Algebraic & Geometric Topology

Volume 17 (2017)

Issue 6 (pages 3213–3852)

msp

ALGEBRAIC & GEOMETRIC TOPOLOGY
msp.org/agt

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

The subscription price for 2017 is US $410/year for the electronic version, and $650/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/

Algebraic & Geometric Topology 17 (2017) 3213–3257

3–manifolds built from injective handlebodies

JAMES COFFEY
HYAM RUBINSTEIN

This paper studies a class of closed orientable 3–manifolds constructed from a gluing of three handlebodies, such that the inclusion of each handlebody is π_1–injective. This construction is the generalisation to handlebodies of the construction for gluing three solid tori to produce non-Haken Seifert fibred 3–manifolds with infinite fundamental group. It is shown that there is an efficient algorithm to decide if a gluing of handlebodies satisfies the disk-condition. Also, an outline for the construction of the characteristic variety (JSJ decomposition) in such manifolds is given. Some non-Haken and atoroidal examples are given.

57N10, 57M10, 57M50

1 Introduction

This paper is concerned with the class of 3–manifolds that meet the disk-condition. These are closed orientable 3–manifolds constructed from the gluing of three handlebodies, such that the induced map on the fundamental group of each of the handlebodies is injective. Thus all manifolds that meet the disk-condition have infinite fundamental group. The disk-condition is an extension to handlebodies of conditions for the gluing of three solid tori to produce non-Haken Seifert fibred manifolds with infinite fundamental group. These manifolds appear to have many nice properties. In this paper, some tools for understanding manifolds that meet the disk-condition are investigated. A number of constructions are given for this class, including some manifolds that are non-Haken and some that are atoroidal. The characteristic variety of manifolds that meet the disk-condition is also investigated. It is shown that the handlebody structure carries all the information for building the characteristic variety.

In Section 2, standard definitions that are used throughout this paper are given. Also, the "disk-condition" is defined and discussed. In particular, it is shown how this condition is a generalisation of the construction of non-Haken Seifert fibred manifolds with infinite fundamental group. We also discuss how, on an intuitive level, the class of manifolds that meet the disk-condition contains many other non-Haken examples.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3213

Section 3 is divided into three subsections. The first develops some basic tools and also shows that all 3–manifolds that meet the disk-condition have infinite fundamental group and are irreducible. In the second subsection, a sufficient condition is given for gluings of handlebodies to meet the disk-condition. This condition is easily checked and useful for constructing examples. We then give a necessary and sufficient condition and an algorithm that can be checked in bounded time. The final part gives some constructions of manifolds that meet the disk-condition, using Dehn fillings along knots in $\mathbb{S}^3$ and n–fold cyclic branched covers of knots in $\mathbb{S}^3$. Some non-Haken examples are produced.

Section 4 is concerned with the construction of the characteristic variety Σ in a manifold M that satisfies the disk-condition. The main theorem proved in Section 4 is:

Theorem 1.1 *Let M be a closed orientable 3–manifold that satisfies the disk-condition, and let T be a torus. If $f\colon T \to M$ is a π_1–injective map, then there is $\Sigma \subseteq M$ a Seifert fibred submanifold with essential boundary and a map $g\colon T \to M$ homotopic to f such that $g(T) \subset \Sigma$.*

If the characteristic variety Σ has nonempty boundary, then the boundary components are essential embedded tori. Therefore, a direct corollary of the above theorem is:

Corollary 1.2 *If M is a closed orientable 3–manifold that satisfies the disk-condition and there is a π_1–injective map of the torus into M, then either there is a π_1–injective embedding of a torus in M, or M is a non-Haken Seifert fibred manifold.*

These are not new results. However, the aim is to examine how the characteristic variety behaves in manifolds that meet the disk-condition. The proof of the torus theorem (Theorem 1.1) is constructive and gives an algorithm for finding the characteristic variety of manifolds that meet the disk-condition. In the construction of the characteristic variety, the components come in two "flavours". The intersection of all three handlebodies in the manifold is a set of injective simple closed curves, called the triple curves. The first flavour is a component which is disjoint from the triple curves. These components are similar to the constructions used by W Jaco and P Shalen to prove the torus theorem for Haken manifolds; see Jaco [6]. The intersections of the components of the characteristic variety with each handlebody are either essential Seifert fibred submanifolds or I–bundles. If we remove an open neighbourhood of the triple curves, we get a manifold with incompressible boundary, which is therefore Haken. What remains of the boundaries of the handlebodies after the triple curves are removed is a set of disjoint spanning surfaces. Therefore, the fact that these carry all the information for the characteristic variety components disjoint from the triple curves is not surprising.

We will refer to the second flavour of characteristic variety as the disk components. The intersections of these disk components with the handlebodies are regular neighbourhoods of intersecting meridian disks. For this flavour of characteristic variety components to occur, the manifold must meet a minimal disk-condition, as described in Section 2. The two flavours of characteristic variety components are not necessarily disjoint. If two such components intersect, their fibrings can always be made to agree. In fact, when they intersect, the disk components are thickened compressing annuli of the characteristic variety components disjoint from the triple curves.

Acknowledgements The authors would like to thank Ian Agol for a very helpful comment on this project. We would also like to thank the referee for extremely diligent work, which has greatly improved this paper. This research was partially supported by the Australian Research Council.

2 Definitions and preliminaries

Throughout this paper, we will assume that, unless stated otherwise, we are working in the PL category of manifolds and maps. We will use standard PL constructions, such as regular neighbourhoods and transversality, defined by C Rourke and B Sanderson in [12]. Other definitions relating to 3–manifolds are given by J Hempel in [5] or Jaco in [6].

A manifold M is *closed* if it is compact and $\partial M = \varnothing$. Also, M is *irreducible* if every embedded S^2 bounds a ball. We will assume, unless otherwise stated, that all 3–manifolds are orientable. The reason for this is that all closed nonorientable $\mathbb{P}^2$–irreducible 3–manifolds are Haken. (A manifold is $\mathbb{P}^2$–irreducible if it is irreducible and does not contain any embedded 2–sided projective planes.) A main motivation for our approach is to find constructions of non-Haken 3–manifolds.

A map $f\colon S \to M$ is *proper* if $f^{-1}(\partial M) = \partial S$. If $F\colon S \times I \to M$ is a homotopy/isotopy such that $F|_{S\times 0}$ is a proper map, then it is assumed, unless otherwise stated, that $F|_{S\times t}$ is a proper map for all $t \in I$. To simplify notation, an isotopy/homotopy of a surface $S \subset M$ is used without defining the map. Here we are assuming that there is a map $f\colon S \to M$, and we are referring to an isotopy/homotopy of f. If M is a 3–manifold and S is a compact surface which is not a sphere, disk or projective plane, the proper map $f\colon S \to M$ is called π_1*–injective* if the induced map $f_*\colon \pi_1(S) \to \pi_1(M)$ is injective. If a π_1–injective map f is not homotopic as a map of pairs $(S, \partial S) \to (M, \partial M)$ into ∂M, then the map is called *essential*.

If H is a handlebody and D is a properly embedded disk in H such that ∂D is essential in ∂H, then D is a *meridian disk* of H. If D is a proper singular disk in H such that ∂D is essential in ∂H, then it is called a *singular meridian disk.*

In this paper, normal curve theory, as defined by S Matveev in [9], is used to list finite classes of curves in surfaces. A triangulation of the surface is required to define normal curves. The surfaces may have polygonal faces. However, a barycentric subdivision will produce the required triangulation.

2.1 The disk-condition

Before we discuss the disk-condition in closed 3–manifolds, we define some useful objects and the disk-condition in handlebodies.

Definition 2.1 Let H be a handlebody, $\mathcal{T}$ a set of curves in ∂H and D a meridian disk. Assuming that ∂D and $\mathcal{T}$ are transverse, $|D|$ will denote the number of intersection points of ∂D and $\mathcal{T}$.

Definition 2.2 If H is a handlebody and $\mathcal{T}$ is a set of essential disjoint simple closed curves in ∂H, then $\mathcal{T}$ satisfies the n *disk-condition* in H if $|D| \geq n$ for every meridian disk D.

This seems a difficult condition to verify, for if H has genus two or higher, there are an infinite number of meridian disks to check. However, later we give some sufficient conditions that are easily checked and an algorithm that determines if the disk-condition is satisfied.

Next we give a construction of 3–manifolds that meet the disk-condition. Please note that even though this description is technically correct, it is not enlightening, so later we discuss different ways of describing these manifolds that are much more useful.

Let H_1, H_2 and H_3 be three handlebodies. Let $S_{i,j}$, for $i \neq j$, be a subsurface of ∂H_i such that:

(1) $\partial S_{i,j} \neq \varnothing$.

(2) The induced map of $\pi_1(S_{i,j})$ into $\pi_1(H_i)$ is injective.

(3) $S_{i,j} \cup S_{i,k} = \partial H_i$ for $j \neq k$.

(4) $\mathcal{T}_i = S_{i,j} \cap S_{i,k} = \partial S_{i,j} = \partial S_{i,k}$ is a set of disjoint essential simple closed curves that meet the n_i disk-condition in H_i.

(5) $S_{i,j} \subset \partial H_i$ is homeomorphic to $S_{j,i} \subset \partial H_j$.

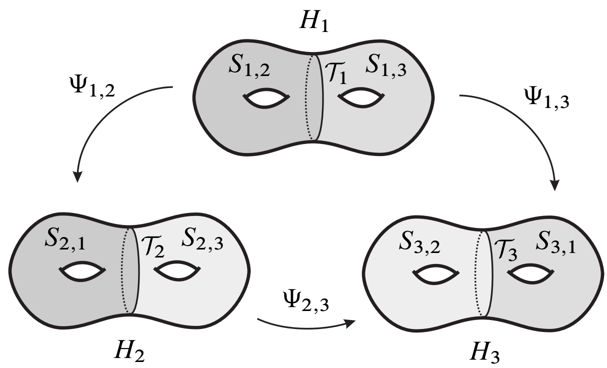


Figure 1: Homeomorphisms between boundaries of handlebodies

Note that $S_{i,j}$ need not be connected. Given that the boundary of each handlebody is cut up into π_1–injective regions, we glue the handlebodies together by homeomorphisms $\Psi_{i,j}\colon S_{i,j} \to S_{j,i}$ that agree along the $\mathcal{T}_i$; see Figure 1. The result is a closed 3–manifold M for which the image of each handlebody is embedded.

Definition 2.3 If M is a manifold constructed from three handlebodies as above such that $\mathcal{T}_i$ satisfies the n_i disk-condition in H_i and

$$\sum_{i=1,2,3} \frac{1}{n_i} \le \frac{1}{2}, \tag{1}$$

then M satisfies the (n_1, n_2, n_3) *disk-condition*. M is simply said to meet the *disk-condition* if the specific (n_1, n_2, n_3) is understood from the context.

As said previously, the above definition is not very enlightening. Thus, from now on, we view 3–manifolds that meet the disk-condition in the following way. Assume that M is a manifold that satisfies the disk-condition and H_1, H_2 and H_3 are the images of the handlebodies in M. Then $M = \bigcup_{i=1,2,3} H_i$, and each H_i is embedded in M. Then $X = \bigcup_{i=1,2,3} \partial H_i$ cuts M up into handlebodies. X can be viewed as a 2–complex by splitting up each of the surfaces forming X into cells. Also, $\mathcal{T} = \bigcap_{i=1,2,3} H_i$ is a set of essential disjoint simple closed curves in M that satisfies the n_i disk-condition in H_i where $\sum_{i=1,2,3} 1/n_i \le \frac{1}{2}$.

It may seem confusing that we are using the same name for the conditions for the construction of 3–manifolds and the curves in the boundary of handlebodies. However, the curve condition is the restriction of the condition on closed 3–manifolds to each of its component handlebodies. When we have an equality in (1), the result is the three "minimal" cases for the disk-condition. These are: $(6,6,6)$, $(4,8,8)$ or $(4,6,12)$. These three cases are of special interest since if a manifold satisfies the disk-condition,

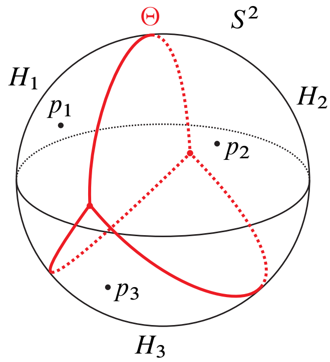


Figure 2: Base space of non-Haken Seifert fibred space with infinite π_1

then it meets at least one of these three conditions. Therefore, these are the key cases to consider. It is also worth noting that unlike Heegaard splittings, we don't require the three handlebodies to have the same genera.

Another way of viewing a 3–manifold M that satisfies the disk-condition is that $X = \bigcup \partial H_i$ is a 2–complex such that the triple curves $\mathcal{T}$ consist of essential curves in X. Therefore, we obtain a manifold M that satisfies the disk condition by gluing handlebodies to X such that each meridian disk of the handlebodies intersects $\mathcal{T}$ enough times. In fact, the disk-condition is an extension of the construction of non-Haken Seifert fibred 3–manifolds with infinite fundamental group. In the latter case, if a Seifert fibred space is non-Haken with infinite fundamental group, then it has a fibring with base space a 2–sphere, and it has three exceptional fibres of multiplicity p_i, where $\sum 1/p_i \le 1$ (*), as in Figure 2. For more details, see P Scott in [13]. If the inequality (*) is made an equality, the exceptional fibres have indices $(3,3,3)$, $(2,4,4)$ or $(2,3,6)$. Another way of viewing this construction is if Θ is the graph in S^2 shown in Figure 2, then $\Theta \times \mathbb{S}^1$ is a 2–complex X consisting of three annuli glued together along two triple curves $\mathcal{T}$. Then glue three solid tori H_i to X so that the boundaries of the meridian disks meet each triple curve p_i times. As there are two triple curves in $\mathcal{T}$, each meridian disk has $2p_i$ intersections with $\mathcal{T}$. Thus, as $\sum 1/(2p_i) \le \frac{1}{2}$, all non-Haken Seifert fibred manifolds with infinite π_1 are in the class of manifolds that meet the disk-condition.

Yet another way of viewing 3–manifolds that meet the disk-condition is if we glue two handlebodies together to form a 3–manifold with a single incompressible boundary component. Then glue a handlebody to this boundary component. A very short hierarchy in a closed Haken manifold, as defined by I Aitchison and H Rubinstein in [1], can be built from a set of handlebodies, gluing each handlebody to itself so that each of the resulting manifolds has incompressible boundary. Then glue these incompressible boundaries together to produce the closed manifold. So the incompressible boundaries become incompressible surfaces in the Haken manifold. This suggests that the disk-condition is a weaker condition than the manifold being Haken. In fact, we already know that the class of manifolds satisfying the disk-condition contains all non-Haken Seifert fibred manifolds with infinite π_1, but it also contains examples of other non-Haken manifolds.

The disk-condition can be easily extended to gluings of four or more handlebodies such that all the statements in this paper follow. Construct a closed manifold M by gluing together $r \geq 3$ handlebodies $H_1, \dots, H_r$ such that, for i, j, k and l different,

- H_i is embedded,
- $H_i \cap H_j \subset \partial H_i \cap \partial H_j$ is a subsurface,
- $H_i \cap H_j \cap H_k$ is a possibly empty set of pairwise disjoint curves, and
- $H_i \cap H_j \cap H_k \cap H_l = \varnothing$.

Then $X = \bigcup_{1 \leq i < j \leq r} H_i \cap H_j$ is a 2–complex which cuts M up into the H_i, and $\mathcal{T} = \bigcup_{1 \leq i < j < k \leq r} H_i \cap H_j \cap H_k$ is a union of pairwise disjoint simple closed curves. Suppose α is a component of $\mathcal{T}$. Let H_{α_1}, H_{α_2} and H_{α_3} be the three handlebodies around α and suppose that $\mathcal{T}$ satisfies the n_{α_i} disk-condition in H_{α_i}. Then M satisfies the *generalised disk-condition* if $\sum_{i=1,2,3} 1/n_{\alpha_i} \leq \frac{1}{2}$ for each $\alpha \in \mathcal{T}$. For the purposes of this paper, we will not consider such manifolds for $r \geq 4$ as they are all Haken. To see this, if $r \geq 4$, then we can choose H_i and H_j such that $H_i \cap H_j \neq \varnothing$ and there is a component M' of $\overline{M - (H_i \cup H_j)}$ that contains at least two of the handlebodies. Let S be the boundary surface between $H_i \cup H_j$ and M'. Then the proof of Lemma 3.2 can be modified to show that no essential simple closed curve in S bounds a disk, and thus S is an embedded incompressible surface. Therefore, the manifold is Haken as claimed.

3 Conditions and examples

For later use, we state a special case of Dehn's lemma and the loop theorem:

Lemma 3.1 *Let H be a handlebody and $\mathcal{T}$ a collection of essential curves in ∂H. If there is a singular meridian disk D of H such that D has n intersections with $\mathcal{T}$, then there exists an embedded meridian disk of H that intersects $\mathcal{T}$ at most n times.*

Let H be a handlebody and $\mathcal{T}$ be a set of disjoint essential simple closed curves in ∂H that satisfies the n disk-condition. A direct result of this lemma is that if α is a possibly singular loop in ∂H that intersects $\mathcal{T}$ less than n times and α contracts in H, then by Lemma 3.1 it follows that α is inessential in ∂H.

Lemma 3.2 *Let M be a manifold that satisfies the disk-condition. If $f\colon D \to M$ is a map of a disk D such that $f(\partial D) \subset \operatorname{int}(H_i)$ for some i, then f can be homotoped to g, keeping the boundary fixed, so that $g(D) \subset \operatorname{int}(H_i)$.*

Proof We can assume that $f(D)$ is transverse to X, where X is the union of the boundaries of the three handlebodies making up M and f is the disk map as in the lemma. Thus $\Gamma = f^{-1}(X)$ is a set of trivalent graphs and simple closed curves Γ_j, $1 \le j \le m$, in D. Note that $\partial D \cap \Gamma = \varnothing$. An *innermost* component of Γ is a component Γ_j such that there is a subdisk $D^* \subset D$ where $\partial D^* \subset \Gamma_j$ and $D^* \cap \Gamma = \Gamma_j$. An easy argument shows that if Γ is nonempty, then it must have at least one innermost component. The reason is that the closure of a component of the complement of Γ_j which does not contain ∂D is a subdisk D'. Clearly we can define a partial order on the components of Γ by $\Gamma_r < \Gamma_j$ if Γ_j has a complementary component which does not meet ∂D and contains Γ_r. A smallest component is then innermost.

If Γ_j is a simple loop, then $\Gamma_j = \partial D'$ and $f(D') \subset H_k$ for $k = 1, 2$ or 3. By the disk-condition, we know that $f(\partial D')$ must be nonessential in ∂H_k as $f(\partial D')$ doesn't intersect $\mathcal{T}$ and thus $f(D')$ is homotopic into ∂H_k. We can thus homotope f so that $f(D') \subset \partial H_k$ and then push $f(D')$ through to remove the component Γ_j altogether.

If Γ_j is a graph, then as it is innermost, there is a disk D^* with $\partial D^* \subset \Gamma_j$ and $\Gamma_j = \Gamma \cap D^*$. Thus any face F bounded by a subset of Γ_j in D^* is an (m,n)–gon, where F has m vertices in its boundary and is mapped by f to a handlebody H_k such that $\mathcal{T}$ satisfies the n disk-condition in H_k. We can put a PL metric on D^* by assuming that all the edges are geodesic arcs of unit length, that the internal angle at each vertex of an (m,n)–gon F is $\pi(1-2/n)$ and all the curvature of F is at a cone point in $\operatorname{int}(F)$. For example, if H_k satisfies the 6 disk-condition, the angle at each corner of an $(m,6)$–gon will be $\frac{2\pi}{3}$. Note that as each vertex of Γ_j in the interior of D^* is adjacent to three faces, each of these faces is mapped to a different handlebody. Assuming that M satisfies the $(6,6,6)$, $(4,6,12)$ or $(4,8,8)$ disk-conditions, then the total angle around each such interior vertex is 2π. If F is an (m,n)–gon, then $\chi(F) = 1$ and the exterior angle sum is $m(2\pi/n)$. If $\boldsymbol{K}(F)$ is the curvature of the cone point in $\operatorname{int}(F)$, then by the Gauss–Bonnet theorem,

$$\boldsymbol{K}(F) = 2\pi - m(2\pi/n) = 2\pi(1-m/n).$$

Thus if F is an (m,n)–gon and $m < n$, then $\boldsymbol{K}(F) > 0$, and if $m \ge n$, then $\boldsymbol{K}(F) \le 0$. Let $\boldsymbol{F}$ be the set of faces of D^* and $\boldsymbol{v}$ be the vertices in ∂D^*. For $v \in \boldsymbol{v}$, there are two faces $F_1, F_2 \in \boldsymbol{F}$ adjacent to v. Let F_i be an (m_i, n_i)–gon. Let the *jump* angle at v be $\theta_v = \pi - \sum_{i=1,2} \pi(1-2/n_i)$. By the disk-condition, $n_i = 4, 6, 8$ or 12, and it is

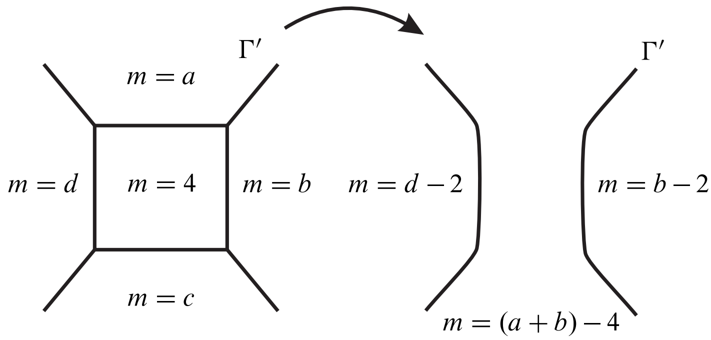


Figure 3: Removing a $(4, n)$–gon from Γ' by homotopy

not possible to have $n_1 = n_2 = 4$. Thus $\theta_v \le -\frac{\pi}{6}$. Then once again by Gauss–Bonnet we know that

$$\sum_{F \in \boldsymbol{F}} \boldsymbol{K}(F) = 2\pi - \sum_{v \in \boldsymbol{v}} \theta_v > 2\pi.$$

This implies that D^* must always have some (m, n)–gon faces such that $m < n$. For example, if the manifold satisfies the $(6, 6, 6)$ disk-condition, then D^* would have some $(2, 6)$–gons and/or some $(4, 6)$–gons, since m is even. If F is an (m, n)–gon of D^* such that $m < n$ and $f(F) \subset H_k$, then by the disk-condition and Lemma 3.1, we know that $f(\partial F)$ is not essential in ∂H_k. Thus we can homotope f so that $f(F)$ lies in ∂H_k. We can then homotope f so $f(F)$ is pushed off ∂H_k. This decreases the total number of faces of D^*, as shown in Figure 3. Thus in a finite number of steps, Γ_j will become a simple closed curve, and we can then homotope f to remove the component Γ_j entirely.

As Γ always contains an innermost component, we can continue this process until all of Γ has been removed, and thus $f(D) \subset \operatorname{int}(H_i)$. □

This lemma yields important corollaries about 3–manifolds that meet the disk-condition.

Corollary 3.3 *Let M be a 3–manifold that satisfies the disk-condition. Then, for any $1 \le i \le 3$, the induced map of $\pi_1(H_i)$ into $\pi_1(M)$ is injective.*

Remark 3.4 Note that $\pi_1(H_i)$ is the free group on g generators, where $g > 0$ is the genus of H_i. This corollary implies that if a 3–manifold satisfies the disk-condition, then its fundamental group is infinite.

Proof Let D be a disk and γ be a simple closed curve in H_i that represents a nontrivial element of $\pi_1(H_i)$. If the element is trivial in $\pi_1(M)$, then there is a map $f \colon D \to M$ such that $f(\partial D) = \gamma$. By Lemma 3.2, we can homotope f so that $f(D) \subset \operatorname{int}(H_i)$, giving us a contradiction. □

Corollary 3.5 *If M is a 3–manifold that satisfies the disk-condition, it is irreducible.*

Proof Let S be a 2–sphere and $f\colon S \to M$ be an embedding. Note that f is an embedding and all the moves in the proof of Lemma 3.2 can be performed as isotopies. Thus we can isotope f so that $f(S) \cap X = \varnothing$; that is, for some i, $f(S) \subset H_i$. Then, as handlebodies are irreducible, $f(S)$ must bound a 3–ball. □

3.1 Test for the n disk-condition in handlebodies

It is not necessary to check every meridian disk of a handlebody H to find out if a set of curves $\mathcal{T}$ in ∂H satisfies the n disk-condition. Let $\mathcal{D}$ be a set made up of a single representative from each isotopy class of meridian disk of H.

The first test is that $\mathcal{T}$ must separate ∂H into subsurfaces that can be 2–coloured. Therefore, all meridian disks must intersect $\mathcal{T}$ an even number of times. From this point on we will assume that $\mathcal{T}$ is separating in ∂H.

Put a Riemannian metric on ∂H. We will assume that the loops in $\mathcal{T}$ are length minimizing geodesics. Note that if $\mathcal{T}$ contains parallel curves, the neighbourhood of the corresponding length minimizing geodesic can be "flattened", so we can have parallel length minimizing geodesics. We will also assume the boundaries of the disks in $\mathcal{D}$ are length minimizing geodesics. Both of these can be done simultaneously. From M Freedman, J Hass and Scott [2], we know that this implies that the number of intersections between the boundary of a disk in $\mathcal{D}$ and $\mathcal{T}$ is minimal, as is the intersection between the boundaries of any two disks in $\mathcal{D}$, after possibly a small perturbation to make these intersections transverse. For any disk $D \in \mathcal{D}$, let $|D|$ be the number of intersections of ∂D with $\mathcal{T}$ and for any set of meridian disks $\boldsymbol{D} = \{D_i\} \subset \mathcal{D}$, let $|\boldsymbol{D}| = \sum_i |D_i|$. From this point on, unless otherwise stated, when discussing meridian disks, we will assume that the number of intersections between their boundaries is minimal.

Lemma 3.6 *Any two disks of $\mathcal{D}$ can be isotoped, leaving their boundaries fixed, so that any curves of intersection are properly embedded arcs.*

Proof This proof uses the standard innermost argument and the fact that handlebodies are irreducible to remove all the components of intersection between two disks that are simple closed curves. □

Definition 3.7 Let H be a genus-g handlebody. We shall call $\boldsymbol{D} \subset \mathcal{D}$ a *system of meridian disks* if all the disks are disjoint, nonparallel and cut H up into a set of 3–balls. If $\partial \boldsymbol{D}$ cuts ∂H up into $2g-2$ pairs of pants (thrice punctured 2–spheres), then it is a *basis* for H.

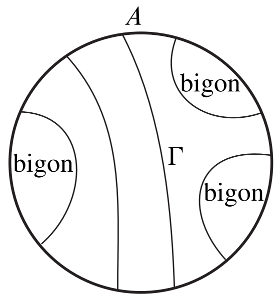


Figure 4: Meridian disk cut up by arcs of intersection

If H has genus g, then a minimal system of meridian disks for H consists of g disjoint meridian disks which cut H up into a single ball.

Definition 3.8 Let P be a punctured sphere and γ be a properly embedded arc in P. If both ends of γ are in one component of ∂P and the arc is not isotopic into ∂P, then it is called a *wave*.

Let H be a handlebody, $\mathcal{T}$ a set of essential disjoint simple closed curves in ∂H, $\boldsymbol{D}$ a system of meridian disks for H and $\{P_1, \dots, P_l\}$ the resulting set of punctured spheres produced when we cut ∂H along $\partial \boldsymbol{D}$. Also, let $\mathcal{T}_i = P_i \cap \mathcal{T}$. Thus $\mathcal{T}_i$ is a set of properly embedded disjoint arcs in P_i.

Definition 3.9 If each $\mathcal{T}_i$ contains no waves, then $\boldsymbol{D}$ is said to be a *waveless* system of meridian disks for H.

Definition 3.10 Let $\boldsymbol{D}$ be a waveless system of disks. If every wave in each P_i intersects $\mathcal{T}_i$ at least $\frac{1}{2}n$ times, then $\boldsymbol{D}$ is called an *n–waveless* system of meridian disks.

If $\boldsymbol{D}$ is an n–waveless basis, then each $\mathcal{T}_i$ has at least $\frac{1}{2}n$ parallel arcs running between each pair of boundaries in P_i.

Lemma 3.11 *Let H be a handlebody, $\mathcal{T} \subset \partial H$ a separating set of essential simple closed curves and $\boldsymbol{D}$ a basis for H. If $\boldsymbol{D}$ is an n–waveless basis, then $\mathcal{T}$ satisfies the n disk-condition in H.*

Proof From the definition of the n–waveless condition we know that $\mathcal{T}$ intersects each disk in $\boldsymbol{D}$ at least $\frac{3}{2}n$ times. If $C \in \mathcal{D}$ is a meridian disk not in $\boldsymbol{D}$, then $C \cap \boldsymbol{D} \neq \varnothing$. By Lemma 3.6, we can isotope C so that $C \cap \boldsymbol{D}$ is a set of disjoint properly embedded arcs. Therefore, if we cut C along $C \cap \boldsymbol{D}$ the faces produced must all be disks and contain at least two bigons, as shown in Figure 4. Therefore, the set $\{P_i \cap \partial C\}$ must contain

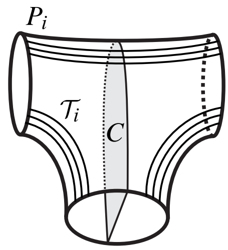


Figure 5: Bigon in a pair of pants

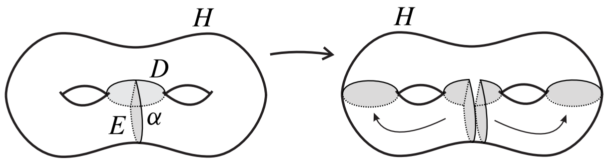


Figure 6: Boundary compressing a meridian disk

at least two waves, coming from bigons. As $\boldsymbol{D}$ satisfies the n–waveless condition, any wave must intersect $\mathcal{T}$ at least $\frac{1}{2}n$ times; see Figure 5. Therefore, ∂C must intersect $\mathcal{T}$ at least n times. □

If $\mathcal{T}$ intersects each disk in $\boldsymbol{D}$ exactly n times, then it must be an n–waveless basis. The reason is that the only pattern of arcs in a pair of pants, where there are the same number n of endpoints on each boundary curve, consists of $\frac{1}{2}n$ arcs joining each pair of boundary loops. This gives us the following corollary.

Corollary 3.12 *Let H be a handlebody, $\mathcal{T} \subset \partial H$ a separating set of simple closed curves and $\boldsymbol{D}$ a basis for H. If $\mathcal{T}$ intersects each disk in $\boldsymbol{D}$ exactly n times, then $\mathcal{T}$ satisfies the n disk-condition in H.*

This test for the n disk-condition is a significant restriction. However, it is an easy enough condition to verify when constructing examples.

Next we describe a specific type of surgery of meridian disks. Let D be a meridian disk of H and let E be an embedded disk in H such that $\partial E \subset D \cup \partial H$, $\partial E \cap \partial D$ is two points, a_1 and a_2 in ∂H, $\alpha = E \cap \partial H$ is an arc in ∂H which is not homotopic through ∂H into ∂D and $D \cap E$ is an arc properly embedded in D, as shown in Figure 6. If we then surger D along E, we produce two disks. As α is an arc which is not homotopic through ∂H into ∂D, both resulting disks are meridian disks isotopic to disks in $\mathcal{D}$. We shall call this surgery a *boundary compression* of a meridian disk.

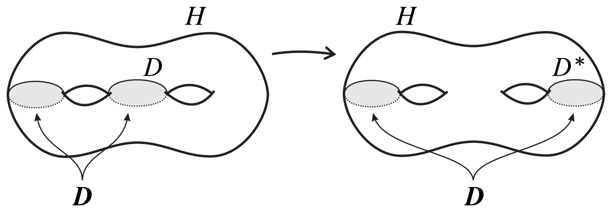


Figure 7: Disk-swap move

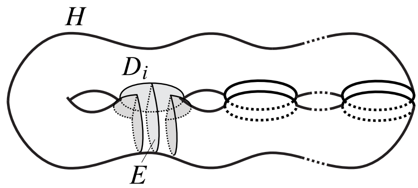


Figure 8: Boundary compressing a disk from a system of meridian disks

Let $\boldsymbol{D}$ be a system of disks for the handlebody H. Let $D^* \in \mathcal{D}$ be a meridian disk disjoint from $\boldsymbol{D}$ such that $(\boldsymbol{D} \setminus D) \cup D^*$ is a system of meridian disks for some $D \in \boldsymbol{D}$. Then if we remove D from $\boldsymbol{D}$ and replace it with D^*, this is called a *disk-swap move* on $\boldsymbol{D}$ as shown in Figure 7.

Lemma 3.13 *For a minimal system of meridian disks $\boldsymbol{D} = \{D_1, \dots, D_n\}$, if we perform a boundary compression on any D_i along a disk disjoint from $\boldsymbol{D} \setminus \{D_i\}$, then one of the resulting disks can be used for a disk-swap move on $\boldsymbol{D}$ removing D_i.*

Remark 3.14 Note that an essential wave in $\overline{\partial H - \boldsymbol{D}}$ defines a disk-swap move on $\boldsymbol{D}$.

Proof Let $\boldsymbol{D}^*$ be the set of all meridian disks disjoint from $\boldsymbol{D}$. Then if a disk $D_i \in \boldsymbol{D}$ is boundary compressed along a disk E disjoint from $\boldsymbol{D} - D_i$, one of the resulting disks will be isotopic to a disk in $\boldsymbol{D} \cup \boldsymbol{D}^*$. If we cut H along $\{D_1, \dots, D_{i-1}, D_{i+1}, \dots, D_n\}$ the result is a solid torus T. Then D_i is a meridian disk of T. Thus a boundary compression on D_i along E will produce two disks, one of which is a meridian disk of T and the other is boundary parallel, as shown in Figure 8. □

Let $\boldsymbol{D} \subset \mathcal{D}$ be a minimal system of meridian disks for the handlebody H. That is, $\boldsymbol{D}$ cuts H up into a single ball. Let $\boldsymbol{D}^* \subset \mathcal{D}$ be the set of disks disjoint from $\boldsymbol{D}$.

Lemma 3.15 *$\mathcal{T}$ satisfies the n disk-condition if and only if there is a minimal system of meridian disks $\boldsymbol{D}$ such that $|D| \geq n$ for all disks $D \in \boldsymbol{D} \cup \boldsymbol{D}^*$ and there are no disk-swap moves between $\boldsymbol{D}$ and $\boldsymbol{D}^*$ that reduce $|\boldsymbol{D}|$.*

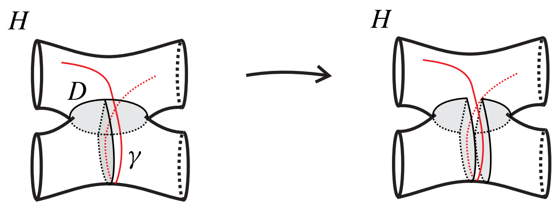


Figure 9: Boundary compression to remove a wave

Proof In the "only if" direction, $\mathcal{T}$ satisfying the n disk-condition in H implies that $|D| \geq n$ for any meridian disk. Given any initial $\boldsymbol{D} \cup \boldsymbol{D}^*$ such that there are disk-swap moves to reduce $|\boldsymbol{D}|$, we can construct a sequence of disk-swaps that reduce $|\boldsymbol{D}|$ with each move. If $\mathcal{T}$ satisfies the n disk-condition, then such a sequence must terminate, thus giving the required basis.

For the proof in the "if" direction, the first thing to note is that if there are no disk-swap moves to reduce $|\boldsymbol{D}|$, then every essential wave in $\overline{\partial H - \boldsymbol{D}}$ must intersect $\mathcal{T}$ at least $\frac{1}{2}n$ times. Let $D \in \mathcal{D}$ be a meridian disk such that $D \notin \boldsymbol{D} \cup \boldsymbol{D}^*$. Then $\Gamma = D \cap \boldsymbol{D} \neq \varnothing$. We are assuming that the intersection between the boundaries of disks is minimal. Thus by Lemma 3.6 we can assume that Γ is a set of pairwise disjoint properly embedded arcs in D, as shown in Figure 4. Thus all the faces of the meridian D, when D is cut along Γ, are disks. Also, there must be at least two bigons, D_1 and D_2 in this meridian disk. $D_i \cap \overline{\partial H - \boldsymbol{D}}$ are essential waves in $\overline{\partial H - \boldsymbol{D}}$ and thus intersect $\mathcal{T}$ at least $\frac{1}{2}n$ times. □

Next we want to use Lemma 3.15 to produce an algorithm to determine whether a boundary pattern satisfies the n disk-condition. Here by a boundary pattern, we mean a family of disjoint essential closed curves in the boundary of a handlebody.

Lemma 3.16 *Assume we are given a handlebody H and a set $\mathcal{T}$ of essential curves in ∂H. There is an algorithm to find, in finite time, a waveless minimal system of meridian disks.*

Proof Suppose we start with an arbitrary minimal system of meridian disks $\boldsymbol{D}$ for H. If $\mathcal{T}$ has a wave when H is cut along $\boldsymbol{D}$, then there is a subarc $\gamma \subset \mathcal{T}$ with both ends in some disk $D \in \boldsymbol{D}$ and $\operatorname{int}(\gamma) \cap \boldsymbol{D} = \varnothing$. Then D has a boundary compression disk E such that the arc $E \cap \partial H = \gamma$. Let D_1 and D_2 be the disks produced by compressing D along E. Then $\Sigma_i |D_i| \leq |D| - 2$, as shown in Figure 9. Thus when a disk-swap move is done swapping D for one of the D_i, we see that $|\boldsymbol{D}|$ will decrease by at least two. Note also that the number of waves does not go up. If there is another wave we can always do another boundary disk compression and a disk-swap move to reduce $|\boldsymbol{D}|$, thus this process must terminate in a finite number of moves. □

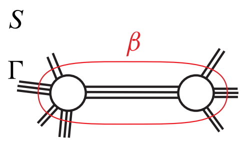


Figure 10: Boundary of meridian disk to add to $\boldsymbol{D}$

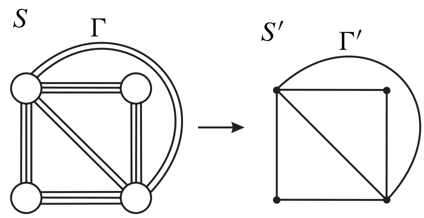


Figure 11: Γ and Γ'

Given that it is possible to find a waveless minimal system of meridian disks $\boldsymbol{D}$, to show that we can find a waveless basis, we proceed as follows. Suppose we have already found a waveless system of disks and want to add new waveless disks, until we get a basis. We can use our initial set of boundary curves of disks to cut ∂H to obtain a punctured sphere $S = \overline{\partial H - \boldsymbol{D}}$. Suppose that there is at least one pair of boundary curves of S such that all the arcs of $\Gamma = \mathcal{T} \cap S$ running between them are parallel. Then there is a simple closed curve β which is essential in S, is not boundary parallel and each curve in Γ intersects β at most once, as shown in Figure 10. Then we can add a disk with boundary β to enlarge our system of waveless disks.

To simplify this problem, collapse each boundary component of S to a vertex and identify parallel copies of edges of Γ. This produces a graph Γ' embedded in a 2–sphere S' such that Γ' is connected, no two edges are parallel and no edge has both ends at one vertex. This means that if we cut S' along Γ' all the resulting faces will be disks and will have degree at least 3.

Definition 3.17 A 2–*cycle* in a graph is a simple closed curve that is the union of two edges.

The problem of finding a waveless basis is now to show that we can always find two vertices of Γ' that are joined by exactly one edge. This means finding a vertex not contained in a 2–cycle. Let c be a 2–cycle in Γ', thus c cuts S' into two disks and as Γ' does not contain any parallel edges, the interior of both disks must contain at least one vertex of Γ'. We now want to show that there is a vertex of Γ' that is not part of a 2–cycle. Let c and c' be two 2–cycles in Γ'. If $c \cap c'$ is empty, a single vertex or edge, then the interior of one of the disks produced when we cut S' along c must be disjoint from c'. If $c \cap c'$ is two vertices, then we can construct a third 2–cycle c'' such that when we cut S' along c'', the interior of one of the disks produced is disjoint from both c and c'. (We obtain c'' by taking one edge from each of c, c'.) By induction on the number of 2–cycles in C, the set of all 2–cycles in Γ', it follows that there must be a 2–cycle $c \in C$ such that when S' is cut along c we get a disk D for which there are no 2–cycles intersecting $\operatorname{int}(D)$. As there are no parallel edges in Γ', we have $\Gamma' \cap \operatorname{int}(D) \neq \varnothing$. Therefore, Γ' has to have a vertex in $\operatorname{int}(D)$ that is not in a 2–cycle. This gives us the following lemma.

Lemma 3.18 *Assume we are given a handlebody H and a set $\mathcal{T}$ of essential curves in ∂H. There is an algorithm to find, in finite time, a waveless basis.*

Note that this means that once a minimal waveless system of meridian disks has been found, most of the work has been done and that to produce a waveless basis, suitable meridian disks are added to the system. This lemma is not expressly used in the rest of this paper, but waveless bases are used in Section 4 in a condition for manifolds to be atoroidal. Thus it is nice to know that given a 3–manifold that satisfies the disk-condition, we can always find a waveless basis for each of its handlebodies.

Lemma 3.19 *Let H be a handlebody and $\mathcal{T}$ a set of essential curves in ∂H. Then there is an algorithm to determine, in finite time, if $\mathcal{T}$ satisfies the n disk-condition.*

Proof Once again let $\boldsymbol{D}$ be a minimal system of disks and $N(\boldsymbol{D})$ be a regular neighbourhood of $\boldsymbol{D}$. Let $S = \overline{\partial H - N(\boldsymbol{D})}$ and $\Gamma = \mathcal{T} \cap S$. Then S is a $2g$–punctured sphere, where g is the genus of H. Also, Γ is a set of arcs properly embedded in S. By Lemma 3.16, we can assume that Γ does not contain any waves. Therefore, Γ cuts S up into polygonal disks of degree at least four. As above let $\boldsymbol{D}^* \subset \mathcal{D}$ be the set of meridian disks disjoint from $\boldsymbol{D}$. For any $D^* \in \boldsymbol{D}^*$, we have that $D^* \cap S = \alpha$ is a simple closed curve in $\operatorname{int}(S)$. Let $|\alpha|$ be the number of times that α intersects Γ. Note that $|\alpha| = |D^*|$. We have therefore reduced the question of looking for meridian disks disjoint from $\boldsymbol{D}$ to studying essential simple closed curves in S. For $D \in \boldsymbol{D}$, we have that $N(D) \cap S$ is two boundary curves, ∂D_1 and ∂D_2, of S. Then if γ is an essential simple closed curve in S that separates ∂D_1 from ∂D_2, the disk bounded by γ can be used for a disk-swap move on D. Let $N = \max\{|D| : D \in \boldsymbol{D}\}$ and L be the set of essential simple closed curves in S of length at most N. Thus as L is a finite set of curves and as each face of S is a polygon, we can list all the elements of L using normal curve theory, using the polygonal disk structure or a triangular subdivision. Therefore, to test whether $\boldsymbol{D}$ satisfies Lemma 3.15 we need to check that; all disks in $\boldsymbol{D}$ intersect $\mathcal{T}$ at least n times, all the curves in L have length at least n, and $|\gamma| \geq |D|$ for $\gamma \in L$ and $D \in \boldsymbol{D}$ such that γ separates the two curves $D \cap S$ in S. If a disk-swap move is found, then we perform the move and then test the new system. As $|D|$ decreases by at least two with each move, the algorithm will terminate in finite time, either when a suitable system is found, meaning $\mathcal{T}$ satisfies the n disk-condition or when a meridian disk is found that intersects $\mathcal{T}$ less than n times. □

Note that this algorithm can be continued until a system is found which has a "locally minimal" intersection. If $n = \min\{|D| : D \in \boldsymbol{D}\}$, then n is the supremum disk-condition satisfied by $\mathcal{T}$. For if there is a meridian disk that intersects $\mathcal{T}$ less than n times that is not in $\boldsymbol{D}$, then the algorithm would not have terminated. An equivalent statement is that $\boldsymbol{D}$ is an n–waveless system of disks. Clearly if there is an essential wave in $\overline{\partial H - \boldsymbol{D}}$ that intersects $\mathcal{T}$ less than $\frac{1}{2}n$ times, then there is a disk-swap move to reduce $|\boldsymbol{D}|$. In the other direction, if $\boldsymbol{D}$ is an n–waveless system and there is a meridian disk $D \in \mathcal{D}$ such that $|D| < n$, then clearly $D \cap \boldsymbol{D} \neq \varnothing$. Thus D gives a boundary compressing disk for some disk in $\boldsymbol{D}$ and thus a wave in $\overline{\partial H - \boldsymbol{D}}$, that intersects $\mathcal{T}$ at less than $\frac{1}{2}n$ points. Therefore, there is an alternative algorithm to test the disk-condition, giving the corollary:

Corollary 3.20 *If H is a handlebody and $\mathcal{T} \subset \partial H$ is a set of essential curves that meet the n disk-condition, then there is an algorithm to find an n–waveless minimal system of meridian disks.*

3.2 Examples

To construct manifolds that meet the disk-condition, we use Dehn surgery or branched covers to build a manifold M which contains a 2–complex that cuts M up into three injective handlebodies.

3.2.1 Dehn filling examples The first class of examples of manifolds that meet the disk-condition are constructed by performing Dehn surgery along suitable knots in $\mathbb{S}^3$. Let $K \subset \mathbb{S}^3$ be the $(3,3,3)$–pretzel knot and F the free spanning surface shown in Figure 12. For $A \subset \mathbb{S}^3$, let $N(A)$ be a regular neighbourhood of A. Let $H_3 = N(K)$ and $H_1 = \overline{N(F) - H_3}$, as shown in Figure 13. Then H_1 is a genus-2 handlebody, and $\mathcal{T} = \partial(H_1 \cap H_3)$ is two copies of K. Furthermore, H_1 is homeomorphic to an I–bundle over F and $\mathcal{T}$ to the boundary curves of the vertical boundary of the I–bundle structure. Given the arcs $\beta_1, \beta_2, \beta_3$ in Figure 12, $\bigcup_i (\beta_i \times I)$ is a basis for H_1. Each wave in the pairs of pants produced when ∂H_1 is cut along the basis intersects $\mathcal{T}$ at least twice. Therefore, the basis is 4–waveless, and by Lemma 3.11, $\mathcal{T}$ satisfies the 4 disk-condition in H_1. Also, $H_2 = \overline{\mathbb{S}^3 - (H_1 \cup H_3)}$ is a genus-2

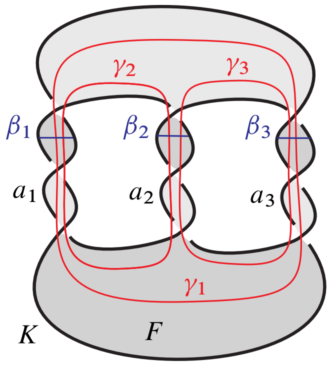


Figure 12: (3, 3, 3)–pretzel knot

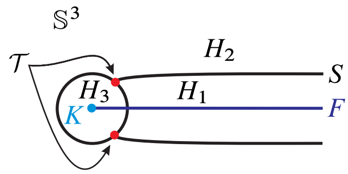


Figure 13: Handlebodies in Dehn filling construction

handlebody, and the curves $\gamma_1, \gamma_2, \gamma_3$ in Figure 12 bound meridian disks of a basis $\boldsymbol{D}$ for H_2. As $\mathcal{T}$ is two copies of K each wave in the two pairs of pants, produced by cutting ∂H_2 along the γ_i, intersects $\mathcal{T}$ six times. Thus $\boldsymbol{D}$ is a 12–waveless basis for H_2, and by Lemma 3.11, $\mathcal{T}$ satisfies the 12 disk-condition in H_2. Therefore, if a Dehn surgery along K is performed such that the meridian disk of the solid torus glued back in intersects $\mathcal{T}$ at least six times, a manifold that satisfies the $(4, 6, 12)$ disk-condition is produced. U Oertel showed in [10] that all but finitely many Dehn surgeries on such pretzel knots produce non-Haken 3–manifolds.

This construction can be generalised to any knot $K \subset \mathbb{S}^3$, that has a free spanning surface F, such that K satisfies the 6 disk-condition in $\overline{\mathbb{S}^3 - F}$. Then any Dehn surgery of type (p, q) with $|p| \geq 6$ will produce a manifold meeting the disk-condition.

3.2.2 Branched cover examples The next method for constructing manifolds which meet the disk-condition is taking cyclic branched covers over knots in $\mathbb{S}^3$. We look at two conditions on knots that are sufficient for the resulting manifolds to meet the disk-condition.

Let B_i, for $i = 1, 2$ or 3, be 3–balls and $\gamma_i = \{\gamma_i^1, \dots, \gamma_i^k\}$, for $k \geq 2$, be a set of properly unknotted pairwise disjoint embedded arcs in B_i. Unknotted means that there

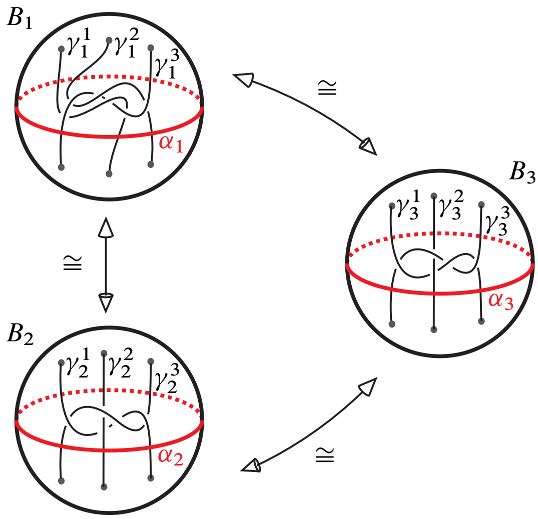


Figure 14: Bubble construction

is a set of pairwise disjoint embedded disks, $D_i = \{D_i^1, \dots, D_i^k\}$, such that

$$\gamma_i^j \subset \partial D_i^j \quad \text{and} \quad \overline{\partial D_i^j - \gamma_i^j} = D_i^j \cap \partial B.$$

Therefore, if we take the p–fold cyclic branched cover of B_i, with γ_i as the branch set, then the result will be a genus-$(p-1)(k-1)$ handlebody H_i. Let $r_i\colon H_i \to B_i$ be the branched covering map and $\alpha_i \subset \partial B_i$ be a simple closed curve disjoint from γ_i such that $\mathcal{T}_i = r^{-1}(\alpha_i)$ satisfies the n_i disk-condition in H_i. Note that α_i can be thought of as cutting ∂D_i up into two hemispheres.

Now glue the three balls by homeomorphisms between their hemispheres, as shown in Figure 14, so that the resulting manifold is $\mathbb{S}^3$ and the endpoints of the γ_i match up. Thus $K = \bigcup \gamma_i$ is a link and $C = \bigcup \partial B_i$ is a 2–complex of three disks glued along a triple curve α, which is the image of the α_i. Let M be the p–fold cyclic branched cover of $\mathbb{S}^3$ with K as the branch set. Let $r\colon M \to \mathbb{S}^3$ be the branched covering map. Then $X = r^{-1}(C)$ is a 2–complex that cuts M up into handlebodies and $\mathcal{T} = r^{-1}(\alpha)$ is a set of triple curves that satisfies the n_i disk-condition in H_i. Thus if $\sum 1/n_i \le \frac{1}{2}$, then M satisfies the disk-condition.

If $k = 2$ or 3 and the intersection of α_i with D_i is minimal under isotopy in $\partial B_i - \gamma_i$, then a sufficient condition for the lift of γ_i to the p–fold cyclic branched cover of B_i to meet the n disk-condition is that any essential wave in $\overline{\partial B_i - D_i}$ intersects $\gamma \cap \overline{\partial B_i - D_i}$ at least $\frac{1}{2}n$ times. Note that this is a slight variation of Lemma 3.11 and the proof is essentially the same. Given the 2–complex shown in Figure 15, it can be seen that any p–fold cyclic branched cover over an (a_1, a_2, a_3)–pretzel knot in $\mathbb{S}^3$ such that $|a_i| \ge 2$ will produce a manifold that satisfies the disk-condition.

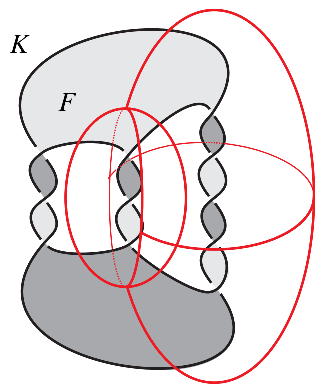


Figure 15: $(3,3,5)$–pretzel knot

Let M be a manifold that satisfies the disk-condition and can be constructed from the gluing of three genus-2 injective handlebodies. Then a simple Euler characteristic argument shows that all the faces of the 2–complex X must either be once punctured tori or twice punctured disks. If all the faces are once punctured tori, then the set of triple curves, $\mathcal{T}$, is a single curve. Thus a free involution of $\mathcal{T}$ can be canonically extended, up to isotopy, to an involution on each of the faces of X with three fixed points. Using a waveless basis for each handlebody, the involution on X can be extended to the whole of M. This means that any such manifold has a $\mathbb{Z}_2$ symmetry and is the 2–fold cyclic branched cover of $\mathbb{S}^3$ over some knot or link. In fact, the quotient of M by the involution is three balls glued together along hemispheres as in Figure 14. If all the faces of X are pairs of pants, then there is no corresponding involution of M.

The second construction involves the 3–fold cyclic branched cover of a knot that meets essentially the same condition as in the Dehn filling construction, so that the lift of the Seifert surface gives the 2–complex X. Let K be a knot in $\mathbb{S}^3$ and F be a free Seifert surface for K. This means that $\overline{\mathbb{S}^3 - F}$ is a handlebody. We construct the 3–fold cyclic branched cover over the knot K in $\mathbb{S}^3$ given by D Rolfsen in [11]. Let $N(K)$ be a regular neighbourhood of K, $\alpha \subset \partial N(K)$ the meridian curve of $N(K)$ and $N = \overline{\mathbb{S}^3 - N(K)}$. Let $\widetilde{N}$ be the 3–fold cyclic cover of N and $p\colon \widetilde{N} \to N$ the covering projection. That is, let $G \subset \pi_1(N)$ be the kernel of the homomorphism mapping $\pi_1(N)$ onto $\mathbb{Z}_3$, where the meridian of $N(K)$ is sent to a generator of $\mathbb{Z}_3$. Then $\widetilde{N}$ is the cover corresponding to G. So $\widetilde{N}$ has a single torus boundary and $\widetilde{\alpha} = p^{-1}(\alpha)$ is a single curve that covers α three times. Therefore, $\widetilde{F} = p^{-1}(F)$ is a set of three properly embedded spanning surfaces in $\widetilde{N}$. As F is free, $\overline{\widetilde{N} - \widetilde{F}}$ is three handlebodies. Let M be the 3–fold cyclic branched cover of $\mathbb{S}^3$ with K as the branch set. Then M can be constructed by gluing a solid torus T to $\partial\widetilde{N}$ so that its meridian matches $\widetilde{\alpha}$. Next extend each surface in $\widetilde{F}$ along an annulus to the spine $\mathcal{T}$ of T to produce a 2–complex X. Thus X is a 2–complex that cuts M into three handlebodies. Thus for M to meet the disk-condition it is sufficient for K to meet the 6 disk-condition in $\overline{\mathbb{S}^3 - F}$. An obvious example of such a knot is the $(3,3,3)$ pretzel knot in Figure 12.

The 3–fold cyclic branched cover of the $(3,3,5)$ pretzel knot K gives an example of a manifold with two distinct splitting 2–complexes that meet the disk-condition. Let M be the 3–fold cyclic branched cover of $\mathbb{S}^3$ with K as the branch set. Let X be the 2–complex produced by lifting the Seifert surface F to M and let X' be the 2–complex produced by lifting the "bubble" 2–complex shown in Figure 15. X and X' are distinct 2–complexes meeting the disk-condition. That is there is no homeomorphism of M that sends X to X', for if there was, M would have a $\mathbb{Z}_3 \times \mathbb{Z}_3$ symmetry and thus K would have a $\mathbb{Z}_3$ symmetry, which is clearly not the case. Note that if each twisted band in K has the same number of crossings, for example the $(3,3,3)$ pretzel knot, then the 3–fold cyclic branched cover does have a $\mathbb{Z}_3 \times \mathbb{Z}_3$ symmetry.

4 Characteristic variety

In this section we prove the torus theorem and construct the characteristic variety in 3–manifolds that meet the disk-condition. The first step is to look at how, in the component handlebodies, properly embedded essential annuli disjoint from the triple curves intersect and how meridian disks that intersect $\mathcal{T}$ exactly n_i times intersect. This allows us to build a picture of the characteristic variety in each of the handlebodies, which we then use to construct the characteristic variety of the manifold.

4.1 Handlebodies, embedded annuli and meridian disks

Throughout this section, let H be a handlebody and $\mathcal{T}$ be a set of disjoint essential simple closed curves in ∂H that meet the n disk-condition in H. We will assume that all intersections between surfaces are transverse. Before we look at the components of the characteristic variety in each handlebody, we need to look at some properties of embedded essential annuli that are disjoint from $\mathcal{T}$.

4.2 Essential annuli

In this section we investigate intersections between embedded essential proper annuli.

Definition 4.1 An intersection curve between two annuli is said to be *vertical* if it is a properly embedded arc which is not boundary parallel in either annulus. The intersection curve is *horizontal* if it is an essential simple closed curve in both annuli.

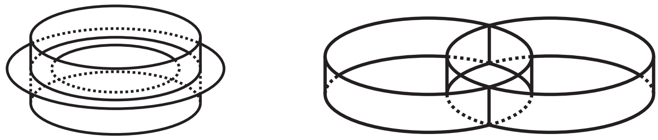


Figure 16: Intersecting embedded annuli: horizontal (left) and vertical (right)

If there is a proper isotopy in $H - \mathcal{T}$ of two annuli which removes their intersections, then the annuli will be said to have *trivial intersection* and if the intersection cannot be removed, the annuli have *nontrivial intersection*. This means that if two embedded annuli have nontrivial intersection they cannot be isotopically parallel. The disk-condition restricts how properly embedded annuli can intersect.

Lemma 4.2 *Let A_1 and A_2 be two essential properly embedded annuli in $H - \mathcal{T}$. Then there is a proper isotopy of them in $H - \mathcal{T}$ such that all their intersections are either vertical or horizontal.*

Remark 4.3 This means that nontrivial intersections between embedded annuli must either be all horizontal or all vertical.

Proof This uses standard innermost curve arguments and the following observations. Let A_1 and A_2 be essential properly embedded annuli in $H - \mathcal{T}$ and let $\Gamma = A_1 \cap A_2$. First note that as the A_i are embedded they cannot have both horizontal and vertical intersections. As H is irreducible there is an isotopy of A_1 to remove components of Γ that are simple closed curves and inessential in both A_i. Also, by irreducibility of H and the disk-condition, there is an isotopy of A_1 to remove components of Γ which are properly embedded arcs and boundary parallel in both A_i. Let γ be a component of Γ which is a simple closed curve and is essential in A_1 and not essential in A_2. Then the disk in A_2 bounded by γ implies that A_1 is not π_1–injective, which is a contradiction. Now let γ be a component of Γ which is a properly embedded arc which has both ends in the same boundary curve of A_1 and runs between the boundary curves of A_2. Then the disk bounded by γ in A_1 is a boundary compression disk for A_2 and the disk produced by compressing A_2 is disjoint from $\mathcal{T}$, thus implying that A_2 is boundary parallel in $H - \mathcal{T}$. □

Lemma 4.4 *Let H be a handlebody and $\mathcal{T}$ a set of curves in ∂H that meet the n disk-condition. Assume a properly embedded essential annulus in $H - \mathcal{T}$ intersects two other properly embedded essential annuli in $H - \mathcal{T}$, one vertically and the other horizontally. Then if there is a nontrivial horizontal intersection, the vertical intersections can be removed by an isotopy.*

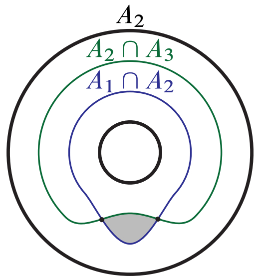


Figure 17: Curves of intersection in A_2

Remark 4.5 This indicates there are three types of essential embedded annuli in $H - \mathcal{T}$: those that have nontrivial horizontal intersections with other annuli, those that have nontrivial vertical intersections with other annuli and those that have no nontrivial intersections with other annuli. Later in this section, we will see that these types of annuli correspond to the flavours of characteristic variety in $H - \mathcal{T}$.

We could follow a least-area argument using a suitable Riemannian metric on the handlebody but use instead a more elementary direct cut-and-paste approach.

Proof Let A_1, A_2 be two properly embedded essential annuli in $H - \mathcal{T}$ that have nontrivial horizontal intersection. Let A_3 be a third embedded essential annulus in $H - \mathcal{T}$ that intersects A_1 vertically. If the vertical intersection between A_1 and A_3 is nonempty, then $(A_1 \cap A_2) \cap A_3 \neq \varnothing$ and thus the intersection between A_2 and A_3 is nonempty. By Lemma 4.2, we can isotope A_3 so that its intersection with A_2 is either vertical or horizontal and its intersection with A_1 is vertical. We will assume that the vertical intersection between A_1 and A_3 is still nonempty. If the intersection between A_2 and A_3 is horizontal, then ∂A_3 is disjoint from ∂A_2, as both $A_2 \cap A_1$ and $A_2 \cap A_3$ are essential simple closed curves in A_2. There is an innermost bigon on A_2 bounded by one arc from each of $A_2 \cap A_1$ and $A_2 \cap A_3$ with common endpoints; see Figure 17. This is clear because each arc of $A_1 \cap A_3$ has to have at least one corresponding vertex of $(A_2 \cap A_1) \cap (A_2 \cap A_3)$. If we assume there is a single vertical arc of $A_1 \cap A_3$ which contains both vertices of the bigon, then by the irreducibility of H there is an isotopy of A_2 over a ball in H bounded by the bigon and disks in A_1 and A_3 to remove the bigon. It is then straightforward to see that A_2 can be isotoped so for any bigon bounded by an arc of $A_2 \cap A_1$ and $A_2 \cap A_3$ there are two vertical arcs of intersection of $A_1 \cap A_3$ which contain the two vertices of this bigon; see Figure 18. We can then isotope A_3 across this bigon to convert these two vertical arcs into two boundary parallel arcs of $A_1 \cap A_3$ which can be removed by a further isotopy. In this way, eventually all the vertical arcs of $A_1 \cap A_3$ can be removed. Thus we can assume that A_3 intersects both A_1 and A_2 vertically.

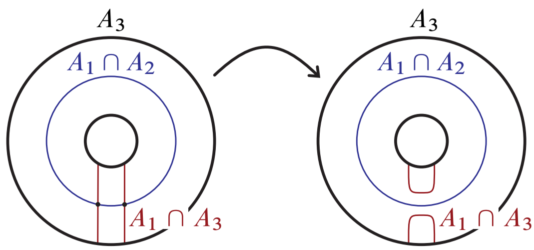


Figure 18: Curves of intersection in A_3

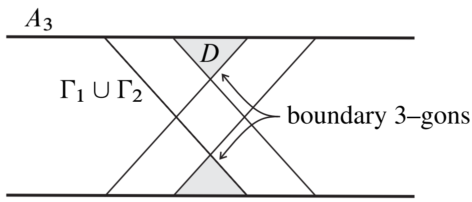


Figure 19: Component of the pullback graph $\Gamma_1 \cup \Gamma_2$

Let $\Gamma_i = A_3 \cap A_i$ for $i \neq 3$. Then Γ_i is a set of properly embedded pairwise disjoint spanning arcs in A_3, where each arc from Γ_1 intersects at least one arc from Γ_2. The faces produced when A_3 is cut up along $\Gamma_1 \cup \Gamma_2$ are all disks. As each connected component of $\Gamma_1 \cup \Gamma_2$ contains at least two arcs, each component will have a boundary 3–gon, D, as shown in Figure 19, such that subarcs of ∂A_3, Γ_1 and Γ_2 make up its three edges. Then the disk D gives an isotopy of A_1 that converts the corresponding essential closed curve of $A_1 \cap A_2$ into a boundary parallel arc. Thus there is a further isotopy to remove the intersection altogether. This process can be repeated to remove all the intersections of $A_1 \cap A_2$, giving a contradiction. □

Therefore, if a proper essential annulus in $H - \mathcal{T}$ has a nontrivial horizontal/vertical intersection with one annulus, then we can arrange that all its nontrivial intersections with all other essential annuli must be horizontal/vertical.

4.3 Meridian disks

Next we want to examine intersecting meridian disks. In particular, if $\mathcal{T}$ satisfies the n disk-condition in H, then there may be meridian disks that intersect $\mathcal{T}$ exactly n times. These disks are important when we are considering the disk flavour of characteristic variety.

Definition 4.6 If F is an n–gon and γ is a properly embedded arc in F such that if F is cut along γ, the result is two disks that have $\frac{1}{2}n$ intersections with $\mathcal{T}$, then γ is said to be a *bisecting* arc of F.

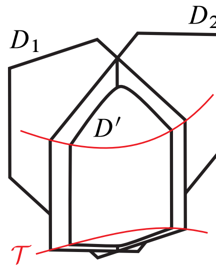


Figure 20: Two trivially intersecting 6–gons

Lemma 4.7 *Let H be a handlebody and $\mathcal{T}$ a set of curves in ∂H that satisfies the n disk-condition. If D_1 and D_2 are meridian disks that have n intersections with $\mathcal{T}$, then there is an isotopy of the disks such that $\Gamma = D_1 \cap D_2$ is a set of properly embedded disjoint bisecting arcs in both D_i or the intersection Γ can be removed.*

Proof This proof uses the usual innermost curve arguments and the following observations, to construct an isotopy to remove arcs of Γ that are not bisecting in both disks. By Lemma 3.6, we can assume that all components of Γ are properly embedded arcs. If such an arc is not bisecting in D_1, it is easy to see there is an arc γ of Γ which bounds an innermost subdisk D in D_1 which intersects $\mathcal{T}$ less than $\frac{1}{2}n$ times. Then one of the disks D' produced by surgering D_2 along D must intersect $\mathcal{T}$ in less than n points, as shown in Figure 20, and thus is boundary parallel in H. So there is an isotopy of D_1 to remove γ. □

Lemma 4.8 *Let H be a handlebody, $\mathcal{T}$ a set of curves in ∂H that meet the n disk-condition and D_1, D_2 and D_3 a set of meridian disks that all have n intersections with $\mathcal{T}$. Then there is an isotopy of the D_i such that $\bigcap D_i = \varnothing$.*

Proof By the previous lemma, we can isotope D_1 and D_2 so that their intersection is a set of parallel arcs in both disks. Assume that D_1 and D_2 have been isotoped so that their intersection has the least possible number of components and that $D_1 \cap D_2 \neq \varnothing$. Let A be a regular neighbourhood of $D_1 \cup D_2$ and B be the frontier of A in H. As no annulus component of B intersects $\mathcal{T}$, B consists of meridian disks that intersect $\mathcal{T}$ exactly n times and essential annuli whose boundary compressing disks intersect $\mathcal{T}$ at least $\frac{1}{2}n$ times.

Let D be a disk and $f\colon D \to H$ be an embedding such that $f(D) = D_3$. Then f can be isotoped so that $\Gamma = f^{-1}(B)$ is a set of properly embedded pairwise disjoint curves. As usual there is an isotopy of f to remove components of Γ that are simple closed curves. If D_3 intersects an annulus of B, then from above, either the intersections are parallel arcs or there is an isotopy of f to remove them. Similarly from Lemma 4.7 if D_3 intersects a disk of B, then either the intersections are bisecting parallel arcs or there is an isotopy of f to remove them. Therefore, there is an isotopy of f such that Γ is a set of parallel bisecting arcs. Thus $f^{-1}(A)$ is a set of 4–gons. Let D' be a 4–gon in $f^{-1}(A)$. Then using the same arguments as in the final step of the proof of Lemma 4.4, there is an isotopy of f such that $D' \cap f^{-1}(D_2 \cup D_1)$ is a set of parallel bisecting arcs. Moreover $f(D') \cap D_1 \cap D_2 = \varnothing$. This process can be repeated for each component of $f^{-1}(A)$ and thus $D_1 \cap D_2 \cap D_3 = \varnothing$. □

4.4 Flavours of characteristic variety in the handlebodies

4.4.1 I–bundle regions Let H be a handlebody and $\mathcal{T}$ a set of essential simple closed curves in ∂H, that meet the n disk-condition in H. Let N be a maximal, up to isotopy, I–bundle in H disjoint from $\mathcal{T}$, with its horizontal boundaries embedded in $\partial H - \mathcal{T}$, each component of N has nontrivial fundamental group and the induced map on the fundamental group is injective. Thus N is an I–bundle with a base space which is an embedded surface in H. Let S be a component of this embedded surface. If S is orientable, then the corresponding component of N has a product structure and its horizontal surface consists of two copies of S embedded in $\partial H - \mathcal{T}$. Alternatively, if S is nonorientable, then the corresponding component of N has a horizontal boundary which is a double cover of S embedded in $\partial H - \mathcal{T}$. In both cases the vertical boundary is a set of essential properly embedded annuli. From this point on these surfaces will be called frontier annuli. Also note that none of the base surfaces can be disks. This means that N is a set of embedded handlebodies in H with genus ≥ 1. N is not unique, for if H contains two embedded annuli that intersect horizontally, in a nontrivial way, then N can contain the regular neighbourhood of one or the other annulus but not both.

Definition 4.9 Let the *I–bundle region*, N_I, be the set of all components N_i from N which have base spaces that are not annuli or Möbius bands.

Later the I–bundle region is shown to be unique up to isotopy.

Lemma 4.10 *If A is a properly embedded essential annulus in $H - \mathcal{T}$ that has a nontrivial vertical intersection with another properly embedded essential annulus, then it is isotopic into N_I.*

Proof Let the map $f_i\colon A \to H - \mathcal{T}$, for $i = 1$ or 2, be an essential proper embedding of an annulus A such that $f_1(A) = A_1$ and $f_2(A) = A_2$ have nontrivial vertical intersections. Let B be the set of frontier annuli of N_I. If $A_1 \cap N_I \neq \varnothing$, then by Lemmas 4.2 and 4.4 we know that there is an isotopy of f_1 such that the intersection between A_1 and the annuli in B is vertical. Thus the pullback $\Gamma_1 = f_1^{-1}(B)$ is a set of properly embedded nonboundary parallel arcs in A and, as B is separating in H, there must be an even number of them. Thus Γ_1 cuts A up into quadrilaterals and every alternate one is mapped by f_1 into $\overline{(H - N_I)}$. Let $A' \subset A$ be a quadrilateral such that $f(A') \subset \overline{(H - N_I)}$. Also, let $N(f_1(A'))$ be the regular neighbourhood of $f_1(A')$ in $\overline{(H - N_I)}$ disjoint from $\mathcal{T}$. Note that $N(f_1(A'))$ can be fibred as an I–bundle over a quadrilateral. Then there must be an isotopy of f_1 to remove the curves $\Gamma_1 \cap A'$ otherwise $N(f_1(A')) \cup N_I$ would be larger than N_I, contradicting maximality. We can repeat this process until $\Gamma_1 = \varnothing$, thus $A_1 \cap B = \varnothing$. This process can be repeated for A_2 so that it is disjoint from B. If $A_1 \cap A_2$ is disjoint from N_I and the annuli have been isotoped so that their intersection is a minimal set of essential arcs, then $N(A_1 \cup A_2)$ can be fibred as an I–bundle and added to N_I, contradicting maximality. Thus $A_1 \cup A_2 \subset N_I$. □

Note that in distinction to the above lemma, if an annulus A meets another annulus horizontally, it may not be possible to isotope A into N_I.

Now let $\breve{H}$ be a regular finite-sheeted cover of H and $\breve{\mathcal{T}}$ be the lift of $\mathcal{T}$. Thus $\breve{H}$ also is a handlebody with $\breve{\mathcal{T}}$ satisfying the n disk-condition. Now let $N_I \subset \breve{H}$ be the I–bundle region, as described above. Also, let G be the group of covering translations of $\breve{H}$, so $\breve{H}/G = H$. Let N_i, for $1 \le i \le n$, be the connected subhandlebodies of N_I and S_i be the base-surface corresponding to N_i.

Lemma 4.11 *If N_i is a component of N_I, then $g(N_i)$ is isotopic to a component of N_I for any $g \in G$.*

Proof Let $\boldsymbol{A}$ be the set of frontier annuli of $g(N_i)$ and $\boldsymbol{B}$ the set of frontier annuli of N_I. If $g(N_i)$ and N_I have a nontrivial intersection, then by Lemma 4.2 there is an isotopy of g such that if any annuli in $\boldsymbol{A}$ and any annuli in $\boldsymbol{B}$ intersect, then the intersection curves are all either vertical or horizontal. Now isotope g to remove all trivial intersections between annuli in $\boldsymbol{A}$ and $\boldsymbol{B}$.

Let $B \in \boldsymbol{B}$ be an annulus such that it intersects at least one annulus in $\boldsymbol{A}$ horizontally. By Lemma 4.4, it can only intersect the other annuli in $\boldsymbol{A}$ horizontally. Thus $B \cap g(N_i)$ is a set of annuli properly embedded in $g(N_i)$. Let $B' \subset B$ be one such annulus.

Isotope B' so that it is transverse to the I–bundle structure. As intersections of B with annuli in $\boldsymbol{A}$ are minimal, B' either projects one-to-one onto the base space or double covers it. This depends on whether the two boundary curves of B' are in different annuli in A or in the same annulus, respectively. Therefore, the base space of $g(N_i)$ and thus N_i is either an annulus or a Möbius band, giving us a contradiction. This means that all horizontal intersections between annuli in $\boldsymbol{A}$ and $\boldsymbol{B}$ can be removed.

Therefore, all intersections between annuli in $\boldsymbol{A}$ and $\boldsymbol{B}$ that are nontrivial are vertical. But by Lemma 4.10 we can isotope all such annuli in $\boldsymbol{A}$ into N_I. Therefore, there is an isotopy of g such that $g(N_i) \cap N_I \neq \varnothing$ and $\boldsymbol{A} \cap \boldsymbol{B} = \varnothing$. Thus we know that we can isotope g so that $g(N_i)$ lies inside N_I, otherwise $g(N_i) \cup N_I$ would be a larger I–bundle than N_I, contradicting maximality.

As $g(N_i)$ is connected we know that it lies in a single component, N_k, of N_I. If $g(N_i)$ is not isotopic to N_k, then $g^{-1}(N_k - g(N_i)) \cup N_I$ is a larger I–bundle region, contradicting maximality. □

From the previous lemma we get the following corollary.

Corollary 4.12 *The regions N_I and $g(N_I)$ are isotopic for any $g \in G$.*

This corollary can be used to show that N_I can be isotoped so that it is preserved by G. Put a Riemannian metric on H, lift it to $\breve{H}$ and then isotope N_I so that the frontier annuli of the N_I are least area. Let $g \in G$ and A be a frontier annulus of N_I. By the arguments used by Freedman, Hass and Scott in [3], $g(A)$ is either a frontier annulus of N_I or disjoint from all frontier annuli of N_I. Let N_I' and N_I'' be components of N_I such that $g(N_I')$ is isotopic to N_I''. If $N_I' \neq N_I''$, then replace N_I'' by $g(N_I')$. Now assume that $N_I' = N_I''$. We need to look at what happens to the frontier annuli under g. Let A and A' be frontier annuli of N_I' such that $g(A)$ is isotopic to A'. If $A \neq A'$, then replace A' by $g(A)$. Now assume that $A = A'$ and $g(A) \neq A$. As each element of G is a periodic homeomorphism, $g(N_I') \not\subset \operatorname{int}(N_I')$. Then by this observation and maximality of N_I, <u>either $g(N_I') \cap N_I'$ is empty</u> or it is isotopic to N_I'. Another way of saying this is that $\overline{g(N_I') - N_I'}$ and $\overline{N_I' - g(N_I')}$ are sets of thickened annuli. We can then assume that <u>$g(A)$ is disjoint from</u> N_I'. Let U_i, for $i \in \mathbb{N}$, be the thickened annulus component of $\overline{g^i(N_I') - g^{i-1}(N_I')}$, where g^0 is the identity. As $\breve{H}$ is a finite-sheeted normal cover, there is some $m \in \mathbb{N}$ such that g^m is the identity. Therefore, $U_1 \cup \cdots \cup U_m$ is an annulus bundle over $\mathbb{S}^1$ properly embedded in $\breve{H}$, which cannot happen, thus $g(A) = A$. This gives us the following corollary.

Corollary 4.13 *There is an isotopy of $N_I \subset \breve{H}$ such that it is preserved by all the covering transformations.*

Lemma 4.10 implies that if H contains two embedded annuli that have nontrivial vertical intersection, then N_I is not empty. Note this is a sufficient condition not a necessary one. For example, if N_I is an I–bundle over a twice punctured disk, then any two embedded annuli contained in N_I are parallel to frontier annuli and thus their intersections can be removed isotopically.

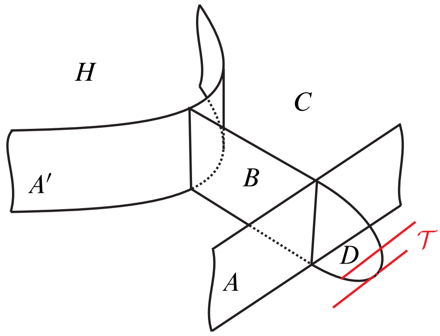


Figure 21: Extending boundary compression disk through an I–bundle component

Lemma 4.14 *N_I is unique up to ambient isotopy of H.*

We will not give the proof for this lemma as the method is the same as Lemma 4.11, the idea being that if we assume that we have two I–bundle regions N_I and N_I' that are not isotopic, then we get a contradiction to their maximality. Another property of N_I we need later is this lemma:

Lemma 4.15 *Let H be a handlebody, $\mathcal{T}$ a set of pairwise disjoint essential simple closed curves in ∂H that meet the n disk-condition and N_I the I–bundle region in H. Then if A is a frontier annulus of N_I and D is a boundary compression disk for A, then $|D| \geq \frac{1}{2}n$.*

Proof Assume that N_I has a frontier annulus A with a boundary compressing disk D such that $|D| < \frac{1}{2}n$. Also, let N_i be the component of N_I that has A as a frontier annulus. If we compress A along D to get a disk E, then $|E| < n$. Therefore, A must be boundary parallel, meaning there is a proper isotopy of A into ∂H. Note that this does not mean there is a proper isotopy of A into $\partial H - \mathcal{T}$. First assume that N_i has more than one frontier annulus. Let A' be another frontier annulus of N_i. As N_i is an I–bundle there is a 4–gon B, properly embedded in N_i, such that $B \cap A = D \cap A$ and $A' \cap B$ is a properly embedded arc in A' that is not boundary parallel, as shown in Figure 21, for suitable choice of D. Let $D' = D \cup B$. Then $|D'| < \frac{1}{2}n$, and if we compress A' along D', we get a disk E' with $|E'| < n$. Therefore, A' is boundary parallel through a region containing A. So A and A' must be parallel and N_i is the regular neighbourhood of a properly embedded annulus and thus can not be contained in N_I. If N_i has a single frontier annulus A, then similarly by the I–bundle structure, there is a properly embedded 4–gon $B \subset N_I$ such that it is not boundary parallel and $A \cap B$ is two arcs that are not parallel into ∂A. Then there are two boundary compression disks for A that can be glued to B along $A \cap B$. This produces a meridian disk that intersects $\mathcal{T}$ less than n times, contradicting the disk-condition. □

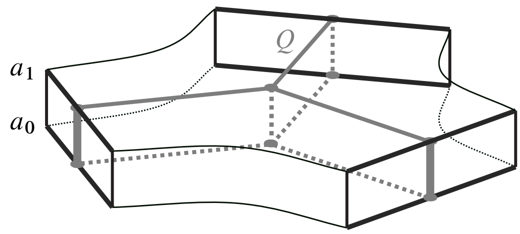


Figure 22: An example of an A_q

4.4.2 Tree regions Now let $N = \{N_i\}$ be a maximal set, up to isotopy, of fibred solid tori embedded in $H - \mathcal{T}$ such that $N_i \cap N_j = \varnothing$ for $i \neq j$ and $\partial H \cap N_i$ is a nonempty set of annuli that are π_1–injective in both ∂N_i and $\partial H - \mathcal{T}$, and the frontier of N_i in H is a nonempty set of annuli essential in $H - \mathcal{T}$ for each i. Then N is a maximal tree region of $H - \mathcal{T}$. The reason for this name will become clearer when we describe it further. Note that by Haken–Kneser finiteness arguments, we can see that N has a finite number of components.

Definition 4.16 Let a *simple q–tree* be a tree that is the cone on $q \geq 2$ points. A vertex of valency one is called an *end vertex*.

Let Q be a simple q–tree. Embed Q in $\mathbb{R}^2 \subset \mathbb{R}^3$. Let P^Q be a $2q$ polygon embedded in $\mathbb{R}^2$ such that every alternate edge intersects Q at an end vertex. Colour the edges of P^Q containing an end vertex of Q thick and all the others thin. Then let $A_q = P^Q \times [0,1]$ and $a_t = P^Q \times \{t\}$, for $t = 0$ or 1. Let Φ_p be a homeomorphism between a_0 and a_1 that twists by $2\pi/p$, such that it maps thick edges to thick edges and thin to thin. This means that $p = q/n$ for $n \in \mathbb{Z}$. Let $A_{(p,q)}$ be A_q with the faces a_0 and a_1 glued according to Φ_p. Therefore, $A_{(p,q)}$ is a solid torus fibred by $\mathbb{S}^1$ with an exceptional fibre of order (p,q). For each $N_i \in N$, there is a unique (p_i, q_i) such that there is a fibre-preserving homeomorphism from $A_{(p_i,q_i)}$ to N_i where the fibring agrees with the boundary curves of the frontier annuli.

Let A_1 and A_2 be two properly embedded essential annuli in $H - \mathcal{T}$ that intersect horizontally and $N(A_1 \cup A_2)$ be a regular neighbourhood disjoint from $\mathcal{T}$. Then the frontier of $N(A_1 \cup A_2)$ in H is a set of properly embedded annuli and tori. Let T be such a torus. The induced map on $\pi_1(T)$ has nontrivial image and $\pi_1(H)$ does not contain any free abelian subgroups of rank 2. Therefore, T bounds a solid torus whose intersection with $N(A_1 \cup A_2)$ is T. Glue solid tori to each torus in the frontier of $N(A_1 \cup A_2)$ in H to produce a submanifold P. Now the frontier of P in H is a set of properly embedded essential annuli and P is a solid torus. Note there is a homeomorphism from P to some $A_{(p,q)}$ that sends the boundary curves of $P \cap \partial H$ to fibres of $A_{(p,q)}$.

Definition 4.17 Let the *tree region* N_T be the union of all components $N_i \in N$ such that $p_i > 2$.

As with the I–bundle region, we are removing the components of N that are homeomorphic to $A_{(1,2)}$ or $A_{(2,2)}$, that is, regular neighbourhoods of properly embedded annuli or Möbius bands, to get N_T. This is because if there are two annuli in $H - \mathcal{T}$ that have a nontrivial vertical intersection, then a maximal tree region can contain the regular neighbourhood of only one of the annuli. Therefore, $H - \mathcal{T}$ may have a number of maximal tree regions. Later it is shown that the tree region is unique up to isotopy.

Lemma 4.18 *If A is a properly embedded annulus in $H - \mathcal{T}$ that has at least one nontrivial horizontal intersection with another properly embedded annulus in $H - \mathcal{T}$, then there is an isotopy of A into N_T.*

This proof is similar to Lemma 4.10.

Proof Let the map $f_i \colon A \to H$, for $i = 1$ or 2, be an essential proper embedding of an annulus A such that $f_i(A) = A_i$ is disjoint from $\mathcal{T}$ for each i and A_1 and A_2 have nontrivial horizontal intersections. Let B be the set of frontier annuli of N_T. If $A_1 \cap N_T \neq \varnothing$, then by Lemmas 4.2 and 4.4, we know that there is an isotopy of f_1 such that the intersection curves between A_1 and the annuli in B are horizontal. Thus the pullback $\Gamma_1 = f_1^{-1}(B)$ is a set of essential simple closed curves in A. Therefore, Γ_1 cuts A up into essential annuli. Let $A' \subset A$ be one of these annuli such that $f_1(A') \subset \overline{H - N_T}$ and let $N(f_1(A'))$ be a regular neighbourhood of $f(A')$ disjoint from $\mathcal{T}$. Then $N(f_1(A'))$ can be fibred as an $A_{(1,2)}$ fibred torus. Thus there must be an isotopy of f_1 to remove the curves $A' \cap \Gamma_1$ (there may be just one if $\partial A \cap \partial A' \neq \varnothing$) otherwise $N_T \cup N(f_1(A'))$ would be larger than N_T, contradicting maximality. So by repeating this process, there is an isotopy of f_1 such that $A_1 \cap B = \varnothing$. This same process produces an isotopy of f_2 so that $A_2 \cap B = \varnothing$. If $A_1 \cup A_2$ is disjoint from N_T, then as above, the torus boundaries of $N(A_1 \cup A_2)$ can be filled in with solid tori so the resulting manifold P is a solid torus. Then $N_T \cup P$ will be a larger tree region contradicting maximality, thus $A_1 \cup A_2 \subset N_T$. □

Once again let $\breve{H}$ be a finite-sheeted normal cover of H, $\breve{\mathcal{T}}$ the lift of $\mathcal{T}$ and G the group of covering translations of $\breve{H}$ such that $\breve{H}/G = H$. Also, let N_T be the tree region in $\breve{H}$. We then get the following lemma.

Lemma 4.19 *Let N_i be a component of N_T. For any $g \in G$, we have that $g(N_i)$ is isotopic to an element of N_T.*

Proof Assume that N_i is a component of N_T and, for some $g \in G$, that $g(N_i)$ is not isotopic to an element of N_T. Let $\boldsymbol{A}$ be the set of frontier annuli of $g(N_i)$ and $\boldsymbol{B}$

be the set of frontier annuli of N_T. By Lemma 4.2, we know that there is an isotopy of g such that any annuli from $\boldsymbol{A}$ and $\boldsymbol{B}$ intersect vertically or horizontally. Also, all trivial intersections are then removed.

Let B be an annulus in $\boldsymbol{B}$ that intersects some annuli from $\boldsymbol{A}$ vertically. Then $B \cap g(N_i)$ is a set of properly embedded squares in $g(N_i)$. Let B' be one such square. As the number of intersections between B and $\boldsymbol{A}$ has been minimized $\partial B'$ is essential in $\partial g(N_i)$. Therefore, $g(N_i)$, and thus N_i, is the regular neighbourhood of an annulus or Möbius band. This implies that $p_i = 2$, contradicting that N_i is a component of N_T. Then any intersections between annuli from $\boldsymbol{A}$ and $\boldsymbol{B}$ must be nontrivial and horizontal. By Lemma 4.18, we can isotope all such annuli from $\boldsymbol{A}$ into N_T.

We have now isotoped g so that $\boldsymbol{A} \cap \boldsymbol{B} = \varnothing$. We can thus isotope g so that $g(N_i)$ lies inside a single component of N_T, otherwise $g(N_i) \cup N_T$ would be a larger tree region, contradicting maximality of N_T. Let $g(N_i)$ lie in $N_k \in N_T$. If $g(N_i)$ is not isotopic to N_k, then $g^{-1}(N_k - g(N_i)) \cup N_T$ is a larger tree region. □

From the previous lemma we get the following corollary.

Corollary 4.20 *For any $g \in G$, we have that $g(N_T)$ is isotopic to N_T.*

From the above corollary and using the same least area arguments as we did with I–bundle regions we get the following corollary.

Corollary 4.21 *There is an isotopy of N_T in $\breve{H}$ that is preserved by the covering transformations.*

This means that N_T will project down to a nontrivial tree region in H. If H contains two embedded annuli that have a nontrivial horizontal intersection, then H has a nonempty tree region. Note this is a sufficient condition but not a necessary one.

Lemma 4.22 *N_T is unique up to ambient isotopy of H.*

We will not give the proof for this lemma as the argument is the same as Lemma 4.11. The idea is that if we assume that there are two tree regions N_T and N'_T that are not isotopic, then we get a contradiction to their maximality.

4.4.3 Annulus regions It is clear from the definitions of N_I and N_T that:

Lemma 4.23 *If H is a handlebody and $\mathcal{T}$ is a set of curves in ∂H that meet the n disk-condition, then there is an isotopy of N_I and N_T such that $N_I \cap N_T = \varnothing$.*

Let A_I be the set of I–bundles in a maximal I–bundle region but not in N_I. That is, they have base spaces that are either annuli or Möbius bands. Let A_T be the set of fibred solid tori that are in a maximal tree region but not in N_T. That is, they are all the components of the maximal tree region whose associated trees have two end vertices. Let N_A be those components of A_T which are ambient isotopic to components of A_I. Components of N_A are regular neighbourhoods of properly embedded annuli or Möbius bands and they can be fibred by intervals or circles. The components of $A_I - N_A$ ($A_T - N_A$) are the components of the maximal I–bundle (maximal tree region) that cause the maximal I–bundle (maximal tree region) to be not unique and, in fact, the components of $A_I - N_A$ ($A_T - N_A$) can be isotoped into N_T (N_I).

Clearly by the definition, N_A can be isotoped to be disjoint from N_I and N_T. Therefore, it is contained in the set of handlebodies $H' = \overline{H - (N_I \cup N_T)}$. Any annulus that can be made to intersect another nonparallel annulus either vertically or horizontally is isotopic into $N_I \cup N_T$. Thus any nonparallel annuli in H' cannot be isotoped to intersect either vertically or horizontally. Therefore, by the maximality of the maximal I–bundle region and the maximal tree region we know that N_A is isotopic to the regular neighbourhood of the maximal set of disjoint and nonparallel properly embedded annuli in H'. Thus we get the following lemma.

Lemma 4.24 *N_A is unique up to ambient isotopy of H and can be isotoped to be disjoint from $N_I \cup N_T$.*

Definition 4.25 If H is a handlebody and $\mathcal{T}$ is a set of essential disjoint simple curves in its boundary that satisfies the n disk-condition, then for the pair $\{H, \mathcal{T}\}$, let the *maximal annulus region* be $N = N_I \cup N_T \cup N_A$, where N_I, N_T and N_A are as defined above.

4.4.4 Disk regions In this section, we want to define the building blocks for the flavour of characteristic variety that intersects the triple curves. In each handlebody H_i, these blocks look like the regular neighbourhood of meridian disks that intersect the triple curves exactly n_i times, where $\sum 1/n_i = \frac{1}{2}$. Hence we will refer to them as *disk regions*. Let H be a handlebody and $\mathcal{T}$ a set of essential curves in its boundary that meet the n disk-condition in H. Let $\boldsymbol{D}$ be a set made up of a single representative from each isotopy class of meridian disks that intersect $\mathcal{T}$ exactly n times. Let S be the resulting punctured sphere when ∂H is cut along a waveless basis for $\mathcal{T}$. Then $\Gamma = \mathcal{T} \cap S$ is a set of pairwise disjoint properly embedded arcs that cut S into n–gons. Therefore, by normal curve theory up to isotopy there is a finite number of simple closed curves in the interior of S that have n intersections with Γ and waves that have $\frac{1}{2}n$ intersections with Γ. Thus $\boldsymbol{D}$ contains a finite number of disks.

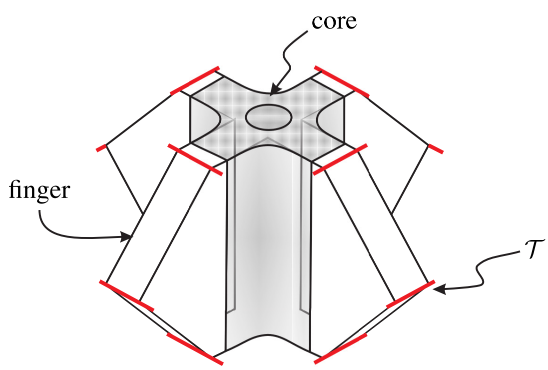


Figure 23: A component of the disk region

Assume that the disks in $\boldsymbol{D}$ have been isotoped so that the intersection between any pair of disks is a set of bisecting arcs and the intersection between any three disks is empty. Let $N(\boldsymbol{D})$ be the regular neighbourhood of $\boldsymbol{D}$. Then the frontier of $N(\boldsymbol{D})$ in H is a set of properly embedded disks that have n intersections with $\mathcal{T}$ and annuli that are disjoint from $\mathcal{T}$. For any of the boundary components that are either nonmeridian disks or nonessential annuli, add the appropriate 3–cell to $N(\boldsymbol{D})$. The resulting submanifold P is the *disk region*.

By Lemma 4.8, we can isotope the disks in $\boldsymbol{D}$ so that the intersection between any pair of disks is a set of parallel bisecting arcs and the intersection between any three is empty. Therefore, for any disk $D_i \in \boldsymbol{D}$, the intersection $\Gamma_i = D_i \cap (\boldsymbol{D} \backslash D_i)$ is a set of parallel bisecting arcs.

Then there are two innermost bisecting arcs in D_i. Therefore, when D_i is cut along the innermost bisecting arcs the result is three disks: two bigons and a third quadrilateral. Let D_i' be the third disk. Let $\boldsymbol{D}'$ be the set of disks produced when this is done to all disks in $\boldsymbol{D}$. Then $\bigcup D_i'$ is an I–bundle over a graph. This fibring can then be extended to the "core" of each component of P. The unfibred parts of each component are the regular neighbourhoods of disks that have $\frac{1}{2}n$ intersections with $\mathcal{T}$ and which boundary compress the frontier annuli of the core. We will call these *fingers*; see Figure 23. Note that each component has at least one finger. Unlike the I–bundle regions defined earlier, the core may have a disk as its base space. The fibring of each component is unique, up to isotopy, except if the component is the regular neighbourhood of a single meridian disk. In the latter case we do not fibre the core until later.

Lemma 4.26 *All possibly singular meridian disks which have n intersections with $\mathcal{T}$ can be homotoped into P.*

Proof Let D be a disk and $f\colon D \to H$ be a possibly singular map such that $A = f(D)$ is a meridian disk, ie has essential boundary. Let P be the maximal disk region, as defined above and $f^{-1}(\mathcal{T})$ be n vertices in ∂D. Then B, the frontier of P in H, is a set of meridian disks and annuli essential in $H - \mathcal{T}$. Then $\Gamma = f^{-1}(B)$ is a set of properly embedded arcs and simple closed curves in D. As H is irreducible there is a homotopy of f to remove all simple closed curves from Γ. Thus Γ is a set of properly embedded disjoint simple arcs in D.

By maximality of P, any boundary compressing disks of a component of B, as described in Section 3.1, must intersect $\mathcal{T}$ more than $\frac{1}{2}n$ times. There must be an innermost disk $D_1 \subset D$ such that $f(D_1)$ intersects $\mathcal{T}$ at most $\frac{1}{2}n$ times. Thus by Dehn's lemma and the loop theorem — see Lemma 3.1 — we can remove any arc from Γ which is in the image of ∂D_1. We can repeat this process until A is disjoint from B. Thus either A is contained in P or disjoint from P. If it is disjoint, then there must be a homotopy of f such that $A \subset P$. Otherwise, using Dehn's lemma and the loop theorem, we get a contradiction to the maximality of P. □

4.5 Handlebodies and singular annuli

In Jaco and Shalen's [7] and K Johannson's [8] proofs of the torus theorem, an essential step is the annulus theorem. In fact, the torus theorem is a consequence of the annulus theorem. Similarly, a lemma that is a slight variation of the annulus theorem is required here. Our annulus theorem is simpler as it is restricted to handlebodies. Namely, suppose a handlebody H has a set of curves in its boundary, $\mathcal{T}$, that satisfies the n disk-condition. Assume also there is a proper essential (possibly singular) map f of an annulus into $H - \mathcal{T}$. Then f is properly homotopic to an essential (possibly singular) map of an annulus into the maximal annulus region. There are two main steps to prove this lemma. The first is to show that if there is a proper singular essential map of an annulus into $H - \mathcal{T}$, then there is a similar embedded one. Next we show any proper essential embedding of an annulus in $H - \mathcal{T}$ is properly isotopic into one of its maximal annulus regions.

Lemma 4.27 *Let H be a handlebody and $\mathcal{T}$ a set of simple closed curves in ∂H that meet the n disk-condition. Let A be an annulus and $f\colon A \to H - \mathcal{T}$ a proper immersion. If f is not properly homotopic into $\partial H - \mathcal{T}$ and the curves $f(\partial A)$ are essential in ∂H, then there is a properly embedded essential annulus in $\partial H - \mathcal{T}$.*

Remark 4.28 The proof for this lemma uses a simplified version of the covering space argument used by Freedman, Hass and Scott [3]. The argument is easier, since we are operating in a handlebody.

Proof The first step is to find another f such that all the lifts of $f(A)$ in the universal cover are embedded. We then use subgroup separability to produce a finite-sheeted cover of H which contains a lift of $f(A)$ that is embedded and does not intersect any of its translates. From this cover we find a regular cover, in which all the lifts of $f(A)$ are embedded. This then implies that the finite regular cover has a nontrivial annulus region and thus so does the original handlebody.

We will assume that the map f is transverse at all times. Let $G=\pi_1(H)$, f_* be the induced map on $\pi_1(A)$ and $f_*(\pi_1(A))=B\subseteq G$. Therefore, B is a cyclic subgroup generated by some $z\in G$.

Let $\bar{H}$ be the cover of H with the projection $\bar{p}\colon \bar{H}\to H$ such that $\bar{p}_*(\pi_1(\bar{H}))=B$. This means there is a lift, $\bar{f}$ of f, which is an immersed annulus such that $\pi_1(\bar{H})\cong \bar{f}_*\pi_1(A)$. Let $\bar{\mathcal{T}}=\bar{p}^{-1}(\mathcal{T})$. As f is not properly homotopic into $\partial H-\mathcal{T}$, we have that $\bar{f}$ is not properly homotopic into $\partial\bar{H}-\bar{\mathcal{T}}$.

We now want to find an embedded annulus in $\bar{H}$ which is π_1–injective and not properly homotopic into $\partial\bar{H}-\bar{\mathcal{T}}$. Let $N(\bar{f}(A))$ be a regular neighbourhood of $\bar{f}(A)$ such that $N(\bar{f}(A))\cap\bar{\mathcal{T}}=\varnothing$. Then the frontier of $N(\bar{f}(A))$ in $\bar{H}$ is a set of embedded surfaces. As $\pi_1(\bar{A})\cong\pi_1(\bar{H})$, we can find two of these embedded surfaces in $\bar{H}$ both of which have (at least) two essential boundary curves in $\bar{H}-\bar{\mathcal{T}}$. (Note that $\bar{H}$ is a missing boundary solid torus, ie has interior which is an open solid torus and compactifies to a solid torus.) Let one of these surfaces be $\bar{A}'$. The boundary curves of $\bar{f}(A)$ are not homotopic in $\partial\bar{H}-\bar{\mathcal{T}}$, that is, $\bar{f}$ is not homotopic into $\partial\bar{H}-\bar{\mathcal{T}}$; thus the two essential boundary curves of $\bar{A}'$ can be chosen to be not homotopic in $\partial\bar{H}-\bar{\mathcal{T}}$.

By Dehn's lemma and the loop theorem, since $\pi_1(\bar{H})$ is infinite cyclic, we know that any handles in $\bar{A}'$ can be compressed until $\bar{A}'$ is an essential embedded annulus in $\partial\bar{H}-\bar{\mathcal{T}}$. Now let $A'=\bar{p}(\bar{A}')$. We can assume that $\bar{p}$ restricted to $\bar{A}'$ is transverse. Let $\bar{A}_i$, for $1\le i\le n$, be the lifts of A' in $\bar{H}$ that intersect $\bar{A}'$ and $\bar{\alpha}_i=\bar{A}'\cap\bar{A}_i$. Thus each $\bar{\alpha}_i$ is a set of singular curves in $\bar{A}'$.

Let $\tilde{H}$ be the universal cover of H and therefore also the universal cover of $\bar{H}$ with the projections $p\colon\tilde{H}\to H$ and $\tilde{p}\colon\tilde{H}\to\bar{H}$, such that $p=\bar{p}\tilde{p}$. As H is a handlebody, $\tilde{H}$ is a missing boundary ball, that is, a ball with a compact set removed from its boundary. As A' is π_1–injective in H, each pullback to $\tilde{H}$ is a universal cover of A', an infinite strip. As $\bar{A}'$ is embedded in $\bar{H}$, each pullback to $\tilde{H}$ is embedded. Then by applying the covering transformation group to $\tilde{H}$ we know that all the lifts of A' in $\tilde{H}$ are infinite strips.

Let $\tilde{A}$ be a lift of $\bar{A}'$ in $\tilde{H}$. Then any lift of A' in $\tilde{H}$, that intersects $\tilde{A}$ must be a lift of one of the $\bar{A}_i$ in $\bar{H}$. Let $\tilde{A}_i$ be some lift of $\bar{A}_i$ that intersects $\tilde{A}$ and $\tilde{\alpha}_i=\tilde{A}\cap\tilde{A}_i$.

Note this means that $\widetilde{p}(\widetilde{\alpha}_i) = \overline{\alpha}_i$. Also, let $\widetilde{G}$ be the group of deck transformations on $\widetilde{H}$ and $\widetilde{B} \subset \widetilde{G}$ the stabilizer of $\widetilde{A}$. Therefore, $\widetilde{G} \cong G$ and $\widetilde{B}$ is the cyclic subgroup of translations along $\widetilde{A}$. Also, let $g_i \in \widetilde{G}$ where $g_i(\widetilde{A}) = \widetilde{A}_i$. This means that $g_i \notin \widetilde{B}$ and that $\widetilde{B}_i = g_i \widetilde{B}$ is the set of transformations taking $\widetilde{A}$ to $\widetilde{A}_i$. So for all $b \in \widetilde{B}$, $\widetilde{p}(b(\widetilde{\alpha}_i)) = \overline{\alpha}_i$.

By Hall [4], we know there is a finite index subgroup $\widetilde{L}_i \subseteq \widetilde{G}$ such that $\widetilde{B} \subseteq \widetilde{L}_i$ but $g_i \notin \widetilde{L}_i$. This property is called subgroup separability. For all $b \in \widetilde{B}$, we have that $bg_i\widetilde{A}$ is a translate that intersects $\widetilde{A}$ and $bg_i \notin \widetilde{L}_i$. This means for any $l \in \widetilde{L}_i$ that $l(\widetilde{A}) \neq b(\widetilde{A}_i) = bg_i(\widetilde{A})$ for all $b \in \widetilde{B}$. In other words none of the deck transformations in $\widetilde{L}_i$ map $\widetilde{A}$ to the lift of $\overline{A}_i$ that intersects $\widetilde{A}$. Let $\widehat{H}_i = \widetilde{H}/\widetilde{L}_i$ be the cover of H with the fundamental group corresponding to $\widetilde{L}_i$ such that $\widehat{p}_i\colon \widetilde{H} \to \widehat{H}_i$. Therefore, $\widehat{p}_i(\widetilde{A})$ is an embedded annulus in $\widehat{H}_i$. Also, $\widehat{p}_i(b\widetilde{A}_i) \cap \widehat{p}_i(\widetilde{A}) = \varnothing$ for any $b \in B$, and as $\widetilde{L}_i$ has finite index in G, we have that $\widetilde{H}_i$ is a finite-sheeted cover of H.

Therefore, $L = \widetilde{L}_1 \cap \cdots \cap \widetilde{L}_n$ is a finite index subgroup of $\widetilde{G}$ such that for $l \in L$, either $\widetilde{A} = l(\widetilde{A})$ or $\widetilde{A} \cap l(\widetilde{A}) = \varnothing$. Let $\widetilde{H}/L = \widehat{H}$ be the finite-sheeted cover of H with the projection $\widehat{p}\colon \widetilde{H} \to \widehat{H}$. Then $\widehat{p}(\widetilde{A}) = \widehat{A}$ is an embedded annulus in $\widehat{H}$ that does not intersect any other lifts of A'.

As L has finite index, it must have a finite number of right cosets, $\{Lx_1, \dots, Lx_n\}$, for $x_1, \dots, x_n \in G$. Assume that $Lx_1 = L$. Thus if S_n is the group of permutations of n elements, there is a map $\phi\colon G \to S_n$, where $\phi(g)$, for $g \in G$, is the element of S_n that sends $\{Lx_i\}$ to $\{Lx_i g\}$. Both $\phi(g_1)\phi(g_2)$ and $\phi(g_1g_2)$ send $\{Lx_i\}$ to $\{Lx_i g_1 g_2\}$, so ϕ is a homomorphism. Let $K \subseteq G$ be the kernel of ϕ. If $g \in K$, then $Lx_i = Lx_i g = Lgx_i$, thus $K \subseteq L$. As S_n has a finite number of elements, the kernel K is a finite index normal subgroup. Therefore, $\breve{H} = \widetilde{H}/K$ is a finite-sheeted normal cover of H. Let $\breve{p}\colon \breve{H} \to H$ be the covering projection. Then $\breve{H}$ is a handlebody and $\breve{\mathcal{T}} = \breve{p}^{-1}(\mathcal{T})$ is a set of curves in $\partial\breve{H}$ that meet the n disk-condition in $\breve{H}$. Also, $\breve{H}$ is a cover of $\widehat{H}$; thus all the lifts of A' are properly embedded essential annuli in $\partial\breve{H} - \breve{\mathcal{T}}$.

Then by Freedman, Hass and Scott [3], if we put a Riemannian metric on H and properly homotope A' to be of least area, then all trivial self intersections between lifts of A' will be removed, and thus by Lemmas 4.2 and 4.4 all the lifts of A' in $\breve{H}$ are either pairwise disjoint or intersect each other vertically or horizontally. If the lifts of A' are pairwise disjoint, A' must be a properly embedded essential annulus in $\partial H - \mathcal{T}$. Otherwise, by Lemmas 4.10 and 4.18, we know that $\breve{H}$ must have a nontrivial region $N_I \cup N_T$. By Lemmas 4.11 and 4.19, we know that $N_I \cup N_T$ can be isotoped so that its frontier annuli are preserved under K and thus project to properly embedded essential annuli in $\partial H - \mathcal{T}$. □

Lemma 4.29 *If H is a handlebody, $\mathcal{T}$ is a set of triple curves in its boundary that satisfies the n disk-condition and $f\colon A \to H$ is a properly embedded annulus, then f is properly isotopic into the maximal annulus region N.*

Proof Let $f\colon A \to H$ be a properly embedded annulus that cannot be properly isotoped into N. By Lemmas 4.18 and 4.10, we know that if $f(A)$ has a nontrivial intersection with another embedded annulus, then f can be isotoped into N_I or N_T. Therefore, we can isotope f so that its image is disjoint from all the frontier annuli of N. This contradicts maximality of N, thus we must be able to properly isotope $f(A)$ into N. □

Lemma 4.30 *Let H be a handlebody, $\mathcal{T}$ a set of curves in its boundary that satisfies the n disk-condition and N the annulus region in H. If A is an annulus and $f\colon A \to H - \mathcal{T}$ is a proper singular essential map, then there is a proper homotopy of f such that $f(A)$ is in N.*

Proof To save on notation, we will refer to $f(A)$ by A as well. Let B be the set of frontier annuli of N and $\mathcal{T}' = \mathcal{T} \cup \partial B$. Then $H' = \overline{H - N}$ is a set of handlebodies such that for any component H'_j, the set of essential simple closed curves $\mathcal{T}' \cap H'_j$ satisfies the 4 disk-condition in H'_j. Also, there is a proper homotopy of f such that $f^{-1}(N)$ is either a set of 4–gons (case 1) or essential embedded annuli (case 2).

Case 1 All the components of N that A intersects are either in N_I or N_A. Assume the singular 4–gons $H' \cap A$ are essential in H'. Then by Dehn's lemma and the loop theorem, we know that there is an embedded essential 4–gon with two boundary arcs in the frontier annuli of N. This contradicts maximality of N.

Case 2 Here, all the components of N that A intersects are either in N_T or N_A. Then by Lemma 4.27 we know that H' must contain an essential properly embedded annulus, contradicting maximality of N.

Thus there must be a proper homotopy of f such that A is disjoint from B. If A is not contained in N, then once again by Lemma 4.27, H' contains essential embedded annuli, contradicting maximality of N. □

4.6 Torus theorem

Let M be a 3–manifold that satisfies the (n_1, n_2, n_3) disk-condition. That is, $H_i \subset M$ is an embedded handlebody for $1 \le i \le 3$ such that $\bigcup H_i = M$, $\bigcup \partial H_i = X$ is a 2–complex that cuts M up into the H_i, and $\bigcap H_i = \mathcal{T}$ is a set of essential simple closed curves that meet the n_i disk-condition in H_i. We will assume that (n_1, n_2, n_3) is either $(6, 6, 6)$, $(4, 6, 12)$ or $(4, 8, 8)$, for if the gluing of the three handlebodies meets some disk-condition, it meets one of these three.

Lemma 4.31 *Let M be a closed 3–manifold that satisfies the disk-condition as described above. Suppose T is a torus and $f\colon T \to M$ is an essential possibly singular map. Then there is a homotopy of f such that either $f(T)$ is disjoint from $\mathcal{T}$ and $H_i \cap f(T)$ is a set of essential annuli for each i, or $H_i \cap f(T)$ is a set of singular disks for each i with essential boundaries that each intersect $\mathcal{T}$ exactly n_i times.*

Proof Assume that f is transverse to X. Thus $\Gamma = f^{-1}(X)$ is a set of simple closed curves and trivalent embedded graphs which separates T. Define an (m,n)–gon to be a face of T that is a disk, has m vertices in its boundary and is mapped by f into the handlebody in which $\mathcal{T}$ satisfies the n disk-condition. Let the Γ_j be the components of Γ. Then Γ_i is a nonessential component if there is a disk $D \subset T$ such that $\Gamma_i \subset D$. So by Lemma 3.2, we know that there is a homotopy of f to remove Γ_i and hence all nonessential components of Γ.

Consequently, there are two cases. Either all faces of Γ are disks or Γ has faces which are essential annuli. Note that $f(T)\cap X \neq \varnothing$ as f is π_1–injective and $\pi_1(H_i)$ doesn't have a free abelian subgroup of rank 2.

If Γ is connected, then all the faces must be (m,n)–gons and all the vertices have order three. Let $\boldsymbol{F}$ be the set of faces of T. We can then put a metric on T, as we did in the proof of Lemma 3.2. So all the edges are geodesics of unit length, and if $F \in \boldsymbol{F}$ is an (m,n)–gon, then the angle at each vertex is $\pi(1-2/n)$ and there is a cone point in $\operatorname{int}(F)$. Once again this means that the curvature around each vertex is 2π. Let $\boldsymbol{K}(F)$ be the curvature at the cone point in F. By the Gauss–Bonnet theorem, we know that

$$\boldsymbol{K}(F) = 2\pi(1-m/n).$$

Therefore, if $m > n$ then $\boldsymbol{K}(F) < 0$, if $m = n$ then $\boldsymbol{K}(F) = 0$ and if $m < n$ then $\boldsymbol{K}(F) > 0$. Also, by the Gauss–Bonnet theorem, we know that

$$\sum_{F \in \boldsymbol{F}} \boldsymbol{K}(F) = 0.$$

Therefore, if $\boldsymbol{F}$ contains an (m,n)–gon such that $m > n$, then it must also contain a face F such that $m < n$. Thus by the disk-condition we know that $f(\partial F)$ is not essential in ∂H_k. So there is a homotopy of f such that $f(F) \subset \partial H_k$. We can then push F off ∂H_k removing the face F from $\boldsymbol{F}$. Note that when we do this, the order of the faces adjacent to F either decreases by two or an (m,n)–gon and an (m',n)–gon merge to become an $(m+m'-4,n)$–gon, as shown in Figure 3. We can repeat this process as long as $\boldsymbol{F}$ contains faces with positive curvature. Each time we do this move, we reduce the number of faces in $\boldsymbol{F}$ by at least one. Therefore, this process must terminate after a finite number of moves, when all the faces are (m,n)–gons such that $m = n$.

Now let's look at the case where Γ contains more than one component. Let Γ_i be a component of Γ. Then Γ_i cuts T up into faces that are a single annulus and a number of disks. Let A be the union of Γ_i and the faces which are disks. Now we know that the Euler characteristic of A is 0. Put a metric on A as we did above. Γ_i must have boundary vertices, that is vertices adjacent to less than three faces of A. Thus using the same arguments using the Gauss–Bonnet theorem we know that A must have some face with positive curvature. This means that such faces are boundary parallel in the handlebody and there is a homotopy of f to remove them. As before this process can be repeated until all the components are simple closed essential loops. $\square$

We are now ready to prove the torus theorem.

Proof of Theorem 1.1 Let N_i be the maximal annulus region for H_i and P_i be the maximal disk region for H_i. The idea of this proof is to find submanifolds of either the N_i or the P_i such that when glued together, the resulting embedded submanifold can be fibred by $\mathbb{S}^1$ and either has essential tori boundary or the fibring can be extended to the whole of M. In the interest of reducing notation, the image of $f(T)$ in M will be denoted as T. Thus when we talk about a homotopy of T, we are implying a homotopy of f.

By Lemma 4.31, there is a homotopy such that either T is disjoint from $\mathcal{T}$ and for each i, $H_i \cap T$ is a set of essential singular annuli not properly homotopic into $\partial H - \mathcal{T}$ or, for each i, $H_i \cap T$ is a set of singular meridian disks that intersect $\mathcal{T}$ exactly n_i times.

The first case is therefore that T is disjoint from the triple curves and $H_i \cap T$ is a set of singular essential annuli. We can also assume that no components of $H_i \cap T$ are properly homotopic into $\partial H_i - \mathcal{T}$. By Lemma 4.30, we can isotope each N_i so that $H_i \cap T \subset N_i$.

Let $A_i = X \cap N_i$, where $X = \bigcup \partial H_j$. Then A_i is a set of essential surfaces in ∂H_i and the boundary of the maximal annulus region N_i. Note that $T \cap \partial H_i \subset A_i$ and thus $T \cap X \subset \bigcup_{i \neq j}(A_i \cap A_j)$. We will first shrink N_1. Let $S_i = A_i \cap (A_j \cup A_k)$, where i, j and k are different. Let N_1' be the maximal subset of N_1 such that $N_1' \cap X \subseteq S_1$ and each component of the frontier of N_1' in H_1 is an essential annulus parallel to the fibring of N_1. There are three cases to discuss corresponding to components of N_I, N_T and N_A.

Let B be a component of N_1 such that B is an I–bundle region and F is its base space. Then let $F' \subseteq F$ be the maximal subsurface such that $B' \cap \partial H_1 \subseteq S_1$, where B' is the I–bundle over F'. Then B' is a component of N_1'. Note that components that do not intersect S_1 are removed.

If B is a tree region, then it is a fibred solid torus and $B \cap \partial H_1$ is a set of essential annuli. Then there is an isotopy of B such that each annulus in $B \cap \partial H_1$ is either contained in S_1 or in $\mathrm{int}(H_1)$. Note that some annuli in ∂H_1 may get pushed into $\mathrm{int}(H_1)$. Let B' be the resulting fibred torus. Note that when the number of annuli in $B \cap \partial H$ is reduced to produce B', the fibring of the torus is still parallel to the boundary curves of the frontier annuli. Then B' is a component of N_1'. If $B' \cap H_1 = \varnothing$ we remove it from N_1'.

If B is a component of N_A, as defined in Section 4.4.3, then either it can be isotoped so that $B \cap H_1 \subseteq S_1$ or it is removed. As $T \cap X \subset \bigcup_{i \neq j}(A_i \cap A_j)$ we know that $N_1' \neq \varnothing$. We now let $N_1 = N_1'$.

We now repeat this process for each N_i in turn until the process stabilises. That is, for $i \neq j$, $i \neq k$ and $k \neq j$, we have $A_i = \partial H_i \cap (A_j \cup A_k)$. We know that it stabilises before $\bigcup N_i = \varnothing$ because $T \subset \bigcup N_i$.

Next we want to change the fibrings of the N_i so that all components that are regular neighbourhoods of embedded annuli or Möbius bands are fibred by $\mathbb{S}^1$. This means that for any component B of N_i such that $B \cap \partial H_i$ is a set of annuli, then B is a fibred solid torus, or an I–bundle. Now when we let $\boldsymbol{N} = \bigcup N_i$ and all the fibrings of components match, then N is a Seifert fibred submanifold of M and $\partial \boldsymbol{N}$ is a set of embedded tori.

By Lemma 4.15, if $\boldsymbol{N}_j$ is a component of $\boldsymbol{N}$ such that $H_i \cap \boldsymbol{N}_j$ is an I–bundle with a base space that is not an annulus or a Möbius band, then the boundary tori of $\boldsymbol{N}_j$ are essential in $\boldsymbol{M}$. The final step in this case is to either make all the boundary tori of $\boldsymbol{N}$ essential or expand $\boldsymbol{N}$ so that $\boldsymbol{N} = \boldsymbol{M}$. If $\boldsymbol{N}_j$ is a component of $\boldsymbol{N}$ and $F \subset \boldsymbol{M}$ is an embedded solid torus such that $\partial F \subseteq \boldsymbol{N}_j$, then either $F \cap \boldsymbol{N}_j = \partial F$ or $F \cap \boldsymbol{N}_j = \boldsymbol{N}_j$. If $F \cap \boldsymbol{N}_j = \partial F$, we then add F to $\boldsymbol{N}$ and extend the fibring to it. This can always be done as the fibres of the component are essential in $\boldsymbol{M}$. Therefore, the meridian disk of the solid torus being added cannot be parallel to the fibring of $\boldsymbol{N}_j$. If $\boldsymbol{N}_j$ is contained in F we remove $\boldsymbol{N}_j$ from $\boldsymbol{N}$. This process is repeated until either all boundary tori are essential or $\boldsymbol{N} = \boldsymbol{M}$. We know the process will terminate before all of $\boldsymbol{N}$ has been removed because $T \subset \boldsymbol{N}$ and T is essential. Thus the component containing T cannot be contained in a solid torus.

The next case is when $H_i \cap T$ is a set of singular n_i–gons. Let P_i be the disk region in the handlebody H_i. Next we want to define a process for shrinking components of P_i until all their boundaries coincide in X and then show that we can expand the "core" fibring to the whole submanifold. Let $A_i = X \cap P_i$. By Lemma 4.26, we know that we can isotope each P_i so that $H_i \cap g(T) \subset P_i$. Thus $T \cap \partial H_i \subset P_i \cap (P_j \cup P_k)$, for $i \neq j$, $j \neq k$ and $k \neq i$.

Reduce P_1 so that $P_1 \subseteq P_2 \cup P_3$. By reducing, we mean chop off fingers that don't match up, reduce base spaces of the cores and possible remove entire components of P_1. This process finishes before P_1 is entirely removed as $T \cap \partial H_i \subset P_i \cap (P_j \cup P_k)$.

Note that if a component of P_1 is reduced to the regular neighbourhood of a single meridian disk we forget the fibring of its core. As we reduce P_1, the frontier of P_1 in H remains a set of essential annuli and meridian disks.

This process is repeated in turn for each P_i. Once again we know that the process stabilises before all the P_i are removed as $T\cap\partial H_i\subset P_i\cap(P_j\cup P_k)$. All the components with fibred cores obviously match up to be fibred tori in $\boldsymbol{P}=\bigcup P_i$. Clearly these do not intersect so the fibring can be extended across $\boldsymbol{P}$. Also, $\boldsymbol{P}$ is a Seifert fibred submanifold of M and each of the boundary tori of $\boldsymbol{P}$ is tiled by either meridian disks or essential annuli that are essential in T. As before if any of the torus boundaries of $\boldsymbol{P}$ are not essential, they are either filled in with a solid torus or removed. □

4.7 Characteristic variety

Finally we show that both flavours of characteristic variety fit together nicely. That is, if the flavours intersect, their $\mathbb{S}^1$ fibrings can always be made to agree. If either component is a $T^2\times I$, this is easy. Thus we want to study the case where each component has a unique fibring.

Let $\boldsymbol{N}$ be the maximal annulus region in M and $\boldsymbol{P}$ be the maximal disk region. By the usual arguments, we can see that both are unique up to isotopy. We can also assume that $\boldsymbol{N}$ is disjoint from $\mathcal{T}$ and that both flavours have nonempty boundary. Thus $\partial\boldsymbol{N}\cup\partial\boldsymbol{P}$ is a set of essential embedded tori. If $\boldsymbol{N}\cap\boldsymbol{P}=\varnothing$, then there is no problem. Therefore, we can assume that $\boldsymbol{N}\cap\boldsymbol{P}\neq\varnothing$. Let N' be a component of $\boldsymbol{N}$ and P' be a component of $\boldsymbol{P}$ such that $N'\cap P'\neq\varnothing$. It is not possible for $P'\subset N'$ and if $N'\subset P'$ there is no problem. Therefore, we can assume that there is a boundary torus $B\subset\partial P'$ such that $B\cap N'\neq\varnothing$. As $\partial N'$ is a set of essential tori, $B\cap N'$ is a set of essential annuli in N'. Thus $H_i\cap(B\cap N')$, for any i, is a set of quadrilaterals. Therefore, if the components of $H_i\cap N'$ are fibred by $\mathbb{S}^1$, then $N'\cong T^2\times I$. Thus we can assume that N' is fibred such that $N'\cap H_i$ is a set of I–bundles. Therefore, it just remains to show that $H_i\cap(N'\cap P')$ is an I–bundle.

Let F and F' be two meridian disks in H_i that have n_i intersections with $\mathcal{T}$ and have a nontrivial intersection and A be an essential properly embedded annulus in $H_i-\mathcal{T}$. We can assume that A has been isotoped so that $F\cap A$ is a set of disjoint properly embedded arcs in F. If any of the arcs in $F\cap A$ are not bisecting, then A is boundary parallel. In this case $F'\cap A$ cannot contain any properly embedded arcs, for if it did, this would provide an isotopy of F to remove that intersection between F and F'. Thus $F\cap A$ must be a set of bisecting arcs in F, similarly $F'\cap A$ is a set of properly embedded bisecting arcs in F' and A is not boundary parallel. If we then let Q be the regular neighbourhood of $F\cup F'$, then B, the frontier of Q in H, is a set of properly embedded annuli and meridian disks that intersect $\mathcal{T}$ exactly n_i times. As in the proof of Lemma 4.8, there is an isotopy of A such that $A \cap B$ is a set of properly embedded parallel arcs that are not boundary parallel in A. Thus there is an isotopy to remove any triple points.

The components of $P' \cap H_i$ can be thought of as regular neighbourhoods of a set of meridian disks that intersect $\mathcal{T}$ exactly n_i times. From above, if there are two meridian disks in H_i that have a nontrivial intersection and that have n_i intersections with $\mathcal{T}$, then any essential annulus can be isotoped so that it is disjoint from their intersection. Lemma 4.15 says any boundary compressing disk of the annuli $N' \cap H_i$ has order at least $\frac{1}{2}n_i$. Therefore, the intersection between frontier annuli of $N' \cap H_i$ and a meridian disk of order n_i must be bisecting in the meridian disk. By these two observations, we can see that $H_i \cap (N' \cap P')$ is an I–bundle.

4.8 Atoroidal manifolds

An interesting question asked us by Cameron Gordon, is to find an additional condition that would result in manifolds satisfying the n disk-condition being atoroidal. By Lemma 4.31, a sufficient condition for a manifold M that satisfies the disk-condition to not contain any essential tori that intersect the triple curves, is the manifold meets a stronger disk-condition with $\sum 1/n_i < \frac{1}{2}$. A sufficient condition that M does not contain any essential tori disjoint from the triple curves is that in at least two of the handlebodies, any essential annuli disjoint from $\mathcal{T}$ are boundary parallel.

Let H be a handlebody and $\mathcal{T}$ an essential set of disjoint simple closed curves in ∂H that meet the n disk-condition. Let A be a properly embedded essential annulus in H disjoint from $\mathcal{T}$. Then by Lemma 3.16, H has a waveless minimal system of disks, $\mathbb{D}$; see Definition 3.9. Let B be the 3–ball produced when H is cut along $\mathbb{D}$, let $S \subset \partial B$ be the punctured sphere produced when ∂H is cut along $\mathbb{D}$ and let $\Gamma = \mathcal{T} \cap S$. As in the proof for Lemma 3.18, let $\Gamma' \subset \mathbb{S}^2$ be the graph produced by letting components of ∂S correspond to vertices and parallel components of Γ correspond to single edges; see Figure 11.

As A is a properly embedded essential annulus, $B \cap A = \{A_1, \dots, A_k\}$ is a set of properly embedded quadrilaterals in B such that $A_i \cap S$ is two properly embedded arcs in S for any i. An equivalent statement to A being boundary parallel is that the curves ∂A are parallel in ∂H or that for each i, the arcs $A_i \cap S$ are parallel in S.

Lemma 4.32 *If Γ' is maximal and contains no 2–cycles* (Definition 3.17), *then all properly embedded annuli in H disjoint from $\mathcal{T}$ are boundary parallel.*

Proof By maximality of Γ', the arcs of $A_i \cap S$, for all i, must be parallel to some arc of Γ and as Γ' contains no 2–cycles, both arcs of $A_i \cap S$ must be parallel to the same arc of Γ and thus parallel. Therefore, from above, any properly embedded essential annulus in $H - \mathcal{T}$ must be boundary parallel. $\square$

Let $K \subset \mathbb{S}^3$ be an (a_1, a_2, a_3) pretzel link such that, for each i, $a_i \geq 4$ and the spanning surface F shown in Figure 12 is orientable. As in Section 3.2.1, let M be the manifold produced by taking the 3–fold branched cover of $\mathbb{S}^3$ with K as the branch set and X be the 2–complex produced by gluing the lifts F in M. Then M satisfies the disk-condition and X is a 2–complex that cuts it up into injective handlebodies. As $a_i \geq 4$, the basis bounded by the curves shown in Figure 12 is an 8–waveless basis (Definition 3.10) for K in the handlebody $\overline{\mathbb{S}^3 - S}$. Therefore, all meridian disks in the handlebody $\overline{\mathbb{S}^3 - S}$ intersect K at least eight times. We can produce a waveless minimal system of meridian disks for the handlebody $\overline{\mathbb{S}^3 - f}$ by removing any one of the disks from the basis. The associated graph Γ', as constructed above satisfies the conditions of Lemma 4.32. Thus the 3–fold branched cover of such a pretzel link satisfies the disk-condition and is atoroidal.

References

[1] **I R Aitchison**, **J H Rubinstein**, *Localising Dehn's lemma and the loop theorem in* 3*–manifolds*, Math. Proc. Cambridge Philos. Soc. 137 (2004) 281–292 MR

[2] **M Freedman**, **J Hass**, **P Scott**, *Closed geodesics on surfaces*, Bull. London Math. Soc. 14 (1982) 385–391 MR

[3] **M Freedman**, **J Hass**, **P Scott**, *Least area incompressible surfaces in* 3*–manifolds*, Invent. Math. 71 (1983) 609–642 MR

[4] **M Hall, Jr**, *Coset representations in free groups*, Trans. Amer. Math. Soc. 67 (1949) 421–432 MR

[5] **J Hempel**, 3*–Manifolds*, Ann. of Math. Studies 86, Princeton Univ. Press (1976) MR

[6] **W Jaco**, *Lectures on three-manifold topology*, CBMS Regional Conference Series in Mathematics 43, Amer. Math. Soc., Providence, RI (1980) MR

[7] **W Jaco**, **P B Shalen**, *A new decomposition theorem for irreducible sufficiently-large* 3*–manifolds*, from "Algebraic and geometric topology, part 2" (R J Milgram, editor), Proc. Sympos. Pure Math. XXXII, Amer. Math. Soc., Providence, RI (1978) 71–84 MR

[8] **K Johannson**, *Homotopy equivalences of* 3*–manifolds with boundaries*, Lecture Notes in Mathematics 761, Springer (1979) MR

[9] **S Matveev**, *Algorithmic topology and classification of* 3*–manifolds*, Algorithms and Computation in Mathematics 9, Springer (2003) MR

[10] **U Oertel**, *Closed incompressible surfaces in complements of star links*, Pacific J. Math. 111 (1984) 209–230 MR

[11] **D Rolfsen**, *Knots and links*, Mathematics Lecture Series 7, Publish or Perish, Berkeley, CA (1976) MR

[12] **C P Rourke**, **B J Sanderson**, *Introduction to piecewise-linear topology*, Ergeb. Math. Grenzgeb. 69, Springer (1972) MR

[13] **P Scott**, *There are no fake Seifert fibre spaces with infinite* π_1, Ann. of Math. 117 (1983) 35–70 MR

Northcote, Victoria, Australia

Department of Mathematics and Statistics, The University of Melbourne
Parkville, Victoria, Australia

jadcoffey@gmail.com, joachim@unimelb.edu.au

http://www.ms.unimelb.edu.au/~rubin/

Received: 21 February 2006 Revised: 19 January 2017

Algebraic & Geometric Topology 17 (2017) 3259–3339

Equivariant iterated loop space theory and permutative G–categories

BERTRAND J GUILLOU
J PETER MAY

We set up operadic foundations for equivariant iterated loop space theory. We start by building up from a discussion of the approximation theorem and recognition principle for V–fold loop G–spaces to several avatars of a recognition principle for infinite loop G–spaces. We then explain what genuine permutative G–categories are and, more generally, what E_∞–G–categories are, giving examples showing how they arise. As an application, we prove the equivariant Barratt–Priddy–Quillen theorem as a statement about genuine G–spectra and use it to give a new, categorical proof of the tom Dieck splitting theorem for suspension G–spectra. Other examples are geared towards equivariant algebraic K–theory.

55P42, 55P47, 55P48, 55P91; 18D10, 18D50

Introduction

Let G be a finite group. We will develop equivariant infinite loop space theory in a series of papers. In this introductory one, we focus on the operadic equivariant infinite loop space machine. This is the most topologically grounded machine, as we illustrate by first focusing on its relationship to V–fold deloopings for G–representations V. Genuine permutative G–categories and, more generally, E_∞–G–categories are also defined operadically. They provide the simplest categorical input needed to construct genuine G–spectra from categorical input.

For background, naive G–spectra are just spectra with actions by G. They have their uses, but they are not adequate for serious work in equivariant stable homotopy theory. The naive suspension G–spectra of spheres S^n with trivial G–action are invertible in the naive equivariant stable homotopy category. In contrast, for all real orthogonal G–representations V, the genuine suspension G–spectra of G–spheres S^V are invertible in the genuine equivariant stable homotopy category, where S^V is the one-point compactification of V. Naive G–spectra represent $\mathbb{Z}$–graded cohomology theories, whereas genuine G–spectra represent cohomology theories graded on the real

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3259

representation ring $RO(G)$. The $RO(G)$–grading is essential for Poincaré duality and, surprisingly, for many nonequivariant applications.

The zeroth space $E_0 = \Omega^\infty E$ of a naive Ω–G–spectrum is an infinite loop G–space in the sense that it is equivalent to an n–fold loop G–space $\Omega^n E_n$ for each $n \geq 0$. The zeroth space E_0 of a genuine Ω–G–spectrum E is an infinite loop G–space in the sense that it is equivalent to a V–fold loop G–space $\Omega^V E(V)$ for all real representations V. The essential point of equivariant infinite loop space theory is to construct G–spectra from space or category level data. Such a result is called a recognition principle since it allows us to recognize infinite loop G–spaces when we see them. A functor that constructs G–spectra from G–space or G–category level input is called an equivariant infinite loop space machine.

As we shall see, a recognition principle for naive G–spectra is obtained simply by letting G act in the obvious way on the input data familiar from the nonequivariant theory. One of our main interests is to construct and apply an equivariant infinite loop space machine that constructs genuine G–spectra from categorical input.

A permutative category is a symmetric strictly associative and unital monoidal category, and any symmetric monoidal category is equivalent to a permutative category. The classifying space of a permutative category $\mathscr{A}$ is rarely an infinite loop space, but infinite loop space theory constructs an Ω–spectrum $\mathbb{K}\mathscr{A}$ whose zeroth space is a group completion of the classifying space $B\mathscr{A}$. A naive permutative G–category is a permutative category that is a G–category with equivariant structure data. It is a straightforward adaptation of the nonequivariant theory to construct naive G–spectra $\mathbb{K}\mathscr{A}$ from naive permutative G–categories $\mathscr{A}$ in such a way that $\mathbb{K}_0\mathscr{A}$ is a group completion of $B\mathscr{A}$, meaning that $(\mathbb{K}_0\mathscr{A})^H$ is a nonequivariant group completion of $B(\mathscr{A}^H)$ for all subgroups H of G.

In this paper, we explain what genuine permutative G–categories are and what E_∞–G–categories are, and we explain how to construct a genuine G–spectrum $\mathbb{K}_G\mathscr{A}$ from a genuine permutative G–category $\mathscr{A}$ or, more generally, from an E_∞–G–category $\mathscr{A}$. A genuine G–spectrum has an underlying naive G–spectrum, and the underlying naive G–spectrum of $\mathbb{K}_G\mathscr{A}$ will be $\mathbb{K}\mathscr{A}$. Therefore, we still have the crucial group completion property relating $B\mathscr{A}$ to the zeroth G–space of $\mathbb{K}_G\mathscr{A}$.

We use this theory to show how to construct suspension G–spectra from categorical data, giving a new equivariant version of the classical Barratt–Priddy–Quillen (BPQ) theorem for the construction of the sphere spectrum from symmetric groups. In Guillou, May, Merling and Osorno [13], we shall use this version of the BPQ theorem as input to a proof of the results from equivariant infinite loop space theory that were promised in Guillou and May [10], where we described the category of G–spectra as an easily understood category of spectral presheaves. Here we use this version of the BPQ theorem to give a new categorical proof of the tom Dieck splitting theorem for the fixed-point spectra of suspension G–spectra. The new proof is simpler and gives more precise information than the classical proof by induction up orbit types.

A complementary interest is to understand the geometry of V–fold loop G–spaces. As we shall explain in this paper, these interests lead to quite different perspectives. They are manifested in point-set level distinctions that would be invisible to a more abstract approach. One way of pinpointing these differences is to emphasize the distinction between the role played by E_V–operads for representations V, which are the equivariant generalizations of E_n–operads, and the role played by (genuine) E_∞–operads of G–spaces.

An E_V–space is a G–space with an action by an E_V–operad. We here develop a machine that constructs V–fold loop G–spaces from E_V–spaces. For future perspective, we envision the possibility of an equivariant version of factorization homology in which E_V–operads will govern local structure of G–manifolds in analogy with the role played by E_n–operads in the existing nonequivariant theory. For such a theory, E_∞–operads would be essentially irrelevant.

In contrast, for infinite loop space theory, E_V–operads serve merely as scaffolding used to build a machine that constructs genuine G–spectra from E_∞–G–spaces, which are spaces with an action by some E_∞–operad. The classifying G–spaces of genuine permutative G–categories are examples of E_∞–G–spaces with actions by a particular E_∞–operad $\mathscr{P}_G$, but E_∞–G–spaces with actions by quite different E_∞–operads abound. We concentrate on such an operadic machine in this paper. The machine we concentrate on in the sequels (with Merling and Osorno [31; 12; 13]) makes no use of E_V–operads and does not recognize V–fold loop G–spaces, but it allows a level of categorical power and multiplicative control that is unobtainable with the machine built here.

This paper offers a number of variant perspectives on the topics it studies. We give recognition principles for V–fold loop spaces (Theorem 1.14), for orthogonal G–spectra (Theorem 1.25 and Definition 2.7) and, preserving space level structure invisible in orthogonal G–spectra, for Lewis–May G–spectra (Definition 2.11 and Theorem 2.13). The geometric input data for Theorem 1.14 consists of algebras over the little disks or Steiner operad, $\mathscr{D}_V$ or $\mathscr{K}_V$. For Theorem 1.25, it consists of compatible algebras over the $\mathscr{K}_V$ for all finite-dimensional V.

In both Definitions 2.7 and 2.11, the input data consists of algebras over an E_∞–operad of G–spaces. These algebras may come by applying the classifying-space functor B to algebras over an E_∞–operad of G–categories. The orthogonal spectrum machine and the Lewis–May spectrum machine are shown to be equivalent by comparing them both to a machine landing in the S_G–modules of Elmendorf, Kriz, Mandell and May (EKMM) [7] and Mandell and May [19]. In effect, the machines landing in Lewis–May G–spectra and in S_G–modules provide highly structured fibrant approximations of the machine landing in orthogonal G–spectra. In retrospect, such fibrant approximation is central to nonequivariant calculational understanding, and one can hope that the same will eventually prove true equivariantly.

The variants have alternative and contradictory good features, which become particularly apparent and relevant when specialized to free E_∞–algebras, where they are all viewed as giving variants of the equivariant BPQ theorem. Thinking unstably and geometrically, Theorem 1.21 shows how the machine recognizes V–fold suspensions $\Sigma^V X$ and shows that the recognition is precisely compatible with the evident G–homeomorphisms $\Sigma^V X \wedge \Sigma^W Y \cong \Sigma^{V\oplus W} Y$. Thinking stably and geometrically, Theorems 1.31 and 2.18 show how the machine recognizes orthogonal or Lewis–May suspension G–spectra $\Sigma^\infty_G X$. In both cases, the recognition is precisely compatible with the standard G–isomorphisms $\Sigma^\infty_G X \wedge \Sigma^\infty_G Y \cong \Sigma^\infty_G (X \wedge Y)$. However, the meaning of Σ^∞_G is quite different in the two cases. For orthogonal G–spectra, $\Sigma^\infty_G X$ is cofibrant if G is cofibrant as a G–space, but it is never fibrant. For Lewis–May or EKMM G–spectra, $\Sigma^\infty_G X$ is always fibrant and often bifibrant.

Theorems 6.1 and 9.9 show how the machine recognizes suspension G–spectra from two variant categorical inputs. Here we do *not* have precise compatibility with smash products, a failure that will be rectified with a hefty dose of 2–category theory in the sequel [13], but instead we have structure that allows our new proof of the tom Dieck splitting theorem.

As already mentioned, there are three sequels to this paper. The first [31] develops a new version of the Segal–Shimakawa infinite loop space machine and proves among other things that it is equivalent both to the original Segal–Shimakawa machine and to the machine landing in orthogonal G–spectra that we develop here. That requires a generalization of the present machine from operads to categories of operators, about which we say nothing here. The second [12] gives a multiplicative elaboration of the Segal–Shimakawa machine, starting from space level input. The third [13] gives a more categorically sophisticated machine. It starts with more general categorical input than we deal with here, and it gives new information even nonequivariantly.

Outline We begin with a machine for recognizing iterated equivariant loop spaces in Section 1. All versions of our iterated loop space machine are based on use of the Steiner operads, whose equivariant versions have not previously appeared. We define them and compare them to the little disks operads in Section 1.1. All versions are also based on an approximation theorem, which is explained in Section 1.2. We use a strengthened version due to Rourke and Sanderson [37], and that allows us to obtain slightly stronger versions of the recognition principle than might be expected. The compatibility with smash products of the geometric versions of the recognition principle is based on pairings between Steiner operads that are defined in Section 1.4; the relevant definition of a pairing is recalled in Appendix A. The promised variants of the recognition principle starting from space level input data are given in Sections 1.3, 1.5, 2.2, and 2.3.

Section 2 gives our machines for recognizing infinite loop G–spaces. After recalling the notion of E_∞–G–operad and giving some examples in Section 2.1, the orthogonal and Lewis–May machines are defined and compared in Sections 2.2–2.4. Examples of E_∞–G–spaces are given in Section 2.5. General properties that must hold for any equivariant infinite loop space machine are described in Section 2.6. A recognition principle for naive G–spectra, with G not necessarily finite, is given in Section 2.7. An interesting detail there shows how to use the recognition principle to construct change of universe functors on the space level. The proof uses a double bar construction described in Appendix B.

The recognition principle starting from categorical input is given in Section 4.5. It is preceded by preliminaries about equivariant universal bundles and equivariant E_∞–operads in Section 3 and by a discussion of operadic definitions of naive and genuine permutative G–categories in Section 4. In the brief and parenthetical Section 4.4, we point out how these ideas and our prequel [11] with Merling specialize to give a starting point for equivariant algebraic K–theory; see also Dress and Kuku [6], Fiedorowicz, Hauschild and May [9], Kuku [17] and Merling [33]. We give an alternative and equivalent starting point in the case of G–rings R in Section 8.2.

We give a precise description of the G–fixed E_∞–categories of free $\mathscr{P}_G$–categories in Section 5. This is a precursor of our first categorical version of the BPQ theorem, which we prove in Section 6.1, and of the tom Dieck splitting theorem for suspension G–spectra, which we reprove in Section 6.2.

Changing focus, in Sections 7 and 8 we give three interrelated examples of E_∞–G–operads, denoted by $\mathscr{V}_G$, $\mathscr{V}_G^\times$, and $\mathscr{W}_G$, and give examples of their algebras. This approach to examples is more intuitive than the approach based on genuine permutative G–categories, and it has some technical advantages. It is new and illuminating even nonequivariantly. It gives a more intuitive categorical hold on the BPQ theorem than does the treatment starting from genuine permutative G–categories, as we explain in Section 9.3. It also gives a new starting point for multiplicative infinite loop space theory, both equivariantly and nonequivariantly, but that is work in progress.

Notational preliminaries A dichotomy between Hom objects with G–actions and Hom objects of equivariant morphisms, often denoted using a G in front, is omnipresent. We start with an underlying category $\mathscr{V}$. A G–object X in $\mathscr{V}$ can be defined to be a group homomorphism $G \to \mathrm{Aut}\, X$. We have the category $\mathscr{V}_G$ of G–objects in $\mathscr{V}$ and all morphisms in $\mathscr{V}$ between them, with G acting by conjugation. We denote the morphism objects of $\mathscr{V}_G$ simply by $\mathscr{V}(X,Y)$.[1] We also have the category $G\mathscr{V}$ of G–objects in $\mathscr{V}$ and G–maps in $\mathscr{V}$. Since objects are fixed by G, we see that $G\mathscr{V}$ is in fact the G–fixed category $(\mathscr{V}_G)^G$, although we shall not use that notation. Thus the Hom object $G\mathscr{V}(X,Y)$ in $\mathscr{V}$ of G–morphisms between G–objects X and Y is the fixed-point object $\mathscr{V}(X,Y)^G$.

One frequently used choice of $\mathscr{V}$ is $\mathscr{U}$, the category of unbased (compactly generated) spaces. We let $\mathscr{T}$ denote the category of based spaces. We assume once and for all that the basepoints $*$ of all given based G–spaces X (or spaces X when $G = e$) are nondegenerate. This means that $* \to X$ is a G–cofibration (satisfies the G–HEP). It follows that $* \to X^H$ is a cofibration for all $H \subset G$.

By an equivalence $f\colon X \to Y$ of G–spaces, we understand a G–map whose fixed-point maps $f^H\colon X^H \to Y^H$ are weak homotopy equivalences for all subgroups H of G. When X and Y have the homotopy types of G–CW complexes, such an f is a G–homotopy equivalence.

By a topological category $\mathscr{C}$, we understand a category internal to $\mathscr{U}$; thus it has an object space and a morphism space such that the structural maps I, S, T, and C are continuous. This is more structure than a topologically enriched category, which would have a discrete space of objects. We also have the based variant of categories internal to $\mathscr{T}$, but $\mathscr{U}$ will be the default.

We let $\mathscr{C}at$ denote the category of (small) topological categories. As above, starting from $\mathscr{C}at$, we obtain the concomitant categories $G\mathscr{C}at$ and $\mathscr{C}at_G$ of G–categories. A G–category is a topological category equipped with an action of G through natural isomorphisms. This is the same structure as a category internal to $G\mathscr{U}$. Similarly, a based G–category is a category internal to $G\mathscr{T}$. That is, an action of G on a topological category $\mathscr{C}$ is given by actions of G on both the object space and the morphism space such that I, S, T, and C are G–maps. In particular, G can and often will act nontrivially on the space of objects. That may be unfamiliar (as the referee noted), but in many of our examples it is essential for proper behavior on passage to H–fixed subcategories for $H \subset G$.

[1] In [19] and elsewhere, we used the notation $\mathscr{V}_G(X,Y)$ instead of $\mathscr{V}(X,Y)$, but some readers found that misleadingly analogous to $\mathrm{Hom}_G(X,Y)$.

For brevity of notation, we shall often but not always write $|-|$ for the composite classifying-space functor $B = |N-|$ from topological categories through simplicial spaces to spaces. It works equally well to construct G–spaces from topological G–categories. We assume that the reader is familiar with operads (as originally defined in May [21]) and especially with the fact that operads can be defined in any symmetric monoidal category $\mathscr{V}$. Brief modernized expositions are given in May [27; 28]. Since it is product preserving, the functor $|-|$ takes operads in $\mathscr{C}at$ or in $G\mathscr{C}at$ to operads in $\mathscr{U}$ or in $G\mathscr{U}$, and it takes algebras over an operad $\mathscr{C}$ in $\mathscr{C}at$ or in $G\mathscr{C}at$ to algebras over the operad $|\mathscr{C}|$ in $\mathscr{U}$ or in $G\mathscr{U}$.

To avoid proliferation of letters, we shall write $\mathbb{O}_G$ for the monad on based G–categories constructed from an operad $\mathcal{O}_G$ of G–categories. We shall write $\boldsymbol{O}_G$ for the monad on based G–spaces constructed from the operad $|\mathscr{O}_G|$ of G–spaces. More generally, for an operad $\mathscr{C}_G$ of unbased G–spaces, we write $\boldsymbol{C}_G$ for the associated monad on based G–spaces.

Acknowledgements Guillou thanks Nat Stapleton for very helpful discussions leading to the rediscovery of the operad $\mathscr{P}_G$, which was in fact first defined, but not used, by Shimakawa [43, Remark, page 255]. May thanks Mona Merling for many conversations and questions that helped clarify ideas. We both thank the referee for helpful suggestions. We also thank Anna Marie Bohmann and Angélica Osorno for pointing out a mistake in the original version. That led to a reworking of this paper and to much of the work in the sequels [31; 12; 13]. It also led to the long delay in the publication of this paper, which is entirely due to the authors and not at all to the referee or editors. We thank them for their patience. Guillou was supported by Simons Collaboration Grant 282316.

1 E_V–operads and V–fold loop G–spaces

In this geometrically focused chapter, we first define E_V–operads and give two examples. We then relate E_V–spaces to V–fold loop G–spaces via the equivariant approximation theorem and recognition principle. The approximation theorem shows how to approximate "free" V–fold loop G–spaces $\Omega^V \Sigma^V X$ in terms of free algebras $\mathbb{D}_V X$ or $\mathbb{K}_V X$ over the E_V–operad $\mathscr{D}_V$ or $\mathscr{K}_V$. The recognition principle shows how to construct V–fold loop spaces from E_V–algebras. We elaborate multiplicatively by showing how machine-built pairings relate to evident pairings between iterated loop G–spaces. We then give a geometric version of a concrete spacewise infinite loop G–space machine that does not use E_∞–operads and is new even nonequivariantly. This gives a geometric precursor of the BPQ theorem that relates well to smash products. As already noted, we envision that the theory here can provide the local data for an as yet undeveloped equivariant factorization homology theory.

1.1 The little disks and Steiner operads

Definition 1.1 Let $D(V)$ be the open unit disk in V. A little V–disk is a map $d\colon D(V) \to D(V)$ of the form $d(v) = rv + v_0$ for some $r \in [0,1)$ and some $v_0 \in V$; $c(d) = v_0$ is the center point of d and r is the radius. For $g \in G$, we have $(gd)(v) = rv + gv_0$. Define $\mathscr{D}_V(j)$ to be the G–space of (ordered) j–tuples of little V–disks whose images have empty pairwise intersections. With the evident structure maps determined by disjoint union and composites of little disks, the $\mathscr{D}_V(j)$ form an operad $\mathscr{D}_V$, called the little disks operad.

For a G–space V, let $F(V,j) \subset V^j$ be the configuration space of (ordered) j–tuples of distinct points of V, with G acting by restriction of the diagonal action on V^j. By convention, $F(V,0)$ is a point, the empty 0–tuple of points in V. We are interested in the special case when V is a real representation of G, by which we understand an orthogonal action of G on a real inner product space. In contrast to the nonequivariant case, very little is known about the (Bredon) homology and cohomology of the G–spaces $F(V,j)$, but we have the following result.

Lemma 1.2 *There is a $(G\times\Sigma_j)$–homotopy equivalence $\mathscr{D}_V(j) \to F(V,j)$ for each $j \geq 0$.*

Proof Choose a decreasing rescaling homeomorphism $\zeta\colon [0,\infty) \to [0,1)$ and also denote by ζ the rescaling homeomorphism $V \to D(V)$ that sends v to $\zeta(|v|/|v|)v$, where $D(V)$ is the open unit disc in V. Then ζ induces a rescaling homeomorphism $\zeta\colon F(V,j) \to F(D(V),j)$. Define a map $c\colon \mathscr{D}_V(j) \to F(D(V),j)$ by sending little disks to their center points. For a point $\underline{v} = (v_1,\dots,v_j)$ in $F(D(V),j)$, define

$$\delta(\underline{v}) = \tfrac{1}{2}\min\bigl\{|v_i - v_j| \bigm| i \neq j\bigr\}.$$

Define $s\colon F(D(V),j) \to \mathscr{D}_V(j)$ by $s(\underline{v}) = (d_1,\dots,d_j)$, where $d_j(v) = \delta(\underline{v})v + v_i$. Then s and c are $(G\times\Sigma_j)$–maps, $c \circ s = \mathrm{id}$, and there is a $(G\times\Sigma_j)$–homotopy $h\colon s\circ c \simeq \mathrm{id}$. If $\underline{d} = (d_1,\dots,d_j) \in \mathscr{D}(j)$, where $d_i(v) = r_i v + v_i$, then $c(\underline{d}) = \underline{v}$ and $h(\underline{d},t)$ has i^{th} little V–disk $d_i(t)$ given by $d_i(t)(v) = ((1-t)\delta(\underline{v}) + tr_i)v + v_i$. □

The following definition is the equivariant generalization of the usual definition of an E_n–operad. We say that a map of operads of G–spaces is a weak equivalence if its j^{th} map is a weak $(G\times\Sigma_j)$–equivalence.

Definition 1.3 An operad $\mathscr{C}_G$ of G–spaces is an E_V–operad if there is a chain of weak equivalences of operads connecting $\mathscr{C}_G$ to $\mathscr{D}_V$.

Of course, we could use any operad weakly equivalent to $\mathscr{D}_V$ as a reference operad in the definition. As explained in [30, Section 3], for inclusions $V \subset W$ of inner product spaces, there is no map of operads $\mathscr{D}_V \to \mathscr{D}_W$ that is compatible with suspension, so that use of the little disks operads is inappropriate for iterated loop space theory. The Steiner operads remedy the defect and will be used in [31] to compare the operadic and Segalic equivariant infinite loop space machines. Their equivariant definition is little different from their nonequivariant definition given in [30], following Steiner [46].

Definition 1.4 Let E_V be the space of embeddings $V \to V$, with G acting by conjugation, and let $\mathrm{Emb}_V(j) \subset E_V^j$ be the G–subspace of (ordered) j–tuples of embeddings with pairwise disjoint images. Regard such a j–tuple as an embedding ${}^jV \to V$, where ${}^jV$ denotes the disjoint union of j copies of V (where ${}^0V$ is empty). The element id in $\mathrm{Emb}_V(1)$ is the identity embedding, the group Σ_j acts on $\mathrm{Emb}_V(j)$ by permuting embeddings, and the structure maps

$$\gamma\colon \mathrm{Emb}_V(k) \times \mathrm{Emb}_V(j_1) \times \cdots \times \mathrm{Emb}_V(j_k) \to \mathrm{Emb}_V(j_1 + \cdots + j_k)$$

are defined by composition and disjoint union in the evident way [30, Section 3]. This gives an operad Emb_V of G–spaces.

Define $R_V \subset E_V = \mathrm{Emb}_V(1)$ to be the G–subspace of distance-reducing embeddings $f\colon V \to V$. This means that $|f(v) - f(w)| \le |v - w|$ for all $v, w \in V$. Define a Steiner path to be a map $h\colon I \to R_V$ such that $h(1) = \mathrm{id}$ and let P_V be the G–space of Steiner paths, with action of G induced by the action on R_V. Define $\pi\colon P_V \to R_V$ by evaluation at 0; that is, $\pi(h) = h(0)$.

Define $\mathscr{K}_V(j)$ to be the G–space of (ordered) j–tuples $(h_1, \dots, h_j)$ of Steiner paths such that the $\pi(h_i)$ have disjoint images. The element id in $\mathscr{K}_V(1)$ is the constant path at the identity embedding, the group Σ_j acts on $\mathscr{K}_V(j)$ by permutations, and the structure maps γ are defined pointwise in the same way as those of Emb_V. This gives an operad of G–spaces, and application of π to Steiner paths gives a map of operads $\pi\colon \mathscr{K}_V \to \mathrm{Emb}_V$. Evaluation of embeddings at $0 \in V$ gives center point $(G\times\Sigma_j)$–maps $c\colon \mathrm{Emb}_V(j) \to F(V, j)$.

The Steiner operads $\mathscr{K}_V$ are reduced, meaning that $\mathscr{K}_V(0)$ is a point, and $\mathscr{K}_0$ is the trivial operad with $\mathscr{K}_0(1) = \mathrm{id}$ and $\mathscr{K}_0(j) = \varnothing$ for $j > 1$. By pullback along π, any space with an action by Emb_V inherits an action by $\mathscr{K}_V$. As in [21, Section 5], [24, Section VII.2], or [30, Section 3], Emb_V acts naturally on $\Omega^V X$ for based G–spaces X.

Proposition 1.5 [46] *There is a weak equivalence of operads $\iota\colon \mathscr{D}_V \to \mathscr{K}_V$.*

Proof For each j, we have a composite $(G\times\Sigma_j)$–map

$$c\circ\pi\colon \mathscr{K}_V(j)\to \mathrm{Emb}_V(j)\to F(V,j).$$

Steiner's nonequivariant proof that $c\circ\pi$ is a Σ_j–homotopy equivalence applies to prove that it is a $(G\times\Sigma_j)$–homotopy equivalence. The argument is a clever and nontrivial variant on the proof above for $\mathscr{D}_V$, but for us the essential point is that it uses the metric on V and the contractibility of I and V in such a way that the construction is clearly G–equivariant.

For a little disk $d(v)=rv+v_0$, define a path of little disks from d to the identity map of $D(V)$ by sending $s\in I$ to the little disk

$$d(s)(v)=(s-rs+r)v+(1-s)v_0.$$

Conjugating d by the rescaling ζ of Lemma 1.2 gives a distance-reducing embedding $\zeta^{-1}d\zeta\colon V\to V$, and conjugating paths pointwise gives an embedding ι of $\mathscr{D}_V$ as a suboperad of $\mathscr{K}_V$. Composing the inverse $(G\times\Sigma_j)$–homotopy equivalence $F(V,j)\to\mathscr{D}_V(j)$ with $\iota\colon\mathscr{D}_V(j)\to\mathscr{K}_V(j)$ gives an inverse $(G\times\Sigma_j)$–homotopy equivalence to $c\circ\pi$, by Steiner's proof, and it follows that ι is a $(G\times\Sigma_j)$–homotopy equivalence. □

Again, one key advantage of the Steiner operads over the little disks operads is that, for an inclusion $V\subset W$ of G–inner product spaces, there is an induced inclusion $\mathscr{K}_V\to\mathscr{K}_W$ of G–operads such that the map

$$\Omega^V\eta\colon \Omega^V X\to\Omega^V\Omega^{W-V}\Sigma^{W-V}X\cong\Omega^W\Sigma^{W-V}X$$

is a map of $\mathscr{K}_V$–spaces for any G–space X. Here $W-V$ is the orthogonal complement of V in W. If $f\colon V\to V$ is a distance-reducing embedding, then $f\oplus\mathrm{id}_{W-V}\colon W\to W$ is also distance reducing, and this construction induces the inclusion.

1.2 The approximation theorem

Write $\boldsymbol{K}_V$ for the monad on based G–spaces associated to the operad $\mathscr{K}_V$. For a G–space X, we have $\boldsymbol{K}_V X=\coprod\mathscr{K}_V(j)\times_{\Sigma_j}X^j/(\sim)$. If $\sigma_i\colon\mathscr{K}_V(j)\to\mathscr{K}_V(j-1)$ deletes the i^{th} Steiner path and $s_i\colon X^{j-1}\to X^j$ inserts the basepoint in the i^{th} position, then $(\sigma^i k,y)\sim(k,s_i y)$ for $k\in\mathscr{K}_V(j)$ and $y\in X^{j-1}$. The monad $\boldsymbol{D}_V$ arising from the operad $\mathscr{D}_V$ is defined the same way.

The unit $\eta\colon\mathrm{Id}\to\Omega^V\Sigma^V$ of the monad $\Omega^V\Sigma^V$ and the action θ of $\boldsymbol{K}_V$ on the G–spaces $\Omega^V\Sigma^V X$ induce a composite natural map

$$\alpha_V\colon \boldsymbol{K}_V X\xrightarrow{\boldsymbol{K}_V\eta}\boldsymbol{K}_V\Omega^V\Sigma^V X\xrightarrow{\theta}\Omega^V\Sigma^V X,$$

and $\alpha_V\colon \boldsymbol{K}_V \to \Omega^V\Sigma^V$ is a map of monads whose adjoint defines a right action of $\boldsymbol{K}_V$ on the functor Σ^V, just as in [21]. The restriction to $\boldsymbol{D}_V$ gives the corresponding map $\alpha_V\colon \boldsymbol{D}_V X \to \Omega^V\Sigma^V X$.

The heart of the operadic recognition principle is the approximation theorem that says that α_V is a group completion. However, already nonequivariantly, we have two variants of what it means for a map $X \to Y$ to be a group completion. Recall that Hopf spaces are spaces with a product with a two-sided unit element up to homotopy.

Definition 1.6 A Hopf space Y is grouplike if $\pi_0(Y)$ is a group. Let X and Y be homotopy associative and commutative Hopf spaces, where Y is grouplike, and let $f\colon X \to Y$ be a Hopf map. Then f is a group completion if $f_*\colon \pi_0(X) \to \pi_0(Y)$ is the Grothendieck construction converting a commutative monoid to an abelian group and if, for any field of coefficients k, the map of commutative k–algebras $H_*(X)[\pi_0(X)^{-1}] \to H_*(Y)$ induced by f_* is an isomorphism.

The second version of group completion drops the commutativity assumption and lives in the setting of A_∞–spaces. For us, an A_∞–space will mean a space with an action of the Steiner operad $\mathscr{K}_{\mathbb{R}}$. An A_∞–map will mean either a map homotopic to a map of $\mathscr{K}_{\mathbb{R}}$–spaces or the homotopy inverse of a map of $\mathscr{K}_{\mathbb{R}}$–spaces that is an underlying homotopy equivalence.

Definition 1.7 An A_∞–map $f\colon X \to Y$ of $\boldsymbol{K}_{\mathbb{R}}$–spaces is a weak group completion if it is equivalent under a chain of A_∞–maps to the natural map $\eta\colon M \to \Omega BM$ for some topological monoid M.

The following classical result has several proofs; see [23, Section 15] for discussion in slightly greater generality.

Theorem 1.8 *If a topological monoid M is homotopy commutative, then the natural map $\eta\colon M \to \Omega BM$ is a group completion.*

Returning to our equivariant context, we have the following definition.

Definition 1.9 A Hopf G–space Y is grouplike if each $\pi_0(Y^H)$ is a group. Let X and Y be homotopy associative and commutative Hopf G–spaces, where Y is grouplike, and let $f\colon X \to Y$ be a Hopf G–map. Then f is a group completion if $f^H\colon X^H \to Y^H$ is a group completion for all subgroups H of G.

For the equivariant notion of weak group completion, note that if X is a $\boldsymbol{K}_{\mathbb{R}}$–$G$–space and $H \subset G$ is a subgroup, then X^H inherits an action of $\boldsymbol{K}_{\mathbb{R}}$.

Definition 1.10 A map $f\colon X \to Y$ of $\boldsymbol{K}_{\mathbb{R}}$–$G$–spaces is a weak group completion if f^H is a weak group completion for all H. By Theorem 1.8, f is then a group completion if X and Y are homotopy commutative.

In the weak case we require no compatibility between the monoids $M(H) \simeq X^H$ as H varies. Recall that we understand equivalences of G–spaces to mean maps that induce (weak) equivalences on passage to fixed points and observe that a group completion is an equivalence if X is grouplike, for example if X is G–connected in the sense that each X^H is (path) connected.

Theorem 1.11 (the approximation theorem) *Let V be a representation of G. If X is G–connected, then $\alpha_V\colon \boldsymbol{K}_V X \to \Omega^V \Sigma^V X$ is an equivalence. If V contains a copy of the trivial representation $\mathbb{R}$, then α_V is a weak group completion. Therefore, if V contains a copy of $\mathbb{R}^2$, then α_V is a group completion.*

We shall not give a proof, only a commentary on the existing proofs. The group completion version was first proven by Hauschild in his unpublished Habilitationschrift [14], but the shorter published version [15] restricts to the case $X = S^0$, remarking that the proof in the general case is essentially the same. Assuming that V contains $\mathbb{R}^\infty$ and not just $\mathbb{R}^2$, Caruso and Waner [3, Theorem 1.18] gave a shorter proof in a paper that concentrated on compact Lie groups G, rather than just finite groups.

Nonequivariantly, there is a proof by direct calculation due to Fred Cohen [18] and a geometric proof due to Segal [42]. Starting from Segal's proof, Rourke and Sanderson [38; 39; 40] gave an elegant proof using their "compression theorem". Following up a suggestion of May, they generalized that proof to give the stated version of the theorem in [37]. However, their notation is quite different from ours. They never work equivariantly and focus instead on G–fixed-point spaces. They use the notation $\Omega^V \Sigma^V X$ for the G–fixed-point space $(\Omega^V \Sigma^V X)^G$. One can replace G by a subgroup H in their proof, and it works just as well.

All known proofs are manifold-theoretic in nature and start with the G–space $\boldsymbol{F}_V X$ of (unordered) configurations of points in V with labels in X. More precisely, $\boldsymbol{F}_V X = \coprod F(V, j) \times_{\Sigma_j} X^j/(\sim)$ is defined in the same way as $\boldsymbol{K}_V X$. In the notation of [37], their $C_V X$ is our $(\boldsymbol{F}_V X)^G$. They work with little disks, and their $C_V^o X$ is our $(\boldsymbol{D}_V X)^G$. Their map j_V is the restriction to $\boldsymbol{F}_V(X)^G$ of our map α_V^G.

Translated to our notation, [37, Theorem 1] proves the first statement of Theorem 1.11, taking X to be G–connected; here there are no Hopf G–space structures in sight. When $W = V \oplus \mathbb{R}$, Rourke and Sanderson observe that $(\boldsymbol{D}_W X)^G$ is equivalent to a monoid, and their [37, Theorem 2] proves that its classifying space is weak homotopy equivalent to $(\Omega^V \Sigma^W X)^G$. The approximation theorem as stated follows by applying Ω as in [37, Corollary 1].

1.3 The recognition principle for V–fold loop spaces

We explain how K_V–spaces, which are based spaces with an action of $\mathscr{K}_V$, give rise to V–fold loop spaces. For fixed V, we can work equally well with $\mathscr{D}_V$. For compatibility as V varies, $\mathscr{K}_V$ is required. The two-sided monadic bar construction is described in [21; 30] and works exactly the same way equivariantly as nonequivariantly.[2] The adjoint of α_V gives a right action $\widetilde{\alpha}_V\colon \Sigma^V \boldsymbol{K}_V \to \Sigma^V$ of the monad $\boldsymbol{K}_V$ on the functor Σ^V.

Definition 1.12 Let Y be a $\mathscr{K}_V$–space. We define

$$\mathbb{E}_V Y = B(\Sigma^V, \boldsymbol{K}_V, Y).$$

We have the diagram of $\mathscr{K}_V$–spaces and $\mathscr{K}_V$–maps

$$Y \xleftarrow{\varepsilon} B(\boldsymbol{K}_V, \boldsymbol{K}_V, Y) \xrightarrow{\bar{\alpha}_V} B(\Omega^V \Sigma^V, \boldsymbol{K}_V, Y) \xrightarrow{\zeta} \Omega^V B(\Sigma^V, \boldsymbol{K}_V, Y), \tag{1.13}$$

where $\bar{\alpha}_V = B(\alpha_V, \mathrm{id}, \mathrm{id})$ and ζ will be defined in the following sketch proof, which is based on arguments in [4; 21; 22].

Theorem 1.14 (from $\mathscr{K}_V$–spaces to V–fold loop spaces) *The following statements hold relating a $\mathscr{K}_V$–space Y to its **V–fold delooping** $\mathbb{E}_V Y$:*

(i) *The map ε is a G–homotopy equivalence with a natural homotopy inverse ν.*

(ii) *The map $\bar{\alpha}_V$ is an equivalence when Y is G–connected and is a weak group completion when $V \supset \mathbb{R}$.*

(iii) *The map ζ is an equivalence.*

Therefore, the composite

$$\xi = \zeta \circ \bar{\alpha}_V \circ \nu\colon Y \to \Omega^V \mathbb{E}_V Y \tag{1.15}$$

is an equivalence if Y is G–connected, a weak group completion if $V \supset \mathbb{R}$, and a group completion if $V \supset \mathbb{R}^2$.

Proof The proof of (i) uses an "extra-degeneracy argument" explained in [21, Proposition 9.8]; note that the homotopy equivalence ν is not a $\mathscr{K}_V$–map. For (ii), it is shown nonequivariantly in [22, Theorem 2.3], that $\bar{\alpha}_V$ is an equivalence when Y is connected

[2] In particular, Reedy cofibrancy (or properness) works the same way; see [31].

and is a group completion when $V=\mathbb{R}^n$ with $n\geq 2$. We use Theorem 1.11 to improve on that equivariantly. Geometric realization of simplicial G–spaces commutes with passage to H–fixed points, so we can work nonequivariantly, one fixed-point space at a time. If Y is G–connected, each $(\mathbb{K}^q_V Y)^H$ is connected, hence $\overline{\alpha}^H$ is the realization of a levelwise equivalence of simplicial spaces and hence an equivalence.

Now assume $V\supset\mathbb{R}$ and let $\mathscr{K}=\mathscr{K}_{\mathbb{R}}$, with associated monad $\boldsymbol{K}$. We then have an inclusion of the nonequivariant A_∞–operad $\mathscr{K}$ in $\mathscr{K}_V$ and can regard Y and each $(\mathbb{K}^q_V Y)^H$ as a $\mathscr{K}$–space. From here we combine arguments from [21, Section 13] and the proof of [22, Theorem 2.3] with the Rourke–Sanderson proof of the approximation theorem. Let $\mathscr{M}$ be the associativity operad that defines monoids; we have a weak equivalence of (G–fixed) operads $\delta\colon \mathscr{K}\to\mathscr{M}$. For a $\mathscr{K}$–space X, we define a topological monoid $\Lambda(X)=B(\boldsymbol{M},\boldsymbol{K},X)$, where the monad $\boldsymbol{M}$ is a $\boldsymbol{K}$–functor via δ. We have a zigzag

$$X \xleftarrow{\varepsilon} B(\boldsymbol{K},\boldsymbol{K},X) \xrightarrow{\overline{\delta}} B(\boldsymbol{M},\boldsymbol{K},X)=\Lambda X$$

in which ε is a $\mathscr{K}$–map and a G–homotopy equivalence and $\overline{\delta}=B(\delta,\mathrm{id},\mathrm{id})$ is an equivalence. Define $\Gamma(X)=\Omega B\Lambda(X)$ and $\gamma=\eta\circ\overline{\delta}\colon B(\boldsymbol{K},\boldsymbol{K},X)\to\Gamma X$. We view γ as a natural choice of a weak group completion. Moreover, γ is an equivalence if X is grouplike. If $f\colon X\to Y$ is a weak group completion between $\mathscr{K}$–spaces, then Γf is an equivalence. To see this, note that by the definition of weak group completion, we may assume without loss of generality that f is the map $\eta\colon M\to\Omega BM$ for some topological monoid M. It suffices to show that $B\Lambda(\eta)\colon B\Lambda M\to B\Lambda(\Omega BM)$ is an equivalence. This follows from [47, Proposition 3.9 and Theorem 3.11].

Now consider the following commutative diagram:

$$\begin{tikzcd}
Y & B(\boldsymbol{K}_V,\boldsymbol{K}_V,Y) \arrow[l, "\varepsilon"'] \arrow[r, "\overline{\alpha}_V"] & B(\Omega^V\Sigma^V,\boldsymbol{K}_V,Y) \\
B(\boldsymbol{K},\boldsymbol{K},Y) \arrow[u, "\varepsilon"] \arrow[d, "\gamma"'] & B(B(\boldsymbol{K},\boldsymbol{K},\boldsymbol{K}_V),\boldsymbol{K}_V,Y) \arrow[l, "\varepsilon"'] \arrow[r, "\overline{\alpha}_V"] \arrow[u, "\varepsilon"'] \arrow[d, "\overline{\gamma}"] & B(B(\boldsymbol{K},\boldsymbol{K},\Omega^V\Sigma^V),\boldsymbol{K}_V,Y) \arrow[u, "\varepsilon"'] \arrow[d, "\overline{\gamma}"] \\
\Gamma Y & B(\Gamma\boldsymbol{K}_V,\boldsymbol{K}_V,Y) \arrow[l, "\varepsilon"'] \arrow[r, "\overline{\Gamma\alpha_V}"] & B(\Gamma\Omega^V\Sigma^V,\boldsymbol{K}_V,Y)
\end{tikzcd}$$

The maps ε are G–homotopy equivalences, hence the middle map $\overline{\gamma}=B(\gamma,\mathrm{id},\mathrm{id})$ is a weak group completion since γ is so. The right map $\overline{\gamma}$ and the bottom map $\overline{\Gamma\alpha_V}$ are equivalences since realization preserves levelwise equivalences. Therefore, $\overline{\alpha}_V$ is a weak group completion.

In (iii), ζ is an instance of the natural G–map $\zeta\colon |\Omega^V K| \to \Omega^V |K|$ for simplicial based G–spaces K; suspensions commute with realization, and the adjoint of ζ is the evident evaluation G–map $\Sigma^V|\Omega^V K| \cong |\Sigma^V \Omega^V K| \to |K|$. The proof of (iii) is due to Hauschild [14] and appears in [4, pages 495–496]. We will not repeat the argument, which reduces the proof to the nonequivariant case treated in [21, Section 12]. The main equivariant input that allows the reduction is the fact if $S(V)$ is the unit sphere in V, then the space $\mathrm{Map}_H(S(V), K_n)$ of H–maps is connected, where $K_n = \Sigma^V \boldsymbol{K}_V^n Y$ is the G–space of n–simplices of the simplicial G–space $B_*(\Sigma^V, \boldsymbol{K}_V, Y)$. This holds since K_n^J is $(\dim(V^J)-1)$–connected for each subgroup $J \subset G$, while $S(V)$ regarded as an H–CW complex only has cells of type $H/J \times e^n$ where $n < \dim(V^J)$. □

Remark 1.16 Equivariant homotopy theory often admits varying generalizations of nonequivariant theorems. A very different and very interesting equivariant recognition principle was proven by Salvatore and Wahl [41].

1.4 The pairing $(\mathscr{K}_V, \mathscr{K}_W) \to \mathscr{K}_{V\oplus W}$ and the recognition principle

The general notion of a pairing of operads is recalled in Appendix A. In [21, Proposition 8.3], a pairing

$$\boxtimes\colon \boldsymbol{C}_m X \wedge \boldsymbol{C}_n Y \to \boldsymbol{C}_{m+n}(X \wedge Y)$$

is defined for based spaces X and Y, where $\boldsymbol{C}_n$ denotes the monad on based spaces induced from the little n–cubes operad $\mathscr{C}_n$. Implicitly, it comes from a pairing of operads $\boxtimes\colon (\mathscr{C}_m, \mathscr{C}_n) \to \mathscr{C}_{m+n}$. The Steiner operad analogue appears in [25, page 337], and we recall it here.

Proposition 1.17 *For finite-dimensional real inner product G–spaces V and W, there is a unital, associative, and commutative system of pairings*

$$\boxtimes\colon (\mathscr{K}_V, \mathscr{K}_W) \to \mathscr{K}_{V\oplus W}$$

of Steiner operads of G–spaces.

Proof The required maps

$$\boxtimes\colon \mathscr{K}_V(j) \times \mathscr{K}_W(k) \to \mathscr{K}_{V\oplus W}(jk)$$

are given by $(c \otimes d) = e$, where, writing $c = (f_1, \dots, f_j)$ and $d = (g_1, \dots, g_k)$, e is the jk–tuple of Steiner paths

$$(f_q, g_r)\colon I \to R_V \times R_W \subset R_{V\oplus W}$$

for $1 \le q \le j$ and $1 \le r \le k$, ordered lexicographically. The formulas required in Definition A.1 are easily verified, as we illustrate in Example A.4.

The pairing is unital in the sense that $\boxtimes\colon \mathscr{K}_V(j) \cong \mathscr{K}_0(1) \times \mathscr{K}_V(j) \to \mathscr{K}_V(j)$ is the identity map. It is associative in the sense that the following diagram commutes for a triple (V, W, Z) of inner product G–spaces and a triple (i, j, k):

$$\begin{tikzcd}
\mathscr{K}_V(i) \times \mathscr{K}_W(j) \times \mathscr{K}_Z(k) \arrow[r, "\boxtimes \times \mathrm{id}"] \arrow[d, "\mathrm{id} \times \boxtimes"'] & \mathscr{K}_{V\oplus W}(ij) \times \mathscr{K}_Z(k) \arrow[d, "\boxtimes"] \\
\mathscr{K}_V(i) \times \mathscr{K}_{W\oplus Z}(jk) \arrow[r, "\boxtimes"'] & \mathscr{K}_{V\oplus W\oplus Z}(ijk)
\end{tikzcd}$$

It is commutative in the sense that the following diagram commutes:

$$\begin{tikzcd}
\mathscr{K}_V(j) \times \mathscr{K}_W(k) \arrow[r, "\boxtimes"] \arrow[d, "t"'] & \mathscr{K}_{V\oplus W}(jk) \arrow[d, "\tau(j,k)"] \\
\mathscr{K}_W(k) \times \mathscr{K}_V(j) \arrow[r, "\boxtimes"'] & \mathscr{K}_{W\oplus V}(kj)
\end{tikzcd}$$

Here t is the interchange map and $\tau(j,k)$ is determined in an evident way by the interchange map for V and W and the permutation $\tau(j,k)$ of jk–letters. □

Passing to monads as in Proposition A.3 below, we obtain a unital, associative, and commutative system of pairings

$$\boxtimes\colon \boldsymbol{K}_V X \wedge \boldsymbol{K}_W Y \to \boldsymbol{K}_{V\oplus W}(X \wedge Y). \tag{1.18}$$

For the unit property, when $V = 0$ the map $\boxtimes\colon X \wedge \boldsymbol{K}_W Y \to \boldsymbol{K}(X \wedge Y)$ is induced by the maps $X \times Y^j \to (X \times Y)^j$ obtained from the diagonal map on X and shuffling. We have the following key observation. Its analogue for the little cubes operads is [21, Proposition 8.3].

Lemma 1.19 *The following diagram commutes:*

$$\begin{tikzcd}
\boldsymbol{K}_V X \wedge \boldsymbol{K}_W Y \arrow[r, "\boxtimes"] \arrow[d, "\alpha_V \wedge \alpha_W"'] & \boldsymbol{K}_{V\oplus W}(X \wedge Y) \arrow[d, "\alpha_{V\oplus W}"] \\
\Omega^V \Sigma^V X \wedge \Omega^W \Sigma^W Y \arrow[r, "\wedge"'] & \Omega^{V\oplus W} \Sigma^{V\oplus W}(X \wedge Y)
\end{tikzcd}$$

The notion of a pairing of a $\mathscr{K}_V$–space X and a $\mathscr{K}_W$–space Y to a $\mathscr{K}_{V\oplus W}$–space Z is defined in Definition A.2, and we have the following recognition principle for pairings.

Note that smashing maps out of spheres gives a natural map

$$\Omega^V X \wedge \Omega^W Y \to \Omega^{V\oplus W}(X \wedge Y).$$

Proposition 1.20 *A pairing $f\colon X \wedge Y \to Z$ of a $\mathscr{K}_V$–space X and a $\mathscr{K}_W$–space Y to a $\mathscr{K}_{V\oplus W}$–space Z induces a G–map*

$$\mathbb{E}f\colon \mathbb{E}_V X \wedge \mathbb{E}_W Y \to \mathbb{E}_{V\oplus W} Z$$

such that the following diagram commutes:

$$\begin{tikzcd}
X \wedge Y \arrow[r, "\xi\wedge\xi"] \arrow[d, "f"'] & \Omega^V \mathbb{E}_V X \wedge \Omega^W \mathbb{E}_W Y \arrow[r] & \Omega^{V\oplus W}(\mathbb{E}_V X \wedge \mathbb{E}_W Y) \arrow[d, "\mathbb{E}f"] \\
Z \arrow[rr, "\xi"'] & & \Omega^{V\oplus W}\mathbb{E}_{V\oplus W} Z
\end{tikzcd}$$

Proof By convention, $\boldsymbol{K}_V^0 = \mathrm{Id}$ for any V. Starting at $q=0$ with the identity map on $X \wedge Y$, the map $\boxtimes$ inductively determines a pairing $\boxtimes^q$ for all q, namely the composite

$$\boldsymbol{K}_V^q X \wedge \boldsymbol{K}_W^q Y \xrightarrow{\boxtimes} \boldsymbol{K}_{V\oplus W}(\boldsymbol{K}_V^{q-1} X \wedge \boldsymbol{K}_W^{q-1} Y) \xrightarrow{\boldsymbol{K}_{V\oplus W}\boxtimes^{q-1}} \boldsymbol{K}_{V\oplus W}^q (X \wedge Y).$$

The map $\mathbb{E}f$ is the geometric realization of a map of simplicial topological spaces that is given on q–simplices by

$$\Sigma^V \boldsymbol{K}_V^q X \wedge \Sigma^W \boldsymbol{K}_W^q Y \cong \Sigma^{V\oplus W}(\boldsymbol{K}_V^q X \wedge \boldsymbol{K}_W^q Y) \xrightarrow{\Sigma^{V\oplus W}\boxtimes^q} \Sigma^{V\oplus W} \boldsymbol{K}_{V\oplus W}^q (X \wedge Y).$$

Commutation with face and degeneracy operators follows from Proposition A.3. The diagram in the statement commutes by a diagram chase from Lemma 1.19, Definition A.2, and the description of ξ given in (1.15). □

We have an unstable precursor of the BPQ theorem.

Theorem 1.21 (the BPQ theorem for V–fold suspensions) *For based G–spaces X, there is a natural G–homotopy equivalence*

$$\omega\colon \Sigma^V X \to \mathbb{E}_V \boldsymbol{K}_V X$$

such that the following diagram commutes for based G–spaces X and Y:

$$\begin{tikzcd}
\Sigma^V X \wedge \Sigma^W Y \arrow[r, "\omega\wedge\omega"] \arrow[d, "\cong"'] & \mathbb{E}_V \boldsymbol{K}_V X \wedge \mathbb{E}_W \boldsymbol{K}_W Y \arrow[d, "\mathbb{E}(\boxtimes)"] \\
\Sigma^{V\oplus W}(X \wedge Y) \arrow[r, "\omega"'] & \mathbb{E}_{V\oplus W} \boldsymbol{K}_{V\oplus W}(X \wedge Y)
\end{tikzcd}$$

Therefore, $\mathbb{E}(\boxtimes)$ is an equivalence.

Proof Since $\mathbb{E}_V \boldsymbol{K}_V X = B(\Sigma^V, \boldsymbol{K}_V, \boldsymbol{K}_V X)$, another extra-degeneracy argument explained in [21, Proposition 9.8] gives the natural homotopy equivalence ω. For the diagram, it suffices to prove commutativity of the adjoint diagram, which features two adjoint maps $X \wedge Y \to \Omega^{V \oplus W}(-)$. These maps are equal by inspection of definitions. □

1.5 The geometric recognition principle for orthogonal G–spectra

As in [10], we let $G\mathscr{S}$ denote the category of orthogonal G–spectra. Briefly, these start with $\mathscr{I}_G$–spaces E, which are continuous functors $E\colon \mathscr{I}_G \to \mathscr{T}_G$, where $\mathscr{I}_G$ is the category of finite-dimensional G–inner product spaces and linear isometric isomorphisms, with G acting by conjugation on morphism spaces $\mathscr{I}_G(V, V')$. The continuous G–maps $E\colon \mathscr{I}_G(V, V') \to \mathscr{T}_G(E(V), E(V'))$ can be specified via adjoint evaluation G–maps $\mathscr{I}_G(V, V')_+ \wedge E(V) \to E(V')$.

An $\mathscr{I}_G$–space E is an orthogonal G–spectrum if there exist structure G–maps $\Sigma^W E(V) \to E(V \oplus W)$ that give a natural transformation $E \barwedge S_G \to E \circ \oplus$ of functors $\mathscr{I}_G \times \mathscr{I}_G \to \mathscr{T}_G$, where $S_G = \{S^V\}$ is the sphere G–spectrum, $\barwedge$ is the external smash product specified by $(D \barwedge E)(V, W) = D(V) \wedge E(W)$ for $\mathscr{I}_G$–spaces D and E, and $\oplus\colon \mathscr{I}_G \times \mathscr{I}_G \to \mathscr{I}_G$ is the direct sum of G–inner product spaces functor. See [19, Section II.2] for details.

Definition 1.22 We define a continuous G–functor $\mathscr{K}_*$ from $\mathscr{I}_G$ to G–operads. It takes a G–inner product space V to the Steiner operad $\mathscr{K}_V$. Linear isometric isomorphisms $i\colon V \to V'$ act by conjugation of embeddings to send R_V to $R_{V'}$. The action extends pointwise to Steiner paths and then applies one at a time to j–tuples of Steiner paths to give G–maps $\mathscr{K}_V(j)$ to $\mathscr{K}_{V'}(j)$. Compatibility with the operad structure is immediate. Composing with the functor that sends the operad $\mathscr{K}_V$ to the associated monad $\boldsymbol{K}_V$ on based G–spaces gives a functor $\boldsymbol{K}$ from $\mathscr{I}_G$ to the category of monads in the category of $\mathscr{I}_G$–spaces. In more detail, for an $\mathscr{I}_G$–space $\mathscr{X}$ with V^{th} space $\mathscr{X}(V)$, we have based evaluation G–maps

$$\mathscr{I}(V, V')_+ \wedge \boldsymbol{K}_V \mathscr{X}(V) \to \boldsymbol{K}_{V'} \mathscr{X}(V').$$

Using the diagonal action of $\mathscr{I}_G(V, V')$, we obtain G–maps

$$\begin{array}{c} \mathscr{I}_G(V, V') \times \mathscr{K}_V(k) \times \mathscr{K}_V(j_1) \times \cdots \times \mathscr{K}_V(j_k) \times \mathscr{X}(V) \\ \Big\downarrow \\ \mathscr{K}_{V'}(k) \times \mathscr{K}_{V'}(j_1) \times \cdots \times \mathscr{K}_{V'}(j_k) \times \mathscr{X}(V'), \end{array}$$

and these give evaluation G–maps

$$\mathscr{I}_G(V, V')_+ \wedge \boldsymbol{K}_V \boldsymbol{K}_V \mathscr{X}(V) \to \boldsymbol{K}_{V'} \boldsymbol{K}_{V'} \mathscr{X}(V').$$

The product and unit maps are compatible with these maps in the sense that the following diagrams commute, where the unlabeled arrows are evaluation G–maps:

(1.23)
$$\begin{tikzcd}
\mathscr{I}_G(V,V')_+ \wedge \mathscr{X}(V) \arrow[r, "\pi"] \arrow[d, "\mathrm{id}\wedge\eta"'] & \mathscr{X}(V') \arrow[d, "\eta"] \\
\mathscr{I}_G(V,V')_+ \wedge \boldsymbol{K}_V\mathscr{X}(V) \arrow[r] & \boldsymbol{K}_{V'}\mathscr{X}(V')
\end{tikzcd}$$
$$\begin{tikzcd}
\mathscr{I}_G(V,V')_+ \wedge \boldsymbol{K}_V\boldsymbol{K}_V\mathscr{X}(V) \arrow[r] \arrow[d, "\mathrm{id}\wedge\mu"'] & \boldsymbol{K}_{V'}\boldsymbol{K}_{V'}\mathscr{X}(V) \arrow[d, "\mu"] \\
\mathscr{I}_G(V,V')_+ \wedge \boldsymbol{K}_V\mathscr{X} \arrow[r] & \boldsymbol{K}_{V'}\mathscr{X}(V')
\end{tikzcd}$$

Note that we can regard based G–spaces X as constant $\mathscr{I}_G$–spaces, $X(V)=X$; the evaluation G–maps $\mathscr{I}_G(V,V')_+\wedge X\to X$ are then the projections.

Definition 1.24 Define a $\mathscr{K}_*$–G–space $\mathscr{Y}$ to be an $\mathscr{I}_G$–space $\mathscr{Y}$ with a structure of $\mathscr{K}_V$–algebra on $\mathscr{Y}(V)$ for each V together with G–maps $i\colon \mathscr{Y}(V)\to\mathscr{Y}(V\oplus W)$ such that the following diagrams commute, where the θ are monad action maps:

$$\begin{tikzcd}
\mathscr{I}_G(V,V')_+ \wedge \boldsymbol{K}_V\mathscr{Y}(V) \arrow[r] \arrow[d, "\mathrm{id}\wedge\theta"'] & \boldsymbol{K}_{V'}\mathscr{Y}(V') \arrow[d, "\theta"] \\
\mathscr{I}_G(V,V')_+ \wedge \mathscr{Y}(V) \arrow[r] & \mathscr{Y}(V')
\end{tikzcd}$$

In the second diagram, we identify $S^V\wedge S^W$ with $S^{V\oplus W}$:

$$\begin{tikzcd}
(\mathscr{I}_G(V,V')\times\mathscr{I}_G(W,W'))_+ \wedge \mathscr{Y}(V)\wedge S^V\wedge S^W \arrow[r] \arrow[d, "\oplus\wedge i\wedge\mathrm{id}"'] & \mathscr{Y}(V')\wedge S^{V'}\wedge S^{W'} \arrow[d, "i\wedge\mathrm{id}"] \\
(\mathscr{I}_G(V\oplus W,V'\oplus W'))_+ \wedge \mathscr{Y}(V\oplus W)\wedge S^{V\oplus W} \arrow[r] & \mathscr{Y}(V'\oplus W')\wedge S^{V'\oplus W'}
\end{tikzcd}$$

The first diagram says that θ is a map of $\mathscr{I}_G$–spaces and, ignoring the sphere coordinates, the second diagram says that $i\colon \mathscr{Y}\circ\pi_1\Rightarrow\mathscr{Y}\circ\oplus$ is a natural transformation of functors $\mathscr{I}_G\times\mathscr{I}_G\to\mathscr{T}_G$.

Theorem 1.25 (from $\mathscr{K}_*$–G–spaces to orthogonal G–spectra) *For a $\mathscr{K}_*$–G–space $\mathscr{Y}$, the based G–spaces $\mathbb{E}_V\mathscr{Y}(V)$ and the based G–maps*

$$\Sigma^W\mathbb{E}_V\mathscr{Y}(V)\to\mathbb{E}_{V\oplus W}\mathscr{Y}(V\oplus W)$$

determined by $i\colon \mathscr{Y}\circ\pi_1\Rightarrow\mathscr{Y}\circ\oplus$ specify an orthogonal G–spectrum $\mathbb{E}_G^{\mathrm{geo}}\mathscr{Y}$.

Proof Regarding the $\mathscr{I}_G(V,V')$ as constant simplicial G–spaces, we see by diagram chases from the definitions that the data of the previous definitions determine G–maps

$$\mathscr{I}_G(V,V')_+ \wedge \Sigma^V \boldsymbol{K}^q_V \mathscr{Y}(V) \to \Sigma^{V'} \boldsymbol{K}^q_{V'} \mathscr{Y}(V')$$

and

$$\Sigma^W \Sigma^V \boldsymbol{K}^q_V \mathscr{Y}(V) \to \Sigma^{V\oplus W} \boldsymbol{K}^q_{V\oplus W} \mathscr{Y}(V \oplus W).$$

On passage to geometric realization, these give the required $\mathscr{I}_G$–space $\mathbb{E}^{\mathrm{geo}}_G \mathscr{Y}$ and the required natural transformation $\mathbb{E}^{\mathrm{geo}}_G \mathscr{Y} \barwedge S_G \to \mathbb{E}^{\mathrm{geo}}_G \mathscr{Y} \circ \oplus$. □

Of course, the recognition principle of (1.13) and Theorem 1.14 applies to describe the relationship between the G–spaces $\mathscr{Y}(V)$ and $\Omega^V(\mathbb{E}^{\mathrm{geo}}_G \mathscr{Y})(V)$. The recognition principle for pairings also adapts directly.

Definition 1.26 Let $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$ be $\mathscr{K}_*$–G–spaces. A pairing

$$f\colon \mathscr{X} \barwedge \mathscr{Y} \to \mathscr{Z} \circ \oplus$$

is a natural transformation of continuous functors $\mathscr{I}_G \times \mathscr{I}_G \to \mathscr{T}_G$ such that each $f\colon \mathscr{X}(V) \wedge \mathscr{Y}(W) \to \mathscr{Z}(V \oplus W)$ is a pairing as in Definition A.2 and the following diagram commutes for all U, V, W:

$$\begin{tikzcd}
\mathscr{X}(U) \wedge \mathscr{Y}(V) \arrow[rr, "\mathrm{id} \wedge i"] \arrow[d, "i \wedge \mathrm{id}"'] \arrow[dr, "f"] & & \mathscr{X}(U) \wedge \mathscr{Y}(V \oplus W) \arrow[dd, "f"] \\
 & \mathscr{Z}(U \oplus V) \arrow[dr, "i"] & \\
\mathscr{X}(U \oplus W) \wedge \mathscr{Y}(V) \arrow[r, "f"'] & \mathscr{Z}(U \oplus W \oplus V) \arrow[r, "\mathscr{Z}(\mathrm{id} \oplus t)"'] & \mathscr{Z}(U \oplus V \oplus W)
\end{tikzcd} \tag{1.27}$$

This diagram expresses that the three composite natural transformations of functors $\mathscr{I}^3_G \to \mathscr{T}_G$ in sight agree.

The smash product of orthogonal G–spectra is obtained by first applying Day convolution to the external smash product $\barwedge$ and then coequalizing the action of the sphere G–spectrum on the two variables. See [19, Section II.3] for details.

Proposition 1.28 *A pairing $f\colon \mathscr{X} \barwedge \mathscr{Y} \to \mathscr{Z} \circ \oplus$ of $\mathscr{K}_*$–G–spaces induces a map*

$$\mathbb{E}^{\mathrm{geo}}_G f\colon \mathbb{E}^{\mathrm{geo}}_G \mathscr{X} \wedge \mathbb{E}^{\mathrm{geo}}_G \mathscr{Y} \to \mathbb{E}^{\mathrm{geo}}_G \mathscr{Z}$$

of orthogonal G–spectra that is given levelwise by specialization of Proposition 1.20.

Proof The definition of a pairing immediately implies that f induces an external pairing

$$\mathbb{E}_G^{\mathrm{geo}}\mathscr{X} \barwedge \mathbb{E}_G^{\mathrm{geo}}\mathscr{Y} \to \mathbb{E}_G^{\mathrm{geo}}\mathscr{Z} \circ \oplus,$$

and the diagram (1.27) ensures that the resulting map from the Day convolution to $\mathbb{E}_G^{\mathrm{geo}}\mathscr{Z}$ factors through the coequalizer defining $\mathbb{E}_G^{\mathrm{geo}}\mathscr{X} \wedge \mathbb{E}_G^{\mathrm{geo}}\mathscr{Y}$. $\square$

The suspension G–spectrum $\Sigma_G^\infty X$ of a based G–space X is given by the G–spaces $\Sigma^V X$; its structure maps are the evident identifications $\Sigma^W \Sigma^V X \cong \Sigma^{V \oplus W} X$. The unstable BPQ theorem of Theorem 1.21 leads to the following "geometric" version of the BPQ theorem.

Definition 1.29 For a based G–space X, define $\boldsymbol{K}_* X$ to be the $\mathscr{K}_*$–G–space given by the $\mathscr{K}_V$–spaces $\boldsymbol{K}_V X$ and the maps $i\colon \boldsymbol{K}_V X \to \boldsymbol{K}_{V\oplus W} X$ induced by the map of operads $\mathscr{K}_V \to \mathscr{K}_{V\oplus W}$ obtained by sending embeddings $e\colon V \to V$ to $e \times \mathrm{id}\colon V \times W \to V \times W$.

It is easily verified that $\boldsymbol{K}_* X$ is a $\mathscr{K}_*$–G–space and the pairings $\boxtimes$ of (1.18) prescribe pairings

$$\boxtimes\colon \mathbb{K}_* X \barwedge \mathbb{K}_* Y \to \mathbb{K}_*(X \wedge Y) \circ \oplus. \tag{1.30}$$

Theorem 1.31 (the geometric BPQ theorem for orthogonal suspension G–spectra) *For based G–spaces X, there is a natural equivalence*

$$\omega\colon \Sigma_G^\infty X \to \mathbb{E}_G^{\mathrm{geo}} \boldsymbol{K}_* X$$

such that the following diagram commutes for based G–spaces X and Y:

$$\begin{tikzcd}
\Sigma_G^\infty X \wedge \Sigma_G^\infty Y \arrow[r, "\omega\wedge\omega"] \arrow[d, "\cong"'] & \mathbb{E}_G^{\mathrm{geo}} \boldsymbol{K}_* X \wedge \mathbb{E}_G^{\mathrm{geo}} \boldsymbol{K}_* Y \arrow[d, "\mathbb{E}_G^{\mathrm{geo}}(\boxtimes)"] \\
\Sigma_G^\infty (X \wedge Y) \arrow[r, "\omega"'] & \mathbb{E}_G^{\mathrm{geo}} \boldsymbol{K}_*(X \wedge Y)
\end{tikzcd}$$

Proof The levelwise equivalence follows from Theorem 1.21. For the diagram, the functor Σ_G^∞ is left adjoint to the 0^{th} G–space functor, and inspection of definitions shows that the adjoint diagram starting with $X \wedge Y$ commutes. $\square$

1.6 A configuration space model for free $\mathscr{K}_V$–spaces

The free $\mathscr{K}_V$–spaces $\mathbb{K}_V X$ can be modeled more geometrically by configuration spaces. To explain this, we first record the nonequivariant analogue in terms of the little cubes operads, since that is relevant folklore which is not in the literature.

Consider the little n–cubes operads $\mathscr{C}_n$ and their associated monads $\boldsymbol{C}_n$. Let $J = (0,1)$ be the interior of I. We have the configuration spaces $F(J^n, j)$ of j–tuples of distinct points in J^n. Sending little n–cubes $c\colon J^n \to J^n$ to their center points $c\big(\frac{1}{2},\dots,\frac{1}{2}\big)$ gives a Σ_n–homotopy equivalence $f\colon \mathscr{C}_n(j) \to F(J^n, j)$.

For based spaces X, we construct spaces $\boldsymbol{F}_n X$ by replacing $\mathscr{C}_n(j)$ by $F(J^n, j)$ in the construction of $\boldsymbol{C}_n X$ as the quotient of $\coprod \mathscr{C}_n(j) \times_{\Sigma_j} X^j$ by basepoint identifications; we now use the evident omit a point projections $F(j,n) \to F(j,n-1)$ rather than the analogous maps $\mathscr{C}_n(j) \to \mathscr{C}_n(j-1)$. The maps f induce a homotopy equivalence

$$f\colon \boldsymbol{C}_n X \to \boldsymbol{F}_n X.$$

That much has been known since [21].

The folklore observation is that although the $F(J^n, j)$ do not form an operad, $\mathscr{C}_n$ acts on $\boldsymbol{F}_n X$ in such a way that f is a map of $\mathscr{C}_n$–spaces. Indeed, we can evaluate little n–cubes $J^n \to J^n$ on points of J^n to obtain maps

$$\mathscr{C}_n(j) \times F(J^n, j) \to F(J^n, j),$$

and any reader of [21] will see how to proceed from there. Moreover, we have pairings

$$\boxtimes\colon \boldsymbol{F}_m X \wedge \boldsymbol{F}_n Y \to \boldsymbol{F}_{m+n}(X \wedge Y)$$

defined as in Definition A.2 and Proposition A.3, starting from the maps

$$F(J^n, j) \times F(J^n, k) \to F(J^n, jk)$$

that send (x, y), where $x = (x_1, \dots, x_j)$ and $y = (y_1, \dots, y_k)$, to the set of pairs (x_q, y_r), for $1 \le q \le j$ and $1 \le r \le k$, ordered lexicographically.

Nonequivariantly, we put this together to obtain an analogue of Theorem 1.31, using the evident variant of the geometric recognition principle that is obtained from the operads $\mathscr{C}_n$ as n varies. Here it is more natural to use symmetric spectra rather than orthogonal spectra, since it is natural to deal with sequences rather than inner product spaces. The relationship between the little cubes operads and symmetric spectra is explained in [19, Section I.8], and we leave details of the relevant retooling of the previous subsections to the interested reader.

Theorem 1.32 (the configuration space BPQ theorem for symmetric spectra) *For based spaces X, there is a natural equivalence*

$$\omega\colon \Sigma^\infty X \to \mathbb{E}^{\mathrm{geo}} \boldsymbol{F}_* X$$

such that the following diagram commutes for based spaces X and Y:

$$\begin{tikzcd}
\Sigma^\infty X \wedge \Sigma^\infty Y \arrow[r, "\omega\wedge\omega"] \arrow[d, "\cong"'] & \mathbb{E}^{\mathrm{geo}} \boldsymbol{F}_* X \wedge \mathbb{E}^{\mathrm{geo}} \boldsymbol{F}_* Y \arrow[d, "\mathbb{E}^{\mathrm{geo}}_G(\boxtimes)"] \\
\Sigma^\infty (X \wedge Y) \arrow[r, "\omega"'] & \mathbb{E}^{\mathrm{geo}} \boldsymbol{F}_* (X \wedge Y)
\end{tikzcd}$$

For fixed V, the discussion generalizes equivariantly to relate $\boldsymbol{D}_V X$ or $\boldsymbol{K}_V X$ to $\boldsymbol{F}_V X$ for based G–spaces X. In the case of $\boldsymbol{K}_V X$, we use the time-0 projections from Steiner paths to embeddings $V \to V$ and the centerpoint map from $\mathrm{Emb}_V(j)$ to $F(V, j)$. Letting V vary, we obtain the following equivariant version of Theorem 1.32.

Theorem 1.33 (the configuration space BPQ theorem for orthogonal G–spectra) *For based G–spaces X, there is a natural equivalence*

$$\omega\colon \Sigma^\infty_G X \to \mathbb{E}^{\mathrm{geo}}_G \boldsymbol{F}_* X$$

such that the following diagram commutes for based G–spaces X and Y:

$$\begin{tikzcd}
\Sigma^\infty_G X \wedge \Sigma^\infty_G Y \arrow[r, "\omega\wedge\omega"] \arrow[d, "\cong"'] & \mathbb{E}^{\mathrm{geo}}_G \boldsymbol{F}_* X \wedge \mathbb{E}^{\mathrm{geo}}_G \boldsymbol{F}_* Y \arrow[d, "\mathbb{E}^{\mathrm{geo}}_G(\boxtimes)"] \\
\Sigma^\infty_G (X \wedge Y) \arrow[r, "\omega"'] & \mathbb{E}^{\mathrm{geo}}_G \boldsymbol{F}_* (X \wedge Y)
\end{tikzcd}$$

2 The recognition principle for infinite loop G–spaces

The equivariant recognition principle shows how to recognize (genuine) G–spectra in terms of category or space level information. It comes in various versions. We shall give two modernized variants of the machine from [21], differing in their choice of the output category of G–spectra. In contrast with the previous section, we are now concerned with infinite loop space machines with input given by E_∞–G–spaces (or G–categories) defined over any (genuine) E_∞–operad. A G–spectrum E is connective if the negative homotopy groups of each of its fixed point spectra E^H are zero, and all infinite loop space machines take values in connective G–spectra.

As in [10], we let $\mathscr{S}$, $\mathscr{S}p$, and $\mathscr{Z}$ denote the categories of orthogonal spectra [20], Lewis–May spectra [18], and EKMM S–modules [7]. Similarly, we let $G\mathscr{S}$, $G\mathscr{S}p$

and $G\mathscr{Z}$ denote the corresponding categories of genuine G–spectra from [19], [18], and again [19]. We start with a machine that lands in $G\mathscr{S}$. It is related to but different from the geometric machine of the previous section, and it is the choice preferred in [10] and in the sequels [31; 12; 13]. The sphere G–spectrum S_G in $G\mathscr{S}$ is cofibrant, and so are the suspension G–spectra $\Sigma_G X$ of cofibrant based G–spaces X. We then give the variant machine that lands in $G\mathscr{S}p$ or $G\mathscr{Z}$, where every object is fibrant, and give a comparison that illuminates homotopical properties of the first machine via its comparison with the second.

2.1 Equivariant E_∞–operads

Since operads make sense in any symmetric monoidal category, we have operads of categories, spaces, G–categories, and G–spaces. Operads in $G\mathscr{U}$ were first used in [18, Chapter VII]. Although we are only interested in finite groups G in this paper, the following definition makes sense for any topological group G and is of interest in at least the generality of compact Lie groups.

Definition 2.1 An E_∞–operad $\mathscr{C}_G$ of G–spaces is an operad in the cartesian monoidal category $G\mathscr{U}$ such that $\mathscr{C}_G(0)$ is a contractible G–space and the $(G\times\Sigma_j)$–space $\mathscr{C}_G(j)$ is a universal principal (G,Σ_j)–bundle for each $j\geq 1$. Equivalently, for a subgroup Λ of $G\times\Sigma_j$, the Λ–fixed-point space $\mathscr{C}_G(j)^\Lambda$ is contractible if $\Lambda\cap\Sigma_j=\{e\}$ and is empty otherwise. We say that $\mathscr{C}_G$ is reduced if $\mathscr{C}_G(0)$ is a point.

As is usual in equivariant bundle theory, we think of G as acting from the left and Σ_j as acting from the right on the spaces $\mathscr{C}_G(j)$. These actions must commute and so define an action of $G\times\Sigma_j$. We shall say nothing more about equivariant bundle theory except to note the following parallel. In [21], an operad $\mathscr{C}$ of spaces was defined to be an E_∞–operad if $\mathscr{C}(j)$ is a free contractible Σ_j–space. Effectively, $\mathscr{C}(j)$ is then a universal principal Σ_j–bundle. If we regard each $\mathscr{C}(j)$ as a G–trivial G–space, such an operad is called a naive E_∞–operad of G–spaces. Analogously, we have defined genuine E_∞–operads by requiring the $\mathscr{C}_G(j)$ to be universal principal (G,Σ_j)–bundles. That dictates the appropriate homotopical properties of the $\mathscr{C}_G(j)$, and it is only those homotopical properties and not their bundle-theoretic consequences that concern us in the theory of operads. The bundle theory implicitly tells us which homotopical properties are relevant to equivariant infinite loop space theory. Our default is that E_∞–operads are understood to be genuine unless otherwise specified.

We give two well-known examples. Recall that a complete G–universe U is a G–inner product space that contains countably many copies of each irreducible representation of G; a canonical choice is the sum of countably many copies of the regular representation ρ_G.

Example 2.2 (the Steiner operad $\mathscr{K}_U$) Inclusions $V \subset W$ induce inclusions of operads $\mathscr{K}_V \to \mathscr{K}_W$. Let $\mathscr{K}_U$ be the union over $V \subset U$ of the operads $\mathscr{K}_V$, where U is a complete G–universe.[3] This is the infinite Steiner operad of G–spaces. It is an E_∞–operad since Σ_j–acts freely on $\mathscr{K}_U(j)$ and $\mathscr{K}_U(j)^\Lambda$ is contractible if $\Lambda \subset G \times \Sigma_j$ and $\Lambda \cap \Sigma_j = e$. Indeed, such a Λ is isomorphic to a subgroup H of G via the projection $G \times \Sigma_j \to G$, and if we let H act on U through the isomorphism, then U is a complete H–universe and U^H is isomorphic to $\mathbb{R}^\infty$. Therefore, by the proof of Proposition 1.5, $\mathscr{K}_U(j)^\Lambda$ is equivalent to the configuration space $F(\mathbb{R}^\infty, j)$, which is contractible.

Example 2.3 (linear isometries operad) The equivariant linear isometries operad $\mathscr{L}_U$ was first used in [18, Section VII.1] and is defined just as nonequivariantly (eg [30, Section 2]). The $(G \times \Sigma_j)$–space $\mathscr{L}_U(j)$ is the space of linear isometries $U^j \to U$, with G acting by conjugation, and $\mathscr{L}_U$ is an E_∞–operad of G–spaces if U is a complete G–universe. Indeed, Σ_j acts freely on $\mathscr{L}_U(j)$ and $\mathscr{L}_U(j)^\Lambda$ is contractible if $\Lambda \subset G \times \Sigma_j$ and $\Lambda \cap \Sigma_j = e$. If $\Lambda \cong H$ and H acts on U through the isomorphism, then U is a complete H–universe and $\mathscr{L}_U(j)^H$ is isomorphic to the space of H–linear isometries $U^j \to U$. The usual argument that $\mathscr{L}(j)$ is contractible (eg [24, Lemma I.1.3]) adapts to prove that this space is contractible.

We define E_∞–operads in G–categories in Section 3.3 and give examples in Section 4.2 and Section 7.

Remark 2.4 We will encounter one naturally occurring operad that is not reduced. When an operad $\mathscr{C}$ acts on a space X via maps θ_i and we choose points $c_i \in \mathscr{C}(i)$, we have a map $\theta_0\colon \mathscr{C}(0) \to X$ and the relation

$$\theta_2(c_2; \theta_0(c_0), \theta_1(c_1, x)) = \theta_1(\gamma(c_2; c_0, c_1), x)$$

for $x \in X$. When the $\mathscr{C}(i)$ are connected, this says that $\theta_0(c_0)$ is a unit element for the product determined by c_2. Reduced operads give a single unit element. The original definition [21, 1.1] required operads to be reduced.

Lemma 2.5 *Let $\mathscr{C}_G$ be an E_∞–operad of G–spaces and define $\mathscr{C} = (\mathscr{C}_G)^G$. Then $\mathscr{C}$ is an E_∞–operad of spaces. If Y is a $\mathscr{C}_G$–space, then Y^G is a $\mathscr{C}$–space.*

Proof $(\mathscr{C}_G)^G$ is an operad since the fixed-point functor commutes with products, and it is an E_∞–operad since the space $\mathscr{C}_G(j)^G$ is contractible and Σ_j–free. □

[3] We denoted the nonequivariant version as $\mathscr{C}$ in [30], but we prefer the notation $\mathscr{K}_U$ here.

2.2 The infinite loop space machine: orthogonal G–spectrum version

In brief, we have a functor $\mathbb{E}_G = \mathbb{E}_G^{\mathscr{S}}$ that assigns an orthogonal G–spectrum $\mathbb{E}_G Y$ to a G–space Y with an action by some chosen E_∞–operad $\mathscr{C}_G$ of G–spaces. We want to start with $\mathscr{C}_G$–algebras and still exploit the Steiner operads, and we use the product of operads trick recalled in Section 2.3 to allow this; compare [30, Section 9]. For simplicity of notation, define $\mathscr{C}_V = \mathscr{C}_G \times \mathscr{K}_V$. We use the following observation.

Lemma 2.6 *If $\mathscr{C}_G$ is an E_∞–operad of G–spaces, then the projection*

$$\mathscr{C}_V(j) = \mathscr{C}_G(j) \times \mathscr{K}_V(j) \to \mathscr{K}_V(j)$$

is a $(G\times\Sigma_j)$–equivalence for each j.

Proof We must show that for each subgroup $\Lambda \subseteq G \times \Sigma_j$, the induced map on fixed points

$$\mathscr{C}_G(j)^\Lambda \times \mathscr{K}_V(j)^\Lambda \to \mathscr{K}_V(j)^\Lambda$$

is an equivalence. If $\Lambda \cap \{e\} \times \Sigma_j = \{e\}$, then $\mathscr{C}(j)^\Lambda \simeq *$, so the projection is an equivalence. If Λ contains a nonidentity permutation, then the fixed points on both sides are empty. Both sides are trivial if $j = 0$. □

We view $\mathscr{C}_G$–spaces as $\mathscr{C}_V$–spaces for all V via the projections $\mathscr{C}_V \to \mathscr{C}_G$, and $\mathscr{C}_V$ acts on V–fold loop spaces via its projection to $\mathscr{K}_V$. Write $\boldsymbol{C}_V$ for the monad on based G–spaces associated to the operad $\mathscr{C}_V$. The categories of $\mathscr{C}_V$–spaces and $\boldsymbol{C}_V$–algebras are isomorphic. As in the V–fold delooping argument, the unit $\eta\colon \mathrm{Id} \to \Omega^V\Sigma^V$ of the monad $\Omega^V\Sigma^V$ and the action θ of $\boldsymbol{C}_V$ on the G–spaces $\Omega^V\Sigma^V X$ induce a composite natural map

$$\alpha_V\colon \boldsymbol{C}_V X \xrightarrow{\boldsymbol{C}_V\eta} \boldsymbol{C}_V\Omega^V\Sigma^V X \xrightarrow{\theta} \Omega^V\Sigma^V X,$$

and $\alpha_V\colon \boldsymbol{C}_V \to \Omega^V\Sigma^V$ is a map of monads whose adjoint defines a right action of $\boldsymbol{C}_V$ on the functor Σ^V.

Definition 2.7 (from $\mathscr{C}_G$–spaces to orthogonal G–spectra) Let Y be a $\mathscr{C}_G$–space. We define an orthogonal G–spectrum $\mathbb{E}_G Y$, which we denote by $\mathbb{E}_G^{\mathscr{S}} Y$ when necessary for clarity. Let

$$\mathbb{E}_G Y(V) = B(\Sigma^V, \boldsymbol{C}_V, Y).$$

Using the action of isometric isomorphisms on the $\mathscr{K}_V$ and Σ^V, as in the previous section but starting with Y regarded as a constant $\mathscr{I}_G$–functor, as we can since its action by $\mathscr{C}_G$ is independent of V, this defines an $\mathscr{I}_G$–space. The structure G–map

$$\sigma\colon \Sigma^W\mathbb{E}_G Y(V) \to \mathbb{E}_G Y(V \oplus W)$$

is the composite

$$\Sigma^W B(\Sigma^V, \boldsymbol{C}_V, Y) \cong B(\Sigma^{V\oplus W}, \boldsymbol{C}_V, Y) \to B(\Sigma^{V\oplus W}, \boldsymbol{C}_{V\oplus W}, Y).$$

obtained by commuting Σ^W with geometric realization and using the map of monads $\boldsymbol{C}_V \to \boldsymbol{C}_{V\oplus W}$ induced by the inclusion $i\colon \mathscr{K}_V \to \mathscr{K}_{V\oplus W}$.

Just as in (1.13), we have the diagram of $\mathscr{C}_V$–spaces and $\mathscr{C}_V$–maps

(2.8) $$Y \xleftarrow{\varepsilon} B(\boldsymbol{C}_V, \boldsymbol{C}_V, Y) \xrightarrow{\overline{\alpha}} B(\Omega^V\Sigma^V, \boldsymbol{C}_V, Y) \xrightarrow{\zeta} \Omega^V B(\Sigma^V, \boldsymbol{C}_V, Y),$$

where $\overline{\alpha} = B(\alpha, \mathrm{id}, \mathrm{id})$. Theorem 1.14 applies verbatim, with the same proof. We let $\xi_V = \zeta \circ \overline{\alpha} \circ \nu$, where ν is the canonical homotopy inverse to ε. Then the following diagram commutes, where $\widetilde{\sigma}$ is adjoint to σ:

$$\begin{tikzcd} & Y \arrow[dl, "\xi_V"'] \arrow[dr, "\xi_{V\oplus W}"] & \\ \Omega^V \mathbb{E}_G Y(V) \arrow[rr, "\Omega^V\widetilde{\sigma}"'] & & \Omega^{V\oplus W}\mathbb{E}_G Y(V\oplus W) \end{tikzcd}$$

Therefore, $\Omega^V\widetilde{\sigma}$ is a weak equivalence if $V \supset \mathbb{R}$. If we replace $\mathbb{E}_G Y$ by a fibrant approximation $\boldsymbol{R}\mathbb{E}_G Y$, there results a group completion $\xi\colon Y \to (\boldsymbol{R}\mathbb{E}_G Y)_0$. We shall shortly use the category $\mathscr{S}p$ to give an explicit way to think about this.

Remark 2.9 Since $\mathscr{K}_0(0) = \{*\}$, $\mathscr{K}_0(1) = \{\mathrm{id}\}$, and $\mathscr{K}_0(j) = \varnothing$ for $j > 1$, we have that $\boldsymbol{C}_0$ is the identity functor if $\mathscr{C}_G(0) = \{*\}$ and $\mathscr{C}_G(1) = \{\mathrm{id}\}$. In that case,

$$\mathbb{E}_G Y(0) = B(\Sigma^0, \boldsymbol{C}_0, Y) = B(\mathrm{Id}, \mathrm{Id}, Y) \cong Y.$$

We comment on an alternative point of view not taken above but relevant below. We can use the product of operads trick from [21] to replace a $\mathscr{C}_G$–space Y by the equivalent $\mathscr{K}_U$–space $B(\boldsymbol{K}_U, \boldsymbol{C}_U, Y)$, where $\boldsymbol{C}_U$ is the monad associated to the E_∞–operad $\mathscr{C}_U = \mathscr{C}_G \times \mathscr{K}_U$ and from there only use Steiner operads. However, there is a catch. A $\mathscr{K}_U$–algebra Y is a $\mathscr{K}_V$–algebra by restriction, but the constant $\mathscr{I}_G$–space Y is *not* a $\mathscr{K}_*$–G–space in the sense of Definition 1.24 since conjugation by isometries is not compatible with the inclusions used to define $\mathscr{K}_U$. Therefore, the $B(\Sigma^V, \mathscr{K}_V, Y)$ do not define an $\mathscr{I}_G$–space. However, ignoring isometries, they do define a coordinate free G–prespectrum, as defined in [19, II.1.2]. That can be viewed as the starting point for the alternative machine we construct next.

2.3 The infinite loop space machine: Lewis–May G–spectrum version

A Lewis–May (henceforward LM) G–spectrum E consists of G–spaces EV for each finite-dimensional G–inner product subspace V in a complete G–universe U together with G–homeomorphisms $EV \to \Omega^{W-V}EW$ whenever $V \subset W$. For a based G–space X we define $Q_G X = \operatorname{colim} \Omega^V \Sigma^V X$. The suspension LM G–spectrum $\Sigma_G^\infty X$ has V^{th} G–space $Q_G \Sigma^V X$, and the functor Σ_G^∞ is left adjoint to the zeroth space functor Ω_G^∞. We sometimes change notation to Σ_U^∞ and Ω_U^∞, allowing change of universe. While $G\mathscr{S}p$ is not symmetric monoidal, that is rectified by passage to the S_G–modules of [7], at the inevitable price of losing the adjunction; see [30, Section 11].

The operad $\mathscr{K}_U$ acts on $\Omega_U^\infty E$ for any LM G–spectrum E. One could not expect such precise structure when working with orthogonal G–spectra. Nonequivariantly, such highly structured infinite loop spaces are central to calculations, and it is to be hoped that the equivariant theory will eventually reach a comparable state. Therefore, it is natural to want an infinite loop space machine that lands in the category $G\mathscr{S}p$ of LM G–spectra.

The operad $\mathscr{K}_U$ plays a privileged role. As noted above, if $\mathscr{C}_G$ is an E_∞–G–operad, we can convert $\mathscr{C}_G$–spaces to equivalent $\mathscr{K}_U$–spaces, so that it suffices to build a machine for $\mathscr{K}_U$–spaces. On the other hand, $\mathscr{C}_G$ spaces inherit actions of $\mathscr{C}_U = \mathscr{C}_G \times \mathscr{K}_U$, so that it suffices to build a machine for $\mathscr{C}_U$–spaces. To encompass both of these approaches in a single machine, we suppose given a map (necessarily an equivalence) of E_∞–G–operads $\mathscr{O}_G \to \mathscr{K}_U$. We can take $\mathscr{O}_G = \mathscr{C}_U$ or $\mathscr{O}_G = \mathscr{K}_U$, but both here and in [12; 13; 31], our primary interest is in $\mathscr{C}_U$. Formally, the equivariant theory now works in the same way as the nonequivariant theory, and we follow the summary in [30, Section 9]. An early version of this machine is in the paper [4] of Costenoble and Waner.

Scholium 2.10 We must use the Steiner operads $\mathscr{K}_V$ and $\mathscr{K}_U$ rather than the little disks operads $\mathscr{D}_V$ and $\mathscr{D}_U$, which was the choice in [4], and our notion of an E_∞–operad of G–spaces should replace the notion of a complete operad used there.

Definition 2.11 (from $\mathscr{O}_G$–spaces to Lewis–May G–spectra) Let Y be an $\mathscr{O}_G$–space. We define a LM G–spectrum $\mathbb{E}_G Y$, which we denote by $\mathbb{E}_G^{\mathrm{Sp}}$ when necessary for clarity, by

$$\mathbb{E}_G Y = B(\Sigma_G^\infty, \boldsymbol{O}_G, Y).$$

Here $\boldsymbol{O}_G$ acts on Σ_G^∞ through its projection to $\mathbb{K}_U$.

We have the diagram of $\mathscr{O}_G$–spaces and $\mathscr{O}_G$–maps

$$(2.12)\quad Y \xleftarrow{\varepsilon} B(\boldsymbol{O}_G, \boldsymbol{O}_G, Y) \xrightarrow{\bar{\alpha}_U} B(Q_G, \boldsymbol{O}_G, Y) \xrightarrow{\zeta} \Omega_G^\infty B(\Sigma_G^\infty, \boldsymbol{O}_G, Y) = \Omega_G^\infty \mathbb{E}_G Y,$$

where $\bar{\alpha}_U = B(\alpha_U, \mathrm{id}, \mathrm{id})$. As explained nonequivariantly in [30, Section 9], the following analogue of Theorem 1.14 holds.

Theorem 2.13 *Let $\mathscr{O}_G$ be an E_∞–G–operad with a map of operads $\mathscr{O}_G \to \mathscr{K}_U$. The following statements hold for an $\mathscr{O}_G$–space Y:*

(i) *The map ε is a G–homotopy equivalence with a natural homotopy inverse ν.*

(ii) *The map $\overline{\alpha}_U$ is an equivalence when Y is connected and is a group completion otherwise.*

(iii) *The map ζ is an equivalence.*

Therefore, the composite

$$\xi = \zeta \circ \overline{\alpha}_U \circ \nu\colon Y \to \Omega_G^\infty \mathbb{E}_G Y$$

is an equivalence if Y is grouplike and is a group completion otherwise.

We shall not pursue this variant of the recognition principle in further detail, but we reemphasize that its much tighter relationship with space level data may eventually aid equivariant calculation. However, it is worth stating the alternative geometric version of the stable BPQ theorem to which it leads. Here we specialize to the case $\mathscr{O}_G = \mathscr{K}_U$. This allows us to use the pairings of Steiner operads described in Section 1.4, which are not available for other E_∞–operads. By passage to colimits, we obtain the following analogue of Proposition 1.17.

Proposition 2.14 *For G–universes U and U', there is a unital, associative, and commutative pairing*

$$\boxtimes\colon (\mathscr{K}_U, \mathscr{K}_{U'}) \to \mathscr{K}_{U\oplus U'}$$

of Steiner operads of G–spaces.

Passing to monads, we obtain a unital, associative, and commutative system of pairings

$$\boxtimes\colon \mathbb{K}_U X \wedge \mathbb{K}_{U'} Y \to \mathbb{K}_{U\oplus U'}(X \wedge Y). \tag{2.15}$$

Passage to colimits from Lemma 1.19 gives the following analogue of that result.

Lemma 2.16 *The following diagram commutes:*

$$\begin{tikzcd}
\boldsymbol{K}_U X \wedge \boldsymbol{K}_U Y \arrow[r, "\boxtimes"] \arrow[d, "\alpha_U \wedge \alpha_U"'] & \boldsymbol{K}_{U\oplus U}(X \wedge Y) \arrow[d, "\alpha_{U\oplus U}"] \\
\Omega_U^\infty \Sigma_U^\infty X \wedge \Omega_U^\infty \Sigma_U^\infty Y \arrow[r, "\wedge"'] & \Omega_{U\oplus U}^\infty \Sigma_{U\oplus U}^\infty (X \wedge Y)
\end{tikzcd}$$

The following recognition principle for pairings can by derived from Proposition 1.20 by passage to colimits or can be proven by the same argument as there. We note that our definition of the machine $\mathbb{E}_G$ depends on a choice of complete G–universe U, and we sometimes write $\mathbb{E}_U$ to indicate that choice.

Proposition 2.17 *A pairing $f\colon X \wedge Y \to Z$ of a $\mathscr{K}_U$–space X and a $\mathscr{K}_U$–space Y to a $\mathscr{K}_{U\oplus U}$–space Z induces a map*

$$\mathbb{E}f\colon \mathbb{E}_U X \barwedge \mathbb{E}_U Y \to \mathbb{E}_{U\oplus U} Z$$

of LM G–spectra indexed on $U \oplus U$ such that the following diagram commutes:

$$\begin{tikzcd}
X \wedge Y \arrow[r, "\xi\wedge\xi"] \arrow[d, "f"'] & \Omega^\infty_U \mathbb{E}_U X \wedge \Omega^\infty_U \mathbb{E}_U Y \arrow[r] & \Omega^\infty_{U\oplus U}(\mathbb{E}_U X \barwedge \mathbb{E}_U Y) \arrow[d, "\Omega^\infty_{U\oplus U}\mathbb{E}f"] \\
Z \arrow[rr, "\xi"'] & & \Omega^\infty_{U\oplus U}\mathbb{E}_{U\oplus U} Z
\end{tikzcd}$$

We can internalize the external smash product, as in [18], by choosing a linear isometry $\phi\colon U \oplus U \to U$. Then ϕ induces a change of universe functor ϕ_* which allows us to replace the right arrow by $\Omega^\infty_U \phi_* \mathbb{E} f$. In the following result we can either stick with Lewis–May G–spectra or pass to the S_G–modules of [7; 19]. We interpret the smash product according to choice.

Theorem 2.18 (the $\mathscr{K}_U$–space BPQ theorem for Lewis–May G–spectra) *For based G–spaces X, there is a natural equivalence*

$$\omega\colon \Sigma^\infty_U X \to \mathbb{E}_U \boldsymbol{K}_U X$$

such that the following diagram commutes for based G–spaces X and Y:

$$\begin{tikzcd}
\Sigma^\infty_U X \wedge \Sigma^\infty_U Y \arrow[r, "\omega\wedge\omega"] \arrow[d, "\cong"'] & \mathbb{E}_U \boldsymbol{K}_U X \wedge \mathbb{E}_U \boldsymbol{K}_U Y \arrow[d, "\mathbb{E}(\boxtimes)"] \\
\Sigma^\infty_U (X \wedge Y) \arrow[r, "\omega"'] & \mathbb{E}_U \boldsymbol{K}_U (X \wedge Y)
\end{tikzcd}$$

Sketch proof The first statement is the usual extra-degeneracy argument [21, Proposition 9.8]. We comment on the diagram. In either $G\mathscr{S}p$ or $G\mathscr{Z}$, it is an internalization of a diagram of G–spectra indexed on $U \oplus U$:

$$\begin{tikzcd}
\Sigma^\infty_U X \barwedge \Sigma^\infty_U Y \arrow[r, "\omega\barwedge\omega"] \arrow[d, "\cong"'] & \mathbb{E}_U \boldsymbol{K}_U X \barwedge \mathbb{E}_U \boldsymbol{K}_U Y \arrow[d, "\mathbb{E}(\boxtimes)"] \\
\Sigma^\infty_{U\oplus U} (X \wedge Y) \arrow[r, "\omega"'] & \mathbb{E}_{U\oplus U} \boldsymbol{K}_{U\oplus U} (X \wedge Y)
\end{tikzcd}$$

The isomorphism on the left is trivial on the prespectrum level (indexing on inner product G–spaces of the form $V \oplus W$) and follows on the spectrum level. After passage to adjoints, to check commutativity it suffices to check starting from $X \wedge Y$ on the bottom left, where an inspection of definitions gives the conclusion. If in $G\mathscr{S}p$, this is internalized by use of a linear isometry $\phi\colon U \oplus U \to U$. If in $G\mathscr{Z}$, this is internalized by use of the definition of the smash product in terms of the linear isometries operad $\mathscr{L}_U$, as in [7; 19]. □

In fact, with the model-theoretic modernization of the original version of the theory that is given nonequivariantly in [1], one can redefine the restriction of $\mathbb{E}_U$ to cofibrant $\mathscr{K}_U$–spaces Y to be

$$\mathbb{E}_U Y = \Sigma^\infty_G \otimes_{\boldsymbol{K}_U} Y,$$

where $\otimes_{\boldsymbol{K}_U}$ is the evident coequalizer. With that reinterpretation and taking X to be a G–CW complex, $\mathbb{E}_U \boldsymbol{K}_U X$ is actually isomorphic to $\Sigma^\infty_G X$.

The nonequivariant statement is often restricted to the case $Y = S^0$. Then $\boldsymbol{K}_U S^0$ is the disjoint union of operadic models for the classifying spaces $B\Sigma_j$. Similarly, $\boldsymbol{K}_U S^0$ is the disjoint union of operadic models for the classifying G–spaces $B(G, \Sigma_j)$.

2.4 A comparison of infinite loop space machines

We compare the $\mathscr{S}$ and $\mathscr{S}p$ machines $\mathbb{E}^{\mathscr{S}}_G$ and $\mathbb{E}^{\mathscr{S}p}_G$ by transporting both of them to the category $G\mathscr{Z}$ of S_G–modules, following [19]. As discussed in [19, Section IV.4] with slightly different notation, there is a diagram of Quillen equivalences:

$$\begin{tikzcd}
G\mathscr{P} \arrow[r, shift left, "L"] \arrow[d, shift left, "\mathbb{P}"] & G\mathscr{S}p \arrow[l, shift left, "\ell"] \arrow[d, shift left, "\mathbb{F}"] \\
G\mathscr{S} \arrow[u, shift left, "\mathbb{U}"] \arrow[r, shift left, "\mathbb{N}"] & G\mathscr{Z} \arrow[l, shift left, "\mathbb{N}^\#"] \arrow[u, shift left, "\mathbb{V}"]
\end{tikzcd}$$

Here $G\mathscr{P}$ is the category of coordinate-free G–prespectra. The left adjoint $\mathbb{N}$ is strong symmetric monoidal, and the unit map $\eta\colon X \to \mathbb{N}^\#\mathbb{N}X$ is a weak equivalence for all cofibrant orthogonal G–spectra X. It can be viewed as a fibrant approximation in the stable model structure on $G\mathscr{S}$. The pair $(\mathbb{N}, \mathbb{N}^\#)$ is a Quillen equivalence with the positive stable model structure on $G\mathscr{S}$; see [19, Sections III.4–5].

We can compare machines using the diagram. In fact, by a direct inspection of definitions, we see the following result, which is essentially a reinterpretation of the original construction of [21] that becomes visible as soon as one introduces orthogonal spectra.

Lemma 2.19 *The functor $\mathbb{E}_G^{\mathscr{S}p}$ from $\mathscr{C}_G$–spaces to the category $G\mathscr{S}p$ of Lewis–May G–spectra is naturally isomorphic to the composite functor $L \circ \mathbb{U} \circ \mathbb{E}_G^{\mathscr{S}}$.*

As explained in [19, Section IV.5], there is a monad $\mathbb{L}$ on $G\mathscr{S}p$ and a category $G\mathscr{S}p[\mathbb{L}]$ of $\mathbb{L}$–algebras. The left adjoint $\mathbb{F}$ in the diagram is the composite of left adjoints

$$\mathbb{L}\colon G\mathscr{S}p \to G\mathscr{S}p[\mathbb{L}] \quad \text{and} \quad \mathbb{J}\colon G\mathscr{S}p[\mathbb{L}] \to G\mathscr{Z}.$$

The functor $L \circ \mathbb{U}\colon G\mathscr{S} \to G\mathscr{S}p$ lands naturally in $G\mathscr{S}p[\mathbb{L}]$, so that we can define

$$\mathbb{M} = \mathbb{J} \circ L \circ \mathbb{U}\colon G\mathscr{S} \to G\mathscr{Z}.$$

By [19, Lemma IV.5.2 and Theorem IV.5.4], $\mathbb{M}$ is lax symmetric monoidal and there is a natural lax symmetric monoidal map $\alpha\colon \mathbb{N}X \to \mathbb{M}X$ that is a weak equivalence when X is cofibrant. Effectively, we have two infinite loop space machines landing in $G\mathscr{Z}$, namely $\mathbb{N} \circ \mathbb{E}_G^{\mathscr{S}}$ and $\mathbb{J} \circ \mathbb{E}_G^{\mathscr{S}p}$. In view of the lemma, the latter is isomorphic to $\mathbb{M} \circ \mathbb{E}_G^{\mathscr{S}}$; hence

$$\alpha\colon \mathbb{N} \circ \mathbb{E}_G^{\mathscr{S}} \to \mathbb{M} \circ \mathbb{E}_G^{\mathscr{S}} \cong \mathbb{J} \circ \mathbb{E}_G^{\mathscr{S}p}$$

compares the two machines, showing that they are equivalent for all practical purposes. Homotopically, these categorical distinctions are irrelevant, and we can use whichever machine we prefer, deducing properties of one from the other.

2.5 Examples of E_∞–spaces and E_∞–ring spaces

Many of the examples from the nonequivariant theory generalize directly to the equivariant setting. To illustrate the point of using varying E_∞–operads and their natural actions on spaces of interest, rather than just using $\mathscr{K}_U$, we focus on actions of the linear isometries operad $\mathscr{L}_U$.

Nonequivariantly, taking $U \cong \mathbb{R}^\infty$, a systematic account of naturally occurring examples of $\mathscr{L}_U$–spaces was already given in [24, Section I.1]. It was revisited briefly in more modern language [30, Section 2]. It includes the infinite classical groups O, SO, Spin, U, SU, Sp, their classifying spaces, constructed either using Grassmannian manifolds or the standard classifying-space functor B, and all of their associated infinite homogeneous spaces. All of these examples are grouplike, and all of them are given infinite loop spaces by application of the nonequivariant infinite loop space machine. The discussion in [24; 30] was in terms of inner-product subspaces V of a universe U. The point to make here is that the entire exposition works verbatim equivariantly, with the V being G–inner-product subspaces of our complete G–universe U. We give a brief account to show the idea.

As explained in [30, Section 2], an $\mathscr{I}_G$–FCP (functor with cartesian product) is a lax symmetric monoidal functor $\mathscr{I}_G \to \mathscr{T}_G$. We say that an $\mathscr{I}_G$–FCP is monoid valued if it factors through the category of equivariant topological monoids and monoid homomorphisms. The classical groups all give group-valued $\mathscr{I}_G$–FCPs:

$$V \mapsto O(V), \quad V \mapsto \mathrm{SO}(V), \quad V \mapsto U(\mathbb{C} \otimes_{\mathbb{R}} V), \quad V \mapsto \mathrm{SU}(\mathbb{C} \otimes_{\mathbb{R}} V), \quad \text{etc.}$$

Any $\mathscr{I}_G$–FCP X extends to a functor on all isometries (not just isometric isomorphisms) as follows: an isometry $\alpha\colon V \to W$ yields an identification $W \cong \alpha(V) \oplus \alpha(V)^{\perp}$. Then $X(\alpha)$ is the composite

$$X(V) \xrightarrow{X(\alpha)\times 0} X(\alpha(V)) \times X(\alpha(V)^{\perp}) \to X(\alpha(V) \oplus \alpha(V)^{\perp}).$$

Then the colimit $X(U) = \operatorname{colim}_V X(V)$ inherits an action of $\mathscr{L}_U$. The classifying space BF of a monoid-valued $\mathscr{I}_G$–FCP F is an $\mathscr{I}_G$–space, and the cited sources show that F is equivalent to ΩBF as an $\mathscr{L}_U$–space when F is group valued.

The formal structure of the operad pair $(\mathscr{K}_U, \mathscr{L}_U)$ works the same way equivariantly as nonequivariantly. It is an E_∞–operad pair in the sense originally defined in [24, VI.1.2] and reviewed in [30, Section 1] and, in more detail, [29, 4.2]. See Section 7.2 below for an example of an operad pair in G–categories. The action of $\mathscr{L}_U$ on $\mathscr{K}_U$ is defined nonequivariantly in [30, Section 3], and it works the same way equivariantly.

From here, multiplicative infinite loop space theory works equivariantly to construct E_∞–ring G–spectra from $(\mathscr{K}_U, \mathscr{L}_U)$–spaces, alias E_∞–ring G–spaces, in exactly the same way as nonequivariantly [24; 30; 29]. In particular, for any $\mathscr{L}_U$–algebra X, the free $\mathscr{K}_U$–algebra $\boldsymbol{K}_U X_+$ is an E_∞–ring G–space, where X_+ is obtained from X by adjoining an additive G–fixed basepoint 0. The group completion $\alpha_U\colon \boldsymbol{K}_U X_+ \to Q_G X_+$ is a map of E_∞–ring G–spaces, and $\mathbb{E}_G \boldsymbol{K}_U X_+$ is equivalent to $\Sigma^\infty_G X_+$ as E_∞–ring G–spectra.

As we intend to show elsewhere, the passage from category level data to E_∞–ring G–spaces, in analogy with [26; 29], generalizes to equivariant multicategories.

We remark that the usual construction of Thom G–spectra, such as MO_G and MU_G, already presents them as E_∞–ring G–spectra, without use of infinite loop space theory, as was explained and generalized in [18, Chapter X].

2.6 Some properties of equivariant infinite loop space machines

Many properties of the infinite loop space machine $\mathbb{E}_G$ follow directly from the group completion property, independent of how the machine is constructed, but it is notationally convenient to work with the machine $\mathbb{E}_G^{\mathscr{S}p}$, for which ξ is a natural group completion without any bother with fibrant approximation. The results apply equally well to $\mathbb{E}_G^{\mathscr{S}}$. It is plausible to hope that the group completion property actually characterizes the machine up to homotopy, as in [32], but the proof there fails equivariantly. A direct point-set level comparison of our machine with a new version of the Segal–Shimakawa machine will be given in [31]. We illustrate with the following two results, some version of which must hold for any equivariant infinite loop space machine $\mathbb{E}_G$. The first says that it commutes with passage to fixed points and the second says that it commutes with products, both up to weak equivalence.

Theorem 2.20 *For $\mathscr{C}_G$–spaces Y, there is a natural map of spectra*

$$\phi\colon \mathbb{E}(Y^G) \to (\mathbb{E}_G Y)^G$$

that induces a natural map of spaces under Y^G

$$\begin{tikzcd} & Y^G \arrow[dl, "\xi"'] \arrow[dr, "\xi^G"] & \\ \Omega^\infty \mathbb{E}(Y^G) \arrow[rr] & & (\Omega_G^\infty \mathbb{E}_G Y)^G \end{tikzcd}$$

in which the diagonal arrows are both group completions. Therefore, the horizontal arrow is a weak equivalence of spaces, and ϕ is a weak equivalence of spectra.

Proof For based G–spaces X, we have natural inclusions $\boldsymbol{C}_{U^G}(X^G) \to (\boldsymbol{C}_U X)^G$ and $\Sigma^\infty(X^G) \to (\Sigma_G^\infty X)^G$. For G–spectra E, we have a natural isomorphism $\Omega^\infty(E^G) \cong (\Omega_G^\infty E)^G$. This gives the required natural map of spectra

$$\mathbb{E}(Y^G) = B(\Sigma^\infty, \boldsymbol{C}_{U^G}, Y^G) \xrightarrow{\phi} (B(\Sigma_G^\infty, \boldsymbol{C}_U, Y))^G = (\mathbb{E}_G Y)^G$$

and the induced natural map of spaces under Y^G. Since the diagonal arrows in the diagram are group completions, the horizontal arrow must be a homology isomorphism and hence a weak equivalence. Since our spectra are connective, ϕ must also be a weak equivalence. □

Theorem 2.21 *Let X and Y be $\mathscr{C}_G$–spaces. Then the map*

$$\mathbb{E}_G(X \times Y) \to \mathbb{E}_G X \times \mathbb{E}_G Y$$

induced by the projections is a weak equivalence of G–spectra.

Proof We are using that the product of $\mathscr{C}_G$–spaces is a $\mathscr{C}_G$–space, the proof of which uses that the category of operads is cartesian monoidal. Working in $G\mathscr{S}p$, the functor Ω^∞_G commutes with products and passage to fixed points, and we have the commutative diagram:

$$\begin{tikzcd}
& (X\times Y)^H \cong X^H\times Y^H \arrow[dl, "\xi^H"'] \arrow[dr, "\xi^H\times\xi^H"] & \\
(\Omega^\infty_G\mathbb{E}_G(X\times Y))^H \arrow[rr] & & (\Omega^\infty_G\mathbb{E}_G X)^H\times(\Omega^\infty_G\mathbb{E}_G Y)^H
\end{tikzcd}$$

Since the product of group completions is a group completion, the diagonal arrows are both group completions. Therefore, the horizontal arrow is a weak equivalence. Since our spectra are connective, the conclusion follows. □

2.7 The recognition principle for naive G–spectra

We elaborate on Theorem 2.20. The functor $\mathbb{E}=\mathbb{E}_e$ in that result is the nonequivariant infinite loop space machine, which is defined using the product of the nonequivariant Steiner operad $\mathscr{K}=\mathscr{K}_{U^G}$ and the fixed-point operad $\mathscr{C}=(\mathscr{C}_G)^G$. We may think of U^G as $\mathbb{R}^\infty$, without reference to U, and start with any (naive) E_∞–operad $\mathscr{C}$ to obtain a recognition principle for naive G–spectra, which are just spectra with G–actions. Again we can use either the category $\mathscr{S}$ of orthogonal spectra or the category $\mathscr{S}p$ of Lewis–May spectra, comparing the two by mapping to the category $\mathscr{Z}$ of EKMM S–modules, but letting G act on objects in all three. We continue to write $\mathbb{E}$ for this construction since it is exactly the same construction as the nonequivariant one, but applied to G–spaces with an action by the G–trivial E_∞–operad $\mathscr{C}$.

It is worth emphasizing that when working with naive G–spectra, there is no need to restrict to finite groups. We can just as well work with general topological groups G. The machine $\mathbb{E}$ still enjoys the same properties, including the group completion property. Working with Lewis–May spectra, the adjunction $(\Sigma^\infty,\Omega^\infty)$ relating spaces and spectra applies just as well to give an adjunction relating based G–spaces and naive G–spectra. For based G–spaces X, the map $\alpha\colon \mathbf{C}X\to\Omega^\infty\Sigma^\infty X$ is a group completion of Hopf G–spaces by the nonequivariant special case since $(\mathbf{C}X)^H=\mathbf{C}(X^H)$ and $(\Omega^\infty\Sigma^\infty X)^H=\Omega^\infty\Sigma^\infty(X^H)$.

Returning to finite groups, we work with Lewis–May spectra and G–spectra in the rest of this section in order to exploit the more precise relationship between spaces and spectra that holds in that context. However, the conclusions can easily be transported to orthogonal spectra. We index genuine G–spectra on a complete G–universe U and we index naive G–spectra on the trivial G–universe $U^G\cong\mathbb{R}^\infty$. The inclusion of universes $i\colon U^G\to U$ induces a forgetful functor $i^*\colon G\mathscr{S}p^U\to G\mathscr{S}p^{U^G}$ from genuine G–spectra to naive G–spectra. It represents the forgetful functor from $RO(G)$–graded cohomology theories to $\mathbb{Z}$–graded cohomology theories. The functor i^* has a left adjoint i_*. The following observations are trivial but important.

Lemma 2.22 *The functors $i_*\Sigma^\infty$ and Σ^∞_G from based G–spaces to genuine G–spectra are isomorphic.*

Proof Clearly $\Omega^\infty \iota^* = \Omega^\infty_G$, since both are evaluation at $V = 0$, hence their left adjoints are isomorphic. □

Remark 2.23 For G–spaces X, the unit of the (i_*, i^*) adjunction gives a natural map $\Sigma^\infty X \to i^* i_* \Sigma^\infty X \cong \iota^* \Sigma^\infty_G$ of naive G–spectra. It is very far from being an equivalence, as the tom Dieck splitting theorem shows; see Theorem 6.5.

The inclusion of universes $i\colon U^G \to U$ induces an inclusion of operads of G–spaces $\iota\colon \mathscr{K}_{U^G} \to \mathscr{K}_U$, where G acts trivially on $\mathscr{K}_{U^G}$. The product of this inclusion and the inclusion $\iota\colon \mathscr{C} = (\mathscr{C}_G)^G \to \mathscr{C}_G$ is an inclusion

$$\iota\colon \mathscr{C}_{U^G} \equiv \mathscr{C} \times \mathscr{K}_{U^G} \to \mathscr{C}_G \times \mathscr{K}_U \equiv \mathscr{C}_U.$$

Pulling actions back along ι gives a functor ι^* from $\mathscr{C}_U$–spaces to $\mathscr{C}_{U^G}$–spaces. The following consistency statement is important since, by definition, the H–fixed-point spectrum E^H of a genuine G–spectrum E is $(i^*E)^H$ and the homotopy groups of E are $\pi^H_*(E) \equiv \pi_*(E^H)$.

Theorem 2.24 *Let Y be a $\mathscr{C}_G$–space. Then there is a natural weak equivalence of naive G–spectra $\mathbb{E}\iota^* Y \to i^* \mathbb{E}_G Y$.*

Proof Again, although we work with $\mathbb{E}^{\mathscr{S}p}_G$, the conclusion carries over to $\mathbb{E}^{\mathscr{S}}_G$. It is easy to check from the definitions that, for G–spaces X, we have a natural commutative diagram of G–spaces:

$$\begin{tikzcd}
\boldsymbol{C}_{U^G} X \arrow[r, "\alpha"] \arrow[d] & \Omega^\infty \Sigma^\infty X \arrow[d] \\
\boldsymbol{C}_U X \arrow[r, "\alpha"'] & \Omega^\infty_G \Sigma^\infty_G X
\end{tikzcd}$$

The vertical arrows both restrict colimits over representations to colimits over trivial representations. Passing to adjoints, we obtain a natural commutative diagram:

$$\begin{tikzcd}
\Sigma^\infty_G \boldsymbol{C}_{U^G} X \arrow[r] \arrow[d] & \Sigma^\infty X \arrow[d] \\
\Sigma^\infty_G \boldsymbol{C}_U X \arrow[r] & \Sigma^\infty_G X
\end{tikzcd}$$

The composite gives a right action of $\boldsymbol{C}_{UG}$ on Σ_G^∞ that is compatible with the right action of $\boldsymbol{C}_U$. Using the natural map $\Sigma^\infty \to i^*\Sigma_G^\infty$ of Remark 2.23, there results a natural map

$$\mu\colon \mathbb{E}\iota^*Y = B(\Sigma^\infty, \boldsymbol{C}_{UG}, \iota^*Y) \to B(i^*\Sigma_G^\infty, \boldsymbol{C}_U, Y) \cong i^*\mathbb{E}_G Y$$

of naive G–spectra. The following diagram commutes by a check of definitions:

$$\begin{tikzcd}
Y \arrow[d, "="'] & B(\boldsymbol{C}_{UG}, \boldsymbol{C}_{UG}, \iota^*Y) \arrow[l, "\varepsilon"'] \arrow[r, "B(\alpha,\mathrm{id},\mathrm{id})"] \arrow[d] & B(Q, \boldsymbol{C}_{UG}, Y) \arrow[r, "\zeta"] \arrow[d] & \Omega^\infty B(\Sigma^\infty, \boldsymbol{C}_{UG}, Y) \arrow[d, "\Omega^\infty\mu"] \\
Y & B(\boldsymbol{C}_G, \boldsymbol{C}_G, Y) \arrow[l, "\varepsilon"'] \arrow[r, "B(\alpha,\mathrm{id},\mathrm{id})"] & B(Q_G, \boldsymbol{C}_U, Y) \arrow[r, "\zeta"] & \Omega_G^\infty B(\Sigma_G^\infty, \boldsymbol{C}_U, Y)
\end{tikzcd}$$

Here the right vertical map is the map of zeroth spaces given by μ. Replacing the maps ε with their homotopy inverses, the horizontal composites become group completions. Therefore, $\Omega^\infty\mu$ is a weak equivalence, hence so is μ. $\square$

We also have the corresponding statement for the left adjoint i_* of i^*. In effect, it gives a space level construction of the change of universe functor i_* on connective G–spectra. We need a homotopically well-behaved version of the left adjoint of the functor ι^* from $\mathscr{C}$–spaces to $\mathscr{C}_G$–spaces, and we define it by $\iota_! X = B(\boldsymbol{C}_G, \boldsymbol{C}, X)$.

