K_0 (also listed in these tables) on which the same surgery slope yields the same spherical space form, and K and K_0 have the same knot Floer homology; see Gu [27].

(4) If a knot K in S^3 admits a T–type or I–type finite surgery with a half-integer slope, then the surgery slope is one of the ten slopes listed in Table 4 in Section 2 (with the two on the trefoil knot omitted as they can only be realized on the trefoil knot), there is a sample knot K_0 (also listed in the table) on which the same surgery slope yields the same spherical space form, and K and K_0 have the same knot Floer homology; see Li and Ni [36].

(5) If a knot K in S^3 admits an integer D–type finite surgery slope $p \leq 32$, then p is one of the slopes listed in Table 5 in Section 2, there is a sample knot K_0 (also listed in the table) on which the same surgery slope yields the same D–type spherical space form, and K and K_0 have the same knot Floer homology; see Doig [18]. (Also see Ballinger, Hsu, Mackey, Ni, Ochse and Vafaee [2] for more recent progress on D–type realization problem.)

(6) If a knot K in S^3 admits a cyclic surgery with an integer slope p, then there is a Berge knot K_0 (given in Berge [4]) such that $S^3_K(p) = S^3_{K_0}(p)$, and K and K_0 have the same knot Floer homology; see Greene [26].

(7) If p is a cyclic surgery slope for a hyperbolic knot K in S^3, then $p = 14$ or $p \geq 18$. Moreover, if $p \geq 4g(K)-1$, then K is a Berge knot; see Baker [1].

The purpose of this paper is to update and improve results on finite surgeries on hyperbolic knots in S^3, applying various techniques and results combined together.

As recalled in Known Facts 1.2(4), there are only ten specific half-integer slopes, listed in Table 4, each of which could possibly be a finite surgery slope for some hyperbolic knot in S^3. Our first result excludes two of them.

Theorem 1.3 *Neither $17/2$ nor $23/2$ can be a finite surgery slope for a hyperbolic knot in S^3.*

A main result of this paper is the following

Theorem 1.4 *Let K be any fixed hyperbolic knot in S^3.*

(1) *The distance between any two finite surgery slopes on K is at most 2. Consequently, there are at most three nontrivial finite surgeries on K.*

(2) *If K admits three nontrivial finite surgeries, then K must be the pretzel knot $P(-2,3,7)$.*

(3) *If K admits two noncyclic finite surgeries, the surgery slopes are one of the following ten pairs, the knot K has the same knot Floer homology as the sample knot K_0 given along with the pair, and the pair of slopes yields the same pair of spherical space forms on K_0:*

$\{43/2, 21, [11,2;3,2]\}$, $\{53/2, 27, [13,2;3,2]\}$, $\{103/2, 52, [17,3;3,2]\}$,
$\{113/2, 56, [19,3;3,2]\}$, $\{22, 23, P(-2,3,9)\}$, $\{28, 29, -K(1,1,0)\}$,
$\{50, 52, [17,3;3,2]\}$, $\{56, 58, [19,3;3,2]\}$, $\{91, 93, [23,4;3,2]\}$,
$\{99, 101, [25,4;3,2]\}$.

(4) *If K admits two finite surgery slopes which are distance 2 apart, then the two slopes are one of the following seventeen pairs, the knot K has the same knot Floer homology as the sample knot K_0 given along with the pair, and the pair of slopes yields the same pair of spherical space forms on K_0:*

$\{43/2, 1/0, [11,2;3,2]\}$, $\{45/2, 1/0, [11,2;3,2]\}$, $\{51/2, 1/0, [13,2;3,2]\}$,
$\{53/2, 1/0, [13,2;3,2]\}$, $\{77/2, 1/0, [19,2;5,2]\}$, $\{83/2, 1/0, [21,2;5,2]\}$,
$\{103/2, 1/0, [17,3;3,2]\}$, $\{113/2, 1/0, [19,3;3,2]\}$, $\{17, 19, P(-2,3,7)\}$,
$\{21, 23, [11,2;3,2]\}$, $\{27, 25, [13,2;3,2]\}$, $\{37, 39, [19,2;5,2]\}$,
$\{43, 41, [21,2;5,2]\}$, $\{50, 52, [17,3;3,2]\}$, $\{56, 58, [19,3;3,2]\}$,
$\{91, 93, [23,4;3,2]\}$, $\{99, 101, [25,4;3,2]\}$.

The notation for the sample knots will be explained in Section 2. Note that a sample knot is not necessarily hyperbolic. When a sample knot is nonhyperbolic, the corresponding case of the described finite surgeries on a hyperbolic knot may never happen.

Parts of the theorem are sharp; on the pretzel knot $P(-2,3,7)$ (which is hyperbolic), 17, 18, 19 are three finite surgery slopes, 17 being I–type and $18, 19$ being cyclic, and on the pretzel knot $P(-2,3,9)$ (which is also hyperbolic), 22 and 23 are two noncyclic finite surgery slopes, 22 being O–type and 23 being I–type.

A D–type spherical space form is also called a prism manifold. Let $P(n,m)$ be the prism manifold with Seifert invariants

$$(-1;(2,1),(2,1),(n,m)),$$

where the base orbifold has genus 0, $n>1$, and $\gcd(n,m)=1$. Every prism manifold can be expressed in this form. As a byproduct of the proof of Theorem 1.4, we have the following:

Theorem 1.5 *If $P(n,m)$ can be obtained by Dehn surgery on a knot K in S^3, then $n<|4m|$.*

Theorem 1.5 improves [19, Theorem 2] of Doig, where no explicit bound on n was given. (In an earlier preprint of [19], the author claimed a bound of $n<|16m|$ without proof.)

On a torus knot $T(2m+1,2)$, recall that $4m$ and $4m+4$ are D–type finite surgery slopes. Our next main result implies that neither of the prism manifolds $S^3_{T(2m+1,2)}(4m)$ nor $S^3_{T(2m+1,2)}(4m+4)$ can be obtained by surgery on any other knot in S^3 besides $\pm T(2m+1,2)$.

Theorem 1.6 *Suppose that $S^3_K(4n)\cong \varepsilon S^3_{T(2m+1,2)}(4n)$ for some $\varepsilon\in\{\pm\}$ and $n=m$ or $m+1$, where $\varepsilon\in\{\pm\}$ stands for an orientation. Then $\varepsilon=+$ and $K=T(2m+1,2)$.*

In the terminology of Ni and Zhang [41], the above theorem implies that $4m$ and $4m+4$ are characterizing slopes for $T(2m+1,2)$; that is, whenever $S^3_K(4n)\cong S^3_{T(2m+1,2)}(4n)$ for $n=m$ or $m+1$, then $K=T(2m+1,2)$.

Combining Theorem 1.6 with Known Facts 1.2(5) and Known Facts 1.1(4), we have:

Corollary 1.7 *Any D–type finite surgery slope of a hyperbolic knot in S^3 is an integer greater than or equal to 28.*

The bound 28 can be realized as a D–type finite surgery slope on two hyperbolic knots in S^3; see Table 5 in Section 2.

The results described above suggest the following updated conjectural picture concerning finite surgeries on hyperbolic knots in S^3.

Conjecture 1.8 *Let K be a hyperbolic knot in S^3.*

(1) **Berge conjecture** *If K admits a nontrivial cyclic surgery, then K is a primitive/primitive knot as defined in* [4] (*ie a Berge knot*).

(2) *If K admits a T–type, O–type or I–type finite surgery, then K is one of the twenty-three hyperbolic sample knots listed in Tables 1–3.* (This was raised in [27].)

(3) *K does not have any half-integral finite surgery slope.*

(4) *If K admits two noncyclic finite surgeries, then K is either $P(-2,3,9)$ or $-K(1,1,0)$.*

(5) *If K admits two finite surgeries at distance 2, then K is $P(-2,3,7)$.*

(6) *If K admits a noncyclic finite surgery, then K is a primitive/Seifert-fibered knot as defined in Dean* [17].

(7) *If the prism manifold $P(n,m)$ can be obtained by surgery on K, then $n < 2|m|-2$.* (This was improved from [19, Conjecture 12].)

The proofs of the above theorems are completed mainly using $\mathrm{PSL}_2(\mathbb{C})$–representation techniques, the correction terms from Heegaard Floer homology and the Casson–Walker invariant besides Known Facts 1.1 and 1.2. In Section 2, we give a more detailed explanation of Known Facts 1.2(3)–(5), which will be convenient to apply in later sections. In Section 3, we recall briefly some machinery for using $\mathrm{PSL}_2(\mathbb{C})$–representations for studying finite surgeries, specialized to the case for hyperbolic knots in S^3. We prove Theorem 1.3 in Section 4, where an outline of the proof will be indicated at the beginning. The method of proof is completed mainly using $\mathrm{PSL}_2(\mathbb{C})$–representation techniques combined with Known Facts 1.2(4) as well as various other results. We then present the proof of Theorem 1.4, which we split into two parts corresponding to the two cases whether a half-integer finite surgery slope is involved or not. Part I of the proof, given in Section 5, is completed mainly using $\mathrm{PSL}_2(\mathbb{C})$–representation techniques combined with Known Facts 1.2 as well as various other results, and part II of the proof, given in Section 6, is completed mainly using the Casson–Walker invariant combined with Known Facts 1.2. Section 6 also contains

the proof of Theorem 1.5. In Section 7, we prove Theorem 1.6 by applying Heegaard Floer homology and some topological arguments.

Acknowledgements Ni was partially supported by NSF grant numbers DMS-1103976, DMS-1252992, and an Alfred P Sloan Research Fellowship.

2 Tables of finite surgeries

Given a rational homology sphere Y and a Spin^c structure $\mathfrak{s} \in \mathrm{Spin}^c(Y)$, Ozsváth and Szabó defined a rational number $d(Y,\mathfrak{s})$ called the *correction term* [42]. Recall that a rational homology sphere Y is an *L-space* if its Heegaard Floer homology $\widehat{\mathrm{HF}}(Y,\mathfrak{s})$ is isomorphic to $\mathbb{Z}$ for each $\mathfrak{s} \in \mathrm{Spin}^c(Y)$. Lens spaces and, more generally, spherical 3–manifolds are L-spaces. The correction terms are the only informative Heegaard Floer invariants for L-spaces.

Let $L(p,q)$ be the lens space obtained by p/q–surgery on the unknot in S^3, and fix a particular identification $\mathrm{Spin}^c(L(p,q)) \cong \mathbb{Z}/p\mathbb{Z}$. The correction terms for lens spaces can be computed inductively as in [42]:

$$d(S^3,0)=0, \quad d(L(p,q),i) = -\frac{1}{4} + \frac{(2i+1-p-q)^2}{4pq} - d(L(q,r),j), \tag{2-1}$$

where $0 \le i < p+q$, and r and j are the reductions modulo q of p and i, respectively.

For a knot K in S^3, suppose its Alexander polynomial normalized by the conditions $\Delta_K(t) = \Delta_K(t^{-1})$ and $\Delta_K(1) = 1$ is

$$\Delta_K(t) = \sum_i a_i t^i.$$

Define a sequence of integers

$$t_i = \sum_{j=1}^{\infty} j a_{i+j}, \quad i \ge 0.$$

Note that $\Delta_K(t)$ can be recovered from the t_i. If K admits a L-space surgery, then one can prove [46; 49] that

$$t_i \ge 0, \quad t_i \ge t_{i+1} \ge t_i - 1, \quad t_{g(K)} = 0, \tag{2-2}$$

and the correction terms of $S^3_K(p/q)$ can be computed in terms of the t_i by the formula

$$d(S^3_K(p/q),i) = d(L(p,q),i) - 2t_{\min\{\lfloor i/q\rfloor, \lfloor (p+q-i-1)/q\rfloor\}}.$$

The above results give us a necessary condition for a spherical space form Y with $H_1(Y) \cong \mathbb{Z}/p\mathbb{Z}$ to be the p/q–surgery on any knot $K \subset S^3$.

Condition 2.1 There exist a sequence of integers $\{t_i\}_{i\geq 0}$ satisfying

$$t_i \geq 0, \quad t_i \geq t_{i+1} \geq t_i - 1 \quad \text{and} \quad t_i = 0 \text{ when } i \gg 0,$$

and a symmetric affine isomorphism $\phi\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ such that

$$d(Y, \phi(i)) = d(L(p,q), i) - 2t_{\min\{\lfloor i/q \rfloor, \lfloor (p+q-i-1)/q \rfloor\}}.$$

Here the fact that ϕ is symmetric means that ϕ commutes with the conjugation of Spin^c structures once we identify $\mathbb{Z}/p\mathbb{Z}$ with the corresponding sets of Spin^c structures.

Condition 2.1 is easy to check using a simple computer program. When it is satisfied, we can recover the Alexander polynomial of the possible knot K admitting the surgery to Y. It is surprising that Condition 2.1 is also sufficient in all known cases. For example, Ozsváth and Szabó [44, Proposition 1.13] confirmed that a lens space $L(p,q)$ with $p \leq 1500$ can be obtained by p–surgery on a knot in S^3 if and only if Condition 2.1 holds, and Rasmussen [50, Section 6] has further extended this confirmation to $p < 100\,000$. Doig [18] proved similar results for prism spaces with $|H_1| \leq 32$, and Gu [27] and Li and Ni [36] proved similar results for T–type, O–type and I–type spherical space forms.

Here we list the T–type, O–type, I–type and some D–type finite surgery slopes and the sample knots mentioned in Known Facts 1.2(3)–(5), which are reproduced from [27; 36; 18] for the reader's convenience. For each of these finite surgeries, we actually list four to six relevant pieces of data: p $\left(\text{or } \frac{1}{2}p\right)$, Y, K_0, g, $\det(K)$, and $\Delta_K''(1)$, which are useful in this and later sections. Here $p > 0$ $\left(\text{or } \frac{1}{2}p > 0\right)$ is the finite surgery slope, Y is the resulting manifold, K_0 is a sample knot which admits the finite surgery, and g, $\det(K)$ and $\Delta_K(t)$ are respectively the Seifert genus, the determinant and the normalized Alexander polynomial for any knot $K \subset S^3$ which admits the finite surgery.

Now we explain the notation for the manifolds $\mathbb{T}(p/q)$ and for the sample knots K_0 we use in these tables. Let $-K$ be the mirror image of K.

- It is not hard to see that every T-, O- or I–type spherical space form, up to orientation reversal, can be obtained by Dehn filling on $\mathbb{T}$, the exterior of the right-hand trefoil knot in S^3. Thus we represent the corresponding surgery manifold Y by Dehn filling on $\mathbb{T}$ and specify the orientation.

p	Y	K_0	g	$\det(K)$	$\Delta_K''(1)$
3	$\mathbb{T}(3)$	$T(3,2)$	1	3	2
9	$\mathbb{T}(9)$	$T(3,2)$	1	3	2
21	$\mathbb{T}(21/4)$	$[11,2;3,2]$	7	11	38
27	$\mathbb{T}(27/4)$	$[13,2;3,2]$	8	13	50
51	$-\mathbb{T}(51/8)$	$K_2^{\#}$	18	1	200
69	$-\mathbb{T}(69/11)$	$-K(2,3,5,-3)$	25	1	368
81	$\mathbb{T}(81/13)$	$K(2,3,5,3)$	31	1	536
93	$\mathbb{T}(93/16)$	$[23,4;3,2]$	37	23	692
99	$\mathbb{T}(99/16)$	$[25,4;3,2]$	40	25	812

Table 1: Integral T–type surgeries [27]

- Many of the knots in the tables are torus knots or iterated torus knots. As in [6], we use $[p,q;r,s]$ to denote the (p,q)–cable of the torus knot $T(r,s)$.
- There are two hyperbolic pretzel knots in the tables: $P(-2,3,7)$ and $P(-2,3,9)$.
- Following [38], let $K(p,q,r,n)$ be the twist torus knot obtained by applying n full twists to r parallel strings in $T(p,q)$, where p,q are coprime integers, $q>|p|\geq 2$, $0\leq r\leq p+q$, and $n\in\mathbb{Z}$. It is proved in [38] that the $pq+n(p+q)^2$–surgery on $K(p,q,p+q,n)$ yields a Seifert-fibered manifold with base orbifold $S^2(|p|,q,|n|)$.
- Let $B(p,q;a)$ be the Berge knot [4] whose dual is the simple knot [50] in the homology class a in $L(p,q)$.
- Three knots $K_2^{\#}$, K_3^*, K_3 are from [9, Section 10].
- One knot $K(1,1,0)$ is from [21, Section 4], which is also the knot K_1 given in [6, Proposition 18].
- Three knots are primitive/Seifert knots from [5]. We will use the notation there, starting with "P/SF$_d$ KIST".

Remark 2.2 In [6], the authors enumerated all finite surgeries on iterated torus knots. However, one case was missed in their list: the 58–surgery on $[19,3;3,2]$ yields the O–type manifold $\mathbb{T}(58/9)$. This mistake was inherited in [27], where the author found a knot on which 58–surgery yields $\mathbb{T}(58/9)$, but she thought the knot was hyperbolic because this case was not listed in [6].

p	Y	K_0	g	$\det(K)$	$\Delta_K''(1)$
2	$\mathbb{T}(2)$	$T(3,2)$	1	3	2
10	$\mathbb{T}(10)$	$T(3,2)$	1	3	2
10	$-\mathbb{T}(10)$	$T(4,3)$	3	3	10
14	$-\mathbb{T}(14/3)$	$T(4,3)$	3	3	10
22	$\mathbb{T}(22/3)$	$P(-2,3,9)$	6	3	34
38	$\mathbb{T}(38/7)$	$B(39,16;16)$	13	3	114
46	$-\mathbb{T}(46/7)$	$B(45,19;8)$	16	3	162
50	$\mathbb{T}(50/9)$	$[17,3;3,2]$	19	3	210
58	$\mathbb{T}(58/9)$	$[19,3;3,2]$	21	3	258
62	$-\mathbb{T}(62/11)$	P/SF$_d$ KIST III $(-5,-3,-2,-1,1)$	23	3	306
70	$\mathbb{T}(70/11)$	$B(71,27;11)$	27	3	402
86	$\mathbb{T}(86/15)$	$-K(3,4,7,-2)$	33	3	586
94	$-\mathbb{T}(94/15)$	$-K(2,3,5,-4)$	35	3	690
106	$\mathbb{T}(106/17)$	$K(2,3,5,4)$	41	3	914
106	$-\mathbb{T}(106/17)$	$[35,3;4,3]$	43	3	906
110	$-\mathbb{T}(110/19)$	$K(3,4,7,2)$	45	3	1002
110	$-\mathbb{T}(110/19)$	$[37,3;4,3]$	45	3	1002
146	$\mathbb{T}(146/25)$	$[29,5;3,2]$	61	3	1730
154	$\mathbb{T}(154/25)$	$[31,5;3,2]$	65	3	1970

Table 2: Integral O–type surgeries [27]

In Table 4, we omit the two half-integer finite surgeries on $T(3,2)$ as it was already known that only $T(3,2)$ can have such surgeries [45].

Lemma 2.3 *In Tables 1–5, any two sample knots expressed in different notation are different knots with different Alexander polynomials.*

Proof By Known Facts 1.2(1), all the knots in S^3 with finite surgeries are fibered. So if two sample knots in the tables have different genera, they must have different Alexander polynomials. Below we will compare the Alexander polynomials of sample knots with the same genera. Since $\det(K) = |\Delta_K(-1)|$, we often just compare $\det(K)$ and $\Delta_K''(1)$.

- $g = 3$ There are two knots $T(4,3)$ and $T(7,2)$. They have different $\det(K)$.
- $g = 4$ There are two knots $T(5,3)$ and $T(9,2)$. They have different $\det(K)$.

p	Y	K_0	g	$\det(K)$	$\Delta_K''(1)$
1	$\mathbb{T}(1)$	$T(3,2)$	1	3	2
7	$\mathbb{T}(7/2)$	$T(5,2)$	2	5	6
11	$\mathbb{T}(11)$	$T(3,2)$	1	3	2
13	$-\mathbb{T}(13/3)$	$T(5,2)$	2	5	6
13	$\mathbb{T}(13/3)$	$T(5,3)$	4	1	16
17	$\mathbb{T}(17/2)$	$T(5,3)$	4	1	16
17	$-\mathbb{T}(17/2)$	$P(-2,3,7)$	5	1	24
19	$\mathbb{T}(19/4)$	$[9,2;3,2]$	6	9	28
23	$\mathbb{T}(23/3)$	$P(-2,3,9)$	6	3	34
29	$\mathbb{T}(29/4)$	$[15,2;3,2]$	9	15	64
29	$-\mathbb{T}(29/4)$	$-K(1,1,0)$	9	11	62
37	$-\mathbb{T}(37/7)$	$-K(-2,5,3,-3)$	11	1	96
37	$\mathbb{T}(37/7)$	$[19,2;5,2]$	13	19	114
43	$-\mathbb{T}(43/8)$	$[21,2;5,2]$	14	21	134
47	$\mathbb{T}(47/7)$	K_3^*	16	1	168
49	$\mathbb{T}(49/9)$	$[16,3;3,2]$	18	9	188
59	$\mathbb{T}(59/9)$	$[20,3;3,2]$	22	9	284
83	$-\mathbb{T}(83/13)$	P/SF$_d$ KIST V$(1,-2,-1,2,2)$	32	1	552
91	$\mathbb{T}(91/16)$	$[23,4;3,2]$	37	23	692
101	$\mathbb{T}(101/16)$	$[25,4;3,2]$	40	25	812
113	$-\mathbb{T}(113/18)$	$-K(3,5,8,-2)$	45	1	1024
113	$\mathbb{T}(113/18)$	$-$P/SF$_d$ KIST V$(-3,-2,-1,2,2)$	46	1	1048
119	$-\mathbb{T}(119/19)$	$-K(2,3,5,-5)$	45	1	1112
131	$\mathbb{T}(131/21)$	$K(2,3,5,5)$	51	1	1392
133	$\mathbb{T}(133/23)$	$[44,3;5,3]$	55	3	1434
137	$-\mathbb{T}(137/22)$	$-K(2,5,7,-3)$	55	3	1506
137	$\mathbb{T}(137/22)$	$[46,3;5,3]$	57	3	1554
143	$\mathbb{T}(143/23)$	$K(3,5,8,2)$	60	1	1696
157	$\mathbb{T}(157/27)$	$[39,4;5,2]$	65	39	1996
157	$-\mathbb{T}(157/27)$	$K(2,5,7,3)$	65	3	2034
163	$-\mathbb{T}(163/28)$	$[41,4;5,2]$	68	41	2196
211	$\mathbb{T}(211/36)$	$[35,6;3,2]$	91	35	3642
221	$\mathbb{T}(221/36)$	$[37,6;3,2]$	96	37	4062

Table 3: Integral I–type surgeries [27]

$\frac{1}{2}p$	Y	K_0	g	$\frac{1}{2}p$	Y	K_0	g
$17/2$	$-\mathbb{T}(17/2)$	$T(5,2)$	2	$53/2$	$\mathbb{T}(53/8)$	$[13,2;3,2]$	8
$23/2$	$\mathbb{T}(23/3)$	$T(5,2)$	2	$77/2$	$-\mathbb{T}(77/12)$	$[19,2;5,2]$	13
$43/2$	$\mathbb{T}(43/8)$	$[11,2;3,2]$	7	$83/2$	$\mathbb{T}(83/13)$	$[21,2;5,2]$	14
$45/2$	$\mathbb{T}(45/8)$	$[11,2;3,2]$	7	$103/2$	$\mathbb{T}(103/18)$	$[17,3;3,2]$	19
$51/2$	$\mathbb{T}(51/8)$	$[13,2;3,2]$	8	$113/2$	$\mathbb{T}(113/18)$	$[19,3;3,2]$	21

Table 4: Half-integral surgeries [36]

p	K_0	g	$\det(K)$	$\Delta_K''(1)$
4	$T(3,2)$	1	3	2
8	$T(3,2)$	1	3	2
8	$T(5,2)$	2	5	6
12	$T(5,2)$	2	5	6
12	$T(7,2)$	3	7	12
16	$T(7,2)$	3	7	12
16	$T(9,2)$	4	9	20
20	$T(9,2)$	4	9	20
20	$T(11,2)$	5	11	30
24	$T(11,2)$	5	11	30
24	$T(13,2)$	6	13	42
28	$T(13,2)$	6	13	42
28	$-K(1,1,0)$	9	11	62
28	K_3	8	5	54
28	$T(15,2)$	7	15	56
32	$T(15,2)$	7	15	56
32	$T(17,2)$	8	17	72

Table 5: Integral D–type p–surgeries with $p \leq 32$ [18]

- $\boldsymbol{g=5}$ There are two knots $P(-2,3,7)$ and $T(11,2)$. They have different $\det(K)$.
- $\boldsymbol{g=6}$ There are three knots $[9,2;3,2]$, $P(-2,3,9)$ and $T(13,2)$. They have different $\det(K)$.
- $\boldsymbol{g=7}$ There are two knots $[11,2;3,2]$ and $T(15,2)$. They have different $\det(K)$.

- $g = 8$ There are three knots $[13, 2; 3, 2]$, $T(17, 2)$ and K_3. They have different $\det(K)$.
- $g = 9$ There are two knots $[15, 2; 3, 2]$ and $-K(1, 1, 0)$. They have different $\det(K)$.
- $g = 13$ There are two knots $B(39, 16; 16)$ and $[19, 2; 5, 2]$. They have different $\det(K)$.
- $g = 16$ There are two knots $B(45, 19; 8)$ and K_3^*. They have different $\det(K)$.
- $g = 18$ There are two knots $K_2^{\#}$ and $[16, 3; 3, 2]$. They have different $\det(K)$.
- $g = 45$ There are four knots. The two knots with O–type surgery can be distinguished from the two with I–type surgery via $\det(K)$. The two knots with I–type surgery have different $\Delta_K''(1)$. The two knots with O–type surgery have different Alexander polynomials

$$\Delta_{K(3,4,7,2)}(t) = 1 - (t^2 + t^{-2}) + (t^3 + t^{-3}) - (t^5 + t^{-5}) + (t^7 + t^{-7}) + \cdots,$$
$$\Delta_{[37,3;4,3]}(t) = 1 - (t^2 + t^{-2}) + (t^3 + t^{-3}) - (t^5 + t^{-5}) + (t^6 + t^{-6}) + \cdots.$$

- $g = 55$ There are two knots $[44, 3; 5, 3]$ and $K(2, 5, 7, -3)$. They have different $\Delta_K''(1)$.
- $g = 65$ There are three knots $[31, 5; 3, 2]$, $[39, 4; 5, 2]$ and $K(2, 5, 7, 3)$. They have different $\Delta_K''(1)$.

This finishes the proof. □

Remark 2.4 In Tables 1–5, any sample knot expressed in a notation different from that of a torus knot or a cable over a torus knot is a hyperbolic knot. This is because that those torus knots and cables over torus knots appearing in Tables 1–5 constitute the set of all nonhyperbolic knots in S^3 which admit T–type, O–type or I–type finite surgeries, or D–type integer p–surgeries with $p \leq 32$. By Lemma 2.3, all other sample knots are different from these nonhyperbolic ones.

3 $\mathrm{PSL}_2(\mathbb{C})$–character variety, Culler–Shalen norm or seminorm, and finite surgery

In this section, we briefly review some machinery and results from [16; 9; 11; 12] used in studying cyclic and finite surgeries but specialized to the case of knots in S^3. In fact, for simplicity, we shall mainly restrict our discussion to the following very special

situation: any hyperbolic knot in S^3 which is assumed to have a half-integer finite surgery. This is sufficient for our purposes in this paper.

For a finitely generated group Γ, we use $R(\Gamma)$ to denote the $\mathrm{PSL}_2(\mathbb{C})$–representation variety of Γ. (The term variety used here means complex affine algebraic set). Let $\Phi\colon \mathrm{SL}_2(\mathbb{C})\to \mathrm{PSL}_2(\mathbb{C})$ be the canonical quotient homomorphism. Given an element Υ in $\mathrm{PSL}_2(\mathbb{C})$, we have that $\Phi^{-1}(\Upsilon)=\{A,-A\}$ for some $A\in \mathrm{SL}_2(\mathbb{C})$, and we often simply write $\Upsilon=\pm A$. In particular, we may define $\mathrm{tr}^2(\Upsilon):=[\mathrm{trace}(A)]^2$, which is obviously well defined on Υ. An element $\Upsilon\in \mathrm{PSL}_2(\mathbb{C})$ is said to be parabolic if it is not the identity element $\pm I$ and satisfies $\mathrm{tr}^2(\Upsilon)=4$.

A representation $\rho\in R(\Gamma)$ is said to be irreducible if it is not conjugate to a representation whose image lies in

$$\left\{\pm\begin{pmatrix} a & b\\ 0 & a^{-1}\end{pmatrix}:\ a,b\in\mathbb{C},\ a\neq 0\right\}.$$

A representation $\rho\in R(\Gamma)$ is said to be strictly irreducible if it is irreducible and is not conjugate to a representation whose image lies in

$$\left\{\pm\begin{pmatrix} a & 0\\ 0 & a^{-1}\end{pmatrix},\pm\begin{pmatrix} 0 & b\\ -b^{-1} & 0\end{pmatrix}:\ a,b\in\mathbb{C},\ a\neq 0,\ b\neq 0\right\}.$$

For a representation $\rho\in R(\Gamma)$, its character χ_ρ is the function $\chi_\rho\colon \Gamma\to\mathbb{C}$ defined by $\chi_\rho(\gamma)=\mathrm{tr}^2(\rho(\gamma))$ for each $\gamma\in\Gamma$. Let $X(\Gamma)=\{\chi_\rho:\rho\in R(\Gamma)\}$ denote the set of characters of representations of Γ. Then $X(\Gamma)$ is also a complex affine algebraic set, usually referred as the $\mathrm{PSL}_2(\mathbb{C})$–character variety of Γ.

A character $\chi_\rho\in X(\Gamma)$ is said to be irreducible or strictly irreducible or discretely faithful or dihedral if the representation ρ has the corresponding property.

Let $t\colon R(\Gamma)\to X(\Gamma)$ denote the natural onto map defined by $t(\rho)=\chi_\rho$. Then t is a regular map between the two algebraic sets. For an element $\gamma\in\Gamma$, the function $f_\gamma\colon X(\Gamma)\to\mathbb{C}$ is defined by $f_\gamma(\chi_\rho)=\chi_\rho(\gamma)-4=\mathrm{tr}^2(\rho(\gamma))-4$ for each $\chi_\rho\in X(\Gamma)$. Each f_γ is a regular function on $X(\Gamma)$. Obviously, $\chi_\rho\in X(\Gamma)$ is a zero point of f_γ if and only if either $\rho(\gamma)=\pm I$ or $\rho(\gamma)$ is a parabolic element. It is also evident that f_γ is invariant when γ is replaced by a conjugate of γ or by the inverse of γ.

If $h\colon \Gamma\to\Gamma'$ is a surjective homomorphism between two finitely generated groups, it naturally induces an embedding of $R(\Gamma')$ into $R(\Gamma)$ and an embedding of $X(\Gamma')$ into $X(\Gamma)$. So we may simply consider $R(\Gamma')$ and $X(\Gamma')$ as subsets of $R(\Gamma)$ and $X(\Gamma)$, respectively, and write $R(\Gamma')\subset R(\Gamma)$ and $X(\Gamma')\subset X(\Gamma)$.

For a connected compact manifold Y, let $R(Y)$ and $X(Y)$ denote $R(\pi_1(Y))$ and $X(\pi_1(Y))$, respectively.

Let M be the exterior of a knot K in S^3. A slope on ∂M is called a boundary slope if there is an orientable properly embedded incompressible and boundary-incompressible surface F in M whose boundary ∂F is a nonempty set of parallel essential curves in ∂M of slope γ. For a slope γ on ∂M, we denote by $M(\gamma)$ the Dehn filling of M with slope γ. Throughout, we let μ denote the meridian slope and λ the canonical longitude slope on ∂M. We use $\Delta(\gamma_1, \gamma_2)$ to denote the distance between two slopes γ_1 and γ_2 on ∂M. We call K or M hyperbolic if the interior of M supports a complete hyperbolic metric of finite volume. Note that for any slope γ, there is a surjective homomorphism from $\pi_1(M)$ to $\pi_1(M(\gamma))$, and thus $R(M(\gamma)) \subset R(M)$ and $X(M(\gamma)) \subset X(M)$.

For the exterior M of a nontrivial knot in S^3, we consider $H_1(\partial M; \mathbb{Z}) \cong \pi_1(\partial M)$ as a subgroup of $\pi_1(M)$ which is well defined up to conjugation. Hence the function f_γ on $X(M)$ is well defined for each class $\gamma \in H_1(\partial M; \mathbb{Z})$. As f_γ is also invariant under the change of the orientation of γ, we see that f_γ is also well defined when γ is a slope in ∂M. For convenience, we will often not make a distinction among a primitive class of $H_1(\partial M; \mathbb{Z})$, the corresponding element of $\pi_1(\partial M)$ and the corresponding slope in ∂M; that is, we shall often use these terms exchangeably under the same notation.

Lemma 3.1 [9, Lemma 5.3] *Let Y be a spherical space form.*

(1) *If Y is T–type, then $X(Y)$ has exactly one irreducible character χ_ρ, and the image of ρ is isomorphic to the tetrahedral group*

$$T_{12} = \{x, y : x^2 = y^3 = (xy)^3 = 1\}.$$

(2) *If Y is O–type, then $X(Y)$ has exactly two irreducible characters χ_{ρ_1} and χ_{ρ_2}, one of ρ_1 and ρ_2 has its image isomorphic to the octahedral group*

$$O_{24} = \{x, y : x^2 = y^3 = (xy)^4 = 1\}$$

(we name this character the O–type character of $X(Y)$), *and the other has image isomorphic to the dihedral group*

$$D_6 = \{x, y : x^2 = y^2 = (xy)^3 = 1\}$$

(we name this character the D–type character of $X(Y)$).

(3) *If Y is I–type, then $X(Y)$ has exactly two irreducible characters χ_{ρ_1} and χ_{ρ_2}, and both ρ_1 and ρ_2 have image isomorphic to the icosahedral group*

$$I_{60} = \{x, y : x^2 = y^3 = (xy)^5 = 1\}.$$

Let M be the exterior of a knot in S^3 and suppose that β is a D–type, T–type, O–type or I–type finite surgery slope on ∂M. Let $\rho \in R(M(\beta)) \subset R(M)$ be an irreducible representation, and we require ρ to have image O_{24} when β is O–type. Let ϕ denote the composition $\pi_1(\partial M) \to \pi_1(M) \to \pi_1(M(\beta)) \xrightarrow{\rho} \mathrm{PSL}_2(\mathbb{C})$, and let $q = |\phi(\pi_1(\partial M))|$. Then q is uniquely associated to the finite surgery slope β, and following [13], we say more specifically that the finite surgery slope β is a $D(q)$–, $T(q)$–, $O(q)$–, or $I(q)$–type, respectively.

Remark 3.2 The number q possesses the following useful property which we shall often apply: for any slope γ on ∂M whose distance from β is divisible by q, the representation ρ factors through $\pi_1(M(\gamma))$; ie $\rho(\gamma) = \pm I$.

The following lemma can be extracted from specializing [13] to exteriors of hyperbolic knots in S^3.

Lemma 3.3 [13] *Let M be the exterior of a hyperbolic knot K in S^3 and β a finite noncyclic surgery slope of K.*

(1) *If β is D–type, it is actually $D(2)$–type, and β is an integer divisible by 4.*

(2) *If β is T–type, it is actually $T(3)$–type, and β is an integer or half-integer whose numerator is an odd integer divisible by 3.*

(3) *If β is I–type, it is $I(2)$–, $I(3)$–, or $I(5)$–type, and β is an integer or half-integer whose numerator is relatively prime to 30.*

(4) *If β is O–type, it is $O(2)$– or $O(4)$–type, and β is an even integer not divisible by 4.*

By a curve in an algebraic set, we mean an irreducible 1–dimensional algebraic subset. It is known (eg [11]) that any curve X_0 in $X(M)$ belongs to one of the following three mutually exclusive types:

(a) for each slope γ on ∂M, the function f_γ is constant on X_0;

(b) there is a unique slope γ_0 on ∂M such that the function f_{γ_0} is constant on X_0;

(c) for each slope γ on ∂M, the function f_γ is nonconstant on X_0.

We call a curve of $X(M)$ in case (a) a constant curve, in case (b) a seminorm curve and in case (c) a norm curve. Indeed as the names indicate, a seminorm curve or norm curve in $X(M)$ can be used to define a seminorm or norm, respectively, on the real 2–dimensional plane $H_1(\partial M;\mathbb{R})$ satisfying certain properties. In this paper, we only need to consider those curves in $X(M)$ which are irreducible components of $X(M)$ and contain irreducible characters. We call such a curve a nontrivial curve component of $X(M)$. In a nontrivial curve component of $X(M)$, all but finitely many characters are irreducible. But to see if a curve component in $X(M)$ is nontrivial, it suffices to check if the curve contains at least one irreducible character. Note that a norm curve component is automatically nontrivial. Also note that a norm curve component always exists for any hyperbolic knot exterior, but a constant curve or a nontrivial seminorm curve may not always exist.

For a curve X_0 in $X(M)$, let $\widetilde{X}_0$ be the smooth projective completion of X_0 and let $\phi\colon \widetilde{X}_0 \to X_0$ be the birational equivalence. The map ϕ is onto and is defined at all but finitely many points of $\widetilde{X}_0$. The points where ϕ is not defined are called ideal points. The map ϕ induces an isomorphism from the function field of X_0 to that of $\widetilde{X}_0$. In particular, every regular function f_γ on X_0 corresponds uniquely to its extension $\widetilde{f_\gamma}$ on $\widetilde{X}_0$ which is a rational function. If $\widetilde{f_\gamma}$ is not a constant function on $\widetilde{X}_0$, its degree, denoted by $\deg(\widetilde{f_\gamma})$, is equal to the number of zeros of $\widetilde{f_\gamma}$ in $\widetilde{X}_0$ counted with multiplicity; ie if $Z_x(\widetilde{f_\gamma})$ denotes the zero degree of $\widetilde{f_\gamma}$ at a point $x \in \widetilde{X}_0$, then

$$\deg(\widetilde{f_\gamma}) = \sum_{x \in \widetilde{X}_0} Z_x(\widetilde{f_\gamma}).$$

Note that if χ_ρ is a smooth point of X_0 then $\phi^{-1}(\chi_\rho)$ is a single point and the zero degree of f_γ at χ_ρ is equal to the zero degree of $\widetilde{f_\gamma}$ at $x = \phi^{-1}(\chi_\rho)$.

From now on, in this section, we make the following *special assumption*: let M be the exterior of a hyperbolic knot in S^3, and suppose M has a half-integer finite surgery slope α. It follows from [16, Theorem 2.0.3] that each of α and the meridian slope μ is not a boundary slope.

We shall identify $H_1(\partial M, \mathbb{R})$ with the real xy–plane so that $H_1(\partial M;\mathbb{Z})$ are integer lattice points (m,n) with $\mu = (1,0)$ being the meridian class and $\lambda = (0,1)$ the longitude class. So each slope p/q in ∂M corresponds to the pair of primitive elements $\pm(p,q) \in H_1(\partial M;\mathbb{Z})$.

Theorem 3.4 *Let X_1 be a norm curve component of $X(M)$. Then X_1 can be used to define a norm $\|\cdot\|_{X_1}$ on $H_1(\partial M;\mathbb{R})$, known as the Culler–Shalen norm, with the following properties:*

(1) *For each nontrivial element $\gamma=(m,n)\in H_1(\partial M;\mathbb{Z})$, we have that $\|\gamma\|_{X_1}=\deg(\widetilde{f_\gamma})\neq 0$ (thus is a positive integer).*

(2) *The norm is symmetric around the origin; ie $\|(a,b)\|_{X_1}=\|(-a,-b)\|_{X_1}$ for all $(a,b)\in H_1(\partial M;\mathbb{R})$. Let*

$$s_1=\min\{\|\gamma\|_{X_1}:\gamma\in H_1(\partial M;\mathbb{Z}),\ \gamma\neq 0\}$$

and B_1 be the set of points in $H_1(\partial M;\mathbb{R})$ with norm less than or equal to s_1. Then B_1 is a convex finite-sided polygon symmetric around the origin whose interior does not contain any nonzero element of $H_1(\partial M;\mathbb{Z})$.

(3) *If (a,b) is a vertex of B_1, then there is a boundary slope p/q in ∂M such that $\pm(p,q)$ lie on the line passing through (a,b) and $(0,0)$.*

(4) *If we normalize the area of a parallelogram spanned by any pair of generators of $H_1(\partial M;\mathbb{Z})$ to be 1, then $\operatorname{Area}(B_1)\leq 4$.*

(5) *If β is a cyclic surgery slope but is not a boundary slope, then $\beta\in\partial B_1$ (so $\|\beta\|_{X_1}=s_1$) but is not a vertex of B_1. More precisely, for each nonzero element $\gamma\in H_1(\partial M;\mathbb{Z})$ and for every point $x\in\widetilde{X}_1$, we have $Z_x(\widetilde{f_\beta})\leq Z_x(\widetilde{f_\gamma})$. In particular, the meridian slope μ has this property.*

(6) *If β is a T–type finite surgery slope but is not a boundary slope, then $\|\beta\|_{X_1}=s_1+2$ or s_1 corresponding to whether the irreducible character χ_ρ of $X(M(\beta))$* (given by Lemma 3.1(1)) *is contained in X_1 or not, respectively.*

(7) *If β is an O–type finite surgery slope but is not a boundary slope, then $\|\beta\|_{X_1}$ is either s_1+3, s_1+2, s_1+1, or s_1 corresponding to whether both, only the O–type, only the D–type, or neither of the two irreducible characters of $X(M(\beta))$* (given by Lemma 3.1(2)) *is contained in X_1, respectively.*

(8) *If β is an I–type finite surgery slope but is not a boundary slope, then $\|\beta\|_{X_1}$ is either s_1+4, s_1+2, or s_1 corresponding to whether both, only one, or neither of the two irreducible characters of $X(M(\beta))$* (given by Lemma 3.1(3)) *is contained in X_1, respectively.*

(9) *The half-integral finite surgery slope α is either T–type or I–type.*

(9a) *If α is T–type, the curve X_1 contains the unique irreducible character of $X(M(\alpha))$, $\|\alpha\|_{X_1} = \|\mu\|_{X_1} + 2 = s_1 + 2$, and $X(M)$ has no other norm curve component.*

(9b) *If α is I–type, the curve X_1 contains at least one of the two irreducible characters of $X(M(\alpha))$, and either $\|\alpha\|_{X_1} = \|\mu\|_{X_1} + 2 = s_1 + 2$ if exactly one of the two irreducible characters of $X(M(\alpha))$ is contained in X_1, or $\|\alpha\|_{X_1} = \|\mu\|_{X_1} + 4 = s_1 + 4$ if both of the two irreducible characters of $X(M(\alpha))$ are contained in X_1. Also, if $\|\alpha\|_{X_1} = \|\mu\|_{X_1} + 4 = s_1 + 4$, then $X(M)$ does not have any other norm curve component.*

Properties (1)–(5) of Theorem 3.4 originate from [16] and properties (6)–(9) from [9; 11]. As (9) was not explicitly stated in [9; 11], we give here a brief explanation. Since each of μ and α is not a boundary slope, each of them is not contained in a line passing a vertex of B_1 and the origin by (4). In particular, each of μ and β is not a vertex of B_1. By (5), we have $\|\mu\|_{X_1} = s_1$. We claim that α is not contained in B_1; ie $\|\alpha\|_{X_1} > \|\mu\|_{X_1} = s_1$. For if $\alpha = (2p+1, 2) \in B_1$, then since $\mu = (1, 0) \in B_1$ and since B_1 is a convex set, B_1 also contains the points $(p, 1)$ and $(p+1, 1)$. This would imply that the area of B_1 is ≥ 4, which by (4) would imply that B_1 is a parallelogram with $\pm\mu$ and $\pm\alpha$ as vertices, contradicting to our early conclusion. Now the conclusion that α is either T–type or I–type follows from Lemma 3.3 and all the conclusions of (9a) and (9b) follow directly from [9; 11], due essentially to the facts that each irreducible character in $X(M(\alpha))$ is a smooth point of $X(M)$ and that the zero degree of f_α at such a character is 2 while the zero degree of f_μ at such a point is 0.

Remark 3.5 Properties (6)–(8) of Theorem 3.4 are also due to similar facts: when β is a finite noncyclic slope of M, each irreducible character of $X(M(\beta))$ is a smooth point of $X(M)$ [9; 11] and thus is contained a unique component of $X(M)$, and when such a character is contained in X_1, then the zero degree of f_β at the character is 2 (except when the character is dihedral in which case the zero degree is 1) while the zero degree of f_μ at the character is 0. Moreover, if the character factors through $M(\gamma)$ for some slope γ then the zero degree of f_γ at this point is also 2 (or 1 when the character is dihedral). We shall say that the character contributes to the norm of γ by 2 (or 1) beyond the minimum norm s_1. This extended property shall also be applied later in this paper.

As M is hyperbolic, any component X_1 of $X(M)$ which contains the character of a discretely faithful representation of $\pi(M)$ is a norm curve component of $X(M)$. To

apply Theorem 3.4 more effectively we consider the set C of all (mutually distinct) norm curve components $X_1, \dots, X_k$ in $X(M)$ and let $\|\cdot\| = \|\cdot\|_{X_1} + \cdots + \|\cdot\|_{X_k}$ be the norm defined by C. In particular, C contains the orbit of X_1 under the $\mathrm{Aut}(\mathbb{C})$–action on $X(M)$; see [13, Section 5] for the $\mathrm{Aut}(\mathbb{C})$–action. In fact, under the special assumption that M has a half-integer finite surgery slope, C has at most two components. Let

$$s = \min\{\|\gamma\| : \gamma \in H_1(M, \partial M),\ \gamma \neq 0\},$$

and let B be the disk in the plane $H_1(\partial M; \mathbb{R})$ centered at origin with radius s with respect to the norm $\|\cdot\|$. Obviously, C, $\|\cdot\|$, s and B are uniquely associated to M. The following theorem follows directly from Theorem 3.4.

Theorem 3.6 *With C, $\|\cdot\|$, s and B defined above for M, we have:*

(1) *The value $s > 0$ is an integer, and B is a convex finite-sided polygon symmetric around the origin.*

(2) *If β is a cyclic slope but is not a boundary slope, then $\beta \in \partial B$ (so $\|\beta\| = s$) but is not a vertex of B. In particular, μ is such a slope.*

(3) *If β is a T–type finite surgery slope but is not a boundary slope, then $\|\beta\| = s + 2$ or s corresponding to whether the irreducible character χ_ρ of $X(M(\beta))$* (given by Lemma 3.1(1)) *is contained in C or not, respectively.*

(4) *If β is an O–type finite surgery slope, then $\|\beta\|$ is either $s+3$, $s+2$, $s+1$ or s corresponding to whether both, only the O–type, only the D–type, or neither of the two irreducible characters of $X(M(\beta))$* (given by Lemma 3.1(2)) *is contained in C, respectively.*

(5) *If β is an I–type finite surgery slope but is not a boundary slope, then $\|\beta\| = s+4$ or s corresponding to whether both or neither of the two irreducible characters of $X(M(\beta))$* (given by Lemma 3.1(3)) *is contained in C, respectively.*

(6) *The half-integral finite surgery slope α is either T–type or I–type.*

(6a) *If α is T–type, then $\|\alpha\| = \|\mu\| + 2 = s + 2$, and the irreducible character of $X(M(\alpha))$ is contained in C.*

(6b) *If α is I–type, then $\|\alpha\| = \|\mu\| + 4 = s + 4$, and both irreducible characters of $X(M(\alpha))$ are contained in C.*

We only need to note that Theorem 3.6(4) and (6b) hold because of the $\mathrm{Aut}(\mathbb{C})$–action; see [13, Remark 9.4].

Theorem 3.7 *Suppose that X_0 is a nontrivial seminorm curve component of $X(M)$* (such a curve may not always exist), *and let γ_0 be the unique slope such that f_{γ_0} is constant on X_0* (we call γ_0 the associated slope to X_0). *The curve X_0 can be used to define a seminorm $\|\cdot\|_{X_0}$ on $H_1(\partial M;\mathbb{R})$, called the Culler–Shalen seminorm, with the following properties:*

(1) $\|\gamma_0\|_{X_0}=0$, *and γ_0 is a boundary slope of M.*

(2) *For each slope $\gamma\neq\gamma_0$, we have $\|\gamma\|_{X_0}=\deg(\widetilde{f}_\gamma)\neq 0$ (so is a positive integer).*

(3) *For the meridian slope $\mu=(1,0)$, we have that $\mu\neq\gamma_0$ and that $\|\mu\|_{X_0}>0$ is minimal among all slopes $\gamma\neq\gamma_0$. More precisely, for every point $x\in\widetilde{X}_0$, $Z_x(\widetilde{f}_\mu)\leq Z_x(\widetilde{f}_\gamma)$ for each slope $\gamma\neq\gamma_0$. Furthermore, for every slope γ, we have $\|\gamma\|_{X_0}=\Delta(\gamma,\gamma_0)\|\mu\|_{X_0}$. In particular, $\Delta(\mu,\gamma_0)=1$; ie γ_0 is an integer slope.*

(4) *For the half-integer finite surgery slope α, we have $\|\alpha\|_{X_0}=\|\mu\|_{X_0}$ and thus $\Delta(\alpha,\gamma_0)=1$.*

Theorem 3.7 is contained in [11]. We only need to note that Theorem 3.7(4) holds because X_0 cannot contain any irreducible character of $X(M(\alpha))$ due to Theorem 3.6(6).

Lemma 3.8 *Let η be any one of the two integer slopes which are distance 1 from the half-integer finite surgery slope α. Then $\|\eta\|\leq s+1$ if α is T–type, and $\|\eta\|\leq s+2$ if α is I–type.*

Proof Write $\alpha=(2p+1,2)$. Then $\eta=(p,1)$ or $(p+1,1)$. We prove the case when $\alpha=(2p+1,2)$ is I–type and $\eta=(p+1,1)$. The other three cases can be treated similarly. So we have $\|\alpha\|=s+4$ by Theorem 3.6(6b). Let $B(r)$ be the norm disk in the plane $H_1(\partial M;\mathbb{R})$ centered at the origin with radius r. Then $B(s)=B$. The point $\mu=(1,0)$ lies in $\partial B(s)$. There is a positive real number a such that the point $(1+2a,0)$ has norm $\|\mu\|+4=s+4$. By the convexity of $B(r)$ with any radius r, the line segment in the plane $H_1(\partial M;\mathbb{R})$ with endpoints $(1+2a,0)$ and $(2p+1,2)$ is contained in the norm disk $B(s+4)$. It follows that the line segment with endpoints $(1+a,0)$ and $(p+1,1)$ is contained in the norm disk $B(s+2)$. □

4 Proof of Theorem 1.3

Proof of Theorem 1.3 Suppose otherwise that K is a hyperbolic knot in S^3 on which $17/2$ or $23/2$ is a finite surgery slope. We will get a contradiction from this assumption. Here is an outline of our strategy. Let α be the finite surgery slope $17/2$ or $23/2$ on K, and let δ be the slope

$$\delta = \begin{cases} 9 & \text{if } \alpha = 17/2, \\ 11 & \text{if } \alpha = 23/2. \end{cases}$$

Our first task is to show:

Proposition 4.1 *Dehn surgery on the given hyperbolic knot K with slope δ does not yield a hyperbolic 3–manifold.*

Note that by Known Facts 1.2(4) and Table 4, the knot K has genus 2. It is known that Dehn surgery with the slope δ on any hyperbolic knot in S^3 of genus 2 can never produce a lens space (Known Facts 1.2(7)) or a reducible manifold [37] or a toroidal manifold [34]. It also follows from Lemma 3.3, Known Facts 1.2(3), and Tables 1 and 3 that the δ–surgery on K cannot yield a spherical space form. Therefore, by Proposition 4.1, the δ–surgery on K must produce an irreducible Seifert-fibered space which has infinite fundamental group but does not contain incompressible tori. But this will contradict our next assertion:

Proposition 4.2 *For the given knot K, Dehn surgery with the slope δ cannot yield an irreducible Seifert-fibered space with infinite fundamental group but containing no incompressible tori.*

The rest of this section is devoted to the proofs of the above two propositions. The main tool is the character variety method.

Recall that the Lie algebra $\mathrm{sl}_2(\mathbb{C})$ of $\mathrm{SL}_2(\mathbb{C})$ consists of all 2×2 complex matrices with zero trace. The group $\mathrm{SL}_2(\mathbb{C})$ acts on $\mathrm{sl}_2(\mathbb{C})$ through the adjoint homomorphism $\mathrm{Ad}\colon \mathrm{SL}_2(\mathbb{C})\to \mathrm{Aut}(\mathrm{sl}_2(\mathbb{C}))$ given by matrix conjugation. As $-I$ acts trivially on $\mathrm{sl}_2(\mathbb{C})$, the adjoint action of $\mathrm{SL}_2(\mathbb{C})$ on $\mathrm{sl}_2(\mathbb{C})$ factors through $\mathrm{PSL}_2(\mathbb{C})$. If $\rho\in R(\Gamma)$ is a $\mathrm{PSL}_2(\mathbb{C})$–representation of a group Γ, we use $\mathrm{Ad}\circ\rho$ to denote the induced action of Γ on $\mathrm{sl}_2(\mathbb{C})$. Let $H^1(\Gamma;\mathrm{Ad}\circ\rho)$ denote the group cohomology with respect to the module $\mathrm{Ad}\circ\rho\colon \Gamma\to\mathrm{Aut}(\mathrm{sl}_2(\mathbb{C}))$. Note that for a connected compact manifold Y and $\rho\in R(Y)$, we have $H^1(Y;\mathrm{Ad}\circ\rho)\cong H^1(\pi_1(Y);\mathrm{Ad}\circ\rho)$.

Lemma 4.3 *Let M be the exterior of a knot in S^3. Suppose that for some slope β, $X(M(\beta))$ has a character χ_{ρ_0} such that*

(i) *χ_{ρ_0} is strictly irreducible,*

(ii) *$H^1(M(\beta); \mathrm{Ad}\circ\rho_0) = 0$,*

(iii) *the image of ρ_0 does not contain parabolic elements.*

Then the following conclusions hold:

(1) *Viewing χ_{ρ_0} as a point in $X(M(\beta))$, it is a 0–dimensional algebraic component of $X(M(\beta))$, and as a point in $X(M)$, it is a smooth point contained in a unique curve component X_0 of $X(M)$.*

(2) *For the curve component X_0 given in* (1)*, the function f_β is not constant on X_0. So in particular, X_0 is not a constant curve.*

(3) *The point χ_{ρ_0} is a zero point of f_β but is not a zero point of f_μ, and moreover, the zero degree of f_β at χ_{ρ_0} is 2.*

Proof The conclusion of part (1) follows from conditions (i) and (ii) and is a special case of [12, Theorem 3] (although the theorem there was stated for $\mathrm{SL}_2(\mathbb{C})$–representations, the same proof applies to $\mathrm{PSL}_2(\mathbb{C})$–representations).

The idea of proof for part (2) is essentially contained in [10] for a similar situation in the $\mathrm{SL}_2(\mathbb{C})$–setting. For the reader's convenience, we give a proof for our current situation. Suppose otherwise that f_β is constant on X_0. Then it is constantly zero on X_0 since χ_{ρ_0} is obviously a zero point of f_β. So $\rho(\beta)$ is either $\pm I$ or a parabolic element for every $\chi_\rho \in X_0$. Note that $\rho(\beta)$ cannot be $\pm I$ for all $\chi_\rho \in X_0$, for otherwise X_0 becomes a curve in $X(M(\beta))$ containing χ_{ρ_0}, which contradicts the fact that χ_{ρ_0} is an isolated point in $X(M(\beta))$. Therefore, $\rho(\beta)$ is parabolic for all but finitely many points χ_ρ in X_0. As the meridian μ commutes with β in $\pi_1(M)$, we have that $\rho(\mu)$ is either $\pm I$ or parabolic for all but finitely many points $\chi_\rho \in X_0$. Hence f_μ is also constantly zero. In particular, χ_{ρ_0} is a zero point of f_μ. But $\rho_0(\mu)$ cannot be $\pm I$, so $\rho_0(\mu)$ is parabolic. This violates condition (iii).

As we have seen in the proof of part (2), χ_{ρ_0} is a zero point of f_β but cannot be a zero point of f_μ. Combined with condition (i), the conclusion of part (3) now follows from [3, Theorem 2.1(2)]. □

Let M be the exterior of the given hyperbolic knot K and $\|\cdot\|$ be the total Culler–Shalen norm on $H_1(\partial M;\mathbb{R})$ defined in Section 3. Recall

$$s=\min\{\|\gamma\|:\gamma\in H_1(\partial M;\mathbb{Z}),\ \gamma\neq 0\}.$$

Let $B(r)$ be the norm disk in the plane $H_1(\partial M;\mathbb{R})$ centered at the origin of radius r. We already knew that each of μ and α is not a boundary slope. By Lemma 3.3, α is an I–type finite surgery slope, and by Theorem 3.6,

$$\|\mu\|=s,\quad \|\alpha\|=s+4.$$

By Lemma 3.8, we have

$$\|\delta\|\leq s+2. \tag{4-1}$$

Proposition 4.1 follows from (4-1) and the following proposition.

Proposition 4.4 *If $M(\delta)$ is a hyperbolic 3–manifold, then*

$$\|\delta\|\geq s+4.$$

Proof Note that by Mostow rigidity the closed hyperbolic 3–manifold $M(\delta)$ has exactly two discretely faithful $\mathrm{PSL}_2(\mathbb{C})$–representations ρ_1 and ρ_2, up to conjugation. Obviously, the two distinct characters χ_{ρ_1} and χ_{ρ_2} satisfy conditions (i) and (iii) of Lemma 4.3. Both of them also satisfy condition (ii) of Lemma 4.3, which is proved in [53]; compare [12, Corollary 5]. Hence Lemma 4.3 applies: each χ_{ρ_i} lies in a unique curve component X_i of $X(M)$ (note that $X_1=X_2$ is possible) as a smooth point, X_i is not a constant curve, and the zero degree of f_δ at χ_{ρ_i} is 2 but f_μ is not zero-valued at χ_{ρ_i}.

Claim *Each X_i is a norm curve component of $X(M)$.*

Proof of claim We just need to show that each X_i is not a seminorm curve. Suppose otherwise that some X_i is a seminorm curve. Let γ_i be the associated slope and $\|\cdot\|_{X_i}$ the corresponding Culler–Shalen seminorm.

By Lemma 4.3(2), $\gamma_i\neq\delta$. By Theorem 3.7(3), $\mu\neq\gamma_i$, μ has the minimal Culler–Shalen seminorm among all slopes $\gamma\neq\gamma_i$ and γ_i is an integer slope. Moreover, by Theorem 3.7(4), γ_i must be slope 8 if $\alpha=17/2$ or slope 10 if $\alpha=23/2$. So we have $\Delta(\gamma_i,\delta)=1$, which implies that $\|\delta\|_{X_i}=\|\mu\|_{X_i}$ by Theorem 3.7(3). But at χ_{ρ_i}, we have that f_δ is zero and f_μ is nonzero. Hence it follows from Theorem 3.7(3) that $\|\delta\|_{X_i}$ is strictly larger than $\|\mu\|_{X_i}$. We get a contradiction and the claim is proved. □

Hence each X_i is a norm curve component of $X(M)$ and thus is a member of the set C, which is the union of all norm curve components of $X(M)$. In particular, both χ_{ρ_1} and χ_{ρ_2} are contained in C which, by Lemma 4.3(3) and Theorem 3.4(5), implies that $\|\delta\| \geq \|\mu\| + 4 = s + 4$. This completes the proof of the proposition. □

Proposition 4.2 follows from (4-1) and the following proposition.

Proposition 4.5 *If $M(\delta)$ is an irreducible Seifert-fibered space with infinite fundamental group but containing no incompressible tori, then*

$$\|\delta\| \geq s + 4.$$

Proof Since $M(\delta)$ is an irreducible Seifert-fibered space with infinite fundamental group but containing no incompressible tori, its base orbifold is a 2–sphere with three cone points whose cone orders do not form an elliptic triple. So the base orbifold is $S^2(a,b,c)$ and $1/a + 1/b + 1/c \leq 1$. Note that the fundamental group of $\pi_1(M(\delta))$ surjects onto the orbifold fundamental group of $S^2(a,b,c)$, which is the triangle group

$$\triangle(a,b,c) = \langle x, y : x^a = y^b = (xy)^c = 1 \rangle.$$

We may assume that $a \geq b \geq c \geq 2$. Note that $M(\delta)$ has cyclic first homology of odd order. It follows that $\gcd(a,b,c) = 1$, $b \geq 3$, $a \geq 5$, and at most one of a, b, c is even. In particular, $S^2(a,b,c)$ must be a hyperbolic 2–orbifold and thus $\triangle(a,b,c)$ has a discretely faithful representation ρ_1 into $\mathrm{PSL}_2(\mathbb{R}) \subset \mathrm{PSL}_2(\mathbb{C})$. Therefore, ρ_1 is strictly irreducible and its image group does not contain parabolic elements.

On the other hand, applying [7, Addendum on page 224], we see that the triangle group $\triangle(a,b,c)$ has a nonabelian representation ρ_2 into $\mathrm{SO}(3) \subset \mathrm{PSL}_2(\mathbb{C})$. Thus ρ_2 is irreducible and its image does not contain parabolic elements. It is easy to check that ρ_2 is also strictly irreducible for otherwise ρ_2 would be a dihedral representation, contradicting the fact that $M(\delta)$ has odd order first homology.

Evidently ρ_1 is not conjugate to ρ_2. So we have two distinct characters χ_{ρ_1} and χ_{ρ_2} in $X(\triangle(a,b,c)) \subset X(M(\delta)) \subset X(M)$. As points in $X(M(\delta))$, the two characters χ_{ρ_1} and χ_{ρ_2} both satisfy conditions (i) and (iii) of Lemma 4.3. They also both meet condition (ii) by [12, Proposition 7] (although the result there is stated for $\mathrm{SL}_2(\mathbb{C})$–representations, a similar argument works for $\mathrm{PSL}_2(\mathbb{C})$–representations). Hence Lemma 4.3 applies: each χ_{ρ_i} lies in a unique curve component X_i of $X(M)$ ($X_1 = X_2$ is possible) as a smooth point, X_i is not a constant curve and the zero degree of f_δ at χ_{ρ_i} is 2 but f_μ is not zero valued at χ_{ρ_i}.

We can now argue exactly as in the proof of Proposition 4.4 to show that each X_i is a norm curve component of $X(M)$, which leads to the conclusion that $\|\delta\| \geq s+4$. □

The proof of Theorem 1.3 is now completed. □

5 Proof of Theorem 1.4, part I

In this section, we prove Theorem 1.4 in case there is a half-integer finite surgery slope on the given knot K. We actually show the following theorem, which provides more information in this case.

Theorem 5.1 *Suppose that K is a hyperbolic knot in S^3 which admits a half-integer finite surgery slope α.*

(1) *α is one of the following slopes:*

$$\{43/2,[11,2;3,2]\},\ \{45/2,[11,2;3,2]\},\ \{51/2,[13,3;3,2]\},\ \ \{53/2,[13,3;3,2]\},$$
$$\{77/2,[19,2;5,2]\},\ \{83/2,[21,2;5,2]\},\ \{103/2,[17,3;3,2]\},\ \{113/2,[19,3;3,2]\}.$$

Here each sample knot attached to a slope in the list plays the same role as before: the same surgery slope on the sample knot yields the same spherical space form, and K has the same knot Floer homology as the sample knot.

(2) *There is at most one other nontrivial finite surgery slope β, and if there is one, it is an integer slope distance 1 from α. The only possible pairs for such α and β are*

$$\{43/2,21,[11,2;3,2]\},\quad \{53/2,27,[13,2;3,2]\},$$
$$\{103/2,52,[17,3;3,2]\},\quad \{113/2,56,[19,3;3,2]\}$$

when β is noncyclic, and

$$\{45/2,23,[11,2;3,2]\},\quad \{51/2,25,[13,2;3,2]\},$$
$$\{77/2,39,[19,2;5,2]\},\quad \{83/2,41,[21,2;5,2]\}$$

when β is cyclic. Here each sample knot attached to a pair plays the same role as before: the same surgery slopes on the sample knot yield the same spherical space forms, and K has the same knot Floer homology as the sample knot.

The proof uses mainly character variety techniques based on Known Facts 1.1 and 1.2. First we need to prepare a few more lemmas.

Lemma 5.2 *Let K be a knot in S^3. Suppose the Alexander polynomial $\Delta_K(t)$ of K has a simple root $\xi = e^{i\theta}$ on the unit circle in the complex plane of order n. Then the knot exterior M of K has a reducible nonabelian $\mathrm{PSL}_2(\mathbb{C})$–representation ρ such that:*

(1) *$\rho(\lambda) = \pm I$, and $\rho(\mu)$ has order n.*

(2) *The character χ_ρ of ρ is contained in a unique nontrivial curve component X_0 of $X(M)$ and is a smooth point of X_0. Moreover, X_0 is either a seminorm curve or a norm curve.*

(3) *For any slope γ on ∂M, if f_γ is nonconstant on X_0 and if the reducible nonabelian character χ_ρ is a zero point of f_γ, then the zero degree of f_γ at χ_ρ is at least 2.*

Proof (1) It was known a long time ago [51; 14] that the exterior of a knot K in S^3 has a reducible, nonabelian $\mathrm{PSL}_2(\mathbb{C})$–representation ρ with $\rho(\lambda) = \pm I$ and $\rho(\mu) = \pm\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right)$ if and only if $\Delta_K(a^2) = 0$. Hence the conclusions of (1) follow from the given conditions.

