# Algebras for enriched ∞-operads

RUNE HAUGSENG

Using the description of enriched ∞-operads as associative algebras in symmetric sequences, we define algebras for enriched ∞-operads as certain modules in symmetric sequences. For **V** a symmetric monoidal model category and **O** a Σ-cofibrant operad in **V** for which the model structure on **V** can be lifted to one on **O**-algebras, we then prove that strict algebras in **V** are equivalent to ∞-categorical algebras in the symmetric monoidal ∞-category associated to **V**. We also show that for an ∞-operad $\mathcal{O}$ enriched in a suitable closed symmetric monoidal ∞-category $\mathcal{V}$, we can equivalently describe $\mathcal{O}$-algebras in $\mathcal{V}$ as morphisms of ∞-operads from $\mathcal{O}$ to a self-enrichment of $\mathcal{V}$.

## 1 Introduction

If **V** is a symmetric monoidal category whose tensor product is compatible with colimits, then (one-object) *operads* enriched in **V** can be described as associative algebras in $\mathrm{Fun}(\mathbb{F}^{\simeq}, \mathbf{V})$, the category of *symmetric sequences* in **V**. (We focus on the one-object case for simplicity, but similar descriptions apply to (∞-)operads with any fixed set (space) of objects.) Here $\mathbb{F}^{\simeq}$ denotes the groupoid $\coprod_n B\Sigma_n$ of finite sets and bijections, and the monoidal structure on symmetric sequences is the *composition product*, which is a monoidal structure given by

$$(X \odot Y)(n) \cong \coprod_{k=0}^{\infty} \Big( \coprod_{i_1 + \cdots + i_k = n} (Y(i_1) \otimes \cdots \otimes Y(i_k)) \times_{\Sigma_{i_1} \times \cdots \times \Sigma_{i_k}} \Sigma_n \Big) \otimes_{\Sigma_k} X(k).$$

In [8] we proved that (one-object) ∞-operads enriched in a suitable symmetric monoidal ∞-category $\mathcal{V}$ admit a similar description, as associative algebras in $\mathrm{Fun}(\mathbb{F}^{\simeq}, \mathcal{V})$ using a monoidal structure given by the same formula.

Our goal in this short paper is to use this description of ∞-operads to study *algebras* for enriched ∞-operads. Classically, if **O** is a (one-object) **V**-operad, then an **O**-*algebra* in **V** consists of an object $A \in \mathbf{V}$ and $\Sigma_n$-equivariant morphisms

$$A^{\otimes n} \otimes \mathbf{O}(n) \to A$$

compatible with the composition and unit of **O**. This data can be packaged in a convenient way using the composition product: an **O**-algebra is the same thing as a right[1] **O**-module $M$ in $\mathrm{Fun}(\mathbb{F}^{\simeq}, \mathbf{V})$ that is concentrated in degree zero, ie $M(n)$ is the initial object $\varnothing$ for $n \neq 0$. Indeed, such a right **O**-module is given by a morphism

$$M \odot \mathbf{O} \to M,$$

and expanding out the composition product we see that (since $M(n)$ vanishes for $n \neq 0$) this is precisely given by a map

$$\coprod_k M(0)^{\otimes k} \otimes_{\Sigma_k} \mathbf{O}(k) \to M(0).$$

Here we take the corresponding modules in the $\infty$-categorical setting (and their analogues for many-object operads) as a *definition* of algebras for enriched $\infty$-operads. For a symmetric monoidal $\infty$-category $\mathcal{V}$ (whose tensor product is compatible with colimits indexed by $\infty$-groupoids) and a $\mathcal{V}$-enriched $\infty$-operad $\mathcal{O}$, this results in an $\infty$-category $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})$ with several pleasant properties, including the expected formula for free $\mathcal{O}$-algebras, as we will see in Section 3 after reviewing the results of [8] in Section 2.

We then prove two main results about this $\infty$-categorical notion of $\mathcal{O}$-algebras. First, in Section 4 we prove a rectification result for algebras over operads enriched in a symmetric monoidal model category:

**Theorem 1.1**  (see Theorem 4.10)   *Let $\mathbf{V}$ be a symmetric monoidal model category (with cofibrant unit) and $\mathbf{O}$ an $S$-colored $\Sigma$-cofibrant $\mathbf{V}$-operad such that the category $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})$ admits a model structure whose weak equivalences and fibrations are detected by the forgetful functor to $\mathrm{Fun}(S, \mathbf{V})$. Then this model category describes the $\infty$-category of algebras for this operad in the $\infty$-categorical localization of $\mathbf{V}$. That is, there is an equivalence of $\infty$-categories*

$$\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})[W_{\mathbf{O}}^{-1}] \simeq \mathrm{Alg}_{\mathcal{O}}(\mathcal{V}),$$

*where on the left $W_{\mathbf{O}}$ denotes the collection of weak equivalences between $\mathbf{O}$-algebras, and on the right $\mathcal{V} := \mathbf{V}[W^{-1}]$ is the localization of $\mathbf{V}$ at its weak equivalences and $\mathcal{O}$ denotes $\mathbf{O}$ viewed an $\infty$-operad enriched in $\mathcal{V}$ via the localization functor.*

The comparison applies, for instance, to *all* $\Sigma$-cofibrant operads in chain complexes over a field of characteristic zero or in simplicial sets.[2] We in fact prove a slightly more general result that avoids the assumption that the unit is cofibrant, which applies to all $\Sigma$-cofibrant operads in symmetric spectra. The proof boils down to a combination of model-categorical results of Pavlov and Scholbach and our formula for free algebras, using the same strategy as [12, Theorems 4.1.4.4; 14, Theorem 7.10] to prove that both sides are $\infty$-categories of algebras for equivalent monads.

---

[1]This is correct under our convention for the ordering of the composition product, chosen to be compatible with our construction of the $\infty$-categorical version; in most references on ordinary operads the reverse ordering is used, so that **O**-algebras are certain *left* **O**-modules.

[2]For more general model categories, including chain complexes in positive characteristic, there is typically only a *semi*model structure on algebras over a $\Sigma$-cofibrant operad; see Remark 4.14 for more discussion of this case.

Another classical description of algebras over (one-object) **V**-operads uses *endomorphism operads*: For $v$ an object of **V** there is an operad $\mathrm{End}_{\mathbf{V}}(v)$ with $n$-ary operations given by the internal Hom $\mathrm{HOM}_{\mathbf{V}}(v^{\otimes n}, v)$ (where the $\Sigma_n$-action permutes the factors in $v^{\otimes n}$). If **O** is a (one-object) **V**-operad then we can describe **O**-algebras in **V** with underlying object $v$ as morphisms of one-object operads $\mathbf{O} \to \mathrm{End}_{\mathbf{V}}(v)$. More generally, we can consider an $S$-colored endomorphism operad $\mathrm{End}_{\mathbf{V}}(f)$ for any map of sets $S \to \mathrm{ob}\,\mathbf{V}$, where operations from $(s_1, \ldots, s_n)$ to $s'$ are given by $\mathrm{HOM}_{\mathbf{V}}(f(s_1) \otimes \cdots \otimes f(s_n), f(s'))$; for an $S$-colored operad **O**, we can then identify **O**-algebras in **V** given by $f$ on objects with morphisms of $S$-colored operads from **O** to $\mathrm{End}_{\mathbf{V}}(f)$.

In Section 5 we will use Lurie's construction of endomorphism algebras [12, Section 4.7.1], following work of Hinich [10] in the case of enriched ∞-categories, to construct endomorphism ∞-operads $\mathrm{End}_{\mathcal{V}}(f)$ for any map of ∞-groupoids $f : X \to \mathcal{V}^{\simeq}$, where $\mathcal{V}$ is a closed symmetric monoidal ∞-category compatible with colimits indexed by small ∞-groupoids. Moreover, we show that these endomorphism ∞-operads can be combined into a self-enrichment of $\mathcal{V}$, which gives our second main result:

**Theorem 1.2** (see Theorem 5.12) *Let $\mathcal{V}$ be a closed symmetric monoidal ∞-category compatible with colimits indexed by small ∞-groupoids. Then there exists a $\mathcal{V}$-∞-operad $\overline{\mathcal{V}}$, whose object of multimorphisms from $(v_1, \ldots, v_n)$ to $w$ is the internal Hom $\mathrm{MAP}_{\mathcal{V}}(v_1 \otimes \cdots \otimes v_n, w)$, such that for a $\mathcal{V}$-∞-operad $\mathcal{O}$ there is a natural equivalence of ∞-groupoids*

$$\{\mathcal{O}\text{-algebras in } \mathcal{V}\} \simeq \{\text{morphisms of } \mathcal{V}\text{-}\infty\text{-operads } \mathcal{O} \to \overline{\mathcal{V}}\}.$$

**Warning 1.3** Throughout this paper we use the term "$\mathcal{V}$-∞-operad" to refer to the *algebraic* notion of an ∞-operad, given by objects of multimorphisms in $\mathcal{V}$ with homotopy-coherently associative and unital composition operations. Thus we have a class of fully faithful and essentially surjective morphisms between enriched ∞-operads that we would have to invert to get the "correct" ∞-category of $\mathcal{V}$-∞-operads. In terms of the description of $\mathcal{V}$-∞-operads we use, this means we are not requiring these to be "complete" (see [4, Section 3]). In the terminology of [1], our enriched ∞-operads can be thought of as being *flagged* enriched ∞-operads, meaning a (complete) enriched ∞-operad equipped with an essentially surjective morphism of ∞-groupoids to its space of objects. As we will see in Remark 3.10, the ∞-categories of algebras we will study are in fact invariant under fully faithful and essentially surjective maps of enriched ∞-operads, so it does not really make a difference whether we use complete objects or not.

## 1.1 Related work

Much of our work here is not particularly reliant on the specific construction of the composition product from [8]. An alternative construction, using the description of symmetric sequences in $\mathcal{V}$ as the free presentably symmetric monoidal ∞-category on $\mathcal{V}$ and generalizing the approach to 1-categorical operads due to Trimble [19] and Carboni, has been worked out by Brantner [3]; however, this construction of ∞-operads has not yet been compared to any of the other approaches. In the setting of dendroidal sets,

Heuts describes algebras valued in spaces and $\infty$-categories in terms of dendroidal versions of left and cocartesian fibrations in [9].

## Acknowledgments

## 2 $\infty$-operads as algebras

In this section we will review the main results on enriched $\infty$-operads from [8], where we showed that enriched $\infty$-operads can be viewed as associative algebras in a double $\infty$-category of symmetric collections. For this the relevant notion of enriched $\infty$-operads is an enriched variant of Barwick's definition of $\infty$-operads [2] as presheaves on a category $\mathbf{\Delta}_{\mathbb{F}}$ satisfying Segal (and completeness) conditions; this definition was first introduced in [4]. We will now briefly review this definition, as well as a slight generalization considered in [8] that we will use to define modules over enriched $\infty$-operads; we start by recalling the definition of Barwick's category $\mathbf{\Delta}_{\mathbb{F}}$:

**Definition 2.1** Let $\mathbb{F}$ denote a skeleton of the category of finite sets, with objects $\mathbf{k} := \{1, \ldots, k\}$, $k = 0, 1, \ldots$. We write $\mathbf{\Delta}_{\mathbb{F}}$ for the category whose objects are pairs $([n], f\colon [n] \to \mathbb{F})$, with a morphism $([n], f) \to ([m], g)$ given by a morphism $\phi\colon [n] \to [m]$ in $\mathbf{\Delta}$ and a natural transformation $\eta\colon f \to g \circ \phi$ such that

(i) the map $\eta_i\colon f(i) \to g(\phi(i))$ is injective for all $i = 0, \ldots, m$,

(ii) the commutative square

$$
\begin{array}{ccc}
f(i) & \xrightarrow{\ \eta_i\ } & g(\phi(i)) \\
\downarrow & & \downarrow \\
f(j) & \xrightarrow{\ \eta_j\ } & g(\phi(j))
\end{array}
$$

is cartesian for all $0 \le i \le j \le m$.

There is an obvious projection $\mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}} \to \mathbf{\Delta}^{\mathrm{op}}$; this is a double $\infty$-category. There is also a functor $V\colon \mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}} \to \mathbb{F}_*$ which takes $([n], f)$ to $\left(\coprod_{i=1}^{n} f(i)\right)_+$; see [4, Definition 2.2.11] for a complete definition.

**Remark 2.2** An object of $\mathbf{\Delta}_{\mathbb{F}}$ is a sequence of maps of finite sets

$$
\mathbf{a}_0 \xrightarrow{f_1} \mathbf{a}_1 \to \cdots \xrightarrow{f_n} \mathbf{a}_n,
$$

which we can think of as a forest of $|\mathbf{a}_n|$ oriented trees with $n$ levels: the edges are the elements of all the sets $\mathbf{a}_i$, and since each vertex has a unique outgoing edge we can also think of the elements of $\mathbf{a}_i$ for $i > 0$ as the vertices; the function $f_i$ assigns to each edge in level $i - 1$ the vertex of which it is an incoming edge. This means the functor $V$ takes each forest to its set of vertices (with a disjoint base point). The morphisms in $\mathbf{\Delta}_{\mathbb{F}}$ are defined so that a vertex in the source is mapped to a subtree of the target with the same number of incoming edges.

**Definition 2.3** For $X \in \mathcal{S}$, we write $\mathbf{\Delta}_{\mathbb{F},X}^{\mathrm{op}} \to \mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$ for the left fibration corresponding to the functor $\mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}} \to \mathcal{S}$ obtained as the right Kan extension of the functor $* \to \mathcal{S}$ with value $X$ along the inclusion $\{([0], \mathbf{1})\} \hookrightarrow \mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$.

**Remark 2.4** The fiber of $\mathbf{\Delta}_{\mathbb{F},X}^{\mathrm{op}} \to \mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$ at an object $F$ is equivalent to a product of copies of $X$ indexed by the number of edges in the forest $F$; we can thus think of an object of $\mathbf{\Delta}_{\mathbb{F},X}^{\mathrm{op}}$ as a forest whose edges are labeled by points of $X$.

**Notation 2.5** If $\mathcal{O}$ is a nonsymmetric ∞-operad[3], we write $\mathcal{O}_{\mathbb{F}} := \mathcal{O} \times_{\mathbf{\Delta}^{\mathrm{op}}} \mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$ and $\mathcal{O}_{\mathbb{F},X} := \mathcal{O} \times_{\mathbf{\Delta}^{\mathrm{op}}} \mathbf{\Delta}_{\mathbb{F},X}^{\mathrm{op}}$.

**Definition 2.6** We say a morphism in $\mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$ is *operadic inert* if it lies over an inert morphism in $\mathbf{\Delta}^{\mathrm{op}}$. We then call a morphism in $\mathbf{\Delta}_{\mathbb{F},X}^{\mathrm{op}}$ *operadic inert* if it is a (necessarily cocartesian) morphism over an operadic inert morphism in $\mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$. If $\mathcal{O}$ is a nonsymmetric ∞-operad, we similarly say a morphism in $\mathcal{O}_{\mathbb{F},X}$ is *operadic inert* if it maps to an inert morphism in $\mathcal{O}$ and an operadic inert morphism in $\mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}}$. The functor $V : \mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}} \to \mathbb{F}_*$ takes operadic inert morphisms to inert morphisms, hence if $\mathcal{V}$ is a symmetric monoidal ∞-category we can define an *operadic algebra* for $\mathcal{O}_{\mathbb{F},X}$ in $\mathcal{V}$ to be a commutative square

$$
\begin{array}{ccc}
\mathcal{O}_{\mathbb{F},X} & \xrightarrow{\;A\;} & \mathcal{V}^{\otimes} \\
\downarrow & & \downarrow \\
\mathbf{\Delta}_{\mathbb{F}}^{\mathrm{op}} & \xrightarrow{\;V\;} & \mathbb{F}_*
\end{array}
$$

such that $A$ takes operadic inert morphisms to inert morphisms in $\mathcal{V}^{\otimes}$. We write $\mathrm{Alg}_{\mathcal{O}_{\mathbb{F},X}}^{\mathrm{opd}}(\mathcal{V})$ for the full subcategory of $\mathrm{Fun}_{/\mathbb{F}_*}(\mathcal{O}_{\mathbb{F},X}, \mathcal{V}^{\otimes})$ spanned by the operadic algebras. We also write $\mathrm{Algd}_{\mathcal{O}_{\mathbb{F}}}^{\mathrm{opd}}(\mathcal{V}) \to \mathcal{S}$ for the cartesian fibration corresponding to the functor $X \mapsto \mathrm{Alg}_{\mathcal{O}_{\mathbb{F},X}}^{\mathrm{opd}}(\mathcal{V})$ and refer to its objects as *operadic $\mathcal{O}_{\mathbb{F}}$-algebroids* in $\mathcal{V}$.

**Remark 2.7** The condition for a commutative triangle

$$
\begin{array}{ccc}
\mathbf{\Delta}_{\mathbb{F},X}^{\mathrm{op}} & \xrightarrow{\qquad F \qquad} & \mathcal{V}^{\otimes} \\
& \searrow \qquad \swarrow & \\
& \mathbb{F}_* &
\end{array}
$$

---

[3]We do not review the definition here, as it will not really play a role in this paper: the only examples we will encounter are $\mathbf{\Delta}^{\mathrm{op}}$ and the nonsymmetric operad for right modules, which we describe explicitly in Definition 3.1. We refer the reader to [8, Section 2.1] for a brief review, or [6, Section 2.2] for more motivation and discussion of the definition.

to be an operadic algebra is essentially that the value of $F$ at a forest whose edges are labeled by points of $X$ consists of a list of the values of $F$ at the corollas (one-vertex subtrees) of the forest. The latter should be thought of as the objects of multimorphisms in a $\mathcal{V}$-enriched $\infty$-operad, whose homotopy-coherent composition is encoded by the rest of the data in $F$. Indeed, operadic algebras for $\mathbf{\Delta}^{\mathrm{op}}_{\mathbb{F},X}$ gives one of the notions of enriched $\infty$-operads introduced in [4], which justifies the following notation:

**Notation 2.8**  For a space $X$, we write

$$\mathrm{Opd}_X(\mathcal{V}) := \mathrm{Alg}^{\mathrm{opd}}_{\mathbf{\Delta}^{\mathrm{op}}_{\mathbb{F},X}}(\mathcal{V}).$$

We also write $\mathrm{Opd}(\mathcal{V}) := \mathrm{Algd}^{\mathrm{opd}}_{\mathbf{\Delta}^{\mathrm{op}}_{\mathbb{F}}}(\mathcal{V})$, so that we have a cartesian fibration $\mathrm{Opd}(\mathcal{V}) \to \mathcal{S}$ whose fiber at $X$ is $\mathrm{Opd}_X(\mathcal{V})$.

**Notation 2.9**  For $X, Y \in \mathcal{S}$, we write $\mathbb{F}^{\simeq}_{X,Y}$ for the $\infty$-groupoid $\coprod_{n=0}^{\infty} X^{\times n}_{h\Sigma_n} \times Y$. For a functor $\Phi \colon \mathbb{F}^{\simeq}_{X,Y} \to \mathcal{V}$ we will denote its value at $((x_1, \dots, x_n), y)$ by $\Phi\binom{x_1, \dots, x_n}{y}$. We also abbreviate $\mathbb{F}^{\simeq}_X := \mathbb{F}^{\simeq}_{X,X}$ and write $\mathrm{Coll}_X(\mathcal{V}) := \mathrm{Fun}(\mathbb{F}^{\simeq}_X, \mathcal{V})$; we refer to the objects of this $\infty$-category as (symmetric) $X$-*collections* in $\mathcal{V}$.

To state the main result we will use from [8], we need to recall one further definition:

**Definition 2.10**  A *double $\infty$-category* is a cocartesian fibration $\mathcal{F} \to \mathbf{\Delta}^{\mathrm{op}}$ that corresponds to a functor $F \colon \mathbf{\Delta}^{\mathrm{op}} \to \mathrm{Cat}_\infty$ that satisfies the *Segal condition*: for every $n$, the functor

$$F([n]) \to F([1]) \times_{F([0])} \cdots \times_{F([0])} F([1]),$$

induced by the value of $F$ at the inert maps $[0], [1] \to [n]$ in $\mathbf{\Delta}$, is an equivalence. We say a double $\infty$-category is *framed* if the functor $F([1]) \to F([0]) \times F([0])$, induced by the two face maps $[0] \to [1]$, is a cocartesian fibration. If $\mathcal{O}$ is a nonsymmetric $\infty$-operad, then an $\mathcal{O}$-*algebra* in $\mathcal{F}$ is a commutative triangle



such that $A$ preserves cocartesian morphisms over inert maps in $\mathbf{\Delta}^{\mathrm{op}}$.

**Remark 2.11**  If $\mathcal{F} \to \mathbf{\Delta}^{\mathrm{op}}$ is a double $\infty$-category, then we think of this as an $\infty$-categorical version of a double category where

- the objects are the objects of $\mathcal{F}_{[0]}$,
- the vertical morphisms are the morphisms in $\mathcal{F}_{[0]}$,
- the horizontal morphisms are the objects of $\mathcal{F}_{[1]}$,
- the squares are the morphisms in $\mathcal{F}_{[1]}$.

**Theorem 2.12** (see [8, Corollary 4.2.8]) *Suppose $\mathcal{V}$ is a symmetric monoidal ∞-category compatible with colimits indexed by small ∞-groupoids. Then there exists a framed double ∞-category $\mathrm{COLL}(\mathcal{V})$ such that*:

(i) $\mathrm{COLL}(\mathcal{V})_0 \simeq \mathcal{S}$, *ie the objects of $\mathrm{COLL}(\mathcal{V})$ are small ∞-groupoids and the vertical morphisms are morphisms thereof.*

(ii) *A horizontal morphism from $X$ to $Y$ is a functor $\mathbb{F}^{\simeq}_{X,Y} \to \mathcal{V}$.*

(iii) *If $\Phi$ is a horizontal morphism from $X$ to $Y$ and $\Psi$ is one from $Y$ to $Z$ then their composite $\Phi \odot_Y \Psi$ is given by*

$$\Phi \odot_Y \Psi \binom{x_1, \ldots, x_n}{z} \simeq \operatorname*{colim}_{\mathbf{n} \to \mathbf{m} \to \mathbf{1}} \operatorname*{colim}_{(y_i) \in Y^{\times m}} \bigotimes_{i \in \mathbf{m}} \Phi \binom{x_k : k \in \mathbf{n}_i}{y_i} \otimes \Psi \binom{y_1, \ldots, y_k}{z}.$$

(iv) *If $\mathcal{O}$ is any nonsymmetric ∞-operad then there is a natural equivalence*

$$\mathrm{Alg}_{\mathcal{O}}(\mathrm{COLL}(\mathcal{V})) \simeq \mathrm{Algd}^{\mathrm{opd}}_{\mathcal{O}_{\mathbb{F}}}(\mathcal{V}).$$

(v) *If $F \colon \mathcal{V} \to \mathcal{W}$ is a symmetric monoidal functor that preserves colimits indexed by small ∞-groupoids, then composition with $F$ induces a morphism of double ∞-categories $\mathrm{COLL}(\mathcal{V}) \to \mathrm{COLL}(\mathcal{W})$.*

**Remark 2.13** In (iii), the outer colimit is more precisely over the groupoid $\mathrm{Fact}(\mathbf{n} \to \mathbf{1})$ of factorizations $\mathbf{n} \to \mathbf{m} \to \mathbf{1}$, with morphisms given by diagrams
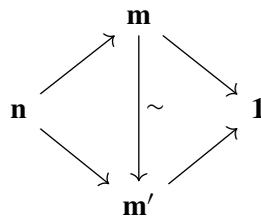


**Remark 2.14** In particular, associative algebras in $\mathrm{COLL}(\mathcal{V})$ are equivalent to ∞-operads enriched in $\mathcal{V}$:

$$\mathrm{Alg}_{\mathbf{\Delta}^{\mathrm{op}}}(\mathrm{COLL}(\mathcal{V})) \simeq \mathrm{Algd}^{\mathrm{opd}}_{\mathbf{\Delta}^{\mathrm{op}}_{\mathbb{F}}}(\mathcal{V}) \simeq \mathrm{Opd}(\mathcal{V}).$$

For $X \in \mathcal{V}$, the ∞-category

$$\mathrm{COLL}(\mathcal{V})(X, X) \simeq \mathrm{Coll}_X(\mathcal{V}) \simeq \mathrm{Fun}(\mathbb{F}^{\simeq}_X, \mathcal{V})$$

of horizontal endomorphisms of $X$ has a monoidal structure given by composition. Moreover, by [8, Proposition 3.4.8] a morphism $f \colon X \to Y$ induces a natural lax monoidal functor $f^* \colon \mathrm{Coll}_Y(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$, given by composition with the induced map $\mathbb{F}^{\simeq}_X \to \mathbb{F}^{\simeq}_Y$. By [8, Corollary 3.4.10] we also have:

**Corollary 2.15** *Let $\mathcal{O}$ be a weakly contractible nonsymmetric $\infty$-operad. Then the functor*

$$\mathrm{Alg}_{\mathcal{O}}(\mathrm{COLL}(\mathcal{V})) \to \mathcal{S}$$

*given by evaluation at $* \in \mathcal{O}_0$ is a cartesian fibration corresponding to the functor $\mathcal{S} \to \mathrm{Cat}_{\infty}$ that takes $X$ to $\mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_X(\mathcal{V}))$ and a morphism $f : X \to Y$ to the functor given by composition with the lax monoidal functor $f^* : \mathrm{Coll}_Y(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$.*

**Remark 2.16** This corollary applies in particular to the weakly contractible nonsymmetric $\infty$-operad $\mathbf{\Delta}^{\mathrm{op}}$, so that by Remark 2.14, enriched $\infty$-operads with $X$ as space of objects are given by associative algebras in $\mathrm{Coll}_X(\mathcal{V})$, ie

$$\mathrm{Opd}_X(\mathcal{V}) \simeq \mathrm{Alg}_{\mathbf{\Delta}^{\mathrm{op}}}(\mathrm{Coll}_X(\mathcal{V})).$$

**Remark 2.17** For $f : X \to Y$, the lax monoidal functor $f^* : \mathrm{Coll}_Y(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$ is given by composition with a morphism of $\infty$-groupoids $f_{\mathbb{F}\simeq} : \mathbb{F}_{\widetilde{X}}^{\simeq} \to \mathbb{F}_{\widetilde{Y}}^{\simeq}$. Since $\mathcal{V}$ has colimits indexed by $\infty$-groupoids, this functor has a left adjoint $f_!$, given by left Kan extension along $f_{\mathbb{F}\simeq}$. Moreover, since $\mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_X(\mathcal{V})) \to \mathrm{Coll}_X(\mathcal{V})$ detects limits and sifted colimits for any nonsymmetric $\infty$-operad $\mathcal{O}$, the functor $f^* : \mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_Y(\mathcal{V})) \to \mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_X(\mathcal{V}))$ preserves limits and sifted colimits, since this is true for $f^* : \mathrm{Coll}_Y(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$. If $\mathcal{V}$ is presentably symmetric monoidal, then the $\infty$-category $\mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_Y(\mathcal{V}))$ is presentable, since it is equivalent to $\mathrm{Alg}_{\mathcal{O}_{\mathbb{F},X}}^{\mathrm{opd}}(\mathcal{V})$, which in turn is equivalent to the $\infty$-category of algebras in $\mathcal{V}$ for some symmetric $\infty$-operad. It then follows from the adjoint functor theorem that $f^* : \mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_Y(\mathcal{V})) \to \mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_X(\mathcal{V}))$ has a left adjoint. This implies:

**Corollary 2.18** *Let $\mathcal{O}$ be a weakly contractible nonsymmetric $\infty$-operad and $\mathcal{V}$ a presentably symmetric monoidal $\infty$-category. Then the functor*

$$\mathrm{Alg}_{\mathcal{O}}(\mathrm{COLL}(\mathcal{V})) \to \mathcal{S}$$

*given by evaluation at $* \in \mathcal{O}_0$ is also a cocartesian fibration.* □

In general the cocartesian morphisms over $f$ are not easily described in terms of the left Kan extension along the map $f_{\mathbb{F}\simeq} : \mathbb{F}_{\widetilde{X}}^{\simeq} \to \mathbb{F}_{\widetilde{Y}}^{\simeq}$. However, we can derive a simple description in the case of monomorphisms of $\infty$-groupoids:

**Proposition 2.19** *Suppose $i : X \hookrightarrow Y$ is a monomorphism of $\infty$-groupoids. Then the left adjoint $i_! : \mathrm{Coll}_X(\mathcal{V}) \to \mathrm{Coll}_Y(\mathcal{V})$ has a canonical monoidal structure, such that composition with $i_!$ and $i^*$ gives for any $\infty$-operad $\mathcal{O}$ an adjunction*

$$i_! : \mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_X(\mathcal{V})) \rightleftarrows \mathrm{Alg}_{\mathcal{O}}(\mathrm{Coll}_Y(\mathcal{V})) : i^*.$$

**Proof** We will prove this by applying [12, Corollary 7.3.2.12], which requires us to show that for $\Phi, \Psi \in \mathrm{Coll}_X(\mathcal{V})$, the canonical map

$$i_!(\Phi \odot_X \Psi) \to i_!\Phi \odot_Y i_!\Psi$$

is an equivalence.

We first describe $i_!\Phi$ more explicitly: For $(y_1, \ldots, y_n)$ in $Y^n_{h\Sigma_n}$, we can identify the fiber of $X^n_{h\Sigma_n}$ over this point as $X_{y_1} \times \cdots \times X_{y_n}$ using the commutative diagram

$$
\begin{array}{ccccc}
X_{y_1} \times \cdots \times X_{y_n} & \longrightarrow & X^n & \longrightarrow & X^n_{h\Sigma_n} \\
\downarrow & \lrcorner & \downarrow & \lrcorner & \downarrow \\
\{(y_1, \ldots, y_n)\} & \longrightarrow & Y^n & \longrightarrow & Y^n_{h\Sigma_n} \\
& & \downarrow & \lrcorner & \downarrow \\
& & * & \longrightarrow & B\Sigma_n
\end{array}
$$

where all three squares are cartesian. Hence the fiber of $\mathbb{F}_X^{\widetilde{\simeq}} \to \mathbb{F}_Y^{\widetilde{\simeq}}$ at $\binom{y_1, \ldots, y_n}{y}$ is equivalent to $\prod_i X_{y_i} \times X_y$, giving

$$
i_!\Phi\binom{y_1, \ldots, y_n}{y} \simeq \operatorname*{colim}_{(x_1, \ldots, x_n, x) \in \prod_i X_{y_i} \times X_y} \Phi\binom{x_1, \ldots, x_n}{x}.
$$

We can then rewrite the formula for $i_!(\Phi \odot_X \Psi)\binom{y_1, \ldots, y_n}{y}$ as

$$
i_!(\Phi \odot_X \Psi)\binom{y_1, \ldots, y_n}{y}
$$
$$
\simeq \operatorname*{colim}_{(x_1, \ldots, x_n, x) \in \prod_i X_{y_i} \times X_y} (\Phi \odot_X \Psi)\binom{x_1, \ldots, x_n}{x}
$$
$$
\simeq \operatorname*{colim}_{(x_1, \ldots, x_n, x) \in \prod_i X_{y_i} \times X_y} \operatorname*{colim}_{\mathbf{n} \to \mathbf{m} \to \mathbf{1}} \operatorname*{colim}_{(x'_j) \in X^m} \bigotimes_j \Phi\binom{x_i : i \in \mathbf{n}_j}{x'_j} \otimes \Psi\binom{x'_1, \ldots, x'_m}{x}
$$
$$
\simeq \operatorname*{colim}_{\mathbf{n} \to \mathbf{m} \to \mathbf{1}} \operatorname*{colim}_{(y'_j) \in Y^m} \operatorname*{colim}_{(x_i) \in \prod_i X_{y_i}} \operatorname*{colim}_{(x'_j) \in \prod_j X_{y'_j}} \operatorname*{colim}_{x \in X_y} \bigotimes_j \Phi\binom{x_i : i \in \mathbf{n}_j}{x'_j} \otimes \Psi\binom{x'_1, \ldots, x'_m}{x}.
$$

On the other hand $(i_!\Phi \odot_Y i_!\Psi)\binom{y_1, \ldots, y_n}{y}$ is equivalent to

$$
\operatorname*{colim}_{\mathbf{n} \to \mathbf{m} \to \mathbf{1}} \operatorname*{colim}_{(y'_j) \in Y^m} \operatorname*{colim}_{(x_i) \in \prod_i X_{y_i}} \operatorname*{colim}_{(x'_j) \in \prod_j X_{y'_j}} \operatorname*{colim}_{(x''_j) \in \prod_j X_{y'_j}} \operatorname*{colim}_{x \in X_y} \bigotimes_j \Phi\binom{x_i : i \in \mathbf{n}_j}{x'_j} \otimes \Psi\binom{x''_1, \ldots, x''_m}{x},
$$

and the canonical map corresponds under these equivalences to the map of colimits arising from the diagonal map $\prod_j X_{y'_j} \to \prod_j X_{y'_j} \times \prod_j X_{y'_j}$. Since these are $\infty$-groupoids, this map is cofinal if and only if it is an equivalence, which holds if and only if the spaces $X_y$ for $y \in Y$ are either contractible or empty, ie if and only if $i$ is a monomorphism. $\square$

**Corollary 2.20** *Let $i\colon X \to Y$ be a monomorphism of $\infty$-groupoids and $\mathcal{O}$ a weakly contractible nonsymmetric $\infty$-operad.*

(i) *For every $A \in \operatorname{Alg}_{\mathcal{O}}(\operatorname{Coll}_X(\mathcal{V}))$, the unit morphism $A \to i^*i_!A$ is an equivalence.*

(ii) *The functor*
$$
i_!\colon \operatorname{Alg}_{\mathcal{O}}(\operatorname{Coll}_X(\mathcal{V})) \to \operatorname{Alg}_{\mathcal{O}}(\operatorname{Coll}_Y(\mathcal{V}))
$$
*is fully faithful.* $\square$

**Remark 2.21** We will also need a more general version of Theorem 2.12, which follows by using part (iii) of [8, Proposition 3.5.6] instead of (vi): If $F \colon \mathcal{V} \to \mathcal{W}$ is a symmetric monoidal functor then composition with $F$ induces a morphism of generalized nonsymmetric $\infty$-operads $F_* \colon \mathrm{COLL}(\mathcal{V}) \to \mathrm{COLL}(\mathcal{W})$, which restricts to lax monoidal functors $F_* \colon \mathrm{Coll}_X(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{W})$. These are compatible with the lax monoidal functors $f^*$ coming from maps of spaces $f \colon X \to Y$: A priori the square

$$
\begin{array}{ccc}
\mathrm{Coll}_Y(\mathcal{V}) & \xrightarrow{\ F_*\ } & \mathrm{Coll}_Y(\mathcal{W}) \\
\downarrow{\scriptstyle f^*} & & \downarrow{\scriptstyle f^*} \\
\mathrm{Coll}_X(\mathcal{V}) & \xrightarrow{\ F_*\ } & \mathrm{Coll}_X(\mathcal{W})
\end{array}
$$

only commutes up to a natural transformation, but this is clearly a natural equivalence since both functors are given by composition.

## 3 Algebras for $\infty$-operads as modules

In this section we define algebras for an enriched $\infty$-operad $\mathcal{O}$ as certain right $\mathcal{O}$-modules in $\mathrm{COLL}(\mathcal{V})$. We first recall the definition of the nonsymmetric $\infty$-operad for right modules, and prove that this is weakly contractible, allowing us to apply Corollary 2.15:

**Definition 3.1** Let **rm** denote the nonsymmetric operad for right modules. This has two objects, $a$ and $m$, and there is a unique multimorphism $(x_1, \ldots, x_n) \to y$ if $x_1 = \cdots = x_n = y = a$ ($n = 0$ allowed) or $x_1 = y = m$ and $x_2 = \cdots = x_n = a$, and no multimorphisms otherwise. We write $\mathrm{RM} \to \boldsymbol{\Delta}^{\mathrm{op}}$ for the corresponding nonsymmetric $\infty$-operad, or in other words the category of operators of **rm**. This has as objects sequences $(x_1, \ldots, x_n)$ with each $x_i$ being either $a$ or $m$, and a morphism $(x_1, \ldots, x_n) \to (y_1, \ldots, y_m)$ is given by a map $\phi \colon [m] \to [n]$ in $\boldsymbol{\Delta}$ and multimorphisms $(x_{\phi(i-1)+1}, \ldots, x_{\phi(i)}) \to y_i$ in **rm**.

**Proposition 3.2** *The category* $\mathrm{RM}$ *is weakly contractible.*

**Proof** In this proof it is convenient to use the notation $(i_0, \ldots, i_n)_{\mathrm{RM}}$ for the object of $\mathrm{RM}$ given by the sequence $(a, \ldots, a, m, \ldots, m, a, \ldots, a)$ where there are $n$ copies of $m$ and $i_t$ copies of $a$ between the $t^{\mathrm{th}}$ and $(t+1)^{\mathrm{st}}$ copy of $m$ (and $i_0$ before the first and $i_n$ after the last). Define a functor $\mu \colon \boldsymbol{\Delta}_{\mathrm{int}}^{\mathrm{op}} \to \mathrm{RM}$ over $\boldsymbol{\Delta}^{\mathrm{op}}$ by taking $[n]$ to the unique object of the form $(0, \ldots, 0)_{\mathrm{RM}} = (m, \ldots, m)$ over $[n]$, and determined on morphisms by the inert morphisms between these objects. We claim that $\mu$ is coinitial, and so in particular a weak homotopy equivalence. To see this, it suffices by [11, Theorem 4.1.3.1] to show that for every object $X \in \mathrm{RM}$ the category $(\boldsymbol{\Delta}_{\mathrm{int}}^{\mathrm{op}})_{/X}$ is weakly contractible. But this category has a terminal object: if $X = (i_0, \ldots, i_n)_{\mathrm{RM}}$ then any morphism $(0, \ldots, 0)_{\mathrm{RM}} \to X$ factors as an inert morphism followed by the (unique) degeneracy $\mu([n]) \to X$. Since $\boldsymbol{\Delta}_{\mathrm{int}}^{\mathrm{op}}$ is weakly contractible (for example, because the inclusion $\boldsymbol{\Delta}_{\mathrm{int}}^{\mathrm{op}} \hookrightarrow \boldsymbol{\Delta}^{\mathrm{op}}$ is cofinal and $\boldsymbol{\Delta}^{\mathrm{op}}$ has an initial object), this implies that $\mathrm{RM}$ is also weakly contractible. $\square$

**Corollary 3.3** *The functor*

$$\mathrm{Algd}^{\mathrm{opd}}_{\mathrm{RM}_{\mathbb{F}}}(\mathcal{V}) \simeq \mathrm{Alg}_{\mathrm{RM}}(\mathrm{COLL}(\mathcal{V})) \to \mathcal{S},$$

*given by evaluation at* $() \in \mathrm{RM}_0$, *is a cartesian fibration corresponding to the functor* $\mathcal{S} \to \mathrm{Cat}_\infty$ *that takes* $X$ *to* $\mathrm{Alg}_{\mathrm{RM}}(\mathrm{Coll}_X(\mathcal{V}))$ *and a morphism* $f \colon X \to Y$ *to the functor given by composition with the lax monoidal functor* $f^* \colon \mathrm{Coll}_Y(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$. □

To define algebras we want to restrict to those modules that are concentrated in degree 0, which will be justified by the next proposition.

**Definition 3.4** We say that $\Phi \in \mathrm{Coll}_X(\mathcal{V})$ is *concentrated in degree* 0 if

$$\Phi\binom{x_1, \ldots, x_n}{y} \simeq \varnothing$$

whenever $n > 0$, where $\varnothing$ denotes the initial object in $\mathcal{V}$.

**Proposition 3.5** *Let* $\mathcal{V}$ *be a symmetric monoidal* ∞*-category compatible with colimits indexed by small* ∞*-groupoids.*

(i)  *The functor* $Z \colon \mathrm{Coll}_X(\mathcal{V}) \to \mathrm{Fun}(X, \mathcal{V})$ *given by composition with* $X \hookrightarrow \mathbb{F}^{\simeq}_X$ *has a fully faithful left adjoint, which identifies* $\mathrm{Fun}(X, \mathcal{V})$ *with the collections that are concentrated in degree* 0.

(ii)  *If* $M \colon \mathbb{F}^{\simeq}_X \to \mathcal{V}$ *is concentrated in degree* 0, *then so is* $M \odot_X N$ *for any* $N \in \mathrm{Coll}_X(\mathcal{V})$.

(iii)  *The composition product induces a right* $\mathrm{Coll}_X(\mathcal{V})$*-module structure on the* ∞*-category* $\mathrm{Fun}(X, \mathcal{V})$.

(iv)  *For* $f \colon X \to Y$, *composition with* $f$ *and the induced functor* $\mathbb{F}^{\simeq}_X \to \mathbb{F}^{\simeq}_Y$ *gives a lax* RM*-monoidal functor*

$$f^* \colon (\mathrm{Fun}(Y, \mathcal{V}), \mathrm{Coll}_Y(\mathcal{V})) \to (\mathrm{Fun}(X, \mathcal{V}), \mathrm{Coll}_X(\mathcal{V}))$$

(v)  *Composition with a symmetric monoidal functor* $F \colon \mathcal{V} \to \mathcal{W}$ *gives a lax* RM*-monoidal functor*

$$F_* \colon (\mathrm{Fun}(X, \mathcal{V}), \mathrm{Coll}_X(\mathcal{V})) \to (\mathrm{Fun}(X, \mathcal{W}), \mathrm{Coll}_X(\mathcal{W})).$$

*If* $F$ *preserves colimits indexed by small* ∞*-groupoids, then* $F_*$ *is an* RM*-monoidal functor.*

**Proof** Part (i) is obvious from the description of $\mathbb{F}^{\simeq}_X$ as $\coprod_n X^{\times n}_{h\Sigma_n} \times X$ and the formula for pointwise left Kan extensions, while part (ii) follows immediately from the description of composition of horizontal morphisms in $\mathrm{COLL}(\mathcal{V})$ in Theorem 2.12. Part (iii) then holds by combining parts (i) and (ii), and parts (iv) and (v) follow by restricting the lax monoidal functors discussed in Section 2. □

**Definition 3.6** Let $\mathcal{O}$ be a $\mathcal{V}$-∞-operad with space of objects $X$, viewed as an associative algebra in $\mathrm{Coll}_X(\mathcal{V})$. An $\mathcal{O}$-*algebra in* $\mathcal{V}$ is a right $\mathcal{O}$-module in $\mathrm{Fun}(X, \mathcal{V})$. We write $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})$ for the ∞-category $\mathrm{RMod}_{\mathcal{O}}(\mathrm{Fun}(X, \mathcal{V}))$ of these right modules.

**Remark 3.7** By Proposition 3.5(iv) we see that for $\mathcal{O} \in \mathrm{Opd}_Y(\mathcal{V})$, composition with $f \colon X \to Y$ gives a functor $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V}) \to \mathrm{Alg}_{f^*\mathcal{O}}(\mathcal{V})$, while composition with a symmetric monoidal functor $F \colon \mathcal{V} \to \mathcal{W}$ gives a functor $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V}) \to \mathrm{Alg}_{F_*\mathcal{O}}(\mathcal{W})$.

Since there is always a formula for free modules, with this definition we immediately get a formula for free algebras over enriched $\infty$-operads:

**Proposition 3.8** *The forgetful functor* $U_{\mathcal{O}} \colon \mathrm{Alg}_{\mathcal{O}}(\mathcal{V}) \to \mathrm{Fun}(X, \mathcal{V})$ *has a left adjoint* $F_{\mathcal{O}}$, *and the endofunctor* $U_{\mathcal{O}} F_{\mathcal{O}}$ *satisfies*

$$U_{\mathcal{O}} F_{\mathcal{O}} M(x) \simeq \coprod_n \operatorname*{colim}_{(x_1,\ldots,x_n) \in X^n_{h\Sigma_n}} M(x_1) \otimes \cdots \otimes M(x_n) \otimes \mathcal{O}\binom{x_1,\ldots,x_n}{x}.$$

*Moreover,* $U_{\mathcal{O}}$ *preserves sifted colimits and the adjunction is monadic.*

**Proof** By [12, Corollary 4.2.4.8] the left adjoint $F_{\mathcal{O}}$ exists, and $U_{\mathcal{O}} F_{\mathcal{O}}(M)$ is given by the composition product $M \odot \mathcal{O}$ (with $M$ viewed as a symmetric sequence concentrated in degree 0). Expanding out this composition product now gives the formula.

It follows from [12, Proposition 4.2.3.1] that $U_{\mathcal{O}}$ detects equivalences and from [12, Corollary 4.2.3.5] that $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})$ has sifted colimits and $U_{\mathcal{O}}$ preserves these, since the composition product preserves sifted colimits in each variable. The adjunction is therefore monadic by the monadicity theorem for $\infty$-categories, [12, Theorem 4.7.3.5]. □

Applying [6, Proposition A.5.9], we get:

**Corollary 3.9** *If* $\mathcal{V}$ *is a presentably symmetric monoidal* $\infty$-*category and* $\mathcal{O}$ *is a* $\mathcal{V}$-*enriched* $\infty$-*operad, then the* $\infty$-*category* $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})$ *is presentable.* □

**Remark 3.10** Let $F \colon \mathcal{O} \to \mathcal{O}'$ be a morphism of $\mathcal{V}$-$\infty$-operads given on spaces of objects by $f \colon X \to Y$, and suppose $f$ is surjective on $\pi_0$ and $F$ is fully faithful in the sense that all the maps

$$\mathcal{O}\binom{x_1,\ldots,x_n}{y} \to \mathcal{O}'\binom{f(x_1),\ldots,f(x_n)}{f(y)}$$

are equivalences in $\mathcal{V}$. Then we have a commutative square

$$
\begin{array}{ccc}
\mathrm{Alg}_{\mathcal{O}'}(\mathcal{V}) & \xrightarrow{\ F^*\ } & \mathrm{Alg}_{\mathcal{O}}(\mathcal{V}) \\
\downarrow & & \downarrow \\
\mathrm{Fun}(Y, \mathcal{V}) & \xrightarrow[f^*]{} & \mathrm{Fun}(X, \mathcal{V})
\end{array}
$$

where the surjectivity of $f$ implies that the composite functor $\mathrm{Alg}_{\mathcal{O}'}(\mathcal{V}) \to \mathrm{Fun}(X, \mathcal{V})$ is a monadic right adjoint. Using the formula from Proposition 3.8 it is easy to see that $F^*$ gives an equivalence of monads on

Fun$(X, \mathcal{V})$ and so gives an equivalence of ∞-categories $\mathrm{Alg}_{\mathcal{O}'}(\mathcal{V}) \simeq \mathrm{Alg}_{\mathcal{O}}(\mathcal{V})$ by [12, Corollary 4.7.3.16]. This applies in particular if $\mathcal{O}'$ is the completion of $\mathcal{O}$, so by a two-out-of-three argument it follows that any fully faithful and essentially surjective morphism of $\mathcal{V}$-∞-operads $F\colon \mathcal{O} \to \mathcal{P}$ induces an equivalence

$$\mathrm{Alg}_{\mathcal{O}}(\mathcal{V}) \simeq \mathrm{Alg}_{\mathcal{P}}(\mathcal{V})$$

on ∞-categories of algebras in $\mathcal{V}$.

We end this section by showing that the nullary operations of a $\mathcal{V}$-∞-operad $\mathcal{O}$ give a canonical $\mathcal{O}$-algebra, using the next observation:

**Proposition 3.11**  $Z\colon \mathrm{Coll}_X(\mathcal{V}) \to \mathrm{Fun}(X, \mathcal{V})$ *is a functor of* $\mathrm{Coll}_X(\mathcal{V})$*-modules.*

**Proof**  By definition of the $\mathrm{Coll}_X(\mathcal{V})$-module structure on $\mathrm{Fun}(X, \mathcal{V})$, the inclusion $\mathrm{Fun}(X, \mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$ is a functor of $\mathrm{Coll}_X(\mathcal{V})$-modules. Using [12, Corollary 7.3.2.7], this implies that its right adjoint $Z$ is a lax RM-monoidal functor. Thus for $M, N \in \mathrm{Coll}_X(\mathcal{V})$ there are natural maps

$$Z(M) \odot_X N \to Z(M \odot_X N);$$

by the formula for $\odot_X$ these maps are equivalences, and so $Z$ is an RM-monoidal functor.  $\square$

**Corollary 3.12**  *If* $\mathcal{O}$ *is an associative algebra in* $\mathrm{Coll}_X(\mathcal{V})$ *and* $M \in \mathrm{Coll}_X(\mathcal{V})$ *is a right* $\mathcal{O}$*-module, then the restriction* $Z(M) \in \mathrm{Fun}(Y, \mathcal{V})$ *is also a right* $\mathcal{O}$*-module.*  $\square$

Since an algebra is canonically a right module over itself, this specializes to:

**Corollary 3.13**  *Suppose* $\mathcal{O}$ *is an algebra in* $\mathrm{Coll}_X(\mathcal{V})$*, ie a* $\mathcal{V}$*-∞-operad with* $X$ *as space of objects. Then the functor* $Z(\mathcal{O})\colon X \to \mathcal{V}$ *picking out the nullary operations is canonically a right* $\mathcal{O}$*-module.*  $\square$

# 4  Comparison with model categories of operad algebras

Let **V** be a symmetric monoidal model category (with cofibrant unit). Then by [12, Proposition 4.1.7.4] the localization $\mathbf{V}[W^{-1}]$ (with $W$ the class of weak equivalences) is a symmetric monoidal ∞-category, and the localization functor $\mathbf{V} \to \mathbf{V}[W^{-1}]$ is symmetric monoidal when restricted to the cofibrant objects. If **O** is a (levelwise cofibrant) operad in **V** then this means we can also view **O** as an operad in $\mathbf{V}[W^{-1}]$. Moreover, in good cases there is a model structure on the category $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})$ of **O**-algebras in **V**. In this section we will give conditions under which the corresponding ∞-category $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})[W_{\mathbf{O}}^{-1}]$ (with $W_{\mathbf{O}}$ the class of weak equivalences of **O**-algebras) is equivalent to the ∞-category $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}])$, defined as in the previous section. In order to do the comparison in sufficient generality to cover examples such as symmetric spectra, we do not want to assume that the unit of the monoidal structure is cofibrant. Instead we consider model categories with a subcategory of *flat* objects in the following sense:

**Definition 4.1** Let $\mathbf{V}$ be a symmetric monoidal model category.[4] A *subcategory of flat objects* is a full subcategory $\mathbf{V}^\flat$ that satisfies the following conditions:

- $\mathbf{V}^\flat$ is a symmetric monoidal subcategory, ie the unit is flat and the tensor product of two flat objects is flat.

- If $X$ is flat and $Y \to Y'$ is a weak equivalence between flat objects, then $X \otimes Y \to X \otimes Y'$ is again a weak equivalence.

- All cofibrant objects are flat.

**Example 4.2** If the unit of $\mathbf{V}$ is cofibrant, then the subcategory $\mathbf{V}^c$ of cofibrant objects is a subcategory of flat objects.

**Proposition 4.3** *Let $\mathbf{V}$ be a symmetric monoidal model category and $\mathbf{V}^\flat$ a subcategory of flat objects. Then the inclusions $\mathbf{V}^c \hookrightarrow \mathbf{V}^\flat \hookrightarrow \mathbf{V}$ induce equivalences of localizations*

$$\mathbf{V}^c[W^{-1}] \xrightarrow{\sim} \mathbf{V}^\flat[W^{-1}] \xrightarrow{\sim} \mathbf{V}[W^{-1}],$$

*where we denote the collections of weak equivalences in the subcategories by $W$ in all cases.*

**Proof** Let $Q \colon \mathbf{V} \to \mathbf{V}$ be a cofibrant replacement functor, with a natural weak equivalence $\eta \colon Q \to \mathrm{id}$. If $i$ denotes the inclusion $\mathbf{V}^c \hookrightarrow \mathbf{V}$ then we may view $Q$ as a functor $\mathbf{V} \to \mathbf{V}^c$ and $\eta$ as a natural transformation $iQ \to \mathrm{id}_\mathbf{V}$. If $X$ is cofibrant, then $\eta_X \colon QX \to X$ is a morphism in $\mathbf{V}^c$, so we may view $\eta i \colon iQi \to i$ as a natural transformation $\eta^c \colon Qi \to \mathrm{id}_{\mathbf{V}^c}$. The functor $Q$ preserves weak equivalences, and both $\eta$ and $\eta^c$ are natural weak equivalences. It follows that $Q$ induces a functor $\mathbf{V}[W^{-1}] \to \mathbf{V}^c[W^{-1}]$ and the transformations $\eta$ and $\eta^c$ induce transformations that exhibit this as an inverse of the functor $\mathbf{V}^c[W^{-1}] \to \mathbf{V}[W^{-1}]$ induced by $i$. The same argument applies to $Q$ restricted to the full subcategory $\mathbf{V}^\flat$; the functor $\mathbf{V}^\flat[W^{-1}] \to \mathbf{V}[W^{-1}]$ is therefore an equivalence by the two-out-of-three property of equivalences. $\square$

**Corollary 4.4** *Let $\mathbf{V}$ be a symmetric monoidal model category and $\mathbf{V}^\flat$ a subcategory of flat objects. Then the $\infty$-category $\mathbf{V}[W^{-1}]$ inherits a symmetric monoidal structure such that the functor $\mathbf{V}^\flat \to \mathbf{V}[W^{-1}]$ is symmetric monoidal.*

**Proof** By assumption, in $\mathbf{V}^\flat$ the tensor product is compatible with weak equivalences, and so the $\infty$-category $\mathbf{V}[W^{-1}] \simeq \mathbf{V}^\flat[W^{-1}]$ inherits a symmetric monoidal structure with this property by [12, Proposition 4.1.7.4]. $\square$

Using Remark 3.7, composition with the symmetric monoidal functor $\mathbf{V}^\flat \to \mathbf{V}[W^{-1}]$ gives a natural functor

$$\mathrm{Alg}_\mathbf{O}(\mathbf{V}^\flat) \to \mathrm{Alg}_\mathbf{O}(\mathbf{V}[W^{-1}]),$$

if $\mathbf{O}$ is a levelwise flat $\mathbf{V}$-operad. Here we can interpret $\mathrm{Alg}_\mathbf{O}(\mathbf{V}^\flat)$ as the classical ordinary category of $\mathbf{O}$-algebras in $\mathbf{V}^\flat$.

---

[4]We assume that model categories have functorial factorizations.

**Definition 4.5** An $S$-colored operad **O** in a symmetric monoidal model category **V** is called *admissible* if there exists a model structure on $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})$ where a morphism is a weak equivalence or a fibration precisely if its underlying morphism in $\mathrm{Fun}(S, \mathbf{V})$ is one (ie it is a weak equivalence or fibration in **V** for each element of $S$).

**Definition 4.6** An $S$-colored **V**-operad **O** is called $\Sigma$-*cofibrant* if the unit map $\mathbb{1}_S \to U(\mathbf{O})$ is a cofibration in the projective model structure on $\mathrm{Fun}(\mathbb{F}_S^{\widetilde{\cong}}, \mathbf{V})$, where $U$ denotes the forgetful functor from operads to collections and $\mathbb{1}_S$ is the monoidal unit for the composition product, given by

$$\mathbb{1}_S \binom{s_1, \ldots, s_n}{s'} = \begin{cases} \mathbb{1}, & n = 1, s_1 = s', \\ \varnothing, & \text{otherwise}, \end{cases}$$

where $\mathbb{1}$ is the monoidal unit in **V**.

**Example 4.7** A one-colored **V**-operad **O** is $\Sigma$-cofibrant precisely if $\mathbb{1} \to \mathbf{O}(1)$ is a cofibration, and the object $\mathbf{O}(n)$ is projectively cofibrant in $\mathrm{Fun}(B\Sigma_n, \mathbf{V})$ for all $n \neq 1$.

**Definition 4.8** Let **V** be a symmetric monoidal model category and $\mathbf{V}^{\flat}$ a subcategory of flat objects. We will say that a **V**-operad **O** is *flat* if it is enriched in the full subcategory $\mathbf{V}^{\flat}$.

**Remark 4.9** Since cofibrant objects are flat, if **O** is $\Sigma$-cofibrant then it is flat precisely if in addition the objects of (unary) endomorphisms $\mathbf{O}(x, x) \in \mathbf{V}$ are all flat.

By [14, Proposition 6.2], if **O** is an admissible $\Sigma$-cofibrant **V**-operad, then cofibrant **O**-algebras have cofibrant underlying objects in **V**. Since cofibrant objects are in particular flat, if **O** is flat, admissible and $\Sigma$-cofibrant we have a functor

$$\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})^c \to \mathrm{Alg}_{\mathbf{O}}(\mathbf{V}^{\flat}) \to \mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}]).$$

This takes weak equivalences in $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})^c$ to equivalences in $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}])$, since the weak equivalences are lifted from the weak equivalences in **V**, and so induces a functor of $\infty$-categories

$$\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})^c[W_{\mathbf{O}}^{-1}] \to \mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}]),$$

where $W_{\mathbf{O}}$ denotes the collection of weak equivalences between **O**-algebras.

**Theorem 4.10** *Let* **V** *be a symmetric monoidal model category equipped with a subcategory* $\mathbf{V}^{\flat}$ *of flat objects. If* **O** *is a flat admissible* $\Sigma$-*cofibrant* **V**-*operad, then the functor*

$$\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})^c[W_{\mathbf{O}}^{-1}] \to \mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}])$$

*is an equivalence of* $\infty$-*categories.*

**Proof**   We follow the proof of [14, Theorem 7.10], which in turn is a variant of those of [12, Theorems 4.1.4.4 and 4.5.4.7]. Let $S$ be the set of objects of $\mathbf{O}$. The right Quillen functor $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V}) \to \mathrm{Fun}(S, \mathbf{V})$ induces a functor of $\infty$-categories $U \colon \mathrm{Alg}_{\mathbf{O}}(\mathbf{V})^c[W_{\mathbf{O}}^{-1}] \to \mathrm{Fun}(S, \mathbf{V}[W^{-1}])$, which is a right adjoint by [13, Theorem 2.1]. As $\mathbf{O}$ is $\Sigma$-cofibrant, the forgetful functor preserves sifted homotopy colimits by [14, Proposition 7.8]. Since it also detects weak equivalences, it follows by [12, Theorem 4.7.3.5] (the monadicity theorem for $\infty$-categories) that $U$ is a monadic right adjoint. The same holds for the forgetful functor $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}]) \to \mathrm{Fun}(S, \mathbf{V}[W^{-1}])$ by Proposition 3.8, so using [12, Corollary 4.7.3.16] we see that to show that the functor $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})^c[W_{\mathbf{O}}^{-1}] \to \mathrm{Alg}_{\mathbf{O}}(\mathbf{V}[W^{-1}])$ is an equivalence it suffices to show that the two associated monads on $\mathrm{Fun}(S, \mathbf{V}[W^{-1}])$ have equivalent underlying endofunctors. This follows from the formula in Proposition 3.8, since the $\Sigma_n$-orbits that appear in the formula for free strict $\mathbf{O}$-algebras are homotopy orbits when $\mathbf{O}$ is $\Sigma$-cofibrant.                                                                                   $\square$

The cases to which this result applies are primarily those where *all* operads are admissible, as more generally we only have semimodel structure on algebras over $\Sigma$-cofibrant operads. This includes the following examples, as discussed in [15, Section 7]:

 (i)   the category $\mathrm{Set}_\Delta$ of simplicial sets, equipped with the Kan–Quillen model structure,

 (ii)  the category Top of compactly generated weak Hausdorff spaces, equipped with the usual model structure,

 (iii) the category $\mathrm{Ch}_k$ of chain complexes of $k$-vector spaces, where $k$ is a field of characteristic 0 (or more generally a ring containing $\mathbb{Q}$), equipped with the projective model structure,

 (iv)  the category $\mathrm{Sp}^\Sigma$ of symmetric spectra, equipped with the positive stable model structure,

In the first three examples the unit is cofibrant, and in the positive stable model structure on symmetric spectra a suitable subcategory of flat objects is supplied by the *S-cofibrant* objects of [17] (see also [16, Chapter 5], where these are called *flat* objects). A $\Sigma$-cofibrant operad in symmetric spectra is necessarily flat, since the flat objects are the cofibrant objects in a model structure whose cofibrations include the usual cofibrations.

Specializing to these cases, we have:

**Corollary 4.11**   (i)   *Let* $\mathbf{O}$ *be a* $\Sigma$-*cofibrant simplicial operad. Then*

$$\mathrm{Alg}_{\mathbf{O}}(\mathrm{Set}_\Delta)[W_{\mathbf{O}}^{-1}] \simeq \mathrm{Alg}_{\mathbf{O}}(\mathcal{S}).$$

(ii)   *Let* $\mathbf{O}$ *be a* $\Sigma$-*cofibrant topological operad. Then*

$$\mathrm{Alg}_{\mathbf{O}}(\mathrm{Top})[W_{\mathbf{O}}^{-1}] \simeq \mathrm{Alg}_{\mathbf{O}}(\mathcal{S}).$$

(iii)   *Let* $\mathbf{O}$ *be a* $\Sigma$-*cofibrant dg-operad over a field* $k$ *of characteristic zero. Then*

$$\mathrm{Alg}_{\mathbf{O}}(\mathrm{Ch}_k)[W_{\mathbf{O}}^{-1}] \simeq \mathrm{Alg}_{\mathbf{O}}(\mathcal{D}(k)),$$

*where* $\mathcal{D}(k)$ *is the derived* $\infty$-*category of* $k$-*modules.*

(iv)   *Let* **O** *be a* $\Sigma$-*cofibrant operad in symmetric spectra. Then*

$$\mathrm{Alg}_{\mathbf{O}}(\mathrm{Sp}^{\Sigma})[W_{\mathbf{O}}^{-1}] \simeq \mathrm{Alg}_{\mathbf{O}}(\mathrm{Sp}),$$

*where* $\mathrm{Sp}$ *is the* ∞-*category of spectra.*

**Remark 4.12**   The case of simplicial operads was already proved as [14, Theorem 7.10].

**Remark 4.13**   According to Spitzweck's thesis [18, Theorem 4], a **V**-operad with a single object that is cofibrant in the semimodel structure on one-object operads in **V** is admissible without further assumptions on **V**. A version for colored operads does not yet seem to appear in the literature, but if this is correct then the comparison of Theorem 4.10 would apply in general for such cofibrant operads.

**Remark 4.14**   If **O** is a $\Sigma$-cofibrant operad in a symmetric monoidal model category **V**, then under much weaker assumptions on **V** there exists a *semimodel* structure on the category $\mathrm{Alg}_{\mathbf{O}}(\mathbf{V})$, by a result of Spitzweck [18, Theorem 5] in the one-object case and White–Yau for colored operads [20, Theorem 6.3.1]. Using results of Cisinski [5], White and Yau have recently extended the results relating structures in model categories to their analogues in ∞-categories needed to carry out the proof of Theorem 4.10 in the setting of semimodel categories, and thereby extended the comparison with ∞-operad algebras to the case where there is only a semimodel structure on algebras over a $\Sigma$-cofibrant operad; see [21, Section 7.3].

# 5 Endomorphism ∞-operads

The first goal of this subsection is to prove that for any morphism of ∞-groupoids $f : X \to \mathcal{V}^{\simeq}$ there exists a corresponding *endomorphism* ∞-*operad* $\mathrm{End}_{\mathcal{V}}(f)$, where $\mathcal{V}$ denotes a closed symmetric monoidal ∞-category compatible with small ∞-groupoid-indexed colimits. Our strategy for obtaining these objects is taken from [10, Section 6.3] and uses the construction of endomorphism algebras from [12, Section 4.7.1], which we first briefly recall:[5]

Suppose $\mathcal{A}$ is a monoidal ∞-category and $\mathcal{M}$ is right-tensored over $\mathcal{A}$. An *endomorphism algebra* for an object $M \in \mathcal{M}$ is an associative algebra $\mathrm{End}(M)$ in $\mathcal{A}$ and a right $\mathrm{End}(M)$-module structure on $M$ with the universal property that for any associative algebra $A$ in $\mathcal{A}$, right $A$-module structures on $M$ are naturally equivalent to morphisms of associative algebras $A \to \mathrm{End}(M)$.

By [12, Proposition 4.7.1.30 and Theorem 4.7.1.34] there exists a monoidal ∞-category $\mathcal{A}[M]$ whose objects are pairs $(X \in \mathcal{A}, M \otimes X \to M$ in $\mathcal{M})$, with the property that an associative algebra in $\mathcal{A}[M]$ corresponds to an associative algebra $A \in \mathcal{A}$ together with a right $A$-module structure on $M$. An endomorphism algebra for $M$ is thus precisely a terminal object in $\mathrm{Alg}_{\mathbf{\Delta}^{\mathrm{op}}}(\mathcal{A}[M])$. Since the terminal object of $\mathcal{A}[M]$ has a unique algebra structure if it exists, we have:

---
[5]We restate it for right instead of left modules.

**Proposition 5.1** [12, Corollary 4.7.1.40]   *If $\mathcal{A}[M]$ has a terminal object $(A, M \otimes A \to M)$ then $A$ is the underlying object of an endomorphism algebra for $M$.* □

By construction the forgetful functor $\mathcal{A}[M] \to \mathcal{A}$ is a right fibration, corresponding to the functor

$$A \mapsto \mathrm{Map}_{\mathcal{M}}(M \otimes A, M).$$

In the case of $\mathrm{Coll}_X(\mathcal{V})$ and its right module $\mathrm{Fun}(X, \mathcal{V})$ we can explicitly identify this functor:

**Proposition 5.2**   *For $M \in \mathrm{Fun}(X, \mathcal{V})$ and $S \in \mathrm{Coll}_X(\mathcal{V})$ there is a natural equivalence*

$$\mathrm{Map}_{\mathrm{Fun}(X, \mathcal{V})}(M \odot S, M) \simeq \mathrm{Map}_{\mathrm{Coll}_X(\mathcal{V})}(S, \mathrm{End}_{\mathcal{V}}(M)),$$

*where $\mathrm{End}_{\mathcal{V}}(M) \colon \mathbb{F}_X^{\widetilde{\simeq}} \to \mathcal{V}$ is the functor given by*

$$\mathrm{End}_{\mathcal{V}}(M)\begin{pmatrix} x_1, \ldots, x_n \\ x \end{pmatrix} \simeq \mathrm{MAP}_{\mathcal{V}}(M(x_1) \otimes \cdots \otimes M(x_n), M(x)),$$

*with $\mathrm{MAP}_{\mathcal{V}}$ denoting the internal Hom in $\mathcal{V}$.*

**Proof**   Since $X$ is an $\infty$-groupoid, the twisted arrow $\infty$-category $\mathrm{Tw}(X)$ is equivalent to $X$, and so [7, Proposition 5.1] yields a natural equivalence

$$\mathrm{Map}_{\mathrm{Fun}(X, \mathcal{V})}(M \odot S, M) \simeq \lim_{x \in X} \mathrm{Map}_{\mathcal{V}}((M \odot S)(x), M(x)).$$

Now the description of $M \odot S$ from Proposition 3.8 shows that this is naturally equivalent to

$$\lim_{x \in X} \mathrm{Map}_{\mathcal{V}}\Big( \coprod_n \operatorname*{colim}_{(x_1, \ldots, x_n) \in X_{h\Sigma_n}^n} M(x_1) \otimes \cdots \otimes M(x_n) \otimes S\begin{pmatrix} x_1, \ldots, x_n \\ x \end{pmatrix}, M(x)\Big).$$

Taking the limit out and applying the universal property of MAP, this becomes

$$\lim_{x \in X}\Big( \prod_n \lim_{(x_1, \ldots, x_n) \in X_{h\Sigma_n}^n} \mathrm{Map}_{\mathcal{V}}\Big( S\begin{pmatrix} x_1, \ldots, x_n \\ x \end{pmatrix}, \mathrm{End}(M)\begin{pmatrix} x_1, \ldots, x_n \\ x \end{pmatrix}\Big)\Big).$$

We can now combine the limits to get a limit over $\coprod_n X \times X_{h\Sigma_n}^n \simeq \mathbb{F}_X^{\widetilde{\simeq}}$, ie

$$\lim_{\xi \in \mathbb{F}_X^{\widetilde{\simeq}}} \mathrm{Map}_{\mathcal{V}}(S(\xi), \mathrm{End}_{\mathcal{V}}(M)(\xi)).$$

Applying [7, Proposition 5.1] once more now identifies this limit (since $\mathbb{F}_X^{\widetilde{\simeq}}$ is again an $\infty$-groupoid) with $\mathrm{Map}_{\mathrm{Fun}(\mathbb{F}_X^{\widetilde{\simeq}}, \mathcal{V})}(S, \mathrm{End}_{\mathcal{V}}(M))$, as required. □

**Corollary 5.3**   *For any $M \colon X \to \mathcal{V}$, the $\infty$-category $\mathrm{Coll}_X(\mathcal{V})[M]$ has a terminal object.*

**Proof**   By Proposition 5.2, the functor $\mathrm{Coll}_X(\mathcal{V})^{\mathrm{op}} \to \mathcal{S}$ corresponding to the right fibration

$$\mathrm{Coll}_X(\mathcal{V})[M] \to \mathrm{Coll}_X(\mathcal{V})$$

is represented by the object $\mathrm{End}_{\mathcal{V}}(M)$. This implies that we have an equivalence

$$\mathrm{Coll}_X(\mathcal{V})[M] \simeq \mathrm{Coll}_X(\mathcal{V})_{/\mathrm{End}_{\mathcal{V}}(M)}.$$

Since the right-hand side clearly has a terminal object, this completes the proof. □

Applying Proposition 5.1, we get:

**Corollary 5.4** *For any $M \in \mathrm{Fun}(X, \mathcal{V})$ there exists an endomorphism $\infty$-operad $\mathrm{End}_\mathcal{V}(M)$ in $\mathrm{Opd}_X(\mathcal{V}) \simeq \mathrm{Alg}_{\mathbf{\Delta}^{\mathrm{op}}}(\mathrm{Coll}_X(\mathcal{V}))$ whose underlying $X$-collection is*

$$\mathrm{End}_\mathcal{V}(M)\binom{x_1, \ldots, x_n}{y} \simeq \mathrm{MAP}(M(x_1) \otimes \cdots \otimes M(x_n), M(y)).$$

*This has the universal property that, for any $\mathcal{O} \in \mathrm{Opd}_X(\mathcal{V})$, morphisms $\mathcal{O} \to \mathrm{End}_\mathcal{V}(M)$ in $\mathrm{Opd}_X(\mathcal{V})$ correspond to $\mathcal{O}$-algebra structures on $M$, ie there is natural equivalence*

$$\mathrm{Map}_{\mathrm{Opd}_X(\mathcal{V})}(\mathcal{O}, \mathrm{End}_\mathcal{V}(M)) \simeq \mathrm{Alg}_\mathcal{O}(\mathcal{V})_{\widetilde{M}}^{\simeq},$$

*where the right-hand side denotes the underlying $\infty$-groupoid of the fiber of $\mathrm{Alg}_\mathcal{O}(\mathcal{V}) \to \mathrm{Fun}(X, \mathcal{V})$ at $M$.*

**Remark 5.5** For $X \simeq *$, so that the functor $* \to \mathcal{V}$ picks out an object $v$ of $\mathcal{V}$, we get an $\infty$-categorical analogue of the classical endomorphism operad: $\mathrm{End}_\mathcal{V}(v)$ is a one-object $\mathcal{V}$-$\infty$-operad with underlying symmetric sequence

$$\mathrm{End}_\mathcal{V}(v)(n) \simeq \mathrm{MAP}_\mathcal{V}(v^{\otimes n}, v).$$

If $\mathcal{O}$ is a one-object $\mathcal{V}$-$\infty$-operad, the universal property says that an $\mathcal{O}$-algebra structure on $v$ is equivalent to a morphism of one-object $\infty$-operads $\mathcal{O} \to \mathrm{End}_\mathcal{V}(v)$.

**Example 5.6** By Corollary 3.13, if $\mathcal{O}$ is any $\mathcal{V}$-$\infty$-operad with space of objects $X$, then the functor $Z(\mathcal{O}) \colon X \to \mathcal{V}$ picking out the nullary operations is canonically a right $\mathcal{O}$-module. This corresponds to a canonical morphism of $\mathcal{V}$-$\infty$-operads $\mathcal{O} \to \mathrm{End}(Z(\mathcal{O}))$, given by maps

$$\mathcal{O}\binom{x_1, \ldots, x_n}{y} \to \mathrm{MAP}_\mathcal{V}(Z(\mathcal{O})(x_1) \otimes \cdots \otimes Z(\mathcal{O})(x_n), Z(\mathcal{O})(y)),$$

adjoint to the composition maps

$$\mathcal{O}\binom{}{x_1} \otimes \cdots \otimes \mathcal{O}\binom{}{x_n} \otimes \mathcal{O}\binom{x_1, \ldots, x_n}{y} \to \mathcal{O}\binom{}{y}$$

for $\mathcal{O}$.

We now observe that the endomorphism algebras are compatible with the lax monoidal functors $f^* \colon \mathrm{Coll}_Y(\mathcal{V}) \to \mathrm{Coll}_X(\mathcal{V})$ induced by morphisms of $\infty$-groupoids $f \colon X \to Y$:

**Proposition 5.7** *For $f \colon X \to Y$ a morphism in $\mathcal{S}$ and $M \colon Y \to \mathcal{V}$, there is a natural equivalence of RM-algebras*

$$f^*(M, \mathrm{End}_\mathcal{V}(M)) \xrightarrow{\sim} (f^*M, \mathrm{End}_\mathcal{V}(f^*M)).$$

**Proof** If $\mathcal{O}$ is a $\mathcal{V}$-$\infty$-operad with $Y$ as space of objects, the lax monoidal functor $f^*$ induces a natural functor

$$\mathrm{Alg}_\mathcal{O}(\mathcal{V}) \to \mathrm{Alg}_{f^*\mathcal{O}}(\mathcal{V}),$$

given on the underlying functors to $\mathcal{V}$ by composition with $f$. Applying this to the $\mathcal{V}$-$\infty$-operad $\mathrm{End}_{\mathcal{V}}(M)$ and the canonical $\mathrm{End}_{\mathcal{V}}(M)$-algebra structure on $M$, we obtain an $f^*\mathrm{End}_{\mathcal{V}}(M)$-algebra structure on $f^*M = M \circ f$. By the universal property of endomorphism $\infty$-operads this corresponds to a morphism of $\infty$-operads $f^*\mathrm{End}_{\mathcal{V}}(M) \to \mathrm{End}_{\mathcal{V}}(f^*M)$. Using the explicit description of the underlying collection of $\mathrm{End}_{\mathcal{V}}(M)$ in terms of internal Homs we see that this is an equivalence. $\qquad\square$

There exists a universal functor from an $\infty$-groupoid to $\mathcal{V}$, namely the inclusion $\mathcal{V}^{\simeq} \to \mathcal{V}$ of the underlying $\infty$-groupoid of $\mathcal{V}$. Our construction does not apply directly to this, since the $\infty$-groupoid $\mathcal{V}^{\simeq}$ is not small. However, by passing to a larger universe we can define a universal endomorphism $\infty$-operad for $\mathcal{V}$:

**Definition 5.8**  Let $\mathcal{V}$ be a large closed symmetric monoidal $\infty$-category compatible with colimits indexed by small $\infty$-groupoids. By [12, Proposition 4.8.1.10] there is a very large presentable $\infty$-category $\widehat{\mathcal{V}}$ compatible with large colimits, with a fully faithful symmetric monoidal functor $\mathcal{V} \hookrightarrow \widehat{\mathcal{V}}$ that preserves colimits over small $\infty$-groupoids. Let $\mathrm{Opd}(\widehat{\mathcal{V}})$ be the $\infty$-category of $\widehat{\mathcal{V}}$-enriched $\infty$-operads with potentially large spaces of objects.

**Remark 5.9**  Since $\mathcal{V}$ is a symmetric monoidal full subcategory of $\widehat{\mathcal{V}}$, we can regard $\mathrm{Opd}(\mathcal{V})$ as a full subcategory of $\mathrm{Opd}(\widehat{\mathcal{V}})$, containing precisely those $\widehat{\mathcal{V}}$-$\infty$-operads whose spaces of objects are small and whose objects of multimorphisms all lie in the full subcategory $\mathcal{V}$. Similarly, for any $\mathcal{O} \in \mathrm{Opd}_X(\mathcal{V})$ we have a pullback square

$$
\begin{array}{ccc}
\mathrm{Alg}_{\mathcal{O}}(\mathcal{V}) & \hookrightarrow & \mathrm{Alg}_{\mathcal{O}}(\widehat{\mathcal{V}}) \\
\downarrow & & \downarrow \\
\mathrm{Fun}(X, \mathcal{V}) & \hookrightarrow & \mathrm{Fun}(X, \widehat{\mathcal{V}})
\end{array}
$$

where the horizontal maps are fully faithful.

**Definition 5.10**  Let $i : \mathcal{V}^{\simeq} \to \widehat{\mathcal{V}}$ denote the inclusion of the space of objects in the full subcategory $\mathcal{V}$. Applying Corollary 5.4 in the enlarged universe to $i$, we get an endomorphism $\widehat{\mathcal{V}}$-$\infty$-operad

$$\overline{\mathcal{V}} := \mathrm{End}_{\widehat{\mathcal{V}}}(i).$$

The formula for its multimorphism objects implies that we can regard $\overline{\mathcal{V}}$ as a (large) $\mathcal{V}$-$\infty$-operad (ie an object in the $\infty$-category $\mathrm{Opd}_{\mathcal{V}^{\simeq}}(\mathcal{V}) \subseteq \mathrm{Opd}_{\mathcal{V}^{\simeq}}(\widehat{\mathcal{V}})$, which makes sense also when the $\infty$-groupoid of objects is large). Moreover, for any map of $\infty$-groupoids $M : X \to \mathcal{V}^{\simeq}$ where $X$ is small, we can regard

$$\mathrm{End}_{\widehat{\mathcal{V}}}(i \circ M) \simeq M^*\overline{\mathcal{V}}$$

as an object of $\mathrm{Opd}_X(\mathcal{V})$.

**Lemma 5.11**  *For any map $M : X \to \mathcal{V}^{\simeq}$ where $X$ is a small $\infty$-groupoid, we have a canonical equivalence*

$$M^*\overline{\mathcal{V}} \simeq \mathrm{End}_{\mathcal{V}}(M).$$

**Proof** For $\mathcal{O} \in \mathrm{Opd}_X(\mathcal{V})$ we have a natural equivalence

$$\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})_M \xrightarrow{\sim} \mathrm{Alg}_{\mathcal{O}}(\widehat{\mathcal{V}})_{iM}.$$

Using the universal property of the endomorphism objects this corresponds to a natural equivalence

$$\mathrm{Map}(\mathcal{O}, M^*\overline{\mathcal{V}}) \xrightarrow{\sim} \mathrm{Map}(\mathcal{O}, \mathrm{End}_{\mathcal{V}}(M)),$$

and so an equivalence $M^*\overline{\mathcal{V}} \xrightarrow{\sim} \mathrm{End}_{\mathcal{V}}(M)$, as required. $\qquad\square$

Let $U$ denote the canonical $\overline{\mathcal{V}}$-algebra structure on $i$, which we can regard as an object of $\mathrm{Alg}_{\overline{\mathcal{V}}}(\mathcal{V})$. For every map $M \colon X \to \mathcal{V}^{\simeq}$ with $X$ small, the pullback $M^*U$ is then the canonical $\mathrm{End}_{\mathcal{V}}(M)$-algebra structure on $M$, which leads to the following:

**Theorem 5.12** *For any small $\mathcal{V}$-$\infty$-operad $\mathcal{O}$, the morphism of $\infty$-groupoids*

$$\mathrm{Map}_{\mathrm{Opd}(\widehat{\mathcal{V}})}(\mathcal{O}, \overline{\mathcal{V}}) \to \mathrm{Alg}_{\mathcal{O}}(\mathcal{V})^{\simeq},$$

*which takes $\phi \colon \mathcal{O} \to \overline{\mathcal{V}}$ to $\phi^*U \in \mathrm{Alg}_{\mathcal{O}}(\mathcal{V})^{\simeq}$, is an equivalence.*

**Proof** Let $X$ be the space of objects of $\mathcal{O}$. Then we have a commutative triangle of $\infty$-groupoids



It suffices to show that we have an equivalence on the fibers over each map $M \colon X \to \mathcal{V}$. But we have an equivalence between $\mathrm{Map}_{\mathrm{Opd}(\widehat{\mathcal{V}})}(\mathcal{O}, \overline{\mathcal{V}})_M$ and

$$\mathrm{Map}_{\mathrm{Opd}_X(\mathcal{V})}(\mathcal{O}, M^*\overline{\mathcal{V}}) \simeq \mathrm{Map}_{\mathrm{Opd}_X(\mathcal{V})}(\mathcal{O}, \mathrm{End}_{\mathcal{V}}(M)),$$

under which the map to $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})^{\simeq}_M$ is equivalent to that taking $\phi \colon \mathcal{O} \to \mathrm{End}_{\mathcal{V}}(M)$ to $\phi^*$ applied to the canonical $\mathrm{End}_{\mathcal{V}}(M)$-algebra structure on $M$. This is an equivalence by the universal property of the endomorphism algebra. $\qquad\square$

**Remark 5.13** In [4] we constructed a natural tensoring of $\mathcal{V}$-$\infty$-operads over $\infty$-categories. This induces an enrichment in $\infty$-categories, given by

$$\mathrm{Map}_{\mathrm{Cat}_{\infty}}(\mathcal{C}, \mathrm{Alg}_{\mathcal{O}}(\mathcal{P})) \simeq \mathrm{Map}_{\mathrm{Opd}_{\infty}^{\mathcal{V}}}(\mathcal{C} \otimes \mathcal{O}, \mathcal{P}).$$

For $\mathcal{P} = \overline{\mathcal{V}}$, we can use Theorem 5.12 to identify the $\infty$-category $\mathrm{Alg}_{\mathcal{O}}(\overline{\mathcal{V}})$ with the Segal space $\mathrm{Alg}_{\Delta^{\bullet} \otimes \mathcal{O}}(\mathcal{V})^{\simeq}$. We expect that this should in fact be equivalent to the $\infty$-category $\mathrm{Alg}_{\mathcal{O}}(\mathcal{V})$, but to prove it we need a better understanding of the tensoring of $\mathcal{V}$-$\infty$-operads and $\infty$-categories. As this was defined rather inexplicitly in [4], we suspect that this requires setting up a new definition of enriched $\infty$-operads where the tensoring can be described more concretely.

In the case where $\mathcal{V}$ is the $\infty$-category $\mathcal{S}$ of spaces, we can identify $\bar{\mathcal{S}}$ explicitly:

**Proposition 5.14** *Let $\mathcal{S}^{\times}$ denote the symmetric monoidal $\infty$-category given by the cartesian product in $\mathcal{S}$, viewed as an $\mathcal{S}$-enriched $\infty$-operad. There is an equivalence $\mathcal{S}^{\times} \xrightarrow{\sim} \bar{\mathcal{S}}$.*

**Proof** For $X \in \mathcal{S}$, we have $Z(\mathcal{S}^{\times})(X) \simeq \mathrm{Map}_{\mathcal{S}}(*, X) \simeq X$, and the functor $Z(\mathcal{S}^{\times}) \colon \mathcal{S}^{\simeq} \to \mathcal{S}$ is the inclusion of the underlying $\infty$-groupoid. Hence, thinking of $\bar{\mathcal{S}}$ as an endomorphism object for $\infty$-operads enriched in large spaces, by Example 5.6 there is a canonical morphism $\mathcal{S}^{\times} \to \bar{\mathcal{S}}$. This is given by equivalences

$$\mathcal{S}^{\times}\binom{X_1, \ldots, X_n}{Y} \xrightarrow{\sim} \mathrm{Map}_{\mathcal{S}}(X_1 \times \cdots \times X_n, Y),$$

and so it is an equivalence of $\mathcal{S}$-$\infty$-operads. $\square$

**Remark 5.15** It follows that for $\mathcal{O}$ an $\mathcal{S}$-$\infty$-operad, the $\infty$-groupoid $\mathrm{Alg}_{\mathcal{O}}(\mathcal{S})^{\simeq}$ in our sense is equivalent to $\mathrm{Map}_{\mathrm{Opd}(\widehat{\mathcal{S}})}(\mathcal{O}, \mathcal{S}^{\times})$. This is the underlying $\infty$-groupoid of the $\infty$-category of $\mathcal{O}$-algebras in $\mathcal{S}$ defined in [12], so for $\mathcal{S}$-enriched $\infty$-operads our notion of $\mathcal{O}$-algebras agrees with that of [12], at least on the level of $\infty$-groupoids.

# References

[1] **D Ayala**, **J Francis**, *Flagged higher categories*, from "Topology and quantum theory in interaction", Contemp. Math. 718, Amer. Math. Soc., Providence, RI (2018) 137–173 MR

[2] **C Barwick**, *From operator categories to higher operads*, Geom. Topol. 22 (2018) 1893–1959 MR

[3] **D L B Brantner**, *The Lubin–Tate theory of spectral Lie algebras*, PhD thesis, Harvard University (2017) Available at `https://people.maths.ox.ac.uk/brantner/brantnerthesis.pdf`

[4] **H Chu**, **R Haugseng**, *Enriched $\infty$-operads*, Adv. Math. 361 (2020) art. id. 106913 MR

[5] **D-C Cisinski**, *Higher categories and homotopical algebra*, Cambridge Stud. Adv. Math. 180, Cambridge Univ. Press (2019) MR

[6] **D Gepner**, **R Haugseng**, *Enriched $\infty$-categories via non-symmetric $\infty$-operads*, Adv. Math. 279 (2015) 575–716 MR

[7] **D Gepner**, **R Haugseng**, **T Nikolaus**, *Lax colimits and free fibrations in $\infty$-categories*, Doc. Math. 22 (2017) 1225–1266 MR

[8] **R Haugseng**, *$\infty$-operads via symmetric sequences*, Math. Z. 301 (2022) 115–171 MR

[9] **G Heuts**, *Algebras over $\infty$-operads*, preprint (2011) arXiv 1110.1776

[10] **V Hinich**, *Yoneda lemma for enriched $\infty$-categories*, Adv. Math. 367 (2020) art. id. 107129 MR

[11] **J Lurie**, *Higher topos theory*, Ann. of Math. Stud. 170, Princeton Univ. Press (2009) MR

[12] **J Lurie**, *Higher algebra*, preprint (2017) Available at `https://url.msp.org/Lurie-HA`

[13] **A Mazel-Gee**, *Quillen adjunctions induce adjunctions of quasicategories*, New York J. Math. 22 (2016) 57–93 MR

[14] **D Pavlov**, **J Scholbach**, *Admissibility and rectification of colored symmetric operads*, J. Topol. 11 (2018) 559–601 MR

[15] **D Pavlov**, **J Scholbach**, *Homotopy theory of symmetric powers*, Homology Homotopy Appl. 20 (2018) 359–397 MR

[16] **S Schwede**, An untitled book project about symmetric spectra (2007) Available at `https://www.math.uni-bonn.de/~schwede/SymSpec.pdf`

[17] **B Shipley**, *A convenient model category for commutative ring spectra*, from "Homotopy theory: relations with algebraic geometry, group cohomology, and algebraic $K$-theory", Contemp. Math. 346, Amer. Math. Soc., Providence, RI (2004) 473–483 MR

[18] **M Spitzweck**, *Operads, algebras and modules in general model categories*, preprint (2001) arXiv math/0101102

[19] **T H Trimble**, *Notes on the Lie operad*, preprint (2019) Available at `https://math.ucr.edu/home/baez/trimble/trimble_lie_operad.pdf`

[20] **D White**, **D Yau**, *Bousfield localization and algebras over colored operads*, Appl. Categ. Structures 26 (2018) 153–203 MR

[21] **D White**, **D Yau**, *Smith ideals of operadic algebras in monoidal model categories*, Algebr. Geom. Topol. 24 (2024) 341–392 MR

Department of Mathematical Sciences, Norwegian University of Science and Technology
Trondheim, Norway

`rune.haugseng@ntnu.no`

# An overtwisted convex hypersurface in higher dimensions

River Chiang

Klaus Niederkrüger-Eid

We show that the germ of the contact structure surrounding a certain kind of convex hypersurface is overtwisted. We then find such hypersurfaces close to any plastikstufe with toric core, thereby proving that the existence of a plastikstufe with toric core implies overtwistedness. All proofs in this article are explicit, and we hope that the methods used here might hint at a deeper understanding of the size of neighborhoods in contact manifolds.

In the appendix we reprove in a concise way that the Legendrian unknot is loose if the ambient manifold contains a large enough neighborhood of a 2-dimensional overtwisted disk. Additionally we prove the folklore result that the singular distribution induced on a hypersurface $\Sigma$ of a contact manifold $(M, \xi)$ determines the germ of the contact structure around $\Sigma$.

53D10

## 1 Introduction

The fundamental distinction between tight and overtwisted contact structures was discovered first by Eliashberg in dimension 3 [5], and then generalized to arbitrary dimensions by Borman, Eliashberg and Murphy [2]. The main feature is that an overtwisted contact structure is flexible. However, the original high-dimensional definition makes it practically unverifiable if a given contact structure is overtwisted. Thanks to Casals, Murphy and Presas [3] we know that most previously existing conjectural definitions for overtwistedness are actually equivalent to the one given in [2].

Still many open questions persist. One of them has been settled by Huang [8]: The second author of this article proposed a definition of overtwistedness called plastikstufe [13]. The results in [12] combined with those of [3] showed that certain very specific plastikstufes imply overtwistedness. Huang explained in [8] that this result extends to *any* plastikstufe. However, we are unable to verify several claims made in his proof.

In this article, we reprove Huang's result for the more restrictive case of plastikstufes with toric core (generalizing results by Adachi [1]).

Our strategy is based on the following observation about a certain type of convex hypersurface. Consider for any $C > 0$ the manifold

$$\Sigma_C = \mathbb{D}^2_{\leq \pi} \times (-C, C)^{2n}$$

carrying a singular distribution $\mathcal{D}_C$ given as the kernel of the 1-form $\beta = r \sin r \, d\vartheta - \sum_{j=1}^{n} t_j \, ds_j$, where $(r, \vartheta)$ are polar coordinates on the disk, and $(s_j, t_j)$ are the natural coordinates on the cube $(-C, C)^n \times (-C, C)^n$.

**Theorem A** *There exists a constant $C_{\text{OT}} > 0$ such that the following holds for every contact manifold $(M, \xi)$ of dimension at least 5. If $(M, \xi)$ admits an embedding of a hypersurface $(\Sigma_C, \mathcal{D}_C)$ with $C > C_{\text{OT}}$ such that $\xi$ induces the singular distribution $\mathcal{D}_C$ on $\Sigma_C$, then $(M, \xi)$ is overtwisted.*

Consequently, we call any embedded hypersurface $(\Sigma_C, \mathcal{D}_C)$ with $C > C_{\text{OT}}$ an *overtwisted convex disk*.

**Remark** (a) This definition is closely related to the characterization of overtwisted contact structures in terms of "large neighborhoods" given in [3; 15]. Our result states that instead of considering large embedded balls, it is already sufficient to find a large hypersurface.

(b) In the model used in Section 4, one sees directly that there is an obvious contact vector field $Z$ ($= \partial_z$ in fact) which is transverse to the overtwisted convex disk. It is tempting to try to characterize the constant $C_{\text{OT}}$ more explicitly by trying to recognize some sort of higher-dimensional Giroux criterion for convex hypersurfaces. Unfortunately, the dividing set for $Z$ is noncompact, and we have not succeeded in finding a more suitable contact vector field.

(c) Even though we are unable to give a specific value for the size parameter that appears in the definition of overtwistedness in [2] (and the equivalent formulation via large neighborhoods in [3]), it follows from our argument that the size parameter can be chosen uniformly for all dimensions.

**Proof of Theorem A** We show in Sections 3 and 4 that every neighborhood of $\mathbb{D}^2_{\leq \pi} \times (-C, C)^{2n}$ contains the embedding of a certain type of open subset $B(h) \times \left(-\frac{5}{6}C, \frac{5}{6}C\right)^{2n}$ of arbitrary height $h > 0$. See Corollary 4.3 for the details. As explained first in [12, page 1813], the Legendrian unknot is loose if $C$ is chosen sufficiently large. We give in Appendix A a streamlined proof of this statement.

It was proved in [3] that any contact manifold in which the unknot is loose is overtwisted. $\square$

**Question** *Is it possible to explicitly show that in any neighborhood of a hypersurface $\Sigma_C$ with $C > C_{\text{OT}}$ one can embed a hypersurface $\Sigma_{C'}$ with $C' > 2C$ as in the analogous claim for loose charts [11, Proposition 4.4]?*

Choose a $\rho > 0$ and define $V_\rho$ as in Appendix A. It is likely that the contact germs around the overtwisted convex disk $(\Sigma_C, \mathcal{D}_C)$ correspond to thick neighborhoods of the form $\left(\mathbb{R}^3 \times V_\rho, \ker(\alpha_{\text{OT}} + \lambda_{\text{can}})\right)$. It was

shown in [3] that these neighborhoods are for sufficiently large $\rho > 0$ overtwisted. However while it is obvious that the contact germs considered here embed into such thick neighborhoods, it is not obvious how to directly prove the converse. By taking instead the detour over the loose unknots, we are able to split our argument cleanly into precise steps.

Even if it might be unclear at the moment if Theorem A is more than a curious observation, it allows us to prove in an extremely elementary way that every contact manifold containing a plastikstufe with *toric* core is overtwisted, as claimed in [8].

In Appendix A, we show that the Legendrian unknot is loose in a large neighborhood of an overtwisted 2-disk; in Appendix B we prove the folklore result that a hypersurface in a contact manifold determines together with the induced singular distribution the germ of the contact structure.

## Acknowledgments

## 2 A plastikstufe with toric core implies overtwistedness

Denote the cylindrical coordinates on $\mathbb{R}^3$ by $(r, \vartheta, z)$. The 1-form

$$\alpha_{\mathrm{OT}} = \cos(r)\, dz + r\sin(r)\, d\vartheta$$

is then a well-defined contact form. The disk $\mathbb{D}_{\mathrm{OT}} := \big\{ (r, \vartheta, z) \mid r \leq \pi,\ z = 0 \big\}$ is overtwisted, and we will call it the *standard* overtwisted disk.

We choose a small cylindrical box of height $h$ around $\mathbb{D}_{\mathrm{OT}}$ of the form

(2-1) $$B(h) := \mathbb{D}^2_{<\pi+\delta} \times (-h, h).$$

Let $(M, \xi)$ be a $(2n+3)$-dimensional contact manifold that contains a plastikstufe of the form $\mathbb{D}_{\mathrm{OT}} \times \mathbb{T}^n$. We will show that we can find an arbitrarily "large" hypersurface $\mathbb{D}^2_{\leq \pi} \times [-C, C]^{2n}$ in any neighborhood of the plastikstufe by successively unwinding each of the $\mathbb{S}^1$-factors of the torus. This way we obtain the following corollary.

**Corollary 2.1** *Every contact manifold that contains a plastikstufe $\mathbb{D}_{\mathrm{OT}} \times \mathbb{T}^n$ with toric core also admits an embedding of an overtwisted convex disk.*

**Proof** There is a neighborhood of the plastikstufe that is contactomorphic to an open neighborhood of $\mathbb{D}_{OT} \times \mathbb{T}^n$ in

$$\left(\mathbb{R}^3 \times T^*\mathbb{T}^n, \alpha_{OT} + \lambda_{can}\right);$$

compare [14, Theorem I.1.3.]. Choosing $\varepsilon > 0$ and $\delta > 0$ small enough, we can assume that this neighborhood is contactomorphic to a product of the form $B(\varepsilon) \times \mathbb{D}_{<\delta}\left(T^*\mathbb{T}^n\right)$, where $B(\varepsilon) \subset \mathbb{R}^3$ is a cylindrical box as defined in (2-1) and $\mathbb{D}_{<\delta}\left(T^*\mathbb{T}^n\right)$ is the disk bundle of radius $\delta$ in $T^*\mathbb{T}^n$.

We can now apply to this neighborhood Lemma 2.3 in dimension 5, or Lemma 2.4 in the general case to find a hypersurface of the form $\mathbb{D}^2_{\leq \pi} \times (-C, C)^n \times (-a, a)^n$ for $a = \delta/\left(2\sqrt{n}\right)$, and where $C > 0$ can be chosen to be arbitrarily large. The singular distribution induced by the contact structure agrees with the kernel of $r \sin r \, d\vartheta - \sum_{j=1}^n t_j \, ds_j$, where $(r, \vartheta)$ are polar coordinates on the disk, and $(s_j, t_j)$ are the natural coordinates on the rectangle $(-C, C) \times (-a, a)$.

If $C > 0$ is chosen larger than $2C_{OT}^2/a$, then it suffices to apply to each coordinate pair $(s_j, t_j)$ the diffeomorphism $(s_j, t_j) \mapsto \left(\mu^{-1}s_j, \mu t_j\right)$ with $\mu = 2C_{OT}/a$ to obtain the desired overtwisted convex disk $(\Sigma_{\widetilde{C}}, \mathcal{D}_{\widetilde{C}})$ for some appropriate $\widetilde{C} > C_{OT}$. $\qquad\square$

**Remark 2.2** The initial version of a plastikstufe $\mathbb{D}^2 \times S$ introduced in [13] differed slightly from the form $\mathbb{D}_{OT} \times S$ used here, because the boundary of $\mathbb{D}_{OT} \times S$ is composed of singular leaves of the foliation. The presence of either type of a plastikstufe in a contact manifold implies by [14, Theorem I.1.3.] the other one: The plastikstufe determines the germ of the contact structure which will be contactomorphic to an open subset of $\left(\mathbb{R}^3 \times T^*S, \alpha_{OT} + \lambda_{can}\right)$. This allows us to modify the plastikstufe along its boundary by a $C^0$-small deformation to move from one model to the other one.

We will now show how to "unwrap" the toric plastikstufe. Consider for simplicity first a contact manifold of dimension 5 so that $\mathbb{T}^n = \mathbb{S}^1$.

**Lemma 2.3** *Choose any $\varepsilon > 0$ and $\delta > 0$, and let $B(\varepsilon) \times \mathbb{D}_{<\delta}\left(T^*\mathbb{S}^1\right)$ be a neighborhood of a plastikstufe $\mathbb{D}_{OT} \times \mathbb{S}^1$ in $\left(\mathbb{R}^3 \times T^*\mathbb{S}^1, \alpha_{OT} + \lambda_{can}\right)$.*

*For any arbitrarily large $C > 0$ it is possible to embed the hypersurface*

$$S_C := \mathbb{D}^2_{\leq \pi} \times (-C, C) \times \left(-\tfrac{\delta}{2}, \tfrac{\delta}{2}\right)$$

*into $B(\varepsilon) \times \mathbb{D}_{<\delta}\left(T^*\mathbb{S}^1\right)$ such that the contact structure induces the singular distribution*

$$\mathcal{D} := \ker\left(r \sin r \, d\vartheta - t \, ds\right)$$

*on $S_C$. Here $(r, \vartheta)$ are polar coordinates on the disk, and $(s, t)$ are the natural coordinates on the rectangle $\left(-C, C\right) \times \left(-\tfrac{\delta}{2}, \tfrac{\delta}{2}\right)$.*

**Proof** Define for any $\hbar > 0$ an embedding

$$\Psi_\hbar \colon \mathbb{D}^2_{\leq \pi} \times \mathbb{R}^2 \to \mathbb{R}^3 \times T^* \mathbb{S}^1, \quad (r, \vartheta; s, t) \mapsto (r, \vartheta, \hbar s; q = e^{is}, p = t + \hbar \cos r),$$

"reeling up" the hypersurface along the $\mathbb{S}^1$-core of the plastikstufe. One easily verifies that

$$\Psi_\hbar^*(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}}) = r \sin r \, d\vartheta - t \, ds$$

so that the singular distribution induced by $\ker(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}})$ is indeed equal to $\mathcal{D}$.

Choose $\hbar = \varepsilon / C$ and suppose that $\hbar < \frac{\delta}{2}$. We see that $\Psi_\hbar(S_C)$ lies in the given neighborhood $B(\varepsilon) \times \mathbb{D}_{<\delta}(T^* \mathbb{S}^1)$. $\qquad\square$

For general dimensions, the embedding is only slightly more complicated. Denote the standard coordinates on $\mathbb{T}^n$ by $\boldsymbol{q} = (q_1, \ldots, q_n)$ and those on $T^* \mathbb{T}^n$ by $(\boldsymbol{q}, \boldsymbol{p}) = (q_1, \ldots, q_n; p_1, \ldots, p_n)$.

**Lemma 2.4** *Any neighborhood* $B(\varepsilon) \times \mathbb{D}_{<\delta}(T^* \mathbb{T}^n)$ *of a plastikstufe* $\mathbb{D}_{\mathrm{OT}} \times \mathbb{T}^n$ *with* $\varepsilon > 0$ *and* $\delta > 0$ *arbitrarily small contains an embedded hypersurface of the form*

$$S_C := \mathbb{D}^2_{\leq \pi} \times \left\{ (s_1, \ldots, s_n; t_1, \ldots, t_n) \in \mathbb{R}^{2n} \;\middle|\; |s_j| < C, \, |t_j| < \frac{\delta}{2\sqrt{n}} \right\}$$

*with* $C > 0$ *arbitrarily large such that the contact structure induces the singular distribution*

$$\mathcal{D} = \ker\left( r \sin r \, d\vartheta - \sum_{j=1}^n t_j \, ds_j \right)$$

*on* $S_C$.

**Proof** Choose constants $\hbar_1, \ldots, \hbar_n > 0$ that are linearly independent over $\mathbb{Q}$, and define a map $\Psi \colon \mathbb{D}^2_{\leq \pi} \times \mathbb{R}^{2n} \to \mathbb{R}^3 \times T^* \mathbb{T}^n$ by

$$(r, \vartheta; s_1, \ldots, s_n; t_1, \ldots, t_n)$$

$$\mapsto \left( r, \vartheta, z = \sum_{j=1}^n \hbar_j s_j; q_1 = e^{is_1}, \ldots, q_n = e^{is_n}; p_1 = t_1 + \hbar_1 \cos r, \ldots, p_n = t_n + \hbar_n \cos r \right).$$

It is easy to verify that $\Psi^*(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}}) = r \sin r \, d\vartheta - (t_1 \, ds_1 + \cdots + t_n \, ds_n)$ induces the distribution $\mathcal{D}$ on $S_C$. It is also immediately clear that $\Psi$ is an immersion.

To see that $\Psi$ is injective, use first that the images of two points $(r, \vartheta; s_1, \ldots, s_n; t_1, \ldots, t_n)$ and $(r', \vartheta'; s'_1, \ldots, s'_n; t'_1, \ldots, t'_n)$ by $\Psi$ can only agree if $r = r'$, $\vartheta = \vartheta'$, and $t_j = t'_j$ for all $j = 1, \ldots, n$, and if $s_j - s'_j$ is for every $j = 1, \ldots, n$ an integer multiple of $2\pi$. The equation $\hbar_1 s_1 + \cdots + \hbar_n s_n = \hbar_1 s'_1 + \cdots + \hbar_n s'_n$ implies that $\hbar_1(s_1 - s'_1) + \cdots + \hbar_n(s_n - s'_n) = 0$, but by our assumption that the $\hbar_j$ are linearly independent over $\mathbb{Q}$ it follows that all coefficients $s_1 - s'_1$ need to vanish so that $\Psi$ is injective.

We still need to verify that the image of $\Psi$ lies in the neighborhood $B(\varepsilon) \times \mathbb{D}_{<\delta}(T^* \mathbb{T}^n)$. To respect the $z$-height, it suffices to choose $\hbar_1 + \cdots + \hbar_n < \varepsilon / C$, so that the $z$-coordinate of $\Psi$ is bounded by $\varepsilon$. For the radius of the fibers in $\mathbb{D}_{<\delta}(T^* \mathbb{T}^n)$ choose $\hbar_j < \delta / (2\sqrt{n})$ so that $\Psi$ also stays inside the $\delta$-disk bundle of $T^* \mathbb{T}^n$. $\qquad\square$

# 3 The standard overtwisted contact structure on $\mathbb{R}^3$

For a cylindrical box of height $h$ around the standard overtwisted disk $\mathbb{D}_{\mathrm{OT}}$ in $(\mathbb{R}^3, \alpha_{\mathrm{OT}})$ of the form

$$B(h) := \mathbb{D}^2_{<\pi+\delta} \times (-h, h),$$

it is well known that the choice of $h$ is not relevant for the contactomorphism type. Below we will give a contact vector field that can be used to prove this fact by hand, but instead one can also easily convince oneself that all $B(h)$ are *overtwisted at infinity* which uniquely characterizes by [6] a contact structure on $\mathbb{R}^3$.

The main technical problem that we will deal with in this article is to show that the choice of the $h$-parameter also remains largely irrelevant for the contactomorphism type when we take the product with a Liouville domain.

We will now discuss a contact vector field $X$ whose flow compresses any large box $B(h)$ into an arbitrarily small neighborhood of the standard overtwisted disk. Ideally we would like $X$ to be a strict contact vector field or at least to have a constant scaling factor $c$ such that $\mathcal{L}_X \alpha_{\mathrm{OT}} = c \cdot \alpha_{\mathrm{OT}}$. Unfortunately such a vector field cannot exist: firstly, $X$ should be contracting and thus it needs to reduce the total volume. This implies that $c$ would have to be strictly negative on a predominant part of its domain. On the other hand, $c$ cannot be *everywhere* strictly negative as this would allow us to squeeze with the strategy of Section 4 a high-dimensional overtwisted chart into an arbitrarily "thin" set, thus contradicting the existence of tight contact manifolds.

The following vector field arose in discussions with Patrick Massot around 2010:

$$(3\text{-}1) \qquad X := -z \, \partial_z - \frac{r \cos(r) \sin(r)}{r + \cos(r) \sin(r)} \, \partial_r$$

is well defined and induces a contact flow on $(\mathbb{R}^3, \alpha_{\mathrm{OT}})$.

Even though this vector field might at first appear overly complicated, note that all coordinates in $X$ are uncoupled. This allows us to see that its time $T$ flow preserves the $\vartheta$-coordinate, and it contracts the $z$-coordinate by the factor $e^{-T}$. The radial coordinate is fixed on every cylinder of radius $r \in \frac{\pi}{2}\mathbb{N}$. The cylinders of radius $r \in \frac{\pi}{2} + \pi\mathbb{N}$ are repelling; the cylinders of radius $r \in \pi\mathbb{N}$ are attracting, in the sense that all points between these cylinders are pushed away from the repelling cylinder towards the attracting one; see Figure 1.

The height of the box $B(h)$ is squeezed by the flow of $X$ by an exponential factor, while the radial direction also shrinks, without becoming ever smaller than $\pi$ though. This way an arbitrarily tall box $B(h)$ can be squeezed by a contactomorphism into an arbitrarily small neighborhood of the standard overtwisted disk. In fact, one can convince oneself that the image of a box $B(h)$ for a certain choice of $\delta > 0$ will be of the form $\mathbb{D}^2_{<\pi+\delta'} \times (-h', h')$ for some smaller $h' > 0$ and $\delta' > 0$.
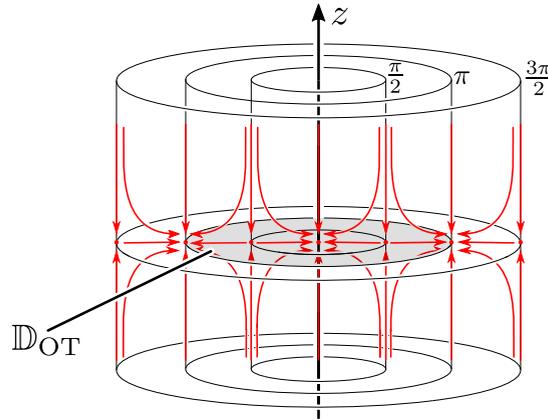
Figure 1: The flow of $X$ preserves the cylinders of radius $r \in \frac{\pi}{2} + \pi \mathbb{N}$. The boundary of the standard overtwisted disk sits on an attracting cylinder.

Finally note that $\mathcal{L}_X \alpha_{\mathrm{OT}} = g \cdot \alpha_{\mathrm{OT}}$, where

$$(3\text{-}2) \qquad\qquad g(r, \vartheta, z) = -\frac{\cos r \, (r \cos r + \sin r)}{r + \cos r \sin r}.$$

As we claimed above, $g$ takes both positive and negative values. More precisely:

**Lemma 3.1** *The function $g: B(h) \to \mathbb{R}$ is everywhere negative on $B(h)$ except for the domain lying between $r_m = \pi/2$ and $r_M \approx 2.03$ such that $r_M = -\tan r_M$. See also the graph in Figure 2.*

**Proof** The function $g$ only depends on the radial coordinate $r$. Its denominator is everywhere positive while the numerator changes once its sign at $r = \pi/2$, where $\cos r$ vanishes, and then again at $r \approx 2.03$, where $r \cos r + \sin r$ vanishes. □

We can also read off from the graph, Figure 2, that $g$ is everywhere smaller than 0.1; ie, even though $g$ becomes positive it only becomes very slightly so.
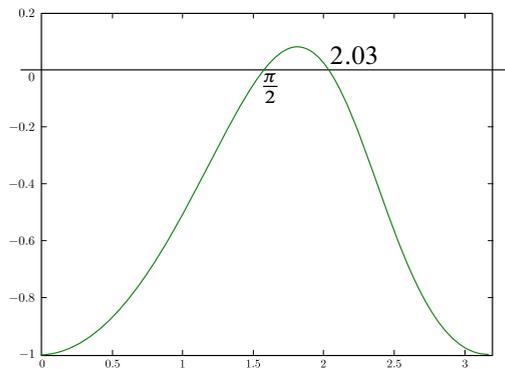


Figure 2: The graph of $g = -\dfrac{\cos r \, (r \cos r + \sin r)}{r + \cos r \sin r}$.

## 4 Contactomorphism on product structure

Let $(M, \xi)$ be a contact manifold with contact form $\alpha$. Assume that $X$ is a contact vector field such that $\mathcal{L}_X \alpha = g \cdot \alpha$ for some function $g\colon M \to \mathbb{R}$.

Choose an exact symplectic manifold $(W, d\lambda)$ that has a Liouville vector field $Y$, then we easily check that the contact form of $(M \times W, \alpha + \lambda)$ is preserved by the vector field

$$\widehat{X} = X + g \cdot Y.$$

Note that even if $Y$ points outwards and is expanding on $(W, d\lambda)$, the behavior of $\widehat{X}$ on the product manifold $M \times W$ is controlled in $W$-direction by the sign of the function $g$ that might take positive or negative values.

We consider now the main example we will be interested in: Let $B(h) \subset (\mathbb{R}^3, \alpha_{OT})$ be a box of height $h$ as defined in (2-1), and let $(L, \langle \cdot, \cdot \rangle)$ be a Riemannian manifold that does not need to be closed or geodesically complete. We denote the disk bundle of radius $c$ in $(T^*L, d\lambda_{can})$ by

$$\mathbb{D}_{<c}(T^*L) := \{v \in T^*L \mid \|v\| < c\}.$$

**Proposition 4.1** *There exists a positive constant $\mu_0 < \frac{7}{6}$ such that every contact domain*

$$\left( B(h) \times \mathbb{D}_{<c}(T^*L), \ker(\alpha_{OT} + \lambda_{can}) \right)$$

*with $c > 0$ can be embedded by a contactomorphism into*

$$\left( B(h') \times \mathbb{D}_{<\mu_0 \cdot c}(T^*L), \ker(\alpha_{OT} + \lambda_{can}) \right)$$

*independently of the choices of $h, h' > 0$.*

**Proof** For $h' \geq h$ the claim is obvious; for $h' < h$ the strategy is to use the flow of the vector field $\widehat{X} := X + g \cdot Y$ with $X$ and $g$ introduced in the previous section, and $Y$ the Liouville vector field on $T^*L$ that is defined by $\iota_Y d\lambda_{can} = \lambda_{can}$.

By writing $Y$ in a coordinate chart, one easily convinces oneself that the time $t$ flow of $Y$ simply consists of multiplying the fiber of $T^*L$ by $e^t$. In particular, even if $L$ is open or has boundary there is no danger that the flow of $\widehat{X}$ escapes transversely through the fibers of $T^*L$. Now let us study the behavior of the flow $\Phi_T^{\widehat{X}}$ in more details.

Recall that the coordinates are uncoupled by the flow of $X$ given in (3-1). We can thus write

$$\Phi_T^X(r, \vartheta, z) = (F(r, T), \vartheta, e^{-T} z),$$

where $F(r, T)$ is the solution of the ODE

$$y'(t) = -\frac{y(t) \cos y(t) \sin y(t)}{y(t) + \cos y(t) \sin y(t)} \quad \text{and} \quad y(0) = r.$$

The flow $\Phi_T^{\widehat{X}}$ is therefore of the form

$$\Phi_T^{\widehat{X}}(r, \vartheta, z; \boldsymbol{q}, \boldsymbol{p}) = \left( \Phi_T^X(r, \vartheta, z); \boldsymbol{q}, e^{G(r,T)} \cdot \boldsymbol{p} \right).$$

That is, the flow on the $\mathbb{R}^3$-factor simply reduces to the corresponding flow of $X$ and can be evaluated independently of the $T^*L$-part; the flow on the cotangent bundle factor is obtained by multiplying the fiber direction by a positive function $e^G$ that can be computed via

(4-1) $$G(r, t) = \int_0^t g\left( \Phi_s^X(r, \vartheta, z) \right) ds = \int_0^t g\left( F(r, s) \right) ds.$$

If $T$ is chosen to be $T = \ln \frac{h}{h'}$, it follows that $\Phi_T^{\widehat{X}}$ squeezes the first factor of $B(h) \times \mathbb{D}_{<c}(T^*L)$ into $B(h')$. By Lemma 4.2 below, $G(r, t) < \ln \frac{7}{6}$ for any point in $B(h)$ and any $t \geq 0$. This implies as desired that the initial domain is squeezed into $B(h') \times \mathbb{D}_{<(7/6)c}(T^*L)$. $\qquad\square$

**Lemma 4.2** *The function $G(r, t)$ given in (4-1) is bounded from above by $\ln \frac{7}{6}$ for all $r \in [0, \pi + \delta)$ and all $t \in [0, \infty)$.*

The sharp upper bound in the lemma is $\ln \frac{2r_M \sin r_M}{\pi}$ with $r_M$ specified in Lemma 3.1.

**Proof** Denote the $r$-coefficient of the vector field $X$ by

$$f(r) = -\frac{r \cos(r) \sin(r)}{r + \cos(r) \sin(r)}.$$

Then $F(r, t)$ is the flow of the field $X_r(r) := f(r) \, \partial_r$ on $[0, \pi + \delta)$; that is, $F$ is the solution of the ordinary differential equation $\partial_t F(r, t) = f\left( F(r, t) \right)$ with initial condition $F(r, 0) = r$.

The only critical points of $X_r$ are the points $r \in \frac{\pi}{2} \mathbb{N}$; see also Figure 1. Furthermore, recall that $r = \frac{\pi}{2}$ and $r = \frac{3\pi}{2}$ are repelling, and that $r = \pi$ is an attracting critical point.

According to Lemma 3.1, the function $g$ is everywhere on $[0, \pi + \delta)$ negative except for the interval $[r_m, r_M]$ with $r_m = \frac{\pi}{2}$, and $r_M \approx 2.03$ given by $r_M = -\tan r_M$.

Since all trajectories of $X_r$ starting in $\left[0, \frac{\pi}{2}\right]$ are trapped inside this interval, the function $G(r, t)$ will be negative for all $r \in \left[0, \frac{\pi}{2}\right]$ and all $t \geq 0$. Similarly, the points in $[\pi, \pi + \delta)$ are pulled by the flow towards $r = \pi$ without ever crossing this point. Thus $G(r, t)$ will also be negative for all $r \in [\pi, \pi + \delta)$ and all $t \geq 0$.

It only remains to understand the behavior of $G(r, t)$ for $r \in \left(\frac{\pi}{2}, \pi\right)$. Since $f$ is strictly positive on this interval, it follows that, for every initial value $r \in \left(\frac{\pi}{2}, \pi\right)$,

$$F(r, \cdot) : \mathbb{R} \to \left(\tfrac{\pi}{2}, \pi\right)$$

is an orientation preserving diffeomorphism, and in particular there is a unique time $T_r \in \mathbb{R}$ such that $F(r, T_r) = r_M$.

For every fixed $r \in \left(\frac{\pi}{2}, r_M\right]$ and all positive $t \geq 0$, the upper bound of $G(r, t)$ in (4-1) is then given by

$$G(r, T_r) = \int_0^{T_r} g\big(F(r, s)\big)\, ds,$$

because $g$ is strictly positive up to $t = T_r$ so that $G(r, \cdot)$ increases up to that moment; for all later times $t > T_r$, the trajectory $F(r, t)$ lies in the zone $[r_M, \pi)$ where $g$ is negative so that $G(r, t) \leq G(r, T_r)$ for all $t \geq T_r$.

To compute $G(r, T_r)$ use that $F(r, \cdot) \colon \mathbb{R} \to \left(\frac{\pi}{2}, \pi\right)$ is for every choice of $r \in \left(\frac{\pi}{2}, \pi\right)$ a diffeomorphism, so that we can substitute $u = F(r, s)$ in the integral using that $\frac{du}{ds} = f\big(F(r, s)\big) = f(u)$, and obtain

$$G(r, T_r) = \int_0^{T_r} g\big(F(r, s)\big)\, ds = \int_{F(r,0)}^{F(r,T_r)} \frac{g(u)}{f(u)}\, du = \int_r^{r_M} \frac{u \cos u + \sin u}{u \sin u}\, du$$

$$= \ln\big(u \sin u\big)\Big|_r^{r_M} = \ln \frac{r_M \sin r_M}{r \sin r}.$$

The denominator $r \sin r$ is increasing on $\left[\frac{\pi}{2}, r_M\right]$ so that its smallest value on this interval is attained at $r = \frac{\pi}{2}$. We obtain the estimate

$$G(r, t) \leq \ln \frac{r_M \sin r_M}{r \sin r} < \ln \frac{r_M \sin r_M}{\frac{\pi}{2} \sin \frac{\pi}{2}} = \ln \frac{2 r_M \sin r_M}{\pi} < \ln \tfrac{7}{6}. \qquad \square$$

In particular, we obtain the following result.

**Corollary 4.3** *Let $L$ be a manifold that does not need to be closed, and let*

$$\Sigma := \mathbb{D}^2_{\leq \pi} \times \mathbb{D}_{<c_0}\big(T^* L\big)$$

*be a hypersurface in a contact manifold $(M, \xi)$ such that the singular distribution induced by $\xi$ on the hypersurface agrees with the kernel of the 1-form $\beta = r \sin r\, d\vartheta + \lambda_{\mathrm{can}}$, where $(r, \vartheta)$ denotes the polar coordinates on $\mathbb{D}^2_{\leq \pi}$.*

*Let $c > 0$ be such that $\mu_0 c < c_0$ with the constant $\mu_0 < \frac{7}{6}$ in Proposition 4.1. Then we can embed the contact domain*

$$\Big( B(h) \times \mathbb{D}_{<c}\big(T^* L\big), \ker\big(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}}\big) \Big)$$

*of "width" $c > 0$ and of any chosen "height" $h > 0$ into an arbitrarily small neighborhood of the hypersurface $\mathbb{D}^2_{\leq \pi} \times \mathbb{D}_{<c_0}\big(T^* L\big)$.*

**Proof** The induced singular distribution of a hypersurface determines by Proposition B.1 the germ of the contact structure on a neighborhood of the hypersurface. We can thus assume that $\Sigma$ has a neighborhood $U$ that is contactomorphic to $\big( B(\varepsilon) \times \mathbb{D}_{<c_0}\big(T^* L\big), \ker\big(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}}\big) \big)$ for small $\varepsilon > 0$. By Proposition 4.1 we can thus embed $B(h) \times \mathbb{D}_{<c}\big(T^* L\big)$ into $U$. $\qquad \square$

# Appendix A   The Legendrian unknot is loose in a sufficiently large overtwisted chart

In this appendix, we provide an elementary proof that Legendrian unknots are loose in ambient manifolds containing a large neighborhood of an overtwisted chart. This was first hinted at in [12]. A proof was given by Huang in [8] but he used piecewise smooth Legendrians so that a lot of the potential clarity was lost.

The key idea here is that for Legendrians that are in product form, any isotopy on the first factor can be trivially extended to the second one to obtain a global isotopy. If the Legendrian is only locally in product form, this construction only provides a local isotopy. To globalize it, we want to smoothen it out so that local isotopy glues to the identity outside the considered neighborhood. This interpolation requires a sufficient amount of space.

Except for a certain 3-dimensional result that is accepted as a black-box, the proof then boils down to a careful inspection of the original definition of looseness given by Murphy [11].

Let $\left(\mathbb{R}^{2n+1}, \xi_0 = \ker\left(dz - \sum_{j=1}^n y_j\, dx_j\right)\right)$ with coordinates $(\boldsymbol{x}, \boldsymbol{y}, z) = (x_1, \ldots, x_n, y_1, \ldots, y_n, z)$ be the standard contact space. The *Legendrian unknot* $\Lambda_0$ in $\mathbb{R}^{2n+1}$ can be given by the embedding

(A-1) $$\mathbb{S}^n \hookrightarrow (\mathbb{R}^{2n+1}, \xi_0),\ (\boldsymbol{x}, s) \mapsto \left(\boldsymbol{x}, -s\boldsymbol{x}, \tfrac{1}{3}s^3\right).$$

By extension, any Legendrian $\Lambda$ in a contact manifold is called a *Legendrian unknot* if there exists a Darboux chart containing $\Lambda$ in its interior such that $\Lambda$ agrees in the chart with $\Lambda_0$.

Let $(M, \xi)$ be a contact manifold. We want to study Legendrians in $M$ that look locally like product submanifolds in the following sense: Suppose that there is an open subset $U \subset M$ that is diffeomorphic to $U_M \times U_W$ where $U_M$ is a manifold that carries a contact form $\alpha$, and $U_W$ is an open Liouville domain with Liouville form $\lambda$ such that $\xi|_U = \ker(\alpha + \lambda)$. Then assume that a Legendrian $\Lambda$ satisfies

$$\Lambda \cap \left(U_M \times U_W\right) = L \times N,$$

where $L$ is a Legendrian in $\left(U_M, \alpha\right)$ and $N$ is an exact Lagrangian in $\left(U_W, \lambda\right)$ with $\lambda|_{TN} = 0$. We do not assume in general that $L$ or $N$ are closed.

The key notion we want to study in this appendix is due to Murphy [11].

**Definition**   Let $\Lambda$ be a Legendrian in the manifold $(M, \xi)$ that is locally in product form $L \times Z_\rho$ in a chart $\left(C \times V_\rho, \ker(\alpha_0 + \lambda_{\mathrm{can}})\right)$, where

- $C = \{(x, y, z) \in \mathbb{R}^3 \mid x, y, z \in (-1, 1)\}$ is a cube with side lengths 2 and standard contact form $\alpha_0 = dz - y\, dx$,
- $V_\rho = \{(\boldsymbol{q}, \boldsymbol{p}) \in \mathbb{R}^{2n-2} \mid \|\boldsymbol{q}\| < \rho,\ \|\boldsymbol{p}\| < \rho\}$ with Liouville form $\lambda_{\mathrm{can}} = -\sum_j p_j\, dq_j$,
- $L$ is a properly embedded Legendrian arc whose front is a zig-zag and which is equal to the set $\{y = z = 0\}$ near the boundary,
- $Z_\rho = \left\{(\boldsymbol{q}, \boldsymbol{p}) \in V_\rho \mid \boldsymbol{p} = \mathbb{0}\right\}$.

We say that $\Lambda$ is *loose* if $\rho > 1$.

**Remark A.1** If we replace the cube $C$ in the definition above with any cube of side lengths smaller than 2, then it can be seen with the argument in [11, Proposition 4.4] that the corresponding Legendrian is still loose.

The result we want to show in this appendix is the following proposition.

**Proposition A.2** *There exists a $\rho_0 > 0$ (that is independent of the dimension of $V_\rho$) such that the Legendrian unknot $\Lambda_0$ is loose in every contact manifold*

$$\left(B(1) \times V_\rho, \ker(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}})\right)$$

*for which $\rho > \rho_0$. Here, $\alpha_{\mathrm{OT}} = \cos(r)\, dz + r\sin(r)\, d\vartheta$ denotes the standard overtwisted contact form on $B(1)$; see also Section 3.*

**Proof** By [4] or [7], see also the details in [12], the Legendrian unknot is in any overtwisted 3-manifold the stabilization of another Legendrian knot $L_1$. More precisely, let $\left(B(1), \alpha_{\mathrm{OT}}\right)$ be the cylindrical box surrounding an overtwisted disk described in Section 3. We can assume that there is a Darboux chart $U_1$ centered around a point of $L_1$ such that

- the restriction of $\alpha_{\mathrm{OT}}$ agrees in the coordinates of the Darboux chart with the standard form $dz - y\, dx$,

- $U_1$ is a cube of size $\varepsilon_1 < 1$,

- $L_1 \cap U_1$ is the Legendrian arc $\{y = z = 0\}$.

We stabilize $L_1$ inside $U_1$ by adding a zig-zag. The resulting knot is then a Legendrian unknot $L_0$.

In particular, there exists a Darboux chart $U_0$ in $B(1)$ such that $L_0$ lies in standard position (A-1) inside $U_0$. The restriction $\alpha_{\mathrm{OT}}|_{U_0}$ with respect to the coordinates of $U_0$ will be of the form $e^{f(x,y,z)}(dz - y\, dx)$ for some smooth function $f$ that probably cannot be chosen to be equal to 0. Nonetheless, we can assume that there are constants $c_0 > 0$, and $C_0 > 1$ such that $c_0 \le e^f \le C_0$.

Using the Darboux chart $U_0$ in $B(1)$, we can find in the product contact manifold

$$\left(B(1) \times V_\rho, \ker(\alpha_{\mathrm{OT}} + \lambda_{\mathrm{can}})\right)$$

a higher-dimensional Darboux chart $U_0 \times V_{\rho'}$ for $\rho' = \rho/C_0$ embedded by

$$U_0 \times V_{\rho'} \to U_0 \times V_\rho, (x, y, z; \boldsymbol{q}, \boldsymbol{p}) \mapsto \left(x, y, z; \boldsymbol{q}, e^{f(x,y,z)}\boldsymbol{p}\right).$$

If $\rho' > 1$, we can embed the Legendrian unknot $\Lambda_0$ into this Darboux chart using the standard map (A-1)

$$\mathbb{S}^{n+1} \hookrightarrow \left(U_0 \times V_{\rho'}, \xi_0\right), (x_0, \boldsymbol{x}, s) \mapsto \left(x_0, -sx_0, \tfrac{1}{3}s^3; \boldsymbol{x}, -s\boldsymbol{x}\right).$$

The intersection of $\Lambda_0$ with the slice $B(1) \times \{\mathbb{0}\}$ is precisely the unknot $L_0$ in $U_0$ that we had singled out in $B(1)$.

By slightly deforming $\Lambda_0$ close to $L_0 \times \{0\}$ we can assume that $\Lambda_0$ is locally in product form $L_0 \times Z_\varepsilon$ for some small constant $\varepsilon > 0$. This can be easily seen using the front projection (even though "seeing" the front projection requires from dimension 7 on some experience). More explicitly, let $g \colon [0, 1] \to [0, 1]$ be a monotonous smooth function such that $g$ is equal to 0 in a neighborhood of 0 and 1 in a neighborhood of 1. We can then consider the deformed sphere $\mathbb{S}_g^{n+1} \subset \mathbb{R}^{n+2}$ given by the equation

$$x_0^2 + s^2 + g(x^2) \cdot x^2 = 1.$$

Interpolating linearly between the equation of the round sphere and the deformed one, one sees that both are isotopic to each other.

We define a Legendrian embedding of $\mathbb{S}_g^{n+1}$ by

$$\mathbb{S}_g^{n+1} \hookrightarrow \left(U_0 \times V_{\rho'}, \xi_0\right), (x_0, \boldsymbol{x}, s) \mapsto \left(x_0, -s x_0, \tfrac{1}{3} s^3; \boldsymbol{x}, -s \boldsymbol{x}\left(g(x^2) + x^2 g'(x^2)\right)\right).$$

We denote this deformed sphere by $\Lambda_0'$. It is Legendrian isotopic to the initial Legendrian unknot, and it is composed of a cylindrical part $L_0 \times Z_\delta$ for small values of $\boldsymbol{x}$ in $U_0 \times V_\delta$.

Recall that the Darboux chart $U_1 \times V_{\rho'}$ had been embedded into the ball $B(1) \times V_\rho$ via the map $(x, y, z; \boldsymbol{q}, \boldsymbol{p}) \mapsto \left(x, y, z; \boldsymbol{q}, e^{f(x,y,z)} \boldsymbol{p}\right)$. By looking at the preimage of this embedding, it follows that $\Lambda_0'$ also has a cylindrical segment in $B(1) \times V_\rho$. More explicitly, we have shown that after an isotopy the Legendrian unknot $\Lambda_0$ in $B(1) \times V_\rho$ has a cylindrical part of the form $L_0 \times Z_{\delta'}$ in the open ball $B(1) \times V_{\delta'}$ with $\delta' = \delta/c_0$.

We stretch out $\Lambda_0'$ in the $V_\rho$-direction of $B(1) \times V_\rho$ using an isotopy $(x, y, z; \boldsymbol{q}, \boldsymbol{p}) \mapsto \left(x, y, z; e^t \boldsymbol{q}, e^{-t} \boldsymbol{p}\right)$ where the maximal size of $t \geq 0$ depends on the width of $V_\rho$. If there is enough space, then we can suppose that the cylindrical part found above expands to be of the form $L_0 \times Z_\eta$ in the open ball $B(1) \times V_\eta$ with $\eta > 1$.

If we now consider the Darboux chart $U_1 \subset B(1)$, we see that the intersection of the deformed Legendrian sphere with $U_1 \times V_\eta$ is $L \times Z_\eta$, where $L$ is a Legendrian arc in $U_1$ whose front is a zig-zag, just as in the definition of looseness. If $\eta > 1$ and if $U_1$ is a cube of size smaller than 1 (see Remark A.1), then the deformed unknot and thus also $\Lambda_0$ are loose. $\square$

# Appendix B  Contact germ along a hypersurface

A folklore result states that a hypersurface in a contact manifold determines the germ of the contact structure surrounding it. Not having found a proof for dimension $> 3$ in the literature we have decided to add it here.

**Proposition B.1**  *Let $(M_0, \xi_0)$ and $(M_1, \xi_1)$ be two $(2n+1)$-dimensional contact manifolds, and let $\Sigma$ be a (not necessarily closed) $2n$-dimensional manifold. Assume that there are two embeddings*

$$\iota_0 \colon \Sigma \hookrightarrow M_0 \quad \text{and} \quad \iota_1 \colon \Sigma \hookrightarrow M_1$$

*such that the singular distributions $\mathcal{D}_0 = (D\iota_0)^{-1}(\xi_0)$ and $\mathcal{D}_1 = (D\iota_1)^{-1}(\xi_1)$ agree.*

*Then there exist a neighborhood $U_0 \subset M_0$ of $\iota_0(\Sigma)$, a neighborhood $U_1 \subset M_1$ of $\iota_1(\Sigma)$, and a contacto-morphism*

$$\Phi\colon (U_0, \xi_0) \to (U_1, \xi_1)$$

*such that $\Phi \circ \iota_0 = \iota_1$.*

**Remark B.2** To be able to apply Proposition B.1 to a hypersurface $\Sigma$ with nonempty boundary, one needs to attach a small collar along $\partial\Sigma$, and extend the embeddings $\iota_0$ and $\iota_1$ in such a way that the singular distributions $\mathcal{D}_0 = (D\iota_0)^{-1}(\xi_0)$ and $\mathcal{D}_1 = (D\iota_1)^{-1}(\xi_1)$ agree.

We split the proof of Proposition B.1 into several lemmas. The first one is due to Giroux, but we learned about it from [9].

**Lemma B.3** *Let $\Sigma$ be a (not necessarily closed) manifold carrying a (cooriented) singular distributions $\mathcal{D}$ that is given as the kernel of a 1-form $\beta$ such that $d\beta$ does not vanish at the singular points of $\mathcal{D}$; that is, at the points where $\beta = 0$.*

*If $\beta'$ is any other 1-form such that $\mathcal{D} = \ker \beta'$ inducing the same coorientation, and such that $d\beta'$ does not vanish either at the singular points of $\mathcal{D}$, then there exists a smooth positive function $f\colon \Sigma \to \,]0, \infty[$ such that*

$$\beta = f \cdot \beta'.$$

**Proof** Denote the set of all regular points of the distributions $\mathcal{D}$ by $U_{\mathrm{reg}} = \{p \in \Sigma \mid \mathcal{D}_p \neq T_p\Sigma\}$. On $U_{\mathrm{reg}}$, we can simply define $f$ to be the quotient $\beta(X)/\beta'(X)$, where $X$ is any vector field on $U_{\mathrm{reg}}$ that is transverse to $\mathcal{D}$. We are thus left with studying the singular points $p \notin U_{\mathrm{reg}}$ of $\mathcal{D}$, where $\beta$ and $\beta'$ both vanish, and proving that $f$ extends to a nonvanishing smooth function.

Use a coordinate chart for $\Sigma$ centered at $p \in \Sigma \setminus U_{\mathrm{reg}}$ with coordinates $\boldsymbol{x} = (x_1, \ldots, x_n)$. We can then write

$$\beta = g_1\,dx_1 + \cdots + g_n\,dx_n \quad \text{and} \quad \beta' = g'_1\,dx_1 + \cdots + g'_n\,dx_n$$

with functions $g_1, \ldots, g_n$ and $g'_1, \ldots, g'_n$ such that all $g_j$ and all $g'_j$ vanish at the origin. In fact for each $j$, the two functions $g_j$ and $g'_j$ vanish precisely on the same subset. By our assumption $d\beta' \neq 0$ at $p$, so that we can assume after possibly permuting the coordinates that $(\partial g'_1/\partial x_2)(\mathbb{0}) \neq 0$.

We will now show that $f$ extends in the chart smoothly to a neighborhood of the origin such that $g_1(\boldsymbol{x}) = f(\boldsymbol{x}) \cdot g'_1(\boldsymbol{x})$. Note that $\{\boldsymbol{x} \mid g'_1(\boldsymbol{x}) \neq 0\}$ is a subset of $U_{\mathrm{reg}}$ so that $f$ is a well-defined function on this subset. The condition $(\partial g'_1/\partial x_2)(\mathbb{0}) \neq 0$ allows us to apply the implicit function theorem to find a new set of coordinates $\boldsymbol{y} = (y_1, \ldots, y_n)$ for which $g'_1$ simplifies to $g'_1(\boldsymbol{y}) = y_2$. In this new chart, we obtain that $f$ is defined in particular for all points $\{y_2 \neq 0\} \subset U_{\mathrm{reg}}$.

Consider now the function $g_1$ represented with respect to the $\boldsymbol{y}$-coordinates. It also vanishes precisely along the hyperplane $\{y_2 = 0\}$ so that there exists a smooth functions $\widetilde{g}_1$ allowing us to write $g_1$ as

$$g_1(\boldsymbol{y}) = y_2\widetilde{g}_1(\boldsymbol{y});$$

see, for example, [10, Lemma 2.1]. Using this representation we see that $f(y) = \tilde{g}_1(y)$ extends to a smooth function on the whole chart so that it obviously satisfies the equation $g_1 = f \cdot g'_1$.

In particular, since $U_{\text{reg}}$ is dense in $\Sigma$ the continuous extension of $f$ is unique and does not depend on our choice of charts. This way, $f$ can be defined smoothly on all of $\Sigma$, and it satisfies $\beta = f \cdot \beta'$ on $\Sigma$.

It remains to prove that $f$ does not vanish anywhere, but this is clear because if $f$ ever vanished at a singular point $p$ of $\mathcal{D}$, we would find from $\beta = f \cdot \beta'$ that $d\beta = 0$ at $p$ — contrary to our assumption that $d\beta \neq 0$ along the singular set of $\mathcal{D}$. $\square$

**Lemma B.4** *Let $(M_0, \xi_0)$ and $(M_1, \xi_1)$ be two $(2n+1)$-dimensional contact manifolds with contact forms $\alpha_0$ and $\alpha_1$, respectively, and let $\Sigma$ be a (not necessarily closed) manifold of dimension $2n$.*

*Suppose that there are two embeddings*

$$\iota_0 \colon \Sigma \hookrightarrow (M_0, \xi_0) \quad \text{and} \quad \iota_1 \colon \Sigma \hookrightarrow (M_1, \xi_1)$$

*of $\Sigma$ into $M_0$ and $M_1$ such that $\iota_0^* \alpha_0 = \iota_1^* \alpha_1$.*

*Then, there is a bundle isomorphism*

$$\Phi \colon TM_0|_\Sigma \to TM_1|_\Sigma$$

*such that:*

(i) $\Phi|_{T\Sigma} = \text{id}_{T\Sigma}$ *(we identify here, and in the rest of the proof, $T\Sigma$ with the tangent spaces of $\iota_0(\Sigma)$ and $\iota_1(\Sigma)$).*

(ii) $\alpha_1 \circ \Phi = \alpha_0$ *on $TM_0|_\Sigma$.*

(iii) *The linear interpolation $(1 - \tau) \, d\alpha_0 + \tau \, (d\alpha_1 \circ \Phi)$ is for every $\tau \in [0, 1]$ a symplectic form on $\xi_0|_\Sigma$.*

**Proof** Denote the 1-form $\iota_0^* \alpha_0 = \iota_1^* \alpha_1$ by $\beta$. To construct the desired bundle isomorphism $\Phi$, we distinguish two types of subsets of $\Sigma$: Define

$$U_{\text{reg}} = \{p \in \Sigma \mid \beta_p \neq 0\} \quad \text{and} \quad U_{\text{symp}} = \{p \in \Sigma \mid d\beta_p^n \neq 0\}.$$

Both sets are open and their union covers all of $\Sigma$, because $d\beta = \iota_j^* d\alpha_j$ is at every point $p \in \Sigma$ where $T_p \Sigma = \xi_j(p)$ a maximally nondegenerate form on $T_p\Sigma$; that is, $d\beta$ is a symplectic form on $T_p\Sigma$ at every point $p$ where $\beta$ vanishes.

We construct now separate bundle isomorphisms over $U_{\text{reg}}$ and over $U_{\text{symp}}$ that we then glue together using a partition of unity.

Over $U_{\text{symp}}$, we can decompose $TM_j$ as $T\Sigma \oplus \text{span}(R_j)$, where $R_j$ is the Reeb vector field of $\alpha_j$. This allows us to define a first bundle isomorphism

$$\Phi_{\text{symp}} \colon TM_0|_{U_{\text{symp}}} \to TM_1|_{U_{\text{symp}}}$$

by $\Phi_{\text{symp}}(v + cR_0) = v + cR_1$ for every $v \in T\Sigma|_{U_{\text{symp}}}$ and every $c \in \mathbb{R}$. It is easy to verify that $\alpha_0$ and $\alpha_1 \circ \Phi_{\text{symp}}$ agree on $TM_0|_{U_{\text{symp}}}$. Furthermore, $d\alpha_0$ and $d\alpha_1 \circ \Phi_{\text{symp}}$ also agree, because for any pair of vectors $v, v' \in T\Sigma$, we have $d\alpha_j(v, v') = d\beta(v, v')$ on one hand, and $d\alpha_j(R_j, v) = 0$ on the other, for both $j = 0, 1$.

To construct a bundle isomorphism over $U_{\text{reg}}$, we define the *characteristic foliation* $\mathcal{F}$ on $\Sigma$. It is characterized over $U_{\text{reg}}$ as the (singular) subdistribution of $\ker \beta = T\Sigma \cap \xi$ on which $d\beta|_{\ker \beta}$ vanishes. A dimension count shows that $\mathcal{F}$ is of dimension 1.

Choose a volume form $d\operatorname{vol}_\Sigma$ on $\Sigma$, and let $X$ be the vector field determined by the equation

$$\iota_X d\operatorname{vol}_\Sigma = \beta \wedge (d\beta)^{n-1}.$$

Since $X$ only vanishes at points where $\beta$ vanishes, it follows that $X$ is everywhere on $U_{\text{reg}}$ nonsingular, and it is easy to convince oneself that the characteristic foliation is generated by $X$.

Choosing compatible complex structures $J_0$ on $(\xi_0, d\alpha_0)$ and $J_1$ on $(\xi_1, d\alpha_1)$, we define two vector fields $Y_0 = J_0 \cdot X$ and $Y_1 = J_1 \cdot X$ along $\iota_0(\Sigma)$ and $\iota_1(\Sigma)$, respectively. These vector fields are everywhere over $U_{\text{reg}}$ transverse to $\Sigma$ and they lie in the kernel of $\alpha_j$. This way, we can split the tangent bundles as

$$TM_j\big|_{U_{\text{reg}}} = T\Sigma\big|_{U_{\text{reg}}} \oplus \operatorname{span}(Y_j)\big|_{U_{\text{reg}}},$$

and use these decompositions to define the bundle isomorphism

$$\Phi_{\text{reg}} \colon TM_0|_{U_{\text{reg}}} \to TM_1|_{U_{\text{reg}}}$$

by $\Phi_{\text{reg}}(v + cY_0) = v + cY_1$ for every $v \in T\Sigma|_{U_{\text{reg}}}$ and every $c \in \mathbb{R}$. Again, we easily check that $\alpha_1 \circ \Phi_{\text{reg}}$ agrees on $TM_0|_{U_{\text{reg}}}$ with $\alpha_0$ so that $\Phi_{\text{reg}}(\xi_0|_{U_{\text{reg}}}) = \xi_1|_{U_{\text{reg}}}$.

To understand the interpolation between $d\alpha_0$ and $d\alpha_1 \circ \Phi_{\text{reg}}$, choose at a point $p \in U_{\text{reg}}$ a basis of $\xi_0(p)$ of the form $v_1, \dots, v_{2n-2}, X(p), Y_0(p)$, where the $v_j$ all lie in $\ker \beta$ and are complementary to $X(p)$. Assume they are ordered in such a way that

$$d\alpha_0^{n-1}(v_1, \dots, v_{2n-2}) = d\beta^{n-1}(v_1, \dots, v_{2n-2}) = d\alpha_1^{n-1}(v_1, \dots, v_{2n-2}) > 0.$$

Note that $d\alpha_0(X, \cdot)$ and $d\alpha_1(X, \cdot)$ vanish on all the vectors $v_1, \dots, v_{2n-2}, X$.

Define $d\alpha_\tau = (1-\tau) d\alpha_0 + \tau (d\alpha_1 \circ \Phi_{\text{reg}})$ for any $\tau \in [0, 1]$. Then we compute for all $\tau \in [0, 1]$ that

$$d\alpha_\tau^n(v_1, \dots, v_{2n-2}, X, Y_0) = n \, d\alpha_\tau(X, Y_0) \cdot d\alpha_\tau^{n-1}(v_1, \dots, v_{2n-2}) > 0,$$

because $d\alpha_\tau(X, Y_0) = (1-\tau) d\alpha_0(X, J_0 X) + \tau \, d\alpha_1(X, J_1 X) > 0$.

We glue now $\Phi_{\text{reg}}$ and $\Phi_{\text{symp}}$ to produce a global bundle isomorphism. Choose a smooth function $\rho \colon \Sigma \to [0, 1]$ with support in $U_{\text{reg}}$ such that $1 - \rho$ has support in $U_{\text{symp}}$ so that $\rho$ and $1 - \rho$ form a partition of unity subordinate to $\{U_{\text{reg}}, U_{\text{symp}}\}$. Define a bundle homomorphism

$$\Phi \colon TM_0|_\Sigma \to TM_1|_\Sigma$$

by mapping a vector $v \in T_p M_0$ at a point $p \in \Sigma$ to $\Phi(v) = \rho(p) \cdot \Phi_{\text{reg}}(v) + (1 - \rho(p)) \cdot \Phi_{\text{symp}}(v)$. It is obvious that $\Phi$ is a bundle homomorphism such that $\Phi|_{T\Sigma} = \text{id}_{T\Sigma}$ and such that $\alpha_0 = \alpha_1 \circ \Phi$ on $TM_0|_\Sigma$ proving properties (i) and (ii) in the lemma.

It remains to verify property (iii). Define the interpolation $d\alpha_\tau := (1 - \tau) \, d\alpha_0 + \tau \, (d\alpha_1 \circ \Phi)$ for $\tau \in [0, 1]$. Since $\Phi$ agrees with $\Phi_{\text{symp}}$ at the points where $\beta = 0$, we obtain that $d\alpha_\tau = d\alpha_0$ is nondegenerate at any such point. We study now the desired property at points at which $\beta \neq 0$ and thus $X \neq 0$.

Since $d\alpha_\tau|_{T\Sigma} = d\beta$ is independent of $\tau$, we see that $d\alpha_\tau(X, \cdot)$ vanishes on every vector that lies in $\ker \beta$. Using the same basis chosen above with $Y_0 = J_0 X$ and $Y_1 = J_1 X$, it follows that the sign of $d\alpha_\tau^n(v_1, \ldots, v_{2n-2}, X, Y_0) = n \, d\alpha_\tau(X, Y_0) \cdot d\beta^{n-1}(v_1, \ldots, v_{2n-2})$ only depends on the sign of the term $d\alpha_\tau(X, Y_0)$.

For this term we obtain $d\alpha_\tau(X, Y_0) = (1 - \tau) \, d\alpha_0(X, J_0 X) + \tau \, d\alpha_1\big(X, \Phi(Y_0)\big)$. The first term is clearly positive, and for the second one write $d\alpha_1\big(X, \Phi(Y_0)\big) = \rho \, d\alpha_1(X, J_1 X) + (1 - \rho) \, d\alpha_1\big(X, \Phi_{\text{symp}}(Y_0)\big)$, where the first term is again positive. Recall that $d\alpha_0 = d\alpha_1 \circ \Phi_{\text{symp}}$ so that we can simplify the second term as $d\alpha_1\big(X, \Phi_{\text{symp}}(Y_0)\big) = d\alpha_1\big(\Phi_{\text{symp}}(X), \Phi_{\text{symp}}(Y_0)\big) = \big(d\alpha_1 \circ \Phi_{\text{symp}}\big)(X, Y_0) = d\alpha_0(X, Y_0)$. Thus $d\alpha_\tau(X, Y_0) = (1 - \tau) \, d\alpha_0(X, Y_0) + \tau\big(\rho \, d\alpha_1(X, Y_1) + (1 - \rho) \, d\alpha_0(X, Y_0)\big)$ is positive as a convex combination of positive terms, and we have shown property (iii). $\qquad \square$

**Proof of Proposition B.1** Let $\alpha_0$ be a contact form for $\xi_0$, and let $\alpha_1$ be a contact form for $\xi_1$. By Lemma B.3 there is a smooth function $f : \Sigma \to \mathbb{R}_{>0}$ such that $\iota_0^* \alpha_0 = f \cdot \iota_1^* \alpha_1$. We would like to extend $f \circ \iota_1^{-1}$ to all of $M_1$ to normalize $\alpha_1$ globally; in general though, if $\Sigma$ is not closed, this might be impossible.

Denote the normal bundle of $\iota_1(\Sigma)$ in $M_1$ by $\nu_1 \Sigma \xrightarrow{\pi} \Sigma$, and recall that there is a tubular neighborhood $U_1$ of $\iota_1(\Sigma)$ that is diffeomorphic to a neighborhood $V_1$ of the 0-section in $\nu_1 \Sigma$ (of course $V_1$ will generally not have uniform radius in the fiber directions, when $\Sigma$ is not closed). The function $f \circ \pi$ is a smooth positive function on $\nu_1 \Sigma$. We will replace $M_1$ by the open subset $U_1$, and use $f \circ \pi$ to rescale $\alpha_1$ on $U_1$ so that we can assume that $\iota_0^* \alpha_0 = \iota_1^* \alpha_1$. This allows us to apply Lemma B.4 to obtain a bundle isomorphism $\Phi$ between $TM_0|_{\iota_0(\Sigma)}$ and $TM_1|_{\iota_1(\Sigma)}$.

Let $U_0$ be a tubular neighborhood of $\iota_0(\Sigma)$ in $M_0$ such that the exponential map $\exp_0$ (with respect to some Riemannian metric) defines a diffeomorphism $\exp_0 : V_0 \to U_0$, where $V_0$ is a neighborhood of the 0-section of the normal bundle of $\Sigma$ in $M_0$. Similarly, let $\exp_1$ be the exponential map on $M_1$. By suitably reducing the size of $U_0$ and $U_1$, we can assume that

$$\Psi := \exp_1 \circ \Phi \circ \exp_0^{-1} : U_0 \to U_1$$

is a diffeomorphism. To simplify our setup, pull-back $\alpha_1$ to $U_0$, and work in the fixed ambient manifold $U_0$. For simplicity we also write $\alpha_1$ for its pull-back. Then it follows that $U_0$ contains the submanifold $\Sigma$, and carries two contact structures given by contact forms $\alpha_0$ and $\alpha_1$ such that $\alpha_0$ and $\alpha_1$ agree at all points of $\Sigma$, and such that the linear interpolation of $d\alpha_0$ and $d\alpha_1$ is a path of symplectic structures on $\xi_0|_\Sigma = \xi_1|_\Sigma$.

The rest of the proof is an application of the Moser trick: Clearly the interpolation $\alpha_\tau := (1-\tau)\alpha_0 + \tau\alpha_1$ satisfies along $\Sigma$ for every $\tau \in [0, 1]$ the contact condition. There is thus a small neighborhood of $\Sigma$ in $U_0$ on which all $\alpha_\tau$ are contact forms.

As in the standard proof of Gray stability, we define a vector field $X_\tau$ on this neighborhood by the equations

$$\alpha_\tau(X_\tau) = 0 \quad \text{and} \quad d\alpha_\tau(X_\tau, \cdot) = f_\tau \alpha_\tau - \dot\alpha_\tau$$

with $f_\tau := \dot\alpha_\tau(R_\tau)$, where $R_\tau$ is the Reeb field of $\alpha_\tau$. Note that the right-hand side of the second equation vanishes along $\Sigma$, thus it follows that $X_\tau(p) = 0$ at every $p \in \Sigma$. By reducing to a smaller neighborhood of $\Sigma$ in $U_0$, the flow of $X_\tau$ will be defined up to time 1 giving a contact isotopy between $\xi_0$ and $\xi_1$.

Composing this isotopy with $\Psi$, we find the desired contactomorphism. $\qquad\square$

# References

[1]  **J Adachi**, *Plastikstufe with toric core*, preprint (2016)  arXiv 1609.00837

[2]  **M S Borman**, **Y Eliashberg**, **E Murphy**, *Existence and classification of overtwisted contact structures in all dimensions*, Acta Math. 215 (2015) 281–361  MR

[3]  **R Casals**, **E Murphy**, **F Presas**, *Geometric criteria for overtwistedness*, J. Amer. Math. Soc. 32 (2019) 563–604  MR

[4]  **K Dymara**, *Legendrian knots in overtwisted contact structures on $S^3$*, Ann. Global Anal. Geom. 19 (2001) 293–305  MR

[5]  **Y Eliashberg**, *Classification of overtwisted contact structures on 3-manifolds*, Invent. Math. 98 (1989) 623–637  MR

[6]  **Y Eliashberg**, *Classification of contact structures on $\mathbb{R}^3$*, Int. Math. Res. Not. 1993 (1993) 87–91  MR

[7]  **J B Etnyre**, *On knots in overtwisted contact structures*, Quantum Topol. 4 (2013) 229–264  MR

[8]  **Y Huang**, *On plastikstufe, bordered Legendrian open book and overtwisted contact structures*, J. Topol. 10 (2017) 720–743  MR

[9]  **P Massot**, *Sur quelques aspects riemanniens des structures de contact en dimension trois*, PhD thesis, École Normale Supérieure de Lyon (2008)  Available at `http://tel.archives-ouvertes.fr/tel-00347962`

[10]  **J Milnor**, *Morse theory*, Ann. of Math. Stud. 51, Princeton Univ. Press (1963)  MR

[11]  **E Murphy**, *Loose Legendrian embeddings in high dimensional contact manifolds*, preprint (2012)  arXiv 1201.2245

[12]  **E Murphy**, **K Niederkrüger**, **O Plamenevskaya**, **A I Stipsicz**, *Loose Legendrians and the plastikstufe*, Geom. Topol. 17 (2013) 1791–1814  MR

[13]  **K Niederkrüger**, *The plastikstufe: a generalization of the overtwisted disk to higher dimensions*, Algebr. Geom. Topol. 6 (2006) 2473–2508  MR

[14]  **K Niederkrüger**, *On fillability of contact manifolds*, habilitation à diriger des recherches, Université Paul Sabatier (2013)  Available at `http://tel.archives-ouvertes.fr/tel-00922320`

[15]  **K Niederkrüger**, **F Presas**, *Some remarks on the size of tubular neighborhoods in contact topology and fillability*, Geom. Topol. 14 (2010) 719–754  MR

*Department of Mathematics, National Cheng Kung University*
*Tainan City, Taiwan*

*Institut Camille Jordan UMR5208, Université Claude Bernard Lyon 1*
*Villeurbanne, France*

riverch@mail.ncku.edu.tw,  niederkruger@math.univ-lyon1.fr

# Presheaves of groupoids as models for homotopy types

LÉONARD GUETTA

We introduce the notion of groupoidal (weak) test category, which is a small category A such that the Grpd-valued presheaves over A model homotopy types in a "canonical and nice" way. The definition does not require *a priori* that A is a (weak) test category, but we prove two important comparison results: (1) every weak test category is a groupoidal weak test category, (2) a category is a test category *if and only if* it is a groupoidal test category.

As an application, we obtain new models for homotopy types, such as the category of groupoids internal to cubical sets with or *without* connections, the category of groupoids internal to cellular sets, the category of groupoids internal to semisimplicial sets, etc.

We also prove, as a by-product result, that the category of groupoids internal to the category of small categories models homotopy types.

18N40, 20L05; 55P99, 55U35, 55U40

## Introduction

In his famous manuscript *Pursuing stacks* from 1983 [10], Grothendieck introduced the theory of *test categories*. Informally speaking, a test category is a small category A such that the category $\hat{A}$ of Set-valued presheaves over A models homotopy types in a canonical way, which was axiomatized by Grothendieck. "Models homotopy types" means here that there is a particular class of morphisms on $\hat{A}$, such that the

localization of $\widehat{A}$ with respect to this class of morphisms is equivalent to the category **Hot** of *homotopy types*, that is, the category of CW-complexes and homotopy classes of continuous maps between them.[1] The archetypal example of a test category is of course the category $\Delta$ of finite nonempty ordinals, for which it has been known since the famous result of Milnor [19] that the category $\widehat{\Delta}$ of simplicial sets models homotopy types.

Examples of test categories abound, such as the cubical category [4, corollaire 8.4.13], the cubical category with connections [18], Joyal's $\Theta$ category [5], the dendroidal category $\Omega$ [1], etc. The point of view of the theory of test categories is that the category of presheaves on *any* test category should be, in some sense, as good a model of homotopy types as the category of simplicial sets. But modeling homotopy types is not the only good homotopical property of the category of simplicial sets. For example, we have the following results:

(i) The category $\mathsf{Ab}(\widehat{\Delta})$ of abelian groups internal to simplicial sets models *homology types*, that is, chain complexes in nonnegative degree up to quasi-isomorphisms[2] [7; 13].

(ii) The category $\mathsf{Grp}(\widehat{\Delta})$ of groups internal to simplicial sets models pointed connected homotopy types [14].

(iii) The category $\mathsf{Grpd}(\widehat{\Delta})$ of groupoids internal to simplicial sets models homotopy types [6, Theorem 8.3; 12, Theorem 10].

In this article, we focus on property (iii) above and prove its generalization, whose precise formulation requires the basic setup of the theory of test categories, which we recall now. For any small category A, Grothendieck considers the canonical functor

$$i_{\mathsf{A}} \colon \widehat{\mathsf{A}} \to \mathsf{Cat},$$

which sends an object $X$ of $\widehat{\mathsf{A}}$ to its category of elements $\mathsf{A}/X$. By a result attributed to Quillen by Illusie [11, chapitre VI, corollaire 3.3.1], the category Cat models homotopy types, when equipped with the class $\mathcal{W}_{\infty}$ of functors $u \colon \mathsf{C} \to \mathsf{D}$ that induce a weak homotopy equivalence between the classifying spaces of C and D. Using this, we can define the *homotopy type* of a presheaf $X$ over A as the category $i_{\mathsf{A}}(X) = \mathsf{A}/X$ seen as an object of $\mathsf{Cat}[\mathcal{W}_{\infty}^{-1}] \simeq \mathbf{Hot}$. This definition has a very natural interpretation: the category $\mathsf{A}/X$ is nothing but the so-called *Grothendieck construction* of $X$, and by a theorem of Thomason [21], this is the homotopy colimit of $X$, seen as a (discrete) space-valued presheaf. More generally, we define a class of weak equivalences in $\widehat{\mathsf{A}}$ as

$$\mathcal{W}_{\widehat{\mathsf{A}}} := i_{\mathsf{A}}^{-1}(\mathcal{W}_{\infty}),$$

and consider the induced functor at the level of homotopy categories

$$\bar{\imath}_{\mathsf{A}} \colon \widehat{\mathsf{A}}[\mathcal{W}_{\widehat{\mathsf{A}}}^{-1}] \to \mathsf{Cat}[\mathcal{W}_{\infty}^{-1}] \simeq \mathbf{Hot}.$$

---

[1]In practice, this equivalence of homotopy categories will always come from an equivalence at the level of $(\infty, 1)$-categories.

[2]The classical "Dold–Kan" equivalence even says that there is an equivalence of categories between $\mathsf{Ab}(\widehat{\Delta})$ and the category of chain complexes in nonnegative degree. However, from a homotopical point view, it is the homotopical result stated above which is relevant.

The category A is a *pseudotest category* if this functor is an equivalence of categories, a *weak test category* if the right adjoint of $i_A$ (which always exists) also preserves weak equivalences and induces an equivalence of homotopy categories, and, finally, a *test category* if it is a weak test category and for every object $a$ of A, $A/a$ is a weak test category.

What about $\mathsf{Grpd}(\widehat{A})$ now? There are two trivial but essential observations: (1) the category of groupoids internal to $\widehat{A}$ is equivalent to the category $[A^{op}, \mathsf{Grpd}]$ of presheaves over A with values in the category of groupoids, (2) the Grothendieck construction is also defined for $\mathsf{Grpd}$-valued presheaves, and so we can define a functor

$$I_A \colon \mathsf{Grpd}(\widehat{A}) \to \mathsf{Cat},$$

where $I_A(X)$ is the Grothendieck construction of $X$. The generalization of the situation for $\widehat{A}$ described earlier is then straightforward. We define a canonical class of weak equivalences on $\mathsf{Grpd}(\widehat{A})$ as

$$\mathcal{W}_{\mathsf{Grpd}(\widehat{A})} := I_A^{-1}(\mathcal{W}_\infty),$$

and consider the functor induced at the level of homotopy categories

$$\overline{I}_A \colon \mathsf{Grpd}(\widehat{A})[\mathcal{W}_{\mathsf{Grpd}(\widehat{A})}^{-1}] \to \mathsf{Cat}[\mathcal{W}_\infty^{-1}] \simeq \mathbf{Hot}.$$

Then, we can define the notions of *groupoidal pseudotest category*, *groupoidal weak test category* and *groupoidal test category* as perfect analogues of the usual notions for Set-valued presheaves.[3] In fact, almost all of the results from the usual theory also work in the groupoidal theory, once they have been appropriately adapted. Once the new theory is well established, we can compare the Set-valued notions with the Grpd-valued notions and we obtain the main results of this paper.

**Main Theorem** (Theorem 7.3 and Theorem 7.11) *Let* A *be a small category.*

(i) A *is a groupoidal test category* **if and only if** *it is a test category.*

(ii) *If* A *is a weak test category, then it is also a groupoidal weak test category.*

Before getting to the immediate applications of this theorem, let us comment on an important point.

Part (ii) of the previous theorem, which implies that for a (weak) test category A, the category $\mathsf{Grpd}(\widehat{A})$ models homotopy types, does not really come as a surprise, and here is a sketch of a very short proof of this mere fact (see Section 9 for details). One can easily show that if we equip the category $\mathsf{Grpd}(\mathsf{Cat})$ of groupoids internal to the category of small categories with the class of weak equivalences induced by the diagonal of a well-defined bisimplicial nerve, then for *every* weak test category A, we have an equivalence of homotopy categories[4]

$$\mathrm{Ho}(\mathsf{Grpd}(\widehat{A})) \simeq \mathrm{Ho}(\mathsf{Grpd}(\mathsf{Cat})).$$

---

[3]We will also define the notions of "groupoidal local test category" and "groupoidal strict test category".

[4]Note however that this does *not* work if we replace (internal) groupoids by (internal) groups. Indeed, simplicial groups model pointed connected homotopy types, but groups internal to categories (ie crossed modules) only model pointed connected homotopy 2-types (see Remark 9.10).

Then, by applying this isomorphism twice, if B is any other weak test category, we have

$$\text{Ho}(\text{Grpd}(\widehat{A})) \simeq \text{Ho}(\text{Grpd}(\widehat{B})).$$

In particular, for $B = \Delta$, we already know that $\text{Grpd}(\widehat{\Delta})$ models homotopy types; hence the desired result.

Nevertheless, the approach taken in this paper is *completely* different and the proof of the Main Theorem comes as an easy consequence of the theory of groupoidal test categories thoroughly developed beforehand. The advantage is at least twofold:

(1) We do not use the previously known case of $\Delta$, and hence we obtain a new (and much simpler) proof that $\text{Grpd}(\widehat{\Delta})$ models homotopy types.

(2) We also obtain part (i) of the Main Theorem, which gives a converse for groupoidal test categories: if A is a groupoidal test category, *then* A is a test category.

This second point is rather astonishing; it says that the fact that $\text{Grpd}(\widehat{A})$ models homotopy types (in a canonical way) is a necessary *and* sufficient condition for $\widehat{A}$ to model homotopy types. For the sake of comparison, let us consider the category $\text{Cat}(\widehat{A})$ of category objects in $\widehat{A}$. It is an easy exercise to show that this category models homotopy types (in a canonical way) if and only if the nerve of A is weakly contractible. In particular, it is *not* a sufficient condition to ensure that $\widehat{A}$ models homotopy types, and the theory of "categorical test categories" is essentially trivial. Somehow, this is not the case for Grpd-valued presheaves and requiring that a category is a groupoidal test category is, unexpectedly, not a weaker condition than requiring that it is a test category. As an application of the Main Theorem, we obtain a plethora of new models of homotopy types:

- the category $\text{Grpd}(\widehat{\square})$ of groupoids internal to cubical sets with or without connections,
- the category $\text{Grpd}(\widehat{\Theta})$ of groupoids internal to cellular sets,
- the category $\text{Grpd}(\widehat{\Omega})$ of groupoids internal to dendroidal sets,
- the category $\text{Grpd}(\widehat{\Delta'})$ of groupoids internal to semisimplicial sets,
- etc.

The example of cubical sets *without* connections is certainly worth noticing. Indeed, it is part of mathematical folklore that cubical groups without connections are not Kan in general (see [22] for example), and consequently it is believed that the category of cubical groups without connections does *not* model pointed connected homotopy types, contrary to simplicial groups. It can then come as a surprise that the category of groupoids internal to cubical sets without connections models homotopy types. Once again, Grpd-valued presheaves seem to behave particularly well and hopefully this justifies their study.

Let us end this introduction with a quick word on what is *not* treated in the present article. In his book [4], Cisinski showed that the category of (Set-valued) presheaves on any test category admits a model structure where the weak equivalences are the ones canonically defined by Grothendieck. The generalization of this for Grpd-valued presheaves (known when $A = \Delta$ [6; 12]) is not addressed at all here and is left as future work.

## Organization of the paper

Section 1 is a preliminary section recalling some basic homotopical algebra needed in the rest of the paper. Section 2 is a quick recollection, without proofs, of the classical theory of test categories. It is in Section 3 that we finally dive into the subject and give the basic setup of the homotopy theory of Grpd-valued presheaves, mimicking Grothendieck's axiomatic for the homotopy theory of Set-valued presheaves. In Sections 4, 5 and 6, we respectively develop the theory of groupoidal weak test categories, groupoidal test categories and groupoidal strict test categories. Section 7 is dedicated to the comparison of the classical theory with the groupoidal theory and in particular we obtain the Main Theorem stated previously in the introduction. Then, in Section 8, we give an alternative definition of the weak equivalences of Grpd-valued presheaves, which makes the link with previous existing works on simplicial groupoids. Finally, Section 9 is a "bonus" and almost independent section from the rest of the paper, where we prove that Grpd(Cat) models homotopy types.

## Acknowledgements

## 1 Preliminaries

**1.1** A *category with weak equivalences* is a pair $(\mathcal{C}, \mathcal{W})$, where $\mathcal{C}$ is a category and $\mathcal{W}$ is a class of morphisms of $\mathcal{C}$, generically referred to as the *weak equivalences*, which contains all isomorphisms and satisfy the 2-out-of-3 property. We say that $\mathcal{W}$ is *weakly saturated* if in addition it satisfies the following closure property: if $i : X \to Y$ and $r : Y \to X$ are morphisms of $\mathcal{W}$ such that $ri = \mathrm{id}_X$ and $ir \in \mathcal{W}$, then $r \in \mathcal{W}$ (and thus so is $i$ by 2-out-of-3).

When $\mathcal{C}$ has pullbacks, we say that a morphism $f : X \to Y$ is a *universal weak equivalence*, if for every pullback square

$$
\begin{array}{ccc}
X' & \longrightarrow & X \\
{\scriptstyle f'}\downarrow & \lrcorner & \downarrow{\scriptstyle f} \\
Y' & \longrightarrow & Y
\end{array}
$$

the morphism $f'$ is a weak equivalence (in particular $f$ is a weak equivalence).

By *homotopy category* of a category with weak equivalences $(\mathcal{C}, \mathcal{W})$, we mean the localization of $\mathcal{C}$ with respect to $\mathcal{W}$ [8]. It is denoted by $\mathrm{Ho}_{\mathcal{W}}(\mathcal{C})$, or simply $\mathrm{Ho}(\mathcal{C})$ when there is no risk of confusion.

**1.2** Let $(\mathcal{C}, \mathcal{W})$ and $(\mathcal{C}', \mathcal{W}')$ be two categories with weak equivalences. A functor $F : \mathcal{C} \to \mathcal{C}'$ is said to *preserve weak equivalences* if $F(\mathcal{W}) \subseteq \mathcal{W}'$. In this case, it induces the functor

$$\bar{F} : \mathrm{Ho}(\mathcal{C}) \to \mathrm{Ho}(\mathcal{C}')$$

at the level of homotopy categories. A *homotopy inverse* of such a functor is a functor $G : \mathcal{C}' \to \mathcal{C}$ which preserves weak equivalences and such that there exist zigzags of natural transformations $FG \leftrightsquigarrow \mathrm{id}_{\mathcal{C}'}$ and $GF \leftrightsquigarrow \mathrm{id}_{\mathcal{C}}$, which are pointwise weak equivalences (see (ii) below). In this case, $F$ induces an equivalence of categories $\mathrm{Ho}(\mathcal{C}) \simeq \mathrm{Ho}(\mathcal{C}')$ and similarly for $G$.[5]

Finally, we say that an adjunction $L : \mathcal{C} \rightleftarrows \mathcal{C}' : R$ is a *homotopical equivalence* if:

(i)   $L$ and $R$ preserve weak equivalences

$$L(\mathcal{W}) \subseteq \mathcal{W}' \quad \text{and} \quad R(\mathcal{W}') \subseteq \mathcal{W}.$$

(ii)   The unit and counit of the adjunction are pointwise weak equivalences: for every object $X$ of $\mathcal{C}'$ and every object $Y$ of $\mathcal{C}$, we have

$$\varepsilon_X : LR(X) \to X \in \mathcal{W}' \quad \text{and} \quad \eta_Y : Y \to RL(Y) \in \mathcal{W}.$$

Note that in this case $L$ and $R$ are homotopy inverses to each other.

The following lemma is a very useful criterion to detect homotopical equivalences.

**1.3 Lemma** *Let $(\mathcal{C}, \mathcal{W})$ and $(\mathcal{C}', \mathcal{W}')$ be two categories with weak equivalences and $L : \mathcal{C} \rightleftarrows \mathcal{C}' : R$ an adjunction. If $\mathcal{W} = L^{-1}(\mathcal{W}')$, then $L \dashv R$ is a homotopical equivalence if and only if the counit of the adjunction is a pointwise equivalence.*

**1.4 Remark** The dual of the previous lemma is also true, but we won't need it in this paper.

**Proof** First, let's prove that $R$ preserves weak equivalences. Let $f : X \to Y$ be a morphism a $\mathcal{C}'$ that belongs to $\mathcal{W}'$. Since $\mathcal{W} = L^{-1}(\mathcal{W}')$, we need to show that $LR(f) \in \mathcal{W}'$, and this follows from the 2-out-of-3 property of $\mathcal{W}'$, the commutativity of the square

$$
\begin{array}{ccc}
LR(X) & \xrightarrow{\ LR(f)\ } & LR(Y) \\
{\scriptstyle \varepsilon_X}\downarrow & & \downarrow{\scriptstyle \varepsilon_Y} \\
X & \xrightarrow{\ \ f\ \ } & Y
\end{array}
$$

and the fact that the counit is a pointwise weak equivalence.

Now let's prove that the unit of the adjunction is also a pointwise weak equivalence. Let $Y$ be an object of $\mathcal{C}'$. Since $\mathcal{W} = L^{-1}(\mathcal{W}')$, we need to show that $L(\eta_Y)$ is a weak equivalence. By the triangle identity, the triangle

$$
L(Y) \xrightarrow{\ L(\eta_Y)\ } LRL(Y) \xrightarrow{\ \varepsilon_{L(Y)}\ } L(Y)
$$
$$
\underset{\mathrm{id}_{L(Y)}}{\underbrace{\qquad\qquad\qquad\qquad}}
$$

is commutative; hence the desired result follows from the 2-out-of-3 property of $\mathcal{W}'$ and the fact that the counit is a pointwise weak equivalence. $\qquad\square$

---

[5]In fact, $F$ and $G$ are then even Dwyer–Kan equivalences [2]; hence they induce an equivalence of $(\infty, 1)$-categories.

# 2 Very brief recollection of the theory of test categories

We now provide only a quick summary of the basic notions and results (without proofs) of the theory of test categories. For a detailed exposition, we refer the reader to Maltsiniotis's book on the subject [17].

**2.1 Notation**   For a small category A, we denote by $\widehat{A}$ the category of Set-valued presheaves over A, that is, the category of functors $A^{op} \to Set$ and natural transformations between them.

We denote by Cat the category of small categories and functors between them. We use the notation $e$ for the terminal object of Cat, that is, the category with one object and no nonidentity morphism.

For a small category A and an object $a$ of A, we denote by $A/a$ the slice category of A over $a$. Explicitly, $A/a$ is the category whose objects are pairs $(a', p : a' \to a)$, where $a'$ and $p$ are respectively an object and a morphism of A, and whose morphisms $(a', p') \to (a'', p'')$ are morphisms $f : a' \to a''$ of A, such that $p'' \circ f = p'$. Note that we have an obvious projection functor $A/a \to A$.

More generally, if $u : A \to B$ is a morphism of Cat and $b$ is an object of B, we denote by $A/b$ the category whose objects are pairs $(a, q : u(a) \to b)$, where $a$ is an object of A and $q$ is a morphism of B, and whose morphisms $(a, q) \to (a', q')$ are the morphisms $f : a \to a'$ of A such that $q' \circ u(f) = q$. We make the abuse of notation of not making $u$ appear in the notation $A/b$, but this category obviously depends on $u$.

**2.2**   Let $\Delta$ be the category whose objects are the ordered sets $\Delta_n := \{0 < \cdots < n\}$ for $n \geq 0$ and whose morphisms are nondecreasing functions between them. The category $\widehat{\Delta}$ is referred to as the category of *simplicial sets*. The canonical inclusion $\Delta \hookrightarrow Cat$ induces the so-called *nerve functor*

$$N : Cat \to \widehat{\Delta}, \quad C \mapsto \big(\Delta_n \mapsto Hom_{Cat}(\Delta_n, C)\big).$$

Let us denote by $\mathcal{W}_\infty$ the class of morphisms $u : A \to B$ of Cat such that $N(u)$ is a weak homotopy equivalence of simplicial sets.[6] Recall now the fundamental result of the homotopy theory of Cat [11, chapitre VI, corollaire 3.3.1]: the nerve functor induces an equivalence

$$Ho(Cat) \simeq Ho(\widehat{\Delta})$$

at the level of homotopy categories. In other words, the category Cat equipped with $\mathcal{W}_\infty$ models homotopy types. In the theory of test categories, the point of view is reversed and $(Cat, \mathcal{W}_\infty)$ is taken as the fundamental model of homotopy types. As it happens, the results of this theory only rely on a few formal properties of the class $\mathcal{W}_\infty$, which are shared by other classes of weak equivalences (not necessarily modeling homotopy types), referred to as *basic localizers*, and whose definition is recalled below.

**2.3 Definition**   A class $\mathcal{W}$ of morphisms of Cat is called a *basic localizer* if it satisfies the following properties:

   (i)   $\mathcal{W}$ is weakly saturated.

---

[6]This means a weak equivalence of the Kan–Quillen model structure on simplicial sets.

(ii) For every small category A with a terminal object, the canonical morphism to the terminal category

$$A \to e$$

is in $\mathcal{W}$.

(iii) If for every object $c$ of C in any commutative triangle

$$A \xrightarrow{\;u\;} B$$
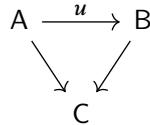


of Cat, the morphism induced by $u$,

$$A/c \to B/c,$$

is in $\mathcal{W}$, then $u$ is also in $\mathcal{W}$.

**2.4 Example** The class $\mathcal{W}_\infty$ is a basic localizer. It is in fact the smallest basic localizer [3]. More generally, for any $n \geq 0$, let $\mathcal{W}_n$ be the class of morphisms $u\colon A \to B$ of Cat such that $N(u)$ induces an equivalence on homotopy groups of simplicial sets, up to dimension $n$. Then, $\mathcal{W}_n$ is a basic localizer [4, section 9.2] and $(\mathsf{Cat}, \mathcal{W}_n)$ models homotopy $n$-types.

We now fix once and for all in this section a basic localizer $\mathcal{W}$ of Cat.

**2.5** A small category A is called $\mathcal{W}$-*aspherical*, or simply *aspherical*, when the canonical morphism to the terminal category $A \to e$ is in $\mathcal{W}$.

More generally, a morphism $u\colon A \to B$ is $\mathcal{W}$-*aspherical*, or simply *aspherical*, if for every object $b$ of B, the category $A/b$ is aspherical. Note that it follows from the axioms of basic localizers that *every aspherical morphism is in $\mathcal{W}$*. Practically, this is very often how we will prove that a morphism of Cat is a weak equivalence.

Finally, we say that a small category A is *totally $\mathcal{W}$-aspherical*, or simply *totally aspherical*, if it is aspherical and the diagonal functor $\delta\colon A \to A \times A$ is aspherical.

For later reference we put here the following lemma. We refer to [17, paragraphe 1.1.15] for the definition of Grothendieck fibrations.

**2.6 Lemma** *Let*

$$
\begin{array}{ccc}
A' & \xrightarrow{\;u\;} & A \\
p' \downarrow & \lrcorner & \downarrow p \\
B' & \xrightarrow[\;v\;]{} & B
\end{array}
$$

*be a pullback square of* Cat. *If $p$ is a Grothendieck fibration and $v$ is aspherical then $u$ is also aspherical.*

**Proof** This is a particular case of the dual of [17, théorème 3.2.15] combined with [17, exemple 3.2.2]. □

The first step of the theory of test categories is to equip $\widehat{A}$, for any small category A, with a canonical class of weak equivalences. For that, we use a canonical functor $\widehat{A} \to \mathsf{Cat}$.

**2.7** Let A be a small category. We write $i_A \colon \widehat{A} \to \mathsf{Cat}$ for the functor

$$i_A \colon \widehat{A} \to \mathsf{Cat}, \quad X \mapsto A/X,$$

where $A/X$ is the category of elements of $X$. In details, the objects of $A/X$ are pairs $(a, x)$, where $a$ is an object of A and $x \in X(a)$. A morphism $(a, x) \to (a', x')$ consists of a morphism $f \colon a \to a'$ in A such that $x = X(f)(x')$.

The functor $i_A$ has a right adjoint given by the formula

$$i_A^* \colon \mathsf{Cat} \to \widehat{A}, \quad C \mapsto \big(a \mapsto \mathrm{Hom}_{\mathsf{Cat}}(A/a, C)\big).$$

**2.8 Definition**  Let A be a small category. A morphism $\varphi \colon X \to Y$ of $\widehat{A}$ is a *$\mathcal{W}$-equivalence*, or simply a *weak equivalence*, if $i_A(\varphi)$ belongs to $\mathcal{W}$. We denote by $\mathcal{W}_{\widehat{A}}$ the class of $\mathcal{W}$-equivalences.

An object $X$ of $\widehat{A}$ is *$\mathcal{W}$-aspherical*, or simply *aspherical*, if $i_A(X)$ is a $\mathcal{W}$-aspherical category.

Finally, an object $X$ of $\widehat{A}$ is *$\mathcal{W}$-locally aspherical*, or simply *locally aspherical*, if for every object $a$ of A, the object $X|_{A/a}$ of $\widehat{A/a}$ defined as the composition

$$(A/a)^{\mathrm{op}} \longrightarrow A^{\mathrm{op}} \xrightarrow{\ X\ } \mathsf{Set},$$

is aspherical.

**2.9 Lemma**  *If* A *is aspherical, then an object $X$ of $\widehat{A}$ is aspherical if and only if the canonical morphism to the terminal presheaf $X \to *$ is a weak equivalence. In particular, every locally aspherical presheaf over an aspherical category* A *is aspherical.*

**Proof**  See [17, section 1.2.6]. $\qquad\square$

**2.10 Example**  As follows from [11, chapitre VI, théorème 3.3(ii)], in the case that $A = \Delta$, the $\mathcal{W}_\infty$-equivalences coincide with the usual weak homotopy equivalences, and thus a simplicial set is $\mathcal{W}_\infty$-aspherical if and only if it is weakly contractible.

**2.11 Definition**  (Grothendieck)  Let A be a small category. We say that:

(a)  A is a *$\mathcal{W}$-pseudotest category*, or simply a *pseudotest category*, if it satisfies both of the following conditions:

  (i)  A is aspherical.[7]

  (ii)  $i_A \colon \widehat{A} \to \mathsf{Cat}$ induces an equivalence $\mathrm{Ho}(\widehat{A}) \simeq \mathrm{Ho}(\mathsf{Cat})$ at the level of homotopy categories.

(b)  A is a *$\mathcal{W}$-weak test category*, or simply a *weak test category*, if the adjunction $i_A \dashv i_A^*$ is a homotopical equivalence (Section 1.2).

(c)  A is a *$\mathcal{W}$-local test category*, or simply a *local test category*, if for every object $a$ of A, the category $A/a$ is weak test.

---

[7]In fact, it follows from a result of Cisinski [4, proposition 4.2.4] and from [17, proposition 1.3.5] that condition (i) is implied by condition (ii).

(d) A is a *W-test category*, or simply a *test category*, if it is both a weak test category and a local test category.

(e) A is a *W-strict test category*, or simply a *strict test category*, if it is both totally aspherical and a test category.

**2.12 Remark**  We have the sequence of implications

$$\text{strict test} \implies \text{test} \implies \text{weak test} \implies \text{pseudotest},$$

but it can be shown that the converse of the first two implications do not hold. For the converse of the third one, it is still an open question.

**2.13 Example**  The archetypal example of a strict test category is $\Delta$ [17, proposition 1.5.13], but it is far from being the only one. The class of strict test categories also contains the cubical category with connections [18], Joyal's $\Theta$ category [5], etc. Examples of test categories which are not strict include the cubical category (without connections) [4, corollaire 8.4.13] and the dendroidal category $\Omega$ [1]. Examples of weak test categories which are not test include the subcategory $\Delta'$ of $\Delta$ with only monomorphisms as the morphisms [17, proposition 1.7.25]. All the examples above are of "shape-like" nature, but it is not always the case. For example, the monoid of nondecreasing functions $\mathbb{N} \to \mathbb{N}$, seen as a category with only one object, is a strict test category [5, exemple 3.16].

We now sum up the classical criteria to detect weak test, local test, test and strict test categories. For details and other characterizations, we refer to the first chapter of [17].

Recall that we denote by $\Delta_1$ the ordered set $\{0 < 1\}$.

**2.14 Proposition**  *Let* A *be a small category. We have the following characterizations*:

(a) A *is a weak test category if and only if for every small category* C *with a terminal object, the presheaf* $i_A^*(C)$ *is aspherical.*

(b) A *is a local test category if and only if the presheaf* $i_A^*(\Delta_1)$ *is locally aspherical.*

(c) A *is a test category if and only if it is aspherical and* $i_A^*(\Delta_1)$ *is locally aspherical.*

(d) A *is a strict test category if and only if it is totally aspherical and* $i_A^*(\Delta_1)$ *is aspherical.*

**Proof**  The four items are respectively proposition 1.3.9, théorème 1.5.6, remarque 1.5.4(a) and proposition 1.6.8 of [17]. □

Let us end this section with a quick word on aspherical functors and locally aspherical functors.

**2.15 Definition**  Let A be a small category, $i : A \to \mathsf{Cat}$ a functor and let $i^* : \mathsf{Cat} \to \widehat{A}$ be the functor defined as

$$i^* : \mathsf{Cat} \to \widehat{A}, \quad C \mapsto \big(a \mapsto \mathrm{Hom}_{\mathsf{Cat}}(i(a), C)\big).$$

We say that $i$ is a *W-aspherical functor*, or simply an *aspherical functor*, if it satisfies the two following conditions:

(a)   $i(a)$ has a terminal object for every object $a$ of A.

(b)   If C is a small category with a terminal object, $i^*(C)$ is an aspherical object of $\widehat{A}$.

We say that $i$ is a *W-locally aspherical functor*, or simply a *locally aspherical functor*, if it satisfies condition (a) above and the following condition instead of (b):

(b′)   $i^*(\Delta_1)$ is a locally aspherical object of $\widehat{A}$.

**2.16 Remark**   The definition of aspherical functor given above is not the one given in [17, définition 1.7.1], but is a particular case of the latter one, as follows from [17, proposition 1.7.6(d)]. This variant will be sufficient for our purpose. (See also Remark 4.7 below.)

**2.17 Remark**   The first item (resp second item) of Proposition 2.14 can be reformulated as: A is a weak test category (resp local test category) if and only if A → Cat, $a \mapsto A/a$ is an aspherical functor (resp locally aspherical functor).

**2.18 Proposition**   *Let A be a small category and $i : A \to$ Cat a functor. We have the following implications*:

(a)   *If A is weak test and $i$ is an aspherical functor, then $i^*:$ Cat $\to \widehat{A}$ is a homotopy inverse of $i_A$.*

(b)   *If $i$ is a locally aspherical functor, then A is a local test category.*

(c)   *If $i$ is a locally aspherical functor and A is aspherical, then A is a test category and $i^*:$ Cat $\to \widehat{A}$ is a homotopy inverse of $i_A$.*

**Proof**   This follows from [17, proposition 1.7.6 and théorème 1.7.13].   □

**2.19 Example**   The archetypal example of aspherical functor (which is even locally aspherical) is the inclusion functor $i : \Delta \hookrightarrow$ Cat [17, exemple 1.7.18]. Then $i^*:$ Cat $\to \widehat{\Delta}$ is the nerve functor. Hence, we recover via the above proposition that the nerve induces an equivalence of homotopy categories $\mathrm{Ho}(\mathrm{Cat}) \simeq \mathrm{Ho}(\widehat{\Delta})$.

**2.20 Remark**   The name *locally aspherical functor* is nonstandard, but by (b) of the previous proposition, it is equivalent to the usual notion of *local test functor*. Similarly, an aspherical functor $i : A \to$ Cat such that A is a (weak) test category, is usually called a (*weak*) *test functor*. We will not use this terminology.

# 3   Homotopy theory of presheaves of groupoids

**3.1 Notation**   We denote by Grpd the category of (small) groupoids and for a small category A, we denote by $\widehat{A}_{\mathrm{Grpd}}$ the category of Grpd-valued presheaves over A. That is, $\widehat{A}_{\mathrm{Grpd}}$ is the category of functors $A^{\mathrm{op}} \to$ Grpd and natural transformations between them.[8] We use the notation $*$ for the terminal object of $\widehat{A}_{\mathrm{Grpd}}$.

---

[8]By that, we mean actual *strict* natural transformations and not *pseudo*natural transformations.

The canonical inclusion $\mathsf{Set} \hookrightarrow \mathsf{Grpd}$, which identifies sets with discrete groupoids, induces a canonical inclusion $\widehat{\mathsf{A}} \hookrightarrow \widehat{\mathsf{A}}_{\mathsf{Grpd}}$. Hence, every $\mathsf{Set}$-valued presheaf can be seen as a $\mathsf{Grpd}$-valued presheaf.

**3.2** Let $\mathsf{A}$ be a small category and $X$ an object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$. We write $\mathsf{A} /\!\!/ X$ for the *category of elements of $X$*, which is defined as the following:

- An object is a pair $(a, x)$, where $a$ is an object of $\mathsf{A}$ and $x$ is an object of $X(a)$.

- A morphism $(a, x) \to (a', x')$ is a pair $(f, k)$ where $f : a \to a'$ is a morphism of $\mathsf{A}$ and $k : x \xrightarrow{\sim} X(f)(x')$ is a morphism of $X(a)$ (which is necessarily an isomorphism).

The identity morphism of $(a, x)$ is given by $(\mathrm{id}_a, \mathrm{id}_x)$ and the composition of

$$(a, x) \xrightarrow{(f,k)} (a', x') \xrightarrow{(f',k')} (a'', x'')$$

is given by

$$(f' \circ f, k'') : (a, x) \to (a'', x''),$$

where $k''$ is the composite of

$$x \xrightarrow{k} X(f)(x') \xrightarrow{X(f)(k')} X(f' \circ f)(x'').$$

For every object $X$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, the category $\mathsf{A} /\!\!/ X$ comes equipped with a canonical morphism

$$\zeta_X : \mathsf{A} /\!\!/ X \to \mathsf{A}, \quad (a, x) \mapsto a,$$

which is easily checked to be a Grothendieck fibration (see also Remark 3.3 below).

Given a morphism $\alpha : X \to X'$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ (ie a natural transformation), we define a functor $\mathsf{A} /\!\!/ X \to \mathsf{A} /\!\!/ X'$ in the following way:

- An object $(a, x)$ of $\mathsf{A} /\!\!/ X$ is sent to the object $(a, \alpha_a(x))$ of $X'(a)$.

- A morphism $(f, k) : (a, x) \to (a', x')$ of $\mathsf{A} /\!\!/ X$ is sent to the morphism

$$(f, \alpha_a(k)) : (a, \alpha_a(x)) \to (a', \alpha_{a'}(x'))$$

of $\mathsf{A} /\!\!/ X'$ (where we used the naturality of $\alpha$ for the target of this morphism to be compatible).

This makes the correspondence $X \mapsto \mathsf{A} /\!\!/ X$ functorial in $X$, and yields a functor denoted by $I_{\mathsf{A}}$:

$$I_{\mathsf{A}} : \widehat{\mathsf{A}}_{\mathsf{Grpd}} \to \mathsf{Cat}, \quad X \mapsto \mathsf{A} /\!\!/ X.$$

We shall see later that $I_{\mathsf{A}}$ admits a right adjoint.

**3.3 Remark** Via the canonical inclusion $\mathsf{Grpd} \hookrightarrow \mathsf{Cat}$, any $\mathsf{Grpd}$-valued presheaf $X$ can be seen as a $\mathsf{Cat}$-valued presheaf. Then, $\mathsf{A} /\!\!/ X$ is nothing but the Grothendieck construction of $X$ (see Section 8.1 for details) and the canonical morphism $\zeta_X : \mathsf{A} /\!\!/ X \to \mathsf{A}$ is the Grothendieck fibration associated to $X$.

**3.4 Remark** When $X$ is an object of $\widehat{\mathsf{A}}$, which we see as an object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ via the canonical inclusion $\widehat{\mathsf{A}} \hookrightarrow \widehat{\mathsf{A}}_{\mathsf{Grpd}}$, we have

$$i_{\mathsf{A}}(X) = I_{\mathsf{A}}(X).$$

In other words, the following triangle is commutative:

$$\widehat{\mathsf{A}} \lhook\joinrel\longrightarrow \widehat{\mathsf{A}}_{\mathsf{Grpd}}$$

$$i_{\mathsf{A}} \searrow \quad \downarrow I_{\mathsf{A}}$$

$$\mathsf{Cat}$$

**3.5 Remark** By an obvious variation of the Yoneda lemma, the category $I_{\mathsf{A}}(X)$ can alternatively be defined as the category whose objects are pairs $(a, p\colon a \to X)$ (we identify an object $a$ of A with the Set-valued presheaf represented by $a$) and whose morphisms $(a, p) \to (a', p')$ are pairs $(f, \sigma)$, where $f\colon a \to a'$ is a morphism of A and $\sigma\colon p \overset{\simeq}{\Longrightarrow} p' \circ f$ is a natural isomorphism.

For the rest of this section, we fix once and for all a basic localizer $\mathcal{W}$ of Cat.

**3.6 Definition** Let A be a small category. A morphism of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$

$$\varphi\colon X \to Y$$

is a $\mathcal{W}$-*equivalence*, or simply a *weak equivalence*, if the induced morphism of Cat

$$I_{\mathsf{A}}(\varphi)\colon I_{\mathsf{A}}(X) \to I_{\mathsf{A}}(Y)$$

is in $\mathcal{W}$. We denote by $\mathcal{W}_{\widehat{\mathsf{A}}_{\mathsf{Grpd}}}$ the class of $\mathcal{W}$-equivalences.

An object $X$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ is $\mathcal{W}$-*aspherical*, or simply *aspherical*, if the category $I_{\mathsf{A}}(X)$ is aspherical.

**3.7 Remark** It follows from Remark 3.4 that a Set-valued presheaf is aspherical in the sense of Definition 2.8 if only if it is aspherical in the sense of the previous definition (using the canonical inclusion $\widehat{\mathsf{A}} \hookrightarrow \widehat{\mathsf{A}}_{\mathsf{Grpd}}$).

In the case that the category A is aspherical, there is an equivalent characterization of aspherical objects of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ as stated in the following lemma.

**3.8 Lemma** *If* A *is aspherical, then for every object* $X$ *of* $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$*, we have the equivalence*

$$X \to * \text{ is a weak equivalence} \iff X \text{ is aspherical}.$$

*Conversely, if this equivalence is true for every* $X$ *in* $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$*, then* A *is aspherical.*

**Proof** Notice that $I_{\mathsf{A}}(*) \simeq \mathsf{A}$; hence $*$ is aspherical if and only if A is aspherical. The equivalence follows then from the obvious fact that for a morphism $\varphi\colon X \to Y$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, if $Y$ is aspherical, then $\varphi$ is a weak equivalence if and only if $X$ is aspherical.

For the second part of the lemma, it suffices to notice that the identity morphism $* \to *$ is always a weak equivalence (as all identity morphisms are) and $I_{\mathsf{A}}(*) \simeq \mathsf{A}$. $\qquad\square$

For later reference, we put here the following result which gives other characterizations of aspherical morphisms of Cat.

**3.9 Proposition** *Let $u\colon A \to B$ be a morphism of $\mathsf{Cat}$. The following conditions are equivalent:*

(a) *$u$ is aspherical.*

(b) *The functor $u^*\colon \widehat{B}_{\mathsf{Grpd}} \to \widehat{A}_{\mathsf{Grpd}}$ preserves and reflects aspherical objects: an object $X$ of $\widehat{B}_{\mathsf{Grpd}}$ is aspherical if and only if $u^*(X)$ is aspherical.*

(c) *The functor $u^*\colon \widehat{B}_{\mathsf{Grpd}} \to \widehat{A}_{\mathsf{Grpd}}$ preserves aspherical objects: for every aspherical object $X$ of $\widehat{B}_{\mathsf{Grpd}}$, $u^*(X)$ is aspherical.*

*All these equivalent conditions imply the following condition:*

(d) *The functor $u^*\colon \widehat{B}_{\mathsf{Grpd}} \to \widehat{A}_{\mathsf{Grpd}}$ preserves and reflects weak equivalences: $(u^*)^{-1}(\mathcal{W}_{\widehat{A}_{\mathsf{Grpd}}}) = \mathcal{W}_{\widehat{B}_{\mathsf{Grpd}}}$.*

*If $A$ and $B$ are aspherical, then conditions (a) to (d) are all equivalent and equivalent to the following condition:*

(e) *The functor $u^*$ preserves weak equivalences: $u^*(\mathcal{W}_{\widehat{B}_{\mathsf{Grpd}}}) \subseteq \mathcal{W}_{\widehat{A}_{\mathsf{Grpd}}}$.*

**Proof** Let us start with some preliminaries. It is easily checked that for every object $X$ of $\widehat{B}_{\mathsf{Grpd}}$, the square

$$
\begin{array}{ccc}
I_A(u^*(X)) & \xrightarrow{\ \lambda_X\ } & I_B(X) \\
{\scriptstyle \zeta_{u^*(X)}}\downarrow & \lrcorner & \downarrow{\scriptstyle \zeta_X} \\
A & \xrightarrow[\ u\ ]{} & B
\end{array}
$$

where $\lambda_X$ is the functor defined on objects as

$$(a, x) \mapsto (u(a), x),$$

and on morphisms as

$$\big((a, x) \xrightarrow{(f, \sigma)} (a', x')\big) \mapsto \big((u(a), x) \xrightarrow{(u(f), \sigma)} (u(a'), x')\big),$$

is a pullback square.

Now, using the fact that the vertical morphisms of the previous pullback square are Grothendieck fibrations, it follows from Lemma 2.6 that if $u$ is aspherical, then $\lambda_X$ is aspherical. In particular, with these conditions, $X$ is aspherical if and only if $u^*(X)$ is aspherical, which proves the implication (a) $\Longrightarrow$ (b). The implication (b) $\Longrightarrow$ (c) is tautological. To prove (c) $\Longrightarrow$ (a), consider an object $b$ of $B$, seen as a $\mathsf{Set}$-valued representable presheaf (and as an object of $\widehat{A}_{\mathsf{Grpd}}$ via the inclusion $\widehat{A} \hookrightarrow \widehat{A}_{\mathsf{Grpd}}$). It is straightforward to check that $I_A(u^*(b)) \simeq A/b$. Hence, if condition (c) is satisfied, then $u$ is aspherical.

To prove (c) $\Longrightarrow$ (d), it suffices to notice that $\lambda_X$ is natural in $X$, and thus, for every morphism $f\colon X \to Y$ of $\widehat{B}_{\mathsf{Grpd}}$, we have a commutative square

$$
\begin{array}{ccc}
I_A(u^*(X)) & \xrightarrow{\ I_A(u^*(f))\ } & I_A(u^*(Y)) \\
{\scriptstyle \lambda_X}\downarrow & & \downarrow{\scriptstyle \lambda_Y} \\
I_B(X) & \xrightarrow[\ I_B(f)\ ]{} & I_B(Y)
\end{array}
$$

We have already seen that if $u$ is aspherical, then the vertical arrows of the previous square are weak equivalences. Hence, in this case, $I_{\mathsf{B}}(f)$ is a weak equivalence if and only if $I_{\mathsf{A}}(u^*(f))$ is. By definition of weak equivalences of Grpd-valued presheaves, this means exactly that $u^*$ preserves and reflects weak equivalences.

The implication (d) $\implies$ (e) is trivial. Finally, let us prove (e) $\implies$ (c). Let $X$ be an aspherical object of $\widehat{\mathsf{B}}_{\mathsf{Grpd}}$. Since $\mathsf{B}$ is aspherical, we have already seen (Lemma 3.8) that this means exactly that the canonical morphism

$$X \to *$$

is a weak equivalence of $\widehat{\mathsf{B}}_{\mathsf{Grpd}}$. If $u^*$ preserves weak equivalences, then

$$u^*(X) \to u^*(*) \simeq *$$

is a weak equivalence of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$. Using that $\mathsf{A}$ is aspherical, we deduce from Lemma 3.8 again that $u^*(X)$ is aspherical. Hence, $u^*$ preserves aspherical objects. $\qquad\square$

**3.10 Definition** A small category $\mathsf{A}$ is $\mathcal{W}$-*groupoidal pseudotest*, or simply *groupoidal pseudotest*, if:

(a) $\mathsf{A}$ is aspherical.

(b) $I_{\mathsf{A}}$ induces an equivalence at the level of localized categories,

$$\mathrm{Ho}(\widehat{\mathsf{A}}_{\mathsf{Grpd}}) \xrightarrow{\sim} \mathrm{Ho}(\mathsf{Cat}).$$

# 4 Groupoidal weak test categories

**4.1 Notation** For two (small) categories $\mathsf{C}$ and $\mathsf{D}$, we denote by $\underline{\mathrm{Hom}}^{\mathrm{iso}}_{\mathsf{Cat}}(\mathsf{C}, \mathsf{D})$ the groupoid whose objects are functors $\mathsf{C} \to \mathsf{D}$ and whose morphisms are natural *isomorphisms* between those functors.

**4.2** Let $i \colon \mathsf{A} \to \mathsf{Cat}$ be a functor with $\mathsf{A}$ a small category. We denote by $I^*$ the functor

$$I^* \colon \mathsf{Cat} \to \widehat{\mathsf{A}}_{\mathsf{Grpd}}, \quad \mathsf{C} \mapsto \underline{\mathrm{Hom}}^{\mathrm{iso}}_{\mathsf{Cat}}(i(-), \mathsf{C}).$$

When $i$ is the functor $\mathsf{A} \to \mathsf{Cat}$, $a \mapsto \mathsf{A}/a$, we use the special notation $I^*_{\mathsf{A}}$ for the functor $I^*$. In other words, for a small category $\mathsf{C}$, $I^*_{\mathsf{A}}$ is the functor

$$I^*_{\mathsf{A}} \colon \mathsf{Cat} \to \widehat{\mathsf{A}}_{\mathsf{Grpd}}, \quad \mathsf{C} \mapsto \underline{\mathrm{Hom}}^{\mathrm{iso}}_{\mathsf{Cat}}(\mathsf{A}/(-), \mathsf{C}).$$

The following lemma is without a doubt part of mathematical folklore and I claim no originality for it.

**4.3 Lemma** *Let* $i \colon \mathsf{A} \to \mathsf{Cat}$ *be a functor, with* $\mathsf{A}$ *a small category. The functor* $I^* \colon \mathsf{Cat} \to \widehat{\mathsf{A}}_{\mathsf{Grpd}}$ *has a left adjoint. In the case that* $i$ *is the functor* $a \mapsto \mathsf{A}/a$ *(and thus* $I^* = I^*_{\mathsf{A}}$*), this left adjoint is* $I_{\mathsf{A}} \colon \widehat{\mathsf{A}}_{\mathsf{Grpd}} \to \mathsf{Cat}$ *(*Section 3.2*).*

**Proof** Denote by $\int^{a \in \mathsf{A}} F(a, a)$ the coend, and by $\int_{a \in \mathsf{A}} F(a, a)$ the end, of a functor $F : \mathsf{A}^{\mathrm{op}} \times \mathsf{A} \to \mathsf{Cat}$. We define a functor $I_! : \widehat{\mathsf{A}}_{\mathsf{Grpd}} \to \mathsf{Cat}$ as

$$X \mapsto \int^{a \in \mathsf{A}} X(a) \times i(a).$$

For every small category $\mathsf{C}$ and every $\mathsf{Grpd}$-valued presheaf $X$, we then have the following sequence of natural isomorphisms:

$$\mathrm{Hom}_{\mathsf{Cat}}\Big(\int^{a \in \mathsf{A}} X(a) \times i(a), \mathsf{C}\Big) \simeq \int_{a \in \mathsf{A}} \mathrm{Hom}_{\mathsf{Cat}}(X(a) \times i(a), \mathsf{C})$$

$$\simeq \int_{a \in \mathsf{A}} \mathrm{Hom}_{\mathsf{Grpd}}\big(X(a), \underline{\mathrm{Hom}}_{\mathsf{Cat}}^{\mathrm{iso}}(i(a), \mathsf{C})\big)$$

$$\simeq \mathrm{Hom}_{\widehat{\mathsf{A}}_{\mathsf{Grpd}}}(X, I^*(\mathsf{C})).$$

Hence, $I_!$ is left adjoint of $I^*$. In the case that $i : \mathsf{A} \to \mathsf{Cat}$ is the functor $a \mapsto \mathsf{A}/a$, we have

$$I_!(X) = \int^{a \in \mathsf{A}} X(a) \times \mathsf{A}/a,$$

which is nothing but the Grothendieck construction of $X$. To see this, recall that the Grothendieck construction of a $\mathsf{Cat}$-valued functor (and, a fortiori, for a $\mathsf{Grpd}$-valued functor) is its oplax colimit [9], and that the oplax colimit of a contravariant functor is computed as the colimit weighted by the slices of the source [20]. The conclusion follows then from Remark 3.3. $\qquad \square$

We now fix, once and for all in this section, a basic localizer $\mathcal{W}$ of $\mathsf{Cat}$.

**4.4 Definition** A small category $\mathsf{A}$ is $\mathcal{W}$-*groupoidal weak test*, or simply *groupoidal weak test*, if the adjunction $I_{\mathsf{A}} \dashv I_{\mathsf{A}}^*$ is a homotopical equivalence (Section 1.2).

**4.5 Remark** An immediate computation shows that $I_{\mathsf{A}} I_{\mathsf{A}}^*(e) \simeq \mathsf{A}$. Thus, if $\mathsf{A}$ is groupoidal weak test, the counit morphism $\mathsf{A} \to e$ is a weak equivalence and so $\mathsf{A}$ is aspherical. This shows that every groupoidal weak test category is a groupoidal pseudotest category.

We would like now to find characterizations of groupoidal weak test categories. For that, we begin by studying a class of homotopically well-behaved functors $\mathsf{A} \to \mathsf{Cat}$.

**4.6 Definition** Let $\mathsf{A}$ be a small category. A functor $i : \mathsf{A} \to \mathsf{Cat}$ is $\mathcal{W}$-*groupoidal aspherical*, or simply *groupoidal aspherical*, if:

 (a)  For every object $a$ of $\mathsf{A}$, the category $i(a)$ has a terminal object.

 (b)  For every small category $\mathsf{C}$ with a terminal object, the $\mathsf{Grpd}$-valued presheaf $I^*(\mathsf{C})$ is aspherical.

**4.7 Remark** A more general notion of groupoidal aspherical functor is obtained by replacing the conditions of the previous definitions by:

 (a′)  $i(a)$ is aspherical for every object $a$ of $\mathsf{A}$.

 (b′)  For every small aspherical category $\mathsf{C}$, $I^*(\mathsf{C})$ is aspherical.

This is the straightforward generalization of [17, définition 1.7.1]. (We will see in Proposition 4.11 that when (a) is satisfied, conditions (b) and (b$'$) are equivalent.) However, the author does not know if the theory fully works for this more general notion of groupoidal aspherical functor, and the restricted version we chose is sufficient for our purpose. For the interested reader, the difficulty is that it seems that [17, lemme 1.7.4(b)] cannot be generalized to Grpd-valued presheaves. (Nevertheless, see Lemma 7.8(b) below for a *partial* generalization of this lemma.)

**4.8**  Let $i \colon \mathsf{A} \to \mathsf{Cat}$ be a functor, with $\mathsf{A}$ a small category, and suppose that for every object $a$ of $\mathsf{A}$, the category $i(a)$ has a terminal object $e_a$. We are going to define a canonical natural transformation

$$\alpha \colon I_\mathsf{A} I^* \Longrightarrow \mathrm{id}_{\mathsf{Cat}}.$$

Let $\mathsf{C}$ be a small category. Spelling out the definitions, we see that the category $I_\mathsf{A} I^*(\mathsf{C})$ has for object pairs $(a, p \colon i(a) \to \mathsf{C})$ where $a$ is an object of $\mathsf{A}$ and $p$ a morphism of $\mathsf{Cat}$. A morphism $(a, p) \to (a', p')$ in $I_\mathsf{A} I^*(\mathsf{C})$ consists of a pair $(f, \sigma)$, where $f \colon a \to a'$ is a morphism of $\mathsf{A}$ and

$$\sigma \colon p \overset{\cong}{\Longrightarrow} p' \circ i(f)$$

is a natural isomorphism.

At the level of objects, we define the morphism $\alpha_\mathsf{C} \colon I_\mathsf{A} I^*(\mathsf{C}) \to \mathsf{C}$ with the formula

$$\alpha_\mathsf{C}(a, p) := p(e_a).$$

At the level of morphisms, the image of a morphism $(f, \sigma) \colon (a, p) \to (a', p')$ is defined as the composite

$$\alpha_\mathsf{C}(f, \sigma) := p(e_a) \xrightarrow{\sigma_{e_a}} p'(i(f)(e_a)) \to p'(e_{a'}),$$

where the morphism on the right is induced by the canonical morphism $i(f)(e_a) \to e_{a'}$.

We leave it to the reader to check that $\alpha_\mathsf{C} \colon I_\mathsf{A} I^*(\mathsf{C}) \to \mathsf{C}$ does indeed define a functor and that it is natural in $\mathsf{C}$.

**4.9  Remark**  When $i \colon \mathsf{A} \to \mathsf{Cat}$ is the functor $a \mapsto \mathsf{A}/a$, then $\alpha \colon I_\mathsf{A} I_\mathsf{A}^* \Longrightarrow \mathrm{id}_{\mathsf{Cat}}$ is nothing but the counit of the adjunction $I_\mathsf{A} \dashv I_\mathsf{A}^*$.

**4.10  Lemma**  *Let $i \colon A \to \mathsf{Cat}$ be a functor where $\mathsf{A}$ is a small category and suppose that for every object $a$ of $\mathsf{A}$, the category $i(a)$ has a terminal object. For every small category $\mathsf{C}$ and for an object $c$ of $\mathsf{C}$, we have a canonical isomorphism*

$$I_\mathsf{A} I^*(\mathsf{C})/c \simeq I_\mathsf{A} I^*(\mathsf{C}/c).$$

**Proof**  Let us denote by $e_a$ the terminal object of $i(a)$. The category $I_\mathsf{A} I^*(\mathsf{C})/c$ is described as follows:

- An object is a triple $(a, \ p \colon i(a) \to \mathsf{C}, \ g \colon p(e_a) \to c)$, where $a$ is an object of $\mathsf{A}$, $p$ is a morphism of $\mathsf{Cat}$, and $g$ is a morphism of $\mathsf{C}$.

- A morphism $(a, p, g) \to (a', p', g')$ is a couple $(f, \sigma)$, where $f : a \to a'$ is a morphism of A and

$$
\begin{array}{ccc}
i(a) & \xrightarrow{\ i(f)\ } & i(a') \\
& \llap{$p$}\searrow \ {}^\sigma\!\!\nearrow{\simeq} \ \downarrow{p'} & \\
& & \mathsf{C}
\end{array}
$$

is a natural isomorphism, such that the triangle

$$
\begin{array}{ccc}
p(e_a) & \xrightarrow{\ \alpha_\mathsf{C}(f,\sigma)\ } & p'(e'_a) \\
& \llap{$g$}\searrow \quad \downarrow{g'} & \\
& & c
\end{array}
$$

is commutative.

The category $I_\mathsf{A} I^*(\mathsf{C}/c)$ is described as:

- An object is a couple $(a, q : i(a) \to \mathsf{C}/c)$, where $a$ is an object of A and $q$ is a morphism of Cat.
- A morphism $(a, q) \to (a', q')$ is a pair $(f, \sigma)$, where $f : a \to a'$ is a morphism of A and

$$
\begin{array}{ccc}
i(a) & \xrightarrow{\ i(f)\ } & i(a') \\
& \llap{$q$}\searrow \ {}^\sigma\!\!\nearrow{\simeq} \ \downarrow{q'} & \\
& & \mathsf{C}/c
\end{array}
$$

is a natural isomorphism.

Let us write $\pi_c : \mathsf{C}/c \to \mathsf{C}$ for the canonical projection functor. We define a morphism of Cat as

$$
\theta : I_\mathsf{A} I^*(\mathsf{C}/c) \to I_\mathsf{A} I^*(\mathsf{C})/c, \quad (a, q : i(a) \to \mathsf{C}/c) \mapsto \big(a, \ \pi_c \circ q : i(a) \to \mathsf{C}, \ q(e_a) : \pi_c(q(e_a)) \to c\big),
$$

the definition on morphisms being the obvious one.

We now leave it to the reader to verify that this morphism is indeed an isomorphism. This mainly amounts to showing the following general fact: let A and B be two categories and suppose that A has a terminal object $t_\mathsf{A}$. For any object $b$ of B, a functor $f : \mathsf{A} \to \mathsf{B}/b$ is entirely determined by the postcomposition $\mathsf{A} \xrightarrow{f} \mathsf{B}/b \to \mathsf{B}$ and by $f(t_\mathsf{A})$, seen as a morphism of B whose target is $b$. (This follows essentially from the universal property of $\mathsf{B}/b$ as a comma-category.) $\qquad\square$

**4.11 Proposition** *Let $i : \mathsf{A} \to \mathsf{Cat}$ be a functor such that for every $a$ in A the category $i(a)$ has a terminal object. The following conditions are equivalent:*

(a) *$i$ is groupoidal aspherical.*

(b) *$I^*$ preserves aspherical objects: for every small aspherical category C, the Grpd-valued presheaf $I^*(\mathsf{C})$ is aspherical.*

(c) *$I^*$ preserves and reflects aspherical objects: a small category C is aspherical if and only if the Grpd-valued presheaf $I^*(\mathsf{C})$ is aspherical.*

(d)  *For every small category* $\mathsf{C}$, *the canonical morphism*

$$\alpha_{\mathsf{C}} \colon I_{\mathsf{A}} I^*(\mathsf{C}) \to \mathsf{C}$$

  *is a weak equivalence.*

(e)  $\mathsf{A}$ *is aspherical and* $I^*$ *preserves and reflects weak equivalences*:

$$(I^*)^{-1}(\mathcal{W}_{\widehat{\mathsf{A}}_{\mathsf{Grpd}}}) = \mathcal{W}.$$

(f)  $\mathsf{A}$ *is aspherical and* $I^*$ *preserves weak equivalences*:

$$I^*(\mathcal{W}) \subseteq \mathcal{W}_{\widehat{\mathsf{A}}_{\mathsf{Grpd}}}.$$

**Proof**  The implications (d) $\implies$ (c) $\implies$ (b) $\implies$ (a) are trivial (the last one comes from the fact that every category with a terminal object is aspherical). For the implication (a) $\implies$ (d), it follows from Lemma 4.10 (and the fact the slices $\mathsf{C}/c$ have a terminal object) that $\alpha_{\mathsf{C}}$ is aspherical. Hence, $\alpha_{\mathsf{C}}$ is a weak equivalence. This proves the equivalence of the first four conditions. For the implication (d) $\implies$ (e), notice first that if $\mathsf{C} = e$ is the terminal category, we have $I_{\mathsf{A}} I^*(e) \simeq \mathsf{A}$, and hence (c) implies that $\mathsf{A}$ is aspherical. The fact that (d) implies that $I^*$ preserves and reflects weak equivalences follows from the naturality of $\alpha$, 2-out-of-3 and the fact that $\mathcal{W}_{\widehat{\mathsf{A}}_{\mathsf{Grpd}}} = I_{\mathsf{A}}^{-1}(\mathcal{W})$ by definition. The implication (e) $\implies$ (f) is tautological. Finally, for the implication (f) $\implies$ (b), let $\mathsf{C}$ be a small aspherical category and consider the canonical morphism $\mathsf{C} \to e$, which is, by definition, a weak equivalence. Since $I^*$ preserves weak equivalences by hypothesis, the induced morphism

$$I^*(\mathsf{C}) \to I^*(e)$$

is a weak equivalence of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$. But $I_{\mathsf{A}} I^*(e) \simeq \mathsf{A}$, which is by hypothesis aspherical. Hence, $I^*(e)$ is aspherical and then so is $I^*(\mathsf{C})$. $\qquad\square$

We then obtain the following characterization of groupoidal weak test categories.

**4.12 Proposition**  *Let* $\mathsf{A}$ *be a small category. The following conditions are equivalent*:

(a)  $\mathsf{A}$ *is groupoidal weak test.*

(b)  *For every small category* $\mathsf{C}$ *with a terminal object*, *the* $\mathsf{Grpd}$-*valued presheaf* $I_{\mathsf{A}}^*(\mathsf{C})$ *is aspherical.*

**Proof**  The second condition means exactly that the functor $\mathsf{A} \to \mathsf{Cat}$, $a \mapsto \mathsf{A}/a$ is groupoidal aspherical (since the slice categories $\mathsf{A}/a$ have a terminal object). The equivalence follows then from condition (d) of Proposition 4.11 combined with Remark 4.9 and Lemma 1.3. $\qquad\square$

Interestingly, we also obtain the following result.

**4.13 Corollary**  *Let* $\mathsf{A}$ *be a groupoidal weak test category and* $i \colon \mathsf{A} \to \mathsf{Cat}$ *a groupoidal aspherical functor. Then*, $I^* \colon \mathsf{Cat} \to \widehat{\mathsf{A}}_{\mathsf{Grpd}}$ *is a homotopy inverse of* $I_{\mathsf{A}} \colon \widehat{\mathsf{A}}_{\mathsf{Grpd}} \to \mathsf{Cat}$.

**Proof** We already know that if $i$ is a groupoidal aspherical functor, we have a natural transformation $\alpha \colon I_A I^* \Longrightarrow \mathrm{id}_{\mathsf{Cat}}$ which is a weak equivalence argument by argument. Consider now the zigzag of natural transformations

$$I^* I_A \Longrightarrow I_A^* I_A I^* I_A \Longrightarrow I_A^* I_A \Longleftarrow \mathrm{id},$$

where the arrow on the left is induced by the unit of the adjunction $I_A \dashv I_A^*$, the middle arrow is obtained by postcomposing $I_A^*$ to $\alpha$ and precomposing with $I_A$ and the arrow on the right is again the unit of the adjunction $I_A \dashv I_A^*$. Since A is groupoidal weak test, the unit of this adjunction is a weak equivalence argument by argument and $I_A^*$ preserves weak equivalences. This proves that the three natural transformations of the previous zigzag are weak equivalences argument by argument. $\square$

**4.14 Remark** Following the terminology from the usual test category theory, a groupoidal aspherical functor whose source is a weak test category ought to be called a *groupoidal weak test functor*.

**4.15 Remark** Even if we remove the hypothesis that A is groupoidal weak test from the previous corollary, it follows trivially from Proposition 4.11(d) that $I^*$ is a homotopical inverse "on one side" of $I_A$, but it does not seem to be a homotopical inverse "on both sides" in general. No counterexample, however, is known by the author.

# 5 Groupoidal test categories

We fix once and for all in this section a basic localizer $\mathcal{W}$ of Cat.

**5.1 Definition** A small category A is *$\mathcal{W}$-groupoidal local test*, or simply *groupoidal local test*, if for every object $a$ of A, the category $A/a$ is groupoidal weak test. We say that A is *$\mathcal{W}$-groupoidal test*, or simply *groupoidal test*, if it is both a groupoidal weak test category and a groupoidal local test category.

We are now going to look for characterizations of groupoidal local test categories and groupoidal test categories. For that, it is useful to introduce first a variation of the notion of aspherical object of $\widehat{A}_{\mathsf{Grpd}}$.

**5.2 Notation** Let A be a small category and $a$ an object of A. Given a Grpd-valued presheaf $X$ over $A$, we denote by $X|_{A/a}$ the Grpd-valued presheaf over $A/a$ defined as the composition

$$(A/a)^{\mathrm{op}} \longrightarrow A^{\mathrm{op}} \xrightarrow{\ X\ } \mathsf{Grpd},$$

where $A/a \to A$ is the canonical projection.

**5.3 Definition** Let A be a small category. An object $X$ of $\widehat{A}_{\mathsf{Grpd}}$ is *$\mathcal{W}$-locally aspherical*, or simply *locally aspherical*, if for every object $a$ of A, the object $X|_{A/a}$ of $\widehat{(A/a)}_{\mathsf{Grpd}}$ is aspherical.

We now have the following reformulation.

**5.4 Proposition** *Let* A *be a small category. The following conditions are equivalent*:

(a)  A *is groupoidal local test.*

(b)  *For every small category* C *with a terminal object,* $I_A^*(C)$ *is locally aspherical.*

**Proof**  This follows immediately from Proposition 4.12 and the fact that for every small category C and every object $a$ of A, we have
$$I_{A/a}^*(C) \simeq I_A^*(C)|_{A/a}. \qquad \square$$

More generally, we can consider the following variation of groupoidal aspherical functors.

**5.5 Definition**  Let A be a small category. A functor $i : A \to$ Cat is $\mathcal{W}$-*groupoidal locally aspherical*, or simply *groupoidal locally aspherical*, if the following conditions are satisfied:

(a)  For every object $a$ of A, $i(a)$ has a terminal object.

(b)  For every small category C with a terminal object, $I^*(C)$ is locally aspherical.

**5.6 Remark**  In other words, Proposition 5.4 says that A is groupoidal local test if and only if A $\to$ Cat, $a \mapsto$ A$/a$ is a groupoidal locally aspherical functor.

We now turn to a couple of key technical results on locally aspherical Grpd-valued presheaves.

**5.7 Lemma**  *Let* A *be a small category,* $a$ *an object of* A *and* $X$ *an object of* $\widehat{A}_{\mathsf{Grpd}}$. *We have a sequence*
$$I_A(a \times X) \simeq I_{A/a}(X|_{A/a}) \simeq I_A(X)/a$$
*of isomorphisms natural in* $X$, *where*:

- *On the left-hand side, we abusively wrote* $a$ *for the* Set-*valued presheaf represented by* $a$.
- *On the right-hand side, we implicitly used the canonical morphism* $\zeta_X : I_A(X) \to$ A *to make sense of the slice category.*

**Proof**  These three categories have the same description (up to canonical isomorphism):

- An object is a triple $(a', p : a' \to a, x)$, where $a'$ is an object of A, $p$ is a morphism of A, and $x$ is an object of $X(a')$.
- A morphism $(a', p, x) \to (a'', p', x')$ is a pair $(f, k)$ where $f : a' \to a''$ is a morphism of A such that $p' \circ f = p$ and $k : x \to X(f)(x')$ is a morphism of $X(a')$. $\qquad \square$

**5.8 Lemma**  *Let* A *be a small category and* $X$ *an object of* $\widehat{A}_{\mathsf{Grpd}}$. *The following conditions are equivalent*:

(a)  $X$ *is locally aspherical.*

(b)  *The image by* $I_A$ *of the morphism to the terminal object* $X \to *$,
$$I_A(X) \to \mathsf{A},$$
*is an aspherical morphism of* Cat.

(c)  *For every object $a$ of $\mathsf{A}$, seen as a representable presheaf (and as an object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, via the canonical inclusion $\widehat{\mathsf{A}} \hookrightarrow \widehat{\mathsf{A}}_{\mathsf{Grpd}}$), the product in $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$*

$$a \times X$$

*is aspherical.*

(d)  *The morphism $X \to *$ is a universal weak equivalence, ie for every object $Y$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, the canonical projection*

$$Y \times X \to Y$$

*is a weak equivalence.*

**Proof**  The equivalence of (a), (b) and (c) follows immediately from Lemma 5.7. For the implication (b) $\implies$ (d), consider the cartesian square

$$
\begin{array}{ccc}
Y \times X & \longrightarrow & X \\
\downarrow & \lrcorner & \downarrow \\
Y & \longrightarrow & *
\end{array}
$$

We leave it to the reader to check that the functor $I_{\mathsf{A}}$ preserves cartesian squares. Hence, we obtain a cartesian square

$$
\begin{array}{ccc}
I_{\mathsf{A}}(Y \times X) & \longrightarrow & I_{\mathsf{A}}(X) \\
\downarrow & \lrcorner & \downarrow {\scriptstyle \zeta_X} \\
I_{\mathsf{A}}(Y) & \xrightarrow{\zeta_Y} & \mathsf{A}
\end{array}
$$

Since $I_{\mathsf{A}}(Y) \to \mathsf{A}$ is a Grothendieck fibration, we deduce from Lemma 2.6 that the left vertical morphism of the above square is an aspherical morphism of $\mathsf{Cat}$, which implies that $Y \times X \to X$ is a weak equivalence. Finally, for the implication (d) $\implies$ (c), let $a$ be an object of $\mathsf{A}$, and consider the canonical projection

$$a \times X \to a,$$

where once again we wrote $a$ for the Set-valued presheaf represented by $a$. Since the presheaf $a$ is aspherical (because the category $I_{\mathsf{A}}(a) = \mathsf{A}/a$ has a terminal object), it follows that $a \times X$ is aspherical. $\square$

**5.9 Remark**  Lemma 5.8 hides a subtle distinction between the classical theory for Set-valued presheaves and the theory for Grpd-valued presheaves. Indeed, let us call *local weak equivalence* a morphism $f : X \to Y$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ such that for every object $a$ of $\mathsf{A}$, the morphism of $\widehat{(\mathsf{A}/a)}_{\mathsf{Grpd}}$

$$f|_{\mathsf{A}/a} : X|_{\mathsf{A}/a} \to Y|_{\mathsf{A}/a}$$

is a weak equivalence. Then, an object $X$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ is locally aspherical if and only if $X \to *$ is a local weak equivalence. Now, Lemma 5.8(d) tells us that $X$ is locally aspherical if and only if $X \to *$ is a universal weak equivalence, and we might think that this characterization is true for all local weak equivalences (as the analogue result is true for Set-valued presheaves [17, paragraphe 1.2.5]). This does not work though. The correct generalization, which goes beyond the scope of this paper, involves what ought to be called "comma-universal weak equivalence", whose definition is the same as universal equivalence except that pullback squares are replaced by comma squares.

**5.10 Lemma** *Let* A *be a small category and* $X$ *an object of* $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$. *If* A *is aspherical, then we have the implication*

$$X \text{ locally aspherical} \implies X \text{ aspherical}.$$

**Proof** Thanks to Lemma 5.8 (for example item (b)), we know that if an object $X$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ is locally aspherical, then $X \to *$ is a weak equivalence. The conclusion follows from Lemma 3.8. $\qquad\square$

From this last lemma, we immediately deduce the following two propositions.

**5.11 Proposition** *Let* A *be a small category and* $i : \mathsf{A} \to \mathsf{Cat}$ *a groupoidal locally aspherical functor. Then* $i$ *is a groupoidal aspherical functor if and only if* A *is aspherical.*

**5.12 Proposition** *Let* A *be a groupoidal local test category. Then* A *is groupoidal test if and only if it is aspherical.*

This last result means that in order to characterize groupoidal test categories, it suffices to characterize groupoidal local test categories.

**5.13** Let M be a category with finite products (this includes a terminal object, which we denote by $e_{\mathsf{M}}$). An *interval* in M is a triple $(\mathbb{I}, i_0, i_1)$, where $\mathbb{I}$ is an object of M and $i_0$ and $i_1$ are morphisms of M:

$$i_0, i_1 : e_{\mathsf{M}} \to \mathbb{I}.$$

A *morphism of intervals* $(\mathbb{I}, i_0, i_1) \to (\mathbb{I}', i_0', i_1')$ consists of a morphism $\varphi : \mathbb{I} \to \mathbb{I}'$ of M such that $i_\varepsilon' = \varphi \circ i_\varepsilon$ for $\varepsilon = 0, 1$.

Now, let $f, g : X \to Y$ be two parallel morphisms of M and $(\mathbb{I}, i_0, i_1)$ an interval in M. An $\mathbb{I}$-*homotopy from $f$ to $g$* is a morphism $h : \mathbb{I} \times X \to Y$ of M such that the following diagram is commutative:



We consider the smallest equivalence relation on the set $\mathrm{Hom}_{\mathsf{M}}(X, Y)$ such that $f$ is equivalent to $g$ if there exists a $\mathbb{I}$-homotopy from $f$ to $g$. If two morphisms $X \to Y$ are in the same equivalence class for this relation, we say that they are $\mathbb{I}$-*homotopic*.

We say that an object $X$ of M is $\mathbb{I}$-*contractible* if $\mathrm{id}_X : X \to X$ is $\mathbb{I}$-homotopic to a constant morphism (ie a morphism which factorizes through the terminal object $e_{\mathsf{M}}$).

Finally, notice that if $F : \mathsf{M} \to \mathsf{M}'$ is a functor preserving finite products, then for any interval $(\mathbb{I}, i_0, i_1)$ of M, $(F(\mathbb{I}), F(i_0), F(i_1))$ is an interval of M', and $F$ sends $\mathbb{I}$-homotopic morphisms to $F(\mathbb{I})$-homotopic morphisms. In particular, $F$ sends $\mathbb{I}$-contractible objects to $F(\mathbb{I})$-contractible objects.

**5.14 Example** Let $\Delta_1$ be the poset $\{0 < 1\}$ seen as an object of Cat, and denote by $e_0, e_1 \colon e \to \Delta_1$ the canonical inclusions of 0 and 1 respectively. Then $(\Delta_1, e_0, e_1)$ is an interval of Cat. A $\Delta_1$-homotopy from a morphism $u \colon A \to B$ to a morphism $v \colon A \to B$ is simply a natural transformation $u \Longrightarrow v$. Notice that a small category with either a terminal object or an initial object is $\Delta_1$-contractible.

**5.15 Example** Let $i \colon \mathsf{A} \to \mathsf{Cat}$ be a functor where $\mathsf{A}$ is a small category. Since $I^*$ preserves limits, $(I^*(\Delta_1), I^*(e_0), I^*(e_1))$ is an interval of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$.

The following lemma relates the notion of $\mathbb{I}$-homotopy with the homotopy theory induced by a class of weak equivalences in the ambient category.

**5.16 Lemma** *Let* $\mathsf{M}$ *be a category with finite products*, $\mathsf{W}$ *a weakly saturated class of morphisms of* $\mathsf{M}$ *and* $\mathbb{I}$ *an interval of* $\mathsf{M}$ *such that the canonical morphism to the terminal object* $\mathbb{I} \to e_{\mathsf{M}}$ *is universally in* $\mathsf{W}$. *Then, for every* $\mathbb{I}$-contractible object $X$ of $\mathsf{M}$, *the canonical morphism* $X \to e_{\mathsf{M}}$ *is universally in* $\mathsf{W}$.

**Proof** This is a reformulation of [17, lemme 1.4.6]. □

We can now apply this to groupoidal locally aspherical functors.

**5.17 Proposition** *Let* $i \colon \mathsf{A} \to \mathsf{Cat}$ *be a functor, with* $\mathsf{A}$ *a small category, such that for every object* $a$ *of* $\mathsf{A}$, *the category* $i(a)$ *has a terminal object. The following conditions are equivalent*:

(a) *$i$ is groupoidal locally aspherical.*

(b) *$I^*(\Delta_1)$ is locally aspherical.*

**Proof** The implication (a) $\Longrightarrow$ (b) is trivial because $\Delta_1$ has a terminal object. For the converse, we need to show that for every small category $\mathsf{C}$ with a terminal object, $I^*(\mathsf{C})$ is locally aspherical. By Lemma 5.8(d), this is equivalent to showing that $I^*(\mathsf{C}) \to *$ is a universal weak equivalence. Notice that $I^*$ preserves limits and so it sends $\Delta_1$-contractible objects of Cat to $I^*(\Delta_1)$-contractible objects of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$. Since every small category with a terminal object is $\Delta_1$-contractible, the result follows from Lemma 5.16. □

We could now apply the previous proposition to the functor $\mathsf{A} \to \mathsf{Cat}$, $a \mapsto \mathsf{A}/a$ and obtain a characterization of groupoidal local test categories. As it happens, we will soon obtain an even finer characterization, but we first need some more results on intervals.

**5.18 Definition** Let $\mathsf{M}$ be a category with finite products, and whose terminal object is denoted by $e_{\mathsf{M}}$. A *multiplicative interval* in $\mathsf{M}$ is an interval $(\mathbb{L}, \lambda_0, \lambda_1)$ together with a binary operation

$$\Lambda \colon \mathbb{L} \times \mathbb{L} \to \mathbb{L},$$

such that $\lambda_0$ is a unit on the left and $\lambda_1$ is absorbing on the left. To wit, these diagrams are commutative:

**5.19 Example** The interval $(\Delta_1, e_0, e_1)$ of Cat is multiplicative when equipped with the binary operation

$$\Delta_1 \times \Delta_1 \to \Delta_1, \quad (a, b) \mapsto a + b - ab.$$

Since $I_A^*$ preserves limits, it follows that the interval $(I_A^*(\Delta_1), I_A^*(e_0), I_A^*(e_1))$, equipped with the image by $I_A^*$ of the above binary operation, is multiplicative.

**5.20 Lemma** *Let* M *be a category with finite products,* $\mathcal{W}$ *a weakly saturated class of maps of* M, $(\mathbb{I}, i_0, i_1)$ *an interval in* M *such that* $\mathbb{I} \to e_M$ *is universally in* $\mathcal{W}$ *and* $(\mathbb{L}, \lambda_0, \lambda_1, \Lambda)$ *a multiplicative interval in* M. *If there exists a morphism of intervals* $(\mathbb{I}, i_0, i_1) \to (\mathbb{L}, \lambda_0, \lambda_1)$, *then* $\mathbb{L} \to e_M$ *is universally in* $\mathcal{W}$.

**Proof** See [17, lemme 1.4.10]. □

For the next definition, recall that a $(2, 1)$-category is a 2-category such that every 2-morphism is invertible (in other words, a Grpd-enriched category). Limits and colimits in a $(2, 1)$-category are the Grpd-enriched ones.

**5.21 Definition** Let $\underline{M}$ be a $(2, 1)$-category with finite products (the terminal object is denoted by $e_M$) and an initial object $\varnothing$. An interval $(\mathbb{I}, i_0, i_1)$ (of the underlying category of) $\underline{M}$ is said to be *strongly separating* if for every 2-square of $\underline{M}$



we necessarily have $X = \varnothing$ (and thus the 2-morphism is the identity as colimits are understood to be the enriched ones).

**5.22 Example** Consider Cat as a $(2, 1)$-category, where the 2-morphisms are the natural isomorphisms between functors. Then $(\Delta_1, e_0, e_1)$ is strongly separating.

**5.23** Let A be a small category. The category $\widehat{A}_{Grpd}$ has a canonical structure of a 2-category where the 2-morphisms are the strict natural 2-transformations. That is, given two parallel morphisms $\varphi, \psi : X \to Y$ of $\widehat{A}_{Grpd}$, a 2-morphism $\alpha : \varphi \Longrightarrow \psi$ consists of a family of natural transformations



such that, for every $f : a \to a'$ in A, the following naturality condition is satisfied:



Notice that since $X$ and $Y$ take values in groupoids, every $\alpha_a$ is invertible, and it follows that every 2-morphism of $\widehat{A}_{Grpd}$ is also invertible. Hence, $\widehat{A}_{Grpd}$ is a $(2, 1)$-category.

**5.24 Lemma** *Let $i\colon \mathsf{A} \to \mathsf{Cat}$ be a functor, with $\mathsf{A}$ a small category, such that for every object $a$ of $\mathsf{A}$, the category $i(a)$ is not empty. Then, the interval $(I^*(\Delta_1), I^*(e_0), I^*(e_1))$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ is strongly separating.*

**Proof** It is obvious that $I^*\colon \mathsf{Cat} \to \widehat{\mathsf{A}}_{\mathsf{Grpd}}$ can be extended to a $(2,1)$-functor, and so is its left adjoint $I_!$ which was defined in the proof of Lemma 4.3. We obtain this way a $(2,1)$-adjunction. The fact that $(\Delta_1, e_0, e_1)$ is strongly separating in $\mathsf{Cat}$ can be expressed as the fact that the commutative square

$$
\begin{array}{ccc}
\varnothing & \longrightarrow & e \\
\downarrow & & \downarrow{\scriptstyle e_0} \\
e & \xrightarrow[e_1]{} & \Delta_1
\end{array}
$$

is a $(2,1)$-comma square.[9] Since $I^*$ preserves $\mathsf{Grpd}$-enriched limits, the square

$$
\begin{array}{ccc}
\varnothing & \longrightarrow & * \\
\downarrow & & \downarrow{\scriptstyle I^*(e_0)} \\
* & \xrightarrow[I^*(e_1)]{} & I^*(\Delta_1)
\end{array}
$$

is also a $(2,1)$-comma square (we used that $I^*(\varnothing) \simeq \varnothing$, which follows from the hypothesis on the nonemptiness of the categories $i(a)$), which means exactly that the interval $(I^*(\Delta_1), I^*(e_0), I^*(e_1))$ is strongly separating. Details are left to the reader. $\qquad\square$

**5.25 Lemma** *Let $\mathsf{A}$ be a small category. For every strongly separating interval $(\mathbb{I}, i_0, i_1)$ of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, there exists a morphism of intervals $(\mathbb{I}, i_0, i_1) \to (I^*_{\mathsf{A}}(\Delta_1), I^*_{\mathsf{A}}(e_0), I^*_{\mathsf{A}}(e_1))$ (not necessarily unique).*

**Proof** By adjunction, we need to find a morphism of $\mathsf{Cat}$, $u\colon I_{\mathsf{A}}(\mathbb{I}) \to \Delta_1$, such that the following diagram is commutative:

(1)
$$
\begin{array}{ccc}
\mathsf{A} & \longrightarrow & e \\
\downarrow & & \downarrow{\scriptstyle e_0} \\
I_{\mathsf{A}}(\mathbb{I}) & \xrightarrow{\ u\ } & \Delta_1 \\
\uparrow & & \uparrow{\scriptstyle e_1} \\
\mathsf{A} & \longrightarrow & e
\end{array}
$$

Here, the map $\mathsf{A} \to I_{\mathsf{A}}(\mathbb{I})$ at the top is defined as $a \mapsto \big(a, a \longrightarrow * \xrightarrow{i_0} \mathbb{I}\big)$, and the other one similarly with $i_1$ instead of $i_0$. Let $\big(a, a \xrightarrow{\ p\ } \mathbb{I}\big)$ be an object of $I_{\mathsf{A}}(\mathbb{I})$.

- If $p$ is such that there exists a natural isomorphism

$$
\begin{array}{ccc}
a & \xrightarrow{\ \simeq\ } & * \\
& {\scriptstyle p}\searrow & \downarrow{\scriptstyle i_0} \\
& & \mathbb{I}
\end{array}
$$

  then we define $u(a, p) = 0$.

- Else we define $u(a, p) = 1$.

---

[9]The definition is the same as the usual notion of comma square in a 2-category, except that every 2-morphism involved is invertible.

Given a morphism $(a', p') \to (a, p)$ of $I_A(\mathbb{I})$, notice that if $u(a, p) = 0$, then $u(a', p') = 0$ too. (In other words, the objects sent to 0 form a *sieve*.) This allows for a unique possible way of defining $u$ on arrows. The upper square of (1) is commutative by definition. For the lower square, we need to prove that

$$u\big(a, a \longrightarrow * \xrightarrow{\; i_1 \;} \mathbb{I}\big) = 1.$$

Suppose that it is not the case: this would mean that there exists a 2-square

$$
\begin{array}{ccc}
a & \longrightarrow & * \\
\downarrow & \overset{\simeq}{\nearrow\!\!\!\!\nearrow} & \downarrow {\scriptstyle i_0} \\
* & \underset{i_1}{\longrightarrow} & \mathbb{I}
\end{array}
$$

which is forbidden since $(\mathbb{I}, i_0, i_1)$ is strongly separating and the initial presheaf $\varnothing$ is never representable. $\square$

**5.26 Proposition** *Let* A *be a small category. The following are equivalent*:

(a)  A *is groupoidal local test.*

(b)  $I_A^*(\Delta_1)$ *is locally aspherical.*

(c)  *There exists a strongly separating interval* $(\mathbb{I}, i_0, i_1)$ *in* $\widehat{A}_{\mathsf{Grpd}}$, *such that* $\mathbb{I}$ *is locally aspherical.*

(d)  *There exists a groupoidal locally aspherical functor* $i : A \to \mathsf{Cat}$.

**Proof**  By Remark 5.6, A is groupoidal local test if and only if $A \to \mathsf{Cat}$, $a \mapsto A/a$ is a groupoidal locally aspherical functor. Hence, the implication (a) $\Longrightarrow$ (d) is trivial and the equivalence (a) $\Longleftrightarrow$ (b) follows from Proposition 5.17. From Lemma 5.24, we know that $(I_A^*(\Delta_1), I_A^*(e_0), I_A^*(e_1))$ is a strongly separating interval; hence the implication (b) $\Longrightarrow$ (c). For the implication (c) $\Longrightarrow$ (b), we know from Example 5.19 that $(I_A^*(\Delta_1), I_A^*(e_0), I_A^*(e_1))$ is a multiplicative interval. Then, Lemma 5.25 implies that there exists a morphism of intervals $(\mathbb{I}, i_0, i_1) \to (I_A^*(\Delta_1), I_A^*(e_0), I_A^*(e_1))$. Since, by hypothesis, $\mathbb{I}$ is locally aspherical, it follows from Lemma 5.20 that $I_A^*(\Delta_1)$ is locally aspherical.

So far, we have shown (c) $\Longleftrightarrow$ (b) $\Longleftrightarrow$ (a) $\Longrightarrow$ (d). Let us conclude with the implication (d) $\Longrightarrow$ (c). If $i : A \to \mathsf{Cat}$ is a groupoidal locally aspherical functor, then by definition $I^*(\Delta_1)$ is locally aspherical. Besides, each category $i(a)$ has a terminal object and in particular is not empty; hence Lemma 5.24 applies and $(I^*(\Delta_1), I^*(e_0), I^*(e_1))$ is a strongly separating interval. $\square$

# 6  Groupoidal strict test categories

We fix once and for all in this section a basic localizer $\mathcal{W}$ of $\mathsf{Cat}$.

**6.1**  Recall that a small category A is *totally aspherical* if:

(i)  A is aspherical.

(ii)  The diagonal functor
$$\delta : A \to A \times A$$
is aspherical.

**6.2 Example** A small category that has finite products (including the empty product) is totally aspherical [17, exemple 1.6.4].

**6.3 Definition** A small category A is $\mathcal{W}$-*groupoidal strict test*, or simply *groupoidal strict test*, if the following conditions are satisfied:

(a) A is totally aspherical.

(b) A is groupoidal test.

In the following proposition, it is important to understand that "finite" includes "empty".

**6.4 Proposition** *Let* A *be a small category. The following are equivalent*:

(a) A *is totally aspherical.*

(b) *The functor* $I_A\colon \widehat{A}_{\mathsf{Grpd}} \to \mathsf{Cat}$ *preserves finite products up to weak equivalence: for every finite family* $(X_i)_{i\in I}$ *of objects of* $\widehat{A}_{\mathsf{Grpd}}$, *the canonical morphism*
$$I_A\Big(\prod_{i\in I} X_i\Big) \to \prod_{i\in I} I_A(X_i)$$
*is a weak equivalence.*

(c) *The class of aspherical objects of* $\widehat{A}_{\mathsf{Grpd}}$ *is stable by finite products: if* $(X_i)_{i\in I}$ *is a finite family of aspherical objects of* $\widehat{A}_{\mathsf{Grpd}}$, *then*
$$\prod_{i\in I} X_i$$
*is also aspherical.*

(d) *For every finite family of* $(a_i)_{i\in I}$ *of objects of* A, *seen as representable presheaves* (*and as objects of* $\widehat{A}_{\mathsf{Grpd}}$ *via the canonical inclusion* $\widehat{A} \hookrightarrow \widehat{A}_{\mathsf{Grpd}}$), *the product in* $\widehat{A}_{\mathsf{Grpd}}$
$$\prod_{i\in I} a_i$$
*is aspherical.*

**Proof** Let us begin with (a) $\Longrightarrow$ (b). For the empty product, this is simply saying that $I_A(*) \simeq A \to e$ is a weak equivalence, which is the case because a totally aspherical category is in particular aspherical. Now let $X$ and $Y$ be two objects of $\widehat{A}_{\mathsf{Grpd}}$, and notice that the square
$$
\begin{array}{ccc}
I_A(X\times Y) & \longrightarrow & I_A(X)\times I_A(Y) \\
\zeta_{X\times Y}\downarrow & & \downarrow \zeta_X\times\zeta_Y \\
A & \xrightarrow{\ \delta\ } & A\times A
\end{array}
$$
is commutative and a pullback square. Since a product of Grothendieck fibrations is again a Grothendieck fibration, the right vertical arrow is a Grothendieck fibration and by Lemma 2.6 we deduce that the top horizontal arrow is aspherical. The general case follows from an immediate induction and the fact that weak equivalences in Cat are stable by finite products [17, proposition 2.1.3].

The implication (b) $\Longrightarrow$ (c) is immediate because a finite product of aspherical categories is aspherical. The implication (c) $\Longrightarrow$ (d) is trivial.

Finally, for the implication (d) $\Longrightarrow$ (a), notice first that condition (d) applied to the empty product gives that the terminal $*$ object of $\widehat{A}_{\mathsf{Grpd}}$ is aspherical, which means exactly that A is aspherical as usual. Now let $a$ and $b$ be the two objects of A, seen as representable presheaves and thus as objects of $\widehat{A}_{\mathsf{Grpd}}$. It is straightforward to check that

$$I_A(a \times b) \simeq A/(a, b),$$

where on the right-hand side, $(a, b)$ has to be understood as an object of $A \times A$ and the slice is relative the diagonal functor $\delta \colon A \to A \times A$. This slice category being aspherical for every $(a, b) \in A \times A$ means exactly that $\delta$ is aspherical. $\qquad\square$

Now, the crucial result is the following.

**6.5 Lemma**  *Let* A *be a totally aspherical category and* $X$ *an object of* $\widehat{A}_{\mathsf{Grpd}}$. *The following conditions are equivalent*:

(a)  $X$ *is aspherical.*

(b)  $X$ *is locally aspherical.*

**Proof**  A totally aspherical category being in particular aspherical, the implication (b) $\Longrightarrow$ (a) has already been proved in Lemma 5.10. For the other implication, let $a$ be an object of A, which we see as a representable presheaf (and then as an object of $\widehat{A}_{\mathsf{Grpd}}$ via the canonical inclusion $\widehat{A} \hookrightarrow \widehat{A}_{\mathsf{Grpd}}$). Thanks to Proposition 6.4 and because representable presheaves are always aspherical, we know that

$$a \times X$$

is an aspherical object of $\widehat{A}_{\mathsf{Grpd}}$, which proves that $X$ is locally aspherical by Lemma 5.8. $\qquad\square$

The following results are straightforward consequences of the previous lemma.

**6.6 Proposition**  *Let* A *be a small category and* $i \colon A \to \mathsf{Cat}$ *a functor. If* A *is totally aspherical, then the following conditions are equivalent*:

(a)  $i$ *is a groupoidal locally aspherical functor.*

(b)  $i$ *is a groupoidal aspherical functor.*

**6.7 Proposition**  *Let* A *be a small category. If* A *is totally aspherical, then the following are equivalent*:

(a)  A *is groupoidal strict test.*

(b)  A *is groupoidal test.*

(c)  A *is groupoidal weak test.*

(d)  $I_A(\Delta_1)$ *is aspherical.*

(e)  *There exists a strongly separating interval* $(\mathbb{L}, i_0, i_1)$ *in* $\widehat{A}_{\mathsf{Grpd}}$ *such that* $\mathbb{L}$ *is aspherical.*

(f)  *There exists a groupoidal aspherical functor* $i \colon A \to \mathsf{Cat}$.

# 7 Test categories vs groupoidal test categories

**7.1** The comparison of the theory of groupoidal test categories and test categories relies on the following trivial but essential observation. If D is a (small) category with no nontrivial isomorphisms, then for any (small) category C, the groupoid $\underline{\mathrm{Hom}}_{\mathsf{Cat}}^{\mathrm{iso}}(\mathsf{C}, \mathsf{D})$ doesn't have any nontrivial morphisms. In other words, $\underline{\mathrm{Hom}}_{\mathsf{Cat}}^{\mathrm{iso}}(\mathsf{C}, \mathsf{D})$ is a set and we have

$$\underline{\mathrm{Hom}}_{\mathsf{Cat}}^{\mathrm{iso}}(\mathsf{C}, \mathsf{D}) = \mathrm{Hom}_{\mathsf{Cat}}(\mathsf{C}, \mathsf{D}).$$

In particular, let $i : \mathsf{A} \to \mathsf{Cat}$ be a functor, with A a small category. For any category D with no nontrivial isomorphisms, we have

$$I^*(\mathsf{D}) = i^*(\mathsf{D}).$$

(Remember that we consider $\widehat{\mathsf{A}}$ as a full subcategory of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$.)

We then immediately have the following result.

**7.2 Proposition** *Let* $i : \mathsf{A} \to \mathsf{Cat}$ *a functor such that for every object $a$ of* A, *the category $i(a)$ has a terminal object. The following are equivalent*:

   (a)  *$i$ is a groupoidal locally aspherical functor.*

   (b)  *$i$ is a locally aspherical functor.*

**Proof** By Proposition 5.17, condition (a) is equivalent to $I^*(\Delta_1)$ being locally aspherical. Condition (b) means that $i^*(\Delta_1)$ is locally aspherical (see Definition 2.15). Since $\Delta_1$ has no nontrivial isomorphism,

$$I^*(\Delta_1) = i^*(\Delta_1).$$

To conclude, let us prove that a Set-valued presheaf is locally aspherical as an object of $\widehat{\mathsf{A}}$ (Definition 2.8) if and only if it is locally aspherical as an object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ (Definition 5.3). First notice that the canonical inclusion $\widehat{\mathsf{A}} \hookrightarrow \widehat{\mathsf{A}}_{\mathsf{Grpd}}$ preserves and reflects limits, and preserves and reflects weak equivalences. Hence, given an object $X$ of $\widehat{\mathsf{A}}$, if $X \to *$ is a universal weak equivalence of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, then it is also a universal weak equivalence of $\widehat{\mathsf{A}}$. The latter is the definition of locally aspherical object of $\widehat{\mathsf{A}}$ and the former is a characterization of locally aspherical object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ (Lemma 5.8(d)). This proves the "if" part. Conversely, if $X \to *$ is a universal weak equivalence of $\widehat{\mathsf{A}}$, then for every object $a$ of A, seen as a representable presheaf, the canonical projection $a \times X \to a$ is a weak equivalence of $\widehat{\mathsf{A}}$. Since $a$ is an aspherical object of $\widehat{\mathsf{A}}$ (because $i_{\mathsf{A}}(a) = \mathsf{A}/a$ has a terminal object), it follows that $a \times X$ is an aspherical object of $\widehat{\mathsf{A}}$. By Remark 3.7, we deduce that it is also an aspherical object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$, which proves that $X$ is locally aspherical as an object of $\widehat{\mathsf{A}}_{\mathsf{Grpd}}$ by Lemma 5.8(c). $\qquad\square$

From this, we deduce our first comparison theorem.