Theorem 2.25 *Let X be a $\mathscr{C}$–space. Then there is a natural weak equivalence of genuine G–spectra $\mathbb{E}_G(i_! X) \simeq i_*\mathbb{E}(X)$.*

We give the proof in Appendix B, using a construction that is of independent interest.

3 Categorical preliminaries on classifying G–spaces and G–operads

We recall an elementary functor $\mathscr{C}at(\mathcal{E}G, -)$ from G–categories to G–categories from our paper [11] with Mona Merling. We explored this functor in detail in the context of equivariant bundle theory in [11], and we refer the reader there for proofs. In Section 4, we shall use it to define a certain operad $\mathscr{P}_G$ of G–categories. The $\mathscr{P}_G$–algebras will be the genuine permutative G–categories.

3.1 Chaotic topological categories and equivariant classifying spaces

For (small) categories $\mathscr{A}$ and $\mathscr{B}$, we let $\mathscr{C}at(\mathscr{A}, \mathscr{B})$ denote the category whose objects are the functors $\mathscr{A} \to \mathscr{B}$ and whose morphisms are the natural transformations between them. When $\mathscr{B}$ has a right action by some group Π, then $\mathscr{C}at(\mathscr{A},\mathscr{B})$ inherits a right Π–action. When a group G acts from the left on $\mathscr{A}$ and $\mathscr{B}$, $\mathscr{C}at(\mathscr{A},\mathscr{B})$ inherits a left G–action by conjugation on objects and morphisms. Then $G\mathscr{C}at(\mathscr{A},\mathscr{B})$ is alternative notation for the G–fixed category $\mathscr{C}at(\mathscr{A},\mathscr{B})^G$ of G–functors and G–natural transformations. We have the G–equivariant version of the standard adjunction

$$\mathscr{C}at(\mathscr{A}\times\mathscr{B},\mathscr{C})\cong\mathscr{C}at(\mathscr{A},\mathscr{C}at(\mathscr{B},\mathscr{C})). \tag{3.1}$$

Definition 3.2 For a space X, the chaotic (topological) category $\mathcal{E}X$ has object space X, morphism space $X\times X$, and structure maps I, S, T, and C given by $I(x)=(x,x)$, $S(y,x)=x$, $T(y,x)=y$, and $C((z,y),(y,x))=(z,x)$. For any point $*\in X$, the map $\eta\colon X\to X\times X$ specified by $\eta(x)=(*,x)$ is a continuous natural isomorphism from the identity functor to the trivial functor $\mathcal{E}X\to *\to\mathcal{E}X$, hence $\mathcal{E}X$ is equivalent to $*$. When $X=G$ is a topological group, $\mathcal{E}G$ is isomorphic to the translation category of G, but the isomorphism encodes information about the group action and should not be viewed as an identification; see [11, Remark 1.7]. We say that a topological category with object space X is chaotic if it is isomorphic to $\mathcal{E}X$.

Definition 3.3 Without changing notation, we regard a topological group Π as a topological category with a single object $*$ and morphism space Π, with composition given by multiplication. Then Π is isomorphic to the orbit category $\mathcal{E}\Pi/\Pi$, where Π acts from the right on $\mathcal{E}\Pi$ via right multiplication on objects and diagonal right multiplication on morphisms. The resulting functor $p\colon\mathcal{E}\Pi\to\Pi$ is given by the trivial map $\Pi\to *$ of object spaces and the map $p\colon\Pi\times\Pi\to\Pi\times\Pi/\Pi\cong\Pi$ on morphism spaces specified by $p(\tau,\sigma)=\tau\sigma^{-1}$.

Theorem 3.4 [11, Theorem 2.7] *For a G–space X and a topological group Π, regarded as a G–trivial G–space, the functor $p\colon\mathcal{E}\Pi\to\Pi$ induces an isomorphism of topological G–categories*

$$\xi\colon\mathscr{C}at(\mathcal{E}X,\mathcal{E}\Pi)/\Pi\to\mathscr{C}at_G(\mathcal{E}X,\Pi).$$

Therefore, passing to G–fixed-point categories,

$$(\mathscr{C}at(\mathcal{E}X,\mathcal{E}\Pi)/\Pi)^G\cong\mathscr{C}at(\mathcal{E}X,\Pi)^G\cong\mathscr{C}at(\mathcal{E}X/G,\Pi).$$

The last isomorphism is clear since G acts trivially on Π. Situations where G is allowed to act nontrivially on Π are of considerable interest, as we shall see in Section 4.4, but otherwise they will only appear peripherally in this paper. The paper [11] works throughout in that more general context. The previous result will not be used directly, but it is the key underpinning for the results of the next section.

3.2 The functor $\mathscr{C}at(\mathcal{E}G,-)$

The functor $\mathscr{C}at(\mathcal{E}G,-)$ from G–categories to G–categories is a right adjoint (3.1), hence it preserves limits and in particular products. The projection $\mathcal{E}G \to *$ to the trivial G–category induces a natural map

$$\iota\colon \mathscr{A} = \mathscr{C}at(*,\mathscr{A}) \to \mathscr{C}at(\mathcal{E}G,\mathscr{A}). \tag{3.5}$$

The map ι is not an equivalence of G–categories in general [11, Proposition 4.19], but the functor $\mathscr{C}at(\mathcal{E}G,-)$ is idempotent in the sense that the following result holds.

Lemma 3.6 *For any G–category $\mathscr{A}$,*

$$\iota\colon \mathscr{C}at(\mathcal{E}G,\mathscr{A}) \to \mathscr{C}at(\mathcal{E}G,\mathscr{C}at(\mathcal{E}G,\mathscr{A}))$$

is an equivalence of G–categories.

Proof This follows from the adjunction (3.1) using that the diagonal $\mathcal{E}G \to \mathcal{E}G \times \mathcal{E}G$ is an equivalence with inverse given by either projection and that the specialization of ι here is induced by the first projection. □

Lemma 3.7 [11, Lemma 3.7] *Let Λ be a subgroup of $G\times\Pi$. The Λ–fixed category $\mathscr{C}at(\mathcal{E}G,\mathcal{E}\Pi)^{\Lambda}$ is empty if $\Lambda\cap\Pi \neq e$ and is nonempty and chaotic if $\Lambda\cap\Pi = e$.*

With G acting trivially on Π, let $H^1(G;\Pi)$ denote the set of isomorphism classes of homomorphisms $\alpha\colon G\to\Pi$. Equivalently, it is the set of Π–conjugacy classes of subgroups $\Lambda = \{(g,\alpha(g)) \mid g\in G\}$ of $G\times\Pi$. Define $\Pi^{\alpha}\subset\Pi$ to be the subgroup of elements σ that commute with $\alpha(g)$ for all $g\in G$.

Theorem 3.8 [11, Theorems 4.14 and 4.18] *For $H\subset G$, the H–fixed category $\mathscr{C}at(\mathcal{E}G,\Pi)^{H}$ is equivalent to the coproduct of the groups Π^{α} (regarded as categories), where the coproduct runs over $[\alpha]\in H^1(H;\Pi)$.*

Definition 3.9 Define $E(G,\Pi) = |\mathscr{C}at(\mathcal{E}G,\mathcal{E}\Pi)|$ and $B(G,\Pi) = |\mathscr{C}at(\mathcal{E}G,\Pi)|$. Let

$$p\colon E(G,\Pi)\to B(G,\Pi)$$

be induced by the passage to orbits functor $\mathcal{E}\Pi\to\Pi$.

Theorem 3.10 [11, Theorems 3.11, 4.23, 4.24] *Let Π be a discrete or compact Lie group and let G be a discrete group. Then $p\colon E(G,\Pi)\to B(G,\Pi)$ is a universal principal (G,Π)–bundle. For a subgroup H of G,*

$$B(G,\Pi)^{H} \simeq \coprod B(\Pi^{\alpha}),$$

where the union runs over $[\alpha]\in H^1(H;\Pi)$.

3.3 E_∞–operads of G–categories

The definition of an E_∞–operad of G–spaces given in Section 2.1 has the following categorical analogue.

Definition 3.11 An E_∞–operad $\mathscr{O}_G$ of (topological) G–categories is an operad in the cartesian monoidal category $G\mathscr{C}at$ such that $|\mathscr{O}_G|$ is an E_∞–operad of G–spaces. We say that $\mathscr{O}_G$ is reduced if $\mathscr{O}_G(0)$ is the trivial category. In practice, the $\mathscr{O}_G(j)$ are groupoids.

The proof of Lemma 2.5 works just as well to give the following analogue.

Lemma 3.12 *Let $\mathscr{O}_G$ be an E_∞–operad of G–categories. Then $\mathscr{O} = (\mathscr{O}_G)^G$ is an E_∞–operad of categories. If $\mathscr{A}$ is an $\mathscr{O}_G$–category, then $\mathscr{A}^G$ is an $\mathscr{O}$–category.*

4 Categorical philosophy: what is a permutative G–category?

4.1 Naive permutative G–categories

We have a notion of a monoidal category $\mathscr{A}$ internal to a cartesian monoidal category $\mathscr{V}$. It is a category internal to $\mathscr{V}$ together with a coherently associative and unital product $\mathscr{A} \times \mathscr{A} \to \mathscr{A}$. It is strict monoidal if the product is strictly associative and unital. It is symmetric monoidal if it has an equivariant symmetry isomorphism satisfying the usual coherence properties. A functor $F\colon \mathscr{A} \to \mathscr{B}$ between symmetric monoidal categories is strict monoidal if $F(A \otimes A') = FA \otimes FA'$ for $A, A' \in \mathscr{A}$ and $FI = J$, where I and J are the unit objects of $\mathscr{A}$ and $\mathscr{B}$.

A permutative category is a symmetric strict monoidal category.[4] Taking $\mathscr{V}$ to be $\mathscr{U}$, these are the topological permutative categories. Taking $\mathscr{V}$ to be $G\mathscr{U}$, these are the *naive* topological permutative G–categories.

Nonequivariantly, there is a standard E_∞–operad of spaces that is obtained by applying the classifying-space functor to an E_∞–operad $\mathscr{P}$ of categories. The following definition goes back to Barratt and Eccles, thought of simplicially [2], and to [22], thought of categorically.

[4] In interesting examples, the product cannot be strictly commutative.

Definition 4.1 We define an E_∞–operad $\mathscr{P}$ of categories. Let $\mathscr{P}(j) = \mathcal{E}\Sigma_j$. Since Σ_j acts freely and $\mathcal{E}\Sigma_j$ is chaotic, the classifying space $|\mathscr{P}(j)|$ is Σ_j–free and contractible, as required of an E_∞–operad. The structure maps

$$\gamma\colon \mathcal{E}\Sigma_k \times \mathcal{E}\Sigma_{j_1} \times \cdots \times \mathcal{E}\Sigma_{j_k} \to \mathcal{E}\Sigma_j,$$

where $j = j_1 + \cdots + j_k$, are dictated on objects by the definition of an operad. If we view the object sets of the $\mathscr{P}(j)$ as discrete categories (identity morphisms only), then they form the associativity operad $\mathscr{M}$.

We can define $\mathscr{M}$–algebras and $\mathscr{P}$–algebras in $\mathscr{C}at$ or in $G\mathscr{C}at$. In the latter case, we regard $\mathscr{M}$ and $\mathscr{P}$ as operads with trivial G–action. The following result characterizes naive permutative G–categories operadically. The proof is easy [22].

Proposition 4.2 *The category of strict monoidal G–categories and strict monoidal G–functors is isomorphic to the category of $\mathscr{M}$–algebras in $G\mathscr{C}at$. The category of naive permutative G–categories and strict symmetric monoidal G–functors is isomorphic to the category of $\mathscr{P}$–algebras in $G\mathscr{C}at$.*

The term "naive" is appropriate since naive permutative G–categories give rise to naive G–spectra on application of an infinite loop space machine. Genuine permutative G–categories need more structure, especially precursors of transfer maps, to give rise to genuine G–spectra. Nonequivariantly, there is no distinction.

4.2 Genuine permutative G–categories

The following observation will play a helpful role in our work. Recall the natural map $\iota\colon \mathscr{A} \to \mathscr{C}at(\mathcal{E}G, \mathscr{A})$ of (3.5).

Lemma 4.3 *For any space X regarded as a G–trivial G–space, $\iota\colon \mathcal{E}X \to \mathscr{C}at(\mathcal{E}G, \mathcal{E}X)$ is the inclusion of the G–fixed category $G\mathscr{C}at(\mathcal{E}G, \mathcal{E}X)$.*

Proof Since $\mathcal{E}X$ is chaotic, functors $\mathcal{E}G \to \mathcal{E}X$ are determined by their object map $G \to X$ and are G–fixed if and only if the object map factors through $G/G = *$. $\square$

Definition 4.4 Let $\mathscr{P}_G$ be the (reduced) operad of G–categories whose j^{th} G–category is $\mathscr{P}_G(j) = \mathscr{C}at(\mathcal{E}G, \mathscr{P}(j))$, where $\mathscr{P}(j) = \mathcal{E}\Sigma_j$ is viewed as a G–category with trivial G–action and is given its usual right Σ_j–action. The unit in $\mathscr{P}_G(1)$ is the unique functor from $\mathcal{E}G$ to the trivial category $\mathscr{P}(1) = \mathscr{P}_G(1)$. The structure maps γ of $\mathscr{P}_G$ are induced from those of $\mathscr{P}$, using that the functor $\mathscr{C}at(\mathcal{E}G, -)$ preserves products. By Theorem 3.10, $\mathscr{P}_G$ is an E_∞–operad of G–categories. The natural map ι of (3.5) induces an inclusion $\iota\colon \mathscr{P} = (\mathscr{P}_G)^G \to \mathscr{P}_G$ of operads of G–categories.

Definition 4.5 A *genuine* permutative G–category is a $\mathscr{P}_G$–algebra in $G\mathscr{C}at$. A map of genuine permutative G–categories is a map of $\mathscr{P}_G$–algebras.

We usually call these $\mathscr{P}_G$–categories. We have an immediate source of examples. Let ι^* be the functor from genuine permutative G–categories to naive permutative G–categories that is obtained by restricting actions by $\mathscr{P}_G$ to its suboperad $\mathscr{P}$.

Proposition 4.6 *The action of $\mathscr{P}$ on a naive permutative G–category $\mathscr{A}$ induces an action of $\mathscr{P}_G$ on $\mathscr{C}at(\mathcal{E}G, \mathscr{A})$. Therefore, $\mathscr{C}at(\mathcal{E}G, -)$ restricts to a functor from naive permutative G–categories to genuine permutative G–categories.*

Proof This holds since the functor $\mathscr{C}at(\mathcal{E}G, -)$ preserves products. □

Proposition 4.7 *The map ι of* (3.5) *restricts to a natural map $\mathscr{A} \to \iota^*\mathscr{C}at(\mathcal{E}G, \mathscr{A})$ of naive permutative G–categories, and ι is an equivalence when $\mathscr{A} = \iota^*\mathscr{C}at(\mathcal{E}G, \mathscr{B})$ for a naive permutative G–category $\mathscr{B}$.*

Proof Since ι is induced by the projection $\mathcal{E}G \to \mathcal{E}\{e\} = *$, the first claim is clear, and the second holds by Lemma 3.6. □

As noted before, the map $\iota\colon \mathscr{A} \to \iota^*\mathscr{C}at(\mathcal{E}G, \mathscr{A})$ is not an equivalence in general [11, Proposition 4.19]. The $\mathscr{P}_G$–categories of interest in this paper are of the form $\mathscr{C}at(\mathcal{E}G, \mathscr{A})$ for a naive permutative G–category $\mathscr{A}$. In fact, we do not yet know how to construct other examples, although we believe that they exist.

Remark 4.8 Shimakawa [43, page 256] introduced the E_∞–operad $\mathscr{P}_G$ under the name $\mathscr{D}$ and demonstrated the first part of Proposition 4.6.

Remark 4.9 One might hope that $(\mathscr{C}at(\mathcal{E}G, -), \iota^*)$ is an adjoint pair. However, regarding ι^* monadically as the forgetful functor from $\mathbb{P}_G$–algebras to $\mathbb{P}$–algebras, its left adjoint is the coend that sends a naive permutative G–category $\mathscr{A}$ to the genuine permutative G–category $\mathbb{P}_G \otimes_{\mathbb{P}} \mathscr{A}$, which is the coequalizer in $G\mathscr{C}at$ of the maps $\mathbb{P}_G\mathbb{P}\mathscr{A} \rightrightarrows \mathbb{P}_G\mathscr{A}$ induced by the action map $\mathbb{P}\mathscr{A} \to \mathscr{A}$ and by the map $\mathbb{P}_G\mathbb{P} \to \mathbb{P}_G\mathbb{P}_G \to \mathbb{P}_G$ induced by the inclusion $\mathbb{P} \to \mathbb{P}_G$ and the product on $\mathbb{P}_G$. The universal property of the coequalizer gives a natural map

$$\tilde{\iota}\colon \mathbb{P}_G \otimes_{\mathbb{P}} \mathscr{A} \to \mathscr{C}at(\mathcal{E}G, \mathscr{A})$$

of genuine permutative G–categories that restricts to ι on $\mathscr{A}$, but $\tilde{\iota}$ is not an isomorphism. We shall say a bit more about this in Remark 4.20.

4.3 E_∞–G–categories

We can generalize the notion of a genuine permutative G–category by allowing the use of E_∞–operads other than $\mathscr{P}_G$. In fact, thinking as algebraic topologists rather than category theorists, there is no need to give the particular E_∞–operad $\mathscr{P}_G$ a privileged role.

Definition 4.10 An E_∞–G–category $\mathscr{A}$ is a G–category together with an action of some E_∞–operad $\mathscr{O}_G$ of G–categories. The classifying space $B\mathscr{A} = |\mathscr{A}|$ is then an $|\mathscr{O}_G|$–space and thus an E_∞–G–space.

We may think of E_∞–G–categories as generalized kinds of genuine permutative G–categories. The point of the generalization is that we have interesting examples of E_∞–operads of G–categories with easily recognizable algebras. We shall later define E_∞–operads $\mathscr{V}_G$, $\mathscr{V}_G^\times$, and $\mathscr{W}_G$ that are interrelated in a way that illuminates the study of multiplicative structures.

Observe that $\mathscr{P}_G$–algebras, like nonequivariant permutative categories, have a canonical product, whereas E_∞–G–categories over other operads do not. The general philosophy of operad theory is that algebras over an operad $\mathscr{C}$ in any suitable category $\mathscr{V}$ have j–fold operations parametrized by the objects $\mathscr{C}(j)$. Homotopical properties of $\mathscr{C}$ relate these operations. In general, in an E_∞ space, there is no preferred choice of a product on its underlying H–space, and none is relevant to the applications; E_∞–G–categories work similarly.

Remark 4.11 Symmetric monoidal categories occur more often "in nature" than permutative categories. We have not specified a notion of a genuine symmetric monoidal G–category in this paper. One approach is to apply the construction $\mathscr{C}at(\mathcal{E}G, -)$ to the tree operad that defines symmetric monoidal categories. Another approach, which we find more useful, is to define a genuine symmetric monoidal G–category to be a pseudoalgebra over $\mathscr{P}_G$. That approach is developed and applied in the categorical sequels [12; 13]. We shall not pursue the topic further here. A first comparison between symmetric monoidal G–categories and (genuine) G–symmetric monoidal categories, whose definition is a priori quite different, is given in Hill and Hopkins [16, Section 3.2], but work in progress shows that there is a good deal more to be said about that comparison and about the comparison between these notions and Tambara functors that is given in [16, Section 5.1].

Up to homotopy, any two choices of E_∞–operads give rise to equivalent categories of E_∞–G–spaces. To see that, we apply the trick from [21] of using products of operads to transport operadic algebras from one E_∞–operad to another. The product of operads $\mathscr{C}$ and $\mathscr{D}$ in any cartesian monoidal category $\mathscr{V}$ is given by

$$(\mathscr{C} \times \mathscr{D})(j) = \mathscr{C}(j) \times \mathscr{D}(j),$$

with the evident permutations and structure maps. With the choices of $\mathscr{V}$ of interest to us, the product of E_∞–operads is an E_∞–operad. The projections

$$\mathscr{C} \leftarrow \mathscr{C} \times \mathscr{D} \rightarrow \mathscr{D}$$

allow us to construct $(\mathscr{C}\times\mathscr{D})$–algebras in $\mathscr{V}$ from either $\mathscr{C}$–algebras or $\mathscr{D}$–algebras in $\mathscr{V}$, by pullback of action maps along the projections.

More generally, for any map $\mu\colon \mathscr{C} \rightarrow \mathscr{D}$ of operads in $\mathscr{V}$, the pullback functor μ^* from $\mathscr{D}$–algebras to $\mathscr{C}$–algebras has a left adjoint pushforward functor $\mu_!$ from $\mathscr{C}$–algebras to $\mathscr{D}$–algebras. One can work out a homotopical comparison model categorically. Pragmatically, use of the two-sided bar construction as in [21; 30] gives all that is needed. One redefines $\mu_! X = B(\mathbb{D}, \mathbb{C}, X)$, where $\mathbb{C}$ and $\mathbb{D}$ are the monads whose algebras are the $\mathscr{C}$–algebras and $\mathscr{D}$–algebras.[5] In spaces, or equally well G–spaces, μ^* and $\mu_!$ give inverse equivalences of homotopy categories between $\mathscr{C}$–algebras and $\mathscr{D}$–algebras when $\mathscr{C}$ and $\mathscr{D}$ are E_∞–operads.

Starting with operads in $\mathscr{C}at$ or in $G\mathscr{C}at$ we can first apply the classifying-space functor and then apply this trick. The conclusion is that all E_∞–categories and E_∞–G–categories give equivalent inputs for infinite loop space machines. In particular, for example, letting $\boldsymbol{O}_G$, $\boldsymbol{P}_G$, and $\boldsymbol{O}_G \times \boldsymbol{P}_G$ denote the monads in the category of G–spaces whose algebras are $|\mathscr{O}_G|$–algebras, $|\mathscr{P}_G|$–algebras, and $|\mathscr{O}_G \times \mathscr{P}_G|$–algebras, we see that after passage to classifying spaces, every $\boldsymbol{P}_G$–algebra Y determines an $\boldsymbol{O}_G$–algebra $X = B(\boldsymbol{O}_G, \boldsymbol{O}_G \times \boldsymbol{P}_G, Y)$ such that X and Y are weakly equivalent as $(\boldsymbol{O}_G \times \boldsymbol{P}_G)$–algebras (and conversely). This says that for purposes of equivariant infinite loop space theory, $\mathscr{P}_G$ and any other E_∞–operad $\mathscr{O}_G$ can be used interchangeably, regardless of how their algebras compare categorically.

4.4 Equivariant algebraic K–theory

The most interesting nonequivariant permutative categories are given by categories $\mathscr{A} = \coprod \Pi_n$, where $\{\Pi_n \mid n \geq 0\}$ is a sequence of groups (regarded as categories with a single object) and where the permutative structure is given by an associative and unital system of pairings $\Pi_m \times \Pi_n \rightarrow \Pi_{m+n}$. Then the pairings give the classifying space $B\mathscr{A} = \coprod B\Pi_n$ a structure of topological monoid, and one definition of the algebraic K–groups of $\mathscr{A}$ is the homotopy groups of the space $\Omega B(B\mathscr{A})$.

Equivariantly, it is sensible to replace the spaces $B\Pi_n$ by the classifying G–spaces $B(G, \Pi_n)$ and proceed by analogy. This definition of equivariant algebraic K–groups was introduced and studied calculationally in [9]. It is the equivariant analogue of

[5] Of course, this is an abuse of notation, since $\mu_!$ here is really a derived functor.

Quillen's original definition in terms of the plus construction. With essentially the same level of generality, the analogue of Quillen's definition in terms of the Q–construction has been studied by Dress and Kuku [6; 17]. Shimada [44] has given an equivariant version of Quillen's "plus $=$ Q" theorem in this context.

Regarding $\mathscr{A}$ as a G–trivial naive permutative G–category, we see that the classifying G–space of the genuine permutative G–category $\mathscr{C}at(\mathcal{E}G,\mathscr{A})$ is the disjoint union of classifying spaces $B(G,\Pi_n)$. Just as nonequivariantly, the functor ΩB can be replaced by the zeroth space functor $\Omega^\infty_G \mathbb{E}_G$ of an infinite loop G–space machine $\mathbb{E}_G$. The underlying equivariant homotopy type is unchanged. Therefore, we may redefine the algebraic K–groups to be the homotopy groups of the genuine G–spectrum $\mathbb{K}_G\mathscr{A} \equiv \mathbb{E}_G B\mathscr{C}at(\mathcal{E}G,\mathscr{A})$. Essentially the same definition is implicit in Shimakawa [43], who focused on an equivariant version of Segal's infinite loop space machine. A different equivariant version of Segal's machine is developed and compared to Shimakawa's in [31]. It is generalized categorically in [12; 13].

Applying the functor $\mathscr{C}at(\mathcal{E}G,-)$ to naive permutative G–categories $\mathscr{A}$ with nontrivial G–actions gives more general input for equivariant algebraic K–theory than has been studied in the literature. This allows for G–actions on the groups Π_n, and we then replace $B(G,\Pi_n)$ by classifying G–spaces $B(G,(\Pi_n)_G)$ for the $(G,(\Pi_n)_G)$–bundles associated to the split extensions $\Pi_n \rtimes G$. Such classifying spaces are studied in [11]. Alternative but equivalent constructions of the associated G–spectra $\mathbb{K}_G(\mathscr{A})$ are given in Section 4.5 and Section 8.2 below. The resulting generalization of equivariant algebraic K–theory is studied in [33].

4.5 The recognition principle for permutative G–categories

We may start with any E_∞–operad $\mathscr{O}_G$ of G–categories and apply the classifying-space functor to obtain an E_∞–operad $|\mathscr{O}_G|$ of G–spaces. If $\mathscr{O}_G$ acts on a category $\mathscr{A}$, then $|\mathscr{O}_G|$ acts on $|\mathscr{A}| = B\mathscr{A}$. We can replace $|\mathscr{O}_G|$ by its product with the Steiner operads $\mathscr{K}_V$ or with the Steiner operad $\mathscr{K}_U$ and apply the functor $\mathbb{E}_G^{\mathscr{S}}$ or $\mathbb{E}_G^{\mathscr{S}p}$ to obtain a (genuine) associated G–spectrum, which we denote ambiguously by $\mathbb{E}_G(B\mathscr{A})$.

Definition 4.12 Define the (genuine) algebraic K–theory G–spectrum of an $\mathscr{O}_G$–category $\mathscr{A}$ by $\mathbb{K}_G(\mathscr{A}) = \mathbb{E}_G(B\mathscr{A})$.

We might also start with an operad $\mathscr{O}$ of categories such that $|\mathscr{O}|$ is an E_∞–operad of spaces and regard these as G–objects with trivial action. Following up the previous section, we then have the following related but less interesting notion.

Definition 4.13 The (naive) algebraic K–theory G–spectrum of an $\mathscr{O}$–category $\mathscr{A}$ is defined by $\mathbb{K}(\mathscr{A}) = \mathbb{E}(B\mathscr{A})$.

Until Section 7, we restrict attention to the cases $\mathscr{O}_G = \mathscr{P}_G$ and $\mathscr{O} = \mathscr{P}$, recalling that the $\mathscr{P}_G$–categories are the genuine permutative G–categories, the $\mathscr{P}$–categories are the naive permutative G–categories, and the inclusion $\iota\colon \mathscr{P} \to \mathscr{P}_G$ induces a forgetful functor ι^* from genuine to naive permutative G–categories. Since the classifying-space functor commutes with products, passage to fixed points, and the functors ι^*, Theorems 2.20, 2.21, and 2.24 have the following immediate corollaries. The first was promised in [10, Theorem 2.2].

Theorem 4.14 *For $\mathscr{P}_G$–categories $\mathscr{A}$, there is a natural weak equivalence of spectra*

$$\mathbb{K}(\mathscr{A}^G) \to (\mathbb{K}_G \mathscr{A})^G.$$

Theorem 4.15 *Let $\mathscr{A}$ and $\mathscr{B}$ be $\mathscr{P}_G$–categories. Then the map*

$$\mathbb{K}_G(\mathscr{A} \times \mathscr{B}) \to \mathbb{K}_G \mathscr{A} \times \mathbb{K}_G \mathscr{B}$$

induced by the projections is a weak equivalence of G–spectra.

Theorem 4.16 *For $\mathscr{P}_G$–categories $\mathscr{A}$, there is a natural weak equivalence of naive G–spectra $\mathbb{K}\iota^*\mathscr{A} \to i^*\mathbb{K}_G\mathscr{A}$.*

The algebraic K–groups of $\mathscr{A}$ are defined to be the groups

$$K_*^H \mathscr{A} = \pi_*^H(\mathbb{K}\iota^*\mathscr{A}) \cong \pi_*^H(\mathbb{K}_G\mathscr{A}). \tag{4.17}$$

We are particularly interested in examples of the form $\mathscr{C}at(\mathcal{E}G, \mathscr{A})$, where $\mathscr{A}$ is a naive permutative G–category. As noted in Proposition 4.6, we then have a natural map $\iota\colon \mathscr{A} \to \iota^*\mathscr{C}at(\mathcal{E}G, \mathscr{A})$ of naive permutative G–categories. We can pass to classifying spaces and apply the functor $\mathbb{E}$ to obtain a natural map

$$\mathbb{K}\mathscr{A} \xrightarrow{\mathbb{K}\iota} \mathbb{K}\iota^*\mathscr{C}at(\mathcal{E}G, \mathscr{A}) \underset{\simeq}{\xrightarrow{\mu}} i^*\mathbb{K}_G\mathscr{C}at(\mathcal{E}G, \mathscr{A}). \tag{4.18}$$

This map is a weak equivalence when $\iota^H\colon \mathscr{A}^H \to (\iota^*\mathscr{C}at(\mathcal{E}G, \mathscr{A}))^H$ is an equivalence of categories for all $H \subset G$. The following example where this holds is important in equivariant algebraic K–theory.

Example 4.19 Let E be a Galois extension of F with Galois group G and let G act entrywise on $\mathrm{GL}(n, E)$ for $n \geq 0$. The disjoint union of the $\mathrm{GL}(n, E)$ is a naive permutative G–category that we denote by $\mathrm{GL}(E_G)$. Its product is given by the block sum of matrices. Write $\mathrm{GL}(R)$ for the nonequivariant permutative general linear category of a ring R. As we proved in [11, Example 4.20], Serre's version of Hilbert's Theorem 90 implies that

$$\iota^H\colon \mathrm{GL}(E^H) \cong \mathrm{GL}(E_G)^H \to (\iota^*\mathscr{C}at(\mathcal{E}G, \mathrm{GL}(E_G)))^H$$

is an equivalence of categories for $H \subset G$. This identifies the equivariant algebraic K–groups of E with the nonequivariant algebraic K–groups of its fixed fields E^H.

Remark 4.20 In the list above of theorems about permutative categories, a consequence of Theorem 2.25 is conspicuous by its absence. Letting $\iota_!\mathscr{A} \equiv \mathbb{P}_G \otimes_{\mathbb{P}} \mathscr{A}$ denote the left adjoint of ι^*, as defined in Remark 4.9, one might hope that $B\iota_!\mathscr{A}$ is equivalent as an $|\mathscr{P}_G|$–space to $\iota_! B\mathscr{A}$ for a naive permutative G–category $\mathscr{A}$. We do not know whether or not that is true.

5 The free $|\mathscr{P}_G|$–space generated by a G–space X

The goal of this section is to obtain a decomposition of the fixed point categories of free permutative G–categories. This decomposition will be the crux of the proof of the tom Dieck splitting theorem given in Section 5.2.

5.1 The monads $\mathbb{P}_G$ and P_G associated to $\mathscr{P}_G$

Recall that $\mathscr{P}_G$ is reduced. In fact, both $\mathscr{P}_G(0)$ and $\mathscr{P}_G(1)$ are trivial categories. As discussed for spaces in [30, Section 4], there are two monads on G–categories whose algebras are the genuine permutative G–categories. The unit object of an $\mathscr{P}_G$–category can be preassigned, resulting in a monad $\mathbb{P}_G$ on based G–categories, or it can be viewed as part of the $\mathscr{P}_G$–algebra structure, resulting in a monad $\mathbb{P}_{G+}$ on unbased G–categories. Just as in [30], these monads are related by

$$\mathbb{P}_G(\mathscr{A}_+) \cong \mathbb{P}_{G+}\mathscr{A},$$

where $\mathscr{A}_+ = \mathscr{A} \sqcup *$ is obtained from an unbased G–category $\mathscr{A}$ by adjoining a disjoint copy of the trivial G–category $*$. Explicitly,

$$\mathbb{P}_G(\mathscr{A}_+) = \coprod_{j\geq 0} \mathscr{P}_G(j) \times_{\Sigma_j} \mathscr{A}^j. \tag{5.1}$$

The term with $j = 0$ is $*$ and accounts for the copy of $*$ on the left. The unit $\eta\colon \mathscr{A} \to \mathbb{P}_G(\mathscr{A}_+)$ identifies $\mathscr{A}$ with the term with $j = 1$. The product $\mu\colon \mathbb{P}_G\mathbb{P}_G\mathscr{A}_+ \to \mathbb{P}_G\mathscr{A}_+$ is induced by the operad structure maps γ. We are only concerned with based G–categories that can be written in the form $\mathscr{A}_+$.

Since we are concerned with the precise point-set relationship between an infinite loop space machine defined on G–categories and suspension G–spectra, it is useful to think of (unbased) G–spaces X as categories. Thus we also let X denote the topological G–category with object and morphism G–space X and with I, S, T, and C all given by the identity map $X \to X$; this makes sense for C since we can identify $X \times_X X$ with X. We can also identify the classifying G–space $|X|$ with X.

By specialization of (5.1), we have an identification of (topological) G–categories

$$\mathbb{P}_G(X_+) = \coprod_{j\geq 0} \mathscr{P}_G(j) \times_{\Sigma_j} X^j. \tag{5.2}$$

The following illuminating result gives another description of $\mathbb{P}_G(X_+)$.

Proposition 5.3 *For G–spaces X, there is a natural isomorphism of genuine permutative G–categories*

$$\mathbb{P}_G(X_+) = \coprod_j \mathscr{C}at(\mathcal{E}G, \mathcal{E}\Sigma_j) \times_{\Sigma_j} X^j \to \coprod_j \mathscr{C}at(\mathcal{E}G, \mathcal{E}\Sigma_j \times_{\Sigma_j} X^j) = \mathscr{C}at(\mathcal{E}G, \mathbb{P}(X_+)).$$

Proof For each j and for $(G\times\Sigma_j)$–spaces Y, such as $Y = X^j$, we construct a natural isomorphism of $(G\times\Sigma_j)$–categories

$$\mathscr{C}at(\mathcal{E}G, \mathcal{E}\Sigma_j) \times Y \to \mathscr{C}at(\mathcal{E}G, \mathcal{E}\Sigma_j \times Y).$$

Here Y is viewed as the constant $(G\times\Sigma_j)$–category at Y. The target is

$$\mathscr{C}at(\mathcal{E}G, \mathcal{E}\Sigma_j) \times \mathscr{C}at(\mathcal{E}G, Y).$$

Since there is a map between any two objects of $\mathcal{E}G$ but the only maps in Y are identity maps $i_y\colon y \to y$ for $y \in Y$, the only functors $\mathcal{E}G \to Y$ are the constant functors c_y at $y \in Y$ and the only natural transformations between them are the identity transformations $\mathrm{id}_y\colon c_y \to c_y$. Sending y to c_y on objects and i_y to id_y on morphisms specifies an identification of $(G\times\Sigma_j)$–categories $Y \to \mathscr{C}at(\mathcal{E}G, Y)$. The product of the identity functor on $\mathscr{C}at(\mathcal{E}G, \mathcal{E}\Sigma_j)$ and this identification gives the desired natural equivalence. With $Y = X^j$, passage to orbits over Σ_j gives the j^{th} component of the claimed isomorphism of G–categories. It is an isomorphism of $\mathscr{P}_G$–categories since on both sides the action maps are induced by the structure maps of the operad $\mathscr{P}$. □

Recall that we write $\boldsymbol{P}_G$ for the monad on based G–spaces associated to the operad $|\mathscr{P}_G|$. Thus $\boldsymbol{P}_G(X_+)$ is the free $|\mathscr{P}_G|$–space generated by the G–space X.

Proposition 5.4 *For G–spaces X, there is a natural isomorphism*

$$\boldsymbol{P}_G(X_+) = \coprod_{j\geq 0} |\mathscr{P}_G(j)| \times_{\Sigma_j} X^j \cong |\mathbb{P}_G X_+|.$$

Proof For a $(G\times\Sigma_j)$–space Y viewed as a G–category, the nerve NY can be identified with the constant simplicial space Y_* with $Y_q = Y$. The nerve functor N

does not commute with passage to orbits in general, but arguing as in [11, Section 2.3] we see that

$$N(\mathscr{P}_G(j)\times_{\Sigma_j} Y)\cong (N\mathscr{P}_G(j))\times_{\Sigma_j} Y_* = N(\mathscr{P}_G(j)\times_{\Sigma_j} NY).$$

Therefore, the classifying-space functor commutes with coproducts, products, and the passage to orbits that we see here. □

5.2 The identification of $(\mathbb{P}_G X_+)^G$

The functor $|-|$ commutes with passage to G–fixed points, and we shall prove the following identification. Let $\mathbb{P}$ denote the monad on nonequivariant based categories associated to the operad $\mathscr{P}$ that defines permutative categories.

Theorem 5.5 *For G–spaces X, there is a natural equivalence of $\mathscr{P}$–categories*

$$\mathbb{P}_G(X_+)^G \simeq \prod_{(H)} \mathbb{P}(\mathcal{E}WH\times_{WH} X^H)_+,$$

where (H) runs over the conjugacy classes of subgroups of G and $WH = NH/H$.

We are regarding $\mathscr{P}$ as the suboperad $(\mathscr{P}_G)^G$ of $\mathscr{P}_G$, and the identification of categories will make clear that the identification preserves the action by $\mathscr{P}$. Of course,

$$\mathbb{P}_G(X_+)^G = \coprod_{j\geq 0}(\mathscr{P}_G(j)\times_{\Sigma_j} X^j)^G \tag{5.6}$$

and

$$\mathbb{P}(\mathcal{E}WH\times_{WH} X^H)_+ = \coprod_{k\geq 0} \mathcal{E}\Sigma_k\times_{\Sigma_k}(\mathcal{E}WH\times_{WH} X^H)^k. \tag{5.7}$$

We shall prove Theorem 5.5 by identifying both (5.6) and (5.7) with a small (but not skeletal) model $\mathscr{F}_G(X)^G$ for the category of finite G–sets over X and their isomorphisms over X. We give the relevant definitions and describe these identifications here, and we fill in the easy proofs in Sections 5.3 and 5.4.

A homomorphism $\alpha\colon G\to\Sigma_j$ is equivalent to the left action of G on the set $\boldsymbol{j}=\{1,\dots,j\}$ specified by $g\cdot i=\alpha(g)(i)$ for $i\in\boldsymbol{j}$. Similarly, an antihomomorphism $\alpha\colon G\to\Sigma_j$ is equivalent to the right action of G on $\boldsymbol{j}$ specified by $i\cdot g=\alpha(g)(i)$ or, equivalently, the left action specified by $g\cdot i=\alpha(g^{-1})(i)$; of course, if we set $\alpha^{-1}(g)=\alpha(g)^{-1}$, then α^{-1} is a homomorphism. We focus on homomorphisms and left actions, and we denote such G–spaces by $(\boldsymbol{j},\alpha)$. When we say that A is a finite G–set, we agree to mean that $A=(\boldsymbol{j},\alpha)$ for a given homomorphism $\alpha\colon G\to\Sigma_j$. That convention has the effect of fixing a small groupoid $G\mathscr{F}$ equivalent to the groupoid of all finite G–sets and isomorphisms of finite G–sets. By a j–pointed G–set, we mean a G–set with j elements.

Definition 5.8 Let X be a G–space and $j \geq 0$.

(i) Let $\mathscr{F}_G(j)$ be the G–groupoid whose objects are the j–pointed G–sets A and whose morphisms $\sigma\colon A \to B$ are the bijections, with G acting by conjugation. Then $\mathscr{F}_G(j)^G$ is the category with the same objects and with morphisms the isomorphisms of G–sets $\sigma\colon A \to B$.

(ii) Let $\mathscr{F}_G(j, X)$ be the G–groupoid whose objects are the maps (not G–maps) $p\colon A \to X$ and whose morphisms $f\colon p \to q$, for $q\colon B \to X$, are the bijections $f\colon A \to B$ such that $q \circ f = p$; the action of G is by conjugation on all maps p, q, and f. We view $\mathscr{F}_G(j, X)^G$ as the category of j–pointed G–sets over X and isomorphisms of j–pointed G–sets over X.

(iii) Let $\mathscr{F}_G = \coprod_{j\geq 0} \mathscr{F}_G(j)$ and $\mathscr{F}_G(X) = \coprod_{j \geq 0} \mathscr{F}_G(j, X)$.

(iv) Let $\mathscr{F}_G^{\mathscr{P}}(j)$ be the full G–subcategory of G–fixed objects of $\mathscr{P}_G(j)/\Sigma_j$ and let $\mathscr{F}_G^{\mathscr{P}}(j,X)$ be the full G–subcategory of G–fixed objects of $\mathscr{P}_G(j)\times_{\Sigma_j} X^j$. Then

$$\mathscr{F}_G^{\mathscr{P}}(j)^G = (\mathscr{P}_G(j)/\Sigma_j)^G \quad \text{and} \quad \mathscr{F}_G^{\mathscr{P}}(j, X)^G = (\mathscr{P}_G(j) \times_{\Sigma_j} X^j)^G.$$

In Section 5.3, we prove that the right side of (5.6) can be identified with $\mathscr{F}_G(X)^G$.

Theorem 5.9 *There is a natural isomorphism of permutative categories*

$$(\mathbb{P}_G(X_+))^G = \coprod_{j\geq 0} \mathscr{F}_G^{\mathscr{P}}(j, X)^G \cong \coprod_{j \geq 0} \mathscr{F}_G(j, X)^G = \mathscr{F}_G(X)^G.$$

We will prove an equivariant variant of this result, before passage to fixed points, in Theorem 9.6. In Section 5.4, we prove that the right side of (5.7) can also be identified with $\mathscr{F}_G(X)^G$. At least implicitly, this identification of fixed-point categories has been known since the 1970s; see for example Nishida [36, Appendix A].

Theorem 5.10 *There is a natural equivalence of categories*

$$\prod_{(H)} \coprod_{k\geq 0} \mathcal{E}\Sigma_k \times_{\Sigma_k} (\mathcal{E}WH \times_{WH} X^H)^k \to \coprod_{j\geq 0} \mathscr{F}_G(j, X)^G = \mathscr{F}_G(X)^G.$$

These two results prove Theorem 5.5.

Remark 5.11 With our specification of finite G–sets as $A = (\boldsymbol{j}, \alpha)$, the disjoint union of A and $B = (\boldsymbol{k}, \beta)$ is obtained via the obvious identification of $\boldsymbol{j} \sqcup \boldsymbol{k}$ with $\boldsymbol{j} + \boldsymbol{k}$. The disjoint union of finite G–sets over a G–space X gives $\mathscr{F}_G(X)$ a structure of naive permutative G–category. By Theorem 5.9, its fixed-point category $\mathscr{F}_G(X)^G$ is a $\mathscr{P}$–category equivalent to $(\mathbb{P}_G(X_+))^G$. One might think that $\mathscr{F}_G(X)$ is a genuine permutative G–category equivalent to the free $\mathscr{P}_G$–category $\mathscr{P}_G(X_+)$. However, its H–fixed subcategory for $H \neq G$ is not equivalent to $\mathscr{F}_H(X)^H$, and one cannot expect an action of $\mathscr{P}_G$ (or any other E_∞–G–operad) on $\mathscr{F}_G(X)$. To see the point, let G be the quaternion group of order 8: $G = \{\pm 1, \pm i, \pm j, \pm k\}$, and let $X = *$. Every nontrivial subgroup of G contains the center $H = Z = \pm 1$. Therefore, the H–set H cannot be obtained by starting with a G–set (a disjoint union of orbits G/K) and restricting along the inclusion $H \to G$.

To compare with our paper [10], we offer some alternative notation.

Definition 5.12 For an unbased G–space X, let $\mathscr{E}_G(X) = \mathscr{E}_G^{\mathscr{P}}(X) = \mathbb{P}_G(X_+)$. It is a genuine permutative G–category, and its H–fixed subcategory $\mathscr{E}_G(X)^H$ is equivalent to $\mathscr{E}_H(X)^H$ and therefore to $\mathscr{F}_H(X)^H$.

Remark 5.13 In [10], we gave a more intuitive definition of a G–category $\mathscr{E}_G(X)$. It will reappear in Section 9, where it will be given the alternative notation $\mathscr{E}_G^U(X)$. It is acted on by an E_∞–operad $\mathscr{V}_G$ of G–categories, and, again, its fixed-point category $\mathscr{E}_G^U(X)^H$ is equivalent to $\mathscr{E}_H^U(X)^H$ and therefore to $\mathscr{F}_H(X)^H$.

5.3 The proof of Theorem 5.9

We first use Theorem 3.8 to identify (5.6) when X is a point. The proof of Theorem 3.8 compares several equivalent categories, and antihomomorphisms appear naturally. To control details of equivariance, it is helpful to describe the relevant categories implicit in our operad $\mathscr{P}_G$ in their simplest forms up to isomorphism. Details are in [11, Sections 2.1, 2.2, 4.1, 4.2].

Lemma 5.14 *The objects of the chaotic $(G \times \Sigma_j)$–category $\mathscr{P}_G(j)$ are the functions $\phi\colon G \to \Sigma_j$. The (left) action of G on $\mathscr{P}_G(j)$ is given by $(g\phi)(h) = \phi(g^{-1}h)$ on objects and the diagonal action on morphisms. The (right) action of Σ_j is given by $(\phi\sigma)(h) = \phi(h)\sigma$ on objects and the diagonal action on morphisms.*

Lemma 5.15 *The objects of the G–category $\mathscr{P}_G(j)/\Sigma_j$ are the functions $\alpha\colon G \to \Sigma_j$ such that $\alpha(e) = e$. The morphisms $\sigma\colon \alpha \to \beta$ are the elements $\sigma \in \Sigma_j$, thought of as the functions $G \to \Sigma_j$ specified by $\sigma(h) = \beta(h)\sigma\alpha(h)^{-1}$. The composite of σ with $\tau\colon \beta \to \gamma$ is $\tau\sigma\colon \alpha \to \gamma$. The action of G is given on objects by*

$$(g\alpha)(h) = \alpha(g^{-1}h)\alpha(g^{-1})^{-1}.$$

In particular, $(g\alpha)(e) = e$. The action on morphisms is given by

$$g(\sigma\colon \alpha \to \beta) = \sigma\colon g\alpha \to g\beta.$$

Lemma 5.16 *For $\Lambda \subset G \times \Sigma_j$, the subcategory $\mathscr{P}_G(j)^\Lambda$ is empty if $\Lambda \cap \Sigma_j \neq e$. It is a nonempty and hence chaotic subcategory of $\mathscr{P}_G(j)$ if $\Lambda \cap \Sigma_j = e$.*

Lemma 5.17 *The objects of $(\mathscr{P}_G(j)/\Sigma_j)^G$ are the antihomomorphisms $\alpha\colon G \to \Sigma_j$. Its morphisms $\sigma\colon \alpha \to \beta$ are the conjugacy relations $\beta = \sigma\alpha\sigma^{-1}$, where $\sigma \in \Sigma_j$. For $H \subset G$, restriction of functions gives an equivalence of categories*

$$(\mathscr{P}_G(j)/\Sigma_j)^H \to (\mathscr{P}_H(j)/\Sigma_j)^H.$$

Now return to a general G–space X. To prove Theorem 5.9, it suffices to prove that $(\mathscr{P}_G(j) \times_{\Sigma_j} X^j)^G$ is isomorphic to $\mathscr{F}_G(j,X)^G$ for all j. Passage to orbits here means that for $\phi \in \mathscr{P}_G(j)$, $y \in X^j$, and $\sigma \in \Sigma_j$ (thought of as acting on the left on $\boldsymbol{j}$ and therefore on j–tuples of elements of X), $(\phi\sigma, y) = (\phi, \sigma y)$ in $\mathscr{P}_G(j) \times_{\Sigma_j} X^j$. Observe that an object $(\phi, z_1, \dots, z_j) \in \mathscr{P}_G(j) \times_{\Sigma_j} X^j$ has a unique representative in the same orbit under Σ_j of the form $(\alpha, x_1, \dots, x_j)$ where $\alpha(e) = e$. It is obtained by replacing ϕ by $\phi\tau$, where $\tau = \phi(e)^{-1}$, and replacing z_i by $x_i = z_{\tau(i)}$.

Lemma 5.18 *An object $(\alpha, y) \in \mathscr{P}_G(j) \times_{\Sigma_j} X^j$, where $\alpha(e) = e$ and $y \in X^j$, is G–fixed if and only if $\alpha\colon G \to \Sigma_j$ is an antihomomorphism and $\alpha(g^{-1})y = gy$ for all $g \in G$.*

Proof Assume that $(\alpha, y) = (g\alpha, gy)$ for all $g \in G$. Then each $g\alpha$ must be in the same Σ_j–orbit as α, where α is regarded as an object of $\mathscr{P}_G(j)$ and not $\mathscr{P}_G(j)/\Sigma_j$, so that $(g\alpha)(h) = \alpha(g^{-1}h)$. Then $(g\alpha)(h) = \alpha(h)\sigma$ for all $h \in G$ and some $\sigma \in \Pi$. Taking $h = e$ shows that $\sigma = \alpha(g^{-1})$. The resulting formula $\alpha(g^{-1}h) = \alpha(h)\alpha(g^{-1})$ implies that α is an antihomomorphism. Now

$$(\alpha, y) = (g\alpha, gy) = (\alpha\alpha(g^{-1}), gy) = (\alpha, \alpha(g)gy),$$

which means that $\alpha(g)gy = y$ and thus $gy = \alpha(g^{-1})y$. □

Use α^{-1} to define a left action of G on $\boldsymbol{j}$ and define $p\colon \boldsymbol{j} \to X$ by $p(i) = x_i$. Then the lemma shows that the G–fixed elements (α, y) are in bijective correspondence with the maps of G–sets $p\colon A \to X$, where A is a j–pointed G–set. Using Lemma 5.17, we see similarly that maps $f\colon A \to B$ of j–pointed G–sets over X correspond bijectively to morphisms in $(\mathscr{P}_G(j) \times_{\Sigma_j} X^j)^G$. These bijections specify the required isomorphism between $\mathscr{F}_G(j,X)^G$ and $(\mathscr{P}_G(j) \times_{\Sigma_j} X^j)^G$.

5.4 The proof of Theorem 5.10

This decomposition is best proven by a simple thought exercise. Every finite G–set A decomposes nonuniquely as a disjoint union of orbits G/H, and orbits G/H and G/J

are isomorphic if and only if H and J are conjugate. Choose one H in each conjugacy class. Then A decomposes uniquely as the disjoint union of the G–sets A_H, where A_H is the set of elements of A with isotropy group conjugate to H. This decomposes the category $G\mathscr{F} \equiv (\mathscr{F}_G)^G$ as the product over H of the categories $G\mathscr{F}(H)$ of finite G–sets all of whose isotropy groups are conjugate to H.

In turn, $G\mathscr{F}(H)$ decomposes uniquely as the coproduct over $k \geq 0$ of the categories $G\mathscr{F}(H,k)$ whose objects are isomorphic to the disjoint union, denoted by kG/H, of k copies of G/H. Up to isomorphism, kG/H is the only object of $G\mathscr{F}(H,k)$. The automorphism group of the G–set G/H is WH, hence the automorphism group of kG/H is the wreath product $\Sigma_k \int WH$. Viewed as a category with a single object, we may identify this group with the category $\mathcal{E}\Sigma_k \times_{\Sigma_k} (WH)^k$. This proves the following result.

Proposition 5.19 *The category $G\mathscr{F}$ is equivalent to the category*

$$\prod_{(H)} \coprod_{k \geq 0} \mathcal{E}\Sigma_k \times_{\Sigma_k} (WH)^k.$$

The displayed category is a skeleton of $G\mathscr{F}$. As written, its objects are sets of numbers $\{k_H\}$, one for each (H), but they are thought of as the finite G–sets $\coprod_H k_H G/H$. Its morphism groups specify the automorphisms of these objects. On objects, the equivalence sends a finite G–set A to the unique finite G–set of the form $\coprod_{(H)} kG/H$ in the same isomorphism class as A. Via chosen isomorphisms, this specifies the inverse equivalence to the inclusion of the chosen skeleton in $G\mathscr{F}$.

We parametrize this equivalence to obtain a description of the category $G\mathscr{F}(X)$ of finite G–sets over X. Given any H and k, a k–tuple of elements $\{x_1, \dots, x_k\}$ of X^H determines the G–map $p\colon kG/H \to X$ that sends eH in the i^{th} copy of G/H to x_i, and it is clear that every finite G–set A over X is isomorphic to one of this form. Similarly, for a finite G–set $q\colon B \to X$ over X and an isomorphism $f\colon A \to B$, we have that f is an isomorphism over X from q to $p = q \circ f$, and every isomorphism over X can be constructed in this fashion. Since we may as well choose A and B to be in our chosen skeleton of $G\mathscr{F}$, this argument proves Theorem 5.10.

6 The Barratt–Priddy–Quillen and tom Dieck splitting theorems

6.1 The Barratt–Priddy–Quillen theorem revisited

The BPQ theorem shows how to model suspension G–spectra in terms of free E_∞–G–categories and G–spaces. It is built tautologically into the equivariant infinite loop space machine in the same way as it is nonequivariantly [22, Theorem 2.3(vii)] or [30, Section 10]. The following result works for either $\mathbb{E}_G = \mathbb{E}_G^{\mathscr{S}p}$ or $\mathbb{E}_G = \mathbb{E}_G^{\mathscr{S}}$, but note that the interpretation of both the source and target are different in the two cases. The proof shows consistency with the versions of the BPQ theorem in Theorems 1.31 and 2.18.

Theorem 6.1 (the E_∞–operad BPQ theorem) *For an E_∞–operad $\mathscr{C}_G$ of G–spaces and based G–spaces X, there is a natural weak equivalence of G–spectra*

$$\Sigma_G^\infty X \to \mathbb{E}_G \boldsymbol{C}_G X.$$

Proof For $\mathbb{E}_G^{\mathscr{S}p}$, recall that $\mathscr{C}_U = \mathscr{K}_U \times \mathscr{C}_G$. The same formal argument as for Theorem 2.18 and use of the projections to $\mathscr{C}_G$ and to $\mathscr{K}_U$ give equivalences of LM G–spectra:

$$\begin{tikzcd}
\Sigma_G^\infty X \arrow[r] \arrow[dr] & B(\Sigma_G^\infty, \boldsymbol{C}_U, \boldsymbol{C}_U X) \arrow[r] \arrow[d] & B(\Sigma_G^\infty, \boldsymbol{C}_U, \boldsymbol{C}_G X) \\
& B(\Sigma_G^\infty, \boldsymbol{K}_U, \boldsymbol{K}_U X) &
\end{tikzcd}$$

For $\mathbb{E}_G^{\mathscr{S}}$, recall that $\mathscr{C}_V = \mathscr{K}_V \times \mathscr{C}_G$. Analogously to Theorem 1.31, there is an orthogonal G–spectrum with V^{th} space $B(\Sigma^V, \boldsymbol{C}_V, \boldsymbol{C}_V X)$. The usual formal argument and the projections to $\mathscr{C}_G$ and $\mathscr{K}_V$ give diagrams

$$\begin{tikzcd}
\Sigma^V X \arrow[r] \arrow[dr] & B(\Sigma^V, \boldsymbol{C}_V, \boldsymbol{C}_V X) \arrow[r] \arrow[d] & B(\Sigma^V, \boldsymbol{C}_V, \boldsymbol{C}_G X) \\
& B(\Sigma^V, \boldsymbol{K}_V, \boldsymbol{K}_V X) &
\end{tikzcd}$$

for all V in which the left horizontal arrow and the vertical arrow are level equivalences of orthogonal G–spectra, and the right horizontal arrow is a weak equivalence (π_*–isomorphism) of orthogonal G–spectra, as we see by forgetting to G–prespectra and passing to colimits over $V \subset U$, where U is a complete G–universe. □

Taking $Y = X_+$ for an unbased G–space X and using (5.2), we can rewrite this version of the BPQ theorem using the infinite loop space machine defined on permutative G–categories.

Theorem 6.2 (the categorical BPQ theorem: first version) *For unbased G–spaces X, there is a natural weak equivalence of G–spectra*

$$\Sigma_G^\infty X_+ \to \mathbb{K}_G \mathbb{P}_G(X_+).$$

Remark 6.3 Diagrams showing compatibility with smash products, like those in Theorems 1.31 and 2.18 are conspicuous by their absence from Theorems 6.1 and 6.2. A previous version of this article erroneously claimed that the operad $\mathscr{P}$ has a self pairing $(\mathscr{P},\mathscr{P}) \to \mathscr{P}$ induced by the homomorphisms

(6.4) $$\otimes\colon \Sigma_j \times \Sigma_k \to \Sigma_{jk},$$

which are made precise in Appendix A by use of lexicographic ordering. However, these do not satisfy the condition in Definition A.1(iii); see Counterexample A.5. For a conceptual understanding of why $\mathscr{P}$ cannot have a self-pairing, consider the free $\mathscr{P}$–algebra $\mathbb{P}(S^0)$. This is a model for the groupoid of finite sets. As explained in [26, Appendix A], a self-pairing on $\mathscr{P}$ would give strict distributivity on both sides in $\mathbb{P}(S^0)$. But the lexicographic ordering on $\boldsymbol{j} \times (\boldsymbol{k} \sqcup \boldsymbol{m})$ does not agree with the lexicographic ordering on $(\boldsymbol{j} \times \boldsymbol{k}) \sqcup (\boldsymbol{j} \times \boldsymbol{m})$.

As we explain in [13], the homomorphisms $\otimes$ exhibit a product that exists in any operad. The categorical operads $\mathscr{P}$ and $\mathscr{P}_G$ are *pseudocommutative*, meaning that certain diagrams of functors defined using these products commute up to natural isomorphism. Putting together Theorem 6.2, the comparison of operadic and Segalic machines in [31], and 2–category machinery developed in [12], we will obtain multicategorical generalizations of the missing diagrams in [13], where we complete the proofs from equivariant infinite loop space theory promised in [10].

6.2 The tom Dieck splitting theorem

The G–fixed-point spectra of suspension G–spectra have a well-known splitting. It is due to tom Dieck [5] on the level of homotopy groups and was lifted to the spectrum level in [18, Section V.11]. The tom Dieck splitting actually works for all compact Lie groups G, but we have nothing helpful to add in that generality. Our group G is always finite. In that case, we have already given the ingredients for a new categorical proof, as we now explain.

Theorem 6.5 *For a based G–space Y,*

$$(\Sigma^\infty_G Y)^G \simeq \bigvee_{(H)} \Sigma^\infty (EWH_+ \wedge_{WH} Y^H).$$

The wedge runs over the conjugacy classes of subgroups H of G, and $WH = NH/H$.

Theorem 6.5 and the evident natural identifications

(6.6) $$EWH_+ \wedge_{WH} X^H_+ \cong (EWH \times_{WH} X^H)_+$$

imply the following version for unbased G–spaces X.

Theorem 6.7 *For an unbased G–space X,*

$$(\Sigma^\infty_G X_+)^G \simeq \bigvee_{(H)} \Sigma^\infty (EWH \times_{WH} X^H)_+.$$

Conversely, we can easily deduce Theorem 6.5 from Theorem 6.7. Viewing S^0 as $\{1\}_+$ with trivial G action, our standing assumption that basepoints are nondegenerate gives a based G–cofibration $S^0 \to Y_+$ that sends 1 to the basepoint of Y, and $Y = Y_+/S^0$. The functors appearing in Theorem 6.7 preserve cofiber sequences, and the identifications (6.6) imply identifications

$$(EWH \times_{WH} Y^H)_+/(EWH \times_{WH} \{1\})_+ \cong EWH_+ \wedge_{WH} Y^H. \tag{6.8}$$

Therefore, Theorem 6.7 implies Theorem 6.5.

We explain these splittings in terms of the categorical BPQ theorem. We begin in the based setting. The nonequivariant case $G = e$ of the BPQ theorem relates to the equivariant case through Theorem 2.20. Explicitly, Theorems 2.20 and 6.1 give a pair of weak equivalences

$$(\Sigma^\infty_G Y)^G \to (\mathbb{E}_G \boldsymbol{C}_G Y)^G \leftarrow \mathbb{E}((\boldsymbol{C}_G Y)^G). \tag{6.9}$$

Since the functor Σ^∞ commutes with wedges, the nonequivariant BPQ theorem gives a weak equivalence

$$\bigvee_{(H)} \Sigma^\infty (EWH_+ \wedge_{WH} Y^H) \to \mathbb{E}\boldsymbol{C}\big(\bigvee_{(H)} (EWH_+ \wedge_{WH} Y^H)\big). \tag{6.10}$$

If we could prove that there is a natural weak equivalence of $\mathscr{C}$–spaces

$$(\boldsymbol{C}_G Y)^G \simeq \boldsymbol{C}\big(\bigvee_{(H)} (EWH_+ \wedge_{WH} Y^H)\big),$$

that would imply a natural weak equivalence

$$\mathbb{E}((\boldsymbol{C}_G Y)^G) \simeq \mathbb{E}\boldsymbol{C}\big(\bigvee_{(H)} (EWH_+ \wedge_{WH} Y^H)\big) \tag{6.11}$$

and complete the proof of Theorem 6.5. However, the combinatorial study of the behavior of C on wedges is complicated by the obvious fact that wedges of based spaces do not commute with products.

We use the following consequence of Theorem 5.5 and the relationship between wedges and products of spectra to get around this. Recall that $\boldsymbol{P}_G$ is the monad on based G–spaces obtained from the operad $|\mathscr{P}_G|$ of G–spaces.

Theorem 6.12 *For unbased G–spaces X, there is a natural equivalence of $|\mathscr{P}|$–spaces*

$$(\boldsymbol{P}_G X_+)^G \simeq \prod_{(H)} \boldsymbol{P}(EWH \times_{WH} X^H)_+,$$

where (H) runs over the conjugacy classes of subgroups of G and $WH = NH/H$.

Proof Remembering that $|\mathcal{E}G| = EG$, we see that the classifying space of the category $\mathcal{E}WH \times_{WH} X^H$ can be identified with $EWH \times_{WH} X^H$. The commutation relations between $|-|$ and the constituent functors used to construct the monads $\boldsymbol{P}_G$ on G–spaces and $\mathbb{P}_G$ on G–categories make the identification clear. □

Remark 6.13 Of course, we can and must replace $\mathscr{P}_G$ and $\mathscr{P}$ by their products with the equivariant and nonequivariant Steiner operad to fit into the infinite loop space machine. There is no harm in doing so since if we denote the product operads by $\mathscr{O}_G$ and $\mathscr{O}$, as before, the projections $\mathscr{O}_G \to \mathscr{P}_G$ and $\mathscr{O} \to \mathscr{P}$ induce weak equivalences of monads that fit into the following commutative diagram:

$$\begin{tikzcd}
(\boldsymbol{O}_G X_+)^G \arrow[r, "\simeq"] \arrow[d, "\simeq"'] & \prod_{(H)} \boldsymbol{O}(EWH \times_{WH} X^H)_+ \arrow[d, "\simeq"] \\
(\boldsymbol{P}_G X_+)^G \arrow[r, "\simeq"] & \prod_{(H)} \boldsymbol{P}(EWH \times_{WH} X^H)_+
\end{tikzcd}$$

The functor Σ^∞_G commutes with wedges, and the natural map of G–spectra

$$E \vee F \to E \times F$$

is a weak equivalence. Theorems 2.21 and 6.1 have the following implication. We state it equivariantly, but we shall apply its nonequivariant special case.

Proposition 6.14 *For based G–spaces X and Y, the natural map*

$$\mathbb{E}_G \boldsymbol{O}_G(X \vee Y) \to \mathbb{E}_G(\boldsymbol{O}_G X \times \boldsymbol{O}_G Y)$$

is a weak equivalence of G–spectra.

Proof The following diagram commutes by the universal property of products:

$$\begin{tikzcd}
\Sigma^\infty_G(X \vee Y) \arrow[r] \arrow[d, "\cong"'] & \mathbb{E}_G \boldsymbol{O}_G(X \vee Y) \arrow[d] \\
\Sigma^\infty_G X \vee \Sigma^\infty_G Y \arrow[d] & \mathbb{E}_G(\boldsymbol{O}_G X \times \boldsymbol{O}_G Y) \arrow[d] \\
\Sigma^\infty_G X \times \Sigma^\infty_G Y \arrow[r] & \mathbb{E}_G \boldsymbol{O}_G X \times \mathbb{E}_G \boldsymbol{O}_G Y
\end{tikzcd}$$

All arrows except the upper right vertical one are weak equivalences, hence that arrow is also a weak equivalence. □

For any nonequivariant E_∞–operad $\mathscr{C}$, we therefore have a weak equivalence

$$\mathbb{E}\boldsymbol{C}\big(\textstyle\bigvee_{(H)}(EWH_+ \wedge_{WH} Y^H)\big) \to \mathbb{E}\displaystyle\prod_{(H)} \boldsymbol{C}(EWH_+ \wedge_{WH} Y^H). \tag{6.15}$$

Together with (6.15), Theorem 6.12 and Remark 6.13 give a weak equivalence (6.11) in the case $Y = X_+$. Together with (6.9) and (6.10), this completes the proof of Theorem 6.7, and Theorem 6.5 follows.

7 The E_∞–operads $\mathscr{V}_G$, $\mathscr{V}_G^\times$, and $\mathscr{W}_G$

The operad $\mathscr{P}_G$ has a privileged conceptual role, but there are other categorical E_∞–G–operads with different good properties. We define three interrelated examples. The objects of the chaotic category $\mathscr{P}_G(j)$ are functions $G \to \Sigma_j$. We give analogous chaotic G–categories in which the objects are suitable functions between well chosen infinite G–sets, with G again acting by conjugation. Their main advantage over $\mathscr{P}_G$ is that it is easier to recognize G–categories on which they act.

7.1 The definitions of $\mathscr{V}_G$ and $\mathscr{V}_G^\times$

We start with what we would like to take as a particularly natural choice for the j^{th} category of an E_∞–G–operad. It is described in more detail in [11, Section 6.1].

Definition 7.1 Let U be a countable ambient G–set that contains countably many copies of each orbit G/H. Let U^j be the product of j copies of U with diagonal action by G, and let ${}^jU$ be the disjoint union of j copies of the G–set U. Here U^0 is a one-point set, sometimes denoted by 1, and ${}^0U$ is the empty set, sometimes denoted by $\varnothing$ and sometimes denoted by 0.

Let $\boldsymbol{j} = \{1, \dots, j\}$ with its natural left action by Σ_j, written $\sigma\colon \boldsymbol{j} \to \boldsymbol{j}$.

Definition 7.2 For $j \geq 0$, let $\widetilde{\mathscr{E}}_G^U(j)$ be the chaotic $(G \times \Sigma_j)$–category whose objects are the pairs (A, ι), where A is a j–element subset of U and $\iota\colon \boldsymbol{j} \to A$ is a bijection. The group G acts on objects by $g(A, \iota) = (gA, g\iota)$, where $(g\iota)(i) = g \cdot \iota(i)$. The group Σ_j acts on objects by $(A, \iota)\sigma = (A, \iota \circ \sigma)$ for $\sigma \in \Sigma_j$. Since $\widetilde{\mathscr{E}}_G^U(j)$ is chaotic, this determines the actions on morphisms.

Proposition 7.3 [11, Proposition 6.3] *For each j, the classifying space $|\widetilde{\mathscr{E}}_G^U(j)|$ is a universal principal (G, Σ_j)–bundle.*

Therefore, $\widetilde{\mathscr{E}}_G^U(j)$ satisfies the properties required of the j^{th} category of an E_∞–G–operad. However, these categories as j varies do not form an operad. The problem is a familiar one. These categories can be thought of as analogous to configuration spaces. Just as we fattened up the configuration space models of Section 1.6 to the little discs operads of Section 1.1, we must fatten up these categories to provide enough room for an operad structure.

Definition 7.4 We define a reduced operad $\mathscr{V}_G$ of G–categories. Let $\mathscr{V}_G(j)$ be the chaotic G–category whose set of objects is the set of injective functions ${}^jU \to U$. Let G act by conjugation and let Σ_j have the right action induced by its left action on ${}^jU$. Let $\mathrm{id} \in \mathscr{V}_G(1)$ be the identity function $U \to U$. Define

$$\gamma\colon \mathscr{V}_G(k) \times \mathscr{V}_G(j_1) \times \cdots \times \mathscr{V}_G(j_k) \to \mathscr{V}_G(j),$$

where $j = j_1 + \cdots + j_k$, to be the composite

$$\mathscr{V}_G(k) \times \mathscr{V}_G(j_1) \times \cdots \times \mathscr{V}_G(j_k) \to \mathscr{V}_G(k) \times \mathscr{V}_G({}^jU, {}^kU) \to \mathscr{V}_G(j)$$

obtained by first taking coproducts of maps and then composing. Here $\mathscr{V}({}^jU, {}^kU)$ is the set of injections ${}^jU \to {}^kU$. The operad axioms [21, Definition 1.1] are easily verified.

Remembering that taking sets to the free $\mathbb{R}$–modules they generate gives a coproduct-preserving functor from sets to $\mathbb{R}$–modules, we see that $\mathscr{V}_G$ is a categorical analogue of the linear isometries operad $\mathscr{L}_U$.

There is a parallel definition that uses products instead of coproducts.

Definition 7.5 We define an unreduced operad $\overline{\mathscr{V}}_G^\times$ of G–categories. Let $\overline{\mathscr{V}}_G^\times(j)$ be the chaotic G–category whose set of objects is the set of injective functions $U^j \to U$. Let G act by conjugation and let Σ_j have the right action induced by its left action on U^j. Let $\mathrm{id} \in \overline{\mathscr{V}}_G^\times(1)$ be the identity function. Define

$$\gamma\colon \overline{\mathscr{V}}_G^\times(k) \times \overline{\mathscr{V}}_G^\times(j_1) \times \cdots \times \overline{\mathscr{V}}_G^\times(j_k) \to \overline{\mathscr{V}}_G^\times(j),$$

where $j = j_1 + \cdots + j_k$, to be the composite

$$\overline{\mathscr{V}}_G^\times(k) \times \overline{\mathscr{V}}_G^\times(j_1) \times \cdots \times \overline{\mathscr{V}}_G^\times(j_k) \to \overline{\mathscr{V}}_G^\times(k) \times \overline{\mathscr{V}}_G^\times(U^j, U^k) \to \overline{\mathscr{V}}_G^\times(j)$$

obtained by first taking products of maps and then composing. Here $\overline{\mathscr{V}}_G^\times(U^j, U^k)$ is the set of injections $U^j \to U^k$. Again, the operad axioms are easily verified.

Observe that the objects of $\bar{\mathscr{V}}_G^{\times}(0)$ are the injections from the point U^0 into U and can be identified with the set U, whereas $\mathscr{V}_G(0)$ is the trivial category given by the injection of the empty set ${}^0U$ into U. As in Remark 2.4, the objects of the zeroth category give unit objects for operad actions, and it is convenient to restrict attention to a reduced variant of $\bar{\mathscr{V}}_G^{\times}$.

Definition 7.6 Choose a G–fixed point $1 \in U$ (or, equivalently, adjoin a G–fixed basepoint 1 to U) and also write 1 for the single point in U^0. Give U^j, $j \geq 0$, the basepoint whose coordinates are all 1. The reduced variant of $\bar{\mathscr{V}}_G^{\times}$ is the operad $\mathscr{V}_G^{\times}$ of G–categories that is obtained by restricting the objects of the $\bar{\mathscr{V}}_G^{\times}(j)$ to consist only of the basepoint-preserving injections $U^j \to U$ for all $j \geq 0$.

Remark 7.7 If $\bar{\mathscr{V}}_G^{\times}$ acts on a category $\mathscr{A}$, then $\mathscr{V}_G^{\times}$ acts on $\mathscr{A}$ by restriction of the action. However, $\mathscr{V}_G^{\times}$ can act even though $\bar{\mathscr{V}}_G^{\times}$ does not. This happens when the structure of $\mathscr{A}$ encodes a particular unit object and the operad action conditions fail for other choices of objects in $\mathscr{A}$.

Proposition 7.8 *The classifying spaces $|\mathscr{V}_G(j)|$, $|\bar{\mathscr{V}}_G^{\times}(j)|$, and $|\mathscr{V}_G^{\times}(j)|$ are universal principal (G, Σ_j)–bundles; hence $\mathscr{V}_G$, $\bar{\mathscr{V}}_G^{\times}$, and $\mathscr{V}_G^{\times}$ are E_∞–operads.*

Proof Since the objects of our categories are given by injective functions, Σ_j acts freely on the objects of $\mathscr{V}_G(j)$ and $\mathscr{V}_G^{\times}(j)$. Since our categories are chaotic, it suffices to show that if $\Lambda \cap \Sigma_j = \{e\}$, where $\Lambda \subset G \times \Sigma_j$, then the object sets $\mathscr{V}_G(j)^\Lambda$ and $\mathscr{V}_G^{\times}(j)^\Lambda$ are nonempty. This means that there are Λ–equivariant injections ${}^jU \to U$ and $U^j \to U$, and in fact, there are Λ–equivariant bijections. We have $\Lambda = \{(h, \alpha(h)) \mid h \in H\}$ for a subgroup H of G and a homomorphism $\alpha\colon H \to \Sigma_j$, and we may regard U as an H–set via the canonical isomorphism $H \cong \Lambda$. Since countably many copies of every orbit of H embed in U, ${}^jU$, and U^j for $j \geq 1$, these sets are all isomorphic as H–sets and therefore as Λ–sets. □

7.2 The definition of $\mathscr{W}_G$ and its action on $\mathscr{V}_G$

This section is parenthetical, aimed towards work in progress on a new version of multiplicative infinite loop space theory. The notion of an action of a *multiplicative* operad $\mathscr{G}$ on an *additive* operad $\mathscr{C}$ was defined in [24, VI.1.6], and $(\mathscr{C}, \mathscr{G})$ was then said to be an *operad pair*. This notion was redefined and discussed in [30; 29]. Expressed in terms of diagrams rather than elements, it makes sense for operads in any cartesian monoidal category, such as the categories of G–categories and of G–spaces. As is emphasized in the cited papers, although this notion is the essential starting point for the theory of E_∞–ring spaces, the only interesting nonequivariant example we know is $(\mathscr{K},\mathscr{L})$, where $\mathscr{K}$ is the Steiner operad. As pointed out in Section 2.5, this example works equally well equivariantly.

The pair of operads $(\mathscr{V}_G,\mathscr{V}_G^\times)$ very nearly gives another example, but we must shrink $\mathscr{V}_G^\times$ and drop its unit object to obtain this.

Definition 7.9 Define $\mathscr{W}_G\subset\mathscr{V}_G^\times$ to be the suboperad such that $\mathscr{W}_G(j)$ is the full subcategory of $\mathscr{V}_G^\times(j)$ whose objects are the based bijections $U^j\to U$. In particular, $\mathscr{W}_G(0)$ is the empty category, so that the operad $\mathscr{W}_G$ does not encode unit object information. By the proof of Proposition 7.8, $\mathscr{W}_G(j)$ for $j\geq 1$ is again a universal principal (G,Σ_j)–bundle. We view $\mathscr{W}_G$ as a restricted E_∞–operad, namely one without unit objects.

Proposition 7.10 *The restricted operad $\mathscr{W}_G$ acts on the operad $\mathscr{V}_G$.*

Proof We must specify action maps

$$\lambda\colon \mathscr{W}_G(k)\times\mathscr{V}_G(j_1)\times\cdots\times\mathscr{V}_G(j_k)\to\mathscr{V}_G(j),$$

where $j=j_1\cdots j_k$ and $k\geq 1$. To define them, consider the set of sequences $I=\{i_1,\dots,i_k\}$, ordered lexicographically, where $1\leq i_r\leq j_r$ and $1\leq r\leq k$. For an injection $\phi_r\colon {}^{j_r}U\to U$, let $\phi_{i_r}\colon U\to U$ denote the restriction of ϕ_r to the i_r^{th} copy of U in ${}^{j_r}U$. Then let

$$\phi_I=\phi_{i_1}\times\cdots\times\phi_{i_k}\colon U^k\to U^k.$$

For a bijection $\psi\colon U^k\to U$, define

$$\lambda(\psi;\phi_1,\dots,\phi_k)\colon {}^jU\to U$$

to be the injection which restricts on the I^{th} copy of U to the composite

$$U\xrightarrow{\psi^{-1}}U^k\xrightarrow{\phi_I}U^k\xrightarrow{\psi}U.$$

It is tedious but straightforward to verify that all conditions specified in [24, Definition VI.1.6], [29, Definition 4.2] that make sense are satisfied.[6] □

Remark 7.11 When all $j_i=1$, so that there is only one sequence I, we can define λ more generally, with $\mathscr{V}_G^\times(k)$ replacing $\mathscr{W}_G(k)$, by letting

$$\lambda(\psi;\phi_1,\dots,\phi_k)\colon U\to U$$

[6] In fact, with the details of [29, Definition 4.2], the only condition that does not make sense would require $\lambda(1)=\mathrm{id}\in\mathscr{V}_G(1)$, where $\{1\}=\mathscr{W}(0)$, and that condition lacks force since it does not interact with the remaining conditions.

be the identity on the complement of the image of the injection $\psi\colon U^k \to U$ and

$$\psi(U) \xrightarrow{\psi^{-1}} U^k \xrightarrow{\phi_I} U^k \xrightarrow{\psi} \psi(U)$$

on the image of ψ. Clearly we can replace $\mathscr{V}_G(1)$ by $\mathscr{V}_G^{\times}(1)$ here.

This allows us to give the following speculative analogue of Definition 4.10. An E_∞–ring space is defined to be a $(\mathscr{C},\mathscr{G})$–space, where $(\mathscr{C},\mathscr{G})$ is an operad pair such that $\mathscr{C}$ and $\mathscr{G}$ are E_∞–operads of spaces. Briefly, a $(\mathscr{C},\mathscr{G})$–space X is a $\mathscr{C}$–space and a $\mathscr{G}$–space with respective basepoints 0 and 1 such that 0 is a zero element for the $\mathscr{G}$–action and the action $\boldsymbol{C}X \to X$ is a map of $\mathscr{G}$–spaces with zero, where $\boldsymbol{C}$ denotes the monad associated to the operad $\mathscr{C}$. Here the action of $\mathscr{G}$ on $\mathscr{C}$ induces an action of $\mathscr{G}$ on the free $\mathscr{C}$–spaces $\boldsymbol{C}X$, so that $\boldsymbol{C}$ restricts to a monad in the category of $\mathscr{G}$–spaces. These notions are redefined in the more recent papers [30; 29]. The definitions are formal and apply equally well to spaces, G–spaces, categories, and G–categories.

Definition 7.12 An E_∞–ring G–category $\mathscr{A}$ is a G–category together with an action by the E_∞–operad pair $(\mathscr{V}_G,\mathscr{W}_G)$ such that the multiplicative action extends from the restricted E_∞–operad $\mathscr{W}_G$ to an action of the E_∞–operad $\mathscr{V}_G^{\times}$.

The notion of a bipermutative category, or symmetric strict bimonoidal category, was specified in [24, Definition VI.3.3]. With the standard skeletal model, the direct sum and tensor product on the category of finite-dimensional free modules over a commutative ring R gives a typical example. Without any categorical justification, we allow ourselves to think of E_∞–ring G–categories as an E_∞ version of genuine operadic bipermutative G–categories. A less concrete but more general version of this notion is defined and developed in [13].

Our notion of an E_∞–G–category $\mathscr{A}$ implies that $B\mathscr{A}$ is an E_∞–G–space. We would like to say that our notion of an E_∞–ring G–category $\mathscr{A}$ implies that $B\mathscr{A}$ is an E_∞–ring G–space, but that is not quite true. However, we believe there is a way to prove the following conjecture that avoids the categorical work of [8; 12; 13; 26; 29]. However, that proof is work in progress.

Conjecture 7.13 *There is an infinite loop space machine that carries E_∞–ring G–categories to E_∞–ring G–spectra.*

8 Examples of E_∞ and E_∞–ring G–categories

We have several interesting examples. We emphasize that these particular constructions are new even when $G = e$. In that case, we may take U to be the set of positive integers, with 1 as basepoint.

We have the notion of a genuine permutative G–category, which comes with a preferred product, and the notion of a $\mathscr{V}_G$–category, which does not. It seems plausible that the latter notion is more general, but to verify that we would have to show how to regard a permutative category as a $\mathscr{V}_G$–algebra. One natural way to do so would be to construct a map of operads $\mathscr{V}_G \to \mathscr{P}_G$, but we do not know how to do that. Of course, the equivalence of $\mathscr{V}_G$–categories and $\mathscr{P}_G$–categories shows that genuine permutative categories give a plethora of examples of $\mathscr{V}_G$–algebras up to homotopy. However, the most important examples can easily be displayed directly, without recourse to the theory of permutative categories.

8.1 The G–category $\mathscr{E}_G^U = \mathscr{E}_G^{\mathscr{V}}$ of finite sets

Recall Remark 5.13. Intuitively, we would like to have a genuine permutative G–category whose product is given by disjoint unions of finite sets, with G relating finite sets (not G–sets) by translations. Even nonequivariantly, this is imprecise due to both size issues and the fact that categorical coproducts are not strictly associative. We make it precise by taking coproducts of finite subsets of our ambient G–set U, but we must do so without assuming that our given finite subsets are disjoint. We achieve this by using injections ${}^jU \to U$ to separate them. We do not have canonical choices for the injections, hence we have assembled them into our categorical E_∞–operad $\mathscr{V}_G$. Recall Definition 7.2 and Proposition 7.3.

Definition 8.1 The G–category $\widetilde{\mathscr{E}}_G^U$ of finite ordered sets is the coproduct over $n \geq 0$ of the G–categories $\widetilde{\mathscr{E}}_G^U(n)$. The G–category $\mathscr{E}_G^U = \mathscr{E}_G^{\mathscr{V}}$ of finite sets is the coproduct over $n \geq 0$ of the orbit categories $\widetilde{\mathscr{E}}_G^U(n)/\Sigma_n$. By Proposition 7.3, $B\mathscr{E}_G^U$ is the coproduct over $n \geq 0$ of classifying spaces $B(G, \Sigma_n)$. Explicitly, by [11, Lemma 6.5], the objects of $\mathscr{E}_G^U$ are the finite subsets (not G–subsets) A of U. Its morphisms are the bijections $\nu\colon A \to B$; if A has n points, the morphisms $A \to A$ give a copy of the set Σ_n. The group G acts by translation on objects, so that $gA = \{ga \mid a \in A\}$, and by conjugation on morphisms, so that $g\nu\colon gA \to gB$ is given by $(g\nu)(g \cdot a) = g \cdot \nu(a)$.

Proposition 8.2 *The G–categories $\widetilde{\mathscr{E}}_G^U$ and $\mathscr{E}_G^U$ are $\mathscr{V}_G$–categories, and passage to orbits over symmetric groups defines a map $\widetilde{\mathscr{E}}_G^U \to \mathscr{E}_G^U$ of $\mathscr{V}_G$–categories.*

Proof Define a G–functor

$$\theta_j\colon \mathscr{V}_G(j) \times (\mathscr{E}_G^U)^j \to \mathscr{E}_G^U$$

as follows. On objects, for $\phi \in \mathscr{V}_G(j)$ and $A_i \in \operatorname{Ob}\mathscr{E}_G^U$, $1 \leq i \leq j$, define

$$\theta_j(\phi; A_1, \ldots, A_j) = \phi(A_1 \sqcup \cdots \sqcup A_j),$$

where A_i is viewed as a subset of the i^{th} copy of U in ${}^jU$. For a morphism

$$(\zeta; \nu_1, \dots, \nu_j)\colon (\phi; A_1, \dots, A_j) \to (\psi; B_1, \dots, B_j),$$

where $\zeta\colon \phi \to \psi$ is the unique morphism, define $\theta_j(\zeta; \nu_1, \dots, \nu_j)$ to be the unique bijection that makes the following diagram commute:

$$\begin{tikzcd}
A_1 \sqcup \cdots \sqcup A_j \arrow[r, "\phi"] \arrow[d, "\nu_1 \sqcup \cdots \sqcup \nu_j"'] & \phi(A_1 \sqcup \cdots \sqcup A_j) \arrow[d, "\theta_j(\zeta; \nu_1, \dots, \nu_j)"] \\
B_1 \sqcup \cdots \sqcup B_j \arrow[r, "\psi"'] & \psi(B_1 \sqcup \cdots \sqcup B_j)
\end{tikzcd}$$

Then the θ_j specify an action of $\mathscr{V}_G$ on $\mathscr{E}_G^U$.

Since the $\widetilde{\mathscr{E}}_G^U(n)$ are chaotic, to define an action of $\mathscr{V}_G$ on $\widetilde{\mathscr{E}}_G^U$ we need only specify the required G–functors

$$\widetilde{\theta}_j\colon \mathscr{V}_G(j) \times (\widetilde{\mathscr{E}}_G^U)^j \to \widetilde{\mathscr{E}}_G^U$$

on objects. A typical object has the form $(\phi; (A_1, \iota_1), \dots, (A_j, \iota_j))$ for $\iota_i\colon \boldsymbol{n}_i \to A_i$. We have the canonical isomorphism $\boldsymbol{n}_1 \sqcup \cdots \sqcup \boldsymbol{n}_j \cong \boldsymbol{n}$, where $n = n_1 + \cdots + n_j$, and $\widetilde{\theta}_j$ sends our typical object to

$$(\phi(A_1 \sqcup \cdots \sqcup A_j), \phi \circ (\iota_1 \sqcup \cdots \sqcup \iota_j)).$$

Again, the $\widetilde{\theta}_j$ specify an action. The compatibility with passage to orbits is verified by use of canonical orbit representatives for objects A that are obtained by choosing fixed reference maps $\eta_A\colon \boldsymbol{n} \to A$ for each n–point set $A \subset U$; compare [11, Proposition 6.3 and Lemma 6.5]. □

Remark 8.3 If we restrict to the full G–subcategory of $\mathscr{E}_G^U$ of G–fixed sets A of cardinality n, we obtain an equivalent analogue of the category $\mathscr{F}_G(n)$ of Definition 5.8: these are two small models of the G–category of all G–sets with n elements and the bijections between them, and they have isomorphic skeleta. Thus the restriction of $\mathscr{E}_G^U$ to its full G–subcategory of G–fixed sets A is an equivalent analogue of $\mathscr{F}_G$. Remember from Remark 5.11 that no E_∞–operad can be expected to act on $\mathscr{F}_G$. The $\mathscr{V}_G$–category $\mathscr{E}_G^U$ gives a convenient substitute.

8.2 The G–category $\mathscr{GL}_G(R)$ for a G–ring R

Let R be a G–ring, that is a ring with an action of G through automorphisms of R. We have analogues of Definitions 7.2 and 8.1 that can be used in equivariant algebraic K–theory. For a set A, let $R[A]$ denote the free R–module on the basis A. Let G act entrywise on the matrix group $\mathrm{GL}(n, R)$ and diagonally on R^n. Our conventions on semidirect products and their universal principal $(G, \mathrm{GL}(n, R)_G)$–bundles are in [11], and [11, Section 6.3] gives more details on the following definitions.

Definition 8.4 We define the chaotic general linear category $\widetilde{\mathscr{GL}}_G(n, R)$. Its objects are the monomorphisms of (left) R–modules $\iota\colon R^n \to R[U]$. The group G acts on objects by $g\iota = g\circ\iota\circ g^{-1}$. The group $\mathrm{GL}(n, R)$ acts on objects by $\iota\tau = \iota\circ\tau\colon R^n \to R[U]$. Since $\widetilde{\mathscr{GL}}_G(n, R)$ is chaotic, this determines the actions on morphisms.

Proposition 8.5 [11, Proposition 6.18] *The actions of G and $\mathrm{GL}(n,R)$ on $\widetilde{\mathscr{GL}}_G(n,R)$ determine an action of $\mathrm{GL}(n, R) \rtimes G$, and the classifying space $|\widetilde{\mathscr{GL}}_G(n, R)|$ is a universal principal $(G, \mathrm{GL}(n, R)_G)$–bundle.*

Definition 8.6 The general linear G–category $\mathscr{GL}_G(R)$ of finite-dimensional free R–modules is the coproduct over $n \geq 0$ of the orbit categories $\widetilde{\mathscr{GL}}_G(n, R)/\,\mathrm{GL}(n, R)$. By Proposition 8.5, $B\mathscr{GL}_G(R)$ is the coproduct over $n \geq 0$ of classifying spaces $B(G, \mathrm{GL}(n, R)_G)$. Explicitly, by [11, Lemma 6.20], the objects of $\mathscr{GL}_G(R)$ are the finite-dimensional free R–submodules M of $R[U]$. The morphisms $\nu\colon M \to N$ are the isomorphisms of R–modules. The group G acts by translation on objects, so that $gM = \{gm \mid m \in M\}$, and by conjugation on morphisms, so that $(g\nu)(gm) = \nu(m)$ for $m \in M$ and $g \in G$.

Proposition 8.7 *The G–categories $\widetilde{\mathscr{GL}}_G(R)$ and $\mathscr{GL}_G(R)$ are $\mathscr{V}_G$–categories and passage to orbits over general linear groups defines a map $\widetilde{\mathscr{GL}}_G(R) \to \mathscr{GL}_G(R)$ of $\mathscr{V}_G$–categories.*

Proof Define a functor

$$\theta_j\colon \mathscr{V}_G(j) \times \mathscr{GL}_G(R)^j \to \mathscr{GL}_G(R)$$

as follows. On objects, for $\phi \in \mathscr{V}_G(j)$ and $M_i \in \mathscr{O}b\,\mathscr{GL}_G(R)$, $1 \leq i \leq j$, define

$$\theta_j(\phi; M_1, \dots, M_j) = R[\phi](M_1 \oplus \cdots \oplus M_j),$$

where $R[\phi]\colon R[{}^jU] \to R[U]$ is induced by $\phi\colon {}^jU \to U$ and M_i is viewed as a submodule of the i^{th} copy of $R[U]$ in $R[{}^jU] = \bigoplus_j R[U]$. For a morphism

$$(\zeta; \nu_1, \dots, \nu_j)\colon (\phi; M_1, \dots, M_j) \to (\psi; N_1, \dots, N_j),$$

define $\theta_j(\iota; \nu_1, \dots, \nu_j)$ to be the unique isomorphism of R–modules that makes the following diagram commute:

$$\begin{tikzcd}
M_1 \oplus \cdots \oplus M_j \arrow[r, "R[\phi]"] \arrow[d, "\nu_1\oplus\cdots\oplus\nu_j"'] & R[\phi](M_1 \oplus \cdots \oplus M_j) \arrow[d, "\theta_j(\zeta;\nu_1,\dots,\nu_j)"] \\
N_1 \oplus \cdots \oplus N_j \arrow[r, "R[\psi]"'] & R[\psi](N_1 \oplus \cdots \oplus N_j)
\end{tikzcd}$$

Then the θ_j specify an action of $\mathscr{V}_G$ on $\mathscr{GL}_G(R)$. Since the $\widetilde{\mathscr{GL}}_G(R,n)$ are chaotic, to define an action of $\mathscr{V}_G$ on $\widetilde{\mathscr{GL}}_G(R)$, we need only specify the required G–functors

$$\widetilde{\theta}_j\colon \mathscr{V}_G(j)\times \widetilde{\mathscr{GL}}_G(R)^j \to \widetilde{\mathscr{GL}}_G(R)$$

on objects. A typical object has the form $(\phi;\iota_1,\dots,\iota_j)$ for $\iota_i\colon R^{n_i}\to R[U]$, and with $n=n_1+\dots+n_j$, we have that $\widetilde{\theta}_j$ sends it to

$$R[\phi]\circ(\iota_1\oplus\cdots\oplus\iota_j)\colon R^n\to R[U].$$

Again, the $\widetilde{\theta}_j$ specify an action. The compatibility with passage to orbits is verified by use of canonical orbit representatives for objects that are obtained by choosing reference maps $\eta_M\colon R^n\to M$ for each M–dimensional free R–module $M\subset R[U]$; compare [11, Proposition 6.18, Lemma 6.20]. □

On passage to classifying spaces and then to G–spectra via our infinite loop space machine $\mathbb{E}_G$, we obtain a model $\mathbb{E}_G B\mathscr{GL}_G(R)$ for the K–theory spectrum $\mathbb{K}_G(R)$ of R. The following result compares the two evident models in sight.

Definition 8.8 Define the naive permutative G–category $\mathrm{GL}_G(R)$ to be the G–groupoid whose objects are the $n\geq 0$ and whose set of morphisms $m\to n$ is empty if $m\neq n$ and is the G–group $\mathrm{GL}(n,R)$ if $m=n$, where G acts entrywise. The product is given by block sum of matrices. Applying the chaotic groupoid functor to the groups $\mathrm{GL}(n,R)$ we obtain another naive permutative G–category $\mathcal{E}\mathrm{GL}_G(R)$ and a map $\mathcal{E}\mathrm{GL}_G(R)\to\mathrm{GL}_G(R)$ of naive permutative G–categories. Applying the functor $\mathscr{C}at(\mathcal{E}G,-)$ from Proposition 4.6, we obtain a map of genuine permutative G–categories $\mathscr{C}at(\mathcal{E}G,(\mathcal{E}\mathrm{GL}_G(R)))\to\mathscr{C}at(\mathcal{E}G,(\mathrm{GL}_G(R)))$.

It is convenient to write $\mathscr{GL}_G^{\mathscr{P}}(R)$ for the $\mathscr{P}_G$–category $\mathscr{C}at(\mathcal{E}G,(\mathrm{GL}_G(R)))$ and $\mathscr{GL}_G^{\mathscr{V}}(R)$ for the $\mathscr{V}_G$–category $\mathscr{GL}_G(R)$, and similarly for their total space variants $\mathscr{C}at(\mathcal{E}G,(\mathcal{E}\mathrm{GL}_G(R)))$ and $\widetilde{\mathscr{GL}}_G(R)$. We have the following comparison theorem.

Theorem 8.9 *The G–spectra $\mathbb{K}_G\mathscr{GL}_G^{\mathscr{P}}(R)$ and $\mathbb{K}_G\mathscr{GL}_G^{\mathscr{V}}(R)$ are weakly equivalent, functorially in G–rings R.*

Proof We again use the product of operads trick from [21]. Projections and quotient maps give the following commutative diagram of $(\mathscr{P}_G\times\mathscr{V}_G)$–categories:

$$\begin{array}{ccccc}
\widetilde{\mathscr{GL}}_G^{\mathscr{P}}(R) & \longleftarrow & \widetilde{\mathscr{GL}}_G^{\mathscr{P}}(R)\times\widetilde{\mathscr{GL}}_G^{\mathscr{V}}(R) & \longrightarrow & \widetilde{\mathscr{GL}}_G^{\mathscr{V}}(R)\\
\downarrow & & \downarrow & & \downarrow\\
\mathscr{GL}_G^{\mathscr{P}}(R) & \longleftarrow & \mathscr{GL}_G^{\mathscr{P}\times\mathscr{V}}(R) & \longrightarrow & \mathscr{GL}_G^{\mathscr{V}}(R)
\end{array}$$

The middle term at the top denotes the diagonal product, namely

$$\coprod_n \widetilde{\mathscr{GL}}_G^{\mathscr{P}}(n,R) \times \widetilde{\mathscr{GL}}_G^{\mathscr{V}}(n,R).$$

The middle term on the bottom is the coproduct over n of the orbits of these products under the diagonal action of $\mathrm{GL}(n,R)$. The product of total spaces of universal principal $(G,\mathrm{GL}(R,n)_G)$–bundles is the total space of another universal principal $(G,\mathrm{GL}(R,n)_G)$–bundle. Therefore, after application of the classifying-space functor, the horizontal projections display two equivalences between universal principal $(G,\mathrm{GL}(R,n)_G)$–bundles. The conclusion follows by hitting the resulting diagram with the functor $\mathbb{K}_G$ defined with respect to $(\mathscr{P}_G \times \mathscr{V}_G)$–categories and using evident equivalences to the functors $\mathbb{K}_G$ defined with respect to $\mathscr{P}_G$–categories and $\mathscr{V}_G$–categories when the input is given by $\mathscr{P}_G$ or $\mathscr{V}_G$–categories. □

8.3 Multiplicative actions on $\mathscr{E}_G^U$ and $\mathscr{GL}_G(R)$

We agree to think of $\mathscr{V}_G^\times$–categories as multiplicative, whereas we think of $\mathscr{V}_G$–categories as additive.

Proposition 8.10 *The G–category $\mathscr{E}_G^U$ is a $\mathscr{V}_G^\times$–category.*

Proof Define a G–functor

$$\xi_j\colon \mathscr{V}_G^\times(j) \times (\mathscr{E}_G^U)^j \to \mathscr{E}_G^U$$

as follows. On objects, for $\phi \in \mathscr{V}_G^\times(j)$ and $A_i \in \mathscr{E}_G$, $1 \le i \le j$, define

$$\xi_j(\phi; A_1,\dots,A_j) = \phi(A_1 \times \cdots \times A_j).$$

For a morphism

$$(\zeta;\nu_1,\dots,\nu_j)\colon (\phi;A_1,\dots,A_j) \to (\psi;B_1,\dots,B_j),$$

define $\xi_j(\zeta;\nu_1,\dots,\nu_j)$ to be the unique bijection that makes the following diagram commute:

$$\begin{tikzcd}
A_1 \times \cdots \times A_j \arrow[r, "\phi"] \arrow[d, "\nu_1\times\cdots\times\nu_j"'] & \phi(A_1 \times \cdots \times A_j) \arrow[d, "\xi_j(\zeta;\nu_1,\dots,\nu_j)"] \\
B_1 \times \cdots \times B_j \arrow[r, "\psi"'] & \psi(B_1 \times \cdots \times B_j)
\end{tikzcd}$$

Then the ξ_j specify an action of $\mathscr{V}_G^\times$ on $\mathscr{E}_G^U$. □

Proposition 8.11 *If R is a commutative G–ring, then $\mathscr{GL}_G(R)$ is a $\mathscr{V}_G^\times$–category.*

Proof Define a functor
$$\xi_j\colon \mathscr{V}_G^{\times}(j)\times\mathscr{GL}(R)_G^j \to \mathscr{GL}_G(R)$$
as follows. Identify $R[U^j]$ with $\bigotimes_j R[U]$, where $\otimes = \otimes_R$. On objects, for $\phi \in \mathscr{V}_G(j)$ and R–modules $M_i \subset R[U]$, $1 \le i \le j$, define
$$\xi_j(\phi; M_1,\dots,M_j) = R[\phi](M_1\times\cdots\times M_j).$$
For a morphism
$$(\zeta;\nu_1,\dots,\nu_j)\colon (\phi;M_1,\dots,M_j)\to(\psi;N_1,\dots,N_j),$$
define $\xi_j(\zeta;\nu_1,\dots,\nu_j)$ to be the unique isomorphism of R–modules that makes the following diagram commute:
$$\begin{tikzcd}
M_1\otimes\cdots\otimes M_j \arrow[r, "R[\phi]"] \arrow[d, "\nu_1\otimes\cdots\otimes\nu_j"'] & \phi(M_1\otimes\cdots\otimes M_j)] \arrow[d, "\xi_j(\zeta;\nu_1,\dots,\nu_j)"] \\
N_1\otimes\cdots\otimes N_j \arrow[r, "R[\psi]"] & \psi(N_1\otimes\cdots\otimes N_j)
\end{tikzcd}$$
Then the ξ_j specify an action of $\mathscr{V}_G^{\times}$ on $\mathscr{GL}_G(R)$. □

Restricting the action from $\mathscr{V}_G^{\times}$ to $\mathscr{W}_G$, the examples above and easy diagram chases prove that the operad pair $(\mathscr{V}_G,\mathscr{W}_G)$ acts on the categories $\mathscr{E}_G$ and $\mathscr{GL}_G(R)$. This proves the following result.

Theorem 8.12 *The categories $\mathscr{E}_G^U$ and $\mathscr{GL}_G(R)$ for a commutative G–ring R are E_∞–ring G–categories in the sense of Definition 7.12.*

Although we have a definition of a genuine permutative G–category, we do not have a comparably simple definition of a genuine bipermutative G–category. The previous examples show that we do have examples of E_∞–ring G–categories. In [13], we will show how to construct E_∞–ring G–categories from general naive bipermutative G–categories, in particular nonequivariant bipermutative categories, and we will show how to construct genuine commutative ring G–spectra from them.

9 The $\mathscr{V}_G$–category $\mathscr{E}_G^U(X)$ and the BPQ theorem

We now return to the categorical BPQ theorem, but thinking in terms of $\mathscr{V}_G$–categories rather than $\mathscr{P}_G$–categories. This gives a more intuitive approach to the G–category of finite sets over a G–space X.

9.1 The G–category $\mathscr{E}_G^U(X)$ of finite sets over X

Definition 9.1 Let X be a G–space. We define the G–groupoid $\mathscr{E}_G^U(X) = \mathscr{E}_G^{\mathscr{V}}(X)$ of finite sets over X. Its objects are the functions $p\colon A \to X$, where A is a finite subset of our ambient G–set U. For a second function $q\colon B \to X$, a map $\nu\colon p \to q$ is a bijection $\nu\colon A \to B$ such that $q \circ \nu = p$. Composition is given by composition of functions over X. The group G acts by translation of G–sets and conjugation on all maps in sight. Thus, for an object $p\colon A \to X$, we have $gp\colon gA \to X$ given by $(gp)(ga) = g(p(a))$. For a map $\nu\colon p \to q$, we have $g\nu\colon gA \to gB$ given by $(g\nu)(ga) = g(\nu(a))$.

To topologize $\mathscr{E}_G^U(X)$, give U and X disjoint basepoints $*$.[7] View the set $\mathscr{O}b$ of objects of $\mathscr{E}_G^U(X)$ as the set of based functions $p\colon U_+ \to X_+$ such that $p^{-1}(*)$ is the complement of a finite set $A \subset U$. Topologize $\mathscr{O}b$ as a subspace of $X_+^{U_+}$. View the set $\mathscr{M}or$ of morphisms of $\mathscr{E}_G^U(X)$ as a subset of the set of functions $\mu\colon U_+ \to U_+$ that send the complement of some finite set $A \subset U$ to $*$ and map A bijectively to some finite set $B \subset U$. Topologize $\mathscr{M}or$ as the subspace of points (p, μ, q) in $\mathscr{O}b \times U_+^{U_+} \times \mathscr{O}b$, where $U_+^{U_+}$ is discrete. When X is a finite set and thus a discrete space (since points are closed in spaces in the category $\mathscr{U}$), $\mathscr{E}_G^U(X)$ is discrete.

Let $\mathscr{E}_G^U(n, X)$ denote the full subcategory of $\mathscr{E}_G^U(X)$ of maps $p\colon A \to X$ such that A has n elements. Then $\mathscr{E}_G^U(X)$ is the coproduct of the groupoids $\mathscr{E}_G^U(n, X)$.

Proposition 9.2 *The operad $\mathscr{V}_G$ acts naturally on the categories $\mathscr{E}_G^U(X)$.*

Proof For $j \geq 0$, we must define functors

$$\theta_j\colon \mathscr{V}_G(j) \times \mathscr{E}_G^U(X)^j \to \mathscr{E}_G^U(X).$$

To define θ_j on objects, let $\phi\colon {}^jU \to U$ be an injective function and $p_i\colon A_i \to X$ be a function, $1 \leq i \leq j$, where A_i is a finite subset of U. We define $\theta_j(\phi; p_1, \ldots, p_j)$ to be the composite

$$\phi(A_1 \sqcup \cdots \sqcup A_j) \xrightarrow{\phi^{-1}} A_1 \sqcup \cdots \sqcup A_j \xrightarrow{\sqcup p_i} {}^jX \xrightarrow{\nabla} X,$$

where ∇ is the fold map, the identity on each of the j copies of X. To define θ on morphisms, let $\psi\colon {}^jU \to U$ be another injective function, and let $\zeta\colon \phi \to \psi$ be the unique map in $\mathscr{V}_G(j)$. For functions $q_i\colon B_i \to X$ and bijections $\nu_i\colon A_i \to B_i$ such that $q_i\nu_i = p_i$, define $\theta_j(\zeta; \nu_1, \ldots, \nu_j)$ to be the unique dotted arrow bijection that

[7] These basepoints are just a convenience for specifying the topology; they play no other role.

makes the following diagram commute:

$$\begin{tikzcd}
\phi(A_1 \sqcup \cdots \sqcup A_j) \arrow[r, "\phi^{-1}"] \arrow[d, dashed, "\theta(\zeta;\nu_1,\dots,\nu_j)"'] & A_1 \sqcup \cdots \sqcup A_j \arrow[d, "\sqcup \nu_i"] \arrow[dr, "\sqcup p_i"] & & \\
& & {}^j X \arrow[r, "\nabla"] & X \\
\psi(B_1 \sqcup \cdots \sqcup B_j) \arrow[r, "\psi^{-1}"'] & B_1 \sqcup \cdots \sqcup B_j \arrow[ur, "\sqcup q_i"'] & &
\end{tikzcd}$$

Then the maps θ_j specify an action of $\mathscr{V}_G$ on the category $\mathscr{E}_G^U(X)$. □

We have a multiplicative elaboration, which is similar to [24, Proposition VI.1.9] but curiously restricted. Regarding a G–space X as a constant G–category with object and morphism space both X, it makes sense to speak of an action of the operad $\mathscr{V}_G^\times$ on the G–category X. For example, $\mathscr{V}_G^\times$ acts on X if X is a commutative topological G–monoid. The following result is closely related to Proposition 7.10. It has the minor advantage that restriction from $\mathscr{V}_G^\times$ to $\mathscr{W}_G$ is unnecessary but the major limitation that it only applies to commutative G–monoids, not to general $\mathscr{V}_G^\times$–algebras.

Proposition 9.3 *If X is a commutative topological G–monoid, then $\mathscr{E}_G^U(X)$ is an E_∞–ring G–category.*

Proof By analogy with the previous proof, for $k \geq 0$, we have functors

$$\xi\colon \mathscr{V}_G^\times(k) \times \mathscr{E}_G^U(X)^k \to \mathscr{E}_G^U(X^k).$$

With notation as in the previous proof, on objects $(\phi; p_1, \dots, p_k)$ for $p_r\colon A_r \to X$, we define $\xi(\phi; p_1, \dots, p_k)$ to be the composite

$$\phi(A_1 \times \cdots \times A_k) \xrightarrow{\phi^{-1}} A_1 \times \cdots \times A_k \xrightarrow{\times p_k} X^k \xrightarrow{\pi} X,$$

where π is the k–fold product on X. On morphisms $(\zeta; \nu_1, \dots, \nu_k)$, where the $\nu_r\colon p_r \to q_r$ are understood to be order preserving, $\xi(\zeta; \nu_1, \dots, \nu_k)$ is defined to be the unique dotted arrow that makes the following diagram commute:

$$\begin{tikzcd}
\phi(A_1 \times \cdots \times A_k) \arrow[r, "\phi^{-1}"] \arrow[d, dashed, "\xi(\zeta;\nu_1,\dots,\nu_k)"'] & A_1 \times \cdots \times A_k \arrow[d, "\times \nu_i"] \arrow[dr, "\times p_i"] & & \\
& & X^k \arrow[r, "\pi"] & X \\
\psi(B_1 \times \cdots \times B_k) \arrow[r, "\psi^{-1}"'] & B_1 \times \cdots \times B_k \arrow[ur, "\times q_i"'] & &
\end{tikzcd}$$

Further details are similar to those in the proof of [24, Proposition VI.1.9] or [30, Proposition 4.9]. □

9.2 Free $\mathscr{V}_G$–categories and the $\mathscr{V}_G$–categories $\mathscr{E}_G^U(X)$

The categories $\mathscr{E}_G^U(X)$ are conceptually simple, and they allow us to give the promised genuinely equivariant variant of Theorem 5.9. To see that, we give a reinterpretation of $\mathscr{E}_G^U(X)$. Regarding X as a topological G–category as before, we have the topological G–category $\widetilde{\mathscr{E}}_G^U(j) \times_{\Sigma_j} X^j$.

Lemma 9.4 *The topological G–categories $\mathscr{E}_G^U(j, X)$ and $\widetilde{\mathscr{E}}_G^U(j) \times_{\Sigma_j} X^j$ are naturally isomorphic.*

Proof For an ordered set $A = (a_1, \dots, a_j)$ of points of U, let a point $(A; x_1, \dots, x_j)$ of $\mathscr{O}b(\widetilde{\mathscr{E}}_G(j) \times_{\Sigma_j} X^j)$ correspond to the function $p\colon A \to X$ given by $p(a_i) = x_i$. Similarly, let a point $(\nu\colon A \to B; x_1, \dots, x_j)$ of $\mathscr{M}or(\widetilde{\mathscr{E}}_G(j) \times_{\Sigma_j} X^j)$ correspond to the bijection $\nu\colon p \to q$ over X, where $q\nu(a_i) = p(a_i) = x_i$. Since we have passed to orbits over Σ_j, our specifications are independent of the ordering of A. These correspondences identify the two categories. □

Recall that we write $\mathbb{V}_G$ for the monad on based G–categories associated to the operad $\mathscr{V}_G$, we write $|\mathscr{V}_G|$ for the operad of G–spaces obtained by applying the classifying-space functor B to $\mathscr{V}_G$, and we write $\boldsymbol{V}_G$ for the monad on based G–spaces associated to $|\mathscr{V}_G|$. Recall too that X_+ denotes the union of the G–category X with a disjoint trivial basepoint category $*$ and that

$$\mathbb{V}_G(X_+) = \coprod_{j \geq 0} \mathscr{V}_G(j) \times_{\Sigma_j} X^j. \tag{9.5}$$

Theorem 9.6 *There is a natural map*

$$\omega\colon \mathbb{V}_G(X_+) \to \mathscr{E}_G^U(X)$$

of $\mathscr{V}_G$–categories, and it induces a weak equivalence

$$B\omega\colon \boldsymbol{V}_G(X_+) \to B\mathscr{E}_G^U(X)$$

of $|\mathscr{V}_G|$–spaces on passage to classifying spaces.

Proof Pick any G–fixed point $1 \in U$.[8] Define an inclusion $i\colon X_+ \to \mathscr{E}_G^U(X)$ of based G–categories by identifying $*$ with $\mathscr{E}_G^U(0, X)$ and mapping X to $\mathscr{E}_G^U(1, X)$ by sending x to the map $1 \to x$ from the 1–point subset 1 of U to X. Since $\mathbb{V}_G(X_+)$ is

[8] This must not be confused with the convenience basepoint $*$ used to define the topology.

the free (based) $\mathscr{V}_G$–category generated by X_+, the inclusion i induces the required natural map ω. Explicitly, it is the composite

$$\mathbb{V}_G(X_+) \xrightarrow{\mathbb{V}_G i} \mathbb{V}_G \mathscr{E}_G^U(X)) \xrightarrow{\theta} \mathscr{E}_G^U(X).$$

More explicitly still, let $\underline{1} \subset {}^jU$ be the j–point subset consisting of the elements 1 in the j summands. Then ω is the coproduct of the maps

$$\omega_j = i_j \times_{\Sigma_j} \mathrm{id}\colon \mathscr{V}_G(j) \times_{\Sigma_j} X^j \to \widetilde{\mathscr{E}}_G^U(j) \times_{\Sigma_j} X^j,$$

where $i_j\colon \mathscr{V}_G(j) \to \widetilde{\mathscr{E}}_G^U(j)$ is the $(G\times\Sigma_j)$–functor that sends an object $\phi\colon {}^jU \to U$ to the set $\phi(\underline{1}) \subset U$ and sends the morphism $\nu\colon \phi \to \psi$ to the bijection

$$\phi(\underline{1}) \xrightarrow{\phi^{-1}} \underline{1} \xrightarrow{\psi} \psi(\underline{1}).$$

Passing to classifying spaces, $|i_j|$ is a map between universal principal (G, Σ_j)–bundles, both of which are $(G\times\Sigma_j)$–CW complexes. Therefore, $|i_j|$ is a $(G\times\Sigma_j)$–equivariant homotopy equivalence. The conclusion follows. □

9.3 The categorical BPQ theorem: second version

We begin by comparing Theorem 9.6, which is about G–categories, with Theorems 5.5, 5.9 and 5.10, which are about G–fixed categories. Clearly $\mathscr{E}_G(X)^G$ is a $\mathscr{V}$–category, where $\mathscr{V} = (\mathscr{V}_G)^G$. By Theorem 9.6, it is weakly equivalent (in the homotopical sense) to the $\mathscr{V}$–category $(\mathbb{V}_G X_+)^G$. We also have the $\mathscr{P}$–category $\mathscr{F}_G(X)^G$, which by Theorem 5.9 and Remark 5.11 is equivalent (in the categorical sense) to the $\mathscr{P}$–category $(\mathbb{P}_G X_+)^G$. Elaborating Remark 8.3, $\mathscr{E}_G^U(X)^G$ and $\mathscr{F}_G(X)^G$ are two small models for the category of all finite G–sets and G–isomorphisms over X and are therefore equivalent. To take the operad actions into account, recall the discussion in Section 4.3. We say that a map of topological G–categories is a weak equivalence if its induced map of classifying G–spaces is a weak equivalence.

Lemma 9.7 *The $\mathscr{P}_G$–category $\mathbb{P}_G X_+$ and the $\mathscr{V}_G$–category $\mathbb{V}_G X_+$ are weakly equivalent as $(\mathscr{P}_G\times\mathscr{V}_G)$–categories. Therefore, the $\mathscr{P}$–category $(\mathbb{P}_G X_+)^G$ and the $\mathscr{V}$–category $(\mathbb{V}_G X_+)^G$–categories are weakly equivalent.*

Proof The projections

$$\mathbb{P}_G X_+ \leftarrow (\mathbb{P}_G \times \mathbb{V}_G)(X_+) \to \mathbb{V}_G X_+$$

are maps of $(\mathscr{P}_G\times\mathscr{V}_G)$–categories that induce weak equivalences of $|\mathscr{P}_G \times \mathscr{V}_G|$–spaces on passage to classifying spaces. □

Theorem 9.8 *The classifying spaces of the $\mathscr{P}$–category $\mathscr{F}_G(X)^G$ and the $\mathscr{V}$–category $\mathscr{E}_G^U(X)^G$ are weakly equivalent as $|\mathscr{P} \times \mathscr{V}|$–spaces.*

The conclusion is that, on the G–fixed level, the categories $\mathscr{E}_G^U(X)^G$ and $\mathscr{F}_G(X)^G$ can be used interchangeably as operadically structured versions of the category of finite G–sets over X. On the equivariant level, $\mathscr{E}_G^U(X)$ but not $\mathscr{F}_G(X)$ is operadically structured. It is considerably more convenient than the categories $\mathbb{P}_G(X_+)$ or $\mathbb{V}_G(X_+)$. With the notation $\mathbb{K}_G\mathbb{V}_G(X_+)=\mathbb{E}_G B\mathbb{V}_G(X_+)=\mathbb{E}_G \boldsymbol{V}_G(X_+)$ and $\mathbb{K}_G\mathscr{E}_G^U(X)=\mathbb{E}_G B\mathscr{E}_G^U(X)$, we have the following immediate consequence of Theorems 6.2 and 9.6. It is our preferred version of the categorical BPQ theorem, since it uses the most intuitive categorical input.

Theorem 9.9 (categorical Barratt–Priddy–Quillen theorem) *For G–spaces X, there is a composite natural weak equivalence*

$$\alpha\colon \Sigma_G^\infty X_+ \to \mathbb{K}_G\mathbb{V}_G X_+ \to \mathbb{K}_G\mathscr{E}_G^U(X).$$

Remark 9.10 It is not known how the tom Dieck splitting theorem behaves with respect to the Mackey functor structure on homotopy groups. It seems likely to us that this could be analyzed using this version of the BPQ theorem and our categorical proof of the splitting.

Appendix A: Pairings of operads

We recall the following definition from [25, 1.4]. It applies equally well equivariantly. We write it elementwise, but written diagrammatically it applies to operads in any symmetric monoidal category $\mathscr{V}$. Write $\boldsymbol{j}=\{1,\dots,j\}$ and let

$$\otimes\colon \Sigma_j\times\Sigma_k\to\Sigma_{jk}$$

be the homomorphism obtained by identifying $\boldsymbol{j}\times\boldsymbol{k}$ with $\boldsymbol{jk}$ by ordering the set of jk elements (q,r), $1\le q\le j$ and $1\le r\le k$, lexicographically. More precisely, let $\lambda_{j,k}\colon \boldsymbol{jk}\to\boldsymbol{j}\times\boldsymbol{k}$ be the lexicographic ordering. Then, given $\rho\in\Sigma_j$ and $\sigma\in\Sigma_k$, $\rho\otimes\sigma$ is defined by

$$\boldsymbol{jk}\xrightarrow{\lambda_{j,k}}\boldsymbol{j}\times\boldsymbol{k}\xrightarrow{\rho\times\sigma}\boldsymbol{j}\times\boldsymbol{k}\xrightarrow{\lambda_{j,k}^{-1}}\boldsymbol{jk}.$$

For nonnegative integers h_q and i_r, let

$$\delta\colon \coprod_{(q,r)}(\boldsymbol{h}_q\times\boldsymbol{i}_r)\to\Big(\coprod_q\boldsymbol{h}_q\Big)\times\Big(\coprod_r\boldsymbol{i}_r\Big)$$

be the distributivity isomorphism viewed as a permutation via block and lexicographic identifications of the source and target sets with the appropriate set $\boldsymbol{n}$. A little more precisely, we define the permutation δ to be the composite

$$\sum_{q,r} \boldsymbol{h}_q \boldsymbol{i}_r \xrightarrow{\cong} \boldsymbol{h}_1 \boldsymbol{i}_1 \sqcup \boldsymbol{h}_1 \boldsymbol{i}_2 \sqcup \cdots \sqcup \boldsymbol{h}_j \boldsymbol{i}_k \xrightarrow{\lambda \sqcup \cdots \sqcup \lambda} \boldsymbol{h}_1 \times \boldsymbol{i}_1 \sqcup \cdots \sqcup \boldsymbol{h}_j \times \boldsymbol{i}_k$$
$$\xrightarrow{\text{dist}} (\boldsymbol{h}_1 \sqcup \cdots \sqcup \boldsymbol{h}_j) \times (\boldsymbol{i}_1 \sqcup \cdots \sqcup \boldsymbol{i}_k) \xrightarrow{\cong} \boldsymbol{h} \times \boldsymbol{i} \xrightarrow{\lambda_{h,i}^{-1}} \boldsymbol{h}\boldsymbol{i}.$$

Definition A.1 Let $\mathscr{C}$, $\mathscr{D}$, and $\mathscr{E}$ be operads in a symmetric monoidal category $\mathscr{V}$ (with product denoted by $\otimes$). A pairing of operads

$$\boxtimes\colon (\mathscr{C}, \mathscr{D}) \to \mathscr{E}$$

consists of maps

$$\boxtimes\colon \mathscr{C}(j) \otimes \mathscr{D}(k) \to \mathscr{E}(jk)$$

in $\mathscr{V}$ for $j \geq 0$ and $k \geq 0$ such that the diagrammatic versions of the following properties hold, where $c \in \mathscr{C}(j)$ and $d \in \mathscr{D}(k)$:

(i) If $\mu \in \Sigma_j$ and $\nu \in \Sigma_k$, then

$$c\mu \boxtimes d\nu = (c \boxtimes d)(\mu \otimes \nu).$$

(ii) With $j = k = 1$, we have $\mathrm{id} \boxtimes \mathrm{id} = \mathrm{id}$.

(iii) If $c_q \in \mathscr{C}(h_q)$ for $1 \leq q \leq j$ and $d_r \in \mathscr{D}(i_r)$ for $1 \leq r \leq k$, then[9]

$$\gamma(c \boxtimes d; \times_{(q,r)} c_q \boxtimes d_r) = [\gamma(c; \times_q c_q) \boxtimes \gamma(d; \times_r d_r)]\delta.$$

When specialized to spaces, the following definition (which is a variant of [25, 1.2]) gives one possible starting point for multiplicative infinite loop space theory.

Definition A.2 Let $\boxtimes\colon (\mathscr{C}, \mathscr{D}) \to \mathscr{E}$ be a pairing of operads in $\mathscr{V}$. A pairing of a $\mathscr{C}$–algebra X and a $\mathscr{D}$–algebra Y to an $\mathscr{E}$–algebra Z is a map $f\colon X \otimes Y \to Z$ such that the following diagram commutes for all j and k, where X^j denotes the j^{th} tensor power in $\mathscr{V}$ and we write θ generically for action maps:

$$\begin{tikzcd}
\mathscr{C}(j) \otimes X^j \otimes \mathscr{D}(k) \otimes Y^k \arrow[rr, "\theta \otimes \theta"] \arrow[d, "\boxtimes"'] & & X \otimes Y \arrow[d, "f"] \\
\mathscr{E}(jk) \otimes (X \otimes Y)^{jk} \arrow[r, "\mathrm{id} \otimes f^{jk}"'] & \mathscr{E}(jk) \otimes Z^{jk} \arrow[r, "\theta"'] & Z
\end{tikzcd}$$

On the left, $\boxtimes$ denotes the composite

$$\mathscr{C}(j) \otimes X^j \otimes \mathscr{D}(k) \otimes Y^k \xrightarrow{\mathrm{id} \otimes t \otimes \mathrm{id}} \mathscr{C}(j) \otimes \mathscr{D}(k) \otimes X^j \otimes Y^k \xrightarrow{\boxtimes \otimes \lambda} \mathscr{E}(jk) \otimes Z^{jk}.$$

[9] The original definition in [25] had δ on the other side in this condition.

Here, in elementwise notation,

$$\lambda((x_1 \otimes \cdots \otimes x_j) \otimes (y_1 \otimes \cdots \otimes y_k)) = ((x_1 \otimes y_1) \otimes \cdots \otimes (x_j \otimes y_k)),$$

where we order the pairs $(x_q \otimes y_r)$, $1 \le q \le j$ and $1 \le r \le k$, lexicographically.

Letting $\mathscr{V}$ be the category of unbased G–spaces, with $\otimes = \times$, but then passing to monads on based G–spaces, we obtain the following observations.

Proposition A.3 *For based G–spaces X and Y, a pairing $\boxtimes\colon (\mathscr{C}_G, \mathscr{D}_G) \to \mathscr{E}_G$ of operads of G–spaces induces a natural pairing*

$$\boxtimes\colon \boldsymbol{C}_G X \wedge \boldsymbol{D}_G Y \to \boldsymbol{E}_G(X \wedge Y)$$

such that the following diagrams commute:

$$\begin{tikzcd}
X \wedge Y \arrow[r, "\eta\wedge\eta"] \arrow[dr, "\eta"'] & \boldsymbol{C}_G X \wedge \boldsymbol{D}_G Y \arrow[d, "\boxtimes"] \\
& \boldsymbol{E}_G(X \wedge Y)
\end{tikzcd}$$

$$\begin{tikzcd}
\boldsymbol{C}_G\boldsymbol{C}_G X \wedge \boldsymbol{D}_G\boldsymbol{D}_G Y \arrow[rr, "\mu\wedge\mu"] \arrow[d, "\boxtimes"'] & & \boldsymbol{C}_G X \wedge \boldsymbol{D}_G Y \arrow[d, "\boxtimes"] \\
\boldsymbol{E}_G(\boldsymbol{C}_G X \wedge \boldsymbol{D}_G Y) \arrow[r, "\boldsymbol{E}_G\boxtimes"'] & \boldsymbol{E}_G\boldsymbol{E}_G(X \wedge Y) \arrow[r, "\mu"'] & \boldsymbol{E}_G(X \wedge Y)
\end{tikzcd}$$

The following diagram commutes for any pairing $f\colon X \otimes Y \to Z$ of a $\mathscr{C}_G$–algebra X and a $\mathscr{D}_G$–algebra Y to an $\mathscr{E}_G$–algebra Z:

$$\begin{tikzcd}
\boldsymbol{C}_G X \wedge \boldsymbol{D}_G Y \arrow[rr, "\theta\wedge\theta"] \arrow[d, "\boxtimes"'] & & X \wedge Y \arrow[d, "f"] \\
\boldsymbol{E}_G(X \wedge Y) \arrow[r, "\boldsymbol{E}_G f"'] & \boldsymbol{E}_G Z \arrow[r, "\theta"'] & Z
\end{tikzcd}$$

Proof The map $\boxtimes$ is induced from the map $\boxtimes$ of the previous definition and the commutativity of the first two diagrams is checked by chases from Definition A.1. The commutativity of the second implies that $\boxtimes$ is a pairing in the sense of Definition A.2. The commutativity of the third follows from Definition A.2. $\square$

Example A.4 The following commutative diagram, in which we ignore the path space variable for simplicity, shows that condition (iii) is satisfied by the pairing $(\mathscr{K}_V, \mathscr{K}_W) \to \mathscr{K}_{V\oplus W}$ defined in Proposition 1.17; this completes the proof of that result:

$$
\begin{tikzcd}
\boldsymbol{h}\boldsymbol{i}\times V\times W \arrow[r] \arrow[d, "\delta\times\mathrm{id}"'] & \Big(\coprod_{q,r}\boldsymbol{h}_q\boldsymbol{i}_r\Big)\times V\times W \arrow[r, "\mathrm{dist}"] \arrow[d, "\sqcup\lambda\times\mathrm{id}"'] & \coprod_{q,r}(\boldsymbol{h}_q\boldsymbol{i}_r\times V\times W) \arrow[r, "\sqcup c_q\otimes d_r"] \arrow[d, "\sqcup\lambda"'] & \boldsymbol{j}\times\boldsymbol{k}\times V\times W \arrow[d, "\mathrm{twist}"] \\
\boldsymbol{h}\boldsymbol{i}\times V\times W \arrow[d, "\lambda\times\mathrm{id}"'] & \Big(\coprod_{q,r}\boldsymbol{h}_q\times\boldsymbol{i}_r\Big)\times V\times W \arrow[r, "\mathrm{dist}"] \arrow[d, "\bar{\delta}\times\mathrm{id}"'] & \coprod_{q,r}(\boldsymbol{h}_q\times\boldsymbol{i}_r\times V\times W) \arrow[d, "\sqcup\,\mathrm{twist}"'] & \boldsymbol{j}\times V\times\boldsymbol{k}\times W \arrow[d, "c\times d"] \\
\boldsymbol{h}\times\boldsymbol{i}\times V\times W \arrow[r] \arrow[d, "\mathrm{twist}"'] & \Big(\coprod_q\boldsymbol{h}_q\Big)\times\Big(\coprod_r\boldsymbol{i}_r\Big)\times V\times W \arrow[d, "\mathrm{twist}"'] & \coprod_{q,r}(\boldsymbol{h}_q\times V\times\boldsymbol{i}_r\times W) \arrow[d, "\bar{\delta}"'] \arrow[uur, "\sqcup(c_q\times d_r)"] & V\times W \\
\boldsymbol{h}\times V\times\boldsymbol{i}\times W \arrow[r] & \Big(\coprod_q\boldsymbol{h}_q\Big)\times V\times\Big(\coprod_r\boldsymbol{i}_r\Big)\times W \arrow[r, "\mathrm{dist}"] & \Big(\coprod_q\boldsymbol{h}_q\times V\Big)\times\Big(\coprod_r\boldsymbol{i}_r\times W\Big) \arrow[r, "\sqcup c_q \times\sqcup d_r"] & \Big(\coprod_q V\Big)\times\Big(\coprod_r W\Big) \arrow[u, "c\times d"'] \arrow[uu, bend left, "\mathrm{id}"]
\end{tikzcd}
$$

The following counterexample was pointed out to us by Anna Marie Bohmann and Angelica Osorno. Using a more sophisticated categorical framework, we shall explain how to get around the difficulty in [12; 13].

Counterexample A.5 We show that the pairing (6.4) is not a self-pairing of $\mathscr{P}$. Letting $\tau\in\mathscr{P}(2)$ be the transposition $\tau=(12)$, we calculate

$$\gamma(\tau\otimes\tau;\ \mathrm{id}_2\otimes\mathrm{id}_1,\ \mathrm{id}_2\otimes\mathrm{id}_1,\ \mathrm{id}_1\otimes\mathrm{id}_1,\ \mathrm{id}_1\otimes\mathrm{id}_1) = (1526)(3)(4)$$

whereas

$$[\gamma(\tau;\mathrm{id}_2,\mathrm{id}_1)\otimes\gamma(\tau;\mathrm{id}_1,\mathrm{id}_1)]\delta = (14526)(3).$$

In this case δ is the transposition (23). Thus condition (iii) fails.

Appendix B: The double bar construction and the proof of Theorem 2.25

The proof of Theorem 2.25 is based on a construction that the senior author has used for decades in unpublished work and whose algebraic analogue has also long been used. Heretofore he has always found alternative arguments that avoid its use in published work, and the topological version seems not to have appeared in print. The construction works in great generality with different kinds of bar constructions, as described in [34; 35; 45], for example. We restrict attention to the monadic bar construction used in this paper. We shall be informal, since it is routine to fill in the missing details.

We assume given two monads $\boldsymbol{C}$ and $\boldsymbol{D}$ in some reasonable category $\mathscr{U}$, and we assume given a morphism of monads $\iota\colon \boldsymbol{C} \to \boldsymbol{D}$. We also assume given a right $\boldsymbol{D}$–functor $\Sigma\colon \mathscr{U} \to \mathscr{V}$ for some other reasonable category $\mathscr{V}$. Then Σ is a right $\boldsymbol{D}$–functor with the pullback action

$$\Sigma\boldsymbol{C} \to \Sigma\boldsymbol{D} \to \Sigma.$$

Let X be a $\boldsymbol{C}$–algebra in $\mathscr{U}$. Reasonable means in particular that we can form "geometric realizations" of simplicial objects X as usual, tensoring X over the category Δ with a canonical (covariant) simplex functor from Δ to $\mathscr{U}$ or $\mathscr{V}$.

We assume that the functor $\boldsymbol{D}$ commutes with geometric realization, so that the realization of a simplicial $\boldsymbol{D}$–algebra is a $\boldsymbol{D}$–algebra. Then the bar construction

$$\iota_! X = B(\boldsymbol{D}, \boldsymbol{C}, X)$$

in $\mathscr{U}$ specifies an "extension of scalars" functor that converts $\boldsymbol{C}$–algebras X to $\boldsymbol{D}$–algebras in a homotopically well-behaved fashion. Since D acts on Σ, we have the bar construction $B(\Sigma, \boldsymbol{D}, \iota_! X)$, and we also have the bar construction $B(\Sigma, \boldsymbol{C}, X)$, both with values in $\mathscr{V}$. Under these assumptions, we have the following result.

Theorem B.1 *There is a natural equivalence* $B(\Sigma, \boldsymbol{D}, \iota_! X) \simeq B(\Sigma, \boldsymbol{C}, X)$.

Proof of Theorem 2.25 We replace $\mathscr{U}$ by $G\mathscr{U}$ and $\mathscr{V}$ by $G\mathscr{S}p$. We take $\boldsymbol{C}$ to be the monad associated to the operad $\mathscr{C}_{U^G} = (\mathscr{C}_G)^G \times \mathscr{K}_{U^G}$ and $\boldsymbol{D}$ to be the monad associated to $\mathscr{C}_U = \mathscr{C}_G \times \mathscr{K}_U$. We take Σ to be Σ^∞_G, and we recall that $\Sigma^\infty_G = i_* \Sigma^\infty$ by Lemma 2.22. By inspection or a commutation of left adjoints argument, the functor i_* commutes with geometric realization. Therefore,

$$\boldsymbol{E}_G(\iota_! X) \equiv B(\Sigma^\infty_G, \boldsymbol{C}_U, \iota_! X) \simeq B(\Sigma^\infty_G, \boldsymbol{C}_{U^G}, X) \cong i_* B(\Sigma^\infty, \boldsymbol{C}_{U^G}, X) \equiv i_* \boldsymbol{E} X,$$

where Theorem B.1 gives the equivalence. $\square$

Proof of Theorem B.1 We construct the double bar construction

$$B(\Sigma, \boldsymbol{D}, \boldsymbol{D}, \boldsymbol{C}, X)$$

as the geometric realization of the bisimplicial object $B_{\bullet,\bullet}(\Sigma, \boldsymbol{D}, \boldsymbol{D}, \boldsymbol{C}, X)$ in $\mathscr{V}$ whose (p,q)–simplex object is $\Sigma \boldsymbol{D}^p \boldsymbol{D} \boldsymbol{C}^q X$. The horizontal face and degeneracy operations are those obtained by applying the simplicial bar construction $B_\bullet(\Sigma, \boldsymbol{D}, Y)$ to the $\boldsymbol{D}$–algebras $Y = \boldsymbol{D}\boldsymbol{C}^q X$. The vertical face and degeneracy operations are those obtained by applying the simplicial bar construction $B_\bullet(\Upsilon, \boldsymbol{C}, X)$ to the $\boldsymbol{C}$–functors $\Upsilon = \Sigma \boldsymbol{D}^p \boldsymbol{D}$. The geometric realization of a bisimplicial object is obtained equivalently as the realization of the diagonal simplicial object, the horizontal realization of its vertical realization, and the vertical realization of its horizontal realization. Realizing first vertically and then horizontally, we obtain

$$B(\Sigma, \boldsymbol{D}, B(\boldsymbol{D}, \boldsymbol{C}, X)) = B(\Sigma, \boldsymbol{D}, i_! X).$$

Realizing first horizontally and then vertically, we obtain the bar construction

$$B(B(\Sigma, \boldsymbol{D}, \boldsymbol{D}), \boldsymbol{C}, X) \simeq B(\Sigma, \boldsymbol{C}, X).$$

Here $B(\Sigma, \boldsymbol{D}, \boldsymbol{D})$ is the right $\boldsymbol{C}$–functor whose value on a $\boldsymbol{C}$–algebra Y is $B(\Sigma, \boldsymbol{D}, \boldsymbol{D}Y)$ with right $\boldsymbol{C}$–action induced by the $\boldsymbol{C}$–action $\boldsymbol{C}Y \to Y$. The equivalence is induced by the standard natural equivalence $B(\Sigma, \boldsymbol{D}, \boldsymbol{D}Y) \to \Sigma Y$. □

Remark B.2 The double bar construction can be defined more generally and symmetrically. Dropping the assumption that there is a map of monads $\boldsymbol{C} \to \boldsymbol{D}$, we have that $B(\Sigma, \boldsymbol{D}, \boldsymbol{F}, \boldsymbol{C}, X)$ is defined if F is a left $\boldsymbol{D}$–functor and a right $\boldsymbol{C}$–functor $\mathbb{U} \to \mathbb{U}$ such that the following diagram commutes:

$$\begin{tikzcd}
\boldsymbol{DFC} \arrow[r] \arrow[d] & \boldsymbol{DF} \arrow[d] \\
\boldsymbol{FC} \arrow[r] & \boldsymbol{F}
\end{tikzcd}$$

This can even work when the domain and target categories of $\boldsymbol{F}$ differ but agree with the categories on which $\boldsymbol{C}$ and $\boldsymbol{D}$ are defined.

References

[1] **M Ando**, **A J Blumberg**, **D Gepner**, **M J Hopkins**, **C Rezk**, *Units of ring spectra, orientations and Thom spectra via rigid infinite loop space theory*, J. Topol. 7 (2014) 1077–1117 MR

[2] **M G Barratt**, **P J Eccles**, *Γ^+–structures, I: A free group functor for stable homotopy theory*, Topology 13 (1974) 25–45 MR

[3] **J Caruso**, **S Waner**, *An approximation theorem for equivariant loop spaces in the compact Lie case*, Pacific J. Math. 117 (1985) 27–49 MR

[4] **S R Costenoble**, **S Waner**, *Fixed set systems of equivariant infinite loop spaces*, Trans. Amer. Math. Soc. 326 (1991) 485–505 MR

[5] **T tom Dieck**, *Orbittypen und äquivariante Homologie, II*, Arch. Math. (Basel) 26 (1975) 650–662 MR

[6] **A W M Dress**, **A O Kuku**, *A convenient setting for equivariant higher algebraic K–theory*, from "Algebraic K–theory, Part I" (R K Dennis, editor), Lecture Notes in Math. 966, Springer (1982) 59–68 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] **A D Elmendorf**, **M A Mandell**, *Rings, modules, and algebras in infinite loop space theory*, Adv. Math. 205 (2006) 163–228 MR

[9] **Z Fiedorowicz**, **H Hauschild**, **J P May**, *Equivariant algebraic K–theory*, from "Algebraic K–theory, Part II" (R K Dennis, editor), Lecture Notes in Math. 967, Springer (1982) 23–80 MR

[10] **B Guillou**, **J P May**, *Models of G–spectra as presheaves of spectra*, preprint (2011) arXiv

[11] **B Guillou**, **J P May**, **M Merling**, *Categorical models for equivariant classifying spaces*, Algebr. Geom. Topol. 17 (2017) 2565–2602 MR

[12] **B Guillou**, **J P May**, **M Merling**, **A Osorno**, *Equivariant infinite loop space theory, II: The multiplicative Segal machine* In preparation

[13] **B Guillou**, **J P May**, **M Merling**, **A Osorno**, *Equivariant infinite loop space theory, III: The categorical story* In preparation

[14] **H Hauschild**, *Klassifizierende Räume, G–Konfiburations-räume und äquivariante Schleifenräume*, Habilitationsschrift, Göttingen (1978)

[15] **H Hauschild**, *Äquivariante Konfigurationsräume und Abbildungsräume*, from "Topology Symposium" (U Koschorke, W D Neumann, editors), Lecture Notes in Math. 788, Springer (1980) 281–315 MR

[16] **M A Hill**, **M J Hopkins**, *Equivariant symmetric monoidal structures*, preprint (2016) arXiv

[17] **A Kuku**, *Representation theory and higher algebraic K–theory*, Pure and Applied Mathematics (Boca Raton) 287, Chapman & Hall, Boca Raton, FL (2007) MR

[18] **L G Lewis, Jr**, **J P May**, **M Steinberger**, **J E McClure**, *Equivariant stable homotopy theory*, Lecture Notes in Math. 1213, Springer (1986) MR

[19] **M A Mandell**, **J P May**, *Equivariant orthogonal spectra and S–modules*, Mem. Amer. Math. Soc. 755, Amer. Math. Soc., Providence, RI (2002) MR

[20] **M A Mandell**, **J P May**, **S Schwede**, **B Shipley**, *Model categories of diagram spectra*, Proc. London Math. Soc. 82 (2001) 441–512 MR

[21] **J P May**, *The geometry of iterated loop spaces*, Lectures Notes in Math. 271, Springer (1972) MR

[22] **J P May**, *E_∞ spaces, group completions, and permutative categories*, from "New developments in topology" (G Segal, editor), London Math. Soc. Lecture Note Ser. 11, Cambridge Univ. Press (1974) 61–93 MR

[23] **J P May**, *Classifying spaces and fibrations*, Mem. Amer. Math. Soc. 155, Amer. Math. Soc., Providence, RI (1975) MR

[24] **J P May**, *E_∞ ring spaces and E_∞ ring spectra*, Lecture Notes in Math. 577, Springer (1977) MR

[25] **J P May**, *Pairings of categories and spectra*, J. Pure Appl. Algebra 19 (1980) 299–346 MR

[26] **J P May**, *Multiplicative infinite loop space theory*, J. Pure Appl. Algebra 26 (1982) 1–69 MR

[27] **J P May**, *Definitions: operads, algebras and modules*, from "Operads: Proceedings of Renaissance Conferences" (J-L Loday, J D Stasheff, A A Voronov, editors), Contemp. Math. 202, Amer. Math. Soc., Providence, RI (1997) 1–7 MR

[28] **J P May**, *Operads, algebras and modules*, from "Operads: Proceedings of Renaissance Conferences" (J-L Loday, J D Stasheff, A A Voronov, editors), Contemp. Math. 202, Amer. Math. Soc., Providence, RI (1997) 15–31 MR

[29] **J P May**, *The construction of E_∞ ring spaces from bipermutative categories*, from "New topological contexts for Galois theory and algebraic geometry" (A Baker, B Richter, editors), Geom. Topol. Monogr. 16, Geom. Topol. Publ., Coventry (2009) 283–330 MR

[30] **J P May**, *What precisely are E_∞ ring spaces and E_∞ ring spectra?*, from "New topological contexts for Galois theory and algebraic geometry" (A Baker, B Richter, editors), Geom. Topol. Monogr. 16, Geom. Topol. Publ., Coventry (2009) 215–282 MR

[31] **J P May**, **M Merling**, **A Osorno**, *Equivariant infinite loop space theory, I: The space level story*, preprint (2017) arXiv

[32] **J P May**, **R Thomason**, *The uniqueness of infinite loop space machines*, Topology 17 (1978) 205–224 MR

[33] **M Merling**, *Equivariant algebraic K-theory of G–rings*, Math. Z. 285 (2017) 1205–1248 MR

[34] **J-P Meyer**, *Bar and cobar constructions, I*, J. Pure Appl. Algebra 33 (1984) 163–207 MR

[35] **J-P Meyer**, *Bar and cobar constructions, II*, J. Pure Appl. Algebra 43 (1986) 179–210 MR

[36] **G Nishida**, *The transfer homomorphism in equivariant generalized cohomology theories*, J. Math. Kyoto Univ. 18 (1978) 435–451 MR

[37] **C Rourke**, **B Sanderson**, *Equivariant configuration spaces*, J. London Math. Soc. 62 (2000) 544–552 MR

[38] **C Rourke**, **B Sanderson**, *The compression theorem, I*, Geom. Topol. 5 (2001) 399–429 MR

[39] **C Rourke**, **B Sanderson**, *The compression theorem, II: Directed embeddings*, Geom. Topol. 5 (2001) 431–440 MR

[40] **C Rourke**, **B Sanderson**, *The compression theorem, III: Applications*, Algebr. Geom. Topol. 3 (2003) 857–872 MR

[41] **P Salvatore**, **N Wahl**, *Framed discs operads and Batalin–Vilkovisky algebras*, Q. J. Math. 54 (2003) 213–231 MR

[42] **G Segal**, *Configuration-spaces and iterated loop-spaces*, Invent. Math. 21 (1973) 213–221 MR

[43] **K Shimakawa**, *Infinite loop G–spaces associated to monoidal G–graded categories*, Publ. Res. Inst. Math. Sci. 25 (1989) 239–262 MR

[44] **K Shimakawa**, *Mackey structures on equivariant algebraic K–theory*, K–Theory 5 (1991/92) 355–371 MR

[45] **M Shulman**, *Homotopy limits and colimits and enriched homotopy theory*, preprint (2006) arXiv

[46] **R Steiner**, *A canonical operad pair*, Math. Proc. Cambridge Philos. Soc. 86 (1979) 443–449 MR

[47] **R W Thomason**, *Uniqueness of delooping machines*, Duke Math. J. 46 (1979) 217–252 MR

Department of Mathematics, The University of Kentucky
Lexington, KY, United States

Department of Mathematics, The University of Chicago
Chicago, IL, United States

bertguillou@uky.edu, may@math.uchicago.edu

Received: 14 July 2012 Revised: 3 March 2017

Algebraic & Geometric Topology 17 (2017) 3341–3373

The localized skein algebra is Frobenius

NEL ABDIEL
CHARLES FROHMAN

When A in the Kauffman bracket skein relation is set equal to a primitive n^{th} root of unity ζ with n not divisible by 4, the Kauffman bracket skein algebra $K_\zeta(F)$ of a finite-type surface F is a ring extension of the $\mathrm{SL}_2\mathbb{C}$–character ring of the fundamental group of F. We localize by inverting the nonzero characters to get an algebra $S^{-1}K_\zeta(F)$ over the function field of the corresponding character variety. We prove that if F is noncompact, the algebra $S^{-1}K_\zeta(F)$ is a symmetric Frobenius algebra. Along the way we prove $K(F)$ is finitely generated, $K_\zeta(F)$ is a finite-rank module over the coordinate ring of the corresponding character variety, and learn to compute the trace that makes the algebra Frobenius.

57M27

1 Introduction

This paper is a step in a program to build a 4–dimensional extended field theory that assigns invariants to manifolds equipped with a homomorphism of their fundamental group into $\mathrm{SL}_2\mathbb{C}$. A *symmetric Frobenius algebra* A over a field k is a k–algebra equipped with a k–linear map $\mathrm{Tr}\colon A \to k$ that is cyclic in the sense that for all $\alpha, \beta \in A$, $\mathrm{Tr}(\alpha\beta) = \mathrm{Tr}(\beta\alpha)$, and for all nonzero $\alpha \in A$, there exists $\beta \in A$ with $\mathrm{Tr}(\alpha\beta) \neq 0$. Frobenius algebras are central to the construction of field theories.

We show that the Kauffman bracket skein algebra of a compact surface with nonempty boundary can be localized to give a symmetric Frobenius algebra over the function field a character variety of the fundamental group of the surface. The trace that makes the localized skein algebra Frobenius is a potent tool for explicating the algebraic structure of $K_\zeta(F)$, as seen in Frohman and Kania-Bartoszynska [10].

A surface F is of finite type if there is a closed oriented surface $\widehat{F}$ and a finite set of points $\{p_i\} \in \widehat{F}$ such that $F = \widehat{F} - \{p_i\}$. In this paper all surfaces are either compact oriented (possibly with boundary) or of finite type. If F is a compact, connected, oriented surface, a punctured disk can be glued into each boundary component to obtain a finite-type surface. There is a one-to-one correspondence between disjoint families of simple closed curves in the two surfaces, so the theorems we prove working with finite-type surfaces apply to surfaces having finitely many boundary components.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3341

A central result of this paper is:

Theorem 3.7 *The Kauffman bracket skein algebra $K(F)$ of a finite-type surface with coefficients in $\mathbb{Z}[A, A^{-1}]$ is finitely generated as an algebra by a finite family of simple closed curves S_i. In fact,*

$$\{S_{\sigma(1)}^{k_1} * S_{\sigma(2)}^{k_2} * \cdots * S_{\sigma(n)}^{k_n}\}, \tag{1-1}$$

where $k_i \in \mathbb{Z}_{\geq 0}$, spans $K(F)$ for any permutation $\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\}$.

This is an extension of a theorem of Bullock [3]. Key to the proof is a well ordering of the simple diagrams on the surface. Given an ideal triangulation with edges E of the surface F, there is an embedding of the isotopy classes of simple diagrams on F into $\mathbb{Z}_{\geq 0}^E$. Letting $\mathcal{S}$ denote the simple diagrams on F, there is an injective map,

$$\iota\colon \mathcal{S} \to \mathbb{Z}_{\geq 0}^E, \tag{1-2}$$

which assigns the tuple $\prod_c i(S,c)$ to the simple diagram S, where $i(S,c)$ is the geometric intersection number of S with the edge c. Choosing an order on E gives rise to the lexicographic ordering of $\mathbb{Z}_{\geq 0}^E$ which in turn induces a well ordering of $\mathcal{S}$. The *geometric sum* of two simple diagrams S and S' is the simple diagram $S + S'$ such that $\iota(S+S') = \iota(S) + \iota(S')$.

Any skein has a unique expression as a linear combination of simple diagrams with nonzero coefficients. The *lead term* of a skein is the term in that expression involving the largest simple diagram. Suppose that we have defined the Kauffman bracket skein algebra $K_{\mathfrak{D}}(F)$ over an integral domain $\mathfrak{D}$, so that the variable ζ in the Kauffman bracket skein relation is a unit in $\mathfrak{D}$. The central tool for proving the theorems in this paper is:

Theorem 3.4 *Let F be a finite-type surface with at least one puncture and negative Euler characteristic. Choose an ideal triangulation with edges E, and order E in order to define the lead term of a skein. The lead term of the product of two simple diagrams S and S' in $K_{\mathfrak{D}}(F)$ is ζ raised to a power times the geometric sum of S and S'.*

If A is set equal to -1 in $K(F)$, the corresponding skein algebra is canonically isomorphic to the coordinate ring of the $\mathrm{SL}_2\mathbb{C}$–character variety of the fundamental group of F; see Bullock [4] and Przytycki and Sikora [15]. If A is set equal to 1, then the corresponding skein algebra is still isomorphic to the $\mathrm{SL}_2\mathbb{C}$–character variety of $\pi_1(F)$, just not canonically; see Barrett [1].

Let ζ be an n^{th} root of unity, and let $m = n/\gcd(n,4)$. For reasons that will become clear, we call m the *index of threading*. Let $\epsilon = \zeta^{m^2}$. Throughout this paper we assume

$n \neq 0 \bmod 4$. In the case that n is odd, $\epsilon = 1$. If $n = 2 \bmod 4$ then $\epsilon = -1$. There is a theorem of Bonahon and Wong that, when $n \neq 0 \bmod 4$, there is a natural embedding

$$\text{Ch}\colon K_\epsilon(F) \to K_\zeta(F). \tag{1-3}$$

We denote the image of Ch with its coefficients extended to $\mathbb{Z}\left[\frac{1}{2}, \zeta\right]$ by $\chi(F)$. This is to remind us that it is canonically isomorphic to the coordinate ring of a character variety. We use the finite generation of $K(F)$ to prove that $K_\zeta(F)$ is a finitely generated module over $K_\epsilon(F)$. Localizing at $S = \chi(F) - \{0\}$, the algebra $S^{-1}K_\zeta(F)$ is finite-dimensional over $S^{-1}\chi(F)$.

Theorem 3.9 *Suppose that ζ is an n^{th} root of unity with $n \neq 0 \bmod 4$, and let m be the index of threading. Let F be a finite-type surface. If S_i is any system of simple diagrams corresponding to an integral basis of the cone of admissible colorings, then the skeins $\prod_i T_{k_i}(S_i)$, where the $k_i \in \{0, 1, \ldots, m-1\}$ span $K_\zeta(F)$ over $\chi(F)$. In particular, $K_\zeta(F)$ is a finite ring extension of $\chi(F)$.*

Frohman and Kania-Bartoszynska [10] prove that $S^{-1}K_\zeta(F)$ is a vector space of dimension $m^{-3e(F)}$ over $S^{-1}\chi(F)$, where $e(F)$ is the Euler characteristic of F and m is the index of threading. Next we prove:

Theorem 3.10 *If F is closed and ζ is a primitive n^{th} root of unity with $n \neq 0 \bmod 4$, then $K_\zeta(F)$ is a finite-rank module over $\chi(F)$.*

This means that all irreducible representations of $K_\zeta(F)$ over the complex numbers are of bounded dimension.

Each $\alpha \in S^{-1}K_\zeta(F)$ induces an $S^{-1}\chi(F)$–linear endomorphism

$$L_\alpha\colon S^{-1}K_\zeta(F) \to S^{-1}K_\zeta(F) \tag{1-4}$$

by left multiplication. The *normalized trace* $\text{Tr}(\alpha)$ of α is the trace of L_α as a linear endomorphism divided by the dimension of the vector space $S^{-1}K_\zeta(F)$ over the field $S^{-1}\chi(F)$. The normalized trace has the properties:

- $\text{Tr}(1) = 1$.
- $\text{Tr}(\alpha * \beta) = \text{Tr}(\beta * \alpha)$.
- Tr is $S^{-1}\chi(F)$–linear.

Hence, if Tr is nondegenerate then $S^{-1}K_\zeta(F)$ equipped with the normalized trace is a symmetric Frobenius algebra over the function field of the character variety of the fundamental group of the surface F.

Along the way we learn to compute Tr: $S^{-1}K_\zeta(F) \to S^{-1}\chi(F)$ with respect to a special basis. A primitive diagram on F is a system of disjoint simple closed curves S_i such that no S_i bounds a disk and no two curves in the system cobound an annulus. The skein $\prod_i T_{k_i}(S_i)$ is the product over all i of the result of threading S_i with the k_i^{th} Chebyshev polynomial of the first kind. These span $S^{-1}K_\zeta(F)$ over $S^{-1}\chi(F)$.

Theorem 4.13 *Suppose that $s = \sum_i \beta_i P_i$, where the β_i are in $S^{-1}\chi(F)$ and the P_i are primitive diagrams whose components have been threaded with Chebyshev polynomials of the first kind. Let J be those indices i such that the components of P_i have only been threaded with Chebyshev polynomials whose index is divisible by the index of threading. Then*

$$\mathrm{Tr}(s) = \sum_{i \in J} \beta_i P_i. \tag{1-5}$$

The derivation of the formula for the trace depends on the following surprising fact. Let $\bigcup_i S_i$ be a simple diagram, made up of the simple closed curves S_i. The extension $S^{-1}\chi(F)[S_1, \dots, S_n]$ of $S^{-1}\chi(F)$ obtained by adjoining the S_i is a field. This extends a result of Muller [14], which says that simple closed curves are not zero divisors.

Since the value of the formula for the trace doesn't have any denominators that didn't appear in the input, the trace is actually defined as a $\chi(F)$–linear map

$$\mathrm{Tr}\colon K_\zeta(F) \to \chi(F). \tag{1-6}$$

Next, the formula for the trace is used to prove that there are no nontrivial principal ideals in the kernel of Tr: $K_\zeta(F) \to \chi(F)$, completing the proof that $S^{-1}K_\zeta(F)$ is a symmetric Frobenius algebra. Essential to the proof is the fact that, given a primitive diagram $\bigcup S_i$, the skeins $\prod_i T_{k_i}(S_i)$ with $0 \le k_i \le m-1$ generate a field extension of $S^{-1}\chi(F)$ in $S^{-1}K_\zeta(F)$.

Acknowledgements The authors would like to thank Pat Gilmer, Thang Lê and Joanna Kania-Bartoszynska for helpful input.

2 Preliminaries

2.1 Kauffman bracket skein module

Let M be an orientable 3–manifold. A *framed link* in M is an embedding of a disjoint union of annuli into M. Throughout this paper $M = F \times [0,1]$ for an orientable surface F. Diagrammatically we depict framed links by showing the core of the annuli lying parallel to F. Two framed links in M are equivalent if they are isotopic. Let $\mathcal{L}$ denote the set of equivalence classes of framed links in M, including the empty link. By $\mathbb{Z}[A, A^{-1}]$ we mean Laurent polynomials with integral coefficients in the formal variable A. Consider the free module over $\mathbb{Z}[A, A^{-1}]$,

$$\mathbb{Z}[A, A^{-1}]\mathcal{L}, \tag{2-1}$$

with basis $\mathcal{L}$. Let S be the submodule spanned by the Kauffman bracket skein relations,

(2-2)  $-A$  $-A^{-1}$ 

and

$$\bigcirc \cup L + (A^2 + A^{-2})L.$$

The framed links in each expression are identical outside the balls pictured in the diagrams, and when both arcs in a diagram lie in the same component of the link, the same side of the annulus is up. The Kauffman bracket skein module $K(M)$ is the quotient

$$\mathbb{Z}[A, A^{-1}]\mathcal{L}/S(M). \tag{2-3}$$

A *skein* is an element of $K(M)$. Let F be a compact orientable surface and let $I = [0, 1]$. There is an algebra structure on $K(F \times I)$ that comes from laying one framed link over the other. The resulting algebra is denoted by $K(F)$ to emphasize that it comes from the particular structure as a cylinder over F. Denote the stacking product with a $*$, so $\alpha * \beta$ means α stacked over β. If it is known the two skeins commute, the $*$ will be omitted.

A *simple diagram* D on the surface F is a system of disjoint simple closed curves such that none of the curves bounds a disk. A simple diagram D is *primitive* if no two curves in the diagram cobound an annulus. A simple diagram can be made into a framed link by choosing a system of disjoint annuli in F so that each annulus has a single curve in the diagram as its core. This is sometimes called the *blackboard framing*. The set of isotopy classes of blackboard framed simple diagrams form a basis for $K(F)$ [5; 12; 16].

2.2 Specializing A

If R is a commutative ring and $\zeta \in R$ is a unit, then R is a $\mathbb{Z}[A, A^{-1}]$–module, where the action

$$\mathbb{Z}[A, A^{-1}] \otimes R \to R \tag{2-4}$$

is given by letting $p \in \mathbb{Z}[A, A^{-1}]$ act by multiplication by the result of evaluating p at ζ. The skein module specialized at $\zeta \in R$ is

$$K_R(M) = K(M) \otimes_{\mathbb{Z}[A,A^{-1}]} R. \tag{2-5}$$

You can think of the specialization as setting A equal to ζ in the Kauffman bracket skein relations.

This is much too general a setting to get nice structure theorems for $K_R(M)$, so we restrict our attention to when the ring R is an integral domain. To emphasize that we are working with an integral domain we denote the ring by $\mathfrak{D}$. Since $\mathbb{Z}[A, A^{-1}]$ is an integral domain and A is a unit, our results hold for $K(A)$ as a special case. For that reason the theorems in this paper are all stated in terms of $K_{\mathfrak{D}}(M)$, the skein module specialized at a unit ζ in an integral domain $\mathfrak{D}$.

We are most interested in the case when ζ is a primitive n^{th} root of unity, where $n \neq 0 \bmod 4$. The integral domain is $\mathbb{Z}\left[\frac{1}{2}, \zeta\right]$. We need 2 to be a unit so that a collection of skeins that are adapted to the computation of the trace will be a basis.

2.3 Threading

The Chebyshev polynomials of the first type T_k are defined recursively by

- $T_0(x) = 2$,
- $T_1(x) = x$, and
- $T_{n+1}(x) = T_1(x) \cdot T_n(x) - T_{n-1}(x)$.

They satisfy some nice properties.

Proposition 2.1 *For $m, n > 0$, $T_m(T_n(x)) = T_{mn}(x)$. Furthermore, for all $m, n \geq 0$, $T_m(x)T_n(x) = T_{m+n}(x) + T_{|m-n|}(x)$.*

For a proof see [8].

We denote the oriented surface of genus g with p punctures by $\Sigma_{g,p}$. It is easy to see that $K_{\mathfrak{D}}(\Sigma_{0,2})$ is isomorphic to $\mathfrak{D}[x]$, where x is the framed link coming from the blackboard framing of the core of the annulus. Hence $1, x, x^2, \dots, x^n, \dots$ is a basis for $K_{\mathfrak{D}}(\Sigma_{0,2})$. Since $T_0(x) = 2$, in order to use the $T_k(x)$ as a basis for $\mathfrak{D}$, 2 must be a unit in $\mathfrak{D}$. If $2 \in \mathfrak{D}$ is a unit then $\{T_k(x) \mid k \in \mathbb{Z}_{\geq 0}\}$ is a basis for $K_{\mathfrak{D}}(\Sigma_{0,2})$.

If the components of the primitive diagram on a finite-type surface F are the simple closed curves S_i and $k_i \in \mathbb{Z}_{\geq 0}$ has been chosen for each component, the result of threading each of the curves S_i with the T_{k_i} is $\prod_i T_{k_i}(S_i)$. Since the S_i are disjoint from one another, they commute, so order doesn't matter in the product. For any compact or finite-type surface F, the primitive diagrams on F up to isotopy, with their components threaded with all possible choices of Chebyshev polynomials, form a basis for $K_{\mathfrak{D}}(F)$ so long as $2 \in \mathfrak{D}$ is a unit. This basis is becoming more commonly used in the study of skein algebras [9; 17; 13].

The following theorem of Bonahon and Wong is the starting point for this investigation. The convention for defining $K_\zeta(M)$ means that it is a module over $\mathbb{Z}\left[\frac{1}{2}, \zeta\right]$. This means that $K_\epsilon(M)$ is a module over $\mathbb{Z}\left[\frac{1}{2}\right]$. For the sake of the following theorem, after choosing an m^{th} root of unity ζ we interpret $K_\epsilon(M)$ to have its coefficients extended to include $\mathbb{Z}\left[\frac{1}{2}, \zeta\right]$; that way we don't have to mess around with extending coefficients in the range. More formally, let

$$\overline{K}_\epsilon(M) = K_\epsilon(M) \otimes_{\mathbb{Z}\left[\frac{1}{2}\right]} \mathbb{Z}\left[\tfrac{1}{2}, \zeta\right] \tag{2-6}$$

Theorem 2.2 (Bonahon and Wong [2; 13]) *If M is a compact oriented three-manifold and we specialize at an n^{th} root of unity ζ such that $n \neq 0 \bmod 4$, there is a $\mathbb{Z}\left[\frac{1}{2}, \zeta\right]$–linear map*

$$\mathrm{Ch}\colon \overline{K}_\epsilon(M) \to K_\zeta(M)$$

given by threading framed links with T_m, where $m = n/\gcd(n,4)$ is the index of threading. Any framed link in the image of Ch *is central in the sense that if $L' \cup K$ differs from $L \cup K$ by a crossing change of L and L' with K, then $T_m(L) \cup K = T_m(L') \cup K$. In the case that $M = F \times [0,1]$, the map*

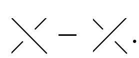
$$\mathrm{Ch}\colon \overline{K}_\epsilon(F) \to K_\zeta(F)$$

is an injective homomorphism of algebras such that the image of Ch *lies in the center of $K_\zeta(F)$.*

The skein module $\overline{K}_\epsilon(M)$ is a ring under disjoint union. At $A = \pm 1$, the Kauffman bracket skein relation

(2-7) 

can be rotated 90 degrees and then subtracted from itself to yield

(2-8) 

This means that, in $K_{\pm 1}(M)$, changing crossings does not change the skein. To take the product of two equivalence classes of framed links, choose representatives that are disjoint from one another and take their union. The product is independent of the representatives chosen, since the results differ by isotopy and changing crossings. The product can be extended distributively to give a product on $K_{\pm 1}(M)$. Let $\sqrt{0}$ denote the nilradical of $K_{-1}(M)$. It is a theorem of Bullock [3], proved independently in [15], that, for any oriented compact 3–manifold, $K_{-1}(M)/\sqrt{0}$ is canonically isomorphic to the coordinate ring of the $\mathrm{SL}_2\mathbb{C}$–character variety of the fundamental group of M. In the case that $M = F \times [0,1]$, the disjoint union product coincides with the stacking product, as stacking is one way to perturb the components of the two links so that they are disjoint. For any oriented finite-type surface F, the ring $\chi(F)$ has basis the isotopy classes of primitive diagrams threaded with $T_{k_i m}$ for all choices $k_i \in \mathbb{Z}_{\geq 0}$.

In Sections 2.4–2.6, some algebraic background is presented that will be applied to $K_\zeta(F)$ in Section 4.

2.4 Specializing at a place

A place of $\chi(F)$ is a homomorphism $\phi\colon \chi(F) \to \mathbb{C}$. The places correspond to evaluation at a point on the character variety. A place defines a $\chi(F)$–module structure on $\mathbb{C}$, where the action

$$\chi(F) \otimes \mathbb{C} \to \mathbb{C} \tag{2-9}$$

is defined by letting $s \in \chi(F)$ act as multiplication by $\phi(s)$ on $\mathbb{C}$. We define the *specialization* of $K_\zeta(F)$ at ϕ to be

$$K_\zeta(F)_\phi = K_\zeta(F) \otimes_{\chi(F)} \mathbb{C}. \tag{2-10}$$

The specialization at a place is an algebra over the complex numbers.

2.5 Localization

Let $R \subset J$ be a ring extension, where R is an integral domain, J is an associative ring with unit and R is a subring of the center of J. Since R has no zero divisors, $S = R - \{0\}$ is multiplicatively closed. Start with the set of ordered pairs $J \times S$, and place an equivalence relation on $J \times S$ by saying (a,s) is equivalent to (b,t) if $at = bs$. Denote the equivalence class of (a,s) under this relation by $[a,s]$. The set of equivalence classes is denoted by $S^{-1}J$, and called the localization of J with respect to S. Denote the set of equivalence classes $[a,s]$ for $a \in R$ by $S^{-1}R$. Define multiplication of equivalence classes by $[a,s][b,t] = [ab, st]$ and addition by $[a,s] + [b,t] = [at + bs, st]$. Under these operations, $S^{-1}R$ is a field, and $S^{-1}J$ is an algebra over that field.

In this paper, J is a subalgebra of $K_\zeta(F)$ and R is $\chi(F)$. This means that $S^{-1}R$ is the function field of the character variety associated to $K_\epsilon(F)$.

2.6 Trace and extension of scalars

Let f be a field and suppose that V is a finite-dimensional vector space over f. If $L \in \mathrm{End}_f(V)$, we use $\mathrm{tr}(L)$ to denote the unnormalized trace of L. The linear map L can be represented with respect to a basis $\{v_j\}$ by a matrix (l_i^j). The trace of L is given by

$$\mathrm{tr}(L) = \sum_i l_i^i. \tag{2-11}$$

If W is also a finite-dimensional vector space over f and $M\colon W \to W$ is an f–linear map, then

$$\mathrm{tr}(L \otimes_f M) = \mathrm{tr}(L)\,\mathrm{tr}(M) \quad \text{and} \quad \mathrm{tr}(L \oplus M) = \mathrm{tr}(L) + \mathrm{tr}(M). \tag{2-12}$$

Suppose that $f \le a$ is a field extension and V is a vector space of dimension n over f; then

$$V \otimes_f a \tag{2-13}$$

is a vector space of dimension n over a. In fact, if $\{v_j\}$ is a basis for V then $\{v_j \otimes 1\}$ is a basis for $V \otimes_f a$ over a.

Under extension of scalars, $L\colon V \to V$ gets sent to $L \otimes_f 1_a$. The matrix of $L \otimes_f 1_a$ with respect to the basis $\{v_j \otimes 1\}$ is the same as the matrix of L with respect to $\{v_j\}$, so

$$\mathrm{tr}(L \otimes_f 1_a) = \mathrm{tr}(L), \tag{2-14}$$

where the trace on the left is taken as an a–linear map, and the trace on the right is taken as an f–linear map, and we are using $f \le a$ to make the identification.

The next proposition gives the method by which we will be computing the trace.

Proposition 2.3 *Suppose that $K \le P \le J$, where K and P are fields, J is an algebra over K and J is a finite-dimensional vector space over k. Thus J is a finite-dimensional vector space over P, and P is a finite field extension of K. If $s \in P$, then it defines a K–linear maps $l_s\colon P \to P$, and $L_s\colon J \to J$ by left multiplication. If d is the dimension of J over P, then*

$$\mathrm{tr}(L_s) = d\,\mathrm{tr}(l_s), \tag{2-15}$$

where the traces are both taken as linear maps over K.

Proof Since $K \le P$ is finite-dimensional it has basis $p_1, \dots p_n$ over K. Since J is a finite-dimensional vector space of dimension d over P it has basis $j_1, \dots j_d$ over P. This implies that $p_a j_c$ is a basis of J over K. Expressing l_s with respect to the basis p_a, we get

$$l_s(p_a) = \sum_b l_b^a p_b. \tag{2-16}$$

Since s acts as scalar multiplication on J,

$$L_s(p_a j_c) = \sum_b l_b^a p_b j_c. \tag{2-17}$$

Hence the matrix for L_s decomposes into d blocks that are all copies of the matrix for l_s. Therefore the trace of L_s is equal to d times the trace of l_s. □

2.7 Geometric intersection numbers

Suppose that X and Z are properly embedded 1–manifolds in the finite-type surface F, where X is compact. We say that X' is a transverse representative of X if X' is ambiently isotopic to X via a compactly supported isotopy and $X' \pitchfork Z$. Define the *geometric intersection number* of X and Z, denoted by $i(X,Z)$, to be the minimum cardinality of $X' \cap Z$ over all transverse representatives of X. We could have instead worked with Z up to compactly supported ambient isotopy and taken the minimum over all Z' isotopic to Z and transverse to X and gotten the same number, so $i(X,Z) = i(Z,X)$.

It is a theorem that a transverse representative of X realizes the geometric intersection number $i(X,Z)$ if and only if there are no *bigons*. A bigon is a disk D embedded in F so that the boundary of D consists of the union of two arcs $a \subset X$ and $b \subset Z$ [7]. If there is a bigon, there is always an innermost bigon, whose interior is disjoint from $X \cup Z$.

3 $K_{\mathfrak{D}}(F)$ is finitely generated

3.1 Parametrizing the simple diagrams

An ideal triangle is a triangle with its vertices removed. An ideal triangulation of a finite-type surface F consists of finitely many ideal triangles Δ_i with their edges identified pairwise, along with a homeomorphism from the resulting quotient space to F. Alternatively an ideal triangulation is defined by a family E of properly embedded lines that cuts F into finitely many ideal triangles. The surface F needs to have at least

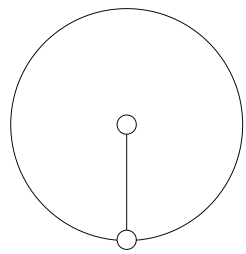


Figure 1: A folded triangle

one puncture, and negative Euler characteristic or it doesn't admit an ideal triangulation. If the Euler characteristic of the surface F is $-e(F)$ then any ideal triangulation of F consists of $2e(F)$ ideal triangles. The cardinality of a set of lines E defining an ideal triangulation is $3e(F)$.

If Δ is an ideal triangle in an ideal triangulation then $\partial\Delta = \{a, b, c\}$, where a, b and c are homeomorphic to $\mathbb{R}$. The lines a, b, and c are the *sides* of Δ. There is a map of Δ to the closure of a component D of the complement of E into F. If this map is an embedding, then Δ is an *embedded ideal triangle*. It could be that two sides c_1 and c_2 of the ideal triangle Δ get mapped to the same line c; in this case Δ is a *folded ideal triangle*. Figure 1 is a picture of a folded ideal triangle. There are two punctures in the picture, and the mapping is $2-1$ along the vertical line joining them. The edge that is covered twice by the mapping has *multiplicity* 2.

Let E denote a disjoint family of properly embedded lines that defines an ideal triangulation of F, and suppose the triangles are the set $\{\Delta_j\}$. An *admissible coloring* $f\colon E \to \mathbb{Z}_{\geq 0}$ is an assignment of a nonnegative integer $f(c)$ to each $c \in E$ such that the following conditions hold:

- If $\{a, b, c\}$ form the boundary of an embedded ideal triangle Δ_j then the sum $f(a) + f(b) + f(c)$ is even and the triple $\{f(a), f(b), f(c)\}$ satisfies the triangle inequality

$$f(a) \leq f(b) + f(c), \quad f(b) \leq f(a) + f(c) \quad \text{and} \quad f(c) \leq f(a) + f(b). \tag{3-1}$$

- If $\{a, b\}$ are the image of the boundary of a folded ideal triangle Δ_j, where b has multiplicity 2, we require that $f(a) + 2f(b)$ be even and $f(a) \leq 2f(b)$.

If $S \subset F$ is a simple diagram then $f_S\colon E \to \mathbb{Z}_{\geq 0}$ given by $f_S(c) = i(S, c)$ is an admissible coloring. Conversely, for each admissible coloring $f\colon E \to \mathbb{Z}_{\geq 0}$ there is an isotopy class of simple diagrams having geometric intersection numbers with the edges given by f. We denote a representative of this isotopy class by $[f]$. In particular, $[f_s] = S$. We use $\mathcal{A}$ to denote the set of all admissible colorings $f\colon E \to \mathbb{Z}_{\geq 0}$.

Proposition 3.1 *The admissible colorings of the edges of an ideal triangulation of F are in one-to-one correspondence with isotopy classes of simple diagrams on F.* □

A *pointed integral polyhedral cone* is a subset $\mathcal{A}$ of some $\mathbb{Z}^k$ that is defined by finitely many equations and inequalities with $\vec{0} \in \mathcal{A}$.

Proposition 3.2 *The admissible colorings of an ideal triangulation of F form a pointed integral polyhedral cone.*

Proof If E is the set of edges of the ideal triangulation then there is a map

$$\mathcal{A} = \{f\colon E \to \mathbb{Z}_{\geq 0} \mid f \text{ is admissible}\} \to \mathbb{Z}^E \tag{3-2}$$

that sends each f to its tuple of values. We still denote the image of this map by $\mathcal{A}$.

The only part of recognizing $\mathcal{A}$ as an integral cone that is tricky is the condition that the sum of colors over the sides of a triangle needs to be even. This can be avoided by using a linearly equivalent description of the admissible colorings via *corner numbers*. An ideal triangle has three corners, determined by a choice of two of the three sides. For instance, if a triangle has three sides a, b and c, then the corners correspond to $\{a,b\}$, $\{b,c\}$ and $\{a,c\}$. If $f\colon E \to \mathbb{Z}_{\geq 0}$ is an admissible coloring, the three corner numbers of this triangle are

$$\tfrac{1}{2}(f(a)+f(b)-f(c)), \quad \tfrac{1}{2}(f(b)+f(c)-f(a)), \quad \tfrac{1}{2}(f(a)+f(c)-f(b)). \tag{3-3}$$

It is easy to see that the corner numbers determine the admissible coloring and vice versa. An assignment of corner numbers corresponds to an admissible coloring of the edges if and only if all the corner numbers are nonnegative and if, for each edge, the sum of the two corner numbers on one side of the edge is equal to the sum of the corner numbers on the other side of that edge. The description in terms of corner numbers allows us to conclude that the admissible colorings are a pointed integral cone. □

An *integral basis* of a pointed integral polyhedral cone is a subset of the cone that has minimal cardinality among all subsets that span the cone additively. It is a classical result [11] that any pointed integral polyhedral cone admits a finite integral basis. The integral basis is unique. If P is a pointed integral polyhedral cone, $p \in P$ is *indivisible* if $s = 0$ or $p = 0$ whenever $s, t \in P$ and $s + t = p$. The set of indivisible elements of P is the integral basis [18]. In the case of the cone of admissible colorings, the diagrams corresponding to indivisible colorings are simple closed curves.

Forgetting positivity, and the triangle inequality, the admissible colorings generate a free module over $\mathbb{Z}$. It makes sense to ask whether a collection $f_{S_i}\colon E \to \mathbb{Z}_{\geq 0}$ are linearly independent. Oddly, the integral basis need not be linearly independent.

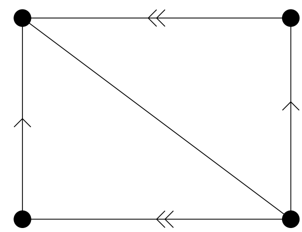


Figure 2: An ideal triangulation of $\Sigma_{1,1}$

Remark 3.3 Decompose the punctured torus $\Sigma_{1,1}$ into two ideal triangles. This requires three edges, which form the boundary of both triangles. In the diagram below we identify the left- and right-hand sides of the rectangle, and the top and bottom of the rectangle with the vertices deleted to obtain a once-punctured torus. The lines defining the triangulation come from the sides of the rectangle and the diagonal, as shown in Figure 2.

The admissible colorings can be seen as triples of counting numbers (m, n, p) whose sum is even that satisfy the triangle inequality. The nonzero indecomposable admissible colorings are $(1,1,0)$, $(1,0,1)$ and $(0,1,1)$. This set is an integral basis. Notice that if (a,b,c) is an admissible coloring and one of the triangle inequalities is strict, say $a < b+c$, we can subtract the corresponding indecomposable $(0,1,1)$ to get a triple $(a, b-1, c-1)$ that still satisfies the triangle inequality and the sum of the colors $a+b+c-2 < a+b+c$. If all three triangle inequalities are equalities $a = b+c$, $b = a+c$ and $c = a+b$, then $(a,b,c) = (0,0,0)$. The three curves corresponding to $(1,1,0)$, $(1,0,1)$ and $(0,1,1)$ are the generators that Bullock and Przytycki [6] obtained for $K(\Sigma_{1,1})$. There are infinitely many ideal triangulations of $\Sigma_{1,1}$ but Euler characteristic forces them all to be two triangles that share all their edges. The argument above goes through, even though the curves on the torus will be different. Since the integral basis is unique, any set of skeins that generates $K_{\mathfrak{D}}(\Sigma_{1,1})$ must have at least three elements.

Choose an ordering of E. Use this to order $\mathbb{Z}_{\geq 0}^E$ lexicographically. Notice that $\mathbb{Z}_{\geq 0}^E$ in the lexicographic ordering is a well-ordered monoid. By that we mean $\mathbb{Z}_{\geq 0}^E$ is well-ordered and, if $a, b \in \mathbb{Z}_{\geq 0}$ have $a < b$, then $a+c < b+c$ for any $c \in \mathbb{Z}_{\geq 0}$. Since $\mathcal{A}$ is a submonoid of $\mathbb{Z}_{\geq 0}^E$, we have that $\mathcal{A}$ is a well-ordered monoid.

If $\alpha \in K_\zeta(F)$ then we can write α as a finite linear combination of simple diagrams with complex coefficients,

$$\alpha = \sum_S z_S S, \tag{3-4}$$

where the S are simple diagrams and the z_S are nonzero elements of $\mathfrak{D}$. The *lead term* of α is $z_S S$, where S is the largest diagram appearing in the sum. We denote the lead term of the skein α as $\mathrm{ld}(\alpha)$.

3.2 The algebra $K_{\mathfrak{D}}(F)$ is finitely generated over $\mathfrak{D}$

If f_S and $f_{S'}$ are admissible colorings, choose simple diagrams S and S' that realize the colorings as the cardinality of their intersections with the $c_i \in E$ and such that S and S' realize their geometric intersection number and $S \cap S'$ is disjoint from all c_i. Up to isotopy there is a unique simple diagram whose associated coloring is $f_S + f_{S'}$, called the *geometric sum* of S and S'. Since addition of admissible colorings is associative, so is the geometric sum. It is worth noting, the geometric sum of two diagrams depends on the choice of ideal triangulation.

Suppose that S and S' transversely represent $i(S, S')$. Furthermore assume that $S \cap S' \cap E = \varnothing$. If there are n points of intersection in $S \cap S'$, there are 2^n ways of smoothing all the crossings of S and S' to get a system of simple closed curves. We call a system of simple closed curves obtained by smoothing all crossings s a *state*. A state might not be a simple diagram as it may contain some trivial simple closed curves. There is a process for writing the product $S * S'$ as a linear combination of simple diagrams. First expand the product as a sum of states using the Kauffman bracket skein relation for crossings, then delete the trivial components of each state, and for each trivial component deleted from a state multiply the coefficient of the state by $-\zeta^2 - \zeta^{-2}$. Order the crossings of $S * S'$. Based on the ordering there is a rooted tree, where the root is the diagram $S * S'$, the vertices are partial smoothings (resolvents) of the diagram, and the directed edges correspond to smoothing the crossings in order. The states are the leaves of this tree. If the shortest path from the root to a state s passes through a resolvent r, we say that s is a *descendent* of r.

Theorem 3.4 *Let S and S' be simple diagrams associated to admissible colorings $f_S, f_{S'}\colon E \to \mathbb{Z}_{\geq 0}$. Assume the product $S * S' \in K_{\mathfrak{D}}(F)$ has been written as $\sum_D z_D D$, where the D are simple diagrams that are distinct up to isotopy and the $z_D \in \mathfrak{D}$ are nonzero. There exists a unique simple diagram E in this sum, so that $f_E = f_S + f'_S$, and all the other simple diagrams appearing with nonzero coefficient in the sum are strictly smaller in the well ordering of diagrams. Furthermore, the coefficient z_E is a power of ζ.*

Lemma 3.5 *Let G be a four-valent graph with at least one vertex, embedded in a disk D^2. Assume that G is the union of two families of properly embedded arcs $A_1 \cup A_2$ and that there are three special points p, q and r in ∂D^2 such that*

- *the endpoints of the* A_1 *and* A_2 *are disjoint from one another and* $\{p,q,r\}$ *in* ∂D^2,
- *if* $a_1 \in A_1$ *and* $a_2 \in A_2$, *then* a_1 *and* a_2 *intersect transversely, and realize their geometric intersection number relative to their boundaries,*
- *if* $a, b \in A_i$ *then* $a \cap b = \varnothing$, *and*
- *for any arc* $a \in A_1 \cup A_2$, *the endpoints of* a *are separated by* $\{p,q,r\}$,

If $A_1 \cap A_2$ *is nonempty, then there is an embedded triangle* Δ *whose sides consist of an arc of* ∂D^2 *that is disjoint from* $\{p,q,r\}$, *an arc contained in some* $a \in A_1$ *that only intersects* A_2 *in a single point which is one of its endpoints, and an arc in some* $b \in A_2$ *that only intersects* A_1 *in a single point which is one of its endpoints.*

(We call this an *outermost triangle*.)

Proof The graph dissects the disk into vertices, edges and faces. The alternating sum of the numbers of vertices, edges and faces is 1, as that is the Euler characteristic of the disk. A face f has two kinds of sides, sides in ∂D^2 and sides in the interior of D^2. Let $e_\partial(f)$ denote the number of sides of f lying in ∂D^2 and $e_i(f)$ the number of sides of f in the interior. Similarly, let $v_\partial(f)$ be the number of vertices of the face that lie in ∂D^2, and $v_i(f)$ be the number of vertices of f that lie in the interior of D^2. The contribution of the face f to the Euler characteristic of the disk is

$$c(f) = 1 - \tfrac{1}{2}e_i(f) - e_\partial(f) + \tfrac{1}{4}v_i(f) + \tfrac{1}{2}v_{\partial(f)}. \tag{3-5}$$

We have that $\sum_f c(f) = 1$. The faces that are contained in the interior of the disk have an even number of sides, as their edges are partitioned into arcs of A_1 and arcs of A_2. Since the arcs of A_1 and A_2 realize their geometric intersection number, the interior faces have at least four sides. Hence the largest contribution of an interior face is 0. A face touching the boundary can have two sides, but these faces are cut off by a single component of A_1 or A_2, and contain a point of $\{p,q,r\}$ in their boundary face by the last condition. They contribute $\frac{1}{2}$ to the Euler characteristic of the disk. However, the edges involved in these pieces can be removed from the families A_1 and A_2 and the remaining curves still satisfy the hypotheses, so do this. Now, the only faces that contribute positively to the Euler characteristic of the disk are triangles with one edge on the boundary. These contribute $\frac{1}{4}$ to the Euler characteristic. There must be at least 4 such triangles. That means if the intersection of A_1 and A_2 is nonempty, then one of those triangles does not contain a point from $\{p,q,r\}$, so it is an outermost triangle. □

Proof of Theorem 3.4 Let S and S' be two simple diagrams, with associated colorings $f_S, f_{S'}\colon E \to \mathbb{Z}_{\geq 0}$, where E is the system of proper lines defining an ideal

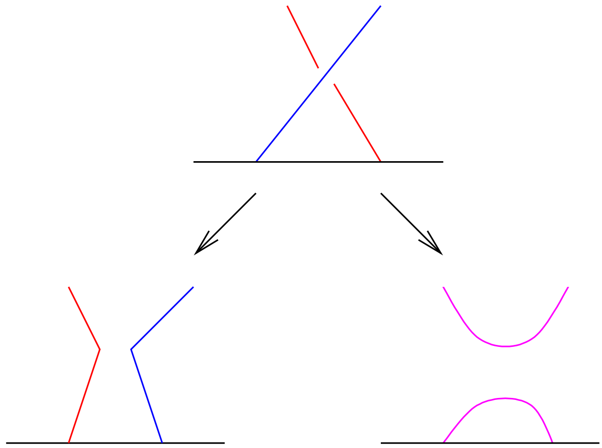


Figure 3: Resolving at an outermost triangle

triangulation with ideal triangles Δ_j. We do not need to distinguish between embedded and folded triangles for this proof, because the combinatorial lemma above is applied in the completed components of the complement of E. Isotope S and S' so that they are transverse to one another, and the lines in E, and realize all geometric intersection numbers $i(S,S')$, $i(S,c)$ and $i(S',c)$ for $c \in E$. Also make sure that $S \cap S' \cap E = \varnothing$. We resolve $S * S'$ one ideal triangle at a time. The four-valent graph $(S \cup S') \cap \Delta_j$ for each Δ_j satisfies the hypotheses of the lemma. To start with, A_1 is made up of the components of $S \cap \Delta_j$ and A_2 is made up of the components of $S' \cap \Delta_j$. Therefore we can find an outermost triangle $\Delta \subset \Delta_j$. If we resolve the crossing of $S * S'$ at the apex of the triangle there are two resolvents. One resolvent forms a bigon with the edge of the triangle, and hence any simple diagram descendent from this resolvent is strictly smaller in the ordering of diagrams than $[f_S + f_{S'}]$. This is shown in Figure 3

The other resolvent doesn't have a bigon. Any state resulting in a simple diagram whose coloring is $f_S + f_{S'}$ is a descendent of this resolvent. The triangle Δ has a face $p \subset S$ and a face $q \subset S'$. Assume that p lies in the component a of the family A_1 and q lies in the component b of A_2. We smooth by forming arcs $a - p \cup q$ and $b - q \cup p$ and then perturb them slightly so that they are disjoint. To continue on inductively, we declare that the perturbed version of $a - p \cup q$ is in A_1, whilst removing a, and the perturbation of $b - q \cup p$ is in A_2 and discard b. This operation does not produce any components of A_1 or A_2 that are simple closed curves inside the triangle Δ, because every component of the new families A_1 and A_2 still have two endpoints. Notice that the assignment of A_1 and A_2 is now just local to the ideal triangle instead of corresponding to the diagrams S and S'. However, we work ideal triangle by ideal triangle, so this isn't a problem. If the new graph has a crossing, it still satisfies the hypotheses of the lemma, so we can continue resolving crossings at the apex of an outermost triangle. There is a unique resolvent that can have a descendent with coloring $f_S + f_{S'}$. Continue until there are no crossings in Δ_j. There is a single resolvent with no bigons in Δ_j, so all the crossings in Δ_j have been resolved. All the other resolvents with no crossings in Δ_j have bigons in Δ_j and will lead to simple diagrams that are strictly smaller in the ordering of diagrams. Do this for each triangle. In the end, there is a single state with no bigons. The state must be a simple diagram. The construction did not produce any simple closed curves contained in a triangle. A simple closed curve that bounds a disk and has nonempty intersection with the edges of the triangulation must have a bigon, since a proper arc in a disk always separates the disk into two subdisks. Since there are no bigons between E and the edges of the triangulation, the admissible coloring associated to E is $f_S + f_{S'}$, so E is the geometric sum of S and S'. The coefficient of E is $\zeta^{p(E)-n(E)}$, where $p(E)$ is the number of positive smoothings and $n(E)$ is the number of negative smoothings that gave rise to the state E. The rest of the expansion is a linear combination of simple diagrams that are strictly smaller. □

Remark 3.6 A collection of skeins $\beta \in B$ spans $K_{\mathfrak{D}}(F)$ over $\mathfrak{D}$ if and only if the lead terms of the elements in β consist of units in $\mathfrak{D}$ times simple diagrams, and each isotopy class of simple diagrams appears at least once in the lead term of some $\beta \in B$.

Theorem 3.7 *Suppose that $\mathfrak{D}$ is an integral domain and $\zeta \in \mathfrak{D}$ is a unit and $2 \in \mathfrak{D}$ is a unit. Let S_i be a family of simple diagrams corresponding to the integral basis of the admissible colorings of an ideal triangulation. The skeins $\left\{\prod_i T_{k_i}(S_i)\right\}$, where the k_i range over all nonnegative integers, spans $K_{\mathfrak{D}}(F)$ over $\mathfrak{D}$.*

Proof The lead term of $T_{k_1}(S_1) * T_{k_2}(S_2) * \cdots * T_{k_n}(S_n)$ is a power of ζ times a simple diagram corresponding to the admissible coloring $\sum_i k_i f_{S_i}$, where f_{S_i} is the admissible coloring corresponding to S_i. Since the lead terms of these skeins correspond to all simple diagrams, we can inductively rewrite any skein as a linear combination of these by starting at the terms of highest weight. □

This extends a theorem of Bullock [4]. In that paper it is proved that the arbitrary products of a finite collection of curves S_i spans. Our theorem is stronger because we can specify the order of the product of the S_i, as no matter what order we work in, the leading terms are the same, though maybe with different powers of ζ as the lead coefficient. It could be that the integral basis of the space of admissible colorings is not linearly independent over $\mathbb{Z}$, so we don't have that the products form a basis.

3.3 The case when ζ is a primitive n^{th} root of unity

Now we go on to study $K_\zeta(F)$, meaning the coefficients are $\mathbb{Z}\big[\frac{1}{2},\zeta\big]$, where ζ is a primitive n^{th} root of unity, $n \neq 0 \bmod 4$, and A is set equal to ζ. Recall $\chi(F)$ is the image of the threading map

$$\text{Ch}\colon \overline{K}_\epsilon(F) \to K_\zeta(F). \tag{3-6}$$

Recall the overline is to indicate that we have extended the scalars of $K_\epsilon(F)$ to include $\frac{1}{2}$ and ζ. The map Ch threads every component of a framed link corresponding to a simple diagram with $T_m(x)$, where $m = n/\gcd(n,4)$, the index of threading. Since

$$T_k(x) = \sum_{i=0}^{\lfloor k/2 \rfloor} (-1)^i \frac{k}{k-i}\binom{k-i}{i} x^{k-2i}, \tag{3-7}$$

the lead term of $\text{Ch}(S)$ where the simple diagram has admissible coloring $f_S\colon E \to \mathbb{Z}_{\geq 0}$ of weight i is $[mf_S]$.

Let S be a simple diagram with associated coloring $f_S\colon E \to \mathbb{Z}_{\geq 0}$. Assume that f_S is not identically zero. The integers $\{f_S(c)\}_{c\in C}$ generate a subgroup of $\mathbb{Z}$, which, being cyclic, has a smallest positive generator, denoted by $\gcd(f_S)$.

Proposition 3.8 *If $n > 0$ is odd and $n \mid \gcd(f_S)$ then $\frac{1}{n} f_S\colon E \to \mathbb{Z}_{\geq 0}$ is an admissible coloring with associated simple diagram S' and $(S')^n = S \in K_\zeta(F)$.*

Proof Since F is orientable, the diagram S' is two-sided, meaning that we can push it completely off of itself to take the product. This means that the admissible coloring of $(S')^n$ is $nf_{S'}\colon E \to \mathbb{Z}_{\geq 0}$. If $f_S\colon E \to \mathbb{Z}_{\geq 0}$ is an admissible coloring, and for all $c \in C$, the odd integer $n \mid f(c)$, then, for any $\{a,b,c\} = \partial\Delta$ of an embedded ideal triangle in the triangulation, $\big\{\frac{1}{n} f_S(a), \frac{1}{n} f_S(b), \frac{1}{n} f_S(c)\big\}$ satisfy all three triangle inequalities as the triangle inequality is linear. The sum $\frac{1}{n}(f_S(a) + f_S(b) + f_S(c))$ is even, as an even number divided by an odd number is even. Similarly, $\frac{1}{n} f_S\colon E \to \mathbb{Z}_{\geq 0}$ satisfies the conditions to be admissible for folded triangles. $\square$

The restriction to $n \neq 0 \bmod 4$ means that m is always odd, so Proposition 3.8 applies. If S is a simple diagram associated to the admissible coloring

$$f_S\colon E \to \mathbb{Z}_{\geq 0} \tag{3-8}$$

then, as noted above, the lead term of $\text{Ch}(S)$ is $[mf_S]$, and m divides $\gcd(mf_S)$.

Theorem 3.9 *Let F be a finite-type surface with at least one puncture, that has been ideally triangulated. If S_i is any system of simple diagrams corresponding to an integral basis of the cone of admissible colorings of the triangulation, then the skeins $\prod_i T_{k_i}(S_i)$, where the $k_i \in \{0, 1, \dots, m-1\}$, span $K_\zeta(F)$ over $\chi(F)$. In particular, $K_\zeta(F)$ is a finite ring extension of $\chi(F)$.*

Proof The proof is by induction on largest diagram appearing with nonzero coefficient in a skein. Start with a skein written in terms of the basis over $\mathbb{Z}\left[\frac{1}{2}, \zeta\right]$ of simple diagrams,

$$\sum_j \alpha_j \prod_i T_{k_{i,j}}(S_{i,j}), \tag{3-9}$$

with $\alpha_j \in \mathbb{Z}\left[\frac{1}{2}, \zeta\right]$. Suppose the lead term of the skein is indexed by j. Since

$$T_{m+k}(x) = T_m(x) * T_k(x) - T_{|m-k|}(x) \tag{3-10}$$

if some $k_{i,j} \geq m$, then, as $\chi(F)$ is central, we can factor out an element of $\chi(F)$ from the term to get a simple diagram of lower weight. Continue on inductively till the skein is written as

$$\sum_j \beta_j \prod_i T_{k_{i,j}}(S_{i,j}), \tag{3-11}$$

where all $k_{i,j} \in \{0, 1, \dots, m-1\}$ and $\beta_j \in \chi(F)$. □

Theorem 3.10 *If F is closed, and ζ is a primitive n^{th} root of unity, where $n \neq 0 \bmod 4$, then $K_\zeta(F)$ is a finite-rank module over $\chi(F)$.*

Proof If F is closed and $p \in F$, then the inclusions $K_\zeta(F - \{p\}) \to K_\zeta(F)$ and $\chi(F - \{p\}) \to \chi(F)$ are surjective homomorphisms that fit into a commutative square

$$\begin{array}{ccc} K_\zeta(F - \{p\}) & \longrightarrow & K_\zeta(F) \\ \uparrow & & \uparrow \\ \chi(F - \{p\}) & \longrightarrow & \chi(F) \end{array} \tag{3-12}$$

After choosing an ideal triangulation for $F - \{p\}$, if the admissible colorings associated with S_i form an integral basis then $T_{k_1}(S_1) * \cdots * T_{k_n}(S_n)$, where the $k_i \in \{0, \dots, m-1\}$, span $K_\zeta(F)$ over $\chi(F)$. □

In [8] we prove that $K_\zeta(\Sigma_{1,0})$ is not free over $\chi(\Sigma_{1,0})$, so there are definitely linear dependencies between the elements of the spanning set produced this way.

Theorem 3.11 *For every* $\phi\colon \chi(F) \to \mathbb{C}$, $K_\zeta(F)_\phi$ *is a finite-dimensional algebra over the complex numbers.*

Proof This follows from the definition of specialization. □

Let F be a finite-type surface of negative Euler characteristic and at least one puncture. Let E be the edges of an ideal triangulation of F. Recall that $\mathcal{A}$ denotes the admissible colorings of E. After ordering the set E, you can view $\mathcal{A} \subset \mathbb{Z}^E_{\geq 0}$. This allows us to define a map

$$\text{res}\colon \mathcal{A} \to \mathbb{Z}^E_m, \tag{3-13}$$

by sending each admissible coloring to the tuple of residues of its values modulo m. This is used to define

$$\text{res}\colon K_\zeta(F) - \{0\} \to \mathbb{Z}^E_m. \tag{3-14}$$

Every nonzero skein α can be written $\sum_S z_S S$ where the z_S are nonzero complex numbers and the S are simple diagrams, and the sum is nonempty. Let $f_S\colon E \to \mathbb{Z}$ be the admissible coloring of the diagram appearing in the lead term of α. Define $\text{res}(\alpha)$ to be $\text{res}(f_S)$.

Lemma 3.12 *If* $\alpha, \beta \in K_\zeta(F) - \{0\}$ *then* $\text{res}(\alpha * \beta) = \text{res}(\alpha) + \text{res}(\beta)$.

Proof Suppose that the lead term of α is $z_S S$ and the lead term of β is $w_T T$. If the admissible colorings corresponding to S and T are f_S and f_T, then, by Theorem 3.4, the diagram underlying the lead term of $\alpha * \beta$ has coloring $f_S + f_T$. Hence $\text{res}(\alpha * \beta) = \text{res}(f_S + f_T) = \text{res}(f_S) + \text{res}(f_T)$. □

Theorem 3.13 *Suppose that* $\{\alpha_i\}$ *is a collection of nonzero skeins and the restriction of* $\text{res}\colon K_\zeta(F) - \{0\} \to \mathbb{Z}^E_m$ *to* $\{\alpha_i\}$ *is one-to-one. The collection of skeins* $\{\alpha_i\}$ *is linearly independent over* $\chi(F)$.

Proof Suppose that $\sum_i \beta_i \alpha_i = 0$ with the $\beta_i \in \chi(F)$. This means that the lead term of $\sum_i \beta_i \alpha_i$ is equal to zero. However if $\beta_i \in \chi(F) - \{0\}$ the coloring corresponding to its lead term is divisible by m by Proposition 3.8. Hence $\text{res}(\beta_i * \alpha_i) = \text{res}(\beta_i) + \text{res}(\alpha_i) = \text{res}(\alpha_i)$. Since the $\text{res}(\alpha_i)$ are all distinct, there can be no cancellation among the leading terms of the sum, and in fact all the $\beta_i = 0$. □

4 Computing the trace

Recall if $\alpha \in K_\zeta(F)$, there is a $S^{-1}\chi(F)$–linear map,

$$L_\alpha\colon S^{-1} K_\zeta(F) \to S^{-1} K_\zeta(F), \tag{4-1}$$

given by left multiplication. If d is the dimension of $S^{-1}K_\zeta(F)$ as a vector space over $S^{-1}\overline{K}_\epsilon(F)$ then, by definition, the normalized trace of L_α is $\frac{1}{d}\operatorname{tr}(L_\alpha)$, where tr denotes the standard trace. The goal of this section is to compute $\operatorname{Tr}(L_\alpha)$. The correct basis for $K_\zeta(F)$ over $\mathcal{D}$ is the primitive diagrams P whose components S_i have been threaded by $T_{k_i}(x)$. If α has been written as a linear combination of such P, then the normalized trace of L_α is the result of crossing out all terms where some component of the diagram has been threaded with a k_i that is not divisible by m.

Since the trace is linear, we only need to compute the trace of skeins of the form $\prod_i T_{k_i}(S_i)$, where the $\{S_i\}$ for $i=1,\dots,n$ are the components of a primitive diagram. The strategy is to prove that the subalgebra of $K_\zeta(F)$ obtained by adjoining the curves S_i to $\chi(F)$, denoted by $\chi(F)[S_1,\dots,S_n]$, is isomorphic to the tensor product of n copies of the skein algebra of the annulus $\Sigma_{0,2}$ with its coefficients extended to $\chi(F)$. The trace of the tensor product of linear endomorphisms is the product of the traces of the endomorphisms. Therefore the trace of $\prod_i T_{k_i}(S_i)$ is the product of the traces of its individual factors. Once we localize at $S=\chi(F)-\{0\}$, the subalgebra $S^{-1}\chi(F)[S_1,\dots,S_n]$ is a commutative ring that is a finite extension of the field $S^{-1}\chi(F)$, having no zero divisors. This means that $S^{-1}\chi(F)[S_1,\dots,S_n]$ is a field. Hence $S^{-1}K_\zeta(F)$ is a finite-dimensional vector space over $S^{-1}\chi(F)[S_1,\dots,S_n]$, which in turn is a finite-dimensional vector space over $S^{-1}\chi(F)$. This is exactly the computational setting for Proposition 2.3.

Since m is not necessarily prime, $\mathbb{Z}_m$ might have zero divisors. Hence, linear independence in a module over $\mathbb{Z}_m$ is subtle. Since $\mathbb{Z}_m^E$ is a free module over $\mathbb{Z}_m$, there are linearly independent subsets. Let $\vec{e}_c\in\mathbb{Z}_m^E$ be the vector whose entries are all 0 except for a 1 in the c^{th} entry. The vectors $\vec{e}_c$ form a basis for $\mathbb{Z}_m$.

Lemma 4.1 *Choose an ordering of E. Let $V=\{\vec{v}_i\}\in\mathbb{Z}_m^E$ be a collection of vectors indexed by an initial segment of the counting numbers. There is a map $I\colon V\to E$ the sends each $\vec{v}_i$ to the index of its first nonzero entry. If I is increasing and the first nonzero entry of each $\vec{v}_i$ is a unit in $\mathbb{Z}_m$, then V is linearly independent over $\mathbb{Z}_m$.*

Proof Adjoin those $\vec{e}_c$ to V that don't appear as a first nonzero entry. The determinant of the $n\times n$ matrix you get this way is a unit. Therefore it is a basis of $\mathbb{Z}_m^E$. The original $\{\vec{v}_i\}$ is independent as any subset of a basis is independent. □

Proposition 4.2 *Suppose that S_i is an ordered collection of disjoint simple closed curves on the finite-type surface F. Suppose further that there is an ideal triangulation of F with ordered set of edges E such that the map $I\colon\{S_i\}\to E$ that sends each curve S_i to the smallest edge in E that it has nonzero geometric intersection number*

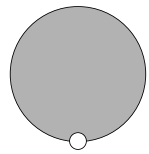


Figure 4: A monogon

with is increasing, and the geometric intersection number of S_i *with* $I(S_i)$ *is always* 1 *or* 2. *The set of skeins* $\left\{\prod_i T_{k_i}(S_i)\right\}$, *where the* k_i *range from* 0 *to* $m-1$, *are linearly independent in* $K_\zeta(F)$.

Proof By Lemma 4.1, the vectors $\{\mathrm{res}(S_i)\}$ are linearly independent in $\mathbb{Z}_m^E$. This implies that the vectors $\left\{\sum_i k_i \,\mathrm{res}(S_i)\right\}$, where the k_i range over 0 to $m-1$, are all distinct. However, the residue of $\prod_i T_{k_i}(S_i)$ is equal to $\sum_i k_i \,\mathrm{res}(S_i)$, hence the residues of the $\left\{\prod_i T_{k_i}(S_i)\right\}$ are distinct. By Theorem 3.13, the set $\left\{\prod_i T_{k_i}(S_i)\right\}$ is independent in $K_\zeta(F)$. □

The next several paragraphs are to prove that if the $\{S_i\}$ are the components of simple diagram, then we can find a triangulation E such that the hypotheses of the last proposition hold true for a choice of orderings for $\{S_i\}$ and E.

Suppose that E is a properly embedded system of disjoint lines in the finite-type surface F. A *monogon* is a component of the complement of E that completes to a closed disk with a single point removed from its boundary. We show a monogon in Figure 4.

A *bigon* is a component of the complement of E that completes to a closed disk with two points removed from its boundary. We show a bigon in Figure 5.

Proposition 4.3 *Suppose that* E *is a properly embedded system of disjoint lines in the finite-type surface* F *whose complement has no monogons or bigons. There exists a collection* D *of properly embedded lines such that* $C \cup D$ *defines an ideal triangulation of* F. □

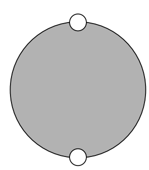


Figure 5: A bigon

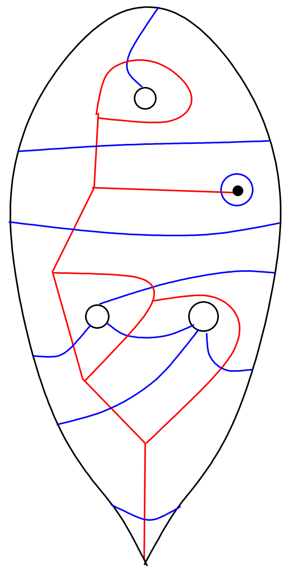


Figure 6: A filling diagram and its dual graph

Suppose that $P \subset F$ is a primitive diagram. We say that P *fills* F if the components of $F - P$ consist of once-punctured disks about the punctures of F and planar surfaces of Euler characteristic -1.

Theorem 4.4 *Suppose that F is a finite-type surface of negative Euler characteristic with at least one puncture and P fills F. There is an ordering of the disjoint curves S_i that make up P and a collection of disjoint embedded lines c_i such that if $i < j$ then $i(S_j, c_i) = 0$, and $i(S_i, c_i)$ is 1 or 2. Since no two of the c_i are parallel and all of the c_i are essential, the collection c_i can be built up to be an ideal triangulation of F.*

If P fills F, there is a dual 1–dimensional CW–complex, with a 0–cell for every component of the complement of P and a 1–cell for every component of P. The trivalent 0–cells of the CW–complex correspond to components of the complement that complete to pants. The monovalent 0–cells correspond to components of the complement that complete to a punctured disk. If a 1–cell has both its endpoints at the same 0–cell, the corresponding simple closed curve is a nonseparating curve lying in the closure of a component of the complement of P that is homeomorphic to $\Sigma_{1,1}$. The CW–complex minus its valence-one vertices can be properly embedded in the surface F, where each edge intersects the corresponding simple closed curve once in a transverse point of intersection and the trivalent vertices embedded in the

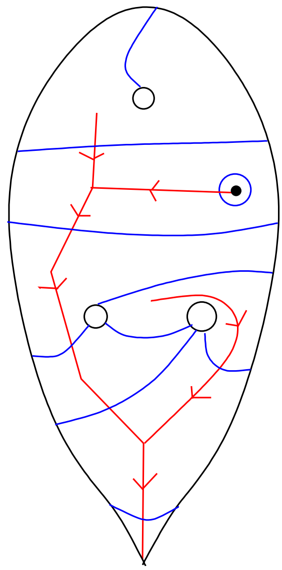


Figure 7: A maximal rooted tree

corresponding components of the complement of P, and the ends of the deleted CW–complex mapped to the ends of the corresponding disk with a point deleted. The edges of the CW–complex are in one-to-one correspondence with the components of P. If the edge e and the component S intersect one another, we say that they are *dual*. The intersection is necessarily a single point of transverse intersection.

In Figure 6 we show a twice-punctured surface of genus three. The filling diagram is in blue and the embedded dual graph is red.

Choose a maximal tree of the CW–complex and a valence-one 0–cell. Orient the tree so that it is rooted at the chosen 0–cell. That is, every edge is oriented so that it points towards the root. The monovalent 0–cells of the tree that are sources are the *leaves* of the tree. The rooted tree is in red.

We will build a train track from this tree as shown in Figure 7.

Figure 8 is color-coded so that each of the following steps is visible. First smooth the vertices of the tree so that the two edges pointing into each interior 0–cell have the same outward-pointing tangent vector. Next, for each component of the diagram that doesn't bound a punctured disk and is dual to an edge of the tree, push it off itself towards the root, and then put a kink in it where it intersects the edge dual to it and smooth the kink to get a switch where both outward normals of the curve at the kink point towards the root. These are in magenta. Next, add the remaining edges of the

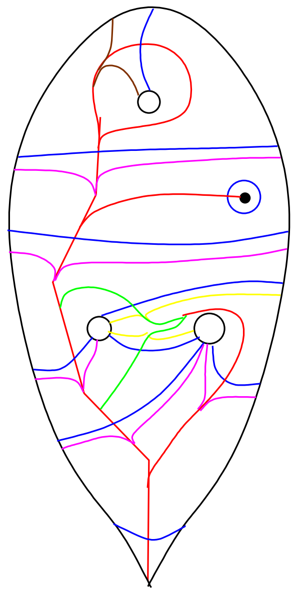


Figure 8: The train track

CW–complex, so that their outward normals, at the switches created, point towards the root. These are in green. If both endpoints of the edge are attached at the same 0–cell, that edge e lies in the closure of a component of the complement of P that is a torus. If S is the dual edge, push it off of itself and add a kink where it intersects e so that the outward tangent vectors point towards the vertex in the torus component. This is in brown. Suppose now that the 0–cells of the tree that e is attached at are distinct. For each one of those 0–cells that is a leaf, add a branch to the track, which is a pushoff of the dual component of P, with a kink in it that makes a switch in the train track pointing at that 0–cell. These are in yellow.

We produce a family of disjoint properly embedded lines by splitting the tree at the switches and cutting open all the way to the root. The switches in the tree point towards the roots, and the switches in the additional edges point towards the tree, so the process of cutting open along switches terminates at the root, and we have produced a family of disjoint properly embedded lines. The train track does not carry any simple closed curves.

Order the components of P so that S and T are dual to edges in the tree then their relative order is consistent with their distance from the root of the tree, and if they aren't dual to edges of the tree then they come after all the components that are dual to edges of the tree. Working in order we prove that, given a component S of P, there is

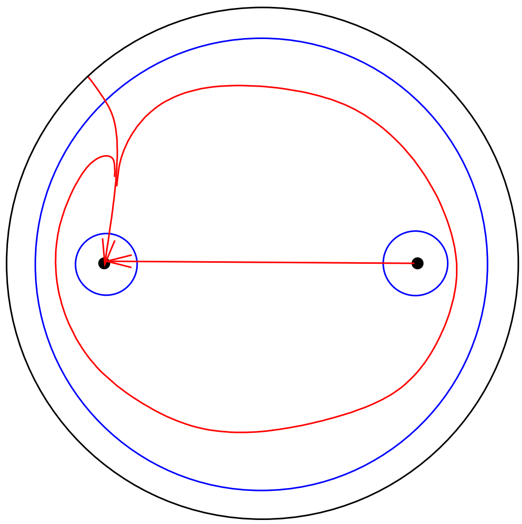


Figure 9: An edge that is simultaneously initial and a leaf

a line c_S in our family such that $i(X, c_S)$ is 1 or 2, and if $T > S$ then $c_S \cap T = \varnothing$, or we exchange order so that we can do so.

Since the lines c_S are indexed by S, the condition on intersections implies that no line is homotopically trivial (bounds a monogon) and no two lines are parallel (cobound a bigon), so the family c_S can be built up to a triangulation. The complication of the construction is that to construct the line for a given edge in the tree we need to understand what immediately follows the edge in the ordering.

We start at the root. If an edge leaving the root is a leaf in the tree, there are three possible cases. The surface could be a once-punctured torus, or a thrice-punctured sphere, or the terminal points of the edge are at punctures, and the punctured disks containing those punctures abut the same pair of pants. The construction for the punctured torus, and thrice-punctured pair of pants can be done by inspection. We focus on the last case, shown in Figure 9.

According to our rules either the edge of the tree dual to the blue curve parallel to the outer boundary or the line joining to the two punctures could come first. You really want the edge dual to the curve parallel to the outer boundary component to be first. If S_1 is the component of P that bounds the punctured disk at the root, let c_{S_1} be the line built from the branch of the track that follows the outer boundary component before heading to the puncture. Notice $i(c_{S_1}, S_1) = 2$. The circle S_2 surrounding the other puncture has geometric intersection number 1 with the line c_{S_2} having one end

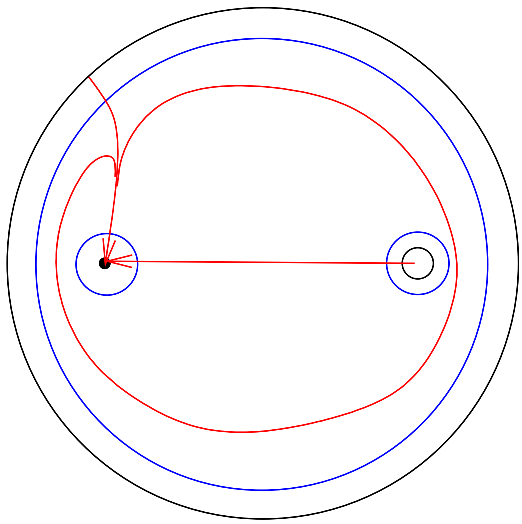


Figure 10: The case when the edge leaving the root is not a leaf

at each puncture. Since the line c_{S_1} is completely inside the diagram we have that it has geometric intersection number 0 with all later curves.

Now suppose that the edge leaving the root is not a leaf. In Figure 10 we show the situation. The line c_{S_1} coming from the branch of the track that runs around the outer boundary component has geometric intersection number 2 with S_1 and misses all the other components of the filling diagram, and for all later components it has geometric intersection number 0.

An edge dual to S_i of the tree is intermediate if there is an edge dual to S_{i-1} before it and an edge dual to S_{i+1} after it in the tree from the ordering. Let c_{S_i} be the line coming from the branch of the track that was built by perturbing S_{i+1}. Notice $i(c_{S_i}, S_i) = 2$. Do all intermediate edges before doing the leaves.

If an edge is a leaf, then it could end at a puncture, it could have both ends of an edge not part of the tree attached at its terminal 0–cell, or it could have two different edges not in the tree attached at its terminal end. These both occur in Figure 7. The highest leaf in the diagram is of the first type, and the lower leaf is of the second kind.

In the first case, the component S of P bounds a punctured disk. The line c_S of the train track that emanates from that puncture and ends at the root has geometric intersection number 1 with S.

In the second case, the vertex of the edge lies in a torus that is the closure of a component of $F-P$. Call the curve dual to the edge with both its ends attached at that 0–cell S'.

The pushoff of S' gives rise to the line c_S that has geometric intersection 2 number with S that is dual to the edge.

In the third case, let S' be the component of P that is dual to one of the edges attached at the leaf. The pushoff of S' towards the leaf gives rise to an embedded line that has geometric intersection number 2 with the curve S dual to the edge.

Throw out any curves that weren't used. Augment to form an ideal triangulation. □

Given a primitive diagram $P = \{S_1, \dots, S_n\}$, we can form the subalgebra of $K_\zeta(F)$,

$$\mathcal{P} = \chi(F)[S_1, \dots, S_n]. \tag{4-2}$$

This means we are taking the smallest subalgebra of $K_\zeta(F)$ that contains all $\chi(F)$–linear combinations of the S_i. Notice that $\mathcal{P}$ is a commutative ring, since the S_i are disjoint from one another. Also,

Form the ring

$$\bigotimes_{i=1}^{n} \chi(F)[S_i], \tag{4-3}$$

where the tensor product is as $\chi(F)$–modules. There is a ring homomorphism

$$\Psi\colon \bigotimes_{i=1}^{n} \chi(F)[S_i] \to \chi(F)[S_1, \dots, S_n] \tag{4-4}$$

given by

$$\Psi(\alpha_1 \otimes \cdots \otimes \alpha_n) = \alpha_1 * \alpha_2 * \cdots * \alpha_n. \tag{4-5}$$

Proof This is an immediate consequence of Proposition 4.2. □

Theorem 4.5 *If S_i for $i \in \{1, \dots, n\}$ is a system of simple closed curves on F that forms a primitive diagram then $\chi(F)[S_1, \dots, S_n]$ is a field extension of $\chi(F)$ of dimension m^n, then*

$$\Psi\colon \bigotimes_{\chi(F)} \chi(F)[S_i] \to \chi(F)[S_1, \dots, S_n] \tag{4-6}$$

is an isomorphism

Proof We can complete $\{S_i\}$ to a filling diagram. We apply Theorem 4.4 to get a system of curves of $\{S_i\}$ that satisfies the hypotheses of Proposition 4.2. □

Corollary 4.6 *The ring $S^{-1}\chi(F)[S_1, \dots, S_n]$, is a field extension of degree m^n over $S^{-1}\chi(F)$.*

Proof By Theorem 4.5, the ring $S^{-1}\chi(F)[S_1,\dots,S_n]$ is a commutative algebra over $S^{-1}\chi(F)$ that has dimension m^n as a vector space. Since $K_\zeta(F)$ has no zero divisors, neither does $S^{-1}\chi(F)[S_1,\dots,S_n]$. A finite commutative extension of a field is a field. $\square$

If $S\subset F$ is a nontrivial simple closed curve, let $\Sigma_{0,2}(S)$ be an annular neighborhood of S in F. There is a left action of $K_\zeta(\Sigma_{0,2}(S))\otimes K_\zeta(F)\to K_\zeta(F)$ by gluing a copy of $\Sigma_{0,2}(S)\times[0,1]$ onto the top of $F\times[0,1]$. Notice that it restricts to give an action $\chi(\Sigma_{0,2}(S))$ on $\chi(F)$ making $S^{-1}\chi(\Sigma_{0,2}(S))\le S^{-1}\chi(F)$ a field extension.

Remark 4.7 It is worth mentioning that

$$S^{-1}\chi(\Sigma_{0,2}(S))[S]=S^{-1}K_\zeta(\Sigma_{0,2}(S)). \tag{4-7}$$

Theorem 4.8 *$S^{-1}\chi(F)[S]$ is the result of extending the coefficients of the ring $S^{-1}\chi(\Sigma_{0,2}(S))[S]$ as a vector space over $S^{-1}\chi(\Sigma_{0,2}(S))$ to a vector space over $S^{-1}\chi(F)$.*

Proof The dimension of $S^{-1}\chi(\Sigma_{0,2}(S))[S]$ over $S^{-1}\chi(\Sigma_{0,2}(S))$ is equal to the dimension of $S^{-1}\chi(F)[S]$ over $S^{-1}\chi(F)$, so the map

$$S^{-1}\chi(\Sigma_{0,2}(S))[S]\otimes_{S^{-1}\chi(\Sigma_{0,2})}S^{-1}\chi(F)\to\chi(F)[S] \tag{4-8}$$

that sends $S\otimes 1$ to S is a linear isomorphism. $\square$

From our last paper:

Proposition 4.9 [8] *If $\Sigma_{0,2}$ is an annulus and x is the skein at its core and*

$$\mathrm{tr}\colon K_\zeta(\Sigma_{0,2})\to\chi(\Sigma_{0,2}) \tag{4-9}$$

is the unnormalized trace, $\mathrm{tr}(L_{T_k(x)})=0$ unless $m\,|\,k$, at which point $\mathrm{tr}(L_{T_k(x)})=mT_k(x)$. $\square$

This implies the same result for $T_k(S)\colon S^{-1}K_\zeta(\Sigma_{0,2}(S))\to S^{-1}K_\zeta(\Sigma_{0,2}(S))$.

Proposition 4.10 *Let $S\subset F$ be a nontrivial simple closed curve. Define the map $L_{T_k(S)}\colon S^{-1}\chi(F)[S]\to S^{-1}\chi(F)[S]$ by left multiplication; then $\mathrm{tr}(L_{T_k(S)})=0$ unless $m\,|\,k$, at which point $\mathrm{tr}(L_{T_{k(S)}})=mT_k(S)$.*

Proof The map $L_{T_k(S)}\colon S^{-1}\chi(F)[S]\to S^{-1}\chi(F)[S]$ comes from

$$L_{T_k(S)}\colon S^{-1}\chi(\Sigma_{0,2}(S))[S]\to S^{-1}\chi(\Sigma_{0,2}(S))[S] \tag{4-10}$$

by extension of scalars and the fact that $\chi(\Sigma_{0,2}(S))[S]=K_\zeta(\Sigma_{0,2}(S))$. $\square$

Proposition 4.11 *Let $\prod_i T_{k_i}(S_i)$ act on $S^{-1}\chi(F)[S_1,\dots,S_n]$ by multiplication,*

$$L_{\prod_i T_{k_i}(S_i)}\colon S^{-1}\chi(F)[S_1,\dots,S_n] \to S^{-1}\chi(F)[S_1,\dots,S_n]. \tag{4-11}$$

Then the unnormalized trace of $L_{k_1,\dots,k_n}$ is zero unless $m\,|\,k_i$ for all i, in which case it is $m^n\prod_i T_{k_i}(S_i)$.

Proof The diagram

$$\begin{array}{ccc} \bigotimes_{\chi(F)} \chi(\Sigma_{0,2}(S_i))[S_i] \otimes_{\chi(\Sigma_{0,2}(S_i))} \chi(F) & \xrightarrow{\psi} & \chi(F)[S_1,\dots,S_n] \\ {\scriptstyle \otimes L_{T_{k_i}(S_i)}}\big\downarrow & & \big\downarrow{\scriptstyle L_{\prod_i T_{k_i}(S_i)}} \\ \bigotimes_{\chi(F)} \chi(\Sigma_{0,2}(S_i))[S_i] \otimes_{\chi(\Sigma_{0,2}(S_i))} \chi(F) & \xrightarrow{\psi} & \chi(F)[S_1,\dots,S_n] \end{array} \tag{4-12}$$

where ψ is the natural isomorphism, commutes. This means that the trace of $L_{\prod_i T_{k_i}(S_i)}$ is the product of the traces of the

$$\begin{aligned} L_{T_{k_i}(S_i)}\colon \chi(\Sigma_{0,2}(S_i))[S_i] &\otimes_{\chi(\Sigma_{0,2}(S_i))} \chi(F) \\ &\to \chi(\Sigma_{0,2}(S_i))[S_i] \otimes_{\chi(\Sigma_{0,2}(S_i))} \chi(F) \end{aligned} \tag{4-13}$$

which are obtained by extension of scalars from

$$L_{T_{k_i}(S_i)}\colon K_\zeta(\Sigma_{0,2}(S_i)) \to K_\zeta(\Sigma_{0,2}(S_i)). \qquad \square$$

Theorem 4.12 *Suppose that $d = [S^{-1}\chi(F)[S_1,\dots,S_n] : S^{-1}K_\zeta(F)]$. The unnormalized trace of*

$$L_{k_1,\dots,k_n}\colon S^{-1}K_\zeta(F) \to S^{-1}K_\zeta(F) \tag{4-14}$$

is zero unless $m\,|\,k_i$ for all i, in which case it is $dm^n\prod_i T_{k_i}(S_i)$.

Proof By Theorem 3.9, $S^{-1}K_\zeta(F)$ is a finite-dimensional vector space over

$$S^{-1}\chi(F) \le S^{-1}\chi(F)[S_1,\dots,S_n], \tag{4-15}$$

so Proposition 2.3 applies. $\square$

We define the normalized trace

$$\mathrm{Tr}\colon S^{-1}K_\zeta(F) \to S^{-1}\chi(F) \tag{4-16}$$

to be the trace divided by dm^n. The map Tr is $S^{-1}\chi(F)$–linear, cyclic, and $\mathrm{Tr}(1) = 1$.

Theorem 4.13 *Suppose that $s=\sum_i \beta_i P_i$ where the $\beta\in S^{-1}\chi(F)$ and the P_i are primitive diagrams whose components have been threaded with T_k. Let J be those indices i such that the components of P_i have only been threaded with T_k for $m\,|\,k$; then*

$$\operatorname{Tr}(s)=\sum_{i\in J}\beta_i P_i. \tag{4-17}$$

Theorem 4.14 *The restriction of* Tr: $S^{-1}K_\zeta(F)\to S^{-1}\chi(F)$ *to* $K_\zeta(F)$, *embedded in* $S^{-1}K_\zeta(F)$ *as fractions having denominator* 1, *yields*

$$\operatorname{Tr}\colon K_\zeta(F)\to\chi(F), \tag{4-18}$$

which is a $\chi(F)$*–linear map, so that* $\operatorname{Tr}(1)=1$ *and, for every* $\alpha,\beta\in K_\zeta(F)$,

$$\operatorname{Tr}(\alpha*\beta)=\operatorname{Tr}(\beta*\alpha). \tag{4-19}$$

Proof From the formula for Tr, the only fractions that appear in the coefficients in the trace come from fractions that are in the coefficients of the skein. □

5 The trace is nondegenerate

Lemma 5.1 *Let F be a finite-type surface with an ideal triangulation cut out by E. Suppose that $\sum_i z_i S_i\in K_\zeta(F)$, where the $z_i\in\mathbb{Z}\left[\frac{1}{2},\zeta\right]$ and the S_i are distinct simple diagrams. If some $[f_S]$ appearing in the symbol of $\sum_i z_i S_i$ with nonzero coefficient z has $m\,|\,\gcd(f_S)$, then*

$$\operatorname{Tr}\Big(\sum_i z_i S_i\Big)\neq 0. \tag{5-1}$$

Proof Suppose that the primitive diagram P underlying S is made up of simple closed curves S'_j. The threaded diagram having lead coefficient S is $\prod_j T_{mk_j}(S'_j)$ for some $k_j\in\mathbb{Z}_{\geq 0}$. Rewriting $\sum_i z_i S_i$ in terms of threaded diagrams, the threaded diagrams appearing in the symbol appear with the same coefficients and are distinct from one another in the sum. Hence $\prod_j T_{mk_j}(S'_j)$ appears in the trace with coefficient $z\neq 0$. This term can't cancel with other highest-weight terms in the trace, as the S_i were distinct, nor can it cancel with lower-weight terms, as that would violate the filtration of $\chi(F)$, so $\operatorname{Tr}\big(\sum_i z_i S_i\big)\neq 0$. □

Theorem 5.2 *Let F be a noncompact, finite-type surface. There are no nontrivial principal ideals in the kernel of*

$$\operatorname{Tr}\colon K_\zeta(F)\to\chi(F). \tag{5-2}$$

Proof Let $\alpha \in K_\zeta(F)$ be nonzero. Choosing an ideal triangulation and an ordering of edges, we can write $\alpha = \sum_i z_i P_i$ where the z_i are nonzero complex numbers and the P_i are threaded primitive diagrams. Suppose that the lead term of α is $z_* P_*$. Let P' be a threaded primitive diagram such that the residue of P' in $\mathbb{Z}_m^E$ is the additive inverse of the residue of P_*. If I is the principal ideal generated by α, then $P' * \alpha$ is in the principal ideal generated by α, and the residue of its lead term is zero. Since α was an arbitrary nonzero skein, there does not exist a nontrivial principal ideal of $K_\zeta(F)$ contained in the kernel of Tr. □

Corollary 5.3 *$S^{-1}K_\zeta(F)$ is a symmetric Frobenius algebra over $S^{-1}\chi(F)$.* □

Corollary 5.4 *There is a proper subvariety of the character variety of $\pi_1(F)$ away from which $K_\zeta(F)_\phi$ is a symmetric Frobenius algebra over $\mathbb{C}$.* □

References

[1] **J W Barrett**, *Skein spaces and spin structures*, Math. Proc. Cambridge Philos. Soc. 126 (1999) 267–275 MR

[2] **F Bonahon**, **H Wong**, *Representations of the Kauffman bracket skein algebra, I: Invariants and miraculous cancellations*, Invent. Math. 204 (2016) 195–243 MR

[3] **D Bullock**, *Rings of* $\mathrm{SL}_2(\mathbf{C})$*–characters and the Kauffman bracket skein module*, Comment. Math. Helv. 72 (1997) 521–542 MR

[4] **D Bullock**, *A finite set of generators for the Kauffman bracket skein algebra*, Math. Z. 231 (1999) 91–101 MR

[5] **D Bullock**, **C Frohman**, **J Kania-Bartoszyńska**, *Understanding the Kauffman bracket skein module*, J. Knot Theory Ramifications 8 (1999) 265–277 MR

[6] **D Bullock**, **J H Przytycki**, *Multiplicative structure of Kauffman bracket skein module quantizations*, Proc. Amer. Math. Soc. 128 (2000) 923–931 MR

[7] **A Fathi**, **F Laudenbach**, **V Poénaru**, *Travaux de Thurston sur les surfaces*, Astérisque 66, Soc. Math. France, Paris (1979) MR

[8] **C Frohman**, **N Abdiel**, *Frobenius algebras derived from the Kauffman bracket skein algebra*, J. Knot Theory Ramifications 25 (2016) art. id. 1650016 MR

[9] **C Frohman**, **R Gelca**, *Skein modules and the noncommutative torus*, Trans. Amer. Math. Soc. 352 (2000) 4877–4888 MR

[10] **C Frohman**, **J Kania-Bartoszynska**, *The structure of the Kauffman bracket Skein algebra at roots of unity*, preprint (2016) arXiv

[11] **P Gordan**, *Ueber die Auflösung linearer Gleichungen mit reellen Coefficienten*, Math. Ann. 6 (1873) 23–28 MR

[12] **J Hoste**, **J H Przytycki**, *The* $(2,\infty)$*–skein module of lens spaces; a generalization of the Jones polynomial*, J. Knot Theory Ramifications 2 (1993) 321–333 MR

[13] **T T Q Lê**, *On Kauffman bracket skein modules at roots of unity*, Algebr. Geom. Topol. 15 (2015) 1093–1117 MR

[14] **G Muller**, *Skein and cluster algebras of marked surfaces*, Quantum Topol. 7 (2016) 435–503 MR

[15] **J H Przytycki**, **A S Sikora**, *On skein algebras and* $\mathrm{Sl}_2(\mathbf{C})$*–character varieties*, Topology 39 (2000) 115–148 MR

[16] **A S Sikora**, **B W Westbury**, *Confluence theory for graphs*, Algebr. Geom. Topol. 7 (2007) 439–478 MR

[17] **D P Thurston**, *Positive basis for surface skein algebras*, Proc. Natl. Acad. Sci. USA 111 (2014) 9725–9732 MR

[18] **J G van der Corput**, *Über Systeme von linear-homogenen Gleichungen und Ungleichungen*, Proc. Akad. Wet. Amsterdam 34 (1931) 368–371

Department of Mathematics, University of Iowa
Iowa City, IA, United States

`nel.abdiel@gmail.com`, `charles-frohman@uiowa.edu`

Received: 11 January 2015 Revised: 11 May 2017

Algebraic & Geometric Topology 17 (2017) 3375–3397

Generalized augmented alternating links and hyperbolic volumes

COLIN ADAMS

Augmented alternating links are links obtained by adding trivial components that bound twice-punctured disks to nonsplit reduced non-2–braid prime alternating projections. These links are known to be hyperbolic. Here, we extend to show that generalized augmented alternating links, which allow for new trivial components that bound n–punctured disks, are also hyperbolic. As an application we consider generalized belted sums of links and compute their volumes.

57M50; 57M25

1 Introduction

By the work of W Thurston [15], a nonsplit link in S^3 is known to be either hyperbolic or to contain an essential torus or annulus in its complement. When the link is hyperbolic, its complement admits a hyperbolic metric that is uniquely determined and, hence, the hyperbolic volume of its complement becomes an invariant that can be used to distinguish it from other links.

Menasco [9] proved that prime alternating non-2–braid links are hyperbolic. In Adams [2], it was further proven that augmented alternating links are hyperbolic. These links are obtained from a prime non-2–braid alternating link projection by adding trivial "vertical" components perpendicular to the projection plane that bound a disk punctured twice by the alternating link. These augmented alternating links have proved useful in a variety of contexts. In particular, they appear as the geometric limits of alternating links that correspond to twisting the two strands around which the augmenting components wrap. As such, together with the alternating links, they form the closure of the collection of alternating links in the geometric topology (see Corollaries 2 and 3 of Lackenby [8]). In particular, the volumes of the links in such a sequence must approach the volume of the augmented link from below.

If a link is not alternating, one can augment it at a subset of the crossings that would need to be changed to make it alternating, and then the result has the same complement as an augmented alternating link, since a full twist along one of the twice-punctured disks can reverse that crossing.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3375

Augmented alternating links are useful in a variety of settings. See for instance Blair, Futer and Tomova [4], Dasbach and Tsvietkova [5], Futer, Kalfagianni and Purcell [6], Lackenby [8] and Purcell [10; 13]. In some papers, the links considered are fully augmented. That is to say, every crossing in the original knot is in a twist sequence around which a vertical component has been added. In that case, one can use Andreev's theorem to prove hyperbolicity (see [13]).

In this paper, we extend the results of [2] to allow the vertical components to bound disks that are punctured more than twice by the alternating link in the projection plane. These new links are called generalized augmented alternating links. A precise definition appears in Section 2. Our main theorem is to prove that, indeed, their complements are always hyperbolic.

In particular, since $(1,q)$–Dehn filling of a vertical component corresponds to adding q full twists to the strands of the original link, Thurston's hyperbolic Dehn surgery theorem and the fact the augmented link is hyperbolic imply that the resulting links are always hyperbolic for high enough values of q.

Note that in several papers, authors have considered generalized augmented links that were also obtained by adding vertical components to a projection, but in this case, a not necessarily alternating projection such that the projection breaks up into generalized twist regions as in Figure 1. But here again, each twist region has to receive a crossing circle. These links have a variety of interesting properties, as discussed in Futer, Kalfagianni and Purcell [6] and Purcell [11; 12; 13].

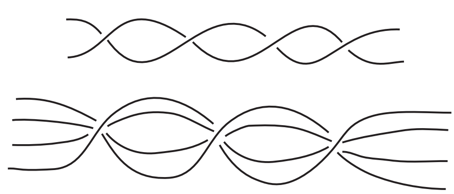


Figure 1: A traditional twist region and a generalized twist region

The presence of twice-punctured disks in link complements and the fact that twice-punctured disks are totally geodesic with a unique hyperbolic structure (see Adams [1]) implies that one can take belted sums of the links, which is the operation depicted in Figure 2, top row. The resulting link $L_1 \#_b L_2$ has volume equal to the sum of the volumes of the two links L_1 and L_2 that are summed. Note that a belt is any component bounding a twice-punctured disk in a hyperbolic link complement.

In Section 3, we generalize the notion of belted sum, and find an explicit formula between the volumes of the two links and their summand. Specifically, we show

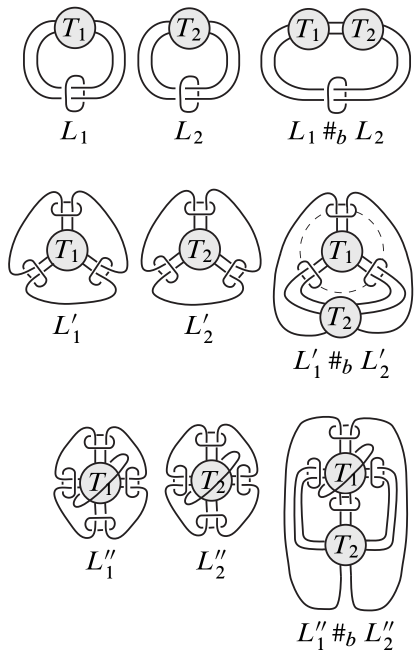


Figure 2: Generalized belted sums of links

that if a hyperbolic link denoted by $L_1' \#_b L_2'$ is constructed from two links L_1' and L_2' as in Figure 2, center, then $\mathrm{vol}(L_1' \#_b L_2') = \mathrm{vol}(L_1') + \mathrm{vol}(L_2') - 4(3.6638\dots)$. Similarly, if $L_1'' \#_b L_2''$ is a link constructed as in Figure 2, bottom, then $\mathrm{vol}(L_1'' \#_b L_2'') = \mathrm{vol}(L_1'') + \mathrm{vol}(L_2'') - 8(3.6638\dots)$. In this case, there are two distinct options for the central belt, wrapping either lower left to upper right or lower right to upper left.

More generally, let L_1 and L_2 be two links, each with a $2n$–string tangle at center, with n belts around adjacent pairs of the exiting strings and $n-3$ belts around the central tangle in the same pattern, as for instance appears in Figure 3 in the case $n=5$. Then $\mathrm{vol}(L_1 \#_b L_2) = \mathrm{vol}(L_1) + \mathrm{vol}l(L_2) - 4(n-2)3.6638\dots$. This construction

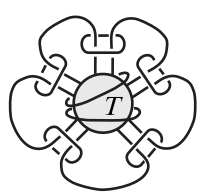


Figure 3: A potential generalized belted sum factor link

answers a question asked by Oliver Dasbach about the behavior of the volumes in Figure 2, center row, and was motivated by that question.

Of course, given a specific link, we would like to add vertical components to obtain either a link that can be decomposed via belted sum into simpler links or composed via belted sum into more complicated links. But we need to know that the resulting link is hyperbolic. This is what the main theorem provides when the initial link is alternating.

In the case of augmented alternating links, twisting a half-twist on the twice-punctured disk bounded by a vertical component, which adds or subtracts a crossing, will yield a new link complement that is hyperbolic with the same volume as the original link. This follows because the links are hyperbolic and the twice-punctured disks are totally geodesic with a unique hyperbolic structure (see Adams [1]). However, in the case of generalized augmented alternating links, if we twist a half-twist on an n–punctured disk bounded by a vertical component for $n \geq 3$, the result need not be hyperbolic, and even if it is, the volume is generally not preserved. As an example, adding one vertical component bounding a thrice-punctured disk in the figure-eight knot complement and then twisting a half-twist yields a Seifert fibered space. Further applications of generalized augmented alternating links to volume bounds for links appear in Adams [3].

2 Hyperbolicity

Given an alternating link J in a reduced alternating projection P, we will often consider it as a 4–regular graph on the projection sphere. That graph cuts the sphere up into complementary regions.

Let J be a prime nonsplit non-2–braid alternating link. Let P be a reduced alternating projection of J. Note that by results of Menasco [9], the projection is connected and there are no simple closed curves in the plane that intersect the projection transversely twice such that there are crossings to either side of the curve. Choose two complementary regions in the projection plane that do not share an edge. Take a trivial component C that intersects the projection sphere in precisely one point in each of the complementary regions. Then we say that $J' = J \cup C$ is a *generalized singly augmented alternating link*. We call the additional component a *vertical component*.

In the projection plane, we keep track of the vertical component as a gray arc γ. While fixing endpoints, we can isotope γ to minimize the number of intersections with the link J. This corresponds to an isotopy of the vertical component in the complement of J. We will assume that such an isotopy has already taken place, and call the corresponding vertical component a *minimal representative* of the isotopy class.

For any other pair of nonadjacent complementary regions, we allow the introduction of additional vertical components, as long as there are minimal representations of all the individual vertical components that are disjoint as arcs in the plane. We call the resulting link a *generalized augmented alternating link*. Note that for any pair of nonadjacent regions there can be at most one corresponding vertical component. There can be quite a few vertical components, as for instance occurs in the generalized augmented figure-eight knot in Figure 4. This link is maximally augmented in the sense that every pair of possible nonadjacent regions corresponds to a vertical component. Note that for the figure-eight knot, there is more than one option for the maximally augmented link that results.

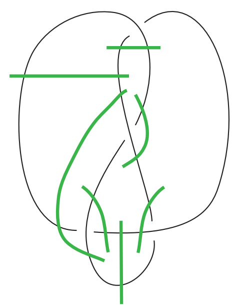


Figure 4: A generalized augmented projection of the figure-eight knot with the maximum possible number of vertical components

It should be noted that there are reduced alternating projections such that not all of the possible vertical components can be added since minimal representations overlap, as in Figure 5.

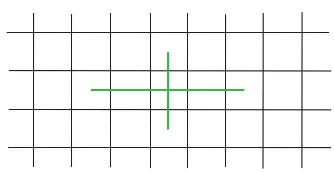


Figure 5: These two vertical components on this alternating grid cannot be made to avoid intersecting while in minimal representations.

Theorem 2.1 *Let J be a prime non-2–braid nonsplit alternating link. Then any generalized augmented alternating link J' constructed from a reduced alternating projection of J is hyperbolic.*

In fact, even in the case of a 2–braid link in a reduced alternating projection, if we add a vertical component that does not correspond to the axis around which the knot is braided, the result is hyperbolic. However, we will not include that case here.

It was proved in [9] that J is hyperbolic. Here, we are proving that the addition of these new vertical component preserves hyperbolicity. We use the machinery developed in [9], as described below.

We consider the projection plane P as a subspace of the projection sphere obtained by the 1–point compactification. We will move from the plane to the sphere without comment.

Let L be a link in a projection. At each crossing of the projection, a bubble is inserted, with the overstrand going over the top of the bubble and the understrand going under the bubble. We let S_+ be the sphere obtained by replacing each equatorial disk of a bubble in the projection sphere with the top hemisphere of the corresponding bubble. We define B_+ to be the ball bounded by S_+. Similarly, we define S_- and B_-, using the bottom hemispheres. In Lemma 1 of [9], Menasco showed that for any projection, and any closed surface that is incompressible and not boundary parallel, the surface can be isotoped so that it intersects bubbles in saddles and each intersection curve with $S_\pm$ intersects each bubble at most once and intersects at least one bubble. We will use a version of this that applies to surfaces with boundary, as appears below. We begin with the following lemma.

Lemma 2.2 *If C is a vertical component in a generalized augmented alternating link J' such that C is a minimal representative of its isotopy class, then the vertical punctured disk D that is bounded by C is incompressible.*

Proof We can treat D as a surface in either $S^3-(J\cup C)$ or S^3-J'. If there are no compression disks for D in $S^3-(J\cup C)$, then there are certainly no compression disks in S^3-J'. Hence we can drop all the vertical components except for C, since incompressibility of D in $S^3-(J\cup C)$ implies incompressibility in S^3-J'. Let γ denote the arc in the projection plane that is the projection of C.

Suppose that D compresses. Let E be a compressing disk, in general position with respect to the projection of J, chosen to have the minimum possible number of saddles for such a compressing disk. We will show, as Menasco did with closed and punctured surfaces in alternating knot and link complements, that the disk E behaves appropriately with respect to the bubbles inserted at each crossing.

As in [9], we isotope E to be transverse to a vertical edge in each bubble, and then push E out from that edge so that E only intersects each bubble in a possibly empty collection of saddle disks. We assume that E has the minimum number of saddles for a compressing disk of D. If the intersection of E with either B_+ or B_-, say B_+ for the argument, contains a subsurface W of E' other than a disk, then there is a nontrivial curve in E' (take an innermost boundary component of W) that bounds a disk in B_+. We can perform surgery on E using this disk to lower the number of saddles in E, contradicting our assumption that E has a minimal number of saddles. Thus, we can assume that the intersections of E with the bubbles and spheres S_+ and S_- decomposes E into saddle disks corresponding to where it intersects bubbles, and over-disks which have interior above S_+ and boundary on S_+ and under-disks, which have interior below S_- and boundary on S_-. We call the resultant graph on E, where saddles are treated as vertices, the *intersection graph*.

We can immediately eliminate any simple closed curves that do not intersect a bubble. Such a curve must bound a disk in either B_+ or B_- and again we can perform surgery on E to eliminate all such curves.

As in Lemma 1 of [9], we show that no intersection arc intersects a bubble more than once. For, if an intersection curve hits both sides of a bubble, we can choose an innermost such intersection curve α on that bubble, which must hit both sides of a single saddle. By taking an arc crossing the saddle and an arc on the disk bounded by the intersection curve, we obtain a closed curve δ on E bounding a disk E' on E, whose boundary can be isotoped to a meridian of J as in Figure 6, left. But this creates a once-punctured sphere in the complement of J, a contradiction.

If an intersection curve hits the same side of a bubble twice, we can again take an innermost such curve α and then perform an isotopy to eliminate the two saddles that touch the curve in this bubble as in Figure 6, right, pulling the blue band through the bubble, again contradicting the minimality of the number of saddles in E. Thus, no intersection curve hits a bubble more than once.

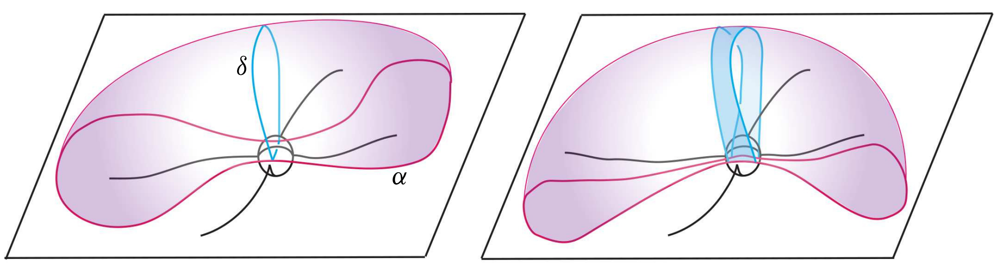


Figure 6: Intersection curves intersecting a bubble more than once

Suppose now that there is a simple closed intersection curve α on E. By this we mean a simple closed curve in the intersection graph that avoids ∂E and that forms a component of the boundary of a region containing no other intersection curves on E. Then it is also a simple closed curve on either S_+ or S_-; for convenience, assume S_+.

Note that all intersection arcs that begin and end on ∂E must begin and end on the same side of γ in the projection plane since the boundary of E must occur on only one of the two sides of D. Since a simple closed intersection curve does not intersect γ, in S_+ all the intersection arcs lie to the side of α containing γ. Take an innermost intersection curve α' to the other side of α. If there are none, take α itself considered innermost to the outside. Since the projection is alternating, each intersection curve must intersect bubbles such that the overstrand of the bubble is alternately on the left and right of the curve. As in [9], this forces α' to hit a bubble twice, a contradiction. Thus, any such simple closed curve must avoid all bubbles. But then we could replace the disk it bounds on E with the disk it bounds on S_+, and push off to lower the number of intersection curves.

Hence, there are no simple closed intersection curves on E. All intersection curves are arcs that begin and end on ∂E. Note that when viewed in the plane, no such intersection arc crosses γ but all intersection arcs start and end on γ, all coming out one side of it. See Figure 7 for the picture.

A *fork* of the intersection graph is a vertex with at least three edges ending on ∂E, keeping in mind that all interior vertices are 4–valent. We show that every intersection graph has at least one fork. Since every complementary region must intersect ∂E in its boundary, the graph obtained by throwing away all edges with an endpoint on the boundary of E is a collection of trees. Every tree of two or more vertices always has at least two leaves, and those leaves will have three edges that must all end on the boundary. So in this case, there are at least two forks. The one exception is if there is only one vertex to one tree, which coincides with the case of there being only one saddle in E. However, then we still have a fork.

We consider what a fork tells us about the projection P. See Figure 7. The two disks bounded by the one saddle and the three curves ending on the boundary of E cause there to be exactly two arcs of the knot that come out of the crossing in question and then pass through E as punctures, without crossing any other arcs of the knot in between. By Theorem 1(b) of [9], which shows that an alternating knot is prime if and only if it is obviously so in any alternating projection, these arcs cannot contain any crossings between when they puncture E and when they pass through the crossing in question.