(2) It was shown in [22] that the given reducible character χ_ρ is an endpoint of a (real) curve of irreducible $\mathrm{SO}(3)$ characters and also an endpoint of a (real) curve of irreducible $\mathrm{PSL}_2(\mathbb{R})$ characters. Later [30], it was shown that for the given reducible nonabelian representation ρ, the space of group 1–cocycles $Z^1(\pi_1(M), \mathrm{Ad}\circ\rho)$ is 4–dimensional, ρ is contained in a unique 4–dimensional component R_0 of $R(M)$ as a smooth point, and χ_ρ is contained in a unique 1–dimensional nontrivial component X_0 of $X(M)$ and is a smooth point of X_0 (although in [30], the above conclusions are given for SL_2–representation and character varieties, the same conclusions also hold in PSL_2 setting; see [8, Theorem 4.1; 29]). The fact that X_0 is not a constant curve is proved in [15, Lemma 7.3(d)]. So X_0 is either a seminorm or a norm curve component of $X(M)$.

(3) First note that since

$$\dim_{\mathbb{C}} Z^1(\pi_1(M), \mathrm{Ad}\circ\rho) = \dim_{\mathbb{C}} R_0 = 4,$$

the Zariski tangent space of R_0 at ρ can be identified with $Z^1(\pi_1(M), \mathrm{Ad}\circ\rho)$. By [8, Theorem 4.1], there is an analytic 2–disk D smoothly embedded in R_0 containing ρ such that $D \cap t^{-1}(\chi_\rho) = \{\rho\}$ and $t|_D$ is an analytic isomorphism onto a smooth 2–disk neighborhood of χ_ρ in X_0. One can choose a smooth path ρ_s in $D \subset R_0$

depending differentiably on a real parameter s close to 0, passing through ρ at $s=0$, of the form

$$\rho_s = \pm \exp(su + O(s^2))\rho$$

for some $u \in Z^1(\pi_1(M), \mathrm{Ad}\circ\rho)$; see [25]. Now letting $\sigma(s) = t(\rho_s) \subset X_0$ and calculating as in [9, Section 4], we get

$$f_\gamma(\sigma(s)) = [\mathrm{trace}(\rho_s(\gamma))]^2 - 4 = 2\,\mathrm{trace}(u(\gamma)^2)s^2 + O(s^3),$$

which implies that the zero degree of f_γ at χ_ρ is at least 2 (applying [9, Lemma 4.8]). Here we have used the fact that $\rho(\pi_1(\partial M))$ is a cyclic group of finite order, and thus $f_\gamma(\chi_\rho) = 0$ means $\rho(\gamma) = \pm I$. □

Remark 5.3 Distinct roots of $\Delta_K(t)$ lying on the upper-half unit circle of the complex plane give rise to distinct reducible nonabelian characters because these characters have distinct real values in $(0,1)$ when valued on the meridian μ of the knot.

Remark 5.4 The curve X_0 given in Lemma 5.2(2) is either a norm or seminorm curve on which f_μ is nonconstant. Note that the reducible nonabelian character χ_ρ given in Lemma 5.2 is not a zero point of f_μ. Now suppose that μ is not a boundary slope, so it has the minimal norm or seminorm. Then if f_γ is nonconstant on X_0 and $f_\gamma(\chi_\rho) = 0$ for some slope γ, then the point χ_ρ contributes to the norm or seminorm of γ at least by 2 beyond the norm or seminorm of μ. Also note that χ_ρ is a zero of f_γ if and only if the numerator of γ is divisible by n (which is the order of $\rho(\mu)$).

Lemma 5.5 *Let M be the exterior of a hyperbolic knot K in S^3. Suppose that M admits two I–type surgery slopes β_1 and β_2. Then as points in $X(M)$, the set of two irreducible characters of $X(M(\beta_1))$ is equal to the set of two irreducible characters of $X(M(\beta_2))$* (see Lemma 3.1(3)). *Hence in particular, β_1 and β_2 are of the same $I(q)$–type, q divides $\Delta(\beta_1,\beta_2)$ and $q = 2$ or 3. Also $q = 3$ if and only if one of β_1 and β_2 is half-integral.*

Proof Because the fundamental group of any I–type spherical space form is of the form $I_{120}\times\mathbb{Z}_j$ and because any irreducible $\mathrm{PSL}_2(\mathbb{C})$–representation of $I_{120}\times\mathbb{Z}_j$ kills the factor $\mathbb{Z}_j$ and sends the factor I_{120} (the binary icosahedral group) onto I_{60} (the icosahedral group), the first conclusion of the lemma follows. The rest of the conclusions of the lemma follow from Lemma 3.3(3) and the distance bound 3 for finite surgery slopes on hyperbolic knots; see Known Facts 1.1(2). □

Lemma 5.6 (1) *If a knot K in S^3 admits a D–type finite surgery, then* $\det(K) > 1$.

(2) *If a knot K in S^3 admits an O–type finite surgery, then* $\det(K) = 3$.

(3) *If a knot K in S^3 admits a cyclic surgery slope with an even numerator, then* $\det(K) = 1$.

Proof The lemma follows from [33, Theorem 10] which states that for any knot K in S^3, its knot group has precisely $\frac{1}{2}(\det(K)-1)$ distinct $\mathrm{PSL}_2(\mathbb{C})$ dihedral representations, modulo conjugation, and moreover any such representation will kill any slope with even numerator. □

Now we proceed to prove Theorem 5.1. Part (1) of the theorem is just the combination of Known Facts 1.2(4) and Theorem 1.3. So the slope α is one of the eight elements in

$$\{43/2,\ 45/2,\ 51/2,\ 53/2,\ 77/2,\ 83/2,\ 103/2,\ 113/2\}.$$

We divide the proof of part (2) correspondingly into eight cases. If β is another nontrivial finite surgery slope on K, then β must be an integer slope by Known Facts 1.1(2). We shall show, in each of the eight cases, that

(1) the assumption $\Delta(\alpha,\beta) > 1$ will lead to a contradiction,

(2) there is at most one such β and $\{\alpha,\beta\}$ is one of the pairs listed in Theorem 5.1.

That each attached sample knot has the said properties will also be checked.

Before we get into the cases, we make some general notes that apply to every case. If $\Delta(\alpha,\beta) > 1$, then $\Delta(\alpha,\beta) \geq 3$, and thus $\Delta(\alpha,\beta) = 3$ and β is noncyclic by Known Facts 1.1(2)–(3). Also if $\Delta(\alpha,\beta) > 1$, then β is not a boundary slope by [16, Theorem 2.0.3]. Recall that by the same reason, each of α and μ is not a boundary slope. By Lemma 3.3, α is either a T–type or I–type slope. In fact, α is T–type if and only if its numerator is divisible by 3. Let $\|\cdot\|$ be the total Culler–Shalen norm defined by the norm curve set C and $B(r)$ the norm disk of radius r. By Theorem 3.6, $\|\mu\| = s$ has minimal norm among all slopes, μ is in $\partial B(s)$ but is not a vertex of $B(s)$, $\|\alpha\| = s+2$ if α is T–type and $\|\alpha\| = s+4$ if α is I–type.

Case 1 $\alpha = 43/2$ By Known Facts 1.2(4) and Table 4, K has the same Alexander polynomial as $[11, 2; 3, 2]$, which is

$$\Delta_K(t) = \Delta_{T(11,2)}(t)\Delta_{T(3,2)}(t^2).$$

In particular, $\det(K) = \det(T(11,2)) = 11$.

If $\Delta(\alpha,\beta)=3$, then β is either 20 or 23. If $\beta=23$, then by Lemma 3.3, β is I–type. But this is impossible by Known Facts 1.2(3), Table 3 and Lemma 2.3. If $\beta=20$, then β is D–type by Lemma 3.3. But this is impossible by Known Facts 1.2(5), Table 5 and Lemma 2.3.

So $\Delta(\alpha,\beta)=1$ and β is either 21 or 22. By Lemma 3.3 and Lemma 5.6, 22 cannot be a finite surgery slope for K. So the only possible value for β is 21.

Claim $\beta=21$ *cannot be a cyclic slope on* K.

Proof of claim By Known Facts 1.2(6), if 21 is a cyclic slope for K, then there is a Berge knot K_0 on which 21 is also a cyclic slope and K_0 has the same Alexander polynomial as K and thus as $[11,2;3,2]$. By [4, Table 1], K_0 is not hyperbolic and thus is a torus knot or a cable over torus knot. From the Alexander polynomial, we see that K_0 has to be $[11,2;3,2]$. But then 21 is not a cyclic slope for $[11,2;3,2]$ (by eg [6, Table 1]). The claim is proved. □

So $\beta=21$ can only possibly be a T–type slope by Lemma 3.3. Finally we note that 21 is a T–type slope for $[11,2;3,2]$; see Table 1. Thus in this case we arrive at $\{43/2, 21, [11,2;3,2]\}$. Theorem 5.1(2) is proved in this case.

Case 2 $\boldsymbol{\alpha=45/2}$ Then α is a T–type finite surgery slope and K has the same Alexander polynomial as $[11,2;3,2]$, which is

$$\Delta_K(t)=\Delta_{T(11,2)}(t)\Delta_{T(3,2)}(t^2).$$

If $\Delta(\alpha,\beta)=3$, then β is either 21 or 24. If $\beta=24$, then by Lemma 3.3, β is D–type. But this is impossible by Table 5 and Lemma 2.3.

When $\beta=21$, it is a T–type slope and $\|\alpha\|=s+2$. As $\Delta(\alpha,\beta)=3$, we have $\|\beta\|=\|\alpha\|$ by Lemma 3.3 and Theorem 3.6(3). By the convexity of the norm disk of any radius, the line segment with endpoints $(21,1)$ and $(45,2)$ is contained in the norm disk of radius $s+2$ and in particular the midpoint $\left(33,\frac{3}{2}\right)$ of the segment has norm less than or equal to $s+2$. As $\left\|\left(33,\frac{3}{2}\right)\right\|=\frac{3}{2}\|(22,1)\|\geq\frac{3}{2}s$, we have $\frac{3}{2}s\leq s+2$, from which we get

$$s\leq 4. \tag{5-1}$$

The integer slope 23 is distance 1 from α and thus $\|(23,1)\|\leq s+1$ by Lemma 3.8. As $s\leq 4$, the norm of the point $(1.25,0)$ is at most $s+1$. Hence the line segment

with endpoints $(23,1)$ and $(1.25,0)$ is contained in $B(s+1)$. It follows that the line segment with endpoints $(24,1)$ and $(2.25,0)$ is contained in $B(2s+1)$. In particular,

$$\|(24,1)\| \leq 2s+1. \tag{5-2}$$

On the other hand, the roots of the Alexander polynomial $\Delta_K(t)$ of K are all simple and are all roots of unity. The factor $\Delta_{T(3,2)}(t^2)$ of $\Delta_K(t)$ contains three roots on the upper-half unit circle in the complex plane: $e^{\pi i/3}$, $e^{\pi i/6}$ and $e^{5\pi i/6}$, of order 6, 12 and 12, respectively, so the corresponding three distinct reducible nonabelian $\mathrm{PSL}_2(\mathbb{C})$ characters of $X(M)$ factor through $M(24)$; see Lemma 5.2 and Remarks 5.3 and 5.4. If they are all contained in the norm curve set C then they will contribute to the norm of $(24,1)$ by at least 6 beyond $s=\|\mu\|$; see Lemma 5.2(3). Hence we have

$$\|(24,1)\| \geq s+6. \tag{5-3}$$

Combining (5-2) and (5-3), we have $s+6 \leq 2s+1$, ie $s \geq 5$, which contradicts (5-1).

Hence at least one of the above three reducible nonabelian characters, which we denote by χ_{ρ_0}, is not contained in C. By Lemma 5.2(2) and the definition of C we see that χ_{ρ_0} is contained in a nontrivial seminorm curve component X_0. Now by Theorem 3.7(3)–(4) we have that the associated boundary slope γ_0 of X_0 is an integer and $\Delta(\alpha,\gamma_0)=1$. Hence $\gamma_0=22$ or 23. But f_{γ_0} must be constantly zero on X_0 (otherwise $\widetilde{f}_\mu$ would have larger zero degree than $\widetilde{f}_{\gamma_0}$ at some point of $\widetilde{X}_0$ which is impossible by [16, Proposition 1.1.3]) and in particular is zero valued at the nonabelian reducible character χ_{ρ_0}, which implies that γ_0 is divisible by 6; see Remark 5.4. We arrive at a contradiction.

So $\Delta(\alpha,\beta)=1$ and β is either 22 or 23. For the same reasons as given in Case 1, we see that 22 cannot be a finite surgery slope and 23 cannot be a noncyclic finite surgery slope. So $\beta=23$ is possibly a cyclic slope for K. In fact, 23 is a cyclic slope for $[11,2;3,2]$. Hence in Case 2, we arrive at the pair $\{45/2,23\}$ with the sample knot $[11,2;3,2]$.

Case 3 $\alpha=51/2$ This case can be handled very similarly to Case 2, and we get the pair $\{51/2,25\}$, where 25 is a possible cyclic slope, with $[13,2;3,2]$ as a sample knot.

Case 4 $\alpha=53/2$ In this case, K has the same Alexander polynomial as the sample knot $[13,2;3,2]$. If $\Delta(\alpha,\beta)=3$, then β is either 25 or 28. By Lemma 3.3, 25 cannot be a finite surgery slope. Using Lemma 3.3, Table 5 and Lemma 2.3, we can easily rule out $\beta=28$. So $\beta=26$ or 27. By Lemma 3.3 and Lemma 5.6(2), $\beta=26$ cannot

be a noncyclic finite surgery slope and by Lemma 5.6(3), $\beta = 26$ cannot be a cyclic surgery slope. Hence $\beta = 27$ which is a T–type slope for the sample knot $[13, 2; 3, 2]$. So we just need to rule out the possibility that $\beta = 27$ is a cyclic slope for K. This can be done by checking that there is no Berge knot which has 27 as cyclic slope and has the same Alexander polynomial as $[13, 2; 3, 2]$. So in this case, we get the member $\{53/2, 27, [13, 2; 3, 2]\}$.

Case 5 $\alpha = 77/2$ The argument is pretty much similar to that for Case 2. The knot K has the same Alexander polynomial as $[19, 2; 5, 2]$, which is

$$\Delta_K(t) = \Delta_{T(19,2)}(t)\Delta_{T(5,2)}(t^2).$$

In particular, $\det(K) = \det(T(19, 2)) = 19$.

If $\Delta(\alpha, \beta) = 3$, then β is either 37 or 40. The Alexander polynomial of K has six simple roots provided by the factor $\Delta_{T(5,2)}(t^2)$ on the upper-half unit circle: $e^{k\pi i/10}$ for $k = 1, 2, 3, 6, 7, 9$; these are of order 10 or 20, and they give rise to six reducible nonabelian $\mathrm{PSL}_2(\mathbb{C})$ characters, all of which factor through $M(40)$. Hence 40 cannot be a finite surgery slope for K.

So suppose $\beta = 37$. Both α and β are of I–type. As $\|\alpha\| = s+4$, we have $\|\beta\| = s+4$ by Lemma 5.5. So the line segment with endpoints $(37, 1)$ and $(77, 2)$ is contained in $B(s + 4)$; in particular, the midpoint $\left(57, \frac{3}{2}\right)$ of the segment has norm less than or equal to $s + 4$. As $\left\|\left(57, \frac{3}{2}\right)\right\| = \frac{3}{2}\|(38, 1)\| \geq \frac{3}{2}s$, we have $\frac{3}{2}s \leq s + 4$, whence $s \leq 8$.

The integer slope 39 is distance 1 from α, and thus $\|(39, 1)\| \leq s + 2$ by Lemma 3.8. Hence the line segment with endpoints $(39, 1)$ and $(1.25, 0)$ is contained in the norm disk of radius $s + 2$. It follows that the line segment with endpoints $(40, 1)$ and $(2.25, 0)$ is contained in $B(2s + 2)$. In particular, $\|(40, 1)\| \leq 2s + 2$.

On the other hand, if the above six reducible nonabelian characters are all contained in the norm curve set C, then they would contribute to the norm of $(40, 1)$ by at least 12 beyond s; ie $\|(40, 1)\| \geq s + 12$. Hence combining the last two inequalities, we have $s+12 \leq 2s+2$; ie $s \geq 10$. We arrive at a contradiction with the earlier inequality $s \leq 8$.

So at least one of the above six reducible nonabelian characters is not contained in C in which case we can get a contradiction exactly as in Case 2.

Hence if β is another nontrivial finite surgery slope, then $\Delta(\alpha, \beta) = 1$ and β is either 38 or 39. By Lemma 3.3 and Lemma 5.6(2)–(3), 38 cannot be a finite surgery slope for K. If 39 is a finite surgery slope, it cannot be noncyclic by Lemma 3.3 and Table 1.

It could be a cyclic slope for K. In fact, it is a cyclic slope of $[19, 2; 5, 2]$ by [6, Table 1]. So in this case, $\alpha = 77/2$ and $\beta = 39$ are the only possible finite surgery slopes for K (the former an I–type and the latter a C–type), with $[19, 2; 5, 2]$ as a sample knot.

Case 6 $\alpha = 83/2$ This case can be treated very similarly to Case 5, and $\alpha = 83/2$ and $\beta = 41$ are the only possible finite surgery slopes for K (with α an I–type and β a C–type), with $[21, 2; 5, 2]$ as a sample knot.

Case 7 $\alpha = 103/2$ This is perhaps the hardest case. We know that α is I–type, and K has the same Alexander polynomial as $[17, 3; 3, 2]$, which is

$$\Delta_K(t) = \Delta_{T(17,3)}(t)\Delta_{T(3,2)}(t^3).$$

If $\Delta(\alpha, \beta) = 3$, then β is either 50 or 53. By Lemma 3.3 and Table 3, 53 cannot be a finite surgery slope for K. If $\beta = 50$ is a finite surgery slope, it is O–type by Lemma 3.3. We have $\|\alpha\| = s + 4$ and $\|\beta\| \le s + 3$ by Theorem 3.6(4) and (6). So the line segment with endpoints $(50, 1)$ and $(103, 2)$ is contained in the norm disk of radius $s + 4$, and moreover, the midpoint $\frac{3}{2}(51, 1)$ of the segment is contained in the interior of $B(s + 4)$ and thus has norm less than $s + 4$. But $\frac{3}{2}\|(51, 1)\| \ge \frac{3}{2}s$. So we have $\frac{3}{2}s < s + 4$, from which we get $s \le 7$.

As K is fibered by Known Facts 1.2(1), we may apply [24, Theorem 5.3], which asserts that there is an essential lamination in M with a degenerate slope γ_0 such that $M(\gamma)$ has an essential lamination and thus has infinite fundamental group if $\Delta(\gamma, \gamma_0) > 1$. Hence γ_0 must be the slope 51. Furthermore, by [54, Theorem 2.5] combined with the geometrization theorem of Perelman, $M(\gamma)$ is hyperbolic if $\Delta(\gamma, \gamma_0) > 2$. Hence $M(54)$ is hyperbolic. In particular, $M(54)$ has two discretely faithful characters corresponding to the hyperbolic structure which must be contained in C (by the proof of Proposition 4.4). So these two points of C contribute to the norm $\|(54, 1)\|$ by 4 beyond s. Since $\Delta(\alpha, \beta) = 3$ and β is O–type, α cannot be $I(3)$–type; see Remark 3.2. Similarly, α cannot be $I(2)$–type since $\Delta(\alpha, \mu) = 2$. Thus α is $I(5)$–type by Lemma 3.3. Hence the two irreducible characters of $M(\alpha)$ factor through $M(54)$ (see Remark 3.2 again), and these two characters are contained in C by Theorem 3.6(6b). Hence these two points of C contribute another 4 to the norm $\|(54, 1)\|$ beyond s; see Remark 3.5.

The Alexander polynomial of K has four simple roots of orders divisible by 6 (they are roots of the factor $\Delta_{T(3,2)}(t^3)$) which provide four reducible nonabelian characters which factor through $M(54)$. Let χ_{ρ_0} be the irreducible character of $M(\beta)$ such that

the image of ρ_0 is the octahedral group. By Lemma 3.3, ρ_0 also factors through $M(54)$. If all the four reducible nonabelian characters and the O–type character χ_{ρ_0} are contained in C, then $\|(54,1)\| \geq s+4+4+8+2 = s+18$. On the other hand, by Lemma 3.8, $\|(52,1)\| \leq s+2$, from which we see that $\|(54,1)\| \leq 3s+2$. So $s+18 \leq 3s+2$, ie $s \geq 8$, yielding a contradiction with the early conclusion that $s \leq 7$.

So some of the four reducible nonabelian characters or the O–type character χ_{ρ_0} are not contained in C. If some of the four reducible nonabelian characters are not contained in C, then we can get a contradiction similar to that in Case 2. Thus we may suppose that the O–type character χ_{ρ_0} is not contained in C and all the four reducible nonabelian characters are contained in C. Then the same argument as above yields $\|(54,1)\| \geq s+16$ and $s+16 \leq 3s+2$; ie $s \geq 7$. Hence $s=7$. Since χ_{ρ_0} is not contained in C, we have $\|\beta\| \leq s+1 = 8$ by Theorem 3.6(4). As $\|\alpha\| = \|(103,2)\| = s+4 = 11$, the point $(824/11, 16/11)$ (which lies in the line segment with end points $(0,0)$ and $(103,2)$) has norm 8. So the line segment with endpoints $(50,1)$ and $(824/11, 16/11)$ is contained in $B(8)$. The intersection point of this line segment with the line passing through $(0,0)$ and $(51,1)$ is $(1224/19, 24/19)$. So $\|(1224/19, 24/19)\| \leq 8$. But $\|(1224/19, 24/19)\| = \left\|\frac{24}{19}(51,1)\right\| = \frac{24}{19}\|(51,1)\| \geq \frac{24}{19}s = 168/19$. So we would have $168/19 \leq 8$, which is absurd. This final contradiction shows that 50 cannot be a finite surgery slope for K.

So β is possibly 51 or 52. If 51 is a finite noncyclic surgery slope for K, then it is T–type. But this cannot happen from Table 1 and Lemma 2.3. With a similar argument as that of the claim in Case 1, 51 cannot be a cyclic surgery slope. So β is possibly 52. In fact, 52 is a D–type surgery slope for $[17,3;3,2]$; see [6, Table 1]. From Lemma 5.6(3), 52 cannot be a cyclic surgery slope. Hence in this case, we have possibly $\alpha = 103/2$ an I–type and $\beta = 52$ a D–type, with $[17,3;3,2]$ as an sample knot.

Case 8 $\alpha = 113/2$ This case can be handled entirely as Case 7, and the only possibility is $\alpha = 113/2$ an I–type and $\beta = 56$ a D–type, with $[19,3;3,2]$ as a sample knot.

The proof of Theorem 5.1 is finished. □

6 Proof of Theorem 1.4, part II, and proof of Theorem 1.5

Theorem 1.4 is included in the combination of Theorem 5.1 and the following theorem, which we prove in this section.

Theorem 6.1 *Let K be a hyperbolic knot in S^3 which does not admit a half-integer finite surgery.*

(1) *The distance between any two integer finite surgery slopes is at most 2. Consequently, there are at most three nontrivial finite surgery slopes, and if three, they are consecutive integers.*

(2) *There are at most two integer noncyclic finite surgery slopes, and all possible such pairs of slopes are*

$$\{22, 23, P(-2,3,9)\}, \quad \{28, 29, -K(1,1,0)\}, \quad \{50, 52, [17, 3; 3, 2]\},$$
$$\{56, 58, [19, 3; 3, 2]\}, \quad \{91, 93, [23, 4; 3, 2]\}, \quad \{99, 101, [25, 4; 3, 2]\}.$$

Also included with each pair is a sample knot which has identical knot Floer homology and the same pair of finite surgeries as K.

(3) *If there are three integer finite surgery slopes on K, they must be the triple $(17, 18, 19)$, and they produce the same spherical space forms as those on the pretzel knot $P(-2,3,7)$. Also, K has the same knot Floer homology as $P(-2,3,7)$.*

(4) *If there are three integer finite surgery slopes on K, then K is the knot $P(-2,3,7)$.*

(5) *If there are two finite surgery slopes on K realizing distance 2, they must be one of the following pairs:*

$$\{17, 19, P(-2,3,7)\}, \quad \{21, 23, [11, 2; 3, 2]\}, \quad \{27, 25, [13, 2; 3, 2]\},$$
$$\{37, 39, [19, 2; 5, 2]\}, \quad \{43, 41, [21, 2; 5, 2]\}, \quad \{50, 52, [17, 3; 3, 2]\},$$
$$\{56, 58, [19, 3; 3, 2]\}, \quad \{91, 93, [23, 4; 3, 2]\}, \quad \{99, 101, [25, 4; 3, 2]\}.$$

Also included with each pair is a sample knot which has identical knot Floer homology and the same pair of finite surgeries as K.

Of course, part (4) of the theorem supersedes part (3), but to get part (4) we need to get part (3) first.

We assumed that K has no half-integer finite surgery slope, so by Known Facts 1.1(2), all nontrivial finite surgery slopes of K are integers, and their mutual distance is at most 3. Also, Known Facts 1.2(3) puts significant restrictions on possible T-, O- and I–type finite surgeries, and Known Facts 1.2(6) on cyclic surgeries. The main issue arises when a D–type finite surgery is involved, in which case our method is to apply the Casson–Walker invariant. We first make some preparations accordingly. Along the way, we shall also give a proof of Theorem 1.5.

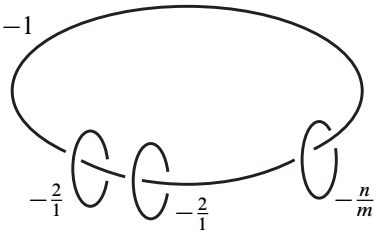


Figure 1: A surgery diagram of $P(n,m)$

Let $P(n,m)$ be the prism manifold with Seifert invariants

$$(-1;(2,1),(2,1),(n,m)),$$

where $n>1$ and $\gcd(n,m)=1$. It is easy to see that $P(n,-m)=-P(n,m)$ and $|H_1(P(n,m))|=|4m|$. As noted earlier, every D–type spherical space form is homeomorphic to some $P(n,m)$.

Given a real number x, let $\{x\}=x-\lfloor x\rfloor$ be the fractional part of x. Given a pair of coprime integers p,q with $p>0$, let $s(q,p)$ be the Dedekind sum

$$s(q,p)=\sum_{i=1}^{p-1}\Big(\Big(\frac{i}{p}\Big)\Big)\Big(\Big(\frac{iq}{p}\Big)\Big),$$

where

$$((x))=\begin{cases}\{x\}-\frac{1}{2} & \text{if } x\in\mathbb{R}\setminus\mathbb{Z},\\ 0 & \text{if } x\in\mathbb{Z}.\end{cases}$$

By [35, Proposition 6.1.1], the Casson–Walker invariant of $-P(n,m)$, when $m>0$, can be computed by the formula

$$\lambda(-P(n,m))=-\frac{1}{12}\Big(-\frac{n}{m}\Big(\frac{1}{n^2}-\frac{1}{2}\Big)-\frac{m}{n}+3+12s(m,n)\Big). \tag{6-1}$$

The Casson–Walker invariant has the following surgery formula for knots in S^3 [52, Theorem 4.2]:

$$\lambda(S^3_K(p/q))=-s(q,p)+\frac{q}{p}\Delta_K''(1). \tag{6-2}$$

Here, the Alexander polynomial $\Delta_K(t)$ is normalized to be symmetric and to obey $\Delta_K(1)=1$.

Note that the Casson–Walker invariant has the property $\lambda(-Y)=-\lambda(Y)$. In our application, it is sufficient to use only $|\lambda(Y)|$, so we do not have to worry about the orientation of the manifold involved.

Lemma 6.2 *If a knot K in S^3 admits a D–type finite surgery slope β, then the numerator of β is an integer $4m$, and the resulting prism manifold is $\varepsilon P(n,|m|)$ for some $\varepsilon\in\{\pm\}$, where n is the determinant of the knot K. In particular, if $\det(K)=1$, then K does not admit D–type finite surgery.*

Proof This lemma is just a refinement of Lemma 3.3(1) and Lemma 5.6(1). Note that $|H_1(\varepsilon P(n,m))|=|4m|$. So we just need to show that $n=\det(K)$. Theorem 10 of [33] says that for any knot K in S^3, its knot group has precisely $\frac{1}{2}(\det(K)-1)$ distinct $\mathrm{PSL}_2(\mathbb{C})$ dihedral representations, modulo conjugation, and moreover, any such representation will kill any slope with even numerator. On the other hand, for a prism manifold $\varepsilon P(n,m)$, it has precisely $\frac{1}{2}(n-1)$ distinct $\mathrm{PSL}_2(\mathbb{C})$ dihedral representations, modulo conjugation [3, Proposition D]. As a D–type finite surgery on a knot in S^3 is actually $D(2)$–type (see Lemma 3.3(1)), the set of dihedral representations of any D–type surgery manifold on a knot in S^3 is precisely the set of dihedral representations of the knot group. The conclusion of the lemma follows. □

Proof of Theorem 1.5 By Lemma 6.2, the surgery slope is $4m/q$, $(4m,q)=1$, and $n=\det(K)$. Up to reversing the orientation of $P(n,m)$, we may assume m is positive and up to replacing K by its mirror image, we may assume that $q>0$. So by Known Facts 1.2(2), $4m/q\geq 2g(K)-1$. Also by Known Facts 1.2(2), the nonzero coefficients of the Alexander polynomial of K are all ±1, and by Known Facts 1.2(1), the knot K is fibered, and thus the degree of the Alexander polynomial is $2g(K)$, which together imply that

$$n=\det(K)=|\Delta_K(-1)|\leq 2g(K)+1\leq\frac{4m}{q}+2\leq 4m+2.$$

Since $n=\det(K)$ must be an odd integer, we just need to rule out the possibility of $n=4m+1$. If $n=4m+1$, then $g(K)=2m$, $q=1$ and

$$\Delta_K(t)=1+\sum_{i=1}^{2m}(-1)^i(t^i+t^{-i}).$$

It follows that $\Delta_K''(1)=4m^2+2m$. Since

$$s(1,4m)=\sum_{i=1}^{4m-1}\Big(\frac{i}{4m}-\frac{1}{2}\Big)^2=\frac{8m^2-6m+1}{24m}, \tag{6-3}$$

using (6-2), we get

$$\lambda(S^3_K(4m))=\frac{16m^2+18m-1}{24m}.$$

On the other hand, we have

$$\begin{aligned}
s(m,4m+1) &= \sum_{i=1}^{4m}\Big(\Big(\frac{i}{4m+1}\Big)\Big)\Big(\Big(\frac{im}{4m+1}\Big)\Big)\\
&= \sum_{i=1}^{4m}\Big(\Big(\frac{4i}{4m+1}\Big)\Big)\Big(\Big(\frac{-i}{4m+1}\Big)\Big)\\
&= \sum_{k=0}^{3}\sum_{j=1}^{m}\Big(\Big(\frac{4(km+j)}{4m+1}\Big)\Big)\Big(\Big(\frac{-km-j}{4m+1}\Big)\Big)\\
&= \sum_{k=0}^{3}\sum_{j=1}^{m}\Big(\frac{4j-k}{4m+1}-\frac{1}{2}\Big)\Big(\frac{(4-k)m+1-j}{4m+1}-\frac{1}{2}\Big)\\
&= \frac{4m-m^2}{12m+3}.
\end{aligned}$$

So it follows from (6-1) that

$$\lambda(-P(4m+1,m)) = \frac{2m^2-18m+1}{24m}.$$

Thus $\lambda(S^3_K(4m)) \neq \pm\lambda(-P(4m+1,m))$ for any positive integer m. We get a contradiction. This shows $n \neq 4m+1$. □

Proof of Theorem 6.1 Suppose α and β are two integer finite surgery slopes on K. To prove part (1) of the theorem we only need to rule out the possibility of $\Delta(\alpha,\beta)=3$. So suppose that $\Delta(\alpha,\beta)=3$. Then one of α and β, say α, is an odd integer and the other, β, is an even integer. We know from Known Facts 1.1(3) that neither α nor β can be C–type. So by Lemma 3.3, α is I–type or T–type and β is O–type or D–type. In fact, α cannot be T–type for otherwise by Lemma 3.3(2) α is $T(3)$–type which means, since $\Delta(\alpha,\beta)=3$, that the irreducible representation of $M(\alpha)$ with image T_{12} also factors through $M(\beta)$; ie $M(\beta)$ has an irreducible representation with image T_{12}. But this is impossible since any O–type or D–type spherical space form does not have such representation. So α is I–type. Now from Tables 2 and 3, one can check quickly that there is no sample knot which admits an integer I–type surgery and an integer O–type surgery, distance 3 apart (one just need to check for those sample knots in Table 3 with $\det(K)=3$ and there are only 7 of them). Hence by Lemma 2.3 and Known Facts 1.2(3), there is no knot in S^3 which admits an integer I–type surgery and an integer O–type surgery, distance 3 apart. So β is a D–type slope.

So we have that α is I–type, β is D–type and they are distance 3 apart. To rule out this case, we shall apply the Casson–Walker invariant. By Known Facts 1.2(3), α is one of the slopes given in Table 3 and K has the same Alexander polynomial (in particular the same determinant) as the corresponding sample knot. We only need to consider those slopes in the table whose sample knots have determinant larger than 1 (by Lemma 6.2). We may express such a slope as $\alpha = 4m+3$ or $4m-3$ for some integer $m > 0$. So we just need to show that $S^3_K(4m)$ is not a prism manifold. To do this, we compute $\lambda(S^3_K(4m))$ using (6-2) and compute $\lambda(-P(n,m))$ for $n = \det(K)$ using (6-1), and check whether $|\lambda(S^3_K(4m))|$ is equal to $|\lambda(-P(n,m))|$. Also note that by Known Facts 1.2(5), if $4m \leq 32$, then K must have the same Alexander polynomial as a corresponding sample knot given in Table 5. This finite process of computation shows that the only possible case is when $m = 1$ and the corresponding sample knot is $T(3,2)$. It follows from [45] that K must be $T(3,2)$, contradicting the assumption that K is hyperbolic. Part (1) of the theorem is proved.

Part (2) is treated with a similar strategy. Suppose α and β are two distinct integer noncyclic finite surgery slopes of K. We are going to show that (α,β) must be one of the pairs listed in part (3) with the corresponding sample knot playing the said role, and that there cannot be a third noncyclic finite surgery on K. By part (1), $\Delta(\alpha,\beta) = 2$ or 1.

Let us first consider the case when $\Delta(\alpha,\beta) = 2$. Then α and β are both odd or both even integers. If they are both odd, then each of α and β is T–type or I–type by Lemma 3.3. Then by Known Facts 1.2(3) and Lemma 2.3, we just need to check which sample knots in Tables 1 and 3 have two slopes listed in these tables distance 2 apart. There are only four such instances:

$$\{1,3,T(3,2)\},\quad \{9,11,T(3,2)\},\quad \{91,93,[23,4;3,2]\},\quad \{99,101,[25,4;3,2]\}.$$

The first two instances can be excluded due to [45]. In the third instance, we just need to show that 92 cannot be a finite noncyclic surgery slope for the same knot K. Suppose otherwise that 92 is a noncyclic finite surgery slope for K. Then it must be a D–type surgery slope by Lemma 3.3. But by Lemma 6.2, the resulting prism manifold would be $\varepsilon P(23,23)$, which does not make sense since 23 and 23 are not relative prime integers. Thus 92 cannot be a D–type slope for K. The fourth instance can be treated exactly as the third one.

If both α and β are even, then each of α and β is O–type or D–type by Lemma 3.3. Note that α and β cannot both be D–type by Known Facts 1.1(4). We see that α and β cannot both be O–type from Known Facts 1.2(3) and Table 2. So we may assume

that α is an O–type slope and β a D–type slope. Each slope α in Table 2 can be expressed as $4m+2$. If $4m$ or $4m+4$ is a D–type slope for K, then we should have $|\lambda(P(3,m))| = |\lambda(S^3_K(4m))|$ or $|\lambda(P(3,m+1))| = |\lambda(S^3_K(4m+4))|$, respectively. Calculation using (6-1) and (6-2) shows that this happens only in four instances:

$$\{2,4,T(3,2)\},\quad \{8,10,T(3,2)\},\quad \{50,52,[17,3;3,2]\},\quad \{56,58,[19,3;3,2]\}.$$

Again the first two instances cannot happen for a hyperbolic knot due to [45]. In the third instance, 52 is indeed a D–type slope for $[17,3;3,2]$, and one can easily rule out the possibility for 51 to be a noncyclic finite surgery slope. The fourth instance can be treated exactly as the third one.

Next we consider the case when $\Delta(\alpha,\beta)=1$. As in part (1), we may assume α is a T–type or I–type slope and β an O–type or D–type slope. With a similar process as used in part (1), we only obtain the following instances (with the trefoil case excluded):

$$\{7,8,T(5,2)\},\quad \{12,13,T(5,2)\},\quad \{22,23,P(-2,3,9)\},\quad \{28,29,-K(1,1,0)\}.$$

We note that each pair of slopes are realized on the sample knot as noncyclic finite surgery slopes, and, by the result obtained in the preceding two paragraphs, there is no third noncyclic finite surgery in each instance for the same hyperbolic knot K. The first two instances with the sample knot $T(5,2)$ can be ruled out by Theorem 1.6. Part (2) of the theorem is proved.

To prove part (3), suppose that K has three integer finite surgery slopes. They are consecutive integers by part (1). At least one of them is noncyclic by Known Facts 1.1(1), and at most two of them are noncyclic by part (2).

If two of them are noncyclic, then the two slopes must be one the pairs listed in part (2) of the theorem with the corresponding sample knot. The case of $\{7,8,T(5,2)\}$ cannot happen since any hyperbolic knot cannot have a cyclic surgery slope 6 or 9 by Known Facts 1.2(7). In each of other cases, the same hyperbolic knot K can no longer have a cyclic surgery slope. For if it does, then by Known Facts 1.2(6) there will be a Berge knot having the same cyclic slope and same Alexander polynomial as the corresponding sample knot attached to the pair of noncyclic finite surgery slopes. But one can check (which is a finite process) that there does not exist such Berge knot.

So we may assume that there are exactly one noncyclic finite slope and two cyclic surgery slopes on K. By Known Facts 1.1(4)–(5), the noncyclic finite surgery slope α cannot be D–type or O–type. So α is a T–type or I–type surgery slope belonging to

Table 1 or Table 3. In particular, it is a positive integer less than or equal to 221. Also the determinant of K is 1 by Lemma 5.6(3) and so there are only 14 possible values for α. In each of the 14 cases, we check that there is no Berge knot K_0 such that K_0 admits two integer cyclic surgery slopes which form consecutive integers with α and that K_0 has the same Alexander polynomial as the sample knot attached to the slope α in Table 1 or Table 3, except for the case when $\alpha = 17$. In fact, we have:

Lemma 6.3 *If for some positive integer $p \leq 222$, we have that p and $p+1$ are cyclic surgery slopes for a nontrivial Berge knot, then p is one of the ten values* 18, 30, 31, 67, 79, 116, 128, 165, 177, 214. *If for such p, we have that p, $p+1$ and a slope α from Tables 1 or 3 form consecutive integers, then the corresponding Berge knot and the corresponding sample knot associated to α have different Alexander polynomials, except when $\alpha = 17$ ($p = 18$).*

Proof We use a Mathematica program to check the following fact: if for some positive integer $p \leq 222$ and integers q_1, q_2, both $L(p, q_1)$ and $L(p+1, q_2)$ satisfy Condition 2.1, then p is one of the ten values in the lemma. For each of these values of p, there is only one possible Alexander polynomial for the corresponding knot, which can be realized by a Berge knot. In fact, for each integer n, the Eudave-Muñoz knot $k(2, 2, n, 0)$ has two lens space surgeries with slopes $49n-18$ and $49n-19$, and the knot $k(2, -1, 2, 0)$ has two lens space surgeries with slopes 30 and 31; see [20]. Moreover, the Alexander polynomials of these knots are different from those of the knots in Tables 1 and 3 except when $p = 18$. □

When $\alpha = 17$, the sample knot is $P(-2, 3, 7)$, which does have 18 and 19 as cyclic surgery slopes. Part (3) of the theorem is proved.

Based on part (3), we can quickly prove part (4). If K admits three nontrivial finite surgery, then by part (3), the surgery slopes are the triple $17, 18, 19$, with 17 an I–type and 18, 19 C–type slopes, and K has the same knot Floer homology as $P(-2, 3, 7)$. In particular, K has genus 5. Now applying Known Facts 1.2(7) to the cyclic slope 19, we see that K is a Berge knot. But among all hyperbolic Berge knots, $P(-2, 3, 7)$ is the only one which admits cyclic slope 18 or 19. This last assertion follows from Lemma 1, Theorem 3 and the table of lens spaces of [4].

To prove part (5), assume that α and β are two integer finite surgery slopes for K with $\Delta(\alpha, \beta) = 2$. If both α and β are noncyclic, then by part (2), they are one of the pairs

$$\{50, 52, [17, 3; 3, 2]\},\ \{56, 58, [19, 3; 3, 2]\},\ \{91, 93, [23, 4; 3, 2]\},\ \{99, 101, [25, 4; 3, 2]\}.$$

So we may assume that exactly one of them, say α, is noncyclic by Known Facts 1.1(1), and α must be a T–type or I–type slope by Known Facts 1.1(4)–(5). So α is one of the slopes in Tables 1 and 3, and K has the same Alexander polynomial as the corresponding sample knot associated to α. Again, in this situation, we only need to check the following lemma.

Lemma 6.4 *If there exists a hyperbolic knot K such that K admits a cyclic surgery slope $p \leq 223$ which is distance 2 from a slope α in Table 1 or Table 3, and that K has the same Alexander polynomial as the sample knot attached to α in Table 1 or Table 3, then α, p are one of the pairs*

$$\{17, 19, P(-2,3,7)\}, \quad \{21, 23, [11, 2; 3, 2]\}, \quad \{27, 25, [13, 2; 3, 2]\},$$
$$\{37, 39, [19, 2; 5, 2]\}, \quad \{43, 41, [21, 2; 5, 2]\},$$

and each pair is realized on the attached sample knot.

Proof Again, this is proved by using a Mathematica program to check Condition 2.1 for each α–surgery in Tables 1 and 3 and a lens space $L(p,q)$ with $\Delta(\alpha, p) = 2$. We get all such pairs α, p satisfying Condition 2.1 along with the recovered Alexander polynomials, which yield corresponding sample knots in Tables 1 and 3. Such a sample knot is either a torus knot ($T(3,2)$ or $T(5,2)$), an iterated torus knot listed in the lemma or $P(-2,3,7)$. The case of torus knots can be ruled out by [45] and Known Facts 1.2(7). □

Part (5) of the theorem is proved. □

7 Proof of Theorem 1.6

We first consider the case of Theorem 1.6 when K is nonhyperbolic. This case follows easily from existing results. In fact, by Thurston's geometrization theorem for Haken manifolds, if K is not hyperbolic, then K is either a torus knot or a satellite knot. The classification of surgeries on torus knots is carried out in [39]. By [9, Corollary 1.4], if a satellite knot admits a finite surgery, then this knot is a cable of a torus knot. Finite surgeries on such cable knots are classified in [6, Theorem 7]. From these classification results, one can readily check that $T(2m+1,2)$ is the only nonhyperbolic knot in S^3 admitting a surgery to $\varepsilon S^3_{T(2m+1,2)}(4n)$ with slope $4n$ (it is easy to see this for torus knots and cables of torus knots; see [6, Table 1]).

From now on, we assume K is hyperbolic. We first get an estimate on the genus of K applying the correction terms from Heegaard Floer homology. In our current situation,

the correction terms of $S^3_K(4n)$ are given by the formula

$$d(S^3_K(4n),i) = -\frac{1}{4} + \frac{(2n-i)^2}{4n} - 2t_{\min\{i,4n-i\}}(K), \quad i = 0,1,2,\dots,4n-1. \tag{7-1}$$

For the torus knot $T(2m+1,2)$, the coefficients of its normalized Alexander polynomial are

$$a_i = \begin{cases} (-1)^{m-i} & \text{if } |i| \le m, \\ 0 & \text{otherwise.} \end{cases}$$

So

$$\begin{aligned} t_i(T(2m+1,2)) &= \begin{cases} \sum_{j=1}^{m-i} (-1)^{m-i-j}\, j & \text{for } 0 \le i < m, \\ 0 & \text{for } i \ge m, \end{cases} \\ &= \begin{cases} \lceil \frac{1}{2}(m-i) \rceil & \text{for } 0 \le i < m, \\ 0 & \text{for } i \ge m. \end{cases} \end{aligned} \tag{7-2}$$

For any $k \in \mathbb{Z}$, let $\theta(k) \in \{0,1\}$ be the reduction of k modulo 2. Let $\zeta = n-m \in \{0,1\}$. Applying (7-1) to $T(2m+1,2)$, we can compute that

$$d(S^3_{T(2m+1,2)}(4n),i) = \begin{cases} -\frac{1}{4} + i^2/(4n) + \zeta - \theta(n-\zeta-i) & \text{if } 0 \le i < n, \\ -\frac{1}{4} + (2n-i)^2/(4n) & \text{if } n \le i \le 2n, \\ d(S^3_{T(2m+1,2)}(4n), 4n-i) & \text{if } 2n < i < 4n. \end{cases} \tag{7-3}$$

Proposition 7.1 *If $S^3_K(4n) \cong \varepsilon S^3_{T(2m+1,2)}(4n)$, then $\varepsilon = +$, and $g(K) \le n$.*

Proof Since $S^3_K(4n) \cong \varepsilon S^3_{T(2m+1,2)}(4n)$, we know there exists an affine isomorphism $\phi\colon \mathbb{Z}/4n\mathbb{Z} \to \mathbb{Z}/4n\mathbb{Z}$ such that

$$d(S^3_K(4n),i) = \varepsilon d(S^3_{T(2m+1,2)}(4n), \phi(i)). \tag{7-4}$$

The map ϕ sends each Spin Spin^c structure of $S^3_K(4n)$ to a Spin Spin^c structure of $S^3_{T(2m+1,2)}(4n)$, namely, $\phi(\{0,2n\}) = \{0,2n\}$. Using (7-4), (7-1) and (7-3) for $i = 0$ and $i = 2n$, one easily sees that $\varepsilon = +$ since otherwise $t_{2n}(K)$ would not be integer valued. Since ϕ is an affine isomorphism of $\mathbb{Z}/4n\mathbb{Z}$, it must send $\{n,3n\}$ to $\{n,3n\}$. Using (7-4), (7-1) and (7-3) for $i = n$ and $i = 3n$, together with $\varepsilon = +$, one sees directly that $t_n(K) = 0$. It follows from (2-2) that $g(K) \le n$. □

Corollary 7.2 *Suppose that $S^3_K(4n) \cong \varepsilon S^3_{T(2m+1,2)}(4n)$ and that K is hyperbolic. Then $n = m = g(K)$, and K has the same Alexander polynomial as $T(2m+1,2)$. Moreover, there is a once-punctured Klein bottle properly embedded in the exterior of K of boundary slope $4m$.*

Proof By Proposition 7.1, $g(K) \le n$. Since $S^3_K(4n)$, being a prism space, contains a Klein bottle, it follows from [31, Corollary 1.3] that $4n \le 4g(K)$. So $n = g(K)$. Again by [31, Corollary 1.3], $4n$ is the boundary slope of a once-punctured Klein bottle in the exterior of K.

Next we prove that $n = m$. If not, we have $n = m+1$. By Known Facts 1.2(2),

$$\Delta_K(t) = (-1)^r + \sum_{i=1}^{r} (-1)^{i-1}(t^{n_i} + t^{-n_i}),$$

where

$$m+1 = n_1 > n_2 > \cdots > n_r > 0.$$

Lemma 6.2 implies that $|\Delta_K(-1)| = \det(K) = \det(T(2m+1,2)) = 2m+1$, so $r = m$ or $m+1$.

If $r = m$, since $|\Delta_K(-1)| = \det(K) = 2m+1$, we see that $n_i + i - 1$ has the same parity as m for any $i \in \{1, 2, \dots, m\}$. This contradicts the assumption that $n_1 = m+1$.

If $r = m+1$, then $\Delta_K(t) = (-1)^{m+1} + \sum_{i=0}^{m} (-1)^i (t^{m+1-i} + t^{-m-1+i})$, and $\det(K) = 2m+3 \neq 2m+1$; we also get a contradiction.

So we have proved $n = m$. Since $|\Delta_K(-1)| = \det(K) = 2m+1$, we must have $\Delta_K(t) = \Delta_{T(2m+1,2)}(t)$. □

Let M be the exterior of K. Let P be a once-punctured Klein bottle in M with boundary slope $4m$, provided by Corollary 7.2. Let H be a regular neighborhood of P in M, then H is a handlebody of genus 2. Let $H' = M \setminus H$. Then $F = H \cap H' = \partial H \cap \partial H'$ is a twice-punctured genus-1 surface properly embedded in M. Each component of ∂F is a simple closed curve in ∂M parallel to ∂P and thus is of slope $4m$. Note that ∂F separates ∂M into two annuli A and A' such that $\partial H = F \cup A$ and $\partial H' = F \cup A'$.

Lemma 7.3 *F is compressible in H' and is incompressible in H.*

Proof Let Q be the Seifert surface for K of genus $g(K) = m$ from Corollary 7.2. By the incompressibility of the surfaces P and Q, we may assume that P and Q intersect transversely, that $P \cap Q$ contains no circle component which bounds a disk in P or Q, and that ∂P intersects ∂Q in exactly $4m$ points. Hence $P \cap Q$ has precisely $2m$ arc components, each of which is essential in P and Q (again because the incompressibility of P and Q).

Now consider the intersection graphs G_P and G_Q determined by the surfaces P and Q as usual (see eg [31]); that is, if $\widehat{P}$ (resp. $\widehat{Q}$) is the closed surface in $M(4m)$ (resp. in $M(0)$) obtained from P (resp. Q) by capping off its boundary by a disk, then G_P (resp. G_Q) is a graph in $\widehat{P}$ (resp. $\widehat{Q}$) obtained by taking the disk $\widehat{P}\setminus P$ (resp. $\widehat{Q}\setminus Q$) as a fat vertex and taking the arc components of $P\cap Q$ as edges. In particular, each of G_P and G_Q has precisely $2m$ edges.

A simple Euler characteristic calculation shows that the graph G_Q must have at least one disk face D. Let $D' = D\setminus H = D\cap H'$. Then D' is a properly embedded disk in H'.

We claim that $\partial D'$ is an essential curve on $\partial H'$. In fact, a component C of ∂F is an essential curve on $\partial H'$ and is also an essential curve in ∂M of slope $4m$. As C has $4m$ intersection points with ∂Q, all with the same sign, if D is a k–gon face of the graph G_Q, then $\partial D'$ has k intersection points with C, all with the same sign. So $\partial D'$ is an essential curve on $\partial H'$.

The claim proved in the last paragraph implies that D' is a compressing disk for $\partial H'$.

If F is incompressible in H', then by the handle addition lemma [32], the manifold obtained by attaching a 2–handle to H' along A' will give a manifold Y' with incompressible boundary (which is a torus). The manifold Y obtained by attaching a 2–handle to H along A gives a twisted I-bundle over Klein bottle, whose boundary is incompressible. Then $M(4m)$ is the union of Y and Y' along their torus boundary and thus is a Haken manifold, a contradiction. Therefore, F is compressible in H'.

Note that H is an I-bundle over P and that F is the horizontal boundary of H with respect to the I-bundle structure. It follows that the composition of the inclusion map $F\hookrightarrow H$, and the projection map $H\to P$ with respect to the I-bundle structure is a 2–fold covering map and is thus π_1–injective. As the fundamental group of H is isomorphic to $\pi_1(P)$, it follows that F is π_1–injective in H and thus is incompressible in H. □

Lemma 7.4 *Let $\widehat{P}\subset M(4m)$ be the Klein bottle obtained by capping off ∂P with a disk, and let $\nu(\widehat{P})$ be its tubular neighborhood. Then $\nu(\widehat{P})$ is a twisted I-bundle over $\widehat{P}$, $V = M(4m)\setminus\nu(\widehat{P})$ is a solid torus, and the dual knot $K'\subset M(4m)$ can be arranged by an isotopy to intersect $\nu(\widehat{P})$ in an I-fiber and intersect V in a boundary parallel arc.*

Proof By Lemma 7.3, $F = \partial H'\setminus A'$ is compressible in H'. Let D_* be a compressing disk for F in H'. Let $\widehat{F}$ be the closed surface in $M(4m)$ obtained from F by capping off each component of ∂F with a disk. Then $\widehat{F}$ is a torus.

If ∂D_* is an inessential curve in the torus $\widehat{F}$, ie bounds a disk B in $\widehat{F}$, then B must contain both components of ∂F since ∂D_* is an essential curve in F and ∂M is incompressible in M. So compressing F with D_* produces the disjoint union of a torus T_* and an annulus A_*. As M is hyperbolic, the torus T_* bounds a solid torus in M or is parallel to ∂M in M. From the construction of T_*, we see that T_* cannot be parallel to ∂M since otherwise P would be contained in the regular neighborhood of ∂M bounded by ∂M and T_*, which is obviously impossible. So T_* bounds a solid torus in M and in fact in H'. Similarly the annulus A_* cannot be essential in M and thus must be parallel to ∂M. In fact, A_* must be parallel to A' in H'. In particular, there exists a proper disk $D' \subset H'$ whose boundary consist of an essential arc in A_* and an essential arc in A'. It follows that the surface S which is ∂H pushed slightly into the interior of M and ∂M bound a compression body. In other words, S is a genus-2 Heegaard surface of M. Attaching a 2–handle to H' along A' will cancel the 1–handle with cocore D', hence we get a solid torus V.

If ∂D_* is an essential curve in the torus $\widehat{F}$, then compressing F with D_* gives an annulus $A_\#$. Again as M is hyperbolic, $A_\#$ must be parallel to ∂M and in fact must be parallel to A' in H'. This implies that the surface S which is ∂H pushed slightly into the interior of M and ∂M bound a compression body and thus S is a genus-2 Heegaard surface of M. Let $D' \subset H'$ be a proper disk whose boundary consist of an essential arc in $A_\#$ and an essential arc in A'. Attaching a 2–handle to H' along A' will cancel the 1–handle with cocore D'; hence we get a solid torus V.

In any case, we have shown that $M(4m)$ is the union of $\nu(\widehat{P})$ and a solid torus V. Some neighborhood of $K' \cap V$ is the 2–handle added to A'; thus D' gives a parallelism between $K' \cap V$ and an arc in ∂V. Some neighborhood of $K' \cap \nu(\widehat{P})$ is the 2–handle added to A consisting of I-fibers of $\nu(\widehat{P})$. Clearly, $K' \cap \nu(\widehat{P})$ can be considered as an I-fiber of $\nu(\widehat{P})$. □

Lemma 7.5 *Let Z be the double branched cover of S^3 with ramification locus K, and let $\widetilde{K} \subset Z$ be the preimage of K. Then $Z_{\widetilde{K}}(2m)$ is a 2–fold cover of $S^3_K(4m)$, which is a lens space. Moreover, let $\widetilde{K}' \subset Z_{\widetilde{K}}(2m)$ be the dual knot of $\widetilde{K}$. Then $\widetilde{K}'$ is a 1–bridge knot with respect to the standard genus-1 Heegaard splitting of $Z_{\widetilde{K}}(2m)$.*

Proof Let $\pi\colon Z \to S^3$ be the branched covering map. Then $\pi\colon Z \setminus \widetilde{K} \to S^3 \setminus K$ is an unramified 2–fold covering map, and π maps the simple loop with slope $2m$ on $\partial\nu(\widetilde{K})$ homeomorphically to the simple loop with slope $4m$ on $\partial\nu(K)$. Thus $\pi\colon Z \setminus \widetilde{K} \to S^3 \setminus K$ can be extended to an unramified 2–fold covering map $Z_{\widetilde{K}}(2m) \to S^3_K(4m)$.

Now we look at the double cover of $S^3_K(4m)$. Since $H_1(S^3_K(4m);\mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$, this cover is unique. Let $U = \nu(\widehat{P})$ be a twisted I-bundle over $\widehat{P}$. Then $S^3_K(4m)$ is the union of U and V, where V is the solid torus in Lemma 7.4. Let $\pi_U\colon \widetilde{U} \to U$ be the 2–fold covering map induced by the covering map $\partial U \to \widehat{P}$. Clearly, $\widetilde{U}$ is homeomorphic to $T^2 \times I$. So we may construct a cover of $S^3_K(4m)$ by gluing two copies of V to $\partial\widetilde{U}$. As a result, $Z_{\widetilde{K}}(2m)$ is a lens space. Since $\widetilde{K}'$ is the preimage of K' under the covering map, it follows from Lemma 7.4 that $\widetilde{K}'$ is a 1–bridge knot. □

Lemma 7.6 *Z is an L-space.*

Proof This follows from a standard fact in Heegaard Floer homology. Notice that m is also the genus of $\widetilde{K}$. By [43, Corollary 4.2 and Remark 4.3], for any Spin^c structure $\mathfrak{s}$ over Z, there exists a Spin^c structure $[\mathfrak{s}, k]$ over $Z_{\widetilde{K}}(2m)$ such that $\widehat{\mathrm{HF}}(Z,\mathfrak{s}) \cong \widehat{\mathrm{HF}}(Z_{\widetilde{K}}(2m),[\mathfrak{s},k])$. Since $Z_{\widetilde{K}}(2m)$ is an L-space, we have $\widehat{\mathrm{HF}}(Z,\mathfrak{s}) \cong \mathbb{Z}$ for all $\mathfrak{s}$, so Z is also an L-space. □

Now we will use a result due to Hedden [28] and Rasmussen [50]. We will use the form in [28, Theorem 1.4(2)]. Although the original statement is only for knots in S^3, the same proof works for null-homologous knots in L-spaces.

Theorem 7.7 [28; 50] *Let Z_1 be an L-space and $L \subset Z_1$ a null-homologous knot with genus g. Suppose that the p–surgery on L yields an L-space Z_2, and $p \geq 2g$. Then the dual knot $L' \subset Z_2$ is Floer simple; ie* $\operatorname{rank}\widehat{\mathrm{HFK}}(Z_2, L') = \operatorname{rank}\widehat{\mathrm{HF}}(Z_2)$.

Let $L(p,q)$ be a lens space, let $V_1 \cup V_2$ be the standard genus-1 Heegaard splitting of $L(p,q)$, and let $D_i \subset V_i$ be a meridian disk such that $D_1 \cap D_2$ consists of exactly p points. A knot L in a lens space $L(p,q)$ is *simple* if it is the union of two arcs a_1, a_2, where a_i is a boundary parallel arc in V_i that is disjoint from D_i for $i = 1, 2$. In each homology class in $H_1(L(p,q))$, there exists a unique (up to isotopy) oriented simple knot.

Corollary 7.8 *$\widetilde{K}'$ is a simple knot in the lens space $Z_{\widetilde{K}}(2m)$.*

Proof Lemmas 7.5 and 7.6 and Theorem 7.7 imply that $\widetilde{K}'$ is Floer simple in $Z_{\widetilde{K}}(2m)$. Lemma 7.5 also tells us that $\widetilde{K}'$ is 1–bridge. Using [28, Proposition 3.3], we see that $\widetilde{K}'$ is simple. □

Lemma 7.9 *Let T' be the knot dual to $T = T(2m+1, 2)$ in $S^3_T(4m) \cong S^3_K(4m)$, and let $\widetilde{T}'$ be its preimage in $Z_{\widetilde{K}}(2m)$. Then*

(1) $[K'] = \pm[T']$ *or* $\pm(2m-1)[T']$ *in* $H_1(S^3_K(4m))$,

(2) $[\widetilde{K}'] = \pm[\widetilde{T}']$ *in* $H_1(Z_{\widetilde{K}}(2m))$.

Proof (1) Recall that $\mathrm{Spin}^c(S^3_K(4m))$ is an affine space over $H^2(S^3_K(4m))$. In other words, $H_1(S^3_K(4m)) \cong H^2(S^3_K(4m))$ acts on $\mathrm{Spin}^c(S^3_K(4m))$. There is a standard way to identify $\mathrm{Spin}^c(S^3_K(4m))$ with $\mathbb{Z}/4m\mathbb{Z}$ in [43, Section 4]: Let W'_{4m} be the 2–handle cobordism from $S^3_K(4m)$ to S^3, let G be a Seifert surface for K and let $\widehat{G} \subset W'_{4m}$ be obtained by capping off ∂G with a disk. For any integer i, let $\mathfrak{t}_i \in \mathrm{Spin}^c(W'_{4m})$ be the unique Spin^c structure satisfying

$$\langle c_1(\mathfrak{t}_i), [\widehat{G}]\rangle = 2i - 4m. \tag{7-5}$$

Then we have an affine isomorphism $\sigma\colon \mathrm{Spin}^c(S^3_K(4m)) \to \mathbb{Z}/4m\mathbb{Z}$ which sends $\mathfrak{t}_i|_{S^3_K(4m)}$ to $i \pmod{4m}$.

Let μ be the meridian of K. Then μ is isotopic to K' in $S^3_K(4m)$. Using (7-5), we see that

$$\sigma(\mathfrak{s} + \mathrm{PD}[\mu]) - \sigma(\mathfrak{s}) = [\mu]\cdot[G] = 1.$$

So the action of $[K']$ on $\mathrm{Spin}^c(S^3_K(4m))$ is equivalent to adding 1 in $\mathbb{Z}/4m\mathbb{Z}$. There is a similar result when we replace K with T.

We identify $S^3_K(4m)$ with $S^3_T(4m)$ by a homeomorphism $f\colon S^3_K(4m) \to S^3_T(4m)$, which induces a symmetric affine isomorphism $\phi\colon \mathrm{Spin}^c(S^3_K(4m)) \to \mathrm{Spin}^c(S^3_T(4m))$. Clearly, ϕ is equivariant with respect to the $H_1(S^3_K(4m)) = H_1(S^3_T(4m))$ action, where we identify $H_1(S^3_K(4m))$ with $H_1(S^3_T(4m))$ using f_*. If $\phi(i) = ai + b$, consider the actions of $[K']$ and $[T']$ on the Spin^c structures, so we get that $[K'] = f_*([K'])$ acts as adding a on $\mathrm{Spin}^c(S^3_T(4m))$. Since $[T']$ acts as adding 1 on $\mathrm{Spin}^c(S^3_T(4m))$, $[K'] = a[T']$.

We should have

$$d(S^3_T(4m), i) = d(S^3_K(4m), i) = d(S^3_T(4m), \phi(i))$$

for any $i \in \mathbb{Z}/4m\mathbb{Z}$, where the first equality holds since $\Delta_K(t) = \Delta_T(t)$. We will use (7-3) to compute $d(S^3_T(4m), i)$. Note that $m = n$ and $\zeta = 0$. Recall from the proof of Proposition 7.1 that $\phi(\{0, 2n\}) = \{0, 2n\}$.

When m is even, it is straightforward to check that the minimal value of $d(S^3_T(4m), i)$ is $-\frac{1}{4} + 1/(4m) - 1$, which is attained if and only if $i = 1$ or $4m-1$. So $\phi(1) = 1$ or $4m-1$. Since $\phi(0) = 0$ or $2m$, we have $a = \phi(1) - \phi(0) \in \{\pm 1, \pm(2m-1)\} \pmod{4m}$.

When m is odd, $d(S^3_T(4m), 0) \neq d(S^3_T(4m), 2m)$, so we must have $\phi(0) = 0$. We have $d(S^3_T(4m), 1) = -\frac{1}{4} + 1/(4m)$. Since m is odd, $4m + 1 \equiv 5 \pmod 8$, so

$-\frac{1}{4}+i^2/(4m)-1 \neq d(S^3_T(4m),1)$ for any integer i. It follows that $d(S^3_T(4m),i) = d(S^3_T(4m),1)$ only when $i \in \{1, 2m\pm 1, 4m-1\}$. Hence $a = \phi(1)-\phi(0) = \phi(1) \in \{\pm 1, \pm(2m-1)\} \pmod{4m}$.

In any case, we proved that $a \in \{\pm 1, \pm(2m-1)\} \pmod{4m}$, thus our conclusion holds.