**7.3 Theorem** *Let* A *be a small category. We have the following equivalences*:

   (a)  A *is groupoidal local test* $\Longleftrightarrow$ A *is local test.*

   (b)  A *is groupoidal test* $\Longleftrightarrow$ A *is test.*

   (c)  A *is groupoidal strict test* $\Longleftrightarrow$ A *is strict test.*

**Proof** By Remark 5.6 (resp Remark 2.17), A is a groupoidal local test category (resp local test category) if and only if A → Cat, $a \mapsto A/a$ is a locally groupoidal aspherical functor (resp locally aspherical functor). The equivalence (a) follows then from Proposition 7.2.

By Proposition 5.12 (resp Proposition 2.14), we know that A is groupoidal test (resp test) if and only if it is groupoidal locally test (resp locally test) and aspherical. Hence, the equivalence (b) follows trivially from (a).

Finally, A is groupoidal strict test (resp strict test) if it is groupoidal test (resp test) and totally aspherical. Hence, the equivalence (c) follows trivially from (b). □

**7.4 Corollary** *If* A *is a test category (or equivalently a groupoidal test category), the canonical inclusion functor* $\widehat{A} \hookrightarrow \widehat{A}_{\mathsf{Grpd}}$ *induces an equivalence at the level of homotopy categories*:

$$\mathrm{Ho}(\widehat{A}) \simeq \mathrm{Ho}(\widehat{A}_{\mathsf{Grpd}}).$$

**Proof** Consider the commutative triangle

$$
\begin{array}{ccc}
\widehat{A} & \lhook\joinrel\longrightarrow & \widehat{A}_{\mathsf{Grpd}} \\
& {\scriptstyle i_A}\searrow & \downarrow {\scriptstyle I_A} \\
& & \mathsf{Cat}
\end{array}
$$

If A is a test category (or equivalently a groupoidal test category), then both vertical arrows of the previous triangle induce equivalences at the level of homotopy categories. The result follows then by a 2-out-of-3 property for equivalences of categories. □

**7.5 Remark** The proof of the previous corollary is straightforwardly generalized to deduce that if A is a test category, then $\widehat{A} \hookrightarrow \widehat{A}_{\mathsf{Grpd}}$ induces a Dwyer–Kan equivalence [2] $(\widehat{A}, \mathcal{W}_{\widehat{A}}) \xrightarrow{\sim} (\widehat{A}_{\mathsf{Grpd}}, \mathcal{W}_{\widehat{A}_{\mathsf{Grpd}}})$; hence it induces an equivalence of $(\infty, 1)$-categories.

**7.6 Example** It follows from Theorem 7.3 that all the examples of (strict) test categories given in Example 2.13 are also groupoidal (strict) test categories. In particular, $\Delta$ is a groupoidal strict test category, and we recover the classical result that the category $\widehat{\Delta}_{\mathsf{Grpd}}$ models homotopy types (using Corollary 7.4 for example). But this is also the case of Grpd-valued presheaves over the cube category with or without connections, over Joyal's $\Theta$ category, over the dendroidal category, etc.

Let us now compare groupoidal weak test categories with weak test categories. For this, first recall the following technical result.

**7.7 Lemma** *Let* $u : A \to B$ *be a morphism of* Cat, *and* $i : A \to$ Cat *and* $j : B \to$ Cat *be such that the triangle*

$$
\begin{array}{ccc}
A & \xrightarrow{u} & B \\
& {\scriptstyle i}\searrow & \downarrow {\scriptstyle j} \\
& & \mathsf{Cat}
\end{array}
$$

*is commutative, and suppose that for every object* $b$ *of* B, *the category* $j(b)$ *has a terminal object.*

(a) *If $u$ is an aspherical morphism of* Cat, *then $i : \mathsf{A} \to$ Cat is an aspherical functor if and only if $j : \mathsf{B} \to$ Cat is.*

(b) *If $j : \mathsf{B} \to$ Cat is fully faithful and $i : \mathsf{A} \to$ Cat is an aspherical functor, then $u$ is an aspherical morphism of* Cat *and $j : \mathsf{B} \to$ Cat is an aspherical functor.*

**Proof** This is [17, lemme 1.7.4] with the only exception that our definition of aspherical functor (Definition 2.15) is stronger than the one used there (see Remark 2.16). The only adaptation required is to notice that with our hypothesis, it is clear that $i(a)$ has a terminal object for every object $a$ in A. □

For Grpd-valued presheaves, we have the following partial generalization.

**7.8 Lemma** *Let $u : \mathsf{A} \to \mathsf{B}$ be a morphism of* Cat, *and $i : \mathsf{A} \to$ Cat and $j : \mathsf{B} \to$ Cat be such that the triangle*

$$\mathsf{A} \xrightarrow{\ u\ } \mathsf{B}$$
$$\overset{i}{\searrow} \quad \downarrow{\scriptstyle j}$$
$$\mathsf{Cat}$$

*is commutative, and suppose that for every object $b$ of* B, *the category $j(b)$ has a terminal object.*

(a) *If $u$ is an aspherical morphism of* Cat, *then $i : \mathsf{A} \to$ Cat is a groupoidal aspherical functor if and only if $j : \mathsf{B} \to$ Cat is.*

(b) *If $j : \mathsf{B} \to$ Cat is fully faithful, $i : \mathsf{A} \to$ Cat is a groupoidal aspherical functor, and for every object $b$ of* B, *the category $j(b)$ does not have any nontrivial isomorphism, then $u$ is an aspherical morphism of* Cat *and $j : \mathsf{B} \to$ Cat is a groupoidal aspherical functor.*

**Proof** Notice first that the hypotheses imply that the category $i(a)$ has a terminal object for every object $a$ of A.

The given commutative triangle induces a commutative triangle

$$\widehat{\mathsf{A}}_{\mathsf{Grpd}} \xleftarrow{\ u^*\ } \widehat{\mathsf{B}}_{\mathsf{Grpd}}$$
$$\overset{I^*}{\nwarrow} \quad \uparrow{\scriptstyle J^*}$$
$$\mathsf{Cat}$$

If $u$ is aspherical, it follows from Proposition 3.9 that for a small aspherical category C, $J^*(\mathsf{C})$ is aspherical if and only if $I^*(\mathsf{C})$ is aspherical. In particular, this proves (a).

For (b), notice that with the hypotheses, we have, for every object $b$ of B,

$$J^*(j(b)) = j^*(j(b)) \simeq b,$$

where the first equality comes from the fact that $j(b)$ does not have any nontrivial isomorphism and the second from the fact that $j$ is fully faithful (note that on the right-hand side of the second equality, we abusively wrote $b$ for the Set-valued presheaf represented by $b$). We then have

$$u^*(b) \simeq u^*\big(J^*(j(b))\big) = I^*(j(b)).$$

Since $i$ is a groupoidal aspherical functor and $j(b)$ is an aspherical category, we have that $I^*(j(b))$ is aspherical and so is $u^*(b)$. This means that the category $i_A(u^*(b))$ is aspherical and an immediate verification shows that we have a canonical isomorphism $i_A(u^*(b)) \simeq A/b$, which proves by definition that $u$ is aspherical. Hence, we can apply (a) and the conclusion follows. □

**7.9 Remark** It is only the (b) of Lemma 7.7 that does not generalize straightforwardly in Lemma 7.8 and for which we added the hypothesis that for every $b$ in B, $j(b)$ does not have nontrivial isomorphisms. We do not know whether this hypothesis is necessary or not.

We can now prove the following result.

**7.10 Proposition** *Let* $i : A \to \mathrm{Cat}$ *be a functor, with* A *a small category, such that* $i(a)$ *has a terminal object for every object $a$ of* A. *We have the following implication*:

$$i \text{ is an aspherical functor } \Longrightarrow i \text{ is a groupoidal aspherical functor.}$$

*If we suppose moreover that for every object $a$ of* A, *the category* $i(a)$ *does not have any nontrivial isomorphism, then we also have the converse implication*:

$$i \text{ is a groupoidal aspherical functor } \Longrightarrow i \text{ is an aspherical functor.}$$

**Proof** We begin by the first implication. Suppose that $i : A \to \mathrm{Cat}$ is an aspherical functor and let B be the smallest full subcategory of Cat such that:

- Every $i(a)$, for $a$ in A, is an object of B.
- B is stable by finite products.

Since a finite product of categories with terminal object has a terminal object, it follows that every object of B has a terminal object. By construction, we have a factorization

$$A \xrightarrow{\ i_0\ } B$$
$$i \searrow \quad \downarrow j$$
$$\mathrm{Cat}$$

where $j$ is the canonical inclusion and $i_0 : A \to B$ is the functor $a \mapsto i(a)$. Since $i$ is an aspherical functor, Lemma 7.7(b) implies that $i_0$ is aspherical and $j : B \hookrightarrow \mathrm{Cat}$ is an aspherical functor. Since B is stable by finite products, it is totally aspherical (Example 6.2), and thus $j$ is also a locally aspherical functor (this follows easily from Lemma 6.5 applied to Set-valued presheaves). Applying Proposition 7.2, we obtain that $j$ is a groupoidal locally aspherical functor, and then a groupoidal aspherical functor because B is aspherical. Finally, using Lemma 7.8(a), we have that $i$ is a groupoidal aspherical functor.

For the converse implication, let $i : A \to \mathrm{Cat}$ be a groupoidal aspherical functor such that each $i(a)$, for $a$ in A, does not have any nontrivial isomorphism and let B be the smallest full subcategory of Cat such that:

- Every $i(a)$, for $a$ in A, is an object of B.
- Every object $\Delta_n$ of $\Delta$, for $n \geq 0$, is an object of B.

Notice that the objects of B do not have any nontrivial isomorphism. By construction, we have a commutative diagram

$$A \xrightarrow{\ i_0\ } B \xleftarrow{\ k_0\ } \Delta$$
$$i \searrow \quad \downarrow j \quad \swarrow k$$
$$\mathsf{Cat}$$

where:

- $i_0$ is the functor $a \mapsto i(a)$.
- $j$ is the full subcategory inclusion.
- $k$ is the canonical fully faithful inclusion of $\Delta$ in Cat.
- $k_0$ is the full subcategory inclusion.

By hypothesis, $i$ is a groupoidal aspherical functor and since $j$ is fully faithful and the objects of B do not have any nontrivial isomorphism, it follows from Lemma 7.8(b) that $i_0$ is an aspherical morphism of Cat. Since $k : \Delta \to \mathsf{Cat}$ is an aspherical functor [17, exemple 1.7.18] (the functor $k^* : \mathsf{Cat} \to \widehat{\Delta}$ is nothing but the nerve functor), and since $j$ is fully faithful, it follows from Lemma 7.7(b) that $j$ is also an aspherical functor. Using that $i_0$ is an aspherical morphism of Cat, we deduce from an application of Lemma 7.7(a) that $i$ is an aspherical functor. $\qquad\square$

From this we deduce the following comparison theorem.

**7.11 Theorem** *Let* A *be a small category. We have the following implication*:

A *is a weak test category* $\implies$ A *is a groupoidal weak test category.*

*If we suppose that* A *does not have any nontrivial isomorphism*, *then we also have the converse implication*:

A *is a groupoidal weak test category* $\implies$ A *is a weak test category.*

**Proof** By Proposition 4.12 (resp Remark 2.17), A is groupoidal weak test (resp weak test) if and only if A $\to$ Cat, $a \mapsto$ A/$a$ is a groupoidal aspherical functor (resp aspherical functor). Hence, the result follows immediately from Proposition 7.10. $\qquad\square$

**7.12 Remark** In light of Remark 7.9, we do not know if the hypothesis that A does not have any nontrivial isomorphism is necessary for the second implication of the previous theorem. If a counterexample exists, then it would necessarily be a small category A with nontrivial isomorphisms which is groupoidal weak test but not groupoidal test (or else Theorem 7.3 applies).

**7.13 Corollary** *If* A *is a weak test category*, *then the canonical inclusion* $\widehat{A} \to \widehat{A}_{\mathsf{Grpd}}$ *induces an equivalence at the level of homotopy categories*

$$\mathrm{Ho}(\widehat{A}) \simeq \mathrm{Ho}(\widehat{A}_{\mathsf{Grpd}}).$$

**Proof** Similar to the proof of Corollary 7.4. $\qquad\square$

**7.14 Remark** Same remark as for Remark 7.5.

**7.15 Example** It follows from Theorem 7.3 that all examples of weak test categories from Example 2.13 are also groupoidal weak test categories. For example, it is the case of the category $\Delta'$ of finite nonempty ordinals and nondecreasing monomorphisms. In particular, the category $(\widehat{\Delta'})_{\mathsf{Grpd}}$ models homotopy types.

**7.16** Finally, let us end this section with a quick word on the comparison of pseudotest categories and groupoidal pseudotest categories. Although the implication

$$(2) \qquad\qquad \text{pseudotest category} \implies \text{groupoidal pseudotest category}$$

seems reasonable to expect, it remains an open question for the author. As for the converse implication, it is not true in general. More precisely, the example below shows that there exists a basic localizer $\mathcal{W}$ (which is not $\mathcal{W}_\infty$!) such that the class of $\mathcal{W}$-groupoidal pseudotest categories strictly contains the class of $\mathcal{W}$-pseudotest categories. (Hence, for this *particular* basic localizer, the implication (2) is true.) The question remains open for an arbitrary basic localizer, in particular for $\mathcal{W}_\infty$.

**7.17 Example** Consider the functor $\pi_1 \colon \mathsf{Cat} \to \mathsf{Grpd}$, left adjoint of the canonical inclusion functor $\iota \colon \mathsf{Grpd} \to \mathsf{Cat}$, and let $\mathcal{W}_1$ be the class of morphisms $f$ of $\mathsf{Cat}$ such that $\pi_1(f)$ is an equivalence of groupoids. We leave it as an exercise to the reader to show that $\mathcal{W}_1$ is a basic localizer of $\mathsf{Cat}$. Since $\iota$ is fully faithful, the counit of the adjunction $\pi_1 \dashv \iota$ is an isomorphism and it follows then from Lemma 1.3 that this adjunction induces a homotopical equivalence between $(\mathsf{Cat}, \mathcal{W}_1)$ and $(\mathsf{Grpd}, \mathcal{W}_{\mathsf{EqGrpd}})$, where $\mathcal{W}_{\mathsf{EqGrpd}}$ is the class of equivalences of groupoids. (This proves in particular that $\mathsf{Cat}$ models homotopy 1-types.) It follows that a small category $\mathsf{A}$ is $\mathcal{W}_1$-groupoidal pseudotest if and only if it is $\mathcal{W}_1$-aspherical and the functor

$$\widehat{\mathsf{A}}_{\mathsf{Grpd}} \to \mathsf{Grpd}, \quad X \mapsto \pi_1(i_A(X)),$$

induces an equivalence at the level of homotopy categories.

Now, let $\mathsf{A} = e$ be the terminal category. The previous functor is nothing but the identity functor of $\mathsf{Grpd}$ and it follows trivially that $e$ is $\mathcal{W}_1$-groupoidal pseudotest. On the other hand, $e$ is not $\mathcal{W}_1$-pseudotest. Indeed, $i_e \colon \mathsf{Set} \to \mathsf{Cat}$ is simply the canonical inclusion functor and so $i_e^{-1}(\mathcal{W}_1)$ is the class of isomorphisms of $\mathsf{Set}$. If $e$ was $\mathcal{W}_1$ pseudotest, this would imply that $(\mathsf{Set}, \mathsf{Iso}) \hookrightarrow (\mathsf{Grpd}, \mathcal{W}_1)$ induces an equivalence at the level homotopy categories (and that $\mathsf{Set}$ models homotopy 1-types), which is easily seen to be false.

**7.18 Remark** Note that the previous example also shows that the class of $\mathcal{W}_1$-groupoidal weak test categories is strictly bigger than the class of $\mathcal{W}_1$-groupoidal pseudotest categories. Indeed, if the terminal category $e$ were a $\mathcal{W}_1$-groupoidal weak test category, then it would be a $\mathcal{W}_1$-weak test category (since it does not have any nontrivial isomorphisms), and in particular a $\mathcal{W}_1$-pseudotest category.

# 8 Weak equivalences via the nerve

We now give an equivalent definition of weak equivalences of $\mathsf{Grpd}$-valued presheaves in terms of nerve functors. In particular, in Example 8.10 below, we recover the usual definition of weak equivalences on $\widehat{\Delta}_{\mathsf{Grpd}}$ used in the literature [6, Section 8; 12].

**8.1**  Let A be a small category. For a functor $X \colon \mathsf{A}^{\mathrm{op}} \to \mathsf{Cat}$, we denote by $\int_\mathsf{A} X$ the *Grothendieck construction of $X$*. This means that $\int_\mathsf{A} X$ is the category such that:

- Objects are pairs $(a, x)$ where $a$ is an object of A and $x$ is an object of $X(a)$.

- A morphism $(a, x) \to (a', x')$ is a pair $(f, k)$ where $f \colon a \to a'$ is a morphism of A and $k \colon x \to X(f)(x')$ is a morphism of $X(a)$.

(For details, we refer to [17, paragraphe 2.2.6], where the notation $\nabla_\mathsf{A}$ for $\int_\mathsf{A}$ is used.) This construction is functorial and provides a functor

$$\int_\mathsf{A} \colon \widehat{\mathsf{A}}_\mathsf{Cat} \to \mathsf{Cat},$$

where we write $\widehat{\mathsf{A}}_\mathsf{Cat}$ for the category of functors $\mathsf{A}^{\mathrm{op}} \to \mathsf{Cat}$ and natural transformations between them.

We have canonical inclusions $\widehat{\mathsf{A}} \hookrightarrow \widehat{\mathsf{A}}_\mathsf{Grpd} \hookrightarrow \widehat{\mathsf{A}}_\mathsf{Cat}$, and, as already observed, for an object $X$ of $\widehat{\mathsf{A}}$ we have $\int_\mathsf{A} X = i_\mathsf{A}(X)$, and for an object $X$ of $\widehat{\mathsf{A}}_\mathsf{Grpd}$ we have $\int_\mathsf{A} X = I_\mathsf{A}(X)$.

**8.2 Proposition**  [17, proposition 2.3.1]  *Let $\mathcal{W}$ be a basic localizer of $\mathsf{Cat}$ and A a small category. The functor $\int_\mathsf{A} \colon \widehat{\mathsf{A}}_\mathsf{Cat} \to \mathsf{Cat}$ sends pointwise $\mathcal{W}$-equivalences to $\mathcal{W}$-equivalences.*

We now fix once and for all a basic localizer $\mathcal{W}$ of $\mathsf{Cat}$.

**8.3**  Let A and B be two small categories and consider the presheaf category $\widehat{\mathsf{A} \times \mathsf{B}}$. Using the identification $\widehat{\mathsf{A} \times \mathsf{B}} \simeq \underline{\mathrm{Hom}}(\mathsf{A}^{\mathrm{op}}, \widehat{\mathsf{B}})$ and the postcomposition by the functor $i_\mathsf{B} \colon \widehat{\mathsf{B}} \to \mathsf{Cat}$ defines a functor

$$i_\mathsf{B} \colon \widehat{\mathsf{A} \times \mathsf{B}} \to \widehat{\mathsf{A}}_\mathsf{Cat}$$

which we abusively denote by $i_\mathsf{B}$ again. Then if we postcompose by $\int_\mathsf{A}$, we obtain a functor

$$\int_\mathsf{A} i_\mathsf{B} \colon \widehat{\mathsf{A} \times \mathsf{B}} \to \mathsf{Cat}.$$

The proof of the following lemma is a straightforward verification, which we leave to the reader.

**8.4 Lemma**  *For every object $X$ of $\widehat{\mathsf{A} \times \mathsf{B}}$, there is a canonical isomorphism*

$$i_{\mathsf{A} \times \mathsf{B}}(X) \simeq \int_\mathsf{A} i_\mathsf{B}(X),$$

*which is natural in $X$.*

**8.5 Remark**  Remember that the Grothendieck construction of a functor with values in $\mathsf{Cat}$ is weakly equivalent to its homotopy colimit (with respect to any basic localizer on $\mathsf{Cat}$) [17, théorème 3.1.7].[10] Since the functor $i_\mathsf{B}$ is just the restriction of the Grothendieck construction to Set-valued presheaves, the previous lemma can be simply restated by saying that the homotopy colimit of a functor of two variables is computed by successively taking the homotopy colimit relative to each variable.

---

[10]When the basic localizer is $\mathcal{W}_\infty$, this is a result of Thomason [21].

**8.6**   Let A and B be small categories and let $i : B \to \mathsf{Cat}$ be a functor. Recall that we denote by $i^* : \mathsf{Cat} \to \widehat{B}$ the functor $C \mapsto \mathrm{Hom}_{\mathsf{Cat}}(i(-), C)$. By considering Grpd as a subcategory of Cat, the functor $i^*$ induces by postcomposition a functor

$$i^* : \widehat{A}_{\mathsf{Grpd}} \to \widehat{A \times B},$$

which we abusively denote by $i^*$ as well.

**8.7  Proposition**   *Let* A *and* B *be small categories and* $i : B \to \mathsf{Cat}$ *a functor such that for every* $b$ *in* B, *the category* $i(b)$ *has a terminal object. Then, there exists a natural transformation*



*If* $i$ *is an aspherical functor, then this natural transformation is a weak equivalence argument by argument.*

**Proof**   For every $b$ in B, let $e_b$ be the terminal object of $i(b)$. By an analogous construction as the one in Section 4.8 in the case of Set-valued presheaves (see [4, paragraphe 3.2.4] for details), for every small category C we define a morphism

$$(3) \qquad\qquad \alpha_C : i_B i^*(C) \to C, \quad (b, p : i(b) \to C) \mapsto p(e_b),$$

which is natural in C. For every $X$ in $\widehat{A}_{\mathsf{Grpd}}$ and $a$ in A, we obtain a map

$$\alpha_{X(a)} : i_B i^*(X(a)) \to X(a),$$

natural in $X$ and $a$, and by applying $\int_A$, we obtain a canonical map

$$i_{A \times B} i^*(X) \simeq \int_A i_B i^*(X) \to \int_A X = I_A(X),$$

which is natural in $X$. Now, if $i$ is an aspherical functor, then by [17, proposition 1.7.6] (which is the analogue of our Proposition 4.11 for Set-valued presheaves), the map (3) is a weak equivalence. We conclude with Proposition 8.2. $\qquad\square$

**8.8  Corollary**   *Let* A *and* B *be small categories, and* $i : B \to \mathsf{Cat}$ *an aspherical functor* (*such that* $i(b)$ *has a terminal object for every object* $b$ *of* B). *Then* $i^* : \widehat{A}_{\mathsf{Grpd}} \to \widehat{A \times B}$ *preserves and reflects weak equivalences*:

$$\mathcal{W}_{\widehat{A}_{\mathsf{Grpd}}} = i^{*-1}(\mathcal{W}_{\widehat{A \times B}}).$$

A particular case where the previous corollary applies is the following.

**8.9  Corollary**   *Let* A *be a totally aspherical small category and* $i : A \to \mathsf{Cat}$ *be an aspherical functor* (*such that* $i(a)$ *has a terminal object for every object* $a$ *of* A). *Then a morphism of* $\widehat{A}_{\mathsf{Grpd}}$ *is a weak equivalence if and only if its image by* $i^* : \widehat{A}_{\mathsf{Grpd}} \to \widehat{A \times A}$ *is a diagonal weak equivalence*:

$$\mathcal{W}_{\widehat{A}_{\mathsf{Grpd}}} = i^{*-1}(\delta^{*-1}(\mathcal{W}_{\widehat{A}})),$$

where $\delta^* : \widehat{A \times A} \to \widehat{A}$ *is the diagonal functor.*

**Proof** If A is totally aspherical, then the diagonal functor $\delta \colon A \to A \times A$ is aspherical and so the induced functor $\delta^* \colon \widehat{A \times A} \to \widehat{A}$ preserves and reflects weak equivalences [17, proposition 1.2.9(d)]. □

**8.10 Example** Let $A = B = \Delta$ and $i \colon \Delta \to Cat$ be the canonical inclusion, so that $i^* \colon Cat \to \widehat{\Delta}$ is the usual nerve functor. The previous corollary implies that the weak equivalences on simplicial groupoids $\widehat{\Delta}_{Grpd}$ are exactly those morphisms that induce diagonal weak equivalences of bisimplicial sets.

# 9 Bonus result: groupoids internal to categories model homotopy types

**9.1** Let $\mathcal{C}$ be a category with finite limits (or even only pullbacks). A groupoid internal to $\mathcal{C}$ consists of a pair of objects $(X_0, X_1)$ of $\mathcal{C}$ equipped with

- source and target maps $s, t \colon X_1 \to X_0$,
- a unit map $i \colon X_0 \to X_1$ such that $s \circ i = t \circ i = \mathrm{id}_{X_0}$,
- an inverse map $\mathrm{inv} \colon X_1 \to X_1$ such that $s \circ \mathrm{inv} = t$ and $t \circ \mathrm{inv} = s$,
- a composition map $m \colon X_1 \times_{X_0} X_1 \to X_1$ such that $s \circ m = s \circ \pi_1$ and $t \circ m = t \circ \pi_2$, where $X \times_{X_0} X_1$, $\pi_1$ and $\pi_2$ are defined as the fibered product

$$
\begin{array}{ccc}
X_1 \times_{X_0} X_1 & \xrightarrow{\ \pi_1\ } & X_1 \\
\big\downarrow{\scriptstyle \pi_2} & \lrcorner & \big\downarrow{\scriptstyle t} \\
X_1 & \xrightarrow[\ s\ ]{} & X_0
\end{array}
$$

All these objects satisfy the usual axioms saying that $m$ is associative, $i$ is the unit on the left and right of the composition, and inv gives the inverse on the left and right of the composition.

We shall often abuse notation and refer to a groupoid internal to $\mathcal{C}$ as a pair $X = (X_0, X_1)$.

An internal morphism of groupoids $f \colon X \to X'$ is a pair of morphisms $(f_0 \colon X_0 \to X'_0, f_1 \colon X_1 \to X'_1)$ of $\mathcal{C}$ which commute with source, target, inverse, unit and composition in the obvious way. Internal groupoids and internal morphisms of groupoids form a category denoted by $\mathsf{Grpd}(\mathcal{C})$.

If $\mathcal{C}'$ is another category with pullbacks and $F \colon \mathcal{C} \to \mathcal{C}'$ is a functor preserving pullbacks, then $F$ sends groupoids internal to $\mathcal{C}$ to groupoids internal to $\mathcal{C}'$, hence there is an induced functor

$$\mathsf{Grpd}(F) \colon \mathsf{Grpd}(\mathcal{C}) \to \mathsf{Grpd}(\mathcal{C}').$$

Furthermore, if $F$ admits a left adjoint $G \colon \mathcal{C}' \to \mathcal{C}$ that *also preserves pullbacks*, then we have an induced adjunction $\mathsf{Grpd}(G) \dashv \mathsf{Grpd}(F)$. The unit and counit of this adjunction are obtained by applying those of the adjunction $G \dashv F$ levelwise. This means that if $(X_1, X_0)$ is an object of $\mathsf{Grpd}(\mathcal{C})$, then the counit of this adjunction is simply

$$(\varepsilon_{X_1} \colon GF(X_1) \to X_1, \ \varepsilon_{X_0} \colon GF(X_0) \to X_0),$$

where $\varepsilon$ is the counit of the adjunction $G \dashv F$, and similarly for the unit.

**9.2 Remark** Even if the left adjoint of $F$ does not preserve pullbacks, under mild conditions (eg $\mathcal{C}$ and $\mathcal{C}'$ are locally presentable and $F$ is accessible) the functor $\mathsf{Grpd}(F)$ still admits a left adjoint. In general, this left adjoint is hard to manipulate and is constructed abstractly. The point of the previous paragraph is that, *when the left adjoint $G$ of $F$ preserves pullbacks*, the left adjoint of $\mathsf{Grpd}(F)$ is simply $\mathsf{Grpd}(G)$, and furthermore the unit and counit have a particularly nice form. This will play an important role later in the proof of Proposition 9.9 (see also Remark 9.10).

**9.3** Let $\mathcal{C}$ be a category with pullbacks and $X = (X_0, X_1)$ an internal groupoid of $\mathcal{C}$. The *nerve of $X$* is the simplicial object $N_*(X)\colon \Delta^{\mathrm{op}} \to \mathcal{C}$ defined by

$$N_n(X) = \underbrace{X_1 \times_{X_0} \cdots \times_{X_0} X_1}_{n \text{ times}},$$

where for $n \geq 2$, this means that $N_n(X)$ is the limit of the diagram
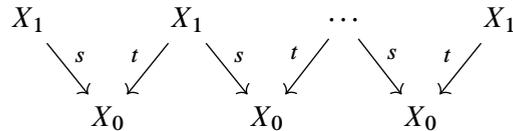


where $X_1$ appears $n$ times, and, by convention, for $n = 0, 1$,

$$N_0(X) = X_0, \quad N_1(X) = X_1.$$

The face and degeneracy maps are defined as follows:

- For $0 < j < n$, $\partial_j\colon N_n(X) \to N_{n-1}(X)$ is induced by $m\colon X_1 \times_{X_0} X_1 \to X_1$ acting on the $(j-1)^{\mathrm{st}}$ and $j^{\mathrm{th}}$ factors of $N_n(X)$.
- For $n \geq 1$, $j = 0$ or $j = n$, $\partial_j\colon N_n(X) \to N_{n-1}(X)$ is the canonical projection that discards the $j^{\mathrm{th}}$ factor of $N_n(X)$.
- For $0 \leq j \leq n$, $\epsilon_j\colon N_n(X) \to N_{n+1}(X)$ is induced by $i\colon X_0 \to X_1$ hitting the $j^{\mathrm{th}}$ factor of $N_{n+1}(X)$.

Note that this construction does not use the inverse map $\mathrm{inv}\colon X_1 \to X_1$ and thus only depends on the underlying category (internal to $\mathcal{C}$) of $X$.

**9.4 Lemma** *Let $\mathcal{C}, \mathcal{C}'$ be categories with pullbacks and $F\colon \mathcal{C} \to \mathcal{C}'$ a functor preserving pullbacks. The following diagram is commutative (up to an isomorphism of functors):*



**Proof** This is a straightforward verification left to the reader. □

We now fix once and for all a basic localizer $\mathcal{W}$ on $\mathsf{Cat}$.

**9.5** Let $\mathsf{Grpd}(\mathsf{Cat})$ be the category of groupoids internal to $\mathsf{Cat}$. We denote by $\widehat{\Delta}_{\mathsf{Cat}}$ the category of $\mathsf{Cat}$-valued presheaves over $\Delta$. The construction from Section 9.3 yields a functor $N_*\colon \mathsf{Grpd}(\mathsf{Cat}) \to \widehat{\Delta}_{\mathsf{Cat}}$. If we postcompose by the Grothendieck construction, we obtain a functor from $\mathsf{Grpd}(\mathsf{Cat})$ to $\mathsf{Cat}$:

$$\mathsf{Grpd}(\mathsf{Cat}) \xrightarrow{N_*} \widehat{\Delta}_{\mathsf{Cat}} \xrightarrow{\int_\Delta} \mathsf{Cat}.$$

**9.6 Definition** A morphism $f\colon X \to Y$ of $\mathsf{Grpd}(\mathsf{Cat})$ is a *weak equivalence* if

$$\int_\Delta N_*(f)\colon \int_\Delta N_*(X) \to \int_\Delta N_*(Y)$$

is in $\mathcal{W}$.

**9.7** Let $\mathsf{A}$ be a small category and consider the adjunction $i_\mathsf{A}\colon \widehat{\mathsf{A}} \rightleftarrows \mathsf{Cat} \colon i_\mathsf{A}^*$. The key observation is that the left adjoint $i_\mathsf{A}$ preserves pullbacks (this can been seen by observing that the Grothendieck construction $\int_\mathsf{A}\colon \widehat{\mathsf{A}}_{\mathsf{Cat}} \to \mathsf{Cat}/\mathsf{A}$ has a left adjoint, and the forgetful functor $\mathsf{Cat}/\mathsf{A} \to \mathsf{Cat}$ preserves pullbacks). In particular, it follows from Section 9.1 that we have an induced adjunction

$$\mathsf{Grpd}(i_\mathsf{A})\colon \mathsf{Grpd}(\widehat{\mathsf{A}}) \rightleftarrows \mathsf{Grpd}(\mathsf{Cat}) \colon \mathsf{Grpd}(i_\mathsf{A}^*).$$

We also have a canonical isomorphism $\mathsf{Grpd}(\widehat{\mathsf{A}}) \simeq \widehat{\mathsf{A}}_{\mathsf{Grpd}}$.

**9.8 Lemma** *Let* $\mathsf{A}$ *be a small category. The functor*

$$\mathsf{Grpd}(i_\mathsf{A})\colon \widehat{\mathsf{A}}_{\mathsf{Grpd}} \simeq \mathsf{Grpd}(\widehat{\mathsf{A}}) \to \mathsf{Grpd}(\mathsf{Cat})$$

*preserves and reflects weak equivalences.*

**Proof** Consider the commutative square (up to isomorphism) from Lemma 9.4

$$
\begin{array}{ccc}
\widehat{\mathsf{A}}_{\mathsf{Grpd}} \simeq \mathsf{Grpd}(\widehat{\mathsf{A}}) & \xrightarrow{\ \mathsf{Grpd}(i_\mathsf{A})\ } & \mathsf{Grpd}(\mathsf{Cat}) \\
{\scriptstyle N_*}\big\downarrow & & \big\downarrow{\scriptstyle N_*} \\
\widehat{\Delta \times \mathsf{A}} \simeq \underline{\mathrm{Hom}}(\Delta^{\mathrm{op}}, \widehat{\mathsf{A}}) & \xrightarrow[\underline{\mathrm{Hom}}(\Delta^{\mathrm{op}}, i_\mathsf{A})]{} & \widehat{\Delta}_{\mathsf{Cat}}
\end{array}
$$

Let us say that a morphism $f$ of $\widehat{\Delta}_{\mathsf{Cat}}$ is a weak equivalence if $\int_\Delta f$ is in $\mathcal{W}$. Then, by definition, the vertical arrow on the right in the above square preserves and reflects weak equivalences, and it follows from Lemma 8.4 that the bottom horizontal arrow of the above square also preserves and reflects weak equivalences. By Corollary 8.8, the vertical arrow on the left of the previous square preserves and reflects weak equivalences, and the desired conclusion follows at once. $\square$

**9.9 Proposition** *Let* $\mathsf{A}$ *be a weak test category. The adjunction*

$$\mathsf{Grpd}(i_\mathsf{A})\colon \widehat{\mathsf{A}}_{\mathsf{Grpd}} \simeq \mathsf{Grpd}(\widehat{\mathsf{A}}) \rightleftarrows \mathsf{Grpd}(\mathsf{Cat}) \colon \mathsf{Grpd}(i_\mathsf{A}^*)$$

*is a homotopical equivalence (Section 1.2). In particular it induces an adjoint equivalence after localization,*

$$\mathrm{Ho}(\widehat{\mathsf{A}}_{\mathsf{Grpd}}) \rightleftarrows \mathrm{Ho}(\mathsf{Grpd}(\mathsf{Cat})).$$

**Proof** From Lemma 9.8, we know that $\mathsf{Grpd}(i_A)$ preserves and reflects weak equivalences. Hence, by Lemma 1.3, we need to show that for every object $X = (X_0, X_1)$ of $\mathsf{Grpd}(\mathsf{Cat})$, the counit

$$\varepsilon_X \colon \mathsf{Grpd}(i_A \circ i_A^*)(X) \to X$$

is a weak equivalence of $\mathsf{Grpd}(\mathsf{Cat})$. By definition, this means that we have to prove that $\int_\Delta N_*(\varepsilon_X)$ is in $\mathcal{W}$. It follows from the fact that $i_A$ preserves pullbacks that $N_*(\mathsf{Grpd}(i_A \circ i_A^*)(X))$ can be identified with the simplicial object in $\mathsf{Cat}$

$$\Delta^{\mathrm{op}} \to \mathsf{Cat}, \quad [n] \mapsto i_A^* i_A (\underbrace{X_1 \times_{X_0} \cdots \times_{X_0} X_1}_{n \text{ times}}),$$

and that $N_*(\epsilon_X)$ can be identified with the morphism of simplicial objects of $\mathsf{Cat}$ which is the counit of the adjunction $i_A \dashv i_A^*$ levelwise,

$$i_A i_A^* (X_1 \times_{X_0} \cdots \times_{X_0} X_1) \to X_1 \times_{X_0} \cdots \times_{X_0} X_1;$$

see Section 9.1. By hypothesis, $A$ is a weak test category and thus these morphisms are all in $\mathcal{W}$. The conclusion follows then from Proposition 8.2. $\qquad\qquad\square$

**9.10 Remark** In particular, when $\mathcal{W} = \mathcal{W}_\infty$, the previous proposition shows that $\mathsf{Grpd}(\mathsf{Cat})$ models homotopy types. This is slightly surprising considering that groups internal to categories only model pointed connected homotopy 2-types [15; 16]. Somehow, restricting to groups internal to categories instead of groupoids internal to categories does not only amount to considering pointed connected objects. This is in contrast with what happens for simplicial sets: groupoids internal to simplicial sets model homotopy types and groups internal to simplicial sets model pointed connected homotopy types [14].

A hint of explanation comes from the fact that the functor $i_A \colon \widehat{A} \to \mathsf{Cat}$ does not preserve products (even if it preserves pullbacks, as we have already seen), and in particular does not preserve group objects. Hence, the strategy to prove Proposition 9.9 cannot be adapted for group objects instead of groupoid objects.

**9.11 Remark** The proof of Proposition 9.9 does not depend on the theory of groupoidal test categories (as long as we define weak equivalences in $\mathsf{Grpd}(\widehat{A})$ via the nerve as in Section 8). In particular, if we already know that $\mathsf{Grpd}(\widehat{\Delta})$ models homotopy types (for example, by appealing to [6, Theorem 8.3; 12, Theorem 10]), then we obtain a second proof of the fact that $\mathsf{Grpd}(\widehat{A})$ models homotopy type for any (weak) test category $A$. Indeed, first we apply Proposition 9.9 for $A = \Delta$ to deduce that $\mathsf{Grpd}(\mathsf{Cat})$ models homotopy types and then we apply the same result again for an arbitrary (weak) test category $A$. Note, however, that we do *not* recover this way the other direction in the equivalence (b) of Theorem 7.3.

# References

[1]  **D Ara**, **D-C Cisinski**, **I Moerdijk**, *The dendroidal category is a test category*, Math. Proc. Cambridge Philos. Soc. 167 (2019) 107–121  MR

[2]   **C Barwick**, **D M Kan**, *A characterization of simplicial localization functors and a discussion of DK equivalences*, Indag. Math. 23 (2012) 69–79  MR

[3]   **D-C Cisinski**, *Le localisateur fondamental minimal*, Cah. Topol. Géom. Différ. Catég. 45 (2004) 109–140  MR

[4]   **D-C Cisinski**, *Les préfaisceaux comme modèles des types d'homotopie*, Astérisque 308, Soc. Math. France, Paris (2006)  MR

[5]   **D-C Cisinski**, **G Maltsiniotis**, *La catégorie Θ de Joyal est une catégorie test*, J. Pure Appl. Algebra 215 (2011) 962–982  MR

[6]   **S E Crans**, *Quillen closed model structures for sheaves*, J. Pure Appl. Algebra 101 (1995) 35–57  MR

[7]   **A Dold**, *Homology of symmetric products and other functors of complexes*, Ann. of Math. 68 (1958) 54–80  MR

[8]   **P Gabriel**, **M Zisman**, *Calculus of fractions and homotopy theory*, Ergebnisse der Math. 35, Springer (1967)  MR

[9]   **J W Gray**, *The categorical comprehension scheme*, from "Category theory, homology theory and their applications, III", Lecture Notes in Math. 99, Springer (1969) 242–312  MR

[10]  **A Grothendieck**, *Pursuing stacks (À la poursuite des champs)*, *I*, Doc. Math. (Paris) 20, Soc. Math. France, Paris (2022)  MR

[11]  **L Illusie**, *Complexe cotangent et déformations, II*, Lecture Notes in Math. 283, Springer (1972)  MR

[12]  **A Joyal**, **M Tierney**, *On the homotopy theory of sheaves of simplicial groupoids*, Math. Proc. Cambridge Philos. Soc. 120 (1996) 263–290  MR

[13]  **D M Kan**, *Functors involving c.s.s. complexes*, Trans. Amer. Math. Soc. 87 (1958) 330–346  MR

[14]  **D M Kan**, *On homotopy theory and c.s.s. groups*, Ann. of Math. 68 (1958) 38–53  MR

[15]  **J-L Loday**, *Spaces with finitely many nontrivial homotopy groups*, J. Pure Appl. Algebra 24 (1982) 179–202  MR

[16]  **S MacLane**, **J H C Whitehead**, *On the 3-type of a complex*, Proc. Nat. Acad. Sci. U.S.A. 36 (1950) 41–48  MR

[17]  **G Maltsiniotis**, *La théorie de l'homotopie de Grothendieck*, Astérisque 301, Soc. Math. France, Paris (2005)  MR

[18]  **G Maltsiniotis**, *La catégorie cubique avec connexions est une catégorie test stricte*, Homology Homotopy Appl. 11 (2009) 309–326  MR

[19]  **J Milnor**, *The geometric realization of a semi-simplicial complex*, Ann. of Math. 65 (1957) 357–362  MR

[20]  **R Street**, *Limits indexed by category-valued 2-functors*, J. Pure Appl. Algebra 8 (1976) 149–181  MR

[21]  **R W Thomason**, *Homotopy colimits in the category of small categories*, Math. Proc. Cambridge Philos. Soc. 85 (1979) 91–109  MR

[22]  **A P Tonks**, *Cubical groups which are Kan*, J. Pure Appl. Algebra 81 (1992) 83–87  MR

*Mathematical Institute, Utrecht University*
*Utrecht, Netherlands*

`l.s.guetta@uu.nl`

`https://leoguetta.github.io/`

# Product and coproduct on fixed point Floer homology of positive Dehn twists

Yuan Yao

Ziwen Zhao

We compute the product and coproduct structures on the fixed point Floer homology of iterations on a single positive Dehn twist, subject to some mild topological restrictions. We show that the resulting product and coproduct structures are determined by the product and coproduct on Morse homology of the complement of the twist region, together with certain sectors of product and coproduct structures on the symplectic homology of $T^*S^1$. The computation is done via a direct enumeration of $J$-holomorphic sections: we use a local energy inequality to show that some of the putative holomorphic sections do not exist, and we use a gluing construction plus some Morse–Bott theory to construct the sections we could not rule out.

## 1 Introduction

In this paper, we calculate the product and coproduct structures on fixed point Floer homologies of iterations of a single positive Dehn twist on a surface.

### 1.1 Fixed point Floer homology

In this section, we briefly review the definition of the fixed point Floer homology of a symplectomorphism on a compact surface. Let $(\Sigma, \omega_0)$ be a compact symplectic surface (possibly with boundary) and $\phi \colon \Sigma \to \Sigma$ a symplectomorphism. If $\partial\Sigma$ is nonempty, we further assume that near each boundary

component of $\partial \Sigma$, we can identify an open neighborhood with[1] $((-\epsilon_i, 0]_{x_i} \times S^1_{y_i}, dx_i \wedge dy_i)$, such that $\phi$ is the time-1 map of the Hamiltonian $H_i(x_i, y_i) = \theta_i x_i$ for a small irrational number $\theta_i$. Notice that under such assumptions, there are no fixed points of $\phi$ near $\partial \Sigma$. We assume that $\phi$ is nondegenerate, that is, for every fixed point $x$ of $\phi$, the linearization $d\phi_x$ does not have 1 as an eigenvalue. The fixed point Floer homology is the homology of the chain complex $(\mathrm{CF}_*(\Sigma, \phi), \partial)$, whose underlying module is generated over $\mathbb{Z}/2$ by all fixed points of $\phi$.

The differential of $\mathrm{CF}_*(\Sigma, \phi)$ is defined by counting $J$-holomorphic cylinders in the symplectization of the mapping torus of $\phi$. More precisely, for any symplectomorphism $\phi \colon \Sigma \to \Sigma$, the mapping torus $Y_\phi$ is defined as

$$Y_\phi = [0, 1]_t \times \Sigma / \big((1, p) \sim (0, \phi(p))\big).$$

The mapping torus comes with a projection $\pi \colon Y_\phi \to S^1_t$, and the symplectic form $\omega_0$ induces a closed 2-form $\omega_\phi \in \Omega^2(Y_\phi)$ which restricts to $\omega_0$ on each fiber (to be more precise, $\omega_0$ pulls back to a closed 2-form on the product $[0, 1] \times \Sigma$, and $\omega_\phi$ is the induced 2-form on the quotient space $Y_\phi$). The vector field $\partial_t$ on $[0, 1]_t \times \Sigma$ descends to a vector field on $Y_\phi$, which we still denote by $\partial_t$. Notice that there is a one-to-one correspondence between fixed points of $\phi$ and closed orbits of $\partial_t$ that cover $S^1_t$ once. We will denote by $\gamma_x$ the closed orbit associated to a fixed point $x$.

The projection $\pi$ now extends to $\mathbb{R} \times Y_\phi \to \mathbb{R} \times S^1$, and the fiberwise symplectic form $\omega_\phi$ extends to the symplectization as well. The symplectization of $Y_\phi$ is the 4-manifold $\mathbb{R}_s \times Y_\phi$ together with the symplectic form $ds \wedge dt + \omega_\phi$. An almost complex structure $J$ on $\mathbb{R} \times Y_\phi$ is called $\phi$-*compatible*, if it is invariant under the natural $\mathbb{R}$-action, sends $\partial_s$ to $\partial_t$, sends $\ker d\pi$ to itself, and that $\omega_\phi(\cdot, J\cdot)$ is a Riemannian metric on $\ker d\pi$. Given 2 fixed points $x$, $y$ of $\phi$, we define the moduli space $\mathcal{M}^J_{x,y}$ to be

$$\mathcal{M}^J_{x,y} := \big\{u \colon \mathbb{R}_s \times S^1_t \to \mathbb{R} \times Y_\phi \mid \partial_s u + J \partial_t u = 0; \lim_{s \to \infty} u(s, \cdot) = \gamma_x, \lim_{s \to -\infty} u(s, \cdot) = \gamma_y\big\}.$$

Now if $\phi$ is monotone (we will clarify this notion later in Definition 5.1), the differential on $\mathrm{CF}_*(\Sigma, \phi)$ is defined by

$$(1) \qquad \qquad \langle \partial x, y \rangle := \#_{\mathbb{Z}/2}(\mathcal{M}_{x,y}/\mathbb{R}).$$

Here $\mathcal{M}_{x,y}/\mathbb{R}$ is the natural quotient of $\mathcal{M}_{x,y}$ induced by the $\mathbb{R}$-translation on $\mathbb{R} \times Y_\phi$. Recall that every $J$-holomorphic section $u$ has a *Fredholm index*:

$$\mathrm{ind}(u) = 2c_1^\tau(u) + \mathrm{CZ}_\tau(\gamma_x) - \mathrm{CZ}_\tau(\gamma_y)$$

where $\tau$ is a trivialization of $\ker d\pi$ over the periodic orbits $\gamma_x$ and $\gamma_y$, $c_1^\tau(u)$ is the relative first Chern number,[2] and $\mathrm{CZ}_\tau$ is the Conley–Zehnder index of the Reeb orbits with respect to $\tau$. For a more detailed

---

[1] Here and throughout the paper, we use notation like $S^1_y$ and $[0, 1]_t$ to indicate the coordinates we choose. For example, $S^1_y$ denotes the circle $S^1$ with coordinates written as $y \in S^1$. Later on we will use the notation $[S^1_y]$ to denote the homology class generated by this circle whose coordinates we have chosen to be $y \in S^1$.

[2] The relative first Chern number $c_1^\tau(u)$ is defined as follows: $u^*(\ker d\pi)$ is a symplectic bundle over the domain of $u$. One chooses a generic section $\xi$ of this bundle which on each end is nonvanishing and constant with respect to the trivialization $\tau$ and $c_1^\tau(u)$ is defined to be the algebraic count of zeroes of $\xi$.

explanation on these terms, see Section 5.2. Under the monotonicity assumption, for a generic almost complex structure $J$, the set of Fredholm index one $J$-holomorphic sections modulo the natural $\mathbb{R}$ action is a compact 0-dimensional manifold, and $\#_{\mathbb{Z}/2}(\mathcal{M}^J_{x,y}/\mathbb{R})$ denotes the mod 2 count of points in the moduli space. If $J$ is generic then $\partial^2 = 0$, and we will denote by $\mathrm{HF}_*(\Sigma, \phi)$ the homology of $(\mathrm{CF}_*(\Sigma, \phi), \partial)$. The homology $\mathrm{HF}_*(\Sigma, \phi)$ is invariant under symplectic isotopies of $\phi$, see for example [23].

Fixed point Floer homology for a symplectomorphism of a surface has been computed in various cases, see for example [7; 18; 21; 10; 8; 6; 16]. Fixed point Floer homology can also be viewed as a special case of periodic Floer homology, which was calculated for iterations of a positive Dehn twist in [12].

## 1.2  Product and coproduct structures

Under suitable monotonicity assumptions[3] (see Definition 5.3), fixed point Floer homology is functorial, in the sense that fiberwise symplectic cobordisms with cylindrical ends induce morphisms between fixed point Floer homologies. In this paper, we focus on (completed) fiberwise symplectic cobordisms coming from the composition of two symplectomorphisms. We begin with a review of the concept of symplectic fiber bundles.

**Definition 1.1** [22, Definition 7.1]  Let $B$ be a smooth manifold. A symplectic fiber bundle $(E, \pi, \omega)$ over $B$ is a smooth proper submersion $\pi \colon E \to B$ together with a closed 2-form $\omega \in \Omega^2(E)$ such that the restriction of $\omega$ to any fiber is nondegenerate.

The mapping torus $Y_\phi$, together with the natural projection $\pi \colon Y_\phi \to S^1$ and the 2-form $\omega_\phi$, is an example of a symplectic fiber bundle. If $\phi, \psi \colon \Sigma \to \Sigma$ are two symplectomorphisms, then there is a symplectic fiber bundle $(X, \pi_X, \omega_X)$ over the thrice punctured sphere $B_0$, which, near the three punctures, is symplectomorphic to $[0, \infty) \times Y_\phi$, $[0, \infty) \times Y_\psi$ and $(-\infty, 0] \times Y_{\psi \circ \phi}$, respectively. A more precise description will appear in Section 2. For now, let us observe that (under assumptions on monotonicity, see Definition 5.3), such a bundle induces a morphism

$$\bullet \colon \mathrm{HF}_*(\Sigma, \phi) \otimes \mathrm{HF}_*(\Sigma, \psi) \to \mathrm{HF}_*(\Sigma, \psi \circ \phi)$$

This is what we call the *product structure* of the fixed-point Floer homology. In particular, if we assume that $\psi$ is isotopic to the identity, then the product structure gives a $H_*(\Sigma; \mathbb{Z}_2)$ module structure on $\mathrm{HF}_*(\Sigma, \phi)$. For computations of this module structure, see [21; 8; 10; 6]. Similarly, one can define, under suitable conditions, the *coproduct structure*:

$$\Delta \colon \mathrm{HF}_*(\Sigma, \psi \circ \phi) \to \mathrm{HF}_*(\Sigma, \phi) \otimes \mathrm{HF}_*(\Sigma, \psi).$$

Like the definition of the differential, the product and coproduct structures have geometric descriptions, this time by counting rigid pseudoholomorphic sections of the bundle $X \to B_0$ with appropriate asymptotes.

---

[3]Or with the use of Novikov rings.

Namely, if $x$, $y$, $z$ are fixed points of $\phi$, $\psi$ and $\psi \circ \phi$ respectively, and $J$ is a tame almost complex structure (see Section 2 for the definition) then the moduli space $\mathcal{M}_{x,y;z}^J$ is defined by

$$\mathcal{M}_{x,y;z}^J = \left\{ u \colon B_0 \to X \,\middle|\, \begin{array}{l} \pi_X \circ u = \text{id}, \ u \text{ is } J\text{-holomorphic, and} \\ u \text{ is asymptotic to } \gamma_x, \gamma_y \text{ and } \gamma_z \text{ over the} \\ \text{three appropriate punctures.} \end{array} \right\}.$$

The product (under suitable monotonicity assumptions; see Definition 5.3) on the chain level is now defined as

$$\langle x \bullet y, z \rangle = \#_{\mathbb{Z}/2} \mathcal{M}_{x,y;z}^J, \tag{2}$$

where $\#_{\mathbb{Z}/2} \mathcal{M}_{x,y;z}^J$ denotes the mod 2 count of Fredholm index 0 sections for a generic almost complex structure. The coproduct structure is defined in a similar way.

## 1.3  Main results

In this paper, we calculate the product and coproduct structures of the fixed-point Floer homology of iterations of a single positive Dehn twist. We fix a symplectic surface (possibly with boundary) $(\Sigma, \omega_0)$ and a noncontractible simple closed curve $\gamma \subset \Sigma$. We assume that $\gamma$ does not intersect $\partial \Sigma$ and choose a tubular neighborhood $N$ of $\gamma$ with coordinates $x \in (-\epsilon, 1 + \epsilon)$ and $y \in S^1 = \mathbb{R}/\mathbb{Z}$, where $\omega_0 = dx \wedge dy$. A (unperturbed) *positive Dehn twist* along $\gamma$ is a symplectomorphism of $\Sigma$, which has the form

$$\phi_0 \colon (x, y) \mapsto (x, y - x) \tag{3}$$

inside $N$, and is the time-1 map of a Morse function $H_0$ outside of $N' = [\epsilon, 1 - \epsilon] \times S^1 \subset N$. To break the degeneracy of the fixed points of $\phi_0$, we perturb $\phi_0$ near each of the $S^1$-family of orbits, breaking the Morse–Bott degeneracy, and denote by $\phi$ the perturbed twist. We further assume, as before, that near each boundary component of $\Sigma$, there are tubular coordinates $x_i \in (-\epsilon_i, 0]$, $y_i \in S^1$ and a small real number $\theta_i$ such that $H_0(x_i, y_i) = \theta_i x_i$ (later we will impose some other conditions on $H_0$; see Section 2).

Let us assume for now that $\partial \Sigma \neq \varnothing$ or the genus of $\Sigma$ is at least 2. Let $\Sigma_0$ denote $\Sigma - N'$. It was shown, for example in [21] or [12], that for all positive integers $m$,

$$\mathrm{HF}_*(\phi^m) \cong H_*(\Sigma_0; \mathbb{Z}_2) \oplus \left( \bigoplus_{i=1}^{m-1} H_*(S^1) \right). \tag{4}$$

The isomorphism can be understood as follows. With our assumption on $\phi$, the chain complex $\mathrm{CF}(\phi^m)$ is generated by two types of fixed points: those corresponding to the critical points of $H_0$ on $\Sigma_0$ and those from the breaking of the Morse–Bott $S^1$-family of the fixed points of $\phi_0^m$ inside $N'$. As it turns out, the two types of fixed points generate subcomplexes of $\mathrm{CF}(\phi^m)$, whose homologies correspond to the summands of the right-hand side of (4). Let $m$ and $n$ be two positive integers. With the previously described isomorphism understood, let proj denote the projection map $\mathrm{HF}_*(\phi^m) \to H_*(\Sigma_0; \mathbb{Z}_2)$, let $\cap$ denote the intersection product on $H_*(\Sigma_0; \mathbb{Z}_2)$ (which is the Poincaré dual of the cup product on $H^*(\Sigma_0, \partial \Sigma_0; \mathbb{Z}_2)$), and let $\iota$ denote the inclusion map $H_*(\Sigma_0; \mathbb{Z}_2) \to \mathrm{HF}_*(\phi^{m+n})$. Our first result is the following:

**Theorem 1.2** *Suppose that*

- *if $\gamma$ is nonseparating, then $\partial\Sigma \neq \varnothing$ or $\Sigma$ is closed with genus at least 2;*

- *if $\gamma$ is separating, then each component of $\Sigma - \gamma$ either contains a component of $\partial\Sigma$ or has genus at least 2.*

*Then for any pair of positive integers $m, n$, the product of $\mathrm{HF}(\phi^m)$ and $\mathrm{HF}(\phi^n)$ is the composition of*

$$\mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n) \xrightarrow{\mathrm{proj} \otimes \mathrm{proj}} H_*(\Sigma_0; \mathbb{Z}_2) \otimes H_*(\Sigma_0; \mathbb{Z}_2) \xrightarrow{\cap} H_*(\Sigma_0; \mathbb{Z}_2) \xrightarrow{\iota} \mathrm{HF}_*(\phi^{m+n}).$$

The counterpart for the coproduct structure is more involved. Let us first make the following remark on the decomposition (4). The $i^{\text{th}}$ component of $\bigoplus_{i=1}^{m-1} H_*(S^1)$ has an explicit description as follows: let $e_i^m$ (resp. $h_i^m$) be the elliptic (resp. hyperbolic) orbit of $\phi^m$ over the tubular coordinate $x = \frac{i}{m}$ ($i = 0, 1, \ldots, m$) that arises from perturbing the Morse–Bott degenerate $\phi$ to $\phi_0$ (see Section 2). Then $e_i^m$ and $h_i^m$ are cycles and the $i^{\text{th}}$ component of $\bigoplus_{i=1}^{m-1} H_*(S^1)$ is spanned by the two homology classes $[e_i^m]$ and $[h_i^m]$. Let us denote by $[e_j^n], [h_j^n], [e_k^{m+n}], [h_k^{m+n}]$ the homology classes appearing in the similar decomposition for $\mathrm{HF}_*(\phi^n)$ and $\mathrm{HF}_*(\phi^{m+n})$, respectively. Finally, let us recall that for any space $M$, a coproduct structure $\Delta_0$ on $H_*(M; \mathbb{Z}/2)$ is defined as the composition of $\mathrm{diag}_* \colon H_*(M; \mathbb{Z}/2) \to H_*(M \times M; \mathbb{Z}/2)$ and $H_*(M \times M; \mathbb{Z}/2) \cong H_*(M; \mathbb{Z}/2) \otimes H_*(M; \mathbb{Z}/2)$. Our second main result is the following:

**Theorem 1.3** *Suppose that*

- *if $\gamma$ is nonseparating, then $\partial\Sigma \neq \varnothing$ or $\Sigma$ is closed with genus at least 2;*

- *if $\gamma$ is separating, then each component of $\Sigma - \gamma$ either contains a component of $\partial\Sigma$ or has genus at least 2.*

*Then for any pair of positive integers $m, n$, the coproduct $\Delta \colon \mathrm{HF}_*(\phi^{m+n}) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n)$ described in the previous section is completely determined by the following:*

(1) *When restricted to $H_*(\Sigma_0; \mathbb{Z}_2) \subset \mathrm{HF}_*(\phi^{m+n})$, $\Delta$ is equal to*

$$H_*(\Sigma_0; \mathbb{Z}_2) \xrightarrow{\Delta_0} H_*(\Sigma_0; \mathbb{Z}_2) \otimes H_*(\Sigma_0; \mathbb{Z}_2) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n).$$

(2) *For each $[e_k^{m+n}] \in \bigoplus_{i=1}^{m+n-1} H_*(S^1)$,*

$$\Delta([e_k^{m+n}]) = \sum_{\substack{i \in \{0,1,\ldots,m\} \\ k-i \in \{0,1,\ldots,n\}}} [e_i^m] \otimes [e_{k-i}^n].$$

(3) *For each $[h_k^{m+n}] \in \bigoplus_{i=1}^{m+n-1} H_*(S^1)$,*

$$\Delta([h_k^{m+n}]) = \sum_{\substack{i \in \{0,1,\ldots,m\}, \\ k-i \in \{0,1,\ldots,n\}}} [e_i^m] \otimes [h_{k-i}^n] + [h_i^m] \otimes [e_{k-i}^n].$$

**Remark 1.4** It was asked in [6] how one could get an understanding of the ring structure of

$$\bigoplus_{n=0}^{\infty} \mathrm{HF}_*(\Sigma, \phi^n)$$

for general $\phi$. Our results compute the algebra and coalgebra structure of $\bigoplus_{n>0}^{\infty} \mathrm{HF}_*(\Sigma, \phi^n)$ when $\phi$ is the positive Dehn twist. One can use the computation in [21] to calculate the case for $n = 0$.

The coproduct structure[4] on $\bigoplus_{n\geq 0} \mathrm{HF}_*(\phi^n)$ can be interpreted as the dual of the product structure on the ring $\bigoplus_{n\leq 0} \mathrm{HF}_*(\phi^n)$ associated with the negative Dehn twist. To see this, we observe that the curves counted in computing the coproduct structure $\Delta : \mathrm{HF}_*(\phi^{m+n}) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n)$ are in one-to-one correspondence with the curves counted in the product structure $\mathrm{HF}^*(\phi^{-m}) \otimes \mathrm{HF}^*(\phi^{-n}) \to \mathrm{HF}^*(\phi^{-m-n})$, where $\mathrm{HF}^*(\phi^{-m})$ can be naturally viewed as the dual of $\mathrm{HF}_*(\phi^{-m})$ since we are using field coefficients. Likewise, the product structure on $\bigoplus_{n\geq 0} \mathrm{HF}_*(\phi^n)$ can be interpreted as the dual of the coproduct structure on $\bigoplus_{n\leq 0} \mathrm{HF}_*(\phi^n)$.

It will be clear from the proof of the main theorems that we can generalize Theorems 1.2 and 1.3 to the case where $\phi$ is the composition of disjoint positive Dehn twists, subject to certain topological conditions.

**Theorem 1.5** *Assume that $\Sigma$ is either closed with genus at least 2 or $\partial\Sigma \neq \varnothing$. Let $\phi$ be the composition of positive Dehn twists along disjoint simple closed curves $C_1, C_2, \ldots, C_k$, such that*

$$\Sigma_0 := \Sigma - (C_1 \cup C_2 \cup \cdots \cup C_k)$$

*is connected. Then the same result in Theorem 1.2 holds.*

**Theorem 1.6** *Let $\phi$ be the composition of positive Dehn twists along disjoint simple closed curves $C_1, C_2, \ldots, C_k$, such that $\Sigma_0 := \Sigma - (C_1 \cup C_2 \cup \cdots \cup C_k)$ is connected. Then the coproduct*

$$\Delta : \mathrm{HF}_*(\phi^{m+n}) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n)$$

*is completely determined by the following:*

(1) *When restricted to $H_*(\Sigma_0; \mathbb{Z}_2) \subset \mathrm{HF}_*(\phi^{m+n})$, $\Delta$ is equal to*

$$H_*(\Sigma_0; \mathbb{Z}_2) \xrightarrow{\Delta_0} H_*(\Sigma_0; \mathbb{Z}_2) \otimes H_*(\Sigma_0; \mathbb{Z}_2) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n).$$

(2) *For each circle $C_j$ and each $[e_{k,j}^{m+n}] \in \bigoplus_{i=1}^{m+n-1} H_*(S^1)$,*

$$\Delta([e_{k,j}^{m+n}]) = \sum_{\substack{i\in\{0,1,\ldots,m\}, \\ k-i\in\{0,1,\ldots,n\}}} [e_{i,j}^m] \otimes [e_{k-i,j}^n].$$

(3) *For each $[h_{k,j}^{m+n}] \in \bigoplus_{i=1}^{m+n-1} H_*(S^1)$,*

$$\Delta([h_{k,j}^{m+n}]) = \sum_{\substack{i\in\{0,1,\ldots,m\}, \\ k-i\in\{0,1,\ldots,n\}}} [e_{i,j}^m] \otimes [h_{k-i,j}^n] + [h_{i,j}^m] \otimes [e_{k-i,j}^n].$$

*In the above, the homology classes $[e_{i,j}^*]$ and $[h_{i,j}^*]$ correspond to the elliptic and hyperbolic orbits inside the twist region around circle $C_j$, over the tubular coordinate $x_j = i/*$.*

---

[4]We thank Tim Perutz for pointing out this fact to us.

## 1.4   Strategy of the proof

In this subsection, we summarize the key ideas behind the proof. As will be explained in more details in Section 2, the symplectic fiber bundle $X$ computing the product or the coproduct can be decomposed into two pieces, which we will call $X_D$ and $X_H$. Roughly speaking, $X_D$ is the union of fibers where iterations of a positive Dehn twist take place, and $X_H$ is (an open neighborhood of) the complement. The first key observation is that, under mild assumptions on the almost complex structure $J$, any $J$-holomorphic section of $X$ that has wrapping number (see Section 3) 0 must be completely contained in either $X_D$ or $X_H$. A key technical lemma used in the argument is the "local energy inequality" (Lemma 3.2) that was inspired by Lemma 3.11 of [12], which is reproved in our setting in Section 3. We next observe that for $J$-holomorphic sections, Fredholm index being 0 implies that the wrapping number is 0. Thus we only need to focus on sections contained in one of the two pieces.

$J$-holomorphic sections that are contained in $X_H$ are relatively easy to understand: by results of [17; 9; 14] (also known as the PSS isomorphism), these sections contribute to the intersection product or the coproduct of $H_*(\Sigma_0; \mathbb{Z}_2)$.

Sections that are contained in $X_D$ are more interesting. When computing the product structure, we are again able to rule out most of them using the local energy inequality. For the remaining sections, we use a translation trick (equation (13)) together with the PSS isomorphism to conclude that they contribute zero. For the coproduct structure, sections contained in $X_D$ do make contributions. To understand the contributions of these sections, we give a concrete description of the (unperturbed) moduli spaces, see Proposition 6.14, whose proof involves a deformation argument and a concrete construction. Finally, a Morse–Bott correspondence theorem (explained in Section 6.2) finishes the calculation.

**Remark 1.7**   For the computation of the coproduct structure, we expect that sections contained in the twist region $X_D$ could also be understood by results of [1; 4]. More precisely, we think that sections contained in $X_D$ could be viewed as part of curves counted in the product of the Floer homology of $T^*S^1$. By the results in [1; 4], the Chas–Sullivan loop product on the singular homology of the loop space of $S^1$ is expected to give us directly the coefficients in the coproduct structure. However, this method is not used in the current paper. Instead, we explicitly determine the Morse–Bott moduli spaces, with the expectation that such constructions will allow us to calculate similar cobordism maps for the periodic Floer homology in the future.

## 1.5   Directions for future work

**General pairs of symplectomorphisms**   While in this paper we computed the product and coproduct structures for iterations of a single positive Dehn twist, it is interesting to investigate the same question for $\phi$, $\psi$ Dehn twists along different circles. In general we expect this to be hard, probably requiring techniques beyond this paper.

For a different direction, one could also investigate the situation where $\phi$, $\psi$ are of finite type. The fixed point Floer homology for finite type maps was computed in [10]. It might be possible that the separating results in the current paper still holds under certain conditions, yielding a computation for the cobordism maps.

**Periodic Floer homology**  Fixed point Floer homology can be viewed as the degree $d = 1$ case of the periodic Floer homology (PFH), which is a more complicated Floer theory one can associate to a symplectomorphism of a surface. While the general construction of cobordism maps in PFH relies on Seiberg–Witten theory, special cases in which cobordism maps can be defined via counts of holomorphic curves or buildings have been worked out. For works in this direction, see for example [3; 19]. To generalize the results in the current paper, one could try to calculate the product and coproduct structures of PFH of a single Dehn twist over a surface. While getting a complete answer could be difficult, one could start by understanding the (multi)sections contained in the twist region. We hope that the detailed analysis presented in the current paper could give a hint on what the moduli spaces should look like.

## 1.6 Organization of the paper

The rest of the paper is organized as follows. In Section 2 we review the relevant geometric setup in detail. In particular, we give an explicit description of the symplectic fiber bundles $X_{m,n}$ and $X^{m,n}$ that are used to define the cobordism maps. In Section 3, we follow an idea of [12] to establish "no crossing" results for a special almost complex structure, Lemmas 3.11 and 3.13. In Section 4, we use an SFT compactness argument to prove Theorem 4.1, which is a generalization of the "no crossing" results in Section 3 for a general almost complex structure $J$. Finally, in Sections 5 and 6, we prove our main results, Theorems 1.2 and 1.3.

## Acknowledgements

## 2 The setup

We fix a symplectic surface, possibly with boundary, $(\Sigma, \omega_0)$. Let $\phi$ be any symplectomorphism. Let $Y_\phi$ be the mapping torus of $\phi$, let $\pi$ be the projection $Y_\phi \to S^1$, and let $\omega_\phi$ be the induced closed two-form.

The mapping torus has the structure of a stable Hamiltonian manifold, where the 1-form is $dt$, the 2-form is $\omega_\phi$ and the associated Reeb vector field is $R = \partial_t$. Closed integral curves of the Reeb vector field are called *Reeb orbits*, and they are called *nondegenerate* if the linearized first return map does not have 1 as an eigenvalue. We call a nondegenerate Reeb orbit *hyperbolic* if the eigenvalues are real, and *elliptic* otherwise. In this paper, we will be mainly interested in Reeb orbits of degree 1, that is, those who cover once under the projection map $\pi$.

We fix a homologically nontrivial simple closed curve $\gamma \subset \Sigma$. As mentioned in the introduction, we choose a tubular neighborhood $N$ of $\gamma$ with coordinates $x \in (-\epsilon, 1 + \epsilon)$ and $y \in S^1 = \mathbb{R}/\mathbb{Z}$, and $\omega_0 = dx \wedge dy$. The *(unperturbed) positive Dehn twist* along $\gamma$ is a symplectomorphism of $\Sigma$, which has the form

$$\phi_0 \colon (x, y) \mapsto (x, y - x)$$

inside $N$, and is the time-1 map of a Hamiltonian $H_0$ outside of $N' = [\epsilon, 1 - \epsilon] \times S^1 \subset N$. We require $H_0$ takes the following form on $N \setminus N'$:

(1)   $H_0(x, y) = \frac{1}{2}x^2$ in $(-\epsilon, \epsilon) \times S^1 \subset N$, and

(2)   $H_0(x, y) = \frac{1}{2}(x - 1)^2$ in $(1 - \epsilon, 1 + \epsilon) \times S^1 \subset N$.

We assume on $\Sigma \setminus N$, the function $H_0$ is a $C^2$ small Morse function, so that the associated time-1 map is nondegenerate. We further assume that near each boundary component of $\Sigma$, there are tubular coordinates $x_i \in (-\epsilon_i, 0]$, $y_i \in S^1$ and a small real number $\theta_i$ such that $H_0(x_i, y_i) = \theta_i x_i$.

We note the unperturbed positive Dehn twist is nondegenerate, except for the Morse–Bott $S^1$ family of periodic orbits corresponding to $x = 0$ and $x = 1$. We shall later consider iterations of $\phi_0$, which we denote by $\phi_0^n$. By the above, on $N$, the map $\phi_0^n$ takes

$$(x, y) \mapsto (x, y - nx)$$

on $N$ and looks like the time-1 map of $nH_0$ outside of $N'$. We assume both $nH_0$ (in $C^2$ norm) and $n\theta_i$ are small.

We note that in order to define the fixed-point symplectic homology, we need the symplectomorphisms to be nondegenerate (equivalently, that the Reeb orbits are cut out transversely). Since the symplectomorphism $\phi_0^n$ on $\Sigma - N$ is the time-1 map of a Hamiltonian $nH_0$, this is achieved outside of $N$ by requiring that $H_0$ be a $C^2$-small Morse function. Inside the tubular region $N$, Reeb orbits come in Morse–Bott $S^1$ families. Following [12], we overcome this technical difficulty by perturbing $\phi_0$ (in a small neighborhood of finitely many values of $x$ over which Reeb orbits exist) in a Hamiltonian way, which amounts to adding a Hamiltonian perturbation term. For example, near $x = 0$, we can modify $H_0$ to be $\left(\frac{1}{2}x^2 + \lambda(x)h(y)\right)$, where $\lambda(x)$ is a cutoff function supported in $(-\delta, \delta)_x$ with $\lambda(0) = 1$ as a nondegenerate local max, and $h \colon S_y^1 \to \mathbb{R}$ is a small perfect Morse function. We perform this kind of perturbation for each $S^1$ family of fixed points in $N$. We always assume that the Hamiltonian perturbation only takes place in the union of all intervals $(x_i - \delta, x_i + \delta)$ (where $x_i$'s are the $x$-coordinates for all possible Morse–Bott $S^1$-families)

for some positive real number $\delta$ much smaller than $\epsilon$ (we'll later call the complement of these intervals the *unperturbed range*). Once this is done, viewed from the perspective of the mapping torus $Y_{\phi_0^n}$, the $S^1$ family of fixed points at $x = \frac{i}{n}$ become perturbed to a pair of Reeb orbits (one elliptic and one hyperbolic).

With the above *perturbed positive Dehn twist*, which we denote by $\phi^n$, we can define its fixed point Floer homology $\mathrm{HF}(\Sigma, \phi^n)$ after we pick a generic $\phi^n$ compatible almost complex structure $J$ on $Y_{\phi^n} \times \mathbb{R}$. We next describe the symplectic fiber bundle that allows us to define product and coproduct structures on $\mathrm{HF}(\Sigma, \phi^n)$. We first describe the construction for the unperturbed positive Dehn twists, then perturb to break the Morse–Bott degeneracy. The reason we describe the Morse–Bott situation in detail is because for the coproduct computation we will be enumerating $J$-holomorphic sections in the Morse–Bott setting, then we will use Morse–Bott theory to convert that to enumerations of $J$-holomorphic sections in the nondegenerate setting.

Recall that, given two symplectomorphisms, there is a symplectic fiber bundle $(X, \pi_X, \omega_X)$ over the thrice punctured sphere $B_0$, which is modelled by the symplectizations of mapping tori over the punctures. We now describe in more details what the bundle $X_{m,n}$ used in computing the product structure

$$\bullet \colon \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n) \to \mathrm{HF}_*(\phi^{m+n})$$

looks like. The description for the bundle $X^{m,n}$ used to compute the coproduct structure is almost identical, and we will mention at the end of this section what changes need to be made.

We designate two of the punctures of $B_0$ as "positive", and the other as "negative". Choose local conformal coordinates $s_i \in [0, \infty)$ and $t_i \in S^1 (i = 1, 2)$ near the 2 positive punctures of $B_0$, and local conformal coordinates $s_{-\infty} \in (-\infty, 0]$, $t_{-\infty} \in S^1$ near the negative puncture. Fix also a smooth map[5] $g_{m,n} \colon B_0 \to S^1$ such that $dg_{m,n} = m \, dt_1$ near the first positive puncture, $dg_{m,n} = n \, dt_2$ near the second positive puncture and $dg_{m,n} = (m + n) dt_{-\infty}$ near the negative puncture. We further assmume that $g_{m,n} = d \cdot g'_{m,n}$ where $d = gcd(m, n)$ (the primitive $g'_{m,n}$ will be used in the proof of Theorem 1.2). Define the closed one-form $\beta_{m,n} = dg_{m,n}$.

Let $\Sigma_0 = \Sigma - N'$. We now describe the fiberwise symplectic cobordism $X_{m,n}$ as the union of two fiberwise symplectic cobordisms $X_D$ and $X_H$ as follows. Topologically, $X_D = B_0 \times N$ and $X_H = B_0 \times \Sigma_0$. In order to describe $\phi_0$ as the time-1 map of the Hamiltonian $H_0$ near the two ends of the tubular region $N$ we choose coordinates $(p, x_L, y_L) \in B_0 \times (-\epsilon, \epsilon) \times S^1$ and $(p, x_R, y_R) \in B_0 \times (1 - \epsilon, 1 + \epsilon) \times S^1$ for the two ends of $B_0 \times (N - N') \subset X_H$ and impose that

(1) $\omega_0 = dx_L \wedge dy_L$ or $dx_R \wedge dy_R$ in the two components of $N - N'$,

(2) $H_0(x_L, y_L) = \frac{1}{2} x_L^2$ in $(-\epsilon, \epsilon) \times S^1 \subset N$, and

(3) $H_0(x_R, y_R) = \frac{1}{2}(x_R - 1)^2$ in $(1 - \epsilon, 1 + \epsilon) \times S^1 \subset N$.

---

[5]To see such a map exists, we can choose a degree $m + n$ branched covering map $\varphi \colon B_0 \to \mathbb{R} \times S^1$ such that near the three punctures $\varphi$ has the standard local form $(s, t_*) \mapsto (s, k t_*)$ (where $k = m, n, m + n$, respectively), and let $g$ be $\varphi$ followed by the projection map to $S^1$.

Topologically, the 4-manifold $X$ is defined to be $X = X_H \cup X_D / \sim$, where we identify points

$$(p, x, y) \in B_0 \times (-\epsilon, \epsilon)_x \times S^1_y \subset X_D \quad \text{with} \quad (p, x, y) \in B_0 \times (-\epsilon, \epsilon)_{x_L} \times S^1_{y_L} \subset X_H$$

and

$$(p, x, y) \in B_0 \times (1-\epsilon, 1+\epsilon)_x \times S^1_y \subset X_D \quad \text{with} \quad (p, x, y + g_{m,n}(p)) \in B_0 \times (1-\epsilon, 1+\epsilon)_{x_R} \times S^1_{y_R} \subset X_H.$$

We define the fiberwise symplectic 2-form $\omega_{X,0}$ to be $dx \wedge dy + d(\frac{1}{2}x^2 \beta_{m,n})$ in $X_D$, and $\omega_0 + d(H_0 \beta_{m,n})$ in $X_H$. It is easy to see that the two definitions agree in

$$X_D \supset B_0 \times (-\epsilon, \epsilon)_x \times S^1_y = B_0 \times (-\epsilon, \epsilon)_{x_L} \times S^1_{y_L} \subset X_H.$$

To see that the two definitions agree in $X_D \supset B_0 \times (1-\epsilon, 1+\epsilon)_x \times S^1_y = B_0 \times (1-\epsilon, 1+\epsilon)_{x_R} \times S^1_{y_R} \subset X_H$, we calculate

$$
\begin{aligned}
dx_R \wedge dy_R + d\left(\tfrac{1}{2}(x_R - 1)^2 \beta_{m,n}\right) &= dx \wedge (dy + \beta_{m,n}) + d\left(\tfrac{1}{2}(x-1)^2 \beta_{m,n}\right) \\
&= dx \wedge dy + dx \wedge \beta_{m,n} + (x-1)dx \wedge \beta_{m,n} \\
&= dx \wedge dy + d\left(\tfrac{1}{2}x^2 \beta_{m,n}\right).
\end{aligned}
$$

We remark that over the positive punctures, the symplectic fiber bundle defined above are isomorphic to $[0, \infty)$ times the mapping tori $Y_{\phi_0^m}$, $Y_{\phi_0^n}$, and over the negative puncture, the above fiber bundle is modelled by $(-\infty, 0]$ times the mapping torus $Y_{\phi_0^{m+n}}$.

The above fiber bundle have Morse–Bott degeneracies in its Reeb orbits at each of its punctures. To arrive at the definition of $X_{m,n}$, we perturb the Reeb orbits to be nondegenerate as before. Since we are working in the language of a symplectic fiber bundle, we achieve this by adding a Hamiltonian perturbation term to $\omega_{X,0} = dx \wedge dy + d(\frac{1}{2}x^2 \beta_{m,n})$.

As before, near $x = 0$, we can modify $\omega_{X,0}$ to be $\omega_X = dx \wedge dy + d\left(\left(\frac{1}{2}x^2 + \lambda(x)h(y)\right)\beta_{m,n}\right)$, where $\lambda(x)$ is a cutoff function supported in $(-\delta, \delta)_x$ with $\lambda(0) = 1$ as a nondegenerate local max, and $h \colon S^1_y \to \mathbb{R}$ is a small perfect Morse function. We assume as before that the Hamiltonian perturbation only takes place in the union of all intervals $(x_i - \delta, x_i + \delta)$ (where $x_i$'s are the $x$-coordinates for all possible Morse–Bott $S^1$-families) for some positive real number $\delta$ much smaller than $\epsilon$ (we'll later call the complement of these intervals the *unperturbed range*). Once this is done, near the first (resp. the second) positive puncture, $x = \frac{i}{m}$ (resp. $x = \frac{j}{n}$) each correspond to a pair of Reeb orbits (one elliptic and one hyperbolic), and near the negative puncture $x = \frac{k}{m+n}$ each correspond to a pair of Reeb orbits (one elliptic and one hyperbolic).

This defines the fiberwise symplectic 2-form $\omega_X$, and we now describe the symplectic structure on $X_{m,n}$. It's illustrated in eg [3] that from the fiberwise symplectic cobordism $(X, \pi_X, \omega_X)$ one can construct a symplectic form $\Omega_X = \omega_X + K\pi_X^* \omega_{B_0}$, where $K$ is a large positive number, and $\omega_{B_0}$ is an area form on $B_0$. Without loss of generality, we assume from now on that $K\pi_X^* \omega_{B_0} = ds_i \wedge dt_i$ near the punctures.

This concludes the definition of $X_{m,n}$, we now equip it with a tame almost complex structure $J$.

**Definition 2.1** An almost complex structure on $(X_{m,n}, \omega_X, \pi)$ is called tame if the following conditions are satisfied:

(1) Near the punctures of $B_0$ where the symplectic fiber bundle is isomorphic to $Y_{\phi^n} \times [0, \infty)$ (resp. $Y_{\phi^m} \times [0, \infty)$ or $Y_{\phi^{m+n}} \times (-\infty, 0]$), the almost complex structure is given by the restriction of a $\phi^n$ (resp. $\phi^m$, or $\phi^{m+n}$) compatible almost complex structure.

(2) Away from the cylindrical neighborhoods around the punctures of $B_0$, the almost complex structure $J$ is tamed by the symplectic form $\Omega_X$.

Then for generic tame $J$, if $x$, $y$, $z$ are fixed points of $\phi^n$, $\phi^m$ and $\phi^{m+n}$ respectively, under suitable topological assumptions (eg when the bundle $X_{m,n}$ is weakly monotone; see Section 5.1) the moduli space $\mathcal{M}^J_{x,y;z}$, defined by

$$\mathcal{M}^J_{x,y;z} = \left\{ u \colon B_0 \to X \,\middle|\, \begin{array}{l} \pi_X \circ u = \mathrm{id}, \ u \text{ is } J\text{- holomorphic, and} \\ u \text{ is asymptotic to } \gamma_x, \gamma_y \text{ and } \gamma_z \text{ over the} \\ \text{three appropriate punctures.} \end{array} \right\}$$

is a manifold whose dimension is given by the Fredholm index formula

$$\mathrm{ind}(u) = 1 + 2\langle c_1^\tau(TX_{m,n}), [u]\rangle + \mathrm{CZ}_\tau(\gamma_x) + \mathrm{CZ}_\tau(\gamma_y) - \mathrm{CZ}_\tau(\gamma_z).$$

Here $\tau$ denotes a choice of fixed trivializations around each Reeb orbit, and $\mathrm{CZ}_\tau$ denotes the Conley–Zehnder indices of Reeb orbits with respect to this trivialization. Similarly the relative first Chern class $c_1^\tau$ is also determined by this choice of trivialization. See Section 5 for our specific choices of $\tau$. The product on the chain level is now defined as

$$(5) \qquad\qquad \langle x \bullet y, z \rangle = \#_{\mathbb{Z}/2} \mathcal{M}^J_{x,y;z},$$

where $\#_{\mathbb{Z}/2} \mathcal{M}^J_{x,y;z}$ denotes the mod 2 count of Fredholm index 0 sections (we will explain in Section 5.1 that the monotonicity condition ensures that the moduli space is compact). See Section 5 for the details of this computation.

For the coproduct structure, the symplectic fiber bundle $X^{m,n}$ is defined almost verbatim, so we only highlight the minor changes that need to be made. We again begin with the thrice punctured sphere $B_0$, but this time choose one of the three punctures as the "positive puncture" with a local conformal coordinate $(s_\infty, t_\infty) \in [0, \infty) \times S^1$, and choose the other two punctures as the two "negative" punctures with local coordinates $(s_i, t_i) \in (-\infty, 0] \times S^1$. Fix a smooth function $g^{m,n} \colon B_0 \to S^1$ such that $dg^{m,n} = (m+n)dt_\infty$ near the positive puncture and $dg^{m,n} = mdt_1$ and $ndt_2$ near the two negative punctures respectively. We again define $X^{m,n}$ by gluing two trivial fiber bundles $X_H = B_0 \times \Sigma_0$ and $X_D = B_0 \times N$, but this time the gluing map for the right side of $X_D$ is

$$B_0 \times (1-\epsilon, 1+\epsilon)_x \times S^1_y \ni (p, x, y) \sim (p, x, y + g^{m,n}(p)) \in B_0 \times (1-\epsilon, 1+\epsilon)_{x_R} \times S^1_{y_R}.$$

Let $\beta^{m,n} = dg^{m,n}$. We similarly define the (unperturbed) fiberwise symplectic 2-form $\omega_{X,0}$ to be $dx \wedge dy + d\left(\frac{1}{2}x^2\beta_{m,n}\right)$ in $X_D$ and $\omega_0 + d(H_0\beta_{m,n})$ in $X_H$. As before, we perturb $\omega_{X,0}$ to be $\omega_X$ in order to break the Morse–Bott degeneracy, and we always assume that such perturbation is supported in a $\delta$ neighborhood of Reeb orbits inside $X_D$.

We also define the symplectic form $\Omega_X$ on $X^{m,n}$, and the notion of *tame* almost complex structures with respect to $\Omega_X$. Then the coproduct is defined by considering the moduli space of $J$-holomorphic sections $\mathcal{M}^J_{z;x,y}$ where $z$ is a fixed point of $\phi^{m+n}$ we think of as the input, and $x$ and $y$ are fixed points of $\phi^n$ and $\phi^m$ we think of as outputs. For generic $J$ this moduli space is a manifold and the coproduct is defined by the mod 2 count of Fredholm index 0 $J$-holomorphic sections.

# 3 The "no crossing" results for unperturbed $J$

Throughout this section, the symplectic fiber bundle $X$ refers to either $X_{m,n}$ or $X^{m,n}$. As explained in the introduction, to define the cobordism map, we count $J$-holomorphic sections that asymptote to appropriate Reeb orbits. We'll prove some key properties about such $J$-holomorphic sections for some particularly nice almost complex structures. Before doing that, let us introduce some terminologies. Following [15], the *vertical distribution* Ver is the kernel of $d\pi_X : TX \to TB_0$. The *horizontal distribution* Hor is defined as $\mathrm{Hor}_x := \{u \in T_x X \mid \omega_X(u,v) = 0 \ \forall v \in \mathrm{Ver}_x\}$.

**Definition 3.1** [15, Definition 8.2.6]  An almost complex structure on $(X, \omega_X)$ is called fibration-compatible if the following holds:

(1)  The projection $\pi_X$ is holomorphic: $d\pi \circ J = j_0 \circ d\pi$.

(2)  For every $p \in B_0$, the restriction $J_p$ of $J$ to $\pi_X^{-1}(p)$ is tamed by $\omega_X|_{\pi_X^{-1}(p)}$.

(3)  The horizontal distribution Hor is preserved by $J$.

Note that by definition, there is a one-to-one correspondence between fibration-compatible almost complex structures and $\omega_X$-tame almost complex structures on the vertical distribution.

In this section we only consider fibration-compatible almost complex structures. For a fibration-compatible $J$, all the horizontal sections are $J$-holomorphic (a section $u : B_0 \to X$ is horizontal, if $du(TB_0) \subset \mathrm{Hor}$).

Following [12, Lemma 3.11], we now establish a local energy inequality for $J$-holomorphic sections. To state the inequality, for any $x \in (-\epsilon, 1+\epsilon)$ we let $F_x$ denote the 3-manifold $B_0 \times \{x\} \times S^1_y \subset X_D$. Likewise, let $F_{[x_1,x_2]}$ denote the 4-manifold $B_0 \times [x_1,x_2]_x \times S^1_y \subset X_D$. The first homology group of $X_D = B_0 \times (-\epsilon, 1+\epsilon)_x \times S^1_y$ is $\mathbb{Z}^3$, generated by $[S^1_{t_1}]$, $[S^1_{t_2}]$ and $[S^1_y]$. We identify

$$p[S^1_y] + q_1[S^1_{t_1}] + q_2[S^1_{t_2}] \in H_1(X_D)$$

with a tuple $(p, q_1, q_2) \in \mathbb{Z}^3$.

**Lemma 3.2** (local energy inequality) *Let $C$ be a $J$-holomorphic section $u\colon B_0 \to X$ which is not horizontal. Assume that $C$ intersects $F_x$ transversely and that $C \cap F_x \neq \varnothing$ for some $x$ in the unperturbed range. Orient each circle in $C \cap F_x$ using the boundary orientation of $C \cap F_{[x-\epsilon', x]}$ (for a small $\epsilon'$) induced by $j_0$. Under this orientation, let $(p, q_1, q_2)$ denote the homology class of $C \cap F_x$, then we have*

$$(6) \qquad\qquad p + x(mq_1 + nq_2) > 0.$$

Before proving the lemma, here are some observations on the vertical energy of $J$-holomorphic sections.

**Definition 3.3** Let $u\colon B_0 \to X$ be a smooth section of the bundle $X \to B_0$, $J$ be an almost complex structure and $g_J$ the metric induced by $\omega_X$ and $J$. Let $(s, t)$ be a local conformal coordinate on $B_0$. The *vertical energy* of $u$ is defined to be

$$E(u) = \frac{1}{2} \int_{B_0} |\partial_s u - \partial_s^\#|_{g_J}^2 + |\partial_t u - \partial_t^\#|_{g_J}^2 \, ds \wedge dt$$

where $\partial_s^\#$ and $\partial_t^\#$ are the horizontal lifts of the vector fields $\partial_s$ and $\partial_t$, respectively.

**Remark 3.4** Our definition, written in local conformal coordinates, coincides with that of [15, equation 8.1.8]. It's also clear from the definition that a smooth map has zero vertical energy if and only if it is a horizontal section. Later in Definition 6.7 we will generalize the notion of vertical energy to include more examples of symplectic fiber bundles that will be useful later. For now, the main observation is the following:

**Lemma 3.5** *Let $u$ be a $J$-holomorphic section of $(X, \omega_X, \pi_X)$ described in Section 2. Let $(s, t)$ be a local conformal coordinate on $B_0$ and $\partial_s^\#$ and $\partial_t^\#$ be the horizontal lifts of the vector fields $\partial_s$ and $\partial_t$. The two-form*

$$\tfrac{1}{2}(|\partial_s u - \partial_s^\#|_{g_J}^2 + |\partial_t u - \partial_t^\#|_{g_J}^2) \, ds \wedge dt$$

*can be rewritten as*

$$u^* \omega_X - \omega_X(\partial_s^\#, \partial_t^\#) \, ds \wedge dt,$$

*which is equal to $u^* \omega_X$ in $X_D$ minus the perturbed region.*

**Proof** If $u$ is $J$-holomorphic, we have

$$\begin{aligned}
\tfrac{1}{2}(|\partial_s u - \partial_s^\#|_{g_J}^2 + |\partial_t u - \partial_t^\#|_{g_J}^2) \, ds \wedge dt &= \omega_X(\partial_s u - \partial_s^\#, \partial_t u - \partial_t^\#) \, ds \wedge dt \\
&= \omega_X(\partial_s u, \partial_t u) - \omega_X(\partial_s^\#, \partial_t^\#) \, ds \wedge dt \\
&= u^* \omega_X - \omega_X(\partial_s^\#, \partial_t^\#) \, ds \wedge dt.
\end{aligned}$$

If we write the two-form $\omega_X$ as $dx \wedge dy + F(x, y, s, t)ds + G(x, y, s, t)dt$ inside $X_D$, then the term $\omega_X(\partial_s^\#, \partial_t^\#)$ is equal to

$$\frac{\partial G}{\partial s} - \frac{\partial F}{\partial t} + \frac{\partial G}{\partial x}\frac{\partial F}{\partial y} - \frac{\partial G}{\partial y}\frac{\partial F}{\partial x},$$

which is equal to zero outside of the perturbed region, where

$$\frac{\partial F}{\partial y} = \frac{\partial G}{\partial y} = 0 \quad \text{and} \quad \frac{\partial G}{\partial s} = \frac{\partial F}{\partial t}$$

by our construction. □

It follows that if $u$ is a $J$-holomorphic section of $X$, then the part of $u$ in $X_D$ minus the perturbed region satisfies

(7) $$u^* \omega_X(v, j_0 v) \geq 0$$

for any $v \in TB_0$, and the equality holds if and only if $du(v) \in \text{Hor}$. Now we are ready to prove Lemma 3.2.

**Proof of Lemma 3.2** Choose $x_1 < x < x_2$ such that

(1) $C$ intersects both $F_{x_1}$ and $F_{x_2}$ transversely,

(2) $x - x_1 = x_2 - x$, and

(3) $[x_1, x_2]$ is contained in the unperturbed range.

We orient $C \cap F_{x_1}$ and $C \cap F_{x_2}$ in the same way as stated in the lemma, so we have $\partial(C \cap F_{[x_1, x_2]}) = C \cap F_{x_2} - C \cap F_{x_1}$. Also notice that with the specified orientations, $C \cap F_{x_1}$, $C \cap F_{x_2}$ and $C \cap F_x$ all have the same homology class in $H_1(X_D)$. Now we note that $C \cap F_{[x_1, x_2]}$ is not horizontal, otherwise $C$ has to be horizontal everywhere by unique continuation (recall that all horizontal sections are $J$-holomorphic).

Using the inequality (7), we have (in the following the one-form $\beta$ refers to either $\beta_{m,n}$ or $\beta^{m,n}$)

$$0 < \int_{C \cap F_{[x_1, x_2]}} u^* \omega_X = \int_{C \cap F_{[x_1, x_2]}} u^* \big(dx \wedge dy + d\big(\tfrac{1}{2}x^2 \beta\big)\big)$$

$$= \int_{\partial(C \cap F_{[x_1, x_2]})} u^* \big(x \, dy + \tfrac{1}{2}x^2 \beta\big)$$

$$= x_2 \int_{C \cap F_{x_2}} u^* \, dy + \tfrac{1}{2}x_2^2 \int_{C \cap F_{x_2}} u^* \beta - x_1 \int_{C \cap F_{x_1}} u^* \, dy - \tfrac{1}{2}x_1^2 \int_{C \cap F_{x_1}} u^* \beta$$

$$= x_2 p + \tfrac{1}{2}x_2^2(mq_1 + nq_2) - x_1 p - \tfrac{1}{2}x_1^2(mq_1 + nq_2)$$

$$= (x_2 - x_1)(p + x(mq_1 + nq_2)). \qquad \square$$

**Remark 3.6** It is clear from the proof that if $u$ is a horizontal section, then the equality

$$p + x(mq_1 + nq_2) = 0$$

holds.

Let's assume for the moment that $\Sigma_0 = \Sigma - N'$ is connected. Following [12], we define the *wrapping number* of $J$-holomorphic sections with cylindrical asymptotes:

**Definition 3.7** The wrapping number of a $J$-holomorphic section $C$ with cylindrical asymptotes is $\eta(C) = \#C \cap (B_0 \times \{P_0\})$, where $P_0 \in \Sigma_0$ is not a critical point of $H_0$.

The algebraic intersection number does not depend on the choice of $P_0$, so the wrapping number is well-defined. We also note that $\eta(C)$ is identically zero if $\partial\Sigma \neq \varnothing$ and under the additional assumption that near each boundary component, $J$ is induced from the vertical almost complex structure that sends $\partial_{x_i}$ to $\partial_{y_i}$. The reason is that once such a $J$ is chosen, no $J$-holomorphic sections can enter the boundary region by the following maximum principle, so one can choose $P_0$ inside one of the boundary region and easily see that $C \cap (B_0 \times \{P_0\}) = \varnothing$.

**Lemma 3.8** (maximum principle) *Let $J$ be a fibration-compatible almost complex structure on $X$ that sends $\partial_{x_i}$ to $\partial_{y_i}$ near each boundary component of $\Sigma$. Here the index $i$ labels the different boundary components of $\Sigma$. Let $V$ denote an open subset of $B_0$ with local coordinates $(s, t)$. Let $\tilde{u}: V \to X_H$ denote a $J$-holomorphic section, which in coordinates look like $(s, t) \mapsto (s, t, x_i(s, t), y_i(s, t))$. We further assume for $(s, t) \in V$, the pair $(x_i(s, t), y_i(s, t))$ is in a neighborhood of the $i^{th}$ boundary component of $\Sigma$. Then $x_i(s, t)$ is a harmonic function.*

**Proof** The setup is almost identical to that of Lemma 6.10, except that the Hamiltonian function is $\theta_i x_i$ instead of $x_i^2/2$. In particular, the horizontal lifts are

$$\partial_s^\# = \partial_s - \theta_i F \partial_y, \quad \partial_t^\# = \partial_t - \theta_i G \partial_y$$

and we have a similar equation,

$$\frac{\partial x_i}{\partial t} + \frac{\partial y_i}{\partial s} + \theta_i F = 0, \quad \frac{\partial y_i}{\partial t} - \frac{\partial x_i}{\partial s} + \theta_i G = 0.$$

Notice that $\beta = F(s, t)ds + G(s, t)dt$ being closed tells us that $\frac{\partial F}{\partial t} = \frac{\partial G}{\partial s}$, so the conclusion follows by a simple calculation. $\qquad\square$

In particular, the above lemma implies that as long as $J$ is chosen in a neighborhood of each of the boundary components of $\Sigma$ to be fibration compatible and sends $\partial_{x_i}$ to $\partial_{y_i}$, no $J$-holomorphic section may approach the boundary components of $\Sigma$.

**Remark 3.9** If $\Sigma_0 = \Sigma - N'$ is not connected, ie $\gamma$ is separating, we can define two wrapping numbers $\eta_1$ and $\eta_2$ for each of the connected components of $\Sigma_0$. It is clear from the above arguments that if each connected components of $\Sigma_0$ contains part of $\partial\Sigma$, then all the wrapping numbers vanish automatically.

**Remark 3.10** Following a similar idea in [12, Lemma 4.3], we will show in the following (Remark 3.12) that the wrapping numbers of any $J$-holomorphic sections are nonnegative.

The first main result of this section is the following "no crossing" lemma:

**Lemma 3.11** *Assume $J$ is a fibration-compatible almost complex structure on $X_{m,n}$. If $C$ is a J-holomorphic section of $X_{m,n}$ such that all wrapping numbers are zero, then $C$ is either contained in $X_H$ or contained in $X_D$.*

**Proof** We first consider the case where $C$ is not horizontal. Suppose there is a nonhorizontal $J$-holomorphic section $C$ that is neither contained in $X_H$ nor in $X_D$, then since $C$ is connected, we can either find some $\epsilon_1 \in (\delta, \epsilon)$ such that $C$ intersects both $F_{\epsilon_1}$ and $F_{-\epsilon_1}$ transversely and $C \cap F_{\pm\epsilon_1} \neq \varnothing$, or some $\epsilon_1 \in (\delta, \epsilon)$ such that $C$ intersects both $F_{1+\epsilon_1}$ and $F_{1-\epsilon_1}$ transversely and $C \cap F_{1\pm\epsilon_1} \neq \varnothing$. Without loss of generality, let us assume the first situation happens. (If the second situation happens, the following proof works almost verbatim; the only change one needs to make is that, if we use $(p^\pm, q_1^\pm, q_2^\pm)$ to denote the homology classes of $C \cap F_{1\pm\epsilon_1}$, then the condition $\eta = 0$ translates to $p^\pm + mq_1^\pm + nq_2^\pm = 0$.)

Let $(p^\pm, q_1^\pm, q_2^\pm)$ denote the homology classes of $C \cap F_{\pm\epsilon_1}$. Since all the wrapping numbers vanish, we observe that $p^\pm = 0$. To see this fact, notice that we can choose $P_0$ to be $(\pm\epsilon_1, y_0) \in \Sigma_0$ for some fixed $y_0 \in S^1$, then $\#C \cap (B_0 \times \{P_0\})$ is precisely the number of times $C \cap F_{\pm\epsilon_1}$ passes through $y_0$, which equals $p^\pm$.

Lemma 3.2 tells us that

$$\epsilon_1(mq_1^+ + nq_2^+) > 0 > \epsilon_1(mq_1^- + nq_2^-)$$

which implies that

$$mq_1^+ + nq_2^+ \geq 1, \quad mq_1^- + nq_2^- \leq -1.$$

Now we consider $C_{[-\epsilon_1, \epsilon_1]} := C \cap F_{[-\epsilon_1, \epsilon_1]}$ which is a surface with boundary $C \cap F_{\epsilon_1} - C \cap F_{-\epsilon_1}$, possibly with positive and negative punctures at $x = 0$. Let $d_1$(resp. $d_2, d_{-\infty}) \in \{0, 1\}$ denote the number of punctures of $C_{[-\epsilon_1, \epsilon_1]}$ that project to the first positive puncture (resp. the second positive puncture, the negative puncture) of $B_0$. Notice that the two Reeb orbits[6] over $x = 0$ at the first positive puncture (resp. the second puncture, the negative puncture) have the homology class $(0, 1, 0)$ (resp. $(0, 0, 1)$ and $(0, 1, 1)$), so we have

$$d_1(0, 1, 0) + d_2(0, 0, 1) + (0, q_1^+, q_2^+) = d_{-\infty}(0, 1, 1) + (0, q_1^-, q_2^-)$$

and hence

$$2 \leq (mq_1^+ + nq_2^+) - (mq_1^- + nq_2^-) = (m + n)d_{-\infty} - md_1 - nd_2$$

which implies that

$$d_{-\infty} = 1.$$

The above equation implies that $C$ has no other negative punctures. So for any $\epsilon_2 \in (\delta, \epsilon)$, the section $C$ cannot intersect both $F_{1-\epsilon_2}$ and $F_{1+\epsilon_2}$, because otherwise, the same argument as above would imply that $C$ has another negative puncture asymptotic to one of the Reeb orbits over $x = 1$, a contradiction. So there are two remaining possibilities:

---

[6]Recall that the unperturbed Reeb vector fields over $x$ are $\partial_{t_1} - mx\partial_y$, $\partial_{t_2} - nx\partial_y$ and $\partial_{t_{-\infty}} - (m+n)x\partial_y$ near the three punctures.

(1) $C \cap F_{1-\epsilon_2} = \varnothing$   Let us consider $C \cap F_{[\epsilon_1, 1-\epsilon_2]}$. For this part of $C$, there are no negative punctures or positive punctures, so we conclude that $\partial(C \cap F_{[\epsilon_1, 1-\epsilon_2]}) = -C \cap F_{\epsilon_1}$ is null homologous in $H_1(X_D)$, which contradicts the fact that $mq_1^+ + nq_2^+ \geq 1$.

(2) $C \cap F_{1+\epsilon_2} = \varnothing$   Let us consider $C \cap F_{[\epsilon_1, 1+\epsilon_2]}$. For this part of $C$, let $a, b \in \{0, 1\}$ denote the number of punctures $C$ has at $x = 1$ that project to the first and second positive puncture, respectively, of $B_0$. Observe that the Reeb orbits at $x = 1$ near the first (resp. second) positive puncture have the homology class $(-1, 1, 0) \in H_1(X_D)$ (resp. $(-1, 0, 1) \in H_1(X_D)$), so we conclude that

$$a(-m, 1, 0) + b(-n, 0, 1) = (0, q_1^+, q_2^+),$$

which in turn implies that $a = b = 0$. But then it follows that $q_1^+ = q_2^+ = 0$, contradicting $mq_1^- + nq_2^- \leq -1$.

Finally, we consider the case where $C$ is horizontal. Suppose there exist such horizontal section $C$ that is neither contained in $X_D$ nor in $X_H$, then again without loss of generality we can assume that there exist some $\epsilon_1 \in (\delta, \epsilon)$ such that $C$ intersects both $F_{\epsilon_1}$ and $F_{-\epsilon_1}$ transversely and $C \cap F_{\pm\epsilon_1} \neq \varnothing$. Recall that inside $X_D$ apart from the perturbed region, $\omega_X = dx \wedge dy + d(\frac{1}{2}x^2 \beta_{m,n})$. We show in the following that outside of the perturbed region, the $x$-coordinate of the section contained in $X_D$ is locally constant, which obviously leads to a contradiction.

To see this fact, we write the one-form $\frac{1}{2}x^2 \beta_{m,n}$ as

$$\tfrac{1}{2}x^2 \beta_{m,n} = f\, ds + g\, dt,$$

where $(s, t)$ is the local coordinate for $B_0$. We next compute that the horizontal lifts $\partial_s^{\#}, \partial_t^{\#}$ of the two vector fields $\partial_s, \partial_t$ are

$$\partial_s^{\#} = \partial_s + \frac{\partial f}{\partial y}\partial_x - \frac{\partial f}{\partial x}\partial_y, \quad \partial_t^{\#} = \partial_t + \frac{\partial g}{\partial y}\partial_x - \frac{\partial g}{\partial x}\partial_y.$$

It follows that if $u$ is horizontal, then the part of $u(s, t) = (s, t, x(s, t), y(s, t))$ in $X_D$ outside of the perturbed region satisfies

$$\frac{\partial x}{\partial s} = \frac{\partial f}{\partial y}, \quad \frac{\partial y}{\partial s} = -\frac{\partial f}{\partial x}, \qquad \frac{\partial x}{\partial t} = \frac{\partial g}{\partial y}, \quad \frac{\partial y}{\partial t} = -\frac{\partial g}{\partial x}.$$

Recall that by our assumption, away from the perturbed region inside $X_D$, we have $f_y = g_y = 0$. It follows that $x$ is locally constant.

This concludes the proof of Lemma 3.11. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

**Remark 3.12**   If we do not assume that the section $C$ has vanishing wrapping number, the above argument still shows that the wrapping number $\eta(C)$ is nonnegative. To see this, using the same notation we have $p^+ + \epsilon(mq_1^+ + nq_2^+) \geq 0$ for all generic $\epsilon > 0$. Notice that the homology class $(p^+, q_1^+, q_2^+)$ does not depend on generic $\epsilon > 0$, so we can let $\epsilon \to 0$ and conclude that $\eta(C) = p^+ \geq 0$.

A parallel result holds for $X^{m,n}$:

**Lemma 3.13** *Assume $J$ is a fibration-compatible almost complex structure on $X^{m,n}$. If $C$ is a $J$-holomorphic section of $X^{m,n}$ such that all wrapping numbers are zero, then $C$ is either contained in $X_H$ or contained in $X_D$.*

The proof of this result, however, is different from the one described above, so we present the details here:

**Proof** As before, we only need to consider the case where $C$ is not horizontal. Suppose there is some $J$-holomorphic section $C$ that is neither contained in $X_H$ nor $X_D$, without loss of generality we assume that there is some $\epsilon_1 \in (\delta, \epsilon)$ such that $C$ intersects both $F_{\pm\epsilon_1}$ transversely and $C \cap F_{\pm\epsilon_1} \neq \varnothing$.

Let $(p^\pm, q_1^\pm, q_2^\pm)$ denote the homology classes of $C \cap F_{\pm\epsilon_1}$. Since $\eta(C) = 0$, we again observe that $p^\pm = 0$. Now the local energy inequality implies that

$$mq_1^+ + nq_2^+ \geq 1, \quad mq_1^- + nq_2^- \leq -1.$$

Let $d_\infty$(resp. $d_1, d_2$) $\in \{0, 1\}$ denote the number of punctures of $C_{[-\epsilon_1, \epsilon_1]}$ that project to the positive puncture (resp. the two negative punctures) of $B_0$. We have

$$d_1(0, 1, 0) + d_2(0, 0, 1) + (0, q_1^-, q_2^-) = d_\infty(0, 1, 1) + (0, q_1^+, q_2^+)$$

and hence

$$2 \leq m(q_1^+ - q_1^-) + n(q_2^+ - q_2^-) = m(d_1 - d_\infty) + n(d_2 - d_\infty).$$

We conclude that $d_\infty = 0$ and that at least one of $d_1$ and $d_2$ is 1. There are two possibilities:

(1) $d_1 = d_2 = 1$ If this is the case, then $C$ does not have other outputs. We conclude that for any small enough $\epsilon_2$, the section $C$ cannot intersect both $F_{1\pm\epsilon_2}$, otherwise the exact same argument would tell us that $C$ has at least another output over $x = 1$. Choose $l \in \{1 - \epsilon_2, 1 + \epsilon_2\}$ such that $C \cap F_l = \varnothing$. We now look at $C \cap F_{[\epsilon_1, l]}$. This part of $C$ can only have a positive puncture (or no punctures at all) with homology class $(-k, 1, 1)$ for some $k \in \{1, 2, \ldots, m + n\}$, but the same homology class should match $(0, q_1^+, q_2^+)$, which means that there's no positive puncture. So we conclude that $C \cap F_{[\epsilon_1, l]}$ is a surface without puncture, whose boundary is $-C \cap F_{\epsilon_1}$, which implies that $q_1^+ = q_2^+ = 0$, contradicting $mq_1^+ + nq_2^+ \geq 1$.

(2) We have either $d_1 = 1$ and $d_2 = 0$ or $d_1 = 0$ and $d_2 = 1$. Without loss of generality let us assume the first case happens. There are two subcases.

**Case 2.1** If there is some small $\epsilon_2 \in (\delta, \epsilon)$ such that $C \cap F_{1-\epsilon_2} = \varnothing$ or $C \cap F_{1+\epsilon_2} = \varnothing$, then as before we fix $l \in \{1 - \epsilon_2, 1 + \epsilon_2\}$ such that $C \cap F_l = \varnothing$, and look at $C \cap F_{[\epsilon_1, l]}$. This part of $C$ can have at most one positive puncture with homology class $(-k, 1, 1)$ where $k \in \{1, 2, \ldots, m + n\}$ and at most one negative puncture with homology class $(-j, 0, 1)$ for some $j \in \{1, 2, \ldots, n\}$. If $C \cap F_{[\epsilon_1, l]}$ has no punctures, then we argue as before to show that $q_1^+ = q_2^+ = 0$, which leads to a contradiction. So $C \cap F_{[\epsilon_1, l]}$ has at least

one puncture, but then again by homology considerations we conclude that $C \cap F_{[\epsilon_1, l]}$ has precisely two punctures, with homology classes $(-k, 1, 1)$ and $(-k, 0, 1)$ for some $k \in \{1, 2, \ldots, n\}$. Now we have

$$(-k, 1, 1) = (0, q_1^+, q_2^+) + (-k, 0, 1)$$

which implies that $q_1^+ = 1$ and $q_2^+ = 0$. Now $d_1 = 1$ and $d_2 = 0$ tells us that $q_1^- = q_1^+ - 1 = 0$ and $q_2^- = q_2^+ = 0$, contradicting $mq_1^- + nq_2^- \leq -1$.

**Case 2.2** The other possibility is that we can find some $\epsilon_2 \in (\delta, \epsilon)$ such that $C$ intersects both $F_{1 \pm \epsilon_2}$ transversely. We use $(p^{1 \pm \epsilon_2}, q_1^{1 \pm \epsilon_2}, q_2^{1 \pm \epsilon_2})$ to denote the homology classes of $C \cap F_{1 \pm \epsilon_2}$. The condition $\eta(C) = 0$ now translates to $p^{1 \pm \epsilon_2} + mq_1^{1 \pm \epsilon_2} + nq_2^{1 \pm \epsilon_2} = 0$, because the wrapping number is now the integral of $dy_R$, which equals the integral of $dy + \beta^{m,n}$. The local energy inequality tells us that

$$p^{1-\epsilon_2} + (1 - \epsilon_2)(mq_1^{1-\epsilon_2} + nq_2^{1-\epsilon_2}) > 0, \quad p^{1+\epsilon_2} + (1 + \epsilon_2)(mq_1^{1+\epsilon_2} + nq_2^{1+\epsilon_2}) > 0$$

which simplifies to

$$mq_1^{1+\epsilon_2} + nq_2^{1+\epsilon_2} \geq 1, \quad mq_1^{1-\epsilon_2} + nq_2^{1-\epsilon_2} \leq -1.$$

Let $d_\infty', d_2' \in \{0, 1\}$ denote the number of punctures of $C_{[1-\epsilon_2, 1+\epsilon_2]}$ that project to the positive puncture and the second negative punctures, respectively, of $B_0$. We have

$$d_2'(-n, 0, 1) + (p^{1-\epsilon_2}, q_1^{1-\epsilon_2}, q_2^{1-\epsilon_2}) = d_\infty'(-m-n, 1, 1) + (p^{1+\epsilon_2}, q_1^{1+\epsilon_2}, q_2^{1+\epsilon_2})$$

which implies that

$$q_1^{1+\epsilon_2} = q_1^{1-\epsilon_2} - d_\infty', \quad q_2^{1+\epsilon_2} = q_2^{1-\epsilon_2} + d_2' - d_\infty'.$$

Again $2 \leq m(q_1^{1+\epsilon_2} - q_1^{1-\epsilon_2}) + n(q_2^{1+\epsilon_2} - q_2^{1-\epsilon_2})$ tells us that $d_\infty' = 0$ and $d_2' = 1$, and hence $p^{1-\epsilon_2} - n = p^{1+\epsilon_2}$.

We now look at $C \cap F_{[-\epsilon_1, 1+\epsilon_2]}$. This part of $C$ has two outputs with the homology classes $(0, 1, 0)$ and $(-n, 0, 1)$, and at most one puncture with homology class $(-k, 1, 1)$ for some $k \in \{1, 2, \ldots, m+n-1\}$. We also have $\partial(C \cap F_{[-\epsilon_1, 1+\epsilon_2]}) = C \cap F_{1+\epsilon_2} - C \cap F_{-\epsilon_1}$. We observe that $C \cap F_{[-\epsilon_1, 1+\epsilon_2]}$ must contain a positive puncture, otherwise $p_1^{1+\epsilon_2} = -n$, and hence $p^{1-\epsilon_2} = 0$, so $\eta = 0$ implies that $mq_1^{1-\epsilon_2} + nq_2^{1-\epsilon_2} = 0$, contradicting the local energy inequality $mq_1^{1-\epsilon_2} + nq_2^{1-\epsilon_2} \leq -1$. Finally, we have

$$(0, q_1^-, q_2^-) + (0, 1, 0) + (-n, 0, 1) = (p^{1+\epsilon_2}, q_1^{1+\epsilon_2}, q_2^{1+\epsilon_2}) + (-k, 1, 1)$$

which implies that $q_1^- = q_1^{1+\epsilon_2}$ and $q_2^- = q_2^{1+\epsilon_2}$, but then the local energy inequalities $mq_1^- + nq_2^- \leq -1$ and $mq_1^{1+\epsilon_2} + nq_2^{1+\epsilon_2} \geq 1$ cannot both be true. This concludes the proof of Lemma 3.13. □

# 4 The "no crossing" results for general $J$

Although the fibration compatible almost complex structures in Definition 3.1 are convenient to work with, they are not suitable for defining the cobordism map. The reason is that for given fibration compatible $J$,

not all *J*-holomorphic sections are cut out transversely, so there is not a well defined count for the cobordism map as defined in Section 2. In this section, we use the SFT compactness theorem developed in [2] to show that we can always perturb the almost complex structure slightly to a tame almost complex structure — not necessarily fibration compatible, in such a way that the no crossing results Lemma 3.11 and Lemma 3.13 continue to hold.

Throughout this section, we let $X$ denote either the bundle $X_{m,n}$ or $X^{m,n}$. We fix a fibration compatible almost complex structure $J$ on $X$, and denote by $J_+^1$, $J_+^2$, $J_-$ its restrictions on the three cylindrical ends of $X$. In the case that $\partial\Sigma \neq \varnothing$, we choose coordinates $(x_i, y_i)$ near each boundary component of $\partial\Sigma$, such that any almost complex structure we choose, even if it is not fibration compatible elsewhere, is fibration compatible near the boundary and sends $\partial_{x_i}$ to $\partial_{y_i}$.

**Theorem 4.1** *Let $\{J_k\}$ be a sequence of tame almost complex structures that $C^\infty$ converges to a fixed fibration-compatible almost complex structure $J$, and $\{C_k\}$ be a sequence of finite-energy $J_k$-holomorphic sections, which we view as maps $u_k : B_0 \to X$, that are asymptotic to fixed Reeb orbits in $Y_{\phi^m}$, $Y_{\phi^n}$ and $Y_{\phi^{m+n}}$. If all wrapping numbers of $\{C_k\}$ vanish, then $C_k$ is contained in $X_H$ or $X_D$ for sufficiently large $k$.*

The proof of Theorem 4.1 relies largely on a careful analysis of *J*-holomorphic sections in $X$ and the symplectizations $Y_{\phi^m}$, $Y_{\phi^n}$ and $Y_{\phi^{m+n}}$, which we take up in the following subsections. To begin the proof of Theorem 4.1, let us make the following simple observation. We could slightly shrink the two open subsets $X_H$ and $X_D$ to $X_{H,\tilde\epsilon}$ and $X_{D,\tilde\epsilon}$, where $\tilde\epsilon \in (\delta, \epsilon)$ and

$$X_{D,\tilde\epsilon} := B_0 \times (-\tilde\epsilon, 1+\tilde\epsilon)_x \times S_y^1, \quad X_{H,\tilde\epsilon} := B_0 \times (\Sigma - (\tilde\epsilon, 1-\tilde\epsilon)_x \times S_y^1),$$

such that Lemmas 3.11 and 3.13 still hold for the new cover $X = X_{D,\tilde\epsilon} \cup X_{H,\tilde\epsilon}$.

## 4.1 *J*-holomorphic cylinders in symplectizations

The next step is to analyze *J*-holomorphic cylinders in the symplectization $\mathbb{R} \times Y_{\phi^m}$. The analysis for the remaining cases of $Y_{\phi^n}$ and $Y_{\phi^{m+n}}$ are analogous. Similar to what we saw in Section 2, there is a decomposition of $Y_{\phi^m}$:

$$Y_{m,D,\tilde\epsilon} := S_t^1 \times (-\tilde\epsilon, 1+\tilde\epsilon)_x \times S_y^1, \quad Y_{m,H,\tilde\epsilon} := S_t^1 \times (\Sigma - (\tilde\epsilon, 1-\tilde\epsilon)_x \times S_y^1).$$

When $\tilde\epsilon = \epsilon$, without risk of confusion, we will abbreviate the two components by $Y_D$ and $Y_H$, respectively. The gluing map of $Y_{m,D,\tilde\epsilon}$ and $Y_{m,H,\tilde\epsilon}$ is defined similarly as in Section 2. For *J*-holomorphic sections in $\mathbb{R} \times Y_{\phi^m}$, the wrapping numbers are defined similarly; see [12, Definition 4.2].

As in Section 3, by slightly abusing the notations, let us denote by $F_x$ the three-dimensional manifold $\mathbb{R} \times S_t^1 \times \{x\} \times S_y^1 \subset Y_D$. Let $F_{(x_1,x_2)}$ denote the four-manifold $\mathbb{R} \times S_t^1 \times (x_1, x_2)_x \times S_y^1 \subset \mathbb{R} \times Y_D$. We also identify the first homology class in $Y_D$ with a pair $(p, q)$. Fix a symplectization compatible (and hence by Definition 3.1, a fibration compatible) almost complex structure $J$ on $\mathbb{R} \times Y_{\phi^m}$. The local energy inequality for *J*-holomorphic sections in $\mathbb{R} \times Y_{\phi^m}$ is the following:

**Lemma 4.2** [12, Lemma 3.11]   *Let $C$ be a $J$-holomorphic section, which we write as a map*

$$u \colon \mathbb{R} \times S^1 \to \mathbb{R} \times Y_{\phi^m}.$$

*Assume that $C$ intersects $F_x$ transversely for some $x$ in the unperturbed range. We have*

$$p + mxq \geq 0.$$

*Furthermore, the equality holds if and only if $C \cap F_x = \varnothing$.*

**Proof**   This is a straightforward generalization of Lemma 3.2. For a different proof, see [12]. □

The above inequality implies the following "no crossing" result for $J$-holomorphic cylinders in $\mathbb{R} \times Y_{\phi^m}$:

**Lemma 4.3**   *Let $J$ be a symplectization-compatible almost complex structure on $\mathbb{R} \times Y_{\phi^m}$. If $C$ is a $J$-holomorphic section of the bundle $\mathbb{R} \times Y_{\phi^m} \to \mathbb{R} \times S^1_t$ with vanishing wrapping numbers, then:*

(1)   *For any $\tilde{\epsilon} \in (\delta, \epsilon)$, the section $C$ is either contained in $\mathbb{R} \times Y_{m,D,\tilde{\epsilon}}$ or $\mathbb{R} \times Y_{m,H,\tilde{\epsilon}}$.*

(2)   *For such a section, if the positive end is one of the two orbits over $x = 0$ (resp. $x = 1$), then for any $\tilde{\epsilon} \in (\delta, \epsilon)$, $C$ is contained in $F_{(-\tilde{\epsilon}, \tilde{\epsilon})}$ (resp. $F_{(1-\tilde{\epsilon}, 1+\tilde{\epsilon})}$).*

(3)   *For such a section, if the negative end is one of the two orbits over $x = 0$ (resp. $x = 1$), then for any $\tilde{\epsilon} \in (\delta, \epsilon)$, the section $C$ is contained in $\mathbb{R} \times Y_{m,H,\tilde{\epsilon}}$.*

(4)   *Finally, if such a section does not have any end over $x = 0$ or $x = 1$, then it is completely contained in $\mathbb{R} \times (Y_D - Y_H)$ or $\mathbb{R} \times (Y_H - Y_D)$.*

The proof is similar to that of Lemma 3.11, but it is worthwhile to write down the details.

**Proof**   For any $\epsilon_1 \in (\delta, \epsilon)$, let us denote the homology classes of $C \cap F_{\pm \epsilon_1}$ by $(p^\pm, q^\pm)$ (the choice of $\epsilon_1$ does not matter here). Since all wrapping numbers of $C$ vanish, we conclude that $p^\pm = 0$.

To prove the first bullet point, suppose $C$ is not contained in either region. Without loss of generality we could assume there is some $\epsilon_1$ such that $C \cap F_{\pm \epsilon_1} \neq \varnothing$. Now Lemma 4.2 tells us that

$$m\epsilon_1 q^+ > 0 > m\epsilon_1 q^-.$$

So $q^+ \geq 1$ and $q^- \leq -1$. Notice that for punctures of $C$ that are contained in $F_{[-\epsilon_1, \epsilon_1]}$, the homology class is $(0, 1)$. Let us assume there are $d_\infty \in \{0, 1\}$ (resp. $d_{-\infty}$) many of such positive (resp. negative) punctures, and we have

$$d_{-\infty}(0, 1) + (0, q^-) = d_\infty(0, 1) + (0, q^+).$$

But this is not possible, because otherwise

$$2 \leq q^+ - q^- = d_{-\infty} - d_\infty \leq 1.$$

To prove the second bullet point, it suffices to show that for any $\epsilon_1 \in (\delta, \epsilon)$, we have $C \cap F_{\pm\epsilon_1} = \varnothing$ and $C \cap F_{1\pm\epsilon_1} = \varnothing$. Without loss of generality suppose $C \cap F_{\epsilon_1} \neq \varnothing$ or $C \cap F_{-\epsilon_1} \neq \varnothing$. By the same argument as in the previous paragraph, we have $q^+ - q^- \geq 1$, but now we have $d_\infty = 1$, so

$$1 \leq q^+ - q^- = d_{-\infty} - d_\infty \leq 0,$$

a contradiction.

To prove the third bullet point, simply notice that otherwise such a section is completely contained in $\mathbb{R} \times Y_D$ by the first bullet point. Now observe that Reeb orbits in $Y_D$ that are over different values of $x$ have different homology classes, it follows that both ends of $C$ are over $x = 0$ or $x = 1$. Now the second bullet point shows that such a section is contained in $\mathbb{R} \times Y_{m,H,\tilde{\epsilon}}$ as well.

Finally, to prove the last bullet point, observe that (in the same notation as before) $d_{\pm\infty} = 0$ forces that $q^\pm = 0$, hence $C \cap F_{\pm\epsilon_1} = \varnothing$. Similarly $C \cap F_{1\pm\epsilon_1} = \varnothing$ for any $\epsilon_1$. $\qquad\square$

## 4.2 More about *J*-holomorphic sections in the twist region

To prove Theorem 4.1, the final ingredient we need is a more detailed understanding of *J*-holomorphic sections that are contained in the twist region $X_D$. Let us recall that, for *J*-holomorphic sections of $X_{m,n}$ that are contained in the twist region $X_D = B_0 \times (-\epsilon, 1+\epsilon)_x \times S^1_y$, the Reeb orbits can occur over

(1)   $x = \frac{i}{m}$ ($i \in \{0, 1, \ldots, m\}$) for the first positive end;

(2)   $x = \frac{j}{n}$ ($j \in \{0, 1, \ldots, n\}$) for the second positive end;

(3)   $x = \frac{k}{m+n}$ ($k \in \{0, 1, \ldots, m+n\}$) for the negative end.

The next lemma tells us that for pseudoholomorphic sections of $X_{m,n}$ that are contained in the twist region, the three ends must in fact lie over the same $x$-coordinate.

**Lemma 4.4**   *Let J be a fibration-compatible almost complex structure on $X_{m,n}$, and C be a J-holomorphic section that is completely contained in $X_D = B_0 \times (-\epsilon, 1+\epsilon)_x \times S^1_y$. Then the x coordinate of the three cylindrical ends of C asymptote to the same value. Furthermore, C itself is completely contained in the $\delta$-neighborhood of the slice $F_x$ (the subscript x denotes the x value to which the ends of C asymptote.)*

**Proof**   The Reeb vector field near the first positive end is $\partial_t - mx\partial_y$, so the homology class of any Reeb orbit over $x = \frac{i}{m}$ is $[S^1_{t_1}] - i[S^1_y] \in H_1(X_D)$. Similarly, for Reeb orbits over the second positive end with the $x$-coordinate $\frac{j}{n}$, the homology class is $[S^1_{t_2}] - j[S^1_y] \in H_1(X_D)$; the homology class for Reeb orbits over the negative end with the $x$-coordinate $\frac{k}{m+n}$ is $[S^1_{t_1}] + [S^1_{t_2}] - k[S^1_y] \in H_1(X_D)$.

It follows that for a *J*-holomorphic section that is completely contained in $X_D$, we have $k = i + j$ for homological reasons. Now if the three ends don't share the same $x$-coordinates, without loss of generality

we can assume that $\frac{i}{m} < \frac{i+j}{m+n} < \frac{j}{n}$. Pick some $x_0 \in \left(\frac{i+j}{m+n}, \frac{j}{n}\right)$ such that $C$ intersects the slice $F_{x_0}$ transversely, then the homology class of $C \cap F_{x_0}$ in $H_1(X_D)$ is $(-j, 0, 1)$. Using Lemma 3.2, we have

$$-j + x_0(m \cdot 0 + n \cdot 1) \geq 0$$

which implies that $x_0 \geq \frac{j}{n}$, a contradiction.

Now suppose $C$ is not contained in the $\delta$-neighborhood of the slice $F_x$, we can choose some $\tilde{\epsilon}$ slightly bigger than $\delta$ such that $C$ intersects $F_{x \pm \tilde{\epsilon}}$ transversely and the intersect is nonempty. But notice that $C \cap F_{x \pm \tilde{\epsilon}}$ are both null-homologous, so this is a violation of Lemma 3.2. $\qquad\square$

**Remark 4.5** The second part of the above lemma also holds for $J$-holomorphic sections of $X^{m,n}$ that are completely contained in the twist region. Namely, if all three ends of such a section share the same $x$-coordinate, then the entire section is contained in the $\delta$-neighborhood of the slice $F_x$.

## 4.3 Proof of "no crossing" for general $J$ (Theorem 4.1)

Now we are ready to prove the main result of this section.

**Proof of Theorem 4.1** Fix some $\tilde{\epsilon} \in (\delta, \epsilon)$. Suppose that the statement of Theorem 4.1 fails, by the SFT compactness theorem, we can find a subsequence of $\{C_k\}$, still denoted by $\{C_k\}$, such that

(1) for every $k$, $C_k$ is not contained in $X_H$ or $X_D$, and

(2) $\{C_k\}$ converges to a $J$-holomorphic building $\mathcal{B}$.

Let us first observe that by our assumptions, $\pi_2(X_{m,n})$, $\pi_2(X^{m,n})$, $\pi_2(Y_{\phi^m})$ are all trivial, so bubbling off of $J$-holomorphic spheres cannot occur in any level of $\mathcal{B}$. Notice also that it is not possible for any component of any level of $\mathcal{B}$ to have only positive or only negative punctures, simply by homological considerations. The above two observations imply that

(1) $\mathcal{B}$ has no nodes;

(2) the main level of $\mathcal{B}$ is a $J$-holomorphic section of $X$, and

(3) every other level of $\mathcal{B}$ is a (resp. pair of) holomorphic cylinder in the symplectization of $Y_{\phi^{m+n}}$ (resp. $Y_{\phi^m} \sqcup Y_{\phi^n}$).

We note that all levels of $\mathcal{B}$ must have vanishing wrapping numbers. The reason is that the wrapping number is homological, so the sum of the wrapping number from all different level is equal to zero. By Remark 3.10, all wrapping numbers are nonnegative, so they have to vanish in each level as well.

If the main level of $\mathcal{B}$ is contained in $X_{H,\tilde{\epsilon}}$, we can use Lemma 4.3 and induction to show that all other levels of $\mathcal{B}$ are contained in $\mathbb{R} \times Y_{m+n,H,\tilde{\epsilon}}$ or $\mathbb{R} \times (Y_{m,H,\tilde{\epsilon}} \sqcup Y_{n,H,\tilde{\epsilon}})$, respectively. For example, the first level above the main level consists of one $J_+$-holomorphic cylinder in $\mathbb{R} \times Y_{\phi^{m+n}}$ (if $X = X^{m,n}$) or a pair of $J_+$-holomorphic cylinders in $\mathbb{R} \times (Y_{\phi^m} \sqcup Y_{\phi^n})$ (if $X = X_{m,n}$). In either case, those $J_+$-holomorphic

cylinders have negative ends which are either over $x = 0, 1$ inside the twist regions, or outside the twist region. Now the third and forth bullets points of Lemma 4.3 tell us that these cylinders are entirely contained in $\mathbb{R} \times Y_{m+n,H,\tilde{\epsilon}}$ or $\mathbb{R} \times (Y_{m,H,\tilde{\epsilon}} \sqcup Y_{n,H,\tilde{\epsilon}})$. We conclude, using induction, that all levels above the main level are contained in the same region. Now let us consider the first level under the main level. For any such $J_-$-holomorphic cylinder, if the positive end is over $x = 0, 1$, then by the second bullet point of Lemma 4.3, they are completely contained in $F_{(-\tilde{\epsilon},\tilde{\epsilon})}$; if the positive end is contained outside of the twist regions, then the third and forth bullet points of Lemma 4.3 imply that the cylinders are contained in $\mathbb{R} \times Y_{m+n,H,\tilde{\epsilon}}$ or $\mathbb{R} \times (Y_{m,H,\tilde{\epsilon}} \sqcup Y_{n,H,\tilde{\epsilon}})$. Again, we can repeat the above analysis to find that all levels under the main level are contained in $\mathbb{R} \times Y_{m+n,H,\tilde{\epsilon}}$ or $\mathbb{R} \times (Y_{m,H,\tilde{\epsilon}} \sqcup Y_{n,H,\tilde{\epsilon}})$. In summary, the entire building is contained in the (slightly shrunk) nontwist region, which implies that for sufficiently large $k$, the section $C_k$ is contained in $X_H$ as well, a contradiction.

If the main level of $\mathcal{B}$ is contained in $X_{D,\tilde{\epsilon}} - X_{H,\tilde{\epsilon}}$, then again we can use the fourth bullet point of Lemma 4.3 and induction to deduce that all other levels of $\mathcal{B}$ are contained in $\mathbb{R} \times Y_{m+n,D,\tilde{\epsilon}}$ or $\mathbb{R} \times (Y_{m,D,\tilde{\epsilon}} \sqcup Y_{n,D,\tilde{\epsilon}})$. It follows that for sufficiently large $k$, the section $C_k$ is completely contained in $X_D$, a contradiction. □

## 5 The product

In this section, we use the no crossing results to calculate the pair-of-pants product defined in Section 2,

$$(8) \qquad \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n) \to \mathrm{HF}_*(\phi^{m+n}),$$

where $\phi$ is the (Hamiltonian perturbed) positive Dehn twist along a homologically nontrivial simple closed curve $\gamma \subset \Sigma$, with the extra conditions stated in Theorem 1.2.

As reviewed in Section 2, we fix the cobordism $X = X_{m,n}$ and a generic Hamiltonian perturbation. We always assume that the almost complex structure $J$ is $C^\infty$ close to a fibration-compatible one, as in Section 3. Furthermore, we require that near each boundary component of $\partial\Sigma$ with local coordinates $(x_i, y_i)$, the almost complex structure $J$ is fibration-compatible, and is induced from the almost complex structure on Ver that sends $\partial_{x_i}$ to $\partial_{y_i}$. Lemma 3.8 tells us that $J$-holomorphic sections cannot approach $\partial\Sigma$ by the maximum principle.

### 5.1 Several remarks on monotonicity

To define fixed point Floer homology without using the Novikov rings, we need a monotonicity condition. In what follows, we will use a slightly stronger version of "weak monotonicity" introduced in [6].

**Definition 5.1** [6, Condition 2.5] Let $\psi$ be a symplectomorphism of $(\Sigma, \omega_0)$. Let $\omega_\psi$ denote the 2-form on the mapping torus $Y_\psi$ induced by $\omega_0$, and Ver the vertical distribution of $Y_\psi \to S^1$. We say $\psi$ is weakly monotone if $[\omega_\psi]$ vanishes on the kernel of

$$c_1(\mathrm{Ver})\colon H_2(Y_\psi) \to \mathbb{R}.$$

We have the following:

**Lemma 5.2** *For the positive Dehn twist $\phi\colon (\Sigma, \omega_0) \to (\Sigma, \omega_0)$, the map $\phi^m$ is weakly monotone for any positive integer $m$.*

**Proof** The proof is almost verbatim to that of [12, Lemma 5.1], and the only difference is that in our setting $\langle [\Sigma], c_1(\mathrm{Ver}) \rangle = 2 - 2g(\Sigma)$ if $\partial\Sigma = \varnothing$. $\qquad\square$

The above lemma tells us that the count in (1) is finite, so the fixed point Floer homology $\mathrm{HF}_*(\phi^m)$ is well-defined without using the Novikov rings. Similarly, we need a weak monotonicity condition for the count (5) to be finite.

**Definition 5.3** Let $\pi\colon (E, \omega) \to B$ be a symplectic fiber bundle, and $\mathrm{Ver} := \mathrm{Ker}(d\pi)$ be the vertical distribution. We say the symplectic bundle is weakly monotone if $[\omega]$ vanishes on the kernel of

$$c_1(\mathrm{Ver})\colon H_2(E) \to \mathbb{R}.$$

Similarly, we have the following lemma, which tells us that the count (5) is finite (see the discussions following the proof of Lemma 5.4), and hence the product and coproduct structures induced by $X_{m,n}$ and $X^{m,n}$ are well-defined without use of Novikov rings.

**Lemma 5.4** *If $\partial\Sigma \neq \varnothing$ or $\Sigma$ is closed with genus at least 2, then both $X_{m,n}$ and $X^{m,n}$ are weakly monotone.*

**Proof** Take a closed surface $C \subset X$ such that $[C]$ lies in the kernel of $c_1(\mathrm{Ver})$. Using the same notation as in Lemma 3.2, let $(p, q_1, q_2)$ denote the homology class of $[C \cap F_0] = [C \cap F_1] \in H_1(X_D)$ (isotope $C$ slightly to make the two intersections transverse).

Let us start with the situation where $\gamma$ is nonseparating and $\Sigma$ is closed. It is not difficult to see that

$$(9) \qquad\qquad \langle [C], c_1(\mathrm{Ver}) \rangle = (2 - 2g(\Sigma))\eta(C),$$

where $\eta$ is the wrapping number. We conclude that the wrapping number of $C$ is zero. As explained in the proof of Lemma 3.11, we have

$$\eta(C) = p = p + mq_1 + nq_2 = 0.$$

Now if $\gamma$ is separating and $\Sigma$ is closed, let $\Sigma_1$, $\Sigma_2$ denote the two components of $\Sigma - [0, 1]_x \times S_y^1$. By our assumption: $g(\Sigma_1), g(\Sigma_2) \geq 1$. Similar to the above, we have:

$$\langle [C], c_1(\mathrm{Ver}) \rangle = (1 - 2g(\Sigma_1))\eta_1(C) + (1 - 2g_2(\Sigma_2))\eta_2(C).$$

So $[C] \in \mathrm{Ker}(c_1(\mathrm{Ver}))$ implies that both wrapping numbers of $C$ vanish, hence we have

$$p = p + mq_1 + nq_2 = 0$$

again.

It is not difficult to calculate, using the explicit expression of $\omega_X$, that $\int_C \omega_X$ is a linear combination of $p$ and $mq_1 + nq_2$. So $[\omega_X]$ indeed vanishes on $[C]$.

Finally, if $\partial\Sigma \neq \varnothing$, then the wrapping number of $C$ is automatically zero if $\gamma$ is nonseparating. If $\gamma$ is separating, then:

(1) If both components of $\Sigma - [0,1]_x \times S^1_y$ contains at least one component of $\partial C$, then both wrapping numbers of $C$ automatically vanishes;

(2) If only one of the components of $\Sigma - [0,1]_x \times S^1_y$, say $\Sigma_2$, contains components of $\partial\Sigma$ (so $\eta_2$ vanishes), then we have

$$\langle [C], c_1(\mathrm{Ver}) \rangle = (1 - 2g(\Sigma_1))\eta_1(C).$$

So $[C] \in \ker(c_1(\mathrm{Ver}))$ implies that $\eta_1$ vanishes as well.

The conclusion is that $p = p + mq_1 + nq_2 = 0$ regardless. Using the exact same argument as above, we conclude that $[\omega_X]$ vanishes on $[C]$ as well. $\qquad\square$

Lemma 5.2 tells us that for any positive integer $m$, $\mathrm{HF}_*(\phi^m)$ is well defined without use of Novikov coefficients. In fact, it is well-known (see for example [21; 12]) that

$$(10) \qquad \mathrm{HF}_*(\phi^m) \cong H_*(\Sigma_0; \mathbb{Z}_2) \oplus \left( \bigoplus_{i=1}^{m-1} H_*(S^1) \right),$$

where the $i^{\text{th}}$ component of $\bigoplus_{i=1}^{m-1} H_*(S^1)$ comes from the Reeb orbits inside the Dehn twist region over $x = \frac{i}{m}$.

Before proving Theorem 1.2, some remarks on the *Hofer energy* of the $J$-holomorphic sections are needed. In the following, we view the cobordism $X_{m,n}$ as the completion of the compact cobordism $K_{m,n}$ with the same symplectic form $\Omega_X$.

**Definition 5.5** [11; 25] Let $J$ be a tame almost complex structure on $X_{m,n}$. The Hofer energy of a $J$-holomorphic section $u$ is defined as

$$E^{\mathrm{Hofer}}(u) = \sup_{f \in \mathcal{T}} \int_{B_0} u^* \omega_f$$

Where $\mathcal{T} = \left\{ f \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \mid f' > 0 \text{ and } f(x) = x \text{ near } [-\delta, \delta]) \right\}$ (for sufficiently small $\epsilon$ and $\delta$), and

$$\omega_f = \begin{cases} \Omega_X & \text{on } K_{m,n}, \\ d(f(s_i)dt_i) + \omega_X & \text{near the three punctures.} \end{cases}$$

The discussions about the monotonicity conditions imply that when the $J$-holomorphic sections have the same indices and the same asymptotes, they share the same integral $\int_{B_0} \omega_X$. Hence we have a uniform bound on the Hofer energy for sections in $\mathcal{M}^J_{x,y;z}$, so the SFT compactness theorem implies that (5) (as well as the corresponding count for the coproduct structure) is a finite count.

## 5.2  All sections have vanishing wrapping numbers

In this section we explain that in our setting, all *J*-holomorphic sections of $X_{m,n} \to B_0$ with Fredholm index 0 have vanishing wrapping numbers. This observation will allow us to use the no crossing results from Sections 3 and 4.

**Theorem 5.6**  *Let $J$ be an almost complex structure on $X_{m,n}$ that is close to a fibration-compatible one. Suppose that*

- *if $\gamma$ is nonseparating, then $\partial\Sigma \neq \varnothing$ or $\Sigma$ is closed with genus at least 2;*
- *if $\gamma$ is separating, then each component of $\Sigma - \gamma$ either contains a component of $\partial\Sigma$ or has genus at least 2.*

*Then for any Fredholm index zero $J$-holomorphic section $C$ with cylindrical ends, $C$ has vanishing wrapping numbers.*

To prove Theorem 5.6, let us recall the Fredholm index formula. Let $C$ be a $J$-holomorphic section in $X_{m,n}$ with positive asymptotes $\alpha_i$ and negative asymptote $\beta$, represented by a map $u : B_0 \to X_{m,n}$. Fix a trivialization $\tau$ of the vertical distribution along each Reeb orbit, and denote by $\langle c_1^\tau(TX_{m,n}), [C]\rangle$ the first Chern number of the complex vector bundle $u^*TX_{m,n}$ over $B_0$ with respect to the trivialization $\tau$ and the natural splitting $TX_{m,n}|_\gamma \cong \mathrm{Ver} \oplus \mathbb{R}\langle R, \partial_s\rangle$ over the ends. Here $\mathbb{R}\langle R, \partial_s\rangle$ denotes the distribution spanned by the Reeb vector field and the symplectization direction. For each asymptotic orbit, let $\mathrm{CZ}_\tau$ be the Conley–Zehnder index with respect to $\tau$. We have the Fredholm index formula:

$$(11) \qquad \mathrm{ind}(C) = 1 + 2\langle c_1^\tau(TX_{m,n}), [C]\rangle + \sum \mathrm{CZ}_\tau(\alpha_i) - \mathrm{CZ}_\tau(\beta).$$

Notice that in our setting, the map $u$ is a section of the fibration $X_{m,n} \to B_0$, so $u^*TX_{m,n}$ naturally splits as $u^*TX_{m,n} \cong TB_0 \oplus u^*\mathrm{Ver}$. In light of this splitting, we have

$$\langle c_1^\tau(TX_{m,n}), [C]\rangle = -1 + \langle c_1^\tau(\mathrm{Ver}), [C]\rangle.$$

So the index formula can be rewritten as

$$\mathrm{ind}(C) = -1 + 2\langle c_1^\tau(\mathrm{Ver}), [C]\rangle + \sum \mathrm{CZ}_\tau(\alpha_i) - \mathrm{CZ}_\tau(\beta).$$

Now recall that the Reeb orbits can be divided into two types: those coming from critical points of $mH_0$, $nH_0$ or $(m+n)H_0$ outside of $N$ and those lying inside the twist region. There is a natural choice of the trivialization $\tau$ of the distribution Ver over these Reeb orbits: for the critical points of $H$, the distribution Ver can be identified with the tangent space $T\Sigma$ at the point; and over the Dehn twist region $N$; we can identify Ver with $TN$. We will always choose $\tau$ as above, and the Conley–Zehnder index $\mathrm{CZ}_\tau$ with respect to such a trivialization is:

$-1$  if the orbit comes from a local minimum of $H$, or is an elliptic orbit inside the Dehn twist region;

$\phantom{-}0$  if the orbit comes from a saddle point of $H$, or is a hyperbolic orbit inside the Dehn twist region;

$\phantom{-}1$  if the orbit comes from a local maximum of $H$.

Following [12], we now demonstrate a lemma relating the relative first Chern number $\langle c_1^\tau(\text{Ver}), [C] \rangle$ to the wrapping number $\eta(C)$ (similar ideas were applied in the proof of Lemma 5.4; we present a proof of the generalization of (9) here):

**Lemma 5.7** *If $\Sigma$ is a closed surface with genus $g$, the loop $\gamma$ is nonseparating, and $C$ is a $J$-holomorphic section, then*

$$\langle c_1^\tau(\text{Ver}), [C] \rangle = (2 - 2g)\eta(C).$$

**Proof** Choose a generic point outside of the twist region, which we denote by $pt \in \Sigma_0$, such that $C$ intersects $B_0 \times \{pt\}$ transversely. Recall that $\eta(C)$ is by definition the algebraic intersection number $\#C \cap (B_0 \times \{pt\})$.

Choose a section $\psi$ of Ver over $X_{m,n}$, with the following property:

(1) When restricted to the Reeb orbits, $\psi$ is constant with respect to $\tau$.

(2) There are $l$ points $p_1, p_2, \ldots, p_l$ concentrated in an arbitrarily small neighborhood of $pt$, such that on each fiber of $X_{m,n} \to B_0$, the section $\psi$ has transverse zeroes at precisely $p_1, p_2, \ldots, p_l$, with total degree $2 - 2g$.

We can also arrange $\psi$ so that $C$ intersects each $B_0 \times \{p_i\}$ transversely. Now by definition, $\langle c_1^\tau(\text{Ver}), [C] \rangle$ is the algebraic count of zeroes of $u^*\psi$. Observe that the zeroes of $\psi|_C$ occurs at precisely

$$C \cap (B_0 \times \{p_1, p_2, \ldots, p_l\}),$$

and the algebraic count of these zeroes is $(2 - 2g)\eta(C)$. $\qquad\square$

Now we are ready to prove Theorem 5.6.

**Proof of Theorem 5.6** Let us start with the case where $\gamma$ is nonseparating. As remarked before, if $\partial\Sigma \neq \varnothing$, then $\eta(C)$ is automatically zero. If $\Sigma$ is closed, by equation (9), we have

$$0 = \text{ind}(C) = -1 + 2(2 - 2g)\eta(C) + \sum \text{CZ}_\tau(\alpha_i) - \text{CZ}_\tau(\beta).$$

But since $\text{CZ}_\tau \in \{-1, 0, 1\}$, we have

$$\text{ind}(C) \leq 2 + (4 - 4g)\eta(C).$$

This, together with the fact that $\eta(C) \geq 0$ and the assumption $g \geq 2$, forces that $\eta(C) = 0$.

Now let us deal with the case where $\gamma$ is separating. As before, let us denote by $\Sigma_1$ and $\Sigma_2$ the two components of $\Sigma - N$. If $\Sigma$ is closed, then similar to Lemma 5.7, we have

$$\langle c_1^\tau(\text{Ver}), [C] \rangle = (1 - 2g(\Sigma_1))\eta_1(C) + (1 - 2g(\Sigma_2))\eta_2(C).$$

So

$$0 = \mathrm{ind}(C)$$

$$= -1 + 2(1 - 2g(\Sigma_1))\eta_1(C) + 2(1 - 2g(\Sigma_2))\eta_2(C) + \sum \mathrm{CZ}_\tau(\alpha_i) - \mathrm{CZ}_\tau(\beta)$$

$$\leq 2 + 2(1 - 2g(\Sigma_1))\eta_1(C) + 2(1 - 2g(\Sigma_2))\eta_2(C).$$

By our assumption, $g(\Sigma_1), g(\Sigma_2) \geq 2$. Combined with the fact that $\eta_1, \eta_2 \geq 0$, we have

$$\eta_1(C) = \eta_2(C) = 0,$$

as desired.

The case where both $\Sigma_1$ and $\Sigma_2$ contain a component of $\partial\Sigma$ is easy: we only need to observe as before that if $\Sigma_i$ contains a component of $\partial\Sigma$, then $\eta_i$ is automatically zero. The only remaining situation is the following: only one of the two components of $\Sigma - N$ contains a component of $\partial\Sigma$, and the other one has genus at least 2. Without loss of generality let us assume $\Sigma_2$ is the one containing $\partial\Sigma$ (so $\eta_2$ vanishes automatically). Similar to what we saw in Lemma 5.7, we have

$$\langle c_1^\tau(\mathrm{Ver}), [C] \rangle = (1 - 2g(\Sigma_1))\eta_1(C)$$

which implies that

$$0 = \mathrm{ind}(C)$$

$$= -1 + 2(1 - 2g(\Sigma_1))\eta_1(C) + \sum \mathrm{CZ}_\tau(\alpha_i) - \mathrm{CZ}_\tau(\beta)$$

$$\leq 2 + 2(1 - 2g(\Sigma_1))\eta_1(C).$$

Again, since $g(\Sigma_1) \geq 2$ and $\eta_1 \geq 0$, we conclude that $\eta_1(C)$ has to vanish as well. $\qquad\square$

## 5.3 Computation of the product (proof of Theorem 1.2)

**Proof of Theorem 1.2**  We choose a generic tame almost complex structure $J$ on $X_{m,n}$ such that all moduli spaces of Fredholm index zero sections are cut out transversely. We further assume that $J$ is $C^\infty$ close to a fibration-compatible almost complex structure so that Theorem 4.1 applies. Theorems 4.1 and 5.6 tell us that all the $J$-holomorphic sections are either contained in $X_H$ or $X_D$.

Now the count of $J$-holomorphic sections contained in $X_H$ precisely corresponds to the intersection product of $H_*(\Sigma_0; \mathbb{Z}_2) \subset \mathrm{HF}_*(\phi^m)$ and $H_*(\Sigma_0; \mathbb{Z}_2) \subset \mathrm{HF}_*(\phi^n)$ in the sense of the decomposition (10). By [17; 9; 14], the cobordism map of the pair-of-pants product in this case can be identified with the intersection pairing (notice that in our case $\pi_2(\Sigma_0) = 0$, so no Novikov rings are needed here),

$$(12) \qquad\qquad H_*(\Sigma_0; \mathbb{Z}_2) \otimes H_*(\Sigma_0; \mathbb{Z}_2) \xrightarrow{\ \cap\ } H_*(\Sigma_0; \mathbb{Z}_2).$$

To finish the proof, we only need to show that the count of sections contained in the twist region contributes to zero in the cobordism map. By Lemma 4.4, any such $J$-holomorphic section must be contained in the $\delta$-neighborhood of some slice $F_x$, where $x = \frac{i}{d}$ for some $i \in \{1, 2 \ldots, d-1\}$, here $d := \gcd(m, n)$ (the reason is that we need $mx$, $nx$ and $(m+n)x$ to be integers simultaneously for the three Reeb orbits to exist inside a $\delta$-neighborhood of $F_x$). Recall that, near $x = \frac{i}{d}$, the symplectic fiber bundle is given by the trivial product $X_i := B_0 \times \left(\frac{i}{d} - \delta, \frac{i}{d} + \delta\right)_x \times S_y^1$ with the fiberwise symplectic closed 2-form $\omega_X$, which is

a small Hamiltonian perturbation of $\omega_{X,0} = dx \wedge dy + d\left(\frac{1}{2}x^2\beta_{m,n}\right)$. Now we use a change-of-coordinate trick to show that the above symplectic fiber bundle is equivalent to another one, which calculates the pair-of-pants product of the Hamiltonian Floer homology of *small* Hamiltonians on $\left(\frac{i}{d} - \delta, \frac{i}{d} + \delta\right)_x \times S^1_y$. To do this, let $X_0$ be the trivial bundle $B_0 \times (-\delta, \delta)_{x'} \times S^1_{y'}$ together with the fiberwise symplectic form $dx' \wedge dy' + d\left(\frac{1}{2}x'^2\beta_{m,n}\right)$. Define a diffeomorphism $\mu \colon X_i \to X_0$ by

(13) $$x' = x - \frac{i}{d}, \quad y' = y + i \cdot g'_{m,n}(p),$$

where $p$ denotes the coordinate on $B_0$. It's easy to see that $\mu$ preserves the fibers, and that $\mu$ pulls $dx' \wedge dy' + d\left(\frac{1}{2}x'^2\beta_{m,n}\right)$ back to $\omega_{X,0}$, because (recall that $d \cdot g'_{m,n} = g_{m,n}$)

$$
\begin{aligned}
\mu^*\left(dx' \wedge dy' + d\left(\tfrac{1}{2}x'^2\beta_{m,n}\right)\right) &= dx \wedge (dy + i \cdot g'_{m,n}) + d\left(\tfrac{1}{2}\left(x - \tfrac{i}{d}\right)^2\beta_{m,n}\right) \\
&= dx \wedge dy + \tfrac{i}{d}dx \wedge \beta_{m,n} + \left(x - \tfrac{i}{d}\right)dx \wedge \beta_{m,n} \\
&= dx \wedge dy + d\left(\tfrac{1}{2}x^2\beta_{m,n}\right) \\
&= \omega_{X,0}.
\end{aligned}
$$

For any almost complex structure $J$ on $X_i$, we have the following one-to-one correspondence:

$$\{J\text{-holomorphic sections of } X_i\} \overset{1:1}{\longleftrightarrow} \{\mu_*(J)\text{-holomorphic sections of } X_0\}.$$

Now, similar to the $J$-holomorphic sections that are contained in $X_H$, it is clear that sections contained in $X_0$ computes the pair-of-pants product of (a small perturbation of) the fixed points of time-1 maps of $m \cdot \frac{1}{2}x'^2$ and $n \cdot \frac{1}{2}x'^2$. This corresponds to the intersection product of $H_*((-\delta, \delta)_{x'} \times S^1_{y'})$, which is identically zero. This concludes the proof of Theorem 1.2. $\qquad\square$

# 6 The coproduct

In this section, we generalize the methods used in Section 5 further to compute the pair-of-pants coproduct of fixed point Floer homology of Dehn twists: $\mathrm{HF}_*(\phi^{m+n}) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n)$, where $m, n$ are positive integers, and $\phi$ is the positive Dehn twist described in the setup. Note that Lemma 5.4 tells us that the cobordism map is well-defined even without the use of Novikov rings. The goal of this section is to prove Theorem 1.3. To begin with, similar to Theorem 5.6, we have the following:

**Theorem 6.1** *Let $J$ be an almost complex structure on $X^{m,n}$ that is close to a fibration-compatible one. Suppose*:

- *If $\gamma$ is nonseparating, then $\partial\Sigma \neq \varnothing$ or $\Sigma$ is closed with genus at least 2.*

- *If $\gamma$ is separating, then each component of $\Sigma - \gamma$ either contains a component of $\partial\Sigma$ or has genus at least 2.*

*Then for any index zero $J$-holomorphic section $C$ with cylindrical ends, $C$ has vanishing wrapping numbers.*

And the proof is almost the same as that of Theorem 5.6 (we only need to slightly modify the Conley–Zehnder index term). Combined with Theorem 4.1, it tells us the following:

**Corollary 6.2** *Suppose the almost complex structure $J$, and the loop $\gamma$ satisfy the same condition as in Theorem 6.1, then all $J$-holomorphic sections of $X^{m,n} \to B_0$ that have Fredholm index zero must be contained in $X_H$ or $X_D$.*

What is different from Section 5 is that the count of sections contained in the twist region does not contribute to zero in the cobordism map. In the following two subsections, we first give a detailed understanding of the moduli space of all $J$-holomorphic sections in the Morse–Bott (unperturbed) setting, then explain what would happen if we perturb the form $\omega_{X,0}$ to break the Morse–Bott degeneracy.

## 6.1 *J*-holomorphic sections inside the twist region

In this subsection, we analyze possible $J$-holomorphic sections inside $X_D$. Unless otherwise specified, the almost complex structure $J$ is assumed to be fibration-compatible, and when restricted to Ver sends $\partial_x$ to $\partial_y$ (so $J$ is completely determined by the fiberwise symplectic 2-form $\omega$). We start with the fiberwise symplectic 2-form $\omega_{X,0} = dx \wedge dy + d\left(\frac{1}{2}x^2\beta_{m,n}\right)$ and view $(X_D, \omega_{X,0})$ as part of $\overline{X}_D := B_0 \times \mathbb{R}_x \times S_y^1$ with the same fiberwise symplectic 2-form. Note that by extending the cobordism we are not introducing new curves: Lemma 3.2 ensures that if the $x$-coordinates of the asymptotic Reeb vector fields of a given $J$-holomorphic section are contained in $(-\epsilon, 1+\epsilon)$, then the entire $J$-holomorphic section of $\overline{X}_D$ is in fact entirely contained in $X_D$.

Notice that with this Morse–Bott setting, the Reeb orbits at the ends come in $S^1$-families. The possible $x$-coordinates of such families are

- $x = \dfrac{k_\infty}{m+n}$ at the positive end,

- $x = \dfrac{k_1}{m}$ at the first negative end, and

- $x = \dfrac{k_2}{n}$ at the second negative end.

Observe that if a $J$-holomorphic section has cylindrical ends at $x = \dfrac{k_\infty}{m+n}, \dfrac{k_1}{m}$, and $\dfrac{k_2}{n}$, then $k_\infty = k_1 + k_2$.

We make the following two basic observations about $J$-holomorphic sections of $\overline{X}_D$ with cylindrical ends. In what follows, for any fiberwise symplectic 2-form $\omega$ on $\overline{X}_D$ that coincides with $\omega_{X,0}$ outside of some compact subset $K \subset \overline{X}_D$, let $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ denote the moduli space of $J$-holomorphic sections of $(\overline{X}_D, \omega)$ whose ends have $x$-coordinates asymptoting to

$$\frac{k_\infty}{m+n}, \quad \frac{k_1}{m}, \quad \text{and} \quad \frac{k_2}{n}, \quad \text{where } k_\infty = k_1 + k_2.$$

**Remark 6.3** We make an observation about $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ that is already implicit in its definition. Note here the Reeb orbits come in $S^1$ families, and in defining $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ we allow the ends of its elements to land on any Reeb orbit on a given $S^1$ family. In other words, the ends of a $J$-holomorphic section are "free". We can also require ends of $J$-holomorphic sections to land on a specific Reeb orbit in a $S^1$ family, in which case the ends are "fixed". This distinction will be very important to us when we pass from Morse–Bott case to the Morse case.

**Lemma 6.4** *There is a one-to-one correspondence between*

$$\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2) \quad \text{and} \quad \mathcal{M}_{\omega_{X,0}}(k_\infty + m + n; k_1 + m, k_2 + n).$$

**Proof** Define a diffeomorphism $\mu \colon \overline{X}_D \to \overline{X}_D$, $(p, x, y) \mapsto (p, x', y')$ by

$$x' = x - 1, \quad y' = y + g^{m,n}(p).$$

Observe that $\mu$ preserves Ver, and a simple calculation shows $\mu^* \omega_{X,0} = \omega_{X,0}$. So $\mu^* J = J$, and hence if $u$ is a $J$-holomorphic section contained in $\mathcal{M}_{\omega_{X,0}}(k_\infty + m + n; k_1 + m, k_2 + n)$, then $\mu \circ u$ is a $J$-holomorphic section contained in $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$, and vice versa. □

**Lemma 6.5** *For any $u \in \mathcal{M}_\omega(k_\infty; k_1, k_2)$, $\mathrm{ind}(u) = 1$ and $u$ is cut out transversely.*

**Proof** Similar to what we did in Section 5, choose the trivialization

$$\tau \colon \mathrm{Ver} \to T(\mathbb{R}_x \times S^1_y) \cong \mathbb{R}^2.$$

Let $\alpha$ and $\beta_1, \beta_2$ denote the three ends of $u$. The index formula is

$$\begin{aligned}
\mathrm{ind}(u) &= 1 + 2\langle c_1^\tau(TX), [C] \rangle + \mathrm{CZ}_\tau^+(\alpha) - \sum \mathrm{CZ}_\tau^-(\beta_i) \\
&= -1 + \mathrm{CZ}_\tau^+(\alpha) - \sum \mathrm{CZ}_\tau^-(\beta_i) \\
&= -1 + 0 - (-1) - (-1) = 1.
\end{aligned}$$

Now the automatic transversality theorem [24, Theorem 1] applies here, because

$$1 = \mathrm{ind}(u) > c_N(u) + Z(du) = 0 + 0. \qquad □$$

Notice that Lemma 6.5 does not require $\omega$ to be $\omega_{X,0}$. Let us consider a 1-parameter family of closed fiberwise symplectic 2-forms $\omega_\lambda$ for $\lambda \in [0, 1]$, where

(1) $\omega_0 = \omega_{X,0}$,

(2) when restricted to each fiber of $\overline{X}_D \to B_0$, we have $\omega_\lambda|_{\mathrm{fiber}} = dx \wedge dy$,

(3) there is a compact subset of $\overline{X}_D$ outside of which all $\omega_\lambda$, with $\lambda \in [0, 1]$, agree.

Observe that any closed fiberwise symplectic 2-form $\omega$ that agrees with $\omega_{X,0}$ outside of a compact set and restricts to $dx \wedge dy$ on each fiber can be connected to $\omega_{X,0}$ using a family $\omega_\lambda$ described above: one can simply put $\omega_\lambda = \lambda\omega + (1-\lambda)\omega_{X,0}$. For each $\lambda$, let $J_\lambda$ denote the fibration compatible almost complex structure on $\overline{X}_D$ determined by $\omega_\lambda$. As before, let $\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)$ denote the moduli space of $J_\lambda$-holomorphic sections of $(\overline{X}_D, \omega_\lambda)$ whose ends have $x$-coordinates $\frac{k_\infty}{m+n}, \frac{k_1}{m}, \frac{k_2}{n}$, where $k_\infty = k_1 + k_2$.

Let $\{\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)\}|_{\lambda \in [0,1]}$ denote the parametrized moduli space:

$$\{\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)\}|_{\lambda \in [0,1]} := \{(u, \lambda)|u \in \mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2), \ \lambda \in [0, 1])\}.$$

We now describe the specific type of deformed fiberwise symplectic 2-form that we will use. Fix a compact subset $K_1 \subset B_0$ such that the complement of $K_1$ is contained in the cylindrical ends. For any $R > 0$, denote $[-R, R]_x \times S_y^1 \subset \mathbb{R}_x \times S_y^1$ by $Q_R$.

**Definition 6.6** A 1-form $\sigma \in \Omega^1(B_0, C^\infty(\mathbb{R} \times S^1))$ is called *admissible* if there exists some compact subset $K_2 \subset B_0$ containing $K_1$, and some $R > 0$, such that $\sigma = \beta^{m,n} \cdot \frac{1}{2}x^2$ outside of $K_2 \times Q_R$.

For any admissible 1-form $\sigma$, we define the corresponding closed fiberwise symplectic 2-form $\omega$ to be $dx \wedge dy + d\sigma$. The following simple observation asserts that if $\omega = dx \wedge dy + d\sigma$ and $\sigma$ is admissible, then all pseudoholomorphic sections in $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ have a uniform upper bound on the vertical energy (see definition below) and the range of its $x$ component is bounded.

**Definition 6.7** Fix an admissible 1-form $\sigma$ and the corresponding almost complex structure $J$. Let $g_J$ denote the metric determined by $\omega = d\sigma + dx \wedge dy$ and $J$. For any smooth map $u \colon B_0 \to \overline{X}_D$, define the vertical energy $E(u)$ to be

$$E(u) = \frac{1}{2} \int_{B_0} |\partial_s u - \partial_s^\#|_{g_J}^2 + |\partial_t u - \partial_t^\#|_{g_J}^2 \ ds \wedge dt,$$

where $\partial_s^\#$ and $\partial_t^\#$ are the horizontal lifts of the vector fields $\partial_s$ and $\partial_t$, respectively.

**Lemma 6.8** *Let $\sigma_\lambda|_{\lambda \in [0,1]}$ be a family of admissible 1-form as in Definition 6.6 (where $R$ is assumed to be sufficiently large compared to $k_\infty$, $k_1$ and $k_2$), and $\omega_\lambda = dx \wedge dy + d\sigma_\lambda$ be the corresponding closed fiberwise symplectic 2-form on $\overline{X}_D$. Let $J_{\omega_\lambda}$ denote the fibration-compatible almost complex structure determined by $\omega_\lambda$. Then for any $u \in \mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)$,*

(1) *$u$ is contained in $\{-2R \leq x \leq 2R\}$, and*

(2) *the vertical energy $E(u)$ has a uniform bound.*

**Proof** Assume that there is some $u \in \mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)$ not contained in $\{-2R \leq x \leq 2R\}$. It's easy to show that such a section cannot be horizontal (otherwise, inside the region $\{R < |x| < 2R\}$ the function $x$ would be locally constant, a contradiction). Outside of $\{-2R \leq x \leq 2R\}$ the 2-form $\omega$ coincides with $\omega_{X,0}$,

so after picking some $x_0$ such that $|x_0| > 2R$ and $u$ intersects $\{x = x_0\}$ transversely in a nonempty way, the homology class of the intersection $u \cap \{x = x_0\}$ satisfies the local energy inequality

$$p + x_0(mq_1 + nq_2) > 0,$$

which contradicts the fact that $u \cap \{x = x_0\}$ is null-homologous. For the bullet point 2, we first observe that for a sequence of subdomains $D_k$ exhausting $B_0$, we have

$$\int_{B_0} u^* \omega_\lambda = \lim_{k \to \infty} \int_{D_k} u^* \omega_\lambda = \lim_{k \to \infty} \int_{\partial D_k} u^*(x\,dy + \sigma_\lambda) = \lim_{k \to \infty} \int_{\partial D_k} u^*\left(x\,dy + \beta^{m,n} \cdot \tfrac{1}{2}x^2\right).$$

And the last limit does not depend on $\lambda$ or $u$. We also observe that given the fact $u \in \mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)$, this term is finite.

Next, similar to what we did in Lemma 3.5, we calculate $E(u)$ for a $J$-holomorphic section $u$ (again using local conformal coordinates $(s, t)$):

$$E(u) = \frac{1}{2} \int_{B_0} |\partial_s u - \partial_s^\#|_{g_J}^2 + |\partial_t u - \partial_t^\#|_{g_J}^2 \, ds \wedge dt = \int_{B_0} \omega_\lambda(\partial_s u - \partial_s^\#, \partial_t u - \partial_t^\#) \, ds \wedge dt$$

$$= \int_{B_0} \omega_\lambda(\partial_s u, \partial_t u) - \omega_\lambda(\partial_s^\#, \partial_t^\#) \, ds \wedge dt$$

$$= \int_{B_0} u^* \omega_\lambda - \int_{B_0} \omega_\lambda(\partial_s^\#, \partial_t^\#) \, ds \wedge dt.$$

So it suffices to estimate the term $\int_{B_0} \omega_\lambda(\partial_s^\#, \partial_t^\#) ds \wedge dt$. Write $\sigma_\lambda$ as $F^\lambda(s, t, x, y) ds + G^\lambda(s, t, x, y) dt$, and we have

$$\omega_\lambda(\partial_s^\#, \partial_t^\#) = \frac{\partial G^\lambda}{\partial s} - \frac{\partial F^\lambda}{\partial t} + \frac{\partial G^\lambda}{\partial x} \frac{\partial F^\lambda}{\partial y} - \frac{\partial G^\lambda}{\partial y} \frac{\partial F^\lambda}{\partial x}.$$

Now we simply observe that on the union of cylindrical ends $Z$ where $\sigma_\lambda = \tfrac{1}{2}x^2 \beta^{m,n}$, the term $F^\lambda$ is identically zero and $G^\lambda$ is independent of $s$. Consequently, $\omega_\lambda(\partial_s^\#, \partial_t^\#)$ is compactly supported on

$$(s, t, x, y, \lambda) \in (B_0 - Z) \times [-2R, 2R] \times S^1 \times [0, 1]$$

and hence has a uniform bound for all $u \in \mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)$. $\qquad \square$

Lemma 6.8, together with the usual SFT compactness argument developed in [2] (see also [25] for a nice account), tells us that for any closed fiberwise symplectic 2-form $\omega = dx \wedge dy + d\sigma$ where $\sigma$ is admissible, the moduli space $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ is compact. To see that no SFT type breaking can occur for a sequence of sections in $\mathcal{M}_\omega(k_\infty; k_1, k_2)$, we observe that levels in the symplectizations are necessarily cylinders, and such cylinders asymptote to orbits in the same Morse–Bott family for homological reasons. Now such cylinders have zero vertical energy, hence are trivial cylinders. We also observe no bubbles appear, as $\pi_2$ of the bundle is trivial.

More generally, for such $\omega$, if we define the 1-parameter family $\omega_\lambda := \lambda\omega + (1-\lambda)\omega_{X,0}$, then the parametric moduli space $\{\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)\}|_{\lambda \in [0,1]}$ is compact as well. Lemma 6.5 tells us that for such $\omega$, both

$\mathcal{M}_\omega(k_\infty; k_1, k_2)$ and $\{\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)\}|_{\lambda \in [0,1]}$ are transversely cut out, so $\{\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)\}|_{\lambda \in [0,1]}$ is a compact cobordism between two closed 1-dimensional manifolds. Observe also that for each fixed $\lambda' \in [0, 1]$, the slice $\mathcal{M}_{\omega_{\lambda'}}(k_\infty; k_1, k_2) \subset \{\mathcal{M}_{\omega_\lambda}(k_\infty; k_1, k_2)\}|_{\lambda \in [0,1]}$ is a closed 1-manifold, so we get the following:

**Corollary 6.9** *For any $\omega = dx \wedge dy + d\sigma$ where $\sigma$ is admissible, $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ is diffeomorphic to $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$.*

The next observation (Corollary 6.13) asserts that $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$ has at most one component. To get started, let us observe that there is an $S^1$ symmetry of $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$. In the following, we will view $J$-holomorphic sections $u$ of $(\overline{X}_D, \omega)$ as maps $\bar{u} \colon B_0 \to \mathbb{R} \times S^1$, so it is handy to establish the following lemma:

**Lemma 6.10** *Choose a local conformal coordinate $(s, t)$ of $B_0$ and suppose locally*

$$\sigma = \tfrac{1}{2}x^2(F(s,t)ds + G(s,t)dt), \quad \omega_{X,0} = dx \wedge dy + d\sigma.$$

*A map $u \colon B_0 \to (\overline{X}_D, \omega)$ is a $J$-holomorphic section if and only if the corresponding map $\bar{u} = (x(s,t), y(s,t)) \colon B_0 \to \mathbb{R} \times S^1$ in our coordinate system solves the following PDE:*

$$\frac{\partial x}{\partial t} + \frac{\partial y}{\partial s} + xF = 0, \quad \frac{\partial y}{\partial t} - \frac{\partial x}{\partial s} + xG = 0.$$

**Proof** Let $v^\#$ denote the horizontal lift (with respect to $\omega$) of any vector $v \in TB_0$. A simple calculation shows that

$$\partial_s^\# = \partial_s - xF\partial_y, \quad \partial_t^\# = \partial_t - xG\partial_y.$$

So by definition, $J(\partial_s - xF\partial_y) = \partial_t - xG\partial_y$. Recall that we required that $J$ always sends $\partial_x$ to $\partial_y$, so this shows

$$J(\partial_s) = \partial_t - xF\partial_x - xG\partial_y.$$

Now suppose $u \colon B_0 \to \overline{X}_D$, $(s,t) \mapsto (s, t, x, y)$ is $J$-holomorphic, ie

$$J\left(\partial_s + \frac{\partial x}{\partial s}\partial_x + \frac{\partial y}{\partial s}\partial_y\right) = \partial_t + \frac{\partial x}{\partial t}\partial_x + \frac{\partial y}{\partial t}\partial_y.$$

Combine the above equations and collect the coefficients of $\partial_x$ and $\partial_y$, and we get the desired PDE. $\square$

The following corollary is immediate:

**Corollary 6.11** *If $u \colon B_0 \to \overline{X}_D$ is a map given by $(s,t) \mapsto (s, t, x(s,t), y(s,t))$, and $u$ is an element of $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$, then for any $y_0 \in S^1$, $(s,t) \mapsto (s, t, x(s,t), y(s,t) + y_0)$ is also an element of $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$.*

In other words, if $\mathcal{M}_{\omega_X,0}(k_\infty; k_1, k_2) \neq \varnothing$, then $S^1$ acts freely on $\mathcal{M}_{\omega_X,0}(k_\infty; k_1, k_2)$ by translating the $y$-coordinate. We now show that this $S^1$ action is also transitive.

**Lemma 6.12** *Suppose $u_1$ and $u_2$ are two different J-holomorphic sections in $\mathcal{M}_{\omega_X,0}(k_\infty; k_1, k_2)$, and let $\bar{u}_i : B_0 \to \mathbb{R} \times S^1$ denote the corresponding maps to the fiber, which we write as $(s, t) \mapsto (x_i, y_i)$. Then $x_1 = x_2$, and there is some $y_0 \in S^1$ such that $y_2(s, t) = y_1(s, t) + y_0$.*

**Proof** Consider $w := \bar{u}_1 - \bar{u}_2 : B_0 \to \mathbb{R} \times S^1$, $(s, t) \mapsto (l(s, t), m(s, t))$. By Lemma 6.10, locally the following PDEs are satisfied:

$$\frac{\partial l}{\partial t} + \frac{\partial m}{\partial s} + lF = 0, \quad \frac{\partial m}{\partial t} - \frac{\partial l}{\partial s} + lG = 0.$$

Observe that by our assumption, $u_1$ and $u_2$ have the same asymptotics on their cylindrical ends, so it follows that $w$ induces the trivial map on $\pi_1$, hence can be lifted to $\tilde{w} : B_0 \to \mathbb{R} \times \mathbb{R}$, $(s, t) \mapsto (l(s, t), \tilde{m}(s, t))$, where $(l, \tilde{m})$ solves the same PDE. Let us denote the covering map $\mathbb{R} \times \mathbb{R} \to \mathbb{R} \times S^1$ by $\pi$.

Now for any $\zeta \in \mathbb{R}$, the map $\bar{u}_2 + \pi(\zeta \cdot \tilde{w})$ solves the same PDE, and hence gives an element in $\mathcal{M}_{\omega_X,0}(k_\infty; k_1, k_2)$. This implies that $l(s, t)$ is identically zero and hence $\tilde{m}(s, t)$ is constant, because otherwise,

$$\bar{u}_2 + \pi(\zeta \cdot \tilde{w} + \xi \cdot (0, 1))$$

solves the same PDE for any $\zeta, \xi \in \mathbb{R}$, giving us a two dimensional family of solutions. This contradicts the fact that $u_2$ is cut out transversely and $\mathrm{ind}(u_2) = 1$. $\square$

A corollary of the above discussion is the following:

**Corollary 6.13** *For any admissible $\sigma$, the moduli space $\mathcal{M}_{dx \wedge dy + d\sigma}(k_\infty; k_1, k_2)$ is either empty, or diffeomorphic to $S^1$.*

We now proceed to show that $\mathcal{M}_{dx \wedge dy + d\sigma}(k_\infty; k_1, k_2)$ is not empty. By Lemma 6.4, we can assume that $k_\infty$, $k_1$ and $k_2$ are all positive. In light of Corollary 6.9, it suffices to find *one* special admissible $\sigma$, and show that $\mathcal{M}_\omega(k_\infty; k_1, k_2)$ is not empty, where $\omega = dx \wedge dy + d\sigma$. The rest of this subsection describes how one can construct an admissible $\sigma$ with a nonempty $\mathcal{M}_{dx \wedge dy + d\sigma}(k_\infty; k_1, k_2)$.

To this end, we make the following observation. Fix cylindrical ends $Z_\infty$, $Z_1$ and $Z_2$ of $B_0$ outside of $K_1$, each with conformal coordinates $(s_i, t_i) \in [N, \infty) \times S^1)$ (or $(-\infty, -N] \times S^1)$, and choose cutoff functions $\chi_i : Z_i \to [0, 1]$ such that $1 - \chi_i$ are compactly supported. Suppose $v : B_0 \to \mathbb{R} \times S^1$ is a smooth map, such that (notice the resemblance to the formulations in [20]):

(1) The restriction of $v = (x, y)$ to the cylindrical end $Z_\infty$ solves the equations

$$\frac{\partial x}{\partial t_\infty} + \frac{\partial y}{\partial s_\infty} = 0, \quad \frac{\partial y}{\partial t_\infty} - \frac{\partial x}{\partial s_\infty} + (m + n) \chi_\infty(s_\infty, t_\infty) x = 0.$$

(2)  The restriction of $v = (x, y)$ to the cylindrical end $Z_1$ solves the equations

$$\frac{\partial x}{\partial t_1} + \frac{\partial y}{\partial s_1} = 0, \quad \frac{\partial y}{\partial t_1} - \frac{\partial x}{\partial s_1} + m\chi_1(s_1, t_1)x = 0.$$

(3)  The restriction of $v = (x, y)$ to the cylindrical end $Z_2$ solves the equations

$$\frac{\partial x}{\partial t_2} + \frac{\partial y}{\partial s_2} = 0, \quad \frac{\partial y}{\partial t_2} - \frac{\partial x}{\partial s_2} + n\chi_2(s_2, t_2)x = 0.$$

(4)  The restriction of $v = (x, y)$ to the complement of $Z_1 \cup Z_2 \cup Z_\infty$ is holomorphic:

$$\frac{\partial x}{\partial t} + \frac{\partial y}{\partial s} = \frac{\partial y}{\partial t} - \frac{\partial x}{\partial s} = 0$$

(5)  The map $v$ approaches the projections (to the fiber $S^1 \times \mathbb{R}$) of three Reeb orbits at

$$x_1 = \frac{k_1}{m}, \quad x_2 = \frac{k_2}{n}, \quad x_\infty = \frac{k_\infty}{m+n}$$

at its corresponding cylindrical ends. Note the Reeb orbits whose projections the map $v$ approaches all live in $S^1$ families. We do not care which orbits in these $S^1$ families $v$ approaches.

Then we can construct an admissible $\sigma$ such that the moduli space

$$\mathcal{M}_{dx \wedge dy + d\sigma}(k_\infty; k_1, k_2)$$

is not empty. The reason is that $\tilde{v} : (s, t) \mapsto (s, t, v(s, t))$ is a smooth section of $\overline{X}_D \to B_0$ with the desired asymptotes, and the image of $v$ is contained in $Q_R$ for some large $R$. If we define an admissible 1-form $\sigma$ such that

$$\sigma = \begin{cases} \frac{1}{2}(m+n)x^2\chi_\infty(s_\infty, t_\infty)dt_\infty, & \text{in } Z_\infty \times Q_R, \\ \frac{1}{2}mx^2\chi_1(s_1, t_1)dt_1, & \text{in } Z_1 \times Q_R, \\ \frac{1}{2}nx^2\chi_2(s_2, t_2)dt_2, & \text{in } Z_2 \times Q_R, \\ 0, & \text{in } (B_0 - Z_1 \cup Z_2 \cup Z_\infty) \times Q_R. \end{cases}$$

Then Lemma 6.10 tells us that $\tilde{v} \in \mathcal{M}_{dx \wedge dy + d\sigma}(k_\infty; k_1, k_2)$.

We now construct $v$ as described above. In the following, inside the cylindrical ends $Z_i$, we will always assume that $\chi_i$'s are $t_i$-independent and monotone, and that

$$y = \begin{cases} -k_\infty t_\infty, & \text{in } Z_\infty, \\ -k_1 t_1, & \text{in } Z_1, \\ -k_2 t_2, & \text{in } Z_2. \end{cases}$$

So inside $Z_i$, $x$ is $t_i$-independent, and the equations simplify to ODEs

$$\frac{\partial x}{\partial s_\infty} + k_\infty - (m+n)\chi_\infty(s_\infty)x = 0, \quad \frac{\partial x}{\partial s_1} + k_\infty - m\chi_1(s_1)x = 0, \quad \frac{\partial x}{\partial s_2} + k_\infty - n\chi_2(s_2)x = 0.$$

If we let

$$a_\infty(s_\infty) := \int_N^{s_\infty} -(m+n)\chi_\infty(\rho)\,d\rho, \quad a_1(s_1) := \int_{-N}^{s_1} -m\chi_1(\rho)\,d\rho \quad \text{and} \quad a_2(s_2) := \int_{-N}^{s_2} -n\chi_2(\rho)\,d\rho,$$

then the solutions to the above ODEs can be explicitly written down:

$$x(s_\infty) = e^{-a_\infty}\left(c_\infty - k_\infty \int_N^{s_\infty} e^{a_\infty(\rho)}\,d\rho\right),$$

$$x(s_1) = e^{-a_1}\left(c_1 - k_1 \int_{-N}^{s_1} e^{a_1(\rho)}\,d\rho\right), \quad x(s_2) = e^{-a_2}\left(c_2 - k_2 \int_{-N}^{s_2} e^{a_2(\rho)}\,d\rho\right).$$

Based on our assumptions, it's easy to show that for the first expression, there is a unique choice of $c_\infty \in \mathbb{R}$ such that $x(s_\infty)$ stays at $\frac{k_\infty}{m+n}$ when $s_\infty$ is large enough. On the other hand, whatever choice of $c_1$ or $c_2$ is, the second (resp. the third expression) will converge to $\frac{k_1}{m}$ (resp. $\frac{k_2}{n}$) when $s_i$ close enough to $-\infty$. Notice also that near the boundaries of $Z_i$ (ie when $s_i$ is close to $\pm N$) the functions $x(s_i)$ are linear with slopes $-k_i$.

For a suitable choice of $c_\infty < c_1 = c_2$, it is not hard to construct a holomorphic $k_\infty$-fold branched cover $v = (x, y) \colon B_0 - Z_1 \cup Z_2 \cup Z_\infty \to [c_\infty, c_1] \times S^1$ such that near $\partial Z_\infty$ (resp. $\partial Z_1, \partial Z_2$), $(x, y) = (c_\infty - k_\infty(s_\infty - N), -k_\infty t_\infty)$ (resp. $(c_1 - k_1(s_1 + N), -k_1 t_1)$ and $(c_2 - k_2(s_2 + N), -k_2 t_2)$). So for such choices of $c_\infty < c_1 = c_2$, the map $v$ glues smoothly with the three solutions on $Z_i$'s. The above discussion shows the following:

**Proposition 6.14** *For any admissible $\sigma$, $\mathcal{M}_{dx \wedge dy + d\sigma}(k_\infty; k_1, k_2)$ is diffeomorphic to $S^1$.*

## 6.2 Morse–Bott theory and enumeration of sections after the perturbation

In this subsection we use some Morse–Bott theory to enumerate sections after we replace $\omega_{X,0}$ with $\omega_X$ by adding a small Hamiltonian perturbation that breaks the Morse–Bott degeneracy (see Section 2). We know from Theorems 6.1 and 4.1 that no holomorphic sections cross from $X_H$ to $X_D$ or vice versa, so we only need to consider what happens to $J$-holomorphic sections that stay in $X_D$ as we break the Morse–Bott degeneracy.

For simplicity let us work instead in $\overline{X}_D$. By abuse of notation, we will use the same letter $\omega_X$ to denote the 2-form on the completion $\overline{X}_D$, and likewise for $\omega_{X,0}$. We first recall some conventions for $J$-holomorphic sections whose ends land on Morse–Bott submanifolds. We fix $J$ the fibration-compatible almost complex structure (we will have a bit more to say about choice of $J$ after we describe cascades). In our case all Reeb orbits come in $S^1$ families, corresponding to $x = \frac{k_1}{m}, \frac{k_2}{m}, \frac{k_1+k_2}{m+n}$ in $\overline{X}_D$. Hence we have tori that are foliated by Reeb orbits (we will call such tori "Morse–Bott tori").

Recall for $J$-holomorphic sections ending on Morse–Bott tori, we can consider moduli spaces of sections with "fixed" end points and moduli spaces of sections with "free" end points. These are conditions we impose on the given cylindrical ends of the holomorphic section. "Fixed" end points condition means the given end of the section must land on a specific Reeb orbit in the Morse–Bott torus, whereas the "free" end condition means that the given end of a section in the moduli space is allowed to freely move around on the Morse–Bott torus. The dimension of the moduli space of course depends on how many ends are specified as fixed

or free. For instance the moduli space we constructed in the previous section, $\mathcal{M}_{dx\wedge dy+d\sigma}(k_\infty; k_1, k_2)$ has all ends free by this convention (we implicitly assumed this in our previous construction).

We shall invoke some standard Morse–Bott theory (which is easily adapted to cobordisms) to figure out how to count our $J$-holomorphic sections. However, the statement of the entire theory is rather cumbersome, instead we will just state the small parts we need. We refer the reader to [26] for a more detailed account.

We recall that in breaking the Morse–Bott degeneracy via perturbations, each torus of Morse Bott orbits breaks into an hyperbolic orbit $h_k$ at maximum of $h(y)$ and an elliptic orbit $e_k$ at minimum of $h(y)$. These orbits form the generators of the fixed point Floer homology.

For our purposes, $J$-holomorphic sections of degree one and Fredholm index zero in $(\overline{X}_D, \omega_X)$ which we need to count correspond to cascades in $(\overline{X}_D, \omega_{X,0})$. The cascades, generally speaking, take the following form:

- There is a main level $u_0 \colon B_0 \to (\overline{X}_D, \omega_{X,0})$.
- There are upper levels indexed by $i \in \mathbb{Z}_{>0}$. We write them as $u_{i>0} \colon S^1 \times \mathbb{R} \to (S^1 \times \mathbb{R} \times S^1 \times \mathbb{R}, \omega_{X,0})$.
- There are lower levels indexed by $j \in \mathbb{Z}_{<0}$ labeling the level number, and $k \in \{1, 2\}$ labeling which negative puncture it corresponds to. We write them as: $u_{j<0}^k \colon S^1 \times \mathbb{R} \to (S^1 \times \mathbb{R} \times S^1 \times \mathbb{R}, \omega_{X,0})$.
- For each map $u_i$ let $\pi$ denote the projection to the base in the codomain, then $\pi \circ u_i$ is the identity; likewise for $u_j^k$.
- Let $ev^\pm(u_i)$ denote the Reeb orbit $u_i$ approaches as $s \to \pm\infty$, let $\phi_T$ denote the gradient flow of $h(y)$ for time $T$, then there exists $T_i \in (0, \infty)$ so that $\phi_{T_i}(ev^-(u_i)) = ev^+(u_{i-1})$.
- For the main level we have numbers $T_1, T_k^0 \in (0, \infty)$ for $k \in \{1, 2\}$ so that
$$\phi_{T_1}(ev^-(u_1)) = ev^+(u_0), \quad \phi_{T_0^k}(ev^-(u_0)) = ev^+(u_{-1}^k).$$
- For the lower levels, there are numbers $T_j^k \in (0, \infty)$ for $k \in \{1, 2\}$ so that
$$\phi_{T_j^k}(ev^-(u_j^k)) = ev^+(u_{j-1}^k).$$

Generally speaking there are more conditions we can achieve for cascades by choosing generic $J$, however for our case we immediately observe for homological reasons $u_{i\neq 0}$ are cylinders, in fact they must all be trivial cylinders by energy considerations, so there is only the main level, which for ease of notation we denote by $u$. Since our entire cascade only has one level, in order to count $u$, it must live in a moduli space of dimension zero.

Thus we arrive at the following description of the cascades we must count:

**Proposition 6.15** *The* (1-*level*) *cascades that we need to count takes the following form*:

- $u \colon B_0 \to (\overline{X}_D, \omega_{X,0})$ *is a $J$-holomorphic section.*

- Let the ends of $B_0$ be labeled $\{1, 2, \infty\}$. *Then one of the ends in $\{1, 2, \infty\}$ is fixed, the other 2 are free. Hence $u$ belongs in a* (*transversely cut out*) *moduli space of index zero.* (*Fixing one end reduces the virtual dimension by one.*)

- *All free ends avoid critical points of $h(y)$* (*this can be achieved for generic $J$*). *If the ends labeled $1, 2$ are fixed, then they land on the maximum of $h(y)$. If the $\infty$ end is fixed, it lands on the minimum of $h(y)$.*

That the cascades above are the ones we need to count to compute the coproduct map is supplied by the following correspondence theorem:

**Proposition 6.16** *Given a perfect Morse function $h(y)$, we make the Morse perturbation smaller by rescaling it to be $\bar{\delta} h(y)$, where $\bar{\delta} > 0$ is a small positive real number. For small enough $\bar{\delta} > 0$, there is a 1–1 correspondence between $J$-holomorphic sections in the nondegenerate case and $J$ holomorphic cascades. Given a cascade $u$ of the form described in the previous proposition, for the positive end of $u$, if it is a free end we assign the generator $h_k$, and if it a fixed end we assign it the generator $e_k$. We reverse the assignments for the two negative ends $1, 2$. Then each cascade gives rise bijectively to one $J$ holomorphic section beginning and ending at the generators we assigned to the ends of cascade.*

**Proof** See [26] for a proof of this statement and how the correspondence works in the general case of multiple level cascades. However since we are only working in the simple case of 1-level cascades, the correspondence theorem that we need is also established in [5]. $\square$

**Remark 6.17** Note by the above correspondence theorem, the only requirement we need to impose on $J$ in the Morse–Bott case is that all moduli spaces of index zero sections listed above are transversely cut out and that the free ends avoid critical points of $h(y)$. It will be apparent from the paragraph following this remark that the fibration compatible $J$ we chose in Section 6.1, with the help of automatic transversality, suffices for the purpose of showing the index zero sections are transversely cut out. To ensure the free ends avoid critical points of $h$, instead of further perturbing $J$, we shall instead choose generic $h$. This kind of strategy was also undertaken when Morse–Bott techniques were employed in [13].

We also remark that in the correspondence between cascades and $J$ holomorphic sections we need to perturb the almost complex structure from $J$ in the Morse–Bott case in $\overline{X}_D$ to a generic $J_{\bar{\delta}}$ in the nondegenerate case to establish a correspondence between cascades and holomorphic sections, see [26] for more details. The difference between $J_{\bar{\delta}}$ and $J$ can be taken to be $C^\infty$ small. Hence for small enough $\bar{\delta} > 0$ we do not need to worry about this change in almost complex structure since we already established in Section 4 that for $C^\infty$ small perturbations of $J$ the no crossing results continues to hold.

Recall that Proposition 6.14 tells us the moduli space of sections with all three ends free in the Morse–Bott case come in $S^1$ families, and this $S^1$ family is precisely rotation around the $\partial_y$ direction, ie rotation

along the Morse–Bott tori. To obtain a cascade of Fredholm index zero in the form specified above, it is readily apparent we restrict one of the three ends fixed and the rest two free. All of our cascades we need arise this way.

Instead of perturbing $J$ further to ensure that when we have fixed an end, all the remaining free ends avoid critical points of $h$, we choose generic Morse perturbations $h$. To be specific, we choose perfect Morse functions $\{h_l^i(y)\}$, where $l \in \{1, 2, \infty\}$ labels which cylindrical end of the base we are referring to, and $i \in \mathbb{Z}$ refers to the specific Morse–Bott torus in that end. (For example when $l = 1$ and $i = 1$, this refers to the Morse Bott torus at $x = 1$ for the negative end labeled by $k = 1$). We then use these functions to break the Morse–Bott degeneracy.

Given some large integer $N$, for a generic choice of $\{h_l^i(y)\}$ we can arrange that for tuples $(k_1, k_2, k_\infty)$ satisfying $0 \leq k_1, k_2, k_\infty \leq N$ and $k_\infty = k_1 + k_2$, if we consider the moduli space $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$ (here all ends are free) which is diffeomorphic to $S^1$, if any element in this $S^1$ family has one end landing on a critical point of any of the Morse functions in $\{h_l^i(y)\}$, then all the other ends avoid critical points of elements in $\{h_l^i(y)\}$. We can arrange for this to happen by picking generic collection $\{h_l^i(y)\}$ because we only need to consider finite number of moduli spaces.

It is then apparent if a section in the $S^1$ family $\mathcal{M}_{\omega_{X,0}}(k_\infty; k_1, k_2)$ has one end hitting a critical point, it is the only section in this $S^1$ family with an end on that critical point. This, combined with the cascade and holomorphic section correspondence (Proposition 6.16) thus produces the computation required for the coproduct structure restricted to $\overline{X}_D$.

**Remark 6.18** Even though all our previous proofs (eg no crossing) assumed we used a single Morse function $h(y)$ to break the Morse–Bott degeneracy, one can check that using a collection of Morse functions $\{h_l^i(y)\}$ to break the degeneracy does not make a difference to our previous results.

Now we are ready to compute the coproduct map induced by $X^{m,n}$. Let $h_{k_\infty}^{m+n}$ (resp. $e_{k_\infty}^{m+n}$) be the hyperbolic (resp. elliptic) orbit over $x = \frac{k_\infty}{m+n}$ at the positive end, $h_{k_i}^m$ (resp. $e_{k_i}^m$) be the hyperbolic (resp. elliptic) orbit over $x = \frac{k_i}{m}$ at the first negative end, and $h_{k_i}^n$ (resp. $e_{k_i}^n$) be the hyperbolic (resp. elliptic) orbit over $x = \frac{k_i}{n}$ at the second negative end. Let $\Delta$ be the coproduct map induced by $(X^{m,n}, \omega_X)$. The above discussion shows the following:

**Corollary 6.19** *For all $k_\infty \in \{0, 1, \ldots, m+n\}$,*

$$\tag{14} \Delta([e_{k_\infty}^{m+n}]) = \sum_{\substack{k_i \in \{0,1,\ldots,m\}, \\ k_\infty - k_i \in \{0,1,\ldots,n\}}} [e_{k_i}^m] \otimes [e_{k_\infty - k_i}^n],$$

$$\tag{15} \Delta([h_{k_\infty}^{m+n}]) = \sum_{\substack{k_i \in \{0,1,\ldots,m\}, \\ k_\infty - k_i \in \{0,1,\ldots,n\}}} [e_{k_i}^m] \otimes [h_{k_\infty - k_i}^n] + [h_i^m] \otimes [e_{k_\infty - k_i}^n].$$

**Remark 6.20**  Notice that in equations (14) and (15), the homology classes $[e_0^{m+n}]$, $[h_0^{m+n}]$, $[e_{m+n}^{m+n}]$, $[h_{m+n}^{m+n}]$, $[e_0^m]$, $[h_0^m]$, $[e_m^m]$, $[h_m^m]$, $[e_n^n]$, $[h_n^n]$, $[e_n^n]$, $[h_n^n]$ refer to homology classes in $H_*(\Sigma_0; \mathbb{Z}_2)$ in the sense of the decomposition (10). It's also not difficult to see that the computations for $\Delta([e_0^{m+n}])$, $\Delta([h_0^{m+n}])$, $\Delta([e_{m+n}^{m+n}])$ and $\Delta([h_{m+n}^{m+n}])$ coincide with that of the coproduct structure $\Delta_0$ on $H_*(\Sigma_0; \mathbb{Z}_2)$.

Finally, We are able to complete the proof of Theorem 1.3.

**Proof of Theorem 1.3**  Let $J$ be an almost complex structure on $X^{m,n}$ sufficiently close to a fibration-compatible one in the sense of Theorem 4.1. By Corollary 6.2, all $J$-holomorphic sections that are counted in the cobordism map are either contained in the twist region $X_D$ or $X_H$. Similar to what we observed in the proof of Theorem 1.2, the count of $J$-holomorphic sections contained in $X_H$ precisely corresponds to the coproduct

$$(16) \qquad \mathrm{HF}_*(\phi^{m+n}) \supset H_*(\Sigma_0; \mathbb{Z}_2) \xrightarrow{\Delta_0} H_*(\Sigma_0; \mathbb{Z}_2) \otimes H_*(\Sigma_0; \mathbb{Z}_2) \to \mathrm{HF}_*(\phi^m) \otimes \mathrm{HF}_*(\phi^n)$$

in the sense of the decomposition (10). The remaining parts of Theorem 1.3 readily follows from Corollary 6.19. $\qquad\qquad\square$

The proofs of Theorems 1.5 and 1.6 follow directly from the arguments in Sections 5 and 6. The reason is that in the generalized situation, the wrapping number is still defined in the same way. And since the "no crossing" results (Lemmas 3.11 and 3.13) only depend on the local result Lemma 3.2, a completely analogous "no crossing" result holds in the generalized setting. The only global result we need is contained in the first half of the proof of Theorem 5.6, which can be easily generalized to the case where $\Sigma - (C_1 \cup C_2 \cup \cdots \cup C_k)$ is connected. Finally, our construction of the $J$-curves in Section 6 are purely local, and hence can be directly generalized to the setups in Theorems 1.5 and 1.6.

**Remark 6.21**  The above two theorems can be further extended to the case where each component of $\Sigma_0 = \Sigma - (C_1 \cup C_2 \cup \cdots \cup C_n)$ contains either a component of $\partial\Sigma$ or has genus at least 2. As explained above, the only difference is that we need to define multiple wrapping numbers for each of the connected components of $\Sigma_0$. We also have to modify the proof of Theorem 5.6 and conclude that all of the wrapping numbers $\eta_l(C)$ vanish for a $J$-holomorphic section $C$.

# References

[1]  **A Abbondandolo**, **M Schwarz**, *Floer homology of cotangent bundles and the loop product*, Geom. Topol. 14 (2010) 1569–1722  MR

[2]  **F Bourgeois**, **Y Eliashberg**, **H Hofer**, **K Wysocki**, **E Zehnder**, *Compactness results in symplectic field theory*, Geom. Topol. 7 (2003) 799–888  MR

[3]  **G Chen**, *On cobordism maps on periodic Floer homology*, Algebr. Geom. Topol. 21 (2021) 1–103  MR

[4]   **K Cieliebak**, **J Latschev**, *The role of string topology in symplectic field theory*, from "New perspectives and challenges in symplectic field theory" (M Abreu, F Lalonde, L Polterovich, editors), CRM Proc. Lecture Notes 49, Amer. Math. Soc., Providence, RI (2009) 113–146  MR

[5]   **V Colin**, **P Ghiggini**, **K Honda**, **Y Yao**, *Morse–Bott gluing*  MR    Appendix to V Colin, P Ghiggini and K Honda, "Embedded contact homology and open book decompositions", Geom. Topol. 29 (2025) 3345–3463

[6]   **A Cotton-Clay**, *Symplectic Floer homology of area-preserving surface diffeomorphisms*, Geom. Topol. 13 (2009) 2619–2674  MR

[7]   **S Dostoglou**, **D A Salamon**, *Self-dual instantons and holomorphic curves*, Ann. of Math. (2) 139 (1994) 581–640  MR

[8]   **E Eftekhary**, *Floer homology of certain pseudo-Anosov maps*, J. Symplectic Geom. 2 (2004) 357–375  MR

[9]   **U Frauenfelder**, **F Schlenk**, *Hamiltonian dynamics on convex symplectic manifolds*, Israel J. Math. 159 (2007) 1–56  MR

[10]  **R Gautschi**, *Floer homology of algebraically finite mapping classes*, J. Symplectic Geom. 1 (2003) 715–765  MR

[11]  **H Hofer**, *Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three*, Invent. Math. 114 (1993) 515–563  MR

[12]  **M Hutchings**, **M Sullivan**, *The periodic Floer homology of a Dehn twist*, Algebr. Geom. Topol. 5 (2005) 301–354  MR

[13]  **M Hutchings**, **M Sullivan**, *Rounding corners of polygons and the embedded contact homology of $T^3$*, Geom. Topol. 10 (2006) 169–266  MR

[14]  **S Lanzat**, *Hamiltonian Floer homology for compact convex symplectic manifolds*, Beitr. Algebra Geom. 57 (2016) 361–390  MR

[15]  **D McDuff**, **D Salamon**, *$J$-holomorphic curves and symplectic topology*, 2nd edition, AMS Colloquium Publications 52, Amer. Math. Soc., Providence, RI (2012)  MR

[16]  **R Pedrotti**, *Fixed point Floer cohomology of disjoint Dehn twists on a $w^+$-monotone manifold with rational symplectic form*, J. Symplectic Geom. 22 (2024) 549–597  MR

[17]  **S Piunikhin**, **D Salamon**, **M Schwarz**, *Symplectic Floer–Donaldson theory and quantum cohomology*, from "Contact and symplectic geometry" (C B Thomas, editor), Publ. Newton Inst. 8, Cambridge Univ. Press (1996) 171–200  MR

[18]  **M Poźniak**, *Floer homology*, *Novikov rings and clean intersections*, PhD thesis, University of Warwick, Coventry (1994) Available at `https://people.math.ethz.ch/~salamon/PREPRINTS/Pozniak.pdf`

[19]  **J H Rooney**, *On cobordism maps in embedded contact homology*, PhD thesis, University of California, Los Angeles (2018) MR  Available at `https://www.proquest.com/docview/2059826475`

[20]  **M Schwarz**, *Cohomology operations from $S^1$-cobordisms in Floer homology*, PhD thesis, ETH Zürich (1995)

[21]  **P Seidel**, *The symplectic Floer homology of a Dehn twist*, Math. Res. Lett. 3 (1996) 829–834  MR

[22]  **P Seidel**, *Floer homology and the symplectic isotopy problem*, PhD thesis, Oxford University (1997) Available at `http://dedekind.mit.edu/~seidel/thesis.pdf`

[23]  **P Seidel**, *Symplectic Floer homology and the mapping class group*, Pacific J. Math. 206 (2002) 219–229  MR

[24]  **C Wendl**, *Automatic transversality and orbifolds of punctured holomorphic curves in dimension four*, Comment. Math. Helv. 85 (2010) 347–407  MR

[25]  **C Wendl**, *Lectures on symplectic field theory*, preprint (2016)  Available at `https://tinyurl.com/wendlSFTlectures`

[26]  **Y Yao**, *From cascades to $J$-holomorphic curves and back* (2022)  arXiv 2206.04334

*Nantes Université*
*Nantes, France*

*University of California, Los Angeles*
*Los Angeles, CA, United States*

`yuan.yao.yy.1995@gmail.com`,  `zhaoziwen2017@gmail.com`

# Multitwists in big mapping class groups

GEORGE DOMAT

FEDERICA FANONI

SEBASTIAN HENSEL

We show that the group generated by multitwists (ie products of powers of twists about disjoint non-accumulating curves) doesn't contain the Torelli group of an infinite-type surface. As a consequence, multitwists don't generate the closure of the compactly supported mapping class group of a surface of infinite type.

20F65, 57K20

## 1 Introduction

The mapping class group of a surface of finite type has been thoroughly studied for decades. In particular, multiple *simple* sets of generators are known. The Dehn–Lickorish theorem [7; 11], in combination with the Birman exact sequence [4], shows that the pure mapping class group of a finite-type surface can be generated by finitely many Dehn twists about nonseparating curves, and we need to add finitely many half-twists to generate the full mapping class group. Humphries [10] proved that, if the surface is closed and of genus $g \geq 2$, $2g + 1$ Dehn twists about nonseparating curves suffice to generate the mapping class group, and moreover this number is optimal: fewer than $2g + 1$ Dehn twists cannot generate. Other results show that mapping class groups can be generated by two elements (see eg [17]), by finitely many involutions or by finitely many torsion elements (see eg [5]).

In the case of surfaces of infinite type, the (pure) mapping class group is uncountable, so in particular it is not finitely (nor countably) generated. For some of these surfaces the mapping class group is generated by torsion elements, or even by involutions (see [6; 12]), while for other surfaces they aren't (see [6; 8; 12]). To the best of our knowledge, no other generating set is known.

Note that the (pure) mapping class group of a surface of infinite type is endowed with an interesting topology, induced by the compact–open topology on the group of homeomorphisms of the surface. So it is interesting to talk about *topological* generating sets (sets whose *closure* of the group they generate is the (pure) mapping class group). It follows from the finite-type results that Dehn twists topologically generate the closure of the compactly supported mapping class group. Moreover, Patel and Vlamis [14] proved that the pure mapping class group of a surface is topologically generated by Dehn twists if the surface has at most one nonplanar end, and by Dehn twists and maps called *handle shifts* otherwise.

Our goal here is to investigate a natural candidate for a set of generators of the closure of the compactly supported mapping class group of a surface: the collection of *multitwists*. A multitwist is a (possibly infinite) product of powers of Dehn twists about a collection of simple closed curves that do not accumulate anywhere in the surface. Our main result is a negative one, and it follows from a nongeneration result for the Torelli group:

**Theorem A**  *Let $S$ be an infinite-type surface. Then the subgroup of the mapping class group of $S$ generated by multitwists doesn't contain the Torelli group. In particular, multitwists don't generate the closure of the compactly supported mapping class group.*

The idea of the proof is to produce an explicit element in the Torelli group that is not in the subgroup generated by multitwists. This element is built by taking an infinite product of increasing powers of partial pseudo-Anosov homeomorphisms supported on disjoint finite-type subsurfaces. We use work of Bestvina, Bromberg and Fujiwara [3] to certify that the mapping class we construct is not in the subgroup generated by multitwists.

Theorem A also begs the following question:

**Question B**  *What is the subgroup generated by the collection of multitwists? Is there an alternative, more explicit description of its elements?*

Furthermore, our theorem shows that the subgroup generated by the collection of multitwists is not a closed subgroup of the mapping class group. Therefore, it does not immediately inherit a Polish topology from the topology on the mapping class group.

**Question C**  *Is the subgroup generated by the collection of multitwists a Polish group?*

## 2  Preliminaries

Here a surface is a connected orientable Hausdorff second countable two-dimensional manifold without boundary unless otherwise stated. One notable exception is any subsurface, which will always have compact boundary. Boundary components of subsurfaces are assumed to be homotopically nontrivial, but are allowed to be homotopic to a puncture.

Surfaces are *of finite type* if their fundamental groups are finitely generated and *of infinite type* otherwise. A surface $S$ is *exceptional* if it has genus zero and at most four punctures, or genus one and at most one puncture, otherwise it is *nonexceptional*.

The *mapping class group* of a surface $S$ is the group $\mathrm{MCG}(S)$ of orientation-preserving homeomorphisms of $S$ up to homotopy. The *pure mapping class group* $\mathrm{PMCG}(S)$ is the subgroup of $\mathrm{MCG}(S)$ fixing all ends and — if there are any — boundary components, and $\overline{\mathrm{MCG}_c(S)}$ denotes the closure of the subgroup generated by compactly supported mapping classes. The *Torelli group* $\mathcal{I}(S)$ is the subgroup of the mapping class group given by elements acting trivially on the first homology group of the surface.

A pseudo-Anosov mapping class $f$ of a finite-type surface is *chiral* if $f^k$ is not conjugate to $f^{-k}$ for every $k \neq 0$. We will need chiral pseudo-Anosovs in the Torelli group. The existence of such is likely well known, but since we could not find a proof in the literature we include the following:

**Lemma 1** *Let $\Sigma$ be a finite-type nonexceptional surface with boundary. Then there is a mapping class $\varphi \in \mathrm{MCG}(\Sigma)$ with the following properties*:

  (i) *$\varphi$ is pseudo-Anosov,*

 (ii) *$\varphi$ is chiral, and*

(iii) *if $\Sigma$ is a subsurface of a surface $S$, so that each boundary component of $\Sigma$ is separating in $S$, then $\varphi$ acts trivially on the first homology of $S$.*

**Proof** Begin by taking an arc $\rho$ joining a boundary component $c$ of $\Sigma$ to itself, which is filling and nullhomologous in $H_1(\Sigma, \partial\Sigma)$. By Scott's theorem [16] there is a cover of $\Sigma$ in which $\rho$ lifts to a simple arc.

By taking a further cover, we can find a cover $\widehat{\Sigma} \to \Sigma$ such that $\rho$ lifts to a simple arc $\hat{\rho}$ in $\widehat{\Sigma}$ joining two different boundary components $\hat{c}_1$ and $\hat{c}_2$ of $\widehat{\Sigma}$, and such that the cover is characteristic (ie every homeomorphism of $\Sigma$ lifts).

Denote by $\hat{\delta}$ the boundary of a regular neighbourhood of $c_1 \cup \hat{\rho} \cup c_2$. Observe that since one complementary component of $\hat{\delta}$ is a pair of pants, and the other one is not, no orientation-preserving homeomorphism of $\widehat{\Sigma}$ can preserve $\hat{\delta}$ setwise while inverting its orientation.

Now, let $\gamma$ be an oriented loop on $\Sigma$ based at a point $p \in c$, which lifts in $\widehat{\Sigma}$ to a curve freely homotopic to $\hat{\delta}$. We claim that the boundary-push $P_\gamma$ defined by $\gamma$ has the desired property.

First, $\gamma$ is filling (since $\rho$ was filling), and so the same proof as that of Kra's theorem (see [9, Theorem 14.6]) shows that $\varphi = P_\gamma$ is pseudo-Anosov, proving (i).

Next, boundary pushes defined by nullhomologous loops are in the Torelli group of $\Sigma$. If $\Sigma$ is a subsurface of a surface $S$ and all of the boundary components of $\Sigma$ are separating in $S$, the Torelli group of $\Sigma$ is contained in the Torelli group of $S$, showing (iii). (See [15] for a study of Torelli groups of subsurfaces without assuming the boundary components to be separating.)

Finally, note that conjugating a boundary-push simply has the effect of changing the pushing curve:

$$f P_\gamma f^{-1} = P_{f(\gamma)}.$$

Furthermore, push maps are equal if and only if the defining curves are homotopic (by the Birman exact sequence). Thus, if $f \varphi^i f^{-1} = \varphi^j$, then $f(\gamma)^i = \gamma^j$. Since $f$ preserves primitivity in the fundamental group, we only have to exclude the case $i = 1$ and $j = -1$. But if $f \varphi f^{-1} = \varphi^{-1}$, there is a lift $\hat{f}$ of $f$ which preserves $\hat{\delta}$ and inverts its orientation, which is impossible, as discussed above. Hence (ii) holds. $\square$

A *curve* on a surface is the homotopy class of an essential (ie not homotopic to a point, a puncture or a boundary component) simple closed curve. Given a curve $\alpha$, we denote by $\tau_\alpha$ the Dehn twist about $\alpha$.

An *integral weighted multicurve* $\mu$ is a formal sum $\sum_{i \in I} n_i \alpha_i$, where the $\alpha_i$ are pairwise disjoint curves not accumulating anywhere and the $n_i$ are integers. Given an integral weighted multicurve $\mu$, we define $\tau_\mu$ to be the mapping class

$$\tau_\mu = \prod_{i \in I} \tau_{\alpha_i}^{n_i}.$$

Such a mapping class is called a *multitwist*.

We say that an integral weighted multicurve is *finite* if $I$ is finite (ie it contains finitely many curves). An integral weighted multicurve $\nu$ is a *submulticurve* of an integral weighted multicurve $\mu = \sum_{i \in I} n_i \alpha_i$ if $\nu = \sum_{i \in J} n_i \alpha_i$, where $J \subset I$.

Given a surface with boundary, an *arc* is the homotopy class (relative to the boundary) of a simple arc that cannot be homotoped into the boundary. We denote by $\mathcal{C}(S)$ the *curve and arc graph* of a surface $S$, where vertices are curves and, if $\partial S \neq \varnothing$, arcs, and two vertices are adjacent if they have disjoint representatives.

For any two subsurfaces $A$ and $B$ of $S$ that have an essential intersection, the *subsurface projection* of $B$ to $A$ is the subset $\partial B \cap A \subset \mathcal{C}(A)$. This projection is denoted by $\pi_A(B)$. For any $\beta \in \mathcal{C}(B)$ we also define $\pi_A(\beta) := \pi_A(B)$. These projections always have bounded diameter [13] and given any intersecting subsurfaces $A, B, C \subset S$ we define the *projection distance* as

$$d_A(B, C) := \operatorname{diam}_{\mathcal{C}(A)}(\pi_A(B) \cup \pi_A(C)).$$

For a subgroup $G < \mathrm{MCG}(S)$, we say that a subsurface $K \subset S$ is *$G$-nondisplaceable* if $gK$ and $K$ cannot be homotoped to be disjoint for any $g \in G$. Note that if a subsurface $K$ is $G$-nondisplaceable, subsurface projections are always defined between $G$-translates of $K$.

# 3 Proof of Theorem A

Fix an infinite-type surface $S$, different from the Loch Ness monster.

Figure 1 shows some examples of $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable subsurfaces that will be used in the following lemma.
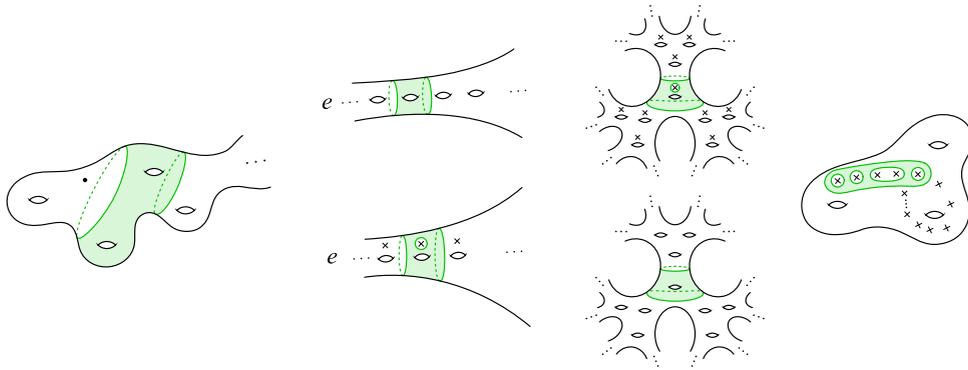
Figure 1: The subsurfaces of Lemma 2.

**Lemma 2** *If $S$ is an infinite-type surface, different from the Loch Ness monster, it contains infinitely many finite-type $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable subsurfaces, which are pairwise disjoint and nonaccumulating. Moreover, the subsurfaces can be chosen to be nonexceptional.*

**Proof** We split the proof into cases. In each we describe a finite-type $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable subsurface such that we can clearly find infinitely many copies with the required properties.

**Case 1** Suppose $S$ is the once-punctured Loch Ness monster. Then note that any separating curve $\alpha$ which separates the two ends cannot be mapped disjointly from itself by any mapping class (because it bounds a nondisplaceable subsurface). As a consequence, for any $g \geq 1$, any genus-$g$ subsurface with two boundary components separating the two ends is nondisplaceable.

**Case 2** Suppose $S$ has at least two nonplanar ends. Note that by the argument in [14, Proposition 6.3], any separating curve such that both complementary components have infinite genus is $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable.

• If $S$ has at least one nonplanar end — denoted by $e$ — which is isolated in $\mathrm{Ends}(S)$, for any $g \geq 1$, any genus-$g$ subsurface with two separating boundary components, each of which cuts off a surface containing only the end $e$, is $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable.

• If $S$ has at least one nonplanar end — denoted by $e$ — which is isolated in $\mathrm{Ends}_g(S)$ but not in $\mathrm{Ends}(S)$, for any $g \geq 1$, any genus-$g$ subsurface with three separating boundary components, two of which cut off a subsurface whose only nonplanar end is $e$ and the third one cuts off a planar surface, is $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable.

• If no nonplanar end is isolated in $\mathrm{Ends}_g(S)$, $\mathrm{Ends}_g(S)$ is a Cantor set. If it contains an end $e$ that is not accumulated by planar ends, we choose a genus-$g$ subsurface with two separating boundary components and no planar ends, so that each complementary component has infinite genus, for $g \geq 1$. Otherwise, we choose a genus-$g$ subsurface with three separating boundary components, so that two complementary components have infinite genus and one is a planar subsurface, for $g \geq 1$.

**Case 3** Suppose $S$ has no nonplanar ends. We can then choose any $n$-holed sphere whose boundary curves are separating, so that there is at least one end in each complementary component, for $n \geq 5$. □

Fix a finite-type $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable subsurface $\Sigma \subset S$ and let $\mathcal{Y}$ be the $\overline{\mathrm{MCG}_c(S)}$-orbit of $\Sigma$. As $\Sigma$ is $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable, any two surfaces in $\mathcal{Y}$ have intersecting boundaries — in particular, subsurface projections $\pi_A$ between surfaces in $\mathcal{Y}$ are always defined. Moreover, by [2; 13], there is some constant $\mu > 0$ such that for every $A, B, C \in \mathcal{Y}$

- at most one of $d_A(B, C)$, $d_B(A, C)$ and $d_C(A, B)$ is bigger than $\mu$, and
- $|\{D \in \mathcal{Y} \mid d_D(A, B) > \mu\}|$ is finite.

See [8, Lemma 3.8] for details on checking these in the infinite-type case. We can therefore run the projection complex machinery to deduce (see [3, Proposition 2.7]):

**Proposition 3** $\overline{\mathrm{MCG}_c(S)}$ *acts by isometries on a hyperbolic graph* $\mathcal{C}(\mathcal{Y})$ *so that for every* $A, B \in \mathcal{Y}$ *with* $A \neq B$:

(1) $\mathcal{C}(A)$ *is isometrically embedded as a convex set in* $\mathcal{C}(\mathcal{Y})$ *and the images of* $\mathcal{C}(A)$ *and* $\mathcal{C}(B)$ *are disjoint.*

(2) *The inclusion*
$$\bigsqcup_{C \in \mathcal{Y}} \mathcal{C}(Y) \hookrightarrow \mathcal{C}(\mathcal{Y})$$
*is* $\overline{\mathrm{MCG}_c(S)}$-*equivariant.*

(3) *The nearest-point projection to* $\mathcal{C}(A)$ *sends* $\mathcal{C}(B)$ *to a bounded set, which is at uniformly bounded distance from* $\pi_A(B)$.

(4) *If* $g \in \overline{\mathrm{MCG}_c(S)}$ *is supported on* $A$ *and the restriction is pseudo-Anosov, and* $\Gamma$ *is the subgroup of* $\overline{\mathrm{MCG}_c(S)}$ *given by elements leaving* $A$ *invariant and preserving the stable and unstable foliations of* $g$, *then* $(\overline{\mathrm{MCG}_c(S)}, \mathcal{C}(\mathcal{Y}), g, \Gamma)$ *satisfies WWPD.*

Furthermore, the same proof as [3, Lemma 2.8] yields:

**Lemma 4** *Let* $\tau$ *be a multitwist about a finite multicurve* $\mu$. *Then for every* $A \in \mathcal{Y}$ *there is a vertex* $v_A$ *of* $\mathcal{C}(\mathcal{Y})$ *such that the nearest-point projection to* $\mathcal{C}(A)$ *of the* $\tau$ *orbit of* $v_A$ *is uniformly bounded. In particular, if* $\tau$ *is hyperbolic, its virtual quasiaxis can intersect* $\mathcal{C}(A)$ *only in a bounded-length segment.*

**Proof** If $\mu \cap A = \varnothing$, $\tau$ fixes any element of $\mathcal{C}(A)$, so it's elliptic. Otherwise, let $v_A$ be an element of $A \cap \mu \neq \varnothing$. Then the nearest-point projection of $\tau^n(v_A)$ to $\mathcal{C}(A)$ is a uniformly bounded distance from $\pi_A(\tau^n(v_A))$, which is defined to be $\pi_A(\partial(\tau^n(A))$. But this is at bounded distance from $\tau^n(\mu) \cap A = \mu \cap A$, so the projection of $\tau^n(v_A)$ is at uniformly bounded distance from $A \cap \mu$ for every $n$. This proves the first statement of the lemma. The second statement follows as in the proof of [3, Lemma 2.8]. □

As a consequence, we can apply [3, Corollary 3.2] to deduce:

**Proposition 5** *Let $\Sigma$ be a $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable subsurface of finite type and $f$ a mapping class that is a pure chiral pseudo-Anosov mapping class of $\Sigma$ of sufficiently large translation length and the identity on the complement. Then there is a homogeneous quasimorphism $\varphi \colon \overline{\mathrm{MCG}_c(S)} \to \mathbb{R}$ of defect $\Delta \le 12$ such that $|\varphi(f^n)| \to \infty$ and $|\varphi(\tau)| \le \Delta$ for every multitwist $\tau$ and every element $\tau$ acting elliptically on $\mathcal{C}(\overline{\mathrm{MCG}_c(S)} \cdot \Sigma)$.*

**Proof** Let $\mathcal{Y} = \overline{\mathrm{MCG}_c(S)} \cdot \Sigma$. The bounds on $\varphi(f)$ and $\varphi(\tau)$, for $\tau$ acting elliptically on $\mathcal{C}(\mathcal{Y})$, are given by [3, Corollary 3.2]. The only thing we need to check is that $|\varphi(\tau)| \le \Delta$ for every multitwist $\tau$. But a multitwist $\tau$ associated to an integral weighted multicurve $\mu = \sum_{i \in I} n_i \alpha_i$ can be written as a product of two multitwists, $\tau_1$ and $\tau_2$, where $\tau_1$ is associated to the integral weighted multicurve

$$\mu_1 = \sum_{i : \alpha_i \cap \Sigma \ne \varnothing} n_i \alpha_i$$

and $\tau_2$ to the integral weighted multicurve

$$\mu_2 = \sum_{i : \alpha_i \cap \Sigma = \varnothing} n_i \alpha_i.$$

Then $\tau_2$ acts elliptically on $\mathcal{C}(\mathcal{Y})$, so $\varphi(\tau_2) = 0$. By Lemma 4, if $\tau_1$ doesn't act elliptically on $\mathcal{C}(\mathcal{Y})$, its virtual quasiaxis has small projections. Thus, by the construction of the quasimorphism $\varphi$ in [3, Proposition 3.1], if the translation length of $f$ is larger than the projection bound from Lemma 4, we have $\varphi(\tau_1) = 0$. As a consequence $|\varphi(\tau)| \le \Delta$. $\qquad\square$

**Proof of Theorem A** Suppose first that $S$ is not the Loch Ness monster. We will construct an element $F$ of $\mathcal{I}(S)$ which is not a finite product of multitwists. Since $F$ is in $\overline{\mathrm{MCG}_c(S)}$ (by construction, or by the fact that $\mathcal{I}(S) \subset \overline{\mathrm{MCG}_c(S)}$ by [1]), this will also show that multitwists don't generate the closure of the compactly supported mapping class group. By Lemma 2, we can find pairwise disjoint nonaccumulating $\overline{\mathrm{MCG}_c(S)}$-nondisplaceable finite-type subsurfaces $\Sigma_n$. For every $n$, fix a mapping class $F_n$ of $S$ supported on $\Sigma_n$ which restricts to a chiral pseudo-Anosov mapping class in the Torelli group of $\Sigma_n$. Then for any $n$, by Proposition 5 (after potentially passing to a power in order to increase the translation length), we can find a homogeneous quasimorphism $\varphi_n \colon \overline{\mathrm{MCG}_c(S)} \to \mathbb{R}$ with defect at most $\Delta = 12$ that is unbounded on powers of $F_n$ and bounded by $\Delta$ on all multitwists or elements acting elliptically on $\mathcal{C}(\overline{\mathrm{MCG}_c(S)} \cdot \Sigma_n)$. Choose powers $k_n$ so that

$$|\varphi_n(F_n^{k_n})| \to \infty,$$

which exist because by assumption $|\varphi_n(F_n)| > 1$ for every $n$. Define

$$F = \prod_{n \in \mathbb{N}} F_n^{k_n}.$$

For every $n$, $\prod_{m \ne n} F_m^{k_m}$ acts elliptically on $\mathcal{C}(\overline{\mathrm{MCG}_c(S)} \cdot \Sigma_n)$, so

$$|\varphi_n(F)| = \left| \varphi_n \left( F_n^{k_n} \circ \prod_{m \ne n} F_m^{k_m} \right) \right| \ge |\varphi_n(F_n^{k_n})| - \Delta \to \infty.$$

If $F$ were a product of $k$ multitwists $\tau_1, \ldots, \tau_k$, then for any $n$,

$$|\varphi_n(F)| = |\varphi_n(\tau_k \circ \cdots \circ \tau_1)| \le 2k\Delta,$$

which gives a contradiction.

Suppose now that $S$ is the Loch Ness monster and fix a point $x \in S$. By the Birman exact sequence [8, Appendix], the kernel of the surjection

$$\mathcal{I}(S \smallsetminus \{x\}) \to \mathcal{I}(S)$$

is the fundamental group of $S$ and is therefore generated by twists. In particular, if $\mathcal{I}(S)$ is contained in the subgroup generated by multitwists, so is the Torelli group of the once-punctured Loch Ness monster, a contradiction. By [14], $\mathrm{MCG}(S) = \overline{\mathrm{MCG}_c(S)}$ and $\mathrm{MCG}(S \smallsetminus \{x\}) = \overline{\mathrm{MCG}_c(S \smallsetminus \{x\})}$, so the same argument applied to $\mathrm{MCG}(S)$ and $\mathrm{MCG}(S \smallsetminus \{x\})$ proves the result for the closure of the compactly supported mapping class group. $\qquad \square$

# References

[1] **J Aramayona**, **T Ghaswala**, **A E Kent**, **A McLeay**, **J Tao**, **R R Winarski**, *Big Torelli groups: generation and commensuration*, Groups Geom. Dyn. 13 (2019) 1373–1399 MR Zbl

[2] **J A Behrstock**, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. 10 (2006) 1523–1578 MR Zbl

[3] **M Bestvina**, **K Bromberg**, **K Fujiwara**, *Stable commutator length on mapping class groups*, Ann. Inst. Fourier (Grenoble) 66 (2016) 871–898 MR Zbl

[4] **J S Birman**, *Mapping class groups and their relationship to braid groups*, Comm. Pure Appl. Math. 22 (1969) 213–238 MR Zbl

[5] **T E Brendle**, **B Farb**, *Every mapping class group is generated by 6 involutions*, J. Algebra 278 (2004) 187–198 MR Zbl

[6] **D Calegari**, **L Chen**, *Normal subgroups of big mapping class groups*, Trans. Amer. Math. Soc. Ser. B 9 (2022) 957–976 MR Zbl

[7] **M Dehn**, *Die Gruppe der Abbildungsklassen: das arithmetische Feld auf Flächen*, Acta Math. 69 (1938) 135–206 MR Zbl

[8] **G Domat**, *Big pure mapping class groups are never perfect*, Math. Res. Lett. 29 (2022) 691–726 MR Zbl With an appendix by R Dickmann and Domat

[9] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) MR Zbl

[10] **S P Humphries**, *Generators for the mapping class group*, from "Topology of low-dimensional manifolds", Lecture Notes in Math. 722, Springer (1979) 44–47 MR Zbl

[11] **W B R Lickorish**, *A finite set of generators for the homeotopy group of a 2–manifold*, Proc. Cambridge Philos. Soc. 60 (1964) 769–778 MR Zbl

[12] **J Malestein**, **J Tao**, *Self-similar surfaces: involutions and perfection*, Michigan Math. J. 74 (2024) 485–508 MR Zbl

[13] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, II: Hierarchical structure*, Geom. Funct. Anal. 10 (2000) 902–974 MR Zbl

[14] **P Patel**, **N G Vlamis**, *Algebraic and topological properties of big mapping class groups*, Algebr. Geom. Topol. 18 (2018) 4109–4142 MR Zbl

[15] **A Putman**, *Cutting and pasting in the Torelli group*, Geom. Topol. 11 (2007) 829–865 MR Zbl

[16] **P Scott**, *Subgroups of surface groups are almost geometric*, J. Lond. Math. Soc. 17 (1978) 555–565 MR Zbl

[17] **B Wajnryb**, *Mapping class group of a surface is generated by two elements*, Topology 35 (1996) 377–383 MR Zbl

*University of Michigan, Department of Mathematics*
*Ann Arbor, MI, United States*

*Laboratoire d'Analyse et de Mathématiques Appliqueées, Université Paris-Est Créteil Val de Marne*
*Créteil, France*

*Department Mathematisches Institut, LMU München*
*Munich, Germany*

domatg@umich.edu, federica.fanoni@u-pec.fr, hensel@math.lmu.de

# Cofibrantly generated model structures for functor calculus

LAUREN BANDKLAYDER
JULIA E BERGNER
RHIANNON GRIFFITHS
BRENDA JOHNSON
REKHA SANTHANAM

Model structures for many kinds of functor calculus can be obtained by applying a theorem of Bousfield to a suitable category of functors. We give a general criterion for when model categories obtained via this approach are cofibrantly generated. Our examples recover the homotopy functor and $n$-excisive model structures of Biedermann and Röndigs, but also include a model structure for the discrete functor calculus of Bauer, Johnson, and McCarthy.

18F50, 18N40, 55U35; 18D15, 18D20

## 1 Introduction

Functor calculi have been used to produce significant results in a wide range of fields, beginning with applications of the homotopy functor calculus of Goodwillie [20] to algebraic $K$-theory [35] and $v_n$-periodic homotopy theory [1; 23]. In another direction, the manifold calculus of Goodwillie and Weiss [21; 48] and orthogonal calculus of Weiss [47] have been used to study embeddings and immersions of manifolds, as well as characteristic classes [38], and spaces of knots [33; 42].

The essential data of a functor calculus is a means of approximating a functor $F$ with a tower of functors under $F$

$$\cdots \leftarrow P_0 F \Longleftarrow P_1 F \Longleftarrow \cdots \Longleftarrow P_n F \longleftarrow P_{n+1} F \Longleftarrow \cdots$$

with $F$ mapping down to each $P_n F$.

in a way analogous to the Taylor series for a function; indeed, we refer to them as *Taylor towers*. In particular, each $P_n F$ should be thought of as some kind of *degree $n$ approximation* to $F$. In fact, in a functor calculus, the natural transformations $p_n F : F \to P_n F$ induce weak equivalences $P_n(p_n F), p_n(P_n F) : P_n F \to P_n^2 F$ which in turn can be used to show that $F \to P_n F$ is, in a categorical sense, the best degree $n$ approximation to $F$.

With a suitable model structure on a category of functors in place, the endofunctor $P_n$ that assigns to a functor its degree $n$ approximation can be used to build a new model structure in which $F \to G$ is a

weak equivalence precisely when $P_n F \to P_n G$ is a weak equivalence in the original model structure. This approach has been employed by Barnes and Oman [5] in the case of the orthogonal calculus, and by Biedermann, Chorny, and Röndigs [9; 10] and by Barnes and Eldred [3] in the case of the homotopy functor calculus. It has led to means by which different functor calculi can be extended to new contexts [45] and compared to one another [4; 44], and provided the context in which to strengthen classifications of homogenous degree $n$ functors [10].

One can also work in a more flexible setting than model categories, such as some well-behaved model for $(\infty, 1)$-categories, which is the approach taken by Lurie [34] and Pereira [37]. As their work demonstrates, one can develop a good general theory for functor calculus in this kind of setting as well, but we do not take that approach here, since we are particularly interested in the nuances of the model structures.

The starting point for this paper was the desire to develop a similar approach for the discrete functor calculus of Bauer, McCarthy, and Johnson [6]. Our hope is to use the model structure for discrete calculus developed here to recast existing comparisons between the homotopy functor and discrete calculi, such as that found in [6], in model category-theoretic terms, and to develop a model category classification of homogeneous degree $n$ functors similar to the one for $n$-homogeneous functors in [10]. In addition, we have sought to define the model structure for discrete calculus in a manner that could be conveniently adapted for general classes of functor calculi. In particular, we would like for it to be applicable to those that can be defined using towers of comonads in the same way as the discrete calculus does. A general approach to such calculi is being developed by Brenda Johnson and Kathryn Hess. To this end, we structure this paper around Theorems 1.1 and 1.2 below, and we show how to use them to build cofibrantly generated model structures corresponding to both the homotopy functor and discrete calculi.

Model structures for abelian versions of the discrete functor calculus were established by Renaudin [39] and Richter [40]. In general, the existence of such functor calculus model structures is guaranteed by a theorem of Bousfield and Friedlander [12], which was modified by Bousfield for his work on telescopic homotopy theory [11]. We use the following version of this theorem, derived from Bousfield's, in the present work; see Theorem 4.2 and Corollary 4.4. The three versions differ slightly in the third axiom.

**Theorem 1.1**  *Let $\mathcal{M}$ be a right proper model category together with an endofunctor $Q \colon \mathcal{M} \to \mathcal{M}$ and a natural transformation $\eta \colon \mathrm{id} \Rightarrow Q$ satisfying the following axioms*:

  (A1)  *the endofunctor $Q$ preserves weak equivalences in $\mathcal{M}$*;

  (A2)  *the maps $\eta_{QX}, Q(\eta_X) \colon QX \to Q^2 X$ are both weak equivalences in $\mathcal{M}$*; *and*

  (A3′)  *$Q$ preserves homotopy pullback squares*.

*Then there exists a right proper model structure, denoted by $\mathcal{M}_Q$, on the same underlying category as $\mathcal{M}$ with the same cofibrations as $\mathcal{M}$ and in which $X \to Y$ is a weak equivalence in $\mathcal{M}_Q$ when $QX \to QY$ is a weak equivalence in the original model structure on $\mathcal{M}$. Furthermore, if $\mathcal{M}$ has the structure of a simplicial model category, then so does $\mathcal{M}_Q$.*

A further version of this result by Stanculescu [43] drops the right properness assumption, but since we have found this condition essential to our arguments here, we retain it. We also note that the second axiom requires precisely the property described above that guarantees that $P_n F$ is, in a categorical sense, the best degree $n$ approximation to the functor $F$. In this sense, Bousfield and Friedlander's theorem and its variants seem tailor-made for constructing model structures for functor calculi. Localizations that are produced via these theorems are sometimes referred to as *Bousfield–Friedlander localizations*, and the endofunctors used to produce them are examples of what are often referred to in the literature as *Quillen idempotent monads*. For specificity here we use the terminology *Bousfield endofunctor* to emphasize that we are assuming the axioms of Theorem 1.1. We review Bousfield's version of the localization theorem and deduce Theorem 1.1 from it in Section 4, then apply it to obtain the hf- and $n$-excisive model structures of [10] in Sections 5 and 6, and the desired discrete degree $n$ model structure in Section 7.

However, applying Theorem 1.1 does not immediately give the additional structure of a cofibrantly generated model category. Facing the same challenge, Biedermann and Röndigs [10] develop a simplicially enriched version of Goodwillie's homotopy functor calculus in such a way that their model structure for $n$-excisive functors is cofibrantly generated. Because this additional structure is quite powerful, we want to employ a similar strategy to develop a degree $n$ model structure for discrete functor calculus.

Indeed, we develop a systematic way to show when model structures obtained from this theorem are cofibrantly generated, via criteria for when arguments of the kind used by Biedermann and Röndigs can be applied. As a result, we are able to present not only the new example of discrete functor calculus, but also to show that the homotopy functor and $n$-excisive model structures arise as consequences of a general theorem. Our hope is that, together with the presentation of each of these three model structures, this general result will shed light on the shared features across these different model structures and the types of technical details that must be checked in specific examples. With this goal in mind, even for the model structures that were already known, we emphasize some of the details in the arguments. At the same time, the general theorem should enable us quickly to establish that the model structures arising from the functor calculi in the forthcoming work of Johnson and Hess, which are constructed using methods similar to those used for the discrete calculus, are also cofibrantly generated.

In particular, in Biedermann and Röndigs' cofibrant generation proofs, they rely on the construction of equivalent formulations of functors such as $P_n$ in terms of representables. Although our proofs do rely on their methods, we are able to circumvent the need for these replacements and use their original formulations instead. In our approach, these representable functors appear in the form of what we call *test morphisms* for a Bousfield endofunctor in the following result, which is restated more precisely as Theorem 8.3 below.

**Theorem 1.2** *Suppose that* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ *is a cofibrantly generated right proper model structure on a category of simplicial functors in which all fibrations are also levelwise fibrations, with some modest additional assumptions on* $\mathcal{C}$ *and* $\mathcal{D}$. *If* $Q$ *is a Bousfield endofunctor of* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ *that admits a collection of test morphisms, then the model structure on* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$ *is cofibrantly generated.*

In practice, test morphisms represent maps in $\mathcal{C}$ that, for any functor $F$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, are converted to weak equivalences by $QF$. After proving this theorem, we identify test morphisms for the hf-, $n$-excisive, and degree $n$ model structures induced by Bousfield's theorem to prove that they are cofibrantly generated.

## 1.1 Outline of the paper

In Section 2 we review properties of right proper and simplicial model categories that will be used in subsequent sections, and we look more specifically at simplicial model categories of functors, which are the examples of interest in this paper. In Section 3 we summarize both the homotopy functor calculus of Goodwillie and the discrete functor calculus of Bauer, Johnson, and McCarthy. We present the localization techniques for model categories that we use in Section 4, and we apply them to get model structures for homotopy functors, $n$-excisive functors, and degree $n$ functors in Sections 5, 6, and 7, respectively. We present our main theorem for cofibrant generation in Section 8 and include the homotopy functor model structure there as a guiding example. We then apply the main theorem to the $n$-excisive model structure in Section 9 and to the degree $n$ model structure in Section 10.

## Acknowledgements

We would like to thank the anonymous referee for several suggestions for the improvement of the paper.

## 2 Background on simplicial model categories

In this section, we review some key facts about model categories that we need, focusing on some features of right proper and simplicial model categories. Of particular importance are some results about simplicial model categories that are difficult to find in the literature.

We begin with a number of properties of homotopy pullbacks in right proper model categories that we use throughout this paper, stated here for ease of reference. The first provides some standard means of constructing and comparing homotopy pullbacks.

**Proposition 2.1** [24, 13.3.4, 13.3.8] *Let $\mathcal{M}$ be a right proper model category, and let*

$$
\begin{array}{ccccc}
X & \longrightarrow & Z & \longleftarrow & Y \\
{\scriptstyle \simeq}\downarrow & & \downarrow{\scriptstyle \simeq} & & \downarrow{\scriptstyle \simeq} \\
X' & \longrightarrow & Z' & \longleftarrow & Y'
\end{array}
$$

be a diagram in $\mathcal{M}$ where the vertical maps are all weak equivalences. The induced map between the homotopy pullback of the diagram $X \to Z \leftarrow Y$ and the homotopy pullback of the diagram $X' \to Z' \leftarrow Y'$ is a weak equivalence. Moreover, if, in any diagram $X \to Z \leftarrow Y$ in $\mathcal{M}$, at least one of the maps is a fibration, then the induced map from the ordinary pullback of the diagram to the homotopy pullback is a weak equivalence.

The next result is a kind of "two-out-of-three property" for homotopy pullbacks.

**Proposition 2.2** [24, 13.3.15] *Let $\mathcal{M}$ be a right proper model category. If the right-hand square in the diagram*

$$
\begin{array}{ccccc}
X & \longrightarrow & Y & \longrightarrow & Z \\
\downarrow & & \downarrow & & \downarrow \\
U & \longrightarrow & V & \longrightarrow & W
\end{array}
$$

*is a homotopy pullback square, then the left-hand square is a homotopy pullback if and only if the composite square is a homotopy pullback.*

Finally, the following result gives a useful criterion for identifying homotopy pullback squares together with a nice property of homotopy pullback squares. It is proved in [36, 3.3.11(1ab)] for the category of topological spaces, but the argument works for any right proper model category.

**Proposition 2.3** *Consider a commutative square in a right proper model category $\mathcal{M}$:*

$$
\begin{array}{ccc}
X & \longrightarrow & Y \\
\downarrow & & \downarrow \\
U & \longrightarrow & V
\end{array}
$$

(1) *If the square is a homotopy pullback square and $Y \to V$ is a weak equivalence, then $X \to U$ is a weak equivalence.*

(2) *If both vertical arrows in the square are weak equivalences, then the square is a homotopy pullback square.*

The model categories we consider in the paper are simplicially enriched model categories. We recall the definition here, with the axioms labeled according to the usual convention. A more detailed overview of simplicially enriched model categories can be found in the prequel to this paper [2].

**Definition 2.4** [24, 9.1.6] A *simplicial model category* is a model category $\mathcal{M}$ that is also a simplicial category, ie a category enriched in the closed monoidal category $\mathcal{S}$ of simplicial sets, such that the following two axioms hold.

(SM6)   The category $\mathcal{M}$ is both tensored and cotensored over $\mathcal{S}$. In particular, there are natural isomorphisms

$$\operatorname{Map}(X \otimes K, Y) \cong \operatorname{Map}(X, Y)^K \cong \operatorname{Map}(X, Y^K).$$

for each pair of objects $X$ and $Y$ in $\mathcal{M}$ and simplicial set $K$.

(SM7)   If $i\colon K \to L$ is a cofibration of simplicial sets and $p\colon X \to Y$ is a fibration in $\mathcal{M}$, then the induced morphism

$$X^L \to X^K \times_{Y^K} Y^L$$

is a fibration in $\mathcal{D}$ that is a weak equivalence if either $i$ or $p$ is.

**Remark 2.5**   [24, 9.3.7]   Assuming axiom (SM6) holds, axiom (SM7) is equivalent to

(SM7$'$)   If $i\colon K \to L$ is a cofibration of simplicial sets and $j\colon X \to Y$ is a cofibration in $\mathcal{M}$, then the induced morphism

$$X \otimes L \amalg_{X \otimes K} Y \otimes K \to Y \otimes L$$

is a cofibration in $\mathcal{D}$ that is a weak equivalence if either $i$ or $j$ is.

**Example 2.6**   [24, 9.1.13]   The model category $\mathcal{S}$ of simplicial sets can be regarded as a simplicial model category. Given any simplicial sets $X$ and $K$, we can take the tensor structure $X \otimes K$ to be the cartesian product $X \times K$ and the cotensor structure $X^K$ to be $\operatorname{Map}(K, X)$.

We recall a result that is well known in the model category of spaces, but holds in any simplicial model category; see, for example, [15, Section 2.4].

**Proposition 2.7**   *Let $\mathcal{M}$ be a simplicial model category and let $f\colon A \to B$ be a morphism in $\mathcal{M}$ with $A$ cofibrant. Then there exists a factorization $f = g\, i_f$ of $f$ where $i_f$ is a cofibration and $g$ is a simplicial homotopy equivalence.*

Our main objects of study are categories whose objects are themselves functors between fixed categories.

**Convention 2.8**   From this point onward we assume that $\mathcal{C}$ is an essentially small simplicial category and that $\mathcal{D}$ is a cofibrantly generated right proper simplicial model category. We denote by $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$ the category whose objects are simplicial functors $\mathcal{C} \to \mathcal{D}$ and whose morphisms are simplicial natural transformations. Simplicial natural transformations are defined analogously to simplicial functors; see [30, Section 1.2] or [2, 2.10].

The following result tells us that there exists a model structure on the category of simplicial functors $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$ induced by the model structure on $\mathcal{D}$. We need some further technical assumptions because we consider simplicial functors between simplicial categories, compared to results such as [24, 11.6.1] for ordinary functors. We omit some of these assumptions in the following theorem, but refer the reader to [22, 4.32] for the precise statement.

**Theorem 2.9** [2, 4.2, 4.3, 5.10]  *Assuming $\mathcal{D}$ satisfies some mild conditions on the set of generating acyclic cofibrations, the category $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ has a cofibrantly generated right proper model structure, called the **projective model structure**, in which a morphism $F \to G$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is a weak equivalence or a fibration if it is one levelwise, ie $FA \to GA$ is a weak equivalence or fibration, respectively, in $\mathcal{D}$ for all objects $A$ of $\mathcal{C}$.*

*Furthermore, it has the structure of a simplicial model category. The tensor and cotensor structures are defined by $(F \otimes K)(A) = FA \otimes K$ and $(F^K)(A) = FA^K$, respectively, for each object $A$ in $\mathcal{C}$.*

Given that the category $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is enriched in $\mathcal{S}$, one might wonder whether it is also enriched in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$. However, such an enrichment would require $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ to be a closed monoidal category, which is not true in general. Nevertheless, we show in Proposition 2.13 and Lemma 2.15 that $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ does satisfy generalizations of axioms (SM6) and (SM7) from Definition 2.4, and therefore enjoys many of the properties of a model category enriched over $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$.

**Definition 2.10**  For each object $C$ of $\mathcal{C}$, the simplicial functor $R^C : \mathcal{C} \to \mathcal{S}$ *represented by $C$* sends each object $A$ of $\mathcal{C}$ to the simplicial set $\mathrm{Map}_{\mathcal{C}}(C, A)$. Dually, the simplicial functor $R_C : \mathcal{C}^{\mathrm{op}} \to \mathcal{S}$ sends each object $A$ of $\mathcal{C}$ to $\mathrm{Map}_{\mathcal{C}}(A, C)$.

The following definition is critical to many of our arguments in this paper.

**Definition 2.11**  Let $F : \mathcal{C} \to \mathcal{D}$ and $X : \mathcal{C} \to \mathcal{S}$ be simplicial functors. The *evaluated cotensor* of the pair $(F, X)$ is the equalizer

$$F^X := \int_A FA^{XA} \to \prod_A FA^{XA} \rightrightarrows \prod_{A,B} FB^{(XA^{\mathrm{Map}_{\mathcal{C}}(A,B)})}$$

in $\mathcal{D}$ whose parallel morphisms are described explicitly in [2, Section 3]. Here, $A$ and $B$ range over objects of $\mathcal{C}$.

**Remark 2.12**  • As suggested by the notation $F^X = \int_A FA^{XA}$, the evaluated cotensor can be thought of as a generalization of an ordinary end. In the special case when $\mathcal{D} = \mathcal{S}$, the assignment $F^X(A, B) = FB^{XA}$ defines a simplicial bifunctor $\mathcal{C}^{\mathrm{op}} \times \mathcal{C} \to \mathcal{S}$ whose end is precisely $F^X$.

• This construction is originally due to Biedermann and Röndigs [10, 2.5], who use the notation $\mathbf{hom}(X, F)$. We have chosen the name "cotensor" and the notation we use here to emphasize the fact that it behaves much like an ordinary cotensor; see Proposition 2.13.

Recall that a simplicial adjunction satisfies the usual condition for an adjunction, but using mapping spaces rather than hom sets. It can be thought of as a special case of a $\mathcal{V}$-adjunction for categories enriched in a general $\mathcal{V}$; see [30, Section 1.11]. The reference for the following proposition, which is an analogue of axiom (SM6) of Definition 2.4 but for the evaluated cotensor, is stated at that level of generality.

**Proposition 2.13** [2, 4.4, 4.5, 4.6] *For any simplicial functors $X : \mathcal{C} \to \mathcal{S}$ and $F : \mathcal{C} \to \mathcal{D}$ and object $D$ of $\mathcal{D}$ there are simplicial adjunctions*

$$
\begin{array}{ccc}
\xymatrix{\mathcal{D} & \perp & \mathrm{Fun}(\mathcal{C},\mathcal{D})} & \xymatrix{\mathcal{D} & \perp & \mathrm{Fun}(\mathcal{C},\mathcal{S})^{\mathrm{op}}} & \xymatrix{\mathrm{Fun}(\mathcal{C},\mathcal{S}) & \perp & \mathrm{Fun}(\mathcal{C},\mathcal{D})}
\end{array}
$$

with upper arrows $X \otimes -$, $\mathrm{Map}_{\mathcal{D}}(-, F)$, $- \otimes D$ and lower arrows $(-)^X$, $F^{(-)}$, $\mathrm{Map}_{\mathcal{D}}(D, -)$

*where $(X \otimes D)(C) := XC \otimes D$ and $\mathrm{Map}_{\mathcal{D}}(D, F)(A) := \mathrm{Map}_{\mathcal{D}}(D, FA)$. In particular, there are simplicial natural isomorphisms*

$$\mathrm{Map}_{\mathrm{Fun}(\mathcal{C},\mathcal{D})}(X \otimes D, F) \cong \mathrm{Map}_{\mathrm{Fun}(\mathcal{C},\mathcal{S})}(X, \mathrm{Map}_{\mathcal{D}}(D, F)) \cong \mathrm{Map}_{\mathcal{D}}(D, F^X).$$

The evaluated cotensor also provides an enriched version of the (co)Yoneda lemma, stated below. The proof follows an argument similar to the one outlined in [30, Section 2.4] for the case where $\mathcal{C}$ is a $\mathcal{V}$-category for some closed monoidal category $\mathcal{V}$ and $\mathcal{D} = \mathcal{V}$.