Fixing its endpoints, we can isotope γ past the resultant crossing. There is a corresponding isotopy of C, D and E, changing the pattern of intersections on E and eliminating at least one saddle on E. We repeat this process until either all saddles are eliminated from E, leaving only simple arcs of intersection, or there is only one saddle remaining. In the first case, taking an outermost arc α'' on E, we cut off a disk E''

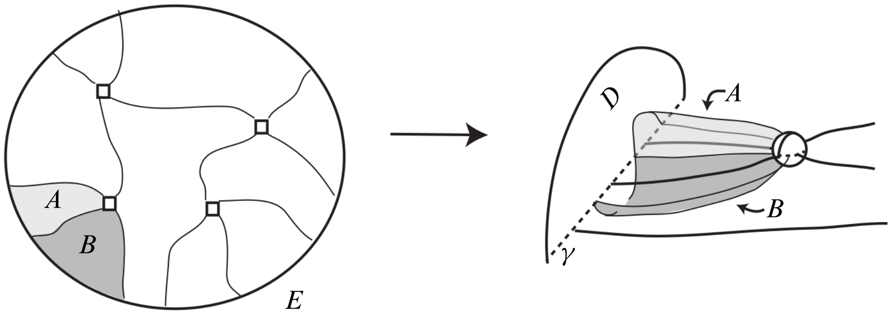


Figure 7: A fork in the compression disk allows us to isotope γ past a crossing.

which intersects the projection plane in an arc that does not intersect a bubble. Hence we can either isotope an arc on γ to this arc, and lower the number of intersections in $\gamma \cap J$, a contradiction to γ corresponding to a minimal representative of C, or if no part of J lies in the region in the projection plane cut off by $\gamma \cup \alpha''$, we can isotope E to lower the number of intersection arcs. In either case, repeating the process if necessary, we obtain a contradiction.

In the second case, if only one saddle remains, then, as in Figure 8, the arc γ can be isotoped with endpoints fixed to lower its number of punctures, and it is therefore not a minimal representative. □

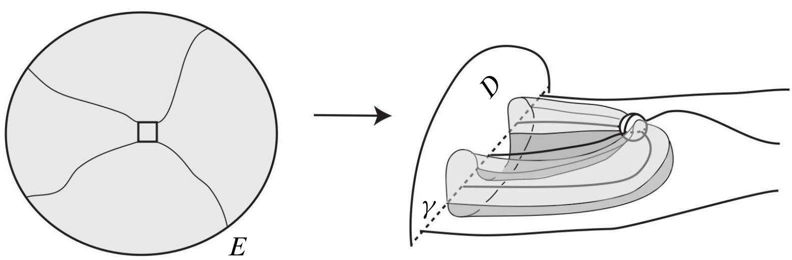


Figure 8: A single saddle in the compression disk implies γ is not a minimal representative.

The following results will put us in position to prove Theorem 2.1. Throughout the rest of this section, J is a prime non-2–braid nonsplit alternating link, and J' is a generalized augmented alternating link complement obtained from J.

Lemma 2.3 *The link complement $S^3 - J'$ is irreducible.*

Proof Proving irreducibility is equivalent to showing that $S^3 - J'$ is not splittable. Since the alternating projection of J is connected, Theorem 1(a) from [9] shows that

J is nonsplittable. Hence, if J' is splittable, there must be a sphere with J to one side and at least one vertical component C to the other side. Let D be the vertical punctured disk bounded by C. Discarding the other vertical components, we work with just this one vertical component. But if C is contained in a sphere, it bounds a disk in the sphere. We can use this disk to obtain a compression disk for D, contradicting Lemma 2.2. Hence, J' is nonsplittable. □

Lemma 2.4 *Any essential torus T in $S^3 - J'$ is meridionally compressible.*

Proof By definition, T is incompressible and not boundary-parallel. To prove meridional compressibility, we must show there is a nontrivial simple closed curve on T that bounds a disk punctured once by the link J'. So suppose T is meridionally incompressible.

We again apply the techniques of [9], which Menasco utilized to prove a similar result for alternating links. First, we flatten each vertical component into the projection plane as in Figure 9.

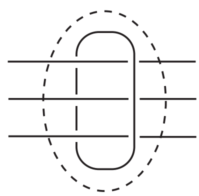


Figure 9: Projecting the vertical components

Since T is incompressible and meridionally incompressible, Lemma 1 of [9] tells us there exists a realization of T such that the intersection curves with S_+ and S_- do not intersect the same crossing bubble more than once. Moreover, every curve must intersect at least one bubble.

We now eliminate the vertical components, without isotoping the surface T. Saddles that appeared in crossing bubbles involving the vertical components disappear and the intersection curves that entered a dotted region as in Figure 9 now connect to one another.

There are two fundamental changes in the system of intersection curves. First of all, each intersection curve need no longer bound a disk above S_+ in the case of B_+ and below S_- in the case of B_-. Instead, a collection of intersection curves can bound a subsurface of T above S_+ or below S_-. Second, for each of the resulting intersection curves on S_+ and S_-, it can either be the case that the curve does or does not intersect a bubble more than once.

We add the vertical components back in, but now they are once again vertical, perpendicular to the projection plane, each puncturing the projection plane in two points. For convenience, we consider intersection curves on S_+, but everything works just as well for intersection curves on S_-. We assume that T is chosen to minimize the resultant number of saddles.

The collection of intersection curves and saddles decompose T into squares on its surface corresponding to the saddles, and the components of intersections with B_+ and B_-. We first show that, with the possible exception of a single n–punctured torus, all of these components are either disks or annuli.

Let R be a planar component of $T \cap B_+$. Suppose an intersection curve α that forms one of the boundaries of R on S_+ bounds a disk F on T. We show that then R is a disk. Let D' be the disk bounded by α on S_+. It may or may not contain additional intersection curves. If there are no vertical components with endpoints in D', then we can isotope the disk F bounded by α on T to D', pushing any other parts of T out of the way in the process. After this isotopy, we have either eliminated α as an intersection curve, simplifying the intersection graph, or R was a disk in $T \cap S_+$. If D' is punctured by some vertical component once, then that vertical component will be nontrivially linked with α, contradicting the fact that α bounds F. Hence, any vertical component C that intersects D' must do so with both of its endpoints. No other vertical components can be linked with this one above the projection plane, as in Figure 10, since if they were, they would also have to be similarly linked below the projection plane, which the existence of F prevents.

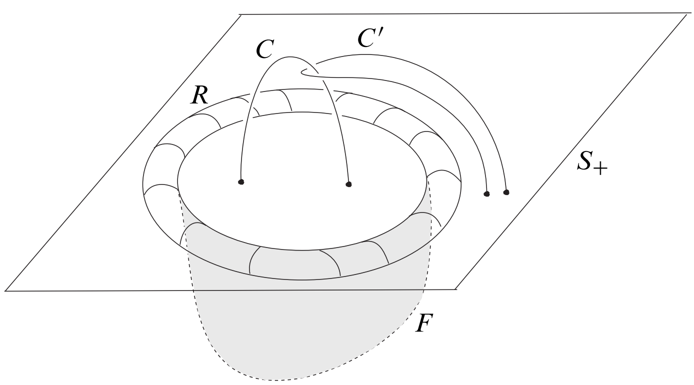


Figure 10: Vertical components cannot be linked above the projection plane when an intersection curve bounds a disk on T.

We can then take a disk D'' that is horizontal and has boundary in R that is a curve parallel to α and slightly above it. Then we can isotope F to D', eliminating the intersection curve α and any saddles that it touches, a contradiction to minimality. Note we are using the fact J' is nonsplittable here. So, the only planar components of $T \cap S_+$ are disks and annuli, with all boundary components of the annuli appearing as parallel nontrivial curves on the torus, as in Figure 11, right.

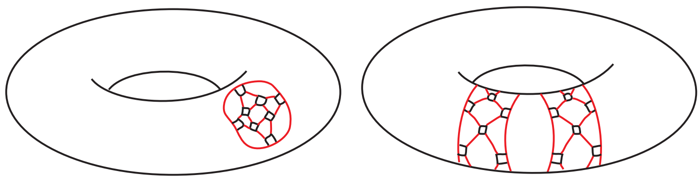


Figure 11: Islands as on the left can only occur on T if there is a component in $T \cap B_\pm$ that is an n–punctured torus. Annular regions appear as on the right.

We first show that there are no innermost curves on S_+ that bound disks in B_+. An innermost curve is one that bounds a disk R in S_+ that contains no other intersection curves. A bubble that intersects an innermost curve is called an inner (outer) bubble if its overstrand lies inside (outside) of R.

Such an innermost curve α bounding a disk G in $T \cap B_+$ must intersect bubbles more than once, as J is alternating, so the bubbles must alternate between bubbles with their overstrand to the right and bubbles with their overstrand to the left as we travel around the curve. Since there are no other curves inside α, the other side of each inner bubble must be hit by α, as in Figure 12.

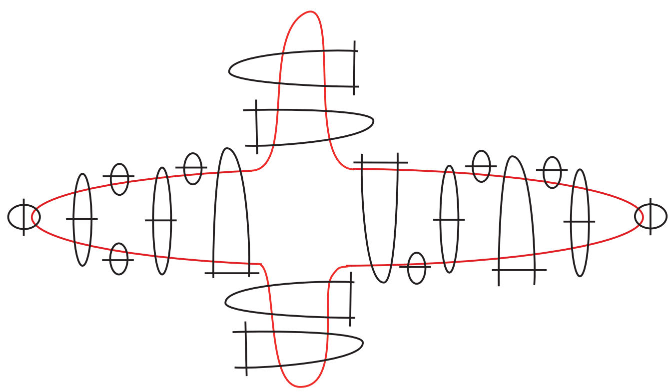


Figure 12: Inner bubbles must be intersected on both sides by α.

There are three types of bubble intersections with α. A bubble of type I intersects α on both sides of the bubble and the overstrand lies to the inside of α. A type II bubble is one that is intersected by α at least twice on one side and the overstrand lies to the outside of α. A bubble of type III intersects α once and has its overstrand to the outside of α. If a bubble is of type II or III, we consider only multiple intersections of that bubble occurring to the inside of the curve. We do not care if distinct bubbles appearing inside α are in fact the same bubble when also considered outside α.

Choosing an innermost pair of intersections of α with a bubble of type I, one on each side of the bubble, we can form a loop δ out of an arc on the corresponding saddle and an arc on G that forms the boundary for a meridional compression to the overstrand of the bubble. Since there are no meridional compressions, it must be the case that one or more vertical components C block this meridional compression, as in Figure 13, left.

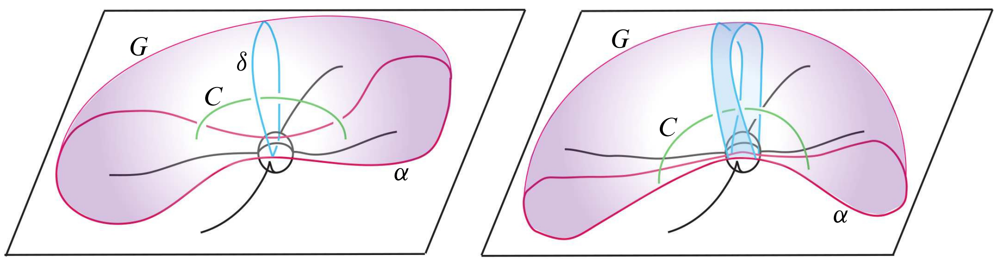


Figure 13: A vertical component blocking meridional compressions and saddle-reducing isotopies

Similarly, for a bubble of type II, we can isotope a band on the disk bounded by α to eliminate two saddles unless the isotopy is blocked by one or more vertical components C, as in Figure 13, right.

Relative to the vertical components, the intersection curve α can wind around the curves, as in Figure 14, so when α is drawn uncomplicated, it could be the case that the vertical components are tangled with one another, as if in a plat.

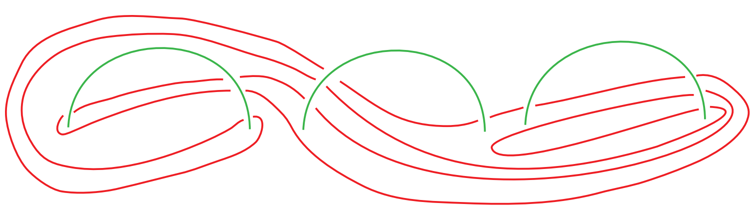


Figure 14: The intersection curve α can wind around the vertical components.

We note the following *pairing property*: Consider all intersection curves on S_+ that intersect a vertical disk D bounded by vertical component C, which is not necessarily in a minimal representation. In B_+, they must form the boundaries of surfaces in $T \cap B_+$. Then $T \cap B_+ \cap D$ is a collection of arcs that pair the points in $\alpha \cap D$, some potentially nested. However, there must be an innermost arc such that it cuts a disk from $D \cap B_+$ that contains no other such arcs. Hence two adjacent intersection curves are connected by an innermost arc on D. We consider those two curves an *innermost pair*.

We now utilize the pairing property. Suppose C is a vertical component bounding a vertical punctured disk D that blocks one or more meridional compressions or saddle-reducing isotopies caused by bubbles to the inside of the innermost curve α bounding the disk G in $T \cap S_+$.

Choose any vertical component C with endpoints in α that blocks crossings from generating either a meridional compression or an isotopy lowering the number of saddles. Regions between bubbles inside α fall into six types, denoted by H, K, L, M, N and P in Figure 15. A region of type H has two bubbles on its boundary, one of type I and one of type III. A region of type K has two bubbles of type I and two bubbles of type III, appearing alternately around its boundary. A region of type L has a bubble of type I, a bubble of type II and a bubble of type III around its boundary. A region of type M has only one bubble of type II on its boundary. A region of type N has two bubbles of type II on its boundary. A region of type P, which we call a *juncture*, has three or more type II bubbles on its boundary (four in the case of Figure 15).

Notice that regions H and M must each have at least one vertical component that has exactly one endpoint in the region, since only two adjacent complementary regions of the projection intersect a region H and only one complementary region of the projection intersects a region of type M. Isotope C so that the arc that represents its projection follows α as closely as possible, as in Figure 15, and let D be the punctured disk that it bounds. Note that we may have to carry along other vertical components with which it is entangled.

Shifting from the intersection curves on S_+ to the corresponding intersection curves on S_-. we find that each pair of adjacent intersection curves passing through D share a bubble. See Figure 16. In every such diagram, each of the six types of regions will have exactly the same pattern of intersection curves shown here, with only a juncture region that could look different, depending on the number of type II crossing around its boundary. If we had left our arc representing the vertical component going straight across a juncture, there would have been a pair of adjacent curves passing under the vertical component that would not have shared a bubble. But, instead, we have C follow α around the outside. This ensures that all pairs of adjacent curves passing through D share a bubble.

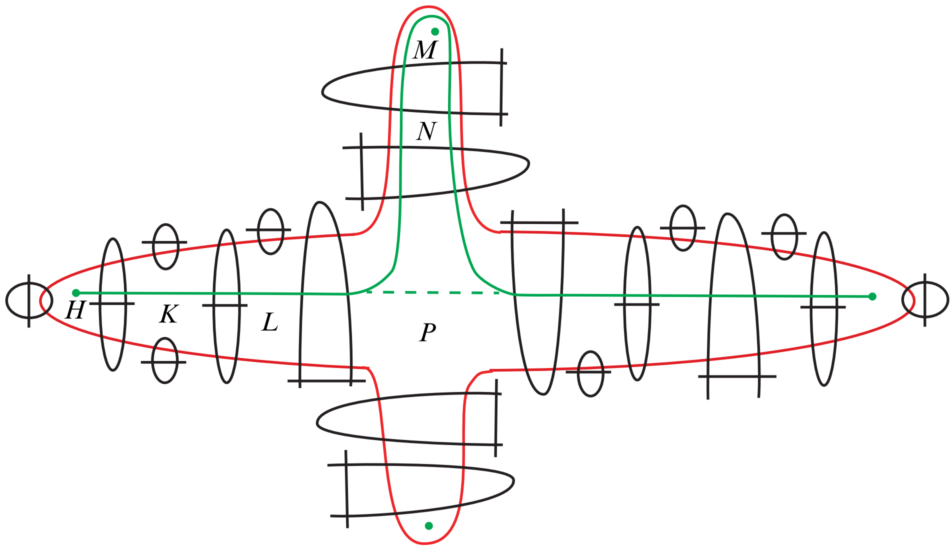


Figure 15: A vertical component blocking a meridional compression

So there exists an innermost pair of adjacent curves passing through D. They must share a bubble. That bubble cannot be blocked by a vertical component because the disk G prevents any such vertical component. Thus, we obtain either a meridional compression if

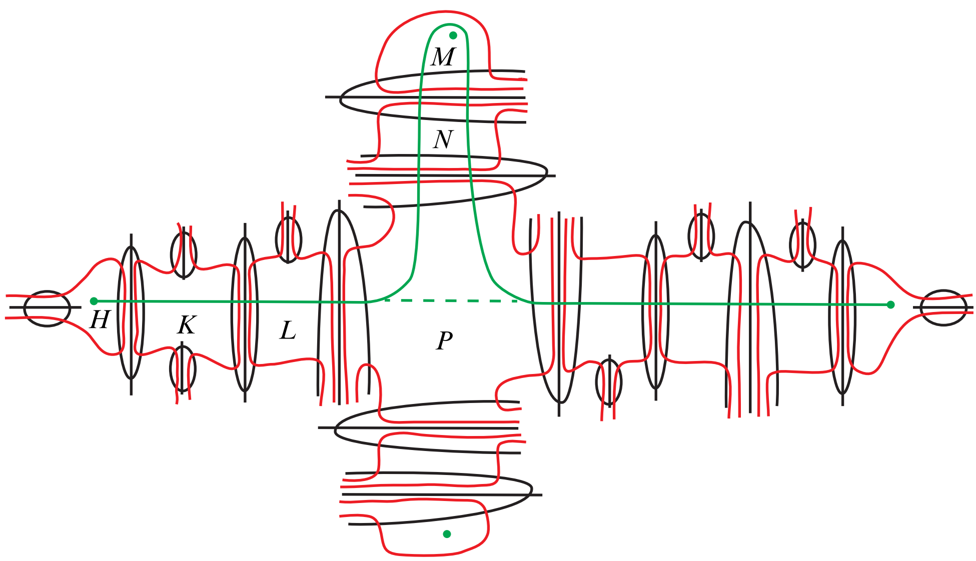


Figure 16: The view from S_-, where all adjacent curves passing under the vertical component share a bubble.

the bubble is a type I or type III bubble, or a saddle reducing isotopy if the bubble is a type II bubble. Hence, there can be no innermost intersection curve bounding a disk in B_+.

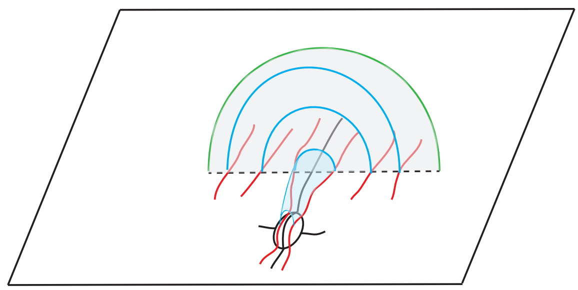


Figure 17: Two intersection curves share both a bubble and a surface, yielding a meridional compression.

We now prove there are no annular components in $T \cap B_+$. Suppose there were such an annulus A. Note that the existence of A precludes the possibility of a punctured torus component in either $T \cap B_+$ or $T \cap B_-$. Each of its boundary components bounds a disjoint disk in S_+ and those disks, together with A bound a ball in B_+. Choose an annulus that is innermost in the sense that its ball contains no other ball corresponding to such an annulus. Note that T then bounds a solid torus V to the ball side of A. Because T is incompressible and meridionally incompressible, there must be a set $\mathcal{C}$ of at least two vertical components that together prevent A from being compressible or meridionally compressible in V. One possibility is that one or more of these components intersect the projection plane in the distinct disks bounded by ∂A. But it could also be the case that two or more vertical components are linked together above the projection plane inside the ball bounded by A. In this case, the same vertical components must be similarly linked beneath the projection plane, since ultimately these vertical components form an unlink when considered as a whole. In either case, the solid torus V must be unknotted, as $\mathcal{C}$ intersects every meridional disk in V and, if V were knotted, this would make $\mathcal{C}$ a nontrivial link by itself, when all other components are dropped, which is a contradiction to how we constructed J'.

Again, in either case, there is a compression disk for A that intersects only vertical components. Dropping all vertical components momentarily, that compression yields a sphere and, since J is nonsplittable, it must lie to one or the other side of that sphere. Hence, J lies to one or the other side of T. If J is in V, then to avoid compressions and meridional compressions, there must be vertical components to the other side. However, then again, we find that by dropping J temporarily, the collection of vertical components forms a nontrivial link, a contradiction.

Hence, it must be the case that J lies to the outside of V. Let A' be any other annulus in $T \cap B_+$. Then its ball must also intersect $\mathcal{C}$, since $\mathcal{C}$ is the only collection of vertical components that together wrap all the way around V. Hence the curves $\partial A'$ must be parallel to the curves in ∂A. This implies that there are no disk components in $T \cap B_+$, as, if there were such, there would need to be an annular component to each side of its boundary on S_+ to not be innermost, which the existence of $\mathcal{C}$ prevents.

Hence, we have only annular components remaining in either $T \cap B_+$ or $T \cap B_-$. However, as in Figure 11, right, if any intersection curve bounding an annulus intersects a bubble, there must be disk components, since each vertex is 4–valent, so an annular boundary hitting a saddle cannot also be an annular boundary of a second annulus. Hence, all intersection curves avoid bubbles. But then the boundaries of the annuli separate the projection of J. We could isotope any annuli away that are not parallel to A, so it must be the case that all annuli are parallel to A. Similarly in B_-, all annuli must be parallel. Hence, there can be only one annulus to either side for T to be connected. Since J is to the outside of V, each disjoint disk on S_+ bounded by ∂A lies in a single complementary region of J. But, since the endpoints of the vertical components in $\mathcal{C}$ lie in these two disks, each such component either has both of its endpoints in the same complementary region or its two endpoints share the same complementary regions as another vertical component. In either case, this contradicts our construction of J'.

The last case to consider is when there is a component of $T \cap B_+$ that is an n–punctured torus. In this case, all other components of $T \cap B_+$ and $T \cap B_-$ must be disks. However, then all components to the B_- side are disks, a possibility we have eliminated.

Thus, we have shown that an essential torus T is meridionally compressible. □

Lemma 2.5 *The link J' is prime.*

Proof That is to say, we show that there are no essential annuli with both boundaries appearing as meridional curves for link components.

We first suppose that the boundary components of A are meridians on $\partial N(J)$. So we think of A as a twice-punctured sphere. Then A demonstrates that J' is a composite link. Since J' is not composite, it must be that the addition of the vertical components prevents A from being boundary-parallel. We assume that A is not meridionally compressible by doing any compressions first and taking only one of the resultant annuli to consider. As we did with T, we use the results of [9] to put the punctured sphere in standard position relative to S_+ and S_- so that no intersection curve intersects a bubble more than once and every curve either has a puncture or a bubble on it. We assume that A has been chosen to minimize the number of saddles in crossing bubbles only involving J. We again throw away the vertical components and obtain intersection curves that no longer need to bound disks in S_+ and S_-, and that can intersect a given bubble more than once. As in $(*)$ preceding the proof of Theorem 2 in [9], since J is alternating, a curve that crosses a bubble with its overstrand to one side of the curve must pass through an odd number of punctures (there are only two total) if it subsequently passes through another bubble with its overstrand to the same side.

The same proof we used in the case of an essential meridionally incompressible torus shows that no components of $A \cap B_\pm$ can be other than disks and also annuli with nontrivial boundaries on A. In fact, no component of $A \cap B_\pm$ can be such an annulus, as any compression disk for such an annulus would intersect only vertical components, but every compression disk for an essential annulus in A must intersect J, since J punctures the sphere corresponding to A twice.

We could have up to four disks in $A \cap B_\pm$ with punctures on their boundaries, which, in addition to the types of regions depicted in Figure 15, also allow for regions Q, R and S, depicted in Figure 18, where the intersection curve (in red) crosses the link at a puncture, and a shaded disk represents a tangle. But again, the argument given previously applies to show that there are no disks innermost on $S_\pm$ in $A \cap B_\pm$ that intersect bubbles, either with or without punctures on their boundary.

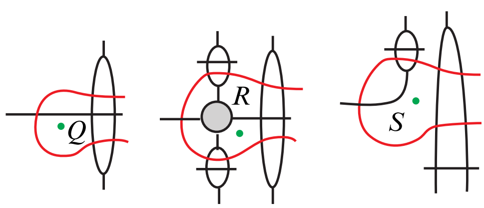


Figure 18: Additional possible regions for an innermost disk with a puncture on its boundary

But then A intersects S_+ and S_- in the same pair of arcs from the first puncture to the second puncture, each of which does not intersect the projection of J. So, to one side, the projection of J is a trivial arc. There can be no vertical components to this side of A since there are only two adjacent regions of the projection plane to this side. This shows that A cannot be an essential annulus with both boundaries meridional on $N(J)$.

We now consider an essential annulus A with both boundary components meridians on boundaries of neighborhoods of the vertical components. It is appropriate to consider A as a twice-punctured sphere. But then both punctures must be from one vertical component C, and J must lie to one side. To the other side is a trivial arc of C

and additional vertical components that prevent A from being boundary-parallel. But individual vertical components inside this twice-punctured sphere bound disks there, contradicting the fact that the n–punctured disks they bound are incompressible.

Note that there can be no annulus with one boundary a meridian on J and another boundary a meridian on a vertical component, as a sphere in S^3 cannot be punctured once by a simple closed curve. Thus, there can be no essential annuli with meridional boundary components in $S^3 - J'$ and J' is prime. □

Corollary 2.6 *There are no essential tori in $S^3 - J'$.*

Proof If an essential torus T exists, Lemma 2.4 implies we can meridionally compress it to obtain a twice-punctured sphere. Since T is not boundary-parallel, this twice-punctured sphere shows that the link J' is composite, contradicting Lemma 2.5. □

Lemma 2.7 *There are no essential annuli in $S^3 - J'$.*

Proof Although we have eliminated essential annuli with both boundaries meridians, we now consider the possibility of other essential annuli. Lemma 1.16 of [7] implies that if there is an essential annulus at all, then $S^3 - J'$ is Seifert fibered, with the boundaries of the annulus as fibers. Moreover, there are either a total of one, two or three torus boundary components in $S^3 - N(J')$. If one boundary of A is a meridian on a vertical component C, and the other is not, then if we fill C in, the annulus becomes a compressing disk on the boundary of the resulting link. If the other boundary component lies on the boundary of a neighborhood of J, this contradicts the hyperbolicity of J. If the other boundary is on a boundary of a neighborhood of a different vertical component, we have that a nontrivial curve on its boundary is trivial in the complement of J. However, this can never occur for a trivial link component unless the curve is a longitude, but then we are contradicting the incompressibility of the vertical disk bounded by the component.

If a boundary component of A is a nonmeridian on a vertical component C, then we can fill in C and extend the Seifert fibration to the solid torus that we filled in, making the result Seifert fibered, which is a contradiction to the hyperbolicity of $S^3 - J$ unless there is a second vertical component. If there is a second vertical component and the other boundary of A is a nonmeridian upon it (as it must be by our previous considerations), then we can fill it in also, and we obtain a Seifert fibration for $S^3 - J$, a contradiction to its hyperbolicity.

The only possibility left is that the boundaries of A is a nonmeridian upon both C_1 and a component K of J, which is all of J, and there is a second vertical component C_2.

Then $S^3 - J'$ is a twice-punctured disk crossed with a circle obtained by taking a regular neighborhood of $A \cup \partial N(K)$. This is embedded in S^3 so that each boundary torus bounds a solid torus to the exterior, which is a neighborhood of the corresponding link component. But then the component corresponding to the outer boundary of the disk, which is one of the vertical components, links both of the other components, one of which is also vertical. However, then two vertical components are linked, a contradiction to the construction of J'.

We now consider an essential annulus A with both boundary components on $N(J)$ but not meridians. Then they must both be on the same component or else A would be essential in the complement of J, a contradiction. In this case, each boundary of A is a (p,q)–curve on the boundary of a neighborhood of the link component K, with $|q| \geq 1$. Hence, C is a (p,q)–cable of K. But then there is an essential annulus with one boundary on $\partial N(C)$ and a second boundary on $N(K)$, a possibility we have already eliminated. □

Proof of Theorem 2.1 By work of Thurston [15], in order to prove that $S^3 - J'$ is hyperbolic, it is enough to show that $S^3 - J'$ is irreducible, and to show there are no essential tori or annuli in $S^3 - J'$. This is the content of Lemma 2.3, Corollary 2.6 and Lemma 2.7. □

3 Generalized belted sums

In this section, we consider generalized belted sums as in Figure 2. We show that if $L_1 \#_b L_2$ is constructed from two links with a $2n$–string tangle at center, with n belts around adjacent pairs of the exiting strings and $n-3$ belts around the central tangle, no two of which are parallel, then $\mathrm{vol}(L_1 \#_b L_2) = \mathrm{vol}(L_1) + \mathrm{vol}(L_2) - 4(n-2)3.6638\ldots$. To see this, we utilize the thrice-punctured spheres that appear in the link complement. Thrice-punctured spheres are known to be totally geodesic with a rigid structure in a hyperbolic 3–manifold (see for instance [1]). In particular, any two are isometric. In the case of a $2n$–string tangle T_1 as in Figure 3, there is a collection of thrice-punctured spheres that shield the part of the manifold corresponding to the $2n$–string tangle from the rest of the manifold. Cutting the manifold open along this collection of thrice-punctured spheres and then for each resulting piece, doubling across the thrice-punctured spheres yields two link complements, one with the $2n$–string tangle appearing twice, once reflected, and the other an untwisted daisy chain with additional components, as appear in Figure 19. Further examples of the untwisted daisy chain with additional components appear in Figure 20. The two halves of the original manifold must have volume exactly half of these, since reflecting across totally geodesic surfaces (the thrice-punctured spheres; see [1]) doubles the volume. So the original manifold has volume exactly half the sum of these two volumes. The same is true for the link with $2n$–string tangle T_2. Now, when we take the two link complements, cut them both open along the collection of thrice-punctured spheres, and throw away the two halves of the untwisted daisy chain, we obtain the volume of the first link plus the volume of the second link minus the volume of the untwisted daisy chain with additional components.

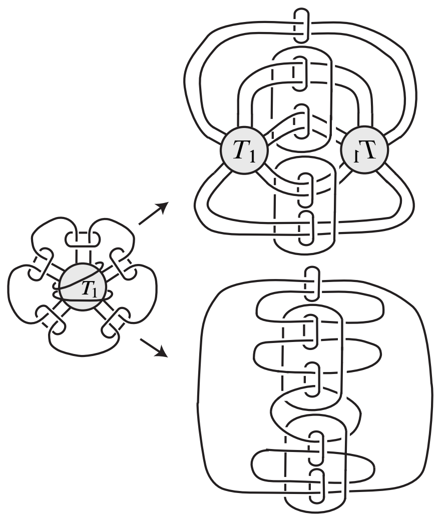


Figure 19: Cutting and doubling

In the case $n = 3$, the volume of the untwisted daisy chain is $4(3.6638\ldots)$, where $3.6638\ldots$ is the volume of an ideal regular octahedron. The manifold is commensurable with the Whitehead link. (See Example 6.8.7 of [14].) For $n > 3$, we can cut the link complement open along the twice-punctured disks bounded by components that are not in the untwisted daisy chain to obtain $n-2$ pieces, each of volume $\frac{1}{2}(4(3.6638\ldots))$, as in Figure 20. When we take the belted sum of the two links, we discard all of these pieces from both link complements, meaning we lose a volume of $4(n-2)3.6638\ldots$.

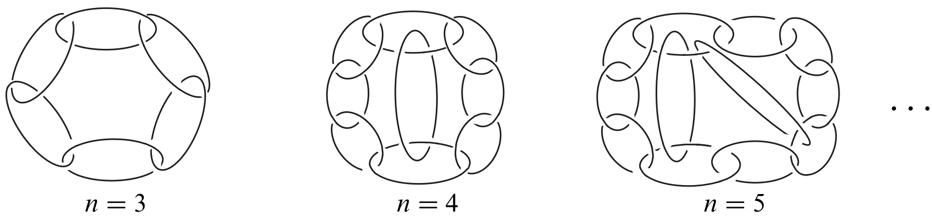


Figure 20: Links of volume $4(n-2)3.6638\ldots$

We can further start with any link and add components to decompose it into pieces, each of which has a volume we can determine as in the case of a generalized belted sum, to obtain the volume of the augmented link, which will bound the volume of the original link since hyperbolic Dehn filling always decreases volume. In the case of an alternating link, Theorem 2.1 tells us that the generalized augmented link that we produce will be hyperbolic, which we need to know for the procedure to apply. As an example, consider the link appearing in Figure 21. We denote the link obtained from a $2n$–tangle T_i by completing it as in Figure 3 by L_i.

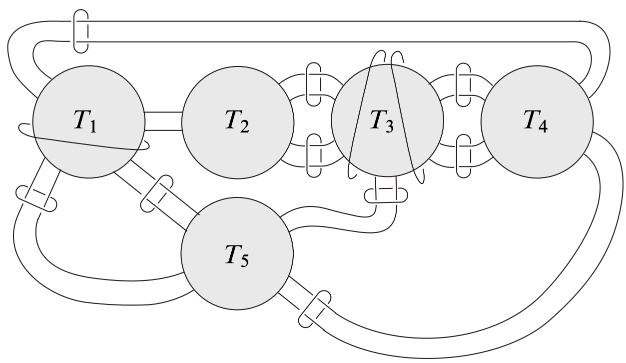


Figure 21: Finding the volume of this link

We can cut along the various thrice-punctured spheres, and realize each of the resulting pieces as a link of the appropriate type, where we have thrown away a volume $(n-2)3.6638\ldots$. In this case, we decompose the link complement into three 8–tangles, one 6–tangle and one 10–tangle. There is also another piece remaining which is a copy of the Borromean rings. Hence the volume is $\mathrm{vol}(L_1)+\mathrm{vol}(L_2)+\mathrm{vol}(L_3)+\mathrm{vol}(L_4)+\mathrm{vol}(L_5)-20(3.6638\ldots)+7.32772\ldots$.

References

[1] **C C Adams**, *Thrice-punctured spheres in hyperbolic 3–manifolds*, Trans. Amer. Math. Soc. 287 (1985) 645–656 MR

[2] **C C Adams**, *Augmented alternating link complements are hyperbolic*, from "Low-dimensional topology and Kleinian groups" (D B A Epstein, editor), London Math. Soc. Lecture Note Ser. 112, Cambridge Univ. Press (1986) 115–130 MR

[3] **C Adams**, *Bipyramids and bounds on volumes of hyperbolic links*, Topology Appl. 222 (2017) 100–114 MR

[4] **R Blair**, **D Futer**, **M Tomova**, *Essential surfaces in highly twisted link complements*, Algebr. Geom. Topol. 15 (2015) 1501–1523 MR

[5] **O Dasbach**, **A Tsvietkova**, *A refined upper bound for the hyperbolic volume of alternating links and the colored Jones polynomial*, Math. Res. Lett. 22 (2015) 1047–1060 MR

[6] **D Futer**, **E Kalfagianni**, **J S Purcell**, *On diagrammatic bounds of knot volumes and spectral invariants*, Geom. Dedicata 147 (2010) 115–130 MR

[7] **A Hatcher**, *Notes on basic 3–manifold topology*, book in progress (2007) Available at `http://www.math.cornell.edu/~hatcher/3M/3Mdownloads.html`

[8] **M Lackenby**, *The volume of hyperbolic alternating link complements*, Proc. London Math. Soc. 88 (2004) 204–224 MR

[9] **W Menasco**, *Closed incompressible surfaces in alternating knot and link complements*, Topology 23 (1984) 37–44 MR

[10] **J S Purcell**, *Volumes of highly twisted knots and links*, Algebr. Geom. Topol. 7 (2007) 93–108 MR

[11] **J S Purcell**, *Slope lengths and generalized augmented links*, Comm. Anal. Geom. 16 (2008) 883–905 MR

[12] **J S Purcell**, *Hyperbolic geometry of multiply twisted knots*, Comm. Anal. Geom. 18 (2010) 101–120 MR

[13] **J S Purcell**, *An introduction to fully augmented links*, from "Interactions between hyperbolic geometry, quantum topology and number theory" (A Champanerkar, O Dasbach, E Kalfagianni, I Kofman, W Neumann, N Stoltzfus, editors), Contemp. Math. 541, Amer. Math. Soc., Providence, RI (2011) 205–220 MR

[14] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at `http://msri.org/publications/books/gt3m`

[15] **W P Thurston**, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. 6 (1982) 357–381 MR

Mathematics and Statistics Department, Williams College
Williamstown, MA, United States

`cadams@williams.edu`

`https://sites.williams.edu/cadams/`

Received: 20 October 2015 Revised: 2 March 2017

Algebraic & Geometric Topology 17 (2017) 3399–3434

Representations of the Kauffman bracket skein algebra II: Punctured surfaces

FRANCIS BONAHON
HELEN WONG

In part I, we constructed invariants of irreducible finite-dimensional representations of the Kauffman bracket skein algebra of a surface. We introduce here an inverse construction, which to a set of possible invariants associates an irreducible representation that realizes these invariants. The current article is restricted to surfaces with at least one puncture, a condition that is lifted in subsequent work relying on this one. A step in the proof is of independent interest, and describes the arithmetic structure of the Thurston intersection form on the space of integer weight systems for a train track.

57M27, 57R56; 57M27

This article is a continuation of [9] and is part of the program described in Bonahon and Wong [6], devoted to the analysis and construction of finite-dimensional representations of the Kauffman bracket skein algebra of a surface.

Let S be an oriented surface of finite topological type without boundary. The *Kauffman bracket skein algebra* $\mathcal{S}^A(S)$ depends on a parameter $A = \mathrm{e}^{\pi \mathrm{i}\hbar} \in \mathbb{C} - \{0\}$, and is defined as follows: One first considers the vector space freely generated by all isotopy classes of framed links in the thickened surface $S \times [0,1]$, and then one takes the quotient of this space by two relations. The first and main relation is the *skein relation*, which states that

$$[K_1] = A^{-1}[K_0] + A[K_\infty]$$

whenever the three links K_1, K_0 and $K_\infty \subset S \times [0,1]$ differ only in a little ball where they are as represented in Figure 1. The second relation is the *trivial knot relation*, which asserts that

$$[K \cup O] = -(A^2 + A^{-2})[K]$$

whenever O is the boundary of a disk $D \subset K \times [0,1]$ disjoint from K, and is endowed with a framing transverse to D. The algebra multiplication is provided by the operation of superposition, for which the product $[K] \cdot [L]$ is represented by the union $[K' \cup L']$ where $K' \subset S \times \left[0, \frac{1}{2}\right]$ and $L' \subset S \times \left[\frac{1}{2}, 1\right]$ are respectively obtained by rescaling the framed links $K \subset S \times [0,1]$ and $L' \subset S \times [0,1]$ in the $[0,1]$ direction.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3399

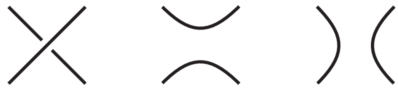


Figure 1: A Kauffman triple

Turaev [33], Bullock, Frohman and Kania-Bartoszyńska [14; 15] and Przytycki and Sikora [29] showed that the skein algebra $\mathcal{S}^A(S)$ provides a quantization of the *character variety*

$$\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) = \{\text{group homomorphisms } r\colon \pi_1(S) \to \mathrm{SL}_2(\mathbb{C})\} /\!\!/ \mathrm{SL}_2(\mathbb{C}),$$

where $\mathrm{SL}_2(\mathbb{C})$ acts on homomorphisms by conjugation, and where the double bar indicates that the quotient is to be taken in the sense of geometric invariant theory; see Mumford, Fogarty and Kirwan [27]. In fact, if one follows the physical tradition that a quantization of a space X replaces the commutative algebra of functions on X by a noncommutative algebra of operators on a Hilbert space, an element of the quantization should be a *representation* of the skein algebra.

When A is a root of unity, a classical example of a finite-dimensional representation of the skein algebra $\mathcal{S}^A(S)$ arises from the Witten–Reshetikhin–Turaev topological quantum field theory associated to the fundamental representation of the quantum group $U_q(\mathfrak{sl}_2)$; see Blanchet, Habegger, Masbaum and Vogel [3], Bonahon and Wong [8], Reshetikhin and Turaev [30], Turaev [34] and Witten [35]. The main purpose of the current article is to provide a wider family of such representations when the surface S has at least one puncture. The case of closed surfaces is considered in our subsequent article [10], which builds on this one.

In part I [9], we identified invariants for irreducible finite-dimensional representations $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ in the case where A^2 is a primitive N^{th} root of unity with N odd. These invariants are a little easier to describe when $A^N = -1$, and most of the current article will be devoted to this case. We indicate in Section 6 how the other possible case when $A^N = +1$ can be deduced from this one. Because N is odd, the property that A^2 is a primitive N^{th} root of unity with $A^N = -1$ is equivalent to the property that A is a primitive N^{th} root of -1.

When $A^N = -1$, our main invariant is a point of the character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$. Its definition involves the n^{th} normalized Chebyshev polynomial $T_n(x)$ of the first kind, determined by the trigonometric identity that $2\cos n\theta = T_n(2\cos\theta)$. Equivalently, $\operatorname{Tr} M^n = T_n(\operatorname{Tr} M)$ for every matrix $M \in \mathrm{SL}_2(\mathbb{C})$.

A character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ associates a trace $\operatorname{Tr} r(K) \in \mathbb{C}$ to each closed curve K on the surface S. This trace is independent of the homomorphism $\pi_1(S) \to \mathrm{SL}_2(\mathbb{C})$

used to represent r, and of the representative chosen in the conjugacy class of $\pi_1(S)$ representing K. In fact, the character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ is defined in such a way that two homomorphisms $r\colon \pi_1(S) \to \mathrm{SL}_2(\mathbb{C})$ correspond to the same character if and only if they induce the same trace function $K \mapsto \operatorname{Tr} r(K)$.

Theorem 1 (Bonahon and Wong [9]) *Suppose that A is a primitive N^{th} root of -1 with N odd, and let $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ be an irreducible finite-dimensional representation of the Kauffman bracket skein algebra. Let $T_N(x)$ be the N^{th} normalized Chebyshev polynomial of the first kind.*

(1) *There exists a unique character $r_\rho \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ such that*
$$T_N(\rho([K])) = -(\operatorname{Tr} r_\rho(K)) \operatorname{Id}_E$$
for every framed knot $K \subset S \times [0,1]$ whose projection to S has no crossing and whose framing is vertical.

(2) *Let P_k be a small simple loop going around the k^{th} puncture of S, and consider it as a knot in $S \times [0,1]$ with vertical framing. Then there exists a number $p_k \in \mathbb{C}$ such that $\rho([P_k]) = p_k \operatorname{Id}_E$.*

(3) *The number p_k of (2) is related to the character $r_\rho \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ of (1) by the property that $T_N(p_k) = -\operatorname{Tr} r_\rho(P_k)$.*

The character $r_\rho \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ associated to the irreducible representation $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ by part (1) of Theorem 1 is the *classical shadow* of ρ. The numbers p_k defined by part (2) are the *puncture invariants* of the representation ρ. Part (3) shows that, once the classical shadow r_ρ is known, there are at most N possible values for each of the puncture invariants p_k.

The classical shadow provides one more example of a situation where a quantum object determines one of the classical objects that are being quantized. See also Lê [25] for another approach to the key results underlying Theorem 1.

The main result of this article is the following converse statement.

Theorem 2 *Assume that the surface S has at least one puncture, that its Euler characteristic is negative, that A is a primitive N^{th} root of -1 with N odd, and that we have*

(1) *a character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ which realizes some ideal triangulation of S in the sense discussed in Section 3;*

(2) *for each puncture of S, a number $p_k \in \mathbb{C}$ such that $T_N(p_k) = -\operatorname{Tr} r(P_k)$, where as in Theorem 1, P_k is a small loop going around the puncture.*

Then there exists an irreducible finite-dimensional representation $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ whose classical shadow is equal to r and whose puncture invariants are the p_k.

The requirement that r realizes some ideal representation is fairly mild. It can be shown to be satisfied by all points outside of an algebraic subset of complex codimension $2|\chi(S)|-1$ in the character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$.

The sequel [10; 11] to this paper greatly improves Theorem 2. In particular, it removes the requirements that r realizes an ideal triangulation, and that S has at least one puncture. It also shows that the representation provided by our construction is independent of the many choices made during the argument, so that its output is natural, in particular with respect to the action of the mapping class group $\pi_0\mathrm{Diff}(S)$ of the surface. The constructions and results of the current article are a key ingredient in the proofs of [10; 11].

The proof of Theorem 2 uses as a fundamental tool the quantum trace homomorphism $\mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathcal{T}^\omega(\lambda)$, constructed in Bonahon and Wong [7], which embeds the skein algebra in the quantum Teichmüller space. The *quantum Teichmüller space* is here incarnated as the Chekhov–Fock algebra $\mathcal{T}^\omega(\lambda)$ of an ideal triangulation λ of the surface, and is a quantization of an object that is closely related to the character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$. It is not as natural as the Kauffman bracket skein algebra, but its algebraic structure is very simple. In particular, its representation theory is relatively easy to analyze; see Bonahon and Liu [5]. The same holds for a smaller algebra $\mathcal{Z}^\omega(\lambda) \subset \mathcal{T}^\omega(\lambda)$ containing the image of the quantum trace homomorphism $\mathrm{Tr}_\lambda^\omega$. Composing representations of $\mathcal{Z}^\omega(\lambda)$ with the homomorphism $\mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathcal{Z}^\omega(\lambda)$ provides an extensive family of representations of the skein algebra $\mathcal{S}^A(S)$, which can then be used to prove Theorem 2.

The main technical challenge in this strategy is to compute the classical shadow of the representations of $\mathcal{S}^A(S)$ so obtained, in terms of the parameters controlling the original representations of $\mathcal{Z}^\omega(\lambda)$. This is provided by the miraculous cancellations discovered in [9]. These properties show that the quantum trace homomorphism $\mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathcal{Z}^\omega(\lambda)$ is well behaved with respect to the Chebyshev homomorphism $\mathcal{S}^{-1}(S) \to \mathcal{S}^A(S)$ used to define the classical shadow of a representation of $\mathcal{S}^A(S)$, and with respect to the Frobenius homomorphism $\mathcal{Z}^\iota(\lambda) \to \mathcal{Z}^\omega(\lambda)$ which computes the invariants of representations of $\mathcal{Z}^\omega(\lambda)$.

One of the steps in the proof, used to determine the algebraic structure of the algebra $\mathcal{Z}^\omega(\lambda)$, may be of interest by itself. This statement describes the structure of the Thurston intersection form on the set $\mathcal{W}(\tau;\mathbb{Z})$ of integer-valued edge weight systems for a train track τ. The result is well known for real-valued weights. However, the integer-valued case has subtler number-theoretic properties, resulting in the unexpected simultaneous occurrence of blocks $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 2 \\ -2 & 0 \end{smallmatrix}\right)$ in the block diagonalization of the Thurston form. See Theorem 26 in the appendix. Because of the ubiquity of the

Thurston intersection form in many geometric problems (for instance, the relationship between complex lengths and the shear-bend cocycle $\beta \in \mathcal{W}(\tau; \mathbb{C}/2\pi\mathrm{i}\mathbb{Z})$ of a pleated surface; see Bonahon [4]), this statement is probably of interest beyond the quantum topology scope of the current article.

Recent works of Abdiel and Frohman [1; 20], Frohman and Kania-Bartoszyńska [21], and Frohman, Kania-Bartoszyńska and Lê [22] develop another construction of representations of $\mathcal{S}^A(S)$ with a given classical shadow $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, valid for r in a Zariski dense open subset of $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$. In particular, the recent preprint [22] abstractly shows that these representations are isomorphic to ours. It would be interesting to compare the two approaches, as the construction pioneered by Abdiel and Frohman in [1; 20] is simple and elegant while ours is more explicit. See also Takenov [31] for an earlier viewpoint on the case of small surfaces.

Acknowledgements This research was partially supported by grants DMS-1105402, DMS-1105692, and DMS-1406559 from the US National Science Foundation. In addition, the article was extensively rewritten and reorganized while Bonahon was a Simons Fellow (grant 301050 from the Simons Foundation) in 2014–15, as well as a Simons Visiting Professor at the Mathematical Sciences Research Institute in Berkeley, California, (NSF grant 09032078000) in the Spring 2015 semester.

1 The Chekhov–Fock algebra and the quantum trace homomorphism

1.1 The Chekhov–Fock algebra

The Chekhov–Fock algebra (introduced in [5] as a reinterpretation of key insights from [19; 17; 18]) is the avatar of the quantum Teichmüller space associated to an ideal triangulation of the surface S. See also [24] for a related construction, and [5; 26] for more discussion.

If S is obtained from a compact surface $\overline{S}$ by removing finitely many points $v_1, \dots, v_s$, an *ideal triangulation* of S is a triangulation λ of $\overline{S}$ whose vertex set is exactly $\{v_1, v_2, \dots, v_s\}$. The surface S admits an ideal triangulation if and only if it is noncompact and if its Euler characteristic is negative; we will consequently assume that these properties are satisfied throughout the article. If the surface has genus g and s punctures, an ideal triangulation then has $n = 6g + 3s - 6$ edges and $4g + 2s - 4$ faces.

Let $e_1, e_2, \dots, e_n$ denote the edges of λ. Let $a_i \in \{0, 1, 2\}$ be the number of times an end of the edge e_j immediately succeeds an end of e_i when going counterclockwise around a puncture of S, and set $\sigma_{ij} = a_{ij} - a_{ji} \in \{-2, -1, 0, 1, 2\}$. The *Chekhov–Fock algebra* $\mathcal{T}^\omega(\lambda)$ of λ is the algebra defined by generators $Z_1^{\pm 1}, Z_2^{\pm 1}, \dots, Z_n^{\pm 1}$ associated to the edges $e_1, e_2, \dots, e_n$ of λ, and by the relations

$$Z_i Z_j = \omega^{2\sigma_{ij}} Z_j Z_i.$$

Remark 3 The actual Chekhov–Fock algebra $\mathcal{T}^q(\lambda)$ that is at the basis of the quantum Teichmüller space uses the constant $q = \omega^4$ instead of ω. The generators Z_i of $\mathcal{T}^\omega(\lambda)$ appearing here are designed to model square roots of the original generators of $\mathcal{T}^q(\lambda)$.

An element of the Chekhov–Fock algebra $\mathcal{T}^\omega(\lambda)$ is a linear combination of monomials $Z_{i_1}^{n_1} Z_{i_2}^{n_2} \cdots Z_{i_l}^{n_l}$ in the generators Z_i, with $n_1, n_2, \dots, n_l \in \mathbb{Z}$. Because of the skew-commutativity relation $Z_i Z_j = \omega^{2\sigma_{ij}} Z_j Z_i$, the order of the variables in such a monomial does matter. It is convenient to use the following symmetrization trick. The *Weyl quantum ordering* for $Z_{i_1}^{n_1} Z_{i_2}^{n_2} \cdots Z_{i_l}^{n_l}$ is the monomial

$$[Z_{i_1}^{n_1} Z_{i_2}^{n_2} \cdots Z_{i_l}^{n_l}] = \omega^{-\sum_{u<v} n_u n_v \sigma_{i_u i_v}} Z_{i_1}^{n_1} Z_{i_2}^{n_2} \cdots Z_{i_l}^{n_l}.$$

The formula is specially designed so that $[Z_{i_1}^{n_1} Z_{i_2}^{n_2} \cdots Z_{i_l}^{n_l}]$ is invariant under any permutation of the $Z_{i_u}^{n_u}$. Note that the algebraic structure of the Chekhov–Fock algebra $\mathcal{T}^\omega(\lambda)$ depends only on the square ω^2, but that the Weyl quantum ordering depends on the choice of ω.

1.2 The quantum trace homomorphism

Theorem 4 [7] *For $A = \omega^{-2}$, there exists an injective algebra homomorphism*

$$\mathrm{Tr}_\lambda^\omega \colon \mathcal{S}^A(S) \to \mathcal{T}^\omega(\lambda).$$

The specific homomorphism $\mathrm{Tr}_\lambda^\omega$ constructed in [7] is the *quantum trace homomorphism.* It is uniquely determined by certain properties stated in that article, but for now we need only use that it exists and satisfies the properties given in Section 1.3 below.

1.3 The Chebyshev and Frobenius homomorphisms

We now assume that A is a primitive N^{th} root of -1 with N odd. Recall that T_N denotes the N^{th} normalized Chebyshev polynomial, defined by the property that $\cos N\theta = \frac{1}{2} T_N(2\cos\theta)$ for every θ.

Theorem 5 [9] *When A is a primitive N^{th} root of -1 with N odd, there is a unique algebra homomorphism $\boldsymbol{T}^A\colon \mathcal{S}^{-1}(S) \to \mathcal{S}^A(S)$ such that*

$$\boldsymbol{T}^A([K]) = T_N([K])$$

for every framed knot $K \subset S \times [0,1]$ whose projection to S has no crossing and whose framing is vertical. In addition, the image of $\boldsymbol{T}^A$ is central in $\mathcal{S}^A(S)$. □

For a framed link $K \subset S \times [0,1]$ whose projection to S is allowed to have crossings, the image $\boldsymbol{T}^A([K])$ is equal to the element $[K^{T_N}] \in \mathcal{S}^A(S)$ defined by threading the Chebyshev polynomial T_N along all components of K; see [9] for a precise definition.

The homomorphism $\boldsymbol{T}^A$ provided by Theorem 5 is the *Chebyshev homomorphism*. It is a key ingredient in the definition of the invariants of Theorem 1.

There is an analogous and much simpler homomorphism at the level of the Chekhov–Fock algebra, namely the following *Frobenius homomorphism*.

Proposition 6 *If $\iota = \omega^{N^2}$, there is an algebra homomorphism*

$$\boldsymbol{F}^\omega\colon \mathcal{T}^\iota(\lambda) \to \mathcal{T}^\omega(\lambda)$$

which maps each generator $Z_i \in \mathcal{T}^\iota(\lambda)$ to $Z_i^N \in \mathcal{T}^\omega(\lambda)$, where in the first instance $Z_i \in \mathcal{T}^\iota(\lambda)$ denotes the generator associated to the i^{th} edge e_i of λ, whereas the second time $Z_i \in \mathcal{T}^\omega(\lambda)$ denotes the generator of $\mathcal{T}^\omega(\lambda)$ associated to the same edge e_i. □

Note that $\iota^2 = \omega^{2N^2} = A^{-N^2} = (-1)^N = -1$, so $\iota = \pm \mathrm{i}$.

The following compatibility statement, which connects the Chebyshev homomorphism to the Frobenius homomorphism through appropriate quantum trace homomorphisms, is fundamental for our arguments. This result encapsulates the miraculous cancellations of [9].

Theorem 7 [9] *The diagram*

$$\begin{tikzcd}
\mathcal{S}^A(S) \arrow[r, "\mathrm{Tr}_\lambda^\omega"] & \mathcal{T}^\omega(\lambda) \\
\mathcal{S}^{-1}(S) \arrow[u, "\boldsymbol{T}^A"] \arrow[r, "\mathrm{Tr}_\lambda^\iota"] & \mathcal{T}^\iota(\lambda) \arrow[u, "\boldsymbol{F}^\omega"']
\end{tikzcd}$$

is commutative. Namely, for every skein $[K] \in \mathcal{S}^{-1}(S)$, the quantum trace $\mathrm{Tr}_\lambda^\omega(\boldsymbol{T}^A([K]))$ of $\boldsymbol{T}^A([K])$ is obtained from the classical trace polynomial $\mathrm{Tr}_\lambda^\iota([K])$ by replacing each generator $Z_i \in \mathcal{T}^\iota(\lambda)$ by $Z_i^N \in \mathcal{T}^\omega(\lambda)$. □

2 The balanced Chekhov–Fock algebra

2.1 Definition of the balanced Chekhov–Fock algebra

The quantum trace homomorphism $\mathrm{Tr}_\lambda^\omega$ of Theorem 4 (and [7]) is far from being surjective. Indeed, for a skein $[K] \in \mathcal{S}^A(S)$ represented by a framed link $K \subset S \times [0,1]$, the exponents of the monomials $Z_1^{k_1} Z_2^{k_2} \cdots Z_n^{k_n}$ appearing in the expression of $\mathrm{Tr}_\lambda^\omega([K])$ are *balanced*, in the sense that they satisfy the following parity condition: for every triangle T_j of the ideal triangulation λ, the sum $k_{i_1} + k_{i_2} + k_{i_3}$ of the exponents of the generators $Z_{i_1}, Z_{i_2}, Z_{i_3}$ associated to the sides of T_j is even.

Let $\mathcal{Z}^\omega(\lambda)$ denote the subalgebra of $\mathcal{T}^\omega(\lambda)$ generated by all monomials satisfying this exponent parity condition. By definition, $\mathcal{Z}^\omega(\lambda)$ is the *balanced Chekhov–Fock algebra* of the ideal triangulation λ. It is designed so that the quantum trace homomorphism restricts to a homomorphism $\mathrm{Tr}_\lambda^\omega \colon \mathcal{S}^A(S) \to \mathcal{Z}^\omega(\lambda)$.

To keep track of the exponent parity condition defining the monomials of $\mathcal{Z}^\omega(\lambda)$, it is convenient to consider a train track τ_λ which, on each triangle T_j of the ideal triangulation λ, looks as in Figure 2. In particular, τ_λ has one switch for each edge of λ, and three edges for each triangle of λ. Let $\mathcal{W}(\tau_\lambda; \mathbb{Z})$ be the set of integer edge weight systems α for τ_λ, assigning a number $\alpha(e) \in \mathbb{Z}$ to each edge e of τ_λ in such a way that, at each switch, the weights of the edges incoming on one side add up to the sum of the weights of the edges outgoing on the other side.

There is a natural map $\mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{Z}^n$ which, given an edge weight system, associates to each of the n switches of τ_λ the sum of the weights of the edges incoming on any side of the switch. Then an element $(k_1, k_2, \dots, k_n) \in \mathbb{Z}^n$ is in the image of this map if and only if it satisfies the parity condition defining the monomials of $\mathcal{Z}^\omega(\lambda)$, namely if and only if the sum of the coordinates associated to the sides of each triangle of λ is even. Also, the map $\mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{Z}^n$ is easily seen to be injective. Since the image of this map has finite index, it follows that $\mathcal{W}(\tau_\lambda; \mathbb{Z})$ is isomorphic to $\mathbb{Z}^n$ as an abelian group.

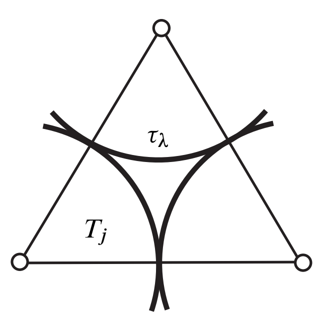


Figure 2

This enables us to give a different description of $\mathcal{Z}^\omega(\lambda)$. For a weight system $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$, which assigns a weight $\alpha_i \in \mathbb{Z}$ to the i^{th} edge e_i of λ (= the i^{th} switch of τ_λ), define

$$Z_\alpha = [Z_1^{\alpha_1} Z_2^{\alpha_2} \cdots Z_n^{\alpha_n}] \in \mathcal{Z}^\omega(\lambda),$$

where the bracket [ ] denotes the Weyl quantum ordering defined in Section 1.1.

The above discussion proves the following fact.

Lemma 8 *As $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ ranges over all weight systems for the train track τ_λ, the associated monomials Z_α form a basis for the vector space $\mathcal{Z}^\omega(\lambda)$.* □

We can elaborate a little on the structure of the group $\mathcal{W}(\tau_\lambda;\mathbb{Z})$. By definition of the parity condition, $\mathcal{W}(\tau_\lambda;\mathbb{Z}) \subset \mathbb{Z}^n$ contains the subset $(2\mathbb{Z})^n$ consisting of all switch weight systems $(\alpha_1,\alpha_2,\dots,\alpha_n) \in \mathbb{Z}^n$ where the α_i are even. Also, given $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$, we can define a chain with coefficients in $\mathbb{Z}_2$ by endowing each edge e of the train track τ_λ with the modulo 2 reduction of the weight $\alpha(e) \in \mathbb{Z}$. The switch relations guarantee that this chain is closed, and this defines a natural homomorphism $\mathcal{W}(\tau_\lambda;\mathbb{Z}) \to H_1(S;\mathbb{Z}_2)$.

Lemma 9 *The inclusion map and homomorphism above define an exact sequence*

$$0 \to (2\mathbb{Z})^n \to \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to H_1(S;\mathbb{Z}_2) \to 0.$$

Proof The homomorphism $\mathcal{W}(\tau_\lambda;\mathbb{Z}) \to H_1(S;\mathbb{Z}_2)$ can also be expressed in terms of the dual graph Γ_λ of the triangulation λ. Indeed, the class $[\alpha] \in H_1(S;\mathbb{Z}_2)$ induced by $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ is also realized by endowing each edge f_i of Γ_λ with the modulo 2 reduction of the switch weight α_i associated by α to the edge e_i of λ that is dual to f_i; the parity condition guarantees that this chain is really closed. The result then immediately follows from the definitions, and from the isomorphism $H_1(\Gamma_\lambda;\mathbb{Z}_2) \cong H_1(S;\mathbb{Z}_2)$ coming from the fact that the surface S deformation retracts to the dual graph Γ_λ. □

Note that the exact sequence of Lemma 9 admits no partial splitting.

2.2 The algebraic structure of the balanced Chekhov–Fock algebra

We first describe the multiplicative structure of the balanced Chekhov–Fock algebra $\mathcal{Z}^\omega(\lambda)$ in the context of Lemma 8.

The weight system space $\mathcal{W}(\tau_\lambda;\mathbb{Z})$ of the train track τ_λ carries a very natural antisymmetric bilinear form

$$\Omega\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \times \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{Z},$$

the *Thurston intersection form* defined by the property that, for $\alpha, \beta \in \mathcal{W}(\tau_\lambda; \mathbb{Z})$,

$$\Omega(\alpha,\beta) = \frac{1}{2} \sum_{e \text{ right of } e'} (\alpha(e)\beta(e') - \alpha(e')\beta(e)),$$

where the sum is over all pairs (e, e') of edges of τ_λ such that e and e' come out of the same side of some switch of τ_λ, with e to the right of e'. See Lemma 28 in the appendix for a more conceptual interpretation of Ω, and for a proof that $\Omega(\alpha,\beta)$ is really an integer.

Lemma 10 *For every $\alpha, \beta \in \mathcal{W}(\tau_\lambda; \mathbb{Z})$,*

$$Z_\alpha Z_\beta = \omega^{2\Omega(\alpha,\beta)} Z_{\alpha+\beta}.$$

In particular, $Z_\alpha Z_\beta = \omega^{4\Omega(\alpha,\beta)} Z_\beta Z_\alpha$.

Proof The second statement, that $Z_\alpha Z_\beta = \omega^{4\Omega(\alpha,\beta)} Z_\beta Z_\alpha$, is a simple computation. After observing that this property holds for any ω (not just roots of unity), the first statement, that $Z_\alpha Z_\beta = \omega^{2\Omega(\alpha,\beta)} Z_{\alpha+\beta}$, then follows by definition of the Weyl quantum ordering. □

This is particularly simple if we replace ω by $\iota = \omega^{N^2}$, with the assumption that $A^{2N} = 1$ so that $\iota^4 = \omega^{4N^2} = A^{-2N^2} = 1$.

Corollary 11 *If $\iota^4 = 1$, the algebra $\mathcal{Z}^\iota(\lambda)$ is commutative.* □

In general, the key to understanding the algebraic structure of $\mathcal{Z}^\omega(\lambda)$ is Lemma 12.

For $k = 1, \dots, s$, the k^{th} puncture of S specifies an element $\eta_k \in \mathcal{W}(\tau_\lambda; \mathbb{Z})$ defined as follows: for every edge e of the train track τ_λ, the edge weight $\eta_k(e)$ is equal to 1 if e is adjacent to the annulus component of $S - \tau_\lambda$ that surrounds this puncture, and is equal to 0 otherwise.

Recall that the surface S has genus g and s punctures.

Lemma 12 *The lattice $\mathcal{W}(\tau_\lambda; \mathbb{Z}) \cong \mathbb{Z}^n$ admits a basis in which the matrix of the Thurston intersection form Ω is block diagonal with g blocks $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $2g+s-3$ blocks $\left(\begin{smallmatrix} 0 & 2 \\ -2 & 0 \end{smallmatrix}\right)$ and s blocks (0). In addition, the kernel of Ω is freely generated by the elements $\eta_1, \eta_2, \dots, \eta_s \in \mathcal{W}(\tau_\lambda; \mathbb{Z})$ associated to the punctures of S as above.*

Proof This is a special case of a result given by Theorem 26 in the appendix, which determines the algebraic structure of the Thurston intersection form for a general train track τ. When applying this result to the train track τ_λ, the numbers h, n_{even} and n_{odd} of Theorem 26 are respectively equal to the genus g of the surface S, to the number s of punctures of S, and to the number $4g+2s-4$ of triangles of the ideal triangulation λ. □

The combination of Lemmas 8, 10 and 12 now provides the complete algebraic structure of the balanced Chekhov–Fock algebra $\mathcal{Z}^\omega(\lambda)$. Let $\mathcal{W}^q$ denote the algebra, known as the quantum torus, defined by generators $X^{\pm 1}$, $Y^{\pm 1}$ and by the relation $XY = qYX$.

Corollary 13 *For $q=\omega^4$, the balanced Chekhov–Fock algebra $\mathcal{Z}^\omega(\lambda)$ is isomorphic to*

$$\mathcal{W}^q_1 \otimes \mathcal{W}^q_2 \otimes \cdots \otimes \mathcal{W}^q_g \otimes \mathcal{W}^{q^2}_{g+1} \otimes \mathcal{W}^{q^2}_{g+2} \otimes \cdots \otimes \mathcal{W}^{q^2}_{3g+s-3} \otimes \mathbb{C}[H_1] \otimes \mathbb{C}[H_2] \otimes \cdots \otimes \mathbb{C}[H_s],$$

where each $\mathcal{W}^q_i$ is a copy of the quantum torus $\mathcal{W}^q$, each $\mathcal{W}^{q^2}_j$ is a copy of $\mathcal{W}^{q^2}$, and each $\mathbb{C}[H_k]$ is a polynomial algebra in the variable H_k.

In addition, the s central generators $H_k = Z_{\eta_k} \in \mathcal{Z}^\omega(\lambda)$ are naturally associated to the punctures of S, and are defined by the edge weight systems $\eta_k \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ generating the kernel of the Thurston intersection form Ω as in Lemma 12. □

2.3 Representations of the balanced Chekhov–Fock algebra

The algebraic structure of the balanced Chekhov–Fock algebra $\mathcal{Z}^\omega(\lambda)$ determined in Corollary 13 is relatively simple. This makes it easy to classify its irreducible finite-dimensional representations.

As usual, we assume that $A=\omega^{-2}$ is a primitive N^{th} root of -1, with N odd.

Proposition 14 *Let $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ be an irreducible finite-dimensional representation of $\mathcal{Z}^\omega(\lambda)$. There exists a map $\zeta_\mu\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ and numbers $h_k \in \mathbb{C}^*$, with $k=1,\dots,s$, associated to the punctures of the surface S such that*

1. *$\mu(Z_\alpha^N) = \zeta_\mu(\alpha)\,\mathrm{Id}_E$ for every edge weight system $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ with associated monomial $Z_\alpha \in \mathcal{Z}^\omega(\lambda)$;*
2. *$\zeta_\mu(\alpha+\beta) = (-1)^{\Omega(\alpha,\beta)}\zeta_\mu(\alpha)\zeta_\mu(\beta)$ for every $\alpha,\beta \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$, where Ω is the Thurston intersection form;*
3. *$\mu(H_k) = h_k\,\mathrm{Id}_E$ for the central element $H_k = Z_{\eta_k} \in \mathcal{Z}^\omega(\lambda)$ associated to the k^{th} puncture of S as in Corollary 13;*
4. *$\zeta_\mu(\eta_k) = h_k^N$ for the weight system $\eta_k \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ associated to the k^{th} puncture of S as in Lemma 12.*

Proof For every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$, Lemma 10 shows that the element $Z_\alpha^N = Z_{N\alpha}$ is central in $\mathcal{Z}^\omega(\lambda)$. In particular, if $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ is an irreducible finite-dimensional representation of $\mathcal{Z}^\omega(\lambda)$, there is a number $\zeta_\mu(\alpha) \in \mathbb{C}^*$ such that $\mu(Z_\alpha^N) = \zeta_\mu(\alpha)\,\mathrm{Id}_E$. In addition, Lemma 10 shows that

$$Z_\alpha^N Z_\beta^N = \omega^{2N^2\Omega(\alpha,\beta)} Z_{\alpha+\beta}^N = (-1)^{\Omega(\alpha,\beta)} Z_{\alpha+\beta}^N,$$

so the map $\zeta_\mu\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ satisfies property (2).

Similarly, Corollary 13 shows that each H_k is central in $\mathcal{Z}^\omega(\lambda)$, so $\mu(H_k) = h_k\,\mathrm{Id}$ for some $h_k \in \mathbb{C}^*$. Then $h_k^N\,\mathrm{Id}_E = \mu(H_k^N) = \mu(Z_{\eta_k}^N) = \zeta_\mu(\eta_k)\,\mathrm{Id}_E$ since $H_k = Z_{\eta_k}$, so $\zeta_\mu(\eta_k) = h_k^N$. □

A map $\zeta\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ that satisfies condition (2) of Proposition 14 is a *twisted homomorphism* twisted by the Thurston form Ω, or more precisely twisted by the symmetric map $(\alpha,\beta) \mapsto (-1)^{\Omega(\alpha,\beta)}$. This notion will probably look less intimidating once one realizes that a twisted homomorphism is completely determined by the assignment of a nonzero complex number to each of the n generators of the group $\mathcal{W}(\tau_\lambda;\mathbb{Z}) \cong \mathbb{Z}^n$.

Proposition 15 *Suppose that we are given a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ twisted by the Thurston form Ω and, for each of the punctures of S, a number $h_k \in \mathbb{C}^*$ such that $h_k^N = \zeta(\eta_k)$. Then, up to isomorphism, there exists a unique irreducible finite-dimensional representation $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ such that*

(1) $\zeta_\mu = \zeta$, *namely* $\mu(Z_\alpha^N) = \zeta(\alpha)\,\mathrm{Id}_E$ *for every* $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$;

(2) $\mu(H_k) = h_k\,\mathrm{Id}_E$ *for* $k = 1,\dots,s$.

In addition, for such a representation, the vector space E has dimension N^{3g+s-3}.

Proof Using elementary linear algebra, this follows immediately from Corollary 13. More precisely, consider the isomorphism

$$\mathcal{Z}^\omega(\lambda) \cong \mathcal{W}_1^q \otimes \cdots \otimes \mathcal{W}_g^q \otimes \mathcal{W}_{g+1}^{q^2} \otimes \cdots \otimes \mathcal{W}_{3g+s-3}^{q^2} \otimes \mathbb{C}[H_1] \otimes \cdots \otimes \mathbb{C}[H_s]$$

provided by Corollary 13.

For $1 \leqslant i \leqslant 3g+s-3$, let $X_i^{\pm1}$ and $Y_i^{\pm1}$ denote the generators of $\mathcal{W}_i^q$ or $\mathcal{W}_i^{q^2}$ (satisfying the relation $X_iY_i = qY_iX_i$ if $1 \leqslant i \leqslant g$ and $X_iY_i = q^2Y_iX_i$ if $g < i \leqslant 3g+s-3$). The proof of Corollary 13 shows that these generators are of the form $X_i = Z_{\alpha_i}$, $Y_i = Z_{\beta_i}$ and $H_k = Z_{\eta_k}$ for some edge weight systems $\alpha_i, \beta_i, \eta_k \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$. In addition, the α_i, β_i and η_k form a basis for $\mathcal{W}(\tau_\lambda;\mathbb{Z}) \cong \mathbb{Z}^n$.

Because N is odd, $q=\omega^4$ and q^2 are both primitive N^{th} roots of unity. Arbitrarily pick N^{th} roots $\zeta(\alpha_i)^{1/N}$ and $\zeta(\beta_i)^{1/N}$, and define $\mu_i\colon \mathcal{W}_i^q \to \mathrm{End}(E_i)$ by the property that, if $v_1, v_2, \dots, v_N$ form a basis for $E_i \cong \mathbb{C}^N$,

$$\mu_i(X_i)(v_j) = \begin{cases} -\zeta(\alpha_i)^{1/N} q^j v_j & \text{if } 1 \le i \le g, \\ \zeta(\alpha_i)^{1/N} q^{2j} v_j & \text{if } g < i \leqslant 3g+s-3, \end{cases}$$

$$\mu_i(Y_i)(v_j) = \zeta(\beta_i)^{1/N} v_{j+1}.$$

Then for $E = E_1 \otimes E_2 \otimes \cdots \otimes E_{3g+s-3}$, define $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ by the property that μ coincides with $\mu_1 \otimes \mu_2 \otimes \cdots \otimes \mu_{3g+s-3}$ on $\mathcal{W}_1^q \otimes \mathcal{W}_2^q \otimes \cdots \otimes \mathcal{W}_{3g+s-3}^q$, and $\mu(H_k) = h_k \,\mathrm{Id}_E$ for every $k = 1, \dots, s$.

It is immediate that μ satisfies the required properties. The fact that μ is irreducible, and that every irreducible representation is isomorphic to μ, is easily proved by elementary linear algebra; see for instance [5, Section 4] for details. □

3 Pleated surfaces and homomorphisms to $\mathrm{SL}_2(\mathbb{C})$

Let us consider the special case of Proposition 15 when $N = 1$. In particular, $A = -1$ and $\iota = \omega = \pm\mathrm{i}$. Since the Chebyshev polynomial $T_1(x)$ is equal to x, the choice of puncture invariants h_k is irrelevant and Proposition 15 associates to any twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{C}^*$ a representation $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathrm{End}(\mathbb{C})$. By composition with the quantum trace homomorphism $\mathrm{Tr}_\lambda^\iota\colon \mathcal{S}^{-1}(S) \to \mathcal{Z}^\iota(\lambda)$ of Theorem 4, we now have a homomorphism

$$\rho_\zeta = \mu_\zeta \circ \mathrm{Tr}_\lambda^\iota\colon \mathcal{S}^{-1}(S) \to \mathrm{End}(\mathbb{C}) = \mathbb{C}.$$

We can then apply the case $N = 1$ of Theorem 1 (which actually is an observation of Doug Bullock, Charlie Frohman, Joanna Kania-Bartoszyńska, Jozef Przytycki and Adam Sikora [12; 13; 14; 15; 29] and plays a crucial rôle in the proof of Theorem 1 in its full generality). It provides a character $r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ such that

$$\rho_\zeta([K]) = -\operatorname{Tr} r_\zeta(K)$$

for every framed knot $K \subset S \times [0,1]$. The property is valid for all knots, not just those whose projection to S has no double point [12; 13; 14; 15; 29].

It is natural to ask which elements of $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ are obtained in this way. The answer involves the following geometric definition.

Let $\widetilde{S}$ be the universal cover of S, and let $\widetilde{\lambda}$ be the ideal triangulation of $\widetilde{S}$ obtained by lifting the edges and faces of λ. Identify $\mathrm{PSL}_2(\mathbb{C})$ to the isometry group of the hyperbolic 3–space $\mathbb{H}^3$. A *pleated surface* with *pleating locus* λ is the data $(\widetilde{f}, \overline{r})$ of a map $\widetilde{f}\colon \widetilde{S} \to \mathbb{H}^3$ and a group homomorphism $\overline{r}\colon \pi_1(S) \to \mathrm{PSL}_2(\mathbb{C})$ such that

(1) $\widetilde{f}$ homeomorphically sends each edge of $\widetilde{\lambda}$ to a complete geodesic of the hyperbolic space $\mathbb{H}^3$, and every face of $\widetilde{\lambda}$ to a totally geodesic ideal triangle of $\mathbb{H}^3$, with vertices on the sphere at infinity $\partial_\infty \mathbb{H}^3$;

(2) $\widetilde{f}$ is $\overline{r}$–equivariant, in the sense that $\widetilde{f}(\gamma\widetilde{x}) = \overline{r}(\gamma)(\widetilde{f}(\widetilde{x}))$ for every $\gamma \in \pi_1(S)$ and every $\widetilde{x} \in \widetilde{S}$.

Following the terminology introduced in [32], we say that the group homomorphism $\overline{r}\colon \pi_1(S) \to \mathrm{PSL}_2(\mathbb{C})$ *realizes* the ideal triangulation λ if there exists a pleated surface $(\widetilde{f}, \overline{r})$ with pleating locus λ. By extension, a point in the character variety $\mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$ *realizes* λ if it can be represented by a homomorphism $\overline{r}\colon \pi_1(S) \to \mathrm{PSL}_2(\mathbb{C})$ realizing λ. Finally, a character in $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ *realizes* λ if it is sent to a point of $\mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$ realizing λ by the canonical projection $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) \to \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$.

We are now ready to state the result promised. At the beginning of this section, we associated a character $r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ to each twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{C}^*$.

Proposition 16 *A character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ is associated to a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{C}^*$ as above if and only it realizes the ideal triangulation λ.*

Proof Suppose that $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ realizes the ideal triangulation λ. By definition, there exists a pleated surface $(\widetilde{f}, \overline{r})$ with pleating locus λ, where the homomorphism $\overline{r}\colon \pi_1(S) \to \mathrm{PSL}_2(\mathbb{C})$ represents the image of r under the projection $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) \to \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$.

The pleated surface $(\widetilde{f}, \overline{r})$ determines, for each edge $\widetilde{e}_i$ of the ideal triangulation $\widetilde{\lambda}$ of $\widetilde{S}$, a complex weight $\widetilde{x}_i \in \mathbb{C}^*$ defined as follows: If $\widetilde{Q}_i \subset \widetilde{S}$ is the quadrilateral formed by the two faces of $\widetilde{\lambda}$ meeting along the edge $\widetilde{e}_i$, then $-\widetilde{x}_i$ is the cross-ratio of the four vertices of $\widetilde{f}(\widetilde{Q}_i)$ in the sphere at infinity $\mathbb{C} \cup \{\infty\}$ of $\mathbb{H}^3$. These edge weights $\widetilde{x}_i$ are equivariant under the action of $\pi_1(S)$, and therefore descend to a system of weights x_i for the edges e_i of λ. The edge weights $x_i \in \mathbb{C}^*$ are the *shear-bend parameters* of the pleated surface $(\widetilde{f}, \overline{r})$.

Choose square roots $z_i = \sqrt{x_i}$. Then for every closed curve K in S, there is an explicit formula that expresses the trace $\mathrm{Tr}\,\overline{r}(K)$ as a Laurent polynomial in the z_i; see for instance [7, Sections 1.3–1.4]. Note that there necessarily is a sign ambiguity in this formula, as the trace of an element of $\mathrm{PSL}_2(\mathbb{C})$ is only defined up to sign. Another sign ambiguity occurs in the choice of the square roots $z_i = \sqrt{x_i}$.

We will use these edge weights $z_i \in \mathbb{C}^*$ to construct representations of $\mathcal{Z}^\iota(\lambda)$ and $\mathcal{S}^{-1}(S)$. Recall that a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{C}^*$ is equivalent to the data of its value on a set of generators of $\mathcal{W}(\tau_\lambda;\mathbb{Z}) \cong Z^n$. We can therefore find such a twisted homomorphism such that

$$\zeta(\alpha) = \pm z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}$$

for every edge weight system $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ assigning weight $\alpha_i \in \mathbb{Z}$ to the edge e_i of λ. The $\pm$ signs are here required by the twisting. In addition, a simple manipulation of the formula for the Thurston intersection form (or a use of Lemma 10) show that $\Omega(\alpha,\beta)$ is even whenever $\alpha \in (2\mathbb{Z})^n \subset \mathcal{W}(\tau_\lambda;\mathbb{Z})$ assigns even weights $\alpha_i \in \mathbb{Z}$ to all edges of λ; in particular, there is no twisting on $(2\mathbb{Z})^n \subset \mathcal{W}(\tau_\lambda;\mathbb{Z})$. Using Lemma 9, we can therefore arrange that, for every $\alpha \in (2\mathbb{Z})^n$,

$$\zeta(\alpha) = +z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}.$$

Note that there are several possible choices for ζ, coming from the signs $\pm$. In fact, Lemma 9 shows that there are exactly 2^d possibilities for ζ, where d is the dimension of $H_1(S;\mathbb{Z}_2)$. We will later adjust the choice of ζ so that it fits our purposes.

Let $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathrm{End}(\mathbb{C}) = \mathbb{C}$ be the representation of $\mathcal{Z}^\iota(\lambda)$ associated to the twisted homomorphism ζ by Proposition 15. Namely, $\mu(Z_\alpha) = \zeta(\alpha)$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$.

The definition of the quantum trace homomorphism $\mathrm{Tr}_\lambda^\iota\colon \mathcal{S}^{-1}(S) \to \mathcal{Z}^\iota(\lambda)$ in [7] was specially designed to copy the formula expressing the trace $\mathrm{Tr}\,\overline{r}(K)$ as a Laurent polynomial in the square roots $z_i = \sqrt{x_i}$ of the shear-bend parameters of the pleated surface $(\widetilde{f},\overline{r})$. In particular, because of the key property that $\mu_\zeta(Z_i^2) = +z_i^2$,

$$\mu_\zeta \circ \mathrm{Tr}_\lambda^\iota([K]) = \pm\,\mathrm{Tr}\,r(K)$$

for every framed knot $K \subset S \times [0,1]$, where the sign $\pm$ depends on K and on the choice of the square roots $z_i = \sqrt{x_i}$; see the discussion in [7, Sections 1.3–1.4].

As discussed at the beginning of this section, the homomorphism $\mu_\zeta \circ \mathrm{Tr}_\lambda^\iota\colon \mathcal{S}^{-1}(S) \to \mathbb{C}$ also defines a character $r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ such that

$$\mu_\zeta \circ \mathrm{Tr}_\lambda^\iota([K]) = -\,\mathrm{Tr}\,r_\zeta(K)$$

for every framed knot $K \subset S \times [0,1]$. As a consequence, $\mathrm{Tr}\,r_\zeta(K) = \pm\,\mathrm{Tr}\,r(K)$ for every knot K.

At this point, there is no reason for the two characters r and $r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ to coincide. However, by construction, they project to the same $\mathrm{PSL}_2(\mathbb{C})$–valued character in $\mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$. Their difference can therefore be encoded by a cohomology class $\varepsilon \in H^1(S;\mathbb{Z}_2)$ such that, for every knot $K \subset S \times [0,1]$,

$$\mathrm{Tr}\,r(K) = (-1)^{\varepsilon(K)}\,\mathrm{Tr}\,r_\zeta(K).$$

Each edge e_i of the ideal triangulation λ is Poincaré dual to a cohomology class $\varepsilon_i \in H^1(S;\mathbb{Z}_2)$. Replacing the square root $z_i = \sqrt{x_i}$ by the other square root $-z_i$ has the effect of replacing r_ζ with $\varepsilon_i r_\zeta$. Since the ε_i generate $H^1(S;\mathbb{Z}_2)$, we can therefore adjust the choice of the square roots $z_i = \sqrt{x_i}$ so that the characters r and $r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ are now equal.

This proves that, if the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ realizes the ideal triangulation λ, there exists a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ whose associated character $r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ is equal to r.

Conversely, suppose that $r = r_\zeta \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ is associated to a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ as above. More precisely, consider the corresponding representation $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathrm{End}(\mathbb{C}) = \mathbb{C}$, defined by the property that $\mu_\zeta(Z_\alpha) = \zeta(\alpha)$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$. Then for every framed knot $K \subset S \times [0,1]$,

$$\mu_\zeta \circ \mathrm{Tr}_\lambda^\iota([K]) = -\operatorname{Tr} r(K).$$

The generator $Z_i \in \mathcal{T}^\iota(\lambda)$ associated to the edge e_i of λ does not satisfy the exponent parity condition defining the balanced Chekhov–Fock algebra $\mathcal{Z}^\iota(\lambda)$, but its square does. We can therefore consider $x_i = \mu_\zeta(Z_i^2) \in \mathbb{C}$, which is different from 0 since Z_i^2 is invertible.

We can then construct a pleated surface $(\widetilde{f}, \overline{r})$ whose pleating locus is equal to λ and whose shear-bend parameters are equal to the edge weights x_i. In particular, this pleated surface is equivariant with respect to a homomorphism $\overline{r}\colon \pi_1(S) \to \mathrm{PSL}_2(\mathbb{C})$, which defines a character $\overline{r} \in \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$.

By our discussion of the geometric interpretation of the trace homomorphism $\mathrm{Tr}_\lambda^\iota$, the character $\overline{r} \in \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$ is the projection of $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$. In particular, r realizes the ideal triangulation λ. This concludes the proof of Proposition 16. □

4 Representations of the skein algebra

We are now ready to prove Theorem 2. We begin with an elementary lemma about the Chebyshev polynomials T_n. Remember that the polynomial T_n is defined by the property that $\operatorname{Tr} M^n = T_n(\operatorname{Tr} M)$ for every $M \in \mathrm{SL}_2(\mathbb{C})$. Applying this to a rotation matrix gives the trigonometric interpretation that $\cos n\theta = \frac{1}{2} T_n(2\cos\theta)$.

Lemma 17 (1) *If $x = a + a^{-1}$, then $T_n(x) = a^n + a^{-n}$.*

(2) *If $y = b + b^{-1}$, the set of solutions to the equation $T_n(x) = y$ consists of the numbers $x = a + a^{-1}$ as a ranges over all n^{th} roots of b.*

Proof For a matrix $M \in \mathrm{SL}_2(\mathbb{C})$, the data of its trace x is equivalent to the data of its spectrum $\{a, a^{-1}\}$. The first property is then a straightforward consequence of the fact that $\mathrm{Tr}\, M^n = T_n(\mathrm{Tr}\, M)$. The second property immediately follows. $\square$

We will also need the following quantum trace computation, which connects the skein $[P_k] \in \mathcal{S}^A(S)$ and the central element $H_k \in \mathcal{Z}^\omega(\lambda)$ that are associated to the same k^{th} puncture of S.

Lemma 18 *For the quantum trace homomorphism* $\mathrm{Tr}_\lambda^\omega \colon \mathcal{S}^A(S) \to \mathcal{Z}^\omega(\lambda)$,

$$\mathrm{Tr}_\lambda^\omega([P_k]) = H_k + H_k^{-1}.$$

Proof Let $e_{i_1}, e_{i_2}, \dots, e_{i_u}$ be the edges of λ that lead to the k^{th} puncture, indexed in counterclockwise order around the puncture; in particular, the e_{i_j} are not necessarily distinct.

The construction of $\mathrm{Tr}_\lambda^\omega([P_k])$ in [7] requires a careful control of elevations (namely $[0,1]$–coordinates) along the knot $P_k \subset S \times [0,1]$. Choose this knot so that it steadily goes up from e_{i_1} to e_{i_u}, and then sharply goes down to return to its starting point in e_{i_1}. In this setup, the formula of [7] yields

$$\mathrm{Tr}_\lambda^\omega([P_k]) = \omega^{-u+2} Z_{i_1} Z_{i_2} \cdots Z_{i_u} + \omega^{-u+2} Z_{i_1}^{-1} Z_{i_2}^{-1} \cdots Z_{i_u}^{-1}.$$

This is relatively straightforward when only one end of the edge e_{i_1} leads to the k^{th} puncture, namely when the projection of P_k to S crosses e_{i_1} only once, but otherwise requires the consideration of correction terms in a bigon neighborhood of e_{i_1}, of the type given by [7, Lemma 22]. Fortunately, these correction terms turn out to be trivial in this case.

We need to connect this formula to $H_k = [Z_{i_1} Z_{i_2} \cdots Z_{i_u}]$. Computing the Weyl quantum ordering is again straightforward when each edge e_{i_j} has only one end leading to the k^{th} puncture. For the general case, we could use a brute force computation as in [10, Lemma 12]. We prefer to give here a more indirect argument, based on the invariance of $\mathrm{Tr}_\lambda^\omega([P_k])$ under isotopy of P_k.

For this, choose the elevation of P_k so that it now goes *down* from e_{i_1} to e_{i_u}, and then goes up near e_{i_1} to return to its starting point. In this setup, the formulas of [7] give

$$\mathrm{Tr}_\lambda^\omega([P_k]) = \omega^{u-2} Z_{i_u} Z_{i_{u-1}} \cdots Z_{i_1} + \omega^{u-2} Z_{i_u}^{-1} Z_{i_{u-1}}^{-1} \cdots Z_{i_1}^{-1}.$$

Comparing the two expressions for $\mathrm{Tr}_\lambda^\omega([P_k]) \in \mathcal{Z}^\omega(\lambda)$ shows in particular that

$$\omega^{-u+2} Z_{i_1} Z_{i_2} \cdots Z_{i_u} = \omega^{u-2} Z_{i_u} Z_{i_{u-1}} \cdots Z_{i_1}.$$

By definition of the Weyl quantum ordering, there exists an integer $a \in \mathbb{Z}$ such that

$$H_k = \omega^a Z_{i_1} Z_{i_2} \cdots Z_{i_u} = \omega^{-a} Z_{i_u} Z_{i_{u-1}} \cdots Z_{i_1}.$$

We can then rephrase the above equality as $\omega^{-a-u+2} H_k = \omega^{a+u-2} H_k$. Although the current article usually focuses on the case where ω is a root of unity, these computations are valid for all ω. It follows that $a = -u + 2$. This proves that $H_k = \omega^{-u+2} Z_{i_1} Z_{i_2} \cdots Z_{i_u}$, and our first computation then shows that $\mathrm{Tr}^\omega_\lambda([P_k]) = H_k + H_k^{-1}$. □

We are now ready to prove Theorem 2, which we repeat here for convenience. Recall that a character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ associates a number $\mathrm{Tr}\, r(P_k)$ to the k^{th} puncture of S, where P_k is a small loop going around the puncture.

Theorem 19 *Assume that the surface S has at least one puncture, that its Euler characteristic is negative, that A is a primitive N^{th} root of -1 with N odd, and that we are given*

(1) *a character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ realizing some ideal triangulation λ of S;*

(2) *a number $p_k \in \mathbb{C}$ such that $T_N(p_k) = -\mathrm{Tr}\, r(P_k)$ for each puncture of S.*

Then there exists an irreducible finite-dimensional representation $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ whose classical shadow is equal to r and whose puncture invariants are the p_k.

Proof Since $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ realizes the ideal triangulation λ, Proposition 16 provides a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{C}^*$ and an associated representation $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathrm{End}(\mathbb{C}) = \mathbb{C}$, such that

$$\mu_\zeta \circ \mathrm{Tr}^\iota_\lambda([K]) = -\mathrm{Tr}\, r(K)$$

for every framed knot $K \subset S \times [0,1]$

By Lemma 18, the image of $[P_k] \in \mathcal{S}^A(S)$ under the quantum trace homomorphism $\mathrm{Tr}^\omega_\lambda\colon \mathcal{S}^A(S) \to \mathcal{Z}^\omega(\lambda)$ is equal to $\mathrm{Tr}^\omega_\lambda([P_k]) = H_k + H_k^{-1}$ in $\mathcal{Z}^\omega(\lambda)$. Similarly, $\mathrm{Tr}^\iota_\lambda([P_k]) = H_k + H_k^{-1}$ in $\mathcal{Z}^\iota(\lambda)$. (Beware that we are using the same symbols to denote the skeins $[P_k] \in \mathcal{S}^A(S)$ and $[P_k] \in \mathcal{S}^{-1}(S)$, and the central elements $H_k \in \mathcal{Z}^\omega(\lambda)$ and $H_k \in \mathcal{Z}^\iota(\lambda)$.) Then for $[P_k] \in \mathcal{S}^{-1}(S)$,

$$\mathrm{Tr}\, r(P_k) = -\mu_\zeta \circ \mathrm{Tr}^\iota_\lambda([P_k]) = -\mu_\zeta(H_k + H_k^{-1}) = -g_k - g_k^{-1}$$

if we set $g_k = \mu_\zeta(H_k) \in \mathrm{End}(\mathbb{C}) = \mathbb{C}$.

For each k, we are given a number $p_k \in \mathbb{C}$ such that $T_N(p_k) = -\mathrm{Tr}\, r(P_k) = g_k + g_k^{-1}$. Lemma 17 then provides an N^{th} root $h_k = \sqrt[N]{g_k}$ of such that $p_k = h_k + h_k^{-1}$.

Proposition 15 associates to the homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda; \mathbb{Z}) \to \mathbb{C}^*$ and to the N^{th} roots $h_k = \mu_\zeta(H_k)^{1/N}$ an irreducible representation $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ such that

(1) $\mu(Z_\alpha^N) = \zeta(\alpha)\,\mathrm{Id}_E$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$;

(2) $\mu(H_k) = h_k\,\mathrm{Id}_E$ for every $k = 1,\dots,s$.

Composing with the quantum trace map $\mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathcal{Z}^\omega(\lambda)$, we now define a representation

$$\rho = \mu \circ \mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathrm{End}(E).$$

To determine the classical shadow of ρ, let K be a framed knot whose projection to S has no crossing and whose framing is vertical. Then for the associated skein $[K] \in \mathcal{S}^A(S)$,

$$T_N(\rho([K])) = \rho(T_N([K])) = \mu\circ\mathrm{Tr}_\lambda^\omega(T_N([K])) = \mu\circ\mathrm{Tr}_\lambda^\omega\circ\boldsymbol{T}^A([K]) = \mu\circ\boldsymbol{F}^\omega\circ\mathrm{Tr}_\lambda^\iota([K])$$

by using the fact that ρ is an algebra homomorphism for the first equality, by definition of the Chebyshev homomorphism $\boldsymbol{T}^A\colon \mathcal{S}^{-1}(S) \to \mathcal{S}^A(S)$ in Section 1.3 for the third equality, and by the miraculous cancellations of Theorem 7 for the last relation. In terms of the Frobenius homomorphism $\boldsymbol{F}^\omega\colon \mathcal{T}^\iota(\lambda) \to \mathcal{T}^\omega(\lambda)$ introduced in Section 1.3 and of the representation $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathrm{End}(\mathbb{C}) = \mathbb{C}$, the property that $\mu(Z_\alpha^N) = \zeta(\alpha)\,\mathrm{Id}_E = \mu_\zeta(Z_\alpha)$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$ can be rephrased as $\mu \circ \boldsymbol{F}^\omega = \mu_\zeta$. Therefore,

$$T_N(\rho([K])) = \mu \circ \boldsymbol{F}^\omega \circ \mathrm{Tr}_\lambda^\iota([K]) = \mu_\zeta \circ \mathrm{Tr}_\lambda^\iota([K])\,\mathrm{Id}_E = -\mathrm{Tr}\,r(K)\,\mathrm{Id}_E\,.$$

Also, for the k^{th} puncture of S, the corresponding puncture invariant is determined by the property that

$$\rho([P_k]) = \mu \circ \mathrm{Tr}_\lambda^\omega([P_k]) = \mu(H_k + H_k^{-1}) = (h_k + h_k^{-1})\,\mathrm{Id}_E = p_k\,\mathrm{Id}_E\,.$$

If we knew that ρ was irreducible, we would be done with the proof of Theorem 19. At this point, there is no reason for this property to hold. However, if ρ is not irreducible, it suffices to consider an irreducible component $\rho'\colon \mathcal{S}^A(S) \to \mathrm{End}(F)$ with $F \subset E$. Restricting the above computations to F shows that the classical shadow of the representation ρ' is equal to the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, and that its puncture invariants are equal to the numbers p_k. □

Remark 20 We conjecture that, when r is sufficiently generic in $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, the representation $\rho = \mu \circ \mathrm{Tr}_\lambda^\omega$ used in the proof of Theorem 19 is already irreducible, and that there is no need to restrict to an irreducible factor. In earlier versions of this article we also conjectured that, for generic $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, there is a unique representation ρ satisfying the conclusions of Theorem 19 up to isomorphism. This second conjecture was recently proved by Frohman, Kania-Bartoszyńska and Lê [22]. See also Takenov [31] for an earlier proof of this second conjecture in the cases of the one-puncture torus and the four-puncture sphere (building on earlier work of Bullock and Przytycki [16] and Havlíček and Pošta [23] for the one-puncture torus).

Remark 21 In the very nongeneric case where $r(P_k)$ is the identity and where $p_k = -\omega^4 - \omega^{-4}$ for some punctures, the representation $\rho = \mu \circ \mathrm{Tr}_\lambda^\omega$ is definitely reducible. This is a key ingredient of the "puncture filling" process developed in [10].

5 A uniqueness property

We made choices in the proof of Theorem 19, and more precisely in its intermediate step the proof of Proposition 16. Indeed, when proving Proposition 16, we first took arbitrary square roots $z_i = \sqrt{x_i}$ for the shear-bend parameters $x_i \in \mathbb{C}^*$ of a pleated surface, and then adjusted these square roots in order to get the desired classical shadow for the representation $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$.

The goal of this section is to show that the output of the construction does not depend on these choices, provided we carefully specify our data and our conclusions. The resulting uniqueness statement will be used in the subsequent article [10]. Indeed, [10] heavily relies on Theorem 19 to construct representations of the skein algebra of a closed surface, by applying this statement to suitably chosen punctured surfaces.

5.1 Pleated surfaces and representations of $\mathcal{Z}^\iota(\lambda)$

The proof of Theorem 19 hinges on Proposition 16 which, given a character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, provides a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$ and its associated representation $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$ such that

$$\mu_\zeta \circ \mathrm{Tr}_\lambda^\iota([K]) = -\operatorname{Tr} r(K)$$

for every framed knot $K \subset S \times [0,1]$. Recall that μ_ζ and ζ are related by the property that $\mu_\zeta(Z_\alpha) = \zeta(\alpha) \in \mathbb{C}^*$ for every basis element $Z_\alpha \in \mathcal{Z}^\iota(\lambda)$ associated to an edge weight system $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$.

For most characters $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, the homomorphism $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$ is uniquely determined by r and by the pleated surface $(\widetilde{f}, \overline{r})$. However this uniqueness fails, in a very specific way, when the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ admits a very special type of internal symmetry which we now describe.

The cohomology group $H^1(S;\mathbb{Z}_2)$ acts on the character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ by the property that, for every homomorphism $r\colon \pi_1(S) \to \mathrm{SL}_2(\mathbb{C})$ and cohomology class $\varepsilon \in H^1(S;\mathbb{Z}_2)$, the homomorphism εr is defined by

$$\varepsilon r(\gamma) = (-1)^{\varepsilon(\gamma)} r(\gamma) \in \mathrm{SL}_2(\mathbb{C})$$

for every $\gamma \in \pi_1(S)$. We say that $\varepsilon \in H^1(S;\mathbb{Z}_2)$ is a *sign-reversal symmetry* for the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ if the action of ε on $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ fixes r. This is equivalent to the property that the trace $\operatorname{Tr} r(\gamma)$ is equal to 0 for every $\gamma \in \pi_1(S)$ with $\varepsilon(\gamma) \neq 0$.

The group $H^1(S;\mathbb{Z}_2)$ also acts on the balanced Chekhov–Fock algebra $\mathcal{Z}^\omega(\lambda)$ by the property that $\varepsilon Z_\alpha = (-1)^{\varepsilon([\alpha])} Z_\alpha$ for every $\varepsilon \in H^1(S;\mathbb{Z}_2)$ and every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$, where $[\alpha] \in H_1(S;\mathbb{Z}_2)$ is the homology class associated to the edge weight system α as in Lemma 9.

Proposition 22 *Suppose the pleated surface $(\widetilde{f},\bar{r})$ has pleating locus the ideal triangulation λ, and let $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ be represented by a group homomorphism $r\colon \pi_1(S) \to \mathrm{SL}_2(\mathbb{C})$ lifting the monodromy $\bar{r}\colon \pi_1(S) \to \mathrm{PSL}_2(\mathbb{C})$ of $(\widetilde{f},\bar{r})$. Then there exists an algebra homomorphism $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$, associated to a twisted homomorphism $\zeta\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$, such that*

(1) *for each edge e_i of λ, we have that $\mu_\zeta(Z_i^2)$ is equal to the shear-bend parameter $x_i \in \mathbb{C}^*$ of e_i in the pleated surface $(\widetilde{f},\bar{r})$;*

(2) $\mu_\zeta \circ \operatorname{Tr}_\lambda^\iota([K]) = -\operatorname{Tr} r(K)$ *for every framed knot* $K \subset S \times [0,1]$.

In addition, μ_ζ is unique up to the action on $\mathcal{Z}^\iota(\lambda)$ of a sign-reversal symmetry $\varepsilon \in H^1(S;\mathbb{Z}_2)$ of the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$.

We say that a homomorphism $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$ satisfying the above conclusions is *compatible* with the pleated surface $(\widetilde{f},\bar{r})$ and the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$.

Proof of Proposition 22 The existence is provided by Proposition 16, or more precisely by its proof to guarantee that $\mu_\zeta(Z_i^2) = x_i$ for every edge e_i of λ.

To prove the uniqueness, suppose that we are given another algebra homomorphism $\mu_{\zeta'}\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$ satisfying the same conclusions, and that this homomorphism is associated to a twisted homomorphism $\zeta'\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{C}^*$. From the property that $\mu_\zeta(Z_i^2) = \mu_{\zeta'}(Z_i^2) = x_i$, we conclude that $\mu_\zeta(Z_\alpha)^2 = \mu_{\zeta'}(Z_\alpha)^2$ and therefore $\mu_\zeta(Z_\alpha) = \pm\mu_{\zeta'}(Z_\alpha)$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$. Since μ_ζ and $\mu_{\zeta'}$ are both algebra homomorphisms, there consequently exists a group homomorphism $\varepsilon\colon \mathcal{W}(\tau_\lambda;\mathbb{Z}) \to \mathbb{Z}_2$ such that $\mu_\zeta(Z_\alpha) = (-1)^{\varepsilon(\alpha)}\mu_{\zeta'}(Z_\alpha)$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$. Another application of the property that $\mu_\zeta(Z_i^2) = \mu_{\zeta'}(Z_i^2)$ shows that ε is trivial on the subgroup $(2\mathbb{Z})^n \subset \mathcal{W}(\tau_\lambda;\mathbb{Z})$ of Lemma 9. This statement then shows that ε comes from a homomorphism $H_1(S;\mathbb{Z}_2) \to \mathbb{Z}_2$, and can therefore be interpreted as a cohomology class $\varepsilon \in H^1(S;\mathbb{Z}_2)$.

In this cohomological interpretation of $\varepsilon \in H^1(S;\mathbb{Z}_2)$, we have that $\mu_\zeta(Z_\alpha) = (-1)^{\varepsilon([\alpha])}\mu_{\zeta'}(Z_\alpha)$ for every $\alpha \in \mathcal{W}(\tau_\lambda;\mathbb{Z})$. Namely, the homomorphisms μ_ζ and $\mu_{\zeta'}\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$ differ by the action of $\varepsilon \in H^1(S;\mathbb{Z}_2)$ on $\mathcal{Z}^\iota(\lambda)$.

Given a framed link $K \subset S \times [0,1]$, the construction of the quantum trace $\mathrm{Tr}^{\iota}_{\lambda}$ in [7] shows that $\mathrm{Tr}^{\iota}_{\lambda}([K]) \in \mathcal{Z}^{\iota}(\lambda)$ is a linear combination of monomials Z_α whose associated homology class $[\alpha] \in H_1(S;\mathbb{Z}_2)$, in the sense of Lemma 9, is the same as the class $[K] \in H_1(S;\mathbb{Z}_2)$ defined by K. As a consequence,

$$\mathrm{Tr}\, r(K) = -\mu_{\xi'} \circ \mathrm{Tr}^{\iota}_{\lambda}([K]) = -(-1)^{\varepsilon(K)} \mu_{\xi} \circ \mathrm{Tr}^{\iota}_{\lambda}([K]) = (-1)^{\varepsilon(K)}\, \mathrm{Tr}\, r(K)$$

for every framed link $K \subset S \times [0,1]$. This proves that $\varepsilon \in H^1(S;\mathbb{Z}_2)$ is a sign-reversal symmetry for the character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$.

As a consequence, the homomorphisms $\mu_\xi, \mu_{\xi'} \colon \mathcal{Z}^{\iota}(\lambda) \to \mathbb{C}$ differ by the action on $\mathcal{Z}^{\iota}(\lambda)$ of a sign-reversal symmetry $\varepsilon \in H^1(S;\mathbb{Z}_2)$ of $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$. □

Characters with nontrivial sign-reversal symmetries exist, but are rare. In fact, the characters that have no (nontrivial) sign-reversal symmetries form a Zariski dense closed subset in $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$. (Hint: Choose a family of simple closed curves $\gamma_1, \gamma_2, \dots, \gamma_k$ in S that generate $H_1(S;\mathbb{Z}_2)$, and consider the set of $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ such that $\mathrm{Tr}\, r(\gamma_i) \neq 0$ for some i.) This subset of $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ includes all injective homomorphisms $\pi_1(S) \to \mathrm{SL}_2(\mathbb{C})$, since their images contain no matrix with trace 0. In particular all "geometric" characters, corresponding to fuchsian or quasifuchsian groups, admit no sign-reversal symmetries.

More precisely, a simple algebraic manipulation shows that every character $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ with a nontrivial sign-reversal symmetry $\varepsilon \in H^1(S;\mathbb{Z}_2)$ is represented by a homomorphism $r\colon \pi_1(S) \to \mathrm{SL}_2(\mathbb{C})$ of the following type: Considering ε as a group homomorphism $\varepsilon\colon \pi_1(S) \to \mathbb{Z}_2$ and for an arbitrary $\gamma_0 \in \pi_1(S)$ with $\varepsilon(\gamma_0) \neq 0$, there exists a group homomorphism $\theta\colon \ker\varepsilon \to \mathbb{C}/2\pi\mathrm{i}\mathbb{Z}$ such that

$$r(\gamma_0) = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} \quad \text{and} \quad r(\gamma) = \begin{pmatrix} \cosh\theta(\gamma) & \sinh\theta(\gamma) \\ \sinh\theta(\gamma) & \cosh\theta(\gamma) \end{pmatrix} \quad \text{for every } \gamma \in \ker\varepsilon.$$

In particular, noting the constraints that $\theta(\gamma_0^2) = \pi\mathrm{i}$ and $\theta(\gamma_0\gamma\gamma_0^{-1}) = -\theta(\gamma)$ for every $\gamma \in \ker\varepsilon$, the space of such characters has complex dimension $2g+s-2$ in the $(6g+3s-6)$–dimensional character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ (where g is the genus of the surface S and s is its number of punctures).

5.2 A strengthening of Theorem 19

Recall that, if the k^{th} puncture of S is adjacent to the edges $e_{i_1}, e_{i_2}, \dots, e_{i_u}$ of the ideal triangulation λ, it determines an element $H_k = [Z_{i_1} Z_{i_2} \cdots Z_{i_u}] \in \mathcal{Z}^{\iota}(\lambda)$.

Proposition 23 *Assume that the surface S has at least one puncture, that its Euler characteristic is negative, that A is a primitive N^{th} root of -1 with N odd, and that we are given*

(i) *a pleated surface* $(\widetilde{f}, \overline{r})$ *with pleating locus* λ, *a character* $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ *lifting* $\overline{r} \in \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$, *and an algebra homomorphism* $\mu_\zeta\colon \mathcal{Z}^\iota(\lambda) \to \mathbb{C}$ *compatible with* $(\widetilde{f}, \overline{r})$ *and* r *as in Proposition 22;*

(ii) *for each puncture of* S, *an* N^{th} *root* h_k *of* $\mu_\zeta(H_k) \in \mathbb{C}^*$.

Then, up to isomorphism, there exists a unique representation $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ *of the balanced Chekhov–Fock algebra* $\mathcal{Z}^\omega(\lambda)$ *with the following properties:*

(1) *the dimension of the vector space* E *is equal to* N^{3g+s-3}, *where* g *is the genus of the surface* S *and* s *its number of punctures;*

(2) $\mu(Z_\alpha^N) = \mu_\zeta(Z_\alpha)$ *for every edge weight system* $\alpha \in \mathcal{W}(\tau_\lambda; \mathbb{Z})$, *where we use the same symbol to represent the associated base elements* $Z_\alpha \in \mathcal{Z}^\omega(\lambda)$ *and* $Z_\alpha \in \mathcal{Z}^\iota(\lambda)$;

(3) $\mu(H_k) = h_k \,\mathrm{Id}_E$ *for the central element* $H_k \in \mathcal{Z}^\omega(\lambda)$ *associated to the* k^{th} *puncture of* S.

Also, the representation μ *is irreducible and the representation* $\rho = \mu \circ \mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ *has classical shadow* $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$, *in the sense that*

$$T_N(\rho([K])) = -\operatorname{Tr} r(K)\,\mathrm{Id}_E$$

for every knot $K \subset S \times [0,1]$ *whose projection to* S *has no crossing and whose framing is vertical* (*where* $T_N(x)$ *is the* N^{th} *Chebyshev polynomial of the first type*).

Proof The existence and uniqueness part is essentially a restatement of the classification of irreducible representations of $\mathcal{Z}^\omega(\lambda)$ in Proposition 15. The fact that ρ has classical shadow r follows from the proof of Theorem 19. □

Although the representation $\mu\colon \mathcal{Z}^\omega(\lambda) \to \mathrm{End}(E)$ of Proposition 23 is irreducible, the representation $\rho = \mu \circ \mathrm{Tr}_\lambda^\omega\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ is not necessarily irreducible; see Remark 21.

6 The case where $A^N = +1$

The case where $A^N = +1$ can be deduced from the case where $A^N = -1$ by the *Barrett isomorphism* $B_\sigma\colon \mathcal{S}^A(S) \to \mathcal{S}^{-A}(S)$ associated to a spin structure σ on the surface S. This isomorphism is defined by the property that, for every framed link $K \subset S \times [0,1]$ with k components,

$$B_\sigma([K]) = (-1)^{k+\sigma(K)}[K] \in \mathcal{S}^{-A},$$

where $\sigma(K) \in \mathbb{Z}_2$ is the monodromy of the framing of K with respect to σ. See [2] and [29, Section 2] for a proof that $B_\sigma\colon \mathcal{S}^A(S) \to \mathcal{S}^{-A}(S)$ is an algebra isomorphism.

If $A^N = +1$, an irreducible finite-dimensional representation $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ defines an irreducible representation $\rho' = \rho \circ B_\sigma\colon \mathcal{S}^{-A}(S) \to \mathrm{End}(E)$, to which we can apply Theorems 1 and 2 since $(-A)^N = -1$ as N is assumed to be odd. This process depends on the choice of a spin structure σ, but we can make it more canonical by the following construction.

Let $\mathrm{Spin}(S)$ denote the set of isotopy classes of spin structures on S. Any two spin structures differ by an obstruction in $H^1(S;\mathbb{Z}_2)$, which defines an action of $H^1(S;\mathbb{Z}_2)$ on $\mathrm{Spin}(S)$. The cohomology group $H^1(S;\mathbb{Z}_2)$ also acts on the character variety $\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ by the property that, if $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ and $\varepsilon \in H^1(S;\mathbb{Z}_2)$, then $\varepsilon r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ is defined by

$$\varepsilon r(\gamma) = (-1)^{\varepsilon(\gamma)} r(\gamma) \in \mathrm{SL}_2(\mathbb{C})$$

for every $\gamma \in \pi_1(S)$.

The *twisted character variety* $\mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ is then defined as the quotient

$$\mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})} = (\mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) \times \mathrm{Spin}(S))/H^1(S;\mathbb{Z}_2).$$

If the twisted character $\hat{r} \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ is represented by $(r,\sigma) \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) \times \mathrm{Spin}(S)$ and if K is a framed knot in $S \times [0,1]$, the definition is designed so that the *trace*

$$\mathrm{Tr}\,\hat{r}(K) = -(-1)^{\sigma(K)}\,\mathrm{Tr}\,r(K)$$

depends only on the twisted character $\hat{r} \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$, and not on its representative $(r,\sigma) \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) \times \mathrm{Spin}(S)$.

The correspondence $\rho \leftrightarrow \rho \circ B_\sigma$ is used in [9] to establish the following result.

Theorem 24 [9] *Suppose that A is a primitive N^{th} root of $+1$ with N odd, and let $\rho\colon \mathcal{S}^A(S) \to \mathrm{End}(E)$ be an irreducible finite-dimensional representation of the Kauffman bracket skein algebra. Let $T_N(x)$ be the N^{th} normalized Chebyshev polynomial of the first kind.*

(1) *There exists a unique twisted character $\hat{r}_\rho \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ such that*

$$T_N(\rho([K])) = -(\mathrm{Tr}\,\hat{r}_\rho(K))\,\mathrm{Id}_E$$

for every framed knot $K \subset S \times [0,1]$ whose projection to S has no crossing and whose framing is vertical.

(2) *Let P_k be a small simple loop going around the k^{th} puncture of S, and consider it as a knot in $S \times [0,1]$ with vertical framing. Then there exists a number $p_k \in \mathbb{C}$ such that $\rho([P_k]) = p_k \operatorname{Id}_E$.*

(3) *The number p_k of (2) is related to the twisted character $\widehat{r}_\rho \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ of (1) by the property that $T_N(p_k) = -\operatorname{Tr} \widehat{r}_\rho(P_k)$.* □

The same correspondence $\rho \leftrightarrow \rho \circ B_\sigma$ can be used to prove the following analogue of Theorem 2. We say that the twisted character $\widehat{r} \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ *realizes* the ideal triangulation λ of S if the image $\overline{r} \in \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$ of $\widehat{r}$ under the natural projection $\mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})} \to \mathcal{R}_{\mathrm{PSL}_2(\mathbb{C})}(S)$ realizes λ in the sense of Section 3.

Theorem 25 *Assume that the surface S has at least one puncture, that its Euler characteristic is negative, that A is a primitive N^{th} root of $+1$ with N odd, and that we are given*

(1) *a twisted character $\widehat{r} \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ which realizes some ideal triangulation λ of S;*

(2) *a number $p_k \in \mathbb{C}$ such that $T_N(p_k) = -\operatorname{Tr} \widehat{r}(P_k)$ for each of the punctures of S.*

Then there exists an irreducible finite-dimensional representation $\rho\colon \mathcal{S}^A(S) \to \operatorname{End}(E)$ whose classical shadow is equal to $\widehat{r}$ and whose puncture invariants are the p_k.

Proof Represent $\widehat{r} \in \mathcal{R}^{\mathrm{Spin}(S)}_{\mathrm{PSL}_2(\mathbb{C})}$ by a pair $(r, \sigma) \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S) \times \mathrm{Spin}(S)$. Theorem 2 provides an irreducible representation $\rho'\colon \mathcal{S}^{-A}(S) \to \operatorname{End}(E)$ with classical shadow $r \in \mathcal{R}_{\mathrm{SL}_2(\mathbb{C})}(S)$ and puncture invariants equal to the p_k. Then $\rho = \rho' \circ B_\sigma\colon \mathcal{S}^A(S) \to \operatorname{End}(E)$ satisfies the required properties. □

Appendix: The Thurston intersection form of a train track

Let τ be a train track in an oriented surface S, and let $\mathcal{W}(\tau;\mathbb{Z})$ be the space of integer-valued edge weights for τ. Namely, an element $\alpha \in \mathcal{W}(\tau;\mathbb{Z})$ assigns a weight $\alpha(e) \in \mathbb{Z}$ to each edge e of τ in such a way that, at each switch s of τ, the sum of the weights of the edges of τ coming in on one side of s is equal to the sum of the weights of the edges going out on the other side. This abelian group comes with an additional structure provided by the Thurston intersection form

$$\Omega\colon \mathcal{W}(\tau;\mathbb{Z}) \times \mathcal{W}(\tau;\mathbb{Z}) \to \mathbb{Z}$$

defined as in Section 2.2. Namely,

$$\Omega(\alpha,\beta) = \frac{1}{2} \sum_{e \text{ right of } e'} (\alpha(e)\beta(e') - \alpha(e')\beta(e)),$$

where the sum is over all pairs (e, e') such that e and e' are two "germs of edges" emerging on the same side of a switch of τ with e to the right of e' (note e and e' are not necessarily adjacent at that switch) for the orientation of S. At this point, $\Omega(\alpha, \beta)$ is only a half-integer, but Theorem 26 below will prove that it is indeed an integer.

We want to determine the algebraic structure of $\mathcal{W}(\tau; \mathbb{Z})$ endowed with Ω. This is a classical property in the case of real-valued edge weights (see for instance [28, Section 3.2] or [4, Section 3]), but the subtleties of the integer-valued case seem less well known. The result is of independent interest because, beyond the scope of this article, integer-valued edge weight do occur in geometric situations where the Thurston intersection form is also relevant. One such instance arises for general pleated surfaces where the pleating locus is allowed to have uncountably many leaves, as opposed to the simpler pleated surfaces considered in Section 3. The bending of such a pleated surface is measured by an edge weight system valued in $\mathbb{R}/2\pi\mathbb{Z}$ for a train track carrying the pleating locus, and this edge weight system is related to rotation numbers by the Thurston intersection form [4].

The complement $S - \tau$ of the train track τ admits a certain number of "spikes", each locally delimited by two edges of τ that approach the same side of a switch of τ. Thicken τ to a subsurface $U \subset S$ that deformation retracts to τ. Each component of $U - \tau$ is then an annulus that contains one component of ∂U and a certain number of spikes of $S - \tau$. We can then consider the genus h of U, and the number n_{even} (resp. n_{odd}) of components of $U - \tau$ that contain an even (resp. odd) number of spikes.

A component U_1 of $U - \tau$ that contains an even number $n_1 > 0$ of spikes of $S - \tau$ determines, up to sign, an element of $\mathcal{W}(\tau; \mathbb{Z})$ as follows. The core of U_1 is homotopic to a closed curve γ_1 in τ that is made up of arcs $k_1, k_2, \dots, k_{n_1}, k_{n_1+1} = k_1$, in this order, such that each arc k_1 is immersed in τ and such that two consecutive arcs k_i and k_{i+1} locally bound a spike of U_1 at their common end point. For each edge e of τ, we can then consider

$$\alpha(e) = \sum_{i=1}^{n_1} (-1)^i \alpha_i(e) \in \mathbb{Z},$$

where $\alpha_i(e) \in \{0, 1, 2\}$ is the number of times the arc k_i passes over the edge e. Because the signs $(-1)^i$ alternate at the spikes of U_1 (using the fact that n_1 is even for $i = n_1$), one easily sees that these edge weights $\alpha(e)$ satisfy the switch conditions, and therefore define an edge weight system $\alpha \in \mathcal{W}(\tau; \mathbb{Z})$.

A component U_1 of $U - \tau$ that contains no spike similarly determines an edge weight system $\alpha \in \mathcal{W}(\tau; \mathbb{Z})$. The core of U_1 is now homotopic to a closed curve γ_1 immersed in τ, and α associates to each edge e the number $\alpha(e)$ of times γ_1 passes over e.

Also, recall that the train track τ is *orientable* if its edges can be oriented in such a way that the orientations match at the switches of τ.

Theorem 26 *For a connected train track τ in the surface S, let the numbers h, n_{even} and n_{odd} be defined as above. Then the lattice $\mathcal{W}(\tau;\mathbb{Z})$ of integer-valued edge weight systems for τ admits a basis in which the Thurston intersection form Ω is block diagonal with*

- *h blocks $\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$, $\frac{1}{2}n_{\mathrm{odd}}-1$ blocks $\left(\begin{smallmatrix}0&2\\-2&0\end{smallmatrix}\right)$, and n_{even} blocks (0) if $n_{\mathrm{odd}}>0$;*
- *h blocks $\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$ and n_{even} blocks (0) if $n_{\mathrm{odd}}=0$ and τ is nonorientable;*
- *h blocks $\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$ and $n_{\mathrm{even}}-1$ blocks (0) if $n_{\mathrm{odd}}=0$ and τ is orientable.*

In addition, in all cases, we can choose the base elements corresponding to the blocks (0) to be the edge weight systems associated as above to the components of $U-\tau$ that contain an even number of spikes.

In particular, n_{odd} is always even.

Proof We will subdivide the proof into several lemmas. The reader may recognize many analogies with the arguments used in the proof of [5, Proposition 5].

We first discuss a classical homological interpretation of the elements of $\mathcal{W}(\tau;\mathbb{Z})$ and of the Thurston intersection form Ω.

Because the edges of τ are not oriented, an edge weight system does not directly define a homology class in $H_1(\tau;\mathbb{Z})$. Instead consider the 2–fold orientation covering $\widehat{\tau}$ of τ, consisting of all pairs (x,o) where $x\in\tau$ and o is a local orientation of the train track τ at x. Note that $\widehat{\tau}$ is a canonically oriented train track, and that the covering involution $\sigma\colon\widehat{\tau}\to\widehat{\tau}$ that exchanges the two sheets of the covering reverses the orientation of $\widehat{\tau}$.

An edge weight system $\alpha\in\mathcal{W}(\tau;\mathbb{Z})$ lifts to a weight system $\widehat{\alpha}\in\mathcal{W}(\widehat{\tau};\mathbb{Z})$. Endowing each (oriented) edge of $\widehat{\tau}$ with the weight assigned by $\widehat{\alpha}$ defines a chain, which is closed because of the switch condition and therefore defines a homology class $[\widehat{\alpha}]\in H_1(\widehat{\tau};\mathbb{Z})$. Note that $\sigma_*([\widehat{\alpha}])=-[\widehat{\alpha}]$ since the covering involution σ reverses the canonical orientation of $\widehat{\tau}$.

Conversely, each homology class $[\widehat{\alpha}]\in H_1(\widehat{\tau};\mathbb{Z})$ is represented by a unique linear combination of the edges of $\widehat{\tau}$, and therefore determines an edge weight system $\widehat{\alpha}\in\mathcal{W}(\widehat{\tau};\mathbb{Z})$. Assuming in addition that $\sigma_*([\widehat{\alpha}])=-[\widehat{\alpha}]$, this edge weight system is invariant under the action of σ, and therefore comes from an edge weight system $\alpha\in\mathcal{W}(\tau;\mathbb{Z})$. This proves:

Lemma 27 *The above correspondence identifies the space $\mathcal{W}(\tau;\mathbb{Z})$ of edge weight systems to the eigenspace*

$$H_1(\hat\tau;\mathbb{Z})^- = \bigl\{[\hat\alpha]\in H_1(\hat\tau;\mathbb{Z}) : \sigma_*([\hat\alpha]) = -[\hat\alpha]\bigr\} \subset H_1(\hat\tau;\mathbb{Z})$$

of the homomorphism $\sigma_*\colon H_1(\hat\tau;\mathbb{Z}) \to H_1(\hat\tau;\mathbb{Z})$ that is induced by the covering involution σ. □

To describe the Thurston intersection form in this homological framework, consider the subsurface U deformation retracting to τ. The covering $\hat\tau\to\tau$ uniquely extends to a 2–fold covering $\widehat U\to U$, whose covering involution $\sigma\colon \widehat U\to\widehat U$ extends our previous involution σ. The orientation of $U\subset S$ lifts to an orientation of $\widehat U$.

Lemma 28 *If $[\hat\alpha],[\hat\beta]\in H_1(\hat\tau)^-$ are associated to the edge weight systems $\alpha,\beta\in\mathcal{W}(\tau_\lambda;\mathbb{Z})$,*

$$\Omega(\alpha,\beta) = \tfrac{1}{2}\,[\hat\alpha]\cdot[\hat\beta],$$

where $\cdot$ denotes the algebraic intersection number of classes of $H_1(\widehat U;\mathbb{Z})\cong H_1(\hat\tau;\mathbb{Z})$. In addition, $[\hat\alpha]\cdot[\hat\beta]$ is even, and $\Omega(\alpha,\beta)$ is an integer.

Proof To prove the first statement push the oriented train track $\hat\tau$ to its left to obtain a train track $\hat\tau'\subset\widehat U$ that is transverse to $\hat\tau$, realize the homology class $[\hat\alpha]$ by $\hat\tau$ endowed with the edge multiplicities coming from α, realize $[\hat\beta]$ by $\hat\tau'$ endowed with the edge multiplicities coming from β, and use this setup to compute the algebraic intersection number $[\hat\alpha]\cdot[\hat\beta]$. Evaluating the contribution to $[\hat\alpha]\cdot[\hat\beta]$ of each point of $\hat\tau\cap\hat\tau'$ then shows that this algebraic intersection number is equal to $2\Omega(\alpha,\beta)$.

The second statement is obtained by a similar but different computation of $[\hat\alpha]\cdot[\hat\beta]$. Perturb τ to a train track τ'' that is transverse to τ, and let $\hat\tau''$ be the preimage of τ'' in $\widehat U$. Now compute $[\hat\alpha]\cdot[\hat\beta]$ by realizing the homology class $[\hat\beta]$ by $\hat\tau''$ endowed with the edge multiplicities coming from β, while still realizing $[\hat\alpha]$ by $\hat\tau$ endowed with the edge multiplicities coming from α. The intersection $\hat\tau\cap\hat\tau''$ splits into pairs of points interchanged by the covering involution σ, and the two points in each pair have the same contribution to $[\hat\alpha]\cdot[\hat\beta]$. It follows that $[\hat\alpha]\cdot[\hat\beta]$ is even. □

We now need to better understand the action of σ_* on the homology group $H_1(\widehat U;\mathbb{Z})$.

It will be convenient to systematically use a notation which already appeared in Lemma 27. If V is a space where some restriction of the covering involution σ induces a homomorphism σ_*, then

$$V^- = \{\alpha\in V : \sigma_*(\alpha) = -\alpha\}.$$

For instance, Lemma 27 provides a natural isomorphism $\mathcal{W}(\tau;\mathbb{Z})\cong H_1(\widehat U;\mathbb{Z})^-$.

Let $\partial_{\mathrm{even}}U$ be the union of the n_{even} components of ∂U that are adjacent to an even number of spikes of $S-\tau$, and set $\partial_{\mathrm{odd}}U = \partial U - \partial_{\mathrm{even}}U$.

Lemma 29 *Let γ_1 be a component of $\partial_{\mathrm{even}}U$, and let $\widehat{\gamma}_1$ be its preimage in $\widehat{U}$. Then $H_1(\widehat{\gamma}_1;\mathbb{Z})^-$ is isomorphic to $\mathbb{Z}$. In addition, the image in $H_1(\widehat{U};\mathbb{Z})^- \cong \mathcal{W}(\tau;\mathbb{Z})$ of one of the generators of $H_1(\widehat{\gamma}_1;\mathbb{Z})^-$ coincides up to sign with the edge weight system that, right before Theorem 26, we associated to the component U_1 of $U-\tau$ that contains γ_1.*

Proof As right above Theorem 26, the curve γ_1 is homotopic to a closed curve γ_1' in τ that is made up of n_1 arcs $k_1, k_2, \dots, k_{n_1}, k_{n_1+1} = k_1$, in this order, such that each arc k_1 is immersed in τ and such that two consecutive arcs k_i and k_{i+1} locally bound a spike of U_1 at their common end point. Because n_1 is even, there are two possible ways to orient these arcs in such a way that consecutive arcs have opposite orientations. This shows that γ_1' has two distinct lifts to $\widehat{\tau}$, and therefore that the preimage $\widehat{\gamma}_1$ of γ_1 in $\widehat{U}$ consists of two components of $\partial\widehat{U}$ that are exchanged by the covering involution. This provides an isomorphism $H_1(\widehat{\gamma}_1;\mathbb{Z}) \cong \mathbb{Z}\oplus\mathbb{Z}$ where σ_* exchanges the two factors. It immediately follows that $H_1(\widehat{\gamma}_1;\mathbb{Z})^- \cong \mathbb{Z}$.

If $\widehat{\gamma}_1' \subset \widehat{\tau}$ denotes one of the two lifts of γ_1' to $\widehat{\tau}$, the image of $H_1(\widehat{\gamma}_1;\mathbb{Z})^-$ in $H_1(\widehat{U};\mathbb{Z})^- \cong H_1(\widehat{\tau};\mathbb{Z})^- \cong \mathcal{W}(\tau;\mathbb{Z})$ is generated by $[\widehat{\gamma}_1'] - \sigma_*([\widehat{\gamma}_1'])$. The second statement easily follows. □

To prove Theorem 26, we will first restrict attention to the case where $n_{\mathrm{odd}} > 0$. This is equivalent to the property that $\partial_{\mathrm{odd}}U$ is nonempty.

We just saw that the restriction of the covering $\widehat{U} \to U$ above $\partial_{\mathrm{even}}U$ is trivial; similarly, its restriction above each component of $\partial_{\mathrm{odd}}U$ is nontrivial. Therefore, the covering $\widehat{U} \to U$ is classified by a cohomology class in $H^1(U;\mathbb{Z}_2)$ which evaluates to 0 on the elements of $\partial_{\mathrm{even}}U$ and to 1 on the components of $\partial_{\mathrm{odd}}U$.

Since the subset $\partial_{\mathrm{odd}}U$ is nonempty, and can therefore realize the cohomology class classifying the covering $\widehat{U} \to U$ as the Poincaré dual of a family $K \subset U$ of disjoint arcs whose boundary $\partial K = K \cap \partial U$ consists of one point in each component of $\partial_{\mathrm{odd}}U$.

Split U along a separating simple closed curve γ to isolate K inside of a planar surface $U_1 \subset S$ with boundary $\partial U_1 = \gamma \cup \partial_{\mathrm{odd}}U$, while the closure U_2 of $U - U_1$ has genus h and boundary $\partial U_2 = \gamma \cup \partial_{\mathrm{even}}U$. Let $\widehat{U}_1$ and $\widehat{U}_2$ be the respective preimages of U_1 and U_2 in $\widehat{U}$.

Since K is disjoint from U_2, the covering $\widehat{U}_2 \to U_2$ is trivial, and $\widehat{U}_2$ consists of two disjoint copies of the surface U_2 which are exchanged by σ.

The covering $\widehat{U}_1 \to U_1$ is nontrivial above each component of $\partial_{\mathrm{odd}} U$ and trivial above γ. Since the surface U_1 is planar, an Euler characteristic computation shows that $\widehat{U}_1$ has genus $\frac{1}{2}n_{\mathrm{odd}} - 1$ and has $n_{\mathrm{odd}} + 2$ boundary components.

Consider the Mayer–Vietoris exact sequence

$$0 \to H_1(\hat{\gamma}; \mathbb{Z}) \to H_1(\widehat{U}_1; \mathbb{Z}) \oplus H_1(\widehat{U}_2; \mathbb{Z}) \to H_1(\widehat{U}; \mathbb{Z}) \to 0,$$

where $\hat{\gamma}$ denotes the preimage of γ in $\widehat{U}$. (To explain the 0 on the right, note that the map $H_0(\hat{\gamma}; \mathbb{Z}) \to H_0(\widehat{U}_2; \mathbb{Z})$ is injective.)

Lemma 30 *Remembering that V^- denotes the (-1)–eigenspace of the action of σ_* over a space V, the above exact sequence induces another exact sequence*

$$0 \to H_1(\hat{\gamma}; \mathbb{Z})^- \to H_1(\widehat{U}_1; \mathbb{Z})^- \oplus H_1(\widehat{U}_2; \mathbb{Z})^- \to H_1(\widehat{U}; \mathbb{Z})^- \to 0.$$

Proof The only point that requires some thought is the fact that the third homomorphism is surjective.

Given $u \in H_1(\widehat{U}; \mathbb{Z})^-$, the first exact sequence provides $u_1 \in H_1(\widehat{U}_1; \mathbb{Z})$ and $u_2 \in H_1(\widehat{U}_2; \mathbb{Z})$ such that $u = u_1 + u_2$ in $H_1(\widehat{U}; \mathbb{Z})$. Since $\sigma_*(u) = -u$, we conclude that there exists $v \in H_1(\hat{\gamma}; \mathbb{Z})$ such that $\sigma_*(u_1) = -u_1 + v$ in $H_1(\widehat{U}_1; \mathbb{Z})$ and $\sigma_*(u_2) = -u_2 - v$ in $H_1(\widehat{U}_2; \mathbb{Z})$. Note that $v \in H_1(\hat{\gamma}; \mathbb{Z})$ is invariant under σ_*. Therefore, for the isomorphism $H_1(\hat{\gamma}; \mathbb{Z}) \cong H_1(\gamma; \mathbb{Z}) \oplus H_1(\gamma; \mathbb{Z})$ coming from the fact that each of the two components of $\hat{\gamma}$ is naturally identified to γ, $v = (w, w)$ for some $w \in H_1(\gamma; \mathbb{Z})$. If we replace u_1 by $u_1' = u_1 - (w, 0)$ and u_2 by $u_2' = u_2 + (w, 0)$, we now have that $u = u_1' + u_2'$ with $\sigma_*(u_1') = -u_1'$ and $\sigma_*(u_2') = -u_2'$, as requested. □

We now analyze the terms of the exact sequence of Lemma 30.

The space $H_1(\widehat{U}_2; \mathbb{Z})^-$ is easy to understand, because $\widehat{U}_2$ is made up of two disjoint copies of U_2, which are exchanged by the covering involution σ. Therefore, $H_1(\widehat{U}_2; \mathbb{Z}) \cong H_1(U_2; \mathbb{Z}) \oplus H_1(U_2; \mathbb{Z})$ and, for this isomorphism, $H_1(\widehat{U}_2; \mathbb{Z})^-$ corresponds to $\{(\alpha, -\alpha) : \alpha \in H_1(U_2; \mathbb{Z})\}$. This defines an isomorphism $H_1(\widehat{U}_2; \mathbb{Z})^- \cong H_1(U_2; \mathbb{Z})$, for which the intersection form of $H_1(\widehat{U}_2; \mathbb{Z})^-$ corresponds to twice the intersection form of $H_1(U_2; \mathbb{Z})$.

Lemma 31 *There exists a basis for $H_1(\widehat{U}_2; \mathbb{Z})^-$ in which the intersection form is block diagonal with h blocks $\left(\begin{smallmatrix} 0 & -2 \\ 2 & 0 \end{smallmatrix}\right)$ and n_{even} blocks (0).*

In addition, we can arrange that the basis elements corresponding to the blocks (0) are the images of generators of $H(\hat{\alpha}; \mathbb{Z})^- \cong \mathbb{Z}$ as $\hat{\alpha} \subset \widehat{U}_2$ ranges over all preimages of components α of $\partial_{\mathrm{even}} U$, and that a generator of $H_1(\hat{\gamma}; \mathbb{Z})^- \cong \mathbb{Z}$ is sent to the sum of these elements.

Proof The surface U_2 has genus h and has $n_{\mathrm{even}}+1$ boundary components, and γ is one of these boundary components. We can therefore find a basis for $H_1(U_2;\mathbb{Z})$ in which the intersection form is block diagonal with h blocks $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ and n_{even} blocks (0). In addition, since $\partial U_2 = \gamma \cup \partial_{\mathrm{even}} U$, we can arrange that the basis elements corresponding to the blocks (0) are the images of generators of $H_1(\alpha;\mathbb{Z})$ as α ranges over all components of $\partial_{\mathrm{even}} U$, while the image of a generator of $H_1(\gamma;\mathbb{Z})$ is sent to the sum of these elements.

The result then follows by considering the isomorphism $H_1(\widehat{U}_2;\mathbb{Z})^- \cong H_1(U_2;\mathbb{Z})$ mentioned above. □

We now consider $H_1(\widehat{U}_1;\mathbb{Z})^-$.

Lemma 32 *There exists a basis for $H_1(\widehat{U}_1;\mathbb{Z})^-$ in which the intersection form is block diagonal with $\frac{1}{2}n_{\mathrm{odd}}-1$ blocks $\left(\begin{smallmatrix}0 & 4\\ -4 & 0\end{smallmatrix}\right)$ and with one block (0). In addition, the block (0) corresponds to the image of the homomorphism $H_1(\widehat{\gamma};\mathbb{Z})^- \to H_1(\widehat{U}_1;\mathbb{Z})^-$ induced by the inclusion map.*

Proof We will use an explicit description of the covering $\widehat{U}_1 \to U_1$, with a specific basis for $H_1(\widehat{U}_1;\mathbb{Z})$.

Recall that this covering is classified by a cohomology class in $H^1(U_1;\mathbb{Z}_2)$ that is dual to a family $K \subset U_1$ of $\frac{1}{2}n_{\mathrm{odd}}$ disjoint arcs, with one boundary point in each component of $\partial_{\mathrm{odd}} U$. Index the components of $\partial_{\mathrm{odd}} U$ as $\alpha_1,\alpha_2,\dots,\alpha_{n_{\mathrm{odd}}}$ and the components of K as $k_1,k_3,k_5,\dots,k_{n_{\mathrm{odd}}-1}$ in such a way that k_{2i-1} joins α_{2i-1} to α_{2i}. Add to K a family of disjoint arcs $k_2,k_4,\dots,k_{n_{\mathrm{odd}}-2}$, disjoint from the k_{2i-1}, such that each k_{2i} joins α_{2i} to α_{2i+1}. See Figure 3.

For $i=1,2,\dots,n_{\mathrm{odd}}-1$, consider a small regular neighborhood of $k_i \cup \alpha_i \cup \alpha_{i+1}$ in U_1 and let β_i be the boundary component of this neighborhood which is neither α_i nor α_{i+1}; endow β_i by the corresponding boundary orientation. Orient each curve α_i by the boundary orientation of $\partial_{\mathrm{odd}} U$.

The preimage of each curve α_i is a single curve $\widehat{\alpha}_i$, which we orient by the orientation of α_i. The preimage of β_j in $\widehat{U}_1$ consists of two disjoint curves. Arbitrarily choose one of these curves $\widehat{\beta}_j$ and orient it by the orientation of β_j. Then the $[\widehat{\alpha}_i]$ and $[\widehat{\beta}_j]$ form a basis for $H_1(\widehat{U}_1;\mathbb{Z})$. See Figure 3.

Consider an element $u \in H_1(\widehat{U}_1;\mathbb{Z})$, uniquely expressed in this basis as

$$u = \sum_{i=1}^{n_{\mathrm{odd}}} a_i[\widehat{\alpha}_i] + \sum_{j=1}^{n_{\mathrm{odd}}-1} b_j[\widehat{\beta}_j]$$

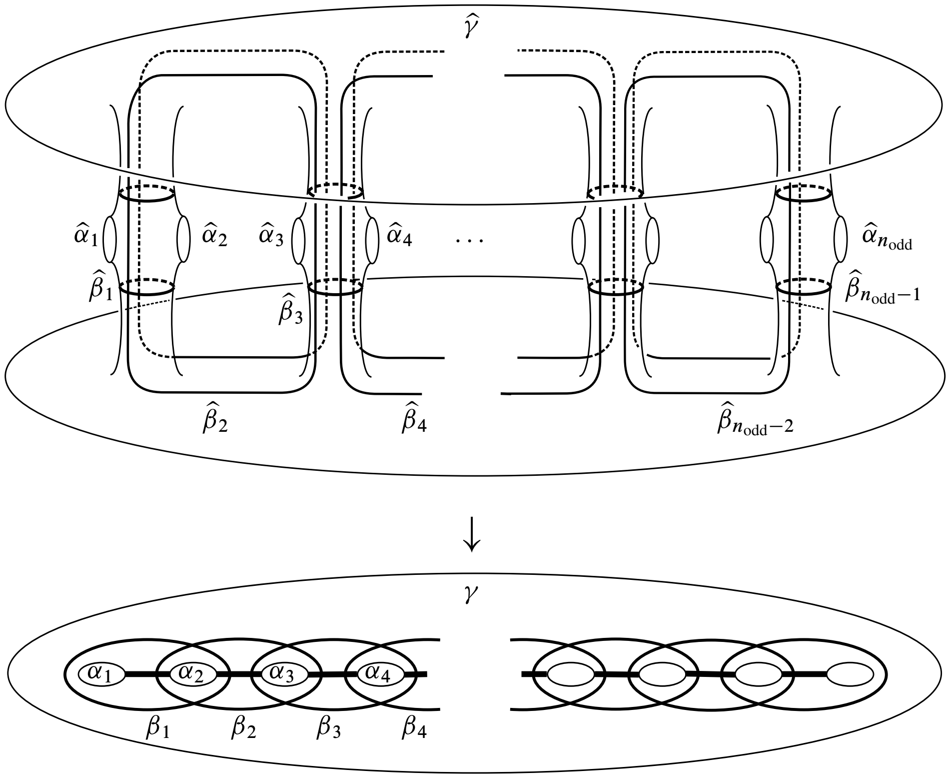


Figure 3

with all $a_i, b_j \in \mathbb{Z}$. By construction of the curves $\widehat{\alpha}_i$ and $\widehat{\beta}_j$,

$$\sigma_*([\widehat{\alpha}_i]) = [\widehat{\alpha}_i] \quad \text{and} \quad \sigma_*([\widehat{\beta}_j]) = -[\widehat{\beta}_j] - [\widehat{\alpha}_j] - [\widehat{\alpha}_{j+1}].$$

If u belongs to $H_1(\widehat{U}_1; \mathbb{Z})^-$, namely if $\sigma_*(u) = -u$, it follows from these observations and from the consideration of the coefficients of each $[\widehat{\alpha}_i]$ that we necessarily have

$$\begin{aligned} b_1 &= 2a_1, \\ b_i + b_{i-1} &= 2a_i \quad \text{for every } i \text{ with } 2 \leqslant i \leqslant n_{\mathrm{odd}} - 1, \\ b_{n_{\mathrm{odd}}-1} &= 2a_{n_{\mathrm{odd}}}. \end{aligned}$$

In particular, the coefficients b_j are all even, and

$$u = \tfrac{1}{2}(u - \sigma_*(u)) = \sum_{j=1}^{n_{\mathrm{odd}}-1} \tfrac{1}{2} b_j \, ([\widehat{\beta}_j] - \sigma_*([\widehat{\beta}_j])).$$

Therefore, the elements $[\widehat{\beta}_j]-\sigma_*([\widehat{\beta}_j])$ generate $H_1(\widehat{U}_1;\mathbb{Z})^-$. Since these elements $[\widehat{\beta}_j]-\sigma_*([\widehat{\beta}_j])=2[\widehat{\beta}_j]+[\widehat{\alpha}_j]+[\widehat{\alpha}_{j+1}]$ are linearly independent, they form a basis for $H_1(\widehat{U}_1;\mathbb{Z})^-$.

Note that $[\widehat{\beta}_j]\cdot[\widehat{\beta}_{j'}]=0$ if $|j-j'|>1$, and $[\widehat{\beta}_j]\cdot[\widehat{\beta}_{j+1}]=\varepsilon_j=\pm 1$, where the sign depends on which lift of β_j we chose for $\widehat{\beta}_j$. Also,

$$\sigma_*([\widehat{\beta}_j])\cdot[\widehat{\beta}_{j'}]=[\widehat{\beta}_j]\cdot\sigma_*([\widehat{\beta}_{j'}])=-\sigma_*([\widehat{\beta}_j])\cdot\sigma_*([\widehat{\beta}_{j'}])=-[\widehat{\beta}_j]\cdot[\widehat{\beta}_{j'}].$$

It follows that, in the basis of $H_1(\widehat{U}_1;\mathbb{Z})^-$ formed by the $[\widehat{\beta}_j]-\sigma_*([\widehat{\beta}_j])$, the matrix of the intersection form is

$$\begin{pmatrix} 0 & 4\varepsilon_1 & 0 & 0 & \cdots & 0 & 0 \\ -4\varepsilon_1 & 0 & 4\varepsilon_2 & 0 & \cdots & 0 & 0 \\ 0 & -4\varepsilon_2 & 0 & 4\varepsilon_3 & \cdots & 0 & 0 \\ 0 & 0 & -4\varepsilon_3 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 4\varepsilon_{n_{\mathrm{odd}}-2} \\ 0 & 0 & 0 & 0 & \cdots & -4\varepsilon_{n_{\mathrm{odd}}-2} & 0 \end{pmatrix}$$

By block diagonalizing this matrix, a final modification of the basis provides a new basis for $H_1(\widehat{U}_1;\mathbb{Z})^-$ in which the intersection form is block diagonal with $\frac{1}{2}n_{\mathrm{odd}}-1$ blocks $\left(\begin{smallmatrix} 0 & 4 \\ -4 & 0\end{smallmatrix}\right)$ and with one block (0).

There remains to show that the block (0) corresponds to the image of $H_1(\widehat{\gamma};\mathbb{Z})^-$. This could be seen by explicitly analyzing the block diagonalization process of the above matrix. However, it is easier to note that $H_1(\widehat{\gamma};\mathbb{Z})^-\cong\mathbb{Z}$ is generated by $[\widehat{\gamma}_1]-[\widehat{\gamma}_2]$, where $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ are the two components of the preimage $\widehat{\gamma}$ of γ and are oriented by the boundary orientation of $\partial\widehat{U}_1$. Then $[\widehat{\gamma}_1]-[\widehat{\gamma}_2]$ is in the kernel of the intersection form of $H_1(\widehat{U}_1;\mathbb{Z})^-$, since $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ are in the boundary of $\widehat{U}_1$, and generate this kernel since it is isomorphic to $\mathbb{Z}$ and since $[\widehat{\gamma}_1]-[\widehat{\gamma}_2]$ is indivisible in $H_1(\widehat{U}_1;\mathbb{Z})$. □

We now only need to combine the computations of Lemmas 30, 31 and 32 to obtain a basis of $H_1(\widehat{U};\mathbb{Z})^-$ in which the intersection form is block diagonal with h blocks $\left(\begin{smallmatrix} 0 & -2 \\ 2 & 0\end{smallmatrix}\right)$, $\frac{1}{2}n_{\mathrm{odd}}-1$ blocks $\left(\begin{smallmatrix} 0 & -4 \\ 4 & 0\end{smallmatrix}\right)$, and n_{even} blocks (0).

Applying Lemmas 27 and 28 to connect this to the Thurston intersection form on the edge weight space $\mathcal{W}(\tau;\mathbb{Z})$, we conclude that $\mathcal{W}(\tau;\mathbb{Z})$ admits a basis in which the intersection form is block diagonal with h blocks $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0\end{smallmatrix}\right)$, $\frac{1}{2}n_{\mathrm{odd}}-1$ blocks $\left(\begin{smallmatrix} 0 & -2 \\ 2 & 0\end{smallmatrix}\right)$, and n_{even} blocks (0). In addition, by the second half of Lemma 31 and using Lemma 29, the generators corresponding to the blocks (0) can be assumed to correspond to the elements of $\mathcal{W}(\tau;\mathbb{Z})$ associated to the components of $\partial_{\mathrm{even}}U$.

This proves Theorem 26, under our assumption that $n_{\mathrm{odd}}>0$.

We now consider the case where $n_{\mathrm{odd}} = 0$, namely where $\partial_{\mathrm{odd}} U = \varnothing$, and where the train track τ is nonorientable. This second property is equivalent to the property that the covering $\widehat{U} \to U$ is nontrivial. We can then realize the cohomology class of $H^1(U;\mathbb{Z}_2)$ classifying the covering $\widehat{U} \to U$ as the Poincaré dual of a nonseparating simple closed curve K. Let $U_1 \subset U$ be a surface of genus 1 containing K and bounded by a simple closed curve γ, and let U_2 be the closure of $U - U_1$. As before, let $\widehat{U}_1$, $\widehat{U}_2$ and $\widehat{\gamma}$ denote the respective preimages of U_1, U_2 and γ in $\widehat{U}$.

The computation of Lemma 31 applies to this case as well, and provides a basis for $H_1(\widehat{U}_2;\mathbb{Z})^-$ in which the intersection form is block diagonal with h blocks $\left(\begin{smallmatrix} 0 & -2 \\ 2 & 0 \end{smallmatrix}\right)$ and n_{even} blocks (0).

The surface $\widehat{U}_1$ is a twice-punctured torus. A simple analysis of the covering $\widehat{U}_1 \to U_1$ shows that $H_1(\widehat{U}_1;\mathbb{Z})^- \cong \mathbb{Z}$ is equal to the image of $H_1(\widehat{\gamma};\mathbb{Z})^-$. The intersection form of $H_1(\widehat{U}_1;\mathbb{Z})^-$ is then 0.

Again, combining these computations with the exact sequence

$$0 \to H_1(\widehat{\gamma};\mathbb{Z})^- \to H_1(\widehat{U}_1;\mathbb{Z})^- \oplus H_1(\widehat{U}_2;\mathbb{Z})^- \to H_1(\widehat{U};\mathbb{Z})^- \to 0$$

provides in this case a basis for $H_1(\widehat{U};\mathbb{Z})^-$ in which the intersection form is block diagonal with h blocks $\left(\begin{smallmatrix} 0 & -2 \\ 2 & 0 \end{smallmatrix}\right)$ and n_{even} blocks (0). Using Lemmas 27 and 28, this provides the result promised in Theorem 26 in this case as well. The fact that the generators corresponding to the blocks (0) can be chosen to be the elements associated to the components of $\partial_{\mathrm{even}} U$ is a byproduct of the proof as in the previous case.

Finally, we need to consider the case where $n_{\mathrm{odd}} = 0$ and the train track τ is orientable. Then the covering $\widehat{U} \to U$ is trivial, so $H_1(\widehat{U};\mathbb{Z})^- \cong H_1(U;\mathbb{Z})$ in such a way that the intersection form of $H_1(\widehat{U};\mathbb{Z})^-$ corresponds to twice the intersection form of $H_1(U;\mathbb{Z})$. By Lemma 28, the last case of Theorem 26 immediately follows. □

References

[1] **N Abdiel**, **C Frohman**, *The localized skein algebra is Frobenius*, preprint (2015) arXiv

[2] **J W Barrett**, *Skein spaces and spin structures*, Math. Proc. Cambridge Philos. Soc. 126 (1999) 267–275 MR

[3] **C Blanchet**, **N Habegger**, **G Masbaum**, **P Vogel**, *Topological quantum field theories derived from the Kauffman bracket*, Topology 34 (1995) 883–927 MR

[4] **F Bonahon**, *Shearing hyperbolic surfaces, bending pleated surfaces and Thurston's symplectic form*, Ann. Fac. Sci. Toulouse Math. 5 (1996) 233–297 MR

[5] **F Bonahon**, **X Liu**, *Representations of the quantum Teichmüller space and invariants of surface diffeomorphisms*, Geom. Topol. 11 (2007) 889–937 MR

[6] **F Bonahon**, **H Wong**, *Kauffman brackets, character varieties and triangulations of surfaces*, from "Topology and geometry in dimension three" (W Li, L Bartolini, J Johnson, F Luo, R Myers, J H Rubinstein, editors), Contemp. Math. 560, Amer. Math. Soc., Providence, RI (2011) 179–194 MR

[7] **F Bonahon**, **H Wong**, *Quantum traces for representations of surface groups in* $\mathrm{SL}_2(\mathbb{C})$, Geom. Topol. 15 (2011) 1569–1615 MR

[8] **F Bonahon**, **H Wong**, *The Witten–Reshetikhin–Turaev representation of the Kauffman bracket skein algebra*, Proc. Amer. Math. Soc. 144 (2016) 2711–2724 MR

[9] **F Bonahon**, **H Wong**, *Representations of the Kauffman bracket skein algebra, I: Invariants and miraculous cancellations*, Invent. Math. 204 (2016) 195–243 MR

[10] **F Bonahon**, **H Wong**, *Representations of the Kauffman bracket skein algebra, III: Closed surfaces and naturality*, preprint (2015) arXiv

[11] **F Bonahon**, **H Wong**, *Representations of the Kauffman bracket skein algebra, IV: Naturality for punctured surfaces*, in preparation

[12] **D Bullock**, *Estimating a skein module with* $\mathrm{SL}_2(\mathbf{C})$ *characters*, Proc. Amer. Math. Soc. 125 (1997) 1835–1839 MR

[13] **D Bullock**, *Rings of* $\mathrm{SL}_2(\mathbf{C})$*–characters and the Kauffman bracket skein module*, Comment. Math. Helv. 72 (1997) 521–542 MR

[14] **D Bullock**, **C Frohman**, **J Kania-Bartoszyńska**, *Understanding the Kauffman bracket skein module*, J. Knot Theory Ramifications 8 (1999) 265–277 MR

[15] **D Bullock**, **C Frohman**, **J Kania-Bartoszyńska**, *The Kauffman bracket skein as an algebra of observables*, Proc. Amer. Math. Soc. 130 (2002) 2479–2485 MR

[16] **D Bullock**, **J H Przytycki**, *Multiplicative structure of Kauffman bracket skein module quantizations*, Proc. Amer. Math. Soc. 128 (2000) 923–931 MR

[17] **L O Chekhov**, **V V Fock**, *Observables in 3D gravity and geodesic algebras*, Czechoslovak J. Phys. 50 (2000) 1201–1208 MR

[18] **V V Fock**, *Dual Teichmüller spaces*, unpublished preprint (1997) arXiv

[19] **V V Fok**, **L O Chekhov**, *Quantum Teichmüller spaces*, Teoret. Mat. Fiz. 120 (1999) 511–528 MR In Russian; translated in Theor. Math. Phys. 120 (1999), 1245–1259

[20] **C Frohman**, **N Abdiel**, *Frobenius algebras derived from the Kauffman bracket skein algebra*, J. Knot Theory Ramifications 25 (2016) art. id. 1650016, 1–25 MR

[21] **C Frohman**, **J Kania-Bartoszyńska**, *The structure of the Kauffman bracket skein algebra at roots of unity*, preprint (2016) arXiv

[22] **C Frohman**, **J Kania-Bartoszyńska**, **T T Q Lê**, *Unicity for representations of the Kauffman bracket skein algebra*, preprint (2017) arXiv

[23] **M Havlíček**, **S Pošta**, *On the classification of irreducible finite-dimensional representations of* $U_q'(\mathrm{so}_3)$ *algebra*, J. Math. Phys. 42 (2001) 472–500 MR

[24] **R M Kashaev**, *Quantization of Teichmüller spaces and the quantum dilogarithm*, Lett. Math. Phys. 43 (1998) 105–115 MR

[25] **T T Q Lê**, *On Kauffman bracket skein modules at roots of unity*, Algebr. Geom. Topol. 15 (2015) 1093–1117 MR

[26] **X Liu**, *The quantum Teichmüller space as a noncommutative algebraic object*, J. Knot Theory Ramifications 18 (2009) 705–726 MR

[27] **D Mumford**, **J Fogarty**, **F Kirwan**, *Geometric invariant theory*, 3rd edition, Ergeb. Math. Grenzgeb. 34, Springer (1994) MR

[28] **R C Penner**, **J L Harer**, *Combinatorics of train tracks*, Annals of Mathematics Studies 125, Princeton Univ. Press (1992) MR

[29] **J H Przytycki**, **A S Sikora**, *On skein algebras and* $\mathrm{Sl}_2(\mathbf{C})$*–character varieties*, Topology 39 (2000) 115–148 MR

[30] **N Reshetikhin**, **V G Turaev**, *Invariants of* 3*–manifolds via link polynomials and quantum groups*, Invent. Math. 103 (1991) 547–597 MR

[31] **N Takenov**, *Representations of the Kauffman skein algebra of small surfaces*, preprint (2015) arXiv

[32] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at `http://msri.org/publications/books/gt3m`

[33] **V G Turaev**, *Skein quantization of Poisson algebras of loops on surfaces*, Ann. Sci. École Norm. Sup. 24 (1991) 635–704 MR

[34] **V G Turaev**, *Quantum invariants of knots and* 3*–manifolds*, De Gruyter Studies in Mathematics 18, de Gruyter, Berlin (1994) MR

[35] **E Witten**, *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. 121 (1989) 351–399 MR

Department of Mathematics, University of Southern California
Los Angeles, CA, United States

Department of Mathematics, Carleton College
Northfield, MN, United States

`fbonahon@math.usc.edu`, `hwong@carleton.edu`

Received: 8 March 2016 Revised: 27 September 2016

Algebraic & Geometric Topology 17 (2017) 3435–3448

The unstabilized canonical Heegaard splitting of a mapping torus

YANQING ZOU

Let S be a closed orientable surface of genus at least 2. The action of an automorphism f on the curve complex of S is an isometry. Via this isometric action on the curve complex, a translation length is defined on f. The geometry of the mapping torus M_f depends on f. As it turns out, the structure of the minimal-genus Heegaard splitting also depends on f: the canonical Heegaard splitting of M_f, constructed from two parallel copies of S, is sometimes stabilized and sometimes unstabilized. We give an example of an infinite family of automorphisms for which the canonical Heegaard splitting of the mapping torus is stabilized. Interestingly, complexity bounds on f provide insight into the stability of the canonical Heegaard splitting of M_f. Using combinatorial techniques developed on 3–manifolds, we prove that if the translation length of f is at least 8, then the canonical Heegaard splitting of M_f is unstabilized.

57M27; 57M50

1 Introduction

Let S be a closed orientable surface of genus at least 2. Then there is a curve complex $\mathcal{C}(S)$ defined by Harvey [5]. Later, Masur and Minsky [6; 7] assigned a metric d on it and then proved that under this metric, the curve complex is δ–hyperbolic. Assume that f is an automorphism of S. Then f is extended to be an isomorphism of $\mathcal{C}(S)$ and hence an isometry on $(\mathcal{C}(S), d_{\mathcal{C}(S)})$. For simplicity, this isometry is still denoted by f. Then there is a translation length $d(f) = \min\{d_{\mathcal{C}(S)}(C, f(C)) \mid C \in \mathcal{C}^0(S)\}$ defined on f. If f is either reducible or periodic, there is an universal upper bound on the translation length of f^n for any $n \in N$. But if f is a pseudo-Anosov map, $d(f^n)$ goes to infinity as n goes to infinity; see [7, Proposition 7.6]. Conversely, if there is an universal upper bound on the translation length of f^n for any n, then by Thurston's result (see Casson and Bleiler [3]), f is either reducible or periodic. Otherwise, f is a pseudo-Anosov map.

Let $M = S \times I$ be an I–bundle of S. It is known that there are two standard Heegaard splittings for M; see Scharlemann and Thompson [8]. One, called the trivial Heegaard

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3435

splitting, is $S \times [0, 0.5] \cup_{S \times \{0.5\}} S \times [0.5, 1]$. The other one is as follows. Assume that there are a point $p \in S$ and an arc $a = p \times I$ in $S \times I$. Let $N(a)$ be the closed regular neighborhood of a in $S \times I$, $V_1 = \overline{S \times [0.3, 0.6] - N(a)}$ and $V_2 = \overline{S \times I - V_1}$. Then both V_1 and V_2 are compression bodies. Hence $V_1 \cup_{\partial_+ V_1} V_2$ is a Heegaard splitting of M.

For the 3–manifold $M = S \times I$, its boundary components consist of two homeomorphic surfaces, $S \times \{0\}$ and $S \times \{1\}$. Thus, gluing these two components by a homeomorphism $f \colon S \times \{1\} \to S \times \{0\}$ produces a closed 3–manifold M_f, called a mapping torus. Here there is a small change in the definition of the translation length of f in M_f, which is $d(f) = \min\{d_{\mathcal{C}(S \times \{0\})}(C \times \{0\}, f(C \times \{1\}))\}$, where $C \times \{0\}$ is an essential simple closed curve in $S \times \{0\}$.

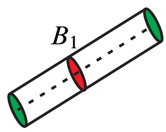


Figure 1: A core disk

It is not hard to see that there is a canonical Heegaard splitting for M_f, as follows. Let $V_2^f = V_2/f$ and let B_1 be the core disk of $N(a)$, as shown in Figure 1. Then V_2^f is homeomorphic to $S \times [0.5, 1] \cup_f B_1 \times [0, 0.5]$, where f maps a disk in $S \times \{1\}$ to $B_1 \times \{0\}$. Let $b \subset V_2^f$ be a properly embedded and unknotted arc connecting $S \times \{0.5\}$ and $S \times \{1\}$ and B_2 be the core disk of $N(b)$. Then

$$H_2 = \overline{V_2^f - N(b)}$$

is a handlebody. Equivalently,

$$H_2 = \overline{S - B_2} \times [0.5, 1] \cup_f B_1 \times [0, 0.5].$$

Moreover, H_1, the complement of H_2 in M_f, is given by

$$H_1 = \overline{S - B_1} \times [0, 0.5] \cup_f B_2 \times [0.5, 1].$$

So it is also a handlebody. Since $\partial H_1 = \Sigma = \partial H_2$, $H_1 \cup_\Sigma H_2$ is a Heegaard splitting of M_f, called the canonical Heegaard splitting.

A Heegaard splitting is *stabilized* if there is a pair of essential disks in two compression bodies such that their boundaries intersect in one point. If a Heegaard splitting is stabilized, then there is a move called a *destabilization* on it, which produces a smaller-genus Heegaard splitting. Thus, to study a Heegaard splitting of a 3–manifold, it is sufficient to study the destabilized one. Furthermore, there are some problems related to a Heegaard splitting, which all require that the Heegaard splitting is unstabilized. For example, the rank-versus-genus problem of a 3–manifold, ie when is $r(M) = g(M)$? Hence, for a given Heegaard splitting, it is a priority to determine its stability.

If f is periodic, then $S \times S^1$ is a finite covering of M_f, so M_f has the geometry of $H^2 \times R$; if f is reducible, then M_f contains at least one essential torus; if f is a pseudo-Anosov map, then Thurston [11, Theorem 0.1] proved that M_f is a hyperbolic 3–manifold. From this point of view, the geometry of M_f is determined by f. Moreover, the stability of its canonical Heegaard splitting is also influenced by f. For example, Schultens [9, Theorem 5.7] proved that if f is isotopic to an identity map, then the canonical Heegaard splitting of M_f is unstabilized; Souto and Biringer [10, Theorem 1.1; 2, Theorem 1.1] proved that if the pseudo-Anosov map f is complicated enough, the canonical Heegaard splitting is unstabilized; Bachmann and Schleimer [1, Corollary 3.2] proved that if the $d(f) \geq 2g(S)$, then the canonical Heegaard splitting is unstabilized and minimal.

With all these supporting results, it seems that the canonical Heegaard splitting of every mapping torus is unstabilized. However, this is not true in general; see Example 1.1.

Example 1.1 Let α and β be two essential simple closed curves in S, where $\alpha \cap \beta$ is a point p. It is known that $\tau_\alpha \circ \tau_\beta$, the concatenation of the two Dehn twists τ_α and τ_β, maps α to β. Let $S_\beta = \overline{S - \beta}$. By Thurston's classification [3] of automorphisms of a surface, there is a pseudo-Anosov map g on S_β fixing its boundary pointwise such that the translation length satisfies $d(g)|_{S_\beta} \geq 6$. Naturally g induces an automorphism on S, still denoted by g. Then $f = g \circ (\tau_\alpha \circ \tau_\beta)$.

Since $\alpha \times [0, 0.5]$ intersects $\beta \times [0.5, 1]$ in one point p on $S \times \{0.5\}$, there are two points $p_1, p_2 \in \alpha \times \{0.5\}$ disjoint from p such that $f(p_2 \times \{1\}) \neq p_1 \times \{0\}$. Let $a = p_1 \times [0, 0.5]$ and $b = p_2 \times [0.5, 1]$. Then both

$$H_1 = \overline{S \times \{0, 0.5\} - N(a)} \cup_f N(b) \quad \text{and} \quad H_2 = \overline{S \times \{0.5, 1\} - N(b)} \cup_f N(a)$$

are handlebodies. Moreover,

$$\overline{\alpha \times [0, 0.5] - N(a)} \quad \text{and} \quad \overline{\beta \times [0.5, 1] - N(b)}$$

are essential disks in H_1 and H_2, respectively, where they intersect in one point p. This means that the Heegaard splitting $H_1 \cup_\Sigma H_2$ is stabilized.

Remark 1.2 In Example 1.1, the translation length of g in S_β is at least 6. It is known that for any $n \in N$, there is an automorphism g of S_β whose translation length restricted to S_β is larger than n. So there are infinitely many choices of g in Example 1.1. Hence there are infinitely many choices of f on S.

So there is a question:

Question 1.3 What is the least value of $d(f)$ such that the canonical Heegaard splitting of M_f is unstabilized?

With tools developed in the curve complex, we give a partial answer to this question.

Theorem 1.4 *If the translation length satisfies $d(f) \geq 8$, then the canonical Heegaard splitting of M_f is unstabilized.*

This paper is organized as follows. We introduce some lemmas in Section 2, and prove the main theorem in Section 3.

2 Some lemmas

Let $\mathcal{C}(S)$ be the curve complex of S. Masur and Minsky proved:

Lemma 2.1 [6, Proposition 4.6] *$(\mathcal{C}(S), d)$ is connected and the diameter is infinite.*

Let $F \subset S$ be a subsurface. Then F is *essential* if there is no incompressible simple closed curve in F bounding a disk in S. If the subsurface F is essential, then Masur and Minsky [7, Section 2.2] introduced the subsurface projection on F for all of those vertices in the curve complex, as follows. For any vertex $\alpha \subset \mathcal{C}^0(S)$, by the bigon criterion [4, Proposition 1.7], there is a representative curve in its isotopy class that intersects ∂F essentially, ie there is no bigon capped by them in S. So the subsurface projection $\pi_F(\alpha)$ is defined to be one essential component of $\partial N(\alpha \cup \partial F)$ in F depending on choice.

An essential simple closed curve α cuts F if $\pi_F(\alpha) \neq \varnothing$. For any two given disjoint essential simple closed curves α and β, if they both cut F, then

$$d_{\mathcal{C}(F)}(\pi_F(\alpha), \pi_F(\beta)) \leq 2.$$

In general, Masur and Minsky proved:

Lemma 2.2 [7, Lemma 2.2] *Let F and S be as above, and let $\mathcal{G} = \{\alpha_0, \dots, \alpha_k\}$ be a geodesic in $\mathcal{C}(S)$ such that α_i cuts F for each $0 \leq i \leq k$. Then $d_{\mathcal{C}(F)}(\alpha_0, \alpha_k) \leq 2k$.*

It is known that when ∂F is connected, no component of $\pi_F(\alpha)$ cuts out a planar surface in F. But if ∂F is not connected, it is possible that some element of $\pi_F(\alpha)$ does cut out a planar subsurface of F. In this case, we introduce the definition of a strongly essential curve in F, which is defined in [12].

Definition 2.3 An essential simple arc or simple closed curve $c \subset F$ is *strongly essential* if no component of $\pi_F(c)$ cuts out a planar subsurface in F.

Let F be a compact orientable surface of genus at least 1 with connected boundary. For the handlebody $F \times [0,1]$, each essential disk intersects ∂F nontrivially. Moreover:

Lemma 2.4 *For any essential disk $D \subset F \times [0,1]$, there is an essential disk D_1 such that:*

(1) $\partial D_1 \cap F \times \{1\}$ *is connected and isotopic to a component of* $\partial D \cap F \times \{1\}$*;*

(2) $D_1 = (\partial D_1 \cap F \times \{1\}) \times [0,1]$*;*

(3) $\partial D_1 \cap F \times \{0\}$ *is disjoint from some component of* $\partial D \cap F \times \{0\}$*.*

Proof Without loss of generality, for any two essential disks in H, it is assumed that their intersection consists of arcs. Since ∂D intersects $F \times \{1\}$ nontrivially, there is an arc $a \subset \partial D \cap F \times \{1\}$ such that the number of components of $(a \times I) \cap D$ is minimal among all arcs in $\partial D \cap F \times \{1\}$.

Let $D_a = a \times I$. An essential arc $\alpha \subset F \times \{0\}$ is called a 0–arc. Similarly, an essential arc $\beta \subset (\partial H - F \times \{0\})$ is called an 1–arc. It is not hard to see that the boundary curve of D consists of alternating 1–arcs and 0–arcs while the boundary curve of D_a consists of one 1–arc and one 0–arc.

If $D_a \cap D = \varnothing$, then the proof is finished. So suppose that $D_a \cap D \neq \varnothing$. Then there is an outermost disk B in D where $B \cap D_a$ is an arc. Since $a \subset D \cap F \times \{1\}$, all of those intersecting arcs between D_a and D have ends in $\partial D_a \cap F \times \{0\}$. Therefore there is a 0–arc of $\partial B \cap F \times \{0\}$ in $\partial D \cap F \times \{0\}$ disjoint from $\partial D_a \cap F \times \{0\}$, for if not, then ∂B contains only one 1–arc and no 0–arc. Doing a boundary compression on D_a along B, D_a is changed into two disks $D_{a,1}$ and $D_{a,2}$. Since D intersects D_a essentially, these two disks are both essential. As one of $D_{a,1}$ and $D_{a,2}$ intersects $F \times \{0\}$ in one arc, one of these two disks is an I–bundle of the 1–arc of $\partial B \cap F \times \{1\}$. Without loss of generality, let $D_{a,1}$ be this disk. By the boundary compression surgery, the 1–arc of $\partial B \cap F \times \{1\}$ lies in ∂B and therefore in $\partial D \cap F \times \{1\}$. So $D_{a,1}$ is an I–bundle of some component of $\partial D \cap F \times \{1\}$. Moreover,

$$|\partial D_{a,1} \cap \partial D| \leq |\partial D_a \cap \partial D| - 2.$$

But this contradicts the choice of D_a. Then $\alpha \times \{0\}$ is disjoint from some 0–arc of ∂B and hence some 0–arc of ∂D. □

Similarly, there is also an essential disk $D_2 \subset F \times [0,1]$ such that:

(1) $\partial D_2 \cap F \times \{0\}$ is connected and isotopic to a component of $\partial D \cap F \times \{0\}$;

(2) $D_2 = (\partial D_2 \cap F \times \{0\}) \times [0,1]$;

(3) $\partial D_2 \cap F \times \{1\}$ is disjoint from some component of $\partial D \cap F \times \{1\}$.

3 Proof of Theorem 1.4

Let f, d, S, M_f, a, b, Σ, H_1, H_2, $H_1 \cup_\Sigma H_2$, B_1 and B_2 be as in Section 1. Then the main theorem is written as follows:

Proposition 3.1 *If the translation length satisfies $d(f) \geq 8$, then $H_1 \cup_\Sigma H_2$ is unstabilized.*

Before proving Proposition 3.1, we need the following lemma:

Lemma 3.2 *For any essential simple closed curve C bounding two essential disks in H_1 and H_2 simultaneously, both $C \cap \partial B_1 \neq \varnothing$ and $C \cap \partial B_2 \neq \varnothing$.*

Proof Since $S \times I$ is irreducible and its boundary components are incompressible, M_f is irreducible and not homeomorphic to S^3.

The construction of $H_1 \cup_\Sigma H_2$ in Section 1 says that

$$H_1 = \overline{S - B_1} \times [0, 0.5] \cup B_2 \times [0.5, 1]$$

and

$$H_2 = \overline{S - B_2} \times [0.5, 1] \cup B_1 \times [0, 0.5].$$

Assume that C bounds an essential disk D (resp. E) in H_1 (resp. H_2). If we consider the intersection between E and B_1 in H_2, then:

Fact 3.3 $$C \cap \partial B_1 \neq \varnothing.$$

Proof Suppose the conclusion is false. Then C is either isotopic to ∂B_1 or disjoint from ∂B_1. Since ∂B_1 bounds no disk in H_1, C is not isotopic to ∂B_1. Thus C is disjoint from ∂B_1. Moreover, C is strongly essential in $\Sigma_{B_1} = \overline{\Sigma - \partial B_1}$, for if not, then C cuts out a pair of pants P in Σ_{B_1} such that ∂P consists of two copies of ∂B_1 and C. Since C bounds an essential disk E in H_2, E cuts out a solid torus $ST \subset H_2$ containing B_1. Similarly, the essential disk D also cuts out a solid torus in H_1. Then the Heegaard splitting $H_1 \cup_\Sigma H_2$ is a connected sum of a genus-1 Heegaard splitting and a smaller-genus Heegaard splitting. Because M_f is irreducible, one of these two Heegaard splittings is of S^3, which implies that the genus-1 Heegaard splitting is not of S^3. The reason is that since the longitude l of the solid torus ST intersects ∂B_1 in one point, l intersects $S \times \{t\}$ in one point for some $t \in (0, 0.5)$. So the representative of l in $\pi_1 M_f$ is nontrivial. Then the Heegaard splittings of genus $(g(\Sigma)-1)$ belongs to S^3. Hence, under this circumstance, M_f is a lens space. Moreover, it contains a closed embedded genus at least 1 incompressible surface. But it contradicts the fact that there is no positive genus closed incompressible surface in a lens space.

After removing $N(a)$ from H_2, H_2 is changed into

$$H_2^{B_1} = \overline{S \times [0.5, 1] - N(b)}.$$

Let

$$H_1^* = \overline{M_f - H_2^{B_1}}.$$

Equivalently, $H_1^* = S \times [0, 0.5] \cup N(b)$. Since C is strongly essential in Σ_{B_1} and $C \cap \partial B_1 = \varnothing$, C is essential in $\partial H_2^{B_1}$. So E is also an essential disk in $H_2^{B_1}$. The I–bundle structure of $H_2^{B_1}$ implies that $C = \partial E$ intersects ∂B_2 nontrivially. Since C (resp. ∂B_2) bounds an essential disk D (resp. B_2) in H_1^*, by the standard outermost disk argument, there is an outermost disk of D in $S \times [0, 0.5] = \overline{H_1^* - B_2}$. By the proof of Lemma 2.6 in [12], this outermost disk is a properly embedded essential disk of $S \times [0, 0.5]$. But this contradicts the fact that $\partial(S \times [0, 0.5])$ is incompressible in $S \times [0, 0.5]$. □

Similarly, $C \cap \partial B_2 \neq \varnothing$. This completes the proof of Lemma 3.2. □

Then the proof of the Proposition 3.1 is written as follows:

Proof of Proposition 3.1 Since $S \times I$ is irreducible and its boundary components are incompressible, M_f is irreducible and not homeomorphic to S^3.

Suppose that the conclusion is false. Then $H_1 \cup_\Sigma H_2$ is stabilized. It is known that each stabilized Heegaard splitting is either reducible or a genus-1 Heegaard splitting of S^3. Since M_f is not homeomorphic to the S^3, the canonical Heegaard splitting $H_1 \cup_\Sigma H_2$ is reducible. Therefore, there is an essential simple closed curve $C \subset \Sigma$ such that C bounds an essential disk D (resp. E) in H_1 (resp. H_2).

It is not hard to see that there is an isotopy on D such that $\partial D \cap \partial E = \varnothing$ (just pushing ∂D away from ∂E). Without loss of generality, it is assumed that ∂D intersects $\partial B_1 \sqcup \partial B_2$ essentially, ie there is no bigon capped by any two of them in Σ. By Lemma 3.2, neither ∂B_1 nor ∂B_2 is disjoint from C. Then $D \cap B_2 \neq \varnothing$. Furthermore, we assume that D intersects B_2 minimally. So $D \cap B_2$ consists of arcs and no closed circle. By the standard outermost disk argument, there is an outermost disk

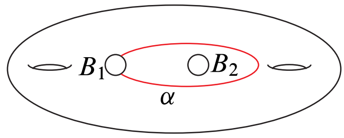


Figure 2: A one-hole bigon

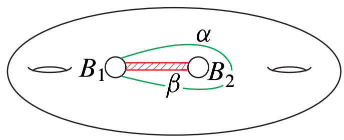


Figure 3: The case where ∂B_1, ∂B_2 and β bound a rectangle

in D bounded by a component $\alpha \subset \partial D$ and an arc of $D \cap B_2$. Similarly, there is an outermost disk in E bounded by a component $\beta \subset \partial E$ and an arc of $E \cap B_1$.

Let $\Sigma_{B_1} = \overline{\Sigma - \partial B_1}$ and $\Sigma_{B_2} = \overline{\Sigma - \partial B_2}$. Then:

Claim 3.4 *The arc α (resp. β) is strongly essential in Σ_{B_2} (resp. Σ_{B_1}).*

Proof We prove this claim for α only; the other case is similar.

Since ∂B_2 is nonseparating in Σ, Σ_{B_2} has two boundary curves C_1 and C_2. Suppose α is not strongly essential in Σ_{B_2}. Then α cuts out an annulus in Σ_{B_2} which contains one boundary component of Σ_{B_2}, for example, C_2. So

$$|C \cap C_2| \leq |C \cap C_1| - 2.$$

But it contradicts the fact that C_1 and C_2 are isotopic in Σ. □

Let $H_1^{B_2} = \overline{H_1 - B_2}$ and $H_2^{B_1} = \overline{H_2 - B_1}$. Since C intersects both ∂B_1 and ∂B_2 essentially, there is no bigon capped by α and ∂B_1 (resp. β and ∂B_2) in Σ_{B_2} (resp. Σ_{B_1}). Furthermore:

Claim 3.5 *There is no one-hole bigon capped by α and ∂B_1 in Σ_{B_2}.*

Note 3.6 A one-hole bigon is shown in Figure 2.

Proof of Claim 3.5 Suppose that the conclusion is false. Then there is a one-hole bigon capped by α and ∂B_1 in Σ_{B_2}. Since $\beta \cap \alpha = \varnothing$ and $\partial\beta \subset \partial B_1$, either $\beta \cap \partial B_2 = \varnothing$ or β intersects ∂B_2 in at most two points. In the latter case, there is a rectangle bounded by ∂B_1, ∂B_2 and β; see Figure 3. For both of these two cases, it is not hard to see that

$\pi_{\Sigma_{B_1}}(\beta)$ is disjoint from $\partial B_1 \cup \partial B_2$ up to isotopy. But since β is in the boundary of the outermost disk in E and strongly essential in Σ_{B_1}, $\pi_{\Sigma_{B_1}}(\beta)$ bounds an essential disk in $H_2^{B_1}$. So $\pi_{\Sigma_{B_1}}(\beta) \cap \partial B_1 \neq \varnothing$ up to isotopy. This is a contradiction. $\square$

Similarly, there is no one-hole bigon capped by β and ∂B_2 in Σ_{B_1}.

Although ∂D intersects ∂B_1 and ∂B_2 minimally, it is possible there is a rectangle bounded by ∂B_1, ∂B_2 and α in Σ_{B_2}; see Figure 4.

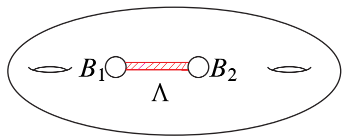


Figure 4: A rectangle

Let

$$\begin{aligned} S_1 &= S_1 \times \{0.5\} = \overline{S \times \{0.5\} - B_1}, \\ S_3 &= S_3 \times \{0.5\} = \overline{S \times \{0.5\} - B_2}, \\ S_2 &= S_1 \cap S_3. \end{aligned}$$

Then $H_1^{B_2} = S_1 \times [0, 0.5]$ and $H_2^{B_1} = S_3 \times [0.5, 1]$.

Claim 3.7 *There is no rectangle bounded by ∂B_1, ∂B_2 and α in Σ_{B_2}.*

Proof Without loss of generality, we assume that both $\partial\alpha$ and $\partial\beta$ are in S_2. The other cases are similar, so we omit them here.

Suppose the conclusion is false. Then there is a rectangle Λ bounded by ∂B_1, ∂B_2 and α in Σ. Although the proof is similar to the proof of Lemma 3.9 in [13], for integrity, it is written here. If $\beta \cap \Lambda \neq \varnothing$, then $\Lambda \cap \beta$ is one or two arcs connecting ∂B_1 and ∂B_2. Otherwise there is at least one point in $\alpha \cap \beta$. Since $\beta \cap \partial B_1 = \partial\beta$ and $\alpha \cap \partial B_2 = \partial\alpha$, there is an isotopy on β such that β is pushed away from Λ. Moreover, $\alpha \cap \beta = \varnothing$. Therefore we may assume that β is disjoint from Λ.

For simplicity, $\pi_{\Sigma_{B_2}}(\alpha)$, disjoint from β, is abbreviated by α. It is not hard to see that there is a bigon capped by α and ∂B_1. Then there is an isotopy on α such that there is no bigon capped by α and ∂B_1 anymore. As a result of this process, by the proof of Claim 3.5, there is no one-hole bigon bounded by α and ∂B_1 in Σ_{B_2}. At the end, there is no bigon or one-hole bigon capped by α and ∂B_1 in Σ_{B_2}. So α intersects ∂B_1 in $\partial H_1^{B_2}$ essentially (for if not, then there is a bigon capped by them, which corresponds to a one-hole bigon or a bigon in Σ_{B_2}). On one hand, since $H_1^{B_2} = S_1 \times [0, 0.5]$, by Lemma 2.4, there is one essential arc $a \subset \alpha \cap S_1 \times \{0.5\}$ such that $a \times \{0\}$ is disjoint from some component $c \subset \alpha \cap S_1 \times \{0\}$. On the other hand, for the subsurface $S_2 \subset \Sigma_{B_2}$, since $S_1 = S_2 \cup B_2$, we have $\alpha \cap S_1 = \alpha \cap S_2$. Then $a \subset S_2$. Since β intersects no bigon bounded by α and ∂B_1 in this isotopy, $\alpha \cap \beta = \varnothing$. Hence $a \cap \beta = \varnothing$.

If the union of β, ∂B_1 and ∂B_2 bound a rectangle in Σ_{B_1}, then $\pi_{\Sigma_{B_1}}(\beta)$, still denoted by β, misses α. Otherwise $\alpha \cap \beta \neq \varnothing$. By the same argument as above, there is also an isotopy on β such that there is no bigon bounded by β and ∂B_2 anymore. As a result of this process, by the proof of Claim 3.5, there is no one-hole bigon bounded by β and ∂B_2 in Σ_{B_1}. Therefore there is no bigon or one-hole bigon capped by β and ∂B_2 in $\partial H_2^{B_1}$. So β intersects ∂B_2 in $\partial H_2^{B_1}$ essentially. On one hand, since $H_2^{B_1} = S_3 \times [0.5, 1]$, by Lemma 2.4, there is one essential arc $b \subset \beta \cap S_1 \times \{0.5\}$ such that $b \times \{1\}$ is disjoint from some component $d \subset \beta \cap S_3 \times \{1\}$. On the other hand, for the subsurface $S_2 \subset \Sigma_{B_1}$, since $S_3 = S_2 \cup B_1$, we have $\beta \cap S_3 = \beta \cap S_2$. Then $b \subset S_2$. Since α intersects no bigon bounded by β and ∂B_2 in the isotopy, $\alpha \cap \beta = \varnothing$. Hence $a \cap b = \varnothing$.

Since $\alpha \cap \beta = \varnothing$, $c \cap f(d) = \varnothing$. Hence

$$\begin{aligned} \pi_{S_2}(a) \cap \pi_{S_2}(b) &= \varnothing; \\ \pi_{S_1 \times \{0\}}(c) \cap f(\pi_{S_3 \times \{1\}}(d)) &= \varnothing; \\ d_{\mathcal{C}(S_3 \times \{1\})}(\pi_{S_3 \times \{1\}}(b \times \{1\}), \pi_{S_3 \times \{1\}}(d)) &\le 2; \\ d_{\mathcal{C}(S_1 \times \{0\})}(\pi_{S_1 \times \{0\}}(a \times \{0\}), \pi_{S_1 \times \{0\}}(c)) &\le 2. \end{aligned}$$

For simplicity, $\pi_{S_2}(a)$ (resp. $\pi_{S_2}(b)$) is abbreviated by a (resp. b). Since $a \times \{1\} \subset S_3 \times \{1\}$ intersects $b \times \{1\}$ trivially, the above equations and inequalities are changed as follows:

$$\begin{aligned} d_{\mathcal{C}(S_1 \times \{0\})}(a \times \{0\}, \pi_{S_1 \times \{0\}}(c)) &\le 2; \\ d_{\mathcal{C}(S_1 \times \{0\})}(\pi_{S_1 \times \{0\}}(c), f(\pi_{S_3 \times \{1\}}(d))) &\le 1; \\ d_{\mathcal{C}(S_3 \times \{1\})}(\pi_{S_3 \times \{1\}}(d), b \times \{1\})) &\le 2; \\ d_{\mathcal{C}(S_3 \times \{1\})}(b \times \{1\}, a \times \{1\}) &\le 1. \end{aligned}$$

It is known that every essential simple closed curve of $S_1 = \overline{S - B_1}$ is essential in S, and similarly for $S_3 = \overline{S - B_2}$. Then by the triangle inequality,

$$\begin{aligned}
d_{\mathcal{C}(S\times\{0\})}(a\times\{0\}, f(a\times\{1\})) &\le d_{\mathcal{C}(S\times\{0\})}(a\times\{0\}, \pi_{S_1\times\{0\}}(c)) \\
&\quad + d_{\mathcal{C}(S\times\{0\})}(\pi_{S_1\times\{0\}}(c), f(\pi_{S_3\times\{1\}}(d))) \\
&\quad + d_{\mathcal{C}(S\times\{0\})}(f(\pi_{S_3\times\{1\}}(d)), f(b\times\{1\})) \\
&\quad + d_{\mathcal{C}(S\times\{0\})}(f(b\times\{1\}), f(a\times\{1\})) \\
&\le d_{\mathcal{C}(S_1\times\{0\})}(a\times\{0\}, \pi_{S_1\times\{0\}}(c)) \\
&\quad + d_{\mathcal{C}(S_1\times\{0\})}(\pi_{S_1\times\{0\}}(c), f(\pi_{S_3\times\{1\}}(d))) \\
&\quad + d_{\mathcal{C}(S\times\{0\})}(f(\pi_{S_3\times\{1\}}(d)), f(b\times\{1\})) \\
&\quad + d_{\mathcal{C}(S\times\{0\})}(f(b\times\{1\}), f(a\times\{1\})) \\
&\le 2+1+d_{\mathcal{C}(S\times\{0\})}(f(\pi_{S_3\times\{1\}}(d)), f(b\times\{1\})) \\
&\quad + d_{\mathcal{C}(S\times\{0\})}(f(b\times\{1\}), f(a\times\{1\})) \\
&\le 3+d_{\mathcal{C}(S\times\{1\})}(\pi_{S_3\times\{1\}}(d), b\times\{1\}) \\
&\quad + d_{\mathcal{C}(S\times\{1\})}(b\times\{1\}, a\times\{1\}) \\
&\le 3+d_{\mathcal{C}(S_3\times\{1\})}(\pi_{S_3\times\{1\}}(d), b\times\{1\}) \\
&\quad + d_{\mathcal{C}(S_3\times\{1\})}(b\times\{1\}, a\times\{1\}) \\
&\le 6.
\end{aligned}$$

But this contradicts the choice of f.

So there is no rectangle bounded by β, ∂B_1 and ∂B_2 in Σ_{B_1}. Moreover, there is no one-hole bigon or bigon capped by $\pi_{\Sigma_{B_1}}(\beta)$ and ∂B_2 in Σ_{B_1}. Otherwise, there is either a rectangle bounded by β, ∂B_1 and ∂B_2 in Σ_{B_1} or a one-hole bigon bounded by β and ∂B_1, which is prohibited by Claim 3.5. Then each component of $\pi_{\Sigma_{B_1}}(\beta)$ intersects ∂B_2 essentially in $\partial H_2^{B_1}$. On one hand, since $H_2^{B_1} = S_3\times[0.5,1]$, by Lemma 2.4, there is one component $b \subset \pi_{\Sigma_{B_1}}(\beta)\cap S_3$ such that $b\times\{1\}$ is disjoint from one component d of $\pi_{\Sigma_{B_1}}(\beta)\cap S_3\times\{1\}$. On the other hand, since $\pi_{\Sigma_{B_1}}(\beta)\cap S_3 = \pi_{\Sigma_{B_1}}(\beta)\cap S_2$, we have $b\subset \pi_{\Sigma_{B_1}}(\beta)\cap S_2$.

Since $\alpha\cap\beta=\varnothing$, $a\cap b$ consists of at most two points, where the worst scenario is that ∂a is not separated by β in ∂B_1. Since $\partial a\subset\partial B_1$ and $\partial b\subset\partial B_2$, $\pi_{S_2}(a)\cap\pi_{S_2}(b)$ consists of at most two points. For simplicity, $\pi_{S_2}(a)$ (resp. $\pi_{S_2}(b)$) is abbreviated by a (resp. b). Then

$$d_{\mathcal{C}(S_3\times\{1\})}(b\times\{1\}, a\times\{1\}) \le 2.$$

By the same argument as above, $d(f)\le 7$. □

Similarly, there is no rectangle bounded by ∂B_1, ∂B_2 and β in Σ.

By Claims 3.5 and 3.7, there is neither a one-hole bigon nor a bigon capped by $\pi_{\Sigma_{B_2}}(\alpha)$ and ∂B_1 in Σ_{B_2}. Otherwise there is a rectangle bounded by the union of α, ∂B_1

and ∂B_2. This means that $\pi_{\Sigma_{B_2}}(\alpha)$ intersects ∂B_1 in $\partial H_1^{B_2}$ essentially without doing any further isotopy. Similarly, $\pi_{\Sigma_{B_1}}(\beta)$ intersects ∂B_2 in $\partial H_2^{B_1}$ essentially without doing any further isotopy too.

Then it is not hard to see that:

Fact 3.8 *Each component of $\pi_{\Sigma_{B_2}}\alpha \cap S_2$ intersects every component of $\pi_{\Sigma_{B_1}}\beta \cap S_2$ in at most two points.*

Proof It is sufficient to prove that there are at most two points in $\pi_{\Sigma_{B_2}}\alpha \cap \pi_{\Sigma_{B_1}}\beta$. Since α is disjoint from β, the worst scenario is that $\alpha \cap \partial B_1$ is separated by $\partial\beta$ while $\beta \cap \partial B_2$ is separated by $\partial\alpha$. Then there are two points in $\pi_{\Sigma_{B_2}}\alpha \cap \pi_{\Sigma_{B_1}}\beta$. So the conclusion holds. □

For simplicity, $\pi_{\Sigma_{B_2}}(\alpha)$ (resp. $\pi_{\Sigma_{B_1}}(\beta)$) is abbreviated by α (resp. β). Then:

Claim 3.9 *There is an essential simple closed curve γ in S such that*

$$d_{\mathcal{C}(S\times\{0\})}(f(\gamma\times\{1\}), \gamma\times\{0\}) \le 7.$$

Proof Since α bounds an essential disk in $S_1 \times [0, 0.5]$, by Lemma 2.4, there is a component a of $\alpha \cap S_1 \times \{0.5\}$ such that $a \times \{0\} \subset S_1 \times \{0\}$ is disjoint from some component $c \subset \alpha \cap S_1 \times \{0\}$. Similarly, there are two such components b and d for β.

By Fact 3.8, a intersects b in at most two points. Since $\partial a \subset \partial B_1$ and $\partial b \subset \partial B_2$, $\pi_{S_2}(a)$ intersects $\pi_{S_2}(b)$ in at most two points. Then since $g(S) \ge 2$, there is a strongly essential simple closed curve γ in S_2 disjoint from both α and β and hence from both $\pi_{S_2}(a)$ and $\pi_{S_2}(b)$. Let $\gamma \times [0, 5, 1]$ and $\gamma \times [0, 0.5]$ be the product I–bundles in $S \times [0.5, 1]$ and $S \times [0, 0.5]$, respectively. Then

$$\gamma \times \{1\} \cap \pi_{S_3}(b \times \{1\}) = \varnothing \quad \text{and} \quad \gamma \times \{0\} \cap \pi_{S_1}(a \times \{0\}) = \varnothing.$$

For simplicity, $\pi_{S_2}(a)$ (resp. $\pi_{S_2}(b)$) is abbreviated by a (resp. b). Therefore $\pi_{S_3} b \times \{1\}$ (resp. $\pi_{S_1} a \times \{0\}$) is isotopic to $b \times \{1\}$ (resp. $a \times \{0\}$). Then by the proof of Claim 3.7,

$$\begin{aligned} d_{\mathcal{C}(S\times\{0\})}(\gamma\times\{0\}, f(\gamma\times\{1\})) \le{} & d_{\mathcal{C}(S\times\{0\})}(\gamma\times\{0\}, a\times\{0\}) \\ & + d_{\mathcal{C}(S\times\{0\})}(a\times\{0\}, \pi_{S_1\times\{0\}}(c)) \\ & + d_{\mathcal{C}(S\times\{0\})}(\pi_{S_1\times\{0\}}(c), f(\pi_{S_3\times\{1\}}(d))) \\ & + d_{\mathcal{C}(S\times\{0\})}(f(\pi_{S_3\times\{1\}}(d)), f(b\times\{1\})) \\ & + d_{\mathcal{C}(S\times\{0\})}(f(b\times\{1\}), f(\gamma\times\{1\})) \end{aligned}$$

$$
\begin{aligned}
&\le 1 + d_{\mathcal{C}(S_1\times\{0\})}(a\times\{0\}, \pi_{S_1\times\{0\}}(c)) \\
&\qquad + d_{\mathcal{C}(S_1\times\{0\})}(\pi_{S_1\times\{0\}}(c), f(\pi_{S_3\times\{1\}}(d))) \\
&\qquad + d_{\mathcal{C}(S\times\{0\})}(f(\pi_{S_3\times\{1\}}(d)), f(b\times\{1\})) \\
&\qquad + d_{\mathcal{C}(S\times\{0\})}(f(b\times\{1\}), f(\gamma\times\{1\})) \\
&\le 1+2+1+d_{\mathcal{C}(S\times\{0\})}(f(\pi_{S_3\times\{1\}}(d)), f(b\times\{1\})) \\
&\qquad + d_{\mathcal{C}(S\times\{0\})}(f(b\times\{1\}), f(\gamma\times\{1\})) \\
&\le 4 + d_{\mathcal{C}(S\times\{1\})}(\pi_{S_3\times\{1\}}(d), b\times\{1\}) \\
&\qquad + d_{\mathcal{C}(S\times\{1\})}(b\times\{1\}, \gamma\times\{1\}) \\
&\le 4 + d_{\mathcal{C}(S_3\times\{1\})}(\pi_{S_3\times\{1\}}(d), b\times\{1\}) \\
&\qquad + d_{\mathcal{C}(S_3\times\{1\})}(b\times\{1\}, \gamma\times\{1\}) \\
&\le 7.
\end{aligned}
$$

This completes the proof of Claim 3.9. □

By Claim 3.9, the translation length of f is at most 7. This contradicts the assumption on f and completes the proof of Proposition 3.1. □

Acknowledgements We would like to thank the referee for helpful comments and suggestions. This work was partially supported by NSFC (No. 11601065) and NSF-DMU 0915150529.

References

[1] **D Bachman**, **S Schleimer**, *Surface bundles versus Heegaard splittings*, Comm. Anal. Geom. 13 (2005) 903–928 MR

[2] **I Biringer**, **J Souto**, *Ranks of mapping tori via the curve complex*, J. Reine Angew. Math. (online publication August 2016) 20 pages

[3] **A J Casson**, **S A Bleiler**, *Automorphisms of surfaces after Nielsen and Thurston*, London Math. Soc. Student Texts 9, Cambridge Univ. Press (1988) MR

[4] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) MR

[5] **W J Harvey**, *Boundary structure of the modular group*, from "Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference" (I Kra, B Maskit, editors), Ann. of Math. Stud. 97, Princeton Univ. Press (1981) 245–251 MR

[6] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, I: Hyperbolicity*, Invent. Math. 138 (1999) 103–149 MR

[7] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, II: Hierarchical structure*, Geom. Funct. Anal. 10 (2000) 902–974 MR

[8] **M Scharlemann**, **A Thompson**, *Heegaard splittings of* (surface) $\times I$ *are standard*, Math. Ann. 295 (1993) 549–564 MR

[9] **J Schultens**, *The classification of Heegaard splittings for (compact orientable surface)* $\times S^1$, Proc. London Math. Soc. 67 (1993) 425–448 MR

[10] **J Souto**, *The rank of the fundamental group of certain hyperbolic* 3*–manifolds fibering over the circle*, from "The Zieschang Gedenkschrift" (M Boileau, M Scharlemann, R Weidmann, editors), Geom. Topol. Monogr. 14, Geom. Topol. Publ., Coventry (2008) 505–518 MR

[11] **W P Thurston**, *Hyperbolic structures on* 3*–manifolds, II: Surface groups and* 3*–manifolds which fiber over the circle*, preprint (1998) arXiv

[12] **Y Zou**, **K Du**, **Q Guo**, **R Qiu**, *Unstabilized self-amalgamation of a Heegaard splitting*, Topology Appl. 160 (2013) 406–411 MR

[13] **Y Zou**, **X Liu**, *Bilateral self-amalgamation of a Heegaard splitting and Hempel distance*, Sci. China Math. 58 (2015) 1499–1510 MR

Department of Mathematics, Dalian Minzu University
Dalian, China

yanqing@dlnu.edu.cn

Received: 23 April 2016 Revised: 14 May 2017

Algebraic & Geometric Topology 17 (2017) 3449–3460

Nine generators of the skein space of the 3–torus

ALESSIO CARREGA

We show that the skein vector space of the 3–torus is finitely generated. We show that it is generated by nine elements: the empty set, some simple closed curves representing the nonzero elements of the first homology group with coefficients in $\mathbb{Z}_2$, and a link consisting of two parallel copies of one of the previous nonempty knots.

57MXX

1 Introduction

An alternative approach to representation theory for *quantum invariants* is provided by *skein theory*. The word "skein" and the notion were introduced by Conway in 1970 for his model of the *Alexander polynomial*. This idea became quite useful after the work of Kauffman [10] which redefined the *Jones polynomial* in a very simple and combinatorial way passing through the *Kauffman bracket*. These combinatorial techniques allow us to reproduce all quantum invariants arising from the representations of $U_q(\mathfrak{sl}_2)$ without any reference to representation theory. This also leads to many interesting and quite easy computations. This skein method was used by Blanchet, Habegger, Masbaum and Vogel [1], Kauffman and Lins [11] and Lickorish [12; 13; 15; 14] to reinterpret and extend some of the methods of representation theory.

The first notion in skein theory is that of a "*skein vector space*" (or *skein module*). These are vector spaces (R–modules) associated to oriented 3–manifolds, where the base field is equipped with a fixed invertible element A. These were introduced independently in 1988 by Turaev [24] and in 1991 by Przytycki [20]. We can think of them as an attempt to get an algebraic topology for knots: they can be seen as homology spaces obtained using isotopy classes instead of homotopy or homology classes. In fact, they are defined taking a vector space generated by subobjects (*framed links*) and then quotienting them by some relations. In this framework, the following questions arise naturally and are still open in general:

Question 1.1
- Are skein spaces (modules) computable?
- How powerful are they to distinguish 3–manifolds and links?

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3449

- Do the vector spaces (modules) reflect the topology/geometry of the 3–manifolds (eg surfaces, geometric decomposition)?
- Does this theory have a functorial aspect? Can it be extended to a functor from a category of cobordisms to the category of vector spaces (modules) and linear maps?

Skein spaces (modules) can also be seen as deformations of the ring of the $\mathrm{SL}_2(\mathbb{C})$–*character variety* of the 3–manifold; see Bullock [3]. Moreover, they are useful to generalize the Kauffman bracket, hence the Jones polynomial, to manifolds other than S^3. Thanks to Hoste and Przytycki [9], Przytycki [22] and (with different techniques) Costantino [4], now we can define the Kauffman bracket also in the connected sum $\#_g(S^1 \times S^2)$ of $g \geq 0$ copies of $S^1 \times S^2$.

Currently, there are only few 3–manifolds whose skein space (module) is known; see for instance Bullock [2], Hoste and Przytycki [7; 8; 9], Marché [16], Mroczkowski [18; 17], Mroczkowski and Dabkowski [19] and Przytycki [21; 22; 23]. Another natural question is:

Question 1.2 Is the skein vector space of a closed oriented 3–manifold always finitely generated?

In this paper, we take as base field the set $\mathbb{Q}(A)$ of all rational functions with rational coefficients and abstract variable A, and we note that every result in this work holds also for the field $\mathbb{C}$ of complex numbers with $A \in \mathbb{C}$ a nonzero number such that $A^{2n} \neq 1$ for every $n > 0$.

Theorem 1.3 *The skein space $K(T^3)$ of the 3–torus $T^3 = S^1 \times S^1 \times S^1$ is finitely generated.*

A set of nine generators is given by the empty set $\varnothing$, some simple closed curves representing the nonzero elements of the first homology group $H_1(T^3; \mathbb{Z}_2) \cong (\mathbb{Z}_2)^3$ with coefficients in $\mathbb{Z}_2$, and a skein element α that is equal to the link consisting of two parallel copies of any previous nonempty knots.

Our main tool is the algebraic work of Frohman and Gelca [5]. The skein space (module) of a (thickened) surface has a natural algebra structure obtained by overlap of framed links. In their work, Frohman and Gelca gave a nice formula that describes the product in the skein space (algebra) $K(T^2)$ of the 2–torus $T^2 = S^1 \times S^1$. A standard embedding of T^2 in T^3 makes this product commutative; hence we can get further relations from the formula of Frohman and Gelca.

A natural question is the following:

Question 1.4 Is 9 the dimension of the skein vector space $K(T^3)$ of the 3–torus?

After this paper was submitted, P Gilmer [6] answered this question positively.

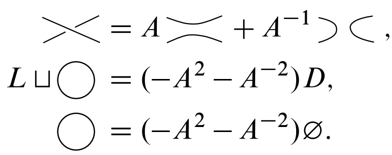
Acknowledgements The author is warmly grateful to Bruno Martelli for his constant support and encouragement.

2 The result

2.1 Definition of skein module

Let M be an oriented 3–manifold, R a commutative ring with unit and $A \in R$ an invertible element of R. Let V be the abstract free R–module generated by all framed links in M (considered up to isotopies) including the empty set $\varnothing$.

Definition 2.1 The (R, A)*–Kauffman bracket skein module* of M, or the R*–skein module*, or simply the *KBSM*, sometimes indicated with $\mathrm{KM}(M; R, A)$, is the quotient of V by all the possible *skein relations*:



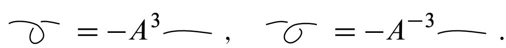
These are local relations where the framed links in an equation differ just in the pictured 3–ball that is equipped with a positive trivialization. An element of $\mathrm{KM}(M; R, A)$ is called a *skein* or a *skein element*. If M is the oriented I–bundle over a surface S (that is, $M = S \times [-1, 1]$ if S is oriented), we simply write $\mathrm{KM}(S; R, A)$ and call it the *skein module* of S.

Let $\mathbb{Q}(A)$ be field of all rational function with rational coefficients and abstract variable A. We set

$$K(M) := \mathrm{KM}(M; \mathbb{Q}(A), A),$$

and we call it the *skein vector space*, or simply the *skein space*, of M.

Remark 2.2 It is easy to verify that if we modify the framing of a component of a framed link, the skein changes by the multiplication of an integer power of $-A^3$:



2.2 The skein algebra of the 2–torus

Definition 2.3 Let S be a surface; the skein module $\mathrm{KM}(S; R, A)$ has a natural structure of an R–algebra that is given by the linear extension of the multiplication defined on framed links. Given two framed links $L_1, L_2 \subset S\times[-1,1]$, the product $L_1\cdot L_2\subset S\times[-1,1]$ is obtained by putting L_1 above L_2, so $L_1\cdot L_2\cap S\times[0,1]=L_1$ and $L_1\cdot L_2\cap S\times[-1,0]=L_2$.

Consider the 2–torus T^2 as the quotient of $\mathbb{R}^2$ modulo the standard lattice of translations generated by $(1,0)$ and $(0,1)$; hence for any nonzero pair (p,q) of integers, we have the notion of (p,q)*–curve*: the simple closed curve in the 2–torus that is the quotient of the line passing trough $(0,0)$ and (p,q).

Definition 2.4 Let p and q be two coprime integers; hence $(p,q)\neq(0,0)$. We denote by $(p,q)_T$ the (p,q)–curve in the 2–torus T^2 equipped with the blackboard framing. Given a framed knot γ in an oriented 3–manifold M and an integer $n\geq 0$, we denote by $T_n(\gamma)$ the skein element defined by induction as follows:

$$\begin{aligned}T_0(\gamma)&:=2\cdot\varnothing,\\ T_1(\gamma)&:=\gamma,\\ T_{n+1}(\gamma)&:=\gamma\cdot T_n(\gamma)-T_{n-1}(\gamma),\end{aligned}$$

where $\gamma\cdot T_n(\gamma)$ is the skein element obtained adding a copy of γ to all the framed links that compose the skein $T_n(\gamma)$. For $p,q\in\mathbb{Z}$ such that $(p,q)\neq(0,0)$, we denote by $(p,q)_T$ the skein element

$$(p,q)_T:=T_{\mathrm{MCD}(p,q)}\left(\left(\frac{p}{\mathrm{MCD}(p,q)},\frac{q}{\mathrm{MCD}(p,q)}\right)_T\right),$$

where $\mathrm{MCD}(p,q)$ is the maximum common divisor of p and q. Finally, we set

$$(0,0)_T:=2\cdot\varnothing.$$

It is easy to show that the set of all the skein elements $(p,q)_T$ with $p,q\in\mathbb{Z}$ generates $\mathrm{KM}(T^2;R,A)$ as R–module.

This is not the standard way to color framed links in a skein module. The colorings $\mathrm{JW}_n(\gamma)$, $n\geq 0$, with the Jones–Wenzl projectors are defined in the same way as $T_n(\gamma)$, but at the 0–level we have $\mathrm{JW}_0(\gamma)=\varnothing$.

Theorem 2.5 (Frohman and Gelca [5]) *For any $p,q,r,s\in\mathbb{Z}$, the following holds in the skein module $\mathrm{KM}(T^2;R,A)$ of the 2–torus T^2:*

$$(p,q)_T\cdot(r,s)_T=A^{\left|\begin{smallmatrix}p&q\\ r&s\end{smallmatrix}\right|}(p+r,q+s)_T+A^{-\left|\begin{smallmatrix}p&q\\ r&s\end{smallmatrix}\right|}(p-r,q-s)_T,$$

where $\left|\begin{smallmatrix}p&q\\ r&s\end{smallmatrix}\right|$ is the determinant $ps-qr$. □

2.3 The abelianization

Definition 2.6 Let B be a R–algebra for a commutative ring with unity R. We denote by $C(B)$ the R–module defined as the quotient

$$C(B) := \frac{B}{[B, B]},$$

where $[B, B]$ is the submodule of B generated by all the elements of the form $ab - ba$ for $a, b \in B$. We call $C(B)$ the *abelianization* of B.

Remark 2.7 Usually in noncommutative algebra, the *abelianization* is the R–algebra defined as the quotient of B modulo the subalgebra (submodule and ideal) generated by all the elements of the form $ab - ba$. In our definition, the denominator is just a submodule and we only get an R–module. We use the word "abelianization" anyway.

Now we work with $C(K(T^2))$, and we still use $(p,q)_T$ and $(p,q)_T \cdot (r,s)_T$ to denote the class of $(p,q)_T \in K(T^2)$ and $(p,q)_T \cdot (r,s)_T \in K(T^2)$ in $C(K(T^2))$.

Lemma 2.8 *Let (p,q) be a pair of integers different from $(0,0)$. Then in the abelianization $C(K(T^2))$ of the skein algebra $K(T^2)$ of the 2–torus T^2, we have*

$$(p,q)_T = \begin{cases} (1,0)_T & \text{if } p \in 2\mathbb{Z}+1 \text{ and } q \in 2\mathbb{Z}, \\ (0,1)_T & \text{if } p \in 2\mathbb{Z} \text{ and } q \in 2\mathbb{Z}+1, \\ (1,1)_T & \text{if } p, q \in 2\mathbb{Z}+1, \\ (2,0)_T & \text{if } p, q \in 2\mathbb{Z}. \end{cases}$$

Hence $C(K(T^2))$ is generated as a $\mathbb{Q}(A)$–vector space by the empty set $\varnothing$, the framed knots $(1,0)_T$, $(0,1)_T$, $(1,1)_T$, and a framed link consisting of two parallel copies of $(1,0)_T$.

Proof By Theorem 2.5, for every $p, q \in \mathbb{Z}$, we have

$$\begin{aligned} A^{-q}(p+2,q)_T + A^{q}(p,q)_T &= (p+1,q)_T \cdot (1,0)_T \\ &= (1,0)_T \cdot (p+1,q)_T \\ &= A^{q}(p+2,q)_T + A^{-q}(-p,-q)_T. \end{aligned}$$

Since $(p,q)_T = (-p,-q)_T$, we have $(A^q - A^{-q})(p,q)_T = (A^q - A^{-q})(p+2,q)_T$. Hence if $q \neq 0$, we get $(p,q)_T = (p+2,q)_T$ (here we use the fact that the base ring is a field and that $A^{2n} \neq 1$ for every $n > 0$). Thus

$$(p,q)_T = \begin{cases} (0,q)_T & \text{if } p \in 2\mathbb{Z} \text{ and } q \neq 0, \\ (1,q)_T & \text{if } p \in 2\mathbb{Z}+1 \text{ and } q \neq 0. \end{cases}$$

Analogously, by using $(0,1)_T$ instead of $(1,0)_T$ for $p \neq 0$, we get

$$(p,q)_T = \begin{cases} (p,0)_T & \text{if } q \in 2\mathbb{Z} \text{ and } q \neq 0, \\ (p,1)_T & \text{if } q \in 2\mathbb{Z}+1 \text{ and } q \neq 0. \end{cases}$$

Therefore, if $p, q \in 2\mathbb{Z}+1$, we have $(p,q)_T = (1,1)_T$. If $p \neq 0$, we get

$$(p,0)_T = (p,2)_T = \begin{cases} (0,2)_T & \text{if } p \in 2\mathbb{Z}, \\ (1,2)_T = (1,0)_T & \text{if } p \in 2\mathbb{Z}+1. \end{cases}$$

In the same way for $q \neq 0$, we get

$$(0,q)_T = (2,q)_T = \begin{cases} (2,0)_T & \text{if } p \in 2\mathbb{Z}, \\ (2,1)_T = (0,1)_T & \text{if } p \in 2\mathbb{Z}+1. \end{cases}$$

In particular, we have

$$(2,0)_T = (2,2)_T = (2,-2)_T = (0,2)_T = (p,q)_T \quad \text{for } (p,q) \neq (0,0),\ p,q \in 2\mathbb{Z}. \qquad \square$$

2.4 The (p,q,r)–type curves

As for the 2–torus T^2, we look at the 3–torus T^3 as the quotient of $\mathbb{R}^3$ modulo the standard lattice of translations generated by $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$.

Definition 2.9 Let (p,q,r) be a triple of coprime integers; that means we have $\mathrm{MCD}(p,q,r) = 1$, where $\mathrm{MCD}(p,q,r)$ is the maximum common divisor of p, q and r, and in particular, we have $(p,q,r) \neq (0,0,0)$. The *(p,q,r)–curve* is the simple closed curve in the 3–torus that is the quotient (under the standard lattice) of the line passing through $(0,0,0)$ and (p,q,r). We denote by $[p,q,r]$ the (p,q,r)–curve equipped with the framing that is the collar of the curve in the quotient of any plane containing $(0,0,0)$ and (p,q,r). The framing does not depend on the choice of the plane.

Definition 2.10 An embedding $e\colon T^2 \to T^3$ of the 2–torus in the 3–torus is *standard* if it is the quotient (under the standard lattice) of a plane in $\mathbb{R}^3$ that is the image of the plane generated by $(1,0,0)$ and $(0,1,0)$ under a linear map defined by a matrix of $\mathrm{SL}_3(\mathbb{Z})$ (a 3×3 matrix with integer entries and determinant 1).

Remark 2.11 There are infinitely many standard embeddings, even up to isotopies. A standard embedding of T^2 in T^3 is the quotient under the standard lattice of the plane generated by two columns of a matrix of $\mathrm{SL}_3(\mathbb{Z})$.

Lemma 2.12 *Let (p,q,r) be a triple of coprime integers. Then the skein element $[p,q,r] \in K(T^3)$ is equal to $[x,y,z]$, where $x,y,z \in \{0,1\}$ and they have respectively the same parities as p, q and r.*

Proof Every embedding $e\colon T^2 \to T^3$ of the 2–torus defines a linear map between the skein spaces

$$e_*\colon K(T^2) \to K(T^3).$$

The map e_* factorizes with the quotient map $K(T^2) \to C(K(T^2))$. In fact, we can slide the framed links in $e(T^2 \times [-1,1])$ from above to below, getting $e_*(L_1 \cdot L_2) = e_*(L_2 \cdot L_1)$ for every two framed links, L_1 and L_2, in $T^2 \times [-1,1]$. As said in Remark 2.11, a standard embedding $e\colon T^2 \to T^3$ corresponds to the plane generated by two columns $(p_1,q_1,r_1),(p_2,q_2,r_2) \in \mathbb{Z}^3$ of a matrix in $\mathrm{SL}_3(\mathbb{Z})$. In this correspondence, $e_*((a,b)_T) = [ap_1+bp_2, aq_1+bq_2, ar_1+br_2]$ for every coprime $a,b \in \mathbb{Z}$. Therefore, by Lemma 2.8, we get

$$\begin{aligned}[][a'p_1+b'p_2, a'q_1+b'q_2, a'r_1+b'r_2] &= e_*((a',b')_T)\\ &= e_*((a,b)_T)\\ &= [ap_1+bp_2, aq_1+bq_2, ar_1+br_2]\end{aligned}$$

for every two pairs $(a,b),(a',b') \in \mathbb{Z}^2$ of coprime integers such that $a+a', b+b' \in 2\mathbb{Z}$.

Let (p,q,r) be a triple of coprime integers. By permuting p, q and r, we get either $(p,q,r) = (1,0,0)$ or $p,q \neq 0$. Consider the case where $p,q \neq 0$. Let d be the maximum common divisor of p and q, and let $\lambda,\mu \in \mathbb{Z}$ such that $\lambda p + \mu q = d$. The following matrix belongs in $\mathrm{SL}_3(\mathbb{Z})$:

$$M_1 := \begin{pmatrix} p/d & -\mu & 0\\ q/d & \lambda & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Let $v_1^{(1)}$ and $v_3^{(1)}$ be the first and the third columns of M_1. We have $(p,q,r) = dv_1^{(1)} + rv_3^{(1)}$. Hence

$$[p,q,r] = \begin{cases} \left[\frac{p}{d},\frac{q}{d},0\right] & \text{if } d \in 2\mathbb{Z}+1 \text{ and } r \in 2\mathbb{Z},\\ [0,0,1] & \text{if } d \in 2\mathbb{Z} \text{ and } r \in 2\mathbb{Z}+1,\\ \left[\frac{p}{d},\frac{q}{d},1\right] & \text{if } d,r \in 2\mathbb{Z}+1.\end{cases}$$

The integers p,q,r cannot be all even because they are coprime; hence d and r cannot be both even. Therefore, we just need to study the cases where $r \in \{0,1\}$.

If $r=0$, we consider the trivial embedding of T^2 in T^3. The corresponding matrix of $\mathrm{SL}_3(\mathbb{Z})$ is the identity. We have $\left(\frac{p}{d},\frac{q}{d},0\right) = \frac{p}{d}(1,0,0) + \frac{q}{d}(0,1,0)$; hence

$$[p,q,0] = \left[\frac{p}{d},\frac{q}{d},0\right] = \begin{cases} [1,0,0] & \text{if } \frac{p}{d} \in 2\mathbb{Z}+1 \text{ and } \frac{q}{d} \in 2\mathbb{Z},\\ [0,1,0] & \text{if } \frac{p}{d} \in 2\mathbb{Z} \text{ and } \frac{q}{d} \in 2\mathbb{Z}+1,\\ [1,1,0] & \text{if } \frac{p}{d},\frac{q}{d} \in 2\mathbb{Z}+1.\end{cases}$$

If $r = 1$, we take the matrix of $\mathrm{SL}_3(\mathbb{Z})$

$$M_2 := \begin{pmatrix} 0 & 0 & 1 \\ q & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Let $v_1^{(2)}$ and $v_3^{(2)}$ be the first and the third columns of M_2. We have $(p, q, 1) = pv_3^{(2)} + v_1^{(2)}$; hence

$$[p, q, 1] = \begin{cases} [1, q, 1] & \text{if } p \in 2\mathbb{Z} + 1, \\ [0, q, 1] & \text{if } p \in 2\mathbb{Z}. \end{cases}$$

By permuting p, q and r, we reduce the case $(p, q, r) = (0, q, 1)$ to the case $p, q \neq 0$, $r = 0$ that we studied before.

It remains to consider the case $p = r = 1$. We consider the matrix of $\mathrm{SL}_3(\mathbb{Z})$

$$M_3 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Let $v_1^{(3)}$ and $v_2^{(3)}$ be the first and the second columns of M_3. We have $(1, q, 1) = v_1^{(3)} + qv_2^{(3)}$. Hence

$$[1, q, 1] = \begin{cases} [1, 0, 1] & \text{if } q \in 2\mathbb{Z}, \\ [1, 1, 1] & \text{if } q \in 2\mathbb{Z} + 1. \end{cases}$$

□

Lemma 2.13 *The intersection of any two different standardly embedded 2–tori in T^3 contains a (p, q, r)–type curve.*

Proof Let T_1 and T_2 be two standardly embedded tori in the 3–torus, and let π_1 and π_2 be two planes in $\mathbb{R}^3$ whose projections under the standard lattice are respectively T_1 and T_2. The intersection $T_1 \cap T_2$ contains the projection of $\pi_1 \cap \pi_2$. We just need to prove that in $\pi_1 \cap \pi_2$, there is a point $(p, q, r) \neq (0, 0, 0)$ with integer coordinates $p, q, r \in \mathbb{Z}$. Every plane defining a standardly embedded torus is generated by two vectors with integer coordinates, and hence it is described by an equation $ax + by + cz = 0$ with integer coefficients $a, b, c \in \mathbb{Z}$. Applying a linear map described by a matrix of $\mathrm{SL}_3(\mathbb{Z})$, we can suppose that π_1 is the trivial plane $\{z = 0\}$. Let $a, b, c \in \mathbb{Z}$ such that $\pi_2 = \{ax + by + cz = 0\}$. The vector $(-b, a, 0)$ is nonzero and lies on $\pi_1 \cap \pi_2$. □

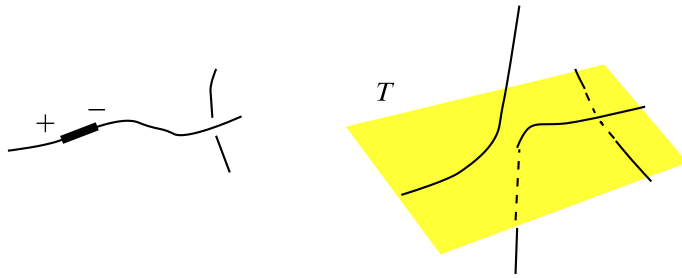


Figure 1: Diagrams of framed links in T^3. The plane is a part of the standardly embedded torus $T \subset T^3$ where the links project. If we look at the framed links in T^3 as framed tangles in $T \times [-1,1]$, the two strands that get out vertically from the plane end in the boundary points $(x,1)$ and $(x,-1)$ for some $x \in T$.

2.5 Diagrams

Framed links in T^3 can be represented by diagrams in the 2–torus T^2. These diagrams are like the usual link diagrams but with further oriented signs on the edges; see Figure 1 (left). Fix a standardly embedded 2–torus T in T^3. After a cut along a parallel copy T' of T, the 3–torus becomes diffeomorphic to $T \times [-1,1]$, and framed links in T^3 correspond to framed tangles of $T \times [-1,1]$. These diagrams are generic projections on T of the framed tangles in $T \times [-1,1]$ via the natural projection $(x,t) \mapsto x$. The further signs on the diagrams represent the intersection of the framed links with the boundary $T \times \{-1,1\}$. In other words, they represent the passages of the links along the (p,q,r)–type curve that, in the Euclidean metric, is orthogonal to T; see Figure 1 (right). If T is the trivial torus $S^1 \times S^1 \times \{x\}$, the further signs represent the passages through the third S^1–factor. We use the proper notion of blackboard framing.

2.6 Generators for the 3–torus

The following is the main theorem proved in this paper. We use all the previous lemmas to get a set of nine generators of $K(T^3)$.

Theorem 2.14 *The skein space $K(T^3)$ of the 3–torus T^3 is generated by the empty set $\varnothing$, $[1,0,0]$, $[0,1,0]$, $[0,0,1]$, $[1,1,0]$, $[1,0,1]$, $[0,1,1]$, $[1,1,1]$ and a skein α that is equal to the framed link consisting of two parallel copies of any (p,q,r)–type curve.*

Proof Let T be the trivial embedded 2–torus: the one containing the (p,q,r)–type curves with $r=0$. Use T to project the framed links and make diagrams. By using the first skein relation on these diagrams, we can see that $K(T^3)$ is generated by the framed links described by diagrams on T without crossings. These diagrams are unions of simple closed curves on T equipped with some signs as the one with $+$

and $-$ in Figure 1. These simple closed curves are either parallel to a (p,q)–curve or homotopically trivial. The framed links described by these diagrams lie in the standardly embedded tori that are the projections (under the standard lattice) of the planes generated by $(0,0,1)$ and $(p,q,0)$ for some p and q. Therefore, $K(T^3)$ is generated by the images of $K(T^2)$ under the linear maps induced by the standard embeddings of T^2 in T^3.

As said in the proof of Lemma 2.12, the linear map e_* induced by any standard embedding $e\colon T^2 \to T^3$ factorizes with the quotient map $K(T^2) \to C(K(T^2))$. Lemma 2.8 applied to the standard embedding e shows that the image $e_*(K(T^2))$ is generated by $\varnothing$, three (p,q,r)–type curves lying on $e(T^2)$, and the skein α_e that is equal to the framed link consisting of two parallel copies of any (p,q,r)–type curve lying on $e(T^2)$.

Let $e_1, e_2\colon T^2 \to T^3$ be two standard embeddings. By Lemma 2.13, $e_1(T^2) \cap e_2(T^2)$ contains a (p,q,r)–type curve γ; hence α_{e_1} and α_{e_2} coincide with the framed link that is two parallel copies of γ. Therefore, the skein element α_e does not depend on the embedding e.

We conclude by using Lemma 2.12, which says that the skein of any (p,q,r)–type curve is equal to the one of a standard representative of a nonzero element of the first homology group $H_1(T^3;\mathbb{Z}_2)$ with coefficient in $\mathbb{Z}_2$, namely a (p,q,r)–type curve with $p,q,r \in \{0,1\}$. □

Remark 2.15 Theorem 2.14, Lemma 2.8 and Lemma 2.12 work for every base pair (R,A) such that $A^{2n}-1$ is an invertible element of R for any $n>0$. In particular, they work for $(\mathbb{C},A)$, where $A^{2n} \neq 1$ for any $n>0$. Unfortunately, we do not know what happens with the base pair $(\mathbb{C},\pm 1)$, which is the one used for the connection with the $\mathrm{SL}_2(\mathbb{C})$–character variety [3]. In fact, in Lemma 2.8, we would get just trivial equalities if $A=\pm 1$.

2.7 Linear independence

Here we talk about the linear independence of our generators of $K(T^2)$. The following proposition shows a direct sum decomposition of $K(T^3)$.

Proposition 2.16 *The skein space $K(T^3)$ is the direct sum of eight subspaces,*

$$K(T^3) = V_0 \oplus V_1 \oplus \cdots \oplus V_7,$$

such that

1. *V_0 is generated by the empty set $\varnothing$ and the skein α (see Theorem 2.14);*
2. *every (p,q,r)–type curve generates a V_j with $j>0$, and every V_j with $j>0$ is generated by one such curve.*

Proof The skein relations relate framed links in the same $\mathbb{Z}_2$–homology class. Hence for every oriented 3–manifold M, we have a direct sum decomposition

$$\mathrm{KM}(M; R, A) = \bigoplus_{h \in H_1(M;\mathbb{Z}_2)} V_h,$$

where V_h is generated by the framed links whose $\mathbb{Z}_2$–homology class is h. The statement follows by this observation and the fact that if $[p, q, r]$ and $[p', q', r']$ represent the same $\mathbb{Z}_2$–homology class, then $[p, q, r] = [p', q', r'] \in K(T^3)$. □

Remark 2.17 Given a triple of integers $(x, y, z) \neq (0, 0, 0)$ such that $x, y, z \in \{0, 1\}$, we can easily find an orientation-preserving diffeomorphism of the 3–torus T^3 sending $[x, y, z]$ to $[1, 0, 0]$. Hence if the skein of one such curve $[x, y, z]$ is null, then also all the other skein elements of such curves are null. Therefore, by Proposition 2.16, the possible dimensions of the skein space $K(T^3)$ are 0, 1, 2, 7, 8 and 9.

After the submission of this paper, P Gilmer [6] showed that the skein of the $(1, 0, 0)$–curve is not null and that the empty set and the skein α are linear independent. This answers Question 1.4 in the affirmative by proving that the set of nine generators is actually a basis for the skein space.

References

[1] **C Blanchet**, **N Habegger**, **G Masbaum**, **P Vogel**, *Topological quantum field theories derived from the Kauffman bracket*, Topology 34 (1995) 883–927 MR

[2] **D Bullock**, *The $(2, \infty)$–skein module of the complement of a $(2, 2p + 1)$ torus knot*, J. Knot Theory Ramifications 4 (1995) 619–632 MR

[3] **D Bullock**, *Rings of $\mathrm{SL}_2(\mathbf{C})$–characters and the Kauffman bracket skein module*, Comment. Math. Helv. 72 (1997) 521–542 MR

[4] **F Costantino**, *Coloured Jones invariants of links and the volume conjecture*, J. Lond. Math. Soc. 76 (2007) 1–15 MR

[5] **C Frohman**, **R Gelca**, *Skein modules and the noncommutative torus*, Trans. Amer. Math. Soc. 352 (2000) 4877–4888 MR

[6] **P M Gilmer**, *On the Kauffman bracket skein module of the 3–torus*, preprint (2016) arXiv To appear in Indiana J. of Math.

[7] **J Hoste**, **J H Przytycki**, *The $(2, \infty)$–skein module of lens spaces; a generalization of the Jones polynomial*, J. Knot Theory Ramifications 2 (1993) 321–333 MR

[8] **J Hoste**, **J H Przytycki**, *The $(2, \infty)$–skein module of Whitehead manifolds*, J. Knot Theory Ramifications 4 (1995) 411–427 MR

[9] **J Hoste**, **J H Przytycki**, *The Kauffman bracket skein module of $S^1 \times S^2$*, Math. Z. 220 (1995) 65–73 MR

[10] **L H Kauffman**, *New invariants in the theory of knots*, Amer. Math. Monthly 95 (1988) 195–242 MR

[11] **L H Kauffman**, **S L Lins**, *Temperley–Lieb recoupling theory and invariants of 3–manifolds*, Annals of Mathematics Studies 134, Princeton Univ. Press (1994) MR

[12] **W B R Lickorish**, *Three-manifolds and the Temperley–Lieb algebra*, Math. Ann. 290 (1991) 657–670 MR

[13] **W B R Lickorish**, *Calculations with the Temperley–Lieb algebra*, Comment. Math. Helv. 67 (1992) 571–591 MR

[14] **W B R Lickorish**, *The skein method for three-manifold invariants*, J. Knot Theory Ramifications 2 (1993) 171–194 MR

[15] **W B R Lickorish**, *Skeins and handlebodies*, Pacific J. Math. 159 (1993) 337–349 MR

[16] **J Marché**, *The skein module of torus knots*, Quantum Topol. 1 (2010) 413–421 MR

[17] **M Mroczkowski**, *Kauffman bracket skein module of a family of prism manifolds*, J. Knot Theory Ramifications 20 (2011) 159–170 MR

[18] **M Mroczkowski**, *Kauffman bracket skein module of the connected sum of two projective spaces*, J. Knot Theory Ramifications 20 (2011) 651–675 MR

[19] **M Mroczkowski**, **M K Dabkowski**, *KBSM of the product of a disk with two holes and S^1*, Topology Appl. 156 (2009) 1831–1849 MR

[20] **J H Przytycki**, *Skein modules of 3–manifolds*, Bull. Polish Acad. Sci. Math. 39 (1991) 91–100 MR

[21] **J H Przytycki**, *Fundamentals of Kauffman bracket skein modules*, Kobe J. Math. 16 (1999) 45–66 MR

[22] **J H Przytycki**, *Kauffman bracket skein module of a connected sum of 3–manifolds*, Manuscripta Math. 101 (2000) 199–207 MR

[23] **J H Przytycki**, *Skein modules*, preprint (2006) arXiv

[24] **V G Turaev**, *The Conway and Kauffman modules of a solid torus*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 167 (1988) 79–89 MR In Russian; translated in J. Soviet Math. 52 (1990) 2799–2805

Dipartimento di Matematica "Tonelli", Università di Pisa
Pisa, Italy

carrega@mail.dm.unipi.it

Received: 28 April 2016 Revised: 8 April 2017

Algebraic & Geometric Topology 17 (2017) 3461–3518

Quasistabilization and basepoint moving maps in link Floer homology

IAN ZEMKE

We analyze the effect of adding, removing, and moving basepoints on link Floer homology. We prove that adding or removing basepoints via a procedure called quasistabilization is a natural operation on a certain version of link Floer homology, which we call CFL^∞_{UV}. We consider the effect on the full link Floer complex of moving basepoints, and develop a simple calculus for moving basepoints on the link Floer complexes. We apply it to compute the effect of several diffeomorphisms corresponding to moving basepoints. Using these techniques we prove a conjecture of Sarkar about the map on the full link Floer complex induced by a finger move along a link component.

57M25, 57M27, 57R58

1 Introduction

Introduced by Ozsváth and Szabó, Heegaard Floer homology associates algebraic invariants to closed three-manifolds. To a three-manifold Y with embedded nullhomologous knot K, there is a refinement of Heegaard Floer homology called knot Floer homology, introduced by Ozsváth and Szabó [8] and independently by Rasmussen [11]. A similar invariant was defined by Ozsváth and Szabó for links [10].

To a nullhomologous knot $K \subseteq Y$ with two basepoints z and w and a relative Spin^c structure $\mathfrak{t} \in \mathrm{Spin}^c(Y, K)$, Ozsváth and Szabó [8] define a $\mathbb{Z} \oplus \mathbb{Z}$–filtered chain complex $\mathrm{CFK}^\infty(Y, K, w, z, \mathfrak{t})$. The $\mathbb{Z} \oplus \mathbb{Z}$–filtered chain homotopy type of $\mathrm{CFK}^\infty(Y, K, w, z, \mathfrak{t})$ is an invariant of the data $(Y, K, w, z, \mathfrak{t})$.

One of the nuances of Heegaard Floer homology is the dependence on basepoints. In the case of closed three-manifolds, if $\boldsymbol{w} \subseteq Y$ is a collection of basepoints, $w \in \boldsymbol{w}$ and γ is a curve in $\pi_1(Y, w)$, then one can consider the diffeomorphism ϕ_γ resulting from a finger move along γ. According to Juhász and Thurston [4], the based mapping class group $\mathrm{MCG}(Y, w)$ acts on $\mathrm{CF}^\circ(Y, \boldsymbol{w}, \mathfrak{s})$ and hence there is an induced map $(\phi_\gamma)_*$ on the closed three-manifold invariant $\mathrm{CF}^\circ(Y, \boldsymbol{w}, \mathfrak{s})$, which is a $\mathbb{Z}_2[U]$–equivariant chain

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3461

homotopy type. In [14], the author computes the equivariant chain homotopy type of $(\phi_\gamma)_*$ to be

$$(\phi_\gamma)_* \simeq \mathrm{id} + [\gamma] \circ \Phi_w,$$

where $[\gamma]$ is the $\Lambda^*(H_1(Y;\mathbb{Z})/\operatorname{Tors})$ action and Φ_w is an analogue of a map appearing also on link Floer homology, which we describe below.

In this paper, we consider the analogous question about basepoint dependence for link Floer homology. In link Floer homology, the basepoints are constrained to be on the link component, so the analogous operation is to consider the map on link Floer homology induced by a finger move ς around a link component, in the positive direction according to the link's orientation. Using grid diagrams, Sarkar [13] computes the map associated to the diffeomorphism ς on a certain version of link Floer homology (the associated graded complex) for links in S^3. For links in arbitrary three-manifolds, and for the induced map on the full link Floer complex, he conjectures the formula. We prove his formula in full generality (Theorem B), but before we state that theorem we will provide a brief description of the complexes and maps which appear.

We will work with a slightly different version of CFL^∞ than the one which most often appears in the literature. For a multibased link $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$ inside of Y and a Spin^c structure $\mathfrak{s} \in \mathrm{Spin}^c(Y)$, we construct a chain complex

$$\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \mathfrak{s}),$$

which is a module over the polynomial ring $\mathbb{Z}_2[U_{\boldsymbol{w}}, V_{\boldsymbol{z}}]$, generated by variables U_w with $w \in \boldsymbol{w}$ and V_z with $z \in \boldsymbol{z}$. The module $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \mathfrak{s})$ has generators of the form

$$\boldsymbol{x} \cdot U_{\boldsymbol{w}}^I V_{\boldsymbol{z}}^J = \boldsymbol{x} \cdot U_{w_1}^{i_1} \cdots U_{w_n}^{i_n} V_{z_1}^{j_1} \cdots V_{z_n}^{j_n}$$

for multi-indices $I = (i_1, \dots, i_n)$ and $J = (j_1, \dots, j_n)$, though we identify two variables V_z and $V_{z'}$ if z and z' are on the same link component. Thus, $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \mathfrak{s})$ has a filtration by $\mathbb{Z}^{|\boldsymbol{w}|} \oplus \mathbb{Z}^{|L|}$ given by filtering over powers of the variables, where $|L|$ denotes the set of components of L.

As with the free stabilization maps from [14], to define functorial maps corresponding to adding or removing basepoints in link Floer homology, we must work with colored complexes. A coloring $(\sigma, \mathfrak{P})$ of a link with basepoints, $(L, \boldsymbol{w}, \boldsymbol{z})$, is a set $\mathfrak{P}$ indexing a collection of formal variables, together with a map $\sigma\colon \boldsymbol{w} \cup \boldsymbol{z} \to \mathfrak{P}$ which maps all $\boldsymbol{z}$–basepoints on a component of L to the same color. Given a coloring $(\sigma, \mathfrak{P})$ of a link $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$, we create a $\mathbb{Z}_2[U_{\mathfrak{P}}]$–chain complex

$$\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s}).$$

The powers of the $U_{\mathfrak{P}}$ variables yield a filtration by $\mathbb{Z}^{\mathfrak{P}}$, which we call the $\mathfrak{P}$–filtration.

In the context of link Floer homology, analogously to adding or removing a free basepoint in a closed 3–manifold, one can add or remove a pair of adjacent basepoints, w and z, on a link component. Some authors refer to this operation as a "special stabilization". Manolescu and Ozsváth [6] consider certain questions about the operation, calling it "quasistabilization", which is the phrase we will use. A full description and proof of naturality of the operation has not been completed, so we do that in this paper:

Theorem A *Suppose z and w are new basepoints on a link $\mathbb{L}=(L,\boldsymbol{w},\boldsymbol{z})$, ordered so that w comes after z, which aren't separated by any basepoints in $\boldsymbol{w}$ or $\boldsymbol{z}$. If $\sigma\colon \boldsymbol{w}\cup\boldsymbol{z}\to\mathfrak{P}$ is a coloring which is extended by $\sigma'\colon \boldsymbol{w}\cup\boldsymbol{z}\cup\{w,z\}\to\mathfrak{P}$, then there are $\mathfrak{P}$–filtered $\mathbb{Z}_2[U_{\mathfrak{P}}]$–chain maps*

$$S^+_{w,z}\colon \mathrm{CFL}^\infty_{UV}(Y,L,\boldsymbol{w},\boldsymbol{z},\sigma,\mathfrak{P},\mathfrak{s})\to \mathrm{CFL}^\infty_{UV}(Y,L,\boldsymbol{w}\cup\{w\},\boldsymbol{z}\cup\{z\},\sigma',\mathfrak{P},\mathfrak{s})$$

and

$$S^-_{w,z}\colon \mathrm{CFL}^\infty_{UV}(Y,L,\boldsymbol{w}\cup\{w\},\boldsymbol{z}\cup\{z\},\sigma',\mathfrak{P},\mathfrak{s})\to \mathrm{CFL}^\infty_{UV}(Y,L,\boldsymbol{w},\boldsymbol{z},\sigma,\mathfrak{P},\mathfrak{s}),$$

which are well-defined invariants, up to $\mathfrak{P}$–filtered, $\mathbb{Z}_2[U_{\mathfrak{P}}]$–equivariant chain homotopy. If z comes after w, there are maps $S^+_{z,w}$ and $S^-_{z,w}$ defined analogously.

Following Sarkar [13], we consider endomorphisms Φ_w and Ψ_z of $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$ ($\Phi_{i,j}$ and $\Psi_{i,j}$ in his notation). We can think of the maps Φ_w and Ψ_z as formal derivatives of the differential ∂ with respect to the variables U_w and V_z, respectively. The maps Φ_w and Ψ_z are invariants of $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s})$ up to $\mathfrak{P}$–filtered chain homotopy.

The maps Ψ_z can be thought of as analogues of the relative homology maps A_λ defined in [14] for the closed three-manifold invariants, since they play the role in the basepoint moving maps for link Floer homology that the relative homology maps introduced in [14] played in the basepoint moving maps for the closed three-manifolds invariants. Indeed the objects $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$ and the maps Ψ_z and $S^\pm_{w,z}$ fit into the framework of a "graph TQFT" for surfaces embedded in four-manifolds with some extra decoration, similar to the TQFT for $\widehat{\mathrm{HFL}}$ constructed using sutured Floer homology by Juhász [2] and considered further by Juhász and Marengon [3] for concordances. Such a TQFT construction for CFL^∞_{UV} will appear in a future paper.

We finally state Sarkar's conjecture, cast into the framework of $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s})$:

Theorem B *Suppose that $\mathbb{L}=(L,\boldsymbol{w},\boldsymbol{z})$ is a multibased link in an arbitrary 3–manifold Y and K is a component of L. Suppose that the basepoints on K are $z_1,w_1,\dots,z_n,w_n$. Letting ς denote the diffeomorphism resulting from a finger move* *around a link component K, the induced map ς_* on* $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s})$ *has the* $\mathfrak{P}$*–filtered equivariant chain homotopy type*

$$\varsigma_* \simeq \mathrm{id} + \Phi_K\Psi_K,$$

where

$$\Phi_K = \sum_{j=1}^n \Phi_{w_j} \quad \textit{and} \quad \Psi_K = \sum_{j=1}^n \Psi_{z_j}.$$

Sarkar's conjecture for the effect on the filtered link Floer complex, which we will denote by $\mathrm{CFL}^\infty(Y,\mathbb{L},\mathfrak{t})$ for $\mathfrak{t}$ a relative $\mathrm{Spin}^c(Y,L)$ structure, follows by setting $(\sigma,\mathfrak{P})$ to be the trivial coloring (ie $\mathfrak{P} = \boldsymbol{w}\cup|L|$ and $\sigma\colon (\boldsymbol{w}\cup\boldsymbol{z})\to(\boldsymbol{w}\cup|L|)$ the natural map) since the complex $\mathrm{CFL}^\infty(Y,\mathbb{L},\mathfrak{t})$ becomes a $\mathbb{Z}_2$–subcomplex of $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s})$ which is preserved by ς_*, where $\mathfrak{s}$ is the underlying Spin^c structure associated to the relative Spin^c structure $\mathfrak{t}$.

There are several other formulations of this conjecture for different versions of link Floer homology. For example, the conjectured formula for ς_* on $\mathrm{CFK}^\infty(S^3,K)$ for $K\subseteq S^3$ is useful for computations in the involutive Heegaard Floer homology theory developed by Hendricks and Manolescu (see [1, Section 6]). In their notation, for a choice of diagrams, the complex $\mathrm{CFK}^\infty(S^3,K)$ for a knot $K\subseteq S^3$ is generated by elements of the form $[\boldsymbol{x},i,j]$ where i and j satisfy $A(\boldsymbol{x}) = i-j$, and A denotes the Alexander grading. In their notation, the U map takes the form $U\cdot[\boldsymbol{x},i,j] = [\boldsymbol{x},i-1,j-1]$. Again, the complex $\mathrm{CFK}^\infty(S^3,K)$ is a $\mathbb{Z}_2$–subcomplex

$$\mathrm{CFK}^\infty(S^3,K)\subseteq \mathrm{CFL}^\infty_{UV}(S^3,K,w,z,\mathfrak{s}_0)$$

which is preserved by ς_*. Recasting Theorem B into this notation and recalling that we are using coefficients in $\mathbb{Z}_2$, we arrive at the following:

Corollary C *For a knot $K\subseteq S^3$, the involution ς_* on* $\mathrm{CFK}^\infty(S^3,K)$ *takes the form*

$$\varsigma_* \simeq 1 + U^{-1}\Bigg(\sum_{\substack{i,j\ge 0\\ i\text{ odd}}} \partial_{ij}\Bigg)\circ\Bigg(\sum_{\substack{i,j\ge 0\\ j\text{ odd}}} \partial_{ij}\Bigg),$$

if we write the differential $\partial = \sum_{i,j\ge 0}\partial_{ij}$. Here ∂_{ij} decreases the first filtration by i and the second filtration by j.

For other flavors, such as $\widehat{\mathrm{CFL}}$ or CFL^-, the formula conjectured by Sarkar also follows, since those cases correspond to setting various variables equal to zero in the formula for ς_*.

In addition, we consider the effect of another diffeomorphism obtained by twisting a link component. Suppose that K is a component of a link $\mathbb{L}$ and suppose that the basepoints of K are $z_1, w_1, \dots, z_n$ and w_n, appearing in that order. We can consider the diffeomorphism $\tau\colon (Y,\mathbb{L}) \to (Y,\mathbb{L})$ which twists $(1/n)^{\text{th}}$ of the way around K. The diffeomorphism τ maps z_i and w_i to z_{i+1} and w_{i+1}, respectively, with indices taken modulo n. If $(\sigma,\mathfrak{P})$ is a coloring of $\mathbb{L}$ which sends all of the $\boldsymbol{w}$–basepoints on K to the same color, then τ naturally induces an automorphism of

$$\mathrm{CFL}^{\infty}_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s}).$$

Using the techniques of this paper, we can compute the following:

Theorem D *Suppose that $\mathbb{L}$ is an embedded link in Y, and K is a component of $\mathbb{L}$ with basepoints $z_1, w_1, \dots, z_n$ and w_n, appearing in that order. Assume that $n > 1$. If τ denotes the diffeomorphism induced by twisting $(1/n)^{th}$ of the way around K, then for a coloring where all $\boldsymbol{w}$–basepoints on K have the same color, we have*

$$\tau_* \simeq (\Psi_{z_1}\Phi_{w_1}\Psi_{z_2}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{z_n}\Phi_{w_n}) + (\Phi_{w_1}\Psi_{z_2}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{z_n}).$$

Organization In Section 2 we define the complexes which will appear in this paper, as well as their algebraic structures as $\mathfrak{P}$–filtered chain complexes over certain modules. In Section 3 we define the maps Φ_w and Ψ_z which feature prominently in this paper. In Sections 5–7 we define quasistabilization maps $S^{\pm}_{w,z}$ and show that they are independent of the choice of diagrams and auxiliary data, proving Theorem A. In Sections 8 and 9 we prove useful relations amongst the maps Ψ_z, Φ_w and $S^{\pm}_{w,z}$. In Section 10 we compute several maps associated with moving basepoints, proving Theorems B and D.

Acknowledgments I would like to thank my advisor, Ciprian Manolescu, for helpful conversations, especially about quasistabilization. I would also like to thank Faramarz Vafaee and Robert Lipshitz for helpful conversations.

2 Background, the complexes $\mathrm{CFL}^{\infty}_{UV}$, and $\mathfrak{P}$–filtrations

In this section we provide some background and describe the complexes $\mathrm{CFL}^{\infty}_{UV}(Y,\mathbb{L},\mathfrak{s})$ which will appear.

2.1 Spin^c structures and relative Spin^c structures

Ozsváth and Szabó [9] define a Spin^c structure on Y to be a homology class of nonvanishing vector fields on Y. For a Heegaard diagram $\mathcal{H} = (\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{w})$ they define a map

$$\mathfrak{s}_{\boldsymbol{w}}\colon \mathbb{T}_\alpha \cap \mathbb{T}_\beta \to \mathrm{Spin}^c(Y).$$

The vector field $\mathfrak{s}_{\boldsymbol{w}}(\boldsymbol{x})$ is obtained by taking an upward gradient-like vector field associated to a Morse function yielding $\mathcal{H}$, and modifying it in a neighborhood of the flowlines passing through the points in $\boldsymbol{w}$ and $\boldsymbol{x}$ to obtain a nonvanishing vector field. Ozsváth and Szabó [10] provide a notion of relative Spin^c structures for a link in a 3–manifold. These are homology classes of vector fields on $Y \setminus N(L)$ which are tangent to the torus $\partial N(L)$. They define a map

$$\mathfrak{s}_{\boldsymbol{w},\boldsymbol{z}}\colon \mathbb{T}_\alpha \cap \mathbb{T}_\beta \to \mathrm{Spin}^c(Y, L).$$

We note that in general there are two natural ways to obtain an absolute Spin^c structure from a relative Spin^c structure. We take the convention that the filling map covers the map $\mathfrak{s}_{\boldsymbol{w}}$ (compare [10, Section 3.7]). In more generality, one has $\mathfrak{s}_{\boldsymbol{w}}(\boldsymbol{x}) - \mathfrak{s}_{\boldsymbol{z}}(\boldsymbol{x}) = \mathrm{PD}[L]$, so if we restrict to links whose total homology class vanishes, then there is no distinction. Since we include the versions of link Floer homology which use relative Spin^c structures only for the sake of comparison, whenever we consider relative Spin^c structures, we will assume that Y is a integer homology sphere. We will primarily be interested in working with the version CFL^∞_{UV}, which uses absolute Spin^c structures.