(2) Let

$$\tau_*\colon H_1(S^3_K(4m)) \cong \mathbb{Z}/(4m\mathbb{Z}) \to H_1(Z_{\widetilde{K}}(2m)) \cong \mathbb{Z}/(2m(2m+1)\mathbb{Z})$$

be the transfer homomorphism, so then $[\widetilde{T}'] = \tau_*([T'])$ and $[\widetilde{K}'] = \tau_*([K'])$. Since $\gcd(4m, 2m+1) = 1$, the order of any element in the image of τ_* is a divisor of $2m$. It follows from (1) that $[\widetilde{K}'] = \pm[\widetilde{T}']$. □

Lemma 7.10 *$\widetilde{T}'$ is a simple knot in the lens space $Z_{\widetilde{K}}(2m)$.*

Proof The knot $T = T(2m+1, 2)$ is also a pretzel knot $P(2, -1, 2m+3)$. Using this pretzel diagram, it is easy to find a once-punctured Klein bottle P_T in the complement of T such that the boundary slope of P_T is $4m$. Moreover, the complement of a neighborhood of P_T is a genus-2 handlebody with respect to which T is primitive. So the same argument as in the proof of Lemma 7.5 shows that $\widetilde{T}'$ is a 1–bridge knot in the lens space $Z_{\widetilde{K}}(2m)$. Then the same argument as in the proof of Corollary 7.8 shows that $\widetilde{T}'$ is simple. □

Proof of Theorem 1.6 As mentioned in the first paragraph of this section, our theorem holds when K is nonhyperbolic. So we assume K is hyperbolic. Proposition 7.1 and Corollary 7.2 imply that $\varepsilon = +$ and $n = m$. Corollary 7.8 and Lemma 7.10 say that both $\widetilde{K}'$ and $\widetilde{T}'$ are simple knots in the lens space $Z_{\widetilde{K}}(2m)$. Now Lemma 7.9 implies that $\widetilde{K}'$ and $\widetilde{T}'$ are isotopic up to orientation reversal. But this is impossible since the complement of $\widetilde{K}'$ is hyperbolic, while the complement of $\widetilde{T}'$ is Seifert fibered. □

References

[1] **K L Baker**, *Small genus knots in lens spaces have small bridge number*, Algebr. Geom. Topol. 6 (2006) 1519–1621 MR

[2] **W Ballinger**, **C C-Y Hsu**, **W Mackey**, **Y Ni**, **T Ochse**, **F Vafaee**, *The prism manifold realization problem*, preprint (2016) arXiv

[3] **L Ben Abdelghani**, **S Boyer**, *A calculation of the Culler–Shalen seminorms associated to small Seifert Dehn fillings*, Proc. London Math. Soc. 83 (2001) 235–256 MR

[4] **J Berge**, *Some knots with surgeries yielding lens spaces*, unpublished manuscript

[5] **J Berge**, **S Kang**, *The hyperbolic P/P, P/SF_d, and P/SF_m knots in S^3*, preprint

[6] **S A Bleiler**, **C D Hodgson**, *Spherical space forms and Dehn filling*, Topology 35 (1996) 809–833 MR

[7] **S Boyer**, *On proper powers in free products and Dehn surgery*, J. Pure Appl. Algebra 51 (1988) 217–229 MR

[8] **S Boyer**, *On the local structure of* $\mathrm{SL}(2,\mathbb{C})$*–character varieties at reducible characters*, Topology Appl. 121 (2002) 383–413 MR

[9] **S Boyer**, **X Zhang**, *Finite Dehn surgery on knots*, J. Amer. Math. Soc. 9 (1996) 1005–1050 MR

[10] **S Boyer**, **X Zhang**, *Cyclic surgery and boundary slopes*, from "Geometric topology" (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc., Providence, RI (1997) 62–79 MR

[11] **S Boyer**, **X Zhang**, *On Culler–Shalen seminorms and Dehn filling*, Ann. of Math. 148 (1998) 737–801 MR

[12] **S Boyer**, **X Zhang**, *On simple points of character varieties of* 3*–manifolds*, from "Knots in Hellas '98" (C M Gordon, V F R Jones, L H Kauffman, S Lambropoulou, J H Przytycki, editors), Ser. Knots Everything 24, World Sci., River Edge, NJ (2000) 27–35 MR

[13] **S Boyer**, **X Zhang**, *A proof of the finite filling conjecture*, J. Differential Geom. 59 (2001) 87–176 MR

[14] **G Burde**, *Darstellungen von Knotengruppen*, Math. Ann. 173 (1967) 24–33 MR

[15] **M Culler**, **N M Dunfield**, *Orderability and Dehn filling*, preprint (2016) arXiv

[16] **M Culler**, **C M Gordon**, **J Luecke**, **P B Shalen**, *Dehn surgery on knots*, Ann. of Math. 125 (1987) 237–300 MR

[17] **J C Dean**, *Small Seifert-fibered Dehn surgery on hyperbolic knots*, Algebr. Geom. Topol. 3 (2003) 435–472 MR

[18] **M I Doig**, *Finite knot surgeries and Heegaard Floer homology*, Algebr. Geom. Topol. 15 (2015) 667–690 MR

[19] **M I Doig**, *On the number of finite p/q–surgeries*, Proc. Amer. Math. Soc. 144 (2016) 2205–2215 MR

[20] **M Eudave-Muñoz**, *Non-hyperbolic manifolds obtained by Dehn surgery on hyperbolic knots*, from "Geometric topology" (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc., Providence, RI (1997) 35–61 MR

[21] **M Eudave-Muñoz**, *On hyperbolic knots with Seifert fibered Dehn surgeries*, Topology Appl. 121 (2002) 119–141 MR

[22] **C D Frohman**, **E P Klassen**, *Deforming representations of knot groups in* SU(2), Comment. Math. Helv. 66 (1991) 340–361 MR

[23] **D Gabai**, *Foliations and the topology of* 3*–manifolds, III*, J. Differential Geom. 26 (1987) 479–536 MR

[24] **D Gabai**, **U Oertel**, *Essential laminations in* 3*–manifolds*, Ann. of Math. 130 (1989) 41–73 MR

[25] **W M Goldman**, *The symplectic nature of fundamental groups of surfaces*, Adv. in Math. 54 (1984) 200–225 MR

[26] **J E Greene**, *The lens space realization problem*, Ann. of Math. 177 (2013) 449–511 MR

[27] **L Gu**, *Integral finite surgeries on knots in S^3*, PhD thesis, California Institute of Technology (2014) MR Available at `https://search.proquest.com/docview/1550893042`

[28] **M Hedden**, *On Floer homology and the Berge conjecture on knots admitting lens space surgeries*, Trans. Amer. Math. Soc. 363 (2011) 949–968 MR

[29] **M Heusener**, **J Porti**, *Deformations of reducible representations of* 3*–manifold groups into* $\mathrm{PSL}_2(\mathbb{C})$, Algebr. Geom. Topol. 5 (2005) 965–997 MR

[30] **M Heusener**, **J Porti**, **E Suárez Peiró**, *Deformations of reducible representations of* 3*–manifold groups into* $\mathrm{SL}_2(\mathbf{C})$, J. Reine Angew. Math. 530 (2001) 191–227 MR

[31] **K Ichihara**, **M Teragaito**, *Klein bottle surgery and genera of knots*, Pacific J. Math. 210 (2003) 317–333 MR

[32] **W Jaco**, *Adding a* 2*–handle to a* 3*–manifold: an application to property R*, Proc. Amer. Math. Soc. 92 (1984) 288–292 MR

[33] **E P Klassen**, *Representations of knot groups in* SU(2), Trans. Amer. Math. Soc. 326 (1991) 795–828 MR

[34] **S Lee**, *Toroidal surgeries on genus-two knots*, J. Knot Theory Ramifications 18 (2009) 1205–1225 MR

[35] **C Lescop**, *Global surgery formula for the Casson–Walker invariant*, Annals of Mathematics Studies 140, Princeton Univ. Press (1996) MR

[36] **E Li**, **Y Ni**, *Half-integral finite surgeries on knots in S^3*, Ann. Fac. Sci. Toulouse Math. 24 (2015) 1157–1178 MR

[37] **D Matignon**, **N Sayari**, *Longitudinal slope and Dehn fillings*, Hiroshima Math. J. 33 (2003) 127–136 MR

[38] **K Miyazaki**, **K Motegi**, *On primitive/Seifert-fibered constructions*, Math. Proc. Cambridge Philos. Soc. 138 (2005) 421–435 MR

[39] **L Moser**, *Elementary surgery along a torus knot*, Pacific J. Math. 38 (1971) 737–745 MR

[40] **Y Ni**, *Knot Floer homology detects fibred knots*, Invent. Math. 170 (2007) 577–608 MR

[41] **Y Ni**, **X Zhang**, *Characterizing slopes for torus knots*, Algebr. Geom. Topol. 14 (2014) 1249–1274 MR

[42] **P Ozsváth**, **Z Szabó**, *Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary*, Adv. Math. 173 (2003) 179–261 MR

[43] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and knot invariants*, Adv. Math. 186 (2004) 58–116 MR

[44] **P Ozsváth**, **Z Szabó**, *On knot Floer homology and lens space surgeries*, Topology 44 (2005) 1281–1300 MR

[45] **P Ozsváth**, **Z Szabó**, *The Dehn surgery characterization of the trefoil and the figure-eight knot*, preprint (2006) arXiv

[46] **P S Ozsváth**, **Z Szabó**, *Knot Floer homology and rational surgeries*, Algebr. Geom. Topol. 11 (2011) 1–68 MR

[47] **G Perelman**, *The entropy formula for the Ricci flow and its geometric applications*, preprint (2002) arXiv

[48] **G Perelman**, *Ricci flow with surgery on three-manifolds*, preprint (2003) arXiv

[49] **J A Rasmussen**, *Floer homology and knot complements*, PhD thesis, Harvard University (2003) MR Available at `https://search.proquest.com/docview/305332635`

[50] **J Rasmussen**, *Lens space surgeries and L-space homology spheres*, preprint (2007) arXiv

[51] **G de Rham**, *Introduction aux polynômes d'un nœud*, Enseign. Math. 13 (1967) 187–194 MR

[52] **K Walker**, *An extension of Casson's invariant*, Annals of Mathematics Studies 126, Princeton Univ. Press (1992) MR

[53] **A Weil**, *On discrete subgroups of Lie groups, II*, Ann. of Math. 75 (1962) 578–602 MR

[54] **Y-Q Wu**, *Sutured manifold hierarchies, essential laminations, and Dehn surgery*, J. Differential Geom. 48 (1998) 407–437 MR

Department of Mathematics, California Institute of Technology
Pasadena, CA, United States

Department of Mathematics, University at Buffalo
Buffalo, NY, United States

`yini@caltech.edu, xinzhang@math.buffalo.edu`

Received: 22 November 2016 Revised: 20 June 2017

msp

Algebraic & Geometric Topology 18 (2018) 493–524

A characterization for asymptotic dimension growth

GOULNARA ARZHANTSEVA
GRAHAM A NIBLO
NICK WRIGHT
JIAWEN ZHANG

We give a characterization for asymptotic dimension growth. We apply it to CAT(0) cube complexes of finite dimension, giving an alternative proof of Wright's result on their finite asymptotic dimension. We also apply our new characterization to geodesic coarse median spaces of finite rank and establish that they have subexponential asymptotic dimension growth. This strengthens a recent result of Špakula and Wright.

20F65, 20F67, 20F69, 51F99

1 Introduction

The concept of asymptotic dimension was first introduced by Gromov [16] in 1992 as a coarse analogue of the classical topological covering dimension. It started to attract much attention in 1998 when Yu [31] proved that the Novikov higher signature conjecture holds for groups with finite asymptotic dimension (FAD). A lot of groups and spaces are known to have finite asymptotic dimension. Among those are, for instance, finitely generated abelian groups, free groups of finite rank, Gromov hyperbolic groups (see Gromov [15] and Roe [25]), mapping class groups (see Bestvina, Bromberg and Fujiwara [6]), CAT(0) cube complexes of finite dimension (see Wright [30]) — see Bell and Dranishnikov [4] for an excellent survey of these and other results. Recently Behrstock, Hagen and Sisto [2] introduced the powerful new notion of hierarchically hyperbolic spaces and showed that these have finite asymptotic dimension, recovering a number of the above results, including, notably, mapping class groups and a number of CAT(0) cube complexes.

On the other hand, there are many groups and spaces with infinite asymptotic dimension. Examples are the wreath product $\mathbb{Z} \wr \mathbb{Z}$, the Grigorchuk group (see Smith [28]), the Thompson groups, etc. Generalizing FAD, Dranishnikov [14] defined the asymptotic dimension growth for a space; if the asymptotic dimension growth function is eventually constant then the space has FAD. Dranishnikov showed that the wreath product of a finitely generated nilpotent group with a finitely generated FAD group has polynomial

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.493

asymptotic dimension growth. He also showed that polynomial asymptotic dimension growth implies Yu's Property A, and, hence, the coarse Baum–Connes conjecture, provided the space has bounded geometry; see Yu [32]. Later, Ozawa [23] extended this result to spaces with subexponential growth; see also Oppenheim [22]. Bell [3] analyzed how the asymptotic dimension growth function is affected by various group-theoretical constructions.

In this paper, we give an alternative characterization for the asymptotic dimension growth function which is inspired by Brown and Ozawa's proof of Property A for Gromov's hyperbolic groups [10, Theorem 5.3.15], which is in turn inspired by Kaimanovich [18]. We use this to study two notable examples: CAT(0) cube complexes of finite dimension and coarse median spaces of finite rank.

The techniques used to study these examples are developments of those used by Špakula and Wright [29] to establish Property A for uniformly locally finite coarse median spaces of finite rank. As a byproduct, we obtain a new proof of finite asymptotic dimension for CAT(0) cube complexes which allows one to explicitly construct the required controlled covers. This compares with Wright's original proof [30], which is discussed below.

CAT(0) cube complexes are a nice class of nonpositively curved spaces, first studied by Gromov [15], who gave a purely combinatorial condition for recognizing the nonpositive curvature of cube complexes. Many well-known groups act properly on CAT(0) cube complexes. For instance, right-angled Artin groups, many small cancellation groups, and Thompson's groups admit such actions. This makes it possible to deduce properties of these groups from the corresponding properties of the CAT(0) cube complexes.

In 2010, Wright [30] proved that the asymptotic dimension of a CAT(0) cube complex X is bounded by its dimension. He proved this by constructing a family of ε–Lipschitz cobounded maps to CAT(0) cube complexes of (at most) the same dimension indexed by $\varepsilon > 0$. We use our characterization for finite asymptotic dimension to give a direct proof of this result. Namely, we construct uniformly bounded covers with suitable properties. Being more explicit, this proof loses, however, the sharp bound on the asymptotic dimension. Thus, we give an alternative proof of the following nonquantitative variant of Wright's theorem:

Theorem 1.1 *Let X be a* CAT(0) *cube complex of finite dimension; then X has finite asymptotic dimension.*

The key point in our approach is to analyze the normal cube path distance on the cube complex, introduced by Niblo and Reeves [19]. We consider the ball with respect

to the normal cube path distance rather than to the ordinary edge-path distance. We decompose such a ball into intervals and use induction on the dimension in order to construct some "separated" net satisfying a suitable consistency property. In the process, we give a detailed analysis of normal balls and normal spheres (ie balls and spheres with respect to the normal cube path distance). See Section 4 for all details.

Our second application is to coarse median spaces, a coarse variant of classical median spaces introduced by Bowditch [7]. The notion of a coarse median group leads to a unified viewpoint on several interesting classes of groups, including Gromov's hyperbolic groups, mapping class groups, and CAT(0) cubical groups. Bowditch showed that hyperbolic spaces are exactly coarse median spaces of rank 1, and mapping class groups are examples of coarse median spaces of finite rank [7]. He also established interesting properties for coarse median spaces such as rapid decay, the property of having quadratic Dehn function, etc.

Intuitively, a coarse median space is a metric space equipped with a ternary operator (called the coarse median), in which every finite subset can be approximated by a finite median algebra. In these approximations the coarse median is approximated by an actual median with the distortion being controlled by the metric. This extends Gromov's observation that, in a δ–hyperbolic space, finite subsets can be well approximated by finite trees.

Recently, Špakula and Wright [29] proved that a coarse median space with finite rank and at most exponential volume growth has Property A. Following their proof and using our characterization for asymptotic dimension growth, we obtain the following result:

Theorem 1.2 *Let X be a geodesic coarse median space with finite rank and at most exponential volume growth, then X has subexponential asymptotic dimension growth.*

Hierarchically hyperbolic spaces are examples of coarse median spaces — see Behrstock, Hagen and Sisto [1] — hence our theorem is broader in scope, though with a weaker conclusion, than the finite asymptotic dimension result proven in Behrstock, Hagen and Sisto [2]. We expect the following general result:

Conjecture 1.3 *Every geodesic coarse median space with finite rank has finite asymptotic dimension.*

By a result of Ozawa [23], subexponential asymptotic dimension growth implies Property A; thus, our theorem strengthens the result of [29].

The paper is organized as follows. In Section 2, we give some preliminaries on asymptotic dimension growth, CAT(0) cube complexes, and coarse median spaces. In Section 3, we provide a characterization of the asymptotic dimension growth function, and, as a special case, give a characterization of finite asymptotic dimension. Sections 4 and 5 deal with CAT(0) cube complexes: in Section 4, we study normal balls and spheres which are essential in our approach to prove Theorem 1.1 in Section 5. Section 6 deals with the coarse median case, and we prove Theorem 1.2 there.

Acknowledgements The authors would like to thank the referee for careful and detailed comments and suggestions, which much improved the paper.

Arzhantseva was partially supported by the Simons Foundation, while visiting the Isaac Newton Institute (Cambridge). Zhang was partially supported by the European Research Council (ERC) grant of Arzhantseva *ANALYTIC* no. 259527 and by the Royal Society Sino–British Fellowship Trust.

2 Preliminaries

2.1 Asymptotic dimension

The notion of asymptotic dimension was first introduced by Gromov in 1993 [16] as a coarse analogue of the classical Lebesgue topological covering dimension. See also [4].

Let (X,d) be a metric space and $r>0$. We call a family $\mathcal{U}=\{U_i\}$ of subsets in X *r–disjoint* if $d(U,U')\geq r$ for any $U\neq U'$ in $\mathcal{U}$, where $d(U,U')=\inf\{d(x,x'):x\in U,x'\in U'\}$. We write

$$\bigsqcup_{r\text{–disjoint}} U_i$$

for the union of $\{U_i\}$. A family $\mathcal{V}$ is said to be *uniformly bounded* if $\operatorname{mesh}(\mathcal{V})=\sup\{\operatorname{diam}(V):V\in\mathcal{V}\}$ is finite. Let $\mathcal{U}=\{U_i\}$ be a cover of X and $r>0$. We define the *r–multiplicity* of $\mathcal{U}$, denoted by $m_r(\mathcal{U})$, to be the minimal integer n such that for any $x\in X$, the ball $B(x,r)$ intersects at most n elements of $\mathcal{U}$. As usual, $m(\mathcal{U})$ denotes the *multiplicity* of a cover $\mathcal{U}$, that is, the maximal number of elements of $\mathcal{U}$ with a nonempty intersection.

The *Lebesgue number* $L(\mathcal{U})$ of the cover $\mathcal{U}$ is defined to be

$$\inf_{x\in X}\sup\{r>0\mid \exists U\in\mathcal{U} \text{ such that } B(x,r)\subseteq U\}.$$

Definition 2.1 [16] We say that the *asymptotic dimension* of a metric space X does not exceed n and we write $\operatorname{asdim} X \leqslant n$ if, for every $r > 0$, the space X can be covered by $n+1$ subspaces $X_0, X_1, \dots, X_n$ and each X_i can be further decomposed into some r–disjoint uniformly bounded subspaces:

$$X = \bigcup_{i=0}^{n} X_i, \quad X_i = \bigsqcup_{\substack{r\text{–disjoint} \\ j\in\mathbb{N}}} X_{ij} \quad \text{and} \quad \sup_{i,j} \operatorname{diam} X_{ij} < \infty.$$

We say $\operatorname{asdim} X = n$ if $\operatorname{asdim} X \leqslant n$ and $\operatorname{asdim} X$ is not less than n.

Example 2.2 [21; 25; 5] (1) $\operatorname{asdim} \mathbb{Z}^n = n$ for all $n \in \mathbb{N}$, where $\mathbb{Z}$ is the group of integers.

(2) Those δ–hyperbolic spaces of bounded geometry, eg word hyperbolic groups, have finite asymptotic dimension. (The result in [5] actually shows something more general with a weakening of the bounded geometry condition.)

From the definition, it is easy to see that the asymptotic dimension of a subspace is at most that of the ambient space. There are other equivalent definitions of asymptotic dimension. We list one here for a later use, and guide the reader to [4] for others.

Proposition 2.3 [4] *Let X be a metric space; then* $\operatorname{asdim} X \leqslant n$ *if and only if for any $r > 0$, there exists a uniformly bounded cover $\mathcal{U}$ of X such that $m_r(\mathcal{U}) \leqslant n+1$.*

2.2 Asymptotic dimension growth

Let us consider the direct sum of infinitely many copies of the integers $G = \bigoplus_{\infty} \mathbb{Z}$. Since for any $n \in \mathbb{N}$ the group $\mathbb{Z}^n$ is contained in G, by the above-mentioned results G has infinite asymptotic dimension. In order to deal with such groups/spaces, Dranishnikov studied the following concept as a generalization of the property of having a finite asymptotic dimension:

Definition 2.4 [14] Let (X, d) be a metric space. Define a function

$$\operatorname{ad}_X(\lambda) = \min\{m(\mathcal{U}) : \mathcal{U} \text{ is a uniformly bounded cover of } X,\ L(\mathcal{U}) > \lambda\} - 1,$$

which is called the *asymptotic dimension function* of X.

Clearly ad_X is nondecreasing, and

$$\lim_{\lambda\to\infty} \operatorname{ad}_X(\lambda) = \operatorname{asdim}(X).$$

By definition, the asymptotic dimension of X is less than or equal to n for some finite n if and only if ad_X is eventually constant and $\operatorname{ad}_X \leqslant n$. If X has infinite asymptotic dimension, then both sides of the above equation are infinite.

Like in the case of the volume function, the growth type of the asymptotic dimension function is more essential than the function itself. Recall that for $f, g\colon \mathbb{R}_+ \to \mathbb{R}_+$, we write $f \preceq g$ if there exists $k \in \mathbb{N}$ such that $f(x) \leqslant k g(kx+k)+k$ for any $x > k$. We write $f \approx g$ if both $f \preceq g$ and $g \preceq f$. It is clear that $\approx$ is an equivalence relation. We define the growth type of f to be the $\approx$–equivalence class of f. Define the asymptotic dimension growth of X to be the growth type of ad_X.

By a result of Bell and Dranishnikov, the growth type of the asymptotic dimension function is a quasi-isometric invariant.

Proposition 2.5 [3; 14] *Let X and Y be two discrete metric spaces with bounded geometry. If X and Y are quasi-isometric, then $\mathrm{ad}_X \approx \mathrm{ad}_Y$. In particular, the asymptotic dimension growth is well-defined for finitely generated groups.*

We give an alternative (equivalent) definition of the asymptotic dimension growth that is used in our characterization.

Lemma 2.6 *Let X be a metric space, and define*

$$\widetilde{\mathrm{ad}}_X(\lambda) = \min\{m_\lambda(\mathcal{U}) : \mathcal{U} \text{ is a uniformly bounded cover of } X\} - 1.$$

Then $\widetilde{\mathrm{ad}}_X \approx \mathrm{ad}_X$.

Proof Given $\lambda > 0$, suppose $\mathcal{U}$ is a uniformly bounded cover of X with $L(\mathcal{U}) > \lambda$. For any $U \in \mathcal{U}$, define the *inner λ–neighbourhood* of U to be

$$N_{-\lambda}(U) = X \setminus N_\lambda(X \setminus U),$$

where N_λ denotes the usual λ–neighbourhood of the set, and we define

$$N_{-\lambda}(\mathcal{U}) = \{N_{-\lambda}(U) : U \in \mathcal{U}\}.$$

Since $L(\mathcal{U}) > \lambda$, the set $N_{-\lambda}(\mathcal{U})$ is still a cover of X and it is uniformly bounded. By definition, it is obvious that $m_\lambda(N_{-\lambda}(\mathcal{U})) \leqslant m(\mathcal{U})$, which yields $\widetilde{\mathrm{ad}}_X \preceq \mathrm{ad}_X$.

Conversely, suppose $\mathcal{U}$ is a uniformly bounded cover of X. Consider $N_\lambda(\mathcal{U})$, which has Lebesgue number not less than λ. It is easy to show $m(N_\lambda(\mathcal{U})) \leqslant m_\lambda(\mathcal{U})$, which implies $\mathrm{ad}_X \preceq \widetilde{\mathrm{ad}}_X$. □

By the preceding lemma, we can use either ad_X or $\widetilde{\mathrm{ad}}_X$ to define the asymptotic dimension growth. Recall that if there exists a polynomial (subexponential) function f such that $\mathrm{ad}_X \preceq f$, then X is said to have polynomial (subexponential) asymptotic dimension growth.

Dranishnikov has shown that polynomial asymptotic dimension growth implies Yu's Property A, and he gave a class of groups having this property.

Proposition 2.7 [14] *Let N be a finitely generated nilpotent group and G be a finitely generated group with finite asymptotic dimension. Then the wreath product $N \wr G$ has polynomial asymptotic dimension growth. In particular, $\mathbb{Z} \wr \mathbb{Z}$ has polynomial asymptotic dimension growth.*

2.3 CAT(0) cube complexes

We recall basic notions and results on the structure of CAT(0) cube complexes. We omit some details and most of the proofs but direct the readers to [9; 13; 15; 19; 27] for more information.

A *cube complex* is a polyhedral complex in which each cell is isometric to a Euclidean cube and the gluing maps are isometries. The *dimension* of the complex is the maximum of the dimensions of the cubes. For a cube complex X, we can associate it with the intrinsic pseudometric d_{int}, which is the minimal pseudometric on X such that each cube embeds isometrically. When X has finite dimension, d_{int} is a complete geodesic metric on X. See [9] for a general discussion on polyhedral complex and the associated intrinsic metric.

There is also another metric associated with X. Let $X^{(1)}$ be the 1–skeleton of X, that is a graph with the vertex set $V = X^{(0)}$. We equip V with the edge-path metric d, which is the minimal number of edges in a path connecting two given vertices. Clearly, when $X^{(1)}$ is connected, d is a geodesic metric on V. For $x, y \in V$, the *interval* is defined by $[x, y] = \{z \in V : d(x, y) = d(x, z) + d(x, y)\}$, that is, it consists of all points on any geodesic between x and y.

A geodesic metric space (X, d) is CAT(0) if all geodesic triangles in X are slimmer than the comparative triangle in the Euclidean space. For a cube complex (X, d_{int}), Gromov has given a combinatorial characterization of the CAT(0) condition [15]: X is CAT(0) if and only if it is simply connected and the link of each vertex is a flag complex (see also [9]).

Another characterization of the CAT(0) condition was obtained by Chepoi [13] (see also [26]): a cube complex X is CAT(0) if and only if for any $x, y, z \in V$, the intersection $[x, y] \cap [y, z] \cap [z, x]$ consists of a single point $\mu(x, y, z)$, which is called the *median* of x, y and z. In this case, we call the graph $X^{(1)}$ a *median graph*, and V equipped with the ternary operator m is indeed a median algebra [17]. In particular, the following equations hold for all $x, y, z, u, v \in V$:

M1. $\mu(x, x, y) = x$;

M2. $\mu(\sigma(x), \sigma(y), \sigma(z)) = \mu(x, y, z)$, where σ is any permutation of $\{x, y, z\}$;

M3. $\mu(\mu(x, y, z), u, v) = \mu(\mu(x, u, v), \mu(y, u, v), z)$.

Obviously, $\mu(x, y, z) \in [x, y]$ and $[x, y] = \{z \in V : \mu(x, y, z) = z\}$.

Lemma 2.8 *Let $x, y, z, w \in V$ be such that $z, w \in [x, y]$. Then $z \in [x, w]$ implies $w \in [z, y]$.*

Proof Since $z \in [x, w]$ and $w \in [x, y]$, we have $\mu(z, x, w) = z$ and $\mu(x, w, y) = w$. So

$$\begin{aligned}\mu(z, w, y) &= \mu(\mu(z, x, w), w, y) = \mu(\mu(z, w, y), \mu(x, w, y), w)\\ &= \mu(\mu(z, w, y), w, w) = w,\end{aligned}$$

which implies $w \in [z, y]$. □

Lemma 2.9 *For $x, y, z \in V$ and $d(z, y) = 1$, either $[x, z] \subseteq [x, y]$ or $[x, y] \subseteq [x, z]$.*

Proof By Chepoi's result [13], $X^{(1)}$ is a median graph, hence it is weakly modular (see [13]). So $d(x, y) \neq d(x, z)$, which implies $d(x, y) = d(x, z) + 1$ or $d(x, z) = d(x, y) + 1$, ie $[x, z] \subseteq [x, y]$ or $[x, y] \subseteq [x, z]$. □

We can equip a CAT(0) cubical complex X with a set of *hyperplanes* [12; 19; 20; 27]. Each hyperplane does not intersect itself, and divides the space into two halfspaces. Given two hyperplanes h and k, if the four possible intersections of halfspaces are all nonempty, then we say h crosses k, denoted by $h \pitchfork k$. This occurs if and only if h and k cross a common cube C (also denoted by $h \pitchfork C$). Furthermore [27], given any finite collection of pairwise intersecting hyperplanes, there exists a cube which all of them cross. Thus, the dimension of X, if it is finite, is the maximal number of pairwise intersecting hyperplanes. We can also define intervals in the language of hyperplanes: $[x, y]$ consists of points which lie in all halfspaces containing both x and y.

We call a subset $Y \subseteq V$ *convex* if $[x, y] \subseteq Y$ for any $x, y \in Y$. Obviously, halfspaces are convex since any geodesic crosses a hyperplane at most once [19; 27]. This also implies

$$d(x, y) = \#\{\text{ hyperplane } h : h \text{ separates } x \text{ from } y\}.$$

2.4 Coarse median spaces

According to Gromov, hyperbolic spaces can be considered locally as a coarse version of trees, in the sense that every finite subset can be approximated by a finite tree in a controlled way [15]. If one wants to approximate a space locally by finite median algebras (graphs), this would turn to the definition of coarse median spaces introduced by Bowditch. See [7; 8; 33] for details.

Definition 2.10 [7] Let (X, ρ) be a metric space, and $\mu\colon X^3 \to X$ be a ternary operation. We say that (X, ρ, μ) is a *coarse median space* and μ is a *coarse median* on X if the following conditions hold:

C1. There exist constants $K, H(0) > 0$ such that for all $a, b, c, a', b', c' \in X$,

$$\rho(\mu(a,b,c), \mu(a',b',c')) \leqslant K(\rho(a,a') + \rho(b,b') + \rho(c,c')) + H(0).$$

C2. There exists a function $H\colon \mathbb{N} \to [0, +\infty)$ with the following property: for a finite subset $A \subseteq X$ with $1 \leqslant |A| \leqslant p$, there exists a finite median algebra (Π, ρ_Π, μ_Π) and maps $\pi\colon A \to \Pi$, $\lambda\colon \Pi \to X$ such that for all $x, y, z \in \Pi$ and $a \in A$,

$$\rho(\lambda\mu_\Pi(x,y,z), \mu(\lambda x, \lambda y, \lambda z)) \leqslant h(p)$$

and

$$\rho(a, \lambda\pi a) \leqslant h(p).$$

We refer to K and H as the *parameters* of (X, ρ, μ). Furthermore, if there exists $d \in \mathbb{N}$ such that we can always choose the median algebra Π in condition (C2) above of rank at most η, then we say X has (coarse) *rank* at most η.

A finitely generated group is said to be *coarse median* if some Cayley graph is a coarse median space.

Note that, by definition, a coarse median on a group is not required to be equivariant under the group action.

Remark 2.11 According to Bowditch, without loss of generality, we may always assume that μ satisfies the median axioms (M1) and (M2): for all $a, b, c \in X$,

M1. $\mu(a,a,b) = a$;

M2. $\mu(a,b,c) = \mu(b,c,a) = \mu(b,a,c)$.

A large class of groups and spaces have been shown to be coarse median, including Gromov's hyperbolic groups, right-angled Artin groups, mapping class groups, CAT(0) cube complexes, etc [7]. Bowditch has proved that coarse median groups have the property of rapid decay [8], quadratic Dehn's function [7], etc. This yielded a unified way to prove these properties for the above-listed groups. Recently, Špakula and Wright have proved that coarse median spaces of finite rank and of at most exponential volume growth have Yu's Property A [29].

3 Characterization for asymptotic dimension growth

In this section, we establish a characterization for asymptotic dimension growth and obtain several interesting consequences of this main result. For instance, we get a characterization for a group to have finite asymptotic dimension.

Recall a metric space (X,d) is *uniformly discrete* if there exists some constant $C>0$ such that $d(x,y)>C$ for $x\neq y\in X$. Without loss of generality, we can always assume the metric only takes integer values.

Theorem 3.1 *Let (X,d) be a uniformly discrete metric space, and $f\colon \mathbb{R}_+\to\mathbb{R}_+$ be a function. Then the following are equivalent:*

(1) $\mathrm{ad}_X \preceq f$.

(2) *There exists a function $g\colon \mathbb{R}_+\to\mathbb{R}_+$ which has the same growth type as f such that $\forall l\in\mathbb{N}$, $\forall k=1,2,\dots,3l$, $\forall x\in X$, we can assign a subset $S(x,k,l)\subseteq X$ satisfying:*

(i) *$\forall l\in\mathbb{N}$, $\exists R_l>0$ such that $S(x,k,l)\subseteq B(x,R_l)$ for all $k=1,\dots,3l$.*

(ii) *$\forall l\in\mathbb{N}$, $\forall k,k'$ with $1\leqslant k\leqslant k'\leqslant 3l$, $\forall x\in X$, we have $S(x,k,l)\subseteq S(x,k',l)$.*

(iii) *$\forall x,y\in X$ with $d(x,y)\leqslant l$, we have:*

- *$S(x,k-d(x,y),l)\subset S(x,k,l)\cap S(y,k,l)$ for $k=d(x,y)+1,\dots,3l$.*
- *$S(x,k+d(x,y),l)\supset S(x,k,l)\cup S(y,k,l)$ for $k=1,\dots,3l-d(x,y)$.*

(iv) *$\forall l\in\mathbb{N}$, $\forall k=1,2,\dots,3l$, $\forall x\in X$, we have $\#S(x,k,l)\leqslant g(l)$.*

Proof (1) $\Longrightarrow$ (2) By Lemma 2.6, we can assume there exists a function $g\colon\mathbb{R}_+\to\mathbb{R}_+$ with $g\approx f$ such that for all $l\in\mathbb{N}$, there exists a uniformly bounded cover $\mathcal{U}_l$ of X with $m_{3l}(\mathcal{U}_l)\leqslant g(l)$. Given such a cover $\mathcal{U}_l=\{U_i : i\in I\}$ we will construct the sets $S(x,k,l)$ as follows. For each $i\in I$, choose $x_i\in U_i$ and, for $k=1,2,\dots,3l$ and $x\in X$, define

$$S(x,k,l)=\{x_i : B(x,k)\cap U_i\neq\varnothing\}.$$

Now let us check the four properties in condition (2):

(i) If $B(x,k)\cap U_i\neq\varnothing$, we can choose $w\in B(x,k)\cap U_i$. Now $d(w,x_i)\leqslant \mathrm{mesh}(\mathcal{U}_l)$, so $d(x,x_i)\leqslant k+\mathrm{mesh}(\mathcal{U}_l)\leqslant 3l+\mathrm{mesh}(\mathcal{U}_l)$. In other words,

$$S(x,k,l)\subseteq B(x,R_l)$$

where $R_l=3l+\mathrm{mesh}(\mathcal{U}_l)$.

(ii) This is immediate by our definition of the sets $S(x,k,l)$.

(iii) For all $x,y\in X$ with $d(x,y)\leqslant l$ and $k=d(x,y)+1,\dots,3l$, we have

$$S(x,k-d(x,y),l)=\{x_i : B(x,k-d(x,y))\cap U_i\neq\varnothing\}.$$

Now if $B(x,k-d(x,y))\cap U_i\neq\varnothing$, we can choose $z\in B(x,k-d(x,y))\cap U_i$, ie $z\in U_i$ and $d(z,x)\leqslant k-d(x,y)$. So $d(z,y)\leqslant k$, ie $z\in B(y,k)\cap U_i$. So

$B(y,k)\cap U_i\neq\varnothing$, which implies

$$S(x,k-d(x,y),l)\subset S(x,k,l)\cap S(y,k,l).$$

On the other hand, suppose that $x_j\in S(x,k',l)\cup S(y,k',l)$ for some $k'=1,\dots,3l-d(x,y)$. We can assume that $x_j\in S(y,k',l)$, ie $B(y,k')\cap U_j\neq\varnothing$, which implies $B(x,k'+d(x,y))\cap U_j\neq\varnothing$. So we have

$$S(x,k'+d(x,y),l)\supset S(x,k',l)\cup S(y,k',l).$$

(iv) For all $l\in\mathbb{N}$ $k=1,2,\dots,3l$ and $x\in X$, we have

$$\#S(x,k,l)=\#\{x_i : B(x,k)\cap U_i\neq\varnothing\}\leqslant m_{3l}(\mathcal{U}_l)\leqslant g(l).$$

(2) $\Longrightarrow$ (1) Now, given sets $S(x,k,l)$, we construct a sequence of covers $\mathcal{U}_l$ with the required properties. For $l\in\mathbb{N}$, let $H=\bigcup_{x\in X}S(x,l,l)$. For $h\in H$, we define $A_h=\{y : h\in S(y,l,l)\}$. We define $\mathcal{U}_l=\{A_h : h\in H\}$. This is a cover of X since $\forall x\in X$, if we take $h\in S(x,l,l)$, then $x\in A_h$. Since $\exists R_l>0$ such that $S(x,l,l)\subseteq B(x,R_l)$, we know that $d(h,y)\leqslant R_l$ for all $y\in A_h$, which implies $\operatorname{mesh}(\mathcal{U}_l)\leqslant R_l$. Finally, let us analyze $m_l(\mathcal{U}_l)$. For $x\in X$, consider $h\in H$ with $B(x,l)\cap A_h\neq\varnothing$. Take $y\in B(x,l)\cap A_h$, ie $d(y,x)\leqslant l$ and $h\in S(y,l,l)$. Now, by the assumptions in condition (2), we have

$$S(y,l,l)\subseteq S(x,l+d(x,y),l)\subseteq S(x,2l,l).$$

So $m_l(\mathcal{U}_l)\leqslant\#S(x,2l,l)\leqslant g(l)$. Finally, by Lemma 2.6, we have

$$\operatorname{ad}_X\approx\widetilde{\operatorname{ad}}_X\leqslant g\approx f.$$ □

Taking in the preceding theorem a constant function f, we obtain a characterization for finite asymptotic dimension.

Corollary 3.2 *Let (X,d) be a uniformly discrete metric space and $n\in\mathbb{N}$. Then the following are equivalent:*

(1) $\operatorname{asdim}X\leqslant n$.

(2) *$\forall l\in\mathbb{N}$, $\forall k=1,2,\dots,3l$, $\forall x\in X$, we can assign a subset $S(x,k,l)\subseteq X$, satisfying:*
   - (i) *$\forall l\in\mathbb{N}$, $\exists R_l>0$ such that $S(x,k,l)\subseteq B(x,R_l)$ for all $k=1,\dots,3l$.*
   - (ii) *$\forall l\in\mathbb{N}$, $\forall k,k'$ with $1\leqslant k\leqslant k'\leqslant 3l$, $\forall x\in X$, we have $S(x,k,l)\subseteq S(x,k',l)$.*
   - (iii) *$\forall x,y\in X$ with $d(x,y)\leqslant l$, we have:*
     - *$S(x,k-d(x,y),l)\subset S(x,k,l)\cap S(y,k,l)$ for $k=d(x,y)+1,\dots,3l$.*
     - *$S(x,k+d(x,y),l)\supset S(x,k,l)\cup S(y,k,l)$ for $k=1,\dots,3l-d(x,y)$.*
   - (iv) *$\forall l\in\mathbb{N}$, $\forall k=1,2,\dots,3l$, $\forall x\in X$, we have $\#S(x,k,l)\leqslant n+1$.*

Now we turn to the case when X is a graph, and obtain a characterization for finite asymptotic dimension which is easier to check.

Corollary 3.3 *Given a graph $X=(V,E)$ with vertices V and edges E, and equipped with the edge-path length metric d, the following are equivalent:*

(1) $\operatorname{asdim} X \leqslant n$.

(2) $\forall l \in \mathbb{N}$, $\forall k = 1,2,\dots,3l$, $\forall x \in X$, *we can assign a subset $S(x,k,l) \subseteq X$, satisfying:*
   (i) $\forall l \in \mathbb{N}$, $\exists R_l > 0$ *such that $S(x,k,l) \subseteq B(x,R_l)$ for all $k = 1,\dots,3l$.*
   (ii) $\forall l \in \mathbb{N}$, $\forall k,k'$ *with* $1 \leqslant k \leqslant k' \leqslant 3l$, $\forall x \in X$, *we have* $S(x,k,l) \subseteq S(x,k',l)$.
   (iii) $\forall x,y \in X$ *with* $d(x,y)=1$ (ie with x and y connected by an edge), *we have $S(y,k,l) \subseteq S(x,k+1,l)$ for all $k = 1,2,\dots,3l-1$.*
   (iv) $\forall l \in \mathbb{N}$, $\#S(x,2l,l) \leqslant n+1$.

Remark 3.4 The only distinction between the above two corollaries is that in Corollary 3.3, assumption (2)(iii) is required only for endpoints of an edge, rather than for an arbitrary pair of points as in Corollary 3.2. We point out that the preceding corollaries can be generalized to the case of arbitrary asymptotic dimension growth. We will not use such a generalization, so we omit it.

Proof of Corollary 3.3 (1) $\Longrightarrow$ (2) is implied directly by Corollary 3.2, so we focus on (2) $\Longrightarrow$ (1).

Following the proof of (2) $\Longrightarrow$ (1) in Theorem 3.1, let $H = \bigcup_{x\in X} S(x,l,l)$ for all $l \in \mathbb{N}$. And for all $h \in H$, define $A_h = \{y : h \in S(y,l,l)\}$. Define $\mathcal{U}_l = \{A_h : h \in H\}$. For all $x \in X$, if we take $h \in S(x,l,l)$ then $x \in A_h$. So $\mathcal{U}_l$ is a cover of X. By condition (2)(i), there is $R_l > 0$ such that $S(x,l,l) \subseteq B(x,R_l)$, we know $d(h,y) \leqslant R_l$ for all $y \in A_h$, which implies $\operatorname{mesh}(\mathcal{U}_l) \leqslant R_l$. Finally, let us analyze $m_l(\mathcal{U}_l)$. For all $x \in X$, consider $h \in H$ with $B(x,l) \cap A_h \neq \varnothing$. Take $y \in B(x,l) \cap A_h$, ie $d(y,x) \leqslant l$ and $h \in S(y,l,l)$. By the definition of the edge-path length metric d, we know that there exists a sequence of vertices $y = y_0, y_1, \dots, y_k = x$ such that $y_i \in V$ for all $i = 0,1,\dots,k$, $d(y_i, y_{i+1}) = 1$ for all $i = 0,1,\dots,k-1$ and $k \leqslant l$. Now, by the hypothesis, we know

$$S(y,l,l) \subseteq S(y_1,l+1,l) \subseteq S(y_2,l+2,l) \subseteq \cdots \subseteq S(y_k,l+k,l) = S(x,k+l,l) \subseteq S(x,2l,l).$$

So $\{h \in H : B(x,l) \cap A_h \neq \varnothing\} \subseteq S(x,2l,l)$, which implies $m_l(\mathcal{U}_l) \leqslant \#S(x,2l,l) \leqslant n+1$. □

4 Normal cube path and normal distance

In the next two sections, we focus on CAT(0) cube complexes, and prove Theorem 1.1. We prove it by constructing a uniformly bounded cover with suitable properties. Such a construction relies deeply on the analysis of normal balls and spheres, which we give in this section.

Normal cube paths, which were introduced by Niblo and Reeves [19] play a key role in the construction of the cover. They determine a distance function on the vertices and the balls and spheres defined in terms of this distance are essential in our proof of Theorem 1.1.

Throughout this section we fix a CAT(0) cube complex X with a fixed vertex x_0. The 1–skeleton $X^{(1)}$ of X is a graph with vertex set $V = X^{(0)}$ and edge set E, which give us the edge metric d on V. This is the restriction of the ℓ^1 metric to the 0–skeleton.

4.1 Normal cube paths

Given a cube $C \in X$, we denote by $\mathrm{St}(C)$ the union of all cubes which contain C as a subcube.

Definition 4.1 [19] Let $\{C_i\}_{i=0}^n$ be a sequence of cubes such that each cube has dimension at least 1, and $C_{i-1} \cap C_i$ consists of a single point, denoted by v_i.

- Call $\{C_i\}_{i=0}^n$ a *cube path* if C_i is the (unique) cube of minimal dimension containing v_i and v_{i+1}, ie v_i and v_{i+1} are diagonally opposite vertices of C_i. Define v_0 to be the vertex of C_0 diagonally opposite to v_1, and v_{n+1} to be the vertex of C_n diagonally opposite to v_n. The so-defined vertices $\{v_i\}_{i=0}^{n+1}$ are called the vertices of the cube path, and we say the cube path is from v_0 to v_{n+1}.
- The *length* of a cube path is the number of the cubes in the sequence.
- A cube path is called *normal* if $C_i \cap \mathrm{St}(C_{i-1}) = v_i$.

Normal cube paths in CAT(0) cube complexes behave like geodesics in trees. More precisely, in [19], the existence and uniqueness of normal cube paths connecting any pair of vertices is established. See also [24].

Proposition 4.2 [19] *For any two vertices $x, y \in V$, there exists a unique normal cube path from x to y.* (Note that the order is important here since in general normal cube paths are not reversible).

Proposition 4.3 [19] *The intersection of a normal cube path and a hyperplane is connected. In other words, a normal cube path crosses a hyperplane at most once.*

Proposition 4.4 [19; 11] *Let $\{C_i\}_{i=0}^n$ and $\{D_j\}_{j=0}^m$ be two normal cube paths in X, and let $\{v_i\}_{i=0}^{n+1}$ and $\{w_j\}_{j=0}^{m+1}$ be the vertices of these normal cube paths. If $d(v_0, w_0) \leqslant 1$ and $d(v_{n+1}, w_{m+1}) \leqslant 1$, then for all k, we have $d(v_k, w_k) \leqslant 1$.*

We omit the proofs for the above three propositions; the readers can find them in the original paper. However, let us recall the construction of the normal cube path from x to y as follows: Consider all the hyperplanes separating x from y and adjacent to x. The key fact is that these hyperplanes all cross a unique cube adjacent to x lying in the interval from x to y. This cube is defined to be the first cube on the normal cube path; then one proceeds inductively to construct the required normal cube path.

We will also need the following lemma, abstracted from [19]. Recall that $h \pitchfork C$ for a hyperplane h and a cube C means the intersection of h and C is nonempty (see Section 2.3).

Lemma 4.5 *Let $\{C_i\}_{i=0}^n$ be the normal cube path and h be a hyperplane. If $h \pitchfork C_i$, then there exists a hyperplane k such that $k \pitchfork C_{i-1}$ and h does not intersect with k.*

Proof Otherwise, $\forall k \pitchfork C_{i-1}$, we have $h \pitchfork k$. Now by Lemma 2.15 in [19], we know that there exists a cube $C \in X$, such that all such $k \pitchfork C$ and $h \pitchfork C$, and C_{i-1} is a face of C. Moreover, C contains an edge e of C_i since $h \pitchfork C$. So $\mathrm{St}(C_{i-1}) \cap C_i$ contains e, which is a contradiction to the definition of normal cube path. □

Now for any two vertices of X, we consider all the hyperplanes separating them, with a partial order by inclusion. More explicitly, for any $x, y \in V$, let $H(x, y)$ be the set of hyperplanes separating x and y. For any $h \in H(x, y)$, let h^- be the halfspace containing x. Define $h \leqslant k$ if $h^- \subseteq k^-$. Note that the definition depends on the vertices we choose, and we may change them under some circumstances, but still write h^- for abbreviation. To avoid ambiguity, we point out the vertices if necessary. We write $h < k$ to mean a strict containment $h^- \subsetneq k^-$.

Lemma 4.6 *For any $h, k \in H(x, y)$, the hyperplanes h and k do not intersect if and only if $h \leqslant k$ or $k \leqslant h$.*

Proof We only need to show the necessity. Let $\{C_i\}_{i=0}^n$ be the normal cube path from x to y, and choose i and j so that $h \pitchfork C_i$ and $k \pitchfork C_j$. Since h and k do not intersect, $i \neq j$. Assume $i < j$. Obviously, $x \in h^- \cap k^-$ and $y \in h^+ \cap k^+$. Since $h \pitchfork C_i$ and $k \pitchfork C_j$, by Proposition 4.3 we have $v_{i+1} \in h^+ \cap k^-$. Since h does not intersect with k, we have $h^- \cap k^+ = \varnothing$, which implies $h^- \subseteq k^-$. □

Combining the above two lemmas, we have the following result on the existence of chains in $H(x, y)$.

Proposition 4.7 *Let $\{C_i\}_{i=0}^n$ be the normal cube path from x to y, and h be a hyperplane such that $h \pitchfork C_l$. Then there exists a chain of hyperplanes $h_0 < h_1 < \cdots < h_{l-1} < h_l = h$ such that $h_i \pitchfork C_i$.*

Proof By Lemma 4.5, there exists a hyperplane k such that $k \pitchfork C_{l-1}$ and h does not intersect with k. Define $h_{l-1} = k$. Inductively, we can define a sequence of hyperplanes as required. Then the conclusion follows by Lemma 4.6. □

Finally, we give a lemma used in the proof of the consistency part of our main theorem.

Lemma 4.8 *Let $x_0, x, y \in V$ with $[x_0, y] \subseteq [x_0, x]$, and let x' and y' be the nth vertex on the normal cube path from x_0 to x, and to y. If $x' \neq y'$, then $x' \notin [x_0, y]$.*

Proof Otherwise, $x' \in [x_0, y]$. By the construction of the normal cube path, we know x' is also the nth vertex on the normal cube path from x_0 to y, since $y \in [x_0, x]$. In other words, $x' = y'$, which is a contradiction to the assumption. □

4.2 Normal metric

We define a new metric on $V = X^{(0)}$ using normal cube paths [19; 24].

Definition 4.9 For any $x, y \in V$, define $d_{\text{nor}}(x, y)$ to be the length of the normal cube path from x to y. We call d_{nor} the *normal metric* on V.

One needs to verify that d_{nor} is indeed a metric. It is easy to see that $d_{\text{nor}}(x, y) \geqslant 0$, and $d_{\text{nor}}(x, y) = 0$ if and only if $x = y$. Note that the normal cube path from x to y is not the one from y to x in general, so the symmetric relation is not that obvious. In order to show the symmetric relation and the triangle inequality, we give the following characterization:

Lemma 4.10 *For $x, y \in V$, let $<$ be the relation defined as above. Then*

$$d_{\text{nor}}(x, y) = \sup\{m + 1 : h_0 < h_1 < \cdots < h_m,\ h_i \in H(x, y)\}.$$

Proof Suppose $\{C_i\}_{i=0}^n$ is the normal cube path from x to y, so $d_{\text{nor}}(x, y) = n + 1$. Denote the right-hand side of the equality in the lemma by n'. Now for any chain $h_0 < h_1 < \cdots < h_m$ in $H(x, y)$, by Proposition 4.3, h_i intersects with just one cube,

denoted by $C_{k(i)}$. Obviously, if $h, k \pitchfork C_i$, then $h \pitchfork k$. So $k(i) \neq k(j)$ if $i \neq j$, which implies $m \leqslant n$, so $n' \leqslant n$.

On the other hand, for any $h \pitchfork C_n$, by Proposition 4.7, we have a chain of hyperplanes $h_0 < h_1 < \cdots < h_{n-1} < h_n = h$ such that $h_i \pitchfork C_i$, which implies $n \leqslant n'$. □

Proposition 4.11 *d_{nor} is indeed a metric on V.*

Proof By Lemma 4.10 $d_{\mathrm{nor}}(x, y)$ is the maximal length of a chain in the poset $H(x, y)$. Since $H(y, x)$ is the opposite poset of $H(x, y)$ we deduce that $d_{\mathrm{nor}}(x, y) = d_{\mathrm{nor}}(y, x)$. For $x, y, z \in V$, $H(x, y) \triangle H(y, z) = H(x, z)$, where $\triangle$ is the symmetric difference operation. The inclusions of $H(x, y) \cap H(x, z)$ into $H(x, y)$ and $H(y, z) \cap H(x, z)$ into $H(y, z)$ are both order-preserving, and therefore, by Lemma 4.10, we have $d_{\mathrm{nor}}(x, z) \leqslant d_{\mathrm{nor}}(x, y) + d_{\mathrm{nor}}(y, z)$. □

4.3 Normal balls and normal spheres

Recall that for any two points x and y in $V = X^{(0)}$, the interval between them is

$$[x, y] = \{z \in V : d(x, y) = d(x, z) + d(z, y)\}.$$

In other words, $[x, y]$ is the set of vertices on the union of all the edge geodesics from x to y. A subset $Y \subseteq V$ is called *convex* if $[x, y] \subseteq Y$ for any $x, y \in Y$.

Now let $B(x, n)$ be the closed ball in the edge metric with centre $x \in V$ and radius n. Generally, $B(x, n)$ is not convex (for example, take $X = \mathbb{Z}^2$). However, as we will see, for the normal metric balls are convex. More precisely, we define the *normal ball* with centre $x \in V$ and radius n to be

$$B_{\mathrm{nor}}(x, n) = \{y \in V : d_{\mathrm{nor}}(x, y) \leqslant n\}$$

and the *normal sphere* with centre $x \in V$ and radius n to be

$$S_{\mathrm{nor}}(x, n) = \{y \in V : d_{\mathrm{nor}}(x, y) = n\}.$$

Lemma 4.12 *$B_{\mathrm{nor}}(x, n)$ is convex for all $x \in V$ and $n \in \mathbb{N}$.*

Proof Given $z, w \in B_{\mathrm{nor}}(x, n)$, and a geodesic γ from z to w, if $\gamma \not\subseteq B_{\mathrm{nor}}(x, n)$, we can assume u is the first vertex on γ which is not in $B_{\mathrm{nor}}(x, n)$, which implies $d_{\mathrm{nor}}(x, u) = n + 1$. Let z' be the vertex preceding u on γ, so $d_{\mathrm{nor}}(x, z') = n$ (since $d_{\mathrm{nor}}(z', u) = 1$). Since $d(z', u) = 1$, there exists a unique hyperplane h separating z' from u, so $H(x, u) = H(x, z') \sqcup \{h\}$. Now, according to Lemma 4.10, there exists a chain $h_0 < \cdots < h_{n-1} < h$ in $H(x, u)$ with $h_i \in H(x, z')$. Since every geodesic

intersects with any hyperplane at most once (see for example [27]), $w \in h^+$, which implies $h_0 < \cdots < h_{n-1} < h$ is also a chain in $H(x, w)$. This is a contradiction to $d_{\mathrm{nor}}(x, w) \leqslant n$, by Lemma 4.10. □

Since the intersection of two convex sets is still convex, we have the following corollary:

Corollary 4.13 *For any $x_0, x \in V$ and $n \in \mathbb{N}$, the set $[x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$ is convex.*

In this section, from now on, we fix a basepoint x_0 in X.

It is well known that for a convex subset Y in a CAT(0) cube complex and a point $v \notin Y$, there is a unique point in Y which is closest to v (see for example [9]). This statement is true both for the intrinsic CAT(0) metric on the cube complex and the edge metric on the vertex set, and we have a dual statement for the normal distance:

Proposition 4.14 *There exists a unique point $v \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$ such that*

$$[x_0, x] \cap B_{\mathrm{nor}}(x_0, n) \subseteq [x_0, v].$$

The point v is characterized by

$$d(x_0, v) = \max\{d(x_0, v') : v' \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)\}.$$

Furthermore, if $d_{\mathrm{nor}}(x_0, x) \geqslant n$, then $v \in [x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$, which implies that v is also the unique point in $[x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$ such that

$$d(x_0, v) = \max\{d(x_0, v') : v' \in [x_0, x] \cap S_{\mathrm{nor}}(x_0, n)\}.$$

Proof If there exist $z \neq w \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$ such that $d(x_0, z) = d(x_0, w)$ attains the maximum, consider the median $m = \mu(z, w, x)$. By Corollary 4.13, $m \in [z, w] \subseteq [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$, so $d(m, x_0) = d(z, x_0) = d(w, x_0)$. On the other hand, $m \in [z, x] \cap [w, x]$, so $m = z = w$, which is a contradiction.

By Corollary 4.13,

$$[x_0, v] \subseteq [x_0, x] \cap B_{\mathrm{nor}}(x_0, n).$$

Conversely, for any $u \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$, let $m = \mu(u, v, x) \in [u, v]$. By Corollary 4.13, $m \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$. Also, $m \in [v, x]$, so $d(m, x_0) \geqslant d(v, x_0)$, which implies $m = v$ by the choice of v, ie $\mu(u, v, x) = v$, so $v \in [u, x]$. Now, by Lemma 2.8, $u \in [x_0, v]$.

Now for x and n satisfying $d_{\mathrm{nor}}(x_0, x) \geqslant n$, if $v \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n-1)$, take a geodesic γ from v to x, and let $v = y_0, y_1, \ldots, y_k = x$ be the vertices on γ. Since $d_{\mathrm{nor}}(x_0, x) \geqslant n$, we have $x \neq v$, which implies $k > 0$. Now, for y_1, since $y_1 \in [v, x]$,

$d(x_0, v) < d(x_0, y_1)$. By the definition of v, we know $y_1 \notin [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$, so $y_1 \notin B_{\mathrm{nor}}(x_0, n)$. However, since $d(v, y_1) = d_{\mathrm{nor}}(v, y_1) = 1$, we have

$$d_{\mathrm{nor}}(x_0, y_1) \leqslant d_{\mathrm{nor}}(x_0, v) + d_{\mathrm{nor}}(v, y_1) \leqslant n,$$

which is a contradiction. □

To use the above proposition more flexibly, we give another characterization of v, which can also be viewed as an alternative definition of v. In the rest of this subsection, we fix $x \in V$ and $n \in \mathbb{N}$ with $d_{\mathrm{nor}}(x_0, x) \geqslant n$, where x_0 is the chosen basepoint.

Proposition 4.15 *Let v be the unique point in $[x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$ such that*

$$[x_0, x] \cap B_{\mathrm{nor}}(x_0, n) \subseteq [x_0, v];$$

then v is the nth vertex on the normal cube path from x_0 to x.

To prove this result, let us focus on subsets in $H(x_0, x)$. Recall that $H(x_0, x)$ is endowed with the relation $\leqslant$, as defined prior to Lemma 4.6.

Definition 4.16 A subset $A \subseteq H(x_0, x)$ is *closed* (under $<$) if $k \in A$ for all $h \in A$ and $k < h$.

Lemma 4.17 *Let $\widetilde{v}$ be the nth vertex on the normal cube path from x_0 to x; then $H(x_0, \widetilde{v})$ is maximal in the following sense: for any closed $A \subseteq H(x_0, x)$ which contains chains only with lengths at most n, $A \subseteq H(x_0, \widetilde{v})$.*

Proof We proceed by induction on n. Suppose that the lemma holds for $n-1$, and let v' be the $(n-1)^{\mathrm{st}}$ vertex on the normal cube path from x_0 to x. Given a closed $A \subseteq H(x_0, x)$ containing chains only with lengths at most n, and a maximal chain $h_0 < h_1 < \cdots < h_m$ in A, if $m \leqslant n-2$, then the closed set $\{h \in A : h \leqslant h_m\}$ contains chains only with lengths at most $n-1$; by induction, it is contained in $H(x_0, v') \subseteq H(x_0, \widetilde{v})$. Now, for $m = n-1$, similarly, $\{h \in A : h \leqslant h_{n-2}\} \subseteq H(x_0, v')$, which implies $h_i \pitchfork C_i$ for $i = 0, 1, \ldots, n-2$. So $h_{n-1} \pitchfork C_k$ for some $k \geqslant n-1$. If $k \neq n-1$, then by Proposition 4.7 and the closeness of A we get a chain in A with length greater than n, which is a contradiction. So $h_{n-1} \pitchfork C_{n-1}$, ie $h_{n-1} \in H(x_0, \widetilde{v})$. □

Proof of Proposition 4.15 Let $\widetilde{v}$ denote the nth vertex on the normal cube path from x_0 to x. By Proposition 4.14, $\widetilde{v} \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n) \subseteq [x_0, v]$, which implies $H(x_0, \widetilde{v}) \subseteq H(x_0, v)$. However, $H(x_0, v)$ is closed and contains chains only with lengths at most n according to Lemma 4.10, so $H(x_0, v) \subseteq H(x_0, \widetilde{v})$ by Lemma 4.17, which implies $H(x_0, v) = H(x_0, \widetilde{v})$. So $H(v, \widetilde{v}) = H(x_0, v) \triangle H(x_0, \widetilde{v}) = \varnothing$, which implies $v = \widetilde{v}$. □

Finally, we characterize those points in $[x_0, x]$ which lie in $[x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$. This will be used in the next subsection to decompose the intersection into a union of intervals.

Let C_{n-1} be the nth cube on the normal cube path from x_0 to x, and let v be the nth vertex on the cube path as above. Let H_n be the set of all hyperplanes intersecting with C_{n-1}.

Proposition 4.18 *For $w \in [x_0, x]$, the following are equivalent:*

(1) $w \in [x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$.

(2) *$\exists h \in H_n$ such that h crosses the last cube on the normal cube path from x_0 to w.*

(3) *$\exists h \in H_n$ such that h separates w from x_0 and $w \in [x_0, v]$.*

Proof (1) $\Longrightarrow$ (3) By Proposition 4.14, $w \in [x_0, v]$. Since $d_{\mathrm{nor}}(x_0, w) = n$, by Lemma 4.10, the maximum length of chains in $H(x_0, w)$ is n. Take such a chain $h_0 < h_1 < \cdots < h_{n-1}$ in $H(x_0, w) \subseteq H(x_0, v)$. Obviously, h_i intersects with different cubes, which implies $h_i \pitchfork C_i$. So $h_{n-1} \in H_n$, and it separates w from x_0.

(3) $\Longrightarrow$ (2) Since h separates w from x_0, h must cross some cube C on the normal cube path from x_0 to w. Since $h \in H_n$, we know there is a chain $h_0 < h_1 < \cdots < h_{n-1} = h$ in $H(x_0, v)$, which is also a chain in $H(x_0, w)$. So h cannot cross the first $n-1$ cubes of the normal cube path from x_0 to w. If h does not cross the last cube, then $d_{\mathrm{nor}}(x_0, w) > n$. However, $w \in [x_0, v]$ implies $H(x_0, w) \subseteq H(x_0, v)$, so by Lemma 4.10, $d_{\mathrm{nor}}(x_0, v) > n$, which is a contradiction.

(2) $\Longrightarrow$ (1) This is immediate, by Lemma 4.10. □

We have another description for H_n, which is implied by Proposition 4.7 directly.

Lemma 4.19 *For $h \in H(x_0, x)$, $h \in H_n$ if and only if the maximal length of chains in $\{k \in H(x_0, x) : k \leqslant h\}$ is n.*

4.4 Decomposition of $[x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$

We want to decompose the set $[x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$, so that we can proceed by the induction on dimension in the proof of Theorem 1.1.

Throughout this subsection, we fix $x \in V$ and $n \in \mathbb{N}$ with $d_{\mathrm{nor}}(x_0, x) \geqslant n$, and let v be as defined in Proposition 4.14. At the end of the preceding subsection, we have

defined H_n to be the set of all hyperplanes intersecting with C_{n-1}, where $\{C_i\}$ is the normal cube path from x_0 to x.

Now we decompose $[x_0, x] \cap S_{\mathrm{nor}}(x_0, n)$ into a union of intervals with dimensions lower than $[x_0, x]$, and the number of these intervals can be controlled by the dimension of $[x_0, x]$. This will make it possible to do induction on the dimension.

Definition 4.20 For $h \in H_n$, we define

$$F_h = \{w \in [x_0, x] \cap S_{\mathrm{nor}}(x_0, n) : h \text{ separates } w \text{ from } x_0\}.$$

By Proposition 4.18, we immediately obtain the following two lemmas:

Lemma 4.21 *We have*

$$\begin{aligned} F_h &= \{w \in [x_0, x] : h \text{ crosses the last cube on the normal cube path from } x_0 \text{ to } w\} \\ &= \{w \in [x_0, v] : h \text{ separates } w \text{ from } x_0\}. \end{aligned}$$

Lemma 4.22 $$[x_0, x] \cap S_{\mathrm{nor}}(x_0, n) = \bigcup_{h \in H_n} F_h.$$

By definition, we know

$$F_h = [x_0, x] \cap B_{\mathrm{nor}}(x_0, n) \cap \{v' : h \text{ separates } v' \text{ from } x_0\},$$

which implies that F_h is convex. Moreover, we will show that F_h is actually an interval.

Lemma 4.23 *Let $x_h \in F_h$ be the point minimizing $d(x_0, x_h)$. Then $F_h = [x_h, v]$.*

Proof Since F_h is convex and $x_h, v \in F_h$, so $[x_h, v] \subseteq F_h$. On the other hand, for all $z \in F_h$, let $m = \mu(x_0, z, x_h)$. So, $m \in F_h$ and $d(x_0, m) \leqslant d(x_0, x_h)$. By the choice of x_h, we know that $d(x_0, m) = d(x_0, x_h)$, which implies $m = x_h$, so $x_h \in [x_0, z]$. By Proposition 4.14, $x_h, z \in [x_0, v]$. Thus, by Lemma 2.8, $z \in [x_h, v]$. □

Proposition 4.24 $[x_0, x] \cap S_{\mathrm{nor}}(x_0, n) = \bigcup_{h \in H_n} [x_h, v]$ *and* $\dim[x_h, v] < \dim[x_0, x]$.

Proof We only need to show $\dim[x_h, v] < \dim[x_0, x]$. For any hyperplane k crossing $[x_h, v]$, by Proposition 4.18, $k \pitchfork h$. So $\dim[x_h, v] < \dim[x_0, x]$. □

Now we give another characterization for x_h, which is useful in the proof of the consistency condition of Theorem 1.1.

Lemma 4.25 *Let x_h be the closest point to x_0 on F_h; then x_h is the unique point in $B_{\mathrm{nor}}(x_0, n)$ such that h separates x_0 from x_h, and for any hyperplane $k \pitchfork h$, k does not separate x_h from x_0.*

Proof Since $x_h \in F_h$, we have $x_h \in [x_0, x] \cap B_{\mathrm{nor}}(x_0, n)$ and h separates x_0 from x_h. Now, for any hyperplane $k \pitchfork h$, if k separates x_h from x_0, we have $x_h \in h^+ \cap k^+$ and $x_0 \in h^- \cap k^-$. Choose $\widetilde{x}_h \in [x_0, x_h]$ such that $\widetilde{x}_h \in h^+ \cap k^-$. Since k does not separate $\widetilde{x}_h$ from x_0, we have $d(x_0, \widetilde{x}_h) < d(x_0, x_h)$. However, by Lemma 4.21, $\widetilde{x}_h \in F_h$, which is a contradiction.