**Lemma 2.14** [2, 3.5, 3.9] *Let $F : \mathcal{C} \to \mathcal{D}$ be a simplicial functor. For each object $C$ of $\mathcal{C}$ there is a natural isomorphism*

$$FC \cong F^{R^C} = \int_A FA^{R^C(A)} = \int_A FA^{\mathrm{Map}_{\mathcal{C}}(C,A)}.$$

*Dually, we have*

$$FC \cong F^{R_C} = \int^A FA^{R_C(A)} = \int^A FA \otimes \mathrm{Map}_{\mathcal{C}}(C, A).$$

The next lemma establishes that a version of the condition (SM7) in Definition 2.4 holds for the evaluated cotensor.

**Lemma 2.15** *Consider $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ and $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$, each equipped with the projective model structure. If $X \to Y$ is a cofibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ and $F \to G$ is a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, then $F^Y \to F^X \times_{G^X} G^Y$ is a fibration in $\mathcal{D}$ that is a weak equivalence if either $X \to Y$ or $F \to G$ is.*

**Proof** Let $F \to G$ be a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, so that $FA \to GA$ is a fibration in $\mathcal{D}$ for any object $A$ of $\mathcal{C}$. If $K \to L$ is a cofibration in $\mathcal{S}$, then by (SM7) the pullback-corner map for the ordinary cotensor

$$FA^L \to FA^K \times_{GA^K} GA^L$$

is a fibration in $\mathcal{D}$. Thus, if $D \to E$ is an acyclic cofibration in $\mathcal{D}$, there exists a lift in any diagram of the form

$$
\xymatrix{
D \ar[r] \ar[d]_{i}^{\simeq} & FA^L \ar[d] \\
E \ar[r] \ar@{-->}[ur] & FA^K \times_{GA^K} GA^L
}
$$

Using the natural isomorphism $\mathrm{Map}_{\mathcal{D}}(E, FA)^L \cong \mathrm{Map}_{\mathcal{D}}(E, FA^L)$ given by taking $\mathcal{M} = \mathcal{D}$ in axiom (SM6), we have a commutative diagram

$$
\begin{array}{ccc}
K & \longrightarrow & \mathrm{Map}_{\mathcal{D}}(E, FA) \\
\uparrow\downarrow & \nearrow & \downarrow \\
L & \longrightarrow & \mathrm{Map}_{\mathcal{D}}(D, FA) \times_{\mathrm{Map}_{\mathcal{D}}(D, GA)} \mathrm{Map}_{\mathcal{D}}(E, GA)
\end{array}
$$

in $\mathcal{S}$. Since $K \to L$ is assumed to be a cofibration, and using the fact that $A$ is an arbitrary object of $\mathcal{C}$, we can conclude that the map

$$
\mathrm{Map}_{\mathcal{D}}(E, F) \to \mathrm{Map}_{\mathcal{D}}(D, F) \times_{\mathrm{Map}_{\mathcal{D}}(D, G)} \mathrm{Map}_{\mathcal{D}}(E, G)
$$

is an acyclic fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$. It follows that if $X \to Y$ is a cofibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$, then a lift exists in any diagram of the form

$$
\begin{array}{ccc}
X & \longrightarrow & \mathrm{Map}_{\mathcal{D}}(E, F) \\
\uparrow\downarrow & \nearrow & \downarrow{\scriptstyle \simeq} \\
Y & \longrightarrow & \mathrm{Map}_{\mathcal{D}}(D, F) \times_{\mathrm{Map}_{\mathcal{D}}(D, G)} \mathrm{Map}_{\mathcal{D}}(E, G)
\end{array}
$$

Applying axiom (SM6) once more, we equivalently obtain a lift

$$
\begin{array}{ccc}
D & \longrightarrow & F^Y \\
i \downarrow{\scriptstyle \simeq} & \nearrow & \downarrow \\
E & \longrightarrow & F^X \times_{G^X} G^Y
\end{array}
$$

so we have thus shown that the map

$$
F^Y \to F^X \times_{G^X} G^Y
$$

is a fibration in $\mathcal{D}$. The proof that $F^Y \to F^X \times_{G^X} G^Y$ is also a weak equivalence if either $X \to Y$ or $F \to G$ is a weak equivalence is analogous. $\qquad\square$

**Corollary 2.16** (1) *If $X \to Y$ is a cofibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ and $F$ is a fibrant object in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, then $F^Y \to F^X$ is a fibration in $\mathcal{D}$ that is a weak equivalence if $X \to Y$ is.*

(2) *If $X$ is a cofibrant object in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ and $F \to G$ is a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, then $F^X \to G^X$ is a fibration in $\mathcal{D}$ that is a weak equivalence if $F \to G$ is.*

**Remark 2.17** Lemma 2.15 and Corollary 2.16 also hold when the model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is a localization of the projective model structure with the same cofibrations, and whose fibrations form a subclass of the projective fibrations. In particular, they apply to model structures obtained by applying the localization given in Theorem 4.2, from which we get our main examples in this paper, as well as any left Bousfield localization in the sense of [24].

Finally, we include a result about the interaction between the evaluated cotensor and homotopy (co)limits.

**Proposition 2.18** [2, 6.5]  *Let $\mathcal{I}$ be a small category and $X$ an $\mathcal{I}$-diagram in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$. Let $\mathcal{D}$ be a cofibrantly generated simplicial model category such that the projective model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ exists. Then for any simplicial functor $F : \mathcal{C} \to D$, there is a natural isomorphism*

$$F^{\mathrm{hocolim}_i \, X_i} \cong \mathrm{holim}_i \, F^{X_i}.$$

# 3   Background on functor calculus

In this section, we give a brief review of two different flavors of functor calculus: Goodwillie's homotopy functor calculus [20] and the discrete functor calculus of Bauer, Johnson, and McCarthy [6]. Each of these functor calculi associates to a functor $F$ a "Taylor" tower of functors under $F$
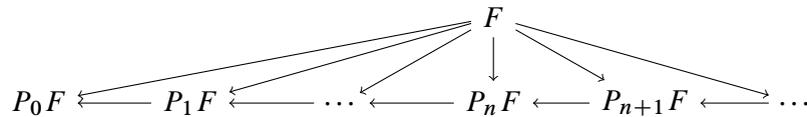


that are analogous to Taylor polynomial approximations to a function; the $n^{\mathrm{th}}$ term $P_n F$ is "degree $n$" in some sense and serves as a universal degree $n$ approximation to $F$. We review the construction of the $n^{\mathrm{th}}$ term in the towers for Goodwillie's calculus and the discrete calculus, and establish some properties that will be needed in subsequent sections.

We use the following notation. For any $n \geq 1$, let $\mathcal{P}(\boldsymbol{n})$ denote the poset of subsets of the set $\boldsymbol{n} = \{1, \ldots, n\}$. Let $\mathcal{P}_0(\boldsymbol{n})$ denote the poset of nonempty subsets of $\boldsymbol{n}$ and $\mathcal{P}_{\leq 1}(\boldsymbol{n})$ the full subcategory of $\mathcal{P}(\boldsymbol{n})$ spanned by the subsets of cardinality less than or equal to 1. An *$n$-cubical diagram* in a category $\mathcal{C}$ is a functor $\mathcal{X} : \mathcal{P}(\boldsymbol{n}) \to \mathcal{C}$. For more details about $n$-cubical homotopy theory, see [19; 31, Section 4.1; 36].

## 3.1   Goodwillie calculus

We describe the first version of functor calculus that we consider, namely, the calculus of homotopy functors as defined by Goodwillie. This calculus grew out of work on pseudoisotopy theory and Waldhausen's algebraic $K$-theory of spaces, and has led to important results in $K$-theory and homotopy theory; see, for example, [1; 8; 14; 32; 35]. The Taylor tower for this calculus is constructed in [20], which serves as the reference for the definitions in this subsection.

Instead of restricting our source and target categories to spaces and spectra as Goodwillie did, we work in the more general setting of simplicial model categories, an approach that Kuhn developed in [31]. In comparison to their approaches, we impose more structure on our category of functors and define $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ to be the category whose objects are simplicial functors, and use a simplicial model structure on it to define an $n$-excisive model structure in Section 6. To start, we describe the construction of the

Taylor tower as developed by Goodwillie and Kuhn and then, at the end of this section, we describe the adjustments that need to be made for our purposes.

**Convention 3.1**  In addition to the assumptions on $\mathcal{C}$ and $\mathcal{D}$ from Convention 2.8, we assume in this subsection that $\mathcal{C}$ is a simplicial subcategory of a simplicial model category that is closed under finite homotopy colimits and has a terminal object $*_{\mathcal{C}}$.

For an object $A$ in $\mathcal{C}$ and a subset $U$ of $\boldsymbol{n}$, the *fiberwise join* $A * U$ is the homotopy colimit in $\mathcal{C}$ of the $\mathcal{P}_{\leq 1}(U)$-diagram that assigns to $\varnothing$ the object $A$ and assigns to each one-element set $\{i\}$ the terminal object $*_{\mathcal{C}}$. If $U$ is a one-element set, then $A * U$ is the simplicial cone on $A$, and in general, $A * U$ consists of $|U|$ copies of the simplicial cone on $A$ glued along $A$. For a fixed object $A$ in $\mathcal{C}$, $A * -$ defines a functor from $\mathcal{P}(\boldsymbol{n})$ to $\mathcal{C}$. This functor plays a key role in the definition of the functors $P_n F : \mathcal{C} \to \mathcal{D}$ appearing in Goodwillie's Taylor tower.

**Definition 3.2**  For any functor $F : \mathcal{C} \to \mathcal{D}$, define the functors $T_n F$ and $P_n F$ by

$$T_n F(A) = \operatorname*{holim}_{U \in \mathcal{P}_0(\boldsymbol{n+1})} F(A * U) \quad \text{and} \quad P_n F(A) = \operatorname*{hocolim}_k T_n^k F(A),$$

where the homotopy colimit is over a sequential diagram obtained by repeatedly applying the transformation $F \xrightarrow{t_n F} T_n F$ induced by the canonical map $A = A * \varnothing \to A * U$ for each $U$. These maps also yield a natural transformation $p_n F : F \to P_n F$.

**Remark 3.3**  The functor $P_n F$ is the *n-excisive approximation* of $F$. More precisely, it is an *n-excisive functor*, that is, it takes strongly homotopy cocartesian $(n+1)$-cubical diagrams in $\mathcal{C}$ to homotopy cartesian $(n+1)$-cubes in $\mathcal{D}$, and the induced natural transformation $p_n F : F \to P_n F$ is appropriately universal with respect to this property [20, 1.8].

## 3.2  Discrete functor calculus

Abelian functor calculus is a functor calculus developed for algebraic settings that builds Taylor towers using certain classical constructions, such as the stable derived functors defined by Dold and Puppe [13] and stable homology of modules over a ring $R$ with coefficients in a ring $S$ as defined by Eilenberg and Mac Lane [16; 17], as the first terms in its Taylor towers. It has been used to study rational algebraic $K$-theory, Hochschild homology, and other algebraic constructions, and to make connections between functor calculus and the tangent and cartesian differential categories of Blute, Cockett, Cruttwell, and Seely [7; 27; 28; 29]. Discrete functor calculus arose as an adaptation of abelian functor calculus for functors of simplicial model categories. In general, the $n^{\text{th}}$ term in the discrete Taylor tower satisfies a weaker degree $n$ condition than the corresponding term in the Goodwillie calculus, but the towers agree for functors that commute with realization [6, 4.11, 5.4].

For this kind of functor calculus, we need the notion of a stable model category. First, we recall that a model category is *pointed* if it has a zero object, that is, if the initial and terminal objects coincide.

**Definition 3.4** A model category $\mathcal{D}$ is *stable* if it is pointed and its homotopy category is triangulated.

We do not go into the details of triangulated categories here, but refer the reader to [26, Chapter 7] for an overview aimed at understanding stable model categories. For our purposes, the most important feature of stable model categories is that homotopy pullback squares are the same as homotopy pushout squares.

The stability condition is used to guarantee the $n^{\text{th}}$ term in the discrete calculus tower for a functor is really a degree $n$ functor, a result that is straightforward using the agreement of homotopy pushout and pullback diagrams in stable model categories; see [6, 4.5, 4.6, 5.4].

**Convention 3.5** In addition to the assumptions on $\mathcal{C}$ and $\mathcal{D}$ from Convention 2.8, we assume in this subsection that $\mathcal{C}$ has finite coproducts and a terminal object $*_{\mathcal{C}}$, and that $\mathcal{D}$ is a stable model category in the sense of Definition 3.4. In particular, $\mathcal{D}$ has a zero object that we denote by $\star_{\mathcal{D}}$.

As we did in the previous subsection, we describe the construction of degree $n$ approximations in the discrete calculus as it was done originally in [6], and then, in the next subsection, explain what is needed to ensure that the terms in the discrete Taylor tower are simplicial functors.

To define the $n^{\text{th}}$ term in the discrete Taylor tower, we use a comonad constructed from the iterated homotopy fiber of a particular $(n+1)$-cubical diagram.

**Definition 3.6** Let $\mathcal{X} \colon \mathcal{P}(2) \to \mathcal{D}$ be a 2-cubical diagram. Its *iterated homotopy fiber* is

$$
\text{ifiber}(\mathcal{X}) := \text{hofiber}\left(\text{hofiber}\left(\begin{array}{c} X_{\varnothing} \\ \downarrow \\ X_2 \end{array}\right) \to \text{hofiber}\left(\begin{array}{c} X_1 \\ \downarrow \\ X_{12} \end{array}\right)\right).
$$

In other words, we take the homotopy fibers of the two vertical maps, which produces the horizontal map between those homotopy fibers; the *iterated* fiber is just the homotopy fiber of this induced map. The iterated fiber of an $n$-cubical diagram $\mathcal{X} \colon \mathcal{P}(\boldsymbol{n}) \to \mathcal{D}$ can be defined analogously. More details can be found in [6, 3.1; 36, Sections 3.4 and 5.5].

For an object $A$ in $\mathcal{C}$ and a functor $F \colon \mathcal{C} \to \mathcal{D}$, let $F^{\boldsymbol{n}}(A, -)$ be the $\mathcal{P}(\boldsymbol{n})$-diagram that assigns to $U \subseteq \boldsymbol{n}$ the object

$$
F^{\boldsymbol{n}}(A, U) := F\Big(\coprod_{i \in \boldsymbol{n}} A_i(U)\Big),
$$

where

**(3.7)**
$$
A_i(U) := \begin{cases} A & \text{if } i \notin U, \\ *_{\mathcal{C}} & \text{if } i \in U. \end{cases}
$$

Furthermore, given a functor $F \colon \mathcal{C} \to \mathcal{D}$, we define the functor $F_n^{\mathcal{P}_0(2)} \colon \mathcal{C} \times \mathcal{P}_0(2)^{\times n} \to \mathcal{D}$ by

$$
F_n^{\mathcal{P}_0(2)}(A; (S_1, \dots, S_n)) := \begin{cases} F^{\boldsymbol{n}}(A, \varphi(S_1, \dots, S_n)) & \text{if } S_i \neq \{2\} \text{ for all } i, \\ \star_{\mathcal{D}} & \text{otherwise,} \end{cases}
$$

where

**(3.8)**
$$
\varphi(S_1, \dots, S_n) = \{i \mid S_i = \{1, 2\}\}.
$$

**Example 3.9** The diagram $F_2^{\mathcal{P}_0(\mathbf{2})}(A, -)$ is given by the diagram

$$
\begin{array}{ccccc}
F(A \amalg A) & \longrightarrow & F(*_{\mathcal{C}} \amalg A) & \longleftarrow & \star_{\mathcal{D}} \\
\downarrow & & \downarrow & & \downarrow \\
F(A \amalg *_{\mathcal{C}}) & \longrightarrow & F(*_{\mathcal{C}} \amalg *_{\mathcal{C}}) & \longleftarrow & \star_{\mathcal{D}} \\
\uparrow & & \uparrow & & \uparrow \\
\star_{\mathcal{D}} & \longrightarrow & \star_{\mathcal{D}} & \longleftarrow & \star_{\mathcal{D}}
\end{array}
$$

**Definition 3.10** For any functor $F : \mathcal{C} \to \mathcal{D}$ and any object $A$ of $\mathcal{C}$, define

$$\perp_n F(A) := \operatorname{holim}_{(S_1, \ldots, S_n)} F_n^{\mathcal{P}_0(\mathbf{2})}(A; (S_1, \ldots, S_n)),$$

where the homotopy limit is taken over the category $\mathcal{P}_0(\mathbf{2})^{\times n}$.

Observe that $\perp_2 F(A)$ is the iterated homotopy fiber of the diagram

$$
\begin{array}{ccc}
F(A \amalg A) & \longrightarrow & F(*_{\mathcal{C}} \amalg A) \\
\downarrow & & \downarrow \\
F(A \amalg *_{\mathcal{C}}) & \longrightarrow & F(*_{\mathcal{C}} \amalg *_{\mathcal{C}})
\end{array}
$$

More generally, $\perp_n F(A)$ is equivalent to the iterated homotopy fiber of the $\mathcal{P}(\mathbf{n})$-diagram that assigns to the set $U \subseteq \mathbf{n}$ the object $F^{\mathbf{n}}(A, U)$; see [6, Section 3.1] for details.

**Definition 3.11** A functor $F : \mathcal{C} \to \mathcal{D}$ is *degree $n$* provided that for any $(n+1)$-tuple of objects $X = (X_1, \ldots, X_{n+1})$ in $\mathcal{C}$, the iterated homotopy fiber of the diagram

$$\mathbf{n} + \mathbf{1} \supseteq U \longmapsto F\left( \coprod_{i \in \mathbf{n}+\mathbf{1}} X_i(U) \right)$$

where

$$X_i(U) = \begin{cases} X_i & \text{if } i \notin U, \\ *_{\mathcal{C}} & \text{if } i \in U, \end{cases}$$

is weakly equivalent to the terminal object in $\mathcal{D}$. This iterated homotopy fiber is called the $(n+1)^{st}$ *cross effect of $F$ at $X$* and is denoted by $cr_{n+1}F(X)$. When $X = (X, \ldots, X)$ for an object $X$ in $\mathcal{C}$, $cr_{n+1}F(X) = \perp_{n+1} F(X)$.

To define a degree $n$ approximation, we note that $\perp_n$ can be given the additional structure of a comonad. Recall that a *comonad* $(\perp, \Delta, \varepsilon)$ acting on a category $\mathcal{A}$ consists of an endofunctor $\perp : \mathcal{A} \to \mathcal{A}$ together with natural transformations $\Delta : \perp \to \perp\perp$ and $\varepsilon : \perp \to \operatorname{id}_{\mathcal{A}}$ satisfying certain identities. For an object $A$ in $\mathcal{A}$, there is an associated simplicial object $\perp^* A$ given by

$$[k] \longmapsto \perp^{k+1} A,$$

whose face and degeneracy maps are defined using $\varepsilon$ and $\Delta$. See [46, Section 8.6] for more details, noting that the author uses the term "cotriple" for what we are calling a "comonad" here.

The indexing category $\mathcal{P}_0(2)^{\times n}$ used for the homotopy limit that defines the functor $\perp_n$ induces both a comultiplication $\Delta_n$ and a counit $\varepsilon_n$. Together with the natural transformations $\Delta_n$ and $\varepsilon_n$, $\perp_n$ defines a comonad. Proving that $\perp_n$ is a comonad requires proving that certain diagrams are strictly commutative. Since $\perp_n$ is defined using a homotopy limit and different models for homotopy limits may only agree up to weak equivalence, one needs to fix a choice of model for homotopy limits. For the proof that $\perp_n$ is a comonad in [6, Section 3], one also needs to ensure that this model satisfies certain standard properties, which are listed in Appendix A of [6], up to isomorphism, rather than weak equivalence. As shown in [6, Appendix A], the standard model for homotopy limits in [24, 18.1.8] satisfies these conditions.

Once one has established the existence of a comonad structure for $\perp_n$ with this choice of model for the homotopy limit, and shown that the functors $\Gamma_n F$ in Definition 3.12 are in fact degree $n$, one can use a different model for homotopy colimit to construct $\Gamma_n F$, provided it is weakly equivalent to the original. Thus we can make the following definition.

**Definition 3.12**   For a functor $F : \mathcal{C} \to \mathcal{D}$, the *degree n approximation* of $F$ at an object $A$ is given by

$$\Gamma_n F(A) := \mathrm{hocofiber}\big(|\perp_{n+1}^{*+1} F(A)| \to F(A)\big),$$

where $\perp_{n+1} F(A)$ is as in Definition 3.10, and $|\perp_{n+1}^{*+1} F(A)|$ is the homotopy colimit over $\Delta^{\mathrm{op}}$ of the standard simplicial object associated to the comonad $\perp_{n+1}$ acting on $F$.

**Remark 3.13**   The functor $\Gamma_n F$ is degree $n$ in the sense of Definition 3.11, which is a weaker condition than the one for $n$-excisive functors, satisfied by the $n^{\mathrm{th}}$ term in Goodwillie's Taylor tower. However, the proof that $\Gamma_n F$ is degree $n$ uses a standard extra degeneracy argument for simplicial objects, whereas Goodwillie describes his proof that $P_n F$ is $n$-excisive as "somewhat opaque" [20, 1.10]. As is the case with the Goodwillie tower, one can show that there is a natural transformation $F \to \Gamma_n F$ that is universal (up to homotopy) among degree $n$ functors with natural transformations from $F$ [6, Section 5].

## 3.3   Polynomial approximations for simplicial functors

To establish the existence of the $n$-excisive and degree $n$ model structures in Sections 6 and 7, we need to know that $P_n F$ and $\Gamma_n F$ are simplicial functors when $F$ is, that is, that $P_n F$ and $\Gamma_n F$ are objects in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ as we have defined it. We also need to know that $P_n$ and $\Gamma_n$ preserve weak equivalences in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. These conditions require knowing that the models we use for the homotopy limit and homotopy colimit in $\mathcal{D}$ preserve weak equivalences and are suitably simplicial.

**Remark 3.14**   Here, we use the models for homotopy limits and homotopy colimits for simplicial model categories as described in [24, 18.1.2, 18.1.8]. To guarantee that our homotopy limits and colimits preserve levelwise weak equivalences, which need not be the case for these models [24, Chapter 18], we use simplicial fibrant and cofibrant replacement functors. Such replacements exist for any cofibrantly generated simplicial model category by [41, 24.2], so it is safe to assume that there are models for homotopy limits and homotopy colimits in $\mathcal{D}$ that are simplicial and preserve levelwise weak equivalences of diagrams.

We conclude the following.

**Proposition 3.15** *For a simplicial functor $F$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, both $P_n F$ and $\Gamma_n F$ are simplicial functors.*

**Proof** As defined, the functor $P_n F$ is a composite of fiberwise joins, the original functor $F$, homotopy limits, and a sequential homotopy colimit. Each of these components is simplicial. In the case of the fiberwise joins, we can use [2, 7.15]. The other cases are consequences of the assumptions we have made. A similar argument gives us the result for $\Gamma_n F$. $\square$

We conclude this section by noting that using models for homotopy (co)limits that involve precomposition with (co)fibrant replacement functors affects the natural transformations from $F$ to $P_n F$ and $\Gamma_n F$; for example, we may only have natural transformations from the cofibrant replacement of $F$ to $P_n F$ and $\Gamma_n F$. We describe how to circumvent such issues in Sections 6 and 7.

# 4 Bousfield's $Q$-theorem

In this section, we recall the approach to localization that we use in this paper, due to Bousfield and Friedlander. Because the basic input is a model category equipped with an endofunctor $Q$, and we use Bousfield's later variant, we refer to this result simply as "Bousfield's $Q$-theorem". As is typical for a localization, we would like to produce a new model structure with more weak equivalences; we begin by defining the weak equivalences that we obtain from the endofunctor $Q$.

**Definition 4.1** Let $\mathcal{M}$ be a model category, together with an endofunctor $Q: \mathcal{M} \to \mathcal{M}$. A morphism $f$ in $\mathcal{M}$ is a *$Q$-equivalence* if it is a weak equivalence after applying the functor $Q$; that is, if $Q(f)$ is a weak equivalence in $\mathcal{M}$.

We now state Bousfield's $Q$-theorem.

**Theorem 4.2** [11, 9.3, 9.5, 9.7] *Let $\mathcal{M}$ be a right proper model category together with an endofunctor $Q: \mathcal{M} \to \mathcal{M}$ and a natural transformation $\eta: \mathrm{id} \Rightarrow Q$ satisfying the following axioms:*

(A1) *the endofunctor $Q$ preserves weak equivalences in $\mathcal{M}$;*

(A2) *the maps $\eta_{QX}, Q(\eta_X): QX \to Q^2 X$ are both weak equivalences in $\mathcal{M}$; and*

(A3) *given a pullback square*

$$
\begin{array}{ccc}
V & \xrightarrow{\ f\ } & X \\
\downarrow & & \downarrow{\scriptstyle h} \\
W & \xrightarrow[\ g\ ]{} & Y
\end{array}
$$

*if $X$ and $Y$ are fibrant, $h: X \to Y$ is a fibration, $\eta_X: X \to QX$ and $\eta_Y: Y \to QY$ are weak equivalences, and $g: W \to Y$ is a $Q$-equivalence, then $f: V \to X$ is a $Q$-equivalence.*

*Then there exists a right proper model structure, denoted by $\mathcal{M}_Q$, on the same underlying category as $\mathcal{M}$ with weak equivalences the $Q$-equivalences and the same cofibrations as $\mathcal{M}$. Furthermore, if $\mathcal{M}$ has the structure of a simplicial model category, then so does $\mathcal{M}_Q$.*

The next two results provide modifications of the hypotheses of this theorem. The first shows that it suffices to check axiom (A2) on fibrant objects, while the second gives us a way to show that (A3) is satisfied.

**Corollary 4.3** *Let $\mathcal{M}$ be a right proper model category together with an endofunctor $Q\colon \mathcal{M} \to \mathcal{M}$ and natural transformation $\eta\colon \mathrm{id} \Rightarrow Q$ satisfying axiom (A1). If (A2) holds for all fibrant objects in $\mathcal{M}$, then (A2) holds for all objects in $\mathcal{M}$.*

**Proof** Let $X$ be an object in $\mathcal{M}$ and let $\beta\colon X \xrightarrow{\;\cong\;} X'$ be a fibrant replacement. By (A1), we know that $Q\beta\colon QX \to QX'$ and $Q^2\beta\colon Q^2X \to Q^2X'$ are weak equivalences. Consider the commutative diagram

$$
\begin{array}{ccc}
QX & \xrightarrow{\;Q\beta\;} & QX' \\
{\scriptstyle Q\eta_X}\downarrow & & \downarrow{\scriptstyle Q\eta_{X'}} \\
Q^2X & \xrightarrow[\;Q^2\beta\;]{} & Q^2X'
\end{array}
$$

obtained by applying $Q$ to the naturality diagram for $\eta$. Since $X'$ is fibrant, the right vertical arrow is a weak equivalence by our assumption that (A2) holds for fibrant objects. As the horizontal arrows are also weak equivalences, we see that $Q\eta_X$ is also a weak equivalence. The fact that $\eta_{QX}$ is a weak equivalence is proved in a similar fashion by replacing the vertical maps in the diagram above with $\eta_{QX}$ and $\eta_{QX'}$, respectively. $\qquad\square$

**Corollary 4.4** *Let $\mathcal{M}$ be a right proper model category together with an endofunctor $Q\colon \mathcal{M} \to \mathcal{M}$ and natural transformation $\eta\colon \mathrm{id} \Rightarrow Q$ satisfying axioms (A1) and (A2) above. If*

(A3′)  *$Q$ preserves homotopy pullback squares,*

*then axiom (A3) holds, and in particular the model structure $\mathcal{M}_Q$ exists and is right proper.*

**Proof** If $Q$ preserves fibrations, this statement is a direct consequence of the right properness condition. To see why it holds more generally, assume that $Q$ preserves homotopy pullback squares and that we have a commutative square as in Theorem 4.2 satisfying the hypotheses of (A3). Since one of those hypotheses is that $h\colon X \to Y$ is a fibration, and $\mathcal{M}$ is right proper, such a square is necessarily a homotopy pullback square by Proposition 2.1. By (A3′), the diagram

$$
\begin{array}{ccc}
QV & \xrightarrow{\;Qf\;} & QX \\
\downarrow & & \downarrow{\scriptstyle Qh} \\
QW & \xrightarrow[\;Qg\;]{} & QY
\end{array}
$$

is a homotopy pullback square. Again by the hypotheses of (A3), $Qg$ is a weak equivalence, and hence $Qf$ is a weak equivalence by Proposition 2.3, establishing (A3). $\qquad\square$

For ease of reference, we make the following definition.

**Definition 4.5** A *Bousfield endofunctor* $Q\colon \mathcal{M} \to \mathcal{M}$ is an endofunctor together with a natural transformation $\eta\colon \mathrm{id} \Rightarrow Q$ satisfying axioms (A1) and (A2) from Theorem 4.2 and axiom (A3′) from Corollary 4.4.

We refer to the fibrations in $\mathcal{M}_Q$, which are completely determined by the cofibrations and weak equivalences as described in Theorem 4.2, as *$Q$-fibrations*. We include the following useful characterizations of them, the first of which was part of Bousfield's original statement of Theorem 4.2. The second part of this result is a consequence of the first via Proposition 2.3. Some similar results are proved in [18, X.4] under the further assumption of left properness.

**Proposition 4.6** [11, 9.3] *Let $\mathcal{M}$ be a right proper model category together with an endofunctor $Q\colon \mathcal{M} \to \mathcal{M}$ and natural transformation $\eta\colon \mathrm{id} \Rightarrow Q$ satisfying axioms* (A1), (A2), *and* (A3).

(1) *A map $f\colon X \to Y$ in $\mathcal{M}$ is a $Q$-fibration if and only if it is a fibration in $\mathcal{M}$ and the induced diagram*

$$
\begin{array}{ccc}
X & \xrightarrow{\ \eta_X\ } & QX \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle Qf} \\
Y & \xrightarrow[\ \eta_Y\ ]{} & QY
\end{array}
$$

*is a homotopy pullback square in $\mathcal{M}$.*

(2) *Assume that $\eta_*\colon * \to Q(*)$ is a weak equivalence where $*$ is the terminal object in $\mathcal{M}$. Then an object $X$ of $\mathcal{M}$ is $Q$-fibrant if and only if it is fibrant in $\mathcal{M}$ and the map $\eta_X\colon X \to QX$ is a weak equivalence in $\mathcal{M}$.*

We refer to the homotopy pullback condition in the first part of the corollary as the *$Q$-fibration condition*, for ease of referring to it in later sections.

# 5 The homotopy functor model structure

In this section, we apply Theorem 4.2 to obtain a model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ in which the fibrant objects are homotopy functors, or functors that preserve weak equivalences. This model structure was established in [10], but we treat it in some detail so that we can refer to the techniques in more generality. While it is a first step in developing a model structure for $n$-excisive functors, this example is also key because it requires us to look closely at many of the ingredients used.

We begin by identifying the necessary assumptions on the categories $\mathcal{C}$ and $\mathcal{D}$ for such a model structure to exist, and one of those conditions requires the existence of a well-behaved simplicial fibrant replacement

functor. While fibrant replacements exist in any model category, and they can be taken to be functorial in nice cases such as cofibrantly generated model structures, it is important for our purposes that they also preserve the simplicial structure. So, our first goal is to describe such a functor, for which we recall the following definition. Although we assume we are working in a model category, since that is the context in which we use it, this definition can be applied to any cocomplete category.

**Definition 5.1** Let $\mathcal{M}$ be a model category. A set $J$ of maps in $\mathcal{M}$ *permits the small object argument* if there exists a cardinal $\kappa$ such that for every regular cardinal $\lambda \geq \kappa$ and every $\lambda$-sequence

$$A_0 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow \cdots \longrightarrow A_\beta \longrightarrow \cdots \qquad (\beta < \lambda)$$

in $\mathcal{B}$ such that each $A_\beta \to A_{\beta+1}$ is a transfinite composition of pushouts of maps in $J$, the map

$$\mathrm{colim}_{\beta < \lambda} \, \mathrm{Hom}_{\mathcal{M}}(U, A_\beta) \to \mathrm{Hom}_{\mathcal{M}}(U, \mathrm{colim}_{\beta < \lambda} A_\beta)$$

is an isomorphism whenever $U$ is the domain of some map in $J$.

**Remark 5.2** If $J$ permits the small object argument and $U$ is in the domain of some map in $J$, then any map $U \to \mathrm{colim}_{\beta < \lambda} A_\beta$ factors through some $A_\alpha$.

The following result is due to Shulman [41, 24.2], but since this case is not proved in full detail in that paper, and we need to use some of the techniques elsewhere, we give a brief outline of the proof here.

**Theorem 5.3** *There exists a simplicial fibrant replacement functor on any cofibrantly generated simplicial model category.*

**Proof** Let $\mathcal{M}$ be a cofibrantly generated simplicial model category, and let $J$ denote a set of generating acyclic cofibrations. For any object $A$ of $\mathcal{M}$, define $A_0 = A$ and, for each $\beta$, define $A_{\beta+1}$ via the pushout

**(5.4)**
$$
\begin{array}{ccc}
\coprod_{U \to V} \mathrm{Map}_{\mathcal{M}}(U, A_\beta) \otimes U & \longrightarrow & A_\beta \\
\downarrow & & \downarrow{\scriptstyle \varphi_\beta^A} \\
\coprod_{U \to V} \mathrm{Map}_{\mathcal{M}}(U, A_\beta) \otimes V & \xrightarrow{\iota_\beta^A} & A_{\beta+1}
\end{array}
$$

where the coproducts are taken over the set of generating acyclic cofibrations. We define the fibrant replacement of $A$ to be the transfinite sequential colimit $\mathbb{F}A := \mathrm{colim}_\beta A_\beta$. By setting $K = \varnothing$ and $L = \mathrm{Map}_{\mathcal{M}}(U, A_\beta)$ in (S7$'$) of Remark 2.5, we see that each $\mathrm{Map}_{\mathcal{M}}(U, A_\beta) \otimes U \to \mathrm{Map}_{\mathcal{M}}(U, A_\beta) \otimes V$ is an acyclic cofibration. Then, using the fact that acyclic cofibrations are closed under small coproducts and are stable under pushouts, we see that each $\varphi_\beta^A \colon A_\beta \to A_{\beta+1}$ is an acyclic cofibration. It follows that the composite $\varphi_A \colon A = A_0 \to \mathrm{colim}_\beta A_\beta = \mathbb{F}A$ is also an acyclic cofibration; see [24, 10.3.4].

It is a now a straightforward exercise to show that the unique map $\mathbb{F}A \to *_{\mathcal{M}}$ is a fibration, that $\mathbb{F}$ is a simplicial functor, and that this factorization is functorial. $\qquad \square$

**Remark 5.5** Observe that for any cofibrantly generated simplicial model category $\mathcal{M}$, the functorial fibrant replacement functor satisfies the following properties.

(i) If $A$ is cofibrant, then $\mathbb{F}A$ is cofibrant, following from the fact that $A \to \mathbb{F}A$ is an acyclic cofibration.

(ii) If $A \to B$ is a weak equivalence, then $\mathbb{F}A \to \mathbb{F}B$ is a weak equivalence by the two-out-of-three property.

(iii) It follows from (i) and (ii) that if $A \to B$ is a weak equivalence between cofibrant objects, then $\mathbb{F}A \to \mathbb{F}B$ is a weak equivalence between objects that are both fibrant and cofibrant, and is therefore a simplicial homotopy equivalence.

(iv) Since fibrant replacements are sequential colimits they commute with all colimits in $\mathcal{M}$, ie $\mathbb{F}(\operatorname{colim}_i A_i) \cong \operatorname{colim}_i \mathbb{F}A_i$.

The remaining piece that we need before stating our assumptions is the following definition.

**Definition 5.6** An object $C$ in a simplicial category $\mathcal{C}$ is *simplicially finitely presentable* if the representable functor $R^C = \operatorname{Map}(C, -) \colon \mathcal{C} \to \mathcal{S}$ preserves filtered colimits, so

$$\operatorname{Map}_{\mathcal{C}}(C, \operatorname{colim}_i A_i) \cong \operatorname{colim}_i \operatorname{Map}_{\mathcal{C}}(C, A_i).$$

**Convention 5.7** In this subsection, we make the following assumptions in addition to those listed in Convention 2.8.

(1) Suppose that $\mathcal{C}$ is a full simplicial subcategory of a cofibrantly generated simplicial model category $\mathcal{B}$, and that the objects of $\mathcal{C}$ are all cofibrant and simplicially finitely presentable.

(2) For every object $C$ of $\mathcal{C}$, the fibrant replacement $\mathbb{F}C$ is a sequential colimit of objects $C_\beta$ in $\mathcal{C}$.

(3) Weak equivalences, fibrations, and homotopy pullbacks are preserved under sequential colimits in $\mathcal{D}$.

**Remark 5.8** Although this list may seem lengthy, we claim that these conditions are satisfied by many familiar model categories. For instance, condition (1) can be satisfied by taking $\mathcal{C}$ to be the full subcategory of $\mathcal{B}$ consisting of the cofibrant simplicially finitely presentable objects.

If we only asked for the objects $C_\beta$ to be in $\mathcal{B}$, then condition (2) is satisfied for any cofibrantly generated model category $\mathcal{B}$, given our construction of fibrant replacements via the small object argument. The issue is whether these objects are in the subcategory $\mathcal{C}$, not in the larger model category $\mathcal{B}$. It does hold in many examples of interest, for example taking $\mathcal{B}$ to be the usual model structure on topological spaces or simplicial sets and $\mathcal{C}$ the subcategory of finite spaces or finite simplicial sets, respectively.

Similarly, condition (3) holds in many nice cases, such as the categories of topological spaces and simplicial sets. For conditions under which sequential colimits preserve weak equivalences and fibrations, see [26, Section 7.4]; for a discussion of filtered colimits commuting with homotopy pullbacks, see [25].

The assumptions on $\mathcal{D}$ have the following consequence.

**Lemma 5.9** *Sequential colimits are weakly equivalent to sequential homotopy colimits in $\mathcal{D}$.*

**Proof** Let colim $D_i$ be a sequential colimit in $\mathcal{D}$ and let $D_0'$ be a cofibrant replacement of $D_0$. We can factor the resulting map $D_0' \to D_1$ as a cofibration $D_0' \to D_1'$ followed by an acyclic fibration $D_1' \to D_1$. We can then repeat the process with the map $D_1' \to D_2$, and so on, giving a commutative diagram

$$
\begin{array}{ccccccccc}
D_0' & \hookrightarrow & D_1' & \hookrightarrow & D_2' & \hookrightarrow & \cdots & \hookrightarrow & \operatorname{colim} D_i' \\
\simeq \downarrow & & \downarrow \simeq & & \downarrow \simeq & & & & \vdots \\
D_0 & \longrightarrow & D_1 & \longrightarrow & D_2 & \longrightarrow & \cdots & \longrightarrow & \operatorname{colim} D_i
\end{array}
$$

where the map $\operatorname{colim} D_i' \to \operatorname{colim} D_i$ is induced by the universal property of colimits and, by condition (3) of Convention 5.7, is a weak equivalence. Since the upper horizontal arrows in the diagram are all cofibrations between cofibrant objects, we have $\operatorname{hocolim} D_i = \operatorname{colim} D_i'$, proving the result.          □

Before proceeding to the homotopy functor model structure, we need to state one more consequence of our conventions. Recall that the simplicial left Kan extension of $F\colon \mathcal{C} \to \mathcal{D}$ along a simplicial functor $p\colon \mathcal{C} \to \mathcal{C}'$ is another simplicial functor $p_!(F)\colon \mathcal{C}' \to \mathcal{D}$ together with a simplicial natural transformation

$$
\begin{array}{ccc}
\mathcal{C} & \xrightarrow{\quad F \quad} & \mathcal{D} \\
 & {}_{p}\searrow \quad {}^{\Downarrow \eta} \quad \nearrow {}_{p_!(F)} & \\
 & \mathcal{C}' &
\end{array}
$$

that is appropriately universal with respect to this property. The functor $p_!(F)$ is given by the coend

$$
p_!(F)(B) = \int^A \operatorname{Map}_{\mathcal{C}'}(p(A), B) \otimes FA
$$

in $\mathcal{D}$ [30, 2.4].

**Remark 5.10** Observe that for the inclusion functor $i\colon \mathcal{C} \to \mathcal{B}$, it follows from Lemma 2.14 that

$$
i_!(F)(C) = \int^A \operatorname{Map}_{\mathcal{B}}(i(A), C) \otimes FA = \int^A \operatorname{Map}_{\mathcal{C}}(A, C) \otimes FA \cong FC
$$

for any object $C$ of $\mathcal{C}$, so $i_!(F) \circ i = F$, and $\eta$ is the identity transformation.

**Proposition 5.11** *For any simplicial functor $F\colon \mathcal{C} \to \mathcal{D}$, the simplicial left Kan extension $i_!(F)$ of $F$ along the inclusion $i\colon \mathcal{C} \to \mathcal{B}$ preserves filtered colimits.*

**Proof** Let $\operatorname{colim}_k B_k$ be a filtered colimit in $\mathcal{B}$. Then

$$
i_!(F)(\operatorname{colim}_k B_k) = \int^A \operatorname{Map}_{\mathcal{B}}(A, \operatorname{colim}_k B_k) \otimes FA \cong \int^A \operatorname{colim}_k (\operatorname{Map}_{\mathcal{B}}(A, B_k)) \otimes FA
$$

$$
\cong \operatorname{colim}_k \int^A \operatorname{Map}_{\mathcal{B}}(A, B_k) \otimes FA = \operatorname{colim}_k i_!(F)(B_k),
$$

where the first isomorphism holds by Convention 5.7(1), and the second isomorphism holds since both tensoring and coends commute with colimits; the former is a left adjoint and the latter is itself a colimit.          □

We now have the ingredients to define a means for replacing any functor $F : \mathcal{C} \to \mathcal{D}$ by a homotopy functor. This definition was first given in [10, 4.10].

**Definition 5.12** Let $F : \mathcal{C} \to \mathcal{D}$ be a simplicial functor. We define the simplicial functor $F^{\mathrm{hf}} : \mathcal{C} \to \mathcal{D}$ as the composite

$$\mathcal{C} \xhookrightarrow{\ i\ } \mathcal{B} \xrightarrow{\ \mathbb{F}\ } \mathcal{B} \xrightarrow{\ i_!(F)\ } \mathcal{D},$$

where $i : \mathcal{C} \to \mathcal{B}$ is the inclusion and $i_!(F)$ is the simplicial left Kan extension.

**Remark 5.13** Since $i_!(F) \circ i = F$, there exists a canonical natural transformation $\theta_F : F \Rightarrow F^{\mathrm{hf}}$ induced by the transformation $\varphi : \mathrm{id}_{\mathcal{B}} \Rightarrow \mathbb{F}$ defined in the proof of Theorem 5.3.

Let us first verify that this construction does in fact produce a homotopy functor, as suggested by the notation.

**Proposition 5.14** *If $F : \mathcal{C} \to \mathcal{D}$ is a simplicial functor, then $F^{\mathrm{hf}} : \mathcal{C} \to \mathcal{D}$ is a simplicial homotopy functor.*

**Proof** Consider a weak equivalence $f : A \to B$ in $\mathcal{C}$. Then, by Remark 5.5(iii) and Convention 5.7(1), $\mathbb{F}(f) : \mathbb{F}(A) \to \mathbb{F}(B)$ is a simplicial homotopy equivalence. Since simplicial homotopy equivalences are preserved by any simplicial functor, $i_!(F)(\mathbb{F}(f)) = F^{\mathrm{hf}}(f)$ is a simplicial homotopy equivalence, and therefore a weak equivalence. $\square$

In light of Proposition 4.6, the following fact is useful.

**Lemma 5.15** *The endofunctor $(-)^{\mathrm{hf}} : \mathrm{Fun}(\mathcal{C}, \mathcal{D}) \to \mathrm{Fun}(\mathcal{C}, \mathcal{D})$ preserves the terminal object.*

**Proof** Consider an object $C$ of $\mathcal{C}$. By Convention 5.7(2) we can express $\mathbb{F}C$ as a sequential colimit of objects $C_\beta$ in $\mathcal{C}$. Then

$$F^{\mathrm{hf}}(C) = (i_!(F) \circ \mathbb{F})(C) = i_!(F)(\operatorname{colim}_\beta C_\beta) \cong \operatorname{colim}_\beta\ i_!(F)(C_\beta),$$

where the isomorphism holds by Proposition 5.11. Now since each $C_\beta$ is an object of $\mathcal{C}$, we have

$$\operatorname{colim}_\beta i_!(F)(C_\beta) = \operatorname{colim}_\beta FC_\beta$$

by Remark 5.10. Therefore, if $FC_\beta$ is the terminal object for all $\beta$, then $F^{\mathrm{hf}}(C)$ is also terminal. $\square$

We now recall the hf-model structure [10, 4.14], which can be proved by showing that $(-)^{\mathrm{hf}}$ satisfies axioms (A1), (A2), and (A3′).

**Theorem 5.16** *Under the assumptions of Convention 5.7 the category $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ has a simplicial right proper model structure, denoted by $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\mathrm{hf}}$, in which $F \to G$ is a weak equivalence if $F^{\mathrm{hf}} \to G^{\mathrm{hf}}$ is a levelwise weak equivalence, and the cofibrations are precisely the projective cofibrations.*

We conclude this section with a few consequences of this result. They are all expected properties for localized model structures, and whose analogues are known for left Bousfield localizations, but that do not appear to be known in generality for Bousfield–Friedlander localization.

**Proposition 5.17**  *Let $\alpha\colon F \to G$ be a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. Then $\alpha$ is a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\mathrm{hf}}$ if and only if for each weak equivalence $A \to B$ in $\mathcal{C}$ the diagram*

$$
\begin{array}{ccc}
FA & \longrightarrow & FB \\
\downarrow & & \downarrow \\
GA & \longrightarrow & GB
\end{array}
$$

*is a homotopy pullback square. Moreover, a functor $F\colon \mathcal{C} \to \mathcal{D}$ is fibrant in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\mathrm{hf}}$ if and only if it is a levelwise fibrant homotopy functor.*

**Proof**  The first statement can be proved just as in [10, 4.15] using Convention 5.7, Proposition 4.6, and the fact that the maps $A_i \to A_{i+1}$ in the sequential colimit defining the fibrant replacement of Theorem 5.3 are acyclic cofibrations. The statement about hf-fibrant objects follows from the first statement and Proposition 2.3.                                                                                  □

The next result is useful for establishing that our model for an $n$-excisive approximation to a functor agrees with that of Goodwillie for all homotopy functors, which we prove in Section 6. In particular, we prove that this result holds for homotopy functors that need not be levelwise fibrant.

**Proposition 5.18**  *If $F$ is a homotopy functor, then $\theta_F\colon F \Rightarrow F^{\mathrm{hf}}$ is a levelwise weak equivalence.*

**Proof**  For any object $C$ of $\mathcal{C}$, we have $F^{\mathrm{hf}}(C) = \mathrm{colim}_\beta\, FC_\beta$ by Convention 5.7(2), Remark 5.10, and Proposition 5.11. We can then write the map $\theta_{FC}$ as

$$
FC = \mathrm{colim}_\beta\, FC \to \mathrm{colim}_\beta\, FC_\beta,
$$

so in particular, $\theta_{FC}$ is the sequential colimit of the morphisms $F(\varphi^C_{\beta-1}\cdots\varphi^C_1\cdot\varphi^C_0)\colon FC \to FC_\beta$. From the proof of Theorem 5.3, we know that each $\varphi^C_i$ is a weak equivalence. Since $F$ is a homotopy functor, it follows from Convention 5.7(3) that $\theta_{FC}$ is a weak equivalence.                                                □

# 6  The *n*-excisive model structure

In this section, we establish the existence of a model structure on the category $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ whose fibrant objects are the $n$-excisive functors for a given $n$. To ensure that the endofunctor $\widehat{P}_n$ we use to obtain this model structure satisfies axiom (A1) of Theorem 4.2, we make use of the functor $(-)^{\mathrm{hf}}$ of Definition 5.12. So, in addition to the assumptions that we made about the category $\mathcal{D}$ for the hf-model structure, we need to add another mild hypothesis to ensure that $\widehat{P}_n$ interacts nicely with the hf-model structure. Thus, we begin this section by making the following definition and then establishing some results for the hf-model structure that we need.

**Definition 6.1** An object $U$ of a category $\mathcal{D}$ is *finite relative to a subcategory* $\mathcal{A}$ if, for all limit ordinals $\lambda$ and $\lambda$-sequences

$$A_0 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow \cdots \longrightarrow A_\beta \longrightarrow \cdots \qquad (\beta < \lambda)$$

in $\mathcal{D}$ such that each $A_\beta \to A_{\beta+1}$ is in $\mathcal{A}$, the map

$$\mathrm{colim}_\beta \, \mathrm{Hom}_\mathcal{D}(U, A_\beta) \to \mathrm{Hom}_\mathcal{D}(U, \mathrm{colim}_\beta \, A_\beta)$$

is an isomorphism.

For this section, we assume the following.

**Convention 6.2** The set $J$ of generating cofibrations of $\mathcal{D}$ can be chosen such that the codomain of each map is finite relative to $J$, in addition to Conventions 3.1 and 5.7.

**Remark 6.3** Since $\mathcal{D}$ is cofibrantly generated by assumption, the set $J$ of generating cofibrations can always be chosen such that the domain of each map is finite relative to $J$. The extra condition on the codomain is satisfied, for example, by any finitely generated model category, as described in [26].

Our assumptions on $\mathcal{D}$ guarantee the following result. The proof is the same as the one for finitely generated model categories; see [26, 7.4.1].

**Lemma 6.4** *Suppose that $\mathcal{D}$ satisfies the conditions of Convention 6.2, $\lambda$ is an ordinal, $X, Y : \lambda \to \mathcal{D}$ are $\lambda$-sequences of acyclic cofibrations, and $p : X \to Y$ is a natural transformation such that $p_\beta : X_\beta \to Y_\beta$ is a fibration for all $\beta < \lambda$. Then $\mathrm{colim}\, p_\beta : \mathrm{colim}\, X_\beta \to \mathrm{colim}\, Y_\beta$ is a fibration that is a weak equivalence if each $p_\beta$ is.*

Now we establish the compatibility with the functor $(-)^{\mathrm{hf}}$ that we need.

**Proposition 6.5** *The endofunctor $(-)^{\mathrm{hf}}$ on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ preserves levelwise fibrations.*

**Proof** Let $\alpha : F \to G$ be a levelwise fibration. Since $F^{\mathrm{hf}}(C) = \mathrm{colim}_\beta \, FC_\beta$ for any object $C$ in $\mathcal{C}$, the component $\alpha_C^{\mathrm{hf}} : F^{\mathrm{hf}}(C) \to G^{\mathrm{hf}}(C)$ at $C$ is induced by the morphisms $\alpha_{C_\beta} : FC_\beta \to GC_\beta$ for all $\beta$. Since each of these maps $\alpha_{C_\beta}$ is a fibration in $\mathcal{D}$ by assumption, the result follows by Convention 5.7(3). $\qquad\square$

We now turn our attention to $n$-excisive approximations of functors, using Definition 3.2. In addition, we replace $F$ with $F^{\mathrm{hf}}$ as defined in Definition 5.12, making it possible to apply some results from [20] that only hold for homotopy functors.

**Definition 6.6** For a functor $F$ in $\text{Fun}(\mathcal{C}, \mathcal{D})$, we define the functor $\widehat{P}_n F$ by

$$\widehat{P}_n F := \text{colim}_k (F^{\text{hf}} \to T_n(F^{\text{hf}}) \to T_n^2(F^{\text{hf}}) \to \cdots \to T_n^k(F^{\text{hf}}) \to \cdots).$$

Using $T_n^\infty$ to represent this colimit, we have

$$\widehat{P}_n F = T_n^\infty F^{\text{hf}}.$$

There is a morphism $F \to \widehat{P}_n F$ given by the composite

$$F \xrightarrow{\ \theta_F\ } F^{\text{hf}} \xrightarrow{\ \iota_{F^{\text{hf}}}\ } T_n^\infty(F^{\text{hf}}) = \widehat{P}_n F,$$

where $\iota$ is the natural transformation whose component at $G$ is induced by the natural maps $G \to T_n G \to T_n^2 G \to \cdots$. Since this construction is natural in $F$, it induces a natural transformation $\widehat{p}_n : \text{id}_{\text{Fun}(\mathcal{C}, \mathcal{D})} \Rightarrow \widehat{P}_n$. We sometimes omit the subscript on $\widehat{p}_n$ when it can be understood from context.

We note that by Lemma 5.9, $\widehat{P}_n F$ is weakly equivalent to $P_n(F^{\text{hf}})$. One can prove that $\widehat{P}_n F$ is an $n$-excisive functor using a proof similar to that of [20, 1.8]. The argument uses the fact that $F^{\text{hf}}$ is a homotopy functor and is the reason we replace $F$ with $F^{\text{hf}}$ in the definition of $\widehat{P}_n F$. By Lemma 5.9 and Proposition 5.18, $\widehat{P}_n F$ is weakly equivalent to Goodwillie's construction of $P_n F$ when $F$ is a homotopy functor.

**Proposition 6.7** *The endofunctor $\widehat{P}_n$ on the category $\text{Fun}(\mathcal{C}, \mathcal{D})$ preserves hf-fibrations.*

**Proof** By Proposition 4.6, it suffices to prove that if $F \to G$ is an hf-fibration, then $\widehat{P}_n F \to \widehat{P}_n G$ is a projective fibration and

$$
\begin{array}{ccc}
\widehat{P}_n F & \longrightarrow & (\widehat{P}_n F)^{\text{hf}} \\
\downarrow & & \downarrow \\
\widehat{P}_n G & \longrightarrow & (\widehat{P}_n G)^{\text{hf}}
\end{array}
$$

is a homotopy pullback square. The fact that $\widehat{P}_n F \to \widehat{P}_n G$ is a projective fibration follows from the definition of $\widehat{P}_n$, Proposition 6.5, the fact that homotopy limits preserve fibrations, and Convention 5.7(3). To confirm that the diagram above is a homotopy pullback square, recall that $\widehat{P}_n F$ and $\widehat{P}_n G$ are homotopy functors by Proposition 5.14. It follows that the horizontal maps are weak equivalences by Proposition 5.18, and the square is a homotopy pullback by Proposition 2.3. $\qquad\square$

With the preceding proposition and definitions in place, we can establish the existence of the $n$-excisive model structure. This theorem is also proved by Biedermann and Röndigs in [10, 5.8] using a different, but naturally equivalent, model for the $n$-excisive approximation of a functor. To obtain cofibrant generation of this model structure in Section 9, we need to take a localization of the hf-model structure, but arguments similar to the ones presented here can be used to place an $n$-excisive model structure on $\text{Fun}(\mathcal{C}, \mathcal{D})$ with the projective model structure instead.

**Theorem 6.8** *Under the assumptions of Convention 6.2, the category* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ *has a right proper model structure in which a morphism* $F \to G$ *is a weak equivalence if* $\widehat{P}_n F \to \widehat{P}_n G$ *is a weak equivalence in the* hf-*model structure, and is a cofibration precisely if it is a cofibration in the* hf-*model structure.*

**Proof** We apply Corollary 4.4 to $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\mathrm{hf}}$ and the endofunctor $\widehat{P}_n$. In particular, we verify axioms (A1), (A2), and (A3′).

To prove (A1), assume that $F \to G$ is an hf-equivalence, ie that $F^{\mathrm{hf}} \to G^{\mathrm{hf}}$ is a levelwise weak equivalence. Our choice of model for homotopy limits preserves such weak equivalences, so $T_n$ does as well. Then Convention 5.7(3) guarantees that $\widehat{P}_n F \to \widehat{P}_n G$ is a levelwise weak equivalence. Since the functor $(-)^{\mathrm{hf}}$ preserves weak equivalences (by the proof of Theorem 5.16), it follows that $\widehat{P}_n F \to \widehat{P}_n G$ is an hf-equivalence.

For (A2), we need to prove that the natural transformations $\widehat{p}_{\widehat{P}_n F} : \widehat{P}_n F \to \widehat{P}_n \widehat{P}_n F$ and $\widehat{P}_n \widehat{p}_F : \widehat{P}_n F \to \widehat{P}_n \widehat{P}_n F$ are weak equivalences in the hf-model structure. Since $(-)^{\mathrm{hf}}$ preserves levelwise weak equivalences, it suffices to prove the stronger result that these natural transformations are levelwise weak equivalences.

For any object $A$ in $\mathcal{C}$, the $(n+1)$-cubical diagram given by $U \mapsto A * U$ is a strongly homotopy cocartesian diagram, so applying any $n$-excisive functor $H$ to it produces a homotopy cartesian diagram. Then the map from the initial object in this diagram to the homotopy limit of the rest of the diagram is a levelwise weak equivalence. By definition, this map is $H \to T_n H$, and since $T_n$ preserves levelwise weak equivalences, Convention 5.7(3) guarantees that the natural transformation $\iota_H : H \to T_n^\infty H$ is a levelwise weak equivalence as well.

To see that $\widehat{p}_{\widehat{P}_n F} : \widehat{P}_n F \to \widehat{P}_n \widehat{P}_n F$ is a levelwise weak equivalence, consider, for an arbitrary functor $X$, the commutative square

$$
\begin{array}{ccc}
X & \xrightarrow{\ \theta_X\ } & X^{\mathrm{hf}} \\
{\scriptstyle \iota_X}\big\downarrow & & \big\downarrow{\scriptstyle \iota_{X^{\mathrm{hf}}}} \\
T_n^\infty X & \xrightarrow[\ T_n^\infty \theta_X\ ]{} & T_n^\infty X^{\mathrm{hf}} = \widehat{P}_n X
\end{array}
$$

When $X = \widehat{P}_n F$, the horizontal maps are weak equivalences by Proposition 5.14 since $\widehat{P}_n F$ is a homotopy functor and $T_n^\infty$ preserves weak equivalences. As noted above, the left vertical map is also a weak equivalence since $\widehat{P}_n F$ is $n$-excisive. It follows that the composition of the right vertical and top horizontal arrows is a weak equivalence, but this composite is $\widehat{p}_{\widehat{P}_n F}$.

To prove that $\widehat{P}_n \widehat{p}_F$ is a levelwise weak equivalence, recall that it is the composite

$$
\widehat{P}_n F \xrightarrow{\ \widehat{P}_n \theta_F\ } \widehat{P}_n (F^{\mathrm{hf}}) \xrightarrow{\ \widehat{P}_n \iota_{F^{\mathrm{hf}}}\ } \widehat{P}_n (T_n^\infty (F^{\mathrm{hf}})).
$$

The map $\widehat{P}_n \theta_F$ is a levelwise weak equivalence by Theorem 5.16, Convention 5.7(3), and the fact that $T_n$ preserves weak equivalences. To see that $\widehat{P}_n \iota_{F^{\mathrm{hf}}}$ is a levelwise weak equivalence, consider the

commutative diagram

$$
\begin{array}{ccc}
T_n^\infty(F^{\mathrm{hf}}) & \xrightarrow{\;T_n^\infty(\iota_{F^{\mathrm{hf}}})\;} & T_n^\infty(T_n^\infty F^{\mathrm{hf}}) \\[2pt]
{\scriptstyle T_n^\infty(\theta_{F^{\mathrm{hf}}})}\Big\downarrow & & \Big\downarrow{\scriptstyle T_n^\infty(\theta_{T_n^\infty F^{\mathrm{hf}}})} \\[2pt]
T_n^\infty(F^{\mathrm{hf}})^{\mathrm{hf}} & \xrightarrow[\;T_n^\infty(\iota_{F^{\mathrm{hf}}})^{\mathrm{hf}}\;]{} & T_n^\infty(T_n^\infty F^{\mathrm{hf}})^{\mathrm{hf}}
\end{array}
$$

and note that the bottom horizontal arrow is $\widehat{P}_n \iota_{F^{\mathrm{hf}}}$. The vertical maps are levelwise weak equivalences by Propositions 5.14 and 5.18, and the fact that $T_n^\infty$ preserves weak equivalences. We can use an argument similar to the one used in the last paragraph of the proof of [20, 1.8] to prove that the top horizontal map is a weak equivalence. In particular, it suffices to show that $T_n^\infty(F^{\mathrm{hf}}) \to T_n^\infty(T_n F^{\mathrm{hf}})$ is a weak equivalence. By Convention 5.7(3) and the fact that homotopy limits commute, the target of this map is weakly equivalent to $T_n T_n^\infty(F^{\mathrm{hf}})$. However, $T_n^\infty(F^{\mathrm{hf}})$ is weakly equivalent to $T_n T_n^\infty(F^{\mathrm{hf}})$ because $T_n^\infty(F^{\mathrm{hf}}) = \widehat{P}_n F$ is an $n$-excisive functor. Hence, the bottom map in the diagram, which is $\widehat{P}_n \widehat{p}_F$, is a levelwise weak equivalence as well, completing the proof that (A2) holds.

Consider a homotopy pullback square

$$
\begin{array}{ccc}
F & \longrightarrow & G \\
\downarrow & & \downarrow \\
H & \longrightarrow & K
\end{array}
$$

in the hf-model structure. Since $(-)^{\mathrm{hf}}$ preserves levelwise weak equivalences, it suffices to show that there is a levelwise weak equivalence from $\widehat{P}_n F$ to the homotopy pullback (in the hf-model structure) of $\widehat{P}_n H \to \widehat{P}_n K \leftarrow \widehat{P}_n G$ to prove (A3′).

We factor $H \to K$ into an hf-equivalence $H \to H'$ followed by an hf-fibration $H' \to K$ so that the homotopy pullback of $G \to H \leftarrow K$ in the hf-model structure, which is right proper, is the strict pullback $H' \times_K G$. Note that $H' \times_K G$ is also the homotopy pullback of $H' \to K \leftarrow G$ in both the projective and hf-model structures since every hf-fibration is a projective fibration.

Since the functor $\widehat{P}_n$ preserves hf-fibrations and hf-weak equivalences (Proposition 6.7 and axiom (A1)), $\widehat{P}_n H \to \widehat{P}_n H' \to \widehat{P}_n K$ is also a factorization via an hf-equivalence and hf-fibration. As above, the strict pullback $\widehat{P}_n H' \times_{\widehat{P}_n K} \widehat{P}_n G$ is also the homotopy pullback of $\widehat{P}_n H \to \widehat{P}_n K \leftarrow \widehat{P}_n G$ in the hf-model structure, and of $\widehat{P}_n H' \to \widehat{P}_n K \leftarrow \widehat{P}_n G$ in both model structures. Hence, to complete the proof, it suffices to show there is a levelwise weak equivalence from $\widehat{P}_n F$ to the homotopy pullback of $\widehat{P}_n H' \to \widehat{P}_n K \leftarrow \widehat{P}_n G$.

We now consider the diagram

**(6.9)**
$$
\begin{array}{ccc}
F & \longrightarrow & G \\
\downarrow & & \downarrow \\
H' & \longrightarrow & K
\end{array}
$$

By assumption, $F^{\mathrm{hf}}$ is levelwise equivalent to $(H' \times_K G)^{\mathrm{hf}}$. But $H' \times_K G$ is the homotopy pullback of $H' \to K \leftarrow G$ in the projective model structure and $(-)^{\mathrm{hf}}$ preserves such homotopy pullbacks, making $F^{\mathrm{hf}}$ levelwise equivalent to the homotopy pullback of $(H')^{\mathrm{hf}} \to K^{\mathrm{hf}} \leftarrow G^{\mathrm{hf}}$. That is, applying $(-)^{\mathrm{hf}}$ to (6.9) yields a homotopy pullback in the projective model structure. Then Convention 5.7 and the fact that $T_n$, as a homotopy limit, preserves homotopy pullback diagrams ensure that applying $\widehat{P}_n$ to (6.9) produces a homotopy pullback in the projective model structure. $\qquad\square$

Proposition 4.6 allows us to characterize fibrations and fibrant objects in the $n$-excisive model structure as follows. Biedermann and Röndigs provide a similar characterization of fibrations in [10, 5.9].

**Proposition 6.10** *A morphism $\alpha\colon F \to G$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is a fibration in the $n$-excisive model structure if and only if it is a fibration in the* hf*-model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ and the diagram*

$$
\begin{array}{ccc}
F & \xrightarrow{\;\widehat{p}_F\;} & \widehat{P}_n F \\
{\scriptstyle\alpha}\big\downarrow & & \big\downarrow{\scriptstyle\widehat{P}_n\alpha} \\
G & \xrightarrow[\;\widehat{p}_G\;]{} & \widehat{P}_n G
\end{array}
$$

*is a homotopy pullback square in the* hf*-model structure. A functor $F$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is fibrant in the $n$-excisive model structure if and only if it is weakly equivalent in the* hf*-model structure to an $n$-excisive functor and is fibrant in the* hf*-model structure.*

**Proof** The first statement is an immediate consequence of Proposition 4.6. For the second part, one can show that $\widehat{p}_*\colon * \to \widehat{P}_n(*)$, where $*$ denotes the terminal object in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, is a weak equivalence in the projective model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ using Lemma 5.15, the facts that $T_n$ preserves weak equivalences and that $T_n(*)$ is weakly equivalent to $*$, and Convention 5.7(3). Since $(-)^{\mathrm{hf}}$ satisfies (A1), the map $\widehat{p}_*\colon * \to \widehat{P}_n(*)$ is also an hf-equivalence. By Proposition 4.6 it suffices to prove that $\widehat{p}_F$ is an hf-equivalence if and only if $F$ is hf-equivalent to an $n$-excisive functor.

Consider the diagram

$$
\begin{array}{ccc}
F & \xrightarrow{\;\widehat{p}_F\;} & \widehat{P}_n F \\
{\scriptstyle\beta}\big\downarrow & & \big\downarrow{\scriptstyle\widehat{P}_n\beta} \\
G & \xrightarrow[\;\widehat{p}_G\;]{} & \widehat{P}_n G
\end{array}
$$

If $\beta$ is an $hf$-weak equivalence, then $\widehat{P}_n\beta$ is an hf-weak equivalences since $\widehat{P}_n$ preserves hf-weak equivalences. Moreover, as noted in the proof of (A2) for Theorem 6.8, if $G$ is $n$-excisive, the maps in the colimit defining $\widehat{P}_n G$ are all weak equivalences and as a result, $\iota_{G^{\mathrm{hf}}}\colon G^{\mathrm{hf}} \to \widehat{P}_n G$ is a levelwise weak equivalence. Since $(-)^{\mathrm{hf}}$ satisfies (A2) and preserves levelwise weak equivalences, it follows that the composite $\widehat{p}_G$ is an hf-equivalence. By Proposition 2.3, $\widehat{p}_F$ is a weak equivalence, establishing one implication. The converse follows immediately from the fact that $\widehat{P}_n F$ is $n$-excisive. $\qquad\square$

# 7 The degree *n* model structure

In this section, we turn our attention to the discrete functor calculus, with the goal of showing that one can equip $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ with a degree $n$ model structure via the degree $n$ approximation $\Gamma_n$. As was the case with the homotopy functor and $n$-excisive model structures, our categories $\mathcal{C}$ and $\mathcal{D}$ must satisfy some conditions, but they are much less complicated in this case.

**Convention 7.1**  In addition to Convention 3.5, we assume that $\mathcal{D}$ is left proper.

Recall from Definition 3.12 that for a functor $F$ and an object $A$ in $\mathcal{C}$,

$$\Gamma_n F(A) := \mathrm{hocofiber}\big(|\perp_{n+1}^{*+1} F(A)| \to F(A)\big),$$

where $|-|$ denotes the homotopy colimit over $\Delta^{\mathrm{op}}$, sometimes referred to as the *fat realization*. We establish the existence of the degree $n$ model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ by proving that $\Gamma_n$ satisfies the conditions of Theorem 4.2, for which we make use of specific models for the homotopy limits and colimits used to define $\Gamma_n$.

For the homotopy limit, we again use the model defined in [24, 18.1.8]. As noted in Section 3.2, we can assume that $\perp_{n+1}$ is a comonad. To guarantee that $\perp_{n+1}^k$ preserves weak equivalences, we replace $F$ with its functorial fibrant replacement $\mathbb{F}(F)$, which exists by Theorem 5.3. That is, we set

$$\perp_{n+1}^k F := \perp_{n+1}^k \mathbb{F}(F)$$

for any functor $F$. Then by [24, 18.5.2, 18.5.3] we know that $\perp_{n+1}^k$ preserves weak equivalences in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$.

For the model of $\Gamma_n$ that we use, we also require a good model for homotopy colimits, for which we use the one described by Hirschhorn [24, 18.1.2]. To ensure that $|-|$ preserves weak equivalences, we precompose the homotopy colimit with the simplicial functorial cofibrant replacement functor guaranteed by [41, 24.2]. Again, [24, 18.5.3] guarantees that this functor preserves levelwise weak equivalences of diagrams.

Finally, to guarantee that the homotopy cofiber preserves weak equivalences, and that there is a natural transformation from the identity functor on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ to $\Gamma_n$, we use the following model for the homotopy cofiber. For a map $F \to G$, we set $\mathrm{hocofiber}(F \to G)$ equal to the pushout of the diagram $E(\star_{\mathcal{D}}) \leftarrow F \to G$ where $\star_{\mathcal{D}}$ is the constant functor on the zero object in $\mathcal{D}$, and $F \to E(\star_{\mathcal{D}}) \to \star_{\mathcal{D}}$ is a functorial factorization of $F \to \star_{\mathcal{D}}$ as a cofibration followed by an acyclic fibration. That this construction is homotopy invariant follows from [24, 13.5.3, 13.5.4] and the fact that $\mathcal{D}$ is left proper.

When using the construction defined above, there is a natural map from $G$ to $\mathrm{hocofiber}(F \to G)$; an application of this fact to the augmentation $|\perp_{n+1}^{*+1} F| \to \mathbb{F}(F)$ yields a natural map $\mathbb{F}(F) \to \Gamma_n(F)$, and precomposing with the natural weak equivalence $\theta_F \colon F \to \mathbb{F}(F)$ yields a natural transformation $F \to \Gamma_n F$ that is natural in $F$. Hence, we have a natural transformation $\gamma \colon \mathrm{id}_{\mathrm{Fun}(\mathcal{C}, \mathcal{D})} \to \Gamma_n$.

Recall that as part of Convention 3.5, we are assuming $\mathcal{D}$ is stable. As noted in Section 3, the proof that the functor $\Gamma_n F$ is a degree $n$ functor in [6, 5.4, 5.6.1] makes use of the fact that $\mathcal{D}$ is stable to ensure that $c r_{n+1}$ commutes with the homotopy cofiber and colimit used to define $\Gamma_n F$. We thus obtain

$$c r_{n+1} \Gamma_n F = c r_{n+1} \operatorname{hocofiber}(|\bot_{n+1}^{*+1} F| \to F)$$
$$\simeq \operatorname{hocofiber}\big(|c r_{n+1}(\bot_{n+1}^{*+1} F)| \to c r_{n+1} F\big)$$
$$\simeq \star_{\mathcal{D}}.$$

The last equivalence is a consequence of [6, 5.5] which uses an extra degeneracy argument to prove that $|c r_{n+1}(\bot_{n+1}^{*+1} F)| \to c r_{n+1} F$ is a weak equivalence.

We now state the main result of this section.

**Theorem 7.2** *Under the assumptions of Convention 7.1, there exists a degree $n$ model structure on the category of functors* $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$ *with weak equivalences given by $\Gamma_n$-equivalences and cofibrations the same as in the projective model structure.*

**Proof** We prove this result via an application of Corollary 4.4, setting $\mathcal{M} = \operatorname{Fun}(\mathcal{C}, \mathcal{D})$ and $Q = \Gamma_n \colon \operatorname{Fun}(\mathcal{C}, \mathcal{D}) \to \operatorname{Fun}(\mathcal{C}, \mathcal{D})$. As described above, the functor $\Gamma_n \colon \operatorname{Fun}(\mathcal{C}, \mathcal{D}) \to \operatorname{Fun}(\mathcal{C}, \mathcal{D})$ is constructed from $F$ via homotopy limits, homotopy colimits, and homotopy cofibers that preserve weak equivalences. Hence, axiom (A1) of Theorem 4.2 is satisfied.

To prove that axiom (A2) holds, we assume that $F$ is fibrant and use Corollary 4.3. By [24, 18.5.2, 18.5.3], the results of applying $\bot_{n+1}^{k}$ to $F$ and $\mathbb{F}(F)$ are weakly equivalent, so we can work directly with $F$ instead of its functorial fibrant replacement.

To prove that $\gamma_{\Gamma_n F}$ is a weak equivalence, we note that for any degree $n$ functor $G$,

$$\bot_{n+1} G = \Delta^* c r_{n+1} G \simeq \star_{\mathcal{D}},$$

where $\Delta^*$ denotes precomposition with the diagonal functor. Hence, the simplicial object $\bot_{n+1}^{*+1} G$ is levelwise weakly equivalent to the constant simplicial object on $\star_{\mathcal{D}}$ and $|\bot_{n+1}^{*+1} G| \simeq \star_{\mathcal{D}}$. Since $\Gamma_n G$ is the homotopy cofiber of $|\bot_{n+1}^{*+1} G| \to G$ and $\mathcal{D}$ is left proper, the dual of Proposition 2.3 guarantees that $\gamma_G$ is a weak equivalence. Setting $G = \Gamma_n F$, we see that $\gamma_{\Gamma_n F}$ is a weak equivalence.

To prove that $\Gamma_n \gamma_F$ is a weak equivalence, it suffices to show that $\Gamma_n |\bot_{n+1}^{*+1} F| \simeq \star_{\mathcal{D}}$. As a comonad, $\bot_{n+1}$ comes equipped with a natural transformation, the comultiplication $\bot_{n+1} \Rightarrow \bot_{n+1} \bot_{n+1}$, which can be used to construct weak equivalences between $|\bot_{n+1}^{k} \bot_{n+1}^{*+1} F|$ and $\bot_{n+1}^{k} F$ for $k \geq 1$, as in [6, 5.5]. Since $\mathcal{D}$ is stable, $\bot_{n+1}$, as a finite homotopy limit, commutes with $|-|$, so we have

$$\big|\bot_{n+1}^{*+1} |\bot_{n+1}^{*+1} F|\big| \simeq \big||\bot_{n+1}^{*+1} \bot_{n+1}^{*+1} F|\big| \simeq |\bot_{n+1}^{*+1} F|.$$

It follows that $\Gamma_n |\bot_{n+1}^{*+1} F| \simeq \star_{\mathcal{D}}$, which implies that $\Gamma_n \gamma_F \colon \Gamma_n F \to \Gamma_n \Gamma_n F$ is a weak equivalence.

It remains to check that $\Gamma_n$ satisfies axiom (A3$'$), ie that homotopy pullback squares are preserved by $\Gamma_n$. Suppose that

$$
\begin{array}{ccc}
F & \longrightarrow & G \\
\downarrow & & \downarrow \\
H & \longrightarrow & K
\end{array}
$$

is a homotopy pullback square in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ and recall that homotopy pullback and homotopy pushout squares agree in a stable model category such as $\mathcal{D}$. Since $\Gamma_n$ is constructed via homotopy limits and colimits, homotopy limits preserve homotopy pullbacks, and homotopy colimits preserve homotopy pushouts, it follows that applying $\Gamma_n$ to this diagram yields a homotopy pullback square. □

We conclude with the following consequence of Proposition 4.6. Its proof is similar to the one of Proposition 6.10 and hence omitted.

**Proposition 7.3**   *A morphism $\alpha\colon F \to G$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is a fibration in the degree $n$ model structure if and only if it is a fibration in the projective model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ and the diagram*

$$
\begin{array}{ccc}
F & \xrightarrow{\ \gamma_F\ } & \Gamma_n F \\
{\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\Gamma_n\alpha} \\
G & \xrightarrow[\ \gamma_G\ ]{} & \Gamma_n G
\end{array}
$$

*is a homotopy pullback square. The object $F$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is fibrant in the degree $n$ model structure if and only if it is degree $n$ and is fibrant in the projective model structure.*

# 8   Cofibrant generation for Bousfield $Q$-model structures

In this section, we establish conditions under which the localizations produced by an endofunctor $Q$ on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ via Theorem 4.2 are cofibrantly generated, expanding on the more specific examples of Biedermann and Röndigs in [10]. A core element of those examples is the strategic creation of additional generating acyclic cofibrations for the model structure $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$. Since our goal is to generalize those examples, we make the following definition to capture the key features of the collections of maps that they use. We assume throughout that $\mathcal{C}$ and $\mathcal{D}$ are as described in Convention 2.8.

**Definition 8.1**   Let $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ be equipped with a right proper model structure and let $Q$ be an endofunctor of $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ satisfying the conditions of Theorem 4.2, including the existence of a natural transformation $\eta\colon \mathrm{id} \Rightarrow Q$. A *collection of test morphisms* for $\eta$ is a collection $T(Q)$ of morphisms in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ such

that, for each fibration $F \to G$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, the diagram

$$
\begin{array}{ccc}
F & \longrightarrow & QF \\
\downarrow & & \downarrow \\
G & \longrightarrow & QG
\end{array}
$$

is a homotopy pullback in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ if and only if the diagram

$$
\begin{array}{ccc}
F^Y & \longrightarrow & F^X \\
\downarrow & & \downarrow \\
G^Y & \longrightarrow & G^X
\end{array}
$$

is a homotopy pullback in $\mathcal{D}$ for every $X \to Y$ in $T(Q)$.

We sometimes omit the endofunctor $Q$ from the notation and simply refer to the collection of test morphisms as $T$. In all the examples in this paper, the functor $Y$ is a representable functor $R^A$, where $A$ is an object of $\mathcal{C}$.

**Example 8.2** Consider the hf-model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ from Theorem 5.16, where $\mathcal{C}$ and $\mathcal{D}$ additionally satisfy the conditions of Convention 5.7. Since it is designed to emphasize weak equivalence-preserving functors, we claim that the collection of morphisms of representable functors $\{R^B \to R^A\}$, where $A \to B$ ranges over all weak equivalences of $\mathcal{C}$, is a collection of test morphisms for $\eta\colon \mathrm{id} \Rightarrow (-)^{\mathrm{hf}}$. Indeed, our definition of a collection of test morphisms is essentially a distillation of the key properties that Biedermann and Röndigs use in [10] to show that the hf-model structure is cofibrantly generated. That these morphisms satisfy Definition 8.1 was proved by Biedermann and Röndigs in [10, 4.15] using the fact that $F^{R^A} \cong F(A)$ by the Yoneda Lemma 2.14.

We revisit this example at the end of this section and give additional examples in Sections 9 and 10, where we consider the $n$-excisive and degree $n$ model structures, respectively.

We can now state and prove our main result.

**Theorem 8.3** *Suppose that $\mathcal{C}$ and $\mathcal{D}$ satisfy the conditions of Convention 6.2, and that $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ is a cofibrantly generated right proper model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ in which all fibrations are also fibrations under the projective model structure. Let $Q\colon \mathrm{Fun}(\mathcal{C}, \mathcal{D}) \to \mathrm{Fun}(\mathcal{C}, \mathcal{D})$ be a Bousfield endofunctor that has a collection $T(Q)$ of test morphisms for the natural transformation $\eta\colon \mathrm{id} \Rightarrow Q$. Then the model structure $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$ is cofibrantly generated.*

**Proof** Let $I$ and $J$ denote sets of generating cofibrations and acyclic cofibrations, respectively, for the model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. We need to identify sets of generating cofibrations and acyclic cofibrations for the model structure $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$. Since the cofibrations are unchanged by the $Q$-localization, we can simply use the set $I$ as a set of the generating cofibrations for $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$.

Assume that $F \to G$ is a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. To identify a set of generating acyclic cofibrations for $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$, we use the $Q$-fibration condition of Proposition 4.6, namely that $F \to G$ is a $Q$-fibration if and only if it is a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$, and the diagram

**(8.4)**
$$
\begin{array}{ccc}
F & \xrightarrow{\eta_F} & QF \\
\downarrow & & \downarrow \\
G & \xrightarrow[\eta_G]{} & QG
\end{array}
$$

is a homotopy pullback square in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. We know that a map is a fibration in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ if and only if it has the right lifting property with respect to the set $J$. Thus, it suffices to identify a set of maps $J_Q$ such that the above diagram is a homotopy pullback square if and only if $F \to G$ has the right lifting property with respect to the maps in $J_Q$; we can then take $J \cup J_Q$ as the set of generating acyclic cofibrations for $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$.

By the definition of test morphism, (8.4) is a homotopy pullback square if and only if

**(8.5)**
$$
\begin{array}{ccc}
F^Y & \longrightarrow & F^X \\
\downarrow & & \downarrow \\
G^Y & \longrightarrow & G^X
\end{array}
$$

is a homotopy pullback square for each $\alpha \colon X \to Y$ in $T$.

However, for each object $X$ in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ we can apply Proposition 2.7 to the morphism from the initial object to $X$ to obtain a simplicial homotopy equivalence

**(8.6)**
$$\widehat{X} \to X,$$

where $\widehat{X}$ is cofibrant. Since simplicial functors preserve simplicial homotopy equivalences [24, 9.6.10], we have that $F^X \to F^{\widehat{X}}$ and $G^X \to G^{\widehat{X}}$ are weak equivalences. For each test morphism $\alpha \colon X \to Y$, we obtain a diagram in which the right-hand square is a homotopy pullback square by Proposition 2.3:

**(8.7)**
$$
\begin{array}{ccccc}
F^Y & \longrightarrow & F^X & \longrightarrow & F^{\widehat{X}} \\
\downarrow & & \downarrow & & \downarrow \\
G^Y & \longrightarrow & G^X & \longrightarrow & G^{\widehat{X}}
\end{array}
$$

Applying Proposition 2.2 to this diagram, we see that (8.5) is a homotopy pullback if and only if the outer square in (8.7) is a homotopy pullback square.

Again by Proposition 2.7 and the fact that simplicial functors preserve simplicial homotopy equivalences, we know that the composite $\widehat{X} \to X \xrightarrow{\alpha} Y$ can be factored as a cofibration $\zeta_\alpha$ followed by a simplicial homotopy equivalence $\widehat{X} \xrightarrow{\zeta_\alpha} \mathrm{Cyl}(\alpha) \xrightarrow{\simeq} Y$, and we obtain weak equivalences of evaluated cotensors

$F^Y \to F^{\mathrm{Cyl}(\alpha)}$ and $G^Y \to G^{\mathrm{Cyl}(\alpha)}$. Then, by Proposition 2.2, the outer square in (8.7) is a homotopy pullback square if and only if

**(8.8)**

$$
\begin{array}{ccc}
F^{\mathrm{Cyl}(\alpha)} & \longrightarrow & F^{\widehat{X}} \\
\downarrow & & \downarrow \\
G^{\mathrm{Cyl}(\alpha)} & \longrightarrow & G^{\widehat{X}}
\end{array}
$$

is a homotopy pullback square.

Since $F \to G$ is a projective fibration by assumption, by Corollary 2.16 we know that $F^{\widehat{X}} \to G^{\widehat{X}}$ is a fibration, so (8.8) is a homotopy pullback if and only if

**(8.9)**
$$ F^{\mathrm{Cyl}(\alpha)} \to F^{\widehat{X}} \times_{G^{\widehat{X}}} G^{\mathrm{Cyl}(\alpha)} $$

is a weak equivalence. Note that (8.9) is a fibration by Lemma 2.15. So we can show that it is a weak equivalence by showing that a lift exists in every commutative diagram of the form

**(8.10)**

$$
\begin{array}{ccc}
C & \longrightarrow & F^{\mathrm{Cyl}(\alpha)} \\
\downarrow & \nearrow & \downarrow \\
D & \longrightarrow & F^{\widehat{X}} \times_{G^{\widehat{X}}} G^{\mathrm{Cyl}(\alpha)}
\end{array}
$$

where $C \to D$ is in the set $I_{\mathcal{D}}$ of generating cofibrations for $\mathcal{D}$. By the first adjunction of Proposition 2.13, a lift exists in (8.10) if and only if a lift exists in

$$
\begin{array}{ccc}
D \otimes \widehat{X} \amalg_{C \otimes \widehat{X}} C \otimes \mathrm{Cyl}(\alpha) & \longrightarrow & F \\
\downarrow & \nearrow & \downarrow \\
D \otimes \mathrm{Cyl}(\alpha) & \longrightarrow & G
\end{array}
$$

Hence, the left-hand vertical maps can be taken as the set $J_Q$. That is,

$$
J_Q = \{ i \,\square\, \zeta_\alpha \colon C \otimes \mathrm{Cyl}(\alpha) \amalg_{C \otimes \widehat{X}} D \otimes \widehat{X} \to D \otimes \mathrm{Cyl}(\alpha) \mid (i \colon C \to D) \in I_{\mathcal{D}},\ \alpha \in T \}.
$$

The preceding argument shows that the maps in $J_Q$ have the left lifting property with respect to all $Q$-fibrations, so that they are indeed acyclic cofibrations in the model structure induced by $Q$. To complete the proof, it remains to show that this set of maps permits the small object argument. We need to show that, given a transfinite composition $\mathrm{colim}_n F_n$ of pushouts along acyclic cofibrations, any map

$$
\langle f, g \rangle \colon C \otimes \mathrm{Cyl}(\alpha) \amalg_{C \otimes \widehat{X}} D \otimes \widehat{X} \to \mathrm{colim}_n F_n
$$

factors through some $F_k$. By the universal property of pushouts, the data of such a map is equivalent to a commutative square

$$
\begin{array}{ccc}
C \otimes \widehat{X} & \xrightarrow{\ i \otimes 1\ } & D \otimes \widehat{X} \\
{\scriptstyle 1 \otimes \zeta_\alpha} \downarrow & & \downarrow {\scriptstyle g} \\
C \otimes \mathrm{Cyl}(\alpha) & \xrightarrow[\ f\ ]{} & \mathrm{colim}_n F_n
\end{array}
$$

and a factorization of such a map through a functor $F_k$ is equivalent to a commutative diagram

$$
\begin{array}{ccc}
C \otimes \widehat{X} & \xrightarrow{\ i \otimes 1\ } & D \otimes \widehat{X} \\
{\scriptstyle 1 \otimes \zeta_\alpha}\big\downarrow & \begin{array}{c} f_k \nearrow\ \ F_k\ \ \nwarrow g_k \end{array} & \big\downarrow {\scriptstyle g} \\
C \otimes \mathrm{Cyl}(\alpha) & \xrightarrow[\ \ f\ \ ]{} & \mathrm{colim}_n\ F_n
\end{array}
$$

Using the first adjunction of Proposition 2.13, this commutative diagram is in turn equivalent to

$$
\begin{array}{ccc}
C & \xrightarrow{\ \widehat{f}\ } & (\mathrm{colim}_n\ F_n)^{\mathrm{Cyl}(\alpha)} \\
{\scriptstyle i}\big\downarrow \quad {\scriptstyle \widehat{f_k}}\nearrow\ \ F_k^{\mathrm{Cyl}(\alpha)} \nearrow & & \big\downarrow {\scriptstyle 1^{\zeta_\alpha}} \\
& \big\downarrow & \\
{\scriptstyle \widehat{g_k}}\nearrow\ \ F_k^{\widehat{X}} & & \\
D & \xrightarrow[\ \widehat{g}\ ]{} & (\mathrm{colim}_n\ F_n)^{\widehat{X}}
\end{array}
$$

The fact that both $C$ and $D$ permit the small object argument, by Convention 6.2, guarantees the existence of maps $\widehat{f_k}$ and $\widehat{g_k}$ such that the diagram above commutes. $\qquad\square$

As a first example of this localized cofibrant generation, the next theorem follows immediately from Theorem 8.3 and Example 8.2; an alternate proof is given in [10, 4.14].

**Theorem 8.11** *Assuming Convention 6.2, the model category* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\mathrm{hf}}$ *of Theorem 5.16 has the structure of a cofibrantly generated model category.*

When $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ has the projective model structure, Theorem 8.3 implies that for any Bousfield endofunctor $Q$ that has a collection of test morphisms, the model structure $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$ of Theorem 4.2 is cofibrantly generated. Moreover, since the fibrations of $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_Q$ must also be projective fibrations by Proposition 4.6, we can use Theorem 8.3 to conclude that the localization of $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ obtained by applying a sequence of Bousfield endofunctors satisfying the appropriate test morphism conditions is cofibrantly generated. We provide an example of this in the next section by building an $n$-excisive model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ from the hf-model structure. To do so, we make use of the next result.

**Proposition 8.12** *Consider the category* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ *with the projective model structure, as well as a localized model structure* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_P$ *induced by a Bousfield endofunctor* $P$ *of* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. *Suppose that* $Q$ *is an endofunctor of* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_P$ *that preserves* $P$*-fibrations. Then for any* $P$*-fibration* $F \to G$,

**(8.13)**

$$
\begin{array}{ccc}
F & \longrightarrow & QF \\
\big\downarrow & & \big\downarrow \\
G & \longrightarrow & QG
\end{array}
$$

*is a homotopy pullback square in the* $P$*-model structure if and only if it is a levelwise homotopy pullback.*

**Proof** By assumption, $QF \to QG$ is a $P$-fibration, and hence, a projective fibration. So the pullback of $G \to QG \leftarrow QF$ is a homotopy pullback in both the projective and $P$-model structures.

If (8.13) is a levelwise homotopy pullback square, then the map $F \to QF \times_{QG} G$ is a levelwise weak equivalence. Since $P$ satisfies axiom (A1) of Theorem 4.2, this map is also a $P$-equivalence, and (8.13) is a homotopy pullback square in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_P$.

Conversely, suppose that the diagram (8.13) is a homotopy pullback square in the $P$-model structure, so that the map $PF \to P(QF \times_{QG} G)$ is a levelwise weak equivalence. Since $P$ satisfies axiom (A3$'$) of Corollary 4.4, the map $P(QF \times_{QG} G) \to PQF \times^h_{PQG} PG$ is a levelwise weak equivalence, and the right-hand square in the diagram

**(8.14)**
$$
\begin{array}{ccccc}
F & \longrightarrow & PF & \longrightarrow & PQF \\
\downarrow & & \downarrow & & \downarrow \\
G & \longrightarrow & PG & \longrightarrow & PQG
\end{array}
$$

is a homotopy pullback square in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. By Proposition 4.6, the left-hand square is a levelwise homotopy pullback as well, from which we can conclude by Proposition 2.2 that the outer square is also a levelwise homotopy pullback square.

However, this outer square can be obtained similarly as a composite diagram with middle vertical map the $P$-fibration $QF \to QG$, so that the left-hand square is the square (8.13). We can conclude that this square is a levelwise homotopy pullback square via another application of Propositions 4.6 and 2.2, completing the proof. $\qquad\square$

# 9 Cofibrant generation and the *n*-excisive model structure

In this section, we apply Theorem 8.3 and Proposition 8.12 to show that the $n$-excisive model structure of Theorem 6.8 is cofibrantly generated. As in Section 6, we assume Convention 6.2.

We first define a candidate set of test morphisms for the $n$-excisive model structure.

**Definition 9.1** For an object $A$ in $\mathcal{C}$, let $\tau_A$ be the morphism

$$
\tau_A \colon \mathrm{hocolim}_{U \subseteq \mathcal{P}_0(\boldsymbol{n+1})} R^{A * U} \to R^A
$$

in $\mathrm{Fun}(\mathcal{C}, \mathcal{S})$ induced by the inclusions $\varnothing \hookrightarrow U$, where $*$ denotes the fiberwise join as in Definition 3.2. We denote by $T(\widehat{P}_n)$ the collection $\{\tau_A\}$ of these morphisms as $A$ ranges over all objects of $\mathcal{C}$.

**Proposition 9.2** *For the* hf*-model structure* $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\mathrm{hf}}$*, the collection* $T(\widehat{P}_n)$ *is a collection of test morphisms for the natural transformation* $\eta \colon \mathrm{id} \Rightarrow P_n$.

**Proof**  Let $F \to G$ be an hf-fibration. By Proposition 6.7, we know that $\widehat{P}_n$ preserves hf-fibrations. Hence, by Proposition 8.12, it suffices to show that

(9.3)
$$
\begin{array}{ccc}
F & \longrightarrow & \widehat{P}_n F \\
\downarrow & & \downarrow \\
G & \longrightarrow & \widehat{P}_n G
\end{array}
$$

is a levelwise homotopy pullback square if and only if

(9.4)
$$
\begin{array}{ccc}
F^{R^A} & \longrightarrow & F^{\operatorname{hocolim}_{U \in \mathcal{P}_0(n+1)} R^{A*U}} \\
\downarrow & & \downarrow \\
G^{R^A} & \longrightarrow & G^{\operatorname{hocolim}_{U \in \mathcal{P}_0(n+1)} R^{A*U}}
\end{array}
$$

is a homotopy pullback square in $\mathcal{D}$ for all objects $A$ in $\mathcal{C}$. Combining the isomorphism of Proposition 2.18, Lemma 2.14, and the definition of $T_n$ (Definition 3.2), we see that (9.4) is a homotopy pullback square for all objects $A$ in $\mathcal{C}$ if and only if

(9.5)
$$
\begin{array}{ccc}
F & \longrightarrow & T_n F \\
\downarrow & & \downarrow \\
G & \longrightarrow & T_n G
\end{array}
$$

is a levelwise homotopy pullback square. So it suffices to show that (9.3) is a levelwise homotopy pullback if and only if (9.5) is.

Suppose (9.3) is a levelwise homotopy pullback and consider the commutative cube

(9.6)
$$
\begin{array}{ccccc}
F & \longrightarrow & & \widehat{P}_n F & \\
& \searrow & & & \searrow \\
\downarrow & & T_n F & \longrightarrow & T_n \widehat{P}_n F \\
& & \downarrow & & \\
G & \dashrightarrow & \widehat{P}_n G & & \downarrow \\
& \searrow & & \searrow & \\
& & T_n G & \longrightarrow & T_n \widehat{P}_n G
\end{array}
$$

The front and back faces are homotopy pullbacks by assumption and the fact that $T_n$ preserves homotopy pullbacks, respectively. If we consider its back and right faces, we see that the right face is a homotopy pullback because $\widehat{P}_n F$ is $n$-excisive, and hence the composite of the back and right faces is a homotopy pullback by Proposition 2.2. As a result, we can conclude that the composite of the left and front faces is a homotopy pullback. Since the front face is a homotopy pullback, we can apply Proposition 2.2 again to see that the left face, which is precisely (9.5), is as well.

Conversely, suppose that (9.5) is a homotopy pullback square. Consider the commutative diagram that defines $\widehat{P}_n$ and the natural transformations in (9.3):

**(9.7)**
$$
\begin{array}{ccccccc}
F & \longrightarrow & F^{\mathrm{hf}} & \longrightarrow & T_n F^{\mathrm{hf}} & \longrightarrow \cdots \longrightarrow & \operatorname{colim}_k T_n^k F^{\mathrm{hf}} = \widehat{P}_n F \\
\downarrow & & \downarrow & & \downarrow & & \vdots \\
G & \longrightarrow & G^{\mathrm{hf}} & \longrightarrow & T_n G^{\mathrm{hf}} & \longrightarrow \cdots \longrightarrow & \operatorname{colim}_k T_n^k G^{\mathrm{hf}} = \widehat{P}_n G
\end{array}
$$

It suffices to show that for each $k \geq 0$, the square

**(9.8)**
$$
\begin{array}{ccc}
F & \longrightarrow & T_n^k F^{\mathrm{hf}} \\
\downarrow & & \downarrow \\
G & \longrightarrow & T_n^k G^{\mathrm{hf}}
\end{array}
$$

whose horizontal maps are given by composites of horizontal maps in (9.7) is a homotopy pullback square. In the case that $k = 0$, this follows from Proposition 4.6.

Assuming (9.8) is a homotopy pullback for some $k \geq 0$, we see that the right-hand square of the commutative diagram

**(9.9)**
$$
\begin{array}{ccccc}
F & \longrightarrow & T_n F & \longrightarrow & T_n^{k+1} F^{\mathrm{hf}} \\
\downarrow & & \downarrow & & \downarrow \\
G & \longrightarrow & T_n G & \longrightarrow & T_n^{k+1} G^{\mathrm{hf}}
\end{array}
$$

is a homotopy pullback since it is obtained by applying the functor $T_n$, which preserves homotopy pullbacks, to (9.8). The outer square is then a homotopy pullback square by Proposition 2.2 since the left-hand square is (9.5). The commutative cube

**(9.10)**


shows us that the outer square in (9.9) is the same square as (9.8) when $k$ is replaced by $k+1$, completing the proof by induction. $\qquad\square$

We can now conclude the main result of this section from Theorem 8.3.

**Theorem 9.11** *The $n$-excisive model structure on $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$ from Theorem 6.8 is cofibrantly generated.*

**Remark 9.12**   Theorem 9.11 is also proved as part of [10, 5.8]. Our proof is essentially a reorganization and generalization of theirs via Theorem 8.3, but our approaches differ in the stage at which the representable functors of Definition 9.1 are introduced. In [10], the authors incorporate them into their definition of $P_n$ by replacing $T_n F$ with an evaluated cotensor

$$T_n^R F(A) := F^{A_n},$$

where

$$A_n \xrightarrow{\simeq} \mathrm{hocolim}_{U \in \mathcal{P}_0(\boldsymbol{n+1})} R^{A*U}$$

is a cofibrant replacement for the homotopy colimit of the representable functors $R^{A*U}$. We proved in [2, 7.4] that $T_n$ and $T_n^R$ agree up to weak equivalence. We have chosen to define $\widehat{P}_n$ without using an evaluated cotensor to highlight the fact that this approach is not needed to establish the existence of the $n$-excisive model structure. It does play a significant role in establishing cofibrant generation, since, in the proof of Theorem 8.3, being able to replace (8.4) with the evaluated cotensor square (8.5) provides the means by which we can identify a set of generating acyclic cofibrations, but now the specific evaluated cotensor approach to $\widehat{P}_n$ only appears concretely in our verification of our set of test morphisms in Proposition 9.2.

## 10   Cofibrant generation and discrete functor calculus

We now revisit the degree $n$ model structure of Section 7 and use Theorem 8.3 to show that it is cofibrantly generated when $\mathcal{D}$ is. As in Section 7 we assume Convention 7.1.

Recall from Definition 3.12 and Section 7 that the functor $\Gamma_n F$ is defined in terms of a comonad $\perp_{n+1}$ that acts on the category $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. More explicitly, it is the homotopy cofiber given by

$$\Gamma_n F := \mathrm{hocofiber}\big(|\perp_{n+1}^{*+1} \mathbb{F}(F)| \to \mathbb{F}(F)\big),$$

where $|\perp_{n+1}^{*+1} F|$ is the fat realization of the standard simplicial object associated to the comonad $\perp_{n+1}$ acting on $F$ and $\mathbb{F}(F)$ is a functorial fibrant replacement of $F$.

**Theorem 10.1**   *The degree $n$ model structure on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ from Theorem 7.2 is cofibrantly generated.*

We want to prove this theorem by an application of Theorem 8.3, for which we need to define a collection of test morphisms for $\Gamma_n$. Recall from (3.7) that for an object $A$ in $\mathcal{C}$, subset $U$ of $\boldsymbol{n+1}$, and element $i \in \boldsymbol{n+1}$, we defined

$$A_i(U) := \begin{cases} A & \text{if } i \notin U, \\ *_{\mathcal{C}} & \text{if } i \in U. \end{cases}$$

Using this definition, we define

$$\sqcup(A, U) := \coprod_{i \in \boldsymbol{n+1}} A_i(U),$$

and note that for $U \subseteq V$ there is a natural map

$$\iota_{U,V} : \sqcup(A, U) \to \sqcup(A, V)$$

induced by the unique morphism $A \to *_{\mathcal{C}}$ on the components indexed by $i \in V \setminus U$.

**Definition 10.2** For an object $A$ in $\mathcal{C}$, let $\rho_A$ be the morphism

$$\rho_A : \operatorname{hocolim}_{U \subseteq \mathcal{P}_0(\boldsymbol{n+1})} R^{\sqcup(A,U)} \to R^{\sqcup(A,\varnothing)}$$

in $\operatorname{Fun}(\mathcal{C}, \mathcal{S})$ induced by the morphisms $\iota_{\varnothing,U}$. We denote by $T(\Gamma_n)$ the collection $\{\rho_A\}$ of these morphisms as $A$ ranges over all objects of $\mathcal{C}$.

**Lemma 10.3** *For the projective model structure on* $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$, *the collection* $T(\Gamma_n)$ *of Definition 10.2 is a collection of test morphisms for the natural transformation* $\gamma : \operatorname{id} \Rightarrow \Gamma_n$.

To prove this lemma, we use the next two results.

**Lemma 10.4** *Let* $\mathcal{C}$ *be a subcategory of a model category that is closed under finite limits. If* $\mathcal{X}$ *is an* $n$-*cube in* $\mathcal{C}$, *then*

$$\operatorname{ifiber}(\mathcal{X}) \simeq \operatorname{hofiber}(\mathcal{X}(\varnothing) \to \operatorname{holim}_{\mathcal{P}_0(\boldsymbol{n})} \mathcal{X}(U)).$$

This lemma was proved in the context of spaces in [36, 3.4.3] for 2-cubes and [36, 5.5.4] for general $n$-cubes; the same line of argument holds in this more general setting.

The proof of the following lemma is a straightforward exercise, using Proposition 2.2 and its dual.

**Lemma 10.5** *For a commutative square*

**(10.6)**

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & B \\ \downarrow & & \downarrow \\ C & \xrightarrow{\gamma} & D \end{array}$$

*in a pointed right proper model category* $\mathcal{D}$, *the induced map of homotopy fibers*

$$\operatorname{hofiber}(\alpha) \to \operatorname{hofiber}(\gamma)$$

*is a weak equivalence if the square* (10.6) *is a homotopy pullback square. If* $\mathcal{D}$ *is stable and proper, the converse is true as well.*

**Proof of Lemma 10.3** Let $F \to G$ be a fibration in $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$. By an argument similar to the one used to start the proof of Proposition 9.2, it suffices to show that

**(10.7)**

$$\begin{array}{ccc} F & \xrightarrow{\gamma_F} & \Gamma_n F \\ \downarrow & & \downarrow \\ G & \xrightarrow{\gamma_G} & \Gamma_n G \end{array}$$

is a homotopy pullback if and only if the diagram

**(10.8)**

$$
\begin{array}{ccc}
F(\sqcup(A,\varnothing)) & \longrightarrow & \operatorname{holim}_{U \subseteq \mathcal{P}_0(\boldsymbol{n+1})} F(\sqcup(A,U)) \\
\downarrow & & \downarrow \\
F(\sqcup(A,\varnothing)) & \longrightarrow & \operatorname{holim}_{U \subseteq \mathcal{P}_0(\boldsymbol{n+1})} G(\sqcup(A,U))
\end{array}
$$

is a homotopy pullback for all objects $A$ in $\mathcal{C}$.

We can write (10.7) as the composite

$$
\begin{array}{ccc}
F & \xrightarrow{\simeq} \mathbb{F}(F) & \longrightarrow \Gamma_n F \\
\downarrow & \downarrow & \downarrow \\
G & \xrightarrow{\simeq} \mathbb{F}(G) & \longrightarrow \Gamma_n G
\end{array}
$$

by definition of the natural transformation $\gamma\colon \operatorname{id}_{\operatorname{Fun}(\mathcal{C},\mathcal{D})} \to \Gamma_n$. Using the dual of Proposition 2.2 and the fact that $\mathcal{D}$ is stable, we see that the right-hand square is a homotopy pullback if and only if the outer square is a homotopy pullback. Similarly, (10.8) is a homotopy pullback if and only if the corresponding square with fibrant replacements of $F$ and $G$ on the left is a homotopy pullback square. Hence, we can restrict to the case where we use fibrant replacements of $F$ and $G$ and for simplicity, we suppress the fibrant replacement notation $\mathbb{F}$ for the remainder of the proof.

By Lemma 10.4, the homotopy fibers of the top and bottom horizontal arrows in (10.8) are $\perp_{n+1} F(A)$ and $\perp_{n+1} G(A)$, respectively. Then by Lemma 10.5, the diagram (10.8) is a homotopy pullback if and only if the induced map of homotopy fibers $\perp_{n+1} F(A) \to \perp_{n+1} G(A)$ is a weak equivalence.

If (10.8) is a homotopy pullback for all objects $A$ in $\mathcal{C}$, it follows that $|\perp_{n+1}^{*+1} F| \to |\perp_{n+1}^{*+1} G|$ is a weak equivalence in $\operatorname{Fun}(\mathcal{C},\mathcal{D})$. Consider the diagram

$$
\begin{array}{ccc}
|\perp_{n+1}^{*+1} F| & \longrightarrow F & \longrightarrow \Gamma_n F \\
\downarrow & \downarrow & \downarrow \\
|\perp_{n+1}^{*+1} G| & \longrightarrow G & \longrightarrow \Gamma_n G
\end{array}
$$

where the top and bottom rows are the homotopy cofiber sequences defining $\Gamma_n F$ and $\Gamma_n G$. respectively. Since $\mathcal{D}$ is stable, the rows are also homotopy fiber sequences, and the right-hand square, which is exactly (10.7), is a homotopy pullback by Lemma 10.5.

Conversely, if (10.7) is a homotopy pullback, then the square

$$
\begin{array}{ccc}
\perp_{n+1} F & \longrightarrow & \perp_{n+1} \Gamma_n F \\
\downarrow & & \downarrow \\
\perp_{n+1} G & \longrightarrow & \perp_{n+1} \Gamma_n G
\end{array}
$$

is also a homotopy pullback as $\perp_{n+1}$ preserves homotopy pullbacks. As noted in the proof of Theorem 7.2, $\perp_{n+1} \Gamma_n F \simeq * \simeq \perp_{n+1} \Gamma_n G$, so $\perp_{n+1} F \to \perp_{n+1} G$ is a weak equivalence by Proposition 2.3, and (10.8) is a homotopy pullback for all objects $A$ in $\mathcal{C}$. $\qquad\square$

**Proof of Theorem 10.1** The proof of Theorem 7.2 establishes that $\Gamma_n$ is a Bousfield endofunctor on $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ with the projective model structure. Proposition 4.6 guarantees that the fibrations in $\mathrm{Fun}(\mathcal{C}, \mathcal{D})_{\Gamma_n}$ are also projective fibrations. Hence, we can apply Theorem 8.3, using the collection of test morphisms for $\Gamma_n$ that we established in Lemma 10.3. $\qquad\square$

# References

[1] **G Arone**, **M Mahowald**, *The Goodwillie tower of the identity functor and the unstable periodic homotopy of spheres*, Invent. Math. 135 (1999) 743–788  MR

[2] **L Bandklayder**, **J E Bergner**, **R Griffiths**, **B Johnson**, **R Santhanam**, *Enriched functor categories for functor calculus*, Topology Appl. 316 (2022) art. id. 108099  MR

[3] **D Barnes**, **R Eldred**, *Capturing Goodwillie's derivative*, J. Pure Appl. Algebra 220 (2016) 197–222  MR

[4] **D Barnes**, **R Eldred**, *Comparing the orthogonal and homotopy functor calculi*, J. Pure Appl. Algebra 220 (2016) 3650–3675  MR

[5] **D Barnes**, **P Oman**, *Model categories for orthogonal calculus*, Algebr. Geom. Topol. 13 (2013) 959–999  MR

[6] **K Bauer**, **B Johnson**, **R McCarthy**, *Cross effects and calculus in an unbased setting*, Trans. Amer. Math. Soc. 367 (2015) 6671–6718  MR

[7] **K Bauer**, **B Johnson**, **C Osborne**, **E Riehl**, **A Tebbe**, *Directional derivatives and higher order chain rules for abelian functor calculus*, Topology Appl. 235 (2018) 375–427  MR

[8] **M Behrens**, *The Goodwillie tower and the EHP sequence*, Mem. Amer. Math. Soc. 1026, Amer. Math. Soc., Providence, RI (2012)  MR

[9] **G Biedermann**, **B Chorny**, **O Röndigs**, *Calculus of functors and model categories*, Adv. Math. 214 (2007) 92–115  MR

[10] **G Biedermann**, **O Röndigs**, *Calculus of functors and model categories, II*, Algebr. Geom. Topol. 14 (2014) 2853–2913  MR

[11] **A K Bousfield**, *On the telescopic homotopy theory of spaces*, Trans. Amer. Math. Soc. 353 (2001) 2391–2426  MR

[12] **A K Bousfield**, **E M Friedlander**, *Homotopy theory of $\Gamma$-spaces, spectra, and bisimplicial sets*, from "Geometric applications of homotopy theory, II", Lecture Notes in Math. 658, Springer (1978) 80–130  MR

[13] **A Dold**, **D Puppe**, *Homologie nicht-additiver Funktoren: Anwendungen*, Ann. Inst. Fourier (Grenoble) 11 (1961) 201–312  MR

[14] **B I Dundas**, **R McCarthy**, *Stable $K$-theory and topological Hochschild homology*, Ann. of Math. 140 (1994) 685–701  MR

[15] **B I Dundas**, **O Röndigs**, **P A Østvær**, *Enriched functors and stable homotopy theory*, Doc. Math. 8 (2003) 409–488  MR

[16]   **S Eilenberg**, **S MacLane**, *Homology theories for multiplicative systems*, Trans. Amer. Math. Soc. 71 (1951) 294–330  MR

[17]   **S Eilenberg**, **S MacLane**, *On the groups $H(\Pi, n)$, II: Methods of computation*, Ann. of Math. 60 (1954) 49–139  MR

[18]   **P G Goerss**, **J F Jardine**, *Simplicial homotopy theory*, Progr. Math. 174, Birkhäuser, Basel (1999)  MR

[19]   **T G Goodwillie**, *Calculus, II: Analytic functors*, $K$-Theory 5 (1992) 295–332  MR

[20]   **T G Goodwillie**, *Calculus, III: Taylor series*, Geom. Topol. 7 (2003) 645–711  MR

[21]   **T G Goodwillie**, **M Weiss**, *Embeddings from the point of view of immersion theory, II*, Geom. Topol. 3 (1999) 103–118  MR

[22]   **B J Guillou**, **J P May**, *Enriched model categories and presheaf categories*, New York J. Math. 26 (2020) 37–91  MR

[23]   **G Heuts**, *Lie algebras and $v_n$-periodic spaces*, Ann. of Math. 193 (2021) 223–301  MR

[24]   **P S Hirschhorn**, *Model categories and their localizations*, Math. Surv. Monogr. 99, Amer. Math. Soc., Providence, RI (2003)  MR

[25]   **P S Hirschhorn**, *Notes on homotopy colimits and homotopy limits*, notes in progress (2014)  Available at `http://www-math.mit.edu/~psh/notes/hocolim.pdf`

[26]   **M Hovey**, *Model categories*, Math. Surv. Monogr. 63, Amer. Math. Soc., Providence, RI (1999)  MR

[27]   **B Johnson**, **R McCarthy**, *Linearization, Dold–Puppe stabilization, and Mac Lane's Q-construction*, Trans. Amer. Math. Soc. 350 (1998) 1555–1593  MR

[28]   **B Johnson**, **R McCarthy**, *Deriving calculus with cotriples*, Trans. Amer. Math. Soc. 356 (2004) 757–803  MR

[29]   **M R Kantorovitz**, **R McCarthy**, *The Taylor towers for rational algebraic K-theory and Hochschild homology*, Homology Homotopy Appl. 4 (2002) 191–212  MR

[30]   **G M Kelly**, *Basic concepts of enriched category theory*, Lond. Math. Soc. Lect. Note Ser. 64, Cambridge Univ. Press (1982)  MR

[31]   **N J Kuhn**, *Goodwillie towers and chromatic homotopy: an overview*, from "Proceedings of the Nishida Fest", Geom. Topol. Monogr. 10, Geom. Topol. Publ., Coventry (2007) 245–279  MR

[32]   **N J Kuhn**, *The Whitehead conjecture, the tower of $S^1$ conjecture, and Hecke algebras of type A*, J. Topol. 8 (2015) 118–146  MR

[33]   **P Lambrechts**, **V Turchin**, **I Volić**, *The rational homology of spaces of long knots in codimension $> 2$*, Geom. Topol. 14 (2010) 2151–2187  MR

[34]   **J Lurie**, $(\infty, 2)$-*categories and the Goodwillie calculus, I*, preprint (2009)  arXiv 0905.0462

[35]   **R McCarthy**, *Relative algebraic K-theory and topological cyclic homology*, Acta Math. 179 (1997) 197–222  MR

[36]   **B A Munson**, **I Volić**, *Cubical homotopy theory*, New Math. Monogr. 25, Cambridge Univ. Press (2015)  MR

[37]   **L Pereira**, *A general context for Goodwillie calculus*, preprint (2013)  arXiv 1301.2832

[38]   **R Reis**, **M Weiss**, *Rational Pontryagin classes and functor calculus*, J. Eur. Math. Soc. 18 (2016) 1769–1811  MR

[39]  **O Renaudin**, *Localisation homotopique et tour de Taylor pour une catégorie abélienne*, Trans. Amer. Math. Soc. 354 (2002) 75–89  MR

[40]  **B Richter**, *Taylor towers for $\Gamma$-modules*, Ann. Inst. Fourier (Grenoble) 51 (2001) 995–1023  MR

[41]  **M Shulman**, *Homotopy limits and colimits and enriched homotopy theory*, preprint (2006)  arXiv math/0610194

[42]  **D P Sinha**, *The topology of spaces of knots: cosimplicial models*, Amer. J. Math. 131 (2009) 945–980  MR

[43]  **A E Stanculescu**, *Note on a theorem of Bousfield and Friedlander*, Topology Appl. 155 (2008) 1434–1438  MR

[44]  **N Taggart**, *Comparing the orthogonal and unitary functor calculi*, Homology Homotopy Appl. 23 (2021) 227–256  MR

[45]  **N Taggart**, *Unitary functor calculus with reality*, Glasg. Math. J. 64 (2022) 197–230  MR

[46]  **C A Weibel**, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. 38, Cambridge Univ. Press (1994)  MR

[47]  **M Weiss**, *Orthogonal calculus*, Trans. Amer. Math. Soc. 347 (1995) 3743–3796  MR

[48]  **M Weiss**, *Embeddings from the point of view of immersion theory, I*, Geom. Topol. 3 (1999) 67–101  MR

LB:   *MIT Plasma Science and Fusion Center*
*Boston, MA, United States*

JEB:   *Department of Mathematics, University of Virginia*
*Charlottesville, VA, United States*

RG:   *Department of Mathematics, Bowdoin College*
*Brunswick, ME, United States*

BJ:   *Department of Mathematics, Union College*
*Schenectady, NY, United States*

RS:   *Department of Mathematics, Indian Institute of Technology Bombay*
*Mumbai, India*

lbandklayder@gmail.com,   jeb2md@virginia.edu,   r.griffiths@bowdoin.edu, johnsonb@union.edu,   reksan@iitb.ac.in

# Endomorphisms of Artin groups of type $D$

FABRICE CASTEL

LUIS PARIS

We determine a classification of the endomorphisms of the Artin group $A[D_n]$ of type $D_n$ for $n \geq 6$. In particular we determine its automorphism group and its outer automorphism group. We also determine a classification of the homomorphisms from $A[D_n]$ to the Artin group $A[A_{n-1}]$ of type $A_{n-1}$ and a classification of the homomorphisms from $A[A_{n-1}]$ to $A[D_n]$ for $n \geq 6$. We show that any endomorphism of the quotient $A[D_n]/Z(A[D_n])$ lifts to an endomorphism of $A[D_n]$ for $n \geq 4$. We deduce a classification of the endomorphisms of $A[D_n]/Z(A[D_n])$, we determine the automorphism and outer automorphism groups of $A[D_n]/Z(A[D_n])$, and we show that $A[D_n]/Z(A[D_n])$ is co-Hopfian for $n \geq 6$. The results are algebraic in nature but the proofs are based on topological arguments (curves on surfaces and mapping class groups).

20F36; 57K20

## 1 Introduction

Let $S$ be a finite set. A *Coxeter matrix* over $S$ is a square matrix $M = (m_{s,t})_{s,t \in S}$ indexed by the elements of $S$, with coefficients in $\mathbb{N} \cup \{\infty\}$, such that $m_{s,s} = 1$ for all $s \in S$ and $m_{s,t} = m_{t,s} \geq 2$ for all $s, t \in S$ with $s \neq t$. Such a matrix is usually represented by a labeled graph $\Gamma$, called a *Coxeter graph*, defined as follows. The set of vertices of $\Gamma$ is $S$. Two vertices $s, t \in S$ are connected by an edge if $m_{s,t} \geq 3$, and this edge is labeled with $m_{s,t}$ if $m_{s,t} \geq 4$.

If $a$ and $b$ are two letters and $m$ is an integer $\geq 2$, then we denote by $\Pi(a, b, m)$ the word $aba \cdots$ of length $m$. In other words $\Pi(a, b, m) = (ab)^{m/2}$ if $m$ is even and $\Pi(a, b, m) = (ab)^{(m-1)/2}a$ if $m$ is odd. Let $\Gamma$ be a Coxeter graph and let $M = (m_{s,t})_{s,t \in S}$ be its Coxeter matrix. With $\Gamma$ we associate a group $A[\Gamma]$, called the *Artin group* of $\Gamma$, defined by the following presentation:

$$A[\Gamma] = \langle S \mid \Pi(s, t, m_{s,t}) = \Pi(t, s, m_{s,t}) \text{ for } s, t \in S, s \neq t, m_{s,t} \neq \infty \rangle.$$

The *Coxeter group* of $\Gamma$, denoted by $W[\Gamma]$, is the quotient of $A[\Gamma]$ by the relations $s^2 = 1$ for $s \in S$.

Despite the popularity of Artin groups, little is known on their automorphisms and even less on their endomorphisms. The most emblematic cases are the braid groups and the right-angled Artin groups. Recall that the *braid group* on $n + 1$ strands is the Artin group $A[A_n]$ where $A_n$ is the Coxeter graph depicted in Figure 1, and an Artin group $A[\Gamma]$ is called a *right-angled Artin group* if $m_{s,t} \in \{2, \infty\}$ for all

Figure 1: The Coxeter graph $A_n$.

$s, t \in S$ with $s \neq t$. The automorphism group of $A[A_n]$ was determined by Dyer and Grossman [26] and the set of its endomorphisms by Castel [12] for $n \geq 5$, by Chen, Kordek and Margalit [17] for $n \geq 3$ and by Orevkov [35] for $n \geq 2$ (see also Bell and Margalit [2] and Kordek and Margalit [31]). On the other hand there are many articles studying automorphism groups of right-angled Artin groups (see Charney and Vogtmann [15; 16], Day [23; 24], Laurence [33] and Bregman, Charney and Vogtmann [8] for example), but almost nothing is known on endomorphisms of these groups.

Apart from these two families little is known on automorphisms of Artin groups. The automorphism groups of two-generator Artin groups were determined by Gilbert, Howie, Metaftsis and Raptis [29], the automorphism groups of the Artin groups of type $B_n$, $\widetilde{A}_n$ and $\widetilde{C}_n$ were determined by Charney and Crisp [14], the automorphisms groups of some 2-dimensional Artin groups were determined by Crisp [20] and by An and Cho [1], the automorphism groups of large-type free-of-infinity Artin groups were determined by Vaskou [43], and the automorphism group of $A[D_4]$ was determined by Soroko [41]. On the other hand, as far as we know the set of endomorphisms of an Artin group is not determined for any Artin group except for those of type $A_n$.

Recall that an Artin group $A[\Gamma]$ is of *spherical type* if $W[\Gamma]$ is finite. The study of spherical-type Artin groups began in the early 1970s with works by Brieskorn [9; 10], Brieskorn and Saito [11] and Deligne [25], which marked in a way the beginning of the theory of Artin groups. This family, and that of right-angled Artin groups, are the two most-studied and best-understood families of Artin groups and, obviously, any question on Artin groups first arises for Artin groups of spherical type and for right-angled Artin groups. Here we are interested in Artin groups of spherical type, and more particularly in those of type $D_n$.

An Artin group $A[\Gamma]$ is called *irreducible* if $\Gamma$ is connected. If $\Gamma_1, \ldots, \Gamma_l$ are the connected components of $\Gamma$, then $A[\Gamma] = A[\Gamma_1] \times \cdots \times A[\Gamma_l]$ and $W[\Gamma] = W[\Gamma_1] \times \cdots \times W[\Gamma_l]$. In particular $A[\Gamma]$ is of spherical type if and only if $A[\Gamma_i]$ is of spherical type for all $i \in \{1, \ldots, l\}$. So to classify Artin groups of spherical type it suffices to classify those which are irreducible. Finite irreducible Coxeter groups, and hence irreducible Artin groups of spherical type, were classified by Coxeter [18; 19]. There are four infinite families, $A_n$ ($n \geq 1$), $B_n$ ($n \geq 2$), $D_n$ ($n \geq 4$) and $I_2(m)$ ($m \geq 5$), and six "sporadic" groups, $E_6$, $E_7$, $E_8$, $F_4$, $H_3$ and $H_4$. As mentioned above, the automorphism group of $A[\Gamma]$ for $\Gamma$ of type $A_n$ ($n \geq 1$), $B_n$ ($n \geq 2$) and $I_2(m)$ ($m \geq 5$) is known. The next step is therefore to understand the automorphism group of $A[D_n]$ for $n \geq 5$ (the case $\Gamma = D_4$ is known by Soroko [41]). The Coxeter graph $D_n$ is illustrated in Figure 2.



Figure 2: The Coxeter graph $D_n$.

Here we determine a complete and precise classification of the endomorphisms of $A[D_n]$ for $n \geq 6$ (see Theorem 2.3). In particular we determine the automorphism group and the outer automorphism group of $A[D_n]$ for $n \geq 6$ (see Corollary 2.6). We also determine a complete and precise classification of the homomorphisms from $A[D_n]$ to $A[A_{n-1}]$ (see Theorem 2.1) and a complete and precise classification of the homomorphisms from $A[A_{n-1}]$ to $A[D_n]$ (see Theorem 2.2). Note that all these results were announced but not proved in Castel [13]; actually the proofs turn out to be much more difficult than the first author thought when he announced them. Note also that our techniques cannot be used to treat the cases $n = 4$ and $n = 5$. In particular we do not know how to determine $\mathrm{Aut}(A[D_5])$.

From our main result we deduce a classification of the endomorphisms of the group $A[D_n]/Z(A[D_n])$ for $n \geq 6$, where $Z(A[D_n])$ denotes the center of $A[D_n]$ (see Theorem 2.8). Then we determine the automorphism group and the outer automorphism group of $A[D_n]/Z(A[D_n])$ (see Corollary 2.10), and we show that $A[D_n]/Z(A[D_n])$ is co-Hopfian (see Corollary 2.11). These results follow from Theorem 2.3 and Proposition 2.7, which states that any endomorphism of $A[D_n]/Z(A[D_n])$ lifts to an endomorphism of $A[D_n]$. Such results were previously known for braid groups, that is, Artin groups of type $A_n$ (see Bell and Margalit [2]). Note that the application of our main result to the study of $A[D_n]/Z(A[D_n])$ was not present in an earlier version of the paper. It was suggested to us by the referee, for which we extend our warm thanks.

A *geometric representation* of an Artin group is a homomorphism from the group to a mapping class group (see Section 3 for more details). In order to achieve our goals we make a study of a particular geometric representation of $A[D_n]$ previously introduced by Perron and Vannier [40] with one boundary component replaced by a puncture. This geometric representation will be the key tool for many of our proofs. Overall, although the results of the paper are algebraic in nature, the proofs are mostly based on topological arguments (on curves on surfaces and mapping class groups).

The paper is organized as follows. In Section 2 we give the main definitions and precise statements of the main results. Section 3 is dedicated to the study of some geometric representations of Artin groups of type $A_n$ and type $D_n$. In Section 4 we determine the homomorphisms from $A[D_n]$ to $A[A_{n-1}]$, in Section 5 we determine the homomorphisms from $A[A_{n-1}]$ to $A[D_n]$, and in Section 6 we determine the endomorphisms of $A[D_n]$. In Section 7 we determine the endomorphisms of $A[D_n]/Z(A[D_n])$.

# 2 Definitions and statements

For $n \geq 4$ we denote by $s_1, \ldots, s_{n-1}$ the standard generators of $A[A_{n-1}]$ numbered as in Figure 1 and by $t_1, \ldots, t_n$ the standard generators of $A[D_n]$ numbered as in Figure 2.

Let $\Gamma$ be a Coxeter graph. For $X \subset S$ we denote by $A_X = A_X[\Gamma]$ the subgroup of $A = A[\Gamma]$ generated by $X$, by $W_X = W_X[\Gamma]$ the subgroup of $W = W[\Gamma]$ generated by $X$, and by $\Gamma_X$ the full subgraph of $\Gamma$ spanned by $X$. We know from van der Lek [34] that $A_X$ is the Artin group of $\Gamma_X$ and from Bourbaki [7] that $W_X$ is the Coxeter group of $\Gamma_X$. A subgroup of the form $A_X$ is called a *standard parabolic subgroup* of $A$ and a subgroup of the form $W_X$ is called a *standard parabolic subgroup* of $W$.

For $w \in W$ we denote by $\lg(w)$ the word length of $w$ with respect to $S$. A *reduced expression* for $w$ is an expression $w = s_1 s_2 \cdots s_l$ of minimal length, that is, such that $l = \lg(w)$. Let $\omega \colon A \to W$ be the natural epimorphism which sends $s$ to $s$ for all $s \in S$. This epimorphism has a natural set-section $\tau \colon W \to A$ defined as follows. Let $w \in W$ and let $w = s_1 s_2 \cdots s_l$ be a reduced expression for $w$. Then $\tau(w) = s_1 s_2 \cdots s_l \in A$. We know from Tits [42] that the definition of $\tau(w)$ does not depend on the choice of its reduced expression.

Assume $\Gamma$ is of spherical type. Then $W$ has a unique element of maximal length, denoted by $w_S$, which satisfies $w_S^2 = 1$ and $w_S S w_S = S$. The *Garside element* of $A$ is defined to be $\Delta = \Delta[\Gamma] = \tau(w_S)$. We know that $\Delta S \Delta^{-1} = S$ and, if $\Gamma$ is connected, then the center $Z(A)$ of $A$ is an infinite cyclic group generated by either $\Delta$ or $\Delta^2$ (see Brieskorn and Saito [11]). For $X \subset S$ we denote by $w_X$ the element of maximal length in $W_X$ and by $\Delta_X = \Delta_X[\Gamma] = \tau(w_X)$ the Garside element of $A_X$.

If $\Gamma = A_{n-1}$, then

$$\Delta = (s_{n-1} \cdots s_1)(s_{n-1} \cdots s_2) \cdots (s_{n-1} s_{n-2}) s_{n-1},$$

$\Delta s_i \Delta^{-1} = s_{n-i}$ for all $1 \le i \le n-1$ and $Z(A)$ is generated by $\Delta^2$. If $\Gamma = D_n$, then

$$\Delta = (t_1 \cdots t_{n-2} t_{n-1} t_n t_{n-2} \cdots t_1)(t_2 \cdots t_{n-2} t_{n-1} t_n t_{n-2} \cdots t_2) \cdots (t_{n-2} t_{n-1} t_n t_{n-2})(t_{n-1} t_n).$$

If $n$ is even, then $\Delta t_i \Delta^{-1} = t_i$ for all $1 \le i \le n$ and $Z(A)$ is generated by $\Delta$. If $n$ is odd, then $\Delta t_i \Delta^{-1} = t_i$ for all $1 \le i \le n-2$, $\Delta t_{n-1} \Delta^{-1} = t_n$, $\Delta t_n \Delta^{-1} = t_{n-1}$ and $Z(A)$ is generated by $\Delta^2$.

If $G$ is a group and $g \in G$, then we denote by $\mathrm{ad}_g \colon G \to G$, $h \mapsto ghg^{-1}$, the conjugation map by $g$. We say that two homomorphisms $\varphi_1, \varphi_2 \colon G \to H$ are *conjugate* if there exists $h \in H$ such that $\varphi_2 = \mathrm{ad}_h \circ \varphi_1$.

A homomorphism $\varphi \colon G \to H$ is called *abelian* if its image is an abelian subgroup of $H$. A homomorphism $\varphi \colon G \to H$ is called *cyclic* if its image is a cyclic subgroup of $H$. If $G = A[A_{n-1}]$, then $\varphi \colon A[A_{n-1}] \to H$ is abelian if and only if it is cyclic, if and only if there exists $h \in H$ such that $\varphi(s_i) = h$ for all $1 \le i \le n-1$. Similarly, if $G = A[D_n]$, then $\varphi \colon A[D_n] \to H$ is abelian if and only if it is cyclic, if and only if there exists $h \in H$ such that $\varphi(t_i) = h$ for all $1 \le i \le n$.

Two automorphisms $\zeta, \chi \in \mathrm{Aut}(A[D_n])$ play a central role in our study. These are defined by

$$\zeta(t_i) = t_i \quad \text{for } 1 \le i \le n-2, \qquad \zeta(t_{n-1}) = t_n, \qquad \zeta(t_n) = t_{n-1}, \qquad \chi(t_i) = t_i^{-1} \quad \text{for } 1 \le i \le n.$$

Both are of order 2 and commute, and hence they generate a subgroup of $\mathrm{Aut}(A[D_n])$ isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. If $n$ is odd, then $\zeta$ is the conjugation map by $\Delta = \Delta[D_n]$. On the other hand, if $n$ is even, then $\zeta$ is not an inner automorphism (see Paris [36]). The automorphism $\chi$ is never inner.

Two other homomorphisms play an important role in our study. The first, $\pi\colon A[D_n] \to A[A_{n-1}]$, is defined by

$$\pi(t_i) = s_i \quad \text{for } 1 \leq i \leq n-2, \qquad \pi(t_{n-1}) = \pi(t_n) = s_{n-1}.$$

The second, $\iota\colon A[A_{n-1}] \to A[D_n]$, is defined by

$$\iota(s_i) = t_i \quad \text{for } 1 \leq i \leq n-1.$$

Observe that $\pi \circ \iota = \mathrm{id}_{A[A_{n-1}]}$, and hence $\pi$ is surjective, $\iota$ is injective and $A[D_n] \simeq \mathrm{Ker}(\pi) \rtimes A[A_{n-1}]$. We refer to Crisp and Paris [21] for a detailed study on this decomposition of $A[D_n]$ as a semidirect product.

Let $n \geq 4$. For $p \in \mathbb{Z}$ we define a homomorphism $\alpha_p\colon A[D_n] \to A[A_{n-1}]$ by

$$\alpha_p(t_i) = s_i \Delta^{2p} \quad \text{for } 1 \leq i \leq n-2, \qquad \alpha_p(t_{n-1}) = \alpha_p(t_n) = s_{n-1}\Delta^{2p},$$

where $\Delta = \Delta[A_{n-1}]$ is the Garside element of $A[A_{n-1}]$. Note that $\alpha_0 = \pi$.

Set $Y = \{t_1, \dots, t_{n-1}\}$. For $p, q \in \mathbb{Z}$ we define a homomorphism $\beta_{p,q}\colon A[A_{n-1}] \to A[D_n]$ by

$$\beta_{p,q}(s_i) = t_i \Delta_Y^{2p} \Delta^{\kappa q} \quad \text{for } 1 \leq i \leq n-1,$$

where $\Delta = \Delta[D_n]$ is the Garside element of $A[D_n]$, $\Delta_Y = \Delta_Y[D_n]$, $\kappa = 2$ if $n$ is odd, and $\kappa = 1$ if $n$ is even. Note that $\beta_{0,0} = \iota$. Note also that, by Paris [36, Theorem 1.1], the centralizer of $Y$ in $A[D_n]$ is the free abelian group of rank 2 generated by $\Delta_Y^2$ and $\Delta^\kappa$.

For $p \in \mathbb{Z}$ we define the homomorphism $\gamma_p\colon A[D_n] \to A[D_n]$ by

$$\gamma_p(t_i) = t_i \Delta^{\kappa p} \quad \text{for } 1 \leq i \leq n,$$

where $\Delta = \Delta[D_n]$ is the Garside element of $A[D_n]$, $\kappa = 2$ if $n$ is odd, and $\kappa = 1$ if $n$ is even. Note that $\gamma_0 = \mathrm{id}$.

Concerning $A[A_{n-1}]$, we define an automorphism $\bar{\chi}\colon A[A_{n-1}] \to A[A_{n-1}]$ by

$$\bar{\chi}(s_i) = s_i^{-1} \quad \text{for } 1 \leq i \leq n-1,$$

and for $p \in \mathbb{Z}$ we define an endomorphism $\bar{\gamma}_p\colon A[A_{n-1}] \to A[A_{n-1}]$ by

$$\bar{\gamma}_p(s_i) = s_i \Delta^{2p} \quad \text{for } 1 \leq i \leq n-1,$$

where $\Delta$ is the Garside element of $A[A_{n-1}]$.

The main results of this paper are the following.

**Theorem 2.1** *Let $n \geq 5$. Let $\varphi\colon A[D_n] \to A[A_{n-1}]$ be a homomorphism. Then up to conjugation we have one of the following two possibilities:*

  (1)  *$\varphi$ is cyclic.*

  (2)  *There exist $\psi \in \langle \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi = \alpha_p \circ \psi$.*

**Theorem 2.2** *Let $n \geq 6$. Let $\varphi \colon A[A_{n-1}] \to A[D_n]$ be a homomorphism. Then up to conjugation we have one of the following two possibilities*:

  (1)  *$\varphi$ is cyclic.*

  (2)  *There exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi = \psi \circ \beta_{p,q}$.*

**Theorem 2.3** *Let $n \geq 6$. Let $\varphi \colon A[D_n] \to A[D_n]$ be a homomorphism. Then up to conjugation we have one of the following three possibilities*:

  (1)  *$\varphi$ is cyclic.*

  (2)  *There exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi = \psi \circ \beta_{p,q} \circ \pi$.*

  (3)  *There exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi = \psi \circ \gamma_p$.*

From Theorem 2.3 we deduce a classification of the injective endomorphisms and of the automorphisms of $A[D_n]$ as follows.

**Corollary 2.4** *Let $n \geq 6$. Let $\varphi \colon A[D_n] \to A[D_n]$ be an endomorphism. Then $\varphi$ is injective if and only if there exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi$ is conjugate to $\psi \circ \gamma_p$.*

**Proof** Let $\varphi \colon A[D_n] \to A[D_n]$ be an endomorphism. By Theorem 2.3 we have one of the following three possibilities, up to conjugation:

  (1)  $\varphi$ is cyclic.

  (2)  There exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi = \psi \circ \beta_{p,q} \circ \pi$.

  (3)  There exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi = \psi \circ \gamma_p$.

If $\varphi$ is cyclic, then $\varphi(t_{n-1}) = \varphi(t_n)$, and hence $\varphi$ is not injective. If there exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi = \psi \circ \beta_{p,q} \circ \pi$, then, again, $\varphi(t_{n-1}) = \varphi(t_n)$, and hence $\varphi$ is not injective. So, if $\varphi$ is injective, then there exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi$ is conjugate to $\psi \circ \gamma_p$.

It remains to show that, if $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$, then $\psi \circ \gamma_p$ is injective. Since the elements of $\langle \zeta, \chi \rangle$ are automorphisms, it suffices to show that $\gamma_p$ is injective. We denote by $z \colon A[D_n] \to \mathbb{Z}$ the homomorphism which sends $t_i$ to 1 for all $1 \leq i \leq n$. It is easily seen that $\gamma_p(u) = u \Delta^{\kappa p z(u)}$ for all $u \in A[D_n]$. Let $u \in \operatorname{Ker}(\gamma_p)$. Then $1 = \gamma_p(u) = u \Delta^{\kappa p z(u)}$, and hence $u = \Delta^q$ where $q = -\kappa p z(u)$. We have $z(\Delta) = n(n-1)$, and hence $z(u) = q n(n-1)$, thus

$$1 = \gamma_p(u) = \Delta^q \Delta^{\kappa p q n(n-1)} = \Delta^{q(1 + \kappa p n(n-1))}.$$

Since $1 + \kappa p n(n-1) \neq 0$, this equality implies that $q = 0$, and hence $u = 1$. So $\gamma_p$ is injective. $\qquad\square$

**Corollary 2.5** *Let $n \geq 6$. Let $\varphi \colon A[D_n] \to A[D_n]$ be an endomorphism. Then $\varphi$ is an automorphism if and only if it is conjugate to an element of $\langle \zeta, \chi \rangle$.*

**Proof** Clearly, if $\varphi$ is conjugate to an element of $\langle \zeta, \chi \rangle$, then $\varphi$ is an automorphism. Conversely, suppose that $\varphi$ is an automorphism. We know from Corollary 2.4 that there exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi$ is conjugate to $\psi \circ \gamma_p$. Thus, up to conjugation and up to composing on the left by $\psi^{-1}$, we can assume that $\varphi = \gamma_p$. It remains to show that $p = 0$.

Again let $z \colon A[D_n] \to \mathbb{Z}$ be the homomorphism which sends $t_i$ to 1 for all $1 \le i \le n$. Recall that $\gamma_p(u) = u \Delta^{\kappa p z(u)}$ for all $u \in A[D_n]$. For $u \in A[D_n]$, we have

$$(z \circ \gamma_p)(u) = (1 + n(n-1)\kappa p)z(u) \in (1 + n(n-1)\kappa p)\mathbb{Z}.$$

Since $\gamma_p$ is an automorphism, $z \circ \gamma_p$ is surjective, and hence $\mathbb{Z} = \mathrm{Im}(z \circ \gamma_p) \subset (1 + n(n-1)\kappa p)\mathbb{Z}$. It follows that $(1 + n(n-1)\kappa p) \in \{\pm 1\}$, and hence $p = 0$. $\qquad \square$

By combining Corollary 2.5 with Crisp and Paris [21, Theorem 4.9] we immediately obtain the following.

**Corollary 2.6** *Let $n \ge 6$.*

(1) *If $n$ is even, then*

$$\mathrm{Aut}(A[D_n]) = \mathrm{Inn}(A[D_n]) \rtimes \langle \zeta, \chi \rangle \simeq (A[D_n]/Z(A[D_n])) \rtimes (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}),$$

*and $\mathrm{Out}(A[D_n]) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, where $Z(A[D_n])$ denotes the center of $A[D_n]$.*

(2) *If $n$ is odd, then*

$$\mathrm{Aut}(A[D_n]) = \mathrm{Inn}(A[D_n]) \rtimes \langle \chi \rangle \simeq (A[D_n]/Z(A[D_n])) \rtimes (\mathbb{Z}/2\mathbb{Z}),$$

*and $\mathrm{Out}(A[D_n]) \simeq \mathbb{Z}/2\mathbb{Z}$.*

We denote by $Z(A[D_n])$ the center of $A[D_n]$, we set $A_Z[D_n] = A[D_n]/Z(A[D_n])$ and we denote by $\xi \colon A[D_n] \to A_Z[D_n]$ the canonical projection. For each $1 \le i \le n$, we set $t_{Z,i} = \xi(t_i)$. Note that an endomorphism $\varphi \colon A[D_n] \to A[D_n]$ induces an endomorphism $\varphi_Z \colon A_Z[D_n] \to A_Z[D_n]$ if and only if $\varphi(Z(A[D_n])) \subset Z(A[D_n])$. We say that an endomorphism $\psi \colon A_Z[D_n] \to A_Z[D_n]$ *lifts* if there exists an endomorphism $\varphi \colon A[D_n] \to A[D_n]$ such that $\varphi_Z = \psi$. Then we call $\varphi$ a *lift* of $\psi$. In Section 7 we prove the following.

**Proposition 2.7** *Let $n \ge 4$. Then every endomorphism of $A_Z[D_n]$ lifts.*

From this proposition combined with Theorem 2.3 we will deduce the following.

**Theorem 2.8** *Let $n \ge 6$. Let $\varphi_Z \colon A_Z[D_n] \to A_Z[D_n]$ be an endomorphism. Then we have one of the following two possibilities, up to conjugation:*

(1) *$\varphi_Z$ is cyclic.*

(2) *$\varphi_Z \in \langle \zeta_Z, \chi_Z \rangle$.*

In addition to Theorem 2.8 we have the following.

**Proposition 2.9** *Let $n \geq 4$. There are only finitely many conjugacy classes of cyclic endomorphisms of $A_Z[D_n]$.*

**Proof** Let $\varphi_Z \colon A_Z[D_n] \to A_Z[D_n]$ be a cyclic endomorphism. There exists $g_Z \in A_Z[D_n]$ such that $\varphi_Z(t_{Z,i}) = g_Z$ for all $1 \leq i \leq n$. We denote by $\Delta$ the Garside element of $A[D_n]$, and we set $\kappa = 2$ if $n$ is odd and $\kappa = 1$ if $n$ is even. We have $1 = (\varphi_Z \circ \xi)(\Delta^\kappa) = g_Z^{\kappa n(n-1)}$, and hence $g_Z$ is of finite order. By Bestvina [3, Theorem 4.5] there are finitely many conjugacy classes of finite subgroups in $A_Z[D_n]$. Since $\langle g_Z \rangle$ is a finite subgroup of $A_Z[D_n]$, it follows that there are finitely many choices for $g_Z$, up to conjugation. $\square$

In Lemma 7.1 we will show that if $n$ is even then $\langle \zeta_Z, \chi_Z \rangle \cap \mathrm{Inn}(A_Z[D_n]) = \{\mathrm{id}\}$, and if $n$ is odd then $\langle \chi_Z \rangle \cap \mathrm{Inn}(A_Z[D_n]) = \{\mathrm{id}\}$. Furthermore, it is well known and can be easily proved (arguing as in the proof of Cumplido and Paris [22, Proposition 3.1(4)], for example) that the center of $A[\Gamma]/Z(A[\Gamma])$ is trivial for any $A[\Gamma]$ of spherical type. These two remarks combined with Theorem 2.8 imply the following.

**Corollary 2.10** *Let $n \geq 6$.*

(1) *If $n$ is even, then*

$$\mathrm{Aut}(A_Z[D_n]) = \mathrm{Inn}(A_Z[D_n]) \rtimes \langle \zeta_Z, \chi_Z \rangle \simeq A_Z[D_n] \rtimes (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}) \simeq \mathrm{Aut}(A[D_n]),$$

*and $\mathrm{Out}(A_Z[D_n]) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \simeq \mathrm{Out}(A[D_n])$.*

(2) *If $n$ is odd, then*

$$\mathrm{Aut}(A_Z[D_n]) = \mathrm{Inn}(A_Z[D_n]) \rtimes \langle \chi_Z \rangle \simeq A_Z[D_n] \rtimes (\mathbb{Z}/2\mathbb{Z}) \simeq \mathrm{Aut}(A[D_n]),$$

*and $\mathrm{Out}(A_Z[D_n]) \simeq \mathbb{Z}/2\mathbb{Z} \simeq \mathrm{Out}(A[D_n])$.*

A group $G$ is said to be *co-Hopfian* if every injective endomorphism of $G$ is an isomorphism. Another direct consequence of Theorem 2.8 is the following.

**Corollary 2.11** *Let $n \geq 6$. Then $A_Z[D_n]$ is co-Hopfian.*

In addition to the case $D_n$ for $n \geq 6$ shown in Corollary 2.11, the Coxeter graphs $\Gamma$ for which we know that $A[\Gamma]/Z(A[\Gamma])$ is co-Hopfian are the Coxeter graphs $A_n$, $B_n$, $\widetilde{A}_n$ and $\widetilde{C}_n$ for $n \geq 2$ (see Bell and Margalit [2]). Note that, for $\widetilde{A}_n$ and $\widetilde{C}_n$, the center $Z(A[\Gamma])$ is trivial, and hence the above remark means that the Artin group itself is co-Hopfian.

# 3 Geometric representations

Let $\Sigma$ be an oriented compact surface possibly with boundary, and let $\mathcal{P}$ be a finite set of punctures in the interior of $\Sigma$. We denote by $\mathrm{Homeo}^+(\Sigma, \mathcal{P})$ the group of homeomorphisms of $\Sigma$ that preserve the orientation, that are the identity on a neighborhood of the boundary of $\Sigma$ and that setwise leave invariant $\mathcal{P}$.

The *mapping class group* of the pair $(\Sigma, \mathcal{P})$, denoted by $\mathcal{M}(\Sigma, \mathcal{P})$, is the group of isotopy classes of elements of $\mathrm{Homeo}^+(\Sigma, \mathcal{P})$. If $\mathcal{P} = \varnothing$, then we write $\mathcal{M}(\Sigma, \varnothing) = \mathcal{M}(\Sigma)$, and if $\mathcal{P} = \{x\}$ is a singleton, then we write $\mathcal{M}(\Sigma, \mathcal{P}) = \mathcal{M}(\Sigma, x)$. We only give definitions and results on mapping class groups that we need for our proofs and we refer to Farb and Margalit [28] for a complete account on the subject.

Recall that a *geometric representation* of an Artin group $A$ is a homomorphism from $A$ to a mapping class group. Their study is the main ingredient of our proofs. Important tools for constructing and understanding them are Dehn twists and essential reduction systems. So, we start by recalling their definitions and their main properties.

A *circle* of $(\Sigma, \mathcal{P})$ is the (nonoriented) image of an embedding $a \colon \mathbb{S}^1 \hookrightarrow \Sigma \setminus (\partial \Sigma \cup \mathcal{P})$. It is called *generic* if it does not bound any disk containing 0 or 1 puncture and if it is not parallel to any boundary component. The isotopy class of a circle $a$ is denoted by $[a]$. We denote by $\mathcal{C}(\Sigma, \mathcal{P})$ the set of isotopy classes of generic circles of $(\Sigma, \mathcal{P})$. The *intersection number* of two classes $[a], [b] \in \mathcal{C}(\Sigma, \mathcal{P})$ is $i([a], [b]) = \min\{|a' \cap b'| \mid a' \in [a] \text{ and } b' \in [b]\}$. The set $\mathcal{C}(\Sigma, \mathcal{P})$ is endowed with a simplicial complex structure, where a finite set $\mathcal{A}$ is a simplex if $i([a], [b]) = 0$ for all $[a], [b] \in \mathcal{A}$. This complex is called the *curve complex* of $(\Sigma, \mathcal{P})$.

By a *Dehn twist* we mean a right Dehn twist and the (right) Dehn twist along a circle $a$ of $(\Sigma, \mathcal{P})$ will be denoted by $T_a$. The following is an important tool for constructing and understanding geometric representations of Artin groups. Its proof can be found in Farb and Margalit [28, Section 3.5].

**Proposition 3.1** *Let $\Sigma$ be a compact oriented surface and let $\mathcal{P}$ be a finite collection of punctures in the interior of $\Sigma$. Let $a$ and $b$ be two generic circles of $(\Sigma, \mathcal{P})$.*

   (1)  *We have $T_a T_b = T_b T_a$ if and only if $i([a], [b]) = 0$.*
   (2)  *We have $T_a T_b T_a = T_b T_a T_b$ if and only if $i([a], [b]) = 1$.*

Let $f \in \mathcal{M}(\Sigma, \mathcal{P})$. A simplex $\mathcal{A}$ of $\mathcal{C}(\Sigma, \mathcal{P})$ is called a *reduction system* for $f$ if $f(\mathcal{A}) = \mathcal{A}$. In that case any element of $\mathcal{A}$ is called a *reduction class* for $f$. A reduction class $[a]$ is an *essential reduction class* if, for all $[b] \in \mathcal{C}(\Sigma, \mathcal{P})$ such that $i([a], [b]) \neq 0$ and for all $m \in \mathbb{Z} \setminus \{0\}$, we have $f^m([b]) \neq [b]$. In particular, if $[a]$ is an essential reduction class and $[b]$ is any reduction class, then $i([a], [b]) = 0$. We denote by $\mathcal{S}(f)$ the set of essential reduction classes for $f$. The following gathers some key results on $\mathcal{S}(f)$ that will be useful later.

**Theorem 3.2** (Birman, Lubotzky and McCarthy [6]) *Let $\Sigma$ be a compact oriented surface and let $\mathcal{P}$ be a finite set of punctures in the interior of $\Sigma$. Let $f \in \mathcal{M}(\Sigma, \mathcal{P})$.*

   (1)  *If $\mathcal{S}(f) \neq \varnothing$, then $\mathcal{S}(f)$ is a reduction system for $f$. In particular, if $\mathcal{S}(f) \neq \varnothing$, then $\mathcal{S}(f)$ is a simplex of $\mathcal{C}(\Sigma, \mathcal{P})$.*
   (2)  *We have $\mathcal{S}(f^n) = \mathcal{S}(f)$ for all $n \in \mathbb{Z} \setminus \{0\}$.*
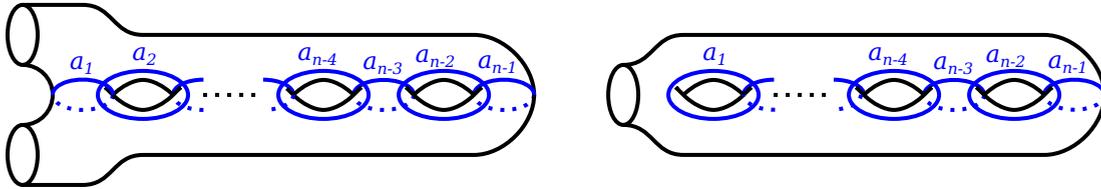   (3)  *We have $\mathcal{S}(gfg^{-1}) = g(\mathcal{S}(f))$ for all $g \in \mathcal{M}(\Sigma, \mathcal{P})$.*

Figure 3: The geometric representation of $A[A_{n-1}]$ for $n$ even (left) and n odd (right).

The following is well known and is a direct consequence of Birman, Lubotzky and McCarthy [6] (see also Castel [12, Corollaire 3.45]). It will be often used in our proofs.

**Proposition 3.3** *Let $\Sigma$ be an oriented compact surface of genus $\geq 2$ and let $\mathcal{P}$ be a finite set of punctures in the interior of $\Sigma$. Let $f_0 \in Z(\mathcal{M}(\Sigma, \mathcal{P}))$ be a central element of $\mathcal{M}(\Sigma, \mathcal{P})$, let $\mathcal{A} = \{[a_1], \dots, [a_p]\}$ be a simplex of $\mathcal{C}(\Sigma, \mathcal{P})$ and let $k_1, \dots, k_p$ be nonzero integers. Let $g = T_{a_1}^{k_1} T_{a_2}^{k_2} \cdots T_{a_p}^{k_p} f_0$. Then $\mathcal{S}(g) = \mathcal{A}$.*

Let $n \geq 4$. If $n$ is even, then $\Sigma_n$ denotes the surface of genus $\frac{1}{2}(n-2)$ with two boundary components, and if $n$ is odd, then $\Sigma_n$ denotes the surface of genus $\frac{1}{2}(n-1)$ with one boundary component. Consider the circles $a_1, \dots, a_{n-1}$ drawn in Figure 3. Then by Proposition 3.1 we have a geometric representation $\rho_A : A[A_{n-1}] \to \mathcal{M}(\Sigma_n)$ which sends $s_i$ to $T_{a_i}$ for all $1 \leq i \leq n-1$. The following is well known; it is a direct consequence of Birman and Hilden [5], and its proof is explicitly given in Perron and Vannier [40].

**Theorem 3.4** (Birman and Hilden [5]) *Let $n \geq 4$. Then $\rho_A : A[A_{n-1}] \to \mathcal{M}(\Sigma_n)$ is injective.*

The following is proved in Castel [12] for $n \geq 6$ using the geometric representation $\rho_A$ defined above. It is proved in Chen, Kordek and Margalit [17] for $n \geq 5$ with a different method.

**Theorem 3.5** (Castel [12], Chen, Kordek and Margalit [17] and Orevkov [35]) *Let $n \geq 5$. Let $\varphi : A[A_{n-1}] \to A[A_{n-1}]$ be a homomorphism. Then up to conjugation we have one of the following two possibilities*:

(1) $\varphi$ *is cyclic.*

(2) *There exist $\psi \in \langle \bar{\chi} \rangle$ and $p \in \mathbb{Z}$ such that $\varphi = \psi \circ \bar{\gamma}_p$.*

Let $n \geq 6$. Pick a puncture $x$ in the interior of $\Sigma_n$ and consider the circles $d_1, \dots, d_n$ drawn in Figure 4. Then by Proposition 3.1 we have a geometric representation $\rho_D : A[D_n] \to \mathcal{M}(\Sigma_n, x)$ which sends $t_i$ to $T_{d_i}$ for all $1 \leq i \leq n$. On the other hand, the embedding of $\mathrm{Homeo}^+(\Sigma_n, x)$ into $\mathrm{Homeo}^+(\Sigma_n)$ induces a surjective homomorphism $\theta : \mathcal{M}(\Sigma_n, x) \to \mathcal{M}(\Sigma_n)$ whose kernel is naturally isomorphic to $\pi_1(\Sigma_n, x)$ (see Birman [4]). It is easily seen that

$$\theta(T_{d_i}) = T_{a_i} \quad \text{for } 1 \leq i \leq n-2, \qquad \theta(T_{d_{n-1}}) = \theta(T_{d_n}) = T_{a_{n-1}},$$
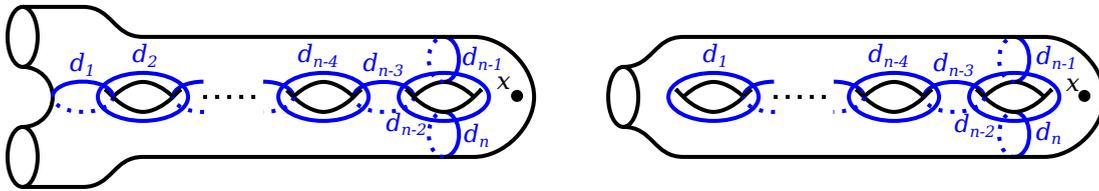
Figure 4: The geometric representation of $A[D_n]$ for $n$ even (left) and $n$ odd (right).

and hence we have the commutative diagram

(3-1)
$$
\begin{array}{ccccccccc}
1 & \longrightarrow & \mathrm{Ker}(\pi) & \longrightarrow & A[D_n] & \xrightarrow{\pi} & A[A_{n-1}] & \longrightarrow & 1 \\
& & \downarrow{\bar\rho} & & \downarrow{\rho_D} & & \downarrow{\rho_A} & & \\
1 & \longrightarrow & \mathrm{Ker}(\theta) & \longrightarrow & \mathcal{M}(\Sigma_n, x) & \xrightarrow{\theta} & \mathcal{M}(\Sigma_n) & \longrightarrow & 1
\end{array}
$$

where we denote by $\bar\rho \colon \mathrm{Ker}(\pi) \to \mathrm{Ker}(\theta)$ the restriction of $\rho_D$ to $\mathrm{Ker}(\pi)$.

The proof of the following can be found in Perron and Vannier [40, Theorem 1] with few modifications. As this result is central in our paper, for the sake of completeness we give a proof. Note that our proof is a little shorter than that of Perron and Vannier [40] because it uses results from Crisp and Paris [21] which were not known and it does not need to deal with some Dehn twist along a boundary component, but our arguments are essentially the same.

**Theorem 3.6** (Perron and Vannier [40]) *Let* $n \geq 4$.

(1) *The homomorphism* $\bar\rho \colon \mathrm{Ker}(\pi) \to \mathrm{Ker}(\theta)$ *is an isomorphism.*

(2) *The geometric representation* $\rho_D \colon A[D_n] \to \mathcal{M}(\Sigma_n, x)$ *is injective.*

**Proof** Part (2) is a consequence of (1) because of the following. Suppose $\bar\rho$ is an isomorphism. Then, since $\rho_A$ is injective, $\rho_D$ is injective by the five lemma applied to (3-1).

Now, we prove (1). We know from Crisp and Paris [21, Proposition 2.3] that $\mathrm{Ker}(\pi)$ is a free group of rank $n - 1$. We also know from Birman [4] that $\mathrm{Ker}(\theta) = \pi_1(\Sigma_n, x)$, which is also a free group of rank $n - 1$. Recall that a group $G$ is Hopfian if every surjective endomorphism $G \to G$ is an isomorphism. It is well known that free groups of finite rank are Hopfian (see de la Harpe [30, Chapter III, Section 19]), and hence in order to show that $\bar\rho$ is an isomorphism it suffices to show that $\bar\rho$ is surjective.

Set $f_{n-1} = T_{d_{n-1}}^{-1} T_{d_n}$. Note that $t_{n-1}^{-1} t_n \in \mathrm{Ker}(\pi)$ and $f_{n-1} = \bar\rho(t_{n-1}^{-1} t_n)$. In particular $f_{n-1} \in \mathrm{Im}(\bar\rho) \subset \mathrm{Ker}(\theta) = \pi_1(\Sigma_n, x)$. This element, seen as an element of $\pi_1(\Sigma_n, x)$, is represented by the loop drawn in Figure 5. For $2 \leq i \leq n - 1$ we define $f_{n-i} \in \pi_1(\Sigma_n, x) \subset \mathcal{M}(\Sigma_n, x)$ by induction on $i$ by setting $f_{n-i} = T_{d_{n-i}} f_{n-i+1} T_{d_{n-i}}^{-1} f_{n-i+1}^{-1}$. The element $f_{n-i}$, viewed as an element of $\pi_1(\Sigma_n, x)$, is represented by the loop drawn in the left-hand side of Figure 6 if $i = 2j$ is even, and by the loop drawn in the right-hand side of Figure 6 if $i = 2j + 1$ is odd, where we compose paths from right to left. Observe that $f_1, \ldots, f_{n-1}$
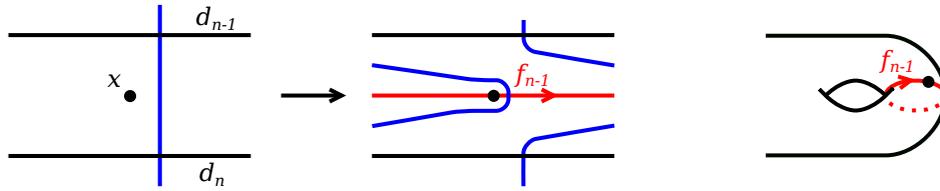
Figure 5: The loop $f_{n-1} \in \pi_1(\Sigma_n, x)$.

generate $\pi_1(\Sigma_n, x)$. So, in order to show that $\bar{\rho}$ is surjective, it suffices to show that $f_{n-i} \in \mathrm{Im}(\bar{\rho})$ for all $i \in \{1, \ldots, n-1\}$. We argue by induction on $i$. We already know that $f_{n-1} = \bar{\rho}(t_{n-1}^{-1} t_n) \in \mathrm{Im}(\bar{\rho})$. Suppose $i \geq 2$ and $f_{n-i+1} \in \mathrm{Im}(\bar{\rho})$. Let $u \in \mathrm{Ker}(\pi)$ such that $f_{n-i+1} = \bar{\rho}(u)$. Since $\mathrm{Ker}(\pi)$ is a normal subgroup of $A[D_n]$, we have $t_{n-i} u t_{n-i}^{-1} \in \mathrm{Ker}(\pi)$; hence $t_{n-i} u t_{n-i}^{-1} u^{-1} \in \mathrm{Ker}(\pi)$, and therefore

$$f_{n-i} = T_{d_{n-i}} f_{n-i+1} T_{d_{n-i}}^{-1} f_{n-i+1}^{-1} = \bar{\rho}(t_{n-i} u t_{n-i}^{-1} u^{-1}) \in \mathrm{Im}(\bar{\rho}). \qquad \square$$

Our last preliminary on geometric representations is a result implicitly proved in Castel [13, Section 3.2], and it is in this theorem that we need the assumption $n \geq 6$.

**Theorem 3.7** (Castel [13]) *Let $n \geq 6$. Let $\varphi \colon A[A_{n-1}] \to \mathcal{M}(\Sigma_n, x)$ be a noncyclic homomorphism. Then there exist generic circles $c_1, \ldots, c_{n-1}$ in $\Sigma_n \setminus \{x\}$, $\varepsilon \in \{\pm 1\}$ and $g \in \mathcal{M}(\Sigma_n, x)$ such that*

(a) $|c_i \cap c_j| = 1$ *if* $|i - j| = 1$ *and* $|c_i \cap c_j| = 0$ *if* $|i - j| \geq 2$, *for all* $1 \leq i, j \leq n-1$,

(b) $g$ *commutes with* $T_{c_i}$ *for all* $1 \leq i \leq n-1$,

(c) $\varphi(s_i) = T_{c_i}^{\varepsilon} g$ *for all* $1 \leq i \leq n-1$.

**Proof** Assume $n$ is even. Let $\partial_1$ and $\partial_2$ be the two boundary components of $\Sigma_n$. We denote by $\hat{\Sigma}_n$ the closed surface obtained from $\Sigma_n$ by gluing a disk $D_1$ along $\partial_1$ and a disk $D_2$ along $\partial_2$. Moreover, we choose a point $\hat{x}_1$ in the interior of $D_1$ and a point $\hat{x}_2$ in the interior of $D_2$, and we set $\hat{\mathcal{P}} = \{x, \hat{x}_1, \hat{x}_2\}$. Assume $n$ is odd. Let $\partial$ be the boundary component of $\Sigma_n$. We denote by $\hat{\Sigma}_n$ the closed surface obtained from $\Sigma_n$ by gluing a disk $D$ along $\partial$. Moreover, we choose a point $\hat{x}$ in the interior of $D$ and we set $\hat{\mathcal{P}} = \{x, \hat{x}\}$. For each $n$ we denote by $\mathcal{PM}(\hat{\Sigma}_n, \hat{\mathcal{P}})$ the subgroup of $\mathcal{M}(\hat{\Sigma}_n, \hat{\mathcal{P}})$ formed by the isotopy classes of elements in $\mathrm{Homeo}^+(\hat{\Sigma}_n, \hat{\mathcal{P}})$ which pointwise fix $\hat{\mathcal{P}}$. The embedding of $\Sigma_n$ into $\hat{\Sigma}_n$ induces a surjective homomorphism $\varpi \colon \mathcal{M}(\Sigma_n, x) \to \mathcal{PM}(\hat{\Sigma}_n, \hat{\mathcal{P}})$. If $n$ is even, then the kernel of $\varpi$ is the free abelian group of rank 2 generated by $T_{\partial_1}$ and $T_{\partial_2}$, and if $n$ is odd, then the kernel of $\varpi$ is the cyclic group generated by $T_{\partial}$. In both cases $\mathrm{Ker}(\varpi)$ is contained in the center of $\mathcal{M}(\Sigma_n, x)$.



Figure 6: The loop $f_{n-i} \in \pi_1(\Sigma_n, x)$.

Let $\varphi\colon A[A_{n-1}] \to \mathcal{M}(\Sigma_n, x)$ be a noncyclic homomorphism. Assume that $\varpi \circ \varphi$ is cyclic. Then there exists $\hat{g} \in \mathcal{PM}(\hat{\Sigma}_n, \hat{\mathcal{P}})$ such that $(\varpi \circ \varphi)(s_i) = \hat{g}$ for all $1 \le i \le n-1$. Let $g \in \mathcal{M}(\Sigma_n, x)$ be such that $\varpi(g) = \hat{g}$. For each $1 \le i \le n-1$ there exists $h_i \in \mathrm{Ker}(\varpi) \subset Z(\mathcal{M}(\Sigma_n, x))$ such that $\varphi(s_i) = g h_i$. Let $1 \le i \le n-2$. Then

$$g^3 h_i^2 h_{i+1} = \varphi(s_i s_{i+1} s_i) = \varphi(s_{i+1} s_i s_{i+1}) = g^3 h_i h_{i+1}^2.$$

Hence $h_i = h_{i+1}$. This shows that $\varphi(s_i) = g h_1$ for all $1 \le i \le n-1$, and hence that $\varphi$ is cyclic, which is a contradiction. So $\varpi \circ \varphi$ is not cyclic.

To differentiate Dehn twists in $\mathcal{M}(\Sigma_n, x)$ from those in $\mathcal{PM}(\hat{\Sigma}_n, \hat{\mathcal{P}})$, for a circle $c$ in $\hat{\Sigma}_n \setminus \hat{\mathcal{P}}$ we denote by $\hat{T}_c$ the Dehn twist in $\mathcal{PM}(\hat{\Sigma}_n, \hat{\mathcal{P}})$ along $c$. By Castel [13, Theorem 1] there exist generic circles $c_1, \ldots, c_{n-1}$ in $\hat{\Sigma}_n \setminus \hat{\mathcal{P}}$, $\varepsilon \in \{\pm 1\}$ and $\hat{g} \in \mathcal{PM}(\hat{\Sigma}_n, \hat{\mathcal{P}})$ such that

(1) $|c_i \cap c_j| = 1$ if $|i - j| = 1$ and $|c_i \cap c_j| = 0$ if $|i - j| \ge 2$, for all $1 \le i, j \le n-1$,

(2) $\hat{g}$ commutes with $\hat{T}_{c_i}$ for all $1 \le i \le n-1$,

(3) $(\varpi \circ \varphi)(s_i) = \hat{T}_{c_i}^{\varepsilon} \hat{g}$ for all $1 \le i \le n-1$.

Clearly, we can choose each $c_i$ siting in the interior of $\Sigma_n$. Let $g \in \mathcal{M}(\Sigma_n, x)$ be such that $\varpi(g) = \hat{g}$. It is easily shown with Castel [13, Lemma 3.2.1] that $g$ and $T_{c_i}$ commute for all $1 \le i \le n-1$. Furthermore, for each $1 \le i \le n-1$, there exists $h_i \in \mathrm{Ker}(\varpi) \subset Z(\mathcal{M}(\Sigma_n, x))$ such that $\varphi(s_i) = T_{c_i}^{\varepsilon} g h_i$. Let $1 \le i \le n-2$. Then

$$T_{c_i}^{\varepsilon} T_{c_{i+1}}^{\varepsilon} T_{c_i}^{\varepsilon} g^3 h_i^2 h_{i+1} = \varphi(s_i s_{i+1} s_i) = \varphi(s_{i+1} s_i s_{i+1}) = T_{c_{i+1}}^{\varepsilon} T_{c_i}^{\varepsilon} T_{c_{i+1}}^{\varepsilon} g^3 h_i h_{i+1}^2$$
$$= T_{c_i}^{\varepsilon} T_{c_{i+1}}^{\varepsilon} T_{c_i}^{\varepsilon} g^3 h_i h_{i+1}^2,$$

and hence $h_{i+1} = h_i$. So there exists $h \in \mathrm{Ker}(\varpi)$ such that $\varphi(s_i) = T_{c_i}^{\varepsilon} g h$ and $gh$ commutes with $T_{c_i}$ for all $1 \le i \le n-1$. $\qquad\square$

# 4  Homomorphisms from $A[D_n]$ to $A[A_{n-1}]$

**Proof of Theorem 2.1**  Let $n \ge 5$. Let $\varphi\colon A[D_n] \to A[A_{n-1}]$ be a homomorphism. By precomposing $\varphi$ with $\iota\colon A[A_{n-1}] \to A[D_n]$, we obtain a homomorphism $\varphi \circ \iota\colon A[A_{n-1}] \to A[D_n] \to A[A_{n-1}]$, and hence, by Theorem 3.5, one of the following two possibilities holds:

- $\varphi \circ \iota$ is cyclic.

- There exist $\psi \in \langle \bar{\chi} \rangle$ and $p \in \mathbb{Z}$ such that $\varphi \circ \iota$ is conjugate to $\psi \circ \bar{\gamma}_p$.

Suppose $\varphi \circ \iota$ is cyclic. Then there exists $u \in A[A_{n-1}]$ such that $(\varphi \circ \iota)(s_i) = \varphi(t_i) = u$ for all $1 \le i \le n-1$. Moreover,

$$\varphi(t_n) = \varphi(t_{n-2} t_n)\varphi(t_{n-2})\varphi(t_n^{-1} t_{n-2}^{-1}) = \varphi(t_{n-2} t_n)\varphi(t_1)\varphi(t_n^{-1} t_{n-2}^{-1}) = \varphi(t_1) = u,$$
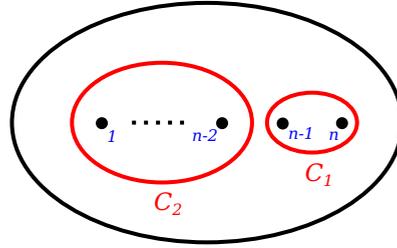
and hence $\varphi$ is cyclic.

Figure 7: Circles in the punctured disk.

So, up to conjugating and replacing $\varphi$ by $\varphi \circ \chi$ if necessary, we can assume that there exists $p \in \mathbb{Z}$ such that $\varphi \circ \iota = \bar{\gamma}_p$. This means that $\varphi(t_i) = (\varphi \circ \iota)(s_i) = s_i \Delta^{2p}$ for all $1 \le i \le n-1$, where $\Delta$ is the Garside element of $A[A_{n-1}]$. Now we turn to showing that $\varphi = \alpha_p$.

Set $Y = \{s_1, \ldots, s_{n-3}\}$. By Paris [37, Theorem 5.1] the centralizer of the group $\langle s_1, \ldots, s_{n-3}, s_{n-1} \rangle$ in $A[A_{n-1}]$ is generated by $\Delta^2$, $\Delta_Y^2$ and $s_{n-1}$, where $\Delta_Y = \Delta_Y[A_{n-1}]$. These three elements pairwise commute and generate a copy of $\mathbb{Z}^3$. Set $u = \varphi(t_n)$. Since $u$ commutes with $\varphi(t_i) = s_i \Delta^{2p}$ for all $i \in \{1, \ldots, n-3, n-1\}$ and $\Delta^2$ is central in $A[A_{n-1}]$, $u$ belongs to the centralizer of $\langle s_1, \ldots, s_{n-3}, s_{n-1} \rangle$, and hence there exist $k_1, k_2, k_3 \in \mathbb{Z}$ such that $u = s_{n-1}^{k_1} \Delta_Y^{2k_2} \Delta^{2k_3}$.

It is well known that $A[A_{n-1}]$ is naturally isomorphic to the mapping class group $\mathcal{M}(\mathbb{D}, \mathcal{P})$, where $\mathbb{D}$ denotes the disk and $\mathcal{P} = \{x_1, \ldots, x_n\}$ is a set of $n$ punctures in the interior of $\mathbb{D}$. In this identification $s_{n-1}^2$ corresponds to the Dehn twist along the circle $c_1$ depicted in Figure 7, $\Delta_Y^2$ corresponds to the Dehn twist along the circle $c_2$ depicted in the same figure and $\Delta^2$ corresponds to the Dehn twist along a circle parallel to $\partial \mathbb{D}$. By Proposition 3.3 we have $\mathcal{S}(u^2) \subseteq \{c_1, c_2\}$, where $c_1 \in \mathcal{S}(u^2)$ if and only if $k_1 \ne 0$ and $c_2 \in \mathcal{S}(u^2)$ if and only if $k_2 \ne 0$. We know that $\varphi(t_1^2) = s_1^2 \Delta^{4p}$, and hence $\mathcal{S}(\varphi(t_1^2))$ is formed by a single circle containing two marked points in its interior. Since $t_1^2$ and $t_n^2$ are conjugate $\varphi(t_1^2)$ and $\varphi(t_n^2) = u^2$ are conjugate, and hence, by Theorem 3.2, $\mathcal{S}(u^2)$ is also formed by a single circle containing two marked points in its interior. It follows that $\mathcal{S}(u^2) = \{c_1\}$, and hence $k_1 \ne 0$ and $k_2 = 0$. It remains to show that $k_1 = 1$ and $k_3 = p$.

From the equality $t_{n-2} t_n t_{n-2} = t_n t_{n-2} t_n$ it follows that $s_{n-2} s_{n-1}^{k_1} s_{n-2} \Delta^{4p+2k_3} = s_{n-1}^{k_1} s_{n-2} s_{n-1}^{k_1} \Delta^{2p+4k_3}$, and hence

$$(s_{n-2} s_{n-1}^{k_1} s_{n-2})(s_{n-1}^{k_1} s_{n-2} s_{n-1}^{k_1})^{-1} = \Delta^{2k_3 - 2p}.$$

We know from Paris [38, Corollary 2.6] that $A_{\{s_{n-2}, s_{n-1}\}}[A_{n-1}] \cap \langle \Delta \rangle = \{1\}$, and hence

$$(s_{n-2} s_{n-1}^{k_1} s_{n-2})(s_{n-1}^{k_1} s_{n-2} s_{n-1}^{k_1})^{-1} = \Delta^{2k_3 - 2p} = 1.$$

Let $z \colon A[A_{n-1}] \to \mathbb{Z}$ be the homomorphism which sends $s_i$ to 1 for all $1 \le i \le n-1$. We have

$$0 = z(1) = z((s_{n-2} s_{n-1}^{k_1} s_{n-2})(s_{n-1}^{k_1} s_{n-2} s_{n-1}^{k_1})^{-1}) = 1 - k_1,$$

and hence $k_1 = 1$. Moreover, $\Delta^{2k_3 - 2p} = 1$ and $\Delta$ is of infinite order; thus $k_3 = p$. $\qquad \square$

# 5 Homomorphisms from $A[A_{n-1}]$ to $A[D_n]$

The formula in the following lemma is a crucial point in various proofs, including those of Lemma 5.4 and Theorem 2.8.

**Lemma 5.1** *Let $n \geq 1$. Then*

$$\Delta[A_n]^2 = (s_1 \cdots s_{n-1} s_n^2 s_{n-1} \cdots s_1)(s_2 \cdots s_{n-1} s_n^2 s_{n-1} \cdots s_2) \cdots (s_{n-1} s_n^2 s_{n-1}) s_n^2.$$

**Proof** We argue by induction on $n$. The case $n = 1$ is trivial, and hence we can assume that $n \geq 2$ and that the induction hypothesis holds. Recall that

$$\Delta[A_n] = (s_1 \cdots s_n)\Delta[A_{n-1}] = \Delta[A_{n-1}](s_n \cdots s_1).$$

Moreover, it is easily checked that $s_i(s_n \cdots s_1) = (s_n \cdots s_1)s_{i+1}$ for all $1 \leq i \leq n-1$. By the induction hypothesis,

$$\Delta[A_{n-1}]^2 = (s_1 \cdots s_{n-2} s_{n-1}^2 s_{n-2} \cdots s_1) \cdots (s_{n-2} s_{n-1}^2 s_{n-2}) s_{n-1}^2.$$

Hence

$$
\begin{aligned}
\Delta[A_n]^2 &= (s_1 \cdots s_n)\Delta[A_{n-1}]^2(s_n \cdots s_1) \\
&= (s_1 \cdots s_n)\big((s_1 \cdots s_{n-2} s_{n-1}^2 s_{n-2} \cdots s_1) \cdots (s_{n-2} s_{n-1}^2 s_{n-2}) s_{n-1}^2\big)(s_n \cdots s_1) \\
&= (s_1 \cdots s_n)(s_n \cdots s_1)\big((s_2 \cdots s_{n-1} s_n^2 s_{n-1} \cdots s_2) \cdots (s_{n-1} s_n^2 s_{n-1}) s_n^2\big) \\
&= (s_1 \cdots s_{n-1} s_n^2 s_{n-1} \cdots s_1) \cdots (s_{n-1} s_n^2 s_{n-1}) s_n^2. \qquad \square
\end{aligned}
$$

Now, Lemmas 5.2–5.8 are preliminaries to the proof of Theorem 2.2.

**Lemma 5.2** *Let $n \geq 6$. Let $\varphi \colon A[A_{n-1}] \to A[D_n]$ be a homomorphism. If $\pi \circ \varphi \colon A[A_{n-1}] \to A[A_{n-1}]$ is cyclic, then $\varphi$ is cyclic.*

**Proof** Assume $\pi \circ \varphi$ is cyclic. Then there exists $u \in A[A_{n-1}]$ such that $(\pi \circ \varphi)(s_i) = u$ for all $1 \leq i \leq n-1$. For $3 \leq i \leq n-1$ we set $v_i = \varphi(s_i s_1^{-1})$. We have $\pi(v_i) = uu^{-1} = 1$, and hence $v_i \in \mathrm{Ker}(\pi)$. We have

$$(s_3 s_1^{-1})(s_4 s_1^{-1})(s_3 s_1^{-1}) = s_3 s_4 s_3 s_1^{-3} = s_4 s_3 s_4 s_1^{-3} = (s_4 s_1^{-1})(s_3 s_1^{-1})(s_4 s_1^{-1}),$$

and hence $v_3 v_4 v_3 = v_4 v_3 v_4$. Since $\mathrm{Ker}(\pi)$ is a free group (see Crisp and Paris [21, Proposition 2.3]) and two elements in a free group either freely generate a free group or commute, the existence of such equality implies that $v_3 v_4 = v_4 v_3$. It follows that $v_3 v_4 v_3 = v_3 v_4^2$; hence $v_3 = v_4$, and therefore

$$\varphi(s_3)\varphi(s_1)^{-1} = v_3 = v_4 = \varphi(s_4)\varphi(s_1)^{-1}.$$

So $\varphi(s_3) = \varphi(s_4)$. We conclude by Castel [13, Lemma 3.1.1] that $\varphi$ is cyclic. $\qquad \square$

Let $n \geq 6$. If $n$ is odd then $\Sigma_n$ has one boundary component, which we denote by $\partial$, and we denote by $T_\partial$ the Dehn twist along $\partial$. If $n$ is even then $\Sigma_n$ has two boundary components, which we denote by $\partial_1$ and $\partial_2$, and we denote by $T_{\partial_1}$ and $T_{\partial_2}$ the Dehn twists along $\partial_1$ and $\partial_2$, respectively. It is known that the center of $\mathcal{M}(\Sigma_n)$, denoted by $Z(\mathcal{M}(\Sigma_n))$, is the cyclic group generated by $T_\partial$ if $n$ is odd, and it is a free abelian group of rank 2 generated by $T_{\partial_1}$ and $T_{\partial_2}$ if $n$ is even (see Paris and Rolfsen [39, Theorem 5.6], for example).

**Lemma 5.3** *Let $n \geq 2$. Let $f \in \mathcal{M}(\Sigma_n)$ such that $f T_{a_i}^2 = T_{a_i}^2 f$ for all $1 \leq i \leq n-1$. Then $f^2 \in Z(\mathcal{M}(\Sigma_n))$.*

**Proof** Assume $n$ is odd. The case where $n$ is even can be proved in the same way. Let $f \in \mathcal{M}(\Sigma_n)$ such that $f T_{a_i}^2 = T_{a_i}^2 f$ for all $1 \leq i \leq n-1$. Since $f T_{a_i}^2 f^{-1} = T_{a_i}^2$ we have $f([a_i]) = [a_i]$ (see Farb and Margalit [28, Section 3.3]). The mapping class $f$ may reverse the orientation of each $a_i$ up to isotopy, but $f^2$ preserves the orientation of all $a_i$ up to isotopy, and hence $f^2$ can be represented by an element of $\mathrm{Homeo}^+(\Sigma_n)$ which is the identity on a (closed) regular neighborhood $\Sigma'$ of $\bigcup_{i=1}^{n-1} a_i$. We observe that $\Sigma'$ is a surface of genus $\frac{1}{2}(n-1)$ with one boundary component, $\partial'$, and that $\partial \cup \partial'$ bounds a cylinder $C$. This implies that $f^2 \in \mathcal{M}(C) \subset \mathcal{M}(\Sigma_n)$. Since $\mathcal{M}(C) = \langle T_\partial \rangle = Z(\mathcal{M}(\Sigma_n))$, we conclude that $f^2 \in Z(\mathcal{M}(\Sigma_n))$. $\square$

**Lemma 5.4** *Let $n \geq 3$. We set $m = n-1$ if $n$ is odd and $m = n-2$ if $n$ is even. Let $1 \leq k \leq m$. Let $c$ be a generic circle of $\Sigma_n \setminus \{x\}$ such that $c \cap d_i = \varnothing$ for $1 \leq i \leq k-2$, $|c \cap d_{k-1}| = 1$ if $k \geq 2$, $c \cap d_k = \varnothing$ and $c$ is isotopic to $d_k$ in $\Sigma_n$. Then there exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq k-1$ and $g([c]) = [d_k]$.*

**Proof** We identify $D_3$ with $A_3$ in this proof to treat the cases $k = 2$ and $k = 1$. We first assume that $k$ is even. If $c$ is isotopic in $\Sigma_n \setminus \{x\}$ to $d_k$, then it suffices to take $g = \mathrm{id}$. So we can assume that $c$ and $d_k$ are not isotopic in $\Sigma_n \setminus \{x\}$. Since $c$ and $d_k$ are isotopic in $\Sigma_n$, by Epstein [27, Lemma 2.4] there exists a cylinder $C$ in $\Sigma_n$ whose boundary components are $d_k$ and $c$. Since $c$ and $d_k$ are not isotopic in $\Sigma_n \setminus \{x\}$, this cylinder must contain the puncture $x$.

Let $\Sigma'$ be a regular neighborhood of $\left( \bigcup_{i=1}^{k-1} d_i \right) \cup C$. The surface $\Sigma'$ contains the cylinder $C$ with boundaries $c$ and $d_k$, having the puncture $x$ in it, and $d_{k-1}$ intersects $c$ and $d_k$ once. Hence an arc of the curve $d_{k-1}$ connects a point on $c$ with a point on $d_k$ within the cylinder $C$, and it may wind around the cylinder in different ways (see Figure 8). However, by applying suitable Dehn twists about $c$ and $d_k$, one can unwind this arc to the simplest case, shown in Figure 9. Hence, up to homeomorphism of the surface $\Sigma_n$, we may assume that the circles $d_1, \ldots, d_k, c$ are arranged as in Figure 9.

By Proposition 3.1 there are homomorphisms $\psi_1 \colon A[D_{k+1}] \to \mathcal{M}(\Sigma_n, x)$ and $\psi_2 \colon A[A_k] \to \mathcal{M}(\Sigma_n, x)$ defined by

$$\psi_1(t_i) = T_{d_i} \quad \text{for } 1 \leq i \leq k, \qquad \psi_1(t_{k+1}) = T_c,$$
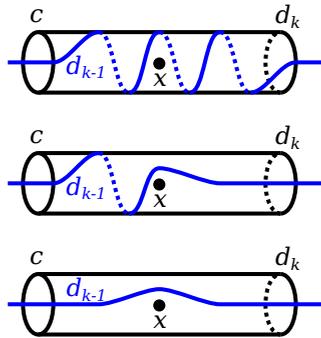$$\psi_2(s_i) = T_{d_i} \quad \text{for } 1 \leq i \leq k-1, \qquad \psi_2(s_k) = T_c.$$

Figure 8: The intersection of $C$ with $d_{k-1}$.

We denote by $\Delta_{D,k}$ the Garside element of $A[D_{k+1}]$ and by $\Delta_{A,k}$ the Garside element of $A[A_k]$, and we set $g = \psi_1(\Delta_{D,k})\psi_2(\Delta_{A,k}^{-2})$. We have $\Delta_{D,k} t_i \Delta_{D,k}^{-1} = t_i$ for all $1 \le i \le k-1$, $\Delta_{D,k} t_{k+1} \Delta_{D,k}^{-1} = t_k$ and $\Delta_{A,k}^2 s_i \Delta_{A,k}^{-2} = s_i$ for all $1 \le i \le k$. Hence $g T_{d_i} g^{-1} = T_{g(d_i)} = T_{d_i}$ for all $1 \le i \le k-1$ and $g T_c g^{-1} = T_{g(c)} = T_{d_k}$. It follows that $g([d_i]) = [d_i]$ for all $1 \le i \le k-1$ and $g([c]) = [d_k]$ (see Farb and Margalit [28, Fact 3.6]).

Since $c$ and $d_k$ are isotopic in $\Sigma_n$, the corresponding Dehn twists $T_c$ and $T_{d_k}$ are equal in $\mathcal{M}(\Sigma_n)$, and hence for $T_c$ and $T_{d_k}$, viewed on the surface $\Sigma_n \setminus \{x\}$, we have $\theta(T_c) = \theta(T_{d_k})$. Moreover,

$$\Delta_{D,k} = (t_1 \cdots t_{k-1} t_k t_{k+1} t_{k-1} \cdots t_1) \cdots (t_{k-1} t_k t_{k+1} t_{k-1})(t_k t_{k+1}),$$
$$\Delta_{A,k}^2 = (s_1 \cdots s_{k-1} s_k^2 s_{k-1} \cdots s_1) \cdots (s_{k-1} s_k^2 s_{k-1}) s_k^2,$$

(see Lemma 5.1 for the second equality); hence $\theta(\psi_1(\Delta_{D,k})) = \theta(\psi_2(\Delta_{A,k}^2))$, and therefore $\theta(g) = 1$. So $g \in \operatorname{Ker}(\theta)$.

Now assume $k$ is odd. If $c$ is isotopic in $\Sigma_n \setminus \{x\}$ to $d_k$, then we can take $g = \mathrm{id}$. So we can assume that $c$ and $d_k$ are not isotopic in $\Sigma_n \setminus \{x\}$. Since $c$ and $d_k$ are isotopic in $\Sigma_n$, there exists a cylinder $C$ in $\Sigma_n$ whose boundary components are $d_k$ and $c$. Since $c$ and $d_k$ are not isotopic in $\Sigma_n \setminus \{x\}$, this cylinder must contain the puncture $x$. Let $\Sigma'$ be a closed regular neighborhood of $\left(\bigcup_{i=1}^{k-1} d_i\right) \cup C$. Then $\Sigma'$ is a surface of genus $\frac{1}{2}(k-1)$ with two boundary components and the circles $d_1, \ldots, d_{k-1}, d_k, c$ are arranged as shown in Figure 10. Since $k \le m$ and $k$ is odd, $\frac{1}{2}(k-1)$ is strictly less than the genus of $\Sigma_n$; hence we can choose a subsurface $\Sigma''$ of $\Sigma_n$ of genus $\frac{1}{2}(k+1)$, with one boundary component, and containing $\Sigma'$. We can also choose a generic circle $e$ in $\Sigma'' \setminus \{x\}$ such that $|e \cap d_1| = 1$, $|e \cap c| = 1$ if $k = 1$, $e \cap d_i = \varnothing$ for all $2 \le i \le k$ and $e \cap c = \varnothing$ if $k \ge 2$ (see Figure 10). By Proposition 3.1 there are homomorphisms
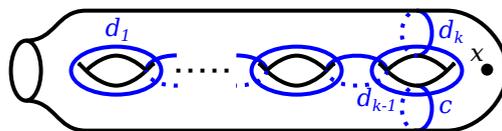


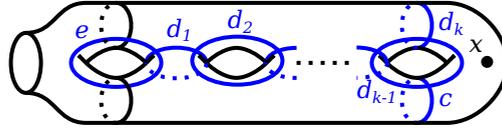Figure 9: The regular neighborhood of $\left(\bigcup_{i=1}^{k-1} d_i\right) \cup C$ when $k$ is even.

Figure 10: The regular neighborhood of $\left(\bigcup_{i=1}^{k-1} d_i\right) \cup C$ when $k$ is odd

$\psi_1 \colon A[D_{k+2}] \to \mathcal{M}(\Sigma_n, x)$ and $\psi_2 \colon A[A_{k+1}] \to \mathcal{M}(\Sigma_n, x)$ defined by

$$\psi_1(t_1) = T_e, \qquad \psi_1(t_i) = T_{d_{i-1}} \quad \text{for } 2 \le i \le k+1, \qquad \psi_1(t_{k+2}) = T_c,$$
$$\psi_2(s_1) = T_e, \qquad \psi_2(s_i) = T_{d_{i-1}} \quad \text{for } 2 \le i \le k, \qquad \psi_2(s_{k+1}) = T_c.$$

We denote by $\Delta_{D,k+1}$ the Garside element of $A[D_{k+2}]$ and by $\Delta_{A,k+1}$ the Garside element of $A[A_{k+1}]$, and we set $g = \psi_1(\Delta_{D,k+1})\psi_2(\Delta_{A,k+1}^{-2})$. Then, as in the case where $k$ is even, we have $g([d_i]) = [d_i]$ for all $1 \le i \le k-1$, $g([c]) = [d_k]$ and $g \in \mathrm{Ker}(\theta)$. $\qquad\square$

The following lemma is the extension of Lemma 5.4 to the case $c \cap d_k \ne \varnothing$.

**Lemma 5.5** *Let $n \ge 3$. Set $m = n-1$ if $n$ is odd and $m = n-2$ if $n$ is even. Let $1 \le k \le m$. Let $c$ be a generic circle of $\Sigma_n \setminus \{x\}$ such that $c \cap d_i = \varnothing$ for $1 \le i \le k-2$, $|c \cap d_{k-1}| = 1$ if $k \ge 2$, and $c$ is isotopic to $d_k$ in $\Sigma_n$. Then there exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \le i \le k-1$ and $g([c]) = [d_k]$.*

**Proof** We argue by induction on $i([c], [d_k])$, which is computed on the surface $\Sigma_n \setminus \{x\}$ and not on $\Sigma_n$. The case $i([c], [d_k]) = 0$ is proved in Lemma 5.4, and hence we can assume that $i([c], [d_k]) \ge 1$ and that the induction hypothesis holds. Note that now $c$ and $d_k$ cannot be isotopic in $\Sigma_n \setminus \{x\}$ since $i([c], [d_k]) \ne 0$. We can assume without loss of generality that $i([c], [d_k]) = |c \cap d_k|$. Since $c$ and $d_k$ are isotopic in $\Sigma_n$, there exists a bigon $D$ in $\Sigma_n$ cobounded by an arc of $d_k$ and an arc of $c$ as shown in Figure 11. We can choose this bigon to be minimal in the sense that its interior intersects neither $c$ nor $d_k$. The bigon $D$ cannot intersect $d_i$ for $1 \le i \le k-2$ and one can easily modify $c$ so that $D$ does not intersect $d_{k-1}$ either. Since $c$ and $d_k$ are not isotopic in $\Sigma_n \setminus \{x\}$, $D$ necessarily contains the puncture $x$ in its interior. We choose a circle $c'$ parallel to $c$ except in the bigon $D$, where it follows the arc of $d_k$ which borders $D$ as illustrated in Figure 11. By construction $c' \cap d_i = \varnothing$ for $1 \le i \le k-2$, $|c' \cap d_{k-1}| = 1$ if $k \ge 2$, and $c'$ is isotopic to $d_k$ in $\Sigma_n$. Moreover $i([c'], [d_k]) \le |c' \cap d_k| < |c \cap d_k| = i([c], [d_k])$. By the induction hypothesis there exists $g_1 \in \mathrm{Ker}(\theta)$ such that $g_1([d_i]) = [d_i]$ for all $1 \le i \le k-1$ and $g_1([c']) = [d_k]$.
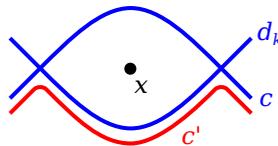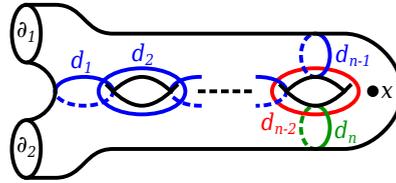


Figure 11: The bigon cobounded by $c$ and $d_k$.

Figure 12: The circles $d_1, \ldots, d_n$.

By Farb and Margalit [28, Lemma 2.9], we can choose $G_1 \in \mathrm{Homeo}^+(\Sigma_n, x)$ which represents $g_1$ such that $G_1(d_i) = d_i$ for all $1 \le i \le k-1$ and $G_1(c') = d_k$. We set $c'' = G_1(c)$. Then $c'' \cap d_i = \varnothing$ for $1 \le i \le k-2$, $|c'' \cap d_{k-1}| = 1$ if $k \ge 2$, $c'' \cap d_k = \varnothing$ and $c''$ is isotopic to $d_k$ in $\Sigma_n$. By Lemma 5.4 there exists $g_2 \in \mathrm{Ker}(\theta)$ such that $g_2([d_i]) = [d_i]$ for all $1 \le i \le k-1$ and $g_2([c'']) = [d_k]$. We set $g = g_2 \circ g_1$. Then $g \in \mathrm{Ker}(\theta)$, $g([d_i]) = [d_i]$ for all $1 \le i \le k-1$ and $g([c]) = [d_k]$. $\qquad\square$

**Lemma 5.6** *Let* $n \ge 4$ *be even. Let* $c$ *be a generic circle of* $\Sigma_n \setminus \{x\}$ *such that* $c \cap d_i = \varnothing$ *for all* $1 \le i \le n-3$, $|c \cap d_{n-2}| = 1$, $c \cap d_{n-1} = \varnothing$ *and* $c$ *is isotopic to* $d_{n-1}$ *in* $\Sigma_n$. *Then we have one of the following two possibilities:*

(1) *$c$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.*

(2) *There exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \le i \le n-1$ and $g([c]) = [d_n]$.*

**Proof** The surface $\Sigma_n$ is a surface of genus $\frac{1}{2}(n-2)$ with two boundary components $\partial_1$ and $\partial_2$. We assume that the circles $d_1, \ldots, d_{n-1}, d_n$ are arranged as in Figure 12. Let $\Omega$ be the surface obtained by cutting $\Sigma_n$ along $\bigcup_{i=1}^{n-1} d_i$. Then $\Omega$ has two connected components $\Omega_1$ and $\Omega_2$. Each of these components is a cylinder that we represent by a square with a hole in the middle, as shown in Figure 13. Two opposite sides of each square represent arcs of $d_{n-2}$, one side represents an arc of $d_{n-1}$ and the last side represents a union of arcs of $d_1, \ldots, d_{n-3}$. The boundary of the hole represents $\partial_1$ for $\Omega_1$ and $\partial_2$ for $\Omega_2$. The puncture $x$ sits inside $\Omega_2$. The trace of the circle $c$ in $\Omega$ is a simple arc $\ell$, either in $\Omega_1$ or in $\Omega_2$.

Suppose $\ell$ is in $\Omega_1$. Let $q$ be the intersection point of $c$ with $d_{n-2}$. Then $q$ is represented in $\Omega_1$ by two points $q_1$ and $q_2$ on two opposite sides of $\Omega_1$, as shown in Figure 13, and $\ell$ is a simple arc connecting $q_1$ with $q_2$. Up to isotopy pointwise fixing the boundary of $\Omega_1$, there exist exactly two simple arcs in $\Omega_1$



Figure 13: The surface $\Omega$ with components $\Omega_1$ (left) and $\Omega_2$ (right).
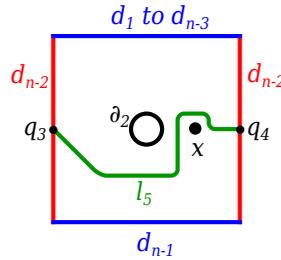
Figure 14: The arc $\ell_5$.

connecting $q_1$ to $q_2$ that are represented by the arcs $\ell_1$ and $\ell_2$ depicted in Figure 13. The arc $\ell$ cannot be isotopic to $\ell_1$, otherwise $c$ would not be isotopic to $d_{n-1}$ in $\Sigma_n$. So $\ell$ is isotopic to $\ell_2$ in $\Omega_1$, which implies that $c$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

Now suppose $\ell$ is in $\Omega_2$. Let $q$ be the intersection point of $c$ with $d_{n-2}$. Then $q$ is represented in $\Omega_2$ by two points $q_3$ and $q_4$ on two opposite sides of $\Omega_2$, as shown in Figure 13, and $\ell$ is a simple arc connecting $q_3$ with $q_4$. Up to isotopy (in $\Omega_2$ and not in $\Omega_2 \setminus \{x\}$) pointwise fixing the boundary of $\Omega_2$, there exist exactly two simple arcs in $\Omega_2$ connecting $q_3$ to $q_4$ that are represented by the arcs $\ell_3$ and $\ell_4$ depicted in Figure 13. The arc $\ell$ cannot be isotopic to $\ell_3$ in $\Omega_2$, otherwise $c$ would not be isotopic to $d_{n-1}$ in $\Sigma_n$. So $\ell$ is isotopic to $\ell_4$ in $\Omega_2$. Let $\{F_t \colon \Omega_2 \to \Omega_2\}_{t \in [0,1]}$ be an isotopy such that $F_0 = \mathrm{id}$, $F_1(\ell) = \ell_4$ and $F_t$ is the identity on the boundary of $\Omega_2$ for all $t \in [0,1]$. The arc $\ell_4$ divides $\Omega_2$ into two parts: the lower one, which does not contain the hole bordered by $\partial_2$ and the puncture $x$, and the upper one, which contains the hole bordered by $\partial_2$ and the puncture $x$, as shown in Figure 13.

Suppose $F_1(x)$ is in the upper part. Let $C$ be the domain of $\Omega_2$ bounded by $\ell_4$, two arcs of $d_{n-2}$ and an arc of $d_{n-1}$, as shown in Figure 13. Let $C' = F_1^{-1}(C)$. Then $C'$ is a domain of $\Omega_2$ bounded by $\ell$, two arcs of $d_{n-2}$ and an arc of $d_{n-1}$, and $C'$ does not contain the puncture $x$. The existence of such a domain implies that $c$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

Now suppose $F_1(x)$ is in the lower part. We can assume without loss of generality that the trace of $d_n$ on $\Omega_2$ is the simple arc $\ell_5$ drawn in Figure 14. We can choose an isotopy $\{F'_t \colon \Omega_2 \to \Omega_2\}_{t \in [0,1]}$ such that $F'_0 = \mathrm{id}$, $F'_1(\ell_4) = \ell_5$, $F'_t$ is the identity on the boundary of $\Omega_2$ for all $t \in [0,1]$, and $F'_1(F_1(x)) = x$. Let $\widetilde{F} \colon \Sigma_n \to \Sigma_n$ be the homeomorphism which is $F'_1 \circ F_1$ on $\Omega_2$ and is the identity outside $\Omega_2$, and let $g \in \mathcal{M}(\Sigma_n, x)$ be the mapping class represented by $\widetilde{F}$. Then $g \in \mathrm{Ker}(\theta)$, $g([d_i]) = [d_i]$ for all $1 \le i \le n-1$, and $g([c]) = [d_n]$. $\square$

**Remark** The element $g$ at the end of the proof of Lemma 5.6 is not necessarily trivial. For example, $\ell$ can be as shown in Figure 15 up to isotopy and, in this case, $g$ must be nontrivial. In fact, $g$ can be any element of the fundamental group $\pi_1(\Omega_2, x)$, which is an infinite cyclic group, seen as a subgroup of $\mathcal{M}(\Sigma_n, x)$.

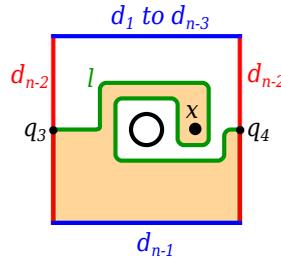The following lemma is the extension of Lemma 5.6 to the case $c \cap d_k \ne \varnothing$.

Figure 15: An arc $\ell$ nonisotopic to $\ell_5$.

**Lemma 5.7** *Let $n \geq 4$ be even. Let $c$ be a generic circle of $\Sigma_n \setminus \{x\}$ such that $c \cap d_i = \varnothing$ for all $1 \leq i \leq n-3$, $|c \cap d_{n-2}| = 1$ and $c$ is isotopic to $d_{n-1}$ in $\Sigma_n$. Then there exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, and either $g([c]) = [d_{n-1}]$ or $g([c]) = [d_n]$.*

**Proof** In this proof the intersection number of two circles is computed on the surface $\Sigma_n \setminus \{x\}$ and not on $\Sigma_n$. We can assume that $|c \cap d_{n-1}| = i([c], [d_{n-1}])$ and $|c \cap d_n| = i([c], [d_n])$. We argue by induction on $|c \cap d_{n-1}| + |c \cap d_n| = i([c], [d_{n-1}]) + i([c], [d_n])$. The case $|c \cap d_{n-1}| = 0$ follows directly from Lemma 5.6, and the case $|c \cap d_n| = 0$ is proved in the same way by replacing $d_{n-1}$ with $d_n$. So we can assume that $i([c], [d_{n-1}]) = |c \cap d_{n-1}| \geq 1$, $i([c], [d_n]) = |c \cap d_n| \geq 1$ and that the induction hypothesis holds. Note that now $c$ and $d_{n-1}$ cannot be isotopic in $\Sigma_n \setminus \{x\}$. Since $c$ and $d_{n-1}$ are isotopic in $\Sigma_n$, there exists a bigon $D$ in $\Sigma_n$ cobounded by an arc of $d_{n-1}$ and an arc of $c$ (see Figure 16). Since $c$ and $d_{n-1}$ are not isotopic in $\Sigma_n \setminus \{x\}$, this bigon necessarily contains the puncture $x$. We can choose $D$ to be minimal in the sense that its interior does not intersect $c$ and $d_{n-1}$. Moreover, up to exchanging the roles of $d_{n-1}$ and $d_n$ if necessary, we can also assume that $d_n$ does not intersect the interior of $D$. Clearly $D$ does not intersect $d_i$ for any $1 \leq i \leq n-3$ and, up to replacing $c$ with an isotopic circle, we can assume that $D$ does not intersect $d_{n-2}$ either. Let $c'$ be a circle parallel to $c$ except in the bigon $D$, where it follows the arc of $d_{n-1}$, which borders $D$ as illustrated in Figure 16. We have $c' \cap d_i = \varnothing$ for all $1 \leq i \leq n-3$, $|c' \cap d_{n-2}| = 1$ and $c'$ is isotopic to $d_{n-1}$ in $\Sigma_n$. We also have $i([c'], [d_{n-1}]) < i([c], [d_{n-1}])$ and $i([c'], [d_n]) \leq i([c], [d_n])$; hence by the induction hypothesis there exists $g_1 \in \mathrm{Ker}(\theta)$ such that $g_1([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, and either $g_1([c']) = [d_{n-1}]$ or $g_1([c']) = [d_n]$. Without loss of generality we can assume that $g_1([c']) = [d_{n-1}]$. We choose $G_1 \in \mathrm{Homeo}^+(\Sigma_n, x)$ which represents $g_1$ such that $G_1(d_i) = d_i$ for all $1 \leq i \leq n-2$ and $G_1(c') = d_{n-1}$. We set $c'' = G_1(c)$. Then $c'' \cap d_i = \varnothing$ for all $1 \leq i \leq n-3$, $|c'' \cap d_{n-2}| = 1$, $c'' \cap d_{n-1} = \varnothing$ and $c''$ is isotopic to $d_{n-1}$ in $\Sigma_n$. By Lemma 5.6 there exists $g_2 \in \mathrm{Ker}(\theta)$ such that $g_2([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, and either $g_2([c'']) = [d_{n-1}]$ or $g_2([c'']) = [d_n]$. We set $g = g_2 \circ g_1$. Then $g \in \mathrm{Ker}(\theta)$, $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, and either $g([c]) = [d_{n-1}]$ or $g([c]) = [d_n]$. $\qquad\square$

**Lemma 5.8** *Let $n \geq 6$. Let $c_1, \ldots, c_{n-1}$ be generic circles in $\Sigma_n \setminus \{x\}$ such that*

(a) $|c_i \cap c_j| = 1$ if $|i - j| = 1$ and $|c_i \cap c_j| = 0$ if $|i - j| \geq 2$, for all $1 \leq i, j \leq n-1$,

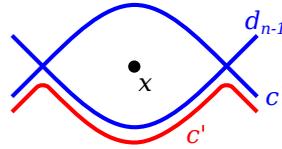(b) $c_i$ is isotopic to $d_i$ in $\Sigma_n$ for all $1 \leq i \leq n-1$.

Figure 16: The bigon cobounded by $c$ and $d_{n-1}$.

*Then*:

(1)  *If $n$ is odd, then there exists $g \in \mathrm{Ker}(\theta)$ such that $g([c_i]) = [d_i]$ for all $1 \le i \le n-1$.*

(2)  *If $n$ is even, then there exists $g \in \mathrm{Ker}(\theta)$ such that $g([c_i]) = [d_i]$ for all $1 \le i \le n-2$, and either $g([c_{n-1}]) = [d_{n-1}]$ or $g([c_{n-1}]) = [d_n]$.*

**Proof**  For $1 \le k \le n-2$ we construct by induction on $k$ an element $g_k \in \mathrm{Ker}(\theta)$ such that $g_k([c_i]) = [d_i]$ for all $1 \le i \le k$. Assume $k = 1$. Then, by Lemma 5.5 applied to $k = 1$, there exists $g_1 \in \mathrm{Ker}(\theta)$ such that $g_1([c_1]) = [d_1]$. Suppose $2 \le k \le n-1$ and $g_{k-1}$ is constructed. We choose $G_{k-1} \in \mathrm{Homeo}^+(\Sigma_n, x)$ which represents $g_{k-1}$ such that $G_{k-1}(c_i) = d_i$ for all $1 \le i \le k-1$, and we set $c_k' = G_{k-1}(c_k)$. Note that, since $g_{k-1} \in \mathrm{Ker}(\theta)$, the circle $c_k'$ is isotopic to $c_k$ in $\Sigma_n$. Then, by Lemma 5.5, there exists $h_k \in \mathrm{Ker}(\theta)$ such that $h_k([d_i]) = [d_i]$ for all $1 \le i \le k-1$ and $h_k([c_k']) = [d_k]$. We set $g_k = h_k \circ g_{k-1}$. Then $g_k([c_i]) = [d_i]$ for all $1 \le i \le k$. Note that when $n$ is odd we can extend the induction to $k = n-1$ and conclude the proof here by setting $g = g_{n-1}$. The case where $n$ is even requires an extra argument.

Assume $n$ is even. We choose $G_{n-2} \in \mathrm{Homeo}^+(\Sigma_n, x)$ which represents $g_{n-2}$ and such that $G_{n-2}(c_i) = d_i$ for all $1 \le i \le n-2$, and we set $c_{n-1}' = G_{n-2}(c_{n-1})$. Again, since $g_{n-2} \in \mathrm{Ker}(\theta)$, the circle $c_{n-1}'$ is isotopic to $c_{n-1}$ in $\Sigma_n$. By Lemma 5.7 there exists $h_{n-1} \in \mathrm{Ker}(\theta)$ such that $h_{n-1}([d_i]) = [d_i]$ for all $1 \le i \le n-2$, and either $h_{n-1}([c_{n-1}']) = [d_{n-1}]$ or $h_{n-1}([c_{n-1}']) = [d_n]$. We set $g = h_{n-1} \circ g_{n-2}$. Then $g([c_i]) = [d_i]$ for all $1 \le i \le n-2$, and either $g([c_{n-1}]) = [d_{n-1}]$ or $g([c_{n-1}]) = [d_n]$.  □

**Proof of Theorem 2.2**  Let $n \ge 6$ and let $\varphi \colon A[A_{n-1}] \to A[D_n]$ be a homomorphism. Composing $\varphi$ with $\pi$, we get a homomorphism $\pi \circ \varphi \colon A[A_{n-1}] \to A[D_n] \to A[A_{n-1}]$. We know by Theorem 3.5 that we have one of the following possibilities:

- $\pi \circ \varphi$ is cyclic.

- There exist $\psi \in \langle \bar{\chi} \rangle$ and $p \in \mathbb{Z}$ such that $\pi \circ \varphi$ is conjugate to $\psi \circ \bar{\gamma}_p$.

By Lemma 5.2, if $\pi \circ \varphi$ is cyclic, then $\varphi$ is cyclic. So we can assume that there exist $\psi \in \langle \bar{\chi} \rangle$ and $p \in \mathbb{Z}$ such that $\pi \circ \varphi$ is conjugate to $\psi \circ \bar{\gamma}_p$. Up to conjugating and composing $\varphi$ on the left by $\chi$ if necessary, we can assume that $\pi \circ \varphi = \bar{\gamma}_p$, that is, $(\pi \circ \varphi)(s_i) = s_i \Delta_A^{2p}$, where $\Delta_A$ denotes the Garside element of $A[A_{n-1}]$.

Set $U = \rho_A(\Delta_A^2)$. If $n$ is odd, then $U^2 = T_\partial$, where $\partial$ is the boundary component of $\Sigma_n$, and if $n$ is even, then $U = T_{\partial_1} T_{\partial_2}$, where $\partial_1$ and $\partial_2$ are the two boundary components of $\Sigma_n$ (see Labruère and Paris [32, Proposition 2.12]). In particular $U^2 \in Z(\mathcal{M}(\Sigma_n))$ in both cases.

By Theorem 3.7 there exist generic circles $c_1, \ldots, c_{n-1}$ in $\Sigma_n \setminus \{x\}$, $\varepsilon \in \{\pm 1\}$ and $f_0 \in \mathcal{M}(\Sigma_n, x)$ such that

(a)  $|c_i \cap c_j| = 1$ if $|i - j| = 1$ and $|c_i \cap c_j| = 0$ if $|i - j| \geq 2$, for all $1 \leq i, j \leq n-1$,

(b)  $f_0$ commutes with $T_{c_i}$ for all $1 \leq i \leq n-1$,

(c)  $(\rho_D \circ \varphi)(s_i) = T_{c_i}^\varepsilon f_0$ for all $1 \leq i \leq n-1$.

For $1 \leq i \leq n-1$ we denote by $b_i$ the circle in $\Sigma_n$ obtained by composing $c_i : \mathbb{S}^1 \to \Sigma_n \setminus \{x\}$ with the embedding $\Sigma_n \setminus \{x\} \hookrightarrow \Sigma_n$. In addition we set $g_0 = \theta(f_0)$. Then $(\theta \circ \rho_D \circ \varphi)(s_i) = T_{b_i}^\varepsilon g_0$ for all $1 \leq i \leq n-1$. Note that, since $\theta \circ \rho_D = \rho_A \circ \pi$ (see (3-1)), we also have $(\theta \circ \rho_D \circ \varphi)(s_i) = (\rho_A \circ \bar{\gamma}_p)(s_i) = \rho_A(s_i \Delta_A^{2p}) = T_{a_i} U^p$ for all $1 \leq i \leq n-1$, where the $a_i$ are the circles depicted in Figure 3.

**Claim**  *We have $\varepsilon = 1$, $g_0 = U^p$ and $b_i$ is isotopic to $a_i$ in $\Sigma_n$ for all $1 \leq i \leq n-1$.*

**Proof of the claim**  Note that $g_0 = \theta(f_0)$ commutes with $T_{b_i} = \theta(T_{c_i})$ and $U = \rho_A(\Delta_A^2)$ commutes with $T_{a_i} = \rho_A(s_i)$; hence $T_{b_i}^{2\varepsilon} g_0^2 = (T_{b_i}^\varepsilon g_0)^2 = (T_{a_i} U^p)^2 = T_{a_i}^2 U^{2p}$. Since $g_0^2$ commutes with $T_{b_i}^{2\varepsilon} g_0^2 = T_{a_i}^2 U^{2p}$ and $U^2 \in Z(\mathcal{M}(\Sigma_n))$, $g_0^2$ commutes with $T_{a_i}^2$ for all $1 \leq i \leq n-1$. By Lemma 5.3 it follows that $g_0^4 \in Z(\mathcal{M}(\Sigma_n))$. By Proposition 3.3 applied to $\mathcal{M}(\Sigma_n)$ we deduce that $\mathcal{S}(T_{a_i}^4 U^{4p}) = \mathcal{S}(T_{b_i}^{4\varepsilon} g_0^4) = \{[a_i]\} = \{[b_i]\}$, and hence $[a_i] = [b_i]$ for all $1 \leq i \leq n-1$. Then $T_{a_i}^{4-4\varepsilon} = U^{-4p} g_0^4$; hence, by Proposition 3.3, $4 - 4\varepsilon = 0$, and therefore $\varepsilon = 1$. Finally, from the equality $T_{a_i} U^p = T_{a_i} g_0$ it follows that $g_0 = U^p$. $\square$

From the claim it follows that $c_i$ is isotopic to $d_i$ in $\Sigma_n$. Hence, by Lemma 5.8, there exists $g \in \mathrm{Ker}(\theta)$ such that $g([c_i]) = [d_i]$ for all $1 \leq i \leq n-2$, $g([c_{n-1}]) = [d_{n-1}]$ if $n$ is odd, and either $g([c_{n-1}]) = [d_{n-1}]$ or $g([c_{n-1}]) = [d_n]$ if $n$ is even. These equalities imply that $g T_{c_i} g^{-1} = T_{d_i}$ for $1 \leq i \leq n-2$, $g T_{c_{n-1}} g^{-1} = T_{d_{n-1}}$ if $n$ is odd, and either $g T_{c_{n-1}} g^{-1} = T_{d_{n-1}}$ or $g T_{c_{n-1}} g^{-1} = T_{d_n}$ if $n$ is even. By Theorem 3.6(1) there exists $v \in \mathrm{Ker}(\pi)$ such that $\rho_D(v) = g$. So, up to composing $\varphi$ on the left by $\mathrm{ad}_v$ first, and composing on the left by $\zeta$ if necessary after, we can assume that $(\rho_D \circ \varphi)(s_i) = T_{d_i} f_0$ for all $1 \leq i \leq n-1$, where $f_0$ commutes with $T_{d_i}$ for all $1 \leq i \leq n-1$. Since $T_{d_1} = \rho_D(t_1) \in \mathrm{Im}(\rho_D)$, we have $f_0 \in \mathrm{Im}(\rho_D)$, and hence there exists $u_0 \in A[D_n]$ such that $\rho_D(u_0) = f_0$. Since $\rho_D$ is injective (see Theorem 3.6), we deduce that $\varphi(s_i) = t_i u_0$ for all $1 \leq i \leq n-1$ and $u_0$ commutes with $t_i$ for all $1 \leq i \leq n-1$. We set $Y = \{t_1, \ldots, t_{n-1}\}$, $\Delta_Y = \Delta_Y[D_n]$, $\Delta_D = \Delta[D_n]$, $\kappa = 2$ if $n$ is odd, and $\kappa = 1$ if $n$ is even. By Paris [36, Theorem 1.1] the centralizer of $Y$ in $A[D_n]$ is generated by $\Delta_Y^2$ and $\Delta_D^\kappa$, and hence there exist $q, r \in \mathbb{Z}$ such that $u_0 = \Delta_Y^{2q} \Delta_D^{\kappa r}$. We conclude that $\varphi = \beta_{q,r}$. $\square$

# 6  Endomorphisms of $A[D_n]$

The following lemma is a counterpart of Lemma 5.8 for the case of odd $n$, and it is a preliminary to the proof of Theorem 2.3.
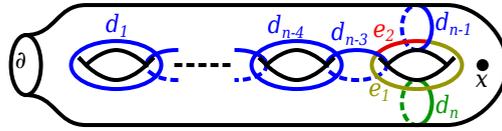
Figure 17: The circles $d_1, \dots, d_n$.

**Lemma 6.1** Let $n \geq 5$ be odd. Let $c$ be a generic circle of $\Sigma_n \setminus \{x\}$ such that $c \cap d_i = \varnothing$ for $1 \leq i \leq n-3$, $|c \cap d_{n-2}| = 1$, $c \cap d_{n-1} = \varnothing$ and $c$ is isotopic to $d_{n-1}$ in $\Sigma_n$. Then we have one of the following three possibilities:

(1) $c$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

(2) There exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-1$ and $g([c]) = [d_n]$.

(3) There exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, $g([d_{n-1}]) = [d_n]$ and $g([c]) = [d_{n-1}]$.

**Proof** The surface $\Sigma_n$ is of genus $\frac{1}{2}(n-1)$ with one boundary component, $\partial$. We assume that the circles $d_1, \dots, d_{n-1}, d_n$ are arranged as shown in Figure 17. The circles $d_{n-3}$ and $d_{n-1}$ divide $d_{n-2}$ into two arcs, $e_1$ and $e_2$, where the arc $e_1$ intersects $d_n$ and the arc $e_2$ does not intersect $d_n$ (see Figure 17). Let $\Omega$ be the surface obtained by cutting $\Sigma_n$ along $\bigcup_{i=1}^{n-1} d_i$. Then $\Omega$ is a cylinder represented by an octagon with a hole in the middle (see Figure 18). Two opposite sides of this octagon represent arcs of $d_{n-1}$ and two opposite sides represent arcs of $d_1, \dots, d_{n-3}$, as shown in the figure. Two other sides represent arcs of $e_1$ and the last two sides represent arcs of $e_2$, arranged as shown in Figure 18. The boundary of the hole represents $\partial$.

The circle $c$ intersects $d_{n-2}$ in a point $q$, and $q$ is either on the arc $e_1$ or on the arc $e_2$. Suppose first that $q$ is on the arc $e_1$. Then $q$ is represented on $\Omega$ by two points $q_1$ and $q_2$ lying on two different sides of $\Omega$ that represent $e_1$, and the trace of $c$ in $\Omega$ is a simple arc $\ell$ connecting $q_1$ to $q_2$. Up to isotopy (in $\Omega$ and not in $\Omega \setminus \{x\}$) pointwise fixing the boundary of $\Omega$, there are exactly two simple arcs in $\Omega$ connecting $q_1$ to $q_2$, represented by the arcs $\ell_1$ and $\ell_2$ depicted in Figure 18. The arc $\ell$ cannot be isotopic to $\ell_2$, otherwise $c$ would not be isotopic to $d_{n-1}$ in $\Sigma_n$. So $\ell$ is isotopic to $\ell_1$ in $\Omega$. Let $\{F_t \colon \Omega \to \Omega\}_{t \in [0,1]}$ be an isotopy such that $F_0 = \mathrm{id}$, $F_1(\ell) = \ell_1$ and $F_t$ is the identity on the boundary of $\Omega$ for all $t \in [0,1]$. The arc $\ell_1$
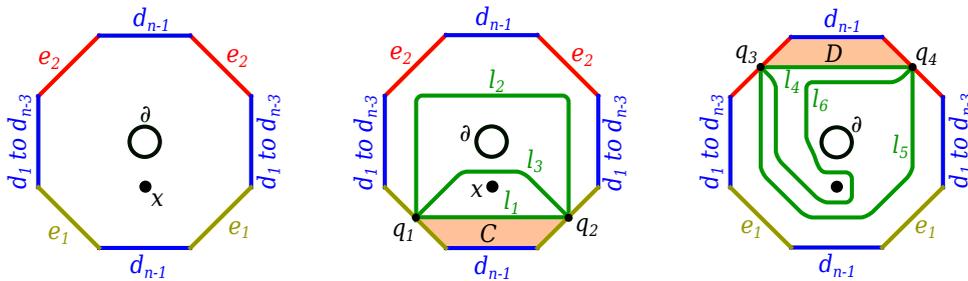


Figure 18: The surface $\Omega$.

divides $\Omega$ into two parts: the lower one, which does not contain the hole bounded by $\partial$ and the puncture $x$, and the upper one, which contains the hole bounded by $\partial$ and the puncture $x$, as shown in Figure 18.

Suppose $F_1(x)$ is in the upper part. Let $C$ be the domain of $\Omega$ bounded by $\ell_1$, two arcs of $e_1$ and an arc of $d_{n-1}$, as shown in Figure 18. Let $C' = F_1^{-1}(C)$. Then $C'$ is a domain of $\Omega$ bounded by $\ell$, two arcs of $e_1$ and an arc of $d_{n-1}$ which does not contain the puncture $x$. The existence of such a domain implies that $c$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

Suppose $F_1(x)$ is in the lower part. We can suppose that the trace of $d_n$ on $\Omega$ is the arc $\ell_3$ depicted in Figure 18. We can choose an isotopy $\{F'_t \colon \Omega \to \Omega\}_{t \in [0,1]}$ such that $F'_0 = \mathrm{id}$, $F'_1(\ell_1) = \ell_3$, $F'_t$ is the identity on the boundary of $\Omega$ for all $t \in [0,1]$, and $F'_1(F_1(x)) = x$. Let $\widetilde{F} \colon \Sigma_n \to \Sigma_n$ be the homeomorphism which is $F'_1 \circ F_1$ on $\Omega$ and is the identity outside $\Omega$, and let $g \in \mathcal{M}(\Sigma_n, x)$ be the mapping class represented by $\widetilde{F}$. Then $g \in \mathrm{Ker}(\theta)$, $g([d_i]) = [d_i]$ for all $1 \le i \le n-1$, and $g([c]) = [d_n]$.

Suppose now that $q$ is on the arc $e_2$. Then $q$ is represented on $\Omega$ by two points $q_3$ and $q_4$ lying on two different sides of $\Omega$ which represent $e_2$, and the trace of $c$ in $\Omega$ is a simple arc $\ell$ connecting $q_3$ to $q_4$. Up to isotopy (in $\Omega$ and not in $\Omega \setminus \{x\}$) pointwise fixing the boundary of $\Omega$, there are exactly two simple arcs in $\Omega$ connecting $q_3$ to $q_4$ represented by the arcs $\ell_4$ and $\ell_5$ depicted in Figure 18. The arc $\ell$ cannot be isotopic to $\ell_5$, otherwise $c$ would not be isotopic to $d_{n-1}$ in $\Sigma_n$. So $\ell$ is isotopic to $\ell_4$ in $\Omega$. Let $\{F_t \colon \Omega \to \Omega\}_{t \in [0,1]}$ be an isotopy such that $F_0 = \mathrm{id}$, $F_1(\ell) = \ell_4$ and $F_t$ is the identity on the boundary of $\Omega$ for all $t \in [0,1]$. The arc $\ell_4$ divides $\Omega$ into two parts: the upper one, which does not contain the hole bounded by $\partial$ and the puncture $x$, and the lower one, which contains the hole bounded by $\partial$ and the puncture $x$, as shown in Figure 18.

Suppose $F_1(x)$ is in the lower part. Let $D$ be the domain of $\Omega$ bounded by $\ell_4$, two arcs of $e_2$ and an arc of $d_{n-1}$ as shown in Figure 18. Let $D' = F_1^{-1}(D)$. Then $D'$ is a domain of $\Omega$ bounded by $\ell$, two arcs of $e_2$ and an arc of $d_{n-1}$ which does not contain the puncture $x$. The existence of such a domain implies that $c$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

Suppose $F_1(x)$ is in the upper part. Let $c'$ be the circle drawn in Figure 19. We can assume that the trace of $c'$ on $\Omega$ is the arc $\ell_6$ drawn in Figure 18. We can choose an isotopy $\{F'_t \colon \Omega \to \Omega\}_{t \in [0,1]}$ such that $F'_0 = \mathrm{id}$, $F'_1(\ell_4) = \ell_6$, $F'_t$ is the identity on the boundary of $\Omega$ for all $t \in [0,1]$, and $F'_1(F_1(x)) = x$. Let $\widetilde{F} \colon \Sigma_n \to \Sigma_n$ be the homeomorphism which is $F'_1 \circ F_1$ on $\Omega$ and is the identity outside $\Omega$, and let $g_1 \in \mathcal{M}(\Sigma_n, x)$ be the mapping class represented by $\widetilde{F}$. Then $g_1 \in \mathrm{Ker}(\theta)$, $g_1([d_i]) = [d_i]$ for all $1 \le i \le n-1$, and $g_1([c]) = [c']$.
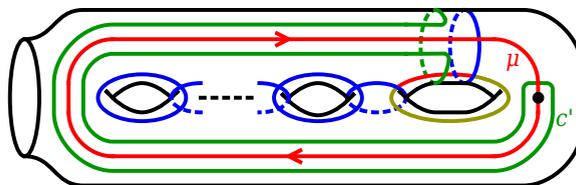


Figure 19: The circle $c'$ and the loop $\mu$.

Let $g_2 \in \pi_1(\Sigma_n, x) = \mathrm{Ker}(\theta)$ be the element represented by the loop $\mu$ drawn in Figure 19. Let us mention here that $g_2$ is not the Dehn twist $T_\mu$ along $\mu$, but rather the image of the point-pushing map applied to $\mu$, which is equal to $T_{\mu_1} T_{\mu_2}^{-1}$ for $\mu_1$ and $\mu_2$ the two boundary curves of a small regular neighborhood of $\mu$, as explained in Farb and Margalit [28, Section 4.2.2]. We have $g_2([d_i]) = [d_i]$ for all $1 \le i \le n-2$, $g_2([d_{n-1}]) = [d_n]$ and $g_2([c']) = [d_{n-1}]$. Set $g = g_2 \circ g_1$. Then $g \in \mathrm{Ker}(\theta)$, $g([d_i]) = [d_i]$ for all $1 \le i \le n-2$, $g([d_{n-1}]) = [d_n]$ and $g([c]) = [d_{n-1}]$. $\qquad\square$

**Proof of Theorem 2.3** Let $n \ge 6$. Let $\varphi \colon A[D_n] \to A[D_n]$ be an endomorphism. Consider the composition homomorphism $\varphi \circ \iota \colon A[A_{n-1}] \to A[D_n] \to A[D_n]$. We know from Theorem 2.2 that we have one of the following two possibilities up to conjugation:

(1)  $\varphi \circ \iota$ is cyclic.

(2)  There exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi \circ \iota = \psi \circ \beta_{p,q}$.

Suppose $\varphi \circ \iota$ is cyclic. Then there exists $u \in A[D_n]$ such that $\varphi(t_i) = (\varphi \circ \iota)(s_i) = u$ for all $1 \le i \le n-1$. We also have

$$\varphi(t_n) = \varphi(t_{n-2} t_n t_{n-2} t_n^{-1} t_{n-2}^{-1}) = \varphi(t_{n-2} t_n)\varphi(t_{n-2})\varphi(t_n^{-1} t_{n-2}^{-1}) = \varphi(t_{n-2} t_n)\varphi(t_1)\varphi(t_n^{-1} t_{n-2}^{-1}) = \varphi(t_1)$$

$$= u,$$

and hence $\varphi$ is cyclic.

So we can assume that there exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi \circ \iota$ is conjugate to $\psi \circ \beta_{p,q}$. We set $Y = \{t_1, \dots, t_{n-2}, t_{n-1}\}$, $\Delta_Y = \Delta_Y[D_n]$, $\Delta_D = \Delta[D_n]$, $\kappa = 2$ if $n$ is odd, and $\kappa = 1$ if $n$ is even. Up to conjugating and composing $\varphi$ on the left by $\zeta$ if necessary, we can assume that there exist $\varepsilon \in \{\pm 1\}$ and $p, q \in \mathbb{Z}$ such that $\varphi(t_i) = (\varphi \circ \iota)(s_i) = t_i^\varepsilon \Delta_Y^{2p} \Delta_D^{\kappa q}$ for all $1 \le i \le n-1$. The remainder of the proof is divided into four cases depending on whether $p$ is zero or not and whether $n$ is even or odd.

**Case 1** ($n$ is even and $p \ne 0$) Then $\Sigma_n$ is a surface of genus $\frac{1}{2}(n-2)$ with two boundary components, $\partial_1$ and $\partial_2$, and $\kappa = 1$. We have $\rho_D(t_i) = T_{d_i}$ for $1 \le i \le n-1$ and, by Labruère and Paris [32, Proposition 2.12], $\rho_D(\Delta_Y^2) = T_e T_{\partial_1}$ and $\rho_D(\Delta_D) = T_{\partial_1} T_{\partial_2}$, where $e$ is the circle drawn in Figure 20. Set $f_i = (\rho_D \circ \varphi)(t_i)$ for all $1 \le i \le n$. Then, by the above,

$$f_i = T_{d_i}^\varepsilon T_e^p T_{\partial_1}^{p+q} T_{\partial_2}^q \quad \text{for all } 1 \le i \le n-1.$$

In particular, $\mathcal{S}(f_i) = \{[d_i], [e]\}$ for all $1 \le i \le n-1$. Since $t_n$ is conjugate in $A[D_n]$ to $t_1$, $f_n$ is conjugate to $f_1$ in $\mathcal{M}(\Sigma_n, x)$; hence $f_n$ is of the form $f_n = T_{d'}^\varepsilon T_{e'}^p T_{\partial_1}^{p+q} T_{\partial_2}^q$, where $d'$ is a nonseparating circle
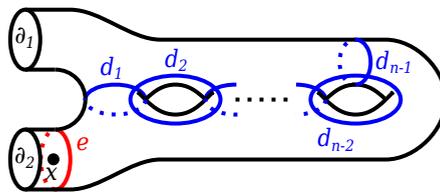


Figure 20: Circles in $\Sigma_n$ when $n$ is even and $p \ne 0$.

and $e'$ is a circle that separates $\Sigma_n$ into two components, one being a cylinder containing $x$ and the other being a surface of genus $\frac{1}{2}(n-2)$ with two boundary components, $\partial_1$ and $e'$, which does not contain $x$. Moreover, by Theorem 2.1, $(\pi \circ \varphi)(t_{n-1}) = (\pi \circ \varphi)(t_n)$, and hence

$$T_{d_{n-1}}^{\varepsilon} T_e^p T_{\partial_1}^{p+q} T_{\partial_2}^q = \theta(f_{n-1}) = \theta(f_n) = T_{d'}^{\varepsilon} T_{e'}^p T_{\partial_1}^{p+q} T_{\partial_2}^q$$

on $\Sigma_n$, that is, $T_{d_{n-1}}^{\varepsilon} T_e^p = T_{d'}^{\varepsilon} T_{e'}^p$ as multitwists on $\Sigma_n$. Now we can invoke Farb and Margalit [28, Lemma 3.14] to conclude that each curve of the set $\{d_{n-1}, e\}$ is isotopic to a curve from the set $\{d', e'\}$ in $\Sigma_n$. To decide which curve of one set is isotopic to which curve in the other set we observe that removing a puncture does not change the property of a curve being nonseparating, but can make a separating curve peripheral. Since both $d_{n-1}$ and $d'$ are nonseparating, whereas $e$ and $e'$ are both separating or peripheral in $\Sigma_n$, we conclude that $d_{n-1}$ is isotopic to $d'$ in $\Sigma_n$ (and also that $e$ is isotopic to $e'$ in $\Sigma_n$).

We have $f_1 f_n = f_n f_1$, and hence by Theorem 3.2(3) we have $f_n(\mathcal{S}(f_1)) = \mathcal{S}(f_1)$. Thus $[e]$ is a reduction class for $f_n$, and therefore $i([e], [e']) = 0$, because $[e']$ is an essential reduction class for $f_n$. We can choose representatives $e$ and $e'$ such that $e \cap e' = \varnothing$ either by eliminating bigons, or by choosing geodesic representatives. Let $C, C' \subset \Sigma_n$ be cylinders containing $x$ and having boundaries $\partial_2 \cup e$ and $\partial_2 \cup e'$, respectively. Then either $C \subset C'$ if $e \subset C'$, or $C' \subset C$ if $e' \subset C$, with $x \in C \cap C'$. Say $C \subset C'$. Being a separating circle on $\Sigma_n$, $e$ separates $C'$ into two subsurfaces, one containing $\partial_2$ and $x$, and the other containing $e'$. Being a subsurface with two boundary components lying inside a cylinder, the latter must be a cylinder itself. This cylinder establishes an isotopy between $e$ and $e'$ in $\Sigma_n \setminus \{x\}$, and hence $[e] = [e']$. So we can assume that $e = e'$.

Choose representatives $d_{n-1}$ and $d'$ in minimal position in $\Sigma_n \setminus \{x\}$. Denote by $C_0$ and $\Sigma'$ the two components into which the curve $e$ separates $\Sigma_n$, with $C_0$ being a cylinder containing $x$, and $\Sigma'$ being the rest of the surface $\Sigma_n$, containing $d_1, \ldots, d_{n-1}$. Suppose $d_{n-1} \cap d' \neq \varnothing$. Then $d_{n-1}$ and $d'$ cobound a bigon. Since $d_{n-1}$ and $d'$ were chosen to be in minimal position in $\Sigma_n \setminus \{x\}$, such a bigon must contain $x$. This implies that $d'$ has nonempty intersection with the cylinder $C_0$ which $e$ separates from the rest of the surface $\Sigma_n$, and since $e$ and $d'$ are disjoint, $d'$ lies entirely in $C_0$. This is not possible because any generic circle in $C_0$ is peripheral in $\Sigma_n$ and $d'$ is nonseparating in $\Sigma_n$. So $d_{n-1} \cap d' = \varnothing$. Then there exists an embedded cylinder $C$ in $\Sigma_n$ with boundary components $d_{n-1}$ and $d'$. Since $e$ is disjoint from $d'$ and $d_{n-1}$, $e$ either lies entirely in $C$ or is disjoint from $C$. The circle $e$ cannot lie entirely in $C$ because $e$ is peripheral in $\Sigma_n$ and, since both $d_{n-1}$ and $d'$ are nonseparating in $\Sigma_n$, any generic circle lying in $C$ must be nonseparating. So $e$ is disjoint from $C$, and hence $C$ lies in $\Sigma'$. Therefore $d_{n-1}$ is isotopic to $d'$ in $\Sigma_n \setminus \{x\}$. Thus we can also assume $d' = d_{n-1}$.

In conclusion we have $(\rho_D \circ \varphi)(t_{n-1}) = (\rho_D \circ \varphi)(t_n) = T_{d_{n-1}}^{\varepsilon} T_e^p T_{\partial_1}^{p+q} T_{\partial_2}^q$, and hence $\varphi(t_{n-1}) = \varphi(t_n) = t_{n-1}^{\varepsilon} \Delta_Y^{2p} \Delta_D^q$. We conclude that $\varphi = \beta_{p,q} \circ \pi$ if $\varepsilon = 1$ and $\varphi = \chi \circ \beta_{-p,-q} \circ \pi$ if $\varepsilon = -1$.

**Case 2** (*n is odd and $p \neq 0$*) Then $\Sigma_n$ is a surface of genus $\frac{1}{2}(n-1)$ with one boundary component, $\partial$, and $\kappa = 2$. We have $\rho_D(t_i) = T_{d_i}$ for $1 \leq i \leq n-1$ and, by Labruère and Paris [32, Proposition 2.12],
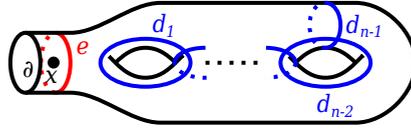
Figure 21: Circles in $\Sigma_n$ when $n$ is odd and $p \neq 0$.

$\rho_D(\Delta_Y^4) = T_e$ and $\rho_D(\Delta_D^2) = T_\partial$, where $e$ is the circle drawn in Figure 21. Set $f_i = (\rho_D \circ \varphi)(t_i)$ for all $1 \leq i \leq n$. Then, by the above,

$$f_i^2 = T_{d_i}^{2\varepsilon} T_e^p T_\partial^{2q} \quad \text{for all } 1 \leq i \leq n-1.$$

In particular, $\mathcal{S}(f_i) = \mathcal{S}(f_i^2) = \{[d_i], [e]\}$ for all $1 \leq i \leq n-1$. The element $t_n$ is conjugate to $t_1$ in $A[D_n]$; hence $\varphi(t_n)$ is conjugate to $\varphi(t_1)$ in $A[D_n]$, and therefore there exists $v \in A[D_n]$ such that $\varphi(t_n) = v\varphi(t_1)v^{-1} = (vt_1^\varepsilon v^{-1})(v\Delta_Y^{2p}v^{-1})\Delta_D^{2q}$. The element $\rho_D(vt_1v^{-1})$ is conjugate to $\rho_D(t_1) = T_{d_1}$, and hence $\rho_D(vt_1v^{-1}) = T_{d'}$, where $d'$ is a nonseparating circle. The element $\rho_D(v\Delta_Y^4 v^{-1})$ is conjugate to $\rho_D(\Delta_Y^4) = T_e$, and hence $\rho_D(v\Delta_Y^4 v^{-1}) = T_{e'}$, where $e'$ is a circle that separates $\Sigma_n$ into two components, one being a cylinder containing $x$ and the other being a surface of genus $\frac{1}{2}(n-1)$ with one boundary component which does not contain $x$. We also have $f_n^2 = T_{d'}^2 T_{e'}^p T_\partial^{2q}$ and $\mathcal{S}(f_n) = \mathcal{S}(f_n^2) = \{[d'], [e']\}$. By Theorem 2.1 $(\pi \circ \varphi)(t_{n-1}) = (\pi \circ \varphi)(t_n)$, and hence $\theta(f_{n-1}^2) = \theta(f_n^2)$. This implies that $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n$.

Since $f_1 f_n = f_n f_1$, by Theorem 3.2(3) we have $f_n^2(\mathcal{S}(f_1)) = \mathcal{S}(f_1)$; hence $[e]$ is a reduction class for $f_n^2$, and therefore $i([e], [e']) = 0$, because $[e']$ is an essential reduction class for $f_n^2$. As in Case 1, we can choose representatives $e$ and $e'$ such that $e \cap e' = \varnothing$. Let $C, C' \subset \Sigma_n$ be cylinders containing $x$ and having boundaries $\partial \cup e$ and $\partial \cup e'$, respectively. Then either $C \subset C'$ if $e \subset C'$, or $C' \subset C$ if $e' \subset C$, with $x \in C \cap C'$. Say $C \subset C'$. Being a separating circle on $\Sigma_n$, $e$ separates $C'$ into two subsurfaces, one containing $\partial$ and $x$, and the other containing $e'$. Being a subsurface with two boundary components lying inside a cylinder, the latter must be a cylinder itself. This cylinder establishes an isotopy between $e$ and $e'$ in $\Sigma_n \setminus \{x\}$, and hence $[e] = [e']$. So we can assume that $e = e'$, and hence $\rho_D(v\Delta_Y^4 v^{-1}) = T_{e'} = T_e = \rho_D(\Delta_Y^4)$. Since $\rho_D$ is injective, it follows that $v\Delta_Y^4 v^{-1} = \Delta_Y^4$.

Using the same argument as in Case 1, from the fact that $d'$ does not intersect $e' = e$ and that $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n$, it follows that $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$; hence we can also assume that $d' = d_{n-1}$. Then $\rho_D(vt_1v^{-1}) = T_{d'} = T_{d_{n-1}} = \rho_D(t_{n-1})$, and hence, since $\rho_D$ is injective, $vt_1v^{-1} = t_{n-1}$. At this stage of the proof we have that $\varphi(t_n) = t_{n-1}^\varepsilon(v\Delta_Y^{2p}v^{-1})\Delta_D^{2q}$ and $(v\Delta_Y^{2p}v^{-1})^2 = v\Delta_Y^{4p}v^{-1} = \Delta_Y^{4p}$. It remains to show that $v\Delta_Y^{2p}v^{-1} = \Delta_Y^{2p}$.

By Theorem 2.2 there exist $\psi \in \langle \zeta, \chi \rangle$ and $r, s \in \mathbb{Z}$ such that $\varphi \circ \zeta \circ \iota$ is conjugate to $\psi \circ \beta_{r,s}$. The automorphism $\zeta$ is inner since $n$ is odd, and hence we can assume that $\psi \in \langle \chi \rangle$. So there exist $w \in A[D_n]$, $\mu \in \{\pm 1\}$ and $r, s \in \mathbb{Z}$ such that $\varphi(t_i) = wt_i^\mu \Delta_Y^{2r} \Delta_D^{2s} w^{-1}$ for all $1 \leq i \leq n-2$ and $\varphi(t_n) = wt_{n-1}^\mu \Delta_Y^{2r} \Delta_D^{2s} w^{-1}$. Set $g = \rho_D(w)$. We have $(\rho_D \circ \varphi)(t_i^2) = T_{d_i}^{2\varepsilon} T_e^p T_\partial^{2q} = gT_{d_i}^{2\mu} T_e^r T_\partial^{2s} g^{-1}$ for all $1 \leq i \leq n-2$ and $(\rho_D \circ \varphi)(t_n^2) = T_{d_{n-1}}^{2\varepsilon} T_e^p T_\partial^{2q} = gT_{d_{n-1}}^{2\mu} T_e^r T_\partial^{2s} g^{-1}$. So $g^{-1}(\mathcal{S}(T_{d_i}^{2\varepsilon} T_e^p T_\partial^{2q})) =$

$\mathcal{S}(T_{d_i}^{2\mu} T_e^r T_{\partial}^{2s})$, and hence $g^{-1}(\{[d_i], [e]\}) \subset \{[d_i], [e]\}$ for all $1 \leq i \leq n-1$. This implies $g^{-1}([d_i]) = [d_i]$, and hence $g$ commutes with $T_{d_i}$; therefore $w$ commutes with $t_i$ for all $1 \leq i \leq n-1$. Since $\Delta_Y$ is in the subgroup of $A[D_n]$ generated by $Y = \{t_1, \ldots, t_{n-1}\}$ and $\Delta_D^2$ is central, it follows that $\varphi(t_i) = t_i^\mu \Delta_Y^{2r} \Delta_D^{2s}$ for all $1 \leq i \leq n-2$ and $\varphi(t_n) = t_{n-1}^\mu \Delta_Y^{2r} \Delta_D^{2s}$. Consider the equality $\varphi(t_1) = t_1^\varepsilon \Delta_Y^{2p} \Delta_D^{2q} = t_1^\mu \Delta_Y^{2r} \Delta_D^{2s}$. Then $t_1^{\varepsilon-\mu} \Delta_Y^{2(p-r)} = \Delta_D^{2(s-q)}$. The right-hand side of this equality lies in the center of $A[D_n]$, the left-hand side lies in $A_Y[D_n]$ and, by Paris [38, Corollary 2.6], the intersection of $A_Y[D_n]$ with the center of $A[D_n]$ is trivial; hence $s = q$ and $t_1^{\varepsilon-\mu} = \Delta_Y^{2(r-p)}$. The element $\Delta_Y^{2(r-p)}$ lies in the center of $A_Y[D_n]$ and $\langle t_1 \rangle$ is a proper parabolic subgroup of $A_Y[D_n]$; hence, again by Paris [38, Corollary 2.6], $t_1^{\varepsilon-\mu} = \Delta_Y^{2(r-p)} = 1$, and therefore $\varepsilon = \mu$ and $r = p$. Here we use that $A[D_n]$ is torsion-free, which follows from Deligne [25], where it is proved that $A[D_n]$ has a finite-dimensional classifying space. So $\varphi(t_n) = t_{n-1}^\varepsilon \Delta_Y^{2p} \Delta_D^{2q}$. We conclude that $\varphi = \beta_{p,q} \circ \pi$ if $\varepsilon = 1$ and $\varphi = \chi \circ \beta_{-p,-q} \circ \pi$ if $\varepsilon = -1$.

**Case 3** ($n$ is even and $p = 0$) Then, again, $\Sigma_n$ is a surface of genus $\frac{1}{2}(n-2)$ with two boundary components, $\partial_1$ and $\partial_2$, and $\kappa = 1$. We have $\rho_D(t_i) = T_{d_i}$ for $1 \leq i \leq n-1$ and, by Labrùere and Paris [32, Proposition 2.12], $\rho_D(\Delta_D) = T_{\partial_1} T_{\partial_2}$. Set $f_i = (\rho_D \circ \varphi)(t_i)$ for all $1 \leq i \leq n$. Then, by the above,

$$f_i = T_{d_i}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q \quad \text{for all } 1 \leq i \leq n-1.$$

In particular, $\mathcal{S}(f_i) = \{[d_i]\}$ for all $1 \leq i \leq n-1$. Since $t_n$ is conjugate in $A[D_n]$ to $t_1$, $f_n$ is of the form $f_n = T_{d'}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q$, where $d'$ is a nonseparating circle.

For $1 \leq i \leq n-3$ we have $t_i t_n = t_n t_i$; hence $T_{d_i} T_{d'} = T_{d'} T_{d_i}$, and therefore, by Proposition 3.1, $i([d_i], [d']) = 0$. Similarly, $i([d_{n-1}], [d']) = 0$. Since $t_{n-2} t_n t_{n-2} = t_n t_{n-2} t_n$, we have $T_{d_{n-2}} T_{d'} T_{d_{n-2}} = T_{d'} T_{d_{n-2}} T_{d'}$, and hence, by Proposition 3.1, $i([d_{n-2}], [d']) = 1$. So we can assume that $d_i \cap d' = \varnothing$ for $1 \leq i \leq n-3$, $d_{n-1} \cap d' = \varnothing$ and $|d_{n-2} \cap d'| = 1$. Moreover, by Theorem 2.1, $(\pi \circ \varphi)(t_{n-1}) = (\pi \circ \varphi)(t_n)$; hence $\theta(f_{n-1}) = \theta(f_n)$, and therefore $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n$. By Lemma 5.6 it follows that we have one of the following two possibilities:

(1) $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

(2) There exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-1$ and $g([d']) = [d_n]$.

Suppose $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$. Then $(\rho_D \circ \varphi)(t_n) = T_{d_{n-1}}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q$, and hence, since $\rho_D$ is injective, $\varphi(t_n) = t_{n-1}^\varepsilon \Delta_D^q$. We conclude that $\varphi = \beta_{0,q} \circ \pi$ if $\varepsilon = 1$ and $\varphi = \chi \circ \beta_{0,-q} \circ \pi$ if $\varepsilon = -1$.

Suppose there exists $g \in \mathrm{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-1$ and $g([d']) = [d_n]$. We have

$$(\rho_D \circ \varphi)(t_i) = T_{d_i}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q = g^{-1} T_{d_i}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q g$$

for all $1 \leq i \leq n-1$ and

$$(\rho_D \circ \varphi)(t_n) = T_{d'}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q = g^{-1} T_{d_n}^\varepsilon T_{\partial_1}^q T_{\partial_2}^q g.$$

By Theorem 3.6 there exists $v \in \mathrm{Ker}(\pi) \subset A[D_n]$ such that $\rho_D(v) = g$. Since $\rho_D$ is injective, it follows that

$$\varphi(t_i) = v^{-1} t_i^\varepsilon \Delta_D^q v \quad \text{for all } 1 \leq i \leq n.$$

We conclude that $\varphi = \mathrm{ad}_{v^{-1}} \circ \gamma_q$ if $\varepsilon = 1$ and $\varphi = \mathrm{ad}_{v^{-1}} \circ \chi \circ \gamma_{-q}$ if $\varepsilon = -1$.