2.2 The complex $\mathbf{CFL}^\infty_{UV}(Y, \mathbb{L}, \mathfrak{s})$

Here we describe the uncolored complex $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \mathfrak{s})$. We first describe an intermediate object, $\mathrm{CFL}^\infty_{UV,0}(Y, \mathbb{L}, \mathfrak{s})$.

Let $\mathbb{Z}_2[U_{\boldsymbol{w}}, U_{\boldsymbol{w}}^{-1}, V_{\boldsymbol{z}}, V_{\boldsymbol{z}}^{-1}]$ denote the ring generated by variables U_w, V_z and their inverses U_w^{-1}, V_z^{-1} for $w \in \boldsymbol{w}$ and $z \in \boldsymbol{z}$. Given a diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ for $(Y, L, \boldsymbol{w}, \boldsymbol{z})$, we define $\mathrm{CFL}^\infty_{UV,0}(\mathcal{H}, \mathfrak{s})$ to be the free $\mathbb{Z}_2[U_{\boldsymbol{w}}, U_{\boldsymbol{w}}^{-1}, V_{\boldsymbol{z}}, V_{\boldsymbol{z}}^{-1}]$–module generated by $\boldsymbol{x} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta$ with $\mathfrak{s}_{\boldsymbol{w}}(\boldsymbol{x}) = \mathfrak{s}$. We refer the reader to eg [10] for the definition of a Heegaard diagram for a link, though we emphasize that in light of the results of [4], we must assume that

$$\boldsymbol{w} \cup \boldsymbol{z} \subseteq \Sigma \subseteq Y$$

and that the embedding of Σ in Y is part of the data of a Heegaard splitting.

We now define a map

$$\partial\colon \mathrm{CFL}^\infty_{UV,0}(\mathcal{H}, \mathfrak{s}) \to \mathrm{CFL}^\infty_{UV,0}(\mathcal{H}, \mathfrak{s})$$

by

$$\partial(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \mathbb{T}_\alpha \cap \mathbb{T}_\beta} \sum_{\substack{\phi \in \pi_2(\boldsymbol{x}, \boldsymbol{y}) \\ \mu(\phi) = 1}} \# \widehat{\mathcal{M}}(\phi) U_{\boldsymbol{w}}^{n_{\boldsymbol{w}}(\phi)} V_{\boldsymbol{z}}^{n_{\boldsymbol{z}}(\phi)} \cdot \boldsymbol{y}.$$

The map ∂ does not square to zero, but we do have the following:

Lemma 2.1 *The map* $\partial\colon \mathrm{CFL}^\infty_{UV,0}(\mathcal{H},\mathfrak{s}) \to \mathrm{CFL}^\infty_{UV,0}(\mathcal{H},\mathfrak{s})$ *satisfies*

$$\partial^2 = \sum_{K\in|L|} (U_{w_{K,1}} V_{z_{K,1}} + V_{z_{K,1}} U_{w_{K,2}} + U_{w_{K,2}} V_{z_{K,2}} + \cdots + V_{z_{K,n_K}} U_{w_{K,1}}),$$

where $w_{K,1}, z_{K,1}, \dots, w_{K,n_K}, z_{K,n_K}$ *are the basepoints on the link component* K*, in the order that they appear on* K*.*

Proof This follows from the usual proof that the differential squares to zero, now just counting boundary degenerations carefully. If there are exactly two basepoints, there are no boundary degenerations by [10, Theorem 5.5], and the above formula is satisfied. If there are more than two, then each $\boldsymbol{\alpha}$– and $\boldsymbol{\beta}$–degeneration has a unique holomorphic representative by [10, Theorem 5.5] and each crosses over a $\boldsymbol{w}$–basepoint and a $\boldsymbol{z}$–basepoint. The formula follows. □

To get a chain complex, we must color $\mathrm{CFL}^\infty_{UV,0}(Y,\mathbb{L},\mathfrak{s})$ by setting certain variables equal. Let $C_{\mathbb{L}}$ denote the ideal generated by elements of the form $V_{z_i} - V_{z_j}$, where z_i and z_j are in the same link component. We let $\mathcal{L} = \mathbb{Z}_2[U_{\boldsymbol{w}}, U_{\boldsymbol{w}}^{-1}, V_{\boldsymbol{z}}, V_{\boldsymbol{z}}^{-1}]/C_{\mathbb{L}}$.

We now define

$$\mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s}) = \mathrm{CFL}^\infty_{UV,0}(\mathcal{H},\mathfrak{s}) \otimes_{\mathbb{Z}_2[U_{\boldsymbol{w}},U_{\boldsymbol{w}}^{-1},V_{\boldsymbol{z}},V_{\boldsymbol{z}}^{-1}]} \mathcal{L}.$$

We have the following:

Lemma 2.2 *The map* ∂ *defined above is a differential on* $\mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s})$*, ie* $\partial^2 = 0$*.*

Proof This follows from the formula in Lemma 2.1 since the module $\mathcal{L}$ simply identifies all V_z variables for z which lie in the same link component. □

Remark 2.3 There are other modules that we could tensor with to make the differential square to zero. The module $\mathcal{L}$ is actually a quite natural choice. As we will see in the proof of Proposition 5.3, terms of the form $V_z + V_{z'}$ appear in the differential after quasistabilization, and these terms must be zero for $S^{\pm}_{w,z}$ to be chain maps.

The $\mathbb{Z}_2[U_{\boldsymbol{w}}, V_{\boldsymbol{z}}]$–module $\mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s})$ has a natural $\mathbb{Z}^{|\boldsymbol{w}|} \oplus \mathbb{Z}^{|L|}$ filtration given by filtering over powers of the variables $U_{\boldsymbol{w}}$ and $V_{\boldsymbol{z}}$.

There are of course many different Heegaard diagrams $\mathcal{H}$ for a given multibased link $(L, \boldsymbol{w}, \boldsymbol{z})$. As in the case of closed three-manifolds, using [4], given two diagrams $\mathcal{H}$ and $\mathcal{H}'$, there is a $\mathbb{Z}^{|\boldsymbol{w}|} \oplus \mathbb{Z}^{|L|}$–filtered map

$$\Phi_{\mathcal{H}\to\mathcal{H}'}\colon \mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s}) \to \mathrm{CFL}^\infty_{UV}(\mathcal{H}',\mathfrak{s})$$

which is a filtered chain homotopy equivalence, and is an invariant up to $\mathbb{Z}^{|\boldsymbol{w}|} \oplus \mathbb{Z}^{|L|}$–filtered chain homotopy. The maps $\Phi_{\mathcal{H}\to\mathcal{H}'}$ are functorial in the sense that if $\mathcal{H}, \mathcal{H}'$ and $\mathcal{H}''$ are three diagrams, then

$$\Phi_{\mathcal{H}'\to\mathcal{H}''} \circ \Phi_{\mathcal{H}\to\mathcal{H}'} \simeq \Phi_{\mathcal{H}\to\mathcal{H}''}.$$

The strongest invariant, which we will occasionally refer to as the *coherent filtered chain homotopy type*, is the collection of all of the complexes $\mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s})$ for all admissible diagrams $\mathcal{H}$ for $(Y, L, \boldsymbol{w}, \boldsymbol{z})$, as well as the maps $\Phi_{\mathcal{H}\to\mathcal{H}'}$. We let

$$\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$$

denote this invariant. Note that since we are working with embedded Heegaard surfaces, the set of Heegaard diagrams for a link is a set, and not a proper class.

Remark 2.4 As we remarked earlier, since $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$ is generated by intersection points with $\mathfrak{s}_{\boldsymbol{w}}(\boldsymbol{x}) = \mathfrak{s}$, there is some asymmetry between the $\boldsymbol{w}$ and $\boldsymbol{z}$ basepoints in the construction of CFL^∞_{UV}. We note that $\mathfrak{s}_{\boldsymbol{w}}(\boldsymbol{x}) - \mathfrak{s}_{\boldsymbol{z}}(\boldsymbol{x}) = \mathrm{PD}[L]$, so if L is null-homologous, this doesn't affect the chain complexes. As a toy example, one can consider $(S^1\times S^2, S^1\times\{\mathrm{pt}\})$ to see how the modules change over different choices of $\mathfrak{s}$.

2.3 Other versions of the link Floer complex

Supposing for simplicity that Y is an integer homology sphere, we briefly describe a complex $\mathrm{CFL}^\infty(Y,\mathbb{L},\mathfrak{t})$, for a relative Spin^c structure $\mathfrak{t}\in\mathrm{Spin}^c(Y,L)$. It will not feature in any of the sections after this, but we describe it as a comparison with CFL^∞_{UV}. Let μ_i be a positive meridian of the i^{th} link component. The complex $\mathrm{CFL}^\infty(Y,\mathbb{L},\mathfrak{t})$ is defined as the subcomplex of $\mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s})$ generated over $\mathbb{Z}_2$ by elements of the form $\boldsymbol{x}\cdot U_{\boldsymbol{w}}^{\boldsymbol{I}} V_{\boldsymbol{z}}^{\boldsymbol{J}}$, where

$$J\cdot\mathrm{PD}[M] = (\mathfrak{t} - \mathfrak{s}_{\boldsymbol{w},\boldsymbol{z}}(\boldsymbol{x})) + I\cdot\mathrm{PD}[M], \tag{1}$$

where PD denotes Poincaré duality. Here, if $J=(j_1,\dots,j_\ell)$, then $J\cdot\mathrm{PD}[M]$ is defined to be $j_1\cdot\mathrm{PD}[\mu_1]+\cdots+j_\ell\cdot\mathrm{PD}[\mu_\ell]$, and $I\cdot\mathrm{PD}[M]$ is defined similarly.

In the case that $\mathbb{L}=(K,w,z)$ is a knot with exactly two basepoints, we see that $\mathrm{CFL}^\infty(Y,K,w,z,\mathfrak{t})$ is generated by elements of the form $\boldsymbol{x}\cdot U_w^i V_z^j$ with

$$j\cdot\mathrm{PD}[\mu] = (\mathfrak{t}-\mathfrak{s}_{\boldsymbol{w},\boldsymbol{z}}(\boldsymbol{x})) + i\cdot\mathrm{PD}[\mu],$$

which is exactly the complex $\mathrm{CFK}^\infty(K,\mathfrak{t})$ found in [8]. More often one writes $[\boldsymbol{x},i,j]$ for what we write $\boldsymbol{x}\cdot U_w^{-i}V_z^{-j}$. Most authors also write U for the action defined by $U\cdot[\boldsymbol{x},i,j]=[\boldsymbol{x},i-1,j-1]$, though in our notation this action corresponds to multiplication by U_wV_z. It's also common to consider an object $\mathrm{CFK}^\infty(Y,K)$, generated by monomials $U_w^i V_z^j \cdot \boldsymbol{x}$ satisfying $A(\boldsymbol{x}) = i - j$, where A is the symmetrized Alexander grading.

Given a relative Spin^c structure $\mathfrak{t} \in \mathrm{Spin}^c(Y, L)$, we write

$$\iota_{\mathfrak{t}}\colon \mathrm{CFL}^\infty(Y, L, \mathfrak{t}) \hookrightarrow \mathrm{CFL}^\infty_{UV}(Y, L, \mathfrak{s})$$

for inclusion.

As a direct sum of $\mathbb{Z}_2$–modules, we have

$$\mathrm{CFL}^\infty_{UV}(\mathcal{H}, \mathfrak{s}) = \bigoplus_{\mathfrak{t} \in \mathrm{Spin}^c(Y, L)} \mathrm{CFL}^\infty(\mathcal{H}, \mathfrak{t}).$$

Write $\pi_{\mathfrak{t}}\colon \mathrm{CFL}^\infty_{UV}(\mathcal{H}, \mathfrak{s}) \to \mathrm{CFL}^\infty(\mathcal{H}, \mathfrak{t})$ for the projection onto $\mathrm{CFL}^\infty(\mathcal{H}, \mathfrak{t})$.

Finally, we note that in $\mathrm{CFL}^\infty(Y, \mathbb{L}, \mathfrak{t})$ the multi-index J in a monomial $U_{\boldsymbol{w}}^I V_{\boldsymbol{z}}^J \cdot \boldsymbol{x} \in \mathrm{CFL}^\infty(Y, \mathbb{L}, \mathfrak{t})$ is determined by the multi-index I, as well as the choice of $\mathfrak{t}$. Thus the full link Floer complex in [10] is described instead as the module generated by monomials $U_{\boldsymbol{w}}^I \cdot \boldsymbol{x}$, but with a filtration by $\mathrm{Spin}^c(Y, L)$. Given a $\mathfrak{t} \in \mathrm{Spin}^c(Y, L)$, it is straightforward to write down an isomorphism of filtered chain complexes between these two objects.

2.4 Colorings and $\mathfrak{P}$–filtrations

As was the case in [14], to define functorial maps it is important to work in a category of chain complexes over a fixed ring. As the link Floer complexes are modules over a ring which depends on the link, we need to formally "color" the complexes to make them modules over a fixed ring. Different choices of base rings will be useful for different applications, but for a single computation, a single ring must be fixed.

If $\mathfrak{P}$ is a finite set, we let $\mathbb{Z}_2[U_{\mathfrak{P}}, U_{\mathfrak{P}}^{-1}]$ denote the ring generated by the formal variables U_p and their inverses U_p^{-1} for $p \in \mathfrak{P}$.

Definition 2.5 If $\mathfrak{P}$ is a finite set, a *$\mathfrak{P}$–filtered chain complex* is a chain complex with a filtration of $\mathbb{Z}^{\mathfrak{P}}$, ie if C is a chain complex, then a $\mathfrak{P}$–filtration is a collection of subcomplexes $\mathcal{F}_I \subseteq C$ ranging over $I \in \mathbb{Z}^{\mathfrak{P}}$ such that if $I \le I'$, then $\mathcal{F}_{I'} \subseteq \mathcal{F}_I$. A $\mathfrak{P}$–filtered homomorphism is a homomorphism $\phi\colon C \to C'$ where C and C' are $\mathfrak{P}$–filtered with filtrations $\mathcal{F}_I$ and $\mathcal{F}'_I$ such that

$$\phi(\mathcal{F}_I) \subseteq \mathcal{F}'_I.$$

Definition 2.6 A *coloring of a multibased link* $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$ in Y is a pair $(\sigma, \mathfrak{P})$ where $\mathfrak{P}$ is a finite set and $\sigma\colon \boldsymbol{w} \cup \boldsymbol{z} \to \mathfrak{P}$ is a map which sends all of the $\boldsymbol{z}$ basepoints on a given link component to the same color (this condition ensures that the differential squares to zero).

If $(\sigma, \mathfrak{P})$ is a coloring of a link $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$, let $\mathcal{C}_{\sigma,\mathfrak{P}}$ denote the module

$$\mathbb{Z}_2[U_{\boldsymbol{w}}, U_{\boldsymbol{w}}^{-1}, V_{\boldsymbol{z}}, V_{\boldsymbol{z}}^{-1}, U_{\mathfrak{P}}, U_{\mathfrak{P}}^{-1}]/I_{\sigma,\mathfrak{P}},$$

where $I_{\sigma,\mathfrak{P}}$ is the submodule generated by elements of the form $U_w - U_{\sigma(w)}$ and $V_z - U_{\sigma(z)}$. The colored complex is then defined as

$$\mathrm{CFL}^\infty_{UV}(\mathcal{H}, \sigma, \mathfrak{P}, \mathfrak{s}) = \mathrm{CFL}^\infty_{UV,0}(\mathcal{H}, \sigma) \otimes_{\mathbb{Z}_2[U_{\boldsymbol{w}}, U_{\boldsymbol{w}}^{-1}, V_{\boldsymbol{z}}, V_{\boldsymbol{z}}^{-1}]} \mathcal{C}_{\sigma,\mathfrak{P}}.$$

Given a coloring $(\sigma, \mathfrak{P})$ of $\boldsymbol{w} \cup \boldsymbol{z}$ of a link $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$ in Y, the colored complexes $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s})$ naturally obtain a $\mathfrak{P}$–filtration by powers of the variables U_p. An element of $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s})$ is uniquely written as a sum of elements of the form $\boldsymbol{x} \cdot U_{\mathfrak{P}}^I$, and given an $I \in \mathbb{Z}^{\mathfrak{P}}$ we define $\mathcal{F}_I$ to be the $\mathbb{Z}_2[U_{\mathfrak{P}}]$–submodule generated by $\boldsymbol{x} \cdot U_{\mathfrak{P}}^J$ with $J \geq I$.

Remark 2.7 Asking that a $\mathbb{Z}_2[U_{\mathfrak{P}}]$–equivariant map

$$F\colon \mathrm{CFL}^\infty_{UV}(\mathcal{H}, \sigma, \mathfrak{P}, \mathfrak{s}) \to \mathrm{CFL}^\infty_{UV}(\mathcal{H}', \sigma', \mathfrak{P}, \mathfrak{s}')$$

be $\mathfrak{P}$–filtered is just asking that F can be written as

$$F(\boldsymbol{x}) = \sum_{I \geq 0} U_{\mathfrak{P}}^I \cdot H_I(\boldsymbol{x}),$$

where the maps H_I do not involve the $U_{\mathfrak{P}}$ variables. Most maps which appear in Heegaard Floer homology are $\mathfrak{P}$–filtered. The differential, the triangle maps, the quadrilateral maps, and the maps Φ_w, Φ_z and $S^{\pm}_{w,z}$ are all $\mathfrak{P}$–filtered.

Given an arbitrary coloring $(\sigma, \mathfrak{P})$ of basepoints $\boldsymbol{w} \cup \boldsymbol{z}$, we may not always be able to define submodules corresponding to relative Spin^c structures $\mathfrak{t}$. However, if no two basepoints from distinct link components are given the same color, then one can use a modification of (1) to define a $\mathfrak{P}$–filtered $\mathbb{Z}_2$–submodule $\mathrm{CFL}^\infty(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{t})$. For our purposes, we just observe that in the case that $(\sigma, \mathfrak{P})$ is the trivial coloring (ie $\mathfrak{P} = \boldsymbol{w} \cup |L|$ and σ is the map sending $w \in \boldsymbol{w}$ to w and $z \in \boldsymbol{z}$ to the link component containing it), then $\mathrm{CFL}^\infty_{UV}(\mathcal{H}, \sigma, \mathfrak{P}, \mathfrak{s})$ is equal to just $\mathrm{CFL}^\infty_{UV}(\mathcal{H}, \mathfrak{s})$ and the maps $\pi_{\mathfrak{t}}$ and $\iota_{\mathfrak{t}}$ are still defined. The following lemma is essentially trivial, though it is useful for relating endomorphisms of $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \mathfrak{s})$ to endomorphisms of the subcomplexes $\mathrm{CFL}^\infty(Y, \mathbb{L}, \mathfrak{t})$:

Lemma 2.8 *Suppose that $(L, \boldsymbol{w}, \boldsymbol{z})$ is a link in an integer homology sphere Y and that $(\sigma, \mathfrak{P})$ is the trivial coloring. Suppose f and g are $\mathfrak{P}$–filtered $\mathbb{Z}_2[U_{\mathfrak{P}}]$–module endomorphisms of $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s})$ such that f and g are chain homotopic via a chain homotopy which is $\mathfrak{P}$–filtered on $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s})$. Then $\pi_{\mathfrak{t}} \circ f \circ \iota_{\mathfrak{t}}$ and $\pi_{\mathfrak{t}} \circ g \circ \iota_{\mathfrak{t}}$ are $\mathbb{Z}^{|\boldsymbol{w}|} \oplus \mathbb{Z}^{|L|}$–filtered $\mathbb{Z}_2[\bar{U}_{\boldsymbol{w}}]$–chain homotopic.*

Proof First note that the filtration on $\mathrm{CFL}^\infty(\mathcal{H},\mathfrak{t})$ is just the pullback under $\iota_{\mathfrak{t}}$ of the $\mathfrak{P}$–filtration on $\mathrm{CFL}^\infty_{UV}(\mathcal{H},\sigma,\mathfrak{P},\mathfrak{s})$. If f and g are $\mathfrak{P}$–filtered chain homotopic, we have that

$$f-g=\partial H+H\partial$$

for a $\mathfrak{P}$–filtered map H. Pre- and postcomposing with the $\mathfrak{P}$–filtered maps $\iota_{\mathfrak{t}}$ and $\pi_{\mathfrak{t}}$ yields a $\mathbb{Z}^{|\boldsymbol{w}|}\oplus\mathbb{Z}^{|L|}$–filtered chain homotopy between $\pi_{\mathfrak{t}}\circ f\circ\iota_{\mathfrak{t}}$ and $\pi_{\mathfrak{t}}\circ g\circ\iota_{\mathfrak{t}}$ because $\pi_{\mathfrak{t}}$ and $\iota_{\mathfrak{t}}$ are $\mathfrak{P}$–filtered chain maps. The chain homotopy is $\mathbb{Z}_2[\bar{U}_{\boldsymbol{w}}]$–equivariant since $\iota_{\mathfrak{t}}$ and $\pi_{\mathfrak{t}}$ are $\mathbb{Z}_2[\bar{U}_{\boldsymbol{w}}]$–equivariant, as we mentioned above (recall that $\bar{U}_w=U_wV_z$, where z is any base point on the link component containing z). □

2.5 Why we use the larger $\mathrm{CFL}^\infty_{UV}(Y,L,\mathfrak{s})$ instead of other versions

We briefly explain why we use the object $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$ to prove formulas for basepoint moving maps, instead of other versions of link Floer homology. In the next sections, we will define maps Φ_w and Ψ_z, which are endomorphisms of $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$. However, due to the extra factors of U_w^{-1} or V_z^{-1} in the definitions, these do not preserve $\mathrm{CFL}^\infty(Y,\mathbb{L},\mathfrak{t})$ for a relative Spin^c structure. Instead they change the relative Spin^c structure by $\pm\,\mathrm{PD}[\mu]$, where μ is the meridian of the component containing w and z (note, however, that the composition $\Phi_w\Psi_z$ does actually preserve relative Spin^c structure). Although this is not insurmountable, what's worse is that the maps $S^+_{w,z}$ and $S^-_{w,z}$ are not even endomorphisms of the same complex, and since $S^+_{w,z}S^-_{w,z}\simeq\Phi_w$, we know that they can't preserve relative Spin^c structures. Similarly, one could try to use the version of link Floer homology described in [10] as a $\mathrm{Spin}^c(Y,L)$–filtration on $\mathrm{CF}^\infty(Y)$, but we have the same problem since the map Φ_w is not $\mathrm{Spin}^c(Y,L)$–filtered.

The solution is clearly to work with the larger complexes $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$.

There are also other algebraic advantages to working with $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\mathfrak{s})$. For instance, we can think of Φ_w and Ψ_z as formal derivatives of the differential. Using our expression for ∂^2, we can quickly derive many relations between various Φ_w and Ψ_z maps.

3 The maps Φ_w and Ψ_z

We now define maps Φ_w and Ψ_z, which are endomorphisms of $\mathrm{CFL}^\infty_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s})$. These are denoted by $\Phi_{i,j}$ and $\Psi_{i,j}$ in [13]. We define $\Phi_w:\mathrm{CFL}^\infty_{UV}(\mathcal{H},\sigma,\mathfrak{P},\mathfrak{s})\to\mathrm{CFL}^\infty_{UV}(\mathcal{H},\sigma,\mathfrak{P},\mathfrak{s})$ by the formula

$$\Phi_w(\boldsymbol{x})=U_w^{-1}\sum_{\boldsymbol{y}\in\mathbb{T}_\alpha\cap\mathbb{T}_\beta}\sum_{\substack{\phi\in\pi_2(\boldsymbol{x},\boldsymbol{y})\\ \mu(\phi)=1}} n_w(\phi)\#\widehat{\mathcal{M}}(\phi)U_{\boldsymbol{w}}^{n_{\boldsymbol{w}}(\phi)}V_{\boldsymbol{z}}^{n_{\boldsymbol{z}}(\phi)}\cdot\boldsymbol{y},$$

which we can alternatively think of as $(d\partial/dU_w)$. Similarly we define

$$\Psi_z(\boldsymbol{x}) = V_z^{-1} \sum_{\boldsymbol{y}\in\mathbb{T}_\alpha\cap\mathbb{T}_\beta} \sum_{\substack{\phi\in\pi_2(\boldsymbol{x},\boldsymbol{y})\\ \mu(\phi)=1}} n_z(\phi)\#\widehat{\mathcal{M}}(\phi) U_{\boldsymbol{w}}^{n_{\boldsymbol{w}}(\phi)} V_{\boldsymbol{z}}^{n_{\boldsymbol{z}}(\phi)} \cdot \boldsymbol{y},$$

which we can alternatively write as $(d\partial/dV_z)$, where the derivative is taken on $\mathrm{CFL}^\infty_{UV,0}$ (ie before tensoring with the module $\mathcal{L}$ and thus setting all of the V_z on a link component equal to each other).

We have the following (compare [13, Lemma 4.1]):

Lemma 3.1 *On $\mathrm{CFL}^\infty_{UV}(\mathcal{H},\mathfrak{s})$, we have $\Phi_w\partial + \partial\Phi_w = 0$. Also $\Psi_z\partial + \partial\Psi_z = U_w + U_{w'}$, where w and w' are the $\boldsymbol{w}$ basepoints adjacent to z.*

Proof One takes the derivative of $\partial\circ\partial$ with respect to either V_z or U_w, before one tensors $\mathrm{CFL}^\infty_{UV,0}$ with $\mathcal{L}$. The map ∂^2 is computed in Lemma 2.1. After tensoring with $\mathcal{L}$, one immediately arrives at the equalities described above. □

In addition, we have the following (compare [13, Theorem 4.2]):

Lemma 3.2 *The maps Φ_w and Ψ_z commute with change of diagram maps $\Phi_{\mathcal{H}_1\to\mathcal{H}_2}$ up to $\mathfrak{P}$–filtered, $\mathbb{Z}_2[U_{\mathfrak{P}}]$–chain homotopy.*

Proof Consider the complex $\mathrm{CFL}^\infty_{UV,0}$ (ie the complex before we set all of the V_z variables on each link component equal to each other). The differential doesn't square to zero, though we can still consider the maps $\Phi_{\mathcal{H}_1\to\mathcal{H}_2}$. These can be written as a composition of maps associated to changing the almost complex structure, triangle maps (corresponding to $\boldsymbol{\alpha}$– or $\boldsymbol{\beta}$–isotopies or handleslides), $(1,2)$–stabilization maps, and maps corresponding to isotoping the Heegaard surface inside of Y via an isotopy which fixes $\mathbb{L}$. We claim that the maps $\Phi_{\mathcal{H}_1\to\mathcal{H}_2}$ satisfy

$$\Phi_{\mathcal{H}_1\to\mathcal{H}_2}\partial + \partial\Phi_{\mathcal{H}_1\to\mathcal{H}_2} = 0,$$

even before tensoring with $\mathcal{L}$. The maps $\Phi_{\mathcal{H}_1\to\mathcal{H}_2}$ are defined as a composition of maps which count holomorphic triangles (handleslides or isotopy maps), holomorphic disks with dynamic almost complex structure (change of almost complex structure maps) or maps which are defined via simple, explicit formulas (stabilization and diffeomorphism). The maps which associated to $(1,2)$–stabilizations and diffeomorphisms obviously satisfy $\partial\phi + \phi\partial = 0$ before tensoring with $\mathcal{L}$. The maps induced by counting disks with dynamic almost complex structure also satisfy $\partial\phi + \phi\partial = 0$ before tensoring with ∂, since that follows from a Gromov compactness argument. Maps induced by handleslides or isotopies of the $\boldsymbol{\alpha}$–curves take the form

$$\boldsymbol{x} \mapsto F_{\boldsymbol{\alpha}'\boldsymbol{\alpha}\boldsymbol{\beta}}(\Theta \otimes \boldsymbol{x}),$$

where Θ is the top-degree generator of a complex $\mathrm{CFL}^-_{UV}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z})$ and $F_{\boldsymbol{\alpha}'\boldsymbol{\alpha}\boldsymbol{\beta}}$ is the map which counts holomorphic triangles. For this to be a chain map before tensoring with $\mathcal{L}$, it is sufficient that $\partial\Theta = 0$ before tensoring with $\mathcal{L}$, since the triangle map

$$F_{\boldsymbol{\alpha}'\boldsymbol{\alpha}\boldsymbol{\beta}} \colon \mathrm{CFL}^-_{UV,0}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z}) \otimes \mathrm{CFL}^\infty_{UV,0}(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z}) \\ \to \mathrm{CFL}^\infty_{UV,0}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$$

is a chain map by a Gromov compactness argument. We note now that the diagram $(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z})$ represents an unlink embedded in $(S^1 \times S^2)^{\#n}$ for some n, and this unlink has exactly two basepoints per link component. By the differential computation in Lemma 2.1, the complex

$$\mathrm{CFL}^-_{UV,0}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z})$$

is a chain complex before tensoring with anything, and in particular the homology group $\mathrm{HFL}^-_{UV,0}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z})$, is well defined even before tensoring with anything. An easy computation shows that if $\mathrm{HFL}^-_{UV,0,\max}$ denotes the subset of maximal homological grading (here the homological grading is obtained by ignoring the $\boldsymbol{z}$–basepoints, and assigning U variables degree -2, and V variables degree 0), then we have an isomorphism

$$\mathrm{HFL}^-_{UV,0,\max}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z}) \cong \mathbb{Z}_2[V_{\boldsymbol{z}}],$$

and in particular $\mathrm{HFL}^-_{UV,0}(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{w}, \boldsymbol{z})$ admits a "generator" Θ which is distinguished by the property of generating the maximally graded subset as a $\mathbb{Z}_2[V_{\boldsymbol{z}}]$–module. In particular, $\partial\Theta = 0$ even before tensoring with $\mathcal{L}$, as we needed.

Hence

$$\Phi_{\mathcal{H}_1 \to \mathcal{H}_2}\partial + \partial\Phi_{\mathcal{H}_1 \to \mathcal{H}_2} = 0,$$

even before tensoring with $\mathcal{L}$. Differentiating with respect to U_w yields that

$$\Phi'_{\mathcal{H}_1 \to \mathcal{H}_2}\partial + \Phi_{\mathcal{H}_1 \to \mathcal{H}_2}\Phi_w + \Phi_w\Phi_{\mathcal{H}_1 \to \mathcal{H}_2} + \partial\Phi'_{\mathcal{H}_1 \to \mathcal{H}_2} = 0,$$

immediately implying that $\Phi_{\mathcal{H}_1 \to \mathcal{H}_2}\Phi_w + \Phi_w\Phi_{\mathcal{H}_1 \to \mathcal{H}_2} \simeq 0$. The only point to check is that the chain homotopy $H = \Phi'_{\mathcal{H}_1 \to \mathcal{H}_2}$ is $\mathfrak{P}$–filtered and $\mathbb{Z}_2[U_{\mathfrak{P}}]$–equivariant. The equivariance condition is trivial. The filtration condition is also easy to check, since whenever F has a decomposition with only nonnegative powers of U_w and V_z, the map (dF/dU_w) also has a decomposition with nonnegative powers of U_w and V_z. $\square$

Remark 3.3 Using the Leibniz rule, we have that

$$\Phi_w = \frac{d}{dU_w} \circ \partial + \partial \circ \frac{d}{dU_w},$$

as long as U_w doesn't share the same color as any other basepoint. Similarly $\Psi_z \simeq 0$ if z doesn't share the same color with any other basepoint, though in both cases the chain homotopy $H = d/dU_w$ or $H = d/dV_z$ is neither $\mathfrak{P}$–filtered nor $\mathbb{Z}_2[U_{\mathfrak{P}}]$–equivariant.

4 Cylindrical boundary degenerations

We consider holomorphic curves whose boundary is mapped to only the $\boldsymbol{\alpha}$–curves, or only the $\boldsymbol{\beta}$ curves. These will be called *cylindrical* $\boldsymbol{\alpha}$*–boundary degenerations* or *cylindrical* $\boldsymbol{\beta}$*–boundary degenerations*.

We now define cylindrical $\boldsymbol{\alpha}$–boundary degenerations. Suppose that S is a Riemann surface with d punctures $\{p_1, \dots, p_d\}$ on its boundary. We consider holomorphic maps

$$u\colon S \to \Sigma \times (-\infty, 1] \times \mathbb{R}$$

such that the following hold:

(1) u is smooth;
(2) $u(\partial S) \subseteq (\boldsymbol{\alpha} \times \{1\} \times \mathbb{R})$;
(3) $\pi_{\mathbb{D}} \circ u$ is nonconstant on each component of S;
(4) $u^{-1}(\alpha_i \times \{1\} \times \mathbb{R})$ consists of exactly one component of $\partial S \setminus \{p_1, \dots, p_d\}$, for each i;
(5) the energy of u is finite;
(6) u is an embedding;
(7) if $z_i \in S$ is a sequence of points approaching a puncture p_j, then $(\pi_{\mathbb{D}} \circ u)(z_i)$ approaches -1 in the compactification of $(-\infty, 1] \times \mathbb{R}$ as the unit complex disk (with the point at ∞ identified with -1).

We organize such curves into moduli spaces $\mathcal{N}(\phi)$ for $\phi \in \pi_2^{\alpha}(\boldsymbol{x})$, modding out by automorphisms of the source, as usual. There is an action of $\mathrm{PSL}_2(\mathbb{R})$ on $\mathcal{N}(\phi)$, which is just the action on the $(-\infty, 1] \times \mathbb{R}$ coordinate of a disk u, and we denote the quotient space $\widehat{\mathcal{N}}(\phi)$. One defines cylindrical $\boldsymbol{\beta}$–boundary degenerations analogously. We now discuss transversality. In the original setup (singly pointed diagrams and disks mapped into $\mathrm{Sym}^g(\Sigma)$), a generic almost complex structure J on $\mathrm{Sym}^g(\Sigma)$ in a neighborhood of $\mathrm{Sym}^g(\mathrm{j})$ achieves transversality for Maslov index 2 holomorphic $\boldsymbol{\alpha}$–degenerate disks [9, Proposition 3.14]. In the cylindrical setup, if a sequence of holomorphic strips for the almost complex structures considered in [5] has a cylindrical boundary degeneration in its limit, the boundary degeneration will be $j_\Sigma \times j_{\mathbb{D}}$–holomorphic. Thus we need transversality for cylindrical boundary degenerations for split almost complex structures. For the standard proof that $\partial^2 = 0$ and for the purposes of this paper, we only need transversality for Maslov index 2 boundary degenerations. Each of these domains has multiplicity 1 in one component of $\Sigma \setminus \boldsymbol{\alpha}$, and zero everywhere else. If $u\colon S \to \Sigma \times (-\infty, 1] \times \mathbb{R}$ is a component of a holomorphic curve representing an element of $\mathcal{N}(\phi)$ for a $\phi \in \pi_2^{\alpha}(\boldsymbol{x})$ such that $\pi_\Sigma \circ u$ is nonconstant, then by easy complex analysis $u|_C$ is injective, where $C \subseteq \partial S$ is the part of S mapping to $\partial \mathcal{D}(\phi)$. Adapting the strategy of perturbing boundary conditions instead of almost complex structures, as in [5, Proposition 3.9], [9, Proposition 3.9] or [7], for generic choice of $\boldsymbol{\alpha}$–curves, we can thus achieve transversality for Maslov index 2 cylindrical boundary degenerations.

An important result for our purposes is a count of Maslov index 2 boundary degenerations produced by Ozsváth and Szabó:

Theorem 4.1 [10, Theorem 5.5] *Consider a surface Σ of genus g equipped with a set of attaching circles $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_{g+\ell-1}\}$ which span a g–dimensional lattice in $H_1(\Sigma; \mathbb{Z})$. If $\mathcal{D}(\phi) \geq 0$ and $\mu(\phi) = 2$, then $\mathcal{D}(\phi) = A_i$ for some i, and indeed*

$$\#\widehat{\mathcal{N}}(\phi) = \begin{cases} 0 \pmod 2 & \text{if } \ell = 1, \\ 1 \pmod 2 & \text{if } \ell > 1. \end{cases}$$

Here A_i denotes a component of $\Sigma \setminus \boldsymbol{\alpha}$.

5 Preliminaries on the quasistabilization operation

Suppose that $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$ is an oriented link in Y and that w and z are two points, both in a single component of $L \setminus (\boldsymbol{w} \cup \boldsymbol{z})$, such that $(L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\})$ has basepoints which alternate between $\boldsymbol{w}$ and $\boldsymbol{z}$ as one traverses the link. We assume that the point w comes after z according to the orientation of L. In Section 7 we prove invariance for quasistabilization maps

$$S^+_{w,z}\colon \mathit{CFL}^\infty_{UV}(Y, L, \boldsymbol{w}, \boldsymbol{z}, \sigma, \mathfrak{P}, \mathfrak{s}) \to \mathit{CFL}^\infty_{UV}(Y, L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\}, \sigma', \mathfrak{P}, \mathfrak{s})$$

and

$$S^-_{w,z}\colon \mathit{CFL}^\infty_{UV}(Y, L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\}, \sigma', \mathfrak{P}, \mathfrak{s}) \to \mathit{CFL}^\infty_{UV}(Y, L, \boldsymbol{w}, \boldsymbol{z}, \sigma, \mathfrak{P}, \mathfrak{s}),$$

which are defined up to $\mathfrak{P}$–filtered $\mathbb{Z}_2[U_{\mathfrak{P}}]$–chain homotopy. Here σ' is a coloring which extends σ.

Though it will take several sections to construct the maps and prove they are well defined, we now summarize that the maps will be defined by the formulas

$$S^+_{w,z}(\boldsymbol{x}) = \boldsymbol{x} \times \theta^+$$

and

$$S^-_{w,z}(\boldsymbol{x} \times \theta^+) = 0, \quad S^-_{w,z}(\boldsymbol{x} \times \theta^-) = \boldsymbol{x},$$

for suitable choices of Heegaard diagrams and almost complex structures.

To define the quasistabilization map, we use the special connected sum operation from [6]. There, Manolescu and Ozsváth describe a way of adding new w and z basepoints to Heegaard multidiagrams. They prove that for multidiagrams with at least three sets of attaching curves (eg Heegaard triples or quadruples), there is an identification of certain moduli spaces of holomorphic curves on the unstabilized diagram and certain moduli spaces of holomorphic curves on the stabilized diagram. They conjecture an analogous result for the holomorphic curves on a Heegaard diagram with two sets of attaching curves (ie for the differentials of quasistabilized diagrams), but only prove the result for grid diagrams using somewhat ad hoc techniques, since in general they run into transversality issues. We will soon prove Proposition 5.3, computing the differential on quasistabilized diagrams for appropriate almost complex structures, showing how to avoid any transversality issues and using no more gluing technology than is used in showing that $\partial^2 = 0$ on multipointed diagrams.

5.1 Topological preliminaries on quasistabilization

Suppose that $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ is a diagram for $(Y, L, \boldsymbol{w}, \boldsymbol{z})$. Given new basepoints w, z in the same component of $L \setminus (\boldsymbol{w} \cup \boldsymbol{z})$, such that w occurs after z, we now describe a new diagram $\bar{\mathcal{H}}_{p,\alpha_s}$, which depends on a choice of point $p \in \Sigma$ and curve $\alpha_s \subseteq \Sigma \setminus \boldsymbol{\alpha}$ which passes through the point p. For fixed α_s and p, the diagram $\bar{\mathcal{H}}_{p,\alpha_s}$ will be defined up to an isotopy of Y which fixes $\boldsymbol{w} \cup \boldsymbol{z} \cup \{w, z\}$ and maps L to L.

Given a diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ as above, let A denote the component of $\Sigma \setminus \boldsymbol{\alpha}$ which contains the basepoints adjacent to w and z on L. Let $p \in A \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta} \cup \boldsymbol{w} \cup \boldsymbol{z})$ be a point. If U_α denotes the handlebody component of $Y \setminus \Sigma$ such that the $\boldsymbol{\alpha}$–curves bound compressing disks in U_α, then there is a path λ in U_α from p to a point on L between w and z. Such a curve λ is specified up to an isotopy fixing $\boldsymbol{w} \cup \boldsymbol{z} \cup \{w, z\}$ which maps L to L by requiring that λ be isotopic in $\bar{U}_\alpha$ to a segment of L concatenated with an embedded arc on $\Sigma \setminus \boldsymbol{\alpha}$.

Let $N(\lambda)$ denote a regular neighborhood of λ inside of U_α such that $\partial N(\lambda)$, the boundary of $N(\lambda)$ inside of U_α, satisfies

$$\partial N(\lambda) \cap L = \{w, z\}.$$

Topologically $\partial N(\lambda)$ is just a disk, which we denote by D_1. Also let D_2 denote $N(\lambda)\cap\Sigma$, which we note is also a disk. We can assume that $D_2\cap(\boldsymbol{\alpha}\cup\boldsymbol{\beta}\cup\boldsymbol{w}\cup\boldsymbol{z})=\varnothing$. Define

$$\overline{\Sigma}_p=(\Sigma\setminus D_2)\cup D_1.$$

The surface $\overline{\Sigma}_p$ is specified up to an isotopy which fixes $(\boldsymbol{\alpha}\cup\boldsymbol{\beta}\cup\boldsymbol{w}\cup\boldsymbol{z})$. Figure 1 shows the situation schematically.

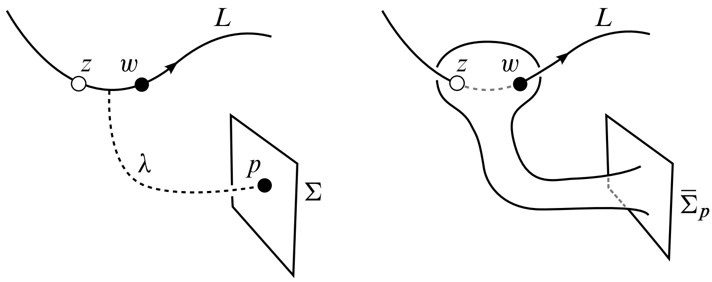


Figure 1: The path λ and the surfaces Σ and $\overline{\Sigma}_p$

We wish to extend the arc $\alpha_s\setminus D_2$ over all of D_1 to get a curve $\overline{\alpha}_s$ on $\overline{\Sigma}_p$. As is demonstrated in Figure 2, there is not an isotopically unique way to do this relative to the new basepoints w and z.

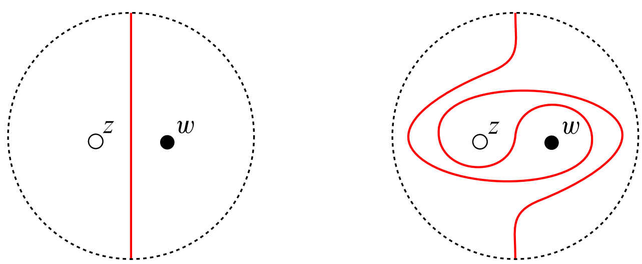


Figure 2: Different choices of $\overline{\alpha}_s$ curve on D_1 interpolating $\alpha_s\setminus D_2$. There is a unique isotopy class of such curves such that the resulting $\overline{\alpha}_s$ curve on $\overline{\Sigma}_p$ bounds a compressing disk which doesn't intersect L.

The set of such curves is easily seen to consist of those generated by the images of the curve on the left in Figure 2 under finger moves of w around z. Fortunately, the arc $\alpha_s\setminus D_2$ can be extended over D_1 uniquely (up to isotopy) by requiring the resulting curve $\overline{\alpha}_s\subseteq\overline{\Sigma}_p$ to bound a compressing disk in $U_{\overline{\alpha}}$ which doesn't intersect L, where here $U_{\overline{\alpha}}$ denotes the component of $Y\setminus\overline{\Sigma}_p$ in which the $\boldsymbol{\alpha}$–curves bound compressing disks.

Suppose $\mathcal{H}=(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{w},\boldsymbol{z})$ is a diagram, and $p\in\Sigma\setminus\boldsymbol{\alpha}$ is a chosen point, and let $N(p_0)$ denote a neighborhood of the connected sum point p_0 on S^2, which intersects α_0 in an arc and doesn't intersect β_0. Given a diffeomorphism $\psi\colon \Sigma \to \overline{\Sigma}_p$ which is the identity outside of D_2 and an embedding $\iota\colon S^2 \setminus N(p_0) \to D_1$ which sends $\alpha_0 \setminus N(p_0)$ to $\overline{\alpha}_s$, we can form a diagram

$$\overline{\mathcal{H}}_{p,\alpha_s} = (\overline{\Sigma}_p, \overline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\beta}}, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\}),$$

where $\overline{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \cup \{\overline{\alpha}_s\}$ and $\overline{\boldsymbol{\beta}} = \boldsymbol{\beta} \cup \{\iota(\beta_0)\}$. Such a choice of ψ and ι will be part of a larger collection of data $\mathcal{J}$ which we will consider in the next section and will call "gluing data". If we need to emphasize the distinction, we will write $\overline{\boldsymbol{\alpha}}$ for $\boldsymbol{\alpha} \cup \{\overline{\alpha}_s\}$, the curves on $\overline{\Sigma}_p$, and we will write $\boldsymbol{\alpha}^+$ for $\boldsymbol{\alpha} \cup \{\alpha_s\}$, the curves on Σ. By abuse of notation, we will often write α_s to denote both $\alpha_s \subseteq \Sigma$ and $\overline{\alpha}_s \subseteq \overline{\Sigma}_p$. Similarly, when no confusion will arise, we will write β_0 for both the curve β_0 on S^2 and the curve $\iota(\beta_0)$ on $\overline{\Sigma}_p$.

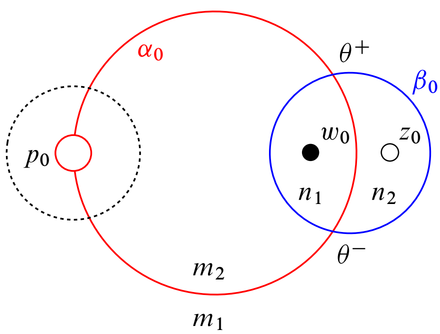


Figure 3: The diagram $\mathcal{H}_0$ used for quasistabilization, with multiplicities labeled. The dashed circle denotes where we will perform the neck stretching in the special connected sum.

5.2 Gluing data and almost complex structures

In [14] the author describes a systematic way of constructing and proving invariance of maps corresponding to adding or removing a basepoint from a closed 3–manifold. A key ingredient was a choice of auxiliary data which we call "gluing data" for patching two almost complex structures together in a systematic way. Here we introduce the analogous idea for quasistabilization.

Suppose that $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ is a diagram for $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$ and w, z are two new consecutive basepoints on L with z following w. Suppose $p \in A \setminus (\boldsymbol{w} \cup \boldsymbol{z} \cup \boldsymbol{\alpha} \cup \boldsymbol{\beta})$ is a distinguished point, where here A denotes the component of $\Sigma \setminus \boldsymbol{\alpha}$ containing the basepoints on L adjacent to w and z. Let $\overline{\Sigma}_p$ denote the Heegaard surface described in the previous subsection.

Definition 5.1 We define *gluing data* to be a collection

$$\mathcal{J} = (\psi, J_s, J_{s,0}, B, B_0, r, r_0, p_0, \iota, \phi),$$

where

(1) $\psi\colon \Sigma \to \overline{\Sigma}_p$ is a diffeomorphism which is fixed outside of D_2 and maps α_s to $\overline{\alpha}_s$;

(2) $B \subseteq \Sigma$ is a closed ball containing p which doesn't intersect $(\boldsymbol{w} \cup \boldsymbol{z} \cup \boldsymbol{\alpha} \cup \boldsymbol{\beta})$ and such that $\alpha_s \cap B$ is a closed arc;

(3) the point $p_0 \in \alpha_0 \setminus \beta_0$ is the connected sum point;

(4) $B_0 \subseteq S^2$ is a closed ball containing p_0 which doesn't intersect β_0 and such that $B_0 \cap \alpha_0$ is a closed arc;

(5) J_s is an almost complex structure on $\Sigma \times [0,1] \times \mathbb{R}$ which is split on B;

(6) $J_{s,0}$ is an almost complex structure on $S^2 \times [0,1] \times \mathbb{R}$ which is split on B_0;

(7) r and r_0 are real numbers such that $0 < r, r_0 < 1$;

(8) using the unique (up to rotation) conformal identifications of (B, p) and (B_0, p_0) as $(\mathbb{D}, 0)$, where $\mathbb{D}$ denotes the unit complex disk, ι is an embedding of $S^2 \setminus r_0 \cdot B_0$ into $r \cdot B \subseteq \Sigma$ such that

$$\iota(\alpha_0) \subseteq \alpha_s, \quad (\psi \circ \iota)(z_0) = z, \quad (\psi \circ \iota)(w_0) = w,$$

and

$$(r \cdot B) \setminus \iota(S^2 \setminus r_0 \cdot B_0)$$

is a closed annulus;

(9) letting $\widetilde{A}$, A and A_0 denote the closures of the annuli $B \setminus \iota(S^2 \setminus B_0)$, $B \setminus r \cdot B$ and $B_0 \setminus r_0 \cdot B_0$, respectively,

$$\phi\colon \widetilde{A} \to S^1 \times [-a, 1+b]$$

is a diffeomorphism which sends the annulus A to $[-a, 0]$ and $\iota(A_0)$ to $[1, 1+b]$ and is conformal on A and A_0.

The space of embeddings ι is connected since if a denotes the arc on the left side of Figure 2, the space of diffeomorphisms $f\colon B \to B$ mapping $\partial B \cup a$ to itself and fixing $\{z, w\}$ is connected. That the space of diffeomorphisms $\psi\colon \Sigma \to \overline{\Sigma}_p$ in the definition is also connected follows for similar reasons.

Gluing data $\mathcal{J}$ and a choice of neck length T determines an almost complex structure $\mathcal{J}(T)$ on $\overline{\Sigma}_p \times [0,1] \times \mathbb{R}$.

5.3 Computing the quasistabilized differential

We now wish to compute the differential after performing quasistabilization. We have the following Maslov index computation:

Lemma 5.2 *If ϕ_0 is a homology class of disks on $\mathcal{H}_0$, then using the multiplicities in Figure 3, we have*

$$\mu(\phi_0) = (n_1 + n_2 + m_1 + m_2)(\phi_0).$$

Proof The formula is easily verified for the constant disk and respects splicing in any of the Maslov index 1 strips. □

Let $\mathcal{J}$ denote gluing data as in the previous section, and let $\mathcal{J}(T)$ denote the almost complex structure on $\overline{\mathcal{H}}_{p,\alpha_s}$ determined by $\mathcal{J}$ for a choice of neck length T. We have the following:

Proposition 5.3 *Suppose that $\mathcal{H}$ is a strongly $\mathfrak{s}$–admissible diagram and that $\mathcal{J}$ is gluing data with almost complex structure J_s on $\Sigma \times [0,1] \times \mathbb{R}$. Then $\overline{\mathcal{H}}_{p,\alpha_s}$ is also strongly $\mathfrak{s}$–admissible and for sufficiently large T there is an identification of uncolored differentials (ie before we tensor with $\mathcal{L}$)*

$$\partial_{\overline{\mathcal{H}}_{p,\alpha_s},\mathcal{J}(T)} = \begin{pmatrix} \partial_{\mathcal{H},J_s} & U_w + U_{w'} \\ V_z + V_{z'} & \partial_{\mathcal{H},J_s} \end{pmatrix},$$

where the basepoints w and z are placed between w' and z' on $\mathbb{L}$.

Proof Suppose that u_i is a sequence of Maslov index 1 holomorphic curves on $\overline{\mathcal{H}}_{p,\alpha_s}$ in a fixed homology class for the almost complex structure $\mathcal{J}(T_i)$ for a sequence of T_i with $T_i \to \infty$. From the sequence u_i we can extract a weak limit of curves on the diagrams $(\Sigma, \boldsymbol{\alpha}^+, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ and $\mathcal{H}_0$. Let U_Σ, and U_0 denote these collections of curves. The curves in U_Σ consist of flowlines on $(\Sigma, \boldsymbol{\alpha}^+, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ as well as $\boldsymbol{\alpha}$– and $\boldsymbol{\beta}$–boundary degenerations on $(\Sigma, \boldsymbol{\alpha}^+)$ and $(\Sigma, \boldsymbol{\beta})$, and closed surfaces mapped into Σ. The holomorphic curves are now allowed to have a puncture along the α–boundary which is mapped asymptotically to p.

We first note that any flowline in the limit (ie a map $u: S \to \Sigma \times [0,1] \times \mathbb{R}$ which maps ∂S to $(\boldsymbol{\beta} \times \{0\} \cup \boldsymbol{\alpha}^+ \times \{1\} \times \mathbb{R}$ such that each component of S has both $\boldsymbol{\alpha}^+$ and $\boldsymbol{\beta}$ components) on the diagram $\mathcal{H}^+ = (\Sigma, \boldsymbol{\alpha}^+, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ must actually be a legitimate flow line on $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$. This is because if u is any holomorphic curve which is part of a weak limit of the curves u_{T_i}, then u cannot have a puncture asymptotic to an intersection point $\alpha_s \cap \beta_j$ for $\beta_j \in \boldsymbol{\beta}$. Hence if u is part of the weak limit of u_i, and if S denotes the source of u, then if ∂S has any points mapped to α_s, then S must have boundary component with a single puncture which is mapped to α_s. Projecting to $[0,1]\times\mathbb{R}$, we note that either $u|_{\partial S}$ attains a local extremum, or u is asymptotic to both $+\infty$ and $-\infty$ as one approaches the puncture. If $u|_{\partial S}$ attains a local extremum, then one can use the doubling trick to create an analytic function mapping $\mathbb{D}$ into $\mathbb{D}$ (where here $\mathbb{D}=\{z:|z|<1\}$) which maps $\mathbb{D}\cap\{\mathrm{im}(z)\geq 0\}$ to $\mathbb{D}\cap\{\mathrm{im}(z)\geq 0\}$ but which satisfies $f'(z)=0$ for some $z\in\mathbb{R}$, which is impossible by writing down a local model. The case that u is asymptotic to both $+\infty$ and $-\infty$ at the puncture is impossible since u must extend to a continuous function over the punctures. Hence any such u must be constant in the $[0,1]\times\mathbb{R}$ component, which implies that u cannot have any portion of ∂S mapped onto a $\boldsymbol{\beta}$–curve. Hence the curves in the weak limit can be taken to be holomorphic disks on $(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{w},\boldsymbol{z})$, $\boldsymbol{\alpha}^+$– or $\boldsymbol{\beta}$–degenerations, or closed surfaces.

Though not essential for our argument, to avoid "annoying" curves (ie maps into $\Sigma\times[0,1]\times\mathbb{R}$ which are constant in the $[0,1]\times\mathbb{R}$–component) among the $\boldsymbol{\beta}$– or $\boldsymbol{\alpha}^+$–degenerations, we observe that by rescaling the $[0,1]\times\mathbb{R}$ component, we could instead get curves that map into $\Sigma\times(-\infty,1]\times\mathbb{R}$ or $\Sigma\times S^2$ such that the $(-\infty,1]\times\mathbb{R}$ or S^2 components are nonconstant. Maps into $\Sigma\times(-\infty,1]\times\mathbb{R}$ are cylindrical $\boldsymbol{\alpha}^+$–boundary degenerations.

We now wish to compute exactly which of the above degenerations can occur in a weak limit of the sequence u_i of Maslov index 1 $\mathcal{J}(T_i)$–holomorphic curves. Assume without loss of generality that all of the u_i are in the same homology class ϕ.

Suppose that U_Σ consists of a collection U_Σ' of curves on $(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta})$ (flowlines, boundary degenerations, closed surfaces) and a collection of curves $\mathcal{A}$ in $(\Sigma,\boldsymbol{\alpha}^+,\boldsymbol{\beta})$ which have a boundary component which maps to α_s. As we've already remarked, the collection $\mathcal{A}$ consists exactly of cylindrical $\boldsymbol{\alpha}^+$–boundary degenerations. Letting ϕ_Σ' denote the underlying homology class of U_Σ', we define a combinatorial Maslov index for U_Σ by

$$\mu(U_\Sigma)=\mu(\phi_\Sigma')+m_1(\mathcal{A})+m_2(\mathcal{A})+2\sum_{\substack{\mathcal{D}\in C(\Sigma\setminus\boldsymbol{\alpha}),\\ \alpha_s\cap\mathcal{D}=\varnothing}} n_{\mathcal{D}}(\mathcal{A}),$$

where $C(\Sigma\setminus\boldsymbol{\alpha})$ denotes the connected components of $\Sigma\setminus\boldsymbol{\alpha}$. By Lemma 5.2 the Maslov index of U_0 satisfies

$$\mu(U_0)=m_1(\phi)+m_2(\phi)+n_1(\phi)+n_2(\phi).$$

The formula for $\mu(U_\Sigma)$ does not necessarily count the expected dimension of anything since we've only defined it combinatorially, though we can compute the Maslov index

$\mu(\phi)$ using these formulas, as follows:

(2) $$\begin{aligned}\mu(\phi) &= \mu(U_\Sigma)+\mu(U_0)-m_1(\phi)-m_2(\phi)\\ &= \mu(\phi'_\Sigma)+n_1(\phi)+n_2(\phi)+m_1(\mathcal{A})+m_2(\mathcal{A})+2\sum_{\substack{\mathcal{D}\in C(\Sigma\setminus\boldsymbol{\alpha})\\ \alpha_s\cap\mathcal{D}=\varnothing}} n_{\mathcal{D}}(\mathcal{A}).\end{aligned}$$

Each term in this sum is nonnegative, and by assumption the total sum is equal to 1. If $\mu(\phi)=1$, we immediately have that $n_{\mathcal{D}}(\mathcal{A})=0$ for $\mathcal{D}\in C(\Sigma\setminus\boldsymbol{\alpha})$, with $\alpha_s\cap\mathcal{D}=\varnothing$. Now $\mu(\phi'_\Sigma)$ does actually count the expected dimension of the moduli space of ϕ'_Σ, and in particular if ϕ'_Σ has a representative as a broken curve, we must have $\mu(\phi'_\Sigma)\geq 0$ with equality if and only if ϕ'_Σ is the constant disk.

As a consequence we see that if $\mu(\phi)=1$, we have that exactly one of $\mu(\phi'_\Sigma)$, $n_1(\phi)$, $n_2(\phi)$, $m_1(\mathcal{A})$ or $m_2(\mathcal{A})$ is equal to 1, and the rest are zero. The cases where $n_1(\phi)=1$ or $n_2(\phi)=1$ are easy to analyze, and those possibilities contribute summands of

$$\begin{pmatrix}0 & U_w\\ 0 & 0\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}0 & 0\\ V_z & 0\end{pmatrix},$$

respectively, to $\partial_{\bar{\mathcal{H}}_{p,\alpha_s}}$.

We now consider broken curves in the limit with $\mu(\phi'_\Sigma)=1$ and the remaining terms zero. In this case we have that $m_1(\mathcal{A})=m_2(\mathcal{A})=0$ (so $\mathcal{A}=0$) and $n_1(\phi)=n_2(\phi)=0$. In this case, we observe that ϕ'_Σ is represented by U_Σ and $\mu(\phi'_\Sigma)=1$, so the limit cannot contain any boundary degenerations or closed surfaces. Furthermore, by Maslov index considerations we have that U_Σ consists of a single Maslov index 1 flowline, which we denote by u_Σ.

We now consider the curves in U_0. There must be a component of U_0 which satisfies a matching condition with u_Σ. Note also that since U_Σ consists only of a single disk on $(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta})$, there cannot be any curves in U_0 which have a point on the boundary mapped to p_0 or have a puncture along their boundary which is asymptotic to p_0.

Let u_0 denote the component of U_0 which satisfies the matching condition

$$\rho^p(u_\Sigma)=\rho^{p_0}(u_0).$$

In particular, this forces $m_1(u_0)=m_2(u_0)$. We also have $n_1(u_0)=n_2(u_0)=0$. Here, if $u\colon S\to\Sigma\times[0,1]\times\mathbb{R}$ is a holomorphic disk, $\rho^q(u)$ is the divisor

$$(\pi_{\mathbb{D}}\circ u)(\pi_\Sigma\circ u)^{-1}(q)\in\mathrm{Sym}^{n_q(u)}(\mathbb{D}).$$

Any additional components u'_0 of U_0 must also satisfy $n_1(u'_0)=n_2(u'_0)=0$ and also can't have an interior point or boundary point mapped to p_0, and hence must be constant. Hence U_0 consists exactly of a holomorphic strip u_0 with $m_2(u_0)=m_1(u_0)$ which satisfies a matching condition with u_Σ, ie (u_Σ, u_0) are prematched strips.

We thus have shown that if $\mu(\phi'_\Sigma)=1$, then the weak limit of the curves u_i is a prematched strip, so following standard gluing arguments (see eg [5, Appendix A]), the count of $\widehat{\mathcal{M}}_{\mathcal{J}(T)}(\phi)$ is equal to the count of prematched strips with total homology class ϕ, for sufficiently large T. If ϕ_0 is a homology class of disks in $\pi_2(\theta^+,\theta^+)$ or $\pi_2(\theta^-,\theta^-)$, let $\mathcal{M}(\phi_0,\boldsymbol{d})$ denote the set of holomorphic strips u representing ϕ_0 with $\rho^{p_0}(u)=\boldsymbol{d}$. Note that there is a unique homology class of disks $\phi_0\in\pi_2(\theta^+,\theta^+)$ with $m_1(\phi_0)=m_2(\phi_0)=|\boldsymbol{d}|$ and $n_1(\phi_0)=n_2(\phi_0)=0$.

We claim that

$$\mathcal{M}(\phi_0,\boldsymbol{d})\equiv 1 \pmod 2$$

if $m_1(\phi_0)=m_2(\phi_0)=|\boldsymbol{d}|$ and $n_1(\phi_0)=n_2(\phi_0)=0$. We consider a path $\boldsymbol{d}_t$ between two divisors $\boldsymbol{d}_0$ and $\boldsymbol{d}_1$ and consider the 1–dimensional space

$$\mathcal{M}=\bigsqcup_{t\in[0,1]}\mathcal{M}(\phi_0,\boldsymbol{d}_t).$$

We count the ends of $\mathcal{M}$. There are ends corresponding to $\mathcal{M}(\phi_0,\boldsymbol{d}_0)$ and $\mathcal{M}(\phi_0,\boldsymbol{d}_1)$. On the other hand, there are ends corresponding to strip breaking or other types of degenerations. No curve in the degeneration can have p_0 in its boundary, which constrains any degeneration to be into disks of the form $\pi_2(\theta^+,\theta^+)$ or $\pi_2(\theta^-,\theta^-)$. But if any nontrivial strip breaking occurs, the Maslov index of the matching component drops, contradicting the formula for the Maslov index. Hence the only ends of $\mathcal{M}$ correspond to $\mathcal{M}(\phi_0,\boldsymbol{d}_0)$ and $\mathcal{M}(\phi_0,\boldsymbol{d}_1)$, implying that

$$\#\mathcal{M}(\phi_0,\boldsymbol{d}_0)\equiv\#\mathcal{M}(\phi_0,\boldsymbol{d}_1) \pmod 2.$$

We now consider a path of divisors $\boldsymbol{d}_T$ consisting of k points in $[0,1]\times\mathbb{R}$ spaced at least T apart which approach the line $\{0\}\times\mathbb{R}$ as $T\to\infty$. Letting $T\to\infty$, since p_0 is not on β_0, we know that the Gromov limit of the curves in $\mathcal{M}(\phi_0,\boldsymbol{d}_T)$ consists of k cylindrical β_0–degenerations, and a single constant holomorphic strip. Applying Theorem 4.1, we get that the total count of the boundary of

$$\overline{\bigsqcup_T \mathcal{M}(\phi_0,\boldsymbol{d}_T)}$$

is $\#\mathcal{M}(\phi_0,\boldsymbol{d}_1)+1$, implying the claim.

Disks which consist of preglued flowlines (u_Σ,u_0) glued together thus provide a total contribution of

$$\begin{pmatrix}\partial_{\mathcal{H},J_s} & 0\\ 0 & \partial_{\mathcal{H},J_s}\end{pmatrix}$$

to the differential.

We now consider the last contributions to the differential. These correspond to having $\mathcal{D}(\phi_1')=0$ and $n_1=n_2=0$, and $m_1(\mathcal{A})+m_2(\mathcal{A})=1$. In this case $\mathcal{A}$ is an $\boldsymbol{\alpha}^+$–boundary degeneration on $(\Sigma,\boldsymbol{\alpha}^+,\boldsymbol{\beta})$. Since

$$\sum_{\substack{\mathcal{D}\in C(\Sigma\setminus\boldsymbol{\alpha})\\ \alpha_s\cap\mathcal{D}=\varnothing}} n_{\mathcal{D}}(\mathcal{A})=0,$$

we have that $\mathcal{D}(\mathcal{A})$ is constrained to one of two domains. For each of the two possible domains of $\mathcal{D}(\mathcal{A})$, there is exactly one corresponding choice of domain for ϕ_0, the homology class of U_0, which is just the domain with exactly one of m_1 and m_2 equal to 1, and the other equal to zero, and n_1 and n_2 also zero.

On $\overline{\mathcal{H}}_{p,\alpha_s}$ these correspond to exactly two homology classes of disks. We now describe two strategies to count such disks. The first would be to perform a gluing argument to glue holomorphic representatives of the bigon on $\mathcal{H}_0$ to Maslov index 2 $\boldsymbol{\alpha}$–boundary degeneration on $(\Sigma,\boldsymbol{\alpha}^+)$ at punctures along their boundaries. As we remarked, one could rescale the curves so that they were genuine cylindrical $\boldsymbol{\alpha}^+$, and by perturbing the $\boldsymbol{\alpha}^+$ curves we could achieve transversality since the domains of such curves are α–injective. By a gluing argument, one could prove that the count on $\overline{\mathcal{H}}_{p,\alpha_s}$ was equal to the product of the counts for the two pieces, for a sufficiently stretched neck. Although the author isn't aware of any obstruction to do this, we will describe another approach which uses more established gluing results and a nice trick.

Let $\boldsymbol{x}$ be an intersection point on the unstabilized diagram. By our work up to now, there are two homology classes we have left to count: a disk $\phi_{z'}\in\pi_2(\boldsymbol{x}\times\theta^+,\boldsymbol{x}\times\theta^-)$ which goes over z' once, and a disk $\phi_{w'}\in\pi_2(\boldsymbol{x}\times\theta^-,\boldsymbol{x}\times\theta^+)$, which goes over w' once. To count the number of representatives of $\phi_{w'}$ and $\phi_{z'}$, we consider the ends of the moduli spaces associated to certain Maslov index 2 homology classes in $\pi_2(\boldsymbol{x}\times\theta^+,\boldsymbol{x}\times\theta^+)$. On $\overline{\mathcal{H}}_{p,\alpha_s}$, we consider the two components of $\overline{\Sigma}_p\setminus\overline{\boldsymbol{\alpha}}$ which have boundary along α_s. For an intersection point $\boldsymbol{x}$ on the unstabilized diagram, each of these two domains yields a homotopy class $\pi_2(\boldsymbol{x}\times\theta^+,\boldsymbol{x}\times\theta^+)$. Let us call these homotopy classes $A_{w'}$ and $A_{z'}$, depending on whether they go over w' or z'. Let us consider the ends of the 1–dimensional space of holomorphic disks $\widehat{\mathcal{M}}(A_{w'})$. The ends correspond to boundary degenerations and strip breaking. By our work so far, for sufficiently stretched almost complex structure, there is a single domain which can appear as the domain of a Maslov index 1 homology class with $n_1\neq 0$ and which admits a holomorphic representative, namely the bigon going over w once. Let us call this bigon b_w. Hence if a 1-parameter family of holomorphic disks in $\widehat{\mathcal{M}}(A_{z'})$ breaks into a pair of holomorphic disks, one of them must be have domain equal to b_w. This forces the other to have domain $A_{z'}-b_w$, ie the homology class must be $\phi_{z'}$. As no other homotopy classes can appear in strip breaking, due to our degenerated almost complex structure, we conclude that

$$\#\widehat{\mathcal{N}}(A_{z'}) + \#\widehat{\mathcal{M}}(\phi_{z'})\#\widehat{\mathcal{M}}(b_w) = 0,$$

as the latter count is the number of ends of $\widehat{\mathcal{M}}(A_{z'})$. Since $\#\widehat{\mathcal{N}}(A_{z'}) = 1$ by Theorem 4.1, and the bigon certainly has a unique holomorphic representative, we conclude that $\widehat{\mathcal{M}}(\phi_{z'}) = 1 \in \mathbb{Z}_2$. By an analogous argument, we conclude that $\widehat{\mathcal{M}}(\phi_{w'}) = 1$, as well.

With the above count we see that such curves make contributions of

$$\begin{pmatrix} 0 & U_{w'} \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ V_{z'} & 0 \end{pmatrix}.$$

Summing together all of the contributions, we see that the differential takes the form

$$\partial_{\overline{\mathcal{H}}_{p,\alpha_s},\mathcal{J}(T)} = \begin{pmatrix} \partial_{\mathcal{H},J_s} & U_w + U_{w'} \\ V_z + V_{z'} & \partial_{\mathcal{H},J_s} \end{pmatrix}. \qquad \square$$

Note that we tensor $\mathrm{CFL}^{\infty}_{UV,0}(\mathcal{H},\mathfrak{s})$ with $\mathcal{L}$ so that the differential squares to zero. When we do this, we set $V_z = V_{z'}$, and the bottom left entry of the differential vanishes.

Example 5.4 We now briefly give an example which helps to illustrate the technique we used to count some of the disks appearing in the off-diagonal entries of the differential. We consider a nested quasistabilization, shown in Figure 4, where we stretch along the dashed curve on the inside quasistabilization. We haven't drawn any basepoints in the figure. We've illustrated a homology class $A \in \pi_2(\boldsymbol{x} \times \theta, \boldsymbol{x} \times \theta)$ whose domain is just a component of $\Sigma \setminus (\overline{\alpha})$ and the ends of $\widehat{\mathcal{M}}(A)$. We've illustrated how the homology class can split into either pairs of Maslov index 1 disks, or a boundary degeneration. When we stretch the neck sufficiently, the weak limits argument from the previous proposition prohibits the middle pair of homology disks from both having a representative. Hence the boundary of $\widehat{\mathcal{M}}(\phi)$ consists of exactly $\widehat{\mathcal{N}}(A)$ (which has a unique representative) and $\widehat{\mathcal{M}}(b) \times \widehat{\mathcal{M}}(\phi)$, where b is a bigon and $\phi \in \pi_2(\boldsymbol{x} \times \theta, \boldsymbol{x} \times \theta')$ is one of the disks we were trying to count at the end of the previous proposition.

5.4 Dependence of quasistabilization on gluing data

In this subsection we prove some initial results about quasistabilization and change of almost complex structure maps. The reader should compare this to [14, Section 6], where the analogous arguments are presented for free stabilization.

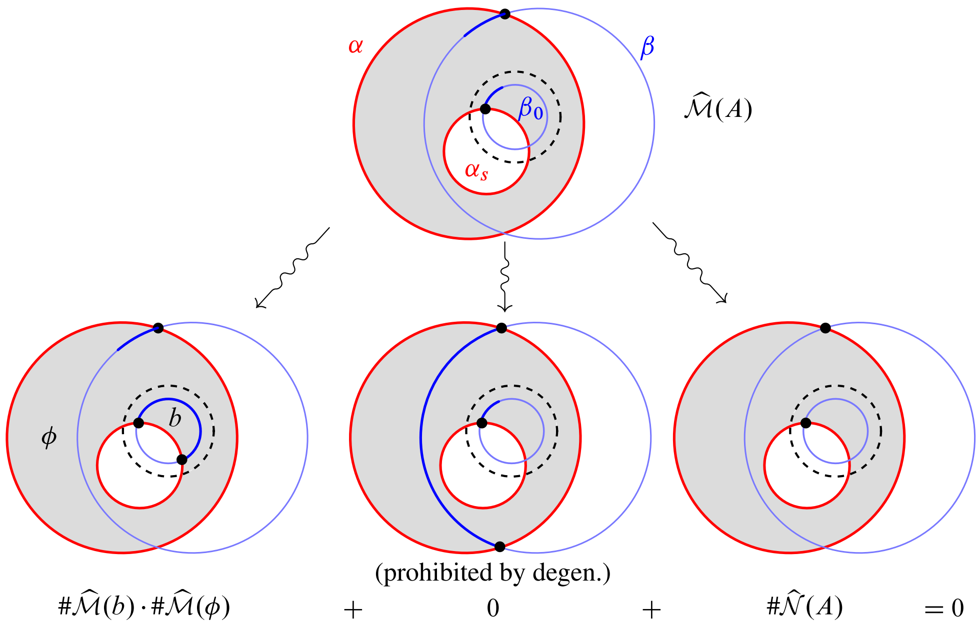


Figure 4: An example of the possible strip breaking which can occur for the homology class A. By degenerating the almost complex structure, our weak limits argument in the previous proposition rules out the middle degeneration. This allows us to count the representatives for the Maslov index 1 homology class ϕ, appearing on the left, which is otherwise hard to count.

Lemma 5.5 *Suppose that $\mathcal{J}$ is gluing data. Then there is an N such that if $T, T' > N$ and if $\mathcal{J}(T)$ and $\mathcal{J}(T')$ achieve transversality, then*

$$\Phi_{\mathcal{J}(T)\to\mathcal{J}(T')} \simeq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The proof is analogous to the proof of [14, Lemma 6.8], using the techniques of Proposition 5.3. Note that in [14], the upper right entry appeared as a $*$. Due to the Maslov index computation in (2), both off-diagonal entries are forced to be zero in our case. Philosophically this is because quasistabilization is a stronger degeneration than free stabilization. As in [14], we make the following definition:

Definition 5.6 We say that N is *sufficiently large* for gluing data $\mathcal{J}$ if $\Phi_{\mathcal{J}(T)\to\mathcal{J}(T')}$ is of the form in the previous lemma for all $T, T' \geq N$. We say T is *large enough to compute* $S^{\pm}_{w,z}$ for the gluing data $\mathcal{J}$ if $T > N$ for some N which is sufficiently large.

Adapting the proofs of [14, Lemmas 6.10–6.16], we have the following:

Lemma 5.7 *If $\mathcal{J}$ and $\bar{\mathcal{J}}$ are two choices of gluing data with almost complex structures J_s and $\bar{J}_s$ on $\Sigma\times[0,1]\times\mathbb{R}$, respectively, then there is an N such that if $T>N$ and if $\mathcal{J}(T)$ and $\bar{\mathcal{J}}(T)$ achieve transversality, then*

$$\Phi_{\mathcal{J}(T)\to\bar{\mathcal{J}}(T)}\simeq\begin{pmatrix}\Phi_{J_s\to\bar{J}_s} & 0\\ 0 & \Phi_{J_s\to\bar{J}_s}\end{pmatrix}.$$

The previous lemma will be used to show that the quasistabilization maps are independent of the choice of gluing data.

6 Quasistabilization and triangle maps

In this section we prove several results about quasistabilizing Heegaard triples, which we will use to prove invariance of the quasistabilization maps. The results for quasistabilization of Heegaard triples along a single α_s curve are established in [6], so we focus on quasistabilizing a Heegaard triple along two curves, α_s and β_s. To compute the quasistabilization maps $S^{\pm}_{w,z}$, we pick a curve α_s in the surface Σ, but there are many choices of such an α_s curve, so in order to address invariance of the quasistabilization maps, we need to show that the maps $S^{\pm}_{w,z}$ commute with the change of diagram map corresponding to moving α_s to α_s', which can be computed using a Heegaard triple which has been quasistabilized along two curves. Our main result is Theorem 6.5, which is an analogue of our computation of the differential after quasistabilizing in Proposition 5.3, but for certain Heegaard triples which we have quasistabilized along two curves which are allowed to travel throughout the diagram.

6.1 Quasistabilizing Heegaard triples along a single curve

We now consider Heegaard triples which are quasistabilized along a single α_s curve which is allowed to run through the diagram. This was first considered in [6]. We state a result from that paper, which considers the quasistabilized configuration shown in Figure 5. The result will be useful in showing that the quasistabilization maps are invariant under $\boldsymbol{\beta}$–handleslides and $\boldsymbol{\beta}$–isotopies.

Lemma 6.1 [6, Proposition 5.2] *Suppose that $\mathcal{T}=(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{w},\boldsymbol{z})$ is a strongly $\mathfrak{s}$–admissible triple and suppose that α_s is a new $\boldsymbol{\alpha}$–curve, passing through the point $p\in\Sigma$. Let $\overline{\mathcal{T}}_{\alpha_s,p}$ denote the Heegaard triple resulting from quasistabilizing along α_s at p, as in Figure 5. If $\mathcal{J}$ is gluing data, then for sufficiently large T, with the almost complex structure $\mathcal{J}(T)$ we have the following identifications:*

$$F_{\overline{\mathcal{T}}_{\alpha_s,p},\bar{\mathfrak{s}}}(\boldsymbol{x}\times\cdot,\boldsymbol{y}\times y^+)=\begin{pmatrix}F_{\mathcal{T},\mathfrak{s}}(\boldsymbol{x},\boldsymbol{y}) & 0\\ 0 & F_{\mathcal{T},\mathfrak{s}}(\boldsymbol{x},\boldsymbol{y})\end{pmatrix}$$

where $\cdot \in \{x^+, x^-\} = \alpha_s \cap \beta_0$ *and the matrix on the right denotes the expansion into the upper and lower generator components of* $\alpha_s \cap \beta_0$ *and* $\alpha_s \cap \gamma_0$*.*

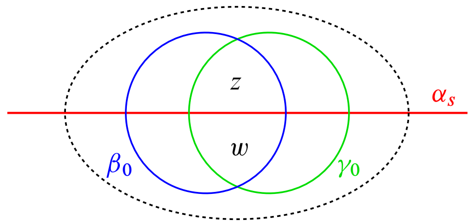


Figure 5: The version of quasistabilization discussed in Lemma 6.1

6.2 Strong positivity condition for diagrams of $(S^1 \times S^2)^{\#k}$

In this subsection we describe a class of simple diagrams for $(S^1 \times S^2)^{\#k}$ which we will use in a technical condition in Theorem 6.5 for quasistabilizing Heegaard triples along two curves, α_s and β_s, passing through a Heegaard triple.

Suppose that $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w})$ is a diagram for $(S^1 \times S^2)^{\#k}$ such that

$$|\alpha_i \cap \beta_j| = \begin{cases} 1 \text{ or } 2 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases}$$

and that if $\alpha_i \cap \beta_i = \{p_i^-, p_i^+\}$, then p_i^- and p_i^+ differ by Maslov grading 1. We do not assume that the α_i curves are small isotopies of the β_j curves. Let $\theta^+ = p_1^+ \times \cdots \times p_{n-1}^+$ denote the top graded (partial) intersection point. Assume that $|\alpha_n \cap \beta_n| = 2$ and write p_n^- and p_n^+ for the two points of $\alpha_n \cap \beta_n$.

Definition 6.2 Under the same assumptions as the previous paragraph, we say that $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w})$ is *strongly positive* with respect to p_n^+ if for every nonnegative disk $\phi \in \pi_2(\theta^+ \times p_n^+, \boldsymbol{y} \times p_n^+)$ we have that

$$(\mu - (m_1 + m_2))(\phi) = (\mu - (n_1 + n_2))(\phi) \geq 0,$$

with equality to zero if and only if ϕ is the constant disk. Here m_1, m_2, n_1, n_2 denote the multiplicities adjacent to the point p_n^+, appearing in the following counterclockwise order: n_1, m_1, n_2, then m_2.

Note that $m_1 + m_2 = n_1 + n_2$ for any disk $\phi \in \pi_2(\boldsymbol{x} \times p_n^+, \boldsymbol{y} \times p_n^+)$, by the vertex relations.

We now describe a class of diagrams which are strongly positive at an intersection point, which will be sufficient for our purposes:

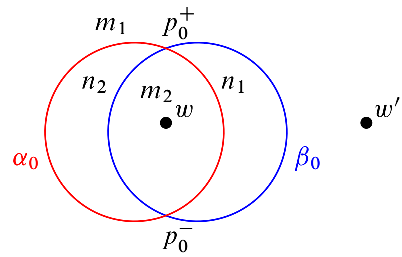


Figure 6: The diagram $\mathcal{H}_0 = (S^2, \alpha_0, \beta_0, w, w')$ in Lemma 6.3 which is strongly positive at p_0^+, and the multiplicities m_1, n_1, m_2 and n_2

Lemma 6.3 *The diagram $\mathcal{H}_0 = (S^2, \alpha_0, \beta_0, w, w')$ in Figure 6 is strongly positive with respect to p_0^+, the intersection point of α_0 and β_0 with higher relative grading. If $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w})$ is a diagram with a distinguished intersection point $p_n^+ \in \alpha_n \cap \beta_n$ where $|\alpha_n \cap \beta_n| = 2$, and $\mathcal{H}' = (\Sigma', \boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{w}')$ is the result of any of the following moves, then $\mathcal{H}'$ is strongly positive with respect to p_n^+ if and only if $\mathcal{H}$ is strongly positive with respect to p_n^+:*

1. *$(1,2)$–stabilization*[1]*;*
2. *taking the disjoint union of $\mathcal{H}$ with the standard diagram $(\mathbb{T}^2, \alpha_0, \beta_0, w)$ for (S^3, w);*
3. *performing surgery on an embedded 0–sphere $\{q_1, q_2\} \subseteq \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$ by removing small disks from Σ, and connecting the resulting boundary components with an annulus with new α_0 and β_0 curves with $|\alpha_0 \cap \beta_0| = 2$, and which are isotopic to each other, and homotopically nontrivial in the annulus.*

Proof We first note that $\mathcal{H}_0$ is strongly positive with respect to p_0^+, because the Maslov index of any disk is given by

$$\mu(\phi) = (m_1 + m_2 + n_1 + n_2)(\phi),$$

by Lemma 5.2, where m_1, m_2, n_1, n_2 are multiplicities appearing in the counterclockwise order m_1, n_1, m_2, n_2 around p_0^+, as in Figure 6. Hence for any disk ϕ, we have

$$(\mu - (n_1 + n_2))(\phi) = (m_1 + m_2)(\phi),$$

which is certainly nonnegative. For a disk $\phi \in \pi_2(p_0^+, p_0^+)$ we also have

$$(m_1 + m_2)(\phi) = (n_1 + n_2)(\phi),$$

so the above quantity is positive if and only if ϕ has positive multiplicities.

[1] If $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w})$ is a Heegaard diagram for Y, a $(1,2)$–stabilization of $\mathcal{H}$ is obtained by taking an embedded torus $\mathbb{T}^2$ inside of a 3–ball in $Y \setminus \Sigma$, together with curves α_0 and β_0 with $|\alpha_0 \cap \beta_0| = 1$, which bound compressing disks with boundary on Σ, and letting $\mathcal{H}'$ be the diagram $(\Sigma \# \mathbb{T}^2, \boldsymbol{\alpha} \cup \{\alpha_0\}, \boldsymbol{\beta} \cup \{\beta_0\}, \boldsymbol{w})$, where $\Sigma \# \mathbb{T}^2$ is the internal connected sum.

We now address moves (1)–(3).

Move (1) If $\mathcal{H}$ is a diagram, and $\mathcal{H}'$ is the result of $(1,2)$–stabilization, there is an isomorphism

$$\sigma_*\colon \pi_2^{\mathcal{H}}(\boldsymbol{x}\times p_n^+,\boldsymbol{y}\times p_n^+)\to \pi_2^{\mathcal{H}'}(\boldsymbol{x}\times c\times p_n^+,\boldsymbol{y}\times c\times p_n^+),$$

where c is the intersection of the new $\boldsymbol{\alpha}$– and $\boldsymbol{\beta}$–curves. Furthermore,

$$(\mu-(n_1+n_2))(\phi)=(\mu-(n_1+n_2))(\sigma_*\phi),$$

from which the claim follows easily.

Move (2) Suppose $\mathcal{H}'$ is formed from $\mathcal{H}$ by taking the disjoint union of $\mathcal{H}$ with a diagram $(\mathbb{T}^2,\alpha_0,\beta_0,w)$. Homology classes on $\mathcal{H}'$ are of the form $\phi\sqcup k\cdot[\mathbb{T}^2]$, where ϕ is a homology disk on $\mathcal{H}$. One has

$$\mu(\phi\sqcup k\cdot[\mathbb{T}^2])=\mu(\phi)+2k,$$

from which the claim follows easily.

Move (3) This move corresponds to surgering on an embedded 0–sphere $\{q_1,q_2\}\subseteq \Sigma\setminus(\boldsymbol{\alpha}\cup\boldsymbol{\beta}\cup\boldsymbol{w})$. Write α_0 and β_0 for the new curves on the annulus, and $\{\theta_0^+,\theta_0^-\}=\alpha_0\cap\beta_0$. Suppose that $y\in\alpha_0\cap\beta_0$ is a choice of intersection point. We can define (noncanonically) an injection

$$\iota_y\colon \pi_2^{\mathcal{H}}(\theta^+\times p_n^+,\boldsymbol{y}\times p_n^+)\to \pi_2^{\mathcal{H}'}(\theta^+\times\theta_0^+\times p_n^+,\boldsymbol{y}\times y\times p_n^+).$$

For $y=\theta_0^+$, we define $\iota_{\theta_0^+}(\phi)$ to be the disk on the surgered diagram which has no change across the curve β_0, but which agrees with the disk ϕ away from the α_0 and β_0 curves. For $y=\theta_0^-$, a map $\iota_{\theta_0^-}$ can be defined by defining it to be the map $\iota_{\theta_0^+}$, defined above, composed with the map on disks obtained by splicing in a choice of one of the bigons from θ_0^+ to θ_0^-. An easy computation shows that

$$(\mu-(m_1+m_2))(\iota_{\theta_0^+}(\phi))=(\mu-(m_1+m_2))(\phi),$$

while

$$(\mu-(m_1+m_2))(\iota_{\theta_0^-}(\phi))=(\mu-(m_1+m_2))(\phi)+1.$$

Any disk in $\pi_2(\theta^+\times\theta_0^+\times p_n^+,\boldsymbol{y}\times y\times p_n^+)$ is equal to one which is in the image of ι_y, with $n\cdot\mathcal{P}$ spliced in, where $\mathcal{P}$ is the periodic domain which is $+1$ in one of the small strips between α_0 and β_0, and -1 in the other. We note that

$$\mu(\mathcal{P})=m_1(\mathcal{P})=m_2(\mathcal{P})=0.$$

From these observations it follows easily that $\mathcal{H}$ is strongly positive with respect to p_n^+ if and only if $\mathcal{H}'$ is. $\square$

Remark 6.4 Suppose $\mathcal{H}=(\Sigma,\boldsymbol{\alpha}',\boldsymbol{\alpha},\boldsymbol{w})$ is obtained by taking attaching curves $\boldsymbol{\alpha}$ and letting $\boldsymbol{\alpha}'$ be small Hamiltonian isotopies of the curves in $\boldsymbol{\alpha}$. Let α_s be a new curve in $\Sigma\setminus\boldsymbol{\alpha}$ which doesn't intersect any $\boldsymbol{\alpha}'$–curves. If α_s' is the result of handlesliding α_s across a curve in $\boldsymbol{\alpha}$, then $(\Sigma,\boldsymbol{\alpha}'\cup\{\alpha_s'\},\boldsymbol{\alpha}\cup\{\alpha_s\},\boldsymbol{w}\cup\{w\})$ is strongly positive at p^+, the intersection point of $\alpha_s'\cap\alpha_s$ with higher grading. Here w is a new basepoint in one of the regions adjacent to p^+.

Similarly, if $\mathcal{H}=(\Sigma,\boldsymbol{\alpha}',\boldsymbol{\alpha},\boldsymbol{w})$ is the result of handlesliding a curve in $\boldsymbol{\alpha}$ across another curve in $\boldsymbol{\alpha}$, and α_s and α_s' are two new curves which are Hamiltonian isotopies of each other, then $(\Sigma,\boldsymbol{\alpha}'\cup\{\alpha_s'\},\boldsymbol{\alpha}\cup\{\alpha_s\},\boldsymbol{w}\cup\{w\})$ is strongly positive with respect to the intersection point of $\alpha_s'\cap\alpha_s$ of higher relative grading.

Note that a diagram $(\Sigma_g,\boldsymbol{\alpha}',\boldsymbol{\alpha},w)$ where the curves in $\boldsymbol{\alpha}'$ are small isotopies of the curves in $\boldsymbol{\alpha}$ with $g(\Sigma_g)=|\boldsymbol{\alpha}'|=|\boldsymbol{\alpha}|=g$ is not a strongly positive diagram at any point, since $[\Sigma_g]$ represents a positive homology class in $\pi_2(\theta^+,\theta^+)$ with

$$\mu([\Sigma_g])-m_1-m_2=2-1-1=0.$$

Strongly positive diagrams always have multiple basepoints. The prototypical example is the one resulting from handlesliding a quasistabilization curve α_s across an $\boldsymbol{\alpha}$–curve, as in Figure 7.

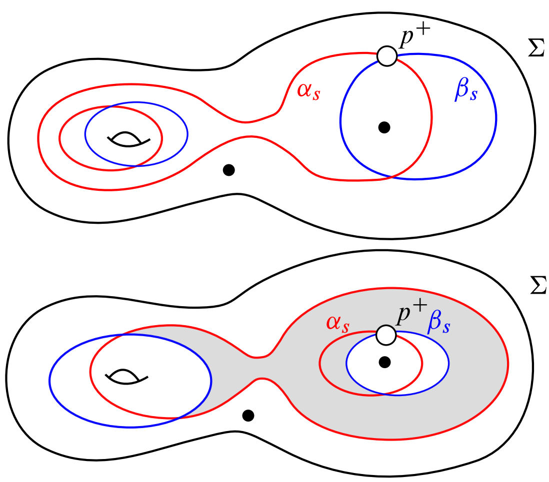


Figure 7: The diagram on the top is strongly positive with respect to the point p^+. The curves α_s and β_s are curves on which one could perform the quasistabilization operation of triangles in Theorem 6.5. The diagram on the bottom is not, and the nonzero, nonnegative domain of a disk ϕ with $\mu(\phi)-(n_1+n_2)(\phi)=0$ is shown.

A more interesting example of a diagram which doesn't satisfy the strong positivity condition would be the pair $(\Sigma, \boldsymbol{\alpha}' \cup \{\alpha_s'\}, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{w})$ that arises in a triple $(\Sigma, \boldsymbol{\alpha}' \cup \{\alpha_s'\}, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{\beta} \cup \{\beta_0\}, \boldsymbol{w})$ for handlesliding an $\boldsymbol{\alpha}$ curve across α_s. Fortunately, such a move is not required in the proof of invariance of the quasistabilization maps.

6.3 Quasistabilizing Heegaard triples along two curves

We now consider the effect on the triangle maps of quasistabilizing along two curves. Our analysis follows a similar spirit to the proof of [6, Proposition 5.2]. Suppose that $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{w}, \boldsymbol{z})$ is a Heegaard triple with a distinguished point

$$p^+ \in \Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta} \cup \boldsymbol{\gamma} \cup \boldsymbol{w} \cup \boldsymbol{z}).$$

Suppose also that α_s and β_s are choices of curves in $\Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{w} \cup \boldsymbol{z})$ and $(\Sigma \setminus (\boldsymbol{\beta} \cup \boldsymbol{w} \cup \boldsymbol{z})$, respectively, which intersect only at p^+ and another point $p^- \in \Sigma$. We can form the diagram $\overline{\mathcal{T}}_{\alpha_s, \beta_s, p^+}$, obtained by quasistabilizing along both α_s and β_s, simultaneously, at the point p^+. This corresponds to removing a small disk containing p^+, and inserting the diagram shown in Figure 8, with a disk centered around p_0^- removed.

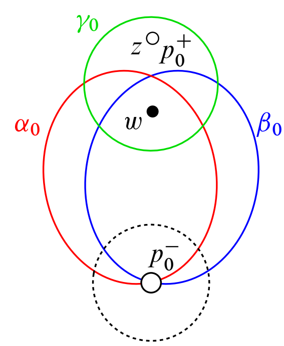


Figure 8: The diagram we insert into a Heegaard triple diagram $\mathcal{T}$ along the curves α_s and β_s to form the diagram $\overline{\mathcal{T}}_{\alpha_s, \beta_s, p^+}$. We cut out the solid circle marked with p_0^- and stretch the almost complex structure along the dashed circle.

Theorem 6.5 *Suppose that $\mathcal{T} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{w}, \boldsymbol{z})$ is a Heegaard triple with curves α_s and β_s, intersecting at two points p^+ and p^-, and let $\overline{\mathcal{T}}_{\alpha_s, \beta_s, p^+}$ be the Heegaard triple resulting from quasistabilization, as described above. If $(\Sigma, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{\beta} \cup \{\beta_s\}, \boldsymbol{w} \cup \{w\})$ is a strongly positive diagram for $(S^1 \times S^2)^{\# k}$ with respect to p^+ (Definition 6.2), and $\mathcal{J}$ is gluing data for stretching along the dashed circle, then for sufficiently large T, with respect to the almost complex structure $\mathcal{J}(T)$, there are identifications*

$$F_{\mathcal{T}^+_{\alpha_s,\beta_s,p^+},\bar{\mathfrak{s}},\mathcal{J}(T)}(\Theta^+_{\alpha\beta}\times p_0^+,\boldsymbol{x}\times\cdot)=\begin{pmatrix} F_{\mathcal{T},\mathfrak{s},J}(\Theta^+_{\alpha\beta},\boldsymbol{x}) & 0\\ 0 & F_{\mathcal{T},\mathfrak{s},J}(\Theta^+_{\alpha\beta},\boldsymbol{x})\end{pmatrix},$$

where $\cdot$ denotes $x^\pm\in\beta_0\cap\gamma_0$ and the matrix on the right denotes the matrix decomposition of the map based on the decompositions given $x^\pm$ and $y^\pm$.

As usual, the argument proceeds by a Maslov index calculation, which we use to put constraints on the homology classes of holomorphic curves which can appear in a weak limit as we let the parameter T approach $+\infty$. Once we determine which homology classes of triangles can appear, we can use standard gluing results to explicitly count holomorphic curves.

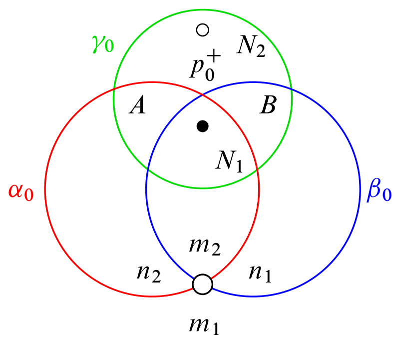


Figure 9: Multiplicities for a triangle on the diagram $(S^2,\alpha_0,\beta_0,\gamma_0,w,z)$

Lemma 6.6 *Suppose $\psi\in\pi_2(x,y,z)$ is a homology disk on $(S^2,\alpha_0,\beta_0,\gamma_0,w,z)$, shown in Figure 9. Then*

$$\mu(\psi)=n_1+n_2+N_1+N_2.$$

Proof The formula is easily checked for any of the Maslov index 0 small triangles, and respects splicing in any Maslov index 1 strip. Since any two triangles on this diagram differ by splicing in some number of the Maslov index 1 strips, the formula follows in full generality. □

We can now prove Theorem 6.5.

Proof of Theorem 6.5 Suppose that u_i is a sequence of holomorphic triangles of Maslov index 0 representing a class $\psi\in\pi_2(\Theta^+_{\alpha\beta}\times p_0^+,\boldsymbol{x}\times x,\boldsymbol{y}\times y)$, for almost complex structure $\mathcal{J}(T_i)$, where T_i is a sequence of neck lengths approaching $+\infty$. Adapting the proof of Proposition 5.3, the limiting curves which appear can be arranged into three classes of broken holomorphic curves:

(1) a broken holomorphic triangle u_Σ which represents a homology class ψ_Σ on $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ which has no boundary components on α_s or β_s;

(2) a broken holomorphic disk $u_{\alpha\beta}$ on $(\Sigma, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{\beta} \cup \{\beta_s\})$ which represents a class $\phi_{\alpha\beta} \in \pi_2(\Theta^+_{\alpha\beta} \times p^+, \boldsymbol{y} \times p^+)$, for some $\boldsymbol{y} \in \mathbb{T}_\alpha \cap \mathbb{T}_\gamma$;

(3) a broken holomorphic triangle u_0 on $(S^2, \alpha_0, \beta_0, \gamma_0)$ which represents a homology triangle $\psi_0 \in \pi_2(p_0^+, x, y)$.

We now wish to write down the Maslov index of ψ in terms of the Maslov indices and multiplicities of ψ_Σ, ψ_0 and $\phi_{\alpha\beta}$. Let $m_1(\cdot)$, $m_2(\cdot)$, $n_1(\cdot)$ and $n_2(\cdot)$ denote the multiplicities of a homology curve in the regions surrounding p^+ or p_0^-, as in Figure 9. In [12], Sarkar derives a formula for the Maslov index of a homology triangle $\rho \in \pi_2(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ which can be computed entirely from the domain $\mathcal{D}(\rho)$. Writing $\mathcal{D} = \mathcal{D}(\rho)$, the formula reads

$$\mu(\rho) = e(\mathcal{D}) + n_{\boldsymbol{x}}(\mathcal{D}) + n_{\boldsymbol{y}}(\mathcal{D}) + a(\mathcal{D}).c(\mathcal{D}) - \tfrac{1}{2}d,$$

where $d = |\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = |\boldsymbol{\gamma}|$. Here $a(\mathcal{D})$ is defined to be the intersection $\partial\mathcal{D} \cap \boldsymbol{\alpha}$ (viewed as a 1–chain), and $c(\mathcal{D})$ is defined similarly, using the $\boldsymbol{\gamma}$–curves. The quantity $a(\mathcal{D}).c(\mathcal{D})$ is defined as the average of the four algebraic intersection numbers of $a'(\mathcal{D})$ and $c(\mathcal{D})$, where $a'(\mathcal{D})$ is a translate of $a(\mathcal{D})$ in any of the four "diagonal directions". If $s \in \alpha_i \cap \beta_j$, then $n_s(\mathcal{D})$ is the average of the multiplicities in the regions surrounding s, and if $\boldsymbol{s}$ is a set of such intersection points, then $n_{\boldsymbol{s}}(\mathcal{D})$ is the sum of the $n_s(\mathcal{D})$ ranging over $s \in \boldsymbol{s}$.

For a homology triangle $\psi \in \pi_2(\Theta^+_{\alpha\beta} \times p_0^+, \boldsymbol{x} \times x, \boldsymbol{y} \times y)$ which can be decomposed into homology classes ψ_Σ, $\phi_{\alpha\beta}$ and ψ_0 as above (as any homology class admitting holomorphic representatives for arbitrarily large neck length can) we observe that $\mu(\psi)$ can be computed by adding up $\mu(\psi_\Sigma)$, $\mu(\phi_{\alpha\beta})$ and $\mu(\psi_0)$, then subtracting the quantities which are over-counted. This corresponds to subtracting

$$\tfrac{1}{2}(m_1 + m_2 + n_1 + n_2)(\psi),$$

which is the excess of Euler measure resulting from removing balls centered at p^+ and at p_0^- $\left(\text{note that the Euler measure of a quarter disk is } \tfrac{1}{4}\right)$, and subtracting

$$2n_{p^+}(\phi_{\alpha\beta}) = \tfrac{1}{2}(n_1 + n_2 + m_1 + m_2)(\phi_{\alpha\beta}),$$

which is the quantity in the expression

$$\begin{aligned}\mu(\phi_{\alpha\beta}) &= e(\mathcal{D}(\phi_{\alpha\beta})) + n_{\Theta^+_{\alpha\beta} \times p^+}(\phi_{\alpha\beta}) + n_{\boldsymbol{y} \times p^+}(\phi_{\alpha\beta}) \\ &= e(\mathcal{D}(\phi_{\alpha\beta})) + n_{\Theta^+_{\alpha\beta}}(\phi_{\alpha\beta}) + n_{\boldsymbol{y}}(\phi_{\alpha\beta}) + \tfrac{1}{2}(n_1 + n_2 + m_1 + m_2)(\phi_{\alpha\beta})\end{aligned}$$

which does not contribute to $\mu(\psi)$. Adding these contributions, we get

$$\mu(\psi) = \mu(\psi_\Sigma) + \mu(\psi_0) + \mu(\phi_{\alpha\beta}) - \tfrac{1}{2}(n_1+n_2+m_1+m_2)(\psi) - \tfrac{1}{2}(n_1+n_2+m_1+m_2)(\phi_{\alpha\beta}). \tag{3}$$

Writing $\psi_0 \in \pi_2(p_0^+, x, y)$, using the vertex multiplicity relations around p_0^-, it is an easy computation that

$$(m_1+m_2)(\psi) = (n_1+n_2)(\psi).$$

Note also that $m_i(\psi_0) = m_i(\psi)$ and similarly for the multiplicities n_i, since we are grouping all holomorphic curves on $(S^2, \alpha_0, \beta_0, \gamma_0)$ appearing in the weak limit into the homology class ψ_0. Using the Maslov index formula from Lemma 6.6 for ψ_0, we get from (3) that

$$\mu(\psi) = \mu(\psi_\Sigma) + \mu(\phi_{\alpha\beta}) + (m_1+m_2+N_1+N_2)(\psi_0) - \tfrac{1}{2}(m_1+m_2+n_1+n_2)(\psi_0) - \tfrac{1}{2}(n_1+n_2+m_1+m_2)(\phi_{\alpha\beta}),$$

which reduces to

$$\mu(\psi) = \mu(\psi_\Sigma) + (N_1+N_2)(\psi_0) + \mu(\phi_{\alpha\beta}) - \tfrac{1}{2}(n_1+n_2+m_1+m_2)(\phi_{\alpha\beta}), \tag{4}$$

since ψ_0 does not have p^+ as a vertex, so the vertex relations at p^+ yield

$$\tfrac{1}{2}(n_1+n_2+m_1+m_2)(\psi_0) = (n_1+n_2)(\psi_0) = (m_1+m_2)(\psi_0).$$

We now use the assumption that $(\Sigma, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{\beta} \cup \{\beta_s\}, \boldsymbol{w})$ is strongly positive with respect to p^+, and hence

$$\mu(\phi_{\alpha\beta}) - \tfrac{1}{2}(n_1+n_2+m_1+m_2)(\phi_{\alpha\beta}) \geq 0,$$

with equality if and only if $\phi_{\alpha\beta}$ is a constant disk. We note that $\mu(\psi_\Sigma) \geq 0$, since ψ_Σ admits a broken holomorphic representative. Hence the formula for $\mu(\psi)$ in (4) can be written as a sum of nonnegative expressions, and hence each must be zero if $\mu(\psi) = 0$. Thus

$$0 = \mu(\psi) = \mu(\psi_\Sigma) = \mu(\phi_{\alpha\beta}) = N_1 = N_2.$$

Note first that this implies that $\phi_{\alpha\beta}$ is a constant disk, since generically there are no nonconstant, Maslov index 0 disks which admit broken holomorphic representatives. From here the argument proceeds in a familiar manner. Note that ψ_Σ is a Maslov index 0 broken holomorphic triangle, and hence must be a genuine holomorphic triangle, as nonconstant holomorphic disks, boundary degenerations, and closed surfaces all have strictly positive Maslov index. Hence, since ψ_Σ is represented by a holomorphic triangle with no boundary components mapped to α_s or β_s, we must have $m_1(\psi) = m_2(\psi) = n_1(\psi) = n_2(\psi)$.

We now claim that this implies that the off-diagonal entries of the matrix representing the triangle map in the statement are zero. Writing $\psi_0 \in \pi_2(p_0^+, x, y)$, and using the multiplicities in Figure 9, the vertex relations at p_0^+ read

$$A + B = 1,$$

and hence exactly one of A and B is 1, and the other is 0, since $A, B \geq 0$. Since $n_1 = n_2 = m_1 = m_2$, we know that by subtracting some number of copies of the γ_0–boundary degeneration of Maslov index 2 with $N_1 = N_2 = 0$, we get a homology triangle in $\pi_2(p_0^+, x, y)$ with N_1, N_2, m_1, m_2, n_1 and n_2 all zero. There are only two homology triangles satisfying that condition. One is in $\pi_2(p_0^+, x^+, y^+)$ and the other is in $\pi_2(p_0^+, x^-, y^-)$, implying that ψ_0 itself must be in one of those sets. Hence the off-diagonal entries of the matrix are zero.

Since u_Σ is a genuine Maslov index 0 holomorphic triangle, there must be a curve in u_0' in the broken holomorphic triangle u_0 which matches, ie which satisfies

$$\rho^{p^+}(u_\Sigma) = \rho^{p_0}(u_0').$$

Recall that if $u\colon S \to \Sigma \times \Delta$ is a holomorphic map and $q \in \Sigma$ is a point, we define

$$\rho^q(u) = (\pi_\Delta \circ u)(\pi_\Sigma \circ u)^{-1}(q) \in \mathrm{Sym}^{n_q(u)}(\Delta).$$

Since this in particular forces $n_{p^+}(u_\Sigma) = n_{p_0^-}(u_0')$, it is easy to see that there can be no other curves in the broken curve u_0 since there are no multiplicities on $(S^2, \alpha_0, \beta_0, \gamma_0)$ which could be increased without increasing n_1, m_1, n_2, m_2, N_1 and N_2 while still preserving the vertex relations. By standard gluing arguments (see eg [5, Appendix A]) the count of prematched triangles[2] is equal to the count of holomorphic triangles in $\#\mathcal{M}_{\mathcal{J}(T)}(\psi_\Sigma \# \psi_0)$ for sufficiently large T.

For x and y of the same relative grading (both the top intersection points or both the lower intersection points), for each k there is a unique homology class ψ_k on $(S^2, \alpha_0, \beta_0, \gamma_0)$ in $\pi_2(p_0^+, x, y)$ with $n_1 = n_2 = m_1 = m_2 = k$. Thus, adapting the proof of Proposition 5.3, it is sufficient to count holomorphic triangles on $(S^2, \alpha_0, \beta_0, \gamma_0)$ of homology class ψ_k which match a fixed divisor $\boldsymbol{d} \in \mathrm{Sym}^k(\Delta)$. The count of such holomorphic triangles matching $\boldsymbol{d}$ is generically 1 (mod 2), as can be seen from a Gromov compactness argument nearly identical to the one done at the

[2]Recall that a prematched triangle is a pair (u_Σ, u_0) where u_Σ and u_0 are holomorphic triangles representing ψ_Σ and ψ_0 on $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ and $(S^2, \alpha_0, \beta_0, \gamma_0)$ respectively and $\rho^{p^+}(u_\Sigma) = \rho^{p_0}(u_0')$.

end of Proposition 5.3. Hence the diagonal entries of the triangle map matrix are as claimed, completing the proof. □

Remark 6.7 Without the "strongly positive" condition, the counts of the previous theorem are false. The bottom of Figure 7 shows a diagram which is not strongly positive, along with a Maslov index 1 disk which could appear on $(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha})$ when we degenerate. Fortunately such pairs $(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha})$ don't appear when proving invariance of the quasistabilization maps, as we don't need to handleslide other $\boldsymbol{\alpha}$ curves across α_s.

7 Invariance of the quasistabilization maps

In this section, we combine the results of the previous section to prove invariance of the quasistabilization maps:

Theorem A *Assume that $(\sigma, \mathfrak{P})$ is a coloring of the basepoints $\boldsymbol{w} \cup \boldsymbol{z}$ which is extended by the coloring $(\sigma', \mathfrak{P})$ of the basepoints $\boldsymbol{w} \cup \boldsymbol{z} \cup \{w, z\}$. The quasistabilization operation induces well-defined maps*

$$S_{w,z}^{+}\colon \mathrm{CFL}_{UV}^{\infty}(Y, L, \boldsymbol{w}, \boldsymbol{z}, \sigma, \mathfrak{P}, \mathfrak{s}) \to \mathrm{CFL}_{UV}^{\infty}(Y, L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\}, \sigma', \mathfrak{P}, \mathfrak{s})$$

and

$$S_{w,z}^{-}\colon \mathrm{CFL}_{UV}^{\infty}(Y, L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\}, \sigma', \mathfrak{P}, \mathfrak{s}) \to \mathrm{CFL}_{UV}^{\infty}(Y, L, \boldsymbol{w}, \boldsymbol{z}, \sigma, \mathfrak{P}, \mathfrak{s}),$$

which are well-defined invariants up to $\mathfrak{P}$–filtered $\mathbb{Z}_2[U_{\mathfrak{P}}]$–chain homotopy.

The proof is to construct the maps for choices of Heegaard diagram and auxiliary data, and show that the maps we describe are independent of that auxiliary data and the choice of diagram.

If $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ is a diagram for $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$, recall from Section 5.1 that if A denotes the component of $\Sigma \setminus \boldsymbol{\alpha}$ containing the basepoints of $(L, \boldsymbol{w}, \boldsymbol{z})$ adjacent to w and z, then we pick a point $p \in A \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta} \cup \boldsymbol{w} \cup \boldsymbol{z})$ and a simple closed curve $\alpha_s \subseteq A \setminus \boldsymbol{\alpha}$ to form a diagram $\overline{\mathcal{H}}_{p,\alpha_s}$. Let $\mathcal{J}$ denote gluing data (see Section 5.2) for performing the special connected sum operation at $p \in \Sigma$ and $p_0 \in S^2$ and gluing almost complex structures on $\mathcal{H}$ and $\mathcal{H}_0$ together.

We now define maps

$$S_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J},T}^{+}\colon \mathrm{CFL}_{UV,J_s}^{\infty}(\mathcal{H}, \mathfrak{s}) \to \mathrm{CFL}_{UV,\mathcal{J}(T)}^{\infty}(\overline{\mathcal{H}}_{p,\alpha_s}, \mathfrak{s})$$

and

$$S_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J},T}^{-}\colon \mathrm{CFL}_{UV,\mathcal{J}(T)}^{\infty}(\overline{\mathcal{H}}_{p,\alpha_s}, \mathfrak{s}) \to \mathrm{CFL}_{UV,J_s}^{\infty}(\mathcal{H}, \mathfrak{s})$$

by the formulas

$$S^{+}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}(\boldsymbol{x}) = \boldsymbol{x} \times \theta^{+}$$

and

$$S^{-}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}(\boldsymbol{x} \times \theta^{+}) = 0, \quad S^{+}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}(\boldsymbol{x} \times \theta^{-}) = \boldsymbol{x},$$

where J_s denotes the almost complex structure on Σ associated to the gluing data $\mathcal{J}$ and T is sufficiently large. Here θ^{+} denotes the top-degree intersection point with respect to the Maslov grading (the grading obtained by using the $\boldsymbol{w}$–basepoints and ignoring the $\boldsymbol{z}$–basepoints).

The maps $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J},T}$ can be extended to the entire $\mathfrak{P}$–filtered chain homotopy-type invariant by pre- and postcomposing with change of diagram maps and change of almost complex structure maps. By functoriality of the change of diagrams maps, we get well-defined maps $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J},T}$ between the coherent chain homotopy type invariants $\mathit{CFL}^{\infty}_{UV}(Y, L, \boldsymbol{w}, \boldsymbol{z}, \sigma, \mathfrak{P}, \mathfrak{s})$ and $\mathit{CFL}^{\infty}_{UV}(Y, L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\}, \sigma', \mathfrak{P}, \mathfrak{s})$, though of course we need to show independence from the choices of $\mathcal{H}$, p, α_s, $\mathcal{J}$ and T.

Any diagram $\overline{\mathcal{H}}$ for $(Y, L, \boldsymbol{w} \cup \{w\}, \boldsymbol{z} \cup \{z\})$ can be connected to one of the form $\overline{\mathcal{H}}_{p,\alpha_s}$ for a diagram $\mathcal{H}$ and a choice of p and α_s by a sequence of handleslides and isotopies of the attaching curves, $(1,2)$–stabilizations, and isotopies of Y relative the basepoints and preserving the link, since we can always find a diagram for the unstabilized link and quasistabilize it, and any two diagrams for the same multibased link can be connected by a sequence of Heegaard moves by [4, Proposition 2.37], for example. This is somewhat unsatisfying since it would be nice to have an actual algorithm for reducing an arbitrary Heegaard diagram for the quasistabilized link to a quasistabilized diagram, but for our purposes it is sufficient to know that such a path exists.

We now begin our proof of invariance of the maps $S^{\pm}_{w,z}$. We first address independence from $\mathcal{J}$ and the parameter T. Recalling Lemma 5.5, there is an N such that if $T, T' > N$ we have

$$\Phi_{\mathcal{J}(T) \to \mathcal{J}(T')} \simeq \begin{pmatrix} \mathrm{id} & 0 \\ 0 & \mathrm{id} \end{pmatrix}.$$

Hence, as with the free stabilization maps from [14], we define the maps $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}$ to be between the complexes $\mathit{CFL}^{\infty}_{J_s,UV}(\mathcal{H}, \mathfrak{s})$ and $\mathit{CFL}^{\infty}_{\mathcal{J}(T),UV}(\overline{\mathcal{H}}_{p,\alpha_s}, \mathfrak{s})$ for any T greater than any such N.

Lemma 7.1 *The maps $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}$ are independent of $\mathcal{J}$ and the parameter T.*

Proof Lemma 5.5 implies that $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J},T}$ is independent of T for any T which is sufficiently large. Let $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}$ denote the common map. Lemma 5.7 implies that the maps $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s,\mathcal{J}}$ are independent of $\mathcal{J}$. We denote the common map from now on by $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s}$. □

Lemma 7.2 *For a fixed diagram $\mathcal{H}$ with fixed $p \in \Sigma$, the maps $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s}$ are independent of α_s.*

Proof This follows from the triangle map computation in Theorem 6.5, which allows one to change the α_s curve through a sequence of isotopies of the α_s curve, and handleslides of the α_s curve over other $\boldsymbol{\alpha}$–curves, each of which can be realized by quasistabilizing a Heegaard triple $(\Sigma, \boldsymbol{\alpha}', \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ along two curves α_s' and α_s with $\alpha_s' \cap \alpha_s = \{p^+, p^-\}$, such that $(\Sigma, \boldsymbol{\alpha}' \cup \{\alpha_s'\}, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{w})$ is strongly positive at p^+ (see Remark 6.4). For each of these moves, by Theorem 6.5 the change of diagrams map can be written as

$$\Phi^{\boldsymbol{\alpha}\cup\{\alpha_s\}\to\boldsymbol{\alpha}'\cup\{\alpha_s'\}}_{\boldsymbol{\beta}\cup\{\beta_s\}} = \begin{pmatrix} \Phi^{\boldsymbol{\alpha}\to\boldsymbol{\alpha}'}_{\boldsymbol{\beta}} & 0 \\ 0 & \Phi^{\boldsymbol{\alpha}\to\boldsymbol{\alpha}'}_{\boldsymbol{\beta}} \end{pmatrix},$$

where the curves in $\boldsymbol{\alpha}'$ are small Hamiltonian isotopies of the curves in $\boldsymbol{\alpha}$. Similarly, by Theorem 6.5 we also have

$$\Phi^{\boldsymbol{\alpha}\cup\{\alpha_s\}\to\boldsymbol{\alpha}'\cup\{\alpha_s\}}_{\boldsymbol{\beta}\cup\{\beta_s\}} = \begin{pmatrix} \Phi^{\boldsymbol{\alpha}\to\boldsymbol{\alpha}'}_{\boldsymbol{\beta}} & 0 \\ 0 & \Phi^{\boldsymbol{\alpha}\to\boldsymbol{\alpha}'}_{\boldsymbol{\beta}} \end{pmatrix},$$

completing the proof. □

We now let $S^{\pm}_{w,z,\mathcal{H},p}$ denote the map $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s}$ for any choice of α_s. We now prove independence from the choice of point $p \in \Sigma$.

Lemma 7.3 *Given a fixed diagram $\mathcal{H}$, the maps $S^{\pm}_{w,z,\mathcal{H},p}$ are independent of the choice of point $p \in \Sigma$.*

Proof Let A denote the component of $\Sigma \setminus \boldsymbol{\alpha}$ containing the basepoints on $\mathbb{L}$ which are adjacent to w and z. Let p and p' be two choices of points in $A \setminus (\boldsymbol{\beta} \cup \boldsymbol{w} \cup \boldsymbol{z})$. Let ϕ_t be an isotopy $\phi_t\colon \Sigma \to \Sigma$ which fixes $\Sigma \setminus A$ and maps p to p'. Recall that the surfaces $\overline{\Sigma}_p$ were well defined up to an isotopy fixing $\boldsymbol{\alpha} \cup \boldsymbol{\beta} \cup \boldsymbol{w} \cup \boldsymbol{z}$ and mapping L to L. Extend $\phi = \phi_1$ to an isotopy of Y which fixes $(\Sigma \setminus A) \cup \boldsymbol{w} \cup \boldsymbol{z} \cup \{w, z\}$ and maps L to L. By definition

$$(\phi)_*(\overline{\Sigma}_p) = \overline{\Sigma}_{p'},$$

as embedded surfaces. The diffeomorphism ϕ fixes all the curves in $\boldsymbol{\alpha}$, but moves some of the $\boldsymbol{\beta}$–curves which pass through the region A.

Observe the factorizations

$$\Phi_{(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},J_s)\to(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\phi_* J_s)} \simeq \Phi^{\boldsymbol{\alpha}}_{\phi_*\boldsymbol{\beta}\to\boldsymbol{\beta}} \circ \phi_* \tag{5}$$

and similarly

$$\Phi_{(\overline{\Sigma}_p,\boldsymbol{\alpha}\cup\{\alpha_s\},\boldsymbol{\beta}\cup\{\beta_0\},\mathcal{J}(T))\to(\overline{\Sigma}_{p'},\boldsymbol{\alpha}\cup\{\phi_*\alpha_s\},\boldsymbol{\beta}\cup\{\phi_*\beta_0\},\phi_*\mathcal{J}(T))} \simeq \Phi^{\boldsymbol{\alpha}\cup\{\phi_*\alpha_s\}}_{(\phi_*\boldsymbol{\beta})\cup\{\phi_*\beta_0\}\to\boldsymbol{\beta}\cup\{\phi_*\beta_0\}}\circ\phi_*.$$

Using Theorem 6.5, for sufficiently stretched almost complex structure we can write

$$\Phi^{\boldsymbol{\alpha}\cup\{\phi_*\alpha_s\}}_{(\phi_*\boldsymbol{\beta})\cup\{\phi_*\beta_0\}\to\boldsymbol{\beta}\cup\{\phi_*\beta_0\}} \simeq \begin{pmatrix}\Phi^{\boldsymbol{\alpha}}_{\phi_*\boldsymbol{\beta}\to\boldsymbol{\beta}} & 0\\ 0 & \Phi^{\boldsymbol{\alpha}}_{\phi_*\boldsymbol{\beta}\to\boldsymbol{\beta}}\end{pmatrix},$$

and hence

$$\Phi^{\boldsymbol{\alpha}\cup\{\phi_*\alpha_s\}}_{(\phi_*\boldsymbol{\beta})\cup\{\phi_*\beta_0\}\to\boldsymbol{\beta}\cup\{\phi_*\beta_0\}}\circ\phi_* \simeq \begin{pmatrix}\Phi^{\boldsymbol{\alpha}}_{\phi_*\boldsymbol{\beta}\to\boldsymbol{\beta}}\circ\phi_* & 0\\ 0 & \Phi^{\boldsymbol{\alpha}}_{\phi_*\boldsymbol{\beta}\to\boldsymbol{\beta}}\circ\phi_*\end{pmatrix}.$$

Combining this with (5), we see that for sufficiently large T, we have

$$\Phi_{(\overline{\Sigma}_p,\boldsymbol{\alpha}\cup\{\alpha_s\},\boldsymbol{\beta}\cup\{\beta_0\},\mathcal{J}(T))\to(\overline{\Sigma}_{p'},\boldsymbol{\alpha}\cup\{\phi_*\alpha_s\},\boldsymbol{\beta}\cup\{\phi_*\beta_0\},\phi_*\mathcal{J}(T))} \simeq \begin{pmatrix}\Phi_{(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},J_s)\to(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\phi_*J_s)} & 0\\ 0 & \Phi_{(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},J_s)\to(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\phi_*J_s)}\end{pmatrix}.$$

Since this map is upper triangular with diagonal entries equal to the change of diagrams maps, the change of diagrams maps commute with the maps $S^{\pm}_{w,z,\mathcal{H},p}$ and $S^{\pm}_{w,z,\mathcal{H},p'}$ in the appropriate sense. Since any two choices of points p and p' at which we can perform quasistabilization are in the region A, we know that the maps on the filtered chain homotopy invariant $\mathit{CFL}^{\infty}_{UV}$ induced by $S^{\pm}_{w,z,\mathcal{H},p,\alpha_s\mathcal{J}}$ and $S^{\pm}_{w,z,\mathcal{H},p',\phi_*\alpha_s,\phi_*\mathcal{J}}$ are equal. Since we proved invariance from the gluing data $\mathcal{J}$ in Lemma 5.7, and we proved invariance from the curve α_s in Lemma 7.2, the proof is thus complete. □

We let $S^{\pm}_{w,z,\mathcal{H}}$ denote the map $S^{\pm}_{w,z,\mathcal{H},p}$ for any choice of p in the component of $\Sigma\setminus\boldsymbol{\alpha}$ containing the basepoints of L adjacent to w and z.

Lemma 7.4 *If $\mathcal{H}$ and $\mathcal{H}'$ are two diagrams for $\mathbb{L}=(L,\boldsymbol{w},\boldsymbol{z})$, then the maps $S^{\pm}_{w,z,\mathcal{H}}$ and $S^{\pm}_{w,z,\mathcal{H}'}$ are filtered chain homotopic.*

Proof Suppose that $\mathcal{H}=(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{w},\boldsymbol{z})$ and $\mathcal{H}'=(\Sigma',\boldsymbol{\alpha}',\boldsymbol{\beta}',\boldsymbol{w},\boldsymbol{z})$ are two diagrams for $(L,\boldsymbol{w},\boldsymbol{z})$. The diagrams $\mathcal{H}$ and $\mathcal{H}'$ are related by a sequence of the following moves:

(1) $\boldsymbol{\alpha}$– and $\boldsymbol{\beta}$–handleslides and isotopies;

(2) (1,2)–stabilizations away from L;

(3) isotopies of Σ inside of Y which fix $\boldsymbol{w}\cup\boldsymbol{z}$, and map L to L.

The maps corresponding to $\boldsymbol{\alpha}$– and $\boldsymbol{\beta}$–handleslides on the unstabilized diagram $\mathcal{H}$ can be computed using triangle maps. For moves of the $\boldsymbol{\beta}$–curves, we simply apply Lemma 6.1 to see that the maps $S^{\pm}_{w,z,\mathcal{H}}$ are invariant under $\boldsymbol{\beta}$–isotopies and handleslides. Theorem 6.5 implies independence under $\boldsymbol{\alpha}$–moves of $\mathcal{H}$ for which there are curves α_s and α'_s in Σ, with top graded intersection point $p \in \alpha'_s \cap \alpha_s$, such that $(\Sigma, \boldsymbol{\alpha}' \cup \{\alpha'_s\}, \boldsymbol{\alpha} \cup \{\alpha_s\}, \boldsymbol{w})$ is strongly positive with respect to p. An arbitrary $\boldsymbol{\alpha}$–move can be realized as a sequence of such moves, along with moves of the point p inside of the region of $\Sigma \setminus \boldsymbol{\alpha}$. Since we've already shown invariance under each of these smaller moves, the maps $S^{\pm}_{w,z,\mathcal{H}}$ are unchanged by handleslides and isotopies of the $\boldsymbol{\alpha}$– and $\boldsymbol{\beta}$–curves.

The maps $S^{\pm}_{w,z,\mathcal{H}}$ obviously commute with the $(1,2)$–stabilization maps.

We now consider isotopies $\phi_t\colon Y \to Y$ which fix $\boldsymbol{w} \cup \boldsymbol{z}$ and map L to L. We note that tautologically we have that

$$\phi_* \circ S^{\pm}_{w,z,\mathcal{H},p,\mathcal{J}} = S^{\pm}_{w,z,\phi_*\mathcal{H},\phi_* p,\phi_*\mathcal{J}} \circ \phi_*.$$

Since we already know that $S^{\pm}_{w,z,\mathcal{H},p,\mathcal{J}}$ is independent from p and $\mathcal{J}$, we thus conclude that $S^{\pm}_{w,z,\mathcal{H}}$ and $S^{\pm}_{w,z,\phi_*\mathcal{H}}$ agree. □

We can now write $S^{\pm}_{w,z}$ for the quasistabilization maps, completing the proof of Theorem A.

Remark 7.5 Given that the triangle map computations in Lemma 6.1 and Theorem 6.5 showed that change of diagrams maps were not only upper triangular, but diagonal, one may ask why it is natural to define $S^{+}_{w,z}$ by $\boldsymbol{x} \mapsto \boldsymbol{x} \times \theta^{+}$ and not $\boldsymbol{x} \mapsto \boldsymbol{x} \times \theta^{-}$. We remark that $\boldsymbol{x} \mapsto \boldsymbol{x} \times \theta^{-}$ is only a chain map when w is given the same color as the other $\boldsymbol{w}$–basepoint adjacent to z, and indeed $\boldsymbol{x} \mapsto \boldsymbol{x} \times \theta^{-}$ is equal to $\Psi_z S^{+}_{w,z}$. The map Ψ_z is only a chain map if the $\boldsymbol{w}$–basepoints adjacent to z have the same color.

Remark 7.6 We have defined quasistabilization maps $S^{\pm}_{w,z}$ in the case that w comes after z and showed that such maps were invariants, ie that they yielded well-defined maps $S^{\pm}_{w,z}$ on the coherent filtered chain homotopy type invariant. These maps were only constructed if w came after z on the link component. In the case that z comes after w, we can define quasistabilization maps $S^{\pm}_{z,w}$ analogously, picking a choice of β_s. We could call such an operation β quasistabilization. There is no ambiguity between α quasistabilizations or β quasistabilizations because $S^{\pm}_{w,z}$ is always an α quasistabilization and $S^{\pm}_{z,w}$ is always a β quasistabilization.

8 Commutation of quasistabilization maps

In this section we show that if $\{w, z\} \cap \{w', z'\} = \varnothing$, then the maps $S^{\pm}_{w,z}$ and $S^{\pm}_{w',z'}$ all commute. In [14] we showed that the free stabilization maps commute, though commutation was easier to show in that setting, since we could just pick a diagram and an almost complex structure where both free stabilization maps could be computed, and by simply looking at the formulas, one could observe that the maps commuted. In the case of quasistabilization, we cannot always pick an almost complex structure which can be used to compute both maps. Nevertheless, we can compute enough components of the change of almost complex structure map to show that quasistabilization maps commute:

Theorem 8.1 *Suppose that $(L, \boldsymbol{w}, \boldsymbol{z})$ is a multibased link in Y^3 and that w, z, w' and z' are new basepoints such that (w, z) and (w', z') are each pairs of adjacent basepoints on $(L, \boldsymbol{w} \cup \{w, w'\}, \boldsymbol{z} \cup \{z, z'\})$. Then*

$$S^{\circ_1}_{w,z} \circ S^{\circ_2}_{w',z'} \simeq S^{\circ_2}_{w',z'} \circ S^{\circ_1}_{w,z}$$

for any $\circ_1, \circ_2 \in \{+, -\}$.

Pick a diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ for $(L, \boldsymbol{w}, \boldsymbol{z})$, and let α_s and α_s' be curves in $\Sigma \setminus \boldsymbol{\alpha}$ along which we can perform quasistabilization for (w, z) and (w', z'), respectively. Let β_0 and β_0' denote the new $\boldsymbol{\beta}$–curves. Let $\mathcal{J}$ denote gluing data for stretching along circles bounding β_0 and β_0'. There are two distinct cases to consider, corresponding to whether the pairs (w, z) and (w', z') are adjacent or not: either α_s and α_s' lie in the same component of $\Sigma \setminus \boldsymbol{\alpha}$ (this case corresponds to having the pair (w, z) be adjacent to the pair (w', z')), or α_s and α_s' lie in different components of $\Sigma \setminus \boldsymbol{\alpha}$ (this case corresponds to the pair (w, z) not being adjacent to (w', z')).

The first case is the easier to consider. In this case, we now show that we can pick an almost complex structure which computes both quasistabilization maps. To this end, we have the following lemma:

Lemma 8.2 *Suppose that $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ is a diagram as in the previous paragraph with new curves α_s and α_s' for quasistabilizing at (w, z) and (w', z'), respectively. If α_s and α_s' are not in the same component of $\Sigma \setminus \boldsymbol{\alpha}$, then for all sufficiently large T_1, T_1', T_2, T_2', we have*

$$\Phi_{\mathcal{J}(T_1, T_1') \to \mathcal{J}(T_2, T_2')} \simeq \mathrm{id}$$

with respect to the obvious identification between the two complexes.

Proof To show this, we will perform a computation similar to the one in Proposition 5.3, but for Maslov index 0 disks. Let A be the component of $\Sigma\setminus\boldsymbol{\alpha}$ which contains α_s and let A' denote the component of $\Sigma\setminus\boldsymbol{\alpha}$ which contains α_s'. By assumption $A\neq A'$. Let A_1 and A_2 denote the two components of $A\setminus\alpha_s$. Let A_1' and A_2' denote the components of $A'\setminus\alpha_s'$. Suppose that ϕ is a Maslov index 0 homology disk on the doubly quasistabilized diagram

$$(\overline{\Sigma},\boldsymbol{\alpha}\cup\{\alpha_s,\alpha_s'\},\boldsymbol{\beta}\cup\{\beta_0,\beta_0'\},\boldsymbol{w}\cup\{w,w'\},\boldsymbol{z}\cup\{z,z'\}).$$

Write $\phi=\phi_\Sigma\#\phi_0\#\phi_0'$, where ϕ_Σ is a homology class on $(\Sigma,\boldsymbol{\alpha}\cup\{\alpha_s,\alpha_s'\},\boldsymbol{\beta},\boldsymbol{w},\boldsymbol{z})$, ϕ_0 is a homology class on $(S^2,\alpha_0,\beta_0,w,z)$ and $(S^2,\alpha_0',\beta_0',w,z)$. Suppose that $T_{1,n}$, $T_{1,n}'$, $T_{2,n}$ and $T_{2,n}'$ are sequences of neck lengths all approaching $+\infty$ and let $\widehat{\mathcal{J}}_n$ denote interpolating almost complex structures between $\mathcal{J}(T_{1,n},T_{1,n}')$ and $\mathcal{J}(T_{2,n},T_{2,n}')$. Pick $\widehat{\mathcal{J}}_n$ so that as $n\to\infty$ the almost complex structures $\widehat{\mathcal{J}}_n$ split into $J_s\vee J_{S^2}\vee J_{S^2}$ on $(\Sigma\vee S^2\vee S^2)\times[0,1]\times\mathbb{R}$. If u_n is a sequence of Maslov index 0 $\widehat{\mathcal{J}}_n$–holomorphic curves representing ϕ, we can extract a weak limit to broken curves U_Σ, U_0 and U_0' on $(\Sigma,\boldsymbol{\alpha}\cup\{\alpha_s,\alpha_s'\},\boldsymbol{\beta},\boldsymbol{w},\boldsymbol{z})$, (S^2,α_0,β_0) and (S^2,α_0',β_0') representing ϕ_Σ,ϕ_0 and ϕ_0' respectively. As in Proposition 5.3, the curves in U_Σ consist of a broken holomorphic strip U_Σ' on $(\Sigma,\boldsymbol{\alpha},\boldsymbol{\beta})$ and a collection $\mathcal{A}$ of cylindrical $(\boldsymbol{\alpha}\cup\{\alpha_s\}\cup\{\alpha_s'\})$–boundary degenerations. Let ϕ_Σ' denote the underlying homology class of U_Σ'. Let m_1, m_2, n_1, n_2, m_1', m_2', n_1' and n_2' be multiplicities as in Figure 10.

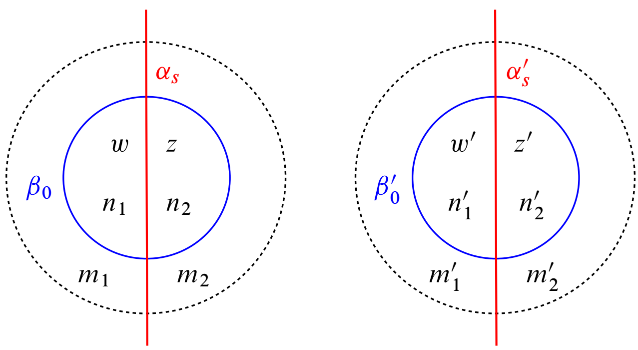


Figure 10: Multiplicities of a disk ϕ near new basepoints w,z,w' and z' on a diagram which has been quasistabilized twice

Adapting the Maslov index computation from Proposition 5.3, we see that

$$\begin{aligned}\mu(\phi)=\mu(\phi_\Sigma')+n_1(\phi)+n_2(\phi)+n_1'(\phi)+n_2'(\phi)&\\ +m_1(\mathcal{A})+m_2(\mathcal{A})+m_1'(\mathcal{A})+m_2'(\mathcal{A})+2&\sum_{\substack{\mathcal{D}\in C(\Sigma\setminus\boldsymbol{\alpha})\\ \mathcal{D}\neq A}}n_{\mathcal{D}}(\mathcal{A}),\end{aligned}$$

where $C(\Sigma\setminus\boldsymbol{\alpha})$ denotes the connected components of $\Sigma\setminus\boldsymbol{\alpha}$.

Since U'_Σ is a broken holomorphic curve for an $\mathbb{R}$–invariant almost complex structure, we conclude that U'_Σ consists only of constant flowlines. Since all of the other summands are zero, it's easy to see that this forces all multiplicities to be zero. Hence only constant disks are counted by the change of almost complex structures map, concluding the proof of the lemma. □

In the case that α_s and α'_s are in the same component, the change of almost complex structure maps will often be nontrivial. Nevertheless, we have the following:

Lemma 8.3 *Suppose (w',z') and (w,z) are adjacent on $(L,\boldsymbol{w}\cup\{w,w'\},\boldsymbol{z}\cup\{z,z'\})$ and that (w',z') comes after (w,z). Let $\theta^{\pm}$ and $(\theta')^{\pm}$ denote the intersection points corresponding to quasistabilization. If T_1, T'_1, T_2, T'_2 are all sufficiently large, then, writing $F=\Phi_{\mathcal{J}(T_1,T'_1)\to\mathcal{J}(T_2,T'_2)}$, we have*

$$\begin{aligned}
F(\boldsymbol{x}\times\theta^+\times(\theta')^+)&=\boldsymbol{x}\times\theta^+\times(\theta')^+,\\
F(\boldsymbol{x}\times\theta^+\times(\theta')^-)&=\boldsymbol{x}\times\theta^+\times(\theta')^-+C\cdot\boldsymbol{x}\times\theta^-\times(\theta')^+,\\
F(\boldsymbol{x}\times\theta^-\times(\theta')^+)&=\boldsymbol{x}\times\theta^-\times(\theta')^+,\\
F(\boldsymbol{x}\times\theta^-\times(\theta')^-)&=\boldsymbol{x}\times\theta^-\times(\theta')^-
\end{aligned}$$

for some C (which is not independent of T_i and T'_i).

Proof We proceed similarly to the previous lemma. Now a single component, which we denote by A, of $\Sigma\setminus\boldsymbol{\alpha}$ contains both α_s and α'_s. Write A_1, A_2 and A_3 for the three different components of $A\setminus(\alpha_s\cup\alpha'_s)$. Two of the A_i share boundary with exactly one of the other A_j, and one of the A_i shares boundary with both of the other A_i. Without loss of generality assume that A_1 shares boundary with A_2, and that A_2 also shares boundary with A_3.

As before, as we simultaneously stretch the necks, a sequence of Maslov index 0 disks u_i has a weak limit as before. Now, however, the Maslov index computation is different. Let $a_i(\mathcal{A})$ denote the multiplicity of the $(\boldsymbol{\alpha}\cup\{\alpha_s\}\cup\{\alpha'_s\})$–degeneration $\mathcal{A}$ in the region A_i. One computes that the Maslov index now satisfies

$$\mu(\phi)=\mu(\phi'_\Sigma)+n_1(\phi)+n_2(\phi)+n'_1(\phi)+n'_2(\phi)+a_1(\mathcal{A})+a_3(\mathcal{A})+2\sum_{\substack{\mathcal{D}\in C(\Sigma\setminus\boldsymbol{\alpha})\\ \mathcal{D}\neq A}}n_{\mathcal{D}}(\mathcal{A}).$$

As usual, this implies that all of the terms above are zero. Hence ϕ'_Σ, which has a broken representative for a cylindrical almost complex structure, must be the constant disk by transversality. The only multiplicities which may be nonzero are $a_2(\mathcal{A})$, $m_i(\phi)$ and $m'_i(\phi)$, none of which appear in the sum above. As is easily observed, this constrains the disk ϕ to be in $\pi_2(\boldsymbol{x} \times \theta^+ \times (\theta')^-, \boldsymbol{x} \times \theta^- \times (\theta')^+)$, completing the proof. An example of a disk which might appear in the change of almost complex structure map is shown in Figure 11. □

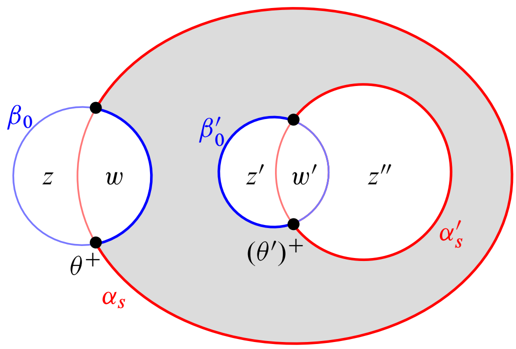


Figure 11: An example of a Maslov index 0 disk which might be counted by $\Phi_{\mathcal{J}(T_1,T_1')\to\mathcal{J}(T_2,T_2')}$ in Lemma 8.3 for arbitrarily large T_i and T_i'

Proof of Theorem 8.1 The proof is easy algebra in all cases using Lemmas 8.2 and 8.3. □

9 Further relations between the maps Ψ_z, Φ_w and $S^{\pm}_{w,z}$

In this section we prove several relations between the maps $S^{\pm}_{w,z}$, Ψ_z and Φ_w. We highlight the convenience of viewing Φ_w and Ψ_z as formal derivatives of the differential, since basically all of the relations in this section are derived by either formally differentiating the expression for $\partial \circ \partial$ from Lemma 2.1, or by differentiating our expression of the quasistabilized differential in Proposition 5.3.

Lemma 9.1 *If w and z are not adjacent, or if w and z are the only basepoints on a link component, then*

$$\Phi_w \Psi_z + \Psi_z \Phi_w \simeq 0.$$

If w and z are adjacent and there are other basepoints on the link component, then

$$\Phi_w \Psi_z + \Psi_z \Phi_w \simeq \mathrm{id}.$$

Proof Take the expression for ∂^2 from Lemma 2.1 and differentiate it with respect to U_w. We obtain

$$\partial \Phi_w + \Phi_w \partial = V_{z'} + V_{z''},$$

where z' and z'' are the variables adjacent to w on its link component. Suppose first that $z' \neq z''$, ie that w and z are not the only basepoints on their link component.

Differentiating the expression above with respect to V_z, we see that

$$\Psi_z \Phi_w + \Phi_w \Psi_z \simeq \begin{cases} \mathrm{id} & \text{if } w \text{ is adjacent to } z, \\ 0 & \text{if } w \text{ is not adjacent to } z, \end{cases}$$

from which the claim follows as long as w and z are not the only basepoints on their link component.

If w and z are the only basepoints on their link component, then $z = z' = z''$ and the argument above is easily modified to give the stated result. □

Lemma 9.2 *We have*

$$\Psi_z \Psi_{z'} + \Psi_{z'} \Psi_z \simeq 0 \quad \textit{and} \quad \Phi_w \Phi_{w'} + \Phi_{w'} \Phi_w \simeq 0$$

for any choice of z, z', w and w'.

Proof This is proven identically to the previous lemma. □

As with the free stabilization maps in [14], we have the following:

Lemma 9.3 *The following relation holds:*

$$S^+_{w,z} S^-_{w,z} = \Phi_w.$$

Proof The differential on the uncolored quasistabilized diagram takes the form

$$\partial_{\overline{\mathcal{H}}} = \begin{pmatrix} \partial_{\mathcal{H}} & U_w + U_{w'} \\ V_z + V_{z'} & \partial_{\mathcal{H}} \end{pmatrix}$$

where $\partial_{\mathcal{H}}$ is the differential on the unstabilized diagram. After taking the U_w derivative we get

$$\Phi_w = \begin{pmatrix} 0 & \mathrm{id} \\ 0 & 0 \end{pmatrix},$$

which is exactly $S^+_{w,z} S^-_{w,z}$. □

We now consider commutators of the quasistabilization maps and the maps Φ_w and Ψ_z.

Lemma 9.4 *Suppose that w and z are two new basepoints for a link. If w is not adjacent to z', then*

$$S^{\pm}_{w,z} \Psi_{z'} + \Psi_{z'} S^{\pm}_{w,z} \simeq 0.$$

With no assumptions on adjacency, we have

$$S^{\pm}_{w,z} \Phi_{w'} + \Phi_{w'} S^{\pm}_{w,z} \simeq 0.$$

Proof Suppose that w and z are inserted between basepoints z'' and w'' on the link $\mathbb{L}$. The quasistabilized differential takes the form

$$\partial_{\overline{\mathcal{H}}} = \begin{pmatrix} \partial_{\mathcal{H}} & U_w + U_{w''} \\ V_z + V_{z''} & \partial_{\mathcal{H}} \end{pmatrix},$$

by Proposition 5.3. By assumption $z' \neq z''$. Differentiating with respect to $V_{z'}$ thus yields

$$\widetilde{\Psi}_{z'} = \begin{pmatrix} \Psi_{z'} & 0 \\ 0 & \Psi_{z'} \end{pmatrix},$$

where $\widetilde{\Psi}_{z'}$ denotes the map on the stabilized diagram and $\Psi_{z'}$ denotes the map on the unstabilized diagram. In matrix notation, the maps $S^{\pm}_{w,z}$ take the form

$$S^{+}_{w,z} = \begin{pmatrix} \mathrm{id} \\ 0 \end{pmatrix} \quad \text{and} \quad S^{-}_{w,z} = \begin{pmatrix} 0 & \mathrm{id} \end{pmatrix}. \tag{6}$$

The stated equality involving $\Psi_{z'}$ now follows from matrix multiplication.

The equality involving $\Phi_{w'}$ follows similarly. □

We also have the following:

Lemma 9.5 *Suppose that z' is adjacent to w and that $z' \neq z$. Then we have*

$$S^{+}_{w,z}\Psi_{z'} \simeq (\Psi_{z'} + \Psi_z)S^{+}_{w,z} \quad \textit{and} \quad \Psi_{z'}S^{-}_{w,z} \simeq S^{-}_{w,z}(\Psi_{z'} + \Psi_z).$$

Proof Once again we consider the quasistabilized differential, which is

$$\partial_{\overline{\mathcal{H}}} = \begin{pmatrix} \partial_{\mathcal{H}} & U_w + U_{w'} \\ V_z + V_{z'} & \partial_{\mathcal{H}} \end{pmatrix}.$$

Differentiating with respect to z' yields

$$\widetilde{\Psi}_{z'} = \begin{pmatrix} \Psi_{z'} & 0 \\ \mathrm{id} & \Psi_{z'} \end{pmatrix} \quad \text{and} \quad \widetilde{\Psi}_{z} = \begin{pmatrix} 0 & 0 \\ \mathrm{id} & 0 \end{pmatrix}.$$

Here $\widetilde{\Psi}_z$ denotes the map on the complex after quasistabilization, and Ψ_z denotes the map on the complex before quasistabilization. Using the matrix notation from (6), the desired relations follow from matrix multiplication. □

The reader can compare the following lemma to [14, Lemma 7.7], the analogous result for the closed three-manifold invariants.

Lemma 9.6 *Suppose that $\mathbb{L} = (L, \boldsymbol{w}, \boldsymbol{z})$ is a multibased link in Y^3 and w and z are two new, consecutive basepoints on $\mathbb{L}$ such that w follows z. If z' is one of the two $\boldsymbol{z}$–basepoints adjacent to w, then we have*

$$S^-_{w,z}\Psi_{z'}S^+_{w,z} \simeq \mathrm{id}.$$

Proof This follows from our usual strategy. Pick a diagram $\mathcal{H}$ for $(L, \boldsymbol{w}, \boldsymbol{z})$ and let $\bar{\mathcal{H}}$ denote a diagram which has been quasistabilized at w and z. Let z'' and w'' denote the basepoints adjacent to w and z on $\mathbb{L}$. Using Proposition 5.3, we have that

$$\partial_{\bar{\mathcal{H}}} = \begin{pmatrix} \partial_{\mathcal{H}} & U_w + U_{w''} \\ V_z + V_{z''} & \partial_{\mathcal{H}} \end{pmatrix}.$$

By assumption, either $z' = z$ or $z' = z''$ (but not both). In both cases, we have that

$$\Psi_{z'} = \Big(\frac{d}{dV_{z'}}\partial_{\bar{\mathcal{H}}}\Big) = \begin{pmatrix} * & * \\ \mathrm{id} & * \end{pmatrix},$$

where the $*$ terms are unimportant. Using the matrix notation from (6), we get the desired equality immediately from matrix multiplication. □

The reader should compare the following to [13, Lemma 4.4].

Lemma 9.7 *We have $\Psi_z^2 \simeq 0$ and $\Phi_w^2 \simeq 0$ as $\mathfrak{P}$–filtered maps of $\mathbb{Z}_2[U_{\mathfrak{P}}]$–modules.*

Proof The proof follows identically to the proof of [14, Lemma 14.19]. □

10 Basepoint moving maps

In this section, we compute several basepoint moving maps. The procedure for computing maps induced by moving basepoints is in a similar spirit to the author's computation of the π_1–action on the Heegaard Floer homology of a closed three-manifold in [14]. We first compute the effect of moving basepoints along a small arc on a link component via a model computation. We then use this to prove Theorem B, the effect of moving all of the basepoints on a link component in one full loop. The final computation is Theorem D, where we compute the effect on certain colored complexes of moving each $\boldsymbol{w}$–basepoint to the next $\boldsymbol{w}$–basepoint, and moving each $\boldsymbol{z}$–basepoint to the next $\boldsymbol{z}$–basepoint.

10.1 Moving basepoints along an arc

Suppose that Y^3 is a three-manifold with embedded multipointed link $\mathbb{L}_0 = (L, \boldsymbol{w}_0, \boldsymbol{z}_0)$, though we allow the case that one of the components of $\mathbb{L}_0$ has no basepoints. Suppose that z, w, z', w' are all points on a single component of $L \setminus (\boldsymbol{w}_0 \cup \boldsymbol{z}_0)$, appearing in that order according to the orientation of $\mathbb{L}$. Let

$$\boldsymbol{w} = \boldsymbol{w}_0 \cup \{w\}, \quad \boldsymbol{z} = \boldsymbol{z}_0 \cup \{z\}, \quad \boldsymbol{w}' = \boldsymbol{w}_0 \cup \{w'\}, \quad \boldsymbol{z}' = \boldsymbol{z}_0 \cup \{z'\}.$$

Finally, assume that $(L, \boldsymbol{w}, \boldsymbol{z})$ has basepoints in each component of L. There is an isotopically unique diffeomorphism of Y which maps L to itself and fixes $\boldsymbol{w}_0 \cup \boldsymbol{z}_0$ and maps w to w' and z to z', which is isotopic to the identity relative to $\boldsymbol{w}_0 \cup \boldsymbol{z}_0$ through isotopies which map L to itself. Let ς_0 denote this diffeomorphism. It induces a map

$$(\varsigma_0)_*\colon \mathrm{CFL}^\infty_{UV}(Y, L, \boldsymbol{w}, \boldsymbol{z}, \mathfrak{s}) \to \mathrm{CFL}^\infty_{UV}(Y, L, \boldsymbol{w}', \boldsymbol{z}', \mathfrak{s}).$$

The map $(\varsigma_0)_*$ is defined as a tautology. That is, if $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{z})$ is a diagram for $(Y, L, \boldsymbol{w}, \boldsymbol{z})$ with almost complex structure J_s, we just apply the map ς_0 to the diagram $\mathcal{H}$ to get a new diagram

$$\varsigma_0(\mathcal{H}) = (\varsigma_0(\Sigma), \varsigma_0(\boldsymbol{\alpha}), \varsigma_0(\boldsymbol{\beta}), \varsigma_0(\boldsymbol{w}), \varsigma_0(\boldsymbol{z})),$$

with almost complex structure $(\varsigma_0)_* J_s$. The diffeomorphism ς_0 tautologically determines a chain map

$$(\varsigma_0)_*\colon \mathrm{CFL}^\infty_{UV,J_s}(\mathcal{H}, \mathfrak{s}) \to \mathrm{CFL}^\infty_{UV,\varsigma_0(J_s)}(\varsigma_0(\mathcal{H}), \mathfrak{s})$$

defined by

$$(\varsigma_0)_*(\boldsymbol{x}) = \varsigma_0(\boldsymbol{x}).$$

By the naturality results of [4], this yields a well-defined morphism on the coherent chain homotopy type invariants (ie it commutes with change of diagrams maps, in the appropriate sense).

Note that $(\varsigma_0)_*$ "appears" like the identity map since it just maps an intersection point $\boldsymbol{x}$ to its image under ς_0. With this in mind, we prove the following:

Lemma 10.1 *The induced map $(\varsigma_0)_*$ is filtered chain homotopic to*

$$(\varsigma_0)_* \simeq S^-_{w,z} \Psi_{z'} S^+_{w',z'}.$$

Proof We first prove the result in the case that the link component containing w and z has at least one extra pair of basepoints. Let z'' denote the basepoint occurring immediately after w'. In this case, we can pick a diagram like the one shown in Figure 12, where the dashed lines show two circles along which the almost complex structure will be stretched. In this diagram, we assume that α_s and α_s' each bound disks on Σ and that α_s, α_s', β_0 and β_0' do not intersect any other $\boldsymbol{\alpha}$– or $\boldsymbol{\beta}$–curves. With this diagram, we can compute all of the maps $\Psi_{z'}$, $S^-_{w,z}$, and $S^+_{w',z'}$ explicitly. We must be careful, though, since we cannot use the same almost complex structure for all of the maps. Instead we will need to use the change of almost complex structure computation from Lemma 8.3. Let J_s be an almost complex structure which is sufficiently stretched along c to compute $S^{\pm}_{w,z}$, and let J_s' be an almost complex structure which is sufficiently stretched along c' to compute $S^{\pm}_{w',z'}$, and assume that both are stretched sufficiently so that the change of almost complex structure map $\Phi_{J_s'\to J_s}$ takes the form described in Lemma 8.3. We wish to compute $S^{-}_{w,z}\circ\Psi_{z'}\circ\Phi_{J_s'\to J_s}\circ S^{+}_{w',z'}$.

Let $\theta^{\pm}$ denote the intersection points of $\alpha_s\cap\beta_0$ and let $(\theta')^{\pm}$ denote the intersection points of α_s' and β_0'.

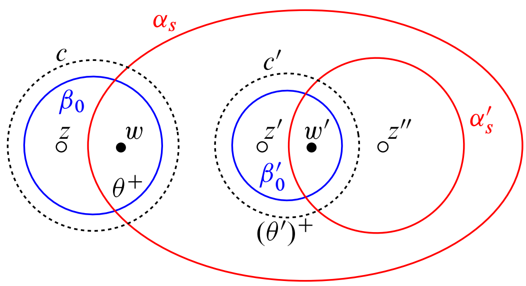


Figure 12: A diagram for Lemma 10.1 when we have another basepoint z'' on the link component containing w,z,w' and z'. The curves α_s and α_s' each bound disks, and $\alpha_s,\alpha_s',\beta_0,\beta_0'$ do not intersect any other $\boldsymbol{\alpha}$– or $\boldsymbol{\beta}$–curves.

Using the analysis in Proposition 5.3, we see that for J_s there are exactly two domains which are the domain of Maslov index 1 disks ϕ which support holomorphic representatives with $n_{z'}(\phi)>0$. These domains are shown in Figure 13. Also every homology disk ϕ which has one of these domains has $\#\widehat{\mathcal{M}}_{J_s}(\phi)=1$.

We wish to show that $S^{-}_{w,z}\circ\Psi_{z'}\circ\Phi_{J_s'\to J_s}\circ S^{+}_{w',z'}=(\varsigma_0)_*$, where ς_0 is the diffeomorphism induced by simply pushing w and z to w' and z', respectively. To this end, it is sufficient to show that

$$(S^{-}_{w,z}\circ\Psi_{z'}\circ\Phi_{J_s'\to J_s}\circ S^{+}_{w',z'})(\boldsymbol{x}\times\theta^{\pm})=\boldsymbol{x}\times(\theta')^{\pm},$$

since $(\varsigma_0)_*(\boldsymbol{x}\times\theta^{\pm})=\boldsymbol{x}\times(\theta')^{\pm}$.

Homology disks with the left domain in Figure 13 yield a contribution to $\Psi_{z'}$ of

$$\boldsymbol{x}\times\theta^{\pm}\times(\theta')^{+}\longrightarrow\boldsymbol{x}\times\theta^{\pm}\times(\theta')^{-}.$$

Homology disks with the right domain in Figure 13 yield a contribution to $\Psi_{z'}$ of

$$\boldsymbol{x}\times\theta^{+}\times(\theta')^{\pm}\longrightarrow\boldsymbol{x}\times\theta^{-}\times(\theta')^{\pm}.$$

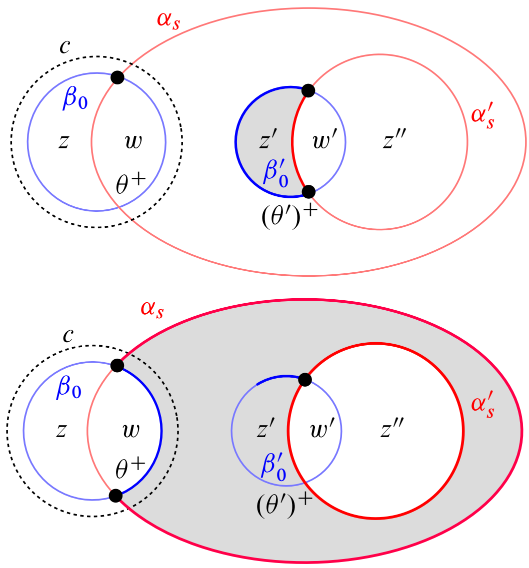


Figure 13: The two domains contributing to $\Psi_{z'}$ in Lemma 10.1 for the almost complex structure J_s stretched first on c', and then stretched on c (possibly much more than on c'). Also drawn in are two examples of holomorphic disks with those domains.

We first compute $(S^-_{w,z} \circ \Psi_{z'} \circ \Phi_{J_s' \to J_s} \circ S^+_{w',z'})(\boldsymbol{x} \times \theta^+)$. Using the computation of $\Psi_{z'}$ above and the computation of $\Phi_{J_s' \to J_s}$ from Lemma 8.3, we have that

$$\begin{aligned}(S^-_{w,z} \circ \Psi_{z'} \circ \Phi_{J_s' \to J_s} \circ S^+_{w',z'})(\boldsymbol{x} \times \theta^+) &= (S^-_{w,z} \circ \Psi_{z'} \circ \Phi_{J_s' \to J_s})(\boldsymbol{x} \times \theta^+ \times (\theta')^+) \\ &= (S^-_{w,z} \circ \Psi_{z'})(\boldsymbol{x} \times \theta^+ \times (\theta')^+) \\ &= S^-_{w,z}(\boldsymbol{x} \times \theta^+ \times (\theta')^- + \boldsymbol{x} \times \theta^- \times (\theta')^+) \\ &= \boldsymbol{x} \times (\theta')^+.\end{aligned}$$

We now compute $(S^-_{w,z} \circ \Psi_{z'} \circ \Phi_{J_s' \to J_s} \circ S^+_{w',z'})(\boldsymbol{x} \times \theta^-)$. Once again using our previous computation of $\Psi_{z'}$ and Lemma 8.3, we have that

$$\begin{aligned}(S^-_{w,z} \circ \Psi_{z'} \circ \Phi_{J_s' \to J_s} \circ S^+_{w',z'})(\boldsymbol{x} \times \theta^-) &= (S^-_{w,z} \circ \Psi_{z'} \circ \Phi_{J_s' \to J_s})(\boldsymbol{x} \times \theta^- \times (\theta')^+) \\ &= (S^-_{w,z} \circ \Psi_{z'})(\boldsymbol{x} \times \theta^- \times (\theta')^+) \\ &= S^-_{w,z}(\boldsymbol{x} \times \theta^- \times (\theta')^-) \\ &= \boldsymbol{x} \times (\theta')^-,\end{aligned}$$

completing the proof of the claim if $\boldsymbol{w}$ and $\boldsymbol{z}$ each have at least two basepoints on L.

We now consider the case that $\mathbb{L}$ doesn't have any basepoints other than w and z. In this case we just introduce two new basepoints w'', z'' which are on the component of $L \setminus \{w, w', z, z'\}$ which goes from w' to z. Note that ς_0 is isotopic relative to $\{w, z\}$ to a diffeomorphism which fixes w'' and z''. Hence $(\varsigma_0)_* S^-_{w'',z''} = S^-_{w'',z''}(\varsigma_0)_*$. We just compute that

$$\begin{aligned}
(\varsigma_0)_* &\simeq (\varsigma_0)_*(S^-_{w'',z''}\Psi_{z''}S^+_{w'',z''}) && \text{(Lemma 9.6)}\\
&\simeq S^-_{w'',z''}(\varsigma_0)_*\Psi_{z''}S^+_{w'',z''} && \text{(observation above)}\\
&\simeq S^-_{w'',z''}(S^-_{w,z}\Psi_{z'}S^+_{w',z'})\Psi_{z''}S^+_{w'',z''} && \text{(previous case)}\\
&\simeq (S^-_{w,z}\Psi_{z'}S^+_{w',z'})(S^-_{w'',z''}\Psi_{z''}S^+_{w'',z''}) && \text{(Theorem 8.1, Lemma 9.4)}\\
&\simeq S^-_{w,z}\Psi_{z'}S^+_{w',z'} && \text{(Lemma 9.6)}
\end{aligned}$$

as we wanted. □

10.2 Sarkar's formula for moving basepoints in a full twist around a link component

In this section, we prove Theorem B, which is Sarkar's conjectured formula for the effect of moving basepoints on a link component in a full twist around the link component for the full link Floer complex. The main technical tool is Lemma 10.1, which computes the effect of moving basepoints on a small arc on a link component. By writing the diffeomorphism of a full twist as a composition of many smaller moves of the previous form, we will obtain Sarkar's formula.

Theorem B *Suppose ς is the diffeomorphism corresponding to a positive Dehn twist around a link component K of $\mathbb{L}$. Suppose that the basepoints on K are $w_1, z_1, \dots, w_n, z_n$. The induced map ς_* on $\mathrm{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s})$ has the $\mathfrak{P}$–filtered $\mathbb{Z}_2[U_{\mathfrak{P}}]$ chain homotopy type*

$$\varsigma_* \simeq \mathrm{id} + \Phi_K \Psi_K,$$

where

$$\Phi_K = \sum_{j=1}^n \Phi_{w_j} \quad \textit{and} \quad \Psi_K = \sum_{j=1}^n \Psi_{z_j}.$$

To the reader who is not interested in colorings, we note that one can just take $\mathfrak{P} = \boldsymbol{w} \cup |L|$, where $|L|$ denotes the components of L.

In this section, we also introduce some new formalism to make the computation easier. The maps $\Psi_{z'}$ and $S^{\pm}_{w,z}$ interact strangely (eg Lemma 9.5), which leads to challenging and messy algebra if we are not careful. Suppose that A is an arc on L between two $\boldsymbol{w}$–basepoints which share the same color. We define the map

$$\Psi_A = \sum_{z \in A \cap \boldsymbol{z}} \Psi_z.$$

The maps Ψ_A can be thought of as defining an action of

$$\Lambda^* H_1(L/(\boldsymbol{w}, \sigma); \mathbb{Z})$$

on $\mathit{CFL}^\infty_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s})$, where $L/(\boldsymbol{w}, \sigma)$ denotes the space obtained by identifying two $\boldsymbol{w}$–basepoints if they share the same color. This formalism is intriguing, but we will only have use for maps Ψ_A for arcs A between $\boldsymbol{w}$–basepoints of the same color.

Given an arc A between two $\boldsymbol{w}$–basepoints, we define an *endpoint* of A to be a basepoint w such that the sets $\overline{K \setminus A}$ and $\overline{A}$ both contain w (so A has no endpoints if $A = K$).

We now proceed to prove some basic properties of the maps Ψ_A, all of which are recastings of previous lemmas proven about the maps Ψ_z.

Lemma 10.2 *We have*

$$S^{\pm}_{w,z} \Psi_A + \Psi_A S^{\pm}_{w,z} \simeq 0$$

as long as w is not an endpoint of A.

Proof This follows immediately from Lemmas 9.4 and 9.5. □

Lemma 10.3 *If A and A' are two arcs between $\boldsymbol{w}$–basepoints, then*

$$\Psi_A \Psi_{A'} + \Psi_{A'} \Psi_A \simeq 0.$$

Proof This follows from Lemma 9.2. □

Lemma 10.4 *If A is an arc on L, then we have*

$$\Psi_A^2 \simeq 0$$

as filtered equivariant maps.

Proof Simply write $\Psi_A = \sum_{z \in A \cap \boldsymbol{z}} \Psi_z$, multiply out Ψ_A^2, then apply Lemmas 9.2 and 9.7. □

Lemma 10.5 *Suppose $A \subseteq K$ is an arc between $\boldsymbol{w}$–basepoints and let $c(A)$ denote the arc $\overline{K \setminus A}$. Then*

$$\Psi_K \Psi_A = \Psi_{c(A)} \Psi_A.$$

Proof Write $\Psi_K = \Psi_A + \Psi_{c(A)}$ and then use the previous lemma to compute that

$$\Psi_K\Psi_A = (\Psi_A + \Psi_{c(A)})\Psi_A = \Psi_A^2 + \Psi_{c(A)}\Psi_A = \Psi_{c(A)}\Psi_A. \qquad \square$$

Lemma 10.6 *If w is an endpoint of A then we have*

$$\Psi_A\Phi_w + \Phi_w\Psi_A \simeq \mathrm{id}\,.$$

If w is not an endpoint of A, then we have

$$\Psi_A\Phi_w + \Phi_w\Psi_A \simeq 0.$$

Proof The first claim follows from Lemma 9.1. The second claim follows from Lemma 10.2 since we can always write $\Phi_w = S_{w,z}^+ S_{w,z}^-$ for z the basepoint immediately preceding w on $\mathbb{L}$. $\square$

We can now prove Theorem B:

Proof of Theorem B Let $w_1, z_1, \dots w_n, z_n$ be the basepoints on K, in the reverse order that they appear on K according to the orientation of K. Let $w_1', z_1', \dots, w_n', z_n'$ be new basepoints on K in the interval between z_n and w_1. Let A_j be the arc on K from w_j to w_j', as in Figure 14.

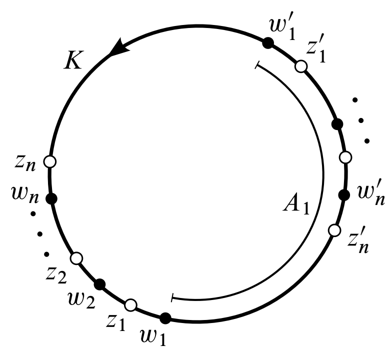


Figure 14: The basepoints $z_1, w_1, \dots, z_n, w_n$ and $z_1', w_1', \dots z_n', w_n'$, and the arcs A_i

Write

$$\boldsymbol{w} = \{w_1, \dots, w_n\}, \quad \boldsymbol{z} = \{z_1, \dots, z_n\}, \quad \boldsymbol{w}' = \{w_1', \dots, w_n'\}, \quad \boldsymbol{z}' = \{z_1', \dots, z_n'\}.$$

As usual, we write ς as a composition of two diffeomorphisms $\varsigma = \varsigma_2 \circ \varsigma_1$, where ς_1 moves the basepoints $\boldsymbol{w}$ and $\boldsymbol{z}$ to $\boldsymbol{w}'$ and $\boldsymbol{z}'$, respectively, and ς_2 moves the basepoints

$\boldsymbol{w}'$ and $\boldsymbol{z}'$ to $\boldsymbol{w}$ and $\boldsymbol{z}$, respectively. Let $c(A_i) = \overline{K \setminus A_i}$. By Lemma 9.6 we have

$$\prod_{j=1}^{n} (S^-_{w'_j,z'_j} \Psi_{A_j} S^+_{w'_j,z'_j}) \simeq \mathrm{id}. \tag{7}$$

Write $S^{\pm}_{\boldsymbol{w},\boldsymbol{z}}$ for $\prod_{j=1}^n S^{\pm}_{w_j,z_j}$, and similarly for $S^{\pm}_{\boldsymbol{w}',\boldsymbol{z}'}$. We compute as follows:

$$\begin{aligned}
\varsigma_* &= (\varsigma_2)_* \circ (\varsigma_1)_* \\
&= \Bigl(\prod_{j=1}^n S^-_{w'_j,z'_j} \Psi_{c(A_j)} S^+_{w_j,z_j}\Bigr)\Bigl(\prod_{j=1}^n S^-_{w_j,z_j} \Psi_{A_j} S^+_{w'_j,z'_j}\Bigr) && \text{(Lemma 10.1)} \\
&= S^-_{\boldsymbol{w}',\boldsymbol{z}'}\Bigl(\prod_{j=1}^n \Psi_{c(A_j)}\Bigr) S^+_{\boldsymbol{w},\boldsymbol{z}} S^-_{\boldsymbol{w},\boldsymbol{z}}\Bigl(\prod_{j=1}^n \Psi_{A_j}\Bigr) S^+_{\boldsymbol{w}',\boldsymbol{z}'} && \text{(Lemmas 10.2, 8.1)} \\
&= S^-_{\boldsymbol{w}',\boldsymbol{z}'}\Bigl(\prod_{j=1}^n \Psi_{c(A_j)}\Bigr)\Bigl(\prod_{j=1}^n \Phi_{w_j}\Bigr)\Bigl(\prod_{j=1}^n \Psi_{A_j}\Bigr) S^+_{\boldsymbol{w}',\boldsymbol{z}'} && \text{(Lemmas 9.3, 8.1)} \\
&= S^-_{\boldsymbol{w}',\boldsymbol{z}'} \sum_{s\in\{0,1\}^n}\Bigl(\prod_{j=1}^n \Phi^{s_j}_{w_j}\Bigr)\Bigl(\prod_{j=1}^n \Psi^{s_j}_{c(A_j)}\Bigr)\Bigl(\prod_{j=1}^n \Psi_{A_j}\Bigr) S^+_{\boldsymbol{w}',\boldsymbol{z}'} && \text{(Lemma 10.6)} \\
&= S^-_{\boldsymbol{w}',\boldsymbol{z}'} \sum_{s\in\{0,1\}^n}\Bigl(\prod_{j=1}^n \Phi^{s_j}_{w_j}\Bigr)\Bigl(\prod_{j=1}^n \Psi^{s_j}_{K}\Bigr)\Bigl(\prod_{j=1}^n \Psi_{A_j}\Bigr) S^+_{\boldsymbol{w}',\boldsymbol{z}'} && \text{(Lemmas 10.3, 10.5)} \\
&= \sum_{s\in\{0,1\}^n}\Bigl(\prod_{j=1}^n \Phi^{s_j}_{w_j}\Bigr)\Bigl(\prod_{j=1}^n \Psi^{s_j}_{K}\Bigr) S^-_{\boldsymbol{w}',\boldsymbol{z}'}\Bigl(\prod_{j=1}^n \Psi_{A_j}\Bigr) S^+_{\boldsymbol{w}',\boldsymbol{z}'} && \text{(Lemmas 10.2, 9.4)} \\
&= \sum_{s\in\{0,1\}^n}\Bigl(\prod_{j=1}^n \Phi^{s_j}_{w_j}\Bigr)\Bigl(\prod_{j=1}^n \Psi^{s_j}_{K}\Bigr) && \text{(Equation (7)).}
\end{aligned}$$

By Lemma 10.4, if $s \in \{0,1\}^n$ then

$$\Bigl(\prod_{j=1}^n \Psi^{s_j}_K\Bigr) \simeq 0$$

if s_j is nonzero for more than one j. Hence the above sum reduces to

$$\varsigma_* \simeq \mathrm{id} + \sum_{j=1}^n \Phi_{w_i} \Psi_K = \mathrm{id} + \Phi_K \Psi_K,$$

completing the proof. □

10.3 The map associated to a partial twist around a link component

In this section, we perform an additional basepoint moving map computation and prove Theorem D. Suppose that $\mathbb{L}$ is a multibased link and K is a component with basepoints z_1, w_1, z_2, $w_2, \dots, z_n$ and w_n, appearing in that order. Let τ be the diffeomorphism induced by twisting $(1/n)^{\text{th}}$ of the way around K, sending z_i to z_{i+1} and w_i to w_{i+1} (with indices taken modulo n). In the case that we pick a coloring $(\sigma, \mathfrak{P})$ where all of the $\boldsymbol{w}$–basepoints have the same color, the map τ induces a map on the complex

$$\mathrm{CFL}^{\infty}_{UV}(Y, \mathbb{L}, \sigma, \mathfrak{P}, \mathfrak{s}).$$

We have the following:

Theorem D *Suppose that $\mathbb{L}$ is an embedded multibased link in Y and K is a component of $\mathbb{L}$ with basepoints $z_1, w_1, \dots, z_n$ and w_n, appearing in that order. Assume that $n > 1$. If τ denotes the $(1/n)^{th}$–twist map, then for a coloring where all $\boldsymbol{w}$–basepoints on K have the same color, we have*

$$\tau_* \simeq (\Psi_{z_1}\Phi_{w_1}\Psi_{z_2}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{z_n}\Phi_{w_n}) + (\Phi_{w_1}\Psi_{z_2}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{z_n}).$$

Proof Let A_i be the arc from w_i to w_{i+1}, respecting the orientation of K. Let w' and z' be new basepoints in the region between z_n and w_1. Let A' denote the arc from w_n to w' and let A'' denote the arc from w' to w_1. This is illustrated in Figure 15.

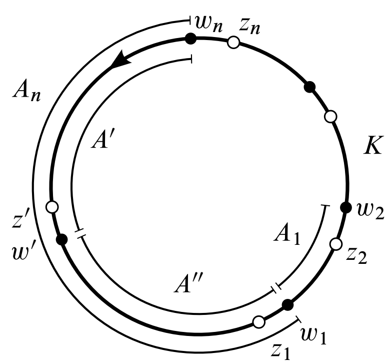


Figure 15: The basepoints $z_1, w_1, \dots, z_n, w_n, z', w'$, and the arcs A_i, A' and A'' from Theorem D

Using Lemma 10.1 repeatedly, we have that

$$\tau_* \simeq (S^-_{w'}\Psi_{A''}S^+_{w_1})(S^-_{w_1}\Psi_{A_1}S^+_{w_2})\cdots(S^-_{w_{n-1}}\Psi_{A_{n-1}}S^+_{w_n})(S^-_{w_n}\Psi_{A'}S^+_{w'}).$$

Using this, we perform the following computation:

$$
\begin{aligned}
\tau_* &\simeq S^-_{w'}\Psi_{A''}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}\Phi_{w_n}\Psi_{A'}S^+_{w'} && \text{(Lemma 9.3)}\\
&\simeq S^-_{w'}\Psi_{A''}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}(\Psi_{A'}\Phi_{w_n}+1)S^+_{w'} && \text{(Lemma 10.6)}\\
&\simeq S^-_{w'}\Psi_{A''}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}\Psi_{A'}\Phi_{w_n}S^+_{w'}\\
&\qquad + S^-_{w'}\Psi_{A''}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}S^+_{w'}\\
&\simeq S^-_{w'}(\Psi_{A''}\Psi_{A'})\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}\Phi_{w_n}S^+_{w'}\\
&\qquad + S^-_{w'}\Psi_{A''}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}S^+_{w'} && \text{(Lemmas 10.6, 10.3)}\\
&\simeq S^-_{w'}(\Psi_{A''}\Psi_{A_n})\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}\Phi_{w_n}S^+_{w'}\\
&\qquad + S^-_{w'}\Psi_{A''}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}S^+_{w'} && \text{(Lemma 10.4)}\\
&\simeq (S^-_{w'}\Psi_{A''}S^+_{w'})\Psi_{A_n}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}\Phi_{w_n}\\
&\qquad + (S^-_{w'}\Psi_{A''}S^+_{w'})\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}} && \text{(Lemma 10.2)}\\
&\simeq \Psi_{A_n}\Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}}\Phi_{w_n}\\
&\qquad + \Phi_{w_1}\Psi_{A_1}\Phi_{w_2}\cdots\Phi_{w_{n-1}}\Psi_{A_{n-1}} && \text{(Lemma 9.6),}
\end{aligned}
$$

completing the proof since $\Psi_{A_i} = \Psi_{z_{i+1}}$ on the complex $\mathit{CFL}^\infty_{UV}(Y,\mathbb{L},\sigma,\mathfrak{P},\mathfrak{s})$, by definition. □

References

[1] **K Hendricks**, **C Manolescu**, *Involutive Heegaard Floer homology*, preprint (2015) arXiv

[2] **A Juhász**, *Cobordisms of sutured manifolds and the functoriality of link Floer homology*, Adv. Math. 299 (2016) 940–1038 MR

[3] **A Juhász**, **M Marengon**, *Concordance maps in knot Floer homology*, Geom. Topol. 20 (2016) 3623–3673 MR

[4] **A Juhász**, **D P Thurston**, *Naturality and mapping class groups in Heegaard Floer homology*, preprint (2012) arXiv

[5] **R Lipshitz**, *A cylindrical reformulation of Heegaard Floer homology*, Geom. Topol. 10 (2006) 955–1097 MR

[6] **C Manolescu**, **P Ozsváth**, *Heegaard Floer homology and integer surgeries on links*, preprint (2011) arXiv

[7] **Y-G Oh**, *Fredholm theory of holomorphic discs under the perturbation of boundary conditions*, Math. Z. 222 (1996) 505–520 MR

[8] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and knot invariants*, Adv. Math. 186 (2004) 58–116 MR

[9] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and topological invariants for closed three-manifolds*, Ann. of Math. 159 (2004) 1027–1158 MR

[10] **P Ozsváth**, **Z Szabó**, *Holomorphic disks, link invariants and the multi-variable Alexander polynomial*, Algebr. Geom. Topol. 8 (2008) 615–692 MR

[11] **J A Rasmussen**, *Floer homology and knot complements*, PhD thesis, Harvard (2003) MR Available at `http://search.proquest.com/docview/305332635`

[12] **S Sarkar**, *Maslov index formulas for Whitney n–gons*, J. Symplectic Geom. 9 (2011) 251–270 MR

[13] **S Sarkar**, *Moving basepoints and the induced automorphisms of link Floer homology*, Algebr. Geom. Topol. 15 (2015) 2479–2515 MR

[14] **I Zemke**, *Graph cobordisms and Heegaard Floer homology*, preprint (2015) arXiv

Department of Mathematics, University of California, Los Angeles
520 Portola Plaza, Los Angeles, CA 90025, United States

`ianzemke@math.ucla.edu`

Received: 3 June 2016 Revised: 10 October 2016

Algebraic & Geometric Topology 17 (2017) 3519–3545

Cosimplicial groups and spaces of homomorphisms

BERNARDO VILLARREAL

Let G be a real linear algebraic group and L a finitely generated cosimplicial group. We prove that the space of homomorphisms $\mathrm{Hom}(L_n, G)$ has a homotopy stable decomposition for each $n \geq 1$. When G is a compact Lie group, we show that the decomposition is G–equivariant with respect to the induced action of conjugation by elements of G. In particular, under these hypotheses on G, we obtain stable decompositions for $\mathrm{Hom}(F_n/\Gamma_n^q, G)$ and $\mathrm{Rep}(F_n/\Gamma_n^q, G)$, respectively, where F_n/Γ_n^q are the finitely generated free nilpotent groups of nilpotency class $q-1$.

The spaces $\mathrm{Hom}(L_n, G)$ assemble into a simplicial space $\mathrm{Hom}(L, G)$. When $G = U$ we show that its geometric realization $B(L, U)$, has a nonunital E_∞–ring space structure whenever $\mathrm{Hom}(L_0, U(m))$ is path connected for all $m \geq 1$.

22E15; 55U10, 20G05

1 Introduction

Let G be a topological group and Γ a finitely generated group. The set of homomorphisms $\mathrm{Hom}(\Gamma, G)$ can be identified with the ordered tuples $(\rho(a_1), \dots, \rho(a_r))$ in G^r, where $\rho\colon \Gamma \to G$ is a homomorphism and $a_1, \dots, a_r$ is a generating set for Γ. Computing the homotopy type of $\mathrm{Hom}(\Gamma, G)$ has proven to be rather complicated. Nevertheless, there has been recognition of the stable homotopy type in several cases. When $G \subset \mathrm{GL}_n(\mathbb{C})$ is a closed subgroup, A Adem and F Cohen [2] gave a homotopy stable decomposition for $\mathrm{Hom}(\mathbb{Z}^n, G)$ as wedges of the quotient spaces $\mathrm{Hom}(\mathbb{Z}^k, G)/S_1(\mathbb{Z}^k, G)$ with $1 \leq k \leq n$. Here $S_1(\mathbb{Z}^k, G)$ stands for the k–tuples with at least one entry equal to the identity matrix I in G. For an arbitrary finitely generated abelian group π, Adem and J M Gómez [5] gave a similar stable decomposition for $\mathrm{Hom}(\pi, G)$, but in this case, G is a finite product of the compact Lie groups $\mathrm{SU}(r)$, $\mathrm{Sp}(k)$ and $U(m)$. We show that the homotopy stable decomposition in [2] (in general, a version of it) still holds if we replace the family $\{\mathbb{Z}^n\}_{n\geq 1}$ with a broader object, namely, a finitely generated cosimplicial group L, that is, for every $n \geq 0$, we have a finitely generated group L_n with its coface and codegeneracy homomorphisms. A specific example of L comes from the finitely generated free nilpotent groups F_n/Γ_n^q. Here F_n denotes the free group on n–generators and Γ_n^q

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3519

is the q^{th} stage of its descending central series. We work under the assumption that G is a real linear algebraic group, and with it we can prove that the inclusions $S_{k+1}(F_n/\Gamma_n^q, G) \hookrightarrow S_k(F_n/\Gamma_n^q, G)$ are closed cofibrations for every $0 \le k \le n$, where $S_k(F_n/\Gamma_n^q, G)$ denotes the subspace of $\mathrm{Hom}(F_n/\Gamma_n^q, G)$ with at least k entries equal to the identity element in G. This condition, as stated by Adem, Cohen and E Torres-Giese [4, page 102], says that real linear algebraic groups have *cofibrantly filtered elements*, and they show that this implies $\mathrm{Hom}(F_n/\Gamma_n^q, G)$ splits after one suspension as wedges of $\mathrm{Hom}(F_k/\Gamma_k^q, G)/S_1(F_k/\Gamma_k^q, G)$, where $0 \le k \le n$. This decomposition was known before for some compact and connected Lie groups G (see for example Cohen and M Stafa [11, Remark 1, page 387 and Theorem 2.13, page 388]).

In a more general setting, let Δ stand for the category whose objects are the natural numbers and morphisms are order-preserving maps. If $L\colon \Delta \to \mathbf{Grp}$ is a finitely generated cosimplicial group and G is a topological group, we get the simplicial space $\mathrm{Hom}(L, G)\colon \Delta^{\mathrm{op}} \to \mathbf{Top}$, where $\mathrm{Hom}(L, G)_n := \mathrm{Hom}(L_n, G)$. We give a homotopy stable decomposition for the n–simplices of $\mathrm{Hom}(L, G)$ as follows. Let X be a simplicial space. Define $S^t(X_n)$ as the subspace of X_n in which any element is in the image of the composition of at least t degeneracy maps. We say X is simplicially NDR when all pairs $(S^{t-1}(X_n), S^t(X_n))$ are neighborhood deformation retracts. It was proven by Adem, A Bahri, M Bendersky, Cohen, and S Gitler [1] that when X is simplicially NDR, each X_n is homotopy stable equivalent to wedges of $S^t(X_n)/S^{t+1}(X_n)$ with $0 \le t \le n$. When G is a real linear algebraic group, the simplicial space $X = \mathrm{Hom}(L, G)$ is simplicially NDR. We prove this by showing that the subspaces $S^t(X_n)$ are real affine subvarieties of X_n for all $0 \le t \le n$ and therefore can be simultaneously triangulated. Define $S_t(L_n, G) := S^t(\mathrm{Hom}(L_n, G))$.

Theorem 1.1 *Let G be a real linear algebraic group, and L a finitely generated cosimplicial group. For each n, there are natural homotopy equivalences*

$$\Theta(n)\colon \Sigma\,\mathrm{Hom}(L_n, G) \simeq \bigvee_{0 \le k \le n} \Sigma(S_k(L_n, G)/S_{k+1}(L_n, G)).$$

The free groups F_n assemble into a cosimplicial group, which we denote by F. In this case $\mathrm{Hom}(F, G)$ is NG, the nerve of G seen as a topological category with one object, which is also the underlying simplicial space of a model of the classifying space BG. For each n, we take quotients F_n/K_n by normal subgroups K_n that are compatible with coface and codegeneracy homomorphisms of F to get finitely generated cosimplicial groups, denoted by F/K. The induced simplicial spaces are more easily described since there is a simplicial inclusion $\mathrm{Hom}(F/K, G) \subset NG$. Fixing $q > 0$, the family of subgroups Γ_n^q arising from the descending central series of F_n is compatible with F, and induces the finitely generated cosimplicial group F/Γ^q. Adem, Cohen and Torres-Giese [4] conjectured that the closed subgroups of $\mathrm{GL}_n(\mathbb{C})$ have cofibrantly filtered elements and thus the homotopy stable decomposition holds. Applying Theorem 1.1 to F/Γ^q allows us to prove the following version of the conjecture.

Corollary 1.2 *If G is a Zariski closed subgroup of $\mathrm{GL}_n(\mathbb{C})$, then there are homotopy equivalences for the cosimplicial group F/Γ^q,*

$$\Sigma\,\mathrm{Hom}(F_n/\Gamma_n^q, G) \simeq \bigvee_{1\le k\le n}\Sigma\big(\bigvee^{\binom{n}{k}}\mathrm{Hom}(F_k/\Gamma_k^q, G)/S_1(F_k/\Gamma_k^q, G)\big)$$

for all n and q.

For any finitely generated cosimplicial group L, conjugation under elements of G gives $\mathrm{Hom}(L_n, G)$ a G–space structure. Moreover, if G is a real algebraic linear group, then it has a G–variety structure. The subspaces $S^t(\mathrm{Hom}(L_n, G))$ are subvarieties that are invariant under the action of G for all $0\le t\le n$. Using techniques from D H Park and D Y Suh [21], when G is a compact Lie group, we show that $\mathrm{Hom}(L_n, G)$ has a G–CW–complex structure, where each $S^t(\mathrm{Hom}(L_n, G))$ is a G–subcomplex. This allows us to prove the equivariant version of the previous theorem. Let $\mathrm{Rep}(L_n, G)$ and $\overline{S}_t(L_n, G)$ denote the orbit spaces of $\mathrm{Hom}(L_n, G)$ and $S^t(\mathrm{Hom}(L_n, G))$, respectively.

Theorem 1.3 *Let G be a compact Lie group. Then, for each n, $\Theta(n)$ in Theorem 1.1 is a G–equivariant homotopy equivalence, and in particular we get homotopy equivalences*

$$\Sigma\,\mathrm{Rep}(L_n, G) \simeq \bigvee_{1\le k\le n}\Sigma(\overline{S}_k(L_n, G)/\overline{S}_{k+1}(L_n, G)).$$

Applying this to the cosimplicial group F/Γ^q as in Corollary 1.2, we obtain

$$\Sigma\,\mathrm{Rep}(F_n/\Gamma_n^q, G) \simeq \bigvee_{1\le k\le n}\Sigma\big(\bigvee^{\binom{n}{k}}\mathrm{Rep}(F_k/\Gamma_k^q, G)/\overline{S}_1(F_k/\Gamma_k^q, G)\big).$$

In the second part of this paper we study the geometric realization of $\mathrm{Hom}(L, G)$ for a finitely generated cosimplicial group, which we denote by $B(L, G)$. We show that the set of 1–cocycles of L, denoted by $Z^1(L)$, is in one-to-one correspondence with cosimplicial morphisms $F\to L$. With this we show that any 1–cocycle of L defines a principal G–bundle over $B(L, G)$.

When $G = U = \mathrm{colim}_m\, U(m)$ we show that $B(L, U)$ has an $\mathbb{I}$–rig structure, that is, if $\mathbb{I}$ stands for the category of finite sets and injections, the functor $B(L, U(\_))\colon \mathbb{I}\to\mathbf{Top}$ is symmetric monoidal with respect to both symmetric monoidal structures on $\mathbb{I}$. Using the machinery developed in [7] by Adem, Gómez, J Lind and U Tillman, we prove:

Theorem 1.4 *Let L be a finitely generated cosimplicial group and suppose that the space $\mathrm{Hom}(L_0, U(m))$ is path connected for all $m\ge 1$. Then $B(L, U)$ is a nonunital E_∞–ring space.*

This theorem is also true if we replace U by SU, Sp, SO or O.

Acknowledgements I would like to thank A Adem for his supervision and advice throughout this work. O Antolín Camarena for his useful input with the details of this paper. I would also like to thank M Bergeron, F Cohen and J M Gómez for their comments on an earlier version. I was supported by a CONACyT fellowship.

2 Homotopy stable decompositions

2.1 Spaces of homomorphisms

Let G be a topological group and Γ a finitely generated group. Any homomorphism $\rho\colon \Gamma \to G$ is uniquely determined by $(\rho(\gamma_1), \dots, \rho(\gamma_n)) \in G^n$ when $\gamma_1, \dots, \gamma_n \in \Gamma$ is a set of generators. On the other hand, if we fix a presentation of Γ, then an n–tuple $(g_1, \dots, g_n) \in G^n$ will induce an element in $\mathrm{Hom}(\Gamma, G)$ whenever $\{g_i\}_{i=1}^n$ satisfy the relations in the presentation of Γ. Thus, there is a one-to-one correspondence between the subset of such n–tuples in G^n and $\mathrm{Hom}(\Gamma, G)$. Topologize $\mathrm{Hom}(\Gamma, G)$ with the subspace topology on G^n.

Lemma 2.1 *Let $\varphi\colon \Gamma \to \Gamma'$ be a homomorphism of finitely generated groups. If G is a topological group, then $\varphi^*\colon \mathrm{Hom}(\Gamma', G) \to \mathrm{Hom}(\Gamma, G)$ is continuous.*

Proof Suppose $\Gamma = \langle a_1, \dots, a_r \mid R\rangle$ and $\Gamma' = \langle b_1, \dots, b_m \mid R'\rangle$. Recall that the induced map $\varphi^*\colon \mathrm{Hom}(\Gamma', G) \to \mathrm{Hom}(\Gamma, G)$ is given by

$$(\rho(b_1), \dots, \rho(b_m)) \mapsto \big(\rho(\varphi(a_1)), \dots, \rho(\varphi(a_r))\big)$$

for $\rho\colon \Gamma' \to G$. For any i, $\varphi(a_i) = b_{i_1}^{n_{i_1}} \cdots b_{i_{q_i}}^{n_{i_{q_i}}}$. By fixing one presentation for each $\varphi(a_i)$ we get that φ^* is given by

$$\begin{aligned}(\rho(b_1), \dots, \rho(b_m)) &\mapsto (\rho(b_{1_1}^{n_{1_1}} \cdots b_{1_{q_1}}^{n_{1_{q_1}}}), \dots, \rho(b_{r_1}^{n_{r_1}} \cdots b_{r_{q_r}}^{n_{r_{q_r}}})) \\ &= (\rho(b_{1_1})^{n_{1_1}} \cdots \rho(b_{1_{q_1}})^{n_{1_{q_1}}}, \dots, \rho(b_{r_1})^{n_{r_1}} \cdots \rho(b_{r_{q_r}})^{n_{r_{q_r}}}).\end{aligned}$$

Therefore φ^* is the restriction of the map $G^m \to G^r$ given by

$$(g_1, \dots, g_m) \mapsto (g_{1_1}^{n_{1_1}} \cdots g_{1_{q_1}}^{n_{1_{q_1}}}, \dots, g_{r_1}^{n_{r_1}} \cdots g_{r_{q_r}}^{n_{r_{q_r}}}),$$

which is continuous. □

In particular, this lemma tells us that given any two presentations of Γ, we get an isomorphism $\varphi\colon \Gamma \to \Gamma$ and hence a homeomorphism φ^* between the induced spaces of homomorphisms. Therefore the topology on the space of homomorphisms does not depend on the choice of presentations.

Recall that an affine variety is the zero locus in k^n of a family of polynomials on n variables over a field k. Throughout this paper we will focus only on $k=\mathbb{R}$. An affine variety that has a group structure with group operations given by polynomial maps, ie maps $f=(f_1,\dots,f_n)$, where each f_i is a polynomial, is called a linear algebraic group. For example, consider any matrix group. It is easy to check that matrix multiplication is in fact a polynomial map. For the inverse operation of matrices, it is easier to think of matrix groups as subgroups of $\mathrm{SL}(n,\mathbb{R})$. Any matrix A in $\mathrm{SL}(n,\mathbb{R})$ satisfies $A^{-1}=C^t$, the transpose of the cofactor matrix C of A. Since the cofactor matrix is described only in terms of minors of A, the map $A\mapsto C^t$ is a polynomial map. In fact, this is the general example, since it can be shown that any linear algebraic group is isomorphic to a group of matrices (see for example [15, page 63]).

Lemma 2.2 *Let G be a linear algebraic group; then for any finitely generated group Γ, $\mathrm{Hom}(\Gamma,G)$ is an affine variety. Moreover, if φ is a homomorphism of finitely generated groups, then φ^* is a polynomial map.*

Proof Suppose Γ is generated by $\gamma_1,\gamma_2,\dots,\gamma_r$ and has a presentation $\{p_\alpha\}_{\alpha\in\Lambda}$. Each p_α is of the form $\gamma_{k_1}^{n_1}\cdots\gamma_{k_q}^{n_q}=e$ with $n_j\in\mathbb{Z}$ and $\gamma_{k_l}\in\{\gamma_1,\dots,\gamma_r\}$ for all $1\le j\le q$. For any homomorphism $\rho\colon\Gamma\to G$ and any such relation p_α, we have

$$\rho(p_\alpha)=\rho(\gamma_{i_1}^{n_1}\cdots\gamma_{i_q}^{n_q})=\rho(\gamma_{i_1})^{n_1}\cdots\rho(\gamma_{i_q})^{n_q}=I,$$

the identity matrix in G. Since products and inverses in G are given in terms of polynomials, this sets up a family of polynomial relations $\{y_{\alpha,i,j}\}_{\alpha,i,j}$, where each $y_{\alpha,i,j}$ is induced by $\rho(p_\alpha)_{i,j}=\delta_{ij}$, the (i,j)–entry of the matrix equality $\rho(p_\alpha)=I$. These relations do not depend on ρ, only on p_α, in the sense that any r–tuple $(g_1,\dots,g_r)\in G$ satisfying $\{y_{\alpha,i,j}\}_{\alpha,i,j}$, ie

$$(g_{i_1}^{n_1}\cdots g_{i_q}^{n_q})_{i,j}=\delta_{ij}$$

for all $\alpha\in\Lambda$ and $1\le i,j\le n$, is an element of $\mathrm{Hom}(\Gamma,G)$. Adding the polynomial relations $\{y_{\alpha,i,j}\}_{\alpha\in\Lambda}$ to the ones describing G^r defines $\mathrm{Hom}(\Gamma,G)$ as an affine variety.

For the second part, recall from the proof of Lemma 2.1, that φ^* is defined in terms of products and inverses of matrices and thus is a polynomial map. □

Similarly, this lemma tells us that the affine variety structure on $\mathrm{Hom}(\Gamma,G)$ does not depend on the presentation of Γ. Indeed, any isomorphism of groups will induce an isomorphism of affine varieties.

2.2 Triangulation of semialgebraic sets

Definition 2.3 A real semialgebraic set is a finite union of subsets of the form

$$\{x \in \mathbb{R}^n \mid f_i(x) > 0 \text{ and } g_j(x) = 0 \text{ for all } i \text{ and } j\},$$

where $f_i(x)$ and $g_j(x)$ are a finite number of polynomials with real coefficients.

Using Hilbert's basis theorem, all affine varieties over $\mathbb{R}$ are real semialgebraic sets. Indeed, the zero locus ideal of an affine variety will be finitely generated and thus the affine variety can be carved out by finitely many polynomials.

What makes semialgebraic sets more interesting is that images of semialgebraic sets in $\mathbb{R}^n$ under a polynomial map $\mathbb{R}^n \to \mathbb{R}^m$ are semialgebraic sets in $\mathbb{R}^m$ (see [14, page 167]), as opposed to affine varieties and regular maps.

Let M and N be semialgebraic subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. A continuous map $f\colon M \to N$ is said to be semialgebraic if its graph is a semialgebraic set in $\mathbb{R}^m \times \mathbb{R}^n$. The next result is proven in [14, page 170].

Proposition 2.4 *Given a finite system of bounded semialgebraic sets M_i in $\mathbb{R}^n$, there is a simplicial complex K in $\mathbb{R}^n$ and a semialgebraic homeomorphism $k\colon |K| \to \bigcup \overline{M}_i$, where each M_i is a finite union of sets $k(\operatorname{int}|\sigma|)$ with $\sigma \in K$.*

Remark 2.5 Proposition 2.4 can be stated without the boundedness condition and the details can be found in [21, Theorem 2.12], where they add the hypothesis that $\bigcup M_i$ is closed in $\mathbb{R}^n$.

In the next sections, we will be using this last result in its full power, but a first application is that any affine variety Z can be triangulated, that is, there exists a simplicial complex K and a homeomorphism $|K| \cong Z$. With Lemma 2.2 and Proposition 2.4 we prove the following.

Corollary 2.6 *Let Γ be a finitely generated group and G a real linear algebraic group. Then $\operatorname{Hom}(\Gamma, G)$ is a triangulated space.*

2.3 Simplicial spaces and homotopy stable decompositions

Let Δ be the category of finite sets $[n] = \{0, 1, \dots, n\}$ with morphisms order preserving maps $f\colon [n] \to [m]$. It can be shown that all morphisms in this category are generated by composition of maps, denoted by $\boldsymbol{d}^i\colon [n-1] \to [n]$ and $\boldsymbol{s}^i\colon [n+1] \to [n]$ for $0 \le i \le n$. These maps are determined by the relations

$$\begin{aligned}
\boldsymbol{d}^j\boldsymbol{d}^i &= \boldsymbol{d}^i\boldsymbol{d}^{j-1} && \text{if } i<j,\\
\boldsymbol{s}^j\boldsymbol{s}^i &= \boldsymbol{s}^{i-1}\boldsymbol{s}^j && \text{if } i>j,\\
\boldsymbol{s}^j\boldsymbol{d}^i &= \begin{cases}\boldsymbol{d}^i\boldsymbol{s}^{j-1} & \text{if } i<j,\\ \mathrm{Id} & \text{if } i=j \text{ or } i=j+1,\\ \boldsymbol{d}^{i-1}\boldsymbol{s}^j & \text{if } i>j+1,\end{cases}
\end{aligned}$$

which are called cosimplicial identities. For any category $\boldsymbol{C}$, let $\boldsymbol{C}^{\mathrm{op}}$ denote its opposite category. A functor

$$X\colon \Delta^{\mathrm{op}} \to \mathbf{Top}$$

is called a simplicial space. Here **Top** stands for k–spaces, ie topological spaces where each compactly closed subset is closed. We write $X_n := X([n])$ and the maps $d_i = X(\boldsymbol{d}^i)$ and $s_i = X(\boldsymbol{s}^i)$ are called face and degeneracy maps, respectively.

Fix n. Define $S^0(X_n) = X_n$ and, for $0 < t \le n$,

$$S^t(X_n) = \bigcup_{J_{n,t}} s_{i_1} \circ \cdots \circ s_{i_t}(X_{n-t}),$$

where $s_{i_j}\colon X_{n-j} \to X_{n-j+1}$ is a degeneracy map, $1 \le i_1 < \cdots < i_t \le n$ is a sequence of t numbers between 1 and n, and $J_{n,t}$ stands for all possible sequences. This defines a decreasing filtration of X_n,

$$S^n(X_n) \subset S^{n-1}(X_n) \subset \cdots \subset S^0(X_n) = X_n.$$

For each n there is a homotopy decomposition of ΣX_n in terms of the quotient spaces $S^k(X_n)/S^{k+1}(X_n)$ with $k \le n$. To do this we need the following.

Let $A \subset Z$ be topological spaces. Recall that (Z, A) is an NDR pair if there exist continuous functions

$$h\colon Z \times [0,1] \to Z, \quad u\colon Z \to [0,1]$$

such that the following conditions are satisfied:

1. $A = u^{-1}(0)$.
2. $h(z,0) = z$ for all $z \in Z$.
3. $h(a,t) = a$ for all $a \in A$ and all $t \in [0,1]$.
4. $h(z,1) \in A$ for all $z \in u^{-1}([0,1))$.

Examples of NDR pairs are pairs consisting of CW–complexes and subcomplexes. Indeed, if Z is a CW–complex and $A \subset Z$ a subcomplex, then the inclusion $A \hookrightarrow Z$ is a cofibration, which is equivalent to a retraction $X \times I$ to $A \times I \cup X \times \{0\}$ relative to $A \times \{0\}$.

When X is a simplicial space, we call X *simplicially* NDR if $(S^{t-1}(X_n), S^t(X_n))$ is an NDR pair for every n and $t \geq 1$. The following result can be found in [1, Theorem 1.6].

Proposition 2.7 *Let X be a simplicial space, and suppose X is simplicially* NDR. *Then for every $n \geq 0$ there is a natural homotopy equivalence*

$$\Theta(n)\colon \Sigma X_n \simeq \bigvee_{0\leq k\leq n} \Sigma(S^k(X_n)/S^{k+1}(X_n)).$$

For each n, the map $\Theta(n)$ is natural with respect to morphisms of simplicial spaces, that is, natural transformations $X \to Y$.

2.4 Cosimplicial groups, 1–cocycles and $\mathrm{Hom}(L, G)$

Definition 2.8 Let **Grp** denote the category of groups. A functor $L\colon \Delta \to \mathbf{Grp}$ is called a *cosimplicial group*. The homomorphisms $d^i = L(\boldsymbol{d}^i)$ and $s^i = L(\boldsymbol{s}^i)$ are called coface and codegeneracy homomorphisms, respectively. We say that L is a *finitely generated cosimplicial group* if each L_n is finitely generated.

There are two canonical finitely generated cosimplicial groups that arise from finitely generated free groups.

Definition 2.9 Define $F\colon \Delta \to \mathbf{Grp}$ as follows: set $F_0 = \{e\}$ and for $n \geq 1$ let $F_n = \langle a_1, \dots, a_n\rangle$, the free group on n generators. The coface homomorphisms $d^i\colon F_{n-1} \to F_n$ are given on the generators by

$$\begin{aligned} d^0(a_j) &= a_{j+1},\\ d^i(a_j) &= \begin{cases} a_j & \text{if } j < i,\\ a_j a_{j+1} & \text{if } j = i,\\ a_{j+1} & \text{if } j > i,\end{cases} \qquad \text{for } 1 \leq i \leq n-1,\\ d^n(a_j) &= a_j; \end{aligned}$$

and the codegeneracy homomorphisms $s^i\colon F_{n+1} \to F_n$ by

$$s^i(a_j) = \begin{cases} a_j & \text{if } j \leq i,\\ e & \text{if } j = i+1,\\ a_{j-1} & \text{if } j > i+1,\end{cases}$$

for $0 \leq i \leq n$.

Definition 2.10 Define $\overline{F}\colon \Delta \to \mathbf{Grp}$ as $\overline{F}_n := \langle a_0, \dots, a_n\rangle$ for any $n \geq 0$; coface and codegeneracy homomorphisms $\overline{d}^i\colon \overline{F}_{n-1} \to \overline{F}_n$ and $\overline{s}^i\colon \overline{F}_{n+1} \to \overline{F}_n$, respectively, are given on the generators by

$$\overline{d}^i(a_j) = \begin{cases} a_j & \text{if } j < i, \\ a_{j+1} & \text{if } j \geq i, \end{cases} \qquad \text{and} \qquad \overline{s}^i(a_j) = \begin{cases} a_j & \text{if } j \leq i, \\ a_{j-1} & \text{if } j > i, \end{cases}$$

for all $0 \leq i \leq n$.

Definition 2.11 We will say that a family of normal subgroups $K_n \subset F_n$ is *compatible* with F if $d^i(K_{n-1}) \subset K_n$ and $s^i(K_{n+1}) \subset K_n$ for all n and all i. Similarly we define *compatible* families of $\overline{F}$.

Given $\{K_n\}_{n\geq 0}$ a compatible family with F, we get induced homomorphisms

$$\begin{array}{ccc} F_{n-1} & \xrightarrow{d^i} & F_n \\ \downarrow & & \downarrow \\ F_{n-1}/K_{n-1} & \xrightarrow{d^i} & F_n/K_n \end{array} \qquad \begin{array}{ccc} F_{n+1} & \xrightarrow{s^i} & F_n \\ \downarrow & & \downarrow \\ F_{n+1}/K_{n+1} & \xrightarrow{s^i} & F_n/K_n \end{array}$$

Define

$$F/K\colon \Delta \to \mathbf{Grp}$$

as $(F/K)_n = F_n/K_n$ with coface and codegeneracy maps the quotient homomorphisms d^i and s^i, respectively. This way F/K is a finitely generated cosimplicial group. Similarly, with a compatible family $\{\overline{K}_n\}_{n\geq 0}$ of $\overline{F}$, we can define $\overline{F}/\overline{K}\colon \Delta \to \mathbf{Grp}$.

Example 2.12 We describe two families of finitely generated cosimplicial groups that can be constructed using F/K and $\overline{F}/\overline{K}$ through the commutator subgroup.

- Let A be a group, define inductively $\Gamma^1(A) = A$ and $\Gamma^{q+1}(A) = [\Gamma^q(A), A]$ for $q > 1$. The descending central series of A is

$$\Gamma^q(A) \trianglelefteq \cdots \trianglelefteq \Gamma^2(A) \trianglelefteq \Gamma^1(A) = A.$$

Given a homomorphism of groups $\phi\colon A \to B$, we have $\phi[a, a'] = [\phi(a), \phi(a')]$ for all $a, a' \in A$, so that

$$\phi(\Gamma^q(A)) \subset \Gamma^q(B).$$

Taking $A = F_n$, and writing $\Gamma^q_n := \Gamma^q(F_n)$, we have that the family of normal subgroups $\{\Gamma^q_n\}_{n\geq 0}$ is compatible with d_i and s_i. Thus we can define F/Γ^q as $(F/\Gamma^q)_n = F_n/\Gamma^q_n$ for all q and $1 \leq i \leq n$. In particular, for $q = 2$, we obtain $F_n/\Gamma^2_n = \mathbb{Z}^n$ for all $n \geq 0$.

- Another example using the commutator is the derived series of a group A,

$$A^{(q)} \trianglelefteq \cdots \trianglelefteq A^{(1)} \trianglelefteq A^{(0)} = A,$$

where $A^{(i+1)} = [A^{(i)}, A^{(i)}]$. Again, $\phi(A^{(q)}) \subset B^{(q)}$ for any homomorphism $\phi\colon A \to B$. Thus $F/F^{(q)}$, where $(F/F^{(q)})_n = F_n/F_n^{(q)}$ defines a finitely generated cosimplicial group.

Similarly, $F_{n+1}^{(q)}, \Gamma_{n+1}^q \subset \overline{F}_n$ define compatible families of $\overline{F}$ and we obtain the finitely generated cosimplicial groups $\overline{F}/\Gamma_{*+1}^q$ and $\overline{F}/F_{*+1}^{(q)}$.

Example 2.13 Here is one example of a cosimplicial group that does not come from a compatible family. Let $L_0 = \Sigma_2 = \langle \tau \rangle$ and $L_1 = \Sigma_3 = \langle \sigma_1, \sigma_2 \rangle$ and define coface homomorphisms

$$L_0 \xrightarrow{d^i} L_1 \quad \text{for } i = 0, 1$$

by $d^0(\tau) = \sigma_2$ and $d^1(\tau) = \sigma_1$. The codegeneracy homomorphism $s_0\colon L_1 \to L_0$ is given by $s^0(\sigma_1) = s^0(\sigma_2) = \tau$. This defines a 1–truncated cosimplicial group, which we denote by $\Sigma_{2,3}$, that is, a functor $\Sigma_{2,3}\colon \Delta_{\leq 1} \to \mathbf{Grp}$. Here $\Delta_{\leq 1}$ stands for the full subcategory of Δ with objects [0] and [1]. We can extend $\Sigma_{2,3}$ to Δ by using its left Kan extension.

For our purposes we describe the second stage of this extension: We have that

$$L_2 = \langle a, b, c \mid a^2 = b^2 = c^2,\ aba = bab,\ aca = cac,\ bcb = cbc \rangle,$$

coface homomorphisms

$$L_1 \xrightarrow{d^i} L_2 \quad \text{for } i = 0, 1, 2$$

are given by

$$\begin{aligned} d^0(\sigma_1) &= a, & d^1(\sigma_1) &= c, & d^2(\sigma_1) &= c, \\ d^0(\sigma_2) &= b, & d^1(\sigma_2) &= b, & d^2(\sigma_2) &= a, \end{aligned}$$

and codegeneracy homomorphisms

$$L_1 \xleftarrow{s^i} L_2 \quad \text{for } i = 0, 1$$

by

$$\begin{aligned} s^0(a) = s^0(c) &= \sigma_1, & s^1(a) = s^1(b) &= \sigma_2 \\ s^0(b) &= \sigma_2, & s^1(c) &= \sigma_1. \end{aligned}$$

Remark 2.14 The symmetric groups Σ_n can not be assembled all together as a cosimplicial group. This is because there are no surjective homomorphisms $\Sigma_n \to \Sigma_{n-1}$ for $n \geq 5$ to use as codegeneracy homomorphisms. Indeed, given a homomorphism

$\varphi\colon \Sigma_n \to \Sigma_{n-1}$, $\ker\varphi$ is a normal subgroup of Σ_n, that is, A_n or Σ_n. Thus the image of φ is either the identity element or a subgroup of order 2.

We describe another method of constructing new cosimplicial groups that arise from a given one. To do this, we recall a concept that was originally introduced in [10, page 284] to define cohomotopy groups (and pointed sets) for a cosimplicial group.

Definition 2.15 Let L be a cosimplicial group. The elements b in L_1 satisfying

$$d^2(b)d^0(b) = d^1(b) \tag{1}$$

are called 1–cocycles of L. The set of 1–cocycles is denoted by $Z^1(L)$.

If b is a 1–cocycle, then applying s^0 to (1), we obtain $s^0d^2(b) = e$ and, using the cosimplicial identities, $d^1s^0(b) = e$, which implies $b \in \ker s^0$. Define inductively $b_n \in L_n$ as $b_{n+1} = d^{n+1}(b_n)$, where $b_1 := b$. These elements will satisfy

$$d^2(b_n)d^0(b_n) = d^1(b_n), \tag{2}$$

$$b_n \in \ker s^0 \tag{3}$$

for all $n \geq 1$. Given a 1–cocycle b, we build a new cosimplicial group.

Construction of L^b Define $L^b\colon \Delta \to \mathbf{Grp}$ as follows. For each $n \geq 0$, $L_n^b := \overline{F}_0 * L_n$ with codegeneracy homomorphisms $s_b^i := \mathrm{Id} * s^i$, $i \geq 0$. The coface homomorphisms are $d_b^i := \mathrm{Id} * d^i$ for $i > 0$. To define d_b^0 consider the homomorphism $k_n\colon \overline{F}_0 \to \overline{F}_0 * L_n$ given by $k_n(a_0) = a_0b_n$ for all $n \geq 0$, then $d_b^0 := k_n * d^0$. There is a canonical inclusion

$$\iota_b\colon L \hookrightarrow L^b$$

induced by the inclusions $L_n \hookrightarrow \overline{F}_0 * L_n$.

Example 2.16 • When $b = e$, $L^e = \overline{F}_0 * L$, where $\overline{F}_0$ represents the constant cosimplicial group with value $\overline{F}_0$.

• Consider the finitely generated free cosimplicial group F. The codegeneracy homomorphism $s^0\colon F_1 \to F_0$ is the constant map and thus $\ker s^0 = F_1 = \langle a_1 \rangle$. Also

$$d^1(a_1) = a_1a_2 = d^2(a_1)d^0(a_1),$$

and hence $a_1 \in Z^1(F)$. Note that any other power of a_1 will fail to satisfy the cocycle condition (1), that is $Z^1(F) = \{e, a_1\}$. Let $F^+ = F^{a_1}$. We denote the canonical inclusion by $\iota_+\colon F \hookrightarrow F^+$. A similar argument shows that $Z^1(F/\Gamma^q) = \{e, a_1\}$ for $q > 2$. We also let $(F/\Gamma^q)^{a_1} = F/\Gamma^{q+}$ and $\iota_+\colon F/\Gamma^q \hookrightarrow F/\Gamma^{q+}$.

- Consider F/Γ^2. As in the previous example, $\ker s_0 = \langle a_1 \rangle$, but since $F_2/\Gamma_2^2 = \mathbb{Z}^2$ all powers of a_1 will satisfy the cocycle condition, that is, $Z^1(F/\Gamma^2) = \mathbb{Z}$. Thus for each positive $m \in \mathbb{Z}$ we get nonisomorphic cosimplicial groups $(F/\Gamma^2)^m$ and inclusions $\iota_m\colon F/\Gamma^2 \hookrightarrow (F/\Gamma^2)^m$. When $m=1$, we write $(F/\Gamma^2)^1 = F/\Gamma^{2+}$.

- Consider $\Sigma_{2,3}$, defined in Example 2.13. The product $\sigma_1\sigma_2 \in (\Sigma_{2,3})_1$ satisfies
$$d^2(\sigma_1\sigma_2)d^0(\sigma_1\sigma_2) = caab = cb = d^1(\sigma_1\sigma_2)$$
and thus $\sigma_1\sigma_2$ is a 1–cocycle and we get the cosimplicial group $\Sigma_{2,3}^{\sigma_1\sigma_2}$.

Now we turn our attention to spaces of homomorphisms. For any topological group G, its underlying group structure defines the functor
$$\mathrm{Hom}_{\mathbf{Grp}}(\_, G)\colon \mathbf{Grp}^{\mathrm{op}} \to \mathbf{Set}.$$
If L is a cosimplicial group, the composition of functors $\mathrm{Hom}_{\mathbf{Grp}}(\_, G)L$, which we denote by $\mathrm{Hom}(L, G)$, defines a simplicial set. Whenever L is finitely generated, for each n we can topologize $\mathrm{Hom}(L_n, G)$ in a way that the induced face and degeneracy maps are continuous. Therefore we get the simplicial space $\mathrm{Hom}(L, G)\colon \Delta^{\mathrm{op}} \to \mathbf{Top}$. We list some known simplicial spaces:

- $\mathrm{Hom}(F, G) = NG$, the nerve of G as a category with one object.
- $\mathrm{Hom}(F^+, G) = (EG)_*$, Steenrod's model for the total space of the universal principal G–bundle $p\colon EG \to BG$, where p is induced by the simplicial map $\iota_+^*\colon \mathrm{Hom}(F^+, G) \to \mathrm{Hom}(F, G)$.
- $\mathrm{Hom}(F/\Gamma^q, G) = (B(q, G))_*$, the underlying simplicial space of the classifying space $B(q, G)$ defined in [4], [6] (for $q=2$) and [7].
- $\mathrm{Hom}(F/\Gamma^{q+}, G) = (E(q, G))_*$, the underlying simplicial space of the total space of the universal bundle $p\colon E(q, G) \to B(q, G)$ also defined in [4], [6] (for $q=2$) and [7]. Again, p is induced by $\iota_+^*\colon \mathrm{Hom}(F/\Gamma^{q+}, G) \to \mathrm{Hom}(F/\Gamma^q, G)$.
- $\mathrm{Hom}(\overline{F}, G) = N\overline{G}$, the nerve of the category $\overline{G}$ that has G as space of objects and a unique morphism between any two objects.

Remark 2.17 Consider the morphism of cosimplicial groups $\gamma\colon F \to \overline{F}$ given on generators by $\gamma^n(a_i) = a_{i-1}a_i^{-1}$, where $\gamma^n\colon F_n \to \overline{F}_n$. Let G be a topological group. The induced map $\gamma\colon N\overline{G} \to NG$ is the underlying simplicial map of Segal's fat geometric realization model for the universal G–bundle. A detailed version of this can be seen in [13, page 66].

2.5 Homotopy stable decomposition of $\mathrm{Hom}(L_n, G)$

Lemma 2.18 *Let G be a linear algebraic group and L a finitely generated cosimplicial group. Let $s^i\colon L_{n+1} \to L_n$ be a codegeneracy map. Then, the image of $s_i := (s^i)^*\colon \mathrm{Hom}(L_n, G) \to \mathrm{Hom}(L_{n+1}, G)$ is a subvariety for all $0 \le i \le n$.*

Proof Suppose $L_{n+1} = \langle a_1, \dots, a_r \rangle$. The homomorphism $s^i\colon L_{n+1} \to L_n$ is surjective, so $L_n \cong L_{n+1}/\ker s^i$. We can describe $\ker s^i = \langle \{b_\alpha\}_{\alpha\in\Lambda} \rangle$, where each b_α is a fixed product of powers of generators a_k. Let $\rho\colon L_{n+1} \to G$ be a homomorphism. Then $(\rho(a_1), \dots, \rho(a_r))$ is in $s_i(\mathrm{Hom}(L_n, G))$ if and only if $\rho(b_\alpha) = I$ for all $\alpha \in \Lambda$. That is, these r–tuples in $\mathrm{Hom}(L_{n+1}, G)$ are determined by the polynomial equations $\{\rho(b_\alpha) = I\}_{\alpha\in\Lambda}$ and hence they build up an affine variety. □

For a cosimplicial group L, write $S_t(L_k, G) := S^t(\mathrm{Hom}(L_k, G))$.

Theorem 2.19 *Let G be a real algebraic linear group, and L a finitely generated cosimplicial group. Then for each n we have homotopy equivalences*

$$\Theta(n)\colon \Sigma\,\mathrm{Hom}(L_n, G) \simeq \bigvee\nolimits_{0\le k\le n} \Sigma(S_k(L_n, G)/S_{k+1}(L_n, G)).$$

Proof Fix n. Using Proposition 2.7, we only need to show that

$$(S_{t-1}(L_n, G), S_t(L_n, G))$$

is a strong NDR pair for all $0 \le t \le n$. By Lemma 2.18, each $s_j(\mathrm{Hom}(L_k, G))$ is an affine variety for all $0 \le j, k \le n$. Then, for all $t \ge 1$, the finite union

$$S_t(L_n, G) = \bigcup_{J_{n,t}} s_{i_1} \circ \dots \circ s_{i_t}(\mathrm{Hom}(L_{n-t}, G))$$

is also an affine variety. Consider the natural filtration

$$S_n(L_n, G) \subset S_{n-1}(L_n, G) \subset \dots \subset S_0(L_n, G) = \mathrm{Hom}(L_n, G).$$

The union $\bigcup_t S_t(L_n, G) = \mathrm{Hom}(L_n, G)$ is an affine variety, and therefore is a closed subspace of some euclidean space. By Remark 2.5, $\mathrm{Hom}(L_n, G)$ can be triangulated in a way that each $S_t(L_n, G)$ is a finite union of interiors of simplices. Since $S_t(L_n, G)$ are closed subspaces, it follows that under the triangulation they are subcomplexes. This way the inclusions $S_t(L_n, G) \subset S_{t-1}(L_n, G)$ are cofibrations and hence NDR pairs. Therefore $\mathrm{Hom}(L, G)$ is simplicially NDR. □

Lemma 2.20 *Let G be a topological group and consider the cosimplicial group F/Γ^q. Then*

$$S_k(F_n/\Gamma^q_n, G)/S_{k+1}(F_n/\Gamma^q_n, G) \cong \bigvee\nolimits^{\binom{n}{k}} \mathrm{Hom}(F_k/\Gamma^q_k, G)/S_1(F_k/\Gamma^q_k, G)$$

for all $1 \le k \le n$.

Proof Let $1 \le i_1 < \cdots < i_{n-k} \le n$. Consider the projections

$$P_{i_1,\dots,i_m}\colon G^n \to G^{n-k}$$

given by $(x_1,\dots,x_n) \mapsto (x_{i_1},\dots,x_{i_{n-k}})$. We claim that the image of $\mathrm{Hom}(F_n/\Gamma_n^q, G)$ under this projection lies on $\mathrm{Hom}(F_{n-k}/\Gamma_{n-k}^q, G)$. Indeed, each projection $P_{i_1,\dots,i_m}$ is induced by the homomorphism $\varphi\colon F_{n-k} \to F_n$ given on generators by $a_j = a_{i_j}$. Since $\varphi(\Gamma_{n-k}^q) \subset \Gamma_n^q$ we get the homomorphism $\bar{\varphi}\colon F_{n-k}/\Gamma_{n-k}^q \to F_n/\Gamma_n^q$ which proves our claim. Assemble the restrictions of $P_{i_1,\dots,i_m}$ to $\mathrm{Hom}(F_n/\Gamma_n^q, G)$ so that we build up a continuous map

$$\eta_n\colon \mathrm{Hom}(F_n/\Gamma_n^q, G) \to \prod_{J_{n,k}} \mathrm{Hom}(F_{n-k}/\Gamma_{n-k}^q, G)$$

given by

$$(x_1,\dots,x_n) \mapsto \{P_{i_1,\dots,i_m}(x_1,\dots,x_n)\}_{(i_1,\dots,i_{n-k})\in J_{n,k}},$$

where $J_{n,k}$ runs over all possible sequences $1 \le i_1 < \cdots < i_{n-k} \le n$ of length $n-k$. Since all sequences $(i_1,\dots,i_{n-k}) \in J_{n,k}$ are disjoint, the restriction

$$\eta_n|_k\colon S_k(F_n/\Gamma_n^q, G) \to \bigvee\nolimits_{J_{n,k}} \mathrm{Hom}(F_{n-k}/\Gamma_{n-k}^q, G)/S_1(F_k/\Gamma_k^q, G)$$

has a continuous inverse $\bigvee_{J_{n,k}} s_{j_1} \circ \cdots \circ s_{j_k}$, where $1 \le j_1 < \cdots < j_k \le n$ and $\{j_1,\dots,j_k\} \cap \{i_1,\dots,i_{n-k}\} = \varnothing$. Therefore $\eta_n|_k$ is a homeomorphism for every k. Finally, note that $S_{k+1}(F_n/K_n, G)$ is mapped to $\bigvee S_1(F_{n-k}/\Gamma_{n-k}^q, G)$. Taking quotients we get the desired homeomorphism. □

The next corollary was first conjectured in [4, page 12] for closed subgroups of $\mathrm{GL}_n(\mathbb{C})$. Since any real linear algebraic group is Zariski closed we have the following version of the conjecture, which follows from Theorem 2.19 and Lemma 2.20.

Corollary 2.21 *If G is a Zariski closed subgroup of $\mathrm{GL}_n(\mathbb{C})$, then there are homotopy equivalences for the cosimplicial group F/Γ^q,*

$$\Theta(n)\colon \Sigma\,\mathrm{Hom}(F_n/\Gamma_n^q, G) \simeq \bigvee\nolimits_{1\le k\le n} \Sigma\big(\bigvee\nolimits^{\binom{n}{k}} \mathrm{Hom}(F_k/\Gamma_k^q, G)/S_1(F_k/\Gamma_k^q, G)\big)$$

for all n and q.

Example 2.22 Let $G = \mathrm{SU}(2)$ and consider F/Γ^q.

- The case $q = 2$ ($F_n/\Gamma_n^2 = \mathbb{Z}^n$) has been largely studied (for example [3; 9; 12]). In this example we follow [3, pages 482–484]. First a few preliminaries. Let T be the maximal torus of G that consists of all diagonal matrices $\left(\begin{smallmatrix}\lambda & 0\\ 0 & \bar{\lambda}\end{smallmatrix}\right)$ with $\lambda \in S^1$ and $W = N(T)/T = \{[w], e\}$ its Weyl group, where $w = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$. The group W acts on T

via $[w]\cdot t = \overline{wtw^{-1}} = t^{-1}$ and using left translation on G/T we get a diagonal action on $G/T \times T^n$. Let $\mathfrak{t} \cong i\mathbb{R}$ be the Lie algebra of T with the induced action of W. There is an equivariant homeomorphism $\mathfrak{t} \to T - \{I\}$ so that $G/T \times_W (T - \{I\})^n \cong G/T \times_W \mathfrak{t}^n$. The quotient map $G/T \to (G/T)/W \cong \mathbb{RP}^2$ is a principal W–bundle and we can take the associated vector bundle $p_n\colon G/T \times_W \mathfrak{t}^n \to \mathbb{RP}^2$. Let λ_2 be the canonical vector bundle over $\mathbb{RP}^2$; then we can identify p_n with $n\lambda_2$, the Whitney sum of n copies of λ_2. The pieces in the homotopy stable decomposition before suspending are

$$\mathrm{Hom}(\mathbb{Z}^n, \mathrm{SU}(2))/S_1(\mathbb{Z}^n, \mathrm{SU}(2)) \cong \begin{cases} S^3 & \text{if } n = 1, \\ (\mathbb{RP}^2)^{n\lambda_2}/s_n(\mathbb{RP}^2) & \text{if } n \geq 2, \end{cases}$$

where $(\mathbb{RP})^{n\lambda_2}$ is the associated Thom space of $n\lambda_2$ and s_n is its zero section. Therefore,

$$\Sigma\,\mathrm{Hom}(\mathbb{Z}^n, \mathrm{SU}(2)) \simeq \Sigma \bigvee_n S^3 \vee \bigvee_{2 \leq k \leq n} \Sigma\bigl(\bigvee^{\binom{n}{k}} (\mathbb{RP}^2)^{k\lambda_2}/s_k(\mathbb{RP}^2)\bigr).$$

- Let $q = 3$. An n–tuple $(g_1, \dots, g_n)$ lies in $\mathrm{Hom}(F_n/\Gamma_n^3, \mathrm{SU}(2))$ if and only if $[[g_i, g_j], g_k] = I$ for all $1 \leq i, j, k \leq n$, ie the commutators $[g_i, g_j]$ are central in the subgroup generated by $g_1, \dots, g_n$. We claim that the center of every nonabelian subgroup of SU(2) sits inside $\{\pm I\}$. Suppose a and b are two elements in SU(2) such that $[a, b] \neq I$. Then, the cyclic groups $\langle a\rangle$ and $\langle b\rangle$ are contained in different tori, say T_1 and T_2, respectively. Since the center of $\langle a, b\rangle$ is abelian it must lie in the intersection $T_1 \cap T_2$. These two circles can only intersect at $\{\pm I\}$, which proves our claim. Therefore the central elements $[g_i, g_j]$ are in $\{\pm I\}$ for all $1 \leq i, j \leq n$. Consider

$$B_n(\mathrm{SU}(2), \{\pm I\}) = \bigl\{(g_1, \dots, g_n) \in \mathrm{SU}(2)^n \mid [g_i, g_j] \in \{\pm I\}\bigr\},$$

the space of almost commuting tuples in SU(2). By the previous observation,

$$\mathrm{Hom}(F_n/\Gamma_n^3, \mathrm{SU}(2)) = B_n(\mathrm{SU}(2), \{\pm I\}).$$

In [3, pages 485–486], they show that

$$B_n(\mathrm{SU}(2), \{\pm I\})/S_1(\mathrm{SU}(2), \{\pm I\}) \\ \cong \mathrm{Hom}(\mathbb{Z}^n, \mathrm{SU}(2))/S_1(\mathbb{Z}^n, \mathrm{SU}(2)) \vee \bigvee_{K(n)} \mathrm{PU}(2)_+,$$

where $K(1) = 0$ and $K(n) = \frac{7^n}{24} - \frac{3^n}{8} + \frac{1}{12}$ for $n \geq 2$. Here $S_1(\mathrm{SU}(2), \{\pm I\})$ are the n–tuples in $B_n(\mathrm{SU}(2), \{\pm I\})$ with at least one coordinate equal to I. Since $\mathrm{PU}(2) \cong \mathbb{RP}^3$ and $S_1(\mathrm{SU}(2), \{\pm I\}) = S_1(F_n/\Gamma_n^3, \mathrm{SU}(2))$ we conclude that

$$\Sigma\,\mathrm{Hom}(F_n/\Gamma_n^3, \mathrm{SU}(2)) \\ \simeq \Sigma \bigvee_n S^3 \vee \bigvee_{2 \leq k \leq n} \Sigma\bigl(\bigvee^{\binom{n}{k}} (\mathbb{RP}^2)^{k\lambda_2}/s_k(\mathbb{RP}^2) \vee \bigvee_{K(k)} \mathbb{RP}^3_+\bigr).$$

Remark 2.23 For $q \geq 4$ we can find nilpotent subgroups of $\mathrm{SU}(2)$ of class q. Indeed, if ξ_n is a representative in $\mathrm{SU}(2)$ of a primitive n^{th} root of unity, then the subgroup generated by the set $\{\xi_{2q}, w\}$ with w as above, is of nilpotency class q. With this we can show that the spaces $\mathrm{Hom}(F_n/\Gamma_n^q, \mathrm{SU}(2))$ for $q \geq 4$ have more connected components than $\mathrm{Hom}(F_n/\Gamma_n^3, \mathrm{SU}(2))$. See [8] for more details.

2.6 Equivariant homotopy stable decomposition of $\mathrm{Hom}(L_n, G)$

Let G and H be topological groups and $f\colon G \to H$ a continuous homomorphism. If L is a finitely generated cosimplicial group then for each n, we have the commutative diagrams

$$\begin{array}{ccc} \mathrm{Hom}(L_n, G) & \xrightarrow{d_i} & \mathrm{Hom}(L_{n-1}, G) \\ {\scriptstyle f_*}\downarrow & & \downarrow{\scriptstyle f_*} \\ \mathrm{Hom}(L_n, H) & \xrightarrow{d_i} & \mathrm{Hom}(L_{n-1}, H) \end{array} \qquad \begin{array}{ccc} \mathrm{Hom}(L_n, G) & \xrightarrow{s_i} & \mathrm{Hom}(L_{n+1}, G) \\ {\scriptstyle f_*}\downarrow & & \downarrow{\scriptstyle f_*} \\ \mathrm{Hom}(L_n, H) & \xrightarrow{s_i} & \mathrm{Hom}(L_{n+1}, H) \end{array}$$

for all $n \geq 0$ and all $0 \leq i \leq n$, so that f_* is a simplicial map. Conjugation by elements of G defines a homomorphism $G \to G$, so that $\mathrm{Hom}(L_n, G)$ is a G–space and each $S_t(L_n, G)$ is a G–subspace.

Definition 2.24 Let M be a G–space. We say that M has a G–CW–structure if there exists a pair (X, ξ) such that X is a G–CW–complex and $\xi\colon X \to M$ is a G–equivariant homeomorphism.

We want to show that for all n, $\mathrm{Hom}(L_n, G)$ has a G–CW–complex structure for which $S_n(L_n, G) \subset S_{n-1}(L_n, G) \subset \cdots \subset \mathrm{Hom}(L_n, G)$ are G–subcomplexes. To show this we slightly generalize some results in [21].

We continue using the techniques of the previous section, so we require G to be a real linear algebraic group. It is known that any compact Lie group has a unique algebraic group structure (see [19, page 247]). Assuming G is a compact Lie group, every representation space of G has finite orbit types (see [20]), so when M is an algebraic G–variety, the equivariant algebraic embedding theorem [21, Proposition 3.2] implies that M has finite orbit types. Also, this theorem guarantees the existence of a G–invariant algebraic map $f\colon M \to \mathbb{R}^d$ for some d such that the induced map $\bar{f}\colon M/G \to f(M)$ is a homeomorphism and $f(M)$ is a closed semialgebraic set in $\mathbb{R}^d$ [21, Lemma 3.4]. If $\tau\colon |K| \to M/G$ is a triangulation, we say that τ is compatible with a family of subsets $\{D_i\}$ of M if $\pi(D_i)$ is a union of some $\tau(\mathrm{int}\,|\sigma|)$, where $\sigma \in K$ and $\pi\colon M \to M/G$ is the quotient map.

Proposition 2.25 *Let G be a compact Lie group, M_0 an algebraic G–variety and $\{M_j\}_{j=1}^n$ a finite system of G–subvarieties of M_0. Then there exists a semialgebraic triangulation $\tau\colon |K| \to M/G$ compatible with the collection*

$$\{M_{j\,(H)} \mid H \text{ is a subgroup of } G\}_{j=0}^n,$$

where $M_{j\,(H)} = \{x \in M_j \mid G_x = gHg^{-1} \text{ for some } g \in G\}$.

Proof Let $H_1, \dots, H_s \subset G$ be the orbit types of G on M and $f\colon M \to \mathbb{R}^d$ as above. By [21, Lemma 3.3], all $M_{j\,(H_i)}$ are semialgebraic sets, and therefore all $f(M_{j\,(H_i)})$ are also semialgebraic. Since i and j vary on finite sets, we can use Remark 2.5 and obtain a semialgebraic triangulation

$$\lambda\colon |K| \to f(M) = \bigcup_{ij} f(M_{j\,(H_i)})$$

such that each $f(M_{j\,(H_i)})$ is a finite union of some $\lambda(\operatorname{int}|\sigma|)$, where $\sigma \in K$. Take $\tau = \bar{f}^{-1} \circ \lambda$. □

Proposition 2.26 *Let G be a compact Lie group. Let M_0 be an algebraic G–variety and $\{M_j\}_{j=1}^k$ a finite system of G–subvarieties. Then M_0 has a G–CW–complex structure such that each M_j is a G–subcomplex of M.*

Proof Let $\tau\colon |K| \to M/G$ be as in Proposition 2.25 and $\pi\colon M \to M/G$ the orbit map. Let K' be a barycentric subdivision of K, which guarantees that, for any simplex Δ^n of K', $\pi^{-1}(\tau(\Delta^n - \Delta^{n-1})) \subset M_{j\,(H_n)}$ for some $H_n \subset G$ and $0 \le j \le k$. Since $\tau|\colon \pi^{-1}(\tau(\Delta^n))/G \to \Delta^n$ is a homeomorphism and the orbit type of $\pi^{-1}(\tau(\Delta^n - \Delta^{n-1}))$ is constant, by [16, Lemma 4.4] there exists a continuous section $s\colon \tau(\Delta^n) \to M_j$ such that $s \circ \tau(\Delta^n - \Delta^{n-1})$ has a constant isotropy subgroup H_n. Consequently there is an equivariant homeomorphism

$$\pi^{-1}\tau(\Delta^n - \Delta^{n-1}) \cong G/H_n \times (\Delta^n - \Delta^{n-1}).$$

Collecting G–cells $Gs \circ \tau(\Delta^n)$ for all simplices of K' we get a G–CW–structure over all M_j for $0 \le j \le k$. □

For a finitely generated cosimplicial group L, let $\operatorname{Rep}(L_n, G) := \operatorname{Hom}(L_n, G)/G$ and $\bar{S}_t(L_n, G) := S_t(L_n, G)/G$.

Theorem 2.27 *Let G be a compact Lie group and L a finitely generated cosimplicial group. Then, for each n, $\Theta(n)$ from* Theorem 2.19 *is a G–equivariant homotopy equivalence, and in particular we get homotopy equivalences*

$$\Sigma \operatorname{Rep}(L_n, G) \simeq \bigvee\nolimits_{1 \le k \le n} \Sigma(\bar{S}_k(L_n, G)/\bar{S}_{k+1}(L_n, G)).$$

Proof Assume $G \subset \mathrm{GL}_N(\mathbb{R})$. Under conjugation by elements of G, $\mathrm{Hom}(L_n, G)$ is an affine G–variety and by Lemma 2.18 the subspaces $S_j(L_n, G)$ are G–subvarieties for all $1 \le j \le n$. Hence, by Proposition 2.26, $\mathrm{Hom}(L_n, G)$ can be given a G–CW–complex structure, where each $S_j(L_n, G)$ is a G–subcomplex. Similarly, the quotient $S_k(L_n, G)/S_{k+1}(L_n, G)$ has a G–CW–complex structure.

To prove that the map $\Theta(n)$ is a G–equivariant homotopy equivalence, first recall that conjugation by elements of G defines a simplicial action on $\mathrm{Hom}(L, G)$, and by the naturality of each $\Theta(n)$, the G–equivariance follows. Let $H \subset G$ be a closed subgroup. The fixed point spaces $\mathrm{Hom}(L_n, G)^H$ and $S_k(L_n, G)^H$ inherit a CW–complex structure, so that $\mathrm{Hom}(L, G)^H$ is simplicially NDR. By Proposition 2.7, we have homotopy equivalences

$$\Theta(n, H)\colon \Sigma(\mathrm{Hom}(L_n, G)^H) \to \bigvee_{0\le k\le n} \Sigma(S_k(L_n, G)^H / S_{k+1}(L_n, G)^H)$$

for each $n \ge 1$. The fixed points map $\Theta(n)^H$ agrees by naturality with $\Theta(n, H)$ and thus is a homotopy equivalence. The result now follows from the equivariant Whitehead theorem. $\square$

Corollary 2.28 *Let G be a compact Lie group. Then the homotopy equivalences in* Corollary 2.21 *are G–equivariant homotopy equivalences, and in particular we get*

$$\Sigma\,\mathrm{Rep}(F_n/\Gamma_n^q, G) \simeq \bigvee_{1\le k\le n} \Sigma\big(\bigvee^{\binom{n}{k}} \mathrm{Rep}(F_k/\Gamma_k^q, G)/\bar{S}_1(F_k/\Gamma_k^q, G)\big).$$

Example 2.29 Let $G = \mathrm{SU}(2)$ and $L = F/\Gamma^q$.

- For $q = 2$, it was proven in [3, page 484] that

$$\mathrm{Rep}(\mathbb{Z}^n, \mathrm{SU}(2))/\bar{S}_1(\mathbb{Z}^n, \mathrm{SU}(2)) \simeq T^{\wedge n}/W = S^n/\Sigma_2,$$

where the action of the generating element on Σ_2 is given by

$$(x_0, x_1, \dots, x_n) \mapsto (x_0, -x_1, \dots, -x_n)$$

for any $(x_0, x_1, \dots, x_n)$. Identifying $S^n = \Sigma S^{n-1}$, we can see the orbit space S^n/Σ_2 as first taking antipodes, and then suspending, that is, $S^n/\Sigma_2 \cong \Sigma\,\mathbb{RP}^{n-1}$. Thus

$$\Sigma\,\mathrm{Rep}(\mathbb{Z}^n, \mathrm{SU}(2)) \simeq \bigvee_{1\le k\le n} \Sigma\big(\bigvee^{\binom{n}{k}} \Sigma\,\mathbb{RP}^{k-1}\big).$$

- Let $q = 3$. We have shown that $\mathrm{Rep}(F_n/\Gamma_n^3, \mathrm{SU}(2)) = B_n(\mathrm{SU}(2), \{\pm I\})/G$ and using the description of these spaces given in [3, p. 486], the stable pieces are

$$\mathrm{Rep}(F_n/\Gamma_n^3, \mathrm{SU}(2))/\bar{S}_1(F_n/\Gamma_n^3, \mathrm{SU}(2)) \simeq \big(\bigvee_{K(n)} S^0\big) \vee \Sigma\,\mathbb{RP}^{n-1},$$

where $K(n)$ is as in Example 2.22. Therefore

$$\Sigma\,\mathrm{Rep}(F_n/\Gamma_n^3, \mathrm{SU}(2)) \simeq \bigvee_{1\le k\le n} \Sigma\big(\bigvee^{\binom{n}{k}} \big(\bigvee_{K(k)} S^0\big) \vee \Sigma\,\mathbb{RP}^{k-1}\big).$$

3 Homotopy properties of $B(L, G)$

3.1 Geometric realization of $\mathrm{Hom}(L, G)$

Definition 3.1 Let L be a finitely generated cosimplicial group and G a topological group. Let

$$B(L, G) := |\mathrm{Hom}(L, G)|.$$

For the cosimplicial groups F/Γ^q we get that $B(F/\Gamma^q, G) = B(q, G)$, the classifying space for G–bundles of transitional nilpotency class less than q. A natural question is whether or not the space $B(L, G)$ is a classifying space for a specific class of G–bundles.

Lemma 3.2 *Let L be a cosimplicial group. The 1–cocycles of L are in one-to-one correspondence with cosimplicial morphisms $F \to L$.*

Proof Suppose b is a 1–cocycle. Any generator $a_j \in F_n$ is in the image of $a_1 \in F_1$ under composition of coface homomorphisms, eg $a_j = (d^0)^{j-1}(d^2)^{n-j}(a_1)$ for all $j \geq 1$. Define $h^n\colon F_n \to L_n$ as $h^n(a_j) = (d^0)^{j-1}(d^2)^{n-j}(b)$. To show that h is cosimplicial, consider the diagrams

$$\begin{array}{ccc} F_{n-1} & \xrightarrow{h^{n-1}} & L_{n-1} \\ {\scriptstyle d^i}\downarrow & & \downarrow{\scriptstyle d^i} \\ F_n & \xrightarrow{h^n} & L_n \end{array} \qquad \begin{array}{ccc} F_{n+1} & \xrightarrow{h^{n+1}} & L_{n+1} \\ {\scriptstyle s^i}\downarrow & & \downarrow{\scriptstyle s^i} \\ F_n & \xrightarrow{h^n} & L_n \end{array}$$

We prove the case of coface homomorphisms. Let $a_j \in F_{n-1}$. On one side we get

$$h_n d^i(a_j) = \begin{cases} (d^0)^{j-1}(d^2)^{n-j}(b) & \text{if } j < i, \\ (d^0)^{j-1}(d^2)^{n-j}(b)(d^0)^j(d^2)^{n-j-1}(b) & \text{if } j = i, \\ (d^0)^j(d^2)^{n-j-1}(b) & \text{if } j > i, \end{cases}$$

and applying the cosimplicial identity $d^k d^l = d^l d^{k-1}$, where $k > l$, we obtain

$$d^i h^{n-1}(a_j) = \begin{cases} (d^0)^{j-1}(d^{i-j+1})(d^2)^{n-j-1}(b) & \text{if } j < i, \\ (d^0)^{j-1}(d^1)(d^2)^{n-j-1}(b) & \text{if } j = i, \\ (d^0)^j(d^2)^{n-j-1}(b) & \text{if } j > i. \end{cases}$$

We need to analyze two cases:

- $j < i$ implies that $i - j + 1 \geq 2$, thus $d^{i-j+1}(d^2)^{n-j-1} = (d^2)^{n-j}$.
- $j = i$. Then the equality follows from (2) applied to $(d^2)^{n-j-1}(b) = b_{n-j}$.

Commutativity for the codegeneracy homomorphisms is similar, but using the cosimplicial identity $s^k d^l = d^l s^{k-1}$ with $k > l$ and condition (3) above. Hence h is uniquely determined by b. Given a morphism $F \to L$, the element b is given by the image of $a_1 \in F_1$. □

Proposition 3.3 *Let L be a cosimplicial group and $h_b\colon F \to L$ be the morphism defined on $F_1 \to L_1$ as $a_1 \mapsto b$. Then the diagram*

$$\begin{tikzcd} F \arrow[r, hook, "\iota_+"] \arrow[d, "h_b"'] & F^+ \arrow[d, "\mathrm{Id} * h_b"] \\ L \arrow[r, hook, "\iota_b"] & L^b \end{tikzcd}$$

is a pushout of cosimplicial groups.

Proof Suppose $f\colon F^+ \to K$ and $g\colon L \to K$ are morphisms such that $f \circ \iota = g \circ h_b$. Define $h\colon L^b \to K$ on each $L_n^b = \overline{F}_0 * L_n$ as $h^n(a_0) = f^n(a_0)$ (here f^n is evaluated on $a_0 \in \overline{F}_0 * F_n$) and $h^n(x) = g^n(x)$ for any $x \in L_n$. To check that h is in fact a cosimplicial homomorphism, by construction of L^b and h, we just need to verify commutativity with coface maps at level $i = 0$. Consider

$$\begin{tikzcd} L_{n-1}^b \arrow[r, "h^{n-1}"] \arrow[d, "d_b^0"'] & K_{n-1} \arrow[d, "d^0"] \\ L_n^b \arrow[r, "h^n"] & K_n \end{tikzcd}$$

We only need to see what happens at $a_0 \in L_{n-1}^b$:

$$\begin{aligned} d^0 h^{n-1}(a_0) &= d^0 f^{n-1}(a_0) = f^n d_+^0(a_0) = f^n(a_0 a_1) = f^n(a_0) f^n(a_1), \\ h^n d_+^0(a_0) &= h^n(a_0 b_n) = f^n(a_0) g^n(b_n). \end{aligned}$$

Let $b = b_1$. By hypothesis, $g^1(b_1) = f^1(a_1)$. Since

$$g^n(b_n) = (d^2)^{n-1} g^1(b_1) \quad \text{and} \quad f^n(a_1) = (d^2)^{n-1} f^1(a_1),$$

the desired equality holds. □

Corollary 3.4 *Let G be a well-based topological group and L a finitely generated cosimplicial group. Using the notation above, suppose $h_b\colon F \to L$ is a morphism. Then the inclusion $\iota_b\colon L \hookrightarrow L^b$ defines a principal G–bundle $|\iota_b^*|\colon B(L^b, G) \to B(L, G)$.*

Proof From the pushout diagram in Proposition 3.3, and applying the functors $\mathrm{Hom}(\_, G)$ and geometric realization, we obtain the pullback diagram

$$\begin{tikzcd} B(L^b, G) \arrow[r] \arrow[d, "|\iota_b^*|"'] & EG \arrow[d] \\ B(L, G) \arrow[r, "|h_b^*|"] & BG \end{tikzcd}$$

and hence $|\iota_b^*|$ is a principal G–bundle. □

Example 3.5 We have seen that there is only one nonconstant homomorphism $h_{a_1} = \mathrm{Id}\colon F \to F$. For $q > 2$ it can be shown that the same is true for $L = F/\Gamma^q$, where $h_{a_1}\colon F \to F/\Gamma^q$ at each n is the quotient homomorphism. The corresponding $B((F/\Gamma^q)^+, G)$ is the space $E(q, G)$ defined in [4, page 94], and $|h_{a_1}^*|\colon B(q, G) \to BG$ is the inclusion. The bundle $E(q, G) \to B(q, G)$ classifies transitionally nilpotent bundles of class less than q (see [7, Section 5]). The case $q = 2$ is more interesting since $Z^1(F/\Gamma^2) = \mathbb{Z}$. For $m = 1$ we obtain $B((F/\Gamma^2)^+, G) = E(2, G)$ and $E(2, G) \to B(2, G)$ classifies transitionally commutative bundles (see [6, Section 2]). Since multiplication by -1 induces a cosimplicial automorphism of F/Γ^2, all constructions are equivalent for $m = -1$. Now let $m > 1$. The bundle $B((F/\Gamma^2)^m, G) \to B(2, G)$ will classify G–bundles whose transition functions $g_{\alpha\beta}\colon U_\alpha \cap U_\beta \to G$ factor through

$$\begin{tikzcd} U_\alpha \cap U_\beta \arrow[r, "\rho_{\alpha,\beta}"] \arrow[dr, "g_{\alpha\beta}"'] & G \arrow[d, "m"] \\ & G \end{tikzcd}$$

where the $\rho_{\alpha\beta}$ are transitionally commutative and m denotes taking the m^{th} power of elements in G.

3.2 Relation between commutative $\mathbb{I}$–monoids and infinite loop spaces

In this section we recall briefly the notion of $\mathbb{I}$–monoid and how it is related to infinite loop spaces. This is more widely covered in [7]. Our goal is to use this machinery to show that, for a finitely generated cosimplicial group L, $B(L, U) = \mathrm{colim}_n\, B(L, U(n))$ is a nonunital E_∞–ring space when $\mathrm{Hom}(L_0, U)$ is path connected.

Let $\mathbb{I}$ stand for the category whose objects are the sets $[0] = \varnothing$ and $[n] = \{1, \dots, n\}$ for each $n \geq 1$, and morphisms are injective functions. Any morphism $j\colon [n] \to [m]$ in $\mathbb{I}$ can be factored as a canonical inclusion $[n] \hookrightarrow [m]$ and a permutation $\sigma \in \Sigma_m$. This category is symmetric monoidal under two different operations. One is concatenation $[n] \sqcup [m] = [n+m]$ with symmetry morphism the permutation $\tau_{m,n} \in \Sigma_{n+m}$ defined as

$$\tau_{m,n}(i) = \begin{cases} n+i & \text{if } i \le m, \\ i-m & \text{if } i > m, \end{cases}$$

and identity object $[0]$. The second operation is the Cartesian product $[m] \times [n] = [mn]$ with $\tau^{\times}_{m,n} \in \Sigma_{mn}$ given by

$$\tau^{\times}_{m,n}((i-1)n + j) = (j-1)m + i,$$

where $1 \le i \le m$ and $1 \le j \le n$. In this case the identity object is $[1]$. Cartesian product is distributive under concatenation (both left and right).