It remains to show that x_h is the unique point satisfying these conditions. Otherwise, let $\hat{x}_h$ be another point satisfying the hypothesis in the lemma and $\hat{x}_h \neq x_h$. Let k be a hyperplane separating $\hat{x}_h$ from x_h, and assume $x_h \in k^-$. Obviously, $k \neq h$. If $k \pitchfork h$, by hypothesis, k does not separate x_h from x_0, as well as $\hat{x}_h$ from x_0, which is a contradiction since k separates $\hat{x}_h$ from x_h. So k does not cross h, which implies $h^- \subsetneq k^-$ by Lemma 4.6. However, $\hat{x}_h \in k^+$, so by Lemma 4.10, $d_{\mathrm{nor}}(x_0, x_h) < d_{\mathrm{nor}}(x_0, \hat{x}_h)$. This is a contradiction since $d_{\mathrm{nor}}(x_0, x_h) \geqslant n$ as h separates x_0 from x_h. □

4.5 Špakula and Wright's construction

We conclude this section with a recent application of normal cube paths, which were invoked by Špakula and Wright [29] in order to provide a new proof that finite-dimensional CAT(0) cube complexes have Yu's Property A. The key to their proof was the construction of a family of maps h_l with the property that, for any interval and any neighbourhood of an endpoint of the interval, the maps push that neighbourhood into the interval itself. These maps were defined in terms of the normal cube paths as follows:

Definition 4.26 (the h maps) Given $l \in \mathbb{N}$, we define $h_l \colon X \to X$ as follows. For $x \in X$, let $h_l(x)$ be the $3l^{\text{th}}$ vertex on the normal cube path from x to x_0 if $d_{\mathrm{nor}}(x, x_0) \geqslant 3l$, and let it be x_0 if $d_{\mathrm{nor}}(x, x_0) < 3l$.

Lemma 4.27 [29] *Let h_l be defined as above and $y \in B(x, 3l)$. Then $h_l(y) \in [x_0, x]$.*

Proof We only need to show that every halfspace containing x and x_0 contains also $z = h_l(y)$. For any hyperplane h such that one of the associated halfspaces, say h^+, contains x and x_0, either $y \in h^+$ or $y \in h^-$. In the former case, $z \in h^+$, so we only need to check the case that h separates x and x_0 from y.

Denote by $C_0, C_1, \ldots, C_m$ the normal cube path from y to x_0, and denote by $y = v_0, v_1, \ldots, v_m = x_0$ the vertices on this cube path. We shall argue that any hyperplane separating y from x and x_0 is "used" within the first $d(x, y)$ steps on the cube path. Suppose that the cube C_i does not cross any hyperplane h with h separating y from x and x_0. Hence every hyperplane $k \pitchfork C_i$ separates y and x from x_0 and v_{i+1}. If there was a hyperplane l separating y from x and x_0 before C_i, then necessarily l separates y and v_{i+1} from x and x_0, hence l crosses all the hyperplanes k crossing C_i. This contradicts the maximality of this step on the normal cube path. Thus, there is no such l, and so all the hyperplanes h separating y from x and x_0 must be crossed within the first $d(x, y)$ steps.

Since z is the $3l^{\text{th}}$ vertex on the cube path and $d(x, y) \leqslant 3l$, all the hyperplanes h separating y from x and x_0 must have been crossed before z. Thus, any such h actually also separates y from x, x_0 and z. □

We will use the remarkable properties of the h maps to construct the S sets defined in our characterization of finite asymptotic dimension in the next section.

5 Finite-dimensional CAT(0) cube complexes

Throughout this section, we fix a CAT(0) cube complex X of finite dimension η and equipped with a basepoint $x_0 \in X$. We will make use of the characterization obtained in Corollary 3.3 in order to prove Theorem 1.1.

5.1 Constructing the sets $S(x, k, l)$

By Corollary 3.3, in order to prove X has finite asymptotic dimension, we need to find a constant $N \in \mathbb{N}$ such that for all $l \in \mathbb{N}$, $k = 1, 2, \ldots, 3l$ and $x \in X$, we can assign a subset $S(x, k, l) \subseteq X$ satisfying:

(i) $\forall l \in \mathbb{N}$, $\exists R_l > 0$ such that $S(x, k, l) \subseteq B(x, R_l)$ for $k = 1, \ldots, 3l$.

(ii) $\forall l \in \mathbb{N}$, $\forall k, k'$ with $1 \leqslant k \leqslant k' \leqslant 3l$, $\forall x \in X$, $S(x, k, l) \subseteq S(x, k', l)$.

(iii) $\forall x, y \in X$ with $d(x, y) = 1$, $S(y, k, l) \subseteq S(x, k+1, l)$ for $k = 1, 2, \ldots, 3l-1$.

(iv) $\forall l \in \mathbb{N}$, $\#S(x, 2l, l) \leqslant N + 1$.

Now, for $l \in \mathbb{N}$, $k = 1, 2, \ldots, 3l$ and $x \in X$, we define

$$\widetilde{S}(x, k, l) = h_l(B(x, k)).$$

It is easy to show that $\{\widetilde{S}(x, k, l)\}$ satisfies (i)–(iii), but it does not satisfy (iv) above, so we need some modification. Intuitively, we construct $S(x, k, l)$ as a uniformly separated net in $\widetilde{S}(x, k, l)$. To be more precise, we require the following lemma:

Lemma 5.1 *There exist two constants N and K, only depending on the dimension η, such that for all $l \in \mathbb{N}$ and $x \in V$, there are subsets $C_x \subseteq [x_0, x]$ and maps $p_x\colon [x_0, x] \to \mathcal{P}(C_x)$, where $\mathcal{P}(C_x)$ denotes the power set of C_x, satisfying:*

- *If $d(x, y) = 1$ and $y \in [x_0, x]$, then $C_x \cap [x_0, y] = C_y$, and $p_x|_{[x_0,y]} = p_y$.*
- *For $z \in [x_0, x]$ and $w \in p_x(z)$, we have $d(z, w) \leqslant Kl$.*
- *$\forall z \in [x_0, x]$, $\#(B(z, Ml) \cap C_x) \leqslant N$, where $M = 3\eta + 3 + K$.*

We postpone the proof of the above lemma and first show how to use it to construct $S(x, k, l)$ (and, hence, to conclude the proof of Theorem 1.1).

Proof of Theorem 1.1 Let N and K be the constants in Lemma 5.1. For all $l \in \mathbb{N}$, $k = 1, 2, \ldots, 3l$ and $x \in X$, let $\widetilde{S}(x, k, l) = h_l(B(x, k))$ be as above, and by Lemma 4.27, we know $\widetilde{S}(x, k, l) \subseteq [x_0, x]$. Recall that $p_x\colon [x_0, x] \to \mathcal{P}(C_x)$ assigns a subset of C_x to each point y of the interval $[x_0, x]$. Now given a *subset* Y of the interval, $p_x(Y)$ is a collection of subsets of C_x and we denote by $\bigcup p_x(Y)$ the union of these subsets $p_x(y)$ for $y \in Y$.

Now we define

$$S(x, k, l) = \bigcup p_x(\widetilde{S}(x, k, l)).$$

Now the only thing left to complete the proof is to verify the conditions in Corollary 3.3:

(i) By the definition of h_l, we know $d(y, h_l(y)) \leqslant 3\eta l$ for all $y \in B(x, k)$. So $d(z, x) \leqslant (3\eta + 3)l$ for any $z \in \widetilde{S}(x, k, l)$. For such z and any $w \in p_x(z)$, by Lemma 5.1, we know $d(z, w) \leqslant Kl$, which implies

$$S(x, k, l) \subseteq B(x, (3\eta + 3 + K)l) = B(x, Ml).$$

(ii) For all $l \in \mathbb{N}$, k and k' with $1 \leqslant k \leqslant k' \leqslant 3l$, and $x \in X$, we have $\widetilde{S}(x, k, l) \subseteq \widetilde{S}(x, k', l)$. Now, immediately from the definition, $S(x, k, l) \subseteq S(x, k', l)$.

(iii) For all $x, y \in X$ with $d(x, y) = 1$, by Lemma 2.9, $y \in [x_0, x]$ or $x \in [x_0, y]$. Assume the former. Let $k = 1, 2, \ldots, 3l-1$. Obviously, $\widetilde{S}(y, k, l) \subseteq \widetilde{S}(x, k+1, l)$, so we have

$$\begin{aligned} S(y, k, l) &= \bigcup p_y(\widetilde{S}(y, k, l)) = \bigcup p_x|_{[x_0,y]}(\widetilde{S}(y, k, l)) = \bigcup p_x(\widetilde{S}(y, k, l)) \\ &\subseteq \bigcup p_x(\widetilde{S}(x, k+1, l)) = S(x, k+1, l). \end{aligned}$$

Here we use the first part of Lemma 5.1 in the second equation. On the other hand, $\widetilde{S}(x, k, l) \subseteq \widetilde{S}(y, k+1, l)$, so we have

$$\begin{aligned} S(x,k,l) &= \bigcup p_x(\widetilde{S}(x,k,l)) \subseteq \bigcup p_x(\widetilde{S}(y,k+1,l)) \\ &= \bigcup p_x|_{[x_0,y]}(\widetilde{S}(y,k+1,l)) \\ &= \bigcup p_y(\widetilde{S}(y,k+1,l)) = S(y,k+1,l). \end{aligned}$$

Here we use the first part of Lemma 5.1 in the fourth equality.

(iv) By (i), we know that $S(x,k,l) \subseteq B(x,Ml)$ for all $k = 1,2,\dots,3l$. Hence, by definition, $S(x,k,l) \subseteq B(x,Ml) \cap C_x$. Now, by the third part of Lemma 5.1, we have $\#S(x,k,l) \leqslant N$. □

The last thing is to prove Lemma 5.1. We use the analysis in Section 4 to construct C_x and p_x inductively. Recall that in Section 4 (Proposition 4.24), for any $l, n \in \mathbb{N}$ and any $x \in X$, we have

$$[x_0,x] \cap S_{\mathrm{nor}}(x_0,nl) = \bigcup_{h \in H_{nl}} [x_h,v],$$

with $\#H_{nl} \leqslant \eta$ and $\dim[x_h,v] < \dim[x_0,x]$. In order to carry out induction on the dimension of $[x_0,x]$, we require a stronger version of Lemma 5.1, which is more flexible on the choice of endpoints of intervals. More explicitly, we have:

Lemma 5.2 *There exist two constants N and K, only depending on the dimension η, such that $\forall l \in \mathbb{N}$, $\forall \overline{x}, x \in V$, $\exists C_{\overline{x},x} \subseteq [\overline{x},x]$ and a map $p_{\overline{x},x}\colon [\overline{x},x] \to \mathcal{P}(C_{\overline{x},x})$ satisfying:*

- *If $d(x,y) = 1$ and $y \in [\overline{x},x]$, then $C_{\overline{x},x} \cap [\overline{x},y] = C_{\overline{x},y}$, and $p_{\overline{x},x}|_{[\overline{x},y]} = p_{\overline{x},y}$.*
- *For $z \in [\overline{x},x]$ and $w \in p_{\overline{x},x}(z)$, we have $d(z,w) \leqslant Kl$.*
- *$\forall z \in [\overline{x},x]$, $\#(B(z,Ml) \cap C_{\overline{x},x}) \leqslant N$, where $M = 3\eta + 3 + K$.*

It is obvious that Lemma 5.1 is implied by Lemma 5.2 (one just needs to take $\overline{x} = x_0$). Now we prove Lemma 5.2.

Proof of Lemma 5.2 Fix an $l \in \mathbb{N}$. We will carry out induction on $\dim[\overline{x},x]$.

Given any $\overline{x}, x \in V$ with $\dim[\overline{x},x] = 1$, we define

$$C_{\overline{x},x} = \{y \in [\overline{x},x] : d_{\mathrm{nor}}(\overline{x},y) \in l\mathbb{N}\},$$

where $l\mathbb{N} = \{0,l,2l,3l,\dots\}$. Since $\dim[\overline{x},x] = 1$, $[\overline{x},x]$ is indeed isometric to an interval in $\mathbb{R}$. We define $p_{\overline{x},x}\colon [\overline{x},x] \to \mathcal{P}(C_{\overline{x},x})$ as follows: for any $y \in [\overline{x},x]$, $p_{\overline{x},x}(y)$ consists of a single point which is at distance $l\lfloor d_{\mathrm{nor}}(\overline{x},y)/l \rfloor$ from $\overline{x}$ in $[\overline{x},y]$, where $\lfloor\cdot\rfloor$ is the function of taking integer part. Now it is obvious that:

- If $d(x,y)=1$ and $y\in[\overline{x},x]$, then $C_{\overline{x},x}\cap[\overline{x},y]=C_{\overline{x},y}$ and $p_{\overline{x},x}|_{[\overline{x},y]}=p_{\overline{x},y}$.
- For $z\in[\overline{x},x]$ and $w\in p_{\overline{x},x}(z)$, we have $d(z,w)\leqslant l$.
- $\forall z\in[\overline{x},x]$, $\#(B(z,Ml)\cap C_{\overline{x},x})\leqslant 3M$.

Suppose for any $\overline{x},x\in V$ with $\dim[\overline{x},x]\leqslant\eta-1$, we have defined $C_{\overline{x},x}\subseteq[\overline{x},x]$ and a map $p_{\overline{x},x}\colon[\overline{x},x]\to\mathcal{P}(C_{\overline{x},x})$ satisfying:

- If $d(x,y)=1$ and $y\in[\overline{x},x]$, then $C_{\overline{x},x}\cap[\overline{x},y]=C_{\overline{x},y}$ and $p_{\overline{x},x}|_{[\overline{x},y]}=p_{\overline{x},y}$.
- For $z\in[\overline{x},x]$ and $w\in p_{\overline{x},x}(z)$, we have $d(z,w)\leqslant\frac{1}{2}(\eta-1)\eta l$.
- $\forall z\in[\overline{x},x]$, $\#(B(z,Ml)\cap C_{\overline{x},x})\leqslant(3M)^{\eta-1}(\eta-1)!$.

Now we focus on $\overline{x},x\in V$ with $\dim[\overline{x},x]=\eta$. For any $n\in\mathbb{N}$ with $nl\leqslant d_{\mathrm{nor}}(\overline{x},x)$, by Proposition 4.24,

$$[\overline{x},x]\cap S_{\mathrm{nor}}(\overline{x},nl)=\bigcup_{h\in H^x_{nl}}F^x_h=\bigcup_{h\in H^x_{nl}}[x_h,v^x_{nl}],$$

where v^x_{nl} is the farthest point from $\overline{x}$ in $[\overline{x},x]\cap S_{\mathrm{nor}}(\overline{x},nl)$, H^x_{nl} is the set of hyperplanes crossing the nl^{th} cube of the normal cube path from $\overline{x}$ to x, and we also have $\dim[x_h,v^x_{nl}]<\dim[\overline{x},x]$. By induction, $C_{x_h,v^x_{nl}}$ and $p_{x_h,v^x_{nl}}$ have already been defined. Now we define

$$C^n_{\overline{x},x}=\bigcup_{h\in H^x_{nl}}C_{x_h,v^x_{nl}}$$

and

$$C_{\overline{x},x}=\bigcup_{n=0}^{\lfloor d_{\mathrm{nor}}(\overline{x},x)/l\rfloor}C^n_{\overline{x},x}.$$

For any $z\in[\overline{x},x]$, let $\widetilde{z}$ be the nl^{th} vertex on the normal cube path from $\overline{x}$ to z, where $n=\lfloor d_{\mathrm{nor}}(\overline{x},z)/l\rfloor$, so $d_{\mathrm{nor}}(\widetilde{z},z)\leqslant l$, which implies $d(\widetilde{z},z)\leqslant\eta l$ and

$$\widetilde{z}\in[\overline{x},x]\cap S_{\mathrm{nor}}(\overline{x},nl)=\bigcup_{h\in H^x_{nl}}[x_h,v^x_{nl}].$$

Now we define

$$p_{\overline{x},x}(z)=\bigcup\{p_{x_h,v^x_{nl}}(\widetilde{z}):h\in H^x_{nl}\text{ and }\widetilde{z}\in[x_h,v^x_{nl}]\},$$

and we need to verify the requirements hold for $C_{\overline{x},x}$ and $p_{\overline{x},x}$.

First, suppose $d(x,y)=1$ and $y\in[\overline{x},x]$, and let h' be the hyperplane separating x from y. Given $n\in\mathbb{N}$ such that $[\overline{x},y]\cap S_{\mathrm{nor}}(\overline{x},nl)\neq\varnothing$, by Proposition 4.15, v^x_{nl} is the nl^{th} vertex on the normal cube path from $\overline{x}$ to x, and v^y_{nl} is the nl^{th} vertex on the

normal cube path from $\overline{x}$ to y. Due to the fellow-traveller property, Proposition 4.4, $d(v_{nl}^x, v_{nl}^y) \leqslant 1$. By Proposition 4.14, we have

$$v_{nl}^y \in S_{\mathrm{nor}}(\overline{x}, nl) \cap [\overline{x}, y] \subseteq S_{\mathrm{nor}}(\overline{x}, nl) \cap [\overline{x}, x] \subseteq [\overline{x}, v_{nl}^x].$$

Recall that $H(z, w)$ denotes the set of all hyperplanes separating z from w. Obviously,

$$H(\overline{x}, x) = H(\overline{x}, y) \cup \{h'\},$$

which implies $H_{nl}^y \subseteq H_{nl}^x \subseteq H_{nl}^y \cup \{h'\}$, by Lemmas 4.10 and 4.19.

If $h' \in H_{nl}^x$ then $F_{h'}^x \cap [\overline{x}, y] = \varnothing$ by Proposition 4.18. On the other hand, for all $h \in H_{nl}^y$, by Lemma 4.25, y_h is the unique point in $B_{\mathrm{nor}}(\overline{x}, nl)$ such that h separates $\overline{x}$ from y_h, and for any hyperplane $k \pitchfork h$, k does not separate y from $\overline{x}$. This implies $y_h = x_h$ since $H_{nl}^y \subseteq H_{nl}^x$, so we can do induction for the new "basepoint" $y_h = x_h$ and v_{nl}^y, v_{nl}^x, since $d(v_{nl}^x, v_{nl}^y) \leqslant 1$ and $v_{nl}^y \in [x_h, v_{nl}^x]$. This implies

$$C_{x_h, v_{nl}^x} \cap [x_h, v_{nl}^y] = C_{x_h, v_{nl}^y}.$$

Since $C_{x_h, v_{nl}^x} \subseteq [x_h, v_{nl}^x]$, we have

$$C_{x_h, v_{nl}^x} \cap [\overline{x}, y] = C_{x_h, v_{nl}^x} \cap [x_h, v_{nl}^x] \cap [\overline{x}, y] = C_{x_h, v_{nl}^x} \cap [x_h, \mu(x_h, v_{nl}^x, y)].$$

Claim $$\mu(x_h, v_{nl}^x, y) = v_{nl}^y.$$

Indeed, if $v_{nl}^x = v_{nl}^y$, then it holds naturally. If $v_{nl}^x \neq v_{nl}^y$, then by Lemma 4.8, $v_{nl}^x \notin [\overline{x}, y]$. Since $d(v_{nl}^x, v_{nl}^y) = 1$, so $v_{nl}^x \in [v_{nl}^y, y]$ or $v_{nl}^y \in [v_{nl}^x, y]$. But the former cannot hold since $[v_{nl}^y, y] \subseteq [\overline{x}, y]$, so $v_{nl}^y \in [v_{nl}^x, y]$, which implies

$$v_{nl}^y \in [x_h, y] \cap [x_h, v_{nl}^x] \cap [v_{nl}^x, y],$$

ie $v_{nl}^y = \mu(x_h, v_{nl}^x, y)$.

By the claim,

$$C_{x_h, v_{nl}^x} \cap [\overline{x}, y] = C_{x_h, v_{nl}^x} \cap [x_h, v_{nl}^y] = C_{x_h, v_{nl}^y}.$$

Now, for the above n, we have

$$\begin{aligned} C_{\overline{x},x}^n \cap [\overline{x}, y] &= \bigcup_{h \in H_{nl}^x} (C_{x_h, v_{nl}^x} \cap [\overline{x}, y]) = \bigcup_{h \in H_{nl}^y} (C_{x_h, v_{nl}^x} \cap [\overline{x}, y]) \\ &= \bigcup_{h \in H_{nl}^y} C_{x_h, v_{nl}^y} = \bigcup_{h \in H_{nl}^y} C_{y_h, v_{nl}^y} = C_{\overline{x},y}^n. \end{aligned}$$

Since $C_{\overline{x},x}^n \subseteq [\overline{x}, x] \cap S_{\mathrm{nor}}(\overline{x}, nl)$, we have

$$C_{\bar{x},x}\cap[\bar{x},y]=\bigcup_{n=0}^{\lfloor d_{\mathrm{nor}}(\bar{x},x)/l\rfloor} C^n_{\bar{x},x}\cap[\bar{x},y]=\bigcup_{n\in J_x} C^n_{\bar{x},x}\cap[\bar{x},y]$$
$$=\bigcup_{n\in J_y} C^n_{\bar{x},x}\cap[\bar{x},y]=\bigcup_{n\in J_y} C^n_{\bar{x},y}=C_{\bar{x},y},$$

where $J_x=\{n:[\bar{x},x]\cap S_{\mathrm{nor}}(\bar{x},nl)\neq\varnothing\}$ and $J_y=\{n:[\bar{x},y]\cap S_{\mathrm{nor}}(\bar{x},nl)\neq\varnothing\}$. For all $z\in[\bar{x},y]$, one needs to show that $p_{\bar{x},x}(z)=p_{\bar{x},y}(z)$. Let $\widetilde{z}$ be the nl^{th} vertex on the normal cube path from $\bar{x}$ to z, where $n=\lfloor d_{\mathrm{nor}}(\bar{x},z)/l\rfloor$. By the analysis above, we know

$$d(v^x_{nl},v^y_{nl})\leqslant 1,\quad v^y_{nl}\in[v^x_{nl},y],\quad x_h=y_h,\quad H^y_{nl}\subseteq H^x_{nl}\subseteq H^y_{nl}\cup\{h'\}.$$

For $h\in H^x_{nl}$ with $\widetilde{z}\in[x_h,v^x_{nl}]$, we have $h\in H^y_{nl}$, ie $h\neq h'$ since $\widetilde{z}\in[\bar{x},y]$. Now, for such h,

$$\widetilde{z}\in[\bar{x},y]\cap[x_h,v^x_{nl}]=[x_h,v^y_{nl}]=[y_h,v^y_{nl}],$$

where the first equality comes from the claim above. Inductively, we know, for such h,

$$p_{x_h,v^x_{nl}}(\widetilde{z})=p_{y_h,v^y_{nl}}(\widetilde{z}).$$

Now, by definition,

$$\begin{aligned}p_{\bar{x},x}(z)&=\bigcup\{p_{x_h,v^x_{nl}}(\widetilde{z}):h\in H^x_{nl}\text{ and }\widetilde{z}\in[x_h,v^x_{nl}]\}\\&=\bigcup\{p_{x_h,v^x_{nl}}(\widetilde{z}):h\in H^y_{nl}\text{ and }\widetilde{z}\in[x_h,v^x_{nl}]\}\\&=\bigcup\{p_{y_h,v^y_{nl}}(\widetilde{z}):h\in H^y_{nl}\text{ and }\widetilde{z}\in[y_h,v^y_{nl}]\}\\&=p_{\bar{x},y}(z).\end{aligned}$$

Second, for any $z\in[\bar{x},x]$ and $w\in p_{\bar{x},x}(z)$, assume that $w\in p_{x_h,v^x_{nl}}(\widetilde{z})$ for some $h\in H^x_{nl}$ and $\widetilde{z}\in[x_h,v^x_{nl}]$ as in the definition. By induction, we know $d(\widetilde{z},w)\leqslant\frac{1}{2}(\eta-1)\eta l$ since $\dim[x_h,v^x_{nl}]\leqslant\eta-1$. So

$$d(z,w)\leqslant d(z,\widetilde{z})+d(\widetilde{z},w)\leqslant\eta l+\tfrac{1}{2}(\eta-1)\eta l=\tfrac{1}{2}\eta(\eta+1)l.$$

Third, for any $z\in[\bar{x},x]$, consider $B(z,Ml)\cap C_{\bar{x},x}$. Suppose $n\in\mathbb{N}$ satisfying $B(z,Ml)\cap C_{\bar{x},x}\cap S_{\mathrm{nor}}(\bar{x},nl)\neq\varnothing$, so $B(z,Ml)\cap\big(\bigcup_{h\in H^x_{nl}}[x_h,v^x_{nl}]\big)\neq\varnothing$, which means there exists some $h\in H^x_{nl}$ such that $B(z,Ml)\cap[x_h,v^x_{nl}]\neq\varnothing$. For such n and h, let $z'=\mu(z,x_h,v^x_{nl})\in[x_h,v^x_{nl}]$. Obviously,

$$B(z,Ml)\cap[x_h,v^x_{nl}]\subseteq B(z',Ml)\cap[x_h,v^x_{nl}].$$

By induction, we have

$$\#(B(z,Ml)\cap C_{x_h,v^x_{nl}})\leqslant\#(B(z',Ml)\cap C_{x_h,v^x_{nl}})\leqslant(3M)^{\eta-1}(\eta-1)!.$$

Now, for the above z, there exist at most $3M$ values of n such that

$$B(z, Ml) \cap [\overline{x}, x] \cap S_{\mathrm{nor}}(\overline{x}, nl) \neq \varnothing;$$

and for such n, since $\#H^x_{nl} \leqslant \eta$, there exist at most η hyperplanes h such that $B(z, Ml) \cap [\overline{x}, x] \cap [x_h, v^x_{nl}] \neq \varnothing$. So we have

$$\begin{aligned}
\#(B(z, Ml) \cap C_{\overline{x},x}) &\leqslant \sum_{\substack{n \\ B(z,Ml)\cap[\overline{x},x]\cap S_{\mathrm{nor}}(\overline{x},nl)\neq\varnothing}} \#(B(z, Ml) \cap C^n_{\overline{x},x}) \\
&\leqslant \sum_{n \text{ as above}} \sum_{h \in H^x_{nl}} \#(B(z, Ml) \cap C_{x_h, v^x_{nl}}) \\
&\leqslant 3M \cdot \eta \cdot (3M)^{\eta-1}(\eta-1)! = (3M)^\eta \eta!.
\end{aligned}$$

Now we take $K = \frac{1}{2}(\eta-1)\eta$ and $N = (3M)^\eta \eta! = (3K + 9\eta + 9)^\eta \eta!$; then the lemma holds for these constants. □

6 Coarse median spaces

In this section, we discuss the coarse median case, and prove Theorem 1.2. We fix a coarse median space X with geodesic metric ρ and coarse median μ with parameters K and H and finite rank η. The definitions and notations are the same as in Section 2.4. According to Remark 2.11, we also assume that the coarse median μ satisfies (M1) and (M2). We recall:

Theorem 6.1 [29] *Any geodesic uniformly locally finite coarse median space of finite rank and at most exponential growth has Property A.*

Our result, Theorem 1.2, says that any coarse median space as above has subexponential asymptotic dimension growth. Thus, combining with Ozawa's result [23], our theorem yields a strengthening of Theorem 6.1.

To prove Theorem 1.2, we use several notations and lemmas from [29]. We use the notation $x \sim_s y$ for $\rho(x, y) \leqslant s$. Given $r > 0$ and $a, b \in X$, the coarse interval $[a, b]_r$ is defined to be

$$[a, b]_r = \{z \in X : \mu(a, b, z) \sim_r z\}.$$

By a result of Bowditch [8], there exists a constant $\lambda > 0$, depending only on the parameters K and H, such that $\mu(x, y, z) \in [x, y]_\lambda$ for all $x, y, z \in X$.

Also recall that the median axiom (M3) holds in the coarse median case up to a constant $\gamma \geqslant 0$ depending only on the parameters K and H: for all $x, y, z, u, v \in X$, we have

$$\mu(\mu(x, y, z), u, v) \sim_\gamma \mu(\mu(x, u, v), \mu(y, u, v), z).$$

Actually we can take $\gamma = 3K(3K+2)H(5) + (3K+2)H(0)$.

Given $r, t, \kappa \geqslant 0$, denote

$$\begin{aligned} L_1(r) &= (K+1)r + K\lambda + \gamma + 2H(0), \\ L_2(r,\kappa) &= (K+1)r + \kappa + H(0), \\ L_3(r,t) &= 3^\eta K^\eta rt + r. \end{aligned}$$

We need the following lemmas from [29]:

Lemma 6.2 [29] *Let X be a coarse median space, $r \geqslant 0$, and let $a, b \in X$ and $x \in [a,b]_\lambda$. Then $[a,x]_r \subset [a,b]_{L_1(r)}$.*

Lemma 6.3 [29] *Let X be a geodesic coarse median space of rank at most η. For every $\kappa > 0$ and $t > 0$, there exists $r_t > 0$, such that for all $r \geqslant r_t$ and $a, b \in X$, there exists $h \in [a,b]_{L_1(r)}$ such that*

- *$\rho(a,h) \leqslant L_3(r,t)$, and*
- *$B(a, rt) \cap [a,b]_\kappa \subset [a,h]_{L_2(r,\kappa)}$.*

Lemma 6.4 [29] *Let X be a coarse median space. Fix $\kappa > 0$. There exist constants $\alpha, \beta \geqslant 0$, depending only on the parameters of the coarse median structure and κ, such that the following holds: Let $a, b, h, m \in X$ and $r \geqslant 0$ satisfy $m \in [a,h]_{L_2(r,\kappa)}$ and $h \in [a,b]_{L_1(r)}$. Then $p = \mu(m,b,h)$ satisfies $\rho(h,p) \leqslant \alpha r + \beta$.*

Proof of Theorem 1.2 The proof is based on the construction used in [29] to prove Property A, and for the reader's convenience, we give a sketch of their proof. In fact we will verify the stronger conditions on the S sets required to apply Theorem 3.1. Fix a basepoint $x_0 \in X$, and let α and β be the constants from Lemma 6.4. First apply Lemma 6.3 for $\kappa = \lambda$ and all $t \in \mathbb{N}$ to obtain a sequence $r_t \in \mathbb{N}$ such that the conclusion of the lemma holds. Furthermore, we can choose the r_t inductively to arrange the sequence $t \mapsto l_t = (tr_t - H(0))/(3K)$ for $t \in \mathbb{N}$ is increasing.

Now fix $x \in X$, $t \in \mathbb{N}$ and $k \in \{1, 2, \ldots, 3l_t\}$. For any $y \in B(x,k)$, Lemma 6.3 applied for $a = y$, $b = x_0$ and $r = r_t$ produces a point $h_y \in [y, x_0]_{L_1(r_t)}$. We define

$$S(x,k,l_t) = \{h_y \in X : y \in B(x,k)\}.$$

We need to verify these sets satisfy condition (2) in the statement of Theorem 3.1, ie we need to show there exists a subexponential function $f\colon \mathbb{R} \to \mathbb{R}$, satisfying:

(i) $\forall t \in \mathbb{N}$, $\exists R_t > 0$ such that $S(x,k,l_t) \subseteq B(x, R_t)$ for all $k = 1, \ldots, 3l_t$.

(ii) $\forall t \in \mathbb{N}$, $\forall k, k'$ with $1 \leqslant k \leqslant k' \leqslant 3l_t$, $\forall x \in X$, we have $S(x,k,l_t) \subseteq S(x,k',l_t)$.

(iii) $\forall x, y \in X$ with $\rho(x,y) \leqslant l_t$, we have:
- $S(x,k-\rho(x,y),l_t) \subset S(x,k,l_t) \cap S(y,k,l_t)$, for $k = \rho(x,y)+1, \ldots, 3l_t$.
- $S(x,k+\rho(x,y),l_t) \supset S(x,k,l_t) \cap S(y,k,l_t)$, for $k = 1, \ldots, 3l_t - \rho(x,y)$.

(iv) $\forall t \in \mathbb{N}$, $\forall k = 1,2,\ldots,3l_t$, $\forall x \in X$, we have $\#S(x,k,l_t) \leqslant f(l_t)$.

By the construction, (ii)–(iii) hold naturally. For (i), by Lemma 6.3, we know

$$S(x,k,l_t) \subseteq B(x, l_t + L_3(r_t,t)).$$

The only thing left is to find a subexponential function f such that condition (iv) holds. The following argument follows totally from the proof in [29], and we omit some calculation. The readers can turn to their original paper for more details.

Take $y \in B(x,k)$, with the notation as above. Let $m_y = \mu(x,y,x_0)$. Then, by Lemma 6.3, one can deduce that $m_y \in [y,h_y]_{L_2(r_t,\lambda)}$. Now, since $h_y \in [y,x_0]_{L_1(r_t)}$, Lemma 6.4 implies the point $p_y = \mu(m_y,x_0,h_y) \in [m_y,x_0]_\lambda$ satisfies $\rho(h_y,p_y) \leqslant \alpha r_t + \beta$. As $m_y = \mu(x,y,x_0) \in [x,x_0]_\lambda$, Lemma 6.2 now implies $p_y \in [x,x_0]_{L_1(\lambda)}$. Consequently, we have $\rho(x,p_y) \leqslant 3l_t + 3^\eta K^\eta t r_t + r_t + \alpha r_t + \beta$, which depends linearly on l_t. Now, by Proposition 9.8 in [8], the number of possible points p_y is bounded by $P(l_t)$ for some polynomial P depending only on H, K, η and the uniform local finiteness of X. Since X has at most exponential growth, it follows that $\#S(x,k,l_t)$ is at most $P(l_t)c'c^{r_t}$ for some constants $c, c' \geqslant 1$. Take $f(l_t) = P(l_t)c'c^{r_t}$ and recall that, in the limit, $r_t/l_t \to 0$. We extend f to a function on $\mathbb{R}^+$ by setting $f(r) := f(l_t)$ for $r \in (l_{t-1}, l_t]$. This completes the proof. □

References

[1] **J Behrstock**, **M F Hagen**, **A Sisto**, *Hierarchically hyperbolic spaces, I: Curve complexes for cubical groups*, preprint (2014) arXiv

[2] **J Behrstock**, **M F Hagen**, **A Sisto**, *Asymptotic dimension and small-cancellation for hierarchically hyperbolic spaces and groups*, Proc. Lond. Math. Soc. 114 (2017) 890–926 MR

[3] **G C Bell**, *Growth of the asymptotic dimension function for groups*, preprint (2005) arXiv

[4] **G Bell**, **A Dranishnikov**, *Asymptotic dimension in Będlewo*, Topology Proc. 38 (2011) 209–236 MR

[5] **G C Bell**, **K Fujiwara**, *The asymptotic dimension of a curve graph is finite*, J. Lond. Math. Soc. 77 (2008) 33–50 MR

[6] **M Bestvina**, **K Bromberg**, **K Fujiwara**, *Constructing group actions on quasi-trees and applications to mapping class groups*, preprint (2010) arXiv

[7] **B H Bowditch**, *Coarse median spaces and groups*, Pacific J. Math. 261 (2013) 53–93 MR

[8] **B H Bowditch**, *Embedding median algebras in products of trees*, Geom. Dedicata 170 (2014) 157–176 MR

[9] **M R Bridson**, **A Haefliger**, *Metric spaces of non-positive curvature*, Grundl. Math. Wissen. 319, Springer (1999) MR

[10] **N P Brown**, **N Ozawa**, *C^*–algebras and finite-dimensional approximations*, Graduate Studies in Mathematics 88, Amer. Math. Soc., Providence, RI (2008) MR

[11] **S Campbell**, **G A Niblo**, *Hilbert space compression and exactness of discrete groups*, J. Funct. Anal. 222 (2005) 292–305 MR

[12] **I Chatterji**, **G Niblo**, *From wall spaces to* CAT(0) *cube complexes*, Internat. J. Algebra Comput. 15 (2005) 875–885 MR

[13] **V Chepoi**, *Graphs of some* CAT(0) *complexes*, Adv. in Appl. Math. 24 (2000) 125–179 MR

[14] **A N Dranishnikov**, *Groups with a polynomial dimension growth*, Geom. Dedicata 119 (2006) 1–15 MR

[15] **M Gromov**, *Hyperbolic groups*, from "Essays in group theory" (S M Gersten, editor), Math. Sci. Res. Inst. Publ. 8, Springer (1987) 75–263 MR

[16] **M Gromov**, *Asymptotic invariants of infinite groups*, from "Geometric group theory, II" (G A Niblo, M A Roller, editors), London Math. Soc. Lecture Note Ser. 182, Cambridge Univ. Press (1993) 1–295 MR

[17] **J R Isbell**, *Median algebra*, Trans. Amer. Math. Soc. 260 (1980) 319–362 MR

[18] **V A Kaimanovich**, *Boundary amenability of hyperbolic spaces*, from "Discrete geometric analysis" (M Kotani, T Shirai, T Sunada, editors), Contemp. Math. 347, Amer. Math. Soc., Providence, RI (2004) 83–111 MR

[19] **G A Niblo**, **L D Reeves**, *The geometry of cube complexes and the complexity of their fundamental groups*, Topology 37 (1998) 621–633 MR

[20] **B Nica**, *Cubulating spaces with walls*, Algebr. Geom. Topol. 4 (2004) 297–309 MR

[21] **P W Nowak**, **G Yu**, *Large scale geometry*, Eur. Math. Soc., Zürich (2012) MR

[22] **I Oppenheim**, *An intermediate quasi-isometric invariant between subexponential asymptotic dimension growth and Yu's property A*, Internat. J. Algebra Comput. 24 (2014) 909–922 MR

[23] **N Ozawa**, *Metric spaces with subexponential asymptotic dimension growth*, Internat. J. Algebra Comput. 22 (2012) art. id. 1250011 MR

[24] **L D Reeves**, *Biautomatic structures and combinatorics for cube complexes*, PhD thesis, University of Melbourne (1995)

[25] **J Roe**, *Hyperbolic groups have finite asymptotic dimension*, Proc. Amer. Math. Soc. 133 (2005) 2489–2490 MR

[26] **M Roller**, *Poc sets, median algebras and group actions: an extended study of Dunwoody's construction and Sageev's theorem*, preprint (2016) arXiv

[27] **M Sageev**, *Ends of group pairs and non-positively curved cube complexes*, Proc. London Math. Soc. 71 (1995) 585–617 MR

[28] **J Smith**, *The asymptotic dimension of the first Grigorchuk group is infinity*, Rev. Mat. Complut. 20 (2007) 119–121 MR

[29] **J Špakula**, **N Wright**, *Coarse medians and Property A*, preprint (2016) arXiv

[30] **N Wright**, *Finite asymptotic dimension for* CAT(0) *cube complexes*, Geom. Topol. 16 (2012) 527–554 MR

[31] **G Yu**, *The Novikov conjecture for groups with finite asymptotic dimension*, Ann. of Math. 147 (1998) 325–355 MR

[32] **G Yu**, *The coarse Baum–Connes conjecture for spaces which admit a uniform embedding into Hilbert space*, Invent. Math. 139 (2000) 201–240 MR

[33] **R Zeidler**, *Coarse median structures on groups*, Master's thesis, University of Vienna (2013) Available at `http://othes.univie.ac.at/28999/`

GA: *Universität Wien, Fakultät für Mathematik*
Vienna, Austria

GN, NW, JZ: *School of Mathematics, University of Southampton*
Southampton, United Kingdom

`goulnara.arzhantseva@univie.ac.at, g.a.niblo@soton.ac.uk,`
`n.j.wright@soton.ac.uk, jiawen.zhang@soton.ac.uk`

Received: 6 January 2017 Revised: 30 May 2017

msp

Algebraic & Geometric Topology 18 (2018) 525–546

Classifying spaces for 1–truncated compact Lie groups

CHARLES REZK

A 1–truncated compact Lie group is any extension of a finite group by a torus. In this note we compute the homotopy types of $\mathrm{Map}_*(BG, BH)$, $\mathrm{Map}(BG, BH)$, and $\mathrm{Map}(EG, B_GH)$ for compact Lie groups G and H with H 1–truncated, showing that they are computed entirely in terms of spaces of homomorphisms from G to H. These results generalize the well-known case when H is finite, and the case when H is compact abelian due to Lashof, May, and Segal.

55R91; 55P92, 55R35, 55R37

1 Introduction

By a 1–*truncated* compact Lie group H, we mean one whose homotopy groups vanish in dimensions 2 and greater. Equivalently, H is a compact Lie group with identity component H_0 a torus (isomorphic to some $U(1)^d$), ie an extension of a finite group by a torus.

The class of 1–truncated compact Lie groups includes (i) all finite groups, and (ii) all compact abelian Lie groups, both of which are included in the class of (iii) all groups which are isomorphic to a product of a compact abelian Lie group with a finite group, or equivalently, a product of a torus with a finite group.

The goal of this paper is to extend certain results, which were already known for finite groups, compact abelian Lie groups, or products thereof, to all 1–truncated compact Lie groups.

We write $\mathrm{Hom}(G, H)$ for the space of continuous homomorphisms, equipped with the compact-open topology. Our first theorem relates this to the space of based maps between classifying spaces.

1.1 Theorem *For G, H compact Lie groups with H 1–truncated, the evident map*

$$B\colon \mathrm{Hom}(G, H) \to \mathrm{Map}_*(BG, BH)$$

is a weak equivalence.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.525

Using this, we will derive an unbased variant.

1.2 Theorem *For G, H compact Lie groups with H 1–truncated, there is a weak equivalence*

$$\mathrm{Hom}(G, H) \times_H EH \to \mathrm{Map}(BG, BH).$$

Here H acts on $\mathrm{Hom}(G, H)$ *by conjugation: $h \cdot \phi = h\phi h^{-1}$.*

When H is discrete, these are well known and classical results. The case of H an abelian compact Lie group is proved by Lashof, May and Segal [7]; both the finite and compact abelian Lie cases are discussed by May in [10, Theorems 5 and 10].

1.3 Remark For G and H compact, there is a homeomorphism (see Proposition 7.1)

$$\mathrm{Hom}(G, H) \approx \coprod_{[\phi: G \to H]} H/C_H(\phi),$$

where the coproduct is over conjugacy classes of homomorphisms, and $C_H(\phi)$ is the centralizer of $\phi(G)$ in H. When H is a 1–truncated compact Lie group, we see from Theorem 1.1 that $\mathrm{Map}_*(BG, BH)$ is therefore weakly equivalent to this coproduct, and from Theorem 1.2 that there is a weak equivalence

$$\mathrm{Map}(BG, BH) \approx \coprod_{[\phi: G \to H]} BC_H(\phi).$$

Finally, we will give a description of the fixed points of the equivariant classifying space $B_G H$, which represents G–equivariant H–principal bundles, in the case that G and H are compact Lie groups and H is 1–truncated.

1.4 Theorem *For G, H compact Lie groups with H 1–truncated, the map*

$$\pi^*\colon B_G H \to \mathrm{Map}(EG, B_G H)$$

induced by restriction along $\pi\colon EG \to *$ is a G–equivariant weak equivalence.

The case when H is finite or compact abelian is proved in [10].

1.5 Remark For any closed subgroup $G' \leq G$, taking G' fixed points gives rise to maps

$$(B_G H)^{G'} \to \mathrm{Map}(EG, B_G H)^{G'} \xleftarrow{\sim} \mathrm{Map}(EG, BH)^{G'} \approx \mathrm{Map}(BG', BH),$$

and Theorem 1.4 amounts to saying that for any G', the first map in this sequence is a weak equivalence. (The middle map arises from a G–equivariant weak equivalence $\mathrm{Map}(EG, BH) \to \mathrm{Map}(EG, B_GH)$; see the proof of Lemma 6.4.) It is standard (see Lashof and May [6, Theorem 10]) that, for arbitrary compact G and H, $(B_GH)^{G'}$ is weakly equivalent to $\coprod_{[\phi\colon G'\to H]} BC_H(\phi)$, while if H is also 1–truncated, Theorem 1.2 and Remark 1.3 imply that $\mathrm{Map}(BG', BH)$ is also weakly equivalent to the same coproduct, thus giving an abstract weak equivalence $(B_GH)^{G'} \approx \mathrm{Map}(BG', BH)$. The point of Theorem 1.4 is to show that the map π^* exhibits this equivalence.

The map of Theorem 1.4 in a certain sense classifies the formation of the G–Borel quotient. That is, given a G–equivariant map $f\colon X \to B_GH$ classifying a G–equivariant principal H–bundle $P \to X$, the G–equivariant map $\pi^* f\colon X \to \mathrm{Map}(EG, B_GH)$ is adjoint to a nonequivariant map $X \times_G EG \to B_GH$ which classifies the bundle $P \times_G EG \to X \times_G EG$; see [10]. As a consequence of the theories of classifying spaces, we obtain the following.

1.6 Corollary *Let G and H be compact Lie groups with H 1–truncated. Then for a paracompact G–space X, formation of the G–Borel quotient gives rise to a bijection between* (i) *equivalence classes of G–equivariant principal H–bundles over X, and* (ii) *equivalence classes of principal H–bundles over $X \times_G EG$.*

1.7 Remark For comparison, there are well-known results (stemming from work of Dwyer and Zabrodsky [3] and Notbohm [12]) on spaces of maps *from* (rather than *to*) the classifying space of a p–toral group (a p–toral group is an extension of a *finite p–group* by a torus). For instance, [12, Theorem 1.3] may be interpreted as saying that for an arbitrary compact Lie group H and p–toral G, the map $B_GH \to \mathrm{Map}(EG, B_GH)$ (see Theorem 1.4) induces an isomorphism in mod p homology on fixed points for all closed subgroups of G (this interpretation is given as [10, Theorem 9]).

Organization of the paper The proof of Theorem 1.1 is the probably the most interesting part of the paper. It is carried out in Sections 2–4. The key ingredient is the use of the nerve $N(H, V)$ of the "exponential crossed module" (see Example 2.3) of the 1–truncated compact Lie group H. We first show that the simplicial space $N(H, V)$ is a Reedy fibrant model for the usual simplicial nerve NH of H (see Corollary 3.8), and so can be used to compute maps from BG to BH in terms of maps of simplicial spaces from NG to $N(H, V)$. The proof is completed in Section 4 by showing that, in a certain sense, the difference between $\mathrm{Hom}(G, H)$ and the space $\mathrm{Map}_{s\mathrm{Top}}(NG, N(H, V))$ of

maps between simplicial spaces is measured precisely by the continuous 2–cocycles on G with values in V, modulo boundaries of 1–cocycles. Because G is compact, Haar measure gives a contracting homotopy (see Proposition 4.2) on the complex of continuous chains on G. A sketch by the author of this proof originally appeared as an answer to a question on the site MathOverflow.[1]

Our approach gives a uniform proof of Theorem 1.1 for all 1–truncated compact Lie groups H; furthermore, even in the case of abelian H, it is somewhat more direct than the one given in [7].

We derived the unbased Theorem 1.2 from the based version Theorem 1.1 in Section 5 by comparing associated fibrations over BH.

The result on equivariant classifying spaces, Theorem 1.4, is proved in Section 6. The proof relies on an explicit model, built as the nerve of a certain topological category, of the restriction of the universal (G, H)–bundle to the fixed-point subspace $(B_G H)^G \subseteq B_G H$. The explicit model we use appears to be essentially of the type described by Guillou, May and Merling [4].

In Section 7, we give for the convenience of the reader a proof of the identification of $\mathrm{Hom}(G, H)$ as mentioned in Remark 1.3.

Conventions In this paper, we write Top for the category of *compactly generated weak Hausdorff spaces* (CGWH), the standard convenient category of spaces. This category is cartesian closed, and we write $\mathrm{Map}(X, Y)$ for the internal function object, ie continuous maps with the k–ification of the compact-open topology. We make use of the "usual" model structure on Top, in which weak equivalences are weak equivalences on homotopy groups, and fibrations are Serre fibrations.

Acknowledgments I thank Peter May for comments on a draft of this paper. The author was supported under NSF grant DMS-1406121.

2 Nerve of a topological crossed module

2.1 Crossed modules

Recall that a *crossed module* consists of

[1] "Equivariant classifying spaces from classifying spaces", `http://mathoverflow.net/q/223546`.

- groups H and V,
- a homomorphism $\epsilon\colon V \to H$,
- a homomorphism $\alpha\colon H \to \mathrm{Aut}(V)$ such that
  (i) $\epsilon(\alpha(h)(v)) = h\epsilon(v)h^{-1}$, and
  (ii) $\alpha(\epsilon(v))(v') = vv'v^{-1}$ for $h \in H$ and $v, v' \in V$.

A *topological crossed module* is one in which V and H are topological groups and ϵ and α are continuous. We typically write (H, V) for the crossed module, leaving ϵ and α understood. Note that we will often consider crossed modules in which V is an abelian group, in which case we will switch to additive notation for V, though not for H.

2.2 Example Given any group H, there is a unique crossed module, denoted $(H, 0)$, for which V is the trivial group.

2.3 Example (exponential crossed module) The following example is the crucial one for this paper. Suppose H is a 1–truncated compact Lie group. We set:

- $V := T_eH$, the Lie algebra of H, which is a group under addition of vectors.
- $\epsilon := \exp\colon V \to H$, the exponential map; this is a homomorphism since H_0 is abelian.
- $\alpha := \mathrm{ad}\colon H \to \mathrm{GL}(V)$, the adjoint action.

We typically write the group law of V additively, so the identities for the crossed module structure become

$$\exp(\mathrm{ad}(h)(v)) = h\exp(v)h^{-1}, \quad \mathrm{ad}(\exp(v))(v') = v + v' - v = v'.$$

The following features of this case will be significant:

(1) $\alpha = \mathrm{ad}\colon H \to \mathrm{GL}(V)$ factors through the quotient group H/H_0;
(2) $\epsilon = \exp\colon V \to H$ is a covering map;
(3) the underlying space of V is contractible.

2.4 Nerve of a crossed module

The *nerve* of a topological crossed module $N(H, V)$ is the simplicial space defined as follows; except for the topology, this is as in [1, Section 3.1]. The space $N(H, V)_n$ in degree n is the space of tuples

$$\big((h_{ij})_{0\le i\le j\le n}, (v_{ijk})_{0\le i\le j\le k\le n}\big), \quad h_{ij} \in H,\ v_{ijk} \in V$$

satisfying the identities

(1) $h_{ii} = e$ and $v_{iij} = v_{ijj} = e$ for all $i \le j$,

(2) $h_{ik} = \epsilon(v_{ijk}) h_{ij} h_{jk}$ for all $i \le j \le k$,

(3) $v_{ik\ell} v_{ijk} = v_{ij\ell}\, \alpha(h_{ij})(v_{jk\ell})$ for all $i \le j \le k \le \ell$.

The action of simplicial operators $\delta\colon [n] \to [m]$ is the evident one: $(\delta h)_{ij} = h_{\delta(i),\delta(j)}$ and $(\delta v)_{ijk} = v_{\delta(i),\delta(j),\delta(k)}$. A standard argument shows that, as a space, $N(H,V)_n \approx H^n \times V^{\binom{n}{2}}$, eg via the projection to coordinates h_{0i}, $1 \le i \le n$, and v_{0ij}, $1 \le i < j \le n$. Note that $N(H,V)_0 = *$; ie $N(H,V)$ is a *reduced* simplicial space.

2.5 Example The nerve of $N(H,0)$ is precisely the usual nerve of the group H; we write $N(H) := N(H,0)$.

2.6 Simplicial spaces and the Reedy model structure

We write $s\mathrm{Top}$ for the category of simplicial spaces, ie functors $\Delta^{\mathrm{op}} \to \mathrm{Top}$. We are going to use the *Reedy model structure* on $s\mathrm{Top}$. We will need to use the following features of this model structure:

(1) Weak equivalences $f\colon X \to Y$ in $s\mathrm{Top}$ are precisely the levelwise weak equivalences; ie $f_n\colon X_n \to Y_n$ is a weak equivalence for all $n \ge 0$.

(2) An object X is cofibrant (*Reedy cofibrant*) if and only if the latching space inclusions $\gamma_n\colon L_n X \to X_n$ are cofibrations in Top.

(3) An object Y is fibrant (*Reedy fibrant*) if and only if the matching space projections $\delta_n\colon Y_n \to M_n Y$ are fibrations in Top.

(4) The model structure is topological. In particular, if X is a cofibrant simplicial space and $Y \to Y'$ is a weak equivalence between fibrant simplicial spaces, then $\mathrm{Map}_{s\mathrm{Top}}(X,Y) \to \mathrm{Map}_{s\mathrm{Top}}(X,Y')$ is a weak equivalence of spaces.

We will need to examine latching and matching spaces in a bit more detail.

2.7 Latching and matching spaces

We recall the notion of latching and matching spaces. For simplicial spaces $X\colon \Delta^{\mathrm{op}} \to \mathrm{Top}$ and all $n \ge 0$, we have natural maps of spaces

$$L_n X \xrightarrow{\gamma_n} X_n \xrightarrow{\delta_n} M_n X,$$

where

$$L_n X = \mathrm{colim}_{(\Delta^{\mathrm{op}}_{/[n]})_{<n}} X, \quad M_n = \lim_{(\Delta^{\mathrm{op}}_{[n]/})_{<n}} X,$$

called the *latching* and *matching* spaces of X.

2.8 Latching spaces for the nerve of a group

2.9 Proposition *Let G be a topological group, and $NG \in s\mathrm{Top}$ its nerve. Then for each $n \geq 0$, the latching inclusion $\gamma_n\colon L_n(NG) \to (NG)_n$ is isomorphic to the inclusion*

$$\{(g_1, \dots, g_n) \mid g_i = e \text{ for some } i\} \to G^n.$$

In particular, NG is Reedy cofibrant if $\{e\} \to G$ is a cofibration in Top.

The proof of this is standard.

2.10 Matching spaces for the nerve of a crossed module

We describe the matching projections for the nerve of a topological crossed module.

2.11 Proposition *Consider $N := N(H, V)$ the nerve of a topological crossed module. We write $M_n := M_n N$ for its matching spaces.*

(0) *$\delta_0\colon N_0 \to M_0$ is the isomorphism of* 1*–point spaces.*

(1) *$\delta_1\colon N_1 \to M_1$ is the projection $H \to *$.*

(2) *$M_2 \approx H^{\times 3}$, and there is a pullback square*

$$\begin{array}{ccc} N_2 & \longrightarrow & V \\ {\scriptstyle \delta_2}\downarrow & & \downarrow{\scriptstyle \epsilon} \\ M_2 & \xrightarrow[(h_{01},h_{02},h_{12})\mapsto h_{02}h_{12}^{-1}h_{02}^{-1}]{} & H \end{array}$$

(3) *There is a commutative diagram*

$$\begin{array}{ccccc} N_3 & \longrightarrow & \{e\} & & \\ {\scriptstyle \delta_3}\downarrow & & \downarrow & & \\ M_3 & \longrightarrow & \operatorname{Ker}\epsilon & \longrightarrow & \{e\} \\ \downarrow & & \downarrow & & \downarrow \\ H^{\times 3} \times V^{\times 3} & \xrightarrow[\pi]{} & V & \xrightarrow[\epsilon]{} & H \end{array}$$

in which all squares are pullback squares, and π is given by

$$(h_{01}, h_{12}, h_{23}, v_{012}, v_{013}, v_{023}, v_{123}) \mapsto v_{023}v_{012}\alpha(h_{01})(v_{123})^{-1}v_{013}^{-1}.$$

(≥ 4) *$\delta_n\colon N_n \to M_n$ is an isomorphism for $n \geq 4$.*

Proof This is straightforward. In (3), one shows directly that the right-hand lower square, bottom rectangle, and left rectangle are pullbacks. □

Recall that a simplicial space X is *Reedy fibrant* if each of the maps $\delta_n\colon X_n \to M_n X$ is a fibration of spaces.

2.12 Corollary *If (H, V) is a topological crossed module such that ϵ is a covering map, then $N(H, V)$ is Reedy fibrant.*

Proof This is immediate using Proposition 2.11. Note that the condition that ϵ be a covering map in Proposition 2.11(3) implies that δ_3 is an open and closed embedding. □

In particular, Corollary 2.12 applies to our main Example 2.3.

3 Maps between reduced simplicial spaces

A simplicial space $X \in s\mathrm{Top}$ is said to be *reduced* if $X_0 \approx *$. We write $s\mathrm{Top}^{\mathrm{red}} \subset s\mathrm{Top}$ for the full subcategory of reduced simplicial spaces. Note that reduced simplicial spaces are canonically based, so we may in fact regard $s\mathrm{Top}^{\mathrm{red}}$ as a full subcategory of simplicial based spaces $s\mathrm{Top}_*$.

3.1 Realization of reduced simplicial spaces

We recall the geometric realization functor $\|-\|\colon s\mathrm{Top} \to \mathrm{Top}$, defined so that $\|X\|$ is the coend of the functor $\Delta^{\mathrm{op}} \times \Delta \to \mathrm{Top}$ given by $([m],[n]) \mapsto X_m \times \Delta^n$, where Δ^n is the topological n–simplex.

3.2 Proposition *The restriction of the geometric realization functor to a functor $\|-\|\colon s\mathrm{Top}^{\mathrm{red}} \to \mathrm{Top}_*$ admits a right adjoint $\nabla\colon \mathrm{Top}_* \to s\mathrm{Top}^{\mathrm{red}}$ defined by*

$$(\nabla Y)_n := \mathrm{Map}_*(\Delta^n / \operatorname{Sk}_0 \Delta^n, Y),$$

where $\operatorname{Sk}_0 \Delta^n \subseteq \Delta^n$ is the set of vertices of the simplex. The adjunction is compatible with the topological enrichment, and so gives a natural homeomorphism

$$\mathrm{Map}_{s\mathrm{Top}}(X, \nabla Y) \approx \mathrm{Map}_*(\|X\|, Y)$$

for $X \in s\mathrm{Top}^{\mathrm{red}}$ and $Y \in \mathrm{Top}_*$.

Proof This is a straightforward consequence of the observation that, for reduced simplicial spaces X, we see that $\|X\|$ is isomorphic to the coend (in Top_*) of $([m],[n]) \mapsto X_m \wedge (\Delta^n / \operatorname{Sk}_0 \Delta^n)$. □

3.3 Proposition *For any $Y \in \mathrm{Top}_*$, the simplicial space ∇Y is Reedy fibrant.*

Proof The matching space projection has the form

$$\mathrm{Map}_*(\Delta^n/\operatorname{Sk}_0 \Delta^n, Y) \to \mathrm{Map}_*(\partial\Delta^n/\operatorname{Sk}_0 \Delta^n, Y),$$

which is clearly a fibration. □

For a topological group H, we consider the classifying space $BH := \|NH\|$.

3.4 Proposition *If H is a topological group with identity element a nondegenerate basepoint (ie $\{e\} \to H$ has the HEP), then the map*

$$\eta\colon NH \to \nabla\|NH\| = \nabla BH$$

given by the unit map of the adjunction of Proposition 3.2 is a levelwise weak equivalence of simplicial spaces.

Proof In degree 0, the map η is the isomorphism of one-point spaces. In degree 1, it has the form

$$H \to \mathrm{Map}_*(\Delta^1/\{0,1\}, \|NH\|) \approx \Omega BH.$$

A standard argument (eg using the usual simplicial model for the universal fibration [9]) shows that this is a weak equivalence.

For $n \geq 2$, we reduce to the $n = 1$ case using the fact that $I_n/\operatorname{Sk}_0 \Delta^n \to \Delta^n/\operatorname{Sk}_0 \Delta^n$ is a homotopy equivalence of pointed spaces, and thus

$$\mathrm{Map}_*(\Delta^n/\operatorname{Sk}_0 \Delta^n,\ \|NH\|) \to \mathrm{Map}_*(I_n/\operatorname{Sk}_0 \Delta^n,\ \|NH\|) \approx (\Omega BG)^{\times n}$$

is a weak equivalence, where $I_n \subseteq \Delta^n$ is the union of the edges with vertices $\{k-1, k\}$ for all $k = 1, \ldots, n$. □

3.5 Map$(X, N(H,V))$ computes Map$_*(\|X\|, BH)$

Now we fix a 1–truncated compact Lie group H and the corresponding exponential crossed module (H, V) of Example 2.3. We have a map of reduced simplicial spaces

$$NH \xrightarrow{(\iota,\eta)} N(H,V) \times \nabla\|NH\|$$

in which ι is the evident inclusion $NH = N(H,0) \subseteq N(H,V)$ and η the unit map of the adjunction of Proposition 3.2. Observe that both ι and η are levelwise weak

equivalences (ι because V is contractible, η by Proposition 3.4). Furthermore, both $N(H,V)$ (see Corollary 2.12) and $\nabla\|NH\|$ (see Proposition 3.3) are Reedy fibrant.

Using the Reedy model structure on simplicial spaces, we can factor the above map as

$$NH \xrightarrow{j} (NH)^f \xrightarrow{(\iota',\eta')} N(H,V)\times\nabla\|NH\| \tag{3.6}$$

so that $(NH)^f$ is Reedy fibrant and j is a levelwise weak equivalence, whence ι' and η' are also levelwise weak equivalences.

3.7 Proposition *For X a Reedy cofibrant simplicial space with $X_0 = *$, and (H,V) the exponential crossed module of a 1–truncated compact Lie group H, we have that $\mathrm{Map}_{s\mathrm{Top}}(X, N(H,V))$ is weakly equivalent to $\mathrm{Map}_*(\|X\|, \|NH\|)$. Furthermore, $\iota_*\colon \mathrm{Map}_{s\mathrm{Top}}(X, NH) \to \mathrm{Map}_{s\mathrm{Top}}(X, N(H,V))$ is a weak equivalence of spaces if and only if $\eta_*\colon \mathrm{Map}_{s\mathrm{Top}}(X, NH) \to \mathrm{Map}_{s\mathrm{Top}}(X, \nabla\|NH\|)$ is.*

Proof This is straightforward using the factorization (3.6), the fact that Reedy model structure is compatible with the topological enrichment, and the adjunction from Proposition 3.2. □

3.8 Corollary *If (H,V) is as above, and G is a topological group such that $\{e\} \to G$ is a cofibration, then $\mathrm{Map}_*(BG, BH)$ is weakly equivalent to $\mathrm{Map}_{s\mathrm{Top}}(NG, N(H,V))$, and*

$$B\colon \mathrm{Hom}(G,H) \to \mathrm{Map}_*(BG, BH)$$

is a weak equivalence if and only if

$$\iota_*\colon \mathrm{Map}_{s\mathrm{Top}}(NG, NH) \to \mathrm{Map}_{s\mathrm{Top}}(NG, N(H,V))$$

is a weak equivalence.

Proof Use Proposition 3.7 with $X = NG$, which is Reedy cofibrant by Proposition 2.9. It is straightforward to see that $\mathrm{Map}_{s\mathrm{Top}}(NG, NH) \to \mathrm{Hom}(G,H)$ (evaluation at spaces in degree 1) is a homeomorphism, and so the map B coincides with ι_*. □

3.9 Remark If H is a discrete group, then NH is already Reedy fibrant, in which case we can immediately derive the well-known fact that $B\colon \mathrm{Hom}(G,H) \to \mathrm{Map}_*(BG, BH)$ is a weak equivalence for any such topological group G.

4 Proof of Theorem 1.1 Based mapping space

As above, we assume that H is a 1–truncated compact Lie group. We will now also assume that G is a compact Lie group. By Corollary 3.8, we have reduced Theorem 1.1 to showing that $\mathrm{Map}_{s\mathrm{Top}}(NG, NH) \to \mathrm{Map}_{s\mathrm{Top}}(NG, N(H,V))$ is a weak equivalence.

Let $E := \mathrm{Map}_{s\mathrm{Top}}(NG, N(H,V))$. Using Proposition 2.11 and the identification of the latching inclusions $L_nNG \to G^n$ (see Proposition 2.9), we see that E is precisely the space of pairs

$$(\zeta, \nu) \in \mathrm{Map}(G, H) \times \mathrm{Map}(G \times G, V)$$

such that

(1) $\zeta(e) = e$ and $\nu(g,e) = 0 = \nu(e,g)$ for $g \in G$,

(2) $\zeta(g_1g_2) = \exp[\nu(g_1,g_2)]\zeta(g_1)\zeta(g_2)$ for $g_1, g_2 \in G$,

(3) $\nu(g_1g_2,g_3)+\nu(g_1,g_2) = \nu(g_1,g_2g_3)+\mathrm{ad}(\zeta(g_1))[\nu(g_2,g_3)]$ for $g_1,g_2,g_3 \in G$.

Explicitly, this corresponds to the map $NG \to N(H,V)$ which (in the notation of Section 2.1) sends $(g_{ij}) \in (NG)_n$ to $(h_{ij}, v_{ijk}) \in N(H,V)_n$ with $h_{ij} = \zeta(g_{ij})$ and $v_{ijk} = \nu(g_{ij}, g_{jk})$.

Let $E^0 := \mathrm{Map}_{s\mathrm{Top}}(NG, NH)$. The map $E^0 \to E$ is precisely inclusion into the subspace consisting of points of the form $(\zeta, 0)$.

For a continuous map $\zeta\colon G \to H$, we write $\overline{\zeta}\colon G \to H/H_0$ for the composite with the quotient map $H \to H/H_0$. Note that if $(\zeta,\nu) \in E$, then $\overline{\zeta}$ is a continuous homomorphism of groups. Since H/H_0 is discrete, we obtain coproduct decompositions

$$E = \coprod_{\gamma} E_\gamma, \quad E^0 = \coprod_{\gamma} E^0_\gamma, \quad \gamma \in \mathrm{Hom}(G, H/H_0).$$

Thus, we must show that for each such γ, the inclusion $E^0_\gamma \subseteq E_\gamma$ is a weak equivalence. In fact, we can give an explicit (strong) deformation retraction of E_γ to E^0_γ, which relies on the existence of a contracting homotopy of the complex $C^*(G, V_{\mathrm{ad}\,\gamma})$ of normalized continuous cochains on G with values in the representation $\mathrm{ad}\,\gamma\colon G \to \mathrm{Aut}(V)$, which may be constructed explicitly using an invariant measure on the compact group G. We spell out the details we need below.