**Case 4** (*n* is odd and $p = 0$)   Then, again, $\Sigma_n$ is a surface of genus $\frac{1}{2}(n-1)$ with one boundary component, $\partial$, and $\kappa = 2$. We have $\rho_D(t_i) = T_{d_i}$ for $1 \leq i \leq n-1$ and, by Labruère and Paris [32, Proposition 2.12], $\rho_D(\Delta_D^2) = T_\partial$. Set $f_i = (\rho_D \circ \varphi)(t_i)$ for all $1 \leq i \leq n$. Then, by the above,

$$f_i = T_{d_i}^\varepsilon T_\partial^q \quad \text{for all } 1 \leq i \leq n-1.$$

In particular, $\mathcal{S}(f_i) = \{[d_i]\}$ for all $1 \leq i \leq n-1$. Since $t_n$ is conjugate in $A[D_n]$ to $t_1$, $f_n$ is conjugate to $f_1$ in $\mathcal{M}(\Sigma_n, x)$, and hence $f_n$ is of the form $f_n = T_{d'}^\varepsilon T_\partial^q$ where $d'$ is a nonseparating circle.

For $1 \leq i \leq n-3$ we have $t_i t_n = t_n t_i$, and hence $T_{d_i} T_{d'} = T_{d'} T_{d_i}$. Therefore, by Proposition 3.1, $i([d_i], [d']) = 0$. Similarly, $i([d_{n-1}], [d']) = 0$. Since $t_{n-2} t_n t_{n-2} = t_n t_{n-2} t_n$, we have $T_{d_{n-2}} T_{d'} T_{d_{n-2}} = T_{d'} T_{d_{n-2}} T_{d'}$, and hence, by Proposition 3.1, $i([d_{n-2}], [d']) = 1$. So we can assume that $d_i \cap d' = \varnothing$ for $1 \leq i \leq n-3$, $d_{n-1} \cap d' = \varnothing$ and $|d_{n-2} \cap d'| = 1$. Moreover, by Theorem 2.1, $(\pi \circ \varphi)(t_{n-1}) = (\pi \circ \varphi)(t_n)$; hence $\theta(f_{n-1}) = \theta(f_n)$, and therefore $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n$. By Lemma 6.1 it follows that we have one of the following three possibilities:

(1)   $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$.

(2)   There exists $g \in \text{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-1$ and $g([d']) = [d_n]$.

(3)   There exists $g \in \text{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, $g([d_{n-1}]) = [d_n]$ and $g([d']) = [d_{n-1}]$.

If $d'$ is isotopic to $d_{n-1}$ in $\Sigma_n \setminus \{x\}$, then we prove as in the case where *n* is even that $\varphi = \beta_{0,q} \circ \pi$ if $\varepsilon = 1$ and $\varphi = \chi \circ \beta_{0,-q} \circ \pi$ if $\varepsilon = -1$. Similarly, if there exists $g \in \text{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-1$ and $g([d']) = [d_n]$, then we prove as in the case where *n* is even that $\varphi = \text{ad}_{v^{-1}} \circ \gamma_q$ if $\varepsilon = 1$ and $\varphi = \text{ad}_{v^{-1}} \circ \chi \circ \gamma_{-q}$ if $\varepsilon = -1$, where $v$ is an element of $\text{Ker}(\pi) \subset A[D_n]$.

Suppose there exists $g \in \text{Ker}(\theta)$ such that $g([d_i]) = [d_i]$ for all $1 \leq i \leq n-2$, $g([d_{n-1}]) = [d_n]$ and $g([d']) = [d_{n-1}]$. We have

$$(\rho_D \circ \varphi)(t_i) = T_{d_i}^\varepsilon T_\partial^q = g^{-1} T_{d_i}^\varepsilon T_\partial^q g \quad \text{for } 1 \leq i \leq n-2,$$

$$(\rho_D \circ \varphi)(t_{n-1}) = T_{d_{n-1}}^\varepsilon T_\partial^q = g^{-1} T_{d_n}^\varepsilon T_\partial^q g, \qquad (\rho_D \circ \varphi)(t_n) = T_{d'}^\varepsilon T_\partial^q = g^{-1} T_{d_{n-1}}^\varepsilon T_\partial^q g.$$

By Theorem 3.6 there exists $v \in \text{Ker}(\pi) \subset A[D_n]$ such that $\rho_D(v) = g$. Since $\rho_D$ is injective, it follows that

$$\varphi(t_i) = v^{-1} t_i^\varepsilon \Delta_D^{2q} v \quad \text{for } 1 \leq i \leq n-2, \qquad \varphi(t_{n-1}) = v^{-1} t_n^\varepsilon \Delta_D^{2q} v, \qquad \varphi(t_n) = v^{-1} t_{n-1}^\varepsilon \Delta_D^{2q} v.$$

We conclude that $\varphi = \text{ad}_{v^{-1}} \circ \zeta \circ \gamma_q$ if $\varepsilon = 1$ and $\varphi = \text{ad}_{v^{-1}} \circ \zeta \circ \chi \circ \gamma_{-q}$ if $\varepsilon = -1$.   □

# 7   Endomorphisms of $A[D_n]/Z(A[D_n])$

**Proof of Proposition 2.7**   Let $\Delta$ be the Garside element of $A[D_n]$. We set $\kappa = 2$ if *n* is odd and $\kappa = 1$ if *n* is even. Recall that $Z(A[D_n])$ is the cyclic group generated by $\Delta^\kappa$. Let $\varphi_Z \colon A_Z[D_n] \to A_Z[D_n]$ be an

endomorphism. For each $1 \leq i \leq n-2$ we define $u_i \in A[D_n]$ by induction on $i$ as follows. First choose any $u_1 \in A[D_n]$ such that $\xi(u_1) = \varphi_Z(t_{Z,1})$. Now assume that $2 \leq i \leq n-2$ and that $u_{i-1}$ is defined. Choose $u_i' \in A[D_n]$ such that $\xi(u_i') = \varphi_Z(t_{Z,i})$. Since $\varphi_Z(t_{Z,i-1}t_{Z,i}t_{Z,i-1}) = \varphi_Z(t_{Z,i}t_{Z,i-1}t_{Z,i})$, there exists $k_i \in \mathbb{Z}$ such that $u_{i-1}u_i'u_{i-1} = u_i'u_{i-1}u_i'\Delta^{\kappa k_i}$. Then set $u_i = u_i'\Delta^{\kappa k_i}$. Note that $\xi(u_i) = \xi(u_i') = \varphi_Z(t_{Z,i})$ and

$$u_{i-1}u_iu_{i-1} = u_{i-1}u_i'u_{i-1}\Delta^{\kappa k_i} = u_i'u_{i-1}u_i'\Delta^{2\kappa k_i} = u_iu_{i-1}u_i.$$

Define in the same way $u_{n-1}, u_n \in A[D_n]$ such that $\xi(u_{n-1}) = \varphi_Z(t_{Z,n-1})$, $\xi(u_n) = \varphi_Z(t_{Z,n})$, $u_{n-2}u_{n-1}u_{n-2} = u_{n-1}u_{n-2}u_{n-1}$ and $u_{n-2}u_nu_{n-2} = u_nu_{n-2}u_n$.

Let $i, j \in \{1, \ldots, n\}$ be such that $i \neq j$ and $t_it_j = t_jt_i$. We have $\varphi_Z(t_{Z,i}t_{Z,j}) = \varphi_Z(t_{Z,j}t_{Z,i})$, and hence there exists $l \in \mathbb{Z}$ such that $u_iu_j = u_ju_i\Delta^{\kappa l}$. Recall the homomorphism $z \colon A[D_n] \to \mathbb{Z}$ which sends $t_i$ to 1 for all $1 \leq i \leq n$. Since $z(\Delta) = n(n-1)$, the previous equality implies that

$$z(u_i) + z(u_j) = z(u_j) + z(u_i) + \kappa l n(n-1).$$

Hence $l = 0$, and therefore $u_iu_j = u_ju_i$.

By the above we have an endomorphism $\varphi \colon A[D_n] \to A[D_n]$ which sends $t_i$ to $u_i$ for all $1 \leq i \leq n$, and this endomorphism is a lift of $\varphi_Z$. $\square$

**Proof of Theorem 2.8** Let $n \geq 6$. Let $\varphi_Z \colon A_Z[D_n] \to A_Z[D_n]$ be an endomorphism. We know from Proposition 2.7 that $\varphi_Z$ admits a lift $\varphi \colon A[D_n] \to A[D_n]$. By Theorem 2.3 we have one of the following three possibilities up to conjugation:

(1) $\varphi$ is cyclic.

(2) There exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi = \psi \circ \beta_{p,q} \circ \pi$.

(3) There exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi = \psi \circ \gamma_p$.

Clearly, if $\varphi$ is cyclic then $\varphi_Z$ is cyclic.

Now we show that the second case cannot occur. Suppose there exist $\psi \in \langle \zeta, \chi \rangle$ and $p, q \in \mathbb{Z}$ such that $\varphi = \psi \circ \beta_{p,q} \circ \pi$. As ever, we set $\kappa = 2$ if $n$ is odd and $\kappa = 1$ if $n$ is even. Recall that the center of $A[D_n]$ is generated by $\Delta^\kappa$, where $\Delta$ is the Garside element of $A[D_n]$. We need to show that $\varphi(\Delta^\kappa) \notin Z(A[D_n]) = \langle \Delta^\kappa \rangle$, which leads to a contradiction. Since $\psi \in \mathrm{Aut}(A[D_n])$, we have $\psi(Z(A[D_n])) = Z(A[D_n])$, and hence we can assume that $\varphi = \beta_{p,q} \circ \pi$. Let $Y = \{t_1, \ldots, t_{n-1}\}$ and let $\Delta_Y = \Delta_Y[D_n]$ be the Garside element of $A_Y[D_n]$. Since

$$\Delta = (t_1 \cdots t_{n-2}t_{n-1}t_nt_{n-2} \cdots t_1) \cdots (t_{n-2}t_{n-1}t_nt_{n-2})(t_{n-1}t_n),$$
$$\Delta[A_{n-1}]^2 = (s_1 \cdots s_{n-2}s_{n-1}^2s_{n-2} \cdots s_1) \cdots (s_{n-2}s_{n-1}^2s_{n-2})s_{n-1}^2,$$

(see Lemma 5.1 for the second equality), we have $\pi(\Delta) = \Delta[A_{n-1}]^2$, and hence

$$\varphi(\Delta^\kappa) = (\beta_{p,q} \circ \pi)(\Delta^\kappa) = \beta_{p,q}(\Delta[A_{n-1}]^{2\kappa}) = \Delta_Y^{2\kappa(1+pn(n-1))}\Delta^{\kappa^2qn(n-1)}.$$

This element does not belong to $Z(A[D_n]) = \langle \Delta^\kappa \rangle$, because $\kappa(1+pn(n-1)) \neq 0$ and $\langle \Delta_Y^2 \rangle \cap \langle \Delta^\kappa \rangle = \{1\}$.

Suppose we are in the third case. So there exist $\psi \in \langle \zeta, \chi \rangle$ and $p \in \mathbb{Z}$ such that $\varphi = \psi \circ \gamma_p$. We have

$$\gamma_p(\Delta^\kappa) = \Delta^{\kappa(1+\kappa pn(n-1))} \in \langle \Delta^\kappa \rangle,$$

and hence $\gamma_p$ induces an endomorphism $\gamma_{Z,p} \colon A_Z[D_n] \to A_Z[D_n]$. Moreover, for all $1 \leq i \leq n$,

$$\gamma_{Z,p}(t_{Z,i}) = \xi(t_i \Delta^{\kappa p}) = \xi(t_i) = t_{Z,i},$$

so $\gamma_{Z,p} = \mathrm{id}$. Clearly $\psi$ is the lift of an element $\psi_Z \in \langle \zeta_Z, \chi_Z \rangle$, and hence $\varphi_Z = \psi_Z \circ \gamma_{Z,p} = \psi_Z$. $\square$

Now, as promised in Section 2, we prove the following.

**Lemma 7.1** *Let $n \geq 4$. If $n$ is even, then $\langle \zeta_Z, \chi_Z \rangle \cap \mathrm{Inn}(A_Z[D_n]) = \{\mathrm{id}\}$, and if $n$ is odd, then $\langle \chi_Z \rangle \cap \mathrm{Inn}(A_Z[D_n]) = \{\mathrm{id}\}$.*

**Proof** We first show that, if $\varphi \colon A[D_n] \to A[D_n]$ is an automorphism such that $\varphi_Z \in \mathrm{Inn}(A_Z[D_n])$, then $\varphi \in \mathrm{Inn}(A[D_n])$. Let $\varphi \in \mathrm{Aut}(A[D_n])$ be such that $\varphi_Z \in \mathrm{Inn}(A_Z[D_n])$. There exists $g_Z \in A_Z[D_n]$ such that $\varphi_Z(t_{Z,i}) = g_Z t_{Z,i} g_Z^{-1}$ for all $1 \leq i \leq n$. Again, we denote by $\Delta$ the Garside element of $A[D_n]$, and we set $\kappa = 2$ if $n$ is odd and $\kappa = 1$ if $n$ is even. Let $g \in A[D_n]$ be such that $\xi(g) = g_Z$. For every $1 \leq i \leq n$, there exists $k_i \in \mathbb{Z}$ such that $\varphi(t_i) = g t_i g^{-1} \Delta^{\kappa k_i}$. Let $i, j \in \{1, \ldots, n\}$ be such that $\{t_i, t_j\}$ is an edge of $D_n$. From the equality $t_i t_j t_i = t_j t_i t_j$ it follows that

$$g t_i t_j t_i g^{-1} \Delta^{\kappa(2k_i + k_j)} = \varphi(t_i t_j t_i) = \varphi(t_j t_i t_j) = g t_j t_i t_j g^{-1} \Delta^{\kappa(k_i + 2k_j)}.$$

Hence $2k_i + k_j = k_i + 2k_j$, and therefore $k_i = k_j$. Since $D_n$ is a connected graph, it follows that $k_i = k_j$ for all $i, j \in \{1, \ldots, n\}$. So there exists $k \in \mathbb{Z}$ such that $\varphi(t_i) = g t_i g^{-1} \Delta^{\kappa k}$ for all $1 \leq i \leq n$. Recall the homomorphism $z \colon A[D_n] \to \mathbb{Z}$ which sends $t_i$ to 1 for all $1 \leq i \leq n$. Since $\varphi$ is an automorphism, we have $\mathrm{Im}(z \circ \varphi) = \mathrm{Im}(z) = \mathbb{Z}$. Furthermore, since $z(\Delta) = n(n-1)$, we have $(z \circ \varphi)(t_i) = 1 + \kappa k n(n-1)$ for all $1 \leq i \leq n$, and hence $\mathrm{Im}(z \circ \varphi) = (1 + \kappa k n(n-1))\mathbb{Z}$. This implies that $k = 0$, and hence $\varphi = \mathrm{ad}_g \in \mathrm{Inn}(A[D_n])$.

Arguing in a similar way we can see that lifts of $\zeta_Z$ and $\chi_Z$ in $\mathrm{Aut}(A[D_n])$ are unique. Since we know that $\langle \zeta, \chi \rangle \cap \mathrm{Inn}(A[D_n]) = \{\mathrm{id}\}$ if $n$ is even and $\langle \chi \rangle \cap \mathrm{Inn}(A[D_n]) = \{\mathrm{id}\}$ if $n$ is odd, it follows that $\langle \zeta_Z, \chi_Z \rangle \cap \mathrm{Inn}(A_Z[D_n]) = \{\mathrm{id}\}$ if $n$ is even and $\langle \chi_Z \rangle \cap \mathrm{Inn}(A_Z[D_n]) = \{\mathrm{id}\}$ if $n$ is odd. $\square$

# References

[1] **B H An**, **Y Cho**, *The automorphism groups of Artin groups of edge-separated CLTTF graphs*, J. Korean Math. Soc. 60 (2023) 1171–1213  MR

[2] **R W Bell**, **D Margalit**, *Braid groups and the co-Hopfian property*, J. Algebra 303 (2006) 275–294  MR

[3] **M Bestvina**, *Non-positively curved aspects of Artin groups of finite type*, Geom. Topol. 3 (1999) 269–302  MR

[4] **J S Birman**, *Mapping class groups and their relationship to braid groups*, Comm. Pure Appl. Math. 22 (1969) 213–238 MR

[5] **J S Birman**, **H M Hilden**, *On isotopies of homeomorphisms of Riemann surfaces*, Ann. of Math. 97 (1973) 424–439 MR

[6] **J S Birman**, **A Lubotzky**, **J McCarthy**, *Abelian and solvable subgroups of the mapping class groups*, Duke Math. J. 50 (1983) 1107–1120 MR

[7] **N Bourbaki**, *Groupes et algèbres de Lie, Chapitres 4–6*, Actualités Scientifiques et Industrielles 1337, Hermann, Paris (1968) MR

[8] **C Bregman**, **R Charney**, **K Vogtmann**, *Outer space for RAAGs*, Duke Math. J. 172 (2023) 1033–1108 MR

[9] **E Brieskorn**, *Die Fundamentalgruppe des Raumes der regulären Orbits einer endlichen komplexen Spiegelungsgruppe*, Invent. Math. 12 (1971) 57–61 MR

[10] **E Brieskorn**, *Sur les groupes de tresses (d'après V I Arnold)*, from "Séminaire Bourbaki, 1971/1972", Lecture Notes in Math. 317, Springer (1973) exposé 201, pp. 21–44 MR

[11] **E Brieskorn**, **K Saito**, *Artin-Gruppen und Coxeter-Gruppen*, Invent. Math. 17 (1972) 245–271 MR

[12] **F Castel**, *Représentations géométriques des groupes de tresses*, PhD thesis, Université de Bourgogne (2009) Available at `https://theses.hal.science/tel-00552007v1/document`

[13] **F Castel**, *Geometric representations of the braid groups*, Astérisque 378, Soc. Math. France, Paris (2016) MR

[14] **R Charney**, **J Crisp**, *Automorphism groups of some affine and finite type Artin groups*, Math. Res. Lett. 12 (2005) 321–333 MR

[15] **R Charney**, **K Vogtmann**, *Finiteness properties of automorphism groups of right-angled Artin groups*, Bull. Lond. Math. Soc. 41 (2009) 94–102 MR

[16] **R Charney**, **K Vogtmann**, *Subgroups and quotients of automorphism groups of RAAGs*, from "Low-dimensional and symplectic topology" (M Usher, editor), Proc. Sympos. Pure Math. 82, Amer. Math. Soc., Providence, RI (2011) 9–27 MR

[17] **L Chen**, **K Kordek**, **D Margalit**, *Homomorphisms between braid groups*, preprint (2019) arXiv 1910.00712

[18] **H S M Coxeter**, *Discrete groups generated by reflections*, Ann. of Math. 35 (1934) 588–621 MR

[19] **H S M Coxeter**, *The complete enumeration of finite groups of the form $R_i^2 = (R_i R_j)^{k_{ij}} = 1$*, J. London Math. Soc. 10 (1935) 21–25 MR

[20] **J Crisp**, *Automorphisms and abstract commensurators of 2-dimensional Artin groups*, Geom. Topol. 9 (2005) 1381–1441 MR

[21] **J Crisp**, **L Paris**, *Artin groups of type B and D*, Adv. Geom. 5 (2005) 607–636 MR

[22] **M Cumplido**, **L Paris**, *Commensurability in Artin groups of spherical type*, Rev. Mat. Iberoam. 38 (2022) 503–526 MR

[23] **M B Day**, *Peak reduction and finite presentations for automorphism groups of right-angled Artin groups*, Geom. Topol. 13 (2009) 817–855 MR

[24] **M B Day**, *On solvable subgroups of automorphism groups of right-angled Artin groups*, Internat. J. Algebra Comput. 21 (2011) 61–70 MR

[25] **P Deligne**, *Les immeubles des groupes de tresses généralisés*, Invent. Math. 17 (1972) 273–302 MR

[26] **J L Dyer**, **E K Grossman**, *The automorphism groups of the braid groups*, Amer. J. Math. 103 (1981) 1151–1169 MR

[27] **D B A Epstein**, *Curves on 2-manifolds and isotopies*, Acta Math. 115 (1966) 83–107 MR

[28] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) MR

[29] **N D Gilbert**, **J Howie**, **V Metaftsis**, **E Raptis**, *Tree actions of automorphism groups*, J. Group Theory 3 (2000) 213–223 MR

[30] **P de la Harpe**, *Topics in geometric group theory*, Univ. Chicago Press (2000) MR

[31] **K Kordek**, **D Margalit**, *Homomorphisms of commutator subgroups of braid groups*, Bull. Lond. Math. Soc. 54 (2022) 95–111 MR

[32] **C Labruère**, **L Paris**, *Presentations for the punctured mapping class groups in terms of Artin groups*, Algebr. Geom. Topol. 1 (2001) 73–114 MR

[33] **M R Laurence**, *A generating set for the automorphism group of a graph group*, J. London Math. Soc. 52 (1995) 318–334 MR

[34] **H van der Lek**, *The homotopy type of complex hyperplane complements*, PhD thesis, Katholieke Universiteit Nijmegen (1983)

[35] **S Y Orevkov**, *Homomorphisms of commutator subgroups of braid groups with small number of strings*, Ann. Fac. Sci. Toulouse Math. 33 (2024) 105–121 MR

[36] **L Paris**, *Centralizers of parabolic subgroups of Artin groups of type $A_l$, $B_l$, and $D_l$*, J. Algebra 196 (1997) 400–435 MR

[37] **L Paris**, *Parabolic subgroups of Artin groups*, J. Algebra 196 (1997) 369–399 MR

[38] **L Paris**, *Artin groups of spherical type up to isomorphism*, J. Algebra 281 (2004) 666–678 MR

[39] **L Paris**, **D Rolfsen**, *Geometric subgroups of mapping class groups*, J. Reine Angew. Math. 521 (2000) 47–83 MR

[40] **B Perron**, **J P Vannier**, *Groupe de monodromie géométrique des singularités simples*, Math. Ann. 306 (1996) 231–245 MR

[41] **I Soroko**, *Artin groups of types $F_4$ and $H_4$ are not commensurable with that of type $D_4$*, Topology Appl. 300 (2021) art. id. 107770 MR

[42] **J Tits**, *Le problème des mots dans les groupes de Coxeter*, from "Symposia Mathematica, I", Academic, London (1969) 175–185 MR

[43] **N Vaskou**, *Automorphisms of large-type free-of-infinity Artin groups*, Geom. Dedicata 219 (2025) art. id. 16 MR

FC, LP: *IMB, UMR 5584, CNRS, Université Bourgogne Europe*
*Dijon, France*

fabrice.castel@gmail.com, lparis@u-bourgogne.fr

# Linear bounds of the crosscap number of knots

ROB MCCONKEY

Kalfagianni and Lee found two-sided bounds for the crosscap number of an alternating link in terms of certain coefficients of the Jones polynomial. We show here that we can find similar two-sided bounds for the crosscap number of Conway sums of strongly alternating tangles. Then we find families of links for which these coefficients of the Jones polynomial and the crosscap number grow independently. These families will enable us to show that neither linear bound generalizes for all links.

57K10, 57K14, 57K16

## 1 Introduction

In [17] Kalfagianni and Lee showed that the crosscap number of an alternating link admits two-sided linear bounds in terms of certain coefficients of the Jones polynomial of the link. The purpose of this paper is twofold; we first generalize the result of [17] for links that are the Conway sum of strongly alternating tangles. Second, we construct families of knots obstructing the generalization of the result of [17] to arbitrary knots.

For a link $L$ let

$$V(L) = \alpha_L t^n + \beta_L t^{n-1} + \cdots + \beta'_L t^{m+1} + \alpha'_L t^m,$$

where $\alpha_L$ and $\alpha'_L$ are nonzero, be the Jones polynomial, and let $T_L = |\beta_L| + |\beta'_L|$.

**Definition 1.1** For a nonorientable connected surface $S$, bounded by a link $L$, the *crosscap number* is defined to be

$$C(S) = 2 - \chi(S) - k_L,$$

where $k_L$ is the number of components of $L$. Then the *crosscap number* for a link $L$, denoted $C(L)$ will be the minimum crosscap number of all nonorientable surfaces bounded by the link.

Quantum knot invariants, such as the crosscap number, have historically been used to better understand classical knot invariants and the geometry of knots and links. In the eighties Kauffman [18] and Murasugi [25] showed that the span of the Jones polynomial realizes the crossing number for alternating links. More recent results from Futer, Kalfagianni, and Purcell showed how the coefficients of the colored Jones polynomial store information about the geometry of incompressible surfaces in the link complement and their strong relations to geometric structures and in particular hyperbolic geometry [11; 12; 13]. For example, specific coefficients of the colored Jones polynomial can coarsely define the volume of large

families of hyperbolic links [8; 10]. Further, Dasbach and Lin [7] did this for all hyperbolic alternating links. The volume conjecture [23] predicts that the asymptotics of the colored Jones polynomial can be used to calculate the volume of all hyperbolic links.

More recently, for alternating links Kalfagianni and Lee showed that there exist linear bounds for the crosscap number with respect to certain coefficients of the Jones polynomial (see Theorem 1.1 in [17]). Before their work, the best known lower bound for all alternating links was $C(K) \geq 1$. On the other side, Clark [4] observed that for any knot the crosscap number is bounded by

$$C(K) \leq 2g(K) + 1,$$

where $g(K)$ is the orientable genus of the knot. Kalfagianni and Lee's results gave an exact calculation of the crosscap number for 283 prime alternating knots on Knotinfo and improvements of 1472 prime alternating knots. Kindred [19] has also made progress in calculating the crosscap number, calculating it for alternating knots with less than 13 crossings. We hope our results in this paper will pave the way for similar calculations for nonalternating knots which are less understood.

Another source of motivation comes from the work that has been done concerning the orientable genus of links and the Alexander polynomial and Heegard Floer homology of a link. Crowell [5] and Murasugi [24] have independently shown that the orientable genus of an alternating link is half the degree span of the Alexander polynomial of the knot. The Heegard Floer homology is now known to detect the genus of links [22]. There is also an algorithm using normal surface theory to calculate the orientable knot genus [2; 15]. The hope with the crosscap number is to find parallel results to Heegard Floer and orientable genus using quantum invariants such as the Jones polynomial and Khovanov homology.

In this paper we will start by finding two-sided linear bounds for $C(L)$, where $L$ is a Conway sum of tangles in terms of $T_L$. We also define many of the necessary terms for Theorem 1.2 in that section, which we state here.

**Theorem 1.2**  *Let $T_1$ and $T_2$ be nonsplittable, twist reduced, strongly alternating tangles whose Conway sum is a link $L$. Let $C(L)$ be the crosscap number of $L$ and $k_L$ be the number of components of $L$. Then*

$$\left\lceil \tfrac{1}{6} T_L \right\rceil - k_L \leq C(L) \leq 2T_L + k_L + 8.$$

A key ingredient in the proof of Theorem 1.2 is Theorem 1.3 which we state below. Theorem 1.3 gives us bounds for $C(L)$ in terms of the crosscap numbers of the closures of the tangles which sum to $L$.

**Theorem 1.3**  *Let $T_1$ and $T_2$ be nonsplittable, twist reduced, strongly alternating tangles, and let $L$ be the link formed by the Conway sum of $T_1$ and $T_2$. Let $K_{iN}$ and $K_{iD}$ be the links formed by the numerator and denominator closures of $T_i$, respectively, $i \in \{1, 2\}$. We have*

$$m - 2 \leq C(L) \leq m + 2,$$

*where $m = \min\{C(K_{1N}) + C(K_{2N}), C(K_{1D}) + C(K_{2D})\}$.*

Having Theorem 1.3 at hand we use a result of [17] and the additivity of twist numbers for strongly alternating tangles to find bounds for $C(L)$ in terms of the twist number of $L$. Then using these bounds and a generalization of Theorem 1.6 from [9] by Futer, Kalfagianni, and Purcell gives us Theorem 1.2.

In Section 5 we show that for arbitrary knots the crosscap number and $T_L$ are independent. Specifically we show:

**Theorem 1.4** (a) *There exists a family of links for which $T_L \leq 2$, but $C(L)$ is arbitrarily large.*

(b) *There exists a family of links for which $C(L) \leq 3$, but $T_L$ is arbitrarily large.*

To show part (a) of Theorem 1.4, we use work by Teragaito [27] to find a family of torus knots $T(p,q)$ where $C(T(p,q))$ grows with $q$ and $q$ can be made arbitrarily large. On the other hand we will show that $T_{T(p,q)} \leq 2$.

For part (b) we will introduce a family of Whitehead doubles for which the crosscap number is always bounded by 3 but $|\beta'_W|$ can be made arbitrarily large. Clark [4] showed that for all links $C(L) \leq 2g(L)+1$ which shows that for all Whitehead doubles, $C(W) \leq 3$. On the other hand, using work by Stoimenow [26], we are able to compute $|\beta'_W|$ for B-adequate links. Then we find a family of B-adequate Whitehead doubles for which $|\beta'_W|$ can be made arbitrarily large.

All the links constructed in Theorem 1.4 are nonhyperbolic. This leaves the question of whether Theorem 1.2 may be generalized to all hyperbolic links. See Section 6 for more details.

### Acknowledgements

## 2 Crosscap bounds on connection of two strongly alternating tangles

In this section we will work to prove Theorem 1.3.

### 2.1 Preliminaries and the upper bound of Theorem 1.3

We start with a couple definitions.

**Definition 2.1** A *tangle* is a graph in the plane contained within a box which intersects the box at the four corners with one-valent vertices, with all other vertices, contained inside the box, four-valent, and given over/under crossing data. We label the four one-valent vertices NW, NE, SE, SW, positioned according to Figure 1.
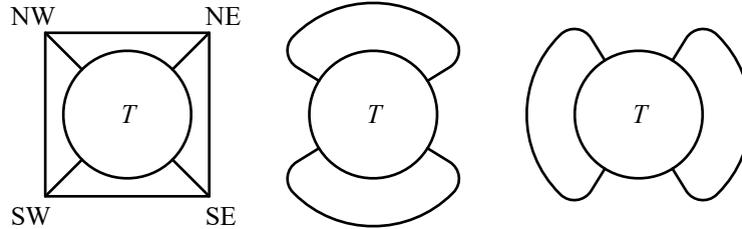
Figure 1: Left: a tangle inside a box with directional strands labeled. Center: numerator closure. Right: denominator closure.

**Definition 2.2** The *closure* of a tangle is the link which results when we connect the NW and NE points along the box and SW and SE points along the box as seen in the center panel of Figure 1, this is called the *numerator closure*. If we close as in the right-hand panel of Figure 1 we call it the *denominator closure*. A tangle is *strongly alternating* if both closures are prime, alternating, and contain no nugatory crossings.

**Definition 2.3** A *Conway sphere* is a 2-sphere which intersects a knot or link transversely in four points. A *Conway sum* is a sum of tangles as shown in Figure 2. For our purposes, a Conway sphere $\Sigma$ will be positioned such that it intersects a Conway sum at the four one-valent vertices for one of the tangles in the sum. If we let $S$ be a spanning surface for our Conway sum, then $S \cap \Sigma$ will contain two arcs and a possibly empty collection of simple closed curves as in Figure 3.

We are now ready to begin proving Theorem 1.3. We separate it into the upper and lower bounds, beginning with the upper bound:

**Lemma 2.4** *Let $T_1$ and $T_2$ be a pair of nonsplittable, strongly alternating tangles. Let $L$ be the link formed by the Conway sum of $T_1$ and $T_2$. Let $K_{iN}$ and $K_{iD}$ be the numerator and denominator closures, respectively, of $T_1$ and $T_2$. If $C(L)$ is the crosscap number of $L$, then*

$$C(L) \leq \min\{C(K_{1N}) + C(K_{2N}) + 2, C(K_{1D}) + C(K_{2D}) + 2\}.$$

**Proof** Start with a pair of strongly alternating tangles, $T_1$ and $T_2$. Let $K_{1N}$ and $K_{2N}$ be the links acquired by the numerator closures of the tangles. Let $S_1$ and $S_2$ be nonorientable spanning surfaces
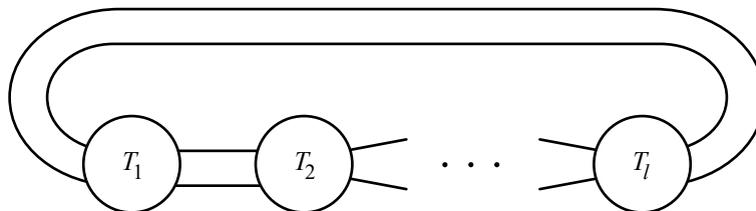


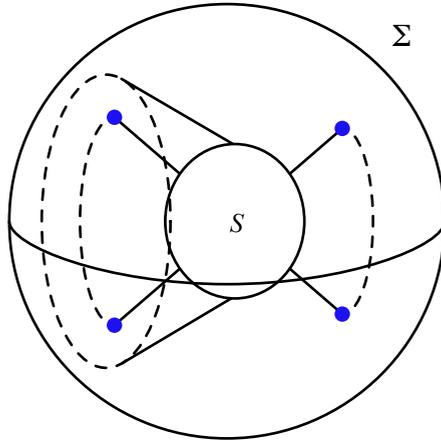Figure 2: An example of a Conway sum of $l$ tangles.

Figure 3: Here we see a tangle contained within a Conway sphere. The dots represent the intersection of $T_i$ with $\Sigma$. Then the dotted lines are the intersections of $S$ with $\Sigma$.

which realize the crosscap numbers of $K_{1N}$ and $K_{2N}$, respectively. We find a spanning surface $S$ of $L$ by attaching $S_1$ and $S_2$ with a pair of bands bounded by the strands along which the Conway sum was taken. As we do not cut $S_1$ and $S_2$, $S$ will also be nonorientable.

Now we study the relationship between $C(S)$ and the sum of $C(S_1)$ and $C(S_2)$. We remind the reader that $C(S) = 2 - \chi(S) - k$. The difference between $S$ and the disjoint union of $S_1$ and $S_2$ is the two connecting bands used to construct $S$. So, $\chi(S) = \chi(S_1) + \chi(S_2) - 2$.

Next we compare the number of link components in $L$ with the total in $K_{1N}$ and $K_{2N}$. The gluing of the East strands of $K_{1N}$ to the West strands of $K_{2N}$ will reduce the number of components by 1 as we are connecting two disjoint links. The other attachment can increase or decrease the number of components by 1, or keep it the same. Then $k_L = k_{K_{1N}} + k_{K_{2N}} - \epsilon$ where $\epsilon = 0, 1,$ or 2.

Now we substitute for $\chi(S)$ and $k_L$ to find

$$C(S) = 2 - \chi(S_1) - \chi(S_2) + 2 - k_{K_{1N}} - k_{K_{2N}} + \epsilon = C(K_{1N}) + C(K_{2N}) + \epsilon.$$

Hence

$$C(L) \le C(S) = C(K_{1N}) + C(K_{2N}) + 2,$$

as $\epsilon = 2$, will give the weakest upper bound. By the same argument with the denominator closures of $T_1$ and $T_2$, we find that

$$C(L) \le C(K_{1D}) + C(K_{2D}) + 2. \qquad \square$$

Here we remark that Lemma 2.4 will hold even if we take two general tangles. Notice in the proof that we do not use the fact that $T_1$ or $T_2$ are nonsplittable or strongly alternating. As this will not hold true for the other statements we included these hypotheses for uniformity.
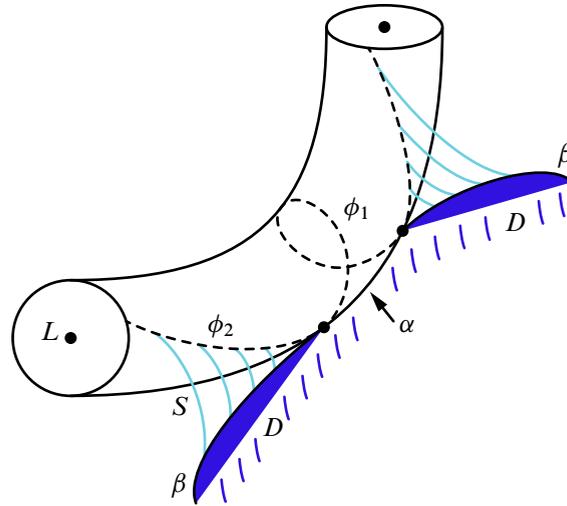
Figure 4: This figure shows a case where $L$ is meridianally boundary compressible.

## 2.2 Technical lemmas

Before we show the lower bound, we will need some more background, as well as some technical results. Lemma 2.5 below was discussed in the proof of Lemma 2.4.

**Lemma 2.5** *Let $L$ be the Conway sum of the tangles $T_1$ and $T_2$, and let $K_1$ and $K_2$ be closures of $T_1$ and $T_2$. If $k_L$ is the number of link components for $L$, and $k_1$ and $k_2$ the number of link components for $K_1$ and $K_2$, respectively, then $k_L = k_1 + k_2 - \epsilon$ for $\epsilon = 0, 1, 2$.*

**Definition 2.6** Let $L$ be a link in $S^3$ and let $N(L)$ be a neighborhood of $L$. A spanning surface $S$ of $L$ in $S^3$ is defined to be *meridianally boundary compressible* if there exists a disk $D$ embedded in $S^3 \setminus N(L)$ such that $\partial D = \alpha \cup \beta$ where $\alpha = D \cap \partial N(L)$ and $\beta = D \cap S$. Notice both $\alpha$ and $\beta$ are arcs, $\beta$ does not cut off a disk of $S$, $\partial D \cap \partial S$ cuts $\partial S$ into two arcs $\phi_1$, $\phi_2$ and $\alpha \cup \phi_i$ is a meridian of the link for one of $i = 1, 2$ as shown in Figure 4. A spanning surface is said to be *meridianally boundary incompressible* if no such disk exists.

**Definition 2.7** Given an alternating projection of a link $L$ on $S^2$, we modify it so that in a neighborhood of each crossing, we have a ball whose equator lies on $S^2$ such that the over strand runs over the ball and the under goes underneath; see Figure 5. We call every such ball a *Menasco ball*, and we call such an embedding of $L$ relative to $S^2$ a *Menasco projection $P$* with $n$ crossings.

**Definition 2.8** We say that a surface $S$ intersects a Menasco ball $B_i$ in a *crossing band* if $S \cap B_i$ consists of a disc bounded by the over and under strands on $\partial B_i$ along with opposite arcs along the equator of $B_i$. We refer the reader to Figures 28–30 in [1].
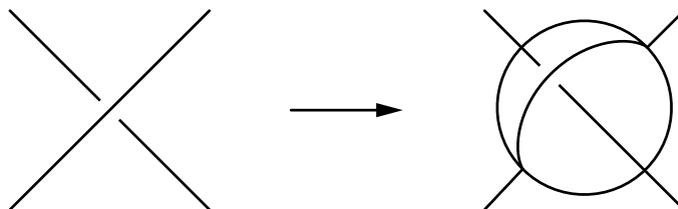
Figure 5: Here we put a crossing into Menasco form, with the over strand running across the top of the ball and the lower strand running along the bottom.

Let $F = S^2 \setminus \bigcup_i B_i$, where the $B_i$ are the Menasco balls for $L$. Given an incompressible (not necessarily meridianally) surface $S$ spanning $L$, we can isotope $S$ so that:

(i)  $S \cap F$ is a collection of simple closed curves and arcs with endpoints on $L$ or the equator of a Menasco ball.

(ii)  $S$ is disjoint wherever possible from the interior of the $B_i$, including along $N(L)$. The only exception will be at crossing bands.

We say such a surface isotoped in this way is in *Menasco form*.

We will also need Lemma 5.1 from [1] by Adams and Kindred which we state here as Lemma 2.9. Lemma 2.9 is required to prove Lemma 2.10, which is proven as Corollary 5.2 in [1]. Lemma 2.10 is essential to our proof of the lower bound, as it guarantees the connectedness of any spanning surface of the numerator or denominator closures of tangles that we consider.

**Lemma 2.9**  *An incompressible and meridianally boundary incompressible surface $S$ spanning an alternating link $L$ can be isotoped relative to a given nontrivial Menasco projection $P$ to obtain a crossing band.*

**Lemma 2.10**  *Any spanning surface for a nonsplittable, alternating link is connected.*

**Proof**  This proof uses induction on the number of crossings a nonsplittable, alternating link contains. As the unknot contains only one link component, any spanning surface for the unknot must be connected. Now consider a nonsplittable alternating link $L$ which has $n$ crossings, and let $S$ be a spanning surface for $L$. Then $S$ is either incompressible and meridianally boundary incompressible or a finite sequence of compressions take $S$ to an incompressible and meridianally boundary incompressible surface $S'$. Choose a reduced alternating diagram of $L$ and put $S'$ into Menasco form relative to $L$. By Lemma 2.9, we can isotope $S'$ such that $S'$ contains a crossing band in at least one of the Menasco balls, $M$. Further, when we cut open the link along $M$, we find a spanning surface $S''$ for a nonsplittable alternating link with fewer crossings $L'$. Part of the equator of $M$ replaces the crossing strands and guarantees that $S''$ is a spanning surface of $L'$. Then, by induction, as $S''$ is a spanning surface for a link of $n-1$ crossings, it is connected. Regluing in the crossing band does not disconnect our surface, showing that $S'$ and $S$ are connected. $\square$

**Lemma 2.11**  *Let $\Sigma$ be a Conway sphere which intersects a Conway sum $L$ of two strongly alternating tangles $T_1$ and $T_2$ in $S^3$ such that $\Sigma$ separates $T_1$ and $T_2$. If we let $S$ be a spanning surface of $L$ and $S \cap \Sigma$ contains a simple closed curve $\gamma$ such that $\gamma$ does not separate the two arcs in $S \cap \Sigma$ on $\Sigma$, then there exists an isotopy on $S$ which will eliminate $\gamma$.*

**Proof**  Assume there is only one closed curve $\gamma$ contained in $S \cap \Sigma$. First we consider the case where cutting $S$ along $\gamma$ and gluing in disks along the resulting boundary components results in two disconnected closed components. This implies that $S$ has a disconnected closed component, contradicting that $S$ is a crosscap realizing surface.

Next assume that cutting along $\gamma$ and gluing in disks does not result in a closed surface component. Then cutting $S$ along $\Sigma$ and gluing disks along the copies of $\gamma$ will result in spanning surfaces for a closure of $T_1$ and a closure of $T_2$, both of which are connected by Lemma 2.10. Reversing this procedure everywhere but $\gamma$ results in a new connected surface $S'$ which also spans $L$. But as $S'$ has two additional disks, $\chi(S') = \chi(S) + 2$, showing that $C(S') < C(S)$ contradicting that $S$ is a crosscap realizing spanning surface.

The only remaining possibility is if cutting $S$ along $\gamma$ and gluing disks to the two resulting boundaries results in a single closed surface component, $U$. Then $U$ will separate $S^3$ into two disjoint spaces. As $\gamma$ does not separate the two arcs on $\Sigma$, one side of $U$ must not contain any part of $S$. But then we can move $U$ to the opposite side of $\Sigma$ and reglue it to $S$ along $\gamma$ to find an isotopy of $S$ for which $\gamma$ is no longer in $\Sigma \cap S$.

In the case that we have multiple such closed curves along $\Sigma$ we do the same as above starting with the innermost closed curve. The innermost closed curve in this case is the one that bounds an empty disk on $\Sigma$.  $\square$

By Lemma 2.11, we can choose $S$ such that the only simple closed curves in $\Sigma \cap S$ are those which bound two disks each containing an arc. Next we show that $S$ can be chosen so that $S \cap \Sigma$ contains at most one such simple closed curve.

**Lemma 2.12**  *There exists a spanning surface $S$ for $L$, where $L$ is the Conway sum of two strongly alternating tangles, such that $C(S) = C(L)$ and $\Sigma \cap S$ contains at most one closed curve.*

**Proof**  By Lemma 2.11 we can assume that if $\Sigma \cap S$ contains closed curves $\gamma_1, \ldots, \gamma_n$, they each split $\Sigma$ such that the two arcs lie on opposite disks. Assume we have $n > 1$ such closed curves in $\Sigma \cap S$, and let $\gamma_1$ and $\gamma_2$ be such that $\gamma_1$ bounds a disk on $\Sigma$ such that no other $\gamma_i$ are in the disk and $\gamma_2$ bounds a disk where the only closed curve in it is $\gamma_1$.

We now find a spanning surface $S'$ such that $C(S') = C(L)$ and $S' \cap \Sigma$ contains $n - 2$ closed curves. We start with $S$ and cut along $\gamma_1$ and $\gamma_2$ and then glue in annuli whose boundaries are a copy of $\gamma_1$ and a copy of $\gamma_2$. Then as the Euler characteristic of an annulus is 0, this cutting and gluing operation will result in $\chi(S) = \chi(S')$.

It remains to show that $S'$ will be a connected surface. As in the proof of Lemma 2.11 we cut $S$ along $\Sigma$ and glue in disks along each $\gamma_i$ except for $\gamma_1$ and $\gamma_2$ which we glue a pair of annuli. This results in spanning surfaces $S_1$ and $S_2$ for closures of $T_1$ and $T_2$ which by Lemma 2.10 are connected. If we reverse this procedure everywhere except $\gamma_1$ and $\gamma_2$ the result will be $S'$ and as we only remove disks before regluing, $S'$ will be connected. Hence, we have found a spanning surface $S'$ for $L$ such that $C(S') = C(L)$ and $S' \cap \Sigma$ contains two fewer closed curves. Hence, repeating for all such pairs of closed curves in $S \cap \Sigma$ we will find the claim. $\square$

**Lemma 2.13** *We can choose a surface $S$ which spans a link $L$ such that $C(S) = C(L)$ and cutting along $\Sigma$ will not give us a closed surface component.*

**Proof** This was shown in the proof of Lemma 2.11. $\square$

## 2.3 Lower bound of Theorem 1.3

In this subsection we will prove the lower bound of Theorem 1.3, which will be restated as Lemma 2.14. We start by discussing what happens when we cut our link $L$ along $\Sigma$. Assume that $S$ is a nonorientable spanning surface for $L$ with $C(S) = C(L)$. The two arcs on $\Sigma$ will define how we close $T_1$ and $T_2$ after cutting. We let $K_1$ and $K_2$ be these closures. To see that $K_1$ and $K_2$ are the numerator or denominator closures consider a crossing which as a vertex in the tangle graph is adjacent to a one-valent vertex. If the exterior regions for a tangle are the faces bounded by the box in the graph then one of the two exterior regions adjacent to the crossing must be included in $S$. This means the boundary of this region will result in the numerator or denominator closures.

We are now ready to prove the lower bound of Theorem 1.3.

**Lemma 2.14** *Let $T_1$ and $T_2$ be nonsplittable, strongly alternating tangles and $L$ the link resulting from the Conway sum of $T_1$ and $T_2$. Also, let $S$ be a spanning surface of $L$ such that $C(L) = C(S)$. Then*

$$C(K_1) + C(K_2) - 2 \le C(L).$$

**Proof** Let $S$ be a nonorientable spanning surface for $L$ such that $C(L) = C(S)$. By Lemma 2.12 $S$ can be chosen such that the intersection of $S$ with $\Sigma$ contains at most one closed curve $\gamma$ and that both disks $\gamma$ bounds contain arcs. We cut $S$ along $\Sigma$ and if $\gamma$ exists we glue a disk to each copy to get spanning surfaces $S_1$ and $S_2$ for $K_1$ and $K_2$, respectively. By Lemma 2.13 we know that $S_1$ and $S_2$ will not have closed components and by Lemma 2.10 $S_1$ and $S_2$ must be connected as $K_1$ and $K_2$ are alternating.

If $k_1$ and $k_2$ are the number of link components for $K_1$ and $K_2$, respectively, then by Lemma 2.5 $k_1 + k_2 - \epsilon = k_L$ where $\epsilon = 0, 1, 2$.

Next we consider how the Euler characteristics of $S$ and the sum of the Euler characteristics of $S_1$ and $S_2$ will be related. We know that $\Sigma \cap S$ contains two arcs and at most one closed curve by Lemma 2.12. Then

cutting the two arcs along $\Sigma$ will increase the Euler characteristic by 2. Assume there are $n$ closed curves. Gluing disks along the two copies after cutting will further increase the Euler characteristic by $2n$. The final consideration we have to make is whether $S_1$ and $S_2$ are nonorientable. Let $t = 0, 1, 2$ be the number of $S_i$ which are orientable. We will have to add $t$ half twist bands to make sure all the $S_i$ are nonorientable decreasing the Euler characteristic by $t$. Now we see that $\chi(S) = \chi(S_1) + \chi(S_2) - 2 - 2n + t$. Then

$$C(S_1) + C(S_2) = 4 - \chi(S) - 2 - 2n + t - k_L - \epsilon = C(L) - 2n + t - \epsilon.$$

Simplifying we find

$$C(K_1) + C(K_2) + 2n - t + \epsilon \leq C(L).$$

This will be the weakest when $n = 0$, $t = 2$, and $\epsilon = 0$. Hence we find $C(K_1) + C(K_2) - 2 \leq C(L)$. $\square$

We restate Theorem 1.3:

**Theorem 1.3**  *Let $T_1$ and $T_2$ be nonsplittable, twist reduced, strongly alternating tangles. Let $L$ be the link formed by the Conway sum of $T_1$ and $T_2$. Let $K_{iN}$ be the link formed by the numerator closure of $T_i$, $i \in \{1, 2\}$, similarly $K_{iD}$ will be the link formed by the denominator closure. If we let $m = \min\{C(K_{1N}) + C(K_{2N}), C(K_{1D}) + C(K_{2D})\}$, then*

$$C(K_1) + C(K_2) - 2 \leq C(L) \leq m + 2.$$

Theorem 1.3 follows directly from Lemmas 2.4 and 2.14. A result similar to Theorem 1.3 exists for the cross cap number of connected sums. In particular, Clark [4] showed with a strategy similar to our own, that if $K_1$ and $K_2$ are knots, then

$$C(K_1) + C(K_2) - 1 \leq C(K_1 \# K_2) \leq C(K_1) + C(K_2).$$

# 3  Crosscap number, twist number, and the Jones polynomial

## 3.1  Twist number bounds

Now we have a relationship between the crosscap numbers of the Conway sum of two tangles and the closures of the tangles which compose it. Unfortunately, the bounds depend upon the tangles and which closures we take. But we can use Theorem 1.3 to find bounds for $C(L)$ entirely dependent upon $L$. Before proceeding with the statements, we will need a definition.

**Definition 3.1**  The *twist number* of a link diagram or a tangle diagram is the number of twist regions a link diagram contains, where a *twist region* is a maximal collection of bigon regions contained end to end. We call a link diagram *twist reduced* if any simple closed curve which meets the link diagram transversely at four points, with two points adjacent to one crossing and the other two another crossing, bounds a possibly empty collection of bigons arranged end to end between the two crossings.

We take a brief pause to mention that we can take the Conway sum of more than two tangles. In particular, for tangles $T_1, T_2, \ldots, T_n$, we can glue the eastern strands of $T_i$ to the western strands of $T_{i+1}$. Then we glue the eastern strands of $T_n$ to the western strands of $T_1$. See Figure 2 for an example.

**Lemma 3.2** *Let $T_1, T_2, \ldots, T_n$ be strongly alternating tangle diagrams whose Conway sum is a link diagram $D(L)$. Then $\mathrm{tw}(D(L)) = \sum_{i=1}^{n} \mathrm{tw}(T_i)$, where $\mathrm{tw}(D(L))$ is the twist number for $D(L)$ and $\mathrm{tw}(T_i)$ the twist number for the tangle diagram $T_i$.*

**Proof** First notice that taking the sum of tangles will not result in new twist regions. This is because, when taking a Conway sum, crossings that shared a twist region will still share a twist region and we introduce no new crossings. Therefore, $\mathrm{tw}(D(L)) \leq \sum_{i=1}^{n} \mathrm{tw}(T_i)$.

Now assume that $\mathrm{tw}(D(L)) < \sum_{i=1}^{n} \mathrm{tw}(T_i)$. Then for some $i$, a twist region in $T_i$ and a twist region in $T_{i+1}$ become one region in $L$. This implies there exists a simple closed curve $\gamma$ which transversely intersects $D(L)$ twice in $T_i$ and twice in $T_{i+1}$. If we think back to $T_i$ lying in a unit square, then $\gamma$ must intersect the north and south edges or the east and west edges of the square. In the first case, this shows that the denominator closure is not prime, and the second, the numerator closure is not prime. But as $T_i$ is strongly alternating, this would be a contradiction, therefore $\mathrm{tw}(D(L)) = \sum_{i=1}^{n} \mathrm{tw}(T_i)$. $\square$

Next we consider the relationship between the twist number of a tangle diagram and the twist numbers of diagrams of its closures.

**Lemma 3.3** *Let $T$ be a strongly alternating tangle diagram, and let $D(K)$ be the link diagram which comes from the numerator or denominator closure. Then*

$$\mathrm{tw}(T) - 2 \leq \mathrm{tw}(D(K)) \leq \mathrm{tw}(T).$$

**Proof** The upper bound is true as we are not adding crossings when closing a tangle, and hence cannot create new twist regions.

The lower bound stems from the fact that when we choose a closure for $T$ we create two new potential bigons. If either region is a bigon, then it joins two twist regions. If both regions are bigons, the twist number is reduced by 2. In Figure 6 we see an example of a tangle where the lower bound is sharp for both closures. $\square$

We will also need Theorem 3.8 from [17] which we state here. This theorem allows us to relate the cross cap numbers of the closures of strongly alternating tangles to twist numbers of their diagrams.

**Theorem 3.4** *Let $L \subset S^3$ be a link of $k_L$ components with a prime, twist reduced, alternating diagram $D(L)$. Suppose that $D(L)$ has $\mathrm{tw}(D(L)) \geq 2$ twist regions. Let $C(L)$ denote the crosscap number of $L$. We have*

$$\left\lceil \tfrac{1}{3} \mathrm{tw}(D(L)) \right\rceil + 2 - k_L \leq C(L) \leq \mathrm{tw}(D(L)) + 2 - k_L.$$

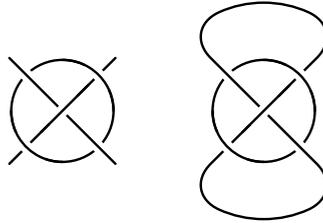*Furthermore, both bounds are sharp.*

Figure 6: Left: a strongly alternating tangle with 5 twist regions. Right: the numerator closure with 3 twist regions. The denominator also results in 3 twist regions, showing the sharpness of the lower bound.

Now that we have that the twist number is additive for strongly alternating tangles by Lemma 3.2 and can relate $C(K_i)$ and $\mathrm{tw}(T_i)$ by Lemma 3.3 and Theorem 3.4, we are ready to state Theorem 3.5

**Theorem 3.5** *Let $T_1$ and $T_2$ be diagrams of nonsplittable, strongly alternating, twist reduced tangles whose Conway sum is a link diagram $D(L)$. Let $C(L)$ be the crosscap number of $L$, $\mathrm{tw}(D(L))$ be the twist number of $D(L)$ and $k_L$ be the number of link components in $L$. Then*

$$\left\lceil \tfrac{1}{3}\mathrm{tw}(D(L)) \right\rceil - k_L \le C(L) \le \mathrm{tw}(D(L)) + 4 - k_L.$$

**Proof** We start with a lower bound in the proof of Lemma 2.14 with ambiguity on $\epsilon$ where $k_1 + k_2 - \epsilon = k_L$. Then $C(K_1) + C(K_2) - 2 + \epsilon \le C(L)$. Let $D(K_1)$ and $D(K_2)$ be the diagrams of $K_1$ and $K_2$ that arise from cutting $L$ as in Theorem 1.3. Notice that $\mathrm{tw}(D(K_i)) \ge 2$ for $i = 1, 2$ as $T_i$ is strongly alternating, and tangle diagrams with twist number 1 will have a nonprime closure. Then by Theorem 3.4 we find for $i \in \{1, 2\}$ that $\left\lceil \tfrac{1}{3}\mathrm{tw}(D(K_i)) \right\rceil + 2 - k_i \le C(K_i)$ where $K_i$ has $k_i$ link components. So

$$\left\lceil \tfrac{1}{3}\mathrm{tw}(D(K_1)) \right\rceil + \left\lceil \tfrac{1}{3}\mathrm{tw}(D(K_2)) \right\rceil + 2 + \epsilon - k_1 - k_2 \le C(L).$$

By Lemma 3.3 $\mathrm{tw}(T_i) - 2 \le \mathrm{tw}(D(K_i))$ and from Lemma 3.2 we know $\mathrm{tw}(T_1) + \mathrm{tw}(T_2) = \mathrm{tw}(D(L))$. Combining the two lemmas shows

$$(1) \qquad \left\lceil \tfrac{1}{3}\mathrm{tw}(D(L)) \right\rceil - 2 \le \left\lceil \tfrac{1}{3}(\mathrm{tw}(T_1) - 2) \right\rceil + \left\lceil \tfrac{1}{3}(\mathrm{tw}(T_2) - 2) \right\rceil$$

$$(2) \qquad \le \left\lceil \tfrac{1}{3}\mathrm{tw}(D(K_1)) \right\rceil + \left\lceil \tfrac{1}{3}\mathrm{tw}(D(K_2)) \right\rceil.$$

Finally substituting in $k_1 + k_2 = k_l + \epsilon$, we find

$$\left\lceil \tfrac{1}{3}\mathrm{tw}(D(L)) \right\rceil - k_L \le C(L).$$

Now we consider the upper bound. Similar to the lower bound we start with a step from Lemma 2.4, $C(L) \le \min\{C(K_{1N}) + C(K_{2N}) + \epsilon, C(K_{1D}) + C(K_{2D}) + \epsilon\}$. By Lemma 3.3 we see that $\mathrm{tw}(D(K_{iN})) \le \mathrm{tw}(T_i)$ and $\mathrm{tw}(D(K_{iD})) \le \mathrm{tw}(T_i)$ and then by Theorem 3.4,

$$C(L) \le \mathrm{tw}(T_1) + \mathrm{tw}(T_2) + 4 + \epsilon - k_1 - k_2.$$

Then substituting for $k_1 + k_2 = k_L + \epsilon$ and considering Lemma 3.2,

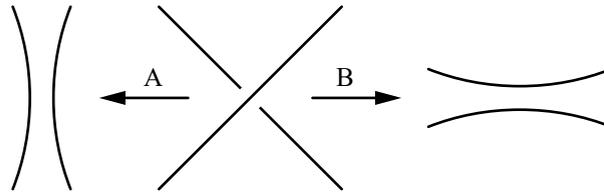$$C(L) \le \mathrm{tw}(D(L)) + 4 - k_L. \qquad \square$$

Figure 7: Given a crossing, we can resolve it to either the A- or B-resolution.

## 3.2 Jones polynomial bounds

From here we work to find bounds in terms of $T_L$, but first we have to generalize Theorem 1.6 in [9] which will allow us to relate the twist number of the diagram of a link $L$ to $T_L$. Theorem 1.6 from [9] only considers knots but we want a similar result for links. We start with some necessary definitions and then a generalization of Lemma 5.4 from [9] which is a necessary piece of our generalization of Theorem 1.6.

**Definition 3.6** If we let $D(L)$ be the diagram of a link $L$ we define the *A-resolution* and *B-resolution* as shown in Figure 7. Then a *Kauffman state* is a choice of resolutions for each crossing in a link diagram.

**Definition 3.7** Next we construct a *all A (resp all B) state graph* $\mathbb{G}_A(D(L))$ (resp $\mathbb{G}_B(D(L))$) by adding edges where we performed resolutions and then contracting the simple closed curves to vertices. Let $e_A$ (resp $e_B$) be the number of edges in $\mathbb{G}_A(D(L))$ (resp $\mathbb{G}_B(D(L))$). Further, if we identify all parallel edges (edges which share two vertices) we find the *reduced state graph* $\mathbb{G}'_A(D(L))$ (resp $\mathbb{G}'_B(D(L))$). Let $e'_A$ (resp $e'_B$) be the number of edges in $\mathbb{G}'_A(D(L))$ (resp $\mathbb{G}'_B(D(L))$).

Now we are ready to define what it means for a link to be adequate.

**Definition 3.8** We call a link diagram *A-adequate* (resp *B-adequate*) if the A-state (resp B-state) graph of the diagram has no one-edge loops. A link diagram is called *adequate* if it is both A-adequate and B-adequate, and a link is adequate if it has a diagram which is adequate.

Here we recall some terminology from [9].

**Definition 3.9** Let $D(L)$ be the link diagram obtained by taking the Conway sum of strongly alternating tangles $T_1, \ldots, T_n$. Let $\ell_{\text{in}}(D(L))$ denote the loss of edges in $\mathbb{G}_A(D(L))$ and $\mathbb{G}_B(D(L))$ as we pass from $e_A + e_B$ to $e'_A + e'_B$ which come from equivalent crossings in the same tangle $T_i$. Then let $\ell_{\text{ext}}(D(L))$ be the number of edges we lose from identification when we take the Conway sum. It follows that $\ell_{\text{in}}(D(L)) + \ell_{\text{ext}}(D(L)) = e_A + e_B - e'_A - e'_B$.

For an alternating tangle diagram $T$, the vertices of $\mathbb{G}_A(T)$ and $\mathbb{G}_B(T)$ are in one-to-one correspondence with the regions of $T$. For the state graph of a tangle, if we consider the tangle lying within a disk, we have four exterior regions bounded by the disk. This means our state graphs have *interior vertices* those whose region lie entirely within the interior of the disk and two *exterior vertices* with corresponding region, with sides on the boundary of the disk.
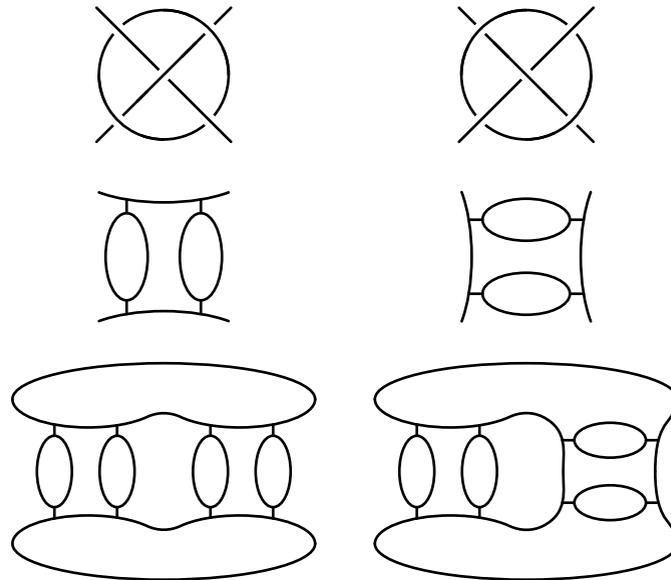
Figure 8: Top: two strongly alternating tangles. Middle: the all A-resolution diagrams for the tangles. Bottom left: the all A-state diagram for the Conway sum of two of the left tangles which has all admissible bridges. Bottom right: the all A-state diagram for the Conway sum of one of both tangles; here the bridges are all inadmissible.

For a tangle $T_i$, a *bridge* of $\mathbb{G}_A(T_i)$ or $\mathbb{G}_B(T_i)$ is a subgraph consisting of an interior vertex $v$, and edges $e'$, $e''$ which connect $v$ to the exterior vertices $v'$ and $v''$. We call the bridge *inadmissible* if the vertices become identified in $\mathbb{G}_A(D(L))$ or $\mathbb{G}_B(D(L))$; see Figure 8.

Now we find an upper bound for $\ell_{\text{ext}}(D(L))$ in terms of the twist number. This work will largely follow the proof of Lemma 5.4 in [9].

**Lemma 3.10** *Let $T_1$ and $T_2$ be strongly alternating tangles whose Conway sum is a link diagram $D(L)$. Let $k_L$ be number of link components in $L$. Then*

$$\ell_{\text{ext}}(D(L)) \le \tfrac{1}{2}\operatorname{tw}(D(L)) + k_L + 4.$$

**Proof** For $T \in \{T_1, T_2\}$ let $b_A(T)$, $b_B(T)$ be the number of bridges in $\mathbb{G}_A(T)$ and $\mathbb{G}_B(T)$, respectively. Then the contribution of $T$ to $\ell_{\text{ext}}$ will be at most $b_A(T) + b_B(T)$. Any other edge identification from moving to the reduced graph will still be counted by $\ell_{\text{int}}$.

If $b$ is a bridge there are two possibilities:

  (i)  The edges $e'$, $e''$ do not come from the resolutions of a single twist region.

  (ii) The edges $e'$, $e''$ come from the resolutions of a single twist region.

For type-(ii) bridges the two crossings which result in the edges are the only two in their respective twist region. Otherwise the two edges will not be adjacent to the exterior vertices or this would no longer constitute a twist region.

For type-(i) bridges when we pass from $\mathbb{G}_A(T)$ and $\mathbb{G}_B(T)$ to $\mathbb{G}'_A(T)$ and $\mathbb{G}'_B(T)$ the contributions to $\ell_{\text{ext}}$ is half the number of twist regions involved in such bridges. Unlike in [9] we can have more than one type-(ii) bridge, as each additional type-(ii) bridge creates a new link component.

**Case 1** Suppose that $b_A(T) \geq 3$ or $b_B(T) \geq 3$. Without loss of generality let $b_A(T) \geq 3$. Then $b_B(T) = 0$. If $b_B(T)$ were not zero then the $B$ state bridge would cross the $A$ state bridges, implying two internal vertices which is not a bridge.

There can be any number of type-(ii) bridges, but each bridge beyond the first will add a new link component. If we have only type-(ii) bridges $b_A(T) \leq k_T$ where $k_T$ is the number of tangle components. On the other hand if we only have type-(i) bridges $b_A(T) \leq \frac{1}{2} \operatorname{tw}(T)$. Then for any mix of bridges we find that

$$b_A(T) + b_B(T) \leq \tfrac{1}{2} \operatorname{tw}(T) + k_T.$$

**Case 2** In this case we will consider $b_A(T) = b_B(T) = 2$. Then $k$ must be at least 2, as the bridges in $\mathbb{G}_A(T)$ and $\mathbb{G}_B(T)$ will create a square resulting in a tangle second component. Also, as $\mathbb{G}_A(T)$ has two bridges there are at least two twist regions in $T$. Then

$$b_A(T) + b_B(T) \leq \tfrac{1}{2} \operatorname{tw}(T) + k_T + 1.$$

**Case 3** Either $b_A(T) \leq 2$ and $b_B(T) \leq 1$ or $b_A(T) \leq 1$ and $b_B(T) \leq 2$. Without loss of generality consider the first possibility. Then $b_A(T) + b_B(T)$ is at most three. If it's less than three we see that $k_T + 1 \geq 2$ so we need only consider when they sum to three. But as with the previous case we will have at least two twist regions as $\mathbb{G}_A(T)$ has two bridges. Thus,

$$b_A(T) + b_B(T) \leq \tfrac{1}{2} \operatorname{tw}(T) + k_T + 1.$$

Then by Lemma 3.2 we know that the twist number is additive over Conway sums. By Lemma 2.5 $k_1 + k_2 \leq k_L + 2$. Then we find the bound

$$\ell_{\text{ext}} \leq \sum_{i=1}^{2} b_A(T_i) + b_B(T_i) \leq \tfrac{1}{2} \operatorname{tw}(D(L)) + k_L + 4. \qquad \square$$

Now we have the tools necessary to prove the main lemma needed to find bounds for $C(L)$ in terms of $T_L$.

**Lemma 3.11** *Let $T_1, \ldots, T_n$ be strongly alternating tangles whose Conway sum is a link diagram $D(L)$ for a link $L$. Then letting $\beta_L$ and $\beta'_L$ be the second and second-to-last coefficients of the Jones polynomial of $L$, $T_L = |\beta_L| + |\beta'_L|$, and $k_L$ the number of link components, we have*

$$\tfrac{1}{2} \operatorname{tw}(D(L)) - k_L - 2 \leq T_L \leq 2 \operatorname{tw}(D(L)).$$

**Proof** We start by noting that we can mutate the link $L$ in such a way that it either is alternating or the sum of $T$ and $T'$ where $T$ is a positive strongly alternating tangle and $T'$ is a negative strongly alternating tangle, without changing its Jones polynomial [20]. The positive and negative refer to whether the northwest strand originates from an overcrossing or an undercrossing. In the former case we have a stronger result by Dasbach and Lin [7] that $T_L = \mathrm{tw}(L)$.

We will assume that $L$ is not alternating. Then work by Lickorish and Thistlethwaite [21] shows that $D(L)$ is adequate. Further, by Propositions 1 and 5 of [21] we have $v_A + v_B = c$ where $v_A$ is the number of vertices in $\mathbb{G}_A(D(L))$, $v_B$ the same in $\mathbb{G}_B(D(L))$ and $c$ the number of crossings in $D(L)$. Every edge we lose when passing from $\mathbb{G}_A(D(L))$ and $\mathbb{G}_B(D(L))$ to $\mathbb{G}'_A(D(L))$ and $\mathbb{G}'_B(D(L))$ comes from either multiple edges in a twist region or an inadmissible bridge. By Lemma 5.2 in [9] we see that the number of edges lost due to twist regions is $c - \mathrm{tw}(D(L))$ which is equivalent to $\ell_{\mathrm{in}}$. This is because the edges which share a twist region are equivalent to those who become identified in the reduction before taking the Conway sum. On the other hand type-(i) inadmissible bridges for strongly alternating tangles will be precisely those which we lose from identification when taking the Conway sum, $\ell_{\mathrm{ext}}(D(L))$. Then the first part of the move from (4) to (5) below will come from the equality $e'_A + e'_B - e_A - e_B = -(c - \mathrm{tw}(D(L)) + \ell_{\mathrm{ext}})$. Work by Stoimenow shows that for an adequate link diagram (see [7] for a proof)

$$(3) \qquad T_L = e'_A + e'_B - v_A - v_B + 2$$

$$(4) \qquad = (e'_A + e'_B - e_A - e_B) + e_A + (e_B - v_A - v_B) + 2$$

$$(5) \qquad = -(c - \mathrm{tw}(D(L)) + \ell_{\mathrm{ext}}) + c + (c - v_A - v_B) + 2$$

$$(6) \qquad \geq \mathrm{tw}(D(L)) - \ell_{\mathrm{ext}} + 2$$

$$(7) \qquad \geq \mathrm{tw}(D(L)) - \tfrac{1}{2}\mathrm{tw}(D(L)) - k_L - 4 + 2 = \tfrac{1}{2}\mathrm{tw}(D(L)) - k_L - 2.$$

The upper bound on $T_L$ was shown by Futer, Kalfagianni and Purcell in [8]. $\qquad\square$

**Theorem 1.2** *Let $T_1$ and $T_2$ be nonsplittable, twist reduced, strongly alternating tangles whose Conway sum is a link $L$. If $C(L)$ is the crosscap number of $L$, $T_L = |\beta_L| + |\beta'_L|$ and $k_L$ is the number of link components in $L$ we find that*

$$\left\lceil \tfrac{1}{6}T_L \right\rceil - k_L \leq C(L) \leq 2T_L + k_L + 8.$$

**Proof** This follows immediately from Theorem 3.5 and Lemma 3.11. $\qquad\square$

**Corollary 3.12** *Let $T_1$ and $T_2$ be twist reduced, nonsplittable, strongly alternating tangles whose Conway sum is a link $L$. Assume that $\mathrm{tw}(T_i) = \mathrm{tw}(K_{iN}) = \mathrm{tw}(K_{iD})$. If $C(L)$ is the crosscap number of $L$, $T_L = |\beta_L| + |\beta'_L|$ and $k_L$ is the number of link components in $L$, we have*

$$\left\lceil \tfrac{1}{6}T_L \right\rceil + 2 - k_L \leq C(L) \leq 2T_L + k_L + 8.$$

# 4   Generalizing to larger Conway sums of tangles

Our goal in this section is to generalize Theorem 1.2 to Conway sums of more than two tangles. A *Conway sum* of more than 2 tangles is a closure where we connect diagrams of the tangles $T_1, T_2, \ldots, T_l$ linearly west to east shown in Figure 2. As with the case of the sum of two tangles, if we let $L$ be our Conway sum and $S$ a spanning surface, cutting $S$ along a Conway sphere intersecting $T_i$ will result in a spanning surface for either $K_{iN}$ or $K_{iD}$, dictating the closure for the tangle. When we cut $L$, we position $l$ Conway spheres such that $\Sigma_i$ intersects $L$ at the directional strands of $T_i$. We note that $S^3 \setminus \bigcup_i \Sigma_i$ will not be a sphere, but we are concerned with the surfaces within the interior of each $\Sigma_i$. $S \setminus \cup_i \Sigma_i$ will be a collection of bands and tubes which we consider in the Euler characteristic change. This section will have similar results to the previous sections but with a factor for the number of tangles. We start with the following lemma which is a generalization of Lemma 2.5.

**Lemma 4.1** *Let $L$ be the Conway sum of the tangles $T_1, T_2, \ldots, T_l$, and $K_1, K_2, \ldots, K_l$ are closures of the tangles. Then if $k_L$ is the number of link components for $L$, and $k_1, k_2, \ldots, k_l$ the number of link components for each link, respectively, then $k_L = \sum_{i=1}^{l} k_i - l + \epsilon$ for $\epsilon = 0, 1, 2$.*

**Theorem 4.2** *Let $T_1, T_2, \ldots, T_l$ be nonsplittable, strongly alternating tangles, and let $L$ be the Conway sum that results from the $l$ tangles. If $K_i$ is the closure of $T_i$ resulting from cutting the crosscap realizing spanning surface for $L$ for all $i \in \{1, 2, \ldots, l\}$ and $K_{iN}$ is the numerator closure of $T_i$ and $K_{iD}$ the denominator closure, then we have*

$$\sum_{i=1}^{l} C(K_i) - l \leq C(L) \leq \min\left\{ \sum_{i=1}^{l} C(K_{iN}) + l, \sum_{i=1}^{l} C(K_{iD}) + 2 \right\}.$$

**Proof**   This proof will largely follow the work we did in Lemmas 2.14 and 2.4. We will start by considering the upper bound.

First we consider the case where we have $K_{iN}$ for all $i \in \{1, 2, \ldots, l\}$, and spanning surfaces $S_i$ for each $K_{iN}$ such that $C(S_i) = C(K_{iN})$. Unlike in Lemma 2.4 the NW and NE strands connect to different tangles and we will find the same for the SW and SE strands. Then the spanning surface resulting from the Conway sum will have northern and southern disks attached to each of the $S_i$ by a band as seen in Figure 9. Let this resulting surface be $S$.

By this construction $\chi(S) = \sum_{i=1}^{l} \chi(S_i) - 2l + 2$ where the $2l$ comes from the bands connecting each surface to the disks, and the 2 from the disks themselves. Then by Lemma 4.1 we see that $k_L = \sum_{i=1}^{l} k_{iN} - l + \epsilon$. Then

$$C(S) = 2 - \left( \sum_{i=1}^{l} \chi(S_i) - 2l + 2 \right) - \left( \sum_{i=1}^{l} k_{iN} - l + \epsilon \right) = \sum_{i=1}^{l} C(K_{iN}) + l - \epsilon.$$
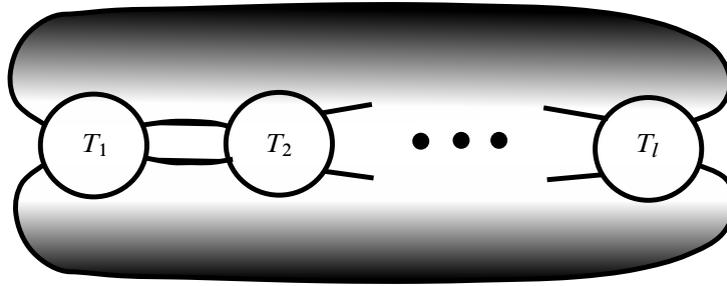
Figure 9: We see here that a surface would have to fill the shaded areas to connect the numerator closures of the tangles since the western and eastern boundaries of the $S_i$ connect to separate tangles.

Then we see the weakest upper bound is when $\epsilon = 0$, so

$$C(L) \le C(S) = \sum_{i=1}^{l} C(K_{iN}) + l.$$

Meanwhile the all denominator closure case will be similar to when $l = 2$. In particular $\chi(S) = \sum_{i=1}^{l} \chi(S_i) - l$ as we add bands to connect each of the $S_i$ to their neighboring surfaces. By Lemma 4.1 and similar computations to the numerator closure case we find $C(L) \le \sum_{i=1}^{l} C(K_{iD}) + 2$. We have a 2 instead of an $l$ as we added half the number of bands in constructing $S$. Then we take the minimum of the denominator and numerator bounds to find an upper bound for $C(L)$.

Now we consider the lower bound. Let $S$ be a spanning surface for $L$ such that $C(S) = C(L)$. Similar to Lemma 2.14 we will be considering the surfaces $S_1, S_2, \dots, S_l$ that result from cutting along the Conway spheres $\Sigma_i$. By a similar argument to the one for Lemma 2.12, $S$ can be chosen so that $\Sigma_i \cap S$ contains at most one closed curve for all $i$. Notice that if any of the steps in Lemma 2.12 were to disconnect the surface outside $\Sigma_i$, then for some other $\Sigma_j$, $i \ne j$, $T_j$ would span a disconnected surface which is a contradiction to Lemma 2.10. Then when we cut along the Conway spheres we see that we are at most cutting along two arcs and a closed curve.

We know from Lemma 4.1 that $k_L = \sum_{i=1}^{l} k_i - l + \epsilon$ for $\epsilon = 0, 1, 2$. If we have closed curves along a $\Sigma_i$, when we cut we will have to add disks to both resulting boundary components which increases the Euler characteristic by 2. For any surface resulting from cutting that is orientable we will have to add in a half twist band to make it nonorientable. Each such half twist band reduces the Euler characteristic by 1. Then $\sum_{i=1}^{l} \chi(S_i) = \chi(S) + t + c - b$ where $t = l$ or $t = 2l - 2$ depending on if the $K_i$ are the denominator or numerator closures, $c$ the number of closed curves on the $\Sigma_i$ which can be as large as $l$ and $b$ the number of twist bands added to make the $S_i$ nonorientable which also has maximum $l$. The value of $t$ arises from the bands which sit in $S^3 \setminus \bigcup_i \Sigma_i$.

Now we see that

$$\sum_{i=1}^{l} C(S_i) = 2l - \sum_{i=1}^{l} \chi(S_i) - \sum_{i=1}^{l} k_i = 2l - \chi(S) - t - c + b - k_L - l + \epsilon.$$

The weakest upper bound for $\sum_{i=1}^{l} C(S_i)$ is when $t$ and $c$ are minimal and $b$ and $\epsilon$ are maximal. So this will be when we do not have any closed curves and each of the resulting $S_i$ need half twist bands to make them nonorientable. So

$$\sum_{i=1}^{l} C(S_i) = 2l - \chi(S) - l + l - k_L - l + 2 = C(L) + l.$$

Then moving the $l$ to the other side we see that $\sum_{i=1}^{l} C(K_i) - l \leq C(L)$, showing the claim. $\qquad\square$

As in Section 3 we will now use Theorem 3.4 to find bounds for $C(L)$ in terms of $\mathrm{tw}(D(L))$ where $\mathrm{tw}(D(L))$ is the twist number for a diagram of $L$.

**Theorem 4.3** *Let* $T_1, T_2, \ldots, T_l$ *be nonsplittable, twist reduced, strongly alternating tangles and let* $D(L)$ *be the link diagram for the link* $L$ *which results from taking the Conway sum of the tangles. Let* $\mathrm{tw}(D(L))$ *denote the twist number of* $D(L)$ *and* $C(L)$ *the crosscap number for* $L$. *Then*

$$\left\lceil \tfrac{1}{3} \mathrm{tw}(D(L)) \right\rceil + 2 - k_L \leq C(L) \leq \mathrm{tw}(D(L)) + l + 2 - k_L.$$

**Proof** We start with the bounds from Theorem 4.2. From here we use Theorem 3.4 and Lemma 3.3 to get bounds on $C(L)$ with respect to the twist numbers of diagrams of $T_1, \ldots, T_l$. Combining these two statements we find

$$\sum_{i=1}^{l} \left\lceil \tfrac{1}{3}(\mathrm{tw}(T_i) - 2) \right\rceil + l - \sum_{i=1}^{l} k_i \leq C(L) \leq \sum_{i=1}^{l} \mathrm{tw}(T_i) + 2l + 2 - \sum_{i=1}^{l} k_i,$$

where $k_i$ is the number of link components for each $K_i$.

We know that the twist number of strongly alternating tangles is additive over a Conway sum by Lemma 3.2. So the only detail left to consider is the relationship between $k_L$ and $\sum_{i=1}^{l} k_i$. By Lemma 4.1 and a similar argument to that in Theorem 3.5 we find the claim:

$$\left\lceil \tfrac{1}{3}\big(\mathrm{tw}(D(L)) - 2l\big) \right\rceil + 2 - k_L \leq c(L) \leq \mathrm{tw}(D(L)) + l + 2 - k_L. \qquad\square$$

The final piece of our puzzle is to find bounds in terms of $T_L$. To do this we use Lemma 3.11 and Theorem 4.3 and the result follows.

**Theorem 4.4** *Let* $T_1, T_2, \ldots, T_l$ *be nonsplittable, twist reduced, strongly alternating tangles and let* $L$ *be the link which results from taking the Conway sum. Then let* $C(L)$ *be the crosscap number and* $k_L$ *the number of link components in* $L$. *Then*

$$\left\lceil \tfrac{1}{6}(T_L - 2l) \right\rceil + 2 - k_L \leq C(L) \leq 2T_L + l + 6 + k_L.$$

With an additional constraint on our tangles we find the following corollary:

**Corollary 4.5** *Let $T_1, T_2, \ldots, T_l$ be nonsplittable, twist reduced, strongly alternating tangles such that* $\mathrm{tw}(T_i) = \mathrm{tw}(D(K_{iN})) = \mathrm{tw}(D(K_{iD}))$ *for all* $i \in \{1, \ldots, l\}$. *Let* $L$ *be the link which results from taking the Conway sum,* $C(L)$ *the crosscap number, and* $k_L$ *the number of link components in* $L$. *Then*

$$\left\lceil \tfrac{1}{6} T_L \right\rceil + 2 - k_L \leq C(L) \leq 2T_L + l + 6 + k_L.$$

# 5  Families where $T_L$ and the crosscap number are independent

We begin by recalling the following theorem from [17] which gives linear bounds for the crosscap number of an alternating link in terms of $T_L$, where $T_L = |\beta_L| + |\beta'_L|$ and $\beta_L$ and $\beta'_L$ are the second and second-to-last coefficients of the Jones polynomial of $L$, respectively.

**Theorem 5.1** *Let* $L$ *be a nonsplit, prime alternating link with* $k$ *components and with crosscap number* $C(L)$. *Suppose that* $K$ *is not a* $(2, p)$ *torus link. We have*

$$\left\lceil \tfrac{1}{3} T_L \right\rceil + 2 - k \leq C(L) \leq T_L + 2 - k,$$

*where* $T_L$ *is as above. Furthermore, both bounds are sharp.*

In the previous sections we showed Theorem 5.1 generalizes to Conway sums of strongly alternating tangles. In this section we will show that Theorem 5.1 does not generalize to arbitrary knots.

**Theorem 1.4** (a)  *There exists a family of links for which* $T_L \leq 2$, *but* $C(L)$ *is arbitrarily large.*

(b)  *There exists a family of links for which* $C(L) \leq 3$, *but* $T_L$ *is arbitrarily large.*

## 5.1  Part (a) of Theorem 1.4

In this section we will consider the family of torus knots $T(p, q)$, where $q = j$ and $p = 2 + 2jk$ for odd $j > 1$ and all natural numbers $k$. This family will allow us to prove part (a) of Theorem 1.4. We start with the following definition from Teragaito [27].

**Definition 5.2** We define the value $N(p, q)$ from [27] for fractions $\frac{p}{q}$, where $p$ and $q$ are coprime, to begin write $\frac{p}{q}$ as a continued fraction

$$\frac{p}{q} = [a_0, a_1, a_2, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \frac{1}{a_n}}}},$$

where the $a_i$ are integers, $a_0 \geq 0$, $a_i > 0$ for $1 \leq i \leq n$, and $a_n > 1$. A continued fraction of this form is unique (see [14]). Now we recursively define $b_i$ as

$$b_0 = a_0, \qquad b_i = \begin{cases} a_i & \text{if } b_{i-1} \neq a_{i-1} \text{ or if } \sum_{j=0}^{i-1} b_j \text{ is odd,} \\ 0 & \text{if } b_{i-1} = a_{i-1} \text{ and } \sum_{j=0}^{i-1} b_j \text{ is even.} \end{cases}$$

Then, $N(p, q) = \frac{1}{2} \sum_{i=1}^{n} b_i$. We say a torus knot $K$ is even if the product of $p$ and $q$ is even and we say $K$ is odd otherwise. Using these definitions we can state Theorem 1.1 from [27].

**Theorem 5.3** *Let $K$ be the nontrivial torus knot of type $(p, q)$, where $p, q > 0$ and let $F$ be a nonorientable spanning surface of $K$ with $C(F) = C(K)$.*

  (i)  *If $K$ is even, then $C(K) = N(p, q)$ and the boundary slope of $F$ is $pq$.*

 (ii)  *If $K$ is odd, then $C(K) = N(pq - 1, p^2)$ (resp $N(pq + 1, p^2)$) and the boundary slope of $F$ is $pq - 1$ (resp $pq + 1$) if $xq \equiv -1 \pmod{p}$ has an even (resp odd) solution $x$ satisfying $0 < x < p$.*

We take advantage of (i) from Theorem 5.3 to both construct our family of torus links and prove Proposition 5.5 below. We also need the following lemma, which gives an explicit formula for the Jones polynomial of a torus knot originally given in Proposition 11.9 of [16], which allows us to calculate $T_L$ for torus knots.

**Lemma 5.4** *The Jones polynomial for a torus knot $T(p, q)$ is given by*

$$V(T(p, q)) = t^{(p-1)(q-1)/2} \frac{1 - t^{p+1} - t^{q+1} + t^{p+q}}{1 - t^2}.$$

**Proposition 5.5** *Let $L = T(p, q)$ be the family of torus knots where $q > 1$ is odd and $p = 2 + 2qk$ for $k \in \mathbb{N}$. Then $T_L \leq 2$ but $C(T(p, q))$ can be made arbitrarily large.*

**Proof** Let $q > 1$ be odd, and $p = 2 + 2qk$ where $k$ is a natural number. To show that for all such torus knots, $L = T(p, q)$, $C(L)$ does not have a universal upper bound with respect to $T_L$, we will show that as $k$ goes to $\infty$, $C(L)$ also goes to $\infty$, but $T_L \leq 2$. We start by computing the crosscap number of $T(p, q)$ using Theorem 5.3.

First we notice that

$$\frac{p}{q} = \frac{2 + 2qk}{q} = 2k + \frac{1}{1 + \frac{1}{2}}.$$

Then $A = [2k, 1, 2]$. Then by Definition 5.2 $B = [2k, 0, 2]$. Finally, as $pq$ is even,

$$C(L) = N(p, q) = \tfrac{1}{2}(2k + 0 + 2) = k + 1.$$

Then as $k \to \infty$, $C(L)$ also goes to $\infty$.

Next by Lemma 5.4 we know that

$$
\begin{aligned}
V(L) &= t^{(p-1)(q-1)/2} \frac{1 - t^{p+1} - t^{q+1} + t^{p+q}}{1 - t^2} \\
&= t^{((2+2qk)q-(2+2qk)-q+1)/2} \frac{1 - t^{2+2qk+1} - t^{q+1} + t^{2+2qk+q}}{1 - t^2} \\
&= t^{((2+2qk)q-(2+2qk)-q+1)/2} (-t^{2qk+q} - t^{2qk+q-2} - \cdots - t^{2+2qk+1} + t^{q-1} + \cdots + t^2 + 1).
\end{aligned}
$$

The last step arises from taking the polynomial division. Therefore, given our choices of $p$ and $q$ we see that $T_L \leq 2$. $\qquad\square$

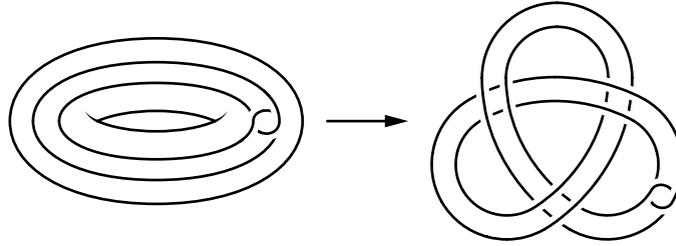Then Proposition 5.5 shows part (a) of Theorem 1.4.

Figure 10: Here we see the unknot with a clasp contained inside the torus, then we see the resulting *positive* Whitehead double with the blackboard framing when map the torus to the trefoil.

## 5.2 Part (b) of Theorem 1.4

In this section we will work to prove part (b) of Theorem 1.4 and the following theorem:

**Theorem 5.6** *There does not exist a universal linear lower bound on $C(L)$ for all links $L$ in terms of $T_L$.*

To prove this we will introduce a family of links for which $C(L)$ is uniformly bounded but $T_L$ can be made arbitrarily large. These links will be constructed by using the Whitehead double defined here:

**Definition 5.7** The *Whitehead double* of a knot $L$ is the satellite of the unknot clasped inside of the torus. We call it a *positive* Whitehead double if the clasp is as in Figure 10 and a *negative* Whitehead double if not.

The particular family is defined in this next theorem:

**Theorem 5.8** *Let $K_1, K_2, \ldots, K_n$ be alternating knots such that $\beta'_{K_i} \neq 0$. Then we let $K$ be the connect sum of $K_1, K_2, \ldots, K_n$ such that $K$ is alternating and $W_-(K)$ be the negative Whitehead double of $K$ using the blackboard framing. Then $C(L) \leq 3$ and $|\beta'_L| \geq n$.*

**Lemma 5.9** *If a link is B-adequate then the negative Whitehead double of the link using the blackboard framing is also B-adequate.*

A similar statement was proven in [3] as Proposition 7.1. They show it for the untwisted negative Whitehead double of a knot with nonnegative writhe. The writhe of the knot introduces extra twists into the diagram of the untwisted Whitehead double, which can interfere with adequacy around the clasp.

**Proof** We start by showing that the blackboard two-cabling will be B-adequate. This is shown by Lickorish in [20] for $n$-cablings. Let $D$ be a B-adequate diagram for our link and $D^2$ the two-cabling. Notice in $D^2$ there will be four copies of each crossing in $D$. Then when we have the all B-resolution state we will end up with four parallel strands instead of two as we did in $D$. If we were to have a one-edge loop, then two of the strands are part of the same state circle. But these state circles are copies of the state circles for $D$ so this would contradict that $D$ is adequate.
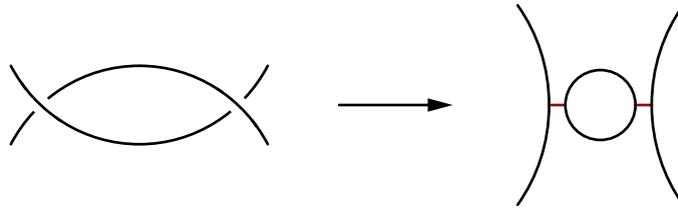
Figure 11: The result of the B-resolution on the clasp of a negative Whitehead double.

Now we want to look at the negative Whitehead double of $D$ using the blackboard framing. If we let the Whitehead double be $W_-(D)$ we will see that $G'(W_-(D))$ will be the same as $G'(D^2)$ but with an additional vertex and 2 new edges as we see in Figure 11. As resolving the clasp does not create a one-edge loop we see that $W_-(D)$ is B-adequate.                                                                          $\square$

Here we remind the reader that $G'_B(D(L))$ for a link diagram $D(L)$ is the reduced all B-state graph. We continue with the following lemma:

**Lemma 5.10** *If $W_-(D(L))$ is the negative Whitehead double of a B-adequate link diagram $D(L)$ using the blackboard framing, and $b(G)$ the first Betti number for a graph, then $b\big(G'_B\big(W_-(D(L))\big)\big) = b\big(G'_B(D(L))\big) + 1.$*

**Proof** Dasbach and Lin [6] showed in Lemma 2.5 that if $D^2$ is the two-cabling of a B-adequate link diagram then $b(\mathbb{G}'_B(D^2)) = b(\mathbb{G}'_B(D(L)))$. For a graph $G$, $b(G) = e - v + 1$, where $e$ is the number of edges and $v$ the number of vertices. In the reduced graph when we take the two-cabling every parallel copy of a state circle will also produce a new edge. Hence, the change in $v$ and $e$ will be the same between $\mathbb{G}'_B(D^2)$ and $\mathbb{G}'_B(D(L))$. Then when we move to $W_-(D(L))$ the clasp will add 2 edges and 1 vertex as we see in Figure 11. Then we see that $b\big(\mathbb{G}'_B\big(W_-(D(L))\big)\big) = b(\mathbb{G}'_B(D^2)) + 1 = b\big(\mathbb{G}'_B(D(L))\big) + 1.$ $\square$

The two previous lemmas allow us to see that the blackboard framing of the negative Whitehead double of an alternating link will be B-adequate. Also, we have a formula for the Betti number of the Whitehead double in relation to the first Betti number of the original link. The only remaining piece of the puzzle is to get from the first Betti number of the reduced B-state graph to the second-to-last coefficient of the Jones polynomial. This comes from the following result proven by Stoimenow in Proposition 3.1 of [26].

**Lemma 5.11** *If $D(L)$ is a B-adequate, connected diagram for a link, then in the representation of the Jones polynomial, $V(D(L))$, we have $\alpha'_{D(L)} = \pm 1$, $\alpha'_{D(L)}\beta'_{D(L)} \leq 0$, and*

$$|\beta'_{D(L)}| = e' - v' + 1 = b\big(G'_B(D(L))\big),$$

*where $\mathbb{G}'_B$ is the reduced all B-state graph and $e'$ and $v'$ are the number of edges and vertices of the graph $G'_B(D(L))$, respectively.*
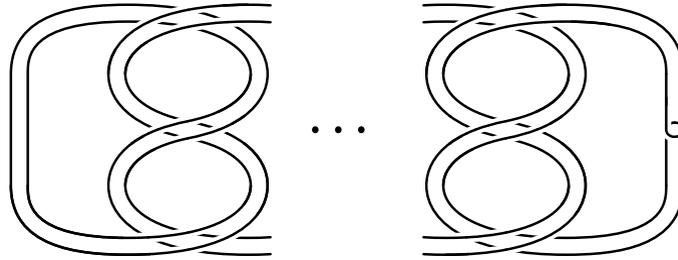
Figure 12: The negative Whitehead double of the connect sum of $m$ trefoil knots.

We now have the tools to prove Theorem 5.8. But first we show a more specific example of a family which satisfies Theorem 1.4(b).

**Proposition 5.12** *Let $W_-(K_m)$ be the negative Whitehead double using the blackboard framing of the connect sum of $m$ trefoils as in Figure 12. Then for all $m$, $C(W_-(K_m)) \le 3$ and $T_{W_-(K_m)}$ grows with $m$. Therefore, $T_{W_-(K_m)}$ can be made arbitrarily large across the family of knots.*

**Proof** The first part of the lemma is a direct result of [4] by Clark where he shows that $c(K) \le 2g(K)+1$ where $g(K)$ is the genus of the knot. For any Whitehead double we can find an oriented spanning surface with genus exactly one by taking the annulus with a double twisted band at the clasp. Then $C(W_-(K_m)) \le 3$ as $g(W_-(K_m)) = 1$.

Now we will compute $\beta'_{W_-(K_m)}$ by finding $\mathbb{G}'_B(W_-(K_m))$. By Lemma 5.11 we only need to find the number of vertices and edges as $W_-(K_m)$ is B-adequate. By Lemma 5.11 and the graph $\mathbb{G}'_B(W_-(K_m))$ shown in Figure 13

$$(8) \quad |\beta'_{W_-(K_m)}| = e\big(\mathbb{G}'_B(W_-(K_m))\big) - v\big(\mathbb{G}'_B(W_-(K_m))\big) + 1 = (5m+3) - (4m+3) + 1 = m.$$

Hence, we have shown that $T_{W_-(K_m)} \ge m$ for all $k$, proving the claim. □

Here we will introduce a more general family of knots for which Theorem 1.4(b) holds true:
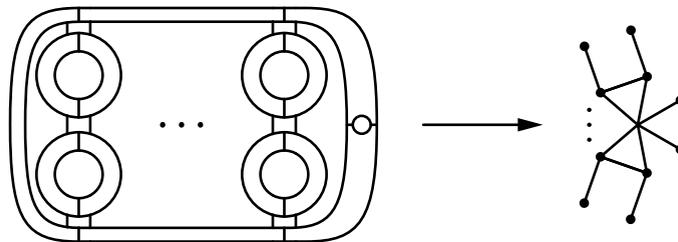


Figure 13: Left: the all B-state circle diagram. Right: the reduced state graph $G'_B(L)$. Notice the disjoint dots are not nodes for the graph but represent that we have $k$ copies of the subgraph on the left.

**Theorem 5.8** Let $K_1, K_2, \ldots, K_n$ be alternating knots such that $\beta'_{K_i} \neq 0$. Then let $K$ be the connect sum of $K_1, K_2, \ldots, K_n$ such that $K$ is alternating, and let $W_-(K)$ the negative Whitehead double of $K$ using the blackboard framing. Then $C(W_-(K)) \leq 3$ and $|\beta'_{W_-(K)}| \geq n$.

**Proof** As in Proposition 5.12 for a Whitehead double such as $W_-(K)$, $C(W_-(K)) \leq 3$. Now we will work to compute $T_{W_-(K)}$. From [20] we know that the Jones polynomial for $K$ will be the product of the Jones polynomials of the $K_i$. Then as all of the $K_i$ are alternating, $\alpha'_{K_i} = \pm 1$ so $\beta'_K = \sum_{i=1}^n \pm \beta'_{K_i}$. From Lemma 5.11 we know that $\alpha'_{K_i} \beta'_{K_i} \leq 0$ which tells us that the signs of $\alpha'_{K_i}$ and $\beta'_{K_i}$ do not match. If we let $m$ be the number of the $\alpha'_i$ which are negative, then we see that $\beta'_K = \sum_{i=1}^n (-1)^{m \pm 1} |\beta'_{K_i}|$. Hence, in our sum the signs match so $|\beta'_K| = \sum_{i=1}^n |\beta'_{K_i}|$. By our hypothesis $|\beta'_{K_i}| > 0$ for all $i$, hence $|\beta'_K| \geq n$.

By Lemma 5.9 we know that $W_-(K)$ will be B-adequate as $K$ is alternating and therefore B-adequate. Then by Lemmas 5.10 and 5.11 we see that $|\beta'_{W_-(K)}| = |\beta'_K| + 1$ and as $|\beta'_K|$ is at least as large as the number of knots in the connect sum so is $|\beta'_{W_-(K)}|$ and further $T_{W_-(K)}$. Then $T_{W_-(K)}$ will grow with $n$ showing that it is unbounded across the family. $\square$

Combining Proposition 5.12 and Theorem 5.8 shows Theorem 1.4(b).

# 6 Future directions

In Sections 2, 3, and 4, we generalized the work from [17] to bound the crosscap number of sums of strongly alternating tangles. Then in Section 5, introduced infinite families of knots for which their crosscap number and $T_L$ grow independently. The links we considered in Sections 2, 3, and 4 are all hyperbolic, meanwhile those that we constructed in Section 5 are not hyperbolic. This leads to the following question:

**Question 6.1** Does Theorem 4.4 generalize for all hyperbolic knots?

A first step for Question 6.1 would be to relax the requirement that the individual tangles be strongly alternating. At the time of writing, this seems reasonable for the first step of our proof, but the uncertainty arises in moving from bounds in terms of the crosscap numbers of individual tangles to the twist number. In particular, alternating is a requirement for our usage of Theorem 3.4. Another potential way to move forward with this question would be to look at adequate links in general, which will be studied in future work.

# References

[1] **C Adams**, **T Kindred**, *A classification of spanning surfaces for alternating links*, Algebr. Geom. Topol. 13 (2013) 2967–3007 MR

[2] **I Agol**, **J Hass**, **W Thurston**, *The computational complexity of knot genus and spanning area*, Trans. Amer. Math. Soc. 358 (2006) 3821–3850 MR

[3] **K L Baker**, **K Motegi**, **T Takata**, *The strong slope conjecture for twisted generalized Whitehead doubles*, Quantum Topol. 11 (2020) 545–608  MR

[4] **B E Clark**, *Crosscaps and knots*, Int. J. Math. Math. Sci. 1 (1978) 113–123  MR

[5] **R Crowell**, *Genus of alternating link types*, Ann. of Math. 69 (1959) 258–275  MR

[6] **O T Dasbach**, **X-S Lin**, *On the head and the tail of the colored Jones polynomial*, Compos. Math. 142 (2006) 1332–1342  MR

[7] **O T Dasbach**, **X-S Lin**, *A volumish theorem for the Jones polynomial of alternating knots*, Pacific J. Math. 231 (2007) 279–291  MR

[8] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Dehn filling, volume, and the Jones polynomial*, J. Differential Geom. 78 (2008) 429–464  MR

[9] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Symmetric links and Conway sums: volume and Jones polynomial*, Math. Res. Lett. 16 (2009) 233–253  MR

[10] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Cusp areas of Farey manifolds and applications to knot theory*, Int. Math. Res. Not. 2010 (2010) 4434–4497  MR

[11] **D Futer**, **E Kalfagianni**, **J Purcell**, *Guts of surfaces and the colored Jones polynomial*, Lecture Notes in Math. 2069, Springer (2013)  MR

[12] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Jones polynomials, volume and essential knot surfaces: a survey*, from "Knots in Poland, III: Part I", Banach Center Publ. 100, Polish Acad. Sci. Inst. Math., Warsaw (2014) 51–77  MR

[13] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Quasifuchsian state surfaces*, Trans. Amer. Math. Soc. 366 (2014) 4323–4343  MR

[14] **G H Hardy**, **E M Wright**, *An introduction to the theory of numbers*, 5th edition, Oxford Univ. Press (1979)  MR

[15] **J Hass**, **J C Lagarias**, **N Pippenger**, *The computational complexity of knot and link problems*, J. ACM 46 (1999) 185–211  MR

[16] **V F R Jones**, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. 126 (1987) 335–388  MR

[17] **E Kalfagianni**, **C R S Lee**, *Crosscap numbers and the Jones polynomial*, Adv. Math. 286 (2016) 308–337  MR

[18] **L H Kauffman**, *State models and the Jones polynomial*, Topology 26 (1987) 395–407  MR

[19] **T Kindred**, *Crosscap numbers of alternating knots via unknotting splices*, Internat. J. Math. 31 (2020) art. id. 2050057  MR

[20] **W B R Lickorish**, *An introduction to knot theory*, Graduate Texts in Math. 175, Springer (1997)  MR

[21] **W B R Lickorish**, **M B Thistlethwaite**, *Some links with nontrivial polynomials and their crossing-numbers*, Comment. Math. Helv. 63 (1988) 527–539  MR

[22] **C Livingston**, **A H Moore**, *KnotInfo: table of knot invariants*, electronic reference (2021)  Available at https://knotinfo.math.indiana.edu

[23] **H Murakami**, *An introduction to the volume conjecture*, from "Interactions between hyperbolic geometry, quantum topology and number theory", Contemp. Math. 541, Amer. Math. Soc., Providence, RI (2011) 1–40  MR

[24]  **K Murasugi**, *On the Alexander polynomial of the alternating knot*, Osaka Math. J. 10 (1958) 181–189  MR
Correction in 11 (1959) 95

[25]  **K Murasugi**, *Jones polynomials and classical conjectures in knot theory*, Topology 26 (1987) 187–194  MR

[26]  **A Stoimenow**, *Coefficients and non-triviality of the Jones polynomial*, J. Reine Angew. Math. 657 (2011) 1–55  MR

[27]  **M Teragaito**, *Crosscap numbers of torus knots*, Topology Appl. 138 (2004) 219–238  MR

*Department of Mathematics and Physics, Colorado State University Pueblo*
*Pueblo, CO, United States*

`rob.mcconkey@csupueblo.edu`

# Atiyah–Segal completion for the Hermitian K-theory of symplectic groups

JENS HORNBOSTEL

HERMAN ROHRBACH

MARCUS ZIBROWIUS

We study equivariant Hermitian K-theory for representations of symplectic groups, especially $\mathrm{SL}_2$. The results are used to establish an Atiyah–Segal completion theorem for Hermitian K-theory and symplectic groups.

## 1  Introduction

In [38], a completion theorem for Hermitian K-theory of schemes with trivial torus action is established. Let $X$ be a regular Noetherian separated scheme over $\mathrm{Spec}\left(\mathbb{Z}\left[\frac{1}{2}\right]\right)$ with a trivial action of a split torus $T \cong \mathbb{G}_m^t$. We let $\mathrm{IO} = \ker(\mathrm{GW}_0^{T,[0]}(X) \to \mathrm{GW}_0^{[0]}(X))$ be the Hermitian version of the augmentation ideal. Define

$$\mathrm{GW}^{T,[n]}(X)^{\wedge}_{\mathrm{IO}}$$

as the derived completion of the $\mathrm{GW}^{T,[0]}(X)$-module $\mathrm{GW}^{T,[n]}(X)$ with respect to IO; see [39] for details. Also define

$$\mathrm{GW}^{[n]}(\mathrm{B}T_X) = \lim_n \mathrm{GW}^{[n]}((\mathbb{P}_X^n)^{\times t}).$$

In [37, Proposition 8.2.2], it is shown that this definition is in fact an equivalence of GW-spectra of motivic spaces. In [38, Theorem 3.2.3], the following completion theorem for Hermitian K-theory is established.

**Theorem 1.0.1** *For all $i, n \in \mathbb{Z}$, the natural map*

$$\pi_i(\mathrm{GW}^{T,[n]}(X)^{\wedge}_{\mathrm{IO}}) \overset{\cong}{\longrightarrow} \mathrm{GW}_i^{[n]}(\mathrm{B}T_X)$$

*is an isomorphism.*

In the case of algebraic K-theory, the corresponding result for split tori $T$, due to Totaro [47] for $K_0$ and Knizel–Neshitov [26] for higher K-groups, was the first important step for establishing more general versions of an Atiyah–Segal style completion theorem for linear algebraic groups, due to Krishna [28], Tabuada–van den Bergh [45] and Carlsson–Joshua [12], and their proofs all rely on a reduction to the case of a split torus. The proof for algebraic K-theory and tori $T$ is easier than Rohrbach's theorem above on Hermitian K-theory because the classifying space $BT$ is a product of copies of $\mathbb{P}^\infty$, and algebraic K-theory is orientable in the sense of Levine–Morel and Panin. As of yet, there is no completion theorem for Hermitian K-theory in the case of general linear groups $\mathrm{GL}_n$, although [40] contains a partial result in this direction by computing the Hermitian K-theory of even-dimensional Grassmannians.

We restrict our attention to the Hermitian K-theory of symmetric forms and symplectic forms in degree zero, which is the $\mathbb{Z}/2\mathbb{Z}$-graded Grothendieck–Witt group $\mathrm{GW}^\pm$. The results can likely be extended to higher Hermitian K-theory over general base schemes using the machinery of derived completion as in [38], in line with the recent work of Tabuada–van den Bergh [45] and Carlsson–Joshua [12]. An extension of the completion theorem for Hermitian K-theory to schemes with nontrivial actions seems more difficult, as all the known proofs in algebraic K-theory rely on the equivariant localization theorem, whose analogue in Hermitian K-theory has not yet been established, and on a suitable geometric equivariant decomposition theorem. We refer to the last subsection for further details.

In a series of fundamental articles, Panin and Walter establish the theory of symplectic oriented cohomology theories on smooth algebraic varieties. In this setting, the one-dimensional torus $\mathrm{GL}_1$ is replaced by $\mathrm{Sp}_2 = \mathrm{SL}_2$. They also show that Hermitian K-theory is a symplectic orientable theory, which in particular implies the following computation of $B\mathrm{Sp}_{2n}$ over a base field $k$ (see [32; 34, Theorem 9.1]):

$$\mathrm{BO}^{*,*}(B\mathrm{Sp}_{2n}) \cong \mathrm{BO}^{*,*}(\mathrm{Spec}\,k)[\![b_1,\ldots,b_n]\!],$$

where the $b_i$ are the *Borel classes* of [34, Section 8] and the right-hand side is the ring of graded power series over $\mathrm{BO}^{*,*}(\mathrm{Spec}\,k)$; compare, eg, [27, Section 6.3]. From this point of view, one might argue that the conjectural Atiyah–Segal completion result for $\mathrm{Sp}_{2n}$ and Hermitian K-theory is the correct analogue of the completion theorem for $\mathrm{GL}_n$ and algebraic K-theory, and should provide computations involving free polynomial rings that are easier than Rohrbach's theorem above. We will show that this is indeed the case and prove the following theorem as Corollary 3.3.13 below.

**Theorem 1.0.2** *Let $k$ be a field of characteristic not two. There is a canonical map of $\mathrm{GW}^\pm(k)$-algebras*

$$\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2r})) \to \mathrm{GW}^\pm(B\mathrm{Sp}_{2r}),$$

*which exhibits $\mathrm{GW}^\pm(B\mathrm{Sp}_{2r})$ as the completion of $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2r}))$ with respect to $\mathrm{IO}_{\mathrm{Sp}_{2r}}$.*

The strategy for GW-theory consists in some sense in systematically replacing $\mathrm{GL}_1$ by $\mathrm{Sp}_2$, and more generally $\mathrm{GL}_n$ by $\mathrm{Sp}_{2n}$ in all steps of the proof for complex or algebraic K-theory.

This approach has already been used in computations for Chow–Witt groups, see, eg, [19], and is motivated in part by the fact that real realization of $\mathrm{Sp}_{2n}$ has $U(n)$ as a maximal compact subgroup.

The techniques used in this proof, which build on and slightly generalize those of Morel–Voevodsky and Panin–Walter, can be used to construct geometric classifying spaces for other groups $G$ and other Grothendieck topologies $\tau$. A general result about geometric descriptions of classifying spaces $\mathrm{B}_{\tau}G$ in motivic homotopy theory is given in Theorem 3.2.17.

Readers familiar with the proof of [4] might consider still another strategy for proving general Atiyah–Segal completion theorems for real and Hermitian K-theory, respectively: rather than $\mathrm{Sp}_{2n}$ and its subgroup $\mathrm{Sp}_2^{\times n}$, one should work with $\mathrm{GL}_n$ and its maximal split torus, but both equipped with a suitable involution that up to homotopy becomes complex conjugation under complex realization. There are indeed nice models for these in the algebro-geometric setting, which are explained in the introduction of Section 2. In particular, the results in Section 2.3 apply to these algebraic groups with involution. However, we are not yet able to perform all necessary equivariant GW-computations for these algebraic models of $\mathrm{GL}_n$ and split tori with involution, and the authors hope to return to this topic in the future.

This article, particularly Section 2, contains results about the *Real* (sometimes *quadratic* or *Hermitian*) representation ring of a large class of reductive groups with involution, such as $\mathrm{Sp}_{2n}$ and $\mathrm{GL}_n$ with involution. This builds on previous work of Calmès–Hornbostel [11] and Zibrowius [49] and is of independent interest. As an example of how some of the techniques of this paper extend to such Real groups, we have included in Section 3.4 a short construction of a classifying space for the multiplicative group $\mathbb{G}_m$ with a nontrivial involution.

## 2 The Grothendieck–Witt rings of some Real split reductive groups

Recall that the proof of Atiyah–Segal completion for topological KO-theory in [4] relies on the splitting of [3, Proposition (5.2)] for topological KR-theory for Real groups, ie for groups with an involution. We now discuss the algebraic analogue of such groups.

A *Real* group is a an algebraic group $G$ with an involution $\iota\colon G \to G$. This involution is assumed to be a group homomorphism. We denote by **Rep**$(G)$ the abelian category of finite-dimensional representations

of $G$, and equip it with a duality $\vee_\iota$ as follows. Recall first that, irrespective of any involution on $G$, associating with a representation $E$ of $G$ the dual representation $E^\vee := \mathrm{Hom}_G(E, k)$ defines a duality $\vee$ on $\mathbf{Rep}(G)$. More precisely, we obtain a category with duality $(\mathbf{Rep}(G), \vee, \eta)$, where $\eta$ denotes the canonical double-dual identification. Given the involution $\iota$ on $G$, we define the associated duality $\mathbf{Rep}(G)$ as the composition $\vee_\iota := \vee \circ \iota^*$, ie $E^{\vee_\iota} := (\iota^* E)^\vee$. We thus obtain a category with duality $\mathbf{Rep}(G, \iota) := (\mathbf{Rep}(G), \vee_\iota, \eta_\iota)$.

In this section, we compute the Grothendieck–Witt ring

$$\mathrm{GW}^\pm(\mathbf{Rep}(G, \iota)) := \mathrm{GW}^+(\mathbf{Rep}(G, \iota)) \oplus \mathrm{GW}^-(\mathbf{Rep}(G, \iota))$$

in cases when $G$ is split reductive, the involution $\iota$ restricts to a maximal torus $T$, and all irreducible representations are self-dual with respect to $\vee_\iota$.

**Example 2.0.1** Consider $(T, \mathrm{inv})$, where $T$ is a split torus of rank $r$ and inv is the involution given by $z \mapsto z^{-1}$. All representations of $T$ are symmetric with respect to $\vee_{\mathrm{inv}}$.

**Example 2.0.2** Consider $(\mathrm{GL}_n, \iota)$ with $\iota(A) := (A^{-1})^T$, an involution that restricts to the involution inv on the standard maximal torus of $\mathrm{GL}_n$. All representations of $\mathrm{GL}_n$ are symmetric with respect to $\vee_\iota$.

**Example 2.0.3** Consider $(\mathrm{Sp}_{2n}, \mathrm{id})$. All representations of $\mathrm{Sp}_{2n}$ are self-dual with respect to $\vee_{\mathrm{id}}$. Some are symmetric, some are antisymmetric (see Example 2.3.15 for more details).

## 2.1 Reductive groups — the setup

Let $k$ be a field. Let $(G, B, T)$ be a triple consisting of a connected split reductive group $G$ over $k$ with $T \subset B \subset G$, where $B$ is a Borel subgroup and $T$ is a maximal torus of rank $t$. We write $X^* := \mathrm{Hom}(T, \mathbb{G}_m)$ and $X_* := \mathrm{Hom}(\mathbb{G}_m, T)$ for the character lattice and the cocharacter lattice, respectively, both isomorphic to $\mathbb{Z}^t$, and $\langle -, - \rangle \colon X^* \times X_* \to \mathbb{Z}$ for the canonical pairing between them. Let $\Phi = \Phi(G, T)$ denote the associated set of roots, $\Phi^+ = \Phi(B, T)$ the set of positive roots associated with the Borel subgroup $B$ as in [13, Proposition 1.4.4], and $\Delta \subset \Phi^+$ the set of simple positive roots. We write $\alpha^{\mathrm{co}}$ for the coroot associated with a root $\alpha$.[1] Let $X^+ \subset X^*$ denote the cone of dominant characters determined by $\Phi^+$, also known as the (closed, integral) fundamental Weyl chamber [13, Equation (1.5.3)]. Explicitly, the relation between $\Phi^+$ and $X^+$ can be described as

$$(2.1.1) \qquad X^+ = \{x \in X^* \mid \langle x, \alpha^{\mathrm{co}} \rangle \geq 0 \text{ for all } \alpha \in \Delta\},$$

$$(2.1.2) \qquad \Phi^+ = \{\alpha \in \Phi \mid \langle x, \alpha^{\mathrm{co}} \rangle \geq 0 \text{ for all } x \in X^+\}.$$

The dominant characters parametrize the irreducible representations of $G$ [43, lemme 5]: for each $x \in X^+$, there is an irreducible $G$-representation $E_x$ of highest weight $x$, unique up to isomorphism. We fix one

---

[1] We deviate from the universally agreed notation $\alpha^\vee$ for $\alpha^{\mathrm{co}}$ to avoid any confusion with the duality $\vee$ on $\mathbf{Rep}(G)$.

| here | [13] | [43] | |
|------|------|------|---|
| $X^*$ | $X(T)$ | $M$ | character lattice |
| $X^+$ | $C$ | $P$ | dominant cone/ [fundamental] Weyl chamber |
| $\Phi$ | $\Phi$ | $R$ | set of roots |
| $\Phi^+$ | $\Phi^+$ | $R^+$ | set of positive roots |
| $\Delta$ | $\Delta$ | — | set of simple roots |
| $X_*$ | $X^\vee(T)$ | — | cocharacter lattice |

Table 1: Translation between notation used here and notation in some of our references.

such representation for each $x \in X^+$. When $G$ is simply connected, $X^+ \cong \mathbb{N}_0^n$, with basis given by fundamental weights $\omega_1, \ldots, \omega_n$. Following [43, section 3.6], we define a partial order on $X^*$ as follows:

**Definition 2.1.1** For $x, y \in X^*$, we write $x \leq y$ if and only if $y - x$ can be written as a $\mathbb{Z}$-linear combination of elements of $\Phi^+$ with nonnegative coefficients.

Other definitions abound in the literature. The ordering defined here is slightly different from both orderings defined in [9, Chapter VI, Definition 2.2].

Finally, let $W = W(G, T) = W(\Phi)$ be the associated Weyl group. We refer to any translate $wX^+ \subset X^*$ as a Weyl chamber. The Weyl group acts simply transitively on the set of Weyl chambers. For the convenience of the reader, a partial translation of the notation used here and the notation used in some of the references is provided by Table 1.

## 2.2 The representation ring

**Lemma 2.2.1** *For dominant weights $x, y \in X^+$,*

$$E_x \cdot E_y = E_{x+y} + \sum_z E_z$$

*in $\mathrm{K}_0(\mathbf{Rep}(G))$, where the sum is over a finite number of $z \in X^+$ such that $z < x + y$.*

**Proof** By [43, théorème 4], there is an injective ring morphism

$$\mathrm{ch}_G : \mathrm{K}_0(\mathbf{Rep}(G)) \to \mathbb{Z}[X^*]$$

with image $\mathbb{Z}[X^*]^W$. By [43, lemme 5], $\mathrm{ch}_G(E_x) = e^x + \sum_i e^{x_i}$ for a finite number of $x_i \in X^*$ such that $x_i < x$, and similarly for $\mathrm{ch}_G(E_y)$ and $\mathrm{ch}_G(E_{x+y})$. It follows that

$$\mathrm{ch}_G(E_x)\,\mathrm{ch}_G(E_y) = e^{x+y} + (\text{terms smaller than } x + y),$$

where by "terms smaller than $x + y$" we mean a $\mathbb{Z}$-linear combination of terms $e^z$ with $z \in X^*$ such that $z < x + y$. On the other hand, $\mathrm{ch}_G(E_x)\,\mathrm{ch}_G(E_y) = \mathrm{ch}_G(E_x \otimes E_y)$, and $E_x \otimes E_y$ can be decomposed into

a sum of irreducible representations $E_1, \ldots, E_N$. Writing $\mathrm{ch}_G(E_k) = e^{w_k} + (\text{terms smaller than} w_k)$, we find that

$$\mathrm{ch}_G(E_x)\,\mathrm{ch}_G(E_y) = \sum_{k=1}^{N} (e^{w_k} + \text{terms smaller than } w_k).$$

It follows by comparison that $w_k \leq x + y$ for each $k$, and that $w_k = x + y$ for exactly one $k$. $\qquad\square$

**Lemma 2.2.2** *There is no $z \in X^+$ such that $z < 0$.*

**Proof** Let $\Delta \subset \Phi^+$ denote the set of simple roots. If $z < 0$ then $z = \sum_{\alpha \in \Delta} a_\alpha \alpha$ with $a_\alpha \leq 0$ for all $\alpha \in \Delta$ and $a_\alpha < 0$ for at least one $\alpha$. In particular, $z$ lies in the span of $\Phi$, so we can pass from the root system $\Phi \subset X^* \otimes_{\mathbb{Z}} \mathbb{R}$ to the *reduced* root system $\Phi \subset \Phi \otimes_{\mathbb{Z}} \mathbb{R}$, to which the results of [7, chapitre VI, section 1, n°10] apply. Let $\sigma$ be the half-sum of all coroots $\alpha^{\mathrm{co}}$ corresponding to roots $\alpha \in \Phi^+$. Then by [7, chapitre VI, section 1, n°10, corollaire], $\langle \alpha, \sigma \rangle = \sum_{\alpha \in \Delta} a_\alpha$. This number is smaller than zero by assumption, so $z \notin X^+$. $\qquad\square$

**Proposition 2.2.3** *Suppose $X^+$ can be split into a direct sum $X^+ \cong \mathbb{N}_0^n \oplus \mathbb{Z}^m$. Pick elements $\omega_1, \ldots, \omega_n \in X^+$ and $\zeta_1, \ldots, \zeta_m \in X^+$ corresponding to an $\mathbb{N}_0$-basis of $\mathbb{N}_0^n$ and a $\mathbb{Z}$-basis of $\mathbb{Z}^m$, respectively. There is a unique ring homomorphism*

$$\mathbb{Z}[w_1, \ldots, w_n, z_1^{\pm 1}, \ldots, z_m^{\pm 1}] \to \mathrm{K}_0(\mathbf{Rep}(G))$$

*taking each $w_i$ to $E_{\omega_i}$ and each $z_j$ to $E_{z_j}$, and this ring homomorphism is an isomorphism.*

**Example 2.2.4** For simply connected $G$, $X^+ \cong \mathbb{N}_0^n$ with basis $\omega_1, \ldots, \omega_n$ the fundamental weights. The proposition shows that $\mathrm{K}_0(\mathbf{Rep}(G))$ is a polynomial ring over $\mathbb{Z}$ on generators $E_{\omega_1}, \ldots, E_{\omega_n}$.

**Example 2.2.5** For $G = T$ a split torus of rank $m$, $X^+ = X^* \cong \mathbb{Z}^m$. Pick a $\mathbb{Z}$-basis $\zeta_1, \ldots, \zeta_m$ of $X^*$. The proposition shows that $\mathrm{K}_0(\mathbf{Rep}(T))$ is a ring of Laurent polynomials with the one-dimensional representations $E_{\zeta_i}$ as generators.

**Proof of Proposition 2.2.3** It suffices to show that there is a well-defined ring map

$$f : \frac{\mathbb{Z}[w_1, \ldots, w_n, z_1, \ldots, z_m, z_1', \ldots, z_m']}{(z_j z_j' - 1 \mid j = 1, \ldots, m)} \to \mathrm{K}_0(\mathbf{Rep}(G))$$

that sends $w_i$ to $E_{\omega_i}$, $z_j$ to $E_{\zeta_j}$ and $z_j' \to E_{-\zeta_j}$, and that this map is an isomorphism.

To see that $f$ is well defined, note that by Lemma 2.2.1

$$f(z_j) f(z_j') = E_{\zeta_j} \cdot E_{-\zeta_j} = E_0 + \sum_z E_z,$$

where the sum is over certain $z \in X^+$ such that $z < 0$. By Lemma 2.2.2, this is the empty sum. So $f(z_j) f(z_j') = E_0 = 1$, as required.

To see that $f$ is an isomorphism, we associate the Laurent monomial $M^x := w_1^{a_1} \cdots w_n^{a_n} \cdot z_1^{b_1} \cdots z_m^{b_m}$ with the element $x = \sum_i a_i \omega_i + \sum_j b_j \zeta_j \in X^+$. This defines a bijection between a $\mathbb{Z}$-basis of the Laurent ring and $X^+$. Under $f$, the basis element $M^x$ maps to $\prod_i (E_{\omega_i})^{a_i} \prod_j (E_{\zeta_j})^{b_j}$. By Lemma 2.2.1, we can rewrite this element as

$$(2.2.1) \qquad f(M^x) = E_x + \sum_z E_z,$$

for certain $z \in X^+$ with $z < x$. Using [7, chapitre VI, section 3, n°4, lemme 4] as in the proof of [43, lemme 6], we deduce that the elements $f(M^x)$ for $x \in X^+$ form a $\mathbb{Z}$-basis of $\mathbb{Z}[X^*]^W$. $\qquad\square$

**Remark 2.2.6** (alternative generators) More generally, under the assumptions of Proposition 2.2.3, we can choose generators for $K_0(\mathbf{Rep}(G))$ as follows. Take $\omega_i$ and $\zeta_j \in X^+$ as before. For $i \in \{1, \ldots, n\}$, pick classes $e_{\omega_i} \in K_0(\mathbf{Rep}(G))$ such that

$$e_{\omega_i} = E_{\omega_i} + \text{(smaller terms)},$$

in $K_0(\mathbf{Rep}(G))$, where (smaller terms) refers to a $\mathbb{Z}$-linear combination of irreducible representations $E_x$ indexed by finitely many $x \in X^+$ with $x < \omega_i$. Then again we have a ring isomorphism

$$\mathbb{Z}[w_1, \ldots, w_n, z_1^{\pm 1}, \ldots, z_m^{\pm 1}] \xrightarrow{\cong} K_0(\mathbf{Rep}(G))$$

taking each $w_i$ to $e_{\omega_i}$ and each $z_j$ to $E_{\zeta_j}$. Indeed, this follows with the same proof, as the key identity (2.2.1) still holds for these more general generators.

## 2.3 The ±-symmetric representation ring

From now on, we always assume $\text{char}(k) \neq 2$.

We now assume that $G$ is equipped with an involution $\iota$ (possibly trivial) *that restricts to the chosen maximal torus $T$*, and study the associated duality $\vee_\iota$ on $\mathbf{Rep}(G)$ introduced at the beginning of the section.

The following lemmas describe the highest weight of the $\vee_\iota$-dual of an irreducible representation. As we have assumed that $\iota$ restricts to $T$, we have induced involutions $\iota^*$ and $\iota_*$ on the character lattice $X^*$ and the cocharacter lattice $X_*$, respectively, compatible with the canonical pairing $\langle -, - \rangle$. For example, the involutions $\text{inv}^*$ and $\text{inv}_*$ induced by the involution from Example 2.0.1 are given by $-\text{id}$. Interpreting characters as one-dimensional representations, we see that the dualities $\vee$ and $\vee_\iota$ also define involutions on $X^*$. Explicitly, $\vee = -\text{id}$ on $X^*$, hence $\vee_\iota = -\iota^*$ on $X^*$.

**Lemma 2.3.1** *There is a unique involution $w_\iota \in W$ whose action on Weyl chambers agrees with this action of $\vee_\iota = -\iota^*$ on the Weyl chambers. In particular, $-\omega_\iota \iota^*$ defines an involution of the fundamental Weyl chamber $X^+$.*

**Example 2.3.2** If $\iota|_T = \mathrm{id}$, then $\vee_\iota = -\mathrm{id}$ on $X^*$, and $w_\iota = w_0$, the longest element of the Weyl group. If $\iota|_T = \mathrm{inv}$ is the inversion from Example 2.0.1, then $\vee_\iota = \mathrm{id}$ on $X^*$, and $w_\iota = 1$.

**Proof** The case when $\iota = \mathrm{id}$ is standard. For the general case, note that both involutions $\iota^*$ and $\vee$ on $X^*$ send roots to roots. (For $\iota^*$, this follows from $\iota$ is a group homomorphism compatible with the inclusion $T \subset G$; see [13, above Theorem 1.3.15].) The claim then follows from the fact that the Weyl group acts simply transitively on the set of Weyl chambers. $\qquad\square$

**Lemma 2.3.3** *Given an irreducible $G$-representation $E_x$ with highest weight $x$, the dual representation $(\iota^* E_x)^\vee$ has highest weight $-w_\iota \iota^* x$, where $w_\iota$ is as in the previous lemma.*

**Proof** Let $\Omega_E \subset X^*$ be the set of weights of $E := E_x$. As $\vee_\iota = -\iota^*$ on $X^*$, we have $\Omega_{(\iota^* E)^\vee} = -\iota^* \Omega_E$. Now consider an arbitrary weight $z \in \Omega_{(\iota^* E)^\vee}$. As $\Omega_{(\iota^* E)^\vee}$ is invariant under the action of the Weyl group, we also have $\omega_\iota z \in \Omega_{(\iota^* E)^\vee}$. So $\omega_\iota z \in (-\iota^* \Omega_E)$. As $x$ is the highest weight of $E$, this means that $\omega_\iota z = -\iota^*(x - \sum_{b \in \Delta} m_b b)$ for certain $m_b \in \mathbb{Z}_{\geq 0}$. So $z = -\omega_\iota \iota^* x + \sum_{b \in w_\iota \iota_*(\Delta)} m'_b b$ for certain $m'_b \in \mathbb{Z}_{\geq 0}$. As noted in Lemma 2.3.1, $\omega_\iota \iota^*(X^+) = -X^+$. In view of (2.1.2), this implies that $w_\iota \iota_*(\Delta) \subseteq -\Phi^+$. So $z = -\omega_\iota \iota^* x - \sum_{b \in \Phi^+} m''_b b$, for certain $m''_b \in \mathbb{Z}_{\geq 0}$. It follows that $-w_\iota \iota^* x$ is the highest weight of $(\iota^* E)^\vee$. $\qquad\square$

**Definition 2.3.4** A dominant character $x \in X^+$ is $(\vee_\iota$-$)$*self-dual* if $x = -w_\iota \iota^* x$, where $w_\iota$ is as in Lemma 2.3.1.

**Definition 2.3.5** We call a representation $E$ of $G$ *self-dual* if there exists an isomorphism $\phi: E \to E^{\vee_\iota}$. We call a self-dual representation $E$ *symmetric* or *antisymmetric* if $\phi$ can be chosen to be symmetric or antisymmetric, respectively.

By Lemma 2.3.3, the irreducible representation with highest weight $x$ is $\vee_\iota$-self-dual if and only if $x$ is $\vee_\iota$-self-dual. In general, a representation may be both symmetric and antisymmetric. However, in our setting a self-dual *irreducible* representation is always either symmetric or antisymmetric, but never both:

**Lemma 2.3.6** *Let $E$ be an irreducible representation of a split reductive group $G$. If there exists an isomorphism $\phi: E \to (\iota^* E)^\vee$, then $\phi$ is either symmetric or antisymmetric, and any isomorphism $E \to (\iota^* E)^\vee$ is a multiple of $\phi$ by an (invertible) scalar.*

**Proof** This is essentially [11, Corollary 2.3], which relies on [11, Lemma 1.21]. Note that [11, Section 1] is phrased in the generality of abelian categories with duality and hence applies to the category of $G$-representations regardless of our choice of duality on this category. $\qquad\square$

In light of this lemma, we make the following definition.

**Definition 2.3.7** The *sign* $s(x)$ of a dominant character $x \in X^+$ is an element of $\{-1, 0, 1\}$, which is 1 if $E_x$ is symmetric, $-1$ if $E_x$ is antisymmetric, and 0 if $E_x$ is not $\vee_\iota$-self-dual.

For an abelian category $\mathcal{A}$ with duality, ie with a fixed involution $\vee$ and double-dual identification $\omega$, we write $\mathrm{GW}^+(\mathcal{A}, \vee, \omega)$ and $\mathrm{GW}^-(\mathcal{A}, \vee, \omega)$ for the Grothendieck–Witt groups of symmetric and antisymmetric forms over $\mathcal{A}$, respectively. Taking $\mathcal{A}$ to be the category of finite-dimensional $k$-vector spaces with its usual duality, this construction yields the usual groups $\mathrm{GW}^+(k)$ and $\mathrm{GW}^-(k)$ of symmetric and antisymmetric bilinear forms over $k$. The tensor product furnishes us with a $\mathbb{Z}/2$-graded ring structure on the direct sum $\mathrm{GW}^\pm(k) := \mathrm{GW}^+(k) \oplus \mathrm{GW}^-(k)$. As it is convenient to think of the grading group $\mathbb{Z}/2$ multiplicatively as $\{\pm 1\}$, we will refer to this grading as a $\pm$-grading in all that follows. Note that $\mathrm{GW}^-(k)$ does not depend on the field $k$. Explicitly, $\mathrm{GW}^-(k) = \mathbb{Z} \cdot H^-$ as a group, and

$$(2.3.1) \qquad\qquad \mathrm{GW}^\pm(k) = \frac{\mathrm{GW}^+(k)[H^-]}{((H^-)^2 - 2H^+)}$$

as a ring, where $H^\pm \in \mathrm{GW}^\pm(k)$ are the respective hyperbolic planes.

More generally, for the category with duality $(\mathbf{Rep}(G), \iota)$ introduced at the beginning of this section, we obtain a $\pm$-graded algebra $\mathrm{GW}^\pm(\mathbf{Rep}\,G, \iota)$ over the $\pm$-graded ring $\mathrm{GW}^\pm(k)$. The following theorem is an analogue of [11, Theorem 2.10] for Grothendieck–Witt theory, under the assumption that all characters are self-dual so as to eliminate the existence of hyperbolic elements.

**Proposition 2.3.8** *Assume that all $x \in X^+$ are $\vee_\iota$-self-dual. Choose an isomorphism $\phi_x : E_x \to (\iota^* E_x)^\vee$ for each $x \in X^+$. Then the classes $[E_x, \phi_x]$ form a basis of the $\mathrm{GW}^\pm(k)$-module $\mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$.*

**Proof** The proof is the same as for [11, Theorem 2.10], noting that the obstruction from [11, Remark 2.11] has been removed with the assumption that all dominant characters are self-dual. The main ingredients are [11, Corollary 1.14], which provides additive decompositions of both $\mathrm{GW}^+(\mathbf{Rep}(G))$ and $\mathrm{GW}^-(\mathbf{Rep}(G))$, and [11, Corollary 1.38], which identifies the summands of $\mathrm{GW}^+(\mathbf{Rep}(G)$ corresponding to symmetric characters and the summands of $\mathrm{GW}^-(\mathbf{Rep}(G))$ corresponding to antisymmetric characters. For a full proof, [11, Corollary 1.38] and the preceding [11, Proposition 1.37] need to be mildly generalized to include all signs. (For example, strictly speaking the identification of $\mathrm{GW}(\mathcal{A}_i)$ with $\mathrm{GW}_-(\mathcal{A}_\mathbb{1})$ as a $\mathrm{GW}(\mathcal{A}_\mathbb{1})$-module is missing from [11, Corollary 1.38 (i)], as the tensor unit $\mathbb{1}$ is assumed to be symmetric, not just $\delta$-symmetric, throughout [11, Section 1.4].) $\qquad\square$

**Remark 2.3.9** Even in the presence of non-self-dual characters, it is easy to describe the $\mathrm{GW}^\pm(k)$-module structure of $\mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$. Let $H^+(E)$ and $H^-(E)$ denote the symmetric and antisymmetric hyperbolic space associated with a representation $E$, respectively. The results quoted from [11] in the proof above show that every pair of non-self-dual dominant characters $(z, -w_\iota z)$ contributes a copy of $\mathbb{Z}$ generated by $H^+(E_z)$ to $\mathrm{GW}^+(\mathbf{Rep}(G, \iota))$ and a copy of $\mathbb{Z}$ generated by $H^-(E_z)$ to $\mathrm{GW}^-(\mathbf{Rep}(G, \iota))$. However, we want to concentrate on the case when $\mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$ is a *free* $\mathrm{GW}^\pm(k)$-module here.

**Lemma 2.3.10** *Consider the duality $\vee_\iota$ on **Rep** $G$. Suppose $(V, \phi)$ is an $\epsilon$-symmetric representation of $G$ such that $V = E_x +$ (smaller terms) in $\mathrm{K}_0(\mathbf{Rep}(G))$, for some $x \in X^+$ and $\epsilon \in \{\pm\}$. Then there exists an $\epsilon$-symmetric isomorphism $\phi_x$ on $E_x$ such that*

$$(V, \phi) = (E_x, \phi_x) + \text{(smaller terms)}$$

*in $\mathrm{GW}^\epsilon(\mathbf{Rep}(G, \iota))$, where (smaller terms) refers to a $\mathrm{GW}(k)$-linear combination of terms $T_z$ indexed by finitely many $z \in X^+$ with $z < x$ such that, for each $z$, $T_z$ is either of the form $(E_z, \phi_z)$ (in case $z$ is $\epsilon$-symmetric) or of the form $H^\epsilon(E_z)$ (in case $z$ is not).*

**Proof** Using the decomposition of $\mathrm{GW}^\epsilon(\mathbf{Rep}(G, \iota))$ of [11, Corollaries 1.14 and 1.38] as in the proof of Proposition 2.3.8 above, we can write $(V, \phi)$ as a $\mathrm{GW}(k)$-linear combination of elements $(E_z, \phi_z)$ with $z \in X^+$ such that $\phi_z$ is $\epsilon$-symmetric, and of elements $H^\epsilon(E_v)$ with $v \in X^+$ such that $E_v$ is not $\epsilon$-self-dual:

$$(V, \phi) = \sum_{\substack{z:s(z)=\epsilon}} \alpha_z \cdot (E_z, \phi_z) + \sum_{\substack{v:s(v)=-\epsilon \\ \text{or } s(v)=0}} a_v \cdot H^\epsilon(E_v).$$

The coefficients $a_v$ may be taken in $\mathbb{Z}$, as $a \cdot H^\epsilon(b) = H^{\delta\epsilon}(F(a)b)$ for any $a \in \mathrm{GW}^\delta(\mathbf{Rep}(G, \iota))$ and any $b \in \mathrm{K}_0(\mathbf{Rep}(G))$. By [11, Remark 1.15], we even know that the coefficients $\alpha_z$ of $(E_z, \phi_z)$ can be chosen to be actual symmetric forms over $k$ of positive rank, and that the coefficients $a_v \in \mathbb{Z}$ of $H^\epsilon(E_v)$ are positive. Applying the forgetful map, we thus find that

$$V = \sum_z (\mathrm{rank}\, \alpha_z) E_z + \sum_v 2 a_v E_v$$

in $\mathrm{K}_0(\mathbf{Rep}(G))$. As the irreducible representations $E_x$ with $x$ ranging over $X^+$ form a $\mathbb{Z}$-basis of $\mathrm{K}_0(\mathbf{Rep}(G))$, we deduce by comparing this decomposition with the given decomposition of $(V, \phi)$ that one of the $z$'s in the first sum must be equal to $x$, with $\alpha_x$ of rank one, and that for all other $z$'s and all $v$'s we have $z < x$ and $v < x$, respectively. $\qquad\square$

The following is a careful restatement of [11, Lemma 2.14]:

**Lemma 2.3.11** *For any two self-dual $x, y \in X^+$, the sum $x + y \in X^+$ is also self-dual. Moreover, given $\epsilon_x$- and $\epsilon_y$-symmetric isomorphisms $\phi_x : E_x \to (\iota^* E_x)^\vee$ and $\phi_y : E_y \to (\iota^* E_y)^\vee$, respectively, there exists an $\epsilon_x \epsilon_y$-symmetric isomorphism $\phi : E_{x+y} \to (\iota^* E_{x+y})^\vee$ such that*

$$(E_x, \phi_x) \cdot (E_y, \phi_y) = (E_{x+y}, \phi) + \text{(smaller terms)},$$

*in $\mathrm{GW}^{\epsilon_x \epsilon_y}(\mathbf{Rep}(G, \iota))$, where (smaller terms) is to be read as in Lemma 2.3.10.*

**Proof** The claim that $x + y$ is self-dual is immediate from the definitions. Also, $(E_x, \phi_x) \otimes (E_y, \phi_y)$ is an $\epsilon_x \epsilon_y$-symmetric representation and hence defines an element of $\mathrm{GW}^{\epsilon_x \epsilon_y}(\mathbf{Rep}(G, \iota))$. From Lemma 2.2.1, we know that $E_x \cdot E_y = E_{x+y} +$ (smaller terms), so we can apply Lemma 2.3.10. $\qquad\square$

The following result should be compared to [11, Theorem 2.16]. It is an analogue, with more restrictive hypotheses, of the description of the usual representation ring provided in Proposition 2.2.3.

**Proposition 2.3.12** *Suppose that all $x \in X^+$ are $\vee_\iota$-self-dual. Assume in addition that $X^+$ can be split into a direct sum $X^+ \cong \mathbb{N}_0^n \oplus \mathbb{Z}^m$. Pick elements $\omega_1, \ldots, \omega_n \in X^+$ and $\zeta_1, \ldots, \zeta_m \in X^+$ corresponding to an $\mathbb{N}_0$-basis of $\mathbb{N}_0^n$ and a $\mathbb{Z}$-basis of $\mathbb{Z}^m$, respectively. Pick (anti)symmetric isomorphisms $\phi_{\omega_i}$ and $\phi_{\zeta_j}$ for each of the representations $E_{\omega_i}$ and $E_{\zeta_j}$, respectively. Consider the ring $\mathrm{GW}^\pm(k)[w_1, \ldots, w_n, z_1^{\pm 1}, \ldots, z_m^{\pm 1}]$ as a graded $\mathrm{GW}^\pm(k)$-algebra with generators $w_i$ and $z_j$ in degrees $s(w_i)$ and $s(z_j)$, respectively. Then there is a unique graded $\mathrm{GW}^\pm(k)$-algebra homomorphism*

$$f : \mathrm{GW}^\pm(k)[w_1, \ldots, w_n, z_1^{\pm 1}, \ldots, z_m^{\pm 1}] \to \mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$$

*taking each $w_i$ to $(E_{\omega_i}, \phi_{\omega_i})$ and each $z_j$ to $(E_{\zeta_j}, \phi_{\zeta_j})$, and this homomorphism is an isomorphism.*

For the one-dimensional representations $E_{\zeta_j}$ we have $E_{\zeta_j}^\vee \cong E_{-\zeta_j}$ with respect to the usual duality $\vee = \vee_{\mathrm{id}}$. So, for $m > 0$, the assumption that each $E_\zeta$ be self-dual with respect to $\vee_\iota$ cannot hold with respect to the usual duality $\vee$.

**Example 2.3.13** Take $(G, \iota) = (T, \mathrm{inv})$, a split torus of rank $m$ equipped with the inversion from Example 2.0.1. In this case, $X^+ = X^* \cong \mathbb{Z}^m$, and all characters are self-dual symmetric. The proposition shows that $\mathrm{GW}^\pm(\mathbf{Rep}(T), \mathrm{inv})$ is a ring of Laurent polynomials over $\mathrm{GW}^\pm(k)$ with one-dimensional symmetric representations as generators.

**Example 2.3.14** Take $(G, \iota) = (\mathrm{GL}_n, \iota)$, the general linear group equipped with the involution from Example 2.0.2. The fundamental representations of $\mathrm{GL}_n$ are given by the exterior powers of the standard $n$-dimensional representation $V$ [24, Part II, Section 2.15]. More precisely, in the notation from above, $E_{\omega_i} = \Lambda^i(V)$ for $i = 1, \ldots, n-1$ and $E_{\zeta_1} = \Lambda^n(V)$. As noted in Example 2.0.2, we can choose a nondegenerate $\vee_\iota$-symmetric form $\phi$ on $V$, so that we obtain elements $\lambda_i := (\Lambda^i(V), \Lambda^i(\phi))$ in $\mathrm{GW}^+(\mathbf{Rep}(\mathrm{GL}_n, \iota))$. The above proposition then shows that

$$\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{GL}_n, \iota)) \cong \mathrm{GW}^\pm(k)[\lambda_1, \ldots, \lambda_{n-1}, \lambda_n, \lambda_n^{-1}].$$

**Example 2.3.15** Take $(G, \iota) = (\mathrm{Sp}_{2n}, \mathrm{id})$, as in Example 2.0.3. The irreducible representation $E_{\omega_i}$ is symmetric for even $i$ and antisymmetric for odd $i$ [8, Chapter VIII, Table 1], so we can choose a nondegenerate $(-1)^i$-symmetric form $\phi_i$ on each $E_{\omega_i}$. The proposition then shows that $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2n}), \mathrm{id})$ is a polynomial algebra over $\mathrm{GW}^\pm(k)$ on $n$ generators $(E_{\omega_1}, \phi_{\omega_1}), \ldots, (E_{\omega_n}, \phi_n)$, with $(E_{\omega_i}, \phi_i)$ of degree $(-1)^i$. (We will see in Lemma 2.4.2, using the generalization of Proposition 2.3.12 discussed in Remark 2.3.16, that alternative polynomial generators are again given by the exterior powers of the standard representation.)

**Proof of Proposition 2.3.12**   We have seen in the proof of Proposition 2.2.3 that $E_{\zeta_j} \cdot E_{-\zeta_j} = 1$ in $\mathrm{K}_0(\mathbf{Rep}(G))$. In particular, each $E_{\zeta_j}$ is one-dimensional. We can therefore pick symmetric isomorphisms $\phi_{-\zeta_j}$ on each $E_{-\zeta_j}^\vee$ such that

$$(E_{\zeta_j}, \phi_{\zeta_j}) \cdot (E_{-\zeta_j}, \phi_{-\zeta_j}) = 1$$

in $\mathrm{GW}(\mathbf{Rep}(G, \iota))$. It follows that we have a well-defined $\mathrm{GW}^\pm(k)$-algebra homomorphism

$$f : \mathrm{GW}^\pm(k)[w_1, \ldots, w_n, z_1^{\pm 1}, \ldots, z_m^{\pm 1}] \to \mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$$

sending $w_i$ to $(E_{\omega_i}, \phi_{\omega_i})$, $z_j$ to $(E_{z_j}, \phi_{z_j})$, and $z_j'$ to $(E_{-z_j}, \phi_{-z_j})$.

To see that this map is an isomorphism, we argue exactly as in the proof of Proposition 2.2.3 and associate the Laurent monomial $M^x := w_1^{a_1} \cdots w_n^{a_n} \cdot z_1^{b_1} \cdots z_m^{b_m}$ with the element $x = \sum_i a_i \omega_i + \sum_j b_j \zeta_j \in X^+$. Under the map $f$ above, the basis element $M^x$ maps to

$$\prod_{i=1}^n (E_{\omega_i}, \phi_{\omega_i})^{a_i} \cdot \prod_{j=1}^m (E_{\zeta_j}, \phi_{\zeta_j})^{b_j},$$

and it suffices to show that these classes form a basis of $\mathrm{GW}^\pm(\mathbf{Rep}(G))$ considered as $\mathrm{GW}^\pm(k)$-module. By Lemma 2.3.11, we can rewrite $f(M^x)$ as

$$(2.3.2) \qquad\qquad f(M^x) = (E_x, \phi_x) + \sum_z a_z (E_z, \phi_z),$$

where $\phi_x$ is an appropriately chosen (anti)symmetric isomorphism on $E_x$, the sum is over certain $z \in X^+$ such that $z < x$, and $a_z \in \mathrm{GW}^\pm(k)$. (Remember all $z \in X^+$ are self-dual by assumption.) Proposition 2.3.8 tells us that, if we equip each irreducible representation $E_x$ with the isomorphism $\phi_x$ that appears here, the elements $(E_x, \phi_x)$ form a $\mathrm{GW}^\pm(k)$-module basis for $\mathrm{GW}^\pm(\mathbf{Rep}(G))$. The claim then again follows from [7, chapitre VI, section 3, n°4, lemme 4]. □

Just as in the case of the usual representation ring (Remark 2.2.6), we can and sometimes will pick more general generators than the ones specified in Proposition 2.3.12.

**Remark 2.3.16**   (alternative generators)   Under the assumptions of Proposition 2.3.12, we can more generally choose generators for $\mathrm{GW}(\mathbf{Rep}(G, \iota))$ as follows. As before, pick (anti)symmetric isomorphisms $\phi_{\omega_i}$ and $\phi_{\zeta_j}$ for each of the representations $E_{\omega_i}$ and $E_{\zeta_j}$, respectively. Pick homogeneous classes $e_{\omega_i} \in \mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$ for $i \in \{1, \ldots, n\}$ such that

$$e_{\omega_i} = (E_{\omega_i}, \phi_{\omega_i}) + \text{(smaller terms)},$$

where (smaller terms) is to be read as in Lemma 2.3.10. Then again we have an isomorphism of graded $\mathrm{GW}^\pm(k)$-algebras

$$f : \mathrm{GW}^\pm(k)[w_1, \ldots, w_n, z_1^{\pm 1}, \ldots, z_m^{\pm 1}] \to \mathrm{GW}^\pm(\mathbf{Rep}(G, \iota))$$

taking each $w_i$ to $e_{\omega_i}$ and each $z_j$ to $(E_{\zeta_j}, \phi_{\zeta_j})$. Indeed, this follows with the same proof, as the key identity (2.3.2) still holds for these more general generators.

## 2.4 Some details on symplectic groups

Now let $G = \mathrm{Sp}_{2n}$ be the symplectic group, ie the group of automorphisms of the antisymmetric form

$$(2.4.1) \qquad J := \begin{pmatrix} 0 & \mathrm{id}_n \\ -\mathrm{id}_n & 0 \end{pmatrix},$$

and $T$ the maximal torus, of rank $n$, consisting of matrices

$$(2.4.2) \qquad \begin{pmatrix} D & 0 \\ 0 & D^{-1} \end{pmatrix}$$

with $D \in \mathrm{GL}_n$ is diagonal; see, eg, [13, Exercise 2.4.6]. The Weyl group of $G$ is the wreath product $S_2 \wr S_n$, that is, the semidirect product $S_2^n \rtimes S_n$ with the permutation action of $S_n$ on $S_2^n$. The character lattice $X^*$ of $G$ is a free $\mathbb{Z}$-module of rank $n$, with dominant characters $X^+ \cong \mathbb{N}_0^n$.

If $n = 1$, then $G = \mathrm{SL}_2 = \mathrm{Sp}_2$ and $X^+ \cong \mathbb{N}_0$. The standard 2-dimensional representation $V$ of $G$ preserves the standard nondegenerate antisymmetric bilinear form $\phi$, yielding an element $[V, \phi] \in \mathrm{GW}^-(\mathbf{Rep}(G))$. Note that $V$ is a simple representation of highest weight 1, which is fundamental. By Proposition 2.3.12, $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2))$ therefore is a polynomial algebra over $\mathrm{GW}^\pm(k)$ on one generator $(V, \phi)$. It will be convenient for us to use the rank-zero element $b := (V, \phi) - H^-$ as a generator instead, so that we have

$$(2.4.3) \qquad \mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2)) \cong \mathrm{GW}^\pm(k)[b].$$

We think of $b$ as a "first Borel class", and note that $b$ generates the graded augmentation ideal of $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2))$ (see Definition 3.3.1).

For $n > 1$, it will be convenient for us to replace the standard polynomial generators $E_{\omega_i}$ of $\mathrm{K}_0(\mathbf{Rep}(\mathrm{Sp}_{2n}))$ and the corresponding generators of $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2n}))$ from Example 2.3.15 by alternative polynomial generators given by the exterior powers $\Lambda^i(V)$ and $\Lambda^i(V, \phi)$, respectively, where $V$ is the $2n$-dimensional standard representation.

**Lemma 2.4.1** *In* $\mathrm{K}_0(\mathbf{Rep}(\mathrm{Sp}_{2n}))$, $\Lambda^i V = E_{\omega_i} + \sum_{x:x<\omega_i} n_x E_x$ *for all* $i \in \{1, \dots, n\}$, *where the sum runs over a finite number of dominant weights* $x$ *such that* $x < \omega_i$ *in the partial order of Definition 2.1.1, and* $n_x \in \mathbb{N}$. *In particular,* $\mathrm{K}_0(\mathbf{Rep}(\mathrm{Sp}_{2n})) \cong \mathbb{Z}[V, \Lambda^2 V, \dots, \Lambda^n V]$.

**Proof** We first note that the fundamental weights $\omega_i$ of $\mathrm{Sp}_{2n}$ satisfy

$$(2.4.4) \qquad \omega_{i+2} > \omega_i$$

for all $i$, ie for all $i \in \{1, \dots, n-2\}$. Indeed, we see from [7, planche III] that

$$\omega_{i+2} - \omega_i = \alpha_{i+1} + \left( \sum_{j=i+2}^{r-1} \alpha_j \right) + \alpha_r,$$

where $\alpha_1, \dots, \alpha_n$ are the simple roots, so $\omega_{i+2} - \omega_i > 0$.

In characteristic zero, $\Lambda^i V$ decomposes as

$$(2.4.5) \qquad \Lambda^i V \cong E_{\omega_i} \oplus E_{\omega_{i-2}} \oplus \cdots \oplus \begin{cases} E_{\omega_1} & \text{if } i \text{ is odd,} \\ k & \text{if } i \text{ is even,} \end{cases}$$

for $i = 1, \ldots, n$ *as a representation*, not just as an element of $K_0(\mathbf{Rep}(\mathrm{Sp}_{2n}))$. (See [8, Chapter VIII, Section 13.3 (IV)] for a proof of the corresponding statement for the Lie algebra of $\mathrm{Sp}_{2n}$, and recall that in characteristic zero we can pass from decompositions of a representation of a split semisimple simply connected algebraic group to decompositions of the associated representation of its Lie algebra and vice versa using [22, Section 3.2].) The claim of the lemma is therefore immediate from (2.4.4) in characteristic zero.

Suppose now that $k$ is a field of positive characteristic. Note that $\mathrm{Sp}_{2n}$ lifts to a split reductive group scheme $\mathrm{Sp}_{2n,\mathbb{Z}}$ over $\mathbb{Z}$ [10, théorème 7.2.0.46], hence the results of [43, section 3.7] apply. Let us temporarily write $\mathrm{Sp}_{2n,\mathbb{Q}}$ and $\mathrm{Sp}_{2n,k}$ for the symplectic groups over $\mathbb{Q}$ and $k$, which can be obtained from $\mathrm{Sp}_{2n,\mathbb{Z}}$ via base change, and let us denote by $E_{\omega_i,\mathbb{Q}}$ and $E_{\omega_i,k}$ the respective irreducible representations of highest weight $\omega_i$. Consider the category $\mathbf{Rep}(\mathrm{Sp}_{2n,\mathbb{Z}})$ of $\mathrm{Sp}_{2n,\mathbb{Z}}$-representations that are finitely generated and free over $\mathbb{Z}$, and the associated K-ring $K_0(\mathbf{Rep}\,\mathrm{Sp}_{2n,\mathbb{Z}})$, denoted by $R_{\mathbb{Z}}(\mathrm{Sp}_{2n})$ in [43, section 3.3].

The base-change homomorphisms

$$i_{\mathbb{Q}} \colon K_0(\mathrm{Sp}_{2n,\mathbb{Z}}) \to K_0(\mathrm{Sp}_{2n,\mathbb{Q}}) \quad \text{and} \quad i_k \colon K_0(\mathrm{Sp}_{2n,\mathbb{Z}}) \to K_0(\mathrm{Sp}_{2n,k})$$

are homomorphisms of $\lambda$-rings, essentially because exterior powers are compatible with base change [16, Proposition A2.2(b)]. By [43, théorème 5], $i_{\mathbb{Q}}$ is an isomorphism. Better still, we see from [43, lemme 2(a)] that for each representation $E_{\mathbb{Q}}$ of $\mathrm{Sp}_{2n,\mathbb{Q}}$ there exists a representation $E_{\mathbb{Z}}$, not just a virtual representation, such that $i_{\mathbb{Q}}(E_{\mathbb{Z}}) = E_{\mathbb{Q}}$. In particular, for each dominant weight $\omega$, we have a representation $E_{\omega,\mathbb{Z}}$ of $\mathrm{Sp}_{2n,\mathbb{Z}}$ with $i_{\mathbb{Q}}(E_{\omega,\mathbb{Z}}) = E_{\omega,\mathbb{Q}}$. Note, however, that there is no reason to assume that $i_k(E_{\omega,\mathbb{Z}}) = E_{\omega,k}$ in general. Rather, arguing as in [43, section 3.8, remarque 3], we find that

$$(2.4.6) \qquad i_k(E_{\omega,\mathbb{Z}}) = E_{\omega,k} + \sum_{x < \omega} n_x E_{x,k},$$

where the sum is over all dominant $x$ with $x < \omega$, and only finitely many of the coefficients $n_x \in \mathbb{N}_0$ are nonzero. Indeed, let us write $\mathrm{ch}_k$ for the character homomorphism $\mathrm{ch}_{\mathrm{Sp}_{2n,k}}$ (see Lemma 2.2.1). As we have a maximal torus of $\mathrm{Sp}_{2n}$ defined over $\mathbb{Z}$, and as restriction to this maximal torus is compatible with base change, we find that $\mathrm{ch}_{\mathrm{Sp}_{2n,k}}(i_k(E_{\omega,\mathbb{Z}})) = \mathrm{ch}_{\mathrm{Sp}_{2n,\mathbb{Q}}}(E_{\omega,\mathbb{Q}})$. Thus, [43, lemme 5(a)], applied both over $\mathbb{Q}$ and over $k$, yields the following two identities, for certain coefficients $a_x, b_x \in \mathbb{N}_0$:

$$\mathrm{ch}_k(i_k(E_{\omega,\mathbb{Z}})) = e^{\omega} + \sum_{x < \omega} a_x e^x \quad \text{and} \quad \mathrm{ch}_k(E_{\omega,k}) = e^{\omega} + \sum_{x < \omega} b_x e^x.$$

Equation (2.4.6) follows from a comparison of these two identities and the injectivity of $\mathrm{ch}_k$.

On the other hand, the standard representation of $\mathrm{Sp}_{2n}$ is already defined over $\mathbb{Z}$, so we have a representation $V_{\mathbb{Z}}$ of $\mathrm{Sp}_{2n,\mathbb{Z}}$ with $i_k(V_{\mathbb{Z}}) = V_k$ for arbitrary fields $k$. Equation (2.4.5), appropriately decorated

with subscripts $(-)_{\mathbb{Z}}$, is therefore equally valid in $K_0(Sp_{2n,\mathbb{Z}})$. The claimed decomposition of $\Lambda^i(V)$ in $K_0(\mathbf{Rep}(Sp_{2n}))$ now follows by combining (2.4.4), this integral version of (2.4.5), and (2.4.6).

The final claim is immediate from this decomposition and Remark 2.2.6. □

**Lemma 2.4.2** *Let $(V, \phi)$ denote the $2n$-dimensional standard representation of $Sp_{2n}$ with its canonical antisymmetric form. The exterior powers $\Lambda^i(V, \phi)$ for $i \in \{1, \ldots, n\}$ define homogeneous polynomial generators of the $\pm$-graded $GW^{\pm}(k)$-algebra $GW^{\pm}(\mathbf{Rep}\, Sp_{2n})$:*

$$GW^{\pm}(\mathbf{Rep}(Sp_{2n})) = GW^{\pm}(k)[(V, \phi), \Lambda^2(V, \phi), \ldots, \Lambda^n(V, \phi)].$$

*The generator $\Lambda^i(V, \phi)$ is of degree $(-1)^i$.*

**Proof** Applying Lemma 2.3.10 to the decomposition of $\Lambda^i(V)$ given in Lemma 2.4.1, we find that $\Lambda^i(V, \phi) = (E_{\omega_i}, \phi_{\omega_i}) + \text{(smaller terms)}$ for certain symmetric isomorphism $\phi_{\omega_i}$ on $E_{\omega_i}$. So the claim is immediate from Remark 2.3.16. □

**Restriction to diagonal** The group $Sp_{2n}$ has a canonical subgroup isomorphic to $Sp_2^{\times n} = Sp_2 \times \cdots \times Sp_n$. This is most easily seen by replacing the antisymmetric form $J$ from (2.4.1) by the isometric form $nH^- = H^- \perp \cdots \perp H^-$; then $Sp_2^{\times n}$ is simply given by $n$ diagonally concatenated copies of $Sp_2$ inside $Sp_{2n}$. The inclusion of this subgroup induces a commutative diagram

$$
\begin{array}{ccc}
K_0(\mathbf{Rep}(Sp_{2n})) & \xrightarrow{\ \text{res}\ } & K_0(\mathbf{Rep}(Sp_2^{\times n})) \\
& \searrow & \downarrow \\
& & K_0(\mathbf{Rep}(T))
\end{array}
$$

which shows that the restriction map $\text{res}: K_0(\mathbf{Rep}(Sp_{2n})) \to K_0(\mathbf{Rep}(Sp_2^{\times n}))$ is injective. Its image can easily be identified, as follows (see explicit calculations in [25, Appendix D]). Consider the symmetric group $S_n$ acting by permutation on $Sp_2^{\times n}$, and its induced action on $K_0(\mathbf{Rep}(Sp_2^{\times n}))$. The image of the restriction map is precisely the fixed ring of this action:

$$K_0(\mathbf{Rep}(Sp_{2n})) \xrightarrow[\text{res}]{\cong} K_0(\mathbf{Rep}(Sp_2^{\times n}))^{S_n}$$

We will now show that the corresponding statement for $GW^{\pm}$ also holds.

**Lemma 2.4.3** *The restriction map $GW^{\pm}(\mathbf{Rep}(Sp_{2n})) \to GW^{\pm}(\mathbf{Rep}(Sp_2^{\times n}))$ is injective, with image the invariant subring under the permutation action of the symmetric group $S_n$.*

**Proof** Note that $Sp_2^{\times n}$ is a simply connected reductive group, and that Proposition 2.3.12 applies just as it applies to $Sp_{2n}$. So $GW^{\pm}(\mathbf{Rep}(Sp_2^{\times n}))$ is a polynomial ring in $n$ variables, which we can identify with $\prod_{i=1}^n GW^{\pm}(\mathbf{Rep}(Sp_2))$. Writing $v^{(i)} := (V^{(i)}, \phi^{(i)})$ for the standard representation of the $i^{\text{th}}$ factor in $Sp_2^{\times n}$, equipped with its canonical antisymmetric form, we thus obtain

$$(2.4.7) \qquad GW^{\pm}(\mathbf{Rep}(Sp_2^{\times n})) \cong \mathbb{Z}[v^{(1)}, \ldots, v^{(n)}].$$

Under this isomorphism, the permutation action of $S_n$ on the left corresponds to the obvious permutation action on the generators $v^{(i)}$ on the right.

Now let $(V, \phi)$ denote the $2n$-dimensional standard representation of $\mathrm{Sp}_{2n}$, equipped with its canonical antisymmetric form. As we have seen in Lemma 2.4.2 we can take the exterior powers $\Lambda^k(V, \phi)$ for $k = 1, \ldots, n$ as polynomial generators of $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2n}))$. Write $\sigma^k(v^\bullet) \in \mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2^{\times n}))$ for the $k^{\text{th}}$ elementary symmetric function in the classes $v^{(i)} = (V^{(i)}, \phi^{(i)})$ introduced above. We claim that

(2.4.8) $$\mathrm{res}\, \Lambda^k(V, \phi) = \sigma_k(v^\bullet) + \text{a polynomial in } \sigma_j(v^\bullet) \text{ with } j < k.$$

To this end, recall that the operations $\lambda^k$ defined in terms of the exterior powers $\Lambda^k$ give $\mathrm{GW}^\pm(\mathbf{Rep}(G))$ the structure of a pre-$\lambda$-ring (for any affine group scheme $G$); see [17, Proposition 4.2.1; 49]. For $k = 1$, equation (2.4.8) simply follows from the fact that $(V, \phi)$ restricts to the direct sum of the representations $(V^{(i)}, \phi^{(i)})$:

$$\mathrm{res}(V, \phi) = \sum_{i=0}^n v(i) = \sigma_1(v^\bullet).$$

For higher $k$, consider the power series expansion $\lambda_t(x) := \sum_k \lambda^k(x) t^k$. For the two-dimensional standard representation $(V_2, \phi_2)$ of $\mathrm{Sp}_2$, we find $\Lambda^2(V_2, \phi_2) = (k, \det(\phi)) = (k, \mathrm{id})$, and $\Lambda^k(V_2, \phi_2) = 0$ for $k > 2$, so

$$\lambda_t(V_2, \phi_2) = 1 + (V_2, \phi_2)t + t^2.$$

As the restriction commutes with the $\lambda$-operations, this implies

$$\mathrm{res}\, \lambda_t(V, \phi) = \lambda_t\left(\sum_i v^{(i)}\right) = \prod_{i=1}^n \lambda_t v^{(i)} = \prod_{i=1}^n (1 + v^{(i)}t + t^2).$$

Now $\mathrm{res}\, \lambda^k(V, \phi)$ is the coefficient of $t^k$ in the above power series. As the power series is invariant under the permutation action of $S_n$, so is each coefficient. So $\mathrm{res}\, \lambda^k(V, \phi)$ is a polynomial in the symmetric polynomials $\sigma^i(v^\bullet)$. Moreover, the highest-degree monomials in the $V^{(i)}$'s that occur in the coefficient of $t^k$ are precisely the monomials that occur $\sigma^k(v^\bullet)$, and they occur with multiplicity one. This proves (2.4.8), and hence the lemma. $\qquad\square$

For the completion at the augmentation ideal, it is again convenient to reformulate the above description in terms of the "first Borel classes" defined as in (2.4.3) above. So let us write $b^{(i)} := (V^{(i)}, \phi^{(i)}) - H^-$ for the first Borel class of the $i^{\text{th}}$ copy of $\mathrm{Sp}_2$, and $\sigma_i(b^\bullet)$ for the $i^{\text{th}}$ elementary symmetric polynomial in these classes. Clearly, the isomorphism (2.4.7) from the proof above can be rewritten as

(2.4.9) $$\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2^{\times n})) \cong \mathbb{Z}[b^{(1)}, \ldots, b^{(n)}].$$

**Corollary 2.4.4** *The restriction induces an isomorphism between $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2n}))$ and the subring $\mathrm{GW}^\pm[\sigma_1(b^\bullet), \ldots, \sigma_n(b^\bullet)]$ of $\mathrm{GW}^\pm(\mathrm{Sp}_2^{\times n})$. Under this isomorphism, the graded augmentation ideal $\mathrm{IO}_{\mathrm{Sp}_{2n}}^\pm$ (see Definition 3.3.1) corresponds to the ideal generated by $\sigma_1(b^\bullet), \ldots, \sigma_n(b^\bullet)$. In particular,*

$$\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2n}))^\wedge_{\mathrm{IO}_{\mathrm{Sp}_{2n}}^\pm} \cong \mathrm{GW}^\pm(k)[\![\sigma_1(b^\bullet), \ldots, \sigma_n(b^\bullet)]\!].$$

# 3 Grothendieck–Witt rings of some classifying spaces

The main purpose of this chapter is to prove Theorem 1.0.2. First, we introduce the specific model category we will be working with to model motivic spaces. Next, in Section 3.2, we recall the notion of *acceptable gadgets* as introduced in [32, Definition 8.3] and establish some important properties of these. In Section 3.3, we finally use the results of Section 2.4 to construct the Atiyah–Segal completion map, and to prove the main theorem. In the final two subsections, we discuss the construction of a classifying space for the multiplicative group with nontrivial involution, and conjectural tools that may be used to generalize our main results to base schemes with nontrivial action and to arbitrary closed subgroups of $\mathrm{Sp}_{2n}$.

## 3.1 Čech localization

Let $S$ be a scheme of finite type over a field $k$. Our standing assumption that $\mathrm{char}(k) \neq 2$ is not necessary for any of the results of this section. Let $\mathbf{sPre}(\mathbf{Sm}_S)$ be the model category of simplicial presheaves on $\mathbf{Sm}_S$ with the global injective model structure, and $L_{\mathrm{mot}}\,\mathbf{sPre}(\mathbf{Sm}_S)$ its motivic localization, which is defined as $L_{\mathrm{mot}} = (L_{\mathbb{A}^1}L_{\mathrm{Nis}})^\infty$, where the repeated localizations are necessary to ensure that $L_{\mathrm{mot}}F$ is both Nisnevich local and $\mathbb{A}^1$-local for any simplicial presheaf $F$. This is a presheaf variant of the model category constructed in [30], with homotopy category the unstable motivic homotopy category $\mathcal{H}(S)$. The weak equivalences in $L_{\mathrm{mot}}\,\mathbf{sPre}(\mathbf{Sm}_S)$ will be called *motivic weak equivalences*; these are precisely the maps that become isomorphisms in $\mathcal{H}(S)$.

More generally, for a subcanonical topology $\tau$ on $\mathbf{Sm}_S$, let $L_\tau\,\mathbf{sPre}(\mathbf{Sm}_S)$ be the (left) Bousfield localization with respect to covering sieves for $\tau$, as considered in [1, Section 3.1]. This localization is also considered in [15, Appendix A], where it is called the Čech localization because it is a localization with respect to Čech covers in the topology $\tau$; see [15, Theorem A5]. This is the naming convention we will follow.

The fibrant objects of $L_\tau\,\mathbf{sPre}(\mathbf{Sm}_S)$ are those fibrant simplicial presheaves in $\mathbf{sPre}(\mathbf{Sm}_S)$ that satisfy $\tau$-descent; we will refer to these as $\tau$-fibrant. The weak equivalences in $\mathbf{sPre}(\mathbf{Sm}_S)$ will be called *objectwise weak equivalences* and the weak equivalences in $L_\tau\,\mathbf{sPre}(\mathbf{Sm}_S)$ will be called $\tau$-*local weak equivalences*. We denote by $L_\tau$ the $\tau$-fibrant replacement functor, viewed as endofunctor on $\mathbf{sPre}(\mathbf{Sm}_S)$.

**Remark 3.1.1** (relation to model structure used by Morel and Voevodsky) Note that $L_\tau\,\mathbf{sPre}(\mathbf{Sm}_S)$ is *not* the same as the Jardine model structure $L_{\mathrm{hyp}\text{-}\tau}\,\mathbf{sPre}(\mathbf{Sm}_S)$, which is Quillen equivalent to the Joyal model structure on simplicial sheaves used in [30]. In general, we only have successive Bousfield localizations

$$\mathbf{sPre}(\mathbf{Sm}_S) \to L_\tau\,\mathbf{sPre}(\mathbf{Sm}_S) \to L_{\mathrm{hyp}\text{-}\tau}\,\mathbf{sPre}(\mathbf{Sm}_S).$$

However, the Bousfield localization $L_\tau\,\mathbf{sPre}(\mathbf{Sm}_S) \to L_{\mathrm{hyp}\text{-}\tau}\,\mathbf{sPre}(\mathbf{Sm}_S)$ is a Quillen equivalence for $\tau$ the Nisnevich topology and $S$ any reasonable base scheme. See [1, Remark 3.1.4] for more details.

**Remark 3.1.2** ($\infty$-language) We may also view $\mathbf{sPre}(\mathbf{Sm}_S)$ as an $\infty$-category as in, eg, [20]. The fibrant replacement functor $L_\tau$ is analogous to the localization endofunctor defined in [20, Proposition 3.4] and has the same formal properties. Lemma 2.1 from [21] will play a key role in our discussion of classifying spaces (see the proof of Proposition 3.2.16 below).

We will need very few specifics about the $\tau$-local model structure. The proof of the following proposition relies on the two subsequent lemmas, both presumably well known.

**Proposition 3.1.3** (a) *For any subcanonical topology $\tau$, every $S$-scheme is $\tau$-fibrant.*

(b) *For any topology $\tau$ at least as coarse as the fppf topology, any filtered colimit in $\mathbf{sPre}(\mathbf{Sm}_S)$ of $S$-schemes is $\tau$-fibrant.*

*In particular, for any scheme $X$ as in* (a) *or any filtered colimit $X$ as in* (b), *the $\tau$-fibrant replacement map $X \to L_\tau(X)$ is an objectwise weak equivalence.*

**Proof** For (a), let $X$ be an $S$-scheme. We need to show that $X$ is fibrant in the injective model structure on $\mathbf{sPre}(\mathbf{Sm}_S)$ and $\tau$-local. For the first assertion, see Lemma 3.1.4 below. For the second assertion, see [1, Example 3.1.2]: a simplicially constant presheaf is $\tau$-local if and only if it is a $\tau$-sheaf. The presheaf $X$ in question is a $\tau$-sheaf precisely because $\tau$ is assumed to be subcanonical. For the filtered colimit in (b), the claim follows by the same argument plus the observation, spelled out in Lemma 3.1.5 below, that such filtered colimits are indeed sheaves. The final assertion is just an application of Ken Brown's lemma. □

As remarked above, the following two lemmas are probably well known to the experts. Lemma 3.1.4 is stated without proof in [36, Section 2.6]. We include full details here for future reference.

**Lemma 3.1.4** *Let $\mathcal{C}$ be a small category. A simplicially constant presheaf $F \in \mathbf{sPre}(\mathcal{C})$ is fibrant in any model structure on $\mathbf{sPre}(\mathcal{C})$ whose weak equivalences are objectwise weak equivalences. In particular, constant simplicial presheaves are fibrant in the injective model structure.*

**Proof** Let $i\colon A \to B$ be an objectwise weak equivalence in $\mathbf{sPre}(\mathcal{C})$ and $a\colon A \to F$ a map of simplicial presheaves. To prove the statement, it suffices to prove the existence of a map $b\colon B \to F$ such that $bi = a$. Note that $F$ is an objectwise Kan complex and therefore projective fibrant, so that $a$ factors through a functorial projective fibrant replacement $RA$ of $A$, as in the diagram

$$
\begin{array}{ccccc}
A & \overset{\sim}{\longrightarrow} & RA & \overset{a'}{\longrightarrow} & F \\
\downarrow{\scriptstyle i} & & \downarrow{\scriptstyle Ri} & \nearrow & \\
B & \overset{\sim}{\longrightarrow} & RB & {\scriptstyle b'} &
\end{array}
$$

and a lift $b$ of $a$ along $i$ exists if a lift $b'\colon RB \to F$ of $a'$ along $Ri$ exists. Therefore, we may and do assume both $A$ and $B$ to be projective fibrant, that is, they are both objectwise Kan complexes.

We first produce a collection of maps $b_C : B(C) \to F(C)$ such that $a_C = b_C i_C$ for each $C \in \mathcal{C}$. Let $C \in \mathcal{C}$. Since $i_C : A(C) \to B(C)$ is a weak equivalence of Kan complexes, it is a simplicial homotopy equivalence. Let $j_C : B(C) \to A(C)$ be a homotopy inverse to $i_C$ and define $b_C : B(C) \to F(C)$ as the composition $a_C j_C$. Then the composition $a_C j_C i_C : A(C) \to F(C)$ is homotopic to the map $a_C$, and since homotopic maps into a constant simplicial set are equal, $a_C = b_C i_C$. Thus, for each $C \in \mathcal{C}$ we have $b_C = a_C j_C$ for a fixed choice of homotopy inverse $j_C$ of $i_C$.

We now need to show that the square

$$
\begin{array}{ccc}
B(D) & \xrightarrow{\ b_D\ } & F(D) \\
\downarrow{\scriptstyle B(f)} & & \downarrow{\scriptstyle F(f)} \\
B(C) & \xrightarrow{\ b_C\ } & F(C)
\end{array}
$$

commutes for any arrow $f : C \to D$ in $\mathcal{C}$. Consider the diagram

$$
\begin{array}{ccccc}
B(D) & \xrightarrow{\ j_D\ } & A(D) & \xrightarrow{\ a_D\ } & F(D) \\
\downarrow{\scriptstyle B(f)} & & \downarrow{\scriptstyle A(f)} & & \downarrow{\scriptstyle F(f)} \\
B(C) & \xrightarrow{\ j_C\ } & A(C) & \xrightarrow{\ a_C\ } & F(C)
\end{array}
$$

and note that the right square commutes. Again because homotopic maps into a constant simplicial set are equal, it suffices to show that the left square commutes up to homotopy. Note that

$$ j_C i_C A(f) j_D = j_C B(f) i_D j_D $$

because $i$ is a map of simplicial presheaves. Moreover, the left-hand side is homotopic to $A(f) j_D$ and the right-hand side is homotopic to $j_C B(f)$ because $j_C i_C \sim \mathrm{id}_{A(C)}$ and $i_D j_D \sim \mathrm{id}_{B(D)}$ by definition. Hence, $F(f) b_D = b_C B(f)$. $\qquad\square$

**Lemma 3.1.5** *Consider a topology $\tau$ coarser than the fppf topology. Given a family $(F_i)_{i \in I}$ of representable $\tau$-sheaves on a quasicompact scheme $S$, where $I$ is a filtered index category, the filtered colimit $F_\infty := \mathrm{colim}_{i \in I} F_i$ is also a $\tau$-sheaf.*

**Proof** As $S$ is quasicompact, we have to check the sheaf condition for $F_\infty$ only for finite $\tau$-covers $\mathcal{U}$ of a scheme $U$, by the following argument. Let $\mathcal{U}$ be an arbitrary $\tau$-cover of $U$. As $U$ is quasicompact, there exists a finite Zariski cover $\mathcal{V}$ of $U$. Note that $\mathcal{U} \cap \mathcal{V}$ is a refinement of $\mathcal{U}$. For $V \in \mathcal{V}$, we consider the $\tau$-cover $\mathcal{U} \cap V$ of $V$. By (the proof of) [44, Tag 021P], this refines to a finite $\tau$-cover of $V$. Hence $\mathcal{U} \cap \mathcal{V}$ refines to a finite $\tau$-cover of $U$, and from now on we assume that $\mathcal{U}$ is a finite $\tau$-cover of $U$. We denote the associated covering sieve by $h_{\mathcal{U}}$, which is a subpresheaf of the representable presheaf $h_U$. We are reduced to showing that if the canonical map $\mathrm{Hom}(h_U, F_i) \to \mathrm{Hom}(h_{\mathcal{U}}, F_i)$ is a bijection for all $i \in I$, then it also is a bijection for $i = \infty$. The representable sheaf $h_U$ is compact, hence commutes with filtered colimits.

Moreover, for a finite covering $\mathcal{U} = (U_j \to U)_{j \in J}$, $h_{\mathcal{U}}$ can be described by the explicit small coequalizer

$$\coprod_{j,j'} U_j \times_U U_{j'} \rightrightarrows \coprod_j U_j$$

of representables, hence is also compact. □

## 3.2 Acceptable gadgets and classifying spaces

Let $S$ be a scheme of finite type over a field $k$. Our standing assumption that $\mathrm{char}(k) \neq 2$ is not necessary for any of the results of this section.

**Classifying spaces** We now fix a $\tau$-sheaf of groups $G$ over $S$. In applications, $G$ will often be a linear algebraic group over $\mathrm{Spec}\, k$ base changed to $S$.

The *simplicial classifying space* $\mathrm{B}G$ is defined in [30] as the nerve of $G$ viewed as a presheaf of groupoids with a single object (so the $n$-simplices are $\mathrm{B}G_n := G^{\times n}$ and face and degeneracy maps are given by composition and inserting identities, as usual).

**Definition 3.2.1** For a topology $\tau$ at least as coarse as the fppf topology, the $\tau$-*classifying space* $\mathrm{B}_\tau G$ of $G$ is defined as $\mathrm{B}_\tau G := L_\tau(\mathrm{B}G)$.

Note that $\mathrm{B}_\tau G$ is well defined — irrespective of any particular choice of fibrant replacement functor — up to objectwise weak equivalence. We now briefly discuss the dependency of $\mathrm{B}_\tau G$ on the topology $\tau$. For the definition of $\tau$-locally trivial $G$-torsors, we refer to [2, Definition 2.3.1].

**Lemma 3.2.2** *The following conditions on $G$ and $\tau$ are equivalent*:

(1) *Every (fppf-locally trivial) $G$-torsor in $\mathbf{Sm}_S$ is already $\tau$-locally trivial.*

(2) *The canonical map $\mathrm{B}_\tau G \to \mathrm{B}_{\mathrm{fppf}} G$ is an objectwise weak equivalence.*

**Proof** This equivalence follows from [2, Lemma 2.3.2(i) and (ii)]. (See also [30, Lemma 4.1.18]). □

**Example 3.2.3** (smooth affine groups) For smooth affine group schemes $G$, every $G$-torsor is already étale-locally trivial, so $\mathrm{B}_{\mathrm{fppf}} G$ is objectwise weakly equivalent to $\mathrm{B}_{\mathrm{ét}} G$. Indeed, smoothness is preserved by faithfully flat descent, so any $G$-torsor is smooth over the base, and smooth morphisms admit sections étale locally.

**Example 3.2.4** (special groups) A linear algebraic group $G$ over $S$ is called *special* if every étale-locally trivial $G$-torsor over a (not necessarily smooth) $S$-scheme is already Zariski-locally trivial. Thus, for a special group $G$, $\mathrm{B}_{\mathrm{Zar}} G$ is objectwise weakly equivalent to $\mathrm{B}_{\mathrm{ét}} G$, and to $\mathrm{B}_\tau G$ for any intermediate topology such as $\tau = \mathrm{Nis}$. Special groups include, in particular, split tori, $\mathrm{GL}_r$, $\mathrm{SL}_{2r}$, $\mathrm{Sp}_{2r}$, and finite products of these [35, Lemmas 3.1 and 3.2]. (In [35; 42], the notion is introduced and discussed only for groups defined over a field $k$, but note that the defining property is stable under base change along $S \to \mathrm{Spec}(k)$. Special groups over fields have been fully classified [18; 23; 29].)

Of course, for $\tau = \text{Nis}$ or any coarser topology, $B_\tau G$ is *motivically* equivalent to the simplicial classifying space $BG$, for any $G$, as $\tau$-local equivalences are motivic equivalences in this case. Only for finer choices of topologies $\tau$, $B_\tau G$ may be motivically distinct.

**Remark 3.2.5** (comparison with [30]) By construction, $B_\tau G$ only satisfies $\tau$-descent. However, it also satisfies $\tau$-hyperdescent for dimension reasons by the argument given in the proof of [2, Lemma 2.3.2], so by Ken Brown's lemma, the canonical map $B_\tau G \to L_{\text{hyp-}\tau} BG$ is an objectwise weak equivalence. In view of Remark 3.1.1, this shows, in particular, that the definition of $B_{\text{ét}} G$ used here agrees with the definition in [30, page 130].

**Outline** Using the *admissible gadgets with a nice $G$-action* of [30, Definitions 4.2.1 and 4.2.4], one constructs a geometric approximation of a universal $G$-torsor over $B_{\text{ét}} G$, namely a colimit of $G$-torsors in $\mathbf{Sm}_S$. We will instead use the more flexible concept of *acceptable gadgets*, following [32, Definition 8.3], and adding a version of a nice $G$-action in this setting. Suppose that we have a commutative diagram of $G$-torsors

(3.2.1)
$$
\begin{array}{ccc}
X_1 \longrightarrow & X_2 \longrightarrow & \cdots \\
\downarrow & \downarrow & \\
Y_1 \longrightarrow & Y_2 \longrightarrow & \cdots
\end{array}
$$

in $\mathbf{Sm}_S$, where the horizontal arrows are closed immersions, such that the colimit of this diagram should be thought of a motivic approximation of $EG \to B_{\text{ét}} G$, with $EG$ being a motivically contractible space with free $G$-action. We will make this idea precise further in a more general setting for any Grothendieck topology $\tau$ at least as coarse as the fppf topology, and for $\tau$-locally trivial $G$-torsors and $\tau$-classifying spaces $B_\tau G$.

Namely, we will establish a chain of weak equivalences

$$
Y_\infty \underset{(1)}{\simeq} L_\tau(X_\infty /\!\!/ G) \underset{(2)}{\simeq} L_\tau(S /\!\!/ G) \underset{(3)}{\cong} B_\tau G.
$$

All quotients $- /\!\!/ G$ appearing here are stacky; see Definition 3.2.9. The objectwise weak equivalence (1) is established below in Lemma 3.2.13. Equivalence (2) is a motivic weak equivalence; see Proposition 3.2.16. The isomorphism (3) is completely general and clear from the definitions; see Lemma 3.2.10. Combined, these equivalences yield *a motivic weak equivalence $Y_\infty \simeq B_\tau G$ for any sequence of $\tau$-locally trivial $G$-torsors as in* (3.2.1) *in which the sequence of $X_i$'s forms an acceptable gadget.* This result, which is stated as Theorem 3.2.17 below, will be used in the next section to show that the products $(\mathbb{HP}^n)^{\times r}$ approximate $BSp_2^{\times r}$, and similarly certain quaternionic Grassmannians approximate $BSp_{2r}$. These are crucial steps in the proof of Atiyah–Segal completion for $BSp_{2r}$.

We begin by recalling the definition of an acceptable gadget.

**Definition 3.2.6** An *acceptable gadget* $(X_n)_{n \in I}$ over a scheme $S$ is a countable totally ordered set $I$ together with a diagram $X : I \to \mathbf{Sm}_S$ such that each $X_n = X(n)$ is quasiprojective and for each $n < m$ in $I$, the map $X(n < m) : X_n \to X_m$ is a closed immersion of $S$-schemes, and such that for any Henselian regular local ring $R$ and any commutative square

$$
\begin{array}{ccc}
\partial\Delta^i_R & \longrightarrow & X_n \\
\downarrow & & \downarrow \\
\Delta^i_R & \longrightarrow & S
\end{array}
$$

there exists an $m \geq n$ and a map $\Delta^i_R \to X_m$ making this diagram commute:

(3.2.2)
$$
\begin{array}{ccccc}
\partial\Delta^i_R & \longrightarrow & X_n & \longrightarrow & X_m \\
\downarrow & & & \nearrow & \downarrow \\
\Delta^i_R & & \longrightarrow & & S
\end{array}
$$

Here is a lemma that shows that acceptable gadgets yield contractible motivic spaces, which will make them useful to construct contractible spaces with free actions, whose quotients are classifying spaces as we will see in Proposition 3.2.16. Some version of Lemma 3.2.7 is implied in [32, page 954 after Proposition 8.5].

**Lemma 3.2.7** Let $(X_i)_{i \in I}$ be an acceptable gadget over a scheme $S$. Then $X_\infty = \mathrm{colim}_i\, X_i$ is contractible in the category $\mathcal{H}(S)$ of motivic spaces over $S$.

**Proof** By [30, Lemmas 2.3.8 and 3.1.11], it suffices to show that the simplicial set $\mathrm{Sing}(X_\infty)(R)$ is contractible for every regular Henselian local ring $R$ over $S$, in other words, for every Nisnevich point of $S$; see also Remark 3.1.1.

By construction, a map $\partial\Delta^n \to \mathrm{Sing}(X_\infty)(R)$ is given by a morphism $\partial\Delta^n_R \to X_\infty$. Let $f$ be such a morphism. Since $\partial\Delta^n_R$ is representable by an object of $\mathbf{Sm}_S$, it is compact, and $f$ factors through a finite stage $f' : \partial\Delta^n_R \to X_i$ for some $i \in I$. Since $(X_i)_{i \in I}$ is an acceptable gadget, there exists $j > i$ in $I$ such that the composition $\partial\Delta^n_R \to X_i \to X_j$ extends to a morphism $\Delta^n_R \to X_j$, which shows that $\mathrm{Sing}(X_\infty)(R)$ is contractible. $\square$

We prove a lemma that allows us to construct acceptable gadgets from existing ones. The first of these shows that acceptable gadgets are stable under base change.

**Lemma 3.2.8** Let $(X_n)_{n \in \mathbb{N}}$ and $(Y_n)_{n \in \mathbb{N}}$ be acceptable gadgets over a scheme $S$.

(a) For any $Y \in \mathbf{Sm}_S$, the base change $(X_n \times_S Y)_{n \in \mathbb{N}}$ is an acceptable gadget over $Y$.

(b) For any cofinal $I \subset \mathbb{N}$, the sequence $(X_n)_{n \in I}$ is an acceptable gadget over $S$.

(c) The product $(X_n \times_S Y_n)_{n \in \mathbb{N}}$ is an acceptable gadget over $S$.

**Proof** We only prove (c); the proofs of (a) and (b) are similar. Consider a commutative diagram

$$
\begin{array}{ccc}
\partial\Delta_R^i & \longrightarrow & X_n \times_S Y_n \\
\downarrow & & \downarrow \\
\Delta_R^i & \longrightarrow & S
\end{array}
$$

The projection maps $X_n \times_S Y_n \to X_n$ and $X_n \times_S Y_n \to Y_n$ yield commutative diagrams

$$
\begin{array}{ccccc}
\partial\Delta_R^i & \longrightarrow & X_n & \longrightarrow & X_{m_1} \\
\downarrow & & & & \downarrow \\
\Delta_R^i & & \longrightarrow & & S
\end{array}
\qquad
\begin{array}{ccccc}
\partial\Delta_R^i & \longrightarrow & Y_n & \longrightarrow & Y_{m_2} \\
\downarrow & & & & \downarrow \\
\Delta_R^i & & \longrightarrow & & S
\end{array}
$$

as $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are acceptable gadgets. Let $m = \max\{m_1, m_2\}$. Then the compositions $\Delta_R^i \to X_{m_1} \to X_m$ and $\Delta_R^i \to Y_{m_2} \to Y_m$ yield a map $\Delta_R^i \to X_m \times_S Y_m$ making this diagram commute:

$$
\begin{array}{ccccc}
\partial\Delta_R^i & \longrightarrow & X_n \times_S Y_n & \longrightarrow & X_m \times_S Y_m \\
\downarrow & & & & \downarrow \\
\Delta_R^i & & \longrightarrow & & S
\end{array}
\qquad\square
$$

All quotients $-\!/\!/\,G$ used in the following will be stacky quotients, defined as follows. Recall that the action groupoid of a set $X$ with a group action is the groupoid with objects the elements of $X$, and morphisms from $x_1 \in X$ to $x_2 \in X$ given by the group elements that send $x_1$ to $x_2$.

**Definition 3.2.9** For a presheaf of sets $X$ on $\mathbf{Sm}_S$ with $G$-action, the *stacky quotient $X \!/\!/\, G$* is the simplicial presheaf given by the nerve of the action groupoid of $X$.

We will mainly be interested in two extremal cases: the stacky quotient of the (trivial) $G$-action on the base scheme $S$, and the stacky quotient of the (free) $G$-action on a $G$-torsor. These are discussed in the following two lemmas, respectively.

**Lemma 3.2.10** *For any subcanonical topology $\tau$, we have a canonical isomorphism of simplicial presheaves $\mathrm{B}_\tau G \cong L_\tau(S \!/\!/\, G)$.*

**Proof** For any $U \in \mathbf{Sm}_S$, the action groupoid of the $G(U)$-action on the one-point set $S(U)$ is isomorphic to the groupoid $G(U)$ used in the definition of $BG$. $\qquad\square$

**Lemma 3.2.11** *For any $\tau$-locally trivial $G$-torsor $\pi\colon X \to Y$ in $\mathbf{Sm}_S$, we have a canonical objectwise weak equivalence $L_\tau(X \!/\!/\, G) \to Y$.*

**Proof** For any presheaf of sets $X$ with a free $G$-action, the stacky quotient $X /\!\!/ G$ is objectwise weakly equivalent to the presheaf quotient $U \mapsto X(U)/G(U)$, which we temporarily denote by $X/_{\text{pre}} G$. Indeed, when the $G(U)$-action on $X(U)$ is free, the action groupoid $X(U)$ is canonically equivalent to the orbit set $X(U)/G(U)$, viewed as a discrete groupoid, so we obtain a weak equivalence when passing to nerves.

For a smooth $S$-scheme $X$ with a free $G$-action, we hence have an objectwise weak equivalence from $L_\tau(X /\!\!/ G)$ to the $\tau$-sheafification $(X/_{\text{pre}} G)^\tau$ of the presheaf quotient. Indeed, note first that the objectwise weak equivalence $X /\!\!/ G \to X/_{\text{pre}} G$ described above induces an objectwise weak equivalence $L_\tau(X /\!\!/ G) \to L_\tau(X/_{\text{pre}} G)$ (by two-out-of-three and Ken Brown's lemma). Secondly, as $(X/_{\text{pre}} G)^\tau$ is $\tau$-fibrant (by Lemma 3.1.4 and [1, Example 3.1.2]), the canonical map $X/_{\text{pre}} G \to (X/_{\text{pre}} G)^\tau$ factorizes through a morphism $L_\tau(X/_{\text{pre}} G) \to (X/_{\text{pre}} G)^\tau$, which also is an objectwise weak equivalence by [1, Example 3.1.2]. So altogether we obtain the objectwise weak equivalence claimed above. Finally, in the situation at hand, we have an isomorphism $(X/_{\text{pre}} G)^\tau \cong Y$. $\qquad\square$

We now provide a variant of [30, Definition 4.2.4].

**Definition 3.2.12** An *acceptable $G$-gadget* $(X_n)_{n \in I}$ *over $S$* is an acceptable gadget over $S$ satisfying

(i) for each $n \in I$, $X_n$ is endowed with a $G$-action; and

(ii) for each $n < m$ in $I$, the corresponding closed immersion $X_n \to X_m$ is $G$-equivariant.

The following result generalizes Lemma 3.2.11 from a single torsor to a sequences of torsors.

**Lemma 3.2.13** *Let $(X_n \to Y_n)_{n \in \mathbb{N}}$ be a sequence of $G$-torsors as in* (3.2.1), *with all torsors $\tau$-locally trivial and all maps $X_n \to X_{n+1}$ monomorphisms. Then the induced map*

$$L_\tau(X_\infty /\!\!/ G) \to Y_\infty$$

*is an objectwise weak equivalence. In particular, this holds if $(X_n)_{n \in \mathbb{N}}$ is an acceptable $G$-gadget with a $\tau$-locally trivial $G$-torsor $X_n \to Y_n$ for each $n \in \mathbb{N}$.*

**Proof** All objects are cofibrant, and the maps $X_n \to X_m$, $X_n /\!\!/ G \to X_m /\!\!/ G$ and $Y_n \to Y_m$ are monomorphisms, hence cofibrations in our model structure. So $X_\infty$, $\text{colim}(X_n /\!\!/ G)$ and $Y_\infty$ are colimits of cofibrations and thus homotopy colimits. The canonical maps $X_n /\!\!/ G \to Y_n$, which are $\tau$-local weak equivalences by Lemma 3.2.11, therefore induce a $\tau$-local weak equivalence $\text{colim}_n(X_n /\!\!/ G) \to Y_\infty$. Finally, we observe that the action groupoid of $X_\infty$ is the strict colimit of the action groupoids of the $X_n$, so that we have an isomorphism

$$\operatorname*{colim}_n(X_n /\!\!/ G) \cong X_\infty /\!\!/ G.$$

Hence, we obtain a $\tau$-local weak equivalence $X_\infty /\!\!/ G \to Y_\infty$. As $Y_\infty$ is $\tau$-fibrant by Proposition 3.1.3(b), this $\tau$-local weak equivalence factors through $L_\tau(X_\infty /\!\!/ G)$, and by Ken Brown's lemma we thus obtain the desired objectwise weak equivalence

$$L_\tau(X_\infty /\!\!/ G) \to Y_\infty.$$

If $(X_n)_{n \in \mathbb{N}}$ is an acceptable $G$-gadget with a $\tau$-locally trivial $G$-torsor $X_n \to Y_n$ for each $n \in \mathbb{N}$, the maps $X_n \to X_m$ are closed immersions and therefore monomorphisms by assumption. As $X_n(U) \to X_m(U)$ is a $G(U)$-equivariant monomorphism for every $U \in \mathbf{Sm}_X$, the induced maps $X_n(U)/G(U) \to X_m(U)/G(U)$ are also monomorphisms. We argue at the end of the proof of Lemma 3.2.11 that $Y_n$ and $Y_m$ differ from the presheaf quotients only by sheafification. As sheafification is exact, this shows that $Y_n \to Y_m$ is also a monomorphism. $\qquad \square$

We now verify a crucial generalized contractibility condition for acceptable $G$-gadgets. Given a $\tau$-locally trivial torsor $T \to Y$ in $\mathbf{Sm}_S$, consider the map $L_\tau(X_\infty \times_S T) \to Y$ given by applying first $(-/\!\!/ G)$ and then $L_\tau$ to the $G$-equivariant projection map $X_\infty \times_S T \to T$ and then composing with the map $L_\tau(T /\!\!/ G) \to Y$ of Lemma 3.2.11. This map is canonical up to the choice of $L_\tau$.

**Proposition 3.2.14** *Let $(X_n)_{n \in I}$ be an acceptable $G$-gadget over $S$. Then, for each $\tau$-locally trivial $G$-torsor $T \to Y$ in $\mathbf{Sm}_S$, the map*

$$L_\tau((X_\infty \times_S T) /\!\!/ G) \to Y$$

*is a motivic weak equivalence.*

**Proof** Apply Lemma 3.2.13 to the sequence of $\tau$-locally trivial $G$-torsors $X_n \times_S T \to X_n \times_S Y$ to obtain an objectwise weak equivalence

$$L_\tau((X_\infty \times_S T) /\!\!/ G) \to X_\infty \times_S Y.$$

This objectwise weak equivalence is, in particular, a motivic weak equivalence. As the map in the statement of the proposition can be obtained by composing this map with the projection map $p \colon X_\infty \times_S Y \to Y$, it suffices to show that $p$ is a motivic equivalence over $S$.

We know that $p$ is a motivic weak equivalence over $Y$ by Lemmas 3.2.7 and 3.2.8. We now use the functor $u_\sharp \colon L_{\mathrm{mot}} \, \mathbf{sPre}(\mathbf{Sm}_Y) \to L_{\mathrm{mot}} \, \mathbf{sPre}(\mathbf{Sm}_S)$ from [5, Proposition 4.5.4] for the structure map $u \colon Y \to S$ to deduce that $p$ is also a motivic weak equivalence over $S$. (These general results do not use the assumption on the "coefficients" discussed at the beginning of section 4.4 of loc cit. Compare also [20, Section 4.1] for a related discussion in the $\infty$-setting.) More precisely, $u_\sharp$ maps a smooth scheme over $Y$ to the same smooth scheme over $S$ by the proof of loc cit., and is a left Quillen functor for the projective model structures by [5, Theorem 4.5.14]. As left adjoints preserve colimits, $u_\sharp$ also maps the morphism $p \colon X_\infty \times_S Y \to Y$ over $Y$ to the same morphism considered over $S$. Hence, as $u_\sharp$ is left Quillen, $p$ is a motivic weak equivalence over $S$ for the projective motivic model structure, and thus also for the injective model structure, which has the same weak equivalences. $\qquad \square$

**Remark 3.2.15** There are definitions of $G$-torsors which are more general than [2, Definition 2.3.1], see, eg, [30, page 128], but which have in common that the conclusion of Lemma 3.2.11 holds. The proofs of Lemma 3.2.13 and Proposition 3.2.14 immediately generalize to these more general torsors provided the bases $Y_n$ are representable.

**Proposition 3.2.16** *For any acceptable $G$-gadget $(X_n)_{n \in I}$ over $S$, $L_\tau(X_\infty /\!\!/ G)$ is motivically equivalent to $L_\tau(S /\!\!/ G)$.*

**Proof** By [21, Lemma 2.1] it suffices to verify the condition established in Proposition 3.2.14. Though [21, Lemma 2.1] is only stated for the fppf-topology, it holds equally for any finer topology $\tau$, as is evident from its (very short) proof. Also, while [21, Lemma 2.1] is stated in terms of general $\infty$-$G$-torsors, for a simplicially discrete group $G$ any such torsor over a simplicially discrete base $X$ is again discrete, as it is locally isomorphic to $X \times_S G$. So in our setting it suffices to verify the assumptions for simplicially discrete $G$-torsors as in [2, Definition 2.3.1]. (Alternatively, we could use Remark 3.2.15.) Finally, the $\infty$-quotients appearing in [21, Lemma 2.1] coincide with our stacky quotients. See Lemmas 3.2.10 and 3.2.11 above for special cases, and [31, Section 3] for details in the general case. $\square$

The following theorem is the promised precise version of the "outline" at the beginning of this section.

**Theorem 3.2.17** *Let $(X_n)_{n \in \mathbb{N}}$ be an acceptable $G$-gadget such that we have a $\tau$-locally trivial $G$-torsor $X_n \to Y_n$ and induced monomorphisms $Y_n \to Y_m$ for all $n, m \in \mathbb{N}$. Then we have motivic weak equivalences*

$$Y_\infty \simeq \mathrm{B}_\tau G \simeq \mathrm{B}_{\mathrm{fppf}} G.$$

**Proof** As already explained in the outline, Lemma 3.2.13, Proposition 3.2.16 and Lemma 3.2.10 yield a chain of motivic weak equivalences $Y_\infty \simeq L_\tau(X_\infty /\!\!/ G) \simeq L_\tau(S /\!\!/ G) \simeq \mathrm{B}_\tau G$. As the same chain of equivalences also applies for any topology finer than $\tau$, we also obtain the motivic equivalence with $\mathrm{B}_{\mathrm{fppf}} G$. $\square$

As recalled in Example 3.2.4, for *special* affine algebraic group schemes $G$ any torsor is Zariski-locally trivial, so that Theorem 3.2.17 will yield motivic equivalences $Y_\infty \simeq \mathrm{B}G \simeq \mathrm{B}_{\mathrm{fppf}} G$.

## 3.3 Hermitian ASC for symplectic groups

Let $G$ be a linear algebraic group over a field $k$ of characteristic not two. In particular, $G$ is smooth and an étale sheaf on $\mathbf{Sm}_k$.

**Definition 3.3.1** We define $\mathrm{IO}_G$ and $\mathrm{IO}_G^\pm$ as kernels of restriction maps:

$$\mathrm{IO}_G := \ker(\mathrm{GW}^+(\mathbf{Rep}(G)) \to \mathrm{GW}^+(k)) \quad \text{and} \quad \mathrm{IO}_G^\pm := \ker(\mathrm{GW}^\pm(\mathbf{Rep}(G)) \to \mathrm{GW}^\pm(k)).$$

Our definition of $\mathrm{IO}_G$ agrees with the definition given in [38]. The graded ideal $\mathrm{IO}_G^\pm$ is not considered there.

**Lemma 3.3.2** *The $\mathrm{IO}_G^\pm$-adic topology on $\mathrm{GW}^\pm(\mathbf{Rep}(G))$ agrees with the $\mathrm{IO}_G$-adic topology.*

**Proof** This is a general fact about graded ideals. Consider a $\pm$-graded ring $R = R^+ \oplus R^-$ and a graded ideal $\mathfrak{a} = \mathfrak{a}^+ \oplus \mathfrak{a}^- \subset R$. Clearly $(\mathfrak{a}^+)^i \cdot R \subset \mathfrak{a}^i$. On the other hand, as $(\mathfrak{a}^-)^2 \subset \mathfrak{a}^+$, we find that

$$\mathfrak{a}^{2i} = (\mathfrak{a}^+ + \mathfrak{a}^-)^{2i} \subseteq \sum_{j=0}^{2i} (a^+)^{2i-j}(a^-)^j \subseteq (\mathfrak{a}^+)^i \cdot R^+ \oplus (\mathfrak{a}^+)^i \cdot R^-.$$

So $\mathfrak{a}^{2i} \subset (\mathfrak{a}^+)^i \cdot R$. This shows that the $\mathfrak{a}$-adic topology on $R$ agrees with the $\mathfrak{a}^+$-adic topology. □

Recall from [32; 33] the $\mathrm{Sp}_{2r}$-torsors $\mathrm{HU}(r, n)$ over the quaternionic Grassmannians $\mathrm{HGr}(r, n)$ associated with the tautological symplectic bundle, defined over a smooth quasiprojective base scheme $S$.

**Proposition 3.3.3** *The sequence of $\mathrm{Sp}_{2r}$-torsors $\mathrm{HU}(r, n) \to \mathrm{HGr}(r, n)$ with $n \in \mathbb{N}$ of [32, Proposition 8.5] defines an acceptable $\mathrm{Sp}_{2r}$-gadget $(\mathrm{HU}(r, n))_{n \in \mathbb{N}}$ over $S$.*

**Proof** It is shown in [32, Proposition 8.5] that $(\mathrm{HU}(r, n))_{n \in \mathbb{N}}$ is an acceptable gadget over $S$. It obviously also satisfies the conditions of an acceptable $\mathrm{Sp}_{2r}$-gadget, where the compatible actions come from the structure as $\mathrm{Sp}_{2r}$-torsors. □

**Theorem 3.3.4** (Panin–Walter) *There are motivic weak equivalences*

$$\mathrm{B}_{\mathrm{Nis}}\,\mathrm{Sp}_{2r} \simeq \mathrm{B}_{\text{ét}}\,\mathrm{Sp}_{2r} \simeq \mathrm{HGr}(r, \infty).$$

**Proof** Both equivalences are stated in [32, after Proposition 8.5], along with a brief indication on how to modify the arguments of [30] to obtain a proof, using their concept of an acceptable gadget. Section 3.2 provides more details for this argument. The first equivalence is evident from the fact that $\mathrm{Sp}_{2r}$ is special (see Example 3.2.4). The second is immediate from Theorem 3.2.17 applied to the acceptable gadget of Proposition 3.3.3. □

**Remark 3.3.5** The motivic equivalence $\mathrm{B}_{\text{ét}}\,\mathrm{Sp}_{2r} \simeq \mathrm{HGr}(r, \infty)$ of Theorem 3.3.4 was proven in [41, Proposition 5] for $r = \infty$ using a different technique. However, [41, Proposition 3], which is an ingredient of this alternative proof, has no obvious analogue for finite $r$.

**Remark 3.3.6** For the reader's convenience, let us compare the discussion above to some arguments in [30]. Theorem 3.3.4 is an analogue for $\mathrm{Sp}_{2r}$ of [30, Proposition 4.3.7] for $\mathrm{GL}_r$. In either case, the first equivalence is established by noting that $\mathrm{GL}_r$ and $\mathrm{Sp}_{2r}$, respectively, are special, and by using [30, Proposition 4.1.18]. The second equivalence rests on [30, Proposition 4.2.6], which has its parallel in our Proposition 3.2.16. Both of these intermediate results rely on the contractibility of colimits of gadgets. For admissible gadgets, this is proved in [30, Proposition 4.2.3], using the ambient vector bundles, while for acceptable gadgets this is Lemma 3.2.7 above. In [30], this contractibility enters via [30, Lemma 4.2.9]; see the proof of [30, Proposition 4.2.6] spelled out below [30, Lemma 4.2.9]. In our case, the contractibility

enters via the corresponding Proposition 3.2.14. Despite all these parallels, the definition of acceptable $G$-gadget employed here is significantly simpler than the definition of nice admissible $G$-gadget used by Morel and Voevodsky: there is no analogue of part (iii) of [30, Definition 4.2.4] in our Definition 3.2.12. The main reason why this simplification is possible is [21, Lemma 2.1].

For $\mathrm{Sp}_2$, Proposition 3.3.3 yields an acceptable gadget $(\mathrm{HU}(1, n+1))_{n \geq 1}$ built from $\mathrm{Sp}_2$-torsors $\mathrm{HU}(1, n+1) \to \mathrm{HP}^n$, where $\mathrm{HP}^n$ is the quaternionic projective space defined in [33]. In combination with Lemma 3.2.8(c), we moreover obtain acceptable gadgets $(\mathrm{HU}(1, n+1)^{\times r})_{n \geq 1}$ built from $\mathrm{Sp}_2^{\times r}$-torsors over products of quaternionic projective spaces $(\mathrm{HP}^n)^{\times r}$, for any $r \in \mathbb{N}$.

**Proposition 3.3.7** *The following maps are isomorphisms*:

(i) *the canonical map* $\mathrm{GW}^{\pm}(\mathrm{B} \, \mathrm{Sp}_2) \to \lim_n \mathrm{GW}^{\pm}(\mathrm{HP}^n)$;

(ii) *the canonical map* $\mathrm{GW}^{\pm}(\mathrm{B} \, \mathrm{Sp}_2^{\times r}) \to \lim_n \mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r})$; *and*

(iii) *the canonical map* $\mathrm{GW}^{\pm}(\mathrm{B} \, \mathrm{Sp}_{2r}) \to \lim_n \mathrm{GW}^{\pm}(\mathrm{HGr}(r, n))$.

**Proof** By [32, Theorem 13.4], bigraded GW-theory is represented by a commutative ring spectrum BO, which is $(8, 4)$-periodic by [32, Theorem 7.5]. Restricting BO to bidegrees $(0, 0)$ and $(4, 2)$ yields the commutative ring $\mathrm{GW}^{\pm}(X)$ for motivic spaces $X$ which are not schemes. For an acceptable $G$-gadget $(X_n)_{n \in \mathbb{N}}$ such that we have compatible $G$-torsors $X_n \to Y_n$ for each $n$, we may consider the motivic Milnor exact sequences (see, eg, [34, Theorem 5.7]):

$$0 \to \lim_n^1 \mathrm{BO}^{-1,0}(Y_n) \to \mathrm{GW}^+(Y_\infty) \to \lim_n \mathrm{GW}^+(Y_n) \to 0,$$

$$0 \to \lim_n^1 \mathrm{BO}^{3,2}(Y_n) \to \mathrm{GW}^-(Y_\infty) \to \lim_n \mathrm{GW}^-(Y_n) \to 0.$$

These yield a canonical isomorphisms $\mathrm{GW}^{\pm}(Y_\infty) \to \lim_n \mathrm{GW}^{\pm}(Y_n)$ provided the $\lim^1$-terms vanish. When $Y_n = \mathrm{HGr}(r, n)$, they do vanish, as [32, Theorem 9.5] provides the necessary surjections of BO-groups, thus proving (i) and (iii). To prove (ii), we argue similarly, using [32, Theorem 9.5] inductively by viewing $(\mathrm{HP}^n)^{\times r}$ as a trivial quaternionic projective bundle over $(\mathrm{HP}^n)^{\times r-1}$. $\square$

In view of this proposition, the reader preferring motivic spaces to actual varieties may replace several $\lim_n \mathrm{GW}^{\pm}(\mathrm{HP}^n)$ below by $\mathrm{GW}^{\pm}(\mathrm{B}_{\text{ét}} \, \mathrm{Sp}_2)$ or $\mathrm{GW}^{\pm}(\mathrm{BSp}_2)$.

We are now ready to examine the case $G = \mathrm{Sp}_2$ in detail. The gadget $(\mathrm{HU}(1, n+1))_{n \geq 1}$ has nice formal properties for *Borel classes*. Let $(\mathcal{U}_n, \psi_n)$ be the tautological rank-2 symplectic bundle on $\mathrm{HP}^n$. It defines an element in $\mathrm{GW}^-(\mathrm{HP}^n)$. We define the universal first Borel class

$$b_1 = (b_{1,(n)})_{n \in \mathbb{N}} \in \lim_n \mathrm{GW}^-(\mathrm{HP}^n)$$

to be the limit of the first Borel classes $b_{1,(n)} \in \mathrm{GW}^-(\mathrm{HP}^n)$, where $b_{1,(n)} = b_1(\mathcal{U}_n, \psi_n)$ is the first Borel class of $(\mathcal{U}_n, \psi_n)$ on $\mathrm{HP}^n$, as defined in [33, Definition 8.3]. By [32, Proposition 9.9], we have

$$b_{1,(n)} = [\mathcal{U}_n, \psi_n] - H^- \in \mathrm{GW}^-(\mathrm{HP}^n),$$

where $H^-$ is the symplectic hyperbolic space of rank 2 with trivial $\mathrm{Sp}_2$-action. Let $(V, \phi)$ be a trivial symplectic vector space of rank $2n + 2$ in which $(V_0, \phi_0)$ is a distinguished rank-2 symplectic subspace, and write $\mathrm{Sp}_{2n+2} = \mathrm{Sp}(V, \phi)$. In this case $(V, \phi)$ is the standard representation of $\mathrm{Sp}_{2n+2}$ discussed in Example 2.3.15. There is a decomposition

$$(V, \phi) \cong (V_0, \phi_0) \perp (V_0^\perp, \phi_0^\perp).$$

Then $\mathrm{Sp}_{2n+2}$ acts on $\mathrm{HP}^n = \mathrm{HP}(V, \phi)$, and the stabilizer of the distinguished point $(V_0, \phi_0)$ in $\mathrm{HP}^n$ is given by those $a \in \mathrm{Sp}(V, \phi)$ that fix $(V_0, \phi_0)$ and $(V_0^\perp, \phi_0^\perp)$, and is therefore given by $\mathrm{Sp}_{2n} \times \mathrm{Sp}_2$, where we make the identifications $\mathrm{Sp}_{2n} = \mathrm{Sp}(V_0^\perp, \phi_0^\perp)$ and $\mathrm{Sp}_2 = \mathrm{Sp}(V_0, \phi_0)$. Under the identification

$$\mathrm{HP}^n \cong \mathrm{Sp}_{2n+2} / (\mathrm{Sp}_{2n} \times \mathrm{Sp}_2),$$

the canonical $\mathrm{Sp}_2$-torsor $\mathrm{HU}(1, n+1)$ is given by $\mathrm{Sp}_{2n+2} / \mathrm{Sp}_{2n}$, which corresponds to the universal bundle $(\mathcal{U}_n, \psi_n)$ under the correspondence between $\mathrm{Sp}_2$-torsors and rank-2 symplectic bundles as outlined in, for instance, [2, Section 3.3, page 1025]. We let $X_n$ denote $\mathrm{HU}(1, n+1)$ for ease of notation. Let $\pi_n \colon X_n \to \mathrm{Spec}\, k$ be the $G$-equivariant structure map. Since the pullback functor $\gamma_n^* \colon \mathrm{Vect}(\mathrm{HP}^n) \to \mathrm{Vect}^G(X_n)$ along the projection $\gamma_n \colon X_n \to \mathrm{HP}^n$ is an equivalence of exact categories with duality, $\gamma_n^*$ induces a canonical isomorphism $\gamma_n^* \colon \mathrm{GW}^\pm(\mathrm{HP}^n) \to \mathrm{GW}_G^\pm(X_n)$. The following lemma studies the Atiyah–Segal map $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2)) \to \mathrm{GW}^\pm(B\mathrm{Sp}_2)$ induced by approximations $\pi_n \colon X_n \to \mathrm{Spec}\, k$ of the pullback along $E\mathrm{Sp}_2 \to \mathrm{Spec}\, k$.

**Proposition 3.3.8** *Let $H^-$ be the trivial symplectic plane bundle equipped with the trivial $\mathrm{Sp}_2$-action and $[V_0, \phi_0]$ the $\mathrm{Sp}_2$-representation from above. The composition*

$$(\gamma_n^*)^{-1} \pi_n^* \colon \mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2)) \to \mathrm{GW}_{\mathrm{Sp}_2}^\pm(X_n) \to \mathrm{GW}^\pm(\mathrm{HP}^n)$$

*sends $[V_0, \phi_0] - H^-$ to the (first) Borel class $b_{1,(n)} = [\mathcal{U}_n, \psi_n] - H^-$.*

**Proof** The pullback map $\pi_n^* \colon \mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2)) \to \mathrm{GW}_{\mathrm{Sp}_2}^\pm(X_n)$ sends $[V_0, \phi_0]$ to the class $[V_n, \phi_n]$ of the trivial rank-2 symplectic bundle on $X_n$ with the standard $\mathrm{Sp}_2$-action. Since $(\mathcal{U}_n, \psi_n)$ is a rank-2 symplectic bundle corresponding to the torsor $X_n = \mathrm{Symp}_{(\mathcal{U}_n, \psi_n)}$, its pullback to $X_n$ is also the trivial symplectic bundle $[V_n, \phi_n]$ with the standard $\mathrm{Sp}_2$-action. As $X_n \to \mathrm{HP}^n$ is an $\mathrm{Sp}_2$-torsor, the map

$$\mathrm{Sp}_2 \times_k X_n \to X_n \times_{\mathrm{HP}^n} X_n$$

given on points by $(a, x) \mapsto (ax, x)$ is an isomorphism. Hence,

$$\mathrm{Aut}_{\mathrm{Symp}}(H_{X_n}^-) \cong \mathrm{Sp}_2 \times_k X_n \cong \gamma_n^* \mathrm{Symp}_{(\mathcal{U}_n, \psi_n)} \cong \mathrm{Symp}_{(f^* \mathcal{U}_n, f^* \psi_n)}$$

and we deduce

$$\pi_n^*([V_0, \phi_0] - H^-) = [V_n, \phi_n] - H^- = \gamma_n^*([\mathcal{U}_n, \psi_n] - H^-). \qquad \square$$

As an immediate corollary, we obtain Atiyah–Segal completion for $G = \mathrm{Sp}_2$ and $\mathrm{GW}^\pm$.

**Corollary 3.3.9** *For $G = \mathrm{Sp}_2$, the map $\mathrm{GW}^{\pm}(\mathbf{Rep}(G)) \to \lim_n \mathrm{GW}^{\pm}(\mathrm{HP}^n)$ from Proposition 3.3.8 above is a completion of $\mathrm{GW}^{\pm}(\mathbf{Rep}(G))$ with respect to the Hermitian augmentation ideal $\mathrm{IO}_G$.*

**Proof**　The computation (2.4.3) immediately implies that $\mathrm{IO}_{\mathrm{Sp}_2}^{\pm}$ is generated by $[V_0, \phi_0] - H^-$. By [32, Theorem 9.5],

$$\lim_n \mathrm{GW}^{\pm}(\mathrm{HP}^n) \cong \mathrm{GW}^{\pm}(k)[\![b_1]\!].$$

Hence the claim follows from Proposition 3.3.8 and Lemma 3.3.2. $\qquad\square$

Now we prove Atiyah–Segal completion for general $\mathrm{Sp}_{2r}$. We will use the diagram

(3.3.1)

$$
\begin{array}{ccc}
\mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_{2r})) & \xrightarrow{\ \mathrm{res}\ } & \mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_2^{\times r})) \\
 & & \downarrow \\
\mathrm{GW}^{\pm}(B\mathrm{Sp}_{2r}) & \xrightarrow{\hspace{2cm}} & \mathrm{GW}^{\pm}(B\mathrm{Sp}_2^{\times r})
\end{array}
$$

in which the vertical arrow is a completion by Proposition 3.3.12. We start with the following result of Panin and Walter.

**Proposition 3.3.10** *For $1 \le i \le r$, let $y_i \in \lim_n \mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r})$ be the element defined by the inverse limit*

$$\lim_{n \in \mathbb{N}} b_1(\mathcal{U}_n^{(i)}, \phi_n^{(i)})$$

*of the first Borel classes $b_1(\mathcal{U}_n^{(i)}, \phi_n^{(i)})$ of the $i^{\text{th}}$ tautological rank-2 bundle on $(\mathrm{HP}^n)^{\times r}$. Then*

$$\lim_n \mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r}) = \mathrm{GW}^{\pm}(k)[\![y_1, \dots, y_r]\!].$$

**Proof**　For $r = 1$, this is a consequence of [33, Section 11] or [32, Theorem 9.5] as already recalled above. Note that $(\mathrm{HP}^n)^{\times r} \to (\mathrm{HP}^n)^{\times r - 1}$ is a trivial $\mathrm{HP}^n$-bundle, so by [32, Theorem 9.4],

$$\mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r}) \cong \frac{\mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r - 1})[b_1(\mathcal{U}_n^{(r)}, \phi_n^{(r)})]}{(b_1(\mathcal{U}_n^{(r)}, \phi_n^{(r)})^n)}.$$

Iterating, we obtain

$$\mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r}) \cong \frac{\mathrm{GW}^{\pm}(k)[b_1(\mathcal{U}_n^{(1)}, \phi_n^{(1)}), \dots, b_1(\mathcal{U}_n^{(r)}, \phi_n^{(r)})]}{(b_1(\mathcal{U}_n^{(1)}, \phi_n^{(1)})^n, \dots, b_1(\mathcal{U}_n^{(r)}, \phi_n^{(r)})^n)}$$

and it follows that

$$\lim_n \mathrm{GW}^{\pm}((\mathrm{HP}^n)^{\times r}) = \mathrm{GW}^{\pm}(k)[\![y_1, \dots, y_r]\!]. \qquad\square$$

For $n \in \mathbb{N}$, let $f_n \colon (\mathrm{HP}^n)^{\times r} \to \mathrm{HGr}(r, rn)$ be the canonical map such that the pullback of the tautological rank-$2r$ symplectic bundle on $\mathrm{HGr}(r, rn)$ is the orthogonal sum of the rank-2 symplectic bundles $(\mathcal{U}_n^{(i)}, \phi_n^{(i)})$ on $(\mathrm{HP}^n)^{\times r}$; such a map exists by the universal property of quaternionic Grassmannians discussed in [33, 10–11]. Note that $(\mathrm{HGr}(r, rn))_{n \in \mathbb{N}}$ is an acceptable gadget by Lemma 3.2.8. Recall also that we have $\lim_n \mathrm{GW}^{\pm}(\mathrm{HGr}(r, rn)) \cong \mathrm{GW}^{\pm}(k)[\![b_1, \dots, b_r]\!]$ by [33, Theorem 11.4].

**Theorem 3.3.11** *Let $b_i \in \lim_n \mathrm{GW}^\pm(\mathrm{HGr}(r, rn))$ be the element defined by the inverse limit*

$$b_1(\mathcal{U}_{r,rn}, \phi_{r,rn})_{n \in \mathbb{N}}$$

*of Borel classes of the tautological rank-$2r$ bundle on $\mathrm{HGr}(r, rn)$. The limit*

$$f^* \colon \lim_n \mathrm{GW}^\pm(\mathrm{HGr}(r, rn)) \to \lim_n \mathrm{GW}^\pm((\mathrm{HP}^n)^{\times r})$$

*of the pullback maps $f_n^*$ sends $b_i$ to the $i^{\mathrm{th}}$ symmetric elementary polynomial in the variables $y_j$ defined in Proposition 3.3.10.*

**Proof** The $i^{\mathrm{th}}$ tautological symplectic bundle of rank 2 on $(\mathrm{HP}^n)^{\times r}$ is an orthogonal direct summand of $f_n^*(\mathcal{U}_{r,rn}, \phi_{r,rn})$ by definition of $f_n$. We can now run the same argument as in the first part of the proof of [33, Theorem 10.2] to show that the image of $b_i$ is the $i^{\mathrm{th}}$ symmetric elementary polynomial in the variables $y_j$. $\qquad\square$

Recall from Proposition 3.3.7 that we have an isomorphism

$$\mathrm{GW}^\pm(\mathrm{BSp}_2^{\times r}) \cong \lim_n \mathrm{GW}^\pm((\mathrm{HP}^n)^{\times r})$$

with the ring structure given by Proposition 3.3.10. Considering the composition $(\gamma_n)^{-1} \circ \pi_n^*$ and $n \to \infty$ for $G = \mathrm{Sp}_2^{\times r}$ and $G = \mathrm{Sp}_{2r}$, we obtain morphisms of $\mathrm{GW}^\pm(k)$-algebras $\mathrm{GW}^\pm(\mathbf{Rep}(G)) \to \mathrm{GW}^\pm(\mathrm{B}G)$ as well, both generalizing (take $r = 1$) the map of Corollary 3.3.9.

**Proposition 3.3.12** *The map of $\mathrm{GW}^\pm(k)$-algebras*

$$\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2^{\times r})) \to \mathrm{GW}^\pm(\mathrm{BSp}_2^{\times r})$$

*defined above exhibits $\mathrm{GW}^\pm(\mathrm{BSp}_2^{\times r})$ as the completion of $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$ with respect to $\mathrm{IO}_{\mathrm{Sp}_2^{\times r}}$.*

**Proof** By Proposition 3.3.8, the generator $b^{(i)} \in \mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$ of (2.4.9) is mapped to the Borel class $y_i \in \mathrm{GW}^\pm(\mathrm{BSp}_2^{\times r})$ of Proposition 3.3.10 for each $1 \le i \le r$. Since $\mathrm{IO}_{\mathrm{Sp}_2^{\times r}} \subset \mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$ is generated, as an ideal, by the classes $b^{(i)}$, the result follows. $\qquad\square$

We are now ready to prove the following Atiyah–Segal completion result for Hermitian K-theory and symplectic groups:

**Corollary 3.3.13** *The map of $\mathrm{GW}^\pm(k)$-algebras*

$$\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2r})) \to \mathrm{GW}^\pm(\mathrm{BSp}_{2r})$$

*defined above exhibits $\mathrm{GW}^\pm(\mathrm{BSp}_{2r})$ as the completion of $\mathrm{GW}^\pm(\mathbf{Rep}(\mathrm{Sp}_{2r}))$ with respect to $\mathrm{IO}_{\mathrm{Sp}_{2r}}$.*

**Proof** By Corollary 2.4.4, the map $\mathrm{res}\colon \mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_{2r})) \to \mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$ from the upper line of (3.3.1) is injective and maps (higher) Borel classes to elementary symmetric polynomials in the generators $b^{(i)} \in \mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$. Similarly, by Theorem 3.3.11 and Proposition 3.3.7, the map $\mathrm{GW}^{\pm}(\mathrm{BSp}_{2r}) \to \mathrm{GW}^{\pm}(\mathrm{BSp}_2^{\times r})$ from (3.3.1) is injective and maps (higher) Borel classes to elementary symmetric polynomials in the generators $y_i \in \mathrm{GW}^{\pm}(\mathrm{BSp}_2^{\times r})$. It follows that the image of $\mathrm{IO}_{\mathrm{Sp}_{2r}}$ in $\mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$ is

$$\mathrm{IO}_{\mathrm{Sp}_{2r}} = \mathrm{IO}_{\mathrm{Sp}_2^{\times r}} \cap \mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_{2r}))$$

if we consider $\mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_{2r}))$ as a subalgebra of $\mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_2^{\times r}))$ via the restriction map. Thus it follows from Proposition 3.3.12 that the map

$$\mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_{2r})) \to \mathrm{GW}^{\pm}(\mathrm{BSp}_{2r})$$

exhibits $\mathrm{GW}^{\pm}(\mathrm{BSp}_{2r})$ as the completion of $\mathrm{GW}^{\pm}(\mathbf{Rep}(\mathrm{Sp}_{2r}))$ with respect to $\mathrm{IO}_{\mathrm{Sp}_{2r}}$. $\square$

### 3.4 Classifying space for multiplicative group with a nontrivial involution

Consider the multiplicative group $\mathbb{G}_m$ with the involution $\iota\colon t \mapsto t^{-1}$, as in Example 2.0.1. In this section, we show how to approximate the classifying space $\mathrm{B}\mathbb{G}_m$ in a way that is compatible with the involution. That is, we will construct $\mathbb{G}_m$-torsors $U_n \to B_n$ such that the torsors $U_n$ form an acceptable gadget, and are equipped with involutions $\iota\colon U_n \to U_n$ that are compatible with the $\mathbb{G}_m$-action in the sense that

$$(3.4.1) \qquad\qquad \iota(t.x) = \iota(t).\iota(x)$$

for $t \in \mathbb{G}_m$, $x \in U_n$. This condition ensures, in particular, that the involution on $U_n$ descends to an involution on $B_n$.

**Remark 3.4.1** Usually, $\mathrm{B}\mathbb{G}_m$ is approximated by the projective spaces $\mathbb{P}^n$. We have $\mathbb{G}_m$-torsors $\mathbb{A}^{n+1} \setminus 0 \to \mathbb{P}^n$, and these torsors form an acceptable gadget. However, there seems to be no involution $\iota$ on $\mathbb{A}^{n+1} \setminus 0$ satisfying (3.4.1).

Concretely, we will use the principal $\mathbb{G}_m$-torsors

$$U_n := \{(x, y) \in \mathbb{A}^{n+1} \times \mathbb{A}^{n+1} \mid x^T y = 1\}$$

with $\mathbb{G}_m$-action $t.(x, y) := (tx, t^{-1}y)$ and involution $\iota\colon (x, y) \mapsto (y, x)$. This involution clearly satisfies (3.4.1). Quotienting by the $\mathbb{G}_m$-action, we obtain the following open subschemes of $\mathbb{P}^n \times \mathbb{P}^n$:

$$B_n := \{([x], [y]) \in \mathbb{P}^n \times \mathbb{P}^n \mid x^T y \neq 0\}$$

The induced involution on $B_n$ is given by $\iota\colon ([x], [y]) \mapsto ([y], [x])$.

**Remark 3.4.2** To verify that the obvious projection $U_n \to B_n$ is a principal $\mathbb{G}_m$-torsor, we can identify it with the canonical projection

$$\mathrm{GL}_{1+n}/(1 \times \mathrm{GL}_n) \to \mathrm{GL}_{1+n}/(\mathbb{G}_m \times \mathrm{GL}_n)$$

via the map that sends the left coset represented by a matrix $A$ to the pair $(x, y)$ consisting of the first column of $A$, and of the first row of $A^{-1}$:

$$
\begin{array}{ccc}
\frac{\mathrm{GL}_{1+n}}{1 \times \mathrm{GL}_n} & \xrightarrow{\cong} & U_n \\
\downarrow \mathbb{G}_m & & \downarrow \mathbb{G}_m \\
\frac{\mathrm{GL}_{1+n}}{\mathbb{G}_m \times \mathrm{GL}_n} & \xrightarrow{\cong} & B_n
\end{array}
$$

The involutions on $U_n$ and $B_n$ are induced by the involution $A \mapsto (A^{-1})^T$ on $\mathrm{GL}_{1+n}$.

**Remark 3.4.3** The fixed points of the scheme $B_n$ under the involution can be identified with the scheme $\{[x] \in \mathbb{P}^n \mid x^T x \neq 0\}$, ie with the complement of a quadric in $\mathbb{P}^n$. This complement is often used as an algebro-geometric replacement for real projective space $\mathbb{RP}^n$, for example in [14] or [48]. It is a special case of the algebraic orthogonal Grassmannians of Schlichting and Tripathi [41].

**Lemma 3.4.4** *The schemes $U_n$ with the obvious inclusions $U_n \subset U_{n+1} \subset \cdots$ form an acceptable gadget in the sense of Definition 3.2.6.*

**Proof** Let $R$ be an arbitrary commutative ring, and $g \in R$. As in the proof of [32, Proposition 8.5], it suffices to show that, given an arbitrary morphism $\mathrm{Spec}(R/g) \to U_n$, we can fill in the dashed arrow in the diagram

$$
\begin{array}{ccc}
\mathrm{Spec}(R/g) & \longrightarrow & U_n \hookrightarrow U_{n+1} \\
\downarrow & \nearrow & \\
\mathrm{Spec}(R) & &
\end{array}
$$

The given arrow corresponds to a tuple $(\bar{a}_0, \ldots, \bar{a}_n, \bar{b}_0, \ldots, \bar{b}_n)$ in $(R/g)^{2n+2}$ such that $\sum_i \bar{a}_i \bar{b}_i = 1$ in $R/g$. Pick an arbitrary lift $(a_0, \ldots, a_n, b_0, \ldots, b_n) \in R^{2g+2}$. Then the previous equality tells us that there exists some element $r \in R$ such that $\sum_i a_i b_i = 1 + gr$ in $R$. The composition from $\mathrm{Spec}(R/g)$ to $U_{n+1}$ corresponds to the tuple $(\bar{a}_0, \ldots, \bar{a}_n, \bar{0}, \bar{b}_0, \ldots, \bar{b}_n, \bar{0})$ in $(R/g)^{2n+4}$. In order to construct the dashed arrow in a way that the diagram commutes, we need to construct a lift of this tuple, say,

$$
(a'_0, \ldots, a'_n, gc, b'_0, \ldots, b'_n, gd)
$$

such that $\sum_i a'_i b'_i + g^2 cd = 1$. Pick $a'_i := a_i$, $b'_i := b_i - grb_i$, and $c := d := r$. A quick calculation shows that this choice fits the bill. $\qquad\square$

## 3.5 Tools for generalizations to base schemes with nontrivial group actions

Recall that if we have Atiyah–Segal completion for $\mathrm{GL}_n$ for schemes $X$ with arbitrary $\mathrm{GL}_n$-action, then applying it to $X = \mathrm{GL}_n / H$ yields the completion theorem for $H$ over a point $\mathrm{Spec}(k)$. The same applies to $\mathrm{Sp}_{2n}$, and hence either case would cover all $H$ which are, eg, split reductive, and in particular all finite groups.

We now briefly recall two techniques that have been successfully used to generalize results on schemes $X$ with trivial $G$-action to general $G$-schemes $X$. Both only apply to $G = T$ a torus, and both have the same underlying idea: under suitable assumptions, there is a big open $G$-subscheme $U$ of $X$ on which the action of $X$ has a very simple product description. Using this product description, we may prove the desired result for $U$, and then using a finite number of induction steps also for $X$, assuming that the equivariant cohomology theory we care about (here: equivariant Hermitian K-theory) satisfies a suitable equivariant localization theorem.

The first technique is to work with $T$-filtrable schemes for $G = T$ an "algebraic torus", and has been used, eg, by Brion and Krishna, and more recently by [45]. This goes back to Białynicki-Birula. The main ingredient is probably [6, Theorem 2.5], which states that $U$ is $T$-equivariantly isomorphic to $(U \cap X^T) \times V$, where $V$ is a finite-dimensional $T$-module.

The second technique uses the "torus generic slice theorem" of Thomason; see [46, Proposition 4.10]. Here $X$ is very general, and $G = T$ is a "diagonalizable torus". The main geometric result here is that we have a $T$-equivariant isomorphism $U \cong T/T' \times U/T$ with $T'$ a diagonalizable subtorus, and the induction is then done essentially on page 804 of loc cit.

Both techniques would be useful if we could generalize them from tori $T$ to $\mathrm{GL}_n$ or $\mathrm{Sp}_{2n}$, or to products of $\mathrm{Sp}_2$, or to tori with involution. We unfortunately don't have this yet. Still, the above method is expected to yield the completion theorem for subgroups of $T$, eg, products of groups of roots of unity $\mu_l$. If we assume that these roots $\mu_l$ are contained in the base field, we could deduce the completion theorem for the corresponding direct sums of (constant) cyclic groups.

# References

[1] **A Asok**, **M Hoyois**, **M Wendt**, *Affine representability results in $\mathbb{A}^1$-homotopy theory, I: Vector bundles*, Duke Math. J. 166 (2017) 1923–1953  MR

[2] **A Asok**, **M Hoyois**, **M Wendt**, *Affine representability results in $\mathbb{A}^1$-homotopy theory, II: Principal bundles and homogeneous spaces*, Geom. Topol. 22 (2018) 1181–1225  MR

[3] **M F Atiyah**, *Bott periodicity and the index of elliptic operators*, Quart. J. Math. Oxford Ser. 19 (1968) 113–140  MR

[4] **M F Atiyah**, **G B Segal**, *Equivariant K-theory and completion*, J. Differential Geometry 3 (1969) 1–18  MR

[5] **J Ayoub**, *Les six opérations de Grothendieck et le formalisme des cycles évanescents dans le monde motivique, II*, Astérisque 315, Soc. Math. France, Paris (2007)  MR

[6] **A Białynicki-Birula**, *Some theorems on actions of algebraic groups*, Ann. of Math. 98 (1973) 480–497  MR

[7] **N Bourbaki**, *Groupes et algèbres de Lie, Chapitres 4–6*, Actualités Scientifiques et Industrielles 1349, Hermann, Paris (1972)  MR

[8] **N Bourbaki**, *Lie groups and Lie algebras, Chapters 7–9*, Springer (2005)  MR

[9] **T Bröcker**, **T tom Dieck**, *Representations of compact Lie groups*, Graduate Texts in Mathematics 98, Springer, New York (1995) MR

[10] **B Calmès**, **J Fasel**, *Groupes classiques*, from "Autours des schémas en groupes, II" (B Edixhoven, P Gille, G Prasad, P Polo, editors), Panor. Synthèses 46, Soc. Math. France, Paris (2015) 1–133 MR

[11] **B Calmès**, **J Hornbostel**, *Witt motives, transfers and reductive groups*, preprint (2005) Available at https://www.math.uni-bielefeld.de/LAG/man/143.pdf

[12] **G Carlsson**, **R Joshua**, *Equivariant algebraic K-theory, G-theory and derived completions*, Adv. Math. 430 (2023) art. id. 109194 MR

[13] **B Conrad**, *Reductive group schemes*, from "Autour des schémas en groupes, I", Panor. Synthèses 42/43, Soc. Math. France, Paris (2014) 93–444 MR

[14] **A Delzant**, *Définition des classes de Stiefel–Whitney d'un module quadratique sur un corps de caractéristique différente de* 2, C. R. Acad. Sci. Paris 255 (1962) 1366–1368 MR

[15] **D Dugger**, **S Hollander**, **D C Isaksen**, *Hypercovers and simplicial presheaves*, Math. Proc. Cambridge Philos. Soc. 136 (2004) 9–51 MR

[16] **D Eisenbud**, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics 150, Springer (1995) MR

[17] **J Fasel**, **O Haution**, *The stable Adams operations on Hermitian K-theory*, Geom. Topol. 29 (2025) 127–169 MR

[18] **A Grothendieck**, *Torsion homologique et sections rationnelles*, Paris (1958) MR Available at https://eudml.org/doc/110353

[19] **J Hornbostel**, **M Wendt**, *Chow–Witt rings of classifying spaces for symplectic and special linear groups*, J. Topol. 12 (2019) 916–966 MR

[20] **M Hoyois**, *The six operations in equivariant motivic homotopy theory*, Adv. Math. 305 (2017) 197–279 MR

[21] **M Hoyois**, *Cdh descent in equivariant homotopy K-theory*, Doc. Math. 25 (2020) 457–482 MR

[22] **J E Humphreys**, *Linear algebraic groups*, Graduate Texts in Mathematics 21, Springer (1975) MR

[23] **M Huruguen**, *Special reductive groups over an arbitrary field*, Transform. Groups 21 (2016) 1079–1104 MR

[24] **J C Jantzen**, *Representations of algebraic groups*, 2nd edition, Mathematical Surveys and Monographs 107, Amer. Math. Soc., Providence, RI (2003) MR

[25] **N A Karpenko**, **A S Merkurjev**, *Chow filtration on representation rings of algebraic groups*, Int. Math. Res. Not. 2021 (2021) 6691–6716 MR

[26] **A Knizel**, **A Neshitov**, *Algebraic analogue of the Atiyah completion theorem*, Homology Homotopy Appl. 16 (2014) 289–306 MR

[27] **A Krishna**, *Equivariant cobordism of schemes*, Doc. Math. 17 (2012) 95–134 MR

[28] **A Krishna**, **C Ravi**, *Algebraic K-theory of quotient stacks*, Ann. K-Theory 3 (2018) 207–233 MR

[29] **A Merkurjev**, *Classification of special reductive groups*, Michigan Math. J. 72 (2022) 529–541 MR

[30] **F Morel**, **V Voevodsky**, $\mathbb{A}^1$*-homotopy theory of schemes*, Inst. Hautes Études Sci. Publ. Math. 90 (1999) 45–143 MR

[31]  **T Nikolaus**, **U Schreiber**, **D Stevenson**, *Principal ∞-bundles: general theory*, J. Homotopy Relat. Struct. 10 (2015) 749–801  MR

[32]  **I Panin**, **C Walter**, *On the motivic commutative ring spectrum* $\mathbb{BO}$, Algebra i Analiz 30 (2018) 43–96  MR

[33]  **I Panin**, **C Walter**, *Quaternionic Grassmannians and Borel classes in algebraic geometry*, Algebra i Analiz 33 (2021) 97–140  MR

[34]  **I Panin**, **C Walter**, *On the algebraic cobordism spectra* **MSL** *and* **MSp**, St. Petersburg Math. J. 34 (2023) 109–141  MR

[35]  **Z Reichstein**, **D Tossici**, *Special groups, versality and the Grothendieck–Serre conjecture*, Doc. Math. 25 (2020) 171–188  MR

[36]  **C Rezk**, *A Cartesian presentation of weak n-categories*, Geom. Topol. 14 (2010) 521–571  MR

[37]  **H Rohrbach**, *On Atiyah–Segal completion for Hermitian K-theory*, PhD thesis, Bergische Universität Wuppertal (2021)  Available at `https://tinyurl.com/DissertationRohrbach`

[38]  **H Rohrbach**, *On Atiyah–Segal completion for T-equivariant Hermitian K-theory*, preprint (2022)  arXiv 2203.15518

[39]  **H Rohrbach**, *The projective bundle formula for Grothendieck–Witt spectra*, J. Pure Appl. Algebra 226 (2022) art. id. 106917  MR

[40]  **H Rohrbach**, *Hermitian K-theory of Grassmannians*, preprint (2023)  arXiv 2309.14896

[41]  **M Schlichting**, **G S Tripathi**, *Geometric models for higher Grothendieck–Witt groups in* $\mathbb{A}^1$*-homotopy theory*, Math. Ann. 362 (2015) 1143–1167  MR

[42]  **J-P Serre**, *Espaces fibrés algébriques*, Séminaire Claude Chevalley 3 (1958) 1–37

[43]  **J-P Serre**, *Groupes de Grothendieck des schémas en groupes réductifs déployés*, Inst. Hautes Études Sci. Publ. Math. 34 (1968) 37–52  MR

[44]  *Stacks project* (2018)  Available at `http://stacks.math.columbia.edu`

[45]  **G Tabuada**, **M Van den Bergh**, *Motivic Atiyah–Segal completion theorem*, Int. Math. Res. Not. 2024 (2024) 3497–3550  MR

[46]  **R W Thomason**, *Comparison of equivariant algebraic and topological K-theory*, Duke Math. J. 53 (1986) 795–825  MR

[47]  **B Totaro**, *The Chow ring of a classifying space*, from "Algebraic *K*-theory", Proc. Sympos. Pure Math. 67, Amer. Math. Soc., Providence, RI (1999) 249–281  MR

[48]  **M Zibrowius**, *Witt groups of curves and surfaces*, Math. Z. 278 (2014) 191–227  MR

[49]  **M Zibrowius**, *Symmetric representation rings are λ-rings*, New York J. Math. 21 (2015) 1055–1092  MR

*Fakultät IV, Mathematik und Informatik, Bergische Universität Wuppertal*
*Wuppertal, Germany*
*Faculty of Mathematics, Universität Duisburg-Essen*
*Essen, Germany*
*Mathematisches Institut, Heinrich Heine University Düsseldorf*
*Düsseldorf, Germany*

`hornbost@uni-wuppertal.de`,  `hermanrohrbach@gmail.com`,  `marcus.zibrowius@hhu.de`

# A minimality property for knots without Khovanov 2-torsion

ONKAR SINGH GUJRAL

JOSHUA WANG

A conjecture of Shumakovitch states that every nontrivial knot has 2-torsion in its Khovanov homology. We show that if a knot $K$ has no 2-torsion in its Khovanov homology, then the rank of its reduced Khovanov homology is minimal among all knots obtainable from $K$ by a proper rational tangle replacement. It follows, for example, that unknotting number 1 knots have 2-torsion in their Khovanov homology.

Shumakovitch conjectured that every nontrivial knot has an element of order 2 in its Khovanov homology [10, Conjecture 1]. The conjecture has been verified for some infinite families of knots (see, for example, [2; 9; 10]) and has withstood large computational searches. In this note, we provide topological evidence for the conjecture, and we verify the conjecture for a large class of knots that include all unknotting number 1 knots.

Two links differ by a *rational tangle replacement* if they agree outside of a ball, and if within the ball, each is a rational tangle. A rational tangle replacement is *proper* if the arcs of the two rational tangles connect the same end points [4; 8]. Changing a crossing is an example of a proper rational tangle replacement, while resolving a crossing is an example of a nonproper rational tangle replacement. In the following statement, $\mathrm{Kh}(K)$ and $\overline{\mathrm{Kh}}(K)$ denote the unreduced and reduced Khovanov homology groups of $K$, respectively, thought of as abelian groups with bigradings suppressed.

**Theorem 1** *Suppose $K$ is a knot such that there is no 2-torsion in $\mathrm{Kh}(K)$. If $J$ is a knot that differs from $K$ by a proper rational tangle replacement, then*

$$\mathrm{rank}\,\overline{\mathrm{Kh}}(K) \le \mathrm{rank}\,\overline{\mathrm{Kh}}(J).$$

**Corollary 2** *Any knot whose unknotting number is 1 has 2-torsion in its Khovanov homology. More generally, if $K$ is a nontrivial knot that can be obtained from the unknot or a trefoil by a proper rational tangle replacement, then $\mathrm{Kh}(K)$ contains 2-torsion.*

**Proof of Corollary 2** Let $J$ be the unknot or a trefoil, and let $K$ be obtained from $J$ by a proper rational tangle replacement. If there is no 2-torsion in $\mathrm{Kh}(K)$, then $\mathrm{rank}\,\overline{\mathrm{Kh}}(K) \le \mathrm{rank}\,\overline{\mathrm{Kh}}(J) \le 3$ by Theorem 1. The rank of $\overline{\mathrm{Kh}}(K)$ cannot be 3 since then $K$ would be a trefoil [3, Theorem 1.4], which has 2-torsion in its Khovanov homology. Since the rank of $\overline{\mathrm{Kh}}(K)$ is odd, it must be 1, and so $K$ is the unknot [6]. □

Our proof of Theorem 1 combines the main result of Iltgen, Lewark, and Marino [4] with an observation of Kotelskiy, Watson, and Zibrowius [5, Proposition 9.3] using the following lemma.

**Lemma 3** *Let $\mathbb{F}$ be a field, and suppose $M$ and $N$ are finitely generated modules over the polynomial ring $\mathbb{F}[X]$ of the form*

$$M = (\mathbb{F}[X])^r \oplus \bigoplus_{i=1}^{m} \frac{\mathbb{F}[X]}{X^{a_i}}, \quad N = (\mathbb{F}[X])^s \oplus \bigoplus_{i=1}^{n} \frac{\mathbb{F}[X]}{X^{b_i}},$$

*where $r, m, s, n \geq 0$ and $a_1, \ldots, a_m, b_1, \ldots, b_n \geq 1$. Furthermore, suppose $f \colon M \to N$ and $g \colon N \to M$ are $\mathbb{F}[X]$-module maps for which $f \circ g = X$ and $g \circ f = X$. If the numbers $a_1, \ldots, a_m$ are all at least 2, then $m \leq n$.*

**Proof** Let $X_M$ and $X_N$ denote the structural maps $X \colon M \to M$ and $X \colon N \to N$, respectively. Our aim is to establish $m = \dim_{\mathbb{F}} \ker X_M \leq \dim_{\mathbb{F}} \ker X_N = n$.

Setting $C := g^{-1}(\ker X_M)$, we first claim that $g|_C \colon C \to \ker X_M$ is surjective. Since the numbers $a_1, \ldots, a_m$ are all at least two, any element $y$ in the kernel of $X_M$ lies in the image of $X_M$, and therefore may be written as $y = X_M z = g(f(z))$, which proves the claim. Next, note that $g$ sends $X_N C$ to zero, and so $g|_C$ induces a surjection $C/X_N C \to \ker X_M$. Thus

$$\dim_{\mathbb{F}} \ker X_M \leq \dim_{\mathbb{F}} C - \dim_{\mathbb{F}} X_N C = \dim_{\mathbb{F}} \ker(X_N|_C) \leq \dim_{\mathbb{F}} \ker X_N. \qquad \square$$

**Proof of Theorem 1** Let $\overline{\mathrm{BN}}(K)$ denote the reduced Bar-Natan homology of $K$ with rational coefficients. It is a rank-1 finitely generated graded module over $\mathbb{Q}[H]$ where $H$ has nonzero degree, so we may write

$$\overline{\mathrm{BN}}(K) \cong \mathbb{Q}[H] \oplus \bigoplus_{i=1}^{m} \frac{\mathbb{Q}[H]}{H^{a_i}}, \quad \overline{\mathrm{BN}}(J) \cong \mathbb{Q}[H] \oplus \bigoplus_{i=1}^{n} \frac{\mathbb{Q}[H]}{H^{b_i}},$$

where $a_1, \ldots, a_m, b_1, \ldots, b_n$ are positive. By hypothesis, there is no 2-torsion in $\mathrm{Kh}(K)$, so Proposition 9.3 of [5] implies that the numbers $a_1, \ldots, a_m$ are all at least 2. Furthermore, [5, Proof of Proposition 9.3] also gives $\mathrm{rk}\,\overline{\mathrm{Kh}}(K) = 1 + 2m$ and $\mathrm{rk}\,\overline{\mathrm{Kh}}(J) = 1 + 2n$.