Definition 3.6 An $\mathbb{I}$–*space* is a functor $X\colon \mathbb{I} \to \mathbf{Top}$. This functor is determined by the following:

1. A family of spaces $\{X[n]\}_{n \ge 0}$, where each $X[n]$ is a Σ_n–space;
2. Σ_n–equivariant structural maps $j_n\colon X[n] \to X[n+1]$ (here we consider $X[n+1]$ is a Σ_n–space under the restriction of the Σ_{n+1}–action to the canonical inclusion $\Sigma_n \hookrightarrow \Sigma_{n+1}$) with the property: for any $j\colon [n] \to [m]$ and any $\sigma, \sigma' \in \Sigma_m$ whose restrictions in Σ_n are equal, we have $\sigma \cdot x = \sigma' \cdot x \in X(j)(X[n])$.

We say that an $\mathbb{I}$–space X is a *commutative* $\mathbb{I}$*–monoid* if it is a symmetric monoidal functor $X\colon (\mathbb{I}, \sqcup, [0]) \to (\mathbf{Top}, \times, \{\mathrm{pt}\})$. Additionally, we say that X is a *commutative* $\mathbb{I}$*–rig* if X is also symmetric monoidal with respect to $(\mathbb{I}, \times, [1])$. For the latter definition we also require X to preserve distributivity.

Definition 3.7 Let $\boldsymbol{C}$ be a small category and $Y\colon \boldsymbol{C} \to \mathbf{Top}$ a functor. Denote by $\boldsymbol{C} \ltimes Y$ the category of elements of Y, that is, objects are pairs (c, x) consisting of an object c of $\boldsymbol{C}$ and a point $x \in Y(c)$. A morphism in $\boldsymbol{C} \ltimes Y$ from (c, x) to (c', x') is a morphism $\alpha\colon c \to c'$ in $\boldsymbol{C}$ satisfying the equation $Y(\alpha)(x) = x'$.

Given $Y\colon \boldsymbol{C} \to \mathbf{Top}$, with the notation above, if we consider $\boldsymbol{C} \ltimes Y$ as a topological category whose space of objects and space of morphisms are

$$\coprod_{c \in \mathrm{obj}(\boldsymbol{C})} Y(c) \quad \text{and} \quad \coprod_{f \in \mathrm{mor}(\boldsymbol{C})} Y(f),$$

respectively, then we have that the homotopy colimit of Y is the classifying space $B(\boldsymbol{C} \ltimes Y) = \mathrm{hocolim}_{\boldsymbol{C}} Y$, that is, the realization of the nerve of the category $\boldsymbol{C} \ltimes Y$.

Let X denote a commutative $\mathbb{I}$–monoid. The category of elements $\mathbb{I} \ltimes X$ is a permutative category, that is, a symmetric monoidal category where associativity and unit relations hold strictly. According to [18], the classifying space of a permutative category has an E_∞–space structure, and so we get that $\operatorname{hocolim}_{\mathbb{I}} X$ has an E_∞–space structure. Here we think of an E_∞–space as a space with an operation that is associative and commutative up to a system of coherent homotopies. Thus, the group completion $\Omega B(\operatorname{hocolim}_{\mathbb{I}} X)$ is an infinite loop space. If X is a commutative $\mathbb{I}$–rig, then $\mathbb{I} \ltimes X$ is a bipermutative category and its classifying space is an E_∞–ring space (as explained in [7]), that is, an E_∞–space with an operation that is associative and commutative (up to coherent homotopy) and distributive (up to coherent homotopy) over the E_∞–space operation.

Consider the subcategory of $\mathbb{I}$ consisting of the same set of objects and all isomorphisms. We denote it by $\mathbb{P}$. The (bi)permutative structure on $\mathbb{I} \ltimes X$ restricts to $\mathbb{P} \ltimes X$, so that $\operatorname{hocolim}_{\mathbb{P}} X$ is also an E_∞–space (E_∞–ring space) and its group completion $\Omega B(\operatorname{hocolim}_{\mathbb{P}} X)$ is an infinite loop space (E_∞–ring space). The maps $X[n] \to *$ induce a map of (bi)permutative categories $\mathbb{P} \ltimes X \to \mathbb{P} \ltimes *$ and therefore a map of infinite loop spaces (E_∞–ring spaces)

$$\rho^X \colon \Omega B(\operatorname{hocolim}_{\mathbb{P}} X) \to \Omega B(\operatorname{hocolim}_{\mathbb{P}} *).$$

It follows that the homotopy fiber $\operatorname{hofib} \rho^X$ is an infinite loop space (nonunital E_∞–ring space). Let $X_\infty := \operatorname{hocolim}_{\mathbb{N}} X$, where $\mathbb{N}$ denotes the subcategory of $\mathbb{I}$ with same set of objects and as arrows the canonical inclusions, and X_∞^+ its Quillen plus construction applied with respect to the maximal perfect subgroup of $\pi_1(X_\infty)$. The following proposition is proved in [7, Theorem 3.1].

Proposition 3.8 *Let $X \colon \mathbb{I} \to$ **Top** be a commutative $\mathbb{I}$–monoid. Assume that*

- *the action of Σ_∞ on $H_*(X_\infty)$ is trivial;*
- *the inclusions induce natural isomorphisms $\pi_0(X[n]) \simeq \pi_0(X_\infty)$ of finitely generated abelian groups with multiplication compatible with the Pontrjagin product and in the center of the homology Pontrjagin ring;*
- *the commutator subgroup of $\pi_1(X_\infty)$ is perfect (for each component) and X_∞^+ is abelian.*

Then $\operatorname{hofib} \rho^X \simeq X_\infty^+$*, and in particular X_∞^+ is an infinite loop space.*

Note that the last two conditions of the previous Proposition are satisfied when each $X[n]$ is connected and X_∞ is abelian. Under these hypothesis X_∞ has an infinite loop space structure.

3.3 Nonunital E_∞–ring space structure of $B(L, U)$

Our first example and application of the machinery described in the previous section is showing the classical result

$$\operatorname*{colim}_m U(m) = U$$

has an infinite loop space structure. This will allow us to prove nonunital E_∞–ring space structures on our spaces of interest.

First, we show that $U(\_)$ is a commutative $\mathbb{I}$–rig. Recall that $\Sigma_m \subset U(m)$ as permutation matrices, so that $U(m)$ has a Σ_m action. Consider the inclusions

$$i_m\colon U(m) \to U(m+1), \quad A \mapsto \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix},$$

which are continuous and preserve group structure. The maps i_m restrict to the canonical inclusions $\Sigma_m \hookrightarrow \Sigma_{m+1}$, therefore

$$i_m(\sigma \cdot A) = i_m(\sigma) i_m(A) i_m(\sigma)^{-1} = \sigma \cdot i_m(A),$$

where $\sigma \in \Sigma_m$, $A \in U(m)$ and on the right-hand side $\sigma \in \Sigma_{m+1}$. Now let $\sigma, \sigma' \in \Sigma_r$ with $m < r$ and suppose both restrictions to the subset $\{1, \dots, m\}$ determine equal permutations in Σ_m. Let $i = i_r \circ i_{r-1} \circ \cdots \circ i_m$. Then, for $A \in U(m)$,

$$\sigma \cdot i(A) = \begin{pmatrix} (\sigma|_m) A (\sigma|_m^{-1}) & 0 \\ 0 & I_{r-m} \end{pmatrix} = \begin{pmatrix} (\sigma'|_m) A (\sigma'|_m^{-1}) & 0 \\ 0 & I_{r-m} \end{pmatrix} = \sigma' \cdot i(A).$$

Therefore $U(\_)\colon \mathbb{I} \to \mathbf{Top}$ is a functor. This $\mathbb{I}$–space has a commutative $\mathbb{I}$–rig structure as follows. Let $\oplus_{m,n}\colon U(m) \times U(n) \to U(m+n)$ denote the block sum of matrices, which is a group homomorphism. The (m,n) shuffle map $U(n+m) \to U(n+m)$ is given by $A \mapsto \tau_{m,n} \cdot A$. We have the commutative diagram

$$\begin{array}{ccc} U(m) \times U(n) & \xrightarrow{\oplus_{m,n}} & U(m+n) \\ \Big\downarrow{\scriptstyle \tau} & & \Big\downarrow{\scriptstyle \tau_{m,n}} \\ U(n) \times U(m) & \xrightarrow{\oplus_{n,m}} & U(m+n) \end{array}$$

where $\tau(A, B) = (B, A)$. Therefore $U(\_)$ is a commutative $\mathbb{I}$–monoid. The other monoidal structure is given by $\otimes_{m,n}\colon U(m) \times U(n) \to U(mn)$ the tensor product of matrices. Indeed, by definition $\tau^{\times}_{m,n} \cdot \otimes_{m,n}(A, B) = \otimes_{n,m} \tau(A, B)$, where $A \in U(m)$ and $B \in U(n)$. Since the image $\oplus_{m,n}(U(m) \times U(n))$ correspond to direct sum, then associativity, left and right distributivity over $\otimes_{m,n}$ hold.

Now we check the conditions of Proposition 3.8: the action of Σ_m on $U(m)$ is homologically trivial since conjugation action on $U(m)$ is trivial up to homotopy, $U(m)$ being path connected. The inclusions i_m are cellular and hence $U(\_)_\infty \simeq U$ and since U is an H–space under block sum of matrices, it is abelian. Therefore $U(\_)_\infty \simeq U$ is an infinite loop space (nonunital E_∞–ring space).

Lemma 3.9 *Let L be a finitely generated cosimplicial group and G, H real algebraic linear groups. Let $p_1\colon G\times H\to G$ and $p_2\colon G\times H\to H$ be the projections. Then*

$$B(L,p_1)\times B(L,p_2)\colon B(L,G\times H)\to B(L,G)\times B(L,H)$$

is a natural homeomorphism.

Proof Since $G\times H$ is a direct product, p_1 and p_2 are continuous homomorphism and therefore

$$p=(p_1)_*\times(p_2)_*\colon \mathrm{Hom}(L,G\times H)\to \mathrm{Hom}(L,G)\times\mathrm{Hom}(L,H)$$

is a simplicial map. Its easy to check that in fact p is a simplicial isomorphism. Both G and H being real algebraic imply that $\mathrm{Hom}(L_n,G)$ and $\mathrm{Hom}(L_n,H)$ have a CW–complex structure, and therefore are k–spaces. By [17, Theorem 11.5], the composition

$$B(L,G\times H)\xrightarrow{|p|}\mathrm{Hom}(L,G)\times\mathrm{Hom}(L,H)|\xrightarrow{|\pi_1|\times|\pi_2|}B(L,G)\times B(L,H)$$

is a natural homeomorphism, where $|\pi_1\circ p|\times|\pi_2\circ p|=B(L,p_1)\times B(L,p_2)$. □

Proposition 3.10 *Let L be a finitely generated cosimplicial group; then $B(L,U(\_))$ is a commutative $\mathbb{I}$–rig.*

Proof Consider the $\mathbb{I}$–rig $U(\_)$. Both the structural maps i_m and the action by elements of Σ_m are continuous group homomorphisms and hence $B(L,U(\_))=B(L,\_)U(\_)$ is an $\mathbb{I}$–space. Also, block sum of matrices and tensor product are topological group morphisms, so that with Lemma 3.9 we can define

$$\mu_{m,n}=B(L,\oplus_{m,n})\circ(B(L,p_1)\times B(L,p_2))^{-1},$$
$$\pi_{m,n}=B(L,\otimes_{m,n})\circ(B(L,p_1)\times B(L,p_2))^{-1},$$

where $p_1\colon U(m)\times U(n)\to U(m)$ and $p_2\colon U(m)\times U(n)\to U(n)$ are the projections. Let $p_1'\colon U(n)\times U(m)\to U(n)$ and $p_2'\colon U(n)\times U(m)\to U(m)$ also denote projections. Notice that

$$\tau\circ B(L,p_2')\times B(L,p_1')=B(L,p_1)\times B(L,p_2)\circ B(L,\tau)$$

(where τ, as before, is the symmetry morphism in **Top**). This implies that all properties satisfied by $\oplus_{m,n}$ and $\otimes_{m,n}$ will be preserved by $\mu_{m,n}$ and $\pi_{m,n}$. □

Theorem 3.11 *Let L be a finitely generated cosimplicial group and suppose that the space $\mathrm{Hom}(L_0,U(m))$ is path connected for all $m\ge 1$. Then $B(L,U)$ is a nonunital E_∞–ring space.*

Proof By Proposition 3.10, $B(L, U(\_))$ is a commutative $\mathbb{I}$–rig. It remains to check the conditions of Proposition 3.8. Note that the conjugation action of Σ_n is homologically trivial since it factors through conjugation action on $U(m)$. Since all $\mathrm{Hom}(L_0, U(m))$ are path connected, $|\mathrm{Hom}(L, U(m))| = B(L, U(m))$ is path connected for all $m \geq 1$. The colimit $B(L, U)$ is also an H–space under block sum of matrices, and therefore abelian. □

Example 3.12 The property $\pi_0\big(\mathrm{Hom}(L_0, U(m))\big) = 0$ for all $m \geq 1$ is satisfied by the following cosimplicial groups:

- $L = F/\Gamma^q$ and $L = F/F^{(q)}$ since $L_0 = \{e\}$ in both cases.
- $L = \overline{F}/\Gamma^q$ and $L = \overline{F}/F^{(q)}$ since $\mathrm{Hom}(L_0, U(m)) = U(m)$ in both cases.
- Consider $\Sigma_{2,3}$, and the cosimplicial morphism $h_{\sigma_1\sigma_2}\colon F \to \Sigma_{2,3}$. The image $h_{\sigma_1\sigma_2}(F)$ defines a cosimplicial subgroup of $\Sigma_{2,3}$, such that $h_{\sigma_1\sigma_2}(F)_0 = \{e\}$.

Remark 3.13 The results in this section also apply to the groups SU and Sp. For SO and O the proofs are not exactly similar, but still true. The arguments used in [7, Theorem 4.1] also apply in our case.

References

[1] **A Adem**, **A Bahri**, **M Bendersky**, **F R Cohen**, **S Gitler**, *On decomposing suspensions of simplicial spaces*, Bol. Soc. Mat. Mexicana 15 (2009) 91–102 MR

[2] **A Adem**, **F R Cohen**, *Commuting elements and spaces of homomorphisms*, Math. Ann. 338 (2007) 587–626 MR

[3] **A Adem**, **F R Cohen**, **J M Gómez**, *Stable splittings, spaces of representations and almost commuting elements in Lie groups*, Math. Proc. Cambridge Philos. Soc. 149 (2010) 455–490 MR

[4] **A Adem**, **F R Cohen**, **E Torres Giese**, *Commuting elements, simplicial spaces and filtrations of classifying spaces*, Math. Proc. Cambridge Philos. Soc. 152 (2012) 91–114 MR

[5] **A Adem**, **J M Gómez**, *On the structure of spaces of commuting elements in compact Lie groups*, from "Configuration spaces" (A Bjorner, F Cohen, C De Concini, C Procesi, M Salvetti, editors), CRM Series 14, Ed. Norm., Pisa (2012) 1–26 MR

[6] **A Adem**, **J M Gómez**, *A classifying space for commutativity in Lie groups*, Algebr. Geom. Topol. 15 (2015) 493–535 MR

[7] **A Adem**, **J M Gómez**, **J A Lind**, **U Tillmann**, *Infinite loop spaces and nilpotent K–theory*, Algebr. Geom. Topol. 17 (2017) 869–893 MR

[8] **O Antolín Camarena**, **B Villarreal**, *Nilpotent n–tuples in* $\mathrm{SU}(2)$, preprint (2016) arXiv

[9] **T Baird**, **L C Jeffrey**, **P Selick**, *The space of commuting n–tuples in* $\mathrm{SU}(2)$, Illinois J. Math. 55 (2011) 805–813 MR

[10] **A K Bousfield**, **D M Kan**, *Homotopy limits, completions and localizations*, Lecture Notes in Mathematics 304, Springer (1972) MR

[11] **F R Cohen**, **M Stafa**, *On spaces of commuting elements in Lie groups*, Math. Proc. Cambridge Philos. Soc. 161 (2016) 381–407 MR

[12] **M C Crabb**, *Spaces of commuting elements in* $\mathrm{SU}(2)$, Proc. Edinb. Math. Soc. 54 (2011) 67–75 MR

[13] **J L Dupont**, *Curvature and characteristic classes*, Lecture Notes in Mathematics 640, Springer (1978) MR

[14] **H Hironaka**, *Triangulations of algebraic sets*, from "Algebraic geometry" (R Hartshorne, editor), Amer. Math. Soc., Providence, RI (1975) 165–185 MR

[15] **J E Humphreys**, *Linear algebraic groups*, Graduate Texts in Mathematics 21, Springer (1975) MR

[16] **T Matumoto**, **M Shiota**, *Unique triangulation of the orbit space of a differentiable transformation group and its applications*, from "Homotopy theory and related topics" (H Toda, editor), Adv. Stud. Pure Math. 9, North-Holland, Amsterdam (1987) 41–55 MR

[17] **J P May**, *The geometry of iterated loop spaces*, Lectures Notes in Mathematics 271, Springer (1972) MR

[18] **J P May**, *E_∞ spaces, group completions, and permutative categories*, from "New developments in topology", London Math. Soc. Lecture Note Ser. 11, Cambridge Univ. Press (1974) 61–93 MR

[19] **A L Onishchik**, **E B Vinberg**, *Lie groups and algebraic groups*, Springer (1990) MR

[20] **R S Palais**, *The classification of G–spaces*, Mem. Amer. Math. Soc. 36, Amer. Math. Soc., Providence, RI (1960) MR

[21] **D H Park**, **D Y Suh**, *Equivariant semi-algebraic triangulations of real algebraic G–varieties*, Kyushu J. Math. 50 (1996) 179–205 MR

Department of Mathematics, University of British Columbia
Vancouver, BC, Canada

bernvh@math.ubc.ca

http://www.math.ubc.ca/~bernvh

Received: 11 July 2016 Revised: 26 March 2017

Algebraic & Geometric Topology 17 (2017) 3547–3619

Gorenstein duality for real spectra

J P C GREENLEES
LENNART MEIER

Following Hu and Kriz, we study the C_2–spectra $BP\mathbb{R}\langle n\rangle$ and $E\mathbb{R}(n)$ that refine the usual truncated Brown–Peterson and the Johnson–Wilson spectra. In particular, we show that they satisfy Gorenstein duality with a representation grading shift and identify their Anderson duals. We also compute the associated local cohomology spectral sequence in the cases $n = 1$ and 2.

55P91, 55U30; 55P43, 55Q91

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3547

1 Introduction

1A Background

Philosophy For us, *real spectrum* is a loose term for a C_2–spectrum built from the C_2–spectrum $M\mathbb{R}$ of real bordism, considered by Araki [2], Araki and Murayama [3], Landweber [22], and Hu and Kriz [18]. The present article shows that bringing together real spectra and Gorenstein duality reveals rich and interesting structures.

It is part of our philosophy that theorems about real spectra can often be shown in the same style as theorems for the underlying complex oriented spectra, although the details might be more difficult, and groups needed to be graded over the real representation ring $RO(C_2)$ (indicated by $\star$) rather than over the integers (indicated by $*$). This extends a well known phenomenon: complex orientability of equivariant spectra makes it easy to reduce questions to integer gradings, and we show that even in the absence of complex orientability, good behaviour of coefficients can be seen by grading with representations.

Bordism with reality In studying these spectra, the real regular representation $\rho = \mathbb{R}C_2$ plays a special role. We write σ for the sign representation on $\mathbb{R}$, so $\rho = 1 + \sigma$. One of the crucial features of $M\mathbb{R}$ is that it is *strongly even* in the sense of Meier and Hill [27], ie

(1) the restriction functor $\pi^{C_2}_{k\rho} M\mathbb{R} \to \pi_{2k} MU$ is an isomorphism for all $k \in \mathbb{Z}$, and

(2) the groups $\pi^{C_2}_{k\rho-1} M\mathbb{R}$ are zero for all $k \in \mathbb{Z}$.

We now localize at 2, and (with two exceptions) all spectra and abelian groups will henceforth be 2–local. The Quillen idempotent has a C_2–equivariant refinement, and this defines the C_2–spectrum $BP\mathbb{R}$ as a summand of $M\mathbb{R}_{(2)}$. The isomorphism (1) allows us to lift the usual v_i to classes $\bar{v}_i \in \pi^{C_2}_{(2^i-1)\rho} BP\mathbb{R}$. The real spectra we are interested in are quotients of $BP\mathbb{R}$ by sequences of $\bar{v}_i$ and localizations thereof. For example, we can follow [18] and Hu [17] and define

$$BP\mathbb{R}\langle n\rangle = BP\mathbb{R}/(\bar{v}_{n+1}, \bar{v}_{n+2}, \dots) \quad \text{and} \quad E\mathbb{R}(n) = BP\mathbb{R}\langle n\rangle[\bar{v}_n^{-1}].$$

These spectra are still strongly even, as we will show. Apart from the extensive literature on K-theory with reality (eg Atiyah [4], Dugger [8] and Bruner and Greenlees [7]), real spectra have been studied by Hu and Kriz, in a series of papers by Kitchloo and Wilson (see eg [21] for one of the latest instalments), by Banerjee [5], by Ricka [28] and by Lorman [24]. A crucial point is that a morphism between strongly even C_2–spectra is an equivalence if it is an equivalence of underlying spectra [27, Lemma 3.4].

We are interested in two dualities for real spectra: Anderson duality and Gorenstein duality. These are closely related (see Greenlees and Stojanoska [13]) but apply to different classes of spectra.

Anderson duality The Anderson dual $\mathbb{Z}^X$ of a spectrum X is an integral version of its Brown–Comenetz dual (in accordance with our general principle, $\mathbb{Z}$ denotes the 2–local integers). The homotopy groups of the Anderson dual lie in a short exact sequence

$$0 \to \mathrm{Ext}^1_{\mathbb{Z}}(\pi_{-*-1}X, \mathbb{Z}) \to \pi_*(\mathbb{Z}^X) \to \mathrm{Hom}_{\mathbb{Z}}(\pi_{-*}X, \mathbb{Z}) \to 0. \tag{1.1}$$

One reason to be interested in the computation of Anderson duals is that they show up in universal coefficient sequences; see Anderson [1] or Section 3B. The situation is nicest for spectra that are Anderson self-dual in the sense that $\mathbb{Z}^X$ is a suspension of X. Many important spectra like KU, KO, Tmf (see Stojanoska [31]) or $Tmf_1(3)$ are indeed Anderson self-dual. These examples are all unbounded as the sequence (1.1) nearly forces them to be.

Anderson duality also works C_2–equivariantly as first explored in [28]; the only change in the above short exact sequence is that equivariant homotopy groups are used. The C_2–spectra $K\mathbb{R}$ (see Heard and Stojanoska [14]) and $Tmf_1(3)$ [27] are also C_2–equivariantly Anderson self-dual, at least if we allow suspensions by *representation spheres*.

One simpler example is essential background: if $\underline{\mathbb{Z}}$ denotes the constant Mackey functor (ie with restriction being the identity and induction being multiplication by 2) then the Anderson dual of its Eilenberg–Mac Lane spectrum is the Eilenberg–Mac Lane spectrum for the dual Mackey functor $\underline{\mathbb{Z}}^* = \mathrm{Hom}_{\mathbb{Z}}(\underline{\mathbb{Z}}, \mathbb{Z})$ (ie with restriction being multiplication by 2 and induction being the identity). It is then easy to check that in fact $H(\underline{\mathbb{Z}}^*) \simeq \Sigma^{2(1-\sigma)} H\underline{\mathbb{Z}}$. (From one point of view this is the fact that $\mathbb{R}P^1 = S(2\sigma)/C_2$ is equivalent to the circle). The dualities we find are in a sense all dependent on this one.

Gorenstein duality By contrast with Anderson self-duality, many connective ring spectra are Gorenstein in the sense of Dwyer, Greenlees and Iyengar [9]. We sketch the theory here, explaining it more fully in Sections 6 and 7.

The starting point is a connective commutative ring C_2–spectrum R, whose 0^{th} homotopy Mackey functor is constant at $\mathbb{Z}$:

$$\underline{\pi}_0^{C_2}(R) \cong \underline{\mathbb{Z}}.$$

This gives us a map $R \to H\underline{\mathbb{Z}}$ of commutative ring spectra by killing homotopy groups. We say that R is *Gorenstein* of shift $a \in RO(C_2)$ if there is an equivalence of R–modules

$$\mathrm{Hom}_R(H\underline{\mathbb{Z}}, R) \simeq \Sigma^a H\underline{\mathbb{Z}}.$$

We are interested in the duality this often entails. Note that the Anderson dual $\mathbb{Z}^R$ obviously has the Matlis lifting property

$$\mathrm{Hom}_R(H\underline{\mathbb{Z}}, \mathbb{Z}^R) \simeq H\underline{\mathbb{Z}}^*,$$

where $\mathbb{Z}^* = \mathrm{Hom}_{\mathbb{Z}}(\underline{\mathbb{Z}}, \mathbb{Z})$ as above. Thus if R is Gorenstein, in view of the equivalence $H(\underline{\mathbb{Z}}^*) \simeq \Sigma^{2(1-\sigma)} H\underline{\mathbb{Z}}$, we have equivalences

$$\begin{aligned}\mathrm{Hom}_R(H\underline{\mathbb{Z}}, \mathrm{Cell}_{H\underline{\mathbb{Z}}} R) &\simeq \mathrm{Hom}_R(H\underline{\mathbb{Z}}, R)\\ &\simeq \Sigma^a H\underline{\mathbb{Z}}\\ &\simeq \Sigma^{a-2(1-\sigma)} H(\underline{\mathbb{Z}}^*)\\ &\simeq \mathrm{Hom}_R(H\underline{\mathbb{Z}}, \Sigma^{a-2(1-\sigma)}\mathbb{Z}^R).\end{aligned}$$

Here, $\mathrm{Cell}_{H\underline{\mathbb{Z}}}$ denotes the $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellularization as in Section 2B. We would like to remove the $\mathrm{Hom}_R(H\underline{\mathbb{Z}}, \cdot)$ from the composite equivalence above.

Definition 1.2 We say that R has *Gorenstein duality* of shift b if we have an equivalence of R–modules

$$\mathrm{Cell}_{H\underline{\mathbb{Z}}} R \simeq \Sigma^b \mathbb{Z}^R.$$

As in the nonequivariant setting, the passage from Gorenstein to Gorenstein duality requires showing that the above composite equivalence is compatible with the right action of $\mathcal{E} = \mathrm{Hom}_R(H\underline{\mathbb{Z}}, H\underline{\mathbb{Z}})$. This turns out to be considerably more delicate than the nonequivariant counterpart because connectivity is harder to control; but if one can lift the R–equivalence to an $\mathcal{E}$–equivalence, the conclusion is that if R is Gorenstein of shift a, then it has Gorenstein duality of shift $b = a - 2(1-\sigma)$.

Local cohomology The duality statement becomes more interesting when the cellularization can be constructed algebraically. For any finitely generated ideal J of the $RO(C_2)$–graded coefficient ring $R^{C_2}_\star$, we may form the stable Koszul complex $\Gamma_J R$, which only depends on the radical of J. In our examples, this applies to the augmentation ideal $J = \ker(R^{C_2}_\star \to H\underline{\mathbb{Z}}^{C_2}_\star)$, which may be radically generated by finitely many elements $\bar{v}_i$ in degrees which are multiples of ρ. Adapting the usual proof to the real context, Proposition 3.8 shows that $\Gamma_J R \to R$ is an $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellularization:

$$\mathrm{Cell}_{H\underline{\mathbb{Z}}} R \simeq \Gamma_J R.$$

The $RO(C_2)$–graded homotopy groups of $\Gamma_J R$ can be computed using a spectral sequence involving local cohomology.

Conclusion In favourable cases, the Gorenstein condition on a ring spectrum R implies Gorenstein duality for R; this in turn establishes a strong duality property on the $RO(C_2)$–graded coefficient ring, expressed using local cohomology.

1B Results

Our main theorems establish Gorenstein duality for a large family of real spectra. We present in this introduction the particular cases of $BP\mathbb{R}\langle n\rangle$ and $E\mathbb{R}(n)$, deferring the more general theorem to Section 5. Let again σ denote the nontrivial representation of C_2 on the real line and $\rho = 1+\sigma$ the real regular representation. Furthermore, set $D_n = 2^{n+1}-n-2$ so that $D_n\rho = |\bar{v}_1| + \cdots + |\bar{v}_n|$. Other terms in the statement will be explained in Section 3.

Theorem 1.3 *For each $n \geq 1$, the C_2–spectrum $BP\mathbb{R}\langle n\rangle$ is Gorenstein of shift $-D_n\rho - n$, and has Gorenstein duality of shift $-D_n\rho - n - 2(1-\sigma)$. This means that*

$$\mathbb{Z}_{(2)}^{BP\mathbb{R}\langle n\rangle} \simeq \Sigma^{D_n\rho+n+2(1-\sigma)}\Gamma_{\bar{J}_n} BP\mathbb{R}\langle n\rangle,$$

where $\bar{J}_n = (\bar{v}_1, \ldots, \bar{v}_n)$. This induces a local cohomology spectral sequence

$$H^*_{\bar{J}_n}(BP\mathbb{R}\langle n\rangle^{C_2}_\star) \implies \pi^{C_2}_\star(\Sigma^{-D_n\rho-n-2(1-\sigma)}\mathbb{Z}_{(2)}^{BP\mathbb{R}\langle n\rangle}).$$

Theorem 1.4 *For each $n \geq 1$, the C_2–spectrum $E\mathbb{R}(n)$ has Gorenstein duality of shift $-D_n\rho - (n-1) - 2(1-\sigma)$. This means that*

$$\begin{aligned}\mathbb{Z}_{(2)}^{E\mathbb{R}(n)} &\simeq \Sigma^{D_n\rho+(n-1)+2(1-\sigma)}\Gamma_{\bar{J}_{n-1}} E\mathbb{R}(n)\\ &\simeq \Sigma^{(n+2)(2^{2n+1}-2^{n+2})+n+3}\,\Gamma_{J_{n-1}} E\mathbb{R}(n)\end{aligned}$$

for $J_{n-1} = \bar{J}_{n-1} \cap \pi^{C_2}_* E\mathbb{R}(n)$. This induces likewise a local cohomology spectral sequence.

We note that this has implications for the C_2–fixed point spectrum $(BP\mathbb{R}\langle n\rangle)^{C_2} = BPR\langle n\rangle$. The graded ring

$$\pi_*(BPR\langle n\rangle) = \pi^{C_2}_*(BP\mathbb{R}\langle n\rangle)$$

is the integer part of the $RO(C_2)$–graded coefficient ring $\pi^{C_2}_\star(BP\mathbb{R}\langle n\rangle)$. However, since the ideal $\bar{J}_n$ is not generated in integer degrees, the statement for $BPR\langle n\rangle$ is usually rather complicated, and one of our main messages is that working with the equivariant spectra gives more insight. On the other hand, $ER(n) = E\mathbb{R}(n)^{C_2}$ has integral Gorenstein duality because one can use the additional periodicity to move the representation suspension and the ideal $\bar{J}_n$ to integral degrees.

We will discuss the general result in more detail later, but the two first cases are about familiar ring spectra.

Example 1.5 (see Sections 6 and 11) For $n=1$, connective K-theory with reality $k\mathbb{R}$ is 2–locally a form of $BP\mathbb{R}\langle 1\rangle$. For this example, we can work without 2–localization, so that $\mathbb{Z}$ means the integers. Our first theorem states that $k\mathbb{R}$ is Gorenstein of shift $-\rho-1=-2-\sigma$ and that it has Gorenstein duality of shift $-4+\sigma$. This just means that

$$\mathbb{Z}^{k\mathbb{R}} \simeq \Sigma^{4-\sigma}\,\mathrm{fib}(k\mathbb{R}\to K\mathbb{R}).$$

The local cohomology spectral sequence collapses to a short exact sequence associated to the fibre sequence just mentioned. We will see in Section 11 that the sequence is not split, even as abelian groups.

Theorem 1.4 recovers the main result of [14], ie that $\mathbb{Z}^{K\mathbb{R}} \simeq \Sigma^4 K\mathbb{R}$, which also implies $\mathbb{Z}^{KO} \simeq \Sigma^4 KO$. It is a special feature of the case $n=1$ that we also get a nice duality statement for the fixed points in the connective case. Indeed, by considering the $RO(C_2)$–graded homotopy groups of $k\mathbb{R}$, one sees [7, Corollary 3.4.2] that

$$(k\mathbb{R}\otimes S^{-\sigma})^{C_2} \simeq \Sigma^1 ko.$$

This implies that connective ko has untwisted Gorenstein duality of shift -5, ie that

$$\mathbb{Z}^{ko} \simeq \Sigma^5\,\mathrm{fib}(ko\to KO).$$

This admits a closely related nonequivariant proof, combining the fact that ku is Gorenstein (clear from coefficients) and the fact that complexification $ko\to ku$ is relatively Gorenstein (connective version of Wood's theorem [7, Lemma 4.1.2]).

Example 1.6 (see Examples 4.13 and 5.12 or Lemma 7.1 and Corollary 7.5) The 2–localization of the C_2–spectrum $tmf_1(3)$ is a form of $BP\mathbb{R}\langle 2\rangle$, and the theorem is closely related to results in [27]. It states that $tmf_1(3)$ is Gorenstein of shift $-4\rho-2=-6-4\sigma$ and has Gorenstein duality of shift $-8-2\sigma$. We show in Section 13 that there are nontrivial differentials in the local cohomology spectral sequence.

Passing to fixed points, we obtain the 2–local equivalence

$$BPR\langle 2\rangle = (BP\mathbb{R}\langle 2\rangle)^{C_2} = tmf_0(3).$$

By contrast with the $n=1$ case, as observed in [27], $tmf_0(3)$ does not have untwisted Gorenstein duality of any integer degree.

A variant of Theorem 1.4 also computes the C_2–equivariant Anderson dual of $TMF_1(3)$, and the computation of the Anderson dual of $Tmf_1(3)$ from [27] follows as well.

The results apply to $tmf_1(3)$ and $TMF_1(3)$ themselves (ie with just 3 inverted, and not all other odd primes).

Our main theorem also recovers the main result of [28] about the Anderson self-duality of integral real Morava K-theory.

1C Guide to the reader

While the basic structure of this paper is easily visible from the table of contents, we want to comment on a few features.

The precise statements of our main results can be found in Section 5. We will give two different proofs of them. One (Part III) might be called "the hands-on approach" which is elementary and explicit, and one (Part II) uses Gorenstein techniques inspired by commutative algebra. The intricacy and dependence on specific calculations in the explicit approach make the conceptual approach valuable. The subtlety of the structural requirements of the conceptual approach make the reassurance of the explicit approach valuable. The results from the latter approach are also a bit more general: In Part III, we prove a version of Gorenstein duality for a quite general class of quotients of $BP\mathbb{R}$, but we treat only $BP\mathbb{R}\langle n\rangle$ itself in Part II.

While the Gorenstein approach only relies on the knowledge of the homotopy groups of $H\underline{\mathbb{Z}}$ and the reduction theorem Corollary 4.7, we need detailed information about the homotopy groups of quotients of $BP\mathbb{R}$ for the hands-on approach. In the appendix, we give a streamlined version of the computation of $\pi^{C_2}_\star BP\mathbb{R}$ (which appeared first in [18]). In Section 4, we give a rather self-contained account of the homotopy groups of $BP\mathbb{R}\langle n\rangle$ and of other quotients of $BP\mathbb{R}$, which can also be read independently of the rest of the paper. While some of this is rather technical, most of the time we just have to use Corollary 4.6 whose statement (though not proof, perhaps) is easy to understand.

We give separate arguments for the computation of the Anderson dual of $k\mathbb{R}$ so that this easier case might illustrate the more complicated arguments of our more general theorems. Thus, if the reader is only interested in $k\mathbb{R}$, he or she might ignore most of this paper. More precisely, under this assumption one might proceed as follows: First one looks at Section 11B for a quick reminder on $\pi^{C_2}_\star k\mathbb{R}$, then one skims through Sections 2 and 3 to pick up the relevant definitions, and then one proceeds directly to Section 6 or Section 8 to get the proof of the main result in the case of $k\mathbb{R}$. Afterwards, one may look at the pictures and computations in the rest of Section 11 to see what happens behind the scenes of Gorenstein duality.

Acknowledgements We are grateful to the Hausdorff Institute of Mathematics in Bonn for providing us the opportunity in Summer 2015 for the discussions starting this work. We also thank Vitaly Lorman and Nicolas Ricka for helpful discussions.

Part I Preliminaries and results

2 Basics and conventions about C_2–spectra

2A Basics and conventions

We will work in the homotopy category of genuine G–spectra (ie stable for suspensions by S^V for any finite dimensional representation V) for $G = C_2$, the group of order 2. We will denote by $\otimes$ the derived smash product of spectra.

We may combine the equivariant and nonequivariant homotopy groups of a C_2–spectrum into a Mackey functor, which we denote by $\underline{\pi}_*^{C_2} X$ and denote C_2–equivariant and underlying homotopy groups correspondingly by $\pi_*^{C_2} X$ and $\pi_*^e X$. For an abelian group A, we write $\underline{A}$ for the constant Mackey functor (ie restriction maps are the identity), and $\underline{A}^*$ for its dual (ie induction maps are the identity). We write HM for the Eilenberg–Mac Lane spectrum associated to a Mackey functor M.

Another C_2–spectrum of interest to us is $k\mathbb{R}$, the C_2–equivariant connective cover of Atiyah's K-theory with reality [4]. The term "real spectra" derives from this example. The examples of real bordism and the other C_2–spectra derived from it will be discussed in Section 4.

We will usually grade our homotopy groups by the real representation ring $RO(C_2)$, and we write $M_\star$ for $RO(C_2)$–graded groups. In addition to the real sign representation σ and the regular representation ρ, the virtual representation $\delta = 1 - \sigma$ is also significant. Examples of $RO(C_2)$–graded homotopy classes are the geometric Euler classes $a_V\colon S^0 \to S^V$; in particular, $a = a_\sigma$ will play a central role. In addition to a, we will also often have a class $u = u_{2\sigma}$ of degree 2δ.

We often want to be able to discuss gradings by certain subsets of $RO(C_2)$. To start with, we often want to refer to gradings by multiples of the regular representation (where we write $M_{*\rho}$), but we also need to discuss gradings of the form $k\rho - 1$. For this, we use the notation

$$*\rho- = \{k\rho \mid k \in \mathbb{Z}\} \cup \{k\rho - 1 \mid k \in \mathbb{Z}\}.$$

Following [27], we call an $RO(C_2)$–graded object M *even* if $M_{k\rho-1} = 0$ for all k. An $RO(C_2)$–graded Mackey functor is *strongly even* if it is even and all the Mackey functors in gradings $k\rho$ are constant. We call a C_2–spectrum (strongly) even if its homotopy groups are (strongly) even.

If X is a strongly even C_2–spectrum and $x \in \pi_{2k} X$, we denote by $\bar{x}$ its counterpart in $\pi_{k\rho}^{C_2} X$. If we want to stress that we consider a certain spectrum as a C_2–spectrum, we will also sometimes indicate this by a bar; for example, we may write $\overline{tmf_1(3)}$ if we want to stress the C_2–structure of $tmf_1(3)$.

2B Cellularity

In a general triangulated category, it is conventional to say M is *K–cellular* if M is in the localizing subcategory generated by K (or equivalently by all integer suspensions of K). A reference in the case of spectra is [9, Section 4.1]. We say that a C_2–spectrum M is *K-$\mathbb{R}$–cellular* (for a C_2–spectrum K) if it is in the localizing subcategory generated by the suspensions $S^{k\rho} \otimes K$ for all integers k. We note that this is the same as the localizing subcategory generated by integer suspensions of K and $(C_2)_+ \otimes K$ because of the cofibre sequence

$$(C_2)_+ \to S^0 \to S^\sigma.$$

We say that a map $N \to M$ is a *K-$\mathbb{R}$–cellularization* if N is K-$\mathbb{R}$–cellular and the induced map

$$\mathrm{Hom}(K, N) \to \mathrm{Hom}(K, M)$$

is an equivalence of C_2–spectra. The K-$\mathbb{R}$–cellularization is clearly unique up to equivalence.

We note that cellularity and $\mathbb{R}$–cellularity are definitely different. For example, $(C_2)_+$ is not S^0–cellular, but it is S^0-$\mathbb{R}$–cellular.

In this article, we will only use $\mathbb{R}$–cellularity.

2C The slice filtration

Recall from [16, Section 4.1] or [15] that the *slice cells* are the C_2–spectra of the form

- $S^{k\rho}$ of dimension $2k$,
- $S^{k\rho-1}$ of dimension $2k-1$, and
- $S^k \otimes (C_2)_+$ of dimension k.

A C_2–spectrum X is $\leq k$ if for every slice cell W of dimension $\geq k+1$, the mapping space $\Omega^\infty \mathrm{Hom}_{\mathbb{S}}(W, X)$ is equivariantly contractible. As explained in [16, Section 4.2], this leads to the definition of $X \to P^k X$, which is the universal map into a C_2–spectrum that is $\leq k$. The fibre of

$$X \to P^k X$$

is denoted by $P_{k+1}X$. The *k-slice* $P^k_k X$ is defined as the fibre of

$$P^k X \to P^{k-1} X,$$

or equivalently, as the cofibre of the map $P_{k+1}X \to P_k X$. We have the following two useful propositions:

Proposition 2.1 [15, Corollary 2.12, Theorem 2.18] *If X is an even C_2–spectrum, then $P^{2k-1}_{2k-1} X = 0$ for all $k \in \mathbb{Z}$.*

Proposition 2.2 [15, Corollary 2.16, Theorem 2.18] *If X is a C_2–spectrum such that the restriction map in $\underline{\pi}^{C_2}_{k\rho}$ is injective, then $P^{2k}_{2k} X$ is equivalent to the Eilenberg–Mac Lane spectrum $\underline{\pi}^{C_2}_{k\rho} X$.*

This allows us to give a characterization of an Eilenberg–Mac Lane spectrum based on regular representation degrees.

Corollary 2.3 *Any even C_2–spectrum X with*

$$\underline{\pi}^{C_2}_{k\rho}(X) = \begin{cases} \underline{A} & \text{if } k = 0, \\ 0 & \text{otherwise,} \end{cases}$$

for an abelian group A, is equivalent to $H\underline{A}$.

Proof By the last two propositions, we have

$$P^k_k X \simeq \begin{cases} H\underline{A} & \text{if } k = 0, \\ 0 & \text{otherwise.} \end{cases}$$

By convergence of the slice spectral sequence [16, Theorem 4.42], the result follows. □

3 Anderson duality, Koszul complexes and Gorenstein duality

3A Duality for abelian groups

It is convenient to establish some conventions for abelian groups to start with, so as to fix notation.

Pontrjagin duality is defined for all graded abelian groups A by

$$A^\vee = \mathrm{Hom}_{\mathbb{Z}}(A, \mathbb{Q}/\mathbb{Z}).$$

Similarly, the rational dual is defined by

$$A^{\vee\mathbb{Q}} = \mathrm{Hom}_{\mathbb{Z}}(A, \mathbb{Q}).$$

Since $\mathbb{Q}$ and $\mathbb{Q}/\mathbb{Z}$ are injective abelian groups these two dualities are homotopy invariant and pass to the derived category. Finally the Anderson dual A^* is defined by applying $\mathrm{Hom}_{\mathbb{Z}}(A, \cdot)$ to the exact sequence

$$0 \to \mathbb{Z} \to \mathbb{Q} \to \mathbb{Q}/\mathbb{Z} \to 0,$$

so we have a triangle

$$A^* \to A^{\vee\mathbb{Q}} \to A^{\vee}.$$

If M is a free abelian group, then the Anderson dual is simply calculated by

$$M^* = \mathrm{Hom}_{\mathbb{Z}}(M, \mathbb{Z})$$

(since M is free, the Hom need not be derived).

If M is a graded abelian group which is an $\mathbb{F}_2$–vector space then up to suspension the Anderson dual is the vector space dual:

$$M^{\vee} = \mathrm{Hom}_{\mathbb{F}_2}(M, \mathbb{F}_2) \simeq \Sigma^{-1} M^*$$

(since vector spaces are free, Hom need not be derived).

3B Anderson duality

Anderson duality is the attempt to topologically realize the algebraic duality from the last subsection. It appears that it was invented by Anderson (only published in mimeographed notes [1]) and Kainen [19], with similar ideas by Brown and Comenetz [6]. For brevity and consistency, we will only use the term Anderson duality instead of Anderson–Kainen duality or Anderson–Brown–Comenetz duality throughout. We will work in the category of C_2–spectra, where Anderson duality was first explored by Ricka in [28].

Let I be an injective abelian group. Then we let $I^{\mathbb{S}}$ denote the C_2–spectrum representing the functor

$$X \mapsto \mathrm{Hom}(\pi_*^{C_2} X, I).$$

For an arbitrary C_2–spectrum, we define I^X as the function spectrum $F(X, I^{\mathbb{S}})$. For a general abelian group A, we choose an injective resolution

$$A \to I \to J$$

and define A^X as the fibre of the map $I^X \to J^X$. For example, we get a fibre sequence

$$\mathbb{Z}^X \to \mathbb{Q}^X \to (\mathbb{Q}/\mathbb{Z})^X.$$

In general, we get a short exact sequence of homotopy groups

$$0 \to \mathrm{Ext}_{\mathbb{Z}}(\pi_{-k-1}^{C_2}(X), A) \to \pi_k^{C_2}(A^X) \to \mathrm{Hom}(\pi_{-k}^{C_2}(X), A) \to 0.$$

The analogous exact sequence is true for $RO(C_2)$–graded Mackey functor valued homotopy groups by replacing X by $(C_2/H)_+ \wedge \Sigma^V X$. Our most common choices will be $A = \mathbb{Z}$ and $A = \mathbb{Z}_{(2)}$.

From time to time we use the following property of Anderson duality: If R is a strictly commutative C_2–ring spectrum and M an R–module, then $\mathrm{Hom}_R(M, A^R) \simeq A^M$ as R–modules as can easily be seen by adjunction.

One of the reasons to consider Anderson duality is that it provides universal coefficient sequences. In the C_2–equivariant world, this takes the following form [28, Proposition 3.11]:

$$0 \to \mathrm{Ext}^1_{\mathbb{Z}}(E^{C_2}_{\alpha-1}(X), A) \to (A^E)^{C_2}_{\alpha}(X) \to \mathrm{Hom}_{\mathbb{Z}}(E^{C_2}_{\alpha}(X), A) \to 0,$$

where E and X are C_2–spectra, $\alpha \in RO(C_2)$ and A is an abelian group.

Our first computation is the Anderson dual of the Eilenberg–Mac Lane spectrum of the constant Mackey functor $\underline{\mathbb{Z}}$.

Lemma 3.1 *The Anderson dual of the Eilenberg–Mac Lane spectrum $H\underline{\mathbb{Z}}$ (as an $\mathbb{S}$–module) is given by the following, where $\delta = 1 - \sigma$:*

$$\mathbb{Z}^{H\underline{\mathbb{Z}}} \simeq H\underline{\mathbb{Z}}^* \simeq \Sigma^{2\delta} H\underline{\mathbb{Z}}.$$

Proof The first equivalence follows from the isomorphisms

$$\underline{\pi}^{C_2}_*(\mathbb{Z}^{H\underline{\mathbb{Z}}}) \cong \mathrm{Hom}_{\mathbb{Z}}(\underline{\pi}^{C_2}_{-*} H\underline{\mathbb{Z}}, \mathbb{Z}) \cong \underline{\mathbb{Z}}^*.$$

Since

$$\pi_*^{C_2}(S^{2-2\sigma} \otimes H\underline{\mathbb{Z}}) = H^*_{C_2}(S^{2\sigma-2}; \underline{\mathbb{Z}}) = H^*(S^{2\sigma-2}/C_2; \mathbb{Z}),$$

and $S^{2\sigma} = S^0 * S(2\sigma)$ is the unreduced suspension of $S(2\sigma)$, the second equivalence is a calculation of the cohomology of $\mathbb{R}P^1$. □

Remark 3.2 This proof shows that if C_2 is replaced by a cyclic group of any order, we still have

$$\mathbb{Z}^{H\underline{\mathbb{Z}}} = H\underline{\mathbb{Z}}^* \simeq \Sigma^{\lambda} H\underline{\mathbb{Z}},$$

where $\lambda = \epsilon - \alpha$ (with ϵ the trivial one dimensional complex representation and α a faithful one dimensional representation).

Anderson duality works, of course, also for nonequivariant spectra. We learnt the following proposition comparing the equivariant and nonequivariant version in a conversation with Nicolas Ricka.

Proposition 3.3 *Let A be an abelian group. We have $(A^X)^{C_2} \simeq A^{(X^{C_2})}$ for every C_2–spectrum X.*

Proof Let $\mathrm{inf}_e^{C_2} Y$ denote the inflation of a spectrum Y to a C_2–spectrum with "trivial action", ie the left derived functor of first regarding it as a naive C_2–spectrum with trivial action and then changing the universe. This is the (derived) left adjoint for the fixed point functor [25, Proposition 3.4].

Let I be an injective abelian group. Then there is for every spectrum Y a natural isomorphism

$$\begin{aligned}[Y, (I^X)^{C_2}] &\cong [\mathrm{inf}_e^{C_2} Y, I^X]^{C_2} \\ &\cong \mathrm{Hom}(\pi_0^{C_2}(\mathrm{inf}_e^{C_2} Y \otimes X), I) \\ &\cong \mathrm{Hom}(\pi_0(Y \otimes X^{C_2}), I) \\ &\cong [Y, I^{(X^{C_2})}].\end{aligned}$$

Here, we use that fixed points commute with filtered homotopy colimits and cofibre sequences and therefore also with smashing with a spectrum with trivial action. Thus, there is a canonical isomorphism in the homotopy category of spectra between $I^{(X^{C_2})}$ and $(I^X)^{C_2}$ that is also functorial in I (by Yoneda). For a general abelian group A, we can write $A^{(X^{C_2})}$ as the fibre of $(I^0)^{X^{C_2}} \to (I^1)^{X^{C_2}}$ (and similarly for the other side) for an injective resolution $0 \to A \to I^0 \to I^1$. Thus, we obtain a (possibly noncanonical) equivalence between $A^{(X^{C_2})}$ and $(A^X)^{C_2}$. □

Remark 3.4 An analogous result holds, of course, for every finite group G.

3C Koszul complexes and derived power torsion

Let R be a nonequivariantly E_∞ C_2–ring spectrum and M an R–module. In this section, we will recall two versions of stable Koszul complexes. Among their merits is that they provide models for cellularization or $\mathbb{R}$–cellularization in cases of interest for us. A basic reference for the material in this section is [11].

As classically, the r–power torsion in a module N can be defined as the kernel of $N \to N[1/r]$, we define the *derived J–power torsion* of M with respect to an ideal

$J = (x_1, \dots, x_n) \subseteq \pi_\star^{C_2}(R)$ as

$$\Gamma_J M = \mathrm{fib}\Big(R \to R\Big[\frac{1}{x_1}\Big]\Big) \otimes_R \cdots \otimes_R \mathrm{fib}\Big(R \to R\Big[\frac{1}{x_n}\Big]\Big) \otimes_R M.$$

This is also sometimes called the *stable Koszul complex*, also denoted by $K(x_1, \dots, x_n)$. As shown in [11, Section 3], this only depends on the ideal J and not on the chosen generators. As algebraically, the derived functors of J–power torsion are the local cohomology groups, we might expect a spectral sequence computing the homotopy groups of $\Gamma_J M$ in terms of local cohomology. As in [11, Section 3], this takes the form

$$H_J^s(\pi_{\star+V}^{C_2} M) \implies \pi_{V-s}^{C_2}(\Gamma_J M). \tag{3.5}$$

Our second version of the Koszul complex can be defined in the one-generator case as

$$\kappa_R(x) = \underset{\rightarrow}{\mathrm{holim}}\, \Sigma^{(1-l)|x|} R/x^l$$

for $x \in \pi_\star^{C_2}(R)$. Here, the map $R/x^l \to \Sigma^{-|x|} R/x^{l+1}$ is induced by the diagram of cofibre sequences:

$$\begin{tikzcd}
\Sigma^{|x^l|} R \arrow[r, "x^l"] \arrow[d, "="] & R \arrow[r] \arrow[d, "x"] & R/x^l \arrow[d, dashed] \\
\Sigma^{|x^l|} R \arrow[r, "x^{l+1}"] & \Sigma^{-|x|} R \arrow[r] & \Sigma^{-|x|} R/x^{l+1}
\end{tikzcd}$$

More generally, we can make, for a sequence $\boldsymbol{x} = (x_1, \dots, x_n)$ in $\pi_\star^{C_2}(R)$, the definition

$$\begin{aligned}
\kappa_R(\boldsymbol{x}; M) &:= \kappa_R(x_1) \otimes_R \cdots \otimes_R \kappa_R(x_n) \otimes_R M \\
&\simeq \underset{\rightarrow}{\mathrm{holim}}\, \Sigma^{-((l_1-1)+\cdots+(l_n-1))|x|} M/(x_1^{l_1}, \dots, x_n^{l_n}).
\end{aligned}$$

The spectrum $\kappa_R(x)$ comes with an obvious filtration by $\Sigma^{(1-l)|x|} R/x^l$ with filtration quotients $\Sigma^{-l|x|} R/x$. We can smash these filtrations together to obtain a filtration of $\kappa_R(\boldsymbol{x})$ with filtration quotients wedges of summands of the form

$$\Sigma^{-l_1|x_1|-\cdots-l_n|x_n|} R/(x_1, \dots, x_n);$$

see [32, Definition 1.3.11, Proposition 12] or [33, Remark 2.8, Lemma 2.12]. Using the following lemma, we obtain also a corresponding filtration on $\Gamma_J R$.

Lemma 3.6 *For $\boldsymbol{x}$ as above, we have*

$$\kappa_R(\boldsymbol{x}) \simeq \Sigma^{|x_1|+\cdots+|x_n|+n}\, \Gamma_J R.$$

Proof See [11, Lemma 3.6]. □

We can also define $\kappa_{\boldsymbol{R}}(\boldsymbol{x}; M)$ (and likewise the other versions of Koszul complexes) for an infinite sequence of x_i by just taking the filtered homotopy colimit over all finite subsequences. Usually Lemma 3.6 breaks down in the infinite case.

Remark 3.7 The homotopy colimit defining $\kappa_{\boldsymbol{R}}(\boldsymbol{x}; M)$ has a directed cofinal subsystem, both in the finite and in the infinite case. Indeed, the colimit ranges over all sequences $(l_1, l_2, \dots)$ with only finitely many entries nonzero. For the directed subsystem, we start with $(0, 0, \dots)$ and raise in the n^{th} step the first n entries by 1. Directed homotopy colimit are well known to be weak colimits in the homotopy category of R–modules, ie every system of compatible maps induces a (possibly nonunique) map from the homotopy colimit [26, Section 3.1; 29, Section II.5].

One of the reasons for introducing $\Gamma_J M$ is that it provides a model for the $\mathbb{R}$–cellularization of M with respect to $R/J = (R/x_1) \otimes_{\boldsymbol{R}} \cdots \otimes_{\boldsymbol{R}} (R/x_n)$ in the sense of Section 2B.

Proposition 3.8 *Suppose that $x_1, \dots, x_n \in \pi_{*\rho}^{C_2} R$, and set $J = (x_1, \dots, x_n)$. Then $\Gamma_J M \to M$ is an $\mathbb{R}$–cellularization with respect to R/J in the (triangulated) category of R–modules.*

Proof Clearly, $\kappa_{\boldsymbol{R}}(x_1, \dots, x_n; M)$ is $\mathbb{R}$–cellular with respect to M/J; furthermore M/J is R/J-$\mathbb{R}$–cellular as clearly M is R–cellular. To finish the proof, we have to show that

$$\mathrm{Hom}_{\boldsymbol{R}}(R/J, \Gamma_J M) \to \mathrm{Hom}_{\boldsymbol{R}}(R/J, M)$$

is an equivalence. Note that $\Gamma_J M = \Gamma_{x_n}(\Gamma_{(x_1,\dots,x_{n-1})} M)$. Thus, it suffices by induction to show that

$$\mathrm{Hom}_{\boldsymbol{R}}(A/x, \Gamma_x B) \to \mathrm{Hom}_{\boldsymbol{R}}(A/x, B)$$

is an equivalence for all R–modules A, B. This is equivalent to

$$\mathrm{Hom}_{\boldsymbol{R}}(A/x, B[x^{-1}]) = 0,$$

which is true as multiplication by x induces an equivalence

$$\mathrm{Hom}_{\boldsymbol{R}}(A, B[x^{-1}]) \xrightarrow{x^*} \mathrm{Hom}_{\boldsymbol{R}}(\Sigma^{|x|} A, B[x^{-1}]). \qquad \square$$

Corollary 3.9 *Let M be a connective R–module and A an abelian group. Then the Anderson dual A^M is $\mathbb{R}$–cellular with respect to R/J for every ideal $J \subset \pi_\star^{C_2}$ finitely generated in degrees $a + b\sigma$ with $a \geq 1$ and $a + b \geq 1$.*

Proof By the last proposition, we have to show that $\Gamma_J A^M \simeq A^M$. For this, it suffices to show that $A^M[x^{-1}]$ is contractible for every generator x of J. As M is connective, we know that $\pi_{a+b\sigma} M = 0$ if $a < 0$ and $a + b < 0$ (this follows, for example, using the cofibre sequence $(C_2)_+ \to S^0 \to S^\sigma$). Thus, $\pi_{a+b\sigma} A^M = 0$ if $a > 0$ and $a + b > 0$. The result follows. □

4 Real bordism and the spectra $BP\mathbb{R}\langle n\rangle$

4A Basics and homotopy fixed points

The C_2–spectrum $M\mathbb{R}$ of *real bordism* was originally defined by Araki and Landweber. In the naive model of C_2–spectra, where a C_2–spectrum is just given as a sequence (X_n) of pointed C_2–spaces together with maps

$$\Sigma^\rho X_n \to X_{n+1},$$

it is just given by the Thom spaces $M\mathbb{R}_n = BU(n)^{\gamma_n}$ with complex conjugation as C_2–action. Defining it as a strictly commutative C_2–orthogonal spectrum requires more care and was done in [30, Example 2.14] and [16, Section B.12]. An important fact is that the geometric fixed points of $M\mathbb{R}$ are equivalent to MO (first proven in [3] and reproven in [16, Proposition B.253]).

As shown in [2] and [18, Theorem 2.33], there is a splitting

$$M\mathbb{R}_{(2)} \simeq \bigoplus_i \Sigma^{m_i \rho} BP\mathbb{R},$$

where the underlying spectrum of $BP\mathbb{R}$ agrees with BP. This splitting corresponds on geometric fixed points to the splitting

$$MO \simeq \bigoplus_i \Sigma^{m_i} H\mathbb{F}_2.$$

As shown in [18] (see also the appendix), the restriction map

$$\pi^{C_2}_{*\rho} BP\mathbb{R} \to \pi_{2*} BP$$

is an isomorphism. Choose now arbitrary indecomposables $v_i \in \pi_{2(2^i-1)} BP$ and denote their lifts to $\pi^{C_2}_{(2^i-1)\rho} BP\mathbb{R}$ and their images in $\pi^{C_2}_{(2^i-1)\rho} M\mathbb{R}$ by $\bar{v}_i$. We denote by $BP\mathbb{R}\langle n\rangle$ the quotient

$$BP\mathbb{R}/(\bar{v}_{n+1}, \bar{v}_{n+2}, \dots)$$

in the homotopy category of $M\mathbb{R}$–modules. At least a priori, this depends on the choice of v_i.

We want to understand the homotopy groups of $BP\mathbb{R}\langle n\rangle$. This was first done by Hu in [17] (beware though that Theorem 2.2 is not correct as stated there) and partially redone in [20]. For the convenience of the reader, we will give the computation again. Note that our proofs are similar but not identical to the ones in the literature. The main difference is that we do not use ascending induction and prior knowledge of $H\mathbb{Z}$ to compute $\Phi^{C_2}BP\mathbb{R}\langle n\rangle$, but precise knowledge about $\pi_\star^{C_2}BP\mathbb{R}$; this is not simpler than the original approach, but gives extra information about other quotients of $BP\mathbb{R}$, which we will need later. We recommend that the reader looks at the appendix for a complete understanding of the results that follow.

We will use the Tate square [12] and consider the following diagram in which the rows are cofibre sequences:

$$\begin{tikzcd}
BP\mathbb{R}\langle n\rangle\otimes(EC_2)_+ \arrow[r] \arrow[d, "\simeq"] & BP\mathbb{R}\langle n\rangle \arrow[r] \arrow[d] & BP\mathbb{R}\langle n\rangle\otimes\widetilde{E}C_2 \arrow[r] \arrow[d] & \Sigma BP\mathbb{R}\langle n\rangle\otimes(EC_2)_+ \arrow[d] \\
BP\mathbb{R}\langle n\rangle^{(EC_2)_+}\otimes(EC_2)_+ \arrow[r] & BP\mathbb{R}\langle n\rangle^{(EC_2)_+} \arrow[r] & BP\mathbb{R}\langle n\rangle^{(EC_2)_+}\otimes\widetilde{E}C_2 \arrow[r] & \Sigma BP\mathbb{R}\langle n\rangle\otimes(EC_2)_+
\end{tikzcd}$$

After taking fixed points, this becomes the diagram:

$$\begin{tikzcd}
BP\mathbb{R}\langle n\rangle_{hC_2} \arrow[r] \arrow[d, "="] & BP\mathbb{R}\langle n\rangle^{C_2} \arrow[r] \arrow[d] & BP\mathbb{R}\langle n\rangle^{\Phi C_2} \arrow[r] \arrow[d] & \Sigma BP\mathbb{R}\langle n\rangle_{hC_2} \arrow[d] \\
BP\mathbb{R}\langle n\rangle_{hC_2} \arrow[r] & BP\mathbb{R}\langle n\rangle^{hC_2} \arrow[r] & BP\mathbb{R}\langle n\rangle^{tC_2} \arrow[r] & \Sigma BP\mathbb{R}\langle n\rangle_{hC_2}
\end{tikzcd}$$

First, we compute the homotopy groups of the homotopy fixed points. For this, we use the $RO(C_2)$–graded homotopy fixed point spectral sequence, described for example in [27, Section 2.3].

Proposition 4.1 *The $RO(C_2)$–graded homotopy fixed point spectral sequence*

$$E_2 = H^*(C_2;\pi_\star^e BP\mathbb{R}\langle n\rangle)\cong \mathbb{Z}_{(2)}[\bar{v}_1,\dots,\bar{v}_n,u^{\pm1},a]/2a \implies \pi_\star^{C_2}(BP\mathbb{R}\langle n\rangle^{(EC_2)_+})$$

has differentials generated by $d_{2^{i+1}-1}(u^{2^{i-1}})=a^{2^{i+1}-1}\bar{v}_i$ for $i\le n$ and $E_{2^{n+1}}=E_\infty$.

Proof The description of $E_{2^{n+1}}$ is entirely analogous to the proof of A.2, using that $a^{2^{i+1}-1}\bar{v}_i=0$ in $\pi_\star^{C_2}BP\mathbb{R}\langle n\rangle^{(EC_2)_+}$. Now we need to show that there are no further differentials: As every element in filtration f is divisible by a^f in $E^{2^{n+1}}$, the existence of a nonzero d_m (with $m\ge 2^{n+1}$) implies the existence of a nonzero d_m with source in the 0–line. Moreover, a nonzero d_m of some element $u^l\bar{v}$ (for $\bar{v}$ a polynomial in the $\bar{v}_i$) on the 0–line implies a nonzero d_m on u^l as $\bar{v}$ is a permanent cycle (in the image from $BP\mathbb{R}$). The image of such a differential must be of the form $a^m u^{l'} \bar{v}'$, where $\bar{v}'$ is a polynomial in $\bar{v}_1, \dots, \bar{v}_n$. As $a^m \bar{v}_i = 0$ for $1 \le i \le n$ in $E^{2^{n+1}}$, the polynomial $\bar{v}'$ must be constant. Counting degrees, we see that

$$(2l-1) - 2l\sigma = |u^l| - 1 = |a^m u^{l'}| = 2l' - (2l'+m)\sigma,$$

and thus $m = 2l - 2l' = 1$. This is clearly a contradiction. $\square$

Corollary 4.2 *We have*

$$\pi_\star^{C_2}\big(BP\mathbb{R}\langle n\rangle^{(EC_2)_+} \otimes \widetilde{E}C_2\big) \cong \mathbb{F}_2[u^{\pm 2^n}, a^{\pm 1}].$$

In particular, we get $\pi_* BP\mathbb{R}\langle n\rangle^{tC_2} \cong \mathbb{F}_2[x^{\pm 1}]$, where $x = u^{2^n} a^{-2^{n+1}}$ and $|x| = 2^{n+1}$. These are understood to be additive isomorphisms.

Proof Recall that

$$\pi_\star^{C_2}\big(BP\mathbb{R}\langle n\rangle^{(EC_2)_+} \otimes \widetilde{E}C_2\big) = \pi_\star^{C_2}\big(BP\mathbb{R}\langle n\rangle^{(EC_2)_+}\big)[a^{-1}].$$

as $S^{\infty\sigma}$ is a model of $\widetilde{E}C_2$. The result follows as all $\bar{v}_i$ are a–power torsion, but $u^{2^n m}$ is not. $\square$

4B The homotopy groups of $BP\mathbb{R}\langle n\rangle$

Computing the homotopy groups of the fixed points is more delicate than the computation of the homotopy fixed points. We first have to use our detailed knowledge about the homotopy groups of $BP\mathbb{R}$. Given a sequence $\underline{l} = (l_1, \dots)$, we denote by $BP\mathbb{R}/\bar{v}^{\underline{l}}$ the spectrum $BP\mathbb{R}/(\bar{v}_{i_1}^{l_{i_1}}, \bar{v}_{i_2}^{l_{i_2}}, \dots)$, where i_j runs over all indices such that $l_{i_j} \neq 0$. Similarly $BP\mathbb{R}/\bar{v}_i^j$ is understood to be $BP\mathbb{R}$ if $j = 0$. We use the analogous convention when we have algebraic quotients of homotopy groups.

We recommend the reader skips the proof of the following result for first reading, as the technical detail is not particularly illuminating.

Proposition 4.3 *Let $k \ge 1$ and let $\underline{l} = (l_1, l_2, \dots)$ be a sequence of nonnegative integers with $l_k = 0$. Then the element $\bar{v}_k$ acts injectively on $(\pi_{*\rho-c}^{C_2} BP\mathbb{R})/\bar{v}^{\underline{l}}$ if $0 \le c \le 2^{k+1}+1$, with a splitting on the level of $\mathbb{Z}_{(2)}$–modules.*

Proof Recall from the appendix that $\pi_\star^{C_2} BP\mathbb{R}$ is isomorphic to the subalgebra of

$$P/(2a, \bar{v}_i a^{2^{i+1}-1})$$

(where i runs over all positive integers) generated by $\bar{v}_i(j) = u^{2^i j}\bar{v}_i$ (with $i, j \in \mathbb{Z}$ and $i \ge 0$) and a, where $P = \mathbb{Z}_{(2)}[a, \bar{v}_i, u^{\pm 1}]$. The degrees of the elements are $|a| = 1-\rho$ and

$$|\bar{v}_i(j)| = (2^i - 1)\rho + 2^i j(4-2\rho) = (2^i - 1 - 2^{i+1} j)\rho + 2^{i+2} j.$$

We add the relations $\bar{v}_i^{l_i} = 0$ if $l_i \neq 0$.

We will first show that the ideal of $\bar{v}_k$–torsion elements in $(\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}}$ is contained in the ideal generated by $a^{2^{k+1}-1}$ and $\bar{v}_s^{l_s-1}\bar{v}_s(j)$ for s with $l_s \neq 0$ and j divisible by 2^{k-s} if $s < k$. Indeed, because the ideal $(2a, \bar{v}_i a^{2^{i+1}-1}, \underline{\bar{v}}^{\underline{l}}) \subset P$ is generated by monomials, a polynomial in P defines a $\bar{v}_k$–torsion element in $(\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}}$ if and only if each of its monomials define $\bar{v}_k$–torsion elements. A monomial x_P in P can only define a nonzero $\bar{v}_k$–torsion element in $(\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}}$ if it is divisible by $a^{2^{k+1}-1}$ or $\bar{v}_s^{l_s}$. In the latter case, x_P is of the form $\bar{v}\bar{v}_s^{l_s}u^m$, where $\bar{v}$ is a polynomial in the $\bar{v}_i$. This is divisible by $\bar{v}_s^{l_s}$ in $\pi_\star^{C_2} BP\mathbb{R}$ if and only if m is divisible by 2^i for some $\bar{v}_i$ in $\bar{v}$. Thus, x_P defines a nonzero element x in $(\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}}$ such that $\bar{v}_k x$ defines 0 only if $2^k | m$, which corresponds to the condition above.

Let x be a nonzero $\bar{v}_k$–torsion element in $(\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}}$, represented by a monomial in P. First assume that x is divisible by a^n with $n \geq 2^{k+1}-1$, but not by a^{n+1}. Then, x is not divisible by any $\bar{v}_i(j)$ with $i \leq k$ as $a^n\bar{v}_i(j) = 0$. We demand that x is in degree $*\rho - c$ with $c \geq 0$; in particular, $x \neq a^n$. Let $\bar{v}_i(j)$ a divisor of x with minimal i. Thus, the degree of x must be of the form $*\rho + 2^{i+2}m + n$. We know that $n \leq 2^{i+1}-2$. The largest negative value the non-ρ-part can take is $-2^{i+2} + 2^{i+1} - 2 = -2^{i+1} - 2$. The statement about injectivity follows in this case as $i > k$.

Now assume that x is a $\bar{v}_k$–torsion element not divisible by a^n for $n \geq 2^{k+1}-1$. Then x must be of the form $\bar{v}_s^{l_s-1}\bar{v}_s(j)$, where j is divisible by 2^{k-s} if $s < k$. Observe that

$$\bar{v}_s^{l_s-1}\bar{v}_s(j)\bar{v}_t(m) = \bar{v}_s^{l_s}\bar{v}_t(2^{s-t}j + m) = 0 \in (\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}}$$

for $t < s$, so y is not divisible by any $\bar{v}_t(m)$ for $t < s$. Likewise observe that if $s \leq t \leq k$, then

$$\bar{v}_s^{l_s-1}\bar{v}_s(j)\bar{v}_t(m) = \bar{v}_s^{l_s}\bar{v}_t(m + 2^{k-t}j') = 0 \in (\pi_\star^{C_2} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}},$$

where $j = 2^{k-s}j'$. Thus, y is also not divisible by any $\bar{v}_t(m)$ with $s \leq t \leq k$. As $|\bar{v}_s(j)| = *\rho + d$, where d is divisible by 2^{k+2}, and the same is true for $|\bar{v}_t(j)|$ with $t > k$, we see that if $|x|$ is of the form $*\rho - c$ with $c \geq 0$, then we have

$$c \geq 2^{k+2} - (2^{k+1} - 2) = 2^{k+1} + 2.$$

The statement about injectivity follows also in this case.

We still have to show the split injectivity. If $\bar{v}_k y = 2z$, but y is not divisible by 2, then y must be of the form $2\bar{v}u^{2^k j}$ in P, where $\bar{v}$ is a polynomial in the $\bar{v}_i$. Thus, $|y| = 2^{k+2}j + *\rho$, so we are fine in degree $*\rho - c$ for $0 \leq c \leq 2^{k+1}+1 \leq 2^{k+2}-1$. □

Remark 4.4 The exact bounds in the preceding proposition are not very important. The only important part for later inductive arguments is that the bound grows with k.

Corollary 4.5 *Let $\underline{l} = (l_1, l_2, \dots)$ be a sequence with only finitely many nonzero entries, and let j be the smallest index such that $l_j \neq 0$. Then the map*

$$(\pi^{C_2}_{*\rho-c} BP\mathbb{R})/\underline{\bar{v}}^{\underline{l}} \to \pi^{C_2}_{*\rho-c}(BP\mathbb{R}/\underline{\bar{v}}^{\underline{l}})$$

is an isomorphism for $0 \le c \le 2^{j+1}$.

Proof We use induction on the number n of nonzero indices in $\underline{l}$. If $n = 0$ (and $j = \infty$), the statement is clear.

For the step, define $\underline{l}'$ to be the sequence obtained from $\underline{l}$ by setting $l_j = 0$. Consider the short exact sequence

$$0 \to \big(\pi^{C_2}_{*\rho-c}(B/\underline{\bar{v}}^{\underline{l}'})\big)/\bar{v}_j^{l_j} \to \pi^{C_2}_{*\rho-c}(B/\underline{\bar{v}}^{\underline{l}}) \to \big\{\pi^{C_2}_{*\rho-(c+1)}(B/\underline{\bar{v}}^{\underline{l}'})\big\}_{\bar{v}_j^{l_j}} \to 0.$$

Here, the notation $\{X\}_z$ denotes the elements in X killed by z.

Assume $c \le 2^{j+1}$. By the induction hypothesis, $\pi^{C_2}_{*\rho-c}(B/\underline{\bar{v}}^{\underline{l}'}) \cong (\pi^{C_2}_{*\rho-c} B)/\underline{\bar{v}}^{\underline{l}'}$ as $c \le 2^{j+2}$, so $(\pi^{C_2}_{*\rho-c}(B/\underline{\bar{v}}^{\underline{l}'}))/\bar{v}_j^{l_j} \cong (\pi^{C_2}_{*\rho-c} B)/\underline{\bar{v}}^{\underline{l}}$. Furthermore,

$$\big\{\pi^{C_2}_{*\rho-(c+1)}(B/\underline{\bar{v}}^{\underline{l}'})\big\}_{\bar{v}_j^{l_j}} \cong \big\{(\pi^{C_2}_{*\rho-(c+1)} B)/\underline{\bar{v}}^{\underline{l}'}\big\}_{\bar{v}_j^{l_j}} \cong 0,$$

as follows from $c+1 \le 2^{j+2}$ and $c+1 \le 2^{j+1}+1$ by the induction hypothesis and Proposition 4.3. Thus $(\pi^{C_2}_{*\rho-c} B)/\underline{\bar{v}}^{\underline{l}} \to \pi^{C_2}_{*\rho-c}(B/\underline{\bar{v}}^{\underline{l}})$ is an isomorphism. □

The following corollary is crucial:

Corollary 4.6 *Let $I \subset \mathbb{Z}_{(2)}[\bar{v}_1, \dots]$ be an ideal generated by powers of the $\bar{v}_i$. Then $BP\mathbb{R}/I$ is strongly even.*

Proof As being strongly even is a property closed under filtered homotopy colimits, we are reduced to the case that I is finitely generated. By the last corollary, it suffices to show that $BP\mathbb{R}$ itself is strongly even. That the Mackey functor $\underline{\pi}^{C_2}_{*\rho}(BP\mathbb{R})$ is constant is clear from Theorem A.4.

Assume that x is a nonzero element in $\pi^{C_2}_{*\rho-1} BP\mathbb{R}$. We can assume that x is represented by $a^k u^l \bar{v}$ in the E_2–term of the homotopy fixed point spectral sequence for $BP\mathbb{R}$, where $\bar{v}$ is a monomial in the $\bar{v}_i$ (with $\bar{v}_0 = 2$), $k \ge 0$ and $l \in \mathbb{Z}$. The element x is in degree $k + 4l + *\rho$. Let $j \ge 0$ be the minimal number such that $\bar{v}_j | \bar{v}$. Then $2^j | l$ and $k \le 2^{j+1} - 2$. This is clearly in contradiction with $k + 4l = -1$. □

We recover the C_2–case of the reduction theorem of [18, Proposition 4.9] and [16, Theorem 6.5].

Corollary 4.7 *There is an equivalence* $BP\mathbb{R}/(\bar{v}_1, \bar{v}_2, \dots) \simeq H\underline{\mathbb{Z}}_{(2)}$.

Proof This follows directly from the last corollary and Corollary 2.3. □

Corollary 4.8 *Let* $I \subset \mathbb{Z}_{(2)}[\bar{v}_1, \dots]$ *be an ideal generated by powers of the* $\bar{v}_i$. *Then*

$$\pi^{C_2}_{*\rho+1} BP\mathbb{R}/I \cong \mathbb{F}_2\{a\} \otimes \mathbb{Z}_{(2)}[\bar{v}_1, \bar{v}_2, \dots]/I.$$

Proof As $BP\mathbb{R}/I$ is strongly even, this follows from [27, Lemma 2.15]. □

This allows us to compute $\pi^{C_2}_{\star} BP\mathbb{R}\langle n\rangle$.

Proposition 4.9 *The spectrum* $BP\mathbb{R}\langle n\rangle$ *is the connective cover of its Borel completion* $BP\mathbb{R}\langle n\rangle^{(EC_2)_+}$. *The cofibre* C *of* $BP\mathbb{R}\langle n\rangle \to BP\mathbb{R}\langle n\rangle^{(EC_2)_+}$ *has homotopy groups*

$$\pi^{C_2}_{\star} C \cong \mathbb{F}_2[a^{\pm 1}, u^{-2^n}]u^{-2^n},$$

with the naming of the elements indicating the map $\pi^{C_2}_{\star} BP\mathbb{R}\langle n\rangle^{(EC_2)_+} \to \pi^{C_2}_{\star} C$.

Proof This is clear on underlying homotopy groups. Thus, we have only to show that $BP\mathbb{R}\langle n\rangle^{C_2} \to BP\mathbb{R}\langle n\rangle^{hC_2}$ is a connective cover. For that purpose, it is enough to show that $BP\mathbb{R}\langle n\rangle^{\Phi C_2}$ is connective and that the fibre of $BP\mathbb{R}\langle n\rangle^{\Phi C_2} \to BP\mathbb{R}\langle n\rangle^{tC_2}$ has homotopy groups only in negative degrees.

We have $BP\mathbb{R}\langle n\rangle^{\Phi C_2} \simeq BP\mathbb{R}^{\Phi C_2}/(\bar{v}_{n+1}, \dots)$. As noted in the proof of Proposition A.1, the image of $\bar{v}_i$ in $M\mathbb{R}^{\phi C_2}$ is 0. As the quotient $BP\mathbb{R}^{\Phi C_2}/(\bar{v}_{n+1}, \dots)$ can be taken in the category of $M\mathbb{R}^{\Phi C_2}$–modules, we are only quotienting out by 0. It follows easily that $(BP\mathbb{R}/(\bar{v}_{n+1}, \dots, \bar{v}_{n+m}))^{\Phi C_2}$ has in the homotopy groups an $\mathbb{F}_2$ in all degrees of the form $\sum_{i=n+1}^{n+m} \varepsilon_i(|v_i|+1) = \sum_{i=n+1}^{n+m} \varepsilon_i 2^i$ with $\varepsilon_i = 0$ or 1. As geometric fixed point commute with homotopy colimits, we see that $\pi_* BP\mathbb{R}\langle n\rangle^{\Phi C_2} \cong \mathbb{F}_2[y]$ (additively) with $|y| = 2^{n+1}$. It remains to show that y^k maps nonzero to $\pi_* BP\mathbb{R}\langle n\rangle^{tC_2}$ (and hence maps to x^k).

It is enough to show that $a^{-|y^k|-1} y^k$ maps nonzero to $\pi^{C_2}_{\star} \Sigma BP\mathbb{R}\langle n\rangle \otimes (EC_2)_+$ in the sequence coming from the Tate square, ie that $a^{-|y^k|-1} y^k$ is not in the image from (the fixed points of) $BP\mathbb{R}\langle n\rangle$. But $a^{-|y^k|-1} y^k$ is in degree $(|y^k|+1)\rho - 1$ and $\pi^{C_2}_{(|y^k|+1)\rho-1} BP\mathbb{R}\langle n\rangle = 0$ by Corollary 4.6. □

Let us describe the homotopy groups of $BP\mathbb{R}\langle n\rangle$ in more detail. We set $\bar{v}_0 = 2$ for convenience. Denote by BB (for *basic block*) the $\mathbb{Z}_{(2)}[a, \bar{v}_1, \dots, \bar{v}_n]/2a$–submodule of

$$\mathbb{Z}_{(2)}[\bar{v}_1, \dots, \bar{v}_n]/(a^{2^{k+1}-1}\bar{v}_k)_{0 \le k \le n}$$

generated by 1 and by $\bar{v}_k(m) = u^{2^k m}\bar{v}_k$ for $0 \le k < n$ and $0 < m < 2^{n-k}$. By Proposition 4.1, we see that

$$\pi_\star^{C_2} BP\mathbb{R}\langle n\rangle^{(EC_2)_+} \cong BB[U^{\pm 1}]$$

with $U = u^{2^n}$. Note that this isomorphism is not claimed to be multiplicative; in general, $BP\mathbb{R}\langle n\rangle$ is not even known to have any kind of (homotopy unital) multiplication.

Define BB' to be the kernel of the map $BB \to \mathbb{F}_2[a]$ given by sending all $\bar{v}_k$ and $\bar{v}_k(m)$ to zero. Set $NB = \Sigma^{\sigma-1}\mathbb{F}_2[a]^\vee \oplus BB'$, where NB stands for *negative block*. Then it is easy to see from the last proposition that

$$\pi_\star^{C_2} BP\mathbb{R}\langle n\rangle \cong BB[U] \oplus U^{-1}NB[U^{-1}],$$

where this isomorphism is again only meant additively. We will be a little bit more explicit about the homotopy groups of $BP\mathbb{R}\langle n\rangle$ in the cases $n = 1$ and 2 in Part IV.

4C Forms of $BP\mathbb{R}\langle n\rangle$

Our next goal is to identify certain spectra as forms of $BP\mathbb{R}\langle n\rangle$. We take the following definition from [27]:

Definition 4.10 Let E be an even 2–local commutative and associative C_2–ring spectrum up to homotopy. By [27, Lemma 3.3], E has a real orientation, and after choosing one, we obtain a formal group law on $\pi_{*\rho}^{C_2} E$. The 2–typification of this formal group law defines a map $\pi_{2*}^e BP \cong \pi_{*\rho}^{C_2} BP\mathbb{R} \to \pi_{*\rho}^{C_2} E$. We call E a *form of* $BP\mathbb{R}\langle n\rangle$ if the map

$$\underline{\mathbb{Z}_{(2)}[\bar{v}_1, \dots, \bar{v}_n]} \subset \underline{\pi}_{*\rho} BP\mathbb{R} \to \underline{\pi}_{*\rho} E$$

is an isomorphism of constant Mackey functors.

This depends neither on the choice of $\bar{v}_i$ nor on the chosen real orientation, as can be seen using that $\bar{v}_i$ is well defined modulo $(2, \bar{v}_1, \dots, \bar{v}_{i-1})$.

Equivalently, one can say that E is a form of $BP\mathbb{R}\langle n\rangle$ if and only if E is strongly even and its underlying spectrum is a form of $BP\langle n\rangle$. We want to show that every form of $BP\mathbb{R}\langle n\rangle$ is also of the form $BP\mathbb{R}/(\bar{v}_{n+1}, \bar{v}_{n+2}, \dots)$ for some choice of elements $\bar{v}_i$. For this, we need the following lemma from [27, Lemma 3.4]:

Lemma 4.11 *Let $f\colon E \to F$ be a map of C_2–spectra. Assume that f induces isomorphisms*

$$\pi_{k\rho}^{C_2} E \to \pi_{k\rho}^{C_2} E \quad \text{and} \quad \pi_k E \to \pi_k F$$

for all $k \in \mathbb{Z}$. Assume furthermore that $\pi_{k\rho-1}^{C_2} E \to \pi_{k\rho-1}^{C_2} F$ is an injection for all $k \in \mathbb{Z}$ (for example, if $\pi_{k\rho-1}^{C_2} E = 0$). Then f is an equivalence of C_2–spectra.

Proposition 4.12 *Let E be a form of $BP\mathbb{R}\langle n\rangle$. Then one can choose indecomposables $\bar{v}_i \in \pi^{C_2}_{(2^i-1)\rho} BP\mathbb{R}$ for $i \geq n+1$ such that $E \simeq BP\mathbb{R}/(\bar{v}_{n+1}, \bar{v}_{n+2}, \dots)$.*

Proof First choose any system of $\bar{v}_i$. Also choose a real orientation $f\colon BP\mathbb{R} \to E$ and denote $f(\bar{v}_i)$ by x_i. Define a multiplicative section

$$s\colon \pi^{C_2}_{*\rho} E \to \pi^{C_2}_{*\rho} BP\mathbb{R}$$

by $s(x_i) = \bar{v}_i$ for $1 \leq i \leq n$.

Now define a new system of $\bar{v}_i$ by

$$\bar{v}_i^{\mathrm{new}} = \bar{v}_i - s(f_*(\bar{v}_i))$$

for $i \geq n+1$. As these agree with $\bar{v}_i$ mod $(\bar{v}_1, \dots, \bar{v}_n)$, they are still indecomposable. Furthermore, the $\bar{v}_i^{\mathrm{new}}$ are for $i \geq n+1$ clearly in the kernel of f_*. Thus, we obtain a map $BP\mathbb{R}\langle n\rangle/(\bar{v}_{n+1}^{\mathrm{new}}, \bar{v}_{n+2}^{\mathrm{new}}, \dots) \to E$ that is an isomorphism on $\pi^{C_2}_{*\rho}$. By Corollary 4.6, the source is strongly even. By Lemma 4.11, the map is an equivalence. □

Examples 4.13 We consider real versions of the classical examples ku and $tmf_1(3)$.

(1) The connective real K-theory spectrum $k\mathbb{R}_{(2)}$ is a form of $BP\mathbb{R}\langle 1\rangle$. Indeed, the underlying spectrum $ku_{(2)}$ is well known to be a form of $BP\langle 1\rangle$ and $k\mathbb{R}_{(2)}$ is also strongly even (as can be seen by the results from [7, Section 3.7.D] or from the computation in Section 11).

(2) Define $\overline{tmf_1(3)}$ as the equivariant connective cover of the spectrum $\overline{Tmf_1(3)}$, ie $Tmf_1(3)$ with the algebro-geometrically defined C_2–action (see [27, Section 4.1] for details). As shown in [27, Corollary 4.17], $\overline{tmf_1(3)}_{(2)}$ is a form of $BP\mathbb{R}\langle 2\rangle$. By Proposition 4.12, we can construct $\overline{tmf_1(3)}_{(2)}$ by killing a sequence $\bar{v}_2, \bar{v}_3, \dots$ in $BP\mathbb{R}$. This construction is used in [23] to define a C_2–equivariant version of $tmf_1(3)_{(2)}$. In particular, we see (using the discussion before Proposition 4.23 in [27]) that $\overline{TMF_1(3)}_{(2)}$ (with the algebro-geometrically defined C_2–action) agrees with the $\mathbb{TMF}_1(3)_{(2)}$ of [23].

5 Results and consequences

In this section, we want to discuss our main results in more detail than in the introduction and we will also derive some consequences and give some examples. Recall to that purpose the notation from Sections 3C and 4A. Furthermore, we will implicitly localize everything at 2, so $\mathbb{Z}$ means $\mathbb{Z}_{(2)}$, etc. Our main theorem is the following:

Theorem 5.1 *Let $(m_1, m_2, \dots)$ be a sequence of nonnegative integers with only finitely many entries bigger than 1, and let M be the quotient $BP\mathbb{R}/(\bar{v}_1^{m_1}, \bar{v}_2^{m_2}, \dots)$, where we only quotient by the positive powers of $\bar{v}_i$. Denote by $\underline{\bar{v}}$ the sequence of $\bar{v}_i$ in $\pi_\star^{C_2} M\mathbb{R}$ such that $m_i = 0$, by $|\underline{\bar{v}}|$ the sum of their degrees and by m' the sum of all $(m_i - 1)|\bar{v}_i|$ for $m_i > 1$. Then*

$$\mathbb{Z}^M \simeq \Sigma^{-m'+4-2\rho} \kappa_{M\mathbb{R}}(\underline{\bar{v}}; M).$$

The most important case is where $m_{n+1} = m_{n+2} = \cdots = 1$, so

$$M = BP\mathbb{R}\langle n\rangle/(\bar{v}_1^{m_1}, \dots, \bar{v}_n^{m_n}).$$

If k is the number of elements in $\underline{\bar{v}}$, we also get

$$\mathbb{Z}^M \simeq \Sigma^{-m'+k+|\underline{\bar{v}}|+4-2\rho} \Gamma_{\underline{\bar{v}}} M,$$

where we view M as an $M\mathbb{R}$–module.

The first form will be proved as Theorem 10.1 and the second follows from it using Lemma 3.6. The second form also follows from Corollary 7.5 (using that $\Gamma_{\underline{\bar{v}}}$ preserves cofibre sequences to pass to quotients of $BP\mathbb{R}\langle n\rangle$).

Example 5.2 $\mathbb{Z}^{BP\mathbb{R}\langle n\rangle} \simeq \Sigma^{n+D_n\rho+4-2\rho} \Gamma_{(\bar{v}_1,\dots,\bar{v}_n)} BP\mathbb{R}\langle n\rangle$ for $D_n = |v_1| + \cdots + |v_n|$. This says $BP\mathbb{R}\langle n\rangle$ has Gorenstein duality with respect to $H\underline{\mathbb{Z}} \simeq BP\mathbb{R}\langle n\rangle/(\bar{v}_1, \dots, \bar{v}_n)$. (The last equivalence follows from Corollary 4.7.)

Example 5.3 Set $k\mathbb{R}(n) = BP\mathbb{R}\langle n\rangle/(\bar{v}_1, \dots, \bar{v}_{n-1})$ to be connective integral real Morava K-theory and $K\mathbb{R}(n) = k\mathbb{R}(n)[\bar{v}_n^{-1}]$ its periodic version. Then

$$\begin{aligned} \mathbb{Z}^{k\mathbb{R}(n)} &\simeq \Sigma^{1+|\bar{v}_n|+4-2\rho} \Gamma_{\bar{v}_n} k\mathbb{R}(n) \\ &\simeq \Sigma^{(2^n-3)\rho+4} \operatorname{cof}(k\mathbb{R}(n) \to K\mathbb{R}(n)). \end{aligned}$$

This includes for $n = 1$ the case of usual (2–local) connective real K-theory.

Example 5.4 To have a slightly stranger example, take $M = BP\mathbb{R}\langle 3\rangle/(\bar{v}_1^4, \bar{v}_3^2)$. Then

$$\mathbb{Z}^M \simeq \Sigma^{5-9\rho} \Gamma_{\bar{v}_2} M.$$

So far, we have only talked about *quotients* of $BP\mathbb{R}$. This does not include important real spectra like the real Johnson–Wilson theories $E\mathbb{R}(n) = BP\mathbb{R}\langle n\rangle[\bar{v}_n^{-1}]$ or the (integral) real Morava K-theories $K\mathbb{R}(n)$. For this, we have to study the behaviour of our constructions under localizations.

Let M be an $RO(C_2)$–graded $\mathbb{Z}[v]$–module, where v has some degree $|v| \in RO(C_2)$. We say that M has *bounded v–divisibility* if for every degree $a + b\sigma$, there is a k such that

$$v^k \colon M_{a+b\sigma-|v^k|} \to M_{a+b\sigma}$$

is zero. We will also apply the concept to modules that are just $\mathbb{Z}|v|$–graded.

Lemma 5.5 *The class of $RO(C_2)$–graded $\mathbb{Z}[v]$–modules of bounded v–divisibility is closed under submodules, quotients and extensions.*

Proof This is clear for submodules and quotients. Let

$$0 \to K \to M \to N \to 0$$

be a short exact sequence of $\mathbb{Z}[v]$–modules where K and N are of bounded v–divisibility. For a given degree $\alpha \in RO(C_2)$, we know that there is a k such that v^k maps trivially into K_α. Furthermore, there is an n such that v^n maps trivially into $N_{\alpha-k|v|}$. Thus, multiplication by v^{n+k} is the zero map $M_{\alpha-(k+n)|v|} \to M_\alpha$. □

Let M be an $M\mathbb{R}$–module. We say that M is of *bounded $\bar{v}_n$–divisibility* if both $\pi^{C_2}_\star M$ and $\pi^e_* M$ are of bounded $\bar{v}_n$–divisibility. This is, for example, true if M is connective.

Lemma 5.6 *We have the following two properties of $\bar{v}_n$–divisibility.*

1. *Being of bounded $\bar{v}_n$–divisibility is closed under cofibres and suspensions.*
2. *An $M\mathbb{R}$–module M is of bounded $\bar{v}_n$–divisibility if and only if $\pi^{C_2}_{*\rho} M$ and $\pi^e_* M$ are of bounded $\bar{v}_n$–divisibility.*

Proof Both statements follow from the last lemma. For the second item, we additionally use the exact sequence

$$\pi^e_{a+b+1} M \to \pi^{C_2}_{a+(b+1)\sigma} M \to \pi^{C_2}_{a+b\sigma} M \to \pi^e_{a+b} M$$

induced by the cofibre sequence

$$(C_2)_+ \to S^0 \to S^\sigma. \qquad \square$$

Lemma 5.7 *If M has bounded $\bar{v}_n$–divisibility, then there is a natural equivalence*

$$M[\bar{v}_n^{-1}] \simeq \Sigma \operatorname{holim}_{\leftarrow}(\cdots \to \Sigma^{|\bar{v}_n|} \Gamma_{\bar{v}_n} M \xrightarrow{\bar{v}_n} \Gamma_{\bar{v}_n} M)$$

of $M\mathbb{R}$–modules.

Proof We apply the endofunctor $H\colon N \mapsto \operatorname{holim}_{\leftarrow}(\cdots \to \Sigma^{|\bar{v}_n|}N \xrightarrow{\bar{v}_n} N)$ of $M\mathbb{R}$–modules to the cofibre sequence

$$\Gamma_{\bar{v}_n}M \to M \to M[\bar{v}_n^{-1}].$$

Clearly $H(M[\bar{v}_n^{-1}]) \simeq M[\bar{v}_n^{-1}]$. Thus, we just have to show that $H(M) \simeq 0$. This follows by the $\lim^1$–sequence and bounded $\bar{v}_n$–divisibility. □

Lemma 5.8 *Let B be a quotient of $BP\mathbb{R}$ by powers of the $\bar{v}_i$. Then $B[\bar{v}^{-1}]$ has bounded $\bar{v}_n$–divisibility if $\bar{v}$ is a product of $\bar{v}_i$ not containing $\bar{v}_n$. Hence, the same is also true for the stable Koszul complex $\Gamma_{\underline{\bar{v}}}B$, where $\underline{\bar{v}}$ is a sequence of $\bar{v}_i$ not containing $\bar{v}_n$.*

Proof By Lemma 5.6, it is enough to check the first statement on $\pi_{*\rho}^{C_2}$ and on π_*^e. On the latter, it is clear and the former is isomorphic to it by Corollary 4.6. For the second statement, we use that $\Gamma_{\underline{\bar{v}}}B$ is the fibre of $B \to \check{C}(\underline{\bar{v}}; B)$, where $\check{C}(\underline{\bar{v}}; B)$ has a filtration with subquotients $M\mathbb{R}$–modules of the form $\Sigma^? B[x^{-1}]$ for some $x \in \pi_\star^{C_2} M\mathbb{R}$ [11, Lemma 3.7]. Thus, the second statement follows from Lemma 5.6. □

Theorem 5.9 *Let the notation be as in Theorem 5.1, and assume for simplicity that only finitely many m_i are zero and that $m_n = 0$. Then*

$$\mathbb{Z}^{M[\bar{v}_n^{-1}]} \simeq \Sigma^{-m'+|\underline{\bar{v}}|+(k-1)+4-2\rho}\Gamma_{\underline{\bar{v}}\setminus\bar{v}_n}M.$$

Here $\underline{\bar{v}} \setminus \bar{v}_n$ denotes the sequence of all $\bar{v}_i$ such that $m_i = 0$ and $i \neq n$.

Proof The preceding lemmas imply the following chain of equivalences:

$$\begin{aligned}
\mathbb{Z}^{M[\bar{v}_n^{-1}]} &\simeq \mathbb{Z}^{\operatorname{holim}_{\to}(M \xrightarrow{\bar{v}_n} \Sigma^{-|\bar{v}_n|}M \xrightarrow{\bar{v}_n} \cdots)} \\
&\simeq \operatorname{holim}_{\leftarrow}(\cdots \xrightarrow{\bar{v}_n} \mathbb{Z}^M) \\
&\simeq \Sigma^{-m'+|\underline{\bar{v}}|+k+4-2\rho}\operatorname{holim}_{\leftarrow}(\cdots \xrightarrow{\bar{v}_n} \Gamma_{\underline{\bar{v}}}M) \\
&\simeq \Sigma^{-m'+|\underline{\bar{v}}|+k+4-2\rho}\operatorname{holim}_{\leftarrow}(\cdots \xrightarrow{\bar{v}_n} \Gamma_{\bar{v}_n}(\Gamma_{\underline{\bar{v}}\setminus\bar{v}_n}M)) \\
&\simeq \Sigma^{-m'+|\underline{\bar{v}}|+(k-1)+4-2\rho}(\Gamma_{\underline{\bar{v}}\setminus\bar{v}_n}M)[\bar{v}_n^{-1}] \\
&\simeq \Sigma^{-m'+|\underline{\bar{v}}|+(k-1)+4-2\rho}\Gamma_{\underline{\bar{v}}\setminus\bar{v}_n}(M[\bar{v}_n^{-1}]).
\end{aligned}$$
□

Example 5.10 We recover the following result by Ricka [28]:

$$\mathbb{Z}^{K\mathbb{R}(n)} \simeq \Sigma^{4-2\rho}K\mathbb{R}(n).$$

Here, $K\mathbb{R}(n)$ denotes integral Morava K-theory $E\mathbb{R}(n)/(\bar{v}_1, \ldots, \bar{v}_{n-1})$.

Example 5.11 In the following, we will use the fact that there are invertible classes $x, \bar{v}_n \in \pi_\star^{C_2} E\mathbb{R}(n)$ of degree $-2^{2n+1} + 2^{n+2} - \rho$ and $(2^n - 1)\rho$, respectively, where $x = \bar{v}_n^{1-2^n} u^{2^n(1-2^{n-1})}$:

$$\begin{aligned}\mathbb{Z}^{E\mathbb{R}(n)} &\simeq \Sigma^{D_{n-1}\rho+(n-1)+4-2\rho}\Gamma_{(\bar{v}_1,\dots,\bar{v}_{n-1})} E\mathbb{R}(n)\\ &\simeq \Sigma^{-(n+2)\rho+(n+3)}\Gamma_{(\bar{v}_1,\dots,\bar{v}_{n-1})} E\mathbb{R}(n)\\ &\simeq \Sigma^{(n+2)(2^{2n+1}-2^{n+2})+n+3}\Gamma_{(\bar{v}_1,\dots,\bar{v}_{n-1})} E\mathbb{R}(n).\end{aligned}$$

This says that $E\mathbb{R}(n)$ has Gorenstein duality with respect to $E\mathbb{R}(n)/(\bar{v}_1, \dots, \bar{v}_{n-1}) = K\mathbb{R}(n)$. Note that we can replace the ideal $(\bar{v}_1, \dots, \bar{v}_{n-1})$ by an ideal generated in integral degrees, namely $(\bar{v}_1 x, \dots, \bar{v}_{n-1} x^{2^{n-1}-1})$.

Example 5.12 Recall from [27] the spectra $tmf_1(3)$, $Tmf_1(3)$ and $TMF_1(3)$ and the corresponding C_2–spectra $\overline{tmf_1(3)}$, $\overline{Tmf_1(3)}$ and $\overline{TMF_1(3)}$. Recall that we have $\pi_* tmf_1(3) = \mathbb{Z}[a_1, a_3]$, where a_1 and a_3 can be identified with the images of the Hazewinkel generators v_1 and v_2, and that $\overline{tmf_1(3)}$ is a form of $BP\mathbb{R}\langle 2\rangle$ (as already discussed in Examples 4.13). This gives the Anderson dual of $\overline{tmf_1(3)}$. Tweaking the last theorem a little bit allows us also to show that

$$\mathbb{Z}^{\overline{TMF_1(3)}} \simeq \Sigma^{5+2\rho}\Gamma_{\bar{v}_1}\overline{TMF_1(3)}.$$

We can also recover one of the main results of [27], namely that

$$\mathbb{Z}^{\overline{Tmf_1(3)}} \simeq \Sigma^{5+2\rho}\overline{Tmf_1(3)}.$$

Indeed, $Tmf_1(3)$ is by [27, Section 4.3] the cofibre of the map

$$\Gamma_{\bar{v}_1,\bar{v}_2}\overline{tmf_1(3)} \to \overline{tmf_1(3)}.$$

As the source is equivalent to $\Sigma^{-6-2\rho}\mathbb{Z}^{\overline{tmf_1(3)}}$, applying Anderson duality shows that $\mathbb{Z}^{\overline{Tmf_1(3)}}$ is the fibre of

$$\Sigma^{6+2\rho}\overline{tmf_1(3)} \to \Sigma^{6+2\rho}\Gamma_{\bar{v}_1,\bar{v}_2}\overline{tmf_1(3)}.$$

This is equivalent to $\Sigma^{5+2\rho}\overline{Tmf_1(3)}$. This example does not require 2–localization, only that 3 is inverted.

Remark 5.13 By Proposition 3.3, all the results in this section have direct implications for the Anderson duals of the fixed point spectra. These are easiest to understand in the case of $ER(n) = (E\mathbb{R}(n))^{C_2}$, where we get

$$\mathbb{Z}^{ER(n)} \simeq \Sigma^{(n+2)(2^{2n+1}-2^{n+2})+n+3}\Gamma_{(\bar{v}_1 x,\dots,\bar{v}_{n-1}x^{2^n-1})} ER(n).$$

Part II The Gorenstein approach

In this part, we explain the Gorenstein approach to prove Gorenstein duality, first for $k\mathbb{R}$ and then for $BP\mathbb{R}\langle n\rangle$.

6 Connective K-theory with reality

The present section considers K-theory with reality, which is more familiar than $BP\mathbb{R}\langle n\rangle$ for general n, and no 2–localization is necessary. The arguments are especially simple, firstly because $k\mathbb{R}$ is a commutative ring spectrum, and secondly because we only need to consider principal ideals. Simple as the argument is, we see in Section 11 that the consequences for coefficient rings are interesting.

6A Gorenstein condition and Matlis lift

It is well known that there is a cofibre sequence

$$\Sigma^{\epsilon} ku \xrightarrow{v} ku \to H\mathbb{Z}.$$

If one knows the coefficient ring $ku_* = \mathbb{Z}[v]$, this is easy to construct, since we can identify ku/v as the Eilenberg–Mac Lane spectrum from its homotopy groups.

There is a version with reality [8]. Indeed, we may construct the cofibre sequence

$$\Sigma^{\rho} k\mathbb{R} \xrightarrow{\overline{v}} k\mathbb{R} \to H\underline{\mathbb{Z}},$$

where $k\mathbb{R}/\overline{v}$ is identified using Corollary 2.3

Since the Dugger sequence is self dual we immediately deduce that $k\mathbb{R}$ is Gorenstein.

Lemma 6.1 $$\mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, k\mathbb{R}) = \Sigma^{-\rho-1} H\underline{\mathbb{Z}},$$

and $k\mathbb{R} \to H\underline{\mathbb{Z}}$ is Gorenstein.

Proof Apply $\mathrm{Hom}_{k\mathbb{R}}(\,\cdot\,, k\mathbb{R})$ to the Dugger sequence. □

To actually get Gorenstein duality we need to construct a Matlis lift (adapted from [9, Section 6]), which is a counterpart in topology of the injective hull of the residue field.

Definition 6.2 If M is an $H\underline{\mathbb{Z}}$–module, we say that a $k\mathbb{R}$–module $\widetilde{M}$ is a *Matlis lift* of M if $\widetilde{M}$ is $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellular, and

$$\mathrm{Hom}_{k\mathbb{R}}(T, \widetilde{M}) \simeq \mathrm{Hom}_{H\underline{\mathbb{Z}}}(T, M)$$

for all $H\underline{\mathbb{Z}}$–modules T.

The Anderson dual provides one such example.

Lemma 6.3 *The $k\mathbb{R}$–module $\Sigma^{-2(1-\sigma)}\underline{\mathbb{Z}}^{k\mathbb{R}}$ is a Matlis lift of $H\underline{\mathbb{Z}}$. Indeed,*

(i) $\underline{\mathbb{Z}}^{k\mathbb{R}}$ *is $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellular, and*

(ii) *there is an equivalence*

$$\Sigma^{2\delta} H\underline{\mathbb{Z}} \simeq H\underline{\mathbb{Z}}^* = \mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, \mathbb{Z}^{k\mathbb{R}}),$$

where $\delta = 1-\sigma$.

Proof One could prove the first part from the slice tower, but it also follows directly from Corollary 3.9.

The second statement is immediate from Lemma 3.1. □

6B Gorenstein duality

We next want to move on to Gorenstein duality, so we write

$$\mathcal{E} = \mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, H\underline{\mathbb{Z}}).$$

Combining Lemmas 6.1 and 6.3, we have

$$\mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, k\mathbb{R}) \simeq \Sigma^{-\rho-1} H\underline{\mathbb{Z}} \simeq \mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, \Sigma^{-4+\sigma}\underline{\mathbb{Z}}^{k\mathbb{R}}). \tag{6.4}$$

We now want to remove the $\mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, \cdot)$ from this equivalence.

Lemma 6.5 (effective constructibility) *The evaluation map*

$$\mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, M) \otimes_{\mathcal{E}} H\underline{\mathbb{Z}} \to M$$

is an $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellularization for every left $k\mathbb{R}$–module M.

Proof Since the domain is clearly $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellular, it is enough to show the map is an equivalence for all cellular modules M.

This is clear for $M = H\underline{\mathbb{Z}}$. The class of M for which the statement is true is closed under (i) triangles, (ii) coproducts (since $H\underline{\mathbb{Z}}$ is small) and (iii) suspensions by representations. This gives all $\mathbb{R}$–cellular modules. □

Local cohomology gives an alternative approach to cellularization. Recall that we define the $\bar{v}$–power torsion of a $k\mathbb{R}$–module M by the fibre sequence

$$\Gamma_{\bar{v}} M \to M \to M[1/\bar{v}].$$

The following lemma is a special case of Proposition 3.8.

Lemma 6.6 *The map*

$$\Gamma_{\bar{v}} M \to M$$

is an $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellularization.

It remains to check that the two $\mathcal{E}$–actions on $H\underline{\mathbb{Z}}$ coincide.

Lemma 6.7 *There is a unique right $\mathcal{E}$–module structure on $H\underline{\mathbb{Z}}$.*

Proof Suppose that $H\underline{\mathbb{Z}}'$ is a right $\mathcal{E}$–module whose underlying C_2–spectrum is equivalent to the Eilenberg–Mac Lane spectrum $H\underline{\mathbb{Z}}$. We first claim that $H\underline{\mathbb{Z}}'$ can be constructed as an $\mathcal{E}$–module with cells in degrees $k\rho$ for $k \leq 0$:

$$H\underline{\mathbb{Z}}' \simeq_{\mathcal{E}} S^0_{\mathcal{E}} \cup e^{-\rho}_{\mathcal{E}} \cup e^{-2\rho}_{\mathcal{E}} \cup \cdots .$$

Once that is proved, we argue as follows. If $H\underline{\mathbb{Z}}''$ is another right $\mathcal{E}$–module with underlying C_2–spectrum $H\underline{\mathbb{Z}}$, we may construct a map $H\underline{\mathbb{Z}}' \to H\underline{\mathbb{Z}}''$ skeleton by skeleton in the usual way. We start with the $\mathcal{E}$–module map $\mathcal{E} = (H\underline{\mathbb{Z}}')^{(0)} \to H\underline{\mathbb{Z}}'$ giving the unit, and successively extend the map over the cells of $H\underline{\mathbb{Z}}'$. At each stage the obstruction to the existence of an extension over $(H\underline{\mathbb{Z}}')^{-k\rho}$ lies in $\pi^{C_2}_{-k\rho-1}(H\underline{\mathbb{Z}}'')$. These groups are zero. We end with a map which is an isomorphism on 0^{th} homotopy Mackey functors and therefore an equivalence.

For the cell-structure, it is enough to show that for every right $\mathcal{E}$–module $H\underline{\mathbb{Z}}'$ of the homotopy type of the Eilenberg–Mac Lane spectrum $H\underline{\mathbb{Z}}$, there is a map $\mathcal{E} \to H\underline{\mathbb{Z}}'$ of right $\mathcal{E}$–modules whose fibre has the homotopy type of $\Sigma^{-\rho-1} H\underline{\mathbb{Z}}$. Indeed, suppose we have already constructed a right $\mathcal{E}$–module $(H\underline{\mathbb{Z}}')^{(n)}$ with an $\mathcal{E}$–map to $H\underline{\mathbb{Z}}'$ with fibre of the homotopy type $\Sigma^{-(n+1)\rho-1} H\underline{\mathbb{Z}}$. Then it is easy to see that the cofibre $(H\underline{\mathbb{Z}}')^{(n+1)}$ of the map $\Sigma^{-(n+1)\rho-1}\mathcal{E} \to \Sigma^{-(n+1)\rho-1} H\underline{\mathbb{Z}} \to (H\underline{\mathbb{Z}}')^{(n)}$ has the analogous property. Taking the homotopy colimit, we get a map $\operatorname{holim}_{\to}(H\underline{\mathbb{Z}}')^{(n)} \to H\underline{\mathbb{Z}}'$ with fibre $\operatorname{holim}_{\to} \Sigma^{-(n+1)\rho-1} H\underline{\mathbb{Z}}$, which is clearly zero (eg by Lemma 4.11 and the fact that $H\underline{\mathbb{Z}}$ is even; we refer to [28, Section 3.4] for a table of $\underline{\pi}^{C_2}_{\star} H\underline{\mathbb{Z}}$).

We choose the map $f\colon \mathcal{E} \to H\underline{\mathbb{Z}}'$ representing $1 \in \pi^{C_2}_0 H\underline{\mathbb{Z}}'$ and call the fibre F. We want to show that f agrees with the canonical map $\mathcal{E} \to H\underline{\mathbb{Z}}$ on homotopy groups of the form $\pi^{C_2}_{k-\sigma}$ for $k \in \mathbb{Z}$. Indeed, the only nonzero class in $H\underline{\mathbb{Z}}'$ in these degrees is $a \in \pi^{C_2}_{-\sigma} H\underline{\mathbb{Z}}'$, which has to be hit by $a \in \pi^{C_2}_{-\sigma}\mathcal{E}$ as it comes from the sphere. Thus, $\pi^{C_2}_{k-\sigma} F \cong \pi^{C_2}_{k-\sigma} \Sigma^{-1-\rho} H\underline{\mathbb{Z}}$ for all k and hence $F \simeq \Sigma^{-1-\rho} H\underline{\mathbb{Z}}$ as C_2–spectra, as we needed to show. □

From this the required statement follows.

Corollary 6.8 (Gorenstein duality) *There is an equivalence of $k\mathbb{R}$–modules*

$$\Gamma_{\bar{v}} k\mathbb{R} \simeq \Sigma^{-4+\sigma} \mathbb{Z}^{k\mathbb{R}}.$$

Proof By (6.4) and Lemma 6.7, we know that

$$\mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, k\mathbb{R}) \otimes_{\mathcal{E}} H\underline{\mathbb{Z}} \simeq \mathrm{Hom}_{k\mathbb{R}}(H\underline{\mathbb{Z}}, \Sigma^{-4+\sigma}\mathbb{Z}^{k\mathbb{R}}) \otimes_{\mathcal{E}} H\underline{\mathbb{Z}}.$$

By Lemma 6.5, the two sides are the cellularizations of $k\mathbb{R}$ and $\Sigma^{-4+\sigma}\mathbb{Z}^{k\mathbb{R}}$ respectively. By Lemmas 6.6 and 6.3, the former is $\Gamma_{\bar{v}} k\mathbb{R}$ and the latter is $\Sigma^{-4+\sigma}\mathbb{Z}^{k\mathbb{R}}$ itself. $\square$

The implications of this equivalence for the coefficient ring are investigated in Section 11.

7 $BP\langle n\rangle$ with reality

We now turn to the case of $BP\mathbb{R}\langle n\rangle$ for a general n. The counterpart of the argument of Section 6 is a little simpler when $BP\mathbb{R}\langle n\rangle$ is a commutative ring spectrum. For $n=1$ and $n=2$, the spectra $k\mathbb{R}$, and $tmf_1(3)$, are both known to be a commutative ring spectra, and their 2–localizations give $BP\mathbb{R}\langle n\rangle$ when $n=1$ and $n=2$ respectively. However for higher n it is not known that $BP\mathbb{R}\langle n\rangle$ is a commutative ring spectrum. This is a significant technical issue, but one that is familiar when working with nonequivariant BP–related theories since BP is not known to be a commutative ring. The established method for getting around this is to use the fact that BP and $BP\langle n\rangle$ are modules over the commutative ring MU. We will adopt precisely the same method by working with $M\mathbb{R}$–modules. The only real complication is that we are forced to work with spectra whose homotopy groups are bigger than we might like, but if we focus on the relevant part, it causes no real difficulties.

7A Gorenstein condition and Matlis lift

As mentioned in the introduction of this section, we will work in the setting of $M\mathbb{R}$–modules. More precisely, we will always (implicitly) localize at 2 and set $S = M\mathbb{R}_{(2)}$. As discussed in Section 4A, we can define S–modules $BP\mathbb{R}\langle n\rangle$, once we have chosen a sequence of $\bar{v}_i$ (for example, the Hazewinkel or Araki generators).

The ideal

$$\bar{J}_n = (\bar{v}_1, \dots, \bar{v}_n)$$

plays a prominent role, and we will abuse notation by writing

$$S/\bar{J}_n := \mathrm{cof}(S \xrightarrow{\bar{v}_1} S) \otimes_S \mathrm{cof}(S \xrightarrow{\bar{v}_2} S) \otimes_S \cdots \otimes_S \mathrm{cof}(S \xrightarrow{\bar{v}_n} S),$$

and then

$$M/\bar{J}_n := M \otimes_S S/\bar{J}_n.$$

In particular,

$$BP\mathbb{R}\langle n\rangle/\bar{J}_n = BP\mathbb{R}\langle n\rangle/\bar{v}_n/\bar{v}_{n-1}/\cdots/\bar{v}_1 \simeq H\underline{\mathbb{Z}}$$

by the C_2–case of the reduction theorem, here proved as Corollary 4.7.

If $BP\mathbb{R}\langle n\rangle$ is a ring spectrum,

$$\mathrm{Hom}_{BP\mathbb{R}\langle n\rangle}(H\underline{\mathbb{Z}}, M) = \mathrm{Hom}_{BP\mathbb{R}\langle n\rangle}(BP\mathbb{R}\langle n\rangle \otimes_S S/\bar{J}_n, M) = \mathrm{Hom}_S(S/\bar{J}_n, M).$$

The right-hand side gives a way for us to express the fact that certain $BP\mathbb{R}\langle n\rangle$–modules (such as $BP\mathbb{R}\langle n\rangle$ and $\mathbb{Z}^{BP\mathbb{R}\langle n\rangle}$) are Matlis lifts, using only module structures over S.

Applying this when $M = BP\mathbb{R}\langle n\rangle$, we obtain the Gorenstein condition.

Lemma 7.1 *The map $BP\mathbb{R}\langle n\rangle \to H\underline{\mathbb{Z}}$ is Gorenstein of shift $-D_n\rho - n$ in the sense that*

$$\mathrm{Hom}_S(S/\bar{J}_n, BP\mathbb{R}\langle n\rangle) \simeq \Sigma^{-D_n\rho-n} H\underline{\mathbb{Z}},$$

where

$$D_n\rho = |\bar{v}_n| + |\bar{v}_{n-1}| + \cdots + |\bar{v}_1| = [2^{n+1} - n - 2]\rho.$$

Proof Since each of the maps $\bar{v}_i\colon \Sigma^{|\bar{v}_i|}S \to S$ is self-dual, for any S–module M, we have

$$\mathrm{Hom}_S(S/\bar{J}_n, M) \simeq \Sigma^{-D_m\rho-n} S/\bar{J}_n \otimes_S M. \qquad \square$$

Applying this when $M = \mathbb{Z}^{BP\mathbb{R}\langle n\rangle}$, we obtain the Anderson Matlis lift.

Lemma 7.2 *The Anderson dual of $BP\mathbb{R}\langle n\rangle$ is a Matlis lift of $H\underline{\mathbb{Z}}^*$ in the sense that*

(i) $\mathbb{Z}^{BP\mathbb{R}\langle n\rangle}$ *is $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellular, and*

(ii) *there is an equivalence*

$$\Sigma^{2-2\sigma} H\underline{\mathbb{Z}} \simeq H\underline{\mathbb{Z}}^* \simeq \mathrm{Hom}_S(S/\bar{J}_n, \mathbb{Z}^{BP\mathbb{R}\langle n\rangle}).$$

Proof One could prove the first part from the slice tower, but it also follows directly from Corollary 3.9.

For the second statement, observe that

$$\mathrm{Hom}_S(S/\bar{J}_n, \mathbb{Z}^{BP\mathbb{R}\langle n\rangle}) \simeq \mathrm{Hom}_S(S/\bar{J}_n \otimes_S BP\mathbb{R}\langle n\rangle, \mathbb{Z}^S) \simeq \mathbb{Z}^{H\underline{\mathbb{Z}}}.$$

Thus, Lemma 3.1 implies the statement. $\square$

7B Gorenstein duality

Throughout this section, we will write $R = BP\mathbb{R}\langle n\rangle$ for brevity. Combining Lemmas 7.1 and 7.2, we have an equivalence of S–modules

$$\mathrm{Hom}_S(S/\bar{J}_n, R) \simeq \Sigma^{-D_n\rho-n} H\underline{\mathbb{Z}} \simeq \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{-(D_n+n+2)-(D_n-2)\sigma}\mathbb{Z}^R).$$

We want to remove the $\mathrm{Hom}_S(S/\bar{J}_n, \cdot)$ from this equivalence. The endomorphism ring

$$\tilde{\mathcal{E}}_n = \mathrm{Hom}_S(S/\bar{J}_n, S/\bar{J}_n)$$

of the small S–module $S/\bar{J}_n$, replaces $\mathcal{E}_n = \mathrm{Hom}_R(H\underline{\mathbb{Z}}, H\underline{\mathbb{Z}})$ from the case that $R = BP\mathbb{R}\langle n\rangle$ is a ring spectrum. We note that

$$\tilde{\mathcal{E}}_n \otimes_S R = \mathrm{Hom}_S(S/\bar{J}_n, S/\bar{J}_n) \otimes_S R \simeq \mathrm{Hom}_S(S/\bar{J}_n, S/\bar{J}_n) \otimes_S R).$$

If $R = BP\mathbb{R}\langle n\rangle$ were a commutative ring, then this would be a ring equivalent to $\mathrm{Hom}_R(H\underline{\mathbb{Z}}, H\underline{\mathbb{Z}})$.

In any case, the following is proved exactly like Lemma 6.5.

Lemma 7.3 (effective constructibility) *The evaluation map*

$$\mathrm{Hom}_S(S/\bar{J}_n, M) \otimes_{\tilde{\mathcal{E}}_n} S/\bar{J}_n \to M$$

is an $S/\bar{J}_n$-$\mathbb{R}$–cellularization. □

Of course local cohomology gives an alternative approach to cellularization. Recall that we define

$$\Gamma_{\bar{J}_n} M = \Gamma_{\bar{v}_1} S \otimes_S \Gamma_{\bar{v}_2} S \otimes_S \cdots \otimes_S \Gamma_{\bar{v}_n} S \otimes_S M.$$

Then Proposition 3.8 gives the following lemma.

Lemma 7.4
$$\Gamma_{\bar{J}_n} M \to M$$
is an $H\underline{\mathbb{Z}}$-$\mathbb{R}$–cellularization.

It remains to check that the two $\tilde{\mathcal{E}}_n$ actions on $H\underline{\mathbb{Z}}$ coincide. For $k\mathbb{R}$ (ie $n = 1$), we showed there was a unique right $\mathcal{E}_n$–module structure on $H\underline{\mathbb{Z}}$. This may be true for $\tilde{\mathcal{E}}_n$–module structures, but we will instead just prove in the next subsection that the two particular $\tilde{\mathcal{E}}_n$–modules that arose from the left and right-hand ends of the first display of this subsection are equivalent.

The required Gorenstein duality statement follows. Its implications for the coefficient ring for $n = 2$ are investigated explicitly in Section 13.

Corollary 7.5 (Gorenstein duality) *There is an equivalence of $M\mathbb{R}$–modules*

$$\Gamma_{\bar{J}_n} R \simeq \Sigma^{-(D_n+n+2)-(D_n-2)\sigma} \underline{\mathbb{Z}}^R$$

with $R = BP\mathbb{R}\langle n\rangle$.

Proof We will argue in Section 7C that the equivalence

$$\mathrm{Hom}_S(S/\bar{J}_n, R) \simeq \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{-D_n\rho-n-2\delta}\underline{\mathbb{Z}}^R),$$

is in fact an equivalence of right modules over $\widetilde{\mathcal{E}}_n$. By Lemma 7.3, we see that R and $\Sigma^{-(D_n+n+2)-(D_n-2)\sigma}\underline{\mathbb{Z}}^R$ have equivalent $S/\bar{J}_n$ cellularizations. We have seen above that the cellularization of R is $\Gamma_{\bar{J}_n} BP\mathbb{R}\langle n\rangle$ and that $\Sigma^{-D_n\rho-n-2\delta}\underline{\mathbb{Z}}^R$ itself is cellular. □

7C The equivalence of induced and coinduced Matlis lifts of $H\underline{\mathbb{Z}}$

For brevity, we will still write $R = BP\mathbb{R}\langle n\rangle$, and note that we have a map $S = M\mathbb{R} \to BP\mathbb{R}\langle n\rangle = R$. The two S–modules that concern us are of a very special sort, one looks as if it is obtained from an S–module by "extension of scalars from S to R" and one looks as if it is obtained by "coextension of scalars from S to R".

Lemma 7.6 *We have equivalences of right $\widetilde{\mathcal{E}}_n$–modules*

$$\mathrm{Hom}_S(S/\bar{J}_n, R) \simeq \mathrm{Hom}_S(S/\bar{J}_n, S) \otimes_S R,$$
$$\mathrm{Hom}_S(S/\bar{J}_n, \underline{\mathbb{Z}}^R) = \mathrm{Hom}_S(R, \mathrm{Hom}_S(S/\bar{J}_n, \underline{\mathbb{Z}}^S)).$$

Proof The first equivalence is immediate from the smallness of $S/\bar{J}_n$.

The second equivalence is a consequence of the following equivalence of S–modules:

$$\underline{\mathbb{Z}}^R \simeq \mathrm{Hom}_S(R, \underline{\mathbb{Z}}^S). \qquad \square$$

Suspending the equivalences from Lemma 7.6 so that we are comparing two $\widetilde{\mathcal{E}}_n$–modules equivalent to $H\underline{\mathbb{Z}}$ (see Lemma 7.2), we have

$$Y_1 = \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{D_n\rho+n} R) \simeq \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{D_n\rho+n} S) \otimes_S R = X_1 \otimes_S R$$

and

$$Y_2 = \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{2\delta}\underline{\mathbb{Z}}^R) \simeq \mathrm{Hom}_S(S, \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{2\delta}\underline{\mathbb{Z}}^S)) = \mathrm{Hom}_S(R, X_2).$$

In Section 7D, we will construct an $\widetilde{\mathcal{E}}_n$–map $\alpha : X_1 \to Y_2$ and then argue in Section 7E that this extends along $X_1 = X_1 \otimes_S S \to X_1 \otimes_S R = Y_1$ to give a map $\tilde{\alpha}\colon Y_1 \to Y_2$

which is easily seen to be an equivalence: it is clearly a $*\rho-$ isomorphism and hence an equivalence by Lemma 4.11.

To see our strategy, note that the extension problem

$$\begin{tikzcd} X_1 \arrow[r, "\alpha"] \arrow[d] & \mathrm{Hom}_S(S/\bar{J}_n, \mathrm{Hom}_S(R, \mathbb{Z}^S)) \\ X_1 \otimes_S R \arrow[ur, dashed, "\tilde{\alpha}"'] & \end{tikzcd}$$

in the category of $\tilde{\mathcal{E}}_n$–modules is equivalent to the extension problem

$$\begin{tikzcd} X_1 \otimes_{\tilde{\mathcal{E}}_n} S/\bar{J}_n \otimes_S R \arrow[r, "\alpha'"] \arrow[d] & \mathbb{Z}^S \\ X_1 \otimes_{\tilde{\mathcal{E}}_n} S/\bar{J}_n \otimes_S R \otimes_S R \arrow[ur, dashed, "\tilde{\alpha}'"'] & \end{tikzcd}$$

in the category of S–modules. The point is that by the defining property of the Anderson dual, this latter extension problem can be tackled by looking in $\pi_0^{C_2}$. The 0^{th} homotopy groups of the spectra on the left are easily calculated from the known ring $\pi_\star^{C_2}(H\underline{\mathbb{Z}})$.

7D Construction of the map α

We construct the map α using a similar method as in the proof of Lemma 6.7.

Lemma 7.7 *There is a map*

$$\alpha\colon X_1 \to Y_2$$

of right $\tilde{\mathcal{E}}_n$–modules that takes the image of $1 \in \pi_0^{C_2}(S)$ to a generator of $\pi_0^{C_2}(H\underline{\mathbb{Z}}) = \mathbb{Z}$.

Proof First we claim that X_1 has a $\tilde{\mathcal{E}}_n$–cell structures with one 0–cell and other cells in dimensions which are negative multiples of ρ. More precisely, there is a filtration

$$\tilde{\mathcal{E}}_n \simeq X_1^{[0]} \to X_1^{[1]} \to X_1^{[2]} \to \cdots \to X_1$$

such that $X_1 \simeq \underrightarrow{\mathrm{holim}}_d X_1^{[d]}$, and there are cofibre sequences

$$X_1^{[d-1]} \to X_1^{[d]} \to \bigvee \Sigma^{-d\rho} \tilde{\mathcal{E}}_n.$$

By definition, $X_1 = \mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{D_n\rho+n} S)$. By Proposition 3.8 and Lemma 3.6, this is equivalent to

$$\mathrm{Hom}_S(S/\bar{J}_n, \Sigma^{D_n\rho+n} \Gamma_{\bar{J}_n} S) \simeq \mathrm{Hom}_S(S/\bar{J}_n, \kappa_S(\bar{v}_1, \ldots, \bar{v}_n))$$

because $\Gamma_{\bar{J}_n} S \to S$ is an $S/\bar{J}_n$-$\mathbb{R}$–cellularization. The usual construction of the stable Koszul complex from the unstable Koszul complex recalled in Section 3C, shows that

$$\kappa_S(\bar{v}_1, \ldots, \bar{v}_n)$$

has a filtration with subquotients sums of $(-k\rho)$–fold suspensions of $S/\bar{J}_n$. This induces a corresponding filtration on X_1.

As in Lemma 6.7 we may construct α by obstruction theory. Indeed, we start by choosing a map $\widetilde{\mathcal{E}}_n = X_1^{[0]} \to Y_2^{[0]}$ taking the unit to a generator. At the d^{th} stage we have a problem:

$$\begin{tikzcd} X_1^{[d-1]} \arrow[r] \arrow[d] & Y_2 \\ X_1^{[d]} \arrow[ur, dashed] & \end{tikzcd}$$

The obstruction to extension is in a finite product of groups

$$[\Sigma^{-d\rho-1}\widetilde{\mathcal{E}}_n, Y_2]^{\widetilde{\mathcal{E}}_n} = \pi^{C_2}_{-d\rho-1}(H\underline{\mathbb{Z}}) = 0,$$

where the vanishing is from the known value of $\pi^{C_2}_\star(H\underline{\mathbb{Z}})$. □

7E The map $\widetilde{\alpha}$

Referring to the second extension problem diagram above, we note $S/\bar{J}_n \otimes_S R \simeq H\underline{\mathbb{Z}}$ as S–modules. Thus, we have to solve the lifting problem

$$\begin{tikzcd} X_1 \otimes_{\widetilde{\mathcal{E}}_n} H\underline{\mathbb{Z}} \otimes_S S \arrow[r, "\alpha'"] \arrow[d, "1\otimes 1\otimes \pi"'] & \mathbb{Z}^S \\ X_1 \otimes_{\widetilde{\mathcal{E}}_n} H\underline{\mathbb{Z}} \otimes_S R \arrow[ur, dashed, "\widetilde{\alpha}'"'] & \end{tikzcd}$$

where $H\underline{\mathbb{Z}}$ is equipped with some $\widetilde{\mathcal{E}}_n$–module structure. Denote the upper left corner by T. The map $T \to T \otimes_S R$ is a split inclusion on underlying MU–modules. Indeed,

$$T \simeq X_1 \otimes_{\widetilde{\mathcal{E}}_n} S/\bar{J}_n \otimes_S R,$$

and the map $R \to R \otimes_S R$ is a split inclusion on underlying spectra because $BP\langle n\rangle$ has the structure of a homotopy unital MU–algebra [10, Theorem V.2.6].

By the definition of Anderson duals, we have a diagram of short exact sequences:

$$\begin{tikzcd}
0 \arrow[r] & \mathrm{Ext}^1_{\mathbb{Z}}(\pi^{C_2}_{-1}(T\otimes_S R),\mathbb{Z}) \arrow[r] \arrow[d] & {[T\otimes_S R,\mathbb{Z}^S]^S} \arrow[r] \arrow[d] & \mathrm{Hom}_{\mathbb{Z}}(\pi^{C_2}_0(T\otimes_S R),\mathbb{Z}) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathrm{Ext}^1_{\mathbb{Z}}(\pi^{C_2}_{-1}(T),\mathbb{Z}) \arrow[r] & {[T,\mathbb{Z}^S]^S} \arrow[r] & \mathrm{Hom}_{\mathbb{Z}}(\pi^{C_2}_0(T),\mathbb{Z}) \arrow[r] & 0
\end{tikzcd}$$

We want to show that the maps $\pi^{C_2}_k T \to \pi^{C_2}_k T\otimes_S R$ are split injections for $k=0,-1$, which solves the problem. For the computation of $\pi^{C_2}_* T$, recall from the last section that X_1 has a filtration starting with $X_1^{[x]} = \widetilde{\mathcal{E}}_n$ and with subquotients sums of terms of the form $\Sigma^{-d\rho}\widetilde{\mathcal{E}}_n$. Thus, T obtains a filtration starting with $T^{[1]} = H\underline{\mathbb{Z}}$ and with subquotients sums of terms of the form $\Sigma^{-d\rho}H\underline{\mathbb{Z}}$. The map $H\underline{\mathbb{Z}} = T^{[1]} \to T$ clearly induces isomorphisms on $\underline{\pi}^{C_2}_k$ for $k=0,-1$ by the known homotopy groups of $H\underline{\mathbb{Z}}$; see eg [28, Section 3.4] for a table. Thus, $\underline{\pi}^{C_2}_{-1}T = 0$ and $\underline{\pi}^{C_2}_0 T = \underline{\mathbb{Z}}$.

If we have a map $\underline{\mathbb{Z}} \to M$ from the constant Mackey functor, it is a split injection on (C_2/C_2) if it is one on (C_2/e). But we have already seen above that on underlying spectra $T \to T\otimes_S R$ is a split inclusion. Thus, we have shown that $\pi^{C_2}_k T \to \pi^{C_2}_k(T\otimes_S R)$ is split injective, which provides the map $\widetilde{\alpha}'$.

Part III The hands-on approach

In this part, we give a different way to compute the Anderson dual of $BP\mathbb{R}\langle n\rangle$ by first computing the Anderson dual of $BP\mathbb{R}$ itself. Again, we will first do the case of $k\mathbb{R}$.

8 The case of $k\mathbb{R}$ again

To illustrate our strategy, we give an alternative calculation of the Anderson dual of $k\mathbb{R}$. This can also be deduced from our main theorem below, but it might be helpful to see the proof in this simpler case first. General references for the $RO(C_2)$–graded homotopy groups of $k\mathbb{R}$ are [7, Section 3.7] or Section 11B.

We want to show the following proposition:

Proposition 8.1 *There is an equivalence $\kappa_{k\mathbb{R}}(\bar{v}) \to \Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}$.*

Recall here that $\bar{v} \in \pi^{C_2}_\rho k\mathbb{R}$ is the Bott element for real K-theory, and

$$\kappa_{k\mathbb{R}}(\bar{v}) = \operatorname*{hocolim}_n \Sigma^{-(n-1)\rho} k\mathbb{R}/\bar{v}^n.$$

Our idea is simple: to get a map from the homotopy colimit, we have just to give maps

$$\Sigma^{-(n-1)\rho} k\mathbb{R}/\bar{v}^n \to \Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}$$

that are compatible in the homotopy category (see Remark 3.7). We will show in the next lemma that these maps are essentially unique: the Mackey functor of homotopy classes of $k\mathbb{R}$–linear maps $\Sigma^{-(n-1)\rho}k\mathbb{R}/\bar{v}^n \to \Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}$ is isomorphic to $\underline{\mathbb{Z}}$ and the precomposition with the map $\Sigma^{-(n-1)\rho}k\mathbb{R}/\bar{v}^n \to \Sigma^{-n\rho}k\mathbb{R}/\bar{v}^{n+1}$ induces the identity on $\underline{\mathbb{Z}}$. Choosing the C_2–equivariant map $\kappa_{k\mathbb{R}}(\bar{v}) \to \Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}$ that corresponds to $1 \in \mathbb{Z}$ for every n induces an equivalence on underlying homotopy groups. By Lemma 4.11, the result follows as soon as we have established that $\kappa_{k\mathbb{R}}(\bar{v})$ is strongly even and that the Mackey functor $\underline{\pi}_{*\rho}\Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}$ is constant. These two facts will also be shown in the following lemma, finishing the proof of the proposition.

Lemma 8.2 *For a $\mathbb{Z}[\bar{v}]$ module M, denote by $\{M\}_{\bar{v}^n}$ its $\bar{v}^n$–torsion. Then we have:*

(1) *$k\mathbb{R}/\bar{v}^n$ is strongly even, and hence the same is true for $\kappa_{k\mathbb{R}}(\bar{v})$.*

(2) *$\underline{\pi}^{C_2}_{n\rho}\Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}} \cong \underline{\pi}^{C_2}_{(n-2)\rho+4}\mathbb{Z}^{k\mathbb{R}}$ is constant for all $n \in \mathbb{Z}$.*

(3) *$\underline{[\Sigma^{-(n-1)\rho}k\mathbb{R}/\bar{v}^n, \Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}]}^{C_2}_{k\mathbb{R}} \cong \{\underline{\pi}^{C_2}_{-(n-1)\rho}\Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}\}_{\bar{v}^n} \cong \underline{\mathbb{Z}}$.*

Proof The first part follows as

$$\underline{\pi}^{C_2}_{k\rho-i}(k\mathbb{R}/\bar{v}^n) = \underline{\pi}^{C_2}_{k\rho-i}(k\mathbb{R})/\bar{v}^n$$

for $i = 0, 1$ because $\pi^{C_2}_{k\rho-i}k\mathbb{R} = 0$ for $i = 1, 2$.

For the second part, consider the short exact sequence

$$0 \to \operatorname{Ext}(\underline{\pi}^{C_2}_{k\rho-5}k\mathbb{R}, \mathbb{Z}) \to \underline{\pi}^{C_2}_{-k\rho+4}\mathbb{Z}^{k\mathbb{R}} \to \operatorname{Hom}(\underline{\pi}^{C_2}_{k\rho-4}k\mathbb{R}, \mathbb{Z}) \to 0.$$

We have $\underline{\pi}^{C_2}_{k\rho-5}k\mathbb{R} = 0$ for all $k \in \mathbb{Z}$. For $k < 2$, the Mackey functor $\underline{\pi}^{C_2}_{k\rho-4}k\mathbb{R}$ vanishes as well and for $k \geq 2$, we have $\underline{\pi}^{C_2}_{k\rho-4}k\mathbb{R} \cong \underline{\mathbb{Z}}^*$, generated by v^{k-2} and $2\bar{v}^{k-2}u$. Thus,

$$\underline{\pi}^{C_2}_{-k\rho+4}\mathbb{Z}^{k\mathbb{R}} \cong \begin{cases} 0 & \text{if } k < 2, \\ \underline{\mathbb{Z}} & \text{if } k \leq 2. \end{cases}$$

This shows part (2). As multiplication by $\bar{v}^n$ does not hit $\underline{\pi}^{C_2}_{(n+1)\rho-4}k\mathbb{R}$, the whole Mackey functor $\underline{\pi}^{C_2}_{-(n+1)\rho+4}\mathbb{Z}^{k\mathbb{R}}$ is $\bar{v}^n$–torsion. This gives the second isomorphism of the third part.

For the remaining isomorphism, note that the cofibre sequence

$$\Sigma^{\rho}k\mathbb{R} \xrightarrow{\bar{v}^n} \Sigma^{-(n-1)\rho}k\mathbb{R} \to \Sigma^{-(n-1)\rho}k\mathbb{R}/\bar{v}^n \to \Sigma^{\rho+1}k\mathbb{R}$$

induces a short exact sequence

$$0 \to (\underline{\pi}^{C_2}_{\rho+1}\Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}})/\bar{v}_n \to \underline{[\Sigma^{-(n-1)\rho}k\mathbb{R}/\bar{v}^n, \Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}]}^{C_2}_{k\mathbb{R}} \to \{\underline{\pi}^{C_2}_{-(n-1)\rho}\Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}}\}_{\bar{v}^n} \to 0.$$

We have $\underline{\pi}^{C_2}_{\rho+1}\Sigma^{2\rho-4}\mathbb{Z}^{k\mathbb{R}} \cong \underline{\pi}^{C_2}_{5-\rho}\mathbb{Z}^{k\mathbb{R}}$, which sits in a short exact sequence

$$0 \to \mathrm{Ext}_{\mathbb{Z}}(\underline{\pi}^{C_2}_{\rho-6}k\mathbb{R}, \mathbb{Z}) \to \underline{\pi}^{C_2}_{5-\rho}\mathbb{Z}^{k\mathbb{R}} \to \mathrm{Hom}_{\mathbb{Z}}(\underline{\pi}^{C_2}_{\rho-5}k\mathbb{R}, \mathbb{Z}) \to 0.$$

But because of connectivity, $\underline{\pi}^{C_2}_{\rho-c}k\mathbb{R} = 0$ for $c \geq 3$. □

9 Duality for $BP\mathbb{R}$

We will use throughout the abbreviation $B = BP\mathbb{R}$ and will furthermore implicitly localize everything at 2, so $\mathbb{Z} = \mathbb{Z}_{(2)}$ etc, and all Hom and Ext groups are over $\mathbb{Z} = \mathbb{Z}_{(2)}$ unless marked otherwise. Denote by $\underline{\bar{v}}$ a sequence of indecomposable elements $\bar{v}_i \in \pi^{C_2}_{(2^i-1)\rho}B$. The aim of this section is to show that $\Sigma^{2\rho-4}\mathbb{Z}^B \simeq \kappa_{M\mathbb{R}}(\underline{\bar{v}}; B)$.

Recall that $\kappa_{M\mathbb{R}}(\underline{\bar{v}}; B)$ is defined as follows: Given a sequence $\underline{l} = (l_1, l_2, \dots)$ with $l_i \geq 0$, we denote by $B/\underline{\bar{v}}^{\underline{l}}$ the spectrum $B/(\bar{v}_{i_1}^{l_{i_1}}, \bar{v}_{i_2}^{l_{i_2}}, \dots)$, where i_j runs over all indices such that $l_{i_j} > 0$. Set

$$|\underline{l}| = l_1|\bar{v}_1| + l_2|\bar{v}_2| + \cdots .$$

Then

$$\kappa_{M\mathbb{R}}(\underline{\bar{v}}; B) = \operatorname*{hocolim}_{\underline{l}} \Sigma^{-|\underline{l}-\underline{1}|}B/\underline{\bar{v}}^{\underline{l}},$$

where $\underline{l}$ runs over all sequences such that all but finitely many l_i are zero, and $\underline{1}$ denotes the constant sequence of ones. Furthermore, the i^{th} entry of $\underline{l}-\underline{1}$ is defined to be the maximum of 0 and $l_i - 1$.

Thus, to get a map $\kappa_{M\mathbb{R}}(\underline{\bar{v}}; B) \to \Sigma^{2\rho-4}\mathbb{Z}^B$, we have to understand the homotopy classes of maps $B/\underline{\bar{v}}^{\underline{l}} \to \Sigma^{2\rho-4}\mathbb{Z}^B$. This will be the content of the next subsection.

9A Preparation

Recall the Mackey functor $\underline{\mathbb{Z}}^*$ defined by

$$\underline{\mathbb{Z}}^*(C_2/C_2) \cong \underline{\mathbb{Z}}^*(C_2/e) \cong \mathbb{Z}$$

with transfer equalling 1 while restriction is multiplication by 2.

Lemma 9.1 *As $\underline{\mathbb{Z}}[\bar{v}_1, \bar{v}_2, \dots]$–modules, we have the following isomorphisms:*

(1) $\underline{\pi}^{C_2}_{*\rho-4}B \cong \underline{\mathbb{Z}}^* \otimes_{\mathbb{Z}} \mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]$, *where $\underline{\mathbb{Z}}^*$ is generated by 1 on underlying and by $2u^{-1}$ on C_2–equivariant homotopy groups.*

(2) $\underline{\pi}^{C_2}_{*\rho-5}B = 0$.

(3) $\pi^{C_2}_{*\rho-6}B \cong \mathbb{F}_2\{a^2\bar{v}_1(-1)\} \otimes_{\mathbb{Z}} \mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]$.

Proof By Theorem A.4, the groups $\pi^{C_2}_{*\rho-c}B$ are additively generated by nonzero elements of the form $x = a^l\bar{v}$ with $\bar{v}$ a monomial in the $\bar{v}_i(j)$. Let $\bar{v}_i(j)$ be the one occurring with minimal i, where j is chosen such that $\bar{v} = \bar{v}_i(j)\bar{v}'$ with $\bar{v}'$ a monomial in the $\bar{v}_k$ (this is possible by the third relation in Theorem A.4). Then $|x| = *\rho + j2^{i+2} + l$ and $0 \le l < 2^{i+1} - 1$.

For $c = 4$, this implies $j = -1$, $i = 0$ and $l = 0$. Thus, x is of the form $\bar{v}_0(-1)\bar{v}'$. As the restriction of $\bar{v}_0(-1)$ to $\pi^e_0 B$ equals 2, the result follows.

For $c = 5$, we must have $l \ge 2^{i+2} - 5$, which implies $l \ge 2^{i+1} - 1$ or $i = 0$; in the latter case l must be zero, which is not possible.

For $c = 6$, we must have $l = -j2^{i+2} - 6$, which implies $l \ge 2^{i+1} - 1$ or $i \le 1$ and $j = -1$. As $i = 0$ is again not possible, $x = a^2\bar{v}_1(-1)\bar{v}'$ with $\bar{v}' \in \pi^{C_2}_{*\rho}$. $\square$

Lemma 9.2 *For a sequence $\underline{l} = (l_1, l_2, \dots)$, the map*

$$\underline{\pi}^{C_2}_{*\rho+4}\mathbb{Z}^{B/\underline{\bar{v}}^{\underline{l}}} \to \mathrm{Hom}(\underline{\pi}^{C_2}_{-*\rho-4}B/\underline{\bar{v}}^{\underline{l}}, \mathbb{Z}) \cong \underline{\mathbb{Z}} \otimes_{\mathbb{Z}} (\mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]/\underline{\bar{v}}^{\underline{l}})^*$$

is an isomorphism, where $\mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]^* = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots], \mathbb{Z})$ (so that the gradings become nonpositive). Here, the second map is the dual of the map

$$\underline{\mathbb{Z}}^* \otimes_{\mathbb{Z}} \mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]/\underline{\bar{v}}^{\underline{l}} \to \underline{\pi}^{C_2}_{-*\rho-4}B/\underline{\bar{v}}^{\underline{l}}$$

sending $1 \in \underline{\mathbb{Z}}^*(C_2/C_2)$ to the image of u^{-1} under the map $B \to B/\underline{\bar{v}}^{\underline{l}}$ and sending $1 \in \underline{\mathbb{Z}}^*(C_2/e)$ to 1.

Proof We have a short exact sequence

$$0 \to \mathrm{Ext}(\underline{\pi}^{C_2}_{-*\rho-5}B/\underline{\bar{v}}^{\underline{l}}, \mathbb{Z}) \to \underline{\pi}^{C_2}_{*\rho-4}\mathbb{Z}^{B/\underline{\bar{v}}^{\underline{l}}} \to \mathrm{Hom}(\underline{\pi}^{C_2}_{-*\rho-4}B/\underline{\bar{v}}^{\underline{l}}, \mathbb{Z}) \to 0.$$

If $l_1 = 0$, then Corollary 4.5 and Lemma 9.1 directly imply the statement. If $l_1 \neq 0$, Corollary 4.5 only allows us to identify the homotopy Mackey functor in degree $-*\rho-4$, but not the one in degree $-*\rho-5$. We give a separate argument in this case.

If $l_1 \neq 0$, consider the sequence $\underline{l}' = (0, l_2, l_3, \dots)$ and the corresponding cofibre sequence

$$\Sigma^{l_1\rho}B/\underline{\bar{v}}^{\underline{l}'} \xrightarrow{\bar{v}_1^{l_1}} B/\underline{\bar{v}}^{\underline{l}'} \to B/\underline{\bar{v}}^{\underline{l}} \to \Sigma^{l_1\rho+1}B/\underline{\bar{v}}^{\underline{l}'}.$$

This induces a short exact sequence

$$0 \to (\underline{\pi}^{C_2}_{*\rho-5}B/\underline{\bar{v}}^{\underline{l}'})/\bar{v}_1^{l_1} \to \underline{\pi}^{C_2}_{*\rho-5}B/\underline{\bar{v}}^{\underline{l}} \to \{\underline{\pi}^{C_2}_{*\rho-6}B/\underline{\bar{v}}^{\underline{l}'}\}_{\bar{v}_1^{l_1}} \to 0.$$

Here the last term denotes the Mackey subfunctor of $\underline{\pi}^{C_2}_{*\rho-6}B/\underline{\bar{v}}^{\underline{l}'}$ killed by $\bar{v}_1^{l_1}$. By Corollary 4.5 and Lemma 9.1, we see that $\underline{\pi}^{C_2}_{*\rho-5}B/\underline{\bar{v}}^{\underline{l}} = 0$. $\square$

As $B = BP\mathbb{R}$ is not known to have an E_∞–structure, we have to work with $M\mathbb{R}$–linear maps instead, for which the following lemma is useful:

Lemma 9.3 *The map*

$$\mathbb{Z}^B \simeq \mathrm{Hom}_{M\mathbb{R}}(M\mathbb{R}, \mathbb{Z}^B) \to \mathrm{Hom}_{M\mathbb{R}}(B, \mathbb{Z}^B)$$

is an equivalence.

Proof Let $e\colon M\mathbb{R} \to M\mathbb{R}$ be the Quillen–Araki idempotent. Recall that

$$B = \mathrm{hocolim}(M\mathbb{R} \xrightarrow{e} M\mathbb{R} \xrightarrow{e} \cdots).$$

Thus,

$$\mathbb{Z}^B \simeq \underleftarrow{\mathrm{holim}}(\cdots \xrightarrow{e^*} \mathbb{Z}^{M\mathbb{R}} \xrightarrow{e^*} \mathbb{Z}^{M\mathbb{R}}).$$

Hence,

$$\mathrm{Hom}_{M\mathbb{R}}(B, \mathbb{Z}^B) \simeq \underleftarrow{\mathrm{holim}}\big(\cdots \xrightarrow{e^*} \mathrm{Hom}_{M\mathbb{R}}(B, \mathbb{Z}^{M\mathbb{R}}) \xrightarrow{e^*} \mathrm{Hom}_{M\mathbb{R}}(B, \mathbb{Z}^{M\mathbb{R}})\big).$$

As every $\mathrm{Hom}_{M\mathbb{R}}(B, \mathbb{Z}^{M\mathbb{R}})$ is equivalent to a holim over

$$\mathrm{Hom}_{M\mathbb{R}}(M\mathbb{R}, \mathbb{Z}^{M\mathbb{R}}) \simeq \mathbb{Z}^{M\mathbb{R}}$$

connected by e^*, we get that $\mathrm{Hom}_{M\mathbb{R}}(B, \mathbb{Z}^B)$ is the homotopy limit $\mathrm{holim}_{\mathbb{Z}^- \times \mathbb{Z}^-} \mathbb{Z}^{M\mathbb{R}}$, where $\mathbb{Z}^-$ denotes the poset of negative numbers and all connecting maps are e^*. This is equivalent to the homotopy limit indexed over the diagonal, which in turn is equivalent to the homotopy limit indexed over a vertical. $\square$

Recall that we want to show that $X = \Sigma^{2\rho-4}\mathbb{Z}^B$ is equivalent to $\kappa_{M\mathbb{R}}(\underline{\bar{v}}, B)$. The reason for the choice of suspension is essentially (as before) that $H\underline{\mathbb{Z}} \simeq \Sigma^{2\rho-4} H\underline{\mathbb{Z}}^*$.

Proposition 9.4 *For a sequence $\underline{l} = (l_1, l_2, \dots)$, we have an isomorphism*

$$\underline{[\Sigma^{*\rho} B/\underline{\bar{v}}^{\underline{l}}, X]^{C_2}_{M\mathbb{R}}} \cong \underline{\mathbb{Z}} \otimes_{\mathbb{Z}} (\mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]/\underline{\bar{v}}^{\underline{l}})^*,$$

natural with respect to the maps $B/\underline{\bar{v}}^{\underline{l}} \to \Sigma^{-|\underline{l}'-\underline{l}|\rho} B/\underline{\bar{v}}^{\underline{l}'}$ in the defining homotopy colimit for $\kappa_{M\mathbb{R}}(\underline{\bar{v}}; B)$ for $\underline{l}' = (l_1', l_2', \dots)$ a sequence with $l_i' \geq l_i$ for all $i \geq 1$.

Proof The last lemma implies that we also have

$$\mathbb{Z}^{B/\underline{\bar{v}}^{\underline{l}}} \simeq \mathrm{Hom}_{M\mathbb{R}}(B/\underline{\bar{v}}^{\underline{l}}, \mathbb{Z}^B)$$

as the functors $\mathbb{Z}^?$ and $\mathrm{Hom}_{M\mathbb{R}}(?, \mathbb{Z}^B)$ behave the same way with respect to cofibre sequences and (filtered) homotopy colimits. Then we just have to apply Lemma 9.2. $\square$

9B The theorem

We first describe the homotopy groups of $X = \Sigma^{2\rho-4}\mathbb{Z}^B$ with $B = BP\mathbb{R}$ as before. By Lemma 9.2, we get

$$\underline{\pi}^{C_2}_{*\rho}X \cong \mathrm{Hom}(\underline{\pi}^{C_2}_{(*+2)\rho-4}B, \mathbb{Z}) \cong \underline{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{Z}[\bar{v}_1, \bar{v}_2, \dots]^*.$$

Let $\underline{l}$ be a sequence with only finitely many nonzero entries. By Proposition 9.4, the element $(\underline{\bar{v}}^{\underline{l}-\underline{1}})^*$ induces a corresponding $M\mathbb{R}$–linear map $\Sigma^{-|\underline{l}-\underline{1}|}B/\underline{\bar{v}}^{\underline{l}} \to X$, which is unique up to homotopy. By this uniqueness, these maps are also compatible for comparable $\underline{l}$. By Remark 3.7, this induces a map

$$\kappa_{M\mathbb{R}}(\underline{\bar{v}}, B) = \operatorname*{hocolim}_{\underline{l}}(\Sigma^{-|\underline{l}-\underline{1}|}B/\underline{\bar{v}}^{\underline{l}}) \xrightarrow{h} X,$$

where $\underline{l}$ ranges over all sequences where only finitely many l_i are nonzero.

Theorem 9.5 *The map $h\colon \kappa_{M\mathbb{R}}(\underline{\bar{v}}; B) \to X$ is an equivalence of C_2–spectra.*

Proof By Corollary 4.6, we get on $\underline{\pi}_{*\rho}$–level

$$\operatorname*{colim}_{\underline{l}} \Sigma^{-|\underline{l}-\underline{1}|}\underline{\mathbb{Z}}[\bar{v}_1, \bar{v}_2, \dots]/(\bar{v}_1^{l_1}, \dots) \to \underline{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{Z}[\bar{v}_1, \dots]^*,$$

which is an isomorphism. The odd underlying homotopy groups of both sides are zero. To apply Lemma 4.11, it is left to show that $\pi^{C_2}_{k\rho-1}\kappa_{M\mathbb{R}}(\underline{\bar{v}}; B) = 0$ for all $k \in \mathbb{Z}$. Again by Corollary 4.6, it is even true that $\pi^{C_2}_{k\rho-1}(B/\underline{\bar{v}}^{\underline{l}})$ is zero for all $k \in \mathbb{Z}$ and all sequences $\underline{l}$. □

10 Duality for regular quotients

The goal of this section is to prove our main result Theorem 5.1:

Theorem 10.1 *Let $(m_1, m_2, \dots)$ be a sequence of nonnegative integers with only finitely many entries bigger than 1. Denote by $\underline{\bar{v}}'$ the sequence of $\bar{v}_i$ in $\pi^{C_2}_{\star}M\mathbb{R}$ such that $m_i = 0$ and by m' the sum of all $(m_i - 1)|\bar{v}_i|$ for $m_i > 1$. Then there is an equivalence*

$$\mathbb{Z}^{B/\underline{\bar{v}}^{\underline{m}}} \simeq \Sigma^{-m'+4-2\rho}\kappa_{M\mathbb{R}}(\underline{\bar{v}}'; B/\underline{\bar{v}}^{\underline{m}}).$$

Here and for the rest of the section we will implicitly localize everything at 2 again. Before we prove the theorem, we need some preparation.

Lemma 10.2 *Let $\underline{m} = (m_1, \dots)$ be a sequence of nonnegative integers with a finite number n of nonzero entries. Then*

$$\mathbb{Z}^{B/\underline{\bar{v}}^{\underline{m}}} \simeq \Sigma^{-|\underline{m}|-n}(\mathbb{Z}^B)/\underline{\bar{v}}^{\underline{m}}.$$

Proof Let Y be an arbitrary (C_2–)spectrum and $\Sigma^{|v|}Y \xrightarrow{v} Y \to Y/v$ be a cofibre sequence. Then we have an induced cofibre sequence

$$\mathbb{Z}^{Y/v} \to \mathbb{Z}^{Y} \xrightarrow{v} \Sigma^{-|v|}\mathbb{Z}^{Y} \to \Sigma\mathbb{Z}^{Y/v} \simeq \Sigma^{-|v|}(\mathbb{Z}^{Y})/v.$$

Thus, $\mathbb{Z}^{Y/v} \simeq \Sigma^{-|v|-1}(\mathbb{Z}^{Y})/v$. The claim follows by induction. □

Lemma 10.3 *The element $\bar{v}_i^{3k}$ acts trivially on $B/\bar{v}_i^k$ for every $i \geq 1$ and $k \geq 1$.*

Proof By the commutativity of the diagram

$$\begin{tikzcd} \Sigma^{k|\bar{v}_i|}B \arrow[r, "q"] \arrow[d, "\bar{v}_i^k"] & \Sigma^{k|\bar{v}_i|}B/\bar{v}_i^k \arrow[d, "\bar{v}_i^k"] \\ B \arrow[r] & B/\bar{v}_i^k \end{tikzcd}$$

we see that the composite $\bar{v}_i^k q$ is zero, and so the $\bar{v}_i^k$ on the right factors over an $M\mathbb{R}$–linear map $\Sigma^{2k|\bar{v}_i|+1}B \to B/\bar{v}_i^k$. As $[\Sigma^{2k|\bar{v}_i|+1}B, B/\bar{v}_i^k]_{M\mathbb{R}}$ is a retract of $[\Sigma^{2k|\bar{v}_i|+1}M\mathbb{R}, B/\bar{v}_i^k]_{M\mathbb{R}} \cong \pi^{C_2}_{2k|\bar{v}_i|+1}B/\bar{v}_i^k$, we just have to show that $\bar{v}_i^{2k}x = 0$ for every $x \in \pi_{2k|\bar{v}_i|+1}B/\bar{v}_i^k$.

We have a short exact sequence

$$0 \to (\pi_\star^{C_2}B)/\bar{v}_i^k \to \pi_\star^{C_2}(B/\bar{v}_i^k) \to \{\pi^{C_2}_{\star-k|\bar{v}_i|-1}B\}_{\bar{v}_i^k} \to 0.$$

As $\bar{v}_i^k x$ clearly maps to zero, it is the image of a $y \in (\pi_\star^{C_2}B)/\bar{v}_i^k$. But $\bar{v}_i^k y = 0$. □

Lemma 10.4 *We have*

$$B/\bar{v}_i^l \otimes_{M\mathbb{R}} B/\bar{v}_j^m \simeq B/(\bar{v}_i^l, \bar{v}_j^m).$$

Furthermore, there is an equivalence

$$\operatorname*{hocolim}_l \Sigma^{-(l-1)|\bar{v}_i|}B/\bar{v}_i^l \otimes_{M\mathbb{R}} B/\bar{v}_i^m \simeq \Sigma^{|\bar{v}_i|+1}B/\bar{v}_i^m$$

of $M\mathbb{R}$–modules if $m \geq 1$.

Proof We have

$$B \otimes_{M\mathbb{R}} B \simeq \operatorname{hocolim}(B \xrightarrow{e} B \xrightarrow{e} \cdots) \simeq B,$$

where e denotes again the Quillen–Araki idempotent, and thus also

$$B/\bar{v}_i^l \otimes_{M\mathbb{R}} B/\bar{v}_j^m \simeq B/(\bar{v}_i^l, \bar{v}_j^m).$$

Thus, the maps in the homotopy colimit in the lemma are induced by the following diagram of cofibre sequences:

$$\begin{tikzcd}
\Sigma^{|\bar{v}_i|} B/\bar{v}_i^m \arrow[r, "\bar{v}_i^l"] \arrow[d, "\mathrm{id}"] & \Sigma^{-(l-1)|\bar{v}_i|} B/\bar{v}_i^m \arrow[r] \arrow[d, "\bar{v}_i"] & \Sigma^{-(l-1)|\bar{v}_i|} B/\bar{v}_i^l \otimes_{M\mathbb{R}} B/\bar{v}_i^m \arrow[d] \\
\Sigma^{|\bar{v}_i|} B/\bar{v}_i^m \arrow[r, "\bar{v}_i^{l+1}"] & \Sigma^{-l|\bar{v}_i|} B/\bar{v}_i^m \arrow[r] & \Sigma^{-l|\bar{v}_i|} B/\bar{v}_i^{l+1} \otimes_{M\mathbb{R}} B/\bar{v}_i^m
\end{tikzcd}$$

We can assume that the homotopy colimit only runs over $l \geq 3m$ so that, by the last lemma, the two cofibre sequences split, and we get

$$\Sigma^{-(l-1)|\bar{v}_i|} B/\bar{v}_i^l \otimes_{M\mathbb{R}} B/\bar{v}_i^m \simeq \Sigma^{-(l-1)|\bar{v}_i|} B/\bar{v}_i^m \oplus \Sigma^{|\bar{v}_i|+1} B/\bar{v}_i^m.$$

The corresponding map

$$\Sigma^{-(l-1)|\bar{v}_i|} B/\bar{v}_i^m \oplus \Sigma^{|\bar{v}_i|+1} B/\bar{v}_i^m \to \Sigma^{-l|\bar{v}_i|} B/\bar{v}_i^m \oplus \Sigma^{|\bar{v}_i|+1} B/\bar{v}_i^m$$

induces multiplication by $\bar{v}_i$ on the first summand, the identity on the second plus possibly a map from the second summand to the first.

Using this decomposition, it is easy to show that

$$\operatorname*{hocolim}_l \Sigma^{-(l-1)|\bar{v}_i|} B/\bar{v}_i^l \otimes_{M\mathbb{R}} B/\bar{v}_i^m \to \Sigma^{|\bar{v}_i|+1} B/\bar{v}_i^m$$

(defined by the projection on the second summand for $l \geq 3m$) is an equivalence. Indeed, on homotopy groups the map is clearly surjective. And if

$$(x, y) \in \pi_\star^{C_2} \Sigma^{-l|\bar{v}_i|} B/\bar{v}_i^m \oplus \pi_\star^{C_2} \Sigma^{|\bar{v}_i|+1} B/\bar{v}_i^m$$

maps to $0 \in \pi_\star^{C_2} \Sigma^{|\bar{v}_i|+1} B/\bar{v}_i^m$, then $y = 0$ and $(x, 0)$ represents 0 in the colimit because $\bar{v}_i$ acts nilpotently. □

Proof of Theorem 10.1 As in the theorem, let $\underline{v}'$ be the sequence of $\bar{v}_i$ such that $m_i = 0$ and also denote by $\underline{v}'' = (\bar{v}_{i_1}, \bar{v}_{i_2}, \dots)$ the sequence of $\bar{v}_i$ such that $m_i \neq 0$.

We begin with the case that $\underline{m}$ has only finitely many nonzero entries (say n). By Lemma 10.2, we see that

$$\mathbb{Z}^{B/\underline{v}^{\underline{m}}} \simeq \Sigma^{-|\underline{m}|-n} (\mathbb{Z}^B)/\underline{v}^{\underline{m}}.$$

Combining this with Theorem 9.5, we obtain

$$\begin{aligned}
\mathbb{Z}^{B/\underline{v}^{\underline{m}}} &\simeq \Sigma^{-|\underline{m}|-n+4-2\rho} \kappa_{M\mathbb{R}}(\underline{v}, B)/\underline{v}^{\underline{m}} \\
&\simeq \Sigma^{-|\underline{m}|-n+4-2\rho} \kappa_{M\mathbb{R}}(\underline{v}', \kappa_{M\mathbb{R}}(\underline{v}'', B/\underline{v}^{\underline{m}})).
\end{aligned}$$

Thus, we have to show that $\kappa_{M\mathbb{R}}(\underline{v}'', B/\underline{v}^{\underline{m}}) \simeq \Sigma^{|\bar{v}_{i_1}|+\cdots+|\bar{v}_{i_n}|+n} B/\underline{v}^{\underline{m}}$.

By Lemma 10.4, we have an equivalence

$$(B/\underline{v}^{\underline{m}})/(\bar{v}_{i_1}^{l_{i_1}},\dots,\bar{v}_{i_n}^{l_{i_n}}) \simeq (B/\bar{v}_1^{l_{i_1}} \otimes_{M\mathbb{R}} B/\bar{v}_1^{m_{i_1}}) \otimes_{M\mathbb{R}} \dots \otimes_{M\mathbb{R}} (B/\bar{v}_n^{l_{i_n}} \otimes_{M\mathbb{R}} B/\bar{v}_n^{m_{i_n}}).$$

If we let now the homotopy colimit run over the sequences $(l_{i_1},\dots,l_{i_n})$, we can do it separately for each tensor factor. Hence, we obtain again by Lemma 10.4 an equivalence

$$\kappa_{M\mathbb{R}}(\underline{v}'', B/\underline{v}^{\underline{m}}) \simeq \Sigma^{|\bar{v}_{i_1}|+\dots+|\bar{v}_{i_n}|+n} B/\underline{v}^{\underline{m}}.$$

Thus, we have shown the theorem when $\underline{m}$ has only finitely many nonzero entries.

We prove the case that $\underline{m}$ has possibly infinitely many nonzero entries by a colimit argument. Define $\underline{m}_{\le k}$ to be the sequence obtained from $\underline{m}$ by setting $m_{k+1}, m_{k+2}, \dots$ to zero. Then $B/\underline{m} \simeq \operatorname{hocolim}_k B/\underline{m}_{\le k}$ and thus $\mathbb{Z}^{B/\underline{m}} \simeq \operatorname{holim}_k \mathbb{Z}^{B/\underline{m}_{\le k}}$. Denote by $\underline{v}'_{\le k}$ the sequence of $\bar{v}_i$ such that $m_i = 0$ or $i > k$ and by m'_k the quantity $|\underline{m}_{\le k} - \underline{1}|$; note that $m'_k = m'$ for k large.

We have to show that the map

$$h\colon \Sigma^{-m'}\kappa_{M\mathbb{R}}(\underline{v}', B/\underline{v}^{\underline{m}}) \to \operatorname*{holim}_k \Sigma^{-m'_k}\kappa_{M\mathbb{R}}(\underline{v}'_{\le k}, B/\underline{v}^{\underline{m}_{\le k}})$$

is an equivalence. This map is defined as follows: We know that

$$\kappa_{M\mathbb{R}}(\underline{v}', B/\underline{v}^{\underline{m}}) \simeq \operatorname*{hocolim}_k \kappa_{M\mathbb{R}}(\underline{v}', B/\underline{v}^{\underline{m}_{\le k}}).$$

Using this, we get a map induced from the maps

$$\kappa_{M\mathbb{R}}(\underline{v}', B/\underline{v}^{\underline{m}_{\le k}}) \to \kappa_{M\mathbb{R}}(\underline{v}'_{\le k}, B/\underline{v}^{\underline{m}_{\le k}})$$

for k large.

By Corollary 4.6, we can describe what happens on $\pi_{*\rho}^{C_2}$: The left-hand side has as $\mathbb{Z}$–basis monomials of the form $\underline{v}^{\underline{n}}$ with only finitely many n_i nonzero, $n_i \le 0$ and $n_i \ge -m_i + 1$ if $m_i \ne 0$. Likewise,

$$\pi_{*\rho}^{C_2}\big(\Sigma^{m'_k}\kappa_{M\mathbb{R}}(\underline{v}'_{\le k}, B/\underline{v}^{\underline{m}_{\le k}})\big)$$

has as $\mathbb{Z}$–basis monomials of the form $\underline{v}^{\underline{n}}$ with only finitely many n_i nonzero, $n_i \le 0$ and $n_i \ge -m_i + 1$ if $m_i \ne 0$ and $i \le k$. The maps in the homotopy limit induce the obvious inclusion maps. Thus, clearly the map

$$\pi_{*\rho}^{C_2}\big(\Sigma^{m'}\kappa_{M\mathbb{R}}(\underline{v}', B/\underline{v}^{\underline{m}})\big) \to \lim_k \pi_{*\rho}^{C_2}\big(\Sigma^{m'_k}\kappa_{M\mathbb{R}}(\underline{v}'_{\le k}, B/\underline{v}^{\underline{m}_{\le k}})\big)$$

is an isomorphism.

It remains to show $\lim_k^1 \pi_{*\rho+1}^{C_2}\big(\Sigma^{m'_k}\kappa_{M\mathbb{R}}(\underline{v}'_{\le k}, B/\underline{v}^{\underline{m}_{\le k}})\big)$ vanishes. By Corollary 4.8, every term has as $\mathbb{F}_2$–basis monomials of the form $a\underline{v}^{\underline{n}}$ with only finitely many n_i nonzero, $n_i \le 0$ and $n_i \ge -m_i + 1$ if $m_i \ne 0$ and $i \le k$. The system becomes stationary in every degree, more precisely if $* > -2^{k+1}$. Thus, the $\lim^1$–term vanishes.

A similar $\lim^1$–argument also shows that the odd underlying homotopy groups of $\mathrm{holim}_k\, \Sigma^{-m'_k} \kappa_{M\mathbb{R}}(\underline{\bar{v}}'_{\le k}, B/\underline{\bar{v}}^{\underline{m}_{\le k}})$ vanish.

As the source of h is strongly even by Corollary 4.6 and by the arguments we just gave the morphism h induces an isomorphism on $\underline{\pi}^{C_2}_{*\rho}$ and on (odd) underlying homotopy groups, Lemma 4.11 implies that h is an equivalence. □

Part IV Local cohomology computations

In Part IV, we will describe the local cohomology spectral sequence in some detail, and use it to understand the structure of the $H\underline{\mathbb{Z}}$–cellularization of $BP\mathbb{R}\langle n\rangle$. The calculation is not difficult, but on the other hand it is quite hard to follow because it is made up of a large number of easy calculations which interact a little, and because one needs to find a helpful way to follow the $RO(C_2)$–graded calculations.

In contrast, the case of $k\mathbb{R}$ is simple enough to be explained fully without further scaffolding, and it introduces many of the structures that we will want to highlight. Since it may also be of wider interest than the general case of $BP\mathbb{R}\langle n\rangle$ we devote Section 11 to it before returning to the general case in Section 12. Section 13 will then give a more detailed account in the interesting case $n = 2$.

Let us also recall some notation used throughout this part. As in the rest of the paper we work 2–locally, except when speaking about $k\mathbb{R}$ or $tmf_1(3)$ when fewer primes need be inverted. We often write $\delta = 1-\sigma \in RO(C_2)$. We also recall the duality conventions from Section 3A; in particular, for an $\mathbb{F}_2$–vector space $V^\vee$ equals the dual vector space $\mathrm{Hom}_{\mathbb{F}_2}(V, \mathbb{F}_2)$ and for a torsion-free $\mathbb{Z}$–module M, we set $M^* = \mathrm{Hom}(M, \mathbb{Z})$.

If R is a C_2–spectrum, we will use the notation $R^{C_2}_\star$ for its $RO(C_2)$–graded homotopy groups. We will also write $R^{hC_2}_\star = \pi^{C_2}_\star(R^{(EC_2)_+})$ and similarly for geometric fixed points and the Tate construction.

11 The local cohomology spectral sequence for $k\mathbb{R}$

This section focuses entirely on the classical case of $k\mathbb{R}$, where there are already a number of features of interest. This gives a chance to introduce some of the structures we will use for the general case.

11A The local cohomology spectral sequence

Gorenstein duality for $k\mathbb{R}$ (Corollary 6.8) has interesting implications for the coefficient ring, both computationally and structurally. Writing $\star$ for $RO(C_2)$–grading as usual, the local cohomology spectral sequence [11, Section 3] takes the following form.

Proposition 11.1 *There is a spectral sequence of $k\mathbb{R}^{C_2}_\star$–modules*

$$H^*_{(\bar{v})}(k\mathbb{R}^{C_2}_\star) \implies \Sigma^{-4+\sigma}\pi^{C_2}_\star(\mathbb{Z}^{k\mathbb{R}}).$$

The homotopy of the Anderson dual in an arbitrary degree $\alpha \in RO(C_2)$ lies in an exact sequence

$$0 \to \mathrm{Ext}_{\mathbb{Z}}(k\mathbb{R}^{C_2}_{-\alpha-1}, \mathbb{Z}) \to \pi^{C_2}_\alpha(\mathbb{Z}^{k\mathbb{R}}) \to \mathrm{Hom}_{\mathbb{Z}}(k\mathbb{R}^{C_2}_{-\alpha}, \mathbb{Z}) \to 0.$$

Since local cohomology is entirely in cohomological degrees 0 and 1, the spectral sequence collapses to a short exact sequence

$$0 \to \Sigma^{-1}H^1_{(\bar{v})}(k\mathbb{R}^{C_2}_\star) \to \Sigma^{-4+\sigma}\pi^{C_2}_\star(\mathbb{Z}^{k\mathbb{R}}) \to H^0_{(\bar{v})}(k\mathbb{R}^{C_2}_\star) \to 0.$$

This sequence is not split, even as abelian groups.

One should not view Proposition 11.1 as an algebraic formality: it embodies the fact that $k\mathbb{R}^{C_2}_\star$ is a very special ring. To illustrate this, we recall the calculation of $k\mathbb{R}^{C_2}_\star$ in Section 11B. In Section 11C, we calculate its local cohomology, and how the Gorenstein duality isomorphism with the known homotopy of the Anderson dual works.

11B The ring $k\mathbb{R}^{C_2}_\star$

One may easily calculate $k\mathbb{R}^{C_2}_\star$. This has already been done in [7], but we sketch a slightly different method. We will first calculate $k\mathbb{R}^{hC_2}_\star$ and then use the Tate square [12].

In the homotopy fixed point spectral sequence

$$\mathbb{Z}[\bar{v}, a, u^{\pm 1}]/2a \implies k\mathbb{R}^{hC_2}_\star,$$

all differentials are generated by $d_3(u) = \bar{v}a^3$. Indeed, this differential is forced by $\eta^4 = 0$ and there is no room for further ones. It follows that $U = u^2$ is an infinite cycle, and so the whole ring is U–periodic:

$$k\mathbb{R}^{hC_2}_\star = BB[U, U^{-1}],$$

where BB is a certain "basic block". This basic block is a sum

$$BB = BR \oplus (2u)\cdot\mathbb{Z}[\bar{v}]$$

as BR–modules, where

$$BR = \mathbb{Z}[\bar{v}, a]/(2a, \bar{v}a^3).$$

It is worth illustrating BB in the plane (with $BB_{a+b\sigma}$ placed at the point (a, b)); see Figure 1. The squares and circles represent copies of $\mathbb{Z}$, and the dots represent

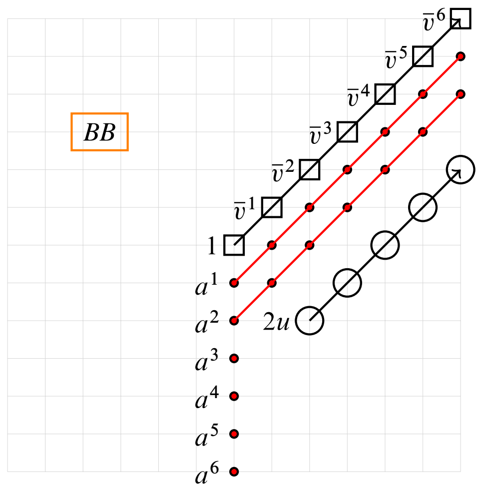


Figure 1: The basic block *BB*

copies of $\mathbb{F}_2$. The left-hand vertical column consists of 1 (at the origin, $(0,0)$) and the powers of a, but the feature to concentrate on is the diagonal lines representing $\mathbb{Z}[\overline{v}]$–submodules. These are either copies of $\mathbb{Z}[\overline{v}]$ or of $\mathbb{F}_2[\overline{v}]$ or simply copies of $\mathbb{F}_2$.

Proceeding with the calculation, we may invert a to find the homotopy of the Tate spectrum $k\mathbb{R}^t = F(E(C_2)_+, k\mathbb{R}) \wedge S^{\infty\sigma}$:

$$k\mathbb{R}^{tC_2}_\star = \mathbb{F}_2[a, a^{-1}][U, U^{-1}].$$

One also sees that the homotopy of the geometric fixed points (the equivariant homotopy of $k\mathbb{R}^\Phi = k\mathbb{R} \wedge S^{\infty\sigma}$) is

$$k\mathbb{R}^{\Phi C_2}_\star = \mathbb{F}_2[a, a^{-1}][U]$$

using the following lemma:

Lemma 11.2 *Let X be a C_2–spectrum which is nonequivariantly connective and such that $X^{C_2} \to X^{hC_2}$ is a connective cover. Then $X^{\Phi C_2} \to X^{tC_2}$ is a connective cover as well.*

Proof This follows from the diagram of long exact sequences

$$\begin{tikzcd}
\pi_k X_{hC_2} \arrow[r] \arrow[d] & \pi_k X^{C_2} \arrow[r] \arrow[d] & \pi_k X^{\Phi C_2} \arrow[r] \arrow[d] & \pi_{k-1} X_{hC_2} \arrow[r] \arrow[d] & \pi_{k-1} X^{C_2} \arrow[d] \\
\pi_k X_{hC_2} \arrow[r] & \pi_k X^{hC_2} \arrow[r] & \pi_k X^{tC_2} \arrow[r] & \pi_{k-1} X_{hC_2} \arrow[r] & \pi_{k-1} X^{hC_2}
\end{tikzcd}$$

the fact that X_{hC_2} is connective, and the five lemma. □

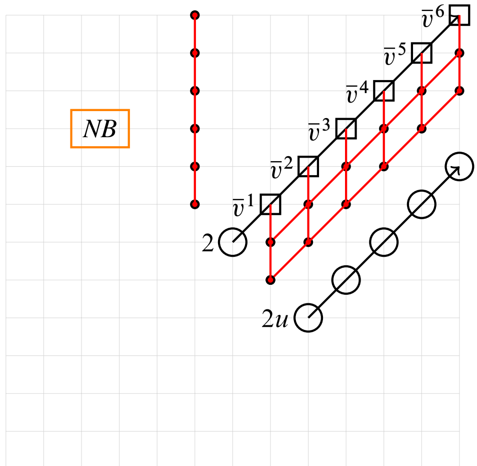


Figure 2: The negative block NB

Now the Tate square

$$\begin{tikzcd} k\mathbb{R} \arrow[r] \arrow[d] & k\mathbb{R} \wedge S^{\infty\sigma} \arrow[d] \\ k\mathbb{R}^{(EC_2)_+} \arrow[r] & k\mathbb{R}^{(EC_2)_+} \wedge S^{\infty\sigma} \end{tikzcd}$$

gives $k\mathbb{R}^{C_2}_\star$.

It is convenient to observe that the two rows are of the form $M \to M[1/a]$, so the fibre is $\Gamma_a M$. Since the two rows have equivalent fibres, we calculate the homotopy of the second and obtain

$$k\mathbb{R}^\star_{hC_2} = NB[U, U^{-1}],$$

where NB is quickly calculated as the (a)–local cohomology $H^*_{(a)}(BB)$ (and named NB for "negative block"). The element a acts vertically and we can immediately read off the answer: the tower $\mathbb{Z}[a]/(2a)$ gives some H^1, and the rest is a–power torsion:

$$NB = BB' \oplus \Sigma^{-\delta}\mathbb{F}_2[a]^\vee,$$

where $BB' \subset BB$ is the BR–submodule generated by 2, $\overline{v}$ and $2u$ (informally, we may say that BB' omits from BB all monomials a^k for $k \geq 1$ and the generator 1). Note that NB is placed so that its element 2 is in degree 0 for ease of comparison to BB; all occurrences of NB in $k\mathbb{R}^{C_2}_\star$ involve nontrivial suspensions.

Again, it is helpful to display the negative block; see Figure 2. This differs from BB in that the powers of a have been deleted, and replaced by a new left-hand column $\Sigma^{-\delta}\mathbb{F}_2[a]^\vee$. The other new feature is that the copy of $\mathbb{Z}[\overline{v}]$ generated by 1 has been replaced by the kernel $(2,\overline{v})$ of $\mathbb{Z}[\overline{v}]\to\mathbb{F}_2$, as indicated by the circle at the origin, labelled by its generator 2.

The Tate square then lets us read off

$$k\mathbb{R}_\star^{C_2}=\bigoplus_{k\leq -1} NB\cdot\{U^k\}\oplus\bigoplus_{k\geq 0} BB\cdot\{U^k\}=(U^{-1}\cdot NB[U^{-1}])\oplus BB[U].$$

The $\mathbb{Z}[U]$ module structure is given by letting U act in the obvious way on the NB and BB parts, and by the maps

$$NB\to BB'\to BB$$

in passage from the U^{-1} factor of NB to the U^0 factor of BB.

Perhaps it is helpful to note that with the exception of the towers $U^{-k}\Sigma^{-\delta}\mathbb{F}_2[a]^\vee$, we have a subring of $BB[U,U^{-1}]$, which consists of blocks $BB\cdot U^i$ for $i\geq 0$ and blocks $BB'\cdot U^i$ for $i<0$.

11C Local cohomology

Recall that we are calculating local cohomology with respect to the principal ideal $(\overline{v})$ so that we only need to consider $k\mathbb{R}_\star^{C_2}$ as a $\mathbb{Z}[\overline{v}]$–module. As such it is a sum of suspensions of the blocks BB and NB, so we just need to calculate the local cohomology of these.

More significantly, $\mathbb{Z}[\overline{v}]$ is graded over multiples of the regular representation, so local cohomology calculations may be performed on one diagonal at a time (ie we fix n and consider gradings $n+*\rho$). The only modules that occur are

$$\mathbb{Z}[\overline{v}],\quad \mathbb{F}_2[\overline{v}],\quad \mathbb{F}_2\quad\text{and the ideal } (2,\overline{v})\subseteq\mathbb{Z}[\overline{v}],$$

each of which has local cohomology that is very easily calculated.

Lemma 11.3 *The local cohomology of the basic block BB is as follows:*

$$H^0_{(\overline{v})}(BB)=a^3\mathbb{F}_2[a],$$
$$H^1_{(\overline{v})}(BB)=\Sigma^{-\rho}\mathbb{Z}[\overline{v}]^*\oplus\Sigma^{-\rho+2\delta}\mathbb{Z}[\overline{v}]^*\oplus\Sigma^{-\rho-\sigma}\mathbb{F}_2[\overline{v}]^\vee\oplus\Sigma^{-\rho-2\sigma}\mathbb{F}_2[\overline{v}]^\vee.$$

Proof The local cohomology is the cohomology of the complex

$$BB\to BB[1/\overline{v}].$$

It is clear that

$$BB[1/\overline{v}]=\mathbb{Z}[\overline{v},\overline{v}^{-1}]\oplus u\cdot\mathbb{Z}[\overline{v},\overline{v}^{-1}]\oplus a\cdot\mathbb{F}_2[\overline{v},\overline{v}^{-1}]\oplus a^2\cdot\mathbb{F}_2[\overline{v},\overline{v}^{-1}].\qquad\square$$

Turning to NB, we recall that $NB = BB' \oplus \Sigma^{-\delta}\mathbb{F}_2[a]^\vee$, and we have a short exact sequence

$$0 \to BB' \to BB \to \mathbb{F}_2[a] \to 0.$$

The local cohomology is thus easily deduced from that of BB.

Lemma 11.4 *The local cohomology of the negative block NB is as follows:*

$$H^0_{(\overline{v})}(NB) = \Sigma^{-\delta}\mathbb{F}_2[a]^\vee,$$
$$H^1_{(\overline{v})}(NB) = \Sigma^{-\rho}\mathbb{Z}[\overline{v}]^* \oplus \mathbb{F}_2 \oplus \Sigma^{-\rho+2\delta}\mathbb{Z}[\overline{v}]^* \oplus \Sigma^{-\sigma}\mathbb{F}_2[\overline{v}]^\vee \oplus \Sigma^{-2\sigma}\mathbb{F}_2[\overline{v}]^\vee.$$

More properly, the $\mathbb{Z}[\overline{v}]$–module structure of the sum of the first two terms is

$$\Sigma^{-\rho}\mathbb{Z}[\overline{v}]^* \oplus \mathbb{F}_2 \cong \mathbb{Z}[\overline{v}]^*/(2\cdot 1^*).$$

Proof The local cohomology is the cohomology of the complex

$$NB \to NB[1/\overline{v}].$$

It is clear that $NB[1/\overline{v}] = BB[1/\overline{v}]$, which makes the part coming from the 2–torsion clear. For the $\mathbb{Z}$–torsion free part, it is helpful to consider the exact sequence

$$0 \to (2, \overline{v}) \to \mathbb{Z}[\overline{v}] \to \mathbb{F}_2 \to 0$$

and then consider the long exact sequence in local cohomology. □

Immediately from the defining cofibre sequence $\Gamma_{\overline{v}}k\mathbb{R} \to k\mathbb{R} \to k\mathbb{R}[1/\overline{v}]$ we see that there is a short exact sequence

$$0 \to H^1_{(\overline{v})}(\Sigma^{-1}k\mathbb{R}^{C_2}_\star) \to \pi^{C_2}_\star(\Gamma_{(\overline{v})}k\mathbb{R}) \to H^0_{(\overline{v})}(k\mathbb{R}^{C_2}_\star) \to 0.$$

This gives $\pi^{C_2}_\star(\Gamma_{(\overline{v})}k\mathbb{R})$ up to extension. The Gorenstein duality isomorphism can be used to resolve the remaining extension issues, and the answer is recorded in the proposition below.

The diagram Figure 3 should help the reader interpret the statement and proof of the calculation of the homotopy of $\Gamma_{(\overline{v})}k\mathbb{R}$. We have omitted dots, circles and boxes except at the ends of diagonals or where an additional generator is required. The vertical lines denote multiplication by a and the dashed vertical line is an exotic multiplication by a that is not visible on the level of local cohomology. The green diamond does not denote a class, but marks the point one has to reflect (nontorsion classes) at to see Anderson duality. Torsion classes are shifted by -1 after reflection (ie shifted one step horizontally to the left).

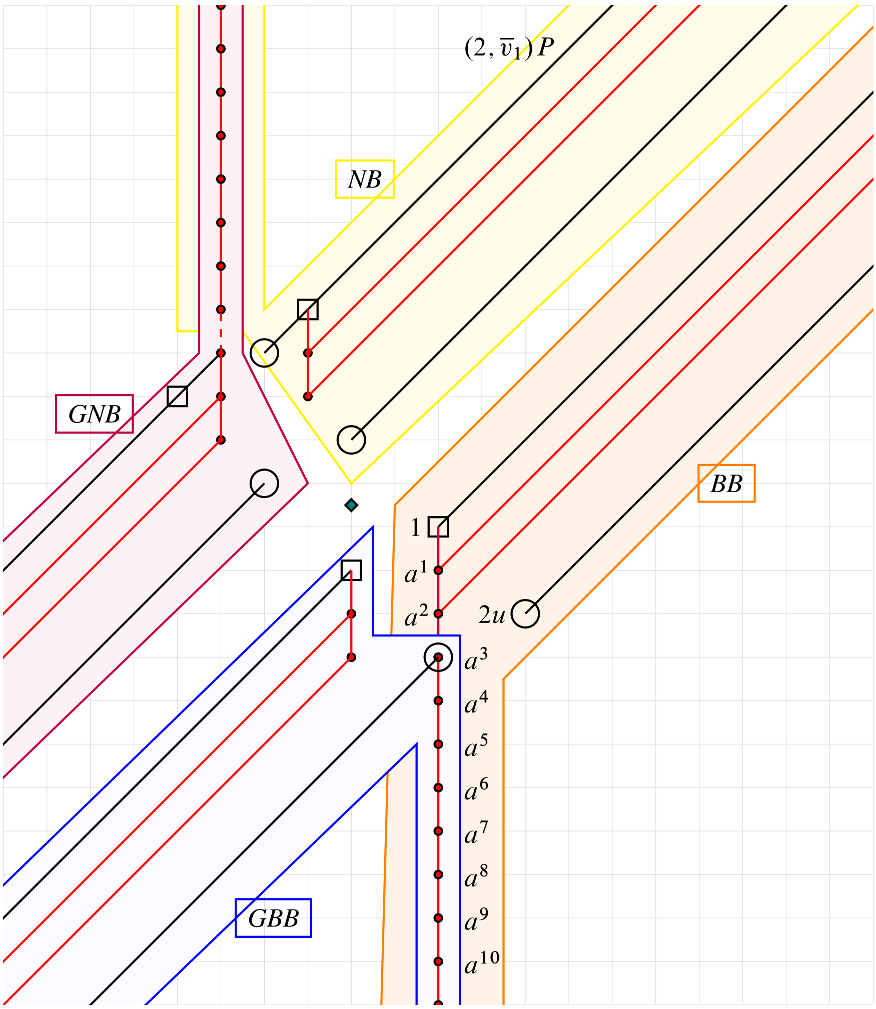


Figure 3: Gorenstein duality for $k\mathbb{R}$

Proposition 11.5 *The homotopy of the derived $\overline{v}$–power torsion is given by*

$$\pi_\star^{C_2}(\Gamma_{(\overline{v})}k\mathbb{R}) \cong (U^{-1}\cdot GNB[U^{-1}]) \oplus GBB[U],$$

where GBB and GNB are based on the local cohomology of BB and NB respectively, and described as follows. We have

$$GBB = \Sigma^{-2-\sigma}\big[\mathbb{Z}[\overline{v}]^* \oplus a\cdot\mathbb{F}_2[\overline{v}]^\vee \oplus a^2\cdot\mathbb{F}_2[\overline{v}]^\vee \oplus u\cdot N\big],$$

where N (with top in degree 0) is given by an exact sequence nonsplit in degree 0:

$$0 \to \mathbb{Z}[\overline{v}]^* \to N \to \mathbb{F}_2[a] \to 0.$$

Similarly,

$$GNB=\Sigma^{-1}\big[\mathbb{Z}[\bar{v}]^*/(2\cdot(1^*))\oplus a\cdot\mathbb{F}_2[\bar{v}]^\vee\oplus a^2\cdot\mathbb{F}_2[\bar{v}]^\vee\oplus\Sigma^{1-3\sigma}\mathbb{Z}[\bar{v}]^*\oplus\Sigma^\sigma\mathbb{F}_2[a]^\vee\big],$$

where the action of a is as suggested by the sum decomposition except that multiplication by a is nontrivial wherever possible (ie when one dot is vertically above another, or where a box is vertically above a dot).

Proof We first note that the contributions from the different blocks do not interact. Indeed, the only time that different blocks give contributions in the same degree come from the $\mathbb{F}_2[a]$ towers of BB: one class in that degree is $\bar{v}$–divisible (and not killed by $\bar{v}$) and the other class is annihilated by $\bar{v}$. We may therefore consider the blocks entirely separately.

The block GBB comes from the local cohomology of BB and therefore lives in a short exact sequence

$$0\to H^1_{(\bar{v})}(\Sigma^{-1}BB)\to GBB\to H^0_{(\bar{v})}(BB)\to 0.$$

The block GNB comes from the local cohomology of NB and therefore lives in a short exact sequence

$$0\to H^1_{(\bar{v})}(\Sigma^{-1}NB)\to GNB\to H^0_{(\bar{v})}(NB)\to 0.$$

Most questions about module structure over $BB[U]$ are resolved by degree, but there are two which remain. These can be resolved by Gorenstein duality (Corollary 6.8) and the known module structure in $\mathbb{Z}^{k\mathbb{R}}$.

In GBB, the additive extension in $\pi^{C_2}_{-3\sigma}$ is nontrivial:

$$\pi^{C_2}_{-3\sigma}(\Gamma_{(\bar{v})}k\mathbb{R})\cong\mathbb{Z}.$$

Also the multiplication by a

$$\mathbb{F}_2\cong GNB_{-1+\sigma}\to GNB_{-1}\cong\mathbb{F}_2$$

is nonzero (where $GNB_{-1+\sigma}$ corresponds to $\pi^{C_2}_{-5+5\sigma}(\Gamma_{(\bar{v})}k\mathbb{R})$ in the U^{-1}–shift). □

Remark 11.6 It is striking that the duality relates the top BB to the bottom NB (ie Anderson duality takes the part of $\Gamma_{\bar{v}}k\mathbb{R}$ coming from the local cohomology of BB to NB), and it takes the bottom NB to the top BB (ie Anderson duality takes the part of $\Gamma_{\bar{v}}k\mathbb{R}$ coming from the local cohomology of NB to BB). Indeed, as commented after Lemma 11.2, since $NB=\Gamma_{(a)}BB$, we have

$$\Sigma^{2+\sigma}\,\Gamma_{(\bar{v})}BB\simeq(\Gamma_{(a)}BB)^*\quad\text{and}\quad\Gamma_{(\bar{v},a)}BB\simeq\Sigma^{-2-\sigma}BB^*,$$

with the second stating that BB is Gorenstein of shift $-2-\sigma$ for the ideal $(a,\bar{v})$.

By extension, Anderson duality takes the part of $\Gamma_{\overline{v}} k\mathbb{R}$ coming from the local cohomology of all copies of BB to all copies of NB and vice versa. This might suggest separating $k\mathbb{R}$ into a part with homotopy $BB[U]$, giving a cofibre sequence

$$\langle BB[U]\rangle \to k\mathbb{R} \to \langle U^{-1}NB[U^{-1}]\rangle,$$

where the angle brackets refer to a spectrum with the indicated homotopy. However one may see that there is no C_2–spectrum with homotopy the Mackey functor corresponding to $BB[U]$ (considering the $b\sigma$ and $(b+1)\sigma$ rows one sees that the nonequivariant homotopy of the spectrum would be zero up to about degree $2b$; taking all rows together it would have to be nonequivariantly contractible and hence a–periodic). Similarly, there is no spectrum with homotopy $U^{-1}NB[U^{-1}]$, so these dualities are purely algebraic.

12 The local cohomology spectral sequence for $BP\mathbb{R}\langle n\rangle$

Gorenstein duality for $BP\mathbb{R}\langle n\rangle$ (Example 5.2) has interesting implications for the coefficient ring, both computationally and structurally. Writing $\star$ for $RO(C_2)$–grading as usual, the local cohomology spectral sequence [11, Section 3] takes the form described in the following proposition. We now revert to our standard assumption of working 2–locally, so $\mathbb{Z}$ means the 2–local integers.

Proposition 12.1 *There is a spectral sequence of $BP\mathbb{R}\langle n\rangle^{C_2}_\star$–modules*

$$H^*_{\overline{J}_n}(BP\mathbb{R}\langle n\rangle^{C_2}_\star) \implies \Sigma^{-(D_n+n+2)-(D_n-2)\sigma}\pi^{C_2}_\star(\mathbb{Z}^{BP\mathbb{R}\langle n\rangle})$$

for $\overline{J}_n = (\overline{v}_1,\dots,\overline{v}_n)$. The homotopy of the Anderson dual in an arbitrary degree $\alpha \in RO(C_2)$ is easily calculated:

$$0 \to \mathrm{Ext}_{\mathbb{Z}}(BP\mathbb{R}\langle n\rangle^{C_2}_{-\alpha-1},\mathbb{Z}) \to \pi^{C_2}_\alpha \mathbb{Z}^{BP\mathbb{R}\langle n\rangle} \to \mathrm{Hom}_{\mathbb{Z}}(BP\mathbb{R}\langle n\rangle^{C_2}_{-\alpha},\mathbb{Z}) \to 0.$$

For $n \geq 2$, the local cohomology spectral sequence has some nontrivial differentials.

One should not view Proposition 12.1 as an algebraic formality: it embodies the fact that $BP\mathbb{R}\langle n\rangle^{C_2}_\star$ is a very special ring.

In the present section, we will discuss the implications of this for the coefficient ring for general n. The perspective is a bit distant so the reader is encouraged to refer back to $k\mathbb{R}$ (ie the case $n=1$) in Section 11 to anchor the generalities.

However the case $n=1$ is too simple to show some of what happens, so we will also illustrate the case $tmf_1(3)$ (ie the case $n=2$) in Section 13.

12A Reduction to diagonals

For brevity, we write $R_\star = BP\mathbb{R}\langle n\rangle^{C_2}_\star$. Because the ideal $\bar{J}_n = (\bar{v}_1, \dots, \bar{v}_n)$ is generated by elements whose degrees are a multiple of ρ, we can do $\bar{J}_n$–local cohomology calculations over the subring $R_{*\rho}$ of elements in degrees which are multiples of ρ.

Thus, for an $R_\star$–module $M_\star$ we have a direct sum decomposition

$$M_\star = \bigoplus_d M_{d+*\rho}$$

as $R_{*\rho}$–modules, where we refer to the gradings $d + *\rho$ as the *d–diagonal*. Hence, we also have

$$H^i_{\bar{J}_n}(M_\star) = \bigoplus_d H^i_{\bar{J}_n}(M_{d+*\rho}).$$

(We have abused notation by also writing $\bar{J}_n$ for the ideal of $R_{*\rho}$ generated by $\bar{v}_1, \dots, \bar{v}_n$.)

12B The general shape of $BP\mathbb{R}\langle n\rangle^{C_2}_\star$

By the description at the end of Section 4B, we have an isomorphism

$$R_\star = U^{-1} \cdot NB[U^{-1}] \oplus BB[U]$$

with *BB* and *NB* as described there. It is easy to see that *BB* and *NB* decompose as $R_{*\rho}$–modules into modules of a certain form we will describe now. We will implicitly 2–localize everywhere.

The modules *BB* and *NB* decompose into are

$$P = R_{*\rho} = \mathbb{Z}[\bar{v}_1, \dots, \bar{v}_n] \quad \text{and} \quad \bar{P}_s = P/(\bar{v}_0, \dots, \bar{v}_s) = \mathbb{F}_2[\bar{v}_{s+1}, \dots, \bar{v}_n]$$

for $s \geq 0$ and the ideals expressed by the exact sequences

$$0 \to (2, \bar{v}_1, \dots, \bar{v}_t) \to P \to \bar{P}_t \to 0 \quad \text{or} \quad 0 \to (\bar{v}_{s+1}, \dots, \bar{v}_t) \to \bar{P}_s \to \bar{P}_t \to 0$$

with $s \geq 0$.

Their local cohomology is easily calculated. In the first two cases, the modules only have local cohomology in a single degree:

$$H^*_{\bar{J}_n}(P) = H^n_{\bar{J}_n}(P) = P^*(-D_n\rho),$$
$$H^*_{\bar{J}_n}(\bar{P}_s) = H^{n-s}_{\bar{J}_n}(\bar{P}_s) = \bar{P}_s^\vee((D_s - D_n)\rho).$$

The top nonzero degree of P^* is zero, so $1^* \in P^*(-D_n\rho)$ is in degree $-D_n\rho = -|\bar{v}_1| - \cdots - |\bar{v}_n|$. We alert the reader to the fact that star is used in two ways:

occasionally in H^* to mean cohomological grading and rather frequently here in P^* to mean the $\mathbb{Z}$–dual of P.

Now we turn to the ideal $(\overline{v}_{s+1},\dots,\overline{v}_t)$. If $t=s+1$ the ideal is principal and $(\overline{v}_{s+1})\cong\overline{P}_s((s+1)\rho)$; thus we get a single local cohomology group

$$H^{n-s}_{\overline{J}_n}((\overline{v}_{s+1})\overline{P}_s)=\overline{P}_s^{\vee}((D_s-D_n+s+1)\rho)$$

as can be seen from the long exact sequence of local cohomology.

Otherwise we get two local cohomology groups

$$H^{n-s}_{\overline{J}_n}((\overline{v}_{s+1},\dots,\overline{v}_t)\overline{P}_s)=\overline{P}_s^{\vee}((D_n-D_s)\rho)$$

and

$$H^{n-t+1}_{\overline{J}_n}((\overline{v}_{s+1},\dots,\overline{v}_t)\overline{P}_s)=\overline{P}_t^{\vee}((D_n-D_t)\rho).$$

The case of $(2,\overline{v}_1,\dots,\overline{v}_t)$ is similar but with an extra case. The case $t=0$ is easy since then $(2)\cong P$ so the local cohomology is all in cohomological degree n where it is $P^*(-D_n\rho)$. If $t=1$ we again get a single local cohomology group

$$H^{n}_{\overline{J}_n}((2,\overline{v}_1)P)=P^*(-D_n\rho)\oplus\overline{P}_1^{\vee}((D_1-D_n)\rho).$$

Otherwise we get two local cohomology groups

$$H^{n}_{\overline{J}_n}((2,\dots,\overline{v}_t)P)=P^*(-D_n\rho)\quad\text{and}\quad H^{n-t+1}_{\overline{J}_n}((2,\dots,\overline{v}_t)P)=\overline{P}_t^{\vee}((D_t-D_n)\rho).$$

12C The special case $n=1$

The best way to make the patterns apparent is to look at the simplest cases. In this section, we begin with $k\mathbb{R}^{C_2}_{\star}$ as treated in Section 11 above, and we encourage the reader to relate the calculations here to the diagrams in Section 11. In that case,

$$P=k\mathbb{R}^{C_2}_{*\rho}=\mathbb{Z}[\overline{v}_1],\quad \overline{P}_0=\mathbb{F}_2[\overline{v}_1]\quad\text{and}\quad\overline{P}_1=\mathbb{F}_2.$$

Table 1 (left) displayes *BB* by d–diagonal. The position of the modules along the d–diagonal can be inferred from the label at the top of the column. Thus the first column has generators in degree $-d\sigma$, and the second column similarly, but in the column of u (namely the 2–column). Noting that u is on the 4–diagonal, the d^{th} row has generators in $|u|-(d-4)\sigma=2-(d-2)\sigma$. For example, along the 4–diagonal we have $a^4\overline{P}_1\oplus(2u)P$.

Taking local cohomology, and shifting $H^s_{\overline{J}_n}$ down by s (as in the local cohomology spectral sequence), we have Table 1 (right). Note that shifting down by s both lowers d by s and adds a shift by $-s\rho$. For example, considering the 3–diagonal of this table,

d	BB: 1	BB: u
0	P	
1	$\overline{P}_0$	
2	$\overline{P}_0$	
3	$\overline{P}_1$	
4	$\overline{P}_1$	$(2)P$
5	$\overline{P}_1$	
6	$\overline{P}_1$	
7	$\overline{P}_1$	
8	$\overline{P}_1$	

d	$H^*_{(\overline{v}_1)}(BB)$: 1	$H^*_{(\overline{v}_1)}(BB)$: u
-1	$P^*(-2\rho)$	
0	$\overline{P}_0^\vee(-2\rho)$	
1	$\overline{P}_0^\vee(-2\rho)$	
2		
3	$\overline{P}_1$	$P^*(-2\rho)$
4	$\overline{P}_1$	
5	$\overline{P}_1$	
6	$\overline{P}_1$	
7	$\overline{P}_1$	
8	$\overline{P}_1$	

Table 1: BB (left) and the local cohomology (right) by d-diagonal for $n=1$. The H^1–groups are coloured brown.

the $\overline{P}_1$ comes directly from the 3–diagonal of BB, whilst the $P^*(-2\rho)$ comes from the $(2)P$ on the 4–diagonal of BB; the local cohomology is $P^*(-\rho)$, but its diagonal is shifted by -1 since it is a first local cohomology, and because it is by reference to the 2–column the shift is $-\rho$. The top of this module is calculated by reference to the column of $|u|$ (ie the 2–column), and has top in degree $2-(3-2)\sigma-2\rho=-3\sigma$.

We saw in Section 11 that the two modules on the 3–diagonal give a nontrivial additive extension (in degree -3σ) after running the spectral sequence.

12D The special case $n = 2$

Continuing our effort to make patterns visible, we consider $tmf_1(3)^{C_2}_\star$ in this subsection (ie the case $n=2$). With $\mathbb{Z}$ denoting the integers with 3 inverted here, this has

$$P=tmf_1(3)^{C_2}_{*\rho}=\mathbb{Z}[\overline{v}_1,\overline{v}_2],\quad \overline{P}_0=\mathbb{F}_2[\overline{v}_1,\overline{v}_2],\quad \overline{P}_1=\mathbb{F}_2[\overline{v}_2]\quad\text{and}\quad \overline{P}_2=\mathbb{F}_2.$$

See Table 2. Once again, the column labelled u^i is the $2i^{\text{th}}$ column, and shifts along the diagonal have as reference point where this column meets the relevant diagonal.

We take local cohomology, again remembering that $H^s_{\overline{J}_n}$ is shifted down by s, which changes the diagonal by s. For example, on the 7–diagonal, $\overline{P}_2$ comes from the 7–diagonal in BB, whereas the $\overline{P}_0^\vee(-5\rho)$ comes from the 2^{nd} local cohomology of the entry $(\overline{v}_1)\overline{P}_0$ on the 9–diagonal; the local cohomology of $\overline{P}_0$ is $\overline{P}_0^\vee(-4\rho)$, this is shifted by a further -2ρ from the change of diagonal, and $+\rho$ because of the $\overline{v}_1$.

	BB			
d	1	u	u^2	u^3
0	P			
1	$\overline{P}_0$			
2	$\overline{P}_0$			
3	$\overline{P}_1$			
4	$\overline{P}_1$	(2)		
5	$\overline{P}_1$			
6	$\overline{P}_1$			
7	$\overline{P}_2$			
8	$\overline{P}_2$		$(2,\overline{v}_1)P$	
9	$\overline{P}_2$		$(\overline{v}_1)\overline{P}_0$	
10	$\overline{P}_2$		$(\overline{v}_1)\overline{P}_0$	
11	$\overline{P}_2$			
12	$\overline{P}_2$			(2)
13	$\overline{P}_2$			

	$H^*_{(\overline{v}_1,\overline{v}_2)}(BB)$			
d	1	u	u^2	u^3
-2	$P^*(-6\rho)$			
-1	$\overline{P}_0^\vee(-6\rho)$			
0	$\overline{P}_0^\vee(-6\rho)$			
1				
2	$\overline{P}_1^\vee(-4\rho)$	$P^*(-6\rho)$		
3	$\overline{P}_1^\vee(-4\rho)$			
4	$\overline{P}_1^\vee(-4\rho)$			
5	$\overline{P}_1^\vee(-4\rho)$			
6			$\overline{P}_1^\vee(-5\rho)\oplus P^*(-6\rho)$	
7	$\overline{P}_2$		$\overline{P}_0^\vee(-5\rho)$	
8	$\overline{P}_2$		$\overline{P}_0^\vee(-5\rho)$	
9	$\overline{P}_2$			
10	$\overline{P}_2$			$P^*(-6\rho)$
11	$\overline{P}_2$			
12	$\overline{P}_2$			
13	$\overline{P}_2$			

Table 2: BB (left) and the local cohomology (right) by d-diagonal for $n=2$. The H^1–groups are coloured in brown and the H^2–groups in teal.

We will see below that there are nontrivial extensions on the 2– and 10–diagonals, and that there are differentials in the local cohomology spectral sequence from the 7–, 8– and 9–diagonals (differentials go from the d–diagonal to the $(d-1)$–diagonal).

12E Moving from the basic block BB to the negative block NB

Moving from BB to NB only affects the 0 column, where in each case M is replaced by $\ker(M\to\mathbb{F}_2)=(2)M$. In effect, this replaces $\overline{P}_n$ by 0. It also adds on a new (-1)–column of $\overline{P}_n=\mathbb{F}_2$ going up from the σ row. We resist the temptation to display a table for NB explicitly, but note that $NB=\Gamma_{(a)}BB$ as for $k\mathbb{R}$.

12F Gorenstein duality

With the above data in mind, we may consider the d–diagonal BB_d, where the lowest value of d is 0 and the highest is $N=4(2^n-1)$. If we ignore the difference between BB and NB (which is at most $\mathbb{F}_2$ in any degree) we find approximately that BB_d has a relationship to BB_{N-d}, namely something like an equality

$$H^n_{\overline{J}_n}(BB_d)^*=BB_{N-d}.$$

There are various ways in which this is inaccurate and needs to be modified. Firstly, if the local cohomology of BB_d is entirely in cohomological degree $n-\epsilon$ with $\epsilon \neq 0$, there will be a shift of ϵ (if it is in several degrees there is a further complication). Secondly, Anderson duality introduces a shift of one diagonal if applied to torsion modules. Thirdly, we have seen that there may be extensions between these local cohomology groups, sometimes removing $\mathbb{Z}$–torsion. Finally, there may be differentials.

In fact, all of these effects are "small" in the sense that the growth rate along a diagonal is bounded by a polynomial of degree $n-1$. Encouraged by this, if we ignore all of these effects, we see that BB is a Gorenstein module in the sense that the reverse-graded version is equivalent to the dual of its local cohomology:

$$H^n_{\bar{J}_n}(BB)^* = \mathrm{rev}(BB).$$

This is rather as if there is a cofibre sequence

$$S \to BP\mathbb{R}\langle n\rangle \to Q$$

where S is Gorenstein and Q is a Poincaré duality algebra of formal dimension $N = 2(1-\sigma)(2^n-1)$.

13 The local cohomology spectral sequence for $tmf_1(3)$

We examine the local cohomology spectral sequence and Gorenstein duality in more detail for $tmf_1(3)$. Actually, our calculations are equally valid for all forms of $BP\mathbb{R}\langle 2\rangle$, but we prefer the more evocative name $tmf_1(3)$ of the most prominent example. More of the general features are visible for $tmf_1(3)$ than for $k\mathbb{R}$.

As usual we will implicitly localize everywhere at 2 (although for $tmf_1(3)$ itself it would actually suffice to just invert 3).

13A The local cohomology spectral sequence

We make explicit the implications for the coefficient ring, both computationally and structurally. Writing $\star$ for $RO(C_2)$–grading as usual, the spectral sequence takes the following form.

Proposition 13.1 *There is a spectral sequence of $tmf_1(3)^{C_2}_\star$–modules*

$$H^*_{\bar{J}_n}(tmf_1(3)^{C_2}_\star) \implies \Sigma^{-8-2\sigma}\pi^{C_2}_\star(\mathbb{Z}^{tmf_1(3)}).$$

The homotopy of the Anderson dual is easily calculated:

$$0 \to \mathrm{Ext}_{\mathbb{Z}}(tmf_1(3)^{C_2}_{-\alpha-1}, \mathbb{Z}) \to \pi^{C_2}_\alpha \mathbb{Z}^{tmf_1(3)} \to \mathrm{Hom}_{\mathbb{Z}}(tmf_1(3)^{C_2}_{-\alpha}, \mathbb{Z}) \to 0.$$

The local cohomology spectral sequence has some nontrivial differentials.

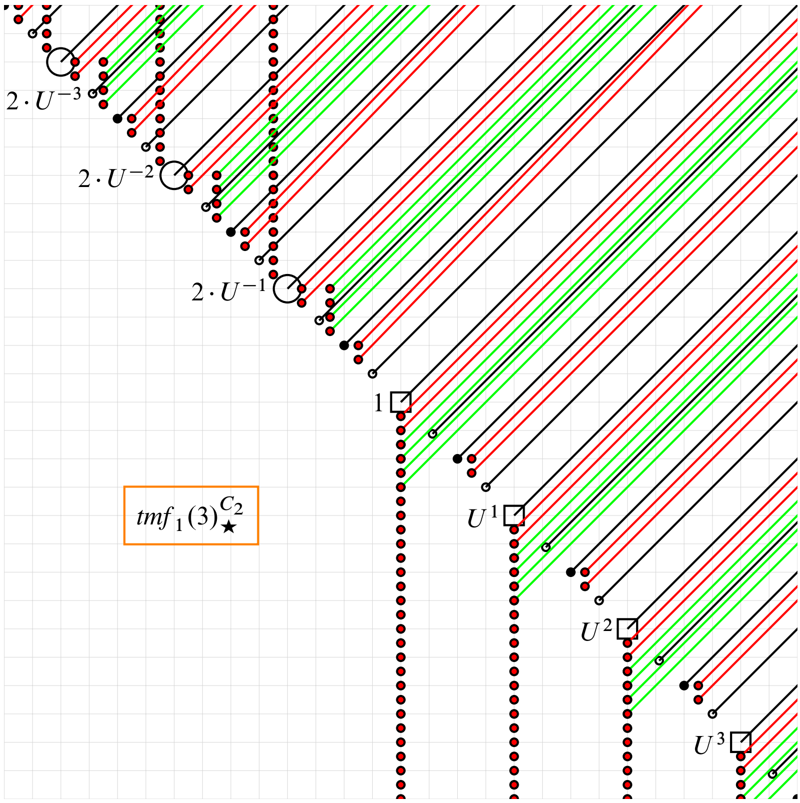


Figure 4: The homotopy of $tmf_1(3)$

13B The ring $tmf_1(3)^{C_2}_\star$

The ring $tmf_1(3)^{C_2}_\star$ is approximately calculated in [27] and more precisely described as

$$BB[U] \oplus U^{-1} NB[U^{-1}]$$

as at the end of Section 4B with $n = 2$. We already tabulated BB in Section 12D, but we want also want to display a bigger chart of $\pi^{C_2}_\star tmf_1(3)$ as Figure 4 to give the reader a feeling of how the blocks piece together.

A black diagonal line means a copy of P when it starts in a box, a copy of $(2)P$ when it starts in a small circle, a copy of $(2, \overline{v}_1)P$ when it starts in a dot and a copy of $(2, \overline{v}_1, \overline{v}_2)$ when it starts in a big circle. In Figure 4, a red diagonal line means a copy of $\overline{P}_0$ and a green diagonal line a copy of $\overline{P}_1$. A red dot is a copy of $\mathbb{F}_2 = \overline{P}_2$.

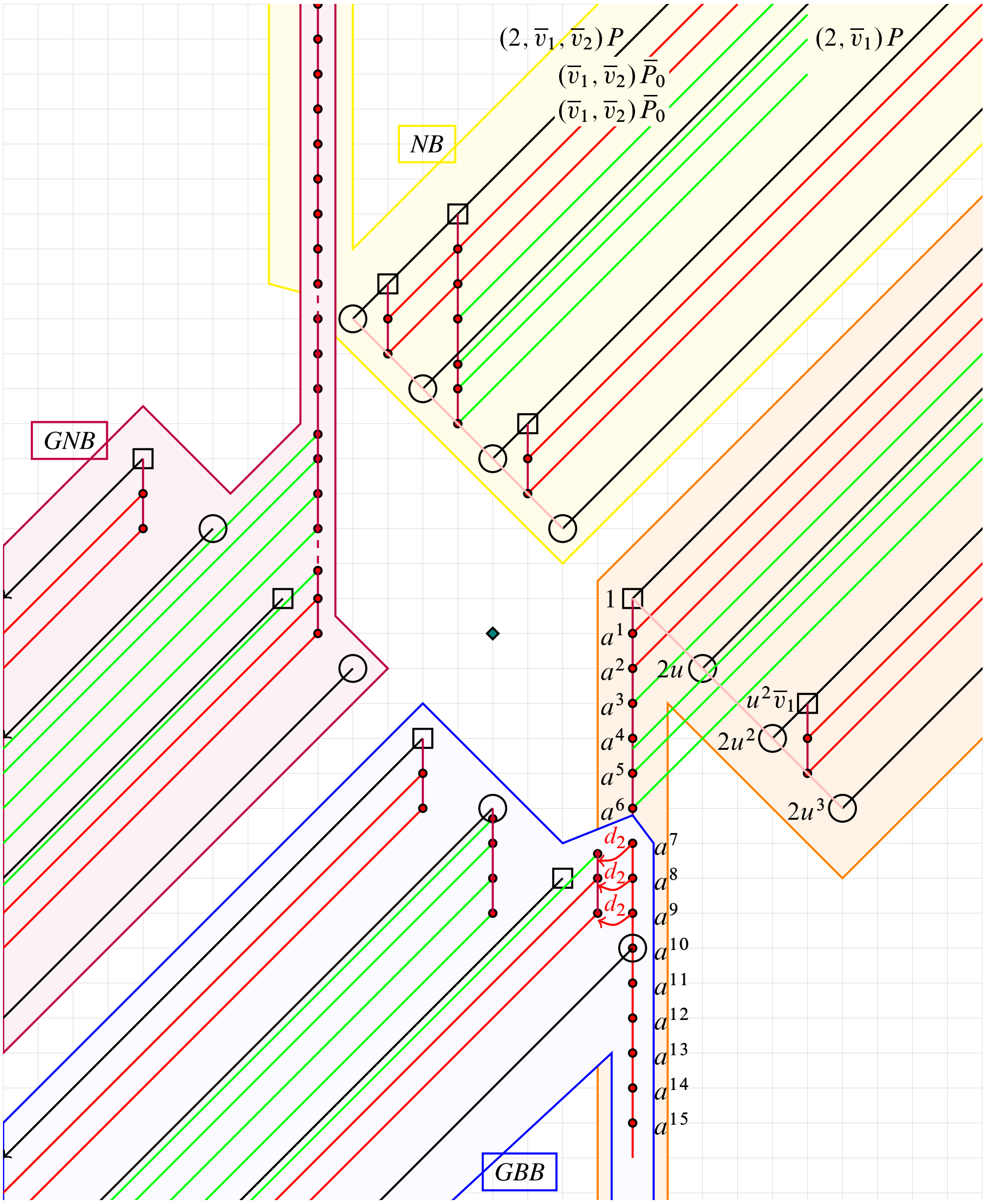


Figure 5: Gorenstein duality for $tmf_1(3)$

13C Local cohomology

We are calculating local cohomology with respect to the ideal $\overline{J}_2 = (\overline{v}_1, \overline{v}_2)$ so that we only need to consider $tmf_1(3)^{C_2}_{\star}$ as a $\mathbb{Z}[\overline{v}_1, \overline{v}_2]$–module. As such it is a sum of suspensions of the blocks BB and NB, so we just need to calculate the local cohomology of these. This was described in Section 12 above. Here we will simply describe the extensions and the behaviour of the local cohomology spectral sequence.

The basis of this discussion are the tables of BB and GBB from Section 12D together with the analogues for NB and GNB. Although these are organized by diagonal, Figure 5 displaying $BB, GBB, U^{-1}NB$ and $U^{-1}GNB$ may help visualize the way the modules are distributed along each diagonal. The vertical lines denote multiplication by a and the dashed vertical line is an exotic multiplication by a that is not visible on the level of local cohomology. The green diamond does not denote a class, but marks the point one has to reflect (nontorsion classes) at to see Anderson duality. Torsion classes are shifted after reflection by -1 (ie one step horizontally to the left).

The strategy is to take the known subquotients from the local cohomology calculation, and resolve the extension problems using Gorenstein duality.

Proposition 13.2 *We have an isomorphism*

$$\pi_\star^{C_2}\Gamma_{\bar{J}_2} tmf_1(3) \cong GBB[U] \oplus U^{-1}GNB[U^{-1}],$$

where GBB and GNB are described in the following. We will simultaneously describe what differentials and extensions in the local cohomology spectral sequence caused the passage from $H^*_{\bar{J}_2}(BB)$ and $H^*_{\bar{J}_2}(NB)$ to GBB and GNB respectively.

(i) *The $\mathbb{Z}[\bar{v}_1, \bar{v}_2]$–modules along the diagonals in GBB are as in Table 3 (left). There are three nontrivial differentials*

$$d_2\colon H^0_{\bar{J}_2}(BB) \to H^2_{\bar{J}_2}(BB)$$

from the groups at $-7\sigma, -8\sigma, -9\sigma$ to the groups at $-7\sigma-1, -8\sigma-1, -9\sigma-1$, which have affected the values on the 6–, 7–, 8– and 9–diagonals in Table 3 (left).

The extensions

$$0 \to P^* \to [(2, \bar{v}_1)P]^* \to \mathbb{F}_2[\bar{v}_2]^\vee \to 0$$

on the 2–diagonal and the 6–diagonal are Anderson dual to the defining short exact sequence

$$0 \to (2, \bar{v}_1)P \to P \to \mathbb{F}_2[\bar{v}_2] \to 0$$

in the following sense: The Anderson dual of the latter exact sequence is a triangle

$$\mathbb{F}_2[\bar{v}_2]^* \to P^* \to [(2, \bar{v}_1)P]^* \to \Sigma\mathbb{F}_2[\bar{v}_2]^* \cong \mathbb{F}_2[\bar{v}_2]^\vee$$

which induces (on homology) the extensions above. The extension

$$0 \to P^* \to [(2, \bar{v}_1, \bar{v}_2)P]^* \to \mathbb{F}_2 \to 0$$

	GBB			GNB	
d	module	top degree	d	module	top degree
			$-k \le -3$	$\mathbb{F}_2$	$-1-k\sigma$
-2	P^*	$-6-4\sigma$	-2	$P^*, \mathbb{F}_2$	$-6-4\sigma, -1+\sigma$
-1	$\overline{P}_0^\vee$	$-6-5\sigma$	-1	$\overline{P}_0^\vee, \mathbb{F}_2$	$-6-5\sigma, -1+0\sigma$
0	$\overline{P}_0^\vee$	$-6-6\sigma$	0	$\overline{P}_0^\vee, \mathbb{F}_2$	$-6-6\sigma, -1-\sigma$
1	0		1	$\mathbb{F}_2$	$-1-2\sigma$
2	$[(2,\overline{v}_1)P]^*$	$-4-6\sigma$	2	$P^*, \overline{P}_1^\vee$	$-4-6\sigma, -1-3\sigma$
3	$\overline{P}_1^\vee$	$-4-7\sigma$	3	$\overline{P}_1^\vee$	$-1-4\sigma$
4	$\overline{P}_1^\vee$	$-4-8\sigma$	4	$\overline{P}_1^\vee$	$-1-5\sigma$
5	$\overline{P}_1^\vee$	$-4-9\sigma$	5	$\overline{P}_1^\vee$	$-1-6\sigma$
6	$[(2,\overline{v}_1)P]^*$	$-2-8\sigma$	6	$[(2,\overline{v}_1)P]^*$	$-1-7\sigma$
7	$(\overline{v}_1,\overline{v}_2)\overline{P}_0$	$-2-9\sigma$	7	$\overline{P}_0^\vee$	$-1-8\sigma$
8	$(\overline{v}_1,\overline{v}_2)\overline{P}_0$	$-2-10\sigma$	8	$\overline{P}_0^\vee$	$-1-9\sigma$
9	0		9	0	
10	$[(2,\overline{v}_1,\overline{v}_2)P]^*$	$0-10\sigma$	10	P^*	$0-10\sigma$
$10+k \ge 11$	$\mathbb{F}_2$	$0-(10+k)\sigma$			

Table 3: $\mathbb{Z}[\overline{v}_1,\overline{v}_2]$–modules as described in Proposition 13.2

on the 10*–diagonal is Anderson dual to the short exact sequence*

$$0 \to (2,\overline{v}_1,\overline{v}_2)P \to P \to \mathbb{F}_2 \to 0.$$

(ii) *The* $\mathbb{Z}[\overline{v}_1,\overline{v}_2]$*–modules along the diagonals in* GNB *are as in Table 3* (*right*) (*take the direct sum of the two entries for the* (-2)*–,* (-1)*–,* 0*– and* 2*–diagonals*)*. The extension*

$$0 \to P^* \to [(2,\overline{v}_1)P]^* \to \mathbb{F}_2[v_2]^\vee \to 0$$

on the 6*–diagonal is Anderson dual to the short exact sequence*

$$0 \to (2,\overline{v}_1)P \to P \to \mathbb{F}_2[v_2] \to 0.$$

Proof We first note that the contributions from the different blocks do not interact. Indeed, the only time that different blocks give contributions in the same degree comes from the $\mathbb{F}_2[a]$ towers of *BB*, and one class in that degree is divisible by $\overline{v}_1$ or $\overline{v}_2$ and not killed by both $\overline{v}_1$ and $\overline{v}_2$. We may therefore consider the blocks entirely separately.

The block *GBB* comes from the local cohomology of *BB* in the sense that there is a spectral sequence

$$H^*_{\overline{J}_2}(BB) \implies GBB.$$

	$H^*_{\bar{J}_2}(NB)$			
d	1	u^2	u^4	u^6
-2	$P^*(-6\rho)$			
-1	$\bar{P}_0^\vee(-6\rho) \oplus \bar{P}_2$			
0	$\bar{P}_0^\vee(-6\rho) \oplus \bar{P}_2$			
1	$\bar{P}_2$			
2	$\bar{P}_1^\vee(-4\rho)$	$P^*(-6\rho)$		
3	$\bar{P}_1^\vee(-4\rho)$			
4	$\bar{P}_1^\vee(-4\rho)$			
5	$\bar{P}_1^\vee(-4\rho)$			
6			$\bar{P}_1^\vee(-5\rho) \oplus P^*(-6\rho)$	
7			$\bar{P}_0^\vee(-5\rho)$	
8			$\bar{P}_0^\vee(-5\rho)$	
9				
10				$P^*(-6\rho)$
11				
12				
13				

Table 4: Local cohomology for $n = 2$ from the proof of Proposition 13.2. Again, the H^1–groups are coloured in brown and the H^2–groups in teal.

Thus there is a filtration

$$GBB = GBB^0 \supseteq GBB^1 \supseteq GBB^2 \supseteq GBB^3 = 0$$

with

$$0 \to GBB^0/GBB^1 \to H^0_{\bar{J}_2}(BB) \xrightarrow{d_2} \Sigma^{-1} H^2_{\bar{J}_2}(BB) \to \Sigma^1 GBB^2 \to 0$$

and

$$GBB^1/GBB^2 \cong \Sigma^{-1} H^1_{\bar{J}_2}(BB).$$

The block GNB comes from the local cohomology of NB in a precisely analogous way.

Most questions about module structure over $BB[U]$ are resolved by degree. The remaining issues are resolved by using Gorenstein duality.

Referring to the table for $H^*_{\bar{J}_2}(BB)$ in Section 12D, the first potential extension is on the 2–diagonal. Using Gorenstein duality to compare with $NB_{\delta=8}$ we see that the actual extension on the 2–diagonal of GBB is

$$0 \to P^* \to [(2, \bar{v}_1)P]^* \to \bar{P}_1^\vee \to 0,$$

δ	δ' s.t. $H^*_{\bar{J}_2}(BB_\delta)^* \sim NB_{\delta'}$	δ	δ' s.t. $H^*_{\bar{J}_2}(NB_\delta)^* \sim BB_{\delta'}$
0	12	0	12
1	10	1	10
2	9	2	9
3	8	3	8
4	8, 6	4	8, 6
5	5	5	5
6	4	6	4
7	2	7	.
8	4, 3	8	4
9	2	9	2
10	1, 0	10	1
11	0	11	.
12	0	12	0

Table 5: Diagonal contributions from Remark 13.3(i)

where we have shifted the modules so they all have top degree 0. There is an additive extension on the 10–diagonal by reference to the Anderson dual. Finally the three nonzero d_2 differentials from $-1-k\sigma$ for $k=7,8$ and 9 are necessary for connectivity (this removes the need to discuss the possible extensions on the 7– and 8–diagonals).

The situation is rather similar for GNB. We will not explicitly display NB since the only effect (apart from the addition of $\mathbb{F}_2[a]^\vee$) is on the first column, where a module is replaced by the kernel of a surjection to $\mathbb{F}_2$. It is perhaps worth displaying $H^2_{\bar{J}_2}(NB)$, where we leave out the big $\mathbb{F}_2[a]^\vee$–tower in $H^0_{\bar{J}_2}NB$. See Table 4. In this case, all extensions are split, except for the one on the 6–diagonal and there are no differentials. The a multiplications in the $\mathbb{F}_2[a]^\vee$ tower are clear from Gorenstein duality and the a–tower $\mathbb{F}_2[a]$ in BB. □

Remark 13.3 (i) In Table 5, we summarize the way a diagonal BB_δ contributes to NB as in

$$H^*_{\bar{J}_2}(BB_\delta)^* \sim NB_{\delta'}$$

as sketched in Section 12F. Because most of the modules are 2–torsion the most common pairing is between δ and $11-\delta$ rather than between δ and $12-\delta$ as happens for the main U–power diagonals.

(ii) We also note as before that since $NB = \Gamma_{(a)}BB$, we have

$$\Sigma^{6+4\sigma}\Gamma_{(\bar{v}_1,\bar{v}_2)}BB \sim (\Gamma_{(a)}BB)^*$$

(where we have written $\sim$ rather than $\simeq$ in recognition of the differentials) and

$$\Sigma^{6+4\sigma}\Gamma_{(\bar{v}_1,\bar{v}_2,a)}BB \simeq BB^*,$$

with the second stating that BB is Gorenstein of shift $-6-4\sigma$ for the ideal $(\bar{v}_1, \bar{v}_2, a)$.

Appendix: The computation of $\pi_\star^{C_2} BP\mathbb{R}$

Our main goal in this appendix is to compute the homotopy fixed point spectral sequence for $BP\mathbb{R}$ and hence for $M\mathbb{R}$. All the results in this appendix and the essential idea of the argument for Proposition A.2 are contained in [18] (see especially Formula 4.16). We just rearranged their arguments and added some details. Our argument for the multiplicative extensions might be considered new though. We have strived for elementary and short proofs though they retain some computational complexity. We hope this is helpful for the reader to understand this crucial computation. Note that even before Hu and Kriz, the computation of $\pi_\star^{C_2} BP\mathbb{R}$ was announced in [3].

We will work throughout 2–locally. As before, we denote by ρ the regular real C_2–representation and by σ the sign representation. We need a few facts, first proven by Araki:

(1) If E is a real-oriented spectrum, then $E_{C_2}^\star(\mathbb{C}P^\infty) \cong E_{C_2}^\star[\![u]\!]$ with $|u| = -\rho$ and $E_{C_2}^\star(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \cong E_{C_2}^\star[\![1\otimes u, u\otimes 1]\!]$. This induces a formal group law on $\pi_{*\rho}^{C_2} E$ and the forgetful map $\pi_{*\rho}^{C_2} E \to \pi_{2*}^e E$ maps it to the usual formal group law from the complex orientation of E. [18, Theorem 2.10]

(2) Thus, we get a ring map $\pi_{2*}^e MU \to \pi_{*\rho}^{C_2} M\mathbb{R}$ from the Lazard ring such that $\pi_{2*}^e MU$ is a retract of $\pi_{*\rho}^{C_2} M\mathbb{R}$. For every class in $x \in \pi_{2*}MU$, we have thus a corresponding class $\bar{x} \in \pi_{*\rho}^{C_2} M\mathbb{R}$. [18, Proposition 2.27]

(3) There is a splitting $M\mathbb{R}_{(2)} \simeq \bigoplus_{m_i} \Sigma^{m_i\rho} BP\mathbb{R}$, where the underlying spectrum of $BP\mathbb{R}$ agrees with BP. This splitting corresponds on geometric fixed points to the splitting $MO \simeq \bigoplus_{m_i} \Sigma^{m_i} H\mathbb{F}_2$. [18, Theorem 2.33]

Define $a\colon S^0 \to S^\sigma$ as before to be the inclusion of the points 0 and ∞; we will denote the image of a in $\pi_\star M\mathbb{R}$ and $\pi_\star BP\mathbb{R}$ by the same symbol. The class a has degree $-\sigma = 1-\rho$.

Proposition A.1 *We have $a^{2^{n+1}-1}\bar{v}_n = 0$ in $\pi_\star^{C_2} M\mathbb{R}$.*

Proof We have a fibre sequence

$$(EC_2)_+ \otimes M\mathbb{R} \to M\mathbb{R} \to \widetilde{E}C_2 \otimes M\mathbb{R}.$$

First, we claim that the image of $\bar{v}_n$ under $M\mathbb{R} \to \widetilde{E}C_2 \otimes M\mathbb{R}$ is zero. Indeed, as a is invertible on $\widetilde{E}C_2 \otimes M\mathbb{R}$, the formal group law on $\pi_{*\rho}^{C_2}(\widetilde{E}C_2 \otimes M\mathbb{R})$ agrees with that on $\pi_*^{C_2}(\widetilde{E}C_2 \otimes M\mathbb{R}) = \pi_* MO$, which is additive. Therefore, the map

$$MU_{2*} \to \pi_{*\rho}^{C_2} M\mathbb{R} \to \pi_{*\rho}^{C_2} \widetilde{E}C_2 \otimes M\mathbb{R}$$

sends all v_n to zero. Thus, $\bar{v}_n$ and hence also $a^{2^{n+1}-1}\bar{v}_n$ are in the image of the map

$$(EC_2)_+ \otimes M\mathbb{R} \to M\mathbb{R}.$$

Observe that

$$|a^{2^{n+1}-1}\bar{v}_n| = -(2^{n+1}-1)\sigma + (2^n - 1)(1+\sigma) = 2^n - 1 - 2^n\sigma.$$

We claim that $\pi_{2^n-1-2^n\sigma}^{C_2}((EC_2)_+ \otimes M\mathbb{R})$ is zero. Indeed, we have

$$\pi_{2^n-1-2^n\sigma}^{C_2}((EC_2)_+ \otimes M\mathbb{R}) \cong \pi_{2^n-1}(\Sigma^{2^n\sigma} M\mathbb{R})_{hC_2}.$$

This can be computed by the homotopy orbit spectral sequence

$$H_p(C_2; \pi_q \Sigma^{2^n\sigma} M\mathbb{R}) \implies \pi_{p+q}(\Sigma^{2^n\sigma} M\mathbb{R})_{hC_2}.$$

But $\pi_q \Sigma^{2^n\sigma} M\mathbb{R} = 0$ for $q < 2^n$, so $\pi_{2^n-1}(\Sigma^{2^n\sigma} M\mathbb{R})_{hC_2} = 0$. Thus, we see that $a^{2^{n+1}-1}\bar{v}_n = 0$ in $\pi_\star^{C_2} M\mathbb{R}$. □

For a C_2–spectrum X, the $RO(C_2)$ graded homotopy fixed point spectral sequence is defined by combining the homotopy fixed point spectral sequences

$$E_2^{p,q}(r) = H^q(C_2, \pi_{p+q}(X \wedge S^{-r\sigma})) \Longrightarrow \pi_p^{C_2}((X \wedge S^{-r\sigma})^{hC_2}) \cong \pi_{p+r\sigma}^{C_2}(X^{(EC_2)_+})$$

into a single spectral sequence with differential

$$d_n \colon E_n^{p,q}(r) \to E_n^{p-1,q+n}(r).$$

Note that we use an Adams grading convention here. We will often call $p + r\sigma$ the *degree* of an element.

The $RO(C_2)$–graded homotopy fixed point spectral sequence (HFPSS) for $BP\mathbb{R}$ has E_2–term

$$\mathbb{Z}_{(2)}[a, u^{\pm 1}, \bar{v}_1, \bar{v}_2, \dots]/2a$$

with

$$|a| = (-\sigma, 1), \quad |u| = (2 - 2\sigma, 0) \quad \text{and} \quad |\bar{v}_i| = ((2^i - 1)\rho, 0).$$

This can be seen, for example, by the identification with the Bockstein spectral sequence for a discussed in [27, Lemma 4.8]. As $BP\mathbb{R}$ is a retract of $M\mathbb{R}_{(2)}$, it has the structure of a (homotopy) ring spectrum and thus the $RO(C_2)$–graded homotopy fixed point spectral sequence is multiplicative by [27, Section 2.3].

By the discussion above, a and the $\bar{v}_i$ are permanent cycles. As $a^{2^{n+1}-1}\bar{v}_n$ is zero, it must be hit by a differential. This is the crucial ingredient for the following central proposition. It is fully formal in the sense that we do not need any other input in addition to the things we already discussed; we argue just with the form of the spectral sequence. We will set $\bar{v}_0 = 2$ for convenience.

Proposition A.2 *In the HFPSS for $BP\mathbb{R}$, we have $E_{2^n} = E_{2^{n+1}-1}$, and it is the subalgebra of*

$$E_2/(a^3\bar{v}_1, \dots, a^{2^n-1}\bar{v}_{n-1})$$

generated by a, $u^{\pm 2^{n-1}}$, the $\bar{v}_i$ for $i \geq 0$ and by the $\bar{v}_i u^{2^i j}$ for $i < n-1$ and $j \in \mathbb{Z}$.

Proof We prove it by induction. It is obviously true for $n = 1$ by the checkerboard phenomenon; indeed, for all generators of the E_2–term in degree $(a + b\rho, q)$ we have $a + q$ even.

Now assume it to be true for a given n. First, we will show that $d_{2^{n+1}-1}(u^{2^{n-1}}) = a^{2^{n+1}-1}\bar{v}_n$. Indeed, as $a^{2^{n+1}-1}\bar{v}_n$ is nonzero in $E_{2^{n+1}-1}$, it must be hit by a $d_{2^{n+1}-1}$. Its source x is in the zero-line in degree $2^{n+1} - 2^n\rho$. As the zero-line in E_2 is generated by u of degree $4 - 2\rho$ and by the $\bar{v}_i$ in regular representation degrees, we see that the exponent of u in x must be 2^{n-1}, so there is no room for further $\bar{v}_i$. Thus, $d_{2^{n+1}-1}(u^{2^{n-1}}) = a^{2^{n+1}-1}\bar{v}_n$.

Next, we want to show that $d_q(\bar{v}_i u^{2^i j}) = 0$ for $2^{n+1} - 1 \leq q < 2^{n+2} - 1$ and $i < n$. Write $d_q(\bar{v}_i u^{2^i j}) = a^q x$. The degree of x is

$$(2^i - 1)\rho + 2^i j(4 - 2\rho) - q(1 - \rho) - 1 = (2^{i+2} j - q - 1) + (2^i - 2^{i+1} j + q - 1)\rho.$$

Thus, $x = u^{2^i j - (q+1)/4}\bar{v}$, where $\bar{v}$ is a polynomial in the $\bar{v}_\nu$. The degree of $\bar{v}$ is $(2^i - 2 + \frac{1}{2}(q+1))\rho$. As $\frac{1}{2}(q+1) < 2^{n+1}$, we have

$$|\bar{v}| < |\bar{v}_{n+1}^2| < |\bar{v}_r|$$

for $r \geq n+2$. Thus, no monomial in $\bar{v}$ is divisible by $\bar{v}_{n+1}^2$ or $\bar{v}_r$. Assume that $|\bar{v}| = |\bar{v}_{n+1}|$. Then $\frac{1}{2}(q+1) = 2^{n+1} - 1 + 2 - 2^i = 2^{n+1} - 2^i + 1$, which is odd; but then $\frac{1}{4}(q+1) \notin \mathbb{Z}$, which is a contradiction. Thus, every monomial in $\bar{v}$ is divisible by some $\bar{v}_k$ for some $k \leq n$ as $\bar{v} \neq 1$ for degree reasons. But $a^q \bar{v}_k = 0$ in E_q. Thus, also $a^q x = 0$ in E_q.

Similarly, write $d_q(u^{2^n}) = a^q x$ for $2^{n+1} - 1 \leq q < 2^{n+2} - 1$ and assume that this is nonzero. The degree of x is

$$2^n(4 - 2\rho) - q(1 - \rho) - 1 = (2^{n+2} - q - 1) + (q - 2^{n+1})\rho.$$

Thus, we can write x in E_2 as $u^{2^n-(q+1)/4}\bar{v}$, where $\bar{v}$ is a polynomial in the $\bar{v}_\nu$. The degree of $\bar{v}$ is $\frac{1}{2}(q-1) < 2^{n+1}-1$. Thus, no monomial in $\bar{v}$ can be divisible by $\bar{v}_r$ for $r \geq n+1$. Thus, every monomial in $\bar{v}$ is divisible by some $\bar{v}_k$ for some $k \leq n$ as $\bar{v} \neq 1$ for degree reasons. But $a^q \bar{v}_k = 0$ in E_q. Thus, $d_q(u^{2^n}) = 0$.

By the Leibniz rule, this implies the proposition. □

Before we solve the multiplicative extension issues, we need a technical lemma.

Lemma A.3 *Assume that there is an element $a^k u^l \bar{v} \neq 0$ above the zero line in the E_∞–term of the $RO(C_2)$–graded HFPSS for $BP\mathbb{R}$ with $\bar{v}$ a monomial in the $\bar{v}_\nu$ and in the same degree as $\bar{v}_i \bar{v}_m u^{2^m j}$. Let p be the minimal index such that $\bar{v}_p$ divides $\bar{v}$ (which we will show to exist). Then $i > p+m$.*

Proof The degree of $\bar{v}_i \bar{v}_m u^{2^m j}$ is

$$2^m j(4-2\rho) + (2^i - 1 + 2^m - 1)\rho = 2^{m+2} j + (2^i + 2^m - 2^{m+1} j - 2)\rho.$$

Let $a^k u^l \bar{v} \neq 0$ be an element in E_∞ in this degree with $\bar{v}$ a monomial in the $\bar{v}_\nu$ of degree $n\rho$ and assume that $k > 0$. (In the following, we will use the notation $\|\bar{v}_p\| = |\bar{v}_p|/\rho$ so that $\|\bar{v}\| = n$.) We get

$$\begin{aligned} 4l + k &= 2^{m+2} j, \\ n - 2l - k &= 2^i + 2^m - 2^{m+1} j - 2. \end{aligned}$$

This implies $n = 2^i + 2^m - 2 + \frac{1}{2}k$. We see that $n \neq 0$. Let p be the minimal index such that $\bar{v}_p | \bar{v}$. Then $2^p | l$ and we set $c = l/2^p$. Then $k = 2^{m+2} j - 2^{p+2} c$. Due to the relation $a^{2^{p+1}-1}\bar{v}_p = 0$, we have $k \leq 2^{p+1} - 2$ and thus $m+2 \leq p$ (as else $2^{p+1} | k$ and thus $k \geq 2^{p+1}$). In particular, 2^{m+1} divides $\frac{1}{2}k$. Now observe that $n \geq \|\bar{v}_p\| = 2^p - 1$, so

$$2^i + 2^m - 1 \geq 2^p - \tfrac{1}{2}k.$$

As $k \leq 2^{p+1} - 2$, the right-hand side is positive; as it is also divisible by 2^{m+1} it is thus it is at least 2^{m+1}. We see that $i \geq m+1$. Thus $n \equiv 2^m - 2 \mod 2^{m+1}$. As $\|\bar{v}_q\| \equiv -1 \mod 2^{m+1}$ for $q \geq p > m+1$, we see that the total exponent of $\bar{v}$ (ie the degree of $\bar{v}$ as a monomial in the $\bar{v}_\nu$) must be $\equiv 2^m + 2 \mod 2^{m+1}$. In particular, $n \geq \|\bar{v}_p\|(2^m + 2) = (2^p - 1)(2^m + 2)$. Thus,

$$\tfrac{1}{2}k = n - 2^i - 2^m + 2 \geq 2^{p+m} - 2^i + (2^{p+1} - 2^{m+1}).$$

If $p + m \geq i$, then the right-hand side is at least 2^p, which would be a contradiction. Thus $i > p+m$. □

Now, we are ready to prove the main result of the appendix. Note that [18, Theorem 4.11] gives a different relation than our last one; our relation implies their relation, but not vice versa. Note also that our arguments for the multiplicative relations are completely algebraic (using the form of the spectral sequence), while [18] uses additionally a C_2–equivariant Adams spectral sequence.

Theorem A.4 *The ring $\pi_\star^{C_2}BP\mathbb{R}$ is isomorphic to the E_∞–term of the homotopy fixed point spectral sequence above, ie to the subalgebra of*

$$\mathbb{Z}_{(2)}[a, \bar{v}_i, u^{\pm 1}]/(2a, \bar{v}_i a^{2^{i+1}-1})$$

(where i runs over all positive integers) generated by $\bar{v}_m(n) = u^{2^m n}\bar{v}_m$ (with $m, n \in \mathbb{Z}$ and $m \geq 0$) and a with $\bar{v}_0 = 2$. Consequently, it is the quotient R of the ring

$$\mathbb{Z}_{(2)}[a,\ \bar{v}_m(n) \mid m \geq 0,\ n \in \mathbb{Z}]$$

by the relations

$$\begin{aligned}
\bar{v}_0(0) &= 2, \\
a^{2^{m+1}-1}\bar{v}_m(n) &= 0, \\
\bar{v}_i(j)\bar{v}_m(n) &= \bar{v}_i\bar{v}_m(2^{i-m}j + n) \quad \text{for } i \geq m,
\end{aligned}$$

with $\bar{v}_i = \bar{v}_i(0)$. Here, $|a| = 1 - \rho$ and $|\bar{v}_m(n)| = 2^{m+2}n + (2^m - 1 - 2^{m+1}n)\rho$.

Proof It suffices to show that the expression above computes the homotopy fixed points $\pi_\star^{C_2}BP\mathbb{R}^{(EC_2)_+}$. Indeed, Proposition A.2 implies that $(a^{-1}BP\mathbb{R}^{(EC_2)_+})^{C_2} \simeq H\mathbb{F}_2$, so the map $BP\mathbb{R}^{\Phi C_2} \to BP\mathbb{R}^{tC_2}$ is an equivalence and hence also $BP\mathbb{R} \to BP\mathbb{R}^{(EC_2)_+}$ by the Tate square.

Set $\bar{v}_0(0) = 2$. By Proposition A.2, the classes $u^{2^m n}\bar{v}_m$ are permanent cycles in the HFPSS; choose element $\bar{v}_m(n) \in \pi_\star^{C_2}BP\mathbb{R}^{(EC_2)_+}$ representing them. Again by Proposition A.2, the $\bar{v}_m(n)$ generate together with a the E_∞–term of the HFPSS. Thus, we get a surjective map $R \to E_\infty$. The third relation defining R allows to define a normal form: Every monomial in the $\bar{v}_i(j)$ equals in R an element of the form $\bar{v}\,\bar{v}_m(k)$, where $\bar{v}$ is a monomial in the $\bar{v}_i$ and m was the smallest index of all $\bar{v}_i(j)$. Thus, two monomials in the $\bar{v}_i(j)$ are equal in R if they are equal in E_∞; hence, the map $R \to E_\infty$ is also injective.