Fix $\gamma \in \mathrm{Hom}(G, H/H_0)$. Let $C^1_\gamma \subseteq \mathrm{Map}(G,V)$ be the subspace of functions $\mu\colon G \to V$ such that

$$\mu(e) = 0.$$

Let $Z^2_\gamma \subseteq \mathrm{Map}(G\times G, V)$ be the subspace of functions $\nu\colon G\times G\to V$ such that

$$\nu(g,e)=0=\nu(e,g),\quad g\in G,$$

and

$$\nu(g_1g_2,g_3)+\nu(g_1,g_2)=\nu(g_1,g_2g_3)+\mathrm{ad}\,\gamma(g_1)[\nu(g_2,g_3)],\quad g_1,g_2,g_3\in G.$$

Both Z^2_γ and C^1_γ are topological real vector spaces. Define continuous and linear maps

$$d\colon C^1_\gamma\to Z^2_\gamma,\quad H\colon Z^2_\gamma\to C^1_\gamma$$

by

$$d\mu(g_1,g_2):=\mu(g_1)-\mu(g_1g_2)+\mathrm{ad}\,\gamma(g_1)\mu(g_2),$$
$$H\nu(g):=\int_G x^{-1}\nu(x,g)\,dx,$$

where we use right-invariant Haar measure on G normalized so that $\int_G dx=1$.

4.1 Lemma *The composite $dH\colon Z^2_\gamma\to Z^2_\gamma$ is the identity map.*

Proof For $g\in G$ and $v\in V$, write $g\,v$ for $\mathrm{ad}\,\gamma(g)(v)$ below. Given $\nu\in Z^2_\gamma$, we have

$$\begin{aligned}
dH\nu(g_1,g_2)&=\int_G\big(x^{-1}\nu(x,g_1)-x^{-1}\nu(x,g_1g_2)+g_1x^{-1}\nu(x,g_2)\big)\,dx\\
&=\int_G\big(x^{-1}\nu(x,g_1)-x^{-1}[\nu(xg_1,g_2)+\nu(x,g_1)-x\nu(g_1,g_2)]\\
&\qquad\qquad +g_1x^{-1}\nu(x,g_2)\big)\,dx\\
&=\nu(g_1,g_2)-\int_G g_1(xg_1)^{-1}\nu(xg_1,g_2)\,dx+\int_G g_1x^{-1}\nu(x,g_2)\,dx\\
&=\nu(g_1,g_2),
\end{aligned}$$

where the last cancellation is by right invariance of the measure. □

4.2 Proposition *The inclusion $E^0_\gamma\subseteq E_\gamma$ admits a strong deformation retraction.*

Proof Define $K_t\colon E_\gamma\to E_\gamma$ for $0\le t\le 1$ by $K_t(\zeta,\nu):=(\zeta_t,\nu_t)$, with

$$\zeta_t(g):=\exp[tH\nu(g)]\zeta(g),$$
$$\nu_t(g_1,g_2):=\nu(g_1,g_2)-t\,dH\nu(g_1,g_2).$$

We have $K_0=\mathrm{id}_{E_\gamma}$, $K_t|E^0_\gamma=\mathrm{id}_{E^0_\gamma}$, and $K_1(E_\gamma)\subseteq E^0_\gamma$ as desired, the last using Lemma 4.1. □

The proof of Theorem 1.1 follows using Proposition 3.7 and the remarks above.

4.3 Remark If H is an abelian group, then ad: $H \to \mathrm{Aut}(V)$ is trivial. In this case, the proof of Proposition 4.2 directly gives a deformation retraction of $E^0 \subseteq E$.

5 Proof of Theorem 1.2 Unbased mapping space

Given simplicial spaces X and Y, we have an internal function object $\underline{\mathrm{Map}}(X, Y) \in s\mathrm{Top}$, characterized so that $\underline{\mathrm{Map}}(X, -)$ is the right adjoint to $(-) \times X$. We have that

$$\underline{\mathrm{Map}}(X, Y)_n = \mathrm{Map}_{s\mathrm{Top}}(X \times N[n], Y),$$

where $[n]$ is the n–arrow category. In particular, $\underline{\mathrm{Map}}(X, Y)_0 \approx \mathrm{Map}_{s\mathrm{Top}}(X, Y)$.

Formation of the internal function object is compatible with realization: there are canonical maps

$$\rho\colon \|\underline{\mathrm{Map}}(X, Y)\| \to \mathrm{Map}(\|X\|, \|Y\|), \tag{5.1}$$

natural in X and Y. This map exists exactly because $\|-\|\colon s\mathrm{Top} \to \mathrm{Top}$, the realization functor, preserves finite products and is characterized as the map adjoint to

$$\|\underline{\mathrm{Map}}(X, Y)\| \times \|X\| \xleftarrow{\sim} \|\underline{\mathrm{Map}}(X, Y) \times X\| \xrightarrow{|\mathrm{eval}|} \|Y\|.$$

Given topological groups G and H, we consider the function object $\underline{\mathrm{Map}}(NG, NH)$. We have an evident isomorphism

$$\underline{\mathrm{Map}}(NG, NH) \approx N\mathrm{Fun}(G, H),$$

where $\mathrm{Fun}(G, H)$ is the internal category in Top of functors and natural transformations from G to H. Explicitly, this has

- objects $\phi \in \mathrm{Hom}(G, H)$, and
- morphisms $\phi_0 \xrightarrow{h} \phi_1$, where $h \in H$ and $\phi_1 = h\phi_0 h^{-1}$,

and thus homeomorphisms $N\mathrm{Fun}(G, H)_n = \underline{\mathrm{Map}}(NG, NH)_n = \mathrm{Hom}(G, H) \times H^{\times n}$.

Write $(H \curvearrowright H)$ for the translation category of the left action of H on itself, viewed as a category object in Top. This has

- objects $h_0 \in H$, and
- morphisms $h_0 \xrightarrow{h} h_1$, where $h \in H$ and $h_1 = hh_0$.

We have homeomorphisms $N(H \curvearrowright H)_n = H^{\times(n+1)}$. The group H acts on the category $(H \curvearrowright H)$ by $\delta \cdot h_0 = h_0\delta^{-1}$ (on objects) and $\delta \cdot (h_0 \xrightarrow{h} h_1) = h_0\delta^{-1} \xrightarrow{h} h_1\delta^{-1}$ (on morphisms), where $\delta \in H$.

We let $EH := \|N(H \curvearrowright H)\|$, a contractible H space with free H–action.

5.2 Lemma *There is a homeomorphism* $\|\underline{\mathrm{Map}}(NG, NH)\| \approx (\mathrm{Hom}(G, H) \times EH)/H$, *where* H *acts on* $\mathrm{Hom}(G, H)$ *by conjugation.*

Proof of Theorem 1.2 We have a commutative diagram

$$\begin{array}{ccc} \|\underline{\mathrm{Map}}(NG, NH)\| & \xrightarrow{\rho} & \mathrm{Map}(\|NG\|, \|NH\|) \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} \\ \|\underline{\mathrm{Map}}(*, NH)\| & \xrightarrow[\approx]{} & \mathrm{Map}(\|*\|, \|NH\|) \end{array}$$

where the vertical maps are induced by restriction along $* \to NG$, and the lower horizontal map is the evident homeomorphism (both source and target are homeomorphic to BH). We claim that ρ is a weak equivalence.

By Lemma 5.2 we see that $\alpha\colon (\mathrm{Hom}(G, H) \times EH)/H \to BH$ is a fiber bundle with fiber $\mathrm{Hom}(G, H)$. Since $\beta\colon \mathrm{Map}(BG, BH) \to BH$ is also a fibration, and the base space BH is path connected, ρ is a weak equivalence if and only if its restriction to the fiber over the base point is, which is precisely the weak equivalence $\mathrm{Hom}(G, H) \to \mathrm{Map}_*(BG, BH)$ of Theorem 1.1. □

6 Proof of Theorem 1.4 Equivariant classifying space

6.1 Recollections on equivariant bundles

A *G–equivariant principal H–bundle* (or *(G, H)–bundle*), is a principal H–bundle $\pi\colon P \to X$, together with actions of G on P and X, compatible with π, such that G acts via maps of principal H–bundles. We will always assume that both G and H are compact Lie groups.

This definition is somewhat anomalous in that (G, H)–bundles are not characterized by a property which is local in X. Thus, we say that a (G, H)–bundle is *locally trivial* if it looks locally like

$$(G \times H) \times_{\Lambda_\phi} U \to G \times_{G'} U,$$

where $G' \leq G$ is a closed subgroup, $\Lambda_\phi := \{(g, \phi(g)) \mid g \in G'\}$ is the graph of some homomorphism $\phi\colon G' \to H$, and $\Lambda \xrightarrow{\sim} G'$ acts on a space U. The key result is that if G and H are compact and X is completely regular,[2] then any (G, H)–bundle over X is locally trivial; see [5, Corollary 1.5].

A (G, H)–bundle $P \to X$ is *numerable* if it admits a locally trivializing cover which itself admits a subordinate partition of unity by G–invariant functions. Over a paracompact base X, every locally trivial bundle is numerable; see [5, Corollary 1.13]. There is a universal (G, H)–bundle $E_G H \to B_G H$, which classifies equivalence classes of numerable bundles; see [5], [6], and also [8] for a recent and more general treatment.

We will be mainly concerned with the case of (G, H)–bundles $\pi\colon P \to X$ such that G acts trivially on X. In such a case, there is a natural function

$$\tau\colon P \to \mathrm{Hom}(G, H)$$

defined so that $\tau(p)(\gamma) \in H$ is the unique $\delta \in H$ such that $(\gamma, \delta) \cdot p = p$. When $P \to X$ is locally trivial, the map τ is seen to be continuous. Observe that τ is $G \times H$–equivariant, where this group acts on $\mathrm{Hom}(G, H)$ by conjugation: $(\gamma, \delta) \cdot \phi = \delta\phi(\gamma)^{-1}\phi\phi(\gamma)\delta^{-1}$.

6.2 Lemma *For any locally trivial (G, H)–bundle $\pi\colon P \to X$ over a G–fixed base X, the map $\tau\colon P \to \mathrm{Hom}(G, H)$ is a Serre fibration.*

Proof This will follow by showing that $(\tau, \pi)\colon P \to \mathrm{Hom}(G, H) \times X$ is actually a fiber bundle. Since π is locally trivial, we can reduce to the case when π has the form $\pi\colon (G \times H)/\Lambda_\phi \times U \to U$, where $\Lambda_\phi \leq G \times H$ is the graph of some homomorphism $\phi\colon G \to H$. Then

$$(\tau, \pi) = \rho \times \mathrm{id}_U\colon (G \times H)/\Lambda_\phi \times U \to \mathrm{Hom}(G, H) \times U,$$

where $\rho\colon (G \times H)/\Lambda_\phi \to \mathrm{Hom}(G, H)$ sends $[\gamma, \delta] \mapsto \delta\phi(\gamma)^{-1}\phi(\delta\phi(\gamma)^{-1})^{-1}$. Because $\mathrm{Hom}(G, H)$ is topologically a coproduct of orbits under H–conjugation (see Remark 1.3 and Proposition 7.1), we see that ρ is isomorphic to the composite of a projection map $(G \times H)/\Lambda_\phi \to H/C_H(\phi)$ (induced by $(\gamma, \delta) \mapsto \delta\phi(\gamma)^{-1}$) with an open and closed immersion, and thus is a fibration. □

[2] Completely regular means that points are closed, and any point and disjoint closed subset are separated by a real valued function.

6.3 Outline of the proof

To prove that the map $B_G H \to \mathrm{Map}(EG, B_G H)$ (induced by restriction along $EG \to *$) is a G–equivariant weak equivalence, it suffices to show that it induces a weak equivalence of spaces $(B_G H)^{G'} \to \mathrm{Map}(EG, B_G H)^{G'}$ for all closed subgroups $G' \leq G$. Without loss of generality, we may assume $G' = G$ since, when the group action is restricted to the subgroup G', we have that $B_G H$ is a $B_{G'} H$ and EG is an EG'. Thus, we will show that $(B_G H)^G \to \mathrm{Map}(EG, B_G H)^G$ is an equivalence, using the following.

6.4 Lemma *Suppose we are given a (G, H)–bundle $P \to X$ over a space X with trivial G–action, together with maps*

(1) *$\alpha\colon X \to (B_G H)^G \subseteq B_G H$ classifying the G–equivariant H–bundle $P \to X$, ie covered by a (G, H)–bundle map $P \to E_G H$, and*

(2) *$\rho\colon X \to \mathrm{Map}(BG, BH)$, whose adjoint $\widetilde{\rho}\colon X \times BG = (X \times EG)/G \to BH$ classifies the H–bundle $(P \times EG)/G \to (X \times EG)/G$, ie is covered by an H–bundle map $(P \times EG)/G \to EH$.*

Then the diagram

$$\begin{array}{ccc} X & \xrightarrow{\quad\alpha\quad} & (B_G H)^G \\ {\scriptstyle\rho}\downarrow & & \downarrow \\ \mathrm{Map}(BG, BH) & \xrightarrow[\sim]{} & \mathrm{Map}(EG, B_G H)^G \end{array} \tag{6.5}$$

commutes up to homotopy, where the bottom map is induced by a G–equivariant map $\upsilon\colon BH \to B_G H$ (with G acting trivially on BH) which classifies the universal H–bundle viewed as a G–equivariant H–bundle with trivial G–action.

Furthermore, the bottom map of the diagram is a weak equivalence.

Proof The adjoints of both composite maps $X \to \mathrm{Map}(EG, B_G H)^G$ are G–equivariant maps $(X \times EG)/G \to B_G H$, which in either case are covered by maps $(P \times EG)/G \to E_G H$ of (G, H)–bundles. The homotopy-commutativity of the diagram follows from the universal property of $B_G H$ as the classifying space for such bundles.

To see that the bottom map of the diagram is a weak equivalence, note that it may be constructed as follows. Choose any $G \times H$ equivariant map $EH \to E_G H$ (unique up to homotopy by the defining property of $E_G H$), and take the quotient with respect to the free H–actions, obtaining a G–equivariant map $\upsilon\colon BH \to B_G H$ which classifies

the universal bundle as stated. By construction, υ is a G–equivariant map which is a weak equivalence on underlying spaces.

For any G–space X and subgroup $G' \leq G$, the fixed-point space $\mathrm{Map}(EG, X)^{G'}$ is the space of ordinary homotopy fixed points of the G'–action on X; as a consequence, $\mathrm{Map}(EG, -)$ takes a map of G–spaces which is a weak equivalence of underlying spaces, to a G–equivariant weak equivalence. In particular, we see that $\mathrm{Map}(EG, BH) \to \mathrm{Map}(EG, B_G H)$ is a G–equivariant weak equivalence. Taking G–fixed points gives an equivalence

$$\mathrm{Map}(BG, BH) \approx \mathrm{Map}(EG, BH)^G \to \mathrm{Map}(EG, B_G H)^G,$$

which is the map in the diagram. □

The strategy is as follows. Fix compact Lie groups G and H, and take $\rho\colon X \to \mathrm{Map}(BG, BH)$ in Lemma 6.4 to be isomorphic to the map $(\mathrm{Hom}(G, H) \times EH)/H \to \mathrm{Map}(BG, BH)$ described in Section 5 which, for 1–truncated H, gives the weak equivalence of Theorem 1.2. We will

(1) construct a certain (G, H)–bundle $P \to X$ (where G acts trivially on X),

(2) prove that a map $\alpha\colon X \to B_G H$ classifying $P \to X$ induces a weak equivalence $X \xrightarrow{\sim} (B_G H)^G \subseteq B_G H$, and

(3) construct a bundle map $(P \times EG)/G \to EH$ covering $\widetilde{\rho}\colon X \times BG \to BH$, the adjoint to ρ.

Thus by Lemma 6.4, both α and ρ fit in a homotopy commutative square (6.5). It follows that if H is 1–truncated, Theorem 1.2 implies that ρ is a weak equivalence, from which it follows that the right-hand vertical arrow is a weak equivalence, which is the desired result. Note: the hypothesis that H is 1–truncated is used only to show that ρ (which exists for arbitrary H) is a weak equivalence.

6.6 Step 1: Construction of $P \to X$

As in the previous section, we consider categories $\mathrm{Fun}(G, H)$ and $(H \curvearrowright H)$ (internal to Top), where G and H are compact Lie groups. Consider the topological category C defined as the fiber product

$$C := \mathrm{Fun}(G, H) \times_H (H \curvearrowright H)$$

via the evident restriction functors $\mathrm{Fun}(G,H)\to\mathrm{Fun}(\{e\},H)=H$ and $(H\curvearrowright H)\to(H\curvearrowright *)=H$. (Here H represents a topological category with one object.)

The group $G\times H$ acts on C via

$$(\gamma,\delta)\cdot(\phi_0,h_0)=(\phi_0,\phi_0(\gamma)h_0\delta^{-1})$$

on objects, and

$$(\gamma,\delta)\cdot(\phi_0\xrightarrow{h}\phi_1,h_0\xrightarrow{h}h_1)=(\phi_0\xrightarrow{h}\phi_1,\phi_0(\gamma)h_0\delta^{-1}\xrightarrow{h}\phi_1(\gamma)h_1\delta^{-1})$$

on morphisms, for $(\gamma,\delta)\in G\times H$. (This works exactly because $h\phi_0(\gamma)=\phi_1(\gamma)h$.) The evident projection functor $C\to\mathrm{Fun}(G,H)$ is invariant under the $G\times H$–action on C, with respect to the trivial action on $\mathrm{Fun}(G,H)$.

We set $P:=\|NC\|$ and $X:=\|N\mathrm{Fun}(G,H)\|$, with $P\to X$ induced by the evident projection functor. It is straightforward to show that the induced $G\times H$–action on P is compatible with the projection map to X and that H acts freely on P with $P/H\approx X$. In particular, $P\to X$ has the structure of a G–equivariant principal H–bundle.

We note an equivalent description of C, and hence of P. Let

$$C':=\mathrm{Hom}(G,H)\times N(H\curvearrowright H),$$

where $\mathrm{Hom}(G,H)$ is viewed as a topological category with only identity maps. There is an isomorphism $C'\to C$ of topological categories, given on objects and morphisms by

$$(\phi,h_0)\mapsto(h_0\phi h_0^{-1},\,h_0),\qquad(\phi,h_0\xrightarrow{h}h_1)\mapsto(h_0\phi h_0^{-1}\xrightarrow{h}h_1\phi h_1^{-1},\,h_0\xrightarrow{h}h_1).$$

The $G\times H$–action on C' induced by this isomorphism is described by

$$\begin{aligned}(\gamma,\delta)\cdot(\phi,h_0)&=(\delta\phi(\gamma)^{-1}\phi\phi(\gamma)\delta^{-1},\,h_0\phi(\gamma)\delta^{-1}),\\(\gamma,\delta)\cdot(\phi,h_0\xrightarrow{h}h_1)&=(\delta\phi(\gamma)^{-1}\phi\phi(\gamma)\delta^{-1},\,h_0\phi(\gamma)\delta^{-1}\xrightarrow{h}h_1\phi(\gamma)\delta^{-1}).\end{aligned}$$

In particular, the projection functor $C'\to\mathrm{Hom}(G,H)$ induces a $G\times H$–equivariant map $P\to\mathrm{Hom}(G,H)$ (using the conjugation $G\times H$–action on $\mathrm{Hom}(G,H)$), and this map is a nonequivariant weak equivalence since $P\approx\|NC'\|\approx\mathrm{Hom}(G,H)\times EH$.

6.7 Step 2: The weak equivalence $\alpha\colon X\to(B_GH)^G$

Choose any $X\to B_GH$ classifying the $P\to X$ constructed above (this exists because X is paracompact and completely regular, so numerable), and so covered by a $G\times H$–equivariant map $P\to E_GH$. Since the action of G on X is trivial, these factor

through $\alpha\colon X \to (B_G H)^G$ and $\alpha'\colon P \to p^{-1}((B_G H)^G)$, where $p\colon E_G H \to B_G H$ is the universal bundle.

6.8 Lemma *The map $\alpha'\colon P \to p^{-1}((B_G H)^G)$ is a weak equivalence of underlying spaces.*

Proof The map α' fits in the commutative diagram

$$\begin{array}{ccc} P & \xrightarrow{\quad\alpha'\quad} & p^{-1}((B_G H)^G) \\ & {\scriptstyle\tau}\searrow \quad \swarrow{\scriptstyle\tau} & \\ & \mathrm{Hom}(G,H) & \end{array}$$

where by Lemma 6.2, both maps marked τ are Serre fibrations. The fibers of these τ over $\phi \in \mathrm{Hom}(G,H)$ are EH and $(E_G H)^{\Lambda_\phi}$, respectively, both of which are contractible spaces. Thus α' is a weak equivalence (as are both τ). □

It follows that $\alpha\colon X \to (B_G H)^G$ is a weak equivalence as it is obtained by the quotient of α' by free H–actions.

6.9 Step 3: The bundle map covering $\rho\colon X \to \mathbf{Map}(BG, BH)$

We have a commutative square of functors

$$\begin{array}{ccc} C \times (G \curvearrowright G) = \mathrm{Fun}(G,H) \times_H (H \curvearrowright H) \times (G \curvearrowright G) & \longrightarrow & (H \curvearrowright H) \\ \downarrow & & \downarrow \\ \mathrm{Fun}(G,H) \times (G \curvearrowright G) & \longrightarrow & H \end{array}$$

where the vertical arrows are the evident projections, the top horizontal arrow is given by

$$(\phi_0, h_0, g_0) \mapsto (\phi_0(g_0)h_0),$$
$$(\phi_0 \xrightarrow{h} \phi_1, h_0 \xrightarrow{h} h_1, g_0 \xrightarrow{g} g_1) \mapsto (\phi_0(g_0)h_0 \xrightarrow{h\phi_0(g)=\phi_1(g)h} \phi_1(g_1)h_1).$$

on objects and morphisms, and the bottom horizontal arrow is given by

$$(\phi_0, g_0) \mapsto *,$$
$$(\phi_0 \xrightarrow{h} \phi_1, g_0 \xrightarrow{g} g_1) \mapsto (* \xrightarrow{h\phi_0(g)=\phi_1(g)h} *).$$

The group G acts on the objects on the left-hand side of the square, where G acts on C as described above, by the tautological right action on $(G \curvearrowright G)$, and trivially on $\mathrm{Fun}(G,H)$. The horizontal arrows are invariant under this G–action.

Thus, taking geometric realizations of nerves and passing to quotients by G–actions, we obtain a commutative square

$$\begin{array}{ccc} (P \times EG)/G & \longrightarrow & EH \\ \downarrow & & \downarrow \\ (X \times EG)/G & \longrightarrow & BH \end{array}$$

which is evidently a map of H–bundles. Under the identification $(X \times EG)/G \approx X \times BG \approx N(\mathrm{Fun}(G, H) \times G)$, we see that the bottom arrow is isomorphic to that obtained from the evaluation functor $\mathrm{Fun}(G, H) \times G \to H$, and thus is adjoint to the map $\rho\colon X \to \mathrm{Map}(BG, BH)$ described earlier.

7 The space of homomorphisms between compact Lie groups

Recall that $\mathrm{Hom}(G, H)$ denotes the space of homomorphisms equipped with the compact-open topology. We give a proof of the following fact, which is standard but not easily read from the literature with which the author is familiar.

7.1 Proposition *Let G and H be Lie groups, with G compact. The map*

$$(\phi, hC(\phi)) \mapsto h\phi h^{-1}\colon \coprod_{[\phi]} H/C(\phi) \to \mathrm{Hom}(G, H),$$

where $C(\phi) = \{h \in H \mid \phi(g)h = h\phi(g) \text{ for all } g \in G\}$ and $[\phi]$ runs over a set of H–conjugacy classes in $\mathrm{Hom}(G, H)$, is a homeomorphism. In particular, $\mathrm{Hom}(G, H)$ is locally compact, and thus a CGWH space.

Proof We quote a classical theorem of Montgomery and Zippin [11, page 216]: for every compact subgroup K of a Lie group L, there exists a neighborhood U of K such that every closed subgroup of L in U is L–conjugate to a subgroup of K. Applied to $L = G \times H$ and $K = \Lambda_\phi = \{(g, \phi(g)) \mid g \in G\}$, the graph of a continuous homomorphism $\phi\colon G \to H$, we obtain a neighborhood $U \subseteq G \times H$ of Λ_ϕ such that if $\Lambda_{\phi'} \in U$ for $\phi' \in \mathrm{Hom}(G, H)$, then ϕ' is H–conjugate to ϕ; see [2, Lemma 38.1].

There exists a neighborhood V of $e \in H$ such that $\Lambda_\phi \subseteq \{(g, h) \mid h\phi(g)^{-1} \in V\} \subseteq U$. To see this, use the homeomorphism $\alpha\colon G \times H \to G \times H$, $\alpha(g, h) = (g, h\phi(g)^{-1})$, together with the tube lemma applied to $G \times \{e\} \subseteq \alpha(U)$.

By definition, the set $V' := \{f\colon G \to H \mid f(G) \subseteq V\}$ is an open subset of $C(G, H)$, the space of continuous maps $G \to H$ equipped with the compact-open topology. The space $C(G, H)$ is a topological group under pointwise multiplication in H; to prove this, use the fact that G, H, and finite products thereof are locally compact, so the relevant evaluation maps are continuous. Therefore, the translated subset $V'\phi$ is open in $C(G, H)$. Tracing through the definitions, we see that any continuous homomorphism $G \to H$ in $V'\phi$ must be conjugate to ϕ.

Thus, we have shown that conjugacy classes are open subsets of $\mathrm{Hom}(G, H)$.

Now consider the map of the proposition. Each $H/C(\phi)$ maps bijectively to a conjugacy class in $\mathrm{Hom}(G, H)$. As H is Hausdorff, so is $C(G, H)$ and hence so is the subspace $\mathrm{Hom}(G, H)$. Therefore, each $H/C(\phi) \to \mathrm{Hom}(G, H)$ gives a homeomorphism to its image, since $H/C(\phi)$ is compact. Because the image is also open, the homeomorphism of the proposition follows.

As an immediate consequence, we see that $\mathrm{Hom}(G, H)$ is a coproduct of compact Hausdorff spaces, and thus locally compact. □

References

[1] **R Brown**, *Groupoids and crossed objects in algebraic topology*, Homology Homotopy Appl. 1 (1999) 1–78 MR

[2] **P E Conner**, **E E Floyd**, *Differentiable periodic maps*, Ergeb. Math. Grenzgeb. 33, Springer (1964) MR

[3] **W Dwyer**, **A Zabrodsky**, *Maps between classifying spaces*, from "Algebraic topology" (J Aguadé, R Kane, editors), Lecture Notes in Math. 1298, Springer (1987) 106–119 MR

[4] **B Guillou**, **P May**, **M Merling**, *Categorical models for equivariant classifying spaces*, Algebr. Geom. Topol. 17 (2017) 2565–2602 MR

[5] **R K Lashof**, *Equivariant bundles*, Illinois J. Math. 26 (1982) 257–271 MR

[6] **R K Lashof**, **J P May**, *Generalized equivariant bundles*, Bull. Soc. Math. Belg. Sér. A 38 (1986) 265–271 MR

[7] **R K Lashof**, **J P May**, **G B Segal**, *Equivariant bundles with abelian structural group*, from "Proceedings of the Northwestern Homotopy Theory Conference" (H R Miller, S B Priddy, editors), Contemp. Math. 19, Amer. Math. Soc., Providence, RI (1983) 167–176 MR

[8] **W Lück**, **B Uribe**, *Equivariant principal bundles and their classifying spaces*, Algebr. Geom. Topol. 14 (2014) 1925–1995 MR

[9] **J P May**, *Classifying spaces and fibrations*, Mem. Amer. Math. Soc. 155, Amer. Math. Soc., Providence, RI (1975) MR

[10] **J P May**, *Some remarks on equivariant bundles and classifying spaces*, from "International Conference on Homotopy Theory", Astérisque 191, Soc. Math. France, Paris (1990) 239–253 MR

[11] **D Montgomery**, **L Zippin**, *Topological transformation groups*, Interscience, New York (1955) MR

[12] **D Notbohm**, *Maps between classifying spaces*, Math. Z. 207 (1991) 153–168 MR

Department of Mathematics, University of Illinois
Urbana, IL, United States

rezk@illinois.edu

http://www.math.illinois.edu/~rezk

Received: 2 February 2017 Revised: 30 June 2017

msp

Algebraic & Geometric Topology 18 (2018) 547–568

Second mod 2 homology of Artin groups

TOSHIYUKI AKITA
YE LIU

In this paper, we compute the second mod 2 homology of an arbitrary Artin group, without assuming the $K(\pi,1)$ conjecture. The key ingredients are (A) Hopf's formula for the second integral homology of a group and (B) Howlett's result on the second integral homology of Coxeter groups.

20F36, 20J06; 20F55

1 Introduction

An Artin group (or an Artin–Tits group) is a finitely presented group with at most one simple relation between a pair of generators. Examples includes finitely generated free abelian groups, free groups of finite rank, Artin's braid groups with finitely many strands, right-angled Artin groups, etc. Artin groups appear in diverse branches of mathematics such as singularity theory, low-dimensional topology, geometric group theory, the theory of hyperplane arrangements, etc.

Artin groups are closely related to Coxeter groups. For a Coxeter graph Γ and the corresponding Coxeter system $(W(\Gamma), S)$, we associate an Artin group $A(\Gamma)$ obtained by, informally speaking, dropping the relations that each generator has order 2 from the standard presentation of $W(\Gamma)$. The symmetric group $\mathfrak{S}_n$ is the Coxeter group associated to the Coxeter graph of type A_{n-1}, and the braid group $\mathrm{Br}(n)$ is the corresponding Artin group. The Coxeter group $W(\Gamma)$ can be realized as a reflection group acting on a convex cone U (called Tits cone) in $\mathbb{R}^n$ with $n = \#S$ the rank of W. Let $\mathcal{A}$ be the collection of reflection hyperplanes. The complement

$$M(\Gamma) = \big(\mathrm{int}(U) + \sqrt{-1}\mathbb{R}\big) \setminus \bigcup_{H \in \mathcal{A}} H \otimes \mathbb{C}$$

admits the free $W(\Gamma)$–action, and the resulting orbit space $N(\Gamma) = M(\Gamma)/W(\Gamma)$ has the fundamental group isomorphic to $A(\Gamma)$; see van der Lek [20]. The celebrated $K(\pi,1)$ conjecture states that $N(\Gamma)$ is a $K(A(\Gamma),1)$ space. See Section 2.3 for a list of Γ for which the $K(\pi,1)$ conjecture is proved.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.547

Existing results about (co)homology of Artin groups all focus on particular types Artin groups for which the $K(\pi,1)$ conjecture has been proved. There are very few properties that can be said for (co)homology of *all* Artin groups (except for their first integral homology, which is simply the abelianization). In this paper, we compute the second mod 2 homology of *all* Artin groups without assuming an affirmative solution of the $K(\pi,1)$ conjecture. Our main tools are Hopf's formula on the second homology (or the Schur multiplier) of groups, together with Howlett's theorem (Theorem 3.2) on the second integral homology of Coxeter groups. We are inspired by Korkmaz and Stipsicz [19], who computed the second integral homology of the mapping class groups of oriented surfaces using Hopf's formula.

Our main result is the following.

Theorem 1.1 *Let $A(\Gamma)$ be the Artin group associated to a Coxeter graph Γ. Then*

$$H_2(A(\Gamma);\mathbb{Z}_2) \cong \mathbb{Z}_2^{p(\Gamma)+q(\Gamma)},$$

where $p(\Gamma)$ and $q(\Gamma)$ are nonnegative integers associated to Γ (see Theorem 2.6 for definitions).

As a corollary, we obtain a sufficient condition that the classifying map $c\colon N(\Gamma)\to K(A(\Gamma),1)$ induces an isomorphism

$$c_*\colon H_2(N(\Gamma);\mathbb{Z})\to H_2(A(\Gamma);\mathbb{Z}).$$

Furthermore, we conclude that the induced homomorphism

$$c_*\otimes \mathrm{id}_{\mathbb{Z}_2}\colon H_2(N(\Gamma);\mathbb{Z})\otimes\mathbb{Z}_2\to H_2(A(\Gamma);\mathbb{Z})\otimes\mathbb{Z}_2$$

is always an isomorphism. This gives affirmative evidence for the $K(\pi,1)$ conjecture.

A part of this paper is based on the second author's PhD thesis.

2 Preliminaries

We collect relevant definitions and properties of Coxeter groups and Artin groups. We refer to [2; 17] for Coxeter groups and [23; 24; 25] for Artin groups.

2.1 Coxeter groups

Let S be a finite set. A *Coxeter matrix* over S is a symmetric matrix $M=(m(s,t))_{s,t\in S}$ such that $m(s,s)=1$ for all $s\in S$, and $m(s,t)=m(t,s)\in\{2,3,\cdots\}\cup\{\infty\}$ for distinct $s,t\in S$. It is convenient to represent M by a labeled graph Γ, called the *Coxeter graph* of M, defined as follows:

- the vertex set is $V(\Gamma) = S$;
- the edge set is $E(\Gamma) = \{\{s,t\} \subset S \mid m(s,t) \geq 3\}$;
- the edge $\{s,t\}$ is labeled by $m(s,t)$ if $m(s,t) \geq 4$.

Let Γ_{odd} be the subgraph of Γ with $V(\Gamma_{\text{odd}}) = V(\Gamma)$ and $E(\Gamma_{\text{odd}}) = \{\{s,t\} \in E(\Gamma) \mid m(s,t) \text{ odd}\}$ inheriting labels from Γ. By abuse of notation, we frequently regard Γ (hence also Γ_{odd}) as its underlying 1–dimensional CW-complex.

Definition 2.1 Let Γ be a Coxeter graph and S its vertex set. The *Coxeter system* associated to Γ is the pair $(W(\Gamma), S)$, where the *Coxeter group* $W(\Gamma)$ is defined by the standard presentation

$$W(\Gamma) = \langle S \mid (st)^{m(s,t)} = 1 \text{ for all } s,t \in S \text{ with } m(s,t) \neq \infty \rangle.$$

Each generator $s \in S$ of W has order 2. For distinct $s,t \in S$, the order of st is precisely $m(s,t)$ if $m(s,t) \neq \infty$. In the case where $m(s,t) = \infty$, the element st has infinite order.

However, in this paper, we adopt an equivalent definition. For two letters s,t and an integer $m \geq 2$, we shall use the following notation of the word of length m consisting of s and t in an alternating order:

$$(st)_m := \overbrace{sts\cdots}^{m}.$$

For example, $(st)_2 = st$, $(st)_3 = sts$ and $(st)_4 = ststs$.

Definition 2.2 Let Γ be a Coxeter graph and S its vertex set. The *Coxeter group* associated to Γ is the group defined by the presentation

$$W(\Gamma) = \langle S \mid \overline{R_W} \cup Q_W \rangle.$$

The sets of relations are $\overline{R_W} = \{R(s,t) \mid m(s,t) < \infty\}$ and $Q_W = \{Q(s) \mid s \in S\}$, where $R(s,t) := (st)_{m(s,t)}(ts)^{-1}_{m(s,t)}$ and $Q(s) := s^2$.

Note that since $R(s,t) = R(t,s)^{-1}$, we may reduce the relation set $\overline{R_W}$ by introducing a total order on S and put $R_W := \{R(s,t) \mid m(s,t) < \infty, s < t\}$. We have the following presentation with fewer relations:

$$W(\Gamma) = \langle S \mid R_W \cup Q_W \rangle.$$

We shall omit the reference to Γ if there is no ambiguity. The rank of W is defined to be $\#S$.

Let (W,S) be a Coxeter system. For a subset $T \subset S$, let W_T denote the subgroup of W generated by T, called a *parabolic subgroup* of W. In particular, $W_S = W$ and $W_\varnothing = \{1\}$. It is known that (W_T, T) is the Coxeter system associated to the Coxeter graph Γ_T (the full subgraph of Γ spanned by T inheriting labels); see Théorème 2 in Chapter IV of [2].

2.2 Artin groups

The Artin group $A(\Gamma)$ associated to a Coxeter graph Γ is obtained from the presentation of $W(\Gamma)$ by dropping the relation set Q_W.

Definition 2.3 Given a Coxeter graph Γ (hence a Coxeter system (W,S)), we introduce a set $\Sigma = \{a_s \mid s \in S\}$ in one-to-one correspondence with S. Then the Artin system associated to Γ is the pair $(A(\Gamma), \Sigma)$, where $A(\Gamma)$ is the *Artin group* of type Γ defined by the presentation

$$A(\Gamma) = \langle \Sigma \mid \overline{R_A} \rangle,$$

where $\overline{R_A} = \{R(a_s, a_t) \mid m(s,t) < \infty\}$ and $R(a_s, a_t) = (a_s a_t)_{m(s,t)} (a_t a_s)^{-1}_{m(s,t)}$.

As in the Coxeter group case, we introduce a total order on S and put $R_A := \{R(a_s, a_t) \mid m(s,t) < \infty, s < t\}$. We have the following presentation with fewer relations:

$$A(\Gamma) = \langle \Sigma \mid R_A \rangle.$$

There is a canonical projection $p\colon A(\Gamma) \to W(\Gamma)$, $a_s \mapsto s$ $(s \in S)$, whose kernel is called the *pure Artin group* of type Γ.

We say that an Artin group $A(\Gamma)$ is of finite type (or spherical type) if the associated Coxeter group $W(\Gamma)$ is finite; otherwise, $A(\Gamma)$ is of infinite type (or nonspherical type).

2.3 $K(\pi, 1)$ conjecture

Consider a Coxeter graph Γ and the associated Coxeter system (W,S) with rank $\#S = n$. Recall that W can be realized as a reflection group acting on a Tits cone $U \subset \mathbb{R}^n$; see [24]. Let $\mathcal{A}$ be the collection of the reflection hyperplanes. Put

$$M(\Gamma) := \big(\mathrm{int}(U) + \sqrt{-1}\mathbb{R}^n\big) \setminus \bigcup_{H \in \mathcal{A}} H \otimes \mathbb{C}.$$

Then W acts on $M(\Gamma)$ freely and properly discontinuously. Denote the orbit space by

$$N(\Gamma) := M(\Gamma)/W. \tag{2-1}$$

It is known that:

Theorem 2.4 [20] *The fundamental group of $N(\Gamma)$ is isomorphic to the Artin group $A(\Gamma)$.*

In general, $N(\Gamma)$ is only conjectured to be a classifying space of $A(\Gamma)$.

Conjecture 2.5 *Let Γ be an arbitrary Coxeter graph. Then the orbit space $N(\Gamma)$ is a $K(\pi,1)$ space, hence a classifying space of the Artin group $A(\Gamma)$.*

This conjecture is proved to hold for a few classes of Artin groups. Here is a list of such classes known so far:

- Artin groups of finite type [12].
- Artin groups of large type [14].
- 2–dimensional Artin groups [7].
- Artin groups of FC type [7].
- Artin groups of affine types $\widetilde{A}_n$, $\widetilde{C}_n$ [21].
- Artin groups of affine type $\widetilde{B}_n$ [5].
- Artin groups $A(\Gamma)$ such that the $K(\pi,1)$ conjecture holds for all $A(\Gamma_T)$ where $T \subset S$ and Γ_T does not contain ∞–labeled edges [13].
- Artin groups whose corresponding Coxeter system (W,S) has the following property: every finite irreducible parabolic subgroup W_T is either $\mathfrak{S}_4$, $\mathbb{Z}_2$ or a dihedral group [6].

2.4 First and second homology of $N(\Gamma)$

Clancy and Ellis [8] computed the second integral homology of $N(\Gamma)$ using the Salvetti complex for an Artin group. We recall their result and follow their notation.

Let us first fix some notation. Let Γ be a Coxeter graph with vertex set S. Define $Q(\Gamma) = \{\{s,t\} \subset S \mid m(s,t) \text{ is even}\}$ and $P(\Gamma) = \{\{s,t\} \subset S \mid m(s,t) = 2\}$. Write $\{s,t\} \equiv \{s',t'\}$ if two such pairs in $P(\Gamma)$ satisfy $s = s'$ and $m(t,t')$ is odd. This generates an equivalence relation on $P(\Gamma)$, denoted by $\sim$. Let $P(\Gamma)/\sim$ be the set of equivalence classes. An equivalence class is called a *torsion* class if it is represented by a pair $\{s,t\} \in P(\Gamma)$ such that there exists a vertex $v \in S$ with $m(s,v) = m(t,v) = 3$. In the above situation, Clancy and Ellis proved the following theorem.

Theorem 2.6 [8] *Let Γ be a Coxeter graph and $N(\Gamma)$ as in* (2-1)*. Then*

$$H_2(N(\Gamma);\mathbb{Z}) \cong \mathbb{Z}_2^{p(\Gamma)} \oplus \mathbb{Z}^{q(\Gamma)},$$

where

$$\begin{aligned}
p(\Gamma) &:= \text{ number of torsion classes in } P(\Gamma)/\sim,\\
q_1(\Gamma) &:= \text{ number of nontorsion classes in } P(\Gamma)/\sim,\\
q_2(\Gamma) &:= \#(Q(\Gamma) - P(\Gamma)) = \#\{\{s,t\} \subset S \mid m(s,t) \geq 4,\ m(s,t) \text{ is even}\},\\
q_3(\Gamma) &:= \operatorname{rank} H_1(\Gamma_{\mathrm{odd}};\mathbb{Z}),\\
q(\Gamma) &:= q_1(\Gamma) + q_2(\Gamma) + q_3(\Gamma).
\end{aligned}$$

Remark Note that $H_1(N(\Gamma);\mathbb{Z}) \cong H_1(A(\Gamma);\mathbb{Z})$ is isomorphic to the abelianization of $A(\Gamma)$, which is a free abelian group with rank equal to rank $H_0(\Gamma_{\mathrm{odd}};\mathbb{Z})$, the number of connected components of Γ_{odd}.

3 Second mod 2 homology of Artin groups

The (co)homology of the orbit space $N(\Gamma)$ coincides with that of the Artin group $A(\Gamma)$, provided the $K(\pi,1)$ conjecture for $A(\Gamma)$ holds. There are many results about (co)homology of $N(\Gamma)$ in the literature, for example, [10; 9; 4; 5]. The $K(\pi,1)$ conjecture is known to hold in these cases.

In this section, nevertheless, we shall work on the second homology of *arbitrary* Artin groups without assuming that the $K(\pi,1)$ conjecture holds. Our main result is the following theorem.

Theorem 3.1 *Let Γ be an arbitrary Coxeter graph and $A(\Gamma)$ the associated Artin group. Then the second mod 2 homology of $A(\Gamma)$ is*

$$H_2(A(\Gamma);\mathbb{Z}_2) \cong \mathbb{Z}_2^{p(\Gamma)+q(\Gamma)},$$

where $p(\Gamma)$ and $q(\Gamma)$ are as in Theorem 2.6.

The outline of our proof is as follows. In Section 3.1, we state Howlett's theorem on the second integral homology group $H_2(W(\Gamma);\mathbb{Z})$ of the Coxeter group $W(\Gamma)$. Next, in Section 3.2, we recall Hopf's formula of the second homology of a group. The key of the proof is that, by virtue of Hopf's formula, we are able to find explicitly a set $\Omega(W)$ of generators of $H_2(W(\Gamma);\mathbb{Z})$ (Section 3.3), as well as a set $\Omega(A)$

of generators of $H_2(A(\Gamma);\mathbb{Z})$ (Section 3.4). On the other hand, Howlett's theorem implies that $\Omega(W)$ forms a basis of $H_2(W(\Gamma);\mathbb{Z})$, which is an elementary abelian 2–group of rank $p(\Gamma)+q(\Gamma)$. Furthermore, we will show that the homomorphism $p_*\colon H_2(A(\Gamma);\mathbb{Z})\to H_2(W(\Gamma);\mathbb{Z})$ induced by the projection $p\colon A(\Gamma)\to W(\Gamma)$ maps $\Omega(A)$ onto $\Omega(W)$. Hence p_* is actually an epimorphism and becomes an isomorphism when tensored with $\mathbb{Z}_2$.

3.1 Howlett's theorem

As previously mentioned, we shall study the homomorphism $p_*\colon H_2(A(\Gamma);\mathbb{Z})\to H_2(W(\Gamma);\mathbb{Z})$ induced by the projection $p\colon A(\Gamma)\to W(\Gamma)$. A reason for doing so is that we have the following theorem of Howlett:

Theorem 3.2 [16] *The second integral homology of the Coxeter group $W(\Gamma)$ associated to a Coxeter graph Γ is*

$$H_2(W(\Gamma);\mathbb{Z})\cong\mathbb{Z}_2^{p(\Gamma)+q(\Gamma)},$$

where $p(\Gamma)$ and $q(\Gamma)$ are as in Theorem 2.6.

Remark The original statement in [16] was

$$H_2(W(\Gamma);\mathbb{Z})\cong\mathbb{Z}_2^{-n_1(\Gamma)+n_2(\Gamma)+n_3(\Gamma)+n_4(\Gamma)},$$

where

$$\begin{aligned}
n_1(\Gamma)&:=\#S,\\
n_2(\Gamma)&:=\#\{\{s,t\}\in E(\Gamma)\mid m(s,t)<\infty\},\\
n_3(\Gamma)&:=\#P(\Gamma)/\sim,\\
n_4(\Gamma)&:=\operatorname{rank} H_0(\Gamma_{\mathrm{odd}};\mathbb{Z}).
\end{aligned}$$

For a Coxeter graph Γ, the above numbers are related to those used by Clancy–Ellis as follows:

$$-n_1(\Gamma)+n_2(\Gamma)+n_3(\Gamma)+n_4(\Gamma)=p(\Gamma)+q(\Gamma).$$

In fact, $n_1(\Gamma)=\#V(\Gamma_{\mathrm{odd}})$, $n_2(\Gamma)=q_2(\Gamma)+\#E(\Gamma_{\mathrm{odd}})$ and $n_3(\Gamma)=p(\Gamma)+q_1(\Gamma)$. The above equation follows from the Euler–Poincaré theorem applied to Γ_{odd}:

$$\#V(\Gamma_{\mathrm{odd}})-\#E(\Gamma_{\mathrm{odd}})=\operatorname{rank} H_0(\Gamma_{\mathrm{odd}};\mathbb{Z})-\operatorname{rank} H_1(\Gamma_{\mathrm{odd}};\mathbb{Z}).$$

Example 3.3 We shall make use of the following example later.

Let $\Gamma = I_2(m)$. Thus $W(\Gamma) = D_{2m}$ is the dihedral group of order $2m$. Theorem 3.2 shows that

$$H_2(W(\Gamma);\mathbb{Z}) \cong \begin{cases} \mathbb{Z}_2 & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd.} \end{cases}$$

See Corollary 10.1.27 of [18] for a complete list of integral homology of dihedral groups.

3.2 Hopf's formula

Hopf's formula gives a description of the second integral homology of a group. We first recall some notation. For a group G, the commutator of $x, y \in G$ is the element $[x, y] = xyx^{-1}y^{-1}$. The commutator subgroup $[G, G]$ of G is the subgroup of G generated by all commutators. In general, for any subgroups H and K of G, we define $[H, K]$ as the subgroup of G generated by all $[h, k]$ for $h \in H$, $k \in K$.

Theorem 3.4 (Hopf's formula) *If a group G has a presentation $\langle S \mid R\rangle$, then*

$$H_2(G;\mathbb{Z}) \cong \frac{N \cap [F, F]}{[F, N]},$$

where $F = F(S)$ is the free group generated by S and $N = N(R)$ is the normal closure of R (the subgroup of F normally generated by the relation set R).

See Section II.5 of [3] for a topological proof. Moreover, Hopf's formula admits the following naturality; see Section II.6, Exercise 3(b) of [3].

Proposition 3.5 *Let $G = F/N = \langle S \mid R\rangle$ and $G' = F'/N' = \langle S' \mid R'\rangle$ as in Theorem 3.4. Suppose a homomorphism $\alpha\colon G \to G'$ lifts to $\widetilde{\alpha}\colon F \to F'$. Then the following diagram commutes:*

$$\begin{array}{ccc} H_2(G;\mathbb{Z}) & \xrightarrow[\cong]{} & N \cap [F, F]/[F, N] \\ {\scriptstyle H_2(\alpha)}\downarrow & & \downarrow{\scriptstyle \alpha_*} \\ H_2(G';\mathbb{Z}) & \xrightarrow[\cong]{} & N' \cap [F', F']/[F', N'] \end{array}$$

where α_* is induced by $\widetilde{\alpha}$.

For simplicity, we denote by $\langle x\rangle_G = x[F,N] \in F/[F,N]$ the coset of $[F,N]$ represented by $x \in F$ and $\langle x,y\rangle_G = [x,y][F,N] \in [F,F]/[N,F]$ for $x, y \in F$. Thanks to Hopf's formula, second homology classes of G can be regarded as $\langle x\rangle_G$ for $x \in N \cap [F,F]$.

To see what the representatives look like, we make the following simple observations, which we learned from [19].

Lemma 3.6 *The group $N/[F,N]$ is abelian.*

Proof Note that $N/[F,N]$ is a quotient group of $N/[N,N]$ and the latter is the abelianization of N. □

Thus we write the group $N/[F,N]$ additively. It is clear $\langle n\rangle_G = -\langle n^{-1}\rangle_G$ for $n \in N$.

Lemma 3.7 *In the abelian group $N/[F,N]$, for $n \in N$ and $f \in F$, we have*

$$\langle n\rangle_G = \langle fnf^{-1}\rangle_G.$$

Proof Since $[f,n] \in [F,N]$, we have

$$\langle f,n\rangle_G = \langle fnf^{-1}n^{-1}\rangle_G = \langle fnf^{-1}\rangle_G - \langle n\rangle_G = 0.$$ □

Therefore, a coset in $N/[F,N]$ is represented by an element of the form $\prod_{r\in R} r^{n(r)}$ for $n(r) \in \mathbb{Z}$. Hopf's formula implies that a second homology class of G can be represented by an element $\prod_{r\in R} r^{n(r)} \in [F,F]$.

Lemma 3.8 *Let $G = F/N$ as in Theorem 3.4. If $x, y, z \in F$ such that $[x,y], [x,z] \in N \cap [F,F]$, then*

$$\langle x, yz\rangle_G = \langle x,y\rangle_G + \langle x,z\rangle_G, \quad \langle x, y^{-1}\rangle_G = -\langle x,y\rangle_G.$$

Proof Note that $[x,yz] = [x,y]y[x,z]y^{-1}$. Then in the abelian group $N/[F,N]$,

$$\langle x,yz\rangle_G = \langle x,y\rangle_G + \langle y[x,z]y^{-1}\rangle_G.$$

For the last term, we have $\langle y[x,z]y^{-1}\rangle_G = \langle x,z\rangle_G$ since

$$[x,z]^{-1}y[x,z]y^{-1} = [[x,z]^{-1}, y] \in [N,F].$$

Hence the first equality holds. The second follows immediately from the first. □

3.3 Hopf's formula applied to Coxeter groups

The aim of this subsection is to construct an explicit set $\Omega(W)$ of generators of $H_2(W(\Gamma);\mathbb{Z})$. Combined with Howlett's theorem (Theorem 3.2), we show that $\Omega(W)$ is a basis of $H_2(W(\Gamma);\mathbb{Z})$.

Let us describe the construction of $\Omega(W)$. Let Γ be a Coxeter graph and (W,S) the associated Coxeter system with S totally ordered. Then $W=\langle S \mid R_W \cup Q_W\rangle$ is as in Definition 2.2. Let $F_W=F(S)$ be the free group on S and $N_W=N(R_W\cup Q_W)$ the normal closure of $R_W\cup Q_W$. Therefore, $W=F_W/N_W$. Using Hopf's formula, we identify $H_2(W;\mathbb{Z})\cong(N_W\cap[F_W,F_W])/[F_W,N_W]$. We shall construct three sets $\Omega_i(W)\subset(N_W\cap[F_W,F_W])/[F_W,N_W]$ $(i=1,2,3)$. In view of Lemma 3.6 and Lemma 3.7, a second homology class of W is of the form $\langle x\rangle_W$ with x expressed by a word $\prod_{R(s,t)\in R_W} R(s,t)^{n(s,t)}\prod_{Q(s)\in Q_W} Q(s)^{n(s)}\in[F_W,F_W]$. We decompose $\langle x\rangle_W=\langle x_1\rangle_W+\langle x_2\rangle_W+\langle x_3\rangle_W$, as in the proof of Theorem 3.15, such that $\langle x_i\rangle_W$ is generated by $\Omega_i(W)$. Then $\Omega(W)=\Omega_1(W)\cup\Omega_2(W)\cup\Omega_3(W)$ generates $H_2(W;\mathbb{Z})$. Now we exhibit respectively the constructions of $\Omega_i(W)$ $(i=1,2,3)$.

3.3.1 Construction of $\Omega_1(W)$

Let

$$\Omega_1(W)=\{\langle s,t\rangle_W \mid s,t\in S,\ s<t,\ m(s,t)=2\}.$$

Recall that $\langle s,t\rangle_W=[s,t][F_W,N_W]\in(N_W\cap[F_W,F_W])/[F_W,N_W]$ and $R(s,t)=[s,t]$ when $m(s,t)=2$. Note that the above expression may have repetitions. In fact, we have the following.

Proposition 3.9 $\#\Omega_1(W)\le p(\Gamma)+q_1(\Gamma)$.

Proof We shall show that $\langle s,t\rangle_W=\langle s,t'\rangle_W$ in $\Omega_1(W)$ if $\{s,t\}\equiv\{s,t'\}$ in $P(\Gamma)$. Suppose $s<t$ and $s<t'$ with $\{s,t\}\equiv\{s,t'\}$ in $P(\Gamma)$; that is, $m(s,t)=m(s,t')=2$ and $m(t,t')$ is odd. Then in $N_W/[F_W,N_W]$,

$$\langle s,R(t,t')\rangle_W=\langle s,(tt'\cdots t)(t't\cdots t')^{-1}\rangle_W=\langle s,t\rangle_W-\langle s,t'\rangle_W,$$

where the latter equality follows from Lemma 3.8. On the other hand,

$$\begin{aligned}\langle s,R(t,t')\rangle_W&=\langle sR(t,t')s^{-1}R(t,t')^{-1}\rangle_W\\&=\langle sR(t,t')s^{-1}\rangle_W+\langle R(t,t')^{-1}\rangle_W\\&=\langle R(t,t')\rangle_W-\langle R(t,t')\rangle_W=0,\end{aligned}$$

where the second equality follows from Lemma 3.8 and the third from Lemma 3.7. Thus we conclude $\langle s,t\rangle_W = \langle s,t'\rangle_W$. Similarly, $\langle s,t\rangle_W = \langle s',t\rangle_W$ in $\Omega_1(W)$ if $\{s,t\} \equiv \{s',t\}$ in $P(\Gamma)$. Hence $\#\Omega_1(W) \le \#(P(\Gamma)/\sim) = p(\Gamma) + q_1(\Gamma)$. □

3.3.2 Construction of $\Omega_2(W)$ Let

$$\Omega_2(W) = \{\langle R(s,t)\rangle_W \mid s,t \in S,\ s<t,\ m(s,t) \ge 4,\ m(s,t) \text{ is even}\}.$$

Recall that $R(s,t) = (st)_{m(s,t)}(ts)^{-1}_{m(s,t)}$. Note that when $m(s,t)$ is even, $R(s,t)$ is in the kernel of the abelianization map Ab: $F_W \to F_W/[F_W,F_W]$, and hence $R(s,t) \in [F_W,F_W]$. The following result is obvious.

Proposition 3.10 $\#\Omega_2(W) \le q_2(\Gamma)$.

3.3.3 Construction of $\Omega_3(W)$ The construction of $\Omega_3(W)$ requires more preparation. Recall that Γ_{odd} is the subgraph of Γ considered as a 1–dimensional CW-complex with 0–cells S and 1–cells $\{\langle s,t\rangle \mid s,t \in S,\ s<t,\ m(s,t) \text{ odd}\}$ oriented by $\partial\langle s,t\rangle = t-s$. We define a group

$$\mathcal{C}_W = \{(\alpha,\beta) \in C_1(\Gamma_{\text{odd}}) \oplus 2C_0(\Gamma_{\text{odd}}) \mid \partial\alpha = \beta\},$$

where $2C_0(\Gamma_{\text{odd}}) = \{2\gamma \mid \gamma \in C_0(\Gamma_{\text{odd}})\}$ is the group of 0–chains with all coefficients even, and $\mathcal{D}_W$ is the subgroup of $\mathcal{C}_W$ generated by $(2\langle s,t\rangle, -2s+2t)$ for all 1–cells $\langle s,t\rangle$.

Consider the homomorphism

$$\Phi_W\colon \mathcal{C}_W \to \frac{N_W \cap [F_W,F_W]}{[F_W,N_W]}$$

defined by

$$\Phi_W\Big(\sum_{\substack{s<t\\ m(s,t)\text{ odd}}} n(s,t)\langle s,t\rangle,\ \sum_{s\in S} 2n(s)s\Big) = \Big\langle \prod_{\substack{s<t\\ m(s,t)\text{ odd}}} R(s,t)^{n(s,t)} \prod_{s\in S} Q(s)^{n(s)} \Big\rangle_W.$$

The definition is indeed valid by the following easy lemma.

Lemma 3.11 *The following are equivalent:*

(A) $$\Big(\sum_{\substack{s<t\\ m(s,t)\text{ odd}}} n(s,t)\langle s,t\rangle,\ \sum_{s\in S} 2n(s)s\Big) \in \mathcal{C}_W,$$

(B) $$\prod_{\substack{s<t\\ m(s,t)\text{ odd}}} R(s,t)^{n(s,t)} \prod_{s\in S} Q(s)^{n(s)} \in [F_W,F_W].$$

Proof We suppress the ranges since they should be clear:

$$\begin{aligned}(\mathrm{A}) &\iff \partial\Big(\sum n(s,t)\langle s,t\rangle\Big) = \sum 2n(s)s \\ &\iff \sum 2n(s)s + \sum n(s,t)(s-t) = 0 \\ &\iff \mathrm{Ab}\Big(\prod R(s,t)^{n(s,t)} \prod Q(s)^{n(s)}\Big) = 0 \iff (\mathrm{B}),\end{aligned}$$

where Ab is the abelianization map as above, and we write $F_W/[F_W, F_W]$ additively. Note that $\mathrm{Ab}(R(s,t)) = s-t$ if $m(s,t)$ is odd. □

The following is a consequence of Example 3.3.

Proposition 3.12 *$\mathcal{D}_W$ lies in the kernel of Φ_W.*

Proof It suffices to show that any generator $(2\langle s,t\rangle, -2s+2t)$ of $\mathcal{D}_W$ is mapped to the identity by Φ_W, or equivalently, the word $((st)_m(ts)_m^{-1})^2(s^2)^{-1}t^2$ lies in $[F_W, N_W]$ when m is odd. Let $s,t \in S$ with $m := m(s,t)$ odd; consider the parabolic subgroup $W' := W_{\{s,t\}}$ of W, which is isomorphic to the dihedral group D_{2m} of order $2m$. From Example 3.3, we know $H_2(W';\mathbb{Z}) = 0$. On the other hand, Hopf's formula applied to W' shows that $H_2(W';\mathbb{Z}) \cong (N_{W'} \cap [F_{W'}, F_{W'}])/[F_{W'}, N_{W'}]$. Therefore, the word $((st)_m(ts)_m^{-1})^2(s^2)^{-1}t^2 \in N_{W'} \cap [F_{W'}, F_{W'}]$ represents the trivial homology class. That is to say,

$$((st)_m(ts)_m^{-1})^2(s^2)^{-1}t^2 \in [F_{W'}, N_{W'}] \subset [F_W, N_W]. \qquad \square$$

As a consequence, the homomorphism Φ_W factors through

$$\mathcal{C}_W \twoheadrightarrow \mathcal{C}_W/\mathcal{D}_W \to (N_W \cap [F_W, F_W])/[F_W, N_W].$$

Let $Z_1(\Gamma_{\mathrm{odd}};\mathbb{Z})$ denote the group of 1–cycles of Γ_{odd} with integral coefficients and $Z_1(\Gamma_{\mathrm{odd}};\mathbb{Z}_2)$ the group of 1–cycles of Γ_{odd} with coefficients in $\mathbb{Z}_2$. Define a homomorphism $\Xi_W\colon \mathcal{C}_W \to Z_1(\Gamma_{\mathrm{odd}};\mathbb{Z}_2)$ by $(\alpha, \partial\alpha) \mapsto \overline{\alpha}$, where $\alpha \in C_1(\Gamma_{\mathrm{odd}})$ such that $\partial\alpha \in 2C_0(\Gamma_{\mathrm{odd}})$ and $\overline{\alpha} \in C_1(\Gamma_{\mathrm{odd}};\mathbb{Z}_2)$ is the mod 2 reduction of α. The condition $\partial\alpha \in 2C_0(\Gamma_{\mathrm{odd}})$ asserts that $\overline{\alpha}$ is indeed a 1–cycle of Γ_{odd} with coefficients in $\mathbb{Z}_2$.

Proposition 3.13 *The homomorphism $\Xi_W\colon \mathcal{C}_W \to Z_1(\Gamma_{\mathrm{odd}};\mathbb{Z}_2)$ factors through an isomorphism:*

$$\begin{array}{ccc} \mathcal{C}_W & & \\ \downarrow & \searrow^{\Xi_W} & \\ \mathcal{C}_W/\mathcal{D}_W & \xrightarrow{\cong} & Z_1(\Gamma_{\mathrm{odd}};\mathbb{Z}_2) \end{array}$$

Proof The homomorphism Ξ_W is obviously an epimorphism, and

$$\begin{aligned}(\alpha, \partial\alpha) \in \operatorname{Ker} \Xi_W &\iff \overline{\alpha} = 0 \in Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2)\\ &\iff \alpha \in 2C_1(\Gamma_{\mathrm{odd}})\\ &\iff (\alpha, \partial\alpha) \in \mathcal{D}_W.\end{aligned}$$

Hence $\operatorname{Ker} \Xi_W = \mathcal{D}_W$. □

Via the isomorphism in Proposition 3.13, we obtain a homomorphism

$$\Psi_W\colon Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2) \to \frac{N_W \cap [F_W, F_W]}{[F_W, N_W]},$$

which fits into the following commutative diagram:

(3-1)
$$\begin{array}{ccc} \mathcal{C}_W & \xrightarrow{\Phi_W} & N_W \cap [F_W, F_W]/[F_W, N_W] \\ \downarrow & \nearrow & \uparrow \Psi_W \\ \mathcal{C}_W/\mathcal{D}_W & \xrightarrow{\cong} & Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2) \end{array}$$

We fix a basis $\Omega(\Gamma_{\mathrm{odd}}; \mathbb{Z})$ of $Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}) \cong \mathbb{Z}^{q_3(\Gamma)}$ once and for all and denote by $\Omega(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2)$ the basis of $Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2) \cong \mathbb{Z}_2^{q_3(\Gamma)}$ obtained from $\Omega(\Gamma_{\mathrm{odd}}; \mathbb{Z})$ by mod 2 reduction $Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}) \twoheadrightarrow Z_1(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2)$.

Define $\Omega_3(W)$ to be the image of $\Omega(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2)$ under Ψ_W:

$$\Omega_3(W) = \Psi_W(\Omega(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2)).$$

To be precise,

$$\Omega_3(W) = \left\{ \left\langle \prod_{\substack{s<t \\ m(s,t)\ \mathrm{odd}}} R(s,t)^{n(s,t)} \right\rangle_W \;\middle|\; \sum_{\substack{s<t \\ m(s,t)\ \mathrm{odd}}} n(s,t)\langle s,t\rangle \in \Omega(\Gamma_{\mathrm{odd}}; \mathbb{Z}_2) \right\}.$$

Proposition 3.14 $\#\Omega_3(W) \le q_3(\Gamma)$.

Let $\Omega(W) = \Omega_1(W) \cup \Omega_2(W) \cup \Omega_3(W)$. We conclude that:

Theorem 3.15 *$\Omega(W)$ is a basis of $H_2(W; \mathbb{Z})$.*

Proof Since $\#\Omega(W) \le p(\Gamma) + q(\Gamma)$ and $H_2(W; \mathbb{Z}) \cong \mathbb{Z}_2^{p(\Gamma)+q(\Gamma)}$ (Theorem 3.2), it suffices to show that $\Omega(W)$ generates $H_2(W; \mathbb{Z})$. An arbitrary homology class

in $H_2(W;\mathbb{Z})$ is represented by the coset $\langle x_1x_2x_3\rangle_W$ with

$$x_1 = \prod_{s<t,\ m(s,t)=2} [s,t]^{n(s,t)},$$
$$x_2 = \prod_{\substack{s<t,\ m(s,t)\geq 4\\ m(s,t)\text{ even}}} R(s,t)^{n(s,t)},$$
$$x_3 = \prod_{s<t,\ m(s,t)\text{ odd}} R(s,t)^{n(s,t)} \prod_{Q(s)\in Q_W} Q(s)^{n(s)},$$

with $x_1, x_2, x_3 \in [F_W, F_W]$. Thus $\langle x_1x_2x_3\rangle_W = \langle x_1\rangle_W + \langle x_2\rangle_W + \langle x_3\rangle_W$. We claim that $\langle x_i\rangle_W$ is generated by $\Omega_i(W)$. In fact, the claim for $i = 1, 2$ is straightforward. For $i = 3$, let

$$\alpha = \sum_{s<t,\ m(s,t)\text{ odd}} n(s,t)\langle s,t\rangle, \quad \beta = \sum_{s\in S} 2n(s)s.$$

Thus $(\alpha,\beta) \in \mathcal{C}_W$ by Lemma 3.11 with $\Phi_W(\alpha,\beta) = \langle x_3\rangle_W$. By the commutative diagram (3-1), the mod 2 reduction $\bar{\alpha} \in Z_1(\Gamma_{\text{odd}};\mathbb{Z}_2)$ of α is mapped to $\langle x_3\rangle_W$ by Ψ_W. This proves the claim. □

Remark It is worth noting that in the previous proof, we have managed to get rid of the relations $Q(s)$ without altering the homology class $\langle x_3\rangle_W$. This will be crucial in the proof of Theorem 3.20.

3.4 Hopf's formula applied to Artin groups

Now we turn to the Artin group case. The arguments here are parallel to those in the Coxeter group case.