By [4, Proof of Theorem 1.1], there are $\mathbb{Q}[H]$-module maps $f \colon \overline{\mathrm{BN}}(K) \to \overline{\mathrm{BN}}(J)$ and $g \colon \overline{\mathrm{BN}}(J) \to \overline{\mathrm{BN}}(K)$ satisfying $f \circ g = H$ and $g \circ f = H$. We note that the complex $[\![D]\!]$ over $\mathbb{Z}[G]$ associated to a diagram $D$ considered in [4] recovers the reduced Bar-Natan complex as $[\![D]\!] \otimes_{\mathbb{Z}[G]} \mathbb{Q}[H]$ where $\mathbb{Z}[G] \to \mathbb{Q}[H]$ sends $G$ to $-H$. By Lemma 3, we obtain

$$\mathrm{rk}\,\overline{\mathrm{Kh}}(K) = 1 + 2m \leq 1 + 2n = \mathrm{rk}\,\overline{\mathrm{Kh}}(J). \qquad \square$$

**Remark 4** Suppose $K$ is a knot such that $\overline{\mathrm{BN}}(K)$ does not contain $\mathbb{Q}[H]/H$ as a direct summand. Our proof of Theorem 1 implies that the conclusion of Theorem 1 holds for $K$. This observation gives evidence in favor of the affirmative for [5, Question 9.4], which suggests that the reduced Bar-Natan homology

of any nontrivial knot contains $\mathbb{Q}[H]/H$ as a direct summand. We note that the existence of such a summand in reduced Bar-Natan homology implies the existence of 2-torsion in Khovanov homology [5, Proposition 9.3]. An analogous question is raised in [1; 7] in the context of the Floer homology of rational homology spheres.

# References

[1] **A Alfieri**, **F Binns**, *Is the geography of Heegaard Floer homology restricted or the L-space conjecture false*?, preprint (2024) arXiv 2404.00490

[2] **M M Asaeda**, **J H Przytycki**, *Khovanov homology: torsion and thickness*, from "Advances in topological quantum field theory", NATO Sci. Ser. II Math. Phys. Chem. 179, Kluwer, Dordrecht (2004) 135–166 MR

[3] **J A Baldwin**, **S Sivek**, *Khovanov homology detects the trefoils*, Duke Math. J. 171 (2022) 885–956 MR

[4] **D Iltgen**, **L Lewark**, **L Marino**, *Khovanov homology and rational unknotting*, preprint (2021) arXiv 2110.15107

[5] **A Kotelskiy**, **L Watson**, **C Zibrowius**, *Immersed curves in Khovanov homology*, preprint (2019) arXiv 1910.14584

[6] **P B Kronheimer**, **T S Mrowka**, *Khovanov homology is an unknot-detector*, Publ. Math. Inst. Hautes Études Sci. 113 (2011) 97–208 MR

[7] **F Lin**, *A remark on taut foliations and Floer homology*, preprint (2023) arXiv 2309.01222

[8] **D McCoy**, **R Zentner**, *The Montesinos trick for proper rational tangle replacement*, Proc. Amer. Math. Soc. 151 (2023) 1811–1822 MR

[9] **J H Przytycki**, **R Sazdanović**, *Torsion in Khovanov homology of semi-adequate links*, Fund. Math. 225 (2014) 277–304 MR

[10] **A N Shumakovitch**, *Torsion of Khovanov homology*, Fund. Math. 225 (2014) 343–364 MR

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

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

onkar@mit.edu, joshuaxw@princeton.edu

# Local indicability of groups with homology circle presentations

AGUSTÍN NICOLÁS BARRETO

ELÍAS GABRIEL MINIAN

We investigate conditions that guarantee local indicability of groups that admit presentations with the homology of a circle, generalizing a result of J Howie for two-relator presentations. We apply our results to investigate local indicability of LOT groups and some classes of non-cycle-free Adian presentations, extending previous results in that direction by J Howie and D Wise.

20F05, 20F65, 57M05, 57M07, 57M10

## 1 Introduction

Recall that a group $G$ is locally indicable if each of its finitely generated nontrivial subgroups admits a homomorphism onto the infinite cyclic group $\mathbb{Z}$. This class of groups appeared in the work of Higman on the units of group rings [8] and has been intensively studied in the last four decades since the works of Brodskii [3] and Howie [9; 10]. The theory of locally indicable groups has connections with equations over groups [3; 9; 14], with complexes of nonpositive curvature and coherence [13; 16; 17; 18], and with dynamics and left orderable groups (see, for example, [4; 5; 15]). They are also related to the study of asphericity of 2-complexes and Whitehead's conjecture, since any connected 2-complex $X$ with $\pi_1(X)$ locally indicable and $H_2(X) = 0$ is aspherical (see [10; 11]). In fact, locally indicable groups are $\mathbb{Z}$-conservative [6; 12].

It is known that torsion-free one-relator groups are locally indicable [3]. Moreover one-relator products of locally indicable groups are locally indicable if the relator is not a proper power (see [10]). In [11, Theorem 6.2] Howie proved local indicability of groups admitting presentations

$$\mathcal{P} = \langle a, b, c \mid a^{-1} w_1^{-1} b w_1, \, b^{-1} w_2^{-1} c w_2 \rangle$$

with three generators and two relators, where $w_1$ is a specific type of word called a sloping word. These presentations arise naturally when studying (weak) labeled oriented trees (LOTs), and the result is used to prove local indicability of LOT groups for LOTs of diameter 3 and some families of LOTs of diameter 4 (see [11]).

One of the main results of this paper is a generalization of Howie's result [11, Theorem 6.2] to homology circle presentations of any length where the relators satisfy some extra hypotheses (weaker than Howie's

original result for two-relator presentations). It is clear that some extra condition is needed to derive local indicability, since one can construct homology circle presentations of nonlocally indicable groups just by adding a new generator to any balanced presentation of a perfect group. The condition that we require depends on the minima or maxima of the total exponents of the initial subwords of all but one of the relators (see Definitions 2.2 and 2.4).

**Theorem 2.5** *Let $\mathcal{P} = \langle a_1, \ldots, a_{k+2} \mid r_1, \ldots, r_k, s \rangle$ be a presentation (of deficiency 1) of a group $G$ with $H_1(G) = \mathbb{Z}$ (for some $k \geq 0$), where all the relators are cyclically reduced and have total exponent 0. If the multisets of minima $m(r_1), \ldots, m(r_k)$ are concatenable, then $G$ is locally indicable.*

The proof essentially follows Howie's original proof of [11, Theorem 6.2], adapted to a more general context. One can recover Howie's result as a particular case.

The condition on the relators having total exponent 0 can be dropped by considering any surjective homomorphism $\varphi \colon G \to \mathbb{Z}$ and the sequences of minima or maxima with respect to $\varphi$.

**Theorem 2.9** *Let $\mathcal{P} = \langle a_1, \ldots, a_{k+2} \mid r_1, \ldots, r_k, s \rangle$ be a presentation (of deficiency 1) of a group $G$ with $H_1(G) = \mathbb{Z}$ (for some $k \geq 0$), where all the relators are cyclically reduced. If there is a surjective homomorphism $\varphi \colon G \to \mathbb{Z}$ such that $m_\varphi(r_1), \ldots, m_\varphi(r_k)$ are concatenable, then $G$ is locally indicable.*

For two-relator presentations of deficiency 1 one can widely extend the result by using relative maxima or minima (see Definition 2.11).

**Corollary 2.12** *Let $\mathcal{P} = \langle a, b, c \mid r, s \rangle$ be a presentation of a group $G$ with $H_1(G) = \mathbb{Z}$ where the relators are cyclically reduced. Let $\varphi \colon G \to \mathbb{Z}$ be a surjective homomorphism with $\varphi(c) \neq 0$. If $r$ attains a unique relative minimum at $a$, then $G$ is locally indicable.*

In Theorem 2.18 we extend our methods to investigate presentations with the homology of a wedge of circles.

Our main applications are related to LOT groups and Adian presentations. Labeled oriented trees give rise to homology circle presentations that generalize Wirtinger presentations for knots. It is well known that knot groups are locally indicable (see [10]). Local indicability of LOT groups is an open problem (that would imply asphericity of the associated presentations). In fact, it is not even known whether all LOT groups are torsion-free. In [11] Howie associated to any LOT $\Gamma$ two graphs. The right graph $I(\Gamma)$ is obtained from $\Gamma$ by interchanging the initial vertex and the label of every edge in $\Gamma$. Analogously, the left graph $T(\Gamma)$ is constructed by interchanging the final vertex with the label. A similar construction was investigated by Gersten in [7] in the more general context of Adian presentations $\mathcal{P} = \langle A \mid u_i = v_i \ (i \in J) \rangle$ (where $u_i$ and $v_i$ are nontrivial positive words). Howie proved that, if $I(\Gamma)$ or $T(\Gamma)$ has no cycles, then the group $G(\Gamma)$ is locally indicable (and, in particular, $\Gamma$ is aspherical) [11, Theorem 10.1]. For

Adian presentations, Gersten showed that if both $T(\mathcal{P})$ and $I(\mathcal{P})$ have no cycles then the associated 2-complex $\mathcal{K}_{\mathcal{P}}$ is diagrammatically reducible and, in particular, aspherical [7, Proposition 4.12]. Moreover, when $\ell(u_i) = \ell(v_i)$ for every $i$ (as in the case of LOT presentations), if either $T(\mathcal{P})$ or $I(\mathcal{P})$ has no cycles then $\mathcal{K}_{\mathcal{P}}$ is diagrammatically reducible (and hence aspherical) [7, Proposition 4.15]. Here $\ell(w)$ denotes the length of the word $w$. More recently, Wise proved that the 2-complexes associated to cycle-free Adian presentations (ie those for which both graphs are forests) have nonpositive sectional curvature, which implies that the corresponding groups are locally indicable [17, Theorem 11.4]. For LOT presentations we obtain the following extension of Howie's result.

**Corollary 3.1** *Let $\Gamma$ be a LOT. If either $I(\Gamma)$ or $T(\Gamma)$ has at most one cycle, then $G(\Gamma)$ is locally indicable.*

In Corollary 3.3 we extend this result to a wider class of Adian presentations, generalizing, for a certain subfamily, Wise's result [17, Theorem 11.4].

Finally, Theorem 3.6 describes a strategy to study local indicability of LOT groups when both graphs $I(\Gamma)$ and $T(\Gamma)$ have more than one cycle. This result is a consequence of Theorem 2.18.

## Acknowledgements

## 2 Main results

The results of this section are based on some ideas and techniques developed by Jim Howie in the proof of [11, Theorem 6.2] and use well-known properties of locally indicable groups. Given a presentation $\mathcal{P}$ of a group $G = G(\mathcal{P})$, we denote by $\mathcal{K}_{\mathcal{P}}$ its standard 2-complex which has a single 0-cell, one 1-cell for each generator and one 2-cell for each relator. Recall from [10] that an elementary reduction $(X, Y)$ is a pair of 2-complexes such that $Y \subset X$ and $X - Y$ consists of exactly one 1-cell $e^1$ and at most one 2-cell $e^2$, where $e^2$ properly involves $e^1$, which means that its attaching map is not homotopic in $Y \cup e^1$ to a map into $Y$. A 2-complex $K$ is reducible if for every finite subcomplex $X$ there exists an elementary reduction $(X, Y)$. Also, a group presentation $\mathcal{P}$ is reducible if its associated 2-complex $\mathcal{K}_{\mathcal{P}}$ is. We will use the following known fact.

**Theorem 2.1** (Howie [10]) *If $G$ has a reducible presentation in which no relator is a proper power, then $G$ is locally indicable.*

Let $F$ be a free group on a set $A$. Recall that the total exponent of a word $S = a_1^{\varepsilon_1} \ldots a_n^{\varepsilon_n}$ (with $\varepsilon_i \in \{\pm 1\}$) is $\exp(S) = \varepsilon_1 + \cdots + \varepsilon_n$. The length of a word $S$ will be denoted by $\ell(S)$. An initial subword of a word $S = a_1^{\varepsilon_1} \ldots a_n^{\varepsilon_n}$ is a word of the form $a_1^{\varepsilon_1} \ldots a_i^{\varepsilon_i}$ for some $1 \leq i \leq n$. Suppose $\mathcal{P}$ is a finite presentation of a group $G$ such that the total exponents of all its relators are zero. Let $\varphi \colon G \to \mathbb{Z}$ be the surjective homomorphism which sends every generator of $\mathcal{P}$ to $1 \in \mathbb{Z}$. Let $\widetilde{\mathcal{K}}$ be the infinite cyclic cover of $\mathcal{K}_\mathcal{P}$ determined by $\varphi$. Similarly as in [10], we index its 0-cells with the integers. Via $\varphi \colon G \to \mathbb{Z}$ every generator $a$ in $\mathcal{P}$ determines one 1-cell $a_j$ oriented from $j$ to $j+1$ for each $j \in \mathbb{Z}$. Given a relator $r$, we denote by $R$ the corresponding 2-cell in $\mathcal{K}_\mathcal{P}$, and by $R_j$, the 2-cell in $\widetilde{\mathcal{K}}$ that covers $R$ and whose minimum traversed 0-cell is $j$.

For any relator $r$, the sequence of total exponents of its initial subwords represents the indices of the 0-cells traversed by the attaching path of the corresponding 2-cell in $\widetilde{\mathcal{K}}$ starting at vertex 0. Suppose that the minimum value attained in the sequence is $m \in \mathbb{Z}$. Note that this minimum can be attained more than once and that the rightmost letter of an initial subword where $m$ is attained must have exponent $-1$. This letter represents a 1-cell joining the vertices $m$ and $m + 1$ (and it is traversed backwards). The letters with exponent 1 where the value $m + 1$ is attained represent 1-cells oriented from $m$ to $m + 1$ and they are traversed forwards. Note that if we invert the relator these roles are interchanged. This motivates the following definition.

**Definition 2.2**  Let $r$ be a relator of a presentation $\mathcal{P}$ and let $m$ be the minimum value of the sequence of total exponents of its initial subwords. We define the *multiset of minima $m(r)$* as the multiset that contains the letters of $r$ with exponent $-1$ that attain the value $m$ and the letters that appear with exponent 1 in $r$ that attain the value $m + 1$.

For example, if $r = a^{-1}b^{-1}cb^{-1}ab$, the total exponents are $-1, -2, -1, -2, -1, 0$ and $m(r) = \{b, c, b, a\}$.

**Remark 2.3**  We can formulate an analogous definition for maxima instead of minima. In fact, every result in this paper that holds for minima has also an analogous version for maxima.

**Definition 2.4**  Given multisets $A_1, \ldots, A_n$ we say that they are *concatenable* if there is an ordering $A_{i_1}, \ldots, A_{i_n}$ such that, for every $1 \leq j \leq n$, there exists an element in $A_{i_j}$ with multiplicity 1 that does not belong to the union of $A_{i_1}, \ldots, A_{i_{j-1}}$.

The following result is our first generalization of Howie's [11, Theorem 6.2].

**Theorem 2.5**  *Let $\mathcal{P} = \langle a_1, \ldots, a_{k+2} \mid r_1, \ldots, r_k, s \rangle$ be a presentation (of deficiency 1) of a group $G$ with $H_1(G) = \mathbb{Z}$ (for some $k \geq 0$), where all the relators are cyclically reduced and have total exponent 0. If the multisets of minima $m(r_1), \ldots, m(r_k)$ are concatenable, then $G$ is locally indicable.*

We recover Howie's result [11, Theorem 6.2] as a particular case.

**Corollary 2.6**  (Howie [11])  *Let $\Gamma$ be a weakly labeled oriented tree with 3 vertices, in which at least one label is a reduced sloping word. Then $G(\Gamma)$ is locally indicable.*

The corollary follows from the fact that the standard presentation of the group of a weakly labeled oriented tree with three vertices $\mathcal{P} = \langle a, b, c \mid r, s \rangle$ is a homology circle and the sloping word condition assures that $m(r)$ attains a unique minimum or a unique maximum in its sequence. In particular, it is concatenable. In fact, for a two-relator presentation $\mathcal{P} = \langle a, b, c \mid r, s \rangle$ it is not required that $m(r)$ has a unique minimum in order to be concatenable. We just need that there exists a generator that appears only once in $m(r)$. In Corollary 2.12 we will show a generalization of this result for two-relator presentations.

The proof of Theorem 2.5 uses the following technical lemma.

**Lemma 2.7** *Let $F$ be a free group on a set $A$ and $T$ a cyclically reduced word in $F$ with $\exp(T) = 0$. Then $T$ cannot be factored as a product*

$$T = w_0 S_0 w_0^{-1} w_r S_r w_r^{-1} w_{r-1} S_{r-1} w_{r-1}^{-1} \ldots w_1 S_1 w_1^{-1}$$

*with $w_0, w_1, \ldots, w_r, S_0, S_1, \ldots, S_r \in F$ $(r \geq 0)$, satisfying*

- $\exp(T) = \exp(S_i) = 0$ for $0 \leq i \leq r$,
- $\exp(w_i) \geq 1$ for $1 \leq i \leq r$ and $\exp(w_0) = 1$,
- *the total exponents of all the initial subwords of $T$, $S_i$ and $w_i$ are nonnegative $(0 \leq i \leq r)$.*

**Proof** Suppose there exists such a factorization for some cyclically reduced word $T$. Over all possible factorizations of $T$ and $T^{-1}$ satisfying these conditions, we consider the cyclically reduced words $w_0, w_1, \ldots, w_r, S_0, S_1, \ldots, S_r$ that minimize the sum

$$L = \sum_{i=0}^{r} (\ell(w_i) + \ell(S_i)).$$

We can assume that the factorization that minimizes $L$ is one of $T$. Otherwise, we change $T$ by $T^{-1}$. We rewrite the equality as

$$w_0^{-1} T w_1 S_1^{-1} w_1^{-1} w_2 S_2^{-1} w_2^{-1} \ldots w_r S_r^{-1} w_r^{-1} w_0 = S_0.$$

Note that $S_i^{-1}$ satisfies the same properties as $S_i$ for every $0 \leq i \leq r$.

Let $c$ be the rightmost letter in $T$. Note that it must appear as $c^{-1}$ (with negative exponent) since the initial subwords of $T$ have nonnegative total exponent and $\exp(T) = 0$. The key observation is that this appearance of $c^{-1}$ in $T$ must cancel with a $c$, since, in any other case, $w_0^{-1} T$ would be an initial subword of $S_0$ with total exponent $-1$, which contradicts the hypotheses. There are five types of possible cancellations for $c^{-1}$.

**Case 1** (the rightmost letter $c^{-1}$ in $T$ cancels with a $c$ in some $S_i^{-1}$ with $i \geq 1$)  In this case we can write $S_i^{-1} = UcV$ for some (possibly empty) words $U$ and $V$. The word located between the letter $c^{-1}$ in $T$ and the $c$ in $S_i^{-1}$ must be trivial, so $w_1 S_1^{-1} w_1^{-1} \ldots w_{i-1} S_{i-1}^{-1} w_{i-1}^{-1} w_i U$ is trivial, and in particular this word has total exponent 0. Since the total exponent of $w_1 S_1^{-1} w_1^{-1} \ldots w_{i-1} S_{i-1}^{-1} w_{i-1}^{-1}$ is 0 and $\exp(w_i) \geq 1$, we have that $\exp(U) \leq -1$, but this is a contradiction since the subwords of $S_i^{-1}$ have nonnegative total exponent.

**Case 2** (the rightmost letter $c^{-1}$ in $T$ cancels with a $c$ in some $w_i = UcV$ with $i \geq 1$)  In this case, the hypothesis implies that $w_1 S_1^{-1} w_1^{-1} \ldots w_{i-1} S_{i-1}^{-1} w_{i-1}^{-1} U$ is trivial, so we can erase this subword from the original equality, replace $w_i^{-1} = V^{-1} c^{-1} U^{-1}$, and obtain

$$w_0^{-1} T c V S_i^{-1} V^{-1} c^{-1} U^{-1} w_{i+1} \ldots w_r S_r^{-1} w_r^{-1} w_0 = S_0.$$

Note that $\exp(cV) = 1$, since $\exp(w_i) = \exp(UcV) \geq 1$ and $\exp(U) = 0$. Also, the total exponents of the initial subwords of $cV$ are nonnegative because this property holds for $w_i$ and $\exp(U) = 0$. If $U$ is empty, we obtain a strictly shorter writing of $T$ where $cV$ plays the role of $w_i$, which is a contradiction. If $U$ is not empty, we can replace in the last equality $U^{-1} = w_1 S_1^{-1} w_1^{-1} \ldots w_{i-1} S_{i-1}^{-1} w_{i-1}^{-1}$ obtaining a new factorization for $T$ where $cV$ plays the role of $w_i$. Since $cV$ has a strictly shorter length than $w_i$, this would be a contradiction.

**Case 3** (the rightmost letter $c^{-1}$ in $T$ cancels with a $c$ in some $w_i^{-1} = V^{-1} c U^{-1}$ with $i \geq 1$)  This is similar to case 2.

**Case 4** (the rightmost letter $c^{-1}$ in $T$ cancels with a $c$ in $w_0^{-1} = V^{-1} c U^{-1}$)  This condition implies $c U^{-1} = T^{-1}$. After replacing, we get

$$V^{-1} w_1 S_1^{-1} w_1^{-1} \ldots w_r S_r^{-1} w_r^{-1} T V = S_0.$$

Since $\exp(w_0) = \exp(TV) = 1$ and $\exp(T) = 0$ we get that $\exp(V) = 1$. Also, the total exponents of initial subwords of $V$ are nonnegative because $w_0$ satisfies that condition and $\exp(T) = 0$. Taking inverses at both sides of the equality we get

$$V^{-1} T^{-1} w_r S_r w_r^{-1} \ldots w_1 S_1 w_1^{-1} V = S_0^{-1},$$

and since $T$ and the $S_i$ have total exponent 0, this provides a factorization of $T^{-1}$ satisfying the conditions, with a strictly smaller value of $L$, which is again a contradiction.

**Case 5** (the rightmost letter $c^{-1}$ in $T$ cancels with a $c$ in $w_0 = UcV$)  This is similar to case 4.  □

**Proof of Theorem 2.5**  We follow the proof of [11, Theorem 6.2]. If $k = 0$, $G$ is a torsion-free one-relator group and the result follows from [3] (see also [10, Corollary 4.3]). Now suppose that $k \geq 1$. Let $R_1, \ldots, R_k, S$ be the 2-cells of $\mathcal{K}_\mathcal{P}$ corresponding to $r_1, \ldots, r_k$ and $s$, respectively. Let $\varphi \colon G \to \mathbb{Z}$ be the (surjective) homomorphism that sends every generator to 1 and let $\widetilde{\mathcal{K}}$ be the corresponding infinite cyclic cover of $\mathcal{K}_\mathcal{P}$. It suffices to prove that $\pi_1(\widetilde{\mathcal{K}})$ is locally indicable.

We index the 0-cells of $\widetilde{\mathcal{K}}$ with the integers and denote $a_{i,j}$ the 1-cell that covers $a_i$ and joins the 0-cells $j$ and $j + 1$. It is oriented from $j$ to $j + 1$. We denote by $R_{i,j}$ and $S_j$ the 2-cells that cover $R_i$ and $S$, respectively, for which the minimum 0-cell is $j$. Since $\pi_1(\widetilde{\mathcal{K}})$ is the direct limit of the fundamental groups of its finite connected subcomplexes, it suffices to prove that for every finite connected subcomplex $K \subset \widetilde{\mathcal{K}}$, there is connected subcomplex $L \subset \widetilde{\mathcal{K}}$ with $K \subset L$ and $\pi_1(L)$ locally indicable. The strategy is to include $K$ in a finite subcomplex $L$ that is homotopy equivalent to a 2-complex $L'$, which in turn collapses to a reducible 2-complex and then use Theorem 2.1.

We define first the iterated rewrites of the cells $S_l$. Since $m(r_1), \ldots, m(r_k)$ are concatenable, we can assume without loss of generality that $a_i \in m(r_i)$ with multiplicity 1 and $a_i \notin \bigcup_{j=1}^{i-1} m(r_j)$ ($1 \le i \le k$). Equivalently, for every $1 \le i \le k$, $a_{i,0}$ appears exactly once in the attaching path of $R_{i,0}$ and does not appear in the attaching path of $R_{j,0}$ for any $j < i$. Also $a_{i,0}$ passes through the 0-cell 0 once. We can then write the attaching path of $R_{i,0}$ as $a_{i,0} p_{i,0}^{-1}$ where $p_{i,0}$ is a path which does not involve $a_{i,0}$. Let $S_0^{(1)}$ be the path obtained after replacing, for each $1 \le i \le k$ in decreasing order, every occurrence of $a_{i,0}$ in the attaching path of $S_0$ by $p_{i,0}$ and cyclically reducing. Note that $S_0^{(1)}$ does not pass through $a_{i,0}$ for any $1 \le i \le k$. Analogously, $a_{i,1}$ appears just once in the attaching path of $R_{i,1}$ and so it can be written as $a_{i,1} p_{i,1}^{-1}$ where $p_{i,1}$ is a path that does not involve $a_{i,1}$. Let $S_0^{(2)}$ be the path obtained after replacing in $S_0^{(1)}$ every occurrence of $a_{i,1}$ by $p_{i,1}$, for each $1 \le i \le k$ in decreasing order, and cyclically reducing. Denote $S_1^{(1)}$ the path obtained by doing the same in $S_1$. By repeating this procedure, we define the iterated rewrite $S_l^{(j)}$ for each $l \ge 0$ and $j \ge 1$.

Let $\alpha_1$ be the minimum 0-cell traversed by $S_0^{(1)}$. Note that $\alpha_1 \ge 0$. In general, for every $j$, we denote by $\alpha_j$ the minimum 0-cell of $S_0^{(j)}$. Note that $\alpha_j \ge \alpha_{j-1}$ and that the minimum 0-cell of $S_l^{(j)}$ is $\alpha_j + l$. We will show first that for $j \in \mathbb{N}$ big enough, the sequence $\{\alpha_j\}_{j \ge 1}$ stabilizes. Since the sequence is increasing, we only need to prove that it is bounded above. We show that $\alpha_n < \rho + \sigma + 1$ for every $n$, where $\rho = \sum_{i=1}^k \rho_i$ and $\rho_1, \ldots, \rho_k, \sigma$ are the maximum 0-cells traversed by the attaching paths of $R_{1,0}, \ldots, R_{k,0}, S_0$, respectively.

Suppose that there exists $j \in \mathbb{N}$ such that $\alpha_j \ge \rho + \sigma + 1$. Take $n \in \mathbb{N}$ big enough such that $\alpha_{n-\rho-\sigma} \ge \rho + \sigma + 1$. Let $Y$ be the full subcomplex of $\widetilde{\mathcal{K}}$ containing all 0-cells $j$ for $0 \le j \le n + \rho + \sigma$. This complex contains the 1-cells $a_{i,j}$ for each $1 \le i \le k+2$ and $0 \le j \le n + \rho + \sigma - 1$, the 2-cells $R_{i,j}$ for $1 \le i \le k$ and $0 \le j \le n + \sigma + \rho - \rho_i$ and the 2-cells $S_j$ with $0 \le j \le n + \rho$. Note that the number of 0-cells in $Y$ is $n + \rho + \sigma + 1$, the number of 1-cells is $(k+2)(n + \rho + \sigma)$, for each $1 \le i \le k$ it has $n + \sigma + \rho - \rho_i + 1$ 2-cells corresponding to the $R_{i,j}$ with $0 \le j \le n + \sigma + \rho - \rho_i$, and $n + \rho + 1$ 2-cells corresponding to the $S_j$ with $0 \le j \le n + \rho$. Therefore its Euler characteristic is $\chi(Y) = -\rho - \sigma + k + 2$.

We construct a 2-complex $Y'$ homotopy equivalent to $Y$ by replacing each $S_j$ by a 2-cell attached via the path $S_j^{(n-j)}$ ($0 \le j \le \rho + \sigma$). Note that this 2-complex collapses to a subcomplex $Y''$ by removing $R_{i,j}$ and $a_{i,j}$ for every $1 \le i \le k$ and $0 \le j \le \rho + \sigma$. Let $Z$ be the 1-subcomplex of $Y''$ that contains the 0-cells from 0 to $\rho + \sigma + 1$ and the 1-cells $a_{i,j}$ with $i = k+1, k+2$ and $0 \le j \le \rho + \sigma$. Since, for every $j$, the minimum 0-cell of $S_j^{(n-j)}$ is $j + \alpha_{n-j} \ge j + \alpha_{n-\rho-\sigma} \ge j + \rho + \sigma + 1$, the identity $Z \mapsto Z$ can be continuously extended to a retraction $Y'' \mapsto Z$ which sends all of $Y'' - Z$ to the 0-cell $\rho + \sigma + 1$. It follows that $\beta_1(Y) = \beta_1(Y'') \ge \beta_1(Z) = \rho + \sigma + 1$. Here $\beta_1$ denotes the first Betti number. Note that $\beta_0(Y) = 1$ since $Y$ is connected. Note also that $\mathcal{K}_\mathcal{P}$ is a homology circle because $H_1(\mathcal{K}_\mathcal{P}) = H_1(G) = \mathbb{Z}$ and the deficiency of the presentation is 1. This implies that $\beta_2(Y) = 0$ since $\widetilde{\mathcal{K}}$ is an infinite cyclic cover (see [1, Proposition 1]). It follows that $\chi(Y) \le -\rho - \sigma$, which is a contradiction. Therefore, the sequence $\{\alpha_j\}_{j \ge 1}$ is bounded above by $\rho + \sigma + 1$.

Now let $K \subset \widetilde{\mathcal{K}}$ be a finite connected subcomplex. We can assume that the minimum 0-cell in $K$ is 0. Let $q$ be the maximum 0-cell in $K$ and take $n$ large enough such that $\alpha_i = \alpha$ for every $i \geq n$. Take $L$ to be the subcomplex of $\widetilde{\mathcal{K}}$ containing every 0-cell between 0 and $q + n + \rho - \sigma$, every 1-cell $a_{i,j}$ with $1 \leq i \leq k+2$ and $0 \leq j \leq q+n+\rho-\sigma-1$, the 2-cells $R_{i,j}$ with $1 \leq i \leq k$ and $0 \leq j \leq q+n+\rho-\rho_i-\sigma$ and the 2-cells $S_j$ with $0 \leq j \leq q - \sigma$. Note that $K \subset L$.

In order to prove that $\pi_1(L)$ is locally indicable, we replace each $S_j$ by a 2-cell attached via $S_j^{(q+n-\sigma-j)}$ and obtain a new 2-complex $L'$ homotopy equivalent to $L$ which collapses to a 2-subcomplex $L''$ by removing every 2-cell $R_{i,j}$ and every 1-cell $a_{i,j}$ for each $1 \leq i \leq k$ and $0 \leq j \leq q + n + \rho - \rho_i - \sigma$. We will show that the fundamental group of the 2-complex $L''$ admits a reducible presentation. For $0 \leq j \leq q - \sigma$, the minimum 0-cell of $S_j^{(q+n-\sigma-j)}$ is $j + \alpha$. The 2-cell $S_j^{(q+n-\sigma-j)}$ involves any of the 1-cells in its attaching path incident to the 0-cell $j + \alpha$, and these 1-cells are not in the attaching paths of the remaining 2-cells $S_l^{(q+n-\sigma-l)}$ (for $l > j$). We take the quotient of $L''$ by a spanning tree in its 1-skeleton. The quotient can be seen as the 2-complex associated to a presentation $\mathcal{P}$ of $\pi_1(L'')$. In order to prove that this presentation is reducible it suffices to show that any subpresentation $\mathcal{P}' \subseteq \mathcal{P}$ with at least one relator has an elementary reduction. We take a subpresentation $\mathcal{P}'$ and consider the relator $s_l$ associated to the 2-cell $S_l^{(q+n-\sigma-l)}$ with minimum $l$. At least one of the 1-cells of $S_l^{(q+n-\sigma-l)}$ incident to the 0-cell $l + \alpha$ is not in the spanning tree and therefore it represents a generator $g$ of the presentation. Lemma 2.7 applied to $T = s_l$ (where the words $S_1, \ldots, S_r$ in Lemma 2.7 play the role of the remaining relators of $\mathcal{P}'$) shows that $s_l$ is not a conjugate in $H * \langle g \rangle$ to an element in the group $H$ whose presentation is obtained from $\mathcal{P}'$ by removing the generator $g$ and the relator $s_l$. This implies that $s_l$ properly involves $g$ (see [10, page 449]). It follows that $\pi_1(L) = \pi_1(L'')$ is locally indicable by Theorem 2.1. Note that the no-proper-power condition follows from [1, Proposition 1] and the fact that $\widetilde{\mathcal{K}}$ is an infinite cyclic cover of a homology circle 2-complex (see also [11, page 290]). $\qquad\square$

**Example 2.8** Consider
$$\mathcal{P} = \langle a,b,c,d \mid a^{-1}c^{-1}b^{-1}a^{-1}badcb^{-1}d, \ c^{-1}a^{-1}b^{-1}d^{-1}badcb^{-1}d, \ a^{-1}b^{-1}c^{-1}acdb^{-1}c^{-1}ad \rangle.$$

Name the relators $r_1, r_2$ and $r_3$, respectively. Note that $m(r_1) = \{a, b\}$ and $m(r_2) = \{d, b\}$. Theorem 2.5 (with $a_1 = a$ and $a_2 = d$) proves that $G(\mathcal{P})$ is locally indicable.

Theorem 2.5 can be generalized to presentations $\mathcal{P}$ where the total exponents of the relators are not necessarily 0. Let $\mathcal{P}$ be a presentation and $\varphi \colon G(\mathcal{P}) \to \mathbb{Z}$ a surjective homomorphism. By changing, if necessary, a generator by its inverse, we can assume that $\varphi(a)$ is nonnegative for every generator $a$ of $\mathcal{P}$. Given a word $w$ on the generators, the weight of $w$ (with respect to $\varphi$) is $\varphi(w)$ (where $w$ is viewed as an element of $G(\mathcal{P})$). Similarly as before, we can consider the sequence of weights (with respect to $\varphi$) of initial subwords of a relator $r$ and define the multiset $m_\varphi(r)$ of minima of $r$. If the minimum of the sequence is $m$, $m_\varphi(r)$ contains a copy of the letter $a$ for every initial subword $wa^{-1}$ with weight $m$ and every initial subword $wa$ with weight $m + \varphi(a)$. In particular, if $\varphi$ sends every generator to 1,

$m_\varphi(r) = m(r)$. For example, if $\mathcal{P} = \langle a, b, c \mid abc^{-1}b^2, ab^{-3}a \rangle$ we can define $\varphi\colon G(\mathcal{P}) \to \mathbb{Z}$ by $\varphi(a) = 3$, $\varphi(b) = 2$ and $\varphi(c) = 9$. In this case, the sequence for the relator $r = abc^{-1}b^2$ is $3, 5, -4, -2, 0$.

**Theorem 2.9** *Let $\mathcal{P} = \langle a_1, \ldots, a_{k+2} \mid r_1, \ldots, r_k, s \rangle$ be a presentation (of deficiency 1) of a group $G$ with $H_1(G) = \mathbb{Z}$ (for some $k \geq 0$), where all the relators are cyclically reduced. If there is a surjective homomorphism $\varphi\colon G \to \mathbb{Z}$ such that $m_\varphi(r_1), \ldots, m_\varphi(r_k)$ are concatenable, then $G$ is locally indicable.*

**Proof** We adapt the proof of Theorem 2.5 to the more general setting. Let $\widetilde{\mathcal{K}}$ be the infinite cyclic cover of $\mathcal{K}_\mathcal{P}$ determined by $\varphi\colon G \to \mathbb{Z}$. The 1-cells $a_{i,j}$ are oriented from $j$ to $j + \varphi(a_i)$ (recall that we assume that the weights $\varphi(a_i)$ are nonnegative). As before, $\rho_i$ and $\sigma$ are the maximum 0-cells of $R_{i,0}$ and $S_0$, respectively, where $R_{i,j}$ and $S_j$ are the 2-cells of $\widetilde{\mathcal{K}}$ covering $R_i$ and $S$ with minimum 0-cell $j$.

We will assume again that $a_i \in m_\varphi(r_i)$ with multiplicity 1 and $a_i \notin \bigcup_{j=1}^{i-1} m_\varphi(r_j)$. This allows us to write the attaching path of $R_{i,j}$ as $a_{i,j} p_{i,j}^{-1}$ where $p_{i,j}$ is a path that does not involve $a_{i,j}$. Let $S_j^{(t)}$ be the iterated rewrite defined as in the proof of Theorem 2.5 and again let $\alpha_j$ be the minimum 0-cell of $S_0^{(j)}$.

In order to prove that the sequence $\{\alpha_j\}_{j \geq 1}$ stabilizes one has to consider a slightly different complex $Y$. We show that $\alpha_j < \rho + \sigma + \varphi(a_s)$ where $\varphi(a_s) = \max\{\varphi(a_i)\}$. Suppose $\alpha_j \geq \rho + \sigma + \varphi(a_s)$ for some $j$. Take $n \in \mathbb{N}$ big enough such that $\alpha_{n-\rho-\sigma} \geq \rho + \sigma + \varphi(a_s)$. Let $\widetilde{Y}$ be the full subcomplex of $\widetilde{\mathcal{K}}$ containing the 0-cells between $0$ and $n + \rho + \sigma$. It contains the 1-cells $a_{i,j}$ for $1 \leq i \leq k+2$, $0 \leq j \leq n+\rho+\sigma-\varphi(a_i)$, the 2-cells $R_{i,j}$ for $1 \leq i \leq k$ and $0 \leq j \leq n+\rho+\sigma-\rho_i$ and the 2-cells $S_j$ for $0 \leq j \leq n+\rho$. The complex $Y$ is obtained from $\widetilde{Y}$ by attaching an extra 1-cell between $j$ and $j+1$ for every $0 \leq j \leq \rho+\sigma+\varphi(a_s)-1$.

Now we construct a homotopy equivalent 2-complex $Y'$ by replacing each $S_j$ by a 2-cell attached via the path $S_j^{(n-j)}$ for each $0 \leq j \leq \rho+\sigma$. Note that $Y'$ collapses to a subcomplex $Y''$ after removing $a_{i,j}$ and $R_{i,j}$ for every $1 \leq i \leq k$ and $0 \leq j \leq \rho+\sigma+\varphi(a_s)-\varphi(a_i)$. The attaching of the extra 1-cells guarantees that the full subcomplex $Z$ of $Y''$ that contains the 0-cells between $0$ and $\rho+\sigma+\varphi(a_s)$ is connected, and so is $Y$. Note that $Z$ is a 1-subcomplex since $\alpha_{n-\rho-\sigma} \geq \rho+\sigma+\varphi(a_s)$. By a similar argument to that of Theorem 2.5, $\beta_1(Y) \geq \beta_1(Z) = 2\rho+2\sigma+2\varphi(a_s)-\varphi(a_{k+1})-\varphi(a_{k+2})+2$. As before, this implies that $\chi(Y) \leq -1-2\rho-2\sigma-2\varphi(a_s)+\varphi(a_{k+1})+\varphi(a_{k+2})$. On the other side, a direct computation shows that

$$\chi(Y) = -2\rho - 2\sigma - \varphi(a_s) + \sum_{i=1}^{k+2} \varphi(a_i),$$

which implies $\sum_{i=1}^{k} \varphi(a_i) + \varphi(a_s) \leq -1$, a contradiction.

Now the proof proceeds similarly as in Theorem 2.5. $\qquad\square$

**Example 2.10** Let $\mathcal{P} = \langle a, b, c \mid c^{-1}b^{-1}c^{-1}abca, b^{-1}c^{-1}aca^{-1}bab^{-1} \rangle$. Using the homomorphism $\varphi\colon G(\mathcal{P}) \to \mathbb{Z}$ defined by $\varphi(a) = \varphi(b) = 1$ and $\varphi(c) = 2$, the multiset of minima of the first relator is $\{c, a\}$. By Theorem 2.9, $G_\mathcal{P}$ is locally indicable.

## A generalization for two-relator presentations

For presentations $\mathcal{P} = \langle a, b, c \mid r, s \rangle$, Theorem 2.9 requires the existence of a surjective homomorphism $\varphi \colon G \to \mathbb{Z}$ such that $m_\varphi(r)$ (or $m_\varphi(s)$) is concatenable, which essentially means that there exists a generator that attains the minimum of the whole word (and only once). We show now that, by applying some Andrews–Curtis moves (or extended Nielsen transformations) to $\mathcal{P}$, this hypothesis can be relaxed: we only require the existence of a generator with a unique relative minimum.

**Definition 2.11** Let $\mathcal{P}$ be a presentation and $\varphi \colon G(\mathcal{P}) \to \mathbb{Z}$ a surjective homomorphism. By changing, if necessary, a generator by its inverse, we assume that $\varphi(a)$ is nonnegative for every generator $a$ of $\mathcal{P}$. We say that a relator $r$ attains a *unique relative minimum at $a$* if it has an initial subword of the form $w = \widetilde{w}a^{-1}$ or $w = \widetilde{w}a$, with $\varphi(w) = m_a$, that satisfies the following condition. If $w = \widetilde{w}a^{-1}$, then $\varphi(\widehat{w}a^{-1}) > m_a$ and $\varphi(\widehat{w}a) > m_a + \varphi(a)$ for every other initial subwords $\widehat{w}a^{-1}$ and $\widehat{w}a$. If $w = \widetilde{w}a$, then $\varphi(\widehat{w}a) > m_a$ and $\varphi(\widehat{w}a^{-1}) > m_a - \varphi(a)$ for every other initial subwords $\widehat{w}a^{-1}$ and $\widehat{w}a$.

For example, if $r = a^{-1}bac^{-1}bbc^{-1}b$ and $\varphi(a) = 1, \varphi(b) = 2, \varphi(c) = 4$, the sequence of weights is $-1, 1, 2, -2, 0, 2, -2, 0$. Note that $m_\varphi(r) = \{c, c, b, b\}$ is not concatenable but there is a unique relative minimum at $a$ (with value $-1$).

**Corollary 2.12** *Let $\mathcal{P} = \langle a, b, c \mid r, s \rangle$ be a presentation of a group $G$ with $H_1(G) = \mathbb{Z}$ where the relators are cyclically reduced. Let $\varphi \colon G \to \mathbb{Z}$ be a surjective homomorphism with $\varphi(c) \neq 0$. If $r$ attains a unique relative minimum at $a$, then $G$ is locally indicable.*

**Proof** Let $m \in \mathbb{N}$ be arbitrary. Suppose that $r$ attains the relative minimum in an initial subword of the form $\widetilde{w}a^{-1}$. We apply the following Nielsen transformations to $\mathcal{P}$: we add a new generator $d$ together with a new relator $ac^{-m}d^{-1}c^m$, then we replace all the occurrences of the generator $a$ in $r$ and $s$ by $c^{-m}dc^m$ (and cyclically reduce, if necessary) and obtain new relators $\tilde{r}$ and $\tilde{s}$. Finally we remove the relator $ac^{-m}d^{-1}c^m$ together with the generator $a$. We end up with an equivalent presentation $\mathcal{P}' = \langle b, c, d \mid \tilde{r}, \tilde{s} \rangle$. Since the presentations are equivalent, we have $H_*(\mathcal{P}') = H_*(\mathcal{P})$ and $G(\mathcal{P}') = G(\mathcal{P})$. Note that $\varphi(d) = \varphi(a)$. We will show that, for $m \in \mathbb{N}$ big enough, $\mathcal{P}'$ satisfies the hypotheses of Theorem 2.9.

For every subword $wa^{-1}$ with weight $k$ in $r$ there are $2m + 1$ subwords

$$wc^{-1}, \ldots, wc^{-m}, wc^{-m}d^{-1}, wc^{-m}d^{-1}c, \ldots, wc^{-m}d^{-1}c^m$$

in $\tilde{r}$ with weights $k + \varphi(a) - \varphi(c), \ldots, k + \varphi(a) - m\varphi(c), k - m\varphi(c), k - (m-1)\varphi(c), \ldots, k$. Similarly, for every subword $wa$ with weight $k$ in $r$, we have $2m + 1$ subwords

$$wc^{-1}, \ldots, wc^{-m}, wc^{-m}d, wc^{-m}dc, \ldots, wc^{-m}dc^m$$

with weights $k - \varphi(a) - \varphi(c), \ldots, k - \varphi(a) - m\varphi(c), k - m\varphi(c), k - (m-1)\varphi(c), \ldots, k$. Since there is no initial subword of the form $\widetilde{w}a$ with weight $m_a + \varphi(a)$ or less, we can take $m$ big enough such that the minimum in the sequence of weights of $\tilde{r}$ is unique and is attained at the $d$ that appears after replacing the rightmost $a^{-1}$ of the unique subword $\widetilde{w}a$ of weight $m_a$. The corresponding subword has weight $m_a - m\varphi(c)$.

If $r$ attains the relative minimum in an initial subword of the form $\widetilde{w}a$, replace $r$ by $r^{-1}$. $\qquad \square$

**Example 2.13** Consider the presentation

$$\mathcal{P} = \langle a, b, c \mid a^{-1}b^{-1}abaab^{-1}cb^{-1}a^{-1}a^{-1}a^{-1}baac^{-1}, \ c^{-1}b^{-1}cbccb^{-1}cb^{-1}c^{-1}c^{-1}c^{-1}baca^{-1} \rangle.$$

If we take $\varphi \colon G(\mathcal{P}) \to \mathbb{Z}$ defined by $\varphi(a) = \varphi(b) = \varphi(c) = 1$, the first relator attains a unique relative minimum at $c$. By Corollary 2.12, $G(\mathcal{P})$ is locally indicable. One can check that the previous methods do not work for this presentation.

## Weakly concatenable relators and $I$-test

Theorems 2.5 and 2.9 apply to presentations with the homology of $S^1$. Both results impose conditions on all the relators of the presentation but one. This fact will be used in the next section to extend a well-known criterion of Howie on LOT groups (see Corollary 3.1). We now investigate local indicability of groups with presentations with the homology of a wedge of circles, ie groups $G$ with nontrivial free abelian $H_1(G)$ that admit presentations $\mathcal{P} = \langle a_1, \ldots, a_n \mid r_1, \ldots, r_k \rangle$ where $n - k = \mathrm{rank}(H_1(G))$. The next result imposes weaker conditions to all the relators of the presentation.

It will be convenient to use the notion of $I$-*values*. This notion is inspired in the $I$-test developed in [2]. We will work with sequences of minima. As before, one can formulate and prove analogous results for maxima.

**Definition 2.14** Let $\mathcal{P}$ be a presentation of a group $G$ and $\varphi \colon G \to \mathbb{Z}$ a surjective homomorphism. By changing, if necessary, a generator by its inverse, we can assume that $\varphi(a)$ is nonnegative for every generator $a$ of $\mathcal{P}$. Given a word $r$ on the generators, the sequence of $I$-values of $r$ with respect to $\varphi$ is obtained by assigning to each initial subword $w$ with the form $w = \widetilde{w}x$ (where $x$ is a generator) the value $\varphi(\widetilde{w})$ and to each initial subword of the form $w = \widetilde{w}x^{-1}$ the value $\varphi(\widetilde{w}x^{-1})$.

This sequence is useful when considering the covering of $\mathcal{K}_\mathcal{P}$ corresponding to $\varphi$. The $I$-values of a relator $r$ are precisely the initial vertices of the 1-cells traversed by the attaching path of $R_m$ (where $m$ is the minimum of its sequence of weights).

**Remark 2.15** The unique relative minimum condition of Definition 2.11 is equivalent to the relator having an initial subword $wa^{-1}$ or $wa$ with minimum $I$-value $m_a$ over all initial subwords of the form $\widetilde{w}a$ or $\widetilde{w}a^{-1}$.

**Remark 2.16** Given a relator $r$, the multiset of letters in which the minima of the sequence of $I$-values of $r$ are attained is exactly $m_\varphi(r)$.

**Definition 2.17** A family of relators $\{r_1, \ldots, r_k\}$ is *weakly concatenable* (with respect to $\varphi \colon G \to \mathbb{Z}$), if there is an ordering $m_\varphi(r_{i_1}), \ldots, m_\varphi(r_{i_k})$ of their multisets of minima such that, for every $1 \le j \le k$, there exists an element $x \in m_\varphi(r_{i_j})$ such that $x \notin \bigcup_{s=1}^{j-1} m_\varphi(r_{i_s})$ and, if the minimum $I$-value of $r_i$ is $m$, the number of subwords $wx^{-1}$ of $r_i$ with $I$-value $m$ is different from the number of subwords $wx$ with $I$-value $m$.

The condition on the number of subwords with rightmost letter $x$ in the previous definition will be used in the next result to prove that a certain 2-cell $R_{i,j}$ corresponding to the relator $r_i$ properly involves a 1-cell corresponding to the generator $x$.

**Theorem 2.18** *Let $\mathcal{P} = \langle a_1, \ldots, a_n \mid r_1, \ldots, r_k \rangle$ be a group presentation with cyclically reduced relators of a group $G$ with $H_1(G)$ nontrivial free abelian of rank $n - k$. Let $\varphi \colon G \to \mathbb{Z}$ be a surjective homomorphism. If $\{r_1, \ldots, r_k\}$ is weakly concatenable with respect to $\varphi$, then $G$ is locally indicable.*

**Proof** We consider as before the covering $\widetilde{\mathcal{K}}$ of $\mathcal{K}_\mathcal{P}$ corresponding to $\varphi$. Again, the 1-cells $a_{i,j}$ are oriented from $j$ to $j + \varphi(a_i)$ (recall that we assume that the weights $\varphi(a_i)$ are nonnegative) and $R_{i,j}$ are the 2-cells covering $R_i$ with minimum 0-cell $j$. By Theorem 2.1 it suffices to prove that every finite connected 2-subcomplex $K \subset \widetilde{\mathcal{K}}$ is reducible. Note again that the no-proper-power condition follows from [1, Proposition 1] and the fact that $\widetilde{\mathcal{K}}$ is an infinite cyclic cover of a 2-complex $\mathcal{K}_\mathcal{P}$ with the homology of a wedge of circles.

Given a finite and connected 2-subcomplex $K$, let $j \in \mathbb{Z}$ be the minimum 0-cell in $K$ that is incident to some 2-cell of $K$. Let $R_{i_1,j}, \ldots, R_{i_s,j}$ with $i_1 < \ldots < i_s$ be the 2-cells incident to the 0-cell $j$. By hypothesis, $a_{i_s,j}$ is a face of the 2-cell $R_{i_s,j}$ and is not face of any other 2-cell $R_{i,l}$ of $K$. This 1-cell is properly involved in $R_{i_s,j}$ by definition of weakly concatenable, and therefore it is an elementary reduction. $\square$

**Example 2.19** Consider the presentation $\mathcal{P} = \langle a, b, c \mid r_1, r_2 \rangle$ with relators

$$r_1 = a^{-1}c^{-1}aaab^{-1}b^{-1}c^{-1}ab, \quad r_2 = c^{-1}b^{-1}cb^{-1}caba^{-1}ba^{-1}.$$

Let $\varphi \colon G \to \mathbb{Z}$ be the homomorphism defined by $\varphi(a) = \varphi(b) = \varphi(c) = 1$. It is easy to verify that the multisets of minima are $m_\varphi(r_1) = \{c, a, c, a\}$ and $m_\varphi(r_2) = \{b, c, b, c\}$, and that $\{r_1, r_2\}$ is weakly concatenable with the ordering $m_\varphi(r_2), m_\varphi(r_1)$. By Theorem 2.18, $G(\mathcal{P})$ is locally indicable.

# 3 Applications to labeled oriented trees

We now apply the results of the previous section to derive new results on LOT groups. The first one is a straightforward application of Theorem 2.5 and is one of the main results of this paper. It is an extension of a well known result of Howie [11, Theorem 10.1] on the left and right graphs associated to a labeled oriented tree, and is also related to similar results by Gersten [7, Proposition 4.15] and Wise [17] on Adian presentations.

Recall that a labeled oriented graph (LOG) $\Gamma$ consists of two sets $V(\Gamma)$ and $E(\Gamma)$ of vertices and edges, and three functions $i, t, \lambda \colon E \to V$ that map each edge to its initial vertex, terminal vertex and label, respectively. It is called a labeled oriented tree (LOT) when the underlying graph is a tree. The standard presentation $P(\Gamma)$ (of a group $G(\Gamma)$) associated to each LOG $\Gamma$ is

$$P(\Gamma) = \langle V(\Gamma) \mid \{t(e)^{-1}\lambda(e)^{-1}i(e)\lambda(e) : e \in E(\Gamma)\}\rangle$$

Note that all relators have total exponent 0 and, if $\Gamma$ is a LOT, $P(\Gamma)$ has deficiency 1 and it is a homology circle. Every LOT can be changed, by a finite sequence of transformations, to a reduced one with the same homotopy type (see [11, Section 3]). In particular one can assume that $\lambda(e) \neq i(e)$ and $\lambda(e) \neq t(e)$ for every edge $e$, which implies that all the relators of $P(\Gamma)$ are cyclically reduced.

In this section we will use the homomorphism $\varphi \colon G(\Gamma) \to \mathbb{Z}$ that sends every generator of $P(\Gamma)$ to 1. The sequence of $I$-values with respect to $\varphi$ will be called $I$-sequence. Note that, given any relator $r$ (corresponding to an edge $e$ of $\Gamma$), its $I$-sequence attains exactly two minima, one at $\lambda(e)$ and the other one at $i(e)$, then $m(r) = \{\lambda(e), i(e)\}$. Analogously, it attains two maxima, one at $\lambda(e)$ and the other one at $t(e)$.

In [11] Howie introduced two graphs $I(\Gamma)$ and $T(\Gamma)$ associated to any LOG $\Gamma$. The vertex set of the left graph $T(\Gamma)$ is $V(\Gamma)$ and, for every edge $e$ in $\Gamma$, we put an (unoriented) edge $\{\lambda(e), i(e)\}$ (these are precisely the letters in which the minima of the $I$-sequence are attained). Similarly, the right graph $I(\Gamma)$ has an edge $\{\lambda(e), t(e)\}$ for every edge $e$ in $\Gamma$. In [7] Gersten investigated these graphs in the more general context of Adian presentations. An Adian presentation is one of the form

$$\mathcal{P} = \langle A \mid u_i = v_i \ (i \in J)\rangle,$$

where $u_i$ and $v_i$ are nontrivial positive words on $A$. In this context, the edges of the graph $T(\mathcal{P})$ (resp $I(\mathcal{P})$) connect the first (resp last) letter of $u_i$ to that of $v_i$. Howie proved that, if $I(\Gamma)$ or $T(\Gamma)$ is a tree then $G(\Gamma)$ is locally indicable (and, in particular, $\Gamma$ is aspherical) [11, Theorem 10.1]. In the case of Adian presentations, Gersten proved that if the presentation is cycle-free (ie if both graphs are forests) then $\mathcal{K}_\mathcal{P}$ is DR (diagrammatically reducible) and, in particular, aspherical [7, Proposition 4.12]. Moreover, when $\ell(u_i) = \ell(v_i)$ for every $i$ (as in the case of LOTs), if either $T(\mathcal{P})$ or $I(\mathcal{P})$ has no cycles then $\mathcal{K}_\mathcal{P}$ is DR (and hence aspherical) [7, Proposition 4.15]. More recently, Wise showed that the 2-complexes associated to cycle-free Adian presentations have nonpositive sectional curvature, which implies that the corresponding groups are locally indicable [17, Theorem 11.4].

Note that Theorem 2.5 imposes conditions to all the relators of the presentation but one. If either $I(\Gamma)$ or $T(\Gamma)$ has only one cycle, we can remove an edge of the cycle to get a forest. Then, as an immediate application of Theorem 2.5 (or the analogous result for maxima), we obtain the following generalization of [11, Theorem 10.1].

**Corollary 3.1** *Let $\Gamma$ be a LOT. If either $I(\Gamma)$ or $T(\Gamma)$ has at most one cycle, then $G(\Gamma)$ is locally indicable.*

**Example 3.2** Consider the following LOT $\Gamma$:
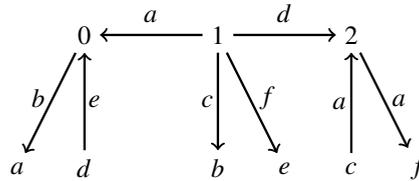


Neither $I(\Gamma)$ nor $T(\Gamma)$ is a tree, so Howie's (or Gersten's) result do not apply. Note that $T(\Gamma)$ has exactly one cycle and then, by Corollary 3.1, the group $G(\Gamma)$ is locally indicable (and $\Gamma$ is aspherical).

In [11] Howie studied some subfamilies of LOTs of diameter 4 with 3 nonextremal vertices. Using Corollary 3.1 one can show local indicability for most examples of the remaining subfamilies. Example 3.2 belongs to one of them, in which one of the extremal vertices appears three times as a label.

Corollary 3.1 can be generalized to a more general class of Adian presentations. Suppose

$$\mathcal{P} = \langle a_1, \ldots, a_n \mid u_1 = v_1, \ldots, u_{n-1} = v_{n-1} \rangle$$

is an Adian presentation of deficiency 1 with $H_1(G(\mathcal{P})) = \mathbb{Z}$ (where the relators are cyclically reduced). Suppose further that there exists a surjective homomorphism $\varphi \colon G(\mathcal{P}) \to \mathbb{Z}$ with $\varphi(a_i) > 0$ for every $a_i$. For any such homomorphism $\varphi$, the multiset of minima $m_\varphi(r_i)$ of each relator $r_i$ is the set consisting of the first letter of $u_i$ and the first letter of $v_i$. Then, if either $I(\mathcal{P})$ or $T(\mathcal{P})$ has at most one cycle, as an immediate application of Theorem 2.9, we obtain the following result, which extends, for a subfamily of Adian presentations, Wise's result [17, Theorem 11.4].

**Corollary 3.3** Let $\mathcal{P} = \langle a_1, \ldots, a_n \mid u_1 = v_1, \ldots, u_{n-1} = v_{n-1} \rangle$ be an Adian presentation of deficiency 1 with $H_1(G(\mathcal{P})) = \mathbb{Z}$. Suppose further that there exists a surjective homomorphism $\varphi \colon G(\mathcal{P}) \to \mathbb{Z}$ with $\varphi(a_i) > 0$ for every $a_i$. If either $I(\mathcal{P})$ or $T(\mathcal{P})$ has at most one cycle, then $G(\mathcal{P})$ is locally indicable.

The condition on the existence of the homomorphism $\varphi$ in the previous corollary is automatically satisfied when $\ell(u_i) = \ell(v_i)$ for every $i$.

We finish the paper by analyzing a strategy that can be useful to study local indicability of LOT groups for which $I(\Gamma)$ and $T(\Gamma)$ have more than one cycle. The main idea is to apply convenient extended Nielsen transformations to the LOT presentation $P(\Gamma)$ to obtain an equivalent presentation $\mathcal{P}$ which satisfies the hypotheses of Theorem 2.18.

For now on we consider the labeled oriented (original) versions of $I(\Gamma)$ and $T(\Gamma)$: for every edge $a \xrightarrow{c} b$ of a LOT $\Gamma$, we put an edge $a \xrightarrow{b} c$ in $T(\Gamma)$ and an edge $c \xrightarrow{a} b$ in $I(\Gamma)$. We state the result for $T(\Gamma)$, the corresponding result for $I(\Gamma)$ is analogous.

**Definition 3.4** Let $\Gamma$ be a LOT. Let $C$ be a cycle in $T(\Gamma)$ and $e$ an edge in $C$. We say that $e$ is *properly labeled* in $C$ if the number of edges in $C$ with the same label as $e$ oriented clockwise is different from the number of edges with the same label as $e$ oriented counterclockwise.

**Lemma 3.5** *Let $\Gamma$ be a LOT. Let $e$ be an edge in a simple cycle $C$ of $T(\Gamma)$ with label $x$. Suppose further that $x$ does not appear as a vertex in $C$. Then, by applying extended Nielsen transformations, one can replace the relator $r$ corresponding to $e$ in $P(\Gamma)$ by another relator $\tilde{r}$ which contains the generator $x$ and other generators that appear as labels or vertices in $C$, the new relator $\tilde{r}$ has a constant $I$-sequence (it attains a minimum at every letter), and in particular it attains a minimum at $x$. Moreover, if $e$ is properly labeled in $C$, the number of initial subwords of $\tilde{r}$ of the form $wx$ and the number of initial subwords of the form $wx^{-1}$ are different.*

**Proof** We say that a word is a *zigzag* if its length is even and the exponent of the $i^{\text{th}}$ letter is $(-1)^{i+1}$ (ie it has the form $x_1 x_2^{-1} \ldots x_{2s}^{-1}$). Suppose $r_1$ and $r_2$ are relators and $a$ is a letter in which both relators attain a minimum in the $I$-sequence. Let $b$ and $c$ be the other letters in which $r_1$ and $r_2$ attain the other minimum. Up to cyclic permutation and inversions, we can assume $r_1 = bw_1 a^{-1}$ and $r_2 = aw_2 c^{-1}$ with $w_1, w_2$ zigzags. Now, we can change $r_1$ by $r_1 r_2 = bw_1 w_2 c^{-1}$ which attains exactly two minima in its $I$-sequence, one at $b$ and the other one at $c$.

Let $C = a_1, a_2, a_3, \ldots, a_l, a_1$ be the cycle of $T(\Gamma)$ and $e = (a_1, a_2)$. Let $r_i$ be the relator corresponding to $(a_i, a_{i+1})$ for every $1 \le i \le l$ (where the indices are taken modulo $l$). Up to cyclic permutations and inversions, we can assume that $r_i = a_i w_i a_{i+1}^{-1}$ where $w_i$ is a zigzag word. Now we can replace $r_1$ by $r_1 r_2 \ldots r_l = a_1 w_1 w_2 \ldots w_l a_1^{-1}$, which, after a cyclic permutation, is $w_1 w_2 \ldots w_l$. Note that $r = w_1 w_2 \ldots w_l$ is a zigzag. Then every letter attains the same $I$-value (the minimum). Finally, since $(a_1, a_2)$ is properly labeled, the number of occurrences of $x$ in this new relator is different from the number of occurrences of $x^{-1}$. $\qquad\square$

**Theorem 3.6** *Let $\Gamma$ be a LOT. Denote by $X_0, X_1, \ldots, X_k$ the connected components of $T(\Gamma)$. Suppose that $C_1, \ldots, C_k$ are generating simple cycles in $X_0$ and the other connected components $X_1, \ldots, X_k$ are trees. If for every $1 \le i \le k$ there is at least one properly labeled edge of $C_i$ that has a vertex of $X_i$ as a label and for $j \ge i$ there is no other vertex of $X_j$ that appears as a label in the edges of $C_i$, then $G_\Gamma$ is locally indicable.*

**Proof** To illustrate the idea of the proof, we show the case $k = 1$. Let $e_1$ be a properly labeled edge in the cycle $C_1$, which is labeled with a vertex $x_1$ of $X_1$. By Lemma 3.5 we can replace the corresponding relator $r_1$ by a relator $\widetilde{r_1}$, which attains a minimum at the vertex $x_1$ and in which the number of initial subwords $wx_1$ with minimum $I$-value and the number of initial subwords $wx_1^{-1}$ with minimum $I$-value are different. Note that the new presentation $\mathcal{P}$ is Andrews–Curtis equivalent to $P(\Gamma)$.

Now, the relators of the new presentation $\mathcal{P}$ are weakly concatenable and the result follows from Theorem 2.18. The ordering for a weak concatenation from back to front is the following: first follow the order of collapses of all the edges of the tree $X_1$ to the vertex $x_1$, then continue with the new relator $\widetilde{r_1}$, and finally use any ordering of weak concatenation for the remaining relators of $X_0$ (note that $X_0 - e_1$ is a tree).
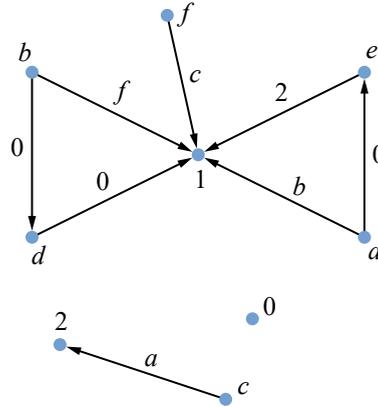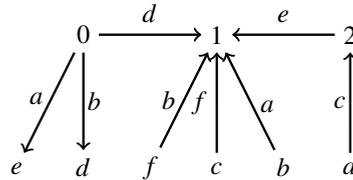
Figure 1: $I(\Gamma)$.

The general case follows by induction in a similar way. We require that for $j \geq i$ there is no other vertex of $X_j$ that appears as a label in the edges of $C_i$. This condition is necessary to guarantee weak concatenability, since by Lemma 3.5 the new relators $\widetilde{r_i}$ attain a minimum at every letter that appears in the cycle $C_i$. $\square$

**Example 3.7** Consider the following LOT $\Gamma$:



Both $I(\Gamma)$ and $T(\Gamma)$ have two cycles, so we cannot apply Corollary 3.1. The graph $I(\Gamma)$ is displayed in Figure 1.

Note that $I(\Gamma)$ satisfies the hypotheses of Theorem 3.6 with $X_0$ the subgraph with vertices $\{b, d, 1, e, a, f\}$, $X_1 = \{0\}$ and $X_2$ the tree with vertices $\{2, c\}$. In the cycle $C_1 = 1, b, d, 1$ we consider the edges properly labeled with $x_1 = 0$ and in the cycle $C_2 = 1, a, e, 1$ we take the edge properly labeled with $x_2 = 2$. This shows that $G(\Gamma)$ is locally indicable.

# References

[1]  **J F Adams**, *A new proof of a theorem of W H Cockcroft*, J. Lond. Math. Soc. 30 (1955) 482–488  MR

[2]  **J A Barmak**, **E G Minian**, *A new test for asphericity and diagrammatic reducibility of group presentations*, Proc. Roy. Soc. Edinburgh Sect. A 150 (2020) 871–895  MR

[3]  **S D Brodskii**, *Equations over groups and groups with one defining relation*, Uspekhi Mat. Nauk 35 (1980) 183  MR  In Russian; translated in Russian Math. Surv. 35 (1980) 165

[4]  **A Clay**, **D Rolfsen**, *Ordered groups and topology*, Graduate Studies in Math. 176, Amer. Math. Soc., Providence, RI (2016)  MR

[5] **B Deroin**, **A Navas**, **C Rivas**, *Groups*, *orders*, *and dynamics*, preprint (2014) arXiv 1408.5805

[6] **S M Gersten**, *Conservative groups, indicability, and a conjecture of Howie*, J. Pure Appl. Algebra 29 (1983) 59–74 MR

[7] **S M Gersten**, *Reducible diagrams and equations over groups*, from "Essays in group theory", Math. Sci. Res. Inst. Publ. 8, Springer (1987) 15–73 MR

[8] **G Higman**, *The units of group-rings*, Proc. Lond. Math. Soc. 46 (1940) 231–248 MR

[9] **J Howie**, *On pairs of 2-complexes and systems of equations over groups*, J. Reine Angew. Math. 324 (1981) 165–174 MR

[10] **J Howie**, *On locally indicable groups*, Math. Z. 180 (1982) 445–461 MR

[11] **J Howie**, *On the asphericity of ribbon disc complements*, Trans. Amer. Math. Soc. 289 (1985) 281–302 MR

[12] **J Howie**, **H R Schneebeli**, *Homological and topological properties of locally indicable groups*, Manuscripta Math. 44 (1983) 71–93 MR

[13] **J Howie**, **H Short**, *Coherence and one-relator products of locally indicable groups*, Groups Geom. Dyn. 17 (2023) 1085–1113 MR

[14] **A A Klyachko**, **M A Mikheenko**, *Yet another Freiheitssatz: mating finite groups with locally indicable ones*, Glasg. Math. J. 65 (2023) 337–344 MR

[15] **A Rhemtulla**, **D Rolfsen**, *Local indicability in ordered groups: braids and elementary amenable groups*, Proc. Amer. Math. Soc. 130 (2002) 2569–2577 MR

[16] **H Wilton**, *Rational curvature invariants for 2-complexes*, Proc. A. 480 (2024) art. id. 20240025 MR

[17] **D T Wise**, *Sectional curvature, compact cores, and local quasiconvexity*, Geom. Funct. Anal. 14 (2004) 433–468 MR

[18] **D T Wise**, *Coherence, local indicability and nonpositive immersions*, J. Inst. Math. Jussieu 21 (2022) 659–674 MR

*Departamento de Matematica, Universidad de Buenos Aires*
*Buenos Aires, Argentina*

*Departamento de Matematica, Universidad de Buenos Aires*
*Buenos Aires, Argentina*

abarreto@dm.uba.ar, gminian@dm.uba.ar

# Cubulating drilled bundles over graphs

Mahan Mj

Biswajit Nag

We start with a Gromov-hyperbolic surface bundle $E$ over a graph, and drill out essential simple closed curves from fibers to obtain a drilled bundle $F$. We prove that for such drilled bundles $F$, the fundamental group $\pi_1(F)$ is relatively hyperbolic with $\mathbb{Z} \oplus \mathbb{Z}$ peripheral groups. Combining the relative hyperbolicity of $\pi_1(F)$ thus obtained with a theorem of Wise, we establish virtually special cubulability of $\pi_1(F)$ provided that the maximal undrilled subbundles of $F$ are cubulable.

20F65, 20F67; 22E40, 57M50

## 1 Introduction

In [24], Manning, Mj, and Sageev investigated the following question:

**Question 1.1** *Let $1 \to H \to G \to F_n \to 1$ be an exact sequence of hyperbolic groups with $F_n$ the free group on $n$ generators. Is $G$ virtually special cubulable?*

They gave sufficient conditions guaranteeing an affirmative answer. Surface-by-free groups as in Question 1.1 naturally arise as fundamental groups of surface bundles $E$ over graphs $\mathcal{G}$. The main contribution of [24] was an explicit construction of a codimension-one quasiconvex subgroup of $G$ along which $G$ splits. Wise's quasiconvex hierarchy theorem [34] then furnishes cubulability of such groups. Groups $G$ arising as in Question 1.1 may alternatively be described as a free product with amalgamation of hyperbolic groups along *nonquasiconvex hyperbolic subgroups*. For instance, if $n = 2$ and $H$ is a

hyperbolic surface group, then $G = G_1 *_H G_2$, where $G_1$ and $G_2$ are fundamental groups of hyperbolic 3-manifolds fibering over the circle. Note that $H$ is not quasiconvex in $G_1$ and $G_2$. Thus the aim of Question 1.1 is to understand how far the quasiconvexity hypothesis in Wise's quasiconvex hierarchy theorem [34] can be relaxed. In other words, the class of hyperbolic groups arising as fundamental groups of surface bundles, as over graphs, ie $G$ in Question 1.1, arise as a natural test case to relax the quasiconvexity hypothesis in Wise's theorem.

When $\mathcal{G}$ is a circle, Question 1.1 has an affirmative answer thanks to Kahn and Markovic's work on the surface subgroup problem [17], Bergeron and Wise's cubulability result [4], and Agol's theorem [1]. In this case, $E$ is a hyperbolic 3-manifold $M$ fibering over the circle. The existence of an embedded quasiconvex surface in such an $M$ is thus not guaranteed when the first Betti number of $M$ is one. On the other hand, for hyperbolic 3-manifolds with toroidal boundary components, the construction of embedded geometrically finite surfaces is much easier. In particular, if one drills out a simple closed curve $\sigma$ from a fiber $S$ of a fibered manifold $M$ as above, then $M \setminus \sigma$ admits a complete hyperbolic structure by Thurston's theorem [18, Chapter 15], and each component of $S \setminus \sigma$ is geometrically finite. The fact that each component of $S \setminus \sigma$ is geometrically finite allows us to cut $M$ along these components, and hence construct a geometrically finite hierarchy. This allows us to use Wise's relatively quasiconvex hierarchy theorem [34] even without assuming the existence of a plentiful supply of codimension-one geometrically finite subgroups.

We shall adopt the point of view that a surface bundle $E$ over a graph $\mathcal{G}$ generalizes hyperbolic 3-manifolds fibering over the circle. We shall then create a setup where one can apply the above strategy for drilled hyperbolic 3-manifolds. Thus the main objects of study will be drilled surface bundles over graphs, where drilling corresponds to removing open neighborhoods of simple closed curves in fibers. More precisely (see Figure 1):

**Definition 1.2** Let $\mathcal{G}$ be a connected graph, thought of as a 1-complex, and consider a bundle $\Pi \colon E \to \mathcal{G}$ with fiber $S$ a surface. We refer to $\Pi \colon E \to \mathcal{G}$ as a *surface bundle over the graph* $\mathcal{G}$. The fiber $\Pi^{-1}(x)$ over $x \in \mathcal{G}$ will be denoted by $S_x$. If $x$ is a vertex of $\mathcal{G}$, $S_x$ will be called a singular fiber; otherwise it will be called a *regular fiber*.

Let $\sigma_i \subset S_{x_i}$ be a finite collection of essential simple closed curves in regular fibers $S_{x_i}$, so that for each regular fiber $S_x$, the collection of simple closed curves contained in $S_x$ are disjoint. Let $\{N_\epsilon(\sigma_i)\}$ be a collection of small open tubular neighborhoods, missing singular fibers. We assume that:

(1) The closures $\{\overline{N_\epsilon(\sigma_i)}\}$ are disjoint.

(2) For $i \neq j$, $\sigma_i$ and $\sigma_j$ are not freely homotopic in $E$, nor is there a nontrivial free homotopy between $\sigma_i$ and itself. This can equivalently be described as follows. Let $F = \left( E \setminus \bigcup_i N_\epsilon(\sigma_i) \right)$. Then any $\pi_1$-injective smooth map $(S^1 \times [0,1], S^1 \times \{0,1\}) \to \left( F, \bigcup_i \partial N_\epsilon(\sigma_i) \right)$ is homotopic, rel. boundary, into $\partial N_\epsilon(\sigma_i)$ for some $i$.
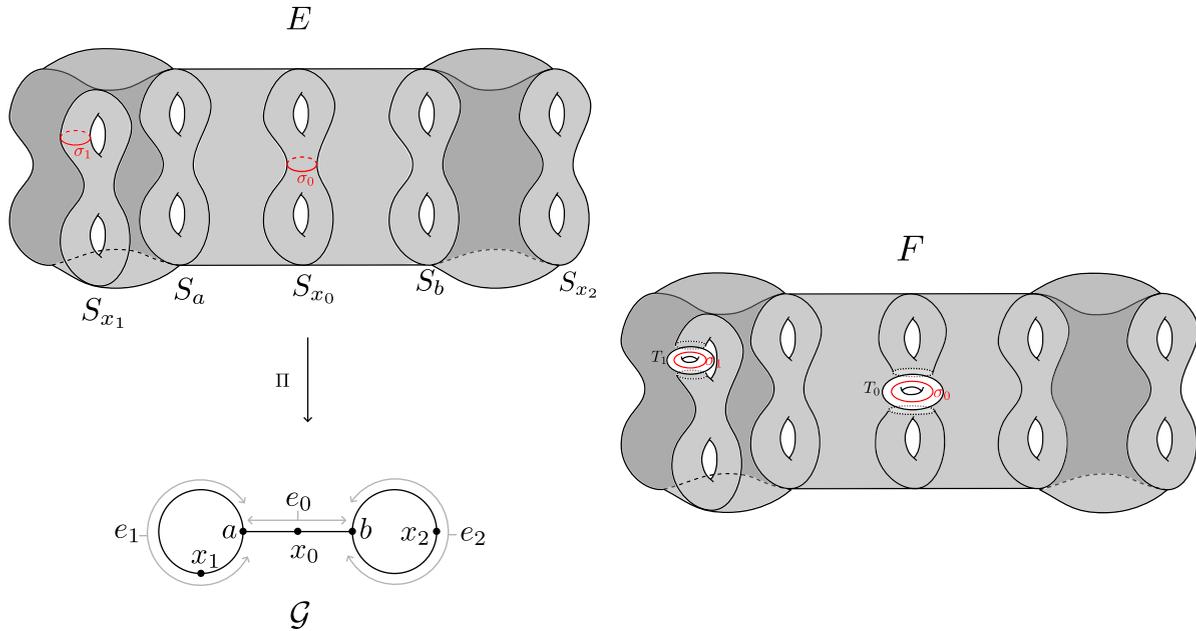
Figure 1: A drilled surface bundle over a graph: $e_0$ and $e_1$ are drilled edges, $x_0$ and $x_1$ are drilled points, $S_{x_0}$ and $S_{x_1}$ are drilled fibers, $\sigma_0$ and $\sigma_1$ are drilled curves, $e_2$ is an undrilled edge, and the closure $\bar{e}_2$ is an undrilled component. Also, $S_a$ and $S_b$ are singular fibers.

The complement $F = \left( E \setminus \bigcup_i N_\epsilon(\sigma_i) \right)$ will be referred to as a *drilled surface bundle over a graph*.

Drilled surface bundles over a graph are our principal objects of study. Our objective is to find sufficient conditions to cubulate them. Each torus $\{\partial N_\epsilon(\sigma_i)\}$ will be referred to as a *boundary torus* of $F$, and denoted by $T_i$. The union $\bigcup_i T_i$ will be called the *boundary* of $F$. The surfaces $S_{x_i}$ (containing some $\sigma_i$) will be called *drilled fibers* and the curves $\sigma_i$ will be called *drilled curves*. The points $x_i \in \mathcal{G}$ will be called *drilled points*. An edge $e$ containing a drilled point will be called a *drilled edge*. Otherwise, we call it an *undrilled edge*.

A $\pi_1$-injective smooth map $(S^1 \times [0, 1], S^1 \times \{0, 1\}) \to \left( F, \bigcup_i \partial N_\epsilon(\sigma_i) \right)$ will be referred to as an *essential annulus* (see Figure 4). The second condition in Definition 1.2 thus says that $F$ has no essential annuli.

We shall be particularly interested in the case that $\mathcal{G}$ is finite, and $\Gamma = \pi_1(E)$ is Gromov hyperbolic. In this special case, we note the further restriction on drilled curves that follows from the second condition in Definition 1.2. Let $E_S$ be the cover of $E$ corresponding to $\pi_1(S)$. Then we have:

**Condition 1.3** *Let $\theta_1$ and $\theta_2$ be any two distinct elevations of the drilled curves $\sigma_i \subset E$ to $E_S$. Then $\theta_1$ and $\theta_2$ are not freely homotopic in $E_S$ in the complement of other elevated curves. Equivalently, $E_S$ has* **no essential annuli** *(see Definition 3.4).*

Condition 1.3 follows from Definition 1.2(2), viz that $E$ has no essential annuli (see Definition 3.4). However, we make this explicit as much of this paper will involve $E_S$.

The bulk of this paper goes into establishing strong relative hyperbolicity of fundamental groups of drilled bundles (see Theorem 3.1):

**Theorem 1.4** *Let $E$ be a surface bundle over a graph $\mathcal{G}$ (see Definition 1.2) such that $\Gamma = \pi_1(E)$ is hyperbolic. Let $F$ be a drilled surface bundle over $\mathcal{G}$ obtained by drilling $E$. Then $G(= \pi_1(F))$ is strongly hyperbolic relative to the collection of peripheral subgroups $\{\mathcal{P}_i = \pi_1(\partial N_\epsilon(\sigma_i))\}$.*

The drilling operation in this paper, and specifically in Theorem 1.4, is motivated, in part, by the drilling of simple closed geodesics from closed hyperbolic 3-manifolds [32, Chapter 4]. We also refer to [3; 2; 13] for related, but different, drilling constructions: in the first two, totally geodesic codimension-2 submanifolds are drilled out of CAT($-1$) manifolds, and in the last, closed geodesics are drilled out of hyperbolic PD(3) groups.

With Theorem 1.4 in place, we can cut $F$ along the components $\Sigma$ of $S \setminus \bigcup_i \sigma_i$. The fundamental group $\pi_1(\Sigma)$ of each such component $\Sigma$ is relatively quasiconvex in $G$ (see Proposition 4.21). We can now apply Wise's relatively hyperbolic version of the quasiconvex hierarchy theorem [34, Theorem 15.1]. To state our main theorem, a bit more terminology needs to be set up. Let $\mathcal{K}_1, \ldots, \mathcal{K}_n$ denote maximal subgraphs of $\mathcal{G}$ that contain no points $x$ such that the fiber $S_x$ is drilled. We refer to the $\mathcal{K}_i$ as undrilled components of $\mathcal{G}$ (see Figure 1). The restrictions of $E$ to the undrilled components $\mathcal{K}_1, \ldots, \mathcal{K}_n$ will be denoted by $E(\mathcal{K}_1), \ldots, E(\mathcal{K}_n)$ of $F$ (note that each such $E(\mathcal{K}_n)$ is naturally contained in $F$). Our main theorem is the following (see Theorem 5.5).

**Theorem 1.5** *Let $\{E(\mathcal{K}_i)\}$ denote the restrictions of $E$ to the undrilled components $\mathcal{K}_1, \ldots, \mathcal{K}_n$ as above. If for each $\mathcal{K}_i$, $\pi_1(E(\mathcal{K}))$ is cubulable virtually special, then so is $G = \pi_1(F)$.*

Thus, if Question 1.1 has a positive answer (the undrilled case), then for the drilled groups $G$, cubulability follows. However, even in the absence of a definitive answer to Question 1.1, Theorem 1.5 furnishes a number of examples, as given below:

(1) when each edge of $\mathcal{G}$ contains an $x$ such that $S_x$ is drilled (Example 5.6),

(2) when undrilled components of $\mathcal{G}$ are either contractible or homotopy equivalent to a circle (Example 5.7),

(3) when the restrictions of $E$ to the undrilled components satisfy the sufficient conditions in [24] (Example 5.8).[1]

Finally, we use Theorem 1.5 in conjunction with Kielak's theorem [20] to deduce that the cubulable virtually special groups virtually algebraically fiber (see Theorem 6.4).

**Theorem 1.6** *Let $F$ be drilled surface bundle over a finite graph $\mathcal{G}$ (see Definition 1.2) satisfying the hypotheses of Theorem 1.5. Then $G = \pi_1(F)$ virtually algebraically fibers.*

---

[1]We have removed the term "undrilled constituent" as it only occurred locally around this point only thrice.

To prove Theorem 1.6, we first establish that for any drilled surface bundle $F$ over a graph, $\pi_1(F)$ has vanishing first $l^2$ Betti number (Proposition 6.3). This is done using work of Lott and Lück [21] and Fernós and Valette [12].

**Scheme**   As indicated above, the main difficulty in establishing the cubulation Theorem 1.5 is the relative hyperbolicity result Theorem 1.4. Sections 3 and 4 are devoted to proving this result. Section 3 proves hyperbolicity of the universal cover of a *partially electrified space* $(F, d_{\mathrm{pel}})$. Here, only the meridians of boundary tori in a drilled surface bundle $F$ over $\mathcal{G}$ (see Definition 1.2) are electrified instead of the whole boundary tori to give a pseudometric $d_{\mathrm{pel}}$. Theorem 3.14, the main theorem of Section 3, then shows that $(\widetilde{F}, d_{\mathrm{pel}})$ is hyperbolic (here, the lifted pseudometric is also denoted by $d_{\mathrm{pel}}$).

Using Theorem 3.14, we then use a guessing geodesics argument in Section 4 and a theorem of Sisto to establish the relative hyperbolicity of $G$, ie Theorem 1.4.

**Recurring notation**   Before we proceed, we collect below for the ease of reference the recurring notation we will use throughout the paper. Other notation, specific to a section, will be listed at the beginning of the relevant section.

- $\mathcal{G}$ is a finite graph. $S$ is a closed hyperbolic surface. $\Pi \colon E \to \mathcal{G}$ is an $S$-bundle. $H$ is the fundamental group of $S$. $S_x$ is the fiber of $\Pi$ over a point $x$. See Definition 1.2.

- $N_\epsilon(\sigma)$ is a small tubular neighborhood in $E$ of a curve $\sigma$ in a fiber $S$. $\partial N_\epsilon(\sigma)$ is the torus boundary of $N_\epsilon(\sigma)$. This torus is also denoted by $T$.

- $F$ is the drilled bundle obtained by drilling $E$ along the curves $\{\sigma_i\}$. See Definition 1.2.

- $\Gamma$ denotes the fundamental group of $E$. $G$ denotes the fundamental group of $F$.

- $P_i$ denotes the $\mathbb{Z}^2$ subgroup of $G$ corresponding to the torus $T_i = \partial N_\epsilon(\sigma_i)$ and $\mathcal{P}$ denotes the collection $\{P_i\}$.

- $E_S$ is the cover of $E$ corresponding to $H$. $F_S$ is $E_S$ but drilled along all the lifts of $\{\sigma_i\}$.

- $M_r$ denotes a drilled atom. See Definition 3.2.

- $d_{\mathrm{pel}}$ is the partially electrocuted pseudometric on an appropriate space, defined in Section 3.4.

- $\mathcal{A}$ and $\mathcal{H}$ are used to denote annuli and hallways, respectively. See Definition 3.7.

- Given two lifts of fibers $\widetilde{S}_i$ and $\widetilde{S}_j$ inside $\widetilde{F}$, $\pi_{ij}$ denotes the projection map from the former to the latter.

- $\Sigma$ denotes a maximal connected subsurface of $S$ lying in the complement of the drilled curves.

**Note to the reader**   The nature of the problem necessitates a fair bit of case-by-case analysis, especially in Section 4, where we deal with a host of topological and geometric objects. Some of these need to be named for quick referencing later. To make this somewhat easier to handle, we have hyperlinked many of the technical terms in subsequent appearances. Clicking on these will take the reader to the line where the term is introduced.

## 2 Bundles, drilled bundles, and graphs of groups

Consider the exact sequence

(2-1) $$1 \to H \to \Gamma \to Q \to 1,$$

where $H = \pi_1(S)$ and $Q = \pi_1(\mathcal{G})$. Note that $\Gamma$ has a graph of groups structure, where each edge and vertex group equals $H$ and all edge-to-vertex maps are isomorphisms. Also, $E$ has a graph of spaces structure, where each edge and vertex space equals $S$, and all edge-to-vertex maps are homeomorphisms. A description of $\pi_1 E$ in general may then be given as follows (see [24, Section 2] for instance). Choose a maximal tree $T \subset \mathcal{G}$. Assume, without loss of generality, that for any edge $e \subset T$, the gluing maps $f_e^-$ are identity maps on $S$. Let $e_1, \ldots, e_n$ be the edges in $\mathcal{G} \setminus T$. For each $i \in \{1, \ldots, n\}$, write $f_i = f_{e_i}^+$. Also, set $\phi_i = (f_i)_* : \pi_1 S \to \pi_1 S$. Then $\pi_1 E$ is given by

$$\pi_1 E \cong \langle \pi_1 S, t_1, \ldots, t_n \mid t_i^{-1} s t_i = \phi_i(s), \ \forall s \in \pi_1 S, \ i \in \{1, \ldots, n\} \rangle.$$

**Hyperbolic $\Gamma$** It was shown by Hamenstädt [15] (see also [19; 28]) that, in the exact sequence (2-1), $\Gamma$ is hyperbolic if and only if $Q$ is a convex cocompact subgroup of the mapping class group in the sense of Farb and Mosher [11]. A useful fact that we will need is the following Scott–Swarup type theorem [29] due to Dowdall, Kent, and Leininger [9, Theorem 1.3]; see [26] for a different proof of the same result.

**Theorem 2.1** *Let*

$$1 \to H \to \Gamma \to Q \to 1$$

*be an exact sequence as above, so that $\Gamma$ is hyperbolic, and $H = \pi_1(S)$. Then any finitely generated infinite-index subgroup of $H$ is quasiconvex in $\Gamma$.*

### A graph of groups structure on $G = \pi_1(F)$

Recall that $F$ is the drilled bundle obtained from the bundle $E \to \mathcal{G}$. We now describe a standard graph of groups description for $\pi_1(F)$. We shall define $G := \pi_1(F)$. Denote the barycentric subdivision of $\mathcal{G}$ by $\mathcal{G}^*$.

The vertex set $\mathcal{V}(\mathcal{G}^*)$ thus consists of

(1) edges $e$ of $\mathcal{G}$, and

(2) vertices $v$ of $\mathcal{G}$.

The edge set $\mathcal{E}(\mathcal{G}^*)$ of $\mathcal{G}^*$ is given by incidences between edges $e$ and vertices $v$ of $\mathcal{G}$, ie if the terminal vertex $e^+$ of $e \in \mathcal{G}$ is $v$, then we introduce an edge in $\mathcal{G}^*$ between $e$ and $v$. Similarly for $e^-$.

Then $G(= \pi_1(F))$ has a graph of groups structure (see Figure 2), where:

(1) The underlying graph is $\mathcal{G}^*$.

(2) The edge groups are all equal to $H = \pi_1(S)$.

(3) The vertex groups $G_e$ (indexed by $e \in \mathcal{G}$) are given as follows. Since the bundle over the interior of $e \subset \mathcal{G}$ is $S \times (0, 1)$, $G_e$ equals $\pi_1(M_e)$, where $M_e$ is a possibly drilled copy of $S \times [0, 1]$.
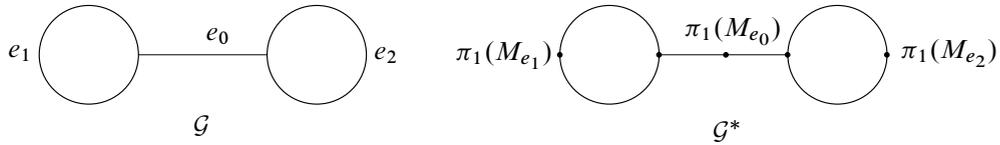
Figure 2: The graph of groups $\mathcal{G}^*$ for a given $\mathcal{G}$.

(4)  The vertex groups $G_v$ (indexed by the vertices $v$ of $\mathcal{G}$) are all $H$.

(5)  In particular, the edge-group to vertex-group maps are injective.

A convenient way to think of the graph of spaces associated to the above graph of groups description is as follows. For any edge $e \in \mathcal{G}$, the corresponding vertex space over $e \in \mathcal{G}^*$ is $M_e$, the possibly drilled copy of $S \times [0, 1]$. For any vertex $v \in \mathcal{G}$, the corresponding vertex space over $v \in \mathcal{G}^*$ is $S \times \text{star}(v)$, where $\text{star}(v)$ denotes a small closed neighborhood of $v$ in $\mathcal{G}$. The edge spaces are all given by $S$, and edge-to-vertex space inclusions are given by the inclusions of $S$ into either a drilled copy of $S \times [0, 1]$ or a copy of $S \times \text{star}(v)$.

Let $P_i = \pi_1(T_i) = \mathbb{Z} + \mathbb{Z}$. Then there exists $e$ such that $P_i \subset G_e = \pi_1(M_e)$. Whenever we are dealing with $G$ equipped with the above graph of groups structure, we shall *implicitly* reindex the collections $\{T_i\}$ and $\{\mathbb{P}_i\}$ so that the collection of tori contained in $M_e$ are indexed as $\{T_e^j\}$. (This notation will be used in Section 4 at the beginning of which we shall recall it.)

**Lemma 2.2** *Suppose that $\Gamma$ is hyperbolic. Let $\sigma_i$ denote a finite collection of curves that are drilled from $E$ to obtain $F$. Let $\mathcal{Q}_i \subset \Gamma$ denote the (conjugacy class of the) cyclic subgroup corresponding to $\sigma_i$. Then $\Gamma$ is strongly hyperbolic relative to the collection $\{\mathcal{Q}_i\}$.*

**Proof**  Since the $\sigma_i$ are simple closed curves, and since $\Gamma$ is hyperbolic, they denote primitive elements of $\Gamma$. Hence the collection $\{\mathcal{Q}_i\}$ denotes a malnormal quasiconvex family of subgroups. By a theorem of Bowditch [7, Theorem 7.11], $\Gamma$ is strongly hyperbolic relative to the collection $\{\mathcal{Q}_i\}$. $\qquad\square$

## 3  Relative hyperbolicity

Notation to be used in this section:

- $\Pi_S$ is the projection $F_S \to \widetilde{\mathcal{G}}$.

- $\mathcal{A}_S$ denotes an essential annulus (see Definition 3.7) in $F_S$. Whenever applicable, $B_S$ denotes the unique atomic wrapping annulus (see Definition 3.5) in $\mathcal{A}_S$. $\mathcal{A}_S^+$ and $\mathcal{A}_S^-$ are the two subannuli of $\mathcal{A}_S$ in the complement of $B_S$. The intersections of $\mathcal{A}_S^\pm$ with $B_S$ are denoted by $\theta^\pm$.

- In the context of Theorem 3.17, $Y$ is a graph of spaces with universal cover $X$ with the tree of spaces structure $\Pi \colon X \to \mathcal{T}$. Then $\rho$, $\lambda$, $H$, and $2m$ stand respectively for constants for thinness, hyperbolicity, girth, and length of a hallway or an annulus.

- $\widetilde{\Pi} \colon \widetilde{F}_S \to \mathcal{T}$ is the tree of spaces structure for $\widetilde{F}_S$. $X_v$ is used to denote vertex spaces of both $X$ and $\widetilde{F}_S$.

We refer the reader to [10; 7] for generalities on relative hyperbolicity.

This section and the next are devoted to proving the following:

**Theorem 3.1** *If $\Gamma$ is hyperbolic, then $G$ is (strongly) hyperbolic relative to the collection $\{P_i\}$.*

Theorem 3.1 says roughly that the result of drilling a hyperbolic 3-complex fibering over a graph gives a relatively hyperbolic 3-complex. This result is along the lines of earlier work of Belegradek and Hruska [3] (see also [2]), who prove similar results in a manifold context. The proof of Theorem 3.1 occupies the rest of this section and the next.

## 3.1 Scheme of the proof of Theorem 3.1

The proof of Theorem 3.1 consists of two steps as indicated at the end of the introduction.

**Step 1 (Section 3)** (hyperbolicity of $(\widetilde{F}, d_{\text{pel}})$) The aim of this section is to prove Theorem 3.14: hyperbolicity of the partially electrified space $(\widetilde{F}, d_{\text{pel}})$. Since the logic behind the proof has a number of ingredients involved, we lay out a sketch below.

**Partial electrification (Section 3.4)** We refer the reader to Section 3.4 for the precise notion of partial electrification. The construction that is relevant to this paper is the following. For each boundary torus $T$ of $F$, there is a distinguished meridional direction, so that $T$ is foliated by these meridians. Partial electrification electrifies each such meridian. This might be easier to see in the universal cover $\widetilde{T}$, where the elevations of each meridian are electrified. Since $\widetilde{T}$ may be identified with $\mathbb{R}^2$ foliated by parallel copies of $\mathbb{R}$ corresponding to elevations of the meridian, it follows that after electrifying each such copy of $\mathbb{R}$ in $\widetilde{T}$, the latter becomes quasi-isometric to $\mathbb{R}$, a hyperbolic space. This is in contrast with the usual electrification operation, where all of $\widetilde{T}$ is electrified to a space quasi-isometric to a point.

**Decomposition into atoms** We note next that $E$ and $F$ are built out of 3-manifolds in a natural way. The following definition captures this decomposition.

**Definition 3.2** (drilled and undrilled atoms) Recall that $\Pi \colon E \to \mathcal{G}$ denotes the surface bundle $E$ over the graph $\mathcal{G}$. Let $M_e = \Pi^{-1}(\bar{e}) \cap F$. If $e$ is a drilled edge, we refer to $M_e$ as a *drilled atom* of $F$; otherwise we refer to it as an *undrilled atom* of $F$ (see Figure 3). Elevations of drilled (resp. undrilled) atoms of $F$ to $\widetilde{F}$ will be referred to as drilled (resp. undrilled) atoms of $\widetilde{F}$.

If $\mathcal{L}$ is a maximal subgraph of $\mathcal{G}$, such that each edge of $\mathcal{L}$ is undrilled, $E(\mathcal{L}) = \Pi^{-1}(\mathcal{L}) \cap F$ will be called a *maximal undrilled subbundle* of $F$ (as well as of $E$). Elevations of such $E(\mathcal{L})$ to $\widetilde{E}_r$ or $\widetilde{F}$ will be called *elevated maximal undrilled subbundles*.

**Atoms and a combination theorem** In Section 3.2 we shall see that work of Thurston ensures that drilling out geodesics in hyperbolic 3-manifolds gives new noncompact 3-manifolds with a complete hyperbolic structure of finite volume. These drilled 3-manifolds are the atoms used in building the drilled
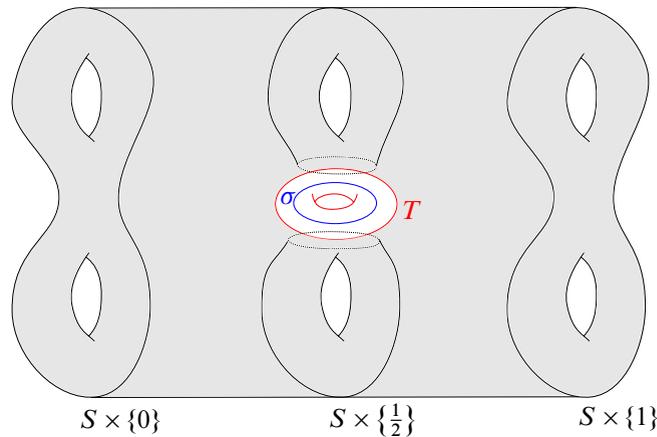
Figure 3: $S \times [0, 1]$ drilled along $\sigma$, or in other words, a drilled atom.

bundle $F$. It is not hard to see (Lemma 3.12) that after partial electrification, each atom has hyperbolic universal cover. It remains then to combine the partially electrified atoms using the underlying graph $\mathcal{G}^*$ to prove Theorem 3.14.

For this, we fall back upon the classical Bestvina–Feighn combination theorem (see Section 3.5.1) in Section 3.5 to piece the partially electrified atoms together. We shall use the annuli flare condition of Bestvina and Feighn to implement this.

It should be borne in mind here that the known combination theorems in the specific context of relative hyperbolicity such as [27] are inadequate for the purposes of this paper. This is the basic reason for using the more elaborate method of first proving a hyperbolic combination theorem in this section and then using it in the next using Sisto's generalization of Bowditch's guessing geodesics lemma to prove relative hyperbolicity.

**Annuli** In order to apply the Bestvina–Feighn combination theorem however, one needs to identify the essential annuli in $F$ (see Definition 3.7). This is what is done in Sections 3.2 and 3.3. Of these, Section 3.2 identifies essential annuli in atoms (see Definition 3.4), and Section 3.3 concatenates them together.

After having identified essential annuli in $F$ (see Definition 3.7), the combination theorem is applied to the partially electrified metric in Section 3.5.2 to prove Theorem 3.14.

**Step 2 (Section 4)** (guessing geodesics) The subsections of Section 4 follow a more straightforward linear logical order than Section 3. Our starting point is a necessary and sufficient condition for relative hyperbolicity due to Sisto building on earlier work of Bowditch and Hamenstädt. We recall this in Section 4.1.

Next, using the hyperbolic geodesics obtained as an output of Theorem 3.14 in Section 3, we guess a family of paths in $\widetilde{F}$ in Section 4.2. The key idea used here is that of electroambient quasigeodesics from [25]: a geodesic in the partially electrified $(\widetilde{F}, d_{\text{pel}})$ can be lifted in a more or less canonical way to a path in $\widetilde{F}$

(with the ordinary, nonelectrified metric). In each atom, the lifted path is a genuine quasigeodesic. This follows from the fact that each drilled atom (see Definition 3.2) has a universal cover that is hyperbolic *relative* to the elevations of its boundary tori. However, relative hyperbolicity of $\widetilde{F}$ turns out to be considerably more difficult to establish. We need to piece together the electroambient quasigeodesics in atoms carefully to guess a path family in $\widetilde{F}$. A fair bit of case-by-case analysis is necessary to get the path family. We therefore provide the reader with a short guide through the cases in Section 4.2.1.

Section 4.3 then establishes that the path family constructed in Section 4.2 satisfies the usual property of quasigeodesics in a relatively hyperbolic space. With the stability property in place, Sections 4.4 and 4.5 check the conditions of Section 4.1 needed to establish relative hyperbolicity. Of these subsections, the conditions checked in Section 4.4 are fairly routine, whereas Section 4.5 is devoted to proving the analog of the thin triangles property. This completes the proof of Theorem 3.1.

Finally, Section 4.6 establishes the relative quasiconvexity of essential subsurfaces of drilled fibers.

## 3.2   Drilling 3-manifolds

As mentioned in Section 3.1, the aim of this subsection is to define, identify, and classify essential annuli (see Definition 3.4) in atoms (cf Definition 3.2). We refer the reader to Figure 4 for a quick idea of the various annuli that may arise.

Let $M$ be either $S \times I$ or a hyperbolic 3-manifold fibering over the circle with fiber $S$. Thus, $M$ fibers over a compact 1-manifold (possibly with boundary) with fiber $S$. We proceed to drill simple closed curves $\sigma_i$ in some finite collection of fibers. A fiber $S'$ will be referred to as *undrilled* if $S'$ does not intersect any of the neighborhoods $N_\epsilon(\sigma_i)$ of the drilled curves. Let $M_r$ denote the drilled manifold. Henceforth, in this paper, $M_r$ will be equipped with

(1)   a complete hyperbolic structure of finite volume when $M$ fibers over the circle,

(2)   a geometrically finite hyperbolic structure with convex boundary otherwise.

In the latter case, we shall think of $M_r$ as the quotient of the convex hull of a geometrically finite representation $\rho \colon \pi_1(M_r) \to \mathrm{PSL}_2(\mathbb{C})$ such that the parabolics of $\rho(\pi_1(M_r))$ correspond precisely to the boundary tori of $M_r$. Since no two of the drilled curves are freely homotopic in $M$ by Definition 1.2, it follows that $M_r$ is atoroidal. Since $M_r$ is an atoroidal Haken manifold, the existence of such hyperbolic structures in both cases follows from Thurston's theorem [18, Chapter 15].

**Convention**   We shall choose hyperbolic structures as above on $M_r$ and fix them for the rest of the discussion.

For such a hyperbolic structure on $M_r$, $M_r$ is finite volume without boundary if $M$ fibers over the circle. Otherwise the boundary consists of two copies of $S$, corresponding to the (singular) fibers (see Definition 1.2) over $0, 1 \in I$. Further, $M_r$ has finitely many rank-two cusps (corresponding to the toroidal boundaries of regular neighborhoods of drilled curves). We start with the following observation:

**Lemma 3.3** *Let $M$ and $M_r$ be as above where the set of drilled curves correspond to a finite nonempty collection $\sigma_i \in S_i$ of simple closed curves on fibers $S_i$. Any undrilled fiber $S'$ is geometrically finite in $M_r$. Any connected components of $S_x \setminus \sigma$ of a drilled fiber is also geometrically finite in $M_r$.*

**Proof** This follows from the covering theorem [32, Theorem 9.2.2; 8]. □

**Essential annuli in atoms** We now describe the essential annuli in atoms.

**Definition 3.4** An *immersed essential annulus* or simply an *essential annulus* $\mathcal{A}$ in a 3-manifold $M$ with nonempty boundary is an immersion $i : (\mathcal{A}, \partial\mathcal{A}) \to (M, \partial M)$ such that

(1) $i_* : \pi_1(\mathcal{A}) \to \pi_1(M)$ is injective,

(2) $i$ is not homotopic rel. boundary into $\partial M$.

When $M$ is a 3-manifold fibering over the circle with a distinguished singular fiber $S_x$ (see Definition 1.2), an *immersed essential annulus* or simply an *essential annulus* $\mathcal{A}$ is an immersion $i : (\mathcal{A}, \partial\mathcal{A}) \to (M, S_x)$ such that

(1) $i_* : \pi_1(\mathcal{A}) \to \pi_1(M)$ is injective,

(2) $i$ is not homotopic rel. boundary into $S_x$.

Henceforth, we shall refer to immersed essential annuli simply as essential annuli. In the 3-manifold literature, essential annuli often refer to embedded essential annuli in $(M, \partial M)$. However, since we shall not have any special use for embedded annuli, we shall use the terminology from Definition 3.4 in this paper. The usage is consistent with the more group-theoretic notion in [5]; see Definition 3.7 below. Note that essential annuli $\mathcal{A}$ in fibered $M$ are allowed to intersect $S_x$ in the interior of $\mathcal{A}$ as well, ie $\mathcal{A}$ is allowed to "wrap around transverse to the fibers" of $M$ multiple times so long as $\partial\mathcal{A}$ maps to $S_x$. Essential annuli in undrilled atoms (see Definition 3.2) are given by the following, *up to homotopy*:

(1) If $M = S \times I$, then essential annuli are of the form $\sigma \times I$, where $\sigma$ is an essential, possibly immersed curve in $S$. After homotopy, we may assume that $\sigma$ is a geodesic in some auxiliary hyperbolic structure on $S$.

(2) If $M$ fibers over the circle, let $M_{\mathbb{Z}}$ denote the cover of $M$ corresponding to $\pi_1(S)$, so that $M_{\mathbb{Z}}$ is homeomorphic to $S \times \mathbb{R}$. Then essential annuli in $M_{\mathbb{Z}}$ are concatenations of annuli $\mathcal{A}_i \subset S \times [i, i+1]$ as in item (1). Further, $\sigma_{i+1} = \mathcal{A}_i \cap S \times \{i+1\} = \mathcal{A}_{i+1} \cap S \times \{i+1\}$, so that $\mathcal{A}_i$ and $\mathcal{A}_{i+1}$ may be concatenated along the essential, possibly immersed curve $\sigma_{i+1}$.

Essential annuli in drilled atoms (see Definition 3.2) are a bit more involved. Let $M = S \times I$, and $M_r$ be a drilled atom obtained from $M$ (see Definition 3.2). Let $\sigma_1, \ldots, \sigma_m$ denote the drilled curves on $S$. Realize the $\sigma_i$ by geodesics in an auxiliary hyperbolic structure on $S$. Let $\Sigma_0$ denote the subsurface of $S$ filled by $\bigcup_i \sigma_i$, ie adjoin to $\bigcup_i \sigma_i$ all simply connected complementary regions. Then essential annuli in $M_r$ (see Definition 3.4) are of three kinds after homotopy:
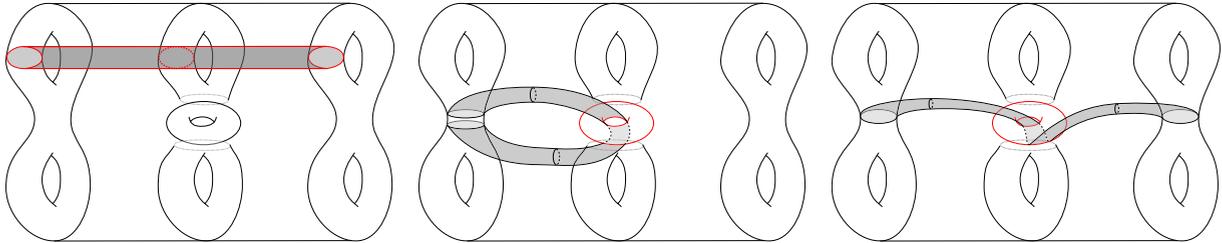
Figure 4: The three types of annuli in a drilled $S \times I$.

(1)   essential annuli of the form $\gamma \times I$, respecting the product structure of $S \times I$, where $\gamma$ is an essential, possibly immersed curve in $S$; in this case, $\gamma$ necessarily lies in (a component of) $S \setminus \Sigma_0$,

(2)   essential annuli with both boundary curves on either $S \times \{0\}$ or $S \times \{1\}$ with core curve homotopic to a multiple of one of the $\sigma_i$; in this case, $\mathcal{A}$ wraps around $N_\epsilon(\sigma_i)$ finitely many times,

(3)   essential annuli with one boundary curve on $S \times \{0\}$, and one on $S \times \{1\}$ with core curve homotopic to a multiple of one of the $\sigma_i$, such that $\mathcal{A}$ wraps around $N_\epsilon(\sigma_i)$ finitely many times.

**Definition 3.5**   An annulus from case (2) above is referred to as an *atomic wrapping backtracking annulus*. An annulus from case (3) is referred to as an *atomic wrapping nonbacktracking annulus*.

Next, suppose that $M$ fibers over the circle with monodromy $\Phi$. Let $S_x \subset M$ denote the unique singular fiber (see Definition 1.2). Let $S_1, \ldots, S_m$ denote drilled fibers and $\sigma_i \subset S_i$ for $i = 1, \ldots, m$ denote drilled simple closed curves. Let $M_r$ denote $M$ after drilling. Let $M_\mathbb{Z}$ denote the cover of $M$ corresponding to $\pi_1(S)$ and $M_S$ denote the elevation of $M_r$ to $M_\mathbb{Z}$. Let $\ldots, S_x^{-1}, S_x^0, S_x^1, \ldots$ denote the elevations of $S_x$ to $M_S$. Without loss of generality, assume that the elevated annulus $\mathcal{A}_1$ starts on $S_x^0$. Also, let $\Sigma_n$ denote the subsurface of $S$ filled by $\bigcup_{i=1,\ldots,m} \bigcup_{j=0,\ldots,n-1} \Phi^j(\sigma_i)$, ie adjoin to $\bigcup_{i=1,\ldots,m} \bigcup_{j=0,\ldots,n-1} \Phi^j(\sigma_i)$ all simply connected complementary regions of $S$.

**Lemma 3.6**   *There exists $N \in \mathbb{N}$ such that $\Sigma_n = S$ for all $n \geq N$.*

**Proof**   Since $\Phi$ is pseudo-Anosov, there exists $N \in \mathbb{N}$ such that

$$d_{\mathcal{C}(S)}(\sigma_i, \Phi^n(\sigma_i)) \geq 3$$

for all $n \geq N$, where $d_{\mathcal{C}(S)}$ denotes distance in the curve complex of $S$. Hence, $\sigma_i$ and $\Phi^n(\sigma_i)$ fill $S$ for all for all $n \geq N$.    □

Essential annuli (see Definition 3.4) in $M_r$ for $M$ a fibered 3-manifold will be described in greater detail and in a more general setting in Section 3.3 below. For now, it suffices to say that an essential annulus in $M_r$ lifts to an essential annulus in $M_S$ starting and ending on $S_x^i$ and $S_x^j$ for some $i, j \in \mathbb{Z}$.

### 3.3 Essential annuli in $F$

As mentioned in Section 3.1, the aim of this subsection is to concatenate the atomic essential annuli (see Definition 3.4) of Section 3.2 above to obtain essential annuli in $F$ (see Definition 3.7).

We use the classification of essential annuli (see Definition 3.7) in $M_r$ described above to identify essential annuli in the drilled bundle $F$. Since $F \subset E$, an essential annulus $\mathcal{A} \subset F$ may also be regarded as a subset of $E$. The formalism we use is due to Bestvina and Feighn [5].

**Definition 3.7** [5, page 87] Let $\mathcal{G}$ be a graph, and $Y$ a graph of spaces with base graph $\mathcal{G}$, so that the maps of edge-spaces to vertex spaces are injective at the level of the fundamental group. Let $X = \tilde{Y}$ be the universal cover, and $\mathcal{T}$ the resulting tree of spaces, whose vertex and edge spaces are universal covers of vertex and edge spaces of $Y$. Let $m$ be a positive integer and $I$ denote the closed unit interval. A *hallway* in $X$ is a map $\mathcal{H} \colon [-m, m] \times I \to X$ with the following properties:

(1)  $\mathcal{H}^{-1}(\cup X_e) = \{-m, -m+1, \ldots, m\} \times I$,

(2)  $\mathcal{H}$ is transverse to $\bigcup X_e$ relative to the previous condition,

(3)  for each $i \in \{-m, -m+1, \ldots, m\}$, the image of $\mathcal{H}$ restricted to $\{i\} \times I$ is a geodesic in the corresponding edge space.

We say that such a hallway has *length* $2m$. The *girth* of the hallway is the length of the curve $\mathcal{H}(\{0\} \times I)$. A hallway is *essential* if the projection of $\mathcal{H}$ onto the base tree $\mathcal{T}$ is a path which does not backtrack.

A map $\mathcal{A} \colon [-m, m] \times S^1 \to Y$ is an *annulus* if the lift $\tilde{\mathcal{A}} \colon [-m, m] \times I \to X$ is a hallway. An annulus $\mathcal{A}$ is *essential* if the hallway $\tilde{\mathcal{A}}$ is essential.

An essential annulus $\mathcal{A} \colon [0, m] \times S^1 \to Y$ is said to *start at* $Y_v$ if $\mathcal{A}(\{0\} \times S^1) \subset Y_v$.

We now relate the notion of essential annuli (see Definition 3.4) in 3-manifolds with those in Definition 3.7 using Definition 3.2. Recall from Section 2 that, $F$ admits a graph of spaces structure, with underlying graph $\mathcal{G}^*$, where the vertex spaces of $F$ are drilled or undrilled atoms (see Definition 3.2). Let $\tilde{\mathcal{G}}$ denote the universal cover of $\mathcal{G}^*$.

Let $F_S$ be the cover of $F$ corresponding to the kernel of the map $\Pi_* \colon \pi_1(F) \to \pi_1(\mathcal{G})$. Let $\Pi_S \colon F_S \to \tilde{\mathcal{G}}$ denote the projection map from $F_S$ to $\tilde{\mathcal{G}}$. Then $F_S$ admits a graph of spaces structure, where:

(1)  Each vertex space is a copy of $S \times I$ with, possibly, some curves drilled. The interior of each vertex space is given by one of the following:

  •  $M \setminus S_x$ if $M$ is an undrilled atom of $F$ fibering over the circle (see Definition 3.2),

  •  $M_r \setminus S_x$ if $M_r$ is a drilled atom of $F$ (see Definition 3.2) fibering over the circle,

  •  $M \setminus S \times \{0, 1\}$ if $M$ is an undrilled atom of $F$ fibering over $I$ (see Definition 3.2),

  •  $M_r \setminus S \times \{0, 1\}$ if $M_r$ is a drilled atom of $F$ (see Definition 3.2) fibering over $I$.

(2)  Each edge space is $S$.

Thus, $F_S$ is a drilled graph bundle over the graph $\widetilde{\mathcal{G}}$ (see Definition 1.2). Note that $E_S$ also has a similar graph of spaces decomposition over $\widetilde{\mathcal{G}}$ where all the vertex spaces are copies of $S \times I$, and each edge space is $S$. With this structure in mind, $F_S$ embeds into $E_S$ as a graph of spaces over $\widetilde{\mathcal{G}}$. We now describe essential annuli in $F_S$ in terms of the description of in atoms given in Section 3.2.

Let $\mathcal{A}$ denote an essential annulus in $F$ (see Definition 3.7). Let $\mathcal{A}_S$ denote an elevation of $\mathcal{A}$ to $F_S$. Then the image $\Pi_S(\mathcal{A}_S)$ is an unparametrized geodesic $[a, b]$ in $\widetilde{\mathcal{G}}$. Let $a = v_0, v_1, \ldots, v_n = b$ denote the vertices of $[a, b]$. Let $F_S^i$ denote the vertex space of $F_S$ corresponding to $v_i$. Then $\mathcal{A}_S$ is given by a concatenation of essential annuli in the atoms $F_S^i$. We shall show now that $\mathcal{A}_S$ comes broadly in two flavors: backtracking and nonbacktracking.

**Definition 3.8** Suppose $\mathcal{A}_S$ is an essential annulus (see Definition 3.7) of the form $\gamma \times [p, q]$, respecting the topological product structure of $F_S$, such that

(1) $\gamma$ is an essential, possibly immersed curve in $S$,

(2) $\gamma \times \{0\}$ lies on $S_x^0$, and $\gamma \times \{n\}$ lies on $S_x^n$.

Then we shall refer to such an essential annulus (see Definition 3.7) as a *nonbacktracking annulus*.

In this case, $\sigma$ necessarily lies in (a component of) $S \setminus \Sigma_n$ (for instance, by the annulus theorem). In this case, there are no atomic wrapping annuli in the sense of Section 3.2 contained in $\mathcal{A}_S$.

Next, suppose $\mathcal{A}_S$ is an essential annulus (see Definition 3.7) containing an atomic wrapping annulus in the sense of Definition 3.5. Recall that $E_S$ denotes the surface bundle over $\widetilde{\mathcal{G}}$ (see Definition 1.2) corresponding to the cover of $E$ corresponding to $\pi_1(S)$, so that $F_S$ is obtained from $E_S$ by drilling a family of nonhomotopic simple closed curves. Then, there exists a drilled curve $\eta$ in $E_S$ such that $\mathcal{A}_S$ wraps around the boundary of $N_\epsilon(\eta)$ in $F_S$. Hence, the core curve of $\mathcal{A}_S$ is homotopic to a nontrivial power of $\eta$. We observe the following:

**Lemma 3.9** *Suppose $\mathcal{A}_S$ is an essential annulus (see Definition 3.7) containing an atomic wrapping annulus (see Definition 3.5). Then $\mathcal{A}_S$ contains exactly one atomic wrapping annulus.*

**Proof** Since $\mathcal{A}_S$ is an essential annulus containing an atomic wrapping annulus, it has core curve freely homotopic to a (nontrivial power of a) drilled curve $\eta$. Note next that, by Definition 1.2 and Condition 1.3, no two distinct elevations of drilled curves to $E_S$ are freely homotopic. Hence, $\mathcal{A}_S$ contains exactly one atomic wrapping annulus. $\square$

Essential annuli (see Definition 3.7) $\mathcal{A}_S$ in $F_S$ containing an atomic wrapping annulus (see Definition 3.5) are of two kinds, depending on the nature of the atomic wrapping annulus. Let $B_S \subset \mathcal{A}_S$ denote the unique atomic wrapping annulus contained in $\mathcal{A}_S$ with core curve homotopic to $\eta^m$ for $m \neq 0$.

(1) Both boundary curves of $B_S$ lie on a single singular fiber (see Definition 1.2) $S_y$. In this case, $B_S$, and hence $\mathcal{A}_S$, wraps around $N_\epsilon(\eta)$ finitely many times. (Note that $B_S$ is an atomic wrapping backtracking annulus in the sense of Definition 3.5.) We refer to $\mathcal{A}_S$ as an *annulus with backtracking*.

(2) There exists a drilled atom $M_e$ (see Definition 3.2) with boundary surfaces $S_y$ and $S_z$, such that the distinct boundary curves of $B_S$ lie on the distinct boundary surfaces $S_y$ and $S_z$. Again, $B_S$, and hence $\mathcal{A}_S$, wraps around $N_\epsilon(\eta)$ finitely many times, but otherwise is required to respect the product structure on $M_e$. (Note that $B_S$ is an atomic wrapping, nonbacktracking annulus tracking annulus.) We shall refer to $\mathcal{A}_S$ as an annulus *with wrapping but no backtracking*.

**Definition 3.10** An essential annulus (see Definition 3.7) $\mathcal{A}_S$ in $F_S$ is said to *intersect* a drilled atom (see Definition 3.2) $M_r \subset F_S$ if $\mathcal{A}_S$ contains an elementary annulus $B_S \subset M_r$. (Note that $B_S$ may be a nonbacktracking annulus (see Definition 3.8), an atomic wrapping backtracking annulus, or an atomic wrapping nonbacktracking annulus (see Definition 3.5).)

**Lemma 3.11** *There exists $D \in \mathbb{N}$ such that the following holds*: *Let $\mathcal{A}_S \subset F_S \subset E_S$ be an essential annulus intersecting (see Definition 3.10) drilled atoms $M_a$ and $M_b$ (see Definition 3.2), where $a, b \in \widetilde{\mathcal{G}}$. Then*

$$d_{\widetilde{\mathcal{G}}}(a, b) \le D.$$

**Proof** This is similar to Lemma 3.6. It was shown by Kent and Leininger [19] and Hamenstädt [15] that the following are equivalent:

(1) $\pi_1(E)$ is hyperbolic,

(2) $\pi_1(\mathcal{G})$ acts on the curve complex $\mathcal{C}(S)$ with qi-embedded orbits,

(3) $\pi_1(\mathcal{G})$ is a convex-cocompact subgroup of the mapping class group $\mathrm{MCG}(S)$ in the sense of [11].

Since $\pi_1(E)$ is hyperbolic by assumption, it follows that $\pi_1(\mathcal{G})$ acts on the curve complex $\mathcal{C}(S)$ with qi-embedded orbits. Let $i : \widetilde{\mathcal{G}} \to \mathcal{C}(S)$ be the induced qi-embedding. Hence, there exists $D \in \mathbb{N}$ such that if $d_{\widetilde{\mathcal{G}}}(a, b) > D$, then $d_{\mathcal{C}(S)}(i(a), i(b)) > 3$. Let $\sigma_a$ and $\sigma_b$ denote drilled curves in $M_a$ and $M_b$, respectively, so that $\sigma_a$ and $\sigma_b$ fill $S$.

If the core curve of $\mathcal{A}_S$ intersects both $M_a$ and $M_b$ in the sense of Definition 3.10, it must be disjoint from the subsurface of $S$ filled by $\sigma_a$ and $\sigma_b$, an impossibility. $\square$

## 3.4 Partial electrification

Recall that the bundle $\Pi : E \to \mathcal{G}$ restricts to a map $\Pi_r : F \to \mathcal{G}$.

**Partial electrification of $\partial N_\epsilon(\sigma_i)$** Let $M_r$ denote a drilled atom in $F$ (see Definition 3.2), and let $\partial N_\epsilon(\sigma_i)$ denote the boundary of a drilled curve in $M_r$. Choose a homeomorphism of $\partial N_\epsilon(\sigma_i)$ with $S^1 \times S^1$, and assume, without loss of generality, that each $S^1 \times \{t\}$ is a meridian. Equip $S^1 \times S^1$ with a product metric, where the first factor is given the zero metric, and the second factor the standard round metric of radius one. We refer to the resulting path-pseudometric on $\partial N_\epsilon(\sigma_i)$ as *the partially electrified path-pseudometric* and denote it by $d_{\mathrm{pel}}$.

The reader at this stage might wish to refer back to Section 3.1 (paragraph titled "Partial electrification") for the heuristic idea behind the notion of partial electrification introduced above. Further, the paragraph titled "Atoms and a combination theorem" in Section 3.1 says briefly why we need this technique.

**Partial electrification of drilled atoms**   Recall that any drilled $M_r$ of $F$ has been equipped with the restriction of a complete hyperbolic metric. If $M_r$ is obtained by drilling $S \times I$, then the surface boundary components $S \times \{0, 1\}$ are convex. Removing a small neighborhood of the cusps we obtain boundary tori $\partial N_\epsilon(\sigma_i)$ equipped with flat Euclidean metrics. Abusing notation slightly, we continue to refer to the resulting compact 3-manifold with boundary also as $M_r$. Rescaling the hyperbolic metric if necessary, we may assume that $\partial N_\epsilon(\sigma_i)$ has a product metric as in the previous paragraph. Replace each such metric by the partially electrified path-pseudometric $d_{\text{pel}}$ described in the previous paragraph. We now consider a family of paths, each of which is given by a concatenation of pieces that either

(1)  have interior disjoint from the boundary tori $\{\partial N_\epsilon(\sigma_i)\}$ (the length of such a piece is given by its hyperbolic length), or

(2)  lie entirely in some boundary torus $\{\partial N_\epsilon(\sigma_i)\}$ (the length of such a path is given by its length with respect to $d_{\text{pel}}$ on $\{\partial N_\epsilon(\sigma_i)\}$).

Then the length of a path is given by the sum of the above pieces. The resulting path-pseudometric on $M_r$ is referred to as *the partially electrified path-pseudometric on $M_r$* and is also denoted by $d_{\text{pel}}$.

If the atom $M_r$ of $F$ is obtained from an atom $M$ of $E$ that fibers over the circle, then any elevation $M_S$ of $M_r$ to $F_S$ is a cyclic cover of $M_r$ corresponding to the natural epimorphism to $\mathbb{Z}$ inherited from $M$. Then $M_S$ is a concatenation of a $\mathbb{Z}$'s worth of atoms of $F_S$. Each atom $M_e$ of $F_S$ is equipped with the inherited path metric from $M_S$, and the resulting path-pseudometric on $M_e$ is referred to as *the partially electrified path-pseudometric on $M_e$* and is also denoted by $d_{\text{pel}}$.

The universal cover of $M_r$ will be denoted by $\widetilde{M}_r$. The lift of the partially electrified path-pseudometric on $M_r$ to $\widetilde{M}_r$ is referred to as *the partially electrified path-pseudometric on $\widetilde{M}_r$* and is also denoted by $d_{\text{pel}}$. Similarly for $M_e$.

**Lemma 3.12**   *There exist $\delta \geq 0$, $K \geq 1$, and $\epsilon \geq 0$ such that the following hold: Let $M_e$ denote an atom of $F_S$. Let $S_x$ denote a surface boundary component of $M_e$. Let $\mathcal{P}$ denote the collection of elevations of the tori $\{\partial N_\epsilon(\sigma_i)\}$ to $\widetilde{M}_e$. Then:*

(1)  *$\widetilde{M}_e$ is strongly hyperbolic relative to $\mathcal{P}$.*

(2)  *$(\widetilde{M}_e, d_{\text{pel}})$ is $\delta$-hyperbolic.*

(3)  *The inclusion of $\widetilde{S}_x$ into $(\widetilde{M}_e, d_{\text{pel}})$ is a $(K, \epsilon)$-qi-embedding for any elevation $\widetilde{S}_x$ of $S_x$.*

**Proof**   Since a hyperbolic structure on $M_e$ may be chosen so that it has convex boundary, $\widetilde{M}_e$ is strongly hyperbolic relative to $\mathcal{P}$. (This is a consequence of the main theorem of [10].) The second conclusion is then a special case of [25, Lemma 1.20].

Let $d$ denote the metric on $\widetilde{M}_e$ lifted from the intrinsic path-metric on $M_e$. Let $d_e$ denote the electric metric on $\widetilde{M}_e$ after electrifying the elements of $\mathcal{P}$ as in [10]. For $u, v \in \widetilde{S}_x$, let $\gamma_{uv}$, $\gamma_{uv}^e$, and $\gamma_{uv}^p$ denote geodesics with respect to $d$, $d_e$, and $d_{\text{pel}}$, respectively. The second conclusion will follow from two facts:

(1) By [25, Lemma 1.21; 10, Lemma 4.5 and Proposition 4.6], $\gamma_{uv}$, $\gamma_{uv}^e$, and $\gamma_{uv}^p$ track each other (uniformly, independent of $u$ and $v$) away from $\mathcal{P}$. (See Lemma 4.4 below for a slightly more general statement.)

(2) The nearest-point projections of elements of $\mathcal{P}$ equipped with $d_{\text{pel}}$ onto $\widetilde{S}_x$ are either uniformly bounded in diameter, or uniform quasi-isometric embeddings.

In fact, any $P \in \mathcal{P}$ is stabilized by a conjugate of $\pi_1(N_\epsilon(\sigma_i)) = \mathbb{Z}^2$ for some $i$. Let $\text{stab}(P)$ denote the stabilizer of $P$. Then, $\text{stab}(P) \cap \widetilde{S}_x$ is either trivial or infinite cyclic.

In the former case, $\gamma_{uv}^e$ (after a small isotopy if necessary) may be assumed to be disjoint from $P$. In the latter case, if $\gamma_{uv}^e$ does enter and exit $P$ at $y$ and $z$, respectively, then there exist $y_1, z_1 \in \widetilde{S}_x$ such that the geodesic joining $y_1$ and $z_1$ lies uniformly close to an elevation of $\sigma_i$. It follows that $\gamma_{uv}$ and $\gamma_{uv}^p$ track each other throughout their lengths. The third conclusion follows. $\square$

**Partial electrification of $F$ and $\widetilde{F}$**   The metric on each of the atoms after drilling (and before partial electrification) is denoted by $d$. Equip $F$ with a $C^0$ piecewise Riemannian metric that is bi-Lipschitz to the metric $d$ on each of the atoms. We refer to this metric on $F$ also by $d$. Now, consider rectifiable paths $\sigma : [0, 1] \to F$ consisting of finitely many pieces, each of which is contained in an atom. The length of $\sigma$ is declared to be the sum of the lengths of these subpaths. Replacing $d$ on each atom by the partially electrified metric $d_{\text{pel}}$ on that atom, we obtain a partially electrified path pseudometric, also denoted by $d_{\text{pel}}$, on $F$.

Lifting $d$ and $d_{\text{pel}}$ to the universal cover $\widetilde{F}$, we obtain a metric $d$ and a partially electrified path pseudometric $d_{\text{pel}}$, respectively. The distance between a pair of point $u, v$ is then obtained by passing to an infimum over all paths $\sigma$ as above joining $u$ and $v$.

**Remark 3.13**   The partial electrification construction above is adapted from [27] (see [25, Section 1.3] for a summary).

## 3.5   Partial electrification and relative hyperbolicity

As mentioned in Section 3.1, the aim of this subsection is to use the classification of essential annuli from Sections 3.2 and 3.3 and the Bestvina–Feighn combination theorem (Theorem 3.17 below) to prove Theorem 3.14 below. Note that the statement of Theorem 3.14 involves the partially electrified metric defined in Section 3.4.

Let $\widetilde{F}$ denote the universal cover of $F$. We lift the pseudometric $d_{\text{pel}}$ to $\widetilde{F}$ to obtain a partially electrified metric on $\widetilde{F}$ denoted again by $d_{\text{pel}}$. The following is the main theorem of this section.

**Theorem 3.14**   $(\widetilde{F}, d_{\text{pel}})$ *is hyperbolic.*

To prove Theorem 3.14, we shall use the Bestvina–Feighn combination theorem (Theorem 3.17).

### 3.5.1 The Bestvina–Feighn combination theorem

**Definition 3.15** [5] Let $\mathcal{G}$ be a graph, and $Y$ a graph of spaces with base graph $\mathcal{G}$, so that the maps of edge-spaces to vertex spaces are injective at the level of the fundamental group. Let $X = \widetilde{Y}$ be the universal cover, and $\mathcal{T}$ the resulting tree of spaces, whose vertex and edge spaces are universal covers of vertex and edge spaces of $Y$. Suppose that the following hold:

(1) Vertex spaces $\{X_v\}$ and edge spaces $\{X_e\}$ are all $\delta$-hyperbolic metric spaces for some $\delta > 0$.

(2) There exist $K \geq 1$ and $\epsilon \geq 0$ such that the maps of edge-spaces to vertex spaces for $X$ are all $(K, \epsilon)$-quasi-isometric embeddings.

Then $Y$ is said to be a graph of hyperbolic metric spaces satisfying the qi-embedded condition.

Recall the notion of hallways and annuli from Definition 3.7.

**Definition 3.16** [5, page 87] For $\lambda > 1$, a hallway $\mathcal{H}$ is said to be $\lambda$-*hyperbolic* if

$$l(\mathcal{H}(\{0\} \times I)) \leq \frac{1}{\lambda} \max\{l(\mathcal{H}(\{-m\} \times I)), l(\mathcal{H}(\{m\} \times I))\},$$

where $l$ denotes the length of the path in the corresponding edge space.

Let $\rho > 0$. Given $i \in \{-m, -m+1, \ldots, m\}$, denote the vertex space that $\mathcal{H}([i, i+1] \times I)$ lies in by $X_{v_i}$. The hallway is $\rho$-*thin* if for all such $i$ and for any $t \in I$, $d_{X_{v_i}}(\mathcal{H}(i, t), \mathcal{H}(i+1, t)) < \rho$.

The *girth* (resp. *length*) of the annulus $\mathcal{A}$ is the girth (resp. *length*) of the induced hallway $\tilde{\mathcal{A}}$.

Similarly, the rest of the terminology, ie hyperbolicity, thinness, essentiality, for the annulus $\mathcal{A}$, is defined via $\tilde{\mathcal{A}}$.

**The annuli flare condition** The graph of spaces $Y$ (with base graph $\mathcal{G}$) satisfies the *annuli flare condition* if there exist $\lambda > 1$ and $m \geq 1$ such that the following holds: given any $\rho > 0$, there exists a constant $H(\rho)$ so that whenever $\mathcal{A}$ is a $\rho$-thin essential annulus (see Definition 3.7) of length $2m$ and girth at least $H(\rho)$, $\mathcal{A}$ is $\lambda$-hyperbolic. The graph of spaces $Y$ satisfies the *weak annuli flare condition* if there are numbers $\lambda > 1, m > 1$, and $H$ such that any $4\delta$-thin essential annulus (see Definition 3.7) of length $2m$ and girth at least $H$ is $\lambda$-hyperbolic. The following statement gives the Bestvina–Feighn combination theorem in a consolidated form.

**Theorem 3.17** [5, Theorem 3.2 of the correction] *Let* $\Pi \colon Y \to \mathcal{G}$ *be a graph of spaces whose vertex and edge spaces have* $\delta$-*hyperbolic universal covers for some* $\delta > 0$. *If* $\Pi \colon Y \to \mathcal{G}$ *satisfies the following conditions, then the universal cover* $X$ *of* $Y$ *is hyperbolic*:

(1) *the quasi-isometrically embedded condition (see Definition 3.15),*

(2) *the annuli flare condition or the weak annuli flare condition.*

The following definition adapts [23, Definition 4.26] to our context.

**Definition 3.18**  We say that a vertex space $Y_v$ flares in all directions weakly if there are numbers $\lambda > 1$, $m > 1$, and $H$ such that if $\mathcal{A} \colon S^1 \times [0, m] \to Y$ is any $4\delta$-thin essential annulus (see Definition 3.7) of length $m$ and girth at least $H$ *starting at $Y_v$* (in the sense of Definition 3.7), then the associated lifted hallway $\mathcal{H}$ satisfies

$$l(\mathcal{H}(\{0\} \times I)) \leq \frac{1}{\lambda} l(\mathcal{H}(\{m\} \times I)).$$

We shall need a modification of Theorem 3.17 to guarantee global quasiconvexity of a vertex space. The following now adapts [23, Proposition 4.27] to our context.

**Corollary 3.19**  *Let $\Pi \colon Y \to \mathcal{G}$ be a graph of spaces satisfying the conditions of Theorem 3.17. Further, let $Y_v \subset Y$ be a vertex space that flares in all directions (see Definition 3.18) weakly. Then $\widetilde{Y}_v$ is quasiconvex in $\widetilde{Y}$.*

**Proof**  Since the context of [23, Proposition 4.27] is slightly different, we sketch the mild modifications necessary. We recall a construction from [22, Section 3] here. It follows from [22, Theorem 3.8] that there exists $C > 0$ such that the following holds. Given any geodesic $\lambda \subset (\widetilde{Y}_v, d_v)$, there exists a "ladder" $\mathcal{L}_\lambda \subset \widetilde{Y}$ containing $\lambda$, such that $\mathcal{L}_\lambda$ is $C$-quasiconvex. Hyperbolicity of $(\widetilde{Y}, d)$ now guarantees that for all such geodesics $\lambda \subset (\widetilde{Y}_v, d_v)$, $\mathcal{L}_\lambda$ is hyperbolic. The construction of $\mathcal{L}_\lambda$ in [22, Section 3] now shows that $\widetilde{\Pi} \colon \widetilde{Y} \to \widetilde{\mathcal{G}}$ restricts to $\widetilde{\Pi}_\lambda \colon \mathcal{L}_\lambda \to \mathcal{T}$, where $\mathcal{T} \subset \widetilde{\mathcal{G}}$ is a tree. Further, for any vertex $w \in \mathcal{T}$, $\widetilde{\Pi}_\lambda^{-1}(w)$ is a geodesic segment in the vertex space $\widetilde{\Pi}^{-1}(w) \subset \widetilde{Y}$. Thus, $\mathcal{L}_\lambda$ is a tree of spaces, where each vertex space is isometric to an interval. The hypothesis that $Y_v \subset Y$ be a vertex space that flares in all directions weakly guarantees that $\mathcal{L}_\lambda$ flares in all directions also. Hence $\lambda$ is a quasigeodesic (with uniform constants) in $\mathcal{L}_\lambda$. Since $\mathcal{L}_\lambda$ is $C$-quasiconvex, $\lambda$ is a quasigeodesic (with uniform constants) in $\widetilde{Y}$.  $\square$

**Remark 3.20**  More generally, if $Z \subset Y_v$ is a subspace such that

(1)  the inclusion induces an injection at the level of fundamental groups,

(2)  $\widetilde{Z}$ is qi-embedded in $\widetilde{Y}_v$,

then an auxiliary vertex $w$ and an edge $e = [w, v]$ may be added to the base graph $\mathcal{G}$, so that $Y_w = Y_e = Z$. Definition 3.18 and Corollary 3.19 may thus be applied to such subspaces $Z$ of $Y_v$ as well. If $Z$ flares in all directions (see Definition 3.18) weakly, then $\widetilde{Z}$ is quasiconvex in $\widetilde{Y}$ by Corollary 3.19.

We also record the following, where we assume implicitly that there is a cocompact group action so that the annuli flare condition makes sense. (We also refer the reader to [5, page 90 and Section 4 of the correction] and for an analogous hallways flare condition.)

**Lemma 3.21**  *Let $\Pi_\mathcal{T} \colon X \to \mathcal{T}$ be a tree of spaces obtained as a universal cover of a graph of compact spaces. Suppose that each vertex space $X_v$ and edge space $X_e$ of $X$ is $\delta$-hyperbolic. Further, suppose that the following conditions are satisfied:*

(1)  *the quasi-isometrically embedded condition* (*see Definition 3.15*),

(2)  *the annuli flare condition or the weak annuli flare condition.*

*Let $\mathcal{T}_0$ be a subtree of $\mathcal{T}$ and $X_0 = \Pi_{\mathcal{T}}^{-1}(\mathcal{T}_0)$. Then $X_0$ is hyperbolic.*

**Proof**  We note that any hallway in $X_0$ is also a hallway in $X$. In particular, $\Pi_{\mathcal{T}} \colon X_0 \to \mathcal{T}_0$ is a tree of spaces satisfying

(1)  the quasi-isometrically embedded condition (see Definition 3.15),

(2)  the annuli flare condition or the weak annuli flare condition.

The corollary is now a direct consequence of Theorem 3.17.  □

**3.5.2  Hyperbolicity of partially electrified bundle: proof of Theorem 3.14**  To prove Theorem 3.14, it suffices to check the two conditions of Theorem 3.17.

**Hyperbolicity of vertex spaces and the quasi-isometrically embedded condition**  This follows from Lemma 3.12.

**Identifying $\rho$-thin annuli**  It remains to prove the annuli flare condition. We recall the description of essential annuli (see Definition 3.7) in $F_S$ from Section 3.3. Let $D$ be as in Lemma 3.11. We choose $m$ in the annuli flaring condition so that $2m > D$. Hence, any essential annulus (see Definition 3.7) $\mathcal{A}_S$ in $F_S$ can intersect (see Definition 3.10) *at most one drilled atom* (see Definition 3.2). Thus, any essential annulus (see Definition 3.7) $\mathcal{A}_S$ in $F_S$ of length $2m$ can be of exactly one of the following three mutually exclusive types:

**Case 1**  $\mathcal{A}_S$ is a nonbacktracking annulus (see Definition 3.8) with core curve having free homotopy type distinct from any of the drilled curves,

**Case 2**  $\mathcal{A}_S$ contains an atomic wrapping nonbacktracking annulus $B_S$. Here, $B_S$ wraps around $\partial N_\epsilon(\sigma)$ for some $\sigma$. The core curve of $\mathcal{A}_S$ is then freely homotopic to a (nontrivial power of) $\sigma$.

**Case 3**  $\mathcal{A}_S$ contains an atomic wrapping backtracking annulus $B_S$ wrapping around $\partial N_\epsilon(\sigma)$ for some $\sigma$. The core curve of $\mathcal{A}_S$ is then freely homotopic to a (nontrivial power of) $\sigma$.

In cases 2 and 3 above, $\mathcal{A}_S$ is the concatenation of three pieces:

(1)  The first is the atomic wrapping annulus $B_S$ (with or without backtracking); see Definition 3.5.

(2)  The other two are nonbacktracking annuli (see Definition 3.8) $\mathcal{A}_S^{\pm}$, such that $\mathcal{A}_S^{\pm} \cap B_S$ consist of curves $\theta^{\pm}$ that are freely homotopic in $F_S$. If $B_S$ is an atomic wrapping nonbacktracking annulus, then there exist distinct singular fibers (see Definition 1.2) $S_x^{\pm}$ of $F_S$ (bounding the atom of $F_S$ containing $B_S$) such that $\theta^{\pm} \subset S_x^{\pm}$. If $B_S$ is an atomic wrapping backtracking annulus (see Definition 3.5), then there exists a single singular fiber (see Definition 1.2) $S_x$ of $F_S$ (a boundary component of the atom of $F_S$ containing $B_S$) such that $\theta^{\pm} \subset S_x$, and $\theta^{\pm}$ are freely homotopic within $S_x$.

Since $B_S$ wraps around $\partial N_\epsilon(\eta)$ for some drilled curve $\eta$, and since $\partial N_\epsilon(\eta)$ is an elevation of one of finitely many tori in $F$, the core curve of $\mathcal{A}_S$ is the same as the core curve of $B_S$, and hence is of the form $\gamma^n$ for some $n \in \mathbb{N}$, where $\gamma$ is one of the drilled curves in $E$.

**Checking the annuli flare condition   Case 1** ($\mathcal{A}_S$ is a nonbacktracking annulus (see Definition 3.8) with core curve having free homotopy type distinct from any of the drilled curves)   We start with the following converse to the Bestvina–Feighn combination theorem.

**Proposition 3.22** $E_S$ *satisfies the weak annuli-flare condition.*

**Proof**   Since $E$ is hyperbolic, this is a special case of [28, Proposition 5.8], where it is shown that $E$ satisfies a flaring condition. This is equivalent to hyperbolicity of hallways in $E_S$ and implies the weak annuli-flare condition. □

**Corollary 3.23** *Nonbacktracking annuli* (see Definition 3.8) *with core curve having free homotopy type distinct from any of the core curves satisfy the annuli-flare condition*: *more precisely, there exist* $\lambda > 1$ *and* $m > 1$ *such that if* $\mathcal{A}_S$ *is a nonbacktracking annulus* (see Definition 3.8) *with girth at least 1, it satisfies the weak annuli flare condition.*

**Proof**   Note that $\mathcal{A}_S \subset F_S \subset E_S$. Since $\mathcal{A}_S$ is a nonbacktracking annulus (see Definition 3.8) in $F_S$, it is an essential annulus (see Definition 3.7) in $E_S$. Proposition 3.22 now gives the desired conclusion. □

**Cases 2 and 3** ($\mathcal{A}_S$ is a wrapping annulus with core curve freely homotopic to a power of one of the drilled curves)   We shall give a unified proof of these two cases. Let $B_S$ denote the atomic wrapping annulus (see Definition 3.5) contained in $\mathcal{A}_S$. Let $\sigma$ denote the drilled curve such that $B_S$ wraps around $\partial N_\epsilon(\sigma)$. Choose an orientation of $\sigma$ and $n \in \mathbb{N}$ such that the core curve of $B_S$ (and hence that of $\mathcal{A}_S$) is freely homotopic to $\sigma^n$. Let $M_e$ denote the drilled atom (see Definition 3.2) of $F_S$ containing $B_S$. Let $S_x^\pm$ denote the boundary components of $M_e$, and $\mathcal{A}_S^\pm$ denote the two components of $\mathcal{A}_S \setminus \text{Int}(B_S)$. Then each of the annuli $\mathcal{A}_S^\pm$ is an essential annulus (see Definition 3.7) in $E_S$ with one boundary curve in $S_x^+ \cup S_x^-$. (If $\mathcal{A}_S$ is backtracking, then both boundary curves lie in the same surface boundary component. If $\mathcal{A}_S$ is without backtracking, then the boundary curves lie in different surface boundary components.) Let $\theta^\pm$ denote the boundary curve of $\mathcal{A}_S^\pm$ on $S_x^+ \cup S_x^-$.

It will help to explicate the special case where $n = 1$, ie the core curve of $\mathcal{A}_S$ is freely homotopic to $\sigma$. Then each of $\mathcal{A}_S^\pm$ is a flaring annulus, with $l(\theta^\pm)$ uniformly close to the girth of $\mathcal{A}_S$. This follows from the fact that the length $l(\theta^\pm)$ is uniformly bounded. Hence, in the formulation of Definition 3.16,

$$l(\mathcal{H}(\{0\} \times I)) \leq \frac{1}{\lambda} l(\mathcal{H}(\{m\} \times I)),$$

for all $m$ large enough (ie we can ignore the negative direction of the hallway from Definition 3.16). Further, note that the $d_{\text{pel}}$-length of the annulus $B_S$ is uniformly bounded. This is because the meridian of $\partial N_\epsilon(\sigma)$ that $B_S$ wraps around has length zero. Hence the concatenation $\mathcal{A}_S^+ \cup B_S \cup \mathcal{A}_S^-$ satisfies the weak annuli flare condition.

For general $n \in \mathbb{N}$, any elevation of $\theta^{\pm}$ (freely homotopic to $\sigma^n$) to the universal cover $\widetilde{E}_S$ gives a uniform quasigeodesic. This follows from the fact that $\sigma$ is an elevation of one of the (finitely many) drilled curves. Hence, again, $l(\theta^{\pm})$ is uniformly close to the girth of $\mathcal{A}_S$. Again, $\mathcal{A}_S^{\pm}$ satisfies the one-sided flare condition

$$l(\mathcal{H}(\{0\} \times I)) \leq \frac{1}{\lambda} l(\mathcal{H}(\{m\} \times I))$$

in the formulation of Definition 3.16. The same argument from the previous paragraph now shows that the concatenation $\mathcal{A}_S^{+} \cup B_S \cup \mathcal{A}_S^{-}$ satisfies the weak annuli flare condition.

Thus, the sufficient conditions of Theorem 3.17 are satisfied by essential annuli $\mathcal{A}_S$ (see Definition 3.7) in $F_S$, equipped with partially electrified pseudometric $d_{\mathrm{pel}}$ (see Section 3.4). Theorem 3.14 follows. $\square$

**Remark 3.24** One place where the partial electrification metric $d_{\mathrm{pel}}$ is essential in the above proof is to conclude that the $d_{\mathrm{pel}}$-length of $B_S$ is uniformly bounded.

**3.5.3 Consequences of Theorem 3.14** The proof of Theorem 3.14 above gives some additional information that we shall need below. Let $\widetilde{\Pi} \colon \widetilde{F}_S \to \mathcal{T}$ be the tree of spaces for the universal cover $\widetilde{F}_S$ $(=\widetilde{F})$ of $F_S$. Let $\mathcal{T}_0 \subset \mathcal{T}$ denote a subtree. Let $\widetilde{\Pi}^{-1}(\mathcal{T}_0) = X_0$, so that $\widetilde{\Pi} \colon X_0 \to \mathcal{T}_0$ is a tree of spaces. The following definition collects together some notions that will be used in Corollary 3.26 below.

**Definition 3.25** Let $v \in \mathcal{T}_0$ be a vertex such that

(1) $\widetilde{M}_v$ is an atom,

(2) $\widetilde{S}_x$ is a boundary component of $\widetilde{M}_v$ that is not contained in the boundary of any other vertex space of $\widetilde{\Pi} \colon X_0 \to \mathcal{T}_0$.

Then we say that $\widetilde{M}_v$ is a *boundary atom of $X_0$* and $\widetilde{S}_x$ is a *boundary component of $X_0$*. If $\widetilde{M}_v$ is drilled (resp. undrilled), we say that $\widetilde{M}_v$ is a *drilled* (*resp. undrilled*) *boundary atom of $X_0$* and $\widetilde{S}_x$ is a *drilled* (*resp. undrilled*) *boundary component of $X_0$*.

**Corollary 3.26** *Let $X_0$ be as above. Then $X_0$ is hyperbolic, and any drilled boundary component $\widetilde{S}_x$ of $X_0$ is $d_{\mathrm{pel}}$-quasiconvex. Here $d_{\mathrm{pel}}$ denotes the restriction of the partially electrified pseudometric $d_{\mathrm{pel}}$ (see Section 3.4) from $X$.*

**Proof** Hyperbolicity of $X_0$ follows from Lemma 3.21 after noting that we have checked the weak annuli flare condition for $F_S$. Next, we attach an auxiliary vertex space $\widetilde{S}_x \times I$ to $X_0$ along $\widetilde{S}_x$ and an auxiliary edge $e$ to $\mathcal{T}_0$ to $v$ to get a tree of spaces $X_0 \cup_{\widetilde{S}_x} \widetilde{S}_x \times I \to \mathcal{T}_0 \cup_v e$. Here, $\widetilde{S}_x \times \{1\} \subset \widetilde{S}_x \times I$ is attached to $X_0$ along $\widetilde{S}_x \subset X_0$, so that $\widetilde{S}_x \subset X_0$ becomes the edge-space $X_e$ corresponding to the new edge $e$. Further, let $v_0$ denote the extra vertex introduced in $\mathcal{T}_0 \cup_v e$ (corresponding to $\{0\} \in I$). Thus, the new vertex space $\widetilde{S}_x \times I$ is $X_{v_0}$. To prove $d_{\mathrm{pel}}$-quasiconvexity of $\widetilde{S}_x$, it suffices, by Corollary 3.19, to show that essential annuli starting on $X_{v_0}$ (in the sense of Definition 3.7) flare in all directions in the sense of

Definition 3.18. But such annuli are simply essential annuli (see Definition 3.7) in the bundle $E \to \mathcal{G}$ (before drilling). Further, their core curves lie in $S_x \setminus \sigma_i$ for one of the finitely many drilled curves $\sigma_i$. The conclusion now follows from Theorem 2.1, which guarantees that each of these finitely many proper essential subsurfaces of $S_x$ have uniformly quasiconvex elevations, and hence that any essential annulus (see Definition 3.7) starting on such a subsurface flares in all directions. $\qquad\square$

Another consequence of Corollary 3.19 that we record is the following.

**Corollary 3.27** *Let $\mathcal{T}_1 \subset \mathcal{T}$ denote a subtree of finite diameter (but not necessarily locally finite). Let $\widetilde{\Pi}^{-1}(\mathcal{T}_1) = X_1$, so that $\widetilde{\Pi}\colon X_1 \to \mathcal{T}_1$ is a tree of spaces. Then $X_1$ is hyperbolic, and for any singular fiber $S_x$ (see Definition 1.2) with $\widetilde{S}_x \subset X_1$, $\widetilde{S}_x$ is $d_{\mathrm{pel}}$-quasiconvex in $X_1$.*

**Proof** Since any essential hallway has length bounded by the diameter of $\mathcal{T}_1$, hyperbolicity of $X_1$ follows from Theorem 3.17. The same reason guarantees $d_{\mathrm{pel}}$-quasiconvexity of $\widetilde{S}_x$ by Corollary 3.19. $\qquad\square$

**Lemma 3.28** *Let $X_1$, $S_x$, and $\widetilde{S}_x$ be as above. Then $\widetilde{S}_x$ with its intrinsic metric is qi-embedded in $X_1$. Also, $\widetilde{S}_x$ is properly embedded in $(\widetilde{F}, d_{\mathrm{pel}})$.*

**Proof** Suppose that $S_x$ is a boundary component of an atom $M_e$ (drilled or undrilled) of $F_S$ (in the sense of Definition 3.25). By Lemma 3.12, $\widetilde{S}_x$ is qi-embedded in $\widetilde{M}_e$ equipped with $d_{\mathrm{pel}}$ (the case where $M_e$ is undrilled is obvious). Let $\widetilde{M}_e$ denote the vertex space $X_v$ for $v$ a vertex of $\mathcal{T}$. Next, let $\mathcal{T}_1 \subset \mathcal{T}$ denote a subtree of finite diameter (but not necessarily locally finite) containing $v$. Then Corollary 3.27 shows that $\widetilde{S}_x$ is $d_{\mathrm{pel}}$-quasiconvex in $\widetilde{\Pi}^{-1}(\mathcal{T}_1) = X_1$.

A reprise of the proof of Lemma 3.12 now shows that $\widetilde{S}_x$, equipped with its intrinsic metric is qi-embedded in $X_1$. Since the finite diameter of $\mathcal{T}_1$ was arbitrary, the second conclusion follows. $\qquad\square$

**Corollary 3.29** *There exist $K \geq 1$ and $\epsilon > 0$ such that the following hold. Let $\gamma$ be a geodesic in $(\widetilde{F}, d_{\mathrm{pel}})$. Let $\widetilde{S}_x$ be the elevation of a singular fiber bounding a drilled atom (see Definition 3.2) $\widetilde{M}_r$ of $\widetilde{F}$. Suppose further that $u$ and $v$ are entry and exit points on $\widetilde{S}_x$ of $\gamma$ into and out of $\widetilde{M}_r$, respectively. Let $\gamma\,|\,[u, v]$ denote the subpath of $\gamma$ between $u$ and $v$, and $\beta(u, v)$ the geodesic in $\widetilde{S}_x$ (with its intrinsic metric) between $u$ and $v$. Let*

$$\gamma'(u, v) = (\gamma \setminus \gamma\,|\,[u, v]) \cup \beta(u, v)$$

*be obtained from $\gamma$ by replacing $\gamma\,|\,[u, v]$ by $\beta(u, v)$. Then $\gamma'(u, v)$ is a $(K, \epsilon)$-quasigeodesic in $(\widetilde{F}, d_{\mathrm{pel}})$.*

**Proof** Let $X_1$ denote the closure of the component of $X \setminus \widetilde{S}_x$ containing $\widetilde{M}_r$. By Corollary 3.26, $\widetilde{S}_x$ is quasiconvex in $(X_1, d_{\mathrm{pel}})$ with uniform constants (independent of $\widetilde{S}_x$ and $X_1$). Hence, there exists $D$ such that $\gamma\,|\,[u, v]$ lies in a $D$-neighborhood of $\widetilde{S}_x$ in $(X_1, d_{\mathrm{pel}})$. Hence, by Lemma 3.28, $\beta(u, v)$ is a $(K_1, \epsilon_1)$-quasigeodesic in $(X_1, d_{\mathrm{pel}})$. The corollary follows. $\qquad\square$

Corollary 3.29 allows us to replace $d_{\mathrm{pel}}$-geodesics by uniform $d_{\mathrm{pel}}$-quasigeodesics that do not backtrack from drilled atoms (see Definition 3.2) in the following sense.

**Definition 3.30** A $d_{\mathrm{pel}}$-quasigeodesic $\gamma'$ in $(\widetilde{F}, d_{\mathrm{pel}})$ is said to be a $d_{\mathrm{pel}}$-quasigeodesic *without backtracking from drilled atoms* (see Definition 3.2) if it satisfies the following: if $\gamma'$ enters a drilled atom $\widetilde{M}_r$ (see Definition 3.2) through a boundary component $\widetilde{S}_1$, then it can only leave $\widetilde{M}_r$ through a boundary component $\widetilde{S}_2 \neq \widetilde{S}_1$.

Corollary 3.29 now allows us to observe that for any $d_{\mathrm{pel}}$-geodesic $\gamma$, there exists a $(K, \epsilon)$-quasigeodesic $\gamma'$ without backtracking from drilled atoms (see Definition 3.2) joining the endpoints of $\gamma$. In fact, $\gamma'$ is obtained from $\gamma$ by

(1) carrying out replacements of all backtracking segments in drilled atoms (see Definition 3.2),

(2) isotoping the resulting path slightly to make it disjoint from $\widetilde{S}_x$ as in Corollary 3.29.

The paragraph titled "Checking the annuli flare condition" in the proof of Theorem 3.14 gives the following further conclusion.

**Corollary 3.31** *There exists $L \geq 1$ such that the following holds: Let $[a, b] \subset \widetilde{\mathcal{G}}$ denote a geodesic of length at least $L$ such that $M_a$ and $M_b$ are drilled atoms (see Definition 3.2) of $F_S$. Let $M_{[a,b]}$ denote the 3-manifold given by $\Pi_S^{-1}([a, b])$. Also, let $S_a$ and $S_b$ denote the boundary components of $M_{[a,b]}$ (see Definition 3.25). Let $\pi_1(S_a), \pi_1(S_b)$ denote the subgroups of $\pi_1(M_{[a,b]})$ carried by $S_a$ and $S_b$. Then $\pi_1(S_a) \cap \pi_1(S_b) = \{1\}$.*

**Proof** Since $M_a$ and $M_b$ are drilled atoms (see Definition 3.2), $\widetilde{S}_a$ and $\widetilde{S}_b$ are $d_{\mathrm{pel}}$-quasiconvex in $\widetilde{M}_{[a,b]}$, equipped with the $d_{\mathrm{pel}}$-metric by Corollary 3.26. Further, the $d_{\mathrm{pel}}$-quasiconvexity constant is uniform, independent of $a$ and $b$.

If $\pi_1(S_a) \cap \pi_1(S_b) \neq \{1\}$, then there exists a loop $\alpha_a \subset S_a$ freely homotopic to an $\alpha_b \subset S_b$. By uniform $d_{\mathrm{pel}}$-quasiconvexity, any geodesic in the free homotopy class of $\alpha_a$ (resp. $\alpha_b$) must lie close to $S_a$ (resp. $S_b$). This forces the existence of an $L$ such that $d_{\widetilde{\mathcal{G}}}(a, b) < L$. $\qquad\square$

# 4 Guessing geodesics

We refer the reader back to the paragraph titled "Step 2" in Section 3.1 for a summary of the contents of the various subsections of this section and the logical order we follow.

Notation to be used in this section:

• Given a hyperbolic drilled atom $M_r$ (see Definition 3.2), $\lambda$, $\hat{\lambda}$, $\lambda_{ea}$, and $\lambda_p$ denote in $\widetilde{M}_r$ respectively the geodesic, the electric geodesic, the electroambient quasigeodesic (see Definition 4.3), and the $d_{\mathrm{pel}}$-geodesic, all with the same endpoints.

• We use $P$ to denote a lift of a boundary torus of $F$ to $\widetilde{F}$, and $\mathcal{P}$ is used to denote the collection of all such lifts. $\mathcal{P}_r$ denotes the collection of such lifts contained in $\widetilde{M}_r$.

• For $z$ lying on a drilled curve $\sigma$ in $M_r$, $m_z$ is a specially chosen meridian of $\partial N_\epsilon(\sigma)$.

- For two lifts of boundary surfaces $\widetilde{S}_1$ and $\widetilde{S}_2$ in $\widetilde{M}_r$, $Z_1$ and $Z_2$ denote respectively the image of the projection of $\widetilde{S}_2$ onto $\widetilde{S}_1$ and vice versa.

- For $[a, b]$ a geodesic in the tree $\mathcal{T}$ for the tree of spaces $\widetilde{F}_S \to \mathcal{T}$, $X_{[a,b]}$ is the subtree of spaces over $[a, b]$.

- $\mathcal{P}_{[a,b]}$ is the collection of lifts of $\{\partial N_\epsilon(\sigma_i)\}$ in $X_{[a,b]}$.

- $\mathcal{B}$ denotes an elevation of a maximal undrilled subbundle (see Definition 3.2) of $F_S$.

- $\mathcal{F}$ is the path family defined in Section 4.2.7 to which the guessing geodesics lemma (Theorem 4.2) is applied. Paths in this family are denoted by $\eta$.

- $A$ is used to denote undrilled atoms (see Definition 3.2), and $B$ is used for *undrilled molecules*.

## 4.1 The guessing geodesics lemma

We shall need a necessary and sufficient condition for relative hyperbolicity due to Sisto [30] building on earlier work of Bowditch [6] and Hamenstädt [16].

**Definition 4.1** Let $(X, d_X)$ be a geodesic metric space, and $\mathcal{P}$ a collection of subsets. The collection $\mathcal{P}$ is said to be *mutually bounded* if for each $K \geq 0$, there exists $B$ such that $\mathrm{diam}(N_K(P) \cap N_K(Q)) \leq B$, for all $P \neq Q \in \mathcal{P}$.

In [30], Sisto refers to the mutual boundedness criterion above as condition $(\alpha_1)$. We shall need the following:

**Theorem 4.2** [30, Theorem 4.2] *Let $(X, d_X)$ and $\mathcal{P}$ be as above. Suppose that for all $x, y \in X$ we are given*

(1) *a path $\eta(x, y)$ connecting them,*

(2) *a closed subset $\theta(x, y) \subset \eta(x, y)$.*

*Suppose that there exists $D \geq 0$ such that the following are satisfied:*

(1) *If $d_X(x, y) \leq 2$ then $\mathrm{diam}(\theta(x, y)) \leq D$.*

(2) *Let $d_H$ denote Hausdorff distance. Then for all $x', y' \in \eta(x, y)$, we have*

$$d_H(\theta(x', y'), \theta(x, y)|[x', y'] \cup \{x', y'\}) \leq D,$$

*where $\theta(x, y)|[x', y'] = \theta(x, y) \cap \eta(x, y)|[x', y']$.*

(3) *For all $x, y, z \in X$, $\theta(x, y) \subset N_D(\theta(x, z) \cup \theta(z, y))$.*

(4) *If $x', y' \in \eta(x, y)$ do not both lie on the same $P \in \mathcal{P}$, then there exists $z \in \theta(x, y)$ between $x'$ and $y'$.*

(5) *The elements of $\mathcal{P}$ are mutually bounded.*

(6) *For all $k \geq 0$, there exists $K \geq 0$ such that the following holds. If for some $P \in \mathcal{P}$*

- $d_X(x, P) \leq k$,
- $d_X(y, P) \leq k$, *and*
- $d_X(x, y) \geq K$,

*then*

- $\theta(x, y) \subset B_K(x) \cup B_K(y)$, *and*
- *there exists $z \in \theta(x, y) \cap N_D(P)$.*

*Then $X$ is strongly hyperbolic relative to $\mathcal{P}$.*

Theorem 4.2 says roughly that if we can guess a family of paths in $X$ that satisfy the conditions required of geodesics in a relatively hyperbolic space, then $X$ itself is relatively hyperbolic.

## 4.2 Path families

Recall that $\mathcal{P}$ denotes the collection of elevations of $\partial N_\epsilon(\sigma_i)$ in $\widetilde{F}$, where $\sigma_i$ ranges over the finitely many drilled curves in $E$. Recall that $\widetilde{F}$ admits three natural pseudometrics in our setup:

(1) the path-metric $d$ lifted from $F$ (recall that the metric $d$ on $F$ is the natural path-metric induced from $E$),

(2) the electric metric $d_e$ obtained from $d$ by electrifying the collection $\mathcal{P}$ [10],

(3) the partially electrified pseudometric $d_{\text{pel}}$ constructed in Section 3.4.

We now recall a construction from [25, Definition 1.13].

**Definition 4.3** Let $\hat{\lambda}$ denote an electric geodesic in $(\widetilde{F}, d_e)$ joining $a, b \in \widetilde{F}$. Modify $\hat{\lambda}$ to a path $\lambda_{ea}$ as follows. First, $\lambda_{ea}$ coincides with $\hat{\lambda}$ away from the elevations of $\partial N_\epsilon(\sigma_i)$ to $\widetilde{F}$. Next, for any $\widetilde{\partial N_\epsilon(\sigma_i)}$ that $\hat{\lambda}$ intersects, let $x_i$ and $y_i$ denote the entry and exit points. Join $x_i$ and $y_i$ by a geodesic in $\widetilde{\partial N_\epsilon(\sigma_i)}$ equipped with its flat Euclidean metric. The resulting path $\lambda_{ea}$ will be called an *electroambient quasigeodesic* in $\widetilde{F}$.

We recall the following consequence of [25, Lemma 1.21; 10, Lemma 4.5 and Proposition 4.6] for easy reference.

**Lemma 4.4** Let $\widetilde{M_r}$ denote a drilled atom of $\widetilde{F}$ (see Definition 3.2). Let $a, b \in \widetilde{M_r}$. Let $\mathcal{P}_r$ denote the collection of elevations of $\widetilde{\partial N_\epsilon(\sigma_i)}$ to $\widetilde{M_r}$. Then $\widetilde{M_r}$ is strongly hyperbolic relative to the collection $\mathcal{P}_r$.

Let $d$, $d_e$, and $d_{\text{pel}}$ denote the metric, electric (pseudo)metric, and the partially electrified (pseudo)metric (see Section 3.4) respectively on $\widetilde{M_r}$. Let $\lambda$, $\hat{\lambda}$, $\lambda_{ea}$, and $\lambda_p$ denote respectively the geodesic, the electric geodesic, the electroambient quasigeodesic (see Definition 4.3), and the geodesic with respect to the $d_{\text{pel}}$ on $\widetilde{M_r}$ joining $a$ and $b$. Then $\lambda$, $\hat{\lambda}$, $\lambda_{ea}$, and $\lambda_p$ track each other away from $\mathcal{P}_r$.

**Proof** It follows from [10] that $\widetilde{M}_r$ is strongly hyperbolic relative to the collection $\mathcal{P}_r$, since $M_r$ admits the structure of a complete hyperbolic manifold with convex boundary.

Lemma 4.5 and Proposition 4.6 of [10] guarantee that $\lambda$ and $\hat{\lambda}$ track each other away from $\mathcal{P}_r$. The construction of $\lambda_{ea}$ guarantees that $\lambda_{ea}$ and $\hat{\lambda}$ agree exactly, away from $\mathcal{P}_r$. Finally, [25, Lemma 1.21] guarantees that $\lambda_p$ and $\hat{\lambda}$ track each other away from $\mathcal{P}_r$. $\qquad\square$

**4.2.1 Connectors: a scheme** The rest of this subsection is devoted towards constructing a family of paths $\mathcal{F}$. In Section 4.2.2 we construct paths that approximate the shortest paths connecting two different elevations of boundary components of a drilled atom (see Definition 3.2). The paths thus constructed are called *connectors*.

In Section 4.2.3, we extend this family to one for *every* pair of points in elevations of boundary components of a drilled atom (see Definition 3.2). The paths thus constructed are called *extended connectors*.

In Section 4.2.4, we are interested in a concatenation of atoms along a geodesic segment in the Bass–Serre tree. We describe how to concatenate extended connectors in atoms to such concatenations of atoms.

Apart from drilled atoms (see Definition 3.2), one needs to consider elevations $\mathcal{B}$ of maximal undrilled subbundle (see Definition 3.2). Let $\widetilde{S}$ be a boundary component of the elevation $\widetilde{M}_r$ of a drilled atom (see Definition 3.2). Further assume that $\widetilde{S}$ is contained in an elevation $\mathcal{B}$ of a maximal undrilled subbundle (see Definition 3.25). Then we need to consider geodesics from $\widetilde{S}$ to itself lying within $\mathcal{B}$. These are geodesics in the intrinsic hyperbolic metric on $\mathcal{B}$ starting and ending on $\widetilde{S}$. Section 4.2.5 is devoted to such paths.

Section 4.2.6 generalizes the discussion of Section 4.2.5 to the case where there are distinct drilled atoms (see Definition 3.2) abutting $\mathcal{B}$ in distinct boundary components (see Definition 3.25) $\widetilde{S}$ and $\widetilde{S}'$.

With the above constituent connectors in place, we finally define the path family in Section 4.2.7.

**4.2.2 Connectors in drilled atoms** Recall that for any drilled atom $M_r$ (see Definition 3.2) of $F_S$, $\widetilde{M}_r$ is strongly hyperbolic relative to $\mathcal{P}_r$ by Lemma 3.12 (see Lemma 3.12 for notation). We assume that $M_r$ has been equipped with a complete hyperbolic metric with convex boundary (as in the proof of the first conclusion of Lemma 3.12). Let $\widetilde{S}_1$ and $\widetilde{S}_2$ denote two boundary components of $\widetilde{M}_r$. We are interested in the nearest-point projections of $\widetilde{S}_1$ and $\widetilde{S}_2$ on each other. The nearest-point projection of $\widetilde{S}_1$ (resp. $\widetilde{S}_2$) onto $\widetilde{S}_2$ (resp. $\widetilde{S}_1$) will be denoted by $\pi_{12}$ (resp. $\pi_{21}$). Three possible cases arise:

**Case 1** (product region connectors) There exists a maximal essential proper subsurface $\Sigma$ of one of the surface boundary components of $M_r$ such that $\Sigma \times [0,1]$, equipped with the standard product structure, embeds in $M_r$, with $\Sigma \times \{0,1\} \subset \partial M_r$. Further, there exists an elevation $\widetilde{\Sigma} \times [0,1] \subset \widetilde{M}_r$, such that $\widetilde{\Sigma} \times \{0\} \subset \widetilde{S}_1$, and $\widetilde{\Sigma} \times \{1\} \subset \widetilde{S}_2$. In this case, $\pi_{12}(\widetilde{S}_1)$ lies in a uniformly bounded neighborhood of $\widetilde{\Sigma} \times \{1\} \subset \widetilde{S}_2$, and $\pi_{21}(\widetilde{S}_2)$ lies in a uniformly bounded neighborhood of $\widetilde{\Sigma} \times \{0\} \subset \widetilde{S}_1$. For any $z \in \Sigma$, we refer to an elevation of $\{z\} \times [0,1]$ to $\widetilde{M}_r$ as a *product region connector* between $\widetilde{S}_1$ and $\widetilde{S}_2$. Note that in this case, $\widetilde{S}_1$ and $\widetilde{S}_2$ are necessarily elevations of distinct boundary components of $M_r$.
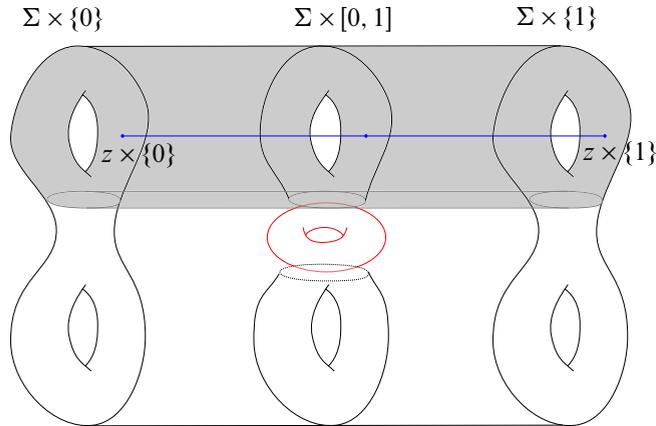
Figure 5: Product region connector in $M_r$.

**Definition 4.5** If $\widetilde{S}_1$ and $\widetilde{S}_2$ are connected by a product region connector, we say that $\widetilde{S}_1$ and $\widetilde{S}_2$ are a *product pair* of elevations.

We emphasize that for $\widetilde{S}_1$ and $\widetilde{S}_2$ a product pair, the nearest-point projection of one onto the other lies in a uniformly bounded neighborhood of $\widetilde{\Sigma} \times \{0\}$ for a maximal essential proper subsurface $\Sigma$.

**Case 2** (annular connectors) Recall that $M_r$ is obtained from $S \times I$ after drilling finitely many curves $\{\sigma_i\}$. For any $z \in \sigma_i \times \{0, 1\}$, for some $i$, suppose that $z_I = z \times I \setminus \mathrm{Int}(N_\epsilon(\sigma_i)) \subset S \times I$ is contained in $M_r$, ie $z_I$ does not intersect any $N_\epsilon(\sigma_j)$ for $j \neq i$. Let $z_I^{\pm}$ denote the two components of $z_I$, and let $m_z$ denote the meridian of $N_\epsilon(\sigma_i)$ passing through $z_I^{\pm} \cap \partial N_\epsilon(\sigma_i)$. Then an *annular connector* $\eta$ in $M_r$ starts and ends at $\{z\} \times \{0, 1\}$, traverses a connected component of $z_I$, wraps around $m_z$ finitely many times, and finally traverses a (possibly same) connected component of $z_I$. If $\eta$ starts and ends at the same point $\{z\} \times \{0\}$ (or $\{z\} \times \{1\}$), then it wraps around $m_z$ (the meridian) $k$ times for some $k \in \mathbb{Z}$; otherwise if it starts and ends at distinct points in $\{z\} \times \{0, 1\}$, it wraps around $m_z$ (the meridian) $k + \frac{1}{2}$ times for some $k \in \mathbb{Z}$.
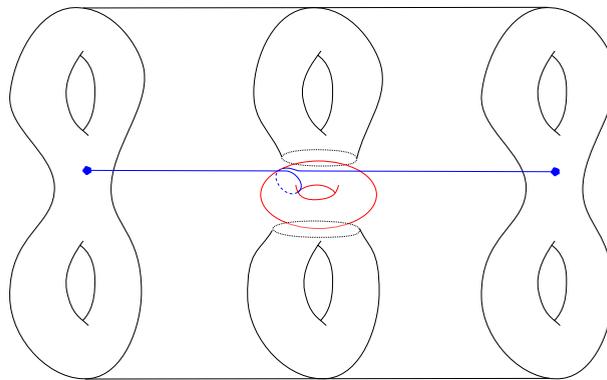


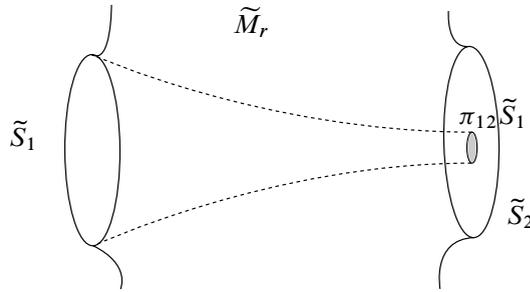Figure 6: An annular connector in $M_r$.

Figure 7: A mutually bounded pair.

Any elevation of an annular connector in $M_r$ is called an *annular connector in $\widetilde{M}_r$*. If the number of times an annular connector wraps around $m_z$ does not equal $\pm\frac{1}{2}$, we call it and its elevations *strict annular connectors*. Let $\widetilde{S}_1$ and $\widetilde{S}_2$ denote the elevations of the boundary components of $M_r$ passing through the endpoints of an annular connector $\tilde{\eta}$. We observe the following:

**Lemma 4.6** *Suppose that $\tilde{\eta}$ is a strict annular connector between $\widetilde{S}_1$ and $\widetilde{S}_2$ joining $\tilde{z}_1 \in \widetilde{S}_1$ to $\tilde{z}_2 \in \widetilde{S}_2$. Let $\tilde{\sigma}$ and $\tilde{\sigma}'$ denote the elevations of $\sigma_i$ through $\tilde{z}_1$ and $\tilde{z}_2$, respectively. Then $\pi_{12}(\widetilde{S}_1)$ (resp. $\pi_{21}(\widetilde{S}_2)$) lies in a uniformly bounded neighborhood of $\tilde{\sigma}'$ (resp. $\tilde{\sigma}$).*

**Proof** Let $z_1$ (the image of $\tilde{z}_1$ under the covering projection) be the basepoint for $\pi_1(M_r)$, and let $z_\sigma$ be an annular connector that wraps around $m_z$ only half a time. Identify $\pi_1(S_2)$ with loops based at $z_1$ preceded and succeeded by $z_\sigma$ with appropriate orientations.

Then the result follows from the fact that $\pi_1(S_l) \cap \pi_1(S_k)^g = \mathbb{Z}$ for $g \in \pi_1(M_r)$ representing any strict annular connector starting and ending at $z_1$, and $1 \leq l, k \leq 2$. $\qquad\square$

**Definition 4.7** If $\widetilde{S}_1$ and $\widetilde{S}_2$ are connected by a strict annular connector, we say that $\widetilde{S}_1$ and $\widetilde{S}_2$ are an *annular pair* of elevations.

**Case 3** (the mutually bounded case)

**Lemma 4.8** *If $\widetilde{S}_1$ and $\widetilde{S}_2$ is a pair of distinct elevations that are neither a product pair (see Definition 4.5), nor an annular pair (see Definition 4.7), then $\pi_{12}(\widetilde{S}_1)$ is uniformly bounded in diameter, ie $\widetilde{S}_1$ and $\widetilde{S}_2$ are mutually bounded in the sense of Definition 4.1.*

**Proof** Let $H_i < \pi_1(M_r)$ denote the stabilizers of $\widetilde{S}_i$ for $i = 1, 2$. The Lemma then follows from the fact that $H_1 \cap H_2 = \{1\}$ (see, for instance, the proof of [31, Corollary 3]) (see). $\qquad\square$

Figure 7 illustrates this case.

**4.2.3 Extended connectors in drilled atoms** We now describe a preferred family of quasigeodesics in $\widetilde{M}_r$ connecting $x \in \widetilde{S}_1$ to $y \in \widetilde{S}_2$. Denote $\pi_{12}\widetilde{S}_1$ by $Z_2$ and $\pi_{21}\widetilde{S}_2$ by $Z_1$. Then there are the following preferred product regions in $\widetilde{M}_r$:

(1)  If $\widetilde{S}_1$ and $\widetilde{S}_2$ are a boundary components (see Definition 3.25), then there exists a proper essential subsurface $\Sigma$ of $S$ and elevations $\widetilde{\Sigma}_i$ for $i = 1, 2$ of boundary components of the product region $\Sigma \times [0, 1]$ such that $Z_1$ (resp. $Z_2$) is coarsely $\widetilde{\Sigma}_1$ (resp. $\widetilde{\Sigma}_2$). Also $Z_i$ are (coarsely) the boundary components of $\widetilde{\Sigma} \times [0, 1]$, where the $[0, 1]$ direction has uniformly bounded length. We normalize the length of $z \times [0, 1]$ to one for all $z \in \widetilde{\Sigma}$.

(2)  If $\widetilde{S}_1, \widetilde{S}_2$ is an annular pair, then $Z_1$ and $Z_2$ are coarsely elevations of (curves parallel to) $\sigma_i$. Further, $Z_1$ and $Z_2$ are boundary components of a flat strip $\widetilde{\sigma}_i \times [0, kl(m_z) + a]$, where $k$ equals the number of times the annular connector wraps around the meridian $m_z$, $l(m_z)$ denotes the length of the meridian $m_z$, and $a$ equals the sum of the lengths of $z_I^{\pm}$, assuming without loss of generality that they are equal.

(3)  Otherwise, if $\widetilde{S}_1, \widetilde{S}_2$ is a mutually bounded pair (as in Lemma 4.8), then $Z_1$ and $Z_2$ are coarsely points, ie $Z_1$ and $Z_2$ have uniformly bounded diameter.

In all three cases, $\widetilde{S}_i$ is strongly hyperbolic relative to $Z_i$ for $i = 1, 2$. Further, there is a natural product $Z \times [a_1, a_2]$ embedded in $\widetilde{M}_r$ with $Z \times \{a_i\} = Z_i$ for $i = 1, 2$. The interval $[a_1, a_2]$ has the following properties:

(1)  If $\widetilde{S}_1$ and $\widetilde{S}_2$ are boundary components (see Definition 3.25), $[a_1, a_2]$ has length one.

(2)  If $\widetilde{S}_1, \widetilde{S}_2$ is an annular pair, $[a_1, a_2]$ has length equal to $kl(m_z) + a$.

(3)  If $\widetilde{S}_1, \widetilde{S}_2$ is a mutually bounded pair, $[a_1, a_2]$ has length equal to the length of an electroambient quasigeodesic (see Definition 4.3) in $\widetilde{M}_r$ joining $\widetilde{z}_1$ and $\widetilde{z}_2$.

Next, for any $x_i \in \widetilde{S}_i$ with $i = 1, 2$, let $y_i \in Z_i$ denote a nearest-point projection (in the intrinsic metric on $\widetilde{S}_i$) of $x_i$ onto $Z_i$. Identifying $Z_i$ for $i = 1, 2$ with $Z \times \{0\} \subset Z \times [0, 1]$ and $Z \times \{1\} \subset Z \times [0, 1]$, respectively, we have the following preferred family of paths joining $x_1$ and $x_2$ in $\widetilde{M}_r$:

Projecting both $y_1$ and $y_2$ to the $Z$-factor, we get points that we call $y_1$ and $y_2$ again to avoid cluttering notation. Let $[y_1, y_2]$ denote the geodesic in $Z$ joining $y_1$ and $y_2$. Let $p$ be any point on $[y_1, y_2]$. Then the preferred collection of paths joining $x_1, x_2 \in \widetilde{M}_r$ are given by the concatenation of the following segments:

(1)  the geodesic $[x_1, y_1] \subset \widetilde{S}_1$ joining $x_1$ and $y_1$,

(2)  the geodesic $[y_1, p \times \{0\}] \subset (Z \times \{0\})(= Z_1) \subset \widetilde{S}_1$ joining $y_1$ and $p \times \{0\}$ in $(Z \times \{0\})$,

(3)  the vertical interval $p \times [0, 1]$ traveling from $p \times \{0\}$ to $p \times \{1\}$,

(4)  the geodesic $[p \times \{1\}, y_2] \subset (Z \times \{1\})(= Z_2) \subset \widetilde{S}_2$ joining $p \times \{1\}$ and $y_2$ in $(Z \times \{1\})$,

(5)  the geodesic $[y_2, x_2] \subset \widetilde{S}_2$ joining $y_2$ and $x_2$.

Note that there is only one vertical interval $p \times [0, 1]$ traveling from $p \times \{0\}$ to $p \times \{1\}$ in each member of the family given above. Let $\mathcal{F}(M_r, x_1, x_2)$ denote the above family.

The construction above shows the following:

**Lemma 4.9** *Each $\alpha \in \mathcal{F}(M_r, x_1, x_2)$ tracks the $d_{\text{pel}}$-geodesic and the electroambient quasigeodesic (see Definition 4.3) between $x_1$ and $x_2$ in the intrinsic metric on $\widetilde{M}_r$.*

**Proof** The fact that $\alpha$ tracks the electroambient quasigeodesic (see Definition 4.3) *along elements of $\mathcal{P}_r$* follows from the fact that the nearest-point projection of any $P \in \mathcal{P}_r$ onto $\widetilde{S}_1$ is either uniformly bounded or (coarsely) an elevation of $\sigma_i$ to $\widetilde{S}_1$. More precisely, in the second case, there exists an elevation $\widetilde{\sigma}_i \subset \widetilde{S}_1$ such that the nearest-point projection of $P \in \mathcal{P}_r$ onto $\widetilde{S}_1$ lies in a uniformly bounded neighborhood of $\widetilde{\sigma}_i$, and hence the concatenations $[x_1, y_1] \cup [y_1, p \times \{0\}]$ and $[p \times \{1\}, y_2] \cup [y_2, x_2]$ used to define $\alpha$ have a maximal subpath each parallel to $P$.

Away from elements of $\mathcal{P}_r$, this is a consequence of Lemmas 3.12 and 4.4. □

**4.2.4 Extended connectors in concatenated drilled atoms** Lemma 4.9 can be extended to a 3-manifold obtained by concatenating finitely many atoms. As before, let $\Pi \colon \widetilde{F} \to \mathcal{T}$ denote the Bass–Serre tree of $\widetilde{F} = \widetilde{F}_S$, with vertex spaces $X_v$ given by elevated atoms.

Corollary 3.31 can be strengthened slightly as follows:

**Corollary 4.10** *There exists $L \geq 1$ such that the following holds: Let $[a, b] \subset \mathcal{T}$ denote a geodesic of length at least $L$ such that $X_a$ and $X_b$ are elevated drilled atoms (see Definition 3.2) of $F_S$. Let $X_{[a,b]}$ denote the 3-manifold given by $\Pi^{-1}([a, b])$. Also, let $\widetilde{S}_a \subset X_a$ and $\widetilde{S}_b \subset X_b$ denote boundary components (see Definition 3.25) of $X_{[a,b]}$ in the sense of Definition 3.25. Then:*

(1) *$X_{[a,b]}$ is strongly hyperbolic relative to the collection $\mathcal{P}_{[a,b]}$ (the elements of $\mathcal{P}$ contained in it).*

(2) *There exists $D$ depending only on $d_{\mathcal{T}}(a, b)$ such that $\pi_{ab}(\widetilde{S}_a)$ has diameter bounded by $D$ (here, $\pi_{ab}$ denotes the nearest-point projection of $\widetilde{S}_a$ on $\widetilde{S}_b$).*

**Proof** The first conclusion has the same proof as the first conclusion of Lemma 3.12. The second now follows from Corollary 3.31. □

The construction of extended connectors in $X_{[a,b]}$ can now be carried out exactly as in Section 4.2.3. We denote the family thus constructed by $\mathcal{F}(X_{[a,b]}, x_1, x_2)$. Corollary 4.10 gives us the following immediate consequence.

**Lemma 4.11** *We continue with the setup of Corollary 4.10. There exists $L \geq 1$ such that if $d_{\mathcal{T}}(a, b) \geq L$, then preferred extended connectors $\alpha \in \mathcal{F}(X_{[a,b]}, x_1, x_2)$ are concatenations of three pieces, where the middle piece $\alpha_m$ is necessarily a cobounded connector, and the first and last ones are geodesics in $\widetilde{S}_a$ and $\widetilde{S}_b$.*

*Thus, the endpoints of $\alpha_m$ are coarsely well-defined, ie there exists $D$ depending only on $d_{\mathcal{T}}(a, b)$, and $z_1$ and $z_2$ in $\widetilde{S}_a$ and $\widetilde{S}_b$, respectively, such that the endpoints of $\alpha_m$ lie on $\widetilde{S}_a$ and $\widetilde{S}_b$ at a distance of at most $D$ from $z_1$ and $z_2$.*

Since $X_{[a,b]}$ is strongly hyperbolic relative to the collection $\mathcal{P}_{[a,b]}$, as are all drilled atoms $\widetilde{M}_c$ (see Definition 3.2) for $c \in [a,b]$, the restriction of $\gamma \in \mathcal{F}(X_{[a,b]}, x_1, x_2)$ to $\widetilde{M}_c$ may be perturbed by a uniformly bounded amount, so that $(\gamma \cap \widetilde{M}_c) \in \mathcal{F}(M_c, x_1, x_2)$ for some $x_1, x_2 \in \partial \widetilde{M}_c$.

**Remark 4.12** Lemma 4.11 implies that $\alpha_m$ is a coarsely well-defined electroambient quasigeodesic (see Definition 4.3) in $X_{[a,b]}$, and *any preferred connector* between $\widetilde{S}_a$ and $\widetilde{S}_b$ coarsely contains it. Note however that the parameter $D$ determining coarseness depends on $d_{\mathcal{T}}(a,b)$.

Let $[a,b] \subset \mathcal{T}$ denote a geodesic of length at least $L$ such that $X_a$ and $X_b$ are elevated drilled atoms of $F_S$ (see Definition 3.2). Let $X_{[a,b]}$ denote the 3-manifold given by $\Pi^{-1}([a,b])$. Assume further that the only drilled atoms in $X_{[a,b]}$ are $X_a$ and $X_b$. Let $\widetilde{S}_a$ and $\widetilde{S}_b$ denote the boundary components (see Definition 3.25) of $X_{[a,b]} \setminus \mathrm{Int}(X_a \cup X_b)$, so that $\widetilde{S}_a$ and $\widetilde{S}_b$ are also boundary components (see Definition 3.25) of $X_a$ and $X_b$, respectively. We refer to $\widetilde{S}_a$ and $\widetilde{S}_b$ as *internal boundary components* of $X_a$ and $X_b$, respectively. All other boundary components will be referred to as *external boundary components*.

**Lemma 4.13** *Let $X_{[a,b]}$, $X_a$, and $X_b$ be as above. Then there exists $L \geq 2$ such that if $d_{\mathcal{T}}(a,b) \geq L$, the following holds. For any external boundary components $\widetilde{S}_1$ and $\widetilde{S}_2$ of $X_{[a,b]}$, there exists a coarsely well-defined geodesic $\lambda$ joining internal boundary components $\widetilde{S}_a$ and $\widetilde{S}_b$ in $X_{[a,b]} \setminus \mathrm{Int}(X_a \cup X_b)$ such that any $\alpha \in \mathcal{F}$ connecting $\widetilde{S}_1$ and $\widetilde{S}_2$ coarsely contains $\lambda$. Further, the coarseness is uniform, independent of $X_{[a,b]}$.*

**Proof** Let $\pi_{1a}$ and $\pi_{2b}$ denote nearest-point projections of $\widetilde{S}_1$ onto $\widetilde{S}_a$, and $\widetilde{S}_2$ onto $\widetilde{S}_b$, respectively. Then the images of $\pi_{1a}$ and $\pi_{2b}$, given by $W_1$ and $W_2$, respectively, are given by an elevation each of a proper essential subsurfaces of $\widetilde{S}_a$ and $\widetilde{S}_b$, respectively. Hence, by Theorem 2.1, there exists $L \geq 2$ such that if $d_{\mathcal{T}}(a,b) \geq L$, there exists a coarsely well-defined shortest geodesic $\sigma$ in $X_{[a,b]} \setminus \mathrm{Int}(X_a \cup X_b)$ joining $W_1$ and $W_2$. Any $\alpha \in \mathcal{F}$ connecting $\widetilde{S}_1$ and $\widetilde{S}_2$ must coarsely join points in $W_1$ and $W_2$, and hence must coarsely contain $\sigma$. $\square$

**4.2.5 Controlling backtracking in elevated subbundles** Suppose that $\widetilde{M}_r$ is a drilled atom of $\widetilde{F}$ (see Definition 3.2). Let $\widetilde{S}_1$, $\widetilde{S}_2$, and $\widetilde{S}_3$ denote three distinct boundary components of $\widetilde{M}_r$, and let $\mathcal{B}$ be the elevation of a maximal undrilled subbundle of $F_S$ (Definition 3.2) such that $\mathcal{B} \cap \widetilde{M}_r = \widetilde{S}_3$. Let $\pi_{13}$ and $\pi_{23}$ be nearest-point projections of $\widetilde{S}_1$ and $\widetilde{S}_2$ onto $\widetilde{S}_3$. Then the images of $\pi_{13}$ and $\pi_{23}$ are either

(1) given by an elevation of a proper essential subsurface of $S_3$ (as in Section 4.2.2), or

(2) uniformly bounded in diameter.

Further, by Theorem 2.1, $\pi_{13}(\widetilde{S}_1)$ and $\pi_{23}(\widetilde{S}_2)$ are uniformly quasiconvex in $\mathcal{B}$, ie there exists $C \geq 1$ such that for any $\mathcal{B}$, $\widetilde{S}_1$, $\widetilde{S}_2$, and $\widetilde{S}_3$ as above, $\pi_{13}(\widetilde{S}_1)$ and $\pi_{23}(\widetilde{S}_2)$ are $C$-quasiconvex in $\mathcal{B}$. Hence, there exists $D_0$ such that if $d(\pi_{13}(\widetilde{S}_1), \pi_{23}(\widetilde{S}_2)) \geq D_0$, there is a coarsely unique shortest path $\alpha$ in $\mathcal{B}$ joining $\pi_{13}(\widetilde{S}_1)$ and $\pi_{23}(\widetilde{S}_2)$. Thus, due to quasiconvexity of $\pi_{13}(\widetilde{S}_1)$ and $\pi_{23}(\widetilde{S}_2)$, there exists $C_1 \geq 0$ such that exactly one of the following holds:
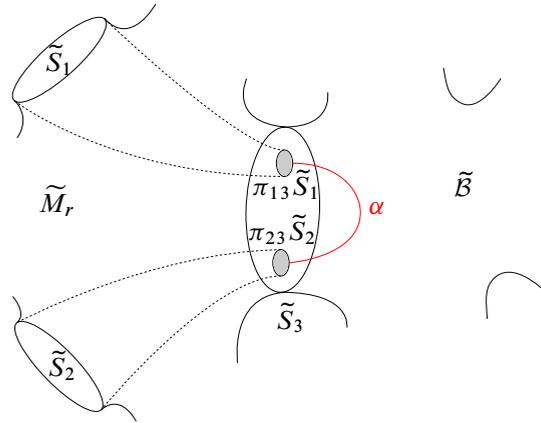
669666

66666

Figure 8: Allowable backtrack in $\mathcal{B}$.

(1) There is a coarsely well-defined geodesic $\alpha$ joining $\pi_{13}(\widetilde{S}_1)$ and $\pi_{23}(\widetilde{S}_2)$ in $\mathcal{B}$. In this case, any $d_{\mathrm{pel}}$-geodesic between $\widetilde{S}_1$ and $\widetilde{S}_2$ in $\widetilde{M}_r \cup \mathcal{B}$ $C_1$-coarsely contains $\alpha$, ie any $d_{\mathrm{pel}}$-geodesic between $\widetilde{S}_1$ and $\widetilde{S}_2$ in $\widetilde{M}_r \cup \mathcal{B}$ contains a subpath tracking $\alpha$ within distance $C_1$ of it. In such a case the $d_{\mathrm{pel}}$-geodesic is said to *have an allowable backtrack in $\mathcal{B}$*.

(2) $\pi_{13}(\widetilde{S}_1) \cup \pi_{23}(\widetilde{S}_2)$ is $C_1$-quasiconvex in $\mathcal{B}$. In this case, a $d_{\mathrm{pel}}$-geodesic $\lambda_p$ between $\widetilde{S}_1$ and $\widetilde{S}_2$ is said to *have nonallowable backtracks in $\mathcal{B}$* if $\lambda_p \cap \mathcal{B} \neq \varnothing$. By perturbing any such $d_{\mathrm{pel}}$-geodesic between $\widetilde{S}_1$ and $\widetilde{S}_2$ by a bounded amount, we obtain a $d_{\mathrm{pel}}$-quasigeodesic that does not intersect $\mathcal{B}$ at all.

**Remark 4.14** The constant $C_1$ above may be chosen to be uniform, ie independent of the choice of $\mathcal{B}$, $\widetilde{M}_r$, $\widetilde{S}_1$, $\widetilde{S}_2$, and $\widetilde{S}_3$ (as there are only finitely many such possibilities up to the action of $\pi_1(F)$).

**4.2.6 Controlling small connectors in elevated subbundles** We generalize the above discussion in Section 4.2.5 to the case where there are two elevations $\widetilde{S}_3$ and $\widetilde{S}'_3$ coming from *different* elevated drilled atoms (see Definition 3.2) abutting a common elevated maximal undrilled subbundle $\mathcal{B}$ (see Definition 3.2).
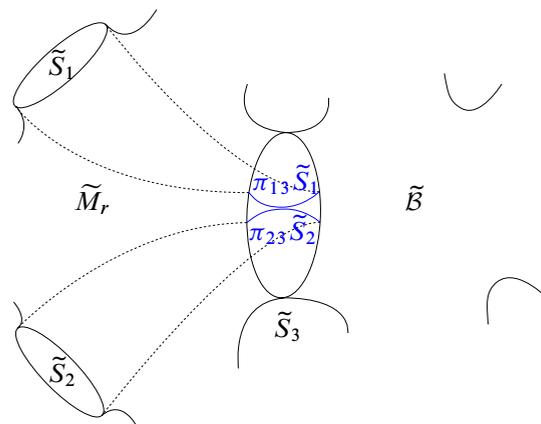


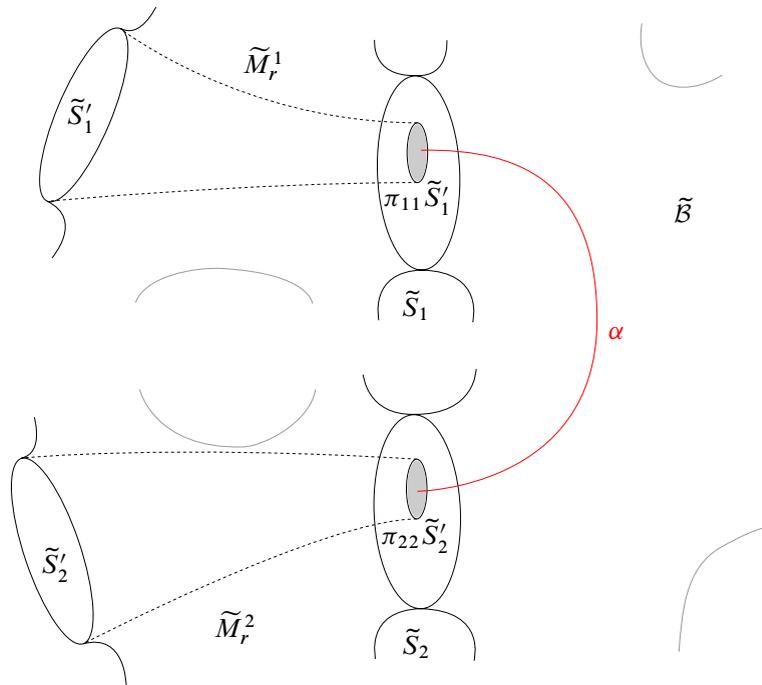Figure 9: Nonallowable backtrack in $\mathcal{B}$.

Figure 10: A long connector in an undrilled elevation.

Let $\widetilde{M}_r^1 \neq \widetilde{M}_r^2$ be drilled atoms of $\widetilde{F}$. Let $\mathcal{B}$ be the elevation of a maximal undrilled subbundle of $F_S$ (Definition 3.2) such that $\mathcal{B} \cap \widetilde{M}_r^i = \widetilde{S}_i$ for $i = 1, 2$. Let $\widetilde{S}_i' \subset \widetilde{M}_r^i$ for $i = 1, 2$ denote boundary components of $\widetilde{M}_r$ (different from $\widetilde{S}_i$). Let $\pi_{11}$ and $\pi_{22}$ be nearest-point projections of $\widetilde{S}_1'$ and $\widetilde{S}_2'$ onto $\widetilde{S}_1$ and $\widetilde{S}_2$, respectively. Then the images of $\pi_{ii}$ for $i = 1, 2$ are either

(1) given by an elevation of a proper essential subsurface of $S_i$ (as in Section 4.2.2), or

(2) uniformly bounded in diameter.

By Theorem 2.1, $\pi_{11}(\widetilde{S}_1')$ and $\pi_{22}(\widetilde{S}_2')$ are uniformly quasiconvex in $\mathcal{B}$. As in Section 4.2.5 above, there exist $C_1 \geq 0$ and $D_1 \geq 0$ such that exactly one of the following holds:

(1) There is a coarsely well-defined geodesic $\alpha$ joining $\pi_{11}(\widetilde{S}_1')$ and $\pi_{22}(\widetilde{S}_2')$ in $\mathcal{B}$. In this case, any $d_{\mathrm{pel}}$-geodesic between $\widetilde{S}_1'$ and $\widetilde{S}_2'$ in $\widetilde{M}_r \cup \mathcal{B}$ $C_1$-coarsely contains $\alpha$.

(2) $\pi_{11}(\widetilde{S}_1') \cup \pi_{22}(\widetilde{S}_2')$ is $C_1$-quasiconvex in $\mathcal{B}$. Further, the distance between $\pi_{11}(\widetilde{S}_1')$ and $\pi_{22}(\widetilde{S}_2')$ in $\mathcal{B}$ is at most $D_1$.

We shall refer to paths $\alpha$ as in item (1) above as *long connectors in undrilled elevations*.

If, on the other hand $\pi_{11}(\widetilde{S}_1') \cup \pi_{22}(\widetilde{S}_2')$ is $C_1$-quasiconvex in $\mathcal{B}$ as in item (2) above, then the triple $\widetilde{M}_r^1, \mathcal{B}, \widetilde{M}_r^2$ will be called a *short connector triple*.

**4.2.7 Defining the path family** We now define the family $\mathcal{F}$ of paths that will feed into Theorem 4.2. Recall that $(\widetilde{F}, d_{\mathrm{pel}})$ is hyperbolic by Theorem 3.14. For $x, y \in \widetilde{F}$, define $\eta_p(x, y)$ to be a geodesic in $(\widetilde{F}, d_{\mathrm{pel}})$. Now, replace $\eta_p(x, y)$ by a $d_{\mathrm{pel}}$-quasigeodesic that

(1) does not backtrack from drilled atoms (in the sense of Definition 3.30) as in Corollary 3.29,

(2) does not have nonallowable backtracks in elevations of maximal undrilled subbundles.

Abusing notation slightly, we continue to denote the nonbacktracking $d_{\mathrm{pel}}$-quasigeodesic by $\eta_p$. For each drilled atom $\widetilde{M}_r$ (see Definition 3.2) and each connected component $\zeta_r$ of $\eta_p \cap \widetilde{M}_r$, replace $\zeta_r$ by an electroambient geodesic joining its endpoints. Denote the resulting path in $\widetilde{F}$ joining the endpoints of $\eta_p$ by $\eta$.

We refer to $\eta$ as the electroambient path obtained from $\eta_p$ by *de-electrification*. Next, define $\mathcal{F}_p$ to be a collection $\{\eta_p(x, y)\}$ of uniform quasigeodesics in $(\widetilde{F}, d_{\mathrm{pel}})$ without backtracking in drilled atoms (cf Definition 3.30), one for every pair $x, y \in \widetilde{F}$. Define $\mathcal{F}$ to be the collection $\{\eta(x, y)\}$ obtained from the collection $\{\eta_p(x, y)\}$ by de-electrification. For each such $\eta(x, y)$ define $\theta(x, y)$ to be the closure of $\eta(x, y) \setminus \bigcup_{P \in \mathcal{P}} P$. Thus, $\theta(x, y)$ is obtained from $\eta(x, y)$ by removing the interiors of the intersections with elements of $\mathcal{P}$.

In short, $\mathcal{F}$ is obtained from geodesics in $(\widetilde{F}, d_{\mathrm{pel}})$ by

(1) removing backtracking in drilled atoms (see Definition 3.30), and removing nonallowable backtracks in elevations of maximal undrilled subbundle (see Definition 3.2),

(2) subsequent de-electrification.

**Remark 4.15** The purpose of replacing a $d_{\mathrm{pel}}$-geodesic by a $d_{\mathrm{pel}}$-quasigeodesic that does not backtrack from drilled atoms is to minimize intersections with elevated singular fibers (see Definition 1.2) $\widetilde{S}_x$ in $\widetilde{F}$. This ensures that the only allowable backtracking is in elevations of maximal undrilled subbundles (see Definition 3.2).

## 4.3 Stability

The aim of this subsection is to prove the following stability condition, which is the main technical result of this section:

**Proposition 4.16** *Given $D > 0$, there exists $C > 0$ such that the following holds*: *Let $\eta(x, y), \eta(u, v) \in \mathcal{F}$ such $d(x, u) \leq D$ and $d(y, v) \leq D$. Then $\eta(x, y)$ and $\eta(u, v)$ track each other in a $C$-neighborhood of each other.*

We first observe a version of Proposition 4.16 in atoms.

**Lemma 4.17** *Given $D > 0$, there exists $C > 0$ such that the following holds*: *Let $\eta(x, y), \eta(u, v) \in \mathcal{F}$ be such that*:

(1) *There exists an atom $\widetilde{\mathbb{M}}$ equal to $\widetilde{M}_r$ (drilled) or $\widetilde{M}$ (undrilled) such that $\eta(x, y), \eta(u, v)$ are contained in $\widetilde{\mathbb{M}}$.*

(2) *$x, u \in \widetilde{S}_x$ and $y, v \in \widetilde{S}_y$, where $\widetilde{S}_x, \widetilde{S}_y$ are elevations of singular fibers (see Definition 1.2), or equivalently, boundary components of $\widetilde{\mathbb{M}}$.*

(3) *$d(x, u) \leq D$ and $d(y, v) \leq D$.*

*Then $\eta(x, y)$ and $\eta(u, v)$ track each other in a $C$-neighborhood of each other.*

**Proof** This is a consequence of Lemma 4.4, which guarantees that $\eta(x, y)$ and $\eta(u, v)$ track each other away from the elements of $\mathcal{P}$ in $\widetilde{\mathbb{M}}$. Further, since each element of $\mathcal{P}$ is a flat $\mathbb{R}^2$, geodesics in each element of $\mathcal{P}$ track each other provided they start and end nearby. The lemma now follows from the construction of electroambient quasigeodesics (see Definition 4.3). □

Before starting with the proof of Proposition 4.16, we point out that the main idea below is to divide an element $\eta \in \mathcal{F}$ into pieces that satisfy the property that its endpoints are coarsely well-defined.

**Proof of Proposition 4.16** Assume, without loss of generality, that $x, u \in \widetilde{S}_1 \subset X_a$ and $y, v \in \widetilde{S}_2 \subset X_b$, where $a, b \in \mathcal{T}$. There exists an indexing set $\mathcal{I}$ giving a sequence of vertices $a = a_0, \ldots, a_n = b$, possibly with repetition, such that $\eta(x, y)$ traverses $\widetilde{M}_{a_i}$ in order. Since $\eta(x, y) \in \mathcal{F}$, it does not

(1) backtrack in drilled atoms (see Definition 3.30),

(2) have nonallowable backtracks in elevations of maximal undrilled subbundle (see Definition 3.2).

If there exists a maximal undrilled subbundle $\mathcal{B}$ (see Definition 3.2) in which $\eta(x, y)$ has an allowable backtrack (Section 4.2.5), we collect together all the (necessarily consecutive) vertices in the sequence $\{a_i\}$ that correspond to atoms contained in $\mathcal{B}$ and replace them by a single vertex $B_j$ for some $j$. We refer to such $B_j$ as an *undrilled molecule*. Thus, if some such $B_j$ occurs, then there exists $a_i \in \mathcal{I}$ such that $B_j$ occurs in a unique subsequence of the form $a_i B_j a_i$ in $\mathcal{I}$.

Again, if there exists a maximal undrilled subbundle $\mathcal{B}$ (see Definition 3.2) in which $\eta(x, y)$ has a long connector (Section 4.2.6), then also we collect together all the (necessarily consecutive) vertices in the sequence $\{a_i\}$ that correspond to (elevated) atoms contained in $\mathcal{B}$ and replace them by a single vertex $B_k$ for some $k$. We also refer to such a $B_k$ as an *undrilled molecule*. Thus, if some such $B_k$ occurs, then there exists $a_i \neq a_s \in \mathcal{I}$ such that $B_k$ occurs in a unique subsequence of the form $a_i B_k a_s$ in $\mathcal{I}$.

The construction of undrilled molecules now allows us obtain a new finite sequence

$$\mathcal{J} = a_{1,1}, \ldots, a_{1,m_1}, B_1, a_{2,1}, \ldots, a_{2,m_2}, B_2, \ldots.$$

Note that the only possible repetition allowable in this sequence are of the following form. If $\eta$ has an allowable backtrack in $B_j$ for some $j$, then there is a triple of the form $a_{j,m_j}, B_j, a_{j+1,1}$ with $a_{j,m_j} = a_{j+1,1}$ corresponding to the same elevated drilled atom (see Definition 3.2).

By the properties of an allowable backtrack (Section 4.2.5) or long connectors in undrilled elevations (Section 4.2.6), both $\eta(x, y)$ and $\eta(u, v)$ coarsely contain a subpath in $B_j$ for all undrilled molecules $B_j$ (here we are conflating the index $B_j$ with the elevated maximal undrilled subbundle (see Definition 3.2) it indexes). Let $\widetilde{S}_j^-$ and $\widetilde{S}_j^+$ denote the boundary components of $B_j$ (see Definition 3.25) through which $\eta(x, y)$ and $\eta(u, v)$ enter and leave $B_j$. Then there exist

(1) $z_j^-(x, y) \in \eta(x, y) \cap \widetilde{S}_j^-$ and $z_j^+(x, y) \in \eta(x, y) \cap \widetilde{S}_j^+$,

(2) $z_j^-(u, v) \in \eta(u, v) \cap \widetilde{S}_j^-$ and $z_j^+(u, v) \in \eta(u, v) \cap \widetilde{S}_j^+$,

such that

(1) $z_j^-(x, y)$ and $z_j^-(u, v)$ lie at a distance of at most $2C_1$ from each other on $\widetilde{S}_j^-$,

(2) $z_j^+(x, y)$ and $z_j^+(u, v)$ lie at a distance of at most $2C_1$ from each other on $\widetilde{S}_j^+$.

Let $\eta_j(x, y)$ denote the subpath of $\eta(x, y)$ between $z_j^-(x, y)$ and $z_j^+(x, y)$. Let $\eta_j(u, v)$ denote the subpath of $\eta(u, v)$ between $z_j^-(u, v)$ and $z_j^+(u, v)$. By hyperbolicity of each $B_j$, $\eta_j(x, y)$ and $\eta_j(u, v)$ track each other in a $C_1'$-neighborhood of each other, where $C_1'$ depends only on $C_1$ and the hyperbolicity constant of $B_j$. Hence $C_1'$ is uniform.

To prove Proposition 4.16, it therefore suffices to assume that the finite sequence $\mathcal{J}$ constructed from $\mathcal{I}$ *does not contain any undrilled molecule $B_j$*.

A caveat is in order. We note that if there is a short connector triple $\widetilde{M}_r^1, \mathcal{B}, \widetilde{M}_r^2$ as in Section 4.2.6, then the atoms of the elevated maximal undrilled subbundle (see Definition 3.2) $\mathcal{B}$ are *not* combined into a single undrilled molecule. Further, there exists uniform $C_2 \geq 1$ (independent of $\mathcal{B}$, $\eta(x, y)$, and $\eta(u, v)$) such that after a uniformly bounded perturbation if necessary, both $\eta(x, y)$ and $\eta(u, v)$

(1) intersect the same set of undrilled atoms (see Definition 3.2) $A_1, \ldots, A_k$ of $\mathcal{B}$ in order without backtracking in any of the atoms (ie after leaving any of the undrilled atoms $A_i$, $\eta(x, y)$, and $\eta(u, v)$ do not return to it),

(2) $k \leq C_2$ (this follows from the quasiconvexity property of the union of projections used to define short connector triples in Section 4.2.6).

We summarize this by saying that $\eta(x, y)$ and $\eta(u, v)$ have no backtracking in short connector triples.

We now return to the sequence of atoms $\mathcal{J} = a_0, \ldots, a_m$ where

(1) $\mathcal{J}$ has no undrilled molecule, and hence

(2) $\mathcal{J}$ is the vertex sequence of a *geodesic* in the Bass–Serre tree $\mathcal{T}$ of $\widetilde{F}$.

The absence of backtracking in short connector triples along with no backtracking in drilled atoms (see Definition 3.30) guarantees that $\mathcal{J}$ is the vertex sequence of a geodesic.

Let $L$ be the maximum of $C_2$ and the constants in Lemmas 4.11 and 4.13. If there is a sequence of more than $L$ contiguous undrilled blocks in $\mathcal{J}$, then choose a maximal sequence $a_k, \ldots, a_l$ with $l - k \geq L$ indexing such blocks. Then, by Lemma 4.13, there exists a coarsely well-defined $\alpha$ connecting boundary components $\widetilde{S}_k$ and $\widetilde{S}_l$ (the *internal boundary components* occurring in Lemma 4.13) coarsely contained in both $\eta(x, y)$ and $\eta(u, v)$. Hence there exist $z_k(x, y) \in \eta(x, y) \cap \widetilde{S}_k$ and $z_k(u, v) \in \eta(u, v) \cap \widetilde{S}_k$ such that $z_k(x, y)$ and $z_k(u, v)$ lie at a distance of at most $2C_1$ from each other on $\widetilde{S}_k$, where $C_1$ is the uniform constant of coarseness from Lemma 4.13. We may therefore assume henceforth that there does not exist a sequence of more than $L$ contiguous undrilled blocks in $\mathcal{J}$.

Next, let $a_{k_1}, \ldots, a_{k_L}$ be a subsequence of $\mathcal{J}$ such that

(1) each $a_{k_i}$ indexes a drilled atom (see Definition 3.2),

(2) $k_{i+1} > k_i$,

(3) for any $i \in 1, \ldots, L$, every $j$ strictly between $a_{k_i}$ and $a_{k_{i+1}}$ indexes an undrilled atom (see Definition 3.2).

Then, by the simplification in the above paragraph, $a_{k_{i+1}} - a_{k_i} \leq L$. Thus, any subsequence of $L$ drilled atoms in $\mathcal{J}$ interpolated only by undrilled atoms (see Definition 3.2) has length at most $L^2$. We shall refer to such a subsequence of $\mathcal{J}$ as a *subsequence of successive $L$ drilled atoms*. Note that in a subsequence of successive $L$ drilled atoms, the drilled atoms (see Definition 3.2) need not be contiguous.

For such a subsequence of successive $L$ drilled atoms, let $a = a_{k_1}$, $b = a_{k_L}$, and $X_{[a,b]}$, $\widetilde{S}_a$, and $\widetilde{S}_b$ be as in Lemma 4.11. Then, by Lemma 4.11, there exists a coarsely well-defined $\alpha$ connecting boundary components (see Definition 3.25) $\widetilde{S}_a$ and $\widetilde{S}_b$ coarsely contained in both $\eta(x, y)$ and $\eta(u, v)$.

Finally divide $\mathcal{J}$ into a subsequence of successive $L$ drilled atoms as follows: let

$$\mathcal{J} = a_0 = a_{n_0}, \ldots, a_{n_1}, \ldots, a_{n_s}, \ldots, a_m$$

be such that

(1) each $a_{n_i}$ for $i = 1, \ldots, s$ is a drilled atom (see Definition 3.2),

(2) the subsequence of $\mathcal{J}$ between $a_{n_i}$ and $a_{n_{i+1}}$ (both included) has exactly $L$ drilled atoms for $i = 0, \ldots, s$.

We also call the initial sequence $a_{n_0}, \ldots, a_{n_1}$ a subsequence of successive $L$ drilled atoms. For each such subsequence of successive $L$ drilled atoms, there exists a coarsely well-defined $\alpha$ connecting boundary components (see Definition 3.25) $\widetilde{S}_a$ and $\widetilde{S}_b$ as before, and coarsely contained in both $\eta(x, y)$ and $\eta(u, v)$.

Setting $c = a_0$ and $d = a_{n_s}$, it follows that there exist $C'$, and $y', v' \in \widetilde{S}_d$, and subpaths $\eta(x, y')$ and $\eta(u, v')$ of $\eta(x, y)$ and $\eta(u, v)$, respectively,

(1) starting at $x$ and $u$, respectively,

(2) ending at $y'$ and $v'$, respectively,

such that

(1)  $y'$ and $v'$ lie at a distance of at most $C'$ from each other on $\widetilde{S}_d$,

(2)  $\eta(x, y')$ and $\eta(u, v')$ track each other in a $C'$-neighborhood of each other,

(3)  $C'$ is independent of $x$, $y$, $u$, $v$, and $\mathcal{J}$.

Let $\eta(y', y)$ and $\eta(v', v)$ denote the subpaths of $\eta(x, y)$ and $\eta(u, v)$, respectively,

(1)  starting at $y'$ and $v'$, respectively,

(2)  ending at $y$ and $v$, respectively.

Let $p = a_{n_s}$ and $q = a_m$. Then $X_{[p,q]}$ is a 3-manifold given by a concatenation of at most $L^2$ atoms. The proof of Lemma 3.12 now furnishes relative hyperbolicity for each such $X_{[p,q]}$. Since there are only finitely many possibilities, the constants of relative hyperbolicity are uniform. The tracking properties of $\eta(y', y)$ and $\eta(v', v)$ now follow from relative hyperbolicity. Combining this with the tracking properties of $\eta(x, y')$ and $\eta(u, v')$ already established, Proposition 4.16 follows.  □

## 4.4  Checking conditions 1, 2, and 4–6 of Theorem 4.2

We shall now show that the family $\mathcal{F}$ defined above satisfy the conditions of Theorem 4.2.

**Theorem 4.2(1)**  After rescaling $F$ if necessary, we might as well assume that the distance between any singular fiber (see Definition 1.2) of $F$ and the boundary $\partial(N_\epsilon(\sigma_i))$ of the neighborhood of any drilled curve is at least 4. Hence, if $d_X(x, y) \leq 2$, then the condition follows from strong relative hyperbolicity of $\widetilde{M}_r$ where $M_r$ is any drilled atom (see Definition 3.2) in $F$ (Lemma 4.4).

**Theorem 4.2(2)**  This is a consequence of the proof of stability of elements of $\mathcal{F}$, Proposition 4.16.

**Theorem 4.2(4)**  This follows from the fact that any element of $\mathcal{F}$ starting on $P_1 \in \mathcal{P}$ and ending on an element $P_2 \neq P_1$ of $\mathcal{P}$ necessarily has points in the complement of $\bigcup_{P \in \mathcal{P}} P$.

**Theorem 4.2(5)**  This follows from Corollary 4.10. Indeed, for any $P_1, P_2 \in \mathcal{P}$, there exists $[a_1, a_2] \in \mathcal{T}$ such that $P_i \in X_{a_i}$ for $i = 1, 2$. The strong relative hyperbolicity of $X_{[a_1,a_2]}$ relative to the collection of elements of $\mathcal{P}$ contained in it (with constants depending only on $d_\mathcal{T}(a_1, a_2)$) furnishes (5).

**Theorem 4.2(6)**  This also follows from Corollary 4.10. Indeed, as in (5) above, we can choose $[a_1, a_2] \in \mathcal{T}$ such that $d_\mathcal{T}(a_1, a_2) \leq 2k$, where $k$ is as in (6). Then strong relative hyperbolicity of $X_{[a_1,a_2]}$ relative to the collection of elements of $\mathcal{P}$ contained in it furnishes the constant $K$ required by (6).

## 4.5  Thin triangles in $\mathcal{F}$

It remains to prove the thin triangle condition, ie Theorem 4.2(3). Let $a, b, c \in \widetilde{F}$.
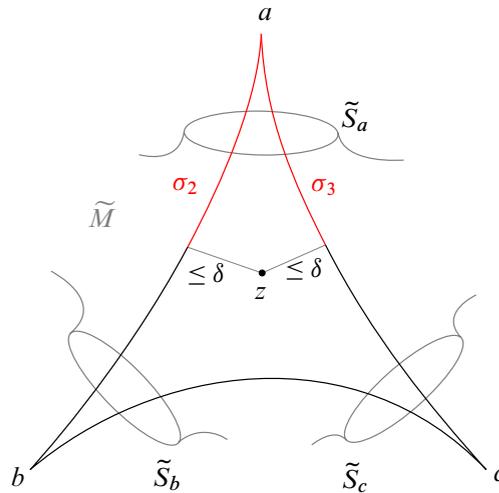
Figure 11: The case when the median lies in an undrilled atom.

Let $\gamma_1^d$, $\gamma_2^d$, and $\gamma_3^d$ be sides of a quasigeodesic triangle in $(\widetilde{F}, d_{\mathrm{pel}})$ with vertices $a$, $b$, and $c$ used for constructing elements of $\mathcal{F}$. Let $\gamma_1, \gamma_2, \gamma_3 \in \mathcal{F}$ denote the elements of the path family constructed from them. Let $a$ (resp. $b, c$) be the vertex opposite $\gamma_1$ (resp. $\gamma_2, \gamma_3$). Let $z$ denote a centroid.

**Case 1** ($z$ lies in an undrilled atom $\widetilde{M}$ in $\widetilde{F}$; see Figure 11)  Thinness of triangles follows from stability, Proposition 4.16. Indeed $z$ lies close to each of $\gamma_1$, $\gamma_2$, and $\gamma_3$ (in the usual unelectrified metric on $\widetilde{F}$), as each boundary component, and hence $\widetilde{M}$, is properly embedded in $(\widetilde{F}, d_{\mathrm{pel}})$; see Lemma 3.28. Hence for $a$ (or $b$, or $c$) there exist a pair of paths $\sigma_2, \sigma_3$ (given by subpaths of $\gamma_2$ and $\gamma_3$, respectively) starting from $a$ and ending close to $z$ in $\widetilde{M}$. By Proposition 4.16, $\sigma_2$ and $\sigma_3$ track each other (with uniform constants). A similar argument works for $b$ and $c$ completing this case.

**Case 2** ($z$ lies in a drilled atom $\widetilde{M}_r$ in $\widetilde{F}$; see Figure 12)  There are three boundary components $\widetilde{S}_a$, $\widetilde{S}_b$, and $\widetilde{S}_b$ of $\widetilde{M}_r$ (recall that the vertices of the triangle are $a$, $b$, and $c$) and subpaths $\beta_1^d$, $\beta_2^d$, and $\beta_3^d$ of $\gamma_1^d$, $\gamma_2^d$, and $\gamma_3^d$ such that the following holds. In $\widetilde{M}_r$ equipped with $d_{\mathrm{pel}}$, the subpaths $\beta_1^d$, $\beta_2^d$, and $\beta_3^d$ pass close to $z$. We assume that $\beta_1^d$ and $\beta_2^d$ have one endpoint each on $\widetilde{S}_c$, $\beta_3^d$ and $\beta_1^d$ have one endpoint each on $\widetilde{S}_b$, and $\beta_2^d$ and $\beta_3^d$ have one endpoint each on $\widetilde{S}_a$.

This gives a hexagon in $\widetilde{M}_r$, where the other three sides (other than $\beta_1^d$, $\beta_2^d$, and $\beta_3^d$) are geodesics in $\widetilde{S}_a$, $\widetilde{S}_b$, and $\widetilde{S}_c$, joining the two intersection points of the segments. Call these sides the *complementary geodesics*, and denote them by $\alpha_a$, $\alpha_b$, and $\alpha_c$, respectively. Call the subpaths $\beta_1^d$, $\beta_2^d$, and $\beta_3^d$ the *internal geodesics*.

Suppose that at least one of the complementary geodesics, say $\alpha_a$, is long. Note that $\alpha_a \subset \widetilde{S}_a$ joins $x_2$ and $x_3$, where $x_2$ and $x_3$ are respectively the intersection points of $\beta_2^d$ and $\beta_3^d$ with $\widetilde{S}_a$. Let $p_2 \in \beta_2^d$ and $p_3 \in \beta_3^d$ be points that are $\delta$-close to $z$ in the $d_{\mathrm{pel}}$-metric, where $\delta$ is a constant depending only on the hyperbolicity constant of $(\widetilde{F}, d_{\mathrm{pel}})$; see Theorem 3.14. Since $\widetilde{S}_a$ is quasi-isometrically embedded
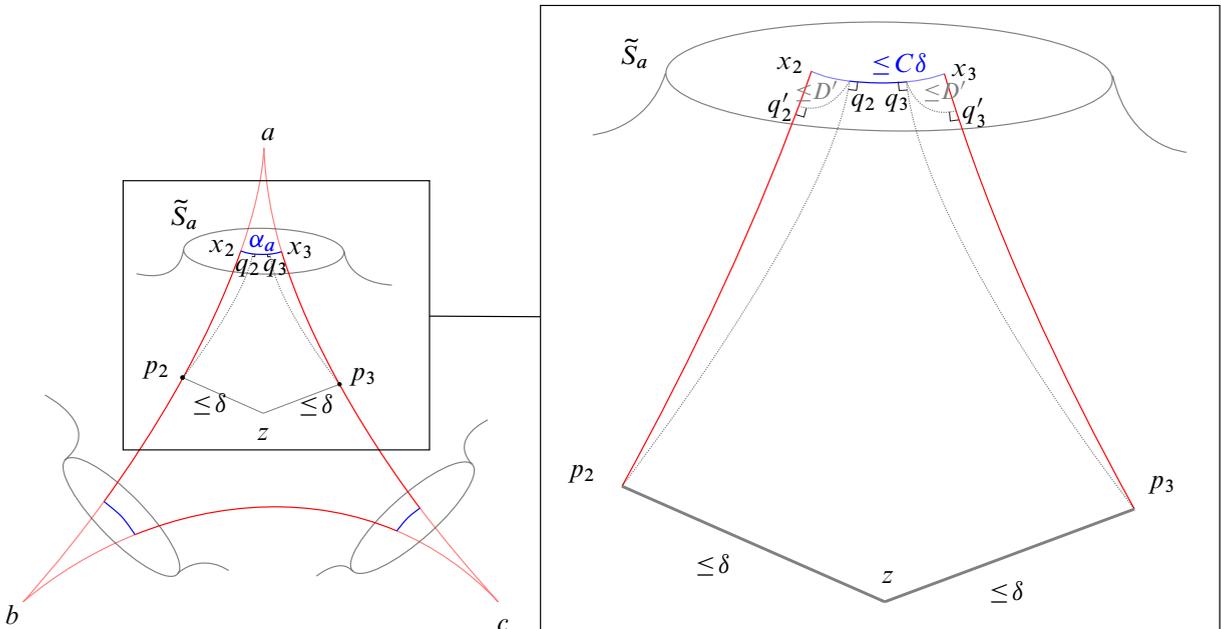
Figure 12: The case when the median lies in a drilled atom.

(Lemma 3.12) in $(\widetilde{M}_r, d_{\text{pel}})$, $\alpha_a$ is a quasigeodesic in $(\widetilde{M}_r, d_{\text{pel}})$. (The quasigeodesic constants are uniform as there are only finitely many possibilities for drilled atoms; see Definition 3.2.) Let $\pi_a$ denote a nearest-point projection of $(\widetilde{M}_r, d_{\text{pel}})$ onto $\alpha_a$. Let $\pi_a(p_i) = q_i$ for $i = 2, 3$. Then the $d_{\text{pel}}$-length of the subsegment $\alpha_a(q_2, q_3)$ of $\alpha_a$ joining $q_2$ and $q_3$ is at most $8\delta$ (see [22, Lemma 3.2] for instance for a proof of this standard fact). Since $\alpha_a$ is a quasigeodesic in $(\widetilde{M}_r, d_{\text{pel}})$, the length of $\alpha_a(q_2, q_3)$ in the intrinsic metric on $\widetilde{S}_a$ is at most $C\delta$, where $C$ depends only on the uniform quasigeodesic constant for $\alpha_a$, and is therefore uniform.

Let $[x_2, q_2]$ (resp. $[x_3, q_3]$) denote the geodesic in $\widetilde{S}_a$ joining $x_2$ and $q_2$ (resp. $x_3$ and $q_3$). Also, let $[q_2, p_2]_{\text{pel}}$ (resp. $[q_3, p_3]_{\text{pel}}$) denote the $d_{\text{pel}}$-geodesic segments in $\widetilde{M}_r$ joining $q_2$ and $p_2$ (resp. $q_3$ and $p_3$). Then

(1)  $\beta'_2 = [x_2, q_2] \cup [q_2, p_2]_{\text{pel}}$ is a $d_{\text{pel}}$-quasigeodesic with uniform constants (cf [22, Lemma 3.2]),

(2)  $\beta'_3 = [x_3, q_3] \cup [q_3, p_3]_{\text{pel}}$ is a $d_{\text{pel}}$-quasigeodesic with uniform constants.

Let $\beta_2$ (resp. $\beta_3$) denote the subpaths of $\beta_2^d$ (resp. $\beta_3^d$) joining $x_2$ and $p_2$ (resp. $x_3$ and $p_3$). By relative hyperbolicity of $\widetilde{M}_r$ (Lemma 3.12), and the tracking properties in Lemma 4.4, there exists $q'_2 \in \beta_2$ (resp. $q'_3 \in \beta_3$) such that the distance (in the unelectrified metric $d$ on $\widetilde{M}_r$) between $q_2$ and $q'_2$ (resp. $q_3$ and $q'_3$) is uniformly bounded by a uniform constant $D'$. Hence $d(q'_2, q'_3) \leq C\delta + 2D'$.

Let $\gamma'_2$ (resp. $\gamma'_3$) be the subpath of $\gamma_2$ from $a$ to $q'_2$ (resp. $a$ to $q'_3$). By Proposition 4.16, $\gamma'_2$ and $\gamma'_3$ track each other within a uniform distance $C''$ of each other.

Removing the initial subpath of $\beta_2^d$ (resp. $\beta_3^d$) between $x_2$ and $q_2'$ (resp. $x_3$ and $q_3'$), we can replace the complementary geodesic $\alpha_a$ by a geodesic $\alpha_a'$ of length at most $C\delta + 2D'$ joining $q_2'$ and $q_3'$. Carrying out such replacements for all the long complementary geodesics, we obtain a $d_{\mathrm{pel}}$-quasigeodesic hexagon whose complementary geodesics are uniformly bounded in length. Let $\beta_f^1$, $\beta_f^2$, and $\beta_f^3$ be the internal geodesics of the resulting hexagon. By strong relative hyperbolicity of $\widetilde{M}_r$ with respect to $\mathcal{P}_r$, the internal geodesics $\beta_f^1$, $\beta_f^2$, and $\beta_f^3$ satisfy Theorem 4.2(3).

Combining this with the tracking properties of pairs such as $\gamma_2'$ and $\gamma_3'$ proved above, using Proposition 4.16, it follows that $\gamma_1, \gamma_2, \gamma_3 \in \mathcal{F}$ satisfy Theorem 4.2(3).

**Proof of Theorem 3.1**   That the family $\mathcal{F}$ constructed in Section 4.2.7 satisfies Theorem 4.2(1), (2), and (4)–(6) has been established in Section 4.4. The thin triangles condition, ie Theorem 4.2(3), has been checked above in this subsection. Hence, by Theorem 4.2, $\widetilde{F}$ is strongly hyperbolic relative to the collection $\mathcal{P}$.                                                                                   $\square$

## 4.6   Relative quasiconvexity

Let $S_0$ be a drilled fiber in $E$. Further, after isotoping the drilled curves if necessary, we can assume that the collection of drilled curves $\sigma_1, \ldots, \sigma_k$ in $S_0$ is maximal, ie no other drilled curve may be isotoped into $S_0$ in the complement of $\bigcup_{i=1,\ldots,k} N_\epsilon(\sigma_i)$.

**Definition 4.18**   The collection of drilled fibers is *reduced* if the collection of drilled curves in any singular fiber (see Definition 1.2) is maximal in the above sense.

Then $S_0 \setminus \left( \bigcup_i \sigma_i \right)$ consists of finitely many components $\Sigma_1, \ldots, \Sigma_m$. By Theorem 2.1, we have the following:

**Lemma 4.19**   *Each $\pi_1(\Sigma_i)$ is quasiconvex in $\pi_1(E)$.*

Next, we consider $F$ and $(\widetilde{F}, d_{\mathrm{pel}})$. Then we have:

**Lemma 4.20**   *Let $\Sigma$ be a component of $S_0 \setminus \left( \bigcup_i \sigma_i \right)$ as above for $S_0$ a drilled fiber. Then there exists $C \geq 1$ such that any elevation $\widetilde{\Sigma}$ is quasiconvex $(\widetilde{F}, d_{\mathrm{pel}})$.*

**Proof**   The same argument as in Section 3.5.2 in the paragraph "Identifying $\rho$-thin annuli" identifies the collection of essential annuli (see Definition 3.7) with core curve homotopic to a curve in $\Sigma$. Corollary 3.23 now establishes that $\Sigma$ flares in all directions in the sense of Definition 3.18. Hence, by Corollary 3.19 and Remark 3.20, there exists $C \geq 1$ such that any elevation $\widetilde{\Sigma}$ is quasiconvex $(\widetilde{F}, d_{\mathrm{pel}})$.                         $\square$

We finally have the following:

**Proposition 4.21**   *Let $\Sigma$ be a component of $S_0 \setminus \left( \bigcup_i \sigma_i \right)$ as above for $S_0$ a drilled fiber. Then there exists $C' \geq 1$ such that any elevation $\widetilde{\Sigma}$ is relatively $C'$-quasiconvex.*

**Proof** Let $\mathcal{T}$ denote the Bass–Serre tree of $\widetilde{F}$, $\Pi\colon \widetilde{F} \to \mathcal{T}$ denote the tree of spaces structure, and $v$ be the vertex such that $\widetilde{S}_0 \subset \widetilde{M}_v$. Let $\mathcal{T}_0$ denote the $C$-neighborhood of $v$ in $\mathcal{T}$, where $C$ is as in Lemma 4.20. Also, let $X_0 = \Pi^{-1}(\mathcal{T}_0)$. By Lemma 4.20 the $d_{\mathrm{pel}}$-geodesic joining a pair of points $x, y \in \widetilde{\Sigma}$ lies in $X_0$. Further, the proof of Lemma 3.28 establishes that $\widetilde{\Sigma}$ with its intrinsic metric is qi-embedded in $(X_0, d_{\mathrm{pel}})$.

The proof of Theorem 3.1 applied to $X_0$ establishes strong-relative hyperbolicity of $X_0$ relative to the collection $\mathcal{P}_0$ consisting of the elements of $\mathcal{P}$ contained in $X_0$. Since $\widetilde{\Sigma}$ with its intrinsic metric is qi-embedded in $(X_0, d_{\mathrm{pel}})$, the construction of the path family $\mathcal{F}$ in Section 4.2.7 shows that we can take $\eta(x, y)$ to lie on $\widetilde{\Sigma}$ for $x, y \in \widetilde{\Sigma}$. Hence, $\widetilde{\Sigma}$ is relatively $C'$-quasiconvex for some $C'$. $\qquad\square$

# 5 Cubulation

Notation to be used in this section:

- $\eta$ is used to denote subdivision intervals. $F_\eta$ is the drilled bundle restricted to $\eta$. $\{\Sigma_{\eta,i}\}$ are the subsurface components in the complement of the drilled curves of $F_\eta$.

- For $\mathcal{K}$ a maximal subgraph of $\mathcal{G}$ not containing any drilled edges, $F(\mathcal{K})$ denotes the drilled subbundle of $F$ restricted to $\mathcal{K}$.

- $\mathcal{G}_\mathcal{K}$ is the canonical reduced form (see Definition 5.2) of $\mathcal{G}$.

We refer the reader to [14; 33; 34] for details on virtually special CAT(0) cube-complexes. However, before attempting to cubulate $G = \pi_1(F)$, we first describe another graph of groups structure on $G$.

## 5.1 Another graph of groups structure on $G$

We construct a new graph of groups structure on $\pi_1(F)$, with a new underlying graph $\mathcal{G}_0$ as follows.

Assume that the collection of drilled fibers is reduced in the sense of Definition 4.18. The vertex groups of $\mathcal{G}_0$ are all isomorphic to $H = \pi_1(S)$, and consist of the following:

(1) For every singular fiber $S_v$ (see Definition 1.2), necessarily undrilled by construction, we have a vertex group $G_v$ isomorphic to $H$.

(2) For a drilled edge $e$, let $S_{x_1}, \ldots, S_{x_p}$ denote the drilled fibers where $x_i$ are points in order along $e$, where $e$ has initial and final vertices $v_i$ and $v_f$. Interpolate vertices $v_i = v_0, v_1, \ldots, v_p = v_f$ where $v_i$ lies between $x_i$ and $x_{i+1}$ for $1 \le i \le p-1$. The points $x_1, \ldots, x_p$ will be referred to as *special drilled points*, and $v_1, \ldots, v_{p-1}$ will be referred to as *interpolating points*.

The vertices of $\mathcal{G}_0$ consist of the vertices of $\mathcal{G}$ *along with interpolating points*.

We now define the edge spaces. The intervals $[v_i, v_{i+1}]$ will be termed *subdivision intervals*. Note that each subdivision interval contains a unique special drilled point $x_{i+1}$. Let $\eta = [c, d]$ be a subdivision interval, so that the vertex groups $G_c$ and $G_v$ are isomorphic to $H$. Now, restrict the drilled bundle $F$ to $\eta$ to obtain $F_\eta$. Then $F_\eta$ is homeomorphic to $S \times [0, 1]$ after removing $\epsilon$-neighborhoods of a nonempty
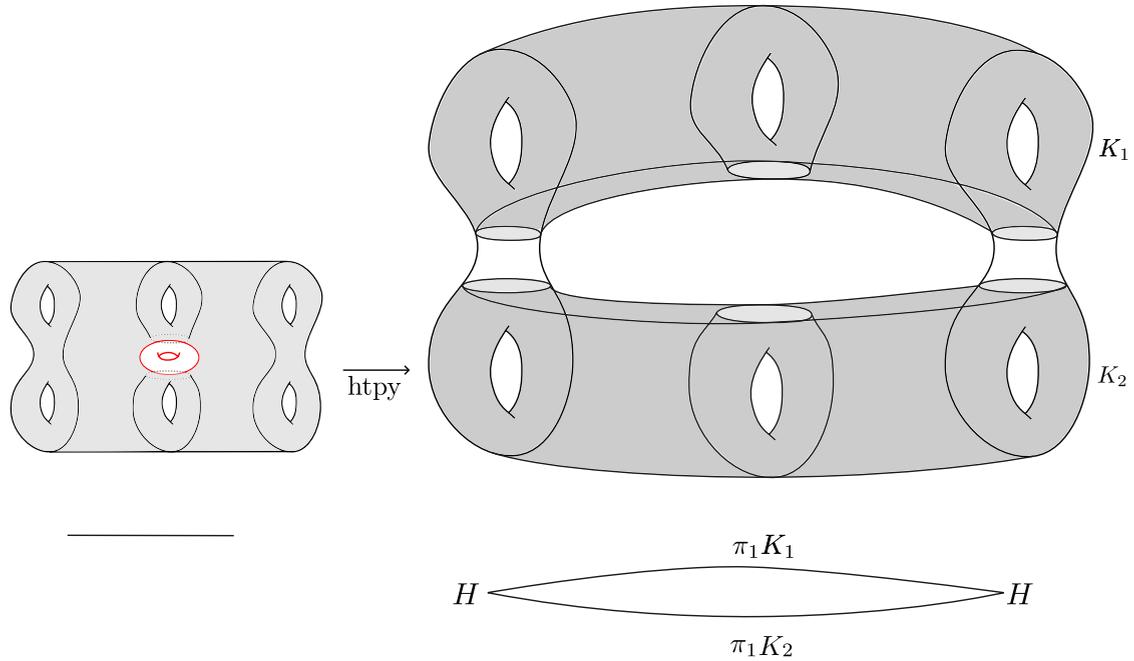
Figure 13: The graph of groups structure for a drilled atom.

family $\{\sigma_i\}$ of disjoint, homotopically distinct, essential simple closed curves on a unique drilled fiber $S_{x_i}$. Let $\Sigma_{\eta,i}$ for $i = 1, \ldots, m$ denote the components of $S \setminus \bigcup_i \sigma_i$. Then $\pi_1(F_\eta)$ admits a graph of groups description with two vertex groups $G_c$ and $G_d$, each isomorphic to $H$, and $m$ edge groups, isomorphic to $\pi_1(\Sigma_{\eta,i})$ for $i = 1, \ldots, m$. Let $\{e_{\eta,i}, i = 1, \ldots, m\}$ denote the resulting edges "living over" the subdivision on interval $\eta$. We call these *subdivision edges*. The groups $\pi_1(\Sigma_{\eta,i})$ for $i = 1, \ldots, m$ are called the *subdivision edge groups*.

The edges of $\mathcal{G}_0$ are obtained as follows:

(1) If there is an edge of $\mathcal{G}$ that has no special drilled point, leave it as it is in $\mathcal{G}_0$. We refer to these as *undrilled edges* of $\mathcal{G}_0$.

(2) Next, replace each interval with at least one special drilled point by subdivision intervals, and finally each subdivision interval by the subdivision edges that live over it.

Note that the above discussion goes through even when the initial and final vertices of the drilled edge $e$ coincide, so that there is a monodromy map $\phi$ for the bundle $E$ restricted to the (closed) $e$. We explicate the inclusion maps for $\pi_1(\Sigma_{\eta,i})$ for $i = 1, \ldots, m$ into the vertex groups $H_i$ and $H_f$ corresponding to the vertices $v_i$ and $v_f$ (the initial and final vertices of $e$). We identify $\pi_1(\Sigma_{\eta,i})$ for $i = 1, \ldots, m$ with subgroups of $H_i$ via the product structure on $S \times [c, d]$. Then, modulo this identification, the edge-to-vertex group maps for $\pi_1(\Sigma_{\eta,i})$ with $i = 1, \ldots, m$ into $G_c$ is the inclusion. The same holds for $d \neq v_f$. For $v_f = d$, the edge-to-vertex group maps for $\pi_1(\Sigma_{\eta,i})$ with $i = 1, \ldots, m$ into $G_d$ is given by inclusion followed by $\phi_*$, the map induced by the monodromy $\phi$.

It remains to identify the edges and edge groups for the underlying graph $\mathcal{G}_0$. The edges are of exactly two kinds: undrilled edges, and subdivision edges. The edge groups for undrilled edges correspond to $H$. The edge groups for subdivision edges correspond to the subdivision edge groups $\pi_1(\Sigma_{\eta,i})$ for $i = 1, \ldots, m$. Finally, the edge-to-vertex group maps are given as above.

**Definition 5.1** A maximal connected undrilled subgraph $\mathcal{K}$ of $\mathcal{G}$ will be called an *undrilled component of $\mathcal{G}$*.

We finally modify $\mathcal{G}_0$ to another graph $\mathcal{G}_{\mathbb{K}}$ by collapsing each undrilled component (see Definition 5.1) $\mathcal{K}$ of $\mathcal{G}$ to a single vertex, ie $\mathcal{G}_{\mathbb{K}}$ is the quotient space obtained from $\mathcal{G}_0$ under the equivalence relation $x \sim y$ if and only if $x$ and $y$ belong to the same undrilled component (see Definition 5.1) $\mathcal{K}$ of $\mathcal{G}$. Let $v_{\mathcal{K}}$ be the resulting vertex of $\mathcal{G}_{\mathcal{K}}$.

The vertex space associated to $v_{\mathcal{K}}$ is then declared to be $\mathcal{F}_{\mathcal{K}}$. The edge to vertex inclusions are given by the composition of

(1) edge to vertex inclusion maps over $\mathcal{G}_0$, composed with

(2) inclusions of vertex spaces over $v \in \mathcal{G}_0$ to the vertex spaces over $v_{\mathcal{K}} \in \mathcal{G}_{\mathbb{K}}$, where $\mathcal{K}$ is the undrilled component (see Definition 5.1) of $\mathcal{G}$ containing $v$.

**Definition 5.2** $\mathcal{G}_{\mathcal{K}}$ will be called the *canonical reduced form* of $\mathcal{G}$ for $F$.

We conclude this subsection with the following observation, which follows from the above construction.

**Lemma 5.3** *The edge spaces of $\mathcal{G}_{\mathcal{K}}$ are precisely the components of $S_d \setminus \bigcup_i \sigma_i$, where $S_d$ ranges over a reduced collection of drilled fibers* (see Definition 4.18).

## 5.2 Cubulating drilled bundles

We shall need the following theorem due to Wise:

**Theorem 5.4** [34, Theorem 15.1] *Let $G$ be a group satisfying the following*:

(1) *$G$ is hyperbolic relative to virtually abelian subgroups.*

(2) *$G$ splits as a graph of groups $\Gamma$ where each edge group is relatively quasiconvex in $G$.*

(3) *Each vertex group is virtually special.*

(4) *For each edge $e$, the edge group $G_e$ has trivial intersection with each $Z^2$ in the fundamental group of the graph of groups $\Gamma - e$.*

*Then $G$ is the fundamental group of a virtually special cube complex.*

For us, $G = \pi_1(F)$, and let $\Pi_0 \colon F \to \mathcal{G}$ denote the natural projection.

**Theorem 5.5** *Suppose that the graph $\mathcal{G}_{\mathbb{K}}$ is the canonical reduced form of $\mathcal{G}$. Suppose further that for every undrilled component $\mathcal{K}$ (see Definition 5.1) of $\mathcal{G}$, $\pi_1(\mathcal{F}_{\mathcal{K}})$ is virtually special cubulable. Then the group $G = \pi_1(F)$ is virtually special cubulable.*

**Proof** Theorem 3.1 proves that $\widetilde{F}$ is strongly hyperbolic relative to $\mathcal{P}$. Since the stabilizer of each $P \in \mathcal{P}$ is $\mathbb{Z} + \mathbb{Z}$, Theorem 5.4(1) is satisfied.

By Lemma 5.3, the edge spaces over $\mathcal{G}_{\mathcal{K}}$ are given by the components of $S_d \setminus \bigcup_i \sigma_i$, where $S_d$ ranges over a reduced collection of drilled fibers (see Definition 4.18). The fundamental groups of these components are relatively quasiconvex in $G$ by Proposition 4.21. Hence Theorem 5.4(2) is satisfied.

Theorem 5.4(3) follows from the hypothesis that $\pi_1(\mathcal{F}_{\mathcal{K}})$ is virtually special cubulable.

Theorem 5.4(4) follows from the hypothesis that $\mathcal{G}_{\mathbb{K}}$ is the canonical reduced form (see Definition 5.2) of $\mathcal{G}$. Indeed, this hypothesis guarantees that there are no accidental parabolics in the components $\Sigma$ of $S_d \setminus \bigcup_i \sigma_i$, where $S_d$ ranges over a reduced collection of drilled fibers (see Definition 4.18), ie no nonperipheral essential curve in any $\Sigma$ is freely homotopic to a drilled curve.

Hence, by Theorem 5.4, $G$ is virtually special cubulable. $\qquad\qquad\square$

## 5.3 Examples

We now give examples of surface bundles $E$ over graphs $\mathcal{G}$ (see Definition 1.2) and drilled curves such that the hypotheses of Theorem 5.5 are satisfied:

**Example 5.6** Each edge of $\mathcal{G}$ contains a drilled point. In this case, the undrilled components $\mathcal{K}$ are precisely the vertices of $\mathcal{G}$, and the associated spaces are given by the fiber $S$. Since $\pi_1(S)$ is special cubulable, the hypotheses of Theorem 5.5 are satisfied.

**Example 5.7** Undrilled components $\mathcal{K}$ are either non-self-intersecting loops in $\mathcal{G}$ or vertices. The associated vertex spaces are either hyperbolic 3-manifolds $M$ fibering over the circle, or the fiber $S$. By [1], $\pi_1(M)$ is virtually special and hence the hypotheses of Theorem 5.5 are satisfied. More generally, undrilled components $\mathcal{K}$ could be homotopy equivalent to circles or contractible.

A simple example for Example 5.7 is given by a graph $\mathcal{G}$ with two vertices $v_1$ and $v_2$, an edge $[v_1, v_2]$, and a loop at each vertex. Further, the edge $[v_1, v_2]$ has a single drilled fiber $S_w$ with one simple closed curve $\sigma \subset S_w$ drilled.

**Example 5.8** Undrilled components $\mathcal{K}$ are homotopy equivalent to a wedge of circles, and the restriction of $E$ over any such $\mathcal{K}$ are examples from [24]. The main theorem of [24] then guarantees that the hypotheses of Theorem 5.5 are satisfied.

More generally, components $\mathcal{K}$ could be a mixture of these cases, ie they could be

(1) contractible, in which case the associated vertex space is homotopy equivalent to $S$,

(2) homotopy equivalent to a circle, in which case the associated vertex space is homotopy equivalent to hyperbolic 3-manifolds $M$ fibering over the circle,

(3) homotopy equivalent to a wedge of circles, with the associated vertex space homotopy equivalent to one of the examples from [24].

# 6 Virtual algebraic fibering

Notation to be used in this section:

- $K_\eta$ denotes a homotoped copy of an elementary drilled atom $F_\eta$, being the union of one torus $T_i$ glued to $S$ along $\sigma_i$ each for every drilled curve $\sigma_i$ of $F_\eta$.

**Definition 6.1** A finitely generated group $G$ is said to *virtually algebraically fiber* if there exists a finite-index subgroup $G_1$ of $G$ such that $G_1$ admits a surjective homomorphism to $\mathbb{Z}$ with finitely generated kernel.

A theorem of Kielak [20] gives the following criterion for virtual algebraic fibering.