We now check that the relations are also satisfied in $\pi_\star^{C_2}BP\mathbb{R}^{(EC_2)_+}$. This is clear or was already shown for the first two relations. Let now i be the least number such that $m \leq i$ and

$$\bar{v}_i(j)\bar{v}_m(n) \neq \bar{v}_i\bar{v}_m(2^{i-m}j + n)$$

for some j, m, n if such an i exists. The difference must be detected by a class $a^k u^l \bar{v}$, where $\bar{v}$ is a polynomial in the $\bar{v}_\nu$. Let p the minimal index such that every monomial in $\bar{v}$ is divisible by a $\bar{v}_r$ with $r \le p$. From Lemma A.3, we know that $p \le i-1$ (and in particular $i \ge 1$). Thus,

$$\bar{v}_i(j)\bar{v}_m(n)\bar{v}_{i-1} \neq \bar{v}_i\bar{v}_m(2^{i-m}j+n)\bar{v}_{i-1}$$

as their difference is detected by a nonzero class $a^k u^l \bar{v}\bar{v}_{i-1}$ (indeed, this could only be zero if $k \ge 2^i - 1$, but $k < 2^{p+1}-1$). By the minimality of i, we have

$$\bar{v}_m(2^{i-m}j+n)\bar{v}_{i-1} = \bar{v}_{i-1}(2j)\bar{v}_m(n).$$

In addition, $\bar{v}_i\bar{v}_{i-1}(2j) = \bar{v}_i(j)\bar{v}_{i-1}$ because there is no element of higher filtration in the same degree as $\bar{v}_{i-1}\bar{v}_i(j)$ by Lemma A.3. The last two equations combine to the chain of equalities

$$\begin{aligned}\bar{v}_i(j)\bar{v}_m(n)\bar{v}_{i-1} &= \bar{v}_i\bar{v}_{i-1}(2j)\bar{v}_m(n)\\ &= \bar{v}_i\bar{v}_m(2^{i-m}j+n)\bar{v}_{i-1}.\end{aligned}$$

This is a contradiction to the inequality above. Thus,

$$\bar{v}_i(j)\bar{v}_m(n) = \bar{v}_i\bar{v}_m(2^{i-m}j+n)$$

is always true for $i \ge m$. □

Remark A.5 We remark that all the work above for the multiplicative extensions was actually necessary. For example, we get from the homotopy fixed point spectral sequence only that $\bar{v}_5\bar{v}_1(1) - \bar{v}_5(1)\bar{v}_1(-15)$ has filtration at least 1. But there are indeed classes in this degree of higher filtration, for example, $a^8\bar{v}_3^3\bar{v}_4$.

References

[1] **D W Anderson**, *Universal coefficient theorems for K-theory*, mimeographed notes, Univ. of California, Berkeley (1969)

[2] **S Araki**, *Orientations in τ–cohomology theories*, Japan. J. Math. 5 (1979) 403–430 MR

[3] **S Araki**, **M Murayama**, *τ–cohomology theories*, Japan. J. Math. 4 (1978) 363–416 MR

[4] **M F Atiyah**, *K–theory and reality*, Quart. J. Math. Oxford Ser. 17 (1966) 367–386 MR

[5] **R Banerjee**, *On the $ER(2)$–cohomology of some odd-dimensional projective spaces*, Topology Appl. 160 (2013) 1395–1405 MR

[6] **E H Brown, Jr**, **M Comenetz**, *Pontrjagin duality for generalized homology and cohomology theories*, Amer. J. Math. 98 (1976) 1–27 MR

[7] **R R Bruner**, **J P C Greenlees**, *Connective real K–theory of finite groups*, Mathematical Surveys and Monographs 169, Amer. Math. Soc., Providence, RI (2010) MR

[8] **D Dugger**, *An Atiyah–Hirzebruch spectral sequence for KR–theory*, K-Theory 35 (2005) 213–256 MR

[9] **W G Dwyer**, **J P C Greenlees**, **S Iyengar**, *Duality in algebra and topology*, Adv. Math. 200 (2006) 357–402 MR

[10] **A D Elmendorf**, **I Kriz**, **M A Mandell**, **J P May**, *Rings, modules, and algebras in stable homotopy theory*, Math. Surv. Monogr. 47, Amer. Math. Soc., Providence, RI (1997) MR

[11] **J P C Greenlees**, **J P May**, *Completions in algebra and topology*, from "Handbook of algebraic topology" (I M James, editor), North-Holland, Amsterdam (1995) 255–276 MR

[12] **J P C Greenlees**, **J P May**, *Generalized Tate cohomology*, Mem. Amer. Math. Soc. 543, Amer. Math. Soc., Providence, RI (1995) MR

[13] **J P C Greenlees**, **V Stojanoska**, *Anderson and Gorenstein duality*, preprint (2017) arXiv

[14] **D Heard**, **V Stojanoska**, *K–theory, reality, and duality*, J. K-Theory 14 (2014) 526–555 MR

[15] **M A Hill**, *The equivariant slice filtration: a primer*, Homology Homotopy Appl. 14 (2012) 143–166 MR

[16] **M A Hill**, **M J Hopkins**, **D C Ravenel**, *On the nonexistence of elements of Kervaire invariant one*, Ann. of Math. 184 (2016) 1–262 MR

[17] **P Hu**, *On Real-oriented Johnson–Wilson cohomology*, Algebr. Geom. Topol. 2 (2002) 937–947 MR

[18] **P Hu**, **I Kriz**, *Real-oriented homotopy theory and an analogue of the Adams–Novikov spectral sequence*, Topology 40 (2001) 317–399 MR

[19] **P C Kainen**, *Universal coefficient theorems for generalized homology and stable cohomotopy*, Pacific J. Math. 37 (1971) 397–407 MR

[20] **N Kitchloo**, **W S Wilson**, *Unstable splittings for real spectra*, Algebr. Geom. Topol. 13 (2013) 1053–1070 MR

[21] **N Kitchloo**, **W S Wilson**, *The $ER(n)$–cohomology of $BO(q)$ and real Johnson–Wilson orientations for vector bundles*, Bull. Lond. Math. Soc. 47 (2015) 835–847 MR

[22] **P S Landweber**, *Conjugations on complex manifolds and equivariant homotopy of MU*, Bull. Amer. Math. Soc. 74 (1968) 271–274 MR

[23] **G Laures**, **M Olbermann**, *$TMF_0(3)$–characteristic classes for string bundles*, Math. Z. 282 (2016) 511–533 MR

[24] **V Lorman**, *The real Johnson–Wilson cohomology of $\mathbb{CP}^\infty$*, Topology Appl. 209 (2016) 367–388 MR

[25] **M A Mandell**, **J P May**, *Equivariant orthogonal spectra and S–modules*, Mem. Amer. Math. Soc. 755, Amer. Math. Soc., Providence, RI (2002) MR

[26] **H R Margolis**, *Spectra and the Steenrod algebra*, North-Holland Mathematical Library 29, North-Holland, Amsterdam (1983) MR

[27] **L Meier**, **M A Hill**, *The C_2–spectrum $Tmf_1(3)$ and its invertible modules*, preprint (2015) arXiv

[28] **N Ricka**, *Equivariant Anderson duality and Mackey functor duality*, Glasg. Math. J. 58 (2016) 649–676 MR

[29] **S Schwede**, *Symmetric spectra*, book draft (2012) `http://tinyurl.com/SchSS12`

[30] **S Schwede**, *Lectures on equivariant stable homotopy theory*, lecture notes (2016) `http://tinyurl.com/SchESHT16`

[31] **V Stojanoska**, *Duality for topological modular forms*, Doc. Math. 17 (2012) 271–311 MR

[32] **S M J Tilson**, *Power operations in the Künneth and C_2–equivariant Adams spectral sequences with applications*, PhD thesis, Wayne State University, Detroit, MI (2013) MR `http://search.proquest.com/docview/1459447634`

[33] **S Tilson**, *Power operations in the Künneth spectral sequence and commutative $H\mathbb{F}_p$–algebras*, preprint (2016) arXiv

School of Mathematics and Statistics, University of Sheffield
Sheffield, United Kingdom

Mathematics Institute, University of Bonn
Bonn, Germany

`j.greenlees@sheffield.ac.uk, lmeier@math.uni-bonn.de`

Received: 13 July 2016 Revised: 17 January 2017

Algebraic & Geometric Topology 17 (2017) 3621–3664

Slice implies mutant ribbon for odd 5–stranded pretzel knots

KATHRYN BRYANT

A pretzel knot K is called *odd* if all its twist parameters are odd and *mutant ribbon* if it is mutant to a simple ribbon knot. We prove that the family of odd 5–stranded pretzel knots satisfies a weaker version of the slice-ribbon conjecture: all slice odd 5–stranded pretzel knots are *mutant ribbon*, meaning they are mutant to a ribbon knot. We do this in stages by first showing that 5–stranded pretzel knots having twist parameters with all the same sign or with exactly one parameter of a different sign have infinite order in the topological knot concordance group and thus in the smooth knot concordance group as well. Next, we show that any odd 5–stranded pretzel knot with zero pairs or with exactly one pair of canceling twist parameters is not slice.

32S55, 57-XX

1 Introduction

A knot $K \subset S^3$ is *smoothly slice* if it bounds a smoothly embedded disk in the 4–ball. Similarly, a knot $K \subset S^3$ is said to be *topologically slice* if it bounds a locally flat embedded disk $D \subset B^4$, where D is a locally flat submanifold of B^4, if for every point $x \in D$ there exists a neighborhood $U \subset B^4$ of x such that the pair $(U, U \cap D)$ is homeomorphic to the pair $(\mathbb{R}^4, \mathbb{R}^2)$. The notions of smoothly slice and topologically slice knots can be used to define the smooth and topological knot concordance groups $\mathcal{C}$ and $\mathcal{T}$, respectively, under the operation of connected sum. These are widely studied groups for which the corresponding slice knot represents the identity element. For explicit information about the concordance relations, see Livingston [12]. Fine details of the group structure of $\mathcal{C}$ and $\mathcal{T}$ continue to elude mathematicians, but concordance order is one small way of gaining insights into these groups. The topic of determining smoothly slice knots and concordance order for knots within families of pretzel knots has also been studied with increasing frequency over the past 30 years and various results can be found in Greene and Jabuka [3], Lecuona [10], Miller [14], Herald, Kirk and Livingston [4] and Long [13]. This work will focus almost entirely on slice knots and concordance in the smooth case, except where "topological" is explicitly stated.

The slice-ribbon conjecture hypothesizes that if a knot is slice then it is also ribbon. Given that ribbon knots are easily seen to be slice, this is ultimately a conjecture about

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3621

the equivalence of the notions "slice" and "ribbon". Previous work by Joshua Greene and Stanislav Jabuka in [3] on the slice-ribbon conjecture for odd 3–stranded pretzel knots and work by Ana Lecuona in [10] on even pretzel knots inspired this project. This paper studies sliceness and concordance order for odd 5–stranded pretzel knots.

A *k–stranded pretzel link*, denoted by $P(p_1, p_2, \dots, p_k)$, where the $p_i \in \mathbb{Z} - \{0\}$ are called the *twist parameters*, is a knot in two cases: when exactly one of the twist parameters is even, or when k is odd and all the twist parameters are odd. A pretzel knot will be called *even* in the former case and *odd* in the latter. A *0–pair pretzel knot* is a pretzel knot for which there are no canceling pairs of twist parameters satisfying $p_i = -p_j$. A *1–pair pretzel knot* is a pretzel knot for which there exists a canceling pair of twist parameters, but when the pair is removed from the k–tuple defining the knot, the resulting $(k-2)$–stranded knot is 0–pair. Generally, a *t–pair pretzel knot* is one for which removing a single canceling pair of twist parameters results in a $(t-1)$–pair pretzel knot with two fewer strands. With this definition, 5–stranded pretzel knots $P(a, b, c, d, e)$ can be 0–pair, 1–pair or 2–pair. See Figure 1.

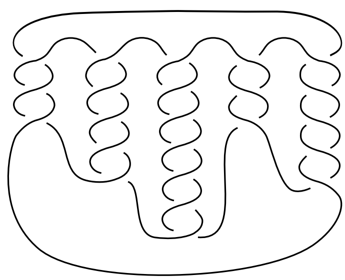


Figure 1: Pretzel knot $P(3, 5, 7, -3, -5)$

When proving statements about pretzel knots, it is often necessary to differentiate between the knots that contain twist parameters equal to ± 1 and those that do not. If for $K = P(p_1, \dots, p_k)$ there exists $i \in \{1, \dots, k\}$ such that $p_i = \pm 1$, then we say K is a *pretzel knot with single-twists*; otherwise, we say K is a *pretzel knot without single-twists*.

The classification of pretzel knots appears in Zieschang [20], a work that classifies the much larger class of Montesinos knots of which pretzel knots are a special case. The classification gives that two pretzel knots without single-twists are smoothly isotopic if their twist parameters differ by cyclic permutations, reflections, or compositions thereof. Two pretzel knots *with* single-twists are smoothly isotopic if their twist parameters differ by cyclic permutations, reflections and/or transpositions involving ± 1–twisted strands. Two k–stranded pretzel knots whose twist parameters are equal as *unordered*

k–tuples but not equal as *ordered* k–tuples are called *pretzel knot mutants*. This specific kind of mutation is the only type considered here, so "mutation" from this point on will always mean "pretzel knot mutation".

Mutation is a crucial topic for the problem of determining sliceness for k–stranded pretzel knots when $k \geq 4$ because many knot invariants used to obstruct sliceness are unable to detect pretzel knot mutants. In fact, any knot invariant based on the double branched cover of S^3 along the knot will fail to detect pretzel knot mutants; Bedient shows in [2] that any two pretzel knots defined by the same unordered k–tuple of twist parameters have the same double branched cover. Given a k–tuple $(p_1, \dots, p_k)$ of twist parameters, $P\{p_1, \dots, p_k\}$ will denote the set of pretzel knots having $\{p_1, \dots, p_k\}$ as twist parameters, and also all mirrors of such knots.

Among pretzel knots is a subset of knots we will call *simple ribbon*. A *simple ribbon move* on a pretzel knot is the ribbon move shown in Figure 2, performed always on the topmost twist of two adjacent strands of K having canceling numbers of twists. We say a pretzel knot K is *simple ribbon* if there exists a sequence of simple ribbon moves that reduces K to a 1–stranded pretzel knot (if K is odd) or to a 2–stranded pretzel knot $P(a,b)$ where $a = -b-1$ (if K is even). A prerequisite for a pretzel knot to be simple ribbon is that if K is k–stranded, then K must be $\frac{1}{2}(k-1)$–pair. But, while all 1–pair 3–stranded pretzel knots are simple ribbon, not all 2–pair 5–stranded pretzel knots are simple ribbon. For example, the 2–pair knot $P(3,5,-3,-5,7)$ is not simple ribbon because no two adjacent strands have canceling numbers of twists. This phenomenon extends for all $k \geq 4$.

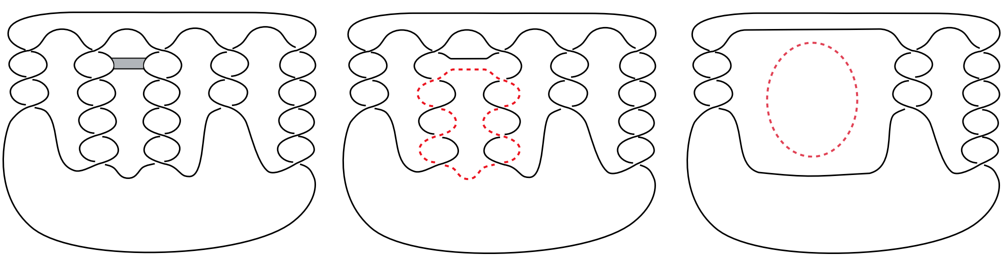


Figure 2: Simple ribbon move on pretzel knot $P(-3,-5,5,3,-5)$

The remainder of this paper is structured as follows: Section 2 presents our main results, the strongest of which is Corollary 2.5, a weak version of the slice-ribbon conjecture for generic odd 5–stranded pretzel knots. Section 3 gives foundational information on branched covers, framed links, weighted graphs and plumbings in the specific context of 4–dimensional topology. Section 4 describes a classical slice obstruction, the signature condition, and it gives the proof of Theorem 2.1. Section 5 gives details about Donaldson's diagonalization theorem and a resulting slice obstruction we call the lattice embedding condition. Section 6 addresses the final slice obstructions utilized in this work, d–invariants and the coset counting conditions. The proofs of the main results following Theorem 2.1 are given in Sections 7–10 and are organized by increasing number of slice obstructions needed to obtain the desired result.

Acknowledgements The author would like to thank her advisor Paul Melvin for his patience and guidance throughout this project. She is forever indebted to him for not only passing along his knowledge of this subject, but also for his invaluable edits of her work, both in style and content. She would also like to thank the referee for helpful suggestions and great attention to detail.

2 Results

As previously mentioned, this project was motivated by work of Greene and Jabuka in [3] on the slice-ribbon conjecture for odd 3–stranded pretzel knots and by work of Ana Lecuona in [10] on even pretzel knots. Lecuona writes down the following conjecture:

Pretzel ribbon conjecture (Lecuona) Let K be a pretzel knot whose twist parameters are all greater than 1 in absolute value. If K is ribbon, then K is simple ribbon.

For odd 3–stranded pretzel knots, the pretzel ribbon conjecture posits that the only ribbon knots are the simple ribbon knots, ie the 1–pairs. Similarly for odd 5–stranded pretzel knots, it says that the only ribbon knots are the simple ribbon knots which are 2–pairs *for which at least one of the canceling pairs is adjacent*. Greene and Jabuka show in [3] that odd 3–stranded pretzel knots satisfy both the pretzel ribbon conjecture and the slice-ribbon conjecture by proving that a knot of this type is slice if and only if it is 1–pair. This result, which proves the two aforementioned conjectures in a particular case, hints to the following possible strengthening of the slice-ribbon conjecture in the specific case of pretzel knots:

Pretzel slice-ribbon conjecture If K is a slice pretzel knot, then K is simple ribbon.

Of course, if the pretzel ribbon conjecture is true then the above is equivalent to the original version of the slice-ribbon conjecture. There is evidence that supports the pretzel slice-ribbon conjecture in the odd 5–stranded case. Herald, Kirk and Livingston [4] prove that $P(3,5,-3,-5,7)$ is *not* slice despite being mutant to the two simple ribbon knots $P(3,-3,5,-5,7)$ and $P(3,5,-5,-3,7)$. See Figure 2 for an illustration of the ribbon move on two adjacent strands in a 5–stranded pretzel knot whose twist numbers cancel.

This present work applies the techniques used by Jabuka and Greene on odd 3–stranded pretzel knots to odd 5–stranded pretzel knots, in the hope of showing that this new class of knots also satisfies the pretzel slice-ribbon conjecture as well. It should be noted that Greene and Jabuka went a step farther and proved that all nonslice odd 3–stranded pretzel knots have infinite order in $\mathcal{C}$. To obtain these results, they used three tools: the knot signature from classical knot theory, Donaldson's diagonalization theorem from gauge theory, and the d–invariant from Heegaard Floer theory.

The main results of this project are given below, accompanied by brief explanations as to where each of the above three tools comes into play. In Theorem 2.1 and its corollary, $\sigma(K)$ denotes the signature of K; s is the difference between the number of positive twist parameters and the number of negative twist parameters of K; $\hat{e}$ is the orbifold Euler characteristic of K given by the sum of the reciprocals of the twist parameters; and sgn( ) is the function returning -1, 0, or $+1$ according to whether the input is negative, zero, or positive, respectively. The first result is about the class of odd pretzel knots:

Theorem 2.1 *If K is an odd pretzel knot, then $\sigma(K) = s - \operatorname{sgn}(\hat{e})$. In particular, $\sigma(K) = 0$ if and only if $s = \operatorname{sgn}(\hat{e})$.*

Corollary 2.2 *All odd pretzel knots with $s \neq \operatorname{sgn}(\hat{e})$ have infinite order in the topological knot concordance group $\mathcal{T}$.*

The corollary follows from the fact that σ is a homomorphism from $\mathcal{T} \to \mathbb{Z}$, and it implies infinite order in the smooth knot concordance group $\mathcal{C}$ as well. It is a well-known fact that we call on later that if a knot K is slice, then $\sigma(K) = 0$. An implicit implication of Theorem 2.1 is that all odd pretzel knots for which $s \neq \pm 1$ are not slice, which is particularly easy to read off from the k–tuple defining the knot. For odd 5–stranded pretzel knots this tells us that if all or all but one of the twist parameters have the same sign, then K is not slice.

Powerful as the signature is as a concordance invariant, the signature alone is insufficient for determining sliceness in odd pretzel knots for which $s = \pm 1$. For example, the pretzel knot $K = P(-3, -5, -7, 9, 27)$ has vanishing signature, but the pretzel slice-ribbon conjecture gives us reason to think that K may not be slice. Such occurrences in the odd 3–stranded case prompted Jabuka and Greene to turn to an obstruction based on Donaldson's diagonalization theorem, which is ultimately phrased as a lattice embedding condition necessary for sliceness. This same obstruction was originally used by Paolo Lisca in [11] to classify slice knots within the family of 2–bridge knots.

The use of Donaldson's diagonalization theorem to define a "lattice embedding condition" for sliceness is based on the construction of a (potentially hypothetical) closed, definite 4–manifold X, created as follows: Assume K is a slice knot. Let Y be the double branched cover of S^3 along K; let W be the double branched cover of the 4–ball with branching set the slice disk for K, so that W is a rational homology 4–ball with $\partial W = Y$; let P be a canonical definite 4–dimensional plumbing[1] with $\partial P = Y$. Define $X = P \cup_Y (-W)$. The lattice embedding condition arises by applying the diagonalization theorem to X, for which it is necessary to verify that the intersection form on X, Q_X, can in fact be diagonalized over the integers. We do this in Section 5.

The lattice embedding condition for sliceness puts great restrictions on the possible k–tuples that can define a slice odd pretzel knot, so it enables us to conclude that all but a very select subset of such knots are not slice. Unfortunately, the knots that satisfy both the vanishing signature condition and the lattice embedding condition are not easily differentiated from the knots satisfying the signature condition but *not* the lattice embedding condition. For example, sliceness is obstructed for $P(-3,-17,27)$ and $P(-3,-7,-19,17,55)$ by the lattice embedding condition, but *not* obstructed for $P(-3,-17,29)$ and $P(-3,-7,-19,19,55)$.

For this reason Jabuka and Greene introduced a third slice obstruction based on the d–invariant from Heegaard Floer theory. It assumes the same construction used above involving K, Y, W, P, and X, but it boils down to a comparison of two different "counts" obtained by analysis on the homology long exact sequences of the pairs (X, W) and (P, Y). We refer to it here as "coset condition I". Combining the signature obstruction, the lattice embedding condition, and coset condition I, Jabuka and Greene were able to prove their full result. With these same tools, we obtain the following results for odd 5–stranded pretzel knots with signature zero:

Theorem 2.3 *If K is a 0–pair odd 5–stranded pretzel knot, then K is not slice.*

Coset condition I fails to obstruct sliceness in t–pair odd k–stranded pretzel knots K if $t \geq 1$, k is odd, and $\sigma(K) = 0$, so yet another tool is required to prove analogous results in the present case. When $k \geq 5$, the increased number of twist parameters introduces complexity not present when $k = 3$, requiring a more refined "count" than Jabuka and Greene needed when implementing the d–invariant obstruction. With just

[1] A canonical definite plumbing P is one for which the weights of the vertices in the corresponding plumbing graph are either all ≥ 2 or ≤ -2. It should be noted that not all knots have such plumbings, but that pretzel knots do.

a little bit of work we derive a stronger version of coset condition I and uncreatively call it coset condition II. Combining the signature obstruction, the lattice embedding condition, and coset condition II, we prove:

Theorem 2.4 *If K is a 1–pair odd 5–stranded pretzel knot without single-twists, then K is not slice.*

Theorem 2.4 avoids mention of odd 5–stranded pretzel knots *with* single-twists because they behave slightly differently from those without single-twists for the following reason: any strand with exactly one positive or negative half twist can be transposed with an adjacent strand through a *flype* as in Figure 3. Such a move preserves the smooth knot type thus, for example, $P(1,3,-5,1,-7)$ and $P(1,1,3,-5,-7)$ are not only mutants of one another but also members of the same smooth isotopy class.

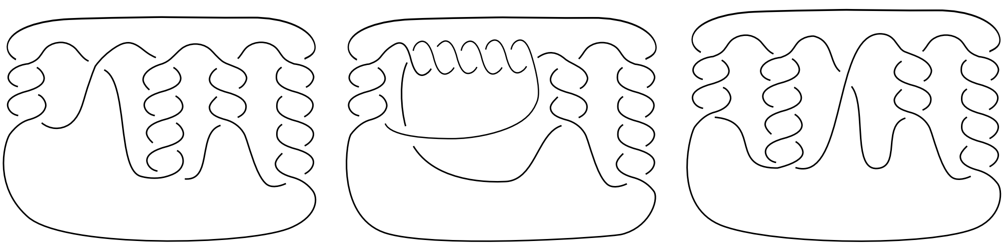


Figure 3: Transposition of a single-twist strand, turning $P(3,1,5,-3,-5)$ into $P(3,5,1,-3,-3)$

Furthermore, by flyping we can always "collect" all strands with ± 1 twists so that they occur in succession. This has the greatest impact on 1– and 2–pair pretzel knots for which at least one of the pairs is $\{-1,1\}$. If K is defined by $\{-1,1,b,c,d\}$, then K is not only concordant to $P(b,c,d)$ but also smoothly isotopic, regardless of the initial locations of 1 and -1 in the 5–tuple. It follows that every 2–pair odd 5–stranded pretzel knot containing the pair $\{-1,1\}$ is simple ribbon. To contrast, if $K \in P\{-a,a,b,c,d\}$ with $|a| \geq 3$, then K is smoothly concordant to $P(b,c,d)$ if and only if the pair $\{-a,a\}$ is adjacent; it is precisely this that leads to $P(3,5,-3,-5,7)$ and $P(3,-3,5,-5,7)$ having different smooth concordance order.

Theorems 2.3 and 2.4 together imply that odd 5–stranded pretzel knots without single-twists satisfy a weaker version of the slice-ribbon conjecture:

Corollary 2.5 *If K is a slice odd 5–stranded pretzel knot without single-twists, then K is mutant to a simple ribbon knot.*

For 2–pair, odd, 5–stranded pretzel knots (with or without single-twists) not containing the pair $\{-1,1\}$ and for 1–pairs with single-twists and pair $\{-a,a\} \neq \{-1,1\}$, the signature condition, the lattice embedding condition, and coset conditions I and II all fail to obstruct sliceness in the knots that are not simple ribbon because these slice obstructions, at their cores, obstruct the double branched covers of the knots from having certain properties. As previously mentioned, all mutants of a given pretzel knot share the same double branched cover and hence there is no hope of obstructing sliceness for a knot $K \in P\{a,b,c,d,e\}$ if any member of $P\{a,b,c,d,e\}$ is slice. Since 2–pair knots of the form $P(a,-a,b,-b,c,)$ and $P(a,b,-b,-a,c)$ are simple ribbon and therefore slice, we cannot use the aforementioned tools to say that $P(a,b,-a,-b,c)$ is not slice. Similarly, Remark 1.3 in [3] gives that the 1–pair knots with single-twists and pair $\{-a,a\} \neq \{-1,1\}$ of the form $P(a,-a,1,b,c)$ with $b+c=4$ are slice, and therefore again there is no way to distinguish between slice and suspected nonslice members of $P\{a,-a,1,b,c\}$.

In [4], Herald, Kirk and Livingston used twisted Alexander polynomials to show that the 2–pair knot $P(3,5,-3,-5,7)$ is not slice, despite being mutant to the simple ribbon knot $P(3,-3,5,-5,7)$. Twisted Alexander polynomials are able to distinguish mutants and, in fact, they can reveal when a knot is not *topologically* slice. The issue in using twisted Alexander polynomials to show that pretzel knots satisfy the slice-ribbon conjecture is that their computation relies on number-theoretic choices that often make it difficult to find general formulas for infinite families of knots. Of the examples computed for pretzel knots to date, there is only one infinite family of pretzel knots whose slice status has been determined using twisted Alexander polynomials. It is a subfamily of the 4–stranded family $K = P(2n, m, -2n \pm 1, -m)$, done by Allison Miller in [14].

3 Branched covers, framed links, weighted graphs and plumbings

Let $K = P(a_1, \dots, a_p, -b_1, \dots, -b_n)$ be an odd k–stranded pretzel knot with $k = p+n$ odd, $a_i, b_j > 0$, and let Y be the double branched cover of S^3 along K. As a 3–manifold, we will describe Y by two framed links L_0 and L_+ which differ by a sequence of moves in the Kirby calculus for framed links. The links L_0 and L_+ are equivalently represented by weighted star-shaped graphs Γ_0 and Γ_+, shown in Figure 4. In Γ_0 and Γ_+, each vertex v_i with weight $w(v_i)$ represents an unknot component K_i with framing $r_i = w(v_i)$; two components K_i and K_j link once in L_0 (resp. L_+) if the corresponding vertices v_i and v_j share an edge in Γ_0 (resp. Γ_+).

To obtain the double cover of S^3 branched along a pretzel knot, we use the following algorithm of Montesinos: start with an unknot in S^3 and the double cover of S^3

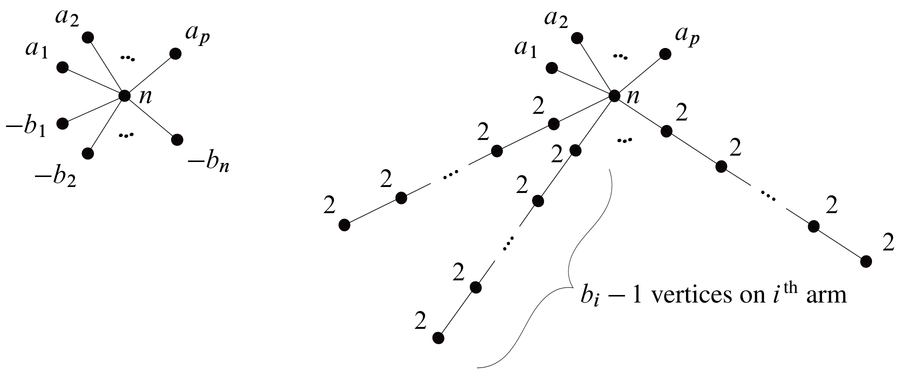


Figure 4: Weighted plumbing graphs Γ_{L_0} (left) and Γ_{L_+} (right)

branched along the unknot. Recall that the double cover of S^3 branched along the unknot is, again, S^3, which can be described as surgery on an unknot with 0–framing. The unknot can be turned into a $P(p_1, \dots, p_k)$ pretzel knot by replacing a 0–tangle in the unknot by a p_i–tangle for each i and replacing a single ∞–tangle by a 0–tangle, as shown in Figure 5 for a 5–stranded pretzel knot. Determining the double branched

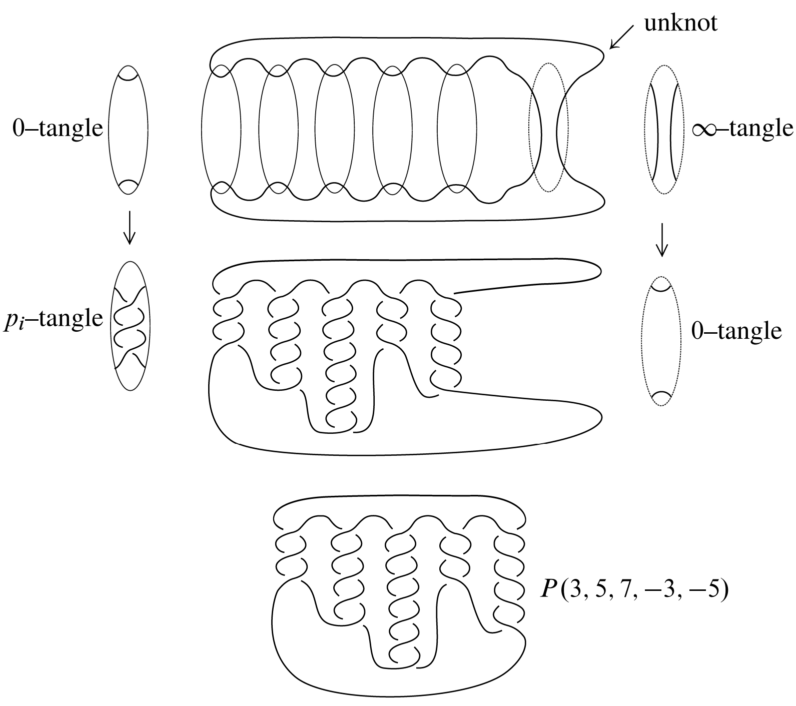


Figure 5: Obtaining a 5–stranded pretzel knot by replacing tangles

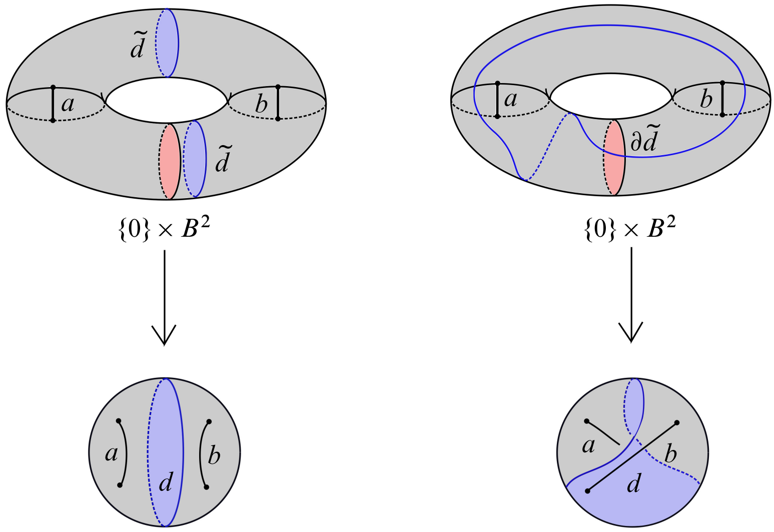
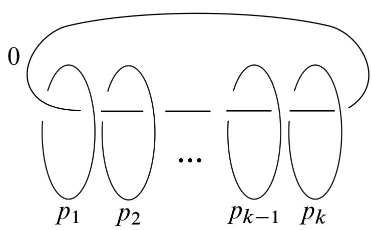


Figure 6: Double covers of B^3 branched along the ∞–tangle and over the $\frac{1}{2}$–tangle

cover of S^3 along $P(p_1, \dots, p_k)$ now amounts to accounting for the changes affected in the cover above by the tangle replacements in the knot below.

The double cover of B^3 branched along a p_i–tangle is a solid torus T_{p_i} parametrized by $S^1 \times B^2$ such that the lift $\widetilde{d}$ of the disk d in B^3 separating the two arcs of the tangle satisfies (i) $\partial\widetilde{d} \cap (\{0\} \times B^2) = 1$ and (ii) $\mathrm{lk}(\partial\widetilde{d}, S^1 \times \{0\}) = p_i$. See Figure 6. As such, the effect in the double branched cover of replacing a 0–tangle by a p_i–tangle is p_i–surgery on an unknot in S^3; the effect in the cover of replacing the ∞–tangle with the 0–tangle is 0–surgery on an unknot in S^3 that links once with each of the other unknots. Considering all surgeries as being done simultaneously, the double cover of S^3 branched along the pretzel knot $P(p_1, \dots, p_k)$ is thus obtained by surgery on S^3 described by this framed link:[2]



[2]From this construction, we see that any two pretzel knots defined by the same unordered k–tuple of twist parameters have the same double branched cover. As a result, any knot invariant based on the double branched cover of S^3 along the knot will fail to detect pretzel knot mutants.

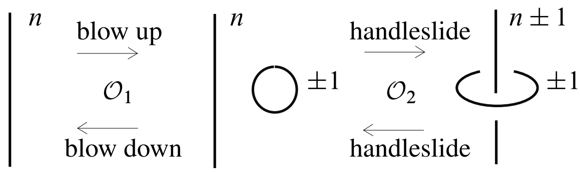


Figure 7: “Slam dunk” shortcut

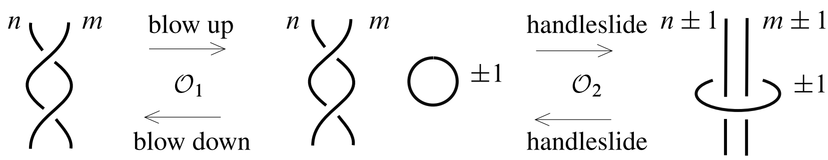


Figure 8: “Kirby twist” shortcut

To transform L_0 into L_+ via Kirby moves, we operate on the negatively framed components of L_0 using the shortcuts shown in Figures 7 and 8. The course of Kirby moves needed to change L_0 into L_+ is described in Figure 9. The accompanying

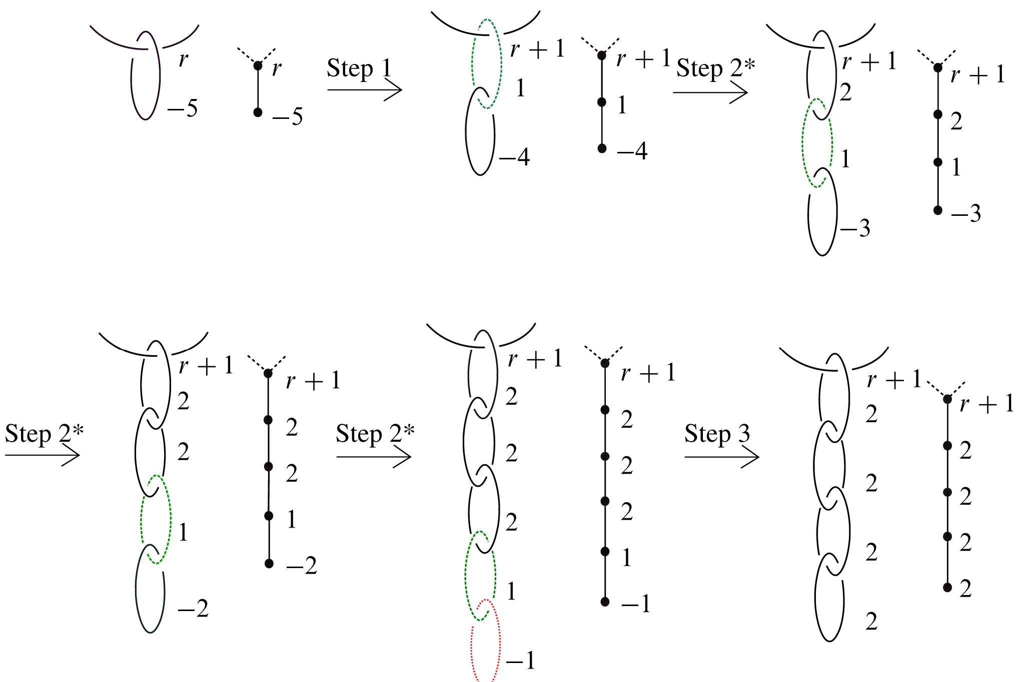


Figure 9: Kirby calculus sequence $L_0 \to L_+$. The asterisk on Step 2 indicates that it will be repeated until the bottommost component has framing -1.

weighted graphs are shown as well. If the prescribed sequence of moves is performed on a component with framing $-b_i$, then Step 2 is repeated b_i-3 times and the original component is ultimately replaced by b_i-1 new components, all of which are unknots with framing 2. In the corresponding weighted graphs, this translates into replacing a single arm of length one, whose lone vertex has weight $-b_i$, by an arm of length b_i-1 containing all weight-2 vertices; the weight of the central vertex increases by 1 for each arm altered.

Any sequence of Kirby moves used to transform one framed link into another is mimicked in the corresponding graphs by adding (resp. deleting) vertices and edges to the graph at the expense of subtracting 1 (resp. adding 1) to the weights of the vertices sharing an edge with the added (resp. deleted) vertex.

In addition to describing the double branched cover Y of S^3 along a pretzel knot, the framed links L_0 and L_+ and weighted graphs Γ_0 and Γ_+ define 4–dimensional plumbings P_0 and P_+ (respectively) bounded by Y. The plumbings P_0 and P_+ can be viewed as 4–dimensional handlebodies whose handle decompositions consist entirely of a single zero handle and a collection of 2–handles. For P_0, each component $C_i \subset L_0$ corresponds to a 4–dimensional 2–handle G_i that is attached to B^4 (the single 0–handle) by a map $f\colon \bigcup_i(\widetilde{G}_i) \to \partial B^4$, where $\widetilde{G}_i = S^1 \times B^2$ is the attaching region of G_i. The map f identifies $\widetilde{G}_i$ with a tubular neighborhood N_i of C_i, such that the core of $\widetilde{G}_i$ lies along C_i and a meridian of $\partial\widetilde{G}_i$ is identified with a curve in N_i that links n_i times with C_i. Plumbing P_+ is given similarly.

The effect on $\partial B^4 = S^3$ of attaching 2–handles to B^4 corresponds exactly to performing surgery on S^3 by viewing the identification of $\widetilde{G}_i$ with N_i instead as removing N_i and replacing it with the $\widetilde{G}_i$ according to the framing. Since P_0 and P_+ are described by L_0 and L_+ as plumbings/4–dimensional handlebodies and Y is described by L_0 and L_+ as surgery on S^3, it follows that $\partial P_0 = \partial P_+ = Y$.

The intersection form Q_{P_0} for P_0, represented as a matrix with basis equal to the set of classes represented by the zero-section of each plumbed disk bundle, is equal to the incidence matrix for Γ_0. Likewise, with an analogous choice of basis the intersection form of P_+ is equal to the incidence matrix for Γ_+. Knowing an exact sequence of Kirby moves between L_0 and L_+ allows one to compute the overall change in the signature from P_0 to P_+ by analyzing how the signature changes with each step.

At this point, it is worth detailing a labeling scheme for the vertices of Γ_0 and Γ_+ so that the bases for the incidence matrices are ordered consistently. Given the vertex labelings pictured in Figure 10, Γ_0 will have ordered basis $\{v_0, v_1, \dots, v_p, v_{p+1}, \dots, v_{p+n}\}$ and Γ_+ will have ordered basis $\{s_0, s_1, \dots, s_p, s_{1,1}, \dots, s_{1,r_1}, \dots, s_{n,1}, \dots, s_{n,r_n}\}$. Written succinctly, the basis for Γ_+ can be written $\{s_i, s_{j,r_j}\}$ where $0 \le i \le p$,

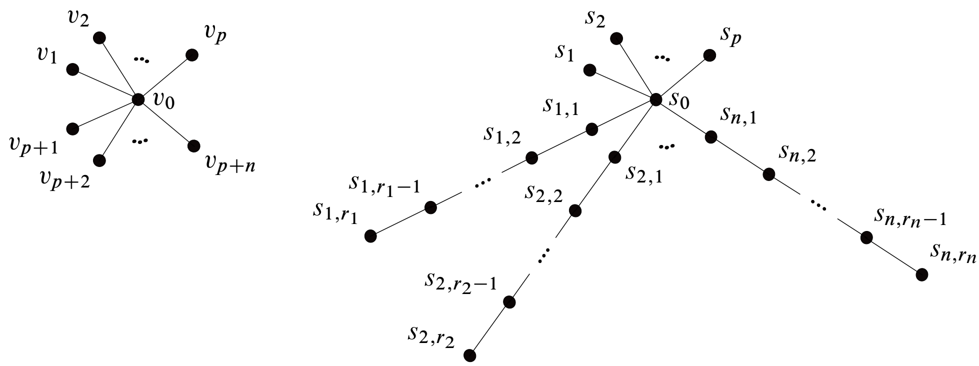


Figure 10: Vertex labeling scheme for Γ_0 (left) and Γ_+ (right)

$1 \le j \le n$, and r_j is equal to one less than the number of vertices in the j^{th} arm of Γ_+. It is with these ordered bases for Γ_0 and Γ_+ that the matrices for Q_{P_0} and Q_{P_+} are given later.

4 The signature condition and proof of Theorem 2.1

The *signature* of a symmetric matrix Q, denoted $\sigma(Q)$, is the difference between the number of positive diagonal entries and the number of negative diagonal entries of Q, after Q has been diagonalized over $\mathbb{R}$. The *signature of a knot* K is defined as $\sigma(K) = \sigma(V^T + V)$, where V is a Seifert matrix for K. Given a 4–manifold X with intersection form Q_X, the *signature of* X is defined as the signature of Q_X: $\sigma(X) = \sigma(Q_X)$. The signature is an abelian invariant based on the double branched cover of the knot, and therefore it cannot detect pretzel mutants.

The signature is a homomorphism $\sigma\colon \mathcal{T} \to \mathbb{Z}$, where $\mathcal{T}$ is the topological knot concordance group. Hence

(1) $\sigma(-K) = -\sigma(K)$, where $-K$ is the mirror of K, and

(2) $\sigma(K_1 \# K_2) = \sigma(K_1) + \sigma(K_2)$.

The signature is also invariant under mutation; see [8]. For pretzel knots, if we combine this fact with (1) we see that computation of the signature of $K = P(p_1, \dots, p_k)$ may be obtained using any knot in $P\{p_1, \dots, p_k\}$. Often, a specific K is chosen in order to simplify computations. Homomorphism property (2) implies that if $\sigma(K) > 0$, then the knot K will have infinite order in the topological (and therefore smooth) knot concordance group. A classical theorem (a proof of which can be found in [18]) states that *any slice knot has signature zero*.

With this result we are now ready to prove Theorem 2.1.

Proof of Theorem 2.1 Let $K = P(a_1, \dots, a_p, -b_1, \dots, -b_n)$ be an odd pretzel knot. For simplicity throughout this proof, we will make two notational simplifications: First, we will call a framed link and the 4–manifold it describes by the same name; in particular, we will use the framed link L_0 and the corresponding plumbing manifold P_0 from Section 3 in this proof but we will resort to using P_0 to describe both, letting context dictate whether we are referring to the framed link or the 4–manifold. This notational simplification will be implicit in all other framed links and 4–manifolds defined within this proof. The second notational simplification we will make is to write $\sigma(M)$ rather than $\sigma(Q_M)$ for the signature of a 4–manifold M; in fact, this is less of an idiosyncrasy in notation and more appropriately an understated definition of the signature of a 4–manifold as the signature of its intersection form.

Kauffman and Taylor prove in [7] that $\sigma(K) = \sigma(T)$, where T is the double branched cover of B^4 along any pushed-in Seifert surface of K. In [1], Akbulut and Kirby give an algorithm for computing the p–fold ($p \geq 2$) cyclic cover of B^4 branched along a pushed-in Seifert surface for a given knot, where the Seifert surface used is one that can be described as a disk with possibly twisted and possibly knotted bands attached. Such a Seifert surface F can be obtained for a pretzel knot by an isotopy of the "standard" Seifert surface, ie the Seifert surface obtained from Seifert's algorithm via the standard pretzel knot diagram. Using this particular Seifert surface F for K together with Akbulut and Kirby's algorithm for the double branched cover of B^4, we get a framed link that describes the handlebody structure of the 2–fold cover of B^4 branched along F pushed-in. We call this particular cover T. See Figure 11 for an example of F and T for the pretzel knot $P(3, 5, -5, -3, -5)$.

Rather than compute $\sigma(T) = \sigma(K)$ directly, we will instead show that $\sigma(T) = \sigma(T \# (S^2 \mathbin{\tilde{\times}} S^2)) = \sigma(P_0)$, where P_0 is the plumbing manifold from Section 3. By choosing the basis for Q_{P_0} to be the set of spheres obtained from the cores of the attaching 2–handles together with hemispheres in B^4 (alternatively, the spheres are the 0–sections of the disk bundles used to create P_0), Q_{P_0} is given by the incidence matrix of the plumbing graph Γ_0 from Section 3. A straightforward diagonalization of Q_{P_0} shows that

$$Q_{P_0} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & a_1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & a_p & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -b_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 0 & -b_n \end{bmatrix} \sim \begin{bmatrix} -\hat{e} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -b_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & -b_n \end{bmatrix}.$$

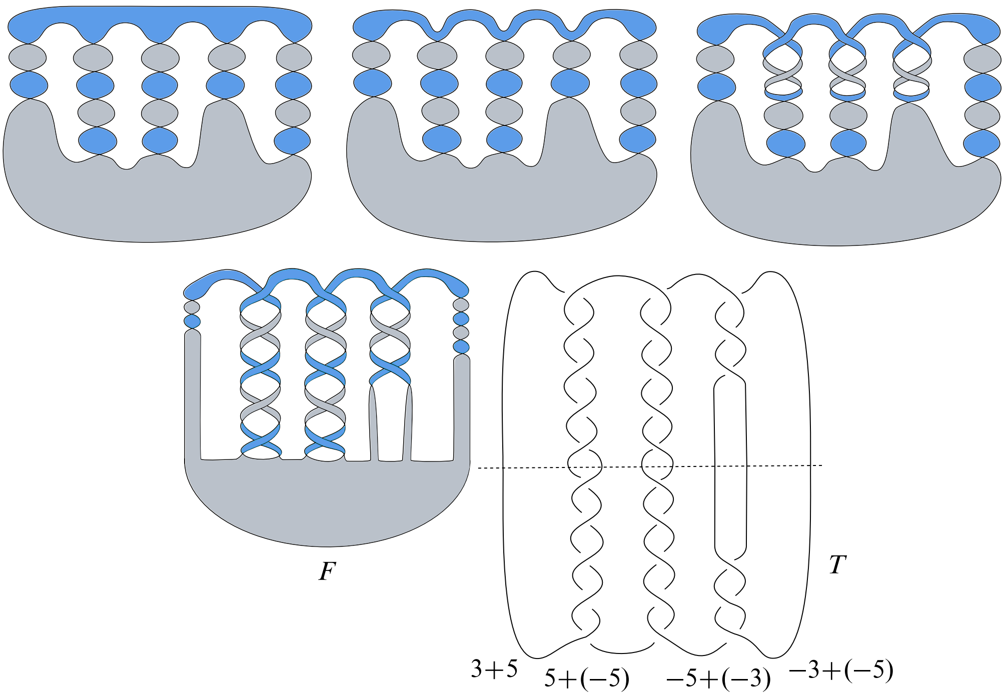


Figure 11: The Seifert surface F and the double cover T of B^4 branched along F pushed-in for the pretzel knot $P(3,5,-5,-3,-5)$ according to Akbulut and Kirby's algorithm.

Therefore, $\sigma(P_0) = s - \operatorname{sgn}(\widehat{e})$.

Now, P_0 can be seen to be equivalent to $T \# (S^2 \widetilde{\times} S^2)$ by performing handle slides: for each adjacent pair of nonzero framed handles $((h_i, \alpha_i), (h_{i+1}, \alpha_{i+1}))$ of P_0, where h_i is the i^{th} nonzero-framed handle[3] and a_i its framing, we slide h_i over h_{i+1}. The result is a new pair of adjacent handles $((\widetilde{h}_i, \alpha_i + \alpha_{i+1}), (h_{i+1}, \alpha_{i+1}))$ that link α_{i+1} times. This is performed for the pairs of handles $((h_i, a_i), (h_{i+1}, a_{i+1}))$ for $1 \le i \le p-1$, for the pair $((h_p, a_p), (h_{p+1}, b_1))$, and lastly for the pairs $((h_{p+j}, b_j), (h_{p+j+1}, b_{j+1}))$ for $1 \le j \le n-1$. The result of performing these handle slides is a framed link diagram of T linked with a Hopf link, where one component of the Hopf link has framing 0 and the other has odd framing. This entire process is summarized in Figure 12.

By Lemma 4.4 in [9], a Hopf link with such framings is equivalent to a Hopf link H with a single 0–framed component and a single 1–framed component, which gives a description of the twisted sphere bundle $S^2 \widetilde{\times} S^2$. By Lemma 4.5 in [9], H can be

[3] Handles ordered from left to right in the standard diagram of P_0.

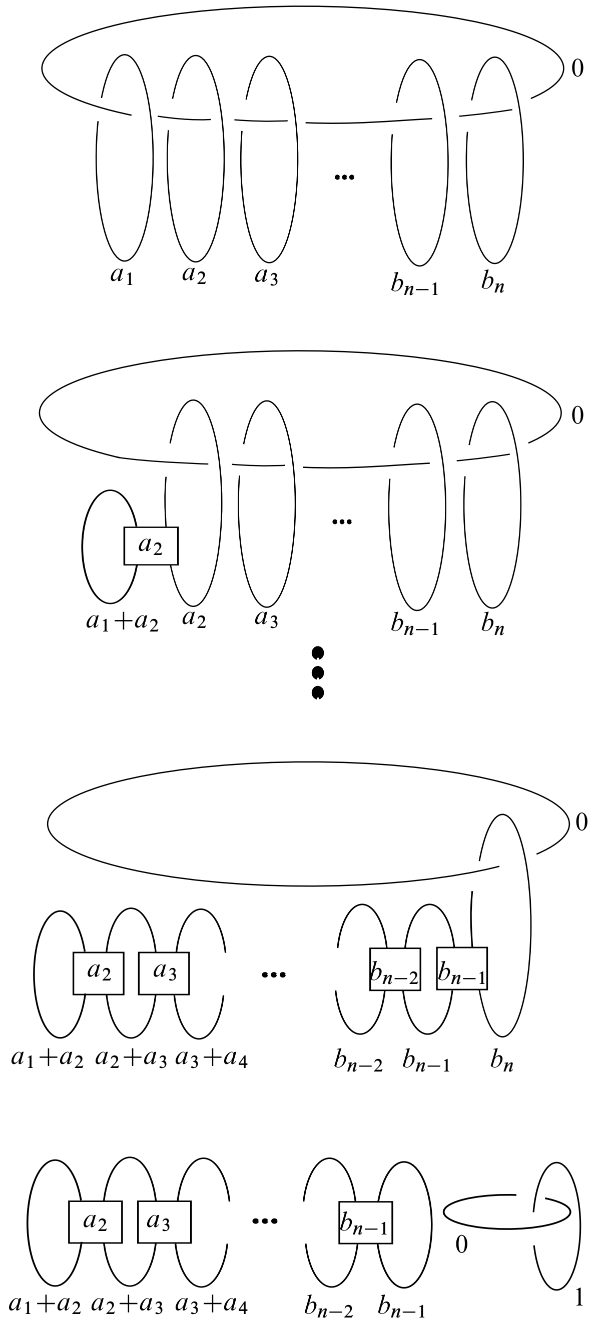


Figure 12: The sequence of handle slides showing that P_0 is equivalent to $T \# (S^2 \widetilde{\times} S^2)$ as 4–manifolds

unlinked from T through handle slides to yield $T \sqcup H$, which describes the 4–manifold $T \,\#\, (S^2 \mathbin{\tilde{\times}} S^2)$. Given that we moved from P_0 to $T \,\#\, (S^2 \mathbin{\tilde{\times}} S^2)$ exclusively through handle slides (no blow-ups or blow-downs), P_0 and $T \,\#\, (S^2 \mathbin{\tilde{\times}} S^2)$ are the same 4–manifold and thus $\sigma(P_0) = \sigma(T \,\#\, (S^2 \mathbin{\tilde{\times}} S^2))$.

To finish, recall that the signature is additive under connected sums and $\sigma(S^2 \mathbin{\tilde{\times}} S^2) = 0$. Hence

$$\begin{aligned}\sigma(K) &= \sigma(T) = \sigma(T) + 0 = \sigma(T) + \sigma(S^2 \mathbin{\tilde{\times}} S^2) \\ &= \sigma(T \,\#\, (S^2 \mathbin{\tilde{\times}} S^2)) \\ &= \sigma(P_0) \\ &= s - \operatorname{sgn}(\hat{e}). \qquad \square\end{aligned}$$

Let $a, b, c, d, e \geq 3$. As mentioned in Section 2, Theorem 2.1 shows nonsliceness for all odd 5–stranded pretzel knots K in $P\{a, b, c, d, e\}$ and in $P\{a, b, c, d, -e\}$, since s fails to equal ± 1. But, it also shows nonsliceness for all K in $P\{a, b, c, -d, -e\}$ for which $1/a + 1/b + 1/c > 1/d + 1/e$. For example, $K = P(5, 5, 5, -3, -3)$ has nonvanishing signature by Theorem 2.1 and is therefore not slice.

5 Donaldson's diagonalization theorem and the lattice embedding condition

Donaldson's diagonalization theorem constitutes a small piece of the larger topic of Yang–Mills gauge theory. It remains one of the most significant results in 4–manifold topology, and it has useful applications in many other areas of low-dimensional topology. Donaldson's diagonalization theorem can be used to obstruct knot sliceness and it is with this goal in mind that we call on it here. Recall that a closed, oriented 4–manifold X has a unimodular intersection form[4]

$$Q_X\colon H_2(X)/\operatorname{Tor} \otimes H_2(X)/\operatorname{Tor} \to \mathbb{Z},$$

and that Q_X is *definite* if $|\sigma(Q_X)| = \operatorname{rk}(Q_X)$. Then:

Theorem (Donaldson 1987) *Let X be a smooth, closed, oriented, 4–manifold with positive definite intersection form Q_X. Then Q_X is equivalent over the integers to the standard diagonal form, so in some base,*

$$Q_X(u_1, u_2, \ldots, u_r) = u_1^2 + u_2^2 + \cdots + u_r^2.$$

[4]Here, Tor denotes the torsion part of $H_2(X)$.

Remark Donaldson's diagonalization theorem was originally phrased for Q_X negative definite, making all the u_i^2 terms negative. Also, Q_X being definite and diagonalizable means that the pair $(\mathbb{Z}^{b_2(X)}, Q_X)$ can be viewed geometrically as a lattice that is isomorphic to $\mathbb{Z}^{b_2(X)}$ with the standard dot product.

Donaldson's diagonalization theorem is used to obstruct sliceness of a knot in the following way: Assume the knot $K \subset S^3$ is slice and that Y is the 2–fold branched cover of S^3 along K. Let P be a canonical definite 4–dimensional plumbing manifold satisfying $\partial P = Y$, and let W be the double branched cover of B^4 along a slicing disk for K. Since K is a knot, Y is a rational homology 3–sphere. Furthermore, W is a rational homology 4–ball with $\partial W = Y$, which follows from the more general fact that the double branched cover of a $\mathbb{Z}/2\mathbb{Z}$–homology ball branched along a codimension-2 $\mathbb{Z}/2\mathbb{Z}$–homology ball is again a $\mathbb{Z}/2\mathbb{Z}$–homology ball. For a proof of this, see [6, Lemma 17.2]. A new 4–manifold X is formed by gluing P and W together along their common boundary Y in the usual, orientation-preserving way. This new manifold X will be compact, smooth, oriented, and have definite intersection form, and thus the diagonalization theorem applies. This gives that $(\mathbb{Z}^{b_2(X)}, Q_X)$ is lattice isomorphic to $(\mathbb{Z}^{b_2(X)}, \mathrm{Id})$, the standard n–dimensional integer lattice.

The Mayer–Vietoris sequence involving $X = P \cup_Y (-W)$ with rational coefficients shows that $H_2(P)$ includes into $H_2(X)$, and therefore $(\mathbb{Z}^{b_2(P)}, Q_P)$ must embed into $(\mathbb{Z}^{b_2(X)}, Q_X)$ as a sublattice of full rank. Algebraically, $(\mathbb{Z}^{b_2(P)}, Q_P)$ embeds into $(\mathbb{Z}^{b_2(X)}, \mathrm{Id})$ if there exists an injection $\alpha\colon \mathbb{Z}^{b_2(P)} \to \mathbb{Z}^{b_2(X)}$ such that $Q_P(a,b) = \mathrm{Id}(\alpha(a), \alpha(b))$ [3]. If this embedding does not exist then the conclusion is that X, as constructed, does not exist. The only assumption made in this construction was that K is slice; therefore the contradiction implies this cannot be the case. Thus, the existence of an embedding α of the lattice $(\mathbb{Z}^{b_2(P)}, Q_P)$ into $(\mathbb{Z}^{b_2(X)}, Q_X)$ is a necessary condition for the knot K to be slice, which is precisely the obstruction to sliceness utilized in [11] and [3]. We call this the lattice embedding condition.

In practice, showing the embedding α exists amounts to writing down a matrix A for α that satisfies $A^T A = Q_P$. This requires a choice basis for $H_2(P)$ and for $H_2(X)/\operatorname{Tor}$. The basis $\{s_i\}$ chosen for $H_2(P)/\operatorname{Tor}$ is the set of classes represented by the zero-sections in the disk bundles used to create P; the basis $\{e_i\}$ for $H_2(X)/\operatorname{Tor}$ is chosen to be one that makes Q_X diagonal by Donaldson's theorem. As such, each column of A corresponds to one of those 2–spheres in P whose intersection information is recorded by the plumbing graph of P. That is, the columns of A must have standard dot products consistent with the information given by the plumbing graph for P.

In an attempt to use Donaldson's diagonalization theorem to obstruct sliceness of an odd pretzel knot K, we refer back to the Section 3 and take $P = P_+$, which has plumbing graph Γ_+ and intersection form Q_{P_+}, with matrix equal to the incidence matrix for Γ_+ with respect to the above bases. By the signature obstruction to sliceness, we need only consider odd pretzel knots K for which $\sigma(K)=0$. In order to utilize the positive definite version of Donaldson's diagonalization theorem, we need to prove that Q_X is positive definite for $X=P_+\cup_Y(-W)$. This is done with the help of the following lemma:

Lemma 5.1 *If K is an odd k–stranded pretzel knot with k odd and $\sigma(K)=0$, then either $Q_{P_+}(K)$ or $Q_{P_+}(-K)$ is positive definite.*

Proof From Theorem 2.1, we know that $\hat{e}\neq 0$ for pretzel knots K with $\sigma(K)=0$. Then, Theorem 5.2 in [15] tells us that Q_{P_+} is either positive definite or negative definite, according to whether $\hat{e}>0$ or $\hat{e}<0$, respectively. Taking the mirror $-K$ of a knot K will change Q_{P_+} from positive definite to negative definite, or vice versa. Thus after mirroring if necessary, it is always possible to choose K so that Q_{P_+} is positive definite. □

With our eye on applying the diagonalization theorem to X and the help of Lemma 5.1, we argue that Q_X is also positive definite for $X=P_+\cup_Y(-W)$. Consider the following portion of the Mayer–Vietoris sequence for X with rational coefficients:

$$0\longrightarrow \mathbb{Q}^n\oplus 0\xrightarrow{\ i_*\ } H_2(X;\mathbb{Q})\longrightarrow 0.$$

The map i_* is an isomorphism, which implies every element $x\in H_2(X)$ is a $\mathbb{Q}$–linear combination of basis elements $\{s_i\}$ for $H_2(P_+)$ and torsion elements of $H_2(W)$. Bilinearity of Q_X and positive-definiteness of Q_{P_+} yield that Q_X is positive definite. Thus, we are free to utilize the previously described construction using Donaldson's diagonalization theorem, with $P=P_+$, to obtain the embedding criterion for sliceness on odd, 5–stranded pretzel knots.

In all the results that follow, we use Theorem 2.1 to immediately reduce to considering only those odd, 5–stranded pretzel knots of the form $P(-a,-b,-c,d,e)$ for which $\mathrm{sgn}(\hat{e})=-1$. We use $P(-a,-b,-c,d,e)$ rather than its mirror in order to use the positive definite formulation of Donaldson's theorem. As stated in the explanation of the lattice embedding condition, we wish to write down a matrix A satisfying $A^TA=Q_{P_+}$. This condition can be phrased as a collection of conditions on the column vectors of A:

Embedding conditions For a slice odd 5–stranded pretzel knot K of the form $P(-a,-b,-c,d,e)$, there exist vectors $v_i, v_{j,r}\in\mathbb{Z}^m$, with $m=a+b+c$, satisfying

$$
\begin{array}{llll}
(0) & v_0 \cdot v_0 = 3, & (1) & v_1 \cdot v_0 = 1, \\
(2) & v_2 \cdot v_0 = 1, & (3) & v_1 \cdot v_2 = 0, \\
(4) & v_1 \cdot v_1 = d, & (5) & v_2 \cdot v_2 = e, \\
(6) & v_{j,1} \cdot v_0 = 1, & (7) & v_{j,r} \cdot v_{j,r} = 2, \\
(8) & v_{j,r} \cdot v_{j,r\pm 1} = 1 & \multicolumn{2}{l}{\text{for } r \geq 2,} \\
(9) & v_{j,r} \cdot v_* = 0 & \multicolumn{2}{l}{\text{for } r \geq 2 \text{ and all vectors } v_* \neq v_{j,r\pm 1}.}
\end{array}
$$

The embedding conditions impose severe restrictions on the form each v_i and v_{j,r_j} can take. Condition (0) for example, implies that v_0 must have exactly three entries equal to ± 1 and zeros otherwise; similarly, condition (7) implies that each vector v_{j,r_j} must have exactly two entries equal to ± 1 and zeros otherwise. It can be verified using conditions (0)–(7) that up to a change of basis, A will have the following form, with $\alpha, \beta, \gamma, x, y, z \in \mathbb{Z}$:

$$
\left[
\begin{array}{ccc|cccc|cccc|cccc}
v_0 & v_1 & v_2 & v_{a,1} & v_{a,2} & \cdots & v_{a,a-1} & v_{b,1} & v_{b,2} & \cdots & v_{b,b-1} & v_{c,1} & v_{c,2} & \cdots & v_{c,c-1} \\
\hline
1 & \alpha & x & 1 & 0 & \cdots & 0 & & & & & & & & \\
0 & \alpha & x & -1 & -1 & & 0 & & & & & & & & \\
0 & \alpha & x & 0 & 1 & & 0 & & & & & & & & \\
0 & \alpha & x & 0 & 0 & \cdots & 0 & \multicolumn{4}{c|}{0} & \multicolumn{4}{c}{0} \\
 & \vdots & & \vdots & & \ddots & \vdots & & & & & & & & \\
0 & \alpha & x & 0 & 0 & \cdots & -1 & & & & & & & & \\
0 & \alpha & x & 0 & 0 & \cdots & 1 & & & & & & & & \\
\hline
1 & \beta & y & & & & & 1 & 0 & \cdots & 0 & & & & \\
0 & \beta & y & & & & & -1 & -1 & & 0 & & & & \\
0 & \beta & y & & & & & 0 & 1 & & 0 & & & & \\
0 & \beta & y & \multicolumn{4}{c|}{0} & 0 & 0 & \cdots & 0 & \multicolumn{4}{c}{0} \\
 & \vdots & & & & & & \vdots & & \ddots & \vdots & & & & \\
0 & \beta & y & & & & & 0 & 0 & \cdots & -1 & & & & \\
0 & \beta & y & & & & & 0 & 0 & \cdots & 1 & & & & \\
\hline
1 & \gamma & z & & & & & & & & & 1 & 0 & \cdots & 0 \\
0 & \gamma & z & & & & & & & & & -1 & -1 & & 0 \\
0 & \gamma & z & & & & & & & & & 0 & 1 & & 0 \\
0 & \gamma & z & \multicolumn{4}{c|}{0} & \multicolumn{4}{c|}{0} & 0 & 0 & \cdots & 0 \\
 & \vdots & & & & & & & & & & \vdots & & \ddots & \vdots \\
0 & \gamma & z & & & & & & & & & 0 & 0 & \cdots & -1 \\
0 & \gamma & z & & & & & & & & & 0 & 0 & \cdots & 1
\end{array}
\right]
$$

Having A in this explicit form allows us to put restrictions on the unordered 5–tuples $\{a,b,c,d,e\}$ ensuring the embedding conditions are satisfied. For fixed a, b, and c, we enumerate the embedding conditions in terms of the entries of the column vectors of A, which reduces the problem of finding the desired embedding to the problem of finding integers α, β, γ, x, y, z that satisfy the following system of nonlinear equations. Each new condition is numbered to correspond to the original embedding condition that implies it. In references by number, no distinction is made between the original and updated conditions since the updated conditions are direct implications of the originals.

(Updated) embedding conditions For a slice odd 5–stranded pretzel knot K of the form $P(-a,-b,-c,d,e)$, there exist integers α, β, γ, x, y, z satisfying

(1) $\alpha+\beta+\gamma=1$,

(2) $x+y+z=1$,

(3) $a\alpha x+b\beta y+c\gamma z=0$,

(4) $a\alpha^2+b\beta^2+c\gamma^2=d$,

(5) $ax^2+by^2+cz^2=e$.

In fact, these updated embedding conditions are exactly the contents of Theorem 4.1.6 in [13], so a more detailed account of these facts can be found there.[5]

6 The d–invariants and the coset conditions

Peter Ozsváth and Zoltán Szabó defined the d–invariant $d(Y,\mathfrak{s})\in\mathbb{Q}$ in the setting of Heegaard Floer homology for a rational homology 3–sphere Y equipped with a Spin^c structure $\mathfrak{s}$. While the d–invariant has an important function as a correction term for the grading in Heegaard Floer homology, it is significant in 4–manifold topology because it is a Spin^c rational homology bordism invariant. As stated in [16], if $(Y_1,\mathfrak{s}_1)$ and $(Y_2,\mathfrak{s}_2)$ are two pairs such that Y_i is a rational homology 3–sphere and $\mathfrak{s}_i$ is a Spin^c structure on Y_i, then if there exists a connected, oriented, smooth cobordism W from Y_1 to Y_2 with $H_i(W;\mathbb{Q})=0$ for $i=1,2$ which can be endowed with a Spin^c structure $\mathfrak{t}$ whose restriction to Y_i is $\mathfrak{s}_i$, then $d(Y_1,\mathfrak{s}_1)=d(Y_2,\mathfrak{s}_2)$. The proof of this highly nontrivial fact is given in Proposition 9.9 of [16], and it has the following corollary:

[5] Warning: Long's approach to the problem of sliceness in 5–stranded pretzel knots uses a negative definite convention rather than the positive definite convention of this paper.

Corollary 6.1 (Ozsváth and Szabó) *Let Y be a rational homology 3–sphere with* Spin^c *structure $\mathfrak{s}$, and let W be a rational homology 4–ball with $\partial W = Y$ and* Spin^c *structure $\mathfrak{t}$. If $\mathfrak{s}$ can be extended over W so that $\mathfrak{t}|_Y = \mathfrak{s}$, then $d(Y, \mathfrak{s}) = 0$.*

In general $d(Y, \mathfrak{s})$ may be hard to compute, but in [17] Ozsváth and Szabó give a formula for $d(Y, \mathfrak{s})$ when Y is the boundary of a 4–dimensional plumbing manifold P. Their formula holds in more generality than the version presented below, but the formula is stated here in the special case relevant to the present situation of determining sliceness of pretzel knots. Throughout this section, we refer to K, Y, P, W, and X as defined in Section 5. To remind the reader of these definitions: K is assumed to be a slice odd pretzel knot; Y is the double branched cover of S^3 along K; W is the double branched cover of B^4 along a fixed slice disk for K with $\partial W = Y$; $P = P_+$ is a positive definite plumbing manifold with $\partial P = Y$; and $X = P \cup_Y (-W)$ is a closed positive definite manifold. Under these assumptions, W is a rational homology 4–ball and Y is a rational homology 3–sphere. To state the aforementioned formula easily and to give a more geometric flavor to the material that follows, we first discuss an identification of $\mathrm{Spin}^c(Y)$ with $H_1(Y)$.

If Y is a 3–manifold such that $H_1(Y)$ is odd torsion, then there is a natural identification of $\mathrm{Spin}^c(Y)$ with $H_1(Y)$. In our current work Y is the double branched cover of S^3 along a knot K and a bit of straightforward algebraic topology reveals that $H_1(Y)$ is always odd torsion in this case. The first step in the identification shows a one-to-one correspondence between $\mathrm{Spin}^c(Y)$ and vect(Y), the set of Euler structures on Y. An Euler structure on a smooth closed connected oriented 3–manifold Y is an equivalence class of nonsingular tangent vector fields on Y, where two vector fields u and v on Y are deemed equivalent if u and v are homotopic as nonsingular vector fields outside of some closed 3–dimensional ball. This particular identification of $\mathrm{Spin}^c(Y)$ with vect(Y) is due to Vladimir Turaev and constitutes Lemma 1.4 in [19], so the reader is directed there for details. The salient feature of this step is that it allows us to view a Spin^c structure on Y as a vector field over Y under some notion of equivalence.

Assuming Turaev's identification of $\mathrm{Spin}^c(Y)$ with vect(Y), the second step is to identify vect(Y) with $H_1(Y)$. Start by fixing a trivialization τ of the tangent bundle TY. Since $H_1(Y)$ has only odd torsion, τ is unique off a 3–ball up to homotopy. Let $[Y, S^2]$ denote the space of smooth maps from Y to S^2 up to homotopy. The identification of vect(Y) with $H_1(Y)$ will be done via a composition vect(Y) $\to [Y, S^2] \to H_1(Y)$.

For each equivalence class of nonvanishing vector fields on Y, we choose a representative vector field $\mathcal{V}$. By a straight line homotopy, we can assume that each vector in $\mathcal{V}$ is a unit vector, where the length is measured according to the trivialization τ. For each point $p \in Y$, the tangent space T_pY at p is isomorphic to R^3 and thus provides a way to give Euclidean coordinates to $v_p \in \mathcal{V}$, the vector based at p. Let (x_p, y_p, z_p) be the Euclidean coordinates for v_p obtained from τ. With this we define a smooth map $\phi_{\mathcal{V}}\colon Y \to S^2$ by sending $p \in Y$ to the vector $(x_p, y_p, z_p) \in S^2$. The map $g\colon \mathrm{vect}(Y) \to [Y, S^2]$ is then defined as $g([\mathcal{V}]) = \phi_{\mathcal{V}}$.

For the second map, recall that any map $\phi \in [Y, S^2]$ has the property that the preimage $\phi^{-1}(z)$ of a regular value $z \in S^2$ will be a submanifold of Y of codimension 2, namely, a curve γ. From this fact, define $h\colon [Y, S^2] \to H_1(Y)$ by $h(\phi) = [\gamma]$, where γ is the preimage of any regular value of ϕ. The Pontryagin–Thom construction shows that this map is well defined. It follows that Spin^c is identified with $H_1(Y)$ via Turaev's identification of Spin^c with $\mathrm{vect}(Y)$ followed by the composition $h \circ g \circ f$.

A second topic necessary to discuss before stating the d–invariant formula is that of characteristic elements of $H_2(X)/\mathrm{Tor}$, $H_2(P)$, and $H_2(P, Y)$. These definitions involve intersection numbers, and in all cases we will abbreviate the intersection number of two elements a, b in $H_2(X)/\mathrm{Tor}$, $H_2(P)$, or $H_2(P, Y)$ by $a \cdot b$ and let the definition of $a \cdot b$ be given by context. As before, Q_X and Q_P are the intersection forms on X and P, respectively. The map Q_P^{-1} is the relative intersection form on (P, Y) given as the inverse of Q_P over $\mathbb{Q}$. We define:

- $a \cdot b = Q_X(a, b)$ if $a, b \in H_2(X)/\mathrm{Tor}$.
- $a \cdot b = Q_P(a, b)$ if $a, b \in H_2(P)$.
- $a \cdot b = Q_P(x, b)$ if $a \in H_2(P, Y)$ and $b \in H_2(P)$, where $x = Q_P^{-1}(a) \in H_2(P)$.
- $a \cdot b = Q_P(x, y)$ if $a, b \in H_2(P, Y)$, where again $x = Q_P^{-1}(a) \in H_2(P)$ and $y = Q_P^{-1}(b) \in H_2(P)$.

By choosing bases for $H_2(X)/\mathrm{Tor}$, $H_2(P)$, and $H_2(P, Y)$, homology classes in these groups can be represented by column vectors and the intersection forms Q_X and Q_P can be represented by matrices. We choose bases as follows: the basis $\{e_i\}$ for $H_2(X)/\mathrm{Tor}$ is the one that makes Q_X diagonal by Donaldson's theorem; the basis $\{s_i\}$ for $H_2(P)$ is the set of homology classes represented by the zero-sections of the disk bundles used to create P; lastly, the basis $\{d_i\}$ for $H_2(P, Y)$ is the set of classes represented by single fiber disks in each of the disk bundles of P. Note that the fiber disks $\{d_i\}$ are the Hom-duals of the $\{s_i\}$.

With fixed bases the above intersection numbers can be computed using column vector representatives for homology classes and the matrix representatives for Q_X and Q_P. As matrices with the above bases, recall that Q_P is equal to the incidence matrix of the weighted graph representing P and Q_X is equal to the identity matrix of rank $b_2(X)$. By an abuse of notation, we use Q_P to denote both the intersection form for P and its matrix representative in this case. This allows us to write and compute the above intersection numbers in terms of column vectors a, b as follows:

- If $a, b \in H_2(X)/\operatorname{Tor}$, then $a \cdot b = a^T b$.

- If $a, b \in H_2(P)$, then $a \cdot b = a^T Q_P b$.

- If $a \in H_2(P, Y)$ and $b \in H_2(P)$, then $a \cdot b = x^T Q_P b$, where $x = Q_P^{-1}(a) \in H_2(P)$. This simplifies to $a \cdot b = a^T b$.

- If $a, b \in H_2(P, Y)$, then $a \cdot b = Q_P(x, y)$, where again $x = Q_P^{-1}(a) \in H_2(P)$ and $y = Q_P^{-1}(b) \in H_2(P)$. This simplifies to $a \cdot b = a^T Q_P^{-1} b$.

Now, we say that an absolute class $w \in H_2(X)/\operatorname{Tor}$ is a *characteristic class of X* if $w \cdot x \equiv x \cdot x \pmod 2$, for all $x \in H_2(X)/\operatorname{Tor}$; we say a characteristic class w is *minimal* if $w \cdot w \leq z \cdot z$ for all characteristic classes z. Characteristic and minimal characteristic elements of $H_2(P)$ are defined similarly. A relative class $w \in H_2(P, Y)$ is *characteristic in X with respect to $\mathfrak{s}$*, where $\mathfrak{s}$ is regarded as an element of $H_1(Y)$, if $\partial w = \mathfrak{s}$ and $w \cdot u \equiv u \cdot u \pmod 2$, for all $u \in H_2(P)$. The set of characteristic elements in $H_2(P, Y)$ relative to $\mathfrak{s}$ is denoted by $\operatorname{Char}_{\mathfrak{s}}(P)$, which makes an appearance in the formula below.

We are now ready to state Ozsváth and Szabó's formula for $d(Y, \mathfrak{s})$ in the case that Y bounds a certain type of 4–dimensional plumbing:

Theorem 6.2 (Ozsváth and Szabó) *Let P be a 4–dimensional plumbing with positive definite intersection form Q_P, such that the weighted graph of P has at most two vertices whose weights are less than their valences. Then under the identification* $\operatorname{Spin}^c(Y) \to H_1(Y)$,

$$d(Y, \mathfrak{s}) = \min_{w \in \operatorname{Char}_{\mathfrak{s}}(P)} \frac{w \cdot w - \sigma(P)}{4}. \tag{1}$$

In [3], Greene and Jabuka use Theorem 6.2 and Corollary 6.1 to give an obstruction to sliceness for odd pretzel knots through some analysis of the cohomology long exact sequences of the pairs (P, Y) and (X, W). Here, we derive their results in terms of homology and obtain the following commutative diagram at the top of the next page. In the diagram the horizontal maps arise from the long exact sequences of the pairs (P, Y) and (X, W); the vertical maps r and γ are induced by inclusions; β is an isomorphism due to excision; and q is the usual quotient map.

$$\begin{tikzcd}
0 \arrow[r] & H_2(P) \arrow[r, "\lambda"] \arrow[d, "r"] & H_2(P,Y) \arrow[r, "\partial"] \arrow[d, "\beta"', "\cong"] & H_1(Y) \arrow[r] \arrow[d, "\gamma"] & 0 \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & H_2(X) \arrow[r, "g"] \arrow[d, "q"] & H_2(X,W) \arrow[r] & H_1(W) \arrow[r] & H_1(X) \arrow[r] & 0 \\
\cdots \arrow[r] & H_2(X)/\operatorname{Tor} \arrow[ur, "\mu"] & & & &
\end{tikzcd}$$

Because $H_2(P)$ is free and r is a homomorphism, the image of r lies entirely in the free part of $H_2(X)$. Let $\alpha = qr$ and $\alpha^* = \beta^{-1}\mu$; this allows us to use the first isomorphism theorem to eliminate $H_2(X)$ from the diagram. By commutativity, λ can be seen to have the factorization $\lambda = \alpha^*\alpha$, converting the previous diagram into:

$$\begin{tikzcd}
0 \arrow[r] & H_2(P) \arrow[r, "\lambda"] \arrow[d, "\alpha"'] & H_2(P,Y) \arrow[r, "\partial"] \arrow[d, "\beta"', "\cong"] & H_1(Y) \arrow[r] \arrow[d, "\gamma"] & 0 \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & H_2(X)/\operatorname{Tor} \arrow[r, "\mu"] \arrow[ur, "\alpha^*"] & H_2(X,W) \arrow[r, "\nu"] & H_1(W) \arrow[r] & H_1(X) \arrow[r] & 0
\end{tikzcd}$$

To use this diagram in conjunction with the lattice embedding condition, it is advantageous to work with matrix representatives of the maps α, α^*, and λ. We choose the bases for $H_2(P)$, $H_2(P,Y)$, and $H_2(X)/\operatorname{Tor}$ as before, we let A be the matrix representative for the map α (induced by the embedding of P into X), and we let A^* be the matrix for α^*.

The columns of A express the basis elements $\{e_i\}$ of $H_2(X)/\operatorname{Tor}$ in terms of the basis disks $\{d_i\}$ for $H_2(P,Y)$. Consequently, the *rows* of A^T express the spheres $\{s_i\}$ in terms of the $\{e_i\}$, which implies that the ij^{th} entry in A^TA gives the intersection number between the spheres s_i and s_j. Thus A^TA is the matrix of the intersection form Q_P of P with respect to the basis $\{s_i\}$.

Recall that each basis element d_i of $H_2(P,Y)$ is the Hom-dual of the basis element s_i of $H_2(P)$, and therefore $\lambda(s_i) = \sum_j (s_i \cdot s_j) d_j$. This implies that with respect to the chosen bases, λ (as a linear map from $H_2(P)$ to $H_2(P,Y)$) is represented by the same matrix as is Q_P (regarded as a bilinear map from $H_2(P) \times H_2(P)$ to $\mathbb{Z}$). Namely, λ is also represented by A^TA. Given that $\lambda = \alpha^*\alpha$, it follows that $A^*A = A^TA$ as matrices. Since Q_P is invertible over $\mathbb{Q}$, so is A; whence $A^* = A^T$. By reinstating the abuse in notation whereby we use Q_P to denote both the intersection form on P

and its matrix representative in this case, we let the matrix Q_P represent λ with respect to the chosen bases.

Dropping the less relevant maps, the previous commutative diagram becomes:

$$\begin{tikzcd}
0 \arrow[r] & H_2(P) \arrow[r, "Q_P"] \arrow[d, "A"'] & H_2(P,Y) \arrow[r] \arrow[d, "\cong"] & H_1(Y) \arrow[r] \arrow[d, "\gamma"] & 0 \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & H_2(X)/\operatorname{Tor} \arrow[r] \arrow[ur, "A^T"] & H_2(X,W) \arrow[r] & H_1(W) \arrow[r] & H_1(X) \arrow[r] & 0
\end{tikzcd}$$

We use this to restate and reprove — in a homological setting — Greene and Jabuka's d–invariant obstruction to sliceness in odd pretzel knots:

Theorem 6.3 (Greene and Jabuka) *Let K be a slice odd pretzel knot with Y, W, $P = P_+$, and X as in the above commutative diagram. Then every coset of* $\operatorname{coker}(\alpha)$ *contains a minimal characteristic class of $H_2(X)/\operatorname{Tor}$.*

Proof Under the assumption that K is slice, it follows that K satisfies the embedding conditions and $\sigma(P) = \operatorname{rk}(Q_P) = \operatorname{rk}(Q_X) = b_2(X) := m$. It also follows from Corollary 6.1 that $d(Y,\mathfrak{s}) = 0$ for every $\mathfrak{s}$ that extends over W. In general, the Spin^c structures on a rational homology 3–sphere Y that extend over a rational homology 4–ball W are identified with precisely those elements in $H_1(Y)$ that bound relative homology classes in $H_2(W,Y)$. As such, they are in one-to-one correspondence with the elements of $\ker(\gamma)$, where $\gamma\colon H_1(Y) \to H_1(W)$ is induced by inclusion.

Theorem 6.2 applies to Y since the plumbing graph of P will have exactly one vertex whose weight is less than its valence, namely, the central node. Ozsváth and Szabó's formula

$$d(Y,\mathfrak{s}) = \min_{w \in \mathrm{Char}_{\mathfrak{s}}(P)} \frac{w\cdot w - \sigma(P)}{4}$$

implies that $d(Y,\mathfrak{s}) = 0$ if and only if there exists $w \in \mathrm{Char}_{\mathfrak{s}}(P)$ such that $w\cdot w = m$. A straightforward diagram chase shows that for each $w \in \mathrm{Char}_{\mathfrak{s}}(P)$ there exists an element $x \in H_2(X)/\operatorname{Tor}$ such that $\alpha^*(x) = w$. In addition, x is characteristic in X and $x\cdot x = w\cdot w$, so in general the characteristic classes of P relative to $\mathfrak{s}$ correspond to absolute characteristic classes of X with equal intersection number. This fact, which is verified below, allows us to compute $w\cdot w$, which appears in formula (1), by using $x\cdot x$ instead.

Fix the bases for $H_2(P)$, $H_2(P,Y)$, and $H_2(X)/\operatorname{Tor}$ as before, and let $A = (a_{ij})$ again be the matrix representative of α with respect to these bases. Let $w \in \mathrm{Char}_{\mathfrak{s}}(P)$

and $x = (x_1, \dots, x_m) \in H_2(X)/\operatorname{Tor}$ such that $\alpha^*(x) = w$. Recall that α^* is represented by the matrix A^T with respect to these bases. To show that x is characteristic in X, it suffices to show that $x \cdot e_j \equiv e_j \cdot e_j \pmod 2$, for all basis elements e_j. Since $e_j \cdot e_j = 1$, one need only show that $x \cdot e_j$ —that is, the j^{th} component of x —is odd for all j. Stated differently, it must be shown that every component of x is odd.

By definition of $\operatorname{Char}_{\mathfrak{s}}(P)$, $w \cdot u \equiv u \cdot u \pmod 2$ for all $u \in H_2(P)$, in particular for u equal to a basis element s_j for $H_2(P)$: $w \cdot s_j \equiv s_j \cdot s_j \pmod 2$. Observe that for all j,

$$w \cdot s_j = A^T x \cdot s_j = x^T A s_j = \sum_i x_i a_{ij},$$

$$s_j \cdot s_j = (Q_P)_{jj} = (A^T A)_{jj} = \sum_i a_{ij} a_{ij} \equiv \sum_i a_{ij} \pmod 2.$$

Hence, $\sum_i x_i a_{ij} \equiv \sum_i a_{ij} \pmod 2$ for all j. Letting $x_i \equiv 1 \pmod 2$ yields a solution to this equation, which in fact is the unique solution since A is invertible modulo 2. Given that this holds for all j, it has thus been shown that x has all odd entries and is therefore characteristic in X. Furthermore, since $w \in H_2(P, Y)$, it follows from above that $w \cdot w = w^T Q_P^{-1} w$. Making the substitutions $Q_P = A^T A$ and $w = A^T x$ shows that $w \cdot w = x \cdot x$.

In addition, the diagram chase from before shows that $\ker(\gamma) \cong \operatorname{coker}(\alpha)$. Combining this with the preceding information implies that $d(Y, \mathfrak{s}) = 0$ for all $\mathfrak{s} \in \ker(\gamma)$ with corresponding $k \in \operatorname{coker}(\alpha)$ if and only if there exists $w \in \operatorname{Char}_{\mathfrak{s}}(P)$ and $x \in \operatorname{Char}(X)$ such that $w = A^T x$, $x \cdot x = m$, and $x + \operatorname{im}(\alpha) = k$. Clearly, $x \cdot x = m$ only if $x_i = \pm 1$ for all i, which implies that x is a *minimal* characteristic class of X. Hence, K slice implies that every element of $\operatorname{coker}(\alpha)$, ie every coset of $\operatorname{im}(\alpha)$, contains a minimal characteristic class of X. □

Theorem 6.3 gives a necessary condition for sliceness for odd 5–stranded pretzel knots that can be rephrased in a simpler, more geometric way by analyzing the quotient $\operatorname{coker}(\alpha) = (H_2(X)/\operatorname{Tor})/\operatorname{im}(\alpha)$. We will reduce the problem of finding minimal characteristic vectors in each coset of $\operatorname{im}(\alpha)$ to a more visualizable problem of finding lattice points in $\mathbb{Z}^2$ with certain properties.

Since $H_2(X)/\operatorname{Tor} \cong \mathbb{Z}^m$, it follows that $\operatorname{coker}(\alpha) \cong \mathbb{Z}^m/\operatorname{im}(\alpha)$. Given that the image of α with the chosen bases is equal to the span of the columns of A, $\operatorname{coker}(\alpha)$ is isomorphic to the quotient of $\mathbb{Z}^m$ by the columns of A. Let $\mathcal{U} = \{v_{j,r_j}\}$ be the set of column vectors of A with standard dot product, where $1 \le j \le n$ and $1 \le r_j \le j-1$. Then the columns of A, as vectors, are given by $\{v_0, v_1, v_2, \mathcal{U}\}$.

Define $B\colon \mathbb{Z}^m \to \mathbb{Z}^2$ by

$$(x_1,\dots,x_a,y_1,\dots,y_b,z_1,\dots,z_c)^T \mapsto \left(\sum_{i=1}^{a} x_i - \sum_{k=1}^{c} z_k, \sum_{j=1}^{b} y_j - \sum_{k=1}^{c} z_k\right)^T.$$

It is straightforward to see that $\ker(B) = \langle v_0, \mathcal{U}\rangle$ and that B is onto, so by the first isomorphism theorem $\mathbb{Z}^2 \cong \mathbb{Z}^m/\langle v_0,\mathcal{U}\rangle$. It follows that

$$\operatorname{coker}(\alpha) \cong \mathbb{Z}^m/\langle v_0, v_1, v_2, \mathcal{U}\rangle \cong \mathbb{Z}^2/\langle B(v_1), B(v_2)\rangle.$$

Let $\overline{v}_1 := B(v_1)$ and $\overline{v}_2 := B(v_2)$. Using the above isomorphisms, the slice condition in Theorem 6.3 can now be rephrased to say that every coset in $\mathbb{Z}^2/\langle \overline{v}_1, \overline{v}_2\rangle$ must have a representative in $B(\{\pm 1\}^m)$. Thus, $\mathbb{Z}^2/\langle \overline{v}_1, \overline{v}_2\rangle$ and $B(\{\pm 1\}^m)$ are analyzed:

$$\begin{aligned}\overline{v}_1 &= B((\alpha,\dots,\alpha,\beta,\dots,\beta,\gamma,\dots,\gamma))^T = (a\alpha - c\gamma, b\beta - c\gamma)^T,\\ \overline{v}_2 &= B((x,\dots,x,y,\dots,y,z,\dots,z))^T\ = (ax - cz, by - cz)^T,\\ B(\{\pm 1\}^m) &= \{(q-s, r-s) \mid -a-c \le q-s \le a+c \text{ and } -b-c \le r-s \le b+c\}.\end{aligned}$$

The vectors $\{\overline{v}_1, \overline{v}_2\}$ define a fundamental domain $R \subset \mathbb{R}^2$. Let $\mathcal{R}$ be the set of lattice points in R that represent unique cosets of $\operatorname{im}(\alpha)$ coming from $\operatorname{coker}(\alpha)$. Note that $\mathcal{R} \subset \mathbb{Z}^2$ and $\mathcal{R} \subset R$. Also, let $\mathcal{H} := B(\{\pm 1\}^m)$. Then $\mathcal{H}$ is a collection of lattice points (x,y) satisfying $-a-c \le x \le a+c$ and $-b-c \le y \le b+c$. Since a, b, and c are odd and positive, every element of $\mathcal{H}$ is an element of $2\mathbb{Z}^2$ and collectively these points lie in a hexagonal region $H \subset \mathbb{R}^2$. Hence, $\mathcal{H} \subset 2\mathbb{Z}^2$ and $\mathcal{H} \subset H$.

Greene and Jabuka observed that Theorem 6.3, which gives the slice condition that every element of $\operatorname{coker}(\alpha)$ contains a minimal characteristic vector of the form $(\{\pm 1\}^n)$, is equivalent to the condition that every lattice point of $\mathcal{R}$ can be translated onto a lattice point of $\mathcal{H}$ by a linear combination of $\overline{v}_1$ and $\overline{v}_2$. Hence, a knot K cannot be slice if there exists an element of $\operatorname{coker}(\alpha)$ that does not contain a minimal characteristic vector of X by Theorem 6.3. The correspondence between cosets of $\operatorname{im}(\alpha)$, minimal characteristic vectors of X, and lattice points implies that K is not slice if there exists a lattice point in $\mathcal{R}$ that can not be translated onto a lattice point in $\mathcal{H}$ by a linear combination of $\overline{v}_1$ and $\overline{v}_2$. By the definition of $\mathcal{R}$, every element of $\mathcal{R}$ represents a distinct coset in the quotient $\mathbb{Z}^2/\langle \overline{v}_1, \overline{v}_2\rangle$, and thus if there are more lattice points in $\mathcal{R}$ than there are in $\mathcal{H}$ for a knot K, then K is not slice. This proves the following:

Coset condition I *If $P(a,b,c,d,e)$ is a slice odd 5–stranded pretzel knot, then $|\mathcal{R}| \le |\mathcal{H}|$.*

It is possible, however, for many points in $\mathcal{H}$ to belong to the same coset in $\mathbb{Z}^2/\langle \overline{v}_1, \overline{v}_2\rangle$. Let $\overline{\mathcal{H}} := \mathcal{H}/\langle \overline{v}_1, \overline{v}_2\rangle$, so that $|\overline{\mathcal{H}}|$ is the number of $\langle \overline{v}_1, \overline{v}_2\rangle$–cosets in $\mathcal{H}$. With this observation and Theorem 6.3 and the above observation, K is not slice if $|\mathcal{R}| > |\overline{\mathcal{H}}|$. This condition is a refinement of coset condition I, which the author unimaginatively calls coset condition II:

Coset condition II *If $P(a,b,c,d,e)$ is a slice odd 5–stranded pretzel knot, then $|\mathcal{R}| \le |\overline{\mathcal{H}}|$.*

7 Proof of Theorem 2.3

Due to the slightly different nature of pretzel knots with single-twists versus those without, the proof of Theorem 2.3 is divided according to this distinction. A technical lemma, Lemma 7.1, is given first and then it is shown that all 0–pair odd 5–stranded pretzel knots *without* single-twists are not slice. Lemma 9.1 refines Lemma 7.1 and is then used to show that all 0–pair odd 5–stranded pretzel knots *with* single-twists are not slice.

Recall from Section 4 that a knot K is slice if and only if its mirror $-K$ is slice. To make the computations in the proof easier, the knot $K = P(-a,-b,-c,d,e)$ in $P\{a,b,c,-d,-e\}$ will be used rather than its mirror $P(a,b,c,-d,-e)$. Lemma 7.1, which is given next, states the conditions on α, β, γ, x, y, and z under which $P(-a,-b,-c,d,e)$ will be a 0–pair pretzel knot. Without loss of generality, assume throughout that $a \le b \le c$.

Lemma 7.1 *If $K \in P\{a,b,c,-d,-e\}$ is 0–pair and satisfies the embedding conditions, then at most one of α, β, γ, x, y, z is zero. Furthermore, if the set $\{\alpha,\beta,\gamma,x,y,z\}$ contains 0, then $d \ge 4a+b$ and $e \ge a+b+c$; otherwise, both d and e are greater than or equal to $a+b+c$.*

Proof Choose $K = P(-a,-b,-c,d,e)$. First it will be shown that if any two of α, β, γ are zero or if any two of x, y, z are zero, then K is not 0–pair. By the symmetry of the embedding conditions on $\{\alpha,\beta,\gamma\}$ and $\{x,y,z\}$, it suffices to prove this only for $\{\alpha,\beta,\gamma\}$. Suppose two of the parameters α, β, γ are zero. Then the third parameter is equal to 1 by embedding condition (1), and thus $d \in \{a,b,c\}$ by embedding condition (4). Consequently, K has at least one pair of canceling twist parameters, a contradiction. It follows that at most one of α, β, γ is zero and at most one of x, y, z is zero. The remainder of the proof consists in showing the stronger statement that the sets $\{\alpha,\beta,\gamma\}$ and $\{x,y,z\}$ cannot *both* contain 0.

Without loss of generality, suppose $\alpha = 0$ and $\beta \ne \gamma \ne 0$. It will be shown that if any of x, y, z is zero, then either K is not 0–pair or there is a contradiction to $x,y,z \in \mathbb{Z}$.

With the assumptions on α, β, and γ, embedding conditions (1) and (4) immediately yield $d = b\beta^2 + c\gamma^2 \leq 4b + c \leq 4a + b$, and embedding condition (3) implies

$$b\beta y = -c\gamma z. \tag{2}$$

From (2), if either one of y or z is 0, then so is the other. Thus, K is not 0–pair by the first paragraph of the proof and therefore y and z are both nonzero.

If $x = 0$, embedding condition (2) implies that $z = 1 - y$. Substituting this into (2) and solving for y yields

$$y = \frac{c\gamma}{c\gamma - b\beta}. \tag{3}$$

Since $\alpha = 0$, embedding condition (1) implies that $\beta + \gamma = 1$. If either one of β or γ is equal to 1, then the other vanishes. This violates the assumption that $\beta \neq \gamma \neq 0$; therefore $\beta, \gamma \notin \{0, 1\}$. Note that β and γ always have different signs. If $\gamma \geq 2$, then $\beta \leq -1$ and thus $-b\beta > 0$. Thus, (3) takes on the form

$$y = \frac{p}{p+q}, \tag{4}$$

where $p, q \in \mathbb{Z}^+$. Thus y cannot be an integer, contradicting the embedding conditions. If instead $\gamma \leq -1$, then $\beta \geq 2$. In this case, one can take (2) and solve for z instead of y, yielding

$$z = \frac{b\beta}{b\beta - c\gamma}. \tag{5}$$

By the same argument given for $\gamma \geq 2$, if $\beta \geq 2$ it follows that z cannot be an integer and the embedding conditions are again contradicted. Thus if $\alpha = 0$, each of x, y, z must be nonzero and $e = ax^2 + by^2 + cz^2 \geq a + b + c$ by embedding condition (5).

If $\beta = 0$, the proof follows similarly with $d = a\alpha^2 + c\gamma^2 \geq 4a + c \geq 4a + b$; if $\gamma = 0$, then again the proof follows similarly with $d = a\alpha^2 + b\beta^2 \geq 4a + b$. In all three cases, $e = ax^2 + by^2 + cz^2 \geq a + b + c$. Lastly, if none of α, β, γ, x, y, z is zero, then embedding conditions (4) and (5) imply that $d = a\alpha^2 + b\beta^2 + c\gamma^2 \geq a + b + c$ and $e = ax^2 + by^2 + cz^2 \geq a + b + c$, since $c \geq b \geq a \geq 1$. $\square$

8 Proof of Theorem 2.3 without single-twists

The proof of Theorem 2.3 will now proceed by showing that if K is a 0–pair odd 5–stranded pretzel knot without single-twists, then coset condition I is violated. Assume K is slice. It follows that K satisfies the signature condition and the lattice embedding condition. Furthermore, it may also be assumed that $K = P(-a,-b,-c,d,e)$ with $a \le b \le c$. Let $\mathcal{R}$ and $\mathcal{H}$ be as in coset condition I.

The fact that K is 0–pair implies that $d \ge 4a+b$ and $e \ge a+b+c$ by Lemma 7.1. Given that $\ker(\gamma) \cong \operatorname{coker}(\alpha)$ (where γ and α here refer to the maps in Section 6) and $|\ker(\gamma)| = \sqrt{|H_1(Y)|} = \sqrt{|\det(K)|}$, it follows that $|\mathcal{R}| = |\operatorname{coker}(\alpha)| = \sqrt{|\det(K)|}$. Theorem 1.4 in [5] gives the following formula for the determinant of odd pretzel knots $P(p_1,\dots,p_k)$:

$$\det(K) = \sum_{i=1}^{k} p_1 \cdots p_{i-1} \widehat{p}_i p_{i+1} \cdots p_k.$$

Using this with the above choice of K, one gets

$$\det(K) = -abcd - abce + abde + acde + bcde.$$

To compute $|\overline{\mathcal{H}}|$, a direct computation shows that the closed hexagonal region H in which $\mathcal{H}$ is contained is a region in $\mathbb{R}^2$ defined by the $2(a+c) \times 2(b+c)$ rectangle centered at the origin, minus the lower-right and upper-left half-square triangular regions with side lengths $2c$. See Figure 13. The set $\mathcal{H}$ contains all lattice points in $2\mathbb{Z}^2$ in the interior of H and on the boundary of H. These can be counted in many different ways but are counted here by observing that there are $a+b+1$ even lattice points in the perpendicular boundary components of H lying in the third quadrant of $\mathbb{R}^2$, and there are $c+1$ copies of this L-shaped collection of even lattice points repeated throughout H; furthermore, H includes a $a \times b$ rectangle of even lattice points. Hence

$$|\mathcal{H}| = (a+b+1)(c+1) + ab = ab + ac + bc + a + b + c + 1.$$

To violate coset condition I, it will be argued that $|\mathcal{R}|^2 > |\mathcal{H}|^2$ using the facts that

(1) $3 \le a \le b \le c$,

(2) $d \ge 4a+b$ and $e \ge a+b+c$ or $d, e \ge a+b+c$, and

(3) $ab > a+b+\frac{1}{2}$ for $a, b \ge 3$.

By [13, Theorem 2.0.3], 0–pair odd 5–stranded pretzel knots $P(-a,-b,-c,d,e)$ are not slice if $d, e \ge a+b+c$, thus that case is omitted here. Hence, we assume $d \ge 4a+b$ and $e \ge a+b+c$. First consider $|\mathcal{H}|^2$:

$$|\mathcal{H}|^2 = (ab + ac + bc + a + b + c + 1)^2 = L + M + N + S,$$

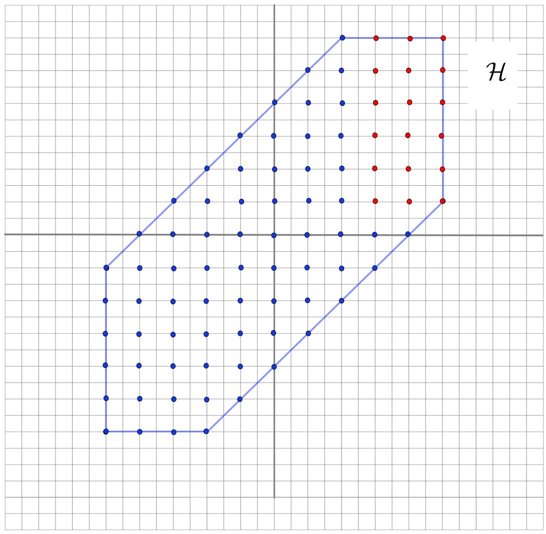


Figure 13: $\mathcal{H}$ for $P(-3,-7,-19,d,e)$

where

$$\begin{aligned}
L &= a^2b^2 + a^2c^2 + b^2c^2 + 2(a^2bc + ab^2c + abc^2),\\
M &= 2(a^2b + a^2c + ab^2 + b^2c),\\
N &= 2c^2\left(a+b+\tfrac{1}{2}\right) + 6abc + 4(ab+ac) + 3bc,\\
S &= a^2 + b^2 + bc + 2(a+b+c) + 1.
\end{aligned}$$

It will be shown that $|\mathcal{R}|^2 > |\mathcal{H}|^2$ by proving equivalently that

$$|\mathcal{R}|^2 - L - M - N > |\mathcal{H}|^2 - L - M - N.$$

Consider $|\mathcal{R}|^2$:

$$\begin{aligned}
|\mathcal{R}|^2 &= |\det(K)| = |-abcd - abce + abde + acde + bcde|\\
&= abd(e-c) + bce(d-a) + acde\\
&\geq 5a^2b^2 + 4a^2c^2 + b^2c^2 + 8a^2bc + 5ab^2c + 4abc^2 + 4a^3(b+c) + b^3(a+c)\\
&=: E_3,
\end{aligned}$$

where the inequality follows from making the substitutions $d \geq 4a+b$ and $e \geq a+b+c$. Thus

$$|\mathcal{R}|^2 - L \geq E_3 - L.$$

Next we have

$$\begin{aligned}E_3 - L &= 4a(a^2b+a^2c)+b(ab^2+b^2c)+4a^2b^2+3a^2c^2+6a^2bc+3ab^2c+2abc^2\\ &\geq 12(a^2b+a^2c)+3(ab^2+b^2c)+4a^2b^2+3a^2c^2+6a^2bc+3ab^2c+2abc^2\\ &=: E_2,\end{aligned}$$

where the inequality comes from the fact that $4a \geq 12$ since $a \geq 3$. Therefore $|\mathcal{R}|^2 - L \geq E_3 - L > E_2$, and so

$$|\mathcal{R}|^2 - L - M \geq E_3 - L - M > E_2 - M.$$

Next we have

$$\begin{aligned}E_2 - M &= 10(a^2b+a^2c)+(ab^2+b^2c)+4a^2b^2+3a^2c^2+6a^2bc+3ab^2c+2abc^2\\ &> 2c^2\big(a+b+\tfrac{1}{2}\big)+6abc+4(ab+ac)+3bc+4a^2b^2+3a^2c^2\\ &\qquad\qquad\qquad\qquad\qquad\qquad +3ab^2c+6(a^2b+a^2c)+ab^2\\ &=: E_1,\end{aligned}$$

where the inequality comes from the following four facts, obtained from the assumption that $c \geq b \geq a \geq 3$:

- $2abc^2 = 2c^2(ab) > 2c^2\big(a+b+\frac{1}{2}\big)$,
- $b^2c > 3bc$,
- $6a^2bc > 6abc$,
- $10(a^2b+a^2c) = 6(a^2b+a^2c)+4(a^2b+a^2c) > 6(a^2b+a^2c)+4(ab+ac)$.

This shows that $E_2 - M > E_1$, so it follows that

$$|\mathcal{R}|^2 - L - M - N \geq E_3 - L - M - N > E_2 - M - N > E_1 - N.$$

Now observe that

$$\begin{aligned}E_1 - N &= 4a^2b^2+3a^2c^2+3ab^2c+6(a^2b+a^2c)+ab^2\\ &= 6a^2b+ab^2+3a^2c^2+(4a^2b^2+3ab^2c+6a^2c)\\ &> a^2+b^2+bc+2(a+b+c)+1\\ &= S = |\mathcal{H}|^2 - L - M - N,\end{aligned}$$

where the inequality comes from the following six facts, again obtained from the assumption that $c \geq b \geq a \geq 3$:

$$6a^2b > a^2, \quad ab^2 > b, \quad 3ab^2c > bc, \quad 3a^2c^2 > 2a, \quad 4a^2b^2 > 2b, \quad 6a^2c > 2c+1.$$

Combining everything, one sees that

$$|\mathcal{R}|^2 - L - M - N > |\mathcal{H}|^2 - L - M - N,$$

which implies that $|\mathcal{R}|^2 > |\mathcal{H}|^2$, as desired. This completes the proof that all 0–pair odd 5–stranded pretzel knots without single-twists are not slice.

9 Proof of Theorem 2.3 with single-twists

Next, the knots with single-twists are addressed. As before, assume all knots in question are slice and therefore satisfy the signature condition, the lattice embedding condition, and both coset conditions. Just a bit of thought reveals that, possibly after mirroring, the signature condition yields only three cases to consider for 0–pair odd 5–stranded pretzel knots with single-twists. Note that $d, e \geq a$ in all cases due to embedding conditions (1), (2), (4), and (5). For $K \in P\{-a,-b,-c,d,e\}$, the cases are

(1) $a = b = c = 1$ and $d, e \geq 3$,

(2) $a = b = 1$ and $c, d, e \geq 3$,

(3) $a = 1$ and $b, c, d, e \geq 3$.

Since the lattice embedding conditions hold, there exist α, β, γ, x, y, $z \in \mathbb{Z}$ satisfying the system of equations given in Section 5. Thus, this proof for nonsliceness of 0–pair pretzel knots with single-twists has the same starting point as the previous proof for nonsliceness of 0–pair knots without single-twists. Lemma 7.1 still applies here for all three cases of 0–pair pretzel knots with single-twists. To obstruct sliceness for 0–pair knots $P(-a,-b,-c,d,e)$ with single twists, however, it is necessary to get more precise lower bounds on d and e than are obtained in Lemma 9.1.

Lemma 9.1 *If $K \in P\{-a,-b,-c,d,e\}$ is 0–pair and d is equal to its lower bound (either $d = 4a + b$ or $d = a + b + c$), then $e \geq 4a + 4b + c$.*

Proof First, suppose $d = 4a + b$ and $e = a + b + c$. By the embedding conditions, it follows that $\alpha = 2, \beta = -1$, and $\gamma = 0$, and $|x| = |y| = |z| = 1$. Embedding condition (3) says $a\alpha x + b\beta y + c\gamma z = 0$, which reduces to $\pm 2a = b$ after substitutions. But, b is odd so this is a contradiction. If instead one supposes that $e = a + b + c$, then by the embedding conditions, $|\alpha| = |\beta| = |\gamma| = |x| = |y| = |z| = 1$. After substitutions, embedding condition (3) becomes $c = \pm a \pm b$, which is again a contradiction since all three of a, b, c are odd.

Hence when $d = 4a + b$ or $d = a + b + c$, we have $e \neq a + b + c$. In words, both d and e cannot simultaneously achieve their lower bounds as given in Lemma 7.1. It follows that at least one of $|x|$, $|y|$, or $|z|$ must be ≥ 2. But, in fact, it will be shown presently that at least two of $|x|$, $|y|$, and $|z|$ must be ≥ 2. If $x = 2$, embedding condition (2) implies that $y + z = -1$; if $x = -2$, then $y + z = 3$. In both cases, it is impossible for both $|y| = 1$ and $|z| = 1$, and therefore $|y| \geq 2$ or $|z| \geq 2$. By the symmetry in x, y, z of embedding condition (2), similar results follow if $y = \pm 2$ or if $z = \pm 2$. Hence, at least two of $|x|$, $|y|$, or $|z|$ must be ≥ 2.

The choices of $|x|, |y|, |z|$ that satisfy the above discovery and that minimize e are $|x| = |y| = 2$ and $|z| = 1$, which yields $e = ax^2 + by^2 + cz^2 = 4a + 4b + c$. Thus, if d is equal to a lower bound then $e \geq 4a + 4b + c$. □

The proof of Theorem 2.3 will now proceed. The goal in each of the following cases is to arrive at a contradiction to coset condition I by showing that $|\mathcal{R}|^2 > |\mathcal{H}|^2$.

Case 1 $K \in P\{-a, -b, -c, d, e\}$ with $a = b = c = 1$.

By Lemma 7.1, $d \geq 4a + b = 5$ or $d \geq a + b + c = 3$. Assume $d = 3$. By Lemma 7.1, it follows that $e \geq 4a + 4b + c = 9$. Then

$$\begin{aligned}
|\mathcal{R}|^2 &= |\det(K)| = |-abcd - abce + abde + acde + bcde| \\
&= d(e-1) + e(d-1) + de \geq 69 \\
&> 49 = (ab + ac + bc + a + b + c + 1)^2 \\
&= |\mathcal{H}|^2,
\end{aligned}$$

as desired.

Case 2 $K \in P\{-a, -b, -c, d, e\}$ with $a = b = 1$ and $c \geq 3$.

By Lemma 7.1, $d \geq 4a + b = 5$ or $d \geq a + b + c = 2 + c$. But, $c \geq 3$ so in any case we have $d \geq 5$ and thus $e \geq 4a + 4b + c = 8 + c$ by Lemma 9.1. Then

$$\begin{aligned}
|\mathcal{R}|^2 &= |\det(K)| = |-abcd - abce + abde + acde + bcde| \\
&= d(e-c) + ce(d-1) + cde \\
&\geq 5(8 + c - c) + c(8 + c)(5 - 1) + 5c(8 + c) = 9c^2 + 72c + 40 \\
&> 9c^2 + 24c + 16 = (3c + 4)^2 \\
&= (ab + ac + bc + a + b + c + 1)^2 \\
&= |\mathcal{H}|^2,
\end{aligned}$$

as desired.

Case 3 $K \in P\{-a, -b, -c, d, e\}$ with $a = 1$ and $b, c \geq 3$.

By Lemma 7.1, $d \geq 4a + b = 4 + b$ or $d \geq a + b + c = 1 + b + c$. Assuming that $d \geq b + 4$ accounts for both situations. By Lemma 9.1, $e \geq 4a + 4b + c = 4 + 4b + c$.

Then

$$\begin{aligned}|\mathcal{R}|^2 &= |\det(K)| = |-abcd - abce + abde + acde + bcde| \\ &= bd(e-c) + bce(d-1) + cde \\ &\geq b(b+4)(4+4b+c-c) + bc(4+4b+c)(b+4-1) + c(b+4)(4+4b+c) \\ &= 4b^3c + b^2c^2 + 4b^3 + 20b^2c + 4bc^2 + 20b^2 + 32bc + 4c^2 + 16b + 16c.\end{aligned}$$

Also

$$\begin{aligned}|\mathcal{H}|^2 &= (ab + ac + bc + a + b + c + 1)^2 \\ &= b^2c^2 + 4b^2c + 4bc^2 + 4b^2 + 12bc + 4c^2 + 8b + 8c + 4.\end{aligned}$$

Let $L = b^2c^2 + 4b^2c + 4bc^2 + 4b^2 + 12bc + 4c^2 + 8b + 8c$. Then

$$|\mathcal{R}|^2 - |\mathcal{H}|^2 = 4b^3c + 16b^2c + 20bc + 8c + 4b^3 + 16b^2 + 8b - 4 > 0$$

Thus, $|\mathcal{R}|^2 > |\mathcal{H}|^2$. This concludes the proof that 0–pair odd pretzel knots with single-twists are not slice, and therefore all 0–pair odd 5–stranded pretzel knots are not slice. □

10 Proof of Theorem 2.4

Theorem 2.4 asserts that if K is a 1–pair odd 5–stranded pretzel knot without single-twists, then K is not slice. This will be shown by proving that coset condition II is violated for the knots in question. It suffices to consider only the 1–pair pretzel knots $P(a,b,c,d,e)$ for which the signature vanishes and both the lattice embedding condition and coset condition I are satisfied. Let $a,b,c,d,e > 0$ such that $a \leq b \leq c$, and assume that $K = P(-a,-b,-c,d,e)$ throughout. Let Y, $P = P_+$, W, X, and the embedding map $\alpha\colon H_2(P) \to H_2(X)/\operatorname{Tor}$ be as usual.

Theorem 6.3 gives that if K is slice, then every coset of $\operatorname{im}(\alpha)$ coming from $\operatorname{coker}(\alpha)$ has a coset representative in the set $\{\pm 1\}^m$, where $m = a+b+c$. Let v_1 and v_2 be the second and third columns (respectively) in the matrix A of α with respect to the bases chosen in Sections 5 and 6; lastly, let B be the map outlined in Section 6.

Recall from the coset conditions the sets $\mathcal{R}$ and $\mathcal{H}$ associated with A. The set $\mathcal{R}$ consists of the integer lattice points in a fundamental region $R \subset \mathbb{R}^2$ defined by $\overline{v}_1$ and $\overline{v}_2$ that correspond to unique cosets of $\operatorname{im}(\alpha)$; $\mathcal{H}$ is the set of lattice points $(x,y) \in 2\mathbb{R}^2$ such that $-a-c \leq x \leq a+c$ and $-b-c \leq y \leq b+c$, lying in a hexagonal region $H \subset \mathbb{R}^2$. The argument now reduces to determining $|\mathcal{R}|$ and $|\overline{\mathcal{H}}|$, where $\overline{\mathcal{H}} = \mathcal{H}/\langle \overline{v}_1, \overline{v}_2 \rangle$. An important note is that the computation of $|\mathcal{R}|$ will be done differently here from how it is done in Chapter 7.

Observe that $\mathcal{R}$ consists of all lattice points in the interior of R as well as all lattice points on the boundary of R modulo $\overline{v}_1$ and $\overline{v}_2$. The action of modding out the boundary of R by $\overline{v}_1$ and $\overline{v}_2$ has the effect of removing half of all boundary lattice points, plus one more. The extra lattice point that must be removed is, without loss of generality, the one at the tip of $\overline{v}_1$ that gets identified with $(0,0)$. Hence, if i is the number of interior lattice points of R and b is the total number of lattice points lying on the boundary, then

$$|\mathcal{R}| = i + \frac{b}{2} - 1.$$

In a lucky turn of events, Pick's theorem equates the right hand side of this expression with the area of R. In it's general form, Pick's theorem states that the area A of any polygon P in $\mathbb{R}^2$ with vertices at integer lattice points is given by

$$A(P) = i + \frac{b}{2} - 1,$$

where i is the number of integer lattice points in the interior of the polygon and b is the number of integer lattices points lying on the boundary of the polygon. Thus,

$$|\mathcal{R}| = i + \frac{b}{2} - 1 = A(R).$$

Given that R is a parallelogram in $\mathbb{R}^2$ defined by $\overline{v}_1, \overline{v}_2 \in \mathbb{Z}^2$, its area $A(R)$ — and thus $|\mathcal{R}|$ — is equal to the absolute value of the determinant of the 2×2 matrix whose column vectors are $\overline{v}_1$ and $\overline{v}_2$:

$$|\mathcal{R}| = A(R) = \begin{vmatrix} a\alpha - c\gamma & ax - cz \\ b\beta - c\gamma & by - cz \end{vmatrix}.$$

Also, recall from Section 8 that

$$|\mathcal{H}| = (a+b+1)(c+1) + ab = ab + ac + bc + a + b + c + 1.$$

In obstructing sliceness for 1–pair pretzel knots (still under the assumption $a \le b \le c$), three cases must be considered: (1) when the pair is $\{a, -a\}$, (2) when the pair is $\{b, -b\}$, and (3) when the pair is $\{c, -c\}$. By assumption, the twist parameters in all three cases satisfy the embedding criterion.

Case 1 $K \in P\{-a, -b, -c, a, e\}$ with $e \notin \{b, c\}$.

When the twist parameters contain the pair $\{a, -a\}$, we obtain $\alpha = 1$, $\beta = \gamma = x = 0$, and that y and z are nonzero. This yields $\overline{v}_1 = (a, 0)$ and $\overline{v}_2 = (-cz, by - cz)$, hence

$$|\mathcal{R}| = \begin{vmatrix} a & -cz \\ 0 & by - cz \end{vmatrix} = a|by - cz|.$$

As $y \to \infty$, it follows that $z \to -\infty$ by embedding condition (2) which says that $1 = x + y + z = 0 + y + z$; thus $|\mathcal{R}| \to \infty$. Similarly, as $y \to -\infty$ (and $z \to \infty$), $|\mathcal{R}| \to \infty$. For this reason, $|\mathcal{R}|$ is minimized when y and z are both small in absolute value, ie when $\overline{v}_2$ is short. Given that $b < c$, we have $\overline{v}_2$ shortest when $y = 2$ and $z = -1$. In this case

$$|\mathcal{R}| = a|2b + c| = 2ab + ac. \tag{6}$$

An upper bound for $|\overline{\mathcal{H}}|$ will now be computed. Due to the shape and dimensions of $\mathcal{H}$, we can see that many of the lattice points of $\mathcal{H}$ lie in the same $\langle \overline{v}_1, \overline{v}_2 \rangle$–coset because any two lattice points in $\mathcal{H}$ that differ by multiple of $\overline{v}_1 = (a, 0)$ will be identified. Furthermore, modding out by $\overline{v}_2$ would only result in more identification among the lattice points of $\mathcal{H}$. Hence

$$|\overline{\mathcal{H}}| = |\mathcal{H}/\langle \overline{v}_1, \overline{v}_2 \rangle \le \mathcal{H}/\langle \overline{v}_1 \rangle.$$

Note: Figures 14–19 show the lattice points in $2\mathbb{Z}^2$, that is, each grid square is 2×2.

From Figure 14, we see that each of the $b + c + 1$ rows in $\mathcal{H}$ has a distinct $\langle \overline{v}_1, \overline{v}_2 \rangle$–cosets. The result of eliminating repeated representatives from each coset to obtain $\overline{\mathcal{H}}$

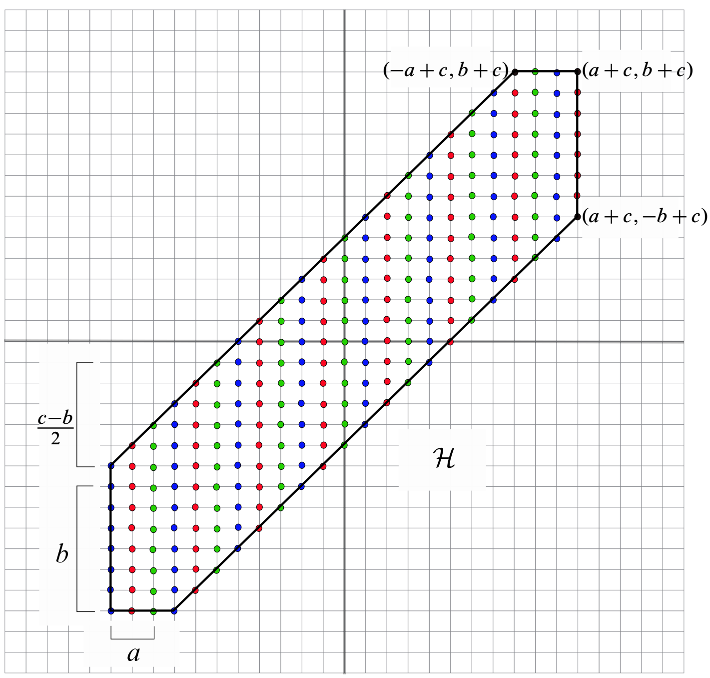


Figure 14: $\mathcal{H}$ and its $\mathcal{R}$–cosets for $P(-3, -7, -19, 3, 47)$. Points of the same color *with the same* y*–coordinate* represent the same $\mathcal{R}$–coset.

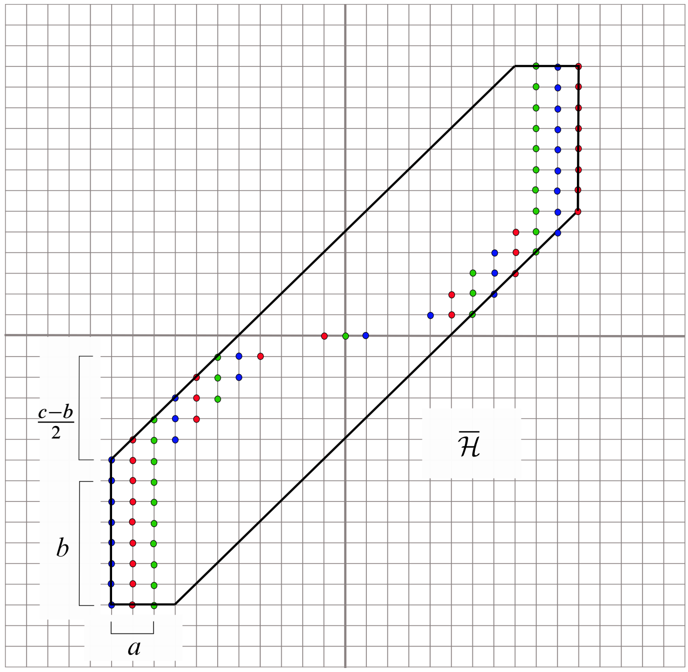


Figure 15: $\mathcal{H}$ and its $\mathcal{R}$–cosets for $P(-3,-7,-19,3,47)$, repeat representatives removed

is shown in Figure 15. Thus, an upper bound for $|\overline{\mathcal{H}}|$ is given by

$$|\overline{\mathcal{H}}| \leq \mathcal{H}/\langle \overline{v}_1 \rangle = a(b+c+1) = ab+ac+a.$$

Comparing this to (6), the desired result is achieved:

$$|\overline{\mathcal{H}}| \leq ab+ac+a < 2ab+ac = |\mathcal{R}|,$$

since $a \leq b \leq c$. Hence, an odd 5–stranded pretzel knot $K \in P\{-a,-b,-c,a,d\}$, with $a,b,c,d \geq 3$, is not slice by coset condition I.

Case 2 $K \in P\{-a,-b,-c,b,d\}$ with $e \notin \{a,c\}$.

When the twist parameters contain the pair $\{b,-b\}$, we obtain $\beta = 1$, $\alpha = \gamma = y = 0$, and that x and z are nonzero. With this, $\overline{v}_1 = (0,b)$ and $\overline{v}_2 = (-cz, by-cz)$, hence

$$|\mathcal{R}| = \begin{vmatrix} 0 & ax-cz \\ b & -cz \end{vmatrix} = b|ax-cz|.$$

Following the logic of Case 1, it suffices to show that $|\mathcal{R}| > |\overline{\mathcal{H}}|$ when the length of $\overline{v}_2$

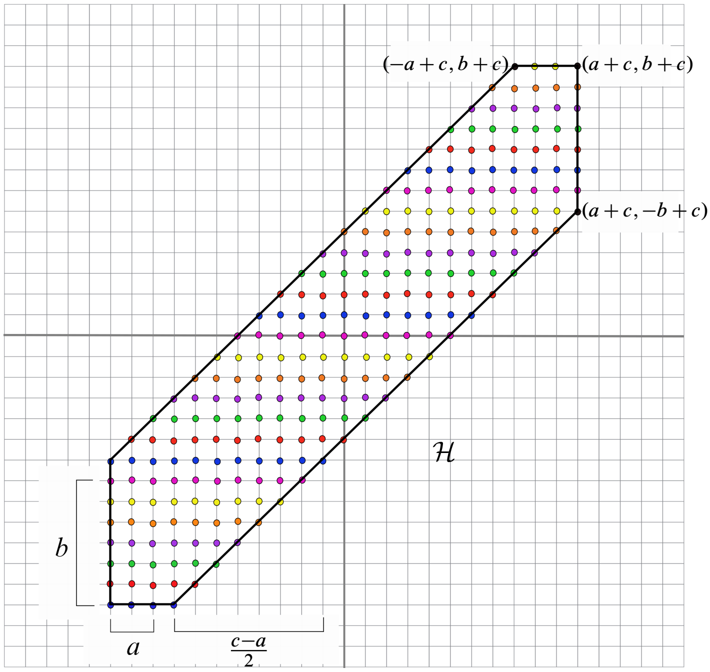


Figure 16: $\mathcal{H}$ and its $\mathcal{R}$–cosets for $P(-3,-7,-19,7,31)$. Points of the same color *with the same x–coordinate* represent the same $\mathcal{R}$–coset.

is minimized. Since $a<c$, we have $\overline{v}_2$ shortest when $x=2$ and $z=-1$, so

$$|\mathcal{R}| = b\,|2a+c| = 2ab+bc. \tag{7}$$

The upper bound for $|\overline{\mathcal{H}}|$ is computed for Case 2 in a similar manner as for Case 1, the only difference being that $\overline{v}_1=(0,b)$, and thus lattice points in $\mathcal{H}/\langle\overline{v}_1\rangle$ are in the same coset when they differ by multiple of $(0,b)$ (vertical translations), as seen in Figure 16. Each of the $a+c+1$ columns in $\mathcal{H}$ always has b distinct $\mathcal{R}$–cosets. The result of eliminating repeated representatives from each coset to obtain $\overline{\mathcal{H}}$ is shown in Figure 17. Thus, an upper bound for $|\overline{\mathcal{H}}|$ is given by

$$|\overline{\mathcal{H}}| \le \mathcal{H}/\langle\overline{v}_1\rangle = b(a+c+1) = ab+bc+b.$$

Comparing this to (7), again the desired result is achieved:

$$|\overline{\mathcal{H}}| \le ab+bc+b < 2ab+bc = |\mathcal{R}|,$$

since $b\ge a\ge 3$. Hence, 5–stranded pretzel knots $K\in P\{-a,-b,-c,b,d)\}$, with $a,b,c,d\ge 3$, are not slice.

Case 3 $K\in P\{-a,-b,-c,c,d\}$ with $e\notin\{a,b\}$.

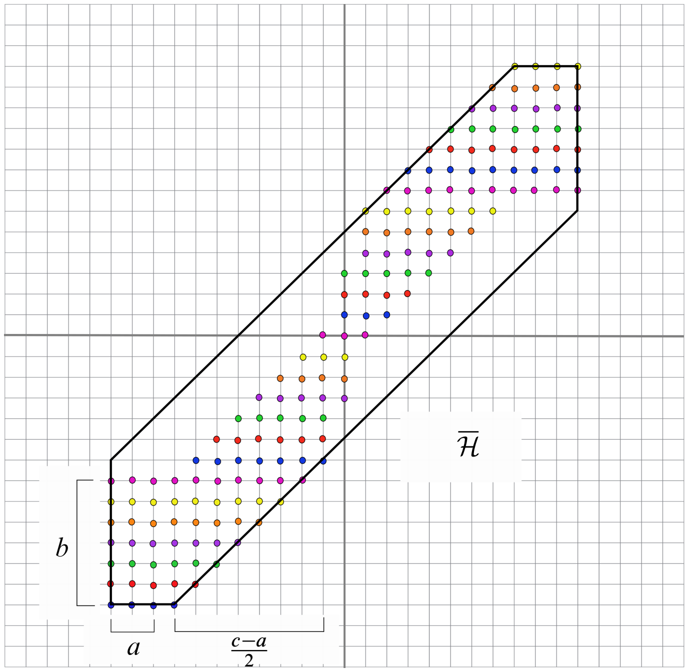


Figure 17: $\mathcal{H}$ and its $\mathcal{R}$–cosets for $P(-3,-7,19,7,31)$, repeat representatives removed.

The final case of 5–stranded odd 1–pair pretzel knots has $\{c,-c\}$ as the pair in the twist parameters. Unlike for the 1–pair cases where the pair of canceling twist parameters is $\{a,-a\}$ or $\{b,-b\}$, the case with $\{c,-c\}$ does not necessarily imply that $\gamma=1$, $\alpha=\beta=z=0$, with x and y nonzero. Since $c\geq b\geq a$, it is possible that both α and β are nonzero and embedding condition (4) is satisfied by $c=a\alpha^2+b\beta^2$. In this case, however, the proof of Lemma 9.1 shows we would have $c\geq 4a+b$ and $e=ax^2+by^2+cz^2\geq a+b+c$, which implies that $P(-a,-b,-c,c,d)$ is not slice by the proof of Theorem 2.3.

Hence, the only case that need be considered is the case in which $\gamma=1$, $\alpha=\beta=z=0$, with x and y nonzero. Under these conditions, $\overline{v}_1=(-c,-c)$ and $\overline{v}_2=(ax,by)$, and therefore

$$|\mathcal{R}|=\begin{vmatrix}-c & ax\\ -c & by\end{vmatrix}=c\,|ax-by|.$$

Again by following the logic from Case 1 and Case 2, it suffices to show that $|\mathcal{R}|>|\overline{\mathcal{H}}|$ when $\overline{v}_2$ is at its shortest. Since $a<b$, the length of $\overline{v}_2$ is minimized when $x=2$ and

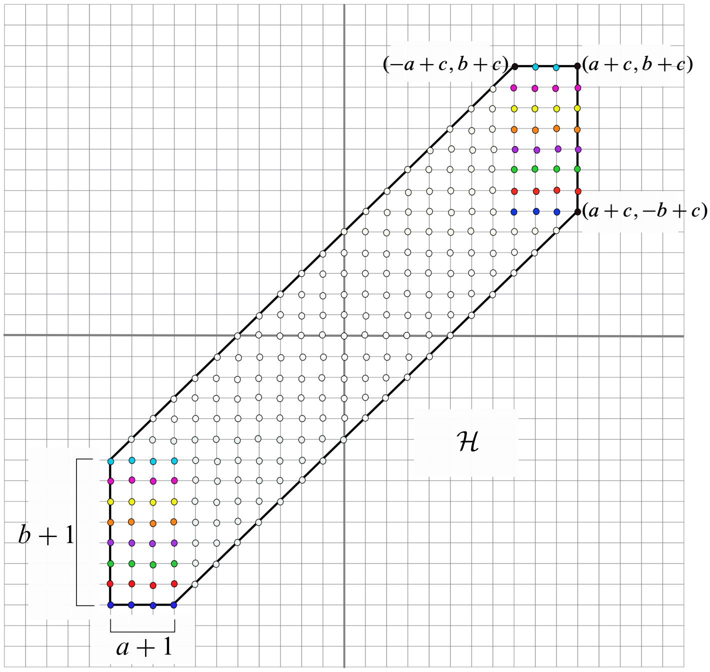


Figure 18: $\mathcal{H}$ and its $\mathcal{R}$–cosets for $P(-3,-7,-19,19,55)$. Each white point represents a distinct $\mathcal{R}$–coset; colored points *lying along the same 45–degree diagonal* represent the same $\mathcal{R}$–coset.

$y=-1$. In this case,

$$|\mathcal{R}| = c\,|2a+b| = 2ac+bc. \tag{8}$$

The computation of an upper bound for $|\overline{\mathcal{H}}|$ in Case 3 is similar to those in Cases 1 and 2. Namely, it is computed by identifying lattice points in $\mathcal{H}$ via multiples of $\overline{v}_1=(-c,-c)$ (45–degree diagonal translations). The computations are also done as before using the well-understood region $\mathcal{H}$, however it is more efficient now to subtract off the number of repeat $\langle \overline{v}_1\rangle$–coset representatives from $|\mathcal{H}|$, rather than count the cosets directly as in Cases 1 and 2. Figure 18 indicates that

$$\begin{aligned} |\overline{\mathcal{H}}| &\le |\mathcal{H}| - (a+1)(b+1) \\ &= ab+ac+bc+a+b+c+1-(ab+a+b+1) \\ &= ac+bc+c. \end{aligned}$$

Since $c \ge a \ge 3$, comparing this result with (8) gives the result

$$|\overline{\mathcal{H}}| \le ac+bc+c < 2ac+bc = |\mathcal{R}|.$$

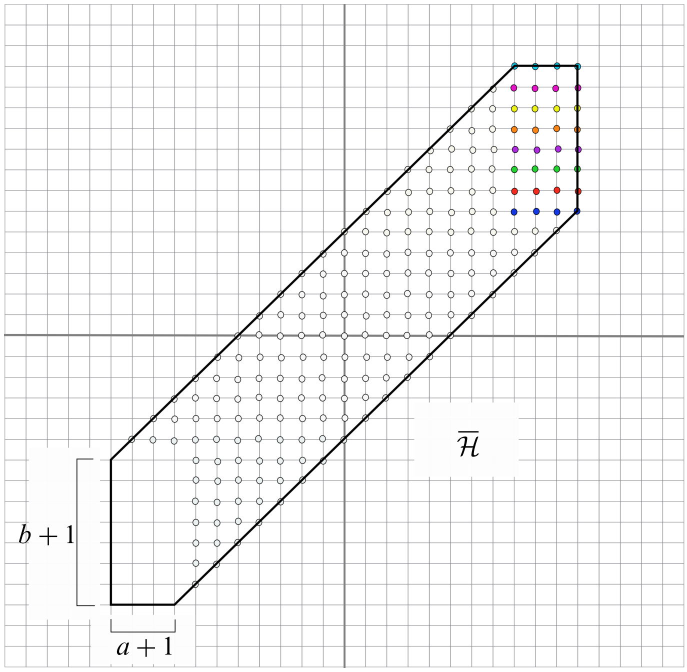


Figure 19: $\mathcal{H}$ and its $\mathcal{R}$–cosets for $P(-3,-7,19,19,55)$, repeat representatives removed.

Refer to Figure 19 for $\overline{\mathcal{H}}$ in this case. Thus, 5–stranded odd pretzel knots of the form $P(-a,-b,-c,c,d)$, with $a,b,c,d \geq 3$, are not slice. The proof is complete. □

References

[1] **S Akbulut**, **R Kirby**, *Branched covers of surfaces in* 4*–manifolds*, Math. Ann. 252 (1979/80) 111–131 MR

[2] **R E Bedient**, *Double branched covers and pretzel knots*, Pacific J. Math. 112 (1984) 265–272 MR

[3] **J Greene**, **S Jabuka**, *The slice-ribbon conjecture for* 3*–stranded pretzel knots*, Amer. J. Math. 133 (2011) 555–580 MR

[4] **C Herald**, **P Kirk**, **C Livingston**, *Metabelian representations, twisted Alexander polynomials, knot slicing, and mutation*, Math. Z. 265 (2010) 925–949 MR

[5] **S Jabuka**, *Rational Witt classes of pretzel knots*, Osaka J. Math. 47 (2010) 977–1027 MR

[6] **L H Kauffman**, *On knots*, Annals of Mathematics Studies 115, Princeton Univ. Press, Princeton, NJ (1987) MR

[7] **L H Kauffman**, **L R Taylor**, *Signature of links*, Trans. Amer. Math. Soc. 216 (1976) 351–365 MR

[8] **C Kearton**, *Mutation of knots*, Proc. Amer. Math. Soc. 105 (1989) 206–208 MR

[9] **R C Kirby**, *The topology of* 4*–manifolds*, Lecture Notes in Mathematics 1374, Springer (1989) MR

[10] **A G Lecuona**, *On the slice-ribbon conjecture for Montesinos knots*, Trans. Amer. Math. Soc. 364 (2012) 233–285 MR

[11] **P Lisca**, *Lens spaces, rational balls and the ribbon conjecture*, Geom. Topol. 11 (2007) 429–472 MR

[12] **C Livingston**, *A survey of classical knot concordance*, from "Handbook of knot theory" (W Menasco, M Thistlethwaite, editors), Elsevier B. V., Amsterdam (2005) 319–347 MR

[13] **L Long**, *Slice ribbon conjecture, pretzel knots and mutation*, PhD thesis, University of Texas at Austin (2014) Available at `http://hdl.handle.net/2152/27145`

[14] **A N Miller**, *Distinguishing mutant pretzel knots in concordance*, J. Knot Theory Ramifications 26 (2017) art. id. 1750041, 24 pp. MR

[15] **W D Neumann**, **F Raymond**, *Seifert manifolds, plumbing, μ–invariant and orientation reversing maps*, from "Algebraic and geometric topology" (K C Millett, editor), Lecture Notes in Math. 664, Springer (1978) 163–196 MR

[16] **P Ozsváth**, **Z Szabó**, *Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary*, Adv. Math. 173 (2003) 179–261 MR

[17] **P Ozsváth**, **Z Szabó**, *On the Floer homology of plumbed three-manifolds*, Geom. Topol. 7 (2003) 185–224 MR

[18] **D Rolfsen**, *Knots and links*, Mathematics Lecture Series 7, Publish or Perish, Berkeley, CA (1976) MR

[19] **V Turaev**, *Torsion invariants of* Spin^c*–structures on* 3*–manifolds*, Math. Res. Lett. 4 (1997) 679–695 MR

[20] **H Zieschang**, *Classification of Montesinos knots*, from "Topology" (L D Faddeev, A A Mal'cev, editors), Lecture Notes in Math. 1060, Springer (1984) 378–389 MR

Department of Mathematics and Computer Science, Colorado College
Colorado Springs, CO, United States

`kathryn.bryant@coloradocollege.edu`

Received: 21 September 2016 Revised: 7 February 2017

Algebraic & Geometric Topology 17 (2017) 3665–3701

Axioms for higher twisted torsion invariants of smooth bundles

CHRISTOPHER OHRT

This paper attempts to investigate the space of various characteristic classes for smooth manifold bundles with local system on the total space inducing a finite holonomy covering. These classes are known as twisted higher torsion classes. We will give a system of axioms that we require these cohomology classes to satisfy. Higher Franz–Reidemeister torsion and twisted versions of the higher Miller–Morita–Mumford classes will satisfy these axioms. We will show that the space of twisted torsion invariants is two-dimensional or one-dimensional depending on the torsion degree and is spanned by these two classes. The proof will greatly depend on results about the equivariant Hatcher constructions developed in Goodwillie, Igusa and Ohrt (2015).

19J10, 55R40; 57R80, 55R10

1 Introduction

Higher torsion invariants have been developed by J Wagoner, J R Klein, K Igusa, M Bismut, J Lott, W Dwyer, M Weiss, E B Williams, S Goette and many others; see Wagoner [12], Igusa [5], Igusa and Klein [9], Bismut and Lott [2], Dwyer, Weiss and Williams [3] and Bismut and Goette [1].

In [7], Igusa defined a higher torsion invariant of degree $2k$ to be a characteristic class $\tau(E) \in H^{4k}(B;\mathbb{R})$ of a smooth bundle $E \to B$ satisfying an additivity and a transfer axiom; see [7, Section 2]. He proved that the set of higher torsion invariants forms a two-dimensional vector space spanned by the higher Reidemeister torsion and the Miller–Morita–Mumford class.

But higher Reidemeister torsion or Igusa–Klein torsion can be defined in a more general way: it is a characteristic class $\tau^{\mathrm{IK}}(E,\rho) \in H^{2k}(B;\mathbb{R})$ for a smooth bundle with an unitary representation $\rho\colon \pi_1 E \to U(m)$ factorizing through a finite group; see for example Igusa [5]. For our purposes it will be better to look at finite complex local systems on E instead. After a choice of a base point, this corresponds to a representation of the fundamental group as can be found for example in T Szamuely's book [11, Corollary 2.6.2]. Regarding that, we will define a twisted higher torsion invariant in degree k to be a characteristic class $\tau(E;\mathcal{F}) \in H^{2k}(B;\mathbb{R})$ depending on

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3665

a finite local complex system $\mathcal{F}$ on E inducing a finite holonomy covering satisfying six axioms: the first two are versions of the original two axioms for nontwisted torsion invariants, which will respect the local system; the remaining four axioms will determine the dependence of the torsion class on the local system.

The goal of this paper is to show an analogous result to Igusa's on twisted torsion invariants. For this we will generalize Igusa's paper [7] step by step:

In Section 2, we will define twisted higher torsion invariants.

In Section 3, we will repeat why the Igusa–Klein torsion τ^{IK} satisfies the axioms, introduce a twisted version of the Miller–Morita–Mumford classes M^{2k} and show that these also satisfy the axioms. The MMM classes will be zero in degree $4l+2$. Then we will state our main theorem:

Theorem 1.1 (main theorem) *The space of higher twisted torsion invariants in degree $4l$ on bundles with simple fibers and base having a finite fundamental group is two-dimensional and spanned by the twisted MMM class and the twisted Igusa–Klein torsion, and one-dimensional in degree $4l+2$ and spanned by the Igusa–Klein torsion. In other words, for any twisted torsion invariant of even degree τ, there exist unique $a, b \in \mathbb{R}$ such that*

$$\tau = a\tau^{\mathrm{IK}} + bM,$$

and for every twisted torsion invariant τ of odd degree there exists a unique $a \in \mathbb{R}$ such that

$$\tau = a\tau^{\mathrm{IK}}.$$

The scalars a and b can be calculated as follows: For torsion in degree $4l$ we look at the universal line bundle $\lambda\colon ES^1 \to \mathbb{CP}^\infty$. Since the cohomology groups $H^{2k}(\mathbb{CP}^\infty;\mathbb{R})$ are one-dimensional, the torsion invariant of the associated S^1–bundle $S^1(\lambda)$ and the associated S^2–bundle $S^2(\lambda)$ over $\mathbb{CP}^\infty$ will determine the scalars a and b. In degree $4l+2$ we only have to calculate a by looking at a fiberwise quotient $S^1(\lambda)/(\mathbb{Z}/n)$ of the n–action on S^1. This admits a nontrivial finite complex local system and therefore has a nontrivial higher twisted torsion.

Before we prove the main theorem, we will extend a higher twisted torsion invariant to have values on bundles with vertical boundaries and then define a relative torsion for bundle pairs (see Section 4), which we will use to deconstruct any bundle into easier pieces and keep control over the torsion.

In Section 5, we will show that the main theorem holds on S^1–bundles. Then we will define the difference torsion to be

$$\tau^\delta := \tau - a\tau^{\mathrm{IK}} - bM,$$

and we will see that $\tau^\delta = 0$ for every sphere bundle, disc bundle and odd-dimensional lens space bundle. In Goodwillie, Igusa and Ohrt [4] we give an explicit base for the space of h–cobordism bundles of a lens space, and the calculations in Section 6 of this paper show that the difference torsion will be zero on these basis elements. From this crucial observation we can deduce that the difference torsion will be a fiber homotopy invariant, and in Section 7 we will show that this fiber homotopy invariant must be trivial if it is restricted to bundles with simple fiber and base having finite fundamental group.

Acknowledgements This paper is the product of my work with Kiyoshi Igusa during my stay at Brandeis University in the academic year 2011/2012, which led up to further work on the equivariant Hatcher constructions with Thomas Goodwillie and Kiyoshi Igusa in the academic year 2012/2013. I want to thank Kiyoshi for the great support, advice, and guidance he offered me. I also want to thank Ulrich Bunke from my home university in Regensburg, Germany, for the many comments and corrections he contributed.

2 Axioms and definitions

2.1 Preliminaries

Throughout the whole paper, let $F \hookrightarrow E \xrightarrow{p} B$ be a smooth fiber bundle, where E and B are compact smooth manifolds, p is a smooth submersion, and F is a compact orientable n–dimensional manifold with or without boundary. In the boundary case, there is a subbundle $\partial F \to \partial^v E \to B$ of E. We call $\partial^v E$ the vertical boundary of E. We assume that B is connected and that the action of $\pi_1 B$ on F preserves the orientation of F. We also assume that $\pi_1 B$ is finite, which immediately implies that the bundle E is unipotent (as required in [7]).

These are all similar assumptions to the ones for considering nontwisted higher torsion classes. Additionally to those, we assume that E comes equipped with a finite complex local system $\mathcal{F}$. By "finite" we mean that there exists a finite covering $\widetilde{E} \to E$ such that the pull-back of the local system is trivializable. These local systems are sometimes also called hermitian local coefficient systems because they induce a well defined hermitian inner product on each fiber. We will often call $\mathcal{F}$ just local coefficient system.

If $F \hookrightarrow E \to B$ is a smooth bundle we have the transfer map

$$\mathrm{tr}_B^E \colon H^*(E;\mathbb{Z}) \to H^*(B;\mathbb{Z}).$$

For an exact definition, one can consult [7, Section 2]. The most important property we will need is that we always have

$$\mathrm{tr}_E^B = (-1)^n \,\mathrm{tr}_B^E,$$

where n is the dimension of the fiber F. In particular, this implies $\mathrm{tr}_B^E = 0$ if $\dim F$ is odd and we only consider it on cohomology with real coefficients.

2.2 Higher twisted torsion invariants

We are ready to give the definition of a twisted higher torsion invariant. Most of the axioms were proposed by Igusa in [8, Section 4].

Definition 2.1 A higher twisted torsion invariant in degree $2k$ with $k \in \mathbb{N}$ is a rule τ_k, which assigns to any bundle $F \hookrightarrow E \to B$ with closed fiber F and local coefficient system $\mathcal{F}$ on E a cohomology class $\tau_k(E, \mathcal{F}) \in H^{2k}(B; \mathbb{R})$ subject to the axioms beneath. We will drop the degree out of the notation most of the time and just write τ.

Remark 2.2 We consider higher twisted torsion invariants as real cohomology classes (rather than rational ones) since our main example is Igusa–Klein torsion which can only be defined with real coefficients.

Axiom 1 (naturality) τ_k is a characteristic class in degree $2k$. That means for a map $f\colon B' \to B$ and a bundle $F \hookrightarrow E \to B$ with local coefficient system $\mathcal{F}$ on E we have

$$\tau_k(f^*(E), f^*\mathcal{F}) = f^*\tau(E, \mathcal{F}) \in H^{2k}(B'; \mathbb{R}),$$

where f^* denotes the pull-back along f.

Remark 2.3 The naturality axiom immediately implies triviality on trivial bundles $\tau_k(B \times F, \mathcal{F}) = 0$, if $\mathcal{F} = \mathbf{1}$ is the constant local system. Furthermore, if B is simply connected, a local system $\mathcal{F}$ on $B \times F$ will pull back from a local system $\mathcal{F}_F$ on F under the projection $B \times F \to F$. So if we view F as a trivial bundle over a point, naturality gives that $\tau(B \times F, \mathcal{F}) = 0$ for any local system $\mathcal{F}$ if B is simply connected.

If B is a space with finite fundamental group and $B \times F \to B$ is a trivial bundle with local system $\mathcal{F}$, we can look at the pull-back $\widetilde{B} \times F \to \widetilde{B}$ of $B \times F$ to the universal covering space $\pi\colon \widetilde{B} \to B$. By the previous paragraph we know that the twisted torsion of $\widetilde{B} \times F$ is trivial with respect to any finite local system and since the π is a finite covering the map $\pi^*\colon H^*(B; \mathbb{R}) \to H^*(B; \mathbb{R})$ is a monomorphism. By naturality we see that the torsion of a trivial bundle over a base with finite fundamental group is 0 with respect to any local system.

Let E_1 and E_2 be bundles over B with local coefficient systems $\mathcal{F}_1$ and $\mathcal{F}_2$, such that there is an isomorphism $\phi\colon \partial^v E_1 \to \partial^v E_2 \neq \varnothing$ and such that we have, for the restrictions of the local systems,

$$(\mathcal{F}_1)_{|\partial^v E_1} \cong \phi^*(\mathcal{F}_2)_{|\partial^v E}.$$

Then we can glue them together to a local coefficient system $\mathcal{F} := \mathcal{F}_1 \cup_\phi \mathcal{F}_2$ on $E_1 \cup_\phi E_2$.

Axiom 2 (geometric additivity) In the setting from above we have, for any twisted torsion invariant τ,

$$\tau(E_1 \cup_\phi E_2, \mathcal{F}) = \tfrac{1}{2}\big(\tau(DE_1, \mathcal{F}_1^l \cup_{\mathrm{id}} \mathcal{F}_1^r) + \tau(DE_2, \mathcal{F}_2^l \cup_{\mathrm{id}} \mathcal{F}_2^r)\big),$$

where DE_i denotes the fiberwise double $E_i^l \cup_{\mathrm{id}} E_i^r$ with a left copy E_i^l and a right copy E_i^r glued together along their isomorphic boundaries and the induced local coefficient system $\mathcal{F}_i^l \cup_{\mathrm{id}} \mathcal{F}_i^r$.

Now suppose again that $p\colon E \to B$ is a bundle with closed fiber F and local coefficient system $\mathcal{F}$ on E. Let $q\colon D \to E$ be an S^n–bundle which is isomorphic to the sphere bundle of a vector bundle. We get the local coefficient system $q^*\mathcal{F}$ on D by pulling back $\mathcal{F}$ along q.

Axiom 3 (geometric transfer) In the situation above, for a twisted torsion invariant τ, we have the following relation between the torsion class $\tau_B(D, q^*\mathcal{F}) \in H^{2k}(B;\mathbb{R})$ of D as a bundle over B and the torsion class $\tau_E(D, q^*\mathcal{F}) \in H^{2k}(E;\mathbb{R})$ of D as a bundle over E:

$$\tau_B(D, q^*\mathcal{F}) = \chi(S^n)\tau_B(E, \mathcal{F}) + \mathrm{tr}_B^E(\tau_E(D, q^*\mathcal{F})),$$

where χ denotes the Euler class, $\mathrm{tr}_B^E\colon H^{2k}(E;\mathbb{R}) \to H^{2k}(B;\mathbb{R})$ the transfer, and $\tau_E(D, q^*\mathcal{F})$ the twisted torsion class of D over E.

Remark 2.4 We have $\chi(S^n) = 2$ or 0 depending on whether n is even or odd.

Remark 2.5 If we take a twisted torsion class τ_{2k} with $k = 2l$ even, we will get a nontwisted torsion class in the sense of Igusa [7],

$$\tau_{\mathrm{nontw}}(E) := \tau(E, \mathbf{1}) \in H^{4k}(B;\mathbb{R}),$$

where $E \to B$ is a bundle and $\mathbf{1}$ the constant local system on E. We will denote this nontwisted torsion invariant simply by $\tau(E)$ without any local system in the argument.

Since according to Igusa's definition there are no higher torsion invariants in degree $4l + 2 = 2k$, we also need the following axiom:

Axiom 4 (triviality) For a twisted torsion invariant in degree $4l + 2$, we have, for every bundle $E \to B$ and the constant local system $\mathbf{1}$ on E,

$$\tau(E, \mathbf{1}) = 0 \in H^{4l+2}(B;\mathbb{R}).$$

These axioms so far were only modifications of the axioms for nontwisted torsion invariants. We also need some axioms concerning the local system $\mathcal{F}$ on E:

Axiom 5 (additivity for coefficients) If $\mathcal{F} = \bigoplus_i \mathcal{F}_i$ for local systems $\mathcal{F}_i$ on E, with $E \to B$ a bundle, then we have, for every twisted torsion invariant τ,

$$\tau(E, \mathcal{F}) = \sum_i \tau(E, \mathcal{F}_i).$$

Axiom 6 (transfer/induction for coefficients) If $\widetilde{E} \to B$ and $E \to B$ are bundles and $\pi\colon \widetilde{E} \to E$ is a finite fiberwise covering, then we have, for every local system $\mathcal{F}$ on $\widetilde{E}$,

$$\tau(\widetilde{E}, \mathcal{F}) = \tau(E, \pi_* \mathcal{F}),$$

where π_* denotes the push-down operator for local systems.

Remark 2.6 Igusa [8, Section 4.7] proposed this axiom originally in the following form, which corresponds to our formulation:

If G is a group that acts freely and fiberwise on $E \to B$, H is a subgroup of G, and V is a unitary representation of H, then the torsions of the orbit bundles E/G, $E/H \to B$ are related by

$$\tau(E/G, \operatorname{Ind}_H^G V) = \tau(E/H, V).$$

Lastly we need a continuity axiom. It roughly states that if we fix a bundle $E \to B$ then the values of a twisted torsion invariant on E depend continuously on the different local systems $\mathcal{F}$ we might choose. More explicitly we can look at the universal linear S^1–bundle $S^\infty \to \mathbb{CP}^\infty$. If we identify the quotient $\mathbb{Q}/\mathbb{Z}$ with the roots of unity in $\mathbb{C}$ we get a local system $\mathcal{F}_\zeta$ on $S^\infty/(\mathbb{Z}/n)$ for every $\zeta \in \mathbb{Q}/\mathbb{Z}$ of degree n. We can use this and a fixed torsion invariant τ to define a map

$$f_\tau\colon \mathbb{Q}/\mathbb{Z} \to H^{2k}(\mathbb{CP}^\infty, \mathbb{R}) \cong \mathbb{R}$$

given by $f_\tau(\zeta) := \tau(S^\infty/(\mathbb{Z}/n), \mathcal{F}_\zeta)$. Details will provided in Section 5 where we need to use the following axiom:

Axiom 7 (continuity) The map $f_\tau\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{R}$ is continuous.

3 Statement of main theorem

3.1 Examples of higher twisted torsion invariants

Our main example of higher twisted torsion is the higher Franz–Reidemeister torsion or Igusa–Klein torsion

$$\tau_k^{\mathrm{IK}}(E, \partial_0 E, \mathcal{F}) \in H^{2k}(B; \mathbb{R}),$$

which is defined for any unipotent bundle pair $(F, \partial_0 F) \to (E, \partial_0 E) \to B$ with $\partial_0 E \subseteq \partial^v E$ and local system $\mathcal{F}$ on E; for details, see [5].

Igusa proved the following result:

Theorem 3.1 [7, Theorem 9.4; 5, Theorem 2.4.7 and Theorem 2.7.1] *The Igusa–Klein torsion invariants are higher twisted torsion invariants for bundles with closed fibers.*

Besides this torsion, we also have the Miller–Morita–Mumford classes in degree $4l$ with $l \in \mathbb{N}$

$$M^{2l}(E) := \mathrm{tr}_B^E((2l!)\,\mathrm{ch}_{4l}(T^v E)),$$

where $\mathrm{ch}_{4l}(T^v E) = \frac{1}{2}\,\mathrm{ch}_{4l}(T^v E \otimes \mathbb{C})$. We will consider this to be a real characteristic class. Igusa also showed that this class is a higher nontwisted torsion invariant [7, Proposition 9.1]. To make it a higher twisted torsion invariant we simply define, for an $m-$dimensional local system $\mathcal{F}$ on E,

$$M^{2l}(E, \mathcal{F}) := mM^{2l}(E) \in H^{4l}(B; \mathbb{R}).$$

Furthermore we set

$$M^{2l+1}(E, \mathcal{F}) := 0,$$

since there is no nontwisted torsion in degree $2k = 2(2l+1)$, and the twisted MMM torsion always induces nontrivial nontwisted torsion. Knowing that the MMM class is a nontwisted torsion invariant (and therefore fulfills the first three axioms) it is now easy to see:

Theorem 3.2 *The twisted MMM class is a higher twisted torsion invariant.*

We also know that for any bundle $F \to E \to B$ with closed l–dimensional fiber F, twice the transfer map tr_B^E is rationally trivial, if l is odd. Therefore we get:

Proposition 3.3 $M^k(E, \mathcal{F}) = 0$ *for closed odd-dimensional fiber* F.

3.2 The space of twisted torsion invariants

We are moving on to the space of higher twisted torsion invariants in degree $2k$. We begin with an elementary observation:

Lemma 3.4 *For each* k*, the set of all twisted torsion invariants of degree* $2k$ *is a vector space over* $\mathbb{R}$.

Of course, the same statement holds for the set of nontwisted higher torsion invariants. Igusa proved for the space of nontwisted higher torsion invariants:

Theorem 3.5 [7, Theorem 4.4] *For any l the space of higher nontwisted torsion invariants in degree $4l$ is two-dimensional and spanned by the nontwisted MMM class M^{4l} and the nontwisted Igusa–Klein torsion τ_{2l}^{IK}. In other words, for any nontwisted torsion invariant τ there exist unique $a, b \in \mathbb{R}$ such that*

$$\tau = a\tau^{\mathrm{IK}} + bM.$$

Now, let $\mathrm{Top}_{\mathrm{fin}}$ be the full subcategory of Top of topological spaces with finite fundamental group and $\mathrm{Top}_{\mathrm{sim}}$ the full subcategory of simple topological spaces. A space F is called simple if the fundamental group $\pi_1 F$ acts trivially on the higher homotopy groups $\pi_* F$. If we restrict a twisted torsion invariant to bundles with fibers in $\mathrm{Top}_{\mathrm{sim}}$ and base in $\mathrm{Top}_{\mathrm{fin}}$ we get the main theorem:

Theorem 3.6 (main theorem) *In the setting above, the space of higher twisted torsion invariants in degree $2k$ on bundles with simple fibers and base having finite fundamental group is two-dimensional and spanned by the twisted MMM class and the twisted Igusa–Klein torsion, if k is even, and one-dimensional and spanned by the Igusa–Klein torsion, if k is odd. In other words, for any twisted torsion invariant τ of degree $4l$, there exist unique $a, b \in \mathbb{R}$ such that*

$$\tau = a\tau^{\mathrm{IK}} + bM,$$

and for every twisted torsion invariant τ of odd degree $4l + 2$ there exists a unique $a \in \mathbb{R}$ such that

$$\tau = a\tau^{\mathrm{IK}}.$$

Remark 3.7 If k is even, we get a nontwisted torsion invariant from the twisted one by always inserting the trivial representation. Then the numbers a and b used in both theorems above will be the same.

The proof of the main theorem is developed in Sections 4 to 7. In the very technical Section 4 we will introduce relative torsion of bundles with vertical boundary and we will turn the geometric additivity axiom into two eye-pleasing formulas that will allow us to dissect the fiber F into easier pieces meeting along a common vertical boundary. Section 5 is devoted to investigating the higher twisted torsion of linear S^1–bundles. Concretely, we show that the continuity, geometric additivity, and geometric transfer axioms together imply that the space of twisted torsion invariants restricted to only the universal bundle $S^1 \to S^\infty \to \mathbb{CP}^\infty$ is one-dimensional. This together with the results in the untwisted case implies that the difference torsion $\tau^\delta := \tau - a\tau^{\mathrm{IK}} - bM$ is trivial on all linear disc and sphere bundles. The goal of Section 6 is to use this to show that τ^δ is a fiber homotopy invariant which will follow from τ^δ being trivial on any lens space bundle. The proof of this last assertion relies on the twisted Hatcher example we defined in [4]. Armed with the fiber homotopy invariance we then proceed in Section 7 to use homotopical tools to replace any fiber bundle $E \to B$ with another one that has homologically trivial fibers and the same difference torsion as E and prove triviality on those.

3.3 The scalars *a* and *b*

Before we get to the proof of the main theorem let us assume for now that it is true. This section aims to explain how given a torsion invariant one can calculate the scalars of the equation $\tau = a\tau^{\mathrm{IK}} + bM$. We need to distinguish between τ having degree $2k = 4l$ or $2k = 4l + 2$.

3.3.1 In degree $2k = 4l$ First we first look at a twisted torsion invariant in degree $2k = 4l$. In this case the scalars must be the same as the ones we get for the corresponding nontwisted torsion. To determine them we follow Igusa's approach [7, Section 4.2] and look at the universal $S^1 \cong U(1) \cong \mathrm{SO}(2)$–bundle λ over $\mathbb{CP}^\infty \cong \mathrm{BU}(1)$. Furthermore, let $S^1(\lambda)$ be the associated circle bundle with λ and $S^2(\lambda)$ the S^2–bundle associated with $S^1(\lambda)$ (by fiberwise suspension of $S^1(\lambda)$). Since the cohomology ring of $\mathbb{CP}^\infty$ is a polynomial algebra generated by $c_1(\lambda)$, the cohomology group $H^{2k}(\mathbb{CP}^\infty; \mathbb{R}) \cong \mathbb{R}$ is generated by $\mathrm{ch}_{2k}(\lambda) = c_1^k/k!$.

From this, we immediately get scalars $s_1, s_2 \in \mathbb{R}$ for any twisted torsion invariant in degree $2k = 4l$ with

$$\tau(S^1(\lambda)) = s_1 \,\mathrm{ch}_{2k}(\lambda) \quad \text{and} \quad \tau(S^2(\lambda)) = s_2 \,\mathrm{ch}_{2k}(\lambda).$$

Furthermore we have the following two propositions:

Proposition 3.8 [5, Chapter 2.7] *We get*

$$\tau_{2l}^{\mathrm{IK}}(S^n(\lambda)) = (-1)^{l+n}\zeta(2l+1)\,\mathrm{ch}_{4l}(\lambda).$$

Proposition 3.9 [7, Proposition 9.2] $M_k(S^2(\lambda)) = 2k!\,\mathrm{ch}_{2k}(\lambda)!.$

Now we are taking into account that the MMM class is trivial on odd-dimensional fibers, and therefore we get that $\tau(S^1(\lambda)) = a\tau^{\mathrm{IK}}(S^1(\lambda))$. From this we get

$$a = s_1/((-1)^{1+l}\zeta(2l+1)).$$

Looking at the $S^2(\lambda)$ case, we have

$$s_2 = a(-1)^l\zeta(2l+1) + b2k! = -s_1 + b2k!$$

and therefore

$$b = \frac{s_1 + s_2}{2k!}.$$

3.3.2 In degree $2k = 4l + 2$ Now let the degree be $2k = 4l + 2$. In this case τ does not define a nontrivial nontwisted torsion invariant. On the other hand we also just need to determine a since the MMM class vanishes in this degree.

Furthermore, we cannot use the standard universal bundle for linear S^1–bundles $ES^1 \to BS^1$, since ES^1 is contractible and therefore will not admit a nonconstant local system. But we can replace it by a very similar construction. First, recall that ES^1 can be constructed as follows: Take $S^1 \subseteq \mathbb{C}$ and $S^{2N-1} \subseteq \mathbb{C}^N$. Then we have a fibration $S^1 \hookrightarrow S^{2N-1} \to \mathbb{CP}^{N-1}$. Taking the direct limit of this will yield an S^1–bundle with total space S^∞, which is contractible and therefore the universal S^1–principal bundle $S^1 \hookrightarrow S^\infty \to \mathbb{CP}^\infty$.

We can look at a $\mathbb{Z}/n$–action on S^1 given by multiplication with the primitive n^{th} root of unity $e^{2\pi i/n}$. This will give rise to a fiberwise $\mathbb{Z}/n$–action on the bundle $S^1 \hookrightarrow S^{2N-1} \to \mathbb{CP}^{N-1}$. The action of $\mathbb{Z}/n$ on S^{2N-1} is by construction the same as the one being taken to get a lens space L_n^{2N-1} as quotient out of S^{2N-1}. Therefore taking the fiberwise quotient under the given $\mathbb{Z}/n$–action gives a bundle (since $S^1/n \cong S^1$)

$$S^1 \hookrightarrow L_n^{2N-1} \to \mathbb{CP}^{N-1},$$

which yields in the limit to

$$S^1 \hookrightarrow L_n^\infty \to \mathbb{CP}^\infty.$$

We will refer to this bundle as $S^1(\lambda)/n$, since it has the S^1–bundle associated with the universal line bundle as its n–fold covering. The n–fold Galois covering $S^{2N-1} \to L_n^{2N-1}$ gives a bundle $S^{2N-1} \times \mathbb{C} \to L_n^{2N-1}$ where a fixed generator of $\mathbb{Z}/n$ acts on $\mathbb{C}$ by multiplication with an n^{th} root of unity ζ_n. Using this we can make the following important definition.

Definition 3.10 In the setting above, the nonconstant local system $\mathcal{F}_{\zeta_n}$ on L_n^{2N-1} is defined to be the nonconstant local system of the sections of the bundle $S^{2N-1} \times \mathbb{C} \to L_n^{2N-1}$. The nonconstant local system $\mathcal{F}_{\zeta_n}$ on L_n^∞ is defined as the direct limit of these local systems on L_n^{2N-1}.

Again, we can use the fact that the cohomology of $\mathbb{CP}^\infty$ is a group ring and that $H^{2k}(\mathbb{CP}^\infty; \mathbb{R})$ will be spanned by $\operatorname{ch}_{2k}(\lambda)$ and therefore

$$\tau(S^1(\lambda)/n, \mathcal{F}_{\zeta_n}) = s_1 \operatorname{ch}_{2k}(\lambda).$$

Furthermore we have again the following result from Igusa [6]:

Proposition 3.11 *For the Igusa–Klein torsion we have*

$$\tau^{\mathrm{IK}}(S^1(\lambda)/n, \mathcal{F}_{\zeta_n}) = -n^k L_{k+1}(\zeta_n)\,\mathrm{ch}_{2k}(\lambda),$$

where L_{k+1} denotes the polylogarithm

$$L_{k+1}(\zeta) := \mathrm{Re}\left(\frac{1}{i^k}\sum_{l=1}^{\infty}\frac{\zeta_n^l}{n^{k+1}}\right).$$

Putting this together we get

$$a = -s_1/(n^k L_{k+1}(\zeta)).$$

We will prove later that a is independent of the choice of the local system.

4 Extension of higher twisted torsion

Now we present some easy consequences of the geometric additivity and transfer axioms. More precisely, we introduce twisted torsion and calculations thereof for bundles with vertical boundary and bundle pairs. This is completely parallel to the corresponding Section 5 in [7] and all the proofs can be translated word-by-word and will be skipped. While the material is very technical the formulas to keep in mind are Lemma 4.2 and Example 4.7.

First we define the higher twisted torsion on bundles with vertical boundary:

Definition 4.1 (higher twisted torsion for bundles with vertical boundary) Suppose $F \hookrightarrow E \to B$ is a bundle with vertical boundary $\partial^v E \to B$ and local coefficient system $\mathcal{F}$ on E and τ is a higher twisted torsion invariant. Then the twisted torsion of the bundle with boundary is defined by

$$\tau(E,\mathcal{F}) := \tfrac{1}{2}\big(\tau(DE, \mathcal{F}^l \cup_{\mathrm{id}} \mathcal{F}^r) + \tau(\partial^v E, \mathcal{F}_{|\partial^v E})\big),$$

where $DE := E^l \cup_{\mathrm{id}} E^r$ denotes the fiberwise double as before.

Building onto two lemmas one can prove the following formula (compare to [7], Proposition 5.4):

Lemma 4.2 (additivity in the boundary case) *Suppose E is a bundle over B and (E_1, ∂_0) and (E_2, ∂_0) are bundle pairs such that $E_1, E_2 \subseteq E$, $\partial_0 E_1 = \partial_0 E_2 = E_1 \cap E_2$ and $E = E_1 \cup E_2$. Let $\mathcal{F}$ be a local system on E and $\mathcal{F}_1 := \mathcal{F}_{|E_1}$ and $\mathcal{F}_2 := \mathcal{F}_{|E_2}$. Then*

$$\tau(E_1 \cup E_2, \mathcal{F}) = \tau(E_1, \mathcal{F}_1) + \tau(E_2, \mathcal{F}_2) - \tau(E_1 \cap E_2, \mathcal{F}_{|E_1 \cap E_2}).$$

Furthermore, we get the transfer formula (compare [7], Proposition 5.5):

Lemma 4.3 (transfer in the boundary case) *Let $X \to D \xrightarrow{q} E$ be an oriented disc or sphere bundle over a bundle $F \to E \to B$ with local coefficient system $\mathcal{F}$ on E. As for the transfer axiom this pulls up to a local coefficient system $q^*\mathcal{F}$ on D and we get*

$$\tau_B(D, q^*\mathcal{F}) = \chi(X)\tau(E,\mathcal{F}) + \mathrm{tr}_B^E(\tau_E(D), q^*\mathcal{F}).$$

Now we turn to bundle pairs.

Definition 4.4 A pair of bundles $(F, \partial_0) \to (E, \partial_0) \to B$ is called a bundle pair if the vertical boundary $\partial^v E$ is the union $\partial^v E = \partial_0 E \cup \partial_1 E$ of two subbundles which meet along their common boundary $\partial_0 E \cap \partial_1 E = \partial^v \partial_0 E = \partial^v \partial_1 E$.

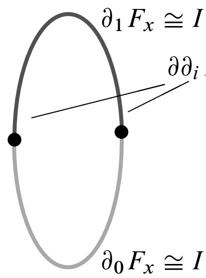


Figure 1: The fiber over x of a bundle pair with fiber $F \cong D^2$

Definition 4.5 (relative torsion) For a bundle pair $(F, \partial_0) \to (E, \partial_0) \to B$ with local coefficient system $\mathcal{F}$ on E we define the relative torsion to be

$$\tau(E, \partial_0, \mathcal{F}) := \tau(E,\mathcal{F}) - \tau(\partial_0 E, \mathcal{F}_{|\partial_0 E}).$$

We get the following proposition (compare [7], Proposition 5.7):

Proposition 4.6 (relative additivity) *Suppose $E \to B$ is a smooth bundle with local coefficient system $\mathcal{F}$, which can be written as the union of two subbundles $E = E_1 \cup E_2$, which meet along a subbundle of their respective vertical boundaries $E_1 \cap E_2 = \partial_0 E_2 \subseteq \partial^v E_1$. Let $\partial^v E_1 = \partial_0 E_1 \cup \partial_1 E_1$ be a decomposition of $\partial^v E_1$, so that $\partial_0 E_2 \subseteq \partial_1 E_1$ and $(E_i, \partial_0) \to B$ for $i = 1, 2$ are smooth bundle pairs. Then $(E, \partial_0 E) \to B$ is a smooth bundle pair and*

$$\tau(E_1 \cup E_2, \partial_0 E_1, \mathcal{F}) = \tau(E_1, \partial_0, \mathcal{F}_{|E_1}) + \tau(E_2, \partial_0, \mathcal{F}_{|E_2}).$$

Example 4.7 The example to keep in mind here are h–cobordism bundles. That is bundle pairs $B \times M \subset E \to B$ such that the fibers are h–cobordisms of M with 0–end the fiber of the trivial bundle. Often we have two h–cobordism bundles $B\times M\subset E\to B$ and $B\times M'\subset E'\to B$ and an inclusion into the 1–end $B\times M'\hookrightarrow E$. This is exactly the situation in which we want to apply the relative additivity, and we get, with an appropriate local system $\mathcal{F}$,

$$\tau(E\cup E', B\times M, \mathcal{F}) = \tau(E, B\times M, \mathcal{F}) + \tau(E', B\times M', \mathcal{F}),$$

where we regard $E\cup E'$ as the h–cobordism bundle obtained by "gluing E'" on top of E.

To state the transfer axiom in the relative case, we need the relative transfer:

$$\mathrm{tr}_B^{(E,\partial_0)}\colon H^*(E;\mathbb{R})\to H^*(B;\mathbb{R}),$$

which is also introduced in [7, Section 5].

Proposition 4.8 (relative transfer; compare [7, Proposition 5.9]) *Let $(F,\partial_0)\to(E,\partial_0)\to B$ and $(X,\partial_0)\to(D,\partial_0)\xrightarrow{q}E$ be bundle pairs with local system $\mathcal{F}$ on E, so that the second bundle is an oriented linear S^n or D^n bundle with $\partial_0 X = S^{n-1}, D^{n-1}$ or $\varnothing$. Then*

$$\tau_B(D,\partial_0 D\cup q^{-1}\partial_0 E, q^*\mathcal{F}) = \chi(X,\partial_0)\tau(E,\partial_0,\mathcal{F}) + \mathrm{tr}_B^{E,\partial_0}(\tau_E(D,\partial_0,q^*\mathcal{F})).$$

Remark 4.9 Note that we do not have a result analogous to the product formula [7, Corollary 5.10]. However, we still have the following corollary.

Corollary 4.10 (stability theorem) *If $(E,\partial_0)\to B$ is a smooth bundle pair with local system $\mathcal{F}$ on E, then so is $(E\times D^n, \partial_0 E\times D^n)$ and the relative torsion is the same:*

$$\tau(E\times D^n, \partial_0 E\times D^n, \mathcal{F}\times\mathbf{1}) = \tau(E,\partial_0,\mathcal{F}),$$

where $\mathcal{F}\times\mathbf{1}$ is the local system constant on D^n.

5 Higher twisted torsion of sphere bundles

The goal of this section is to calculate the higher twisted torsion of linear S^1–bundles only using the axioms. Before we can do this we will discuss why we can always restrict our calculations to finite cyclic local systems on bundles with simply connected base.

5.1 Reduction of the representation

In the following we will simplify the local systems:

Proposition 5.1 *To prove the main theorem, Theorem 3.6, it is enough to only consider bundles with simply connected base instead of base having finite fundamental group. We can also restrict to only considering local systems (on the fiber) that induce n–fold holonomy covers with transition group $\mathbb{Z}/n$ instead of just finite local systems.*

Remark 5.2 Let $F \hookrightarrow E \to B$ be a fiber bundle and $\mathcal{F}$ a finite local system on E. This corresponds to its holonomy cover $\widetilde{E} \to E$ with finite transition group G and representation $\rho\colon G \to U(m)$. On the other hand every finite covering $\widetilde{E} \xrightarrow{G} E$ with representation $\rho\colon G \to U(m)$ gives us a local system $\mathcal{F}_\rho$ as the sections of the bundle $\widetilde{E} \times_G \mathbb{C}^m \to E$ where G acts on $\mathbb{C}^m$ via ρ. This construction is a one-to-one correspondence. Now let $H \subseteq G$ be a subgroup. From the covering $\widetilde{E} \xrightarrow{G} E$ we get coverings $\pi_H\colon \widetilde{E}/H \to E$ and $\widetilde{E} \xrightarrow{H} \widetilde{E}/H$. Suppose we have a representation $\rho_H\colon H \to U(m)$ and thereby get a local system $\mathcal{F}_{\rho_H}$ on $\widetilde{E}/H$. Then we can either form the induced representation $\operatorname{Ind}_H^G(\rho_H)\colon G \to U(m)$ and its corresponding local system $\mathcal{F}_{\operatorname{Ind}_H^G(\rho_H)}$ on E or the local system $\pi_*\mathcal{F}_{\rho_H}$ on E given by the push-down of the local system $\mathcal{F}_\rho$. It follows from an easy calculation that

$$\mathcal{F}_{\operatorname{Ind}_H^G(\rho_H)} = \pi_*\mathcal{F}_{\rho_H}.$$

Proof of the proposition Let $\mathcal{F}$ be again a local system on E corresponding to a finite covering $\widetilde{E} \xrightarrow{G} E$ with representation $\rho\colon G \to U(m)$. Let $\mathcal{H} = \{H_i\}$ be the finite set of cyclic subgroups H_i of G. By Artin's induction theorem, we can write the character of ρ rationally as linear combination of characters of one-dimensional representations. Since we are working over $\mathbb{C}$, we therefore can write ρ rationally as a linear combination of one-dimensional representations $\lambda_i\colon H_i \to U(1)$ and inductions thereof. Concretely we have

$$n\rho \cong \bigoplus_i n_i \operatorname{Ind}_{H_i}^G(\lambda_i) \quad \text{with } n, n_i \in \mathbb{Z}.$$

Let τ be a twisted torsion invariant and $\pi_i\colon \widetilde{E}/H \to E$ be a covering. Then we have, using the transfer of coefficient axiom and the calculation above,

$$\begin{aligned} n\tau(E,\mathcal{F}) &= \sum_i n_i \tau(E, \mathcal{F}_{\operatorname{Ind}_{H_i}^G(\lambda_i)}) = \sum_i n_i \tau(E, \pi_*\mathcal{F}_{\lambda_i}) \\ &= \sum_i n_i \tau(\widetilde{E}/H_i, \mathcal{F}_{\lambda_i}) \in H^{2k}(B;\mathbb{R}). \end{aligned}$$

Therefore it suffices for the rest of the paper to work with local systems with n–fold holonomy covers with cyclic transition group $\mathbb{Z}/n$.

Now let $F \to E \to B$ be a bundle with local system $\mathcal{F}$ on E and the base B having a finite fundamental group. We have the universal covering $q\colon \widetilde{B} \to B$ and pulling back

E along q gives a bundle $\widetilde{E} := q^*E \to \widetilde{B}$ with local system $\widetilde{\mathcal{F}} := q^*\mathcal{F}$. Naturality implies

$$\tau(\widetilde{E},\widetilde{\mathcal{F}}) = q^*\tau(E,\mathcal{F}) \in H^{2k}(\widetilde{B};\mathbb{R}).$$

Furthermore we know that $q^*\colon H^{2k}(B;\mathbb{R}) \to H^{2k}(\widetilde{B};\mathbb{R})$ is injective. By this construction it suffices to prove the main theorem only on bundles with simply connected base. □

5.2 Twisted torsion for S^1–bundles

We want to show the following theorem:

Theorem 5.3 *For every S^1–bundle $S^1 \hookrightarrow E \to B$ with B simply connected and local system $\mathcal{F}$ on E with $\mathbb{Z}/n$–fold holonomy cover $\widetilde{E}_n \to E$ every twisted torsion invariant τ is given by*

$$\tau(E,\mathcal{F}) = a\tau^{\mathrm{IK}}(E,\mathcal{F}),$$

where a is the scalar defined earlier.

We will follow an approach Igusa introduced in [8, Section 4]. Since $B\mathrm{Diff}(S^1) \simeq B\mathrm{SO}(2)$ it suffices to look at linear S^1–bundles. These pull back from the universal S^1–bundle $S^1(\lambda)$ given by $S^1 \hookrightarrow S^\infty \to \mathbb{CP}^\infty$.

Let $E \to B$ be an S^1–bundle with local system $\mathcal{F}$ on E inducing a finite holonomy covering. At first we look at the following n–fold holonomy Galois covering:

$$\begin{tikzcd} S^1 \arrow[r, "n"] \arrow[d] & S^1 \arrow[d] \\ \widetilde{E}_n \arrow[r, "n"] \arrow[d] & E \arrow[d] \\ B \arrow[r, equal] & B \end{tikzcd}$$

Now $\widetilde{E}_n$ is again a linear S^1–bundle with fiberwise $\mathbb{Z}/n$–action. This will pull back equivariantly from the universal S^1–bundle $S^1(\lambda)$ given by $S^\infty \to \mathbb{CP}^\infty$, which also admits an $\mathbb{Z}/n$–action. Therefore E will pull back from the quotient $S^1(\lambda)/(\mathbb{Z}/n)$. Also the local system $\mathcal{F}$ on E will pull back from the local system $\mathcal{F}_{\zeta_n}$ on S^∞ for some n^{th} root of unity ζ_n. We defined this earlier (Definition 3.10) to be given by the bundle $S^1(\lambda)\times\mathbb{C} \to S^1(\lambda)/n$ where the action on $\mathbb{C}$ is given by multiplication by ζ_n. So because of naturality it is enough to show:

Theorem 5.4 *For all n and ζ_n,*

$$\tau(S^1(\lambda)/n,\mathcal{F}_{\zeta_n}) = a\tau^{\mathrm{IK}}(S^1(\lambda)/n,\mathcal{F}_{\zeta_n}) \in H^{2k}(\mathbb{CP}^\infty;\mathbb{R}).$$

First we prove two important lemmas already introduced in [8] (Lemmas 4.11 and 4.12). These will isolate certain properties of $\tau(S^1(\lambda)/n, \mathcal{F}_\zeta)$ thought of as a function of ζ.

Lemma 5.5 *Suppose we have a bundle $E \to B$ and a free fiberwise nm–action on E, where $n, m \in \mathbb{N}$. Then we have, for any twisted torsion invariant and n^{th} root of unity ζ_n,*

$$\tau(E/n, \mathcal{F}_{\zeta_n^m}) = \sum_{\xi^m=1} \tau(E/(nm), \mathcal{F}_{\xi\zeta_n}),$$

where the local systems $\mathcal{F}_{\zeta_n}$ on E/n are given by the construction above.

Proof Denote the projection by $\pi\colon E/n \to E/(nm)$. We get

$$\pi_*\mathcal{F}_{\zeta_n^m} = \bigoplus_{\xi^m=1} \mathcal{F}_{\xi\zeta_n}.$$

Now we can use the transfer of coefficients and the additivity axiom to get

$$\tau(E/n, \mathcal{F}_{\zeta_n^m}) = \tau(E/(nm), \pi_*\mathcal{F}_{\zeta_n^m}) = \sum_{\xi^m=1} \tau(E/(nm), \mathcal{F}_{\xi\zeta_n}). \qquad \square$$

Lemma 5.6 *For every linear S^1–bundle $E \to B$ and any n^{th} root of unity ζ_n, we have, for every twisted torsion class in degree $2k$,*

$$\tau(E/(nm), \mathcal{F}_{\zeta_n}) = m^k \tau(E/n, \mathcal{F}_{\zeta_n}).$$

Proof Again we look at the universal circle bundle $S^1(\lambda)$, and by the naturality axiom it is enough to show the lemma only on $E = S^1(\lambda)$. We have that $S^1(\lambda)/m$ is again a circle bundle over $\mathbb{CP}^\infty$ and therefore classified by a map

$$f_m\colon \mathbb{CP}^\infty \to \mathbb{CP}^\infty.$$

In degree 2 we can see (by looking at circle bundles over spheres S^2) that this map is multiplication by m on H^2. Then it follows that f_m^* is multiplication by m^k on $H^{2k}(\mathbb{CP}^\infty; \mathbb{R})$. The classifying maps for $S^1(\lambda)/nm$ and $S^1(\lambda)/n$ are related by

$$f_{mn} = f_n \circ f_m.$$

The lemma now follows from naturality. $\square$

Now let $f\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{C}$ a function. It is said to satisfy the Kubert identity if

$$f(x) = m^{s-1} \sum_{k=0}^{m-1} f\left(\frac{x+k}{m}\right)$$

for fixed s and all integers m and all $x \in \mathbb{Q}/\mathbb{Z}$. Identifying $\mathbb{Q}/\mathbb{Z}$ with the roots of unity in $\mathbb{C}$ (by $x \mapsto e^{2\pi i x}$), we can write $f(x) = L(e^{2\pi i x})$ and the Kubert identity becomes

$$L(\zeta^m) = m^{s-1} \sum_{\xi^m = 1} L(\zeta\xi).$$

The following result can be proved by considering Fourier coefficients:

Theorem 5.7 (Milnor 1983 [10, Section 3, Theorem 1]) *Let $\mathbb{Q}/\mathbb{Z}$ have the quotient topology. The space of continuous functions $f\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{C}$ satisfying the Kubert identity is two-dimensional and splits into two one-dimensional spaces, the first of which contains all the functions with $L(\zeta) = L(\overline{\zeta})$ and the second, the ones with $L(\zeta) = -L(\overline{\zeta})$.*

Remark 5.8 Milnor states this theorem for continuous functions $\mathbb{R}/\mathbb{Z} \to \mathbb{C}$ rather than $\mathbb{Q}/\mathbb{Z} \to \mathbb{C}$, but since $\mathbb{Q} \subset \mathbb{R}$ is dense this does not impact the statement. Furthermore the original theorem is formulated without identifying $\mathbb{R}/\mathbb{Z}$ with the unit sphere in $\mathbb{C}$. The two one-dimensional subspaces are then formed by the functions satisfying $f(x) = -f(1-x)$ and $f(x) = f(1-x)$, which correspond exactly to the equations $L(\zeta) = \pm L(\overline{\zeta})$ on the unit sphere.

Proof of Theorem 5.4 To any higher twisted torsion invariant τ we get, for any n^{th} root of unity, a coefficient $s_1(\tau, \zeta)$ defined by

$$\tau(S^1(\lambda)/n, \mathcal{F}_\zeta) = s_1(\tau, \zeta)\,\mathrm{ch}_{2k}(\lambda) \in H^{2k}(\mathbb{CP}^\infty; \mathbb{R}) \cong \mathbb{R}.$$

Identifying $\mathbb{Q}/\mathbb{Z}$ with the roots of unity, we get a function $f_\tau\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{R}$ defined by

$$f_\tau(\zeta) := \frac{1}{n^k} s_1(\tau, \zeta),$$

where $\zeta^n = 1$. This is well defined, since by the previous lemma we have

$$\tau(S^1(\lambda)/(nm), \mathcal{F}_\zeta) = m^k \tau(S^1(\lambda)/n, \mathcal{F}_\zeta),$$

so $f_\tau(\zeta)$ is by construction independent from the choice of n with $\zeta^n = 1$.

Our goal is to show that this satisfies the Kubert identity and then to use Milnor's result to prove our theorem. But for this, f_τ needs to be continuous, a fact which we cannot prove, but must assume. Therefore we need the following last axiom:

Axiom 7 (continuity) For any twisted torsion invariant, the function $f_\tau\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{R}$ constructed above is continuous.

As explained earlier in the paper, this axiom basically states that for a fixed bundle $E \to B$ the twisted torsion depends continuously on the local system $\mathcal{F}$ on E.

Continuation of the proof Now we calculate for $\zeta \in \mathbb{Q}/\mathbb{Z}$ with $\zeta^n = 1$ using the two lemmas from above:

$$\begin{aligned} f_\tau(\zeta^m)\,\mathrm{ch}_{2k}(\lambda) &= \frac{1}{n^k}\tau(S^1(\lambda)/nm, \mathcal{F}_{\zeta^m}) \\ &= \frac{1}{n^k}\sum_{\xi^m=1}\tau(S^1(\lambda)/nm, \mathcal{F}_{\xi\zeta}) \\ &= m^k \sum_{\xi^m=1} f_\tau(\xi\zeta)\,\mathrm{ch}_{2k}(\lambda). \end{aligned}$$

So f_τ satisfies the Kubert identity (with $s = k+1$) for any τ.

We note that the change of representation from ζ to $\overline{\zeta}$ represents a change of orientation in the fiber. Therefore, it corresponds to a map $g\colon \mathbb{CP}^\infty \to \mathbb{CP}^\infty$, giving $g_*\colon \pi_1 S^1 \to \pi_1 S^1$ as multiplication by -1. Using that $\pi_1 S^1(\lambda)/n \cong \mathbb{Z}/n$, we get the following commutative diagram relating the exact sequence of the homotopy groups of the fibration $S^1 \hookrightarrow S^1(\lambda)/n \to \mathbb{CP}^\infty$ to itself under g_*:

$$\begin{tikzcd} \pi_2\mathbb{CP}^\infty \arrow[r, "\cdot n"] \arrow[d, "g_*"] & \mathbb{Z} \arrow[r, two heads] \arrow[d, "-1"] & \mathbb{Z}/n \arrow[r] \arrow[d, "g_*"] & 0 \\ \pi_2\mathbb{CP}^\infty \arrow[r, "\cdot n"] & \mathbb{Z} \arrow[r, two heads] & \mathbb{Z}/n \arrow[r] & 0 \end{tikzcd}$$

From this one can see that $g_*\colon \pi_2\mathbb{CP}^\infty \to \pi_2\mathbb{CP}^\infty$ is multiplication by -1. Since $\mathbb{CP}^\infty$ is simply connected, g_* is also multiplication by -1 in homology of degree 2. Since $\mathbb{CP}^\infty$ is an Eilenberg–Mac Lane space, g^* must be multiplication by -1 on degree-2 cohomology and thus multiplication by $(-1)^k$ on degree-$2k$ cohomology.

This yields

$$f_\tau(\zeta) = (-1)^k f_\tau(\overline{\zeta})$$

for any τ with degree $2k$. So f_τ is in one specific one-dimensional subspace of the space of functions satisfying the Kubert identity for any torsion invariant τ of degree $2k$, and therefore we have, for an arbitrary torsion invariant τ and the Igusa–Klein torsion τ^{IK},

$$f_\tau = a f_{\tau^{\mathrm{IK}}}$$

for a certain $a \in \mathbb{R}$. This translates to

$$\tau(S^1(\lambda)/n, \mathcal{F}_\zeta) = a\tau^{\mathrm{IK}}(S^1(\lambda)/n, \mathcal{F}_\zeta)$$

for any root of unity ζ and proves the theorem. □

Remark 5.9 This also shows that the scalar a that we calculated earlier by choosing an arbitrary local system is well-defined and does not depend on this choice.

Remark 5.10 It is an unproven conjecture by Milnor [10] that any function satisfying the Kubert identity is already continuous. If this conjecture was proven we could drop the continuity axiom.

6 The difference torsion

Given a twisted torsion invariant τ, we can now form the twisted difference torsion

$$\tau^\delta := \tau - a\tau^{\mathrm{IK}} - bM,$$

where the scalars a and b are the ones from Theorem 3.6 (and b is 0 if the torsion has degree $4l+2$). Clearly, τ^δ is a twisted torsion invariant.

Our goal in this section and the next is to show $\tau^\delta(E,\mathcal{F}) = 0$ for every bundle $E \to B$ with every local coefficient system $\mathcal{F}$ on E and base B having finite fundamental group. In this section we will show that τ^δ is a fiber homotopy invariant. Here is a sketch of our approach: Given two fiber bundles $E \to B$ and $E' \to B'$ with appropriate local system $\mathcal{F}$ and fiber homotopy equivalence $g\colon E \to E'$ (which we can without restriction assume to be an embedding) we can view $E' \backslash g(E)$ as a bundle with fibers h–cobordisms (this is not necessarily an h–cobordism bundle) the torsion of which is exactly the difference of the torsions of E and E'. (This is done in the proof of Theorem 6.12.) To show that the difference torsion of bundles with h–cobordisms as fibers is trivial in Lemma 6.11 we embed one end of the bundle in a trivial lens space bundle (we have to use lens spaces instead of discs or spheres to preserve a nontrivial first homotopy group) making it a bundle fiber homotopy equivalent to a trivial lens space bundle. Now we use the fiber homotopy from such a bundle to the trivial lens space bundle to get an h–cobordism bundle of a lens space (done in Lemma 6.10). Finally, in our paper [4] we essentially classified all h–cobordism bundles of a lens space and showed that their Igusa–Klein torsion can be calculated only using the axioms. So their difference torsion is zero.

6.1 Lens spaces

Any cyclic group $\mathbb{Z}/n$ acts on the complex numbers $\mathbb{C}$ by rotation. For the rest of the paper, we will pick a generator $1 \in \mathbb{Z}/n$ and have it act by multiplication with $e^{2\pi i/n}$ on $\mathbb{C}$. Then we get a componentwise action on the odd-dimensional sphere $S^{2N+1} \subset \mathbb{C}^{N+2}$.

Definition 6.1 The odd-dimensional lens space L_n^{2N+1} is defined to be the quotient $S^{2N+1}/(\mathbb{Z}/n)$ by the action defined above.

It is well known that the CW–structure on L_n^{2N+1} has a cell in every dimension and its associated chain complex is given by

$$0 \to \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{n} \mathbb{Z} \xrightarrow{0} \cdots \xrightarrow{0} \mathbb{Z} \xrightarrow{n} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \to 0.$$

In particular we see that L_n^{2N+1} is rationally spherical with fundamental group $\mathbb{Z}/n$. Recall that we are interested in twisted torsion invariants and thereby require our manifolds to have nontrivial fundamental group, so the odd-dimensional lens spaces will play the role of finite-dimensional spheres in some sense.

In Section 7 we will also need spaces with nontrivial fundamental group that are rationally contractible to provide a twisted analogue of the infinite-dimensional sphere $S^\infty \simeq *$ or large-dimensional discs D^N. The even-dimensional lens spaces are exactly going to fulfill this condition:

Definition 6.2 The even-dimensional lens space $L_n^{2N} \subset L_n^{2N+1}$ is obtained from the odd-dimensional one by omitting the top cell in the CW decomposition described above.

It follows immediately that L_n^{2N} is rationally acyclic with fundamental group $\mathbb{Z}/n$. We choose a universal covering $\widetilde{L_n^{2N}} \to L_n^{2N}$. This comes equipped with a $(2N-1)$–connected map $\tilde{i}\colon \widetilde{L_n^{2N}} \to S^{2N+1}$.

Lastly, note that there is a chain of inclusions

$$\cdots \subset L_n^{2N-2} \subset L_n^{2N-1} \subset L_n^{2N} \subset L_n^{2N+1} \subset \cdots.$$

6.2 Lens space bundles

Following the outline above, we first want to show that the difference torsion is zero on every linear odd-dimensional lens space bundle $L_n^{2N+1} \hookrightarrow E_n^{2N+1} \to B$ with local coefficient system $\mathcal{F}$ on E_n^{2N+1}. We already know from the base case that the difference torsion is zero on every S^1–bundle. Furthermore, if we take an S^l–bundle with $l > 1$ or disc bundle, we know that the fundamental group of the fiber is trivial and it therefore admits no nonconstant local system. So the twisted difference torsion on these bundles is always given by the nontwisted difference torsion. But the nontwisted difference torsion is zero everywhere as Igusa showed in [7]. From this we get the following lemma:

Lemma 6.3 *For the difference torsion τ^δ associated with any higher twisted torsion invariant, we have*

$$\tau^\delta(E, \mathcal{F}) = 0$$

for any disc or sphere bundle $E \to B$ with local system $\mathcal{F}$ on E.

At first we will prove:

Lemma 6.4 *The difference torsion is* 0 *on any linear odd-dimensional lens space bundle* $L_n^{2N+1} \hookrightarrow E_n^{2N+1} \to B$*. By linear we mean that it is covered by a linear sphere bundle* $S^{2N+1} \hookrightarrow \widetilde{E}^{2N+1} \to B$*.*

Remark 6.5 The corresponding statement [7, Lemma 7.3] only deals with linear disc bundles, the proof of which follows swiftly from the product formula for relative torsion. Unfortunately, there is no twisted product formula, so our proof is slightly more difficult.

Proof The covering sphere bundle $\widetilde{E}^{2N+1}$ is a subbundle of an $(N+1)$–dimensional complex vector bundle. By the splitting principle, it suffices to look at the direct sum of $N+1$ complex line bundles. The sphere bundle will become the fiberwise join of the circle bundles associated with the line bundles:

$$S^1 * \cdots * S^1 \hookrightarrow \widetilde{E}_1^1 * \cdots * \widetilde{E}_{N+1}^1 \to B.$$

Now we have

$$\begin{aligned} L_n^{2N+1} &\cong (S^{2N-1} * S^1)/n \\ &= (S^{2N-1} \times D^2)/n \cup_{(S^{2N-1} \times S^1)/n} (D^{2N} \times S^1)/n. \end{aligned}$$

Fiberwise, this gives us

$$E_n^{2N+1} = H_n^{2N-1} \cup H_n^1,$$

where $H_n^{2N-1} \to B$ is an $(S^{2N-1} \times D^2)/n$–bundle and $H_n^1 \to B$ is a $(D^{2N} \times S^1)/n$–bundle, both meeting along their common vertical boundary, which is given by an $(S^{2N-1} \times S^1)/n$–bundle G_n. The $\mathbb{Z}/n$–action is given by the simultaneous action on each component of the products. While the $\mathbb{Z}/n$–action on any disc is not free, the simultaneous action will guarantee that it is free on the product. We can restrict every local coefficient system $\mathcal{F}$ on E_n^{2N+1} to H_n^{2N-1}, H_n^1 and G_n and use the additivity axiom.

Now we will continue the proof by induction. We know that the difference torsion is 0 on every $L_n^1 \cong S^1$–bundle. Let us then assume that the difference torsion is 0 on any linear L_n^{2N-1}–bundle with any representation of the fundamental group. Given a linear L_n^{2N+1}–bundle $E_n^{2N+1} \to B$ with local coefficient system $\mathcal{F}$, the construction above yield, by Lemma 4.2,

$$\tau^\delta(E_n^{2N+1}, \mathcal{F}) = \tau^\delta(H_n^{2N-1}, \mathcal{F}_{|H_n^{2N-1}}) + \tau^\delta(H_n^1, \mathcal{F}_{|H_n^1}) - \tau^\delta(G_n, \mathcal{F}_{|G_n}).$$

We have nontrivial fibrations

$$D^2 \hookrightarrow (S^{2N-1} \times D^2)/n \to L_n^{2N-1},$$

$$\begin{aligned} D^{2N} \hookrightarrow\ & (D^{2N} \times S^1)/n \ \to L_n^1, \\ S^1 \to\ & (S^{2N-1} \times S^1)/n \to L_n^{2N-1}. \end{aligned}$$

The first of these splits the bundle H_n^{2N-1} in the following manner:

$$\begin{tikzcd} D^2 \arrow[r, hook] & (S^{2N-1} \times D^2)/n \arrow[rr] \arrow[d, hook] & & L_n^{2N-1} \arrow[dl, hook] \\ D^2 \arrow[r, hook] & H_n^{2N-1} \arrow[r] \arrow[d] & J_n \arrow[dl] & \\ & B & & \end{tikzcd}$$

where $J_n \to B$ is an L_n^{2N-1}–bundle and $H_n^{2N-1} \to J_n$ is a D^2–bundle. Since D^2 is contractible, we get a local system $\mathcal{F}_J$ on J_n the pull-back of which to H_n^{2N-1} is isomorphic to $\mathcal{F}_{|H_n^{2N-1}}$. Now we can use the geometric transfer and the fact that we already determined the difference torsion to be 0 on L_n^{2N-1}–bundles and D^2–bundles to show

$$\tau^\delta(H_n^{2N-1}, \mathcal{F}_{|H_n^{2N-1}}) = \chi(D^2)\tau(J_n, \mathcal{F}_J) + \operatorname{tr}_B^{J_n}(\tau_{J_n}(H_n^{2N-1}, \mathcal{F}_{|H_n^{2N-1}})) = 0.$$

A similar argument holds for H_n^1 and G_n, and this completes the proof. □

6.3 Difference torsion as a fiber homotopy invariant

In this section, we will prove that the difference torsion τ^δ is a fiber homotopy invariant. By this we mean that for any two bundles $F_1 \hookrightarrow E_1 \to B$ and $F_2 \hookrightarrow E_2 \to B$ and fiber homotopy equivalence $f\colon E_1 \to E_2$ with local coefficient systems $\mathcal{F}_2$ on E_2 and $f^*\mathcal{F}_2 \cong \mathcal{F}_1$ on E_1, we have

$$\tau^\delta(E_1, \mathcal{F}_1) = \tau^\delta(E_2, \mathcal{F}_2) \in H^{2k}(B; \mathbb{R}).$$

This section will greatly rely on the construction of the equivariant Hatcher examples from [4]. We will especially use some techniques involving h–cobordism bundles, for a basic depiction of which the reader is also referred to [4, Section 1].

First we show the following lemmas:

Lemma 6.6 *For any linear disc bundle $D \xrightarrow{q} E$ and any bundle pair $(E, \partial_0) \to B$ with local coefficient system $\mathcal{F}$ we have*

$$\tau_B^\delta(D, \partial_0, q^*\mathcal{F}) = \tau_B^\delta(E, \partial_0, \mathcal{F}),$$

where we pull the system up to D and $\partial_0 D = q^{-1}\partial_0 E$ as usual.

Proof By geometric transfer (Proposition 4.8) we have

$$\tau_B^\delta(D, \partial_0, q^*\mathcal{F}) = \tau^\delta(E, \partial_0, \mathcal{F}) + \mathrm{tr}_B^E(\tau_E^\delta(D, q^*\mathcal{F}))$$

and $\tau_E^\delta(D, q^*\mathcal{F}) = 0$ because D is a disc bundle over E. □

Remark 6.7 The same statement still holds in the nonrelative case.

We will now need to prove three subsequent lemmas before we can prove the fiber homotopy invariance.

Remark 6.8 As before (Proposition 5.1) it is enough to look at local systems that induce holonomy covers with cyclic transformation group. So we will always assume that.

Lemma 6.9 *Let B be a space with finite fundamental group. Then for sufficiently large integers N the difference torsion τ^δ is zero on any h–cobordism bundle of L_n^{2N-1} over B for a given n.*

Proof Since we can assume that B is simply connected, all local systems on an h–cobordism bundle of $L_n^{2N-1} \times D^M$ inducing an n–fold cyclic holonomy are isomorphic to the local systems of the form $\mathcal{F}_\zeta$, where ζ is an n^{th} root of unity. We will now fix such a ζ.

We will follow Igusa [7, Lemma 7.11] closely in his discussion of the untwisted version of this crucial proof. By the stability of higher torsion (Corollary 4.10) we can view the difference torsion as a map

$$\tau^\delta(\_\_, \mathcal{F}_\zeta)\colon [B, B\mathcal{P}(L_n^{2N-1})] = [B, B(\mathrm{colim}_M\, \mathcal{C}(L_n^{2N-1} \times D^M))] \to H^*(B; \mathbb{R})$$

sending an h–cobordism bundle $h \to B$ to $\tau^\delta(h, \mathcal{F}_\zeta)$. Here $\mathcal{C}(M)$ is the concordance space and $\mathcal{P}(M)$ is the stable concordance space; for details see [4, Section 1]. We can give the set $[B, B\mathcal{P}(L_n^{2N-1})]$ a group structure by the fiberwise gluing together of the h–cobordisms as explained in [7]. From the additivity properties of higher twisted torsion (in particular Example 4.7) it follows that $\tau^\delta(\_\_, \mathcal{F}_\zeta)$ is a group homomorphism. So it is enough to give rational generators of $[B, B\mathcal{P}(L_n^{2N-1})]$ and show that the difference torsion is zero on these generators.

For N large enough ($N \gg \dim B$) we have

$$[B, B\mathcal{P}(L_n^{2N-1})] = [B, \mathcal{H}(L_n^{2N-1})] \cong [B, \mathcal{H}(B\mathbb{Z}/n)],$$

where $\mathcal{H}$ denotes the classifying space of h–cobordism bundles. In [4, Section 3.2], we define the twisted Hatcher maps $\Delta^i\colon G_n/U \to \mathcal{H}(B\mathbb{Z}/n)$ and the main theorem thereof uses those to show that the space $\mathbb{Q} \otimes [B, \mathcal{H}(B\mathbb{Z}/n)]$ is spanned by various Hatcher constructions (also defined in [4]) of one nontrivial vector bundle ξ over B with fiber a homotopically trivial sphere bundle. The calculations in [4, Section 4.1] only rely on the axioms and ensure that the difference torsion of these Hatcher constructions is zero. This is because the nontrivial Igusa–Klein torsion of those bundles arises from the torsion of sphere bundles and linear lens space bundles. □

Lemma 6.10 *Let N be an sufficiently large integer and $E \to B$ a bundle with local system $\mathcal{F}$ on E inducing an n–fold cyclic holonomy covering. Then we have $\tau^\delta(E, \mathcal{F}) = 0$ if there is a fiber homotopy equivalence:*

$$\begin{tikzcd} E \arrow[r, "\sim"] \arrow[d] & L_n^{2N-1} \times B \arrow[d] \\ B \arrow[r, "="] & B \end{tikzcd}$$

Proof Denote the fiber homotopy equivalence $H\colon E \to L_n^{2N-1} \times B$. We can take the product of $L_n^{2N-1} \times B$ with a large-dimensional disc D^M and make H into an embedding

$$\overline{H}\colon E \xhookrightarrow{\sim} D^M \times L_n^{2N-1} \times B.$$

Then we can take a tubular neighborhood of $\overline{H}(E) \subseteq D^M \times L_n^{2N-1} \times B$ to get a codimension-0 embedding of an M'–dimensional disc bundle $D(E)$ over E

$$G\colon D(E) \xhookrightarrow{\sim} D^M \times L_n^{2N-1} \times B.$$

Then $(D^M \times L_n^{2N-1} \times B)\backslash G(D^\circ(E))$ is an h–cobordism bundle of $L_n^{2N-1} \times S^{M-1}$ over B and by Lemma 4.2 its difference torsion is given by

$$\begin{aligned} \tau^\delta((D^M \times L_n^{2N-1} \times B)\backslash G(D^\circ(E)), \mathcal{F}) + \tau^\delta(D(E), \mathcal{F}) - \tau^\delta(S(E), \mathcal{F}) \\ = \tau^\delta(D^M \times L_n^{2N-1} \times B, \mathcal{F}) \\ = 0, \end{aligned}$$

since the last bundle is trivial. $S(E)$ denotes the sphere bundle given as the vertical boundary of $D(E)$. We can use the transfer axiom to show that

$$\tau^\delta(D(E), \mathcal{F}) = \tau^\delta(E, \mathcal{F}),$$

and, given that M' is even,

$$\tau^\delta(S(E), \mathcal{F}) = \chi(S^{M'-1})\tau^\delta_B(E, \mathcal{F}) + \operatorname{tr}^E_B \tau^\delta_E(S(E), \mathcal{F}) = 0,$$

because the difference torsion is zero on any disc and sphere bundles. Therefore it suffices to show that the difference torsion is zero on any h–cobordism bundle of $L_n^{2N-1} \times S^{M-1}$ over B for arbitrarily large N. Such a bundle can easily be reduced to an h–cobordism bundle of L_n^{2N-1} without changing its torsion: Let $H \to B$ be an h–cobordism bundle of $L_n^{2N-1} \times S^{M-1}$. We can embed $S^{M-1} \times I$ as a tubular neighborhood of S^{M-1} into D^M and thereby get

$$H \supseteq L_n^{2N-1} \times S^{M-1} \times B \hookrightarrow L_n^{2N-1} \times D^M \times 1 \times B \subseteq L_n^{2N-1} \times D^M \times I \times B,$$

and we can define the h–cobordism bundle of $L_n^{2N-1} \times D^M$ (and thereby of L_n^{2N-1} by stability)

$$H' := H \cup_{L_n^{2N-1} \times S^{M-1} \times B} L_n^{2N-1} \times D^M \times I \times B.$$

Intuitively, we get H' by gluing the h–cobordism bundle H of $L_n^{2N-1} \times S^{M-1}$ on top of a trivial h–cobordism bundle of $L_n^{2N-1} \times D^M$ along the inclusion $S^{M-1} \hookrightarrow D^M$.

We calculate, using the relative additivity properties of higher torsion (Example 4.7), for any local system $\mathcal{F}$ in L_n^{2N-1} extended naturally to H, H' and $L_n^{2N-1} \times D^M \times I \times B$,

$$\begin{aligned}
\tau^\delta(H', \mathcal{F}) &= \tau^\delta(H', L_n \times D^M \times 0 \times B, \mathcal{F}) \\
&= \tau^\delta(H, L_n^{2N-1} \times S^{M-1} \times B, \mathcal{F}) \\
&\qquad + \tau^\delta(L_n^{2N-1} \times D^m \times I \times B, L_n^{2N-1} \times D^m \times 0 \times B, \mathcal{F}) \\
&= \tau^\delta(H, \mathcal{F}).
\end{aligned}$$

With this construction on h–cobordism bundles the proof now follows from the previous lemma. □

Lemma 6.11 *The difference torsion τ^δ is 0 on any bundle pair $(E, \partial_0) \to B$ the fibers (F, ∂_0) of which are h-cobordisms and have a local system $\mathcal{F}$ inducing a cyclic n–fold holonomy covering.*

Proof This proof can be translated directly from the proof of Lemma 8.3 in [7] by replacing the high-dimensional discs D^N with high-dimensional lens spaces L_n^{2N}. □

Theorem 6.12 *The difference torsion τ^δ is a fiber homotopy invariant of smooth bundle pairs with local systems.*

Proof Same as for Theorem 8.4 in [7]. □

Remark 6.13 Since τ^δ is a fiber homotopy equivalence, it is well defined on any fibration $(Z, C) \to B$ with fiber (X, A) and local system $\mathcal{F}$ on X which is smoothable in the sense that it is fiber homotopy equivalent to a smooth bundle pair (E, ∂_0) with compact manifold fiber (F, ∂_0).

7 Triviality of the difference torsion

Using the fiber homotopy invariance of the difference torsion we will first show that we can replace any bundle $E \to B$ with another one with the same torsion and real acyclic fiber. Then we will show that the difference torsion (or more general any torsion invariant that is fiber homotopy invariant) must be zero on any bundle with acyclic fibers.

7.1 Lens space suspensions

As outlined above, our first goal is to eliminate the real homology groups of the fiber F of a bundle $E \to B$. We will use the fact that the stable homotopy groups of F are rationally equivalent to the rational homology groups. This means that sufficiently large k and for an element $\alpha \in H_{m+k}(\Sigma^k F;\mathbb{R}) \cong H_m(F;\mathbb{R})$ there is a map $S^{m+k} \to \Sigma^k F$ representing α as an element of $\pi_{m+k}(\Sigma^k F) \otimes \mathbb{R}$. We then can glue in an $(m+k+1)$–cell along this map to effectively kill off the element α and continue inductively. Unfortunately, this naive construction has a big problem for us: even just one suspension destroys the first homotopy group of F leaving us with only the trivial local system which is not very interesting (or helpful). So we need an alternative suspension construction that shifts up the rational homology groups, exhibits the isomorphism to stable homotopy groups after sufficiently many suspensions and preserves the first homotopy group. We achieve all of this by suspending via a push-out along two high-dimensional (even) lens spaces rather than discs.

Let us recall that the usual suspension ΣF is defined by the (homotopy) push-out:

$$\begin{tikzcd} F \arrow[r] \arrow[d] & D^N \arrow[d] \\ D^N \arrow[r] & \Sigma F \end{tikzcd}$$

Since D^N is contractible, we know that $\pi_1 \Sigma F = 0$, and therefore this construction cannot give us a nonconstant local system on ΣF. Now we make the following definition:

Definition 7.1 (lens space suspension) Let F be a topological space with local system $\mathcal{F}$ on F inducing an n–fold holonomy cover $\widetilde{F} \to F$ with finite cyclic transition group. The cover gives us a mapping $F \to L_n^{2N}$ for a large $N \in \mathbb{N}$ (because $L_n^\infty \cong K(\mathbb{Z}/n,1)$). Using this map, we can define the lens space suspension $\Sigma_n F$ as the homotopy push-out:

$$\begin{tikzcd} F \arrow[r] \arrow[d] & L_n^{2N} \arrow[d] \\ L_n^{2N} \arrow[r] & \Sigma_n(N)F \end{tikzcd}$$

Remark 7.2 We will drop N from the notation and consider it to be very large.

We have the earlier introduced local systems $\mathcal{F}_\zeta$ on L_n^{2N} for an n^{th} root of unity ζ. By choosing the map $i\colon F \to L_n^{2N}$ properly, we can assume $\mathcal{F} = i^*\mathcal{F}_{e^{2\pi i/n}}$. So we get a local system $\Sigma\mathcal{F} = \mathcal{F}_{e^{2\pi i/n}} \cup_{\mathcal{F}} \mathcal{F}_{e^{2\pi i/n}}$ on $\Sigma_n F$. From this we get the holonomy covering $\widetilde{\Sigma_n F} \xrightarrow{n} \Sigma_n F$; but we also have the holonomy covering $\widetilde{F} \xrightarrow{n} F$. These two covering spaces are related by the following lemma:

Lemma 7.3 *In the setting above, we have*

$$\pi_i \widetilde{\Sigma_n F} \cong \pi_i \Sigma \widetilde{F}$$

in low degrees i (smaller than $2N$).

Proof Let $\Sigma(N)\widetilde{F}$ be the suspension of $\widetilde{F}$ along S^{2N} (instead of S^∞). This forms an n–fold covering $\Sigma(N)\widetilde{F} \to \Sigma_n(N)F$, which must be homotopy equivalent to the universal covering of $\widetilde{\Sigma_n(N)F} \to \Sigma_n(N)F$. □

For the usual suspension, it is well known that $H_{k+1}(\Sigma F;\mathbb{R}) \cong H_k(F;\mathbb{R})$ for all $k \geq 1$. For the lens space suspension this becomes:

Lemma 7.4 *For every topological space F with local system inducing an n–fold holonomy covering, we have, for $k \geq 1$,*

$$H_{k+1}(\Sigma_n F;\mathbb{R}) \cong H_k(F;\mathbb{R}).$$

Proof Using the Mayer–Vietoris sequence for the defining push-out of the lens space suspension, we get:

$$\begin{aligned} \cdots \to H_{k+1}(L_n^{2N};\mathbb{R}) \oplus H_{k+1}(L_n^{2N};\mathbb{R}) \to H_{k+1}(\Sigma_n F;\mathbb{R}) \\ \to H_k(F;\mathbb{R}) \to H_k(L_n^{2N};\mathbb{R}) \oplus H_k(L_n^{2N};\mathbb{R}) \to \cdots \end{aligned}$$

The fact that L_n^{2N} is rationally homologically trivial now yields the desired isomorphism. □

Furthermore, we know for the usual suspension that $\pi_m^S(F) \otimes \mathbb{R} \cong \overline{H}_m(F;\mathbb{R})$, where $\pi_m^S(F) := \pi_m(\operatorname{colim}_k \Omega^k \Sigma^k F)$ denotes the stabilized homotopy group. This becomes:

Lemma 7.5 *If $k \in \mathbb{N}$ is large enough, and F is a space with local system inducing an n–fold holonomy covering, we have an isomorphism*

$$\pi_{m+k}(\Sigma_n^k F) \otimes \mathbb{R} \cong \overline{H}_{m+k}(\Sigma^k \widetilde{F};\mathbb{R})$$

for $m + k < N$.

Proof We get the n–fold holonomy covering $\widetilde{F} \to F$. Using Lemma 7.3 several times, we get in low degrees i

$$\pi_i \widetilde{\Sigma_n^k F} \cong \pi_i \Sigma(\widetilde{\Sigma_n^{k-1} F}) \cong \cdots \cong \pi_i \Sigma^k \widetilde{F}.$$

Thus we have, for $N > m + k > 1$,

$$\begin{aligned} \pi_{m+k}(\Sigma_n^k F) \otimes \mathbb{R} &\cong \pi_{m+k}(\Sigma^k \widetilde{F}) \otimes \mathbb{R} \\ &\cong \pi_m^S(\widetilde{F}) \otimes \mathbb{R} \quad \text{for } k \text{ large} \\ &\cong \overline{H}_m(\widetilde{F}; \mathbb{R}) \\ &\cong \overline{H}_{m+k}(\Sigma^k \widetilde{F}; \mathbb{R}). \end{aligned}$$

□

Remark 7.6 Although we require k to be large in the last lemma, it does not depend on N at all, meaning that we can still choose N to be much larger than k.

We will need the following definition and proposition:

Definition 7.7 A topological space F is called simple if $\pi_1 F$ is abelian and acts trivially on every $\pi_i F$ for $i \geq 2$.

Proposition 7.8 *Let F be a path connected, simple space and $\widetilde{F} \xrightarrow{n} F$ an n–fold Galois covering. Then the transition group $\mathbb{Z}/n$ will act trivially on $H_*(\widetilde{F}; \mathbb{R})$*

Proof Let $\{F^l\}$ be the Postnikov tower for F; that is a sequence of spaces with $\lim_l F^l \cong F$ and $\pi_i F^l \cong \pi_i F$ for $0 \leq i \leq l$ and $\pi_i F^l \cong 0$ for $i > l$. Since we have $\pi_1 F^l \cong \pi_1 F$ for every $l > 0$, we have n–fold coverings $\widetilde{F}^l \xrightarrow{n} F^l$. We will prove by induction that $\mathbb{Z}/n$ acts trivially on $H_*(\widetilde{F}^l; \mathbb{R})$. The sequence $\{\widetilde{F}^l\}$ will clearly provide a Postnikov tower for $\widetilde{F}$, and since the real homology of the stages of a Postnikov tower stabilizes in every degree, this will prove the proposition.

To start the induction we look at $F^1 \simeq K(\pi_1 F, 1)$, which will only have the first homotopy group $\pi_1 F^1 \cong \pi_1 F$. The covering $\widetilde{F} \xrightarrow{n} F$ gives a map $\alpha\colon \pi_1 F \to \mathbb{Z}/n$. Using this, we see that the covering $\widetilde{F}^1 \xrightarrow{n} F^1$ will be an Eilenberg–Mac Lane space:

$$\widetilde{F}^1 \simeq K(\ker \alpha, 1).$$

The group $\mathbb{Z}/n$ acts trivially on $\ker \alpha \subseteq \pi_1 F$ because $\pi_1 F$ is abelian, and therefore $\mathbb{Z}/n$ acts trivially on $\widetilde{F}^1 \simeq K(\ker \alpha, 1)$ and $H_*(\widetilde{F}^1; \mathbb{R})$. This starts the induction.

Now assume that $\mathbb{Z}/n$ acts trivially on $H_*(\widetilde{F}^{l-1}; \mathbb{R})$ with $l > 1$. We have the fibration

$$K(\pi_l \widetilde{F}, l) \to \widetilde{F}^l \to \widetilde{F}^{l-1}.$$

Since we know $\pi_l F \cong \pi_l \widetilde{F}$, the group $\mathbb{Z}/n$ will act trivially on $\pi_l \widetilde{F}$ and thereby also trivially on $K(\pi_l\widetilde{F},l)$ and $H_*(K(\pi_l\widetilde{F},l);\mathbb{R})$. By induction assumption it must also act trivially on

$$H_i(\widetilde{F}^{l-1};H_k(K(\pi_l(\widetilde{F}),l);\mathbb{R})),$$

and thereby it acts trivially on the whole Leray–Serre spectral sequence for the fibration $K(\pi_l\widetilde{F},l)\to\widetilde{F}^l\to\widetilde{F}^{l-1}$. From this it follows that $\mathbb{Z}/n$ acts unipotently on $H_*(\widetilde{F}_l;\mathbb{R})$, and since $\mathbb{R}[\mathbb{Z}/n]$ is semisimple, this includes that $\mathbb{Z}/n$ acts trivially on $H_*(\widetilde{F};\mathbb{R})$. □

From this we get the following important corollary.

Corollary 7.9 *If F is a simple topological space with local system inducing an n–fold holonomy covering $\widetilde{F}\xrightarrow{n}F$, then we have*

$$H_l(\Sigma(N)^k\widetilde{F};\mathbb{R})\cong H_l(\Sigma_n^kF;\mathbb{R})$$

for all $l<2N$. (Recall that $\Sigma(N)\widetilde{F}$ is the suspension of $\widetilde{F}$ along S^{2N} instead of S^∞.)

Proof Since F is simple, the group $\mathbb{Z}/n$ will act trivially on $H_*(\widetilde{F};\mathbb{R})$. It is well known that this implies

$$H_*(F;\mathbb{R})\cong H_*(\widetilde{F};\mathbb{R}).$$

The inclusion $S^{2N}\to S^\infty$ gives a map $\Sigma(N)\widetilde{F}\to\Sigma\widetilde{F}$ that is evidently $2N$–connected. By using Lemma 7.4 we see

$$H_*(\Sigma(N)^k\widetilde{F};\mathbb{R})\cong H_{*-k}(\widetilde{F};\mathbb{R})\cong H_{*-k}(F;\mathbb{R})\cong H_*(\Sigma_n^kF;\mathbb{R})$$

up to degree $2N$. □

Now we are turning back to bundles. For a fiber bundle $F\hookrightarrow E\to B$ with local system $\mathcal{F}$ on F inducing a finite cyclic n–fold holonomy covering, we get a fiberwise map $E\to B\times L_n^{2N}$ and can use this to define the fiberwise lens space suspension as the (homotopy) push-out:

$$\begin{array}{ccc} E & \longrightarrow & B\times L_n^{2N} \\ \downarrow & & \downarrow \\ B\times L_n^{2N} & \longrightarrow & \Sigma_{n,B}E \end{array}$$

It is easy to see that $\Sigma_{n,B}E\to B$ is a bundle with fiber Σ_nF and as before we get a local system $\Sigma\mathcal{F}$ on $\Sigma_{n,B}E$. We have the following lemma analogous to [7, Lemma 8.7]:

Lemma 7.10 *The bundle $\Sigma_{n,B}E$ is smoothable (ie fiber homotopy equivalent to a smooth bundle) if E is smoothable, and we have*

$$\tau^\delta(E,\mathcal{F})=-\tau^\delta(\Sigma_{n,B}E,\Sigma\mathcal{F}).$$

7.2 Reducing the homology of the fiber

We now attempt to make the fiber of a bundle $F \hookrightarrow E \to B$ with a local system on F, simply connected base B, and simple fiber F rationally homologically trivial without changing the difference torsion. This is the general strategy: Assume that m is the largest integer such that $H_m(F;\mathbb{R})$ is nontrivial. Picking an element $\alpha \in H_l(F;\mathbb{R})$, Lemma 7.12 asserts that we can find a representative of α

$$B \times L_n^{m+k} \to \Sigma_{n,B}^k E,$$

and then Lemma 7.11 uses this to create a new fiber bundle $E_1 \to B$ with the same difference torsion as E and the fiber F_1 having overall one dimension lower homology than F. Then Lemma 7.13 puts everything together inductively. The basic ideas reflect what has been done by Igusa in [7], yet the fact that we need to preserve the representation of a fundamental group poses some challenges. In the following, let N always be a sufficiently large integer.

Lemma 7.11 *Suppose $F \hookrightarrow E \to B$ is a fibration with local system $\mathcal{F}$ on F inducing a finite cyclic n–fold holonomy covering. Let $m \in \mathbb{N}$ denote the largest integer for which $\overline{H}_m(F;\mathbb{R}) \neq 0$. Suppose that we have $H_l(F;\mathbb{R}) \cong H_l(\widetilde{F};\mathbb{R})$ for $0 < l < m + \dim B$. Suppose further that m is odd and let α be a map*

$$\alpha\colon B \times L_n^m \to E$$

with the following properties: on each fiber we have $\alpha^*\mathcal{F} \cong \mathcal{F}_\zeta$ for some n^{th} root of unity ζ and $\alpha_*\colon \overline{H}_m(L_n^m;\mathbb{R}) \to \overline{H}_m(F;\mathbb{R})$ is nontrivial. Then if we look at the bundle

$$E_1 = E \cup_{B \times L_n^m} B \times L_n^{2N}$$

with fiber F_1 with local system $\mathcal{F}_1 := \mathcal{F} \cup_{\mathcal{F}_\zeta} \mathcal{F}_\zeta$ and corresponding covering $\widetilde{F}_1 \xrightarrow{n} F_1$, we have

$$\dim_{\mathbb{R}} H_*(F_1;\mathbb{R}) < \dim_{\mathbb{R}} H_*(F;\mathbb{R})$$

and

$$H_l(F_1;\mathbb{R}) \cong H_l(\widetilde{F}_1;\mathbb{R}) \quad \text{for } 0 < l < m + \dim B.$$

Proof Assume that we have a map $\alpha\colon B \times L_n^m \to E$ such that the induced map

$$\alpha_*\colon \overline{H}_m(L_n^m;\mathbb{R}) \to \overline{H}_m(F;\mathbb{R})$$

is nontrivial. Note this implies that the integer m is odd. Then the homology of the fiber $F_1 = F \cup_{L_n^m} L_n^{2N}$ will be given by the Mayer–Vietoris sequence as (where $i < 2N$)

$$H_i(L_N^m;\mathbb{R}) \xrightarrow{\alpha_*} H_i(F;\mathbb{R}) \oplus 0 \to H_i(F_1;\mathbb{R}) \to 0$$

and therefore we have $\dim_{\mathbb{R}} H_*(F_1;\mathbb{R}) < \dim_{\mathbb{R}} H_*(F;\mathbb{R})$. This also shows that $H_i(F_1;\mathbb{R}) \cong H_i(F;\mathbb{R})$ for $i \neq m$.

To show that this F_1 will satisfy the second property, we can use a similar sequence and show $H_i(\widetilde{F}_1;\mathbb{R}) \cong H_i(\widetilde{F};\mathbb{R})$. $\square$

Lemma 7.12 *Suppose $F \hookrightarrow E \to B$ is a fibration with simply connected base B and local system $\mathcal{F}$ on F inducing a finite cyclic n–fold holonomy covering. As before let $m \in \mathbb{N}$ denote the largest integer for which $\overline{H}_m(F;\mathbb{R}) \neq 0$ and suppose that we have $H_l(F;\mathbb{R}) \cong H_l(\widetilde{F};\mathbb{R})$ for $0 < l < m + \dim B$. Then there exists an integer $k \in \mathbb{N}$ and a map*

$$\alpha\colon B \times L_n^{m+k} \to \Sigma_{n,B}^k E$$

such that $\alpha^*\Sigma^k\mathcal{F} \cong \mathcal{F}_\zeta$ for some n^{th} root of unity ζ and $\alpha_*\colon \overline{H}_{m+k}(L_n^{m+k};\mathbb{R}) \to \overline{H}_{m+k}(\Sigma_n^k F;\mathbb{R})$ is nontrivial.

Proof Note that in the following, m and n are fixed, already determined integers, whereas k is an sufficiently large integer bounded by the sufficiently large integer N. Furthermore $m+k$ must be odd, such that L_n^{m+k} has a nonvanishing rational homology group in degree $m+k$, but we can choose k in such a way that this is satisfied.

Such a map α will correspond to a section s of the bundle

$$\mathrm{Map}(L_n^{m+k}, \Sigma_n^k F) \hookrightarrow \mathrm{Map}_B(B \times L_n^{m+k}, \Sigma_{n,B}^k E) \to B,$$

which is a homologically nontrivial map in each fiber. In this context the notation $\mathrm{Map}_B(B \times L_n^{m+k}, \Sigma_{n,b}^E)$ will always mean the space of fiberwise maps between $B \times L_n^{m+k}$ and $\Sigma_{n,B}^k E$. We will construct this section using obstruction theory. Let B_l denote the l–skeleton of B. Firstly, we will give $s_1\colon B_1 \to \mathrm{Map}_B(B \times L_n^{m+k}, \Sigma_{n,B}^k E)$. By the choice of m we have a nonzero element

$$\widetilde{\gamma} \in \overline{H}_{m+k}(\Sigma(N)^k\widetilde{F};\mathbb{R}) \cong \overline{H}_{m+k}(\Sigma_n^k F;\mathbb{R}) \cong \overline{H}_m(F;\mathbb{R}).$$

Since the reduced homology is isomorphic to rationalized stabilized homotopy, we can view $\widetilde{\gamma}$ as an element of $\pi_{m+k}(\Sigma(N)^k\widetilde{F}) \otimes \mathbb{R}$, if k is large enough. Now choose a representative $\widetilde{\alpha}_1\colon S^{m+k} \to \Sigma(N)^k\widetilde{F}$ of $\widetilde{\gamma}$. The map $\widetilde{\alpha}_1$ will clearly be nontrivial on homology.

Our goal is now to modify $\widetilde{\alpha}_1$ to $\widetilde{\alpha}\colon S^{m+k} \to \Sigma(N)^k\widetilde{F}$ such that it covers an $\alpha\colon L_n^{m+k} \to \Sigma_n^k F$. Since $H_{m+k}(\Sigma(N)^k\widetilde{F};\mathbb{R}) \cong H_{m+k}(\Sigma_n^k F;\mathbb{R})$, the map α will be nontrivial on homology. Furthermore the covering will ensure $\alpha^*\mathcal{F} \cong \mathcal{F}_\zeta$ for some n^{th} root of unity ζ. To begin, we have from the last lens space suspension an inclusion

$$i\colon L_n^{m+k} \hookrightarrow \Sigma_n^k F$$

trivial on homology. This will be covered by a homologically trivial equivariant inclusion

$$\widetilde{i}\colon S^{m+k} \hookrightarrow \widetilde{\Sigma_n^k F}\,.$$

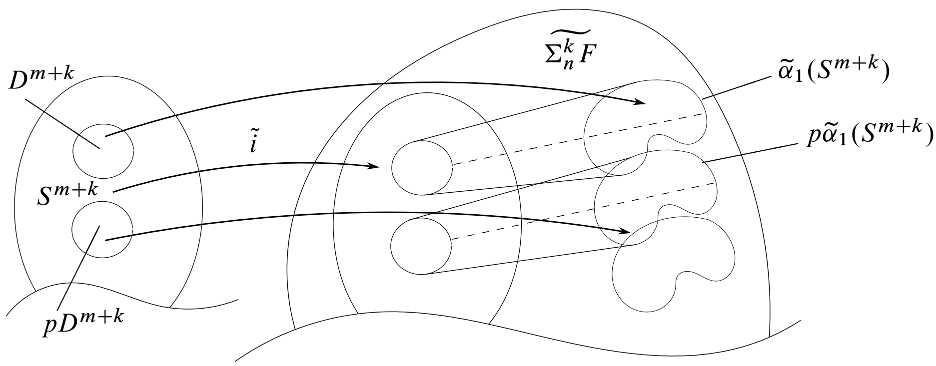


Figure 2: Modifying the inclusion $\tilde{i}\colon S^{m+k} \hookrightarrow \widetilde{\Sigma_n^k F}$

The idea now is to take a small disc D^{m+k} in $S^{m+k} \subseteq \widetilde{\Sigma_n^k F}$ and connect it to the image $\widetilde{\alpha}_1(S^{m+k})$. Then we can map S^{m+k} to this new image instead and this map will be nontrivial on homology because $\widetilde{\alpha}_1$ is nontrivial on homology. To make it equivariant we do the same construction equivariantly to every disc $p^i D^{m+k}$ in the orbit of D^{m+k} under the $\mathbb{Z}/n$ action on S^{m+k}. Here $p \in \mathbb{Z}/n$ denotes a generator. This is illustrated in Figure 2.

The formal construction is the following: Choose a small disc $D^{m+k} \subseteq S^{m+k}$. By doing this in a slightly bigger disc, we can modify the inclusion such that it factorizes

$$D^{m+k} \to * \hookrightarrow \widetilde{\Sigma_n^k F}.$$

Using $D^{m+k}/\partial D^{m+k} \simeq S^{m+k}$, we can glue in $\widetilde{\alpha}_1$ and modify the inclusion again so that it factorizes

$$D^{m+k} \xrightarrow{\widetilde{\alpha}_1} \widetilde{\Sigma_n^k F}.$$

Now let $p \in \mathbb{Z}/n$ be a generator. If we make D^{m+k} small enough, it will not intersect with any of the $p^i D^{m+k} \subseteq S^{m+k}$ for $0 < i < n$. Doing the same construction to every $p^i D^{m+k}$ using $p^i \widetilde{\alpha}_1$, we can modify the inclusion to a map

$$\widetilde{\alpha}\colon S^{m+k} \to \widetilde{\Sigma_n^k F},$$

which will clearly be n–equivariant and thus cover a map

$$\alpha\colon L_n^{m+k} \to \Sigma_n F.$$

The corresponding rationalized homotopy class of $\widetilde{\alpha}$ in $\pi_{m+k}(\Sigma_n^k F) \otimes \mathbb{R}$ is given by

$$[\widetilde{\alpha}] = [\widetilde{\alpha}_1] + p[\widetilde{\alpha}_1] + \cdots + p^{n-1}[\widetilde{\alpha}_1] = n[\widetilde{\alpha}_1] \neq 0,$$

since $\pi_1 F$ acts trivially on

$$\pi_{m+k}\widetilde{\Sigma_n^k F}\otimes\mathbb{R}\cong\pi_{m+k}(\Sigma(N)^k\widetilde{F})\otimes\mathbb{R}\cong H_{m+k}(\Sigma(N)^k\widetilde{F};\mathbb{R})$$
$$\cong H_m(\widetilde{F};\mathbb{R})\cong H_m(F;\mathbb{R})$$

(otherwise the map $H^m(F;\mathbb{R})\hookrightarrow H^m(\widetilde{F};\mathbb{R})$ would not be an isomorphism and thereby $H_m(F;\mathbb{R})$ would not be isomorphic to $H_m(\widetilde{F};\mathbb{R})$ either). So α will be nontrivial in rational homology.

With this we can define $s_0\colon B_0\simeq *\to \mathrm{Map}_B(B\times L_n^{m+k},\Sigma^k_{n,B}E)$ nontrivial in the homology of the fiber. Since B is simply connected, this section, defined over a point of B, can be extended to a section $s_1\colon B_1\to\mathrm{Map}_B(B\times L_n^{m+k},\Sigma^k_{n,B}E)$.

Let us now continue inductively. Suppose we already have a section $s_l\colon B_l\to \mathrm{Map}_B(B\times L_n^{m+k},\Sigma^k_{n,B}E)$ with $1\le l<\dim B$. By restriction, we will get sections

$$s_{l,i}\colon B_l\to\mathrm{Map}_B(B\times L_n^i,\Sigma^k_{n,B}E).$$

Let us first extend $s_{l,1}$ to $s_{l+1,1}\colon B_{l+1}\to\mathrm{Map}_B(B\times L_n^1,\Sigma_n^k E)$: This depends on the obstruction class

$$\theta(s_l,1)\in H^{l+1}(B,B_l;\pi_l(\mathrm{Map}(L_n^1,\Sigma_n^k F)))\cong H^{l+1}(B,B_l;\pi_{l+1}(\Sigma_n^k F)),$$

because $L_n^1\simeq S^1$. So $\theta(s_{l,1})$ is rationally trivial, if k is large enough (larger than $l+1$). This is enough to extend $s_{l,1}$ as Igusa showed in the nontwisted version [7, Lemma 8.9] of this lemma.

We now want to extend $s_{l+1,1}$ to $s_{l+1,2}$ relative to $s_{l,2}$. For this we look at the cofibration sequence

$$L_n^1\hookrightarrow L_n^2\to S^2,$$

which gives us the fibration sequence

$$\Omega^2(\Sigma_n^k F)\hookrightarrow\mathrm{Map}_B(B\times L_n^2,\Sigma^k_{n,B}E)\to\mathrm{Map}_B(B\times L_n^1,\Sigma^k_{n,B}E).$$

From this we get the commutative diagram

$$\begin{tikzcd}
 & \Omega^2(\Sigma_n^k F) \arrow[d, hook] \\
B_l \arrow[r, "s_{l,2}"] \arrow[d, hook] & \mathrm{Map}_B(B\times L_n^2,\Sigma^k_{n,B}E) \arrow[d] \\
B_{l+1} \arrow[r, "s_{l+1,1}"'] \arrow[ur, dotted, "s_{l+1,2}"] & \mathrm{Map}_B(B\times L_n^1,\Sigma^k_{n,B}E)
\end{tikzcd}$$

where the right column is a fibration sequence. Consequently the extension from $s_{l+1,1}$ to $s_{l+1,2}$ depends on the obstruction class

$$\theta(s_{l,1}) \in H^{l+1}(B, B_l; \pi_l(\Omega^2(\Sigma^k_n F))) \cong H^{l+1}(B, B_l; \pi_{l+2}(\Sigma^k_n F)),$$

which is, again, rationally trivial for large k.

Now assume that we have already constructed $s_{l+1,i}$ with $i \in \mathbb{N}$ even. Next, look at the cofibration

$$L^i_n \hookrightarrow L^{i+2}_n \to M(\mathbb{Z}_n, i),$$

where

$$M(\mathbb{Z}_n, i) := \mathrm{cof}(S^i \xrightarrow{n} S^i)$$

is the Moore space. Directly from the definition of the Moore space, we get that $\pi_l(\mathrm{Map}(M(\mathbb{Z}_n, i), X))$ is finite for any space X. Using the fibration

$$\mathrm{Map}(M(\mathbb{Z}_n, i), \Sigma^k_n F) \hookrightarrow \mathrm{Map}_B(B \times L^{i+2}_n, \Sigma^k_{n,B} E) \to \mathrm{Map}_B(B \times L^i_{n,B} \Sigma^k_n E),$$

the commutative diagram

$$\begin{tikzcd}
 & \mathrm{Map}(M(\mathbb{Z}_n, i), \Sigma^k_n F) \arrow[d, hook] \\
B_l \arrow[r, "s_{l,i+2}"] \arrow[d, hook] & \mathrm{Map}_B(B \times L^{i+2}_n, \Sigma^k_{n,B} E) \arrow[d] \\
B_{l+1} \arrow[r, "s_{l+1,i}"'] \arrow[ur, dotted, "s_{l+1,i+2}"] & \mathrm{Map}_B(B \times L^i_n, \Sigma^k_{n,B} E)
\end{tikzcd}$$

tells us that extending $s_{l+1,i}$ to $s_{l+1,i+2}$ depends on the obstruction class

$$\theta(s_{l+1,i}) \in H^{l+1}(B, B_l; \pi_l(\mathrm{Map}(M(\mathbb{Z}_n, i), \Sigma^k_n F))),$$

which is rationally trivial.

Using this inductively, we get $s_{l+1,k+m-1}$. To extend this to $s_{l+1,k+m} = s_{l+1}$, we use again the cofibration sequence

$$L^{k+m-1}_n \hookrightarrow L^{k+m}_N \to S^{k+m},$$

the induced fibration sequence

$$\Omega^{k+m}(\Sigma^k_n F) \hookrightarrow \mathrm{Map}_B(B \times L^{k+m}_n, \Sigma^k_{n,B} E) \to \mathrm{Map}_B(B \times L^{k+m-1}_n, \Sigma^k_{n,B} E)$$

and the commutative diagram

$$\begin{tikzcd}
 & \Omega^{k+m}(\Sigma_n^k F) \arrow[d, hook] \\
B_l \arrow[r, "s_{l,k+m}"] \arrow[d, hook] & \mathrm{Map}_B(B \times L_n^{k+m}, \Sigma_{n,B}^k E) \arrow[d] \\
B_{l+1} \arrow[r, "s_{l+1,k+m-1}"] \arrow[ur, dotted, "s_{l+1,k+m}"] & \mathrm{Map}_B(B \times L_n^{k+m-1}, \Sigma_{n,B}^k E)
\end{tikzcd}$$

making the obstruction class

$$\theta(s_{l+1,k+m-1}) \in H^{l+1}(B, B_l; \pi_{k+m+l}(\Sigma_n^k F)).$$

However, if k is large enough, we have

$$\begin{aligned}
\pi_{k+m+l}(\Sigma_n^k F) \otimes \mathbb{R} &\cong \pi_{k+m+l}(\Sigma(N)^k \widetilde{F}) \otimes \mathbb{R} \\
&\cong \overline{H}_{k+m+l}(\Sigma(N)^k \widetilde{F}; \mathbb{R}) \\
&\cong H_{k+m+l}(\Sigma_n^k F; \mathbb{R}) \\
&\cong H_{m+l}(F; \mathbb{R}) \cong 0
\end{aligned}$$

by assumption because $m + l < m + \dim B$. This guarantees that we can extend $s_{l+1,k+m-1}$ to s_{l+1} and completes the proof. □

Lemma 7.13 *Let $F \hookrightarrow E \to B$ be a fibration with simply connected base B and local system $\mathcal{F}$ on F inducing a finite cyclic n–fold holonomy covering. Suppose further that F is simple. Then there exists a bundle $F' \hookrightarrow E' \to B$ with local coefficient system $\mathcal{F}'$ on F', where F' is rationally homologically trivial such that*

$$\tau^\delta(E, \mathcal{F}) = \pm \tau^\delta(E', \mathcal{F}').$$

Proof Let again m be the largest integer such that $H_m(F; \mathbb{R})$ is nontrivial. Since F is simple we get $H_*(F; \mathbb{R}) \cong H_*(\widetilde{F}; \mathbb{R})$ by Corollary 7.9, and we can use Lemma 7.12 to get, for an integer k, a bundle map

$$\alpha\colon B \times L_n^{m+k} \to \Sigma_{n,B}^k E$$

nontrivial on the $(m+k)^{\text{th}}$ homology. By Lemma 7.3 the n–fold covering of $\Sigma_n^k F$ is given in low degrees by $\Sigma(N)^k \widetilde{F}$. Since both Σ_n^k and $\Sigma(N)^k$ only shift rational homology up by k degrees we have

$$H_l(\Sigma_n^k F; \mathbb{R}) \cong H_l(\Sigma(N)^k \widetilde{F}; R) \cong H_l(\widetilde{\Sigma_n^k F}; \mathbb{R})$$

for all $0 < l < m + k + \dim B$. Furthermore the highest nontrivial homology group of $\Sigma_n^k F$ is in degree $m+k$ and we also have

$$\dim_{\mathbb{R}} H_*(\Sigma_n^k F;\mathbb{R}) = \dim_{\mathbb{R}} H_*(F;\mathbb{R}).$$

Now we can apply the construction of Lemma 7.11 to get a bundle $F_1 \hookrightarrow E_1 \to B$ such that

$$\dim_{\mathbb{R}} H_*(F_1,\mathbb{R}) < \dim_{\mathbb{R}} H_*(\Sigma_n^k F;\mathbb{R}) = \dim_{\mathbb{R}} H_*(F;\mathbb{R}).$$

By definition of E_1, and since the torsion of trivial bundles is zero, we get with additivity and Lemma 7.10

$$\tau^\delta(E_1,\mathcal{F}_1) = \tau^\delta(\Sigma_{n,B}^k E, \Sigma^k\mathcal{F}) = (-1)^k \tau^\delta(E,\mathcal{F}).$$

Since Lemma 7.11 guarantees that $H_l(F_1;\mathbb{R}) \cong H_l(\widetilde{F}_1;\mathbb{R})$ for $0 < l < m+k+\dim B$ we now can repeat this process and decrease the dimension of the rational homology until we will get the bundle $F' \hookrightarrow E' \to B$ with local system $\mathcal{F}'$ on F such that

$$\tau^\delta(E,\mathcal{F}) = \pm\tau^\delta(E',\mathcal{F}')$$

and F' is rationally homologically trivial. □

We are now finally in the position to prove the main theorem. As a consequence of Lemma 7.13 it suffices to only determine τ^δ on bundles with rationally trivial fiber, so we conclude with the following lemma.

Lemma 7.14 *We have $\tau^\delta(Z,\mathcal{F}) = 0$ for any torsion invariant, smoothable bundle $X \hookrightarrow Z \to B$ with $\overline{H}_*(X;\mathbb{R}) = 0$, simply connected base B and local system $\mathcal{F}$ inducing an n–fold holonomy covering.*

Proof This is completely analogous to the proof of Lemma 8.11 in [7]. We will only explain the main points. We replace the bundle by a manifold bundle $M \hookrightarrow E \to B$ and its universal covering $\widetilde{M} \to \widetilde{E} \to B$. Choosing a section of $E \to B$ gives disc bundles $D \subset E \to B$ and $\widetilde{D} \subset \widetilde{E} \to B$. Now there is a universal torsion class

$$\tau^\delta \in H^{2k}(B\mathrm{Diff}_n(\widetilde{M} \operatorname{rel} \widetilde{D});\mathbb{R}),$$

where $B\mathrm{Diff}_n(\widetilde{M} \operatorname{rel} \widetilde{D})$ is the classifying space of $\mathbb{Z}/n$–equivariant diffeomorphisms of $\widetilde{M}$ relative to $\widetilde{D}$.

In the original paper [7], one only has to consider $B\mathrm{Diff}(M \operatorname{rel} D)$, but luckily in the end we are only interested in maps that leave a certain base point fixed and we have

$$B\mathrm{Diff}_{0,n}(\widetilde{M} \operatorname{rel} \widetilde{D}) \cong B\mathrm{Diff}_0(M \operatorname{rel} D),$$

where the subscript 0 indicates the identity component. From here on the proof is parallel to the proof in [7]. □

References

[1] **J-M Bismut**, **S Goette**, *Families torsion and Morse functions*, Astérisque 275, Soc. Math. France, Paris (2001) MR

[2] **J-M Bismut**, **J Lott**, *Flat vector bundles, direct images and higher real analytic torsion*, J. Amer. Math. Soc. 8 (1995) 291–363 MR

[3] **W Dwyer**, **M Weiss**, **B Williams**, *A parametrized index theorem for the algebraic K–theory Euler class*, Acta Math. 190 (2003) 1–104 MR

[4] **T Goodwillie**, **K Igusa**, **C Ohrt**, *An equivariant version of Hatcher's G/O construction*, J. Topol. 8 (2015) 675–690 MR

[5] **K Igusa**, *Higher Franz–Reidemeister torsion*, AMS/IP Studies in Advanced Mathematics 31, Amer. Math. Soc., Providence, RI (2002) MR

[6] **K Igusa**, *Higher complex torsion and the framing principle*, Mem. Amer. Math. Soc. 835, Amer. Math. Soc., Providence, RI (2005) MR

[7] **K Igusa**, *Axioms for higher torsion invariants of smooth bundles*, J. Topol. 1 (2008) 159–186 MR

[8] **K Igusa**, *Outline of higher Igusa–Klein torsion*, lecture notes (2009) Available at `http://people.brandeis.edu/~igusa/Papers/AIM0911.pdf`

[9] **K Igusa**, **J Klein**, *The Borel regulator map on pictures, II: An example from Morse theory*, K–Theory 7 (1993) 225–267 MR

[10] **J Milnor**, *On polylogarithms, Hurwitz zeta functions, and the Kubert identities*, Enseign. Math. 29 (1983) 281–322 MR

[11] **T Szamuely**, *Galois groups and fundamental groups*, Cambridge Studies in Advanced Mathematics 117, Cambridge Univ. Press (2009) MR

[12] **J B Wagoner**, *Diffeomorphisms, K_2, and analytic torsion*, from "Algebraic and geometric topology, 1" (R J Milgram, editor), Proc. Sympos. Pure Math. XXXII, Amer. Math. Soc., Providence, RI (1978) 23–33 MR

Department of Mathematics, University of California
Los Angeles, CA, United States

Current address: *Department of Mathematics, Stanford University*
Stanford, CA, United States

`cohrt@stanford.edu`

Received: 28 September 2016 Revised: 24 March 2017

Algebraic & Geometric Topology 17 (2017) 3703–3749

Super q–Howe duality and web categories

DANIEL TUBBENHAUER
PEDRO VAZ
PAUL WEDRICH

We use super q–Howe duality to provide diagrammatic presentations of an idempotented form of the Hecke algebra and of categories of $\mathfrak{gl}_N$–modules (and, more generally, $\mathfrak{gl}_{N|M}$–modules) whose objects are tensor generated by exterior and symmetric powers of the vector representations. As an application, we give a representation-theoretic explanation and a diagrammatic version of a known symmetry of colored HOMFLY–PT polynomials.