Let Γ be a Coxeter graph with the vertex set S totally ordered and $A = A(\Gamma)$ the Artin group of type Γ with the presentation $A = \langle \Sigma \mid R_A\rangle$ given in Definition 2.3. Let $F_A = F(\Sigma)$ be the free group on Σ and N_A the normal closure of R_A. Hopf's formula yields $H_2(A;\mathbb{Z}) \cong (N_A \cap [F_A, F_A])/[F_A, N_A]$. For the same reason as before, a second homology class of A is represented by a coset $\langle x\rangle_A$ with x of the form $\prod_{R(a_s,a_t)\in R_A} R(a_s,a_t)^{n(s,t)} \in [F_A, F_A]$.

We construct a set $\Omega(A)$ of generators of $H_2(A;\mathbb{Z})$ using the same method as in the previous subsection.

3.4.1 Constructions of $\Omega_1(A)$ and $\Omega_2(A)$ The constructions of $\Omega_1(A)$ and $\Omega_2(A)$ are exactly parallel to those in the Coxeter case. Let

$$\Omega_1(A) = \{\langle a_s, a_t\rangle_A \mid s,t \in S,\ s<t,\ m(s,t)=2\},$$
$$\Omega_2(A) = \{\langle R(a_s,a_t)\rangle_A \mid s,t \in S,\ s<t,\ m(s,t)\geq 4,\ m(s,t) \text{ is even}\}.$$

The same reasoning shows that:

Proposition 3.16 $\#\Omega_1(A) \leq p(\Gamma)+q_1(\Gamma)$, $\#\Omega_2(A) \leq q_2(\Gamma)$.

3.4.2 Construction of $\Omega_3(A)$ Consider the homomorphism

$$\Psi_A\colon Z_1(\Gamma_{\text{odd}};\mathbb{Z}) \to \frac{N_A \cap [F_A,F_A]}{[F_A,N_A]}$$

defined by

$$\Psi_A\Bigg(\sum_{\substack{s<t\\ m(s,t)\text{ odd}}} n(s,t)\langle s,t\rangle\Bigg) = \Bigg\langle \prod_{\substack{s<t\\ m(s,t)\text{ odd}}} R(a_s,a_t)^{n(s,t)}\Bigg\rangle_A.$$

The definition is valid by the following lemma.

Lemma 3.17 *The following are equivalent:*

(A) $$\sum_{\substack{s<t\\ m(s,t)\text{ odd}}} n(s,t)\langle s,t\rangle \in Z_1(\Gamma_{\text{odd}};\mathbb{Z}),$$

(B) $$\prod_{\substack{s<t\\ m(s,t)\text{ odd}}} R(a_s,a_t)^{n(s,t)} \in [F_A,F_A].$$

Proof We suppress again the ranges:

$$\begin{aligned}\text{(A)} &\iff \partial\Big(\sum n(s,t)\langle s,t\rangle\Big)=0\\ &\iff \sum n(s,t)(t-s)=0\\ &\iff \mathrm{Ab}\Big(\prod R(a_s,a_t)^{n(s,t)}\Big)=0 \iff \text{(B)},\end{aligned}$$

where Ab: $F_A \to F_A/[F_A,F_A]$ is the abelianization map. Note that $\mathrm{Ab}(R(a_s,a_t)) = a_s - a_t$ if $m(s,t)$ is odd. □

Recall that we have chosen a basis $\Omega(\Gamma_{\mathrm{odd}};\mathbb{Z})$ for $Z_1(\Gamma_{\mathrm{odd}};\mathbb{Z})$. Let $\Omega_3(A)$ be the image of $\Omega(\Gamma_{\mathrm{odd}};\mathbb{Z})$ under Ψ_A:

$$\Omega_3(A) = \Psi_A(\Omega(\Gamma_{\mathrm{odd}};\mathbb{Z})).$$

To be precise,

$$\Omega_3(A) = \left\{ \left\langle \prod_{\substack{s<t \\ m(s,t)\ \mathrm{odd}}} R(a_s,a_t)^{n(s,t)} \right\rangle_A \;\middle|\; \sum_{\substack{s<t \\ m(s,t)\ \mathrm{odd}}} n(s,t)\langle s,t\rangle \in \Omega(\Gamma_{\mathrm{odd}};\mathbb{Z}) \right\}.$$

Proposition 3.18 $\#\Omega_3(A) \le q_3(\Gamma)$.

Let $\Omega(A) = \Omega_1(A) \cup \Omega_2(A) \cup \Omega_3(A)$; hence $\#\Omega(A) \le p(\Gamma) + q(\Gamma)$.

Theorem 3.19 *$\Omega(A)$ is a set of generators of $H_2(A;\mathbb{Z})$.*

Proof The proof is similar to that of Theorem 3.15, so we omit it. □

3.5 Proof of main results

Theorem 3.1 will follow from the next more precise theorem.

Theorem 3.20 *The projection $p\colon A(\Gamma) \to W(\Gamma)$ induces an epimorphism between the second integral homologies:*

$$p_*\colon H_2(A(\Gamma);\mathbb{Z}) \twoheadrightarrow H_2(W(\Gamma);\mathbb{Z}).$$

Proof The epimorphism $p\colon A \to W$ defined by $p(a_s) = s$ lifts to $\tilde{p}\colon F_A \to F_W$. Then by Proposition 3.5, we obtain the explicit formulation

$$p_*\colon H_2(A;\mathbb{Z}) \cong \frac{N_A \cap [F_A,F_A]}{[F_A,N_A]} \to \frac{N_W \cap [F_W,F_W]}{[F_W,N_W]} \cong H_2(W;\mathbb{Z}),$$

$$\left\langle \prod_{R(a_s,a_t)\in R_A} R(a_s,a_t)^{n(s,t)} \right\rangle_A \mapsto \left\langle \prod_{R(s,t)\in R_W} R(s,t)^{n(s,t)} \right\rangle_W.$$

We claim that p_* maps $\Omega_i(A)$ onto $\Omega_i(W)$. The claim is obvious for $i = 1, 2$. As for the case $i = 3$, consider the following diagram:

$$\begin{array}{ccc} \Omega(\Gamma_{\mathrm{odd}};\mathbb{Z}) & \xrightarrow{\Psi_A} & \Omega_3(A) \\ \big\downarrow{\scriptstyle \mathrm{mod}\ 2} & & \big\downarrow{\scriptstyle p_*} \\ \Omega(\Gamma_{\mathrm{odd}};\mathbb{Z}_2) & \xrightarrow{\Psi_W} & \Omega_3(W) \end{array}$$

Take $\alpha = \sum_{s<t,\ m(s,t)\text{ odd}} n(s,t)\langle s,t\rangle \in \Omega(\Gamma_{\text{odd}};\mathbb{Z})$; then

$$p_* \circ \Psi_A(\alpha) = \left\langle \prod_{\substack{s<t \\ m(s,t)\text{ odd}}} R(s,t)^{n(s,t)} \right\rangle_W \in \frac{N_W \cap [F_W, F_W]}{[F_W, N_W]}.$$

Recall the construction of Φ_W; we have $\Phi_W(\alpha, \partial\alpha) = p_* \circ \Psi_A(\alpha)$. Thus we obtain

$$\Psi_W(\overline{\alpha}) = p_* \circ \Psi_A(\alpha)$$

by the commutative diagram (3-1). This proves that p_* maps $\Omega_3(A)$ into $\Omega_3(W)$, and the above diagram commutes. Since the mod 2 reduction restricts to a bijection $\Omega(\Gamma_{\text{odd}};\mathbb{Z}) \to \Omega(\Gamma_{\text{odd}};\mathbb{Z}_2)$, and since the horizontal maps are onto by definition, $p_*\colon \Omega_3(A) \to \Omega_3(W)$ is onto. The proof is complete. □

Proof of Theorem 3.1 Consider the composition of epimorphisms

$$\mathbb{Z}^{\Omega(A)} \xrightarrow{\phi} H_2(A;\mathbb{Z}) \xrightarrow{p_*} H_2(W;\mathbb{Z}), \tag{3-2}$$

where $\mathbb{Z}^{\Omega(A)}$ is the free abelian group generated by $\Omega(A)$ and $\phi(\omega) = \omega$ for $\omega \in \Omega(A)$. Taking the tensor product with $\mathbb{Z}_2$ for terms in the above sequence (3-2), we have

$$\mathbb{Z}_2^{\Omega(A)} \xrightarrow{\phi\otimes \mathrm{id}_{\mathbb{Z}_2}} H_2(A;\mathbb{Z}) \otimes \mathbb{Z}_2 \xrightarrow{p_*\otimes \mathrm{id}_{\mathbb{Z}_2}} H_2(W;\mathbb{Z}) \otimes \mathbb{Z}_2,$$

where $\mathbb{Z}_2^{\Omega(A)}$ is the elementary abelian 2–group generated by $\Omega(A)$ with rank $\#\Omega(A) \le p(\Gamma) + q(\Gamma)$ and $H_2(W;\mathbb{Z}) \otimes \mathbb{Z}_2 \cong \mathbb{Z}_2^{p(\Gamma)+q(\Gamma)}$ (Theorem 3.2). Since tensoring with $\mathbb{Z}_2$ preserves surjectivity, this forces $\#\Omega(A) = p(\Gamma) + q(\Gamma)$, and both maps are in fact isomorphisms. Thus $H_2(A(\Gamma);\mathbb{Z}) \otimes \mathbb{Z}_2 \cong \mathbb{Z}_2^{p(\Gamma)+q(\Gamma)}$. On the other hand, we have the following exact sequence by the universal coefficient theorem:

$$0 \to H_2(A;\mathbb{Z}) \otimes \mathbb{Z}_2 \to H_2(A;\mathbb{Z}_2) \to \mathrm{Tor}(H_1(A;\mathbb{Z}), \mathbb{Z}_2) \to 0,$$

where $\mathrm{Tor}(H_1(A(\Gamma);\mathbb{Z}), \mathbb{Z}_2) = 0$ since $H_1(A(\Gamma);\mathbb{Z})$ is torsion free (Theorem 2.6 and the Remark following). Now we conclude $H_2(A(\Gamma);\mathbb{Z}_2) \cong \mathbb{Z}_2^{p(\Gamma)+q(\Gamma)}$ and finish the proof of Theorem 3.1. □

As a byproduct of the proof, we have the following corollaries. Recall that $M(\Gamma)$ is the complement of the complexified arrangement of reflection hyperplanes associated to the Coxeter group $W(\Gamma)$. The orbit space $N(\Gamma) = M(\Gamma)/W(\Gamma)$ has fundamental

group $\pi_1(N(\Gamma)) \cong A(\Gamma)$. Let $c\colon N(\Gamma) \to K(A(\Gamma), 1)$ be the classifying map. Then c always induces an isomorphism $c_*\colon H_1(N(\Gamma); \mathbb{Z}) \to H_1(A(\Gamma); \mathbb{Z})$ and an epimorphism $c_*\colon H_2(N(\Gamma); \mathbb{Z}) \to H_2(A(\Gamma); \mathbb{Z})$. We give a sufficient condition on Γ such that c induces an isomorphism $c_*\colon H_2(N(\Gamma); \mathbb{Z}) \to H_2(A(\Gamma); \mathbb{Z})$.

Corollary 3.21 *If Γ is such that*

- *$P(\Gamma)/\sim$ consists of torsion classes,*
- *$\Gamma = \Gamma_{\text{odd}}$,*
- *Γ is a tree,*

then

$$H_2(A(\Gamma); \mathbb{Z}) \cong \mathbb{Z}_2^{p(\Gamma)}.$$

Hence c induces an isomorphism $c_*\colon H_2(N(\Gamma); \mathbb{Z}) \to H_2(A(\Gamma); \mathbb{Z})$.

Proof Since $N(\Gamma)$ is path-connected and has fundamental group $\pi_1(N(\Gamma)) \cong A(\Gamma)$, there is an exact sequence (see for example, Section II.5, Theorem 5.2 of [3])

$$\pi_2(N(\Gamma)) \xrightarrow{h_2} H_2(N(\Gamma); \mathbb{Z}) \xrightarrow{c_*} H_2(A(\Gamma); \mathbb{Z}) \to 0, \tag{3-3}$$

where h_2 is the Hurewicz homomorphism. Suppose that Γ satisfies the three conditions; then $q_1(\Gamma) = q_2(\Gamma) = q_3(\Gamma) = 0$. Theorem 2.6 implies that $H_2(N(\Gamma); \mathbb{Z}) \cong \mathbb{Z}_2^{p(\Gamma)}$. Then by Theorem 3.20, $H_2(A(\Gamma); \mathbb{Z})$ sits in the sequence

$$\mathbb{Z}_2^{p(\Gamma)} \twoheadrightarrow H_2(A(\Gamma); \mathbb{Z}) \twoheadrightarrow \mathbb{Z}_2^{p(\Gamma)}.$$

The composition must be an isomorphism; hence $H_2(A(\Gamma); \mathbb{Z}) \cong \mathbb{Z}_2^{p(\Gamma)}$. As a result, c_* must be an isomorphism. □

Corollary 3.22 *If the three conditions in Corollary 3.21 are satisfied, then $p\colon A \to W$ induces an isomorphism*

$$p_*\colon H_2(A; \mathbb{Z}) \to H_2(W; \mathbb{Z}).$$

Proof This follows from Howlett's Theorem 3.2, Theorem 3.20 and Corollary 3.21. □

Corollary 3.23 *For any Coxeter graph Γ, the induced map $c_*\colon H_2(N(\Gamma); \mathbb{Z}) \to H_2(A(\Gamma); \mathbb{Z})$ becomes an isomorphism after tensoring with $\mathbb{Z}_2$.*

Proof By right-exactness of the tensor functor, taking the tensor product with $\mathbb{Z}_2$ preserves the exactness of (3-3):

$$\pi_2(N(\Gamma))\otimes\mathbb{Z}_2 \xrightarrow{h_2\otimes \mathrm{id}_{\mathbb{Z}_2}} H_2(N(\Gamma);\mathbb{Z})\otimes\mathbb{Z}_2 \xrightarrow{c_*\otimes \mathrm{id}_{\mathbb{Z}_2}} H_2(A(\Gamma);\mathbb{Z})\otimes\mathbb{Z}_2 \to 0.$$

Note that $c_*\otimes \mathrm{id}_{\mathbb{Z}_2}$ is an isomorphism as a consequence of Theorem 3.1 and Clancy–Ellis' Theorem 2.6. □

Example 3.24 The Coxeter graphs of affine type $\widetilde{D}_n$ $(n\geq 4)$ and $\widetilde{E}_i$ $(i=6,7,8)$ all satisfy the conditions in Corollary 3.21. Therefore, we compute the second integral homology of the associated Artin groups as follows:

$$H_2(A(\widetilde{D}_n);\mathbb{Z})\cong\begin{cases}\mathbb{Z}_2^6 & \text{for } n=4,\\ \mathbb{Z}_2^3 & \text{for } n\geq 5,\end{cases}\qquad H_2(A(\widetilde{E}_i);\mathbb{Z})\cong\mathbb{Z}_2 \quad \text{for } i=6,7,8.$$

Besides the above cases, the Coxeter graphs of certain hyperbolic Coxeter groups also provide plenty of examples satisfying the conditions in Corollary 3.21. We point out that to the best of the authors' knowledge, the $K(\pi,1)$ conjecture has not been proved in the above-mentioned cases.

3.6 Homological stability

We mention a corollary concerning homological stability in the end of this paper. Consider a family of Coxeter graphs $\{\Gamma_i\}_{i\geq 1}$ starting from Γ_1 with a base vertex s_1, and each Γ_i $(i\geq 2)$ is obtained by adding a vertex s_i connected to s_{i-1} by an unlabeled edge. The embedding $\Gamma_i\hookrightarrow\Gamma_{i+1}$ of Coxeter graphs induces an inclusion of Coxeter groups $W(\Gamma_i)\hookrightarrow W(\Gamma_{i+1})$, as well as an inclusion of Artin groups $A(\Gamma_i)\hookrightarrow A(\Gamma_{i+1})$; see [20; 22]. It is known that the families of Artin groups $\{A(A_n)\}$, $\{A(B_n)\}$ and $\{A(D_n)\}$ possess integral cohomological stability [1; 11]. Hepworth proved a more general result for Coxeter groups.

Theorem 3.25 [15] *The map $H_k(W(\Gamma_{n-1}))\to H_k(W(\Gamma_n))$ is an isomorphism for $2k\leq n$ with arbitrary constant coefficient.*

As for the sequence of Artin groups $\{A(\Gamma_i)\}$, it is not difficult to see that the first integral homology admits stability. We prove a stability result for the second mod 2 homology of the sequence $\{A(\Gamma_i)\}$.

Theorem 3.26 *The map $H_2(A(\Gamma_{n-1});\mathbb{Z}_2)\to H_2(A(\Gamma_n);\mathbb{Z}_2)$ is an isomorphism for $n\geq 4$.*

Proof Consider the commutative diagram

$$\begin{array}{ccc} H_2(A(\Gamma_{n-1});\mathbb{Z}_2) & \longrightarrow & H_2(A(\Gamma_n);\mathbb{Z}_2) \\ \uparrow & & \uparrow \\ H_2(A(\Gamma_{n-1});\mathbb{Z})\otimes\mathbb{Z}_2 & \longrightarrow & H_2(A(\Gamma_n);\mathbb{Z})\otimes\mathbb{Z}_2 \\ \downarrow{\scriptstyle p_*\otimes \mathrm{id}_{\mathbb{Z}_2}} & & \downarrow{\scriptstyle p_*\otimes \mathrm{id}_{\mathbb{Z}_2}} \\ H_2(W(\Gamma_{n-1});\mathbb{Z})\otimes\mathbb{Z}_2 & \longrightarrow & H_2(W(\Gamma_n);\mathbb{Z})\otimes\mathbb{Z}_2 \end{array} \tag{3-4}$$

where the commutativity of the upper square follows from the naturality of the universal coefficient theorem and that of the lower from tensoring the following commutative diagram with $\mathbb{Z}_2$:

$$\begin{array}{ccc} H_2(A(\Gamma_{n-1});\mathbb{Z}) & \longrightarrow & H_2(A(\Gamma_n);\mathbb{Z}) \\ \downarrow{\scriptstyle p_*} & & \downarrow{\scriptstyle p_*} \\ H_2(W(\Gamma_{n-1});\mathbb{Z}) & \longrightarrow & H_2(W(\Gamma_n);\mathbb{Z}) \end{array} \tag{3-5}$$

Since all vertical maps in (3-4) are isomorphisms and the bottom horizontal map is an isomorphism when $n \geq 4$ (Theorem 3.25), the top horizontal map is an isomorphism when $n \geq 4$. □

Corollary 3.27 *If Γ_n satisfies the three conditions in Corollary 3.21, then the map $H_2(A(\Gamma_{n-1});\mathbb{Z}) \to H_2(A(\Gamma_n);\mathbb{Z})$ is an isomorphism for $n \geq 4$.*

Proof This follows from Corollary 3.22, Theorem 3.25 and the diagram (3-5). □

Acknowledgement Akita was partially supported by JSPS KAKENHI Grant Number 26400077. Liu was supported by JSPS KAKENHI Grant Number 16J00125. They also thank the editor and the referee for valuable suggestions.

References

[1] **V I Arnol'd**, *The cohomology ring of the colored braid group*, Mat. Zametki 5 (1969) 227–231 MR In Russian; translated in Mathematical Notes 5 (1969), 138–140

[2] **N Bourbaki**, *Éléments de mathématique, XXXIV: Groupes et algèbres de Lie, Chapitres 4–6*, Actualités Scientifiques et Industrielles 1337, Hermann, Paris (1968) MR

[3] **K S Brown**, *Cohomology of groups*, Graduate Texts in Mathematics 87, Springer (1982) MR

[4] **F Callegaro**, **D Moroni**, **M Salvetti**, *Cohomology of affine Artin groups and applications*, Trans. Amer. Math. Soc. 360 (2008) 4169–4188 MR

[5] **F Callegaro**, **D Moroni**, **M Salvetti**, *The $K(\pi,1)$ problem for the affine Artin group of type $\widetilde{B}_n$ and its cohomology*, J. Eur. Math. Soc. 12 (2010) 1–22 MR

[6] **R Charney**, *The Deligne complex for the four-strand braid group*, Trans. Amer. Math. Soc. 356 (2004) 3881–3897 MR

[7] **R Charney**, **M W Davis**, *The $K(\pi,1)$–problem for hyperplane complements associated to infinite reflection groups*, J. Amer. Math. Soc. 8 (1995) 597–627 MR

[8] **M Clancy**, **G Ellis**, *Homology of some Artin and twisted Artin groups*, J. K-Theory 6 (2010) 171–196 MR

[9] **C De Concini**, **C Procesi**, **M Salvetti**, *Arithmetic properties of the cohomology of braid groups*, Topology 40 (2001) 739–751 MR

[10] **C De Concini**, **C Procesi**, **M Salvetti**, **F Stumbo**, *Arithmetic properties of the cohomology of Artin groups*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 28 (1999) 695–717 MR

[11] **C De Concini**, **M Salvetti**, *Stability for the cohomology of Artin groups*, Adv. Math. 145 (1999) 291–305 MR

[12] **P Deligne**, *Les immeubles des groupes de tresses généralisés*, Invent. Math. 17 (1972) 273–302 MR

[13] **G Ellis**, **E Sköldberg**, *The $K(\pi,1)$ conjecture for a class of Artin groups*, Comment. Math. Helv. 85 (2010) 409–415 MR

[14] **H Hendriks**, *Hyperplane complements of large type*, Invent. Math. 79 (1985) 375–381 MR

[15] **R Hepworth**, *Homological stability for families of Coxeter groups*, Algebr. Geom. Topol. 16 (2016) 2779–2811 MR

[16] **R B Howlett**, *On the Schur multipliers of Coxeter groups*, J. London Math. Soc. 38 (1988) 263–276 MR

[17] **J E Humphreys**, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics 29, Cambridge Univ. Press (1990) MR

[18] **G Karpilovsky**, *Group representations, Vol 2*, North-Holland Mathematics Studies 177, North-Holland, Amsterdam (1993) MR

[19] **M Korkmaz**, **A I Stipsicz**, *The second homology groups of mapping class groups of oriented surfaces*, Math. Proc. Cambridge Philos. Soc. 134 (2003) 479–489 MR

[20] **H van der Lek**, *The homotopy type of complex hyperplane complements*, PhD thesis, University of Nijmegen (1983)

[21] **C Okonek**, *Das $K(\pi, 1)$–Problem für die affinen Wurzelsysteme vom Typ A_n, C_n*, Math. Z. 168 (1979) 143–148 MR

[22] **L Paris**, *Parabolic subgroups of Artin groups*, J. Algebra 196 (1997) 369–399 MR

[23] **L Paris**, *Braid groups and Artin groups*, from "Handbook of Teichmüller theory, Vol II" (A Papadopoulos, editor), IRMA Lect. Math. Theor. Phys. 13, Eur. Math. Soc., Zürich (2009) 389–451 MR

[24] **L Paris**, *$K(\pi, 1)$ conjecture for Artin groups*, Ann. Fac. Sci. Toulouse Math. 23 (2014) 361–415 MR

[25] **L Paris**, *Lectures on Artin groups and the $K(\pi, 1)$ conjecture*, from "Groups of exceptional type, Coxeter groups and related geometries" (N S N Sastry, editor), Springer Proc. Math. Stat. 82, Springer (2014) 239–257 MR

Department of Mathematics, Hokkaido University
Sapporo, 060-0810, Japan

akita@math.sci.hokudai.ac.jp, liu@math.sci.hokudai.ac.jp

Received: 19 February 2017 Revised: 28 June 2017

msp

Algebraic & Geometric Topology 18 (2018) 569–582

On the third homotopy group of Orr's space

EMMANUEL DROR FARJOUN
ROMAN MIKHAILOV

K Orr defined a Milnor-type invariant of links that lies in the third homotopy group of a certain space K_ω. The problem of nontriviality of this third homotopy group has been open. We show that it is an infinitely generated group. The question of realization of its elements as links remains open.

55Q52, 57M25

1 Introduction

In [12] John Milnor defines his $\overline{\mu}$–invariants of links, which extract numerical values from the lower central series quotients of the link group and the homotopy classes of the link longitudes. His questions motivated generations of research concerning these invariants. While most of Milnor's original questions have been answered (see, for instance, K Orr [13; 14], Igusa and Orr [9], Cochran [5], Habegger and Lin [8]), one key question remains open. Milnor asked if we can extract similar invariants from π/π_ω. The first candidate for a transfinite Milnor invariant was given by Orr in [13], where he suggested a possible domain for these invariants, a space he denoted K_ω. A transfinite Milnor invariant is then an element in the third homotopy group, π_3, of this space. Extensions and refinements were introduced by J P Levine in [11], as mentioned toward the end of this introduction.

The space defined by Orr, denoted by K_ω, is the mapping cone of the natural map

$$K(F,1) \to K(\widehat{F},1).$$

Here $\widehat{F}$ is the pronilpotent completion of F, ie $\widehat{F} = \lim F/\gamma_i(F)$, where $\{\gamma_i(F)\}_{i\geq 1}$ is the lower central series of F. Let $L \subset S^3$ be a link, $G = \pi_1(S^3 \setminus L)$ its group and $f\colon F \to G$ a meridian homomorphism, where F is the free group with rank equal to the number of components of L. Assume that all Milnor's $\overline{\mu}$–invariants of L are zero. This means in fact that the homomorphism f induces isomorphisms of all finite lower central quotients and therefore an isomorphism $\widehat{F} \simeq \widehat{G}$. Thus we have a

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.569

map $S^3 \setminus L \to K(\widehat{F}, 1)$, which extends to a map $S^3 \to K_\omega$. The homotopy class of this map in $\pi_3(K_\omega)$ defines an invariant θ_ω; see Orr [14] for details and discussion. The main aim of the present note is to show that the third homotopy group $\pi_3(K_\omega)$ is nontrivial, in fact that it is infinitely generated.

The space K_ω is simply connected. This follows from the fact that, for a finitely generated group G, the pronilpotent completion $\widehat{G}$ is normally generated by the images of G under the natural injection $G \to \widehat{G}$; see Section 2.2 and observe that this is not true for a free group of infinite rank; see Bousfield [2]. For the case $G = F$, the free group with at least two generators, it is shown in [2] that $H_2(\widehat{F})$ is uncountable; hence the same is true for $\pi_2(K_\omega) \simeq H_2(K_\omega) \simeq H_2(\widehat{F})$. Cochran and Orr conjectured that the relevant third homotopy group is nonvanishing: Cochran [4] wrote: "Bousfield has shown that $\pi_2(K_\omega)$ is infinitely generated and it appears likely that the same will hold for all the homotopy groups of K_ω."

In light of Orr's Milnor-type invariant in $\pi_3(K_\omega)$ the question about nontriviality of higher homotopy groups of K_ω is of interest, and is open. The following is the main result of the present note.

Theorem 1.1 *The third homotopy group $\pi_3(K_\omega)$ is infinitely generated.*

Since we cannot see why $H_3(K_\omega)$ is not zero, we concentrate attention on the kernel of the Hurewicz map in dimension three in order to carry out the argument. In fact, we prove that the kernel of the (surjective) Hurewicz homomorphism $\pi_3(K_\omega) \to H_3(K_\omega)$ is infinitely generated.

Observe that the nontriviality of $\pi_3(K_\omega)$ does not imply the existence of links with nonzero invariants θ_ω. In order to solve that realization problem, another space K_∞ was defined by Levine [11]. The definition of K_∞ is similar to K_ω, the difference being that the algebraic closure is used instead of the pronilpotent completion. Any element from $\pi_3(K_\infty)$ can be realized as an invariant for a certain link; however, the problem of whether the Levine group $\pi_3(K_\infty)$ is nonzero is still very much open. The claim, without proof, in Orr's thesis [13] that all elements in his group $\pi_3(K_\omega)$ are realizable as links is still unsubstantiated. In fact, there is a well-defined map relating Levine's group and Orr's: $l\colon \pi_3(K_\infty) \to \pi_3(K_\omega)$. It is not hard to see from Levine's arguments that the elements in the image of this map are exactly the realizable elements. We do not know whether this map l is surjective. In light of our results below it is probably not surjective.

2 Preliminaries

2.1 Whitehead exact sequence

In order to get a hold on $\pi_3(K_\omega)$ we use an approach due to J H C Whitehead, who defined a certain exact sequence relating the homology and homotopy groups of a space which has a transparent form for 1–connected spaces such as K_ω. Recall the definition of Whitehead's *universal quadratic functor*, Γ^2, introduced in [15, Chapter II]. For an abelian group A, the group $\Gamma^2(A)$ is generated by the symbols $\gamma(x)$, one for each $x \in A$, subject to the defining relations

(1) $\gamma(-x) = \gamma(x)$;

(2) $\gamma(x+y+z) - \gamma(x+y) - \gamma(y+z) - \gamma(x+z) + \gamma(x) + \gamma(y) + \gamma(z) = 0.$

It follows from [7] that this group is naturally isomorphic to the fourth homology group

$$\Gamma^2(A) \cong H_4\, K(A,2).$$

It follows directly form the generator-relation definition that a surjection of abelian groups induces a surjection on the Γ^2 construction. This is used repeatedly in what follows, for example, in the proof in Section 3 of the main theorem.

Let us now recall a natural exact sequence associated to a connected pointed space X, due to Whitehead. Let X be a pointed connected CW-complex with skeletal filtration

$$\mathrm{sk}_1(X) \subset \mathrm{sk}_2(X) \subset \cdots.$$

In [15], Whitehead constructs the long exact sequence

$$\cdots \to \pi_4(X) \xrightarrow{\mathrm{Hur}_4} H_4(X) \to \Gamma_3(X) \to \pi_3(X) \xrightarrow{\mathrm{Hur}_3} H_3(X) \to 0, \tag{2-1}$$

where

$$\Gamma_3(X) := \mathrm{im}\big(\pi_3(\mathrm{sk}_2(X)) \to \pi_3(\mathrm{sk}_3(X))\big)$$

and Hur_i is the ith Hurewicz homomorphism.

There is a natural map $\Gamma^2(\pi_2(X)) \to \Gamma_3(X)$, constructed as follows. Let $\eta\colon S^3 \to S^2$ be the Hopf map and let $x \in \pi_2(X)$ be expressed as the composition

$$S^2 \to \mathrm{sk}_2(X) \hookrightarrow \mathrm{sk}_3(X).$$

Then the composition

$$S^3 \xrightarrow{\eta} S^2 \to \mathrm{sk}_2(X) \hookrightarrow \mathrm{sk}_3(X)$$

defines an element $\eta^*(x) \in \Gamma_3(X)$. According to [15, Section 13], for a simply

connected space X, the natural homomorphism

$$\eta_1\colon \Gamma^2(\pi_2(X)) \to \Gamma_3(X), \quad \gamma(x) \mapsto \eta^*(x) \tag{2-2}$$

is an isomorphism of groups.

In more direct terms, for a simply connected space X, the portion

$$\pi_4(X) \to H_4(X) \to H_4K(\pi_2(X), 2) \to \pi_3(X) \xrightarrow{\mathrm{Hur}_3} H_3(X) \to 0 \tag{2-3}$$

of Whitehead's exact sequence is a portion of the Leray–Serre spectral sequence of the integral homology of the (2–connected cover) fiber sequence $X[2] \to X \to K(\pi_2(X), 2)$. In particular, we have an isomorphism $\Gamma^2(\pi_2(X)) \cong \Gamma_3(X) \cong H_4K(\pi_2(X), 2)$ for any 1–connected space X.

It follows directly from Whitehead's exact sequence that in order to estimate $\pi_3(X)$ for a 1–connected space, we might consider the cokernel of the natural map of Whitehead

$$w_X \equiv w\colon H_4(X) \to \Gamma^2(\pi_2(X)) \cong H_4K(H_2(X), 2).$$

The evaluation of this cokernel for $X = K_\omega$ is thus our main concern from now on.

2.2 Completions

The proof of the following lemma follows the argument of Bousfield given in detail in [1]. We include it here for completeness of exposition, in a more general form.

Lemma 2.1 *Let H be a group and $\{x_1, \dots, x_n\}$ a finite set of elements of H. Let*

$$H \trianglerighteq N_1 \trianglerighteq N_2 \trianglerighteq \cdots$$

be a central series of H such that $[H, N_i] = N_{i+1}$ for $i = 1, 2, \dots$. Suppose that, for every $i \geq 1$, the quotient H/N_i is normally generated by $X \cdot N_i = \{x_1N_i, \dots, x_nN_i\}$. Consider the inverse limit

$$\Gamma := \lim H/N_i$$

and the associated natural map $h\colon H \to \Gamma$. The group Γ is normally generated by the elements $h(X) = \{h(x_1), \dots, h(x_n)\}$.

Proof The normal generation condition implies that, for every $u \in N_i$, where $i \geq 2$, there exist elements $u_1, \dots, u_n \in N_{i-1}$ such that

$$u \equiv [u_1, x_1] \cdots [u_n, x_1] \mod N_{i+1}.$$

Let $a = (a_1N_1, a_2N_2, \dots) \in \Gamma$, where $a_{i+1} \equiv a_i \bmod N_i$, $a_i \in H$, $i \geq 1$. Suppose that $a_1 \in N_1$. We claim that there exist elements $g_1, \dots, g_n \in \Gamma$ such that

$$a = [g_1, h(x_1)] \cdots [g_n, h(x_n)]. \tag{2-4}$$

We construct n elements $(g_i)_{1 \leq i \leq n}$ in the inverse limit as a sequence of elements in the corresponding quotient groups:

$$g_i = (y_1^i N_1, y_2^i N_2, \dots) \in \Gamma \quad \text{for } 1 \leq i \leq n,$$

with the elements y_j^i constructed by induction as follows. Since $a_1 \in N_1$, the element a_2 also lies in N_1. There exist elements $u_1, \dots, u_n \in N_1$ such that

$$a_2 \equiv [u_1, x_1] \cdots [u_n, x_n] \mod N_3.$$

We set $y_1^i = y_2^i := u_i$. Then

$$a_2 \equiv [y_2^1, x_1] \cdots [y_2^n, x_n] \mod N_3.$$

Suppose we have constructed elements $y_1^i, \dots, y_k^i$ for $1 \leq i \leq n$ such that

$$\begin{aligned} y_{j+1}^i &\equiv y_j^i \mod N_j && \text{for } 1 \leq j \leq k-1, \\ a_j &\equiv [y_j^1, x_1] \cdots [y_j^n, x_n] \mod N_{j+1} && \text{for } 1 \leq j \leq k, \end{aligned}$$

and

$$a_{k+1} \equiv [y_k^1, x_1] \cdots [y_k^n, x_n] \mod N_{k+1}.$$

There exists $r_{k+2} \in N_{k+1}$ such that

$$a_{k+1} = [y_k^1, x_1] \cdots [y_k^n, x_n] r_{k+2}.$$

Now we find elements $v_i \in N_k$ such that

$$r_{k+2} \equiv [v_1, x_1] \cdots [v_n, x_n] \mod N_{k+2}.$$

We set $y_{k+1}^i := y_k^i v_i$ for $1 \leq i \leq n$. It follows that

$$a_{k+1} \equiv [y_{k+1}^1, x_1] \cdots [y_{k+1}^n, x_n] \mod N_{k+2}.$$

Now the result follows from the condition that H/N_1 is normally generated by elements $x_1N_1, \dots, x_nN_n$ and the presentation (2-4). □

Remark Let G be a finitely generated group and $\{\gamma_i(G)\}_{i \geq 1}$ its lower central series. The completion map $G \to \widehat{G} := \lim G/\gamma_i(G)$ induces isomorphisms of lower central quotients $G/\gamma_i(G) \xrightarrow{\sim} \widehat{G}/\gamma_i(\widehat{G})$ for $i \geq 1$; see [1]. Hence, by Lemma 2.1, $\widehat{G}$ is normally generated by images of G.

2.3 The second homology of nilpotent completions

We need certain information on the second homology groups, to be used in the proof of the main theorem. In this paragraph we recall a useful relation between the second homology $H_2(G)$ of a finitely generated group and that of its nilpotent completion $H_2(\widehat{G})$. Let F be as above a finitely generated free group. Bousfield shows in [2] that $H_2(\widehat{F})$ is uncountable. In fact, $H_2(\widehat{F})$ maps onto the exterior square of 2–adic integers. An exposition of Bousfield's method of study, the homology of pronilpotent completions, is given in [10]. Let G be a finitely presented group, F a free group and $F \to G$ an epimorphism, and let $R = \ker\{F \to G\}$. The quotients

$$\mathcal{B}_k(G) := \frac{R \cap \gamma_k(F)}{[R, \underbrace{F, \dots, F}_{k-1}]}$$

are known as *Baer invariants* of G and are naturally isomorphic to the first (nonadditive) derived functors of the kth lower central quotients $G/\gamma_k(G)$. They form a tower of groups.

We will need the following property (see [10, Lemma 6.1]):

Proposition 2.2 *Assume that*

$$\lim{}^1 \mathcal{B}_k(G) = 0.$$

Then the cokernel of the natural map $H_2(\widehat{F}) \to H_2(\widehat{G})$ is isomorphic to a quotient of the homology group $H_2(G)$, as in the diagram:

$$\begin{array}{ccccc} & & H_2(G) & & \\ & & \downarrow & \searrow & \\ H_2(\widehat{F}) & \longrightarrow & H_2(\widehat{G}) & \longrightarrow & \operatorname{coker} \end{array}$$

The condition holds, for example, for all finitely presented groups G with finite $H_2(G)$.

3 Proof of Theorem 1.1

We first prove a statement similar to the claim of the main theorem but regarding a simpler, virtually nilpotent group G, rather than the free group F above. We will then compare K_ω to the following space cof_G and complete the proof of the main theorem.

Given any group G we consider the space cof_G defined as the cofiber of

$$K(G,1) \to K(\widehat{G},1).$$

Thus cof_F is just Orr's space K_ω.

In all our cases this cofiber space will be simply connected, since we only consider groups G for which $\widehat{G}$ is normally generated by the images of G. Notice that it is the case for any finitely generated group (see Remark).

Let G be given by

$$G = \langle a, b \mid b^a = b^{-1},\, a^2 = 1\rangle = \mathbb{Z} \rtimes C_2,$$

where $C_2 = \langle a \mid a^2 = 1\rangle$ is the two-element cyclic group acting nontrivially on the integers $\mathbb{Z}$.

We are interested in the third homotopy group of cof_G. Note that by an easy computation below, the third homology $H_3(\mathrm{cof}_G)$ is infinitely generated, and thus the third homotopy group of this space does not vanish, since the space is simply connected. But we need a more precise statement as follows.

Proposition 3.1 *Let G be the semidirect product as above. The kernel of the Hurewicz map*

$$\pi_3(\mathrm{cof}_G) \twoheadrightarrow H_3(\mathrm{cof}_G)$$

is infinitely generated.

Proof Our proof is based on the naturality of Whitehead's map $w\colon H_4(X) \to \Gamma^2(H_2(X))$ with respect to the following zigzagging sequence of maps involving rational completion, denoted by $\mathbb{Q}_\infty$:

$$\mathrm{cof}_G \to \mathrm{cof} \leftarrow \mathbb{Q}_\infty K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1) \to \mathbb{Q}_\infty K(\mathbb{Q}^{\oplus n} \rtimes C_2, 1).$$

Spaces and maps including cof and cof_G in this sequence are discussed below. The functor $\mathbb{Q}_\infty$ is the $\mathbb{Q}$–completion in the sense of Bousfield and Kan [3].

Our interest of course is in the value of the cokernel of w for the space on the left. We estimate the cokernel of w for the left-hand space cof_G by reducing the calculations to an estimate the cokernel of w for the rightmost space.

We will show shortly that in the above sequence of spaces, the two maps on the left induce isomorphisms on the said cokernel of the Whitehead map w while the map on the right induces (again, on the said cokernels) a *surjection* onto an infinitely generated group (= the cokernel for the space on the right-hand side), thus completing the proof.

Recall that to prove the current proposition, using the above exact sequence of Whitehead for the case of 1–connected spaces it is enough to show that the cokernel of the natural map $H_4(\mathrm{cof}_G) \to \Gamma^2(\pi_2(\mathrm{cof}_G))$ is infinitely generated. However, since the space cof_G is simply connected we have $H_2(\mathrm{cof}_G) \cong \pi_2(\mathrm{cof}_G)$; thus we shall consider the second homology in what follows.

First notice that the nilpotent completion of G is just a semidirect product of the 2–adic integers with the cyclic group C_2, namely, $\widehat{G} \cong \mathbb{Z}_2 \rtimes C_2$. This follows from a direct computation of the nilpotent quotients as a semidirect product.

Now we consider the map on the right in the above zigzag of maps. The spaces and map are defined using the following cofiber sequence. Its usefulness follows from the fact that integral homology groups of cof_G turns out to be rational vector spaces, by direct computation as in the lemma below. This allows us to pass to rational completions and work with vector spaces over the rational numbers:

The first map on the left in the zigzag and its domain and codomain are defined in the following cofibration sequences diagram where the commutative square on the left is induced by the maps of C_2–modules $\mathbb{Z} \to \mathbb{Q}$ and $\mathbb{Z}_2 \to \mathbb{Z}_2 \otimes \mathbb{Q}$:

$$\begin{array}{ccccc} K(\mathbb{Z} \rtimes C_2, 1) & \longrightarrow & K(\mathbb{Z}_2 \rtimes C_2, 1) & \longrightarrow & \mathrm{cof}_G \\ \downarrow & & \downarrow & & \downarrow \\ K(\mathbb{Q} \rtimes C_2, 1) & \longrightarrow & K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1) & \longrightarrow & \mathrm{cof} \end{array}$$

Lemma 3.2 *The cofiber spaces* cof_G *and* cof *defined above are simply connected and their homotopy and homology groups are naturally vector spaces over the field of rational numbers* $\mathbb{Q}$*. In addition, the above induced map* $\mathrm{cof}_G \to \mathrm{cof}$ *is a homotopy equivalence.*

Proof Both spaces cof_G and cof are simply connected: this is because their fundamental groups are clearly abelian and their first homology groups vanish, for example by the calculation below.

The homology groups $H_i(\mathbb{Z} \rtimes C_2)$ can be computed using the usual spectral sequence for the group extension $\mathbb{Z} \hookrightarrow \mathbb{Z} \rtimes C_2 \twoheadrightarrow C_2$. The second stage of this spectral sequence stabilizes by the existence of the splitting map $C_2 \hookrightarrow (\mathbb{Z} \rtimes C_2)$:

$$E^2_{i,j} = \begin{cases} H_i(C_2, \langle b \rangle) = \mathbb{Z}/2 & \text{if } (i,j) = (2k, 1),\ k \geq 0 \\ H_i(C_2) = \mathbb{Z}/2 & \text{if } (i,j) = (2k+1, 0),\ k \geq 0, \\ \mathbb{Z} & \text{if } (i,j) = (0,0), \\ 0 & \text{otherwise.} \end{cases}$$

In particular, we get

$$H_i(\mathbb{Z} \rtimes C_2) = \begin{cases} \mathbb{Z}/2 \oplus \mathbb{Z}/2 & \text{for } i \text{ odd}, \\ 0 & \text{for } i \text{ even} > 0. \end{cases}$$

In the same way we compute the homology groups $H_i(\mathbb{Z}_2 \rtimes C_2)$. In this case, the spectral sequence has the form

$$E^2_{i,j} = \begin{cases} \Lambda^{2k}(\mathbb{Z}_2) & \text{if } (i,j) = (0, 2k),\ k \geq 1, \\ H_i(C_2, \langle b \rangle) = \mathbb{Z}/2 & \text{if } (i,j) = (2k, 1),\ k \geq 0, \\ H_i(C_2) = \mathbb{Z}/2 & \text{if } (i,j) = (2k+1, 0),\ k \geq 0, \\ \mathbb{Z} & \text{if } (i,j) = (0,0), \\ 0 & \text{otherwise.} \end{cases}$$

In particular,

$$H_i(\mathbb{Z}_2 \rtimes C_2) = \begin{cases} \mathbb{Z}/2 \oplus \mathbb{Z}/2 & \text{for } i \text{ odd}, \\ \Lambda^{2k}(\mathbb{Z}_2) & \text{for } i = 2k,\ k \geq 1. \end{cases}$$

The monomorphism $\langle b \rangle = \mathbb{Z} \hookrightarrow \mathbb{Z}_2$ induces isomorphisms of homology groups

$$H_i(C_2, \langle b \rangle) \simeq H_i(C_2, \mathbb{Z}_2).$$

Hence, the homology groups of cof_G are

$$H_i(\mathrm{cof}_G) = \begin{cases} \Lambda^{2k}(\mathbb{Z}_2) & \text{for } i = 2k,\ k \geq 1, \\ 0 & \text{for } i = 2k+1. \end{cases}$$

Similarly, the homology groups of cof are

$$H_i(\mathrm{cof}) = \begin{cases} \Lambda^{2k}(\mathbb{Z}_2 \otimes \mathbb{Q}) & \text{for } i = 2k,\ k \geq 1, \\ 0 & \text{for } i = 2k+1. \end{cases}$$

Note that all exterior powers ≥ 2 of $\mathbb{Z}_2$ are $\mathbb{Q}$–vector spaces. To see this, we present the exterior power as a natural quotient of the tensor power, and after tensoring it with $\mathbb{Z}/l$, where $l \geq 2$, one obtains a finite cyclic group, whose image in the exterior power is zero. That is, the exterior powers ≥ 2 of $\mathbb{Z}_2$ are divisible. Since these groups are torsion-free, we conclude that they are $\mathbb{Q}$–vector spaces. Thus the map $\mathrm{cof}_G \to \mathrm{cof}$ is a homology equivalence. Since cof_G and cof are simply connected, $\mathrm{cof}_G \to \mathrm{cof}$ is a homotopy equivalence. □

This completes our discussion of the first map in the above zigzag of maps.

We now turn our attention to the second map. This map arises from the fact proven in Lemma 3.2. As we saw in Lemma 3.2 the homology of cof is a vector space over the field $\mathbb{Q}$.

Spaces here are virtually nilpotent, hence their $\mathbb{Q}$–completions are the same as $\mathbb{Q}$–localizations, in particular, the $\mathbb{Q}$–homology groups of their $\mathbb{Q}$–completions are the same as those of the spaces (see [6, Proposition 3.4]). Thus the map above that defines the space cof as homotopy quotient, namely,

$$K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1) \to \text{cof},$$

factors through the rational completion-localization of the domain. The second map is this canonical factorization.

Lemma 3.3 *The second map $K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1) \to \text{cof}$ above induces an isomorphism homology with rational coefficients.*

Proof This follows directly from the fact that the cobase in the cofibration that defines cof, namely, the space $K(\mathbb{Q} \rtimes C_2, 1)$, has trivial homology with coefficients in $\mathbb{Q}$ by a direct computation via the usual $H(-, \mathbb{Q})$ spectral sequence whose $E^2_{p,q}$ term vanishes anywhere except at $(p, q) = (0, 0)$. Put otherwise, the group G is isomorphic to the free product $C_2 * C_2$ of two cyclic groups whose $\mathbb{Q}$–homology vanishes by the usual pushout Mayer–Vietoris argument. □

It follows from the lemma above that to estimate the cokernel of the Whitehead natural transformation $H_4 \to \Gamma^2 H_2$ for the 1–connected space cof_G we work entirely with vector spaces over $\mathbb{Q}$ and we need to estimate the following cokernel of Whitehead's map w for the third space in the above zigzag. Using the computation in the proof of Lemma 3.2 we need to estimate

$$\text{coker}(\Lambda^4(\mathbb{Z}_2 \otimes \mathbb{Q}) \to \Gamma^2\Lambda^2(\mathbb{Z}_2 \otimes \mathbb{Q})).$$

We show that this cokernel maps *onto* an infinitely generated group by comparing it to the cokernel of the Whitehead map of another space, the fourth in the zigzag above, as follows:

To build the map on the right of our zigzag we note that for any n, there is an epimorphism $\mathbb{Z}_2 \otimes \mathbb{Q} \to \mathbb{Q}^{\oplus n}$ which induces a map between groups

$$(\mathbb{Z}_2 \otimes \mathbb{Q}) \rtimes C_2 \to \mathbb{Q}^{\oplus n} \rtimes C_2,$$

where the action by C_2 in both semidirect products is given by negation. The group $\mathbb{Q}^{\oplus n} \rtimes C_2$ is $\mathbb{Q}$–perfect, ie its abelianization is a torsion group. We have a map between $\mathbb{Q}$–completions

$$\mathbb{Q}_\infty K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1) \to \mathbb{Q}_\infty K(\mathbb{Q}^{\oplus n} \rtimes C_2, 1).$$

As noted above, both spaces $K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1)$ and $K(\mathbb{Q}^{\oplus n} \rtimes C_2, 1)$ are virtually nilpotent, hence their $\mathbb{Q}$–completions are the same as $\mathbb{Q}$–localizations, in particular, the $\mathbb{Q}$–homology groups of their $\mathbb{Q}$–completions are the same as those of the spaces (see [6, Proposition 3.4]).

Recall that the Γ^2 functor sends epimorphisms to epimorphisms. The natural maps between Whitehead sequences

$$\begin{array}{ccc} H_4\mathbb{Q}_\infty K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1) & \longrightarrow & \Gamma^2(H_2\mathbb{Q}_\infty K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1)) \\ \downarrow & & \downarrow \\ H_4\mathbb{Q}_\infty K(\mathbb{Q}^{\oplus n} \rtimes C_2, 1) & \longrightarrow & \Gamma^2(H_2\mathbb{Q}_\infty K(\mathbb{Q}^{\oplus n} \rtimes C_2, 1)) \end{array}$$

are the following:

$$\begin{array}{ccc} \Lambda^4(\mathbb{Z}_2 \otimes \mathbb{Q}) & \longrightarrow & \Gamma^2(\Lambda^2(\mathbb{Z}_2 \otimes \mathbb{Q})) \\ \downarrow & & \downarrow \\ \Lambda^4(\mathbb{Q}^{\oplus n}) & \longrightarrow & \Gamma^2(\Lambda^2(\mathbb{Q}^{\oplus n})) \end{array}$$

Now observe that for $n = 2, 3$, the group $\Lambda^4(\mathbb{Q}^{\oplus n})$ is zero, but the group $\Gamma^2(\Lambda^2(\mathbb{Q}^{\oplus n}))$ is infinitely generated. Therefore, the cokernel of the upper horizontal map in the last diagram must be an infinitely generated group.

We conclude that the cokernel of the map

$$H_4(\mathrm{cof}_G) \to \Gamma^2(\pi_2(\mathrm{cof}_G))$$

is an infinite divisible group as claimed. □

Proof of Theorem 1.1 We use the statement in Proposition 3.1 about cof_G for the semidirect product G as above to deduce a similar statement about the free group F. Consider a free group F of rank 2 and an epimorphism $F \to G$, and construct the map between cofibers:

$$\begin{array}{ccccc} K(F,1) & \longrightarrow & K(\widehat{F},1) & \longrightarrow & K_\omega \\ \downarrow & & \downarrow & & \downarrow \\ K(G,1) & \longrightarrow & K(\widehat{G},1) & \longrightarrow & \mathrm{cof}_G \end{array}$$

In light of $H_1(F) \cong H_1(\widehat{F})$ above we have an isomorphism $H_2(\widehat{F}) \simeq H_2(K_\omega)$. The cokernel of the natural map

$$H_2(\widehat{F}) \simeq H_2(K_\omega) \to H_2(\mathrm{cof}_G) = \Lambda^2(\mathbb{Z}_2)$$

is isomorphic, by Proposition 2.2, to a quotient of $H_2(G) = 0$ (see Section 2); hence we conclude that $H_2(\widehat{F}) \to H_2(\mathrm{cof}_G)$ is an epimorphism. Hence the middle vertical map in the natural map between Whitehead exact sequences is a surjection:

$$\begin{array}{ccccccc} H_4(\widehat{F}) & \longrightarrow & \Gamma^2(H_2(\widehat{F})) & \twoheadrightarrow & \mathrm{coker} & \xrightarrow{1\text{-}1} & \pi_3(K_\omega) \\ \downarrow & & \downarrow & & \downarrow ? & & \downarrow \\ H_4(\mathrm{cof}_G) & \longrightarrow & \Gamma^2(H_2(\mathrm{cof}_G)) & \twoheadrightarrow & \mathrm{coker} & \xrightarrow{1\text{-}1} & \pi_3(\mathrm{cof}_G) \end{array}$$

It is clear now that the arrow with a question mark must be a surjection: since the map $H_2(\widehat{F}) \to H_2(\mathrm{cof}_F) = H_2 K_\omega$ is an epimorphism, the same is true for $\Gamma^2(H_2(\widehat{F})) \to \Gamma^2(H_2(\mathrm{cof}_G))$, and we conclude that the map

$$H_4(\widehat{F}) \to \Gamma^2(H_2(\widehat{F}))$$

has a cokernel which maps onto an infinite divisible group. Theorem 1.1 now follows from the Whitehead exact sequence for the simply connected space K_ω:

$$H_4(\widehat{F}) \to \Gamma^2(H_2(\widehat{F})) \to \pi_3(K_\omega) \to H_3(\widehat{F}) \to 0. \qquad \square$$

4 Localization of virtually nilpotent spaces

Here we deduce from the above results that, for a virtually nilpotent group G, the $H\mathbb{Z}$–localization of $K(G,1)$ is not, in general, a $K(-,1)$. In fact we consider the simplest possible nontrivial example of such a group, namely, the semidirect product G above. Recall from [6] that, for a virtually nilpotent space X, the arithmetic square

$$\begin{array}{ccc} X_{\mathbb{Z}} & \longrightarrow & X_P \\ \downarrow & & \downarrow \\ X_{\mathbb{Q}} & \longrightarrow & X_{P,\mathbb{Q}} \end{array}$$

is, up to homotopy, a fiber square. Here $P = \bigoplus \mathbb{Z}/p$ over all primes and, for a ring R, X_R is the HR–localization.

Consider $X = K(G,1)$, where G is the group from the previous proof, ie $G = \mathbb{Z} \rtimes C_2$. Clearly, G is virtually nilpotent, and it has an infinite cyclic subgroup of index two. Thus the space X is virtually nilpotent and therefore the arithmetic square for X is a fiber square. Since the group G is $\mathbb{Q}$–acyclic, $X_{\mathbb{Q}}$ is contractible and the arithmetic square degenerates to the fiber sequence

$$X_{\mathbb{Z}} \to X_{\mathbb{Z}/2} \to X_{\mathbb{Z}/2,\mathbb{Q}}.$$

Clearly we can ignore all primes from P except for $p = 2$.

The $\mathbb{Z}/2$–localization $X_{\mathbb{Z}/2}$ coincides with the $\mathbb{Z}/2$–completion and is equivalent to $K(\mathbb{Z}_2 \rtimes C_2, 1)$. The space $K(\mathbb{Z}_2 \rtimes C_2, 1)$ is virtually nilpotent; hence its $\mathbb{Q}$–localization coincides with its $\mathbb{Q}$–completion. Therefore

$$\pi_i(X_{\mathbb{Z}}) \simeq \pi_{i+1}(\mathbb{Q}_\infty K(\mathbb{Z}_2 \rtimes C_2, 1)),$$

for all $i \geq 2$. The natural map $\mathbb{Z}_2 \rtimes C_2 \to \mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2$ induces an isomorphism of $\mathbb{Q}$–homology groups; hence

$$\pi_i(X_{\mathbb{Z}}) \simeq \pi_{i+1}(\mathbb{Q}_\infty K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1)).$$

It follows from the previous section that the kernel of the third Hurewicz homomorphism for the space $\mathbb{Q}_\infty K(\mathbb{Z}_2 \otimes \mathbb{Q} \rtimes C_2, 1)$ contains an infinitely generated subgroup; hence $\pi_2(X_{\mathbb{Z}}) \neq 0$.

Acknowledgements

The research is supported by the Russian Science Foundation grant N 16-11-10073. Dror Farjoun is grateful for the hospitality of Chebyshev Laboratory, Saint Petersburg.

References

[1] **G Baumslag**, **U Stammbach**, *On the inverse limit of free nilpotent groups*, Comment. Math. Helv. 52 (1977) 219–233 MR

[2] **A K Bousfield**, *Homological localization towers for groups and Π–modules*, Mem. Amer. Math. Soc. 186, Amer. Math. Soc., Providence, RI (1977) MR

[3] **A K Bousfield**, **D M Kan**, *Homotopy limits, completions and localizations*, Lecture Notes in Mathematics 304, Springer (1972) MR

[4] **T D Cochran**, *Link concordance invariants and homotopy theory*, Invent. Math. 90 (1987) 635–645 MR

[5] **T D Cochran**, *Derivatives of links: Milnor's concordance invariants and Massey's products*, Mem. Amer. Math. Soc. 427, Amer. Math. Soc., Providence, RI (1990) MR

[6] **E Dror**, **W G Dwyer**, **D M Kan**, *An arithmetic square for virtually nilpotent spaces*, Illinois J. Math. 21 (1977) 242–254 MR

[7] **S Eilenberg**, **S Mac Lane**, *On the groups $H(\Pi, n)$, II: Methods of computation*, Ann. of Math. 60 (1954) 49–139 MR

[8] **N Habegger**, **X-S Lin**, *The classification of links up to link-homotopy*, J. Amer. Math. Soc. 3 (1990) 389–419 MR

[9] **K Igusa**, **K E Orr**, *Links, pictures and the homology of nilpotent groups*, Topology 40 (2001) 1125–1166 MR

[10] **S O Ivanov**, **R Mikhailov**, *On lengths of HZ-localization towers*, preprint (2016) arXiv

[11] **J P Levine**, *Link concordance and algebraic closure of groups*, Comment. Math. Helv. 64 (1989) 236–255 MR

[12] **J Milnor**, *Isotopy of links*, from "Algebraic geometry and topology: a symposium in honor of S Lefschetz" (R Fox, D Spencer, A Tucker, editors), Princeton Univ. Press (1957) 280–306 MR

[13] **K E Orr**, *New link invariants and applications*, Comment. Math. Helv. 62 (1987) 542–560 MR

[14] **K E Orr**, *Homotopy invariants of links*, Invent. Math. 95 (1989) 379–394 MR

[15] **J H C Whitehead**, *A certain exact sequence*, Ann. of Math. 52 (1950) 51–110 MR

EDF: *Department of Mathematics, Hebrew University of Jerusalem*
Jerusalem, Israel

Current address for RM: *Chebyshev Laboratory, Saint Petersburg State University*
Saint Petersburg, Russia

RM: *Saint Petersburg Department of Steklov Mathematical Institute*
Saint Petersburg, Russia

RM: *School of Mathematics, Tata Institute of Fundamental Research*
Mumbai, India

farjoun@math.huji.ac.il, romanvm@mi.ras.ru

Received: 30 March 2017 Revised: 2 July 2017

msp

Algebraic & Geometric Topology 18 (2018) 583–612

Groups of homotopy classes of phantom maps

HIROSHI KIHARA

We introduce a new approach to phantom maps which largely extends the rationalization-completion approach developed by Meier and Zabrodsky. Our approach enables us to deal with the set $\mathrm{Ph}(X,Y)$ of homotopy classes of phantom maps and the subset $\mathrm{SPh}(X,Y)$ of homotopy classes of special phantom maps simultaneously. We give a sufficient condition for $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ to have natural group structures, which is much weaker than the conditions obtained by Meier and McGibbon. Previous calculations of $\mathrm{Ph}(X,Y)$ have generally assumed that $[X,\Omega\widehat{Y}]$ is trivial, in which case generalizations of Miller's theorem are directly applicable, and calculations of $\mathrm{SPh}(X,Y)$ have rarely been reported. Here, we calculate not only $\mathrm{Ph}(X,Y)$ but also $\mathrm{SPh}(X,Y)$ in many important cases of nontrivial $[X,\Omega\widehat{Y}]$.

55Q05; 55P60

1 Introduction

Given two pointed CW-complexes X and Y, a map $f\colon X\to Y$ is called a *phantom map* if for any finite complex K and any map $h\colon K\to X$, the composite fh is nullhomotopic. Let $\mathrm{Ph}(X,Y)$ denote the subset of $[X,Y]$ consisting of homotopy classes of phantom maps. Since $\mathrm{Ph}(X,Y)=0$ for any finite complex X, the phantom map concept is a key to understanding maps with infinite-dimensional sources, and has been an important topic in homotopy theory since its discovery; see McGibbon [14] and Roitberg [24].

We briefly review the main results of the theory of phantom maps, focusing on the two most important subjects: natural group structure on $\mathrm{Ph}(X,Y)$ and calculation of $\mathrm{Ph}(X,Y)$.

Recall that a space is called an H_0–space (resp. co-H_0–space) if its rationalization is homotopy equivalent to a product of Eilenberg–Mac Lane spaces (resp. coproduct of Moore spaces). Let $\widehat{\mathbb{Z}}$ denote the product $\prod_p \widehat{\mathbb{Z}}_p$ of the p–completions of $\mathbb{Z}$, in which $\mathbb{Z}$ is diagonally contained.

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.583

Theorem A *Let X be a connected CW-complex and Y a nilpotent CW-complex of finite type with π_1 finite.*

(1) *If X is a co-H_0–space or Y is an H_0–space, $\mathrm{Ph}(X,Y)$ has a natural, divisible, abelian group structure.*

(2) *Suppose that X is a Postnikov space with π_1 locally finite, the classifying space of a compact Lie group, or an infinite loop space with π_1 torsion. Suppose that Y is a nilpotent finite complex or an iterated loop space of such a space. Then there is a noncanonical bijection*

$$\mathrm{Ph}(X,Y) \cong \prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes \widehat{\mathbb{Z}}/\mathbb{Z}).$$

Furthermore, if X is a co-H_0–space or Y is an H_0–space, this bijection can be taken as an isomorphism of groups.

Remark 1.1 (1) Zabrodsky stated [34, Theorem B(d)] that $\mathrm{Ph}(X,Y)$ is a divisible abelian group assuming only that X and Y are 1–connected with finite type. However, it is thought that this assertion is incorrect; see [14, page 1236]. Meier, McGibbon, Roitberg and coauthors have attempted to find a sufficient condition such that $\mathrm{Ph}(X,Y)$ has a natural group structure. Theorem A(1) was essentially proposed by Meier [19] and McGibbon [12, Theorem 4]; see also Roitberg and Touhey [25, page 302].

(2) $\mathrm{Ph}(X,Y)$ is described as the quotient of the set that is (noncanonically) isomorphic to the product $\prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\widehat{\mathbb{Z}}/\mathbb{Z})$ by the group action of $[X,\Omega\widehat{Y}]$ (see the works of Meier [20], Zabrodsky [34], and Roitberg and Touhey [25]). Theorem A(2), which is a central result in the calculation of $\mathrm{Ph}(X,Y)$, follows immediately from this description since $[X,\Omega\widehat{Y}]$ is trivial by generalizations of Miller's theorem (see Proposition 6.3(1)). Though most experts on phantom maps have attempted to calculate $\mathrm{Ph}(X,Y)$, an effective method has not been found in cases where $[X,\Omega\widehat{Y}]$ is nontrivial. Thus, such a method is most desirable.

We are also interested in the subset $\mathrm{SPh}(X,Y)$ of $\mathrm{Ph}(X,Y)$ consisting of homotopy classes of *special phantom maps*, defined by the exact sequence of pointed sets

$$0 \to \mathrm{SPh}(X,Y) \to \mathrm{Ph}(X,Y) \xrightarrow{e_{Y\sharp}} \mathrm{Ph}(X,\check{Y}), \tag{1-1}$$

where $e_Y\colon Y\to\check{Y}=\prod_p Y_{(p)}$ is a natural map called the *local expansion* (see [24, page 150]). The target Y is usually assumed to be nilpotent of finite type.

$\mathrm{SPh}(X,Y)$ is less clearly understood than $\mathrm{Ph}(X,Y)$, and although we know that a result similar to Theorem A(1) holds under the finite-type condition on X (see [12, Theorem 4]), calculations of $\mathrm{SPh}(X,Y)$ have rarely been reported.

In this paper, we develop a new approach to phantom maps which enables us to deal with $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ simultaneously and to address the following two problems:

Problem 1 Identify a weaker sufficient condition such that $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ have natural group structures.

Problem 2 Calculate $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$, especially in the cases where $[X,\Omega\widehat{Y}]$ is nontrivial.

Our solution to Problem 1 is given in Theorem 2.3(1). Our solution to Problem 2 is given in Proposition 2.5 and Theorem 2.7. Proposition 2.5 directly generalizes Theorem A(2) (see Remark 2.6). Theorem 2.7 gives a method for calculating not only $\mathrm{Ph}(X,Y)$ but also $\mathrm{SPh}(X,Y)$ in various important cases of *highly nontrivial* $[X,\Omega\widehat{Y}]$; the power of Theorem 2.7 is illustrated by Corollaries 2.8–2.10.

2 Main results

2.1 Solution to Problem 1

Let $\mathcal{CW}$ denote the category of pointed connected CW-complexes and homotopy classes of maps and let $\mathcal{N}$ denote the full subcategory of $\mathcal{CW}$ consisting of nilpotent CW-complexes of finite type.

Definition 2.1 Let $\mathcal{Q}$ be the full subcategory of $\mathcal{CW}^{\mathrm{op}}\times\mathcal{N}$ consisting of (X,Y) satisfying the following condition:

(Q) For each pair $i,j>0$, the rational cup product

$$\cup\colon H^i(X;\mathbb{Q})\otimes H^j(X;\mathbb{Q})\to H^{i+j}(X;\mathbb{Q})$$

or the rational Whitehead product

$$[\,,]\colon(\pi_{i+1}(Y)\otimes\mathbb{Q})\otimes(\pi_{j+1}(Y)\otimes\mathbb{Q})\to\pi_{i+j+1}(Y)\otimes\mathbb{Q}$$

is trivial.

Remark 2.2 A pair $(X,Y)\in\mathcal{CW}^{\mathrm{op}}\times\mathcal{N}$ is in $\mathcal{Q}$ if X is a co-H_0–space or Y is an H_0–space. $\mathcal{Q}$ contains many other pairs (see Section 4.2).

We have the following fundamental theorem regarding natural group structures on $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$.