**Theorem 6.2** [20] *Let $G$ admit a geometric action on a* CAT(0) *cube complex. Then the following are equivalent*:

(1) *$G$ virtually algebraically fibers.*

(2) *the first $l^2$ Betti number $\beta_1^{(2)}(G)$ equals zero.*

## 6.1 Vanishing first $l^2$ Betti number

The aim of this subsection is to show:

**Proposition 6.3** *Let $F$ be a drilled surface bundle over a finite graph $\mathcal{G}$ (see Definition 1.2). Let $G = \pi_1(F)$. Then $\beta_1^{(2)}(G) = 0$.*

As an immediate consequence of Proposition 6.3 and Theorem 6.2, we have the following:

**Theorem 6.4** *Let $F$ be a drilled surface bundle over a finite graph $\mathcal{G}$ (see Definition 1.2) satisfying the hypotheses of Theorem 5.5. Then $G = \pi_1(F)$ virtually algebraically fibers.*

To prove Proposition 6.3, we shall need a couple of results. A fundamental theorem of Lott and Lück gives the following:

**Theorem 6.5** [21, Theorem 0.1] *Let $M_r$ be a drilled atom of $F$ (see Definition 3.2), and $H_r = \pi_1(M_r)$. Then $\beta_1^{(2)}(H_r) = -\chi(M_r)$.*
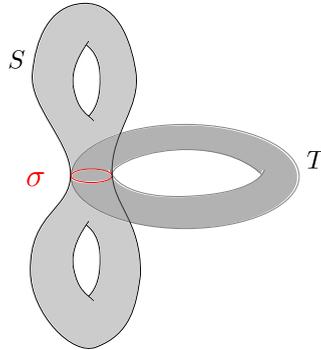
Figure 14: A homotoped elementary drilled atom.

We shall also use the following theorem of Fernós and Valette:

**Theorem 6.6** [12, Theorem 1.1] *Let $G$ be a graph of groups with at least one edge, such that all vertex groups satisfy $\beta_1^{(2)}(G_v) = 0$. If every edge group is infinite, and for every edge-to-vertex group inclusion the edge group is of infinite index in the vertex group, then $\beta_1^{(2)}(G) = 0$.*

Recall from Section 5.1 that if $\eta$ is a subdivision interval, the restriction $F_\eta$ of $F$ to $\eta$ has a simple topology: $F_\eta$ is homeomorphic to $S \times [0,1]$ after removing $\epsilon$-neighborhoods of a nonempty family $\{\sigma_i\}$ of disjoint homotopically distinct essential simple closed curves on a unique drilled fiber $S_x$. In this subsection, we shall refer to such an $F_\eta$ as an *elementary drilled atom*.

Let $\sigma_1, \dots, \sigma_k \subset S_x$ denote the drilled curves. Let $\boldsymbol{T}_1, \dots, \boldsymbol{T}_k$ denote $k$-copies of the standard torus $S^1 \times S^1$. Also, let $K_\eta$ denote the 2-complex obtained as a quotient space of $S_x \sqcup \bigcup_i \boldsymbol{T}_i$ by identifying $S^1 \times \{0\} \subset \boldsymbol{T}_i$ with $\sigma_i$ via a homeomorphism (see Figure 14). Then $F_\eta$ is homotopy equivalent to $K_\eta$. Further, $\pi_1(F_\eta) = \pi_1(K_\eta)$ has a graph of groups description, where the underlying graph $\mathcal{G}$ has $k+1$ vertices, $0, \dots, k$ say, and

(1) $\mathcal{G}$ is a tree with one root vertex 0, and all other vertices $1, \dots, k$ are connected to 0 by an edge each,

(2) the vertex group $G_0$ equals $\pi_1(S_x)$,

(3) for $i = 1, \dots, k$, each vertex group $G_i$ equals $\mathbb{Z} \oplus \mathbb{Z}$,

(4) each edge group is $\mathbb{Z}$.

**Lemma 6.7** *Let $L$ be a group with $\beta_1^{(2)}(L) = 0$. Let $H \subset L$ be a subgroup isomorphic to $\pi_1(S_x)$. Let $G = L *_H \pi_1(K_\eta)$, where $H$ is identified via an automorphism with $G_0 = \pi_1(S_x)$. Then $\beta_1^{(2)}(G) = 0$.*

**Proof** $G$ admits a graph of groups description, where the underlying graph $\mathcal{G}$ has $k+1$ vertices, $0, \dots, k$ say, and

(1) $\mathcal{G}$ is a tree with one root vertex 0, and all other vertices $1, \dots, k$ are connected to 0 by an edge each,

(2)  the vertex group $G_0$ equals $L$,

(3)  for $i = 1, \ldots, k$, each vertex group $G_i$ equals $\mathbb{Z} \oplus \mathbb{Z}$,

(4)  each edge group is $\mathbb{Z}$.

Then $G$ satisfies the hypotheses of Theorem 6.6. Hence, $\beta_1^{(2)}(G) = 0$.   $\square$

**Lemma 6.8**  *Let $F$ be a drilled surface bundle* (see Definition 1.2) *over a finite graph $\mathcal{G}$ such that $\mathcal{G}$ is homotopy equivalent to a circle. Let $G = \pi_1(F)$. Then $\beta_1^{(2)}(G) = 0$.*

**Proof**  Let $\mathcal{C} \subset \mathcal{G}$ be a cycle with no repeated vertices. Then $\mathcal{G} = \mathcal{C} \cup \bigcup_{i=1,\ldots,k} \mathcal{T}_i$, where

(1)  each $\mathcal{T}_i$ is a finite tree

(2)  $\mathcal{T}_i$ intersects $\mathcal{C}$ at a single point $p_i$,

(3)  $\mathcal{T}_i \setminus \{p_i\}$ is disjoint from $\bigcup_{j \neq i} \mathcal{T}_j$.

Each edge of $\mathcal{T}_i$ for $i = 1, \ldots, k$, can be subdivided as in Section 5.1, so that each edge is a subdivision edge. In particular, after such subdivision, for every edge $\eta$ of $\mathcal{T}_i$ for $i = 1, \ldots, k$, $F_\eta$ is an elementary drilled atom.

Let $F_{\mathcal{C}}$ denote the restriction of $F$ to the cycle $\mathcal{C}$. Then $F_{\mathcal{C}}$ is a 3-manifold whose boundary components are all tori. Hence $\chi(F_{\mathcal{C}}) = 0$. By Theorem 6.5, $\beta_1^{(2)}(\pi_1(F_{\mathcal{C}})) = 0$.

Proceeding by induction on the number of elementary drilled atoms in $F$, and applying Theorem 6.6 inductively, we conclude that $\beta_1^{(2)}(G) = 0$.   $\square$

**Proof of Proposition 6.3**  The first part of the argument in Section 5.1 allows us to modify $\mathcal{G}$ to a graph where each edge group is isomorphic to $\pi_1(S)$. Thus, assume without loss of generality that each edge in $\mathcal{G}$ has edge group isomorphic to $\pi_1(S)$.

Let $\mathcal{G}_0$ denote a maximal subgraph of $\mathcal{G}$ such that $\mathcal{G}_0$ is homotopy equivalent to a circle. Let $F_0$ denote the restriction of $F$ to $\mathcal{G}_0$. Let $L = \pi_1(F_0)$. Let $n$ denote the number of edges of $(\mathcal{G} \setminus \mathcal{G}_0)$. Then $G = \pi_1(F)$ admits a new graph of groups decomposition, where the base graph has one vertex $w$ and $n$ loops, with $G_w = L$, and each edge group isomorphic to $\pi_1(S)$.

By Lemma 6.8, $\beta_1^{(2)}(L) = 0$. Since $\mathcal{G}_0$ is homotopy equivalent to a circle, each edge group is of infinite index in the vertex group $L$. Hence, by Lemma 6.7, $\beta_1^{(2)}(G) = 0$.   $\square$

## 6.2  Questions

In this paper, we have only drilled simple closed curves $\sigma$ in fibers $S_x$. A similar drilling operation could be carried out even when a realization of $\sigma$ as a geodesic in the fiber $S$ has self-intersections. This can be done by homotoping $\sigma$ slightly in a product neighborhood of $S_x$ to convert $\sigma$ into a knot in $E$.

Note that this is a noncanonical operation, as we have a choice of over- and under-crossings at every self-intersection point. At any rate, after such a homotopy, the resulting knot can be drilled. We assume, however, that $\sigma$ represents a primitive element of $H = \pi_1(S)$. We call this process *generalized drilling*, and the resulting $F$ a *generalized drilled bundle*. Relative hyperbolicity of the generalized drilled $G$ follows by essentially the same argument as in Theorem 3.1.

**Question 6.9** *Let* $\Gamma = \pi_1(E)$ *be hyperbolic. Let* $F$ *be a generalized drilled bundle obtained by the above generalized drilling operation applied to* $E$. *Is* $G(:= \pi_1(F))$ *cubulable?*

## Acknowledgments

## References

[1]  **I Agol**, *The virtual Haken conjecture*, Doc. Math. 18 (2013) 1045–1087  MR

[2]  **I Belegradek**, *Rigidity and relative hyperbolicity of real hyperbolic hyperplane complements*, Pure Appl. Math. Q. 8 (2012) 15–51  MR

[3]  **I Belegradek**, **G C Hruska**, *Hyperplane arrangements in negatively curved manifolds and relative hyperbolicity*, Groups Geom. Dyn. 7 (2013) 13–38  MR

[4]  **N Bergeron**, **D T Wise**, *A boundary criterion for cubulation*, Amer. J. Math. 134 (2012) 843–859  MR

[5]  **M Bestvina**, **M Feighn**, *A combination theorem for negatively curved groups*, J. Differential Geom. 35 (1992) 85–101  MR  Correction in 43 (1996) 783–788

[6]  **B H Bowditch**, *Intersection numbers and the hyperbolicity of the curve complex*, J. Reine Angew. Math. 598 (2006) 105–129  MR

[7]  **B H Bowditch**, *Relatively hyperbolic groups*, Int. J. Algebra Comput. 22 (2012) art. id. 1250016  MR

[8]  **R D Canary**, *A covering theorem for hyperbolic 3-manifolds and its applications*, Topology 35 (1996) 751–778  MR

[9]  **S Dowdall**, **R P Kent, IV**, **C J Leininger**, *Pseudo-Anosov subgroups of fibered 3-manifold groups*, Groups Geom. Dyn. 8 (2014) 1247–1282  MR

[10]  **B Farb**, *Relatively hyperbolic groups*, Geom. Funct. Anal. 8 (1998) 810–840  MR

[11]   **B Farb**, **L Mosher**, *Convex cocompact subgroups of mapping class groups*, Geom. Topol. 6 (2002) 91–152 MR

[12]   **T Fernós**, **A Valette**, *The Mayer–Vietoris sequence for graphs of groups, property (T), and the first $\ell^2$-Betti number*, Homology Homotopy Appl. 19 (2017) 251–274   MR

[13]   **D Groves**, **P Haïssinsky**, **J F Manning**, **D Osajda**, **A Sisto**, **G S Walsh**, *Drilling hyperbolic groups*, preprint (2024)   arXiv 2406.14667

[14]   **F Haglund**, **D T Wise**, *Special cube complexes*, Geom. Funct. Anal. 17 (2008) 1551–1620   MR

[15]   **U Hamenstädt**, *Word hyperbolic extensions of surface groups*, preprint (2005)   arXiv math/0505244

[16]   **U Hamenstädt**, *Geometry of the complex of curves and of Teichmüller space*, from "Handbook of Teichmüller theory, I", IRMA Lect. Math. Theor. Phys. 11, Eur. Math. Soc., Zürich (2007) 447–467   MR

[17]   **J Kahn**, **V Markovic**, *Immersing almost geodesic surfaces in a closed hyperbolic three manifold*, Ann. of Math. 175 (2012) 1127–1190   MR

[18]   **M Kapovich**, *Hyperbolic manifolds and discrete groups*, Progr. Math. 183, Birkhäuser, Boston, MA (2001) MR

[19]   **R P Kent, IV**, **C J Leininger**, *Shadows of mapping class groups: capturing convex cocompactness*, Geom. Funct. Anal. 18 (2008) 1270–1325   MR

[20]   **D Kielak**, *Residually finite rationally solvable groups and virtual fibring*, J. Amer. Math. Soc. 33 (2020) 451–486   MR

[21]   **J Lott**, **W Lück**, *$L^2$-topological invariants of 3-manifolds*, Invent. Math. 120 (1995) 15–60   MR

[22]   **M Mitra**, *Cannon–Thurston maps for trees of hyperbolic metric spaces*, J. Differential Geom. 48 (1998) 135–164   MR

[23]   **M Mj**, *Tight trees and model geometries of surface bundles over graphs*, J. Lond. Math. Soc. 102 (2020) 1178–1222   MR

[24]   **M Mj**, *Cubulating surface-by-free groups*, J. Topol. 17 (2024) art. id. e70011   MR   With an appendix by J Manning, M Mj, and M Sageev

[25]   **M Mj**, **A Pal**, *Relative hyperbolicity, trees of spaces and Cannon–Thurston maps*, Geom. Dedicata 151 (2011) 59–78   MR

[26]   **M Mj**, **K Rafi**, *Algebraic ending laminations and quasiconvexity*, Algebr. Geom. Topol. 18 (2018) 1883–1916   MR

[27]   **M Mj**, **L Reeves**, *A combination theorem for strong relative hyperbolicity*, Geom. Topol. 12 (2008) 1777–1798   MR

[28]   **M Mj**, **P Sardar**, *A combination theorem for metric bundles*, Geom. Funct. Anal. 22 (2012) 1636–1707 MR

[29]   **G P Scott**, **G A Swarup**, *Geometric finiteness of certain Kleinian groups*, Proc. Amer. Math. Soc. 109 (1990) 765–768   MR

[30]   **A Sisto**, *On metric relative hyperbolicity*, preprint (2012)   arXiv 1210.8081

[31]   **P Susskind**, **G A Swarup**, *Limit sets of geometrically finite hyperbolic groups*, Amer. J. Math. 114 (1992) 233–250   MR

[32]  **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at `https://url.msp.org/gt3m`

[33]  **D T Wise**, *From riches to raags:* 3-*manifolds, right-angled Artin groups, and cubical geometry*, CBMS Region. Conf. Ser. Math. 117, Amer. Math. Soc., Providence, RI (2012)  MR

[34]  **D T Wise**, *The structure of groups with a quasiconvex hierarchy*, Ann. of Math. Stud. 209, Princeton Univ. Press (2021)  MR

MM, BN:  *School of Mathematics, Tata Institute of Fundamental Research*
*Mumbai, India*

mahan@math.tifr.res.in,  biswajit@math.tifr.res.in

# Infinitely many homeomorphic hyperbolic plugs with the same basic sets

FANGFANG CHEN

Using Ghys' work on the geodesic flow of the modular surface, we construct infinitely many smooth flows on $T^2 \times [0, 1]$ transverse to the boundary, whose maximal invariant sets are the same saddle hyperbolic set, but whose global dynamics are not topologically equivalent pairwise. As a corollary of the previous construction, we prove that there exist infinitely many simple Smale flows on $S^3$ with the same basic sets but different global dynamics.

## 1 Introduction

### 1.1 Background and main results

A *hyperbolic plug* is a pair $(W, \psi_t)$, where $W$ is a compact 3-manifold with boundary and $\psi_t$ is a smooth flow on $W$ transverse to $\partial W$ such that the maximal invariant set of $\psi_t$ is a saddle hyperbolic set. Hyperbolic plugs play a fundamental role in the study of structurally stable flows in dimension 3 as they can be used to construct large families of structurally stable flows and also to decompose complex structurally stable flows into simpler ones that can be studied separately.

In [3], Béguin, Bonatti and Yu constructed transitive Anosov flows in dimension 3 by gluing hyperbolic plugs with filling laminations. Here, a *hyperbolic plug with filling laminations* is a hyperbolic plug $(W, \psi_t)$ such that for the maximal invariant set $\Lambda$ of $\psi_t$, the lamination $\partial W \cap (W^s(\Lambda) \cup W^u(\Lambda))$ can be filled into a foliation on $\partial W$. In this context, $W^s(\Lambda)$ (resp $W^u(\Lambda)$) denotes the union of stable (resp unstable) manifolds of the orbits in $\Lambda$. Therefore, exploring the topologically equivalent classes of homeomorphic hyperbolic plugs with filling laminations could provide an answer to the classical question of whether a 3-manifold can admit infinitely many Anosov flows up to topological equivalence.

Nowadays, many experts in the area of 3-dimensional Anosov flows tend to believe that any 3-manifold admits at most finitely many Anosov flows, up to topological equivalence. So, it is highly likely that:

**Conjecture 1.1** *There are at most finitely many homeomorphic hyperbolic plugs with filling laminations, any two of which are not topologically equivalent.*

However, for the hyperbolic plugs without filling laminations, our result (Theorem 1.3), contrary to the claims of Conjecture 1.1, shows that there are infinitely many homeomorphic hyperbolic plugs without filling laminations, whose maximal invariant sets (even with their small neighborhoods) are the same, but whose global dynamics are not topologically equivalent pairwise.
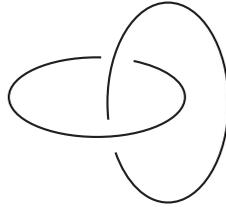
Figure 1: Hopf link.

**Definition 1.2** Let $\varphi_t$ (resp $\phi_t$) be a smooth flow on a 3-manifold $M$ (resp $N$) with a compact invariant set $K$ (resp $L$). The pair $(\varphi_t, K)$ is *equivalent* to $(\phi_t, L)$ if and only if there is a neighborhood $U_K$ of $K$ in $M$ and a neighborhood $U_L$ of $L$ in $N$ such that $\varphi_t|_{U_K}$ and $\phi_t|_{U_L}$ are topologically equivalent via a homeomorphism sending $K$ to $L$. The *germ* $[\varphi_t, K]$ is the equivalence class represented by $(\varphi_t, K)$.

**Theorem 1.3** *There are infinitely many homeomorphic hyperbolic plugs $\{(T^2 \times [0, 1], \psi_t^i)\}_{i \in \tau}$, such that*

(1) *for any $i \in \tau$, the maximal invariant set $\Lambda_i$ of $\psi_t^i$ is a 1-dimensional saddle basic set containing infinitely many closed orbits;*

(2) *for any $j \in \tau$ and $j \neq i$, $[\psi_t^i, \Lambda_i] = [\psi_t^j, \Lambda_j]$, but $\psi_t^i$ is not topologically equivalent to $\psi_t^j$.*

Building on Theorem 1.3, we can explore its implications for Smale flows on $S^3$. A *Smale flow* is a smooth structurally stable flow, whose chain recurrent set is hyperbolic, at most 1-dimensional and satisfies the transversality condition [19]. Sullivan introduced a special class of Smale flows in [20], called *simple Smale flows*. A simple Smale flow is a Smale flow whose chain recurrent set consists of an attracting closed orbit, a repelling closed orbit and a *nontrivial* saddle basic set: a saddle basic set containing infinitely many closed orbits. In [25], Yu showed that every closed orientable 3-manifold admits a simple Smale flow.

Previous studies on simple Smale flows are focused on the templates related to saddle basic sets, including [1; 13; 20; 21; 22] by Adhikari, Haynes, Sullivan, and Yu. Utilizing Theorem 1.3, we can approach this topic from a different perspective (the germs of saddle basic sets) and obtain a surprising result (Corollary 1.4): there exist infinitely many simple Smale flows on $S^3$ with the same basic sets but different global dynamics.

**Corollary 1.4** *There are infinitely many simple Smale flows $\{\phi_t^i\}_{i \in \tau}$ on $S^3$ up to topological equivalence, such that*

(1) *for any $i \in \tau$, $A_i \sqcup R_i$ is a Hopf link (see Figure 1);*

(2) *for any $j \in \tau$, $[\phi_t^i, \Lambda_i] = [\phi_t^j, \Lambda_j]$.*

*Here, $A_i$, $R_i$ and $\Lambda_i$ are the attracting closed orbit, the repelling closed orbit and the saddle basic set of $\phi_t^i$, respectively.*

**Remark 1.5** A paper in preparation by Fan, Lai, and Yu shows that the Whitehead link exterior cannot be the background manifold of a hyperbolic plug (personal communication, 2024). Therefore, the Whitehead link cannot be realized as the attractor and repeller of a simple Smale flow on $S^3$.

**Remark 1.6** If we replace the word "nontrivial" in the definition of simple Smale flows with "trivial," then simple Smale flows will transform into Morse–Smale flows with three basic sets. For certain classes of 3-manifolds, complete classifications of Morse–Smale systems have already been achieved, for instance:

- In [24], Yu proved that up to topological equivalence, there are at most a finite number of nonsingular Morse–Smale flows on $S^3$ with three closed orbits. However, due to the complexity of heteroclinic trajectories, there are infinitely many nonsingular Morse–Smale flows on $S^3$ with four closed orbits up to topological equivalence.

- Pochinka and Shubin introduced an invariant to determine the topological equivalence classes of nonsingular Morse–Smale flows on orientable 3-manifolds with three closed orbits [18].

- The problem of a topological classification of Morse–Smale cascades on 3-manifolds either without heteroclinic points or without heteroclinic curves was solved in [5; 6; 7; 8; 10] by Bonatti, Grines, Laudenbach, Medvedev, Pécou, and Pochinka.

## 1.2 The key ideas

At the 2006 International Congress of Mathematicians, Ghys constructed a compact hyperbolic orbifold $\Sigma$ by deforming the modular surface [11]. Bonatti and Pinsky extended the geodesic flow $g_t$ on the unit tangent bundle $T^1\Sigma$ of $\Sigma$ into a Smale flow $\phi_t$ on $S^3$ [9]. Let $\Lambda$ be the saddle basic set of $\phi_t$ related to $g_t$. A hyperbolic plug $(N, \eta_t)$ is called a *hyperbolic plug for* $\Lambda$ if the germ of $\eta_t$ along the maximal invariant set coincides with the germ $[\phi_t, \Lambda]$. The key idea of this paper is to construct infinitely many homeomorphic hyperbolic plugs for $\Lambda$, such that no two plugs can be topologically equivalent.

In [9], Bonatti and Pinsky constructed a hyperbolic plug $(W, \psi_t)$ for $\Lambda$ where $W \cong S^2 \times [0, 1]$. Let $\Lambda_0$ be the maximal invariant set of $\psi_t$ and $\partial^+ W$ be the connected components of $\partial W$ on which $\psi_t$ points inward. We will prove that $\partial^+ W \setminus W^s(\Lambda_0)$ consists of infinitely many connected components. By performing a surgery-type operation along the connected components of $\partial^+ W \setminus W^s(\Lambda_0)$, we will produce infinitely many homeomorphic hyperbolic plugs for $\Lambda$. By using hyperbolic geometry, we will prove that the topological equivalence of these plugs distinguishes between the different connected components of $\partial^+ W \setminus W^s(\Lambda_0)$ on which we performed surgery, thus proving Theorem 1.3.

## 1.3 Organization of the paper

In Section 2, we introduce some definitions and elementary properties of models for saddle basic sets. In Section 3, we define Ghys' compact orbifold $\Sigma$ from the modular surface and discuss the geodesic flow $g_t$ on the unit tangent bundle $T^1\Sigma$ of $\Sigma$. In Section 4, we discuss the mapping class group of $T^1\Sigma$. In Section 5, we prove Theorem 1.3. In Section 6, we introduce some interesting open questions.

## 2 Preliminaries

### 2.1 Models and templates

Let $\phi_t$ be a Smale flow on a closed orientable 3-manifold and $\Lambda$ be a nontrivial saddle basic set of $\phi_t$.

**Definition 2.1** A *hyperbolic plug for* $\Lambda$ is a hyperbolic plug $(W, \psi_t)$ such that $[\psi_t, \Lambda_0] = [\phi_t, \Lambda]$ where $\Lambda_0$ is the maximal invariant set of $\psi_t$.

To describe the dynamical behavior of $\Lambda$, Béguin and Bonatti defined the following concept in [2].

**Definition 2.2** A *model* for $\Lambda$ is a hyperbolic plug $(W, \psi_t)$ for $\Lambda$, such that if $\Lambda_0$ is the maximal invariant set of $\psi_t$ then

(1) any embedded circle in $\partial^+ W \setminus W^s(\Lambda_0)$ bounds a disk in $\partial^+ W \setminus W^s(\Lambda_0)$, where $\partial^+ W$ denotes the connected components of $\partial W$ on which $\psi_t$ points inward;

(2) any connected component of $W$ contains at least one point of $\Lambda_0$.

The following theorem was proved by Béguin and Bonatti in [2, Theorem 0.3].

**Theorem 2.3** *There exists a unique model for* $\Lambda$ *up to topological equivalence.*

Let $(W, \psi_t)$ be a model for $\Lambda$ and $\Lambda_0$ be the maximal invariant set of $\psi_t$. In [2], Béguin and Bonatti introduced the following method to construct hyperbolic plugs for $\Lambda$. Let $D_1$ and $D_2$ be two disjoint closed disks in $\partial^+ W \setminus W^s(\Lambda_0)$. Since the orbit of any point in $\partial^+ W \setminus W^s(\Lambda_0)$ goes from the entrance to the exit boundary, the orbits of $D_1$, $D_2$ form two cylinders $D^2 \times [0, 1]$ endowed with the vector field $\frac{\partial}{\partial t}$. Remove these two cylinders from $W$ and then glue the two resulting tangent boundary annuli together by preserving flowlines. By doing so, we obtain a new hyperbolic plug for $\Lambda$ (see Figure 2). This dynamical surgery is called *handle attachment*.
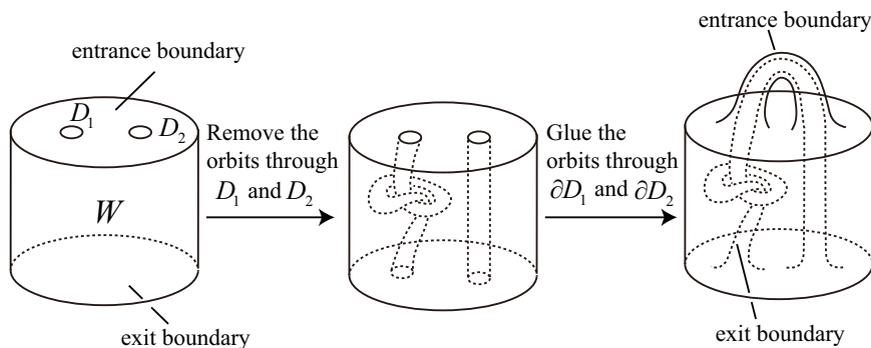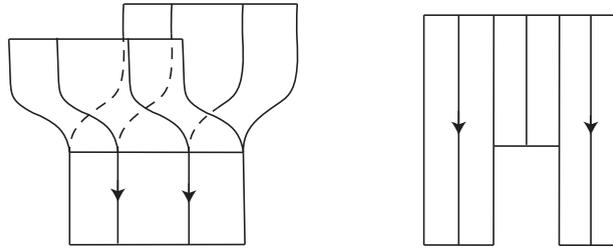


Figure 2: Handle attachment.

Figure 3: Joining and splitting charts.

**Definition 2.4** A *template* is a compact branched 2-manifold with boundary and smooth expansive semiflow built locally from two types of charts: *joining* and *splitting* (see Figure 3). The gluing maps between charts must respect the semiflow and act linearly on the edges.

In [4], Birman and Williams proved that collapsing the strong stable manifolds of a suitable neighborhood of $\Lambda$ yields a template. By reversing the construction of Birman and Williams, we extend a template $T$ in the direction perpendicular to its surface to obtain a *thickened template* $\overline{T}$. Since the semiflow on $T$ is expanding, we can extend the semiflow on each chart, and produce thickened charts as in Figure 4. Figure 5 provides a sectional view of thickened charts. See Meleshuk [15] for more details on thickened templates.

A *dividing curve* of $\overline{T}$ is a closed curve $c$ in $\partial\overline{T}$ such that the natural flow on $\overline{T}$ is tangent to $\partial\overline{T}$ at the curve $c$. For the thickened charts in Figure 4, the dividing curves correspond to the bold curves. We can attach a 2-handle to a dividing curve $c$ of $\overline{T}$ as follows: First, we blow up $c$ to obtain $c \times [0, 1]$ and endow $c \times [0, 1]$ with the vector field $\frac{\partial}{\partial t}$. Next, we glue a 2-handle $D^2 \times [0, 1]$ with the vector field $\frac{\partial}{\partial t}$ to $c \times [0, 1]$ by
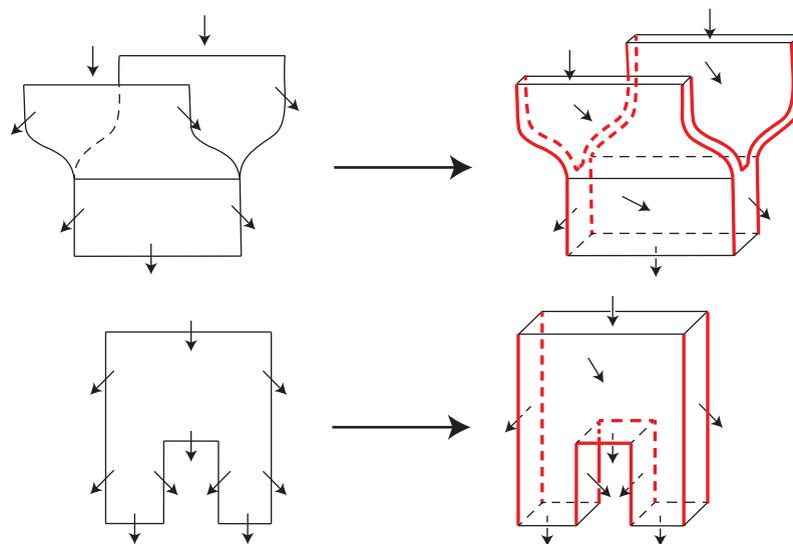


Figure 4: Thickened charts. Top: joining chart thickening. Bottom: splitting chart thickening.
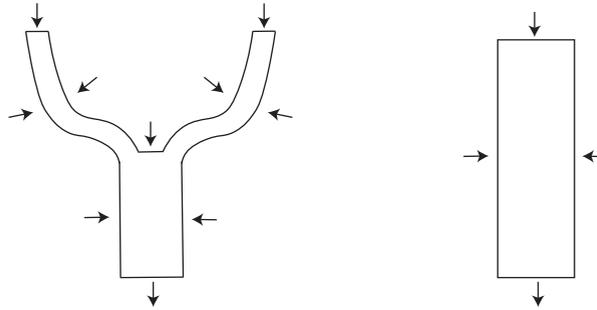
Figure 5: Sectional view of thickened charts.

preserving the flowlines. Figure 6 provides a sectional view of the attachment of a 2-handle. The following theorem of Yu [23, Theorem 4.3], indicates the relationship between thickened templates and models.

**Theorem 2.5**  Let $\overline{T}$ be a thickened template related to $\Lambda$. Then up to topological equivalence, the model for $\Lambda$ can be obtained by attaching a 2-handle to each dividing curve of $\overline{T}$.

## 2.2  The entrance boundary of models

Suppose that $\Lambda$ is a nontrivial saddle basic set of a Smale flow $\phi_t$ on a closed orientable 3-manifold. Let $(W, \psi_t)$ be a model for $\Lambda$ and $\Lambda_0$ be the maximal invariant set of $\psi_t$.

**Definition 2.6**  A *free stable separatrix* of a closed orbit $O \subset \Lambda_0$ is a connected component of $W^s(O) \setminus O$ that is disjoint from $\Lambda_0$.

The following lemma was proved by Béguin, Bonatti, and Yu in [3, Remark 3.5 and Proposition 3.8].

**Lemma 2.7**  (1)  *For each connected component $C$ of $W^s(\Lambda_0) \setminus \Lambda_0$, there are two possible situations*:

- *Either $C$ is a free stable separatrix of some closed orbit. Then $C \cap \partial^+ W$ is homeomorphic to a circle.*

- *Or $C$ is bounded by two orbits in $\Lambda_0$. Then $C \cap \partial^+ W$ is homeomorphic to a line.*

(2)  *In $W^s(\Lambda_0) \cap \partial^+ W$, each half line is asymptotic to a circle.*
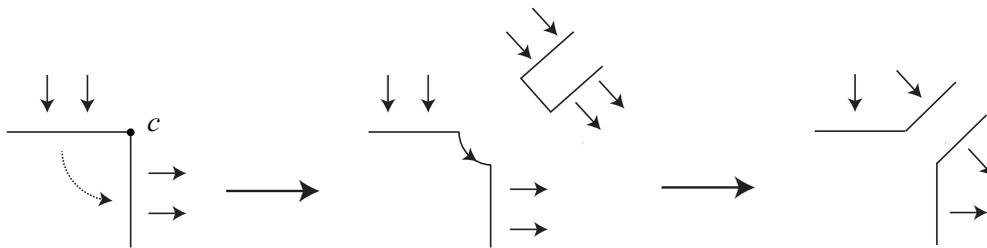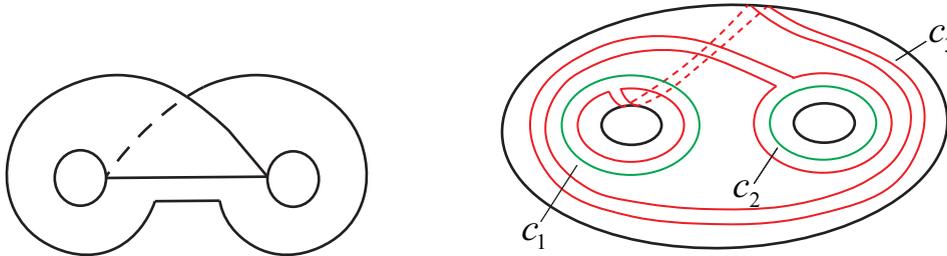


Figure 6: Attaching a 2-handle to a dividing curve $c$.

Figure 7: Left: Lorenz template. Right: thickened Lorenz template.

In this paper, a 2-manifold homeomorphic to $\mathbb{R}^2$ is called an *open disk* if its accessible boundary is a circle. A 2-manifold homeomorphic to $\mathbb{R}^2$ is called a *strip* if its accessible boundary consists of exactly two lines which are asymptotic to each other at both ends.

**Proposition 2.8** *Suppose that the template related to $\Lambda$ is a Lorenz template $T$ (see Figure 7). Then*

(1) $W \cong S^2 \times [0, 1]$;

(2) $\partial^+ W \cap W^s(\Lambda_0)$ *consists of two circles and infinitely many lines*;

(3) $\partial^+ W \setminus W^s(\Lambda_0)$ *consists of two open disks and infinitely many strips.*

**Proof** The thickened Lorenz template $\overline{T}$ has three dividing curves $c_1, c_2, c_3$, as shown in Figure 7, right. By Theorem 2.5, we obtain a model $(W^T, \psi_t^T)$ from $\overline{T}$ by attaching three 2-handles to the dividing curves. By attaching two 2-handles to $c_1, c_2$, the 3-ball is obtained. Thus, $W^T$ is homeomorphic to $S^2 \times [0, 1]$.

Let $\Lambda_0^T$ be the maximal invariant set of $\psi_t^T$. Bonatti and Pinsky showed in [9] that $\Lambda_0^T$ has a section with a first return map $\Gamma$, called the fake horseshoe map (see Figure 8). The map $\Gamma$ was introduced by Smale in [19]. Utilizing Markov partitions and symbolic dynamics, we find that there are only two closed orbits in $\Lambda_0^T$ admitting a free stable separatrix, which correspond to the two fixed points $p_1, p_2$ of $\Gamma$. By Lemma 2.7, $\partial^+ W^T \cap W^s(\Lambda_0^T)$ consists of two circles and infinitely many lines, where each half line is asymptotic to a circle. Therefore, $\partial^+ W^T \setminus W^s(\Lambda_0^T)$ consists of two open disks and infinitely many strips. By Theorem 2.3, $(W^T, \psi_t^T)$ is topologically equivalent to $(W, \psi_t)$. □
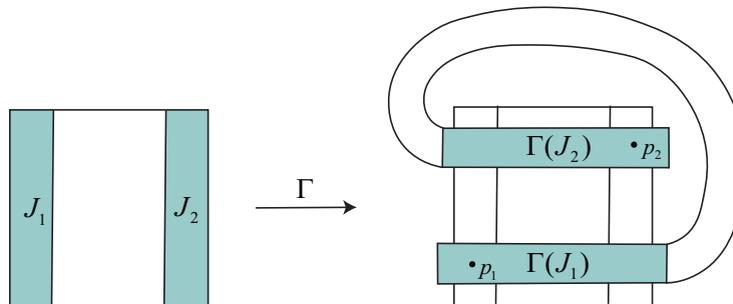


Figure 8: Fake horseshoe map.

# 3 The geodesic flow of the modular surface

Recall that the group $\mathrm{PSL}(2,\mathbb{Z})$ is generated by

$$U = \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad V = \pm \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

As in the discussion of Ghys in [11], we choose two points in the Poincaré disk at distance $\rho > 0$, and take $U_\rho$ to be the rotation of angle $\pi$ around the one and $V_\rho$ to be the rotation of angle $\frac{2\pi}{3}$ around the other. Recall that the group of positive isometries of the Poincaré disk is isomorphic to $\mathrm{PSL}(2,\mathbb{R})$. Thus, we define a homomorphism $i_\rho \colon \mathrm{PSL}(2,\mathbb{Z}) \to \mathrm{PSL}(2,\mathbb{R})$, such that $i_\rho(U) = U_\rho$ and $i_\rho(V) = V_\rho$. If $\rho_0$ is corresponding to the hyperbolic distance between $\sqrt{-1}$ and $\frac{1}{2}(-1 + \sqrt{-3})$ in Poincaré's upper half plane, then we call the orbifold $\mathbb{H}^2 / i_{\rho_0} \mathrm{PSL}(2,\mathbb{Z})$ the *modular surface*, which admits two cone points and a cusp (see Figure 9, left).

Now, we choose a distance $\rho > \rho_0$. Then $\mathbb{H}^2 / i_\rho \mathrm{PSL}(2,\mathbb{Z})$ is a noncompact orbifold with a "funnel". By cutting the orbifold $\mathbb{H}^2 / i_\rho \mathrm{PSL}(2,\mathbb{Z})$ along the geodesic $\gamma$ of Figure 9, right, we get a compact orbifold $\Sigma$. In fact, the geodesic flow $g_t$ of $\Sigma$ contains the same chain recurrent set as the geodesic flow of $\mathbb{H}^2 / i_\rho \mathrm{PSL}(2,\mathbb{Z})$.

Let $M = T^1 \Sigma$, the unit tangent bundle of $\Sigma$. See Montesinos [17] for the definition of the unit tangent bundle of orbifolds. $M$ is the complement of the trefoil $\xi$ in $S^3$ (see Milnor [16]). Let $\alpha_1$ and $\alpha_2$ be the orbits of $g_t$ (endowed with the natural orientation defined by $g_t$) corresponding to $\gamma$ and $-\gamma$, respectively. In [11], Ghys proved that $\alpha_1$ is a meridian of $\xi$, and $\alpha_2$ is isotopic to $-\alpha_1$. Due to the hyperbolicity of the geodesic flow of $\mathbb{H}^2$, $g_t$ is a Smale flow on $M$ with one nontrivial saddle basic set $\Lambda$. In fact, $\Lambda$ is the maximal invariant set of $g_t$.

It is easy to observe that $g_t$ is tangent to $\alpha_1$, $\alpha_2$, and that $\partial M \setminus (\alpha_1 \sqcup \alpha_2)$ consists of two open annuli $A^+$, $A^-$. Here $g_t$ points inward on $A^+$ and points outward on $A^-$. In the following lines, we will explain how one can distinguish the connected components of $A^+ \setminus W^s(\Lambda)$.

For the convenience of description, we provide the following definitions. Let $l_0$, $l_1$ be two orbits of $g_t$ through $A^+ \setminus W^s(\Lambda)$, then $l_i \cap A^- \neq \varnothing$ for $i = 0, 1$. A topological embedding $L_i \colon [0,1] \to M$ is called
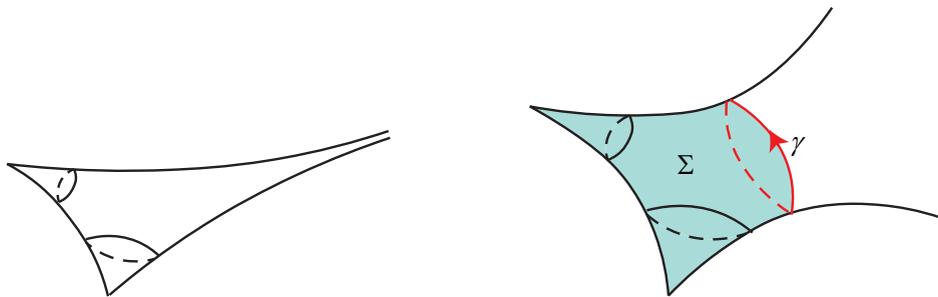


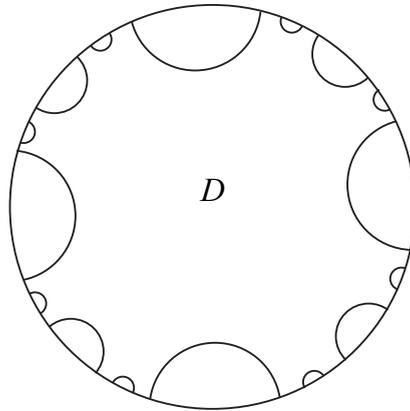Figure 9: Deforming the modular surface.

Figure 10: The universal cover of $\Sigma$.

an *orbit-path* of $l_i$ if the image of $L_i$ is the orbit $l_i$ and $L_i(0) \in A^+ \setminus W^s(\Lambda)$. We say that $l_0$ *is homotopic to $l_1$ preserving $\partial M$* if there is a continuous map $H : [0,1] \times [0,1] \to M$ such that $H_0$ (resp $H_1$) is an orbit-path of $l_0$ (resp $l_1$), and $H_s(0), H_s(1) \in \partial M$ where $H_s = H(\cdot, s)$ for any $s$.

**Lemma 3.1** *Let $l_0, l_1$ be two orbits of $g_t$ through $A^+ \setminus W^s(\Lambda)$ (endowed with the natural orientations defined by $g_t$). The starting points of $l_0$ and $l_1$ are connected in $A^+ \setminus W^s(\Lambda)$ if and only if $l_0$ is homotopic to $l_1$ preserving $\partial M$.*

**Proof  Necessity**  Suppose that the starting points of $l_0$ and $l_1$ are connected in $A^+ \setminus W^s(\Lambda)$. We can connect a path in $A^+ \setminus W^s(\Lambda)$ from the starting point of $l_0$ to the starting point of $l_1$. Note that each positive orbit through this path can reach $A^-$. Therefore, along the orbits through this path, we obtain that $l_0$ is homotopic to $l_1$ preserving $\partial M$.

**Sufficiency**  Suppose that $l_0$ is homotopic to $l_1$ preserving $\partial M$. Then there is a continuous map $H : [0,1] \times [0,1] \to M$ such that $H_0$ (resp $H_1$) is an orbit-path of $l_0$ (resp $l_1$), and $H_s(0), H_s(1) \in \partial M$ where $H_s = H(\cdot, s)$ for any $s$.

Recall that $\Sigma$ is obtained from a noncompact orbifold $\mathbb{H}^2/i_\rho \operatorname{PSL}(2, \mathbb{Z})$ by cutting the funnel, where $\rho > \rho_0$. Then we get the universal cover $D$ of the orbifold $\Sigma$ by removing infinitely many disjoint open semidisks from the Poincaré's disk, as shown in Figure 10. The orbifold-covering map $p : D \to D/i_\rho \operatorname{PSL}(2, \mathbb{Z}) = \Sigma$ is defined by $a \mapsto [a]$. Here, $[a'] = [a]$ if and only if there is an action $B \in i_\rho \operatorname{PSL}(2, \mathbb{Z})$ such that $B(a) = a'$. Obviously, $p$ is a branched covering map.

Define a map $\widetilde{p} : T^1 D \to T^1 D/i_\rho \operatorname{PSL}(2, \mathbb{Z}) = M$ such that $(a, v) \mapsto [(a, v)]$. Here, $[(a', v')] = [(a, v)]$ if and only if there is an action $B \in i_\rho \operatorname{PSL}(2, \mathbb{Z})$ such that $B(a) = a'$ and $dB_a(v) = v'$. It is easy to prove that $\widetilde{p}$ is a covering map. Let $\pi : M \to \Sigma$ and $\widetilde{\pi} : T^1 D \to D$ be the projections of bundles. Then

we get the commutative diagram

$$
\begin{array}{ccc}
T^1 D & \xrightarrow{\ \widetilde{p}\ } & M \\
{\scriptstyle \widetilde{\pi}}\downarrow & & \downarrow{\scriptstyle \pi} \\
D & \xrightarrow{\ p\ } & \Sigma
\end{array}
$$

Let $\widetilde{g}_t$ be the geodesic flow of $D$. According to the definition of the geodesic flow $g_t$,

$$\widetilde{p} \circ \widetilde{g}_t((a, v)) = g_t([(a, v)])$$

for any $(a, v) \in T^1 D$ and any $t$. Then each lift of $H_i$ is an orbit-path in $T^1 D$ for $i = 0, 1$. Let $[(a_i, v_i)] = H_i(0)$, and $\widetilde{H}_0$ be the lift of $H_0$ starting at $(a_0, v_0)$. By the homotopy lifting property (see Hatcher [12]), there is a homotopy $\widetilde{H}_s \colon [0, 1] \to T^1 D$ of $\widetilde{H}_0$ lifts $H_s$ ($s \in [0, 1]$). Without loss of generality, we assume that $\widetilde{H}_1(0) = (a_1, v_1)$.

Using the commutative diagram, $\widetilde{\pi} \circ \widetilde{H}_s(0)$ and $\widetilde{\pi} \circ \widetilde{H}_s(1)$ are two path in $p^{-1}(\partial \Sigma)$. It is not difficult to observe that there is a path $\alpha$ in $\widetilde{p}^{-1}(\partial M)$ from $(a_0, v_0)$ to $(a_1, v_1)$, such that the positive orbits of $\widetilde{g}_t$ through $\alpha$ can reach $\widetilde{p}^{-1}(\partial M)$. Then we get a path $\widetilde{p} \circ \alpha$ in $\partial M$ from $[(a_0, v_0)]$ to $[(a_1, v_1)]$, such that the positive orbits of $g_t$ through $\widetilde{p} \circ \alpha$ can reach $\partial M$. Namely, $[(a_0, v_0)]$ and $[(a_1, v_1)]$ are connected in $A^+ \setminus W^s(\Lambda)$. $\qquad \square$

# 4 The mapping class group of $M$

In the following section, we define $M$, $g_t$, $\Sigma$, $\alpha_1$, $\alpha_2$, $A^+$ as in Section 3. Recall that $\alpha_1$ is a meridian of $\partial M$ and it is isotopic to $-\alpha_2$. Fix an orientation on $M$. This induces an orientation on $\partial M$. We choose a circle $l$ in $\partial M$ intersecting once $\alpha_1$ and $\alpha_2$. Then $l$ is a longitude of $\partial M$. Choose an orientation for $l$ such that $(\alpha_1, l)$ is positively oriented on $\partial M$. Without losing generality, we assume that the positive direction of $l \cap A^+$ is from $\alpha_1$ to $\alpha_2$, as shown in Figure 11.

Let $h$ be a topological equivalence from $(M, g_t)$ to $(M, g_t)$. Then $h(\alpha_1 \sqcup \alpha_2) = \alpha_1 \sqcup \alpha_2$ and $h(A^+) = A^+$. If $h(\alpha_1) = \alpha_1$, then $h(l) \simeq x\alpha_1 + l$. Thus the action of $h|_{\partial M}$ on the homology group of $\partial M$ corresponds
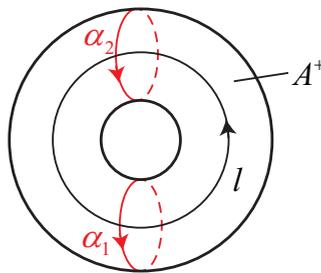


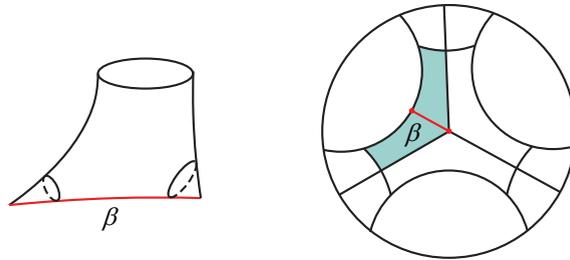Figure 11: The positive direction of $l \cap A^+$ is from $\alpha_1$ to $\alpha_2$.

Figure 12: The orbifold $\Sigma$ and the three fold cover of its fundamental domain (shaded area).

to the matrix $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ for some $x \in \mathbb{Z}$. Similarly, if $h(\alpha_1) = \alpha_2$, then the action of $h|_{\partial M}$ on the homology group of $\partial M$ corresponds to the matrix $\begin{pmatrix} -1 & y \\ 0 & -1 \end{pmatrix}$ for some $y \in \mathbb{Z}$. Therefore, $h|_{\partial M}$ preserves the induced orientation of $\partial M$, which implies that $h$ preserves the orientation of $M$.

Next, let us prove that the image by $h$ of $\alpha_1$ determines the isotopy class of $h$. Denote by $\mathrm{Mod}^+(M)$ *the mapping class group* of $M$, ie, the group of isotopy classes of the orientation-preserving homeomorphisms of $M$. For the orbifold $\Sigma$, let $\mathrm{Mod}^{\pm}(\Sigma)$ be the group of homeomorphisms of $\Sigma$ fixing the singular points, modulo isotopies fixing the singular points. Then $\mathrm{Mod}^{\pm}(\Sigma) \cong \mathbb{Z}_2$.

Johannson [14, Proposition 25.3] proposed a short exact sequence

$$1 \to \mathrm{H}_1(\Sigma, \partial\Sigma) \to \mathrm{Mod}^+(M) \to \mathrm{Mod}^{\pm}(\Sigma) \to 1.$$

Here, $\mathrm{H}_1(\Sigma, \partial\Sigma)$ denotes the first relative homology group, then it is a trivial group. This indicates that $\mathrm{Mod}^+(M) \cong \mathrm{Mod}^{\pm}(\Sigma) \cong \mathbb{Z}_2$.

Let $\beta$ be a geodesic of $\Sigma$ connecting the two singular points; see Figure 12. Let $f$ be the symmetric map on $\Sigma$ about $\beta$. Obviously, $f$ is an isometry, which implies that the map $F = df : M \to M$ induced by $f$ is a topological equivalence from $(M, g_t)$ to $(M, g_t)$. Thus, $F$ preserves the orientation of $M$. In addition, $F(\alpha_1) = \alpha_2$ is isotopic to $-\alpha_1$, then $F$ cannot be isotopic to $\mathrm{id}_M$. Therefore, $\mathrm{Mod}^+(M) = \{[\mathrm{id}_M], [F]\}$. Then, we get the following proposition.

**Proposition 4.1**  *Any topological equivalence $h$ from $(M, g_t)$ to $(M, g_t)$ is isotopic to either $\mathrm{id}_M$ or $F$, according to either $h(\alpha_1) = \alpha_1$ or $h(\alpha_1) = \alpha_2$.*

## 5   Proof of Theorem 1.3

Recall the definitions of $(M, g_t)$, $\Lambda$, $\alpha_1$ and $\alpha_2$ in Section 3. In $\partial M$, $g_t$ is tangent to $\alpha_1 \sqcup \alpha_2$, points inward on $A^+$, and points outward on $A^-$. Here, $A^+$ and $A^-$ are two open annuli which are the connected components of $\partial M \setminus (\alpha_1 \sqcup \alpha_2)$. In [9], Bonatti and Pinsky extended the flow $g_t$ into a Smale flow $\phi_t$ on $S^3$, as follows.
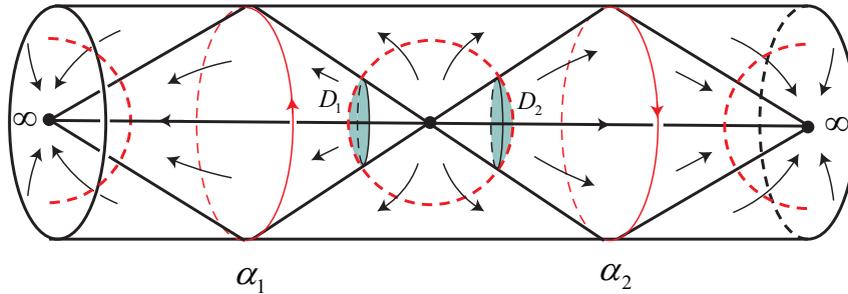
Figure 13: The solid torus is cut along a disk containing $\infty$. The cone-like disks are the stable and the unstable manifolds of the tangent orbits $\alpha_1, \alpha_2$. In the diamond shape regions between the stable and the unstable manifolds of $\alpha_1, \alpha_2$, each point in the past tends to the source and in the future to the sink.

(1)  Construct a flow on a solid torus as shown in Figure 13, whose chain recurrent set consists of a source, a sink and two tangent orbits $\alpha_1, \alpha_2$.

(2)  By suitably gluing the solid torus to $M$, we get a Smale flow $\phi_t$ on $S^3$, whose chain recurrent set consists of $\Lambda$, a source and a sink.

Then $\Lambda$ is the saddle basic set of $\phi_t$. By digging up two open ball neighborhoods of the source and the sink in the solid torus, we get a model $(W, \psi_t)$ for $\Lambda$. Obviously, $g_t$ is the restriction of $\psi_t$ to $M \subset W$, and the maximal invariant set of $\psi_t$ is also $\Lambda$. In the following, we denote by $W^s(\Lambda, W)$ the union of stable manifolds of the orbits in $\Lambda$ for the system $(W, \psi_t)$, and denote by $W^s(\Lambda, M)$ the union of stable manifolds of the orbits in $\Lambda$ for the system $(M, g_t)$.

In [11], Ghys proved that the template related to $\Lambda$ is a Lorenz template. By Proposition 2.8 and Theorem 2.3, $W \cong S^2 \times [0, 1]$ and $\partial^+ W \setminus W^s(\Lambda, W)$ consists of two open disks $D_1, D_2$ and infinitely many strips, where $D_1$ and $D_2$ are shown in Figure 13. Let $\mathcal{S} = \partial^+ W \setminus (W^s(\Lambda, W) \sqcup D_1 \sqcup D_2)$, and $\mathcal{A} = A^+ \setminus W^s(\Lambda, M)$. Along to the orbits of $\psi_t$, we can construct a homeomorphism $\mu \colon \mathcal{S} \to \mathcal{A}$ such that $\mu(x)$ and $x$ lie in a same orbit of $\psi_t$ for each $x \in \mathcal{S}$. Recall that in Section 4, we defined a topological equivalence $F$ from $(M, g_t)$ to $(M, g_t)$. From now on, we distinguish the connected components of $\partial^+ W \setminus W^s(\Lambda, W)$.

**Lemma 5.1**  *Let $h$ be a topological equivalence from $(W, \psi_t)$ to $(W, \psi_t)$. Then there are two possible situations*:

(1)  $h(D_1) = D_1, h(D_2) = D_2$, and $h(C) = C$ for each connected component $C$ of $\mathcal{S}$.

(2)  $h(D_1) = D_2, h(D_2) = D_1$, and $h(C) = \mu^{-1} \circ F \circ \mu(C)$.

**Proof**  Recall that $\partial^+ W \setminus W^s(\Lambda, W)$ consists of two open disks $D_1, D_2$ and infinitely many strips, where $D_i$ is bounded by a circle in $W^s(\alpha_i, W)$ for $i = 1, 2$. Then $h(D_1 \sqcup D_2) = D_1 \sqcup D_2$ and $h(\alpha_1 \sqcup \alpha_2) = \alpha_1 \sqcup \alpha_2$.

It is easy to observe that the set of orbits crossing $A^+$ (resp $A^-$) coincides with the set of orbits crossing $h(A^+)$ (resp $h(A^-)$). Then up to isotopy along the flowlines, $h(\partial M) = \partial M$. This implies that $h$ induces a topological equivalence $h' \colon (M, g_t) \to (M, g_t)$ such that $h'(\alpha_1) = h(\alpha_1)$, $h'(\alpha_2) = h(\alpha_2)$, and $h' \circ \mu(C) = \mu \circ h(C)$ where $C$ is a connected component of $\mathcal{S}$.

Let $l$ be an orbit of $g_t$ through $\mu(C)$ (endowed with the natural orientations defined by $g_t$). Since $\mu$ is a homeomorphism, $\mu(C)$ is a connected component of $\mathcal{A}$.

**Case 1** ($h(D_1) = D_1$ and $h(D_2) = D_2$) Then $h'(\alpha_1) = \alpha_1$ and $h'(\alpha_2) = \alpha_2$. By Proposition 4.1, $h'$ is isotopic to $\mathrm{id}_M$. Then $h'(l)$ is homotopic to $l$ preserving $\partial M$. By Lemma 3.1, the starting points of $h'(l)$ and $l$ are connected in $\mathcal{A}$. Hence, $h' \circ \mu(C) = \mu(C)$ and $h(C) = C$.

**Case 2** ($h(D_1) = D_2$ and $h(D_2) = D_1$) Then $h'(\alpha_1) = \alpha_2$ and $h'(\alpha_2) = \alpha_1$. By Proposition 4.1, $h'$ is isotopic to $F$. Then $h'(l)$ is homotopic to $F(l)$ preserving $\partial M$. By Lemma 3.1, the starting points of $h'(l)$ and $F(l)$ are connected in $\mathcal{A}$. Hence, $h' \circ \mu(C) = F \circ \mu(C)$ and $h(C) = \mu^{-1} \circ F \circ \mu(C)$. $\qquad\square$

**Proof of Theorem 1.3** Let $\{C_i\}_{i \in \tau}$ be an infinite set of connected components of $\mathcal{S}$ such that if $C \in \{C_i\}_{i \in \tau}$, then $\mu^{-1} \circ F \circ \mu(C) \notin \{C_i\}_{i \in \tau}$. For each $i \in \tau$, choose two disjoint closed disks in $C_i$ and then do handle attachment for $(W, \psi_t)$ along them. We obtain a manifold $N_i$ endowed with a flow $\psi_t^i$, such that $(N_i, \psi_t^i)$ is a hyperbolic plug for $\Lambda$. Since $W \cong S^2 \times [0, 1]$ and the orbits through $C_i$ are isotopic, we have $N_i \cong T^2 \times [0, 1]$.

Let $\Lambda_i$ be the maximal invariant set of $\psi_t^i$. Then $\Lambda_i$ is a 1-dimensional saddle basic set containing infinitely many closed orbits. Moreover, for any $j \in \tau$ and $j \neq i$, we have $[\psi_t^i, \Lambda_i] = [\psi_t^j, \Lambda_j]$. Now, we prove that $\psi_t^i$ is not topologically equivalent to $\psi_t^j$ by contradiction.

Suppose that there are two different elements $i_0, j_0 \in \tau$, such that $\psi_t^{i_0}$ is topologically equivalent to $\psi_t^{j_0}$ via a homeomorphism $h_0 \colon (N_{i_0}, \psi_t^{i_0}) \to (N_{j_0}, \psi_t^{j_0})$. By our construction and the fact that $\partial W$ consists of two spheres, there is only one component $T_{i_0}$ (resp $T_{j_0}$) of $\partial^+ N_{i_0} \setminus W^s(\Lambda_{i_0})$ (resp $\partial^+ N_{j_0} \setminus W^s(\Lambda_{j_0})$) with genus one. Since $h_0$ is a topological equivalence, $h_0(T_{i_0}) = T_{j_0}$, which implies that $h_0$ sends an inseparable circle $c_{i_0}$ of $T_{i_0}$ (that is, $T_{i_0} \setminus c_{i_0}$ is connected) to an inseparable circle $c_{j_0}$ of $T_{j_0}$. We cut $N_{i_0}$ (resp $N_{j_0}$) along the orbits through $c_{i_0}$ (resp $c_{j_0}$) and then glue two cylinders $D^2 \times [0, 1]$ endowed with the vector field $\frac{\partial}{\partial t}$ by preserving the oriented orbits. Then, we get the system $(W, \psi_t)$ up to topological equivalence. Thanks to $h_0$, we can construct a topological equivalence $h_0'$ from $(W, \psi_t)$ to $(W, \psi_t)$ such that $h_0'(C_{i_0}) = C_{j_0}$. By Lemma 5.1, $C_{j_0} = C_{i_0}$ or $\mu^{-1} \circ F \circ \mu(C_{i_0})$, which contradicts the definition of $\{C_i\}_{i \in \tau}$. Therefore, $\psi_t^i$ cannot be topologically equivalent to $\psi_t^j$. $\qquad\square$

**Proof of Corollary 1.4** Let $\eta_t^a$ be the flow on $N' = S^1 \times D^2$ induced by the vector field

$$X(\theta, z) = \frac{\partial}{\partial \theta} - z_1 \frac{\partial}{\partial z_1} - z_2 \frac{\partial}{\partial z_2}.$$

Here, $z_1$, $z_2$ denote the standard coordinate functions on $D^2$ and $\frac{\partial}{\partial\theta}$ is a vector field along the factor $S^1$. Let $N''$ be a second copy of $N'$ equipped with the flow $\eta_t^r$ induced by $-X$. By suitably attaching $(N', \eta_t^a)$ and $(N'', \eta_t^r)$ to $(N_i, \psi_t^i)$ for each $i \in \tau$, we can construct infinitely many simple Smale flows on $S^3$ that satisfy the conclusion of Corollary 1.4. □

# 6 Future research

In this paper, utilizing the theory on models of Béguin and Bonatti [2], we constructed infinitely many hyperbolic plugs that are homeomorphic to the Hopf link exterior, such that no two plugs can be topologically equivalent. This prompts a problem: can we do the same trick for the hyperbolic plugs whose background manifolds are other link exteriors?

The forthcoming work by Fan, Lai, and Yu presents a significant finding that the Whitehead link exterior cannot be the background manifold of a hyperbolic plug. Then, a potential starting point for this inquiry could involve investigating which two-component link can be realized as the attractor and repeller of a simple Smale flow on $S^3$, along with determining the model for the corresponding saddle basic set.

# References

[1] **K M Adhikari**, **M C Sullivan**, *Further study of simple Smale flows using four band templates*, Topology Proc. 50 (2017) 21–37 MR

[2] **F Béguin**, **C Bonatti**, *Flots de Smale en dimension* 3: *présentations finies de voisinages invariants d'ensembles selles*, Topology 41 (2002) 119–162 MR

[3] **F Béguin**, **C Bonatti**, **B Yu**, *Building Anosov flows on* 3*-manifolds*, Geom. Topol. 21 (2017) 1837–1930 MR

[4] **J S Birman**, **R F Williams**, *Knotted periodic orbits in dynamical system, II*: *Knot holders for fibered knots*, from "Low-dimensional topology" (S J Lomonaco, Jr, editor), Contemp. Math. 20, Amer. Math. Soc., Providence, RI (1983) 1–60 MR

[5] **C Bonatti**, **V Grines**, *Knots as topological invariants for gradient-like diffeomorphisms of the sphere* $S^3$, J. Dynam. Control Systems 6 (2000) 579–602 MR

[6] **C Bonatti**, **V Grines**, **F Laudenbach**, **O Pochinka**, *Topological classification of Morse–Smale diffeomorphisms without heteroclinic curves on* 3*-manifolds*, Ergodic Theory Dynam. Systems 39 (2019) 2403–2432 MR

[7] **C Bonatti**, **V Grines**, **V Medvedev**, **E Pécou**, *Topological classification of gradient-like diffeomorphisms on* 3*-manifolds*, Topology 43 (2004) 369–391 MR

[8] **C Bonatti**, **V Grines**, **O Pochinka**, *Classification of Morse–Smale diffeomorphisms with the chain of saddles on* 3*-manifolds*, from "Foliations 2005" (P Walczak, R Langevin, S Hurder, T Tsuboi, editors), World Sci., Hackensack, NJ (2006) 121–147 MR

[9] **C Bonatti**, **T Pinsky**, *Lorenz attractors and the modular surface*, Nonlinearity 34 (2021) 4315–4331 MR

[10] **K Bonatti**, **V Z Grines**, **O V Pochinka**, *Classification of Morse–Smale diffeomorphisms with a finite set of heteroclinic orbits on* 3*-manifolds*, Tr. Mat. Inst. Steklova 250 (2005) 5–53 MR

[11] **E Ghys**, *Knots and dynamics*, from "International Congress of Mathematicians, I" (M Sanz-Solé, J Soria, J L Varona, J Verdera, editors), Eur. Math. Soc., Zürich (2007) 247–277 MR

[12] **A Hatcher**, *Algebraic topology*, Cambridge Univ. Press (2002) MR

[13] **E L Haynes**, **M C Sullivan**, *Simple Smale flows with a four band template*, Topology Appl. 177 (2014) 23–33 MR

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

[15] **V Meleshuk**, *Embedding templates in flows*, PhD thesis, Northwestern University (2002) MR Available at https://www.proquest.com/docview/305519791

[16] **J Milnor**, *Hyperbolic geometry*: *the first* 150 *years*, Bull. Amer. Math. Soc. 6 (1982) 9–24 MR

[17] **J M Montesinos**, *Classical tessellations and three-manifolds*, Springer (1987) MR

[18] **O V Pochinka**, **D D Shubin**, *Nonsingular Morse–Smale flows with three periodic orbits on orientable* 3*-manifolds*, Mat. Zametki 112 (2022) 426–443 MR In Russian; translated in Math. Notes 112 (2022), 436–450

[19] **S Smale**, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. 73 (1967) 747–817 MR

[20] **M C Sullivan**, *Visually building Smale flows in* $S^3$, Topology Appl. 106 (2000) 1–19 MR

[21] **M C Sullivan**, *Realizing full n-shifts in simple Smale flows*, Topology Appl. 218 (2017) 19–29 MR

[22] **B Yu**, *Lorenz like Smale flows on three-manifolds*, Topology Appl. 156 (2009) 2462–2469 MR

[23] **B Yu**, *The templates of non-singular Smale flows on three manifolds*, Ergodic Theory Dynam. Systems 32 (2012) 1137–1155 MR

[24] **B Yu**, *Behavior 0 nonsingular Morse Smale flows on* $S^3$, Discrete Contin. Dyn. Syst. 36 (2016) 509–540 MR

[25] **B Yu**, *Every* 3*-manifold admits a structurally stable nonsingular flow with three basic sets*, Proc. Amer. Math. Soc. 144 (2016) 4949–4957 MR

*School of Mathematical Sciences, Tongji University*
*Shanghai, China*

Current address: *School of Mathematics and Statistics, Huangshan University*
*Huangshan, China*

fangfangchen_97@163.com

# One-point compactifications of configuration spaces and the self duality of the little disks operad

CONNOR MALIN

Using configuration space level Pontryagin–Thom constructions, we construct a simple Koszul self duality map for the little disks operad $\Sigma_+^\infty E_n$. For a framed $n$-manifold $M$, we show that a compatible self duality map exists for $\Sigma_+^\infty E_M$.

## 1 Introduction

The problem of whether a specific operad is equivalent to its own Koszul dual has been studied since Ginzburg–Kapranov originally introduced operadic Koszul duality in 1994. In their paper, they demonstrated that the associative operad was its own Koszul dual, up to a shift [13]. Soon after, Jones–Getzler generalized this to the operad $C_*(E_n; \mathbb{Q})$ of rational chains on the little $n$-disks operad, for $0 < n < \infty$ [12].

Around 2005, Ching introduced Koszul duality for operads of spectra via an explicit point-set model [7]. Naturally, it was conjectured that $\Sigma_+^\infty E_n$ was Koszul self dual. On the level of symmetric sequences, this had already been observed by both Ching and Salvatore, but progress was slow to be made on the operadic statement. Around 2010, Fresse used obstruction-theoretic techniques to show $C_*(E_n; \mathbb{Z})$ was Koszul self dual [11], providing even more evidence that $E_n$ should be topologically self dual.

The question of the topological self duality largely took a pause for the next decade, though Lurie and Ayala–Francis made substantial progress on the Koszul duality of $E_n$-algebras themselves, showing that, up to a shift, the derived indecomposables, or Andre–Quillen homology, of an $E_n$-algebra was an $E_n$-coalgebra [3; 15].

In 2020, the self duality of the topological $E_n$-operad was finally resolved when Ching–Salvatore produced an explicit equivalence of operads [8]

$$\Sigma_+^\infty E_n \simeq s_n K(\Sigma_+^\infty E_n).$$

The arguments involved in constructing this equivalence are rather technical and require deep geometric understanding of the $E_n$-operad, operadic bar constructions, and compactifications of configuration spaces. Recently, we made progress toward organizing the theory of topological Koszul self duality using operads

in the category of stable fibrations. As an application, we extended Ching–Salvatore's equivalence to the right modules $E_M$, defined as certain configuration spaces of disks in a framed $n$-manifold $M$ [18].

It is not clear if these techniques are applicable to study the effect of the Koszul self duality of the $E_n$-operad on $E_n$-algebras. This is unsatisfying, since, at the chain level, $E_n$-algebras were the primary motivation of Jones–Getzler for demonstrating the Koszul self duality of $C_*(E_n; \mathbb{Q})$. It has long been suspected that several $\infty$-categorical [15] and geometric [3] approaches to Koszul duality of topological $E_n$-algebras should be equivalent to a combination of the self duality of $E_n$ and the usual theory of Koszul duality of algebras over an operad [10].

In order to begin addressing this question, we first give a simple geometric construction of the Koszul self duality of $E_n$ with respect to a version of Koszul duality recently introduced by Lurie [16] and Espic [9]. Simply stated, the Koszul dual of $O$ is given by the coendomorphism operad (Definition 4.6) of the trivial right $O$-module:

$$K(O) := \mathrm{CoEnd}_{\mathrm{RMod}_O}(1).$$

Espic develops the general theory of Koszul duality in this framework and compares it to Ching's model of Koszul duality, showing they agree in a certain sense.

Using this definition of Koszul duality, one reminiscent of the classical Yoneda algebra approach, the self duality of $E_n$ is remarkably simple. Let $\mathcal{F}_n$ denote the configuration space model of $E_n$ known as the Fulton–MacPherson operad. Let $\mathcal{F}_M$, $\mathcal{F}_{M+}$ denote the right modules of configurations in a framed manifold $M$ and its one-point compactification $M^+$, respectively [20]. We show that the contravariant functoriality of $\mathcal{F}_{M+}$ with respect to embeddings which dilate framings can be used to construct an equivalence of operads

$$\Sigma_+^\infty E_n \xrightarrow{\simeq} \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\mathcal{F}_{(\mathbb{R}^n)+}).$$

There is an equivalence of right modules $\mathcal{F}_{(\mathbb{R}^n)+} \simeq S^n \wedge 1$ coming from "scaling configurations to $\infty$", and so we conclude:

**Theorem 1.1** (Theorem 7.13: self duality of $E_n$)  *There is a zigzag of weak equivalences of operads:*

$$\Sigma_+^\infty E_n \simeq \cdots \simeq s_n K(\Sigma_+^\infty E_n).$$

Similar arguments allow us to extend this result to the modules $\mathcal{F}_U$ where $U \subset \mathbb{R}^n$ is open. Applying the homotopy cosheaf techniques of [18, Section 8], we conclude:

**Theorem 1.2** (Theorem 8.7: self duality of $\mathcal{F}_M$)  *There is a zigzag of weak equivalences of operads*

$$\Sigma_+^\infty \mathcal{F}_n \simeq \cdots \simeq s_n K(\Sigma_+^\infty \mathcal{F}_n).$$

*For any framed manifold $M$, there is a compatible zigzag of weak equivalences of right modules*

$$\Sigma_+^\infty \mathcal{F}_M \simeq \cdots \simeq s_{(n,n)} K(\Sigma^\infty \mathcal{F}_{M+}).$$

## 2  Related work

The Koszul self duality of the spectral $E_n$-operad was recently proven by Ching–Salvatore [8] with respect to a point-set model of Koszul duality introduced in [7]. Building on this, the author constructed a compatible Koszul self duality result for the modules $E_M$ in [18].

In this paper, we demonstrate the Koszul self duality of $E_n$ and $E_M$ using a new model of Koszul duality and different techniques. Though there are comparisons between this model and Ching's model [9], they take place in a larger category of PROPs. As such, there is not yet a comparison with the results of [8; 18]. In this paper, we do make use of [18, Section 8] to extend from submanifolds of $\mathbb{R}^n$ to general framed manifolds.

Koszul duality results for $E_n$-algebras have a longer history. Lurie and Ayala–Francis gave two different constructions of Koszul duality for $E_n$-algebras [3; 15]. To the author's knowledge, the point-set model of Koszul duality used in [8] does not have an obvious extension to algebras, and so the result of Ching–Salvatore does not immediately yield a Koszul duality of $E_n$-algebras.

This paper is greatly inspired by Ayala–Francis's work on Poincaré/Koszul duality [3]. Our construction can be interpreted as Poincaré/Koszul duality "without the algebras". The coendomorphism model of Koszul duality has an extension to algebras, and so when combined with our self duality result, yields a third version of Koszul duality for $E_n$-algebras. It is straightforward to compare this Koszul duality with Ayala–Francis's, but there is a significant amount of work needed to recover the full statement Poincaré/Koszul duality.

### Acknowledgements

## 3  Outline and conventions

In Section 4, we review the category of operads and the model category of $S$-modules. In Section 5, we recall the category of right modules and some of its basic homotopy theory. In Section 6, we discuss Koszul duality. In Section 7, we construct the equivalence $\Sigma_+^\infty E_n \simeq s_n K(\Sigma_+^\infty E_n)$. In Section 8, we prove the analogous result for the right modules $\mathcal{F}_M$.

The heart of this paper is Sections 6, 7, and 8 in which we develop Koszul duality using coendomorphism operads, construct a Pontryagin–Thom map from $E_n$ to its Koszul dual, and extend the construction to right modules. Those familiar with the basics of operads and right modules can read these sections independently with the other sections providing necessary technical support.

Throughout this paper, we make use of model categories. Our model of spectra is the model category of $S$-modules. The utility of $S$-modules lies in the fact that all $S$-modules are fibrant, and so the coendomorphism operads of cofibrant objects are homotopy invariant. The results of this paper can be replicated in any reasonable model of $\infty$-categories which can describe (co)endomorphism operads in a homotopy invariant way.

If $(\mathcal{V}, \otimes)$ is a symmetric monoidal category, one may consider $\mathcal{V}$-enriched categories $C$ which have morphism objects in $\mathcal{V}$ and composition laws and identities defined in terms of $\otimes$ and $1_\otimes$. We will always denote enriched mapping objects by $C(-,-)$, except in the case of Spec where we use $F(-,-)$. For an ordinary category $D$, we denote mapping sets by $\mathrm{Map}_D(-,-)$. From an enriched category $C$, we extract an ordinary category with the same objects obtained by applying $\mathrm{Map}_\mathcal{V}(1,-)$ to the morphism objects of $C$. We call this the underlying category of $C$.

Throughout this paper, all manifolds are tame, ie abstractly diffeomorphic to the interior of a smooth manifold with (possibly empty) boundary.

# 4 Operads and $S$-modules

The category Spec is the category of $S$-modules. Recall that $S$-modules form a symmetric monoidal model category with weak equivalences given by maps which induce isomorphisms on homotopy groups [19]. The internal Hom $F(-,-)$ makes it the prototypical example of a Spec-enriched model category [14, Appendix: Enriched model categories]. It has the excellent property that all objects are fibrant.

Fix spectra $\overline{S}^0, \overline{S}^1, \overline{S}^{-1}$ to be cofibrant replacements of $S^0, S^1, S^{-1}$, respectively. For $n > 0$, we define $\overline{S}^n := (\overline{S}^1)^{\wedge n}$. Similarly, we define $\overline{S}^{-n} := (\overline{S}^{-1})^{\wedge n}$. The axioms of a symmetric monoidal model category imply each of these are cofibrant and weakly equivalent to the evident sphere. Throughout this paper, we will also make use of preferred cofibrant replacements of suspension spectra $\Sigma^\infty X$, at least when $X$ is a CW complex to begin with [1, Lemma 1.5]:

$$\overline{S}^0 \wedge \Sigma^\infty X \to S^0 \wedge X \cong X.$$

Fix a Spec-enriched model category $C$ with all objects fibrant.

**Definition 4.1** An object $c \in C$ "has the correct mapping spectra" if for any $d$ and any equivalence $c' \xrightarrow{\simeq} c$ with $c'$ cofibrant, the map

$$C(c,d) \to C(c',d)$$

induced by precomposition is a weak equivalence.

**Lemma 4.2** *If $c$ has the correct mapping spectra, then for any weak equivalence $d \to d'$,*

$$C(c,d) \to C(c,d')$$

*induced by postcomposition is a weak equivalence.*

**Proof** Since $c$ has the correct mapping spectra, we may replace it by a cofibrant replacement without changing the homotopy type of the map $C(c, d) \to C(c, d')$. The axioms of an enriched model category imply that if $c$ is cofibrant, $C(c, -)$ preserves weak equivalences of fibrant objects, so we get the desired result. $\square$

A curious property of $S$-modules is that although suspension spectra of CW complexes behave poorly with respect to mapping spaces, they behave well with respect to mapping spectra. Fundamentally, this is due to the fact that although $\Sigma^\infty S^0$ is not cofibrant, it is the monoidal unit and so $F(\Sigma^\infty S^0, Z) \cong Z$ by elementary category theory. This implies that $\Sigma^\infty S^0$ has the correct mapping spectra since $F(\overline{S}^0, Z)$ has the weak equivalence type of $Z$ because $\pi_n(F(\overline{S}^0, Z)) \cong [\overline{S}^n \wedge \overline{S}^0, Z] \cong [\overline{S}^n, Z]$.

**Lemma 4.3** *If $Z$ is a CW complex, $\Sigma^\infty Z$ has the correct mapping spectra.*

**Proof** It suffices to demonstrate the hypothesis for the cofibrant replacement $\overline{S}^0 \wedge \Sigma^\infty Z \xrightarrow{\simeq} Z$. Since $\overline{S}^0$ is cofibrant, this is equivalent to showing

$$F(\overline{S}^0, F(\Sigma^\infty Z, Y)) \to F(\overline{S}^0, F(\overline{S}^0 \wedge \Sigma^\infty Z, Y))$$

is a weak equivalence. By adjointness, this is equivalent to showing

$$F(\overline{S}^0 \wedge \Sigma^\infty Z, Y) \to F(\overline{S}^0 \wedge \overline{S}^0 \wedge \Sigma^\infty Z, Y)$$

is a weak equivalence. However, $\overline{S}^0 \wedge \Sigma^\infty Z$ and $\overline{S}^0 \wedge \overline{S}^0 \wedge \Sigma^\infty Z$ are both cofibrant, so this follows from the fact that $F(-, Y)$ preserves weak equivalences of cofibrant objects, a consequence of being an enriched model category with all objects fibrant. $\square$

Let $\mathrm{FinSet}_{\geq 1}^{\cong}$ denote the category of finite, nonempty sets and bijections. Let $(V, \otimes)$ be a symmetric monoidal category.

**Definition 4.4** The category of symmetric sequences in $V$ is

$$\mathrm{SymSeq(V)} := \mathrm{Fun}(\mathrm{FinSet}_{\geq 1}^{\cong}, V).$$

When working with symmetric sequences, we often abbreviate the set $\{1, \dots, n\}$ by $n$. The composition maps for operads are defined in terms of the combinatorics of finite sets. For finite sets $I, J$ with $a \in I$, let $I \cup_a J := I - \{a\} \sqcup J$ denote the infinitesimal composite.

**Definition 4.5** An operad in $(V, \otimes)$ is a symmetric sequence $O$ in $V$ together with morphisms called partial composites:

$$O(I) \otimes O(J) \to O(I \cup_a J)$$

for all $a \in I$. These must satisfy straightforward equivariance and associativity conditions.

We say an operad is *reduced* if $O(1) = 1$ and all partial composites involving $O(1)$ are the identity. In this paper, all operads are reduced, and so when referring to $O(I)$ we always assume $|I| \geq 2$. We now assume $C$ also has a symmetric monoidal product $\otimes$, not necessarily compatible with the model structure.

**Definition 4.6** If $f \colon c \to d$ is a morphism in $C$, the coendomorphism operad of $f$ for $|I| \geq 2$ is

$$\mathrm{CoEnd}_C(f)(I) := C(d, c^{\otimes I}).$$

If $f = \mathrm{Id}_c$, we call this $\mathrm{CoEnd}_C(c)$. The partial composites are given informally by

$$
\begin{array}{c}
d \\
\downarrow r \\
c \otimes \cdots \otimes c \otimes \cdots \otimes c \\
\downarrow f \circ_a r \\
c \otimes \cdots \otimes d \otimes \cdots \otimes c \\
\downarrow s \circ_a f \circ_a r \\
c \otimes \cdots \otimes (c \otimes \cdots \otimes c) \otimes \cdots \otimes c
\end{array}
$$

and more explicitly

$$
\begin{array}{c}
C(d, c^{\otimes I}) \wedge C(d, c^{\otimes J}) \\
\downarrow \\
C(d, c^{\otimes I}) \wedge C(d, c^{\otimes J}) \wedge S^0 \wedge (S^0)^{\wedge I - \{a\}} \\
\downarrow \scriptstyle{\mathrm{Id} \wedge \mathrm{Id} \wedge f \wedge (1_c)^{\wedge I - \{a\}}} \\
C(d, c^{\otimes I}) \wedge C(d, c^{\otimes J}) \wedge C(c, d) \wedge C(c, c)^{\wedge I - \{a\}} \\
\downarrow \\
C(d, c^{\otimes I}) \wedge C(c^{\otimes I}, c^{\otimes I \cup_a J}) \longrightarrow C(d, c^{\otimes I \cup_a J})
\end{array}
$$

where the bottom-most vertical map is given by the identity smashed with the composition of the middle two function spectra followed by the smash product of functions.

**Proposition 4.7** *For a morphism $f \colon c \to d$ in $C$, there are maps of operads*

$$\mathrm{CoEnd}_C(c) \leftarrow \mathrm{CoEnd}_C(f) \to \mathrm{CoEnd}_C(d)$$

*induced by precomposition with $f$ and postcomposition with $f^{\otimes n}$.*

**Proof** These follow from unwinding definitions. $\qquad\square$

We say a weak equivalence $f \colon c \to d$ in $(C, \otimes)$ is $\otimes$-nice if $\otimes n \colon c^{\otimes n} \to d^{\otimes n}$ is a weak equivalence for all $n$. For instance, the standard theory of monoidal model categories implies that all weak equivalences of cofibrant spectra are $\wedge$-nice, though we don't generally require any interaction of $\otimes$ and the model structure in the coming propositions.

**Lemma 4.8** *If $X$ is a CW complex, the weak equivalence $\bar{S}^0 \wedge \Sigma^\infty X \to \Sigma^\infty X$ is $\wedge$-nice.*

**Proof** Since $\Sigma^\infty$ is symmetric monoidal, we may write the $n^{\text{th}}$ power of this map as

$$(\bar{S}^0)^{\wedge n} \wedge \Sigma^\infty(X^{\wedge n}) \to \Sigma^\infty(X^{\wedge n}),$$

which is a cofibrant replacement of $\Sigma^\infty(X^{\wedge n})$ since $(\bar{S}^0)^{\wedge n}$ is cofibrant and $X^{\wedge n}$ is a CW complex.  □

**Lemma 4.9** *If $f: c \xrightarrow{\simeq} c'$ is a $\otimes$-nice weak equivalence in $C$ such that $c$ is cofibrant and $c'$ has correct mapping spectra, then*
$$\mathrm{CoEnd}_C(c) \xleftarrow{\simeq} \mathrm{CoEnd}_C(f) \xrightarrow{\simeq} \mathrm{CoEnd}_C(c').$$

**Proof** The first weak equivalence follows from Definition 4.1. The second follows from Lemma 4.2 combined with the fact $f$ is $\otimes$-nice.  □

Suppose $O$, $P$ are operads. The operad $O \otimes P$ is given by

$$(O \otimes P)(I) := O(I) \otimes P(I),$$

and the partial composites are obtained by taking smash products.

**Definition 4.10** The $n$-sphere operad is

$$S_n = \mathrm{CoEnd}_{\mathrm{Spec}}(S^n).$$

The $n$-sphere operad is used to define operadic suspension:

$$s_n O := S_n \wedge O.$$

**Remark 4.11** There are several uses for operadic suspension. For example, if $C$ is an $O$-coalgebra, $\Sigma^n C$ is an $s_n O$-coalgebra. Dually, if $A$ is an $O$-algebra, $\Sigma^{-n} A$ is an $s_n O$-algebra, up to weak equivalence.

# 5  Right modules over operads

**Definition 5.1** A right module $R$ over an operad $O$ in $(V, \otimes)$ is a symmetric sequence $R$ in $V$ with morphisms called partial composites:

$$R(I) \otimes O(J) \to R(I \cup_a J)$$

for all $a \in I$. These must satisfy straightforward equivariance and associativity conditions.

As before, $(C, \otimes)$ is a symmetric monoidal category which is also a Spec-enriched model category.

**Definition 5.2**   Given $b \in C$ and $f : c \to d$ in $C$, the right $\mathrm{CoEnd}_C(f)$-module $\mathrm{CoEnd}_C^d(b, f)$ is given by

$$\mathrm{CoEnd}_C^d(b, f)(I) := C(b, d^{\otimes I}).$$

The partial composites are given by postcomposition with $\mathrm{CoEnd}_C(f)$, followed by postcomposition with $f$.

The right $\mathrm{CoEnd}_C(f)$-module $\mathrm{CoEnd}_C^c(b, f)$ is given by

$$\mathrm{CoEnd}_C^c(b, f)(I) := C(b, c^{\otimes I}).$$

The partial composites are given by given by postcomposition with $f$, followed by postcomposition with $\mathrm{CoEnd}_C(f)$.

If $f = \mathrm{Id}_c$, we denote these identical modules by $\mathrm{CoEnd}_C(b, c)$.

**Proposition 5.3**   *If $a \xrightarrow{\simeq} b$ is a weak equivalence in $C$ of objects with the correct mapping spectra and $f : c \to d$ in $C$, there are weak equivalences of right modules*

$$\mathrm{CoEnd}_C^d(b, f) \xrightarrow{\simeq} \mathrm{CoEnd}_C^d(a, f), \quad \mathrm{CoEnd}_C^c(b, f) \xrightarrow{\simeq} \mathrm{CoEnd}_C^c(a, f)$$

*induced by precomposition.*

**Proof**   As before, to check these are weak equivalences we may replace $a, b$ by cofibrant objects for which this follows from the axioms of an enriched model category since all objects are assumed fibrant.   □

Given operads $O$, $P$, a right $O$-module $Q$, and a right $P$-module R, we say a map of symmetric sequences $Q \to R$ is *compatible* with a map of operads $O \to P$ if the obvious diagrams commute.

**Proposition 5.4**   *Given $b \in C$ and $f : c \to d$ in $C$, there are compatible maps of operads*

$$\mathrm{CoEnd}_C(c) \leftarrow \mathrm{CoEnd}_C(f) = \mathrm{CoEnd}_C(f) \to \mathrm{CoEnd}_C(d)$$

*and maps of right modules*

$$\mathrm{CoEnd}_C(b, c) \leftarrow \mathrm{CoEnd}_C^c(b, f) \to \mathrm{CoEnd}_C^d(b, f) \to \mathrm{CoEnd}(b, d),$$

*where the first arrow is the identity on symmetric sequences, the second arrow is given by postcomposition with $f$, and the third arrow is the identity on symmetric sequences.*

*If $b$ has the correct mapping spectra and $f : c \xrightarrow{\simeq} d$ is a $\otimes$-nice weak equivalence, then these are levelwise weak equivalences.*

**Proof**   The compatibility is easily checked and the statement about weak equivalences follows from Definition 4.1 and Lemma 4.2.   □

The category of right modules over an operad in (Spec, $\wedge$) has a projective model structure, ie a model structure with levelwise fibrations and weak equivalences [1; 14; 22]. There is a Spec-enrichment of this model structure [1; 14] in terms of the $\circ$-product of symmetric sequences.

**Definition 5.5** For $X, Y \in \mathrm{SymSeq}(\mathrm{Spec})$,

$$(X \circ Y)(I) := \bigvee_{k \geq 1} X(k) \wedge_{\Sigma_k} \bigvee_{U_1 \sqcup \cdots \sqcup U_k = I} (Y(U_1) \wedge \cdots \wedge Y(U_k)).$$

One can think of the $\circ$-product as collecting all of the information about infinitesimal composites $I \cup_a J$ into a single product. We fix an operad $O$ in (Spec, $\wedge$). If $R, R' \in \mathrm{RMod}_O$, the enrichment is of the form

$$\mathrm{RMod}_O(R, R') := \lim(\mathrm{SymSeq}(R, R') \rightrightarrows \mathrm{SymSeq}(R \circ O, R')).$$

If $X$ is a spectrum and $R$ is a right $O$-module, we let $X \wedge R$ denote the right $O$-module,

$$(X \wedge R)(I) := X \wedge R(I).$$

Similarly, if $Q$ is a right $P$-module, we may form the right $O \wedge P$-module $R \wedge Q$,

$$(R \wedge Q)(I) := R(I) \wedge Q(I),$$

in both cases the partial composites are determined by smashing.

**Definition 5.6** For a right $O$-module $R$, the $s_n O$-module $s_{(n,d)} R$ is given by $S^d \wedge S_n \wedge R$.

**Lemma 5.7** *The map* $\mathrm{RMod}_O(R, R') \to \mathrm{RMod}_O(\overline{S}^0 \wedge R, R')$ *is a weak equivalence. As a consequence, if* $\overline{S}^0 \wedge R$ *is cofibrant, then* $R$ *has the correct mapping spectra.*

**Proof** By adjunction, the limit which computes the codomain is

$$\lim\big(F(\overline{S}^0, \mathrm{SymSeq}(R, R')) \rightrightarrows F(\overline{S}^0, \mathrm{SymSeq}(R \circ O, R'))\big).$$

Since $F$ commutes with limits, this is

$$F\big(\overline{S}^0, \lim(\mathrm{SymSeq}(R, R') \rightrightarrows \mathrm{SymSeq}(R \circ O, R'))\big),$$

which is weakly equivalent to $\mathrm{RMod}_O(R, R')$. $\square$

**Definition 5.8** The tensor product of right $O$-modules $R, R'$ is the right $O$-module given by

$$(R \otimes R')(K) := \bigvee_{K = I \sqcup J} R(I) \wedge R'(J),$$

with partial composites determined by smashing.

The categories of right modules behave well with respect to operad morphisms [14, Proposition 2.4]:

**Proposition 5.9** *Given a map of operads $f: O \to P$, there is an enriched Quillen adjunction*

$$\mathrm{RMod}_O \underset{\mathrm{ind}_f}{\overset{\mathrm{res}_f}{\rightleftarrows}} \mathrm{RMod}_P$$

*If $f$ is a weak equivalence, this is a Quillen equivalence.*

When convenient, we will write $\mathrm{ind}_f(-)$ as $\mathrm{ind}_O^P(-)$, that is the induction up to $P$, and $\mathrm{res}_f$ as $\mathrm{res}_O^P$, that is the restriction down from $P$. A convenient model of induction is the relative $\circ$-product $R \circ_O P$ of the right $O$-module $R$ with the left $O$ bimodule $P$, though we will not make use of any particular model.

A computation with the Yoneda lemma shows

**Lemma 5.10** *If $O \to P$ is a map of operads, there is a natural isomorphism of right $P$-modules*

$$\mathrm{ind}_O^P(R \otimes R') \cong \mathrm{ind}_O^P R \otimes \mathrm{ind}_O^P R'.$$

Let 1 denote the initial and terminal reduced operad in spectra which has its first spectrum $S^0$ and all other spectra $*$. It is easy to see

$$\mathrm{RMod}_1 \cong \mathrm{SymSeq}(\mathrm{Spec}).$$

**Definition 5.11** Given $X \in \mathrm{SymSeq}(\mathrm{Spec})$,

$$\mathrm{Free}_O(X) := \mathrm{ind}_1^O X, \quad \mathrm{Triv}_O(X) := \mathrm{res}_O^1 X.$$

**Definition 5.12** The indecomposables of an $O$-module $R$ are

$$\mathrm{Indecom}(R) := \mathrm{ind}_O^1(R).$$

**Lemma 5.13** *For a spectrum $X$, there is an isomorphism $\mathrm{Indecom}(X \wedge R) \cong X \wedge \mathrm{Indecom}(R)$.*

**Proof** The indecomposables may be computed as a cofiber along all the partial composites. Since smashing commutes with cofibers, the result follows. □

Since induction is left Quillen, we have:

**Lemma 5.14** *If $R$ is a cofibrant $O$-module, $\mathrm{ind}_O^P(R)$ is a cofibrant $P$-module. In particular, if $P = 1$, we see $\mathrm{Indecom}(R)$ is a cofibrant symmetric sequence.*

**Proposition 5.15** *If $O$ is an operad in $(\mathrm{Top}_*, \wedge)$, then there is a Quillen adjunction given by levelwise application of the indicated functors*

$$\mathrm{RMod}_O \underset{\bar{S}^n \wedge \Sigma^\infty(-)}{\overset{\mathrm{Map}(\bar{S}^n, -)}{\rightleftarrows}} \mathrm{RMod}_{\Sigma^\infty O}$$

**Proof**   It suffices to show that for an operad $P$ in $(\mathrm{Spec}, \wedge)$ there is a Quillen adjunction of the form

$$
\mathrm{RMod}_P \underset{\bar{S}^n \wedge -}{\overset{F(\bar{S}^n, -)}{\rightleftarrows}} \mathrm{RMod}_P
$$

since the desired adjunction is obtained by taking $P = \Sigma^\infty O$ and composing with the $\Sigma^\infty$-$\Omega^\infty$ adjunction. Since $F(\bar{S}^n, -)$ preserves fibrations and acyclic fibrations as $\mathrm{Spec}$ is a $\mathrm{Spec}$-enriched model category, the main question is whether the levelwise application of $\mathrm{F}(\bar{S}^n, -)$ is still a right $P$-module. This is implied by the statement that $\mathrm{F}(\bar{S}^n, -)\colon \mathrm{Spec} \to \mathrm{Spec}$ is spectrally enriched. Indeed, it has a canonical enrichment

$$
F(X, Y) \cong S^0 \wedge F(X, Y) \to F(F(\bar{S}^n, S^0), F(\bar{S}^n, S^0)) \wedge F(X, Y) \to F(F(\bar{S}^n, X), F(\bar{S}^n, Y)).
$$

It is then easily verified that $\mathrm{Map}_{\mathrm{RMod}_P}(\bar{S}^n \wedge R, R') \cong \mathrm{Map}_{\mathrm{RMod}_P}(R, F(\bar{S}^n, R'))$.                                    □

**Corollary 5.16**   *If $P$ is an operad in $(\mathrm{Top}_*, \wedge)$ and $R$ is a cofibrant right $P$-module, then $\bar{S}^n \wedge \Sigma^\infty R$ is a cofibrant $\Sigma^\infty P$-module. Similarly, if $O$ is an operad in $(\mathrm{Spec}, \wedge)$ and $R$ is a cofibrant right $O$-module, then $\bar{S}^n \wedge R$ is a cofibrant right $O$-module.*

We now study the mapping spectra between right $O$-modules. For an $O$-module $R$, let $R^{\leq i}$ denote the $O$-module:

$$
R^{\leq i}(J) := \begin{cases} R(J) & \text{if } |J| \leq i, \\ * & \text{if } |J| > i. \end{cases}
$$

By inspection, there is a description of $\mathrm{RMod}_O(R, R')$ as a strict inverse limit of the following tower with the given fibers:

$$
\begin{array}{ccc}
& & \vdots \\
& & \downarrow \\
F^{\Sigma_i}(\mathrm{Indecom}(R)(i), R'(i)) & \longrightarrow & \mathrm{RMod}_O(R^{\leq i}, R'^{\leq i}) \\
& & \downarrow \\
F^{\Sigma_{i-1}}(\mathrm{Indecom}(R)(i-1), R'(i-1)) & \longrightarrow & \mathrm{RMod}_O(R^{\leq i-1}, R'^{\leq i-1}) \\
& & \downarrow \\
& & \vdots \\
& & \downarrow \\
& & \mathrm{RMod}_O(R^{\leq 1}, R'^{\leq 1})
\end{array}
$$

**Proposition 5.17**   *If $R$ or $\bar{S}^0 \wedge R$ is cofibrant, the above tower is a tower of fibrations. As such, the inverse limit is a homotopy inverse limit and the fibers are homotopy fibers.*

**Proof** By Lemma 5.7, we reduce to the case $R$ is cofibrant. The map $R'^{\leq i} \to R'^{\leq i-1}$ is a fibration since all $S$-modules are fibrant, so since $R$ is cofibrant the axioms of an enriched model category imply that

$$\mathrm{RMod}_O(R, R'^{\leq i}) \to \mathrm{RMod}_O(R, R'^{\leq i-1})$$

is a fibration. The result follows from observing $\mathrm{RMod}_O(R, R'^{\leq i}) = \mathrm{RMod}_O(R^{\leq i}, R'^{\leq i})$. □

**Corollary 5.18** *If $f \colon O \xrightarrow{\cong} P$ is a weak equivalence of operads and $R$ is a cofibrant $O$-module, then the underlying symmetric sequence of $\mathrm{ind}_f(R)$ is weakly equivalent to $R$. Additionally, if $R'$ is cofibrant the map*

$$\mathrm{RMod}_O(R, R') \to \mathrm{RMod}_P(\mathrm{ind}_f(R), \mathrm{ind}_f(R)')$$

*is a weak equivalence.*

**Proof** The first fact follows from Proposition 5.9. We will prove the second fact by comparing homotopy limit towers; these towers are homotopy limit towers due to Lemma 5.14 and Proposition 5.17. The map on layers is

$$F^{\Sigma_i}(\mathrm{Indecom}(R)(i), R'(i)) \to F^{\Sigma_i}\big(\mathrm{Indecom}(\mathrm{ind}_f(R))(i), \mathrm{ind}_f(R')(i)\big)$$

This map is induced by the isomorphism $\mathrm{Indecom}(R) \cong \mathrm{Indecom}(\mathrm{ind}_f(R))$ which comes from factoring $O \to 1$ as $O \xrightarrow{f} P \to 1$, and the adjunction unit $R' \to \mathrm{res}_f \mathrm{ind}_f(R')$. The adjunction unit is a weak equivalence since all modules are fibrant and $R'$ is cofibrant.

By Lemma 5.14, the domain is cofibrant as a symmetric sequence, so because all symmetric sequences are fibrant in the projective model structure, the axioms of an enriched model category imply the map of mapping spectra is a weak equivalence. □

We now investigate how coendomorphism operads of modules interact under smashing with spectra.

**Lemma 5.19** *There is a map $F(S^n, S^m) \wedge \mathrm{RMod}_O(R, R') \to \mathrm{RMod}_O(S^n \wedge R, S^m \wedge R))$ which is a weak equivalence if $R$ or $\overline{S}^0 \wedge R$ is cofibrant.*

**Proof** The map is given by taking smash products. By Proposition 5.17, it suffices to show that

$$F(S^n, S^m) \wedge \mathrm{RMod}_O(R, R'^{\leq i}) \to \mathrm{RMod}_O(S^n \wedge R, S^m \wedge R'^{\leq i})$$

is a weak equivalence. There is a homotopy fiber sequence $\mathrm{Triv}_O(R'(i)) \to R'^{\leq i} \to R'^{\leq i-1}$ which gives rise to a map of homotopy fiber sequences

$$
\begin{array}{ccc}
F(S^n, S^m) \wedge \mathrm{RMod}_O\big(R, \mathrm{Triv}_O(R'(i))\big) & \longrightarrow & \mathrm{RMod}_O\big(S^n \wedge R, S^m \wedge \mathrm{Triv}_O(R'(i))\big) \\
\downarrow & & \downarrow \\
F(S^n, S^m) \wedge \mathrm{RMod}_O(R, R'^{\leq i}) & \longrightarrow & \mathrm{RMod}_O(S^n \wedge R, S^m \wedge R'^{\leq i}) \\
\downarrow & & \downarrow \\
F(S^n, S^m) \wedge \mathrm{RMod}_O(R, R'^{\leq i-1}) & \longrightarrow & \mathrm{RMod}_O(S^n \wedge R, S^m \wedge R'^{\leq i-1})
\end{array}
$$

By induction, it suffices to prove the statement of the lemma when $R'$ is concentrated in a single degree, ie is $\mathrm{Triv}_O(X)$ for a spectrum $X$ with a $\Sigma_k$-action. This case follows from the adjunction between indecomposables and trivial modules combined with Lemma 5.13. $\qquad\square$

**Proposition 5.20** *If $R$ or $\overline{S}^0 \wedge R$ is cofibrant, then there is a weak equivalence of operads*

$$s_n \mathrm{CoEnd}_{\mathrm{RMod}_O}(R) \to \mathrm{CoEnd}_{\mathrm{RMod}_O}(S^n \wedge R),$$

*and, if $Q$ or $\overline{S}^0 \wedge Q$ is cofibrant, there is a compatible weak equivalence of modules*

$$s_{(n,0)} \mathrm{CoEnd}_{\mathrm{RMod}_O}(Q, R) \xrightarrow{\simeq} \mathrm{CoEnd}_{\mathrm{RMod}_O}(S^n \wedge Q, S^n \wedge R).$$

**Proof** Recall the notation $s_n$ and $s_{(n,0)}$ both denote $\mathrm{CoEnd}_{\mathrm{Spec}}(S^n) \wedge -$. The maps of operads and modules are then given by taking smash products of function spectra, and so are equivalences by the previous lemma. $\qquad\square$

# 6 Koszul duality via categories of right modules

We have now developed enough of the theory of right modules and coendomorphism operads to define the Koszul dual of an operad and study its properties. This theory was developed thoroughly, and in more generality, in [9] where it was also compared to Ching's model of Koszul duality [9, Proposition 3.19]. The author learned of this definition of Koszul duality in a lecture by Lurie [16].

We will often abbreviate the symmetric sequence $\Sigma_+^\infty \Sigma_i$ concentrated in degree $i$ by $\Sigma_i$.

**Lemma 6.1** *There is a cofibrant model $T$ of $\mathrm{Triv}_O(\Sigma_1)$ with the property that $\bigotimes_i T$ is weakly equivalent to $\mathrm{Triv}_O(\Sigma_i)$.*

**Proof** Pick a cofibrant replacement $O'$ of $O$ with respect to the projective model structure on reduced operads [1, Theorem 9.8]. If $T'$ denotes a cofibrant model of $\mathrm{Triv}_{O'}(\Sigma_1)$, then we claim $\bigotimes_i T'$ is weakly equivalent to $\mathrm{Triv}_{O'}(\Sigma_i)$. By [1, Proposition 9.4], the spectra $T'(j)$ are cofibrant, so the smash products defining $\otimes$ are well behaved and the claim follows. By Lemmas 5.10–5.14 and Corollary 5.18, $T := \mathrm{ind}_{O'}^O(T')$ is cofibrant and inherits the weak equivalence type $\bigotimes_i T \simeq \mathrm{Triv}_O(\Sigma_i)$. $\qquad\square$

Fix such a right module $T$.

**Definition 6.2** The Koszul dual of $O$ is

$$K(O) := \mathrm{CoEnd}_{\mathrm{RMod}_O}(T).$$

**Proposition 6.3** *The weak equivalence class of $K(O)$ is independent of the choice of $T$.*

**Proof** Suppose $T'$ is another cofibrant model of $\mathrm{Triv}_O(\Sigma_1)$ such that $\bigotimes_i T' \simeq \mathrm{Triv}_O(\Sigma_i)$. By the cofibrancy of $T$, there is a weak equivalence $T \xrightarrow{\simeq} T'$ which we wish to demonstrate is $\otimes$-nice. Taking tensor powers yields a map $\bigotimes_i T \to \bigotimes_i T'$. Away from $i$, all spectra in these modules are weakly contractible by assumption, and so it suffices to check that $\bigotimes_i T(i) \to \bigotimes_i T'(i)$ is a weak equivalence.

It is simple to argue that for any operad $O$ and cofibrant $O$-module $R$, $R(1)$ is a cofibrant spectrum. Hence, the map $\bigotimes_i T(i) \to \bigotimes_i T'(i)$ is a wedge of smash products of weak equivalences of cofibrant spectra, and so is a weak equivalence. By Lemma 4.9, the resulting coendomorphism operads are connected by a zigzag of weak equivalences. $\qquad\square$

**Theorem 6.4** *Given a weak equivalence of operads $O \xrightarrow{\simeq} P$, there is a weak equivalence of operads*

$$K(O) \xrightarrow{\simeq} K(P).$$

**Proof** Apply Corollary 5.18 to $\mathrm{RMod}_O(T, T^{\otimes I})$ using Lemma 5.10 to commute $\otimes$ and induction. $\quad\square$

**Remark 6.5** Somewhat surprisingly, the equivalence is in the opposite direction one would expect. The construction of this map crucially relies on $O \to P$ being a weak equivalence, otherwise $\mathrm{ind}_O^P(T)$ would not be a model of $\mathrm{Triv}_P(\Sigma_1)$. Indeed, with a suitable symmetric monoidal and spectrally enriched cofibrant replacement functor on $\mathrm{RMod}_O$, there would be the expected contravariant functoriality with respect to all maps $O \to P$.

Although only weakly functorial at the operad level, Koszul duality has an extension to an honest functor

$$K \colon \mathrm{RMod}_O^{\mathrm{op}} \to \mathrm{RMod}_{K(O)}.$$

**Definition 6.6** The Koszul dual of a right $O$-module $R$ is the $K(O)$-module given by

$$K(R)(I) := \mathrm{CoEnd}_{\mathrm{RMod}_O}(R, T).$$

By Proposition 5.3, we have the expected homotopy invariance:

**Proposition 6.7** *If $R \xrightarrow{\simeq} R'$ are $O$-modules with the correct mapping spectra, then*

$$K(R') \xrightarrow{\simeq} K(R).$$

# 7 Fulton–MacPherson and little disks

In this section, we review different models of the little disks operad such as $E_n$ and $\mathcal{F}_n$. Using the functoriality and cofibrancy of some of their modules, we construct a direct map from $\Sigma_+^\infty E_n$ to a model of $K(\Sigma_+^\infty \mathcal{F}_n)$.

**Definition 7.1** The little $n$-disks operad $E_n$ in Operad(Top, $\times$) has its $I^{\text{th}}$ space equal to the configuration space of $I$-labeled open $n$-disks inside the open unit $n$-disk. Partial composites are given by inserting configurations of disks.

**Remark 7.2** Observe that instead of using configurations of disks, we could have described the $I^{\text{th}}$ space as the space of embeddings of a disjoint union of $I$ unit disks $B^n$ with the requirement that the restriction to each component is a combination of scaling and translation. Let $\mathrm{Emb}^{\mathrm{scale}}(M, N)$ denote the embeddings which preserve the tangential framing up to a pointwise scalar. The collection of maps

$$E_n(I) \to \mathrm{Emb}^{\mathrm{scale}}\left(\bigsqcup_I B^n, B^n\right)$$

sends operad composition to function composition.

If $M$ is a manifold, let $\mathrm{Conf}(M, I)$ denote the configurations of $I$-labeled points in $M$. There is an operad $\mathcal{F}_n$, the Fulton–MacPherson operad which has a heuristic but essentially complete definition as follows [20; 21]:

**Definition 7.3** A point in $\mathcal{F}_n(I)$ is represented by a rooted tree with leaves labeled by $I$ with the constraint that a nonroot, nonleaf vertex $v$ is labeled by $\mathrm{Conf}(\mathbb{R}^n, e(v))$ where $e(v)$ denotes the outgoing edge set. This tree should have the property that all nonroot, nonleaf vertices have at least 2 outgoing edges, and the root should have exactly 1 outgoing edge. We then quotient the labels of each vertex by translation and positive scaling. Operad composition is given by grafting trees.

We interpret the elements of $\mathcal{F}_n$ as "infinitesimal configurations" in $\mathbb{R}^n$, modulo the given relations. The label of the vertex adjacent to the root is the "base configuration" and, if the tree branches, we imagine that the single point labeled by that edge is actually a configuration in the tangent space, and this may repeat. It was shown by Salvatore that this operad is related to $E_n$ by a zigzag of weak equivalences [21, Proposition 3.9].

To a framed manifold $M$, we can associate a right module over $\mathcal{F}_n$:

**Definition 7.4** A point in $\mathcal{F}_M(I)$ is represented by a rooted tree with leaves labeled by $I$ such that the root $r$ is labeled by $\mathrm{Conf}(M, e(r))$. If $v$ is a nonleaf vertex adjacent to the root, it is labeled by $\mathrm{Conf}(T_p(M), e(v))$ where $p$ is the point in the root configuration labeled by the edge connecting $v$ to the root. Any nonleaf child $v'$ of $v$ is labeled by $\mathrm{Conf}(T_p(M), e(v'))$ for the same $p$, and so on. This tree should have the property that all nonroot, nonleaf vertices have at least 2 outgoing edges. We then quotient each nonroot, labeled vertex by translation and positive scaling. Module partial composition is given by grafting trees and using the framing to identify $T_p(M)$ with $\mathbb{R}^n$.

We interpret the elements of $\mathcal{F}_M$ as "infinitesimal configurations in $M$". One important observation is that both $\mathcal{F}_M(I)$ and $\mathcal{F}_n(I)$ are manifolds of dimension $n|I|$ and $n|I| - n - 1$, respectively, and the partial composites

$$\mathcal{F}_n(I) \times \mathcal{F}_n(J) \to \mathcal{F}_n(I \cup_a J) \quad \text{and} \quad \mathcal{F}_M(I) \times \mathcal{F}_n(J) \to \mathcal{F}_M(I \cup_a J)$$

are inclusions of codimension-0 portions of the boundary which together cover the boundary. For a detailed account of the topology of these completions, we refer to [23]. An immediate, and yet vital, observation is that the $\mathcal{F}_n$-module $\mathcal{F}_M$ is functorial with respect to embeddings that lie in $\mathrm{Emb}^{\mathrm{scale}}(-, -)$.

**Definition 7.5** If $M$ is a framed $n$-manifold, the $(\mathcal{F}_n)_+$-module $\mathcal{F}_{M+}$ is the objectwise one-point compactification

$$\mathcal{F}_{M+}(I) := (\mathcal{F}_M)^+.$$

The partial composites are given by one-point compactification of the partial composites of $\mathcal{F}_M$, which we note are proper maps.

This module is the module of infinitesimal configurations in $M^+$ which identifies configurations which have a point at $\infty$. To avoid visual confusion we will always denote $\mathcal{F}_M$ with a disjoint basepoint as $(\mathcal{F}_M)_+$. The following proposition is the key to understanding Koszul duality for $\mathcal{F}_n$.

**Proposition 7.6** *The map $\mathcal{F}_{(B^n)+} \to S^n \wedge \mathrm{Triv}_{(\mathcal{F}_n)_+}(\Sigma_1)$ determined by*

$$\mathcal{F}_{(B^n)+}(1) = (B^n)^+ \cong S^n$$

*is a weak equivalence of right modules which is $\otimes$-nice.*

**Proof** We must show $(\mathcal{F}_{B^n})(I)^+$ is contractible if $|I| \geq 2$. Appealing to collar neighborhoods, it suffices to pass to the subspace of configurations which have no infinitesimal components, ie the interior of $\mathcal{F}_{B^n}(I)$, together with the point at $\infty$. This is contractible by scaling to $\infty$ [2, Examples 3.14 and 3.15].[1] The $\otimes$-niceness follows since all the spaces involved are CW complexes. $\qquad\square$

**Definition 7.7** Let $\mathcal{F}^{\mathrm{red}}_{M+}$ denote right $(\mathcal{F}_n)_+$-module constructed as the quotient of $\mathcal{F}_{M+}$ by the submodule of infinitesimal configurations which do not intersect every path component of $M$.

By inspection, one finds the following combinatorial description of reduced configuration spaces of manifolds:

**Lemma 7.8** *If $M$, $N$ are connected, framed $n$-manifolds, there is an isomorphism of right modules*

$$\mathcal{F}_{M+} \otimes \mathcal{F}_{N+} \cong \mathcal{F}^{\mathrm{red}}_{(M \sqcup N)+}.$$

The Fulton–MacPherson operads and modules play an important homotopical role in the study of the $E_n$-operad because they have various cofibrancy properties. Let $\mathrm{Decom}(-)$ denote the decomposables of a right module, ie those points in the image of the right module partial composites for $O(i)$ when $i \geq 2$. The following is well known [21], but we provide a proof.

---

[1] In general, the right modules $\mathcal{F}_{M+}$ are equivalent to the right modules $\mathbb{M}_{M+}$ that Arone–Ching construct in [2, Definition 3.12]. This is because the latter are constructed via Weiss cosheafification and the former satisfy the Weiss cosheaf property seen in Section 8.

**Proposition 7.9** *If $M$ is a tame, framed $n$-manifold, the modules $\mathcal{F}_{M+}$, $(\mathcal{F}_M)_+$ are cofibrant $(\mathcal{F}_n)_+$-modules with respect to the projective model structure.*

**Proof** We deal with the case of $\mathcal{F}_{M+}$; the case of $(\mathcal{F}_M)_+$ is analogous. It is certainly sufficient to show

- any map of modules $\mathcal{F}_{M+}^{\leq i-1} \to R$ extends to $\mathcal{F}_{M+}^{\leq i-1} \vee \mathrm{Decom}(\mathcal{F}_{M+})(i)$;
- the inclusion of $\mathrm{Decom}(\mathcal{F}_{M+})(i)$ into $(\mathcal{F}_{M+})(i)$ is a $\Sigma_i$-cofibration, ie a cofibration in the projective model structure on $\Sigma_i$-spaces.

The first follows directly from the definition of $\mathcal{F}_{M+}(i)$ since the decomposables are formed by formally grafting trees.

Now note that $\mathrm{Decom}(\mathcal{F}_{M+})(i) = (\partial \mathcal{F}_M(i))^+$. Since the action of $\Sigma_i$ is free away from $\infty$, we see that $\mathrm{Decom}(\mathcal{F}_{M+})(i)/\Sigma_i \to \mathcal{F}_{M+}(i)/\Sigma_i$ is the one-point compactification of the inclusion of the boundary of a manifold. Since $M$ is tame, this is a cofibration, and so $\mathcal{F}_{M+}(I)/\Sigma_i$ can be constructed by attaching cells to $\mathrm{Decom}(\mathcal{F}_{M+})(I)/\Sigma_i$. Using covering theory, we can then lift this to a decomposition of $\mathcal{F}_{M+}(i)$ as a sequence of $\Sigma_i$-free cell attachments to $\mathrm{Decom}(\mathcal{F}_{M+})(i)$. Hence, the inclusion is a $\Sigma_i$-cofibration. $\square$

Since $\mathrm{RMod}_O(-,-)$ is a spectral enrichment of $\mathrm{RMod}_O$, applying $\Omega^\infty$ yields the standard mapping spaces of $\mathrm{RMod}_O$. By Remark 7.2, we may identify the elements of $E_n(I)$ with embeddings which pointwise scale the framings, and so there is a map

$$E_n(I)_+ \to \Omega^\infty \mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, \Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}), \quad f \mapsto \Sigma^\infty(\mathcal{F}_f \colon \mathcal{F}_{\bigsqcup_I B^n} \to \mathcal{F}_{B^n})^+.$$

**Definition 7.10** The operadic Pontryagin–Thom construction is the adjoint of the above map:

$$\mathrm{PT}(I) \colon \Sigma_+^\infty E_n(I) \to \mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, \Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}).$$

In order to relate the Pontryagin–Thom construction to Koszul duality, we must first define a reduced version.

**Definition 7.11** The reduced operadic Pontryagin–Thom construction

$$\mathrm{PT}^{\mathrm{red}}(I) \colon \Sigma_+^\infty E_n(I) \to \mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, \Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}^{\mathrm{red}})$$

is obtained by the composition of $\mathrm{PT}(I)$ with the map

$$\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, \Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}) \to \mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, \Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}^{\mathrm{red}}),$$

which is given by postcomposition with the quotient

$$\Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+} \to \Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}^{\mathrm{red}}.$$

Recall by Lemma 7.8, there is an isomorphism $\Sigma^\infty \mathcal{F}_{(\bigsqcup_I B^n)+}^{\mathrm{red}} \cong \bigotimes_I \Sigma^\infty \mathcal{F}_{(B^n)+}$, and so we can write

$$\mathrm{PT}^{\mathrm{red}}(I) \colon \Sigma_+^\infty E_n(I) \to \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{(B^n)+})(I).$$

**Proposition 7.12**  *The reduced Pontryagin–Thom maps assemble into a weak equivalence of operads*

$$\Sigma^\infty_+ E_n \xrightarrow{\simeq} \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma^\infty_+ \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{(B^n)+}).$$

**Proof**  Since the identification in Remark 7.2 sends operad composition to function composition and the Fulton–MacPherson compactifications are functorial with respect to composition of embeddings, the reduced Pontryagin–Thom construction is a map of operads.

By combining Lemma 5.7, Corollary 5.16, and Propositions 7.6–7.9,

$$\mathrm{RMod}_{\Sigma^\infty_+ \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, \Sigma^\infty \mathcal{F}^{\mathrm{red}}_{\bigsqcup_i (B^n)+}) \xrightarrow{\simeq} \mathrm{RMod}_{\Sigma^\infty_+ \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{(B^n)+}, (S^n)^{\wedge i} \wedge \mathrm{Triv}_{\Sigma^\infty_+ \mathcal{F}_n} \Sigma_i)$$

is an equivalence. Since there is an adjunction between indecomposables and trivial modules, the latter spectrum is

$$F^{\Sigma_i}(\mathrm{Indecom}(\Sigma^\infty \mathcal{F}_{(B^n)+})(i), (S^n)^{\wedge i} \wedge \Sigma^\infty_+ \Sigma_i),$$

which is equivalent to $F(\mathrm{Indecom}(\Sigma^\infty \mathcal{F}_{(B^n)+})(i), (S^n)^{\wedge i})$. For formal reasons, the indecomposables agree with $\Sigma^\infty \mathrm{Indecom}(\mathcal{F}_{(B^n)+})(i)$. These are $\Sigma^\infty \mathrm{Conf}(B^n, i)^+$ because the image of the unstable module composition is exactly $(\partial \mathcal{F}_{B^n}(i))^+$. Compiling these observations, our original map is equivalent to

$$\Sigma^\infty_+ E_n(i) \to F(\Sigma^\infty \mathrm{Conf}(B^n, i)^+, (S^n)^{\wedge i}),$$

which is given by the stable adjoint of the pairing

$$E_n(i)_+ \wedge \mathrm{Conf}(B^n, i)^+ \to ((B^n)^+)^{\wedge i} \cong (S^n)^{\wedge i}, \quad \left( \bigsqcup_{i \in I} f_i, (x_i) \right) \to f_i^+(x_i).$$

This pairing is equivalent to the duality pairing of [17, Theorem 5.6] for $B^n$ with its rank-0 normal bundle, so its stable adjoint is an equivalence. □

**Theorem 7.13**  (Self duality of $E_n$)  *There is a zigzag of weak equivalence of operads*

$$\Sigma^\infty_+ E_n \simeq \cdots \simeq s_n K(\Sigma^\infty_+ E_n).$$

**Proof**  By Proposition 7.12, it suffices to demonstrate that $\mathrm{CoEnd}(\Sigma^\infty \mathcal{F}_{(B^n)+})$ is related by a zigzag of weak equivalences of operads to $s_n K(\Sigma^\infty_+ E_n)$. By Proposition 7.6, we already know that $\Sigma^\infty \mathcal{F}_{(B^n)+}$ is weakly equivalent to a trivial module, and so, we just need to apply the homotopical bookkeeping of the previous sections to relate it to Koszul duality. Let $T$ denote a model of $\mathrm{Triv}_{\Sigma^\infty_+ \mathcal{F}_n}(\Sigma_1)$ which satisfies the requirements of Lemma 6.1. There is a zigzag of weak equivalences

$$
\begin{aligned}
\mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma^\infty_+ \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{(B^n)+}) &\simeq \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma^\infty_+ \mathcal{F}_n}}(S^n \wedge T) \quad (4.8, 4.9, 5.7, 5.16, 7.6, 7.9) \\
&\simeq S_n \wedge \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma^\infty_+ \mathcal{F}_n}}(T) \quad (5.20) \\
&\simeq s_n K(\Sigma^\infty_+ \mathcal{F}_n) \quad (6.2) \\
&\simeq s_n K(\Sigma^\infty_+ E_n) \quad (6.4).
\end{aligned}
$$

□

# 8  Extension to the modules $\mathcal{F}_M$

We now prove an analogous self duality result for the modules $\mathcal{F}_M$. If $M$ is an open subset of $\mathbb{R}^n$, the construction is practically identical. However, it is not obvious how to extend the construction to a general framed manifold. Instead we rely on the multilocality properties of $\mathcal{F}_M$ which are inherited from $\mathrm{Conf}(M, -)$. This is expressed through the language of Weiss cosheaves, ie homotopy cosheaves with respect to the covers which contain all finite sets in some open. We briefly recall some properties of Weiss cosheaves in the category $\mathrm{RMod}_O$. For a general treatment we recommend [3; 5] and refer to [18, Section 8] for the specific case of Weiss cosheaves in $\mathrm{RMod}_O$.

**Definition 8.1**  For a smooth $n$-manifold $M$, the poset $\mathrm{Disk}(M)$ has objects the open subsets of $M$ diffeomorphic to $\bigsqcup_I \mathbb{R}^n$, where $I$ is a finite set, and morphisms given by inclusion.

**Definition 8.2**  The category $\mathrm{Mfld}_n^{\mathrm{strict}}$ is the discrete category with objects the framed $n$-manifolds and morphisms given by open embeddings which preserve the framing.

Suppose we have a zigzag of weak equivalences of operads together with functors

$$O_1 \simeq \cdots \simeq O_k, \quad F_i \colon \mathrm{Mfld}_n^{\mathrm{strict}} \to \mathrm{RMod}_{O_i},$$

with compatible natural weak equivalences among their restrictions to $\mathrm{Disk}(\mathbb{R}^n)$. If the $F_i$ satisfy the Weiss cosheaf condition [4, Definition 2.20], by [18, Lemma 8.19] we have, for any framed manifold $M$,

$$F_1(M) \simeq \cdots \simeq F_k(M)$$

as right modules compatible with a, possibly different, zigzag of weak equivalences of operads

$$O_1 \simeq \cdots \simeq O_k.$$

Since homotopy colimits in $\mathrm{RMod}_O$ are computed on the underlying symmetric sequences, as both are computed pointwise, it suffices to verify the Weiss cosheaf condition on the underlying symmetric sequences. In particular, if the underlying functor of symmetric sequences is weakly equivalent to

$$M \to \Sigma_+^\infty \mathrm{Conf}(M, -)$$

then the functors are Weiss cosheaves [6, Lemma 2.5]. We conclude that since $\mathcal{F}_M$ has the homotopy type of configuration space, it suffices to demonstrate the self duality of $\Sigma_+^\infty \mathcal{F}_M$ for open subsets of $\mathbb{R}^n$ in a way that is natural with respect to inclusion.

**Definition 8.3**  For $U \subset \mathbb{R}^n$ an open subset, let $E_U$ denote the right $E_n$-module with $E_U(I)$ equal to the configurations of $I$-labeled open $n$-disks contained in $U$ with partial composites given by inserting configurations of disks.

We can think of $E_U(I)$ as a subspace of $\mathrm{Emb}^{\mathrm{scale}}\big(\bigsqcup_I B^n, U\big)$. As before, we may define a reduced Pontryagin–Thom map by taking one-point compactifications of the maps on Fulton–MacPherson modules:

$$\mathrm{PT}_U^{\mathrm{red}}(I)\colon \Sigma_+^\infty E_U(I)$$
$$\to \mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{U+}, \Sigma^\infty \mathcal{F}^{\mathrm{red}}_{(\bigsqcup_I B^n)+}) \cong \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{U+}, \Sigma^\infty \mathcal{F}_{(B^n)+})(I).$$

**Theorem 8.4** *The reduced Pontryagin–Thom construction for $U \subset \mathbb{R}^n$ is a weak equivalence of modules*

$$\Sigma_+^\infty E_U \xrightarrow{\simeq} \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{U+}, \Sigma^\infty \mathcal{F}_{(B^n)+}),$$

*compatible with the weak equivalence of operads $\Sigma_+^\infty E_n \xrightarrow{\simeq} \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{(B^n)+})$. These weak equivalences are natural with respect to inclusion.*

**Proof** The same arguments as in the proof of Proposition 7.12 show this map is a weak equivalence of modules, compatible with the reduced Pontryagin–Thom construction for operads. To address naturality, we consider the pairing

$$E_U(I)_+ \wedge \mathcal{F}_{U+}(K) \to \mathcal{F}_{(\bigsqcup_I B^n)+}(K)$$

defined as tautological pairing which applies the Pontryagin–Thom map associated to the first coordinate to the configuration in the second coordinate. The adjoints of these pairings were used to define $\mathrm{PT}_U$, and so the naturality of $\mathrm{PT}_U^{\mathrm{red}}$ with respect to inclusion follows from the commutativity of the square for $U \subset V$:

$$
\begin{array}{ccc}
E_U(I)_+ \wedge \mathcal{F}_{V+}(K) & \longrightarrow & E_V(I)_+ \wedge \mathcal{F}_{V+}(K) \\
\downarrow & & \downarrow \\
E_U(I)_+ \wedge \mathcal{F}_{U+}(K) & \longrightarrow & \mathcal{F}_{(\bigsqcup_I B^n)+}(K)
\end{array}
\qquad \square
$$

Salvatore demonstrated that the Fulton–MacPherson operad is a model of the $E_n$-operad [21, Proposition 3.9]. The proof extends directly to the right modules $E_U$ and $\mathcal{F}_U$.

**Proposition 8.5** *There is a zigzag of weak equivalences of operads*

$$E_n \simeq \cdots \simeq \mathcal{F}_n.$$

*For an open subset $U \subset \mathbb{R}^n$, there are compatible weak equivalence of modules*

$$E_U \simeq \cdots \simeq \mathcal{F}_U.$$

*These are natural with respect to inclusion.*

**Corollary 8.6** *There is a zigzag of weak equivalences of operads*

$$\Sigma_+^\infty \mathcal{F}_n \simeq \cdots \simeq s_n K(\Sigma_+^\infty \mathcal{F}_n).$$

*For an open subset $U \subset \mathbb{R}^n$, there is a compatible zigzag of weak equivalences of right modules*

$$\Sigma_+^\infty \mathcal{F}_U \simeq \cdots \simeq s_{(n,n)} K(\Sigma^\infty \mathcal{F}_{U+}).$$

**Proof**　The zigzag of weak equivalences of operads starts from the zigzag $\Sigma_+^\infty \mathcal{F}_n \simeq \Sigma_+^\infty E_n$ of Proposition 8.5 and is extended to the right to $s_n K(\Sigma_+^\infty \mathcal{F}_n)$ using the zigzag appearing in the proof of Theorem 7.13.

The zigzag of weak equivalences of right modules also starts from the zigzag $\Sigma_+^\infty \mathcal{F}_U \simeq \Sigma_+^\infty E_U$ of Proposition 8.5, and is extended to the right by the right module equivalence of Theorem 8.4:

$$\mathrm{PT}_U^{\mathrm{red}} \colon \Sigma_+^\infty E_U \to \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{U+}, \Sigma^\infty \mathcal{F}_{(B^n)+}).$$

Let $T$ denote the model of $\mathrm{Triv}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma_1)$ used in the proof of Theorem 7.13. To finish off, we extend to the right by the zigzag

$$\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}(\Sigma^\infty \mathcal{F}_{U+}, \Sigma^\infty \mathcal{F}_{(B^n)+}) \simeq \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{U+}, S^n \wedge T) \quad (5.4, 5.7, 5.16, 7.6, 7.9)$$

$$\simeq s_{(n,n)} \mathrm{CoEnd}_{\mathrm{RMod}_{\Sigma_+^\infty \mathcal{F}_n}}(\Sigma^\infty \mathcal{F}_{U+}, T) \quad (5.20)$$

$$\simeq s_{(n,n)} K(\Sigma^\infty \mathcal{F}_{U+}) \qquad\qquad (6.6). \qquad\qquad \square$$

Our discussion on Weiss cosheaves then implies the general result.

**Theorem 8.7**　(Self duality of $\mathcal{F}_M$)　*There is a zigzag of weak equivalences of operads*

$$\Sigma_+^\infty \mathcal{F}_n \simeq \cdots \simeq s_n K(\Sigma_+^\infty \mathcal{F}_n).$$

*For any framed manifold $M$, there is a compatible zigzag of weak equivalences of right modules*

$$\Sigma_+^\infty \mathcal{F}_M \simeq \cdots \simeq s_{(n,n)} K(\Sigma^\infty \mathcal{F}_{M+}).$$

# References

[1]　**G Arone**, **M Ching**, *Operads and chain rules for the calculus of functors*, Astérisque 338, Soc. Math. France, Paris (2011)　MR

[2]　**G Arone**, **M Ching**, *Manifolds, K-theory and the calculus of functors*, from "Manifolds and K-theory", Contemp. Math. 682, Amer. Math. Soc., Providence, RI (2017) 1–37　MR

[3]　**D Ayala**, **J Francis**, *Poincaré/Koszul duality*, Comm. Math. Phys. 365 (2019) 847–933　MR

[4]　**D Ayala**, **J Francis**, *A factorization homology primer*, from "Handbook of homotopy theory", CRC Press, Boca Raton, FL (2020) 39–101　MR

[5]　**P Boavida de Brito**, **M Weiss**, *Manifold calculus and homotopy sheaves*, Homology Homotopy Appl. 15 (2013) 361–383　MR

[6]　**R Campos**, **N Idrissi**, **T Willwacher**, *Configuration spaces of surfaces*, preprint (2019)　arXiv 1911.12281

[7]　**M Ching**, *Bar constructions for topological operads and the Goodwillie derivatives of the identity*, Geom. Topol. 9 (2005) 833–933　MR

[8]　**M Ching**, **P Salvatore**, *Koszul duality for topological $E_n$-operads*, Proc. Lond. Math. Soc. 125 (2022) 1–60　MR

[9]　**H Espic**, *Koszul duality for categories with a fixed object set*, preprint (2022)　arXiv 2204.03389

[10] **J Francis**, **D Gaitsgory**, *Chiral Koszul duality*, Selecta Math. 18 (2012) 27–87  MR

[11] **B Fresse**, *Koszul duality of $E_n$-operads*, Selecta Math. 17 (2011) 363–434  MR

[12] **E Getzler**, **J D S Jones**, *Operads*, *homotopy algebra and iterated integrals for double loop spaces*, preprint (1994)  arXiv hep-th/9403055

[13] **V Ginzburg**, **M Kapranov**, *Koszul duality for operads*, Duke Math. J. 76 (1994) 203–272  MR

[14] **B J Guillou**, **J P May**, *Enriched model categories and presheaf categories*, New York J. Math. 26 (2020) 37–91  MR

[15] **J Lurie**, *Higher algebra*, preprint (2017)  Available at `https://url.msp.org/Lurie-HA`

[16] **J Lurie**, *Lie algebras in stable homotopy theory*, lecture at "Homotopy theory in honor of Paul Goerss" conference (2023)

[17] **C Malin**, *An elementary proof of the homotopy invariance of stabilized configuration spaces*, Proc. Amer. Math. Soc. 151 (2023) 3635–3644  MR

[18] **C Malin**, *Koszul self-duality of manifolds*, J. Topol. 17 (2024) art. id. e12334  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 Markl**, *A compactification of the real configuration space as an operadic completion*, J. Algebra 215 (1999) 185–204  MR

[21] **P Salvatore**, *Configuration spaces with summable labels*, from "Cohomological methods in homotopy theory", Progr. Math. 196, Birkhäuser, Basel (2001) 375–395  MR

[22] **S Schwede**, **B Shipley**, *Equivalences of monoidal model categories*, Algebr. Geom. Topol. 3 (2003) 287–334  MR

[23] **D P Sinha**, *Manifold-theoretic compactifications of configuration spaces*, Selecta Math. 10 (2004) 391–428  MR

*Max Planck Institute for Mathematics*
*Bonn, Germany*

`malin@mpim-bonn.mpg.de`

# Braided multitwists

Rodrigo De Pool

We provide a characterization for multitwists satisfying the braid relation in the mapping class group of an orientable surface.

## 1 Introduction

Let $S$ be an orientable surface possibly with punctures or boundary. Let $\mathrm{Map}(S)$ be the mapping class group of $S$, that is, the group of homeomorphisms of $S$ up to homotopy. In this article we characterize *braided multitwists*, ie multitwists $\tau_A, \tau_B \in \mathrm{Map}(S)$ satisfying the braid relation $\tau_A \tau_B \tau_A = \tau_B \tau_A \tau_B$. First, let us consider some examples:

**Example 1.1** Consider $a$ and $b$ two curves in $S$ intersecting once. Denote their Dehn twists by $\delta_a$ and $\delta_b$. It is known that

$$\delta_a \delta_b \delta_a = \delta_b \delta_a \delta_b,$$

and by the same equality the inverses $\delta_a^{-1}, \delta_b^{-1}$ are also braided. It is a lemma that two Dehn twists $\delta_a, \delta_b$ are braided if and only if the curves $a$ and $b$ intersect once [5, Lemma 4.3].

Now, we construct braided multitwists that are not Dehn twists.

**Example 1.2** Let $\tau_A = \delta_{a_1} \delta_{a_2} \delta_{a_3} \delta_{a_4}^{-1}$ and $\tau_B = \delta_{b_1} \delta_{b_2} \delta_{b_3} \delta_{b_4}^{-1}$ be two multitwists along the curves $a_i, b_j$ shown in Figure 1. Since $\delta_{a_i}$ commutes with $\delta_{b_j}$ whenever $i \neq j$ and $a_i, b_i$ intersect once, we deduce $\tau_A \tau_B \tau_A = \tau_B \tau_A \tau_B$ from the previous example.

Note that in Example 1.2 the multitwists $\tau_A$ and $\tau_B$ do not share any common Dehn twist. An easy way to create new braided multitwists is to add common components with the same power. We do so in the next example:

**Example 1.3** Let $\tau_A, \tau_B$ be the multitwists of Example 1.2 and let $d$ be the curve shown in Figure 1. Then, the multitwists $\tau_A \delta_d^m$ and $\tau_B \delta_d^m$ satisfy the braid relation.

The goal of this article is to prove that braided multitwists are essentially as in Example 1.3.
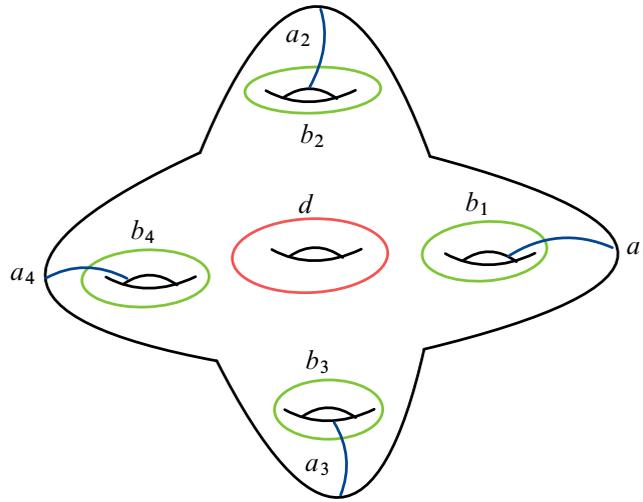
Figure 1: Genus five surface.

**Theorem 1.4** *Let $S$ be an orientable surface. If $\tau_A$, $\tau_B \in \mathrm{Map}(S)$ are two braided multitwists, then*

$$\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k} \tau_C, \quad \tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k} \tau_C,$$

*where $\tau_C$ is a common multitwist, $n_i \in \{-1, 1\}$ and the curves $a_1, \ldots, a_k$, $b_1, \ldots, b_k$ are pairwise disjoint except for $i(a_i, b_i) = 1$.*

As an application of Theorem 1.4, we can describe homomorphisms from braid groups to mapping class groups that send standard generators to multitwists. Recall the braid group on $n$-strands is given by the presentation

$$B_n = \langle \sigma_1, \ldots, \sigma_{n-1} \mid \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \; \forall i, \; [\sigma_i, \sigma_j] = 1 \; \forall |i - j| > 1 \rangle,$$

and the $\sigma_i$'s are said to be the *standard braid generators*.

A natural way to produce an embedding $\varphi \colon B_n \hookrightarrow \mathrm{Map}(S)$ is to send each standard generator $\sigma_i$ to a Dehn twist $\delta_{c_i}$, where the curves $c_i, c_j$ intersect once if $|i - j| = 1$ and are disjoint otherwise. Any such $\varphi$ (or $\varphi^{-1}$) is known as a *geometric* embedding. If we replace $\delta_{c_i}$ by a multitwist, then Theorem 1.4 yields a description of $\varphi$:

**Corollary 1.5** *Let $S$ be an orientable surface and let $B_n$ be the braid group on $n$-strands. If*

$$\varphi \colon B_n \to \mathrm{Map}(S)$$

*sends a standard generator to a multitwist, then $\varphi$ factors as $\varphi = \left( \prod_{i=1}^{k} \varphi_i \right) \circ d$, where $d \colon B_n \to \prod_{i=1}^{k} B_n$ is the diagonal homomorphism, $\varphi_i$ is a geometric embedding for $1 \leq i < k$, and $\varphi_k$ has cyclic image.*

One could hope that every embedding $B_n \hookrightarrow \mathrm{Map}(S)$ comes from a diagonal decomposition of geometric embeddings; however this is not the case. For instance, Szepietowski constructed in [6] an embedding $B_n \hookrightarrow \mathrm{Map}(S)$ sending a standard generator to a nontrivial root of a Dehn twist.

Theorem 1.4 fits into the bigger picture of results studying relations between multitwists. A closely related result is [2, Theorem 3.4] of Hamidi-Tehrani, which provides sufficient conditions for two *positive* multitwists to generate a free group. Leininger improved this result in [4, Theorem 6.1] by classifying pairs of positive multitwists generating a free group. We highlight that Theorem 1.4 for positive multitwists can be proved using Leininger's classification. However, Leininger's techniques do require the multitwists to be positive and it is not clear how to generalize them to the case of arbitrary multitwists. Thus, our proof of Theorem 1.4 uses a different approach.

The original motivation for Theorem 1.4 is to further understand homomorphisms between pure mapping class groups. We will study this application in a forthcoming paper.

## Acknowledgements

## 2 Preliminaries

Let $S$ be a connected orientable surface. By a curve on $S$ we will mean the homotopy class of an unoriented simple closed curve that does not bound a disk or a punctured disk. Brackets around an oriented curve will be used to denote the homology class of the curve.

For two curves $a$ and $b$ the (*geometric*) *intersection number* $i(a, b)$ is minimum number of intersection points between representatives of $a$ and $b$. When the curves $a$ and $b$ are oriented, the *algebraic intersection number* $\hat{\imath}([a], [b])$ is the sum of the signs at each intersection point. Importantly, this sum remains invariant regardless of the choice of representatives for $[a]$ and $[b]$.

Throughout the article, we will blur the difference between curves and their representatives. Additionally, we will often consider representatives that intersect pairwise minimally. It is a theorem that such representatives always exist (see [1, Chapter 1.2]).

If $a$ is a curve on $S$ and $A$ is a regular neighbourhood of $a$ given in (oriented) coordinates by

$$\{(h, \theta) \mid h \in [0, 1], \ \theta \in \mathbb{R}/(x \sim x + 2\pi)\},$$

we can define the twisting map $\delta_a \colon S \to S$ that sends $(h, \theta) \mapsto (h, \theta + h\theta)$ on $A$ and is the identity elsewhere. The resulting mapping class $\delta_a \in \mathrm{Map}(S)$ is the *Dehn twist* along the curve $a$. Note that as a matter of convention, we take our Dehn twists to be twists to the *right*.

A *multitwist* $\tau_A \in \mathrm{Map}(S)$ is a finite product of Dehn twists

$$\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k},$$

where the curves $a_1, \ldots, a_k$ are distinct, pairwise disjoint and $n_i \in \mathbb{Z}$. The multitwist is *positive* if $n_i \geq 0$ for every $i$ and it is *negative* if $n_i \leq 0$ for every $i$. The subindex $A$ will be used to denote the set $A = \{a_1, \ldots, a_k\}$, and we call $A$ the set of *curves* of $\tau_A$. Abusing notation, we will refer to $n_i$ as the power of $a_i$ in $\tau_A$.

The mapping class group $\mathrm{Map}(S)$ acts on the set of curves of $S$ and the action preserves the intersection number. We will write $h \cdot a$ to denote the action of the mapping class $h$ on the curve $a$. Similarly, $\mathrm{Map}(S)$ acts linearly on the homology of $S$ and preserves the algebraic intersection number. We will write $h \cdot [a]$ for the action of the mapping class $h$ on the homology class $[a]$.

The rest of this section is devoted to studying formulas for the geometric intersection number and algebraic intersection number. For other fundamental results and definitions on mapping class groups, we refer the reader to [1].

## 2.1 Geometric intersection formulas

Let $a$, $b$ and $c_j$ be curves on $S$ and consider a positive multitwist $\tau_C = \delta_{c_1}^{n_1} \cdots \delta_{c_k}^{n_k}$. The following is a well-known bound for the intersection number

$$(1) \qquad i(a, b) \geq \left| i(a, \tau_C \cdot b) - \sum_{i=1}^{k} |n_i| \cdot i(a, c_i) \cdot i(b, c_i) \right|.$$

Naturally, the same bound applies for negative multitwists. See [1, Proposition 3.4] for a proof of this inequality.

The formula above requires the multitwist to be either positive or negative. However, Theorem 1.4 deals with multitwists that (possibly) have mixed signs. The next bound on intersection number was proved by Ivanov for general multitwists; see [3, Proposition 4.2].

**Proposition 2.1** *Consider two curves $a, b$ and the multitwist $\tau_C = \delta_{c_1}^{n_1} \cdots \delta_{c_k}^{n_k}$ with $n_i \in \mathbb{Z}$. Then,*

$$i(a, b) \geq -i(a, \tau_C(b)) + \sum_{j=1}^{k} \widetilde{n}_j \cdot i(a, c_j) \cdot i(b, c_j),$$
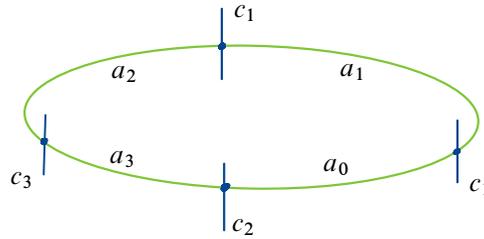
*where $\widetilde{n}_j = \max\{|n_j| - 2, 0\}$.*

Figure 2: Curve $a = \bigcup_{i=0}^{3} a_i$ and the intersecting curves $c_1, c_2, c_3$.

The proof given by Ivanov hides another formula, which will be extensively used in this article. Before introducing this formula, we need one more piece of notation.

Let $a$ be a curve and $\tau_C = \delta_{c_1}^{n_1} \cdots \delta_{c_k}^{n_k}$ a multitwist. The intersection points of $a$ with curves in $C$ define a partition $a = a_0 \cup a_1 \cup \cdots \cup a_m$, where each $a_i$ is an arc between consecutive intersection points. Since we have an arc for each intersection point, there is a total of $m + 1 = \sum_{j=1}^{k} i(a, c_j)$ arcs. Now, for each $a_i$ set $x_i = 1$ if the subarc $a_i$ intersects $c_k, c_l \in C$ and $n_k \cdot n_l > 0$. If $n_k \cdot n_l < 0$, set $x_i = 0$. We define the function

$$X(a, \tau_C) = \sum_{i=1}^{m} x_i.$$

The function $X(a, \tau_C)$ is nonnegative and it does not depend on the choice of (minimally intersecting) representatives. This is a consequence of the next proposition.

Although Ivanov does not explicitly state the following result, it follows immediately from the first part of his proof of [3, Proposition 4.2].

**Proposition 2.2** (Ivanov's hidden formula)  *Consider a curve $a$ and the multitwist $\tau_C = \delta_{c_1}^{n_1} \cdots \delta_{c_k}^{n_k}$ with $n_i \in \mathbb{Z}$. Then,*

$$i(a, \tau_C \cdot a) = \sum_{j=1}^{k} (|n_j| i(a, c_j) - 1) i(a, c_j) + X(a, \tau_C).$$

The next example illustrates how to compute $X(a, \tau_C)$ and $i(a, \tau_C \cdot a)$ using Proposition 2.2.

**Example 2.3**  Consider the curves in Figure 2 and let $\tau_C = \delta_{c_1}^{2} \delta_{c_2}^{-1} \delta_{c_3}$. We compute $X(a, \tau_C)$ and $i(a, \tau_C \cdot a)$.

The intersection points induce the partition $a = a_0 \cup a_1 \cup a_2 \cup a_3$. Since $a_0$ intersects $c_1, c_2$ and they have powers $2 \cdot -1 < 0$, then $x_0 = 0$. Since $a_1$ intersects only $c_1$ and $2 \cdot 2 > 0$, then $x_1 = 1$. Similarly, we compute $x_2 = 1$ and $x_3 = 0$. It follows that $X(a, \tau_C) = \sum_{i=0}^{3} x_i = 2$ and we obtain $i(a, \tau_C \cdot a) = 8$ using Proposition 2.2.

## 2.2 Algebraic intersection formulas

The action of a Dehn twist on the homology of a curve is described in [1, Proposition 6.3] by the formula

$$\delta_c^n \cdot [a] = [a] + n \cdot \hat{\imath}([a], [c]) \cdot [c].$$

Recall that $\operatorname{Map}(S)$ acts linearly on the homology of $S$, thus we may iteratively use this equality to obtain a formula for the action of a multitwist. Let $\tau_C = \delta_{c_1}^{n_1} \cdots \delta_{c_k}^{n_k}$ be a multitwist and $[a]$ the homology class of an oriented curve, it follows that

$$(2) \qquad \tau_C \cdot [a] = [a] + \sum_{i=1}^{k} n_k \cdot \hat{\imath}([a], [c_i]) \cdot [c_i].$$

From (2) and the bilinearity of $\hat{\imath}(\cdot, \cdot)$, we derive the following result:

**Lemma 2.4** *Consider $a$ and $b$ two oriented curves on $S$. Let $\tau_C = \delta_{c_1}^{n_1} \cdots \delta_{c_k}^{n_k}$ be a multitwist in $\operatorname{Map}(S)$. Then,*

$$\hat{\imath}(\tau_C \cdot [a], [b]) = \hat{\imath}([a], [b]) + \sum_{i=1}^{k} n_k \cdot \hat{\imath}([a], [c_i]) \cdot \hat{\imath}([c_i], [b]).$$

## 3 Proof of the main result

Consider two braided multitwists

$$\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}, \quad \tau_B = \delta_{b_1}^{m_1} \cdots \delta_{b_l}^{m_l}.$$

Recall the subindices $A$, $B$ denote the sets $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_l\}$. The first observation is that the braid relation implies

$$(\tau_A \tau_B) \tau_A (\tau_A \tau_B)^{-1} = \tau_B.$$

In particular, for each curve $a_i$ there exists a curve $b_j$ such that $\tau_A \tau_B \cdot a_i = b_j$ and $n_i = m_j$. Up to reindexing $B$, we may assume $\tau_A \tau_B \cdot a_i = b_i$ and $n_i = m_i$. Taking this into account, we write

$$\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}, \quad \tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}.$$

**Remark 3.1** We emphasize that reindexing ensures $\tau_A \tau_B \cdot a_i = b_i$. Also, we know $\tau_B \tau_A \cdot b_i \in A$, but a priori $\tau_B \tau_A \cdot b_i$ could be different from $a_i$.

As a first step towards Theorem 1.4, the next lemma shows that braided multitwists decompose as a common part and two braided multitwists sharing no curve.

**Lemma 3.2** *Let $S$ be an orientable surface. If $\tau_A$, $\tau_B \in \operatorname{Map}(S)$ are two braided multitwists, then $\tau_A = \tau_{A'} \tau_C$ and $\tau_B = \tau_{B'} \tau_C$, where $\tau_C$ is a common multitwist and $\tau_{A'}$, $\tau_{B'}$ are two braided multitwists sharing no curves.*

**Proof**  Let $C = A \cap B$ be the set of common curves of $A$ and $B$. If $C$ is empty, then we take $\tau_C = \mathrm{id}$ and we are done. If $C$ is nonempty:

We may assume $\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}$, $\tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}$ such that $\tau_A \tau_B \cdot a_i = b_i$ and $C = \{a_l, \ldots, a_k\}$. Note that every $a_i \in C$ is a curve in $B$, therefore $a_i$ is disjoint from every curve in $B$ and so $b_i = \tau_A \tau_B \cdot a_i = \tau_A \cdot a_i = a_i$. That is, we may also write $C = \{b_l, \ldots, b_k\}$. Now, we define

$$\tau_{A'} = \delta_{a_1}^{n_1} \cdots \delta_{a_{l-1}}^{n_{l-1}}, \quad \tau_{B'} = \delta_{b_1}^{n_1} \cdots \delta_{b_{l-1}}^{n_{l-1}}$$

and

$$\tau_C = \delta_{a_l}^{n_l} \cdots \delta_{a_k}^{n_k}.$$

By definition $\tau_A = \tau_{A'} \tau_C$ and $\tau_B = \tau_{B'} \tau_C$, where $\tau_C$ is a common multitwist and $A'$, $B'$ share no curves. To finish observe that $\tau_{A'}$, $\tau_{B'}$ are braided multitwists. Indeed, the relation $\tau_A \tau_B \tau_A = \tau_B \tau_A \tau_B$ implies

$$\tau_{A'} \tau_C \tau_{B'} \tau_C \tau_{A'} \tau_C = \tau_{B'} \tau_C \tau_{A'} \tau_C \tau_{B'} \tau_C.$$

Since $\tau_C$ commutes with both $\tau_{A'}$ and $\tau_{B'}$, the equation $\tau_{A'} \tau_{B'} \tau_{A'} = \tau_{B'} \tau_{A'} \tau_{B'}$ follows immediately.  □

It is worth noting that the previous lemma simplifies the proof of Theorem 1.4 to the case where the multitwists share no curves. A preliminary observation on braided multitwists sharing no curves is the following:

**Lemma 3.3**  *Let $S$ be an orientable surface and consider two braided multitwists $\tau_A$, $\tau_B \in \mathrm{Map}(S)$. If $A$ and $B$ share no curves, then every $a_i \in A$ intersects at least one curve in $B$.*

**Proof**  Take $\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}$ and $\tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}$, so that $\tau_A \tau_B \cdot a_i = b_i$.

Seeking a contradiction, suppose there is a curve $a_i \in A$ disjoint from every curve in $B$. Then, $b_i = \tau_A \tau_B \cdot a_i = \tau_A \cdot a_i = a_i$ and $a_i \in B$. But this is not possible since $A$ and $B$ share no curves.  □

The goal of the next section is to understand how the curves in $\tau_A$ and $\tau_B$ may intersect each other.

## 3.1  Curves in braided multitwists

Let $\tau_A, \tau_B \in \mathrm{Map}(S)$ be two braided multitwists. In this section we describe how a curve $a \in A$ may intersect the curves in $B$. This description will come as a consequence of Proposition 2.2.

As in Section 3, we write $\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}$ and $\tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}$, so that $\tau_A \tau_B \cdot a_i = b_i$. Using Ivanov's formula, we deduce

$$(3) \qquad i(a_i, b_i) = i(a_i, \tau_A \tau_B \cdot a_i) = i(a_i, \tau_B \cdot a_i) = \sum_{j=1}^{k} (|n_j| i(b_j, a_i) - 1) i(b_j, a_i) + X(a_i, \tau_B).$$

In Table 1 we list the possible (nonnegative) values of $i(a_i, b_i)$, $X(a, \tau_B)$ and $|n_i|$ for which (3) is satisfied. Notice each curve $a_i \in A$ is of a single type in Table 1. We collect these facts in the following proposition:

| type | $i(a_i, b_i)$ | $|n_i|$ | $X(a_i, \tau_B)$ | intersecting curves |
|------|-----------|---------|--------------|---------------------|
| 1 | 0 | — | 0 | if $i(a_i, b_i) \neq 0$ and $i \neq j$, then $i(a_i, b_j) = 1$ and $|n_j| = 1$ |
| 2 | 1 | 2 | 0 | if $i(a_i, b_i) \neq 0$ and $i \neq j$, then $i(a_i, b_j) = 1$ and $|n_j| = 1$ |
| 3 | 1 | 1 | 1 | if $i(a_i, b_j) \neq 0$, then $i(a_i, b_j) = 1$ and $|n_j| = 1$ |
| 4 | 1 | 1 | 0 | there exists $b_l$ with $i(a_i, b_l) = 1$ and $|n_l| = 2$, while if $b_j \in B \setminus \{b_l\}$ and $i(a_i, b_j) \neq 0$, then $i(a_i, b_j) = 1$ and $|n_j| = 1$ |
| 5 | 2 | 1 | 0 | if $i(a_i, b_j) \neq 0$ and $i \neq j$, then $i(a_i, b_j) = 1$ and $|n_j| = 1$ |

Table 1: Description of curves $a_i \in A$ in braided multitwists $\tau_A, \tau_B$.

**Proposition 3.4** *Let $S$ be an orientable surface and consider two braided multitwists $\tau_A, \tau_B \in \mathrm{Map}(S)$. If*

$$\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}, \quad \tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}$$

*and $\tau_A \tau_B \cdot a_i = b_i$. Then, every curve $a_i \in A$ is of Type 1, 2, 3, 4 or 5 (see Table 1).*

Notice Proposition 3.4 already provides some restrictions on the curves $A \cup B$. For instance, we immediately know that $i(a, b) \in \{0, 1, 2\}$ for any two curves $a, b \in A \cup B$.

## 3.2 Reducing braided multitwists

Lemma 3.2 reduces the proof of Theorem 1.4 to the case where braided multitwists $\tau_A, \tau_B$ share no curves. In this section we further simplify $\tau_A, \tau_B$ to the case where every curve in $A$ intersects at least two curves in $B$. To start, the following lemma describes curves $a \in A$ that intersect a single curve in $B$.

**Lemma 3.5** *Let $S$ be an orientable surface and $\tau_A, \tau_B \in \mathrm{Map}(S)$ two braided multitwists. If $a \in A$ intersects only one curve $b \in B$. Then, $b = \tau_A \tau_B \cdot a$, $i(a, b) = 1$ and $|n| = 1$, where $n$ is the power of $a$ in $\tau_A$. Moreover, $b$ is disjoint from every curve in $A \setminus \{a\}$.*

**Proof** We may assume $\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}$ and $\tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}$, so that $\tau_A \tau_B \cdot a_i = b_i$. Further assume $a = a_1$ and $n = n_1$.

If $a_1$ intersects a single curve, then $X(a_1, \tau_B) > 0$. By Proposition 3.4 the curve $a_1$ is of type 3. It follows that $X(a_1, \tau_B) = 1$, $|n_1| = 1$ and $i(a_1, b_1) = 1$. Since $a_1$ intersects a single curve $b \in B$, then $b = b_1$. This establishes the first part of the lemma.

Recall that $\tau_A \tau_B \tau_A \cdot a_i = b_i$ and $\tau_A \tau_B \tau_A \cdot B = A$. Therefore, if $a_1$ intersects a single curve in $B$, then $b_1$ intersects a single curve in $A$. $\qquad \square$

If the curve $a_i$ is under the conditions of the previous lemma, the pair $a_i, b_i$ can be "deleted" from the multitwists $\tau_A, \tau_B$. Indeed, consider

$$\tau_{A'} = \delta_{a_1}^{n_1} \cdots \delta_{a_{i-1}}^{n_{i-1}} \cdot \delta_{a_{i+1}}^{n_{i+1}} \cdots \delta_{a_k}^{n_k}, \quad \tau_{B'} = \delta_{b_1}^{n_1} \cdots \delta_{b_{i-1}}^{n_{i-1}} \cdot \delta_{b_{i+1}}^{n_{i+1}} \cdots \delta_{b_k}^{n_k}.$$

Then, it is easy to verify that $\tau_{A'} \tau_{B'} \tau_{A'} = \tau_{B'} \tau_{A'} \tau_{B'}$.

The previous process simplifies the braided multitwists $\tau_A, \tau_B$ by eliminating the components that are clearly braided and do not interact with any other curves in $\tau_A, \tau_B$. By simplifying until there are no such pairs, we obtain *reduced* multitwists:

**Definition 3.6** (Reduced multitwists)  Let $\tau_A, \tau_B \in \mathrm{Map}(S)$ be two braided multitwists. We will say $\tau_A, \tau_B$ are *reduced* if every curve $a \in A$ intersects at least two distinct curves in $B$.

**Remark 3.7**  If $\tau_A, \tau_B$ are reduced multitwists, then $A, B$ have no common curves.

A key ingredient in the proof of Theorem 1.4 is that reduced multitwists are trivial. We record this fact in the next proposition, and dedicate most of the next sections to proving it.

**Proposition 3.8**  *Let $S$ be an orientable surface. If $\tau_A, \tau_B$ are reduced multitwists, then*

$$\tau_A = \tau_B = 1 \in \mathrm{Map}(S).$$

To prove Proposition 3.8 we will need to understand the action of $\tau_A, \tau_B$ on the curves $A \cup B$.

## 3.3 Action on curves and their homology classes

Let $\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}$ and $\tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k}$ be two braided multitwists. We will consider the action of $f = \tau_B \tau_A \tau_B$ on the curves $A \cup B$ and the action on their homology classes.

Recall that up to reindexing $B$, we may assume $f \cdot a_i = b_i$. We will write $a_{i_1}$ for the image $f \cdot b_i = a_{i_1}$ and, in general, we write $a_{i_k}$ for the image $f \cdot b_{i_{k-1}} = a_{i_k}$. The orbits of curves will be denoted by

$$\mathrm{Orb}_f(a_i) = \{f^k \cdot a_i \mid k \in \mathbb{Z}\}.$$

The action of $f$ by conjugation on $\tau_A$ and $\tau_B$ permutes the curves $A \cup B$ and preserves the powers of the associated Dehn twists. In other words, any two curves $a, b \in A \cup B$ in the same $f$-orbit have associated Dehn twists with the same powers in $\tau_A, \tau_B$. As a consequence, two curves $a, a' \in A$ in the same $f$-orbit have the same type in Table 1.

We will also consider the action of $f$ on the homology classes of curves in $A \cup B$. To achieve this, we need to assign orientations to the curves in $A \cup B$: for each orbit $\mathrm{Orb}_f(\cdot)$ take a representative of the orbit $a \in \mathrm{Orb}_f(\cdot)$ and fix an arbitrary orientation on $a$. Then, any other curve $c \in \mathrm{Orb}_f(a)$ in the orbit inherits

an induced orientation. This orientation comes from considering the smallest $k \in \mathbb{N}$ such that $f^k \cdot a = c$. Since $a$ is oriented it induces an orientation on $f^k \cdot a$, which we consider as the orientation of $c$.

The selected orientations allows us to talk about homology classes $[c]$ for curves $c \in A \cup B$. The following diagram schematizes our choice of orientations:

$$[a_i] \xrightarrow{f} [b_i] \xrightarrow{f} [a_{i_1}] \xrightarrow{f} [b_{i_1}] \xrightarrow{f} [a_{i_2}] \xrightarrow{f} \cdots \xrightarrow{f} [b_{i_m}],$$

where the orientation $[a_i]$ is arbitrary.

**Remark 3.9**   Notice there could be an $m \in \mathbb{N}$ such that $f^m \cdot [a_i] = -[a_i]$.

The next step is to show that the orbit of every curve $c \in A \cup B$ has size two, ie $|\mathrm{Orb}_f(c)| = 2$.

### 3.4   Orbits have size two

**Proposition 3.10**   *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. Then, $|\mathrm{Orb}_f(c)| = 2$ for every curve $c \in A \cup B$.*

To prove Proposition 3.10 we treat each type of curve $a_i \in A$ separately (see Table 1). Thus, this proof is split among various lemmas and propositions below.

**Lemma 3.11**   *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a curve of type 5, then $|\mathrm{Orb}_f(a_i)| = 2$.*

**Proof**   Looking for a contradiction, assume $|\mathrm{Orb}_f(a_i)| > 2$. In that case, there exists $b \in \mathrm{Orb}_f(a_i)$ such that $f \cdot b = a_i$ and $b \neq b_i$. Notice

$$i(b, a_i) = i(f \cdot b, f \cdot a_i) = i(a_i, b_i) = 2.$$

Thus, $a_i$ has intersection two with two distinct curves $b, b_i \in B$. But this is not possible (see Table 1 for type 5 curves).   □

Before continuing with other cases, we require the following lemma.

**Lemma 3.12**   *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. For any curves $a_i \in A$ and $b \in B$, we have that*

$$\hat{\imath}([a_i], [b]) = -\hat{\imath}([a_i], f^2 \cdot [b]).$$

*In particular, if $\hat{\imath}([a_i], [b]) = \pm 1$ then $i(a_i, f^{2k} \cdot b) = 1$ for every $k \in \mathbb{Z}$.*

In words, a curve $a_i \in A$ has the same algebraic intersection number (up to sign) with every curve in $\mathrm{Orb}_f(b) \cap B$.

**Proof** Let $a \in A$ be the curve $a = \tau_B \tau_A \cdot b$. Note that $\tau_B \tau_A \cdot [b] = \pm[a]$ and so it follows that

$$0 = \hat{\imath}([a_i], [a]) = \hat{\imath}([a_i], \pm \tau_B \tau_A \cdot [b]) = \hat{\imath}([a_i], \tau_B \tau_A \cdot [b]).$$

Now, by (2) we have

$$0 = \hat{\imath}([a_i], \tau_B \tau_A \cdot [b]) = \hat{\imath}\left([a_i], \tau_A \cdot [b] + \sum_{b_k \in B} n_k \hat{\imath}(\tau_A \cdot [b], [b_k])[b_k]\right).$$

From here, linearity yields

(4)
$$\hat{\imath}([a_i], [b]) = -\sum_{b_k \in B} n_k \hat{\imath}(\tau_A \cdot [b], [b_k]) \hat{\imath}([a_i], [b_k]).$$

On the other hand, using the braid relation

$$
\begin{aligned}
\hat{\imath}([a_i], f^2 \cdot [b]) &= \hat{\imath}([a_i], (\tau_A \tau_B \tau_A)(\tau_B \tau_A \tau_B) \cdot [b]) \\
&= \hat{\imath}([a_i], \tau_B \tau_A \tau_B \tau_A \cdot [b]) \\
&= \hat{\imath}([a_i], \tau_B^2 \tau_A \cdot [b]) \\
&= \hat{\imath}\left([a_i], \tau_A \cdot [b] + 2 \sum_{b_k \in B} n_k \hat{\imath}(\tau_A \cdot [b], [b_k])[b_k]\right) \\
&= \hat{\imath}([a_i], [b]) + 2 \sum_{b_k \in B} n_k \hat{\imath}(\tau_A \cdot [b], [b_k]) \hat{\imath}([a_i], [b_k]).
\end{aligned}
$$

Substituting (4) in the previous equality, we obtain

$$\hat{\imath}([a_i], f^2 \cdot [b]) = -\hat{\imath}([a_i], [b]).$$

For the second part of the lemma, it is enough to observe that Proposition 3.4 implies $i(a, b) \in \{0, 1, 2\}$ for any two curves $a, b \in A \cup B$. Thus, $\hat{\imath}(a_i, b) = \pm 1$ implies $i(a_i, b) = 1$ and, since the algebraic intersection number only changes sign, we conclude $i(a_i, f^{2k} \cdot b) = 1$. $\qquad\square$

The next lemma proves Proposition 3.10 for type 2 curves $a_i \in A$.

**Lemma 3.13** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 2 curve, then $|\mathrm{Orb}_f(a_i)| = 2$.*

**Proof** Recall we denoted $b_i = f \cdot a_i$, $a_{i_1} = f \cdot b_i$ and $b_{i_1} = f \cdot a_{i_1}$. To proceed by contradiction, we assume $|\mathrm{Orb}_f(a_i)| > 2$. In this case, $b_{i_1} \neq b_i$.

Since $a_i$ is a type 2 curve, then $i(a_i, b_i) = 1$ and it follows from Lemma 3.12 that $i(a_i, b_{i_1}) = 1$. Moreover, $a_i, b_i$ and $b_{i_1}$ have exponent $|n_i| = 2$ in $\tau_A, \tau_B$, since $a_i$ is of type 2 and the three curves are in the same orbit. Summarizing, $a_i \in A$ is a curve that intersects two distinct curves $b_i, b_{i_1} \in B$ both having exponent $|n_i| = 2$ in $\tau_B$. However, such $a_i$ cannot be a curve in a braided multitwist (see Table 1). $\qquad\square$
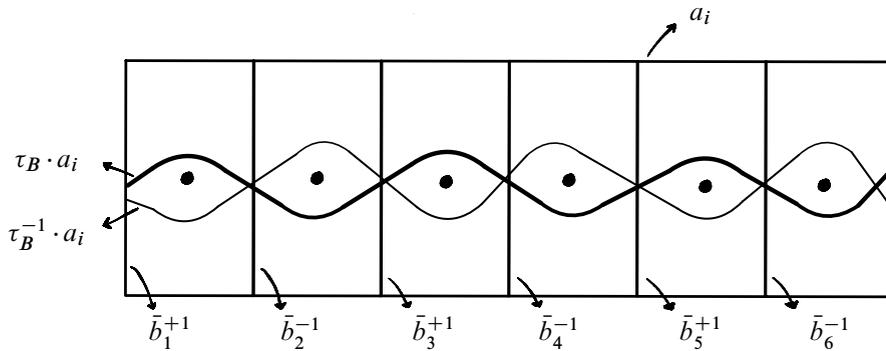
Figure 3: Planar torus $T_i$ for $i(a_i, b_i) = 0$.

From here, the cases of Proposition 3.10 become more subtle. The next result deals with type 1 curves.

**Lemma 3.14** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 1 curve, then $|\mathrm{Orb}_f(a_i)| = 2$.*

**Proof** Consider the curve $a_{i_1} = f^2 \cdot a_i$ and the set of curves $B_i = \{b \in B \mid i(a_i, b) \neq 0\}$. Given that $a_i$ is a type 1 curve, we know every $b \in B_i$ has exponent $\pm 1$ in $\tau_B$ and satisfies $i(a_i, b) = 1$ (see Table 1). Furthermore, $X(a_i, \tau_B) = 0$ and it follows that $|B_i|$ is even.

Let $T_i$ be an open regular neighbourhood of $a_i \cup \bigcup_{b \in B_i} b$. Note the subsurface $T_i$ is homeomorphic to a torus with punctures. Since $X(a_i, \tau_B) = 0$, if two curves $\bar{b}_j, \bar{b}_{j+1} \in B_i$ bound a once punctured annulus in $T_i$, then their powers in $\tau_B$ have opposite signs. We represent $T_i$ as a planar torus in Figure 3, where we denote by $\bar{b}_j$ the curves in $B_i$ and the exponents correspond to those in $\tau_B$.

Given that $a_i$ is a type 1 curve, we have that $i(a_i, b) \in \{0, 1\}$ for every $b \in B$. Since $a_{i_1}$ is in the orbit of $a_i$, then $a_{i_1}$ is also of type 1 and $i(a_{i_1}, b) \in \{0, 1\}$ for every $b \in B$. It is now a consequence of Lemma 3.12 that $i(a_i, b) = i(a_{i_1}, b)$ for every $b \in B$. In particular, we know that $i(a_{i_1}, b) = i(a_i, b) = i(a_i, b_i) = 0$ for every $b \in \mathrm{Orb}_f(a_i)$. This implies

$$\begin{cases} i(a_{i_1}, \tau_B \cdot a_i) = i(a_{i_1}, b_i) = 0, \\ i(a_{i_1}, \tau_B^{-1} \cdot a_i) = i(a_i, \tau_B \cdot a_{i_1}) = i(a_i, b_{i_1}) = 0. \end{cases}$$

Notice the conditions ensure that $a_{i_1}$ is either contained in $T_i$ or disjoint from $T_i$. Indeed, any curve $c$ that intersects $T_i$ and is not contained in $T_i$ either satisfies $i(c, \tau_B \cdot a_i) \neq 0$ or $i(c, \tau_B^{-1} \cdot a_i) \neq 0$ (see Figure 3). Now, $i(a_{i_1}, \bar{b}_1) = i(a_i, \bar{b}_1) = 1$ implies $a_{i_1}$ intersects $T_i$ and therefore $a_{i_1}$ is contained in $T_i$.

To finish, we note that the only curve $a_{i_1}$ in $T_i$ satisfying $i(a_{i_1}, \tau_B \cdot a_i) = i(a_{i_1}, \tau_B^{-1} \cdot a_i) = 0$ is the curve $a_{i_1} = a_i$. So we conclude $|\mathrm{Orb}_f(a_i)| = 2$. $\square$

The next lemma proves Proposition 3.10 for type 4 curves.

**Lemma 3.15** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 4 curve, then $|\mathrm{Orb}_f(a_i)| = 2$.*
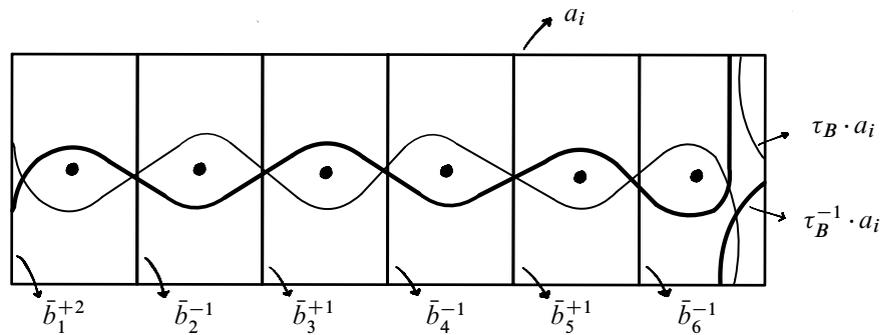
Figure 4: Planar torus $T_i$ for $i(a_i, b_i) = 1$ and $X(a_i, \tau_B) = 0$.

**Proof** The strategy of the proof is analogous to that of Lemma 3.14. Consider $a_{i_1} = f^2 \cdot a_i$ and the set of curves $B_i = \{b \in B \mid i(b, a_i) \neq 0\}$. Since $a_i$ is of type 4, it follows that $i(b, a_i) = 1$ for every $b \in B_i$, that $|B_i|$ is even and every curve in $B_i$ has power $\pm 1$ in $\tau_B$, except for a single curve which has power $\pm 2$ (see Table 1).

Let $T_i$ be an open regular neighbourhood of $a_i \cup \bigcup_{b \in B_i} b$. The above paragraph implies $T_i$ is a torus with punctures. Since $X(a_i, \tau_B) = 0$, then two curves $\bar{b}_j, \bar{b}_k \in B_i$ bounding a single punctured annulus in $T_i$ have powers of opposite sign in $\tau_B$. In Figure 4 we represent $T_i$ as a planar torus, where we label $\bar{b}_j$ the curves in $B_i$ and the exponents correspond to those in $\tau_B$.

Given that $a_i$ and $a_{i_1}$ are type 4 curves, then $i(a_i, b), i(a_{i_1}, b) \in \{0, 1\}$ for every $b \in B$. Thus, Lemma 3.12 implies $i(a_i, b) = i(a_{i_1}, b)$ for every $b \in B$. It follows $a_{i_1}$ is a curve satisfying

$$\begin{cases} i(a_{i_1}, \bar{b}_j) = 1, \\ i(a_{i_1}, a_i) = 0, \\ i(a_{i_1}, \tau_B \cdot a_i) = i(a_{i_1}, b_i) = 1, \\ i(a_{i_1}, \tau_B^{-1} \cdot a_i) = i(a_i, \tau_B \cdot a_{i_1}) = i(a_i, b_{i_1}) = 1. \end{cases}$$

The only curve satisfying above equations is $a_{i_1} = a_i$ (see Figure 4). One may check this fact by considering the point of intersection between $a_{i_1}$ and $\bar{b}_k$, where $\bar{b}_k \in B_i$ is the only curve in $B_i$ with exponent two. Then, by inspection, it is clear $a_{i_1}$ is contained in $T_i$. Readily, we obtain that $a_{i_1}$ and $a_i$ are homotopic, that is, $a_{i_1} = a_i$. Thus, $|\mathrm{Orb}_f(a_i)| = 2$. $\qquad\square$

Lastly, we consider the longer case of type 3 curves. We work through this case in the next three lemmas.

**Lemma 3.16** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 3 curve and $|\mathrm{Orb}_f(a_i)| > 2$, then there exist $b_k \in B$ that intersects $a_i$ and satisfies $|\mathrm{Orb}_f(b_k)| = 2$. Also, any other $b \in B \setminus \{b_k\}$ intersecting $a_i$ satisfies $|\mathrm{Orb}_f(b)| = 4$.*

**Proof** Let $B_i = \{b \in B \mid i(a_i, b) \neq 0\}$ and $T_i$ an open regular neighbourhood of the curves $a_i \cup \bigcup_{b \in B_i} b$. Given that $a_i$ is of type 3, we deduce $|B_i|$ is odd (see Table 1) and $T_i$ is a torus with $|B_i|$ punctures.
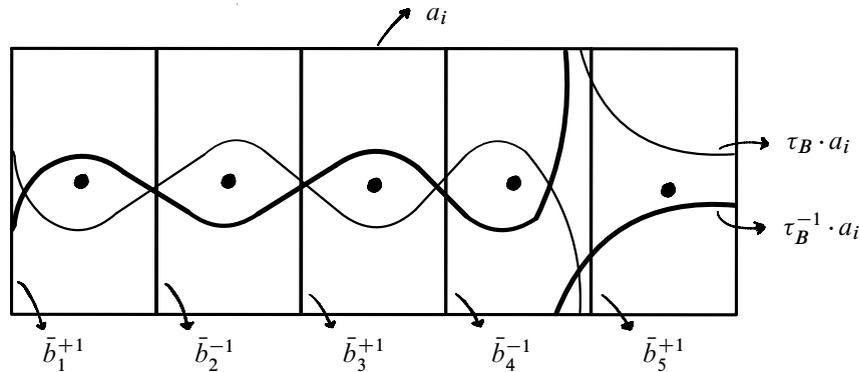
Figure 5: The planar torus $T_i$ for $i(a_i, b_i) = 1$ and $X(a_i, \tau_B) = 1$.

The intersection points of $a_i$ with curves in $B_i$ induce a cyclic order in $B_i$. Considering this order, we may write $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_{|B_i|}$ to denote the curves in $B_i$. Since $X(a_i, \tau_B) = 1$, we can arrange the subindices in $\bar{b}_j$ so that the powers in $\tau_B$ alternate sign. Even though the signs alternate, the curves $\bar{b}_1$ and $\bar{b}_{|B_i|}$ have the same sign since $|B_i|$ is odd. In Figure 5 we represent $T_i$ as a planar torus.

Let $a_{i_1} = f^2 \cdot a_i \in A$. By hypothesis $a_{i_1} \neq a_i$. As in previous lemmas, we may use Proposition 3.4 and Lemma 3.12 to check that $a_{i_1}$ satisfies the equations

$$\begin{cases} i(a_{i_1}, a_i) = 0, \\ i(a_{i_1}, b) = i(a_i, b) \quad \text{for all } b \in B, \\ i(a_{i_1}, \tau_B \cdot a_i) = 1, \\ i(a_{i_1}, \tau_B^{-1} \cdot a_i) = 1. \end{cases}$$

From these conditions it follows that either $a_{i_1} = a_i$ or $a_{i_1} \cap T_i$ has a representative as in Figure 6. Namely, $a_{i_1}$ has a representative in minimal position with curves in $A \cup B$, therefore disjoint from $a_i$ and intersecting once every $b \in B_i$; also, $a_{i_1} \cap T_i$ is an arc with endpoints in a single puncture of $T_i$ and $T_i \setminus (a_i \cup a_{i_1})$ is the union of an annulus and a $|B_i| - 1$ punctured annulus. To check that $a_{i_1} = a_i$ or $a_{i_1} \cap T_i$ has a representative as in Figure 6, we may consider the point of intersection of $a_{i_1}$ with $\bar{b}_{|B_i|}$ (see Figure 5). Then, by inspection, we see that a curve $a_{i_1}$ satisfying above equations must be as claimed.

(We warn the reader that in the next paragraphs we silently use the Alexander method [1, Proposition 2.8]. We are taking representatives of $a_i, a_{i_1}$, every $b \in B_i$ and a representative of $f^2$, so that $f^2$ permutes the curves $b \in B_i$ and sends $a_i$ to $a_{i_1}$.)

Now, denote $\gamma_i$ the subarc of $a_i$ that goes from $a_i \cap \bar{b}_{|B_i|}$ to $a_i \cap \bar{b}_1$ and is disjoint from the rest of $\bar{b}_j$'s. In the same style, define $\gamma_{i_1}$ to be the subarc of $a_{i_1}$ going from $a_{i_1} \cap \bar{b}_{|B_i|}$ to $a_{i_1} \cap \bar{b}_1$ and is disjoint from the rest of $\bar{b}_j$'s (see Figure 6). Observe that $f^2 \cdot \gamma_i$ is a subarc of $a_{i_1}$ with interior disjoint from the curves $\bar{b}_j$'s. The endpoints of $f^2 \cdot \gamma_i$ are points of intersection of $a_{i_1}$ with $f^2 \cdot \bar{b}_1$ and with $f^2 \cdot \bar{b}_{|B_i|}$. Since $\bar{b}_1$ and $\bar{b}_{|B_i|}$ have the same power in $\tau_B$ and $f^2$ preserves the exponents, the curves $f^2 \cdot \bar{b}_1$ and $f^2 \cdot \bar{b}_{|B_i|}$ have the same exponent. But the only subarc of $a_{i_1}$ as above is $\gamma_{i_1}$, that is, $\gamma_{i_1}$ is the only subarc of $a_{i_1}$
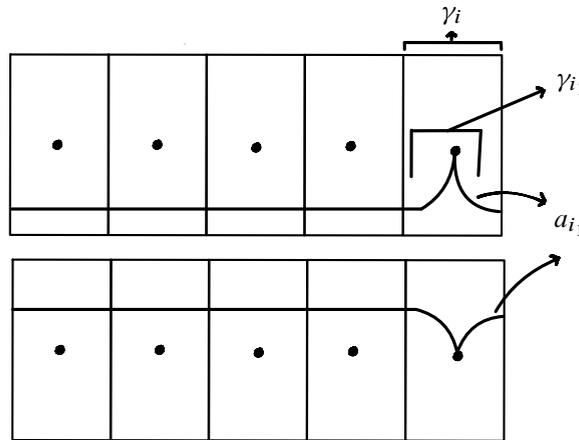
Figure 6: The two possible arcs for $a_{i_1} \cap T_i$.

between consecutive intersection points such that the intersecting curves have the same power. Thus, $f^2 \cdot \gamma_i = \gamma_{i_1}$ and either $f^2$ fixes $\bar{b}_1$ and $\bar{b}_{|B_i|}$, or it interchanges them.

If $f^2$ fixes both $\bar{b}_1$ and $\bar{b}_{|B_i|}$, we have that $f^2$ fixes every $\bar{b}_j$. To see this, consider the subarc of $a_i$ from $\bar{b}_1$ to $\bar{b}_2$ (disjoint from every other $\bar{b}_j$). Naturally, $f^2$ must send this subarc to the subarc of $a_{i_1}$ from $\bar{b}_1$ to $\bar{b}_2$. Since $f^2$ fixes $\bar{b}_1$, then it also fixes $\bar{b}_2$. By repeating the argument on adjacent subarcs, we conclude $f^2$ fixes every $\bar{b}_j$. However, this implies $|\mathrm{Orb}_f(\bar{b}_j)| = 2$ for every $\bar{b}_j \in B_i$ and, since $b_i \in B_i$, this contradicts the hypothesis $|\mathrm{Orb}_f(a_i)| > 2$. Thus, $f^2$ cannot fix $\bar{b}_1$ and $\bar{b}_{|B_i|}$.

If, on the contrary, $f^2$ interchanges $\bar{b}_1$ and $\bar{b}_{|B_i|}$, ie $f^2 \cdot \bar{b}_1 = \bar{b}_{|B_i|}$ and $f^2 \cdot \bar{b}_{|B_i|} = \bar{b}_1$. Again, using the previous argument on subarcs, it is an easy exercise to check that $f^2 \cdot \bar{b}_j = \bar{b}_{|B_i|+1-j}$ and $f^2 \cdot \bar{b}_{|B_i|+1-j} = \bar{b}_j$. Therefore, $|\mathrm{Orb}_f(b)| = 4$ for every $b \in B_i \setminus \{\bar{b}_{\frac{1}{2}(|B_i|+1)}\}$ and $|\mathrm{Orb}_f(\bar{b}_{\frac{1}{2}(|B_i|+1)})| = 2$. The $b_k$ of the statement corresponds to $b_k = \bar{b}_{\frac{1}{2}(|B_i|+1)}$. $\qquad\square$

**Lemma 3.17** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 3 curve and $|\mathrm{Orb}_f(a_i)| > 2$, then there exists a unique $b_k \in B$ that intersects $a_i$ and such that $|\mathrm{Orb}_f(b_k)| = 2$. Furthermore, let $a_k$ be the curve $f \cdot a_k = b_k$. Then $i(a_i, b) = i(a_k, b)$ for every $b \in B$.*

**Proof** The first part of the lemma follows from Lemma 3.16. We only need to see that $i(a_i, b) = i(a_k, b)$ for every $b \in B$.

Let $B_i = \{b \in B \mid i(b, a_i) \neq 0\}$ and $B_k = \{b \in B \mid i(b, a_k) \neq 0\}$. First, we will see $B_i \subset B_k$:

Let $b_j \in B_i$. From Table 1, we know that $i(a_i, b_j) = 1$. Now, using Lemma 3.12 we get $i(b_i, a_j) = 1$ and thus

$$(5) \quad \pm 1 = \hat{\imath}([b_i], [a_j]) = \hat{\imath}(\tau_A \tau_B \cdot [a_i], [a_j]) = \hat{\imath}(\tau_B \cdot [a_i], [a_j]) = \sum_{b_l \in B} n_l \hat{\imath}([a_i], [b_l]) \hat{\imath}([b_l], [a_j])$$

$$= \sum_{b_l \in B_i \cap B_j} n_l \hat{\imath}([a_i], [b_l]) \hat{\imath}([b_l], [a_j]).$$

Now, considering Lemma 3.16, we can partition the set $B_i$ as

$$B_i = B \cap \left( \mathrm{Orb}_f(\bar{b}_2) \sqcup \mathrm{Orb}_f(\bar{b}_3) \sqcup \cdots \sqcup \mathrm{Orb}_f(\bar{b}_{\frac{1}{2}(|B_i|-1)}) \sqcup \{b_k\} \right),$$

where each orbit has two curves in $B$ (except for $\{b_k\}$). This induces a partition of $B_i \cap B_j$ as disjoint union of $\mathrm{Orb}_f(\bar{b}_m) \cap B_j$ and $\{b_k\} \cap B_j$. We are looking to prove $b_k \in B_j$, so seeking a contradiction assume $b_k \notin B_j$. By means of Lemma 3.12, we have

$$n_l \, \hat{\imath}([a_i],[b_l]) \hat{\imath}([b_l],[a_j]) = n_l \, \hat{\imath}([a_i], f^2 \cdot [b_l]) \hat{\imath}(f^2 \cdot [b_l],[a_j]) = \pm n_l \, \hat{\imath}([a_i],[f^2 \cdot b_l]) \hat{\imath}([f^2 \cdot b_l],[a_j]).$$

Note that for $\bar{b}_m \neq b_k$ each orbit $B_j \cap \mathrm{Orb}_f(\bar{b}_m)$ has two elements, which contribute either as

$$2 \cdot n_l \, \hat{\imath}([a_i],[b_l]) \hat{\imath}([b_l],[a_j])$$

or as zero to the sum. Therefore, if $b_k \notin B_j$ the sum

$$\sum_{b_l \in B_i \cap B_j} n_l \, \hat{\imath}([a_i],[b_l]) \hat{\imath}([b_l],[a_j])$$

is even. But this clearly contradicts that it is equal to $\pm 1$ (see equation (5)). Hence, $b_k \in B_j$ and even more the argument yields that $n_k \, \hat{\imath}([a_i],[b_k]) \hat{\imath}([b_k],[a_j]) = \pm \hat{\imath}([b_k],[a_j])$ is odd. Given that $i(b_k, a_j) \in \{0,1,2\}$, it follows that $i(b_k, a_j) = 1$ and by Lemma 3.12 we have $i(a_k, b_j) = 1$. We conclude $1 = i(a_i, b) = (a_k, b)$ for every $b \in B_i$ and therefore $B_i \subset B_k$.

To complete the lemma, we need to see $B_k \subset B_i$. Pursuing a contradiction, assume there exists $b_j \in B$ with $i(a_k, b_j) \neq 0$ and $i(a_i, b_j) = 0$. First, observe that $i(a_i, b_i) = i(a_i, b_k) = 1$ implies $j \neq i$ and $j \neq k$. Now, $i(a_k, b_j) = 1$ follows from $j \neq k$ (see Table 1). Moreover, $0 = i(a_i, b_j) = i(b_i, a_{j_1})$ and by Lemma 3.12 we have $\hat{\imath}([b_i],[a_j]) = 0$. After this, consider

$$0 = \hat{\imath}([b_i],[a_j]) = \hat{\imath}(\tau_A \tau_B \cdot [a_i],[a_j]) = \sum_{b_l \in B_i} n_l \, \hat{\imath}([a_i],[b_l]) \hat{\imath}([b_l],[a_j]) = \sum_{b_l \in B_i \cap B_j} n_l \, \hat{\imath}([a_i],[b_l]) \hat{\imath}([b_l],[a_j]).$$

Replicating the argument above yields that the previous sum is odd. Indeed, each orbit $\mathrm{Orb}_f(\bar{b}_m) \cap B_j$ contributes an even number to the sum, plus

$$n_k \, \hat{\imath}([a_i],[b_k]) \hat{\imath}([b_k],[a_j]) = \pm 1.$$

But the sum cannot be odd and equal to zero. Thus, no such $b_j$ can exist. In other words, $B_k = B_i$ and it follows that $i(a_i, b) = i(a_k, b)$ for every $b \in B$. $\qquad\square$

Finally, we prove Proposition 3.10 for type 3 curves.

**Lemma 3.18** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 3 curve, then $|\mathrm{Orb}_f(a_i)| = 2$.*

**Proof** Assume $|\mathrm{Orb}_f(a_i)| > 2$. By Lemma 3.17, there exist $b_k \in B_i$ with $|\mathrm{Orb}_f(a_k)| = 2$ and $i(a_k, b) = i(a_i, b)$ for every $b \in B$. Since $|B_i| = |B_k|$ is odd, we know from Table 1 that $a_k$ is a type 3 curve and so $i(a_k, b_k) = 1$, $|n_k| = 1$ and $X(a_k, \tau_B) = 1$.
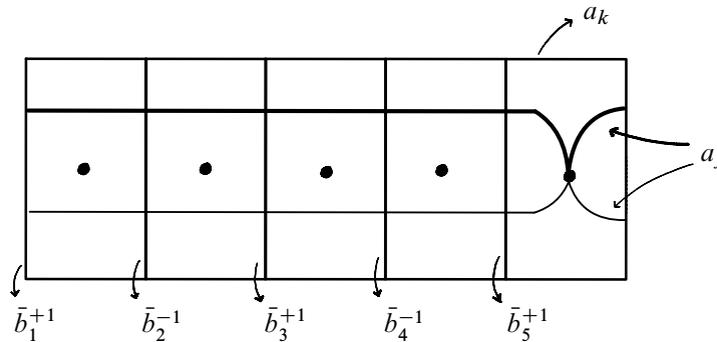
Figure 7: The planar torus $T_k$. The thick line and the thin line represent the two possible arcs for $a_j \cap T_i$.

Take any curve $b_j \in B_k \setminus \{b_k\}$, from Lemma 3.16 we have that $|\mathrm{Orb}_f(b_j)| = 4$. Notice that $a_j$ is a type 3 curve, since Lemmas 3.11, 3.13, 3.14 and 3.15 rule out any other type in Table 1. Now, considering Lemma 3.17, we deduce $i(a_k, b) = i(a_j, b)$ for every $b \in B$. As a consequence, any two curves $b_j, b_l \in B_k = B_i$ satisfy $i(a_j, b) = i(a_l, b)$ for every $b \in B$.

We continue by considering $T_k$ an open regular neighbourhood of the curves $a_k \cup \bigcup_{b \in B_k} b$ (see Figure 7). For any $b_j \in B_k$ consider the curve $a_j$ such that $f \cdot a_j = b_j$. By the previous paragraph, we know $a_j$ satisfies the equations

$$\begin{cases} i(a_j, a_k) = 0, \\ i(a_j, b) = i(a_k, b) \quad \text{for all } b \in B, \\ i(a_j, \tau_B \cdot a_k) = 1, \\ i(a_j, \tau_B^{-1} \cdot a_k) = 1. \end{cases}$$

Since $a_j$ satisfies the same conditions as $a_{i_1}$ in the proof of Lemma 3.16, we may use the same argument to find a representative of $a_j$ disjoint from $a_k$ such that $a_j \cap T_k$ is an arc with endpoints in a single puncture, and $T_k \setminus (a_k \cup a_j)$ is the union of an annulus and a $|B_k| - 1$ punctured annulus. Simpler, $a_j \cap T_k$ is as in Figure 7.

We emphasize that $a_j \cap T_k$ is as in Figure 7 for any curve $a_j \in A$ with $f \cdot a_j = b_j \in B_k$. By taking appropriate representatives of each $a_j$, we may assume that $\bar{b}_1$ and $\bar{b}_{|B_i|}$ induce the same cyclic order on the $a_j$'s (see Figure 8).

Now, consider $T$ the regular neighbourhood of $\bigcup_{b_j \in B_k} b_j \cup \bigcup_{b_j \in B_k} a_j$. The neighbourhood $T$ is homeomorphic to a torus with punctures. We represent $T$ as a planar torus in Figure 9, where we denote $\bar{a}_j$ the curves that satisfy $f \cdot \bar{a}_j = \bar{b}_j$. To check that $T$ is a torus with punctures, it is enough to consider a regular neighbourhood of the curves as depicted in Figure 8.

Recall that the intersection of $a_k$ with curves $\bar{b}_j$ induces the cyclic order $\bar{b}_1, \ldots, \bar{b}_{|B_i|}$. In the same fashion, the intersection of $b_k$ with curves $\bar{a}_j$ induces the cyclic order $\bar{a}_1, \ldots, \bar{a}_{|B_i|}$ (see Figure 9). This simply follows from the fact that $f \cdot \bar{a}_j = \bar{b}_j$ and $f \cdot a_k = b_k$.
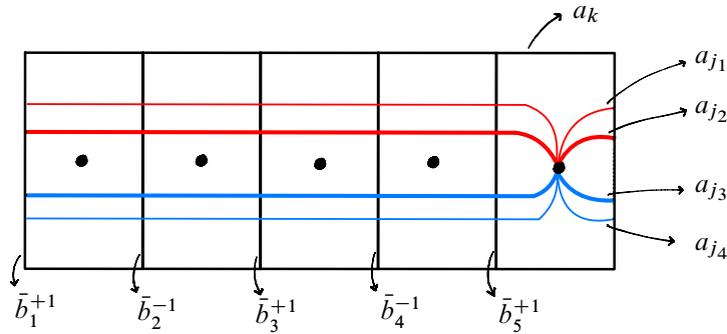
Figure 8: The planar torus $T_k$. The arcs $a_{j_l} \cap T_k$ for $a_{j_l} \in A$ are represented with different colours or thickness.

Finally, one may directly check that $i(\tau_B \tau_A \cdot \bar{b}_{|B_i|}, \bar{a}_1) \neq 0$ by taking a representative of $\tau_B \tau_A \cdot \bar{b}_{|B_i|}$ in $T$ and using the bigon criterion. This clearly yields a contradiction, as $\tau_B \tau_A \cdot \bar{b}_{|B_i|} \in A$ and any two curves in $A$ are disjoint. □

Collecting the previous lemmas we conclude $|\mathrm{Orb}_f(a_i)| = 2$ for any curve $a_i \in A$. This ends up the proof of Proposition 3.10.

**Proof of Proposition 3.10**   Consider a curve $c \in A \cup B$. To prove $|\mathrm{Orb}_f(c)| = 2$, it is enough to prove it for $c \in A$. By Proposition 3.4 the curve $c$ is of type 1, 2, 3, 4 or 5, so it satisfies the conditions of one of Lemmas 3.11, 3.13, 3.14, 3.15 or 3.18. In any case, the conclusion is $|\mathrm{Orb}_f(c)| = 2$. □

## 3.5   Reduced multitwists are trivial

We are close to proving Proposition 3.8. Before doing so, we devote the next lemmas to understand three concrete cases.
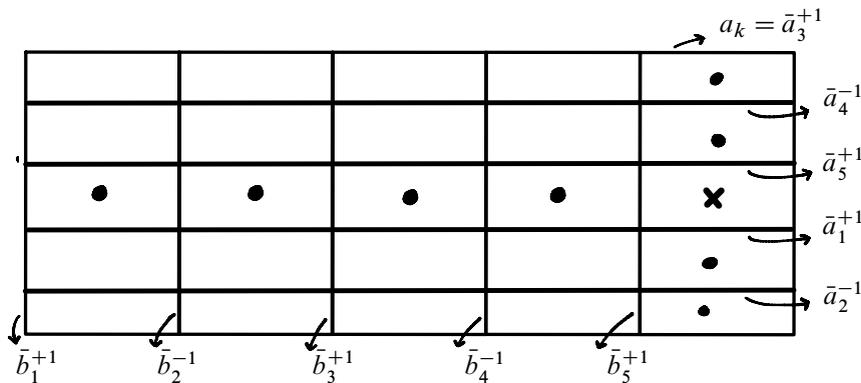


Figure 9: The planar torus $T$. The puncture marked with a cross might be a disk.
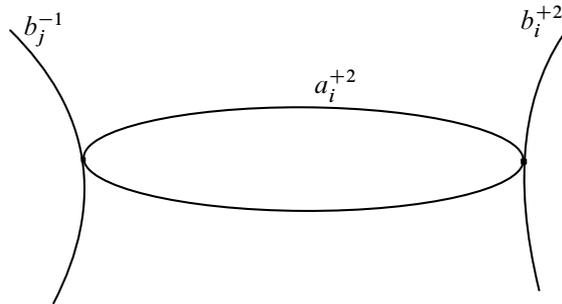
Figure 10: Regular neighbourhood of $i(a_i, b_i) = 1$ with $|n_i| = 2$ such that $a_i$ intersects only two curves in $B$.

**Lemma 3.19** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 2 curve, then $a_i$ intersects at least three distinct curves in $B$.*

**Proof** Seeking a contradiction, assume $a_i$ intersects at most two curves in $B$. Since the multitwists are reduced, $a_i$ intersects exactly two curves in $B$.

Take $b_i, b_j \in B$ to be the two curves that intersect $a_i$. We may assume without loss of generality that $n_i = 2$ and $n_j = -1$ (see Figure 10). We have that

$$\pm 1 = \hat{\imath}(\tau_A \tau_B \cdot [a_i], [a_j]) = \hat{\imath}(\tau_B \cdot [a_i], [a_j]) = \hat{\imath}([a_i] + 2\hat{\imath}([a_i], [b_i])[b_i] - \hat{\imath}([a_i], [b_j])[b_j], [a_j])$$

$$= 2\hat{\imath}([a_i], [b_i])\hat{\imath}([b_i], [a_j]) - \hat{\imath}([a_i], [b_j])\hat{\imath}([b_j], [a_j])$$

$$= \pm 2 \pm \hat{\imath}([b_j], [a_j]).$$

Notice that for the previous equation to be satisfied, it is required that $i(b_j, a_j) = 1$.

To finish, consider $T_i$ a regular neighbourhood of $a_i \cup b_i \cup b_j$. We represent $T_i$ as a planar torus in Figure 11. Now, we know $a_j$ satisfies the equations

$$\begin{cases} i(a_j, b_i) = 1, \\ i(a_j, a_i) = 0, \\ i(a_j, \tau_B \cdot a_i) = 1, \\ i(a_j, \tau_B^{-1} \cdot a_i) = 1. \end{cases}$$

These conditions imply that $a_j$ is a curve contained in $T_i$. But the only curve in $T_i$ satisfying above equations is $a_i$, so $a_j = a_i$. However, this is impossible since we assumed $b_j \neq b_i$. □

**Lemma 3.20** *Consider two reduced multitwists $\tau_A, \tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 1 curve, then $a_i$ intersects at least three curves in $B$.*

**Proof** Seeking a contradiction, assume $a_i$ intersects at most two curves in $B$. Since the multitwists are reduced, $a_i$ intersects exactly two curves in $B$.
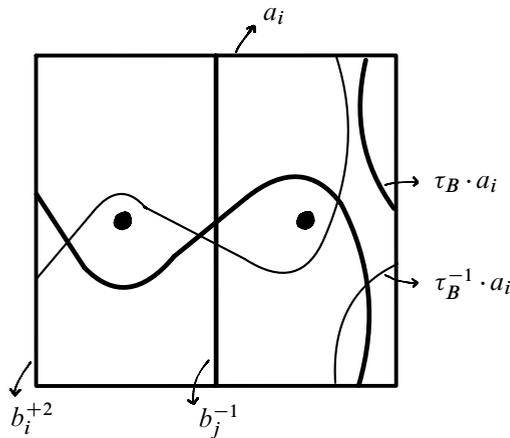
Figure 11: Planar torus $T_i$ with $i(a_i, b_i) = 1$ and $|n_i| = 2$, such that $a_i$ intersects only two curves in $B$.

Take $b_j, b_k \in B$ to be the two curves that intersect $a_i$. Without loss of generality, we may assume the Dehn twists in $\tau_B$ have exponents $n_i = n_k = 1$ and $n_j = -1$ (see Figure 12).

Let $T_i$ be a regular neighbourhood of $a_i \cup b_j \cup b_k$. The subsurface $T_i$ is homeomorphic to a torus with two punctures. We represent $T_i$ as a planar torus in Figure 13. Note that $a_j$ satisfies the equations

$$\begin{cases} i(a_j, \tau_B \cdot a_i) = 1, \\ i(a_j, \tau_B^{-1} \cdot a_i) = 1. \end{cases}$$

Any curve $a_j$ satisfying the conditions above is either $a_i$ or is a curve disjoint from $T_i$ (see Figure 13). If $i(a_j, b_j) \neq 0$, then $a_j$ intersects $T_i$ and so $a_j = a_i$. However, this is not possible since $b_j \neq b_i$. Therefore, we can assume $i(a_j, b_j) = 0$.

Notice that

$$\pm 1 = \hat{\imath}(\tau_B \cdot [a_i], [a_j]) = \hat{\imath}([a_i] + \hat{\imath}([a_i], [b_k])[b_k] - \hat{\imath}([a_i], [b_j])[b_j], [a_j])$$
$$= \hat{\imath}([a_i], [b_k])\hat{\imath}([b_k], [a_j]) - \hat{\imath}([a_i], [b_j])\hat{\imath}([b_j], [a_j])$$
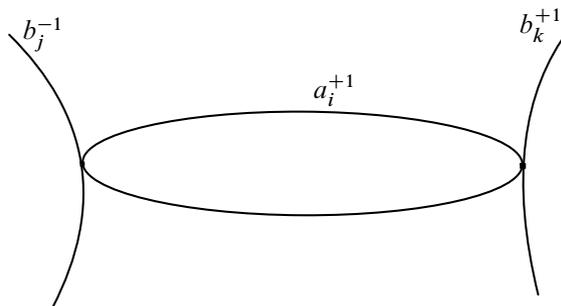$$= \pm \hat{\imath}([b_k], [a_j]) \pm \hat{\imath}([b_j], [a_j]).$$



Figure 12: Regular neighbourhood of $a_i$ with $i(a_i, b_i) = 0$ such that $a_i$ intersects exactly two curves in $B$.
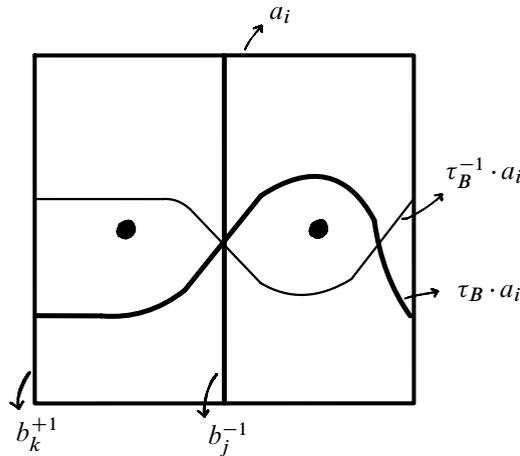
Figure 13: Regular neighbourhood of $i(a_i, b_i) = 0$, where $a_i$ intersects two curves in $B$.

If $i(a_j, b_j) = 0$ then the previous equation implies $i(a_j, b_k) = 1$. Again, it follows that $a_j$ is not disjoint from $T_i$ and so we have the contradiction $a_i = a_j$. $\qquad\square$

**Lemma 3.21** *Consider two reduced multitwists $\tau_A$, $\tau_B$ and let $f = \tau_A \tau_B \tau_A$. If $a_i \in A$ is a type 5 curve and $a_i$ intersects at most three curves in $B$, then, there exists $a_j \in A$ such that*

- *either $a_j$ is a type 1 curve,*
- *or $i(a_j, b_j) = 1$ (type 2, 3 or 4),*
- *or $i(a_j, b_j) = 2$ and $a_j$ intersects at least four curves in $B$.*

**Proof** Since $\tau_A$, $\tau_B$ are reduced, $a_i$ must intersect at least two curves. Using that $i(a_i, b_i) = 2$ and $X(a_i, \tau_B) = 0$ (see Table 1 for type 5 curves), we deduce $a_i$ intersects exactly three curves.

Let $b_i, b_j, b_k \in B$ be the curves intersecting $a_i$. We may assume that $n_j = n_k = -1$ and $n_i = 1$. Now, we have

$$\pm 1 = \hat{\imath}([a_j], [b_i]) = \hat{\imath}([a_j], \tau_B \cdot [a_i])$$
$$= \hat{\imath}([a_j], [a_i] + \hat{\imath}([a_i], [b_i])[b_i] - \hat{\imath}([a_i], [b_j])[b_j] - \hat{\imath}([a_i], [b_k])[b_k])$$
$$= \hat{\imath}([a_i], [b_i])\hat{\imath}([a_j], [b_i]) - \hat{\imath}([a_i], [b_j])\hat{\imath}([a_j], [b_j]) - \hat{\imath}([a_i], [b_k])\hat{\imath}([a_j], [b_k])$$
$$= \pm \hat{\imath}([a_i], [b_i]) \pm \hat{\imath}([a_j], [b_j]) \pm \hat{\imath}([a_j], [b_k]).$$

If $i(a_j, b_j) \in \{0, 1\}$, then we already have the statement. Suppose that $i(a_j, b_j) = 2$.

Using that $i(a_i, b_i) = 2$, the previous equation implies that $i(a_j, b_k) = 1$. Since $i(a_j, b_j) = 2$, $X(a_j, \tau_B) = 0$ and $n_k = n_j = -1$, it follows that $a_j$ must intersect at least four curves in $B$ as otherwise $X(a_j, \tau_B) \neq 0$. $\qquad\square$

Lastly, we are ready to prove Proposition 3.8.

**Proof of Proposition 3.8** Assume $\tau_A, \tau_B$ are nontrivial and let $f = \tau_A \tau_B \tau_A$. For each curve $a_i \in A$ consider the subset $B_i \subset B$ of curves intersecting $a_i$. Since $\tau_A$ is nontrivial, there exists an $a_i \in A$.

If $i(a_i, b_i) \leq 1$ and $|B_i| \geq 3$, consider the graph $\mathcal{G}$ isomorphic to the union of curves $a_i \cup \bigcup_{b \in B_i} b$. Recall that Proposition 3.10 implies that $f^2 \cdot c = c$ for every $c \in A \cup B$, so $f^2$ defines an automorphism of the graph $\mathcal{G}$ by means of the Alexander method. Even more, $f^2$ fixes every vertex in $\mathcal{G}$ and so it follows that $f^2$ preserves the orientation of $a_i$. In other words, $f^2 \cdot [a_i] = [a_i]$. Notice there exists $b \in B_i$ with $i(a_i, b) = 1$ and

$$\hat{\imath}([a_i], [b]) = \hat{\imath}(f^2 \cdot [a_i], f^2 \cdot [b]) = \hat{\imath}([a_i], f^2 \cdot [b]),$$

but this clearly contradicts Lemma 3.12. As a result, there cannot exist $a_i \in A$ with $i(a_i, b_i) \leq 1$ and $|B_i| \geq 3$.

Suppose $a_i \in A$ is a curve with $i(a_i, b_i) = 2$ and $|B_i| \geq 4$, then the previous argument yields the same contradiction. For $i(a_i, b_i) = 2$, we emphasize $|B_i| \geq 4$ is used to ensure $f^2$ fixes every vertex in $\mathcal{G}$.

To finish, we list the rest of the cases according to Proposition 3.4:

- If $a_i$ is a type 1 curve, then Lemma 3.20 ensures $|B_i| \geq 3$ and we already proved this leads to contradiction.

- If $a_i$ is a type 2 curve, then Lemma 3.19 implies $|B_i| \geq 3$. Again, we know this leads to contradiction.

- If $a_i$ is a type 4 curve, then Table 1 guarantees the existence of another curve $a_j \in A$ of type 2. However, this puts us in the previous case, that has already been discarded.

- If $a_i$ is a type 3 curve, then $|B_i|$ is odd. Since the multitwists are reduced, it follows that $|B_i| \geq 3$ and we already checked this case drives to contradiction.

- At last, if $a_i$ is a type 5 curve and $|B_i| \leq 3$, then Lemma 3.21 implies the existence of an $a_j$ of type 1, 2, 3, 4, or 5, and if $a_j$ is of type 5 then $|B_j| \geq 4$. However, we proved in the previous cases that such $a_j$ cannot exist. Thus, $a_i$ of type 5 and $|B_i| \leq 3$ also leads to contradiction.

We have seen that the existence of $a_i \in A$ leads to a contradiction independently of its type. Therefore, $A$ is empty and so is $B$. In other words, $\tau_A = \tau_B = 1 \in \mathrm{Map}(S)$. $\qquad\square$

With this in hand, we may quickly prove Theorem 1.4.

**Proof of Theorem 1.4** Consider two braided multitwists $\tau_A, \tau_B$. Following Section 3 we can write

$$\tau_A = \delta_{a_1}^{n_1} \cdots \delta_{a_k}^{n_k}, \quad \tau_B = \delta_{b_1}^{n_1} \cdots \delta_{b_k}^{n_k},$$

where $\tau_A \tau_B \tau_A \cdot a_i = b_i$. Additionally, we write $A = \{a_1, \dots, a_k\}$ and $B = \{b_1, \dots, b_k\}$. Using Lemma 3.2, we may assume that $A$ and $B$ share no curves and by Lemma 3.3 we may assume every curve $a \in A$ intersects some curve $b \in B$.

Let $a_i \in A$ be a curve intersecting a single curve $b$ in $B$. In this case, Lemma 3.5 guarantees $b = b_i = f \cdot a_i$, $i(a_i, b_i) = 1$ and $|n_i| = 1$. Moreover, we can decompose $\tau_A = \tau_{A'} \circ \delta_{a_i}^{n_i}$ and $\tau_B = \tau_{B'} \circ \delta_{b_i}^{n_i}$, so that $\tau_{A'}, \tau_{B'}$ are still braided and share no curves.

Repeating the above process until every curve $a_i \in A'$ intersects at least two curves in $B'$, leaves us with two reduced multitwists $\tau_{A'}, \tau_{B'}$. But by Proposition 3.8, we have that $\tau_{A'} = \tau_{B'} = 1$. Thus, by Lemma 3.5 we conclude that the curves $a_1, \ldots, a_k$, $b_1, \ldots, b_k$ are pairwise disjoint except for $i(a_i, b_i) = 1$ and that $n_i \in \{-1, 1\}$, just as we claimed in Theorem 1.4. $\qquad\square$

# References

[1] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Mathematical Series 49, Princeton Univ. Press (2012) MR

[2] **H Hamidi-Tehrani**, *Groups generated by positive multi-twists and the fake lantern problem*, Algebr. Geom. Topol. 2 (2002) 1155–1178 MR

[3] **N V Ivanov**, *Subgroups of Teichmüller modular groups*, Translations of Mathematical Monographs 115, Amer. Math. Soc., Providence, RI (1992) MR

[4] **C J Leininger**, *On groups generated by two positive multi-twists: Teichmüller curves and Lehmer's number*, Geom. Topol. 8 (2004) 1301–1359 MR

[5] **J D McCarthy**, *Automorphisms of surface mapping class groups: A recent theorem of N Ivanov*, Invent. Math. 84 (1986) 49–71 MR

[6] **B Szepietowski**, *Embedding the braid group in mapping class groups*, Publ. Mat. 54 (2010) 359–368 MR

*University of Notre Dame du Lac*
*South Bend, IN, United States*

rdepoola@nd.edu

# Positive intermediate Ricci curvature on connected sums

Philipp Reiser

David J Wraith

We consider the problem of performing connected sums in the context of positive $k^{\text{th}}$-intermediate Ricci curvature. We show that such connected sums are possible if the manifolds involved possess "$k$-core metrics" for some $k$. Here, a $k$-core metric is a generalisation of the notion of core metric introduced by Burdick for positive Ricci curvature. Further, we show that connected sums of linear sphere bundles over bases admitting such metrics admit positive $k^{\text{th}}$-intermediate Ricci curvature for $k$ in a particular range. This follows from a plumbing result we establish, which generalises other recent plumbing results in the literature and is possibly of independent interest. As an example of a manifold admitting a $k$-core metric, we prove that $\mathbb{H}P^n$ admits a $(4n-3)$-core metric and that $\mathbb{O}P^2$ admits a 9-core metric, and we show that in both cases these are optimal.

53C20

## 1 Introduction

Given two (or more) closed manifolds with the same dimension, the operation of connected sum is perhaps the most basic topological operation one can perform that again yields a closed manifold with the same dimension. If one wants to understand the interplay between the topology of manifolds and some aspect of geometry, investigating how that geometry behaves under connected sums is thus a fundamental task.

In Riemannian geometry, it is natural to ask what kind of curvature bounds can be preserved under connected sums. In general, this turns out to be a very hard question.

Given two manifolds, both with dimension $d \geq 3$ and Riemannian metrics of positive scalar curvature, Gromov and Lawson [10] famously showed that the connected sum also admits a metric of positive scalar curvature. Moreover, this metric can be chosen so as to agree with the original metrics outside a neighbourhood of the connected sum.

At the other extreme, it is so difficult to find metrics of positive sectional curvature that the question of behaviour under connected sums is unreasonable. In general, it follows from Gromov's Betti number bound [9] that arbitrary connected sums cannot preserve positive sectional curvature. Cheeger [7] showed

that the connected sum of a pair of compact rank-one symmetric spaces admits a metric with nonnegative sectional curvature. However it is unknown, for example, whether the connected sum of three such spaces admits nonnegative sectional curvature. Indeed the Bott conjecture on the rational ellipticity of simply connected closed manifolds admitting nonnegative sectional curvature implies that such metrics should not exist.

In between the scalar and sectional curvatures lies the Ricci curvature. The question of whether the connected sum of two manifolds with positive Ricci curvature also supports a metric of positive Ricci curvature turns out to be intriguing. By the theorem of Bonnet and Myers the connected sum of two closed, non-simply-connected Ricci-positive manifolds cannot admit a metric of positive Ricci curvature. However, if at least one of the manifolds is simply connected, the question is open. This problem was systematically studied by Burdick [4; 5; 6], who, based on earlier work by Perelman [14], introduced the notion of *core metrics* and showed that the connected sum of manifolds with core metrics admits a metric of positive Ricci curvature.

In this article, we consider a natural family of positive curvature conditions which interpolate between positive Ricci curvature and positive sectional curvature:

**Definition 1.1**  A Riemannian manifold $(M^n, g)$ has *positive $k^{th}$-intermediate Ricci curvature* for some $k \in \{1, \ldots, n-1\}$, denoted $\mathrm{Ric}_k > 0$, if for every unit tangent vector $v \in \mathrm{TM}$ and any orthonormal $k$-frame $(e^i)$ in $v^\perp$ the sum $\sum_{i=1}^{k} K(v, e^i)$ is positive, where $K$ denotes the sectional curvature.

For $k = 1$ and $k = n - 1$, we recover the conditions of positive sectional curvature and positive Ricci curvature, respectively. Although intermediate curvatures have appeared in the literature for several decades, in recent times there has been a dramatic increase in interest in these curvatures. For an up-to-date list of papers which feature intermediate curvatures, see [13].

The main goal of this paper is to establish conditions under which connected sums admit metrics with $\mathrm{Ric}_k > 0$.

Our first main result, Theorem A, provides a generalisation of Burdick's results to intermediate Ricci curvatures. This theorem requires a generalisation of Burdick's notion of core metric, and we illustrate this new notion with reference to projective spaces (Theorem B).

The plumbing of disc bundles has proved to be a very important topological construction in the realm of positive Ricci curvature. See, for example, [8]. We prove a plumbing result for $\mathrm{Ric}_k > 0$ (Theorem D), and then illustrate this by providing examples of connected sums between linear sphere bundles which admit metrics with positive intermediate Ricci curvatures (Corollary E).

In order to give a precise statement of the results, we must begin by defining our generalisation of Burdick's core metrics:

**Definition 1.2** Let $M$ be an $n$-dimensional manifold and let $k \in \{1, \ldots, n-1\}$. A Riemannian metric $g$ on $M$ is called a *$k$-core metric* if $g$ has $\mathrm{Ric}_k > 0$ and if there exists an embedding $\varphi \colon D^n \hookrightarrow M$ such that

(i)  the induced metric $g|_{\varphi(S^{n-1})}$ is the round metric of radius one, and

(ii)  $\mathrm{II}_{\varphi(S^{n-1})}$ is positive semidefinite with respect to the outward normal of $S^{n-1} \subseteq D^n$.

For $k = n-1$ we recover the original definition given in [5] except for the fact that the second fundamental form is required to be strictly positive in [5]. However, a core metric in the sense of Definition 1.2 can always be deformed into a core metric in the sense of [5]; see, eg, [4, Proposition 1.2.11].

In [5, Theorem B], it is shown that connected sums of manifolds with $(n-1)$-core metrics support positive Ricci curvature. We can now generalise this as follows.

**Theorem A** *Let $M_1, \ldots, M_\ell$ be $n$-dimensional manifolds that admit $k$-core metrics, where $k \geq 2$. Then $M_1 \# \cdots \# M_\ell$ admits a metric with $\mathrm{Ric}_k > 0$.*

The main ingredients in the proof of Theorem A are the gluing theorem for positive intermediate Ricci curvature established in [17], together with the construction of a metric with $\mathrm{Ric}_2 > 0$ on $S^n \setminus \bigsqcup_\ell (D^n)^\circ$, (which is called the *docking station* in [5]), whose second fundamental form on the boundary can be made arbitrarily small; see Theorem 3.1.

**Remark 1.3** Since the metric on the docking station is invariant under the action of $O(n-1)\, O(2) \subseteq O(n+1)$, we can take quotients by finite subgroups of $O(n-1)\, O(2)$ that act freely as in [5, Corollary 4.7]. In this way we obtain in the situation of Theorem A that $\mathbb{R}P^n \# M_1 \# \cdots \# M_\ell$ and $L \# M_1 \# \cdots \# M_\ell$ admits a metric of $\mathrm{Ric}_k > 0$, where $L$ is any $n$-dimensional lens space (and $n$ is assumed to be odd in this case). By [4, Lemma 1.2.9], lens spaces and real projective spaces are the only additional summands we can obtain in this way.

Concerning the existence of $k$-core metrics, by a result of Wu [22], the boundary condition (ii) in Definition 1.2 imposes the following topological obstruction.

**Proposition 1.4** *Let $M$ be a closed $n$-dimensional manifold that admits a $k$-core metric. Then $M$ is $(n-k)$-connected. In particular, if $k \leq \lfloor \frac{n+1}{2} \rfloor$, then $M$ is a homotopy sphere.*

We immediately obtain the following restrictions in low dimensions: Every closed 3-manifold with a $k$-core metric is diffeomorphic to the standard sphere and the same holds in dimension 5 when $k \leq 3$. In dimension 4 every closed manifold with a $k$-core metric is homeomorphic to the standard sphere when $k \leq 2$.

On the other hand, it is easy to see that the round metric on $S^n$ is a 1-core metric. Further, by [5], complex and quaternionic projective spaces and the Cayley plane of dimension $n$ admit $(n-1)$-core

metrics, where $n$ denotes the real dimension of the corresponding manifold. By Proposition 1.4, this is optimal for complex projective spaces. For quaternionic projective spaces and the Cayley plane we obtain the following improvement, which again is optimal by Proposition 1.4.

**Theorem B** $\mathbb{H}P^n$ *admits a* $(4n-3)$-*core metric and* $\mathbb{O}P^2$ *admits a* $9$-*core metric.*

In [18] it was shown that a Betti number bound as in the case of nonnegative sectional curvature [9] cannot hold for $\mathrm{Ric}_k > 0$ for all $k \geq \lfloor \frac{n}{2} \rfloor + 2$, where $n$ denotes the dimension. By considering connected sums of copies of $\mathbb{H}P^2$ and $\mathbb{O}P^2$ using Theorems A and B, we can slightly improve this result as follows.

**Corollary C** *For any* $\ell \in \mathbb{N}$ *the manifold* $\#_\ell \mathbb{H}P^2$ *admits a metric of* $\mathrm{Ric}_5 > 0$ *and the manifold* $\#_\ell \mathbb{O}P^2$ *admits a metric of* $\mathrm{Ric}_9 > 0$. *In particular, Gromov's Betti number bound does not hold in dimension 8 for* $\mathrm{Ric}_5 > 0$ *and in dimension 16 for* $\mathrm{Ric}_9 > 0$.

By using manifolds with $k$-core metrics as base manifolds of fibre bundles, we can also consider plumbings as in the following theorem, which generalises results for positive Ricci curvature in [15; 20], and for positive intermediate Ricci curvature in [19].

**Theorem D** *Let* $W$ *be the manifold obtained by plumbing linear disc bundles with compact base manifolds according to a simply connected graph. Suppose the following*:

(i) *For a fixed bundle in this graph the base admits a metric with* $\mathrm{Ric}_{k_1} > 0$ *for some* $k_1$. *Denote the base dimension by* $q + 1$ *and the fibre dimension by* $p + 1$.

(ii) *Every other bundle in this graph with base dimension* $q + 1$ *admits a* $k_1$-*core metric.*

(iii) *Every bundle with base dimension* $p + 1$ *admits a* $k_2$-*core metric for some* $k_2$.

*Then, if* $p, q \geq 2$, *the manifold* $\partial W$ *admits a metric of* $\mathrm{Ric}_k > 0$ *for all* $k \geq \max\{p+2, p+k_1, q+2, q+k_2\}$.

We can use plumbings as in Theorem D to construct connected sums of sphere bundles as follows.

**Corollary E** *Let* $E_i \to B_i^q$, $1 \leq i \leq \ell$, *be linear* $S^p$-*bundles with compact base manifolds such that* $B_1$ *admits a metric of* $\mathrm{Ric}_k > 0$ *and each* $B_i$, $2 \leq i \leq \ell$, *admits a* $k$-*core metric. Then the connected sum* $E_1 \# \cdots \# E_\ell$ *admits a metric of* $\mathrm{Ric}_k > 0$ *for all* $k \geq \max\{p+2, p+k, q+1\}$.

This paper is laid out as follows. In Section 2 we prove a generalisation of the main technical result in [18]. The aim is to establish criteria which identify when metrics (of the type under consideration in this paper) have $\mathrm{Ric}_k > 0$. In Section 3 we prove that the neck construction from [14] actually gives a metric with $\mathrm{Ric}_2 > 0$, and we use this to prove Theorem A. The remaining results (Theorems B, D, and Corollary E) are then established in Section 4.

## 2 Preliminaries

Let $(M^n, g)$ be a Riemannian manifold. To characterise the condition $\mathrm{Ric}_k > 0$ we consider the curvature operator $\mathcal{R} \colon \Lambda^2 TM \to \Lambda^2 TM$ defined by

$$g(\mathcal{R}(v_1 \wedge v_2), v_3 \wedge v_4) = g(R(v_1, v_2)v_4, v_3),$$

where $\Lambda^2 TM$ is equipped with the Riemannian metric which is the natural extension of $g$ to $\Lambda^2 TM$, ie

$$g(v_1 \wedge v_2, v_3 \wedge v_4) = g(v_1, v_3)g(v_2, v_4) - g(v_1, v_4)g(v_2, v_3).$$

We recall some definitions of [18]: For an inner product space $V$ the set $\{v_0 \wedge v_1, \dots, v_0 \wedge v_k\} \subseteq \Lambda^2 V$, where $(v_0, \dots, v_k)$ is an orthonormal $(k{+}1)$-frame in $V$, is called a $k$-*chain* with *base* $v_0$. For a linear map $A \colon \Lambda^2 V \to \Lambda^2 V$ and a $k$-chain $\{v_0 \wedge v_1, \dots, v_0 \wedge v_k\}$ the sum

$$\sum_{i=1}^{k} \langle A(v_0 \wedge v_i), v_0 \wedge v_i \rangle$$

is the *value* of $A$ on this $k$-chain. Note that $(M, g)$ has $\mathrm{Ric}_k > 0$ if and only if at every point in $M$ the value of $\mathcal{R}$ on every $k$-chain is positive.

In [18] we considered the condition $\mathrm{Ric}_k > 0$ for doubly warped product metrics. In this case each tangent space splits orthogonally into a direct sum $V_1 \oplus V_2 \oplus V_3$ such that each subspace $V_i \wedge V_j$ is an eigenspace for $\mathcal{R}$. Below we will be interested in the following more general situation.

**Proposition 2.1** *Let $(V, \langle \cdot, \cdot \rangle)$ be a finite-dimensional inner product space of dimension $n$ and let $A \colon \Lambda^2 V \to \Lambda^2 V$ be a linear self-adjoint map. Suppose that $V$ splits orthogonally as*

$$V = V_1 \oplus V_2 \oplus V_3$$

*so that $V_1$ and $V_2$ are one-dimensional and $A$ is given by*

$$\begin{aligned}
A(v_1 \wedge v_2) &= \lambda_{12} v_1 \wedge v_2, \\
A(v_1 \wedge w_1) &= \lambda_{13} v_1 \wedge w_1 + \tilde{\lambda} v_2 \wedge w_1, \\
A(v_2 \wedge w_1) &= \lambda_{23} v_2 \wedge w_1 + \tilde{\lambda} v_1 \wedge w_1, \\
A(w_1 \wedge w_2) &= \lambda_3 w_1 \wedge w_2,
\end{aligned}$$

*for some $\lambda_{12}, \lambda_{13}, \lambda_{23}, \tilde{\lambda}, \lambda_3 \in \mathbb{R}$, where $v_1$ and $v_2$ are unit vectors in $V_1$ and $V_2$, respectively, and $w_1, w_2 \in V_3$. Then for $2 \leq k \leq n - 3$ the value of $A$ on every $k$-chain is positive if and only if*

(i) $\lambda_{12} + \frac{1}{2}(k-1)(\lambda_{13} + \lambda_{23}) > 0$,

(ii) $(\lambda_{12} + (k-1)\lambda_{13})(\lambda_{12} + (k-1)\lambda_{23}) > (k-1)^2 \tilde{\lambda}^2$,

(iii) $\lambda_{13}\lambda_{23} > \tilde{\lambda}^2$,

(iv) $\lambda_{13}, \lambda_{23}, \lambda_3 > 0$.

*For $k = n - 2, n - 1$ these inequalities are still sufficient, but not necessary.*

**Proof** First note that if $\tilde{\lambda} = 0$, then the spaces $V_i \wedge V_j$ are eigenspaces for $A$, so we are in the situation of [18, Proposition 2.3]. Observe that (i)–(iv) in this case now become

$$\lambda_{12} + (k-1)\lambda_{13} > 0,$$
$$\lambda_{12} + (k-1)\lambda_{23} > 0,$$
$$\lambda_{13}, \lambda_{23}, \lambda_3 > 0,$$

and these are the inequalities appearing in [18, Proposition 2.3] for $k \leq n-3$, and for $k = n-2, n-1$ these inequalities are easily seen to be implied by those appearing in [18, Proposition 2.3].

From now on we can therefore assume that $\tilde{\lambda} \neq 0$. We modify the vectors $v_1$ and $v_2$ as follows. First, let

$$\mu = \frac{\lambda_{13} - \lambda_{23} + \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2}}{2\tilde{\lambda}}.$$

and define

$$v_1' = \mu v_1 + v_2, \quad v_2' = -v_1 + \mu v_2.$$

Let $V_1'$ and $V_2'$ be the subspaces generated by $v_1'$ and $v_2'$, respectively, and set $V_3' = V_3$. Then $V_1'$ and $V_2'$ are orthogonal and $V_1' \oplus V_2' = V_1 \oplus V_2$. A calculation shows that the spaces $V_i' \wedge V_j'$ are eigenspaces for $A$ with eigenvectors $\lambda_{ij}'$ given by

$$\lambda_{12}' = \lambda_{12},$$
$$\lambda_{13}' = \tfrac{1}{2}\left(\lambda_{13} + \lambda_{23} + \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2}\right),$$
$$\lambda_{23}' = \tfrac{1}{2}\left(\lambda_{13} + \lambda_{23} - \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2}\right),$$
$$\lambda_{33}' = \lambda_3.$$

By [18, Proposition 2.3], the value of $A$ on every $k$-chain is positive if and only if the sum of any $k$ nondiagonal elements in each row of the following $(n \times n)$ matrix is positive:

$$\begin{pmatrix} 0 & \lambda_{12}' & \lambda_{13}' & \cdots & \cdots & \lambda_{13}' \\ \lambda_{12}' & 0 & \lambda_{23}' & \cdots & \cdots & \lambda_{23}' \\ \lambda_{13}' & \lambda_{23}' & 0 & \lambda_{33}' & \cdots & \lambda_{33}' \\ \vdots & \vdots & \lambda_{33}' & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \lambda_{33}' \\ \lambda_{13}' & \lambda_{23}' & \lambda_{33}' & \cdots & \lambda_{33}' & 0 \end{pmatrix}.$$

For $2 \leq k \leq n-3$ this is equivalent to the system of inequalities

$$\lambda_{12}' + (k-1)\lambda_{13}' > 0, \quad \lambda_{12}' + (k-1)\lambda_{23}' > 0,$$
$$\lambda_{13}' + \lambda_{23}' + (k-2)\lambda_{33}' > 0, \quad \lambda_{13}' + (k-1)\lambda_{33}' > 0,$$
$$\lambda_{23}' + (k-1)\lambda_{33}' > 0, \quad \lambda_{13}', \lambda_{23}', \lambda_{33}' > 0.$$

For $k = n-2, n-1$ these inequalities are still sufficient, but not all are necessary.

The inequalities $\lambda'_{13} + \lambda'_{23} + (k-2)\lambda'_{33} > 0$, $\lambda'_{13} + (k-1)\lambda'_{33} > 0$ and $\lambda'_{23} + (k-1)\lambda'_{33} > 0$ are superfluous. Hence, we arrive at the system of inequalities

$$\lambda_{12} + \tfrac{1}{2}(k-1)\left(\lambda_{13} + \lambda_{23} + \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2}\,\right) > 0, \tag{1}$$

$$\lambda_{12} + \tfrac{1}{2}(k-1)\left(\lambda_{13} + \lambda_{23} - \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2}\,\right) > 0, \tag{2}$$

$$\lambda_{13} + \lambda_{23} + \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2} > 0, \tag{3}$$

$$\lambda_{13} + \lambda_{23} - \sqrt{(\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2} > 0, \tag{4}$$

$$\lambda_3 > 0. \tag{5}$$

First note that (2) implies (1), and (4) implies (3). Hence, we are left with (2), (4) and (5). Next, observe that (4) implies $\lambda_{13} + \lambda_{23} > 0$ and is therefore equivalent to

$$\lambda_{13}\lambda_{23} > \tilde{\lambda}^2.$$

In particular, $\lambda_{13}\lambda_{23} > 0$. This observation, together with $\lambda_{13} + \lambda_{23} > 0$, is equivalent to

$$\lambda_{13}, \lambda_{23} > 0.$$

Hence, (4) is equivalent to

$$\lambda_{13}, \lambda_{23} > 0, \quad \lambda_{13}\lambda_{23} > \tilde{\lambda}^2.$$

Finally, (2) is equivalent to (i) and (ii), since it is equivalent to

$$\lambda_{12} + \tfrac{1}{2}(k-1)(\lambda_{13} + \lambda_{23}) > 0$$

and

$$\left(\lambda_{12} + \tfrac{1}{2}(k-1)(\lambda_{13} + \lambda_{23})\right)^2 > \tfrac{1}{4}(k-1)^2\left((\lambda_{13} - \lambda_{23})^2 + 4\tilde{\lambda}^2\right).$$

A calculation now shows that the second inequality is equivalent to (ii). $\qquad\square$

**Remark 2.2** By adapting the arguments in the proof of Proposition 2.1, one can also obtain equivalent characterisations in the cases $k = 1, n-2, n-1$. We omit this as it is not needed in this article.

## 3 Perelman's neck construction

In this section we prove the following result, which is the main ingredient in the proof of Theorem A.

**Theorem 3.1** *For any $v > 0$ sufficiently small, $\ell \in \mathbb{N}$, $n \geq 3$ and all $k \geq 2$ there exists a metric of $\mathrm{Ric}_k > 0$ on $S^n \setminus \bigsqcup_\ell (D^n)^\circ$ such that the induced metric on each boundary component is the round metric of radius one and the principal curvatures are all given by $-v$.*

The construction of the metric in Theorem 3.1 follows that of [14] and consists of two parts: First, the *ambient space*, which is a metric of positive sectional curvature on $S^n \setminus \bigsqcup_\ell (D^n)^\circ$, where the metric on each boundary component is a warped product metric whose "waist" can be chosen arbitrarily small and with principal curvatures all at least -1. It is already established in [14] that the metric has positive sectional curvature. Second, the *neck*, which is a metric on $S^{n-1} \times [0, 1]$ connecting the metrics on the boundary components of the ambient space to round metrics with constant and arbitrarily small second fundamental form. This metric on the neck is shown to have positive Ricci curvature in [14] and we show below that it has in fact $\mathrm{Ric}_2 > 0$:

**Proposition 3.2** *Let $g$ be a metric on $S^n$, $n \geq 2$, of the form*

$$g = dt^2 + B^2(t) ds_{n-1},$$

*where $t \in [0, \pi R]$, and we set $r = \max_t B(t)$. Assume that $g$ has sectional curvatures greater than 1 and suppose that $r < R^2$. Let $\rho \in (r^{1/2}, R)$. Then there exists a metric of $\mathrm{Ric}_2 > 0$ on $S^n \times [0, 1]$ such that*

  (i)  *the induced metric on $S^n \times \{0\}$ is the round metric of radius $\frac{\rho}{\lambda}$ and satisfies $\mathrm{II} \equiv -\lambda$ for some $\lambda > 0$,*

  (ii) *the induced metric on $S^n \times \{1\}$ is isometric to $g$ and satisfies $\mathrm{II} > 1$.*

The metric we will construct in the proof of Proposition 3.2 is of the form

$$dt^2 + A(t, x)^2 dx^2 + B(t, x)^2 ds_m^2,$$

where $dx^2$ denotes the standard metric on $S^1$. We first compute the curvatures of such a metric.

**Lemma 3.3** *Let $t_0 < t_1$ and denote by $dt^2$ the standard metric on $[t_0, t_1]$, and by $dx^2$ the standard metric on $S^1$. Let $A, B: [t_0, t_1] \times S^1 \to \mathbb{R}_{>0}$ be smooth positive functions and define the metric $g$ on $[t_0, t_1] \times S^1 \times S^m$ by*

$$g = dt^2 + A(t, x)^2 dx^2 + B(t, x)^2 ds_m^2.$$

*Let $v_1, v_2$ denote vectors tangent to $S^m$. Then the curvature tensor of $g$ is given by*

$$\mathcal{R}(\partial_t \wedge \partial_x) = -\frac{A_{tt}}{A} \partial_t \wedge \partial_x,$$

$$\mathcal{R}(\partial_t \wedge v_1) = -\frac{B_{tt}}{B} \partial_t \wedge v_1 + \left( -\frac{B_{xt}}{A^2 B} + \frac{A_t B_x}{A^3 B} \right) \partial_x \wedge v_1,$$

$$\mathcal{R}(\partial_x \wedge v_1) = \left( -\frac{B_{xt}}{B} + \frac{A_t B_x}{AB} \right) \partial_t \wedge v_1 + \left( -\frac{A_t B_t}{AB} - \frac{B_{xx}}{A^2 B} + \frac{A_x B_x}{A^3 B} \right) \partial_x \wedge v_1,$$

$$\mathcal{R}(v_1 \wedge v_2) = \left( \frac{1 - B_t^2}{B^2} - \frac{B_x^2}{A^2 B^2} \right) v_1 \wedge v_2.$$

**Proof** By using the Koszul formula one easily verifies that the Levi-Civita connection of $g$ is given by

$$\nabla_{\partial_t} \partial_t = 0,$$

$$\nabla_{\partial_t} \partial_x = \nabla_{\partial_x} \partial_t = \frac{A_t}{A} \partial_x,$$

$$\nabla_{\partial_t} v_1 = \nabla_{v_1} \partial_t = \frac{B_t}{B} v_1,$$

$$\nabla_{\partial_x} \partial_x = -AA_t \partial_t + \frac{A_x}{A} \partial_x,$$

$$\nabla_{\partial_x} v_1 = \nabla_{v_1} \partial_x = \frac{B_x}{B} v_1,$$

$$\nabla_{v_1} v_2 = -ds_m^2(v_1, v_2) \left( BB_t \partial_t + \frac{BB_x}{A^2} \partial_x \right) + \nabla_{v_1}^{S^m} v_2.$$

From this one can now calculate the full curvature tensor. □

**Proof of Proposition 3.2** We use the same metric as constructed in [14, Section 2]. This metric is constructed as follows.

We rewrite the metric $g$ as

$$g = r^2 \cos^2(x) ds_{n-1}^2 + A^2(x) dx^2,$$

$x \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right]$, where $A$ satisfies $A\left( \pm \frac{\pi}{2} \right) = r$, $A'\left( \pm \frac{\pi}{2} \right) = 0$. Then, since

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} A(x) \, dx = \pi R,$$

there exists $x \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right]$ with $A(x) \geq R \, (> r) \, (R < 1$ by the theorem of Bonnet and Myers). Hence, we can rewrite $A$ as

$$A(x) = r(1 - \eta(x) + \eta(x)a_\infty),$$

where $\eta$ is a function satisfying $\max_x \eta(x) = 1$, $\eta\left( \pm \frac{\pi}{2} \right) = 0$ and $\eta'\left( \pm \frac{\pi}{2} \right) = 0$, and $a_\infty \in \mathbb{R}$ with $a_\infty \geq \frac{R}{r}$.

For $t_0 < t_\infty$ we define the metric $g_{t_0, t_\infty}$ on $S^n \times [t_0, t_\infty]$ by

$$g_{t_0, t_\infty} = dt^2 + A^2(t, x) dx^2 + B^2(t, x) ds_{n-1}^2,$$

where

$$B(t, x) = tb(t) \cos(x), \quad A(t, x) = tb(t)(1 - \eta(x) + \eta(x)a(t))$$

and $a, b$ are functions satisfying

$$a(t_0) = 1, \quad a'(t_0) = 0, \quad b(t_0) = \rho, \quad b'(t_0) = 0, \quad a(t_\infty) = a_\infty > 1, \quad b(t_\infty) > r.$$

This metric will later be rescaled by $r/(t_\infty b(t_\infty))$ to satisfy the required properties.

Using Lemma 3.3 we see that the curvature operator $\mathcal{R}_{g_{t_0,t_\infty}}$ of this metric has the form of the map $A$ in Proposition 2.1 with

$$\lambda_{12} = -\frac{A_{tt}}{A} = K(\partial_t \wedge \partial_x),$$

$$\lambda_{13} = -\frac{B_{tt}}{B} = K(\partial_t \wedge v),$$

$$\tilde{\lambda} = -\frac{B_{xt}}{A^2 B} + \frac{A_t B_x}{A^3 B} = \frac{1}{n-1} \operatorname{Ric}(\partial_t, \partial_x),$$

$$\lambda_{23} = -\frac{A_t B_t}{AB} - \frac{B_{xx}}{A^2 B} + \frac{A_x B_x}{A^3 B} = K(\partial_x \wedge v),$$

$$\lambda_3 = \frac{1 - B_t^2}{B^2} - \frac{B_x^2}{A^2 B^2} = K(v_1 \wedge v_2).$$

Here $v, v_1, v_2$ are tangent to $S^n$.

The functions $a$ and $b$ are now explicitly defined by

$$\frac{b'}{b} = -\frac{\beta(t - t_0)}{2t_0^2 \ln(2t_0)}, \quad t_0 \le t \le 2t_0,$$

$$\frac{b'}{b} = -\frac{\beta \ln(2t_0)}{t \ln(t)^2}, \quad t \ge 2t_0,$$

$$\frac{a'}{a} = -\alpha \frac{b'}{b}, \quad t \ge t_0.$$

The constants $\alpha$ and $\beta$ are defined by

$$\beta = (1 - \epsilon) \frac{\ln(\rho) - \ln(r)}{1 + \frac{1}{4 \ln(2t_0)}},$$

$$\alpha = \frac{(1 + \delta)}{\beta} \frac{\ln(a_\infty)}{1 + \frac{1}{4 \ln(2t_0)}} = \frac{(1 + \delta)}{(1 - \epsilon)} \frac{\ln(a_\infty)}{\ln(\rho/r)}$$

for some $\epsilon, \delta > 0$ small. These values imply that $\int_{t_0}^\infty b'/b = (1-\epsilon)(\ln r - \ln \rho)$ and $\int_{t_0}^\infty a'/a = (1+\delta) \ln a_\infty$.

Similarly as in [14] we estimate $\alpha$ as follows: At a maximum point of $\eta$ we have $\eta(x) = 1$ and $\eta'(x) \tan(x) = 0$. Hence, the sectional curvatures of $g$ at this point satisfy (eg by applying Lemma 3.3)

$$K_g(\partial_x \wedge v) = \frac{1}{A(x)^2} \left( 1 - \frac{\sin(x) A'(x)}{\cos(x) A(x)} \right) = \frac{1}{r^2 a_\infty^2}.$$

Since $K_g > 1$, it follows that $a_\infty < 1/r$. Thus,

$$\ln(a_\infty) < \ln\left(\frac{1}{r}\right) < \ln\left(\frac{1}{r} \frac{\rho^2}{r}\right) = 2 \ln\left(\frac{\rho}{r}\right).$$

Hence, $\ln(a_\infty)/\ln(\rho/r) < 2$.

We also have $a_\infty \ge R/r > \rho/r$, so that $\ln(a_\infty) > \ln(\rho/r)$. Hence, for $\epsilon$ and $\delta$ sufficiently small, $\alpha \in (1, 2)$.

By choosing $\epsilon$ smaller if necessary, we can assume that $(\rho/r)^\epsilon g$ still has sectional curvatures at least 1. The following estimates are established in [14, Section 2] for $t_0$ sufficiently large (see also [4, Lemma 2.6, Corollaries C.2.9 and C.3.3]):

$$\lambda_{23}, \lambda_3 \geq \frac{c_1}{t^2},$$

$$|\lambda_{12}|, |\lambda_{13}|, |\tilde{\lambda}| \leq \frac{c_2 \ln(t_0)}{t^2 \ln(t)^2}$$

for some $c_1, c_2 > 0$. To estimate $\lambda_{13}$ a calculation now shows that

$$\lambda_{13} = -\left(\frac{b''}{b} + \frac{2b'}{tb}\right) \geq \frac{c_3 \ln(t_0)}{t^2 \ln(t)^2}$$

for some $c_3 > 0$.

By Proposition 2.1 we need to satisfy

(6) $$\lambda_{12} + \tfrac{1}{2}(\lambda_{13} + \lambda_{23}) > 0,$$

(7) $$(\lambda_{12} + \lambda_{13})(\lambda_{12} + \lambda_{23}) > \tilde{\lambda}^2,$$

(8) $$\lambda_{13}\lambda_{23} > \tilde{\lambda}^2,$$

(9) $$\lambda_{13}, \lambda_{23}, \lambda_3 > 0.$$

From the above estimates it follows directly that (6), (8) and (9) are satisfied for $t_0$ sufficiently large. For (7) we show that

$$\lambda_{12} + \lambda_{13} > \frac{c_4 \ln(t_0)}{t^2 \ln(t)^2}$$

for some $c_4 > 0$, from which (7) follows. We calculate

$$\lambda_{12} + \lambda_{13} = \left(\frac{\alpha\eta a}{1 - \eta + \eta a} - 2\right)\left(\left(\frac{b'}{b}\right)' + \frac{2b'}{tb}\right) - 2\left(\frac{b'}{b}\right)^2 - \frac{\eta a}{1 - \eta + \eta a}\left(2\frac{a'b'}{ab} + \left(\frac{a'}{a}\right)^2\right).$$

Similarly as in [14, end of page 161] we see that, since $\alpha < 2$ and $\eta \leq 1$, the first factor in the first summand is negative and uniformly bounded from above. Hence, the first summand is bounded from below by $(c_5 \ln(t_0))/(t^2 \ln(t)^2)$ for some $c_5 > 0$ and the absolute value of the remaining terms is bounded from above by $(c_6 \ln(t_0))/(t^2 \ln(t)^4)$ for some $c_6 > 0$. It follows that the required estimate holds for $t_0$ sufficiently large. Thus, the metric has $\mathrm{Ric}_2 > 0$ for $t_0$ sufficiently large.

Note that $\delta$ can still be chosen freely (which then determines $t_\infty$ via $a(t_\infty) = a_\infty$). This is now done as in [14] to ensure that the required conditions on the principal curvatures are satisfied. $\square$

We can now give the proof of Theorems 3.1 and A. For this, we recall the following gluing theorem which was established in [17].

**Theorem 3.4** [17, Theorem A] *Let $(M_1^n, h_1)$ and $(M_2^n, h_2)$ be Riemannian manifolds of $\mathrm{Ric}_k > 0$ for some $1 \le k \le n-1$ with compact boundaries, and let $\phi : (\partial M_1, h_1|_{\partial M_1}) \to (\partial M_2, h_2|_{\partial M_2})$ be an isometry. If the sum of second fundamental forms $\mathrm{II}_{\partial M_1} + \phi^* \mathrm{II}_{\partial M_2}$ is positive semidefinite, then $M_1 \cup_\phi M_2$ admits a smooth metric of $\mathrm{Ric}_k > 0$ which coincides with the $C^0$-metric $h = h_1 \cup_\phi h_2$ outside an arbitrarily small neighbourhood of the gluing area.*

We will also need the following result of Perelman:

**Proposition 3.5** [14, Section 3] *For every $n \ge 3$, $\ell \ge 0$, $R_0 \in (0,1)$ and $r > 0$ sufficiently small there exists a metric $g$ on $S^n \setminus \bigsqcup_\ell (D^n)^\circ$ such that*

(i) *$g$ has positive sectional curvature,*

(ii) *the induced metric on each boundary component is of the form $dt^2 + B(t)^2 ds_{n-2}^2$ with $t \in [0, \pi \cos(r)]$ and $\max_t B(t) = \cos(r) R_0 \sin(r + r^4/4)/\sin(r)$, and has sectional curvature at least 1, and*

(iii) *the principal curvatures at each boundary are all at least $-1$.*

**Proof of Theorem 3.1** We equip $S^n \setminus \bigsqcup_\ell (D^n)^\circ$ with the metric provided by Proposition 3.5, where $R_0$ is so small that $R_0 < \nu^2$, and $r$ is so small so that $\cos(r) > \nu$ and $\cos(r) R_0 \sin(r + r^4)/\sin(r) < \nu^2$. Hence, using Theorem 3.4, we can glue a copy of the neck obtained in Proposition 3.2 to each of the $\ell$ boundary components of $S^n \setminus \bigsqcup_\ell (D^n)^\circ$ to obtain a metric of $\mathrm{Ric}_2 > 0$ on the resulting manifold. Note that $\cos(r)$ in Proposition 3.5 corresponds to $R$ in Proposition 3.2 and $\cos(r) R_0 \sin(r + r^4)/\sin(r)$ in Proposition 3.5 corresponds to $r$ in Proposition 3.2, and we choose $\rho = \nu$. Finally, we rescale the metric by $\lambda/\rho$ so that the induced metric on each boundary component is the round metric of radius 1 and the principal curvatures are all given by $-\rho = -\nu$. $\square$

**Proof of Theorem A** The proof is essentially similar to the proof of [5, Theorem B]. We denote by $\varphi_i : D^n \hookrightarrow M_i$ the embedding provided by Definition 1.2. We now slightly perturb the $k$-core metric on each $M_i \setminus \varphi_i(D^n)^\circ$, eg as in [4, Proposition 1.2.11], such that the second fundamental form is strictly positive. Let $\nu_0 > 0$ be the smallest principal curvature of all these metrics. Thus, by Theorem 3.4, we can glue each $M_i \setminus \varphi_i(D^n)^\circ$ to $S^n \setminus \bigsqcup_\ell (D^n)^\circ$ with the metric provided by Theorem 3.1 by choosing $\nu < \nu_0$. Hence, we obtain a metric of $\mathrm{Ric}_k > 0$ on the connected sum $M_1 \# \cdots \# M_\ell$. $\square$

# 4 $k$-core metrics

In this section we consider $k$-core metrics. We begin by restating Proposition 1.4.

**Proposition 4.1** *Let $M$ be a closed $n$-dimensional manifold that admits a $k$-core metric. Then $M$ is $(n-k)$-connected. In particular, if $k \le \lfloor \frac{n+1}{2} \rfloor$, then $M$ is a homotopy sphere.*

|  | $G$ | $K$ | $H$ |
|---|---|---|---|
| $\mathbb{C}P^n \setminus D^{2n}{}^\circ$ | $U(n)$ | $U(n-1)\,U(1)$ | $U(n-1)$ |
| $\mathbb{H}P^n \setminus D^{4n}{}^\circ$ | $\mathrm{Sp}(n)$ | $\mathrm{Sp}(n-1)\,\mathrm{Sp}(1)$ | $\mathrm{Sp}(n-1)$ |
| $\mathbb{O}P^2 \setminus D^{16}{}^\circ$ | $\mathrm{Spin}(9)$ | $\mathrm{Spin}(8)$ | $\mathrm{Spin}(7)$ |

Table 1: Cohomogeneity-one structure of projective spaces with a disc removed.