57M25, 81R50

1 Introduction

Let $U_q(\mathfrak{gl}_N)$ be the quantum enveloping $\mathbb{C}_q = \mathbb{C}(q)$–algebra for $\mathfrak{gl}_N$ with q being generic. Let $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ denote the braided monoidal category of $U_q(\mathfrak{gl}_N)$–modules[1] tensor generated by *exterior* $\bigwedge_q^k \mathbb{C}_q^N$ *and symmetric* $\mathrm{Sym}_q^l \mathbb{C}_q^N$ powers and $U_q(\mathfrak{gl}_N)$–intertwiners between them.

We denote by $\check{H}$ an *idempotented version* of the direct sum of all Iwahori–Hecke algebras $H_\infty(q) = \bigoplus_{K \in \mathbb{Z}_{\geq 0}} H_K(q)$ of type A. Roughly, $\check{H}$ is the category obtained from the one-object category $H_\infty(q)$ by adding formal Gyoja–Aiston idempotents corresponding to column and row Young diagrams as new objects.[2] By *quantum Schur–Weyl duality*, the categories $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ are quotients of $\check{H}$ and the added idempotents can be thought of as lifts of the exterior $\bigwedge_q^k \mathbb{C}_q^N$ and the symmetric $\mathrm{Sym}_q^l \mathbb{C}_q^N$ powers.

We construct diagrammatic presentations of $\check{H}$ and $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ by using the *green–red* web categories ∞–$\mathbf{Web}_{\mathrm{gr}}$ and N–$\mathbf{Web}_{\mathrm{gr}}$. Morphisms in these $\mathbb{C}_q$–linear categories are combinations of planar, upward-oriented, trivalent graphs with edges labeled by positive integers and colored black, green or red[3] modulo local relations. Objects are

[1] We only consider finite-dimensional, left modules (of type 1) throughout the paper.

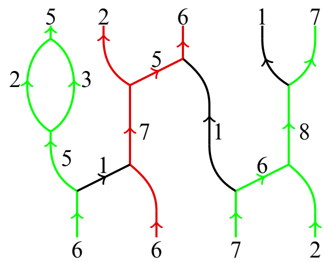
[2] Adding only column idempotents, one obtains the type A Schur algebroids introduced by Williamson in [30].

[3] We use colored diagrams in this paper. The colors (black, green and red) are important and we recommend to read the paper in color. If the reader has a black-and-white version, then green will appear lightly shaded and black and red can be distinguished since black edges are always labeled 1.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3703

boundaries of such green–red webs, ie finite sequences of positive integers, each of which additionally carries the color black, green or red, indicated either by an actual coloring or by a subscript.

An example of a green–red web is:



A green integer k in a boundary sequence is meant to correspond to the $\boldsymbol{U}_q(\mathfrak{gl}_N)$–module $\bigwedge_q^k \mathbb{C}_q^N$, a red integer l to $\mathrm{Sym}_q^l \mathbb{C}_q^N$, and sequences of integers correspond to tensor products of such. Vertical edges are identities on these $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules and trivalent vertices encode more interesting $\boldsymbol{U}_q(\mathfrak{gl}_N)$–intertwiners. The integer 1 should be $\mathbb{C}_q^N \cong \bigwedge_q^1 \mathbb{C}_q^N \cong \mathrm{Sym}_q^1 \mathbb{C}_q^N$ independent of the color green or red, so we color it black.

Our main result is:

Theorem (The diagrammatic presentation) *The additive closures of* ∞–$\mathbf{Web}_{\mathrm{gr}}$ *and of* N–$\mathbf{Web}_{\mathrm{gr}}$ *are braided monoidally equivalent to* $\check{\boldsymbol{H}}$ *and* $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$*, respectively.*

We will see that ∞–$\mathbf{Web}_{\mathrm{gr}}$ admits an involution interchanging the colors green and red. An almost direct consequence of this is a symmetry between the HOMFLY–PT polynomial $\mathcal{P}^{a,q}(\cdot)$ of a link $\mathcal{L}$ colored with $\vec{\lambda} = (\lambda^1, \dots, \lambda^d)$ and the HOMFLY–PT polynomial of $\mathcal{L}$ colored with $\vec{\lambda}^{\mathrm{T}} = ((\lambda^1)^{\mathrm{T}}, \dots, (\lambda^d)^{\mathrm{T}})$:

Proposition (The colored HOMFLY–PT symmetry) *We have*

$$\mathcal{P}^{a,q}(\mathcal{L}(\vec{\lambda})) = (-1)^c \mathcal{P}^{a,q^{-1}}(\mathcal{L}(\vec{\lambda}^{\mathrm{T}})). \tag{1-1}$$

Here c *is the sum of the number of nodes in the Young diagrams* λ^i *for* $1 \leq i \leq d$*.*

Our results might help to understand symmetries observed within the homologies that categorify the colored HOMFLY–PT polynomials; see Gukov and Stošić [10, Section 5].

Moreover, we show that a straightforward generalization of our approach also leads to diagrammatic presentations for categories $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$ of $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$–modules tensor generated by exterior and symmetric powers of the vector representation. The presentations are given by quotients $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ of ∞–$\mathbf{Web}_{\mathrm{gr}}$, which are obtained by killing Gyoja–Aiston idempotents corresponding to box-shaped Young diagrams.

1.1 The framework

A prototypical diagrammatic presentation result (with roots in the work of Rumer, Teller and Weyl [26]) states that the *Temperley–Lieb category* gives a presentation of the full subcategory of $\boldsymbol{U}_q(\mathfrak{sl}_2)$–modules tensor generated by the vector representation $\mathbb{C}_q^2$. Kuperberg [15] extended this to all rank-2 Lie algebras. In particular, he described a presentation of the full subcategory of $\boldsymbol{U}_q(\mathfrak{sl}_3)$–modules tensor generated by the *exterior powers* $\bigwedge_q^1 \mathbb{C}_q^3 \cong \mathbb{C}_q^3$ and $\bigwedge_q^2 \mathbb{C}_q^3$. More generally, Cautis, Kamnitzer and Morrison [3] gave a presentation of $\mathfrak{gl}_N$–$\mathbf{Mod}_e$, the full subcategory of $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules tensor generated by the *exterior powers* $\bigwedge_q^k \mathbb{C}_q^N$ for $k = 0, \dots, N$.

One of their key ideas in [3] is the usage of *skew quantum Howe duality* (or, short, *skew q–Howe duality*). In order to explain their approach, let $\vec{k} \in \mathbb{Z}_{\geq 0}^m$ be such that $k_1 + \cdots + k_m = K$. By skew q–Howe duality, the commuting actions of $\boldsymbol{U}_q(\mathfrak{gl}_m)$ and $\boldsymbol{U}_q(\mathfrak{gl}_N)$ on

$$\textstyle\bigwedge_q^K (\mathbb{C}_q^m \otimes \mathbb{C}_q^N) \cong \displaystyle\bigoplus_{\vec{k} \in \mathbb{Z}_{\geq 0}^m} \textstyle\bigwedge_q^{k_1} \mathbb{C}_q^N \otimes \cdots \otimes \bigwedge_q^{k_m} \mathbb{C}_q^N$$

give rise to a functor $\Phi_{\mathrm{skew}}^m \colon \dot{\boldsymbol{U}}_q(\mathfrak{gl}_m) \to \mathfrak{gl}_N$–$\mathbf{Mod}_e$, where $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_m)$ is the idempotented form of $\boldsymbol{U}_q(\mathfrak{gl}_m)$. Then Cautis, Kamnitzer and Morrison construct a commutative diagram, which takes the following form in our notation:[4]

$$\begin{tikzcd} \dot{\boldsymbol{U}}_q(\mathfrak{gl}_m) \arrow[r, "\Phi_{\mathrm{skew}}^m"] \arrow[dr, "\Upsilon_{\mathrm{skew}}^m"'] & \mathfrak{gl}_N\text{–}\mathbf{Mod}_e \\ & N\text{–}\mathbf{Web}_{\mathrm{g}} \arrow[u, "\Gamma"'] \end{tikzcd} \tag{1-2}$$

Here $\Upsilon_{\mathrm{skew}}^m$ is a certain *ladder functor* realizing an action of $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_m)$ on the diagram category N–$\mathbf{Web}_{\mathrm{g}}$. The *presentation functor* Γ is constructed so that (1-2) commutes. The functor Φ_{skew}^m is full and its kernel is generated by killing $\mathfrak{gl}_m$–weights with entries not in $\{0, \dots, N\}$. That Γ is an equivalence follows since N–$\mathbf{Web}_{\mathrm{g}}$ is defined to be the quotient of a "free" web category by relations coming from $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_m)$ (to make the ladder functor $\Upsilon_{\mathrm{skew}}^m$ well-defined) and the $\Upsilon_{\mathrm{skew}}^m$ image of the kernel of Φ_{skew}^m. $\mathfrak{sl}_N$–$\mathbf{Mod}_e$ can be recovered by identifying $\bigwedge_q^k \mathbb{C}_q^N \cong (\bigwedge_q^{N-k} \mathbb{C}_q^N)^*$ as $\boldsymbol{U}_q(\mathfrak{sl}_N)$–modules.

Rose and the first-named author [25] studied the situation of *symmetric quantum Howe duality* (for short, symmetric q–Howe duality).[5] That is, there is an analogue of (1-2) where $\mathfrak{gl}_N$–$\mathbf{Mod}_e$ is replaced by $\mathfrak{gl}_N$–$\mathbf{Mod}_s$, the full subcategory of $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules tensor generated by the *symmetric powers* $\mathrm{Sym}_q^l \mathbb{C}_q^N$ for $l \in \mathbb{Z}_{\geq 0}$. In the $N = 2$ case,

[4] We consider $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ instead of $\mathfrak{sl}_N$–$\mathbf{Mod}_{\mathrm{es}}$; see also Remark 1.1.

[5] In fact, the observations made in [25] were one of the main motivations to start this project.

the kernel of Φ^m_{sym} is generated by killing $\mathfrak{gl}_m$–weights with negative entries and one additional *dumbbell relation*, which encodes the relation $\mathbb{C}^2_q \otimes \mathbb{C}^2_q \cong \mathbb{C}_q \oplus \mathrm{Sym}^2_q\mathbb{C}^2_q$ in $\mathfrak{gl}_2$–$\mathbf{Mod}_s$. A direct generalization for $N > 2$ would require additional complicated relations besides killing $\mathfrak{gl}_m$–weights.

In this paper we give a diagrammatic presentation of the category $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$, the full subcategory of $U_q(\mathfrak{gl}_N)$–modules tensor generated by both exterior *and* symmetric powers of the vector representation. This diagrammatic presentation gives a common generalization of the web categories of [3] (only black–green webs) and [25] (only black–red webs). We see Cautis, Kamnitzer and Morrison's approach as a *machine* that takes dualities and produces diagrammatic presentations of the related representation-theoretical categories. Specifically, we start with *super quantum Howe duality* (for short, super q–Howe duality) between the superalgebra $U_q(\mathfrak{gl}_{m|n})$ and $U_q(\mathfrak{gl}_N)$. We obtain a full super q–Howe functor $\Phi^{m|n}_{\mathrm{su}}$, which we attempt to factor as a composite of a ladder functor $\Upsilon^{m|n}_{\mathrm{su}}$ — mapping into an appropriate web category — and a diagrammatic presentation functor Γ_N, to give an analogue of the commutative diagram (1-2):[6]

$$\begin{tikzcd}
\dot{U}_q(\mathfrak{gl}_{m|n}) \arrow[r, "\Phi^{m|n}_{\mathrm{su}}"] \arrow[dr, "\Upsilon^{m|n}_{\mathrm{su}}"'] & \mathfrak{gl}_N\text{–}\mathbf{Mod}^{\mathrm{sort}}_{\mathrm{es}} \\
& N\text{–}\mathbf{Web}^{\mathrm{sort}}_{\mathrm{gr}} \arrow[u, "\Gamma^{\mathrm{sort}}_N"']
\end{tikzcd}$$

Having decided to follow this strategy, the definition of the appropriate web category is already determined. Two aspects are important:

(I) In order to make $\Upsilon^{m|n}_{\mathrm{su}}$ well-defined, the web category needs to satisfy ladder images of $\dot{U}_q(\mathfrak{gl}_{m|n})$ relations. Remarkably, it suffices to consider relations coming from the subalgebra $\dot{U}_q(\mathfrak{gl}_m) \oplus \dot{U}_q(\mathfrak{gl}_n)$ and only one additional *super commutation* relation $[2]1_{\vec{k}} = F_m E_m 1_{\vec{k}} + E_m F_m 1_{\vec{k}}$ for $\mathfrak{gl}_{m|n}$–weights with $k_m = k_{m+1} = 1$. This corresponds to the *dumbbell relation* on webs and to $\mathbb{C}^N_q \otimes \mathbb{C}^N_q \cong \textstyle\bigwedge^2_q\mathbb{C}^N_q \oplus \mathrm{Sym}^2_q\mathbb{C}^N_q$ in $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$.

(II) In order to make the diagrammatic presentation functor an equivalence, we need to impose the ladder image of $\ker(\Phi^{m|n}_{\mathrm{su}})$ as relations in the web category. In fact, $\ker(\Phi^{m|n}_{\mathrm{su}})$ is spanned by idempotents corresponding to $\mathfrak{gl}_{m|n}$–weights $\vec{k} = (k_1, \dots, k_{m+n})$ with $k_1, \dots, k_m \notin \{0, \dots, N\}$ or $k_{m+1}, \dots, k_{m+n} \notin \mathbb{Z}_{\geq 0}$. It is remarkable that no extra relations, aside from killing these $\mathfrak{gl}_{m|n}$–weights, are necessary.

[6]Here the superscript "sort" indicates subcategories in which exterior powers are *sorted* to the left of symmetric powers in tensor products. This small technical restriction stems from the use of super q–Howe duality, but will be removed later on.

We impose the ladder images of $\ker(\Phi_{\mathrm{su}}^{m|n})$ in two steps: first we kill all $\mathfrak{gl}_{m|n}$–weights with negative entries by allowing only nonnegative labels on web edges. This produces the web category ∞–$\mathbf{Web}_{\mathrm{gr}}$, which is symmetric under exchanging green and red. On this we further quotient by setting $\mathfrak{gl}_{m|n}$–weights $\vec{k} = (k_1, \dots, k_{m+n})$ to zero if one of $k_1, \dots, k_m$ is greater than N. This produces the web category N–$\mathbf{Web}_{\mathrm{gr}}$ and in Theorem 3.20 we show that its additive closure is equivalent to $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$. Note that, although our graphical calculus is finer than the one in [3] in the sense that it contains more objects, the Karoubi envelopes of these diagrammatic categories agree for each N.

In Theorem 3.22 we use *quantum Schur–Weyl duality* to derive from Theorem 3.20 that ∞–$\mathbf{Web}_{\mathrm{gr}}$ gives a diagrammatic presentation of the idempotented Iwahori–Hecke algebra $\check{\boldsymbol{H}}$ from above.

Remark 1.1 We describe $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ and not $\mathfrak{sl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ because of the algebraic form of super q–Howe duality. In particular, our web categories do not contain duality isomorphisms $\bigwedge_q^k \mathbb{C}_q^N \cong (\bigwedge_q^{N-k} \mathbb{C}_q^N)^*$, which would be necessary for a diagrammatic presentation of $\mathfrak{sl}_N$–$\mathbf{Mod}_{\mathrm{es}}$. In $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$, on the other hand, there are no such hidden duals, as we have $\bigwedge_q^k \mathbb{C}_q^N \cong \bigwedge_q^N \mathbb{C}_q^N \otimes (\bigwedge_q^{N-k} \mathbb{C}_q^N)^*$ as $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules. Here $\bigwedge_q^N \mathbb{C}_q^N \cong L((1, \dots, 1))$ is the $\boldsymbol{U}_q(\mathfrak{gl}_N)$–module of highest weight $\lambda = (1, \dots, 1) \in \mathbb{Z}_{\geq 0}^N$.

Last, but not least, we use the more general *super q-Howe duality between $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ and $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$* to describe $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$. Feeding this duality into the "diagrammatic presentation machine" shows that this representation category is equivalent to the quotient $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ of ∞–$\mathbf{Web}_{\mathrm{gr}}$, which is obtained by killing the Gyoja–Aiston idempotent corresponding to the size $(N+1) \times (M+1)$ box-shaped Young diagram. This is a generalization, since, for $M = 0$, $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$ is equivalent to $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ and $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ is equal to N–$\mathbf{Web}_{\mathrm{gr}}$, because the box idempotent corresponds exactly to an $(N+1)$–labeled green edge.

This generalizes Grant's [9] and Sartori's [28] presentations of the category $\mathfrak{gl}_{1|1}$–$\mathbf{Mod}_{\mathrm{e}}$, and the diagrammatic calculus for $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{e}}$ given by Queffelec and Sartori [23] (see also Grant [8]). Compared to the latter, our generalization, which also takes the symmetric powers of $\mathbb{C}_q^{N|M}$ into account, does not need any extra relations aside from the dumbbell relation. In fact, the one extra relation needed to make the diagrammatic calculus given in [23] faithful — see [23, Remark 6.19] — has a very compact and natural description in our green–red web category $N|M$–$\mathbf{Web}_{\mathrm{gr}}$.

Finally, we sketch how our presentation of $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$ extends to take duals of exterior and symmetric powers into account. This closely follows [23, Section 6]. The resulting diagrammatic category allows the computation of the colored Reshetikhin–Turaev $\mathfrak{gl}_{N|M}$–link invariants. In Corollary 5.13, we interpret the colored HOMFLY–PT symmetry (1-1) as a stable version of a symmetry between colored Reshetikhin–Turaev $\mathfrak{gl}_{N|M}$– and $\mathfrak{gl}_{M|N}$–link invariants.

1.2 Outline of the paper

Section 2 is the diagrammatic heart of our paper, where we introduce ∞–$\mathbf{Web}_{\mathrm{gr}}$ and its subquotients N–$\mathbf{Web}_{\mathrm{gr}}$, N–$\mathbf{Web}_{\mathrm{g}}$ and N–$\mathbf{Web}_{\mathrm{r}}$.

Section 3 contains the proof of our main theorems and splits into three subsections: We first introduce super q–Howe duality. Then we show an equivalence between "sorted" subcategories of N–$\mathbf{Web}_{\mathrm{gr}}$ and $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$. These subcategories are induced by the algebraic form of super q–Howe duality. By using the "sorted" equivalence and the fact that the braiding gives a way to "shuffle" the "sorted" subcategories, we prove our main theorems.

In Section 4 we discuss one application of our diagrammatic presentation: we give a procedure to recover the colored HOMFLY–PT polynomial from ∞–$\mathbf{Web}_{\mathrm{gr}}$. A direct consequence of the green–red symmetry is a symmetry within the colored HOMFLY–PT polynomial obtained by transposing Young diagrams, see (1-1). The colored Reshetikhin–Turaev $\mathfrak{sl}_N$–link polynomials can be recovered from our approach as well, as we sketch in the last subsection.

Finally, in Section 5 we generalize the diagrammatic presentation of $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ to the super case $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$, and we sketch an extension of our diagrammatic calculus to include dual representations. The required arguments are — mutatis mutandis — contained in the previous sections and in [23, Section 6], which allows a very compact exposition in Section 5.

Acknowledgements We especially thank Antonio Sartori for a careful reading of a draft version of this paper and many helpful comments, and David Rose — some of the ideas underlying this paper came up in the joint work between him and DT. We also thank Jonathan Brundan, Jonathan Grant, Jonathan Kujawa, Marco Mackaay, Weiwei Pan, Jake Rasmussen, Marko Stošić, Catharina Stroppel, Geordie Williamson and Oded Yacobi for helpful discussion, comments, and probing questions, and two referees for some further helpful comments. We also like to thank Skype for many useful conversations.

DT was supported by a research funding of the "Deutsche Forschungsgemeinschaft (DFG)" during the main part of this work. DT and PW thank the center of excellence grant "Centre for Quantum Geometry of Moduli Spaces (QGM)" from the "Danish National Research Foundation (DNRF)" for sponsoring a research visit which started this project. PV was financially supported by the "Université catholique de Louvain, Fonds Spéciaux de Recherche (FSR) 12 J A". PW was supported by an EPSRC doctoral training grant.

2 The diagrammatic categories

In the present section we introduce the category ∞–$\mathbf{Web}_{\mathrm{gr}}$ and its quotient N–$\mathbf{Web}_{\mathrm{gr}}$. These provide diagrammatic presentations of $\check{\boldsymbol{H}}$ and its quotient categories $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ respectively. Other subquotients of ∞–$\mathbf{Web}_{\mathrm{gr}}$ are N–$\mathbf{Web}_{\mathrm{g}}$ and N–$\mathbf{Web}_{\mathrm{r}}$ (and later in Section 5, $N|M$–$\mathbf{Web}_{\mathrm{gr}}$) which are related to categories studied in [3] and [25], respectively.

2.1 Definition of the category ∞–$\mathbf{Web}_{\mathrm{gr}}$ and its subquotients

We first introduce the *free green–red web category* ∞–$\mathbf{Web}^f_{\mathrm{gr}}$. To this end, we denote by X the set

$$X = X_b \cup X_g \cup X_r = \{0_b, 1_b\} \cup \{2_g, 3_g, \ldots\} \cup \{2_r, 3_r, \ldots\},$$

where we think of the elements of X_b as being colored *black*, of the elements of X_g as being colored *green* and of the elements of X_r as being colored *red*. We usually omit the subscripts, since the colors on the boundary can be read off from the diagrams.

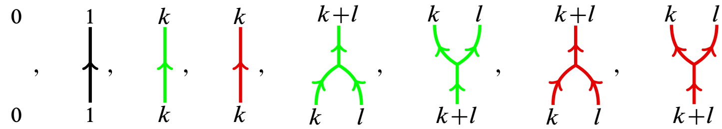
Definition 2.1 The *free green–red web category*, which we denote by ∞–$\mathbf{Web}^f_{\mathrm{gr}}$, is the category determined by the following data:

- The objects of ∞–$\mathbf{Web}^f_{\mathrm{gr}}$ are finite (possibly empty) sequences $\vec{k} \in X^L$ with entries from X for some $L \in \mathbb{Z}_{\geq 0}$, together with a zero object. We display the entries of $\vec{k}$ ordered from left to right according to their appearance in $\vec{k}$.

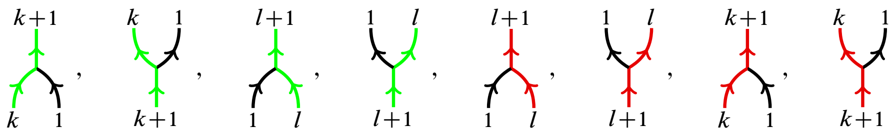
- The morphism space $\mathrm{Hom}_{\infty\text{–}\mathbf{Web}^f_{\mathrm{gr}}}(\vec{k}, \vec{l})$ from $\vec{k}$ to $\vec{l}$ is the $\mathbb{C}_q$–vector space spanned by isotopy classes[7] of planar, upward-oriented, trivalent graphs with edges labeled by positive integers and colored black, green or red, with bottom boundary $\vec{k}$ and top boundary $\vec{l}$. More precisely, we only allow webs that can be obtained by composition $\circ$ (vertical gluing) and taking the monoidal product $\otimes$ (horizontal juxtaposition) of the following basic pieces (including the empty diagram).

[7] We require that isotopies preserve the upward orientations and the boundary of green–red webs.

Let $k,l \in \mathbb{Z}_{\geq 2}$; then the generators are

(2-1)


called (from left to right) *empty identity*, *black identity*, *green identity*, *red identity*, *green merge*, *green split*, *red merge* and *red split*, together with (here $k,l \in \mathbb{Z}_{\geq 0}$)

(2-2)


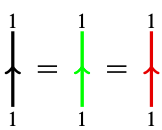
called *mixed merges* and *mixed splits,* respectively. (We also include versions of these involving edges labeled 0, which we, as in (2-1), do not illustrate.)

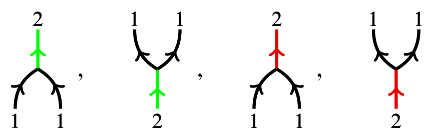
We call webs obtained by composition of generators with only black and green edges or only black and red edges *monochromatic*; cf (2-3). $\diamond$

Remark 2.2 Note the following conventions and properties of ∞–$\mathbf{Web}^{f}_{\mathrm{gr}}$:

- The category is $\mathbb{C}_q$–linear, ie the spaces $\mathrm{Hom}_{\infty\text{–}\mathbf{Web}^{f}_{\mathrm{gr}}}(\vec{k},\vec{l})$ are $\mathbb{C}_q$–vector spaces and the composition $\circ$ is $\mathbb{C}_q$–bilinear. Moreover, the category is monoidal by juxtaposition $\otimes$ of objects and morphisms. $\otimes$ is also $\mathbb{C}_q$–bilinear on morphism spaces.

- It is sometimes convenient in illustrations to allow green and red edges with label 1. By convention, these edges are to be read as being black:

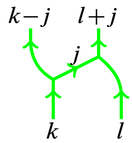
(2-3)


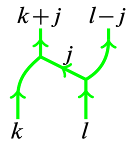
and


For example, the diagrams on the right are obtained by setting $k=1$ or $l=1$ in (2-2).

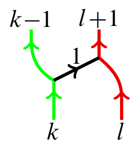
- The reading conventions for all webs are from *bottom to top* and *left to right*: if u and v are webs, then $v \circ u$ is obtained by gluing v on top of u and $u \otimes v$ is given by putting v to the right of u. Moreover, if any of the top boundary labels of u differs from the corresponding bottom boundary label of v, then, by convention, $v \circ u = 0$.

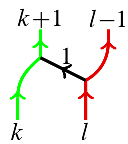
- For $j \in \mathbb{Z}_{\geq 1}$, we define the so-called *monochromatic* $F^{(j)}1_{(k,l)}$*– and* $E^{(j)}1_{(k,l)}$*–ladders* as

(2-4) $F^{(j)}1_{(k,l)} =$ 

, $E^{(j)}1_{(k,l)} =$ 

and analogously in red. (The notation $1_{(k,l)}$ is motivated by the "dual side", as we will see in Section 3.1. For the green–red web calculus it is just a shorthand to indicated the underlying objects.) Sometimes we draw such ladder rungs horizontally. We also have the *mixed* $F1_{(k,l)}$*– and* $E1_{(k,l)}$*–ladders*

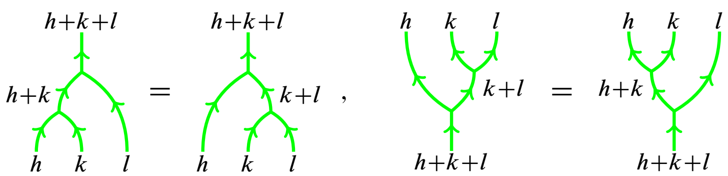
(2-5) $F1_{(k,l)} =$ 

, $E1_{(k,l)} =$ 

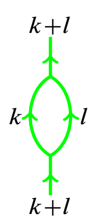
and similarly by exchanging green and red. Note that the ladders from (2-4) exist for all $j \in \mathbb{Z}_{\geq 1}$, while the mixed ladders from (2-5) exist only for $j = 1$.

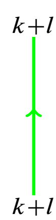
- We usually omit the object 0 as well as edges labeled 0 from illustrations; cf (2-1).

Definition 2.3 The *green–red web category* ∞–$\mathbf{Web}_{\mathrm{gr}}$ is the quotient of ∞–$\mathbf{Web}^{f}_{\mathrm{gr}}$ obtained by imposing the following local relations on morphisms. The *monochromatic relations*, which hold for green webs as well as for red webs: (*co*)*associativity*

(2-6) 

where we use the shorthand notation from (2-3) if some of the labels are 1. Next, the *digon removal relations*

(2-7) 

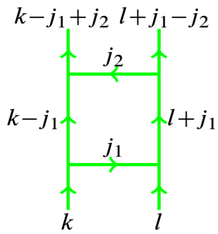
$= \begin{bmatrix} k+l \\ l \end{bmatrix}$ 

for which k and l might be 1. In these relations the (s,t)*–quantum binomial* is given by

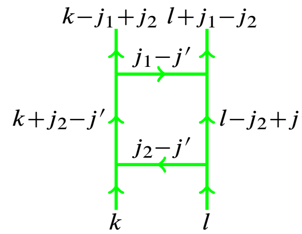
$$\begin{bmatrix} s \\ t \end{bmatrix} = \frac{[s][s-1]\cdots[s-t+2][s-t+1]}{[t]!} \in \mathbb{C}_q.$$

Here $[s]=(q^s-q^{-s})/(q-q^{-1})\in\mathbb{C}_q$ is the *quantum number* and $[t]!=[1][2]\cdots[t]\in\mathbb{C}_q$ is the *quantum factorial* for $s\in\mathbb{Z}$ and $t\in\mathbb{Z}_{\geq 0}$. Finally, the *square switch* relations

(2-8)



$$= \sum_{j'\geq 0}\begin{bmatrix} k-j_1-l+j_2 \\ j' \end{bmatrix}$$

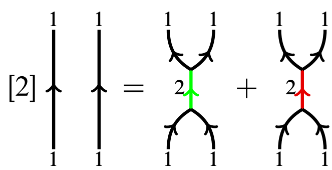


Here we allow j_1 or j_2 to be 1 (we will get *mixed* square switch relations, with one green and one red side, in Lemma 2.10).

To write these relations in a uniform manner, we allow negative labels on edges and set webs with such edges equal to zero.

The defining relation between green and red edges is

(2-9)

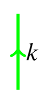


which we call the *dumbbell relation*. ◇

Remark 2.4 The category ∞–$\mathbf{Web}_{\mathrm{gr}}$ is symmetric under exchanging green and red. In the following we will often refer to this symmetry to shorten arguments.

Definition 2.5 The category N–$\mathbf{Web}_{\mathrm{gr}}$ is the quotient category obtained from the category ∞–$\mathbf{Web}_{\mathrm{gr}}$ by imposing the *exterior relations*, that is,

(2-10)



$$=0 \quad \text{if } k>N.$$

The exterior relations hold only for green edges. These relations mean that any web u with a green edge labeled $k>N$ is zero. In contrast, red edges labeled $k>N$ are usually not zero.

The sorted web category N–$\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ is the full (nonmonoidal) subcategory of N–$\mathbf{Web}_{\mathrm{gr}}$ whose object set consists of $\vec{k}\in X^L$ with no red boundary point left of a green boundary point: if $k_i\in X_r$ for some i, then $k_{>i}\in X_b\cup X_r$. ◇

Remark 2.6 The relations (2-10) are diagrammatic versions of $\bigwedge_q^{>N}\mathbb{C}_q^N\cong 0$.

Definition 2.7 The category N–$\mathbf{Web}_{\mathrm{g}}$ is the subcategory of N–$\mathbf{Web}_{\mathrm{gr}}$ consisting of only black and green objects and whose morphism spaces are spanned as $\mathbb{C}_q$–vector spaces by webs that contain only black or green edges.

Similarly, the category N–$\mathbf{Web}_{\mathrm{r}}$ is the subcategory of N–$\mathbf{Web}_{\mathrm{gr}}$ consisting of only black and red objects and whose morphism spaces are spanned as $\mathbb{C}_q$–vector spaces by webs that contain only black or red edges.

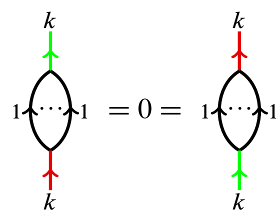
We call these categories *monochromatic*. $\diamond$

Remark 2.8 We will see in Corollary 2.16 that N–$\mathbf{Web}_{\mathrm{g}}$ is equivalent to the web category given in [3, Definition 2.2] (without tags and downward-pointing arrows). The category N–$\mathbf{Web}_{\mathrm{r}}$ is a generalization of the one given in [25, Definition 1.4]. In fact, Proposition 2.15 shows that both monochromatic subcategories are full in N–$\mathbf{Web}_{\mathrm{gr}}$.

2.2 The diagrammatic super relations

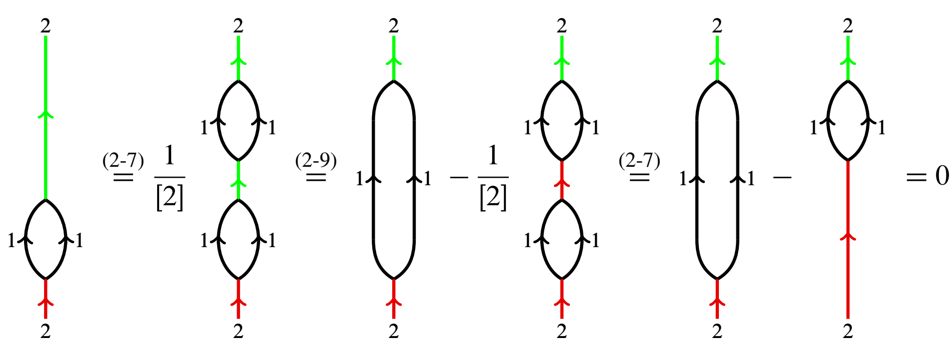
We show in this subsection that diagrammatic versions of the relations (3-1) in the Howe dual quantum group $\dot{U}_q(\mathfrak{gl}_{m|n})$ from Definition 3.1 hold in our diagrammatic categories ∞–$\mathbf{Web}_{\mathrm{gr}}$ and N–$\mathbf{Web}_{\mathrm{gr}}$.

Lemma 2.9 *We have the relations*



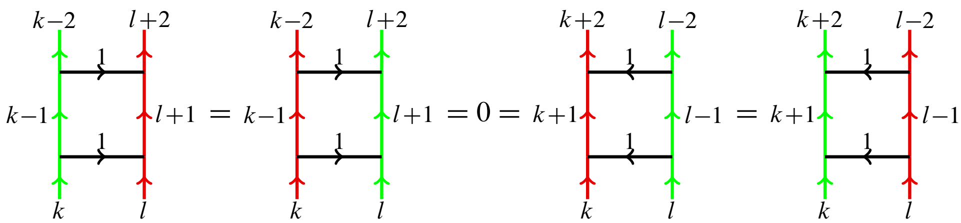
where the dots indicate k parallel black edges with label 1 which split off the bottom and merge with the top in any order (the order does not matter because of (2-6)).

Proof It suffices by associativity (2-6) to show the statement for $k=2$. We have

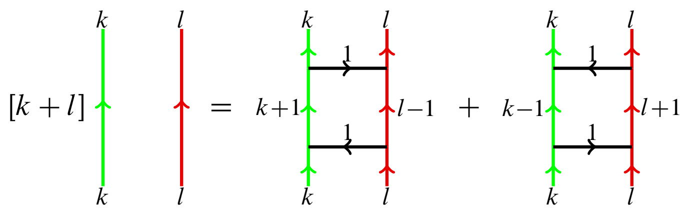


The other $k=2$ relation follows by symmetry. $\square$

Lemma 2.10 (a) *We have, for all $k, l \in \mathbb{Z}_{\geq 0}$,*

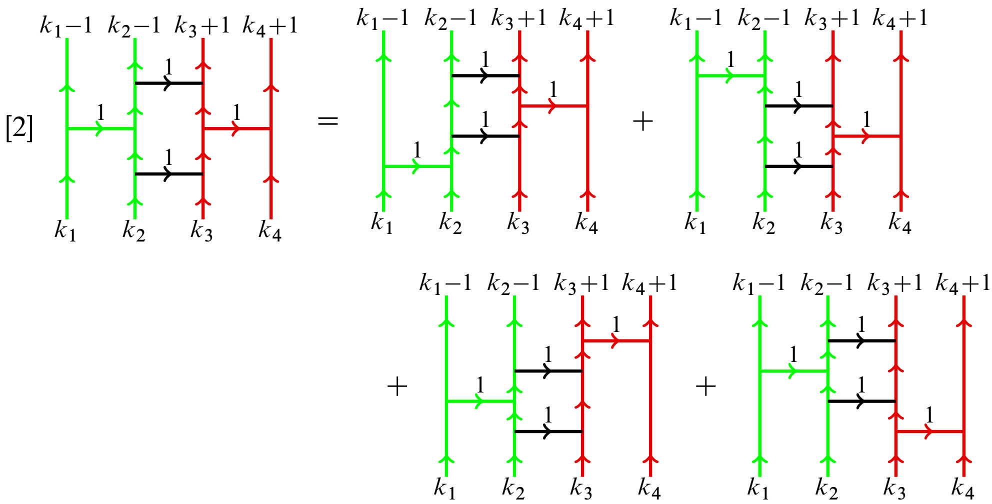


(b) *We have, for all $k, l \in \mathbb{Z}_{\geq 0}$,*



and similarly for exchanged roles of green and red.

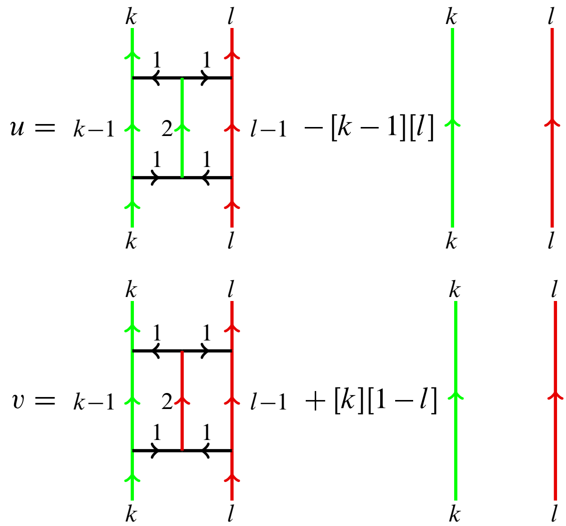
(c) *We have, for all $k, l \in \mathbb{Z}_{\geq 0}$,*



and similarly for exchanged roles of green and red, and flipped horizontal orientations.

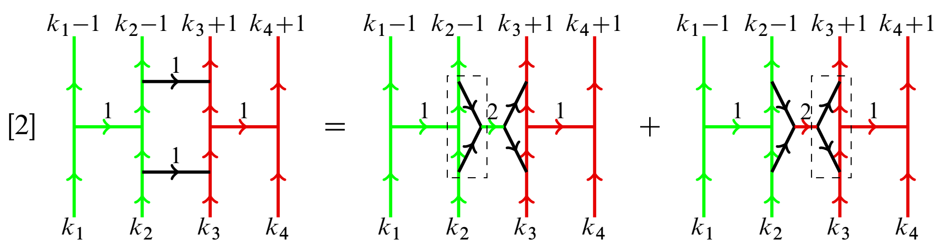
Proof (a) This follows directly from (2-6), Lemma 2.9 and symmetry.

(b) Let u and v denote the two webs on the right-hand side of (b) above. Using (2-8) for the edges labeled $k+1$ and $l+1$ in u, respectively v, we get

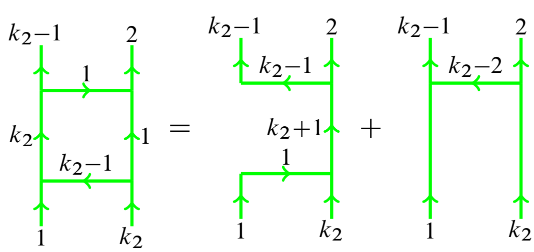


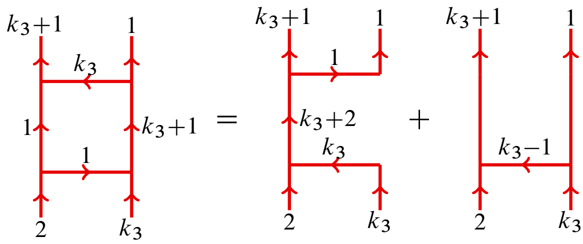
after collapsing appearing digons. By using (2-9) on the central vertical edges in the expansions, we see that $u+v=s\cdot\mathrm{id}_{(k,l)}$. The scalar is $s=[2][k][l]+[k][1-l]-[k-1][l]=[k+l]$. The other cases follow by symmetry.

(c) We start with the web on the left-hand side and first use (2-9) on the middle two horizontal edges. Thus, we obtain (our drawings are simplified and the orientations pointing down could be isotoped to point up)

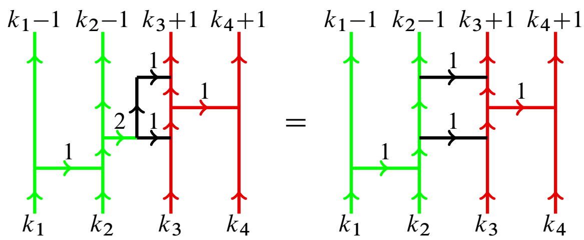


The two marked parts above are monochromatic squares, which can be switched to give





Plugging these four terms back in, we get the four webs from the right-hand side of the equation in (c) (in the indicated order), which can be seen by using (2-6), as for example



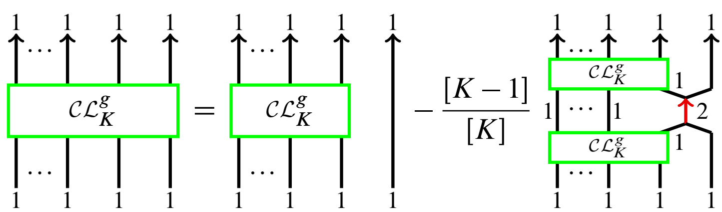
The other three cases in (c) follow by symmetry. □

2.3 Green and red clasps

We show now that our calculus contains web analogues of the *Jones–Wenzl projectors* of the Temperley–Lieb algebra. We call them *clasps*, following [15].

From now on, we denote by capital vectors such as $\vec{K} \in X^K$ special objects of ∞–$\mathbf{Web}_{\mathrm{gr}}$ of the form $\vec{K} = (1_b, \dots, 1_b)$ with K entries equal 1_b and no other entries.

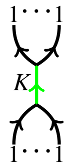
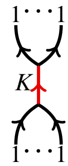
Definition 2.11 Let $K \in \mathbb{Z}_{>0}$. We define the K^{th} *green clasp* $\mathcal{CL}_K^g \in \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$ recursively: $\mathcal{CL}_1^g$ is the black identity strand and for $K \in \mathbb{Z}_{>1}$ set



and similarly for the *red clasp* $\mathcal{CL}_K^r$ by exchanging green and red. ◇

The following lemma identifies the clasps, avoiding the recursive definition.

Lemma 2.12 *We have, for all $K \in \mathbb{Z}_{>0}$,*

$$\mathcal{CL}^g_K = \frac{1}{[K]!}$$  , $$\mathcal{CL}^r_K = \frac{1}{[K]!}$$ 

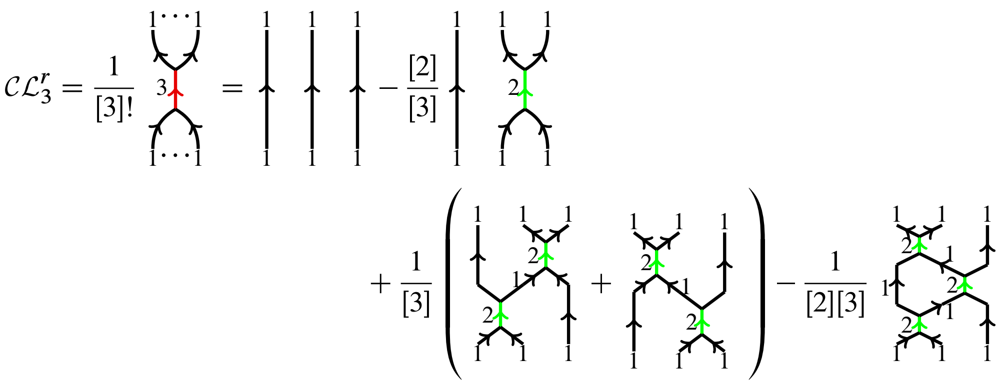
where we repeatedly split an edge labeled K until all of the top and bottom edges are black.

Proof Up to signs and drawing conventions as in [25, Lemma 2.12] and left to the reader. □

Corollary 2.13 *For all $K \in \mathbb{Z}_{>0}$, the projector $\mathcal{CL}^g_K$ can be expressed as a linear combination of webs with only black and red edges of label 2, and similarly for $\mathcal{CL}^r_K$.*

Proof This follows directly from (2-9) and Lemma 2.12. □

Example 2.14 The projector $\mathcal{CL}^r_1$ is just the black identity strand, the projector $\mathcal{CL}^r_2$ is $1/[2]$ times the red dumbbell, as in (2-9), and

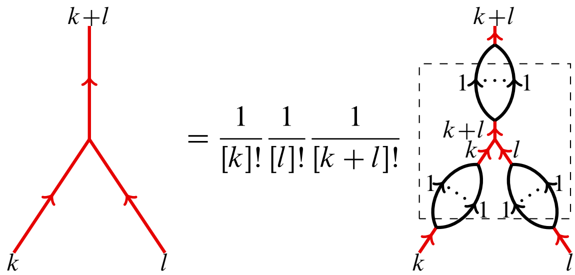


Note that all edges appearing on the right-hand side are black or green with label 2. ◁

Proposition 2.15 *Let $\vec{k}$ and $\vec{l}$ be sequences of black and green boundary points. Every web $u \in \mathrm{Hom}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k}, \vec{l})$ can be expressed as a sum of webs with only black and green edges, and similarly by exchanging green and red.*

Proof We start by exploding[8] every red edge. Around internal vertices of u with no outgoing green edges we get

[8] We "explode" by using (2-7) — the the order does not matter by (2-6). We indicate "explosions" with dots.



Note that the marked part above is $\mathcal{CL}^r_{k+l}$ up to a nonzero scalar. This can be seen by using (co)associativity (2-6) and the expression in Lemma 2.12. Thus, we can use Corollary 2.13 to replace $\mathcal{CL}^r_{k+l}$ by a nonzero sum of webs with only black and green edges. Repeating this for all purely red internal vertices shows the statement, since all outer edges are assumed to be black or green. The other statement follows by symmetry. $\square$

Denote by N–$\mathbf{Web}_{\mathrm{CKM}}$ the subcategory given in [3, Definition 2.2] with only upward-pointing strands, tags replaced by (untruncated) N–labeled edges and additionally allowing 0–labeled objects. As a consequence of Proposition 2.15 we see that interpreting webs in N–$\mathbf{Web}_{\mathrm{CKM}}$ as green webs in N–$\mathbf{Web}_{\mathrm{gr}}$ gives a full functor ι_1^∞ between these categories. In Lemma 3.13 we will see that it is also faithful and we get the following corollary.

Corollary 2.16 *The functor $\iota_1^\infty\colon N$–$\mathbf{Web}_{\mathrm{CKM}} \to N$–$\mathbf{Web}_{\mathrm{gr}}$, given by coloring webs green, is an inclusion of a full, monoidal subcategory. In particular, N–$\mathbf{Web}_{\mathrm{CKM}}$ and N–$\mathbf{Web}_{\mathrm{g}}$ are equivalent as monoidal categories.*

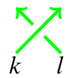
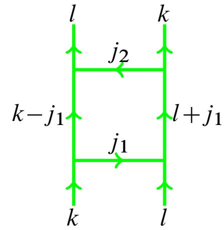
Proof The functor is well-defined since all relations in N–$\mathbf{Web}_{\mathrm{CKM}}$ hold in N–$\mathbf{Web}_{\mathrm{gr}}$. That ι_1^∞ is monoidal is clear, fullness follows from Proposition 2.15 and faithfulness from Lemma 3.13. Thus, we see that N–$\mathbf{Web}_{\mathrm{CKM}}$ and N–$\mathbf{Web}_{\mathrm{g}}$ are monoidally equivalent. $\square$

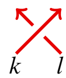
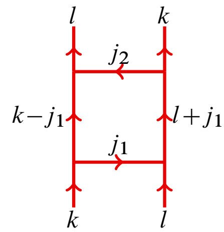
2.4 Braidings

We define now a *braided* monoidal structure on ∞–$\mathbf{Web}_{\mathrm{gr}}$.

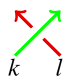
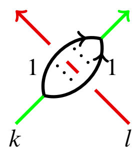
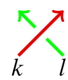
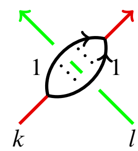
Definition 2.17 Define for $k,l \in \mathbb{Z}_{\geq 0}$ an *elementary crossing* depending on four cases. The *monochromatic crossings* (note the different powers of q)

(2-11)

$$\beta^g_{k,l} =$$  $$= (-1)^{k+kl} q^k \sum_{\substack{j_1, j_2 \geq 0 \\ j_1 - j_2 = k-l}} (-q)^{-j_1}$$ 

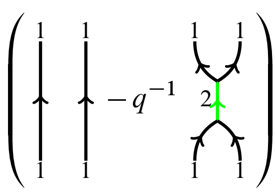
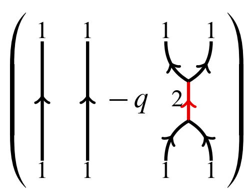
$$\beta^r_{k,l} =$$  $$= (-1)^{k} q^{-k} \sum_{\substack{j_1, j_2 \geq 0 \\ j_1 - j_2 = k-l}} (-q)^{+j_1}$$ 

The *mixed crossings* are defined via explosion of the strand going over:

(2-12) $\beta^m_{k,l} =$  $= \frac{1}{[k]!}$  and $\beta^{\tilde{m}}_{k,l} =$  $= \frac{1}{[k]!}$ 

where the remaining crossings are of the form $\beta^r_{1,l}$ or $\beta^g_{1,l}$, respectively. ◇

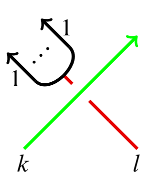
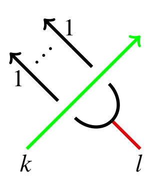
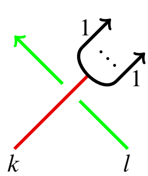
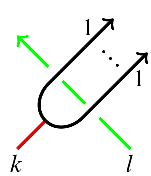
Example 2.18 The case $k = l = 1$ is not ambiguous, since we have

$$\beta^g_{1,1} = q$$  $$\overset{(2\text{-}9)}{=} -q^{-1}$$  $$= \beta^r_{1,1},$$

as a small calculation shows. ◁

As shorthand notation, we write $\beta^\bullet_{k,l}$, where $\bullet$ stands for either g, r, m or $\tilde{m}$ from now on. Note that the sums in (2-11) are finite, because webs with negative labels are zero.

Lemma 2.19 (Pitchfork relations) *We have*

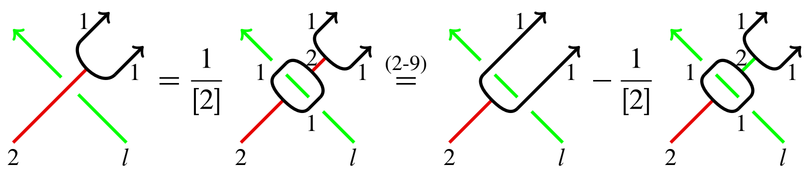
 $=$  ,  $=$ 

and similar with exchanged roles of green and red, for the monochromatic cases and with merges.

Note that the pitchfork lemma directly implies that (2-12) could also be done by exploding the edges going underneath instead of the edges going over (or exploding both).

Proof The pitchfork lemma with only green colored edges follows as in Lemma 5.3 of [22]. By symmetry, the arguments go through for the monochromatic red case as well.

The mixed, left-hand equation is easy to verify by the above, since we explode the overcrossing edge and we thus can directly use the monochromatic case. It remains to prove the mixed, right-hand equation. We only need to check the case $k=2$; the case $k \in \mathbb{Z}_{>2}$ then follows easily from this case by using Lemma 2.9. We write



The rightmost diagram is zero by Lemma 2.9 and the monochromatic pitchfork relations. This proves the mixed right-hand equation. The other cases are analogous. □

Let $\vec{k} \in X^L_{\geq 0}$ be an object in ∞–$\mathbf{Web}_{\mathrm{gr}}$. We define for $i = 1, \dots, L-1$ the crossing $\beta_i^{\bullet} 1_{\vec{k}}$ to be the corresponding elementary crossing $\beta^{\bullet}_{k_i,k_{i+1}}$ between the strands i and $i+1$ and the identity elsewhere. Clearly, it suffices to indicate the rightmost $1_{\vec{k}}$ in a sequence of the $\beta_i^{\bullet} 1_{\vec{k}}$.

Lemma 2.20 *The crossings $\beta_i^{\bullet} 1_{\vec{k}}$ satisfy the braid relations, that is, they are invertible, they satisfy the commutation relations $\beta_i^{\bullet}\beta_j^{\bullet} 1_{\vec{k}} = \beta_j^{\bullet}\beta_i^{\bullet} 1_{\vec{k}}$ for $|i-j| > 2$ and the Reidemeister 3 relations $\beta_i^{\bullet}\beta_j^{\bullet}\beta_i^{\bullet} 1_{\vec{k}} = \beta_j^{\bullet}\beta_i^{\bullet}\beta_j^{\bullet} 1_{\vec{k}}$ for $|i-j| = 1$.*

The inverses $(\beta_i^{\bullet})^{-1}$ are given as in (2-11), but with $q \to q^{-1}$. See also [22, Section 5].

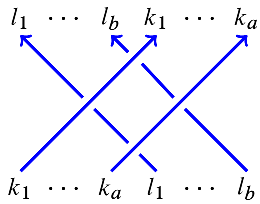
Proof This follows from Lemma 2.19, since the black case can be verified as in [22, Section 5]. □

Remark 2.21 Let S_K denote the symmetric group on K letters. Moreover, let $w \in S_K$ and let $\beta_w^{\bullet} \in \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$ be the permutation braid associated to w (this is a well-defined assignment by Lemma 2.20). Let $\ell(w)$ be the length of w. Following [14, Chapter 3, Section 2], one can show that

$$\mathcal{CL}_K^g = q^{\frac{K(K-1)}{2}} \frac{1}{[K]!} \sum_{w \in S_K} (-q)^{-\ell(w)} \beta_w^{\bullet}, \quad \mathcal{CL}_K^r = q^{-\frac{K(K-1)}{2}} \frac{1}{[K]!} \sum_{w \in S_K} q^{\ell(w)} \beta_w^{\bullet}.$$

The factors $q^{\frac{K(K-1)}{2}}$ and $q^{-\frac{K(K-1)}{2}}$ come from our conventions for crossings.

Define $\beta^{\bullet}_{\vec{k},\vec{l}}$ for objects $\vec{k}=(k_1,\dots,k_a)$ and $\vec{l}=(l_1,\dots,l_b)$ via

$\beta^{\bullet}_{\vec{k},\vec{l}} =$  $\in \mathrm{Hom}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k}\otimes\vec{l},\vec{l}\otimes\vec{k}),$

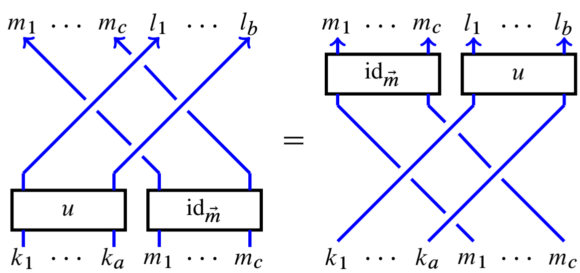
where blue stands for all suitable color possibilities.

Recall that a *braided monoidal category* (with an underlying strict monoidal category) is a pair $(\mathcal{C},\beta^{\mathcal{C}}_{\cdot,\cdot})$ consisting of a monoidal category $\mathcal{C}$ and a collection of natural isomorphisms $\beta^{\mathcal{C}}_{\vec{k},\vec{l}}\colon \vec{k}\otimes\vec{l}\to\vec{l}\otimes\vec{k}$ such that the *hexagon identities* hold for any objects $\vec{k},\vec{l},\vec{m}$ of $\mathcal{C}$:

$$\beta^{\mathcal{C}}_{\vec{k},\vec{l}\otimes\vec{m}}=(\mathrm{id}_{\vec{l}}\otimes\beta^{\mathcal{C}}_{\vec{k},\vec{m}})\circ(\beta^{\mathcal{C}}_{\vec{k},\vec{l}}\otimes\mathrm{id}_{\vec{m}}),\qquad \beta^{\mathcal{C}}_{\vec{k}\otimes\vec{l},\vec{m}}=(\beta^{\mathcal{C}}_{\vec{k},\vec{m}}\otimes\mathrm{id}_{\vec{l}})\circ(\mathrm{id}_{\vec{k}}\otimes\beta^{\mathcal{C}}_{\vec{l},\vec{m}}). \tag{2-13}$$

Proposition 2.22 *The pair $(\infty\text{–}\mathbf{Web}_{\mathrm{gr}},\beta^{\bullet}_{\cdot,\cdot})$ is a braided monoidal category.*

Proof Since $\infty\text{–}\mathbf{Web}_{\mathrm{gr}}$ is a monoidal category and the $\beta^{\bullet}_{\vec{k},\vec{l}}$ are isomorphisms that clearly satisfy (2-13), we only need to prove that they are natural. That is, we need to show that, for each web $u\in\mathrm{Hom}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k},\vec{l})$ and each other object $\vec{m}=(m_1,\dots,m_c)$ of $\infty\text{–}\mathbf{Web}_{\mathrm{gr}}$, we have (we again use blue as a generic color)



The equality follows from Lemma 2.19. This proves the statement. □

The braiding $\beta^{\bullet}_{\cdot,\cdot}$ descends to the subquotients $N\text{–}\mathbf{Web}_{\mathrm{gr}}$, $N\text{–}\mathbf{Web}_{\mathrm{g}}$ and $N\text{–}\mathbf{Web}_{\mathrm{r}}$ and we denote all induced braidings also by $\beta^{\bullet}_{\cdot,\cdot}$. They are all given by the formulas in Definition 2.17, but some diagrams might be zero due to (2-10).

Corollary 2.23 $(N\text{–}\mathbf{Web}_{\mathrm{gr}},\beta^{\bullet}_{\cdot,\cdot})$, $(N\text{–}\mathbf{Web}_{\mathrm{g}},\beta^{\bullet}_{\cdot,\cdot})$ *and* $(N\text{–}\mathbf{Web}_{\mathrm{r}},\beta^{\bullet}_{\cdot,\cdot})$, *with the braiding $\beta^{\bullet}_{\cdot,\cdot}$ induced from $(\infty\text{–}\mathbf{Web}_{\mathrm{gr}},\beta^{\bullet}_{\cdot,\cdot})$, are braided monoidal categories.* □

Note that N–$\mathbf{Web}_{\mathrm{CKM}}$ is also a braided monoidal category; see [3, Corollary 6.2.3]. We rescale their braiding by multiplying it with $q^{kl/N}$ and we denote the resulting braided monoidal category by $(N\text{–}\mathbf{Web}_{\mathrm{CKM}}, \beta^{\bullet}_{\cdot,\cdot})$. The following corollary is immediate from Corollary 2.16.

Corollary 2.24 *The functor $\iota_1^{\infty}\colon (N\text{–}\mathbf{Web}_{\mathrm{CKM}}, \beta^{\bullet}_{\cdot,\cdot}) \to (N\text{–}\mathbf{Web}_{\mathrm{gr}}, \beta^{\bullet}_{\cdot,\cdot})$ is an inclusion of a full, braided monoidal subcategory.* □

2.5 A collection of diagrammatic idempotents

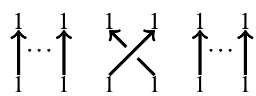
Recall that the *Iwahori–Hecke algebra* $H_K(q)$ is the q–deformation of the symmetric group algebra $\mathbb{C}[S_K]$ on K letters. It is generated by $\{H_i \mid s_i \in S_K\}$ for all transpositions $s_i = (i, i+1) \in S_K$, subject to the relations

$$\begin{aligned} H_i^2 &= (q - q^{-1})H_i + 1 && \text{for } i = 1, \dots, K-1,\\ H_i H_j &= H_j H_i && \text{for } |i-j| > 1,\\ H_i H_j H_i &= H_j H_i H_j && \text{for } |i-j| = 1. \end{aligned}$$

There is a representation $p_K\colon \mathbb{C}_q(B_K) \to H_K(q)$ of the group algebra $\mathbb{C}_q(B_K)$ of the braid group B_K with K strands given by sending the braid group generators b_i (between the strands i and $i+1$) to H_i. Thinking of the generators H_i of $H_K(q)$ as crossings also makes sense from the perspective of the webs, as the next lemma shows.

Lemma 2.25 *Given $K \in \mathbb{Z}_{\geq 0}$, there is an isomorphism of $\mathbb{C}_q$–algebras*

$$\Phi_{q\mathrm{SW}}^{\infty}\colon H_K(q) \xrightarrow{\cong} \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K}), \qquad H_i \mapsto$$



In order to prove Lemma 2.25, which will be used in Section 4, we need Theorem 3.20.

Proof A direct computation shows that $\Phi_{q\mathrm{SW}}$ is a well-defined $\mathbb{C}_q$–algebra homomorphism. In fact, the composite $\Gamma \circ \Phi_{q\mathrm{SW}}^{\infty}$ is the isomorphism induced by quantum Schur–Weyl duality. To see this, let $V = (\mathbb{C}_q^N)^{\otimes K}$ and recall that quantum Schur–Weyl duality states that

$$\begin{aligned} &\Phi_{q\mathrm{SW}}^{N}\colon H_K(q) \twoheadrightarrow \mathrm{End}_{\mathbf{U}_q(\mathfrak{gl}_N)}(V),\\ &\Phi_{q\mathrm{SW}}^{N}\colon H_K(q) \xrightarrow{\cong} \mathrm{End}_{\mathbf{U}_q(\mathfrak{gl}_N)}(V) \quad \text{if } N \geq K. \end{aligned} \tag{2-14}$$

Here $\Phi_{q\mathrm{SW}}^{N}$ is the $\mathbb{C}_q$–algebra homomorphism induced by the action of $H_K(q)$ on the K–fold tensor product V. By Theorem 3.20, we will get an isomorphism $H_K(q) \cong$

$\mathrm{End}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$ if $N \geq K$. By using Proposition 2.15, there is a basis of $\mathrm{End}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$ for $N \geq K$ given by webs with only black edges or green edges with labels at most K. Since K is fixed, a direct comparison shows that $\Phi^{\infty}_{q\mathrm{SW}}$ has to be an isomorphism as well. $\square$

Let $K \in \mathbb{Z}_{\geq 0}$ and let $\Lambda^{+}(K)$ denote the set of all *Young diagrams with K nodes*, eg

$$\lambda = (4,3,1,1) \in \Lambda^{+}(9) \leftrightsquigarrow \lambda = \ydiagram{4,3,1,1}\,,$$

$$\lambda^{T} = (4,2,2,1) \in \Lambda^{+}(9) \leftrightsquigarrow \lambda^{\mathrm{T}} = \ydiagram{4,2,2,1}\,,$$

where we use the English notation for our Young diagrams. Here we have also displayed the *transpose* Young diagram λ^{T} of λ. Next, the following definition is motivated by [11; 1]. (It is best explained via examples — cf Example 2.27 and Example 2.29 — which the reader might want to check while reading the definition.)

Definition 2.26 (Gyoja–Aiston idempotents) Given $\lambda \in \Lambda^{+}(K)$, we associate to it a primitive *idempotent* $e_q(\lambda) \in \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$. First we define two idempotents as tensor products of green or red clasps:

$$e_{\mathrm{col}}(\lambda) = \mathcal{CL}^{g}_{\mathrm{col}_1} \otimes \cdots \otimes \mathcal{CL}^{g}_{\mathrm{col}_c}, \quad e_{\mathrm{row}}(\lambda) = \mathcal{CL}^{r}_{\mathrm{row}_1} \otimes \cdots \otimes \mathcal{CL}^{r}_{\mathrm{row}_r},$$

where c and r are the number of columns and rows of λ respectively, and col_i and row_i denote the number of nodes in the i^{th} column and row.

Denote by $T^{\rightarrow}_{\lambda}$ and by $T^{\downarrow}_{\lambda}$ the two tableaux of shape λ obtained by filling the numbers $1, \dots, K$ into the Young diagram λ in order: $\rightarrow$ means rows before columns and $\downarrow$ means columns before rows (both from left to right). Pick any shortest presentation of the permutation $w(\lambda) \in S_K$ permuting $T^{\rightarrow}_{\lambda}$ to $T^{\downarrow}_{\lambda}$. Then we define the *quasi-idempotent associated to* λ via

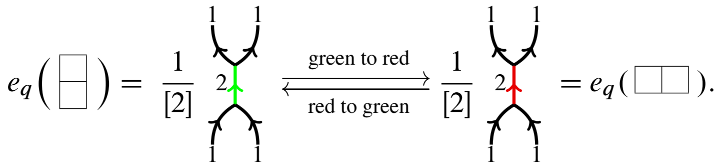
$$\widetilde{e}_q(\lambda) = e_{\mathrm{col}}(\lambda) \circ \beta^{\bullet}_{w(\lambda)} \circ e_{\mathrm{row}}(\lambda) \circ (\beta^{\bullet}_{w(\lambda)})^{-1}.$$

By [1, Theorem 4.7] (and the fact that their definition agrees with ours by Lemma 2.25 and Remark 2.21), there exists a nonzero scalar $a(\lambda) \in \mathbb{C}_q$ such that $\widetilde{e}_q(\lambda)^2 = a(\lambda)\widetilde{e}_q(\lambda)$. Thus, we define the *idempotent associated to* λ to be

$$e_q(\lambda) = \frac{1}{a(\lambda)}\widetilde{e}_q(\lambda). \qquad \diamond$$

These idempotents are primitive and orthogonal by [11, Theorem 4.5; 1, Theorem 4.7].

Example 2.27 If $K = 2$, then there are two primitive idempotents, namely



Note that $a(\lambda) = 1$ for only one column or only one row Young diagrams λ. ◁

Lemma 2.28 *Exchanging green and red sends $e_q(\lambda)$ to $e_q(\lambda^{\mathrm{T}})$ modulo a commutator.*

Proof Note that $e_{\mathrm{col}}(\lambda)$ and $e_{\mathrm{row}}(\lambda)$ differ from $e_{\mathrm{row}}(\lambda^{\mathrm{T}})$ and $e_{\mathrm{col}}(\lambda^{\mathrm{T}})$, respectively, only in exchanging the colors green and red. On black crossings the green–red symmetry acts by $\beta^{\bullet}_{1,1} \mapsto -(\beta^{\bullet}_{1,1})^{-1}$, on permutation braids as $\beta^{\bullet}_{w} \mapsto (-1)^{\ell(w)}(\beta^{\bullet}_{w^{-1}})^{-1}$ and on the quasi-idempotent $\widetilde{e}_q(\lambda)$ as

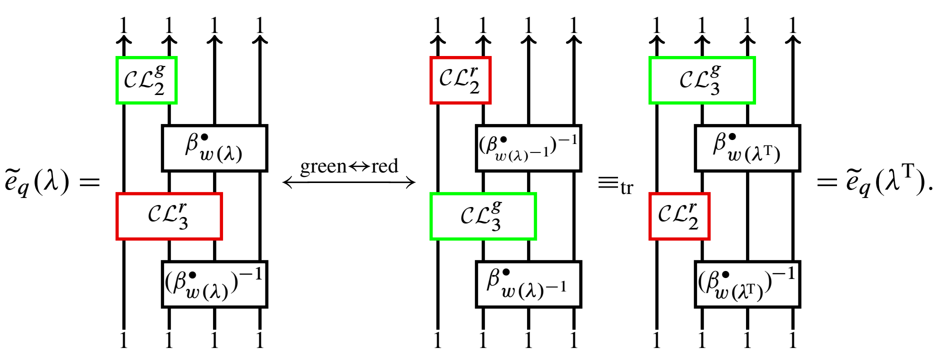
$$\begin{aligned}\widetilde{e}_q(\lambda) = e_{\mathrm{col}}(\lambda) \circ \beta^{\bullet}_{w(\lambda)} \circ e_{\mathrm{row}}(\lambda) \circ (\beta^{\bullet}_{w(\lambda)})^{-1} \\ \mapsto e_{\mathrm{row}}(\lambda^{\mathrm{T}}) \circ (\beta^{\bullet}_{w(\lambda)^{-1}})^{-1} \circ e_{\mathrm{col}}(\lambda^{\mathrm{T}}) \circ \beta^{\bullet}_{w(\lambda)^{-1}} \\ = e_{\mathrm{row}}(\lambda^{\mathrm{T}}) \circ (\beta^{\bullet}_{w(\lambda^{\mathrm{T}})})^{-1} \circ e_{\mathrm{col}}(\lambda^{\mathrm{T}}) \circ \beta^{\bullet}_{w(\lambda^{\mathrm{T}})}.\end{aligned}$$

In the first line, the signs from the crossing inversions cancel, and in the second line we use $w(\lambda)^{-1} = w(\lambda^{\mathrm{T}})$. The result agrees with $\widetilde{e}_q(\lambda^{\mathrm{T}})$ up to a commutator. This proves the statement of the lemma for the quasi-idempotents. Applying the green–red symmetry to both sides of the equation $\widetilde{e}_q(\lambda)^2 = a(\lambda)\widetilde{e}_q(\lambda)$ shows that $a(\lambda) = a(\lambda^{\mathrm{T}})$ and the lemma follows. □

Example 2.29 For $\lambda = (3, 1) \in \Lambda^{+}(4)$, we have

$$\lambda = \begin{array}{|c|c|c|}\hline & & \\ \hline & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}, \quad T^{\rightarrow}_{\lambda} = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}, \quad T^{\downarrow}_{\lambda} = \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}.$$

Thus, $w = (243) = (23)(34) \in S_4$ permutes $T^{\rightarrow}_{\lambda}$ to $T^{\downarrow}_{\lambda}$. Then



Here $\equiv_{\mathrm{tr}}$ means equal modulo a commutator and the scaling factor in this case is $a(\lambda) = [4]/([2][3]) = a(\lambda^{\mathrm{T}})$. $\triangleleft$

Remark 2.30 For $N \geq K$, the $H_K(q)$–module $(\mathbb{C}_q^N)^{\otimes K}$ decomposes into

$$\bigoplus_{\lambda \in \Lambda^+(K)} (S^\lambda)^{\oplus m_\lambda},$$

where the S^λ are the irreducible *Specht modules* for $H_K(q)$ and m_λ are their multiplicities. The primitive idempotents $e_q(\lambda)$ from Definition 2.26 are quantizations of Young symmetrizers that project onto S^λ. Note that a braid-conjugate of $e_q(\lambda)$ might project onto a different copy of S^λ in the above decomposition.

3 Proofs of the diagrammatic presentations

This section contains the proof of our main theorems.

3.1 Super q–Howe duality

Let $m, n \in \mathbb{Z}_{\geq 0}$. We start by recalling the *quantum general linear superalgebra* $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ and its *idempotented form* $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$. We follow the conventions used in [33], but adapt Zhang's notation to be closer to the one from [3].

To this end, recall that the $\mathfrak{gl}_{m|n}$–weight lattice is isomorphic to $\mathbb{Z}^{m+n}$ and we denote the $\mathfrak{gl}_{m|n}$–weights usually by vectors $\vec{k} = (k_1, \dots, k_m, k_{m+1}, \dots, k_{m+n})$. For $\mathbb{I} = \mathbb{I}_0 \cup \mathbb{I}_1$ with $\mathbb{I}_0 = \{1, \dots, m\}$ (even part) and $\mathbb{I}_1 = \{m+1, \dots, m+n\}$ (odd part), define

$$|i| = \begin{cases} 0 & \text{if } i \in \mathbb{I}_0 = \{1, \dots, m\}, \\ 1 & \text{if } i \in \mathbb{I}_1 = \{m+1, \dots, m+n\}. \end{cases}$$

The notation $|\cdot|$ means the *super* degree (which is a $\mathbb{Z}/2$–degree). We use a similar notation for all $\mathbb{Z}/2$–graded spaces, where we, by convention, always consider degrees modulo 2 in the following. Moreover, let $\epsilon_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{Z}^{m+n}$, with 1 being in the i^{th} coordinate, and denote by $\alpha_i = \epsilon_i - \epsilon_{i+1} = (0, \dots, 1, -1, \dots, 0) \in \mathbb{Z}^{m+n}$ for $i \in \mathbb{I} - \{m+n\}$ the i^{th} simple root. Recall that the *super* Euclidean inner product on $\mathbb{Z}^{m+n}$ is given by $(\epsilon_i, \epsilon_j)_{\mathrm{su}} = (-1)^{|i|}\delta_{i,j}$.

Definition 3.1 Let $m, n \in \mathbb{Z}_{\geq 0}$. The *quantum general linear superalgebra* $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ is the associative, $\mathbb{Z}/2$–graded, unital $\mathbb{C}_q$–algebra generated by $L_i^{\pm 1}$ for $i \in \mathbb{I}$, and F_i and E_i for $i \in \mathbb{I} - \{m+n\}$, subject to the *nonsuper* relations

$$L_i L_j = L_j L_i, \qquad L_i L_i^{-1} = L_i^{-1} L_i = 1,$$
$$L_i F_j = q^{-(\epsilon_i, \alpha_j)_{\mathrm{su}}} F_j L_i, \qquad L_i E_j = q^{(\epsilon_i, \alpha_j)_{\mathrm{su}}} E_j L_i,$$

$$\begin{aligned}
E_iF_j - F_jE_i &= (-1)^{|i|}\delta_{i,j}\frac{L_iL_{i+1}^{-1} - L_i^{-1}L_{i+1}}{q-q^{-1}} && \text{if } i \neq m,\\
[2]F_iF_jF_i &= F_i^2F_j + F_jF_i^2 && \text{if } |i-j| = 1,\ i \neq m,\\
[2]E_iE_jE_i &= E_i^2E_j + E_jE_i^2 && \text{if } |i-j| = 1,\ i \neq m,\\
F_iF_j - F_jF_i &= 0 && \text{if } |i-j| > 1,\\
E_iE_j - E_jE_i &= 0 && \text{if } |i-j| > 1
\end{aligned}$$

(for suitable $i, j \in \mathbb{I}$) and the *super relations*

$$F_m^2 = 0 = E_m^2, \quad E_mF_m + F_mE_m = \frac{L_mL_{m+1}^{-1} - L_m^{-1}L_{m+1}}{q-q^{-1}},$$

$$[2]F_mF_{m+1}F_{m-1}F_m =$$
$$F_mF_{m+1}F_mF_{m-1} + F_{m-1}F_mF_{m+1}F_m + F_{m+1}F_mF_{m-1}F_m + F_mF_{m-1}F_mF_{m+1},$$
$$[2]E_mE_{m+1}E_{m-1}E_m =$$
$$E_mE_{m+1}E_mE_{m-1}+E_{m-1}E_mE_{m+1}E_m+E_{m+1}E_mE_{m-1}E_m+E_mE_{m-1}E_mE_{m+1}.$$

Also, $|L_i| = 0$ for $i \in \mathbb{I}$, $|F_i| = |E_i| = 0$ for $i \in \mathbb{I} - \{m\}$ and $|F_m| = |E_m| = 1$. ◇

We recover $\boldsymbol{U}_q(\mathfrak{gl}_N)$ by setting $m = N$ and $n = 0$. We write $\mathbb{I}_N = \{1, \dots, N\}$ in the following to distinguish it from $\mathbb{I}$ as above. Note that $\boldsymbol{U}_q(\mathfrak{gl}_N)$ is concentrated in degree 0.

The algebra $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ is a $\mathbb{Z}/2$–graded Hopf algebra with coproduct Δ, antipode S and the counit ε given by

$$\Delta(F_i) = F_i\otimes 1 + L_i^{-1}L_{i+1}\otimes F_i, \quad \Delta(E_i) = E_i\otimes L_iL_{i+1}^{-1} + 1\otimes E_i, \quad \Delta(L_i) = L_i\otimes L_i,$$
$$S(F_i) = -L_iL_{i+1}^{-1}F_i, \quad S(E_i) = -E_iL_i^{-1}L_{i+1}, \quad S(L_i) = L_i^{-1},$$
$$\varepsilon(F_i) = \varepsilon(E_i) = 0, \quad \varepsilon(L_i) = 1.$$

In the spirit of Lusztig [20, Chapter 23], we now adjoin, for all $\vec{k} \in \mathbb{Z}^{m+n}$, idempotents $1_{\vec{k}}$ of super degree $|1_{\vec{k}}| = 0$ to $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$. Denote by I the ideal generated by

$$1_{\vec{k}}1_{\vec{l}} = \delta_{\vec{k},\vec{l}}1_{\vec{k}}, \qquad 1_{\vec{k}-\alpha_i}F_i1_{\vec{k}} = F_i1_{\vec{k}} = 1_{\vec{k}-\alpha_i}F_i,$$
$$L_i1_{\vec{k}} = q^{k_i(\epsilon_i,\epsilon_i)_{\mathrm{su}}}1_{\vec{k}}, \quad 1_{\vec{k}+\alpha_i}E_i1_{\vec{k}} = E_i1_{\vec{k}} = 1_{\vec{k}+\alpha_i}E_i.$$

Definition 3.2 Define by

$$\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n}) = \Bigg(\bigoplus_{\vec{k},\vec{l}\in\mathbb{Z}^{m+n}} 1_{\vec{l}}\boldsymbol{U}_q(\mathfrak{gl}_{m|n})1_{\vec{k}}\Bigg)\Big/ I$$

the *idempotented* quantum general linear superalgebra. ◇

Remark 3.3 One can view $\dot{U}_q(\mathfrak{gl}_{m|n})$ as generated by the *divided powers*

$$F_i^{(j)} = \frac{F_i^j}{[j]!} \quad \text{and} \quad E_i^{(j)} = \frac{E_i^j}{[j]!} \quad \text{for } i \in \mathbb{I} - \{m+n\}.$$

This allows the definition of an integral version of $\dot{U}_q(\mathfrak{gl}_{m|n})$. For simplicity, we work over $\mathbb{C}_q$ in this paper and we do not consider the integral version.

The relations in $\dot{U}_q(\mathfrak{gl}_{m|n})$ are obtained from the relations of $U_q(\mathfrak{gl}_{m|n})$. For convenience we list the new versions of the *super* relations:

$$\begin{aligned}
&F_m^2 1_{\vec{k}} = 0 = E_m^2 1_{\vec{k}},\\
&E_m F_m 1_{\vec{k}} + F_m E_m 1_{\vec{k}} = [k_m + k_{m+1}] 1_{\vec{k}},\\
&[2] F_m F_{m+1} F_{m-1} F_m 1_{\vec{k}} = F_m F_{m+1} F_m F_{m-1} 1_{\vec{k}} + F_{m-1} F_m F_{m+1} F_m 1_{\vec{k}}\\
&\qquad\qquad + F_{m+1} F_m F_{m-1} F_m 1_{\vec{k}} + F_m F_{m-1} F_m F_{m+1} 1_{\vec{k}},
\end{aligned} \tag{3-1}$$

the second of which we call the *super commutation relation* (the third type of relation holds for E as well).

It is convenient for us hereinafter to view $\dot{U}_q(\mathfrak{gl}_{m|n})$ as a category whose objects are the $\mathfrak{gl}_{m|n}$–weights $\vec{k} \in \mathbb{Z}^{m+n}$ and $\mathrm{Hom}_{\dot{U}_q(\mathfrak{gl}_{m|n})}(\vec{k}, \vec{l}) = 1_{\vec{l}} \dot{U}_q(\mathfrak{gl}_{m|n}) 1_{\vec{k}}$.

Recall that the vector representation $\mathbb{C}_q^{m|n}$ of $U_q(\mathfrak{gl}_{m|n})$ has a basis given by $\{x_i \mid i \in \mathbb{I}\}$ with super degrees $|x_i| = |i|$ for $i \in \mathbb{I}$, where the $U_q(\mathfrak{gl}_{m|n})$–action is defined via

$$F_i(x_j) = \begin{cases} x_{j+1} & \text{if } i = j,\\ 0 & \text{otherwise,}\end{cases} \qquad E_i(x_j) = \begin{cases} x_{j-1} & \text{if } i = j-1,\\ 0, & \text{otherwise,}\end{cases}$$
$$L_i(x_j) = q^{(\epsilon_i, \epsilon_j)_{\mathrm{su}}} x_j.$$

We need to consider the *quantum exterior superalgebra* $\bigwedge_q^{\bullet}(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$. Recall that a vector space $V = V_0 \oplus V_1$ with a $\mathbb{Z}/2$–grading is called a *super* vector space. Here V_0 and V_1 are its degree 0 and 1 parts. These graded parts of $\mathbb{C}_q^{m|n}$ have bases given by $\{x_i \mid i \in \mathbb{I}_0\}$ and $\{x_i \mid i \in \mathbb{I}_1\}$, respectively. In contrast, $\mathbb{C}_q^N = (\mathbb{C}_q^N)_0$ is concentrated in degree zero and we denote its basis by $\{y_j \mid j \in \mathbb{I}_N\}$. Additionally, the tensor product $V \otimes W$ of two super vector spaces V and W is a super vector space with $v \otimes w$ of degree $|v| + |w|$ for two homogeneous elements v and w. Specifically, $\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N$ is a super vector space with $(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)_0$ spanned by $\{z_{ij} = x_i \otimes y_j \mid i \in \mathbb{I}_0,\ j \in \mathbb{I}_N\}$ and $(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)_1$ spanned by $\{z_{ij} = x_i \otimes y_j \mid i \in \mathbb{I}_1,\ j \in \mathbb{I}_N\}$. Here $|z_{ij}| = |i|$. Note that $(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)^{\otimes K}$ is a $\mathbb{Z}/2$–graded $U_q(\mathfrak{gl}_{m|n}) \otimes U_q(\mathfrak{gl}_N)$–module for all $K \in \mathbb{Z}_{\geq 0}$ by using the Hopf algebras structures of $U_q(\mathfrak{gl}_{m|n})$ and $U_q(\mathfrak{gl}_N)$.

We denote by $\mathrm{Sym}^2_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$ the *second symmetric super power* as in [23, (4.1)], but with q inverted in their formulas. Armed with this notation, we define the *quantum exterior superalgebra*

$$\textstyle\bigwedge^\bullet_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N) = T(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)/\mathrm{Sym}^2_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N),$$

where $T(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N) = \bigoplus_{K\in\mathbb{Z}_{\geq 0}}(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)^{\otimes K}$ denotes the *super tensor algebra* of $\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N$. This is a $\boldsymbol{U}_q(\mathfrak{gl}_{m|n}) \otimes \boldsymbol{U}_q(\mathfrak{gl}_N)$–module and decompose as

$$\textstyle\bigwedge^\bullet_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N) \cong \bigoplus_{K\in\mathbb{Z}_{\geq 0}} \bigwedge^K_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N).$$

The space $\bigwedge^K_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$ is called the *degree K part* of $\bigwedge^\bullet_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$.

Remark 3.4 We can recover the degree K part of the quantum exterior algebra $\bigwedge^K_q(\mathbb{C}_q^m \otimes \mathbb{C}_q^N)$ by setting $n = 0$ and, by [28, Remark 2.1], the degree K part of the quantum symmetric algebra $\mathrm{Sym}^K_q(\mathbb{C}_q^n \otimes \mathbb{C}_q^N)$ by setting $m = 0$. These were originally defined in [2, Definition 2.7] and used in [3, Section 4.2; 25, Section 2.1] to study skew and symmetric q–Howe duality.

Example 3.5 Write $z_{\vec{ij}} = z_{i_1 j_1} \otimes \cdots \otimes z_{i_K j_K}$ and $z_{i_k j_k} \preceq z_{i_{k+1} j_{k+1}}$ for the anti-lexicographical order on the indices of the z_{ij}. Then $\bigwedge^K_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$ has a basis given by (cf [23, Lemma 4.1])

$$\begin{aligned}\text{(3-2)}\quad \big\{z_{\vec{ij}} \mid z_{i_k j_k} \preceq z_{i_{k+1} j_{k+1}},\ 1 \leq i_1 \leq \cdots \leq i_K \leq m+n,\ 1 \leq j_1 \leq \cdots \leq j_K \leq N,\\ \text{and } |i_k| = 1, \text{ if } i_k = i_{k+1} \text{ and } j_k = j_{k+1}\big\}.\end{aligned}$$

By setting $m = 1$ and $n = 0$, we obtain the (usual) basis for $\bigwedge^K_q \mathbb{C}_q^N$ of the form

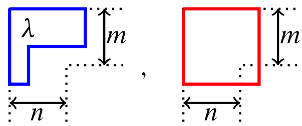
$$\{y_{i_1} \otimes \cdots \otimes y_{i_K} \mid 1 \leq y_1 < \cdots < y_K \leq N\}, \tag{3-3}$$

while setting $m = 0$ and $n = 1$ gives the (usual) basis for $\mathrm{Sym}^K_q \mathbb{C}_q^N$ of the form

$$\{y_{i_1} \otimes \cdots \otimes y_{i_K} \mid 1 \leq y_1 \leq \cdots \leq y_K \leq N\}. \tag{3-4}$$

These are precisely the usual (nonsuper) bases; see for example [2, Section 2.4]. $\lhd$

We call a $\mathfrak{gl}_{m|n}$–weight $\lambda = (\lambda_1, \dots, \lambda_{m+n}) \in \mathbb{Z}^{m+n}$ a *dominant integral* $\mathfrak{gl}_{m|n}$–weight if it is a dominant integral $\mathfrak{gl}_m \oplus \mathfrak{gl}_n$–weight. We only need λ that are $(m|n)$–hook Young diagrams, ie diagrams that fit into a hook-shaped region with one horizontal arm of height m and one vertical arm of width n (here we use the conventions from [4, Definition 2.10]). The following figure shows an $(m|n)$–hook Young diagram λ and a box-shaped Young diagram that is not an $(m|n)$–hook:



We call a dominant integral $\mathfrak{gl}_{m|n}$–weight λ an $(m|n,N)$*–supported* $\mathfrak{gl}_{m|n}$*–weight* if it corresponds to an $(m|n)$–hook Young diagram with at most N columns. For each such λ there exists an irreducible $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$–module $L_{m|n}(\lambda)$ and an irreducible $\boldsymbol{U}_q(\mathfrak{gl}_N)$–module $L_N(\lambda^{\mathrm{T}})$; see eg [16, Section 2.5].

Theorem 3.6 (Super q–Howe duality) *We have the following:*

(a) *Let $K \in \mathbb{Z}_{\geq 0}$. The actions of $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ and $\boldsymbol{U}_q(\mathfrak{gl}_N)$ on $\bigwedge_q^K(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$ commute and generate each others commutant.*

(b) *There exists an isomorphism*

$$\bigwedge_q^{\bullet}(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N) \cong (\bigwedge_q^{\bullet}\mathbb{C}_q^N)^{\otimes m} \otimes (\mathrm{Sym}_q^{\bullet}\mathbb{C}_q^N)^{\otimes n}$$

of $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules under which the $\vec{k}$–weight space of $\bigwedge_q^{\bullet}(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N)$ (considered as a $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$–module) is identified with

$$\bigwedge_q^{\vec{k}_0}\mathbb{C}_q^N \otimes \mathrm{Sym}_q^{\vec{k}_1}\mathbb{C}_q^N = \bigwedge_q^{k_1}\mathbb{C}_q^N \otimes \cdots \otimes \bigwedge_q^{k_m}\mathbb{C}_q^N \otimes \mathrm{Sym}_q^{k_{m+1}}\mathbb{C}_q^N \otimes \cdots \otimes \mathrm{Sym}_q^{k_{m+n}}\mathbb{C}_q^N. \tag{3-5}$$

Here $\vec{k} = (k_1, \dots, k_{m+n})$, $\vec{k}_0 = (k_1, \dots, k_m)$ and $\vec{k}_1 = (k_{m+1}, \dots, k_{m+n})$.

(c) *As $\boldsymbol{U}_q(\mathfrak{gl}_{m|n}) \otimes \boldsymbol{U}_q(\mathfrak{gl}_N)$–modules, we have a decomposition of the form*

$$\bigwedge_q^K(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^N) \cong \bigoplus_{\lambda} L_{m|n}(\lambda) \otimes L_N(\lambda^{\mathrm{T}}),$$

where we sum over all $(m|n,N)$–supported $\mathfrak{gl}_{m|n}$–weights λ whose entries sum up to K. This induces a decomposition

$$\bigwedge_q^{\bullet}(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^n) \cong \bigoplus_{\lambda} L_{m|n}(\lambda) \otimes L_N(\lambda^{\mathrm{T}}),$$

where we sum over all $(m|n,N)$–supported $\mathfrak{gl}_{m|n}$–weights λ.

Remark 3.7 Symmetric and skew Howe duality for the pair $(\mathrm{GL}_m, \mathrm{GL}_N)$ is originally due to Howe; see [12, Sections 2 and 4]. Note that the nonquantum version of Theorem 3.6 can be found for example in [4, Theorem 3.3] or [28, Proposition 2.2]. Moreover, the "dual" of Theorem 3.6, given by considering $\boldsymbol{U}_q(\mathfrak{gl}_N)$ as the Howe dual group instead of $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$, can be found in [23, Proposition 4.3].

Proof Parts (a) and (c) are proven in [31, Theorem 2.2] or in [23, Theorem 4.2] and only (b) remains to be verified. For this purpose, we use the bases from (3-2), (3-3) and (3-4) to define

$$T_i^e\colon \textstyle\bigwedge_q^k(\mathbb{C}_q^N) \to \bigwedge_q^k(\mathbb{C}_q^{m|n}\otimes\mathbb{C}_q^N), \qquad y_{j_1}\otimes\cdots\otimes y_{j_k}\mapsto z_{ij_1}\otimes\cdots\otimes z_{ij_k}, \quad i\in\mathbb{I}_0,$$
$$T_i^s\colon \mathrm{Sym}_q^k(\mathbb{C}_q^n) \to \textstyle\bigwedge_q^k(\mathbb{C}_q^{m|n}\otimes\mathbb{C}_q^N), \quad y_{j_1}\otimes\cdots\otimes y_{j_k}\mapsto z_{ij_1}\otimes\cdots\otimes z_{ij_k}, \quad i\in\mathbb{I}_1.$$

That these maps are well-defined $\boldsymbol{U}_q(\mathfrak{gl}_N)$–intertwiners follows from the explicit description in Example 3.5. Injectivity was shown in [3, Theorem 4.2.2] for the first and in [25, Theorem 2.6] for the second map. Thus, for $\vec{k}\in\mathbb{Z}^{m+n}$ with $k_1+\cdots+k_{m+n}=K$, we see that

$$T\colon \bigoplus_{\vec{k}\in\mathbb{Z}^{m+n}_{\geq 0}} \textstyle\bigwedge_q^{\vec{k}_0}\mathbb{C}_q^N\otimes\mathrm{Sym}_q^{\vec{k}_1}\mathbb{C}_q^N \to \bigwedge_q^K(\mathbb{C}_q^{m|n}\otimes\mathbb{C}_q^N)$$

given by

$$T(v_1\otimes\cdots\otimes v_{m+n}) = T_1^e(v_1)\otimes\cdots\otimes T_m^e(v_m)\otimes T_{m+1}^s(v_{m+1})\otimes\cdots\otimes T_{m+n}^s(v_{m+n})$$

is a $\boldsymbol{U}_q(\mathfrak{gl}_N)$–module isomorphism by comparing the sizes of the bases from Example 3.5. This clearly induces the isomorphism of $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules we are looking for.

It remains to verify the $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$–weight space decomposition from (3-5). To this end, we only have to see that the action on $\bigwedge_q^{\vec{k}_0}\mathbb{C}_q^N\otimes\mathrm{Sym}_q^{\vec{k}_1}\mathbb{C}_q^N$ of the $L_{i'}$ of $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ under the inverse of T is just a multiplication with $q^{k_i(\epsilon_i,\epsilon_{i'})_{\mathrm{su}}}$. The action of $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ is given by

$$L_{i'}(z_{ij_1}\otimes\cdots\otimes z_{ij_{m+n}}) = L_{i'}(z_{ij_1})\otimes\cdots\otimes L_{i'}(z_{ij_{m+n}}) = q^{k_i(\epsilon_i,\epsilon_{i'})_{\mathrm{su}}} z_{ij_1}\otimes\cdots\otimes z_{ij_{m+n}}.$$

Hence, the $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$–weight space decomposition follows. □

By Theorem 3.6(b), we get linear maps

$$f_{\vec{k}}^{\vec{l}}\colon 1_{\vec{l}}\,\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})1_{\vec{k}} \to \mathrm{Hom}_{\boldsymbol{U}_q(\mathfrak{gl}_N)}\big(\textstyle\bigwedge_q^{\vec{k}_0}\mathbb{C}_q^N\otimes\mathrm{Sym}_q^{\vec{k}_1}\mathbb{C}_q^N, \bigwedge_q^{\vec{l}_0}\mathbb{C}_q^N\otimes\mathrm{Sym}_q^{\vec{l}_1}\mathbb{C}_q^N\big)$$

for any two $\vec{k},\vec{l}\in\mathbb{Z}^{m+n}_{\geq 0}$ such that $\sum_{i=0}^{m+n}k_i=\sum_{i=0}^{m+n}l_i$. Using Theorem 3.6(a), we see that the homomorphisms $f_{\vec{k}}^{\vec{l}}$ are all surjective. Thus, we get the following.

Corollary 3.8 *There exists a full functor $\Phi_{\mathrm{su}}^{m|n}\colon \dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})\to\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$, which we call the super q–Howe functor, given on objects and morphisms by*

$$\Phi_{\mathrm{su}}^{m|n}(\vec{k}) = \textstyle\bigwedge_q^{\vec{k}_0}\mathbb{C}_q^N\otimes\mathrm{Sym}_q^{\vec{k}_1}\mathbb{C}_q^N, \quad \Phi_{\mathrm{su}}^{m|n}(1_{\vec{l}}x1_{\vec{k}}) = f_{\vec{k}}^{\vec{l}}(x).$$

Everything else is sent to zero. □

3.2 The sorted equivalences

In this subsection we construct a full and faithful functor

$$\Gamma_N^{\mathrm{sort}}\colon N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}},$$

where $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ is the sorted web category from Definition 2.5 and $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$ denotes the full subcategory of $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}$ whose objects are sorted as in (3-5).

As already explained in the introduction, we essentially define Γ_N^{sort} such that there is a commuting diagram:

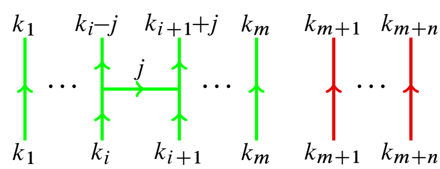
(3-6)
$$\begin{tikzcd}
\dot{U}_q(\mathfrak{gl}_{m|n}) \arrow[r, "\Phi_{\mathrm{su}}^{m|n}"] \arrow[dr, "\Upsilon_{\mathrm{su}}^{m|n}"'] & \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}} \\
& N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \arrow[u, "\Gamma_N^{\mathrm{sort}}"']
\end{tikzcd}$$

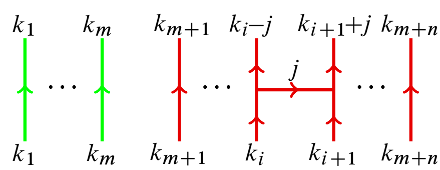
The functor $\Upsilon_{\mathrm{su}}^{m|n}$ is a *ladder functor*, whose definition is motivated by [3, Section 5.1].

Lemma 3.9 *Let $m,n \in \mathbb{Z}_{\geq 0}$. There exists a functor*

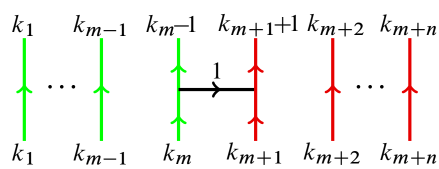
$$\Upsilon_{\mathrm{su}}^{m|n}\colon \dot{U}_q(\mathfrak{gl}_{m|n}) \to N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$$

which sends a $\mathfrak{gl}_{m|n}$–weight $\vec{k} \in \mathbb{Z}_{\geq 0}^{m+n}$ to $((k_1)_g,\dots,(k_m)_g,(k_{m+1})_r,\dots,(k_{m+n})_r)$ in $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ and all other $\mathfrak{gl}_{m|n}$–weights to the zero object. On morphisms, $\Upsilon_{\mathrm{su}}^{m|n}$ is given by

$F_i^{(j)}1_{\vec{k}} \mapsto$ 

$F_i^{(j)}1_{\vec{k}} \mapsto$ 

for $i \in \mathbb{I}_0 - \{m\}$ or $i \in \mathbb{I}_1 - \{m+n\}$, respectively, and

$F_m 1_{\vec{k}} \mapsto$ 

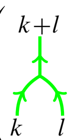
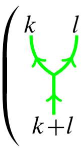
and similarly, but with reversed horizontal orientations, for the generators $E_i^{(j)}1_{\vec{k}}$ and $E_m 1_{\vec{k}}$.

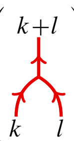
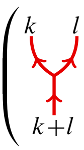
Proof To show that $\Upsilon_{\mathrm{su}}^{m|n}$ is well-defined, it suffices to show that all relations in $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$ are satisfied in N–$\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$. For monochromatic relations we can copy [3, Proposition 5.2.1]. Lemma 2.10 shows that the super relations (3-1) hold in N–$\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$. □

Definition 3.10 (The diagrammatic presentation functor Γ_N^{sort}) We define a functor $\Gamma_N^{\mathrm{sort}}\colon N$–$\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \to \mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$ as follows:

- On objects: to each $\vec{k} = ((k_1)_g, \dots, (k_m)_g, (k_{m+1})_r, \dots, (k_{m+n})_r)$, we assign
$$\Gamma_N^{\mathrm{sort}}(\vec{k}) = \textstyle\bigwedge_q^{\vec{k}_0} \mathbb{C}_q^N \otimes \mathrm{Sym}_q^{\vec{k}_1} \mathbb{C}_q^N,$$
where $\vec{k}_0 = (k_1, \dots, k_m)$ and $\vec{k}_1 = (k_{m+1}, \dots, k_{m+n})$. Moreover, we send the empty tuple to the trivial $\boldsymbol{U}_q(\mathfrak{gl}_N)$–module $\mathbb{C}_q$ and the zero object to the $\boldsymbol{U}_q(\mathfrak{gl}_N)$–module 0.
- On morphisms: we use the functor $\Phi_{\mathrm{su}}^{m|n}$ from Corollary 3.8 to define Γ_N^{sort} on the generating trivalent vertices in N–$\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ (here we assume that the diagrams are the identities outside of the illustrated part). For this, let $i \in \mathbb{I}$ and we use the notation $k = k_i, l = k_{i+1}$ and $(k, l) = (k_1, \dots, k_i = k, k_{i+1} = l, \dots, k_{m+n})$.

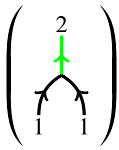
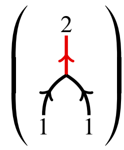
(3-7)

Γ_N^{sort} () $= \Phi_{\mathrm{su}}^{m|n}(E_i^{(l)} 1_{(k,l)})$, Γ_N^{sort} () $= \Phi_{\mathrm{su}}^{m|n}(F_i^{(l)} 1_{(k+l,0)})$,

Γ_N^{sort} () $= \Phi_{\mathrm{su}}^{m|n}(F_i^{(k)} 1_{(k,l)})$, Γ_N^{sort} () $= \Phi_{\mathrm{su}}^{m|n}(E_i^{(k)} 1_{(0,k+l)})$.

Note that these definitions include the mixed case, where we either have $l = 1$ (and colored black) or $k = 1$ (and colored black) and we use the odd generators F_m and E_m. ◇

Remark 3.11 There are certain choices for the images of monochromatic merges and splits, but these choices do not matter; see [25, Remark 2.18]. In contrast, there is no other choice for the mixed merges and splits. For example, take $l = 1$ in the top left in (3-7). The green edge labeled $k+1$ should represent $\bigwedge_q^{k+1} \mathbb{C}_q^N$. Thus, we have to see the top boundary of the left-hand side as $1_{(k+1,0)}$ and not as $1_{(0,k+1)}$, which determines our choices, and similarly for the other mixed generators. For example, if $m = n = 1$, and $k = 1$ or $l = 1$, then

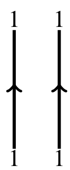
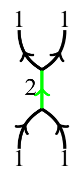
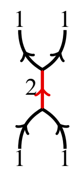
Γ_N^{sort}  $= \Phi_{\mathrm{su}}^{1|1}(E_1 1_{(1,1)}) \neq \Phi_{\mathrm{su}}^{1|1}(F_1 1_{(1,1)}) = \Gamma_N^{\mathrm{sort}}$ .

Lemma 3.12 *Γ_N^{sort} is a well-defined functor $\Gamma_N^{\mathrm{sort}}\colon N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$ making the diagram* (3-6) *commutative.*

Proof First we note that $\Gamma_N^{\mathrm{sort}} \circ \Upsilon_{\mathrm{su}}^{m|n} = \Phi_{\mathrm{su}}^{m|n}$ on generators $F_i^{(j)} 1_{\vec{k}}$ and $F_m 1_{\vec{k}}$ (and analogously for E) with $i \in \mathbb{I} - \{m\}$, $j \in \mathbb{Z}_{\geq 0}$ and $\vec{k} \in \mathbb{Z}^{m+n}$. This follows from the definition of Γ_N^{sort} via $\Phi_{\mathrm{su}}^{m|n}$ and the observation that ladders can be written as compositions of merges and splits; see also [25, Lemma 2.20].

We need to check that the images of the relations from $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ under Γ_N^{sort} hold in $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$. Corollary 3.8 guarantees that all relations in $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$ are induced via $\Phi_{\mathrm{su}}^{m|n}$ from relations in $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$ and the fact that $\Phi_{\mathrm{su}}^{m|n}$ kills certain $\mathfrak{gl}_{m|n}$–weights. It remains to check that the relations in $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ are, likewise, induced via $\Upsilon_{\mathrm{su}}^{m|n}$ from relations in $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$. For the monochromatic and isotopy relations, this follows as in [25, Lemma 2.20].

The dumbbell relation (2-9) can be recovered from $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$ as follows. Without loss of generality we work with $m = n = 1$:

$[2]$  $= \Upsilon_{\mathrm{su}}^{1|1}([2]1_{(1,1)}) = \Upsilon_{\mathrm{su}}^{1|1}(FE1_{(1,1)} + EF1_{(1,1)}) =$  $+$ 

Relation (2-10) is a consequence of killing $\mathfrak{gl}_{m|n}$–weights $\vec{k} = (k_1, \dots, k_{m+n})$, one of whose first m entries is larger than N. $\square$

Lemma 3.13 *The functor $\iota_1^{\infty}\colon N\text{–}\mathbf{Web}_{\mathrm{CKM}} \to N\text{–}\mathbf{Web}_{\mathrm{gr}}$ is faithful.*

Proof By Lemma 3.12 and a comparison of definitions, we have a commuting diagram

$$\begin{tikzcd}
\mathfrak{gl}_N\text{–}\mathbf{Mod}_e \arrow[r, "\iota_{\mathrm{e}}^{\mathrm{es}}"] & \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}} \\
N\text{–}\mathbf{Web}_{\mathrm{CKM}} \arrow[u, "\Gamma_{\mathrm{CKM}}"] \arrow[r, "\iota_1^{\infty}"'] & N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \arrow[u, "\Gamma_N^{\mathrm{sort}}"']
\end{tikzcd}$$

where Γ_{CKM} is the functor considered in [3, Section 3.2] and $\iota_{\mathrm{e}}^{\mathrm{es}}$ is the evident embedding of a full subcategory. Γ_{CKM} is faithful by [3, Theorem 3.3.1] and, thus, ι_1^{∞} is faithful as well. □

Remark 3.14 Let $\mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}})$ be the *additive closure* of $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$: objects are finite, formal direct sums of the objects of $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$ and morphisms are matrices (whose entries are morphisms from $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$). We can extend Γ_N^{sort} additively to a functor

$$\Gamma_N^{\mathrm{sort}}\colon \mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}) \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}},$$

and similarly for Γ_N later on.

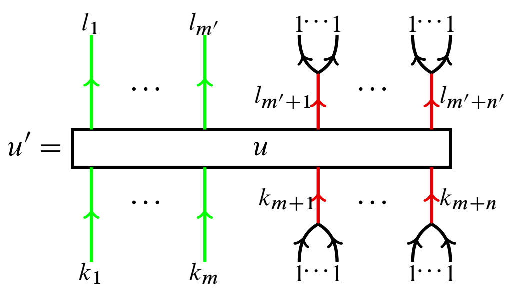
Proposition 3.15 *The functor $\Gamma_N^{\mathrm{sort}}\colon N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$ gives rise to an equivalence of categories $\Gamma_N^{\mathrm{sort}}\colon \mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}) \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$.*

Proof Since $\Gamma_N^{\mathrm{sort}}\colon \mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}) \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$ is well-defined by Lemma 3.12 and Remark 3.14, it remains to show that Γ_N^{sort} is essentially surjective, full and faithful.

Essentially surjective This follows directly from the definitions of Γ_N^{sort}, $N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}$, its additive closure $\mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}})$ and $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}}$.

Full It suffices to verify fullness for morphisms between objects of the form $\vec{k} \in X^{m+n}$, where $X^{m+n} = (X_b \cup X_g)^m \cup (X_b \cup X_r)^n$. That it holds is clear from diagram (3-6), since $\Phi_{\mathrm{su}}^{m|n}$ is full by Corollary 3.8.

Faithful Again it suffices to verify faithfulness for morphisms between objects of the form $\vec{k} \in X^{m+n}$. Given any web $u \in \mathrm{Hom}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}}}(\vec{k}, \vec{l})$ for $\vec{k} \in X^{m+n}$ and $\vec{l} \in X^{m'+n'}$, we can compose u from the bottom and the top with merges and splits, respectively, to obtain



Recall that exploding edges is, by (2-7), a reversible operation. Hence, we have

$$\Gamma_N^{\mathrm{sort}}(u) = \Gamma_N^{\mathrm{sort}}(v) \quad \text{if and only if} \quad \Gamma_N^{\mathrm{sort}}(u') = \Gamma_N^{\mathrm{sort}}(v'),$$

which together with Corollary 2.16 reduces the verification of faithfulness to the case where all web edges are black or green. Such webs lie in $\iota_1^\infty(N\text{–}\mathbf{Web}_{\mathrm{CKM}})$ and faithfulness follows as in the proof of Lemma 3.13. □

3.3 Proofs of the equivalences

Remark 3.16 Recall that the *universal R–matrix* for $\mathfrak{gl}_N$ gives a braiding on the category $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}$ as follows (see eg [29, Chapter XI, Sections 2 and 7]). For any pair of $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules V and W in $\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}$, let $\mathrm{Perm}_{V,W}\colon V \otimes W \to W \otimes V$ be the permutation $\mathrm{Perm}_{V,W}(v \otimes w) = w \otimes v$ and define $\beta^R_{V,W} = \mathrm{Perm}_{V,W} \circ R$. We scale $\beta^R_{V,W}$ as

$$\widetilde{\beta}^R_{V,W} = q^{-\frac{kl}{N}} \beta^R_{V,W}$$

whenever V and W are exterior or symmetric power $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules of exponent k and l, respectively. This induces a scaling $\widetilde{\beta}^R_{V,W}$ of $\beta^R_{V,W}$ for all $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules $V, W \in \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}$. Then $(\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}, \widetilde{\beta}^R_{\cdot,\cdot})$ is a braided monoidal category.

The goal of this subsection is to finally prove our main theorems. To this end, we extend (3-6) to a diagram

$$\begin{tikzcd}
\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n}) \arrow[r, "\Phi^{m|n}_{\mathrm{su}}"] \arrow[dr, "\Upsilon^{m|n}_{\mathrm{su}}"'] & \mathfrak{gl}_N\text{–}\mathbf{Mod}^{\mathrm{sort}}_{\mathrm{es}} \arrow[r, hook, "\iota_{\mathrm{alg}}"] & \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}} \\
& N\text{–}\mathbf{Web}^{\mathrm{sort}}_{\mathrm{gr}} \arrow[u, "\Gamma^{\mathrm{sort}}_N"'] \arrow[r, hook, "\iota_{\mathrm{dia}}"'] & N\text{–}\mathbf{Web}_{\mathrm{gr}} \arrow[u, "\Gamma_N"']
\end{tikzcd} \tag{3-8}$$

where ι_{alg} and ι_{dia} are the evident inclusions of full subcategories. We will define the functor Γ_N such that the diagram (3-8) commutes.

Definition 3.17 (The diagrammatic presentation functor Γ_N) We define a functor $\Gamma_N\colon N\text{–}\mathbf{Web}_{\mathrm{gr}} \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}$ as follows:

- On objects, Γ_N sends an object $\vec{k} \in X^L$ of $N\text{–}\mathbf{Web}_{\mathrm{gr}}$ to the tensor product of exterior and symmetric powers of $\mathbb{C}_q^N$ specified by the entries of $\vec{k}$; green and red integers encode exterior and symmetric powers respectively, and a black entry 1 corresponds to $\mathbb{C}_q^N$ itself.
- On morphisms, for an object $\vec{k} \in X^L$ let $w(\vec{k}) \in S_L$ be a shortest length permutation that sorts green integers in $\vec{k}$ to the left of red integers. We define Γ_N on an arbitrary web $u \in \mathrm{Hom}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k}, \vec{l})$ by precomposing and postcomposing with elementary crossings and the universal R–matrix intertwiners:

$$\Gamma_N(u) = (\widetilde{\beta}^R_{w(\vec{l})})^{-1} \circ \Gamma_N^{\mathrm{sort}}(\beta^\bullet_{w(\vec{l})} \circ u \circ (\beta^\bullet_{w(\vec{k})})^{-1}) \circ \widetilde{\beta}^R_{w(\vec{k})}.$$

Clearly, Γ_N restricts to Γ_N^{sort}. ◇

Lemma 3.18 $\Gamma_N\colon N\text{–}\mathbf{Web}_{\mathrm{gr}} \to \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}$ *is a monoidal functor making* (3-8) *commutative.*

Proof By Lemma 3.12 and the fact that $\beta^{\bullet}_{\cdot,\cdot}$ and $\widetilde{\beta}^{R}_{\cdot,\cdot}$ are braidings (see Proposition 2.22 and Remark 3.16), we see that Γ_N is well-defined. That Γ_N is monoidal and makes (3-8) commutative is clear from its construction. □

Proposition 3.19 *The functor* $\Gamma_N\colon (N\text{–}\mathbf{Web}_{\mathrm{gr}}, \beta^{\bullet}_{\cdot,\cdot}) \to (\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}, \widetilde{\beta}^{R}_{\cdot,\cdot})$ *is a functor of braided monoidal categories.*

Proof By Lemma 3.18, it remains to verify

$$\Gamma_N(\beta^{\bullet}_{\vec{k}\otimes\vec{l}}) = \widetilde{\beta}^{R}_{\Gamma_N(\vec{k}),\Gamma_N(\vec{l})} \quad \text{for all objects } \vec{k} \text{ and } \vec{l} \text{ of } N\text{–}\mathbf{Web}_{\mathrm{gr}}.$$

The green–red symmetry and the fact that the mixed crossings are defined via the monochromatic crossings, together with Corollary 2.24, reduce this problem to the situation studied in [3, Theorem 6.2.1 and Lemma 6.2.2]. It remains to show

$$\Gamma_N(\beta^{g}_{1,1}) = \Gamma_N(\beta^{r}_{1,1}) = \Gamma_N^{\mathrm{sort}}(\beta^{g}_{1,1}) = \Gamma_N^{\mathrm{sort}}(\beta^{r}_{1,1}) = \widetilde{\beta}^{R}_{\mathbb{C}_q^N,\mathbb{C}_q^N}.$$

This follows since $\Gamma_N^{\mathrm{sort}}(\beta^{g}_{1,1}) = \Gamma_N^{\mathrm{sort}}(\beta^{r}_{1,1})$ acts on

$$\mathbb{C}_q^N \otimes \mathbb{C}_q^N \cong {\textstyle\bigwedge}_q^2(\mathbb{C}_q^N) \oplus \mathrm{Sym}_q^2(\mathbb{C}_q^N)$$

as $-q^{-1}$ on the first summand and as q on the second (see Example 2.18). □

Theorem 3.20 (The diagrammatic presentations) *The functor*

$$\Gamma_N\colon (\mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}), \beta^{\bullet}_{\cdot,\cdot}) \to (\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}, \widetilde{\beta}^{R}_{\cdot,\cdot})$$

is an equivalence of braided monoidal categories.

Proof By Proposition 3.19, Γ_N extends to a braided monoidal functor on the additive closure and it remains to show that Γ_N is essentially surjective, full and faithful.

Essentially surjective This follows directly from the definitions; see also Remark 3.14.

Full and faithful As before, it suffices to verify this on morphisms between objects of the form $\vec{k} \in X^L$. Consider the commuting diagram

$$\begin{tikzcd} \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}^{\mathrm{sort}} & \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}} \arrow[l, "\omega_R"'] \\ N\text{–}\mathbf{Web}_{\mathrm{gr}}^{\mathrm{sort}} \arrow[u, "\Gamma_N^{\mathrm{sort}}"] & N\text{–}\mathbf{Web}_{\mathrm{gr}} \arrow[l, "\omega_\bullet"] \arrow[u, "\Gamma_N"'] \end{tikzcd}$$

where ω_R and $\omega_\bullet$ are the functors that order $f \in \mathrm{Hom}_{\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}}(\Gamma_N(\vec{k}), \Gamma_N(\vec{l}))$ and webs $u \in \mathrm{Hom}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k}, \vec{l})$ by using the R–matrix braiding $\widetilde{\beta}^R_{\cdot,\cdot}$ and the braiding $\beta^{\bullet}_{\cdot,\cdot}$, respectively, via a permutation of shortest length. Since sorting is invertible, we get

$$\begin{aligned}\dim\big(\mathrm{Hom}_{\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}}(\Gamma_N(\vec{k}), \Gamma_N(\vec{l}))\big)& \\ &= \dim\big(\mathrm{Hom}_{\mathfrak{gl}_N\text{–}\mathbf{Mod}^{\mathrm{sort}}_{\mathrm{es}}}\big(\Gamma^{\mathrm{sort}}_N(\omega_\bullet(\vec{k})), \Gamma^{\mathrm{sort}}_N(\omega_\bullet(\vec{l}))\big)\big)\\ &= \dim\big(\mathrm{Hom}_{N\text{–}\mathbf{Web}^{\mathrm{sort}}_{\mathrm{gr}}}(\omega_\bullet(\vec{k}), \omega_\bullet(\vec{l}))\big)\\ &= \dim(\mathrm{Hom}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k}, \vec{l})),\end{aligned}$$

where the second equality follows from Proposition 3.15. □

Remark 3.21 For now we restrict ourselves to working with webs with only upward-oriented edges. Downward-oriented edges, as for example in [3], can be used to represent the duals of the $\boldsymbol{U}_q(\mathfrak{gl}_N)$–modules $\textstyle\bigwedge^k_q \mathbb{C}^N_q$ and $\mathrm{Sym}^l_q \mathbb{C}^N_q$. With respect to such an enriched web calculus, the statement of Theorem 3.20 extends to an equivalence of pivotal categories; see [23, Section 6] and Remark 5.12.

Let $\check{\boldsymbol{H}}$ denotes the monoidal, $\mathbb{C}_q$–linear category obtained from the collection $H_\infty(q)$ of Iwahori–Hecke algebras as follows. The objects e and e' of $\check{\boldsymbol{H}}$ are tensor products of Iwahori–Hecke algebra idempotents corresponding to $e_{\mathrm{col}}(\lambda)$ and $e_{\mathrm{row}}(\lambda)$ (as in Definition 2.26) under the isomorphism in Lemma 2.25. The morphism spaces are given by $\mathrm{Hom}_{\check{\boldsymbol{H}}}(e, e') = e' H_\infty(q) e$. The category $\check{\boldsymbol{H}}$ is braided with braiding $\widetilde{\beta}^H_{\cdot,\cdot}$ induced from $H_\infty(q)$.

Theorem 3.22 (The diagrammatic presentation) *For large N the functors Γ_N stabilize to a functor*

$$\Gamma_\infty\colon (\mathbf{Mat}(\infty\text{–}\mathbf{Web}_{\mathrm{gr}}), \beta^{\bullet}_{\cdot,\cdot}) \to (\mathbf{Mat}(\check{\boldsymbol{H}}), \widetilde{\beta}^H_{\cdot,\cdot}),$$

which is an equivalence of braided monoidal categories.

Proof By Schur–Weyl duality (2-14) and by the construction of the categories N–$\mathbf{Web}_{\mathrm{gr}}$ as quotients of ∞–$\mathbf{Web}_{\mathrm{gr}}$, we have quotient functors π^N_∞ and π^N for $N \in \mathbb{Z}_{\geq 0}$ such that

$$\begin{tikzcd} \mathbf{Mat}(\check{\boldsymbol{H}}) \arrow[r, "\pi^N"] & \mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}} \\ \mathbf{Mat}(\infty\text{–}\mathbf{Web}_{\mathrm{gr}}) \arrow[u, "\Gamma_\infty"] \arrow[r, "\pi^N_\infty"'] & \mathbf{Mat}(N\text{–}\mathbf{Web}_{\mathrm{gr}}) \arrow[u, "\Gamma_N"'] \end{tikzcd} \tag{3-9}$$

commutes. Here the functor Γ_∞ is an idempotented version of the inverse of the isomorphism $\Phi^\infty_{q\mathrm{SW}}$ from Lemma 2.25.

Fix two objects $\vec{k}\in X^L$ and $\vec{l}\in X^L$ of ∞–$\mathbf{Web}_{\mathrm{gr}}$ and suppose that N is greater than the sum of the integer values of the entries of $\vec{k}$ (ignoring their colors). Then, by (2-14), Theorem 3.20, the commutativity of (3-9) and the fullness of π^N_∞, we have

$$\begin{aligned}\dim\bigl(\mathrm{Hom}_{\check{\boldsymbol{H}}}(\Gamma_\infty(\vec{k}),\Gamma_\infty(\vec{l}))\bigr) &= \dim\bigl(\mathrm{Hom}_{\mathfrak{gl}_N\text{–}\mathbf{Mod}_{\mathrm{es}}}\bigl(\pi^N(\Gamma_\infty(\vec{k})),\pi^N(\Gamma_\infty(\vec{l}))\bigr)\bigr)\\ &= \dim\bigl(\mathrm{Hom}_{N\text{–}\mathbf{Web}_{\mathrm{gr}}}(\pi^N_\infty(\vec{k}),\pi^N_\infty(\vec{l}))\bigr)\\ &= \dim(\mathrm{Hom}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{k},\vec{l})).\end{aligned}$$

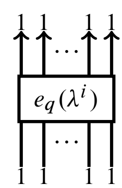
Γ_∞ is clearly essentially surjective and a braided monoidal functor, and the theorem follows. $\square$

4 Applications

In this section we write $\mathcal{L}_D$ for *diagrams of framed, oriented links* $\mathcal{L}$, b^K_D for *diagrams of braids in* K *strands* and $\overline{b}^K_D$ for *closures* of such braid diagrams. We consider labelings of the connected components of $\mathcal{L}$ and of braids by Young diagrams λ^i. If $\mathcal{L}$ is a d–component link, then we write $\mathcal{L}(\vec{\lambda})$ for its labeling by a vector of Young diagrams $\vec{\lambda}=(\lambda^1,\dots,\lambda^d)$, and use an analogous notation for labeled link and braid diagrams. If not mentioned otherwise, then all appearing links and related concepts are assumed to be framed and oriented from now on.

Let $\mathcal{L}_D(\vec{\lambda})=\overline{b}^K_D(\vec{\lambda})$ be a diagram of a framed, oriented, labeled link given as a braid closure. The following process associates to $b^K_D(\vec{\lambda})$ an element $p_{K'}(\widetilde{b}^{K'}_D)e_q(\vec{\lambda})$ of $H_{K'}(q)\cong \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K}')$:

$$\underset{\lambda^i\in\Lambda^+(K_i)}{\lambda^i\uparrow}\xrightarrow{\text{cable}}\underset{K_i\text{ strands}}{\uparrow\uparrow\cdots\uparrow\uparrow}\xrightarrow{p_{K_i}(\cdot)} p_{K_i}\left(\underset{K_i\text{ strands}}{\uparrow\uparrow\cdots\uparrow\uparrow}\right)e_q(\lambda^i)=$$

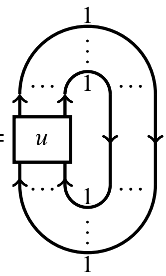


where the last equality follows from Lemma 2.25 and we write p_{K_i} for the Iwahori–Hecke algebra representation of the braid group on K_i strands. The first step replaces strands labeled by a Young diagram λ^i with K_i nodes in the braid diagram b^K_D by K_i parallel strands. This results in a new braid $\widetilde{b}^{K'}_D$, where K' indicates the number of strands. In the second step this cabled braid is interpreted as an element of the Iwahori–Hecke algebra, or, equivalently, as a web in ∞–$\mathbf{Web}_{\mathrm{gr}}$, with an idempotent $e_q(\lambda^i)$ placed on the cable of each previously λ^i labeled strand.

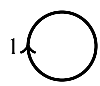
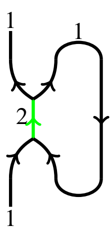
4.1 The colored HOMFLY–PT polynomial via ∞–$\mathbf{Web}_{\mathrm{gr}}$

In this subsection we work over the ground field $\mathbb{C}_{a,q}=\mathbb{C}_q(a)$, with a being a generic parameter. We will use the $\mathbb{C}_{a,q}$–valued *Jones–Ocneanu trace* $\mathrm{tr}(\cdot)$ on the direct sum of all Iwahori–Hecke algebras $H_\infty(q) = \bigoplus_{K \in \mathbb{Z}_{\geq 0}} H_K(q)$. The definition of $\mathrm{tr}(\cdot)$ can be found in [13, Section 5] (which can be easily adapted to our notation). We will use it in the form of the following lemma.

Lemma 4.1 *Given a web $u \in \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$,*

$\mathrm{tr}(u) =$  $\in \mathbb{C}_{a,q}$,

where the closed diagram can be evaluated by using the relations in ∞–$\mathbf{Web}_{\mathrm{gr}}$ and, additionally,

(4-1)  $= \dfrac{a - a^{-1}}{q - q^{-1}}$,  $= \dfrac{aq^{-1} - a^{-1}q}{q - q^{-1}}$ 

Proof By Proposition 2.15 and Corollary 2.13: any given web $u \in \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$ can be expressed using black or green edges with labels at most 2. Using Lemma 2.25 and additionally [24, Section 4.2], where Rasmussen's singular crossings correspond to green dumbbells with label 2, provides the statement. Note that Rasmussen's relations II and III are already part of our diagrammatic calculus. □

Definition 4.2 (The colored HOMFLY–PT polynomial) Let $\mathcal{L}_D(\vec{\lambda}) = \bar{b}_D^K(\vec{\lambda})$ be a diagram of a framed, oriented, labeled link $\mathcal{L}(\vec{\lambda})$ given as a braid closure.

The *colored* HOMFLY–PT *polynomial* of $\mathcal{L}(\vec{\lambda})$, denoted by $\mathcal{P}^{a,q}(\mathcal{L}(\vec{\lambda}))$, is defined via

$$\mathcal{P}^{a,q}(\mathcal{L}(\vec{\lambda})) = \mathrm{tr}(p_{K'}(\tilde{b}_D^{K'})e_q(\vec{\lambda})) \in \mathbb{C}_{a,q},$$

where $e_q(\vec{\lambda})$ is a tensor product of the $e_q(\lambda^i)$, as described above. ◇

This polynomial is independent of all choices involved and an invariant of framed, oriented, colored links. Up to different conventions, this is shown for example in [17, Corollary 4.5].

Remark 4.3 In fact, Definition 4.2 gives the framing dependent, unnormalized version of the colored HOMFLY–PT polynomial. As usual, the polynomial can be normalized by fixing the value of the unknot to be 1 (instead of $(a-a^{-1})/(q-q^{-1})$ as in our convention) and one can get rid of the framing dependence by scaling with a factor coming from Reidemeister 1 moves; see for example [13, Definition 6.1]. We suppress these distinctions in the following.

Note that Lemma 4.1 provides a method to calculate the colored HOMFLY–PT polynomials $\mathcal{P}^{a,q}(\cdot)$ using the web category ∞–$\mathbf{Web}_{\mathrm{gr}}$.

Proposition 4.4 (The colored HOMFLY–PT symmetry) *We have*

$$\mathcal{P}^{a,q}(\mathcal{L}(\vec{\lambda})) = (-1)^c \mathcal{P}^{a,q^{-1}}(\mathcal{L}(\vec{\lambda}^{\mathrm{T}})),$$

where $\vec{\lambda}^{\mathrm{T}} = ((\lambda^1)^{\mathrm{T}}, \dots, (\lambda^d)^{\mathrm{T}})$ *and* c *is the sum of the number of nodes in the* λ^i *for* $1 \le i \le d$.

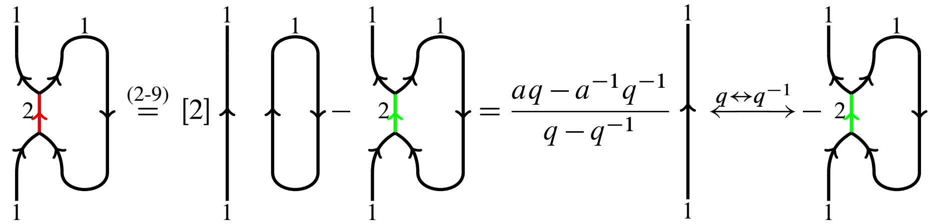
This symmetry is not new: it can be deduced from [19, Section 9] and has been studied in [18; 6, Proposition 4.4]. In our framework it follows directly from the green–red symmetry in ∞–$\mathbf{Web}_{\mathrm{gr}}$.

Proof We only give a proof for the case of knots $\mathcal{K}$. The proof for links is analogous, but the notation is more involved. We denote by I_{gr} the involution on ∞–$\mathbf{Web}_{\mathrm{gr}}$ given by the green–red symmetry, and by I_q the involution on $\mathbb{C}_{a,q}$ which inverts the variable q.

Claim *For* $u \in \mathrm{End}_{\infty\text{–}\mathbf{Web}_{\mathrm{gr}}}(\vec{K})$ *we have*

$$\mathrm{tr}(u) = (-1)^K I_q\big(\mathrm{tr}(I_q(I_{\mathrm{gr}}(u)))\big). \tag{4-2}$$

It suffices to prove $\mathrm{tr}(u) = (-1)^K I_q(\mathrm{tr}(I_{\mathrm{gr}}(u)))$ in the case where u is a primitive web (a morphism that consists of a single web with coefficient 1, which is thus invariant under I_q). In Lemma 4.1 we have met evaluation relations for monochromatic green webs of edge label at most 2, but clearly analogous relations can be derived for red and mixed webs. In fact, all necessary evaluation relations are invariant under I_{gr} and I_q, except the two relations in (4-1). The circle relation is I_{gr}–invariant, but acquires a sign under I_q. The following computation shows that the green and red bubble relations also respect (4-2):



We note that in the computation of $\mathrm{tr}(u)$ via Lemma 4.1 strands can only be removed by circle moves and bubble moves. Both of these acquire a sign under I_q, which causes the factor $(-1)^K$ in (4-2). This proves the claim.

Let b_D^K be a braid diagram that closes to a diagram of $\mathcal{K}$ and suppose that $\mathcal{K}$ is labeled by a Young diagram λ of with L nodes. Let $\widetilde{b}_D^{KL}$ be the L–fold cable of the braid diagram b_D^K.

Now we have

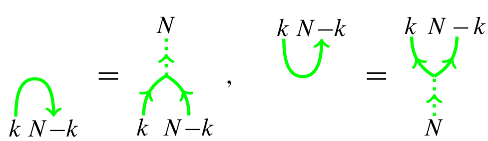
$$\begin{aligned}\mathcal{P}^{a,q}(\mathcal{K}(\lambda)) &= \mathrm{tr}(p_{KL}(\widetilde{b}_D^{KL})e_q(\lambda)^{\otimes K})\\ &= (-1)^{KL} I_q\big(\mathrm{tr}\big(I_q(I_{\mathrm{gr}}(p_{KL}(\widetilde{b}_D^{KL})e_q(\lambda)^{\otimes K}))\big)\big)\\ &= (-1)^{KL+\mathrm{cr}L^2} I_q(\mathrm{tr}(p_{KL}(\widetilde{b}_D^{KL})e_q(\lambda^{\mathrm{T}})^{\otimes K})) = (-1)^L \mathcal{P}^{a,q^{-1}}(\mathcal{K}(\lambda^{\mathrm{T}})),\end{aligned}$$

where cr is the number of crossings of b_D^K. Here we have used (4-2) and that $I_q I_{\mathrm{gr}}$ acts as -1 on black crossings — see Example 2.18 — while sending $e_q(\lambda)$ to $e_q(\lambda^{\mathrm{T}})$ plus a commutator (which is zero in the trace), see Lemma 2.28. Moreover, $(-1)^{KL+\mathrm{cr}L^2} = (-1)^L$ since $\mathrm{cr} \equiv K-1 \bmod 2$ as b_D^K closes into a knot. □

4.2 The colored $\mathfrak{sl}_N$–link polynomials via the categories N–$\mathbf{Web}_{\mathrm{gr}}$

Recall that the *colored Reshetikhin–Turaev $\mathfrak{sl}_N$–link polynomial* $\mathcal{RT}^{q^N,q}(\mathcal{L}(\vec{\lambda}))$ are determined by the corresponding colored HOMFLY–PT polynomials $\mathcal{P}^{a,q}(\mathcal{L}(\vec{\lambda}))$ by specializing $a = q^N$. Alternatively, they can be computed directly inside the categories N–$\mathbf{Web}_{\mathrm{gr}}$ from a framed, oriented, labeled link diagram as follows:

- First we replace all λ–labeled strands in the link diagram by cables equipped with the diagrammatic idempotent $e_q(\lambda)$, written in monochromatic green webs.
- The resulting diagram will contain downward-oriented green edges of label k, which we replace by upward-oriented green edges of label $N-k$. Simultaneously, caps and cups are replaced by splits and merges



- The result is a morphism in N–$\mathbf{Web}_{\mathrm{gr}}$ between objects consisting only of entries 0 and N_g. It follows from Theorem 3.20 that this Hom–space is one-dimensional. Thus, the framed, oriented, labeled link diagram determines a polynomial, which is the desired colored Reshetikhin–Turaev $\mathfrak{sl}_N$–link polynomial.

Recall from Remark 1.1 that this approach relies on the fact that $\mathfrak{sl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ contains the duality isomorphisms $\bigwedge_q^k \mathbb{C}_q^N \cong \big(\bigwedge_q^{N-k} \mathbb{C}_q^N\big)^*$. In Remark 5.12 we sketch how to include duals in diagrammatic presentations of $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$ and $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$ and, thus, to compute the corresponding Reshetikhin–Turaev $\mathfrak{gl}_N$ or $\mathfrak{gl}_{m|n}$–link invariants.

5 Generalization to webs for $\mathfrak{gl}_{N|M}$

We now give a diagrammatic presentation of $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$, the (additive closure of the) braided monoidal category of $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$–modules tensor generated by the exterior $\bigwedge_q^k \mathbb{C}_q^{N|M}$ and the symmetric $\mathrm{Sym}_q^l \mathbb{C}_q^{N|M}$ powers of the vector representation $\mathbb{C}_q^{N|M}$ of $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$. The diagrammatic presentation is given by the following quotient of ∞–$\mathbf{Web}_{\mathrm{gr}}$.

Definition 5.1 The category $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ is the quotient category obtained from ∞–$\mathbf{Web}_{\mathrm{gr}}$ by imposing the *not-a-hook relation*, that is,

$$e_q(\mathrm{box}_{N+1,M+1}) = 0,$$

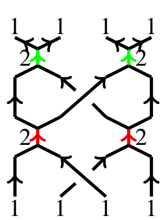
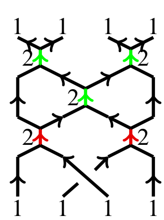
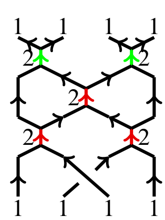
where $\mathrm{box}_{N+1,M+1}$ is the box-shaped Young diagram with $N+1$ rows and $M+1$ columns. $\diamond$

Note that $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ inherits the braiding $\beta^{\bullet}_{\cdot,\cdot}$ from ∞–$\mathbf{Web}_{\mathrm{gr}}$.

Example 5.2 If we take $M = 0$, then $\mathrm{box}_{N+1,1}$ is a column Young diagram with $N+1$ nodes and the corresponding not-a-hook relation is just the exterior relation (2-10). In this case we have that $N|0$–$\mathbf{Web}_{\mathrm{gr}}$ is N–$\mathbf{Web}_{\mathrm{gr}}$ and $\mathfrak{gl}_{N|0}$–$\mathbf{Mod}_{\mathrm{es}}$ is isomorphic to $\mathfrak{gl}_N$–$\mathbf{Mod}_{\mathrm{es}}$. $\triangleleft$

Example 5.3 If we take $M = N = 1$, then we have

$$\tilde{e}_q(\mathrm{box}_{2,2}) = \tilde{e}_q\left(\begin{array}{|c|c|}\hline & \\ \hline & \\ \hline \end{array}\right)$$

$= \frac{1}{[2]^4}$  $= -\frac{1}{[2]^4}$  $= \frac{1}{[2]^4}$ 

It is easy to see that $e_q(\mathrm{box}_{2,2}) = 0$ is equivalent to the relations [27, (3.3.13a) and (3.3.13b)], [9, Section 3.6] and [23, Corollary 6.18], which are used to describe the "purely exterior" representation category $\mathfrak{gl}_{1|1}$–$\mathbf{Mod}_{\mathrm{e}}$. This category could be presented as monochromatic green subcategory of $1|1$–$\mathbf{Web}_{\mathrm{gr}}$, defined analogously as in Definition 2.7. $\triangleleft$

To prove that $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ gives a diagrammatic presentation of $\mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$, we use a version of super q–Howe duality between $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ and $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$. For this, we say a dominant integral $\mathfrak{gl}_{m|n}$–weight λ is *$(m|n, M|N)$–supported* if it corresponds to a Young diagram which is simultaneously an $(m|n)$–hook as well as an $(M|N)$–hook.[9]

Theorem 5.4 (Super q–Howe duality, super–super version) *We have the following:*

(a) *Let $K \in \mathbb{Z}_{\geq 0}$. The actions of $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$ and $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$ on $\bigwedge_q^K(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M})$ commute and generate each others commutant.*

(b) *There exists an isomorphism*

$$\textstyle\bigwedge_q^\bullet(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M}) \cong (\bigwedge_q^\bullet \mathbb{C}_q^{N|M})^{\otimes m} \otimes (\mathrm{Sym}_q^\bullet \mathbb{C}_q^{N|M})^{\otimes n}$$

of $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$–modules under which the $\vec{k}$–weight space of $\bigwedge_q^\bullet(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M})$ (considered as a $\boldsymbol{U}_q(\mathfrak{gl}_{m|n})$–module) is identified with

$$\textstyle\bigwedge_q^{\vec{k}_0} \mathbb{C}_q^{N|M} \otimes \mathrm{Sym}_q^{\vec{k}_1} \mathbb{C}_q^{N|M} = \bigwedge_q^{k_1} \mathbb{C}_q^{N|M} \otimes \cdots \otimes \bigwedge_q^{k_m} \mathbb{C}_q^{N|M} \otimes \mathrm{Sym}_q^{k_{m+1}} \mathbb{C}_q^{N|M} \otimes \cdots \otimes \mathrm{Sym}_q^{k_{m+n}} \mathbb{C}_q^{N|M}.$$

Here $\vec{k} = (k_1, \ldots, k_{m+n})$, $\vec{k}_0 = (k_1, \ldots, k_m)$ and $\vec{k}_1 = (k_{m+1}, \ldots, k_{m+n})$.

(c) *As $\boldsymbol{U}_q(\mathfrak{gl}_{m|n}) \otimes \boldsymbol{U}_q(\mathfrak{gl}_{N|M})$–modules, we have a decomposition of the form*

$$\textstyle\bigwedge_q^K(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M}) \cong \bigoplus_\lambda L_{m|n}(\lambda) \otimes L_{N|M}(\lambda^{\mathrm{T}}),$$

where we sum over all $(m|n, M|N)$–supported $\mathfrak{gl}_{m|n}$–weights λ whose entries sum up to K. This induces a decomposition

$$\textstyle\bigwedge_q^\bullet(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M}) \cong \bigoplus_\lambda L_{m|n}(\lambda) \otimes L_{N|M}(\lambda^{\mathrm{T}}),$$

where we sum over all $(m|n, M|N)$–supported $\mathfrak{gl}_{m|n}$–weights λ.

[9] This is really intended to be $(M|N)$.

Proof As before, (a) and (c) are proven in [23, Theorem 4.2] and only (b) remains to be verified. This works similarly as in the proof of Theorem 3.6 and is left to the reader. For a nonquantized version see [28, Proposition 2.2]. □

In the statement of this theorem, $\bigwedge_q^k \mathbb{C}_q^{N|M}$, $\mathrm{Sym}_q^l \mathbb{C}_q^{N|M}$ and $\bigwedge_q^K (\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M})$ are defined similarly as in Section 3.1; see also [23, Section 3]. As before we then get:

Corollary 5.5 *There exists a full functor $\Phi_{\mathrm{su}}^{m|n}\colon \dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n}) \to \mathfrak{gl}_{N|M}$–$\mathbf{Mod}_{\mathrm{es}}$, which we again call the super q–Howe functor, given on objects and morphisms by*

$$\Phi_{\mathrm{su}}^{m|n}(\vec{k}) = \textstyle\bigwedge_q^{\vec{k}_0} \mathbb{C}_q^{N|M} \otimes \mathrm{Sym}_q^{\vec{k}_1} \mathbb{C}_q^{N|M}, \quad \Phi_{\mathrm{su}}^{m|n}(1_{\vec{l}}\, x 1_{\vec{k}}) = f_{\vec{k}}^{\vec{l}}(x).$$

Everything else is sent to zero. □

In what follows, we denote by $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})^{\geq 0}$ the quotient of $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$ obtained by killing all $\mathfrak{gl}_{m|n}$–weights with negative entries.

Corollary 5.6 *The super q–Howe functor $\Phi_{\mathrm{su}}^{m|n}$ from Corollary 5.5 induces an algebra epimorphism (denoted by the same symbol) as in the diagram:*

$$\begin{tikzcd}
\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})^{\geq 0} \arrow[r, "\cong"] \arrow[d, two heads, "\Phi_{\mathrm{su}}^{m|n}"'] & \bigoplus_{\substack{(m|n)\text{-}\\ \text{hooks } \lambda}} \mathrm{End}_{\mathbb{C}_q}(L_{m|n}(\lambda)) \arrow[d, two heads, "\pi"] \\
\mathrm{End}_{\boldsymbol{U}_q(\mathfrak{gl}_{N|M})}\big(\bigwedge_q^{\bullet}(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M})\big) \arrow[r, "\cong"'] & \bigoplus_{\substack{(m|n,M|N)\text{-}\\ \text{supported } \lambda}} \mathrm{End}_{\mathbb{C}_q}(L_{m|n}(\lambda))
\end{tikzcd}$$

Under Artin–Wedderburn decompositions, $\Phi_{\mathrm{su}}^{m|n}$ corresponds to an algebra epimorphism π, which acts on the summand $\mathrm{End}_{\mathbb{C}_q}(L_{m|n}(\lambda))$ either as an isomorphism or as zero, depending on whether the Young diagram λ is $(m|n, M|N)$–supported or not.

Proof First, note that by Theorem 3.22, $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})^{\geq 0}$ is isomorphic to $\check{\boldsymbol{H}}_{m+n}^{\mathrm{sort}}$, the sorted version of $\check{\boldsymbol{H}}$ with exactly m exterior strands and n symmetric strands. The Artin–Wedderburn decomposition in the top row of the diagram is then given in [21, Theorem 5.1]. The bottom Artin–Wedderburn decomposition follows directly from part (c) of Theorem 5.4. □

Remark 5.7 We obtain from Corollary 5.6 an alternative proof of the presentation of the q–Schur superalgebra $S_q(N|M, K) \cong \mathrm{End}_{H_K(q)}((\mathbb{C}_q^{N|M})^{\otimes K})$ from [5, Theorem 3.13.1].

Lemma 5.8 *Under the correspondence*

$$\check{\boldsymbol{H}}^{\mathrm{sort}}_{m+n} \xleftrightarrow{\cong} \dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})^{\geq 0} \xleftrightarrow{\cong} \bigoplus_{(m|n)\text{–hooks } \lambda} \mathrm{End}_{\mathbb{C}_q}(L_{m|n}(\lambda)),$$

the kernel of the super q–Howe functor $\Phi^{m|n}_{\mathrm{su}}$ from Corollary 5.5 is given by the tensor ideal I_{box} in $\check{\boldsymbol{H}}^{\mathrm{sort}}_{m+n}$ generated by the primitive idempotent $e_q(\mathrm{box}_{N+1,M+1})$.

Proof From the right isomorphism we know that the kernel of $\Phi^{m|n}_{\mathrm{su}}$ is generated by all $e_q(\lambda^{\mathrm{T}})$ where λ is an $(m|n)$–hook, but not an $(M|N)$–hook. Every such λ corresponds to a simple $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$–module which appears in a tensor product $L_{N|M}((\mathrm{box}_{N+1,M+1})^{\mathrm{T}}) \otimes (\mathbb{C}_q^{N|M})^{\otimes K}$ for some $K \in \mathbb{Z}_{\geq 0}$. Accordingly, $e_q(\lambda^T)$ is contained in the ideal I_{box}. □

Proposition 5.9 *There is an equivalence of categories*

$$\mathbf{Mat}(N|M\text{–}\mathbf{Web}^{\mathrm{sort}}_{\mathrm{gr}}) \cong \mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}^{\mathrm{sort}}_{\mathrm{es}}.$$

Proof Lemma 5.8 shows that the sorted web category $N|M\text{–}\mathbf{Web}^{\mathrm{sort}}_{m+n}$, in which webs have m green and n red boundary points both on the bottom and on the top, is equivalent to $\mathrm{End}_{\boldsymbol{U}_q(\mathfrak{gl}_{N|M})}\big(\bigwedge^{\bullet}_q(\mathbb{C}_q^{m|n} \otimes \mathbb{C}_q^{N|M})\big)$, considered as a category. Via the $\dot{\boldsymbol{U}}_q(\mathfrak{gl}_{m|n})$–weight space decomposition in Theorem 5.4(b), $N|M\text{–}\mathbf{Web}^{\mathrm{sort}}_{m+n}$ gives a presentation of the morphism spaces in $\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}^{\mathrm{sort}}_{\mathrm{es}}$ between objects of the form

$$\textstyle\bigwedge_q^{k_1} \mathbb{C}_q^{N|M} \otimes \cdots \otimes \bigwedge_q^{k_m} \mathbb{C}_q^{N|M} \otimes \mathrm{Sym}_q^{k_{m+1}} \mathbb{C}_q^{N|M} \otimes \cdots \otimes \mathrm{Sym}_q^{k_{m+n}} \mathbb{C}_q^{N|M}.$$

Any object in $\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}^{\mathrm{sort}}_{\mathrm{es}}$ is a formal sum of such objects for suitable $m, n \in \mathbb{Z}_{\geq 0}$, and the conclusion follows. □

Remark 5.10 Recall that $\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}_{\mathrm{es}}$ is a braided monoidal category, where the braiding $\beta^R_{\cdot,\cdot}$ is given by the *universal R–matrix for* $\mathfrak{gl}_{N|M}$; see [32]. As before, we use a rescaled braiding $\widetilde{\beta}^R_{\cdot,\cdot}$, where we follow the conventions from [23, (3.12)] except that we substitute q by q^{-1} in their formulas. In particular, $\widetilde{\beta}^R_{\mathbb{C}_q^{N|M},\mathbb{C}_q^{N|M}}$ acts as $-q^{-1}$ on $\bigwedge_q^2 \mathbb{C}_q^{N|M}$ and as q on $\mathrm{Sym}_q^2 \mathbb{C}_q^{N|M}$.

Theorem 5.11 (The diagrammatic presentation) *There is an equivalence of braided monoidal categories*

$$(\mathbf{Mat}(N|M\text{–}\mathbf{Web}_{\mathrm{gr}}), \beta^{\bullet}_{\cdot,\cdot}) \cong (\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}_{\mathrm{es}}, \beta^R_{\cdot,\cdot}).$$

Proof The equivalence from Proposition 5.9 can be extended to a monoidal functor between the categories $\mathbf{Mat}(N|M\text{–}\mathbf{Web}_{\mathrm{gr}})$ and $\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}_{\mathrm{es}}$ as in Definition 3.17. We can also copy the proof of Proposition 3.19, where we use Remark 5.10 to prove that this functor respects the braiding. Equivalence via this functor follows then as in Theorem 3.20. □

Remark 5.12 In [23, Section 6] the authors show how to extend a diagrammatic presentation of $\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}_{\mathrm{e}}$ to diagrammatically encode the full subcategory of $U_q(\mathfrak{gl}_{N|M})$–modules tensor generated by exterior powers and their duals. Graphically, this involves the introduction of additional objects corresponding to the duals of exterior powers, downward-oriented edges (to represent identity morphisms on duals) and cap and cup webs (which represent coevaluation and evaluation morphisms). Additional web relations including analogues of (4-1) are introduced to encode basic relationships between exterior powers and their duals. The extension of the diagrammatic presentation to include duals is then tautological and [23, Theorem 6.5 and Proposition 6.16] show that the extended presentation functor is fully faithful.

They further show in [23, Proposition 6.15] that their graphical calculus allows the computation of the Reshetikhin–Turaev $\mathfrak{gl}_{N|M}$–tangle invariants for tangles labeled with exterior powers of the vector representation.

The same *spiderization strategy* — with minimal changes in proofs — gives an extension of our diagrammatic presentation $N|M\text{–}\mathbf{Web}_{\mathrm{gr}}$ of $\mathfrak{gl}_{N|M}\text{–}\mathbf{Mod}_{\mathrm{es}}$ to one for the full subcategory of $U_q(\mathfrak{gl}_{N|M})$–modules tensor generated by exterior and symmetric powers and their duals. This spiderized green–red web category directly allows the computation of Reshetikhin–Turaev $\mathfrak{gl}_{N|M}$–tangle invariants for tangles labeled with exterior as well as symmetric powers of the vector representation. The cabling strategy from Section 4 can then be used to compute these invariants with respect to arbitrary irreducible representations.

Lastly, we have a direct consequence of the discussion in this section and Proposition 4.4. It is based on the facts that $N|M\text{–}\mathbf{Web}_{\mathrm{gr}}$ is defined as a quotient of $\infty\text{–}\mathbf{Web}_{\mathrm{gr}}$ and that the spiderization in [23, Section 6] respects the specialization $a = q^{N-M}$ of the relations (4-1), which are sufficient to compute colored HOMFLY–PT polynomials of braid closures.

Corollary 5.13 *We have:*

(1) *The Reshetikhin–Turaev $\mathfrak{gl}_{N|M}$–tangle invariant of a labeled tangle depends only on $N - M$. In the case of a labeled link, it agrees with the specialization $a = q^{N-M}$ of the corresponding colored* HOMFLY–PT *polynomial.*

(2) *The green–red symmetry on* ∞–$\mathbf{Web}_{\mathrm{gr}}$ *descends to a symmetry between the categories* $N|M$–$\mathbf{Web}_{\mathrm{gr}}$ *and* $M|N$–$\mathbf{Web}_{\mathrm{gr}}$*. Hence, there is a symmetry between the representation categories of* $\boldsymbol{U}_q(\mathfrak{gl}_{N|M})$ *and* $\boldsymbol{U}_q(\mathfrak{gl}_{M|N})$ *that transposes Young diagrams indexing irreducibles.*

(3) *The symmetry of* HOMFLY–PT *polynomials described in Proposition 4.4 is a stabilized version of the symmetry between colored Reshetikhin–Turaev* $\mathfrak{gl}_{N|M}$*–link invariants and* $\mathfrak{gl}_{M|N}$*–link invariants which transposes Young diagrams and inverts* q*.* □

This confirms decategorified analogues of predictions about relationships between colored HOMFLY–PT homology and conjectural colored $\mathfrak{gl}_{N|M}$–link homologies; see [7].

References

[1] **A K Aiston**, **H R Morton**, *Idempotents of Hecke algebras of type* A, J. Knot Theory Ramifications 7 (1998) 463–487 MR

[2] **A Berenstein**, **S Zwicknagl**, *Braided symmetric and exterior algebras*, Trans. Amer. Math. Soc. 360 (2008) 3429–3472 MR

[3] **S Cautis**, **J Kamnitzer**, **S Morrison**, *Webs and quantum skew Howe duality*, Math. Ann. 360 (2014) 351–390 MR

[4] **S-J Cheng**, **W Wang**, *Howe duality for Lie superalgebras*, Compositio Math. 128 (2001) 55–94 MR

[5] **H El Turkey**, **J R Kujawa**, *Presenting Schur superalgebras*, Pacific J. Math. 262 (2013) 285–316 MR

[6] **N Geer**, **B Patureau-Mirand**, *On the colored HOMFLY–PT, multivariable and Kashaev link invariants*, Commun. Contemp. Math. 10 (2008) 993–1011 MR

[7] **E Gorsky**, **S Gukov**, **M Stošić**, *Quadruply-graded colored homology of knots*, preprint (2013) arXiv

[8] **J Grant**, *A categorification of the skew Howe action on a representation category of* $U_q(\mathfrak{gl}(m|n))$, preprint (2015) arXiv

[9] **J Grant**, *A generators and relations description of a representation category of* $U_q(\mathfrak{gl}(1|1))$, Algebr. Geom. Topol. 16 (2016) 509–539 MR

[10] **S Gukov**, **M Stošić**, *Homological algebra of knots and BPS states*, from "Proceedings of the Freedman Fest" (R Kirby, V Krushkal, Z Wang, editors), Geom. Topol. Monogr. 18, Geom. Topol. Publ., Coventry (2012) 309–367 MR

[11] **A Gyoja**, *A* q*–analogue of Young symmetrizer*, Osaka J. Math. 23 (1986) 841–852 MR

[12] **R Howe**, *Perspectives on invariant theory: Schur duality, multiplicity-free actions and beyond*, from "The Schur lectures" (I Piatetski-Shapiro, S Gelbart, editors), Israel Math. Conf. Proc. 8, Bar-Ilan Univ., Ramat Gan (1995) 1–182 MR

[13] **V F R Jones**, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. 126 (1987) 335–388 MR

[14] **L H Kauffman**, **S L Lins**, *Temperley–Lieb recoupling theory and invariants of* 3*–manifolds*, Annals of Mathematics Studies 134, Princeton Univ. Press (1994) MR

[15] **G Kuperberg**, *Spiders for rank* 2 *Lie algebras*, Comm. Math. Phys. 180 (1996) 109–151 MR

[16] **J-H Kwon**, *Crystal bases of q–deformed Kac modules over the quantum superalgebras* $U_q(\mathfrak{gl}(m|n))$, Int. Math. Res. Not. 2014 (2014) 512–550 MR

[17] **X-S Lin**, **H Zheng**, *On the Hecke algebras and the colored HOMFLY polynomial*, Trans. Amer. Math. Soc. 362 (2010) 1–18 MR

[18] **K Liu**, **P Peng**, *Proof of the Labastida–Mariño–Ooguri–Vafa conjecture*, J. Differential Geom. 85 (2010) 479–525 MR

[19] **S Lukac**, *Homfly skeins and the Hopf link*, PhD thesis, University of Liverpool (2001) Available at `http://www.liv.ac.uk/~su14/papers/lukacthesis.pdf`

[20] **G Lusztig**, *Introduction to quantum groups*, Progress in Mathematics 110, Birkhäuser, Boston, MA (1993) MR

[21] **H Mitsuhashi**, *Schur–Weyl reciprocity between the quantum superalgebra and the Iwahori–Hecke algebra*, Algebr. Represent. Theory 9 (2006) 309–322 MR

[22] **H Murakami**, **T Ohtsuki**, **S Yamada**, *Homfly polynomial via an invariant of colored plane graphs*, Enseign. Math. 44 (1998) 325–360 MR

[23] **H Queffelec**, **A Sartori**, *Mixed quantum skew Howe duality and link invariants of type* A, preprint (2015) arXiv

[24] **J Rasmussen**, *Some differentials on Khovanov–Rozansky homology*, Geom. Topol. 19 (2015) 3031–3104 MR

[25] **D E V Rose**, **D Tubbenhauer**, *Symmetric webs, Jones–Wenzl recursions, and q–Howe duality*, Int. Math. Res. Not. 2016 (2016) 5249–5290 MR

[26] **G Rumer**, **E Teller**, **H Weyl**, *Eine für die Valenztheorie geeignete Basis der binären Vektorinvarianten*, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. 1932 (1932) 499–504

[27] **A Sartori**, *Categorification of tensor powers of the vector representation of* $U_q(\mathfrak{gl}(1|1))$, PhD thesis, Universiät Bonn (2014) Available at `http://hss.ulb.uni-bonn.de/2014/3635/3635.pdf`

[28] **A Sartori**, *Categorification of tensor powers of the vector representation of* $U_q(\mathfrak{gl}(1|1))$, Selecta Math. 22 (2016) 669–734 MR

[29] **V G Turaev**, *Quantum invariants of knots and 3–manifolds*, revised edition, De Gruyter Studies in Mathematics 18, Walter de Gruyter, Berlin (2010) MR

[30] **G Williamson**, *Singular Soergel bimodules*, Int. Math. Res. Not. 2011 (2011) 4555–4632 MR

[31] **Y Wu**, **R B Zhang**, *Unitary highest weight representations of quantum general linear superalgebra*, J. Algebra 321 (2009) 3568–3593 MR

[32] **H Yamane**, *Quantized enveloping algebras associated with simple Lie superalgebras and their universal R–matrices*, Publ. Res. Inst. Math. Sci. 30 (1994) 15–87 MR

[33] **R B Zhang**, *Structure and representations of the quantum general linear supergroup*, Comm. Math. Phys. 195 (1998) 525–547 MR

Mathematisches Institut, Universität Bonn
Bonn, Germany

Institut de Recherche en Mathématique et Physique, Université Catholique de Louvain
Louvain-la-Neuve, Belgium

Department of Mathematics, Imperial College London
London, United Kingdom

dtubben@math.uni-bonn.de, pedro.vaz@uclouvain.be, p.wedrich@gmail.com

http://www.math.uni-bonn.de/people/dtubben/,
http://perso.uclouvain.be/pedro.vaz/, http://paul.wedrich.at/

Received: 16 October 2016 Revised: 1 March 2017

Algebraic & Geometric Topology 17 (2017) 3751–3778

Uniform fellow traveling between surgery paths in the sphere graph

MATT CLAY
YULAN QING
KASRA RAFI

We show that the Hausdorff distance between any forward and any backward surgery paths in the sphere graph is at most 2. From this it follows that the Hausdorff distance between any two surgery paths with the same initial sphere system and same target sphere system is at most 4. Our proof relies on understanding how surgeries affect the Guirardel core associated to sphere systems. We show that applying a surgery is equivalent to performing a Rips move on the Guirardel core.

20E36

1 Introduction

We study surgery paths in the sphere graph. Let $\underline{M}$ be the connected sum of n copies of $S^1 \times S^2$ (we reserve the notation M for the universal cover of $\underline{M}$, which is used more frequently in the body of the paper). The vertices of the sphere graph are essential sphere systems in $\underline{M}$ and edges encode containment (see Section 2 for precise definitions). We denote the sphere graph by $\mathcal{S}$ and the associated metric with $d_{\mathcal{S}}$. It is known that the sphere graph $(\mathcal{S}, d_{\mathcal{S}})$ is hyperbolic in the sense of Gromov [11; 16]. The relationship between the optimal hyperbolicity constant and the rank of the fundamental group of $\underline{M}$ (which is isomorphic to $\mathbb{F}_n$, the free group of rank n) is unknown.

Given a pair of (filling) sphere systems $\underline{S}$ and $\underline{\Sigma}$, there is a natural family of paths, called surgery paths, connecting them. They are obtained by replacing larger and larger portions of spheres in $\underline{S}$ with pieces of spheres in $\underline{\Sigma}$. This process is not unique. Also, families of paths that start from $\underline{S}$ with target $\underline{\Sigma}$ are different from those starting from $\underline{\Sigma}$ with target $\underline{S}$. It follows from Hilion and Horbez [16] that surgery paths are quasigeodesics. Together with the hyperbolicity of the sphere graph, this implies that different surgery paths starting with $\underline{S}$ and with target $\underline{\Sigma}$ have bounded Hausdorff distance in the sphere graph. The bound depends on the optimal hyperbolicity constant, which, as stated above, does not have a good qualitative estimate.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3751

However, in this paper we show that, in any rank, any two surgery paths are within Hausdorff distance at most 4 of each other. This follows by comparing a surgery path that starts from $\underline{S}$ with target $\underline{\Sigma}$ to a surgery path starting from $\underline{\Sigma}$ with target $\underline{S}$.

Theorem 1.1 *Let $\underline{S}$ and $\underline{\Sigma}$ be two filling sphere systems and let*

$$\underline{S} = \underline{S}_1, \underline{S}_2, \dots, \underline{S}_m, \quad d_{\mathcal{S}}(\underline{S}_m, \underline{\Sigma}) \le 1,$$

be a surgery sequence starting from $\underline{S}$ towards $\underline{\Sigma}$ and

$$\underline{\Sigma} = \underline{\Sigma}_1, \underline{\Sigma}_2, \dots, \underline{\Sigma}_\mu, \quad d_{\mathcal{S}}(\underline{\Sigma}_\mu, \underline{S}) \le 1,$$

be a surgery sequence in the opposite direction. Then, for every $\underline{S}_i$ there is a $\underline{\Sigma}_j$ such that $d_{\mathcal{S}}(\underline{S}_i, \underline{\Sigma}_j) \le 2$.

Using this, we get the bound of 4 between paths with the same initial sphere system and same target sphere system.

Theorem 1.2 *Let $\underline{S}$ and $\underline{\Sigma}$ be two filling sphere systems and let*

$$\underline{S} = \underline{S}_1, \underline{S}_2, \dots, \underline{S}_m, \quad d_{\mathcal{S}}(\underline{S}_m, \underline{\Sigma}) \le 1,$$
$$\underline{S} = \underline{S}'_1, \underline{S}'_2, \dots, \underline{S}'_n, \quad d_{\mathcal{S}}(\underline{S}'_n, \underline{\Sigma}) \le 1,$$

be two surgery sequences starting from $\underline{S}$ towards $\underline{\Sigma}$. Then, for every $\underline{S}_i$ there is an $\underline{S}'_j$ such that $d_{\mathcal{S}}(\underline{S}_i, \underline{S}'_j) \le 4$.

Proof Fix two filling sphere systems $\underline{S}$ and $\underline{\Sigma}$ and surgery paths as in the statement of the theorem. Let

$$\underline{\Sigma} = \underline{\Sigma}_1, \underline{\Sigma}_2, \dots, \underline{\Sigma}_\mu, \quad d_{\mathcal{S}}(\underline{\Sigma}_\mu, \underline{S}) \le 1,$$

be a surgery sequence starting at $\underline{\Sigma}$ towards $\underline{S}$. Given $\underline{S}_i$, by Theorem 1.1, there is a $\underline{\Sigma}_k$ such that $d_{\mathcal{S}}(\underline{S}_i, \underline{\Sigma}_k) \le 2$. Applying Theorem 1.1 again, there is an $\underline{S}'_j$ such that $d_{\mathcal{S}}(\underline{\Sigma}_k, \underline{S}_j) \le 2$. Thus $d_{\mathcal{S}}(\underline{S}_i, \underline{S}'_j) \le 4$, as desired. □

The sphere graph is a direct analogue of the graph of arcs on a surface with boundary. In fact, there is an embedding of the arc graph into the sphere graph. The arc graph is known to be uniformly hyperbolic; see Aougab [1], Bowditch [5], Clay, Rafi and Schleimer [6], Hensel, Przytycki and Webb [15] and Przytycki and Sisto [19]. Since the solutions of many algorithmic problems for mapping class groups or hyperbolic 3–manifold that fibers over a circle rely on the action of mapping class group on various curve and arc complexes, the uniform hyperbolicity clarifies which constant depend on the genus and which ones are genus-independent. The uniform hyperbolicity of the sphere graph (or one of the other combinatorial complexes associated to $\mathrm{Out}(\mathbb{F}_n)$) is a central open question in the study of the group $\mathrm{Out}(\mathbb{F}_n)$, the group of outer automorphisms of the free group. Note that Theorem 1.2 is not sufficient to prove that the sphere graph is uniformly hyperbolic.

Summary of other results

The Guirardel core [10] is a square complex associated to two trees equipped with isometric actions by a group, in our case $\mathbb{F}_n$. This is an analogue of a quadratic differential in the surface case; the area of the core is the intersection number between the two associated sphere systems. Following Behrstock, Bestvina and Clay [3], in Section 3, we describe how to compute the core using the change of marking map between the two trees. Lemma 3.7 gives a simple condition on when a product of two edges is in the core, which will be used in future work to study the core. Also, in Section 4 we define the core for two sphere systems, $\mathrm{Core}(\underline{S}, \underline{\Sigma})$, directly, using the intersection pattern of the spheres and show this object is isomorphic to the Guirardel core for the associated tree (Theorem 4.9). Much of what is contained in these two sections is known to the experts, however, we include a self-contained exposition of the material since it is not written in an easily accessible way in the literature.

Applying a surgery to a sphere system amounts to applying a splitting move to the dual tree (see Example 5.6), however, not all splittings towards a given tree come from surgeries. In general, applying a splitting move could change the associated core in unpredictable ways potentially increasing the volume of the core. We will show that (Theorem 5.5) applying a surgery is equivalent to performing a Rips move on the Guirardel core. That is, there is a subset of all splitting paths between two trees that is natural from the point of view of the Guirardel core and it matches exactly with the set of splitting sequences that are associated to surgery paths.

Outline of the proof

Our proof of Theorem 1.1 analyses the Guirardel core $\mathrm{Core}(\underline{S}_i, \underline{\Sigma}_j)$. Generally, this does not have to be related to $\mathrm{Core}(\underline{S}, \underline{\Sigma})$. However, we show that, for small values of i and j, the spheres $\underline{S}_i$ and $\underline{\Sigma}_j$ are still in normal form and $\mathrm{Core}(\underline{S}_i, \underline{\Sigma}_j)$ can be obtained from $\mathrm{Core}(\underline{S}, \underline{\Sigma})$ via a sequence of vertical and horizontal Rips moves (Proposition 5.8). For every i, there is a smallest j where this breaks down, which is exactly the moment the surgery path from $\underline{\Sigma}$ to $\underline{S}$ passes near $\underline{S}_i$. The proof of Theorem 1.1 is completed in Section 6: for every $\underline{S}_i$, apply enough surgery on $\underline{\Sigma}$ until $\mathrm{Core}(\underline{S}_i, \underline{\Sigma}_j)$ has a free edge, which implies $d_{\mathcal{S}}(\underline{S}_i, \underline{\Sigma}_j) \le 2$.

Acknowledgements We would like to thank Mladen Bestvina for helpful conversations, particularly, for giving us the idea to examine how the core changes along splitting sequence. Clay was partially funded by the Simons Foundation (award number 316383). Rafi was partially funded by NSERC Discovery grant RGPIN 435885.

2 Sphere systems and free splittings

Let $\underline{M}$ be the connected sum of n copies of $S^1 \times S^2$ and fix an identification of $\pi_1(\underline{M})$ with $\mathbb{F}_n$. There is a well-known correspondence between spheres in $\underline{M}$ and graph of group decompositions of $\mathbb{F}_n$ with trivial edge groups. We explain this correspondence now.

Definition 2.1 A *sphere system* $\underline{S} \subset \underline{M}$ is a finite union of disjoint, essential (does not bound a 3–ball), embedded 2–spheres in $\underline{M}$. We specifically allow for the possibility that a sphere system contains parallel, ie isotopic, spheres. A sphere system is *filling* if each of the complementary regions $\underline{M} - \underline{S}$ are simply connected.

We define a preorder on the set of sphere systems by $\underline{S} \preceq \underline{\Sigma}$ if every sphere in $\underline{S}$ is isotopic to a sphere in $\underline{\Sigma}$. This induces an equivalence relation: $\underline{S} \sim \underline{\Sigma}$ if $\underline{S} \preceq \underline{\Sigma}$ and $\underline{\Sigma} \preceq \underline{S}$. The set of equivalence classes of sphere systems in $\underline{M}$ is denoted by $\mathcal{S}$; the subset of equivalence classes of filling sphere systems is denoted by $\mathcal{S}_{\text{fill}}$. When there can be no confusion, we denote the equivalence class of $\underline{S}$ again by $\underline{S}$.

The preorder induces a partial order on $\mathcal{S}$ that we continue to denote by $\preceq$. The *sphere graph* is the simplicial graph with vertex set $\mathcal{S}$ and edges corresponding to domination $\underline{S} \preceq \underline{\Sigma}$. For $\underline{S}, \underline{\Sigma} \in \mathcal{S}$, we denote by $d_{\mathcal{S}}(\underline{S}, \underline{\Sigma})$ the distance between $\underline{S}$ and $\underline{\Sigma}$ in the sphere graph. This is the fewest edges in an edge path between the two vertices.

We denote the universal cover of $\underline{M}$ by M and the lift of the sphere system $\underline{S}$ to M by S; we will refer to S as *a sphere system in* M. To simplify notation, we use $\mathcal{S}$ and $\mathcal{S}_{\text{fill}}$, respectively, to denote (equivalence classes of) sphere systems and filling sphere systems, respectively, in M. Let $\text{Map}(\underline{M}) = \text{Homeo}(\underline{M})/_{\text{homotopy}}$. The natural map

$$\text{Map}(\underline{M}) \to \text{Out}(\mathbb{F}_n)$$

is surjective and has finite kernel generated by Dehn twists about embedded 2–spheres in $\underline{M}$ [18]. Such homomorphisms act trivially on spheres systems and hence there is a left action of $\text{Out}(\mathbb{F}_n)$ by automorphisms on the sphere graph. Specifically, realize the given outer automorphism by a homeomorphism of $\underline{M}$ and apply this homeomorphism to the members of a given equivalence class of sphere systems.

Definition 2.2 A *free splitting* G is a simplicial tree equipped with a cocompact action of $\mathbb{F}_n$ by automorphisms (without inversions) such that the stabilizer of every edge is trivial. We specifically allow for the possibility that a free splitting contains vertices of valence two. A free splitting is *filling* if the stabilizer of every vertex is trivial.

We define a preorder on the set of free splittings by $G \preceq \Gamma$ if there is an $\mathbb{F}_n$–equivariant cellular map $\Gamma \to G$ with connected point preimages. This induces an equivalence relation $G \sim \Gamma$ if $G \preceq \Gamma$ and $\Gamma \preceq G$. The set of equivalence classes of free splittings is denoted by $\mathcal{X}$; the subset of equivalence classes of filling free splittings is denoted by $\mathcal{X}_{\text{fill}}$.[1] When there can be no confusion, we denote the equivalence class of G again by G.

The preorder induces a partial order on $\mathcal{X}$ that we continue to denote by $\preceq$. The *free splitting graph* is the simplicial graph with vertex set $\mathcal{X}$ and edges corresponding to domination $G \preceq \Gamma$. For $G, \Gamma \in \mathcal{X}$, we denote by $d_{\mathcal{X}}(G, \Gamma)$ the distance between G and Γ in the free splitting graph. This the fewest edges in an edge path between the two vertices.

Suppose that G is a free splitting and let $\rho\colon \mathbb{F}_n \to \mathrm{Aut}(G)$ be the action homomorphism. Given $\Phi \in \mathrm{Aut}(\mathbb{F}_n)$, the homomorphism $\rho \circ \Phi\colon \mathbb{F}_n \to \mathrm{Aut}(G)$ defines a new free splitting, which we denote by $G \cdot \Phi$. This defines a right action by $\mathrm{Aut}(\mathbb{F}_n)$ on the free splitting graph. As $\mathrm{Inn}(\mathbb{F}_n)$ acts trivially, this induces an action of $\mathrm{Out}(\mathbb{F}_n)$ by automorphisms on the free splitting graph.

There is a natural $\mathrm{Out}(\mathbb{F}_n)$–equivariant map from the sphere graph to the free splitting graph. Given a sphere system $S \subset M$, we define a tree G with vertex set consisting of the components of $M - S$ and edges corresponding to nonempty intersection between the closures of the components. The action of $\mathbb{F}_n$ on M induces a cocompact action of $\mathbb{F}_n$ on G by automorphisms such that the stabilizer of every edge is trivial, ie G is a free splitting. This map is a simplicial isomorphism [2, Lemma 2].

3 The Guirardel core

In this section we give the definition of Guirardel core of two trees as it is presented in [10] specialized to the case of trees in $\mathcal{X}_{\text{fill}}$.

3.1 A core for a pair of tree actions

A *ray* in $G \in \mathcal{X}_{\text{fill}}$ is an isometric embedding $\vec{r}\colon \mathbb{R}_+ \to G$. An *end* of G is an equivalence class of rays under the equivalence relation of having finite Hausdorff distance. The set of all ends is called the *boundary* of G and is denoted by ∂G.

[1] Experts may recognize $\mathcal{X}_{\text{fill}}$ as the vertices in the spine of the Culler–Vogtmann outer space [9].

A *direction* is a connected component of $G-\{x\}$, where x is a point in G. A direction $\delta \subset G$ determines a subset $\partial\delta \subset \partial G$ consisting of all ends for which every representative ray intersects δ in a nonempty (equivalently unbounded) subset. Given an edge $e \subset G$, we denote by $\vec{e}$ the edge with a specific orientation. This determines a direction $\delta_{\vec{e}} \subset G$ by taking the component of $G-\{x\}$ that contains e, where x is the initial vertex of $\vec{e}$. We will denote by $\vec{e}_\infty \subset \partial_\infty G$ the set of ends with a representative that crosses $\vec{e}$ with the specified orientation, ie $\vec{e}_\infty = \partial\delta_{\vec{e}}$.

A *quadrant* in $G \times \Gamma$ is the product $\delta_1 \times \delta_2$ of two directions $\delta_1 \subset G$ and $\delta_2 \subset \Gamma$.

Definition 3.1 Fix a basepoint $(*_1, *_2) \in G \times \Gamma$ and consider a quadrant $Q = \delta_1 \times \delta_2 \subset G \times \Gamma$. We say that Q is *heavy* if there exists a sequence $g_k \in \mathbb{F}_n$ such that:

(1) $(g_k *_1, g_k *_2) \in Q$.

(2) $d_G(g_k *_1, *_1) \to \infty$ and $d_\Gamma(g_k *_2, *_2) \to \infty$ as $k \to \infty$.

Otherwise, we say that Q is *light*.

The core of $G \times \Gamma$ is what remains when one has removed the light quadrants.

Definition 3.2 (the Guirardel core) Suppose that $G, \Gamma \in \mathcal{X}_{\text{fill}}$ and let $\mathcal{L}(G,\Gamma)$ be the collection of light quadrants of $G \times \Gamma$. The (*Guirardel*) *core* of G and Γ is the subset

$$\mathrm{Core}(G,\Gamma) = (G \times \Gamma) - \Bigg(\bigcup_{Q \in \mathcal{L}(G,\Gamma)} Q \Bigg).$$

It follows from the definition that $\mathrm{Core}(G,\Gamma)$ is isomorphic to $\mathrm{Core}(\Gamma,G)$ via the swap $(x,y) \mapsto (y,x)$. For more details and examples, see [10; 3].

3.2 Computing the core

There is an algorithm to compute the core for trees $G, \Gamma \in \mathcal{X}_{\text{fill}}$. This suffices to compute the core for any free splittings $G_0, \Gamma_0 \in \mathcal{X}$. Indeed, if the given trees are not filling, they can be "blown up" to filling trees $G, \Gamma \in \mathcal{X}_{\text{fill}}$ by replacing vertices with nontrivial stabilizer in the quotient graph of groups $G_0/\mathbb{F}_n$ and $\Gamma_0/\mathbb{F}_n$ with roses of the appropriate rank. There are domination maps $p\colon G \to G_0$ and $\pi\colon \Gamma \to \Gamma_0$ and we have that $\mathrm{Core}(G_0,\Gamma_0) = (p \times \pi)(\mathrm{Core}(G,\Gamma))$. This material appears in [3, Section 2] with slightly different terminology and notation. We provide proofs of the most relevant parts necessary for the sequel.

Definition 3.3 Suppose that $G, \Gamma \in \mathcal{X}_{\text{fill}}$. An $\mathbb{F}_n$–equivariant map $f\colon G \to \Gamma$ is called a *morphism* if

(1) f linearly expands every edge across a tight edge path, and

(2) at each vertex of G there are adjacent edges e and e' such that $f(e) \cap f(e')$ is trivial, ie there is more than one gate at each vertex.

Such a map f induces an $\mathbb{F}_n$–equivariant homeomorphism $f_\infty\colon \partial G \to \partial \Gamma$. Indeed, this follows by bounded cancellation; for instance see [7]. Next, we state the criterion provided in [3] regarding the existence of squares in the core.

Lemma 3.4 [3, Lemma 2.3] *Let $f\colon G \to \Gamma$ be a morphism between $G, \Gamma \in \mathcal{X}_{\mathrm{fill}}$. Given two edges $e \subset G$ and $\eta \subset \Gamma$, the square $e \times \eta$ is in the core* $\mathrm{Core}(G,\Gamma)$ *if and only if for every choice of orientations of the edges e and η the subset $f_\infty(\vec{e}_\infty) \cap \vec{\eta}_\infty$ is nonempty.*

This condition is very natural in the following way: Given a curve $\underline{\alpha}$ on a closed surface $\underline{X}$, each lift α of $\underline{\alpha}$ to the universal cover X determines a partition of ∂X (which is homeomorphic to S^1) into two subsets α_+ and α_-. For two curves $\underline{\alpha}$ and $\underline{\beta}$ on $\underline{X}$ that intersect minimally, lifts α and β to X intersect if and only if for every choice of $*, *' \in \{+,-\}$, the set $\alpha_* \cap \beta_{*'}$ is nonempty.

Using f, it is a simple matter to determine when this condition is met for a given pair of edges. We discuss this now. By the *interior* of a simplicial subtree we mean all nonextremal edges.

Definition 3.5 Suppose that $G, \Gamma \in \mathcal{X}_{\mathrm{fill}}$, $f\colon G \to \Gamma$ is a morphism and $\eta \in \Gamma$ is an edge. We let $\mathcal{P}^f_\eta$ be the set of edges in G whose image under f traverses η. In other words, $\mathcal{P}^f_\eta$ is the set of edges containing $f^{-1}(\eta)$. Since f is a morphism, by bounded cancellation, the set $\mathcal{P}^f_\eta$ is finite.

Let $\mathcal{H}^f_\eta$ be the interior of the convex hull of $\mathcal{P}^f_\eta$ and let $\widehat{\mathcal{P}}^f_\eta = \mathcal{P}^f_\eta - \mathcal{H}^f_\eta$; see Figure 1. Notice that the interior of the convex hull of $\widehat{\mathcal{P}}^f_\eta$ is also $\mathcal{H}^f_\eta$. Suppose $e \in \mathcal{H}^f_\eta$ and $\vec{e}$ is an orientation of e. We say that $\vec{e}$ *can escape* $\mathcal{P}^f_\eta$ if there is an embedded ray of the form $\vec{e} \cdot \vec{r}$ such that $\vec{r}$ does not cross any edge of $\widehat{\mathcal{P}}^f_\eta$. Define the *consolidated convex hull* $\mathcal{CH}^f_\eta$ of $\mathcal{P}^f_\eta$ to be the set of edges in $e \in \mathcal{H}^f_\eta$ such that both orientations of e can escape $\mathcal{P}^f_\eta$.

Lemma 3.6 *For every vertex $v \in G$ there is a ray $\vec{r}$ originating at v that is disjoint from $\mathcal{P}^f_\eta$.*

Proof If the lemma were false, then for every edge e adjacent to v the image $f(e)$ would contain the initial edge in the path connecting $f(v)$ to η. This violates condition (2) in Definition 3.3. □

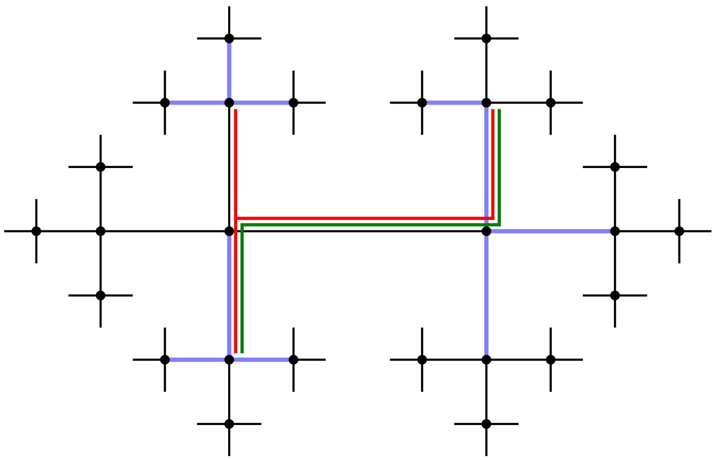


Figure 1: A schematic of the sets $\mathcal{P}^f_\eta$ (blue), $\widehat{\mathcal{P}}^f_\eta$, $\mathcal{H}^f_\eta$ (red) and $\mathcal{CH}^f_\eta$ (green)

The following simple condition tells exactly when a square is in the core.

Lemma 3.7 *Let $G, \Gamma \in \mathcal{X}_{\mathrm{fill}}$, fix morphisms $f\colon G \to \Gamma$ and $\phi\colon \Gamma \to G$ and consider a pair of edges $e \subset G$ and $\eta \subset \Gamma$. The square $e \times \eta$ is in $\mathrm{Core}(G, \Gamma)$ if and only if one of the two following equivalent conditions holds:*

- $e \subseteq \mathcal{CH}^f_\eta$.
- $\eta \subseteq \mathcal{CH}^\phi_e$.

Proof We prove the first of the two equivalent statements; the fact that they are equivalent follows from the symmetry of the construction of the core. For simplicity,

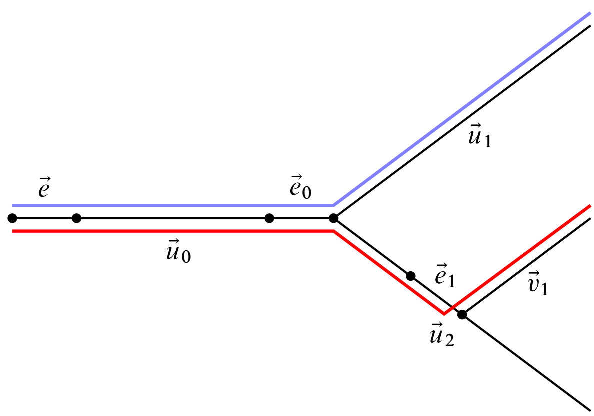


Figure 2: Rays $\vec{r}_0$ (blue) and $\vec{r}_1$ (red) witnessing $e \subseteq \mathcal{CH}_\eta$ in Lemma 3.7

we omit the superscript f on the various subsets from Definition 3.5 during the proof of this lemma.

By Lemma 3.4, what needs to be shown is that $e \subseteq \mathcal{CH}_\eta$ if and only if for each orientation $\vec{e}$ of e and orientation $\vec{\eta}$ of η there is a ray $\vec{r}$ crossing $\vec{e}$ with the specified orientation such that $f_\infty(\vec{r}\,) \in \vec{\eta}_\infty$.

First suppose that $e \subseteq \mathcal{CH}_\eta$ and fix an orientation $\vec{e}$ on e; see Figure 2. As $e \subseteq \mathcal{CH}_\eta$, there is a ray $\vec{r}_0 = \vec{e} \cdot \vec{u}$ such that $\vec{u}$ is disjoint from $\widehat{\mathcal{P}}_\eta$. Let e_0 be the last edge on $\vec{r}_0$ that is in $\mathcal{CH}_\eta$ and decompose $\vec{r}_0 = \vec{u}_0 \cdot \vec{e}_0 \cdot \vec{u}_1$ where $\vec{u}_0$ may be trivial. It is easy to verify that the ray $\vec{u}_1$ is disjoint from $\mathcal{P}_\eta$. As $e_0 \subseteq \mathcal{H}_\eta$, there is a ray of the form $\vec{e}_0 \cdot \vec{u}_2$, where $\vec{u}_2$ is not disjoint from $\mathcal{P}_\eta$. (It may be that $\vec{u}_1$ and $\vec{u}_2$ have nontrivial intersection.) Let e_1 be the first edge on $\vec{u}_2$ that is contained in $\mathcal{P}_\eta$ and $\vec{p}$ the oriented edge path from $\vec{e}$ to $\vec{e}_1$. By Lemma 3.6, there is a ray $\vec{v}_1$ originating at the terminal vertex of $\vec{p}$ that is disjoint from $\mathcal{P}_\eta$. Let $\vec{r}_1 = \vec{p} \cdot \vec{v}_1$. We now see that

$$\#|\vec{r}_1 \cap \mathcal{P}_\eta| = \#|\vec{r}_0 \cap \mathcal{P}_\eta| + 1.$$

Since $\vec{r}_0$ and $\vec{r}_1$ originate from the same vertex, their f_∞–images lie in $\vec{\eta}_\infty$ for opposite choices of orientation of η. By Lemma 3.4, this shows that $e \times \eta \subseteq \mathrm{Core}(G, \Gamma)$.

For the converse we suppose that $e \not\subseteq \mathcal{CH}_\eta$. If, further, $e \not\subseteq \mathcal{H}_\eta$, then there is a choice of orientation $\vec{e}$ such that for every ray of the form $\vec{e} \cdot \vec{r}$, the ray $\vec{r}$ misses $\mathcal{P}_\eta$. Therefore, there is an orientation on η, say $\vec{\eta}$, such that $f_\infty(\vec{e}_\infty) \cap \vec{\eta}_\infty = \varnothing$. By Lemma 3.4, $e \times \eta \not\subseteq \mathrm{Core}(G, \Gamma)$.

Thus we can assume that $e \subseteq \mathcal{H}_\eta - \mathcal{CH}_\eta$. Hence, there is a choice of orientation $\vec{e}$ that cannot escape, ie for every ray form $\vec{e} \cdot \vec{r}$, the ray $\vec{r}$ must contain some edge in $\widehat{\mathcal{P}}_\eta$. By Lemma 3.6, we see that each such ray $\vec{r}$ can only contain a single edge of $\mathcal{P}_\eta$. Again, there is an orientation on η, say $\vec{\eta}$, such that $f_\infty(\vec{e}_\infty) \cap \vec{\eta}_\infty = \varnothing$. By Lemma 3.4, $e \times \eta \not\subseteq \mathrm{Core}(G, \Gamma)$. □

Since $\mathrm{Core}(G, \Gamma)$ is defined without reference to the morphism $f\colon G \to \Gamma$, Lemma 3.7 shows that $\mathcal{CH}^f_\eta$ and $\mathcal{CH}^\phi_e$ do not depend on the actual morphism used to compute them. As such, we will drop the superscripts from these sets for the remainder.

4 Sphere systems and the core

In Section 2 we described an $\mathrm{Out}(\mathbb{F}_n)$–equivariant association between sphere systems and free splittings respecting the notion of filling: $(\mathcal{S}, \mathcal{S}_{\mathrm{fill}}) \leftrightarrow (\mathcal{X}, \mathcal{X}_{\mathrm{fill}})$. In Section 3, given a pair of free splittings $G, \Gamma \in \mathcal{X}_{\mathrm{fill}}$, we described how to construct their Guirardel core $\mathrm{Core}(G,\Gamma)$. The goal of this section is, given a pair of filling sphere systems $S,\Sigma \in \mathcal{S}_{\mathrm{fill}}$, to construct a 2–dimensional square complex $\mathrm{Core}(S,\Sigma)$. We then show that when the pair of sphere systems (S,Σ) is associated to the pair of free splittings (G,Γ) there is a $\mathbb{F}_n$–equivariant isomorphism from $\mathrm{Core}(S,\Sigma) \to \mathrm{Core}(G,\Gamma)$ of square complexes. Moreover, this association is $\mathrm{Out}(\mathbb{F}_n)$–equivariant with respect to the actions on $\mathcal{X}_{\mathrm{fill}}$ and $\mathcal{S}_{\mathrm{fill}}$. This association is implicit in the proof of Proposition 2.1 in [17]. We explain the connection in more detail here and provide an alternative proof. The in-depth description is necessary for understanding the effect of surgery on the core that we describe in Section 5.

4.1 Hatcher's normal form

Central to the understanding of sphere systems in M is Hatcher's notion of normal form. He originally defined normal form only with respect to a maximal sphere system Σ [12] and extended this to filling sphere systems in subsequent work with Vogtmann [13]. We recall this definition now. The sphere system S is said to be in *normal form* with respect to Σ if every sphere $s \in S$ either belongs to Σ, or intersects Σ transversely in a collection of circles that split s into components called *pieces* such that for each component $\Pi \subset M - \Sigma$ one has

(1) each piece in Π meets each boundary sphere in $\partial\Pi$ in at most one circle, and

(2) no piece in Π is a disk that is isotopic relative to its boundary to a disk in $\partial\Pi$.

Hatcher proved that a sphere system S can always be homotoped into normal form with respect to the maximal sphere system Σ and that such a form is unique up to homotopy [12; 13]. Hensel, Osajda and Przytycki generalized Hatcher's definition of normal form to nonfilling sphere systems and in a way that is obviously symmetric with respect to the two sphere systems [14]. With their notion, two sphere systems S and Σ are in *normal form* if for all $s \in S$ and $\sigma \in \Sigma$ one has

(1) s and σ intersect transversely in at most one circle, and

(2) none of the disks in $s - \sigma$ is isotopic relative to its boundary to a disk in σ.

These notions are equivalent when Σ is filling [14, Section 7.1].

4.2 A core for a pair of sphere systems

Suppose that S and Σ are filling sphere systems in M and that they are in normal form. An *S–piece* is the closure of a component of $S - \Sigma$. Likewise, a *Σ–piece* is the closure of a component of $\Sigma - S$. By *piece*, we mean either an S–piece or a Σ–piece (this agrees with the use of "piece" in Section 4.1).

Lemma 4.1 *Suppose that X is the intersection of a component of $M - S$ and a component of $M - \Sigma$. Then:*

(1) *X is connected.*

(2) *∂X is the union of S–pieces and Σ–pieces and moreover, different pieces are subsets of different spheres.*

(3) *If Y is also the intersection of a component of $M - S$ and a component of $M - \Sigma$, then either $X = Y$, their closures are disjoint, or $\partial X \cap \partial Y$ is a piece.*

Proof This follows from the description of normal form; the details are left to the reader. $\square$

The first item in Lemma 4.1 implies that the intersection of a component of $M - S$ and a component of $M - \Sigma$ is either empty or a component of $M - (S \cup \Sigma)$.

Definition 4.2 Suppose that S and Σ are filling sphere systems in M and that they are in normal form. The *core* of S and Σ, denoted by $\mathrm{Core}(S, \Sigma)$, is the square complex defined as follows:

- Vertices correspond to components of $M - (S \cup \Sigma)$. Such a region corresponds to the intersection of a component $P \subset M - S$ and a component $\Pi \subset M - \Sigma$. We denote the vertex by (P, Π).
- There is an edge between two vertices when the closures of the corresponding components of $M - (S \cup \Sigma)$ have nontrivial intersection. By Lemma 4.1, each edge corresponds to a piece. If it is an S–piece, then it is the closure of $s \cap \Pi$ for some sphere $s \in S$ and component $\Pi \subset M - \Sigma$. We denote the edge by (s, Π). Likewise, if it is an Σ–piece, then it is the closure of $P \cap \sigma$ for some component $P \subset M - S$ and sphere $\sigma \in \Sigma$. In this case, we denote the edge by (P, σ).
- Suppose that $s \in S$ and $\sigma \in \Sigma$ have nonempty intersection. Let $P_1, P_2 \subset M - S$ be the components whose boundary contains s and let $\Pi_1, \Pi_2 \subset M - \Sigma$ be the components whose boundary contains σ. Then four edges (s, Π_1), (P_1, σ), (s, Π_2) and (P_2, σ) form the boundary of a square with vertices (P_1, Π_1), (P_2, Π_1), (P_2, Π_2) and (P_1, Π_2), which is then filled in. The square is denoted by $s \times \sigma$; see Figure 3.

Remark 4.3 We always assume that S and Σ do not share a sphere. Otherwise Theorem 1.1 is trivial. However, the core in this case would be disconnected and make the exposition more complicated. There is a procedure to add diagonal edges resulting in the *augmented core*, which is connected. See [10] for details.

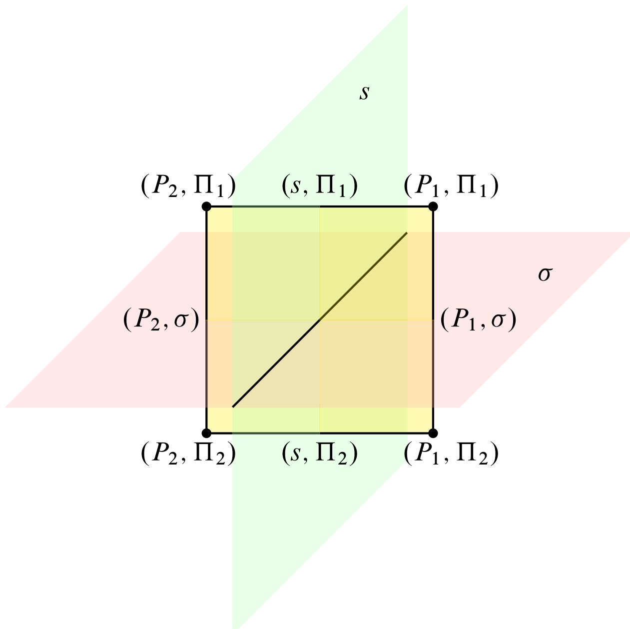


Figure 3: Edges (s, Π_1), (P_1, σ), (s, Π_2) and (P_2, σ) form the boundary of a square $s \times \sigma$.

Let G, $\Gamma \in \mathcal{X}_{\mathrm{fill}}$ be the free splittings corresponding to S and Σ, respectively. We will show that the two notions of the core, $\mathrm{Core}(S, \Sigma)$ and $\mathrm{Core}(G, \Gamma)$, are isomorphic as $\mathbb{F}_n$–square complexes. We will do so by showing that their horizontal hyperplanes agree. To this end we make the following definition:

Definition 4.4 The *shadow of* $\sigma \in \Sigma$ is the union of the edges $e \subset G$ whose associated sphere in S intersects σ. We denote the shadow by $\mathrm{Shadow}(\sigma) \subset G$.

Observe that the shadow of σ is isomorphic to the tree in σ that is dual to the intersection circles between σ and S. Now will show how to relate the two definitions of the core. We will make use of the following notion:

Definition 4.5 If $G \in \mathcal{X}$ corresponds to a sphere system $S \in \mathcal{S}$, and $\iota\colon G \hookrightarrow M$ is an $\mathbb{F}_n$–equivariant embedding, we say $\iota(G)$ is *dual to* S if each sphere $s \in S$ intersects exactly one edge of $\iota(G)$, namely the image of the corresponding edge, and this intersection is transverse and a single point. We say that ι is a *dual embedding* (*for* S).

It is a routine matter to construct a dual embedding for a given free splitting. We need to show that we can make it in some sense normal to Σ.

Lemma 4.6 *There exists a dual embedding $\iota\colon G \hookrightarrow M$ for S such that, for each edge $e \in G$ and sphere $\sigma \in \Sigma$, $\iota(e)$ and σ are either disjoint or intersect transversely at a single point in the interior of $\iota(e)$.*

Proof Let $\iota_0\colon G \hookrightarrow M$ be a dual embedding. By general position, we can assume that $\iota_0(G) \cap S \cap \Sigma = \varnothing$ and that Σ is disjoint from the vertices of $\iota_0(G)$.

Suppose that for some edge $e \subset G$, the image $\iota_0(e)$ intersects some sphere in Σ in more than one point. Let $s \in S$ be the sphere corresponding to e. Fix some innermost pair of intersection points $x, y \in \iota_0(e)$ and let $\sigma \in \Sigma$ be the corresponding sphere. Let I be the subsegment of $\iota_0(e)$ with endpoints x and y.

Notice that any circle of intersection of $S \cap \sigma$ that separates x and y in σ must correspond to a sphere $s' \in S$ such that $s' \cap I \neq \varnothing$. Indeed, if not, since S and Σ are in normal form, there would a loop consisting of I and an arc in σ that intersects some sphere in S exactly once. This is a contradiction as spheres in S are separating.

Therefore, there is an arc $J \subset \sigma$ that intersects exactly the same set of spheres of S as I, which is either s or the empty set. We can then homotope I to J and continue pushing in this direction to reduce the number of intersection points between $\iota_0(e)$ and Σ by two. Equivariantly perform this process to obtain a new dual embedding $\iota_1\colon G \to M$ that has fewer F_n–orbits of intersect between the image of G and Σ.

Iterating this procedure we arrive at $\iota\colon G \to M$ as in the statement of the lemma. □

If $\iota\colon G \hookrightarrow M$ is an $\mathbb{F}_n$–equivariant embedding so that $\iota(G)$ is transverse to Σ we can create a map $k_\iota\colon G \to \Gamma$ by sending an edge e to the edge path in Γ corresponding to the spheres in Σ crossed by $\iota(e)$.

Lemma 4.7 *There exists a dual embedding $\iota\colon G \hookrightarrow M$ for S such that the associated map $k_\iota\colon G \to \Gamma$ is a morphism.*

Proof Whenever a dual embedding $\iota_0\colon G \hookrightarrow M$ satisfies Lemma 4.6, the image of each edge $e \subset G$ is a tight edge path in Γ. Thus the only way that such a dual embedding $\iota_0\colon G \to M$ fails to produce a morphism is if there is some vertex $v \in G$ with adjacent edges $e_1, \dots, e_\ell$ (oriented to have v as their initial vertex) and sphere $\sigma \in \Sigma$ such that the first intersection point of $\iota_0(e_i) \cap \Sigma$ lies in σ for each $i = 1, \dots, \ell$. Arguing as in Lemma 4.6, we can equivariantly homotope ι_0 to locally reduce the number of intersections between the image of G and Σ by pushing the image of v across σ and pushing subarcs of edges with both endpoints on σ across σ as well.

Iterating this procedure we arrive at $\iota\colon G \to M$ as in the statement of the lemma. □

A dual embedding $\iota\colon G \to M$ satisfying the conclusions of Lemmas 4.6 and 4.7 is said to be *normal to* Σ.

Proposition 4.8 *Suppose that S and Σ are filling sphere systems in M and that G and Γ are the associated trees. Fix an edge $\eta \in \Gamma$ and let $\sigma \in \Sigma$ be the associated sphere. If $\iota\colon G \hookrightarrow M$ is a dual embedding that is normal to Σ, then* $\mathrm{Shadow}(\sigma) = \mathcal{CH}_\eta = \mathcal{H}_\eta^{k_\iota}$.

Proof Suppose that $\iota\colon G \hookrightarrow M$ is a dual embedding that is normal to Σ. Let $e \subset G$ be an edge and $s \in S$ the sphere corresponding to e.

First suppose that $e \subseteq \mathrm{Shadow}(\sigma)$. Thus $s \cap \sigma$ is nonempty and, as the sphere systems are in normal form, this intersection is a single circle. Let X be one of the four components of $M - (s \cup \sigma)$. Decompose $\partial X = d \cup \delta$, where d is a subdisk of s and δ is a subdisk of σ.

We claim that $\iota^{-1}(X) \subseteq G$ contains an infinite subtree. Suppose otherwise; thus $\iota^{-1}(X)$ is a finite subforest T. At most one extremal vertex of T corresponds to an intersection of $\iota(e)$ and s (which is in d); the remaining extremal vertices correspond to intersections of δ with edges in $\iota(G)$.

If T is empty or has some component contained in an edge of G then an innermost disk of δ (with respect to the intersection circles $\delta \cap S$) is homotopic relative to its boundary to a disk in S, violating the assumption that S and Σ are in normal form. Otherwise, if for some component $T_0 \subseteq T$ we have that $\iota(T_0)$ does not intersect s, then for any interior vertex of T, as we saw in the proof of Lemma 4.7, the map k_ι only has one gate, violating the assumption that ι is not normal.

Thus we may assume that T is connected and has some interior vertex v, which we assume is adjacent to some extremal edge of T that is not contained in e. We label the edges $e_0, e_1, \dots, e_\ell$ adjacent to v, where $\iota(e_i)$ intersects σ for $i = 1, \dots, \ell$. Let s_i be the spheres of S corresponding to e_i for $i = 0, \dots, \ell$. Then σ must be disjoint from s_i for $i = 1, \dots, \ell$ for otherwise there is a component of $M - (s_i \cup \sigma)$ whose preimage in G contains a component that is contained in a single edge, which we already ruled out. But in this case we have that $\sigma - s_0$ contains a disk isotopic relative to its boundary to a disk in s_0, which again violates the assumption that S and Σ are in normal form.

Hence we can find a ray $\vec{r}$ starting with e such that $\iota(\vec{r})$ is eventually contained in X. Since X was arbitrary, this shows that for each orientation of $\vec{e}$ for e and $\vec{\eta}$ for η we can find a ray $\vec{r}$ crossing $\vec{e}$ with the specified orientation such that $k_\iota(\vec{r}) \in \vec{\eta}_\infty$. By Lemma 3.4, we have that $e \times \eta \subseteq \mathrm{Core}(G, \Gamma)$ and so $e \subseteq \mathcal{CH}_\eta$ by Lemma 3.7. Hence $\mathrm{Shadow}(\sigma) \subseteq \mathcal{CH}_\eta \subseteq \mathcal{H}_\eta^{k_\iota}$.

Now suppose that $e \subseteq \mathcal{H}_\eta^{k_\iota}$. Then, for each orientation $\vec{e}$, there is a ray of the form $\vec{e} \cdot \vec{r}$ such that $k_\iota(\vec{r})$ intersects η and hence $\iota(\vec{r})$ intersects σ. Since s separates M and σ is connected, this shows that s intersects σ, ie $e \subseteq \mathrm{Shadow}(\sigma)$. Hence $\mathcal{H}_\eta^{k_\iota} \subseteq \mathrm{Shadow}(\sigma)$, completing the proof. □

In other words, Proposition 4.8 states that $\mathrm{Shadow}(\sigma) \subset G$ is the interior of the convex hull of $k_\iota^{-1}(\sigma)$.

Recall the relation between a sphere system S and the corresponding free splitting G mentioned in Section 2: vertices of G correspond to connected components $M - S$ and edges correspond to nonempty intersection between the closures of the components, ie spheres in S. We can define a map $\mathrm{Core}(S, \Sigma) \to G \times \Gamma$ as follows:

- The image of a vertex (P, Π) is the vertex $(v, \nu) \in G \times \Gamma$, where v is the vertex corresponding to $P \subset M - S$ and ν is the vertex corresponding to $\Pi \subset M - \Sigma$.
- The image of an edge (s, Π) is the edge $(e, \nu) \subset G \times \Gamma$, where e is the edge corresponding to $s \in S$ and ν is the vertex corresponding to $\Pi \subset M - \Sigma$. Likewise, the image of an edge (P, σ) is the edge $(v, \eta) \subset G \times \Gamma$, where v is the vertex corresponding to $P \subset M - S$ and η is the edge corresponding to $\sigma \in \Sigma$.
- The image of the square $s \times \sigma$ is $e \times \eta \subset G \times \Gamma$, where e is the edge corresponding to $s \in S$ and η is the edge corresponding to $\sigma \in \Sigma$.

The following theorem is implicit in the proof of [17, Proposition 2.1]. There, Horbez uses a characterization by Guirardel of the core as the minimal closed, connect, $\mathbb{F}_n$–invariant subset of $G \times \Gamma$ that has connected fibers [10, Proposition 5.1]. We avoid using this characterization by using Lemma 3.7 and Proposition 4.8.

Theorem 4.9 *If* $G, \Gamma \in \mathcal{X}_{\mathrm{fill}}$ *correspond to* $S, \Sigma \in \mathcal{S}_{\mathrm{fill}}$ *which do not share a sphere, then the map* $\mathrm{Core}(S, \Sigma) \to G \times \Gamma$ *induces an* $\mathbb{F}_n$*–equivariant isomorphism of square complexes* $\mathrm{Core}(S, \Sigma) \to \mathrm{Core}(G, \Gamma)$.

Proof It is clear that the map is injective, $\mathbb{F}_n$–equivariant and preserves the square structure. We just need to show that the image is $\mathrm{Core}(G, \Gamma)$. For each $\sigma \in \Sigma$, let $S_\sigma = \{s \in S \mid s \cap \sigma \neq \varnothing\}$. Notice that the edges in G corresponding to S_σ is $\mathrm{Shadow}(\sigma)$ by definition. We can decompose the core $\mathrm{Core}(S, \Sigma)$ vertically into horizontal slices $C_\sigma = \{s \times \sigma \mid s \in S_\sigma\}$. Now fix a σ and let η by the corresponding edges of Γ. Then image of the strip C_σ is exactly the set of squares $\{e \times \eta \mid e \subseteq \mathrm{Shadow}(\sigma)\}$. By Proposition 4.8 we can also write this as $\{e \times \eta \mid e \subseteq \mathcal{CH}_\eta\}$. By Lemma 3.7 we can further write this as $\{e \times \eta \mid e \times \eta \subseteq \mathrm{Core}(G, \Gamma)\}$. Hence the image of the map is as claimed. □

5 Surgery and the core

The purpose of this section is to show how the core changes along a surgery path in the sphere graph.

5.1 Surgery sequences

Suppose that $S, \Sigma \in \mathcal{S}$ and assume that they are in normal form. We now describe a path from S to Σ in $\mathcal{S}$ using a surgery procedure introduced by Hatcher [12]. It is exactly these paths that appear in the main theorem of this paper.

Fix a sphere $\sigma \in \Sigma$ that intersects some spheres of S. The intersection circles define a pattern of disjoint circles on σ, each of which bounds two disks on σ. Choose an innermost disk δ in this collection, ie a disk that contains no other disk from this collection, and let α be its boundary circle. The sphere $s \in S$ containing α is the union of two disks d_+ and d_- that share the boundary circle α. Briefly, surgery replaces the sphere σ with new spheres $d_+ \cup \delta$ and $d_- \cup \delta$. One problem that arises is that the new sphere system and S are not in normal form. This happens when some innermost disk δ' in a sphere $\sigma' \in \Sigma$ is parallel rel s to δ. To address this, we remove all such disks at once, so that the resulting sphere system and S are in normal form (Lemma 5.1).

Let $\{\alpha_i\}_{i=k}^{\ell}$ be the maximal family of intersection circles in $s \cap \Sigma$ such that

(1) $k \le 0 \le \ell$,

(2) $\alpha_i \subset d_-$ for $i \le 0$ and $\alpha_i \subset d_+$ for $i \ge 0$ (this implies that $\alpha_0 = \alpha$), and

(3) for $k \le i < \ell$, the circles α_i and α_{i+1} cobound an annulus $A_i \subset s$ whose interior is disjoint from Σ.

Related to these circles, we let $\{\delta_i\}_{i=\kappa}^{\lambda}$ be the maximal family of innermost disks in Σ such that

(1) $\kappa \le 0 \le \lambda$,

(2) $\partial\delta_i = \alpha_i$, and

(3) for $\kappa \le i < \lambda$, the sphere $\delta_i \cup A_i \cup \delta_{i+1}$ bounds an embedded 3–ball, ie δ_i and δ_{i+1} are parallel rel s.

See Figure 4 for an example illustrating this set-up and notation.

Using this set-up we can now describe a surgery of S. Let δ_- be a parallel copy of δ_κ rel s such that $\partial\delta_-$ and α_κ cobound an annulus whose interior is disjoint from Σ and A_κ. Similarly let δ_+ be a parallel copy of δ_λ rel s such that $\partial\delta_+$ and α_λ cobound an annulus whose interior is disjoint from Σ and A_λ. Set $\widehat{d}_-$ to be the subdisk of d_- with boundary $\partial\delta_-$ and set $\widehat{d}_+$ to be the subdisk of d_+ with boundary $\partial\delta_+$ We get two new

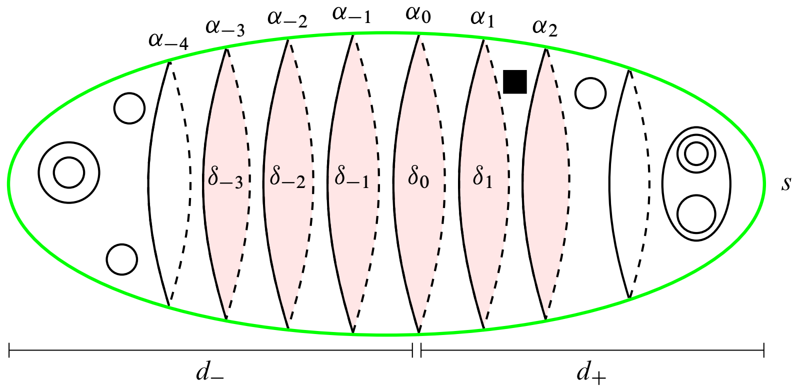


Figure 4: An example illustrating the curves $\{\alpha_i\}$ and the disks $\{\delta_i\}$. The green sphere is s, its intersection with Σ is in black and the red disks are the innermost disks in Σ. The small black box represents an obstruction to isotoping the disk bounded by α_2 to δ_1 relative to s. In this example, $\kappa = -3$ and $\lambda = 1$.

spheres $s_- = \hat{d}_- \cup \delta_-$ and $s_+ = \hat{d}_+ \cup \delta_+$. We say that $\hat{S} = (S - \mathbb{F}_n\{s\}) \cup \mathbb{F}_n\{s_+, s_-\}$ is obtained from S by performing a *surgery on* S *with respect to* Σ.

In what follows, it is important to record the history of the portions of the new spheres and so we introduce notion to this effect. Suppose that $\hat{S}$ is the result of a surgery of S with respect to Σ and that $\hat{s} \in \hat{S}$ is (a translate of) one of the newly created spheres $s_* = \hat{d}_* \cup \delta_*$ for $* \in \{+, -\}$. We call d_* the *portion of* $\hat{s}$ *from* S; denote it by $\hat{s}^S$. Similarly, we call δ_* the *portion of* $\hat{s}$ *from* Σ; denote it by $\hat{s}^\Sigma$. Thus $\hat{s} = \hat{s}^S \cup \hat{s}^\Sigma$. Notice that $\hat{s}^S \subseteq S$ and also that $\hat{s}^\Sigma$ is parallel rel s to a disk in Σ. For all other spheres $s \in \hat{S}$, we set $s^S = s$ and $s^\Sigma = \varnothing$.

Our definition of surgery differs slightly from the standard in three ways: one, we do not remove parallel spheres in $\hat{S}$; two, we perform surgery along parallel innermost disks in a single step; and three, we do not isotope S' to be in normal form with respect to Σ. That we do not remove parallel spheres is in keeping with our definition of sphere systems from Section 2. Justification of the latter two differences is the following lemma, which shows that by performing surgery along the parallel innermost disks we can eliminate the need to perform a subsequent isotopy.

Lemma 5.1 *Let* $\hat{S}$ *be the result of a surgery on* S *with respect to* Σ. *Then* $\hat{S}$ *and* Σ *are in normal form.*

Proof Suppose otherwise. As S and Σ are in normal form by assumption and normal form is a local condition, it must be that one of the newly created spheres is not in normal form with respect to Σ. Denote this sphere by $\widehat{s}=\widehat{d}\cup\delta$, where $\widehat{d}$ is a subdisk of the surgered sphere $s\in S$ and δ is a disk parallel rel s to a disk in Σ. Any intersection between $\widehat{s}$ and some sphere of Σ must lie in $\widehat{d}\subset s$, and hence $\widehat{s}$ and a given sphere in Σ intersect transversely in at most one circle, as the same held for $s\in S$.

Therefore, if $\widehat{s}$ is not in normal form with respect to Σ, then there is a sphere $\sigma\in\Sigma$ such that one of the disks in $\widehat{s}-\sigma$, say d, is isotopic relative to its boundary to a disk in σ, say δ'. Without loss of generality, we can assume that this disk is innermost on $\widehat{s}$, ie no subdisk of d is isotopic relative to its boundary to a disk in some sphere of Σ. The disk d cannot lie entirely in $\widehat{d}$ since s and Σ are in normal form by assumption. Hence d contains δ. Let A be the annulus such that $d=A\cup\delta$. Since $d\cup\delta'$ bounds a 3–ball, the assumptions that S and Σ are in normal form and that d is innermost imply that Σ is disjoint from the interior of A. This contradicts the maximality assumption on the family of disks $\{\delta_i\}_{i=\kappa}^{\lambda}$. Indeed, without loss of generality we can assume that $\delta=\delta_+$. Then $A_\lambda\cup A$ is an annulus in s whose interior is disjoint from Σ and so $\partial\delta'=\alpha_{\lambda+1}$, and further $\delta_\lambda\cup(A_\lambda\cup A)\cup\delta'$ bounds an embedded 3–ball. Hence $\widehat{S}$ and Σ are in normal form. □

Definition 5.2 A *surgery sequence from S to Σ* is a finite sequence of sphere systems

$$S=S_1,\dots,S_m$$

such that S_{i+1} is the result of a surgery of S_i with respect to Σ and $d_{\mathcal{S}}(\Sigma,S_m)\le 1$.

It is a standard fact that if $d_{\mathcal{S}}(S,\Sigma)\ge 2$, then there is a surgery sequence from S; see for instance [16, Lemma 2.2]. Further, $d_{\mathcal{S}}(S_i,S_{i+1})\le 2$ as both S_i and S_{i+1} are dominated by $S_i\cup S_{i+1}$.

The discussion and notion regarding portions from S and from Σ make sense for surgery sequences as well by induction. Indeed, suppose that S_{i+1} is obtained from S_i by a surgery with respect for Σ; specifically, assume that the (orbit of the) sphere $s\in S_i$ is split into (the orbit of) two spheres $s_-=\widehat{d}_-\cup\delta_-$ and $s_+=\widehat{d}_+\cup\delta_+$ in S_{i+1}. Then we have $s=\widehat{d}_-\cup A\cup\widehat{d}_+$ for some annulus A, the boundary curves of which are parallel to circles in Σ rel s. By choosing A sufficiently narrow enough, we can assume that the annuli of s witnessing the isotopy are contained in s^S. We set $s_*^S=\widehat{d}_*\cap s^S$ and $s_*^\Sigma=(\widehat{d}_*\cap s^\Sigma)\cup\delta_*$ for $*\in\{+,-\}$. All other spheres in S_{i+1} are also in $S_i-\{s\}$ and, as such. the portions from S and Σ remain unchanged. See Figure 5.

Lemma 5.3 *Suppose that $S=S_1,\dots,S_m$ is a surgery sequence from S to Σ. Then, for every $s\in S_i$, the subset s^S is connected.*

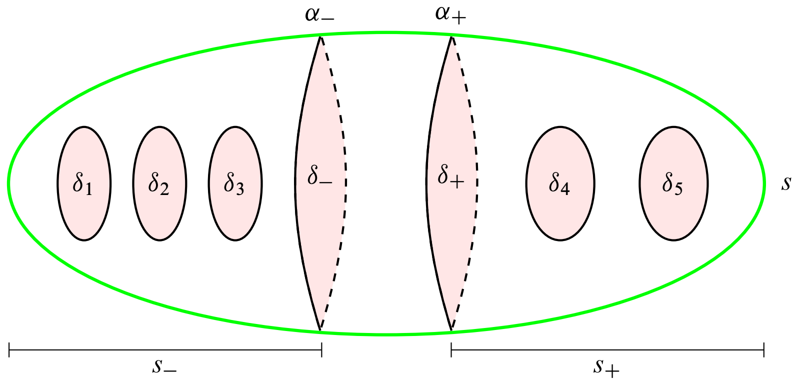


Figure 5: An illustration showing a decomposition of $s \in S_i$ and the resulting spheres $s_-, s_+ \in S_{i+1}$ into their portions from S and Σ. In this example, $s^\Sigma = \{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5\}$, $s_-^\Sigma = \{\delta_1, \delta_2, \delta_3, \delta_-\}$ and $s_-^\Sigma = \{\delta_4, \delta_5, \delta_+\}$. The portions from S are the complements in the respective spheres.

Proof Using induction, we can conclude that the subset s^Σ is a union of disks, each parallel rel s to a disk in Σ. Hence, s^S is the complement of finitely many disks in s and therefore connected. □

We remark that parallel spheres in S_i may have different histories, that is, $s_1, s_2 \in S_i$ may be parallel even though s_1^S and s_2^S are not parallel. For a surgery sequence $S = S_1, \dots, S_m$ from S to Σ, we set

$$S_i^S = \bigcup_{s \in S_i} s_i^S \quad \text{and} \quad S_i^\Sigma = \bigcup_{s \in S_i} s_i^\Sigma.$$

5.2 Rips moves and surgery steps

Suppose that S and Σ are filling sphere systems and assume that they are in normal form. Let $S = S_1, S_2, \dots, S_m$ be a surgery sequence from S to Σ and let $\Sigma = \Sigma_1, \Sigma_2, \dots, \Sigma_\mu$ be a surgery sequence from Σ to S. We describe $\mathrm{Core}(S_i, \Sigma_j)$ as (in some appropriate sense) an intersection of $\mathrm{Core}(S_i, \Sigma)$ and $\mathrm{Core}(S, \Sigma_j)$. We start by giving an embedding of $\mathrm{Core}(S_i, \Sigma)$ and $\mathrm{Core}(S, \Sigma_j)$ into $\mathrm{Core}(S, \Sigma)$. This embedding is constructed inductively. A single step in the construction is reminiscent of one of the elementary moves of the Rips machine [4; 8]. We start with a definition.

Definition 5.4 Suppose that S and Σ are filling sphere systems in M. We define ∂_S, the S–boundary of $\mathrm{Core}(S, \Sigma)$, to be the subset of $\mathrm{Core}(S, \Sigma)$ consisting of the (open) edges (P, σ) that are the face of exactly one square and vertices (P, Π) that are the vertex of exactly 3 edges of the form (P,σ), (P,σ') and (s,Π). A connected component of ∂_S is called an S–*side* of $\mathrm{Core}(S,\Sigma)$. Similarly, we define ∂_Σ, the Σ–boundary of $\mathrm{Core}(S,\Sigma)$, to be the subset of $\mathrm{Core}(S,\Sigma)$ consisting of the (open) edges (s,Π) that are the face of exactly one square and vertices (P,Π) that are the vertex of exactly 3 edges of the form (s,Π), (s',Π) and (P,σ). A connected component of ∂_Σ is called a Σ–*side* of $\mathrm{Core}(S,\Sigma)$.

The union of an S–side with the set of the (open) squares that have a face contained in that side is called a *maximal S–boundary rectangle*. That is, in a S–maximal boundary rectangle, all of the squares are of the form $s_0\times\sigma$ for some fixed $s_0\in S$. A Σ–*maximal boundary rectangle* is similarly defined from a connected component of the Σ–side. A *Rips move* on (S,Σ) is the removal of the $\mathbb{F}_n$–orbit of an S– or Σ–maximal boundary rectangle.

If R is a maximal boundary rectangle in $\mathrm{Core}(S,\Sigma)$, we let $\mathrm{Core}(S,\Sigma)_R$ denote the result of the associated Rips move. We like to think of the removal of the maximal boundary rectangle as collapsing the rectangle by pushing across the adjacent squares.

We postpone presenting an example until after the following theorem.

Theorem 5.5 *Suppose that S and Σ are filling sphere systems in M and let $\widehat{S}$ be the result of a surgery on S with respect to Σ. There is a S–maximal boundary rectangle $R\subseteq\mathrm{Core}(S,\Sigma)$ such that $\mathrm{Core}(S,\Sigma)_R$ is isomorphic to $\mathrm{Core}(\widehat{S},\Sigma)$. Moreover, for each S–maximal boundary rectangle R, there is a sphere $\sigma\in\Sigma$ and innermost disk on σ that defines a surgery $S\mapsto\widehat{S}$ such that $\mathrm{Core}(\widehat{S},\Sigma)$ is isomorphic to $\mathrm{Core}(S,\Sigma)_R$.*

Proof Assume $\widehat{S}$ is obtained from S by a surgery on a sphere $s_0\in S$ and a disk δ that is part of the sphere system Σ, whose boundary α lies on s and is otherwise disjoint from S. By Lemma 5.1, $\widehat{S}$ and Σ are in normal form and so we can use the combinatorics of $\widehat{S}$ and Σ to build $\mathrm{Core}(\widehat{S},\Sigma)$.

We make use of the notation introduced in Section 5.1. Let $\{\delta_i\}_{i=\kappa}^{\lambda}$ be the maximal family of disks in Σ parallel rel s, where $\delta_0=\delta$. Let A be the union of the annuli $A_i\subset s_0$, and d_+ and d_- the components of s_0-A. Thus $\delta_\kappa\cup A\cup\delta_\lambda$ bounds a 3–ball B. The two spheres obtained by surgery of s using this family, s_+ and s_-, are parallel to $d_+\cup\delta_\lambda$ and $d_-\cup\delta_\kappa$, respectively.

Let $P_+\in M-S$ be the component that contains the interior of B and let P_- be the other component with s as a boundary. Each disk δ_i is contained in some sphere $\sigma_i\in\Sigma$. For each $\kappa\le i<\lambda$, there are components $\Pi_i\subset M-\Sigma$ such that both $\sigma_i,\sigma_{i+1}\subset\partial\Pi_i$. We claim that the collections of edges and vertices

$$(P_+,\sigma_\kappa),(P_+,\Pi_\kappa),(P_+,\sigma_{\kappa+1}),\dots,(P_+,\sigma_\lambda)$$

is a side. Indeed, each edge (P_+, σ_i) is the face of only $s \times \sigma_i$ and each vertex (P_+, Π_i) is only also adjacent to (s_0, Π_i). The first of these observations is due to the fact that $P_+ \cap \sigma_i = \delta_i$ is a disk, the second observation due to the fact that $P_+ \cap \Pi_i$ is bounded by $\delta_i \cup A_i \cup \delta_{i+1}$. Maximality of this collection follows from maximality of the collection $\{\delta_i\}_{i=\kappa}^{\lambda}$.

Let R be the corresponding maximal boundary rectangle of $\mathrm{Core}(S, \Sigma)$. We will show that $\mathrm{Core}(\widehat{S}, \Sigma)$ is isomorphic to $\mathrm{Core}(S, \Sigma)_R$. To do so, we will construct an injection of square complexes $\mathrm{Core}(\widehat{S}, \Sigma) \hookrightarrow \mathrm{Core}(S, \Sigma)$ whose image is $\mathrm{Core}(S, \Sigma)_R$.

Components in $M - S$ that are not in the orbit of P_+ and P_- are also components of $M - \widehat{S}$. But $M - \widehat{S}$ has 3 other components: $\widehat{P}_-$, which is obtained from P_- by adding a neighborhood of s and a neighborhood of the 3–ball B bounded by $\delta_\kappa \cup A \cup \delta_\lambda$, and P_+^+ and P_+^-, which are contained in the two components of $P_+ - B$. In other words, we have

$$M - \widehat{S} = \big((M - S) - \mathbb{F}_n\{P_+, P_-\}\big) \cup \mathbb{F}_n\{\widehat{P}_-, P_+^+, P_+^-\}.$$

There is an $\mathbb{F}_n$–equivariant map $\iota\colon M - \widehat{S} \to M - S$, defined by $P_+^+, P_+^- \mapsto P_+$, $\widehat{P}_- \mapsto P_-$ and the identity on the other orbits. Also, there is a $\mathbb{F}_n$–equivariant map $\epsilon\colon \widehat{S} \to S$, defined by $s_+, s_- \mapsto s_0$ and the identity on the other orbits.

Using ι, we get a map on the 0–skeleton of $\mathrm{Core}(\widehat{S}, \Sigma)$, defined by $(P, \Pi) \mapsto (\iota(P), \Pi)$. In order for this to be well-defined, we need to know that if $P \cap \Pi \neq \varnothing$, then $\iota(P) \cap \Pi \neq \varnothing$ also. If P is not in the orbit of $\widehat{P}_-$, then this follows as $P \cap \Pi \subseteq \iota(P) \cap \Pi$. Finally, since $\widehat{P}_- = P_- \cup B$ and no component of $M - \Sigma$ is contained in B, any component of $M - \Sigma$ that intersects $\widehat{P}_-$ necessarily intersects P_- as well.

We extend over the 1–skeleton using $\epsilon\colon (s, \Pi) \mapsto (\epsilon(s), \Pi)$. This map is well-defined since any intersection between s_+ or s_-, with a component of $M - \Sigma$ is contained in the portion of s_+ or s_-, respectively, from S, ie d_+ or d_-, respectively. Notice that this is consistent with the mapping on the 0–skeleton. The edge (s_+, Π) in $\mathrm{Core}(\widehat{S}, \Sigma)$ is sent to (s_0, Π). The vertices of (s_0, Π) are (P_+^+, Π) and $(\widehat{P}_-, \Pi)$, which are the images of (P_+^+, Π) and $(\widehat{P}_-, \Pi)$. Other verifications are similar.

Finally, we extend over the 2–skeleton: $s \times \sigma \mapsto \epsilon(s) \times \sigma$. Since any intersection of s_+ or s_- with Σ is contained in the portion from S, this map is well-defined. Again, the map on the 2–skeleton is consistent with the maps on 1–skeleton and 0–skeleton by construction.

The map $\mathrm{Core}(\widehat{S}, \Sigma) \to \mathrm{Core}(S, \Sigma)$ is $\mathbb{F}_n$–equivariant and preserves the square structure. The map is not surjective as no 2–cell is mapped to the squares associated with $s \times \sigma_i$, ie the image of the map is exactly $\mathrm{Core}(S, \Sigma)_R$.

The converse is similar: if $s_0 \times \sigma_1, \dots, s_0 \times \sigma_\lambda$ forms an S–maximal boundary rectangle R, then one shows that there are disks $\delta_i \subset \sigma_i$ that are parallel rel s_0 and that surgery using the family $\{\delta_i\}_{i=1}^{\lambda}$ results in the sphere system $\widehat{S}$, where $\mathrm{Core}(\widehat{S}, \Sigma)$ is isomorphic to $\mathrm{Core}(S, \Sigma)_R$. □

Example 5.6 Here we describe a Rips move and the corresponding surgery explicitly in an example. Consider a sphere $s \in S$ associated to the edge b in the dual graph G. In the example depicted in Figure 6, s intersects 7 spheres in Σ, spheres $\sigma_0, \dots, \sigma_6$ associated to edges $\eta_0, \dots, \eta_6$ in Γ. We denote the intersection circle between s and σ_i by α_i. The slice over s in $\mathrm{Core}(S, \Sigma)$ consists of squares associated to intersection circles between s and Σ, that is,

$$C_s = \{s \times \sigma_i \mid i = 0, \dots, 6\}.$$

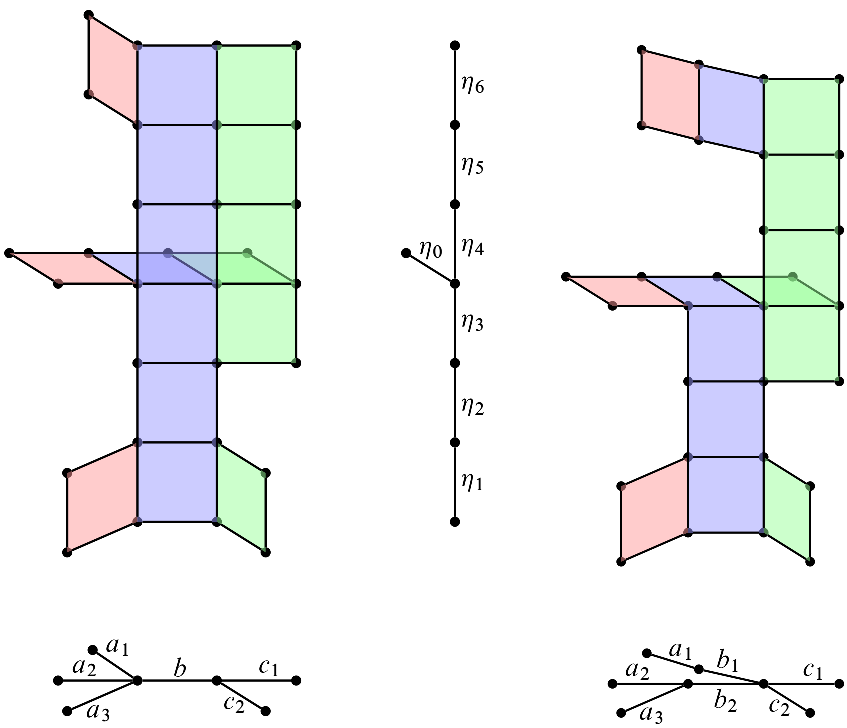


Figure 6: The left-hand side depicts the slice C_s and squares attached to it in $\mathrm{Core}(S, \Sigma)$. Consider the maximal S–boundary rectangle $R = b \times (\eta_4 \cup \eta_5)$. A Rips move along R is associated to a surgery on the sphere s or a splitting of the edge b in the graph G. The right side depicts the associated portion of $\mathrm{Core}(S, \Sigma)_R = \mathrm{Core}(S', R)$.

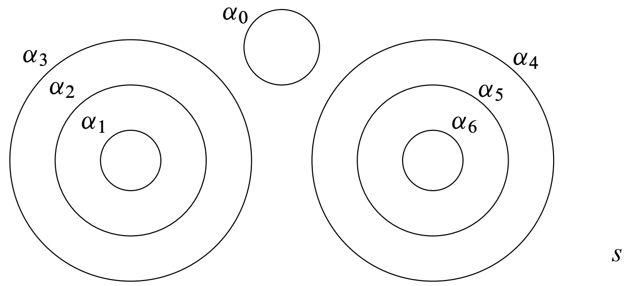


Figure 7: The curves $\alpha_0, \dots, \alpha_6$ are intersection circles between the sphere s and the spheres $\sigma_0, \dots, \sigma_6$, respectively. The circles α_2, α_3, α_4 and α_5 bound a disk in P_l. However, the circle α_3 and α_4 are not parallel.

In the language of trees, C_s is associated to the slice over b, which is

$$b \times \mathrm{Shadow}(s) = \{b \times \eta_i \mid i = 0, \dots, 6\} \subset G \times \Gamma.$$

There are two components of $M - S$ that have s as their boundary sphere. In this example, the component P_l, which we call left, has 3 other boundary spheres (associated to edges a_1, a_2 and a_3) and the component P_r on the right has two other boundary spheres (associated to edges c_1 and c_2).

Note that Figure 6 indicates that the sphere σ_1 intersects spheres in S associated to edges a_1, b and c_2 since the core contains squares $a_1 \times \eta_1$, $b \times \eta_1$ and $c_2 \times \eta_1$. However, the sphere σ_2 does not intersect spheres associated to edges a_1, a_2 and a_3. But, σ_2 intersects s, hence, the circle α_2 must bound a disk δ_2 that is the intersection of σ_2 with P_l. Similarly, circles α_3, α_4 and α_5 bound disks δ_3, δ_4 and δ_5 that are, respectively, intersections of spheres σ_3, σ_4 and σ_5 with P_l (thus the squares $b \times \eta_i$ for $i = 2, \dots, 5$ have boundary edges on their left side). The circle α_2 also bounds a disk in σ_2 in P_r (thus the square $b \times \eta_2$ has a boundary edge on its right side).

The disks δ_2 and δ_3 are parallel and the disks δ_4 and δ_5 are also parallel, however, the two sets of disks are not parallel to each other (see Figure 7). Thus, there are two maximal boundary rectangles from the left, $R = b \times (\eta_2 \cup \eta_3)$ and $R' = b \times (\eta_4 \cup \eta_5)$. More precisely, let Π be the component of $M - \Sigma$ with σ_0, σ_3 and σ_4 as its boundary spheres. Then, referring to Definition 5.4, we see that the vertex (P_l, Π) is not in the S–boundary of $\mathrm{Core}(S, \Sigma)$ because it is the vertex of 5 different edges. Hence, the union of R and R' is not a boundary rectangle.

Define S' to be the sphere system obtained from S by applying the surgery on the set of parallel disks $\{\delta_4, \delta_5\}$ (and their $\mathbb{F}_n$–orbits). The surgery results in two spheres s_1 and s_2 associated to edges b_1 and b_2 and the removal of the maximal boundary rectangle R.

It appears that removal of this rectangle makes the slice over b disconnected. However, the two components are slices over the edges b_1 and b_2.

To summarize, the surgery along the disks $\{\delta_4, \delta_5\}$ changes G by splitting the edge b and changes $\mathrm{Core}(S, \Sigma)$ by removing the maximal boundary rectangle $R = b \times (\eta_4 \cup \eta_5)$, resulting in $\mathrm{Core}(S, \Sigma)_R \cong \mathrm{Core}(S', \Sigma)$.

A different splitting of b into b_1 and b_2 partitioning the a edges into $\{a_1, a_3\}$ and $\{a_2\}$ does not arise as a surgery and could potentially increase the volume of the core.

5.3 The intersection of cores

Applying Theorem 5.5 to the surgery sequence $S = S_1, S_2, \dots, S_m$, we obtain maps, for $i = 1, \dots, m-1$,

$$k_{i,i+1}\colon \mathrm{Core}(S_{i+1}, \Sigma) \to \mathrm{Core}(S_i, \Sigma)$$

that are the composition of the isomorphism $\mathrm{Core}(S_{i+1}, \Sigma) \cong \mathrm{Core}(S_i, \Sigma)_R$ for the corresponding maximal boundary rectangle and the natural inclusion $\mathrm{Core}(S_i, \Sigma)_R \hookrightarrow \mathrm{Core}(S_i, \Sigma)$. By symmetry there are also maps, for $j = 1, \mu - 1$,

$$\kappa_{j,j+1}\colon \mathrm{Core}(S, \Sigma_{j+1}) \to \mathrm{Core}(S, \Sigma_j).$$

Since these maps exist for all $1 \le i \le m-1$ and $1 \le j \le \mu - 1$, we can define the "inclusions" alluded to at the beginning of this section,

$$k_i = k_{1,2}k_{2,3}\cdots k_{i-2,i-1}k_{i-1,i}\colon \mathrm{Core}(S_i, \Sigma) \to \mathrm{Core}(S, \Sigma), \tag{5-1}$$

$$\kappa_j = \kappa_{1,2}\kappa_{2,3}\cdots \kappa_{j-2,j-1}\kappa_{j-1,j}\colon \mathrm{Core}(S, \Sigma_j) \to \mathrm{Core}(S, \Sigma). \tag{5-2}$$

Remark 5.7 On the level of squares, the map $k_i\colon \mathrm{Core}(S_i, \Sigma) \to \mathrm{Core}(S, \Sigma)$ is easy to describe. For each $\hat{s} \in S_i$, we have that $\hat{s}^S \subseteq s$ for a unique $s \in S$. The map is defined by $\hat{s} \times \sigma \to s \times \sigma$.

The following is the fundamental concept essential to the proof of the main theorem.

Proposition 5.8 *With the above set-up, assume*

$$k_i(\mathrm{Core}(S_i, \Sigma)) \cup \kappa_j(\mathrm{Core}(S, \Sigma_j)) = \mathrm{Core}(S, \Sigma); \tag{5-3}$$

then S_i and Σ_j are in normal form. Furthermore, there exists an isomorphism

$$\Phi\colon \mathrm{Core}(S_i, \Sigma_j) \to k_i(\mathrm{Core}(S_i, \Sigma)) \cap \kappa_j(\mathrm{Core}(S, \Sigma_j)).$$

Proof First, we show that every intersection circle between S_i and Σ_j is in fact in $S_i^S \cap \Sigma_j^\Sigma$. This is because a square in $\mathrm{Core}(S, \Sigma)$ associated to an intersection circle in $S_i^\Sigma \cap \Sigma_j^S$ is neither in $k_i(\mathrm{Core}(S_i, \Sigma))$ (S_i^Σ does not intersect Σ) nor in

$\kappa_j(\mathrm{Core}(S, \Sigma_j))$ (Σ_j^S does not intersect S) and, by the assumption (5-3), every square in $\mathrm{Core}(S, \Sigma)$ is in the image of one of these two maps.

This observation implies that S_i and Σ_j are in fact in normal form. In fact, pick spheres $s_i \in S_i$ and $\sigma_j \in \Sigma_j$. We will show that s_i and σ_j intersect at most once. Otherwise, s_i^S and σ_j^Σ intersect more than once. But, by Lemma 5.3, s_i^S and σ_j^Σ are connected, which means there is a sphere $s \in S$ that contains s_i^S and a sphere $\sigma \in \Sigma$ that contains σ_j^Σ. Hence, s and σ intersect more than once. This contradicts the fact that S and Σ are in normal form.

Now consider a square $s_i \times \sigma_j$ in $\mathrm{Core}(S_i, \Sigma_j)$ associated to an intersection circle α. Then α is an intersection circle in $S_i^S \cap \Sigma_j^\Sigma$, which means it is an intersection circle in both $S \cap \Sigma_j$ and $S_i \cap \Sigma$, and thus there are spheres $s \in S$ and $\sigma \in \Sigma$ for which $s \cap \sigma = \alpha$ and $s_i^S \subseteq s$, $\sigma_j^\Sigma \subseteq \sigma$. Hence, $s \times \sigma$ is contained in both $k_i(\mathrm{Core}(S_i, \Sigma))$ and $\kappa_j(\mathrm{Core}(S, \Sigma_j))$ and so we define $\Phi(s_i \times \sigma_j) = s \times \sigma$. Normal form implies that the map is injective.

To prove that Φ is surjective, suppose $s \times \sigma$ is in $k_i(\mathrm{Core}(S_i, \Sigma))$. Then the associated intersection circle in S_i^S. Similarly, the assumption that $s \times \sigma$ is in $\kappa_j(\mathrm{Core}(S, \Sigma_j))$ implies that the associated intersection circle in Σ_j^Σ. Therefore, it also lies in $S_i \cap \Sigma_j$. Hence there are spheres $s_i \in S_i$ and $\sigma_j \in \Sigma_j$ such that $\Phi(s_i \times \sigma_j) = s \times \sigma$. □

For future reference, we record the following corollary:

Corollary 5.9 *If* (5-3) *is satisfied,* $\mathrm{Core}(S_i, \Sigma_{j+1})$ *can be obtained from* $\mathrm{Core}(S_i, \Sigma_j)$ *by a Rips move.*

Proof Let R be a maximal Σ_j–boundary rectangle in $\mathrm{Core}(S, \Sigma_j)$ such that

$$\mathrm{Core}(S, \Sigma_j)_R \cong \mathrm{Core}(S, \Sigma_{j+1}).$$

Thus R consists of squares $s_1 \times \hat{\sigma}, \dots, s_\ell \times \hat{\sigma}$ for some $\hat{\sigma} \in \Sigma_j$ and $s_1, \dots, s_\ell \in S$. Let $\sigma \in \Sigma$ be such that $\hat{\sigma}^\Sigma \subseteq \sigma$ and consider the set $C_{\sigma,i}$ of squares of the form $\hat{s} \times \sigma$ in $\mathrm{Core}(S_i, \Sigma)$. Then $k_i(C_{\sigma,i}) \cap \kappa_j(R)$ corresponds via the isomorphism in Proposition 5.8 to a maximal Σ_j–boundary rectangle in $\mathrm{Core}(S_i, \Sigma_j)$ whose collapse results in $\mathrm{Core}(S_i, \Sigma_{j+1})$. □

6 Proof of Theorem 1.1

To finish the proof of the main theorem, we proceed as follows, using the set-up from the previous section. We start with a lemma giving a necessary condition for two sphere systems to be at a bounded distance. A *free edge* is an edge that does not bound any squares.

Lemma 6.1 *If* $\mathrm{Core}(S_i, \Sigma_j)$ *contains a free edge, then the two sphere systems are of distance at most* 2 *in the sphere graph.*

Proof Edges in the core are associated to spheres in either sphere system S_i or Σ_j and squares are associated to intersection circles between sphere systems. Hence, a free edge in the core is associated to a sphere in either S_i or Σ_j that does not intersect any other spheres from the other system. Thus this sphere can be added to both sphere systems. That is, $\underline{S}_i$ and $\underline{\Sigma}_j$ have distance 2 in the sphere graph. □

We now prove Theorem 1.1. We restate it for convenience.

Theorem 1.1 *Let* $\underline{S}$ *and* $\underline{\Sigma}$ *be two filling sphere systems and let*

$$\underline{S} = \underline{S}_1, \underline{S}_2, \dots, \underline{S}_m, \quad d_{\mathcal{S}}(\underline{S}_m, \underline{\Sigma}) \leq 1,$$

be a surgery sequence starting from $\underline{S}$ *towards* $\underline{\Sigma}$ *and*

$$\underline{\Sigma} = \underline{\Sigma}_1, \underline{\Sigma}_2, \dots, \underline{\Sigma}_\mu, \quad d_{\mathcal{S}}(\underline{\Sigma}_\mu, \underline{S}) \leq 1,$$

be a surgery sequence in the opposite direction. Then, for every $\underline{S}_i$ *there is a* $\underline{\Sigma}_j$ *such that* $d_{\mathcal{S}}(\underline{S}_i, \underline{\Sigma}_j) \leq 2$.

Proof Fix two filling sphere systems $\underline{S}$ and $\underline{\Sigma}$ and surgery paths as in the statement of the theorem. For every $\underline{S}_i$ we need to find $\underline{\Sigma}_j$ with $d_{\mathcal{S}}(\underline{S}_i, \underline{\Sigma}_j) \leq 2$. Fix an $i = 1, \dots, m$ and let j be the largest index where the equality

$$k_i(\mathrm{Core}(S_i, \Sigma)) \cup \kappa_j(\mathrm{Core}(S, \Sigma_j)) = \mathrm{Core}(S, \Sigma) \tag{6-1}$$

still holds. Note that the equation holds when $j = 1$. But, since $\kappa_j(\mathrm{Core}(S, \Sigma_j))$ eventually contains no squares (for instance, when $j = \mu$) and $k_i(\mathrm{Core}(S_i, \Sigma))$ is a proper subset of $\mathrm{Core}(S, \Sigma)$ for each $i > 1$, there exists an index $j + 1$ for which (6-1) does not hold.

We will show that $\mathrm{Core}(S_i, \Sigma_j)$ contains a free edge. By Lemma 6.1, this will complete the proof. Let $s \times \sigma$ be a square in $\mathrm{Core}(S, \Sigma)$ that is not contained in

$$k_i(\mathrm{Core}(S_i, \Sigma)) \cup \kappa_{j+1}(\mathrm{Core}(S, \Sigma_{j+1})).$$

By (6-1), $s \times \sigma$ is contained in $\kappa_j(\mathrm{Core}(S, \Sigma_j))$. Thus a surgery on Σ_j has deleted the intersection circle associated to this square. By Corollary 5.9, $s \times \sigma$ is part of a maximal Σ_j–boundary rectangle. That is, there is a component $\Pi \subseteq M - \Sigma$ for which $\sigma \in \partial \Pi$ and such that the edge (s, Π) is a boundary edge of $s \times \sigma$ but not the boundary edge of any other square in $\kappa_j(\mathrm{Core}(S, \Sigma_j))$.

We also know that $s \times \sigma$ is not contained in $k_i(\mathrm{Core}(S_i, \Sigma))$. Thus, if (s, Π) is an edge in $k_i(\mathrm{Core}(S_i, \Sigma))$ then we have that (s, Π) is a free edge in

$$k_i(\mathrm{Core}(S_i, \Sigma)) \cap \kappa_j(\mathrm{Core}(S, \Sigma_j)) \cong \mathrm{Core}(S_i, \Sigma_j)$$

(Proposition 5.8). If this is not the case, then there is some $i_0 < i$ such that (s, Π) lies between two squares $s \times \sigma'$ and $s \times \sigma''$ that are part of a maximal S_{i_0}–boundary rectangle in $\mathrm{Core}(S_{i_0}, \Sigma)$ that is collapsed in the formation of $\mathrm{Core}(S_{i_0+1}, \Sigma)$. Then neither of these squares are in $k_i(\mathrm{Core}(S_i, \Sigma))$ and at least one of these squares is not in $\kappa_j(\mathrm{Core}(S, \Sigma_j))$. However, this would contradict (6-1). Therefore, (s, Π) is a free edge in $\mathrm{Core}(S_i, \Sigma_j)$. □

References

[1] **T Aougab**, *Uniform hyperbolicity of the graphs of curves*, Geom. Topol. 17 (2013) 2855–2875 MR

[2] **J Aramayona**, **J Souto**, *Automorphisms of the graph of free splittings*, Michigan Math. J. 60 (2011) 483–493 MR

[3] **J Behrstock**, **M Bestvina**, **M Clay**, *Growth of intersection numbers for free group automorphisms*, J. Topol. 3 (2010) 280–310 MR

[4] **M Bestvina**, **M Feighn**, *Stable actions of groups on real trees*, Invent. Math. 121 (1995) 287–321 MR

[5] **B H Bowditch**, *Uniform hyperbolicity of the curve graphs*, Pacific J. Math. 269 (2014) 269–280 MR

[6] **M Clay**, **K Rafi**, **S Schleimer**, *Uniform hyperbolicity of the curve graph via surgery sequences*, Algebr. Geom. Topol. 14 (2014) 3325–3344 MR

[7] **D Cooper**, *Automorphisms of free groups have finitely generated fixed point sets*, J. Algebra 111 (1987) 453–456 MR

[8] **T Coulbois**, **A Hilion**, *Rips induction: index of the dual lamination of an $\mathbb{R}$–tree*, Groups Geom. Dyn. 8 (2014) 97–134 MR

[9] **M Culler**, **K Vogtmann**, *Moduli of graphs and automorphisms of free groups*, Invent. Math. 84 (1986) 91–119 MR

[10] **V Guirardel**, *Cœur et nombre d'intersection pour les actions de groupes sur les arbres*, Ann. Sci. École Norm. Sup. 38 (2005) 847–888 MR

[11] **M Handel**, **L Mosher**, *The free splitting complex of a free group, I: Hyperbolicity*, Geom. Topol. 17 (2013) 1581–1672 MR

[12] **A Hatcher**, *Homological stability for automorphism groups of free groups*, Comment. Math. Helv. 70 (1995) 39–62 MR

[13] **A Hatcher**, **K Vogtmann**, *Isoperimetric inequalities for automorphism groups of free groups*, Pacific J. Math. 173 (1996) 425–441 MR

[14] **S Hensel**, **D Osajda**, **P Przytycki**, *Realisation and dismantlability*, Geom. Topol. 18 (2014) 2079–2126 MR

[15] **S Hensel**, **P Przytycki**, **R C H Webb**, *1–slim triangles and uniform hyperbolicity for arc graphs and curve graphs*, J. Eur. Math. Soc. 17 (2015) 755–762 MR

[16] **A Hilion**, **C Horbez**, *The hyperbolicity of the sphere complex via surgery paths*, J. Reine Angew. Math. (online publication March 2015)

[17] **C Horbez**, *Sphere paths in outer space*, Algebr. Geom. Topol. 12 (2012) 2493–2517 MR

[18] **F Laudenbach**, *Topologie de la dimension trois: homotopie et isotopie*, Astérisque 12, Soc. Math. France, Paris (1974) MR

[19] **P Przytycki**, **A Sisto**, *A note on acylindrical hyperbolicity of mapping class groups*, preprint (2015) arXiv

MC: *Department of Mathematical Sciences, University of Arkansas*
Fayetteville, AR, United States

YQ, KR: *Department of Mathematics, University of Toronto*
Toronto, ON, Canada

`mattclay@uark.edu`, `yulan.qing@utoronto.ca`, `rafi@math.toronto.edu`

`http://comp.uark.edu/~mattclay`, `https://sites.google.com/site/yulanqing/home`, `http://www.math.toronto.edu/~rafi/`

Received: 29 October 2016 Revised: 24 February 2017

Algebraic & Geometric Topology 17 (2017) 3779–3810

On the integral cohomology ring of toric orbifolds and singular toric varieties

ANTHONY BAHRI
SOUMEN SARKAR
JONGBAEK SONG

We examine the integral cohomology rings of certain families of $2n$–dimensional orbifolds X that are equipped with a well-behaved action of the n–dimensional real torus. These orbifolds arise from two distinct but closely related combinatorial sources, namely from characteristic pairs (Q,λ), where Q is a simple convex n–polytope and λ a labeling of its facets, and from n–dimensional fans Σ. In the literature, they are referred as toric orbifolds and singular toric varieties, respectively. Our first main result provides combinatorial conditions on (Q,λ) or on Σ which ensure that the integral cohomology groups $H^*(X)$ of the associated orbifolds are concentrated in even degrees. Our second main result assumes these conditions to be true, and expresses the graded ring $H^*(X)$ as a quotient of an algebra of polynomials that satisfy an integrality condition arising from the underlying combinatorial data. Also, we compute several examples.