Theorem 2.3 *Let (X,Y) be an object of $\mathcal{Q}$.*

(1) *$\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ have natural divisible abelian group structures, for which $\mathrm{SPh}(X,Y)$ is a subgroup of $\mathrm{Ph}(X,Y)$.*

(2) *Let $(f^{\mathrm{op}},g)\colon (K,L)\to(X,Y)$ be a morphism of $\mathcal{CW}^{\mathrm{op}}\times\mathcal{N}$. Then the images $\mathrm{Im}\,\mathrm{Ph}(f,g)$ and $\mathrm{Im}\,\mathrm{SPh}(f,g)$ are divisible abelian subgroups of $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$, respectively.*

(3) *If X is a co-H–space or Y is an H–space, the group structures on $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ are compatible with the multiplicative structures on $[X,Y]$.*

Remark 2.4 (1) By Remark 2.2, Theorem 2.3(1) generalizes Theorem A(1). Furthermore, we have a natural epimorphism

$$\prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\hat{\mathbb{Z}}/\mathbb{Z})\to\mathrm{Ph}(X,Y)$$

and a similar natural epimorphism onto $\mathrm{SPh}(X,Y)$ for $(X,Y)\in\mathcal{Q}$ (Propositions 5.10 and 5.12); such natural epimorphisms are obtained for the first time (see Theorem A(2) and Remark 1.1(2)).

(2) Part 2 of Theorem 2.3 describes a new feature of natural group structures on phantom maps and provides a means of calculating $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ (see Theorem 2.7).

2.2 Solution to Problem 2

We begin with a direct generalization of Part 2 of Theorem A.

A space whose i^{th} homotopy group is zero for $i\le n$ and locally finite for $i=n+1$ is said to be *$n\frac{1}{2}$–connected*. Let us define the classes $\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}'$ and $\mathcal{B}'$ by:

$\mathcal{A}=$ the class of $\frac{1}{2}$–connected Postnikov spaces, the classifying spaces of compact Lie groups, $\frac{1}{2}$–connected infinite loop spaces and their iterated suspensions.

$\mathcal{B}=$ the class of nilpotent finite complexes, the classifying spaces of compact Lie groups and their iterated loop spaces.

$\mathcal{A}'=$ the class of $1\frac{1}{2}$–connected Postnikov spaces of finite type and their iterated suspensions.

$\mathcal{B}'=$ the class of BU, BO, $B\,\mathrm{Sp}$, $B\,\mathrm{SO}$, U/Sp, Sp/U, SO/U, U/SO, and their iterated loop spaces.

A pair $(X,Y)\in\mathcal{A}\times\mathcal{B}$ is in $\mathcal{Q}$ if Y is an iterated loop space of a nilpotent finite complex or if Y is the classifying space of a connected Lie group. Any pair $(X,Y)\in\mathcal{A}'\times\mathcal{B}'$ is in $\mathcal{Q}$.

By showing that $[X,\Omega\widehat{Y}]=0$ if (X,Y) is in $\mathcal{A}\times\mathcal{B}$ or $\mathcal{A}'\times\mathcal{B}'$ (Corollary 6.4), we generalize Theorem A(2). Let $\check{\mathbb{Z}}$ denote the product $\prod_p \mathbb{Z}_{(p)}$ of the p–localizations of $\mathbb{Z}$, in which $\mathbb{Z}$ is diagonally contained.

Proposition 2.5 *Let (X,Y) be in $\mathcal{A}\times\mathcal{B}$ or $\mathcal{A}'\times\mathcal{B}'$. Then there exist bijections*

$$\mathrm{Ph}(X,Y)\cong\prod_{i>0}H^i(X;\pi_{i+1}(Y)\otimes\widehat{\mathbb{Z}}/\mathbb{Z}),$$
$$\mathrm{SPh}(X,Y)\cong\prod_{i>0}H^i(X;\pi_{i+1}(Y)\otimes\check{\mathbb{Z}}/\mathbb{Z}).$$

If (X,Y) is in $\mathcal{Q}$, then these bijections can be taken to be natural isomorphisms of abelian groups.

Remark 2.6 The following notes relate to Proposition 2.5:

(1) We should note the following points:

- Besides Miller's theorem, the theorem of Anderson and Hodgkin can be generalized and used to calculate $\mathrm{Ph}(X,Y)$ (see Proposition 6.3 and Corollary 6.4).
- In the generalized Miller's theorem for calculating $\mathrm{Ph}(X,Y)$, the target spaces can be the classifying spaces of compact Lie groups.

These points have been largely overlooked in the literature; an exception is Meier [20], who proved Proposition 6.3(2) for $A=K(\mathbb{Z},n)$ with $n\geq 3$ and $B=BU$, and hence calculated the group of homotopy classes of phantom maps.

(2) Recall that calculations of $\mathrm{SPh}(X,Y)$ have rarely been reported. We should also note the following points:

- The vanishing of $[X,\Omega\widehat{Y}]$ enables us to calculate not only $\mathrm{Ph}(X,Y)$, but also $\mathrm{SPh}(X,Y)$ (see Remark 6.2(2)).
- The last assertion on the group structures on $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ is novel.

Next, we present a new method for calculating the groups $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ for $(X,Y)\in\mathcal{Q}$ with $[X,\Omega\widehat{Y}]\neq 0$. Note that in the following theorem, $p^\sharp\,\mathrm{Ph}(K,Y)$ and $p^\sharp\,\mathrm{SPh}(K,Y)$ are the subgroups of $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ (Theorem 2.3(2)).

Theorem 2.7 *Let (X,Y) be in $\mathcal{Q}$. Let $X' \xrightarrow{i} X \xrightarrow{p} K$ be a cofibration sequence with $[X', \Omega\widehat{Y}] = 0$, or a fibration sequence with weakly contractible $\mathrm{map}_*(X', \Omega\widehat{Y})$. Then there exist natural split exact sequences of abelian groups given by*

$$0 \to p^\sharp \mathrm{Ph}(K,Y) \to \mathrm{Ph}(X,Y) \to \prod_{i>0} H^i(X; \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}) / p^* H^i(K; \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}) \to 0,$$

$$0 \to p^\sharp \mathrm{SPh}(K,Y) \to \mathrm{SPh}(X,Y) \to \prod_{i>0} H^i(X; \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}) / p^* H^i(K; \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}) \to 0.$$

Using Theorem 2.7, we can produce infinitely many computational examples. We illustrate its power through the following computational results.

Corollary 2.8 *Let X be a connected infinite loop space with $\pi_1(X)/\mathrm{torsion} \cong \mathbb{Z}^m$ for some $m \geq 0$, and let Y be in $\mathcal{B}$. Suppose that (X,Y) is in $\mathcal{Q}$. Then we have the natural isomorphisms of groups*

$$\mathrm{Ph}(X,Y) \cong \prod_{i>0} H^i(X; \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}) / p^* H^i(T; \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}),$$

$$\mathrm{SPh}(X,Y) \cong \prod_{i>0} H^i(X; \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}) / p^* H^i(T; \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}),$$

where the map p from X to the m–dimensional torus T is defined to be the composite $X \to K(\pi_1(X),1) \to K(\pi_1(X)/\mathrm{torsion},1) = T$ of the canonical maps.

Let $K\langle n\rangle$ denote the n–connected cover of K.

Corollary 2.9 *Suppose that K is a $1\frac{1}{2}$–connected finite complex and Y is in $\mathcal{B}$, or that K is a $2\frac{1}{2}$–connected finite complex and Y is in $\mathcal{B}'$. If $(K\langle n\rangle, Y)$ is in $\mathcal{Q}$, there exist natural isomorphisms of groups*

$$\mathrm{Ph}(K\langle n\rangle, Y) \cong \prod_{i>0} H^i(K\langle n\rangle; \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}) / p^* H^i(K; \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}),$$

$$\mathrm{SPh}(K\langle n\rangle, Y) \cong \prod_{i>0} H^i(K\langle n\rangle; \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}) / p^* H^i(K; \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}),$$

where $p\colon K\langle n\rangle \to K$ is the canonical map.

Corollary 2.10 *Let (X,Y) be in $\mathcal{Q}$. Suppose that X is the product of connected CW-complexes X' and K (ie $X'\times K$). Suppose that (X',Y) is in $\mathcal{A}\times\mathcal{B}$ or $\mathcal{A}'\times\mathcal{B}'$. Then there exist natural isomorphisms of groups*

$$\mathrm{Ph}(X,Y)\cong\mathrm{Ph}(K,Y)\oplus\mathrm{Hom}(\widetilde{H}_*(X';\mathbb{Q})\otimes H_*(K;\mathbb{Q}),\pi_{*+1}(Y)\otimes\widehat{\mathbb{Z}}/\mathbb{Z}),$$
$$\mathrm{SPh}(X,Y)\cong\mathrm{SPh}(K,Y)\oplus\mathrm{Hom}(\widetilde{H}_*(X';\mathbb{Q})\otimes H_*(K;\mathbb{Q}),\pi_{*+1}(Y)\otimes\check{\mathbb{Z}}/\mathbb{Z}),$$

where Hom *denotes the module of homomorphisms of graded $\mathbb{Q}$–modules.*

See Example 6.6 for another computational result on $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$.

In the forthcoming paper, we use our approach to study the group of weak identities [23; 16]. We demonstrate that our approach is also useful for investigating the Gray index of phantom maps [5; 6; 7; 18] in the succeeding article.

The paper is organized as follows. For a given map $\varphi\colon Y\to Y'$, the subset $P^{\varphi}(X,Y)$ of $[X,Y]$ is defined as

$$P^{\varphi}(X,Y)=\{[f]\in[X,Y]\mid\varphi_\sharp([f])=0\}.$$

In Section 3, we develop a general theory leading to the double coset formula for $P^{\varphi}(X,Y)$ (Proposition 3.6) that enables us to treat $\mathrm{Ph}(X,Y)$, $\mathrm{SPh}(X,Y)$ and $\mathrm{Ph}(X,\check{Y})$ simultaneously. This formula encourages the study of the conditions under which the group $[X,\Omega Y_{(0)}]\cong[\Sigma X,Y_{(0)}]$ is abelian, to which Section 4 is devoted. In Section 5, the results of Sections 3 and 4 are applied to phantom maps, and Theorem 2.3 is established. In Section 6, we prove Proposition 2.5, Theorem 2.7, and Corollaries 2.8–2.10.

3 Double coset formula

In this section, we develop a general theory that enables simultaneous treatment of $\mathrm{Ph}(X,Y)$, $\mathrm{SPh}(X,Y)$ and $\mathrm{Ph}(X,\check{Y})$; see the exact sequence (1-1) of pointed sets.

For a given map $\varphi\colon Y\to Y'$, the subset $P^{\varphi}(X,Y)$ of $[X,Y]$ is defined as

$$P^{\varphi}(X,Y)=\{[f]\in[X,Y]\mid\varphi_\sharp([f])=0\}.$$

Under certain conditions on φ, we determine the rational homotopy structure of the double looping of the fibration sequence

$$F\to Y\xrightarrow{\varphi}Y'$$

(Proposition 3.1) and establish the double coset formula for $P^{\varphi}(X,Y)$ (Proposition 3.6). By applying these results to the profinite completion $c_Y\colon Y\to\widehat{Y}$, the local expansion $e_Y\colon Y\to\check{Y}$, and the natural map $d_Y\colon\check{Y}\to\widehat{Y}$, we can investigate $\mathrm{Ph}(X,Y)$, $\mathrm{SPh}(X,Y)$ and $\mathrm{Ph}(X,\check{Y})$ simultaneously (see Corollary 5.3, Propositions 5.4 and 5.7).

Notation Here and throughout, the subscript (0) denotes the rationalization of a nilpotent space or a nilpotent group.

Proposition 3.1 *Let Y, Y' be 1–connected CW-complexes. Suppose that $\varphi\colon Y\to Y'$ induces monomorphisms on the rational homotopy groups. Then:*

(1) *The homotopy fiber F of φ is nilpotent and the principal fibration*

$$(\Omega Y)_{(0)}\to(\Omega Y')_{(0)}\to F_{(0)}$$

is trivial.

(2) *F is rationally equivalent to $\prod_{i>0}K(\pi_{i+1}(Y')_{(0)}/\pi_{i+1}(Y)_{(0)},i)$.*

To prove Proposition 3.1, we require the following lemma.

Lemma 3.2 *Let $Z\xrightarrow{\psi}Z'\xrightarrow{p}B$ be a fibration sequence such that Z, Z', and B have the homotopy type of a CW-complex. Suppose that Z and Z' are simple rational spaces with trivial Postnikov invariants and that $\psi\colon Z\to Z'$ induces monomorphisms on the homotopy groups. Then the fibration $p\colon Z'\to B$ is trivial and the base B is a simple rational space with trivial Postnikov invariants.*

Proof First, we construct a homotopy left inverse σ of $\psi\colon Z\to Z'$. Fix homotopy equivalences $Z\simeq\prod K(\pi_i(Z),i)$ and $Z'\simeq\prod K(\pi_i(Z'),i)$ (see Remark 3.3). Choose left inverses s_i of the monomorphisms $\pi_i(\psi)\colon\pi_i(Z)\to\pi_i(Z')$ of $\mathbb{Q}$–modules and define the homotopy equivalence $h\colon Z\to Z$ to be the composite

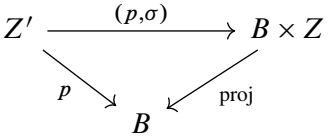
$$Z\xrightarrow{\psi}Z'\simeq\prod K(\pi_i(Z'),i)\xrightarrow{\prod K(s_i,i)}\prod K(\pi_i(Z),i)\simeq Z.$$

Then, the desired homotopy left inverse σ of $\psi\colon Z\to Z'$ is defined to be the composite

$$Z'\simeq\prod K(\pi_i(Z'),i)\xrightarrow{\prod K(s_i,i)}\prod K(\pi_i(Z),i)\simeq Z\xrightarrow{g}Z,$$

where g is a homotopy inverse of h.

Next, consider the morphism of fibrations



We can then easily see that this morphism is a trivialization of the fibration of interest. Since $(p,\sigma)\colon Z' \to B \times Z$ is a homotopy equivalence, B admits an H–space structure, and hence all the Postnikov invariants of B vanish. □

Proof of Proposition 3.1 Since F is nilpotent by [11, Proposition 4.4.1], we have the fibration

$$(\Omega Y)_{(0)} \xrightarrow{(\Omega\varphi)_{(0)}} (\Omega Y')_{(0)} \xrightarrow{\partial_{(0)}} F_{(0)}.$$

We can apply Lemma 3.2 to complete the proof. □

Remark 3.3 Since the product of infinitely many Eilenberg–Mac Lane complexes need not have the homotopy type of a CW-complex, the product in Proposition 3.1 should be interpreted as the weak product (see [32]). We can now operate in the category $\mathcal{CW}$. This interpretation is used throughout the paper.

Remark 3.4 In Proposition 3.1, the assumption that Y and Y' are 1–connected can be replaced by the assumption that $\varphi_\sharp\colon \pi_1(Y) \to \pi_1(Y')$ is a monomorphism. Then, by the universal covering argument, we observe that $\pi_0(F) \cong \pi_1(Y')/\pi_1(Y)$ and that each component of F is rationally equivalent to $\prod_{i>0} K(\pi_{i+1}(Y')_{(0)}/\pi_{i+1}(Y)_{(0)}, i)$.

Recall that for a connected space X and a map $\varphi\colon Y \to Y'$, the subset $P^\varphi(X,Y)$ of $[X,Y]$ is defined by

$$P^\varphi(X,Y) = \{[f] \in [X,Y] \mid \varphi_\sharp([f]) = 0\}.$$

Consider the principal $\Omega Y'$–fibration $\Omega Y' \to F \to Y$, where F is the homotopy fiber of φ. This fibration generates an exact sequence of pointed sets

$$[X, \Omega Y'] \to [X, F] \to P^\varphi(X,Y) \to 0$$

and the group $[X,\Omega Y']$ acts on $[X,F]$ in an obvious way.

Lemma 3.5 (orbit formula) *$P^\varphi(X,Y)$ is isomorphic to the orbit space $[X,F]/[X,\Omega Y']$.*

Proof See [11, Lemma 1.4.7]. □

A more useful formula for $P^{\varphi}(X,Y)$ can be established under certain conditions.

Proposition 3.6 (double coset formula) *Suppose that the map $\varphi\colon Y \to Y'$ is compatible with Proposition 3.1 and that the homotopy fiber F of φ is rational. Then we have the isomorphisms*

$$\begin{aligned} P^{\varphi}(X,Y) &\cong (\Omega\varphi_{(0)})_\sharp[X,\Omega Y_{(0)}]\backslash[X,\Omega Y'_{(0)}]/(\Omega r')_\sharp[X,\Omega Y'] \\ &\cong \varphi_{(0)\sharp}[\Sigma X, Y_{(0)}]\backslash[\Sigma X, Y'_{(0)}]/r'_\sharp[\Sigma X, Y'], \end{aligned}$$

where $r'\colon Y' \to Y'_{(0)}$ denotes the rationalization of Y'.

Proof Rationalizing φ yields the following commutative diagram:

$$\begin{array}{ccc} Y & \xrightarrow{r} & Y_{(0)} \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi_{(0)}} \\ Y' & \xrightarrow[r']{} & Y'_{(0)} \end{array}$$

Since F is rational, this diagram is a homotopy pullback. Thus the natural map

$$(r_\sharp, \varphi_\sharp)\colon [X,Y] \to [X,Y_{(0)}] \times_{[X,Y'_{(0)}]} [X,Y']$$

has a kernel that is isomorphic to

$$(\Omega\varphi_{(0)})_\sharp[X,\Omega Y_{(0)}]\backslash[X,\Omega Y'_{(0)}]/(\Omega r')_\sharp[X,\Omega Y']$$

(see [11, Proposition 2.2.2]). Thus, we merely require that if $\varphi_\sharp([f]) = 0$, then $r_\sharp([f]) = 0$. This is easily seen from Proposition 3.1(1). □

4 The functor $[\cdot, \Omega\, \cdot_{(0)}]$

By Proposition 3.6, if $[X,\Omega Y'_{(0)}] \cong [\Sigma X, Y'_{(0)}]$ is abelian, $P^{\varphi}(X,Y)$ has an abelian group structure. Motivated by this inference, we investigate the functor $[\cdot,\Omega\cdot_{(0)}] \cong [\Sigma\cdot, \cdot_{(0)}]$ and use the results of Scheerer [26] to find a necessary and sufficient condition under which the group $[X,\Omega Y_{(0)}] \cong [\Sigma X, Y_{(0)}]$ is abelian. We also study spaces with trivial rational cup and Whitehead products.

4.1 Necessary and sufficient condition that $[X, \Omega Y_{(0)}]$ is abelian

First we review the relevant background material (see [2, Chapter VI, Section 1]).

Let L be a nilpotent Lie algebra over $\mathbb{Q}$. Then L can be endowed with a group structure via the Baker–Campbell–Hausdorff formula

$$x \cdot y = x + y + \tfrac{1}{2}[x, y] + \tfrac{1}{12}[[x, y], y] + \tfrac{1}{12}[[y, x], x] + \cdots .$$

This group $(L, \cdot)$, denoted by $\exp L$, is a nilpotent rational group [8, Chapter 10].

Let (C, Δ) be a connected cocommutative graded coalgebra over $\mathbb{Q}$, and $(\pi, [\,,\,])$ a reduced graded Lie algebra over $\mathbb{Q}$. The $\mathbb{Q}$–module $\mathrm{Hom}_{\mathbb{Q}}(C, \pi)$ of homomorphisms of graded $\mathbb{Q}$–modules is a (nongraded) Lie algebra over $\mathbb{Q}$ with the Lie bracket defined by

$$[f, g]\colon C \xrightarrow{\Delta} C \otimes C \xrightarrow{f \otimes g} \pi \otimes \pi \xrightarrow{[\,,\,]} \pi.$$

If $C_n = 0$ for sufficiently large n, $\mathrm{Hom}_{\mathbb{Q}}(C, \pi)$ is a nilpotent Lie algebra. Thus we define the group

$$\exp \mathrm{Hom}_{\mathbb{Q}}(C, \pi) = \varprojlim_{n} \exp \mathrm{Hom}_{\mathbb{Q}}(C_{\leq n}, \pi)$$

as the inverse limit of nilpotent groups, where $C_{\leq n}$ is the subcoalgebra of C of all elements of degree $\leq n$.

Recall that the rational homology $H_*(X; \mathbb{Q})$ of an arcwise connected space X is a connected cocommutative graded coalgebra over $\mathbb{Q}$ whose coproduct is induced by the diagonal map of X. Recall also that for a 1–connected space Y, the rational homotopy group $\pi_*(\Omega Y) \otimes \mathbb{Q}$ equipped with the Samelson product is a reduced graded Lie algebra over $\mathbb{Q}$, and that the Samelson product of ΩY is the Whitehead product of Y.

Let 1-$\mathcal{CW}$ denote the full subcategory of $\mathcal{CW}$ consisting of 1–connected CW-complexes, and let Gr denote the category of groups.

Proposition 4.1 *The functor* $[\,\cdot\,, \Omega\,\cdot_{(0)}]\colon \mathcal{CW}^{\mathrm{op}} \times 1\text{-}\mathcal{CW} \to \mathrm{Gr}$ *is naturally isomorphic to the functor* $\exp \mathrm{Hom}_{\mathbb{Q}}(H_*(\,\cdot\,; \mathbb{Q}),\ \pi_*(\Omega\,\cdot\,) \otimes \mathbb{Q})$.

Proof The proof follows from Proposition 1 in [26, Section 0.1], the remark on [26, page 71] and the natural isomorphism $H_*(\Omega\,\cdot\,; \mathbb{Q}) \cong U\pi_*(\Omega\,\cdot_{(0)})$ of graded Hopf algebras, where U denotes the universal enveloping algebra functor [22, Appendix]. □

We introduce the following full subcategory of $\mathcal{CW}^{\mathrm{op}} \times 1\text{-}\mathcal{CW}$.

Definition 4.2 Let $\mathcal{Q}^{\sim}$ be the full subcategory of $\mathcal{CW}^{\mathrm{op}} \times 1\text{-}\mathcal{CW}$ whose members (X, Y) satisfy the condition (Q) (see Definition 2.1).

Theorem 4.3 (1) *Let (X,Y) be an object of $\mathcal{CW}^{\mathrm{op}}\times 1\text{-}\mathcal{CW}$. Then (X,Y) is in $\mathcal{Q}^{\sim}$ if and only if $[X,\Omega Y_{(0)}]$ is abelian.*

(2) *The functor $[\,\cdot\,,\Omega\,\cdot_{(0)}]=[\Sigma\,\cdot\,,\,\cdot_{(0)}]\colon \mathcal{Q}^{\sim}\to \mathrm{Gr}$ is naturally isomorphic to the functor $\prod_{i>0}H^i(\,\cdot\,;\pi_{i+1}(\,\cdot\,)_{(0)})$. In particular, $[\,\cdot\,,\Omega\,\cdot_{(0)}]=[\Sigma\,\cdot\,,\,\cdot_{(0)}]$ on $\mathcal{Q}^{\sim}$ is a $\mathbb{Q}$–module-valued functor.*

Proof (1) By definition, we have

$$\exp \mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q})=\varprojlim_n \exp \mathrm{Hom}_{\mathbb{Q}}(H_{\le n}(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q}).$$

Thus, Proposition 4.1 and the Mal'cev correspondence [8, page 118] imply that the group $[X,\Omega Y_{(0)}]$ is abelian if and only if the Lie algebra $\mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q})$ is abelian. Hence, we show that (X,Y) is in $\mathcal{Q}^{\sim}$ if and only if the Lie algebra $\mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q})$ is abelian.

($\Rightarrow$) We observe that the bracket $[\,,\,]$ of $\mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q})$ is trivial. It is sufficient to observe that the composite

$$\begin{aligned}[f,g]_{p,q}\colon H_n(X;\mathbb{Q})&\xrightarrow{\Delta_{p,q}} H_p(X;\mathbb{Q})\otimes H_q(X;\mathbb{Q})\\ &\xrightarrow{f_p\otimes g_q}(\pi_p(\Omega Y)\otimes\mathbb{Q})\otimes(\pi_q(\Omega Y)\otimes\mathbb{Q})\xrightarrow{[\,,\,]}\pi_n(\Omega Y)\otimes\mathbb{Q}\end{aligned}$$

is zero for any f,g and any p,q,n with $p+q=n$, where $\Delta_{p,q}$ is the composition of Δ and the projection onto $H_p(X;\mathbb{Q})\otimes H_q(X;\mathbb{Q})$. It is easily verified that $\cup\colon H^p(X;\mathbb{Q})\otimes H^q(X;\mathbb{Q})\to H^n(X;\mathbb{Q})$ is zero if and only if $\Delta_{p,q}$ is zero. Thus we have $[f,g]_{p,q}=0$.

($\Leftarrow$) Suppose that

$$\cup\colon H^i(X;\mathbb{Q})\otimes H^j(X;\mathbb{Q})\to H^{i+j}(X;\mathbb{Q})$$

and

$$[\,,\,]\colon(\pi_{i+1}(Y)\otimes\mathbb{Q})\otimes(\pi_{j+1}(Y)\otimes\mathbb{Q})\to\pi_{i+j+1}(Y)\otimes\mathbb{Q}$$

are nonzero for some i and j. Then it is not difficult to find

$$f,g\in\mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q})$$

with $[f,g]\neq 0$.

(2) As is shown in the proof of part (1), the Lie algebra

$$\mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}),\pi_*(\Omega Y)\otimes\mathbb{Q})$$

is abelian for $(X, Y) \in \mathcal{Q}^{\sim}$. Thus we have

$$\begin{aligned}[X, \Omega Y_{(0)}] &\cong \exp \mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}), \pi_*(\Omega Y) \otimes \mathbb{Q}) \\ &= \mathrm{Hom}_{\mathbb{Q}}(H_*(X;\mathbb{Q}), \pi_*(\Omega Y) \otimes \mathbb{Q}) \\ &= \prod_{i>0} H^i(X; \pi_{i+1}(Y)_{(0)}).\end{aligned}$$ □

4.2 C_0–spaces and W_0–spaces

The spaces defined below are important in the context of Theorem 4.3.

Definition 4.4 (1) X is called a C_0–*space* if the cup product on the reduced cohomology $\widetilde{H}^*(X;\mathbb{Q})$ is trivial.

(2) Y is called a W_0–*space* if the Whitehead product on the higher rational homotopy groups $\bigoplus_{i>1} \pi_i(Y) \otimes \mathbb{Q}$ is trivial.

See [9, Proposition 5.5] and [10] for properties characterizing C_0–spaces. Spaces with trivial Whitehead product have been investigated by Siegel [27], who referred to them as W–spaces.

Example 4.5 The classes of C_0– and W_0–spaces properly contain co-H_0–spaces and H_0–spaces, respectively. Many finite Grassmannians (eg $\mathbb{C}P^n$ with $n > 1$) are W_0–spaces but not H_0–spaces.

A pair $(X, Y) \in \mathcal{CW}^{\mathrm{op}} \times 1\text{-}\mathcal{CW}$ is in $\mathcal{Q}^{\sim}$ if X is a C_0–space or Y is a W_0–space. Let us see that $\mathcal{Q}^{\sim}$ contains many pairs (X, Y) in which X is not a C_0–space and Y is not a W_0–space.

Given a pair (X, Y) of connected spaces, define $D_{X,Y}$ by

$$D_{X,Y} = \{i > 0 \mid H^i(X;\mathbb{Q}) \neq 0 \text{ and } \pi_{i+1}(Y) \otimes \mathbb{Q} \neq 0\}.$$

Then the condition (Q) is equivalent to the same vanishing condition for any $i, j \in D_{X,Y}$ with $i + j \in D_{X,Y}$. Note that if Y is a loop space, a classifying space, or a homogeneous space of a Lie group, then $D_{X,Y}$ is finite.

Example 4.6 (1) (X, S^n) is in $\mathcal{Q}^{\sim}$ if and only if either n is odd or the rational cup product

$$\cup\colon H^{n-1}(X;\mathbb{Q}) \otimes H^{n-1}(X;\mathbb{Q}) \to H^{2n-2}(X;\mathbb{Q})$$

is trivial.

(2) Let X be a connected CW-complex with $H_*(X;\mathbb{Q}) = H_{\mathrm{ev}}(X;\mathbb{Q})$. Then (X,Y) is in $\mathcal{Q}^{\sim}$ if the Whitehead product on $\pi_{\mathrm{odd}}(Y)\otimes\mathbb{Q}$ is trivial. In particular, if Y is a homogeneous space of a Lie group, then (X,Y) is in $\mathcal{Q}^{\sim}$.

Proof (1) Recall that

$$\pi_*(S^n)\otimes\mathbb{Q} = \begin{cases} \mathbb{Q}\cdot 1_n & \text{if } n \text{ is odd,} \\ \mathbb{Q}\cdot 1_n \oplus \mathbb{Q}\cdot[1_n,1_n] & \text{if } n \text{ is even,} \end{cases}$$

where 1_n denotes the homotopy class of the identity map of S^n. Then the proof follows immediately from the definition of $\mathcal{Q}^{\sim}$.

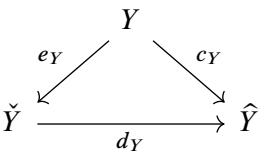
(2) The first clause is obvious. To prove the latter, we set $Y = G/H$ and consider the fiber sequence $H \xrightarrow{i} G \xrightarrow{p} G/H$. Since $\pi_*(H)\otimes\mathbb{Q} = \pi_{\mathrm{odd}}(H)\otimes\mathbb{Q}$, the map $p_\sharp\colon \pi_{\mathrm{odd}}(G)\otimes\mathbb{Q} \to \pi_{\mathrm{odd}}(G/H)\otimes\mathbb{Q}$ is surjective. Since the Whitehead product on $\pi_{\mathrm{odd}}(G)\otimes\mathbb{Q}$ is trivial, the Whitehead product on $\pi_{\mathrm{odd}}(G/H)\otimes\mathbb{Q}$ is also trivial. □

5 Applications to phantom maps

In this section, we apply the results of Sections 3 and 4 to phantom maps, and also prove Theorem 2.3.

The source is considered to be a connected CW-complex and the target is a nilpotent CW-complex of finite type over $\mathbb{Z}_l$, where $\mathbb{Z}_l$ is the localization of $\mathbb{Z}$ at a nonempty set l of primes. (A nilpotent space of finite type over $\mathbb{Z}_l$ is defined in [11, page 79], referred to as an $f\mathbb{Z}_l$–nilpotent space.) A nilpotent space of finite type over $\mathbb{Z}$ defines a nilpotent space of finite type in the usual sense.

For a nilpotent space Y of finite type over $\mathbb{Z}_l$, the profinite l–completion $\widehat{Y}$ and the l–local expansion $\check{Y}$ are defined by $\widehat{Y} = \prod_{p\in l}\widehat{Y}_p$ and $\check{Y} = \prod_{p\in l} Y_{(p)}$, respectively, where $\widehat{Y}_p$ and $Y_{(p)}$ are the p–profinite completion and the p–localization of Y respectively [30]. Thus, we can establish a commutative diagram of natural transformations

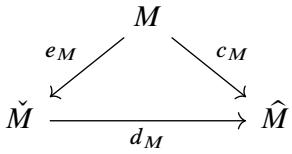


Similarly, for a finitely generated $\mathbb{Z}_l$–module M, the profinite l–completion $\widehat{M}$ and the l–local expansion $\check{M}$ are defined by $\widehat{M} = \prod_{p\in l}\widehat{M}_p$ and $\check{M} = \prod_{p\in l} M_{(p)}$,

respectively. Then there exist natural isomorphisms

$$\widehat{M} \cong M \otimes \widehat{\mathbb{Z}}_l \quad \text{and} \quad \check{M} \cong M \otimes \check{\mathbb{Z}}_l,$$

and the commutative diagram of natural transformations is given as



The following lemma generalizes Meier and Zabrodsky's result [34, page 137] (see also [14, Theorem 5.1] and [25, Section 1]).

Proposition 5.1 *Let X be a connected CW-complex and Y a nilpotent CW-complex of finite type over $\mathbb{Z}_l$. Then a map $f\colon X \to Y$ is a phantom map if and only if the composite $X \xrightarrow{f} Y \xrightarrow{c_Y} \widehat{Y}$ is null-homotopic.*

Proof ($\Rightarrow$) The proof is outlined in [34, page 137].

($\Leftarrow$) Let p be a prime in l. Under the assumption on Y, we have $\pi_i(\widehat{Y}_p) \cong \pi_i(Y)\widehat{\ }_p$ (see the proof of [31, Theorem 3.9]). Thus the p–profinite completion of Y coincides with the completion of Y at p in the context of [11, Definition 10.2.3] (see [11, Theorems 11.1.1 and 11.1.2]). Hence, the reverse implication immediately follows from [11, Theorem 13.1.1]. □

Remark 5.2 (1) Recall that we have assumed that $l \neq \varnothing$ (ie $\mathbb{Z}_l \subsetneq \mathbb{Q}$). In the case where $l = \varnothing$, $\widehat{Y}$ is a singleton, and hence Proposition 5.1 does not hold. But we know that any phantom map to a nilpotent space of finite type over $\mathbb{Q}$ is trivial (see [20, Theorem 1] or [29, Theorem 3.3(b)]).

(2) As implemented by Meier and Zabrodsky [34] and Roitberg and Touhey [25], our approach to phantom maps is based on Proposition 5.1. Though another characterization of phantom maps, Theorem B(b) in [34], is also easily generalizable to an arbitrary nilpotent source, it is not relevant to our approach. See [14] for the other approaches to phantom maps.

(3) Hereafter, for simplicity, we assume that the target Y is nilpotent of finite type. However, all of the results remain valid under appropriate modification, even if the target is regarded as nilpotent of finite type over $\mathbb{Z}_l$ rather than nilpotent of finite type.

Recall the definition of $P^{\varphi}(X,Y)$ in Section 3.

Corollary 5.3 *Let X be a connected CW-complex and Y a nilpotent CW-complex of finite type. Then*

$$\begin{aligned}\mathrm{Ph}(X,Y) &= P^{c_Y}(X,Y),\\ \mathrm{SPh}(X,Y) &= P^{e_Y}(X,Y),\\ \mathrm{Ph}(X,\check{Y}) &= P^{d_Y}(X,\check{Y}).\end{aligned}$$

Proof The first equality follows immediately from Proposition 5.1. Since $d_Y = \prod_p c_{Y_{(p)}}$, the third equality also follows from Proposition 5.1. Futhermore, since $d_Y e_Y = c_Y$, the second equality follows from the definition of $\mathrm{SPh}(X,Y)$ and the first equality. □

Using Corollary 5.3, we apply the results of Section 3 to the study of phantom maps. Let F_Y, F'_Y and F''_Y denote the homotopy fibers of c_Y, e_Y and d_Y, respectively.

Proposition 5.4 *Let Y be a 1–connected CW-complex of finite type.*

(1) *F_Y, F'_Y and F''_Y are homotopically equivalent, respectively, to the products*

$$\begin{aligned}&\prod_{i>0} K(\pi_{i+1}(Y)\otimes\hat{\mathbb{Z}}/\mathbb{Z}, i),\\ &\prod_{i>0} K(\pi_{i+1}(Y)\otimes\check{\mathbb{Z}}/\mathbb{Z}, i),\\ &\prod_{i>0} K(\pi_{i+1}(Y)\otimes\hat{\mathbb{Z}}/\check{\mathbb{Z}}, i).\end{aligned}$$

(2) *The natural sequence*

$$F'_Y \xrightarrow{i_Y} F_Y \xrightarrow{p_Y} F''_Y$$

is a trivial fibration sequence.

Proof (1) Note that

$$\begin{aligned}\pi_i(\hat{Y}) &= \pi_i(Y)\otimes\hat{\mathbb{Z}},\\ \pi_i(\check{Y}) &= \pi_i(Y)\otimes\check{\mathbb{Z}}\end{aligned}$$

and that $\hat{\mathbb{Z}}/\mathbb{Z}$, $\check{\mathbb{Z}}/\mathbb{Z}$ and $\hat{\mathbb{Z}}/\check{\mathbb{Z}}$ are $\mathbb{Q}$–modules. Then the homotopy types of F_Y, F'_Y and F''_Y are determined by Proposition 3.1.

(2) Since $c_Y = d_Y e_Y$, we have the natural fibration sequence

$$F'_Y \to F_Y \to F''_Y.$$

This sequence is trivial as per Lemma 3.2. □

The triviality of the fibration sequence $F'_Y \to F_Y \to F''_Y$ is exploited in the calculation of $\mathrm{SPh}(X, Y)$ in Section 6.

Remark 5.5 For a nilpotent CW-complex Y of finite type, the homotopy fibers F_Y, F'_Y, and F''_Y are determined by Proposition 5.4 and Remark 3.4. In the case of finite $\pi_1(Y)$, the homotopy type of F_Y is determined by the method described by Roitberg and Touhey [25]. Their proof is based on the results of Meier [20] and Steiner [29], whose works are mentioned in Remark 5.2(1).

Remark 5.6 From the definition of a phantom map, it is easily seen that $\mathrm{Ph}(X, Y) = \mathrm{Ph}(X, \widetilde{Y})$ holds (without the finite-type assumption imposed on Y), where $\widetilde{Y}$ denotes the universal cover of Y. Consequently, our approach is valid under the weaker assumption that the target has finitely generated higher homotopy groups. In [17, page 367], such a space is called a *finite-type target*. The classifying space of any compact Lie group is a finite-type target.

The following proposition describes a special case of Proposition 3.6.

Proposition 5.7 (double coset formula) *Let X be a connected CW-complex, and Y a nilpotent CW-complex of finite type. Then there exist natural isomorphisms*

$$\begin{aligned}
\mathrm{Ph}(X, Y) &\cong (\Omega c_{(0)})_\sharp [X, \Omega Y_{(0)}] \backslash [X, \Omega \widehat{Y}_{(0)}] / (\Omega \widehat{\ }r)_\sharp [X, \Omega \widehat{Y}] \\
&\cong c_{(0)\sharp} [\Sigma X, Y_{(0)}] \backslash [\Sigma X, \widehat{Y}_{(0)}] / \widehat{\ }r_\sharp [\Sigma X, \widehat{Y}], \\
\mathrm{SPh}(X, Y) &\cong (\Omega e_{(0)})_\sharp [X, \Omega Y_{(0)}] \backslash [X, \Omega \check{Y}_{(0)}] / (\Omega \check{\ }r)_\sharp [X, \Omega \check{Y}] \\
&\cong e_{(0)\sharp} [\Sigma X, Y_{(0)}] \backslash [\Sigma X, \check{Y}_{(0)}] / \check{\ }r_\sharp [\Sigma X, \check{Y}], \\
\mathrm{Ph}(X, \check{Y}) &\cong (\Omega d_{(0)})_\sharp [X, \Omega \check{Y}_{(0)}] \backslash [X, \Omega \widehat{Y}_{(0)}] / (\Omega \widehat{\ }r)_\sharp [X, \Omega \widehat{Y}] \\
&\cong d_{(0)\sharp} [\Sigma X, \check{Y}_{(0)}] \backslash [\Sigma X, \widehat{Y}_{(0)}] / \widehat{\ }r_\sharp [\Sigma X, \widehat{Y}],
\end{aligned}$$

where $\widehat{\ }r$ and $\check{\ }r$ denote the rationalization of $\widehat{Y}$ and $\check{Y}$ respectively.

Proof By Remark 5.6, it is sufficient to prove that Proposition 5.7 holds when Y is 1–connected. This follows from Proposition 3.6, Corollary 5.3 and Proposition 5.4. □

Remark 5.8 Proposition 5.7 is analogous to the double coset formulas describing local and adelic genera of a group and a space; see [11, pages 147, 169, 242 and 262] and [13].

The following corollary generalizes McGibbon's result [12, page 266].

Corollary 5.9 *Let X be a connected CW-complex and Y a nilpotent CW-complex of finite type. Then the following natural exact sequence of pointed sets exists:*

$$0 \to \mathrm{SPh}(X,Y) \to \mathrm{Ph}(X,Y) \xrightarrow{e_\sharp} \mathrm{Ph}(X,\check{Y}) \to 0.$$

Proof According to Proposition 5.7, both $\mathrm{Ph}(X,Y)$ and $\mathrm{Ph}(X,\check{Y})$ are quotient sets of $[X,(\Omega\widehat{Y})_{(0)}]$, implying that $e_\sharp$ is surjective. The remainder of the proof follows from the definition of $\mathrm{SPh}(X,Y)$. □

We now closely study natural group structures on phantom maps.

Let $0 \to \check{\mathbb{Z}}/\mathbb{Z} \xrightarrow{\iota} \widehat{\mathbb{Z}}/\mathbb{Z} \xrightarrow{\varpi} \widehat{\mathbb{Z}}/\check{\mathbb{Z}} \to 0$ denote the canonical exact sequence of $\mathbb{Q}$–modules.

Proposition 5.10 *Let (X,Y) be an object of $\mathcal{Q}$. Then there are natural isomorphisms*

$$[X,F_Y] \cong \prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\widehat{\mathbb{Z}}/\mathbb{Z}),$$
$$[X,F'_Y] \cong \prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\check{\mathbb{Z}}/\mathbb{Z}),$$
$$[X,F''_Y] \cong \prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\widehat{\mathbb{Z}}/\check{\mathbb{Z}}).$$

Under these isomorphisms, $i_{Y\sharp}\colon [X,F'_Y]\to[X,F_Y]$ and $p_{Y\sharp}\colon [X,F_Y]\to[X,F''_Y]$ are identified with $\prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\iota)$ and $\prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes\varpi)$, respectively.

Proof By Remarks 5.5 and 5.6, we may assume that Y is 1–connected. For a simultaneous treatment of $c_Y\colon Y\to\widehat{Y}$, $e_Y\colon Y\to\check{Y}$ and $d_Y\colon\check{Y}\to\widehat{Y}$, we denote one mapping by $\varphi=\varphi_Y\colon \alpha Y\to\beta Y$ and let F_Y^φ be its homotopy fiber. Then $P^\varphi(X,\alpha Y)$ coincides with $\mathrm{Ph}(X,Y)$, $\mathrm{SPh}(X,Y)$ and $\mathrm{Ph}(X,\check{Y})$ in the respective cases. Let α, β and φ also denote the corresponding functors and natural transformations for finitely generated $\mathbb{Z}$–modules.

Since the homotopy fiber F_Y^φ is rational (Proposition 5.4), Proposition 3.1 implies that the principal fibration

$$(\Omega\alpha Y)_{(0)} \to (\Omega\beta Y)_{(0)} \to F_Y^\varphi$$

is trivial; hence $[X, F_Y^\varphi]$ is isomorphic to the orbit space $[X, (\Omega\alpha Y)_{(0)}]\backslash[X, (\Omega\beta Y)_{(0)}]$ by [11, Lemma 1.4.7]. Note that $1\times\alpha$ and $1\times\beta$ preserve condition (Q). Then, from Theorem 4.3,

$$(\Omega\varphi_Y)_{(0)\sharp}\colon [X, (\Omega\alpha Y)_{(0)}] \to [X, (\Omega\beta Y)_{(0)}]$$

is identified with

$$\prod H^i(X; \pi_{i+1}(Y)\otimes\varphi_{\mathbb{Z}_{(0)}})\colon \prod H^i(X; \pi_{i+1}(Y)\otimes(\alpha\mathbb{Z})_{(0)}) \to \prod H^i(X; \pi_{i+1}(Y)\otimes(\beta\mathbb{Z})_{(0)}).$$

Note that

$$(\beta\mathbb{Z})_{(0)}/(\alpha\mathbb{Z})_{(0)} = (\beta\mathbb{Z}/\alpha\mathbb{Z})_{(0)} = \beta\mathbb{Z}/\alpha\mathbb{Z}.$$

Then we obtain the natural isomorphism

$$[X, F_Y^\varphi] \cong \prod_{i>0} H^i(X; \pi_{i+1}(Y)\otimes\beta\mathbb{Z}/\alpha\mathbb{Z}).$$

Theorem 4.3 also implies that $i_{Y\sharp}$ and $p_{Y\sharp}$ are identified. □

Proof of Theorem 2.3 By Remarks 5.5 and 5.6, we assume that Y is 1–connected, and use the notations introduced in the proof of Proposition 5.10; note that $\alpha = 1$ since $\varphi_Y = c_Y$ or e_Y.

(1) Since, by Theorem 4.3, $[X, (\Omega\beta Y)_{(0)}]$ is a $\mathbb{Q}$–module, we have

$$\begin{aligned} P^\varphi(X,Y) &\cong (\Omega\varphi)_{(0)\sharp}[X, (\Omega Y)_{(0)}]\backslash[X, (\Omega\beta Y)_{(0)}]/(\Omega r')_\sharp[X, \Omega\beta Y] \\ &\cong [X, (\Omega\beta Y)_{(0)}]/(\Omega\varphi)_{(0)\sharp}[X, (\Omega Y)_{(0)}] + (\Omega r')_\sharp[X, \Omega\beta Y] \\ &\cong [X, F_Y^\varphi]/\text{ the image of } [X, \Omega\beta Y] \end{aligned}$$

by the proof of Proposition 5.10, where r' denotes the rationalization of βY. Thus, a natural, divisible, abelian group structure exists on $P^\varphi(X,Y)$.

We can easily see that the inclusion $\mathrm{SPh}(X,Y) \hookrightarrow \mathrm{Ph}(X,Y)$ is a homomorphism of abelian groups (Proposition 5.10).

(2) We treat only the case of $\mathrm{Im}\,\mathrm{Ph}(f,g)$; the same argument applies to $\mathrm{Im}\,\mathrm{SPh}(f,g)$.

By Proposition 5.7, we have the morphism of exact sequences of pointed sets

$$\begin{array}{ccccc} [K, \Omega\widehat{L}_{(0)}] & \longrightarrow & \mathrm{Ph}(K, L) & \longrightarrow & 0 \\ {\scriptstyle [f,\Omega\widehat{g}_{(0)}]}\downarrow & & {\scriptstyle \mathrm{Ph}(f,g)}\downarrow & & \\ [X, \Omega\widehat{Y}_{(0)}] & \longrightarrow & \mathrm{Ph}(X, Y) & \longrightarrow & 0 \end{array}$$

Recall that the map $[X, \Omega\widehat{Y}_{(0)}] \to \mathrm{Ph}(X, Y)$ is an epimorphism of abelian groups which is the composite of the natural epimorphisms $[X, \Omega\widehat{Y}_{(0)}] \to [X, F_Y]$ and $[X, F_Y] \to \mathrm{Ph}(X, Y)$ (see the proof of part (1) above). By Proposition 4.1, Theorem 4.3, and the proof of Proposition 5.10, the image of the composite

$$[K, \Omega\widehat{L}_{(0)}] \xrightarrow{[f,\Omega\widehat{g}_{(0)}]} [X, \Omega\widehat{Y}_{(0)}] \to [X, F_Y]$$

is identified with $\prod_{i>0} \mathrm{Im}\, H^i(f; \pi_{i+1}(g) \otimes \widehat{\mathbb{Z}}/\mathbb{Z})$, which shows that $\mathrm{Im}\,\mathrm{Ph}(f, g)$ is a divisible abelian subgroup of $\mathrm{Ph}(X, Y)$.

(3) We treat first the case of $\mathrm{Ph}(X, Y)$.

Define the maps ι and ∂ by the long fibration sequence

$$\cdots \to \Omega\widehat{Y} \xrightarrow{\partial} F_Y \xrightarrow{\iota} Y \xrightarrow{c_Y} \widehat{Y}.$$

Consider the commutative diagram

$$\begin{array}{ccc} [X, \Omega\widehat{Y}_{(0)}] & & \\ {\scriptstyle \partial_{(0)\sharp}}\downarrow & & \\ [X, F_Y] & \xrightarrow{\iota_\sharp} & [X, Y] \\ {\scriptstyle \pi}\downarrow & \nearrow & \\ \mathrm{Ph}(X, Y) & & \end{array}$$

where π is the identification map by the action of the group $[X, \Omega\widehat{Y}]$. Since the group structure on $\mathrm{Ph}(X, Y)$ is characterized by the property that the composite $\pi\partial_{(0)\sharp}$ is an epimorphism of groups (see the proof of part (1) above), it is sufficient to see that the composite

$$[X, \Omega\widehat{Y}_{(0)}] \xrightarrow{\partial_{(0)\sharp}} [X, F_Y] \xrightarrow{\iota_\sharp} [X, Y] \tag{5-1}$$

preserves multiplication.

First, suppose that X is a co-H–space. Then the result follows from [32, Theorem 5.21 in Chapter III].

Second, suppose that Y is an H–space. Then, F_Y admits an H–structure such that $\partial\colon \Omega\widehat{Y} \to F_Y$ and $\iota\colon F_Y \to Y$ are H–maps (see the proof of [28, Theorem 9.1] and [33, page 15]). Thus, we see that the composite

$$\Omega\widehat{Y}_{(0)} \xrightarrow{\partial_{(0)}} F_Y \xrightarrow{\iota} Y$$

is an H–map, and hence that the composite (5-1) preserves multiplication.

The result for $\mathrm{SPh}(X, Y)$ follows from that for $\mathrm{Ph}(X, Y)$ since $\mathrm{SPh}(X, Y)$ is a subgroup of $\mathrm{Ph}(X, Y)$ for $(X, Y) \in \mathcal{Q}$. □

Remark 5.11 From the proof of Theorem 2.3, it is obvious that a result similar to Theorem 2.3 holds for $\mathrm{Ph}(X, \check{Y})$.

Proposition 5.12 *Let (X, Y) be an object of $\mathcal{Q}$. Then there exists a natural commutative diagram of groups*

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \\
[X, \Omega\check{Y}] & \longrightarrow & [X, F'_Y] & \longrightarrow & \mathrm{SPh}(X, Y) & \longrightarrow & 0 \\
\downarrow & & \downarrow{\scriptstyle i_{Y\sharp}} & & \downarrow & & \\
[X, \Omega\widehat{Y}] & \longrightarrow & [X, F_Y] & \longrightarrow & \mathrm{Ph}(X, Y) & \longrightarrow & 0 \\
{\scriptstyle 1}\downarrow & & \downarrow{\scriptstyle p_{Y\sharp}} & & \downarrow{\scriptstyle e_{Y\sharp}} & & \\
[X, \Omega\widehat{Y}] & \longrightarrow & [X, F''_Y] & \longrightarrow & \mathrm{Ph}(X, \check{Y}) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & &
\end{array}$$

where all rows and all columns except the left column are exact; $[X, F_Y]$, $[X, F'_Y]$ and $[X, F''_Y]$ are regarded as groups by the natural isomorphisms in Proposition 5.10.

Proof The exactness of the rows follows immediately from the proof of Theorem 2.3(1) and Remark 5.11. The exactness of the right and the middle columns follows from Corollary 5.9 and Proposition 5.10 respectively. The commutativity is obvious. □

6 Calculations of $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$

In this section, we prove Proposition 2.5, Theorem 2.7 and Corollaries 2.8–2.10.

6.1 $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ in the case of trivial $[X,\Omega\widehat{Y}]$

In this subsection, we prove Proposition 2.5, generalizing Theorem A(2) not only for $\mathrm{Ph}(X,Y)$, but also for $\mathrm{SPh}(X,Y)$. Though analogous results in all cases can be found for $\mathrm{Ph}(X,\check{Y})$, we omit their statements and proofs, since they follow from the results for $\mathrm{Ph}(X,Y)$, Remark 5.2(3), and the isomorphism $\mathrm{Ph}(X,\check{Y})\cong\prod_p \mathrm{Ph}(X,Y_{(p)})$.

First, let us consider pairs (X,Y) with $[X,\Omega\widehat{Y}]$ locally finite (and, in particular, with $[X,\Omega\widehat{Y}]$ trivial).

Proposition 6.1 *Let X be a connected CW-complex and Y a nilpotent CW-complex of finite type. If the group $[X,\Omega\widehat{Y}]$ is locally finite, we have the isomorphisms*

$$\mathrm{Ph}(X,Y)\cong\prod_{i>0}H^i(X;\pi_{i+1}(Y)\otimes\widehat{\mathbb{Z}}/\mathbb{Z}),$$
$$\mathrm{SPh}(X,Y)\cong\prod_{i>0}H^i(X;\pi_{i+1}(Y)\otimes\check{\mathbb{Z}}/\mathbb{Z}).$$

Further, if (X,Y) is in $\mathcal{Q}$, these isomorphisms can be taken as natural isomorphisms of abelian groups.

Proof We may assume that Y is 1–connected by Remark 5.6.

The case of $\mathrm{Ph}(X,Y)$ By Proposition 5.7, we have

$$\mathrm{Ph}(X,Y)\cong(\Omega c_{(0)})_\sharp[X,\Omega Y_{(0)}]\backslash[X,\Omega\widehat{Y}_{(0)}]/(\Omega\widehat{r})_\sharp[X,\Omega\widehat{Y}].$$

Define the map

$$\psi_k\colon\Omega\widehat{Y}_{(0)}\to\Omega\widehat{Y}_{(0)},\quad \psi_k(\alpha)=\alpha^k\qquad\text{for } k\geq 1.$$

Since ψ_k is a homotopy equivalence, $(\Omega\widehat{r})_\#[X,\Omega\widehat{Y}]=0$, and $\mathrm{Ph}(X,Y)$ is given by

$$\mathrm{Ph}(X,Y)\cong(\Omega c)_{(0)\sharp}[X,\Omega Y_{(0)}]\backslash[X,\Omega\widehat{Y}_{(0)}]\cong[X,F_Y]$$

by Proposition 3.1. The desired isomorphism is implied by Proposition 5.4. If (X,Y) is in $\mathcal{Q}$, the isomorphism can be taken as a natural isomorphism of abelian groups by Proposition 5.10 and the proof of Theorem 2.3(1).

The case of SPh(X, Y) By Corollary 5.3 and Lemma 3.5, we have the commutative diagram

$$\begin{array}{ccc} [X, F'_Y] & \longrightarrow & [X, F_Y] \\ \downarrow & & \downarrow \\ \mathrm{SPh}(X, Y) & \longrightarrow & \mathrm{Ph}(X, Y) \end{array}$$

whose vertical arrows are quotient maps. As previously determined, the right vertical arrow is bijective. Since the upper arrow is injective by Proposition 5.4(2), the left vertical arrow must be bijective. The above argument completes the proof. □

Remark 6.2 The following notes relate to Proposition 6.1.

(1) Meier [20] and Zabrodsky [34] identified the set $\mathrm{Ph}(X, Y)$ when investigating the case of $[X, \Omega\widehat{Y}] = 0$ (see [14, Section 5]). In [25, Theorem 1.5], $\mathrm{Ph}(X, Y)$ is identified in the more general case of finite $[X, \Omega\widehat{Y}]$.

(2) Naturally, the local finiteness of $[X, \Omega\check{Y}]$ implies the same results for $\mathrm{SPh}(X, Y)$. However, the above assertion regarding $\mathrm{SPh}(X, Y)$ is more useful since important vanishing results appear in the form $[X, \Omega\widehat{Y}] = 0$, as demonstrated in the next proposition.

Proposition 6.3 (1) (generalization of Miller [21, Theorem A]) *Let A be in $\mathcal{A}$ and let B be a nilpotent finite complex. Then $\mathrm{map}_*(A, \widehat{B})$ is weakly contractible.*

(2) (generalization of Anderson and Hodgkin [1, Theorem I]) *Let A be in $\mathcal{A}'$ and let B be (the identity component of) a space in $\mathcal{B}'$. Then $\mathrm{map}_*(A, \widehat{B})$ is weakly contractible.*

Proof (1) This follows immediately from the generalizations of Miller's theorem [21] developed by Zabrodsky [34], Friedlander and Mislin [4] and McGibbon [15]; the finite-type condition on the Postnikov space was removed in [15, page 3244].

(2) Here, we may assume that A is a $1\frac{1}{2}$–connected Postnikov space of finite type. We prove the weak contractibility of $\mathrm{map}_*(A, \widehat{B})$ in four steps.

Step 1 We prove the result for $B = BU$; similar proofs can be found for BO and $B\,\mathrm{Sp}$.

Suppose that A is an Eilenberg–Mac Lane space $K(\pi, n)$ with π finite and $n = 2$ or π finitely generated, with $n \geq 3$. The rationalization $\pi \to \pi_{(0)}$ induces the following

morphism of fibration sequences:

$$\begin{array}{ccc}
\mathrm{map}_*(K(\pi,n), F_{BU}) & \xleftarrow{\varphi'} & \mathrm{map}_*(K(\pi_{(0)},n), F_{BU}) \\
\downarrow & & \downarrow \\
\mathrm{map}_*(K(\pi,n), BU) & \xleftarrow{\varphi} & \mathrm{map}_*(K(\pi_{(0)},n), BU) \\
\downarrow & & \downarrow \\
\mathrm{map}_*(K(\pi,n), \widehat{BU}) & \xleftarrow{\varphi''} & \mathrm{map}_*(K(\pi_{(0)},n), \widehat{BU})
\end{array}$$

By [1, Corollary 4.8], φ is a weak homotopy equivalence. Since F_{BU} is rational (Proposition 5.4), φ' is a weak homotopy equivalence [11, Section 19.5]. Since $K(\pi_{(0)},n) \to *$ is an $H\mathbb{Z}/p$–equivalence for any prime p, $\mathrm{map}_*(K(\pi_{(0)},n), \widehat{BU})$ is weakly contractible [11, Sections 19.2 and 19.5], and hence the null component of $\mathrm{map}_*(K(\pi,n), \widehat{BU})$ is weakly contractible. On the other hand, we have

$$\begin{aligned}
[K(\pi,n), \widehat{BU}] &\cong [K(\pi,n), \Omega^2 \widehat{BU}] \\
&\cong \pi_2(\mathrm{map}_*(K(\pi,n), \widehat{BU})) = 0,
\end{aligned}$$

which implies that $\mathrm{map}_*(K(\pi,n), \widehat{BU})$ is weakly contractible.

For general A, the result is proved by induction using Dwyer's version of the Zabrodsky lemma [3, Proposition 3.4].

Step 2 The result for $B = B\,\mathrm{SO}$ is obtained from the result for BO using the fibration sequence

$$\{\pm 1\} \to B\,\mathrm{SO} \to BO.$$

Step 3 The result for $B = U/\,\mathrm{Sp}$ is obtained from the results for BU and $B\,\mathrm{Sp}$ using the fibration sequence

$$U/\,\mathrm{Sp} \to B\,\mathrm{Sp} \to BU.$$

Applying similar arguments, we can obtain results for $B = \mathrm{Sp}\,/U$, $\mathrm{SO}\,/U$ and $U/\,\mathrm{SO}$.

Step 4 We express the identity component of ΩB as $\Omega_0 B$. Observe that if $\mathrm{map}_*(A, \widehat{B})$ is weakly contractible, then

$$\mathrm{map}_*(A, \Omega^k \widehat{B}) \cong \mathrm{map}_*(A, \widehat{\Omega_0^k B})$$

is also weakly contractible. This statement completes the proof. □

Corollary 6.4 *Let (X,Y) be in $\mathcal{A}\times\mathcal{B}$ or $\mathcal{A}'\times\mathcal{B}'$. Then the mapping space* $\operatorname{map}_*(X,\Omega\widehat{Y})$ *is weakly contractible.*

Proof This immediately follows from Proposition 6.3. □

Remark 6.5 Proposition 6.3(2) and its use in proving the results apply equally to the representing spaces of conjugate K–theory KC_* [1].

Proof of Proposition 2.5 The result follows from Proposition 6.1 and Corollary 6.4. □

6.2 $\mathrm{Ph}(X,Y)$ and $\mathrm{SPh}(X,Y)$ in the case of nontrivial $[X,\Omega\widehat{Y}]$

In this subsection we prove Theorem 2.7 and Corollaries 2.8–2.10

Proof of Theorem 2.7 The case of $\mathrm{Ph}(X,Y)$ By Corollary 5.3 and the comment before Lemma 3.5, the morphism of exact sequences of pointed sets is given by

$$\begin{array}{ccccccc}
[X,\Omega\widehat{Y}] & \longrightarrow & [X,F_Y] & \longrightarrow & \mathrm{Ph}(X,Y) & \longrightarrow & 0 \\
\uparrow{\scriptstyle p^\sharp} & & \uparrow{\scriptstyle p^\sharp} & & \uparrow{\scriptstyle p^\sharp} & & \\
[K,\Omega\widehat{Y}] & \longrightarrow & [K,F_Y] & \longrightarrow & \mathrm{Ph}(K,Y) & \longrightarrow & 0
\end{array}$$

We observe that $p^\sharp\colon [K,\Omega\widehat{Y}]\to[X,\Omega\widehat{Y}]$ is surjective. In the cofibration case, the surjectivity follows immediately from the assumption; in the fibration case, the surjectivity is inferred from the assumption via Dwyer's version of the Zabrodsky lemma [3, Proposition 3.4]. Next, consider the induced morphism

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \overline{[X,\Omega\widehat{Y}]} & \longrightarrow & [X,F_Y] & \longrightarrow & \mathrm{Ph}(X,Y) & \longrightarrow & 0 \\
 & & \uparrow{\scriptstyle \psi} & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & \operatorname{Ker}\phi & \longrightarrow & p^\sharp[K,F_Y] & \xrightarrow{\phi} & p^\sharp\,\mathrm{Ph}(K,Y) & \longrightarrow & 0
\end{array}$$

of exact sequences of pointed sets, where $\overline{[X,\Omega\widehat{Y}]}$ denotes the image of $[X,\Omega\widehat{Y}]$ and ϕ denotes the map induced by the natural quotient map $[K,F_Y]\to\mathrm{Ph}(K,Y)$. By Theorem 2.3 and its proof, this is a morphism of exact sequences of abelian groups. Since $p^\sharp\colon[K,\Omega\widehat{Y}]\to[X,\Omega\widehat{Y}]$ is surjective, ψ is also surjective. We regard this morphism of exact sequences as a short exact sequence of chain complexes, and take its homology exact sequence. Thus we find that

$$\mathrm{Ph}(X,Y)/p^\sharp\,\mathrm{Ph}(K,Y)\cong[X,F_Y]/p^\sharp[K,F_Y],$$

which gives the desired exact sequence (see Proposition 5.10 and its proof, Proposition 4.1 and Theorem 4.3). Since $p^\sharp \mathrm{Ph}(K,Y)$ is a divisible abelian group (Theorem 2.3(2)), the short exact sequence splits.

The case of SPh(X,Y) As mentioned in the introduction of Section 6.1, $\mathrm{Ph}(X,Y)$ and $\mathrm{Ph}(X,\check{Y})$ generate analogous results. Thus there exists a morphism of exact sequences of abelian groups

$$\begin{array}{ccccccccc}
0 & \longrightarrow & p^\sharp \mathrm{Ph}(K,Y) & \longrightarrow & \mathrm{Ph}(X,Y) & \longrightarrow & \displaystyle\prod_{i>0} \frac{H^i(X;\pi_{i+1}(Y)\otimes \widehat{\mathbb{Z}}/\mathbb{Z})}{p^*H^i(K;\pi_{i+1}(Y)\otimes \widehat{\mathbb{Z}}/\mathbb{Z})} & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \epsilon} & & \downarrow{\scriptstyle e_{Y\sharp}} & & \downarrow & & \\
0 & \longrightarrow & p^\sharp \mathrm{Ph}(K,\check{Y}) & \longrightarrow & \mathrm{Ph}(X,\check{Y}) & \longrightarrow & \displaystyle\prod_{i>0} \frac{H^i(X;\pi_{i+1}(Y)\otimes \widehat{\mathbb{Z}}/\check{\mathbb{Z}})}{p^*H^i(K;\pi_{i+1}(Y)\otimes \widehat{\mathbb{Z}}/\check{\mathbb{Z}})} & \longrightarrow & 0
\end{array}$$

Regarding this morphism as a short exact sequence of chain complexes, we take its homology exact sequence. Thus we obtain the exact sequence

$$0 \to \mathrm{Ker}\,\epsilon \to \mathrm{SPh}(X,Y) \to \prod_{i>0} \frac{H^i(X;\pi_{i+1}(Y)\otimes \check{\mathbb{Z}}/\mathbb{Z})}{p^*H^i(K;\pi_{i+1}(Y)\otimes \check{\mathbb{Z}}/\mathbb{Z})} \to 0$$

by Corollary 5.9. Define the abelian group $\widetilde{\mathrm{Ker}\,\epsilon}$ by the pullback diagram

$$\begin{array}{ccc}
\widetilde{\mathrm{Ker}\,\epsilon} & \hookrightarrow & \displaystyle\prod_{i>0} H^i(X;\pi_{i+1}(Y)\otimes \check{\mathbb{Z}}/\mathbb{Z}) \\
\downarrow & & \downarrow \\
\mathrm{Ker}\,\epsilon & \hookrightarrow & \mathrm{SPh}(X,Y)
\end{array}$$

in the category of abelian groups, where the right vertical arrow is the natural quotient map (see Proposition 5.10 and the proof of Theorem 2.3(1)). Since the right vertical arrow is an epimorphism, this diagram is also a pushout diagram, implying that the two horizontal arrows share a single cokernel. Thus it is seen that

$$\widetilde{\mathrm{Ker}\,\epsilon} = \prod_{i>0} p^*H^i(K;\pi_{i+1}(Y)\otimes \check{\mathbb{Z}}/\mathbb{Z}),$$

which implies that $\mathrm{Ker}\,\epsilon = p^\sharp \mathrm{SPh}(K,Y)$ by the proof of Theorem 2.3(2), and the desired exact sequence is obtained. The short exact sequence splits by the divisibility of $p^\sharp \mathrm{SPh}(K,Y)$. □

We use Theorem 2.7 to prove Corollaries 2.8–2.10.