**Proof** Since the boundary of $M \setminus \varphi(D^n)^\circ$ has positive semidefinite second fundamental form, it follows from [22, Theorem 1] that $M \setminus \varphi(D^n)^\circ$ is obtained from $\varphi(S^{n-1})$ by attaching cells of dimension at least $n-k+1$. By viewing $\varphi(D^n)$ as a 0-cell, we obtain a CW structure for $M$ with no cells in dimensions between 1 and $n-k$. It follows that $M$ is $(n-k)$-connected.

Now if $k \leq \lfloor \frac{n+1}{2} \rfloor$, we obtain by Poincaré duality that $M$ is a closed simply connected manifold with nontrivial homology groups only in degrees 0 and $n$. Hence, $M$ is a homotopy sphere. $\square$

We will now consider examples of manifolds with $k$-core metrics and applications to plumbing.

## 4.1 Projective spaces

To prove Theorem B, we will adapt the construction in [7], where a metric of nonnegative sectional curvature and round totally geodesic boundary is constructed on $\mathbb{C}P^n$, $\mathbb{H}P^n$ and $\mathbb{O}P^2$ with a disc removed. We will follow [3, Sections 3 and 4] and also include the arguments for $\mathbb{C}P^n$ as they are entirely similar.

The key observation is that, by considering cohomogeneity-one actions on these spaces, they can all be written as a disc bundle $G \times_K D \to G/H$, where $H \subseteq K \subseteq G$ are compact Lie groups. Here $K$ acts by isometries on a Euclidean vector space $V$ with principal isotropy group $H$ via a representation $\rho \colon K \to O(V)$, and $D \subseteq V$ is the unit disc. The corresponding groups are given in Table 1; see [1, Section 6.3; 3, Section 4.1; 12, Example 1].

The representation $\rho$ is given by projection onto $U(1)$ (resp $\mathrm{Sp}(1)$) followed by inclusion into $O(2)$ (resp $O(4)$) for $\mathbb{C}P^n$ (resp $\mathbb{H}P^n$). For $\mathbb{O}P^2$ it is given by the covering map $\mathrm{Spin}(8) \to \mathrm{SO}(8)$.

We will construct a $k$-core metric on $G \times_K D$ by defining a $K$-invariant metric on $G \times D$, which then descends to $G \times_K D$ such that the projection $G \times D \to G \times_K D$ is a Riemannian submersion. On $G \times D$ we consider the metric

$$g = L + (dt^2 + f(t)^2 ds_m^2),$$

where $m = \dim(V) - 1$, $L$ is a left-invariant metric on $G$ which is $\mathrm{Ad}_K$-invariant and $f \colon [0, t_0] \to \mathbb{R}_{\geq 0}$ is a smooth function for some $t_0 > 0$ which is odd at $t = 0$ with $f'(0) = 1$ and $f(t) > 0$ for $t \in (0, t_0]$.

Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$ and $\mathfrak{k} = \mathfrak{h} \oplus \mathfrak{p}$ be $L$-orthogonal decompositions of the Lie algebras. For $X \in \mathfrak{k}$, $t \in [0, t_0]$ and $v \in S^m$ we denote by $X_{tv}^*$ the action field at $tv \in D$ defined by $X$, ie $X_{tv}^* = \frac{d}{ds} \rho(\exp_K(sX))(tv)|_{s=0}$.

Then the vertical and horizontal subspaces $\mathcal{V}_{(e,tv)}$ and $\mathcal{H}_{(e,tv)}$ of $T_{(e,tv)}(G \times D)$ with respect to $g$ are given for $t > 0$ by

$$\mathcal{V}_{(e,tv)} = (\mathfrak{h} \oplus \{0\}) \oplus \{(-X, X_{tv}^*) \mid X \in \mathfrak{p}\},$$

(10)
$$\mathcal{H}_{(e,tv)} = (\mathfrak{m} \oplus \{0\}) \oplus \{(f(t)^2 BY, Y_{tv}^*) \mid Y \in \mathfrak{p}\} \oplus \langle \partial_t \rangle,$$

where $B \colon \mathfrak{p} \to \mathfrak{p}$ is the $L$-symmetric and $\mathrm{Ad}_H$-linear automorphism defined by $L(X, BY) = ds_m^2(X_{tv}^*, Y_{tv}^*)$; compare [3, Equation (3.1)]. For $t = 0$ we have

$$\mathcal{V}_{(e,0)} = \mathfrak{k} \oplus \{0\},$$

(11)
$$\mathcal{H}_{(e,0)} = \mathfrak{m} \oplus T_0 D.$$

With this description we can now give the proof of Theorem B.

**Proof of Theorem B**  We equip $G$ with the left-$G$-invariant and right-$K$-invariant which induces the round metric on $G/H$ (this metric does not need to be normal homogeneous). For $\mathbb{C}P^n$ (resp $\mathbb{H}P^n$) the restriction to $U(1)$ (resp $\mathrm{Sp}(1)$) is then biinvariant, and hence it is the round metric of some radius. In particular, the map $B$ is a multiple of the identity map. For $\mathbb{O}P^2$, the action of $H$ on $\mathfrak{p}$ is irreducible, so $B$ is also a multiple of the identity map by Schur's lemma.

Hence, there exists $b \in \mathbb{R}$ so that $B = b \cdot \mathrm{Id}_\mathfrak{p}$. For $\epsilon > 0$ we now define the metric $L_\epsilon$ on $G$ via

$$L_\epsilon = (1 + \epsilon)L|_\mathfrak{k} + L|_\mathfrak{m},$$

so $L_\epsilon$ is again left-$G$-invariant and right-$K$-invariant and the map $B_\epsilon$ is given by $\frac{1}{1+\epsilon} b \cdot \mathrm{Id}_\mathfrak{p}$. Then the metric

$$g_\epsilon = L_\epsilon + (dt^2 + f(t)^2 ds_m^2)$$

on $G \times D$ induces a metric $\check{g}_\epsilon$ on $G \times_K D$ such that the projection $G \times D \to G \times_K D$ is a Riemannian submersion. The metric induced on a slice $G \times_K S^m = G \times_K (K/H) \cong G/H$ for $t > 0$ is then given by

$$\frac{f(t)^2 \frac{b}{1+\epsilon}}{1 + f(t)^2 \frac{b}{1+\epsilon}} L_\epsilon|_\mathfrak{p} + L_\epsilon|_\mathfrak{m} = (1 + \epsilon) \frac{f(t)^2 \frac{b}{1+\epsilon}}{1 + f(t)^2 \frac{b}{1+\epsilon}} L|_\mathfrak{p} + L|_\mathfrak{m};$$

see eg [3, Lemma 3.1; 7; 11]. In particular, if $f(t) = \sqrt{(1+\epsilon)/(b\epsilon)}$, then this metric coincides with the metric induced from $L$ on $G/H$, ie it is the round metric. Thus, we will assume from now on that for given $\epsilon$, the function $f$ (and the value of $t_0$) is chosen such that $f(t_0) = \sqrt{(1+\epsilon)/(b\epsilon)}$, so that the induced metric on the boundary of $G \times_K D$ is round. Moreover, we assume that $f'(t_0) \geq 0$, so the second fundamental form on the boundary is positive semidefinite.

We will now analyse the curvatures of the metric $\check{g}_\epsilon$ on $G \times_K D$. We assume that $f'' < 0$, so the metric $h_f = dt^2 + f(t)^2 ds_m^2$ on $D$ has positive sectional curvature. We choose $\epsilon$ sufficiently small such that the metric induced on $G/H$ by $L_\epsilon$ also has positive sectional curvature. It then follows that the metric $\check{g}_\epsilon$ has nonnegative sectional curvature; see [3, Lemma 4.1; 7]. Thus, to determine the smallest value $k$ for

which this metric has $\mathrm{Ric}_k > 0$, we only need to identify the 2-planes of vanishing curvature, ie for given $u \in T(G \times_K D)$ we need to determine the set

$$Z_u = \{v \in u^\perp \setminus \{0\} \mid \sec^{\check{g}_\epsilon}(u \wedge v) = 0\}.$$

Let $A$ denote the $A$-tensor of the Riemannian submersion $(G \times D, g_\epsilon) \to (G \times_K D, \check{g}_\epsilon)$ and decompose $A$ into $A = A^1 + A^2$ according to the splitting (10), ie $A^1$ has image in $\mathfrak{h} \oplus \{0\}$ and $A^2$ has image in $\{(-X, X_{tv}^*) \mid X \in \mathfrak{p}\}$. As in [3, Proof of Lemma 4.1] we conclude that for horizontal vectors $u = (u_1, u_2), v = (v_1, v_2)$ in $T_{(e,tv)}(G \times D)$ with $t > 0$, we have

$$A^1{}_u v = A^{G/H}{}_{u_1} v_1,$$

where $A^{G/H}$ is the $A$-tensor of the Riemannian submersion $G \to G/H$ (where we consider $G$ equipped with the metric $L_\epsilon$). It follows from the O'Neill formulas that

$$
\begin{aligned}
\check{g}_\epsilon(R^{\check{g}_\epsilon}(u, v)v, u) &= g_\epsilon(R^{g_\epsilon}(u, v)v, u) + 3\,|A_u v|^2 \\
&= L_\epsilon(R^{L_\epsilon}(u_1, v_1)v_1, u_1) + h_f(R^{h_f}(u_2, v_2)v_2, u_2) + 3\,|A^{G/H}_{u_1} v_1|^2 + 3\,|A^2_u v|^2 \\
&= L_\epsilon(R^{G/H}(u_1, v_1)v_1, u_1) + h_f(R^{h_f}(u_2, v_2)v_2, u_2) + 3\,|A^2_u v|^2.
\end{aligned}
$$

Since both the metric on $G/H$ and the metric $h_f$ have strictly positive sectional curvature, this expression can only vanish if the pairs $(u_1, v_1)$ and $(u_2, v_2)$ are both linearly dependant. If we write, according to (10),

$$
\begin{aligned}
u = (u_1, u_2) &= (X + f(t)^2 B_\epsilon Y, Y_{tv}^* + \lambda \partial_t), \\
v = (v_1, v_2) &= (X' + f(t)^2 B_\epsilon Y', Y'^*_{tv} + \lambda' \partial_t),
\end{aligned}
$$

this is satisfied if and only if there exist $a_1, a_2 \in \mathbb{R}$ such that

$$(X', Y') = a_1(X, Y) \quad \text{or} \quad (X, Y) = (0, 0), \qquad \text{and} \qquad (Y', \lambda') = a_2(Y, \lambda) \quad \text{or} \quad (Y, \lambda) = (0, 0).$$

If $Y \neq 0$, then $a_1 = a_2$, and hence $v = a_1 u$ and $Z_u$ is empty. Hence, we can assume that $Y = 0$. Then, if $X, \lambda \neq 0$, we have $X' = a_1 X$, $\lambda' = a_2 \lambda$ and $Y' = 0$, and hence $Z_u$ is contained in a 1-dimensional subspace. Thus, we are left with the cases $X = 0$, $\lambda \neq 0$ and $X \neq 0$, $\lambda = 0$. In the first case, we have $Y' = 0$ and $\lambda' = a_2 \lambda$, so $Z_u$ is contained in a $\dim(G/K)$-dimensional subspace. In the second case we have $Y' = 0$ and $X' = a_1 X$, so $Z_u$ is contained in a 1-dimensional subspace.

Hence, we have shown that $Z_u$ is contained in a $\dim(G/K)$-dimensional subspace for all $u \in T_{e,tv}(G \times_K D)$ and $t > 0$. By $G$-invariance of the metric $\check{g}_\epsilon$ this holds for all points $(g, tv)$ with $t > 0$. Similar arguments using (11) show that this result extends to the case $t = 0$. Thus, the metric $\check{g}_\epsilon$ has $\mathrm{Ric}_{\dim(G/K)+1} > 0$. For $\mathbb{C}P^n$ this gives a metric of $\mathrm{Ric}_{2n-1} > 0$, for $\mathbb{H}P^n$ a metric of $\mathrm{Ric}_{4n-3} > 0$ and for $\mathbb{O}P^2$ a metric of $\mathrm{Ric}_9 > 0$. $\qquad\square$

## 4.2  Generalised surgery and plumbing

To prove Theorem D we need two additional results: a surgery result extending [15, Theorem A; 19, Theorem 3.2] and a deformation result that ensures that we can satisfy the assumptions of the surgery

theorem in our setting. For $\rho > 0$ we denote by $S^p(\rho)$ the round sphere of radius $\rho$ and for $R, N > 0$ we denote by $D_R^{q+1}(N)$ a geodesic ball of radius $R$ in $S^{q+1}(N)$.

**Theorem 4.2** *Suppose we have*

(i) *a Riemannian manifold $(M^{p+q+1}, g_M)$ of $\mathrm{Ric}_{k_1} > 0$,*

(ii) *an isometric embedding $\iota\colon S^p(\rho) \times D_R^{q+1}(N) \hookrightarrow (M, g_M)$ (which implies $k_1 \geq \max(p+1, q+2)$),*

(iii) *a linear $S^q$-bundle $E \xrightarrow{\pi} B^{p+1}$, where $B$ is compact and admits a $k_2$-core metric $g_B$.*

*Then, if $p, q \geq 2$, for any $r > 0$ sufficiently small, there exists a constant $\kappa = \kappa(p, q, R/N, g_B, r)$, such that if $\frac{\rho}{N} < \kappa$, then the manifold*

$$M_{\iota,\pi} = M \setminus \mathrm{im}(\iota)^\circ \cup_\partial \pi^{-1}(B \setminus \varphi(D^{p+1})^\circ)$$

*admits a metric of $\mathrm{Ric}_k > 0$ for all $k \geq \max(p + 2, q + 2, q + k_2)$. This metric coincides outside the gluing area with a submersion metric on $E$ with totally geodesic round fibres of radius $r$ and a scalar multiple of the metric $g_M$ on $M$.*

**Proof** We equip $E$ with a submersion metric with totally geodesic and round fibres of radius $r$ according to a horizontal distribution which is integrable over $\varphi(D^{p+1}) \subseteq B$. Then, for $r$ sufficiently small, this metric has $\mathrm{Ric}_k > 0$ for all $k \geq \max(p + 2, q + k_2)$ by [19, Corollary 3.1]. Further, over $\varphi(D^{p+1})$, the metric is a product, in particular it is given over $\varphi(S^p)$ by $ds_p^2 + r^2 ds_q^2$. As noted below Definition 1.2, we can slightly deform the metric on $\pi^{-1}(B \setminus \varphi(D^{p+1})^\circ)$ so that the induced metric on the boundary remains unchanged and the second fundamental form on the boundary is strictly positive.

Now, by [18, Theorem C and Remark 4.2] there exists a metric of $\mathrm{Ric}_k > 0$ on the manifold

$$M_\iota = M \setminus \mathrm{im}(\iota)^\circ \cup_\partial (D^{p+1} \times S^q)$$

for all $k \geq \max(p, q) + 2$ such that the metric near the centre of $D^{p+1} \times S^q$ is given by $D_{R'}^{p+1}(N') \times S^q(\rho')$, where the values of $R', N', \rho'$ can be chosen freely — provided $\frac{R'}{N'} < \frac{\pi}{2}$. We choose $\rho' = r$ and $R', N'$ so that the induced metric on $\partial D_{R'}^{p+1}(N')$ is $ds_p^2$ and the principal curvatures at the boundary are at least $-\epsilon$ for given $\epsilon > 0$ — note that they converge to 0 as $\frac{R'}{N'} \to \frac{\pi}{2}$.

It follows that $D_{R'}^{p+1}(N') \times S^q(\rho')$ and $\pi^{-1}(B \setminus \varphi(D^{p+1})^\circ)$ have isometric boundaries, and for $\epsilon$ sufficiently small the principal curvatures of $\pi^{-1}(B \setminus \varphi(D^{p+1})^\circ)$ at the boundary are greater than those of $D_{R'}^{p+1}(N') \times S^q(\rho')$. Hence, by Theorem 3.4, we can replace $D^{p+1} \times S^q$ in $M_\iota$ by $\pi^{-1}(B \setminus \varphi(D^{p+1})^\circ)$ to construct $M_{\iota,\pi}$ while preserving $\mathrm{Ric}_k > 0$. $\square$

To satisfy Theorem 4.2(ii), we need the following deformation result, which generalises [21, Theorem 1.10].

**Lemma 4.3** *Let $(M^n, g_0)$ be a Riemannian manifold of $\mathrm{Ric}_k > 0$ and let $N^p \subseteq M$ be a compact embedded submanifold. Let $g_1$ be a metric of $\mathrm{Ric}_k > 0$ defined in a tubular neighbourhood $U$ of $N$. If the 1-jets of $g_0$ and $g_1$ on $N$ coincide, then there exists a metric $\tilde{g}$ of $\mathrm{Ric}_k > 0$ on $M$ that equals $g_0$ outside $U$ and equals $g_1$ on a (smaller) tubular neighbourhood of $N$.*

**Proof**  We consider for $t \in [0, 1]$ the metric $g_t = (1-t)g_0 + tg_1$ on $U$. Since the 1-jets of $g_0$ and $g_1$ coincide on $N$ and since the sectional curvatures depend linearly on the second derivatives of the metric, we have $K_{g_t} = (1-t)K_{g_0} + tK_{g_1}$ on $N$. In particular, $g_t$ has $\mathrm{Ric}_k > 0$ on $N$ and by compactness this holds in a neighbourhood of $N$. By [2, Theorem 1.2] the local deformation $g_t$ can now be extended to a global deformation of $g_0$, which leaves $g_0$ unchanged outside a neighbourhood of $N$ and coincides with the deformation $g_t$ on a (smaller) tubular neighbourhood of $N$. $\qquad\square$

**Corollary 4.4**  *Let $(M^n, g)$ be a Riemannian manifold of $\mathrm{Ric}_k > 0$ and let $p_1, \ldots, p_\ell \in M$. Then the metric $g$ can be deformed into a metric of $\mathrm{Ric}_k > 0$ that has constant sectional curvature 1 in a neighbourhood of each $p_i$.*

**Proof**  We consider normal coordinates around each $p_i$, ie coordinates $(x_1, \ldots, x_n)$ in which the metric is given by $g_{ab} = \delta_{ab} + O(r^2)$, where $r$ denotes the distance to $p_i$. In particular, the first derivatives $\partial_c g_{ab}$ all vanish at $p_i$. By considering normal coordinates at a point in the round sphere of radius 1, we obtain a second metric around each $p_i$ with the same property. Applying Lemma 4.3 now yields the required deformation. $\qquad\square$

**Proof of Theorem D**  The proof of Theorem D follows the same lines as the proof of [15, Theorem B] by observing that $\partial W$ is obtained by iterated generalised surgeries as in Theorem 4.2. We simply replace [15, Theorem A] by Theorem 4.2, [15, Proposition 2.2] by [19, Corollary 3.1] and the deformation result used in the proof of [15, Theorem B] by Corollary 4.4. $\qquad\square$

**Proof of Corollary E**  Let $\overline{E}_i \to B_i$ denote the disc bundle corresponding to $E_i \to B_i$. We define $W$ as the manifold obtained by plumbing according to the following graph, where we denote by $\underline{D}_M^m$ the trivial disc bundle $M \times D^m \to M$ over a manifold $M$:
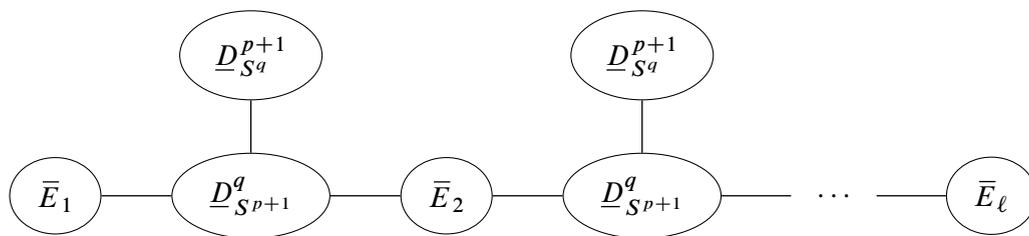


By [16, Propositions 3.2 and 3.3], see also [5, Section 5; 8, Proposition 2.6], the manifold $\partial W$ is diffeomorphic to the connected sum $E_1 \# \cdots \# E_\ell$. By Theorem D, the manifold $\partial W$ admits a metric of $\mathrm{Ric}_k > 0$ for all $k \geq \max\{p+2, p+k, q+1, q\} = \max\{p+2, p+k, q+1\}$. $\qquad\square$

**Remark 4.5**  In [6] it is shown that the manifolds constructed in Theorem D and Corollary E admit a core metric, provided each base manifold of the bundles involved admits a core metric. We conjecture that these manifolds in fact admit $k$-core metrics with $k$ as given in these results. However, this conjecture is open even in the simplest case of a linear sphere bundle over a manifold with a core metric (which can be viewed as a plumbing according to a graph with a single vertex).

# References

[1] **M M Alexandrino**, **R G Bettiol**, *Lie groups and geometric aspects of isometric actions*, Springer (2015) MR

[2] **C Bär**, **B Hanke**, *Local flexibility for open partial differential relations*, Comm. Pure Appl. Math. 75 (2022) 1377–1415 MR

[3] **R G Bettiol**, **R A E Mendes**, *Strongly nonnegative curvature*, Math. Ann. 368 (2017) 971–986 MR

[4] **B L Burdick**, *Metrics of positive Ricci curvature on connected sums*: *projective spaces*, *products*, *and plumbings*, PhD thesis, University of Oregon (2019) MR Available at `https://www.proquest.com/docview/2293989895`

[5] **B L Burdick**, *Ricci-positive metrics on connected sums of projective spaces*, Differential Geom. Appl. 62 (2019) 212–233 MR

[6] **B L Burdick**, *Metrics of positive Ricci curvature on the connected sums of products with arbitrarily many spheres*, Ann. Global Anal. Geom. 58 (2020) 433–476 MR

[7] **J Cheeger**, *Some examples of manifolds of nonnegative curvature*, J. Differential Geometry 8 (1973) 623–628 MR

[8] **D Crowley**, **D J Wraith**, *Positive Ricci curvature on highly connected manifolds*, J. Differential Geom. 106 (2017) 187–243 MR

[9] **M Gromov**, *Curvature, diameter and Betti numbers*, Comment. Math. Helv. 56 (1981) 179–195 MR

[10] **M Gromov**, **H B Lawson, Jr**, *The classification of simply connected manifolds of positive scalar curvature*, Ann. of Math. 111 (1980) 423–434 MR

[11] **K Grove**, **W Ziller**, *Curvature and symmetry of Milnor spheres*, Ann. of Math. 152 (2000) 331–367 MR

[12] **K Iwata**, *Compact transformation groups on rational cohomology Cayley projective planes*, Tohoku Math. J. 33 (1981) 429–442 MR

[13] **L Mouillé**, *Intermediate Ricci curvature database* Available at `https://sites.google.com/site/lgmouille/research/intermediate-ricci-curvature`

[14] **G Perelman**, *Construction of manifolds of positive Ricci curvature with big volume and large Betti numbers*, from "Comparison geometry" (K Grove, P Petersen, editors), Math. Sci. Res. Inst. Publ. 30, Cambridge Univ. Press (1997) 157–163 MR

[15] **P Reiser**, *Generalized surgery on Riemannian manifolds of positive Ricci curvature*, Trans. Amer. Math. Soc. 376 (2023) 3397–3418 MR

[16] **P Reiser**, *Metrics of positive Ricci curvature on simply-connected manifolds of dimension* $6k$, J. Topol. 17 (2024) art. id. e70007 MR

[17] **P Reiser**, **D J Wraith**, *A generalization of the Perelman gluing theorem and applications*, preprint (2023) arXiv 2308.06996

[18] **P Reiser**, **D J Wraith**, *Intermediate Ricci curvatures and Gromov's Betti number bound*, J. Geom. Anal. 33 (2023) art. id. 364 MR

[19] **P Reiser**, **D J Wraith**, *Positive intermediate Ricci curvature on fibre bundles*, Symmetry Integrability Geom. Methods Appl. 21 (2025) art. id. 006 MR

[20] **D Wraith**, *Exotic spheres with positive Ricci curvature*, J. Differential Geom. 45 (1997) 638–649 MR

[21]  **D Wraith**, *Deforming Ricci positive metrics*, Tokyo J. Math. 25 (2002) 181–189  MR

[22]  **H Wu**, *Manifolds of partially positive curvature*, Indiana Univ. Math. J. 36 (1987) 525–548  MR

*Department of Mathematics, University of Fribourg*
*Fribourg, Switzerland*

*Department of Mathematics and Statistics, Maynooth University*
*Maynooth, Ireland*

`philipp.reiser@unifr.ch`,  `david.wraith@mu.ie`

# The $\mathbb{S}_n$-equivariant Euler characteristic of the moduli space of graphs

MICHAEL BORINSKY

JOS VERMASEREN

We prove a formula for the $\mathbb{S}_n$-equivariant Euler characteristic of the moduli space of graphs $\mathcal{MG}_{g,n}$. Moreover, we prove that the rational $\mathbb{S}_n$-invariant cohomology of $\mathcal{MG}_{g,n}$ stabilizes for large $n$. That means, if $n \geq g \geq 2$, then there are isomorphisms $H^k(\mathcal{MG}_{g,n}; \mathbb{Q})^{\mathbb{S}_n} \to H^k(\mathcal{MG}_{g,n+1}; \mathbb{Q})^{\mathbb{S}_{n+1}}$ for all $k$.

14D22, 18G85, 58D29; 05E18, 14T99, 20F28, 20F65, 20J06

## 1 Introduction

A graph $G$ is a finite, at most one-dimensional CW complex. It has *rank* $g$ if its fundamental group is a free group of rank $g$: $\pi_1(G) \cong F_g$. Here, graphs shall be *admissible*, that means they do not have vertices of degree 0 or 2. Univalent vertices have a special role and are called *legs* (often also *hairs*, *leaves* or *marked points*). Similarly, we will reserve the name *edge* to 1-cells that are only incident to nonleg vertices. Legs of graphs are uniquely labeled by integers $\{1, \ldots, n\}$.

A *metric graph* is additionally equipped with a length $\ell(e) \geq 0$ for each edge $e$ such that all edge lengths sum to one, $\sum_e \ell(e) = 1$. Fix $g, n$ such that $g > 0$ and $2g - 2 + n > 0$. The moduli space of graphs, $\mathcal{MG}_{g,n}$, is the space of isometry classes of metric graphs of rank $g$ with $n$ legs, except for graphs that have a cycle in which all edges have length 0. It inherits the topology from the metric and by identifying each graph that has an edge $e$ of length 0 with the respective graph where $e$ is collapsed.

The moduli space of graphs was introduced in [15], where it was shown that $\mathcal{MG}_{g,0}$ serves as a rational classifying space for $\mathrm{Out}(F_g)$, the outer automorphism group of the free group of rank $g$. Further, $\mathcal{MG}_{g,n}$ can be seen as the moduli space of *pure tropical curves* [1] and it is relevant for *Feynman amplitudes*, central objects in quantum field theory [5].

A partition $\lambda \vdash n$ gives rise to both an irreducible representation $V_\lambda$ of the symmetric group $\mathbb{S}_n$ and a Schur polynomial $s_\lambda$, a symmetric polynomial in $\Lambda_n = \mathbb{Q}[x_1, \ldots, x_n]^{\mathbb{S}_n}$. The symmetric group acts on the cohomology of $\mathcal{MG}_{g,n}$ by permuting the leg-labels. So, $H^k(\mathcal{MG}_{g,n}; \mathbb{Q})$ is an $\mathbb{S}_n$-representation that we can decompose into irreducibles, ie there are integers $c_{g,\lambda}^k$ such that

$$H^k(\mathcal{MG}_{g,n}; \mathbb{Q}) \cong \bigoplus_{\lambda \vdash n} c_{g,\lambda}^k V_\lambda.$$

surjection $\Gamma_{g,n} \to \mathrm{Out}(F_g)$ that comes from forgetting the legs. The odd Euler characteristic is equivalent to the Euler characteristic of the moduli space of graphs $\mathcal{MG}_{g,n}$ that is fibered by the local system resulting from the det-representation of $\mathrm{Out}(F_n)$ as pointed out in [19]. Here, we will prove a formula for the $\mathbb{S}_n$-equivariant odd Euler characteristic, denoted as $e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n})$. It is defined analogously to $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ in (1) with the coefficients in $\widetilde{\mathbb{Q}}$ instead of $\mathbb{Q}$.

Morita, Sakasai, and Suzuki previously calculated the Euler characteristic, $e(\mathcal{MG}_{g,0}) = e(\mathrm{Out}(F_g))$ for $g \leq 11$ [21]. The outcome of recent work of Borinsky and Vogtmann was the asymptotic formula $e(\mathrm{Out}(F_g)) \sim -e^{-1/4}(g/e)^g/(g \log g)^2$ for large $g$, implying that the total dimension of the cohomology of $\mathrm{Out}(F_g)$ grows superexponentially with $g$ [8]. In this work, we enhance this result with explicit computations.

We leverage our effective formulas with the FORM programming language by Vermaseren [27], for instance, to compute $e(\mathrm{Out}(F_g))$ for all $g \leq 100$. The resulting values of $e(\mathrm{Out}(F_g))$ and $e^{\mathrm{odd}}(\mathrm{Out}(F_g))$ for $g \leq 15$ were published with [8]. The ancillary files in the online supplement include these numbers for all $g \leq 100$, together with tables of $e(\mathcal{MG}_{g,n})$, $e^{\mathrm{odd}}(\mathcal{MG}_{g,n})$, $e(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$, $e^{\mathrm{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ for all $g + n \leq 60$, and the polynomials $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$, $e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n})$ for all $g + n \leq 30$. Subsets of these data are given in Tables 2–5. In Section 2.7, we comment on our implementation of Theorem 2.18 in FORM, which is included in the online supplement. Based on our computed data, we give some empirical observations on the asymptotic growth rate of the Euler characteristic $e(\mathcal{MG}_{g,n})$ and its variations for large $g$ in Section 2.8.

Our data in Tables 4–5 exhibit an interesting pattern. The $\mathbb{S}_n$-invariant Euler characteristics of $\mathcal{MG}_{g,n}$ *stabilize* for large $n$. The search for an explanation of this pattern led to our second main finding: The $\mathbb{S}_n$-invariant cohomologies $H^\bullet(\mathcal{MG}_{g,n}; \mathbb{Q})^{\mathbb{S}_n}$ and $H^\bullet(\mathcal{MG}_{g,n}; \widetilde{\mathbb{Q}})^{\mathbb{S}_n}$ *stabilize* for large $n$. That means if $n \geq g \geq 2$ and $\mathbb{Q}_\rho \in \{\mathbb{Q}, \widetilde{\mathbb{Q}}\}$, then there are isomorphisms $H^k(\mathcal{MG}_{g,n}; \mathbb{Q}_\rho)^{\mathbb{S}_n} \to H^k(\mathcal{MG}_{g,n+1}; \mathbb{Q}_\rho)^{\mathbb{S}_n}$ for all $k$. This statement is proven as Theorem 3.1 in Section 3, where we also comment on an analogy to the cohomology of the braid group. Theorem 3.1 complements previous stability results for $\mathcal{MG}_{g,n}$ (see [13]). Here, in contrast to these previous results, the range of stabilization does not depend on the cohomological degree.

## Acknowledgements

# 2 The $\mathbb{S}_n$-equivariant Euler characteristic of $\mathcal{MG}_{g,n}$

## 2.1 Forested graph complexes

In this section we will describe chain complexes that compute the homology of $\mathcal{MG}_{g,n}$. These chain complexes are generated by certain graphs.

A subgraph is a subcomplex of a graph that consists of all its vertices, its nonedge 1-cells and a subset of its edges. A subforest is an acyclic subgraph. A pair $(G, \Phi)$ of a graph $G$ and a subforest $\Phi \subset G$ is a *forested graph*. We write $|\Phi|$ for the number of edges in the forest $\Phi$. Figure 1 shows some examples of forested graphs with different rank, forest edge and leg numbers. The forest edges are drawn thicker and in blue. Legs are drawn as labeled half-edges.

A $(+)$-marking $\sigma_\Phi$ of $(G, \Phi)$ is an ordering of the forest edges, ie a bijection $\sigma_\Phi \colon E_\Phi \to \{1, \ldots, |\Phi|\}$. A $(-)$-marking $(\sigma_\Phi, \sigma_{H_1(G, \mathbb{Z})})$ of $(G, \Phi)$ is such an ordering of the forest edges $\sigma_\Phi$ together with a basis for the first homology of the graph, ie a bijection $\sigma_{H_1} \colon H_1(G, \mathbb{Z}) \to \mathbb{Z}^{h_1(G)}$, where $h_1(G) = \mathrm{rk}\, H_1(G, \mathbb{Z})$ is the rank of $H_1(G, \mathbb{Z})$ or equivalently the rank of $G$.

**Definition 2.1** For given $g, n \geq 0$ with $2g - 2 + n \geq 0$, we define two $\mathbb{Q}$-vector spaces:

- $\mathcal{F}_{g,n}^+$ is generated by tuples $(G, \Phi, \sigma_\Phi)$ of a connected admissible forested graph $(G, \Phi)$ of rank $g$ with $n$ legs, which is $(+)$-marked with $\sigma_\Phi$, modulo the relation

$$(G, \Phi, \pi \circ \sigma_\Phi) \sim \mathrm{sign}(\pi) \cdot (G, \Phi, \sigma_\Phi) \quad \text{for all } \pi \in \mathbb{S}_{|\Phi|},$$

and modulo isomorphisms of $(+)$-marked forested graphs.

- $\mathcal{F}_{g,n}^-$ is generated by tuples $(G, \Phi, \sigma_\Phi, \sigma_{H_1})$ of a connected admissible forested graph $(G, \Phi)$ of rank $g$ with $n$ legs, which is $(-)$-marked with $\sigma_\Phi, \sigma_{H_1}$ modulo the relation

$$(G, \Phi, \pi \circ \sigma_\Phi, \rho_{H_1} \circ \sigma_{H_1}) \sim \mathrm{sign}(\pi) \cdot \det \rho_{H_1} \cdot (G, \Phi, \sigma_\Phi) \quad \text{for all } \pi \in \mathbb{S}_{|\Phi|} \text{ and } \rho_{H_1} \in \mathrm{GL}_{h_1(G)}(\mathbb{Z})$$

and modulo isomorphisms of $(-)$-marked forested graphs.

We will discuss the relations imposed in this definition and how they give rise to two different ways to impose *orientations* on graphs in more detail in Section 2.3.
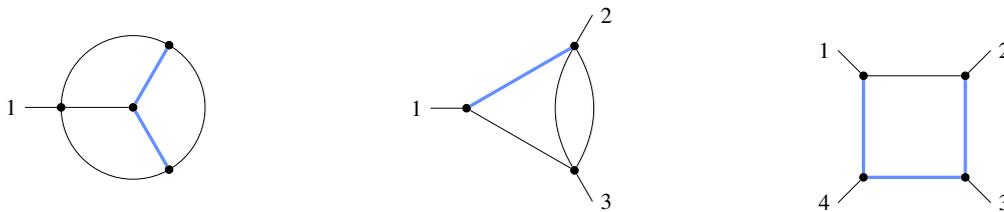


Figure 1: Examples of forested graphs. Left: forested graph of rank 3 with 1 leg and 2 forest edges. Middle: forested graph of rank 2 with 3 legs and 1 forest edge. Right: forested graph of rank 1 with 4 legs and 3 forest edges.

In what follows, we will often cover both the $(+)$- and $(-)$-cases analogously in the same sentence using the $\pm$-notation. Notice that the vector spaces $\mathcal{F}_{g,n}^{\pm}$ are defined for a slightly larger range of pairs $(g, n)$ then the moduli spaces $\mathcal{MG}_{g,n}$. This generalization makes the generating functions for the Euler characteristics in Section 2.6 easier to handle. There are no (admissible) graphs of rank 1 without legs, so $\mathcal{F}_{1,0}^{+} = \mathcal{F}_{1,0}^{-} = \varnothing$, but there is one admissible graph of rank 0 with two legs and no vertices or edges. It just consists of a 1-cell that connects the legs.

**Lemma 2.2**  *For all $g, n \geq 0$ with $2g - 2 + n > 0$, the vector spaces $\mathcal{F}_{g,n}^{\pm}$ are finite dimensional.*

**Proof**  Let $G$ be a connected admissible graph of rank $g$ with $n$ legs, $|E_G|$ edges, and $|V_G|$ vertices. The Euler characteristic of such a graph equals $|V_G| - |E_G| = 1 - g$. If we exclude the aforementioned graph with two legs and no vertices or edges, then each leg of $G$ is incident to a vertex. Each vertex has at most 3 incident edges or legs, so $3 |V_G| \leq n + 2 |E_G|$. It follows that $|E_G| \leq 3g - 3 + n$. There are only finitely many isomorphism classes of graphs of bounded edge number. Hence, there are also only finitely many isomorphism classes of forested graphs of rank $g$ with $n$ legs. The relations in Definition 2.1 guarantee that each such isomorphism class contributes at most one generator to $\mathcal{F}_{g,n}^{\pm}$. $\qquad\square$

The vector spaces $\mathcal{F}_{g,n}^{\pm}$ are graded by the number of edges in the forest. Let $C_k(\mathcal{F}_{g,n}^{\pm})$ be the respective subspace restricted to generators with $k$ forest edges. These spaces form a chain complex, so we will refer to $\mathcal{F}_{g,n}^{\pm}$ as the *forested graph complex* with $(\pm)$ orientation. We will not explicitly state the boundary maps, $\partial_k : C_k(\mathcal{F}_{g,n}^{\pm}) \to C_{k-1}(\mathcal{F}_{g,n}^{\pm})$ here (see [14]), as knowledge of the dimensions of these chain groups suffices for our Euler characteristic considerations.

The following theorem is a consequence of the works of Culler and Vogtmann [15], Kontsevich [19; 20] and Conant and Vogtmann [14] (see in particular [14, Sections 3.1–3.2]):

**Theorem 2.3**  *For $g > 0, n \geq 0$ and $2g - 2 + n > 0$, the chain complexes $\mathcal{F}_{g,n}^{+}$ ($\mathcal{F}_{g,n}^{-}$) compute the homology of $\mathcal{MG}_{g,n}$ with trivial coefficients $\mathbb{Q}$ (with twisted coefficients $\widetilde{\mathbb{Q}}$). The $\mathbb{S}_n$-action on $H^{\bullet}(\mathcal{MG}_{g,n}, \mathbb{Q})$ ($H^{\bullet}(\mathcal{MG}_{g,n}, \widetilde{\mathbb{Q}})$) descents obviously to $\mathcal{F}_{g,n}^{+}$ ($\mathcal{F}_{g,n}^{-}$) by permuting leg-labels.*

In what follows, we will use this theorem to prove a formula for $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ and $e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n})$.

## 2.2  Equivariant Euler characteristics

Recall that a permutation $\pi \in \mathbb{S}_n$ factors uniquely as a product of disjoint cycles. If the orders of these cycles are $\lambda_1, \ldots, \lambda_\ell$, then $(\lambda_1, \ldots, \lambda_\ell)$ is a partition of $n$ called the *cycle type* of $\pi$. For such a permutation $\pi \in \mathbb{S}_n$, we define the *power sum symmetric polynomial*, $p^{\pi} = p_{\lambda_1} \cdots p_{\lambda_\ell} \in \Lambda_n$ with $p_k = \sum_{i=1}^{n} x_i^k$. The *Frobenius characteristic* is a symmetric polynomial associated uniquely to a $\mathbb{S}_n$-representation $V$. It is defined by

$$\mathrm{ch}(\chi_V) = \frac{1}{n!} \sum_{\pi \in \mathbb{S}_n} \chi_V(\pi) p^{\pi},$$

where $\chi_V$ is the *character* associated to $V$ (see, eg, [25, §7.18]). The $\mathbb{S}_n$-*equivariant Euler characteristic* of $\mathcal{F}_{g,n}^\pm$ is the alternating sum over the Frobenius characteristics of $H_k(\mathcal{F}_{g,n}^\pm; \mathbb{Q})$. Note that this is consistent with the definition of $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ in (1) since $\mathrm{ch}(\chi_{V_\lambda}) = s_\lambda$.

As the spaces $\mathcal{F}_{g,n}^\pm$ are finite by Lemma 2.2, we can compute the equivariant Euler characteristic on the chain level:

**Proposition 2.4** *The $\mathbb{S}_n$-equivariant Euler characteristic of $\mathcal{F}_{g,n}^\pm$ is given by*

$$e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm) = \sum_k (-1)^k \, \mathrm{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)}),$$

*where $\mathrm{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)})$ is the Frobenius characteristic of $C_k(\mathcal{F}_{g,n}^\pm)$ as an $\mathbb{S}_n$-representation,*

$$\mathrm{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)}) = \frac{1}{n!} \sum_{\pi \in \mathbb{S}_n} \chi_{C_k(\mathcal{F}_{g,n}^\pm)}(\pi) p^\pi,$$

*and $\chi_{C_k(\mathcal{F}_{g,n}^\pm)}$ is the **character** associated to $C_k(\mathcal{F}_{g,n}^\pm)$.*

**Corollary 2.5** *For $g > 0$, $n \geq 0$ and $2g - 2 + n > 0$,*

$$e_{\mathbb{S}_n}(\mathcal{MG}_{g,n}) = e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^+) \quad \text{and} \quad e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n}) = e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^-).$$

**Proof** Follows directly from Theorem 2.3. $\qquad\square$

The other discussed Euler characteristics of $\mathcal{MG}_{g,n}$ can be obtained by evaluating the polynomials $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm)$ for certain values of $\bar{p} = p_1, p_2, \dots$. We write $\bar{p} = \bar{1}$ for the specification $p_1 = 1$, $p_2 = p_3 = \dots = 0$ and $\bar{p} = \mathbb{1}$ for the specification $p_1 = p_2 = \dots = 1$.

**Proposition 2.6** *For $g > 0$, $n \geq 0$ and $2g - 2 + n > 0$,*

$$e(\mathcal{MG}_{g,n}) = n! \cdot e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^+)|_{\bar{p}=\bar{1}}, \quad e^{\mathrm{odd}}(\mathcal{MG}_{g,n}) = n! \cdot e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^-)|_{\bar{p}=\bar{1}},$$
$$e(\mathcal{MG}_{g,n}^{\mathbb{S}_n}) = e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^+)|_{\bar{p}=\mathbb{1}}, \qquad e^{\mathrm{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n}) = e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^-)|_{\bar{p}=\mathbb{1}}.$$

**Proof** Recall the relationship between the character of the $\mathbb{S}_n$-representation $C_k(\mathcal{F}_{g,n}^\pm)$ and the symmetric polynomials $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm)$ from Proposition 2.4. Observe that $\chi_V(\mathrm{id}) = \dim V$, where id is the trivial permutation. To verify the first two equations, observe that substituting $p_1 = 1$ and $p_2 = p_3 = \dots = 0$, in the formula for $\mathrm{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)})$ of Proposition 2.4 amounts to restricting the sum over $\pi$ to $\pi = \mathrm{id}$, as only the trivial permutation has cycle-type $(1, 1, \dots, 1)$. For the second line recall that $\frac{1}{n!} \sum_{\pi \in \mathbb{S}_n} \chi_V(\pi) = \dim V^{\mathbb{S}_n}$, where $V^{\mathbb{S}_n}$ is the $\mathbb{S}_n$-invariant subspace of $V$. Additionally, use Corollary 2.5. $\qquad\square$

To explicitly compute $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm)$, we hence need to compute the Frobenius characteristic of the chain groups and to compute those we need to know the character $\chi_{C_k(\mathcal{F}_{g,n}^\pm)}$. To compute these characters, we need a more detailed description of the chain groups $C_k(\mathcal{F}_{g,n}^\pm)$ that are generated only by *orientable forested graphs*.

## 2.3 Orientable forested graphs

Roughly, the relations in Definition 2.1 have the effect that a chosen ordering of the forest edges or a chosen homology basis, *only matters up to its sign or its orientation*. We can therefore think of the generators as *oriented* forested graphs. Due to the relations, every forested graphs gives rise to at most one generator of $\mathcal{F}_{g,n}^+$ or $\mathcal{F}_{g,n}^-$. However, not every forested graph is *orientable* in this way, ie gives rise to a generator. Even though Definition 2.1 only involves connected forested graphs, we will define the notion of orientability here also for disconnected graphs. We will need it later.

Let $\mathrm{Aut}(G, \Phi)$ be the group of automorphisms of the forested graph $(G, \Phi)$. Each automorphism is required to fix the leg-labels and the forest. For instance, the graph in Figure 1, left, has one automorphism of order two that mirrors the graph on the horizontal axis. The graph in Figure 1, middle, has one automorphism that flips the two doubled edges and the graph in Figure 1, right, has no (leg-label-preserving) automorphisms.

Each $\alpha \in \mathrm{Aut}(G, \Phi)$ induces a permutation, $\alpha_\Phi \colon E_\Phi \to E_\Phi$ on the set of forest edges and an automorphism on the homology groups $\alpha_{H_0} \colon H_0(G, \mathbb{Z}) \to H_0(G, \mathbb{Z})$ and $\alpha_{H_1} \colon H_1(G, \mathbb{Z}) \to H_1(G, \mathbb{Z})$. For each (possibly disconnected) forested graph $(G, \Phi)$ with an automorphism $\alpha \in \mathrm{Aut}(G, \Phi)$, we define

$$\xi^+(G, \Phi, \alpha) = \mathrm{sign}(\alpha_\Phi) \text{ and } \xi^-(G, \Phi, \alpha) = \det(\alpha_{H_0}) \det(\alpha_{H_1}) \mathrm{sign}(\alpha_\Phi).$$

As connected graphs come with a canonical basis for $H_0(G, \mathbb{Z}) = \mathbb{Z}$ that cannot be changed by automorphisms, we always have $\det(\alpha_{H_0}) = 1$ for them. The other sign factors capture the signs of the relations in Definition 2.1 if the $(\pm)$-marking of the graph $(G, \Phi)$ is acted upon using $\alpha$ in the obvious way.

**Definition 2.7** A forested graph $(G, \Phi)$ is $(\pm)$-*orientable* if it has no automorphism $\alpha \in \mathrm{Aut}(G, \Phi)$ for which $\xi^\pm(G, \Phi, \alpha) = -1$.

For example, the automorphism of the forested graph in Figure 1, left, switches two forest edges and flips the orientation of a chosen homology basis. Hence, the graph is $(-)$-, but not $(+)$-orientable. The automorphism of the graph in Figure 1, middle, does not affect its subforest, but flips the orientation of its homology basis. So, it is $(+)$-, but not $(-)$-orientable. The forested graph in Figure 1, right, is both $(+)$- and $(-)$-orientable as it has no nontrivial automorphism. There are also forested graphs that are neither $(+)$- nor $(-)$-orientable.

**Proposition 2.8** *The chain groups $C_k(\mathcal{F}_{g,n}^\pm)$ are freely generated by isomorphism classes of connected $(\pm)$-orientable forested graphs of rank $g$ with $n$ legs and $k$ forest edges.*

**Proof** Definition 2.1 guarantees that each isomorphism class of a connected forested graph $(G, \Phi)$ contributes at most one generator, as all different $(\pm)$-markings of $(G, \Phi)$ are related via a permutation $\pi \in \mathbb{S}_{|\Phi|}$ (and $\rho_{H_1} \in \mathrm{GL}_{h_1(G)}(\mathbb{Z})$).

It remains to be proven that $(G, \Phi)$ does contribute a free generator if and only if it is $(\pm)$-orientable. We will only prove this for the $(+)$-orientation, as the $(-)$-orientation follows analogously. Recall

that for any isomorphism class of a $(+)$-marked graph $(G, \Phi, \sigma_\Phi)$, Definition 2.1 imposes the relation $(G, \Phi, \pi \circ \sigma_\Phi) \sim \text{sign}(\pi) \cdot (G, \Phi, \sigma_\Phi)$ for each $\pi \in \mathbb{S}_{|\Phi|}$. As $(G, \Phi, \sigma_\Phi)$ is given up to isomorphism, we have $(G, \Phi, \sigma_\Phi \circ \alpha_\Phi) = (G, \Phi, \sigma_\Phi)$ for each $\alpha \in \text{Aut}(G, \Phi)$ with $\alpha_\Phi$ the permutation that $\alpha$ induces on $\Phi$. If there is an automorphism $\alpha \in \text{Aut}(G, \Phi)$ such that $\alpha_\Phi$ is an odd permutation on $\Phi$, then the permutation $\pi := \sigma_\Phi \circ \alpha_\Phi \circ \sigma_\Phi^{-1} \in \mathbb{S}_{|\Phi|}$ is odd as well and we find that $(G, \Phi, \sigma_\Phi) = (G, \Phi, \sigma_\Phi \circ \alpha_\Phi) = (G, \Phi, \pi \circ \sigma_\Phi) \sim -(G, \Phi, \sigma_\Phi) \sim 0$. If there is no such automorphism, then for any $\alpha \in \text{Aut}(G, \Phi)$, $\text{sign}(\sigma_\Phi \circ \alpha_\Phi \circ \sigma_\Phi^{-1}) = 1$. Hence, the sign of a $(+)$-orientable and $(+)$-marked forested graph is fixed by any automorphism which therefore gives a generator. $\qquad\square$

## 2.4 The action of $\mathbb{S}_n$ on orientable graphs

The next step for the explicit evaluation of $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm)$ is to quantify the action of $\mathbb{S}_n$ that permutes the legs of the generators of $\mathcal{F}_{g,n}^\pm$.

Let $\text{UAut}(G, \Phi)$ be the set of automorphisms of a forested graph $(G, \Phi)$ with $n$ legs that are *allowed* to permute the leg-labels. For instance, for the graph in Figure 1, right, which has a trivial leg-label-preserving automorphism group, the group $\text{UAut}(G, \Phi)$ is generated by the automorphism that mirrors the graph vertically and permutes the leg-labels by $(12)(34)$. We have a map $\pi_G \colon \text{UAut}(G, \Phi) \to \mathbb{S}_n$, given by only looking at the induced permutation on the leg-labels. The kernel of this map is equal to $\text{Aut}(G, \Phi)$, the group of leg-label-fixing automorphisms of $(G, \Phi)$.

The $\mathbb{S}_n$-action gives a representation $\rho \colon \mathbb{S}_n \to \text{GL}(C_k(\mathcal{F}_{g,n}^\pm))$. For a specific $\pi \in \mathbb{S}_n$, the linear map $\rho(\pi) \in \text{GL}(C_k(\mathcal{F}_{g,n}^\pm))$ replaces each generator of $C_k(\mathcal{F}_{g,n}^\pm)$ with the respective generator where the leg labels are permuted by $\pi$. The character of $C_k(\mathcal{F}_{g,n}^\pm)$ is the composition of $\rho$ with the trace, $\chi_{C_k(\mathcal{F}_{g,n}^\pm)} = \text{Tr} \circ \rho$. As $\rho(\pi)$ maps each generator to a multiple of another generator, it is sufficient to look at generators that happen to be Eigenvectors of $\rho(\pi)$ to compute $\text{Tr}(\rho(\pi))$. Let $(G, \Phi, \sigma^\pm)$ be a connected $(\pm)$-orientable forested graph with a $(\pm)$-marking $\sigma^\pm$ corresponding to a generator of $C_k(\mathcal{F}_{g,n}^\pm)$. This generator is an Eigenvector of $\rho(\pi)$ if the forested graph $(G, \Phi)$ has a non-leg-label-fixing automorphism $\alpha \in \text{UAut}(G, \Phi)$ such that $\pi_G(\alpha) = \pi$. The following lemma describes the Eigenvalue corresponding to such an Eigenvector. It is either $+1$ or $-1$.

**Lemma 2.9** *If, for given $\pi \in \mathbb{S}_n$, the generator $(G, \Phi, \sigma^\pm) \in C_k(\mathcal{F}_{g,n}^\pm)$ is an Eigenvector of $\rho(\pi)$, then the corresponding Eigenvalue is $\xi^\pm(G, \Phi, \alpha)$, where $\alpha$ is some representative $\alpha \in \pi_G^{-1}(\pi)$.*

**Proof** By Proposition 2.8, $(G, \Phi)$ cannot have a leg-label-fixing automorphism that flips the orientation. However, an automorphism from the larger group $\text{UAut}(G, \Phi)$ can change the sign of the orientation (ie the sign of the ordering and basis given by the $(\pm)$-marking in Definition 2.1 can be flipped). The map $\rho(\pi)$ does so if $\xi^\pm(G, \Phi, \alpha) = -1$ for a representative of the kernel $\alpha \in \pi_G^{-1}(\pi)$. As $\alpha \mapsto \xi^\pm(G, \Phi, \alpha)$ gives a group homomorphism $\text{UAut}(G, \Phi) \to \mathbb{Z}/2\mathbb{Z}$ and as $\xi^\pm(G, \Phi, \alpha) = 1$ for all $\alpha \in \ker \pi_G$, it does not matter which representative we pick. $\qquad\square$

Summing over all such Eigenvalues of $\rho(\pi)$ gives the value of $\operatorname{Tr}(\rho(\pi)) = \chi_{C_k(\mathcal{F}_{g,n}^{\pm})}(\pi)$:

**Corollary 2.10** *For each $\pi \in \mathbb{S}_n$, the character of $C_k(\mathcal{F}_{g,n}^{\pm})$ is*

$$\chi_{C_k(\mathcal{F}_{g,n}^{\pm})}(\pi) = \sum \xi^{\pm}(G, \Phi, \alpha_{(G,\Phi,\pi)}),$$

*where we sum over all isomorphism classes of $(\pm)$-orientable forested graphs $(G, \Phi)$ with $k = |\Phi|$ forest edges, for which the preimage $\pi_G^{-1}(\pi) \subset \operatorname{UAut}(G, \Phi)$ is nonempty (ie $(G, \Phi)$ has at least one non-leg-label-fixing automorphism that permutes the leg labels by $\pi$), and $\alpha_{(G,\Phi,\pi)}$ is some representative of such an automorphism in $\pi_G^{-1}(\pi) \subset \operatorname{UAut}(G, \Phi)$.*

To continue, it is convenient to pass to a sum over all connected forested graphs without restrictions on the orientability to get better combinatorial control over the expression:

**Corollary 2.11** *For $g, n \geq 0$ with $2g - 2 + n \geq 0$ and $k \geq 0$, we have*

$$\chi_{C_k(\mathcal{F}_{g,n}^{\pm})}(\pi) = \sum_{[G,\Phi]} \frac{1}{|\operatorname{Aut}(G, \Phi)|} \sum_{\alpha \in \pi_G^{-1}(\pi)} \xi^{\pm}(G, \Phi, \alpha),$$

*where we sum over all isomorphism classes of (not necessarily orientable) connected forested graphs $[G, \Phi]$ of rank $g$, $n$ legs and $k = |\Phi|$ forest edges.*

**Proof** As $\alpha \mapsto \xi^{\pm}(G, \Phi, \alpha)$ also gives a map $\operatorname{Aut}(G, \Phi) \to \mathbb{Z}/2\mathbb{Z}$, we have by Definition 2.7

$$\sum_{\alpha \in \ker \pi_G} \xi^{\pm}(G, \Phi, \alpha) = \begin{cases} |\operatorname{Aut}(G, \Phi)| & \text{if } (G, \Phi) \text{ is } (\pm)\text{-orientable,} \\ 0 & \text{otherwise.} \end{cases}$$

The statement follows, as $\sum_{\alpha \in \pi_G^{-1}(\pi)} \xi^{\pm}(G, \Phi, \alpha) = \xi^{\pm}(G, \Phi, \alpha') \sum_{\alpha \in \ker \pi_G} \xi^{\pm}(G, \Phi, \alpha)$, for any representative $\alpha' \in \pi_G^{-1}(\pi)$ and the formula from Corollary 2.10 for $\chi_{C_k(\mathcal{F}_{g,n}^{\pm})}(\pi)$. $\qquad \square$

With this we finally obtain our first explicit formula for $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^{\pm})$:

**Theorem 2.12** *For $g, n \geq 0$ with $2g - 2 + n \geq 0$,*

$$e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^{\pm}) = \sum_{[G,\Phi]_U} \frac{(-1)^{|\Phi|}}{|\operatorname{UAut}(G, \Phi)|} \sum_{\alpha \in \operatorname{UAut}(G, \Phi)} \xi^{\pm}(G, \Phi, \alpha) p^{\pi_G(\alpha)},$$

*where we sum over all isomorphism classes of connected forested graphs $[G, \Phi]_U$ of rank $g$ with $n$ **unlabeled** legs.*

**Proof** We can plug the statement of Corollary 2.11 into the definition of the Frobenius characteristic in Proposition 2.4 to get

$$\text{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)}) = \frac{1}{n!} \sum_{\pi \in \mathbb{S}_n} p^\pi \sum_{[G,\Phi]} \frac{1}{|\text{Aut}(G,\Phi)|} \sum_{\alpha \in \pi_G^{-1}(\pi)} \xi^\pm(G,\Phi,\alpha).$$

Next, we can merge the first and the third sum into a summation over the whole group $\text{UAut}(G,\Phi)$, as the preimages of different $\pi \in \mathbb{S}_n$ partition $\text{UAut}(G,\Phi)$:

$$\text{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)}) = \frac{1}{n!} \sum_{[G,\Phi]} \frac{1}{|\text{Aut}(G,\Phi)|} \sum_{\alpha \in \text{UAut}(G,\Phi)} \xi^\pm(G,\Phi,\alpha) \, p^{\pi_G(\alpha)}.$$

Let $L_S(G,\Phi)$ be the set of isomorphism classes of forested graphs $(G,\Phi)$ that only differ by a relabeling of the legs. Obviously, $\mathbb{S}_n$ acts transitively on $L_S(G,\Phi)$ by permuting the leg-labels. The stabilizer of this action is the image of $\pi_G\colon \text{UAut}(G,\Phi) \to \mathbb{S}_n$. By the orbit-stabilizer theorem, $n! = |\mathbb{S}_n| = |\text{im}\,\pi_G||L_S(G,\Phi)|$. By the short exact sequence, $1 \to \text{Aut}(G,\Phi) \to \text{UAut}(G,\Phi) \to \text{im}\,\pi_G \to 1$, we have $|\text{im}\,\pi_G| = |\text{UAut}(G,\Phi)|/|\text{Aut}(G,\Phi)|$. Hence,

$$\text{ch}(\chi_{C_k(\mathcal{F}_{g,n}^\pm)}) = \sum_{[G,\Phi]} \frac{1}{|L_S(G,\Phi)||\text{UAut}(G,\Phi)|} \sum_{\alpha \in \text{UAut}(G,\Phi)} \xi^\pm(G,\Phi,\alpha) \, p^{\pi_G(\alpha)}.$$

The terms in this sum do not depend on the leg-labeling of the graphs, so we can just sum over non-leg-labeled graphs and remove the $|L_S(G,\Phi)|$ in the denominator. $\qquad\square$

## 2.5 Disconnected forested graphs

The expression for $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm)$ in Theorem 2.12 involves a sum over *connected* forested graphs without leg-labels. To eventually get an effective generating function for $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^\pm)$, it is convenient to pass to an analogous formula that sums over *disconnected* forested graphs. Moreover, it is helpful to change from grading the graphs by rank to grading them by their negative Euler characteristic. For connected graphs, this is just a trivial shift as the negative Euler characteristic of such graphs fulfills $\text{rk}\, H_1(G,\mathbb{Z}) - \text{rk}\, H_0(G,\mathbb{Z}) = g - 1$. So, we define for all $t \in \mathbb{Z}$ and $n \geq 0$,

$$(2) \qquad \hat{e}_{t,n}^\pm = \sum_{[G,\Phi]_U} \frac{1}{|\text{UAut}(G,\Phi)|} \sum_{\alpha \in \text{UAut}(G,\Phi)} \xi^\pm(G,\Phi,\alpha) \, p^{\pi_G(\alpha)},$$

where we sum over all isomorphism classes of (possibly disconnected) forested graphs $[G,\Phi]_U$ of Euler characteristic $\text{rk}\, H_0(G,\mathbb{Z}) - \text{rk}\, H_1(G,\mathbb{Z}) = -t$ and $n$ unlabeled legs. There is only a finite number of such graphs, so the sum is finite and $\hat{e}_{t,n}^\pm$ is a symmetric polynomial in $\Lambda_n$.

Let $\hat{\Lambda}$ be the *ring of formal symmetric power series* $\hat{\Lambda} = \lim_{\leftarrow n} \mathbb{Q}[\![x_1,\ldots,x_n]\!]^{\mathbb{S}_n}$. Later in this section we will prove a generating function, expressed as a power series with coefficients in $\hat{\Lambda}$, for the polynomials $\hat{e}_{t,n}^\pm$. Before that, we explain how we can translate them into our desired $\mathbb{S}_n$-equivariant Euler characteristics:

**Proposition 2.13**  *We define, over the ring of formal symmetric power series, the Laurent series*

$$e^{\pm}(\hbar, \bar{p}) = \sum_{\substack{g,n \geq 0 \\ 2g-2+n \geq 0}} e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^{\pm})(\pm\hbar)^{g-1}, \quad E^{\pm}(\hbar, \bar{p}) = \sum_{t \in \mathbb{Z}} \sum_{n \geq 0} \hat{e}_{t,n}^{\pm}(\pm\hbar)^t,$$

*which are elements of $\hat{\Lambda}((\hbar))$. Both are related by the **plethystic exponential***

$$E^{\pm}(\hbar, \bar{p}) = \exp\Big( \sum_{k \geq 1} \frac{e^{\pm}(\hbar^k, \bar{p}_{[k]})}{k} \Big),$$

*where $e^{\pm}(\hbar^k, \bar{p}_{[k]})$ denotes the power series $e^{\pm}$ with the substitutions $\hbar \to \hbar^k$ and $p_i \to p_{ik}$.*

**Proof**  The combinatorial argument for [8, Proposition 3.2] applies to translate between the sum over *connected* forested graphs in Theorem 2.12 to a sum over disconnected forested graphs in (2). The strategy goes back to Pólya [22] (see also [4, Chapter 4.3]). Briefly, each summand $e^{\pm}(\hbar^k, \bar{p}_{[k]})/k$ in the exponent of the stated formula counts pairs consisting of a $k$-tuple of mutually isomorphic forested graphs and an automorphism that cyclically permutes the different graphs. Such automorphisms give rise to different sign factors depending on the orientation of the graphs. Accounting for these signs gives the minus signs in front of $\hbar$ in the $(-)$-orientation case (see [8, Theorem 5.1]).  □

We can solve the equation in the statement above for the generating functions $e^{\pm}(\hbar, \bar{p})$ and therefore obtain the value of $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^{\pm})$ if we know sufficiently many coefficients of $E^{\pm}(\hbar, \bar{p})$. The required inverse transformation of power series is the plethystic logarithm:

**Corollary 2.14**  *We have*

$$e^{\pm}(\hbar, \bar{p}) = \sum_{k \geq 1} \frac{\mu(k)}{k} \log E^{\pm}(\hbar^k, \bar{p}_{[k]}),$$

*where $\mu(k)$ is the number-theoretic Möbius function.*

**Proof**  Let $g_1, g_2, \ldots \in \hat{\Lambda}((\hbar))$ such that $f_n = \sum_{k \geq 1} g_{nk}/(nk)$ converges uniformly in $n$. It follows from the definition of the Möbius function, $\sum_{d|n} \mu(d) = 0$ for all $n \geq 2$ and $\mu(1) = 1$, that

$$g_1 = \sum_{k \geq 1} \mu(k) \frac{f_k}{k}.$$

With $g_n = e^{\pm}(\hbar^n, \bar{p}_{[n]})$ and $f_n = \log E^{\pm}(\hbar^n, \bar{p}_{[n]})$, the statement now follows form Proposition 2.13.  □

## 2.6  Generating functions

In this section we will give the desired generating function for the polynomials $\hat{e}_{t,n}^{\pm}$. Together with the discussion in the last section, this generating function will give us an effective formula for the $\mathbb{S}_n$-equivariant Euler characteristic $e_{\mathbb{S}_n}(\mathcal{F}_{g,n}^{\pm})$.
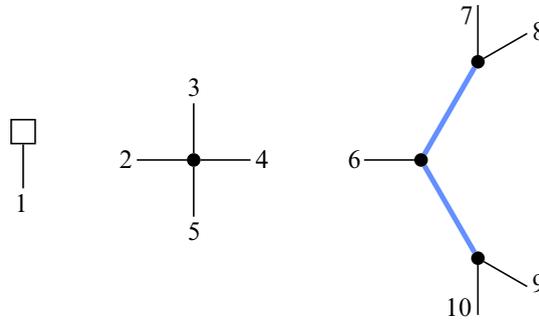
Figure 2: Extended forest with one special component.

The generating function is closely related to the one given in [8, Theorem 3.12] and we refer to the argument given there. As in [8], we define the following power series in $\overline{q} = q_1, q_2, \ldots$:

$$(3) \qquad V(\overline{q}) = q_1 + \tfrac{1}{2}q_1^2 - \tfrac{1}{2}q_2 - (1 + q_1) \sum_{k \geq 1} \frac{\mu(k)}{k} \log(1 + q_k).$$

See [8] for the first coefficients of $V(\overline{q})$. Moreover, we define two power series $F^+$ and $F^-$ both in two infinite sets of variables $\overline{q} = q_1, q_2, \ldots$, $\overline{p} = p_1, p_2, \ldots$ and a single variable $u$,

$$(4) \qquad F^{\pm}(u, \overline{q}, \overline{p}) = \exp\left( \sum_{k \geq 1} (\pm 1)^{k+1} u^{-2k} \frac{V((u \cdot \overline{q})_{[k]}) + u^k q_k p_k}{k} \right),$$

where $V((u \cdot \overline{q})_{[k]})$ means that we replace each variable $q_i$ in $V(\overline{q})$ with $u^{ki} q_{ki}$.

Recall that a forest is an acyclic graph. As graphs, also forests are required to be admissible that means they have no vertices of degree 0 and 2. Again, the univalent vertices are interpreted as the legs of the forest. An *extended* forest is a forests that is additionally allowed to have *special* univalent vertices that are always connected by a 1-cell to a leg. In contrast to graphs, we do not allow extended forests to have components that consist of two adjacent legs or two adjacent special vertices. See also [8, Section 3.7] for a discussion of extended forests. The legs of a forest are labeled by integers $\{1, \ldots, s\}$. The special vertices remain unlabeled. Figure 2 depicts an extended forest with three components; one of which is special. The special vertex is marked as a box. Legs are again drawn as labeled half-edges.

We write $k(\Phi)$, $s(\Phi)$ and $n(\Phi)$ for the total number of connected components, the number of legs and the number of special vertices of an extended forest $\Phi$. An automorphism $\gamma \in \mathrm{UAut}(\Phi)$ that is allowed to permute the leg-labels gives rise to a permutation of the edges of $\Phi$. We write $e_\gamma(\Phi)$ for the numbers of orbits of this permutation. An automorphism $\gamma \in \mathrm{UAut}(\Phi)$ also gives rise to a permutation $\pi_G(\gamma) \in \mathbb{S}_{s(\Phi)}$ of the legs, a permutation $\pi_G^\star(\gamma) \in \mathbb{S}_{n(\Phi)}$ of the special vertices and a permutation $\gamma_{H_0(\Phi)} \in \mathbb{S}_{h_0(\Phi)}$ of the connected components of $\Phi$.

**Proposition 2.15** *The generating functions $F^{\pm}$ count signed pairs of an extended forest $\Phi$ and an automorphism $\gamma$ of $\Phi$. Explicitly,*

$$F^{+}(u, \overline{q}, \overline{p}) = \sum_{(\Phi, \gamma)} (-1)^{e_{\gamma}(\Phi)} u^{s(\Phi) - 2k(\Phi)} p^{\pi_G^{\star}(\gamma)} \frac{q^{\pi_G(\gamma)}}{s(\Phi)!},$$

$$F^{-}(u, \overline{q}, \overline{p}) = \sum_{(\Phi, \gamma)} \operatorname{sign}(\gamma_{H_0(\Phi)})(-1)^{e_{\gamma}(\Phi)} u^{s(\Phi) - 2k(\Phi)} p^{\pi_G^{\star}(\gamma)} \frac{q^{\pi_G(\gamma)}}{s(\Phi)!},$$

*where we sum over all pairs of an extended forest $\Phi$ and all $\gamma \in \operatorname{UAut}(\Phi)$.*

**Proof** This statement is a slight generalization of [8, Propositions 3.10 and 5.5]. Here, we also allow forests to have special vertices and modify the generating function accordingly. The term $(\pm 1)^{k+1} u^{-k} q_k p_k / k$ in (4) accounts for these special components. Explicitly, it stands for $k$ special components that are cyclically permuted by the overall automorphism that is acting on the forested graph (see [8, Lemma 3.3] or [4]). Each component contributes two negative powers of $u$, because it adds one component, and one positive power of $u$ as it adds one leg. So, we mark a cycle of $k$ special components with $u^{-k}$. The different signs for the odd case are a consequence of [8, Lemma 5.3] and the fact that each special vertex counts as a new connected component of $\Phi$. $\qquad\square$

We will use the coefficient extraction operator notation. That means, we denote the coefficient of a power series $f(\overline{q})$ in front of $q^{\lambda}$ as $[q^{\lambda}]f(\overline{q})$.

**Corollary 2.16** *The coefficient $[u^{2t} q^{\mu} p^{\lambda}]F^{\pm}(u, \overline{q}, \overline{p})$ vanishes if $|\mu| = \sum_i \mu_i > 6t + 4|\lambda|$.*

**Proof** A *maximal $(r, n)$-forest* is an extended forest that consists of $r$ copies of a degree-3 vertex with three legs, and $n$ special components of a special vertex and one leg. All extended forests with $n$ special components can be obtained by first starting with a maximal $(r, n)$-forest, subsequently creating new edges by gluing together pairs of legs of different components and finally by contracting edges. The difference $s(\Phi) - 2k(\Phi)$ is left invariant by these gluing and contracting operations, but the number of legs $s(\Phi)$ decreases each time we glue together a pair of legs. It follows that maximal $(r, n)$-forests have the maximal number of legs for fixed $n$. Such a forest contributes a power $u^{3r - 2r + n - 2n} q^{\mu} p^{\lambda} = u^{r-n} q^{\mu} p^{\lambda}$ to the generating function, for some partitions $\mu, \lambda$ with $|\mu| = 3r + n$ and $|\lambda| = n$. Fixing $2t = r - n$ gives $|\mu| = 6t + 4n$.

Alternatively, the statement can also be verified by expanding $V$ and $F^{\pm}$ from (3)–(4). $\qquad\square$

We define two sets of numbers, $\eta_{\lambda}^{+}$ and $\eta_{\lambda}^{-}$, that are indexed by an integer partition, $\lambda \vdash s$. These numbers combine the definitions in Corollaries 3.5 and 5.8 of [8], where $\eta_{\lambda}^{+}$ is denoted as $\eta_{\lambda}$ and $\eta_{\lambda}^{-}$ as $\eta_{\lambda}^{\mathrm{odd}}$. The discussion around these corollaries also includes a detailed combinatorial interpretation of these numbers: they count (signed) fixed-point free involutions that commute with a given permutation of cycle type $\lambda$.

An alternative notation for an integer partitions is $\lambda = [1^{m_1} 2^{m_2} \cdots]$, where $m_k$ denotes the number of parts of size $k$ in $\lambda$. Let

$$\eta_\lambda^\pm = \prod_{k=1}^s \eta_{k,m_k}^\pm, \quad \text{where } \eta_{k,\ell}^\pm = \begin{cases} 0 & \text{if } k \text{ and } \ell \text{ are odd,} \\ k^{\ell/2}(\ell-1)!! & \text{if } k \text{ is odd and } \ell \text{ is even,} \\ \sum_{r=0}^{\lfloor \ell/2 \rfloor}(\pm 1)^{\ell k/2+r}\binom{\ell}{2r}k^r (2r-1)!! & \text{otherwise.} \end{cases}$$

With this we get an effective expression for the polynomials $\widehat{e}_{t,n}^\pm$ from (2):

**Theorem 2.17** *We have*

$$\widehat{e}_{t,n}^\pm = \sum_\mu \eta_\mu^\pm \sum_{\lambda \vdash n} p^\lambda [u^{2t} q^\mu p^\lambda] F^\pm(u, \overline{q}, \overline{p}),$$

*where we sum over all integer partitions $\mu$ and all partitions $\lambda$ of $n$. The terms in the sum are only nonzero for a finite number of such partitions. Moreover, all terms vanish if $t < -\frac{2}{3}n$.*

**Proof** By Corollary 2.16, the sum over $\mu$ has finite support and all terms vanish if $t < -\frac{2}{3}n$.

Matching all legs of an extended forest in pairs gives a graph with a marked forest and special univalent vertices. We get this graph by gluing together the legs as described by a matching. Gluing together two legs creates a new 1-cell between the vertices that are adjacent to the legs and forgets about the legs and the 1-cells that are incident to them. The special vertices are subsequently promoted to (unlabeled) legs of the resulting forested graph. All forested graphs with unlabeled legs can be obtained this way.

As an example consider the extended forest in Figure 2. We can match the legs in the pairs $(1,2)$, $(3,7)$, $(4,6)$, $(5,10)$ and $(8,9)$. Gluing together these legs and promoting the special vertex to a new leg recreates the forested graph in Figure 1, left, without the leg-label.

In general, if the extended forest that we start with has $k$ connected components and $s$ legs, then the graph that we obtain after the gluing has Euler characteristic $k - s/2$. By Proposition 2.15 and the definition of the polynomials $\widehat{e}_{t,n}$ in (2) we therefore extract the correct coefficient as all forests $\Phi$ with $s(\Phi) - 2k(\Phi) = 2t$ contribute to the coefficient of $u^{2t}$. See the proof of [8, Proposition 3.10] and its odd version [8, Proposition 5.5] for the discussions of automorphisms and sign factors which apply also in our generalized case. $\square$

Together Corollaries 2.5–2.14 and Theorem 2.17 give an effective algorithm for the computation of the polynomials $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ and $e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n})$. In summary:

**Theorem 2.18** *Fix $\chi > 1$. To compute $e(\mathcal{F}_{g,n}^\pm)$ for all $2g - 2 + n > 0$ and $g + 1 + n \le \chi$:*

(1) *Compute the coefficients of $V$ up to homogeneous order $6\chi$ in the $\overline{q}$-variables by expanding the power series defined in (3).*

(2) *Compute the coefficients of $F^\pm$ up to homogeneous order $\chi$ in $u^2$ and in the $\overline{p}$ variables by expanding the power series defined in (4).*

(3) *Compute the polynomials $\widehat{e}^{\pm}_{t,n}$ for all pairs $(t,n)$ with $n \geq 0$, $t+n \leq \chi$ and $t \geq -\frac{2}{3}n$ using Theorem 2.17.*

(4) *Compute the polynomials $e_{\mathbb{S}_n}(\mathcal{F}^{\pm}_{g,n})$ using the formula from Corollary 2.14.*

The formulas in Corollary 2.5 and Proposition 2.6 can be used to translate the result into the numbers $e(\mathcal{MG}_{g,n})$, $e^{\text{odd}}(\mathcal{MG}_{g,n})$, $e(\mathcal{MG}^{\mathbb{S}_n}_{g,n})$ and $e^{\text{odd}}(\mathcal{MG}^{\mathbb{S}_n}_{g,n})$. To compute the integers $\sum_k (-1)^k c^k_{g,\lambda}$ that are the alternating sums over the multiplicities of the irreducible representations in the respective cohomology of $\mathcal{MG}_{g,n}$, we have to write the polynomials $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ and $e^{\text{odd}}_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ in terms of Schur polynomials as in (1). We can do so using the Murnaghan–Nakayama rule.

## 2.7 Implementation of Theorem 2.18 in `FORM`

The most demanding computational step in Theorem 2.18 is the expansion of the power series $\boldsymbol{F}^{\pm}$ as defined in (4) in the formal variables $u$, $\overline{p}$ and $\overline{q}$. Conventional computer algebra struggles with such expansions; usually only a small number of terms are accessible. To be able to apply Theorem 2.18 at moderately large values of $\chi$, we use the `FORM` programming language. `FORM` is designed to deal with large analytic expressions that come up in high-energy physics.

A `FORM` program that implements Theorem 2.18 can be found in the online supplement in the file `eMGgn.frm`. It can be run with the command `form eMGgn.frm` after downloading and installing `FORM` from https://github.com/vermaseren/form.git. The syntax and details of the code are described in a `FORM` tutorial [28]. We used `FORM` version 5 *beta* for our computations.

The output of the program is a power series in which each coefficient is a symmetric function that describes the respective $\mathbb{S}_n$-equivariant Euler characteristic. These coefficients are given in the power sum basis of the ring of symmetric functions. To translate the output into Schur symmetric function via the Murnaghan–Nakayama rule, we used `Sage` [23].

**Remark 2.19** By employing the *Feynman transform* introduced by Getzler and Kapranov [16], modified versions of this program can effectively compute the Euler characteristics of *modular operads*. Furthermore, by combining Joyal's theory of species (see, eg, [4]) with findings from Borinsky's thesis [6], the program can be adapted to count various combinatorial objects, such as the number of isomorphism classes of admissible graphs of rank $n$ [18].

## 2.8 Large-$g$ asymptotics of the Euler characteristics of $\mathcal{MG}_{g,n}$

The *virtual* or *rational* Euler characteristic $\chi(G)$ is an invariant of a group $G$ that is often better behaved than the usual Euler characteristic.

Recall that the notation "$f(g) \sim h(g)$ for large $g$" means that $\lim_{g \to \infty} f(g)/h(g) = 1$. By the short exact sequence $1 \to F^n_g \to \Gamma_{g,n} \to \text{Out}(F_g) \to 1$ [13], we have $\chi(\Gamma_{g,n}) = (1-g)^n \chi(\text{Out}(F_g))$. The

numbers $\chi(\text{Out}(F_g))$ can be computed using [7, Proposition 8.5] and the asymptotic growth rate of $\chi(\text{Out}(F_g))$ is known explicitly by [7, Theorem A]. Using this together with the formula for $\chi(\Gamma_{g,n})$ and Stirling's approximation, we find that for fixed $n \geq 0$,

$$\chi(\Gamma_{g,n}) \sim \frac{(-1)^{n+1} g^n \left(\frac{g}{e}\right)^g}{(g \log g)^2} \quad \text{for large } g.$$

In [8], it was proven that $e(\mathcal{MG}_{g,0}) \sim e^{-1/4} \chi(\text{Out}(F_g))$ and $e^{\text{odd}}(\mathcal{MG}_{g,0}) \sim e^{1/4} \chi(\text{Out}(F_g))$ for large $g$. It would be interesting to make similar statements about $\mathcal{MG}_{g,n}$ for $n \geq 1$. Using our data on the Euler characteristics of $\mathcal{MG}_{g,n}$, we empirically verified the following conjecture which generalizes the known asymptotic behavior of $\mathcal{MG}_{g,0}$ to $\mathcal{MG}_{g,n}$ for all $n \geq 0$.

**Conjecture 2.20** *For fixed $n \geq 0$, we have, for large $g$,*

$$e(\mathcal{MG}_{g,n}) \sim e^{-1/4} \chi(\Gamma_{g,n}), \quad e^{\text{odd}}(\mathcal{MG}_{g,n}) \sim e^{1/4} \chi(\Gamma_{g,n}),$$

$$e(\mathcal{MG}_{g,n}^{\mathbb{S}_n}) \sim \frac{e^{-1/4} \chi(\Gamma_{g,n})}{n!}, \quad e^{\text{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n}) \sim \frac{e^{1/4} \chi(\Gamma_{g,n})}{n!}.$$

Proving this conjecture should be feasible by generalizing the analytic argument in [8, Section 4].

# 3 The large-$n$ $\mathbb{S}_n$-invariant cohomological stability of $\mathcal{MG}_{g,n}$

Our data in the Tables 2–5 exhibit an obvious pattern: for fixed $g$ the $\mathbb{S}_n$-*invariant* Euler characteristics $e(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ and $e^{\text{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ appear to be constant for all $n \geq g$. Unfortunately, this is not manifest from our formulas (ie from Theorem 2.18). To explain this pattern, we will prove the stabilization of the associated cohomologies.

**Theorem 3.1** *Fix $\mathbb{Q}_\rho \in \{\mathbb{Q}, \widetilde{\mathbb{Q}}\}$. The cohomology $H^\bullet(\mathcal{MG}_{g,n}; \mathbb{Q}_\rho)^{\mathbb{S}_n}$ stabilizes for $n \to \infty$: if $n \geq g \geq 2$, then there are isomorphisms $H^k(\mathcal{MG}_{g,n}; \mathbb{Q}_\rho)^{\mathbb{S}_n} \to H^k(\mathcal{MG}_{g,n+1}; \mathbb{Q}_\rho)^{\mathbb{S}_n}$ for all $k$.*

This statement is a refinement of the known *representational stability* [12] of $\mathcal{MG}_{g,n}$, which was shown to hold in [13]. In contrast to those previous results, Theorem 3.1 holds independently of the cohomological degree $k$. We will prove this theorem below in Section 3.2 using an argument that is based on the Lyndon–Hochschild–Serre spectral sequence and closely related to lines of thought in [13; 24].

Unfortunately, our proof of Theorem 3.1 gives neither a concrete description of the stable cohomologies $H^\bullet(\mathcal{MG}_{g,\infty}; \mathbb{Q})^{\mathbb{S}_\infty}$ and $H^\bullet(\mathcal{MG}_{g,\infty}; \widetilde{\mathbb{Q}})^{\mathbb{S}_\infty}$ nor gives explicit stabilization maps. A candidate for such a map is a generalization of the injection $H^\bullet(\text{Out}(F_g); \mathbb{Q}) \to H^\bullet(\text{Aut}(F_g); \mathbb{Q})$ from Theorem 1.4 of [13] to arbitrarily many legs. The first few terms of the Euler characteristics associated to these stable cohomologies are tabulated in Table 1. The values are remarkably small in comparison to the value of the Euler characteristic of $\text{Out}(F_g)$ for the respective rank (see the top and bottom row of Table 4 for

| $g$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $e(\mathcal{MG}_{g,\infty}^{\mathbb{S}\infty})$ | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | $-7$ | 71 | $-102$ | 295 |
| $e^{\mathrm{odd}}(\mathcal{MG}_{g,\infty}^{\mathbb{S}\infty})$ | 0 | 1 | $-2$ | 1 | $-2$ | 3 | $-11$ | 0 | $-18$ | $-7$ | $-71$ | $-102$ | $-295$ |

Table 1: Euler characteristics for $H^\bullet(\mathcal{MG}_{n,\infty};\mathbb{Q})^{\mathbb{S}\infty}$ and $H^\bullet(\mathcal{MG}_{n,\infty};\widetilde{\mathbb{Q}})^{\mathbb{S}\infty}$.

a direct comparison). Empirically, the Euler characteristics of the stable cohomologies appear to grow exponentially and not superexponentially. Our argument for Theorem 3.1 does not give any hint why these Euler characteristics are so small. It would also be interesting to prove Theorem 3.1 using graph cohomology methods as this kind of stabilization appears to be a distinguished feature of Lie- and forest graph cohomology. For instance, commutative graph cohomology with legs does not stabilize in the strong sense observed here [26].

## 3.1 Analogy to Artin's braid group

The observed large-$n$ stabilization carries similarities with the cohomology of Artin's *braid group* $B_n$ of equivalence classes of $n$-braids.

For a graph $G$ of rank $g$ with $n$ legs, let $\mathrm{U\Gamma}_{g,n}$ be the group of homotopy classes of self-homotopy equivalences of $G$ that only fix the legs as a set and not pointwise, ie $\mathrm{U\Gamma}_{g,n} = \pi_0(\mathrm{HE}(G,\partial G))$. By looking only at the action of $\mathrm{U\Gamma}_{g,n}$ on the leg-labels $\{1,\dots,n\}$, we get a surjective map to the symmetric group $\mathbb{S}_n$, ie

$$1 \to \Gamma_{g,n} \to \mathrm{U\Gamma}_{g,n} \to \mathbb{S}_n \to 1,$$

where the "pure" group, $\Gamma_{g,n}$, is the subgroup of $\mathrm{U\Gamma}_{g,n}$ that fixes the legs pointwise as defined in the introduction. Analogously, the braid group $B_n$ maps surjectively to $\mathbb{S}_n$. So,

$$1 \to P_n \to B_n \to \mathbb{S}_n \to 1,$$

where the kernel $P_n$ is the *pure braid group*. Whereas the braid groups $B_n$ exhibit *homological stability*, ie if $n \geq 3$, then $H_k(B_n;\mathbb{Q}) \cong H_k(B_{n+1};\mathbb{Q})$ for all $k$ [2], the pure braid groups $P_n$ only satisfy representational stability [12]. For instance, $H^\bullet(P_n;\mathbb{Q})$ is an exterior algebra on $\binom{n}{2}$ generators modulo a 3-term relation [2] and $H^\bullet(\Gamma_{1,n};\mathbb{Q})$ is the even-degree part of an exterior algebra on $n-1$ generators [13]. The numbers of generators of both algebras increase with $n$. The cohomology of $\mathrm{U\Gamma}_{g,n}$ is equal to the $\mathbb{S}_n$-invariant cohomology of $\mathcal{MG}_{g,n}$. So, analogously to the stabilization of $H^\bullet(B_n)$, Theorem 3.1 implies the large-$n$ cohomological stability of $\mathrm{U\Gamma}_{g,n}$.

## 3.2 The Lyndon–Hochschild–Serre spectral sequence

Our proof of Theorem 3.1 makes heavy use of tools from [13]. Following [13], we abbreviate $\mathrm{H} = H^1(F_g)$ and think of it as an $\mathrm{Out}(F_g)$-module. The action of $\mathrm{Out}(F_g)$ on H factors through the action of $\mathrm{GL}_g(\mathbb{Z})$ on H. The $q^{th}$ *exterior power* of H is denoted as $\bigwedge^q \mathrm{H}$. As $\mathrm{rk}\,\mathrm{H} = g$, the determinant representation is recovered by $\bigwedge^g \mathrm{H} = \widetilde{\mathbb{Q}}$. As before, we only work with modules over rational coefficients.

**Proposition 3.2** *Fix $g$, $n$ with $g \geq 2$, $n \geq 0$. There are two Lyndon–Hochschild–Serre spectral sequences with $E_2$ pages given by*

$$E_2^{p,q} = H^p\left(\mathrm{Out}(F_g); \bigwedge\nolimits^q \mathrm{H}\right), \quad \widetilde{E}_2^{p,q} = H^p\left(\mathrm{Out}(F_g); \widetilde{\mathbb{Q}} \otimes \bigwedge\nolimits^q \mathrm{H}\right)$$

*for $p, q \geq 0$ with $p \leq 2g - 3$ and $q \leq \min(g, n)$, and $E_2^{p,q} = \widetilde{E}_2^{p,q} = 0$ for all other values of $p, q$. Both spectral sequences converge: $E_2^{p,q} \Rightarrow H^{p+q}(\Gamma_{g,n}; \mathbb{Q})^{\mathbb{S}_n}$ and $\widetilde{E}_2^{p,q} \Rightarrow H^{p+q}(\Gamma_{g,n}; \widetilde{\mathbb{Q}})^{\mathbb{S}_n}$.*

The argument works along the lines of Section 3.2 of [13] followed by an application of Schur–Weyl duality and the projection to the $\mathbb{S}_n$-invariant cohomology. A similar argument can be found in [24, Lemma 4.1] where it is used to prove representational stability of $H^\bullet(\Gamma_{g,n}; \mathbb{Q})$.

**Proof** For $g \geq 2$ and $n \geq 0$, the Lyndon–Hochschild–Serre spectral sequence associated to the group extension $1 \to F_g^n \to \Gamma_{g,n} \to \mathrm{Out}(F_g) \to 1$ is a first-quadrant spectral sequence with the second page given by $E_2^{p,q} = H^p(\mathrm{Out}(F_g); H^q(F_g^n; \mathbb{Q})) \Rightarrow H^{p+q}(\Gamma_{g,n}; \mathbb{Q})$ (see [13, Section 3.2]). The virtual cohomological dimension of $\mathrm{Out}(F_n)$ is $2g - 3$ [15], so if $p > 2g - 3$, then $E_2^{p,q} = \widetilde{E}_2^{p,q} = 0$. By the Künneth formula, the $\mathbb{S}_n$-module $H^q(F_g^n; \mathbb{Q})$ vanishes if $q > n$ and for $0 \leq q \leq n$ it is obtained by inducing the $\mathbb{S}_q \times \mathbb{S}_{n-q}$-module $\mathrm{H}^{\wedge q} \otimes V_{(n-q)}$ to $\mathbb{S}_n$, (see [13, Lemma 3.4] for details),

$$H^q(F_g^n; \mathbb{Q}) = \mathrm{Ind}_{\mathbb{S}_q \times \mathbb{S}_{n-q}}^{\mathbb{S}_n}(\mathrm{H}^{\wedge q} \otimes V_{(n-q)}),$$

where $V_{(n-q)}$ is the trivial representation of $\mathbb{S}_{n-q}$ and $\mathrm{H}^{\wedge q}$ is $\mathrm{H}^{\otimes q} \otimes V_{(1^q)}$ with $V_{(1^q)}$ the alternating representation of $\mathbb{S}_q$ and $\mathbb{S}_q$ acts on $\mathrm{H}^{\otimes q}$ by permuting the entries of the tensor product.

We can set up a similar spectral sequence with coefficients twisted by $\widetilde{\mathbb{Q}}$ with second page

$$\widetilde{E}_2^{p,q} = H^p(\mathrm{Out}(F_g); H^q(F_g^n; \widetilde{\mathbb{Q}})) \Rightarrow H^{p+q}(\Gamma_{g,n}; \widetilde{\mathbb{Q}}).$$

Note that $F_g^n$ acts trivially on $\widetilde{\mathbb{Q}}$. So, applying the Künneth formula to expand $H^q(F_g^n; \widetilde{\mathbb{Q}})$ as an $\mathbb{S}_n$-module gives

$$H^q(F_g^n; \widetilde{\mathbb{Q}}) = \widetilde{\mathbb{Q}} \otimes \mathrm{Ind}_{\mathbb{S}_q \times \mathbb{S}_{n-q}}^{\mathbb{S}_n}(\mathrm{H}^{\wedge q} \otimes V_{(n-q)}),$$

which only differs by the $GL_g(\mathbb{Z})$-action on $\widetilde{\mathbb{Q}}$. Schur–Weyl duality gives the irreducible decomposition of $\mathrm{H}^{\wedge q}$ as a module over $GL(\mathrm{H}) \times \mathbb{S}_q$ (see, eg, [13, Section 3.1]):

$$\mathrm{H}^{\wedge q} = \bigoplus_{\lambda \vdash q} W_\lambda \otimes V_{\lambda'},$$

where we sum over all partitions $\lambda$ of $q$ with at most $g$ rows and $\lambda'$ is the transposed partition. Here, $W_\lambda$ and $V_\lambda$ are the irreducible $GL(\mathrm{H})$- and $\mathbb{S}_q$-representations associated to $\lambda$.

As $\mathbb{S}_n$ only acts on the coefficients in $E_2^{p,q}$, we get

$$E_2^{p,q} = H^p(\mathrm{Out}(F_g); \mathrm{Ind}_{\mathbb{S}_q \times \mathbb{S}_{n-q}}^{\mathbb{S}_n}(\mathrm{H}^{\wedge q} \otimes V_{(n-q)}))$$

$$= \bigoplus_{\lambda \vdash q} H^p(\mathrm{Out}(F_g); W_\lambda) \otimes \mathrm{Ind}_{\mathbb{S}_q \times \mathbb{S}_{n-q}}^{\mathbb{S}_n}(V_{\lambda'} \otimes V_{(n-q)}).$$

If we project to the trivial $\mathbb{S}_n$-representation, only the partition $\lambda = (1^q)$ contributes. Hence,

$$(E_2^{p,q})^{\mathbb{S}_n} = H^p(\text{Out}(F_g); W_{(1^q)}).$$

The representation $W_{(1^q)}$ is the $q^{\text{th}}$ exterior power of the defining representation H. It vanishes if $q > \text{rk H} = g$. The argument works analogously for $\tilde{E}_2^{p,q}$. $\qquad\square$

**Proof of Theorem 3.1** For $n \geq g \geq 2$, the $E_2$ pages of the spectral sequences in Proposition 3.2 are independent of $n$. Hence, by the spectral sequence comparison theorem, also $H^k(\Gamma_{g,n}, \mathbb{Q})^{\mathbb{S}_n}$ and $H^k(\Gamma_{g,n}, \widetilde{\mathbb{Q}})^{\mathbb{S}_n}$ must be independent of $n$. $\qquad\square$

In the stable regime, the Euler characteristics $e(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ and $e^{\text{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ are equal up to a sign factor (see Table 1). This can be explained as another consequence of Proposition 3.2, and the following two technical lemmas.

For a graph $G$, the first cohomology $H^1(G) = H^1(G; \mathbb{Q})$ is a representation of its automorphism group $\text{Aut}(G)$. Let $H^1(G)^*$ be the corresponding dual representation.

**Lemma 3.3** $H^1(G) \cong H^1(G)^*$ *as $\text{Aut}(G)$-representations.*

**Proof** Each graph $G$ comes with an $\text{Aut}(G)$-invariant symmetric positive definite bilinear form $H_1(G) \otimes H_1(G) \to \mathbb{Q}$, *the graph Laplacian* (see, eg, [3, 3.1.1]). Dually, we have an $\text{Aut}(G)$-invariant pairing on $H^1(G)$ giving the stated isomorphism of $\text{Aut}(G)$-representations. $\qquad\square$

Let H* be the $\text{Out}(F_g)$-representation dual to H and $e(H^\bullet(G; M)) = \sum_p (-1)^p \dim(H^p(G; M))$.

**Lemma 3.4** *For $g \geq 2$ and $q \geq 0$, we have $e\big(H^\bullet\big(\text{Out}(F_g); \bigwedge^q \text{H}\big)\big) = e\big(H^\bullet\big(\text{Out}(F_g); \bigwedge^q \text{H}^*\big)\big)$.*

**Proof** Using standard arguments that relate Culler–Vogtmann outer space to the rational group cohomology of $\text{Out}(F_n)$, we are looking for a cochain model for the stated cohomologies. Let $K_g$ be the *spine of outer space*, ie $K_g$ is a cubical complex in which each $k$-cube corresponds to an isomorphism class of a forested graph $(G, \Phi, \mu)$ which is equipped with a *marking*, an isomorphism $\mu \colon \pi_1(G) \to F_g$. Each cube is oriented by ordering the edges in $\Phi$. The group $\text{Out}(F_g)$ acts on $K_g$ by composition with the marking. This action is not free, but it has finite stabilizers. Most importantly, $K_g$ is contractible [15]. Let $M$ be $\bigwedge^q \text{H}$ or $\bigwedge^q \text{H}^*$, then by standard arguments $H^\bullet(\text{Out}(F_g); M) = H^\bullet(\text{Hom}_{\text{Out}(F_g)}(C_\bullet(K_g), M))$ (see, eg, [9, Chapter 7]). The action of $\text{Out}(F_g)$ on the tuples $(G, \Phi, \mu)$ fixes the isomorphism class of the forested graph and acts transitively on the marking. The stabilizer of this action is $\text{Aut}(G, \Phi)$. Hence,

$$\text{Hom}_{\text{Out}(F_g)}(C_\bullet(K_g), M) = \text{Hom}(C_\bullet(K_g), M)^{\text{Out}(F_g)} = \bigoplus_{[G,\Phi]} (\text{sign}_{(G,\Phi)} \otimes M)^{\text{Aut}(G,\Phi)},$$

where we take the sum over all isomorphism classes of forested graphs of rank $g$ or equivalently over all $\text{Out}(F_g)$-orbits of isomorphism classes of marked forested graphs of rank $g$. Here, $\text{sign}_{(G,\Phi)}$ is the

orientation module, the one-dimensional representation of $\mathrm{Aut}(G, \Phi)$ that is given by $\alpha \mapsto \mathrm{sign}(\alpha_\Phi)$ with $\alpha_\Phi$ the permutation that $\alpha \in \mathrm{Aut}(G, \Phi)$ induces on the forest edges. The group $\mathrm{Aut}(G, \Phi)$ acts on $M$ via the map $\mathrm{Aut}(G, \Phi) \to \mathrm{Out}(F_g)$ that any chosen marking representative $\pi_1(G) \to F_g$ induces. As a $\mathrm{Aut}(G, \Phi)$-representation, H is hence equal to $H^1(G)$. Therefore, due to Lemma 3.3, we may replace the right-most $M$ in the displayed formula above with either $\bigwedge^q H^1(G)$, or $\bigwedge^q (H^1(G)^*) \cong \left( \bigwedge^q H^1(G) \right)^*$. So, in the stated equality, both sides can be computed using cochain complexes with the same generators in each degree. The Euler characteristic only depends on the dimensions of the cochain spaces and not on the differentials.                                                                                    □

**Proposition 3.5**   *If $n \geq g \geq 2$, then $e(\mathcal{MG}_{g,n}^{\mathbb{S}_n}) = (-1)^g e^{\mathrm{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$.*

**Proof**   There is an isomorphism of $\mathrm{GL}_g(\mathbb{Z})$-representations: $\widetilde{\mathbb{Q}} \otimes \bigwedge^q H = \bigwedge^g H \otimes \bigwedge^q H \cong \bigwedge^{g-q} H^*$. This can be seen by computing the characters of both $\bigwedge^g H \otimes \bigwedge^q H$ and $\bigwedge^{g-q} H^*$ and by observing that they differ by a multiple of the character of $\left( \bigwedge^g \right)^{\otimes 2}$, which is the same as the trivial representation of $\mathrm{GL}_g(\mathbb{Z})$, because $\det(h)^2 = 1$ for each $h \in \mathrm{GL}_g(\mathbb{Z})$ (see, eg, [25, Chapter 7.A.2]).

Hence, for $n \geq g \geq 2$, we have, for the $E_2$ pages of the spectral sequences in Proposition 3.2,

$$\widetilde{E}_2^{p,q} = H^p\left( \mathrm{Out}(F_g); \widetilde{\mathbb{Q}} \otimes \bigwedge^q H \right) \cong H^p\left( \mathrm{Out}(F_g); \bigwedge^{g-q} H^* \right).$$

Therefore, it follows from Lemma 3.4 that $\sum_p (-1)^p \dim \widetilde{E}_2^{p,q} = \sum_p (-1)^p \dim E_2^{p,g-q}$. So,

$$e(\mathcal{MG}_{g,n}^{\mathbb{S}_n}) = \sum_{p,q} (-1)^{p+q} \dim E_2^{p,q} = \sum_{p,q} (-1)^{p+g-q} \dim E_2^{p,g-q} = (-1)^g e^{\mathrm{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n}). \qquad \square$$

Apart from the (up-to-sign) equality of the Euler characteristics, our argument does not tell us how even and odd $\mathbb{S}_n$-invariant stable $n \to \infty$ cohomology of $\mathcal{MG}_{g,n}$ are related.

Moreover, the $E_2$ page of the spectral sequences in Proposition 3.2 contains the whole (twisted) cohomology of $\mathrm{Out}(F_g)$ (eg $E_2^{p,0} = H^p(\mathrm{Out}(F_g); \mathbb{Q})$). So, we also do not explain why the Euler characteristics of the large-$n$ stable $\mathbb{S}_n$-invariant cohomologies of $\mathcal{MG}_{g,n}$ are so small in comparison to the Euler characteristic of $\mathrm{Out}(F_g)$.

# Appendix   Tables of Euler characteristics of $\mathcal{MG}_{g,n}$

Larger versions of Tables 2–7 are included in the online supplement: values of the Euler characteristics $e(\mathcal{MG}_{g,n})$, $e^{\mathrm{odd}}(\mathcal{MG}_{g,n})$, $e(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ and $e^{\mathrm{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$ for $g + n \leq 60$ can be found in `eMGgn.tsv`, `eMGgn-odd.tsv`, `eMGgn-modSn.tsv` and `eMGgn-modSn-odd.tsv`. These larger tables are readable as plain text or with standard spreadsheet software. The files `eOutFn.txt` and `eOutFn-odd.txt` list the values $e(\mathrm{Out}(F_g))$ and $e^{\mathrm{odd}}(\mathrm{Out}(F_g))$ for all $g \leq 100$. The full $\mathbb{S}_n$-equivariant Euler characteristic is listed in the files `eMGgn-equiv.tsv` and `eMGgn-equiv-odd.tsv` that contain tables of the polynomials $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ and $e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n})$ for $g + n \leq 30$, respectively.

| $n$ | $g = 1$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 0 | | 1 | 1 | 2 | 1 |
| 1 | 1 | 1 | 1 | 2 | 0 |
| 2 | 1 | 1 | 2 | 3 | −2 |
| 3 | 2 | 1 | 6 | 6 | 0 |
| 4 | 4 | 0 | 18 | 4 | −24 |
| 5 | 8 | −4 | 61 | −3 | 64 |
| 6 | 16 | −19 | 202 | −158 | −69 |
| 7 | 32 | −69 | 701 | −831 | 2905 |
| 8 | 64 | −230 | 2438 | −5135 | 9917 |
| 9 | 128 | −734 | 8721 | −25446 | 112518 |
| 10 | 256 | −2289 | 31602 | −134879 | 552339 |
| 11 | 512 | −7039 | 116821 | −670008 | 4149475 |
| 12 | 1024 | −21460 | 437758 | −3414254 | 22844193 |
| 13 | 2048 | −65064 | 1663481 | −17022549 | 149941792 |
| 14 | 4096 | −196559 | 6388202 | −85672220 | 864112247 |
| 15 | 8192 | −592409 | 24759741 | −427885725 | 5376583485 |
| 16 | 16384 | −1782690 | 96647478 | −2144390153 | 31618003029 |

| $n$ | $g = 6$ | 7 | 8 | 9 |
|---|---|---|---|---|
| 0 | 2 | 1 | 1 | −21 |
| 1 | 3 | 3 | 31 | 154 |
| 2 | 1 | −28 | −153 | −1486 |
| 3 | 36 | 152 | 1423 | 12072 |
| 4 | −86 | −1062 | −9474 | −103392 |
| 5 | 675 | 6421 | 72249 | 845821 |
| 6 | −3453 | −38028 | −506827 | −6971380 |
| 7 | 11182 | 253892 | 3616144 | 56742775 |
| 8 | −124467 | −1306559 | −25952916 | −459328520 |
| 9 | 47564 | 9957185 | 178180002 | 3726950202 |
| 10 | −4683027 | −42459898 | −1308692710 | −29852809180 |
| 11 | −9296152 | 399601667 | 8669878028 | 242171554559 |
| 12 | −188955568 | −1233422911 | −65810827609 | −1924630979085 |
| 13 | −725537667 | 16798999974 | 416483532392 | 15652884150733 |
| 14 | −8133442381 | −25768675818 | −3327304711052 | −123449090799389 |
| 15 | −41758325066 | 753773677302 | 19660027898985 | 1009591434254489 |
| 16 | −367416474589 | 247507159657 | −170405333570573 | −7887831186821342 |

Table 2: Euler characteristic $e(\mathcal{MG}_{g,n})$.

| $n$ | $g=$ 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 0 | | 0 | 0 | −1 | 0 |
| 1 | 0 | 0 | −1 | −1 | −1 |
| 2 | −1 | 0 | −3 | −3 | −5 |
| 3 | −2 | 0 | −8 | −7 | −9 |
| 4 | −4 | −1 | −24 | −27 | −66 |
| 5 | −8 | −5 | −71 | −81 | −137 |
| 6 | −16 | −20 | −228 | −364 | −1185 |
| 7 | −32 | −70 | −743 | −1394 | −3536 |
| 8 | −64 | −231 | −2544 | −6716 | −28509 |
| 9 | −128 | −735 | −8891 | −29974 | −118323 |
| 10 | −256 | −2290 | −32028 | −148035 | −837525 |
| 11 | −512 | −7040 | −117503 | −708621 | −4206038 |
| 12 | −1024 | −21461 | −439464 | −3528385 | −27300597 |
| 13 | −2048 | −65065 | −1666211 | −17361527 | −150529289 |
| 14 | −4096 | −196560 | −6395028 | −86682326 | −934471725 |
| 15 | −8192 | −592410 | −24770663 | −430902388 | −5383163340 |
| 16 | −16384 | −1782691 | −96674784 | −2153412834 | −32735332605 |

| $n$ | $g=$ 6 | 7 | 8 | 9 |
|---|---|---|---|---|
| 0 | −1 | −2 | −8 | −38 |
| 1 | 1 | 3 | 34 | 278 |
| 2 | −7 | −45 | −273 | −2143 |
| 3 | 20 | 172 | 1688 | 16279 |
| 4 | −184 | −1365 | −12037 | −127665 |
| 5 | 340 | 6153 | 80002 | 995425 |
| 6 | −5356 | −47862 | −568014 | −7861828 |
| 7 | 3081 | 210751 | 3814347 | 61973273 |
| 8 | −166300 | −1760727 | −27489461 | −492911760 |
| 9 | −144896 | 6985727 | 182953976 | 3900012883 |
| 10 | −5638835 | −67361954 | −1348453174 | −31209047062 |
| 11 | −13846788 | 214832720 | 8779760089 | 246950368646 |
| 12 | −211328036 | −2704794269 | −66861595436 | −1987593802313 |
| 13 | −834038152 | 5420792406 | 418887979427 | 15685889711601 |
| 14 | −8667253394 | −115615170716 | −3355504362873 | −127079507192857 |
| 15 | −44375995137 | 48464678629 | 19709309010324 | 997347475403633 |
| 16 | −380341928592 | −5327400148291 | −171167953029982 | −8150861890475894 |

Table 3: Euler characteristics $e^{\mathrm{odd}}(\mathcal{MG}_{g,n})$.

| $n$ | $g =$ 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | 1 | 1 | 2 | 1 | 2 | 1 | 1 | −21 | −124 | −1202 | −10738 |
| 1 | 1 | 1 | 1 | 2 | 0 | 3 | 3 | 31 | 154 | 1405 | 12409 | 128198 |
| 2 | 1 | 1 | 2 | 2 | 0 | 1 | −8 | −71 | −645 | −5916 | −60661 | −680524 |
| 3 | 1 | 1 | 2 | 2 | 1 | 4 | 22 | 148 | 1432 | 14933 | 173615 | 2170285 |
| 4 | 1 | 1 | 2 | 1 | 2 | 2 | −6 | −158 | −1911 | −24310 | −324161 | −4610023 |
| 5 | 1 | 1 | 2 | 1 | 2 | 4 | 14 | 114 | 1677 | 26129 | 415109 | 6837661 |
| 6 | 1 | 1 | 2 | 1 | 2 | 3 | 9 | −34 | −900 | −18622 | −368380 | −7231581 |
| 7 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 8 | 296 | 8548 | 223615 | 5456280 |
| 8 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | −20 | −2292 | −88814 | −2878337 |
| 9 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | 268 | 20992 | 1010696 |
| 10 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | −2154 | −212879 |
| 11 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | 71 | 20256 |
| 12 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | 71 | −102 |
| 13 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | 71 | −102 |
| 14 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | 71 | −102 |
| 15 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | 71 | −102 |
| 16 | 1 | 1 | 2 | 1 | 2 | 3 | 11 | 0 | 18 | −7 | 71 | −102 |

| $n$ | $g =$ 13 | 14 | 15 | 16 |
|---|---|---|---|---|
| 0 | −112901 | −1271148 | −15668391 | −208214777 |
| 1 | 1428208 | 17431842 | 229796854 | 3260731764 |
| 2 | −8319674 | −110000218 | −1564363190 | −23810497027 |
| 3 | 29302324 | 423870178 | 6547649971 | 107566687178 |
| 4 | −69547908 | −1112576568 | −18825687993 | −336214369334 |
| 5 | 117254253 | 2100407688 | 39335548893 | 770173833141 |
| 6 | −143971474 | −2934457032 | −61596805980 | −1335762223221 |
| 7 | 129717602 | 3072060291 | 73437715819 | 1786143030025 |
| 8 | −85129652 | −2410044793 | −66988401529 | −1856600124025 |
| 9 | 39693125 | 1399384996 | 46526543233 | 1500185772169 |
| 10 | −12481703 | −584609581 | −24221345542 | −934645878596 |
| 11 | 2377198 | 166410659 | 9165254242 | 440926365563 |
| 12 | −207026 | −28931603 | −2383055405 | −152471494842 |
| 13 | 295 | 2319534 | 381124388 | 36485839046 |
| 14 | 295 | −602 | −28285686 | −5402710802 |
| 15 | 295 | −602 | 1730 | 373201519 |
| 16 | 295 | −602 | 1730 | −3616 |

Table 4: Euler characteristics $e(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$.

| $n$ | $g=$ 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | 0 | 0 | −1 | 0 | −1 | −2 | −8 | −38 | −275 | −2225 | −20358 |
| 1 | 0 | 0 | −1 | −1 | −1 | 1 | 3 | 34 | 278 | 2285 | 20921 | 212777 |
| 2 | 0 | 1 | −1 | −1 | −2 | −1 | −17 | −114 | −918 | −8555 | −88885 | −1010798 |
| 3 | 0 | 1 | −2 | −1 | −2 | 2 | 11 | 158 | 1659 | 18615 | 223544 | 2878235 |
| 4 | 0 | 1 | −2 | 1 | −1 | 0 | −19 | −148 | −1929 | −26136 | −368451 | −5456382 |
| 5 | 0 | 1 | −2 | 1 | −2 | 1 | −8 | 71 | 1414 | 24303 | 415038 | 7231479 |
| 6 | 0 | 1 | −2 | 1 | −2 | 3 | −10 | −31 | −663 | −14940 | −324232 | −6837763 |
| 7 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | −1 | 136 | 5909 | 173544 | 4609921 |
| 8 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −39 | −1412 | −60732 | −2170387 |
| 9 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | 117 | 12338 | 680422 |
| 10 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −1273 | −128300 |
| 11 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −71 | 10636 |
| 12 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −71 | −102 |
| 13 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −71 | −102 |
| 14 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −71 | −102 |
| 15 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −71 | −102 |
| 16 | 0 | 1 | −2 | 1 | −2 | 3 | −11 | 0 | −18 | −7 | −71 | −102 |

| $n$ | $g=$ 13 | 14 | 15 | 16 |
|---|---|---|---|---|
| 0 | −207321 | −2320136 | −28287416 | −373205135 |
| 1 | 2376903 | 28931001 | 381122658 | 5402707186 |
| 2 | −12481998 | −166411261 | −2383057135 | −36485842662 |
| 3 | 39692830 | 584608979 | 9165252512 | 152471491226 |
| 4 | −85129947 | −1399385598 | −24221347272 | −440926369179 |
| 5 | 129717307 | 2410044191 | 46526541503 | 934645874980 |
| 6 | −143971769 | −3072060893 | −66988403259 | −1500185775785 |
| 7 | 117253958 | 2934456430 | 73437714089 | 1856600120409 |
| 8 | −69548203 | −2100408290 | −61596807710 | −1786143033641 |
| 9 | 29302029 | 1112575966 | 39335547163 | 1335762219605 |
| 10 | −8319969 | −423870780 | −18825689723 | −770173836757 |
| 11 | 1427913 | 109999616 | 6547648241 | 336214365718 |
| 12 | −113196 | −17432444 | −1564364920 | −107566690794 |
| 13 | −295 | 1270546 | 229795124 | 23810493411 |
| 14 | −295 | −602 | −15670121 | −3260735380 |
| 15 | −295 | −602 | −1730 | 208211161 |
| 16 | −295 | −602 | −1730 | −3616 |

Table 5: Euler characteristics $e^{\mathrm{odd}}(\mathcal{MG}_{g,n}^{\mathbb{S}_n})$.

| $g$ | $n=$ 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | | | | $s_3$ |
| 1 | | $s_1$ | $s_2$ | $s_{1,1,1}+s_3$ |
| 2 | 1 | $s_1$ | $s_2$ | $s_3$ |
| 3 | 1 | $s_1$ | $2s_2$ | $2s_{2,1}+2s_3$ |
| 4 | 2 | $2s_1$ | $s_{1,1}+2s_2$ | $2s_{1,1,1}+s_{2,1}+2s_3$ |
| 5 | 1 | 0 | $-2s_{1,1}$ | $-s_{1,1,1}+s_3$ |
| 6 | 2 | $3s_1$ | $s_2$ | $10s_{1,1,1}+11s_{2,1}+4s_3$ |

| $g$ | $n=$ 4 | 5 |
|---|---|---|
| 0 | $s_4$ | $s_5$ |
| 1 | $s_{2,1,1}+s_4$ | $s_{1,1,1,1}+s_{3,1,1}+s_5$ |
| 2 | $-s_{2,1,1}+s_{2,2}+s_4$ | $-s_{2,1,1}-s_{3,1,1}+s_{3,2}+s_5$ |
| 3 | $s_{2,1,1}+2s_{2,2}+3s_{3,1}+2s_4$ | $s_{2,1,1,1}+2s_{2,2,1}+3s_{3,1,1}+3s_{3,2}+3s_{4,1}+2s_5$ |
| 4 | $s_{1,1,1,1}+s_{2,2}+s_4$ | $2s_{1,1,1,1,1}+s_{2,1,1,1}-s_{3,1,1}-s_{4,1}+s_5$ |
| 5 | $-5s_{1,1,1,1}-4s_{2,1,1}-3s_{2,2}-s_{3,1}+2s_4$ | $4s_{2,1,1,1}+s_{2,2,1}+4s_{3,1,1}+s_{3,2}+3s_{4,1}+2s_5$ |
| 6 | $-5s_{1,1,1,1}-15s_{2,1,1}-4s_{2,2}-10s_{3,1}+2s_4$ | $34s_{1,1,1,1,1}+51s_{2,1,1,1}+41s_{2,2,1}+22s_{3,1,1}+16s_{3,2}+4s_{4,1}+4s_5$ |

Table 6: $\mathbb{S}_n$-equivariant Euler characteristic $e_{\mathbb{S}_n}(\mathcal{MG}_{g,n})$ in the Schur basis of $\Lambda_n$.

| $g$ | $n=$ 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | | | | $s_3$ |
| 1 | | 0 | $-s_{1,1}$ | $-s_{2,1}$ |
| 2 | 0 | 0 | $-s_{1,1}+s_2$ | $-s_{1,1,1}+s_3$ |
| 3 | 0 | $-s_1$ | $-2s_{1,1}-s_2$ | $-2s_{1,1,1}-2s_{2,1}-2s_3$ |
| 4 | $-1$ | $-s_1$ | $-2s_{1,1}-s_2$ | $-2s_{1,1,1}-2s_{2,1}-s_3$ |
| 5 | 0 | $-s_1$ | $-3s_{1,1}-2s_2$ | $-s_{1,1,1}-3s_{2,1}-2s_3$ |
| 6 | $-1$ | $s_1$ | $-6s_{1,1}-s_2$ | $4s_{1,1,1}+7s_{2,1}+2s_3$ |

| $g$ | $n=$ 4 | 5 |
|---|---|---|
| 0 | $s_4$ | $s_5$ |
| 1 | $-s_{1,1,1}-s_{3,1}$ | $-s_{2,1,1}-s_{4,1}$ |
| 2 | $-s_{1,1,1}-s_{2,1,1}+s_{2,2}+s_4$ | $-s_{1,1,1,1}-s_{2,1,1}-s_{3,1,1}+s_{3,2}+s_5$ |
| 3 | $-2s_{1,1,1,1}-3s_{2,1,1}-s_{2,2}-3s_{3,1}-2s_4$ | $-2s_{1,1,1,1,1}-3s_{2,1,1,1}-2s_{2,2,1}-3s_{3,1,1}-3s_{3,2}-3s_{4,1}-2s_5$ |
| 4 | $-3s_{1,1,1,1}-4s_{2,1,1}-2s_{2,2}-3s_{3,1}+s_4$ | $-3s_{1,1,1,1,1}-4s_{2,1,1,1}-4s_{2,2,1}-4s_{3,1,1}-3s_{3,2}-s_{4,1}+s_5$ |
| 5 | $-7s_{1,1,1,1}-9s_{2,1,1}-5s_{2,2}-7s_{3,1}-s_4$ | $-2s_{1,1,1,1,1}-6s_{2,1,1,1}-5s_{2,2,1}-8s_{3,1,1}-4s_{3,2}-4s_{4,1}-2s_5$ |
| 6 | $-22s_{1,1,1,1}-32s_{2,1,1}-12s_{2,2}-14s_{3,1}$ | $19s_{1,1,1,1,1}+27s_{2,1,1,1}+26s_{2,2,1}+6s_{3,1,1}+10s_{3,2}-s_{4,1}+s_5$ |

Table 7: $\mathbb{S}_n$-equivariant Euler characteristic $e_{\mathbb{S}_n}^{\mathrm{odd}}(\mathcal{MG}_{g,n})$ in the Schur basis of $\Lambda_n$.

# References

[1] **D Abramovich**, **L Caporaso**, **S Payne**, *The tropicalization of the moduli space of curves*, Ann. Sci. Éc. Norm. Supér. 48 (2015) 765–809 MR

[2] **V I Arnold**, *The cohomology ring of the group of dyed braids*, Mat. Zametki 5 (1969) 227–231 MR In Russian; translated in Math. Notes 5 (1969) 138–140

[3] **O Baker**, *The Jacobian map on outer space*, PhD thesis, Cornell University (2011) Available at `https://ecommons.cornell.edu/handle/1813/30769`

[4] **F Bergeron**, **G Labelle**, **P Leroux**, *Combinatorial species and tree-like structures*, Encyclopedia of Mathematics and its Applications 67, Cambridge Univ. Press (1998) MR

[5] **M Berghoff**, *Feynman amplitudes on moduli spaces of graphs*, Ann. Inst. Henri Poincaré D 7 (2020) 203–232 MR

[6] **M Borinsky**, *Graphs in perturbation theory*: *algebraic structure and asymptotics*, Springer (2018) MR

[7] **M Borinsky**, **K Vogtmann**, *The Euler characteristic of* $\mathrm{Out}(F_n)$, Comment. Math. Helv. 95 (2020) 703–748 MR

[8] **M Borinsky**, **K Vogtmann**, *The Euler characteristic of the moduli space of graphs*, Adv. Math. 432 (2023) art. id. 109290 MR

[9] **K S Brown**, *Cohomology of groups*, Graduate Texts in Mathematics 87, Springer (1982) MR

[10] **M Chan**, **C Faber**, **S Galatius**, **S Payne**, *The* $S_n$*-equivariant top weight Euler characteristic of* $\mathcal{M}_{g,n}$, Amer. J. Math. 145 (2023) 1549–1585 MR

[11] **M Chan**, **S Galatius**, **S Payne**, *Tropical curves, graph complexes, and top weight cohomology of* $\mathcal{M}_g$, J. Amer. Math. Soc. 34 (2021) 565–594 MR

[12] **T Church**, **B Farb**, *Representation theory and homological stability*, Adv. Math. 245 (2013) 250–314 MR

[13] **J Conant**, **A Hatcher**, **M Kassabov**, **K Vogtmann**, *Assembling homology classes in automorphism groups of free groups*, Comment. Math. Helv. 91 (2016) 751–806 MR

[14] **J Conant**, **K Vogtmann**, *On a theorem of Kontsevich*, Algebr. Geom. Topol. 3 (2003) 1167–1224 MR

[15] **M Culler**, **K Vogtmann**, *Moduli of graphs and automorphisms of free groups*, Invent. Math. 84 (1986) 91–119 MR

[16] **E Getzler**, **M M Kapranov**, *Modular operads*, Compositio Math. 110 (1998) 65–126 MR

[17] **E Gorsky**, *The equivariant Euler characteristic of moduli spaces of curves*, Adv. Math. 250 (2014) 588–595 MR

[18] **A Howroyd**, **W P Orrick**, *Number of unlabeled connected multigraphs with circuit rank n and degree* $\geq 3$ *at each node, loops allowed*, from "The on-line encyclopedia of integer sequences" (2023) A360864

[19] **M Kontsevich**, *Formal (non)commutative symplectic geometry*, from "The Gelfand Mathematical Seminars, 1990–1992" (L Corwin, I Gelfand, J Lepowsky, editors), Birkhäuser, Boston, MA (1993) 173–187 MR

[20] **M Kontsevich**, *Feynman diagrams and low-dimensional topology*, from "First European Congress of Mathematics, II" (A Joseph, F Mignot, F Murat, B Prum, R Rentschler, editors), Progr. Math. 120, Birkhäuser, Basel (1994) 97–121 MR

[21] **S Morita**, **T Sakasai**, **M Suzuki**, *Integral Euler characteristic of* $\mathrm{Out}\, F_{11}$, Exp. Math. 24 (2015) 93–97 MR

[22] **G Pólya**, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Math. 68 (1937) 145–254 MR

[23] *SageMath*, *version* 9.0 (2020) Available at `https://www.sagemath.org`

[24] **A Saied**, *On the FI-module structure of $H^i(\Gamma_{n,s})$*, preprint (2015) arXiv 1506.05861

[25] **R P Stanley**, *Enumerative combinatorics, I*, Cambridge Studies in Advanced Mathematics 49, Cambridge Univ. Press (1997) MR

[26] **V Turchin**, **T Willwacher**, *Commutative hairy graphs and representations of* $\mathrm{Out}(F_r)$, J. Topol. 10 (2017) 386–411 MR

[27] **J A M Vermaseren**, *New features of FORM*, preprint (2000) arXiv math-ph/0010025

[28] **J A M Vermaseren**, *FORM tutorial on the code that computes the Euler characteristic of $\mathcal{MG}_{g,n}$*, lecture notes (2020) Available at `https://www.nikhef.nl/~form/maindir/courses/uam2022/six.pdf`

*Perimeter Institute for Theoretical Physics*
*Waterloo, Canada*

*Nikhef*
*Amsterdam, Netherlands*

`mborinsky@perimeterinstitute.ca, t68@nikhef.nl`

# Classification of metric fibrations

Yasuhiko Asao

We study a notion of "fibration for metric spaces", called *metric fibration*, that was originally introduced by Leinster (Doc. Math. 18 (2013) 857–905) in the study of *magnitude*. He showed that the magnitude of a metric fibration splits into the product of those of the fiber and the base, which is analogous to the case for the Euler characteristic and topological fiber bundles. His idea and our approach are based on Lawvere's suggestion of viewing a metric space as an enriched category (Rend. Sem. Mat. Fis. Milano 43 (1973) 135–166). Actually, the metric fibrations are the restriction of the enriched *Grothendieck fibrations* (Séminaire Bourbaki 1959/1960 (1966) exposé 190) to metric spaces (arXiv 2303.05677). We give a complete classification of metric fibrations by several means, which are parallel to those used for topological fiber bundles. That is, the classification of metric fibrations is reduced to that of "principal fibrations", which is done by the "1-Čech cohomology" in an appropriate sense. Here we introduce the notion of *torsors* in the category of metric spaces, and the discussions are analogous to those in sheaf theory. Further, we can define the "fundamental group" $\pi_1^m(X)$ of a metric space $X$, which is a group-like object in metric spaces, such that the conjugation classes of homomorphisms $\mathrm{Hom}(\pi_1^m(X), \mathcal{G})$ correspond to the isomorphism classes of "principal $\mathcal{G}$-fibrations" over $X$. In other words, the latter are classified like topological covering spaces.

18F15, 51F30

## 1 Introduction

The idea of *metric fibration* was first introduced by Leinster in the study of magnitude [5]. The magnitude theory that he coined can be considered as a promotion of Lawvere's suggestion of viewing a metric space as a $[0, \infty]$-enriched category. The magnitude of a metric space was defined as a special case of the "Euler characteristic of enriched categories". In fact, he showed that the magnitude of a metric fibration splits into the product of those of the fiber and the base (Theorem 2.3.11 of [5]), which is analogous to the case of topological fiber bundles. Later, the author [1] pointed out that metric fibration can actually be seen as enriched *Grothendieck fibrations*, see [2], when restricted to metric spaces. Here we deal with small categories and metric spaces from a unified viewpoint, namely as *filtered set enriched categories*. By this approach, we can expect to obtain novel ideas for the study of metric spaces by transferring well-understood concepts in category theory, and vice versa.

As an example, Figure 1 is one of the simplest nontrivial metric fibrations. Note that we consider connected graphs as metric spaces by taking the shortest path metric (see also Proposition 2.8). Both graphs are
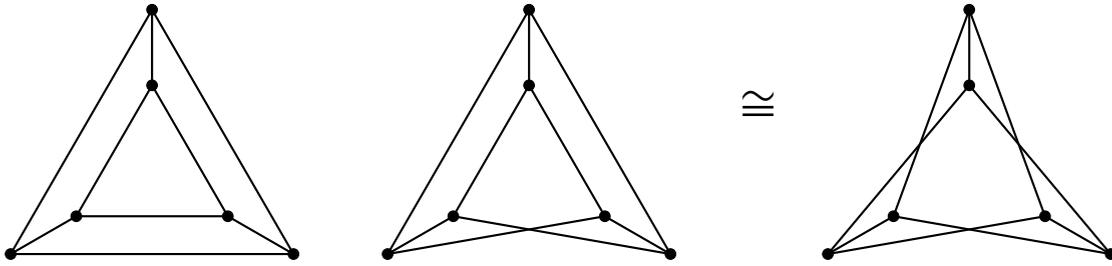
Figure 1: The left is $K_3 \times K_2$, and the right is isomorphic to $K_{3,3}$. They both have magnitude equal to $6/(1 + 3q + 2q^2)$.

metric fibrations over the complete graph $K_3$ with fiber $K_2$ as shown in Example 5.29 of [1]. Further, they have the same magnitude as pointed out in Example 3.7 of [6]. In Proposition 5.30 of [1], it is shown that the right one is the only nontrivial metric fibration over $K_3$ with fiber $K_2$. Here, "trivial" means that it is the cartesian product of graphs. On the other hand, any metric fibration over a four-cycle graph $C_4$, or more generally an even-cycle graph, is shown to be trivial in the same proposition.

In this paper, we give a complete classification of metric fibrations by several means, which are parallel to those used to classify topological fiber bundles. Namely, we define "principal fibrations", "fundamental groups" and "a 1-Čech cohomology" for metric spaces, and obtain an equivalence between categories of these objects. Roughly speaking, we obtain an analogy of the following correspondence in the case of topological fiber bundles with a discrete structure group:

$$\text{Fiber bundles over } X \text{ with structure group } G$$
$$\updownarrow$$
$$\text{Principal } G\text{-bundles over } X \ (G\text{-torsors})$$
$$\updownarrow$$
$$[X, BG] \cong \mathsf{Hom}(\pi_1(X), G)/\text{conjugation}$$
$$\updownarrow$$
$$\mathsf{H}^1(X, G)$$

We give more details below. First recall that any Grothendieck fibration (in the usual sense) over a small category $C$ can be obtained from a lax functor $C \to \mathsf{Cat}$, by a procedure known as the *Grothendieck construction* [3]. In [1], it is shown that any metric fibration over a metric space $X$ can be obtained from a "lax functor" $X \to \mathsf{Met}$ that is called *metric action* (Definition 3.1). Here Met is the category of metric spaces and Lipschitz maps. Whereas metric fibrations can be defined by "a lifting property" in a model categorical spirit, metric actions are best understood by "transformation functions". The following shows that these two viewpoints are in fact equivalent.

**Theorem 1.1** (Proposition 3.9)  *For a metric space $X$, the Grothendieck construction gives an equivalence of categories*

$$\mathsf{Met}_X \simeq \mathsf{Fib}_X,$$

*where we denote the category of metric actions $X \to \mathsf{Met}$ by $\mathsf{Met}_X$ and the category of metric fibrations over $X$ by $\mathsf{Fib}_X$ (Definitions 3.1 and 3.2).*

We can define a category $\mathsf{Tors}_X^{\mathcal{G}}$ that consists of "principal $\mathcal{G}$-fibrations" (Definition 5.7). We call it the *category of $\mathcal{G}$-torsors*. We can also define a subcategory $\mathsf{PMet}_X^{\mathcal{G}}$ of $\mathsf{Met}_X$, that is the counterpart of $\mathsf{Tors}_X^{\mathcal{G}}$ (Definition 5.1). The objects of the category $\mathsf{PMet}_X^{\mathcal{G}}$ consist of a metric action $X \to \mathsf{Met}$ taking a group $\mathcal{G}$, not just a metric space, as the value. Then we have the following.

**Theorem 1.2** (Proposition 5.11)  *For a "metric group" $\mathcal{G}$, the Grothendieck construction gives an equivalence of categories*

$$\mathsf{PMet}_X^{\mathcal{G}} \simeq \mathsf{Tors}_X^{\mathcal{G}}.$$

Here, the group $\mathcal{G}$ is not just a group but is a group-like object in $\mathsf{Met}$, which we call a *metric group* (Definition 4.1). As an example of a metric group, we construct the *fundamental group $\pi_1^m(X)$ of a metric space $X$* (Definition 4.8). We also define a category $\mathsf{Hom}(\pi_1^m(X), \mathcal{G})$ of homomorphisms $\pi_1^m(X) \to \mathcal{G}$, where a morphism between homomorphisms is defined as a conjugation relation (Definition 5.12). Then we have the following.

**Theorem 1.3** (Proposition 5.15)  *We have an equivalence of categories*

$$\mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G}) \simeq \mathsf{PMet}_X^{\mathcal{G}}.$$

As a corollary, we reprove Proposition 5.30 of [1] in the following form. We note that the notion of a metric group is equivalent to that of a "*normed group*" (Proposition 4.5). For a metric group $\mathcal{G}$, we denote the corresponding norm of an element $g \in \mathcal{G}$ by $|g| \in \mathbb{Z}_{\geq 0}$.

**Proposition 1.4** (Proposition 5.16)  *Let $C_n$ be the undirected $n$-cycle graph. Then we have*

$$\pi_1^m(C_n) \cong \begin{cases} \mathbb{Z} \text{ with } |1| = 1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

*Hence we have that*

$$\mathsf{PMet}_{C_n}^{\mathcal{G}} \simeq \begin{cases} \mathsf{Hom}(\mathbb{Z}, \mathcal{G}) & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even,} \end{cases}$$

*for all metric groups $\mathcal{G}$, which implies that there is only a trivial metric fibration over $C_{2n}$ and that there is at most one nontrivial metric fibration over $C_{2n+1}$.*

Now, similarly to the topological case, we can define an "associated bundle construction" from a torsor and a metric space $Y$ (Corollary 6.10). This construction gives the following.

**Theorem 1.5** (Corollary 6.11) *Suppose that $Y$ is a bounded metric space. Then we have an equivalence of categories*

$$\mathsf{PMet}_X^{\mathsf{Aut}\,Y} \simeq \mathsf{core}\,\mathsf{Fib}_X^Y,$$

*where $\mathsf{Fib}_X^Y$ is the full subcategory of $\mathsf{Fib}_X$ that consists of metric fibrations with fiber $Y$ (Definition 6.5), and we denote the core of a category by $\mathsf{core}$ (Definition 2.1(4)).*

Here, we equip the group $\mathsf{Aut}\,Y$ of isometries of $Y$ with a metric group structure by $d_{\mathsf{Aut}\,Y}(f,g) = \sup_{y \in Y} d_Y(fy, gy)$ (Example 4.3). However, for this we should suppose that $Y$ is a bounded metric space so that $d_{\mathsf{Aut}\,Y}$ is indeed a distance function. For the case of general metric fibrations, we must extend our arguments to *extended metric groups*, allowing $\infty$ as the value of the distance function (Definition 6.12). For these we obtain an essentially similar but extended result (Proposition 6.18).

Finally, we define a "1-Čech cohomology category" $\mathsf{H}^1(X, \mathcal{G})$ of a $\mathcal{G}$-torsor $X$ (Definition 7.2). This is analogous to the Čech cohomology constructed from the local sections of a principal bundle. Similarly to the topological case, we can construct a cocycle from a family of local sections (Proposition 7.9), and conversely we can construct a $\mathcal{G}$-torsor by pasting copies of $\mathcal{G}$ along a cocycle (Proposition 7.5). From this correspondence we have the following:

**Theorem 1.6** (Corollary 7.11) *We have an equivalence of categories*

$$\mathsf{H}^1(X; \mathcal{G}) \simeq \mathsf{Tors}_X^{\mathcal{G}}.$$

# 2 Conventions

In this section, we review terms for categories, graphs, weighted graphs and metric spaces that are well known but may not be commonly used. Those who are not familiar with the language of categories and functors may refer to [8]. For categories and metric spaces, [7] may be a good reference.

## 2.1 Categories

In this article, we suppose that categories are locally small. We denote the object class of a category $C$ by $\mathsf{Ob}\,C$, and the set of all morphisms from $a$ to $b$ by $C(a,b)$ for each pair of objects $a, b \in \mathsf{Ob}\,C$. We denote the class of all morphisms in $C$ by $\mathsf{Mor}\,C$.

**Definition 2.1** Let $C$ and $D$ be categories, and $F\colon C \to D$ be a functor.

(1) We say that $F$ is *faithful* if the map $F\colon C(a,b) \to D(Fa, Fb)$ is injective for all objects $a, b \in \operatorname{Ob} C$. We say that $F$ is *full* if the map $F\colon C(a,b) \to D(Fa, Fb)$ is surjective for all objects $a, b \in \operatorname{Ob} C$. We say that $F$ is *fully faithful* if it is faithful and full.

(2) We say that $F$ is *split essentially surjective* if a family of isomorphisms $\{Fc \cong d \mid c \in \operatorname{Ob} C\}_{d \in \operatorname{Ob} D}$ exists.

(3) We say that $F$ is a *category equivalence* if there exists a functor $G\colon D \to C$ and natural isomorphisms $GF \cong \operatorname{id}_C$ and $FG \cong \operatorname{id}_D$. When there exists an equivalence of categories $C \to D$, we say that *$C$ and $D$ are equivalent.*

(4) We define a groupoid core $C$ by $\operatorname{Ob} \operatorname{core} C = \operatorname{Ob} C$ and $\operatorname{core} C(a,b) = \{f \in C(a,b) \mid f \text{ is an isomorphism}\}$ for all $a, b \in \operatorname{Ob} C$.

The following are standard.

**Lemma 2.2** *If a functor $F\colon C \to D$ is fully faithful and split essentially surjective, then it is an equivalence of categories.* $\qquad\square$

**Lemma 2.3** *An equivalence of categories $F\colon C \to D$ induces an equivalence of categories $\operatorname{core} F\colon \operatorname{core} C \to \operatorname{core} D$.* $\qquad\square$

**Remark 2.4** For a classification of objects of a category, we often want to consider "isomorphism classes of objects" and compare it with another category. However, in general, we cannot do that since the class of objects is not necessarily a set. Instead, we consider an equivalence of categories $\operatorname{core} C \to \operatorname{core} D$ that implies a bijection between isomorphism classes of objects if they are small.

## 2.2 Metric spaces

**Definition 2.5** (1) A *pseudometric space* $(X, d)$ is a set $X$ equipped with a function $d\colon X \to \mathbb{R}_{\geq 0}$ satisfying that, for all $x, x', x'' \in X$,

- $d(x, x) = 0$,
- $d(x, x') = d(x', x)$,
- $d(x, x') + d(x', x'') \geq d(x, x'')$.

(2) A *Lipschitz map* $f\colon X \to Y$ between pseudometric spaces $X$ and $Y$ is a map satisfying that $d_Y(fx, fx') \leq d_X(x, x')$ for all $x, x' \in X$. We denote the category of pseudometric spaces and Lipschitz maps by $\psi\mathrm{Met}$. We call an isomorphism in $\psi\mathrm{Met}$ an *isometry*.

(3) A *metric space* $(X, d)$ is a pseudometric space satisfying that

- $d(x, x') = 0$ if and only if $x = x'$.

We denote the full subcategory of $\psi\mathrm{Met}$ that consists of metric spaces by $\mathrm{Met}$.

**Definition 2.6** (1)  A *graph* $G$ is a pair of sets $(V(G), E(G))$ such that $E(G) \subset \{e \in 2^{V(G)} \mid \#e = 2\}$, where we denote the cardinality of a set by #. We call an element of $V(G)$ a *vertex*, and an element of $E(G)$ an *edge*. A *graph homomorphism* $f : G \to H$ between graphs $G$ and $H$ is a map $f : V(G) \to V(H)$ such that $fe \in E(H)$ or $\#fe = 1$ for all $e \in E(G)$. We denote the category of graphs and graph homomorphisms by Grph.

(2)  A *path* in a graph $G$ is a tuple $(x_0, \ldots, x_n) \in V(G)^{n+1}$ for some $n \geq 0$ such that $\{x_i, x_{i+1}\} \in E(G)$ for all $0 \leq i \leq n - 1$. A *connected graph* $G$ is a graph such that for every $x, x' \in V(G)$ there exists a path $(x_0, \ldots, x_n)$ with $x_0 = x$ and $x_n = x'$. We denote the full subcategory of Grph that consists of connected graphs by $\mathsf{Grph_{conn}}$.

(3)  A *weighted graph* $(G, w_G)$ is a graph $G$ equipped with a function $w_G : E(G) \to \mathbb{R}_{\geq 0}$. A *weighted graph homomorphism* $f : G \to H$ between weighted graphs $G$ and $H$ is a graph homomorphism such that $w_H(fe) \leq w_G(e)$ for all $e \in E(G)$, where we stipulate that $w_H(fe) = 0$ if $\#fe = 1$. We denote the category of weighted graphs and weighted graph homomorphisms by wGrph. We also denote the full subcategory of wGrph that consists of weighted graphs $(G, w_G)$ such that the graph $G$ is connected by $\mathsf{wGrph_{conn}}$.

**Definition 2.7**  We define functors $\mathsf{Met} \to \psi\mathsf{Met}$ and $\mathsf{wGrph_{conn}} \to \mathsf{Grph_{conn}}$ by forgetting additional properties. We also define the functor $\psi\mathsf{Met} \to \mathsf{wGrph_{conn}}$ that sends a pseudometric space $(X, d)$ to a weighted graph $(X, w_X)$ defined by $V(X) = X$, $E(X) = \{e \in 2^X \mid \#e = 2\}$ and $w_X\{x, x'\} = d(x, x')$.

**Proposition 2.8**  *The above functors have left adjoints.*

**Proof**  We describe each functor $F$ in the following, and they are the left adjoint functors of each functor $G$ of the above since the unit and the counit give that $FGF = F$ and $GFG = G$.

• We define a functor $\mathsf{Grph_{conn}} \to \mathsf{wGrph_{conn}}$ by sending a connected graph to a weighted graph with $w = 0$.

• We define a functor $\mathsf{wGrph_{conn}} \to \psi\mathsf{Met}$ by sending a weighted graph $(G, w_G)$ to a pseudometric space $(V(G), d_G)$ defined by

$$d_G(x, x') = \inf_{n \geq 0} \bigcup \left\{ \sum_{i=0}^{n-1} w_G\{x_i, x_{i+1}\} \mid (x = x_0, \ldots, x_n = x') \text{ is a path on } G \right\}.$$

• We define a functor $\psi\mathsf{Met} \to \mathsf{Met}$ by sending a pseudometric space $(X, d)$ to a metric space $(\mathsf{KQ}X, \tilde{d})$ defined as follows. We define an equivalence relation $\sim$ on $X$ by $x \sim x'$ if and only if $d(x, x') = 0$. We also define a function $\mathsf{KQ}X := X/\sim \to \mathbb{R}_{\geq 0}$ by $\tilde{d}([x], [x']) = d(x, x')$.  $\square$

**Definition 2.9**  For a pseudometric space $X$, we call the metric space $\mathsf{KQ}X$ the *Kolmogorov quotient* of $X$.

**Definition 2.10** (1)  For pseudometric spaces $(X, d_X)$ and $(Y, d_Y)$, we define a metric space called the $L^1$-*product* $(X \times Y, d_{X \times Y})$ by $d_{X \times Y}((x, y), (x', y')) = d_X(x, x') + d_Y(y, y')$ for all $x, x' \in X$ and $y, y' \in Y$.

(2) For graphs $G$ and $H$, we define a graph called the *cartesian product* $G \times H$ by $V(G \times H) = V(G) \times V(H)$, and $\{(x, y), (x', y')\} \in E(G \times H)$ if and only if one of the following holds:

- $x = x'$ and $\{y, y'\} \in E(H)$,
- $\{x, x'\} \in E(G)$ and $y = y'$,

for all $x, x' \in V(G)$ and $y, y' \in V(H)$.

(3) For weighted graphs $(G, w_G)$ and $(H, w_H)$, we define a weighted graph $(G \times H, w_{G \times H})$ by $w_{G \times H}\{(x, y), (x', y')\} = w_G\{x, x'\} + w_H\{y, y'\}$ for all $\{(x, y), (x', y')\} \in E(G \times H)$, where $G \times H$ is the cartesian product of graphs and we stipulate that $w_G\{x, x\} = w_H\{y, y\} = 0$.

These products make each category a symmetric monoidal category.

**Proposition 2.11**  *The functors* $\mathsf{Met} \to \psi\mathsf{Met} \to \mathsf{wGrph}_{\mathsf{conn}} \to \mathsf{Grph}_{\mathsf{conn}}$ *and their left adjoints are strong monoidal except for the functor* $\psi\mathsf{Met} \to \mathsf{wGrph}_{\mathsf{conn}}$ *that is lax monoidal.*

**Proof**  For the functors $\mathsf{Met} \to \psi\mathsf{Met}$ and $\mathsf{wGrph}_{\mathsf{conn}} \to \mathsf{Grph}_{\mathsf{conn}}$, the claim is obvious since they are inclusions. The claim is also obvious for the functor $\mathsf{Grph}_{\mathsf{conn}} \to \mathsf{wGrph}_{\mathsf{conn}}$ by the definition. For the functor $\psi\mathsf{Met} \to \mathsf{Met}$, we define a map $\mathsf{KQ}(X \times Y) \to \mathsf{KQ}X \times \mathsf{KQ}Y$ by $[(x, y)] \mapsto ([x], [y])$. This is obviously natural and is an isometry since we have that $[(x, y)] \sim [(x', y')]$ if and only if $[x] \sim [x']$ and $[y] \sim [y']$. For the functor $F : \mathsf{wGrph}_{\mathsf{conn}} \to \psi\mathsf{Met}$, the identity on the set $F(G \times H) = F(G) \times F(H)$ is an isometry since

$$
d_{w_{G \times H}}((x, y), (x', y')) = \inf_{n \geq 0} \bigcup \left\{ \sum_{i=0}^{n-1} w_{G \times H}\{(x_i, y_i), (x_{i+1}, y_{i+1})\} \mid \right.
$$
$$
\left. ((x, y) = (x_0, y_0), \dots, (x_n, y_n) = (x', y')) \text{ is a path on } G \times H \right\}
$$
$$
= \inf_{n \geq 0} \bigcup \left\{ \sum_{i=0}^{n-1} w_G\{x_i, x_{i+1}\} + w_H\{y_i, y_{i+1}\} \mid \right.
$$
$$
\left. ((x, y) = (x_0, y_0), \dots, (x_n, y_n) = (x', y')) \text{ is a path on } G \times H \right\}
$$
$$
= \inf_{n \geq 0} \bigcup \left\{ \sum_{i=0}^{n-1} w_G\{x_i, x_{i+1}\} \mid (x = x_0, \dots, x_n = x') \right\}
$$
$$
+ \inf_{m \geq 0} \bigcup \left\{ \sum_{i=0}^{m-1} w_H\{y_i, y_{i+1}\} \mid (y = y_0, \dots, y_m = y') \right\}
$$
$$
= d_{w_G}(x, x') + d_{w_H}(y, y')
$$
$$
= d_{F(G) \times F(H)}((x, y), (x', y')),
$$

for all $x, x' \in V(G)$ and $y, y' \in V(H)$. It is obviously natural. Finally, for the functor $G : \psi\mathsf{Met} \to \mathsf{wGrph}_{\mathsf{conn}}$, the identity on the set $G(X) \times G(Y) = G(X \times Y)$ is a weighted graph homomorphism since it is an inclusion of graphs and preserves weightings. It is obviously natural.  □

**Definition 2.12** (1) An *extended pseudometric space* is a set $X$ equipped with a function $d\colon X \to [0, \infty]$ that satisfies the same conditions for pseudometric spaces. In other words, it is a pseudometric space admitting $\infty$ as a value of distance. A *Lipschitz map* between extended pseudometric spaces is a distance nonincreasing map. We denote the category of extended pseudometric spaces and Lipschitz maps by $\mathsf{E\psi Met}$. We similarly define *extended metric spaces* and we denote the full subcategory of $\mathsf{E\psi Met}$ that consists of them by $\mathsf{EMet}$.

(2) We define the $L^1$-product of extended pseudometric spaces similarly to that of pseudometric spaces. It makes the category $\mathsf{E\psi Met}$ a symmetric monoidal category.

(3) We define functors $\mathsf{EMet} \to \mathsf{E\psi Met}$ and $\mathsf{wGrph} \to \mathsf{Grph}$ by forgetting additional properties. We also define the functor $\mathsf{E\psi Met} \to \mathsf{wGrph}$ similarly to the functor $\mathsf{\psi Met} \to \mathsf{wGrph_{conn}}$ except that $\{x, x'\}$ does not span an edge for $x, x' \in X$ with $d(x, x') = \infty$.

The following is immediate.

**Proposition 2.13** (1) *The functors* $\mathsf{EMet} \to \mathsf{E\psi Met} \to \mathsf{wGrph} \to \mathsf{Grph}$ *have left adjoints. Further*, *the following diagram is commutative*, *where the vertical functors are all inclusions*:

$$
\begin{array}{ccccccc}
\mathsf{EMet} & \longrightarrow & \mathsf{E\psi Met} & \longrightarrow & \mathsf{wGrph} & \longrightarrow & \mathsf{Grph} \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
\mathsf{Met} & \longrightarrow & \mathsf{\psi Met} & \longrightarrow & \mathsf{wGrph_{conn}} & \longrightarrow & \mathsf{Grph_{conn}}
\end{array}
$$

(2) *The functors* $\mathsf{EMet} \to \mathsf{E\psi Met}$ *and* $\mathsf{wGrph} \to \mathsf{Grph}$ *are strong monoidal and the functor* $\mathsf{E\psi Met} \to \mathsf{wGrph}$ *is lax monoidal.* $\qquad\square$

# 3 $\mathsf{Met}_X \simeq \mathsf{Fib}_X$

In this section, we introduce two notions, the *metric action* and the *metric fibration*, and show the equivalence between them. The notion of metric fibration is originally introduced by Leinster [5] in the study of magnitude. The metric action was introduced by the present author in [1], and is the counterpart of *lax functors* in category theory, while the metric fibration is a generalization of the *Grothendieck fibration*. As written in the introduction, we can think of the Grothendieck (or metric) fibration as the definition of fibrations by "the lifting property", while the lax functor is the one by "the transformation functions".

**Definition 3.1** Let $X$ be a metric space.

(1) A *metric action* $F\colon X \to \mathsf{Met}$ consists of metric spaces $Fx \in \mathsf{Met}$ for all $x \in X$ and isometries $F_{xx'}\colon Fx \to Fx'$ for all $x, x' \in X$ satisfying the following for all $x, x', x'' \in X$:

- $F_{xx} = \mathrm{id}_{Fx}$ and $F_{x'x} = F_{xx'}^{-1}$.
- $d_{Fx''}(F_{x'x''}F_{xx'}a, F_{xx''}a) \le d_X(x, x') + d_X(x', x'') - d_X(x, x'')$ for every $a \in Fx$.

(2) A *metric transformation* $\theta\colon F \Rightarrow G$ consists of Lipschitz maps $\theta_x\colon Fx \to Gx$ for all $x \in X$ satisfying that $G_{xx'}\theta_x = \theta_{x'}F_{xx'}$ for all $x, x' \in X$. We can define the composition of metric transformations $\theta$ and $\theta'$ by $(\theta'\theta)_x = \theta'_x\theta_x$. We denote the category of metric actions $X \to \mathsf{Met}$ and metric transformations by $\mathsf{Met}_X$.

**Definition 3.2** (1) Let $\pi\colon E \to X$ be a Lipschitz map between metric spaces. We say that $\pi$ is a *metric fibration over $X$* if it satisfies the following: for all $\varepsilon \in E$ and $x \in X$, there uniquely exists $\varepsilon_x \in \pi^{-1}x$ such that

- $d_E(\varepsilon, \varepsilon_x) = d_X(\pi\varepsilon, x)$,
- $d_E(\varepsilon, \varepsilon') = d_E(\varepsilon, \varepsilon_x) + d_E(\varepsilon_x, \varepsilon')$ for all $\varepsilon' \in \pi^{-1}x$.

We call the point $\varepsilon_x$ the *lift of $x$ along $\varepsilon$*.

(2) For metric fibrations $\pi\colon E \to X$ and $\pi'\colon E' \to X$, a *morphism* $\varphi\colon \pi \to \pi'$ is a Lipschitz map $\varphi\colon E \to E'$ such that $\pi'\varphi = \pi$. We denote the category of metric fibrations over $X$ and their morphisms by $\mathsf{Fib}_X$.

**Example 3.3** For a product of metric spaces $E = X \times Y$, the projection $X \times Y \to X$ is a metric fibration. We call it a *trivial metric fibration*.

**Lemma 3.4** *Let $\pi\colon E \to X$ be a metric fibration, and $x, x' \in X$. Then the correspondence $\pi^{-1}x \ni a \mapsto a_{x'} \in \pi^{-1}x'$ is an isometry, where we equip the sets $\pi^{-1}x$ and $\pi^{-1}x'$ with the induced metric from $E$.*

**Proof** The statement is obviously true if $E = \varnothing$. We suppose that $E \neq \varnothing$ in the following; then every fiber $\pi^{-1}x$ is nonempty. For $a \in \pi^{-1}x$, we have

$$d_E(a_{x'}, a) = d_E(a_{x'}, (a_{x'})_x) + d_E((a_{x'})_x, a) = d_X(x', x) + d_E((a_{x'})_x, a).$$

We also have $d_E(a, a_{x'}) = d_X(x, x')$. Hence we obtain that $d_E((a_{x'})_x, a) = 0$, and thus $(a_{x'})_x = a$ for all $x, x' \in X$. This implies that the correspondence is a bijection. Further, we have

$$d_E(a, b_{x'}) = d_E(a, a_{x'}) + d_E(a_{x'}, b_{x'}) = d_X(x, x') + d_E(a_{x'}, b_{x'})$$

and

$$d_E(b_{x'}, a) = d_E(b_{x'}, b) + d_E(b, a) = d_X(x', x) + d_E(b, a)$$

for all $a, b \in \pi^{-1}x$. We thereby obtain that $d_E(a, b) = d_E(a_{x'}, b_{x'})$ for all $x, x' \in X$ and $a, b \in \pi^{-1}x$, which implies that the correspondence is an isometry. $\square$

**Lemma 3.5** *Let $\varphi\colon \pi \to \pi'$ be a morphism of metric fibrations. For all $x, x' \in X$ and $a \in \pi^{-1}x$, we have $(\varphi a)_{x'} = \varphi a_{x'}$.*

**Proof** We have

$$d_{E'}((\varphi a)_{x'}, \varphi a_{x'}) = d_{E'}(\varphi a, \varphi a_{x'}) - d_X(x, x') \leq d_E(a, a_{x'}) - d_X(x, x') = 0,$$

and hence we obtain that $(\varphi a)_{x'} = \varphi a_{x'}$. $\square$

**Definition 3.6** Let $F\colon X \to \mathsf{Met}$ be a metric action. We define a metric fibration $\pi_F\colon E(F) \to X$ as follows:

(1) $E(F) = \{(x,a) \mid a \in Fx,\ x \in X\}$.

(2) $d_{E(F)}((x,a),(x',b)) = d_X(x,x') + d_{Fx'}(F_{xx'}a,b)$.

(3) $\pi_F(x,a) = x$.

We call the above construction the *Grothendieck construction*.

**Proposition 3.7** *The Grothendieck construction gives a functor $E\colon \mathsf{Met}_X \to \mathsf{Fib}_X$.*

**Proof** Let $\theta\colon F \Rightarrow G$ be a metric transformation. Then we construct Lipschitz maps $\varphi_\theta\colon E(F) \to E(G)$ by $\varphi_\theta(x,a) = (x, \theta_x a)$ for all $x \in X$ and $a \in Fx$. To see that $\varphi_\theta$ is a Lipschitz map, observe that

$$
\begin{aligned}
d_{E(G)}(\varphi_\theta(x,a), \varphi_\theta(x',b)) &= d_{E(G)}((x,\theta_x a),(x',\theta_{x'}b)) \\
&= d_X(x,x') + d_{Gx'}(G_{xx'}\theta_x a, \theta_{x'}b) \\
&= d_X(x,x') + d_{Gx'}(\theta_{x'}F_{xx'}a, \theta_{x'}b) \\
&\le d_X(x,x') + d_{Fx'}(F_{xx'}a, b) \\
&= d_{E(F)}((x,a),(x',b)).
\end{aligned}
$$

It remains to see that the correspondence $\theta \mapsto \varphi_\theta$ is functorial — that is we have $\varphi_{\mathrm{id}_F} = \mathrm{id}_{E(F)}$ and $\varphi_{\theta'\theta} = \varphi_{\theta'}\varphi_\theta$ for all metric transformations $\theta\colon F \Rightarrow G$ and $\theta'\colon G \Rightarrow H$. The former is obvious and the latter is checked as follows:

$$
\varphi_{\theta'\theta}(x,a) = (x, (\theta'\theta)_x a) = (x, \theta'_x \theta_x a) = \varphi_{\theta'}\varphi_\theta(x,a).
$$

Finally, $\varphi_\theta$ is obviously a morphism of the metric fibration. $\qquad\square$

**Proposition 3.8** *We have a functor $F\colon \mathsf{Fib}_X \to \mathsf{Met}_X$ sending a metric fibration $\pi$ to a metric action $F_\pi$ with $F_\pi x = \pi^{-1}x$.*

**Proof** Let $\pi\colon E \to X$ be a metric fibration. We define a metric action $F_\pi\colon X \to \mathsf{Met}$ by $F_\pi x = \pi^{-1}x$ and $(F_\pi)_{xx'}a = a_{x'}$ for all $x, x' \in X$ and $a \in \pi^{-1}x$, where we equip the set $\pi^{-1}x$ with the induced metric from $E$. It follows that $(F_\pi)_{xx} = \mathrm{id}_{F_\pi x}$ by the uniqueness of the lifts, and that $(F_\pi)_{xx'}$ defines an isometry $F_\pi x \to F_\pi x'$ with $(F_\pi)^{-1}_{xx'} = (F_\pi)_{x'x}$ by Lemma 3.4. Further, we have that

$$
\begin{aligned}
d_{F_\pi x''}((F_\pi)_{x'x''}(F_\pi)_{xx'}a, (F_\pi)_{xx''}a) &= d_{F_\pi x''}((a_{x'})_{x''}, a_{x''}) \\
&= d_E(a, (a_{x'})_{x''}) - d_X(x, x'') \\
&\le d_E(a, a_{x'}) + d_E(a_{x'}, (a_{x'})_{x''}) - d_X(x, x'') \\
&= d_X(x, x') + d_X(x', x'') - d_X(x, x''),
\end{aligned}
$$

for all $x, x', x'' \in X$ and $a \in F_\pi x$. Hence $F_\pi$ defines a metric action $X \to$ Met. Next, let $\varphi \colon \pi \to \pi'$ be a morphism of metric fibrations. We define a metric transformation $\theta_\varphi \colon F_\pi \Rightarrow F_{\pi'}$ by $(\theta_\varphi)_x a = \varphi a$ for all $x \in X$ and $a \in F_\pi x$. Then we have that

$$(F_{\pi'})_{xx'}(\theta_\varphi)_x a = (F_{\pi'})_{xx'} \varphi a = (\varphi a)_{x'} = \varphi a_{x'} = (\theta_\varphi)_{x'}(F_\pi)_{xx'},$$

where the third line follows from Lemma 3.5. Thus, $\theta_\varphi$ defines a metric transformation $F_\pi \Rightarrow F_{\pi'}$. Note that we have $\theta_{\mathrm{id}_\pi} = \mathrm{id}_{F_\pi}$ and $(\theta_{\psi\varphi})_x a = \psi \varphi a = (\theta_\psi)_x (\theta_\varphi)_x a$ for morphisms $\varphi$ and $\psi$, which implies the functoriality of $F$. $\square$

The following is the counterpart of the correspondence between lax functors and the Grothendieck fibrations [4, B1], and enhances Corollary 5.26 of [1].

**Proposition 3.9** *The Grothendieck construction functor $E \colon \mathrm{Met}_X \to \mathrm{Fib}_X$ is an equivalence of categories.*

**Proof** We show that $FE \cong \mathrm{id}_{\mathrm{Met}_X}$ and $EF \cong \mathrm{id}_{\mathrm{Fib}_X}$. It is immediate to verify $FE \cong \mathrm{id}_{\mathrm{Met}_X}$ by the definition. We show that $EF_\pi \cong \pi$ for a metric fibration $\pi \colon E \to X$. Note that $EF_\pi$ is a metric space consisting of points $(x, a)$ with $x \in X$ and $a \in \pi^{-1}x$, and we have $d_{EF_\pi}((x, a), (x', a')) = d_X(x, x') + d_{\pi^{-1}x'}(a_{x'}, a')$. We define a map $f \colon EF_\pi \to E$ by $f(x, a) = a$ for all $x \in X$ and $a \in \pi^{-1}x$. Then it is obviously an isometry and preserves fibers, and hence is an isomorphism of metric fibrations. The naturality of this isomorphism is obvious. $\square$

**Remark 3.10** The trivial metric fibration corresponds to the constant metric action, that is $F_{xx'} = \mathrm{id}$ for all $x, x' \in X$.

# 4 The metric fundamental group of a metric space

In this section, we give a concise introduction to *metric groups*. We also give a definition of *metric fundamental group*, which plays a role of $\pi_1$ for metric space in the classification of metric fibrations.

## 4.1 Metric groups

**Definition 4.1** (see [10, Definition 6.1]) (1) A *metric group* is a monoid object in Met that is a group when we forget the metric space structure. That is, a metric space $\mathcal{G}$ equipped with a Lipschitz map $\cdot \colon \mathcal{G} \times \mathcal{G} \to \mathcal{G}$, a function $(-)^{-1} \colon \mathcal{G} \to \mathcal{G}$ and a point $e \in \mathcal{G}$ satisfying the suitable conditions of monoids and groups.

(2) For metric groups $\mathcal{G}$ and $\mathcal{H}$, a *homomorphism* from $\mathcal{G}$ to $\mathcal{H}$ is a Lipschitz map $\mathcal{G} \to \mathcal{H}$ that commutes with the group structure.

(3) We denote the category of metric groups and homomorphisms by MGrp.

**Lemma 4.2** *Let $(\mathcal{G}, d)$ be a metric group.*

(1)  *We have $d(kg, kh) = d(g, h) = d(gk, hk)$ for all $g, h, k \in \mathcal{G}$.*

(2)  *We have $d(g, h) = d(g^{-1}, h^{-1})$ for all $g, h \in \mathcal{G}$.*

**Proof**  (1)  Since the map $\mathcal{G} \to \mathcal{G}$, $g \mapsto kg$, is a Lipschitz map for every $k \in \mathcal{G}$, we have $d(kg, kh) \leq d(g, h)$ and $d(k^{-1}(kg), k^{-1}(kh)) \leq d(kg, kh)$. Hence we obtain that $d(kg, kh) = d(g, h)$. The other identity can be proved similarly.

(2)  By (1), we have $d(g^{-1}, h^{-1}) = d(e, gh^{-1}) = d(h, g) = d(g, h)$.  $\square$

**Example 4.3**  Let $(X, d)$ be a metric space, and let $\mathsf{Aut}^u\, X$ be the set of isometries $f$ on $X$ such that $\sup_{x \in X} d_X(x, fx) < \infty$. We equip $\mathsf{Aut}^u\, X$ with a group structure by compositions. We also define a distance function on $\mathsf{Aut}^u\, X$ by $d_{\mathsf{Aut}^u\, X}(f, g) = \sup_{x \in X} d_X(fx, gx)$. It is straightforward to verify that $(\mathsf{Aut}^u\, X, d_{\mathsf{Aut}^u\, X})$ is a metric group. Note that, if the metric space $X$ is bounded, meaning that

$$\sup_{x, x' \in X} d_X(x, x') < \infty,$$

then the group $\mathsf{Aut}^u\, X$ consists of all isometries of $X$. In this case we denote it by $\mathsf{Aut}\, X$.

**Definition 4.4**  (1)  A *normed group* is a group $G$ equipped with a map $|-| : G \to \mathbb{R}_{\geq 0}$ called a norm, satisfying that

  •  $|g| = 0$ if and only if $g = e$,

  •  $|gh| \leq |g| + |h|$ for all $g, h \in G$.

(2)  A norm on $G$ is called *conjugation invariant* if it satisfies that $|h^{-1}gh| = |g|$ for all $g, h \in G$.

(3)  A norm on $G$ is called *inverse invariant* if it satisfies that $|g^{-1}| = |g|$ for all $g \in G$.

(4)  For normed groups $G$ and $H$, a *normed homomorphism* from $G$ to $H$ is a group homomorphism $\varphi : G \to H$ satisfying that $|\varphi g| \leq |g|$.

(5)  We denote the category of conjugation and inverse invariant normed groups and normed homomorphisms by $\mathsf{NGrp}^{-1}_{\mathrm{conj}}$.

**Proposition 4.5**  (E Roff, [9, Chapter 6])  *The categories $\mathsf{MGrp}$ and $\mathsf{NGrp}^{-1}_{\mathrm{conj}}$ are equivalent.*

**Proof**  Given a metric group $\mathcal{G}$, we can define a conjugation and inverse invariant normed group $\mathsf{N}\mathcal{G}$ by

  •  $\mathsf{N}\mathcal{G} = \mathcal{G}$ as a group,

  •  $|g| = d_{\mathcal{G}}(e, g)$ for all $g \in \mathsf{N}\mathcal{G}$.

This construction is functorial. Conversely, we can define a metric group M$G$ given a conjugation and inverse invariant normed group $G$ by

- M$G = G$ as a group,

- $d_{\mathrm{M}G}(g, h) = |h^{-1}g|$.

This construction is also functorial. It is straightforward to verify that the compositions of these functors are naturally isomorphic to the identities. □

## 4.2 The metric fundamental group

**Definition 4.6** Let $X$ be a metric space and $x \in X$.

(1) For each $n \geq 0$, we define a set $P_n(X, x)$ by

$$P_n(X, x) := \{(x, x_1, \ldots, x_n, x) \in X^{n+2}\}.$$

We also define that $P(X, x) := \bigcup_n P_n(X, x)$.

(2) We define a connected graph $G(X, x)$ with the vertex set $P(X, x)$ as follows. For $u, v \in P(X, x)$, an unordered pair $\{u, v\}$ spans an edge if and only if it satisfies both of the following:

- There is an $n \geq 0$ such that $u \in P_n(X, x)$ and $v \in P_{n+1}(X, x)$.

- There is a $0 \leq j \leq n$ such that $u_i = v_i$ for $1 \leq i \leq j$ and $u_i = v_{i+1}$ for $j + 1 \leq i \leq n$, where we have $u = (x, u_1, \ldots, u_n, x)$ and $v = (x, v_1, \ldots, v_{n+1}, x)$.

(3) We equip the graph $G(X, x)$ with a weighted graph structure by defining a function $w_{G(X,x)}$ on edges by

$$w_{G(X,x)}\{u, v\} = \begin{cases} d_X(v_j, v_{j+1}) + d_X(v_{j+1}, v_{j+2}) - d_X(v_j, v_{j+2}) & \text{if } v_j \neq v_{j+2}, \\ 0 & \text{if } v_j = v_{j+2}, \end{cases}$$

where we use the notation in (2).

(4) We denote the quasimetric space obtained from the weighted graph $G(X, x)$ by $Q(X, x)$. We denote the Kolmogorov quotient of $Q(X, x)$ by $\pi_1^m(X, x)$.

**Lemma 4.7** *Let $X$ be a metric space and $x \in X$.*

(1) *The metric space $\pi_1^m(X, x)$ has a metric group structure with multiplication defined by the concatenation defined as*

$$[(x, u_1, \ldots, u_n, x)] \bullet [(x, v_1, \ldots, v_k, x)] = [(x, u_1, \ldots, u_n, v_1, \ldots, v_k, x)].$$

*The unit is given by $[(x, x)] \in \pi_1^m(X, x)$.*

(2) *For all $x' \in X$, we have an isomorphism $\pi_1^m(X, x) \cong \pi_1^m(X, x')$ given by*

$$[(x, u_1, \ldots, u_n, x)] \mapsto [(x', x, u_1, \ldots, u_n, x, x')].$$

**Proof** (1) We first show that concatenation makes the weighted graph $G(X, x)$ into a monoid object in $\mathsf{wGrph}_{\mathrm{conn}}$. Let $(u, v), (u', v') \in G(X, x) \times G(X, x)$, and suppose that $\{(u, v), (u', v')\}$ spans an edge. Then we have that $u = u'$ and $v \in P_n(X, x), v' \in P_{n+1}(X, x)$, or $v = v'$ and $u \in P_n(X, x), u' \in P_{n+1}(X, x)$ for some $n$. We also have that $w_{G(X,x) \times G(X,x)}\{(u, v), (u', v')\} = w_{G(X,x)}\{u, u'\} + w_{G(X,x)}\{v, v'\}$. Note that $\{u \bullet v, u' \bullet v'\}$ spans an edge in $G(X, x)$. Further, we have $w_{G(X,x)}\{u \bullet v, u' \bullet v'\} = w_{G(X,x)}\{u, u'\} + w_{G(X,x)}\{v, v'\}$. Hence the concatenation map $\bullet \colon G(X, x) \times G(X, x) \to G(X, x)$ is a weighted graph homomorphism. It is immediate to verify that the identity is the element $(x, x)$ and that the multiplication is associative. Thus the weighted graph $G(X, x)$ is a monoid object in $\mathsf{wGrph}_{\mathrm{conn}}$, and by Proposition 2.11, $\pi_1^m(X, x)$ is a monoid object in Met.

Now we show that any element $[(x, x_0, \ldots, x_n, x)]$ has an inverse, namely the element $[(x, x_n, \ldots, x_0, x)]$. It is enough to show that

$$d_{Q(X,x)}((x, x_n, \ldots, x_0, x_0, \ldots, x_n x), (x, x)) = 0.$$

And indeed, it is obvious that the elements $(x, x_n, \ldots, x_0, x_0, \ldots, x_n x)$ and $(x, x)$ can be connected by a path that consists of edges with weight 0 in $G(X, x)$, which implies the desired equality.

(2) This is straightforward. $\qquad\square$

**Definition 4.8** Let $X$ be a metric space and $x \in X$. We call the metric group $\pi_1^m(X, x)$ the *metric fundamental group of* $X$ with respect to the base point $x$. We sometimes omit the base point and denote it by $\pi_1^m(X)$.

**Remark 4.9** As a group, $\pi_1^m(X)$ can be obtained as the fundamental group of a simplicial complex $S_X$ whose $n$-simplices are subsets $\{x_0, \ldots, x_n\} \subset X$ such that any distinct 3 points $x_i, x_j, x_k$ satisfy that $|\Delta(x_i, x_j, x_k)| = 0$ (see Definition 7.1).

**Remark 4.10** Our fundamental group $\pi_1^m(X)$ is not functorial with respect to Lipschitz maps. However, it is functorial with respect to Lipschitz maps that preserve collinearity ($|\Delta(x_i, x_j, x_k)| = 0$), including every embedding of metric spaces.

# 5 $\mathsf{PMet}_X^{\mathcal{G}} \simeq \mathsf{Tors}_X^{\mathcal{G}} \simeq \mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})$

In this section, we introduce the notion of "principal $\mathcal{G}$-bundles" for metric spaces. We define it from two different viewpoints, namely as a metric action and as a metric fibration, which turn out to be equivalent. As a metric action, we call it a $\mathcal{G}$-*metric action*, and as a metric fibration, we call it a $\mathcal{G}$-*torsor*. Then we show that they are classified by the conjugation classes of homomorphisms $\pi_1^m(X, x_0) \to \mathcal{G}$.

## 5.1 PMet$_X^{\mathcal{G}} \simeq$ Tors$_X^{\mathcal{G}}$

**Definition 5.1** Let $X$ be a metric space and $\mathcal{G}$ be a metric group.

(1) A $\mathcal{G}$*-metric action* $F: X \to$ Met is a metric action satisfying the following:

- For all $x \in X$, $F_x = \mathcal{G}$.
- For all $x, x' \in X$, $F_{xx'}$ is a left multiplication by some $f_{xx'} \in \mathcal{G}$.

(2) Let $F, G: X \to$ Met be $\mathcal{G}$-metric actions. A $\mathcal{G}$*-metric transformation* $\theta: F \Rightarrow G$ is a metric transformation such that each component $\theta_x: Fx \to Gx$ is a left multiplication by an element $\theta_x \in \mathcal{G}$. We denote the category of $\mathcal{G}$-metric actions $X \to$ Met and $\mathcal{G}$-metric transformations by PMet$_X^{\mathcal{G}}$.

**Remark 5.2** Obviously, PMet$_X^{\mathcal{G}}$ is a subcategory of Met$_X$ and is also a groupoid.

**Definition 5.3** Let $G$ be a group and $X$ be a metric space. We say that $X$ is a *right $G$-torsor* if $G$ acts on $X$ from the right and satisfies the following:

- It is free and transitive.
- For every $g \in G$, the map $g: X \to X$ is an isometry.
- For all $x, x' \in X$ and every $g \in G$, we have $d_X(x, xg) = d_X(x', x'g)$.

**Lemma 5.4** *Let $(X, d_X)$ be a metric space and $G$ be a group. Suppose that $X$ is a right $G$-torsor. Then there exist a distance function $d_G$ on $G$ and a metric group structure $\cdot_x$ on $X$ for each $x \in X$ such that the map*

$$G \to X, \quad g \mapsto xg,$$

*gives an isomorphism of metric groups $(G, d_G) \cong (X, \cdot_x)$. Furthermore, the unit of the metric group $(X, \cdot_x)$ is $x$.*

**Proof** Fix a point $x \in X$. We define a map $d_G: G \times G \to \mathbb{R}_{\geq 0}$ by $d_G(f, g) = d_X(xf, xg)$, which is independent from the choice of $x \in X$. It is immediate to check that $(G, d_G)$ is a metric space. Further,

$$d_G(ff', gg') = d_X(xff', xgg') \leq d_X(xff', xgf') + d_X(xgf', xgg')$$
$$\leq d_X(xf, xg) + d_X(xf', xg') = d_G(f, g) + d_G(f', g'),$$

and

$$d_G(f^{-1}, g^{-1}) = d_X(xf^{-1}, xg^{-1}) = d_X(x, xg^{-1}f)$$
$$= d_X(xg, (xg)g^{-1}f) = d_X(xg, xf)$$
$$= d_X(xf, xg) = d_G(f, g),$$

for all $f, f', g, g' \in G$. Hence $(G, d_G)$ is a metric group.

Now we define a map $G \to X$ by $g \mapsto xg$. This map is an isometry by the definition. Hence we can transfer the metric group structure on $G$ to $X$ by this map. With respect to this group structure $\cdot_x$ on $X$, we have $x \cdot_x x' = eg' = x'$ and $x' \cdot_x x = g'e = x'$, where we put $x' = xg'$. Hence $x \in X$ is the unit of the group $(X, \cdot_x)$. $\qquad\square$

**Definition 5.5** Let $G$ be a group. A metric fibration $\pi\colon E \to X$ is a $G$-*torsor over* $X$ if it satisfies the following:

- $G$ acts isometrically on $E$ from the right, and preserves each fiber of $\pi$.

- Each fiber of $\pi$ is a right $G$-torsor with respect to the above action.

**Lemma 5.6** *Let* $\pi\colon E \to X$ *be a* $G$-*torsor, and* $x, x' \in X$. *Then the metric group structures on* $G$ *induced from the fibers* $\pi^{-1}x$ *and* $\pi^{-1}x'$ *are identical.*

**Proof** For all $\varepsilon \in \pi^{-1}x$ and $f \in \mathcal{G}$, we have

$$
\begin{aligned}
d_E((\varepsilon f)_{x'}, \varepsilon_{x'} f) &= d_E(\varepsilon f, \varepsilon_{x'} f) - d_E(\varepsilon f, (\varepsilon f)_{x'}) \\
&= d_E(\varepsilon, \varepsilon_{x'}) - d_E(\varepsilon f, (\varepsilon f)_{x'}) \\
&= d_X(x, x') - d_X(x, x') \\
&= 0,
\end{aligned}
$$

and hence we obtain that $(\varepsilon f)_{x'} = \varepsilon_{x'} f$. Let $d_x$ and $d_{x'}$ be the distance function on $G$ induced from the fibers $\pi^{-1}x$ and $\pi^{-1}x'$, respectively. Explicitly, for $\varepsilon \in \pi^{-1}x$ and $f, g \in G$, we have $d_x(f, g) = d_E(\varepsilon f, \varepsilon g)$ and $d_{x'}(f, g) = d_E(\varepsilon_{x'} f, \varepsilon_{x'} g)$. Therefore we obtain that

$$
d_{x'}(f, g) = d_E(\varepsilon_{x'} f, \varepsilon_{x'} g) = d_E((\varepsilon f)_{x'}, (\varepsilon g)_{x'}) = d_E(\varepsilon f, \varepsilon g) = d_x(f, g),
$$

by Lemma 3.4. $\qquad\square$

For a $G$-torsor $\pi\colon E \to X$, we can consider the group $G$ as a metric group that is isometric to a fiber of $\pi$ by Lemma 5.4. Further, such a metric structure is independent of the choice of the fiber by Lemma 5.6. Hence, in the following, we write "$\mathcal{G}$-torsors" instead of "$G$-torsors", where $\mathcal{G}$ denotes the group $G$ equipped with this metric structure.

**Definition 5.7** Let $\pi\colon E \to X$ and $\pi'\colon E' \to X$ be $\mathcal{G}$-torsors. A $\mathcal{G}$-*morphism* $\varphi\colon \pi \to \pi'$ is a $G$-equivariant map $E \to E'$ that is also a morphism of metric fibrations. We denote the category of $\mathcal{G}$-torsors over $X$ and $\mathcal{G}$-morphisms by $\mathsf{Tors}_X^{\mathcal{G}}$.

**Remark 5.8** We can show that any $\mathcal{G}$-morphism is an isomorphism as follows: For all $\varepsilon \in E$, $x \in X$ and $g \in \mathcal{G}$, we have $d_E(\varepsilon, \varepsilon_x g) = d_X(\pi\varepsilon, x) + |g|$ by the definitions. Then the $\mathcal{G}$-equivariance of $\varphi$ and Lemma 3.5 imply that

$$
\begin{aligned}
d_{E'}(\varphi\varepsilon, \varphi(\varepsilon_x g)) &= d_{E'}(\varphi\varepsilon, (\varphi\varepsilon)_x g) \\
&= d_X(\pi'\varphi\varepsilon, x) + |g| \\
&= d_X(\pi\varepsilon, x) + |g| \\
&= d_E(\varepsilon, \varepsilon_x g),
\end{aligned}
$$

which says that $\varphi$ preserves distances. The invertibility of $\varphi$ is immediate from the $G$-equivariance.

Next we show the equivalence of $\mathcal{G}$-metric actions and $\mathcal{G}$-torsors.

**Proposition 5.9** *The Grothendieck construction determines a functor* $E\colon \mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{Tors}_X^{\mathcal{G}}$.

**Proof** Let $F\colon X \to \mathsf{Met}$ be a $\mathcal{G}$-metric action. Let $E(F)$ be the metric fibration given by the Grothendieck construction. We have $d_{E(F)}((x,g),(x',g')) = d_X(x,x') + d_{\mathcal{G}}(g_{xx'}g, g')$. We define a $\mathcal{G}$ action on $E(F)$ by $(x,g)h = (x, gh)$ for all $g, h \in \mathcal{G}$ and $x \in X$. This is obviously compatible with the projection, and also free and transitive on each fiber. We also have that

$$
\begin{aligned}
d_{E(F)}((x,g)h,(x',g')h) &= d_{E(F)}((x,gh),(x',g'h)) \\
&= d_X(x,x') + d_{\mathcal{G}}(g_{xx'}gh, g'h) \\
&= d_X(x,x') + d_{\mathcal{G}}(g_{xx'}g, g') \\
&= d_{E(F)}((x,g),(x',g')),
\end{aligned}
$$

and hence it acts isometrically. Further, we have that

$$
d_{E(F)}((x,g),(x,g)h) = d_{E(F)}((x,g),(x,gh)) = d_{\mathcal{G}}(g, gh) = d_{\mathcal{G}}(e, h),
$$

and hence each fiber is a right $\mathcal{G}$-torsor. Therefore, we obtain that $E(F)$ is a $\mathcal{G}$-torsor.

Now let $\theta\colon F \Rightarrow F'$ be a $\mathcal{G}$-metric transformation. The Grothendieck construction gives a map $\varphi_\theta\colon E(F) \to E(F')$ by $\varphi_\theta(x,g) = (x, \theta_x g)$, which is a morphism of metric fibrations. It is checked that $\varphi_\theta$ is $\mathcal{G}$-equivariant as follows:

$$
(\varphi_\theta(x,g))h = (x, \theta_x gh) = \varphi_\theta(x, gh).
$$

Hence it is a $\mathcal{G}$-morphism. □

**Proposition 5.10** *We have a functor* $F\colon \mathsf{Tors}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ *sending a $\mathcal{G}$-torsor $\pi$ to a $\mathcal{G}$-metric action $F_\pi$ with $F_\pi x = \pi^{-1}x$.*

**Proof** Let $\pi\colon E \to X$ be a $\mathcal{G}$-torsor. We fix points $x_0 \in X$ and $\varepsilon \in \pi^{-1}x_0$. For each $x \in X$, we equip the set $\pi^{-1}x$ with a metric group structure isomorphic to $\mathcal{G}$ with the unit $\varepsilon_x$ by Lemma 5.4. Hence we can identify each fiber with $\mathcal{G}$ by the map $g \mapsto \varepsilon_x g$. Now we put $(\varepsilon_x)_{x'} = \varepsilon_{x'} g_{xx'} \in \pi^{-1}x'$ for $x, x' \in X$ and $g_{xx'} \in \mathcal{G}$. Then, for each $h \in \mathcal{G}$, we have

$$
\begin{aligned}
d_X(x,x') &= d_E(\varepsilon_x h, (\varepsilon_x h)_{x'}) \\
&= d_E(\varepsilon_x, (\varepsilon_x h)_{x'} h^{-1}) \\
&= d_E(\varepsilon_x, \varepsilon_{x'} g_{xx'}) + d_E(\varepsilon_{x'} g_{xx'}, (\varepsilon_x h)_{x'} h^{-1}) \\
&= d_X(x,x') + d_E(\varepsilon_{x'} g_{xx'}, (\varepsilon_x h)_{x'} h^{-1}),
\end{aligned}
$$

and hence we obtain that $(\varepsilon_x h)_{x'} = \varepsilon_{x'} g_{xx'} h$. This implies that the map $\pi^{-1}x \to \pi^{-1}x'$ given by lifts $\varepsilon_x h \mapsto (\varepsilon_x h)_{x'}$ is the left multiplication by $g_{xx'}$ when we identify each fiber with $\mathcal{G}$ as above. Hence the functor $F$ gives a $\mathcal{G}$-metric action.

Next, let $\varphi\colon \pi \to \pi'$ be a $\mathcal{G}$-morphism between $\mathcal{G}$-torsors $\pi\colon E \to X$ and $\pi'\colon E' \to X$. It induces a Lipschitz map $\varphi_x\colon \pi^{-1}x \to \pi'^{-1}x$. Since fibers $\pi^{-1}x$ and $\pi'^{-1}x$ are identified with $\mathcal{G}$ and $\varphi_x$ is $\mathcal{G}$-equivariant, we can identify $\varphi_x$ with the left multiplication by $\varphi_x\varepsilon_x$. This implies that the functor $F$ sends the $\mathcal{G}$-morphism $\varphi$ to a $\mathcal{G}$-metric transformation between $F_\pi$ and $F_{\pi'}$. $\qquad\square$

**Proposition 5.11** *The Grothendieck construction functor* $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{Tors}_X^{\mathcal{G}}$ *is an equivalence of categories.*

**Proof** This is similar to the proof of Proposition 3.9. $\qquad\square$

## 5.2 $\mathsf{PMet}_X^{\mathcal{G}} \simeq \mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})$

First we define the category of homomorphisms of metric groups $\mathcal{G} \to \mathcal{G}'$.

**Definition 5.12** Let $\mathcal{G}$ and $\mathcal{G}'$ be metric groups, and let $\mathsf{Hom}(\mathcal{G}, \mathcal{G}')$ be the set of all homomorphisms $\mathcal{G} \to \mathcal{G}'$. We equip $\mathsf{Hom}(\mathcal{G}, \mathcal{G}')$ with a groupoid structure by defining $\mathsf{Hom}(\mathcal{G}, \mathcal{G}')(\varphi, \psi) = \{h \in \mathcal{G}' \mid \varphi = h^{-1}\psi h\}$ for homomorphisms $\varphi, \psi\colon \mathcal{G} \to \mathcal{G}'$. The identity on $\varphi \in \mathsf{Hom}(\mathcal{G}, \mathcal{G}')$ is the unit $e \in \mathcal{G}'$, and the composition of morphisms $h \in \mathsf{Hom}(\mathcal{G}, \mathcal{G}')(\varphi, \psi)$ and $h' \in \mathsf{Hom}(\mathcal{G}, \mathcal{G}')(\psi, \xi)$ is defined by $h' \circ h = h'h$.

**Lemma 5.13** *Let $X$ be a metric space and $\mathcal{G}$ be a metric group. For each $x_0 \in X$, we have a functor $A\colon \mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G}) \to \mathsf{PMet}_X^{\mathcal{G}}$ sending a homomorphism $\varphi\colon \pi_1^m(X, x_0) \to \mathcal{G}$ to a $\mathcal{G}$-metric action $F_\varphi$ with $F_\varphi x = \mathcal{G}$.*

**Proof** Let $\varphi\colon \pi_1^m(X, x_0) \to \mathcal{G}$ be a homomorphism. We define a $\mathcal{G}$-metric action $F_\varphi\colon X \to \mathsf{Met}$ by $F_\varphi x = \mathcal{G}$ and $(F_\varphi)_{xx'} = \varphi[(x_0, x', x, x_0)]\cdot\colon \mathcal{G} \to \mathcal{G}$, where we denote the left multiplication by $(-)\cdot$. It is verified that this defines a $\mathcal{G}$-metric action as follows.

For all $x, x' \in X$, we have $(F_\varphi)_{xx} = \varphi[(x_0, x, x, x_0)]\cdot = e \cdot = \mathrm{id}_{\mathcal{G}}$, and $(F_\varphi)_{x'x} = \varphi[(x_0, x, x', x_0)]\cdot = (\varphi[(x_0, x', x, x_0)])^{-1} \cdot = (F_\varphi)_{xx'}^{-1}$. Further, we have

$$
\begin{aligned}
d_{\mathcal{G}}((F_\varphi)_{x'x''}(F_\varphi)_{xx'}g, (F_\varphi)_{xx''}g) &= d_{\mathcal{G}}\big(\varphi[(x_0, x'', x', x_0)]\varphi[(x_0, x', x, x_0)], \varphi[(x_0, x'', x, x_0)]\big) \\
&= d_{\mathcal{G}}\big(\varphi[(x_0, x'', x', x, x_0)], \varphi[(x_0, x'', x, x_0)]\big) \\
&= d_{\mathcal{G}}\big((\varphi[(x_0, x'', x, x_0)])^{-1}\varphi[(x_0, x'', x', x, x_0)], e\big) \\
&= d_{\mathcal{G}}\big(\varphi[(x_0, x, x'', x', x, x_0)], e\big) \\
&\leq d_{\pi_1^m(X, x_0)}([(x_0, x, x'', x', x, x_0)], [x_0, x_0]) \\
&\leq d_X(x, x') + d_X(x', x'') - d_X(x, x''),
\end{aligned}
$$

for all $x, x', x'' \in X$ and $g \in \mathcal{G}$. Let $h\colon \varphi \to \psi$ be a morphism in $\mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})$; that is, we have $\varphi = h^{-1}\psi h$ with $h \in \mathcal{G}$. Then we can construct a $\mathcal{G}$-metric transformation $\theta\colon F_\varphi \Rightarrow F_\psi$ by $\theta_x = h \cdot\colon \mathcal{G} \to \mathcal{G}$. It satisfies that $(F_\psi)_{xx'}\theta_x = \theta_{x'}(F_\varphi)_{xx'}$ since we have $\psi[(x_0, x', x, x_0)]h = h\varphi[(x_0, x', x, x_0)]$. $\qquad\square$

**Lemma 5.14** *Let $X$ be a metric space and $\mathcal{G}$ be a metric group. For each $x_0 \in X$, there is a functor $B \colon \mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})$ sending a $\mathcal{G}$-metric action $F$ to a homomorphism $\varphi_F \colon \pi_1^m(X, x_0) \to \mathcal{G}$ defined by*

$$\varphi_F[(x_0, x_1, \ldots, x_n, x_0)] = F_{x_1 x_0} F_{x_2 x_1} \ldots F_{x_n x_{n-1}} F_{x_0 x_n},$$

*for each $[(x_0, x_1, \ldots, x_n, x_0)] \in \pi_1^m(X, x_0)$.*

**Proof** It is immediate to check that this is well defined. Let $F, F' \colon X \to \mathsf{Met}$ be $\mathcal{G}$-metric actions and $\theta \colon F \Rightarrow F'$ be a $\mathcal{G}$-metric transformation. Then we have

$$
\begin{aligned}
\theta_{x_0}^{-1} \varphi_{F'}[(x_0, x_1, \ldots, x_n, x_0)]\theta_{x_0} &= \theta_{x_0}^{-1} F'_{x_1 x_0} F'_{x_2 x_1} \ldots F'_{x_n x_{n-1}} F'_{x_0 x_n} \theta_{x_0} \\
&= F_{x_1 x_0} F_{x_2 x_1} \ldots F_{x_n x_{n-1}} F_{x_0 x_n} \\
&= \varphi_F[(x_0, x_1, \ldots, x_n, x_0)],
\end{aligned}
$$

for each $[(x_0, x_1, \ldots, x_n, x_0)] \in \pi_1^m(X, x_0)$. Hence $\theta_{x_0} \in \mathcal{G}$ gives a morphism $\theta_{x_0} \colon \varphi_F \to \varphi_{F'}$. This correspondence is obviously functorial. $\qquad\square$

**Proposition 5.15** *The functor $A \colon \mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G}) \to \mathsf{PMet}_X^{\mathcal{G}}$ of Lemma 5.13 is an equivalence of categories.*

**Proof** We show the natural isomorphisms $BA \cong \mathrm{id}_{\mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})}$ and $AB \cong \mathrm{id}_{\mathsf{PMet}_X^{\mathcal{G}}}$. For a homomorphism $\varphi \colon \pi_1^m(X, x_0) \to \mathcal{G}$, we have

$$
\begin{aligned}
\varphi_{F_\varphi}[(x_0, x_1, \ldots, x_n, x_0)] &= (F_\varphi)_{x_1 x_0} (F_\varphi)_{x_2 x_1} \ldots (F_\varphi)_{x_0 x_n} \\
&= \varphi[(x_0, x_0, x_1, x_0)]\varphi[(x_0, x_1, x_2, x_0)] \ldots \varphi[(x_0, x_n, x_0, x_0)] \\
&= \varphi[(x_0, x_0, x_1, x_1, \ldots, x_n, x_n, x_0, x_0)] \\
&= \varphi[(x_0, x_1, \ldots, x_n, x_0)],
\end{aligned}
$$

for each $[(x_0, x_1, \ldots, x_n, x_0)] \in \pi_1^m(X, x_0)$. Hence we obtain an isomorphism $BA\varphi = \varphi$ that is obviously natural. Conversely, let $F \colon X \to \mathsf{Met}$ be a $\mathcal{G}$-metric action. Then we have

$$(F_{\varphi_F})_x = \mathcal{G} \quad \text{and} \quad (F_{\varphi_F})_{x x'} = \varphi_F[(x_0, x', x, x_0)] = F_{x' x_0} F_{x x'} F_{x_0 x}.$$

Now we define a $\mathcal{G}$-metric transformation $\theta \colon F_{\varphi_F} \Rightarrow F$ by $\theta_x = F_{x_0 x}$. It is obvious that we have $F_{x x'}\theta_x = \theta_{x'}(F_{\varphi_F})_{x x'}$, and hence it is well defined and obviously an isomorphism. For a $\mathcal{G}$-metric transformation $\tau \colon F \Rightarrow F'$, we have $(AB\tau)_x = \tau_{x_0} \cdot \colon (F_{\varphi_F})_x \to (F'_{\varphi_{F'}})_x$ by the construction. Hence the condition $\tau_x F_{x_0 x} = F'_{x_0 x}\tau_{x_0}$ of the $\mathcal{G}$-metric transformation implies the naturality of this isomorphism. $\qquad\square$

## 5.3 Example

We give the following example of a metric fundamental group.

**Proposition 5.16** *Let $C_n$ be the undirected $n$-cycle graph. Then we have*

$$\pi_1^m(C_n) \cong \begin{cases} \mathbb{Z} \text{ with } |1| = 1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

*Hence we have that*

$$\mathsf{PMet}_{C_n}^{\mathcal{G}} \simeq \begin{cases} \mathrm{Hom}(\mathbb{Z}, \mathcal{G}) & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even,} \end{cases}$$

*for all metric groups $\mathcal{G}$, which implies that there is only a trivial metric fibration over $C_{2n}$ and that there is at most one nontrivial metric fibration over $C_{2n+1}$.*

**Proof** Let $V(C_n) = \{v_1, \dots, v_n\}$ be the vertex set whose numbering is anticlockwise. For $C_{2n}$, it reduces to show that $[(v_1, v_2, \dots, v_{2n}, v_1)] = [(v_1, v_1)]$. Since we have $d_{C_{2n}}(v_i, v_j) = d_{C_{2n}}(v_i, v_k) + d_{C_{2n}}(v_k, v_j)$ for all $i \le k \le j$ with $j - i \le n$, we obtain that

$$\begin{aligned} [(v_1, v_2, \dots, v_{2n}, v_1)] &= [(v_1, \dots, v_{n+1}, \dots, v_{2n}, v_1)] \\ &= [(v_1, v_{n+1}, v_1)] \\ &= [(v_1, v_1)]. \end{aligned}$$

For $C_{2n+1}$, the possible nontrivial element of $\pi_1^m(C_{2n+1})$ is a concatenation or its inverse of the element $[(v_1, \dots, v_{2n+1}, v_1)]$. Now we have $[(v_1, \dots, v_{2n+1}, v_1)] = [(v_1, v_{n+1}, v_{n+2}, v_1)]$, by the same argument as above, and

$$\begin{aligned} &d_{Q(C_{2n+1}, v_1)}((v_1, v_{n+1}, v_{n+2}, v_1), (v_1, v_{n+1}, v_1)) \\ &= d_{C_{2n+1}}(v_{n+1}, v_{n+2}) + d_{C_{2n+1}}(v_{n+2}, v_1) - d_{C_{2n+1}}(v_{n+1}, v_1) \\ &= d_{C_{2n+1}}(v_{n+1}, v_{n+2}) \\ &= 1. \end{aligned}$$

Hence we obtain that $|[(v_1, \dots, v_{2n+1}, v_1)]| = 1$. $\square$

**Remark 5.17** The cycle graph $C_n$ is a metric group $\mathbb{Z}/n\mathbb{Z}$ with $|1| = 1$. Hence the examples in Figure 1 are $\mathbb{Z}/2\mathbb{Z}$-torsors, which are classified by $\mathrm{Hom}(\mathbb{Z}, \mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$.

# 6 Classification of metric fibrations

In this section, we classify general metric fibrations by fixing the base and the fiber. It is analogous to that of topological fiber bundles, namely it reduces to classifying principal bundles whose fiber is the structure group of the concerned fibration. We divide it into two cases, whether the fiber is bounded or not, since we need to consider *expanded metric spaces* for the unbounded case.

## 6.1 The functor $\widehat{(-)}^{x_0}$

Before we show the classification, we introduce a technical functor that will be used later.

**Definition 6.1**  For all metric actions $F: X \to \mathsf{Met}$ and a point $x_0 \in X$, we define a metric action $\widehat{F}^{x_0}: X \to \mathsf{Met}$ as follows. We define that $\widehat{F}^{x_0}x = Fx_0$ and $\widehat{F}^{x_0}_{xx'} = F_{x'x_0}F_{xx'}F_{x_0x}: Fx_0 \to Fx_0$ for all $x, x' \in X$. We verify that this defines a metric action as follows. We have $\widehat{F}^{x_0}_{xx} = F_{xx_0}F_{xx}F_{x_0x} = \mathrm{id}_{Fx_0} = \mathrm{id}_{\widehat{F}^{x_0}x}$. We also have $(\widehat{F}^{x_0}_{x'x})^{-1} = (F_{xx_0}F_{x'x}F_{x_0x'})^{-1} = F_{x'x_0}F_{xx'}F_{x_0x} = \widehat{F}^{x_0}_{xx'}$ and

$$
\begin{aligned}
d_{\widehat{F}^{x_0}x''}(\widehat{F}^{x_0}_{x'x''}\widehat{F}^{x_0}_{xx'}a, \widehat{F}^{x_0}_{xx''}a) &= d_{Fx_0}(F_{x''x_0}F_{x'x''}F_{x_0x'}F_{x'x_0}F_{xx'}F_{x_0x}a, F_{x''x_0}F_{xx''}F_{x_0x}a) \\
&= d_{Fx''}(F_{x'x''}F_{xx'}F_{x_0x}a, F_{xx''}F_{x_0x}a) \\
&\le d_X(x, x') + d_X(x', x'') - d_X(x, x''),
\end{aligned}
$$

for all $x, x', x'' \in X$ and $a \in \widehat{F}^{x_0}x$.

**Lemma 6.2**  *The correspondence $F \mapsto \widehat{F}^{x_0}$ defines a fully faithful functor $\widehat{(-)}^{x_0}: \mathsf{Met}_X \to \mathsf{Met}_X$. Further, this restricts to a fully faithful functor $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ for each metric group $\mathcal{G}$.*

**Proof**  Let $\theta: F \Rightarrow G$ be a metric transformation. We define a metric transformation $\widehat{\theta}^{x_0}: \widehat{F}^{x_0} \Rightarrow \widehat{G}^{x_0}$ by $\widehat{\theta}^{x_0}_x = \theta_{x_0}: \widehat{F}^{x_0}x \to \widehat{G}^{x_0}x, a \mapsto \theta_{x_0}a$. Then we have

$$
\begin{aligned}
\widehat{G}^{x_0}_{xx'}\widehat{\theta}^{x_0}_x &= G_{x'x_0}G_{xx'}G_{x_0x}\theta_{x_0} \\
&= G_{x'x_0}G_{xx'}\theta_x F_{x_0x} \\
&= G_{x'x_0}\theta_{x'}F_{xx'}F_{x_0x} \\
&= \theta_{x_0}F_{x'x_0}F_{xx'}F_{x_0x} \\
&= \widehat{\theta}^{x_0}_x \widehat{F}^{x_0}_{xx'},
\end{aligned}
$$

and hence this defines a metric transformation. It is obvious that $\widehat{\mathrm{id}_F}^{x_0} = \mathrm{id}_{\widehat{F}^{x_0}}$ and $\widehat{(\theta'\theta)}^{x_0} = \widehat{\theta}'^{x_0}\widehat{\theta}^{x_0}$. It is a faithful functor because $G_{xx_0}\theta_x = \theta_{x_0}F_{xx_0}$ implies that $\theta_x = \theta'_x$ for all $x \in X$ if two metric transformation $\theta, \theta'$ satisfy $\theta_{x_0} = \theta'_{x_0}$. By definition, it restricts to a faithful functor $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ for each metric group $\mathcal{G}$. Next we show the fullness. Let $\eta: \widehat{F}^{x_0} \Rightarrow \widehat{G}^{x_0}$ be a metric transformation. Then we have $\widehat{G}^{x_0}_{x_0x}\eta_{x_0} = \eta_x \widehat{F}^{x_0}_{x_0x}$ and $\widehat{F}^{x_0}_{x_0x} = \mathrm{id}_{Fx_0}, \widehat{G}^{x_0}_{x_0x} = \mathrm{id}_{Gx_0}$. Hence we obtain that $\eta_{x_0} = \eta_x$ for all $x \in X$.

Now we define a metric transformation $\widetilde{\eta}: F \Rightarrow G$ by $\widetilde{\eta}_x = G_{x_0x}\eta_{x_0}F_{xx_0}: Fx \to Gx$. Then we have

$$
\begin{aligned}
G_{xx'}\widetilde{\eta}_x &= G_{xx'}G_{x_0x}\eta_{x_0}F_{xx_0} \\
&= G_{x_0x'}\widehat{G}^{x_0}_{xx'}\eta_x F_{xx_0} \\
&= G_{x_0x'}\eta_{x'}\widehat{F}^{x_0}_{xx'}F_{xx_0} \\
&= G_{x_0x'}\eta_{x'}F_{x'x_0}F_{xx'}F_{x_0x}F_{xx_0} \\
&= G_{x_0x'}\eta_{x_0}F_{x'x_0}F_{xx'} \\
&= \widetilde{\eta}_{x'}F_{xx'},
\end{aligned}
$$

and hence this defines a metric transformation. We obviously have $\widehat{(\widetilde{\eta})}^{x_0} = \eta$, which implies that the functor $\widehat{(-)}^{x_0}$ is full. The restriction to $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ is immediate. $\qquad\square$

**Lemma 6.3** *The functor* $\widehat{(-)}^{x_0} \colon \mathsf{Met}_X \to \mathsf{Met}_X$ *is split essentially surjective. Its restriction* $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ *is also split essentially surjective for all metric groups* $\mathcal{G}$.

**Proof** Let $F \colon X \to \mathsf{Met}$ be a metric action. We define a metric transformation $\theta \colon \widehat{F}^{x_0} \Rightarrow F$ by $\theta_x = F_{x_0 x} \colon \widehat{F}^{x_0} x \to Fx, a \mapsto F_{x_0 x} a$. It satisfies that

$$
\begin{aligned}
F_{xx'} \theta_x &= F_{xx'} F_{x_0 x} \\
&= F_{x_0 x'} F_{x' x_0} F_{xx'} F_{x_0 x} \\
&= \theta_{x'} \widehat{F}_{xx'}^{x_0}.
\end{aligned}
$$

Further, we define a metric transformation $\theta^{-1} \colon F \Rightarrow \widehat{F}^{x_0}$ by $\theta_x^{-1} = F_{xx_0} \colon Fx \to \widehat{F}^{x_0} x$ for all $x \in X$. Then we have $\widehat{F}_{xx'}^{x_0} \theta_x^{-1} = \theta_x^{-1} F_{xx'}$ similarly to the above, and hence it defines a metric transformation. It is obviously an isomorphism. The restriction to $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ is immediate. $\qquad\square$

**Corollary 6.4** *The functor* $\widehat{(-)}^{x_0} \colon \mathsf{Met}_X \to \mathsf{Met}_X$ *and its restriction* $\mathsf{PMet}_X^{\mathcal{G}} \to \mathsf{PMet}_X^{\mathcal{G}}$ *for all metric groups* $\mathcal{G}$ *are category equivalences.* $\qquad\square$

**Definition 6.5** (1) We denote the image of the functor $\widehat{(-)}^{x_0} \colon \mathsf{Met}_X \to \mathsf{Met}_X$ by $\widehat{\mathsf{Met}}_X^{x_0}$.

(2) For each metric space $Y$, we denote by $\mathsf{Met}_X^Y$ the full subcategory of $\mathsf{Met}_X$ that consists of metric actions $F \colon X \to \mathsf{Met}$ such that $Fx \cong Y$ for all $x \in X$.

(3) We denote the image of $\widehat{(-)}^{x_0}$ restricted to $\mathsf{Met}_X^Y$ and $\mathsf{PMet}_X^{\mathcal{G}}$ by $\widehat{\mathsf{Met}}_X^{Y, x_0}$ and $\widehat{\mathsf{PMet}}_X^{\mathcal{G}, x_0}$, respectively.

(4) For each metric space $Y$, we denote by $\mathsf{Fib}_X^Y$ the full subcategory of $\mathsf{Fib}_X$ that consists of metric fibrations $\pi \colon E \to X$ such that $\pi^{-1} x \cong Y$ for all $x \in X$.

**Lemma 6.6** (1) *We have category equivalences* $\mathsf{Met}_X^Y \to \widehat{\mathsf{Met}}_X^{Y, x_0}$ *and* $\mathsf{PMet}_X^{\mathcal{G}} \to \widehat{\mathsf{PMet}}_X^{\mathcal{G}, x_0}$.

(2) *The Grothendieck construction functor* $E \colon \mathsf{Met}_X \to \mathsf{Fib}_X$ *restricts to the category equivalence* $\mathsf{Met}_X^Y \to \mathsf{Fib}_X^Y$.

**Proof** (1) follows from Corollary 6.4, and (2) follows from the proof of Proposition 3.9. $\qquad\square$

## 6.2 Classification for the case of bounded fibers

In this subsection, we suppose that $X$ and $Y$ are metric spaces and $Y$ is bounded. In this case the group $\mathrm{Aut}\, Y$ of automorphisms is a metric group (Example 4.3).

**Lemma 6.7** *We have a faithful functor*

$$
- \curvearrowright Y \colon \mathsf{PMet}_X^{\mathrm{Aut}\, Y} \to \mathsf{Met}_X^Y.
$$

**Proof** Let $F \in \mathsf{PMet}_X^{\mathsf{Aut}\,Y}$. We define a metric action $F \curvearrowright Y \colon X \to \mathsf{Met}$ by $(F \curvearrowright Y)x = Y$ and $(F \curvearrowright Y)_{xx'} = F_{xx'} \colon Y \to Y$. It is immediate to verify that this defines a metric action. For an $\mathsf{Aut}\,Y$-metric transformation $\theta \colon F \Rightarrow G$, we define a metric transformation $\theta \curvearrowright Y \colon F \curvearrowright Y \Rightarrow G \curvearrowright Y$ by $(\theta \curvearrowright Y)_x = \theta_x \colon Y \to Y, y \mapsto \theta_x y$. Then it is also immediate to verify that it is a metric transformation. Further, this obviously defines a faithful functor. $\qquad\square$

**Lemma 6.8** *The functor* $- \curvearrowright Y \colon \mathsf{PMet}_X^{\mathsf{Aut}\,Y} \to \mathsf{Met}_X^Y$ *is split essentially surjective.*

**Proof** Let $F \in \mathsf{Met}_X^Y$ and fix isometries $\varphi_x \colon Y \to Fx$ using the axiom of choice. We define an $\mathsf{Aut}\,Y$-metric action $\mathsf{Aut}\,F$ by $(\mathsf{Aut}\,F)x = \mathsf{Aut}\,Y$ and $(\mathsf{Aut}\,F)_{xx'} = \varphi_{x'}^{-1} F_{xx'} \varphi_x \cdot$ that is a left multiplication. We can verify that it is an $\mathsf{Aut}\,Y$-metric action as follows. We have $(\mathsf{Aut}\,F)_{xx} = \varphi_x^{-1} F_{xx} \varphi_x \cdot = \mathrm{id}_{\mathsf{Aut}\,Y}$ and $(\mathsf{Aut}\,F)_{xx'}^{-1} = \varphi_x^{-1} F_{x'x} \varphi_{x'} \cdot = (\mathsf{Aut}\,F)_{x'x}$. We also have that

$$
\begin{aligned}
d_{\mathsf{Aut}\,Y}((\mathsf{Aut}\,F)_{x'x''}(\mathsf{Aut}\,F)_{xx'}, (\mathsf{Aut}\,F)_{xx''}) &= d_{\mathsf{Aut}\,Y}(\varphi_{x''}^{-1} F_{x'x''} \varphi_{x'} \varphi_{x'}^{-1} F_{xx'} \varphi_x, \varphi_{x''}^{-1} F_{xx''} \varphi_x) \\
&= d_{\mathsf{Aut}\,Y}(\varphi_{x''}^{-1} F_{x'x''} F_{xx'} \varphi_x, \varphi_{x''}^{-1} F_{xx''} \varphi_x) \\
&= \sup_{a \in Y} d_Y(\varphi_{x''}^{-1} F_{x'x''} F_{xx'} \varphi_x a, \varphi_{x''}^{-1} F_{xx''} \varphi_x a) \\
&= \sup_{a \in Fx} d_{Fx''}(F_{x'x''} F_{xx'} a, F_{xx''} a) \\
&\le d_X(x, x') + d_X(x', x'') - d_X(x, x'').
\end{aligned}
$$

Now we define a metric transformation $\varphi \colon \mathsf{Aut}\,F \curvearrowright Y \Rightarrow F$ by $\varphi_x \colon (\mathsf{Aut}\,F \curvearrowright Y)x = Y \to Fx$. This satisfies that $F_{xx'} \varphi_x = \varphi_{x'}(\mathsf{Aut}\,F \curvearrowright Y)_{xx'}$ and is an isomorphism by the definition. $\qquad\square$

Since the category $\mathsf{PMet}_X^{\mathsf{Aut}\,Y}$ is a groupoid, the image of the functor $- \curvearrowright Y$ is in $\mathrm{core}\,\mathsf{Met}_X^Y$. (Here core denotes the subcategory consisting of all isomorphisms, as in Definition 2.1(4).)

**Lemma 6.9** *The functor* $\widehat{- \curvearrowright Y}^{x_0} \colon \mathsf{PMet}_X^{\mathsf{Aut}\,Y} \to \mathrm{core}\widehat{\mathsf{Met}}_X^{Y,x_0}$ *is full.*

**Proof** We have $\widehat{- \curvearrowright Y}^{x_0} = \widehat{(-)}^{x_0} \curvearrowright Y$ by the definitions. Since $\widehat{(-)}^{x_0} \colon \mathsf{PMet}_X^{\mathsf{Aut}\,Y} \to \mathsf{PMet}_X^{\mathsf{Aut}\,Y}$ is full by Lemma 6.2, it will suffice to show that the restriction

$$
- \curvearrowright Y \colon \widehat{\mathsf{PMet}}_X^{\mathsf{Aut}\,Y,x_0} \to \mathrm{core}\widehat{\mathsf{Met}}_X^{Y,x_0}
$$

is full. Let $\theta \colon \widehat{F}^{x_0} \curvearrowright Y \Rightarrow \widehat{G}^{x_0} \curvearrowright Y$ be an isomorphism in $\widehat{\mathsf{Met}}_X^{Y,x_0}$, where $F, G \in \mathsf{PMet}_X^{\mathsf{Aut}\,Y}$. Then we have an isometry $\theta_x \colon Y \to Y$ such that $G_{x'x_0} G_{xx'} G_{x_0x} \theta_x = \theta_{x'} F_{x'x_0} F_{xx'} F_{x_0x}$ for all $x, x' \in X$. Since we have $\theta_x \in \mathsf{Aut}\,Y$, we obtain a morphism

$$
\theta' \colon \widehat{F}^{x_0} \Rightarrow \widehat{G}^{x_0} \in \widehat{\mathsf{PMet}}_X^{\mathsf{Aut}\,Y,x_0}
$$

defined by $\theta'_x = \theta_x$. It is obvious that we have $\theta' \curvearrowright Y = \theta$. $\qquad\square$

**Corollary 6.10** *The functor* $\widetilde{-\curvearrowright Y}^{x_0}\colon \mathsf{PMet}_X^{\mathsf{Aut}\,Y} \to \mathsf{core}\widehat{\mathsf{Met}}_X^{Y,x_0}$ *is an equivalence of categories.* □

**Corollary 6.11** *The categories* $\mathsf{PMet}_X^{\mathsf{Aut}\,Y}$ *and* $\mathsf{core}\,\mathsf{Fib}_X^Y$ *are equivalent.*

**Proof** This follows from Corollary 6.10, with $\mathsf{core}\,\mathsf{Fib}_X^Y \simeq \mathsf{core}\,\mathsf{Met}_X^Y \simeq \mathsf{core}\widehat{\mathsf{Met}}_X^{Y,x_0}$ by Lemma 6.6. □

## 6.3 Classification for the case of unbounded fibers

To classify general metric fibrations, we generalize the discussions so far to *extended metric groups*.

**Definition 6.12** (1) An *extended metric group* is a monoid object in $\mathsf{EMet}$ that is a group when we forget the metric structure.

(2) For extended metric groups $\mathcal{G}$ and $\mathcal{H}$, a *homomorphism* from $\mathcal{G}$ to $\mathcal{H}$ is a Lipschitz map $\mathcal{G} \to \mathcal{H}$ that commutes with the group structure.

(3) We denote the category of extended metric groups and homomorphisms by $\mathsf{EMGrp}$. The category $\mathsf{MGrp}$ is a full subcategory of $\mathsf{EMGrp}$.

**Example 6.13** Let $(X, d)$ be a metric space, and let $\mathsf{Aut}\,X$ be the group of isometries on $X$. We define a distance function on $\mathsf{Aut}\,X$ by $d_{\mathsf{Aut}\,X}(f, g) = \sup_{x \in X} d_X(fx, gx)$. Then it is immediate to verify that $(\mathsf{Aut}\,X, d_{\mathsf{Aut}\,X})$ is an extended metric group. The "unit component" of $\mathsf{Aut}\,X$, that is the set of isometries $f$ such that $d_{\mathsf{Aut}\,X}(\mathrm{id}_X, f) < \infty$, is exactly $\mathsf{Aut}^u\,X$ (Example 4.3). If the metric space $X$ has finite diameter, then we have $\mathsf{Aut}\,X = \mathsf{Aut}^u\,X$.

**Definition 6.14** Let $\mathcal{G}$ and $\mathcal{G}'$ be extended metric groups, and let $\mathsf{Hom}(\mathcal{G}, \mathcal{G}')$ be the set of homomorphisms. We equip $\mathsf{Hom}(\mathcal{G}, \mathcal{G}')$ with a groupoid structure similarly to the metric group case by defining $\mathsf{Hom}(\mathcal{G}, \mathcal{G}')(\varphi, \psi) = \{h \in \mathcal{G}' \mid \varphi = h^{-1}\psi h\}$ for all homomorphisms $\varphi, \psi \colon \mathcal{G} \to \mathcal{G}'$.

**Remark 6.15** A statement analogous to Lemma 4.2 holds for extended metric groups. Further, the relationship between extended metric spaces and normed groups similar to Lemma 4.2 holds if we replace the codomain of norms by $[0, \infty]$.

**Definition 6.16** Let $\mathcal{G}$ be an extended metric group and $X$ be a metric space. An *extended $\mathcal{G}$-metric action $F$* is a correspondence $X \ni x \mapsto Fx = \mathcal{G}$ and $F_{xx'} \in \mathcal{G}$ such that

- $F_{xx} = e$, $F_{xx'} = F_{x'x}^{-1}$,
- $d_{\mathcal{G}}(F_{x'x''}F_{xx'}, F_{xx''}) \le d_X(x, x') + d_X(x', x'') - d_X(x, x'')$.

For extended $\mathcal{G}$-metric actions $F$ and $G$, an *extended $\mathcal{G}$-metric transformation* $\theta\colon F \Rightarrow G$ is a family of elements $\{\theta_x \in \mathcal{G}\}_{x \in X}$ such that $G_{xx'}\theta_x = \theta_{x'}F_{xx'}$. We denote the category of extended $\mathcal{G}$-metric actions and extended $\mathcal{G}$-metric transformations by $\mathsf{EPMet}_X^{\mathcal{G}}$.

The following is obtained from the same arguments in Section 5.2 by replacing the term "metric group" by "extended metric group".

**Proposition 6.17** *For an extended metric group $\mathcal{G}$ and a metric space $X$, the categories* $\mathsf{EPMet}_X^{\mathcal{G}}$ *and* $\mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})$ *are equivalent.* □

Further, the arguments in Section 6.2 can be applied to the extended case, and we obtain the following.

**Proposition 6.18** *For all metric spaces $X$ and $Y$, the categories* $\mathsf{EPMet}_X^{\mathsf{Aut}\,Y}$ *and* $\mathsf{core}\,\mathsf{Fib}_X^Y$ *are equivalent. Hence metric fibrations with fiber $Y$ are classified by* $\mathsf{Hom}(\pi_1^m(X, x_0), \mathcal{G})$. □

# 7 Cohomological interpretation

In this section, we give a cohomological classification of $\mathcal{G}$-torsors. It is an analogy of the 1-Čech cohomology. Before giving the definition, we introduce the following technical term.

**Definition 7.1** Let $X$ be a metric space, and $x_1, x_2, x_3 \in X$. We denote the subset $\{x_1, x_2, x_3\} \subset X$ by $\Delta(x_1, x_2, x_3)$ and call it a *triangle*. We define the *degeneracy degree of the triangle* $\Delta(x_1, x_2, x_3)$ by

$$|\Delta(x_1, x_2, x_3)| := \min\{d_X(x_i, x_j) + d_X(x_j, x_k) - d_X(x_i, x_k) \mid \{i, j, k\} = \{1, 2, 3\}\}.$$

It is enough to consider $i, j, k$'s running in the cyclic order to obtain the above minimum.

The following is the definition of our "1-Čech cohomology".

**Definition 7.2** Let $X$ be a metric space and suppose that points of $X$ are indexed as $X = \{x_i\}_{i \in I}$. For a metric group $\mathcal{G}$, we define the 1-*cohomology of $X$ with coefficients in $\mathcal{G}$* as the category $\mathsf{H}^1(X; \mathcal{G})$ by

$$\mathsf{Ob}\,\mathsf{H}^1(X; \mathcal{G}) = \{(a_{ijk}) \in \mathcal{G}^{I^3} \mid a_{ijk}a_{kj\ell} = a_{ij\ell},\ |a_{ijk}a_{jki}a_{kij}| \le |\Delta(x_i, x_j, x_k)|\},$$

and

$$\mathsf{H}^1(X; \mathcal{G})((a_{ijk}), (b_{ijk})) = \{(f_{ij}) \in \mathcal{G}^{I^2} \mid a_{ijk}f_{jk} = f_{ij}b_{ijk}\},$$

where we denote the conjugation invariant norm on $\mathcal{G}$ by $|-|$. We call an object of $\mathsf{H}^1(X; \mathcal{G})$ *a cocycle*. Obviously, the above constructions are independent from the choice of the index $I$.

**Remark 7.3** For a cocycle $(a_{ijk}) \in \mathsf{H}^1(X; \mathcal{G})$, the condition $a_{ijk}a_{kj\ell} = a_{ij\ell}$ implies that $a_{iji} = e$ and $a_{ijk} = a_{kji}^{-1}$ for all $i, j, k \in I$. Further, for a morphism $(f_{ij})$, we have $f_{ij} = f_{ji}$ from the condition $a_{ijk}f_{jk} = f_{ij}b_{ijk}$ and $a_{iji} = b_{iji} = e$.

**Lemma 7.4** *The 1-cohomology of $X$ with coefficients in $\mathcal{G}$ is well defined. That is,* $\mathsf{H}^1(X; \mathcal{G})$ *is indeed a category, and in fact a groupoid.*

**Proof** Let $(a_{ijk}), (b_{ijk}), (c_{ijk}) \in \mathsf{Ob}\,\mathsf{H}^1(X; \mathcal{G})$, and $(f_{ij}): (a_{ijk}) \to (b_{ijk})$ and $(f'_{ij}): (b_{ijk}) \to (c_{ijk})$ be morphisms. Then $(f' \circ f)_{ij} := f_{ij}f'_{ij}$ defines a morphism $((f' \circ f)_{ij}): (a_{ijk}) \to (c_{ijk})$ since

$$a_{ijk}f_{jk}f'_{jk} = f_{ij}b_{ijk}f'_{jk} = f_{ij}f'_{ij}c_{ijk}.$$

It obviously satisfies associativity. The identity on $a_{ijk}$ is obviously defined by $e_{ij} = e$, where $e$ denotes the unit of $\mathcal{G}$. Further, $(f_{ij}^{-1})$ defines a morphism $(f_{ij}^{-1}): b_{ijk} \to a_{ijk}$ that is the inverse of $(f_{ij})$. □

**Proposition 7.5** *We have a faithful functor* $\beta \colon \mathsf{H}^1(X;\mathcal{G}) \to \mathrm{Tors}_X^{\mathcal{G}}$.

**Proof** For $(a_{ijk}) \in \mathrm{Ob}\, \mathsf{H}^1(X;\mathcal{G})$, we define a $\mathcal{G}$-torsor $\beta(a_{ijk})$ as follows. Let $\mathcal{U} = \coprod_{(i,j)\in I^2} \mathcal{G}_{ij}$, where $\mathcal{G}_{ij} = \mathcal{G}_i^{ij} \amalg \mathcal{G}_j^{ij} = \mathcal{G} \amalg \mathcal{G}$. We write an element of $\mathcal{G}_\bullet^{ij}$ as $g_\bullet^{ij}$ and we denote the identification $\mathcal{G} = \mathcal{G}_\bullet^{ij}$ by the map $\mathcal{G} \to \mathcal{G}_\bullet^{ij}$, $g \mapsto g_\bullet^{ij}$, where $\bullet \in \{i,j\}$ for all $i \neq j \in I$. We define an equivalence relation $\sim$ on $\mathcal{U}$ generated by

$$ g_j^{ij} \sim (g a_{ijk})_j^{jk}. $$

We have $g_j^{ij} \sim g_j^{ji}$ for all $i, j \in I$. We denote the quotient set $\mathcal{U}/{\sim}$ by $\beta(a_{ijk})$. Then we have a surjective map $\pi \colon \beta(a_{ijk}) \to X$ defined by $\pi[g_j^{ij}] = x_j$. For this map $\pi$, we have the following.

**Lemma 7.6** *For all* $i, j \in I$, *the map* $\mathcal{G} \to \pi^{-1} x_j$, $g \mapsto [g_j^{ij}]$, *is a bijection.*

**Proof** The surjectivity is clear. We show the injectivity. Suppose that we have $[g_j^{ij}] = [h_j^{ij}]$ for $g, h \in \mathcal{G}$. That is, we have elements $a_{k_0 j k_1}, a_{k_1 j k_2}, \ldots, a_{k_{N-1} j k_N} \in \mathcal{G}$ such that $g a_{k_0 j k_1} \ldots a_{k_{N-1} j k_N} = h$ and $k_0 = k_N = i$. Then the condition $a_{ijk} a_{kj\ell} = a_{ij\ell}$ implies that $g a_{iji} = h$, hence $g = h$. $\qquad\square$

Lemma 7.6 ensures that $[g_j^{ij}] = [h_j^{jk}]$ implies $h = g a_{ijk}$. Now we can define a distance function $d_{\beta(a_{ijk})}$ on $\beta(a_{ijk})$ as follows. Let $\varepsilon_i \in \pi^{-1} x_i$ and $\varepsilon_j \in \pi^{-1} x_j$. Then there uniquely exist $g, h \in \mathcal{G}$ such that $[g_i^{ij}] = \varepsilon_i$ and $[h_j^{ij}] = \varepsilon_j$ by Lemma 7.6. Then we define that

$$ d_{\beta(a_{ijk})}(\varepsilon_i, \varepsilon_j) = d_X(x_i, x_j) + d_{\mathcal{G}}(g, h). $$

The nondegeneracy is clear. The symmetry follows from that $[g_i^{ij}] = [g_i^{ji}]$. The triangle inequality is verified as follows. Let $\varepsilon_i \in \pi^{-1} x_i$, $\varepsilon_j \in \pi^{-1} x_j$ and $\varepsilon_k \in \pi^{-1} x_k$. Suppose that we have $[g_i^{ij}] = \varepsilon_i = [g_i'^{ik}]$, $[h_j^{ij}] = \varepsilon_j = [h_j'^{jk}]$, and $[m_k^{jk}] = \varepsilon_k = [m_k'^{ik}]$. Then we have $g = g' a_{kij}$, $h' = h a_{ijk}$ and $m = m' a_{ikj}$; hence we obtain that

$$
\begin{aligned}
d_{\beta(a_{ijk})}(\varepsilon_i,\varepsilon_j)+d_{\beta(a_{ijk})}(\varepsilon_j,\varepsilon_k) &= d_X(x_i,x_j)+d_{\mathcal{G}}(g,h)+d_X(x_j,x_k)+d_{\mathcal{G}}(h',m) \\
&= d_X(x_i,x_j)+d_X(x_j,x_k)+d_{\mathcal{G}}(g'a_{kij},h)+d_{\mathcal{G}}(ha_{ijk},m'a_{ikj}) \\
&= d_X(x_i,x_j)+d_X(x_j,x_k)+d_{\mathcal{G}}(g'a_{kij},h)+d_{\mathcal{G}}(ha_{ijk}a_{jki}a_{kij},m'a_{kij}) \\
&\qquad\qquad +d_{\mathcal{G}}(h,ha_{ijk}a_{jki}a_{kij})-d_{\mathcal{G}}(h,ha_{ijk}a_{jki}a_{kij}) \\
&\geq d_X(x_i,x_j)+d_X(x_j,x_k)+d_{\mathcal{G}}(g'a_{kij},m'a_{kij})-|a_{ijk}a_{jki}a_{kij}| \\
&\geq d_X(x_i,x_j)+d_X(x_j,x_k)+d_{\mathcal{G}}(g',m')-|\Delta(x_i,x_j,x_k)| \\
&\geq d_X(x_i,x_k)+d_{\mathcal{G}}(g',m') \\
&= d_{\beta(a_{ijk})}(\varepsilon_i,\varepsilon_k).
\end{aligned}
$$

Now a map $\pi \colon \beta(a_{ijk}) \to X$ is obviously 1-Lipschitz. We verify that it is a metric fibration as follows. Let $x_i, x_j \in X$ and $\varepsilon_i \in \pi^{-1} x_i$. Suppose that $\varepsilon_i = [g_i^{ij}]$ for $g \in \mathcal{G}$. Then $\varepsilon_j := [g_j^{ij}] \in \pi^{-1} x_j$ is the unique element in $\pi^{-1} x_j$ such that $d_{\beta(a_{ijk})}(\varepsilon_i, \varepsilon_j) = d_X(x_i, x_j)$. Also, for $\varepsilon_j' := [h_j^{ij}] \in \pi^{-1} x_j$, we

have $d_{\beta(a_{ijk})}(\varepsilon_i, \varepsilon_j') = d_X(x_i, x_j) + d_{\mathcal{G}}(g, h) = d_{\beta(a_{ijk})}(\varepsilon_i, \varepsilon_j) + d_{\beta(a_{ijk})}(\varepsilon_j, \varepsilon_j')$. Finally, we equip the metric fibration $\pi: \beta(a_{ijk}) \to X$ with a right action by $\mathcal{G}$ as $[g_\bullet^{ij}]h = [(h^{-1}g)_\bullet^{ij}]$ for all $i, j \in I$ and $\bullet \in \{i, j\}$. This is well defined since we have that

$$[(ga_{ijk})_j^{jk}]h = [(h^{-1}ga_{ijk})_j^{jk}] = [(h^{-1}g)_j^{ij}] = [g_j^{ij}]h.$$

It is straightforward to verify that this is a $\mathcal{G}$-torsor.

Next we show the functoriality. Let $(f_{ij}): (a_{ijk}) \to (b_{ijk}) \in \mathsf{H}^1(X; \mathcal{G})$. We construct a map $f_*: \beta(a_{ijk}) \to \beta(b_{ijk})$ by $[g_\bullet^{ij}] \mapsto [(gf_{ij})_\bullet^{ij}]$ for all $i, j \in I$ and $\bullet \in \{i, j\}$. It is well defined since we have that

$$[(ga_{ijk})_j^{jk}] \mapsto [(ga_{ijk}f_{jk})_j^{jk}] = [(gf_{ij}b_{ijk})_j^{jk}] = [(gf_{ij})_j^{ij}].$$

The map $f_*$ obviously preserves fibers, and is an isometry since we have that

$$\begin{aligned}
d_{\beta(b_{ijk})}\big(f_*[g_i^{ij}], f_*[h_j^{ij}]\big) &= d_{\beta(b_{ijk})}\big([(gf_{ij})_i^{ij}], [(hf_{ij})_j^{ij}]\big) \\
&= d_X(x_i, x_j) + d_{\mathcal{G}}(gf_{ij}, hf_{ij}) \\
&= d_X(x_i, x_j) + d_{\mathcal{G}}(g, h) \\
&= d_{\beta(a_{ijk})}([g_i^{ij}], [h_j^{ij}]).
\end{aligned}$$

Further, it is $\mathcal{G}$-equivariant since we have that

$$(f_*[g_j^{ij}])m = [(gf_{ij})_j^{ij}]m = [(m^{-1}gf_{ij})_j^{ij}] = f_*([g_j^{ij}]m).$$

The faithfulness is obvious from the construction. $\square$

**Proposition 7.7** *The functor* $\beta: \mathsf{H}^1(X; \mathcal{G}) \to \mathsf{Tors}_X^{\mathcal{G}}$ *is full.*

**Proof** Let $(a_{ijk}), (b_{ijk}) \in \mathrm{Ob}\, \mathsf{H}^1(X; \mathcal{G})$ be cocycles, and suppose that we have a morphism $\varphi: \beta(a_{ijk}) \to \beta(b_{ijk})$ in $\mathsf{Tors}_X^{\mathcal{G}}$. We denote the projections $\beta(a_{ijk}) \to X$ and $\beta(b_{ijk}) \to X$ by $\pi_a$ and $\pi_b$, respectively. For all $i, j \in I$, we have bijections $A_{ij}: \mathcal{G} \to \pi_a^{-1}x_j$ and $B_{ij}: \mathcal{G} \to \pi_b^{-1}x_j$ given by $g \mapsto [g_j^{ij}]$ by Lemma 7.6. Define a map $\varphi_{ij} = B_{ij}^{-1}\varphi A_{ij}: \mathcal{G} \to \mathcal{G}$. We have $\varphi[g_j^{ij}] = [(\varphi_{ij}g)_j^{ij}]$. Now the $\mathcal{G}$-equivariance of $\varphi$ implies

$$\varphi[g_j^{ij}] = \varphi[(ge)_j^{ij}] = (\varphi[e_j^{ij}])g^{-1} = [(\varphi_{ij}e)_j^{ij}]g^{-1} = [(g\varphi_{ij}e)_j^{ij}],$$

which implies that $\varphi_{ij}g = g\varphi_{ij}e$ by Lemma 7.6. From this, we obtain that

$$\varphi[(ga_{ijk})_j^{jk}] = \varphi[(ga_{ijk})_j^{kj}] = [(\varphi_{kj}(ga_{ijk}))_j^{kj}] = [(ga_{ijk}\varphi_{kj}e)_j^{kj}].$$

Since we have $[g_j^{ij}] = [(ga_{ijk})_j^{jk}]$, we obtain that $a_{ijk}\varphi_{kj}e = (\varphi_{ij}e)b_{ijk}$ by Lemma 7.6. Further, since the lift of $x_j$ along $[g_i^{ij}]$ is $[g_j^{ij}]$ and $\varphi$ preserves the lift, the conditions $\varphi[g_i^{ij}] = [(\varphi_{ij}g)_j^{ij}]$ and $\varphi[g_i^{ji}] = [(\varphi_{ji}g)_i^{ji}]$ implies that $\varphi_{ij} = \varphi_{ji}$. Hence we obtain a morphism $(\varphi_{ij}e): (a_{ijk}) \to (b_{ijk})$ in $\mathsf{H}^1(X; \mathcal{G})$, which satisfies that $\beta(\varphi_{ij}e) = \varphi$ by construction. $\square$

**Definition 7.8** Let $\pi\colon E \to X$ be a $\mathcal{G}$-torsor. For $x_i, x_j \in X$, we define a *local section of $\pi$ over a pair $(x_i, x_j)$* to be a pair of points $(\varepsilon_i, \varepsilon_j) \in E^2$ such that $\pi\varepsilon_i = x_i, \pi\varepsilon_j = x_j$ and $\varepsilon_j$ is the lift of $x_j$ along $\varepsilon_i$. We say that $((\varepsilon_i^{ij}, \varepsilon_j^{ij}))_{(i,j) \in I^2}$ is a *local section of $\pi$* if each $(\varepsilon_i^{ij}, \varepsilon_j^{ij})$ is a local section of $\pi$ over a pair $(x_i, x_j)$ and satisfies that $\varepsilon_i^{ij} = \varepsilon_i^{ji}$.

**Proposition 7.9** Let $\pi\colon E \to X$ be a $\mathcal{G}$-torsor. For a local section $s = ((\varepsilon_i^{ij}, \varepsilon_j^{ij}))_{(i,j) \in I^2}$ of $\pi$, we can construct a cocycle $\alpha_s \pi \in \mathrm{Ob}\, \mathsf{H}^1(X; \mathcal{G})$. Further, for all two local sections $s, s'$ of $\pi$, the corresponding cocycles $\alpha_s \pi$ and $\alpha_{s'} \pi$ are isomorphic.

**Proof** We define $a_{ijk} \in \mathcal{G}$ as the unique element such that $\varepsilon_j^{ij} a_{ijk} = \varepsilon_j^{jk}$. Then $(a_{ijk})$ satisfies that $a_{ijk} a_{kj\ell} = a_{ij\ell}$ since we have

$$\varepsilon_j^{ij} a_{ijk} a_{kj\ell} = \varepsilon_j^{jk} a_{kj\ell} = \varepsilon_j^{kj} a_{kj\ell} = \varepsilon_j^{j\ell}.$$

We have $\varepsilon_x g = (\varepsilon g)_x$ for all $\varepsilon \in E, x \in X$ and $g \in \mathcal{G}$. Hence we have that

$$\begin{aligned}
\varepsilon_j^{ij} a_{ijk} a_{jki} a_{kij} &= \varepsilon_j^{jk} a_{jki} a_{kij} = (\varepsilon_k^{jk})_{x_j} a_{jki} a_{kij} \\
&= (\varepsilon_k^{jk} a_{jki})_{x_j} a_{kij} = (\varepsilon_k^{ki})_{x_j} a_{kij} \\
&= ((\varepsilon_i^{ki})_{x_k} a_{kij})_{x_j} = ((\varepsilon_i^{ki} a_{kij})_{x_k})_{x_j} = ((\varepsilon_i^{ij})_{x_k})_{x_j}.
\end{aligned}$$

It follows that

$$\begin{aligned}
|a_{ijk} a_{jki} a_{kij}| &= d_E(\varepsilon_j^{ij}, \varepsilon_j^{ij} a_{ijk} a_{jki} a_{kij}) \\
&= d_E\big(\varepsilon_j^{ij}, ((\varepsilon_i^{ij})_{x_k})_{x_j}\big) \\
&= -d_E(\varepsilon_j^{ij}, \varepsilon_i^{ij}) + d_E\big(\varepsilon_i^{ij}, ((\varepsilon_i^{ij})_{x_k})_{x_j}\big) \\
&\le -d_E(\varepsilon_j^{ij}, \varepsilon_i^{ij}) + d_E(\varepsilon_i^{ij}, (\varepsilon_i^{ij})_{x_k}) + d_E\big((\varepsilon_i^{ij})_{x_k}, ((\varepsilon_i^{ij})_{x_k})_{x_j}\big) \\
&= -d_X(x_j, x_i) + d_X(x_i, x_k) + d_X(x_k, x_j).
\end{aligned}$$

Since the norm $|-|$ on $\mathcal{G}$ is conjugation invariant, the value $|a_{ijk} a_{jki} a_{kij}|$ is invariant under the cyclic permutation on $\{i, j, k\}$, and hence we obtain that $|a_{ijk} a_{jki} a_{kij}| \le |\Delta(x_i, x_j, x_k)|$. Thus we obtain a cocycle $\alpha_s \pi := (a_{ijk}) \in \mathrm{Ob}\, \mathsf{H}^1(X; \mathcal{G})$. Suppose that we have local sections $s = ((\varepsilon_i^{ij}, \varepsilon_j^{ij}))_{(i,j) \in I^2}$ and $s' = ((\mu_i^{ij}, \mu_j^{ij}))_{(i,j) \in I^2}$. Then there exists an element $(f_{ij}) \in \mathcal{G}^{I^2}$ such that $(\varepsilon_i^{ij} f_{ij}, \varepsilon_j^{ij} f_{ij}) = (\mu_i^{ij}, \mu_j^{ij})$. Let $\alpha_s \pi = (a_{ijk})$ and $\alpha_{s'} \pi = (b_{ijk})$. Then we obtain that

$$\varepsilon_j^{ij} a_{ijk} f_{jk} b_{ijk}^{-1} = \varepsilon_j^{jk} f_{jk} b_{ijk}^{-1} = \mu_j^{jk} b_{ijk}^{-1} = \mu_j^{ij},$$

which implies that $f_{ij} = a_{ijk} f_{jk} b_{ijk}^{-1}$. Hence $(f_{ij})$ defines a morphism $(f_{ij})\colon (a_{ijk}) \to (b_{ijk})$ in $\mathsf{H}^1(X; \mathcal{G})$. Since $\mathsf{H}^1(X; \mathcal{G})$ is a groupoid, this is an isomorphism. $\qquad\square$

**Proposition 7.10** The functor $\beta\colon \mathsf{H}^1(X; \mathcal{G}) \to \mathsf{Tors}_X^{\mathcal{G}}$ is split essentially surjective.

**Proof** Let $\pi\colon E \to X$ be a $\mathcal{G}$-torsor. Fix a local section $s = ((\varepsilon_i^{ij}, \varepsilon_j^{ij}))_{(i,j) \in I^2}$ of $\pi$. Let $\alpha_s \pi = (a_{ijk})$ be the cocycle constructed in Proposition 7.9. We show that the $\mathcal{G}$-torsors $\beta(a_{ijk})$ and $\pi$ are isomorphic.

We define a map $\varphi \colon \beta(a_{ijk}) \to E$ by $[g_\bullet^{ij}] \mapsto \varepsilon_\bullet^{ij} g^{-1}$. It is well defined since we have that

$$[(ga_{ijk})_j^{jk}] \mapsto \varepsilon_j^{jk} a_{ijk}^{-1} g^{-1} = \varepsilon_j^{ij} g^{-1}.$$

It obviously preserves fibers and is a bijection. Also, it is an isometry since we have that

$$
\begin{aligned}
d_E(\varphi[g_i^{ij}], \varphi[h_j^{ij}]) &= d_E(\varepsilon_i^{ij} g^{-1}, \varepsilon_j^{ij} h^{-1}) \\
&= d_E(\varepsilon_i^{ij}, \varepsilon_j^{ij} h^{-1} g) \\
&= d_E(\varepsilon_i^{ij}, \varepsilon_j^{ij}) + d_E(\varepsilon_j^{ij}, \varepsilon_j^{ij} h^{-1} g) \\
&= d_X(x_i, x_j) + d_{\mathcal{G}}(g^{-1}, h^{-1}) \\
&= d_{\beta(a_{ijk})}([g_i^{ij}], [h_j^{ij}]).
\end{aligned}
$$

Further, it is immediate that $\varphi$ is $\mathcal{G}$-equivariant. Hence the map $\varphi$ gives an isomorphism in $\mathsf{Tors}_X^{\mathcal{G}}$. ☐

**Corollary 7.11** *The functor* $\beta \colon \mathsf{H}^1(X; \mathcal{G}) \to \mathsf{Tors}_X^{\mathcal{G}}$ *is an equivalence of categories.* ☐

# References

[1] **Y Asao**, *Magnitude and magnitude homology of filtered set enriched categories*, preprint (2023)  arXiv 2303.05677

[2] **A Grothendieck**, *Technique de descente et théorèmes d'existence en géométrie algébrique, I: Généralités: Descente par morphismes fidèlement plats*, from "Séminaire Bourbaki 1959/1960", W A Benjamin, Amsterdam (1966) exposé 190  Reprinted in Séminaire Bourbaki 5 (1958–1960) 299–327, Soc. Math. France, Paris, 1995

[3] **A Grothendieck**, *Revêtements étales et groupe fondamental* (SGA 1), Lecture Notes in Math. 224, Springer (1971)  MR

[4] **P T Johnstone**, *Sketches of an elephant*: *a topos theory compendium*, *I*, Oxford Logic Guides 43, Oxford Univ. Press (2002)  MR

[5] **T Leinster**, *The magnitude of metric spaces*, Doc. Math. 18 (2013) 857–905  MR

[6] **T Leinster**, *The magnitude of a graph*, Math. Proc. Cambridge Philos. Soc. 166 (2019) 247–264  MR

[7] **T Leinster**, **M Shulman**, *Magnitude homology of enriched categories and metric spaces*, Algebr. Geom. Topol. 21 (2021) 2175–2221  MR

[8] **S Mac Lane**, *Categories for the working mathematician*, Graduate Texts in Math. 5, Springer (1971)  MR

[9] **E Roff**, *The size and shape of things*: *magnitude, diversity, homology*, PhD thesis, University of Edinburgh (2022)  Available at https://era.ed.ac.uk/handle/1842/39543

[10] **E Roff**, *Iterated magnitude homology*, preprint (2023)  arXiv 2309.00577

*Department of Applied Mathematics, Fukuoka University*
*Fukuoka, Japan*

asao@fukuoka-u.ac.jp

# Spaces over BO are thickened manifolds

Hiro Lee Tanaka

Consider the topologically enriched category of compact smooth manifolds (possibly with corners), with morphisms given by codimension-zero smooth embeddings. Now formally identify any object $X$ with its thickening $X \times [-1, 1]$. We prove that the resulting $\infty$-category of thickened smooth manifolds is equivalent to the $\infty$-category of finite spaces over BO. (This is one formalization of the philosophy that embedding questions become homotopy-theoretic upon passage to higher dimensions.) The central tool is a geometric construction of pushouts in this $\infty$-category, carried out with an eye toward proving analogous results in exact symplectic geometry. Notably, the proof never invokes smooth approximation nor any $h$-principle.

57N65, 57R40

## 1 Introduction

Throughout this work, we adopt the convention that a smooth manifold may have boundary and corners, and that all connected components of a manifold have equal dimension.

Consider the $\infty$-category $\mathrm{Mfld}^\diamond$ whose objects are smooth, compact manifolds $X$, up to the relation

$$(1\text{-}1) \qquad\qquad X \sim X \times [-1, 1],$$

and where morphisms are codimension-zero smooth embeddings — ie, smooth embeddings whose derivatives pointwise have trivial cokernel. We do not require that the maps respect boundary or corner strata in any way. See Notation 2.8 and 2.9 for a model of this $\infty$-category.

**Remark 1.1** Let $X = [-1, 1]^n$ be the cube and $Y = D^n$ the closed unit disk. Appropriate scalings define a smooth codimension-zero embedding of $X$ inside $Y$, and of $Y$ inside $X$. These are equivalences in $\mathrm{Mfld}^\diamond$, as the composite $X \to Y \to X$ is isotopic to the identity, and likewise for $Y \to X \to Y$. In particular, in $\mathrm{Mfld}^\diamond$, a point — which is equivalent to $[-1, 1]^n$ by (1-1) — is equivalent to a disk of any dimension.

**Remark 1.2** Let $X = \mathbb{R}^0$ be a point and $Y$ be any smooth compact manifold. Then the space of maps in $\mathcal{M}\mathrm{fld}^\diamond$ from $X$ to $Y$ is homotopy equivalent to the stable frame bundle of $Y$. To see this, note that if $X = [-1, 1]^d$ and $\dim Y = d$, then the space of codimension-zero smooth embeddings of $X$ into $Y$ is homotopy equivalent to the frame bundle of $Y$ — ie, the principle $\mathrm{GL}(d)$-bundle associated to the tangent bundle. Further details are in the proof of Proposition 4.5.

The classifying map $X \to \mathrm{BGL} \simeq \mathrm{BO}$ of the stable tangent bundle survives the identification (1-1) up to homotopy. If a map $j\colon X \to Y$ is a smooth embedding of codimension zero, we further know $j$ respects the map to BO. We thus obtain a functor of $\infty$-categories

$$\tag{1-2} \mathcal{M}\mathrm{fld}^\diamond \to \mathcal{S}\mathrm{pc}^{\mathrm{finite}}_{/\mathrm{BO}}$$

to the $\infty$-category of spaces homotopy equivalent to a finite CW complex equipped with a map to BO. (One model of this functor is given in Construction 2.13.) In this work we give a new proof of the following fact:

**Theorem 1.3** (1-2) *is an equivalence.*

This is one manifestation of the philosophy that questions of smooth embeddings become homotopy-theoretic upon thickening. Treated here is the case of smooth embeddings with trivialized normal bundles (ie, thickened codimension-zero embeddings).

**Remark 1.4** (precedents) Theorem 1.3 is of no surprise to experts. Indeed, that (1-2) is fully faithful can alternatively be proven using the $h$-principle. (Here is a sketch: the space of thickened embeddings might as well be a space of immersions with trivialized normal bundle, so invoke a version of the Smale–Hirsch theorem of immersions with trivialized normal bundles, increasing the dimensions of codomains.)

At the level of isomorphism classes of both sides of (1-2) — and in particular, considering the essential surjectivity of (1-2) — there are many precedents. Barry Mazur's works [10; 11] may have been the earliest. (Mazur proves stronger results about objects — up to diffeomorphisms, not up to isotopy equivalences — but weaker results in other respects, making no mention of the space of embeddings.) Mazur's results are referenced and enhanced by Waldhausen [14], as an equivalence of the underlying $\infty$-groupoids in (1-2) (rather than of the $\infty$-categories themselves). Wall [15] also investigated homotopy classes of maps to BO in terms of what he calls "thickenings" of CW complexes.

**Remark 1.5** The $\infty$-category of compact, smooth manifolds — possibly with corners and boundary, and allowing for all smooth maps — is equivalent to the $\infty$-category of spaces homotopy equivalent to finite CW complexes. By instead insisting on maps that are (thickened) codimension-zero embeddings (or, equivalently, smooth maps with trivialized normal bundles) we retain the data of the classifying map to BO, that is, the data of the stable tangent bundle.

**Corollary 1.6** *Suppose two compact manifolds (possibly with corners) $X$ and $Y$ admit a* **continuous** *homotopy equivalence $X \to Y$ respecting the maps to* BO *up to homotopy. Then there exists, possibly after thickening both $X$ and $Y$, a codimension-zero smooth embedding from $X$ to $Y$, and from $Y$ to $X$, which are mutually inverse up to isotopy.*

**Corollary 1.7** *Suppose that $X$ is a compact smooth manifold. Then the space of maps from $X$ to $D^N$ in $\mathcal{M}\mathrm{fld}^\diamond$ is homotopy equivalent to the space of null-homotopies of the classifying map of the stable tangent bundle of $X$.*

*In particular, if $X$ does not have stably trivial tangent bundle, the mapping space to $D^N$ is empty. Otherwise, the space of smooth embeddings into $D^N$ equipped with a trivialization of the normal bundle becomes, as $N$ goes to $\infty$, weakly homotopy equivalent to the space of continuous maps from $X$ to $O$.*

The main geometric fact we use to prove Theorem 1.3 is the following:

**Theorem 1.8** $\mathcal{M}\mathrm{fld}^\diamond$ *has finite colimits.*

While Theorem 1.8 is obvious given Theorem 1.3,[1] at face value Theorem 1.8 can perplex: We know full well that the collection of manifolds is not closed under gluing. Theorem 1.8 states instead that the collection of thickened manifolds is closed under homotopy coherent gluing. Both thickening, and the $\infty$-categorical notion of (homotopy) colimit, are necessary.

We will exhibit the colimits explicitly. As a consequence, we will be able to interpret handle attachments as pushouts; this in turn allows us to show that a single object — the point — generates $\mathcal{M}\mathrm{fld}^\diamond$ under finite colimits (Lemma 4.4). The constructions will also allow us to prove that the functor (1-2) preserves finite colimits (Proposition 4.2).

Now we may sketch the proof of Theorem 1.3: Both the domain and codomain of (1-2) are generated by the point, while the functor on the point is fully faithful by Remark 1.2. More details are given in Section 4.4.

Because $\mathcal{M}\mathrm{fld}^\diamond$ is a small $\infty$-category generated under finite colimits by a single object, Theorem 1.8 makes formal that its Ind-completion is presentable. We may thus remove the finiteness in the codomain of Theorem 1.3:

**Corollary 1.9** *The induced functor*

$$\mathrm{Ind}(\mathcal{M}\mathrm{fld}^\diamond) \to \mathcal{S}\mathrm{pc}_{/\mathrm{BO}}$$

*is an equivalence of $\infty$-categories.*

---

[1] We emphasize that in this note, Theorem 1.8 is proven first.

**Example 1.10** (a model for the terminal object of $\mathrm{Ind}(\mathcal{M}\mathrm{fld}^\diamond)$)  It is easily checked that $\mathcal{M}\mathrm{fld}^\diamond$ does not have a terminal object. For example, the space of maps from $X = \mathbb{R}^0 \sim [-1, 1]^N$ to any manifold $Y$ is homotopy equivalent to the stable frame bundle of $Y$ (Remark 1.2). Any manifold whose stable frame bundle is contractible must be homotopy equivalent to BO — and BO is not homotopy equivalent to the thickening of any compact manifold (with or without corners). After all, the cohomology of BO is not bounded in degree.

Presentability guarantees a terminal object; the corollary tells us the terminal object has the homotopy type of BO. So there is an Ind-object of $\mathcal{M}\mathrm{fld}^\diamond$, homotopy equivalent to BO, which one may model as an increasing union of (thickened) manifolds, and which serves as a terminal object in $\mathrm{Ind}(\mathcal{M}\mathrm{fld}^\diamond)$.

The reader may be tempted to think that this presentation of BO must be identical to the usual one: BO is a colimit $\mathrm{colim}_{n,k} \mathrm{Gr}_k(\mathbb{R}^n)$ of Grassmannians, and the maps in the colimit diagram are smooth embeddings. However, these embeddings do not have trivial normal bundles, even stably. Indeed, when $k = 1$, the embeddings $\mathbb{R}P^n \to \mathbb{R}P^{n+1}$ have normal bundles that pull back to the Möbius bundle along any $\mathbb{R}P^1 \subset \mathbb{R}P^n$, hence these normal bundles have nontrivial characteristic classes.