14M25, 55N91, 57R18; 13F55, 52B11

1 Introduction

There are several advantages to studying topological spaces whose integral cohomology groups $H^*(X)$ are torsion-free and concentrated in even degrees; for example, their complex K–theory and complex cobordism groups may be deduced immediately, because the appropriate Atiyah–Hirzebruch spectral sequences collapse for dimensional

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3779

reasons. For convenience, we call such spaces *even*, where integral coefficients are understood unless otherwise stated. Our fundamental aim is to identify certain families of even spaces within the realms of toric topology, and to explain how their evenness leads to a description of the Borel equivariant cohomology rings $H^*_T(X)$, and thence to the multiplicative structure of $H^*(X)$.

Many even spaces arise from complex geometry, and have been of major importance since the early 20^{th} century. They range from complex projective spaces and Grassmannian manifolds to Thom spaces of complex vector bundles over other even spaces. Examples of the latter include stunted projective spaces, which play an influential and enduring role in homotopy theory, and certain restricted families of weighted projective spaces. In fact *every* weighted projective space is even, thanks to a beautiful and somewhat surprising result of Kawasaki [18], whose calculations lie behind one of our main works in Section 4. In the literature, weighted projective spaces have been viewed as singular toric varieties or as toric orbifolds, which we shall define in Section 3, and our results may be interpreted as an investigation of their generalizations within either context.

We begin in Section 2 by introducing a sequence $\{B_k\}$ of polytopal complexes whose initial term is a simple polytope Q and final term is a vertex of Q. We define the sequence inductively by the rule stated as 2 in Section 2, which is motivated by several spaces called *invariant subspaces*, and *orbifold lens spaces* sitting inside the given toric orbifold.

In Section 3, we summarize the theory of *toric orbifolds* $X = X(Q, \lambda)$,[1] as constructed from an n–dimensional simple convex polytope Q and an $\mathcal{R}$*–characteristic function* λ from its facets to $\mathbb{Z}^n$. The combinatorial data (Q, λ) is called an $\mathcal{R}$*–characteristic pair* associated to the given toric orbifold. The notion of *invariant subspaces* and *orbifold lens spaces* follow from (Q, λ), which we shall explain in the following subsections. Moreover, for each polytopal complex B which appears in a retraction sequence, the $\mathcal{R}$–characteristic function λ may be used to associate a finite group $G_B(v)$ — see (4-8) — to certain vertices v, called free vertices in B, and to define the collection

$$\big\{|G_B(v)| : v \text{ is a } \textit{free vertex} \text{ in } B\big\}. \tag{1-1}$$

Interest in toric orbifolds was stimulated by Davis and Januszkiewicz [9], who saw them as natural extensions to their own smooth toric manifolds.[2] They proved that toric manifolds are always even; however, the best comparable statement for toric *orbifolds* is due to Poddar and the second author [20], who showed that, in general, they are only

[1] In the literature, these orbifolds are sometimes called *quasitoric* orbifolds.

[2] They are renamed in Buchstaber and Panov [4] as *quasitoric* manifolds.

even over the rationals. We introduce our main result of the first part of this paper in Section 4, as follows.

Theorem 1.1 *Given any toric orbifold $X(Q,\lambda)$, assume that the gcd of the collection* (1-1) *is* 1 *for each B which appears in a retraction sequence with* $\dim B > 1$*; then X is even.*

The proof employs a cofiber sequence involving *orbifold lens spaces*, which are a generalization of *lens complexes*, introduced by Kawasaki [18]. Furthermore, Theorem 1.1 automatically applies to weighted projective spaces.

In Section 5, we restrict our emphasis to projective toric orbifolds, which are realized as toric varieties whose details are admirably presented by Cox, Little and Schenck in their encyclopedic book [6]. Every such variety X_Σ is encoded by a fan Σ in $\mathbb{R}^n$, and admits a canonical action by the n–dimensional real torus T^n. If Σ is *smooth*, then the underlying geometry guarantees that X_Σ is always even. Moreover, it is true that the Borel equivariant cohomology ring $H^*_T(X_\Sigma)$ is isomorphic to the Stanley–Reisner ring $\mathcal{SR}[\Sigma]$, which is also concentrated in even degrees, and $H^*(X_\Sigma)$ is its quotient by a linear ideal determined by (5-2). It is important to note that $\mathcal{SR}[\Sigma]$ is isomorphic to the ring $\mathcal{PP}[\Sigma]$ of *integral piecewise polynomials* on Σ for any *smooth* fan.

For a particular class of singular examples, a comparable description of the ring $H^*(X_\Sigma)$ was given in Bahri, Franz and Ray [1], as follows. If Σ is polytopal and X_Σ is even, then $H^*(X_\Sigma)$ is the quotient of $\mathcal{PP}[\Sigma]$ by the ideal generated by all *global* polynomials. It is no longer possible to use the Stanley–Reisner ring, which only agrees with $\mathcal{PP}[\Sigma]$ over the rationals. In these circumstances, when X_Σ is a toric variety over a polytopal fan, we have a major incentive to develop criteria which test whether or not it is even. There also remains the significant problem of presenting $\mathcal{PP}[\Sigma]$ by generators and relations, as exemplified by the calculation for the weighted projective space $\mathbb{CP}^3_{(1,2,3,4)}$ in [1, Section 4]. So the aim of Section 5 is to find an alternative description for the ring of piecewise polynomials. It is accomplished by defining the *weighted Stanley–Reisner ring* $w\mathcal{SR}[\Sigma]$, which turns out to be a subring of $\mathcal{SR}[\Sigma]$, consisting of polynomials that satisfy an *integrality condition*; see Definition 5.2. The main result of Section 5 combines Theorems 1.1 and 5.3, as follows.

Theorem 1.2 *Given any polytopal fan Σ in $\mathbb{R}^n$, assume that the corresponding $\mathcal{R}$–characteristic pair (Q,λ) satisfies the hypothesis of Theorem 1.1; then X_Σ is even, and there exists an isomorphism*

$$H^*(X_\Sigma) \cong w\mathcal{SR}[\Sigma]/\mathcal{J}$$

of graded rings, where $\mathcal{J}$ is an ideal of linear relations determined by the generators of rays of Σ.

So our combinatorial condition on the fan allows us to give an explicit description of the integral cohomology ring of X_Σ.

Several natural questions present themselves for future discussion. For example, Sections 3 and 5 may be linked more closely by establishing a common framework for toric orbifolds and toric varieties over nonsmooth polytopal fans. The theory of multifans is an obvious candidate, but we have been unable to identify an associated ring of piecewise polynomials with sufficient clarity. However, the third author with Darby and Kuroki [8] has recently proposed a definition of piecewise polynomials on an orbifold torus graph, which does allow those two objects to be dealt with simultaneously.

In view of our opening remarks, another reasonable challenge is to extend our study to the complex K–theory and complex cobordism of toric orbifolds. This program was suggested by work of Harada, Henriques and Holm [13], and begun in Harada, Holm, Ray and Williams [14] by the adoption of a categorical approach to piecewise structures; but overall progress has been limited to a small subfamily of weighted projective spaces, and much further work is required. However, some progress has made by the second author and Uma [22].

Acknowledgments We extend our sincere gratitude to Nigel Ray. Most properly, his name belongs among those of the authors. The significance of his mathematical contribution to the results is most certainly not reflected accurately by the omission of his name. Among his many other contributions to the paper is the concise and elegant introduction. We thank Mikiya Masuda and Haozhi Zeng for pointing out a gap in the proof of Theorem 2.12 of the previous version, and Li Cai for helpful comments about retraction sequences. We are also grateful for the hospitality of University of Calgary in July, 2015.

This work has been supported in part by Simons Foundation grants 210368 and 426160. Bahri acknowledges the award of research leave from Rider University and is grateful also for the hospitality of the Princeton University Mathematics Department during the spring of 2015. Sarkar would like to thank Pacific Institute for the Mathematical Sciences and University of Regina for financial support. Lastly, Song was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Korea government (MSIP), No. NRF-2016R1A2B4010823. He also would like to thank his advisor Professor Dong Youp Suh for his encouragements and support throughout the project.

2 A retraction of simple polytopes

In this section, we introduce a natural way of retracting a simple polytope Q to a point, which we call a *retraction sequence*. For each polytope, there are finitely many such retractions, enabling us to develop a sufficient condition for torsion-freeness in the homology of toric orbifolds in the following section. The operation itself is motivated by several spaces which arise in a toric orbifold by decomposing the orbit space. We shall explain this topological interpretation in Section 3. This section is devoted to giving the combinatorial definition and properties of retraction sequences. We begin by introducing the definition of a polytopal complex.

Definition 2.1 [23, Definition 5.1] A *polytopal complex* $\mathcal{C}$ is a finite collection of polytopes in $\mathbb{R}^n$ satisfying:

(1) If E is a face of F and $F \in \mathcal{C}$ then $E \in \mathcal{C}$.

(2) If $E, F \in \mathcal{C}$ then $E \cap F$ is a face of both E and F.

Let $|\mathcal{C}| = \bigcup_{F \in \mathcal{C}} F$ be the underlying set of $\mathcal{C}$.

The elements of $\mathcal{C}$ are called faces and the zero-dimensional faces of $\mathcal{C}$ are called vertices. We denote the set of vertices of $\mathcal{C}$ by $V(|\mathcal{C}|)$. The dimension of $\mathcal{C}$ or $|\mathcal{C}|$ is the maximum of the dimension of its faces. Given a simple polytope Q, let $\mathcal{C}(Q)$ be the collection of all faces of Q and $\mathcal{F}(Q)$ the collection of all facets of Q. Then $\mathcal{C}(Q)$ is a polytopal complex and $|\mathcal{C}(Q)|$ is homeomorphic to Q as manifolds with corners. Throughout this paper, we always write $\ell := |V(Q)|$ for the number of vertices of Q, $m := |\mathcal{F}(Q)|$ for the number of facets of Q and $n := \dim Q$.

Now, given an n–dimensional simple polytope Q, we construct a sequence of triples $\{(B_k, E_k, b_k)\}_{k=1}^{\ell}$, which we call a *retraction sequence* of Q. First, we define $B_1 = Q = E_1$ and $b_1 \in V(B_1)$. The second term (B_2, E_2, b_2) is defined as follows. Consider a subcollection

$$\mathcal{C}_2 = \{E \in \mathcal{C}(Q) \mid b_1 \notin V(E)\}$$

of $\mathcal{C}(Q)$. Then $\mathcal{C}_2$ is an $(n-1)$–dimensional polytopal complex. We define B_2 by the underlying set $|\mathcal{C}_2|$ of $\mathcal{C}_2$. We choose a vertex b_2 of B_2 such that b_2 has a neighborhood diffeomorphic to $\mathbb{R}^N_{\geq 0}$ as manifolds with corners for some $1 \leq N \leq \dim B_2$ and let E_2 be the unique N–dimensional face of B_2 containing b_2. Notice that, in this case, $N = n-1$ and we have n different choices of b_2 because Q is an n–dimensional simple polytope.

Next we construct the sequence of triples inductively. Given (B_k, E_k, b_k), the next term $(B_{k+1}, E_{k+1}, b_{k+1})$ is defined as follows. First we consider a polytopal complex

$$\mathcal{C}_{k+1} = \{E \in \mathcal{C}_k \mid b_k \notin V(E)\}.$$

Then B_{k+1} is defined by its underlying set $|\mathcal{C}_{k+1}|$. We choose a vertex b_{k+1} in $V(B_{k+1})$ satisfying the condition

($\diamond$) b_{k+1} has a neighborhood homeomorphic to $\mathbb{R}^N_{\geq 0}$ as manifolds with corners for some $N \in \{1, \dots, \dim B_{k+1}\}$,

and E_{k+1} defined to be a unique face of B_{k+1} containing b_{k+1} with $\dim E_{k+1} = N$.

Definition 2.2 We call a vertex v in B_k a *free vertex* if it has a neighborhood in B_k that is diffeomorphic to $\mathbb{R}^N_{\geq 0}$ as manifolds with corners for some $N \in \{1, \dots, \dim B_k\}$. We denote the set of free vertices in B_k by $\mathrm{FV}(B_k)$.

Finally, the sequence stops if the sequence reaches a vertex, ie $B_\ell = E_\ell = b_\ell \in V(Q)$. Essentially, we can think of a retraction sequence as an iterated choice of free vertices at each step. Figure 1 shows an example of retraction sequence for the vertex cut of a cube, where the colored face of each B_k indicates E_k for $k = 1, \dots, 10$.

Proposition 2.3 *Every simple polytope has at least one retraction sequence.*

Proof We begin by following the argument of [9, Proposition 3.1]. First, we realize Q as a convex polytope in $\mathbb{R}^n$ and choose a vector $u \in \mathbb{R}^n$ such that

$$\langle u, v\rangle \neq \langle u, w\rangle \quad \text{whenever } v \neq w \in V(Q) \subset \mathbb{R}^n,$$

with respect to the Euclidean inner product $\langle\, ,\rangle$. Let $e := e(vw)$ be the oriented edge with the initial vertex $i(e) = v$ and the terminal vertex $t(e) = w$. Here the direction of $e(vw)$ is given by the rule

$$i(e) = v \quad \text{and} \quad t(e) = w \qquad \text{if and only if} \qquad \langle u, v\rangle < \langle u, w\rangle,$$

which makes the one-skeleton of Q into a directed graph.

Let $\mathrm{ind}(v)$ be the number of inward edges at v and we call $\mathrm{ind}(v)$ the index of v (with respect to the choice of generic vector u). Then, for each face $E \subset Q$, there exists a unique vertex v of E having the maximal index among the vertices in E. Moreover, E is locally diffeomorphic to $\mathbb{R}^{\mathrm{ind}(v)}_{\geq 0}$ around v. Conversely, given a vertex $v \in V(Q)$, there exists a unique face E_v such that $\dim E_v = \mathrm{ind}(v)$.

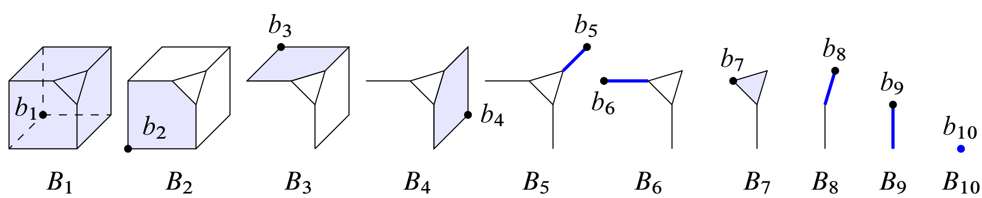


Figure 1: A retraction sequence of a vertex cut of the cube

Let $\{b_k\}_{k=1}^{\ell}$ be a sequence of vertices in Q determined by

$$\langle u, b_1\rangle > \langle u, b_2\rangle > \cdots > \langle u, b_\ell\rangle.$$

Notice that $\operatorname{ind}(b_1) = n = \dim Q$, and $\operatorname{ind}(b_\ell) = 0$. Now we claim that the sequence

$$\left\{\left(B_k := \bigcup_{j\geq k} E_{b_j}, E_{b_k}, b_k\right)\right\}_{k=1,\dots,\ell},$$

where E_{b_k} is a unique face containing b_k with $\dim E_{b_k} = \operatorname{ind}(b_k)$, is a retraction sequence of Q. Indeed, for each $k \in \{1, \dots, \ell-1\}$, we have $\langle u, b_k\rangle > \langle u, v\rangle$ for all $v \in V(B_k) \setminus \{b_k\}$. Hence, there are no outgoing edges from b_k in B_k, which implies that b_k has a neighborhood in $E_{b_k} \subseteq B_k$ homeomorphic to $\mathbb{R}_{\geq 0}^{\operatorname{ind}(b_k)}$ as manifolds with corners. □

We denote by $\mathfrak{R}(Q)$ the set of all retraction sequences of Q and by $\mathfrak{B}(Q)$ the set of all possible B_i which appear in $\mathfrak{R}(Q)$. Evidently, both $\mathfrak{R}(Q)$ and $\mathfrak{B}(Q)$ are finite sets, because we have finitely many choices of free vertices at each step.

Remark The retraction sequence has a strong relation with *shelling* of a simplicial complex. We are preparing an independent article [2] about the exact correspondence and some other interesting properties.

3 Toric orbifolds and orbifold lens spaces

In this section we recall the *characteristic pairs* (Q, λ) of [9; 20], and explain the way in which they are used to construct *toric orbifolds* $X = X(Q, \lambda)$. If λ obeys Davis and Januszkiewicz's condition $(*)$ (see [9, page 423]), then X is smooth and even; so one of the main goals of this paper is to establish Theorem 1.1, which focuses on singular cases, and states a sufficient condition for the orbifold X to be even. In this section, to complete the proof of Theorem 1.1, we commandeer two additional types of spaces, namely the *invariant subspaces* of X which arise as the preimage of faces via the orbit map, and the *orbifold lens spaces* that arise as quotients of odd-dimensional spheres by the actions of certain finite groups associated to λ.

3.1 Toric orbifolds

In this subsection, we discuss a combinatorial definition of toric orbifolds. Let Q be an n–dimensional simple convex polytope in $\mathbb{R}^n$ and $\mathscr{F}(Q) = \{F_1, \dots, F_m\}$ the codimension-one faces of Q, which are called *facets*.

Definition 3.1 A function $\lambda\colon \mathscr{F}(Q) \to \mathbb{Z}^n$ is called a *rational characteristic function* (or *$\mathcal{R}$–characteristic function*) for Q if it satisfies the following condition:

$$\{\lambda(F_{i_1}), \dots, \lambda(F_{i_k})\} \text{ is linearly independent whenever } \bigcap_{j=1}^{k} F_{i_j} \neq \varnothing. \tag{3-1}$$

We write $\lambda_i = \lambda(F_i)$ and call it an *$\mathcal{R}$–characteristic vector* assigned to the facet F_i. The pair (Q, λ) is called an $\mathcal{R}$–characteristic pair.

Remark (1) In the literature about toric manifolds, the pair (Q, λ) satisfying the condition $(*)$ in [9, page 423] is called a *characteristic pair*.

(2) For convenience, we usually express an $\mathcal{R}$–characteristic function λ as an $n \times m$ matrix Λ by listing the λ_i as column vectors. We call Λ an *$\mathcal{R}$–characteristic matrix* associated to λ.

(3) It is easy to check that it suffices to satisfy the linearly independence at each vertex which is an intersection of n facets.

One canonical example of such functions can be given by a *simple lattice polytope*, which is a convex hull of finitely many points in the integer lattice $\mathbb{Z}^n \subset \mathbb{R}^n$ and simple. Namely, we can naturally assign as an $\mathcal{R}$–characteristic vector the primitive normal vector on each facet of a simple lattice polytope. In Section 5, we shall see this again as primitive vectors of 1–dimensional cones in a normal fan associated to a simple lattice polytope.

For $x \in Q$, we denote by $E(x)$ the face of Q which contains x in its interior. If $E(x)$ is a face of codimension k, then it is a unique intersection of k facets $F_{i_1}, \dots, F_{i_k}$. We also denote by $T_{E(x)}$ the subtorus of the standard n–dimensional torus T^n determined by $\lambda_{i_1}, \dots, \lambda_{i_k}$. To be more precise, we may regard the target space $\mathbb{Z}^n$ of λ as the $\mathbb{Z}$–submodule of the Lie algebra of T^n, and $T_{E(x)}$ is the torus generated by the exponential image of the lines determined by the $\mathcal{R}$–characteristic vectors $\lambda_{i_1}, \dots, \lambda_{i_k}$.

Now we define an equivalence relation $\sim_\lambda$ on the product $T^n \times Q$ by

$$(t, x) \sim_\lambda (s, y) \quad \text{if and only if} \quad x = y \quad \text{and} \quad t^{-1}s \in T_{E(x)}. \tag{3-2}$$

The quotient space

$$X(Q, \lambda) = (T^n \times Q)/{\sim_\lambda}$$

has an orbifold structure with a natural T^n–action induced by the group operation; see Section 2 in [20]. Clearly, the orbit space of the T^n–action on $X(Q, \lambda)$ is Q. Let

$$\pi\colon X(Q, \lambda) \to Q, \quad \pi([t, x]_{\sim_\lambda}) = x, \tag{3-3}$$

be the orbit map, where $[t,x]_{\sim_\lambda}$ is the equivalence class of (t,x) with respect to $\sim_\lambda$. The space $X(Q,\lambda)$ is called the toric orbifold associated to the combinatorial pair (Q,λ).

In analyzing the orbifold structure of $X(Q,\lambda)$, Poddar and Sarkar [20, Section 2.2], gave an axiomatic definition of toric orbifolds, which generalizes the axiomatic definition of toric manifolds of [9].

3.2 Invariant subspaces

In this subsection, we study the $\mathcal{R}$–characteristic pair of some invariant subspaces of $X(Q,\lambda)$. Let $E = F_{i_1} \cap \cdots \cap F_{i_k}$ be a face of Q, where $F_{i_1}, \dots, F_{i_k}$ are facets. We can define a natural projection

$$\rho_E\colon \mathbb{Z}^n \to \mathbb{Z}^n / \big((\mathrm{span}\{\lambda_{i_1}, \dots, \lambda_{i_k}\} \otimes_{\mathbb{Z}} \mathbb{R}) \cap \mathbb{Z}^n\big), \tag{3-4}$$

where the target space is isomorphic to $\mathbb{Z}^{n-k}$, because $(\mathrm{span}\{\lambda_{i_1}, \dots, \lambda_{i_k}\} \otimes_{\mathbb{Z}} \mathbb{R}) \cap \mathbb{Z}^n$ is a rank-k direct summand of $\mathbb{Z}^n$. Notice that the rank of the target space of ρ_E is the same as the dimension of E. We consider E as an independent simple polytope, and denote the set of facets of E by

$$\mathscr{F}(E) = \{E \cap F_j \mid F_j \in \mathscr{F}(Q) \text{ and } j \neq i_1, \dots, i_k \text{ and } E \cap F_j \neq \varnothing\}.$$

Now the map ρ_E, together with λ, yields an $\mathcal{R}$–characteristic function

$$\lambda_E\colon \mathscr{F}(E) \to \mathbb{Z}^{n-k} \tag{3-5}$$

on E, defined so that $\lambda_E(E \cap F_j)$ is the primitive vector of $(\rho_E \circ \lambda)(F_j)$. Indeed, the condition (3-1) naturally follows from λ.

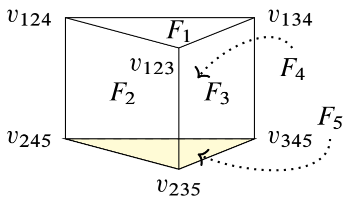
Hence, we get an $\mathcal{R}$–characteristic pair (E, λ_E) from (Q,λ), which yields another toric orbifold

$$X(E,\lambda_E) := (T^{n-k} \times E)/{\sim_{\lambda_E}},$$

where the equivalence relation $\sim_{\lambda_E}$ is defined in a manner similar to (3-2).

Proposition 3.2 [20, Section 2.3] *Let $\pi\colon X(Q,\lambda) \to Q$ and (E,λ_E) be as above. Then $\pi^{-1}(E)$ is a T^n–invariant suborbifold. Moreover, it is a toric orbifold homeomorphic to $X(E,\lambda_E)$ as a topological space.*

The second assertion of the above proposition follows from the fact that the circle subgroups determined by $\lambda_E(E \cap F_j)$ and $(\rho_E \circ \lambda)(F_j)$, respectively, are identical.



$$\Lambda = \begin{array}{c} \begin{matrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \end{matrix} \\ \begin{bmatrix} 2 & 2 & -1 & -1 & 0 \\ 3 & -1 & -1 & 2 & 0 \\ 5 & 0 & -2 & 2 & 1 \end{bmatrix} \end{array}$$

Figure 2

We also remark that the torus T^{n-k} acting on $X(E, \lambda_E)$ can be identified with the image of the map

$$\bar{\rho}_E\colon T^n \to T^{n-k}, \tag{3-6}$$

which is induced from the map ρ_E.

Example 3.3 Suppose we have an $\mathcal{R}$–characteristic pair (Q, λ) as described in Figure 2. Notice that Q is a 3–dimensional polytope with 5 facets, say $\mathscr{F}(Q) = \{F_1, \dots, F_5\}$. Here we assume that the target space $\mathbb{Z}^3$ of λ is generated by the standard basis $\{e_1, e_2, e_3\}$. We choose E to be the facet F_5. So $k = 1$ and $n - k = 2$. Then the projection

$$\rho_E\colon \mathbb{Z}^3 \to \mathbb{Z}^3/\langle e_3\rangle = \langle e_1, e_2, e_3\rangle/\langle e_3\rangle \cong \mathbb{Z}^2$$

is onto the first two coordinates. The facets of E are $F_2 \cap E$, $F_3 \cap E$ and $F_4 \cap E$. Hence, the map

$$\lambda_E\colon \{F_2 \cap E,\ F_3 \cap E,\ F_4 \cap E\} \to \mathbb{Z}^2$$

is defined by

$$\begin{aligned} \lambda_E(F_2 \cap E) &= \rho_E(\lambda(F_2)) = (2, -1) = 2e_1 - e_2, \\ \lambda_E(F_3 \cap E) &= \rho_E(\lambda(F_3)) = (-1, -1) = -e_1 - e_2, \\ \lambda_E(F_4 \cap E) &= \rho_E(\lambda(F_4)) = (-1, 2) = -e_1 + 2e_2. \end{aligned}$$

The orbifold corresponding to (E, λ_E) is known to be a fake weighted projective space with weight $(1, 1, 1)$. We refer to [5; 17] for the details of fake weighted projective space.

3.3 Orbifold lens spaces

Here we introduce a generalization of *lens complexes* and study their homology groups. Let Δ^{n-1} be the $(n-1)$–dimensional simplex and $\mathscr{F}(\Delta^{n-1}) = \{F_1, \dots, F_n\}$ the facets of Δ^{n-1}. We begin by introducing the following definition.

Definition 3.4 A function $\xi\colon \mathscr{F}(\Delta^{n-1}) \to \mathbb{Z}^n$ is called an *$\mathcal{L}$–characteristic function* on Δ^{n-1} if $\{\xi(F_1), \dots, \xi(F_n)\}$ is linearly independent. We set $\xi_i := \xi(F_i)$ for $i = 1, \dots, n$.

Now we define an equivalence relation $\sim_\xi$ on $T^n \times \Delta^{n-1}$ as follows:

$$(t, x) \sim_\xi (s, y) \qquad \text{if and only if} \qquad x = y \quad \text{and} \quad t^{-1}s \in T_{F(x)}, \tag{3-7}$$

where $F(x)$ is the face containing x in its interior and $T_{F(x)}$ denotes the subtorus of T^n determined by $\xi_{i_1}, \dots, \xi_{i_k}$ if $F(x) = F_{i_1} \cap \dots \cap F_{i_k}$. The pair (Δ^{n-1}, ξ), together with the equivalence relation $\sim_\xi$, yields the quotient space

$$L(\Delta^{n-1}, \xi) := T^n \times \Delta^{n-1}/{\sim_\xi},$$

which we call the *orbifold lens space* associated to (Δ^{n-1}, ξ).

Proposition 3.5 *The orbifold lens space $L(\Delta^{n-1}, \xi)$ is homeomorphic to the quotient space of the $(2n-1)$–dimensional sphere S^{2n-1} by the action of a finite group $G_\xi := \mathbb{Z}^n/\mathrm{span}\{\xi_1, \dots, \xi_n\}$.*

Proof The proof is essentially same as the proof of [21, Proposition 2.3]. □

Remark (1) In [21], the function ξ is called a *hypercharacteristic function* if the submodule generated by $\{\xi(F_{i_1}), \dots, \xi(F_{i_k})\}$ is a direct summand of $\mathbb{Z}^{n+1}$ of rank k whenever $F_{i_1} \cap \dots \cap F_{i_k}$ is nonempty. In particular, if $\{\xi(F_{i_1}), \dots, \xi(F_{i_n})\}$ is a linearly independent set, then it becomes an $\mathcal{L}$–characteristic function.

(2) The action of G_ξ is induced from the standard T^n–action on $S^{2n-1} \subset \mathbb{C}^n$.

(3) The order $|G_\xi|$ of G_ξ is exactly same as the determinant of the $n \times n$ matrix $[\xi_1 \mid \cdots \mid \xi_n]$.

Proposition 3.5 leads us to the following lemma.

Lemma 3.6 *Let $p_1, \dots, p_r$ be the prime factors of $|G_\xi|$. Then*

$$H_j(L(\Delta^{n-1}, \xi)) = \begin{cases} \mathbb{Z} & \text{if } j = 0, 2n-1, \\ G_j & \text{if } 1 \le j \le 2n-2, \end{cases}$$

where $G_j = (\mathbb{Z}/p_1^{a_{j1}}\mathbb{Z}) \oplus \dots \oplus (\mathbb{Z}/p_r^{a_{jr}}\mathbb{Z})$ for some nonnegative integers $a_1, \dots, a_r$.

Proof We see $H_0(L(\Delta^{n-1}, \xi)) \cong \mathbb{Z}$ trivially. The isomorphism $H_{2n-1}(L(\Delta^n, \xi)) \cong \mathbb{Z}$ follows because the G_ξ–action on S^{2n-1} is induced from the standard action of T^n on $S^{2n-1} \subset \mathbb{C}^n$, which is orientation-preserving. For $j \in \{1, \dots, 2n-2\}$, recall the following isomorphism, which can be obtained from the classical result for an action of a finite group G on a locally compact Hausdorff space X:

$$H^*(X/G;\boldsymbol{k}) \cong H^*(X;\boldsymbol{k})^G, \tag{3-8}$$

where $\boldsymbol{k}$ is a field of characteristic zero or prime to $|G|$; see [3, III.2].

We apply the isomorphism (3-8) to the orbifold lens space $L(\Delta^{n-1},\xi) \cong S^{2n-1}/G_\xi$. Since $H^j(S^{2n-1};\boldsymbol{k})^{G_\xi} = 0$ for $j = 1,\dots 2n-2$, the claim is proved by the universal coefficient theorem. □

Toric orbifolds, invariant subspaces and orbifold lens spaces motivate the definition of retraction sequences which we introduced in the previous section. For a vertex $v \in V(Q)$, let B_2 be the union of all faces in Q which do not contain v. Next we consider a hyperplane

$$H(v) := \{x \in \mathbb{R}^n \mid \langle x, p_v\rangle = q_v\}, \tag{3-9}$$

where $\langle\,,\rangle$ denotes the Euclidean inner product and $p_v \in \mathbb{R}^n$ and $q_v \in \mathbb{R}$ are chosen in such a way that

- $\{x \in \mathbb{R}^n \mid \langle x, p_v\rangle + q_v \geq 0\} \cap V(Q) = \{v\}$,
- $\{x \in \mathbb{R}^n \mid \langle x, p_v\rangle + q_v \leq 0\} \cap V(Q) = V(Q) \setminus \{v\}$.

Then $\Delta_Q(v) := Q \cap H(v)$ is an $(n-1)$–dimensional simplex, because Q is a simple polytope of dimension n; see Figure 3.

An $\mathcal{L}$–characteristic pair arises naturally from an $\mathcal{R}$–characteristic pair (Q,λ) for each vertex v of Q. Indeed, if $v = F_{j_1} \cap \cdots \cap F_{j_n}$, we denote the set of facets of $\Delta_Q(v)$ by

$$\mathscr{F}(\Delta_Q(v)) = \{\Delta_Q(v) \cap F_{j_1}, \dots, \Delta_Q(v) \cap F_{j_n}\}.$$

Now we define a function

$$\xi_{Q,v}\colon \mathscr{F}(\Delta_Q(v)) \to \mathbb{Z}^n \tag{3-10}$$

by $\xi_{Q,v}(\Delta_Q(v) \cap F_{j_r}) = \lambda(F_{j_r})$ for $r = 1,\dots,n$. Notice that $\dim \Delta_Q(v) = n-1$, but the rank of target space is n. Since $\{\lambda(F_{i_1})\dots,\lambda(F_{i_n})\}$ is a linearly independent set, the function $\xi_{Q,v}$ is an $\mathcal{L}$–characteristic function on $\Delta_Q(v)$.

4 Vanishing odd degree homology and torsion-freeness

Now we combine the ingredients which we introduced in the previous sections to derive a sufficient condition for vanishing odd degree cohomology of toric orbifolds.

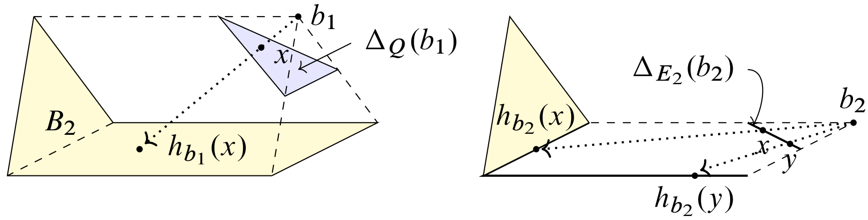


Figure 3: The geometric interpretation of a retraction sequence

In particular, let $X(Q,\lambda)$ be a toric orbifold and the triple $\{(B_k, E_k, b_k)\}_{k=1}^{\ell}$ be a retraction sequence of Q. Given an n–dimensional polytope Q, we begin by defining the map

$$h_{b_1}\colon \Delta_Q(b_1) \to B_2 = \bigcup\{E \mid E \text{ is a face of } Q,\ b_1 \notin V(E)\} \tag{4-1}$$

by $h_{b_1}(x) = B_2 \cap (\text{line passing through } x \text{ and } b_1)$, where $\Delta_Q(b_1)$ is an $(n-1)$–dimensional simplex. The map h_{b_1} is well-defined, because Q is convex. The left picture of Figure 3 shows the map h_{b_1} when Q is a prism.

Define a map

$$f_{b_1}\colon T^n \times \Delta_Q(b_1) \to \bigcup_{E \text{ a face of } B_2} T^{\dim E} \times E \tag{4-2}$$

by $f_{b_1}(t,x) = (\overline{\rho}_E(t), h_{b_1}(x))$, where $\overline{\rho}_E$ is as defined in (3-6). This induces the map

$$\overline{f}_{b_1}\colon L(\Delta_Q(b_1), \xi_{Q,b_1}) \to \bigcup_{E \text{ a face of } B_2} X(E,\lambda_E), \tag{4-3}$$

where ξ_{Q,b_1} is the $\mathcal{L}$–characteristic function defined in (3-10). This map is well-defined by the proof of the following proposition.

Proposition 4.1 *The following diagram commutes:*

$$\begin{tikzcd}
T^n \times \Delta_Q(b_1) \arrow[r, "f_{b_1}"] \arrow[d, "/\sim_{\xi_{Q,b_1}}"] & \bigcup_{E \text{ a face of } B_2} (T^{\dim E} \times E) \arrow[d, "/\sim_{\lambda_E}"] & \\
L(\Delta_Q(b_1), \xi_{Q,b_1}) \arrow[r, "\overline{f}_{b_1}"] & \bigcup_{E \text{ a face of } B_2} X(E,\lambda_E) \arrow[r, hook] & X(Q,\lambda)
\end{tikzcd} \tag{4-4}$$

where the equivalence relations $\sim_{\xi_{Q,b_1}}$ *and* $\sim_{\lambda_E}$ *are defined similarly as in* (3-7) *and* (3-2)*, respectively. Moreover, the bottom row is a cofiber sequence, ie* $X(Q,\lambda)$ *is homotopy equivalent to the mapping cone* $c(\overline{f}_{b_1})$ *of the map* $\overline{f}_{b_1}$*.*

Proof We first show that the map $\overline{f}_{b_1}$ is well-defined. Suppose we choose two different representatives, say $[t,x]_{\sim_{\xi_{Q,b_1}}}$ and $[s,y]_{\sim_{\xi_{Q,b_1}}}$ in $L(\Delta_Q(b_1),\xi_{Q,b_1})$. Then $x=y$, so $h_{b_1}(x)=h_{b_1}(y)$. Moreover, if $x\in\Delta_Q(b_1)\cap F$ for some face F of Q, then $h_{b_1}(x)\in F\cap E$ for some face E of B_2. Hence the map $\overline{\rho}_E$ sends the subtorus $T_{F(x)}$ of T^n to $T_{E(h_{b_1}(x))}$ the subtorus of $T^{\dim E}$. Since the map $\overline{\rho}_E$ is a homomorphism, if $t^{-1}s\in T_{F(x)}$, then

$$\overline{\rho}_E(t)^{-1}\overline{\rho}_E(s)=\overline{\rho}_E(t^{-1}s)\in T_{E(h_{b_1}(x))}.$$

Let $C\Delta_Q(b_1)$ be the cone on $\Delta_Q(b_1)$ in Q with the cone point b_1. Then we can decompose Q into two parts as follows:

$$Q=C\Delta_Q(b_1)\cup_{\Delta_Q(b_1)}\overline{Q\setminus C\Delta_Q(b_1)}. \tag{4-5}$$

Now we define a continuous surjective map

$$g_{b_1}\colon \overline{Q\setminus C\Delta_Q(b_1)}\to B_2$$

in a manner similar to (4-1). We use it to define a straight line homotopy by

$$\phi\colon \overline{Q\setminus C\Delta_Q(b_1)}\times I\to\overline{Q\setminus C\Delta_Q(b_1)},\quad (x,u)\mapsto(1-u)x+u\cdot g_{b_1}(x),$$

which preserves the face structure. Thus, ϕ induces a homotopy

$$\widehat{\phi}\colon (T^n\times\overline{Q\setminus C\Delta_Q(b_1)})/{\sim_\lambda}\times I\to(T^n\times\overline{Q\setminus C\Delta_Q(b_1)})/{\sim_\lambda},$$

defined by

$$([t,x]_{\sim_\lambda},u)\mapsto[t,\phi(x,u)]_{\sim_\lambda}.$$

Note that at $u=0$ the map $\widehat{\phi}$ is the identity and at $u=1$ the image of $\widehat{\phi}$ is $\pi^{-1}(B_2)$.

Then

$$\begin{aligned}X(Q,\lambda)&=\pi^{-1}(C\Delta_Q(b_1))\cup_{L(\Delta_Q(b_1),\xi_{Q,b_1})}\pi^{-1}(\overline{Q\setminus C\Delta_Q(b_1)})\\&\simeq C\big(L(\Delta_Q(b_1),\xi_{Q,b_1})\big)\cup_{L(\Delta_Q(b_1),\xi_{Q,b_1})}\pi^{-1}(B_2)\\&\simeq c(\overline{f}_{b_1}).\end{aligned}$$

Hence, the result follows. □

Now the following isomorphisms are straightforward from the cofiber sequence

$$\begin{aligned}H_*(X(Q,\lambda),\pi^{-1}(B_2))&\cong H_*\big(C\big(L(\Delta_Q(b_1),\xi_{Q,b_1})\big),\pi^{-1}(B_2)\big)\\&\cong\widetilde{H}_{*-1}\big(L(\Delta_Q(b_1),\xi_{Q,b_1})\big).\end{aligned}$$

Those two isomorphisms come from the excision and the long exact sequence of the pair, respectively.

So far, we have considered $B_1 = Q$ and B_2, which is the second term of a retraction sequence starting by choosing $b_1 \in \mathrm{FV}(Q) = V(Q)$. However, we can apply similar arguments to each pair B_i and B_{i+1} in a retraction sequence. This leads us to the following lemma, whose proof is essentially same as that of Proposition 4.1. Before we state the lemma, we first set up the notations: Given a retraction sequence $\{(B_k, E_k, b_k)\}_{k=1}^{\ell}$ of Q:

- $\Delta_{E_k}(b_k) := E_k \cap H(b_k) = B_k \cap H(b_k)$ is the simplex obtained by cutting the vertex b_k from B_k.
- ξ_{E_k,b_k} is an $\mathcal{L}$–characteristic function on $\Delta_{E_k}(b_k)$ defined in a similar manner to (3-10) induced from λ_{E_k}.
- The map
$$\bar{f}_{b_k}\colon L(\Delta_{E_k}(b_k), \xi_{E_k,b_k}) \to \bigcup_{E \text{ a face of } B_{k+1}} X(E, \lambda_E) = \pi^{-1}(B_{k+1})$$
is defined similarly to (4-3) by regarding E_k as a simple polytope.

The right-hand side of Figure 3 illustrates the case of the 3–dimensional prism. The argument above extends to prove the following lemma.

Lemma 4.2 *The sequence*

$$L(\Delta_{E_k}(b_k), \xi_{E_k,b_k}) \xrightarrow{\bar{f}_{b_k}} \pi^{-1}(B_{k+1}) \hookrightarrow \pi^{-1}(B_k) \tag{4-6}$$

is a cofiber sequence. Moreover,

$$H_*(\pi^{-1}(B_k), \pi^{-1}(B_{k+1})) \cong \widetilde{H}_{*-1}\big(L(\Delta_{E_k}(b_k), \xi_{b_k})\big).$$

Recall from Proposition 3.5 that an $\mathcal{L}$–characteristic function

$$\xi\colon \mathscr{F}(\Delta^{n-1}) \to \mathbb{Z}^n$$

defines a finite abelian group $\mathbb{Z}^n/\mathrm{im}(\xi)$. An $\mathcal{R}$–characteristic pair (Q, λ) induces an $\mathcal{R}$–characteristic pair (E, λ_E) as in (3-5) for any face E of Q. Let E be a k–dimensional face of Q for some $k \le n$ and $v \in V(E)$. Then $\Delta_E(v) := E \cap H(v)$ is a $(k-1)$–simplex. These give us an $\mathcal{L}$–characteristic function

$$\xi_{E,v}\colon \mathscr{F}(\Delta_E(v)) \to \mathbb{Z}^k,$$

which is defined in a similar manner to (3-10) associated to $\lambda_E\colon \mathscr{F}(E) \to \mathbb{Z}^k$ and $v \in V(E)$. This $\mathcal{L}$–characteristic function defines the finite group

$$G_E(v) := \mathbb{Z}^k/\mathrm{im}(\xi_{E,v}). \tag{4-7}$$

If $G_E(v)$ is trivial, we call a point $\pi^{-1}(v)$ in $\pi^{-1}(E) \cong X(E, \lambda_E)$ a *smooth* point, and otherwise a *singular* point, where $\pi\colon X(Q, \lambda) \to Q$ is the orbit map defined in (3-3).

Furthermore, for each $B \in \mathfrak{B}(Q)$ and a free vertex $v \in \mathrm{FV}(B)$, there exists a unique maximal face, say E_v, of B containing v. Hence, for each $B \in \mathfrak{B}(Q)$, we write

$$G_B(v) := G_{E_v}(v) \tag{4-8}$$

whenever v is a free vertex in B.

Proposition 4.3 *Given a vertex $v \in V(Q)$, let E and E' be two faces containing v such that E is a face of E'. Then $|G_E(v)|$ divides $|G_{E'}(v)|$.*

Proof From Proposition 3.2, we may assume that $E' = Q$ without loss of generality. Suppose that E is a face of Q with codimension k. For convenience, we further assume that $E = F_1 \cap \cdots \cap F_k$ and $v = F_1 \cap \cdots \cap F_k \cap F_{k+1} \cap \cdots \cap F_n$, where the F_i are facets of Q.

From (3-10) and (4-7), we have $G_Q(v) = \mathbb{Z}^n / \langle \lambda(F_1), \dots, \lambda(F_n) \rangle$ and $G_E(v) = \mathbb{Z}^k / \langle \lambda_E(E \cap F_{k+1}), \dots, \lambda_E(E \cap F_n) \rangle$. Now we consider the composition

$$\mathbb{Z}^n \xrightarrow{\rho_E} \mathbb{Z}^k \twoheadrightarrow \mathbb{Z}^k / \langle \lambda_E(E \cap F_{k+1}), \dots, \lambda_E(E \cap F_n) \rangle,$$

where the map ρ_E is defined in (3-4) and the second map is the natural surjection determined by (3-5). Observe that the kernel of the previous composition contains $\langle \lambda(F_1), \dots, \lambda(F_n) \rangle$. Hence, we get a surjective group homomorphism from $G_Q(v)$ to $G_E(v)$. The result follows from Lagrange's theorem in group theory. □

We are now in a position to prove Theorem 1.1.

Proof of Theorem 1.1 We prove the claim by the induction on the number of vertices of $B \in \mathfrak{B}(Q)$. First, notice that when the retraction sequence reaches an edge or a union of edges, say B_s, then $\pi^{-1}(B_s)$ is $\mathbb{CP}^1$ or homotopic to a finite wedge of $\mathbb{CP}^1$, which implies that $H_*(\pi^{-1}(B_s))$ is torsion-free and concentrated in even degrees. Therefore, if $|V(B)| \leq 2$ for $B \in \mathfrak{B}(Q)$, then the claim is true.

Now we assume that $\pi^{-1}(B)$ is even for $B \in \mathfrak{B}(Q)$ with $|V(B)| \leq i-1$. To complete the induction, we shall prove that the same holds for $B' \in \mathfrak{B}(Q)$ with $|V(B')| = i$. Given such B', there exists $B \in \mathfrak{B}(Q)$ such that B is obtained from B' by deleting all faces containing a free vertex of B'. To be more precise, let $\mathrm{FV}(B') = \{v_{i_1}, \dots, v_{i_r}\}$ be the set of free vertices in B'. Notice that, regarding B' as a generic step of a retraction sequence in $\mathfrak{R}(Q)$, we can produce r many different $B \in \mathfrak{B}(Q)$ with $|V(B)| = i-1$ from B'. According to the induction hypothesis, we assume that for each $t = 1, \dots, r$, the group $H_*(\pi^{-1}(B(v_{i_t})))$ is concentrated in even degrees and torsion-free, where $B(v_{i_t}) \in \mathfrak{B}(Q)$ is obtained from B' by deleting faces containing v_{i_t}. This assumption makes sense, because any retraction sequence reaches a union of edges.

For simplicity, we fix the following notation: For each free vertex $v_{i_t} \in \mathrm{FV}(B')$:

- $X' := \pi^{-1}(B')$ and $\dim B' = d' = \frac{1}{2}\dim_{\mathbb{R}} X'$.
- $X(v_{i_t}) := \pi^{-1}(B(v_{i_t}))$ and $\dim B(v_{i_t}) = d = \frac{1}{2}\dim_{\mathbb{R}} X(v_{i_t})$.
- $L(v_{i_t}) := L(\Delta_{E_{i_t}}(v_{i_t}), \xi_{E_{i_t}, v_{i_t}})$, where E_{i_t} denotes the maximal face of B' containing v_{i_t}.

Notice that $\dim L(v_{i_t}) \leq 2d' - 1$ and $d \leq d'$.

Now we consider the long exact sequence of the homology for the pair $(X', X(v_{i_t})) = \big(\pi^{-1}(B'), \pi^{-1}(B(v_{i_t}))\big)$

$$\begin{aligned}\text{(4-9)} \quad \cdots \to H_{j+1}(X') \to H_{j+1}(X', X(v_{i_t})) \to H_j(X(v_{i_t})) \\ \to H_j(X') \to H_j(X', X(v_{i_t})) \to \cdots .\end{aligned}$$

Suppose that j is odd. By the induction hypothesis and Lemma 4.2, the sequence (4-9) becomes

$$0 \to H_j(X') \to \widetilde{H}_{j-1}(L(v_{i_t})) \xrightarrow{0} H_{j-1}(X(v_{i_t})). \tag{4-10}$$

The rightmost map is the zero map because the domain is a torsion group but the target space is free by assumption. Hence, $H_j(X')$ is isomorphic to $\widetilde{H}_{j-1}(L(v_{i_t}))$, and the latter is zero if $j-1 > \dim L(v_{i_t})$ or a torsion group determined by the prime factors of $|G_{B'}(v_{i_t})|$ if $j-1 \leq \dim L(v_{i_t})$ by Lemma 3.6. This argument holds for each free vertex $v_{i_1}, \dots, v_{i_r}$. Hence we have r many different exact sequences like (4-10). Now the assumption of Theorem 1.1 tells us that

$$\gcd\big\{|\widetilde{H}_{j-1}(L(v_{i_1}))|, \dots, |\widetilde{H}_{j-1}(L(v_{i_r}))|\big\} = 1,$$

but $H_j(X')$ stays same. Hence, we conclude that $H_j(X') = 0$ if j is odd. Moreover, $\widetilde{H}_{j-1}(L(v_{i_t})) = 0$ for all $t = 1, \dots, r$ because of the exactness of (4-10).

Next we assume that j is even. Then the exact sequence (4-9) gives us

$$\widetilde{H}_j(L(v_{i_t})) \xrightarrow{0} H_j(X(v_{i_t})) \to H_j(X') \to \widetilde{H}_{j-1}(L(v_{i_t})) \to 0. \tag{4-11}$$

Then we have the following three cases:

$$\begin{array}{ll} \cdots \xrightarrow{0} H_j(X(v_{i_t})) \to H_j(X') \to 0 & \text{if } j-1 > \dim L(v_{i_t}), \\ \cdots \xrightarrow{0} H_j(X(v_{i_t})) \to H_j(X') \to \mathbb{Z} \to 0 & \text{if } j-1 = \dim L(v_{i_t}), \\ \cdots \xrightarrow{0} H_j(X(v_{i_t})) \to H_j(X') \to G_{j-1} \to 0 & \text{if } j-1 < \dim L(v_{i_t}), \end{array}$$

where G_{j-1} is as defined in Lemma 3.6 and $H_j(X(v_{i_t}))$ is free by the induction hypothesis. The free vertices $v_{i_1}, \dots, v_{i_r}$ in B' gives us r many exact sequences, and each of them is one of the above three cases. If one of the free vertices gives the first or the second type of exact sequence, then $H_j(X')$ cannot have a torsion subgroup because of the exactness. If all of the sequences are of the third type, then $H_j(X')$ has no torsion because of the assumption of the theorem and arguments similar to those used in the case when j is odd. This completes the induction. □

Notice that Kawasaki [18] has shown that the cohomology ring of weighted projective space $\mathbb{CP}^n_\chi$ with weight $\chi = (\chi_0, \dots, \chi_n)$ is concentrated in even degrees and torsion-free, if $\gcd(\chi_0, \dots, \chi_n) = 1$. Theorem 1.1 extends Kawasaki's theorem to the category of toric orbifolds which contains the weighted projective spaces. The following Example 4.4 shows how we can apply this result to a polygon, and Example 4.5 is a practical computation on a higher-dimensional weighed projective space.

Example 4.4 Consider the 4–dimensional toric orbifold X over Q whose $\mathcal{R}$–characteristic pair is described in Figure 4. Let $H(v)$ be an affine hyperplane as defined in (3-9). Then $H(v) \cap Q$ is an 1–simplex. The induced $\mathcal{L}$–characteristic function

$$\xi_{Q,v}\colon \{H(v) \cap F_1,\ H(v) \cap F_m\} \to \mathbb{Z}^2$$

is defined by $\xi_{Q,v}(H(v) \cap F_1) = \lambda(F_1) = (a_1, b_1)$ and $\xi_{Q,v}(H(v) \cap F_m) = \lambda(F_m) = (a_m, b_m)$. Therefore, the orbifold lens space $L(\Delta_Q(v), \xi_{Q,v})$ is homeomorphic to $S^3/G_Q(v)$, where $G_Q(v)$ is a finite abelian group of order $|a_1 b_m - b_1 a_m|$; see Proposition 3.5. Moreover, the prime factors of the order of a torsion element in $H_*(L(\Delta_Q(v), \xi_{Q,v}))$ is a subset of the prime factors of $|a_1 b_m - b_1 a_m|$ by Lemma 3.6.

Now we consider a retraction sequence $\{B_k, E_k, b_k\}_{k=1}^{\ell}$ starting at v. The second space B_2 is the union $F_2 \cup \dots \cup F_{m-1}$ of edges whose preimage $\pi^{-1}(B_2)$ is homotopic to the wedge of $m-2$ copies of $\mathbb{CP}^1$. Hence, $H_*(\pi^{-1}(B_2))$ is torsion-free and $H_{\mathrm{odd}}(\pi^{-1}(B_2))$ vanishes. A cofibration

$$L(\Delta_Q(v), \xi_{Q,v}) \to \pi^{-1}(B_2) \to X$$

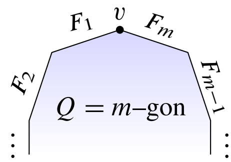
gives an isomorphism $H_j(X, \pi^{-1}(B_2)) \cong \widetilde{H}_{j-1}(L(\Delta_Q(v), \xi_{Q,v}))$. Hence, the long exact sequence of pair $(X, \pi^{-1}(B_2))$ yields

$$\begin{aligned} \cdots \to \widetilde{H}_j(L(\Delta_Q(v), \xi_{Q,v})) \to H_j(\pi^{-1}(B_2)) \\ \to H_j(X) \to \widetilde{H}_{j-1}(L(\Delta_Q(v), \xi_{Q,v})) \to \cdots, \end{aligned}$$

and this shows that, if $H_j(X)$ has a torsion part, then its prime factors must divide $|a_1 b_m - b_1 a_m|$. But the same argument can be applied to all the other vertices in Q. Finally, we may conclude that $H_*(X)$ is torsion-free and concentrated in even degrees if

$$\gcd\bigl\{|a_1 b_2 - b_1 a_2|, \dots, |a_{m-1} b_m - b_{m-1} a_m|, |a_1 b_m - b_1 a_m|\bigr\} = 1, \tag{4-12}$$

which is the assumption of Theorem 1.1.



$$\Lambda = \begin{array}{cccc} \lambda_1 & \lambda_2 & \cdots & \lambda_4 \\ \end{array}$$
$$\Lambda = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \\ b_1 & b_2 & \cdots & b_m \end{bmatrix}$$

Figure 4: An $\mathcal{R}$–characteristic function on a polygon

Example 4.5 We consider an $\mathcal{R}$–characteristic pair (Δ^4, λ), where $\lambda\colon \mathscr{F}(\Delta^4) \to \mathbb{Z}^4$ is defined by

$$\begin{array}{ccccc} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \end{array}$$
$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The column vectors satisfies the relation $\lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4 + 2\lambda_5 = \mathbf{0}$. Then the resulting toric orbifold is a weighted projective space $\mathbb{CP}^4_{(1,1,2,2,2)}$. We refer to [6, Example 3.1.17] or [12, Section 2.2] for more details.

To check the assumption in Theorem 1.1, it suffices to consider all faces of Δ^4 of dimension greater than 1, because the set $\mathfrak{B}(\Delta^4)$ coincides with the set of all faces of Δ^4. First of all, for Δ^4 itself, it is easy to see that

$$\gcd\bigl\{|G_{\Delta^4}(v)| : v \in V(\Delta^4)\bigr\} = \gcd\{2, 2, 2, 1, 1\} = 1.$$

Since the process is essentially the same, we choose $E = F_1 \cap F_2 = \Delta^2$ as a sample. Observe that

$$(\langle \lambda_1, \lambda_2 \rangle \otimes_{\mathbb{Z}} \mathbb{R}) \cap \mathbb{Z}^4 = (\langle -e_1 - 2e_2 - 2e_3 - 2e_4, e_1 \rangle \otimes_{\mathbb{Z}} \mathbb{R}) \cap \mathbb{Z}^4 \cong \langle e_2 + e_3 + e_4, e_1 \rangle.$$

Hence, we may decompose the target space $\mathbb{Z}^4 \cong \langle e_2 + e_3 + e_4 \rangle \oplus \langle e_1 \rangle \oplus \langle e_3 \rangle \oplus \langle e_4 \rangle$. This derives an $\mathcal{R}$–characteristic function

$$\lambda_E\colon \{E \cap F_3,\ E \cap F_4,\ E \cap F_5\} \to \mathbb{Z}^2 \cong \langle e_3 \rangle \oplus \langle e_4 \rangle,$$

defined by $\lambda_E(E \cap F_3) = (-1, -1)$, $\lambda_E(E \cap F_4) = (1, 0)$ and $\lambda_E(E \cap F_5) = (0, 1)$. Hence, $\pi^{-1}(E) = X(\Delta^2, \lambda_E) \cong \mathbb{CP}^2_{(1,1,1)}$. Hence, we have

$$\gcd\bigl\{|G_E(v)| : v \in V(E)\bigr\} = \gcd\{1, 1, 1\} = 1.$$

Sometimes, if the polytope has sufficiently many symmetries, we can analyze all possible retraction sequences efficiently. Proposition 4.3 can then be used to ensure the

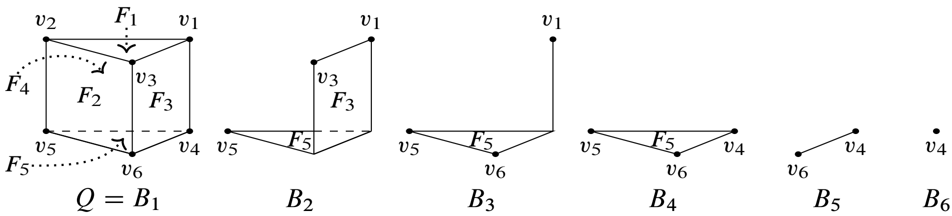


Figure 5: A retraction sequence of a prism

gcd assumption of Theorem 1.1 holds. The main features of the following example are that the polytope has at least two free vertices at each $B \in \mathfrak{B}(Q)$, and that the collection $\{|G_Q(v)| : v \in V(Q)\}$ consists of mutually different prime numbers; in particular, they are pairwise relatively prime.

Example 4.6 Let Q be the 3–dimensional cube whose facets and vertices are illustrated in Figure 5. We assign an $\mathcal{R}$–characteristic function $\lambda\colon \mathscr{F}(Q) \to \mathbb{Z}^3$ as follows:

$$\lambda(F_1) = (p_1, p_2, p_3), \quad \lambda(F_5) = (p_4, p_5, p_6),$$
$$\lambda(F_2) = \boldsymbol{e}_1, \quad \lambda(F_3) = \boldsymbol{e}_2, \quad \lambda(F_4) = \boldsymbol{e}_3,$$

where the p_i are all prime numbers with $p_i \neq p_j$ whenever $i \neq j$, and $\boldsymbol{e}_i$ is the i^{th} standard unit vector in $\mathbb{Z}^3$. Then it is easy to see that $|G_Q(v_i)| = p_i$ for $i = 1, \dots, 6$. Hence, we have

$$\gcd\bigl\{|G_Q(v)| : v \in V(Q)\bigr\} = \gcd\{p_1, \dots, p_6\} = 1.$$

The same property holds for other polytopal complex $B \in \mathfrak{B}(Q)$ from Proposition 4.3. Indeed, for instance,

$$\gcd\bigl\{|G_{B_2}(v)| : v \in \mathrm{FV}(B_2)\bigr\} = \gcd\bigl\{|G_{B_2}(v_1)|, |G_{B_2}(v_3)|, |G_{B_2}(v_5)|\bigr\} = 1$$

because $\gcd\{p_1, p_3, p_5\} = 1$.

5 Cohomology ring of toric orbifolds

The integral equivariant cohomology ring of certain projective toric varieties is given by a ring determined by the fan data. This ring is called the ring of *piecewise polynomials*, which we denote by $\mathcal{PP}[\Sigma]$. For a smooth fan, it uses the fan's combinatorial data only and coincides with the *Stanley–Reisner ring* $\mathcal{SR}[\Sigma]$ of the fan Σ. In general, however, the ring of piecewise polynomials uses all the geometric data in a fan.

To be more precise, let Σ be a fan in $\mathbb{R}^n$ and $\{\lambda_1, \dots, \lambda_m\} \subset \mathbb{Z}^n$ the set of primitive vectors generating 1–dimensional rays in Σ. Then the Stanley–Reisner ring $\mathcal{SR}[\Sigma]$ is defined by the quotient $\mathbb{Z}[x_1, \dots, x_m]/\mathcal{I}$ of polynomial ring with m variables by the following ideal generated by squarefree monomials:

$$\mathcal{I} = \langle x_{i_1} \cdots x_{i_k} \mid \mathrm{cone}\{\lambda_{i_1}, \dots, \lambda_{i_k}\} \notin \Sigma \rangle, \tag{5-1}$$

where $\mathrm{cone}\{\lambda_{i_1}, \dots, \lambda_{i_k}\}$ denotes the cone generated by $\{\lambda_{i_1}, \dots, \lambda_{i_k}\}$. For the case of smooth toric varieties, their odd-degree cohomology always vanishes, which leads us to the following description of the cohomology ring.

Theorem 5.1 [7; 16] *Let X_Σ be a smooth toric variety. Then there exists a ring isomorphism $H^*(X_\Sigma) \cong \mathcal{SR}[\Sigma]/\mathcal{J}$, where $\mathcal{J}$ is the ideal generated by the linear relations*

$$\sum_{i=1}^{m} \langle \lambda_i, \boldsymbol{e}_j \rangle x_i = 0, \quad j = 1, \dots, n, \tag{5-2}$$

where $\boldsymbol{e}_j$ denotes the j^{th} standard unit vector in $\mathbb{Z}^n$.

Notice that, for toric orbifolds, the theorem holds only for $\mathbb{Q}$–coefficients; see for instance [6, Section 12.4]. In order to make the singular theory better resemble the smooth case, we introduce an intermediate ring, which models the Stanley–Reisner ring but is based on a fan $\widehat{\Sigma}$ in $\mathbb{R}^m$ defined from the combinatorial data of Σ, which has m one-dimensional rays. The ring of piecewise polynomials on the original fan Σ is recovered by imposing an *integrality condition*, which leads us to the notion of the *weighted Stanley–Reisner ring* $w\mathcal{SR}[\Sigma]$ of Σ.

5.1 Weighted Stanley–Reisner ring

Let Σ be a simplicial fan in $\mathbb{R}^n$, ie each top-dimensional cone of Σ is generated by n linearly independent primitive vectors in the lattice $\mathbb{Z}^n$. In particular, a simplicial fan Σ is called a *polytopal fan* if it is the normal fan of a simple lattice polytope in $\mathbb{R}^n$; see [6, Chapter 2] or [12, Section 1.5] for more details. Hence, the determinant of generators of each top-dimensional cone is nonzero but not necessarily ± 1, so the corresponding fixed point might be singular. Let $\Sigma^{(j)}$ denotes the set of j–dimensional cones in Σ. To record the singularity of each fixed point in an efficient way, we assign a vector

$$z^\sigma := (z_1^\sigma, \dots, z_m^\sigma) \in \bigoplus_m \mathbb{Q}[u_1, \dots, u_n]$$

to each top-dimensional cone $\sigma = \mathrm{cone}\{\lambda_{i_1}, \dots, \lambda_{i_n}\} \in \Sigma^{(n)}$ by the following rule:

(C1) $z_j^\sigma = 0$ if $j \notin \{i_1, \dots, i_n\}$.

(C2) $\begin{bmatrix} z_{i_1}^\sigma \\ \vdots \\ z_{i_n}^\sigma \end{bmatrix} = \left[\begin{array}{c|c|c} \lambda_{i_1} & \cdots & \lambda_{i_n} \end{array} \right]^{-1} \cdot \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$.

The inverse matrix in the condition (C2) may have rational entries. The following definition is motivated by this observation.

Definition 5.2 Given a fan Σ in $\mathbb{R}^n$ with m one-dimensional rays, we say a polynomial $h(x_1, \dots, x_m) \in \mathbb{Z}[x_1, \dots, x_m]$ satisfies the *integrality condition* with respect to Σ if $h(z^\sigma) \in \mathbb{Z}[u_1, \dots, u_n]$ for all $\sigma \in \Sigma^{(n)}$.

Notice that the collection of polynomials satisfying the integrality condition is closed under addition and multiplication, which induces the natural ring structure on it inherited from that of $\mathbb{Z}[x_1, \dots, x_m]$. Moreover, the polynomials in $\mathcal{I}$ defined in (5-1) satisfy the integrality condition, obviously. Indeed, the condition (C1) leads $h(z^\sigma)$ to be the zero polynomial for all $\sigma \in \Sigma^{(n)}$ whenever $h(x_1, \dots, x_m) \in \mathcal{I}$.

Finally, we define the weighted Stanley–Reisner ring $w\mathcal{SR}[\Sigma]$ as follows:

$$w\mathcal{SR}[\Sigma] := \{h \in \mathbb{Z}[x_1, \dots, x_m] \mid h \text{ satisfies the } \textit{integrality condition}\}/\mathcal{I}. \tag{5-3}$$

Remark When the fan Σ is smooth, $w\mathcal{SR}[\Sigma] = \mathcal{SR}[\Sigma]$. Indeed, the determinant of a smooth top-dimensional cone is ± 1, which implies that its inverse has integer entries.

Now we introduce the second main theorem of this paper. The proof will be given in the next subsection.

Theorem 5.3 *Let X_Σ be a projective toric orbifold over a polytopal fan Σ with $H^{\mathrm{odd}}(X) = 0$. Then there is a ring isomorphism*

$$H^*(X_\Sigma) \cong w\mathcal{SR}[\Sigma]/\mathcal{J},$$

where $\mathcal{J}$ is the ideal generated by the linear relations (5-2).

Consider a simple lattice polytope Q in $\mathbb{R}^n$ whose normal fan is Σ. Then the normal vectors of each facet define an $\mathcal{R}$–characteristic function $\lambda\colon \mathscr{F}(Q) \to \mathbb{Z}^n$. Now we have a natural $\mathcal{R}$–characteristic pair (Q, λ) from Σ, which allows us to apply the results of Sections 2 and 4. Hence, we have a concrete statement, which is Theorem 1.2, with a sufficient condition for $H^{\mathrm{odd}}(X_\Sigma) = 0$.

We complete this subsection by applying Theorem 1.2 to a weighted projective space $\mathbb{CP}^2_{(1,a,b)}$. We shall recover Kawasaki's result [18, Theorem 1].

Example 5.4 Let Σ be a fan in $\mathbb{R}^2$ generated by

$$\lambda_1 = (a,b), \quad \lambda_2 = (-1,0), \quad \lambda_3 = (0,-1) \in \mathbb{Z}^2, \tag{5-4}$$

where a and b are relatively prime. The 2–dimensional cones are σ_{12}, σ_{13} and σ_{23}, where $\sigma_{ij} = \operatorname{cone}\{\lambda_i, \lambda_j\}$. Since $\{\lambda_1, \lambda_2, \lambda_3\}$ generates the lattice $\mathbb{Z}^2$ and satisfies $\lambda_1 + a\lambda_2 + b\lambda_3 = (0,0)$, the toric variety X_Σ is isomorphic to the weighed projective space $\mathbb{CP}^2_{(1,a,b)}$. We refer to [6, Example 3.1.17] or [12, Section 2.2] for the characterization of a fan corresponding to weighted projective spaces.

The direct computation of inverse matrices for $[\lambda_i \mid \lambda_j]$ gives us the following list of vectors:

$$\begin{aligned} z^{\sigma_{12}} &= \Big(\frac{1}{b}u_2,\, -u_1 + \frac{a}{b}u_2,\, 0\Big), \\ z^{\sigma_{13}} &= \Big(\frac{1}{a}u_1,\, 0,\, \frac{b}{a}u_1 - u_2\Big), \\ z^{\sigma_{23}} &= (0,\, -u_1,\, -u_2). \end{aligned}$$

Hence, we have

$$\begin{aligned} &w\mathcal{SR}[\Sigma] = \\ &\qquad \big\{h(x_1,x_2,x_3) \in \mathbb{Z}[x_1,x_2,x_3] \mid h(z^{\sigma_{ij}}) \in \mathbb{Z}[u_1,u_2] \text{ for } 1 \le i < j \le 3\big\}/\mathcal{I}. \end{aligned} \tag{5-5}$$

Finding elements at each degree is straightforward. For instance, for a degree-2 polynomial, $k_1x_1 + k_2x_2 + k_3x_3 \in w\mathcal{SR}[\Sigma]$ if and only if the following three polynomials have integer coefficients:

$$-k_2u_1 + \Big(\frac{1}{b}k_1 + \frac{a}{b}k_2\Big)u_2, \quad \Big(\frac{1}{a}k_1 + \frac{b}{a}k_3\Big)u_1 - k_3u_2, \quad -k_2u_1 - k_2u_2,$$

which is exactly the case when $k_1 + ak_2 \in b\mathbb{Z}$ and $k_1 + bk_3 \in a\mathbb{Z}$. Hence, one can show that the integers (k_1,k_2,k_3) are

$$(a,-1,0), \quad (b,0,-1), \quad (ab,0,0), \quad (0,b,0), \quad (0,0,a),$$

and $\mathbb{Z}$–linear combinations of them. They give us the following degree-2 elements in $w\mathcal{SR}[\Sigma]$:

$$ax_1 - x_2, \quad bx_1 - x_3, \quad abx_1, \quad bx_2, \quad ax_3, \tag{5-6}$$

and $\mathbb{Z}$–linear combinations of them. Similarly, we can find the degree-4 elements:

$$a^2b^2x_1^2, \quad b^2x_2^2, \quad a^2x_3^2, \quad abx_1x_2, \quad a^2x_1x_3, \quad x_2x_3, \tag{5-7}$$

and $\mathbb{Z}$–linear combinations of them.

We continue to calculate the ring structure of $H^*(\mathbb{CP}^2_{(1,a,b)})$ using Theorem 1.2. Indeed, the $\mathcal{R}$–characteristic pair (Δ^2, λ) induced from Σ satisfies the assumption of Theorem 1.1; see Example 4.4. Hence, we conclude that $\mathbb{CP}^2_{(1,a,b)}$ is *even*, which implies that the rank of the integral cohomology group is 1 in each even degree and 0 otherwise.

Remark In general, the integral Betti numbers of a toric manifold or the rational Betti numbers of a toric orbifold are given by the h–vector of its underlying polytope; see [9, Section 3] or [20, Section 4]. Hence, if a toric orbifold is even, then its integral Betti numbers are obtained by the h–vector of the underlying polytope.

Now the characteristic vectors (5-4) and the relation (5-2) determine the ideal $\mathcal{J} = \langle ax_1 - x_2, bx_1 - x_3\rangle$ whose generators are first two items in (5-6). Hence, the elements in (5-6) except the first two are are the same modulo $\mathcal{J}$. Hence, they represent the same element in $H^*(\mathbb{CP}^2_{(1,a,b)})$. We put

$$w_1 := abx_1 = bx_2 = ax_3.$$

Since $\operatorname{rank} H^4(\mathbb{CP}^2_{(1,a,b)}) = 1$, we choose an element in (5-7) which has the minimal divisibility. In this case, we pick

$$w_2 := x_2x_3.$$

Then we have the multiplicative structure $w_1^2 = abw_2$. Finally, we have the presentation

$$H^*(\mathbb{CP}^2_{(1,a,b)}) \cong \mathbb{Z}[w_1, w_2]/\langle w_1^2 - abw_2, w_1w_2\rangle,$$

where $\deg w_1 = 2$ and $\deg w_2 = 4$. Notice that the monomial w_1w_2 comes from the Stanley–Reisner ideal $x_1x_2x_3$.

Remark Even if we can find elements in $w\mathcal{SR}[\Sigma]$ by the direct computation of the integrality condition, finding the minimal set of generators in $w\mathcal{SR}[\Sigma]$ for an arbitrary simplicial fan is not obvious, in general. However, when X_Σ is a weighted projective space, a result of [1] allows us to find generators of the ring of piecewise polynomials $\mathcal{PP}[\Sigma]$ and, hence, generators in $w\mathcal{SR}[\Sigma]$, by a method in the next subsection. Moreover, the identification result, Corollary 5.8, tells us how to interpret those generators in terms of elements in $w\mathcal{SR}[\Sigma]$.

5.2 Piecewise algebra and cohomology ring

We introduce now the ring of piecewise polynomials, which is determined by a fan and describes the equivariant cohomology of a large class of toric orbifolds. As mentioned above, unlike the Stanley–Reisner ring, which encodes combinatorial data only, the ring of piecewise polynomials depends on the full geometric information in a fan.

We begin by introducing piecewise polynomials. Let Σ be a fan in $\mathbb{R}^n$. A function $f\colon \mathbb{Z}^n \to \mathbb{Z}$ is called a *piecewise polynomial* on Σ if, for each cone $\sigma \in \Sigma$, the restriction $f|_\sigma$ is a polynomial function on $\sigma \cap \mathbb{Z}^n$. Such a function can be interpreted as a collection $\{f_\sigma\}_{\sigma \in \Sigma^{(n)}}$, which we denote by $\{f_\sigma\}$ for simplicity, such that

$$f_\sigma|_{\sigma\cap\sigma'} = f_{\sigma'}|_{\sigma\cap\sigma'}. \tag{5-8}$$

In other words, it is enough to consider the polynomials on each top-dimensional cone. The polynomials on lower-dimensional cones are determined by (5-8).

The set $\mathcal{PP}[\Sigma]$ of piecewise polynomial functions on Σ with integer coefficients on Σ has a ring structure under pointwise addition and multiplication. Moreover, the natural inclusion of global polynomials $\mathbb{Z}[u_1, \dots, u_n]$ into $\mathcal{PP}[\Sigma]$ induces a $\mathbb{Z}[u_1, \dots, u_n]$–algebra structure on $\mathcal{PP}[\Sigma]$. Furthermore, by considering $\mathbb{Q}^n$ instead of $\mathbb{Z}^n$, we can define piecewise polynomial functions with rational coefficients $f\colon \mathbb{Q}^n \to \mathbb{Q}$, and we denote the ring of piecewise polynomial functions with rational coefficients by $\mathcal{PP}[\Sigma;\mathbb{Q}]$.

It is well known that the equivariant cohomology ring with rational coefficients of a toric variety over a simplicial fan is isomorphic to $\mathcal{PP}[\Sigma;\mathbb{Q}]$; see [6]. On the other hand, for the case of polytopal fans, Bahri, Franz and Ray [1] proved the following proposition over $\mathbb{Z}$.

Proposition 5.5 [1, Proposition 2.2] *Let Σ be a polytopal fan in $\mathbb{R}^n$, X_Σ the associated compact projective toric variety with $H^{\mathrm{odd}}(X_\Sigma) = 0$, and $T = T^n$ the n–dimensional torus acting on X_Σ. Then $H^*_T(X_\Sigma)$ is isomorphic to $\mathcal{PP}[\Sigma]$ as an $H^*(BT)$–algebra.*

Here $H^*(BT)$–algebra structure on $\mathcal{PP}[\Sigma]$ is obtained by identifying $H^*(BT)$ with the global polynomials $\mathbb{Z}[u_1, \dots, u_n]$, where u_i is the first Chern class of the canonical line bundle given by the i^{th} projection $T \to S^1$.

On the other hand, the combinatorial structure of Σ determines a canonical fan in a higher-dimensional lattice as follows: Let $\Sigma^{(1)} = \{\lambda_1, \dots, \lambda_m\}$ be the set of primitive vectors generating 1–dimensional rays in Σ. We define a linear map $\Lambda\colon \mathbb{Z}^m \to \mathbb{Z}^n$ by $\Lambda(\boldsymbol{e}_i) = \lambda_i$, where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ denote the standard unit vectors in $\mathbb{Z}^m$. By the pull-back of Σ through Λ, we can define a fan

$$\widehat{\Sigma} = \{\hat{\sigma} := \Lambda^{-1}(\sigma) \mid \sigma \in \Sigma\}$$

in $\mathbb{R}^m$. To be more precise, if σ is the cone generated by $\lambda_{i_1}, \dots, \lambda_{i_k}$, then $\hat{\sigma}$ is the cone generated by $\boldsymbol{e}_{i_1}, \dots, \boldsymbol{e}_{i_k}$. Moreover, for a commutative ring $\boldsymbol{k}$, a linear map Λ induces a ring homomorphism

$$\Lambda^*\colon \mathcal{PP}[\Sigma;\boldsymbol{k}] \to \mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}] \tag{5-9}$$

of piecewise polynomial rings, where the map is defined by

$$\Lambda^*(\{f_\sigma\}) = \big\{g_{\widehat{\sigma}}(x_{i_1},\dots,x_{i_n}) := f_\sigma(\Lambda_\sigma\cdot[x_{i_1},\dots,x_{i_n}]^T)\big\}_{\widehat{\sigma}\in\widehat{\Sigma}^{(n)}},$$

where $\Lambda_\sigma = [\lambda_{i_1} \mid \cdots \mid \lambda_{i_n}]$ is a square matrix and $\Lambda_\sigma \cdot [x_{i_1},\dots,x_{i_n}]^T$ is the usual matrix multiplication of $n\times n$ and $n\times 1$ matrices.

Indeed, the map Λ^* is well-defined, since $g_{\widehat{\sigma}}|_{\widehat{\sigma}\cap\widehat{\sigma}'} = g_{\widehat{\sigma}'}|_{\widehat{\sigma}\cap\widehat{\sigma}'}$.

Lemma 5.6 *Given a polytopal fan Σ, as $H^*(BT;\boldsymbol{k})$–algebras:*

(1) *When $\boldsymbol{k}=\mathbb{Q}$, $\mathcal{PP}[\Sigma;\mathbb{Q}]$ is isomorphic to $\mathcal{PP}[\widehat{\Sigma};\mathbb{Q}]$.*

(2) *When $\boldsymbol{k}=\mathbb{Z}$, there is a monomorphism from $\mathcal{PP}[\Sigma]$ to $\mathcal{PP}[\widehat{\Sigma}]$.*

Proof For each top-dimensional cone $\sigma = \mathrm{cone}\{\lambda_{i_1},\dots,\lambda_{i_n}\}\in\Sigma^{(n)}$, we set the following notation:

- $f_\sigma(u_1,\dots,u_n)$, $g_{\widehat{\sigma}}(x_{i_1},\dots,x_{i_n})$ are polynomial functions defined on $\sigma\in\Sigma$ and $\widehat{\sigma}\in\widehat{\Sigma}$, respectively.
- $\{f_\sigma\} := \{f_\sigma(u_1,\dots,u_n) \mid \sigma\in\Sigma^{(n)}\}\in\mathcal{PP}[\Sigma]$.
- $\{g_{\widehat{\sigma}}\} := \{g_{\widehat{\sigma}}(x_{i_1},\dots,x_{i_n}) \mid \widehat{\sigma}\in\widehat{\Sigma}^{(n)}\}\in\mathcal{PP}[\widehat{\Sigma}]$.
- $\Lambda_\sigma := [\lambda_{i_1}\mid\cdots\mid\lambda_{i_n}]$ is an $n\times n$ matrix with column vectors $\lambda_{i_1},\dots,\lambda_{i_n}$.

Recall the ring homomorphism Λ^* introduced in (5-9). If we restrict $\boldsymbol{k}$ to $\mathbb{Q}$, the map Λ^* has the natural inverse

$$\Theta\colon \mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]\to\mathcal{PP}[\Sigma;\boldsymbol{k}], \tag{5-10}$$

defined by

$$\Theta(\{g_{\widehat{\sigma}}\}) = \big\{f_\sigma(u_1,\dots,u_n) := g_{\widehat{\sigma}}(\Lambda_\sigma^{-1}\cdot[u_1,\dots,u_n]^T) \mid \sigma\in\Sigma^{(n)}\big\},$$

where Λ_σ^{-1} is regarded as a linear automorphism of $\mathbb{Q}^n$. Indeed,

$$\begin{aligned}(\Theta\circ\Lambda^*)(\{f_\sigma\}) &= \big\{f_\sigma(\Lambda_\sigma\cdot\Lambda_\sigma^{-1}\cdot[u_1,\dots,u_n]^T) \mid \sigma\in\Sigma^{(n)}\big\}\\ &= \{f_\sigma(u_1,\dots,u_n) \mid \sigma\in\Sigma^{(n)}\} = \{f_\sigma\}.\end{aligned}$$

In particular, Λ^* is a monomorphism in $\mathbb{Z}$–coefficients. Finally, the $H^*(BT;\boldsymbol{k})$–algebra structure on $\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$ is naturally inherited from that of $\mathcal{PP}[\Sigma;\boldsymbol{k}]$ via the map Λ^*. □

Recall that the Stanley–Reisner ring $\mathcal{SR}[\Sigma;\boldsymbol{k}]$ has combinatorial data only, while $\mathcal{PP}[\Sigma;\boldsymbol{k}]$ contains both combinatorial and geometric data. However, $\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$ has only combinatorics, but looks like $\mathcal{PP}[\Sigma;\boldsymbol{k}]$. In this point of view, $\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$ is an intermediate object between $\mathcal{SR}[\Sigma;\boldsymbol{k}]$ and $\mathcal{PP}[\Sigma;\boldsymbol{k}]$. The following lemma, together with Lemma 5.6, concludes the relations among those three objects.

Lemma 5.7 *As an $H^*(BT;\boldsymbol{k})$–algebra, $\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$ is isomorphic to $\mathcal{SR}[\Sigma;\boldsymbol{k}]$ for $\boldsymbol{k}=\mathbb{Z}$ or $\mathbb{Q}$.*

Proof We construct an isomorphism between $\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$ and $\mathcal{SR}[\Sigma;\boldsymbol{k}]$, where $\boldsymbol{k}=\mathbb{Z}$ or $\mathbb{Q}$. Assume that $|\Sigma^{(1)}|=m$. Define a map

$$\phi\colon \boldsymbol{k}[x_1,\dots,x_m]\to\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}] \tag{5-11}$$

by restriction to each cone of $\widehat{\Sigma}$. Then this map ϕ is a surjective ring homomorphism. Indeed, given $\{g_{\widehat{\sigma}}\}\in\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$, we can apply the *inclusion–exclusion principle* to obtain

$$h(x_1,\dots,x_m)=\sum_{j=0}^{n-1}\Biggl((-1)^j\sum_{\substack{\widehat{\tau}\in\widehat{\Sigma}\\ \dim\widehat{\tau}=n-j}} g_{\widehat{\tau}}(x_{i_1},\dots,x_{i_{n-j}})\Biggr) \tag{5-12}$$

which is the desired global function h satisfying $\phi(h)=\{g_{\widehat{\sigma}}\}$, where $\widehat{\sigma}\in\widehat{\Sigma}^{(n)}$.

Moreover, since the zero element in $\mathcal{PP}[\widehat{\Sigma};\boldsymbol{k}]$ is $\{g_{\widehat{\sigma}}=0\mid\widehat{\sigma}\in\widehat{\Sigma}^{(n)}\}$, the kernel is

$$\ker\phi=\operatorname{span}\Biggl\{\prod_{j=1}^{k}x_{i_j}\;\Bigg|\;\operatorname{cone}\{\boldsymbol{e}_{i_1},\dots,\boldsymbol{e}_{i_k}\}\notin\widehat{\Sigma}\Biggr\},$$

which is exactly the Stanley–Reisner ideal $\mathcal{I}$ of Σ. Hence, the result follows. □

Corollary 5.8 *There is an isomorphism $\mathcal{PP}[\Sigma]\cong w\mathcal{SR}[\Sigma]$ (see (5-3)) as $H^*(BT)$–algebras.*

Proof Consider the composition of ring homomorphisms

$$\mathcal{PP}[\Sigma]\xhookrightarrow{\Lambda^*}\mathcal{PP}[\widehat{\Sigma}]\xrightarrow{\Phi^{-1}}\mathcal{SR}[\Sigma],$$

where $\Phi\colon\mathcal{SR}[\Sigma]\to\mathcal{PP}[\widehat{\Sigma}]$ is the isomorphism induced by ϕ. With $\mathbb{Z}$–coefficients, the map Λ^* is injective by Lemma 5.6. Hence, $\mathcal{PP}[\Sigma]$ is isomorphic to its image in $\mathcal{SR}[\Sigma]$ via the composition $\Phi^{-1}\circ\Lambda^*$.

Recall that the composition $\Phi^{-1}\circ\Lambda^*$ is an isomorphism over $\mathbb{Q}$, whose inverse $\Theta\circ\Phi^{-1}$ maps an element $[h]\in\mathcal{SR}[\Sigma;\mathbb{Q}]$ to $\{h(z^\sigma)\}_{\sigma\in\Sigma^{(n)}}\in\mathcal{PP}[\Sigma;\mathbb{Q}]$. Therefore, over integer coefficients, $[h] \in \mathrm{im}(\Phi^{-1} \circ \Lambda^*)$ if and only if the polynomial h satisfies the integrality condition. Hence, the result follows. □

Finally, we conclude this subsection with a proof of Theorem 5.3.

Proof of Theorem 5.3 Since $H^*(X_\Sigma;\mathbb{Z})$ concentrated in even degrees, the Serre spectral sequence for the fibration

$$X_\Sigma \to ET \times_T X_\Sigma \xrightarrow{\pi} BT$$

degenerates at the E_2 level. By the result from Franz and Puppe [11, Theorem 1.1], we get isomorphisms of $H^*(BT)$–algebras,

$$H^*(X_\Sigma) \cong H^*_T(X_\Sigma) \otimes_{H^*(BT)} \mathbb{Z} \cong H^*_T(X_\Sigma)/\mathrm{Im}(\pi^*\colon H^*(BT) \to H^*_T(X_\Sigma)).$$

By Proposition 5.5 and Corollary 5.8, we have $H^*_T(X_\Sigma) \cong w\mathcal{SR}[\Sigma]$. Moreover, for each $u_j \in \mathbb{Z}[u_1,\dots,u_n] \cong H^*(BT)$,

$$(\Phi \circ \Lambda^*)(u_j) = \sum_{i=1}^{m} \langle \lambda_i, e_j \rangle x_i.$$

Hence, we conclude that $\mathrm{im}(\pi^*\colon H^*(BT) \to H^*_T(X_\Sigma)) = \mathcal{J}$. □

6 Example: orbifold Hirzebruch varieties

We finish this paper by illustrating the results of the previous sections with a concrete example which is not a weighted projective space. Consider a primitive vector $(a,b) \in \mathbb{Z}^2$ with $a > 0$. Together with $(-1,0)$, $(0,1)$ and $(0,-1)$, we can make a complete fan Σ in $\mathbb{R}^2$ which gives us a compact toric variety with two singular points. We denote this toric variety by $\mathscr{H}_{(a,b)}$. See Figure 6 for the fan and corresponding $\mathcal{R}$–characteristic pair (Q,λ). When $a = 1$, the toric variety is known as a Hirzebruch surface, say $\mathscr{H}_b$. In this point of view, let us call $\mathscr{H}_{(a,b)}$ an *orbifold Hirzebruch variety*.

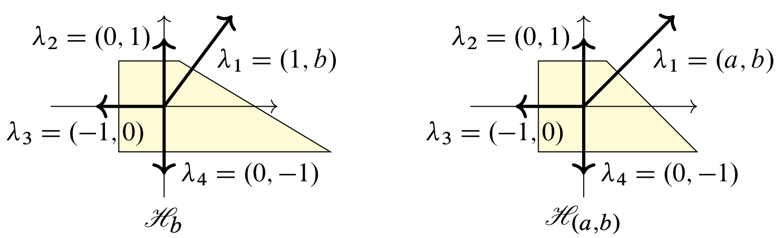


Figure 6: A Hirzebruch surface and an orbifold Hirzebruch variety

Since the collection in (1-1) becomes $\{|G_Q(v)| : v \in V(Q)\} = \{1, 1, a, a\}$ when $B_1 = Q$, its gcd is 1. Moreover, in any retraction sequence, B_2 is given by a union of edges, which guarantees that (Q, λ) satisfies the assumption of Theorem 1.1; see Example 4.4. Moreover, since the underlying polytope is a square, the integral Betti numbers are given by $\beta^0 = \beta^4 = 1$ and $\beta^2 = 2$ by the remark on page 3802.

Remark We may compute the (co)homology groups of low-dimensional toric orbifolds by the spectral sequence whose E_1 page is described by the fan data; see [10; 15]. More generally, the low-dimensional calculations of Kuwata, Masuda and Zeng [19] apply to the category of torus orbifolds.

Let $\sigma_{ij} = \mathrm{cone}\{\lambda_i, \lambda_j\}$, where $\lambda_1, \dots, \lambda_4$ are described in the right-hand side of Figure 6. Then the integrality condition of Definition 5.2 is given by the vectors

$$\begin{aligned} z^{\sigma_{12}} &= \Big(\frac{1}{a}u_1,\ -\frac{b}{a}u_1 + u_2,\ 0,\ 0\Big),\\ z^{\sigma_{14}} &= \Big(\frac{1}{a}u_1,\ 0,\ 0,\ \frac{b}{a}u_1 - u_2\Big),\\ z^{\sigma_{23}} &= (0,\ u_2,\ -u_1,\ 0),\\ z^{\sigma_{34}} &= (0,\ 0,\ -u_1,\ -u_2). \end{aligned}$$

Notice that the last two vectors, $z^{\sigma_{23}}$ and $z^{\sigma_{34}}$, don't contribute to the integrality condition, because their entries have integral coefficients.

A similar computation to Example 5.4 shows that the following polynomials are elements of degree 2 in $w\mathcal{SR}[\Sigma]$:

$$ax_1 - x_3, \quad bx_1 + x_2 - x_4, \quad ax_1, \quad ax_2, \quad x_3, \quad ax_4, \tag{6-1}$$

as are $\mathbb{Z}$–linear combinations of them. The first two elements are actually the linear relations in $\mathcal{J}$, which means that they come from the global polynomials in $\mathcal{PP}[\Sigma]$. Since rank $H^2(\mathscr{X}_{(a,b)}) = 2$, we choose two linearly independent elements as follows:

$$w_1 := ax_1 \quad \text{and} \quad w_2 := ax_4.$$

Next, degree-4 elements in $w\mathcal{SR}[\Sigma]$ are

$$a^2x_1^2, \quad a^2x_2^2, \quad x_3^2, \quad a^2x_4^2, \quad a^2x_1x_2, \quad a^2x_1x_4, \quad x_2x_3 \quad \text{and} \quad x_3x_4, \tag{6-2}$$

and their $\mathbb{Z}$–linear combination. The first four of (6-2) are just the square of degree-2 elements. The remaining four monomials are

- $a^2x_1x_2 = ax_1ax_2 = ax_1a(-bx_1 + x_4) = w_1(-bw_1 + w_2)$,
- $a^2x_1x_4 = ax_1ax_4 = w_1^2$,

Notice that the final two monomials x_2x_3 and x_3x_4 cannot come from degree-2 elements. Hence, we put

$$w_3 := x_3x_4.$$

Then

$$x_2x_3 = (-bx_1 + x_4)x_3 = x_3x_4 = ax_1x_4 = w_3.$$

The second equality holds because of the Stanley–Reisner ideal $\mathcal{I} = \langle x_1x_3, x_2x_4\rangle$. Finally, the ideal $\mathcal{I}$ and $\mathcal{J}$ determine the multiplicative structures as follows:

$$\begin{aligned}
w_1^2 &= (ax_1)^2 = (ax_1)(x_3) = 0,\\
w_1w_2 &= (ax_1)(ax_4) = x_3(ax_4) = aw_3,\\
w_2^2 &= (ax_4)(ax_4) = a(bx_1 + x_2)(ax_4) = abx_3x_4 = abw_3,\\
w_1w_3 &= (ax_1)(x_3x_4) = 0,\\
w_2w_3 &= (ax_4)(x_3x_4) = ax_4x_3(bx_1 + x_2) = 0,\\
w_3^2 &= (x_3x_4)^2 = x_3^2x_4^2 = (ax_1x_3)(x_4^2) = 0.
\end{aligned}$$

Therefore, we get the following presentation for the cohomology ring of orbifold Hirzebruch varieties:

$$\begin{aligned}
\text{(6-3)}\quad & H^*(\mathscr{H}_{(a,b)})\\
&\cong \mathbb{Z}[w_1, w_2, w_3]/\langle w_1^2, w_1w_2 - aw_3, w_2^2 - abw_3, w_1w_3, w_2w_3, w_3^2\rangle,
\end{aligned}$$

where $\deg w_1 = \deg w_2 = 2$ and $\deg w_3 = 4$.

Remark The cohomology ring of Hirzebruch surfaces, by way of comparison, can be computed from the results of [7], [9] or [16]. Indeed it has the presentation

$$H^*(\mathscr{H}_b) \cong \mathbb{Z}[w_1, w_2]/(w_1^2, w_2^2 - bw_1w_2),$$

where $\deg w_1 = \deg w_2 = 2$, which means that it is generated by degree-2 elements. However, $H^*(\mathscr{H}_{(a,b)})$ has the degree-4 generator w_3 which cannot be generated by degree-2 elements, ie $w_1w_2 = aw_3$. Notice that we can recover the presentation of $H^*(\mathscr{H}_b)$ by replacing a by 1 in (6-3).

References

[1] **A Bahri**, **M Franz**, **N Ray**, *The equivariant cohomology ring of weighted projective space*, Math. Proc. Cambridge Philos. Soc. 146 (2009) 395–405 MR

[2] **A Bahri**, **S Sarkar**, **J Song**, *Shellability and retraction sequences*, in preparation

[3] **A Borel**, *Seminar on transformation groups*, Annals of Mathematics Studies 46, Princeton Univ. Press (1960) MR

[4] **V M Buchstaber**, **T E Panov**, *Torus actions and their applications in topology and combinatorics*, University Lecture Series 24, Amer. Math. Soc., Providence, RI (2002) MR

[5] **W Buczynska**, *Fake weighted projective spaces*, preprint (2008) arXiv

[6] **D A Cox**, **J B Little**, **H K Schenck**, *Toric varieties*, Graduate Studies in Mathematics 124, Amer. Math. Soc., Providence, RI (2011) MR

[7] **V I Danilov**, *The geometry of toric varieties*, Uspekhi Mat. Nauk 33 (1978) 85–134 MR In Russian; translated in Russian Math. Surveys 33 (1978) 97–154

[8] **A Darby**, **S Kuroki**, **J Song**, *The equivariant cohomology of torus orbifolds*, in preparation

[9] **M W Davis**, **T Januszkiewicz**, *Convex polytopes, Coxeter orbifolds and torus actions*, Duke Math. J. 62 (1991) 417–451 MR

[10] **S Fischli**, *On toric varieties*, PhD thesis, Universität Bern (1992) Available at `https://www.sws.bfh.ch/~fischli/thesis.pdf`

[11] **M Franz**, **V Puppe**, *Exact cohomology sequences with integral coefficients for torus actions*, Transform. Groups 12 (2007) 65–76 MR

[12] **W Fulton**, *Introduction to toric varieties*, Annals of Mathematics Studies 131, Princeton Univ. Press (1993) MR

[13] **M Harada**, **A Henriques**, **T S Holm**, *Computation of generalized equivariant cohomologies of Kac–Moody flag varieties*, Adv. Math. 197 (2005) 198–221 MR

[14] **M Harada**, **T S Holm**, **N Ray**, **G Williams**, *The equivariant K–theory and cobordism rings of divisive weighted projective spaces*, Tohoku Math. J. 68 (2016) 487–513 MR

[15] **A Jordan**, *Homology and cohomology of toric varieties*, PhD thesis, University of Konstanz (1998) Available at `https://kops.uni-konstanz.de/bitstream/handle/123456789/6176/preprint_057.pdf`

[16] **J Jurkiewicz**, *Torus embeddings, polyhedra, k^*–actions and homology*, Dissertationes Math. (Rozprawy Mat.) 236, Polska Akademia Nauk. Instytut Matematyczny, Warsaw (1985) MR

[17] **A M Kasprzyk**, *Bounds on fake weighted projective space*, Kodai Math. J. 32 (2009) 197–208 MR

[18] **T Kawasaki**, *Cohomology of twisted projective spaces and lens complexes*, Math. Ann. 206 (1973) 243–248 MR

[19] **H Kuwata**, **M Masuda**, **H Zeng**, *Torsion in the cohomology of torus orbifolds*, preprint (2016) arXiv

[20] **M Poddar**, **S Sarkar**, *On quasitoric orbifolds*, Osaka J. Math. 47 (2010) 1055–1076 MR

[21] **S Sarkar**, **D Y Suh**, *A new construction of lens spaces*, preprint (2013) arXiv

[22] **S Sarkar**, **V Uma**, *Equivariant k–theory and cobordism rings of toric orbifolds*, in preparation

[23] **G M Ziegler**, *Lectures on polytopes*, Graduate Texts in Mathematics 152, Springer (1995) MR

Department of Mathematics, Rider University
Lawrenceville, NJ, United States

Department of Mathematics, Indian Institute of Technology Madras
Chennai, India

Department of Mathematical Sciences, KAIST
Daejeon, South Korea

bahri@rider.edu, soumen@iitm.ac.in, jongbaek.song@gmail.com

Received: 20 December 2016 Revised: 22 March 2017 msp

Algebraic & Geometric Topology 17 (2017) 3811–3836

Remarks on coloured triply graded link invariants

Sabin Cautis

We explain how existing results (such as categorical $\mathfrak{sl}_n$ actions, associated braid group actions and infinite twists) can be used to define a triply graded link invariant which categorifies the HOMFLY polynomial of links coloured by arbitrary partitions. The construction uses a categorified HOMFLY clasp defined via cabling and infinite twists. We briefly discuss differentials and speculate on related structures.

57M27; 16T99

1 Introduction

In [14; 16] Khovanov and Rozansky defined a triply graded link invariant using matrix factorizations and subsequently Soergel bimodules. In their case the link is coloured by the partition (1) and the invariant categorifies the HOMFLY polynomial. In this paper we explain how existing tools can be used to extend this construction to links coloured by arbitrary partitions, which categorifies the coloured HOMFLY polynomial.

The idea is as follows. First one defines a 2–category $\mathcal{K}_n$ out of Soergel bimodules and constructs a categorical $(\mathfrak{sl}_n, \theta)$ action on it (Sections 2 and 3). Combining this action with a trace 2–functor (Hochschild (co)homology) one obtains a triply graded invariant for links coloured by partitions with only one part (k) for $k \in \mathbb{N}$.

Finally, to deal with an arbitrary partition $(k_1, \dots, k_i)$, one cables together i strands labeled $k_1, \dots, k_i$ and composes with a certain projector P^-. We will call these (categorified) HOMFLY clasps to differentiate them from those in the Reshetikhin–Turaev context (RT clasps). Apart from some general results on $(\mathfrak{sl}_n, \theta)$ actions and associated braid group actions and projectors (see for instance [5]) I have tried to keep this paper self-contained. Some example computations are worked out in Section 7.

There are many papers in the literature on coloured HOMFLY homology and it is difficult to list them all without forgetting some. We try to recall some of the ones which are more closely related to this paper.

There are several papers defining various generalities of triply graded homologies. In [20] Mackaay, Stošić and Vaz work out the case of links labeled by the one-part

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3811

partition (2). In [24] Webster and Williamson define a triply graded homology of links coloured by partitions with only one part. Their construction, which is geometric, is related to ours via the equivalence between perverse sheaves on finite flag varieties and (singular) Soergel bimodules. The same relationship appears (and is briefly discussed) in Cautis, Dodd and Kamnitzer [6]. More recently, Wedrich [25] examines these constructions in the "reduced" case as well as some associated spectral sequences.

The papers Abel and Hogancamp [1] and Hogancamp [13] discuss the categorified HOMFLY clasps for partitions with parts of size at most one (ie coloured with (1^k) for $k \in \mathbb{N}$). As with our projectors, these are built as infinite twists. As far as I understand, Elias and Hogancamp aim to develop a more systematic, more general construction of such projectors. This will hopefully shed some light on the projectors P^- and the various properties they (are expected to) satisfy.

In Dunfield, Gukov and Rasmussen [11] and Rasmussen [21] it was conjectured (and partially proved) that there exist certain differentials on triply graded link homology which recover $\mathrm{SL}(N)$ link homology. In Section 6 we discuss a differential d_N for $N > 0$ which gives rise to an $\mathrm{SL}(N)$–type link invariant. Somewhat surprisingly, the resulting homology seems to be finite-dimensional while at the same time it categorifies $\mathrm{SL}(N)$–representations of the form $\mathrm{Sym}^k(\mathbb{C}^N)$. A homology of this form is predicted by the physical interpretation of knot homologies as spaces of open BPS states (see for instance Gukov and Stošić [12]) but does not show up in our earlier work on knot homologies (Cautis and Kamnitzer [7] and subsequent papers). In Section 8 we also speculate on defining differentials d_N for $N < 0$ which should categorify $\mathrm{SL}(N)$–representations of the form $\Lambda^k(\mathbb{C}^N)$.

The author was supported by an NSERC discovery/accelerator grant.

2 Background: $(\mathfrak{sl}_n, \theta)$ actions, braid group actions and projectors

2.1 Notation

We work over an arbitrary field $\Bbbk$. By a graded 2–category $\mathcal{K}$ we mean a 2–category whose 1–morphisms are equipped with an auto-equivalence $\langle 1 \rangle$ (so graded means $\mathbb{Z}$–graded). We say $\mathcal{K}$ is idempotent complete if for any 2–morphism f with $f^2 = f$ the image of f exists in $\mathcal{K}$.

For $n \geq 1$ we denote by $[n]$ the quantum integer $q^{n-1} + q^{n-3} + \cdots + q^{-n+3} + q^{-n+1}$, where q is a formal variable. By convention, for negative entries we let $[-n] = -[n]$. Moreover, $[n]! := [1][2]\dots[n]$ and $\begin{bmatrix} n \\ k \end{bmatrix} := \frac{[n]!}{[k]![n-k]!}$.

If $f = f_a q^a \in \mathbb{N}[q, q^{-1}]$ and A is a 1–morphism inside a graded 2–category $\mathcal{K}$ then we write $\bigoplus_f \mathsf{A}$ for the direct sum $\bigoplus_{s\in\mathbb{Z}} \mathsf{A}^{\oplus f_s}\langle s\rangle$. For example,

$$\bigoplus_{[n]} \mathsf{A} = \mathsf{A}\langle n-1\rangle \oplus \mathsf{A}\langle n-3\rangle \oplus \cdots \oplus \mathsf{A}\langle -n+3\rangle \oplus \mathsf{A}\langle -n+1\rangle.$$

We will always assume $\mathbb{N}$ contains 0. Moreover, we will write $\mathrm{End}^i(\mathsf{A})$ as shorthand for $\mathrm{Hom}(\mathsf{A}, \mathsf{A}\langle i\rangle)$, where $i \in \mathbb{Z}$.

Finally, if $\gamma_i\colon \mathsf{A}_i \to \mathsf{B}_i$ is a sequence of 2–morphisms in $\mathcal{K}$ for $i = 1, \dots, k$, we will write $\gamma_1 \cdots \gamma_k\colon \mathsf{A}_1 \cdots \mathsf{A}_k \to \mathsf{B}_1 \cdots \mathsf{B}_k$ for the corresponding 2–morphism. We will denote by I the identity 2–morphism.

2.2 Categorical actions

In [4], $(\mathfrak{g}, \theta)$ actions were introduced in order to simplify some of the earlier definitions from [15; 22; 10]. A $(\mathfrak{g}, \theta)$ action involves a target graded, additive, $\Bbbk$–linear, idempotent complete 2–category $\mathcal{K}$ whose objects are indexed by the weight lattice of $\mathfrak{g}$.

In this paper we only consider the case $\mathfrak{g} = \mathfrak{sl}_n$. The vertex set of the Dynkin diagram of $\mathfrak{sl}_n$ is indexed by $I = \{1, \dots, n-1\}$. However, it will be more convenient if the objects $\mathcal{K}(\underline{k})$ of $\mathcal{K}$ are indexed by $\underline{k} = (k_1, \dots, k_n) \in \mathbb{Z}^n$, which we can identify with the weight lattice of $\mathfrak{gl}_n$. In this notation the root lattice is generated by $\alpha_i = (0, \dots, -1, 1, \dots, 0)$ for $i \in I$ (this notation agrees with that in [5]). We equip $\mathbb{Z}^n$ with the standard nondegenerate bilinear form $\langle \cdot, \cdot \rangle\colon \mathbb{Z}^n \times \mathbb{Z}^n \to \mathbb{Z}$ (so that $\langle \alpha_i, \alpha_j \rangle$ is given by the standard Cartan datum for $\mathfrak{gl}_n$).

We require that the 2–category $\mathcal{K}$ is equipped with the following:

- **1–morphisms** $\mathsf{E}_i \mathbf{1}_{\underline{k}} = \mathbf{1}_{\underline{k}+\alpha_i} \mathsf{E}_i$ and $\mathsf{F}_i \mathbf{1}_{\underline{k}+\alpha_i} = \mathbf{1}_{\underline{k}} \mathsf{F}_i$, where $\mathbf{1}_{\underline{k}}$ is the identity 1–morphism of $\mathcal{K}(\underline{k})$.
- **2–morphisms** For each $\underline{k} \in \mathbb{Z}^n$, a $\Bbbk$–linear map $\mathrm{span}\{\alpha_i : i \in I\} \to \mathrm{End}^2(\mathbf{1}_{\underline{k}})$.

We abuse notation and denote by $\theta \in \mathrm{End}^2(\mathbf{1}_{\underline{k}})$ the image of $\theta \in \mathrm{span}\{\alpha_i : i \in I\}$ under the linear maps above. On this data we impose the following conditions.

(1) $\mathrm{End}^l(\mathbf{1}_{\underline{k}})$ is zero if $l < 0$ and one-dimensional if $l = 0$ and $\mathbf{1}_{\underline{k}} \neq 0$. Moreover, the space of maps between any two 1–morphisms is finite-dimensional.

(2) E_i and F_i are left and right adjoints of each other up to specified shifts. More precisely:
   (a) $(\mathsf{E}_i \mathbf{1}_{\underline{k}})^R \cong \mathbf{1}_{\underline{k}} \mathsf{F}_i \langle \langle \underline{k}, \alpha_i \rangle + 1 \rangle$,
   (b) $(\mathsf{E}_i \mathbf{1}_{\underline{k}})^L \cong \mathbf{1}_{\underline{k}} \mathsf{F}_i \langle -\langle \underline{k}, \alpha_i \rangle - 1 \rangle$.

(3) We have

$$\begin{cases} \mathsf{E}_i\mathsf{F}_i\mathbf{1}_{\underline{k}} \cong \mathsf{F}_i\mathsf{E}_i\mathbf{1}_{\underline{k}} \oplus_{[\langle \underline{k},\alpha_i\rangle]} \mathbf{1}_{\underline{k}} & \text{if } \langle \underline{k},\alpha_i\rangle \geq 0, \\ \mathsf{F}_i\mathsf{E}_i\mathbf{1}_{\underline{k}} \cong \mathsf{E}_i\mathsf{F}_i\mathbf{1}_{\underline{k}} \oplus_{[-\langle \underline{k},\alpha_i\rangle]} \mathbf{1}_{\underline{k}} & \text{if } \langle \underline{k},\alpha_i\rangle \leq 0. \end{cases}$$

(4) If $i \neq j \in I$, then $\mathsf{F}_j\mathsf{E}_i\mathbf{1}_{\underline{k}} \cong \mathsf{E}_i\mathsf{F}_j\mathbf{1}_{\underline{k}}$.

(5) For $i \in I$ we have

$$\mathsf{E}_i\mathsf{E}_i \cong \mathsf{E}_i^{(2)}\langle -1\rangle \oplus \mathsf{E}_i^{(2)}\langle 1\rangle$$

for some 1–morphism $\mathsf{E}_i^{(2)}$. Moreover, if $\theta \in \operatorname{span}\{\alpha_i : i \in I\}$ then the map $I\theta I \in \operatorname{End}^2(\mathsf{E}_i\mathbf{1}_{\underline{k}}\mathsf{E}_i)$ induces a map between the summands $\mathsf{E}_i^{(2)}\langle 1\rangle$ on either side which is

- nonzero if $\langle \theta,\alpha_i\rangle \neq 0$, and
- zero if $\langle \theta,\alpha_i\rangle = 0$.

(6) If $\alpha = \alpha_i$ or $\alpha = \alpha_i + \alpha_j$ for some $i, j \in I$ with $|i-j| = 1$, then $\mathbf{1}_{\underline{k}+r\alpha} = 0$ for $r \gg 0$ or $r \ll 0$.

(7) Suppose $i \neq j \in I$. If $\mathbf{1}_{\underline{k}+\alpha_i}$ and $\mathbf{1}_{\underline{k}+\alpha_j}$ are nonzero, then $\mathbf{1}_{\underline{k}}$ and $\mathbf{1}_{\underline{k}+\alpha_i+\alpha_j}$ are also nonzero.

In [4, Theorem 1.1] we showed that such an $(\mathfrak{sl}_n,\theta)$ action must carry an action of the quiver Hecke algebras (KLR algebras). In particular, this gives us decompositions

$$\mathsf{E}_i^r \cong \bigoplus_{[r]!} \mathsf{E}_i^{(r)} \quad \text{and} \quad \mathsf{F}_i^r \cong \bigoplus_{[r]!} \mathsf{F}_i^{(r)}$$

for certain 1–morphisms $\mathsf{E}_i^{(r)}$ and $\mathsf{F}_i^{(r)}$ (called divided powers). These satisfy

$$(\mathsf{E}_i^{(r)}\mathbf{1}_{\underline{k}})^R \cong \mathbf{1}_\lambda \mathsf{F}_i^{(r)}\langle r(\langle \underline{k},\alpha_i\rangle + r)\rangle,$$
$$(\mathsf{E}_i^{(r)}\mathbf{1}_{\underline{k}})^L \cong \mathbf{1}_\lambda \mathsf{F}_i^{(r)}\langle -r(\langle \underline{k},\alpha_i\rangle + r)\rangle.$$

2.3 (Categorical) braid group actions

The reason we are interested in $(\mathfrak{sl}_n,\theta)$ actions is that they can be used to define braid group actions [8], as we now recall.

Suppose that, as above, we have an $(\mathfrak{sl}_n,\theta)$ action on a 2–category $\mathcal{K}$. Denote by $\mathsf{Kom}(\mathcal{K})$ the bounded homotopy category of $\mathcal{K}$ (where objects are the same as in $\mathcal{K}$, 1–morphisms are complexes of 1–morphisms which are bounded from above and below and 2–morphisms are maps of complexes). We define $\mathsf{T}_i\mathbf{1}_{\underline{k}} \in \mathsf{Kom}(\mathcal{K})$ as

$$\big[\cdots \to \mathsf{E}_i^{(-\langle \underline{k},\alpha_i\rangle+2)}\mathsf{F}_i^{(2)}\langle -2\rangle \to \mathsf{E}_i^{(-\langle \underline{k},\alpha_i\rangle+1)}\mathsf{F}_i\langle -1\rangle \to \mathsf{E}_i^{(-\langle \underline{k},\alpha_i\rangle)}\big]\mathbf{1}_{\underline{k}} \quad \text{if } \langle \underline{k},\alpha_i\rangle \leq 0,$$
$$\big[\cdots \to \mathsf{F}_i^{(\langle \underline{k},\alpha_i\rangle+2)}\mathsf{E}_i^{(2)}\langle -2\rangle \to \mathsf{F}_i^{(\langle \underline{k},\alpha_i\rangle+1)}\mathsf{E}_i\langle -1\rangle \to \mathsf{F}_i^{(\langle \underline{k},\alpha_i\rangle)}\big]\mathbf{1}_{\underline{k}} \quad \text{if } \langle \underline{k},\alpha_i\rangle \geq 0.$$

One can show the differentials must be the unique nonzero maps. Notice that these complexes are actually bounded on the left since $\mathbf{1}_{\underline{k}\pm r\alpha_i} = 0$ if $r \gg 0$. The main result of [8] states that these complexes give us a braid group action. This fact categorifies a classical result of Lusztig [19, 5.2.1].

2.4 Categorified projectors

To obtain projectors let us first consider

$$\mathsf{T}_\omega \mathbf{1}_{\underline{k}} := (\mathsf{T}_{n-1})(\mathsf{T}_{n-2}\mathsf{T}_{n-1})\cdots(\mathsf{T}_2\cdots\mathsf{T}_{n-1})(\mathsf{T}_1\cdots\mathsf{T}_{n-1})\mathbf{1}_{\underline{k}},$$

corresponding to a half-twist in the braid group. In [5, Section 5.2] we constructed a natural map $\mathbf{1}_{\underline{k}} \to \mathsf{T}_\omega^2 \mathbf{1}_{\underline{k}}$ and showed that there is a well-defined limit $\mathsf{P}^-\mathbf{1}_{\underline{k}} := \lim_{\ell\to\infty} \mathsf{T}_\omega^{2\ell}\mathbf{1}_{\underline{k}}$ which lives in a certain subcategory $\mathsf{Kom}_*^-(\mathcal{K}) \subset \mathsf{Kom}^-(\mathcal{K})$ of the bounded-above homotopy category (see [5, Section 3.5] for more details).

Remark To illustrate, if $\mathcal{K}$ was the category of $\mathbb{Z}$–graded $\Bbbk$–vector spaces then $\bigoplus_{i\geq 0} \Bbbk[i]\langle -i\rangle$ would belong to $\mathsf{Kom}_*^-(\mathcal{K})$ because $\sum_{i\geq 0}(-1)^i q^i [\Bbbk]$ converges to $\frac{1}{1+q}[\Bbbk]$ (here $[\Bbbk]$ is the class in K-theory of the one-dimensional vector space). On the other hand, $\bigoplus_{i\geq 0}\Bbbk[i]$ would not belong to $\mathsf{Kom}_*^-(\mathcal{K})$ because $\sum_{i\geq 0}(-1)^i[\Bbbk]$ does not converge.

Having shown that $\mathsf{P}^-\mathbf{1}_{\underline{k}}$ is well-defined it is then easy to see that $\mathsf{P}^-\mathbf{1}_{\underline{k}}$ is idempotent, meaning that $\mathsf{P}^-\mathsf{P}^-\mathbf{1}_{\underline{k}} \cong \mathsf{P}^-\mathbf{1}_{\underline{k}}$. The main result of [5] showed using an instance of skew Howe duality that P^- can be used to categorify all the clasps. The inspiration of using infinite twists to categorify clasps goes back to Rozansky [23], who categorified Jones–Wenzl projectors within Bar-Natan's graphical formulation of Khovanov homology.

3 The category $\mathcal{K}_n$

3.1 Categories and functors

We now define a 2–category $\mathcal{K}_n$ with an $(\mathfrak{sl}_n, \theta)$ action.

For $k \in \mathbb{N}$ consider the affine space $\mathbb{A}^k := \operatorname{Spec} \Bbbk[x_1,\dots,x_k]$, where $\deg(x_\ell) = 2$ for each ℓ (the grading is equivalent to endowing $\mathbb{A}^k$ with a $\Bbbk^\times$ action). The quotient $\mathbb{A}_k := \mathbb{A}^k/S_k$ by the symmetric group S_k on k letters is isomorphic to $\operatorname{Spec}\Bbbk[e_1,\dots,e_k]$, where the e_ℓ are the elementary symmetric functions and $\deg(e_\ell) = 2\ell$. For a sequence $\underline{k}$ we write $\mathbb{A}_{\underline{k}} := \mathbb{A}_{k_1} \times \cdots \times \mathbb{A}_{k_n}$. Finally, we will denote by $D(\mathbb{A}_{\underline{k}})$ the derived category of $\Bbbk^\times$–equivariant quasicoherent sheaves on $\mathbb{A}_{\underline{k}}$. We will denote by $\{\cdot\}$ a shift in the grading induced by the $\Bbbk^\times$ action. In particular, this means that multiplication by e_ℓ induces a map $\mathcal{O}_{\mathbb{A}_{\underline{k}}} \to \mathcal{O}_{\mathbb{A}_{\underline{k}}}\{2\ell\}$ since e_ℓ has degree 2ℓ. This is the same convention as in earlier papers such as [7].

For $n \in \mathbb{N}$ the 2–category $\mathcal{K}_n$ is defined as follows. The objects are the categories $D(\mathbb{A}_{\underline{k}})$. The 1–morphisms are all kernels on products $\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}'}$ (with composition given by the convolution product $\star$) and the 2–morphisms are morphisms between kernels. The grading shift $\langle 1 \rangle$ is by definition $\{1\}$.

Note that for $a, b \in \mathbb{N}$ there exists a natural projection map $\pi\colon \mathbb{A}_{a,b} \to \mathbb{A}_{a+b}$. This map is finite of degree $\binom{a+b}{a}$. More generally, we can consider correspondences such as

$$\begin{tikzcd} & \mathbb{A}_{(\dots,k_i-r,r,k_{i+1},\dots)} \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ \mathbb{A}_{\underline{k}} & & \mathbb{A}_{\underline{k}+r\alpha_i} \end{tikzcd} \tag{1}$$

where $\alpha_i = (0, \dots, -1, 1, \dots, 0)$ with a -1 in position i. We then define the following data:

- 1–morphisms
$$\mathcal{E}_i \mathbf{1}_{\underline{k}} := \mathcal{O}_{\mathbb{A}_{(\dots,k_i-1,1,k_{i+1},\dots)}}\{k_i - 1\} \in D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}+\alpha_i}),$$
$$\mathbf{1}_{\underline{k}} \mathcal{F}_i := \mathcal{O}_{\mathbb{A}_{(\dots,k_i-1,1,k_{i+1},\dots)}}\{k_{i+1}\} \in D(\mathbb{A}_{\underline{k}+\alpha_i} \times \mathbb{A}_{\underline{k}}),$$
where we embed $\mathbb{A}_{(\dots,k_i-1,1,k_{i+1},\dots)}$ into $\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}+\alpha_i}$ using π_1 and π_2 from (1) (taking $r = 1$ in this case).
- A $\Bbbk$–linear map $\theta\colon \mathrm{span}\{\alpha_i : i \in I\} \to \mathrm{End}^2(\mathbf{1}_{\underline{k}})$ where the image of α_i is given by multiplication by $e_1^{(i)} - e_1^{(i+1)}$ where $e_1^{(i)}, e_2^{(i)}, \dots, e_{k_i}^{(i)}$ are the elementary generators of the factor $\mathbb{A}_{k_i}$ inside $\mathbb{A}_{\underline{k}}$.

Remark Although we use derived categories of quasicoherent sheaves, we could restrict everything to abelian categories of coherent sheaves. This is because all the morphisms involved are flat and finite. However, it is natural to work with these larger categories because later we will apply Hochschild cohomology.

Theorem 3.1 *The data above defines an $(\mathfrak{sl}_n, \theta)$ action on $\mathcal{K}_n$.*

Proof The fact that relations of an $(\mathfrak{sl}_n, \theta)$ action are satisfied is not difficult to prove and essentially follows from [15, Section 6]. The fact that θ satisfies relation (5) comes down to the following elementary fact. Consider $\Bbbk[x, y]$ as a $\Bbbk[x, y]^{S_2} \cong \Bbbk[e_1, e_2]$

bimodule where S_2 acts by switching x and y and (following our notation above) $e_1 = x + y$, $e_2 = xy$. Then, as a bimodule,

$$\Bbbk[x,y] \cong \Bbbk[e_1,e_2] \oplus \Bbbk[e_1,e_2]\{2\},$$

and multiplication by x (or y) induces an endomorphism of $\Bbbk[x,y]$ which is an isomorphism between the summands $\Bbbk[e_1,e_2]\{2\}$ on either side.

Perhaps one thing to note is that our choices of shifts in defining the $\mathcal{E}_i$ and the $\mathcal{F}_i$ differ slightly from [15]. However, the specific choice of shifts is not so important and is mainly determined by the fact that the canonical bundle of $\mathbb{A}_{\underline{k}}$ is $\omega_{\mathbb{A}_{\underline{k}}} \cong \mathcal{O}_{\mathbb{A}_{\underline{k}}}\{d_{\underline{k}}\}$, where $d_{\underline{k}} = -\sum_\ell k_\ell(k_\ell+1)$. □

It is not hard to show that the divided powers $\mathsf{E}_i^{(r)}\mathbf{1}_{\underline{k}}$ and $\mathbf{1}_{\underline{k}}\mathsf{F}_i^{(r)}$ are given by kernels

$$\begin{aligned}
\mathcal{E}_i^{(r)}\mathbf{1}_{\underline{k}} &:= \mathcal{O}_{\mathbb{A}_{(\dots,k_i-r,r,k_{i+1},\dots)}}\{r(k_i-r)\} \in D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}+r\alpha_i}),\\
\mathbf{1}_{\underline{k}}\mathcal{F}_i^{(r)} &:= \mathcal{O}_{\mathbb{A}_{(\dots,k_i-r,r,k_{i+1},\dots)}}\{rk_{i+1}\} \in D(\mathbb{A}_{\underline{k}+r\alpha_i} \times \mathbb{A}_{\underline{k}}),
\end{aligned}$$

where again we embed $\mathbb{A}_{(\dots,k_i-r,r,k_{i+1},\dots)}$ using (1) (we will not use this fact).

Remark There are three different gradings that show up. First, there is $\langle 1\rangle = \{1\}$, which corresponds to the grading induced by the $\Bbbk^\times$ action. Second, there is the cohomological grading $[1]$ in $\mathsf{Kom}_*^-(\mathcal{K}_n)$. Third, there is the cohomological grading $[\![1]\!]$, which is internal to $D(\mathbb{A}_{\underline{k}})$. This last grading only shows up when we apply the trace 2–functors described in Section 3.3.

3.2 The braid group action

Following Section 2.3 we define the braid group generators $\mathcal{T}_i\mathbf{1}_{\underline{k}} \in \mathsf{Kom}(\mathcal{K}_n)$ as

$$\Big[\cdots \to \mathcal{E}_i^{(-\langle \underline{k},\alpha_i\rangle+2)} \star \mathcal{F}_i^{(2)}\{-2\} \to \mathcal{E}_i^{(-\langle \underline{k},\alpha_i\rangle+1)} \star \mathcal{F}_i\{-1\} \to \mathcal{E}_i^{(-\langle \underline{k},\alpha_i\rangle)}\Big]\mathbf{1}_{\underline{k}} \text{ if } \langle \underline{k},\alpha_i\rangle \le 0,$$

$$\Big[\cdots \to \mathcal{F}_i^{(\langle \underline{k},\alpha_i\rangle+2)} \star \mathcal{E}_i^{(2)}\{-2\} \to \mathcal{F}_i^{(\langle \underline{k},\alpha_i\rangle+1)} \star \mathcal{E}_i\{-1\} \to \mathcal{F}_i^{(\langle \underline{k},\alpha_i\rangle)}\Big]\mathbf{1}_{\underline{k}} \text{ if } \langle \underline{k},\alpha_i\rangle \ge 0.$$

We also get the corresponding projectors $\mathcal{P}^-\mathbf{1}_{\underline{k}} \in \mathsf{Kom}_*^-(\mathcal{K}_n)$.

Following the construction in [5, Section 7.1] it is useful to also define the elements

$$\mathcal{T}_i'\mathbf{1}_{\underline{k}} := \begin{cases} \mathcal{T}_i\mathbf{1}_{\underline{k}}[-k_{i+1}]\{k_{i+1}+k_ik_{i+1}\} & \text{if } \langle \underline{k},\alpha_i\rangle \le 0,\\ \mathcal{T}_i\mathbf{1}_{\underline{k}}[-k_i]\{k_i+k_ik_{i+1}\} & \text{if } \langle \underline{k},\alpha_i\rangle \ge 0.\end{cases}$$

Notice that in contrast to [5, Section 7.1] we have an extra shift of $\{k_ik_{i+1}\}$. These $\mathcal{T}_i'$ also generate a braid group action but are better behaved with respect to the $\mathcal{E}_i$ and $\mathcal{F}_i$

since, using [5, Corollary 7.3] and [5, Corollary 4.6], we have

(2) $\mathcal{T}_i' \star \mathcal{T}_j' \star \mathcal{E}_i \cong \mathcal{E}_j \star \mathcal{T}_i' \star \mathcal{T}_j'$ and $\mathcal{T}_i' \star \mathcal{T}_j' \star \mathcal{F}_i \cong \mathcal{F}_j \star \mathcal{T}_i' \star \mathcal{T}_j'$ if $|i-j|=1$,

(3) $$\mathcal{T}_i' \star \mathcal{E}_i \cong \mathcal{F}_i \star \mathcal{T}_i' \quad \text{and} \quad \mathcal{T}_i' \star \mathcal{F}_i \cong \mathcal{E}_i \star \mathcal{T}_i'.$$

3.3 Trace 2–functors

For any $\ell \in \mathbb{N}$ we now define a 2–functor $\Psi_\ell\colon \mathcal{K}_n \to \mathcal{K}_{n+1}$. This functor should be thought of as adding a strand labeled ℓ.

At the level of objects Ψ_ℓ takes $D(\mathbb{A}_{\underline{k}})$ to $D(\mathbb{A}_{\underline{k},\ell})$. Given a 1–morphism $\mathcal{M} \in D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}'})$ we define

(4) $$\Psi_\ell(\mathcal{M}) := \Delta_* \pi^*(\mathcal{M}) \in D(\mathbb{A}_{\underline{k},\ell} \times \mathbb{A}_{\underline{k}',\ell}),$$

where π^* and Δ_* are pullback and pushforward with respect to the natural projection and diagonal maps

$$\pi\colon \mathbb{A}_{\underline{k}} \times \mathbb{A}_\ell \times \mathbb{A}_{\underline{k}'} \to \mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}'} \quad \text{and} \quad \Delta\colon \mathbb{A}_{\underline{k}} \times \mathbb{A}_\ell \times \mathbb{A}_{\underline{k}'} \to (\mathbb{A}_{\underline{k}} \times \mathbb{A}_\ell) \times (\mathbb{A}_{\underline{k}'} \times \mathbb{A}_\ell).$$

Given a 2–morphism $f\colon \mathcal{M} \to \mathcal{M}'$ we define $\Psi_\ell(f) := \Delta_* \pi^*(f)$. Using Corollary A.2 this defines a 2–functor $\Psi_\ell\colon \mathcal{K}_n \to \mathcal{K}_{n+1}$. It is not difficult to see that $\Psi_\ell(\mathcal{E}_i \mathbf{1}_{\underline{k}}) \cong \mathcal{E}_i \mathbf{1}_{\underline{k},\ell}$ and $\Psi_\ell(\mathbf{1}_{\underline{k}} \mathcal{F}_i) \cong \mathbf{1}_{\underline{k},\ell} \mathcal{F}_i$.

We can likewise define a 2–functor $\Psi'_\ell\colon \mathcal{K}_{n+1} \to \mathcal{K}_n$. On objects it takes $D(\mathbb{A}_{\underline{k},\ell})$ to $D(\mathbb{A}_{\underline{k}})$. All other objects, meaning $D(\mathbb{A}_{\underline{k},\ell'})$ where $\ell \neq \ell'$, are mapped to zero. On 1–morphisms it acts by

$$D(\mathbb{A}_{\underline{k},\ell} \times \mathbb{A}_{\underline{k}',\ell}) \to D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}'}), \quad \mathcal{N} \mapsto \Psi'_\ell(\mathcal{N}) := \pi_* \Delta^*(\mathcal{N}).$$

By Proposition A.3 we also have

(5) $$\Psi'_\ell(\mathcal{N} \star \Psi_\ell(\mathcal{M})) \cong \Psi'_\ell(\mathcal{N}) \star \mathcal{M},$$

where $\mathcal{M} \in D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}'})$ and $\mathcal{N} \in D(\mathbb{A}_{\underline{k},\ell} \times \mathbb{A}_{\underline{k}',\ell})$.

If we define $\bullet := \operatorname{Spec} \Bbbk$ then $D(\bullet)$ is the category of complexes of (possibly infinite-dimensional) graded vector spaces. For any $\underline{k}$ we define

$$\tau\colon D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}}) \to D(\bullet), \quad \mathcal{M} \mapsto \Psi'_{k_1} \circ \cdots \circ \Psi'_{k_n}(\mathcal{M}).$$

Note that this is just the Hochschild homology $\mathrm{HH}_*(\mathcal{M})$ of $\mathcal{M}$.

4 Link invariants

Consider an oriented link L whose components are coloured by partitions. For now we assume that each such partition has only one part, meaning it is of the form (k) for some $k \in \mathbb{N}$. Such a link can be given as the closure $\widehat{\beta}$ of a coloured braid β, where we visualize the strands of this braid vertically with the top and bottom labeled by the same sequence $\underline{k}$.

To a positive crossing exchanging strands i and $i+1$ (ie the strand starting at i crosses over the one starting at $i+1$) we associate the 1–morphism

$$\mathcal{T}_i' \in \mathsf{Kom}_*^-(D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{s_i \cdot \underline{k}}))$$

as defined in Section 3.2, where s_i acts on $\underline{k}$ by switching k_i and k_{i+1}. Composing these 1–morphisms gives a complex $\mathcal{T}_\beta' \in \mathsf{Kom}_*^-(D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}'}))$. The invariant associated to the closure $\widehat{\beta}$ of the braid is then $\tau(\mathcal{T}_\beta') \in \mathsf{Kom}_*^-(D(\bullet))$.

To deal with partitions with more than one part, we cable strands together and use the projector $\mathcal{P}^-$. More precisely, given a strand labeled by a partition

$$k_i^{(\cdot)} = (k_i^{(1)} \leq \cdots \leq k_i^{(p)}),$$

we replace it with p strands labeled $k_i^{(1)}, \ldots, k_i^{(p)}$ together with the projector $\mathcal{P}^- \mathbf{1}_{k_i^{(\cdot)}}$ on these strands.

Theorem 4.1 *Suppose $L = \widehat{\beta}$, where β is a braid whose strands are coloured by partitions. Then, up to an overall grading shift, $\mathcal{H}(L) := \tau(\mathcal{T}_\beta') \in \mathsf{Kom}_*^-(D(\bullet))$ defines a triply graded link invariant.*

Remark In order to obtain a homology which is invariant on the nose (not just up to shifts) one needs to shift the functor Ψ_ℓ' by $\left[\frac{\ell}{2}\right]\left[\!\left[-\frac{\ell}{2}\right]\!\right]$ and the definition of a $\mathcal{T}$ switching two strands labeled ℓ by $\left[\frac{\ell}{2}\right]\left[\!\left[-\frac{\ell}{2}\right]\!\right]$.

Before we can prove Theorem 4.1 we need the following lemma.

Lemma 4.2 *For $\mathcal{T}_1 \in \mathsf{Kom}_-^*(D(\mathbb{A}_{1,1} \times \mathbb{A}_{1,1}))$ we have*

$$\Psi_1'(\mathcal{T}_1) \cong \mathcal{O}_\Delta[\![1]\!]\{-2\} \quad \textit{and} \quad \Psi_1'(\mathcal{T}_1^{-1}) \cong \mathcal{O}_\Delta\{2\}[-1]$$

inside $\mathsf{Kom}_-^*(D(\mathbb{A}_1 \times \mathbb{A}_1))$.

Remark The key to Lemma 4.2 is the exact triangle $\mathcal{O}_S \to \mathcal{O}_\Delta \to \mathcal{O}_T[\![1]\!]\{-2\}$, where S and T are the loci inside $\mathbb{A}_{\underline{k},1,1} \times \mathbb{A}_{\underline{k},1,1}$ given by (7) on the last two strands. The argument in the proof shows in fact that for $\mathcal{P} \in \mathsf{Kom}(D(\mathbb{A}_{\underline{k},1} \times \mathbb{A}_{\underline{k},1}))$ we have an isomorphism

$$\Psi_1'((\mathcal{O}_S \to \mathcal{O}_\Delta) \star \Psi_1(\mathcal{P})) \xrightarrow{\sim} \Psi_1'(\mathcal{O}_T[\![1]\!]\{-2\} \star \Psi_1(\mathcal{P})) \tag{6}$$

inside $\mathsf{Kom}(D(\mathbb{A}_{\underline{k},1} \times \mathbb{A}_{\underline{k},1}))$.

Proof On $\mathbb{A}_{1,1} \times \mathbb{A}_{1,1}$ consider the following subvarieties:

$$\Delta := \{(x,y,x,y)\}, \quad T := \{(x,y,y,x)\} \quad \text{and} \quad S := T \cup \Delta. \tag{7}$$

Then $\mathcal{T}_1 \cong [\mathcal{O}_S \to \mathcal{O}_\Delta]$ and $\mathcal{T}_1^{-1} \cong [\mathcal{O}_\Delta \to \mathcal{O}_S\{2\}]$, where in both cases $\mathcal{O}_\Delta$ is in cohomological degree zero. The result will follow if we can show that

$$[\Psi_1'(\mathcal{O}_S) \to \Psi_1'(\mathcal{O}_\Delta)] \cong [0 \to \mathcal{O}_\Delta[\![1]\!]\{-2\}],$$
$$[\Psi_1'(\mathcal{O}_\Delta) \to \Psi_1'(\mathcal{O}_S\{2\})] \cong [0 \to \mathcal{O}_\Delta\{2\}]$$

in the homotopy category $\mathsf{Kom}_*^-(D(\mathbb{A}_1 \times \mathbb{A}_1))$.

We will prove the first assertion (the second follows similarly). Note that $S \cap T \subset T$ is the divisor cut out by $x = y$. Thus, $\mathcal{O}_T(-S \cap T) \cong \mathcal{O}_T\{-2\}$ and we have the exact triangle $\mathcal{O}_T\{-2\} \to \mathcal{O}_S \to \mathcal{O}_\Delta$. Recall that $\Psi_1'(\cdot) = \pi_* \Delta^*(\cdot)$, where π and Δ are the natural maps

$$\mathbb{A}_1 \times \mathbb{A}_1 \xleftarrow{\pi} \mathbb{A}_1 \times \mathbb{A}_1 \times \mathbb{A}_1 \xrightarrow{\Delta} \mathbb{A}_{1,1} \times \mathbb{A}_{1,1}.$$

Now, $\mathcal{O}_\Delta \in D(\mathbb{A}_{1,1} \times \mathbb{A}_{1,1})$ has a resolution

$$\mathcal{O}_{\mathbb{A}_{1,1} \times \mathbb{A}_{1,1}}\{-2\} \xrightarrow{\cdot(y_1 - y_2)} \mathcal{O}_{\mathbb{A}_{1,1} \times \mathbb{A}_{1,1}} \to \mathcal{O}_\Delta,$$

which means that $\Delta^* \mathcal{O}_\Delta \cong \mathcal{O}_{\Delta'} \oplus \mathcal{O}_{\Delta'}[\![1]\!]\{-2\}$, where $\Delta' \subset \mathbb{A}_1 \times \mathbb{A}_1 \times \mathbb{A}_1$ is the locus (x,y,x). Moreover, $\pi_*(\mathcal{O}_{\Delta'}) \cong \mathcal{O}_\Delta \otimes_{\Bbbk} \Bbbk[y] \in D(\mathbb{A}_1 \times \mathbb{A}_1)$. Hence

$$\Psi_1'(\mathcal{O}_\Delta) = \pi_* \Delta^* \mathcal{O}_\Delta \cong \pi_*(\mathcal{O}_{\Delta'} \oplus \mathcal{O}_{\Delta'}[\![1]\!]\{-2\}) \cong (\mathcal{O}_\Delta \oplus \mathcal{O}_\Delta[\![1]\!]\{-2\}) \otimes_{\Bbbk} \Bbbk[y]. \tag{8}$$

On the other hand, $\Delta^* \mathcal{O}_T \cong \mathcal{O}_{\{(x,x,x)\}}$, which means that

$$\Psi_1'(\mathcal{O}_T\{-2\}) = \pi_* \Delta^* \mathcal{O}_T\{-2\} \cong \pi_* \mathcal{O}_{\{(x,x,x)\}}\{-2\} \cong \mathcal{O}_\Delta\{-2\}.$$

Thus, the exact triangle $\Psi_1'(\mathcal{O}_S) \to \Psi_1'(\mathcal{O}_\Delta) \to \Psi_1'(\mathcal{O}_T\{-2\}[1])$ becomes

$$\Psi_1'(\mathcal{O}_S) \xrightarrow{f} \mathcal{O}_\Delta \otimes_{\Bbbk} \Bbbk[y] \oplus \mathcal{O}_\Delta[\![1]\!]\{-2\} \otimes_{\Bbbk} \Bbbk[y] \to \mathcal{O}_\Delta[\![1]\!]\{-2\}.$$

Now $\mathcal{H}^0(\Psi_1'(\mathcal{O}_S)) \cong \pi_* L^0 \Delta^* \mathcal{O}_S \cong \pi_* \mathcal{O}_{\Delta'} \cong \mathcal{O}_\Delta \otimes_{\Bbbk} \Bbbk[y]$ and f induces an isomorphism in this degree. Thus, from the long exact sequence we get

$$\mathcal{H}^*(\Psi_1'(\mathcal{O}_S)) = \begin{cases} \mathcal{O}_\Delta \otimes_{\Bbbk} \Bbbk[y] & \text{if } * = 0, \\ \mathcal{O}_\Delta\{-4\} \otimes_{\Bbbk} \Bbbk[y] & \text{if } * = -1, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

It is easy to see that on $\mathbb{A}_1 \times \mathbb{A}_1$ we have $\mathrm{End}^2(\mathcal{O}_\Delta\{\cdot\} \otimes_{\Bbbk} \Bbbk[y]) = 0$, so

$$\Psi_1'(\mathcal{O}_S) \cong (\mathcal{O}_\Delta \oplus \mathcal{O}_\Delta[\![1]\!]\{-4\}) \otimes_{\Bbbk} \Bbbk[y].$$

Hence, using a version of the Gaussian elimination lemma [5, Lemma 3.2], we combine (8) and (9) to obtain

$$[\Psi_1'(\mathcal{O}_S) \to \Psi_1'(\mathcal{O}_\Delta)] \cong [0 \to \mathcal{O}_\Delta[\![1]\!]\{-2\}]. \tag{10}$$ □

Proof of Theorem 4.1 We already know that $\beta \mapsto \mathcal{T}_\beta'$ satisfies the braid relations. It remains to check that $\tau(\mathcal{T}_{\beta_1}' \star \mathcal{T}_{\beta_2}') \cong \tau(\mathcal{T}_{\beta_2}' \star \mathcal{T}_{\beta_1}')$ and the Markov move (stabilization).

The first relation is a standard property of Hochschild homology (in fact the more general trace property $\tau(\mathcal{A} \star \mathcal{B}) \cong \tau(\mathcal{B} \star \mathcal{A})$ holds for any kernels $\mathcal{A}$, $\mathcal{B}$).

To prove the Markov move first note that since projectors $\mathcal{P}^-$ move freely through crossings it suffices to prove the Markov move when the extra strand is coloured by a partition (ℓ). In this case, for any $\mathcal{P} \in \mathsf{Kom}_*^-(D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}})$ we claim that

$$\tau(\mathcal{T}_n' \star \Psi_\ell(\mathcal{P})) \cong \tau(\mathcal{P})[-\ell][\![\ell]\!] \quad \text{and} \quad \tau((\mathcal{T}_n')^{-1} \star \Psi_\ell(\mathcal{P})) \cong \tau(\mathcal{P}). \tag{11}$$

We prove the isomorphism on the left by induction on ℓ (the right one is similar). For $\mathcal{P} \in \mathsf{Kom}_*^-(D(\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}}))$ we have the following algebraic computation:

$$\bigoplus_{[\ell]} \tau(\mathcal{T}_n' \star \Psi_\ell(\mathcal{P})) \cong \bigoplus_{[\ell]} \tau(\mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{T}_{n+2}' \star \mathcal{T}_{n+1}' \star (\Psi_0 \circ \Psi_\ell \circ \Psi_0)(\mathcal{P}))$$
$$\cong \tau(\mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{T}_{n+2}' \star \mathcal{T}_{n+1}' \star \mathcal{F}_{n+2} \star \mathcal{E}_{n+2} \star (\Psi_0 \circ \Psi_\ell \circ \Psi_0)(\mathcal{P})) \tag{12}$$
$$\cong \tau(\mathcal{F}_n \star \mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{T}_{n+2}' \star \mathcal{T}_{n+1}' \star \mathcal{E}_{n+2} \star (\Psi_0 \circ \Psi_\ell \circ \Psi_0)(\mathcal{P})) \tag{13}$$
$$\cong \tau(\mathcal{E}_{n+2} \star \mathcal{F}_n \star \mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{T}_{n+2}' \star \mathcal{T}_{n+1}' \star (\Psi_1 \circ \Psi_{\ell-1} \circ \Psi_0)(\mathcal{P})) \tag{14}$$
$$\cong \tau(\mathcal{F}_n \star \mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{T}_{n+2}' \star \mathcal{T}_{n+1}' \star \mathcal{E}_n \star (\Psi_1 \circ \Psi_{\ell-1} \circ \Psi_0)(\mathcal{P})) \tag{15}$$
$$\cong \tau(\mathcal{F}_n \star \mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{T}_{n+1}' \star \mathcal{E}_n \star (\Psi_{\ell-1} \circ \Psi_0)(\mathcal{P}))[-1][\![1]\!] \tag{16}$$
$$\cong \tau(\mathcal{F}_n \star \mathcal{T}_n' \star \mathcal{T}_{n+1}' \star \mathcal{T}_n' \star \mathcal{E}_n \star (\Psi_{\ell-1} \circ \Psi_0)(\mathcal{P}))[-1][\![1]\!] \tag{17}$$
$$\cong \tau(\mathcal{F}_n \star \mathcal{T}_n' \star \mathcal{T}_n' \star \mathcal{E}_n \star \Psi_0(\mathcal{P}))[-\ell][\![\ell]\!] \tag{18}$$
$$\cong \tau(\mathcal{T}_n' \star \mathcal{E}_n \star \mathcal{F}_n \star \mathcal{T}_n' \star \Psi_0(\mathcal{P}))[-\ell][\![\ell]\!] \tag{19}$$
$$\cong \bigoplus_{[\ell]} \tau(\mathcal{P})[-\ell][\![\ell]\!]. \tag{20}$$

Here we

- added two strands labeled 0 to obtain the first isomorphism,
- used (2) twice to obtain (13),
- used the Markov relation to obtain (14),
- used that

$$\begin{aligned}\mathcal{E}_{n+2} \star \mathcal{F}_n \star \mathcal{T}'_{n+1} \star \mathcal{T}'_n \star \mathcal{T}'_{n+2} \star \mathcal{T}'_{n+1} &\cong \mathcal{F}_n \star \mathcal{E}_{n+2} \star \mathcal{T}'_{n+1} \star \mathcal{T}'_{n+2} \star \mathcal{T}'_n \star \mathcal{T}'_{n+1} \\ &\cong \mathcal{F}_n \star \mathcal{T}'_{n+1} \star \mathcal{T}'_{n+2} \star \mathcal{T}'_n \star \mathcal{T}'_{n+1} \star \mathcal{E}_n\end{aligned}$$

  to get (15),
- used (11) with Ψ_1 to obtain (16) and with $\Psi_{\ell-1}$ to obtain (18),
- applied (3) twice to obtain (20), and
- used that $(\mathcal{T}'_i)^2$ is the identity if one of the strands it acts on is labeled 0 to get (20).

Thus, (11) follows by induction if we can prove the base case $\ell = 1$. In this case we have

$$\tau(\mathcal{T}'_n \star \Psi_1(\mathcal{P})) \cong \tau(\Psi'_1(\mathcal{T}'_n \star \Psi_1(\mathcal{P}))) \cong \tau(\Psi'_1(\mathcal{T}'_n) \star \mathcal{P}),$$

so it suffices to show that $\Psi'_1(\mathcal{T}'_n) \cong \mathcal{O}_\Delta[-1][\![1]\!]$. This follows from Lemma 4.2 (since $\mathcal{T}'_n = \mathcal{T}_n[-1]\{2\}$ in this case). □

Remark For those familiar with webs (see for instance [9]) the algebraic computation above can be summarized as follows. First break up the strand labeled ℓ and then use that "trivalent vertices" move naturally through crossings together with the Markov move. Figure 1 illustrates this procedure, where the box denotes an arbitrary braid (we simplify by omitting the closure of each diagram).

5 K-theory

Recall that to a link L whose strands are coloured by partitions one can associate the coloured HOMFLY polynomial $P_L(q,a) \in \Bbbk(q,a)$. We now explain why the invariant from Theorem 4.1 categorifies the coloured HOMFLY polynomial. This is normalized so that if $L = \bigcirc_{(k)}$ (the unknot labeled by k) then

$$P_L(q,a) = \prod_{\ell=1}^{k} \frac{aq^{-\ell+1} - a^{-1}q^{\ell-1}}{q^{-\ell} - q^{\ell}}. \tag{21}$$

Remark For notational convenience we use the transpose notation, meaning that what we call λ would normally be the transpose partition. For example, our partition (k) would be instead (1^k) (and vice versa).

On the other hand, we can consider the Poincaré polynomial $\mathcal{P}_L(q,a,t)$ of $\mathcal{H}(L)$ from Theorem 4.1. Here the shifts $\{1\}, [\![1]\!]$ and $[1]$ are kept track of by formal variables $q, -a^2, t$ respectively.

Proposition 5.1 *For a coloured link L, the invariants $P_L(q,a)$ and $\mathcal{P}_L(q,aq,-1)$ agree, up to an explicit factor $a^m q^n$.*

Proof In the rest of the proof we will ignore extra factors $a^m q^n$. Let us first suppose that L is the closure of a braid β coloured by partitions (k) with only one part. One can compute $P_L(q,a)$ from β by applying a trace. Moreover, as explained (for instance) in [9, Section 6] one can break down the crossings in L into web diagrams since the crossing element is a linear combination of webs. This reduces the evaluation of $P_L(q,a)$ to evaluating this trace on diagrams.

As usual, one views the trace of a web diagram as the closure of that diagram on the annulus. The algebra of webs on the annulus is generated (as an algebra) by unknots labeled by one-part partitions (where multiplication is given by gluing one annulus inside the other). This reduces the computation of $P_L(q,a)$ to the case $L = \bigcirc_{(k)}$ (which is described in (21)).

Similarly, the evaluation of $\mathcal{P}_L(q,a,-1)$ can be reduced to the case $L = \bigcirc_{(k)}$. This case is computed in Section 7.1 and agrees with (21) once you replace a with aq. This completes the proof when L contains partitions with only one part.

To deal with arbitrary partitions we will show that $P_L(q,q^N) = \mathcal{P}_L(q,q^{N+1},-1)$ for all $N > 0$ (ie the specializations $a = q^N$ for all $N > 0$). Note that $P_L(q,q^N)$

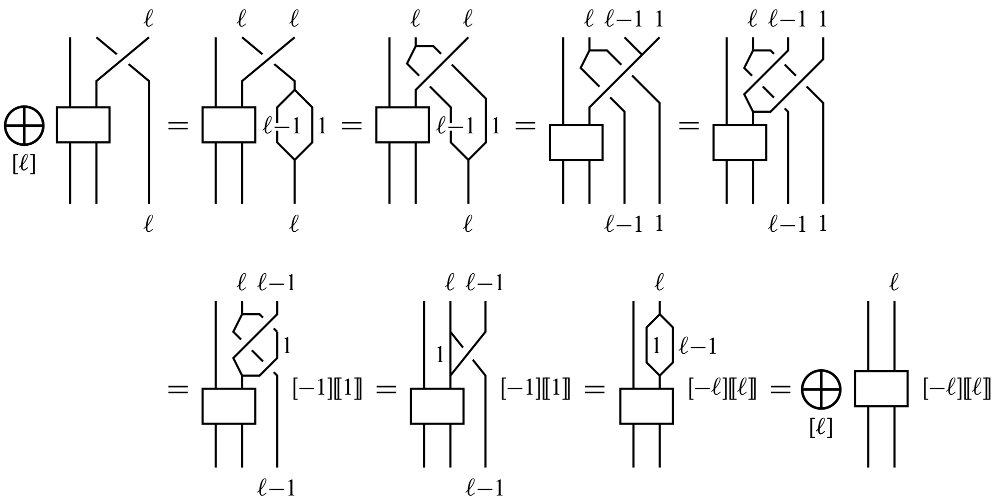


Figure 1: The Markov move involving a strand labeled ℓ

recovers the corresponding $\mathrm{SL}(N)$ Reshetikhin–Turaev (RT) invariant and that we know $P_L(q,q^N) = \mathcal{P}_L(q,q^{N+1},-1)$ if L is coloured by one-part partitions. On the other hand, in [5] we showed that, when evaluating RT invariants, the projectors (clasps) for arbitrary partitions can be constructed as infinite twists. Since this construction only uses the braid group action it follows that $P_L(q,q^N) = \mathcal{P}_L(q,q^{N+1},-1)$ holds for any L. □

6 Some differentials

To simplify notation we will omit the $\{\cdot\}$ grading in this section. We also fix $N > 0$. Note that

$$\mathrm{HH}^1(\mathbb{A}^k) = \mathrm{Ext}^1_{\mathbb{A}^k\times\mathbb{A}^k}(\Delta_*\mathcal{O}_{\mathbb{A}^k}, \Delta_*\mathcal{O}_{\mathbb{A}^k}) \cong \bigoplus_i \Bbbk[x_1,\dots,x_k]\partial_{x_i},$$

so that

$$\gamma_{1^k} := \sum_i x_i^N \partial_{x_i} \in \mathrm{Hom}_{\mathbb{A}^k\times\mathbb{A}^k}(\Delta_*\mathcal{O}_{\mathbb{A}^k}, \Delta_*\mathcal{O}_{\mathbb{A}^k}[1]).$$

Since this element is S_k–invariant it descends to $\mathrm{HH}^1(\Bbbk[e_1,\dots,e_k])$. We denote by

$$\gamma_{\underline{k}} \in \mathrm{Hom}_{\mathbb{A}_{\underline{k}}\times\mathbb{A}_{\underline{k}}}(\Delta_*\mathcal{O}_{\mathbb{A}_{\underline{k}}}, \Delta_*\mathcal{O}_{\mathbb{A}_{\underline{k}}}[1])$$

the corresponding element obtained from $\gamma_{1^{k_1}} \otimes \cdots \otimes \gamma_{1^{k_n}}$ by descent.

Now, given a braid β with endpoints marked $\underline{k}$, we have

$$\tau(\mathcal{T}'_\beta) = \mathrm{HH}_*(\mathcal{T}'_\beta) \cong \mathrm{Ext}^*_{\mathbb{A}_{\underline{k}}\times\mathbb{A}_{\underline{k}}}(\Delta_*\omega^{-1}_{\mathbb{A}_{\underline{k}}}[\![-\dim\mathbb{A}_{\underline{k}}]\!], \mathcal{T}'_\beta),$$

where $\omega_{\mathbb{A}_{\underline{k}}}$ denotes the canonical bundle. Thus, we have an action of $\mathrm{HH}^1(\mathbb{A}_{\underline{k}})$ coming from precomposing on the left with

$$\mathrm{HH}^1(\mathbb{A}_{\underline{k}}) \cong \mathrm{Hom}_{\mathbb{A}_{\underline{k}}\times\mathbb{A}_{\underline{k}}}(\Delta_*\omega^{-1}_{\mathbb{A}_{\underline{k}}}, \Delta_*\omega^{-1}_{\mathbb{A}_{\underline{k}}}[1]).$$

We denote by d_N the action of $\gamma_{\underline{k}}$. Note that $d_N^2 = 0$ since $\gamma_{\underline{k}}$ belongs to HH^1. Moreover, d_N commutes with the differential d used in the definition of the complex $\mathcal{T}'_\beta$ because composition is associative. Thus, we get a bicomplex with differentials d and d_N.

Theorem 6.1 *Suppose $L = \widehat{\beta}$, where β is a coloured braid. If we denote by $\mathcal{H}_N(L)$ the cohomology of $\tau(\mathcal{T}'_\beta)$ equipped with the total differential $d + d_N$, then, up to an overall grading shift, $\mathcal{H}_N(L)$ defines a doubly graded link invariant.*

In the remainder of this section we prove this result. Sometimes we will write $\mathcal{H}_N(\mathcal{T}'_\beta)$ instead of $\mathcal{H}_N(L)$, where $L = \widehat{\beta}$.

If β and β' are equivalent braids, then $\mathcal{T}'_\beta$ is homotopic to $\mathcal{T}'_{\beta'}$, which means that $\mathcal{H}_N(\mathcal{T}'_\beta) \cong \mathcal{H}_N(\mathcal{T}'_{\beta'})$. Next, to prove invariance under conjugation, we must show that $\mathcal{H}_N(\mathcal{T}'_{\beta_1} \star \mathcal{T}'_{\beta_2}) \cong \mathcal{H}_N(\mathcal{T}'_{\beta_2} \star \mathcal{T}'_{\beta_1})$ for any braids β_1 and β_2. This follows as in the proof of Theorem 4.1 together with the fact that for any braid β we have

$$II\gamma_{\underline{k}} = \gamma_{\underline{k}'} II \in \mathrm{Hom}(\Delta_* \mathcal{O}_{\mathbb{A}_{\underline{k}'}} \star \mathcal{T}'_\beta \star \Delta_* \mathcal{O}_{\mathbb{A}_{\underline{k}}}, \Delta_* \mathcal{O}_{\mathbb{A}_{\underline{k}'}} \star \mathcal{T}'_\beta \star \Delta_* \mathcal{O}_{\mathbb{A}_{\underline{k}}}[1]),$$

where $\underline{k}$ and $\underline{k}'$ label the bottom and top strands of β (this equality follows from Lemma 6.2).

Remark Here we use the convention mentioned at the end of Section 2.1. For instance, $II\gamma_{\underline{k}}$ denotes the map induced by the identity on the first two factors of $\Delta_* \mathcal{O}_{\mathbb{A}_{\underline{k}'}} \star \mathcal{T}'_\beta \star \Delta_* \mathcal{O}_{\mathbb{A}_{\underline{k}}}$ and by $\gamma_{\underline{k}}$ on the last (right) one.

Lemma 6.2 *Consider $\mathcal{E} \in D(\mathbb{A}_{i-1,j+1} \times \mathbb{A}_{i,j})$. Then*

$$II\gamma_{i,j} = \gamma_{i-1,j+1} II$$

in $\mathrm{Hom}(\Delta_* \mathcal{O}_{\mathbb{A}_{i-1,j+1}} \star \mathcal{E} \star \Delta_* \mathcal{O}_{\mathbb{A}_{i,j}}, \Delta_* \mathcal{O}_{\mathbb{A}_{i-1,j+1}} \star \mathcal{E} \star \Delta_* \mathcal{O}_{\mathbb{A}_{i,j}}[1])$, and likewise if we replace $\mathcal{E}$ with $\mathcal{F}$.

Proof $\mathcal{E}$ is the kernel inducing the correspondence

$$\mathbb{A}_i \times \mathbb{A}_j \xleftarrow{\pi_1} \mathbb{A}_{i-1} \times \mathbb{A}^1 \times \mathbb{A}_j \xrightarrow{\pi_2} \mathbb{A}_{i-1} \times \mathbb{A}_{j+1}.$$

On the other hand, $II\gamma_{i,j}$ is the element obtained from $\gamma_{1^{i+j}}$ by descent along the map

$$\mathbb{A}^{i+j} \to \mathbb{A}_{i-1} \times \mathbb{A}^1 \times \mathbb{A}_j \xrightarrow{\pi_1} \mathbb{A}_i \times \mathbb{A}_j.$$

Likewise $\gamma_{i-1,j+1} II$ is the map obtained from $\gamma_{1^{i+j}}$ by descent along the map

$$\mathbb{A}^{i+j} \to \mathbb{A}_{i-1} \times \mathbb{A}^1 \times \mathbb{A}_j \xrightarrow{\pi_2} \mathbb{A}_{i-1} \times \mathbb{A}_{j+1}.$$

The result follows. $\square$

Finally, we need invariance under the Markov move. As in the proof of Theorem 4.1, we can significantly reduce what we must show. First, since projectors pass through crossings, we can assume each strand is coloured by a partition (ℓ) with only one part. By breaking up this strand into ℓ strands coloured by 1 and using Lemma 6.2 we can further reduce to the case $\ell = 1$.

Using the homotopy equivalence from (6) we know that

$$\begin{aligned}\mathcal{H}_N(\Psi_1'(\mathcal{T}_n' \star \Psi_1(\mathcal{P}))) = \mathcal{H}_N(\Psi_1'((\mathcal{O}_S \to \mathcal{O}_\Delta) \star \Psi_1(\mathcal{P}))) \\ \xrightarrow{\sim} \mathcal{H}_{-N}(\Psi_1'(\mathcal{O}_T[\![1]\!]\{-2\} \star \Psi_1(\mathcal{P}))),\end{aligned}$$

where $\mathcal{P} \in \mathsf{Kom}(D(\mathbb{A}_{\underline{k},1} \times \mathbb{A}_{\underline{k},1}))$. Recall that we have the standard exact sequence

$$\mathcal{O}_T\{-2\} \to \mathcal{O}_S \to \mathcal{O}_\Delta,$$

where S, T are the varieties corresponding to the last two strands. Moreover,

$$\Psi_1'(\mathcal{O}_T \star \Psi_1(\mathcal{P})) \xrightarrow{\sim} \mathcal{P}$$

and Lemma 6.3 implies that $\mathcal{H}_N(\mathcal{T}_n' \star \Psi_1(\mathcal{P})) \cong \mathcal{H}_N(\mathcal{P})$ (up to a grading shift). This completes the proof of Theorem 6.1.

Lemma 6.3 *For $\mathcal{P} \in \mathsf{Kom}(D(\mathbb{A}_{\underline{k},1} \times \mathbb{A}_{\underline{k},1}))$, the diagram*

$$\begin{tikzcd}
\mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1,1} \times \mathbb{A}_{\underline{k},1,1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1,1}}, \mathcal{O}_T \star \Psi_1(\mathcal{P})) \arrow[r, "\gamma_{\underline{k},1,1}"] \arrow[d, "\varphi"] & \mathrm{Ext}^{j+1}_{\mathbb{A}_{\underline{k},1,1} \times \mathbb{A}_{\underline{k},1,1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1,1}}, \mathcal{O}_T \star \Psi_1(\mathcal{P})) \arrow[d, "\varphi'"] \\
\mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1} \times \mathbb{A}_{\underline{k},1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}, \mathcal{P}) \arrow[r, "\gamma_{\underline{k},1}"] & \mathrm{Ext}^{j+1}_{\mathbb{A}_{\underline{k},1} \times \mathbb{A}_{\underline{k},1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}, \mathcal{P})
\end{tikzcd}$$

commutes, where $S_X^{-1} := \omega_X^{-1}[\![-\dim X]\!]$ for a variety X and where isomorphisms φ and φ' are induced by the isomorphism $\Psi_1'(\mathcal{O}_T \star \Psi_1(\mathcal{P})) \xrightarrow{\sim} \mathcal{P}$.

Proof As before, we will ignore shifts in $\{\cdot\}$. The left adjoint of $\Psi_1' \colon \mathbb{A}_{\underline{k},1} \to \mathbb{A}_{\underline{k}}$ is the functor

$$(\Psi_1')^L(\cdot) = \Delta_*(\pi^*(\cdot) \otimes p^* S^{-1}_{\mathbb{A}^1}),$$

where $p \colon \mathbb{A}_{\underline{k}} \times \mathbb{A}^1 \times \mathbb{A}_{\underline{k}} \to \mathbb{A}^1$ is the projection. Now take

$$\alpha \in \mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1,1} \times \mathbb{A}_{\underline{k},1,1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1,1}}, \mathcal{O}_T \star \Psi_1(\mathcal{P}))$$

and consider the following diagram:

$$\begin{tikzcd}
\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}[\![-1]\!] \arrow[r, "\gamma_{\underline{k},1}"] \arrow[d, "\mathrm{adj}"] & \Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \arrow[r, "\varphi(\alpha)"] \arrow[d, "\mathrm{adj}"] & \Psi_1'(\mathcal{O}_T) \star \mathcal{P}[\![j]\!] \\
\Psi_1'(\Psi_1')^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}})[\![-1]\!] \arrow[r, "\Psi_1'(\Psi_1')^L(\gamma_{\underline{k},1})"] & \Psi_1'(\Psi_1')^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}) \arrow[r, "\Psi_1'(\alpha)"] & \Psi_1'(\mathcal{O}_T \star \Psi_1(\mathcal{P}))[\![j]\!] \arrow[u, "\sim"]
\end{tikzcd}$$

Note that $(\Psi'_1)^L(S^{-1}_{\mathbb{A}_{\underline{k},1}}) = S^{-1}_{\mathbb{A}_{\underline{k},1,1}}$, which explains how $\Psi'_1(\alpha)$ acts. The left square commutes since adjunction is a natural transformation. The square on the right commutes by the definition of φ.

The composition along the top is the map $\alpha \mapsto \varphi(\alpha) \circ \gamma_{\underline{k},1}$. On the other hand, the composition along the bottom row and up the right side is $\alpha \mapsto \varphi'(\alpha \circ (\Psi'_1)^L(\gamma_{\underline{k},1}))$. So it suffices to show that

$$\varphi'(\alpha \circ (\Psi'_1)^L(\gamma_{\underline{k},1})) = \varphi'(\alpha \circ \gamma_{\underline{k},1,1}).$$

The difference $\gamma_{\underline{k},1,1} - (\Psi'_1)^L(\gamma_{\underline{k},1})$ is equal to $x^N_{n+1}\partial_{x_{n+1}}$, so it remains to show that $\varphi'(\alpha \circ x^N_{n+1}\partial_{x_{n+1}})$ vanishes.

The map $\varphi'(\alpha \circ x^N_{n+1}\partial_{x_{n+1}})$ is given by the composition

$$\begin{aligned}\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \xrightarrow{\text{adj}} \Psi'_1(\Psi'_1)^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}) &= \Psi'_1(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1,1}}) \\ &\xrightarrow{\Psi'_1(x^N_{n+1}\partial_{x_{n+1}})} \Psi'_1(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1,1}}[\![1]\!]) \\ &\xrightarrow{\Psi'_1(\alpha)} \Psi'_1(\mathcal{O}_T \star \Psi_1(\mathcal{P}))[\![j+1]\!] \xrightarrow{\sim} \mathcal{P}[\![j+1]\!].\end{aligned}$$

One can check that

$$\Psi'_1(\Psi'_1)^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}) \cong \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\big) \oplus \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}[\![-1]\!]\big).$$

Then the composition of the first two maps is given by

$$\begin{aligned}\text{(22)} \quad \Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} &\to \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}[\![-1]\!]\big) \oplus \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\big) \\ &\to \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}\big) \oplus \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}][\![1]\!]\big),\end{aligned}$$

where, considering the direct sums as column vectors, the maps are respectively

$$\begin{pmatrix} 0 \\ \mathrm{id} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & \cdot x^N_{n+1}\partial_{x_{n+1}} \\ 0 & 0 \end{pmatrix}.$$

On the other hand, to understand $\Psi'_1(\alpha)$, consider the isomorphisms

$$\begin{aligned}&\mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1,1}\times\mathbb{A}_{\underline{k},1,1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1,1}}, \mathcal{O}_T \star \Psi_1(\mathcal{P})) \\ &\qquad \cong \mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1,1}\times\mathbb{A}_{\underline{k},1,1}}((\Psi'_1)^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}), \mathcal{O}_T \star \Psi_1(\mathcal{P})) \\ &\qquad \cong \mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1}\times\mathbb{A}_{\underline{k},1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}, \Psi'_1(\mathcal{O}_T \star \Psi_1(\mathcal{P}))) \\ &\qquad \cong \mathrm{Ext}^j_{\mathbb{A}_{\underline{k},1}\times\mathbb{A}_{\underline{k},1}}(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}, \mathcal{P}).\end{aligned}$$

The image of $\beta\colon \Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \to \mathcal{P}[\![j]\!]$ under these isomorphisms is the composition

$$(\Psi_1')^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}}) \xrightarrow{(\Psi_1')^L(\beta)} (\Psi_1')^L(\mathcal{P})[\![j]\!] = \Delta_* p^*(S^{-1}_{\mathbb{A}^1}) \star \Psi_1(\mathcal{P})[\![j]\!] \xrightarrow{h} (\mathcal{O}_T \star \Psi_1(\mathcal{P}))[\![j]\!],$$

where $p\colon \mathbb{A}_{\underline{k},1,1} \to \mathbb{A}^1$ projects onto the last factor. Here h is induced by the map $\Delta_* p^*(S^{-1}_{\mathbb{A}^1}) \to \mathcal{O}_T$, which comes from the standard exact sequence

$$\mathcal{O}_T\{-2\} \to \mathcal{O}_S \to \mathcal{O}_\Delta$$

after noting that $\Delta_* p^*(S^{-1}_{\mathbb{A}^1}) \cong \mathcal{O}_\Delta[\![-1]\!]\{2\}$. Thus, we can assume that α is such a composition for some β. Applying Ψ_1', we find that $\Psi_1'(\alpha)$ factors as

$$\begin{aligned}&\big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}[\![-1]\!]\big) \oplus \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\big)\\ &\quad\to \big(\mathcal{P} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}[\![j-1]\!]\big) \oplus \big(\mathcal{P} \otimes_{\Bbbk} \Bbbk[x_{n+1}][\![j]\!]\big) \to \Psi_1'(\mathcal{O}_T \star \Psi_1(\mathcal{P}))[\![j]\!],\end{aligned}$$

where, in matrix form, the maps are respectively

$$\begin{pmatrix} \Psi_1'(\Psi_1')^L(\beta) & 0 \\ 0 & \Psi_1'(\Psi_1')^L(\beta)\end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & \Psi_1'(h)\end{pmatrix}.$$

Finally, the composition of $\Psi_1'(h)$ with the isomorphism $\Psi_1'(\mathcal{O}_T \star \Psi_1(\mathcal{P}))[\![j]\!] \xrightarrow{\sim} \mathcal{P}[\![j]\!]$ gives a map which is zero on the summand $\mathcal{P} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}[\![j-1]\!]$ and the natural projection map $\mathcal{P} \otimes_{\Bbbk} \Bbbk[x_{n+1}][\![j]\!] \to \mathcal{P}[\![j]\!]$ on the second summand (which sends x_{n+1} to zero). This fact can be traced back to the map

$$\Delta_* \mathcal{O}_{\mathbb{A}_1} \otimes_{\Bbbk} \Bbbk[x_2][\![-1]\!] \oplus \Delta_* \mathcal{O}_{\mathbb{A}_1} \otimes_{\Bbbk} \Bbbk[x_2] = \Psi_1'(\Delta_* \mathcal{O}_{\mathbb{A}_{1,1}}) \to \Psi_1'(\mathcal{O}_T) \cong \Delta_* \mathcal{O}_{\mathbb{A}_1},$$

which, as we saw in the proof of Lemma 4.2, acts by zero on the first summand and by the natural projection map on the second summand. In conclusion, the composition

$$\Psi_1'(\Psi_1')^L(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}})[\![1]\!] \xrightarrow{\Psi_1'(\alpha)} \Psi_1'(\mathcal{O}_T \star \Psi_1(\mathcal{P}))[\![j+1]\!] \xrightarrow{\sim} \mathcal{P}[\![j+1]\!]$$

is isomorphic to the composition

$$\begin{aligned}&\big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}\big) \oplus \big(\Delta_* S^{-1}_{\mathbb{A}_{\underline{k},1}} \otimes_{\Bbbk} \Bbbk[x_{n+1}][\![1]\!]\big)\\ &\quad\to \big(\mathcal{P} \otimes_{\Bbbk} \Bbbk[x_{n+1}]\partial_{x_{n+1}}[\![j]\!]\big) \oplus \big(\mathcal{P} \otimes_{\Bbbk} \Bbbk[x_{n+1}][\![j+1]\!]\big) \to \mathcal{P}[\![j+1]\!],\end{aligned}$$

where, in matrix form, the maps are respectively

$$\begin{pmatrix} \Psi_1'(\Psi_1')^L(\beta) & 0 \\ 0 & \Psi_1'(\Psi_1')^L(\beta)\end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & \pi\end{pmatrix},$$

and where π is the natural projection map from the second summand. The composition of this with (22) is clearly zero and hence $\varphi'(\alpha \circ x_{n+1}^N \partial_{x_{n+1}}) = 0$. □

Remark In the proof of Lemma 6.3 above, we used the observation that the difference $\gamma_{\underline{k},1,1} - \Psi_1(\gamma_{\underline{k},1})$ is of the form $f\partial_{x_{n+1}}$ for some function f on $\mathbb{A}_{\underline{k},1,1}$.

7 Examples

For a partition $\underline{k} = (k_1, \dots, k_n)$ we denote by $\mathcal{H}(\bigcirc_{\underline{k}})$ the triply graded homology of the unknot labeled by $\underline{k}$. We will compute this invariant when $\underline{k} = (k)$ and $\underline{k} = (1^2)$. Its Poincaré polynomial is denoted by $\mathcal{P}_{\bigcirc_{\underline{k}}}(q,a,t)$, where the shifts $\{1\}$, $[\![1]\!]$ and $[1]$ are kept track of by q, $-a^2$ and t, respectively.

7.1 Cohomology of $\bigcirc_{(k)}$

If $k = 1$ we have

$$\mathcal{H}(\bigcirc_{(1)}) \cong \pi_*\Delta^*(\mathcal{O}_\Delta) \cong \pi_*(\mathcal{O}_{\mathbb{A}_1} \oplus \mathcal{O}_{\mathbb{A}^1}[\![1]\!]\{-2\}) \cong \Bbbk[x] \oplus \Bbbk[x][\![1]\!]\{-2\},$$

where Δ and π are the natural maps $\bullet \xleftarrow{\pi} \mathbb{A}_1 \xrightarrow{\Delta} \mathbb{A}_1 \times \mathbb{A}_1$. Hence

$$\mathcal{P}_{\bigcirc_{(1)}}(q,a,t) = (1 + q^{-2} + q^{-4} + \cdots)(1 - a^2q^{-2}) = \frac{1 - a^2q^{-2}}{1 - q^{-2}}.$$

Note that $\Bbbk[x] \cong \bigoplus_{i \geq 0} \Bbbk\{-2i\}$, which explains why it contributes $(1 + q^{-2} + q^{-4} + \cdots)$. More generally, $\mathbb{A}_k = \operatorname{Spec} \Bbbk[e_1, \dots, e_k]$, and a similar argument shows that

$$\mathcal{H}(\bigcirc_{(k)}) \cong \bigotimes_{\ell=1}^{k} (\Bbbk[e_\ell] \oplus \Bbbk[e_\ell][\![1]\!]\{-2\ell\}).$$

It follows that

$$\mathcal{P}_{\bigcirc_{(k)}}(q,a,t) = \prod_{\ell=1}^{k} \frac{1 - a^2q^{-2\ell}}{1 - q^{-2\ell}}. \tag{23}$$

7.2 Cohomology of $\bigcirc_{(1^2)}$

We need to explicitly identify the projector P^-, which lives in $\mathsf{Kom}_*^-(D(\mathbb{A}_{1,1} \times \mathbb{A}_{1,1}))$. The braid element in this case is isomorphic to

$$\mathsf{T} = [\mathsf{EF}\langle -1\rangle \to \mathrm{id}] \cong [\mathcal{O}_S \to \mathcal{O}_\Delta],$$

where S is the variety described in the proof of Lemma 4.2. If (x, y) are the coordinates of $\mathbb{A}_{1,1}$, then

$$\mathsf{T} = \big[\Bbbk[x,y] \otimes_{\Bbbk[e_1,e_2]} \Bbbk[x,y] \to \Bbbk[x,y]\big]$$

as $\Bbbk[x,y]$–bimodules (where $e_1 = x+y$ and $e_2 = xy$ are the usual elementary symmetric functions). Now, squaring and simplifying gives

$$\begin{aligned}
\mathsf{T}^2 &\cong [\mathsf{EFEF}\langle -2\rangle \to \mathsf{EF}\langle -1\rangle \oplus \mathsf{EF}\langle -1\rangle \to \mathrm{id}] \\
&\cong \Big[\mathsf{EF}\langle -3\rangle \xrightarrow{\uparrow^{\bullet}\downarrow - \uparrow\downarrow^{\bullet}} \mathsf{EF}\langle -1\rangle \xrightarrow{\frown} \mathrm{id}\Big] \\
&\cong [\mathcal{O}_S\{2\} \xrightarrow{x\otimes 1 - 1\otimes x} \mathcal{O}_S \to \mathcal{O}_\Delta].
\end{aligned}$$

The maps in the first and second lines above are encoded using the diagrammatics of [15]. The isomorphism between the first and second lines was proved in [5, Section 10.2]. The isomorphism between the second and third lines follows from the fact that $\uparrow^{\bullet}\downarrow$ corresponds to $x\otimes 1$ and $\uparrow\downarrow^{\bullet}$ to $1\otimes x$ (this follows from the action of the nilHecke defined in [15] or indirectly from the main result in [4]). Now, if we multiply again by T we get

$$\begin{aligned}
\mathsf{T}^3 &\cong [\mathsf{EFEF}\langle -4\rangle \to \mathsf{EF}\langle -3\rangle \oplus \mathsf{EFEF}\langle -2\rangle \to \mathsf{EF}\langle -1\rangle \oplus \mathsf{EF}\langle -1\rangle \to \mathrm{id}] \\
&\cong \Big[\mathsf{EF}\langle -5\rangle \xrightarrow{\uparrow^{\bullet}\downarrow + \uparrow\downarrow^{\bullet} - \uparrow \circlearrowleft^{\bullet 2} \downarrow} \mathsf{EF}\langle -3\rangle \xrightarrow{\uparrow^{\bullet}\downarrow - \uparrow\downarrow^{\bullet}} \mathsf{EF}\langle -1\rangle \xrightarrow{\frown} \mathrm{id}\Big] \\
&\cong [\mathcal{O}_S\{-4\} \xrightarrow{x\otimes 1 - 1\otimes y} \mathcal{O}_S\{-2\} \xrightarrow{x\otimes 1 - 1\otimes x} \mathcal{O}_S \to \mathcal{O}_\Delta].
\end{aligned}$$

Here the 2 beside the dot in the second line indicates that we add two dots. The second line follows again from [5, Section 10.2] while the third isomorphism holds because

$$\uparrow \circlearrowleft^{\bullet 2} \downarrow \quad \text{is given by} \quad 1\otimes 1 \mapsto (x+y)\otimes 1 = 1\otimes (x+y).$$

Continuing this way, one finds that

$$\mathsf{P}^- = \lim_{\ell\to\infty} \mathsf{T}^\ell = \big[\cdots \xrightarrow{g} \mathcal{O}_S\{-6\} \xrightarrow{f} \mathcal{O}_S\{-4\} \xrightarrow{g} \mathcal{O}_S\{-2\} \xrightarrow{f} \mathcal{O}_S \to \mathcal{O}_\Delta\big], \tag{24}$$

where the maps alternate between $f = x\otimes 1 - 1\otimes x$ and $g = x\otimes 1 - 1\otimes y$.

We need to compute $\mathcal{H}(\bigcirc_{(1^2)}) = \Psi_1'\Psi_1'(\mathsf{P}^-)$. Now, using (10) and arguing as in the proof of Lemma 4.2, we find that $\Psi_1'(\mathsf{P}^-)$ is isomorphic to the complex

$$\begin{array}{ccccccc}
\cdots \xrightarrow{0} \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-4\} & \xrightarrow{1\mapsto x-y} & \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-2\} & \xrightarrow{0} & \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y] & \xrightarrow{\sim} & \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y] \\
\oplus & & \oplus & & \oplus & & \oplus \\
\cdots \xrightarrow{0} \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-8\} & \xrightarrow{1\mapsto x-y} & \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-6\} & \xrightarrow{0} & \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-4\} & \xrightarrow{1\mapsto y} & \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-2\}
\end{array}$$

where, going to the left, the differentials alternate. Now, consider the exact triangle

$$\mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y]\{-2\} \xrightarrow{1\mapsto x-y} \mathcal{O}_\Delta \otimes_\Bbbk \Bbbk[y] \to \mathcal{O}_\Delta.$$

Applying Ψ'_1 leaves us with

$$\Psi'_1(\mathcal{O}_\Delta \otimes_{\Bbbk} \Bbbk[y]\{-2\}) \to \Psi'_1(\mathcal{O}_\Delta \otimes_{\Bbbk} \Bbbk[y]) \to \Bbbk[x][\![1]\!]\{-2\} \oplus \Bbbk[x].$$

Thus, applying $\Psi'_1(\cdot)$ to $\Psi'_1(\mathsf{P}^-)$ gives us a complex isomorphic to

$$\begin{array}{ccccccc}
\cdots & 0 & \Bbbk[x]\{-6\} & 0 & \Bbbk[x]\{-2\} & 0 & 0 \\
\cdots & 0 & \Bbbk[x]\{-8,-10\} & 0 & \Bbbk[x]\{-4,-6\} & 0 & \Bbbk[x]\{-2\} \\
\cdots & 0 & \Bbbk[x]\{-12\} & 0 & \Bbbk[x]\{-8\} & 0 & \Bbbk[x]\{-4\}
\end{array}$$

where the top right entry is in cohomology bidegree $(0,0)$. The generating series is then

$$\begin{aligned}
\mathcal{P}_{\bigcirc_{(1^2)}}(q,a,t) &= \frac{1}{1-q^{-2}}\frac{1}{1-q^{-4}t^2}(q^{-2}t^2 - q^{-2}a^2 - q^{-4}a^2t^2 + q^{-4}a^4) \\
&= \frac{q^{-2}t^2(1-q^{-2}a^2)(1-a^2t^{-2})}{(1-q^{-2})(1-q^{-4}t^2)}.
\end{aligned}$$

8 Some remarks and speculation

8.1 SL(N)–homologies

In order to make the differential d_N homogeneous one needs to kill the $[\![\cdot]\!]$ grading. More precisely, one needs to set $[\![-1]\!] = \{-2(-N+1)\}$. Since $[\![1]\!]$ is recorded by $-a^2$ and $\{1\}$ by q this means that the Euler characteristic $\chi_N(L)$ of $\mathcal{H}_N(L)$ satisfies $\chi_N(L) = \mathcal{P}_L(q, iq^{-N+1}, -1)$. But

$$\mathcal{P}_{\bigcirc_{(k)}}(q, iq^{-N+1}, -1) = \prod_{\ell=1}^{k} \frac{1-q^{-2N+2-2\ell}}{1-q^{-2\ell}},$$

which, up to sign and a factor of q, equals $\left[\begin{smallmatrix} N+k-1 \\ k \end{smallmatrix}\right]$. In particular, this means that if L is a link labeled by (k) then $\mathcal{H}_N(L)$ categorifies the RT invariant of SL(N) labeled by the representation $\mathrm{Sym}^k(\mathbb{C}^N)$. Moreover, the homology of the unknot in this case can be shown to be finite-dimensional homology. This implies (using conjugation-invariance of the homology) that $\mathcal{H}_N(L)$ is finite-dimensional for any L labeled by partitions with only one part.

8.2 Batalin–Vilkovisky structures

In Section 6 we defined the differential d_N for $N > 0$. This was based on the fact that $\mathrm{HH}^*(\mathcal{A})$ acts on $\mathrm{HH}_*(\mathcal{M})$ for any algebra $\mathcal{A}$ and $\mathcal{A}$–bimodule $\mathcal{M}$. More generally, under fairly general hypotheses described in [18, Section 1], $\mathrm{HH}^*(\mathcal{A})$ is a Gerstenhaber algebra and $\mathrm{HH}_*(\mathcal{M})$ is a Batalin–Vilkovisky (BV) module.

Without going into details (see [18] for more) this equips $\mathrm{HH}^*(\mathcal{A})$ with the usual cup product as well as a graded Lie algebra structure

$$\{\cdot,\cdot\}\colon \mathrm{HH}^{p+1}(\mathcal{A}) \otimes_{\Bbbk} \mathrm{HH}^{q+1}(\mathcal{A}) \to \mathrm{HH}^{p+q+1}(\mathcal{A}),$$

while $\mathrm{HH}_*(\mathcal{M})$ carries the standard module structure as well as a graded Lie algebra module structure

$$\mathcal{L}\colon \mathrm{HH}^{p+1}(\mathcal{A}) \otimes_{\Bbbk} \mathrm{HH}_n(\mathcal{M}) \to \mathrm{HH}_{n-p}(\mathcal{M}). \tag{25}$$

When $p=-1$ we get a map

$$\mathrm{HH}^0(\mathcal{A}) \otimes_{\Bbbk} \mathrm{HH}_n(\mathcal{M}) \to \mathrm{HH}_{n+1}(\mathcal{M}).$$

If $\mathcal{A}$ is commutative then $\mathrm{HH}^0(\mathcal{A}) \cong \mathcal{A}$ and for $f \in \mathcal{A}$ we denote by df the map $\mathrm{HH}_n(\mathcal{M}) \to \mathrm{HH}_{n+1}(\mathcal{M})$ induced in (25) by f (the condition of being a BV-module implies that $d(fg) = f\,dg + g\,df$). If we take $\mathcal{A} = \Bbbk[x_1,\dots,x_k]$ then we obtain a map

$$\sum_i d(x_i^N)\colon \mathrm{HH}_n(\mathcal{M}) \to \mathrm{HH}_{n+1}(\mathcal{M})$$

for any $\Bbbk[x_1,\dots,x_k]$–bimodule $\mathcal{M}$. One would like this map to give a differential d_{-N} which commutes with d and such that, as in Theorem 6.1, the total differential $d+d_{-N}$ defines a doubly graded link invariant $\mathcal{H}_{-N}(L)$. This would give us a spectral sequence which commences at $\mathcal{H}(L)$ and converges to $\mathcal{H}_N(L)$ for any $N \in \mathbb{Z}$.

On the other hand, if we take $p=0$ then we get a map

$$\mathrm{HH}^1(\mathcal{A}) \otimes_{\Bbbk} \mathrm{HH}_n(\mathcal{M}) \to \mathrm{HH}_n(\mathcal{M}). \tag{26}$$

Since $\mathrm{HH}^1(\Bbbk[x]) \subset \mathrm{HH}^1(\Bbbk[x_1,\dots,x_k])$ can be identified with the so-called Witt algebra one would hope that the resulting action from (26) agrees with the action of the Witt algebra defined in [17] (see the introduction and Theorem 5.6 therein).

Finally, it is worth noting that in [3, Section 2.3] and [2, Corollary 1.1.3] one obtains a Gerstenhaber algebra structure on $\mathrm{Tor}^*_X(\mathcal{O}_Y,\mathcal{O}_Z)$ whenever Y, Z are smooth coisotropic subvarieties inside a smooth Poisson variety X as well as a BV-module structure on $\mathrm{Ext}^*_X(\mathcal{O}_Y,\mathcal{O}_Z)$. In our case each term $\mathcal{M}$ in the complex $\mathcal{T}'_\beta$ is a direct sum of the structure sheaves of nonsmooth Lagrangian subvarieties inside $\mathbb{A}_{\underline{k}} \times \mathbb{A}_{\underline{k}}$, where the latter is equipped with the standard symplectic structure. This suggests that $\mathrm{HH}_*(\mathcal{M})$ might carry the structure of a Gerstenhaber algebra and $\mathrm{HH}^*(\mathcal{M})$ that of a BV-module over it.

Appendix: Ψ–functors

In this section we suppose all varieties are smooth. However, we work over an arbitrary base ring and do not assume properness at any point. The results also hold if we equip all our varieties with an action of $\Bbbk^\times$ and work equivariantly.

Fix a variety Z. For any two varieties Y_1, Y_2 we define

- $\Psi_Z\colon D(Y_1 \times Y_2) \to D((Y_1 \times Z) \times (Y_2 \times Z))$ and
- $\Psi'_Z\colon D((Y_1 \times Z) \times (Y_2 \times Z)) \to D(Y_1 \times Y_2)$

via $\Psi_Z := \Delta_* \pi^*$ and $\Psi'_Z := \pi_* \Delta^!$, where π and Δ are the natural projection and diagonal inclusion maps

$$\pi\colon Y_1 \times Z \times Y_2 \to Y_1 \times Y_2 \quad \text{and} \quad \Delta\colon Y_1 \times Z \times Y_2 \to (Y_1 \times Z) \times (Y_2 \times Z).$$

Recall that if $i\colon Y_1 \to Y_2$ is an inclusion of smooth varieties then $i^!(\cdot) = i^*(\cdot) \otimes \omega_i[-c]$, where $\omega_i = \omega_{Y_1} \otimes \omega_{Y_2}^\vee|_{Y_1}$ and c is the codimension of the inclusion.

Proposition A.1 *Let Y_1, Y_2, Y_3, Z_1, Z_2, Z_3 be six varieties and suppose*

$$\mathcal{P} \in D(Y_1 \times Y_2), \quad \mathcal{Q} \in D(Y_2 \times Y_3), \quad \mathcal{P}' \in D(Z_1 \times Z_2), \quad \mathcal{Q}' \in D(Z_2 \times Z_3).$$

Then $(\mathcal{Q} \boxtimes \mathcal{Q}') \star (\mathcal{P} \boxtimes \mathcal{P}') \cong (\mathcal{Q} \star \mathcal{P}) \boxtimes (\mathcal{Q}' \star \mathcal{P}')$.

Proof This is a fairly straightforward exercise with kernels, which we leave up to the reader. □

Corollary A.2 *Let Y_1, Y_2, Y_3 be three varieties and suppose*

$$\mathcal{P} \in D(Y_1 \times Y_2) \quad \text{and} \quad \mathcal{Q} \in D(Y_2 \times Y_3).$$

Then $\Psi_Z(\mathcal{Q} \star \mathcal{P}) \cong \Psi_Z(\mathcal{Q}) \star \Psi_Z(\mathcal{P})$.

Proof This follows from Proposition A.1 by taking $Z_1 = Z_2 = Z$ and $\mathcal{P}' = \mathcal{Q}' = \Delta_* \mathcal{O}_Z$ because in this case $\Psi_Z(\cdot) \cong (\cdot) \boxtimes \mathcal{O}_Z$. □

Proposition A.3 *Let Y_1, Y_2, Y_3 be three varieties and suppose*

$$\mathcal{P} \in D(Y_1 \times Y_2) \quad \text{and} \quad \mathcal{Q} \in D((Y_2 \times Z) \times (Y_3 \times Z)).$$

Then $\Psi'_Z(\mathcal{Q} \star \Psi_Z(\mathcal{P})) \cong \Psi'_Z(\mathcal{Q}) \star \mathcal{P} \in D(Y_1 \times Y_3)$.

Proof For $i, j \in \{1, 2, 3\}$ denote by $p_{ij}\colon Y_1 \times Y_2 \times Y_3 \to Y_i \times Y_j$ and

$$p'_{ij}\colon (Y_1 \times Z) \times (Y_2 \times Z) \times (Y_3 \times Z) \to (Y_i \times Z) \times (Y_j \times Z)$$

the natural projections. We also denote by

$$\pi_{ij}\colon Y_i \times Z \times Y_j \to Y_i \times Y_j \quad \text{and} \quad \Delta_{ij}\colon Y_i \times Z \times Y_j \to (Y_i \times Z) \times (Y_j \times Z)$$

the projection and diagonal inclusion. Then

$$\begin{aligned}
\Psi'_Z(\mathcal{Q} \star \Psi_Z(\mathcal{P})) &\cong \pi_{13*}\Delta_{13}^*(\mathcal{Q} \star (\Delta_{12*}\pi_{12}^*\mathcal{P})) \\
&\cong \pi_{13*}\Delta_{13}^*(p'_{13*}(p'^*_{12}\Delta_{12*}\pi_{12}^*\mathcal{P} \otimes p'^*_{23}\mathcal{Q})) \\
&\cong \pi_{13*}\Delta_{13}^*(p'_{13*}(\Delta'_{12*}p''^*_{12}\pi_{12}^*\mathcal{P} \otimes p'^*_{23}\mathcal{Q})) \\
&\cong \pi_{13*}\Delta_{13}^*(p'_{13*}\Delta'_{12*}(p''^*_{12}\pi_{12}^*\mathcal{P} \otimes \Delta'^*_{12}p'^*_{23}\mathcal{Q})) \\
&\cong \pi_{13*}\Delta_{13}^*q_{13*}(p''^*_{12}\pi_{12}^*\mathcal{P} \otimes p'''^*_{23}\Delta_{23}^*\mathcal{Q}) \\
&\cong \pi_{13*}q'_{13*}\Delta_Z^*(p''^*_{12}\pi_{12}^*\mathcal{P} \otimes p'''^*_{23}\Delta_{23}^*\mathcal{Q})
\end{aligned}$$

where the third isomorphism follows from the commutative square

$$\begin{tikzcd}
(Y_1 \times Z \times Y_2) \times (Y_3 \times Z) \arrow[r, "\Delta'_{12}"] \arrow[d, "p''_{12}"] & (Y_1 \times Z) \times (Y_2 \times Z) \times (Y_3 \times Z) \arrow[d, "p'_{12}"] \\
Y_1 \times Z \times Y_2 \arrow[r, "\Delta_{12}"] & (Y_1 \times Z) \times (Y_2 \times Z)
\end{tikzcd}$$

the fourth via the projection formula, the fifth using $p'_{23}\Delta'_{12} = \Delta_{23}p'''_{23}$, where p'''_{23} is the map

$$(Y_1 \times Z \times Y_2) \times (Y_3 \times Z) \to Y_2 \times Z \times Y_3, \quad (x_1, z, x_2, x_3, z') \mapsto (x_2, z, x_3),$$

and the last from the commutative square

$$\begin{tikzcd}
(Y_1 \times Z \times Y_2) \times Y_3 \arrow[r, "\Delta_Z"] \arrow[d, "q'_{13}"] & (Y_1 \times Z \times Y_2) \times (Y_3 \times Z) \arrow[d, "q_{13}"] \\
Y_1 \times Z \times Y_3 \arrow[r, "\Delta_{13}"] & (Y_1 \times Z) \times (Y_3 \times Z)
\end{tikzcd}$$

Now, $\pi_{13}q'_{13} = p_{13}(\pi_{12} \times \mathrm{id}_{Y_3})$ and $\pi_{12}p''_{12}\Delta_Z = p_{12}(\pi_{12} \times \mathrm{id}_{Y_3})$, so we get

$$\begin{aligned}
&\pi_{13*}q'_{13*}\Delta_Z^*(p''^*_{12}\pi_{12}^*\mathcal{P} \otimes p'''^*_{23}\Delta_{23}^*\mathcal{Q}) \\
&\quad\cong p_{13*}(\pi_{12} \times \mathrm{id}_{Y_3})_*(\Delta_Z^* p''^*_{12}\pi_{12}^*\mathcal{P} \otimes \Delta_Z^* p'''^*_{23}\Delta_{23}^*\mathcal{Q} \otimes \omega_Z^\vee[-\dim Z]) \\
&\quad\cong p_{13*}(\pi_{12} \times \mathrm{id}_{Y_3})_*((\pi_{12} \times \mathrm{id}_{Y_3})^* p_{12}^*\mathcal{P} \otimes \Delta_Z^* p'''^*_{23}\Delta_{23}^*\mathcal{Q} \otimes \omega_Z^\vee[-\dim Z]) \\
&\quad\cong p_{13*}(p_{12}^*\mathcal{P} \otimes (\pi_{12} \times \mathrm{id}_{Y_3})_*((p'''_{23} \circ \Delta_Z)^*(\Delta_{23}^*\mathcal{Q} \otimes \omega_Z^\vee[-\dim Z]))) \\
&\quad\cong p_{13*}(p_{12}^*\mathcal{P} \otimes p_{23}^*\pi_{23*}\Delta_{23}^*\mathcal{Q}) \\
&\quad\cong \Psi'_Z(\mathcal{Q}) * \mathcal{P},
\end{aligned}$$

where the third isomorphism is via the projection formula and the fourth uses

$$\begin{tikzcd}
(Y_1 \times Z \times Y_2) \times Y_3 \arrow[r, "\pi_{12}\times\mathrm{id}_{Y_3}"] \arrow[d, "p'''_{23}\circ\Delta_Z"] & Y_1 \times Y_2 \times Y_3 \arrow[d, "p_{23}"] \\
Y_2 \times Z \times Y_3 \arrow[r, "\pi_{23}"] & Y_2 \times Y_3
\end{tikzcd}$$

The result follows. □

References

[1] **M Abel**, **M Hogancamp**, *Categorified Young symmetrizers and stable homology of torus links, II*, Selecta Math. 23 (2017) 1739–1801 MR

[2] **V Baranovsky**, **V Ginzburg**, *Gerstenhaber–Batalin–Vilkovisky structures on coisotropic intersections*, Math. Res. Lett. 17 (2010) 211–229 MR

[3] **K Behrend**, **B Fantechi**, *Gerstenhaber and Batalin–Vilkovisky structures on Lagrangian intersections*, from "Algebra, arithmetic, and geometry: in honor of Yu I Manin, I" (Y Tschinkel, Y Zarhin, editors), Progr. Math. 269, Birkhäuser, Boston (2009) 1–47 MR

[4] **S Cautis**, *Rigidity in higher representation theory*, preprint (2014) arXiv

[5] **S Cautis**, *Clasp technology to knot homology via the affine Grassmannian*, Math. Ann. 363 (2015) 1053–1115 MR

[6] **S Cautis**, **C Dodd**, **J Kamnitzer**, *Associated graded of Hodge modules and categorical $\mathfrak{sl}_2$ actions*, preprint (2016) arXiv

[7] **S Cautis**, **J Kamnitzer**, *Knot homology via derived categories of coherent sheaves, I: The $\mathfrak{sl}(2)$–case*, Duke Math. J. 142 (2008) 511–588 MR

[8] **S Cautis**, **J Kamnitzer**, *Braiding via geometric Lie algebra actions*, Compos. Math. 148 (2012) 464–506 MR

[9] **S Cautis**, **J Kamnitzer**, **S Morrison**, *Webs and quantum skew Howe duality*, Math. Ann. 360 (2014) 351–390 MR

[10] **S Cautis**, **A D Lauda**, *Implicit structure in 2–representations of quantum groups*, Selecta Math. 21 (2015) 201–244 MR

[11] **N M Dunfield**, **S Gukov**, **J Rasmussen**, *The superpolynomial for knot homologies*, Experiment. Math. 15 (2006) 129–159 MR

[12] **S Gukov**, **M Stošić**, *Homological algebra of knots and BPS states*, from "Proceedings of the Freedman Fest" (R Kirby, V Krushkal, Z Wang, editors), Geom. Topol. Monogr. 18, Geom. Topol. Publ., Coventry (2012) 309–367 MR

[13] **M Hogancamp**, *Stable homology of torus links via categorified Young symmetrizers, I: One-row partitions, v1*, preprint (2015) arXiv

[14] **M Khovanov**, *Triply-graded link homology and Hochschild homology of Soergel bimodules*, Internat. J. Math. 18 (2007) 869–885 MR

[15] **M Khovanov**, **A D Lauda**, *A categorification of quantum* $\mathrm{sl}(n)$, Quantum Topol. 1 (2010) 1–92 MR

[16] **M Khovanov**, **L Rozansky**, *Matrix factorizations and link homology, II*, Geom. Topol. 12 (2008) 1387–1425 MR

[17] **M Khovanov**, **L Rozansky**, *Positive half of the Witt algebra acts on triply graded link homology*, Quantum Topol. 7 (2016) 737–795 MR

[18] **N Kowalzig**, **U Krähmer**, *Batalin–Vilkovisky structures on Ext and Tor*, J. Reine Angew. Math. 697 (2014) 159–219 MR

[19] **G Lusztig**, *Introduction to quantum groups*, Progress in Mathematics 110, Birkhäuser, Boston (1993) MR

[20] **M Mackaay**, **M Stošić**, **P Vaz**, *The* $1,2$*–coloured HOMFLY-PT link homology*, Trans. Amer. Math. Soc. 363 (2011) 2091–2124 MR

[21] **J Rasmussen**, *Some differentials on Khovanov–Rozansky homology*, Geom. Topol. 19 (2015) 3031–3104 MR

[22] **R Rouquier**, 2*-Kac–Moody algebras*, preprint (2008) arXiv

[23] **L Rozansky**, *An infinite torus braid yields a categorified Jones–Wenzl projector*, Fund. Math. 225 (2014) 305–326 MR

[24] **B Webster**, **G Williamson**, *A geometric construction of colored HOMFLYPT homology*, Geom. Topol. 21 (2017) 2557–2600 MR

[25] **P Wedrich**, *Exponential growth of colored HOMFLY-PT homology*, preprint (2016) arXiv

Department of Mathematics, University of British Columbia
Vancouver, BC, Canada

cautis@math.ubc.ca

Received: 8 February 2017 Revised: 16 May 2017

Algebraic & Geometric Topology 17 (2017) 3837–3852

Correction to the articles Homotopy theory of nonsymmetric operads, I–II

FERNANDO MURO

We correct a mistake in Algebr. Geom. Topol. 11 (2011) 1541–1599 on the construction of push-outs along free morphisms of algebras over a nonsymmetric operad, and we fix the affected results from there and a follow-up article (Algebr. Geom. Topol. 14 (2014) 229–281).

18D50, 55U35; 18D10, 18D35, 18D20

Introduction

In [10, Section 8] we give a wrong construction of push-outs along free maps in the category of algebras over an operad. Contrary to what we intended and claimed in the introduction, it does not generalize Harper [6, Proposition 7.32], which is the correct construction. It does not even yield Schwede and Shipley's description [12] of push-outs along free maps in the category of monoids. As Donald Yau pointed out to us, the trivial ring only maps to itself (since it is characterized by the fact that $0=1$), but our construction yields $\bigotimes_{n\geq 1} Z^{\otimes n}$ for the coproduct of the trivial ring and the tensor algebra on Z. Here, we fix this mistake and its consequences in Muro [10; 11]. The main results of these papers, presented in their introductions, remain true as stated, modulo a modification in the nonsymmetric monoid axiom [10, Definition 9.1] and a strengthening in the hypotheses of [11, Theorem 1.13 and Corollary 1.14]. These changes do not affect the applications. Moreover, the results which are purely on operads, not on algebras, remain completely unaffected.

1 Push-out filtrations in symmetric monoidal categories

In this section we consider operads $\mathcal{O}$ (always nonsymmetric) and their algebras A in a bicomplete closed symmetric monoidal category $\mathscr{V}$ with tensor product $\otimes$ and tensor unit $\mathbb{I}$, as a preliminary step to the more general case in the following section.

Published: 4 October 2017 DOI: 10.2140/agt.2017.17.3837

We start with Harper's description of the $\mathcal{O}$–algebra push-out

$$
\begin{tikzcd}
\mathcal{F}_{\mathcal{O}}(Y) \arrow[r, "\mathcal{F}_{\mathcal{O}}(f)"] \arrow[d, "g"'] \arrow[dr, phantom, "\text{push}"] & \mathcal{F}_{\mathcal{O}}(Z) \arrow[d, "g'"] \\
A \arrow[r, "f'"'] & B
\end{tikzcd}
\tag{1-1}
$$

Here $\mathcal{F}_{\mathcal{O}}$ is the free $\mathcal{O}$–algebra functor, $\mathcal{F}_{\mathcal{O}}(Y) = \coprod_{n\geq 0} \mathcal{O}(n) \otimes Y^{\otimes n}$, and we denote the adjoint of g by $\bar{g}\colon Y \to A$.

The *enveloping operad* $\mathcal{O}_A$ [5; 4; 3] is characterized by the fact that an operad map $\mathcal{O}_A \to \mathcal{P}$ is the same as an operad map $\mathcal{O} \to \mathcal{P}$ together with an $\mathcal{O}$–algebra map $A \to \mathcal{P}(0)$. Aritywise, $\mathcal{O}_A(t)$ is the (reflexive) coequalizer of the following diagram for $t \geq 0$ — compare [6, Proposition 7.28] —

$$
\begin{array}{c}
\displaystyle\coprod_{s\geq 0} \mathcal{O}(s+t) \otimes \left(\frac{\Sigma_{s+t}}{\Sigma_s \times \Sigma_t} \cdot \mathcal{F}_{\mathcal{O}}(A)^{\otimes s} \otimes \mathbb{I}^{\otimes t} \right) \\
\downarrow\downarrow\uparrow \\
\displaystyle\coprod_{s\geq 0} \mathcal{O}(s+t) \otimes \left(\frac{\Sigma_{s+t}}{\Sigma_s \times \Sigma_t} \cdot A^{\otimes s} \otimes \mathbb{I}^{\otimes t} \right)
\end{array}
\tag{1-2}
$$

Here, given a permutation $\sigma \in \Sigma_n$ we write $\sigma \cdot X_1 \otimes \cdots \otimes X_n = X_{\sigma^{-1}(1)} \otimes \cdots \otimes X_{\sigma^{-1}(n)}$, given a subset $S \subset \Sigma_n$ we set $S \cdot X_1 \otimes \cdots \otimes X_n = \coprod_{\sigma \in S} \sigma \cdot X_1 \otimes \cdots \otimes X_n$, and $\Sigma_{s+t}/\Sigma_s \times \Sigma_t$ identifies with the set of (s,t)–shuffles. The two arrows pointing downwards are defined by the operad structure of $\mathcal{O}$ and the $\mathcal{O}$–algebra structure of A, respectively, and the arrow pointing upwards is given by the unit of $\mathcal{O}$. For $t = 0$, the previous formula reduces to the cotriple presentation of $\mathcal{O}_A(0) = A$.

Recall from [10, Section 4] that a map $f\colon Y \to Z$ in $\mathscr{V}$ is the same as a functor $f\colon \mathbf{2} \to \mathscr{V}$ from the poset $\mathbf{2} = \{0 < 1\}$. Given maps $f_i\colon Y_i \to Z_i$ in $\mathscr{V}$ for $1 \leq i \leq n$, their *push-out product* $f_1 \odot \cdots \odot f_n$ is the latching map of the functor

$$\mathbf{2}^n \xrightarrow{f_1 \otimes \cdots \otimes f_n} \mathscr{C}$$

at the final object $(1, \ldots, 1) \in \mathbf{2}^n$ [7, Definition 15.2.5].

The following lemma is a special case of [6, Proposition 7.32].

Lemma 1.1 *The map f' in* (1-1) *is the transfinite composition of a sequence*

$$A = B_0 \xrightarrow{\varphi_1} B_1 \to \cdots \to B_{t-1} \xrightarrow{\varphi_t} B_t \to \cdots$$

in $\mathscr{V}$ such that the morphism φ_t for $t \geq 1$ is given by the push-out square

$$\begin{tikzcd} \bullet \arrow[r, "\mathcal{O}_A(t)\otimes f^{\odot t}"] \arrow[d, "\psi_t"'] \arrow[dr, phantom, "push"] & \bullet \arrow[d, "\bar{\psi}_t"] \\ B_{t-1} \arrow[r, "\varphi_t"'] & B_t \end{tikzcd}$$

where the attaching map ψ_t is defined by the following maps for $1 \leq i \leq t$:

$$\begin{aligned} \mathcal{O}_A(t) \otimes Z^{\otimes(i-1)} \otimes Y \otimes Z^{\otimes(t-i)} \to \mathcal{O}_A(t) \otimes Z^{\otimes(i-1)} \otimes A \otimes Z^{\otimes(t-i)} \\ \xrightarrow{\circ_i} \mathcal{O}_A(t-1) \otimes Z^{\otimes(t-1)} \to B_{t-1}, \end{aligned}$$

where the first map is defined by $\bar{g}\colon Y \to A$ and the last is $\bar{\psi}_{t-1}$ if $t > 1$ and the identity if $t = 1$.

This lemma is also the arity-0 part of the following one. Observe that the enveloping operad $\mathcal{O}_A$ is functorial on A in the obvious way. Moreover, it is a functor of the pair $(\mathcal{O}, A)$ regarded as a object in the Grothendieck construction of the categories of algebras over all operads.

Lemma 1.2 *If we have an $\mathcal{O}$–algebra push-out* (1-1), *$\mathcal{O}_{f'}\colon \mathcal{O}_A \to \mathcal{O}_B$ is the transfinite composition of a sequence of maps in the category of sequences*

$$\mathcal{O}_A = \mathcal{O}_{B,0} \xrightarrow{\Phi_1} \mathcal{O}_{B,1} \to \cdots \to \mathcal{O}_{B,t-1} \xrightarrow{\Phi_t} \mathcal{O}_{B,t} \to \cdots$$

such that $\Phi_t(n)$ for $t \geq 1$ and $n \geq 0$ is given by the push-out square

$$\begin{tikzcd} \bullet \arrow[r, "\mathcal{O}_A(t+n)\otimes((\Sigma_{t+n}/(\Sigma_t\times\Sigma_n))\cdot f^{\odot t}\otimes \mathbb{I}^{\otimes n})"] \arrow[d, "\Psi_t(n)"'] \arrow[dr, phantom, "push"] & \bullet \arrow[d, "\bar{\Psi}_t(n)"] \\ \mathcal{O}_{B,t-1}(n) \arrow[r, "\Phi_t(n)"'] & \mathcal{O}_{B,t}(n) \end{tikzcd}$$

where the attaching map $\Psi_t(n)$ is defined, as in Lemma 1.1, from $\bar{g}$, the composition laws $\circ_i\colon \mathcal{O}_A(t+n) \otimes A \to \mathcal{O}_A(t+n-1)$, and also $\bar{\Psi}_{t-1}(n)$ if $t > 1$.

The universal property of $\mathcal{O}_B$ allows us to obtain it as the push-out in the category of operads

$$\begin{tikzcd} \mathcal{F}(Y) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "\tilde{g}"'] \arrow[dr, phantom, "push"] & \mathcal{F}(Z) \arrow[d] \\ \mathcal{O}_A \arrow[r] & \mathcal{O}_B \end{tikzcd} \tag{1-3}$$

Here $\mathcal{F}$ is the free operad functor, f is regarded as a map of sequences concentrated in arity 0, and $\widetilde{g}$ is the adjoint of $\overline{g}\colon Y \to A \subset \mathcal{O}_A$. Lemma 1.2 follows from the description of operad pushouts in [10, Section 5].

Assume now that we have a push-out square in the category of operads

$$\begin{tikzcd} \mathcal{F}(U) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "g"'] \arrow[dr, phantom, "\text{push}"] & \mathcal{F}(V) \arrow[d, "g'"] \\ \mathcal{O} \arrow[r, "f'"'] & \mathcal{P} \end{tikzcd} \tag{1-4}$$

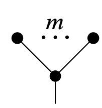
and A is a $\mathcal{P}$–algebra. The universal properties of $\mathcal{P}$ and $\mathcal{P}_A$ show that we have a similar push-out

$$\begin{tikzcd} \mathcal{F}(U)_A \arrow[r, "\mathcal{F}(f)_A"] \arrow[d, "g_A"'] \arrow[dr, phantom, "\text{push}"] & \mathcal{F}(V)_A \arrow[d, "g'_A"] \\ \mathcal{O}_A \arrow[r, "f'_A"'] & \mathcal{P}_A \end{tikzcd} \tag{1-5}$$

The enveloping operad $\mathcal{F}(V)_A$ of an algebra over a free operad admits a description similar to the free operad $\mathcal{F}(V)$; compare [1, Section 3; 10, Section 5]. Namely,

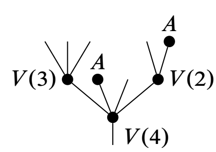
$$\mathcal{F}(V)_A(n) = \coprod_T \bigotimes_{v \in I(T)} \overline{V}(\widetilde{v}). \tag{1-6}$$

Here $\overline{V}$ is the sequence with $\overline{V}(0) = A$ and $\overline{V}(m) = V(m)$ for $m > 0$, T runs over all (isomorphism classes of) trees (planted, planar and with leaves) with n leaves [10, Section 3] which do not contain any *forbidden configuration* for $m \geq 1$,

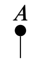


and $I(T)$ is the set of inner vertices of T. Each coproduct factor in (1-6) is usually depicted by labeling each inner vertex v of T with $\overline{V}(\widetilde{v})$, where $\widetilde{v}$ is the number of edges adjacent to v from above, eg

(1-7)



The reason for the forbidden configuration is that we must take into account the $\mathcal{F}(V)$–algebra structure maps $V(m) \otimes A^{\otimes m} \to A$. The operad structure on $\mathcal{F}(V)_A$ is defined by formal tree grafting, applying (repeatedly) if necessary the previous structure maps whenever a forbidden configuration appears, collapsing it to



Hence, the push-out (1-5) admits a filtration description analogous to (1-4). The *level* of a vertex v of T is the number of edges in the shortest path to the root. We say that v is *even* if it has even level. *Odd* vertices are defined similarly. The sets of even and odd inner vertices in T are denoted by $I^e(T)$ and $I^o(T)$, respectively.

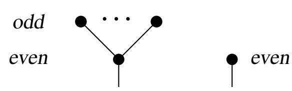
Lemma 1.3 *Given an operad push-out* (1-4) *and a $\mathcal{P}$–algebra A, $f'_A\colon \mathcal{O}_A \to \mathcal{P}_A$ is the transfinite composition of a sequence of maps of sequences*

$$\mathcal{O}_A = \mathcal{P}_{A,0} \xrightarrow{\Phi_1} \mathcal{P}_{A,1} \to \cdots \to \mathcal{P}_{A,t-1} \xrightarrow{\Phi_t} \mathcal{P}_{A,t} \to \cdots$$

such that $\Phi_t(n)$ for $t \geq 1$ and $n \geq 0$ is given by the push-out square

$$\begin{tikzcd}[column sep=huge]
\bullet \arrow[r, "\coprod_T \bigodot_{v\in I^e(T)} f(\tilde{v})\otimes\bigotimes_{w\in I^o(T)} \mathcal{O}_A(\tilde{w})"] \arrow[d, "\Psi_t(n)"'] \arrow[dr, phantom, "push"] & \bullet \arrow[d, "\overline{\Psi}_t(n)"] \\
\mathcal{P}_{A,t-1}(n) \arrow[r, "\Phi_t(n)"'] & \mathcal{P}_{A,t}(n)
\end{tikzcd}$$

where T runs over the isomorphism classes of trees with n leaves concentrated in even levels and t inner even vertices not containing



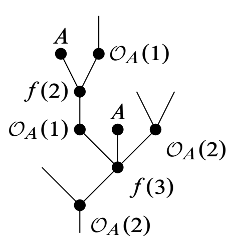
The attaching map $\Psi_t(n)$ is defined by the maps from

$$U(\tilde{u}) \otimes \bigotimes_{v\in I^e(T)\setminus\{u\}} V(\tilde{v}) \otimes \bigotimes_{w\in I^o(T)} \mathcal{O}_A(\tilde{w}), \quad u \in I^e(T),$$

defined by the composite $U \to \mathcal{O} \to \mathcal{O}_A$, composition in $\mathcal{O}_A$, the structure maps $V(m) \otimes A^{\otimes m} \to A$, and the previous $\overline{\Psi}_s(n)$, $s < t$.

Each factor of the coproduct of maps in the statement of the previous lemma is depicted by labeling each even inner vertex v of T with $f(\tilde{v})$, and each odd inner vertex w with $\mathcal{O}_A(\tilde{w})$, eg

(1-8)



The proof of Lemma 1.3 is a slight variation of the explicit construction of the push-out (1-4) given in [10, Section 5]. We believe that this result is new in the literature.

2 Push-out filtrations in nonsymmetric settings

We now turn to our general setting, where operads $\mathcal{O}$ still live in $\mathscr{V}$ but their algebras A live in a bicomplete biclosed monoidal category $\mathscr{C}$ (possibly nonsymmetric) endowed with a strong monoidal left adjoint $z\colon \mathscr{V} \to \mathscr{C}$ which is central, meaning that it is equipped with coherent isomorphisms $z(X) \otimes Y \cong Y \otimes z(X)$. Objects in $\mathscr{V}$ have "underlying" objects in $\mathscr{C}$ via z. We will often drop z from notation. Here we indicate how the three previous lemmas extend to this context.

Enveloping operads do not make sense in this setting since they should live in $\mathscr{C}$, but the definition of operad requires a symmetric tensor product. We must instead consider (always nonsymmetric) *functor-operads* $F = \{F(n)\}_{n \geq 0}$ in $\mathscr{C}$ [9], also known as *multitensors* [2]. They consist of a sequence of functors $F(n)\colon \mathscr{C}^n \to \mathscr{C}$ equipped with composition and unit natural transformations

$$\circ_i\colon F(p)(\overset{i-1}{\cdots}, F(q), \overset{p-i}{\cdots}) \to F(p+q-1), \quad 1 \leq i \leq p,\ q \geq 0,$$
$$u\colon \mathrm{id}_{\mathscr{C}} \to F(1),$$

satisfying relations similar to operads. The values $\mathcal{O}_A(t)(X_1, \dots, X_t)$ of the *enveloping functor-operad* $\mathcal{O}_A$ are defined by replacing $\mathbb{I}^{\otimes t}$ with $X_1 \otimes \cdots \otimes X_t$ in (1-2). Again, $\mathcal{O}_A(0)$, which is a functor from the discrete category on one object $\mathscr{C}^{\mathbf{0}}$, identifies with A. Enveloping functor-operads satisfy the same functoriality properties as enveloping operads do when $\mathscr{C} = \mathscr{V}$.

Consider the $\mathcal{O}$–algebra push-out (1-1), now in our current setting. As above, we regard maps in $\mathscr{C}$ as functors $\mathbf{2} \to \mathscr{C}$. We now present our first amended statement, where the numbering refers to the cited paper.

Lemma 8.1 [10] *The map f' in* (1-1) *is the transfinite composition of a sequence*

$$A = B_0 \xrightarrow{\varphi_1} B_1 \to \cdots \to B_{t-1} \xrightarrow{\varphi_t} B_t \to \cdots$$

in $\mathscr{C}$ such that the morphism φ_t for $t \geq 1$ is given by the push-out square

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d, "\psi_t"'] \arrow[dr, phantom, "push"] & \bullet \arrow[d, "\bar{\psi}_t"] \\ B_{t-1} \arrow[r, "\varphi_t"'] & B_t \end{tikzcd}$$

where the top map is the latching map of $\mathcal{O}_A(t)(f,\dots,f)$ at the final object and the attaching map ψ_t is given by, for $1 \le i \le t$,

$$\mathcal{O}_A(t)(Z,\overset{i-1}{\dots},Z,Y,Z,\overset{t-i}{\dots},Z) \to \mathcal{O}_A(t)(Z,\overset{i-1}{\dots},Z,A,Z,\overset{t-i}{\dots},Z) \xrightarrow{\circ_i} \mathcal{O}_A(t-1)(Z,\overset{t-1}{\dots},Z) \to B_{t-1}$$

where the first map is defined by $\bar g\colon Y \to A$ and the last is $\bar\psi_{t-1}$ if $t>1$ and the identity if $t=1$.

The same proof as in the case $\mathscr{C}=\mathscr{V}$ [6, Proposition 7.32] works here, mutatis mutandis. The lemma actually extends to enveloping functor-operads. Given a functor $F\colon \mathscr{C}^n \to \mathscr{C}$ and a permutation $\sigma\in\Sigma_n$, we let $F\sigma\cdot(X_1,\dots,X_n)=F(X_{\sigma^{-1}(1)},\dots,X_{\sigma^{-1}(n)})$, and given a subset $S\subset\Sigma_n$ we set $FS\cdot(X_1,\dots,X_n)=\coprod_{\sigma\in S}F\sigma\cdot(X_1,\dots,X_n)$.

Lemma 2.1 *If we have an $\mathcal{O}$–algebra push-out* (1-1)*, $\mathcal{O}_{f'}$ is the transfinite composition of a sequence of natural transformations between sequences of functors*

$$\mathcal{O}_A=\mathcal{O}_{B,0}\xrightarrow{\Phi_1}\mathcal{O}_{B,1}\to\cdots\to\mathcal{O}_{B,t-1}\xrightarrow{\Phi_t}\mathcal{O}_{B,t}\to\cdots$$

such that, pointwise, $\Phi_t(n)(X_1,\dots,X_n)$ for $t\ge1$ and $n\ge0$ is given by the push-out

$$\begin{tikzcd}
\bullet \arrow[r] \arrow[d, "\Psi_t(n)(X_1,\dots,X_n)"'] \arrow[dr, phantom, "push"] & \bullet \arrow[d, "\overline{\Psi}_t(n)(X_1,\dots,X_n)"] \\
\mathcal{O}_{B,t-1}(n)(X_1,\dots,X_n) \arrow[r, "\Phi_t(n)(X_1,\dots,X_n)"'] & \mathcal{O}_{B,t}(n)(X_1,\dots,X_n)
\end{tikzcd}$$

where the top map is the latching map of

$$\mathcal{O}_A(t+n)(\Sigma_{t+n}/(\Sigma_t\times\Sigma_n))\cdot(f,\overset{t}{\dots},f,X_1,\dots,X_n)$$

at the final object and the attaching map $\Psi_t(n)(X_1,\dots,X_n)$ is defined, as in Lemma 8.1 above, from $\overline{g}$, the composition laws

$$\circ_i\colon \mathcal{O}_A(t+n)(\dots,A,\dots)\to\mathcal{O}_A(t+n-1),$$

and also $\overline{\Psi}_{t-1}(n)(X_1,\dots,X_n)$ if $t>1$.

This lemma can be proved by fitting Lemma 8.1 above into the coequalizer definition of $\mathcal{O}_B$.

Given a sequence V in $\mathscr{V}$ we identify the object $V(n)$ with the functor $\mathscr{C}^n\to\mathscr{C}\colon (X_1,\dots,X_n)\mapsto V(n)\otimes X_1\otimes\cdots\otimes X_n$. In this way, a sequence in $\mathscr{V}$ can be regarded as a sequence of functors. We similarly identify a map of sequences in $\mathscr{V}$ with the obvious natural transformations. An operad in $\mathscr{V}$ yields a functor-operad in $\mathscr{C}$ through this assignment, and the natural operad map $\mathcal{O}\to\mathcal{O}_A$ becomes a functor-operad map when A is in $\mathscr{C}$.

Lemma 2.2 *Given an operad push-out* (1-4) *and a $\mathcal{P}$–algebra A, the map $f'_A\colon \mathcal{O}_A \to \mathcal{P}_A$ is the transfinite composition of a sequence of natural transformations between sequences of functors*

$$\mathcal{O}_A = \mathcal{P}_{A,0} \xrightarrow{\Phi_1} \mathcal{P}_{A,1} \to \cdots \to \mathcal{P}_{A,t-1} \xrightarrow{\Phi_t} \mathcal{P}_{A,t} \to \cdots$$

such that $\Phi_t(n)(X_1,\dots,X_n)$ for $t \geq 1$ and $n \geq 0$ is given by the push-out square

$$\begin{tikzcd}
\bullet \arrow[rr, "\coprod_T \widetilde{\Phi}_t(T)(X_1,\dots,X_n)"] \arrow[d, "\Psi_t(n)(X_1,\dots,X_n)"'] \arrow[drr, phantom, "\text{push}"] & & \bullet \arrow[d, "\overline{\Psi}_t(n)(X_1,\dots,X_n)"] \\
\mathcal{P}_{A,t-1}(n)(X_1,\dots,X_n) \arrow[rr, "\Phi_t(n)(X_1,\dots,X_n)"'] & & \mathcal{P}_{A,t}(n)(X_1,\dots,X_n)
\end{tikzcd}$$

where T runs over the same set of trees as in Lemma 1.3, $\widetilde{\Phi}_t(T)(X_1,\dots,X_n)$ is the latching map at the final object of the functor $\mathbf{2}^t \to \mathscr{C}$ obtained by composing horizontally the natural transformations $f(\widetilde{v})$ for $v \in I^e(T)$ and the functors $\mathcal{O}_A(\widetilde{w})$ for $w \in I^o(T)$ according to the structure of the tree T (see eg (1-8)), *evaluating at $X_1,\dots,X_n$ in the slots indicated by the leaves, and the attaching map $\Psi_t(n)(X_1,\dots,X_n)$ is defined by the composite $U \to \mathcal{O} \to \mathcal{O}_A$, composition in $\mathcal{O}_A$, the structure maps $V(m) \otimes A^{\otimes m} \to A$, and the previous $\overline{\Psi}_s(n)(X_1,\dots,X_n)$ for $s < t$.*

For the proof of this lemma, we can fit the filtration for the bottom map in (1-4) constructed in [10, Section 5] into the coequalizer definition of $\mathcal{P}_A$.

3 Corrected results

We will sometimes restrict to the following class of operads with homotopically well-behaved enveloping (functor-)operads.

Definition 3.1 Suppose that the tensor unit of $\mathscr{V}$ is cofibrant. An operad $\mathcal{O}$ is *excellent* if the functor $A \mapsto \mathcal{O}_A$ takes an $\mathcal{O}$–algebra A with underlying cofibrant object to a cofibrant sequence, and a weak equivalence between $\mathcal{O}$–algebras with underlying cofibrant objects to a weak equivalence of sequences.

The meaning is clear in case $\mathscr{C} = \mathscr{V}$. In the general case we must consider sequences of functors $\mathscr{C}^n \to \mathscr{C}$ for $n \geq 0$ rather than objects in $\mathscr{V}$. Homotopical notions in this more general context will be defined below. When the tensor unit is not cofibrant, the previous definition makes sense but it is not useful. We will also deal with this more general case below.

An operad which is not excellent, with $\mathscr{C} = \mathscr{V}$ the category of chain complexes over a commutative ring, is the operad whose algebras are nonunital DG–algebras A with

$A^3 = 0$. This operad, which has an underlying cofibrant sequence, can be used to construct examples showing the necessity of the excellence assumption in several statements below.

3.2 Corrections to statements

Note that [10, Lemma 8.1] has already been amended in the previous section. We do not repeat it here.

Proposition 9.2(2) [10] *Consider the push-out diagram* (1-1) *in* $\mathrm{Alg}_{\mathscr{C}}(\mathcal{O})$.

(2) *Suppose that* f *is a cofibration in* $\mathscr{C}$ *and either* A *is cofibrant in* $\mathscr{C}$ *and* $\mathcal{O}$ *is excellent or* A *is cofibrant as an* $\mathcal{O}$*–algebra in* $\mathscr{C}$ *and* $\mathcal{O}(n)$ *is cofibrant in* $\mathscr{V}$ *for* $n \geq 0$. *Then* $f'\colon A \to B$ *is a cofibration in* $\mathscr{C}$.

We will modify not [10, Proposition 9.2(1)] but the definition of the nonsymmetric monoid axiom, so that the statement will be tautologically true by Lemma 8.1 above.

Definition 9.1 [10] The *monoid axiom* in the $\mathscr{V}$–algebra $\mathscr{C}$ says that relative K'–cell complexes are weak equivalences, where K' is the class of morphisms

$$\begin{aligned} K' = \Big\{ & f \otimes X,\ X \otimes f,\ \text{latching map of } \mathcal{O}_A(t)(f, \dots, f) \text{ at the final object} \ \Big| \\ & X \text{ is an object in } \mathscr{C},\ f \text{ is a trivial cofibration in } \mathscr{C},\ \mathcal{O} \text{ is an operad in } \mathscr{V}, \\ & \qquad A \text{ is an } \mathcal{O}\text{–algebra in } \mathscr{C},\ t \geq 1 \Big\}. \end{aligned}$$

This axiom is equivalent to Schwede and Shipley [12, Definition 3.3] if $\mathscr{C} = \mathscr{V}$.

The following two modifications are forced by the previous amendments.

(6-2) [11] Replace this equation with the latching map of $\mathcal{O}_A(t)(f, \dots, f)$ at the final object.

Definition 2.3(3) [11] Replace with the new [10, Definition 9.1] above.

Now, in [11, **Theorems 1.13, 8.1 and D.13, Corollaries 1.14 and 8.2 and Propositions 8.3 and D.14**], we must assume in addition that the operad $\mathcal{O}$ is excellent.

The most general of these results is [11, Theorem D.13], which follows from Proposition 3.4.3 and, if the tensor unit is not cofibrant, the remarks in Section 3.7 below.

Note that [11, Lemmas 6.6 and D.1] are not useful any more, since the map [11, (6-2)] plays no role after the corrections.

Corollary D.2 [11] *Suppose that $\mathscr{C}$ satisfies the strong unit axiom and either A is pseudocofibrant in $\mathscr{C}$ and $\mathcal{O}$ is excellent, or A is pseudocofibrant as an $\mathcal{O}$–algebra in $\mathscr{C}$ and $\mathcal{O}(n)$ is cofibrant in $\mathscr{C}$ for $n \geq 0$. Then any cofibration $\phi\colon A \rightarrowtail B$ in $\mathrm{Alg}_{\mathscr{C}}(\mathcal{O})$ is also a cofibration in $\mathscr{C}$.*

3.3 Correct statements needing new proofs

The nonsymmetric monoidal category $\mathrm{Graph}_S(\mathscr{V})$ of $\mathscr{V}$–graphs with object set S [10, Definition 10.1] still satisfies the amended nonsymmetric monoid axiom.

Proof of [10, Proposition 10.3] It is easy to check (using the symmetry of $\mathscr{V}$) that the latching map of $\mathcal{O}_A(t)(f,\dots,f)$ at the final object is componentwise a coproduct of maps, each of which is the tensor product of a single object in $\mathscr{V}$ with a push-out product of components of f, which are trivial cofibrations in $\mathscr{V}$. Such a push-out product is again a trivial cofibration by the push-out product axiom. Hence, any K'–cell complex is componentwise a K–cell complex in the sense of [10, Definition 6.1], and therefore a weak equivalence by the monoid axiom for $\mathscr{V}$. □

The modifications made to [10, Proposition 9.2] have no impact on [10, Lemma 9.4 and Corollary 9.5], however [10, Lemma 9.6 and Theorem 1.3; 11, Theorems 6.7 and D.4] require a new proof. They follow from the arity-0 part of Proposition 3.4.2 and, if the tensor unit is not cofibrant, the remarks in 3.7 below. The modification in [11, (6-2)] forces us to give new proofs of [11, Propositions 7.3 and D.6]. They follow from the arity-0 part of Proposition 3.4.6 and Section 3.7.

3.4 Auxiliary results

We need the following results to prove the amended statements and to fix proofs of correct statements affected by the amendments.

Proposition 3.4.1 *Let $\mathcal{O}$ be an operad with underlying cofibrant sequence. For any cofibration with cofibrant source $f'\colon A \to B$ in $\mathrm{Alg}_{\mathscr{C}}(\mathcal{O})$, the map $\mathcal{O}_{f'}\colon \mathcal{O}_A \to \mathcal{O}_B$ is a cofibration of sequences. In particular, $\mathcal{O}_A$ is a cofibrant sequence for any cofibrant A in $\mathrm{Alg}_{\mathscr{C}}(\mathcal{O})$.*

Proposition 3.4.2 *Let $\phi\colon \mathcal{O} \xrightarrow{\sim} \mathcal{P}$ be a weak equivalence in $\mathrm{Op}(\mathscr{V})$. Assume that the objects $\mathcal{O}(n)$ and $\mathcal{P}(n)$ are cofibrant in $\mathscr{V}$ for all $n \geq 0$. Given a cofibrant $\mathcal{O}$–algebra A in $\mathscr{C}$, the map $\phi_{\eta_A}\colon \mathcal{O}_A \to \mathcal{P}_{\phi_* A}$ induced by ϕ and by the unit $\eta_A\colon A \to \phi^*\phi_* A$ of the change of operad adjunction $\phi_* \dashv \phi^*$ [10, (1)] is a weak equivalence of sequences.*

Proposition 3.4.3 *If $\mathcal{O}$ is an excellent operad in $\mathscr{V}$ and*

$$\begin{tikzcd} A \arrow[r, tail, "\psi"] \arrow[d, "\varphi"', "\sim"] \arrow[dr, phantom, "push"] & B \arrow[d, "\varphi'"] \\ C \arrow[r, tail, "\psi'"'] & C \cup_A B \end{tikzcd}$$

is a push-out of $\mathcal{O}$–algebras in $\mathscr{C}$ such that the underlying objects of A and C are cofibrant, then φ' is a weak equivalence.

The goal of the two following results is to exhibit a huge class of excellent operads. They are not strictly required to correct the results in [10; 11], but they are essential for applications. We believe that these results, which imply homotopy invariance of enveloping (functor-)operads, are new in the literature in this generality. Similar results for chain complexes have been obtained in [4, Section 17.4].

Proposition 3.4.4 *The initial operad $\mathtt{Ass}^{\mathscr{V}}$, and $\mathtt{uAss}^{\mathscr{V}}$, are excellent.*

Proposition 3.4.5 *If $f'\colon \mathcal{O} \rightarrowtail \mathcal{P}$ is a cofibration in $\mathrm{Op}(\mathscr{V})$ and $\mathcal{O}$ is an excellent operad such that $\mathcal{O}(n)$ is cofibrant for all $n \geq 0$, then so is $\mathcal{P}$.*

In the following result, in addition to our standing context $(\mathscr{V}, \mathscr{C})$ we have another one $(\mathscr{W}, \mathscr{D})$ satisfying the same formal properties. Both of them are related by Quillen pairs, $F\colon \mathscr{V} \rightleftarrows \mathscr{W} :G$ and $\bar{F}\colon \mathscr{C} \rightleftarrows \mathscr{D} :\bar{G}$, with colax monoidal left adjoints F and $\bar{F}$, equipped with a coherent natural map $\tau(X)\colon \bar{F}z(X) \to zF(X)$ which is a weak equivalence for X cofibrant [11, Section 7]. They give rise to a functor between operad categories $F^{\mathrm{oper}}\colon \mathrm{Op}(\mathscr{V}) \to \mathrm{Op}(\mathscr{W})$ and, for each operad $\mathcal{O}$ in $\mathscr{V}$, a functor between algebra categories $\bar{F}_{\mathcal{O}}\colon \mathrm{Alg}_{\mathscr{C}}(\mathcal{O}) \to \mathrm{Alg}_{\mathscr{D}}(F^{\mathrm{oper}}(\mathcal{O}))$. These functors are left adjoint to the obvious functors defined by the lax monoidal functors G and $\bar{G}$. In particular, we obtain a map of sequences $\chi_{\mathcal{O}}\colon F(\mathcal{O}) \to F^{\mathrm{oper}}(\mathcal{O})$ and natural transformations $\chi_{\mathcal{O},A}(n)\colon \bar{F}\mathcal{O}_A(n) \to F^{\mathrm{oper}}(\mathcal{O})_{\bar{F}_{\mathcal{O}}(A)}(n)\bar{F}^{\times n}$ between functors $\mathscr{C}^n \to \mathscr{D}$ for $n \geq 0$ for any $\mathcal{O}$–algebra A.

Proposition 3.4.6 *If $\bar{F} \dashv \bar{G}$ is a weak monoidal Quillen adjunction, $\mathscr{V}$ and $\mathscr{W}$ have cofibrant tensor units, $\mathcal{O}$ is a cofibrant operad in $\mathscr{V}$ and A is a cofibrant $\mathcal{O}$–algebra in $\mathscr{C}$, then $\chi_{\mathcal{O},A}(n)$ is a weak equivalence in $\mathscr{D}$ when evaluated at n cofibrant objects in $\mathscr{C}$ for $n \geq 0$.*

3.5 Proofs for $\mathscr{C} = \mathscr{V}$ and cofibrant tensor units

In this special case our results admit easier proofs which do not need the sophisticated homotopical notions for functors introduced below.

Proof of Proposition 9.2(2) By Lemma 1.1 and the usual inductive transfinite composition and retract argument, this boils down to proving that $\mathcal{O}_A(t)\otimes f^{\odot t}$ is a cofibration for all $t\geq 1$. If the sequence $\mathcal{O}_A$ is cofibrant, this follows from the push-out product axiom. This sequence is cofibrant under the first set of hypotheses, by excellence. The second case is the arity-0 part of Proposition 3.4.1. □

Proof of Proposition 3.4.1 By Lemma 1.2, the usual transfinite composition and retract argument, and the push-out product axiom, it suffices to notice that the map $\mathcal{O}_A(t+n)\otimes f^{\odot t}$ for $t\geq 1$ and $n\geq 0$ is a cofibration provided f is a cofibration and $\mathcal{O}_A$ is a cofibrant sequence, and the sequence $\mathcal{O}_A$ is cofibrant if $A=\mathcal{O}(0)$ is the initial $\mathcal{O}$–algebra, since $\mathcal{O}_{\mathcal{O}(0)}=\mathcal{O}$. □

Proof of Proposition 3.4.2 This follows from the proof of [3, Proposition 5.7], but we here give an argument which extends to our general case. By the aforementioned inductive argument and Lemma 1.2, it is enough to check that the statement holds for $A=\mathcal{O}(0)$ the initial $\mathcal{O}$–algebra and that, assuming the result true for A, the map ϕ_{η_A} induces a weak equivalence of cofibrations $\phi_{\eta_A}(t+n)\otimes f^{\odot t}\colon \mathcal{O}_A(t+n)\otimes f^{\odot t}\to \mathcal{P}_{\phi_*A}(t+n)\otimes f^{\odot t}$, with cofibrant source and target, for f a cofibration as in (1-1). For $A=\mathcal{O}(0)$, $\phi_{\eta_{\mathcal{O}(0)}}=\phi\colon \mathcal{O}\to\mathcal{P}$, which is a weak equivalence by hypothesis. For any cofibrant $\mathcal{O}$–algebra A, $\mathcal{O}_A$ and $\mathcal{P}_{\phi_*A}$ are cofibrant sequences by Proposition 3.4.1. We can assume that $f\colon Y\rightarrowtail Z$ has cofibrant source, replacing it with its push-out $A\rightarrowtail Z\cup_Y A$ along $\bar{g}$ if necessary. Hence, by the push-out product axiom, $\phi_{\eta_A}(t+n)\otimes f^{\odot t}$ is indeed a weak equivalence between cofibrations with cofibrant source and target. □

Proof of Proposition 3.4.3 By the previous inductive argument, we can assume that $\psi=f'$ in (1-1) with f a cofibration. We can also suppose as in the proof of Proposition 3.4.2 that the source of f is cofibrant. By Lemma 1.1, it suffices to notice that

$$\mathcal{O}_\varphi(t)\otimes f^{\odot t}\colon \mathcal{O}_A(t)\otimes f^{\odot t}\to \mathcal{O}_C(t)\otimes f^{\odot t}$$

is a weak equivalence between cofibrations with cofibrant source. Here we use excellence and the push-out product axiom. □

Proof of Proposition 3.4.4 This follows from the fact that $\mathtt{uAss}^{\mathscr{V}}_A(n)=A^{\otimes(n+1)}$ for $n\geq 0$, $\mathtt{Ass}^{\mathscr{V}}_A(0)=A$ and $\mathtt{Ass}^{\mathscr{V}}_A(n)=(A\amalg\mathbb{I})^{\otimes(n+1)}$ for $n\geq 1$, and, for $\mathcal{O}$ the initial operad, $\mathcal{O}_A(0)=A$, $\mathcal{O}_A(1)=\mathbb{I}$ and $\mathcal{O}_A(n)=\varnothing$ for $n\geq 2$. □

Proof of Proposition 3.4.5 As in previous proofs, we can assume that f' fits into a push-out square (1-3) with f a cofibration between cofibrant sequences. Let A be an

$\mathcal{O}$–algebra with underlying cofibrant object and $\varphi\colon A \to B$ a weak equivalence between such $\mathcal{O}$–algebras. By Lemma 1.3, it is enough to notice that, a push-out product of maps $f(n)$ tensored with objects $\mathcal{O}_A(n)$ for $n \geq 0$ is a cofibration between cofibrant objects. Moreover, if we replace $\mathcal{O}_A(n)$ with $\mathcal{O}_\varphi(n)$ we get a weak equivalence between these cofibrations. We are using here the push-out product axiom and the excellence assumption. □

Proof of Proposition 3.4.6 Under the standing hypotheses of this subsection, $\mathscr{C} = \mathscr{V}$, $\mathscr{D} = \mathscr{W}$, and $\chi_{\mathcal{O},A}\colon F(\mathcal{O}_A) \to F^{\mathrm{oper}}(\mathcal{O})_{F_{\mathcal{O}}(A)}$ is a map of sequences in $\mathscr{W}$. If A is the initial $\mathcal{O}$–algebra then $\chi_{\mathcal{O},A} = \chi_{\mathcal{O}}$, so the statement follows from [11, Proposition 4.2]. For general cofibrant $\mathcal{O}$–algebras, using the inductive argument and Lemma 1.2, it suffices to notice that, if the result holds for A and f is a cofibration between cofibrant objects in $\mathscr{V}$, then the map $F(\mathcal{O}_A(t+n)\otimes f^{\odot t}) \to F^{\mathrm{oper}}(\mathcal{O})_{F_{\mathcal{O}}(A)}(t+n)\otimes F(f)^{\odot t}$ for $t \geq 1$ and $n \geq 0$ induced by the comultiplication of F and $\chi_{\mathcal{O},A}$ is a weak equivalence between cofibrations with cofibrant source (and target). Here we are using the push-out product axiom and the cofibrancy results in Proposition 3.4.1 and [11, Corollary 3.8 and Lemma 4.3]. □

3.6 Proofs for $\mathscr{C} \neq \mathscr{V}$ and cofibrant tensor units

The Reedy model structure [7, Section 15.3] on the category of diagrams indexed by $\mathbf{2}^n$ can be generalized as follows.

Proposition 3.6.1 *If $\mathscr{M}$ is a model category and $S \subset \{1, \dots, n\}$, there is a model structure $\mathscr{M}^{\mathbf{2}^n}_S$ on the diagram category $\mathscr{M}^{\mathbf{2}^n}$ such that a map $\tau\colon F \to G$ is*

- *a fibration if $\tau(x_1, \dots, x_n)\colon F(x_1, \dots, x_n) \to G(x_1, \dots, x_n)$ is a fibration in $\mathscr{M}$ for all $(x_1, \dots, x_n)$ in $\mathbf{2}^n$,*
- *a weak equivalence if $\tau(x_1, \dots, x_n)$ is a weak equivalence in $\mathscr{M}$ for all $(x_1, \dots, x_n)$ in $\mathbf{2}^n$ with $x_i = 0$ if $i \in S$, and*
- *a cofibration if the relative latching map of τ at any $(x_1, \dots, x_n)$ in $\mathbf{2}^n$ is a cofibration, and moreover a trivial cofibration if $x_i = 1$ for some $i \in S$.*

Note that $\mathscr{M}^{\mathbf{2}^n}_S$ is a right Bousfield localization of the Reedy model structure $\mathscr{M}^{\mathbf{2}^n}_\varnothing$. Cofibrant diagrams take cofibrant values and have cofibrant latching objects. Moreover, any weak equivalence between cofibrant functors induces weak equivalences between latching objects.

Given model categories $\mathscr{M}$ and $\mathscr{N}$, we introduce some naive homotopical notions for functors of several variables between them. They rely on the previous model structures, hence many facts from ordinary model categories extend to these big functor categories.

Definition 3.6.2 A natural transformation $\tau\colon F \to G$ between functors $\mathscr{N}^n \to \mathscr{M}$ is a *weak equivalence*, *fibration* or *cofibration* if, given cofibrations between cofibrant objects $g_1, \dots, g_n$ in $\mathscr{N}$, $\tau(g_1, \dots, g_n)$ has that property in $\mathscr{M}_S^{\mathbf{2}^n}$ for any $S \subset \{1, \dots, n\}$ such that g_i is a trivial cofibration if $i \in S$. These notions extend aritywise to sequences of functors $F(n)\colon \mathscr{N}^n \to \mathscr{M}$ for $n \geq 0$.

Weak equivalences and fibrations can be just characterized pointwise on cofibrant objects. The condition on cofibrations is stronger. For $n = 0$ we recover the original notions in $\mathscr{M}$. Cofibrant functors preserve cofibrant objects and weak equivalences between them.

When $\mathscr{M} = \mathscr{N}$, we can horizontally compose functors of several variables $\mathscr{M}^n \to \mathscr{M}$, ie $F(\dots, G, \dots)$, and natural transformations between them. Weak equivalences are preserved by horizontal composition if source and target are cofibrant. Cofibrant functors are also preserved, provided we compose at a colimit-preserving slot. All slots in enveloping functor-operads preserve colimits.

The meaning of Definition 3.1 is now clear in the general case for cofibrant tensor units. The proofs in the previous subsection extend straightforwardly, using Lemmas 8.1 above, 2.1 and 2.2 instead of Lemmas 1.1, 1.2 and 1.3, respectively.

3.7 Noncofibrant tensor units

In the proofs of Section 3.5 we have used that the tensor unit is cofibrant at some places. This hypothesis can be relaxed using the theory of *pseudocofibrant* and $\mathbb{I}$*–cofibrant* objects developed in [11, Appendices A and B] provided our monoidal model categories satisfy the *strong unit axiom* and all left Quillen functors satisfy the *pseudocofibrant* and $\mathbb{I}$*–cofibrant axioms*. These will be standing assumptions. We recently learned that pseudocofibrant objects were previously introduced in [8], where they are called semicofibrant.

For $\mathscr{C} = \mathscr{V}$, Definition 3.1 must be modified replacing cofibrancy with pseudocofibrancy. Proposition 3.4.4 holds with the same proof. Essentially the same proofs work for Propositions 3.4.2 and 3.4.3 if we only demand underlying pseudocofibrant objects. Moreover, if we only make pseudocofibrancy hypotheses in Propositions 3.4.1 and 3.4.5, we obtain pseudocofibrant outcomes and honest cofibrations between them. Proposition 3.4.6 holds without $\mathbb{I}$ being cofibrant under our standing assumptions (using [11, Corollary C.3 and Lemma B.14] in the proof). The proof of Corollary D.2 is similar to the proof of Proposition 9.2(2) above.

For $\mathscr{C} \neq \mathscr{V}$, we need new and modified homotopical notions in functor categories. *Pseudocofibrant* and $\mathbb{I}$*–cofibrant* objects F in diagram categories $\mathscr{M}_S^{\mathbf{2}^n}$ are defined by the existence of a cofibration $X \rightarrowtail F$ from a constant diagram on an object X satisfying the corresponding property in $\mathscr{M}$ (which must be monoidal). In Definition 3.6.2, we allow the sources of the g_i to be pseudocofibrant.

The operads in Proposition 3.4.4 are excellent since

$$\mathtt{uAss}^{\mathscr{V}}_A(n)(X_1,\dots,X_n) = A \otimes \bigotimes_{i=1}^{n}(X_i \otimes A) \quad \text{for } n \geq 0,$$

$\mathtt{Ass}^{\mathscr{V}}_A(n)(X_1,\dots,X_n) = (A \amalg \mathbb{I}) \otimes \bigotimes_{i=1}^{n}(X_i \otimes (A \amalg \mathbb{I}))$ for $n \geq 1$ and, for $\mathcal{O}$ the initial operad, $\mathcal{O}_A(1)(X_1) = X_1$ and $\mathcal{O}_A(n)(X_1,\dots,X_n) = \varnothing$ for $n \geq 2$. In Proposition 3.4.1, if we only demand that $\mathcal{O}$ is pseudocofibrant in $\mathscr{C}$ we obtain as outcomes functor-operads with underlying pseudocofibrant sequences and cofibrations between them. In Proposition 3.4.2, the map of sequences in $\mathscr{C}$ underlying ϕ must be a weak equivalence between pseudocofibrant objects. Propositions 3.4.3 and 3.4.5 are true when the underlying objects are pseudocofibrant in $\mathscr{C}$. The analog of Proposition 9.2(2) for noncofibrant $\mathbb{I}$ is Corollary D.2 above.

For the proof of Proposition 3.4.6 without cofibrant tensor units, we must modify again the homotopical notions in Definition 3.6.2, allowing the sources of the g_i to be just ($\mathbb{I}$–)cofibrant. The natural transformations $\chi_{\mathcal{O},A}(n)$ are weak equivalences in this sense, ie when evaluated at ($\mathbb{I}$–)cofibrant objects.

References

[1] **M A Batanin**, *Homotopy coherent category theory and A_∞–structures in monoidal categories*, J. Pure Appl. Algebra 123 (1998) 67–103 MR

[2] **M Batanin**, **M Weber**, *Algebras of higher operads as enriched categories*, Appl. Categ. Structures 19 (2011) 93–135 MR

[3] **C Berger**, **I Moerdijk**, *Axiomatic homotopy theory for operads*, Comment. Math. Helv. 78 (2003) 805–831 MR

[4] **B Fresse**, *Koszul duality of operads and homology of partition posets*, from "Homotopy theory: relations with algebraic geometry, group cohomology, and algebraic K–theory" (P Goerss, S Priddy, editors), Contemp. Math. 346, Amer. Math. Soc., Providence, RI (2004) 115–215 MR

[5] **E Getzler**, **J Jones**, *Operads, homotopy algebra and iterated integrals for double loop spaces*, preprint (1994) arXiv

[6] **J E Harper**, *Homotopy theory of modules over operads and non-Σ operads in monoidal model categories*, J. Pure Appl. Algebra 214 (2010) 1407–1434 MR

[7] **P S Hirschhorn**, *Model categories and their localizations*, Mathematical Surveys and Monographs 99, Amer. Math. Soc., Providence, RI (2003) MR

[8] **L G Lewis, Jr**, **M A Mandell**, *Modules in monoidal model categories*, J. Pure Appl. Algebra 210 (2007) 395–421 MR

[9] **J E McClure**, **J H Smith**, *Cosimplicial objects and little n–cubes, I*, Amer. J. Math. 126 (2004) 1109–1153 MR

[10] **F Muro**, *Homotopy theory of nonsymmetric operads*, Algebr. Geom. Topol. 11 (2011) 1541–1599 MR

[11] **F Muro**, *Homotopy theory of non-symmetric operads, II: Change of base category and left properness*, Algebr. Geom. Topol. 14 (2014) 229–281 MR

[12] **S Schwede**, **B E Shipley**, *Algebras and modules in monoidal model categories*, Proc. London Math. Soc. 80 (2000) 491–511 MR

Facultad de Matemáticas, Departamento de Álgebra, Universidad de Sevilla
Sevilla, Spain

fmuro@us.es

http://personal.us.es/fmuro

Received: 24 July 2015 Revised: 8 March 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 17 Issue 6 (pages 3213–3852) 2017

1472-2747(2017)17:6;1-Q