Proof of Corollary 2.8 Let X' be the fiber of $p\colon X \to T$. Note that X' is a $\frac{1}{2}$–connected infinite loop space. Since $\mathrm{map}_*(X', \Omega\widehat{Y})$ is weakly contractible by Corollary 6.4 and $\mathrm{Ph}(T, Y)$ vanishes, we obtain the result from Theorem 2.7. □

Proof of Corollary 2.9 Consider the fibration sequence

$$\Omega K^{(n)} \to K\langle n\rangle \xrightarrow{p} K,$$

where $K^{(n)}$ is the Postnikov n–stage of K. Since $\mathrm{map}_*(\Omega K^{(n)}, \Omega\widehat{Y})$ is weakly contractible by Corollary 6.4 and $\mathrm{Ph}(K, Y)$ vanishes, we obtain the result from Theorem 2.7. □

Proof of Corollary 2.10 Consider the fibration sequence

$$X' \xrightarrow{i} X \xrightarrow{p} K,$$

where i is the map $(1, 0)\colon X' \to X' \times K = X$ and p is the canonical projection onto K. By Corollary 6.4, we can apply Theorem 2.7 to (X, Y). The canonical section $(0, 1)\colon K \to X' \times K = X$ of the fibration p shows that $p^\sharp\colon \mathrm{Ph}(K, Y) \to \mathrm{Ph}(X, Y)$ is injective and splits the exact sequences of Theorem 2.7. The products of quotients of cohomology groups are easily identified with the Hom-modules in the statement. □

Last, we apply Theorem 2.7 to pairs (X, Y) in $\mathcal{Q}$ admitting a cofibration sequence $X' \xrightarrow{i} X \xrightarrow{p} K$ with $[X', \Omega\widehat{Y}] = 0$.

The Grassmannians $G_n(\mathbb{F})$ and $G_\infty(\mathbb{F})$ are defined by $G_n(\mathbb{F}) = \varinjlim_m G_{n,m}(\mathbb{F})$ and $G_\infty(\mathbb{F}) = \varinjlim_n G_n(\mathbb{F})$ for $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{H}$, where the finite Grassmannian $G_{n,m}(\mathbb{F})$ is the space of n–dimensional subspaces in $\mathbb{F}^{n+m}$.

Example 6.6 Let $G_n(\mathbb{F})/G_{n',m'}(\mathbb{F})$ be the quotient complex of $G_n(\mathbb{F})$ by $G_{n',m'}(\mathbb{F})$ $(n', m' < \infty,\ n' \le n \le \infty)$. Let Y be a space in $\mathcal{B}$ such that $(G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}), Y)$ is in $\mathcal{Q}$. Then there exist natural isomorphisms of groups

$$\mathrm{Ph}(G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}), Y) \cong \prod_{i>0} \frac{H^i\big(G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}); \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}\big)}{p^* H^i\big(\Sigma G_{n',m'}(\mathbb{F}); \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}\big)},$$

$$\mathrm{SPh}(G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}), Y) \cong \prod_{i>0} \frac{H^i\big(G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}); \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}\big)}{p^* H^i\big(\Sigma G_{n',m'}(\mathbb{F}); \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}\big)},$$

where p is the canonical map $G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}) \to \Sigma G_{n',m'}(\mathbb{F})$. Furthermore, if $\mathbb{F} = \mathbb{C}$ or $\mathbb{H}$, then

$$p^* H^i\big(\Sigma G_{n',m'}(\mathbb{F}); \pi_{i+1}(Y) \otimes \widehat{\mathbb{Z}}/\mathbb{Z}\big) = p^* H^i\big(\Sigma G_{n',m'}(\mathbb{F}); \pi_{i+1}(Y) \otimes \check{\mathbb{Z}}/\mathbb{Z}\big) = 0.$$

Proof Consider the cofibration sequence

$$G_n(\mathbb{F}) \to G_n(\mathbb{F})/G_{n',m'}(\mathbb{F}) \xrightarrow{p} \Sigma G_{n',m'}(\mathbb{F}).$$

Since $[G_n(\mathbb{F}), \Omega\widehat{Y}] = 0$ by Corollary 6.4 and $\mathrm{Ph}(\Sigma G_{n',m'}(\mathbb{F}), Y) = 0$, we obtain the result from Theorem 2.7. □

Acknowledgement

I would like to give heartfelt thanks to Professor Nobuyuki Oda, whose comments and suggestions were of inestimable value for my study. I would also like to thank Professor Toshiro Watanabe, who provided technical help and sincere encouragement.

References

[1] **D W Anderson**, **L Hodgkin**, *The K–theory of Eilenberg–MacLane complexes*, Topology 7 (1968) 317–329 MR

[2] **H J Baues**, *Commutator calculus and groups of homotopy classes*, London Math. Soc. Lecture Note Ser. 50, Cambridge University Press (1981) MR

[3] **W G Dwyer**, *The centralizer decomposition of BG*, from "Algebraic topology: new trends in localization and periodicity" (C Broto, C Casacuberta, G Mislin, editors), Progr. Math. 136, Birkhäuser, Basel (1996) 167–184 MR

[4] **E M Friedlander**, **G Mislin**, *Locally finite approximation of Lie groups, I*, Invent. Math. 83 (1986) 425–436 MR

[5] **L M Hà**, *On the Gray index of phantom maps*, Topology 44 (2005) 217–229 MR

[6] **L M Hà**, **J Strom**, *The Gray filtration on phantom maps*, Fund. Math. 167 (2001) 251–268 MR

[7] **K Iriye**, *On the Gray index conjecture for phantom maps*, Topology Appl. 157 (2010) 2059–2068 MR

[8] **E I Khukhro**, *p–automorphisms of finite p–groups*, London Math. Soc. Lecture Note Ser. 246, Cambridge University Press (1998) MR

[9] **H Kihara**, *Commutativity and cocommutativity of cogroups in the category of connected graded algebras*, Topology Appl. 189 (2015) 107–121 MR

[10] **H Kihara**, **N Oda**, *Rational cup product and algebraic K_0–groups of rings of continuous functions* To appear in Proc. Edinburgh Math. Soc

[11] **J P May**, **K Ponto**, *More concise algebraic topology*, Univ. of Chicago Press (2012) MR

[12] **C A McGibbon**, *Clones of spaces and maps in homotopy theory*, Comment. Math. Helv. 68 (1993) 263–277 MR

[13] **C A McGibbon**, *The Mislin genus of a space*, from "The Hilton Symposium 1993" (G Mislin, editor), CRM Proc. Lecture Notes 6, Amer. Math. Soc., Providence, RI (1994) 75–102 MR

[14] **C A McGibbon**, *Phantom maps*, from "Handbook of algebraic topology" (I M James, editor), North-Holland, Amsterdam (1995) 1209–1257 MR

[15] **C A McGibbon**, *A note on Miller's theorem about maps out of classifying spaces*, Proc. Amer. Math. Soc. 124 (1996) 3241–3245 MR

[16] **C A McGibbon**, **J M Møller**, *How can you tell two spaces apart when they have the same n–type for all n?*, from "Adams Memorial Symposium on Algebraic Topology, 1" (N Ray, G Walker, editors), London Math. Soc. Lecture Note Ser. 175, Cambridge Univ. Press (1992) 131–143 MR

[17] **C A McGibbon**, **J Roitberg**, *Phantom maps and rational equivalences, II*, Bol. Soc. Mat. Mexicana 37 (1992) 367–381 MR

[18] **C A McGibbon**, **J Strom**, *Numerical invariants of phantom maps*, Amer. J. Math. 123 (2001) 679–697 MR

[19] **W Meier**, *Localisation, complétion, et applications fantômes*, C. R. Acad. Sci. Paris Sér. A-B 281 (1975) 787–789 MR

[20] **W Meier**, *Pullback theorems and phantom maps*, Quart. J. Math. Oxford Ser. 29 (1978) 469–481 MR

[21] **H Miller**, *The Sullivan conjecture on maps from classifying spaces*, Ann. of Math. 120 (1984) 39–87 MR

[22] **J W Milnor**, **J C Moore**, *On the structure of Hopf algebras*, Ann. of Math. 81 (1965) 211–264 MR

[23] **J Roitberg**, *Weak identities, phantom maps and H–spaces*, Israel J. Math. 66 (1989) 319–329 MR

[24] **J Roitberg**, *Computing homotopy classes of phantom maps*, from "The Hilton Symposium 1993" (G Mislin, editor), CRM Proc. Lecture Notes 6, Amer. Math. Soc., Providence, RI (1994) 141–168 MR

[25] **J Roitberg**, **P Touhey**, *The homotopy fiber of profinite completion*, Topology Appl. 103 (2000) 295–307 MR

[26] **H Scheerer**, *On rationalized H– and co-H–spaces*, Manuscripta Math. 51 (1985) 63–87 MR

[27] **J Siegel**, *G–spaces, H–spaces and W–spaces*, Pacific J. Math. 31 (1969) 209–214 MR

[28] **J Stasheff**, *H–spaces from a homotopy point of view*, Lecture Notes in Math. 161, Springer (1970) MR

[29] **R J Steiner**, *Localization, completion and infinite complexes*, Mathematika 24 (1977) 1–15 MR

[30] **D Sullivan**, *Genetics of homotopy theory and the Adams conjecture*, Ann. of Math. 100 (1974) 1–79 MR

[31] **D P Sullivan**, *Geometric topology: localization, periodicity and Galois symmetry*, K–Monographs in Math. 8, Springer (2005) MR

[32] **G W Whitehead**, *Elements of homotopy theory*, Graduate Texts in Math. 61, Springer (1978) MR

[33] **A Zabrodsky**, *Hopf spaces*, North-Holland Math. Studies 22, North-Holland, Amsterdam (1976) MR

[34] **A Zabrodsky**, *On phantom maps and a theorem of H Miller*, Israel J. Math. 58 (1987) 129–143 MR

Center for Mathematical Sciences, University of Aizu
Aizuwakamatsu, Fukushima, Japan

kihara@u-aizu.ac.jp

Received: 24 April 2017 Revised: 16 June 2017

msp

Algebraic & Geometric Topology 18 (2018) 613–633

Loop homology of some global quotient orbifolds

YASUHIKO ASAO

We determine the ring structure of the loop homology of some global quotient orbifolds. We can compute by our theorem the loop homology ring with suitable coefficients of the global quotient orbifolds of the form $[M/G]$ for M being some kinds of homogeneous manifolds, and G being a finite subgroup of a path-connected topological group $\mathcal{G}$ acting on M. It is shown that these homology rings split into the tensor product of the loop homology ring $\mathbb{H}_*(LM)$ of the manifold M and that of the classifying space of the finite group, which coincides with the center of the group ring $Z(k[G])$.

55N45, 55N91, 55P35, 55P91

1 Introduction

The free loop space of a topological space X is a space of the continuous maps from the circle S^1 to X,

$$LX = \mathrm{Map}(S^1, X), \tag{1}$$

with the compact-open topology. The loop homology of X is the homology of the free loop space, $H_*(LX)$. In the 1990s Moira Chas and Dennis Sullivan [3] discovered a product $\circ$ on the loop homology of a closed oriented smooth manifold $H_*(LM)$,

$$\circ\colon H_p(LM) \otimes H_q(LM) \to H_{p+q-\dim M}(LM), \tag{2}$$

called the loop product, which is a mixture of the intersection product of a manifold and the concatenation operation identifying $S^1 \vee S^1$ with the image of a map from S^1. They also showed in [3] that the product defines a ring structure and a kind of Lie algebra structure called the Batalin–Vilkovisky algebra (BV–algebra) structure on the homology. These rich algebraic structures are called *string topology* of closed oriented manifolds. The string topology of a manifold is related to several areas of mathematics including mathematical physics through the tools of algebraic topology.

The string topology of a wider class of spaces has been developed by several authors:

Published: 19 January 2018 DOI: 10.2140/agt.2018.18.613

- For classifying spaces of connected compact Lie groups, the existence of the loop product and the BV–structure on its homology is proved by Chataur and Menichi in [4].
- For fiberwise monoids including the adjoint bundle of principal bundles, the loop product is constructed by Gruher and Salvatore in [8].
- And more generally, for the Borel construction of a smooth manifold with a smooth action of a compact Lie group, the loop product is constructed by Kaji and Tene in [11].
- In the 2000s, Lupercio, Uribe, and Xicoténcatl defined the loop homology of a global quotient orbifold, which is an orbifold of the form $[M/G]$ for M being a smooth manifold and G being a finite group acting smoothly on M, as the loop homology of the Borel construction $M \times_G EG$. They discovered a product

(3) $$\circ\colon H_p(L(M \times_G EG)) \otimes H_q(L(M \times_G EG)) \to H_{p+q-\dim M}(L(M \times_G EG))$$

on the loop homology of a global quotient orbifold $H_*(L(M \times_G EG))$, and showed that the product defines a BV–algebra structure on $H_*(L(M \times_G EG))$; see [14]. They coined the name *orbifold string topology* for this structure.

In spite of these interesting structural discoveries, concrete computations of loop homology have been achieved for only a few kinds of classes of manifolds. In order to accommodate the change in grading, we define $\mathbb{H}_*(LM) = H_{*+\dim M}(LM)$.

- For a compact Lie group Γ, there is a homeomorphism $L\Gamma \cong \Omega\Gamma \times \Gamma$, and a loop homology ring isomorphism $\mathbb{H}_*(L\Gamma) \cong H_*(\Omega\Gamma) \otimes \mathbb{H}_*(\Gamma)$. Furthermore, the BV–algebra structure with coefficients in $\mathbb{Q}$ and $\mathbb{Z}_2$ of any compact Lie group G is determined by Hepworth in [10].
- For spheres S^n and complex projective spaces $\mathbb{C}\mathrm{P}^n$, the ring structure with coefficients in $\mathbb{Z}$ is determined by Cohen, Jones, and Yan in [6] by constructing the spectral sequence converging to the loop homology.
- The BV–algebra structure for $\mathbb{C}\mathrm{P}^n$ with coefficients in $\mathbb{Z}$ is determined by Hepworth in [9].
- For complex Stiefel manifolds $\mathrm{SU}(n)/\mathrm{SU}(k)$ including odd-dimensional spheres S^{2n+1}, the BV–structure is determined by Tamanoi in [18].
- For arbitrary spheres, Menichi determines the BV–structure for it in [15] by using the Hochschild cohomology.

- For the aspherical manifold $K(\pi,1)$, the BV–structure is determined by Vaintrob in [20] by establishing an isomorphism between the loop homology $\mathbb{H}_*(K(\pi,1))$ and the Hochschild cohomology $\mathrm{HH}^*(\mathbb{Z}[\pi];\mathbb{Z}[\pi])$, and another proof is obtained by Kupers in [12].

The purpose of this paper is to determine the ring structure of some global quotient orbifolds by using the method of the orbifold string topology. We can compute by our theorem the loop homology ring with suitable coefficients of the global quotient orbifolds of the form $[M/G]$ for M a homogeneous manifold of a connected Lie group $\mathcal{G}$, and G a finite subgroup of $\mathcal{G}$.

Now we briefly review a part of the work of Lupercio, Uribe and Xicoténcatl in [14]. For simplicity, we denote $\mathrm{Map}(S^1, M\times_G EG)$ as $L[M/G]$. In [14], the *loop orbifold* of the global quotient orbifold $[M/G]$ is defined as the groupoid $P_GM \rtimes G$, and they show that its Borel construction $P_GM\times_G EG$ is weak homotopy equivalent to the free loop space $L(M\times_G EG)$. To determine the loop homology ring of the lens space $S^{2n+1}/\mathbb{Z}_p$, they constructed a *non-G–equivariant* homotopy equivalence

$$P_{\mathbb{Z}_p}S^{2n+1} \simeq \coprod_{g\in\mathbb{Z}_p} LS^{2n+1} \tag{4}$$

by using the fact that the action of $\mathbb{Z}_p$ extends to an S^1 action. We prove:

Proposition 3.1 [14] *Let M be a closed oriented manifold, and G be a finite subgroup of a path-connected group $\mathcal{G}$ acting continuously on M. Then for each $g\in G$, there exists a homotopy equivalence*

$$P_gX \simeq LM. \tag{5}$$

We prove that this homotopy equivalence can be extended to a wider class of spaces. Furthermore, we find a condition so that the above homotopy equivalence is *G–equivariant at homology level*.

Now we state our main theorem. Let $\mathcal{G}$ be a path-connected group acting continuously on M, and G a finite subgroup of $\mathcal{G}$ acting smoothly on M. Then there is the map $\Phi\colon \Omega\mathcal{G}\times LM\to LM$, with $\Phi(a,l)=(t\mapsto a(t)\cdot l(t))$, and this map induces an action of the Pontrjagin ring $H_*(\Omega\mathcal{G})$, namely $\Phi_*\colon H_*(\Omega\mathcal{G})\otimes H_*(LM)\to H_*(LM)$. If the action Φ_* of $\pi_0(\Omega\mathcal{G})$ on $H_*(LM;k)$ satisfies the equation $\Phi_*(x)=x$ for any $x\in H_*(LM)$, we call the action Φ_* *trivial*. Then we prove the following.

Proposition 3.2 *Assume that the action Φ_* is trivial with coefficient in k. Then the direct sum of the homotopy equivalence of the above proposition,*

$$\coprod_{g\in G} P_g M \simeq \coprod_{g\in G} LM,$$

is G–equivariant at homology level with coefficients in k.

By using this homotopy equivalence, we can compute the loop homology of a certain class of orbifolds. Our main theorem is the following; we denote the order of a group G by $|G|$.

Theorem 4.1 *Let $\mathcal{G}$ be a path-connected topological group acting continuously on an oriented closed manifold M, let G be its finite subgroup, and k be a field whose characteristic is coprime to $|G|$. If the action Φ_* is trivial with coefficients in k, then there exists an isomorphism as k–algebras*

$$\mathbb{H}_*(L[M/G];k) \cong \mathbb{H}_*(LM;k)\otimes Z(k[G]), \tag{6}$$

where $Z(k[G])$ denotes the center of the group ring $k[G]$.

We show that the following are the necessary conditions for the action Φ_* to be trivial.

Proposition 6.1 *If $\mathcal{G}$ is simply connected, then the action Φ_* is trivial for any field k.*

Proposition 6.2 *If $H_{\dim M}(LM;k) = k$, then the action Φ_* is trivial for any pair $(\mathcal{G}, G)$.*

Proposition 6.4 *If the following conditions are satisfied, then the action Φ_* is trivial:*

(i) *M is simply connected.*

(ii) *$|\pi_1\mathcal{G}| < \infty$.*

(iii) *The homomorphism $H_*(\Omega M;k) \to H_*(LM;k)$ induced by the inclusion map $\Omega M \to LM$ is injective, that is, the free loop fibration $\Omega M \to LM \to M$ is **totally noncohomologous to zero** (TNCZ) with coefficients in the field k.*

(iv) *The characteristic of k is coprime to $|\pi_1\mathcal{G}|$.*

The organization of this paper is as follows. After this introduction in Section 1, we briefly review the string topology first developed in [3] and the orbifold string topology constructed in [14] in Section 2. In Section 3, we show some propositions necessary for the proof of the main theorem. In Section 4, we prove the theorem first for vector spaces and second for algebras. In Section 5, we remark on the relation to Hochschild cohomology. Finally in Section 6, we compute concrete examples by applying our theorem.

Acknowledgements The author would like to express his sincere gratitude to his advisor Professor Toshitake Kohno for all advice and encouragement. He also would like to show his appreciation to Takahito Naito for fruitful communication. He is indebted to the referee for providing insightful comments on the relation to Hochschild cohomology.

2 Preliminaries for string topology

In this section, we briefly review the loop product in string topology.

Loop product

Let M be a smooth closed oriented manifold, and $LM = \mathrm{Map}(S^1, M)$ be the free loop space of M, the space of piecewise smooth maps from S^1 to M with compact-open topology. Then we have the following pullback diagram:

$$\begin{array}{ccc} \mathrm{Map}(S^1 \vee S^1, M) = LM \times_M LM & \xrightarrow{\tilde{\iota}} & LM \times LM \\ \downarrow{\scriptstyle \mathrm{ev}_{1/2}} & & \downarrow{\scriptstyle \mathrm{ev} \times \mathrm{ev}} \\ M & \xrightarrow{\iota} & M \times M \end{array} \tag{7}$$

where ι denotes the diagonal embedding, and ev_t denotes the evaluation map with $\mathrm{ev}_t(l) = l(t)$. Then we can consider above $\tilde{\iota}$ as codimension n embedding of the infinite-dimensional manifold, and we have the generalized Pontrjagin–Thom map

$$\tilde{\iota}_{!*}\colon H_*(LM \times LM) \to H_*((LM \times_M LM)^{\mathrm{ev}_{1/2}^* \nu_\iota}) \tag{8}$$

due to Cohen and Klein [7], where ν_ι denotes the normal bundle of the embedding ι, and $(LM \times_M LM)^{\mathrm{ev}_{1/2}^* \nu_\iota}$ denotes the Thom space of the vector bundle $\mathrm{ev}_{1/2}^* \nu_\iota$. The similar but more homotopy-theoretic construction of this umkehr map is considered in [11]. The loop product is formulated in [5] as the composition of maps $\tilde{\iota}_{!*}$, the Thom isomorphism

$$H_*((LM \times_M LM)^{\mathrm{ev}_{1/2}^* \nu_\iota}) \xrightarrow{\cong} H_{*-\dim M}(LM \times_M LM), \tag{9}$$

and the concatenating map $\gamma\colon LM \times_M LM \to LM$ with

$$\gamma(l_1, l_2) = l_1 * l_2 := \begin{cases} l_1(2t) & \text{for } 0 \le t \le \frac{1}{2}, \\ l_2(2t-1) & \text{for } \frac{1}{2} \le t \le 1. \end{cases} \tag{10}$$

Definition 2.1 [5] The following sequence of compositions defines the loop product:

$$\begin{aligned} H_p(LM) \otimes H_q(LM) \xrightarrow{\times} H_{p+q}(LM \times LM) &\xrightarrow{e_{!*}} H_{p+q}((LM \times_M LM)^{\mathrm{ev}_{1/2}^* \nu_\iota}) \\ &\xrightarrow{\text{Thom isom.}} H_{p+q-n}(LM \times_M LM) \xrightarrow{\gamma_*} H_{p+q-n}(LM). \end{aligned}$$

In order to accommodate the change in grading, we define $\mathbb{H}_*(LM) = H_{*+\dim M}(LM)$. Chas and Sullivan prove the following in [3].

Theorem 2.2 *The loop product makes $\mathbb{H}_*(LM)$ an associative graded commutative algebra.*

Remark Cohen and Jones give in [5] an operadic proof of Theorem 2.2, and Tamanoi gives in [19] a more homotopy-theoretic one.

Orbifolds

In this section, we review the basic definitions and properties of orbifolds that we will use in this paper. Following Moerdijk [17], we use the groupoid notion of an orbifold. For more detail, see [1; 17].

Definition 2.3 A *groupoid* is a category whose morphisms are all invertible. In other words, a groupoid G is a pair of sets $(\mathsf{G}_0, \mathsf{G}_1)$ with the structure maps

- **source and target** $s, t\colon \mathsf{G}_1 \to \mathsf{G}_0$,
- **identity** $e\colon \mathsf{G}_0 \to \mathsf{G}_1$,
- **inverse** $i\colon \mathsf{G}_1 \to \mathsf{G}_1$,
- **composition** $m\colon \mathsf{G}_1 \times_{\mathsf{G}_0} \mathsf{G}_1 = \{(a,b) \in \mathsf{G}_1 \times \mathsf{G}_1 | t(a) = s(b)\} \to \mathsf{G}_1$,

satisfying suitable compatibility conditions.

We note here some technical terms on groupoids. A *Lie groupoid* is a groupoid $(\mathsf{G}_0, \mathsf{G}_1)$ such that G_0 and G_1 both have the structure of a smooth manifold, and the structure maps are all smooth. We will also require that s, t are submersions. An *isotropy group* of a groupoid $(\mathsf{G}_0, \mathsf{G}_1)$ at $x \in \mathsf{G}_0$ is the group $G_x = \{f \in \mathsf{G}_1 \mid s(f) = t(f) = x\}$. A Lie groupoid $(\mathsf{G}_0, \mathsf{G}_1)$ is said to be a *proper groupoid* if the map $(s,t)\colon \mathsf{G}_1 \to \mathsf{G}_0 \times \mathsf{G}_0$ is proper. A Lie groupoid $(\mathsf{G}_0, \mathsf{G}_1)$ is said to be a *foliation groupoid* if the isotropy group G_x is discrete for each $x \in \mathsf{G}_0$.

Definition 2.4 A groupoid is said to be an *orbifold groupoid* if it is a proper foliation Lie groupoid.

Example 2.5 Let S be a set, and $\mathcal{G}$ be a group acting on S. Then $(S, S \times \mathcal{G})$ is a groupoid with structure maps

- $s(x,g) = x,\ t(x,g) = gx$,
- $e(x) = (x, \mathrm{id}_{\mathcal{G}})$,

- $i(x,g) = (gx, g^{-1})$,
- $m((x,g),(gx,h)) = (x,hg)$,

for any $x, y \in S$ and $g, h \in \mathcal{G}$. We call this groupoid the *action groupoid*, and denote it by $S \rtimes \mathcal{G}$.

Example 2.6 Let M be a smooth manifold, and Γ be a compact Lie group acting smoothly on M. If the isotropy group Γ_x is finite for each $x \in M$, then the action groupoid $M \rtimes \Gamma$ is an orbifold groupoid.

Orbifolds are defined using the notion of groupoids, as follows.

Definition 2.7 A Morita equivalent class $[\mathsf{G}]$ of orbifold groupoids is called an *orbifold*.

Definition 2.8 An orbifold X is called a *global quotient orbifold* if X has a representation of an action groupoid $M \rtimes G$, where M is a smooth manifold, and G is a finite group. We write $\mathsf{X} = [M/G]$.

Remark The following are fundamental properties of orbifolds; see for example [1] for proof.

(i) Any orbifold groupoid is Morita equivalent to an action groupoid $M \rtimes \Gamma$, where M is a smooth manifold, and Γ is a compact Lie group acting smoothly on M with its isotropy groups being finite.

(ii) Let $\mathsf{X} = [M \rtimes \Gamma]$ be an orbifold. Then the homotopy type of the Borel construction $M \times_\Gamma E\Gamma$ is an orbifold invariant; that is, it is invariant under Morita equivalences.

Orbifold loop product

In this section, we review the construction of orbifold loop product defined in [14].

The following notion of loop orbifold is defined also in [14].

Definition 2.9 [14] Let $[M/G]$ be a global quotient orbifold. The *loop orbifold* of $[M/G]$ is the action groupoid $P_G M \rtimes G$, where

$$P_G M = \coprod_{g \in G} P_g M, \quad P_g M = \{\sigma\colon [0,1] \to M \mid \sigma(1) = \sigma(0)g\} \subset LM,$$

and G acts on $P_G M$ via the map

$$P_g M \times G \to P_{h^{-1}gh} M, \quad (\sigma, h) \mapsto \sigma h.$$

In the same paper, they prove the weak homotopy equivalence $L(X \times_G EG) \simeq P_G X \times_G EG$. Hence we can consider $P_G M \times_G EG$ instead of $L(M \times_G EG)$ for studying the string topology of the orbifold $[M/G]$ because of the Whitehead theorem. They construct the orbifold loop product as follows.

Construction of the orbifold loop product We consider the pullback diagram

$$\begin{array}{ccc} P_g M \times_M P_h M & \xrightarrow{\widetilde{\iota}} & P_g M \times P_h M \\ \downarrow{\scriptstyle \mathrm{ev}_{1/2}} & & \downarrow{\scriptstyle \mathrm{ev}_1 \times \mathrm{ev}_0} \\ M & \xrightarrow{\iota} & M \times M \end{array} \tag{11}$$

for any $g, h \in G$, where ev_t denotes the evaluation map. Then we have the generalized Pontryagin–Thom map

$$\widetilde{\iota}_!\colon P_g M \times P_h M \to (P_g M \times_M P_h M)^{\mathrm{ev}_{1/2}^* \nu_\iota},$$

where ν_ι denotes the normal bundle of the embedding ι, and $(P_g M \times_M P_h M)^{\mathrm{ev}_{1/2}^* \nu_\iota}$ denotes the Thom space of the vector bundle $\mathrm{ev}_{1/2}^* \nu_\iota$. We also have the concatenation map $\gamma\colon P_g M \times_M P_h M \to P_{gh} M$ with

$$\gamma(\sigma_g, \sigma_h) = \begin{cases} \sigma_g(2t) & \text{for } 0 \le t \le \frac{1}{2}, \\ \sigma_h(2t-1) & \text{for } \frac{1}{2} \le t \le 1. \end{cases} \tag{12}$$

Then we obtain a sequence of compositions

$$\begin{aligned} H_p(P_g M) \otimes H_q(P_h M) &\xrightarrow{\times} H_{p+q}(P_g M \times P_h M) \xrightarrow{\widetilde{\iota}_{!*}} H_{p+q}((P_g M \times_M P_h M)^{\mathrm{ev}_{1/2}^* \nu_\iota}) \\ &\xrightarrow{\text{Thom isom.}} H_{p+q-\dim M}(P_g M \times_M P_h M) \xrightarrow{\gamma_*} H_{p+q-\dim M}(P_{gh} M), \end{aligned}$$

and we obtain a map

$$\bullet\colon H_p(P_g M) \otimes H_q(P_h M) \to H_{p+q-\dim M}(P_{gh} M). \tag{13}$$

By summing over $g, h \in G$, we obtain a map, which we also denote by $\bullet$,

$$\bullet\colon H_p(P_G M) \otimes H_q(P_G M) \to H_{p+q-\dim M}(P_G M). \tag{14}$$

To lift the map $\bullet$ to a map

$$\circ\colon H_p(P_G M \times_G EG) \otimes H_q(P_G M \times_G EG) \to H_{p+q-\dim M}(P_G M \times_G EG), \tag{15}$$

we use the following fundamental lemma in algebraic topology.

Lemma 2.10 *Let $p\colon \widetilde{X} \to X$ be a finite Galois covering and G its Galois group. If k is a field whose characteristic is coprime to $|G|$, then there exists an injective homomorphism μ_* called the transfer map,*

$$\mu_*\colon H_*(X;k) \to H_*(\widetilde{X};k), \tag{16}$$

such that

$$\operatorname{Im}\mu_* \cong H_*(\widetilde{X};k)^G, \tag{17}$$

where the right-hand side is the G–invariant subspace of $H_*(X;k)$.

Proof We first define the transfer map $\mu_*\colon H_*(X) \to H_*(\widetilde{X})$. Let μ be a homomorphism between singular chain complexes, $\mu\colon C_*(X) \to C_*(\widetilde{X})$ with

$$\mu(\Delta) = \sum_{\text{lifts of } \Delta} \widetilde{\Delta}.$$

The summation runs over all lifts of the singular simplex Δ. We can see that μ commutes with the differentials, hence we can define $\mu_*\colon H_*(X) \to H_*(\widetilde{X})$ with

$$\mu_*(x) = \sum_{g\in G} g_* x. \tag{18}$$

Then both the maps $p_*\circ\mu_*\colon H_*(X) \to H_*(X)$ and $\mu_*\circ p_*\colon H_*(\widetilde{X})^G \to H_*(\widetilde{X})^G$ are multiplication by $|G| \in k$. As the characteristic of k is coprime to $|G|$, multiplication by $|G|$ is an isomorphism on a k–vector space, which implies μ_* and p_* are also isomorphisms. Thus (16) holds. □

By using this transfer map, we obtain a sequence of compositions

$$\begin{aligned} H_p(P_G M \times_G EG) \otimes H_q(P_G M \times_G EG) &\xrightarrow{\mu_*\otimes\mu_*} H_p(P_G M) \otimes H_q(P_G M) \\ &\xrightarrow{\bullet} H_{p+q-n}(P_G M) \xrightarrow{p_*} H_{p+q-\dim M}(P_G M \times_G EG), \end{aligned}$$

where p_* is the map induced from the covering map $p\colon P_G M \times EG \to P_G M \times_G EG$. The obtained map

$$\circ\colon H_p(P_G M \times_G EG) \otimes H_q(P_G M \times_G EG) \to H_{p+q-\dim M}(P_G M \times_G EG), \tag{19}$$

is the desired orbifold loop product, which coincides with the ordinary loop product when G is the trivial group by the above construction. □

Remark It is shown in [14] that the orbifold loop product $\circ$ is indeed an orbifold invariant.

3 Propositions for the proof of the main theorem

In this section, we prove some propositions that we use for the proof of the main theorem. Unless otherwise stated, we denote by M a smooth closed oriented manifold, by $\mathcal{G}$ a path-connected group acting continuously on M, and by G a finite subgroup of $\mathcal{G}$ acting smoothly on M.

The following proposition is proved in [14] for computing the loop product of some lens spaces. We prove it below because we use the proof of it in this paper.

Proposition 3.1 [14] *Let $\mathcal{G}$ be a path-connected group acting continuously on M, and G a finite subgroup of $\mathcal{G}$ acting smoothly on M. Then for each $g \in G$, there exists a homotopy equivalence*

$$P_g X \simeq LM. \tag{20}$$

Proof By the assumption on G, for each $g \in G$ there exists a path θ_g in $\mathcal{G}$ which starts at the unit in G and ends at g. We fix such paths $\{\theta_g\}_{g\in G}$.

For each $g \in G$ we consider the maps $\tau_g\colon P_g M \to LM$ and $\eta_g\colon LM \to P_g M$ with

$$\tau_g(t) = \begin{cases} \sigma(2t) & \text{for } 0 \le t \le \frac{1}{2}, \\ \sigma(1)\check{\theta}_g(2t-1) & \text{for } \frac{1}{2} \le t \le 1, \end{cases} \tag{21}$$

$$\eta_g(l) = \begin{cases} l(2t) & \text{for } 0 \le t \le \frac{1}{2}, \\ l(1)\theta_g(2t-1) & \text{for } \frac{1}{2} \le t \le 1, \end{cases} \tag{22}$$

where $\check{\theta}_g$ denotes the image of θ_g by the map $\mathcal{G} \to \mathcal{G}$ which sends an element to its inverse. Then we can see $\tau_g \circ \eta_g \simeq \mathrm{id}_{LM}$ and $\eta_g \circ \tau_g \simeq \mathrm{id}_{P_g M}$, hence $P_g M \simeq LM$ for each $g \in G$. □

We consider the map $\Phi\colon \Omega\mathcal{G} \times LM \to LM$ with

$$\Phi(a, l) = (t \mapsto a(t) \cdot l(t)), \tag{23}$$

and the induced map $\Phi_*\colon H_*(\Omega\mathcal{G}) \otimes H_*(LM) \to H_*(LM)$, which defines an action of the Pontrjagin ring $H_*(\Omega\mathcal{G})$ on $H_*(LM)$. If the action Φ_* of $\pi_0(\Omega\mathcal{G})$ on $H_*(LM)$ satisfies $\Phi_*(x) = x$ for any $x \in H_*(LM)$, we call the action Φ_* *trivial*.

Proposition 3.2 *Assume that the action Φ_* is trivial. Then the homotopy equivalence*

$$\coprod_{g\in G} P_g M \simeq \coprod_{g\in G} LM \tag{24}$$

of Proposition 3.1 is G–equivariant at homology level with coefficients in k.

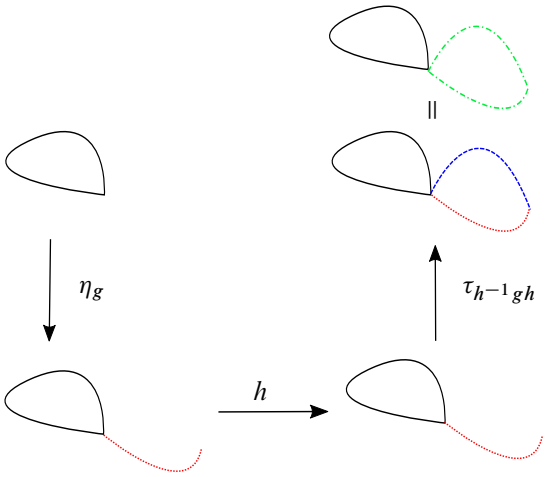


Figure 1: Picture for Proposition 3.2

We use the following lemma for the proof.

Lemma 3.3 *Let $\mathcal{G}$ be a path-connected group acting continuously on M, and G a finite subgroup of $\mathcal{G}$ acting smoothly on M. Then G acts on $H_*(LM)$ trivially.*

Proof The paths $\{\theta_g\}_{g\in G}$ in the proof of Proposition 3.1 make a homotopy between the actions of e and g on $H_*(LM)$. Because $H_*(LM)$ is discrete, they act trivially. □

Proof of Proposition 3.2 We need to show the commutativity of the following diagram for each $g, h \in G$:

$$\begin{array}{ccc} H_*(LM) & \xrightarrow{h_*} & H_*(LM) \\ {\scriptstyle \tau_{g*}}\uparrow & & \uparrow{\scriptstyle \tau_{h^{-1}gh*}} \\ H_*(P_g M) & \xrightarrow{h_*} & H_*(P_{h^{-1}gh} M) \end{array}$$

By Lemma 3.3, the upper h_* is equal to an identity. Thus we need to show

$$(\tau_{h^{-1}gh} \circ h \circ \eta_g)_* = \mathrm{id}\,.$$

Let a be the loop $h^{-1}\theta_g h * \check{\theta}_{h^{-1}gh}$ in $\mathcal{G}$, where $*$ denotes concatenating operation defined by (10); see Figure 1. Then the map $\tau_{h^{-1}gh} \circ h \circ \eta_g$ is equal to the map $\Phi(a, -)\colon LM \to LM$, and the induced homomorphism $(\tau_{h^{-1}gh} \circ h \circ \eta_g)_*$ is equal to the action $\Phi_*([a], -)\colon H_*(LM) \to H_*(LM)$, which is trivial by the assumption. □

4 Main theorem

Let τ be a direct sum

$$\tau = \bigoplus_g \tau_g \colon \coprod_{g\in G} P_g M \to \coprod_{g\in G} LM.$$

Then by Proposition 3.2 and Lemma 3.3, τ induces an isomorphism

$$H_*(L[M/G]) = H_*\Big(\coprod_{g\in G} P_g M\Big)^G \tag{25}$$

$$\overset{\tau_*}{\cong} H_*\Big(\coprod_{g\in G} LM\Big)^G \tag{26}$$

$$\cong H_*(LM)\otimes Z(k[G]). \tag{27}$$

The last isomorphism follows from Lemma 3.3.

Our main theorem, asserting that τ_* preserves loop products, is the following.

Theorem 4.1 *Let $\mathcal{G}$ be a path-connected group acting continuously on M, and G be a finite subgroup of $\mathcal{G}$ acting smoothly on M, and k be a field whose characteristic is coprime to $|G|$. If the action Φ_* is trivial with coefficients in k, then there exists an isomorphism as k–algebras,*

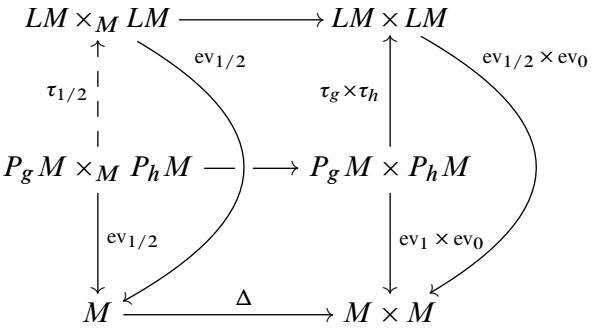
$$\mathbb{H}_*(L[M/G];k) \cong \mathbb{H}_*(LM;k)\otimes Z(k[G]), \tag{28}$$

where $Z(k[G])$ denotes the center of the group ring $k[G]$.

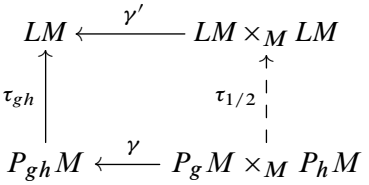
For the proof, we need to show that the diagram

$$\begin{array}{ccc} H_p\big(\coprod_g LM\big)^G \otimes H_q\big(\coprod_g LM\big)^G & \xrightarrow{\circ} & H_{p+q-\dim M}\big(\coprod_g LM\big)^G \\ \uparrow{\scriptstyle \tau_*\otimes\tau_*} & & \uparrow{\scriptstyle \tau_*} \\ H_p\big(\coprod_g P_g M\big)^G \otimes H_q\big(\coprod_g P_g M\big)^G & \xrightarrow{\circ} & H_{p+q-\dim M}\big(\coprod_g P_g M\big)^G \end{array} \tag{29}$$

is commutative. To prove this we need some lemmas. We consider the following commutative diagram of pullbacks, where $\tau_{1/2}$ is a map induced from the universality of the pullback:

(30)


Lemma 4.2 *The diagram*

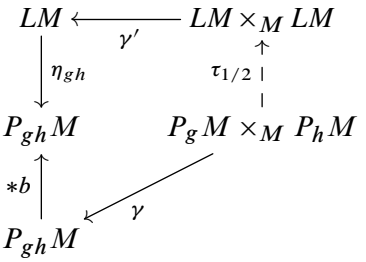
(31)


is commutative at homology level, where γ' and γ are concatenating maps defined by

$$\gamma'((l_1, l_2))(t) = \begin{cases} l_1(2t) & \text{for } 0 \le t \le \frac{1}{4}, \\ l_2\big(2t - \frac{1}{2}\big) & \text{for } \frac{1}{4} \le t \le \frac{3}{4}, \\ l_1(2t-1) & \text{for } \frac{3}{4} \le t \le 1, \end{cases} \tag{32}$$

$$\gamma((\sigma_1, \sigma_2))(t) = \begin{cases} \sigma_1(2t) & \text{for } 0 \le t \le \frac{1}{2}, \\ \sigma_2(2t-1) & \text{for } \frac{1}{2} \le t \le 1. \end{cases} \tag{33}$$

Proof Let b be the loop $h\check{\theta}_h * \check{\theta}_g * g^{-1}\theta_{gh}$ in $\mathcal{G}$ (see Figure 2). Then we have the commutative diagram

(34)


where $*b$ denotes the map $\sigma \mapsto \sigma * \sigma(1)b$. Hence it reduces to showing that $(*b)_*\colon H_*(P_{gh}M) \to H_*(P_{gh}M)$ is equal to an identity, which follows from the similar argument in the proof of Proposition 3.2. □

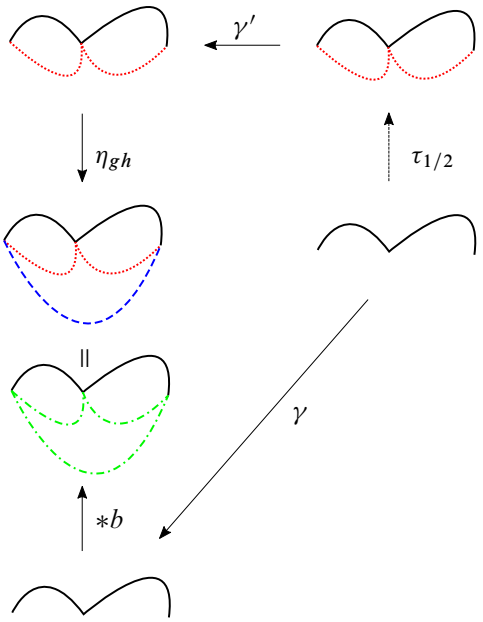


Figure 2: Picture for Lemma 4.2

Lemma 4.3 *The concatenating map γ' defines the same product as the loop product $\circ$; in other words, the following diagram commutes:*

$$
\begin{array}{ccccccc}
H_*(LM\times LM) & \xrightarrow{\mathrm{PT}} & H_*((LM\times_M LM)^{TM}) & \xrightarrow{\mathrm{Thom}} & H_{*-\dim M}(LM\times_M LM) & \xrightarrow{\gamma'_*} & H_{*-\dim M}(LM) \\
\uparrow{\scriptstyle (\rho_{1/2}\times\mathrm{id})_*} & & \uparrow{\scriptstyle \widetilde{(\rho_{1/2}\times_{1/2}\mathrm{id})}_*} & & \uparrow{\scriptstyle (\rho_{1/2}\times_{1/2}\mathrm{id})_*} & & \uparrow{\scriptstyle \rho_{1/4*}} \\
H_*(LM\times LM) & \xrightarrow{\mathrm{PT}} & H_*((LM\times_M LM)^{TM}) & \xrightarrow{\mathrm{Thom}} & H_{*-\dim M}(LM\times_M LM) & \xrightarrow{\gamma_*} & H_{*-\dim M}(LM)
\end{array}
\tag{35}
$$

Proof It follows from the commutative diagram

$$
\begin{array}{ccccc}
LM & \xleftarrow{\gamma'} & LM\times_M LM & \longrightarrow & LM\times LM \\
\uparrow{\scriptstyle \rho_{1/4}} & & \uparrow{\scriptstyle \rho_{1/2}\times_{1/2}\mathrm{id}} & & \uparrow{\scriptstyle \rho_{1/2}\times\mathrm{id}} \\
LM & \xleftarrow[\gamma]{} & LM\times_M LM & \longrightarrow & LM\times LM \\
 & & \downarrow & & \downarrow{\scriptstyle \mathrm{ev}_0\times\mathrm{ev}_0} \\
 & & M & \xrightarrow{\Delta} & M\times M
\end{array}
\qquad \mathrm{ev}_{1/2}\times\mathrm{ev}_0 : LM\times_M LM \to M\times M
\tag{36}
$$

and the naturality of the Pontrjagin–Thom construction and Thom isomorphism, where ρ_a is the parameter transformation map defined by $\rho_a(l) = (t \mapsto l(t+a))$. □

Proposition 4.4 *For any $g, h \in G$, the following diagram commutes:*

$$\begin{array}{ccc} H_p(LM)\otimes H_q(LM) & \xrightarrow{\circ} & H_{p+q-\dim M}(LM) \\ {\scriptstyle \tau_{g*}\otimes\tau_{h*}}\uparrow & & \uparrow{\scriptstyle \tau_{gh*}} \\ H_p(P_gM)\otimes H_q(P_hM) & \xrightarrow{\bullet} & H_{p+q-\dim M}(P_{gh}M) \end{array} \tag{37}$$

Proof Because of the diagram (30) and Lemma 4.2, by the naturality of the Pontrjagin–Thom map and Thom isomorphism we obtain the following commutative diagram:

$$\begin{array}{ccccccc} H_*(LM\times LM) & \longrightarrow & H_*((LM\times_M LM)^{TM}) & \longrightarrow & H_{*-\dim M}(LM\times_M LM) & \longrightarrow & H_{*-\dim M}(LM) \\ \uparrow{\scriptstyle (\tau_g\times\tau_h)_*} & & \uparrow{\scriptstyle \widetilde{\tau_{1/2*}}} & & \uparrow{\scriptstyle \tau_{1/2*}} & & \uparrow{\scriptstyle \tau_{gh*}} \\ H_*(P_gM\times P_hM) & \to & H_*((P_gM\times_M P_hM)^{TM}) & \to & H_{*-\dim M}(P_gM\times_M P_hM) & \to & H_{*-\dim M}(P_{gh}M) \end{array} \tag{38}$$

which implies the diagram (37) is commutative by Lemma 4.3. □

Proof of Theorem 4.1 By the definition of the loop product, diagram (29) is same as the following diagram:

$$\begin{array}{ccc} H_p\big(\coprod_g P_gM\big)^G\otimes H_q\big(\coprod_g P_gM\big)^G & \xrightarrow{\tau_*\otimes\tau_*} & H_p\big(\coprod_g LM\big)^G\otimes H_q\big(\coprod_g LM\big)^G \\ \downarrow & & \downarrow \\ H_p\big(\coprod_g P_gM\big)\otimes H_q\big(\coprod_g P_gM\big) & \xrightarrow{\tau_*\otimes\tau_*} & H_p\big(\coprod_g LM\big)\otimes H_q\big(\coprod_g LM\big) \\ \downarrow{\scriptstyle \bullet} & & \downarrow{\scriptstyle \circ} \\ H_{p+q-\dim M}\big(\coprod_g P_gM\big) & \xrightarrow{\tau_*} & H_{p+q-\dim M}\big(\coprod_g LM\big) \\ \downarrow{\scriptstyle \sigma_{1*}} & & \downarrow{\scriptstyle \sigma_{2*}} \\ H_{p+q-\dim M}\big(\coprod_g P_gM\big)^G & \xrightarrow{\tau_*} & H_{p+q-\dim M}\big(\coprod_g LM\big)^G \end{array} \tag{39}$$

where σ_{1*}, σ_{2*} are defined as

$$\sigma_{1*}\colon H_*\big(\coprod_g P_gM\big) \xrightarrow{\mathrm{pr}_*} H_*\big(\big(\coprod_g P_gM\big)/G\big) \xrightarrow[\cong]{\mu_*} H_*\big(\coprod_g P_gM\big)^G, \tag{40}$$

$$\sigma_{2*}\colon H_*(LM\times G) \xrightarrow{\mathrm{pr}_*} H_*((LM\times G)/G) \xrightarrow[\cong]{\mu_*} H_*(LM\times G)^G, \tag{41}$$

where μ_* is as defined in the proof of Lemma 2.10. The G–equivariance of τ_* implies the commutativity of the upper square, and Proposition 4.4 implies the commutativity

of the middle square. The commutativity of the lower square follows from the definition of μ_* and G–equivariance of τ_*. □

5 Relation to Hochschild cohomology

In [5], the theorem of Cohen and Jones states that if M is simply connected, then its loop homology ring is isomorphic to the Hochschild cohomology of the singular cochains, hence we have $\mathbb{H}_*(LM) \cong \mathrm{HH}^*(C^*M, C^*M)$. In [2], this result is studied for global quotient orbifolds by Ángel, Backelin and Uribe, who show that if M is a simply connected manifold with G action, then its orbifold loop homology ring with coefficients in k is isomorphic to the Hochschild cohomology of the DGA $C^*(M;k)\#_k G$, hence we have $\mathbb{H}_*(L[M/G];k) \cong \mathrm{HH}^*(C^*M\#_k G, C^*M\#_k G)$. Here the DGA structure on $C^*(M;k)\#_k G$ is defined as follows. With $C^*(M;k)\#_k G = C^*(M;k) \otimes_k k[G]$ as a module, the multiplication is defined by $(x \otimes g)(y \otimes h) = xg(y) \otimes gh$, and the differential is defined by $d = d_{C^*M} \otimes \mathrm{id}_{k[G]}$.

By using our theorem, we obtain the following as a corollary.

Corollary 5.1 *Under the situation we consider in the theorem, if Φ_* is trivial and M is simply connected, then it holds that*

$$\mathbb{H}_*(L[M/G];k) \cong \mathrm{HH}^*(C^*M, C^*M) \otimes Z(k[G]) \tag{42}$$

$$\cong \mathrm{HH}^*(C^*M\#_k G, C^*M\#_k G). \tag{43}$$

Proof The first equation follows from our theorem and the theorem of Cohen and Jones, and the second one follows from the above theorem of Ángel, Backelin and Uribe. □

6 Examples

In this section, we will compute some examples of orbifold loop homology. Before that, we will prove some propositions which are useful to apply our main theorem and to compute examples.

Proposition 6.1 *If $\mathcal{G}$ is simply connected, then the action Φ_* is trivial for any field k.*

Proof Because $\pi_0(\Omega\mathcal{G})$ is represented only by the identity loop, its action on $H_*(LM)$ is trivial. □

Proposition 6.2 *If $H_{\dim M}(LM;k)=k$ for a field k, then the action Φ_* is trivial for any pair $(\mathcal{G},G)$.*

To prove this proposition, we use the following lemma by Hepworth.

Lemma 6.3 [9] *Let Φ be the natural map defined by* (23)*. Then the induced linear action $\Phi_*\colon H_0(\Omega\mathcal{G})\otimes H_*(LM)\to H_*(LM)$ is an algebra action, in other words for any $\alpha\in H_0(\Omega\mathcal{G})$ and any $x,y\in H_*(LM)$, they satisfy*

$$\alpha\cdot(x\circ y)=(\alpha\cdot x)\circ y=x\circ(\alpha\cdot y). \tag{44}$$

Proof Let a be a loop in $\mathcal{G}$ which starts and ends at the unit. We need to show that the action $\Phi_{a*}=\Phi_*([a],-)\colon H_*(LM)\to H_*(LM)$ is trivial, where Φ_a denotes the map $\Phi(a,-)$. We also denote by Φ_a the restriction map of the map $\Phi(a,-)\colon LM\to LM$ to $M\subset LM$, where we regard M as the image of the map $c\colon M\to LM$ which assigns the constant loop.

We consider the following sequence of maps:

(45)
$$H_{\dim M}(M)\underset{c_*}{\overset{\Phi_{a*}}{\rightrightarrows}}H_{\dim M}(LM)\xrightarrow{\mathrm{ev}_*}H_{\dim M}(M),\qquad \mathrm{id}_*\colon H_{\dim M}(M)\to H_{\dim M}(M)$$

As $H_{\dim M}(LM)=H_{\dim M}(M)=k$, we obtain $\Phi_{a*}=c_*=\mathrm{id}$. Hence for any $x\in H_*(LM)$, we have by Lemma 6.3 the equality

$$\Phi_{a*}x=\Phi_{a*}([M]\circ x)=(\Phi_{a*}[M])\circ x=x.$$

Therefore the action of $\pi_0(\Omega\mathcal{G})$ on $H_*(LM)$ is trivial. □

Proposition 6.4 *Let k be a field. If the following conditions are satisfied, then the action Φ_* is trivial.*

(i) $\pi_1 M=1$.

(ii) $|\pi_1\mathcal{G}|=:r<\infty$.

(iii) *The homomorphism $H_*(\Omega M;k)\to H_*(LM;k)$ induced by the inclusion map $\Omega M\to LM$ is injective; that is, the free loop fibration $\Omega M\to LM\to M$ is* ***totally noncohomologous to zero*** *(TNCZ) with respect to the field k.*

(iv) *The characteristic of k is coprime to r.*

For the proof of Proposition 6.4, we use the following lemma by Hepworth.

Lemma 6.5 [9] *Let $\mathcal{G}$ be a topological group acting continuously on a closed oriented manifold M. Then the homomorphism $\Phi'_*\colon H_*(\Omega\mathcal{G};k)\to H_*(LM;k)$ with*

$$\Phi'_*[a]=[a]\cdot[M] \tag{46}$$

commutes with the products, ie they satisfy

$$\Phi'_*[a]\circ\Phi'_*[a']=\Phi'_*([a]\cdot[a']), \tag{47}$$

where the product on the right-hand side denotes the Pontrjagin product.

Proof of Proposition 6.4 We denote $\dim M$ by n below. Let a be a loop in $\mathcal{G}$ which starts and ends at the unit. By the arguments in the proof of Proposition 6.2, we need to show that $\Phi_{a*}[M]$ is the unit in $\mathbb{H}_0(LM)$ for the triviality of the action of $\pi_0(\Omega\mathcal{G})$. By the TNCZ assumption, we have the ring isomorphism

$$\mathbb{H}_*(LM)\cong H_*(\Omega M)\otimes\mathbb{H}_*(M), \tag{48}$$

hence we obtain the linear isomorphism

$$H_n(LM)\cong\bigoplus_{p+q=n}H_p(\Omega M)\otimes H_q(M). \tag{49}$$

By the assumption of $\pi_1 M=1$, we have $H_0(\Omega M)\cong k$, hence $H_0(\Omega M)\otimes H_n(M)$ is a 1–dimensional vector space, and we can put

$$\Phi_{a*}[M]=x\cdot 1+\sum_i y_i\cdot\sigma_i\otimes\delta_i, \tag{50}$$

where

$$1=1\otimes[M]\in H_0(\Omega M)\otimes H_n(M),\quad \sigma_i\in H_{>0}(\Omega M),\quad \delta_i\in H_{<n}(M),$$
$$\deg\sigma_i+\deg\delta_i=n,\quad x,y_i\in k.$$

We can see $x=1$ as follows. We have a ring homomorphism induced by the evaluation map $\mathrm{ev}_*\colon\mathbb{H}_*(LM)\to\mathbb{H}_*(M)$. Hence we have

$$\mathrm{ev}_*(\sigma_i\otimes\delta_i)=\mathrm{ev}_*(\sigma_i\otimes[M])\circ\mathrm{ev}_*(1\otimes\delta_i). \tag{51}$$

Because the degree of σ_i is positive, the degree of $\sigma_i\otimes[M]$ is larger than n, hence $\mathrm{ev}_*(\sigma_*\otimes[M])$ is equal to 0. Therefore, we obtain

$$\mathrm{ev}_*\,\Phi_{a*}[M]=x\cdot\mathrm{ev}_*\,1+\sum_i y_i\cdot\mathrm{ev}_*(\sigma_i\otimes\delta_i)=x\cdot\mathrm{ev}_*\,1.$$

Moreover, we have the diagram

(52)
$$\begin{array}{ccccc} & & \overset{\mathrm{id}_*}{\frown} & & \\ H_n(M) & \xrightarrow{\Phi_{a*}} & H_n(LM) & \xrightarrow{\mathrm{ev}_*} & H_n(M) \end{array}$$

Thus we have

$$x\cdot[M] = x\cdot \mathrm{ev}_* 1 = \mathrm{ev}_* \circ \Phi_{a*}[M] = [M], \tag{53}$$

hence we obtain $x = 1$. Furthermore, we can show that the y_i are all 0 as follows. Because the δ_i are all nilpotent, we have

$$(\Phi_{a*}[M] - 1)^N = 0 \tag{54}$$

for sufficiently large N. By the assumption $|\pi_1\mathcal{G}| = r < \infty$ and Lemma 6.5, we have

$$(\Phi_{a*}[M])^r = (\Phi'_*[a])^r = \Phi'_*[a]^r = \Phi'_*[1] = 1. \tag{55}$$

Since the characteristic of k is coprime to r, the common divisor of $(\Phi_{a*}[M]-1)^N$ and $(\Phi_{a*}[M])^r - 1$ is $\Phi_{a*}[M]-1$. Hence we conclude that the y_i are all 0, which implies $\Phi_{a*}[M] = 1$. □

Example 6.6 We consider the case $(M, \mathcal{G}, G, k) = (S^3, \mathrm{Spin}(3), \pi, k)$, where

$$\pi = \langle a, b, c \mid a^2 = b^2 = c^2 = (ab)^2 = (bc)^3 = (ca)^5 \rangle$$

and k is an algebraic closed field whose characteristic is not 2, 3 or 5. The finite group π acts on S^3 and the quotient manifold S^3/π is called the *Poincaré homology sphere*. Then we have an algebra isomorphism

$$\begin{aligned} \mathbb{H}_*(L(S^3/\pi); k) &\cong \mathbb{H}_*(S^3; k) \otimes Z(k[\pi]) \\ &\cong \mathbb{H}_*(S^3; k) \otimes k^9. \end{aligned}$$

Here we use the fact that if k is an algebraic closed field and G is a finite group, then $Z(k[G]) \cong k^{c(G)}$, where $c(G)$ denotes the number of conjugacy classes of G.

Remark If $k = \mathbb{Q}$, Vigué [21] shows that the following are equivalent.

(i) The homomorphism $H_*(\Omega M; k) \to H_*(LM; k)$ induced by the inclusion $\Omega M \to LM$ is injective.

(ii) $H^*(M; k)$ is a free graded commutative algebra.

Remark For $M = S^n$ or $\mathbb{C}\mathrm{P}^n$, Menichi [16] shows that the following are equivalent.

(i) The homomorphism induced by the inclusion $H_*(\Omega M; k) \to H_*(LM; k)$ is injective.

(ii) The Euler number $\chi(M)$ is 0 in k.

Remark The necessary condition for the free loop fibration

$$\Omega M \to LM \to M \tag{56}$$

to be the TNCZ fibration with respect to a finite field $\mathbb{F}_p$ has been studied for many homogeneous spaces by Kuribayashi [13]. For example:

- For $M = \mathbb{C}\mathrm{P}^n$ or $\mathbb{H}\mathrm{P}^n$, the fibration (56) is TNCZ if and only if the Euler number $\chi(M)$ is 0 in $\mathbb{F}_p$.
- For $M = \mathrm{SU}(m+n)/\mathrm{SU}(n)$ or $\mathrm{Sp}(m+n)/\mathrm{Sp}(n)$, the fibration (56) is TNCZ for any $m, n \geq 0,\ p > 0$.
- For $\mathrm{SO}(m+n)/\mathrm{SO}(n)$ and $p > 2$, the fibration (56) is TNCZ if and only if n is odd.
- For $\mathrm{SO}(m+n)/\mathrm{SO}(n)$ and $p = 2$, the fibration (56) is TNCZ if $m \leq 4$ or $1 \leq m \leq 8$ and $n \geq 43$.
- For $U(m+n)/U(m) \times U(n)$ and $\mathrm{Sp}(m+n)/\mathrm{Sp}(m) \times \mathrm{Sp}(n)$, the fibration (56) is *not* TNCZ for any $m, n \geq 0,\ p > 0$.

References

[1] **A Adem**, **J Leida**, **Y Ruan**, *Orbifolds and stringy topology*, Cambridge Tracts in Mathematics 171, Cambridge Univ. Press (2007) MR

[2] **A Ángel**, **E Backelin**, **B Uribe**, *Hochschild cohomology and string topology of global quotient orbifolds*, J. Topol. 5 (2012) 593–638 MR

[3] **M Chas**, **D Sullivan**, *String topology*, preprint (1999) arXiv

[4] **D Chataur**, **L Menichi**, *String topology of classifying spaces*, J. Reine Angew. Math. 669 (2012) 1–45 MR

[5] **R L Cohen**, **J D S Jones**, *A homotopy theoretic realization of string topology*, Math. Ann. 324 (2002) 773–798 MR

[6] **R L Cohen**, **J D S Jones**, **J Yan**, *The loop homology algebra of spheres and projective spaces*, from "Categorical decomposition techniques in algebraic topology" (G Arone, J Hubbuck, R Levi, M Weiss, editors), Progr. Math. 215, Birkhäuser, Basel (2004) 77–92 MR

[7] **R L Cohen**, **J R Klein**, *Umkehr maps*, Homology Homotopy Appl. 11 (2009) 17–33 MR

[8] **K Gruher**, **P Salvatore**, *Generalized string topology operations*, Proc. Lond. Math. Soc. 96 (2008) 78–106 MR

[9] **R A Hepworth**, *String topology for complex projective spaces*, preprint (2009) arXiv

[10] **R A Hepworth**, *String topology for Lie groups*, J. Topol. 3 (2010) 424–442 MR

[11] **S Kaji**, **H Tene**, *Products in equivariant homology*, preprint (2015) arXiv

[12] **A P M Kupers**, *An elementary proof of the string topology structure of compact oriented surfaces*, preprint (2012) arXiv

[13] **K Kuribayashi**, *On the mod p cohomology of the spaces of free loops on the Grassmann and Stiefel manifolds*, J. Math. Soc. Japan 43 (1991) 331–346 MR

[14] **E Lupercio**, **B Uribe**, **M A Xicoténcatl**, *Orbifold string topology*, Geom. Topol. 12 (2008) 2203–2247 MR

[15] **L Menichi**, *String topology for spheres*, Comment. Math. Helv. 84 (2009) 135–157 MR

[16] **L Menichi**, *String topology, Euler class and TNCZ free loop fibrations*, preprint (2013) arXiv

[17] **I Moerdijk**, *Orbifolds as groupoids: an introduction*, from "Orbifolds in mathematics and physics" (A Adem, J Morava, Y Ruan, editors), Contemp. Math. 310, Amer. Math. Soc., Providence, RI (2002) 205–222 MR

[18] **H Tamanoi**, *Batalin–Vilkovisky Lie algebra structure on the loop homology of complex Stiefel manifolds*, Int. Math. Res. Not. 2006 (2006) art. id. 97193 MR

[19] **H Tamanoi**, *Cap products in string topology*, Algebr. Geom. Topol. 9 (2009) 1201–1224 MR

[20] **D Vaintrob**, *The string topology BV algebra, Hochschild cohomology and the Goldman bracket on surfaces*, preprint (2007) arXiv

[21] **M Vigué-Poirrier**, *Dans le fibré de l'espace des lacets libres, la fibre n'est pas, en général, totalement non cohomologue à zéro*, Math. Z. 181 (1982) 537–542 MR

Graduate School of Mathematical Sciences, The University of Tokyo
Tokyo, Japan

asao@ms.u-tokyo.ac.jp

Received: 8 June 2017 Revised: 11 July 2017

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 LaTeX, preferably using the journal's class file. More information on preparing articles in 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 BibTeX 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 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 18 Issue 1 (pages 1–633) 2018

Width of a satellite knot and its companion 1
QILONG GUO and ZHENKUN LI

The closed-open string map for S^1–invariant Lagrangians 15
DMITRY TONKONOG

Heegaard Floer homology and knots determined by their complements 69
FYODOR GAINULLIN

Classification of tight contact structures on small Seifert fibered L–spaces 111
IRENA MATKOVIČ

An infinite family of links with critical bridge spheres 153
DANIEL RODMAN

A second cohomology class of the symplectomorphism group with the discrete topology 187
RYOJI KASAGAWA

On the stability of asymptotic property C for products and some group extensions 221
GREGORY COPELAND BELL and ANDRZEJ NAGÓRKO

Higher cohomology operations and R–completion 247
DAVID BLANC and DEBASIS SEN

On high-dimensional representations of knot groups 313
STEFAN FRIEDL and MICHAEL HEUSENER

DAHA and plane curve singularities 333
IVAN CHEREDNIK and IAN PHILIPP

Inertia groups of high-dimensional complex projective spaces 387
SAMIK BASU and RAMESH KASILINGAM

Nonexistence of boundary maps for some hierarchically hyperbolic spaces 409
SARAH C MOUSLEY

Finite Dehn surgeries on knots in S^3 441
YI NI and XINGRU ZHANG

A characterization for asymptotic dimension growth 493
GOULNARA ARZHANTSEVA, GRAHAM A NIBLO, NICK WRIGHT and JIAWEN ZHANG

Classifying spaces for 1–truncated compact Lie groups 525
CHARLES REZK

Second mod 2 homology of Artin groups 547
TOSHIYUKI AKITA and YE LIU

On the third homotopy group of Orr's space 569
EMMANUEL DROR FARJOUN and ROMAN MIKHAILOV

Groups of homotopy classes of phantom maps 583
HIROSHI KIHARA

Loop homology of some global quotient orbifolds 613
YASUHIKO ASAO

1472-2747(2018)18:1;1-R