What is true is that, in the usual diagram of Grassmannians, one may find a cofinal subdiagram wherein the normal bundle of each smooth embedding is trivializable, and by choosing trivializations, this cofinal sub-diagram lifts to a diagram in $\mathcal{M}\mathrm{fld}^\diamond$. To see this, recall that the tangent bundle of $\mathrm{Gr}_k(\mathbb{R}^n)$ is identified with

$$\hom(\gamma_{k,n}, \gamma_{k,n}^\perp).$$

Here, $\gamma_{k,n}$ is the tautological vector bundle on $\mathrm{Gr}_k(\mathbb{R}^n)$ whose fiber over $V \in \mathrm{Gr}_k(\mathbb{R}^n)$ is $V$ itself, and $\gamma_{k,n}^\perp$ has fibers given by the orthogonal complement of $V$ in $\mathbb{R}^n$. For every $m \geq j \geq 0$, the standard embedding $\mathrm{Gr}_k(\mathbb{R}^n) \to \mathrm{Gr}_{k+j}(\mathbb{R}^{n+m})$ thus has normal bundle

$$(1\text{-}3) \quad \hom(\underline{\mathbb{R}^j}, \gamma_{k,n}^\perp) \oplus \hom(\gamma_{k,n}, \underline{\mathbb{R}^{m-j}}) \oplus \hom(\underline{\mathbb{R}^j}, \underline{\mathbb{R}^{m-j}}) \cong (\gamma_{k,n}^\perp)^{\oplus j} \oplus (\gamma_{k,n}^\vee)^{\oplus m-j} \oplus \hom(\underline{\mathbb{R}^j}, \underline{\mathbb{R}^{m-j}}),$$

where the underlines denote trivial vector bundles, and $\gamma_{k,n}^\vee$ denotes the $\mathbb{R}$-linear dual vector bundle. Using the standard inner product on $\mathbb{R}^n$, one has a short exact sequence of vector bundles on $\mathrm{Gr}_k(\mathbb{R}^n)$

$$0 \to \gamma_{k,n}^\perp \to \underline{\mathbb{R}^n} \to \gamma_{k,n}^\vee \to 0$$

and, because we are working continuously, this short exact sequence splits. In particular, so long as $m = 2j$, the normal bundle (1-3) is trivializable.

Thus, for any $0 \leq k \leq n$, the sequence of smooth manifold embeddings

$$\mathrm{Gr}_k(\mathbb{R}^n) \to \mathrm{Gr}_{k+1}(\mathbb{R}^{n+2}) \to \mathrm{Gr}_{k+2}(\mathbb{R}^{n+4}) \to \cdots$$

lifts to a sequence in $\mathcal{M}\mathrm{fld}^\diamond$. By cofinality, the above sequence is a defining colimit for BO in $\mathcal{S}\mathrm{pc}$. This allows us to lift the sequence to $\mathcal{S}\mathrm{pc}_{/\mathrm{BO}}$, where now the colimit is the terminal object of $\mathcal{S}\mathrm{pc}_{/\mathrm{BO}}$ (see Remark 4.1). By Corollary 1.9, the sequential diagram in $\mathrm{Ind}(\mathcal{M}\mathrm{fld}^\diamond)$ has colimit given by the terminal object of $\mathrm{Ind}(\mathcal{M}\mathrm{fld}^\diamond)$.

## 1.1 Motivations

As indicated above, Theorem 1.3 is a consequence of already available tools, but we had not seen the theorem in the literature. Not only do we record the result here, the proof in the present work is also distinct. Indeed, a main motivation for this work was to give a proof of Theorem 1.8 involving only straightforward constructions of manifolds, using no $h$-principles, and without invoking Theorem 1.3 (which one can now view as a consequence of thickened manifolds admitting finite colimits). This was not for the sake of exercise.

The author's main motivation was as follows.

It turns out that the $\infty$-category of Liouville sectors — a rich class of exact symplectic manifolds modeling cotangent bundles of singular spaces — also admits homotopy pushouts upon thickening. (This was mentioned, though not proven, in [6]. Indeed, we learned the particular model for the pushout in Section 3 from O Lazarev.)

But there are fewer tools for dealing with embeddings of Liouville sectors, as the $h$-principle is not as robustly available, and there is no known analogue of Theorem 1.3. It was with an eye toward understanding colimits of Liouville sectors that we endeavored to write this note.

The central labor of the present work (Section 3) is the labor that carries over to the Liouville setting, and this is the only method we know for exhibiting colimits of Liouville sectors.

**Remark 1.11** The Ind-completion of the $\infty$-category of thickened Liouville sectors also has a terminal object, but we do not know its homotopy type at present. Compare with Example 1.10.

**Remark 1.12** (we do not use any $h$-principles) In a previous draft of this work, we came close to using an $h$-principle by invoking smooth approximation — ie, that any continuous map between smooth manifolds may be continuously homotoped to be smooth — to prove that (1-2) is essentially surjective. As pointed out to us by Branko Juran, the use of smooth approximation is unnecessary. Indeed, in the same draft we had already proven that the functor preserves finite colimits and is fully faithful; and this is enough (see Section 4.4).

There is a second motivation to pursue an $h$-principle-free proof of Theorem 1.3. As pointed out by an anonymous referee and by Branko Juran, it is natural to pursue analogues of Theorem 1.3 for PL manifolds and topological manifolds, where now the stabilized $\infty$-categories would be equivalent to $\mathcal{S}\mathrm{pc}^{\mathrm{finite}}_{/\mathrm{BPL}}$ and $\mathcal{S}\mathrm{pc}^{\mathrm{finite}}_{/\mathrm{BTop}}$, respectively. Indeed, if anything, all our finagling with smoothings becomes unnecessary. (Smoothing is the reason we have to deal with $P$ and $q$ as opposed to the more straightforward $P'$ and $q'$ in Section 3.) The main technical tools are still available via

  (i) the theory of microbundles, and

  (ii) the existence of handle decompositions for PL manifolds (using stars) and for topological manifolds (upon thickening — topological handle decompositions do not exist in all dimensions).

Generalizing Theorem 1.3 to the PL and topological settings would produce rather clean formulations of not just smoothing at the object level, but smoothing thickened embeddings by studying fibers of the functors $\mathcal{Spc}^{\text{finite}}_{/\text{BO}} \to \mathcal{Spc}^{\text{finite}}_{/\text{BPL}}$ and $\mathcal{Spc}^{\text{finite}}_{/\text{BPL}} \to \mathcal{Spc}^{\text{finite}}_{/\text{BTop}}$. (For example, one would have an obstruction-theoretic formulation of when a family of thickened codimension-zero topological embeddings lifts to a family of thickened codimension-zero smooth embeddings.) The *object*-level question of when one can lift a thickened topological manifold to a thickened smooth manifold was already addressed completely in Milnor's original work on microbundles (see Theorem 5.13 of [12]) with answers phrased in terms of topological $K$-theory with respect to $O$ and to $\mathcal{Spc}$.

We do not plan to pursue a full proof of the PL and topological cases, nor a suitable exploration of these implications, at the moment.

## 1.2   Convention: spaces, limits and colimits

Throughout this note, we use the term "mapping space" to refer to a Kan complex of morphisms in a simplicially enriched category, or to a Kan complex of morphisms in an $\infty$-category.

Nowadays it is sometimes common to use the term "anima" when a speaker chooses to remain agnostic about a model for a homotopy theory of spaces (so "anima" could refer to spaces, or to Kan complexes). However, in this work, we must use both Kan complexes and topological spaces (homotopy equivalent to CW complexes) explicitly. For example, our "mapping spaces" are typically Kan complexes, while the underlying space of a manifold is a topological space. Regardless, we do not encounter any topological space that cannot be recovered (up to homotopy equivalence) by its singular Kan complex, so this distinction will be immaterial for us when invoking $\infty$-categorical arguments—the anima aficionado may simply think of our models as anima.

For $\mathcal{C}$ an $\infty$-category and a diagram $I \to \mathcal{C}$, there is only one notion of limit (and only one notion of colimit). This notion captures the intuition of what would traditionally be called a "homotopy (co)limit."

In contrast, for the category of topological spaces or of Kan complexes (and, more generally, any model category) there is utility in distinguishing the categorical notion of (co)limit from the $\infty$-categorical notion. So we may at times emphasize a "point-set" colimit of a diagram of topological spaces, for example in (3-4). This means we take the colimit in the sense of the 1-category of topological spaces—ie, the usual, nonhomotopy-theoretic, notion of gluing topological spaces together. We will also use the term "homotopy (co)limit" in the context of topological spaces or Kan complexes for purposes of emphasis or disambiguation. See Proposition 3.21 for an example of this paragraph and of the preceding paragraph.

## Acknowledgments

at MSRI/SLMATH. We thank Diarmuid Crowley, John Francis, Søren Galatius, John Klein, Sander Kupers, and Jacob Lurie for communications educating me on past works. We also thank Branko Juran for his interest and insight (see Remark 1.12). Finally, we thank the anonymous referees for helpful comments.

# 2 From smooth manifolds to spaces over BO

Here we define $\mathcal{M}\mathrm{fld}^\diamond$ and give a construction of (1-2).

## 2.1 Preliminaries on manifolds with corners

**Remark 2.1** (conventions for manifolds with corners)   There are various definitions of manifolds with corners in the literature, especially when one begins to endow manifolds with stratifying data. We supply no such data and we use the minimal definition of smooth manifolds with corners. We refer to the first paragraphs of Section 2 of [5]; we also note that our notion of smoothness is called *weakly smooth* there.

In particular, an $n$-manifold with corners $X$ in our work is a paracompact[2] Hausdorff topological space equipped with a maximal atlas of smoothly compatible local charts from open subsets of the octant $(\mathbb{R}_{\geq 0})^n$. For any $x \in X$, whether $x$ is the image of a codimension-$k$ face of the octant is invariant under change of charts.

As usual, a map between Euclidean octants is called smooth if it arises as the restriction of a smooth map defined on a small neighborhood of the octants. Through charts, this defines a notion of smoothness for maps between manifolds with corners in the usual way. In particular, smooth maps between manifolds with corners, in this work, need not respect boundary/corner strata — for example, a compact disk may smoothly map to Euclidean space.

**Definition 2.2**   A smooth embedding $j : X \to Y$ is called an *isotopy equivalence* if there exists a smooth embedding $h : Y \to X$ together with smooth isotopies $jh \sim \mathrm{id}_Y$ and $hj \sim \mathrm{id}_X$. (Just as with $h$ and $j$, the smooth isotopies need not respect strata in any way.)

**Remark 2.3** (one can replace corners with boundaries)   Consider a compact smooth manifold $X$ with nonempty corner and boundary strata. There exists a smooth manifold with boundary (but no corners) $X'$ and an isotopy equivalence $X' \hookrightarrow X$. Details for one such construction may be found in Douady and Hérault's appendix to Borel and Serre's compactification paper [2].

We do caution that, it seems to this writer that $X'$ is not unique up to diffeomorphism — indeed, two such $X'$ must have smoothly cobordant boundary, but it is not clear that even the boundaries of $X'$ need

---

[2] All the manifolds we consider in our categories will be compact; but the use of local arguments makes it convenient to define manifolds with corners in noncompact settings. Our manifolds will also have components of equal dimension later on; this is irrelevant for the present, general discussion.

be diffeomorphic (though Douady and Hérault's methods will produce *homeomorphic* boundaries). We have not come up with an example showing that the $X'$ need not be diffeomorphic, though we strongly suspect there exist such examples. It is clear that $X'$ is unique up to isotopy equivalence (as all $X'$ are isotopy equivalent to $X$); this suffices for us.

**Remark 2.4** One may, in light of Remark 2.3, be tempted to consider only smooth manifolds with boundary (and no corners) but this would complicate the coherence of the thickening process that follows. So we allow for corners.

## 2.2 Defining the $\infty$-category of thickened manifolds

**Notation 2.5** ($\mathsf{Mfld}_d$) Let $\mathsf{Mfld}_d$ denote the Kan-complex enriched category whose objects are compact, smooth manifolds of dimension $d$, and whose morphism spaces are (the singular complex of) the spaces of smooth embeddings. (We do not demand that the embeddings respect corner strata in any way.)

**Remark 2.6** One model for the Kan complex of morphisms from $X$ to $Y$ is as follows. Let $\Delta_e^k$ be the subspace of $\mathbb{R}^{k+1}$ whose coordinates sum to 1. A $k$-simplex of $\hom_{\mathsf{Mfld}_d}(X, Y)$ is the data of a smooth embedding $X \times \Delta_e^k \to Y \times \Delta_e^k$ respecting the projections to $\Delta_e^k$. This was also used, for example, in [1, Convention 4.1.5].

**Remark 2.7** The natural notion of equivalence (ie, a map that admits an inverse up to homotopy) in $\mathsf{Mfld}_d$ coincides with the notion of isotopy equivalence.

We have a functor of simplicially enriched categories

$$(2\text{-}1) \qquad\qquad \mathsf{Mfld}_d \to \mathsf{Mfld}_{d+1}, \quad X \mapsto X \times [-1, 1],$$

taking a manifold to a direct product with the compact unit interval.

**Notation 2.8** ($\mathsf{Mfld}^\diamond$) We define

$$(2\text{-}2) \qquad\qquad \mathsf{Mfld}^\diamond := \mathrm{colim}(\mathsf{Mfld}_0 \xrightarrow{-\times[-1,1]} \mathsf{Mfld}_1 \xrightarrow{-\times[-1,1]} \cdots),$$

concretely modeled as an increasing union of simplicially enriched categories.[3] We call $\mathsf{Mfld}^\diamond$ the (simplicially enriched) category of thickened compact manifolds.

Informally, $\mathsf{Mfld}^\diamond$ is a category where any compact manifold $X$ is identified with $X \times [-1, 1]^N$ for any $N \geq 0$, and where a morphism from $X$ to $Y$ is a codimension-zero embedding $X \times [-1, 1]^N \to Y \times [-1, 1]^{N'}$ for some $N$, $N'$.

**Notation 2.9** ($\mathcal{M}\mathrm{fld}^\diamond$) We define the $\infty$-category

$$\mathcal{M}\mathrm{fld}^\diamond := N(\mathsf{Mfld}^\diamond)$$

to be the homotopy coherent nerve of the simplicially enriched category $\mathsf{Mfld}^\diamond$.

---

[3]This is also a homotopy colimit in the model category of simplicially enriched categories.

**Remark 2.10** We recall that the homotopy coherent nerve $N$ commutes with filtered colimits, so one could equivalently define $\mathcal{M}\mathrm{fld}^\diamond$ as the increasing union of $\infty$-categories

$$N(\mathsf{Mfld_0}) \to N(\mathsf{Mfld}^1) \to \cdots$$

obtained by applying $N$ to (2-2).

The mapping spaces in the nerve are homotopy equivalent to the mapping spaces of the simplicially enriched category, so long as the simplicially enriched category has morphisms given by Kan complexes (this hypothesis is met for $\mathsf{Mfld}_d$ for all $d$ and for $\mathsf{Mfld}^\diamond$) — see, for example, Theorems 1.1.5.13 and 2.2.0.1 of [7]. As a result, for any two objects $X, Y \in \mathsf{Mfld}^\diamond$, we have a homotopy equivalence of Kan complexes

(2-3)
$$\hom_{\mathcal{M}\mathrm{fld}^\diamond}(X, Y) \simeq \hom_{\mathsf{Mfld}^\diamond}(X, Y).$$

Any manifold $X$ of dimension $d$ is the base of the frame bundle $\mathrm{Fr}_d(X)$, which is a principle $\mathrm{GL}(d)$-bundle. By thickening, we obtain a space $\mathrm{Fr}(X)$ with a free GL-action, where GL is the infinite general linear group, which we call the stabilized frame bundle of $X$. Let $\mathsf{Top}^{\mathrm{GL}}$ denote the simplicially enriched category of topological spaces (homotopy equivalent to CW complexes) equipped with a continuous GL-action. Then the stable frame bundle construction gives rise to a functor of $\infty$-categories

(2-4)
$$\mathrm{Fr}\colon \mathcal{M}\mathrm{fld}^\diamond \to N(\mathsf{Top}^{\mathrm{GL}})$$

to the $\infty$-category of spaces with GL-action (obtained as the homotopy coherent nerve of $\mathsf{Top}^{\mathrm{GL}}$).

**Remark 2.11** A concrete model of Fr may be given as follows. The commutative (up to natural transformation) diagram of simplicially enriched categories

(2-5)
$$
\begin{array}{ccc}
\mathsf{Mfld}_d & \xrightarrow{\ \mathrm{Fr}_d\ } & \mathsf{Top}^{\mathrm{GL}(d)} \\
{\scriptstyle -\times[-1,1]}\big\downarrow & & \big\downarrow{\scriptstyle -\times_{\mathrm{GL}(d)}\mathrm{GL}(d+1)} \\
\mathsf{Mfld}_{d+1} & \xrightarrow{\ \mathrm{Fr}_{d+1}\ } & \mathsf{Top}^{\mathrm{GL}(d+1)}
\end{array}
$$

where $\mathrm{Fr}_d \times_{\mathrm{GL}(d)} \mathrm{GL}(d+1)$ is the principal $\mathrm{GL}(d+1)$-bundle associated to $\mathrm{Fr}_d$ via the group homomorphism $\mathrm{GL}(d) \to \mathrm{GL}(d+1)$. Because the natural transformation maps are — for all objects $X \in \mathsf{Mfld}_d$ — equivalences in $N(\mathsf{Top}^{\mathrm{GL}(d+1)})$, the homotopy-coherent nerve $N$ renders (2-5) to a diagram $\Delta^1 \times \Delta^1 \to \mathcal{C}\mathrm{at}_\infty$ in the $\infty$-category of $\infty$-categories. Noting the natural map

$$\mathrm{colim}(\mathsf{Top}^{\mathrm{GL}(0)} \to \mathsf{Top}^{\mathrm{GL}(1)} \to \mathsf{Top}^{\mathrm{GL}(2)} \to \cdots) \to \mathsf{Top}^{\mathrm{GL}},$$

and that $N$ commutes with filtered colimits, the colimit (indexed by $d$) of (2-5) induces (2-4).

## 2.3 The functor to spaces over BO

**Remark 2.12** It is classical that the maps $O(d) \to \mathrm{GL}(d)$ are homotopy equivalences of group-like $E_1$-spaces, and that the induced map $O \to \mathrm{GL}$ is an equivalence of group-like $E_1$ spaces. (In fact, though

we will not need this, this map may be promoted to an equivalence of infinite loop spaces.) As a result, the induced functor (restricting a GL-action to an $O$-action)

$$(2\text{-}6) \qquad N(\mathsf{Top}^{\mathrm{GL}}) \to N(\mathsf{Top}^O)$$

is an equivalence of $\infty$-categories.

Let BO denote the Kan complex modeling the classifying space of the simplicial group $O$. (The interested reader may find a concrete model for BO and EO in [9, page 87], where BO is denoted by $\overline{W}(H)$ with $H = O$.) It is classical that one has a functor of $\infty$-categories

$$(2\text{-}7) \qquad N(\mathsf{Top}^O) \to \mathsf{Spc}_{/\mathrm{BO}}$$

from the nerve of the simplicially enriched category $\mathsf{Top}^O$ to the $\infty$-category $\mathsf{Spc}_{/\mathrm{BO}}$ of spaces (Kan complexes) equipped with a map to BO. Informally, the map (2-7) takes an $O$-space, equipped with the $O$-equivariant map to a point, to its homotopy quotient, equipped with the induced map to $\mathrm{pt}/O \simeq \mathrm{BO}$.

Moreover, (2-7) is an equivalence of $\infty$-categories. A proof using simplicially enriched model categories (where equivalences between $O$-spaces are the "coarse", also known as "naive," equivalences) goes back to [3].

**Construction 2.13** Composing (2-4), (2-6) and (2-7), we obtain a functor of $\infty$-categories

$$\mathsf{Mfld}^\diamond \to \mathsf{Spc}_{/\mathrm{BO}}.$$

In fact, it is easy to note that (because all of our manifolds are compact) the above functor factors through the $\infty$-category of (spaces homotopy equivalent to) finite CW complexes equipped with maps to BO:

$$\mathsf{Mfld}^\diamond \to (\mathsf{Spc}^{\mathrm{finite}})_{/\mathrm{BO}}.$$

This is the map (1-2).

**Example 2.14** Given a manifold $X$ of dimension $d$, its image in $(\mathsf{Spc}^{\mathrm{finite}})_{/\mathrm{BO}}$ is computed as follows. First, $\mathrm{Fr}(X)$ is a space with a free $O = O(\infty)$ action. The image of this space under (2-7) is the homotopy quotient of $\mathrm{Fr}(X)$ by the $O$-action, equipped with its natural map to BO. In particular, (2-7) sends $X$ to a space homotopy equivalent to $X$ equipped with a map to BO. We know this map is homotopic to the map classifying the stable frame bundle, as by construction it fits into a homotopy pullback square of spaces

$$
\begin{array}{ccc}
\mathrm{Fr}(X) & \longrightarrow & * \simeq \mathrm{EO} \\
\downarrow & & \downarrow \\
X & \longrightarrow & \mathrm{BO}
\end{array}
$$

where both vertical arrows are quotients by $O$.

**Remark 2.15** (alternative construction of (1-2)) As shown in the proof of Proposition 4.5, one has a natural identification of $\hom_{\mathsf{Mfld}^\diamond}(\mathrm{pt}, X)$ with the stable frame bundle $\mathrm{Fr}(X)$ — as an $O$-space over $X$. Thus, as pointed out to us by an anonymous referee, the assignment $X \mapsto \hom_{\mathsf{Mfld}^\diamond}(\mathrm{pt}, X)$ gives an alternative characterization of (1-2).

# 3 Pushouts of thickened manifolds

**Setup 3.1** Throughout, we fix three compact manifolds $W$, $X$, $Y$. Stabilizing as necessary, we assume $\dim X = \dim W = \dim Y$. By Remark 2.3, we may further assume that $W$, $X$, $Y$ have no corners (but they may have boundary). We equip them with smooth, codimension-zero embeddings

$$(3\text{-}1) \qquad\qquad X \xleftarrow{i_X} W \xrightarrow{i_Y} Y.$$

**Example 3.2** In the special case that $W$ is a manifold with no boundary and no corners, we see that $i_X$ and $i_Y$ are diffeomorphisms onto certain connected components of $X$ and $Y$.

(To see this, note that the interior of $W$ — meaning the subspace of $W$ admitting charts from $\mathbb{R}^n$, or equivalently, the complement of the boundary and corner strata — is an open subset of $W$. If $W$ has no boundary or corners, then $W$ is equal to its interior. On the other hand, any smooth embedding from a manifold with no boundary and no corners to a manifold of the same dimension is an open mapping. Thus when $W$ is compact the image of $i_X$ is both open and closed.)

## 3.1 The model $P$ of a pushout

Given (3-1), consider the topological space

$$(3\text{-}2) \qquad P' := P'(i_X, i_Y) := [-2, -1] \times X \bigcup_{i_X \times \mathrm{id}_{\{-1\}}} [-1, 1] \times W \bigcup_{i_Y \times \mathrm{id}_{\{1\}}} [1, 2] \times Y$$

(which one can recognize as one model for a homotopy pushout of spaces). See Figure 1.

**Remark 3.3** When $W$ has boundary, $P'$ is not canonically a smooth manifold with corners. This is because one must choose a smooth atlas near $\{-1, 1\} \times \partial W$. On the other hand, if we remove the locus
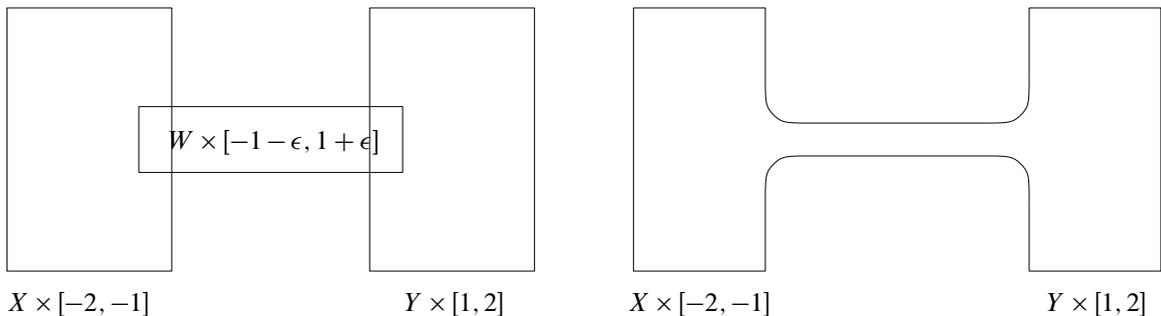


Figure 1: A cartoon of $P'$ on the left (with regions indicated), and of $P$ on the right. The intervals runs horizontally, while the vertical directions are meant to convey movement in the $W$, $X$, $Y$ directions. We have drawn not only $W \times [-1, 1]$, but the whole $W \times [-1 - \epsilon, 1 + \epsilon]$ to emphasize that $W$ admits a collar inside $X \times [-2, -1]$ and inside $Y \times [1, 2]$. The drawing of $P$ (note that the cornered are rounded) is meant to indicate that $P$ is obtained by removing the not-obviously-smooth regions where the rectangles in the $P'$ drawing intersect.

$\{-1, 1\} \times \partial W$, $P'$ has a canonical smooth structure. This is a standard issue. The smooth structure on $P'$ is only as undefined as a smooth structure on the (closed) complement of an (open) quadrant in $\mathbb{R}^2$. Indeed, an outward normal to $\partial W$ and normal vectors to $\{-1, 1\}$ inside $[-2, 2]$ reduce the problem to $(\mathbb{R}^2 \setminus \text{Quadrant}) \subset \mathbb{R}^2$ by considering the embedding

$$(\mathbb{R}^2 \setminus \text{Quadrant}) \times \partial W \hookrightarrow P'.$$

In the following remarks, we present two ways to treat $P'$ as, or alter $P'$ to be, a smooth object. We are careful here because, to avoid notational clutter, it will be very convenient to be able to treat the obvious embeddings

$$(3\text{-}3) \qquad [-2, -1] \times X \hookrightarrow P', \quad [-2, 2] \times W \hookrightarrow P', \quad [1, 2] \times Y \hookrightarrow P'$$

as all smooth. (The embedding $[-2, 2] \times W \hookrightarrow P'$ is set-theoretically defined piecewise, equaling $i_X$ and $i_Y$ along the subintervals $[-2, -1]$ and $[1, 2]$, respectively.)

## 3.2 A smoothing of $P'$

**Choice 3.4** Choose a small positive real number $\epsilon$. (All that matters is that $\epsilon$ is less than 1, and this 1 is only relevant because we play with intervals of length 1 below.)

We may also present $P'$ as the point-set colimit of the diagram of topological spaces

$$(3\text{-}4) \qquad \begin{array}{ccc} & & W \times [1, 1+\epsilon] \xrightarrow{\ i_Y \times \iota\ } Y \times [1, 2] \\ & & \downarrow{\scriptstyle \mathrm{id}_W \times \iota} \\ W \times [-1-\epsilon, -1] \xrightarrow{\ \mathrm{id}_W \times \iota\ } W \times [-1-\epsilon, 1+\epsilon] \\ \downarrow{\scriptstyle i_X \times \iota} \\ X \times [-2, -1] \end{array}$$

where the $\iota$ maps are inclusions of subintervals.

**Notation 3.5** ($u$) We will denote an element of (all of) the intervals in (3-4) by $u$.

**Example 3.6** In the setting of Example 3.2, $P'$ inherits an obvious smooth atlas and is hence a smooth manifold (possibly with boundary and corners) with atlas induced by those of $W$, $X$, $Y$ and $[0, 1]$. If $W$, $X$, $Y$ are furthermore connected, then $P'$ is diffeomorphic to $X \times [-2, 2] \cong X \times [0, 1]$.

We make a smooth model $P$ of $P'$ as follows.

**Choice 3.7** (a coordinatized neighborhood of $\partial W$ in $P'$) As per Setup 3.1, we have assumed that $W$, $X$, $Y$ have no corners (but they may have boundary). This means $W$ admits a vector field strictly outward-pointing along its boundary. So choose one. Flowing along this vector field for time $-2\epsilon$ (which is possible because $W$ is compact) realizes a self-embedding

$$c : W \to W$$

whose image is disjoint from $\partial W$. This induces a diffeomorphism

$$(3\text{-}5) \qquad W \bigcup_{\partial W \times (-\infty, 0]_v} (\partial W \times (-\infty, \epsilon)_v) \xrightarrow{\cong} W \setminus \partial W.$$

The subscript $v$ indicates the variable we will use for elements of the intervals in (3-5); we have parametrized so that the $v = -\epsilon$ locus in the domain is sent to $c(\partial W)$ in the codomain.

**Remark 3.8** We may consider the direct product of (3-5) with $\iota$ to think of

$$(\partial W \times (-\infty, \epsilon)_v) \times [-2, 2]_u \cong \partial W \times [-2, 2]_u \times (-\infty, \epsilon)_v$$

as a subset of $P'$. Doing so, inside $P'$ there is a neighborhood of $\partial W \times \{-1\}_u \times \{0\}_v$ homeomorphic to

$$\partial W \times \big( (-1 - \epsilon, -1 + \epsilon)_u \times (-\infty, 0]_v \bigcup (-1 - \epsilon, -1]_u \times (-\infty, \epsilon)_v \big).$$

Likewise we may think of $\partial W \times \{1\}_u \times (-\infty, \epsilon)_v$ as a subset of $P'$ and we obtain an identification

$$\mathrm{Nbhd}(\partial W \times \{1\}_u \times \{0\}_v) \cong \partial W \times \big( (1 - \epsilon, 1 + \epsilon)_u \times (-\infty, 0]_v \bigcup [1, 1 + \epsilon)_u \times (-\infty, \epsilon)_v \big).$$

**Choice 3.9** (smoothing a concave corner using $\gamma$, $R_X$, $R_Y$) Now choose a set

$$\gamma \subset A := (-1 - \epsilon, -1 + \epsilon)_u \times (-2\epsilon, \epsilon)_v$$

such that:

(i) $\gamma$ is a smooth, connected 1-dimensional submanifold.

(ii) For some small ball $B$ containing the point $\{-1\}_u \times \{-\epsilon\}_v \in A$, we have

$$\gamma \bigcap (A \setminus B) = \big( [-1, \infty)_u \times \{-\epsilon\}_v \bigcup \{-1\}_u \times [-\epsilon, \infty)_v \big) \bigcap (A \setminus B).$$

That is, outside of a tiny neighborhood of $\{-1\}_u \times \{-\epsilon\}_v$, $\gamma$ is equal to two positive rays.

(iii) $\gamma$ is contained in the region where $u \geq -1$ and $v \geq -\epsilon$.

See Figure 2. By (a version of) the Jordan curve theorem, $\gamma$ divides $A$ into two connected regions. We let

$$R_X \subset A$$

denote the region containing $\{-1\}_u \times \{-\epsilon\}_v$ and with boundary given by $\gamma$, so $R_X$ is a smooth 2-manifold with smooth boundary. We informally think of $R_X$ as smoothing the region $\{u \leq -1\} \cup \{v \leq -\epsilon\}$ (or as smoothing $\{u \leq -1\} \cup \{v \leq 0\}$ by removing a chunk of it).
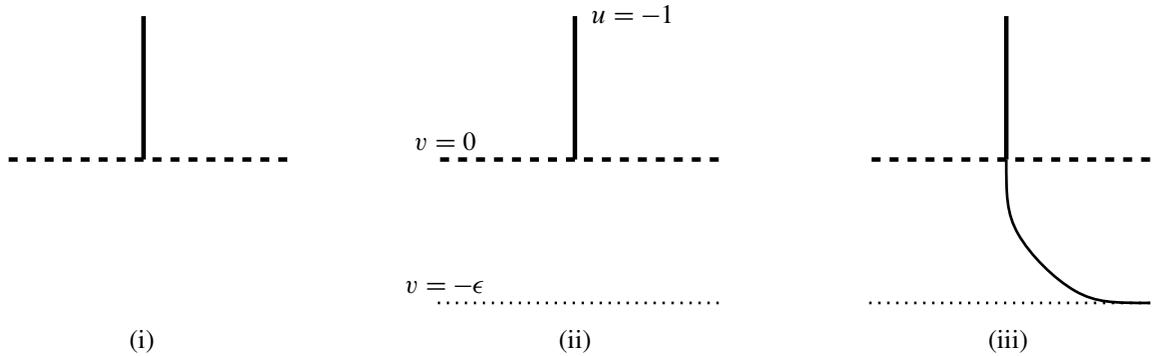
Figure 2: Illustrations of Choice 3.9. In all illustrations is a neighborhood of $\partial W \times \{-1\}_u \times \{0\}_v$ inside $P'$, projected onto the square $(-1-\epsilon, -1+\epsilon)_u \times [-\epsilon, \epsilon)_v$. As in Figure 1, the $u$ direction is drawn horizontally. The $v$ direction is drawn vertically. In (i), $\partial W \times \{-1\}_u \times \{0\}_v$ is the fiber over the point where the solid line intersects the dashed line. The fiber above the dashed line is the locus $\partial W \times (-1-\epsilon, -1+\epsilon)_u \times \{0\}_v$. The region below the dashed locus represents the region $\partial W \times (-1-\epsilon, -1+\epsilon)_u \times [-\epsilon, 0)_v$. In (ii) is now a thin dotted line, indicating the locus $\partial W \times (-1-\epsilon, -1+\epsilon)_u \times \{-\epsilon\}_v$. The new curve in (iii) indicates $\gamma$. The region consisting of points to the left of or below $\gamma$ is $R_X$. Note that $R_X$ intersects some portion of the thick dashed line, but does not contain all of the thick dashed line.

Likewise, there is a smooth manifold with boundary

$$R_Y \subset (1-\epsilon, 1+\epsilon)_u \times (-2\epsilon, \epsilon)_v$$

smoothing the locus $\{u \geq 1\} \cup \{v \leq -\epsilon\}$.

**Construction 3.10** $(P)$  We let $P \subset P'$ denote the smooth manifold with boundary (and no corners)

$$P := (X \times [-2, -1]_u) \cup (\partial W \times R_X) \cup (c(W) \times [-1-\epsilon, 1+\epsilon]_u) \cup (\partial W \times R_Y) \cup (Y \times [1, 2]_u).$$

See Figure 1. We are using Remark 3.8 to identify $\partial W \times R_X$ and $\partial W \times R_Y$ as subsets of $P'$. We think of $c(W) \times [-1-\epsilon, 1+\epsilon]$ as a subset of $W \times [-1-\epsilon, 1+\epsilon]$ in the defining diagram of $P$ (3-4).

**Notation 3.11** $(\alpha)$  By construction, $P$ is a subset of $P'$. The inclusion defines a smooth embedding

(3-6)                                        $$\alpha \colon P \hookrightarrow P',$$

where the smoothness of $a$ makes sense because $a$ avoids the locus $\{-1, 1\} \times \partial W$.

**Choice 3.12** $(\beta)$  Choose a smooth function $\phi$ compactly supported inside $\partial W \times (R_X \amalg R_Y) \subset P'$, and consider the vector field $-\phi \partial_v$ defined on all of $P'$. For a well-chosen $\phi$, we can arrange for the time-1 flow to have image completely contained in $P$. This isotopy is clearly smooth in the coordinate system invoked in Remark 3.8. We will denote by

$$\beta \colon P' \hookrightarrow P$$

the end result (ie, the time-1 map) of a choice of such an isotopy.

**Remark 3.13** By construction, $\beta$ respects the projection to $[-2, 2]_u$.

**Remark 3.14** Note that $\beta\alpha$ is smoothly isotopic to $\mathrm{id}_P$. Also, even though $P'$ is not a smooth manifold a priori, let us define a map out of $P'$ to be smooth if and only if the precomposition with $\beta^{-1} \colon \beta(P') \to P'$ is smooth. Then one also finds that $\alpha\beta$ is smoothly isotopic to $\mathrm{id}_{P'}$. Thus, for any test object $Z$ — that is, any object $Z$ of $\mathsf{Mfld}_{\dim P}$ — we find a homotopy equivalence

(3-7) $$\alpha^* \colon \mathrm{Sing}\,\mathrm{Emb}'(P', Z) \to \mathrm{Sing}\,\mathrm{Emb}(P, Z),$$

where $\mathrm{Emb}'$ is the space of smooth embeddings out of $P'$ (in the sense defined in this remark). By definition, the maps in (3-3) are smooth embeddings; in this way, we conclude that $P$ is a pushout in the $\infty$-category $\mathcal{M}\mathrm{fld}^\diamond$ if $\mathrm{Sing}\,\mathrm{Emb}'(P', -)$ — upon passing to $P' \times [0, 1]^k$ as $k \to \infty$ — models a functor out of $\mathcal{M}\mathrm{fld}^\diamond$ corepresented by a pushout.

**Remark 3.15** The analogues of (3-3) — obtained by postcomposing (3-3) with $\beta$ —

(3-8) $$[-2, -1] \times X \hookrightarrow P, \quad [-2, 2] \times W \hookrightarrow P, \quad [1, 2] \times Y \hookrightarrow P,$$

are all smooth, codimension-zero embeddings.

The following is easy to verify; "smoothing then thickening" is isotopy equivalent to "thickening then smoothing." We omit the proof.

**Proposition 3.16** *There are natural isotopy equivalences*

$$P(i_X \times \mathrm{id}_{[0,1]}, i_Y \times \mathrm{id}_{[0,1]}) \to P(i_X, i_Y) \times [0, 1].$$

**Remark 3.17** Proposition 3.16 states that the construction of $P$ is compatible with thickening. So the restriction map along (3-8) fits into a homotopy-commutative diagram of Kan complexes

$$\begin{array}{ccc} \hom_{\mathsf{Mfld}_{\dim P}}(P, -) & \longrightarrow & \hom_{\mathsf{Mfld}_{\dim P}}([-2, -1] \times X, -) \\ \downarrow & & \downarrow \\ \hom_{\mathsf{Mfld}_{\dim P+1}}(P \times [0, 1], - \times [0, 1]) & \longrightarrow & \hom_{\mathsf{Mfld}_{\dim P+1}}([-2, -1] \times X \times [0, 1], - \times [0, 1]) \end{array}$$

and hence passes to the $\infty$-category of thickened manifolds

$$\hom_{\mathcal{M}\mathrm{fld}^\diamond}(P, -) \to \hom_{\mathcal{M}\mathrm{fld}^\diamond}([-2, -1] \times X, -)$$

(and likewise for $[1, 2] \times Y$ and $[-2, 2] \times W$).

**Remark 3.18** (another approach using diffeological spaces) We may also expand our notion of smooth manifold-with-corners, and treat $P'$ as a diffeological space — ie, a space where we know what we mean by a smooth map into $P'$. Indeed, there is a natural notion of smooth map into $P'$. Consider the map

Quadrant $\times \partial W \hookrightarrow P'$ induced by the normal vectors from the previous paragraph. Then we say a map $f : A \to P'$ is smooth if

(i) when pulled back along the inclusion Quadrant $\times \partial W \hookrightarrow P'$, the composition with the map Quadrant $\times \partial W \hookrightarrow \mathbb{R}^2 \times \partial W$ is smooth, and

(ii) $f$ is smooth when pulled back along the complement of $\{-1, 1\} \times W$.

It is easy to see this is independent of choice of normal coordinates.

For this choice of smooth structure on $P'$, one can arrange for our smoothing $P$ from Construction 3.10 to sit into a diagram $P \hookrightarrow P' \hookrightarrow P \hookrightarrow P'$ where every composition is isotopic to the identity morphism — that is, $P \to P$ is isotopic to the identity of $P$ through smooth embeddings, and $P' \to P'$ is isotopic to the identity through diffeologically smooth embeddings. Thus, we may simply treat $P'$ as a smooth manifold isotopy equivalent to an honest smooth manifold $P$ possibly with corners. We do not take this approach here, at the expense of cluttering some formulas with $a$ and $b$.

## 3.3 Reduction to computing a homotopy pullback of spaces

**Notation 3.19** $(I_s)$ For later notational clarity, we let

$$I_s := [0, 1]$$

denote the interval by which we thicken manifolds. The subscript $s$ is to emphasize that we will denote elements of $I_s$ by $s$.

**Choice 3.20** $(\sigma, h_t)$ Here, we endow every interval in $\mathbb{R}$ with the standard orientation inherited from $\mathbb{R}$. We fix

(a) an orientation-preserving diffeomorphism

$$\sigma : [-2, 2] \to I_s,$$

(b) a smooth isotopy of embeddings

$$\{h_t : I_s \to [-2, 2]\}_{t \in [-2,2]}$$

satisfying the following properties:

(i) $h_t = h_{-1}$ for all $t \leq -1$.

(ii) $h_t = h_1$ for all $t \geq 1$. (These first two conditions amount to a collaring condition on the isotopy.)

(iii) $h_{-1}$ is an orientation-preserving diffeomorphism onto the image $[-2, -1]$.

(iv) $h_1$ is an orientation-preserving diffeomorphism onto the image $[1, 2]$.

(v) $h_t$ is strictly increasing for $-1 < t < 1$, meaning that for every $s \in I_s$, the derivative $\frac{\partial}{\partial t} h_t(s)$ is positive.

Using our choice of isotopies $h_t$, we thus have a diagram $\Delta^1 \times \Delta^1 \to N(\mathsf{Mfld}_{\dim P})$ we informally depict as

(3-9)
$$
\begin{array}{ccc}
I_s \times W & \xrightarrow{\ h_1 \times i_Y\ } & [1,2] \times Y \\
\Big\downarrow{\scriptstyle h_{-1} \times i_X} & & \Big\downarrow \\
[-2,-1] \times X & \xrightarrow{\hspace{2.5cm}} & P
\end{array}
$$

Here, the $i_X$, $i_Y$ are as in (3-1), and unlabeled maps are those from (3-8). The diagram above commutes up to the isotopy parametrized by $t \in [-2,2]$ given as

$$
\{I_s \times W \xrightarrow{\ h_t \times \mathrm{id}_W\ } [-2,2] \times W \xrightarrow{\ (3\text{-}8)\ } P\}_{t \in [-2,2]}.
$$

Thus, for any test object $\bullet \in \mathsf{Mfld}^\diamond$, by applying the functor $\hom_{\mathsf{Mfld}^\diamond}(-,\bullet)$ (and invoking Proposition 3.16) we obtain a homotopy commuting diagram of Kan complexes — or, equivalently, a diagram in the $\infty$-category $\mathcal{S}\mathrm{pc}$ — as

(3-10)
$$
\begin{array}{ccc}
\hom_{\mathsf{Mfld}^\diamond}(P,\bullet) & \longrightarrow & \hom_{\mathsf{Mfld}^\diamond}([-2,-1] \times X, \bullet) \\
\Big\downarrow & & \Big\downarrow \\
\hom_{\mathsf{Mfld}^\diamond}([1,2] \times Y, \bullet) & \longrightarrow & \hom_{\mathsf{Mfld}^\diamond}(I_s \times W, \bullet)
\end{array}
$$

**Proposition 3.21** *Suppose that for every test object $\bullet$, the diagram (3-10) is a homotopy pullback diagram of Kan complexes. Then $P$ is a pushout of (3-1) in the $\infty$-category $\mathsf{Mfld}^\diamond$.*

**Proof** If the hypothesis is satisfied, then the image of (3-9) in $\mathsf{Mfld}^\diamond$ is a pushout diagram. (This follows, for example, from Proposition 7.4.5.13 (03BJ) of [8]. In the notation of loc cit, we take the $\infty$-category $\mathcal{C}$ to be the opposite of the nerve of $\mathsf{Mfld}^\diamond$.) On the other hand, we note that the compositions

$$
I_s \xrightarrow{h_1} [1,2] \hookrightarrow [-2,2] \xrightarrow{\sigma} I_s, \quad I_s \xrightarrow{h_{-1}} [-2,-1] \hookrightarrow [-2,2] \xrightarrow{\sigma} I_s
$$

are both isotopic to $\mathrm{id}_{I_s}$. Thus, in $\mathsf{Mfld}^\diamond$ the diagram (3-1) is equivalent to the diagram

$$
[-2,-1] \times X \xleftarrow{\ h_{-1} \times i_X\ } I_s \times W \xrightarrow{\ h_1 \times i_Y\ } [1,2] \times Y,
$$

so $P$ is a pushout of (3-1). $\qquad\square$

## 3.4 A convenient model for the homotopy pullback of spaces

**Notation 3.22** ($\mathbb{H}^{(a)}$ and $f_X$, $f_{W,t}$, $f_Y$) Let us suppose $Z$ is a test object with $\dim Z = \dim X$. (One may always assume this by thickening either $X$ or $Z$.) For every integer $a \geq 0$, we let

$$
\mathbb{H}^{(a)}
$$

denote the simplicial set where a $k$-simplex $f$ in $\mathbb{H}^{(a)}$ is a triplet

$$
(f_X, \{f_{W,t}\}_{t \in [-2,2]}, f_Y)
$$

of maps where

• $f_X \colon \Delta^k \times (X \times I_s^a) \to \Delta^k \times (Z \times I_s^a)$ is a (codimension-zero) smooth embedding respecting the projections to $\Delta^k$,

• $f_Y \colon \Delta^k \times (Y \times I_s^a) \to \Delta^k \times (Z \times I_s^a)$ is a (codimension-zero) smooth embedding respecting the projections to $\Delta^k$, and

• $\{f_{W,t}\}_{t \in [-2,2]}$ is a $t$-parametrized isotopy of embeddings, which we may think of as a smooth map

$$[-2,2]_t \times \Delta^k \times (W \times I_s^a) \to \Delta^k \times (Z \times I_s^a)$$

respecting the projection to $\Delta^k$, and each of whose restrictions to $t \in [-2,2]$ is a codimension-zero smooth embedding.

The triplet must satisfy the following:

(1)  $\{f_{W,t}\}$ is collared in the $t$-variable in the same sense as $h_t$ in Choice 3.20:

(3-11)      $$f_{W,t} = f_{W,-1} \quad \text{for } t \le -1 \qquad \text{and} \qquad f_{W,t} = f_{W,1} \quad \text{for } t \ge 1.$$

(2)  For $t \le -1$, and above every point of $\Delta^k$, the composition

$$W \times I_s^a \xrightarrow{i_X \times \mathrm{id}} X \times I_s^a \xrightarrow{f_X} Z \times I_s^a$$

is equal to the embedding $f_{W,t}$ (evaluated above the point of $\Delta^k$). By the collaring condition, this condition need only be checked at $t = -1$.

(3)  For $t \ge 1$, and above every point of $\Delta^k$, the composition

$$W \times I_s^a \xrightarrow{i_Y \times \mathrm{id}} Y \times I_s^a \xrightarrow{f_Y} Z \times I_s^a$$

is equal to the embedding $f_{W,t}$ (evaluated above the point of $\Delta^k$). By the collaring condition, this condition need only be checked at $t = 1$.

**Remark 3.23**  Because all constructions will respect the simplicial maps between the $\Delta^k$, we will often try to declutter notation by omitting the $\Delta^k$ variable. In practice, such notation should be interpreted to mean that — given a map $\Delta^k \times A \to \Delta^k \times B$ respecting the projections to $\Delta^k$ — we are testing a condition on the induced maps $A \to B$ for every element of $\Delta^k$.

**Notation 3.24**  ($\mathbb{H}^{(\infty)}$)  There are natural thickening maps $\mathbb{H}^{(a)} \to \mathbb{H}^{(a+1)}$ given by taking the direct product of $f_X$, $f_{W,t}$, $f_Y$ with $\mathrm{id}_{I_s}$. We let

$$\mathbb{H}^{(\infty)} := \mathrm{colim}_{a \to \infty} \mathbb{H}^{(a)}$$

denote the increasing union. Because the thickening maps are cofibrations of simplicial sets, this increasing union models the homotopy colimit of simplicial sets.

**Proposition 3.25** $\mathbb{H}^{(a)}$ *admits a map to the homotopy pullback of the diagram of Kan complexes*

$$\text{hom}_{\text{Mfld}_{\dim X + a}}(X \times I_s^a, Z \times I_s^a)$$

(3-12)

$$\text{hom}_{\text{Mfld}_{\dim X + a}}(Y \times I_s^a, Z \times I_s^a) \longrightarrow \text{hom}_{\text{Mfld}_{\dim X + a}}(W \times I_s^a, Z \times I_s^a)$$

*commuting with thickening, and this map is a homotopy equivalence.*

*In particular,* $\mathbb{H}^{(\infty)}$ *is a homotopy pullback of the diagram of Kan complexes*

$$\text{hom}_{\text{Mfld}^\diamond}(X, Z)$$

(3-13)

$$\text{hom}_{\text{Mfld}^\diamond}(Y, Z) \longrightarrow \text{hom}_{\text{Mfld}^\diamond}(W, Z)$$

**Proof** The standard model for the desired homotopy pullback is a simplicial set of triplets $(f_X, F_W, f_Y)$ where $F_W$ is a map $\Delta^1 \to \text{hom}_{\text{Mfld}_{\dim X + a}}(W \times I_s^a, Z \times I_s^a)$, and the evaluation of $F_W$ at $\{0\}, \{1\} \in \Delta^1$ restrict to $f_X$ and $f_Y$ along $h_{-1} \times i_X$ and $h_1 \times i_Y$, respectively.

Choosing an orientation-preserving diffeomorphism $\Delta^1 \cong [-2, 2]_t$ and smoothly retracting the neighborhood $[-2, -1]$ to $\{-2\}$ and $[1, 2]$ to $\{2\}$, we obtain a homotopy equivalence from the simplicial set of $\{f_{W,t}\}_{t \in [-2,2]}$ to the simplicial set of $F_W$, in a way respecting the evaluation maps at $\{0, 1\} \subset \Delta^1$. Thus we have a homotopy equivalence from $\mathbb{H}^{(a)}$ to the simplicial set of triplets $(f_X, F_W, f_Y)$.

The homotopy equivalence from the previous paragraph respects thickening, so we have a map of sequential diagrams from the sequence $\mathbb{H}^{(0)} \to \mathbb{H}^{(1)} \to \mathbb{H}^{(2)} \to \cdots$ to the sequence obtained by taking the homotopy pullbacks of (3-12). This map of sequences is a homotopy equivalence at the $a$th stage for every $a$, so induces a homotopy equivalence of the homotopy colimits.

On the other hand:

- The sequential colimit of (3-12) (as $a \to \infty$) is precisely the diagram (3-13). This is in fact a homotopy colimit because the thickening maps induce cofibrations of mapping spaces.
- Sequential (in fact, filtered) homotopy colimits of Kan complexes commute with homotopy pullbacks (see Remark 3.26).

Thus the induced map from $\mathbb{H}^{(\infty)}$ to the homotopy pullback of (3-13) is a homotopy equivalence of Kan complexes, proving the claim. $\quad\square$

**Remark 3.26** (sequential homotopy colimits commute with homotopy pullbacks) In $\mathcal{S}\text{pc}$, filtered homotopy colimits commute with finite homotopy limits. (This is not true of all $\infty$-categories, of course.) While this is a well-known fact, we will save the reader some trouble and explain why this is true in the special case of a sequential homotopy colimit and homotopy pullbacks. (This is the case we need in the proof of Proposition 3.25.)

Given a sequential diagram $(A_i) = (A_i)_{i \in I}$ (in our setting, $I$ is the linearly ordered set of natural numbers) of simplicial sets, the homotopy colimit may be computed by replacing $(A_i)$ by a projectively cofibrant diagram $(\mathbb{A}_i)$, and computing the honest colimit (eg, increasing union) of the $\mathbb{A}_i$. This replacement can be made functorially, in that a map of diagrams $(A_i) \to (B_i)$ results in a map of replacements $(\mathbb{A}_i) \to (\mathbb{B}_i)$ while respecting compositions of maps of diagrams — see Construction 7.5.6.8 (03CC) of [8].

(Here, by *replacement* we mean the data of weak homotopy equivalences $\mathbb{A}_i \to A_i$ for which the two compositions $\mathbb{A}_i \to A_i \to A_j$ and $\mathbb{A}_i \to \mathbb{A}_j \to A_j$ are equal. By a projectively cofibrant sequential diagram, we mean a diagram such that, for all $i < j$, $\mathbb{A}_i \to \mathbb{A}_j$ is a cofibration — ie, monomorphism, ie, for all $k$, injections on the set of $k$-simplices. See Example 7.5.6.4 (03C7) of [8]. In particular, if $A_i$ already consists of cofibrations, the homotopy colimit is computed as the honest, point-set colimit.)

Recall that a homotopy pullback of a diagram $A \to B \leftarrow C$ of simplicial sets (in the Quillen model structure) is computed as follows: One replaces $A \to B$ by a weak equivalence followed by a Kan fibration $A \xrightarrow{\sim} A' \to B$, and one then computes the point-set pullback of the diagram $A' \to B \leftarrow C$ of Kan complexes. Up to weak homotopy equivalence, this point-set pullback is independent of the choice of the factorization $A \to A' \to B$. Even better, one can arrange for a factorization for simplicial sets for which:

(i) The factorization is functorial, meaning commutativity of the left-hand diagram below guarantees the commutativity of the right-hand diagram

$$
\begin{array}{ccc}
A_i \longrightarrow B_i \\
\downarrow \qquad\quad \downarrow \\
A_j \longrightarrow B_j
\end{array}
\quad\Longrightarrow\quad
\begin{array}{ccccc}
A_i \xrightarrow{\ \sim\ } A'_i \longrightarrow B_i \\
\downarrow \qquad\quad \downarrow \qquad\quad \downarrow \\
A_j \xrightarrow{\ \sim\ } A'_j \longrightarrow B_j
\end{array}
$$

(ii) Cofibrations are preserved, in that if $A_i \to A_j$ is a cofibration, so is the map $A'_i \to A'_j$ above.

So given diagrams $(A_i \to B_i \leftarrow C_i)_{i \in I}$ indexed by a sequential diagram $I$, let $(\mathbb{A}_i \to \mathbb{B}_i \leftarrow \mathbb{C}_i)_{i \in I}$ denote the sequence of diagrams obtained by functorially replacing the diagrams $(A_i)$, $(B_i)$, $(C_i)$ by projectively cofibrant diagrams. Because homotopy fiber products are preserved under homotopy equivalence (Corollary 8.13 of [4]; alternatively, see Remark 7.5.1.3 (03A2) of [8]), the natural-in-$i$ maps

$$\mathrm{holim}(\mathbb{A}_i \to \mathbb{B}_i \leftarrow \mathbb{C}_i) \to \mathrm{holim}(A_i \to B_i \leftarrow C_i)$$

are all weak homotopy equivalences. Now let us functorially replace the maps $\mathbb{A}_i \to \mathbb{B}_i$ by fibrations $\mathbb{A}'_i \to \mathbb{B}_i$, so that the natural-in-$i$ maps

$$\mathrm{holim}(\mathbb{A}_i \to \mathbb{B}_i \leftarrow \mathbb{C}_i) \to \lim(\mathbb{A}'_i \to \mathbb{B}_i \leftarrow \mathbb{C}_i)$$

are weak homotopy equivalences. Now contemplate the induced arrow

$$(3\text{-}14) \qquad \mathrm{colim}_{i \in I} \lim(\mathbb{A}'_i \to \mathbb{B}_i \leftarrow \mathbb{C}_i) \to \lim((\mathrm{colim}_{i \in I} \mathbb{A}'_i) \to \mathrm{colim}_{i \in I}(\mathbb{B}_i) \leftarrow \mathrm{colim}_{i \in I})\mathbb{C}_i)).$$

Limits and colimits of simplicial sets are computed levelwise — for example, fiber products are given on $k$-simplices as $(A' \times_B C)_k = A'_k \times_{B_k} C_k$. It is of course a classical fact that filtered colimits commute with finite limits in sets, so we may conclude sequential colimits and fiber products commute in simplicial sets. In other words, the above arrow (3-14) is an isomorphism.

On the other hand, by (ii) above, the maps $\mathbb{A}'_i \to \mathbb{A}'_j$ are cofibrations, which means they are levelwise injections — in particular, the maps of fiber products $\mathbb{A}'_i \times_{\mathbb{B}_i} \mathbb{C}_i \to \mathbb{A}'_j \times_{\mathbb{B}_j} \mathbb{C}_j$ are all levelwise injections. It follows that the colimits in (3-14) all compute homotopy colimits. On the other hand, we also know that the arrow $\operatorname{colim}_{i \in I} \mathbb{A}'_i \to \operatorname{colim}_{i \in I} \mathbb{B}_i$ is a fibration because fibrations of simplicial sets are preserved under increasing unions. So all the fiber products in (3-14) are homotopy fiber products.

This proves the claim that sequential homotopy colimits and homotopy fiber products commute for simplicial sets.

**Notation 3.27** $(r)$ For every $a \geq 0$ and every manifold $Z$ with $\dim Z = \dim P - 1$, there is a natural map

$$(3\text{-}15) \qquad r : \hom_{\mathrm{Mfld}_{\dim P + a}}(P \times I_s^a, Z \times I_s^{a+1}) \to \mathbb{H}^{(a+1)},$$

which, given an embedding $j : P \times I_s^a \to Z \times I_s^{a+1}$ (or such a collection smoothly indexed by $\Delta^k$) outputs the triplet obtained by precomposing $j$ with the maps

$$X \times I_s^{a+1} \cong (I_s \times X) \times I_s^a \xrightarrow{(h_{-1} \times \mathrm{id}) \times \mathrm{id}} ([-2,-1] \times X) \times I_s^a \xrightarrow{(3\text{-}8) \times \mathrm{id}} P \times I_s^a,$$

$$\{W \times I_s^{a+1} \cong (I_s \times W) \times I_s^a \xrightarrow{(h_t \times \mathrm{id}) \times \mathrm{id}} ([-2,2] \times W) \times I_s^a \xrightarrow{(3\text{-}8) \times \mathrm{id}} P \times I_s^a\}_{t \in [-2,2]},$$

$$Y \times I_s^{a+1} \cong (I_s \times Y) \times I_s^a \xrightarrow{(h_1 \times \mathrm{id}) \times \mathrm{id}} ([1,2] \times Y) \times I_s^a \xrightarrow{(3\text{-}8) \times \mathrm{id}} P \times I_s^a.$$

Informally, $r$ has the effect of substituting the $(u, x)$, $(u, w)$, $(u, y)$ variables with $(\alpha\beta(x), h_u(s_{a+1}))$, $(\alpha\beta(w), h_u(s_{a+1}))$, and $(\alpha\beta(y), h_u(s_{a+1}))$, respectively.

**Remark 3.28** Let us be punctilious about the first isomorphism in each of the maps above. Write $I_s^{a+1} = I_s^{\{1,2,\ldots,a,a+1\}}$ as the set of functions from the ordered set $\{1,\ldots,a+1\}$ to $I_s$. So, for example, we have the isomorphisms

$$I_s^{a+1} = I_s^{\{1,\ldots,a+1\}} \cong I_s^{\{1,\ldots,a\}} \times I^{\{a+1\}}.$$

This allows us to write

$$X \times I_s^{a+1} \cong X \times I_s^a \times I_s^{\{a+1\}} \cong (X \times I_s^{\{a+1\}}) \times I_s^a \cong (I_s^{\{a+1\}} \times X) \times I_s^a \cong (I_s \times X) \times I_s^a$$

(and likewise for $W$ and $Y$). This pedantry makes it clear that $r$ is *not* compatible with thickening using our convention that thickening occurs by multiplying by $I_s$ on the right.

**Remark 3.29** In $\mathrm{Mfld}^\diamond$ there are particular equivalences

$$[-2,-1] \times X \simeq X, \quad [1,2] \times Y \simeq Y, \quad I_s \times W \simeq W$$

given by choosing thickenings of $X$, $Y$, $W$ and permuting coordinates. Choosing such equivalences induces the map (3-17) in the composition below; it is a homotopy equivalence of Kan complexes:

$$(3\text{-}16) \qquad \hom_{\mathsf{Mfld}^\diamond}(P, Z) \to \hom_{\mathsf{Mfld}^\diamond}([-2, -1] \times X, Z) \times^h_{\hom_{\mathsf{Mfld}^\diamond}(I_s \times W, Z)} \hom_{\mathsf{Mfld}^\diamond}([1, 2] \times Y, Z)$$

$$(3\text{-}17) \qquad\qquad\qquad\qquad \to \operatorname{holim} (3\text{-}13)$$

$$\xrightarrow{\text{Proposition 3.25}} \mathbb{H}^{(\infty)}.$$

We thus see (3-16) is a homotopy equivalence if and only if the composition of the above maps in a homotopy equivalence. So consider the composition

$$(3\text{-}18) \qquad\qquad\qquad\qquad \hom_{\mathsf{Mfld}^\diamond}(P, Z) \to \mathbb{H}^{(\infty)}.$$

Our goal is to now prove:

**Lemma 3.30** *For every $Z \in \mathrm{Ob}\,\mathsf{Mfld}^\diamond$, the map* (3-18) *is a homotopy equivalence of Kan complexes.*

## 3.5 Proof of Lemma 3.30

**Choice 3.31** $(\bar\sigma)$ We fix an orientation-*reversing* diffeomorphism

$$(3\text{-}19) \qquad\qquad\qquad\qquad \bar\sigma \colon [-2, 2] \to I_s.$$

**Remark 3.32** In Choice 3.31 we demand $\bar\sigma$ to be orientation reversing. This demand is not strictly necessary; it merely allows us to simplify the proof of Proposition 3.37. The trade-off is that we then need to show that $rqrq$, rather than just $rq$, is homotopic to a thickening map in Proposition 3.41.

**Proposition 3.33** *The map*

$$(3\text{-}20) \qquad\qquad [-2, 2] \times I_s \to [-2, 2] \times I_s, \quad (u, s) \mapsto (h_u(s), \bar\sigma(u)),$$

*is isotopic to the identity. In fact, one can choose an isotopy of* (3-20) *to the identity through maps for which the locus $\{u \leq -1\}$ has image contained in itself, and likewise for the locus $\{u \geq 1\}$.*

**Proof** This follows from our assumptions on $h$ (Choice 3.20). See Figure 3. □

**Notation 3.34** $(q')$ Fix $a \geq 1$, and fix a triplet $(f = f_X, \{f_{W,t}\}_{t \in [-2,2]}, f_Y) \in \mathbb{H}^{(a)}$. We define an embedding

$$q'(f) \colon P' \times I_s^a \to (Z \times I_s^a) \times I_s$$

to be the union of the three compositions

$$(3\text{-}21) \quad [-2, -1] \times (X \times I_s^a) \cong (X \times I_s^a) \times [-2, -1] \xrightarrow{\mathrm{id} \times \bar\sigma} (X \times I_s^a) \times I_s \xrightarrow{f_X \times \mathrm{id}} (Z \times I_s^a) \times I_s,$$

$$(3\text{-}22) \quad [-1, 1] \times (W \times I_s^a) \cong (W \times I_s^a) \times [-1, 1] \xrightarrow{\mathrm{id} \times \bar\sigma} (W \times I_s^a) \times I_s \xrightarrow{(f_{W,\bar\sigma-1}, \mathrm{id})} (Z \times I_s^a) \times I_s,$$

$$(3\text{-}23) \quad [1, 2] \times (Y \times I_s^a) \cong (Y \times I_s^a) \times [-2, -1] \xrightarrow{\mathrm{id} \times \bar\sigma} (Y \times I_s^a) \times I_s \xrightarrow{f_Y \times \mathrm{id}} (Z \times I_s^a) \times I_s.$$
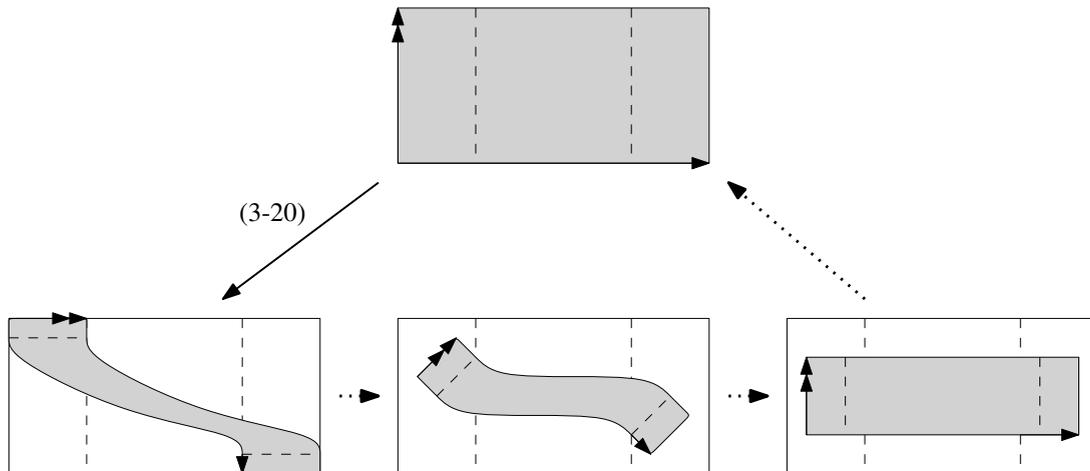
Figure 3: An isotopy of the map (3-20) to the identity. The horizontal direction is the $[-2, 2]_u$ direction, while the vertical direction is the $I_s$ direction, with orientations indicated (in the directions of increasing values of $u$ and of $s$). The dotted arrows indicate isotopies, with the images of each stage of the isotopy indicated as the shaded regions. Note also that the dashed lines indicate the lines $u = \pm 1$ and their images under various stages of the isotopy.

We may define a map out of $P'$ as a union of these three maps by definition of $P'$ (3-2). That the union of these maps is smooth follows from the collaring condition (3-11). That the union is a diffeomorphism follows from the facts that $\bar{\sigma}$ is a diffeomorphism (3-19) and that the maps $f_X$, $f_Y$ and $f_{W,t}$ for each $t$, are embeddings by definition of $\mathbb{H}^{(a)}$.

**Example 3.35** On a triplet $(u, w, \vec{s}) \in [-1, 1] \times (W \times I_s^a)$, (3-22) evaluates as

$$(f_{W,u}(w, \vec{s}), \bar{\sigma}(u)) \in (Z \times I_s^a) \times I_s.$$

**Notation 3.36** $(q)$ We have defined $q'(f)$ as a map with codomain $P' \times I_s^a$. By precomposing $q'(f)$ with the map $\alpha \colon P \to P'$ (3-6), we obtain a map of simplicial sets

$$q \colon \mathbb{H}^{(a)} \to \mathrm{hom}_{\mathrm{Mfld}_{\dim P + a}}(P \times I_s^a, (Z \times I_s^a) \times I_s), \quad f \mapsto (q'(f)) \circ \alpha.$$

(To define $q$ and $q'$ on a $k$-simplex $f$, we follow Remark 3.23 to note that (3-21), (3-22), and (3-23) all make sense as parametrizing functions parametrized over points of the $k$-simplex.)

We now have a sequence of Kan complexes

$$(3\text{-}24) \quad \cdots \xrightarrow{r} \mathbb{H}^{(a)} \xrightarrow{q} \mathrm{hom}_{\mathrm{Mfld}_{\dim P + a}}(P \times I_s^a, Z \times I_s^{a+1})$$
$$\xrightarrow{r} \mathbb{H}^{(a+1)} \xrightarrow{q} \mathrm{hom}_{\mathrm{Mfld}_{\dim P + a + 1}}(P \times I_s^{a+1}, Z \times I_s^{a+2}) \xrightarrow{r} \cdots .$$

**Proposition 3.37**  *For every $a \geq 0$, the composition $qr$ is homotopic to the thickening map. (That is, $qr$ is homotopic to the map of Kan complexes*

$$\hom_{\mathrm{Mfld}_{\dim P+a}}(P \times I_s^a, Z \times I_s^{a+1}) \to \hom_{\mathrm{Mfld}_{\dim P+a+1}}(P \times I_s^{a+1}, Z \times I_s^{a+2}),$$

*defined as in (2-1) by taking the direct products of domain and codomain manifolds with $I_s = [0,1]$ — see Notation 3.19.)*

**Proof of Proposition 3.37**  Fix an embedding $j': P' \times I_s^{a-1} \to Z \times I_s^a$. Consider the restriction of $q'r(j' \circ \alpha)$ along the map

$$(3\text{-}25) \qquad\qquad ([-2,2] \times W \times I_s^{a-1}) \times I_s \xrightarrow{((3\text{-}3)\times\mathrm{id})\times\mathrm{id}} (P' \times I_s^{a-1}) \times I_s.$$

Parsing through the definitions — (3-15), (3-21), (3-22), and (3-23) — we see this restriction is obtained by the composition

$$(3\text{-}26) \qquad\qquad [-2,2] \times (W \times I_s^a) \cong (W \times I_s^a) \times [-2,2]$$

$$\xrightarrow{\mathrm{id}\times\bar{\sigma}} (W \times I_s^a) \times I_t$$

$$(3\text{-}27) \qquad\qquad \cong ((I_s^{\{a\}} \times W) \times I_s^{\{1,\dots,a-1\}}) \times I_t$$

$$(3\text{-}28) \qquad\qquad \xrightarrow{(((h_{\bar{\sigma}^{-1}(t)}\times\mathrm{id})\times\mathrm{id}),\mathrm{id})} \big(([-2,2] \times W) \times I_s^{a-1}\big) \times I_t$$

$$(3\text{-}29) \qquad\qquad \xrightarrow{((3\text{-}3)\times\mathrm{id})\times\mathrm{id}} (P' \times I_s^{a-1}) \times I_t$$

$$\xrightarrow{(\beta\times\mathrm{id})\times\mathrm{id}} (P \times I_s^{a-1}) \times I_t$$

$$\xrightarrow{(j'\circ(\alpha\times\mathrm{id}))\times\mathrm{id}} (Z \times I_s^{a-1}) \times I_t.$$

In the above notation, the intervals $I_s = I_t$ are the same set. The subscripts indicate that (in the relevant lines) we use $t$ to denote a variable representing the elements of the interval $I_t$, to remove the ambiguity regarding the copy of $I$ on which $h_{\bar{\sigma}^{-1}(t)}$ depends. Having disambiguated, we will henceforth revert to writing $I_s$ for all thickening intervals.

In (3-27), we have used notation from Remark 3.28.

In fact, by replacing all instances of $W$ by $X$, and all instances of $[-2,2]$ by $[-2,-1]$ — which has the effect of turning $h_{\sigma^{-1}(t)}$ into $h_{-1}$ — the above lines also describe the restriction of $q'r(j' \circ \alpha)$ along

$$([-2,-1] \times X \times I_s^{a-1}) \times I_s \xrightarrow{((3\text{-}3)\times\mathrm{id})\times\mathrm{id}} (P' \times I_s^{a-1}) \times I_s,$$

and likewise for $Y$ (by replacing the intervals with $[1,2]$).

Now consider the compositions from line (3-26) through line (3-28) for each of $X, W, Y$ (and their relevant intervals $[-2,-1]$, $[-1,1]$ and $[1,2]$, respectively). The compositions define self-embeddings

$$\eta_X, \quad \eta_W, \quad \eta_Y$$

of the manifolds

$$([-2,-1] \times X \times I_s^{a-1}) \times I_s, \quad ([-2,2] \times W \times I_s^{a-1}) \times I_s, \quad ([1,2] \times Y \times I_s^{a-1}) \times I_s.$$

Each of these self-embeddings is a direct product with the identity embeddings of $X, W, Y$, respectively. In particular, the maps $\eta_X, \eta_W, \eta_Y$ glue together to form a single embedding

$$\eta \colon (P' \times I_s^{a-1}) \times I_s \to (P' \times I_s^{a-1}) \times I_s.$$

On the interval components — ie, ignoring the $P'$ factor — $\eta$ acts as

$$[-2,2] \times I_s^a \xrightarrow{\cong} [-2,2] \times I_s^{a-1} \times I, \quad (u, s_1, \ldots, s_a) \mapsto (h_u(s_a), (s_1, \ldots, s_{a-1}), \bar{\sigma}(u)).$$

In other words, the composition acts trivially on all but two factors: It acts on the $u$ and $s_a$ components precisely by the map (3-20) from Proposition 3.33. So the isotopy guaranteed by Proposition 3.33 induces isotopies

$$\eta_X \sim \mathrm{id}, \quad \eta_W \sim \mathrm{id}, \quad \eta_Y \sim \mathrm{id}.$$

(The condition that the isotopy preserves the loci $u \leq -1, u \geq 1$ guarantees that $\eta_X$ is isotoped through $([-2,-1] \times X \times I_s^{a-1}) \times I_s$, and likewise for $Y$.) These isotopies glue together to exhibit an isotopy

(3-30) $$\eta \sim \mathrm{id}.$$

At this point we can appreciate that gluing together the $W$, $X$, $Y$ versions of (3-29) yields the identity map of $P'$. This means

$$q'r(j' \circ \alpha) = \big( (j' \circ (\alpha \times \mathrm{id}_{I_s^{a-1}}) \circ (\beta \times \mathrm{id}_{I_s^{a-1}})) \times \mathrm{id} \big) \circ \eta$$

and applying the isotopy (3-30) yields an isotopy

$$q'r(j' \circ \alpha) \sim (j' \circ (\alpha \times \mathrm{id}_{I_s^{a-1}}) \circ (\beta \times \mathrm{id}_{I_s^{a-1}})) \times \mathrm{id},$$

which, because $\alpha\beta$ is isotopic to id, is isotopic to $j' \times \mathrm{id}$. By varying $j'$, we find that the diagram of Kan complexes

$$\begin{array}{ccccc}
\hom_{\mathrm{Mfld}_{\dim P + a - 1}}(P \times I_s^{a-1}, Z \times I_s^a) & \xrightarrow{r} & \mathbb{H}^{(a)} & \xrightarrow{q} & \hom_{\mathrm{Mfld}_{\dim P + a}}(P \times I_s^a, Z \times I_s^{a+1}) \\
\alpha^* \uparrow & & & \searrow{q'} & \alpha^* \uparrow \\
\mathrm{Sing\,Emb}'(P', \times I_s^{a-1}, Z \times I_s^a) & \xrightarrow{\quad -\times\mathrm{id}_{I_s}\quad} & & & \mathrm{Sing\,Emb}'(P' \times I_s^a, Z \times I_s^{a+1})
\end{array}$$

commutes up to homotopy. (In fact, the triangle on the right commutes on the nose.) By noting that $\alpha^*$ is a homotopy equivalence (3-7) that respects thickening up to homotopy, we are finished. $\qquad\square$

**Notation 3.38** $(H)$ We will now compute the iterate $rqrq(f)$. For this, we introduce the following function for sake of brevity:

$$H \colon I_s \times I_s \times [-2,2]_t \to [-2,2]_t, \quad H(s_{a+1}, s_{a+2}, t) = h_{h_t(s_{a+2})}(s_{a+1}).$$

**Remark 3.39** By the collaring condition on $h_t$ (Choice 3.20), we have

$$H = \begin{cases} h_{-1}(s_{a+1}), & t \in [-2, -1], \\ h_1(s_{a+1}), & t \in [1, 2]. \end{cases}$$

**Proposition 3.40** *The self-map of $[-2, 2]_t \times I_s \times I_s$ given by*

(3-31) $$(t, s_{a+1}, s_{a+2}) \mapsto (H, \overline{\sigma} H, \overline{\sigma} h_t(s_{a+2}))$$

*is homotopic to the identity. Moreover, this homotopy $\{k_\tau\}_{\tau \in [0,1]}$ can be chosen through smooth maps*

$$k_\tau : [-2, 2]_t \times I_s \times I_s \to [-2, 2]_t \times I_s \times I_s$$

*such that*:

(a) *For any value of $t$ and $\tau$, the composition*

(3-32) $$\{t\} \times I_s \times I_s \xrightarrow{k_\tau} [-2, 2]_t \times I_s \times I_s \to I_s \times I_s$$

*is a (codimension-zero) smooth embedding. Here, the first arrow is the restriction of $k_\tau$ to the indicated domain, and the last map is the projection forgetting the $[-2, 2]$ factor.*

(b) *For every value of $\tau$, the composition (3-32) is collared in the $t$-variable. This means that for all $t \leq -1$, the projections of $k_\tau(t, -, -)$ and $k_\tau(-1, -, -)$ to $I_s \times I_s$ are equal. Likewise, for all $t \geq 1$, the projections of $k_\tau(t, -, -)$ and $k_\tau(1, -, -)$ to $I_s \times I_s$ are equal. (However, their projections to the $t$ variable may differ.)*

(c) *For every value of $\tau$, $k_\tau$ restricts to a self-map of $[-2, -1]_t \times I_s \times I_s$ — meaning the image of $[-2, -1]_t \times I_s \times I_s$ under $k_\tau$ is contained in $[-2, -1]_t \times I_s \times I_s$ — and to a self-map of $[1, 2]_t \times I_s \times I_s$.*

**Proof** Let us first focus on the first factor of (3-31), ie, the function $H$ taking value in the $[-2, 2]_t$ component of the image. The conditions (a) and (b) are preserved by homotopies $\{k'_\tau\}_\tau$ that only affect this first factor. On the other hand, it is straightforward to produce a homotopy $\{k'_\tau\}_\tau$ without changing the latter two factors of (3-31), smoothly homotoping the first factor from $H$ to $t$, while preserving (c). Choose such a $\{k'_\tau\}_\tau$; we have homotoped (3-31) to the map

$$(t, s_{a+1}, s_{a+2}) \mapsto (t, \overline{\sigma} H, \overline{\sigma} h_t(s_{a+2})).$$

Now, for every fixed $t$, it is possible to *isotope* the map $(\overline{\sigma} H, \overline{\sigma} h_t(s_{a+2}))$ to the identity map of $I_s \times I_s$. It is clear one can do this smoothly in $t$, so choose such isotopies smoothly in $t$, and constantly along the intervals $-2 \leq t \leq -1$ and $1 \leq t \leq 2$. Such a choice yields a smooth homotopy $\{k''_\tau\}_\tau$ (whose projection to the interval $[-2, 2]_t$ is $\tau$-independent) preserving all the conditions required by the proposition.

Concatenating $\{k''_\tau\}_\tau$ after $\{k'_\tau\}_\tau$, we have obtained the desired homotopy $\{k_\tau\}$. □

**Proposition 3.41** *For every $a \geq 0$, the composition $rqrq$ is homotopic to the square of the thickening map — ie, homotopic to the map $\mathbb{H}^{(a)} \to \mathbb{H}^{(a+2)}$ obtained by performing twice the thickening map in Notation 3.24.*

**Proof of Proposition 3.41**  Let us fix $f \in \mathbb{H}^{(a)}$ and compute the triplet $rq(f)$. For brevity, we let

$$\vec{s} := (s_1, \ldots, s_a) \in I_s^a.$$

The map $q(f)$ sends a point $(p, \vec{s}) \in P \times I_s^a$ to

$$\begin{cases} (f_X(x, \vec{s}), \overline{\sigma}(u)) & \text{if } p = (u, x) \in [-2, -1] \times X, \\ (f_{W,u}(w, \vec{s}), \overline{\sigma}(u)) & \text{if } p = (u, w) \in [-2, 2] \times W, \\ (f_Y(y, \vec{s}), \overline{\sigma}(u)) & \text{if } p = (u, y) \in [1, 2] \times Y. \end{cases}$$

(For points $p$ in $[-2, -1] \times W$, the above is well defined thanks to the collaring conditions on $h$ and on $f_{W,t}$; likewise for points in $[1, 2] \times W$.) For every $t \in [-2, 2]$, we have that

$$(rq(f))_{W,t} : W \times I_s^{a+1} \to (Z \times I_s^a) \times I_s$$

is given by the formula

$$(w, \vec{s}, s_{a+1}) \mapsto \big( f_{W, h_t(s_{a+1})}(\alpha_{h_t(s_{a+1})} \beta_{h_t(s_{a+1})}(w), \vec{s}), \overline{\sigma} h_t(s_{a+1}) \big).$$

Let us explain this notation. Clearly $\alpha$ respects the $u$-coordinate of $P'$, as does $\beta$ (Remark 3.13). Thus it makes sense to fix a $u$-coordinate such as $h_t(s_{a+1})$ and restrict $\alpha$ and $\beta$ as functions with domain and codomain given by fibers above this $u$-value. In fact, $\alpha_u$ is always the identity, so we may omit $\alpha$ in much of what follows.

Likewise, the maps

$$(rq(f))_X : X \times I_s^{a+1} \to (Z \times I_s^a) \times I_s \quad \text{and} \quad (rq(f))_Y : Y \times I_s^{a+1} \to (Z \times I_s^a) \times I_s$$

have formulas

$$(x, \vec{s}, s_{a+1}) \mapsto \big( f_X(\beta_{h_{-1}(s_{a+1})}(x), \vec{s}), \overline{\sigma} h_{-1}(s_{a+1}) \big)$$

and

$$(y, \vec{s}, s_{a+1}) \mapsto \big( f_Y(\beta_{h_1(s_{a+1})}(y), \vec{s}), \overline{\sigma} h_1(s_{a+1}) \big),$$

respectively. In writing the formula for $(rq(f))_X$, we have used that $h_{-1}(s_{a+1}) \in [-2, -1]$ and that $f_{W,t} = f_{W,-1} = f_X \circ i_W$ for $t \in [-2, -1]$; and likewise for the formula for $(rq(f))_Y$.

Now that we have computed $rq$, we may iterate to compute the following.

- $rqrq(f)_X$ sends an element $(x, \vec{s}, s_{a+1}, s_{a+2}) \in X \times I_s^{a+2}$ to
$$\big( f_X(\beta_{h_{-1}(s_{a+1})} \beta_{h_{-1}(s_{a+2})}(x), \vec{s}), \overline{\sigma} h_{-1}(s_{a+1}), \overline{\sigma} h_{-1}(s_{a+2}) \big).$$

- $rqrq(f)_Y$ sends an element $(y, \vec{s}, s_{a+1}, s_{a+2}) \in Y \times I_s^{a+2}$ to
$$\big( f_Y(\beta_{h_1(s_{a+1})} \beta_{h_1(s_{a+2})}(y), \vec{s}), \overline{\sigma} h_1(s_{a+1}), \overline{\sigma} h_1(s_{a+2}) \big).$$

- For all $t \in [-2, 2]$, $rqrq(f)_{W,t}$ sends an element $(w, \vec{s}, s_{a+1}, s_{a+2})$ to
$$\big( f_{W,H}(\beta_H \beta_{h_t(s_{a+2})}(w), \vec{s}), \overline{\sigma} H, \overline{\sigma} h_t(s_{a+2}) \big) \in (Z \times I_s^a) \times I_s \times I_s.$$

Thus, each of these maps can be understood as a composition

$$(3\text{-}33) \quad [-2,2]_t \times \bullet \times I_s^a \times I_s \times I_s \xrightarrow{(3\text{-}34)} ([-2,2]_t \times \bullet \times I_s^a) \times I_s \times I_s \xrightarrow{(f_{\bullet,t}) \times \mathrm{id} \times \mathrm{id}} (Z \times I_s^a) \times I_s \times I_s,$$

where the first map is given in coordinates by

$$(3\text{-}34) \qquad (t, \bullet, \vec{s}, s_{a+1}, s_{a+1}) \mapsto (H, \beta_H \beta_{h_t(s_{a+1})}(\bullet), \vec{s}, \overline{\sigma} H, \overline{\sigma} h_t(s_{a+2})).$$

Here, the $\bullet$ is a stand-in for $X, W, Y$ (or $x, w, y$ in the coordinate formula), it is understood that $f_{X,t} = f_X$ and $f_{Y,t} = f_Y$, and we only evaluate the first map on those elements of $[-2,2] \times \bullet$ that are elements of $P'$.

Because $\beta_u$ is a flow by a vector field (hence isotopic to the identity) we may find a path in $\mathbb{H}^{(a+2)}$ from (3-34) to the map

$$(t, \bullet, \vec{s}, s_{a+1}, s_{a+1}) \mapsto (H, \bullet, \vec{s}, \overline{\sigma} H, \overline{\sigma} h_t(s_{a+2})).$$

By Proposition 3.40, this map enjoys a path in $\mathbb{H}^{(a+2)}$ to the (identity) map

$$(t, \bullet, \vec{s}, s_{a+1}, s_{a+1}) \mapsto (t, \bullet, \vec{s}, s_{a+1}, s_{a+2}).$$

Thus, postcomposing the resulting isotopy from (3-34) to id with $(f_{\bullet,t}) \times \mathrm{id} \times \mathrm{id}$, we witness an isotopy from (3-33) to a two-fold thickening of $f \in \mathbb{H}^{(a)}$. Because the isotopy between $rqrqf$ and $f \times \mathrm{id} \times \mathrm{id}$ is witnessed by precomposing $f \times \mathrm{id} \times \mathrm{id}$ by a series of isotopies independent of $f$, the claim is proven. $\square$

**Proof of Lemma 3.30**  The sequential diagram (3-24) has a cofinal subdiagram

$$(3\text{-}35) \quad \cdots \xrightarrow{qr} \hom_{\mathrm{Mfld}_{\dim P + a}}(P \times I_s^a, Z \times I_s^{a+1}) \xrightarrow{qr} \hom_{\mathrm{Mfld}_{\dim P + a + 1}}(P \times I_s^a, Z \times I_s^{a+2}) \xrightarrow{qr} \cdots.$$

By Proposition 3.37, this sequence is equivalent to the sequence of thickenings, hence has colimit given by $\hom_{\mathrm{Mfld}^\diamond}(P, Z)$.

There is another cofinal subdiagram of the form

$$(3\text{-}36) \qquad\qquad \cdots \xrightarrow{rqrq} \mathbb{H}^{(a)} \xrightarrow{rqrq} \mathbb{H}^{(a+2)} \xrightarrow{rqrq} \cdots.$$

(In fact, there are two such — one could take $a$ to be odd or even. The choice is immaterial to us.) By Proposition 3.41, this sequence is equivalent to the sequence with the same objects, with maps given by (double) thickenings. Thus this subdiagram has colimit given by $\mathbb{H}^{(\infty)}$.

It is now straightforward to check that the induced map on homotopy groups

$$\pi_* \hom_{\mathrm{Mfld}^\diamond}(P, Z) \cong \pi_* \operatorname{hocolim}(3\text{-}35) \xrightarrow{\cong} \pi_* \operatorname{hocolim}(3\text{-}24) \xdashrightarrow{\cong} \pi_* \operatorname{hocolim}(3\text{-}36) \cong \pi_* \mathbb{H}^{(\infty)}$$

(where the dashed arrow is the inverse to the isomorphism induced by the cofinality of (3-36)) is the same map as that induced on homotopy groups by (3-18). $\square$

# 4 Proofs of the main results

## 4.1 Finite colimits: existence and preservation

**Proof of Theorem 1.8** It is clear that $\mathcal{M}\mathrm{fld}^\diamond$ has an initial object — the empty manifold — so to see $\mathcal{M}\mathrm{fld}^\diamond$ has all finite colimits, it suffices to prove that $\mathcal{M}\mathrm{fld}^\diamond$ has pushouts (Corollary 4.4.2.4 of [7]).

Given two morphisms as in (3-1), Proposition 3.21 shows that $P$ is a pushout if and only if the map (3-16) is a homotopy equivalence. By Remark 3.29, we are reduced to showing that (3-18) is a homotopy equivalence. This is Lemma 3.30. □

**Remark 4.1** Let $\mathcal{C}$ be an $\infty$-category and fix an object $B \in \mathrm{Ob}\,\mathcal{C}$. Fix a functor $f\colon D \to \mathcal{C}_{/B}$ to the slice $\infty$-category. Then a diagram $D^\triangleright \to \mathcal{C}_{/B}$ is a colimit diagram if and only if the composition $D^\triangleright \to \mathcal{C}_{/B} \to \mathcal{C}$ is a colimit diagram (see, for example, Corollary 7.1.3.20 (02KC) of [8]).

**Proposition 4.2** *The functor* (1-2) *preserves finite colimits.*

**Proof** The isotopy-commuting diagram of smooth manifolds (3-9) — because it consists of codimension-zero embeddings — induces a homotopy-commuting diagram of frame bundles. By quotienting by the free $O$ action, we thus obtain a homotopy-commuting diagram of spaces over BO. (It may help, or confuse, the reader that this resulting diagram can be notated identically to (3-9).) It is classical that (3-9) is in fact a colimit diagram in the $\infty$-category of topological spaces. (After all, $P$ models a homotopy pushout of (3-1), as $P$ is homotopy equivalent to (3-2).) On the other hand, a diagram in the slice over-category $\mathrm{Top}_{/\mathrm{BO}}$ is a colimit diagram if and only if its image in $\mathrm{Top}$ is (Remark 4.1).

This shows that (1-2) preserves pushouts. The functor also preserves initial objects, as it sends the empty manifold to the empty space. Thus the functor preserves all finite colimits (Corollary 4.4.2.5 of [7]). □

## 4.2 The point generates thickened manifolds

**Proposition 4.3** (handle attachments are pushouts) *Let $X$ be a manifold of dimension $d$, possibly with boundary, and let $X'$ be obtained from $X$ by attaching an index-$k$ handle along $\partial X$. Then the induced diagram $\Delta^1 \times \Delta^1 \to \mathcal{M}\mathrm{fld}^\diamond$, which we informally draw as*

$$
\begin{array}{ccc}
S^{k-1} \times D^{d-k} & \longrightarrow & X \\
\downarrow & & \downarrow \\
D^k \times D^{d-k} & \longrightarrow & X'
\end{array}
$$

*is a pushout square.* (*The top horizontal arrow is modeled as the codimension-zero embedding of $(S^{k-1} \times D^{d-k}) \times [0,1]$ into a collar neighborhood of $\partial X \subset X$.*)

**Proof** We may assume there exists a Morse function $f: X' \to \mathbb{R}$ realizing the handle attachment—concretely, $f^{-1}[b - \delta, b + \delta]$ is the attaching handle $D^k \times D^{n-k}$ inside $X'$, we assume $b$ is the lone critical value in $[b - \delta, \infty]$, and that $X = f^{-1}(-\infty, b - \delta)$. Using the induced map

$$X' \times [-2, 2]_v \xrightarrow{f \times p_v} \mathbb{R} \times [-2, 2]_v$$

one may describe the space $P'$ from (3-4) as the subspace of $X' \times [-2, 2]_v$ given by the union of the three subsets

- $\{(x, v) \in X' \times [-2, 2]_v \mid f(x) \leq b - \delta/2, \ v \leq 1\}$,
- $\{(x, v) \in X' \times [-2, 2]_v \mid f(x) \in [b - \delta, b + \delta], \ v \in [1 - \epsilon, 1 + \epsilon]\}$, and
- $\{(x, v) \in X' \times [-2, 2]_v \mid f(x) \geq b + \delta/2, \ v \geq 1\}$.

It is then clear that the inclusion of the space $P$ (Construction 3.10) into $X' \times [-2, 2]$ is an isotopy equivalence. Because $X' \simeq X' \times [-2, 2]$ in $\mathfrak{Mfld}^\diamond$ and $P$ was already shown to be a pushout, the rest is straightforward. $\square$

The following is a categorical manifestation of the geometric tautology that manifolds are made out of disks:

**Lemma 4.4** *The point generates $\mathfrak{Mfld}^\diamond$ under finite colimits.*

**Proof of Lemma 4.4** Let $X$ be a compact manifold, possibly with corners and boundary. By Remark 2.3 we may assume $X$ has no corners, but possibly nonempty boundary. Then $X$ admits a Morse function $f: X \to \mathbb{R}$ for which $X = f^{-1}[-\infty, b]$ and $f^{-1}(b) = \partial X$, with $b$ a regular value. (See, for example, Theorem 2.5 of [13]. Using the notation there, one considers $V_0 = \varnothing$.) In the usual way, $f$ creates a filtration of $X$

$$X = X_N \supset X_{N-1} \supset \cdots \supset X_1 \supset X_0 \supset X_{-1} = \varnothing,$$

where each $X_i$ is obtained by attaching a single handle of index $a_i$ to $X_{i-1}$. By Proposition 4.3, this realizes $X_i$ as a pushout of $X_{i-1}$, a (thickened) sphere, and a (thickened) point. On the other hand, by induction on $k$, Proposition 4.3 also shows that any sphere is generated in $\mathfrak{Mfld}^\diamond$ under finite colimits by a point. $\square$

## 4.3 Fully faithful

**Proposition 4.5** *For any smooth, compact manifold $X$ (possibly with corners) there exists a natural homotopy equivalence of Kan complexes*

(4-1) $$\hom_{\mathsf{Mfld}^\diamond}(\mathrm{pt}, X) \to \mathrm{Sing}\,\mathrm{Fr}(X).$$

*Here, $\mathrm{Sing}\,\mathrm{Fr}(X)$ is the Kan complex of singular chains associated to the stable frame bundle $\mathrm{Fr}(X)$, and $\mathrm{Fr}$ is modeled as in Remark 2.11.*

**Proof** Set $d = \dim X$. The space of smooth embeddings $[0,1]^d \to X$ fits into a fiber sequence

$$\mathrm{GL}(\mathbb{R}^d, T_{\mathrm{ev}_0}X) \to \mathrm{Emb}([0,1]^d, X) \xrightarrow{\mathrm{ev}_0} X,$$

where the last map is the evaluation at the origin $0 \in [0,1]^d$, and the fiber is identified with the space of linear isomorphisms from the tangent space $\mathbb{R}^d \cong T_0[0,1]^d$ to $T_{\mathrm{ev}_0}X$. (If $X$ has boundary or corners, we may replace $X$ by its interior to obtain a homotopy-equivalent fiber sequence.) It follows that the map

$$\mathrm{Emb}([0,1]^d, X) \to \mathrm{Fr}_d(X)$$

to the frame bundle — given by sending an embedding $j \colon [0,1]^d \to X$ to the derivative-induced framing $\mathbb{R}^d \cong T_0[0,1]^d \cong T_{j(0)}X$ — is a map of fibrations over $X$ (or over the interior of $X$) with homotopy equivalent fibers, hence a homotopy equivalence. By taking the colimit of the induced maps of Kan complexes

$$\mathrm{Sing}\,\mathrm{Emb}([0,1]^{d+k}, X \times [-1,1]^k) \to \mathrm{Sing}\,\mathrm{Fr}_{d+k}(X)$$

as $k \to \infty$, and noting that the transition maps from $k$ to $k+1$ are all cofibrations, we find that the induced map (4-1) is indeed a homotopy equivalence of Kan complexes.

Now, given any smooth codimension-zero embedding $X \times [0,1]^k \to X' \times [0,1]^{k'}$ (note this forces $d + k = d' + k'$), the diagram of Kan complexes below commutes:

$$
\begin{array}{ccc}
\hom_{\mathcal{M}\mathrm{fld}^\diamond}([0,1]^{d+k}, X \times [0,1]^k) & \longrightarrow & \mathrm{Sing}\,\mathrm{Fr}_{d+k}(X \times [0,1]^k) \\
\downarrow & & \downarrow \\
\hom_{\mathcal{M}\mathrm{fld}^\diamond}([0,1]^{d'+k'}, X' \times [0,1]^{k'}) & \longrightarrow & \mathrm{Sing}\,\mathrm{Fr}_{d'+k'}(X' \times [0,1]^{k'})
\end{array}
$$

These maps are compatible with thickening, and is natural with respect to families of codimension-zero embeddings. So the equivalence (4-1) is indeed natural. $\square$

**Proposition 4.6** *For every object $X \in \mathcal{M}\mathrm{fld}^\diamond$, the map $\hom_{\mathcal{M}\mathrm{fld}^\diamond}(\mathrm{pt}, X) \to \hom_{\mathcal{S}\mathrm{pc}_{/\mathrm{BO}}}(\mathrm{pt}, X)$ given by (1-2) is a homotopy equivalence.*

**Proof** Consider the functor (2-4). If $X$ is modeled by a manifold of dimension $d$, a smooth embedding $j \colon [0,1]^d \to X$ is sent to the GL-equivariant map $\mathrm{Fr}([0,1]^d) \to \mathrm{Fr}(X)$ induced by (the derivative of) $j$. Restriction of $j$ to the origin is compatible with the equivalence (4-1), in the sense that the following diagram commutes up to homotopy:

$$
\begin{array}{ccc}
\hom_{\mathcal{M}\mathrm{fld}^\diamond}(\mathrm{pt}, X) & \xrightarrow{\;(4\text{-}1)\;} & \mathrm{Sing}\,\mathrm{Fr}(X) \\
{\scriptstyle(2\text{-}4)}\big\downarrow & \nearrow{\scriptstyle \mathrm{ev}_{0,\mathrm{id}_{\mathbb{R}^d}}} & \\
\hom_{\mathrm{Top}^{\mathrm{GL}}}\big(\mathrm{Fr}([0,1]^d), \mathrm{Fr}(X)\big) & &
\end{array}
$$

Here, $ev_{0,\mathrm{id}_{\mathbb{R}^d}}$ is the evaluation of an GL-equivariant map at the canonical frame of the origin of $[0,1]^d$. This evaluation map is a homotopy equivalence, as it is a map of fiber bundles with homotopy equivalent base spaces (equivalent to $X$) and homotopy equivalent fibers (equivalent to GL).

Now we finish by recalling the fact that mapping spaces in the $\infty$-category $\mathcal{S}\mathrm{pc}_{/\mathrm{BGL}} \simeq N(\mathsf{Top}^{\mathrm{GL}})$ may be computed as equivariant mapping spaces between free GL-spaces, and that (1-2) is defined by precomposing with (2-4). The equivalences (2-6) and (2-7) finish the job.                          □

**Lemma 4.7**  *The functor* (1-2) *is fully faithful.*

**Proof of Lemma 4.7**  Fix smooth manifolds $W$ and $X$. Up to thickening, $W$ is written as a finite colimit of a diagram involving only disks — ie, only the point, up to thickening. (This is the content of Lemma 4.4.) So write $W \simeq \mathrm{colim}_{\mathcal{D}}\,\mathrm{pt}$ where $\mathcal{D}$ is some finite diagram. We have the following homotopy commuting diagram of Kan complexes:

$$\begin{array}{ccccc}
\hom_{\mathcal{M}\mathrm{fld}^\diamond}(W,X) & \xleftarrow{\;\simeq\;} & \hom_{\mathcal{M}\mathrm{fld}^\diamond}(\mathrm{colim}_{\mathcal{D}}\,\mathrm{pt},X) & \xrightarrow{\;\simeq\;} & \mathrm{holim}_{\mathcal{D}}\,\hom_{\mathcal{M}\mathrm{fld}^\diamond}(\mathrm{pt},X) \\
\downarrow & & \downarrow & & \downarrow \\
\hom_{\mathcal{S}\mathrm{pc}_{/\mathrm{BO}}}(W,X) & \xleftarrow{\;\simeq\;} & \hom_{\mathcal{S}\mathrm{pc}_{/\mathrm{BO}}}(\mathrm{colim}_{\mathcal{D}}\,\mathrm{pt},X) & \xrightarrow{\;\simeq\;} & \mathrm{holim}_{\mathcal{D}}\,\hom_{\mathcal{S}\mathrm{pc}_{/\mathrm{BO}}}(\mathrm{pt},X)
\end{array}$$

All vertical arrows are obtained from the functor (1-2).

The upper-left horizontal arrow is a homotopy equivalence by the hypotheses that $W \simeq \mathrm{colim}_{\mathcal{D}}\,\mathrm{pt}$; the two horizontal arrows on the right are equivalences by the definition of colimit. We use Proposition 4.2 to conclude the lower-left horizontal arrow is an equivalence. By Proposition 4.6, the rightmost vertical arrow is a homotopy limit of equivalences — hence an equivalence. It follows that the leftmost vertical arrow is an equivalence, establishing the claim.                          □

## 4.4  The equivalence (proof of Theorem 1.3)

**Proof of Theorem 1.3**  The functor (1-2) is fully faithful by Lemma 4.7. It remains to show it is essentially surjective.

By definition, the $\infty$-category of finite spaces $\mathcal{S}\mathrm{pc}^{\mathrm{finite}} \subset \mathcal{S}\mathrm{pc}$ is the full subcategory generated by a point under finite colimits. (For those readers who prefer other models: $\mathcal{S}\mathrm{pc}^{\mathrm{finite}}$ is equivalent to the $\infty$-category of spaces homotopy equivalent to finite CW complexes. Noting cell attachments are homotopy pushouts along maps $S^{n-1} \to D^n \simeq \mathrm{pt}$ and by applying induction on dimensions of cells, we conclude that all finite CW complexes are generated under finite colimits by a point.)

Thus, given an object $X \to \mathrm{BO}$ of $\mathcal{S}\mathrm{pc}^{\mathrm{finite}}_{/\mathrm{BO}}$, we have a constant finite diagram $f: D \to \mathcal{S}\mathrm{pc}^{\mathrm{finite}}$ (with value pt) having colimit $X$. Choose a functor $f^{\triangleright}: D^{\triangleright} \to \mathcal{S}\mathrm{pc}^{\mathrm{finite}}$ exhibiting $X$ as the colimit. Composition

with the map $X \to \mathrm{BO}$ lifts $f^{\triangleright}$ to a diagram $(\tilde{f})^{\triangleright} \colon D^{\triangleright} \to \mathcal{S}\mathrm{pc}^{\mathrm{finite}}_{/\mathrm{BO}}$. By Remark 4.1, $(\tilde{f})^{\triangleright}$ is a colimit diagram, hence we see that $X \to \mathrm{BO}$ is in the subcategory of $\mathcal{S}\mathrm{pc}^{\mathrm{finite}}_{/\mathrm{BO}}$ generated by the object $\mathrm{pt} \to \mathrm{BO}$.

The functor (1-2) contains the point $\mathrm{pt} \to \mathrm{BO}$ in its image, is fully faithful, and preserves all finite colimits by Proposition 4.2. Thus it is essentially surjective.      $\square$

# References

[1] **D Ayala**, **J Francis**, **H L Tanaka**, *Local structures on stratified spaces*, Adv. Math. 307 (2017) 903–1028 MR

[2] **A Douady**, **L Hérault**, *Arrondissement des variétés à coins* (1973) Appendix to A Borel and J-P Serre, "Corners and arithmetic groups", Comment. Math. Helv. 48 (1973) 436–491

[3] **E Dror**, **W G Dwyer**, **D M Kan**, *Equivariant maps which are self homotopy equivalences*, Proc. Amer. Math. Soc. 80 (1980) 670–672 MR

[4] **P G Goerss**, **J F Jardine**, *Simplicial homotopy theory*, Progr. Math. 174, Birkhäuser, Basel (1999) MR

[5] **D Joyce**, *On manifolds with corners*, from "Advances in geometric analysis", Adv. Lect. Math. 21, International, Somerville, MA (2012) 225–258 MR

[6] **O Lazarev**, **Z Sylvan**, **H L Tanaka**, *The $\infty$-category of stabilized Liouville sectors*, preprint (2021) arXiv 2110.11754

[7] **J Lurie**, *Higher topos theory*, Ann. of Math. Stud. 170, Princeton Univ. Press (2009) MR

[8] **J Lurie**, *Kerodon*: *an online resource for homotopy-coherent mathematics* (2023) Available at `https://kerodon.net`

[9] **J P May**, *Simplicial objects in algebraic topology*, Van Nostrand Math. Stud. 11, Van Nostrand, Princeton, NJ (1967) MR

[10] **B Mazur**, *Stable equivalence of differentiable manifolds*, Bull. Amer. Math. Soc. 67 (1961) 377–384 MR

[11] **B Mazur**, *The method of infinite repetition in pure topology, I*, Ann. of Math. 80 (1964) 201–226 MR

[12] **J Milnor**, *Microbundles, I*, Topology 3 (1964) 53–80 MR

[13] **J Milnor**, *Lectures on the h-cobordism theorem*, Princeton Univ. Press (1965) MR

[14] **F Waldhausen**, *An outline of how manifolds relate to algebraic K-theory*, from "Homotopy theory", Lond. Math. Soc. Lect. Note Ser. 117, Cambridge Univ. Press (1987) 239–247 MR

[15] **C T C Wall**, *Classification problems in differential topology, IV: Thickenings*, Topology 5 (1966) 73–94 MR

*Department of Mathematics, Texas State University*
*San Marcos, TX, United States*

`hiro@txstate.edu`

msp

# Representation-graded Bredon homology of elementary abelian 2-groups

MARKUS HAUSMANN

STEFAN SCHWEDE

We calculate the representation-graded Bredon homology rings of all elementary abelian 2-groups with coefficients in the constant mod-2 Mackey functor. We exhibit minimal presentations for these rings as quotients of the polynomial algebra on the pre-Euler and inverse Thom classes of all nontrivial characters, subject to an explicit finite list of relations arising from orientability properties. Two corollaries of our presentation are the calculation, originally due to Holler and Kriz, of the geometric fixed point rings, and a strengthening of a calculation of Balmer and Gallauer of the localized twisted cohomology ring.

## Introduction

In this paper we establish minimal presentations for the representation-graded Bredon homology rings of all elementary abelian 2-groups, with coefficients in the constant mod-2 Mackey functor $\underline{\mathbb{F}}_2$. More specifically, we determine the "effective cone" of the RO-graded Bredon homology ring, ie, the sector given by the reduced Bredon homology groups of linear representation spheres; we denote this multigraded ring by $H(A, \star)$. The following is our main result, to be proved as Theorem 2.5 below.

**Theorem** *Let $A$ be an elementary abelian 2-group. The $\mathbb{F}_2$-algebra $H(A, \star)$ is generated by the pre-Euler classes $a_\lambda$ and the inverse Thom classes $t_\lambda$ for all nontrivial $A$-characters $\lambda$. The ideal of relations among the classes $a_\lambda$ and $t_\lambda$ is generated by the polynomials*

$$\sum_{\lambda \in T} a_\lambda \cdot \left( \prod_{\mu \in T \setminus \{\lambda\}} t_\mu \right)$$

*for all minimally dependent sets $T$ of nontrivial $A$-characters.*

In [5, Theorem 3.5], we use the presentation of $H(A, \star)$ to establish a "global" universal property of mod-2 Bredon homology, ie, of the system of all representation-graded Bredon homology rings of all elementary abelian 2-groups, including the functoriality in group homomorphisms. In the language of [5], Bredon homology is an initial additively oriented $\mathrm{el}_2^{\mathrm{RO}}$-algebra.

The fact that $H(A, \star)$ is generated by the classes $a_\lambda$ and $t_\lambda$ was previously shown by Holler and Kriz [6]. Our main new contribution is determining the relations between these classes. The relations listed in the theorem are minimal, ie, none of them can be omitted. The origin of the relations is the fact that for a minimally dependent set $T$ of nontrivial $A$-characters, the $A$-representation $\bigoplus_{\lambda \in T} \lambda$ is orientable; see Proposition 1.7.

We illustrate our result for elementary abelian 2-groups of small rank. When $A = C = \{\pm 1\}$ is of order 2, the calculation is classical, originally due to Stong (unpublished), and reproved by several authors. In this case there is only one nontrivial character, and no relations; so $H(C, \star) = \mathbb{F}_2[a, t]$ is a polynomial algebra on the pre-Euler class and the inverse Thom class. The ring $H(C^2, \star)$ was calculated by Ellis-Bloor in [3, Theorem 4.14]. In this case there are three nontrivial characters $p_1$, $p_2$ and $\mu$, and all relations are generated by the single relation

$$a_1 t_2 t_\mu + t_1 a_2 t_\mu + t_1 t_2 a_\mu = 0.$$

So the minimal presentation of $H(C^2, \star)$ is

$$H(C^2, \star) = \mathbb{F}_2[a_1, a_2, a_\mu, t_1, t_2, t_\mu]/(a_1 t_2 t_\mu + t_1 a_2 t_\mu + t_1 t_2 a_\mu).$$

To the best of our knowledge, the presentation of $H(A, \star)$ is new when the rank of $A$ exceeds two. For $A = C^3$ we make our presentation of $H(C^3, \star)$ completely explicit in Example 2.6. In this case there are 14 polynomial generators, namely the classes $a_\lambda$ and $t_\lambda$ for each of the seven nontrivial characters $\lambda$, and 14 minimal relations. Of these relations, seven are cubic in the generators, and of the same general form as in the previous example, ie, $a_\alpha t_\beta t_\gamma + t_\alpha a_\beta t_\gamma + t_\alpha t_\beta a_\gamma = 0$ for all triples of distinct nontrivial characters that satisfy $\alpha \cdot \beta \cdot \gamma = 1$. And there are seven minimal relations that are homogeneous of degree 4 in the generators, of the form

$$a_\alpha t_\beta t_\gamma t_\delta + t_\alpha a_\beta t_\gamma t_\delta + t_\alpha t_\beta a_\gamma t_\delta + t_\alpha t_\beta t_\gamma a_\delta = 0$$

for quadruples of distinct nontrivial characters that satisfy $\alpha \cdot \beta \cdot \gamma \cdot \delta = 1$.

The number of minimal relations grows very quickly in the rank of the elementary abelian 2-group; see the table in Remark 2.7. However, a basic pattern continues as follows: when $A$ has rank $r$, a new family of relations appears that has no predecessor for smaller rank, given by homogeneous polynomials of degree $r + 1$ in the generators.

We use our presentation of $H(A, \star)$ to derive two interesting corollaries. Inverting the pre-Euler classes $a_\lambda$ for all nontrivial $A$-characters and restricting to integer gradings yields the $A$-geometric fixed point ring $\Phi_*^A(H\mathbb{F}_2)$ of mod-2 Bredon homology. This ring was previously calculated by Holler and Kriz in [6, Theorem 2], who also gave a formula for the Poincaré series of the multigraded ring $H(A, \star)$ in [6, Theorem 5], and showed that $H(A, \star)$ maps isomorphically onto the subring of $\Phi_*^A(H\mathbb{F}_2)[a_\lambda]$ generated by the classes $a_\lambda$ and $t_\lambda = x_\lambda \cdot a_\lambda$ for all nontrivial $A$-characters, with the notation as in Corollary 3.4. We explain in Corollary 3.4 how our presentation of $H(A, \star)$ yields the Holler–Kriz presentation of $\Phi_*^A(H\mathbb{F}_2)$ upon localization. A noticeable feature is that inverting the pre-Euler classes makes all polynomial relations of degree at least four redundant.

More generally, we consider a subgroup $B$ of $A$ and determine the "mixed" localization obtained by inverting the pre-Euler classes of all characters that restrict nontrivially to $B$, and the inverse Thom classes of all other characters. The resulting integer-graded ring $H(A|B)$ previously featured in the work of

Balmer and Gallauer [1] on the Balmer spectrum of the tt-category of permutation modules. We obtain an explicit presentation of the ring $H(A|B)$ by generators and relations in Theorem 3.2; here, too, all polynomial relations of degree at least four in the minimal presentation of $H(A, \star)$ become redundant in the localization. Our calculation improves [1, Theorem 17.13], for the prime 2, from a "presentation modulo nilpotence" to an actual presentation. The graded ring $H(A, \star)$ is a domain, see Theorem 2.2 and Remark 3.5, so its localization $H(A|B)$ is a domain, too.

# 1   Representation-graded Bredon homology

In this section we review some basic features of Bredon homology in a form adapted for our purposes. This section does not contain any new mathematics. What is now called *Bredon cohomology* was introduced by Bredon in [2] for finite groups and equivariant CW-complexes. The corresponding equivariant homology theory was introduced by Illman in [7]. Illman develops the theory for arbitrary topological groups, and he uses singular chains to define the equivariant homology and cohomology groups on arbitrary equivariant spaces. In many ways, Bredon homology and cohomology are the correct generalizations of singular (co)homology to the equivariant context, and of fundamental importance in equivariant topology.

**Construction 1.1**   (Bredon homology)   We recall one construction of Bredon homology with coefficients in a constant Mackey functor. We employ a definition that is naturally isomorphic to the original one of Bredon and Illman, namely as the equivariant homology theory represented by the Eilenberg–MacLane $G$-spectrum $H\underline{M}$ of the constant Mackey functor $\underline{M}$ associated to an abelian group $M$. In other words, we define the $m^{\text{th}}$ reduced $G$-equivariant Bredon homology group of a based $G$-space $X$ with $\underline{M}$-coefficients as

$$\widetilde{H}_m^G(X; \underline{M}) = \pi_m^G(H\underline{M} \wedge X).$$

The groups $\widetilde{H}_*^G(-; \underline{M})$ form an equivariant homology theory. In particular, they come with a suspension isomorphism, and a based $G$-map gives rise to a long exact sequence featuring the Bredon homology groups of source, target and the reduced mapping cone. We mostly consider $\mathbb{F}_2$-coefficients in this paper, and we will drop $\mathbb{F}_2$-coefficients from the notation.

The commutative $G$-ring spectrum structure of $H\mathbb{F}_2$ gives rise to associative, commutative and bilinear pairings

$$\cdot \colon \widetilde{H}_m^G(X) \times \widetilde{H}_n^G(Y) \to \widetilde{H}_{m+n}^G(X \wedge Y)$$

for all based $G$-spaces $X$ and $Y$.

In this paper a *representation* of a finite group is a finite-dimensional orthogonal representation. Our results are about the $\underline{\mathbb{F}}_2$-Bredon homology groups of representation spheres, ie, one-point compactifications $S^V$ of such representations $V$. The following proposition collects some well-known general facts about these; we give proofs for the readers' convenience. Part (iii) says that automorphisms of representation spheres are invisible to the eyes of Bredon homology with $\underline{\mathbb{F}}_2$-coefficients. This means that we can — and will — safely ignore the distinction between $G$-representations and their isomorphism classes.

**Proposition 1.2** *Let $V$ be a $d$-dimensional representation of a finite group $G$.*

(i) *The group $\widetilde{H}_k^G(S^V)$ is trivial for $k < 0$ and for $k > d$, and the restriction homomorphism*

$$\mathrm{res}_1^G \colon \widetilde{H}_d^G(S^V) \to \widetilde{H}_d(S^V)$$

*is an isomorphism. Hence the $\mathbb{F}_2$-vector space $\widetilde{H}_d^G(S^V)$ is 1-dimensional.*

(ii) *If $V$ is orientable, then the restriction homomorphism for constant integer coefficients*

$$\mathrm{res}_1^G \colon \widetilde{H}_d^G(S^V; \underline{\mathbb{Z}}) \to \widetilde{H}_d(S^V; \mathbb{Z})$$

*is an isomorphism. Hence the abelian group $\widetilde{H}_d^G(S^V; \underline{\mathbb{Z}})$ is free of rank 1.*

(iii) *For every based $G$-homotopy equivalence $\psi \colon S^V \to S^V$ and every $m \geq 0$, the map*

$$\widetilde{H}_m^G(\psi) \colon \widetilde{H}_m^G(S^V) \to \widetilde{H}_m^G(S^V)$$

*is the identity.*

**Proof** (i) The groups $\widetilde{H}_*^G(S^V)$ can be calculated from the reduced cellular chain complex of a $G$-CW-structure on $S^V$ by taking first $G$-fixed points and then homology. Since the reduced cellular chain complex is concentrated in dimensions 0 through $d$, the vanishing claims follow. Because the underlying chain complex calculates the reduced homology of the underlying space $S^V$, the kernel of the top differential $\delta_d \colon C_d^{\mathrm{cell}}(S^V; \mathbb{F}_2) \to C_{d-1}^{\mathrm{cell}}(S^V; \mathbb{F}_2)$ is 1-dimensional and necessarily with trivial $G$-action. So the kernel of $\delta_d^G \colon (C_d^{\mathrm{cell}}(S^V; \mathbb{F}_2))^G \to (C_{d-1}^{\mathrm{cell}}(S^V; \mathbb{F}_2))^G$ is also 1-dimensional, yielding the second claim.

(ii) The argument for constant integer coefficients in (ii) is essentially the same; the orientability assumption is equivalent to the condition that $G$ acts trivially on the group of $d$-cycles in the underlying chain complex.

(iii) Because $\psi \colon S^V \to S^V$ is a $G$-homotopy equivalence, its class in the $G$-equivariant 0-stem is a unit. Because Bredon homology is represented by the commutative $G$-ring spectrum $H\underline{\mathbb{F}}_2$, the map $\widetilde{H}_*^G(\psi)$ equals multiplication by the Hurewicz image of $\langle \psi \rangle$ in $\pi_0^G(H\underline{\mathbb{F}}_2)$. The ring $\pi_0^G(H\underline{\mathbb{F}}_2)$ is isomorphic to $\mathbb{F}_2$, so 1 is its only unit. □

**Construction 1.3** (the representation-graded ring $H(G, \star)$) We introduce notation to deal with the representation grading. For a finite group $G$, we let $J_G$ be the abelian monoid, under direct sum, of isomorphism classes of $G$-representations with trivial $G$-fixed points. So $J_G$ is freely generated by the isomorphism classes of nontrivial irreducible $G$-representations.

We choose a representative for each element of $J_G$. For $\rho \in J_G$ and $m \geq 0$, we define

$$H_m(G, \rho) = \widetilde{H}_m^G(S^V),$$

where $V$ is the chosen representative of $\rho$. Proposition 1.2(iii) guarantees that this definition is independent of the representative up to preferred isomorphism, induced by any isomorphism of representations.

The pairing of Bredon homology induces an associative, commutative and bilinear pairing

$$\cdot : H_m(G, \rho) \times H_n(G, \kappa) \to H_{m+n}(G, \rho + \kappa)$$

for $\rho, \kappa \in J_G$ and $m, n \geq 0$. This map is defined as the composite

$$\widetilde{H}_m^G(S^V) \times \widetilde{H}_n^G(S^W) \xrightarrow{\;\cdot\;} \widetilde{H}_{m+n}^G(S^V \wedge S^W) \cong \widetilde{H}_{m+n}^G(S^U),$$

where $V$, $W$ and $U$ are the chosen representatives of $\rho$, $\kappa$ and $\rho + \kappa$, respectively. The last isomorphism is induced by a choice of $G$-equivariant isomorphism $V \oplus W \cong U$; it is independent of the choice by Proposition 1.2(iii). We emphasize that this multiplication is strictly commutative, ie, for $x \in H_m(G, \rho)$ and $y \in H_n(G, \kappa)$, the classes $x \cdot y$ and $y \cdot x$ are *equal* in the group $H_{m+n}(G, \rho + \kappa) = H_{n+m}(G, \kappa + \rho)$. This, one more time, uses that automorphisms of representation spheres are invisible in $\widetilde{H}_*^G$. The multiplication maps thus make the collection of groups $H_m(G, \rho)$ into a commutative $(\mathbb{N} \times J_G)$-graded $\mathbb{F}_2$-algebra. We denote this object by $H(G, \star)$ and refer to it as the *representation-graded Bredon homology ring* of the group $G$. We will routinely abuse notation by identifying a $G$-representation $V$ with trivial fixed points with its class in $J_G$; thus we shall write $H_m(G, V)$ for $H_m(G, [V])$.

**Remark 1.4** Bredon homology with coefficients in a Mackey functor is represented by a genuine $G$-spectrum, and hence can be extended to a homology theory for $G$-spaces that is $\mathrm{RO}(G)$-graded. Our results are only about the "effective cone" of the $\mathrm{RO}(G)$-graded coefficient ring, ie, the sector given by Bredon homology of representation spheres. The effective cone has a much nicer algebraic structure than the rest of the $\mathrm{RO}(G)$-graded Bredon homology, which tends to contain many nilpotent classes and trivial products. The effective cone contains the pre-Euler and inverse Thom classes, so it determines the geometric fixed points (obtained by inverting all pre-Euler classes), and various other localizations; see Construction 3.1 below.

We recall two kinds of classes that exist for every $G$-representation, the pre-Euler class and the inverse Thom class. Our pre-Euler class is also called "Euler class" by other authors.

**Construction 1.5** (pre-Euler and inverse Thom classes) We let $V$ be a $G$-representation. The *pre-Euler class*

$$a_V \in H_0(G, V)$$

is the image of the multiplicative unit $1 \in H_0(G, 0)$ under the homomorphism induced by the based $G$-map $S^0 \to S^V$ that sends the point $0$ to the $G$-fixed point $0$ in $S^V$. The pre-Euler class can be zero, for example if $V$ has nonzero $G$-fixed points.

For $d = \dim(V)$, the *inverse Thom class* is the unique nonzero element

$$t_V \in H_d(G, V);$$

compare with Proposition 1.2(i). If $W$ is another $G$-representation, then

$$a_V \cdot a_W = a_{V \oplus W} \quad \text{and} \quad t_V \cdot t_W = t_{V \oplus W}$$

in $H_*(G, V \oplus W)$.

In this paper, a *character* of a finite group $G$ is a group homomorphism $\lambda \colon G \to C = \{\pm 1\}$. We shall routinely confuse a character $\lambda$ with the 1-dimensional $G$-representation on $\mathbb{R}$ in which $g \in G$ acts by multiplication by $\lambda(g)$. Elementary abelian 2-groups are characterized among finite groups by the property that all irreducible real representations are 1-dimensional, and hence given by characters.

We will make use of the Bockstein homomorphism

$$\beta \colon \widetilde{H}_m^G(X) \to \widetilde{H}_{m-1}^G(X)$$

associated to the short exact sequence of constant Mackey functors $\underline{\mathbb{F}}_2 \to \underline{\mathbb{Z}/4} \to \underline{\mathbb{F}}_2$.

**Example 1.6** Let $\lambda \colon G \to C$ be a nontrivial character with kernel $K$. The minimal $G$-CW-structure of $S^\lambda$ with two fixed 0-cells and one 1-cell with isotropy $K$ shows that for every abelian group $M$, the group $\widetilde{H}_n^G(S^\lambda; \underline{M})$ is trivial for $n \neq 0, 1$. And it yields an exact sequence of Bredon homology groups:

$$0 \to \widetilde{H}_1^G(S^\lambda; \underline{M}) \to H_0^G(G/K; \underline{M}) \to H_0^G(G/G; \underline{M}) \xrightarrow{\cdot a_\lambda} \widetilde{H}_0^G(S^\lambda; \underline{M}) \to 0.$$

The middle two groups are coefficient groups of the Mackey functor, and thus equal to $M$. The middle homomorphism is the transfer from $K$ to $G$ in the Mackey functor $\underline{M}$, ie, multiplication by the index $[G : K] = 2$. So the groups $\widetilde{H}_n^G(S^\lambda; \underline{M})$ in dimension 1 and 0 are isomorphic to the kernel and cokernel, respectively, of multiplication by 2 on $M$. For $M = \mathbb{F}_2$ we conclude that $H_1(G, \lambda)$ and $H_0(G, \lambda)$ are 1-dimensional, generated by the inverse Thom class $t_\lambda$ and the pre-Euler class $a_\lambda$, respectively.

We claim that the Bockstein homomorphism takes $t_\lambda$ to $a_\lambda$. Indeed, the short exact sequence of Mackey functors $\underline{\mathbb{F}}_2 \to \underline{\mathbb{Z}/4} \to \underline{\mathbb{F}}_2$ yields an exact sequence of Bredon homology groups:

$$\widetilde{H}_1^G(S^\lambda; \underline{\mathbb{F}}_2) \xrightarrow{\beta} \widetilde{H}_0^G(S^\lambda; \underline{\mathbb{F}}_2) \to \widetilde{H}_0^G(S^\lambda; \underline{\mathbb{Z}/4}) \to \widetilde{H}_0^G(S^\lambda; \underline{\mathbb{F}}_2) \to 0.$$

Since the four nontrivial groups in this sequence are all cyclic of order 2, exactness implies that the Bockstein homomorphism is an isomorphism. Since source and target are spanned by $t_\lambda$ and $a_\lambda$, respectively, the Bockstein satisfies $\beta(t_\lambda) = a_\lambda$.

**Proposition 1.7** *Let $G$ be a finite group.*

(i) *If $V$ is an orientable $G$-representation, then $\beta(t_V) = 0$.*

(ii) *Let $T$ be a set of $G$-characters whose product is 1. Then*

$$\sum_{\lambda \in T} a_\lambda \cdot \left( \prod_{\mu \in T \setminus \{\lambda\}} t_\mu \right) = 0.$$

**Proof** (i) If $V$ is orientable, then by Proposition 1.2(ii), the inverse Thom class $t_V$ lifts to a class in the Bredon homology with constant integral coefficients. So the image of $t_V$ under the Bockstein homomorphism is trivial.

(ii) We let $V = \bigoplus_{\lambda \in T} \lambda$ be the sum of the characters in the set $T$. The determinant of $V$ is the product of the characters in $T$, which is trivial by hypothesis. So the $G$-representation $V$ is orientable, and hence $\beta(t_V) = 0$ by part (i). The Bockstein homomorphism is a derivation, in the sense that $\beta(x \cdot y) = \beta(x) \cdot y + x \cdot \beta(y)$ for all classes $x \in \widetilde{H}_m^G(X)$ and $y \in \widetilde{H}_n^G(Y)$. Applying the derivation property repeatedly and using that $\beta(t_\lambda) = a_\lambda$ shows that

$$0 = \beta(t_V) = \beta\Big( \prod_{\lambda \in T} t_\lambda \Big) = \sum_{\lambda \in T} a_\lambda \cdot \Big( \prod_{\mu \in T \setminus \{\lambda\}} t_\mu \Big). \qquad \square$$

## 2 Bredon homology for elementary abelian 2-groups

In this section we specialize from general finite groups to elementary abelian 2-groups, and we prove our main result, Theorem 2.5. There we exhibit a presentation of $H(A, \star)$ as the quotient of a polynomial $\mathbb{F}_2$-algebra on the classes $a_\lambda$ and $t_\lambda$ for all nontrivial $A$-characters $\lambda$, by an explicit minimal set of homogeneous polynomial relations. Along the way, we give an elementary and self-contained proof that the ring $H(A, \star)$ is a domain; see Theorem 2.2.

Many of our arguments involve bootstrapping information about Bredon homology of a subgroup to the ambient group. In those arguments, we need to restrict representations to subgroups, which typically creates new fixed points. If $V$ is a $G$-representation with $V^G = 0$, and $K$ a subgroup of $G$, we set $k = \dim(V^K)$, and we let $V_K = V - V^K$ be the orthogonal complement of the $K$-fixed points in $V$. The *restriction homomorphism*

$$\mathrm{res}_K^G \colon H_m(G, V) \to H_{m-k}(K, V_K)$$

is the composite

$$\widetilde{H}_m^G(S^V) \xrightarrow{\mathrm{res}_K^G} \widetilde{H}_m^K(S^V) \cong \widetilde{H}_m^K(S^{V_K} \wedge S^k) \xrightarrow[\cong]{(-\wedge S^k)^{-1}} \widetilde{H}_{m-k}^K(S^{V_K}).$$

The first isomorphism is induced by a choice of $K$-equivariant isomorphism $V \cong V_K \oplus \mathbb{R}^k$; it is independent of this choice by Proposition 1.2(iii). The second isomorphism is the inverse of the suspension isomorphism. The restriction homomorphism is multiplicative, and its effect on inverse Thom and pre-Euler classes is given by

$$\mathrm{res}_K^G(t_V) = t_{V_K} \qquad \text{and} \qquad \mathrm{res}_K^G(a_V) = \begin{cases} a_{V_K} & \text{if } V^K = 0, \\ 0 & \text{if } V^K \neq 0. \end{cases}$$

**Proposition 2.1** *Let $A$ be an elementary abelian 2-group and let $W$ be an $A$-representation with trivial fixed points.*

(i)  *For every subgroup $B$ of $A$, the restriction homomorphism*

$$\mathrm{res}_B^A \colon H_m(A, W) \to H_{m-k}(B, W_B)$$

*is surjective, where $k = \dim(W^B)$.*

(ii)  *Suppose that $W = V \oplus \lambda$ for an $A$-representation $V$ and a nontrivial $A$-character $\lambda$ with kernel $K$. Then for $k = \dim(V^K)$, the following sequence is exact:*

$$0 \to H_m(A, V) \xrightarrow{\cdot a_\lambda} H_m(A, V \oplus \lambda) \xrightarrow{\mathrm{res}_K^A} H_{m-k-1}(K, V_K) \to 0.$$

(iii)  *The $\mathbb{F}_2$-vector space $H_m(A, W)$ is spanned by the classes $a_U \cdot t_V$ for all $A$-representations $U$ and $V$ such that $U \oplus V = W$ and $m = \dim(V)$.*

**Proof**  We prove all three statements together by induction over the rank of $A$. The induction starts when $A$ is the trivial group, in which case there is nothing to show. Now we let $A$ be a nontrivial elementary abelian 2-group, and we assume that parts (i)–(iii) hold for all proper subgroups of $A$.

We start by proving (i), where we may assume that $B$ is a proper subgroup of $A$. By part (iii) for $B$, it suffices to show that all classes of the form $a_U \cdot t_V$ are in the image of the restriction homomorphism, whenever $U$ and $V$ are $B$-representations such that $U \oplus V = W_B$ and $m - k = \dim(V)$. Because $W$ is a sum of 1-dimensional $A$-representations, we may choose an $A$-equivariant decomposition $W = \overline{U} \oplus \overline{V} \oplus T$ such that $\mathrm{res}_B^A(\overline{U}) \cong U$, $\mathrm{res}_B^A(\overline{V}) \cong V$, and $B$ acts trivially on $T$. Then $\mathrm{res}_B^A(a_{\overline{U}} \cdot t_{\overline{V}} \cdot t_T) = a_U \cdot t_V$, and we have shown part (i) for $A$.

Now we prove (ii). Smashing the cofiber sequence of based $A$-spaces

$$A/K_+ \to S^0 \to S^\lambda \to A/K_+ \wedge S^1$$

with $S^V$ and applying $A$-equivariant Bredon homology yields a long exact sequence:

$$\cdots \to H_m(A, V) \xrightarrow{\cdot a_\lambda} H_m(A, V \oplus \lambda) \xrightarrow{\partial} \widetilde{H}^A_{m-1}(S^V \wedge A/K_+) \to \cdots.$$

The Wirthmüller and suspension isomorphisms identify the group $\widetilde{H}^A_{m-1}(S^V \wedge A/K_+)$ with the group $\widetilde{H}^K_{m-k-1}(S^{V_K}) = H_{m-k-1}(K, V_K)$. Under this identification, the boundary map $\partial$ becomes the restriction homomorphism $\mathrm{res}_K^A \colon H_m(A, V \oplus \lambda) \to H_{m-k-1}(K, V_K)$, which is surjective by (i). So the long exact sequence decomposes into short exact sequences, showing (ii).

We prove (iii) by induction on the dimension of $W$. For $W = 0$, the groups $H_*(A, 0)$ consist of a single copy of $\mathbb{F}_2$ in dimension 0, spanned by the multiplicative unit $1 = t_0$. If $W$ is nonzero, we write $W = V \oplus \lambda$ for an $A$-representation $V$ and a nontrivial $A$-character $\lambda$, with kernel $K$. By part (i), the restriction map $\mathrm{res}_K^A \colon H_{m-1}(A, V) \to H_{m-k-1}(K, V_K)$ is surjective. So for every class $x \in H_m(A, W)$, there is a class $z \in H_{m-1}(A, V)$ such that $\mathrm{res}_K^A(z) = \mathrm{res}_K^A(x)$. Then

$$\mathrm{res}_K^A(x + z \cdot t_\lambda) = \mathrm{res}_K^A(x) + \mathrm{res}_K^A(z) = 0.$$

Part (ii) provides a class $y \in H_m(A, V)$ such that $y \cdot a_\lambda = x + z \cdot t_\lambda$. Because the dimension of $V$ is smaller than that of $W$, the classes $y$ and $z$ are sums of products of $a$-classes and $t$-classes. Hence the same is true for $x = y \cdot a_\lambda + z \cdot t_\lambda$, and we have shown (iii). □

Our next result shows that the rings $H(A, \star)$ have no zero-divisors. As we explain in Remark 3.5 below, this can also be proven by combining results of [6; 8], so we make no claim to originality. As a service to the reader, we record this key structural result explicitly and give an independent, self-contained and elementary proof.

**Theorem 2.2** *For every elementary abelian 2-group $A$, the representation-graded Bredon homology ring $H(A, \star)$ is a domain.*

**Proof** We call an element of $H(A, \star)$ *regular* if multiplication by it is injective. We will show that all nonzero homogeneous elements of $H(A, \star)$ are regular. We argue by induction over the rank of $A$. For $A = \{1\}$, the ring is the field $\mathbb{F}_2$, hence a domain.

Now we suppose that $A$ is nontrivial. Proposition 2.1(ii) shows that the pre-Euler classes of all nontrivial $A$-characters are regular. We show next that all the inverse Thom classes $t_\mu$ are regular. So we let $y \in H_n(A, W)$ be a homogeneous element such that $t_\mu \cdot y = 0$, for some $A$-representation $W$ with trivial fixed points. We argue by induction on the dimension of $W$. If $W = 0$ there is nothing to show because the integer-graded part of Bredon homology is spanned by the multiplicative unit, and $t_\mu \neq 0$. For $W \neq 0$ we write $W = U \oplus \lambda$ for some nontrivial $A$-character $\lambda$, with kernel $K$. Then

$$\mathrm{res}_K^A(t_\mu) \cdot \mathrm{res}_K^A(y) = \mathrm{res}_K^A(t_\mu \cdot y) = 0.$$

Because $\mathrm{res}_K^A(t_\mu)$ is either 1 (if $\lambda = \mu$), or the inverse Thom class of the restricted character $\mu|_K$ (if $\lambda \neq \mu$), and because $K$ has smaller rank than $A$, the class $\mathrm{res}_K^A(t_\mu)$ is nonzero and regular, so $\mathrm{res}_K^A(y) = 0$. Proposition 2.1(ii) provides a class $u \in H_n(A, U)$ such that $y = u \cdot a_\lambda$. Then $t_\mu \cdot u \cdot a_\lambda = t_\mu \cdot y = 0$, so $t_\mu \cdot u = 0$ because $a_\lambda$ is regular. Because $U$ has smaller dimension than $W$, we deduce that $u = 0$. Hence also $y = 0$, and this concludes the special case.

Now we show that a general nonzero homogeneous element $x \in H_m(A, V)$ is regular, where $V$ is an $A$-representation with trivial fixed points. The assumption that $x$ is nonzero implies that $m \leq \dim(V)$. We argue by induction on $\dim(V) - m$. If $m = \dim(V)$, then $x = t_V$, and so $x$ is a product of inverse Thom classes of $A$-characters. Since all the factors are regular by the special case, so is $t_V$.

Now we suppose that $m < \dim(V)$. We distinguish two cases. In the first case we suppose that there is an index 2 subgroup $K$ of $A$ such that $\mathrm{res}_K^A(x) = 0$. We let $\lambda$ be the $A$-character whose kernel is $K$. Then also $\mathrm{res}_K^A(x \cdot t_\lambda) = 0$. Proposition 2.1(ii) provides a class $u \in H_{m+1}(A, V)$ such that $u \cdot a_\lambda = x \cdot t_\lambda$. Since $x \neq 0$ and $t_\lambda$ is regular, the class $u$ is nonzero. Because the integer degree of $u$ is larger than that of $x$, the induction hypothesis shows that $u$ is regular. Since $u$, $a_\lambda$ and $t_\lambda$ are regular, so is $x$.

In the second case we suppose that for every index 2 subgroup $K$ of $A$, the restriction $\mathrm{res}^A_K(x)$ is nonzero. We consider a homogeneous element $y \in H_n(A, W)$ such that $x \cdot y = 0$, where $W$ is another $A$-representation with trivial fixed points. We perform another induction on the dimension of $W$. If $W = 0$, then $y$ lies in the integer-graded subring, so $y = 0$ or $y = 1$, and we are done. For $W \neq 0$ we write $W = U \oplus \lambda$ for some nontrivial $A$-character $\lambda$, with kernel $K$. Then

$$\mathrm{res}^A_K(x) \cdot \mathrm{res}^A_K(y) = \mathrm{res}^A_K(x \cdot y) = 0.$$

Because $\mathrm{res}^A_K(x) \neq 0$ and $H(K, *)$ is a domain by induction, this implies $\mathrm{res}^A_K(y) = 0$. Proposition 2.1(ii) provides a class $u \in H_n(A, U)$ such that $y = u \cdot a_\lambda$. Thus $x \cdot u \cdot a_\lambda = x \cdot y = 0$. So $x \cdot u = 0$ because $a_\lambda$ is regular. Because the dimension of $U$ is smaller than the dimension of $W$, we conclude that $u = 0$, and hence also $y = 0$. □

Now we move on to our main result, the minimal presentation of the representation-graded Bredon homology ring $H(A, \star)$. Proposition 2.1(iii) shows that $H(A, \star)$ is generated as an $\mathbb{F}_2$-algebra by the classes $a_\lambda$ and $t_\lambda$ for all nontrivial $A$-characters $\lambda$. If $A = C$, there is only one nontrivial character, and then $H(C, \star)$ is well known to be a polynomial algebra on the classes $a$ and $t$. If $A$ is elementary abelian of rank at least 2, however, there are nontrivial polynomial relations between the pre-Euler classes and the inverse Thom classes.

In the following, it will be convenient to use the notation

$$A^\circ = \mathrm{Hom}(A, C) \setminus \{1\}$$

for the set of nontrivial characters of an elementary abelian 2-group.

**Definition 2.3** (minimally dependent sets of characters) Let $A$ be an elementary abelian 2-group. A nonempty subset of $A^\circ$ is *dependent* if it is linearly dependent as a subset of the vector space $\mathrm{Hom}(A, C)$. The set is *minimally dependent* if it is dependent, but no proper subset is dependent.

**Remark 2.4** If $T$ is a minimally dependent subset of $A^\circ$, then the product of all its elements must be the trivial character. Indeed, the linear dependence of $T$ means that some subset of it has product the trivial character; if this were the case for some proper subset of $T$, then that subset would be dependent, contradicting minimality.

Sets of nontrivial $A$-characters that have one or two elements have a nontrivial product. So every dependent set of $A$-characters has a least three elements. If $A$ has rank $r$, then every set of $r + 1$ nontrivial $A$-characters is dependent, and hence has a nonempty subset whose product is the trivial character. So a minimally dependent set of $A$-characters has at most $r + 1$ elements.

The group $C^2$ has precisely one minimally dependent set of nontrivial characters, the set of all three nontrivial characters. In Example 2.6 we enumerate all 14 minimally dependent sets of nontrivial characters of the group $C^3$. In Remark 2.7 we determine the number of minimally dependent subsets of $A^\circ$ as a function of the rank of $A$.

The next theorem is our main result about representation-graded Bredon homology, providing an explicit set of homogeneous polynomial relations between the pre-Euler and inverse Thom classes, parameterized by minimally dependent sets of characters. Moreover, these relations form a minimal generating set for the ideal of all relations. In [5, Theorem 3.5], we use these presentations of $H(A, \star)$ to establish a "global" universal property of the collection of representation-graded Bredon homology rings: we show that mod-2 Bredon homology is an initial additively oriented $\mathrm{el}_2^{\mathrm{RO}}$-algebra. Holler and Kriz determined the Poincaré series of the multigraded ring $H(A, \star)$ in [6, Theorem 5]. We have not investigated how to derive their formula for the Poincaré series from our presentation.

**Theorem 2.5** *Let $A$ be an elementary abelian 2-group.*

(i) *The representation-graded Bredon homology ring $H(A, \star)$ is generated as an $\mathbb{F}_2$-algebra by the classes $a_\lambda$ and $t_\lambda$ for all nontrivial $A$-characters $\lambda$.*

(ii) *The kernel of the surjective homomorphism of $\mathbb{F}_2$-algebras*

$$\epsilon_A \colon \mathbb{F}_2[a_\mu, t_\mu : \mu \in A^\circ] \to H(A, \star)$$

*is the ideal generated by the polynomials*

$$r(T) = \sum_{\lambda \in T} a_\lambda \cdot \left( \prod_{\mu \in T \setminus \{\lambda\}} t_\mu \right)$$

*for all minimally dependent subsets $T$ of $A^\circ$.*

(iii) *Every set of homogeneous elements that generates the kernel of $\epsilon_A$ contains the polynomials $r(T)$ for all minimally dependent subsets $T$ of $A^\circ$.*

**Proof** (i) This was shown in Proposition 2.1(iii), and is repeated here for easier reference.

(ii) For the course of the proof we write $I(A)$ for the ideal of the polynomial ring $\mathbb{F}[a_\mu, t_\mu : \mu \in A^\circ]$ generated by the polynomials $r(T)$ for all minimally dependent subsets $T$ of $A^\circ$. Since minimally dependent sets of characters multiply to 1, Proposition 1.7(ii) shows that $I(A) \subseteq \ker(\epsilon_A)$; so it remains to show the reverse inclusion.

We argue by induction on the rank of $A$. The induction starts with the trivial group, where there is nothing to show. Now we let $A$ be a nontrivial elementary abelian 2-group, and we assume part (ii) for all proper subgroups of $A$. In the inductive step, we shall make use of the polynomials

$$r(T) = \sum_{\lambda \in T} a_\lambda \cdot \left( \prod_{\mu \in T \setminus \{\lambda\}} t_\mu \right)$$

for arbitrary subsets $T$ of $A^\circ$, not necessarily minimally dependent; we alert the reader that if the elements of $T$ do not multiply to the trivial character, then the polynomial $r(T)$ will *not* map to 0 under $\epsilon_A$.

We let $\lambda$ be a nontrivial $A$-character, with kernel $K$. We let $\mathcal{J}$ denote the homogeneous ideal in the ring $\mathbb{F}_2[a_\mu, t_\mu : \mu \in A^\circ \setminus \{\lambda\}]$ consisting of those elements $y$ such that $\mathrm{res}_K^A(\epsilon_A(y)) = 0$. We emphasize that elements of $\mathcal{J}$ are polynomials that do not involve the variables $a_\lambda$ nor $t_\lambda$. We shall prove two properties of this ideal:

(a)  The ideal $\mathcal{J}$ is generated by the polynomials $r(T)$ for all subsets $T$ of $A^{\circ} \setminus \{\lambda\}$ such that either $T$ is minimally dependent, or $T \cup \{\lambda\}$ is minimally dependent.

(b)  In the graded ring $\mathbb{F}_2[a_{\mu}, t_{\mu} : \mu \in A^{\circ}]$, the relation $t_{\lambda} \cdot \mathcal{J} \subset \langle I(A), a_{\lambda} \rangle$ holds.

**Proof of (a)**  We write $J_{A \setminus \lambda}$ for the submonoid of $J_A$ consisting of the classes of $A$-representations $V$ that do not involve the character $\lambda$. Equivalently, $J_{A \setminus \lambda}$ contains the $A$-representations with trivial $A$-fixed points such that $V^K = 0$. The restriction homomorphism $\mathrm{res}_K^A : \mathbb{N} \times J_{A \setminus \lambda} \to \mathbb{N} \times J_K$ lets us inflate $(\mathbb{N} \times J_K)$-graded rings $R$ to $(\mathbb{N} \times J_{A \setminus \lambda})$-graded rings $(\mathrm{res}_K^A)^*(R)$ by setting

$$(\mathrm{res}_K^A)^*(R)(k, V) = R(k, V|_K).$$

We alert the reader that a grading-inflated polynomial algebra is no longer a polynomial algebra. With this grading convention, a morphism of $(\mathbb{N} \times J_{A \setminus \lambda})$-graded $\mathbb{F}_2$-algebras

$$\rho_K^A : \mathbb{F}_2[a_{\mu}, t_{\mu} : \mu \in A^{\circ} \setminus \{\lambda\}] \to (\mathrm{res}_K^A)^*(\mathbb{F}_2[a_{\nu}, t_{\nu} : \nu \in K^{\circ}])$$

is given by sending $a_{\mu}$ to $a_{\mu|K}$ and $t_{\mu}$ to $t_{\mu|K}$. Moreover, the ideal $\mathcal{J}$ is precisely the kernel of the composite homomorphism

$$\mathbb{F}_2[a_{\mu}, t_{\mu} : \mu \in A^{\circ} \setminus \{\lambda\}] \xrightarrow{\rho_K^A} (\mathrm{res}_K^A)^*(\mathbb{F}_2[a_{\nu}, t_{\nu} : \nu \in K^{\circ}])$$

$$\xrightarrow{(\mathrm{res}_K^A)^*(\epsilon_K)} (\mathrm{res}_K^A)^*(H(K, \star)).$$

The kernel of $\rho_K^A$ is the ideal generated by the homogeneous polynomials

$$a_{\mu} t_{\mu\lambda} + t_{\mu} a_{\mu\lambda} = r(\{\mu, \mu\lambda\})$$

for all $\mu \in A^{\circ} \setminus \{\lambda\}$. The set $\{\mu, \mu\lambda\}$ is independent and $\{\mu, \mu\lambda\} \cup \{\lambda\}$ is minimally dependent. So the polynomials $r(\{\mu, \mu\lambda\})$ are among those of the second kind specified in (a).

By the inductive hypothesis for the subgroup $K$, the kernel of $\epsilon_K : \mathbb{F}_2[a_{\nu}, t_{\nu} : \nu \in K^{\circ}] \to H(K, \star)$ is generated by the polynomials $r(S)$ for all minimally dependent subsets $S$ of $K^{\circ}$. So the kernel of the homomorphism

$$(\mathrm{res}_K^A)^*(\epsilon_K) : (\mathrm{res}_K^A)^*(\mathbb{F}_2[a_{\nu}, t_{\nu} : \nu \in K^{\circ}]) \to (\mathrm{res}_K^A)^*(H(K, \star))$$

is the ideal generated by the same polynomials $r(S)$, but each occurring multiple times in different degrees, namely for all subsets $T$ of $A^{\circ} \setminus \{\lambda\}$ such that

$$\bigoplus_{\mu \in T} \mu|_K = \bigoplus_{\nu \in S} \nu.$$

This condition means that for each character $\nu \in S$, the set $T$ contains exactly one of the two extensions of $\nu$ to an $A$-character. Because $S$ is a minimally dependent subset of $K^{\circ}$, the $A$-character

$$\prod_{\mu \in T} \mu \in \mathrm{Hom}(A, C)$$

then restricts to the trivial character on the subgroup $K$. Hence this product is either 1 or $\lambda$. In the first case, $T$ is a minimally dependent subset of $A° \setminus \{\lambda\}$. In the second case, $T$ is an independent subset of $A° \setminus \{\lambda\}$ such that $T \cup \{\lambda\}$ is minimally dependent.

We have now shown that some of the polynomials $r(T)$ with $T$ as specified in (a) generate the kernel of $\rho_K^A$, and the images of the others under $\rho_K^A$ generate the kernel of $(\mathrm{res}_K^A)^*(\epsilon_K)$. So all those polynomials together generate the kernel of the composite, and hence the ideal $\mathcal{J}$. This completes the proof of (a).

**Proof of (b)** It suffices to show that the two types of generating polynomial for $\mathcal{J}$ specified in (a) have the desired property. If $T$ is a minimally dependent subset of $A° \setminus \{\lambda\}$, then the polynomial $r(T)$ belongs to the ideal $I(A)$. Hence also $t_\lambda \cdot r(T) \in I(A)$. If $T$ is a subset of $A° \setminus \{\lambda\}$ such that $T \cup \{\lambda\}$ is minimally dependent, then the relation

$$t_\lambda \cdot r(T) = r(T \cup \{\lambda\}) + a_\lambda \cdot \prod_{\mu \in T} t_\mu$$

holds in the ring $\mathbb{F}_2[a_\mu, t_\mu : \mu \in A°]$, and witnesses that the left hand side lies in the ideal $\langle I(A), a_\lambda \rangle$. This completes the proof of (b).

Now we complete the inductive step, showing that $\ker(\epsilon_A) \subseteq I(A)$. We prove this for all homogeneous pieces $H_m(A, W)$, where $W$ is an $A$-representation with $W^A = 0$, by induction on the dimension of $W$. For $W = 0$ we are considering integer degrees, where source and target of $\epsilon_A$ both consist only of a copy of $\mathbb{F}_2$ in degree 0.

Now we suppose that $W \neq 0$. We let $f$ be a homogeneous polynomial in $\mathbb{F}_2[a_\mu, t_\mu : \mu \in A°]$ of degree $(k, W)$ with $\epsilon_A(f) = 0$. We choose an $A$-character $\lambda$ and an $A$-representation $V$ with $W = V \oplus \lambda$. We let $K$ denote the kernel of $\lambda$. We write $f = t_\lambda \cdot y + a_\lambda \cdot z$ for some homogeneous polynomials $y, z$ of degrees $(k-1, V)$ and $(k, V)$, respectively. Then

$$\mathrm{res}_K^A(\epsilon_A(y)) = \mathrm{res}_K^A(t_\lambda \epsilon_A(y) + a_\lambda \epsilon_A(z)) = \mathrm{res}_K^A(\epsilon_A(f)) = 0.$$

**Case 1** The $A$-character $\lambda$ occurs in $W$ with multiplicity at least 2. Then there is an $A$-representation $U$ such that $V = U \oplus \lambda$. Because $\mathrm{res}_K^A(\epsilon_A(y)) = 0$, there is a class $u \in H_{k-1}(A, U)$ such that $\epsilon_A(y) = a_\lambda \cdot u$. Then

$$a_\lambda \cdot (t_\lambda \cdot u + \epsilon_A(z)) = t_\lambda \cdot \epsilon_A(y) + a_\lambda \cdot \epsilon_A(z) = \epsilon_A(f) = 0.$$

Because multiplication by $a_\lambda$ is injective, we deduce that $t_\lambda \cdot u = \epsilon_A(z)$. We choose a homogeneous polynomial $g$ in $\mathbb{F}_2[a_\mu, t_\mu : \mu \in A°]$ of degree $(k-1, U)$ such that $\epsilon_A(g) = u$. Then $y + a_\lambda g$ and $z + t_\lambda g$ lie in the kernel of $\epsilon_A$. Since $U \oplus \lambda = V$ has smaller dimension than $W$, the classes $y + a_\lambda g$ and $z + t_\lambda g$ are in the ideal $I(A)$, by induction. So

$$f = t_\lambda \cdot y + a_\lambda \cdot z = t_\lambda \cdot (y + a_\lambda g) + a_\lambda \cdot (z + t_\lambda g)$$

also lies in the ideal $I(A)$.

**Case 2** The $A$-character $\lambda$ occurs in $W$ with multiplicity 1. Then $\lambda$ does not occur in $V$, and so the polynomial $y$ does not involve the variable $a_\lambda$ nor $t_\lambda$. Because $\mathrm{res}_K^A(\epsilon_A(y)) = 0$, the polynomial $y$ belongs to the ideal $\mathcal{J}$. By property (b) of the ideal $\mathcal{J}$, the class $t_\lambda \cdot y$ then belongs to the ideal generated by $I(A)$ and $a_\lambda$. Hence there is a homogeneous polynomial $h$ of degree $(k, V)$ such that $t_\lambda \cdot y$ is congruent to $a_\lambda \cdot h$ modulo the ideal $I(A)$. So the polynomial $f$ in the kernel of $\epsilon_A$ satisfies

$$f = t_\lambda \cdot y + a_\lambda \cdot z \equiv a_\lambda \cdot (h + z) \quad \text{modulo } I(A).$$

Because $I(A)$ is contained in the kernel of $\epsilon_A$, we deduce the relation

$$a_\lambda \cdot \epsilon_A(h + z) = \epsilon_A(a_\lambda \cdot (h + z)) = \epsilon_A(f) = 0.$$

Because multiplication by $a_\lambda$ is injective, we conclude that $\epsilon_A(h + z) = 0$. Since $V$ has smaller dimension than $W = V \oplus \lambda$, the class $h + z$ lies in the ideal $I(A)$ by induction. So also $a_\lambda(h + z)$, and hence the class $f$, lies in the ideal $I(A)$. This finishes the inductive step, and hence the proof of part (ii).

(iii)  Source and target of the homomorphism $\epsilon_A$ are graded by the abelian monoid of isomorphism classes of $A$-representations. This abelian monoid is free on the classes of the $A$-characters and thus admits a compatible partial order by declaring $V \leq W$ if $V$ is isomorphic to a direct summand of $W$. Hence all products of a homogeneous polynomial $f$ with other homogeneous elements of $\mathbb{F}_2[a_\mu, t_\mu : \mu \in A^\circ]$ have degrees greater or equal than that of $f$. By (ii), the kernel of $\epsilon_A$ is generated in gradings that contain the sum of a linearly dependent set of characters. So the kernel of $\epsilon_A$ is nontrivial only in degrees that contain the sum of a linearly dependent set of characters. The generating relations $r(T)$ specified in (ii) lie in degrees that are minimal with this property, and they are the unique nontrivial elements in the kernel of $\epsilon_A$ in their degrees. So each of the relations $r(T)$ specified in (ii) is necessary to generate the kernel of $\epsilon_A$. $\square$

We take the time to go through the presentation of Theorem 2.5 for elementary abelian 2-groups of rank at most 3. For the group $C$ with two elements, the representation-graded Bredon homology ring is well studied, and a polynomial algebra on the classes $a$ and $t$,

$$H(C, \star) = \mathbb{F}_2[a, t].$$

This calculation is originally due to Stong (unpublished) and reproved by several authors; the earliest published reference we know of is [9, Section 2].

The ring $H(C^2, \star)$ was calculated by Ellis-Bloor [3, Theorem 4.14]. The group $C^2$ has precisely one dependent set of nontrivial characters, the set $\{p_1, p_2, \mu\}$ of all three nontrivial characters, and this set is minimally dependent. So the presentation of $H(C^2, \star)$ given by Theorem 2.5 has only one relation: the map from $\mathbb{F}_2[a_1, a_2, a_\mu, t_1, t_2, t_\mu]$ that takes the polynomial generators to the classes with the same names factors through an isomorphism of $\mathbb{F}_2$-algebras

$$\mathbb{F}_2[a_1, a_2, a_\mu, t_1, t_2, t_\mu]/(a_1 t_2 t_\mu + t_1 a_2 t_\mu + t_1 t_2 a_\mu) \cong H(C^2, \star).$$

**Example 2.6** (rank 3)  As a proof of concept, we make the minimal presentation given by Theorem 2.5 completely explicit for $A = C^3$. In this case there are 14 polynomial generators, namely the classes $a_\lambda$ and $t_\lambda$ for each of the seven nontrivial characters $\lambda$, and there is also a total of 14 minimal relations. Minimally dependent subsets for $C^3$ have either three or four elements. There are seven minimally dependent subsets of cardinality 3; if $p_i$ denotes the projection of $C^3$ to the $i^{\text{th}}$ factor, then these sets are

$$\{p_1, p_2, p_1 p_2\}, \quad \{p_1, p_3, p_1 p_3\}, \quad \{p_2, p_3, p_2 p_3\}, \quad \{p_1, p_2 p_3, p_1 p_2 p_3\},$$
$$\{p_2, p_1 p_3, p_1 p_2 p_3\}, \quad \{p_3, p_1 p_2, p_1 p_2 p_3\}, \quad \{p_1 p_2, \ p_1 p_3, p_2 p_3\}.$$

This gives seven minimal relations of the form

$$a_\alpha t_\beta t_\gamma + t_\alpha a_\beta t_\gamma + t_\alpha t_\beta a_\gamma,$$

where $\{\alpha, \beta, \gamma\}$ runs over the above seven sets. And there are also seven minimally dependent subsets with four elements, namely

$$\{p_1, p_2, p_3, p_1 p_2 p_3\}, \quad \{p_1, p_2, p_1 p_3, p_2 p_3\}, \quad \{p_1, p_3, p_1 p_2, p_2 p_3\}, \quad \{p_2, p_3, p_1 p_2, p_1 p_3\},$$
$$\{p_1, p_1 p_2, p_1 p_3, p_1 p_2 p_3\}, \quad \{p_2, p_1 p_2, p_2 p_3, p_1 p_2 p_3\}, \quad \{p_3, p_1 p_3, p_2 p_3, p_1 p_2 p_3\}.$$

Each such set $\{\alpha, \beta, \gamma, \delta\}$ gives a minimal relation of the form

$$a_\alpha t_\beta t_\gamma t_\delta + t_\alpha a_\beta t_\gamma t_\delta + t_\alpha t_\beta a_\gamma t_\delta + t_\alpha t_\beta t_\gamma a_\delta.$$

**Remark 2.7** (the number of minimally dependent sets)  We let $A$ be an elementary abelian 2-group of rank $r$. As we explained in Remark 2.4, minimally dependent subsets of $A^\circ$ have at least 3 elements, and at most $r + 1$ elements. We shall now count how many of these there are. We consider $2 \le k \le r$. Every minimally dependent subset with $k + 1$ elements is obtained from a linearly independent $k$-element subset $S$ of $A^\circ$ by adding to it the product of all its members $\sigma = \prod_{\lambda \in S} \lambda$. The resulting minimally dependent subset $S \cup \{\sigma\}$ can be obtained in $k + 1$ different ways from an unordered linearly independent $k$-element set, depending on which of the elements plays the role of the product. An $r$-dimensional $\mathbb{F}_2$-vector space has

$$\frac{(2^r - 1)(2^r - 2) \cdots (2^r - 2^{k-1})}{k!}$$

linearly independent $k$-element subsets. Hence there are

$$\frac{(2^r - 1)(2^r - 2) \cdots (2^r - 2^{k-1})}{(k + 1)!}$$

minimally dependent subsets of $A^\circ$ with $k + 1$ elements. So the total number of relations in the minimal presentation of $H(A, \star)$ given by Theorem 2.5 is

$$\sum_{k=2}^{r} \prod_{i=1}^{k} \frac{2^r - 2^{i-1}}{i + 1}.$$

The following table shows the number $2(2^r - 1)$ of polynomial generators of the representation-graded ring $H(A, \star)$, and the above number of minimal relations for small ranks:

| rank($A$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| number of variables | 2 | 6 | 14 | 30 | 62 | 126 | 254 |
| number of relations | 0 | 1 | 14 | 308 | 20 336 | 4 994 472 | 4 610 816 280 |

While the number of relations grows very quickly with the rank, this is in a sense due to the growing number of automorphisms. Indeed, the automorphism group of $A$ acts transitively on the minimally dependent sets of nontrivial $A$-characters of fixed cardinality. So up to automorphisms of $A$, there is only one relation of degree $k + 1$ for all $2 \leq k \leq \text{rank}(A)$.

# 3  Bredon homology with pre-Euler and inverse Thom classes inverted

In this section we study certain localizations of the representation-graded Bredon homology ring $H(A, \star)$. We fix a subgroup $B$ of $A$ and consider the ring obtained by inverting all pre-Euler classes that restrict nontrivially to $B$, and all inverse Thom classes that restrict trivially to $B$. Our presentation of $H(A, \star)$ yields a presentation of the localized ring; see Theorem 3.2.

The localization by inverting all pre-Euler classes is also known as the geometric fixed point ring of the Eilenberg–MacLane spectrum $H\mathbb{F}_2$; Holler and Kriz [6] previously obtained a presentation of it, which we recover as the special case $B = A$. The mixed localizations were considered by Balmer and Gallauer [1]; we improve their [1, Theorem 17.13], for the prime 2, from a "presentation modulo nilpotence" to an actual presentation; see Remark 3.3.

**Construction 3.1** (localizations)  We let $B$ be a subgroup of an elementary abelian 2-group $A$. We let $H(A|B)$ be the integer-graded part of the localization of the representation-graded ring $H(A, \star)$ obtained by inverting the following classes:

- All classes $a_\lambda$ for all $A$-characters $\lambda$ such that $\lambda|_B$ is nontrivial.
- All classes $t_\lambda$ for all $A$-characters $\lambda$ such that $\lambda|_B$ is trivial.

Every element of $H_k(A|B)$ is then of the form $x/a_V t_W$ for some $A$-representation $V$ with $V^B = 0$, some $A$-representation $W$ with $W^A = 0$ on which $B$ acts trivially, and some $x \in H_{k+\dim(W)}(A, V \oplus W)$. These elements satisfy the relations

$$x/a_V t_W = (x \cdot a_{\overline{V}} t_{\overline{W}})/a_{V \oplus \overline{V}} t_{W \oplus \overline{W}}$$

for all $A$-representations $\overline{V}$ and $\overline{W}$ with the corresponding properties. The localizations introduce enough graded units so that the representation grading effectively collapses to an integer grading. In other words, we are not losing any information by restricting attention to the integer-graded subrings of the localizations.

The following theorem "localizes" the presentation of the representation-graded ring $H(A, \star)$ from Theorem 2.5 to a presentation of the integer-graded ring $H(A|B)$. A noticeable feature is that those relations in the presentation of $H(A, \star)$ that involve minimally dependent sets of four or more characters become redundant in the localization.

**Theorem 3.2** *For every subgroup $B$ of an elementary abelian 2-group $A$, the graded ring $H(A|B)$ is a domain. The ring $H(A|B)$ is generated by*

- *the homogeneous elements $x_\lambda = t_\lambda / a_\lambda$ of degree 1, for $\lambda \in A^\circ$ with $\lambda|_B \neq 1$, and*
- *the homogeneous elements $e_\lambda = a_\lambda / t_\lambda$ of degree $-1$, for $\lambda \in A^\circ$ with $\lambda|_B = 1$.*

*The ideal of relations between these generators is generated by the polynomials*

- $x_\alpha x_\beta + x_\alpha x_\gamma + x_\beta x_\gamma$ *for all triples of nontrivial $A$-characters such that $\alpha \cdot \beta \cdot \gamma = 1$ and such that $\alpha, \beta$ and $\gamma$ are nontrivial on $B$;*

- $x_\alpha + x_\beta + x_\alpha x_\beta e_\gamma$ *for all triples of nontrivial $A$-characters such that $\alpha \cdot \beta \cdot \gamma = 1$, such that $\alpha$ and $\beta$ are nontrivial on $B$, and $\gamma|_B = 1$;*

- $e_\alpha + e_\beta + e_\gamma$ *for all triples of nontrivial $A$-characters such that $\alpha \cdot \beta \cdot \gamma = 1$ and $\alpha|_B = \beta|_B = \gamma|_B = 1$.*

**Proof** The multigraded ring $H(A, \star)$ is a domain by Theorem 2.2, hence so is any localization at a multiplicative subset of homogeneous elements. Since $H(A|B)$ is a subring of such a localization, it is a domain, too.

We write $\chi_+ = \{\lambda \in A^\circ : \lambda|_B \neq 1\}$ and $\chi_- = \{\lambda \in A^\circ : \lambda|_B = 1\}$. If we take the presentation of $H(A, \star)$ given by Theorem 2.5, invert the relevant pre-Euler and inverse Thom classes, and restrict to integer gradings, we recognize $H(A|B)$ as the quotient of the polynomial ring $\mathbb{F}_2[x_\lambda, e_\mu : \lambda \in \chi_+, \mu \in \chi_-]$ by the ideal generated by the polynomials

$$\bar{r}(T) = r(T) / \left( \prod_{\mu \in T \cap \chi_+} a_\mu \cdot \prod_{\mu \in T \cap \chi_-} t_\mu \right) = \left( \sum_{\lambda \in T \cap \chi_+} \prod_{\mu \in (T \cap \chi_+) \setminus \{\lambda\}} x_\mu \right) + \left( \sum_{\lambda \in T \cap \chi_-} e_\lambda \right) \cdot \prod_{\mu \in T \cap \chi_+} x_\mu$$

for all minimally dependent subsets $T$ of $A^\circ$. Among these are the minimally dependent subsets $T = \{\alpha, \beta, \gamma\}$ that have three elements. We note that because $\alpha \cdot \beta \cdot \gamma = 1$, whenever two of $\alpha, \beta$ and $\gamma$ are trivial on $B$, then so is the third. So the relations for minimally dependent subsets with three elements are the ones from the statement of the theorem.

In the rest of the proof we show that the polynomials $\bar{r}(T)$ for minimally dependent subsets $T$ with more than three elements are in the ideal generated by those with three elements. We argue by induction on the cardinality of $T$, and we let $T$ be a minimally dependent subset of $A^\circ$ with at least 4 elements. We pick two distinct elements $\alpha \neq \beta$ from $T$; by minimality, the product $\gamma = \alpha \cdot \beta$ then does not belong to $T$. We set $S = T \setminus \{\alpha, \beta\}$. We claim that the polynomial $\bar{r}(T)$ lies in the ideal generated by $\bar{r}(S \cup \{\gamma\})$ and $\bar{r}(\{\alpha, \beta, \gamma\})$. The sets $S \cup \{\gamma\}$ and $\{\alpha, \beta, \gamma\}$ are both minimally dependent, and both have fewer elements than $T$. So by induction, $\bar{r}(T)$ belongs to the ideal generated by the ternary relations.

It remains to prove the claim. We distinguish two cases, depending on whether the restriction of $\gamma$ to $B$ is trivial or not. If $\gamma|_B = 1$, we exploit the polynomial relation

$$t_\gamma \cdot r(T) = t_\alpha t_\beta \cdot r(S \cup \{\gamma\}) + r(\{\alpha, \beta, \gamma\}) \cdot \prod_{\lambda \in S} t_\lambda$$

that holds by inspection. Because the class $t_\gamma$ is among those being inverted, this relation shows that the polynomial $\bar{r}(T)$ lies in the ideal generated by $\bar{r}(S \cup \{\gamma\})$ and $\bar{r}(\{\alpha, \beta, \gamma\})$. If $\gamma|_B \neq 1$, we exploit the relation

$$a_\gamma \cdot r(T) = (a_\alpha t_\beta + t_\alpha a_\beta) \cdot r(S \cup \{\gamma\}) + r(\{\alpha, \beta, \gamma\}) \cdot r(S)$$

that also holds by inspection. Because the class $a_\gamma$ is among those being inverted, this relation shows that the polynomial $\bar{r}(T)$ lies in the ideal generated by $\bar{r}(S \cup \{\gamma\})$ and $\bar{r}(\{\alpha, \beta, \gamma\})$.                    □

**Remark 3.3** (relation to the work of Balmer and Gallauer) In [1], Balmer and Gallauer also study the representation-graded Bredon homology ring $H(A, \star)$ as input for their computation of the Balmer spectrum of the tt-category of permutation modules; this ring is called the *twisted cohomology ring* of $A$ in [1, Definition 12.16], and denoted $H^{\bullet\bullet}(A)$. The work of Balmer and Gallauer also covers elementary abelian $p$-groups for odd primes $p$, which we do not consider. The connection to Bredon homology comes from the fact that the homotopy category of permutation $A$-modules is equivalent to the homotopy category of modules in genuine $A$-spectra over the Eilenberg–MacLane spectrum $H\mathbb{F}_2$ for the constant Mackey functor. Under this equivalence, the invertible object $u_N$ introduced in [1, Definition 12.3] corresponds to the representation sphere $S^\lambda$. Using our notation, Balmer and Gallauer prove the relation $a_\alpha t_\beta t_\gamma + t_\alpha a_\beta t_\gamma + t_\alpha t_\beta a_\gamma = 0$ in [1, Lemma 17.4], for all triples of $A$-characters whose product is 1.

The localization $H(A|B)$ is introduced as $\mathcal{O}^\bullet_A(B)$ in [1, Definition 14.9]. In [1, Construction 17.5], Balmer and Gallauer define a ring $\underline{\mathcal{O}}^\bullet_A(B)$ by generators and relations as in the presentation of Theorem 3.2. Then they show in [1, Theorem 17.13] that the morphism $\underline{\mathcal{O}}^\bullet_A(B) \to \mathcal{O}^\bullet_A(B)$ becomes an isomorphism after modding out the respective nilradicals. In the notation of Balmer and Gallauer, the content of our Theorem 3.2 is that $\underline{\mathcal{O}}^\bullet_A(B) \to \mathcal{O}^\bullet_A(B)$ is already an isomorphism, without the need to divide out any ideal. In particular, our theorem confirms the expectation formulated in [1, Remark 17.15].

For every genuine equivariant ring spectrum, the localization of the representation-graded homotopy ring at the pre-Euler classes of all nontrivial irreducible representations is isomorphic to the so-called *geometric fixed point ring*; see [4, Proposition 3.20]. In particular, for every elementary abelian 2-group $A$, the localization $H(A|A)$ is isomorphic to the geometric fixed point ring $\Phi^A_*(H\mathbb{F}_2)$ of the Eilenberg–MacLane ring spectrum of the constant Mackey functor $\mathbb{F}_2$. Holler and Kriz gave a presentation of this ring in [6, Theorem 2]. Their result is the special case $A = B$ of Theorem 3.2:

**Corollary 3.4** [6, Theorem 2] *Let $A$ be an elementary abelian 2-group. The graded ring $\Phi^A_*(H\mathbb{F}_2) = H(A|A)$ is generated by the homogeneous elements $x_\lambda = t_\lambda/a_\lambda$ of degree 1, for all nontrivial $A$-*

characters $\lambda$. The ideal of relations between these generators is generated by the quadratic polynomials

$$x_\alpha x_\beta + x_\alpha x_\gamma + x_\beta x_\gamma$$

for all triples of nontrivial $A$-characters such that $\alpha \cdot \beta \cdot \gamma = 1$.

**Remark 3.5** As observed and expanded upon by S Kriz in [8], the ring $H(A|A)$ agrees with the "reciprocal plane of the arrangement of all nontrivial hyperplanes of $A$" previously studied in the algebra literature. We refer to [8] for more information on this point of view. The identification with the reciprocal plane in particular implies that $H(A|A)$ is a domain. The localization maps $H_m(A, W) \to H_m(A|A)$ are injective, see [6, Theorem 5 (ii)] or our Proposition 2.1(ii), so $H(A, \star)$ is a domain, too, which we proved more directly in Theorem 2.2.

If we specialize Theorem 3.2 to the other extreme $B = \{1\}$, we see that $H(A|\{1\})$ is generated by the Euler classes $e_\lambda = a_\lambda/t_\lambda$ for all nontrivial $A$-characters $\lambda$, and the ideal of relations is generated by the linear polynomials $e_\alpha + e_\beta = e_\gamma$ for all triples of nontrivial $A$-characters such that $\alpha \cdot \beta = \gamma$. So if $A$ has rank $r$, and $\lambda_1, \ldots, \lambda_r$ is a basis of the vector space of $A$-characters, then already the Euler classes $e_{\lambda_1}, \ldots, e_{\lambda_r}$ generate $H(A|\{1\})$, and there are no further relations between these. In other words, the ring $H(A|\{1\})$ obtained by inverting all inverse Thom classes is an $\mathbb{F}_2$-polynomial algebra in the classes $e_{\lambda_1}, \ldots, e_{\lambda_r}$, which agrees with the group cohomology ring $H^*(A; \mathbb{F}_2)$.

# References

[1] **P Balmer**, **M Gallauer**, *The geometry of permutation modules*, Invent. Math. 241 (2025) 841–928 MR

[2] **G E Bredon**, *Equivariant cohomology theories*, Lecture Notes in Math. 34, Springer (1967) MR

[3] **B Ellis-Bloor**, *The Klein four homotopy Mackey functor structure of $H\underline{\mathbb{F}_2}$*, honours thesis, Australian National University (2020) Available at `https://openresearch-repository.anu.edu.au/bitstream/1885/258402/1/Honours_thesis.pdf`

[4] **J P C Greenlees**, **J P May**, *Equivariant stable homotopy theory*, from "Handbook of algebraic topology", North-Holland, Amsterdam (1995) 277–323 MR

[5] **M Hausmann**, **S Schwede**, *The universal property of bordism of commuting involutions*, preprint (2024) arXiv 2406.00404

[6] **J Holler**, **I Kriz**, *On RO(G)-graded equivariant "ordinary" cohomology where G is a power of $\mathbb{Z}/2$*, Algebr. Geom. Topol. 17 (2017) 741–763 MR

[7] **S Illman**, *Equivariant singular homology and cohomology, I*, Mem. Amer. Math. Soc. 1956, Amer. Math. Soc., Providence, RI (1975) MR

[8] **S Kriz**, *Equivariant cohomology and the super reciprocal plane of a hyperplane arrangement*, Algebr. Geom. Topol. 22 (2022) 991–1015 MR

[9] **L G Lewis, Jr**, *The RO(G)-graded equivariant ordinary cohomology of complex projective spaces with linear $\mathbb{Z}/p$ actions*, from "Algebraic topology and transformation groups", Lecture Notes in Math. 1361, Springer (1988) 53–122 MR

*Mathematisches Institut, Universität Bonn*
*Bonn, Germany*

*Mathematisches Institut, Universität Bonn*
*Bonn, Germany*

hausmann@math.uni-bonn.de, schwede@math.uni-bonn.de

# The $\mathbb{Z}/p$-equivariant cohomology of the genus-zero Deligne–Mumford space with $1 + p$ marked points

DAIN KIM

NICHOLAS WILKINS

We prove that the Serre spectral sequence of the fibration $\overline{\mathcal{M}}_{0,1+p} \to E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p} \to B\mathbb{Z}/p$ collapses at the $E_2$-page. We use this to prove that a torsion element of the $\mathbb{Z}/p$-equivariant cohomology with $\mathbb{F}_p$-coefficients of genus-zero Deligne–Mumford space with $1 + p$ marked points is lifted from nonequivariant cohomology. We conclude that the only "interesting" $\mathbb{Z}/p$-equivariant operations on quantum cohomology are quantum Steenrod power operations.

## 1 Introduction

### 1.1 Background

The study of equivariant cohomology operations in the context of symplectic geometry is relatively new, but it builds off much older operadic roots. The first notions of these sort of equivariant cohomology operations, specifically useful towards symplectic geometry, appeared in the work independently by Betz [1] and Fukaya [4]. The next appearance of such operations was by Seidel in [8].

More recently, based on the initial definition by Fukaya in [4], Wilkins and Seidel defined and studied in depth the notion of the "quantum Steenrod power operations" [9; 10; 14]. The idea of these papers was to define operations that look like the topological Steenrod power operations of [12], but in the context of quantum cohomology: much as the Steenrod power operations measure the chain-level noncommutativity of the cup product, so do the quantum Steenrod power operations do likewise for the quantum cup product.

All of the above papers follow as their guiding principle the equivariant version of the *symplectic operadic principle*: "The nature of an equivariant symplectic invariant defined by counting holomorphic curves is determined by the operad of the space of domains."

Without delving too deeply into the technical details and definitions of the quantum Steenrod power operations, the general notion for invariants on quantum cohomology is the following: given a natural number $n$, and a permutation group $G < \mathrm{Sym}(n)$, one can consider the Deligne–Mumford space $\overline{\mathcal{M}}_{0,1+n}$.

This space roughly consists of the set of $1 + n$ distinct marked points $(z_0, z_1, \ldots, z_n)$ with $z_i \in S^2$ (with a compactification we describe in Section 2.1), up to Möbius reparametrisations of $S^2$. Then $G$ acts via permuting the last $n$ points. The idea is, for a nice choice of $G$ and $p \mid |G|$ for a prime $p$, that each closed element of $C_*^G(\overline{\mathcal{M}}_{0,1+n}; \mathbb{F}_p)$ should have an associated equivariant quantum operation, and that any boundary in $C_*^G(\overline{\mathcal{M}}_{0,1+n}; \mathbb{F}_p)$ determines a 1-dimensional cobordism (given by a 1-dimensional parametric moduli space) between the moduli spaces determining the chain-level definitions of the two different operations. Hence, one expects that $H_*^G(\overline{\mathcal{M}}_{0,1+n}; \mathbb{F}_p)$ should give a good amount of information about $G$-equivariant quantum operations. One would also wonder the case $p \nmid |G|$, but in this case the $G$-equivariant cohomology of $\mathbb{F}_p$ coefficients is the $G$-invariant part of the ordinary cohomology, which gives no interesting operations.

Now, the quantum Steenrod $p^{\text{th}}$-power operation is defined using $n = p$, $G = \mathbb{Z}/p$, and the class of $H_*^{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ in question is determined by a fixed point of the $G$-action on the Deligne–Mumford space. In particular, viewing $S^2$ as the extended complex plane, say $z_0 = 0$ and $z_1, \ldots, z_p$ are $p^{\text{th}}$-roots of unity (in the obvious order $z_i = \zeta^i$ for any primitive $\zeta$).

When $p = 2$, this fixed point is the entirety of $\overline{\mathcal{M}}_{0,1+2}$, which is just a single point. But when $p$ is larger, the Deligne–Mumford space has more topology: the question of what is the equivariant cohomology of the Deligne–Mumford space is the first step towards answering a question posed by Seidel in [9, Section 5c]. In particular, we study $\mathbb{Z}/p$-actions as opposed to $\text{Sym}_p$-actions. Using the localisation theorem eg [7], we know that the $\mathbb{Z}/p$-equivariant cohomology of $\overline{\mathcal{M}}_{0,1+p}$ matches that of the fixed locus once one inverts the generator $u$ of $H^2(B\mathbb{Z}/p; \mathbb{F}_p)$. In particular, up to $u$-torsion the equivariant cohomology of $\overline{\mathcal{M}}_{0,1+p}$ should look like the fixed-points set of the $\mathbb{Z}/p$-action, which is indeed the case where $z_0 = 0$ and the other $z_i$ are roots of unity, hence quantum Steenrod power operations. Another way of saying this is up to $u$-torsion, the only equivariant operations we obtain in this manner (ie taking classes in $H_*^{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$) are quantum Steenrod power operations.

Finally, this leads us to the core question answered in this paper: "are there any interesting $u$-torsion cycles in $H_{\mathbb{Z}/p}^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$?" A positive answer would have meant potentially interesting new operations, while a negative answer means that every $\mathbb{Z}/p$-equivariant operation is a Steenrod power operation. In this paper, we demonstrate that while there are some $u$-torsion cycles in $H_{\mathbb{Z}/p}^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$, these all correspond to nonequivariant cohomology: basically, they arise as $\mathbb{Z}/p$-invariant cycles in $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$. In conclusion, apart from the resulting nonequivariant operations (which are just defined via standard Gromov–Witten invariants) the only $\mathbb{Z}/p$-equivariant operations we may define are quantum Steenrod power operations. We formalise this in Corollary 4.3.

**Remark 1.1** This should be contrasted with the case where one considers the composition of quantum Steenrod operations (see eg the quantum Adem relation [14, Section 7]): in this setting, $n = p^p$ and $G = \mathbb{Z}/p \int \mathbb{Z}/p$ (the wreath product). At least for $p = 2$ (and expected for larger $p$) there exist interesting and exotic elements of $H_G^*(\overline{\mathcal{M}}_{0,1+n}; \mathbb{F}_p)$ that do not obviously arise as Steenrod power operations.

## 1.2 Summary of paper

We begin in Section 2.1 with recalling some preliminary concepts: in particular, we recall the Deligne–Mumford space, cohomology with local systems, a construction of $\mathbb{Z}/p$-equivariant cohomology, Serre spectral sequence, and the localisation theorem.

In Section 3, we recall a construction of $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ and importantly observe how the action of $\mathbb{Z}/p$ on $\overline{\mathcal{M}}_{0,1+p}$ descends to an action on $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$, noting specifically the fixed classes. We then proceed to construct the associated Serre spectral sequence associated to the fibration

$$\overline{\mathcal{M}}_{0,1+p} \hookrightarrow \overline{\mathcal{M}}_{0,1+p} \times_{\mathbb{Z}/p} E\mathbb{Z}/p \to B\mathbb{Z}/p.$$

Finally, in Section 4 we proceed to prove the main result:

**Theorem 1.2** *Let $\overline{\mathcal{M}}_{0,1+p}$ be the Deligne–Mumford space of genus zero with $1 + p$ marked points. Then the following map is injective:*

$$H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p) \xrightarrow{\rho \oplus i^*} H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p) \oplus H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}^{fix}_{0,1+p}; \mathbb{F}_p),$$

*where $\rho$ is induced by forgetting equivariant parameters and $i^*$ by the inclusion $i: \overline{\mathcal{M}}^{fix}_{0,1+p} \hookrightarrow \overline{\mathcal{M}}_{0,1+p}$.*

To achieve this, we first demonstrate that the Serre spectral sequence collapses on the $E_2$-page. Then we use this collapse to demonstrate that the theorem holds.

### Acknowledgements

## 2 Preliminaries

We first give the definition of the Deligne–Mumford space of genus zero, which is the space of interest. We also recall the definition of equivariant cohomology. To compute the equivariant cohomology of the Deligne–Mumford space $\overline{\mathcal{M}}_{0,1+p}$, we will use a fibration whose corresponding Serre spectral sequence yields this equivariant cohomology. Here, we will need cohomology with twisted coefficients, called a local system. At the end of the section, we will state a localisation theorem, which will be used to show the collapse of a Serre spectral sequence.

## 2.1   The Deligne–Mumford space of genus zero

We give a definition of the Deligne–Mumford space of genus zero with marked points. We use the definition from [6].

A stable nodal curve of genus zero consists of finitely many Riemann spheres $S^2$ where some pairs of two spheres are joined along a nodal singularity so that the adjacency graph of Riemann spheres is a connected tree. An $n$-pointed stable nodal genus-zero curve is a stable nodal genus-zero curve with $n$ marked points on the spheres that are different from nodes so that each of the spheres has at least 3 nodes or marked points. Two stable curves are isomorphic if there is a collection of Möbius transformations from each Riemann sphere to another preserving nodes and marked points. Such a curve is stable in the sense that the identity is the only automorphism.

**Definition 2.1**   (the Deligne–Mumford space of genus zero with $n$ marked points)   For $n \geq 3$, $\overline{\mathcal{M}}_{0,n}$ is the moduli space of isomorphism classes of stable curves of genus zero with $n$ marked points.

We are only interested in the case of $n \geq 3$, for purposes of the applications to equivariant quantum operations. It should also be noted that this $\overline{\mathcal{M}}_{0,n}$ is a smooth compact manifold, for $n \geq 3$.

## 2.2   Cohomology with local systems

Let $X$ be a path-connected and locally path-connected topological space. The $n$-chains of $X$ are finite sums $\sum_i n_i \sigma_i$ where $n_i \in \mathbb{Z}$ and $\sigma_i \colon \Delta^n \to X$ is an $n$-simplex, and the space of $n$-chains is denoted as $C_n(X)$. We may extend the notion of $n$-chains by allowing coefficients $n_i$ to be elements of a fixed abelian group $G$. The cochain complex of $X$ with coefficient $G$ is defined as $\mathrm{Hom}(C_n(X), G)$. We call the cohomology (resp homology) induced by such cochains (resp chains) as cohomology with coefficients (resp homology with coefficients).

We may extend the cohomology with coefficients further by twisting the coefficient, which leads to the notion of a local system. A *local system* on $X$ is a locally constant sheaf of abelian groups on $X$.

**Definition 2.2**   Let $\mathcal{A}$ be a local system on $X$. If $\Delta^n = [v_0, v_1, \dots, v_n]$, then the cochain complex of $X$ with local system $\mathcal{A}$ is defined as

$$C^n(X; \mathcal{A}) = \prod_{\sigma \colon \Delta^n \to X} \mathcal{A}(\sigma(v_0))$$

with differential $\delta \colon C^n(X; \mathcal{A}) \to C^{n+1}(X; \mathcal{A})$ given as

$$(-1)^n (\delta c)(\sigma) = \mathcal{A}(\gamma)^{-1} c(\partial_0 \sigma) + \sum_{i=1}^{n+1} (-1)^i c(\partial_i \sigma)$$

for $c \in C^n(X; \mathcal{A})$ and $\sigma \colon \Delta^{n+1} \to X$ where $\gamma$ is a path $t \mapsto \sigma((1-t)v_0 + t v_1)$ and $\mathcal{A}(\gamma)$ is the map from the stalk at $\sigma(v_0)$ to the stalk at $\sigma(v_1)$ induced by $\gamma$. The *cohomology of $X$ with local system $\mathcal{A}$* is $H^*(X; \mathcal{A}) := \ker \delta / \mathrm{im}\, \delta$.

We refer the reader to [13, Chapter 4] for more details. The local coefficient $G$ may be viewed as the local system that is the trivial bundle $X \times G$.

Even though the local system itself is complicated, if the local system is decomposable as a finite direct sum of local systems, then the cohomology can be written as the direct sum of cohomology with each component of the local system.

**Proposition 2.3** *Let* $X$ *be a path-connected and locally path-connected topological space and* $F$ *a field. If a local system* $\mathcal{A}$ *is decomposed as* $\mathcal{A} = \mathcal{B} \oplus \mathcal{C}$ *as vector spaces over* $F$ *for local systems* $\mathcal{B}$ *and* $\mathcal{C}$ *on* $X$, *then*

$$H^*(X; \mathcal{A}) = H^*(X; \mathcal{B}) \oplus H^*(X; \mathcal{C}).$$

**Proof**   Since $\mathcal{A} = \mathcal{B} \oplus \mathcal{C}$,

$$C^*(X; \mathcal{A}) = C^*(X; \mathcal{B}) \oplus C^*(X; \mathcal{C}), \quad \mathcal{A}(\gamma) = \mathcal{B}(\gamma) \oplus \mathcal{C}(\gamma)$$

canonically. Therefore, $\delta_{\mathcal{A}} = \delta_{\mathcal{B}} \oplus \delta_{\mathcal{C}}$ for differentials corresponding to local systems, which leads to the desired direct decomposition of cohomology.     $\square$

## 2.3   $\mathbb{Z}/p$-equivariant cohomology

We begin by defining the classifying space $BG$, for some group $G$. First we consider $EG$, which is a contractible space equipped with a free $G$-action. In our case, with $G = \mathbb{Z}/p$, we may pick $E\mathbb{Z}/p = S^\infty \subset \mathbb{C}^\infty = \bigcup_j \mathbb{C}^j$, taking the $\mathbb{Z}/p$-action to be diagonal multiplication by $e^{2\pi i/p}$. In general, the classifying space is

$$BG := EG/G.$$

Then, given some topological space $X$ equipped with a $G$-action $\sigma$, we define its homotopy quotient $EG \times_G X$ to be the quotient of $EG \times X$ by the diagonal action (abusively) denoted by

$$\sigma(g) \cdot (v, x) = (e^{2\pi i/p} v, \sigma(g) \cdot x),$$

for $g \in G$.

**Definition 2.4**   Let $X$ be a topological space, and a topological group $G$ acts continuously on $X$. Then *the equivariant cohomology of* $X$ *with action* $G$ *and a coefficient* $F$ is

$$H_G^*(X; F) := H^*(EG \times_G X; F),$$

where $EG$ is the universal cover of the classifying space $BG$.

We notice further that, in the case that $X$ is fixed by the $G$-action, the homotopy quotient $EG \times_G X$ equals $BG \times X$, and hence one can use the Künneth isomorphism to write

$$H_G^*(X; F) = H^*(EG \times_G X; F) \cong H^*(BG; F) \otimes H^*(X; F).$$

We are interested in the case where $G = \mathbb{Z}/p$. Indeed, noting that $\overline{\mathcal{M}}_{0,3}$ is a single point (and therefore has no torsion equivariant cohomology), we may restrict our attention to $n \geq 4$: thus, in this paper we will assume that $p > 2$.

When $F = \mathbb{F}_p$, we follow the conventions of [9, Equation (5.5)] (with the natural extension for general coefficient systems). In particular, we fix some generator $g$ of $\mathbb{Z}/p$ and write

$$C^*_{\mathbb{Z}/p}(X; \mathbb{F}_p) = \mathbb{F}_p[u] \otimes \Lambda(e) \otimes C^*(X; \mathbb{F}_p),$$

where $|e| = 1$ and $|u| = 2$, with

$$d(u^k \otimes c) = eu^k \otimes (gc - c) + u^k \otimes dc,$$
$$d(u^k e \otimes c) = u^{k+1} \otimes (c + gc + \cdots + g^{p-1}c) - u^k e \otimes dc.$$

## 2.4 Serre spectral sequence

**Theorem 2.5** [11, Proposition 6] *If $F \to E \xrightarrow{\pi} B$ is a Serre fibration and $G$ a fixed abelian group, there is a spectral sequence*

$$E_2^{p,q} = H^p(B; \mathcal{H}^q(F; G)) \Rightarrow H^*(E; G),$$

*which is called a Serre spectral sequence. Here, $\mathcal{H}^q(F; G)$ is the local system over $B$ with stalk $H^q(F; G)$ induced by the Serre fibration $\pi$. If $B$ is a CW complex, denote by $B^k$ the $k$-skeleton of $B$. We have a filtration of $E$ as $E^k = \pi^{-1}(B^k)$. Then the filtration of $H^m(E; G)$ is given as $F^k H^m(E; G) = \ker(H^m(E; G) \to H^m(E^{k-1}; G))$.*

For brevity, we define $F_k^m := F^k H^m(E; G)$.

Theorem 2.5 shows that the $E^2$-page of the Serre spectral sequence corresponding to a fibration $X \hookrightarrow EG \times_G X \to BG$ with coefficient $\mathbb{F}_p$ is

$$H^i(BG; \mathcal{H}^j(X; \mathbb{F}_p))$$

and converges to $H^*(EG \times_G X; \mathbb{F}_p) = H_G^*(X; \mathbb{F}_p)$.

## 2.5 Localisation

The intuitive notion of the localisation theorem is that, up to $u$-torsion, all the equivariant cohomology classes of a space $X$ are encoded in the fixed-point space $X^{\text{fix}}$. We will use the localisation theorem on the Deligne–Mumford space $\overline{\mathcal{M}}_{0,1+p}$. The fixed-point space of the Deligne–Mumford space consists of a set of points, and so the localisation theorem provides a powerful restriction on the equivariant cohomology of $\overline{\mathcal{M}}_{0,1+p}$.

We restate the localisation theorem given by Quillen as a version we will use to the Deligne–Mumford space $\overline{\mathcal{M}}_{0,1+p}$ with a group $\mathbb{Z}/p$.

**Theorem 2.6** [7, Theorem 4.2]  *For a compact manifold $X$ with an action of $\mathbb{Z}/p$, the inclusion of the fixed-point set $X^{fix} \hookrightarrow X$ induces an isomorphism*

$$H^*_{\mathbb{Z}/p}(X; \mathbb{F}_p)[u^{-1}] \xrightarrow{\cong} H^*_{\mathbb{Z}/p}(X^{fix}; \mathbb{F}_p)[u^{-1}],$$

*where $u$ is a generator of $H^2(B\mathbb{Z}/p; \mathbb{F}_p)$.*

# 3   The structure of $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ and the Serre spectral sequence

In this section, we investigate how $\mathbb{Z}/p$ acts on $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$, which will be used to compute the Serre spectral sequence. We also compute $\overline{\mathcal{M}}^{\mathbb{Z}/p}_{0,1+p}$, which will appear in the localisation theorem.

## 3.1   The cohomology of $\overline{\mathcal{M}}_{0,1+p}$ and its fixed point under the $\mathbb{Z}/p$-action

The cohomology ring structure of $\overline{\mathcal{M}}_{0,1+p}$ was completely described by Keel in [5], and later an equivalent description but in symmetric basis was given by [3]. We will use the basis by [3] to get a benefit of its symmetry.

**Theorem 3.1** [3, Theorem 5.5]  *For the Deligne–Mumford space $\overline{\mathcal{M}}_{0,1+p}$, let the marked points be $x_1, \dots, x_{p+1}$ and $X = \{x_1, \dots, x_p\}$. For each $S \subset X$ with $|S| \geq 3$, there exists a cohomology class $\Pi_S \in H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$ so that monomials of the form $\prod_{|S| \geq 3} \Pi_S^{d_S}$ satisfy the following conditions freely generate $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$ as a $\mathbb{Z}$-module.*

- *If both $d_S$ and $d_T$ are positive, then $S \cap T \in \{\varnothing, S, T\}$.*

- *For each $S$ with $d_S > 0$, if $S_1, \dots, S_k$ are disjoint and form the maximal proper subsets of $S$ with $d_{S_i} > 0$, then*

$$d_S < k - 1 + |S| - \sum_i |S_i|.$$

*Moreover, if $\sigma \colon X \to X$ is an action induced by $1 \in \mathbb{Z}/p\mathbb{Z}$, ie $\sigma(x_i) = x_{i+1}$ for $1 \leq i \leq p-1$, $\sigma(x_p) = x_1$, and $\sigma(x_{p+1}) = x_{p+1}$, then $\sigma(\Pi_S) = \Pi_{\sigma(S)}$ and is compatible with the ring structure of $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$.*

In the following corollary, we adapt the aforementioned basis for $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$ and find the fixed classes under the $\mathbb{Z}/p$-action.

**Corollary 3.2**  *There is a basis for $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$, on which $\sigma^*$ acts, which contains exactly $p-1$ fixed elements (and the rest partitioning into $\mathbb{Z}/p$ orbits of size $p$).*

**Proof**  We take the basis of $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$ described in Theorem 3.1. Suppose that $\prod_{|S| \geq 3} \Pi_S^{d_S}$ is fixed under $\sigma$, and $d_T > 0$ for some $T$. Let $1 \leq i < j \leq p$ be two elements in $T$.

Since the monomial $\prod_{|S| \geq 3} \Pi_S^{d_S}$ is fixed under $\sigma$ and $d_T > 0$, it follows that $d_{\sigma(T)} > 0$. Repeating the argument, we also have $d_{\sigma^{j-i}(T)} > 0$.

As $i \in T$, by shifting the labels $j - i$ times, $j \in \sigma^{j-i}(T)$ and $|T| = |\sigma^{j-i}(T)|$, so by the first condition in Theorem 3.1, $\sigma^{j-i}(T) = T$. But as $p$ is a prime and $0 < j - i < p$, ie $j - i$ is invertible in $\mathbb{F}_p^\times$, we obtain $\sigma(T) = T$, which implies that $T = X$. Then the second condition of Theorem 3.1 forces that $1 \le d_X \le p - 1$. In the other direction, the cohomology classes $\Pi_X^k$ are fixed since $\sigma(X) = X$.

Any other monomials form several $p$-cycles since for any $S \subset X$, $\sigma^p(S) = S$. This proves the corollary for $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$, ie the set of monomials is decomposed into $p - 1$ fixed points and some $p$-cycles.

Since $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$ is free, the generators of $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ are obtained by reducing the coefficients of the generators in mod $p$. There are no relations among the generators because otherwise we can lift the relation to $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{Z})$. Furthermore, the $\mathbb{Z}/p$ action acts in the same way, so the desired result follows. $\qquad\square$

To use the localisation theorem, we investigate the fixed points of $\overline{\mathcal{M}}_{0,1+p}$ under the $\mathbb{Z}/p$-action. To be explicit, elements of $\overline{\mathcal{M}}_{0,1+p}$ consist of a nodal Riemann surface $S$ along with a $(1+p)$-tuple $(z_1, \ldots, z_p, z_{p+1}) \subset S^{p+1}$, and the $\mathbb{Z}/p$-action acts on the indices of $z_1, \ldots, z_p$, ie $n \cdot z_j = z_{j+n \bmod p}$, while fixing $z_{p+1}$.

**Proposition 3.3** *There are exactly $p - 1$ many fixed points of $\overline{\mathcal{M}}_{0,1+p}$ under the $\mathbb{Z}/p$-action.*

**Proof** We denote by $\overline{\mathcal{M}}_{0,1+p}^{\mathbb{Z}/p}$ the fixed locus under the $\mathbb{Z}/p$-action. Further, the marked points on a stable curve $C \in \overline{\mathcal{M}}_{0,1+p}^{\mathbb{Z}/p}$ are denoted by $x_1, \ldots, x_{p+1}$.

We first show that $C$, as a nodal curve, has exactly one $\mathbb{C}P^1$ component. Suppose for a contradiction that $C$ is composed of more than one sphere. The point $x_{p+1}$ must then be alone, because if $x_j$ is on the same sphere then so too should all of the other $x_i$, which would imply $C$ has only one Riemann sphere. Since the number of marked points on the sphere containing $x_1$ is preserved throughout the action of $\mathbb{Z}/p$, call this number $b$, then for any sphere that contains a marked point in $\{x_1, \ldots, x_p\}$, that sphere must have exactly $b$ marked points from that set. Denote by $c$ the number of such spheres containing a marked point in $\{x_1, \ldots, x_p\}$. Since $p$ is a prime, and $p = b \cdot c$, there are 2 cases ($b = 1$ or $c = 1$): each sphere has at most one marked point or $x_1, \ldots, x_p$ are on one sphere and $x_{p+1}$ is on another sphere. But both cases lead to a contradiction because both lead to trees of spheres with leafs containing less than 2 marked points (hence less than 3 special points overall). Hence $C$ has exactly one Riemann sphere.

Therefore, $x_1, x_2, \ldots, x_{p+1}$ lie on one Riemann sphere. By applying a Möbius transformation, we may assume that $x_1 = 1, x_2 = \eta, x_{p+1} = 0$ for a root of unity $\eta$ with $\eta^p = 1$. Let $\sigma$ be the action of $1 \in \mathbb{Z}/p$. Then

$$\sigma x_{p+1} = x_{p+1} = 0, \tag{3-1}$$

$$\sigma x_1 = x_2 = \eta, \tag{3-2}$$

$$\sigma^{p-1} x_2 = x_1 = 1. \tag{3-3}$$

Since $C$ and $\sigma C$ are isomorphic, there is a Möbius transformation $z \mapsto \frac{az+b}{cz+d}$ which gives an isomorphism between two stable curves. Then (3-1) gives $b = 0$ and by (3-2) we may assume that $a = \eta$ and $c + d = 1$. The Möbius transformation given by $\sigma^{p-1}$ is a rational function whose numerator is constant and denominator is of degree $p - 1$ polynomial in $c$. This can be proven by induction on the exponent of $\sigma$. Therefore, applying (3-3) to this rational function, a solution $\sigma$ satisfying (3-1), (3-2) and (3-3) is in fact a solution of a degree $p - 1$ polynomial over $\mathbb{C}$, hence there are at most $p - 1$ many such Möbius transformations. Every fixed point of $\overline{\mathcal{M}}_{0,1+p}$ must have that the induced $\mathbb{Z}/p$ action is such a Möbius action, and any such action completely determines the points $x_3, \ldots, x_p$. Hence, $p - 1$ is similarly an upper bound of the number of fixed points of $\overline{\mathcal{M}}_{0,1+p}$ under the $\mathbb{Z}/p$-action.

Now we describe these $p - 1$ fixed points. Those points are given as $x_{p+1} = 0$, $x_1 = 1$, and $x_k = \eta^{s(k-1)}$ for $1 \leq s \leq p - 1$ and $2 \leq k \leq p + 1$. Denote the stable curve by $C_s$. Indeed, they are fixed under $\mathbb{Z}/p$ since Möbius transformation $z \mapsto \eta^s z$ maps $C$ to $\sigma C$. To show that they represent distinct isomorphism classes, suppose for a contradiction that there is a Möbius transformation $z \mapsto \frac{az+b}{cz+d}$ from $C_{s_1}$ to $C_{s_2}$ for $1 \leq s_1 < s_2 \leq p-1$. There exists an integer $2 \leq r \leq p-1$ with $rs_1 \equiv s_2 \pmod{p}$. The Möbius transformation maps the marked points of $C_{s_1}$ to the marked points of $C_{s_2}$ and so does the map $z \mapsto z^r$. Therefore,

$$\frac{az + b}{cz + d} = z^r$$

has at least $p+1$ solutions $z = 0, 1, \eta^{s_1}, \ldots, \eta^{s_1(p-1)}$. But as $r \leq p-1$, the equation is equivalent to a polynomial of degree at most $p$, which forces the map to be the identity. Therefore, $s_1 = s_2$, a contradiction. $\square$

## 3.2 The $E_2$-page of the Serre spectral sequence

We compute the $E_2$-page of the Serre spectral sequence for the fibration

$$\overline{\mathcal{M}}_{0,1+p} \to E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p} \to B\mathbb{Z}/p.$$

**Proposition 3.4** *If $A$ is the trivial $\mathbb{Z}/p$-representation of $\mathbb{F}_p$, then*

$$H^i(B\mathbb{Z}/p; A) \cong \mathbb{F}_p$$

*for all $i$.*

**Proof** Section 2a in [10] gives a cell decomposition of $E\mathbb{Z}/p$ with coefficient $\mathbb{F}_p$ as follows. There is $\Delta^i \in C^i(E\mathbb{Z}/p; \mathbb{F}_p)$ such that $\Delta^i, \sigma\Delta^i, \ldots, \sigma^{p-1}\Delta^i$ freely span $C^i(E\mathbb{Z}/p; \mathbb{F}_p)$ where $\sigma$ is the action of $1 \in \mathbb{Z}/p$, and the differential is given as

$$(3\text{-}4) \qquad \delta\Delta^i = \begin{cases} \sigma\Delta^{i+1} - \Delta^{i+1}, & i \text{ even,} \\ \Delta^{i+1} + \sigma\Delta^{i+1} + \cdots + \sigma^{p-1}\Delta^{i+1}, & i \text{ odd.} \end{cases}$$

Therefore each $H^i(B\mathbb{Z}/p; A)$ is spanned by $\Delta^i$, and $\delta\Delta^i = 0$. Since every differential is zero, no $\Delta^i$ is exact. Hence, $H^i(B\mathbb{Z}/p; A) \cong \mathbb{F}_p$. $\square$

**Proposition 3.5** *If $A$ is a $p$-dimensional $\mathbb{F}_p$-vector space $V = \operatorname{Span}(v_1, \ldots, v_p)$ with the induced $\mathbb{Z}/p$-action on indices, then*

$$H^i(B\mathbb{Z}/p; A) \cong \begin{cases} \mathbb{F}_p, & i = 0, \\ 0, & i > 0. \end{cases}$$

**Proof** This is an immediate consequence of the Eckmann–Shapiro lemma, see [2], relating

$$H^*(B\mathbb{Z}/p; A) \cong H^*(\mathbb{Z}/p; A) \cong H^*(1; \mathbb{F}_p)$$

(noting that $A$ is the regular representation of $\mathbb{Z}/p$). $\qquad\square$

**Lemma 3.6** *The $E_2$-page of the Serre spectral sequence for the fibration*

$$\overline{\mathcal{M}}_{0,1+p} \to E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p} \to B\mathbb{Z}/p$$

*is*

$$E_2^{i,j} \cong \begin{cases} \mathbb{F}_p^{r_j} \times \mathbb{F}_p, & i = 0,\ 0 \le j \le 2(p-2),\ j \text{ even}, \\ \mathbb{F}_p, & i \ge 1,\ 0 \le j \le 2(p-2),\ j \text{ even}, \\ 0, & \text{otherwise}, \end{cases}$$

*where $r_j$ is the number of copies of the regular representation in cohomological degree $j$, appearing in the local system associated to $H^j(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ over $B\mathbb{Z}/p$.*

*In particular, as a ring, $E_2^{i,j}$ is generated by*

- *$1 \otimes \alpha \in E_2^{0,2}$,*
- *$u \otimes 1 \in E_2^{2,0}$,*
- *$e \otimes 1 \in E_2^{1,0}$,*
- *some further additional elements of $E_2^{0,*}$,*

*subject to at least the following two relations:*

- *these additional elements of $E_2^{0,*}$ are eliminated by multiplication by $u \otimes 1$ and $e \otimes 1$,*
- *$e \otimes 1$ is eliminated by multiplication with $e \otimes 1$.*

**Proof** By Theorem 2.5,

$$E_2^{i,j} = H^i(B\mathbb{Z}/p; H^j(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)) \Rightarrow H^*(E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p) = H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p).$$

Corollary 3.2 breaks down the local system into $\mathbb{Z}/p$ representations. Proposition 2.3 then tells one that we only need to consider Propositions 3.4 and 3.5.

The ring structure follows because Proposition 2.3 holds for cohomology with coefficients in a local system of rings, and the proof of Corollary 3.2 tells us exactly that the splitting of $H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ as $\mathbb{Z}/p$-representations is also a splitting as rings. $\qquad\square$

# 4 Main result

## 4.1 The spectral sequence collapses

**Theorem 4.1**  *The Serre spectral sequence for the fibration*

$$\overline{\mathcal{M}}_{0,1+p} \to E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p} \to B\mathbb{Z}/p$$

*collapses at the $E_2$-page.*

**Proof**  Lemma 3.6 gives the $E_2$-page of the spectral sequence. In the notation of the previous sections, we define

$$H^*(B\mathbb{Z}/p; \mathbb{F}_p) = \mathbb{F}_p[u] \otimes \Lambda[e] \quad \text{and} \quad H^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)^{\text{fix}} = \mathbb{F}_p[\alpha]/(\alpha^{p-1}),$$

where $\deg \alpha = 2$. Then, when $i > 0$, one can write elements of $E_2^{i,j}$ of the form $e^\epsilon u^l \otimes \alpha^\lambda$. We will abusively write this as $\alpha^\lambda e^\epsilon u^l$.

Suppose for a contradiction that the spectral sequence does not collapse at the $E_2$-page.

**Case 1**  There is some nonvanishing differential on $\alpha^\lambda e^\epsilon u^l$, where $\epsilon \in \{0, 1\}$.

In particular, there is some minimal $r$ and a nontrivial differential $d^r$ on some page $E_r$ such that $d^r(\alpha^\lambda u^i) \neq 0$ or $d^r(\alpha^\lambda e u^i) \neq 0$ for some $i$. If $d^r(\alpha^\lambda e u^i) \neq 0$, then we have

$$d^r(\alpha^\lambda e u^i) = d^r(\alpha^\lambda u^i)e + (d^r e)\alpha^\lambda u^i = d^r(\alpha^\lambda u^i)e \neq 0,$$

which implies that $d^r(\alpha^\lambda u^i) \neq 0$. Hence, we may only consider the case $d^r(\alpha^\lambda u^i) \neq 0$. Since

$$d^r(\alpha^\lambda u^{i+j}) = d^r(\alpha^\lambda u^i)u^j + \alpha^\lambda u^i d^r(u^j) = d^r(\alpha^\lambda u^i)u^j$$

and $u$ acts injectively on the $E_2, \ldots, E_r$ pages except for those generators arising from $p$-cycles of generators by Proposition 3.5, this implies that $d^r(\alpha^\lambda u^{i+j}) \neq 0$. Hence, there are infinitely many nontrivial differentials. In particular, for some sufficiently large $m$, there is an infinite sequence $\{m + 2k\}_{k \geq 0}$ such that:

$$p - 1 = \sum_{i+j=m+2k} \dim_{\mathbb{F}_p} E_2^{i,j} > \sum_{i+j=m+2k} \dim_{\mathbb{F}_p} E_\infty^{i,j} = \dim_{\mathbb{F}_p} H_{\mathbb{Z}/p}^{m+2k}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p).$$

We write $m = 2n + \epsilon$ with $\epsilon \in \{0, 1\}$.

Now consider $\iota^*: H_{\mathbb{Z}/p}^m(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p) \to H_{\mathbb{Z}/p}^m(\overline{\mathcal{M}}_{0,1+p}^{\text{fix}}; \mathbb{F}_p)$. We know that once we invert $u$ this becomes an isomorphism by Theorem 2.6. Denote this isomorphism by

(4-1) $$\iota^*[u^{-1}]: H_{\mathbb{Z}/p}^*(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)[u^{-1}] \to H_{\mathbb{Z}/p}^*(\overline{\mathcal{M}}_{0,1+p}^{\text{fix}}; \mathbb{F}_p)[u^{-1}].$$

However, for the infinite sequence of $\{m + 2k\}_{k \geq 0}$ described above, the $\iota^*$ cannot be surjective due to the fact that $\dim H_{\mathbb{Z}/p}^m(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p) < p - 1$, but

$$\dim H_{\mathbb{Z}/p}^m(\overline{\mathcal{M}}_{0,1+p}^{\text{fix}}; \mathbb{F}_p) = \dim(H^0(\overline{\mathcal{M}}_{0,1+p}^{\text{fix}}; \mathbb{F}_p) \otimes H^m(B\mathbb{Z}/p; \mathbb{F}_p)) = p - 1.$$

Here we use the Künneth isomorphism for the first equality, and Proposition 3.3 for the second (ie $\dim H^0(\overline{\mathcal{M}}^{\text{fix}}_{0,1+p}; \mathbb{F}_p) = p-1$).

By the pigeonhole principle, there exists $v \in H^0(\overline{\mathcal{M}}^{\text{fix}}_{0,1+p}; \mathbb{F}_p) \cong \mathbb{F}_p^{\oplus(p-1)}$, and a set of strictly increasing integers $n_1, n_2, \ldots$ such that $v \otimes e^\epsilon u^{n+n_i} \in H^{m+2n_i}(\overline{\mathcal{M}}^{\text{fix}}_{0,1+p}; \mathbb{F}_p)$ is not hit by $\iota^*$ for each $n_i$.

Now, considering $v \otimes e^\epsilon u^n$ as an element of $H^*(\overline{\mathcal{M}}^{\text{fix}}_{0,1+p}; \mathbb{F}_p)[u^{-1}]$, we see that it must be hit by some element of $H^*_{\mathbb{Z}/p}(X)[u^{-1}]$ under $\iota[u^{-1}]$, as (4-1) is an isomorphism.

Hence, there is some $\sum_j a_j u^{-j} \in H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)[u^{-1}]$ with $a_j \in H^{m+2j}_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ such that

$$\iota^*[u^{-1}]\Big(\sum_j a_j u^{-j}\Big) = v \otimes e^\epsilon u^n.$$

Let $J$ be maximal such that $a_J$ is nonzero. Then if $k > J$, we obtain that

$$\sum_j a_j u^{k-j} \in H^{m+2k}_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p),$$

and in $H^{m+2k}_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)[u^{-1}]$ we see that

$$\sum_j a_j u^{k-j} = \Big(\sum_j a_j u^{-j}\Big)u^k.$$

Then for $\nu: H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}^{\text{fix}}_{0,1+p}) \hookrightarrow H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}^{\text{fix}}_{0,1+p})[u^{-1}]$ and $\eta: H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) \to H^*_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p})[u^{-1}]$,

$$\nu\iota^*\Big(\sum_j a_j u^{k-j}\Big) = \iota^*[u^{-1}]\eta\Big(\sum_j a_j u^{k-j}\Big)$$
$$= \iota^*[u^{-1}]\Big(\Big(\sum_j a_j u^{-j}\Big)u^k\Big)$$
$$= \iota^*[u^{-1}]\Big(\sum_j a_j u^{-j}\Big)u^k$$
$$= v \otimes e^{\epsilon_i} u^{n_i+k}.$$

Now we choose $k > J$ such that $k = n_l$ for some $l$. This yields a contradiction, because $\iota^*$ does not surject onto $v \otimes e^\epsilon u^{n+k} = v \otimes e^\epsilon u^{n+n_l}$.

**Case 2** There is some nonvanishing differential on some element of the spectral sequence arising from $p$-cycles in $E^{0,j}_2$ (the additional generators in the statement of Lemma 3.6).

There is a minimal $r$ and a nontrivial differential $d^r$ on $E_r$ such that $d^r(x) \neq 0$ for $x \in \mathbb{F}_p^{\#\text{ of }p\text{-cycles}} \leq E^{0,j}_2$. Note that $xu = 0$. But since $u$ acts injectively apart from such $x$, we deduce $0 = d^r(0) = d^r(xu) = d^r(x)u$ is nonzero, a contradiction.

From the above two cases, we conclude that the spectral sequence collapses at the $E_2$-page. $\qquad\square$

We now have all of the required components to prove the main theorem.

## 4.2 The proof of Theorem 1.2

Since the spectral sequence collapses at the $E_2$-page, we have

$$\phi \colon H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p) \xrightarrow{\sim} \bigoplus_{i+j=m} E_2^{i,j},$$

an isomorphism of $\mathbb{F}_p$ vector spaces.

We assume, for a contradiction, that there exists some nonzero $x \in H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}; \mathbb{F}_p)$ such that $\rho(x) = 0$ and $\iota^*(x) = 0$. Since $\phi$ is an isomorphism, there is a decomposition of $x = x_0 + x_1 + \cdots + x_m$ so that $\phi(x_i) \in E_2^{i,m-i}$.

We first prove that $x_0 = 0$. By the second statement of Theorem 2.5,

$$F^0 H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) = H^m(E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p}),$$
$$F^1 H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) = \ker(H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) \to H^m(\overline{\mathcal{M}}_{0,1+p})).$$

Hence the map $\phi_0 \colon H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) \to E_2^{0,m}$, the projection of $\phi$ onto $E_2^{0,m}$, is given as

$$H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) \to H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p})/\ker(H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) \to H^m(\overline{\mathcal{M}}_{0,1+p}))$$

induced by the projection. Let $p \colon (E\mathbb{Z}/p \times_{\mathbb{Z}/p} \overline{\mathcal{M}}_{0,1+p}, \overline{\mathcal{M}}_{0,1+p}) \to (B\mathbb{Z}/p, *)$. Then



Since the domain of $d^m$ is $H^m(\overline{\mathcal{M}}_{0,1+p})$, the transgression gives that for every $\alpha \in H^m(\overline{\mathcal{M}}_{0,1+p})$ there exists $\beta \in H^{m+1}(B\mathbb{Z}/p, *)$ with $\delta\alpha = p^*\beta$ and $d^m\alpha = \widetilde{\beta} \in H^{m+1}(B\mathbb{Z}/p)$, where $\widetilde{\beta}$ is the image of $\beta$ in $H^{m+1}(B\mathbb{Z}/p, *) \to H^{m+1}(B\mathbb{Z}/p)$. But as $d^m$ vanishes, $\beta = 0$, and hence $\delta = 0$. Then by the exactness of the first row, $\rho$ is surjective, and therefore $\rho$ induces the same map as $\phi_0$. In particular, $\phi(x_0) = \phi_0(x_0) = \rho(x) = 0$, and hence $x_0 = 0$.

Now we prove that $x_i = 0$ for all $i \geq 1$. Let $j \geq 1$ be the minimal index such that $x_j \neq 0$. Let $u$ be the cohomology class as in the proof of Theorem 4.1, and $v \in F_2^2/\langle 0 \rangle$ the class of $u$ on the level of the spectral sequence. Let $\phi^m_k \colon H^m_{\mathbb{Z}/p}(\overline{\mathcal{M}}_{0,1+p}) \to E_2^{k,m-k}$ be the projection map of $\phi$ onto $E_2^{k,m-k}$.

Since the cup product induces $F^m_j \times F^{2k}_{2k} \to F^{m+2k}_{j+2k}$ and this product induces

$$F^m_j/F^m_{j+1} \times F^{2k}_{2k} \to F^{m+2k}_{j+2k}/F^{m+2k}_{j+2k+1},$$

we obtain $\phi_j^m(x_j)v^k = \phi_{j+2k}^{m+2k}(x_j u^k)$. But as in the proof of Theorem 4.1, $\phi(x_j)v^k \neq 0$, and hence

$$\phi_{j+2k}^{m+2k}(x_j u^k) \neq 0.$$

In the case where $s > j$, we can run a similar argument using the inclusion $(F_s^m, F_{s+1}^m) \to (F_j^m, F_{j+1}^m)$ and considering $F_s^m / F_{s+1}^m \times F_{2k}^{2k} \to F_{j+2k}^{m+2k} / F_{j+2k+1}^{m+2k}$. In particular, we have $\phi_{j+2k}^{m+2k}(x_s u^k) = \phi_j^m(x_s)v^k = 0$. Therefore, $\phi_{j+2k}^{m+2k}(x u^k) \neq 0$. In particular, $x u^k \neq 0$ for all $k$. Then by Theorem 2.6, $\iota^*(x) \neq 0$, which is a contradiction. Therefore, $x = 0$, and the desired result follows.

## 4.3 The connection with quantum operations

In this section, we will describe the relevance of this topological result to symplectic geometry. In order to circumvent the usual requirements of long technical definitions, we will instead provide citations to the relevant technical work for the interested reader.

Fix some prime $p$. Suppose that $(M, \omega)$ is a symplectic manifold, with appropriate technical conditions to define quantum cohomology. As mentioned above, we will not give a formal definition of quantum cochains $\mathrm{QC}^*(M)$ or quantum cohomology $\mathrm{QH}^*(M)$, nor the conditions necessary to define them, but we refer the reader to the definition of quantum cohomology in [6].

**Definition 4.2** We will say that, given a parametrised moduli space of holomorphic maps $\mathcal{M}$, this moduli space is *$G$-equivariantly parametrised of order $a$* if:

• there is a $G$-action on $\{1, \ldots, a\}$, inducing an action on the latter $a$ marked points of $\overline{\mathcal{M}}_{0,1+a}$ (the $1+a$ marked points being denoted by $(m_0, m_1, \ldots, m_a)$),

• there is some $i \in \mathbb{Z}_{\geq 0}$ and some finite-dimensional smooth manifold $A \subset \overline{\mathcal{M}}_{0,1+a} \times_G EG$ representing a homology class in $H_*^G(\overline{\mathcal{M}}_{0,1+a})$

• there is some collection of Hamiltonians $H_w \colon M \to \mathbb{R}$ and almost complex structures $J_w$ on $M$ (for $w \in EG$), with conditions such that a generic choice of these ensures regularity of the moduli space (eg if $M$ is monotone, we can require $H_w \cong 0$ and $J$ to be independent of $w$, whereas if $M$ is weakly monotone we must ensure $J$ is independent of $w$ by $H_w \neq 0$. See [10, Section 4] or [14, Section 4] for more details),

• there is some $\beta \in \mathrm{QH}^*(M) \otimes H_G^*(\mathrm{QC}^*(M)^{\otimes a})$, represented by $B \subset M \times \Pi_1^a M \times EG$ (where $B$ is made as a choice among generic representatives such that regularity of the moduli space holds),

such that the moduli space consists of triples $(m, w, u)$ satisfying

(1) $[m, w] \in A$, where here $[\bullet]$ denotes the equivalence class with respect to the $G$-quotient,

(2) $u \colon \mathrm{Und}(m) \to M$, where here $\mathrm{Und}(m)$ is the underlying nodal genus-zero Riemann surface,

(3) $\bar{\partial}_{J_w} u = X_{H_w}$,

(4) $(u(m_0), u(m_1), \ldots, u(m_n), w) \in B$.

We will define an *operadic $G$-equivariant quantum operation of order $a$* to be some additive homomorphism

$$\mathrm{QH}^*(M; \mathbb{F}_p) \to H_G^*(\mathrm{QC}^*(M)^{\otimes a}; \mathbb{F}_p) \to \mathrm{QH}^*(M; \mathbb{F}_p) \otimes H^*(BG; \mathbb{F}_p),$$

such that the first map is induced by the map $x \mapsto x^{\otimes a}$ and the second map is defined by counting the number of points in 0-dimensional moduli spaces that are $G$-equivariantly parametrised of order $a$, in the usual way. Some examples of such operations include quantum Steenrod operations ([10, Section 4], when $G = \mathbb{Z}/p$ and $a = p$), those involved in the quantum Cartan relation ([14, Section 5], when $G = \mathbb{Z}/p$ and $a = 2p$), and those involved in the quantum Adem relation ([14, Section 7], when $G = D_{2p}$ or $G = S_p$ and $a = p^2$).

An immediate consequence of Theorem 1.2 is then the following:

**Corollary 4.3** *Any operadic $\mathbb{Z}/p$-equivariant quantum operation of order $p$ is a sum of quantum Steenrod operations and some other quantum operation defined by counting (nonequivariant) Gromov–Witten invariants.*

# References

[1]  **M S Betz**, *Operad representations in Morse theory and Floer homology*, PhD thesis, Stanford University (1995) Available at `https://www.proquest.com/docview/304217235`

[2]  **B Eckmann**, *Cohomology of groups and transfer*, Ann. of Math. 58 (1953) 481–493  MR

[3]  **P Etingof**, **A Henriques**, **J Kamnitzer**, **E M Rains**, *The cohomology ring of the real locus of the moduli space of stable curves of genus* 0 *with marked points*, Ann. of Math. 171 (2010) 731–777  MR

[4]  **K Fukaya**, *Morse homotopy and its quantization*, from "Geometric topology, I", AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc., Providence, RI (1997) 409–440  MR

[5]  **S Keel**, *Intersection theory of moduli space of stable n-pointed curves of genus zero*, Trans. Amer. Math. Soc. 330 (1992) 545–574  MR

[6]  **D McDuff**, **D Salamon**, *J-holomorphic curves and symplectic topology*, Amer. Math. Soc. Colloq. Publ. 52, Amer. Math. Soc., Providence, RI (2004)  MR

[7]  **D Quillen**, *The spectrum of an equivariant cohomology ring, I*, Ann. of Math. 94 (1971) 549–572  MR

[8]  **P Seidel**, *The equivariant pair-of-pants product in fixed point Floer cohomology*, Geom. Funct. Anal. 25 (2015) 942–1007  MR

[9]  **P Seidel**, *Formal groups and quantum cohomology*, Geom. Topol. 27 (2023) 2937–3060  MR

[10]  **P Seidel**, **N Wilkins**, *Covariant constancy of quantum Steenrod operations*, J. Fixed Point Theory Appl. 24 (2022) art. id. 52  MR

[11]  **J-P Serre**, *Homologie singulière des espaces fibrés: applications*, Ann. of Math. 54 (1951) 425–505  MR

[12]  **N E Steenrod**, *Products of cocycles and extensions of mappings*, Ann. of Math. 48 (1947) 290–320  MR

[13]  **G W Whitehead**, *Elements of homotopy theory*, Graduate Texts in Math. 61, Springer (1978)  MR

[14] **N Wilkins**, *A construction of the quantum Steenrod squares and their algebraic relations*, Geom. Topol. 24 (2020) 885–970 MR

*Department of Mathematics, MIT*
*Cambridge, MA, United States*
*Max Planck Institute for Mathematics*
*Bonn, Germany*

dain0327@mit.edu, nicholasmichaelwilkins@outlook.com

# Geometry of free extensions of free groups via automorphisms with fixed points on the complex of free factors

PRITAM GHOSH

FUNDA GÜLTEPE

We give conditions of an extension of a free group to be hyperbolic and relatively hyperbolic using the dynamics of the action of $\mathrm{Out}(\mathbb{F})$ on the complex of free factors combined with weak attraction theory. We work with subgroups of exponentially growing outer automorphisms and instead of using a standard ping-pong argument with loxodromics, we allow fixed points for the action and investigate the geometry of the extension group when the fixed points of the automorphisms on the complex of free factors are sufficiently far apart.

20E06, 20F28, 20F67, 57M07

## 1 Introduction

Let $\mathbb{F}$ be a free group of finite rank $\geq 3$. Similar to that of mapping classes for a surface, there is a dichotomy for elements of the group of outer automorphisms $\mathrm{Out}(\mathbb{F})$ of the free group $\mathbb{F}$, in terms of their *growth*. An element $\phi \in \mathrm{Out}(\mathbb{F})$ is called *exponentially growing* if for some conjugacy class $[w]$ of an element $w \in \mathbb{F}$, the word length of $\phi^i([w])$ grows exponentially with $i$ for any fixed generating set of $\mathbb{F}$. We will call a subgroup of $\mathrm{Out}(\mathbb{F})$ exponentially growing if all of its elements are. For an exponentially growing subgroup $\mathcal{Q} < \mathrm{Out}(\mathbb{F})$, we are interested in understanding the geometry of the extension $E_{\mathcal{Q}}$ given by the short exact sequence

$$1 \to \mathbb{F} \to E_{\mathcal{Q}} \to \mathcal{Q} \to 1$$

that is induced from the sequence

$$1 \to \mathbb{F} \to \mathrm{Aut}(\mathbb{F}) \to \mathrm{Out}(\mathbb{F}) \to 1.$$

The extension group $E_{\mathcal{Q}}$ is the pullback of $\mathcal{Q}$ to $\mathrm{Aut}(\mathbb{F})$, and hence a subgroup of $\mathrm{Aut}(\mathbb{F})$.

When $\mathcal{Q}$ is free group of rank $\geq 2$, which is our main focus for this paper, we say that $E_{\mathcal{Q}}$ is a free-by-free group and it is known that $E_{\mathcal{Q}} \cong \mathbb{F} \rtimes \widehat{\mathcal{Q}}$, where $\widehat{\mathcal{Q}}$ is some (any) lift of $\mathcal{Q}$ to $\mathrm{Aut}(\mathbb{F})$. In this paper we give necessary and sufficient conditions for $E_{\mathcal{Q}}$ to be hyperbolic using the dynamical information of the generators of $\mathcal{Q} < \mathrm{Out}(\mathbb{F})$ obtained from train-track maps of free group automorphisms, as well as via their actions on the complex of free factors $\mathcal{FF}$.

An outer automorphism $\phi$ is *atoroidal* (or hyperbolic) if no power of $\phi$ fixes a nontrivial conjugacy class in $\mathbb{F}$. Atoroidal automorphisms are a special class of exponentially growing elements of $\mathrm{Out}(\mathbb{F})$. Using train-track theory, Bestvina–Feighn–Handel [6] developed dynamical invariants for exponentially growing outer automorphisms of $\mathbb{F}$, called *attracting and repelling laminations*. Namely, associated to each exponentially growing outer automorphism $\phi$ (respectively, $\phi^{-1}$), we have finitely many invariant sets of biinfinite *lines* in the (compactified Cayley graph of) Gromov hyperbolic space $\mathbb{F} \cup \partial\mathbb{F}$. These sets of lines are called *attracting laminations* (respectively, *repelling laminations*) and denoted by $\mathcal{L}^+(\phi)$ (respectively, $\mathcal{L}^-(\phi)$). Atoroidal outer automorphisms are characterized by the property that for every conjugacy class $[w]$, the sequence $\phi^i([w])$ converges to some element of $\mathcal{L}^+(\phi)$ as $i \to \infty$ [18, Lemma 3.1]. Ghosh showed that these sets were central to describing the Cannon–Thurston laminations for the extension group $\mathbb{F} \rtimes \langle\phi\rangle$ [18, Lemma 4.4]. We use $\mathcal{L}^\pm(\phi)$ to denote the union of the two lamination sets $\mathcal{L}^+(\phi)$ and $\mathcal{L}^-(\phi)$.

We show that $\mathcal{L}^\pm(\phi)$ is instrumental in characterizing hyperbolicity of the extension group $E_Q$, when $Q$ is a free group of rank $\geq 2$. Our Theorem 1.1 shows that in this case the extension group is hyperbolic, provided we enforce the condition that the attracting and repelling laminations associated to the generators of $Q$ to be pairwise disjoint. We also require throughout the paper that our outer automorphisms are *rotationless* (see Section 2.6), a condition that ensures that we will be able to use special relative train-track maps, which can be achieved by taking a uniform power of automorphisms [16].

**Theorem 1.1** *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing outer automorphisms such that no pair of automorphisms have a common power. Then the following are equivalent:*

(1) *Each $\phi_i$ is an atoroidal outer automorphism and $\mathcal{L}^\pm(\phi_i) \cap \mathcal{L}^\pm(\phi_j) = \varnothing$ for all $i \neq j$.*

(2) *There exists $M > 0$ such that for all $m_i \geq M$, $Q = \langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group and the extension group $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic (where $\widehat{Q}$ is any lift of $Q$).*

Theorem 1.1 is comparable to Theorem 1.3 of Farb and Mosher [15] and the main Theorem of [22] where hyperbolicity of the extension group $E_Q = \pi_1(S) \rtimes Q$ is characterized by the *convex cocompactness* of a free subgroup $Q$ of the mapping class group $\mathrm{MCG}(S)$ of a closed surface $S$.

We say that $\phi_i, \phi_j$, $i \neq j$ are *independent* if generic leaves of elements of $\mathcal{L}^\pm(\phi_i)$ and $\mathcal{L}^\pm(\phi_j)$ are not asymptotic to each other. Hence, Theorem 1.1 can be written as;

**Theorem** *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing outer automorphisms such that no pair of automorphisms have a common power. Then, $\phi_1, \ldots, \phi_k$ is a purely atoroidal (all infinite-order elements are atoroidal), pairwise independent collection if and only if the extension group $\mathbb{F} \rtimes \widehat{Q}$ for a free group $Q = \langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ generated by high enough $m_i$'s is hyperbolic.*

The interest in the geometry of extensions of groups started with Thurston [30] where the $\mathbb{Z}$-extension of $\pi_1(S)$ of a closed surface $S$ is hyperbolic if and only if the cyclic group is generated by a pseudo-Anosov element of the mapping class group. Bestvina–Feighn [2] and Brinkmann [8] have shown that when $Q$ is an infinite cyclic subgroup of $\mathrm{Out}(\mathbb{F})$, $E_Q$ is hyperbolic if and only if $\phi$ is atoroidal.

We do not need $Q$ to be free to prove that the collection is independent and purely atoroidal.

## 1.1   Fixed points on $\mathcal{FF}$ and hyperbolicity of the extension

Free factor complex $\mathcal{FF}$ of the free group $\mathbb{F}$ is a simplicial complex whose vertices are conjugacy classes of nontrivial free factors of $\mathbb{F}$. Two vertices corresponding to free factors are connected by an edge if one includes the other as subgroup, up to conjugation. $\mathcal{FF}$ is Gromov-hyperbolic [3; 24] and in many other ways shows resemblance to the curve complex. Hence the action of $\mathrm{Out}(\mathbb{F})$ on $\mathcal{FF}$ is reminiscent of the action of the mapping class group on the curve complex.

An element of $\mathrm{Out}(\mathbb{F})$ is *fully irreducible* if and only if it has no periodic orbits on $\mathcal{FF}$. It is well known that fully irreducible elements of $\mathrm{Out}(\mathbb{F})$ act loxodromically on $\mathcal{FF}$ [3] and this was exploited by Dowdall–Taylor [13] to produce hyperbolic extensions of free groups. Their theorem can be compared to those of Kent–Leininger [25], and Hamenstädt [22] characterizing convex cocompactness, hence the hyperbolicity of the extension, with quasi-isometric embedding of the group in the curve complex. As such the automorphisms of [13] are fully irreducible.

We take a different approach in this paper and produce hyperbolic extensions of groups whose elements are not necessarily fully irreducible, hence we allow automorphisms to have fixed points on $\mathcal{FF}$. The following result shows that the only condition one needs for hyperbolicity is that they do not have a common fixed point on $\mathcal{FF}$.

**Theorem 1.2**   *Let $\phi_1, \ldots, \phi_k$ be a collection of atoroidal elements which do not have a common power. If no pair $\phi_i, \phi_j$, $i \neq j$, fixes a common vertex in the free factor complex $\mathcal{FF}$, then*

   (1)   *there exists $M > 0$ such that for all $m_i \geq M$, $Q = \langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group,*

   (2)   *the extension group $\mathbb{F} \rtimes \widehat{Q}$ is a hyperbolic group (where $\widehat{Q}$ is any lift of $Q$).*

Our Theorem 1.2 hence responds to the question of what happens when a pair $\phi, \psi$ in a collection of automorphisms do have common fixed points for their action on $\mathcal{FF}$.

## 1.2   Sufficiently different automorphisms, nonattracting sink and the geometry of the extension group

One of the motivations for this work is to establish a connection between the dynamics of the action of $\mathrm{Out}(\mathbb{F})$ on the free factor complex and the dynamical data that we have from the train-track theory. We believe that this is a theme that has barely been explored and a lot can be learned if we are able exploit this connection.

We call a collection of automorphisms *sufficiently different* if no pair of automorphisms have a common power and for each pair of outer automorphisms in the collection, the distance between the fixed vertices in $\mathcal{FF}$ is at least 2 (if such vertices exist) (see Section 3.1). Lemma 3.2 gives a simple criterion for constructing free subgroups of $\mathrm{Out}(\mathbb{F})$ using sufficiently different elements.

Given a short exact sequence $1 \to \mathbb{F} \to E \to Q \to 1$ of finitely generated groups, we say that $E$ has a *cusp-preserving* relatively hyperbolic structure if there exists a collection of finitely generated subgroups $\{H_i\}$ of $\mathbb{F}$ such that $Q$ preserves conjugacy class of each $H_i$ and $E$ is (strongly) hyperbolic relative to the collection $\{N_E(H_i)\}$ of normalizers of $H_i$ in $E$. Our main aim in Section 3.2 is to address the question of when a free subgroup $Q$ of $\mathrm{Out}(\mathbb{F})$ can yield relatively hyperbolic extension $E$ with the cusp-preserving property.

In the proposition below we conclude that being sufficiently different is an obstruction to cusp-preserving relative hyperbolicity of the extension group, except in one particular case. Equivalently, it is an obstruction to having an *admissible subgroup system* [20, Section 2.8] for the collection of $\phi_i$'s. An admissible subgroup system is a malnormal collection of subgroups with some special properties, and the *nonattracting sink* $\mathcal{K}_\phi^*$ of an automorphism $\phi \in \mathrm{Out}(\mathbb{F})$, which is developed in [20] as an example of an admissible subgroup system, carries all conjugacy classes which do not grow exponentially under iteration by $\phi$ (see Lemma 3.6). A nonattracting sink can be computed explicitly using relative train-track maps.

For a collection $\phi_1, \ldots, \phi_k$ of pairwise sufficiently different exponentially growing outer automorphisms, let $\mathcal{K}_i^*$ be the nonattracting sink of $\phi_i$.

**Proposition 3.10** *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing, sufficiently different outer automorphisms. If all the nonattracting sinks are nonempty then exactly one of following are true:*

(1) *$\phi_1, \ldots, \phi_k$ is a collection of fully irreducible geometric outer automorphisms induced by some pseudo-Anosov homemorphisms of the same compact surface with one boundary component. Moreover, all sufficiently high powers of $\phi_i$'s generate a free group $Q$ such that $E_Q$ has a cusp-preserving relatively hyperbolic structure.*

(2) *$\mathcal{K}_i^* \neq \mathcal{K}_j^*$ for some $i \neq j$ and the extension $E_Q$ of a free group $Q$, which is generated by sufficiently high powers of $\phi_i$'s, cannot have a cusp preserving relatively hyperbolic structure.*

We have the following theorem characterizing the hyperbolicity of $E_Q$ using sinks:

**Theorem 1.3** *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing and sufficiently different outer automorphisms. Let $Q$ be a free group generated by sufficiently high powers of $\phi_i$'s. Then, $\mathbb{F} \rtimes \widehat{Q}$ is a hyperbolic group if and only if $\mathcal{K}_i^* = \varnothing$ for all $i \in \{1, \ldots, k\}$.*

In this theorem $Q$ is not necessarily a *convex cocompact* subgroup of $\mathrm{Out}(\mathbb{F})$ in the sense of [11; 12; 21], since our groups do not quasi-isometrically embed in $\mathcal{FF}$. Hence our theorem cannot be obtained via their work. In fact, as far as our knowledge goes, this theorem cannot be obtained by studying the action of $Q$ on any known hyperbolic simplicial complex with a nice $\mathrm{Out}(\mathbb{F})$ action.

In a similar vein, we obtain the following necessary and sufficient condition in terms of nonattracting sinks, regarding cusp-preserving relative hyperbolicity of the extension group when $Q$ is generated by sufficiently different collections of outer automorphisms.

**Theorem 5.1**   *Let $\phi_1, \ldots, \phi_k$ be a collection of pairwise sufficiently different and exponentially growing outer automorphisms. Let $\mathcal{K}_i^*$ be the nonattracting sink of $\phi_i$, $i \in \{1, \ldots, k\}$, and assume that $\mathcal{K}_j^* \neq \varnothing$ for some fixed $j$. Then the following are equivalent:*

(1)  *$\mathbb{F} \rtimes \widehat{Q}$ has a cusp preserving relatively hyperbolic structure where $Q$ is a free group generated by sufficiently large exponents of $\phi_i$'s.*

(2)  *$\mathcal{K}_i^* = \mathcal{K}_j^*$ for all $i$.*

To summarize, given a pair of sufficiently different, exponentially growing elements of $\mathrm{Out}(\mathbb{F})$, all sinks empty gives hyperbolic extensions (Theorem 1.3). No sink empty gives relatively hyperbolic extensions with cusp-preserving structure only for geometric fully irreducibles (Proposition 3.10). Some sink nonempty gives relatively hyperbolic extensions with cusp-preserving structure only if all sinks are nonempty and are equal (Theorem 5.1).

## 1.3  Plan of the paper

Section 2 collects all the definitions and the tools we use.

In Section 3 we investigate obstructions to (relative) hyperbolicity using the dynamics of sufficiently different automorphisms on $\mathcal{FF}$. We also discuss the notion of a sink and include proof of Proposition 3.10.

In Section 4 we continue working with a specific type of exponentially growing automorphism; a reducible automorphism that is fully irreducible on a free factor (a *partial* fully irreducible). We investigate the conditions on the small displacement sets on $\mathcal{FF}$ of partial fully irreducible automorphisms that determine the geometry of the extension groups they generate.

Section 5 includes proofs of Theorems 1.1, 1.2 and 5.1. In this section we discuss further the notion of sink and how the sinks of automorphisms determine the geometry of an extension group generated by those automorphisms and prove Theorem 1.3. We work with partial and relatively fully irreducible elements directly to exhibit examples of relatively hyperbolic free group extensions.

Section 6 discusses further problems such as characterization of nonrelative hyperbolicity; in this section we give an example which could help characterize nonrelative hyperbolicity of a free group extension.

# 2 Preliminaries

## 2.1 Marked graphs, circuits and path

A *marked graph* is a graph $G$ which is a core graph (a graph with no valence-1 vertices) that is equipped with a homotopy equivalence to the rose $m: G \to R_n$ (where $n = \mathrm{rank}(\mathbb{F})$). Thus the fundamental group of $G$ can be identified with $\mathbb{F}$ (up to inner automorphisms). A *circuit* in a marked graph is a locally injective and continuous map of $S^1$ into $G$. The set of circuits in $G$ can be identified with the set of conjugacy classes in $\mathbb{F}$.

A *path* is an immersion of the interval $[0, 1]$ into $G$ with endpoints at vertices of $G$. Every path can be written as a finite concatenation of edges of $G$, so that there is no backtracking.

A *line* $\ell$ is a bi-infinite concatenation $\ell = \ldots E_{i-1} E_i E_{i+1} \ldots, i \in \mathbb{Z}$, of edges of $G$ without backtracking, ie so that $E_{j-1}^{-1} \neq E_j \neq E_{j+1}^{-1}$ (where $E_s^{-1}$ is $E_s$ traveled in opposite direction). A *ray* $\gamma$ is "one-sided" infinite concatenation of edges $\gamma = E_i E_{i+1} \ldots, i \in \mathbb{Z}_{\geq 0}$, in $G$ without backtracking. Two lines are said to be asymptotic if they have a common subray. Any continuous map $f$ from $S^1$ or $[0, 1]$ to $G$ can be *tightened* to a circuit or path, in other words is freely homotopic to a locally injective and continuous map. Tightened image of a path $\alpha$ under $f$ will be denoted by $f_{\#}(\alpha)$ and we will not distinguish between circuits or paths that differ by a homeomorphism of their respective domains.

## 2.2 Topological representative, EG strata, NEG strata

A *filtration* of a marked graph $G$ is a strictly increasing sequence $G_0 \subset G_1 \subset \cdots \subset G_k = G$ of subgraphs $G_r$ with no isolated vertices. The filtration is $f$-*invariant* if $f(G_r) \subset G_r$ for all $r$.

The *stratum of height $r$* is a subgraph $H_r = G_r \smallsetminus G_{r-1}$ together with the endpoints of edges. The minimum $r$ such that a subset of $G$ is contained in $G_r$ is called the *height* of the subset.

For an $f$-invariant filtration, the square matrix $M_r$ whose $j^{\text{th}}$ column records the number of times the image of an edge $e_j$ under $f$ intersects the other edges in $H_r$ is called the *transition matrix* of the stratum $H_r$. $M_r$ is said to be irreducible if for each $i, j$, the $(i, j)$ entry of some power of $M_r$ is nonzero. In this case we say that the associated stratum $H_r$ is also *irreducible*. When $H_r$ is irreducible, the Perron–Frobenius theorem states that the matrix $M_r$ has a unique eigenvalue $\lambda \geq 1$, called the *Perron–Frobenius eigenvalue*, for which some associated eigenvector has positive entries. If $\lambda = 1$ then $H_r$ is a *nonexponentially growing (NEG) stratum* whereas if $\lambda > 1$ we say that $H_r$ is an *exponentially growing (EG) stratum*.

An automorphism $\phi \in \mathrm{Out}(\mathbb{F})$ can be represented by a homotopy equivalence $f: G \to G$ that takes vertices to vertices and edges to edge-paths of a marked graph $G$ with marking $\rho: R_n \to G$, called a *topological representative*. Topological representative $f$ preserves the marking of $G$, in other words $\overline{\rho} \circ f \circ \rho: R_n \to R_n$ represents $R_n$. A nontrivial path $\alpha$ in G is a *periodic Nielsen path* if $f_{\#}^k(\alpha) = \alpha$ for some $k$, where the smallest such $k$ is called the *period*. $\alpha$ is a Nielsen path if $k = 1$. A periodic Nielsen path is *indivisible* (iNP) if it cannot be written as a concatenation of nontrivial periodic Nielsen paths.

Given a topological representative $f : G \to G$ let $Tf$ be such that $Tf(E)$ is the first edge in the edge path associated to $f(E)$. A *turn* is a pair of edges $\{E_i, E_j\}$ of edges and we let $Tf(E_i, E_j) = (Tf(E_i), Tf(E_j))$ making $Tf$ a map that takes turns to turns. We say that a nondegenerate (i.e, $i \neq j$) turn is *illegal* if for some iterate of $Tf$ the turn becomes degenerate; otherwise the turn is legal. A path is said to be *legal path* if it contains only legal turns and it is $r$-legal if it is of height $r$ and all its illegal turns are in $G_{r-1}$.

**Relative train-track map**    Given $\phi \in \mathrm{Out}(\mathbb{F})$ and a topological representative $f : G \to G$ with a filtration $G_0 \subset G_1 \subset \cdots \subset G_k$ which is preserved by $f$, we say that $f$ is a relative train-track map if the following conditions are satisfied:

(1)    $f$ maps $r$-legal paths to $r$-legal paths.

(2)    If $\gamma$ is a path in $G_{r-1}$ with endpoints in $G_{r-1} \cap H_r$ then $f_\#(\gamma)$ is nontrivial.

(3)    If $E$ is an edge in $H_r$ then $Tf(E)$ is an edge in $H_r$. In particular, every turn consisting of a direction of height $r$ and one of height $< r$ is legal.

By [6, Theorem 5.1.5] every $\phi \in \mathrm{Out}(\mathbb{F})$ has a relative train-track map representative which satisfies some useful conditions in addition to the ones we listed above.

## 2.3   Weak topology

Given a graph $G$, we define an equivalence relation on the set of all paths, rays, lines, and circuits in $G$ by declaring two of paths/rays/lines/circuits to be equivalent if they differ by a homeomorphism of their domains.

Let $\widehat{\mathcal{B}}(G)$ denote the compact space of equivalence classes of circuits and finite paths, rays and lines in a graph $G$, whose endpoints (if any) are vertices of $G$. For each finite path $\gamma$ in $G$, we denote by $\widehat{N}(G, \gamma)$ the set of all paths and circuits in $\widehat{\mathcal{B}}(G)$ which have $\gamma$ as its subpath. The collection of all such sets gives a basis for a topology on $\widehat{\mathcal{B}}(G)$ called *weak topology*. Let $\mathcal{B}(G) \subset \widehat{\mathcal{B}}(G)$ be the compact subspace of all lines in $G$ with the induced topology.

Two distinct points in $\partial \mathbb{F}$ determine a *line*, up to reversing the direction. Let $\widetilde{\mathcal{B}} = \{\partial \mathbb{F} \times \partial \mathbb{F} - \Delta\}/(\mathbb{Z}/2\mathbb{Z})$ be the set of pairs of boundary points of $\mathbb{F}$ where $\Delta$ is the diagonal and $\mathbb{Z}_2$ acts by interchanging factors. We can give the weak topology to $\widetilde{\mathcal{B}}$, induced by the Cantor topology on $\partial \mathbb{F}$.

$\mathbb{F}$ acts on $\widetilde{\mathcal{B}}$ and the quotient space $\mathcal{B} = \widetilde{\mathcal{B}}/\mathbb{F}$ is compact but non-Hausdorff as such the topology is called *weak*. The quotient topology is also called the *weak topology*. For any marked graph $G$, there is a natural identification $\mathcal{B} \approx \mathcal{B}(G)$.

We call the elements of $\mathcal{B}$ *lines* as well. A lift of a line $\gamma \in \mathcal{B}$ is an element $\widetilde{\gamma} \in \widetilde{\mathcal{B}}$ that projects to $\gamma$ under the quotient map and the two elements of $\widetilde{\gamma}$ are called its endpoints.

A line (or a path) $\gamma$ is said to be *weakly attracted* to a line (path) $\beta$ under the action of $\phi \in \mathrm{Out}(\mathbb{F})$, if for some $k$ $\phi^k(\gamma)$ converges to $\beta$ in the weak topology, in other words, if any given finite subpath of $\beta$ is

contained in $\phi^k(\gamma)$ for some $k$. Similarly if we have a homotopy equivalence $f : G \to G$, a line(path) $\gamma$ is said to be *weakly attracted* to a line(path) $\beta$ under the action of $f_\#$ if $f_\#^k(\gamma)$ weakly converges to $\beta$. the *accumulation set* of $\gamma$ is the set of lines $l \in \mathcal{B}(G)$ that are elements of the weak closure of a ray $\gamma$ in $G$. This is the set of lines $l$ such that every finite subpath of $l$ occurs infinitely many times as a subpath $\gamma$. Similarly, the weak accumulation set of some point $\xi \in \partial\mathbb{F}$ is the set of lines in the weak closure of any of the asymptotic rays in its equivalence class.

## 2.4 Free factor systems and malnormal subgroup systems

A finite collection $\mathcal{K} = \{[K_1], [K_2], \dots, [K_s]\}$ of conjugacy classes of nontrivial, finite-rank subgroups $K_s < \mathbb{F}$ is called a *subgroup system*. A subgroup $K < \mathbb{F}$ is *malnormal* if $xK_s x^{-1} \cap K$ is trivial for all $x \in \mathbb{F} - K$. A subgroup system is called *malnormal* if each of its subgroups $[K_s] \in \mathcal{K}$ is malnormal and for all $[K_s], [K_t] \in \mathcal{K}$ if $[K_s] \cap [K_t]$ is nontrivial then $s = t$. Given two malnormal subgroup systems $\mathcal{K}, \mathcal{K}'$ we define a partial ordering $\sqsubset$ on the set of subgroup systems by $\mathcal{K} \sqsubset \mathcal{K}'$ if for each conjugacy class of subgroup $[K] \in \mathcal{K}$ there exists some conjugacy class of subgroup $[K'] \in \mathcal{K}'$ such that $K < K'$.

Given a finite collection $\{K_1, K_2, \dots, K_s\}$ of subgroups of $\mathbb{F}$, we say that this collection determines a *free factorization* of $\mathbb{F}$ if $\mathbb{F} = K_1 * K_2 * \cdots * K_s$. A *free factor system* $\mathcal{F} := \{[F_1], [F_2], \dots, [F_p]\}$ of $\mathbb{F}$ is a finite collection of conjugacy classes of subgroups such that there is a free factorization of $\mathbb{F}$ of the form $\mathbb{F} = F_1 * F_2 * \cdots * F_p * B$, where $B$ is some (possibly trivial) finite-rank subgroup of $\mathbb{F}$. Every free factor system is a malnormal subgroup system.

A malnormal subgroup system $\mathcal{K}$ *carries* a conjugacy class $[c] \in \mathbb{F}$ if there exists some $[K] \in \mathcal{K}$ such that $c \in K$. We say that $\mathcal{K}$ carries a line $\gamma$ if one of the following equivalent conditions hold:

(1)   $\gamma$ is the weak limit of a sequence of conjugacy classes carried by $\mathcal{K}$.

(2)   There exists some $[K] \in \mathcal{K}$ and a lift $\widetilde{\gamma}$ of $\gamma$ so that the endpoints of $\widetilde{\gamma}$ are in $\partial K$.

For any marked graph $G$ and any subgraph $H \subset G$, the fundamental groups of the noncontractible components of $H$ form a free factor system, denoted by $[\pi_1(H)]$. Every free factor system $\mathcal{F}$ can be realized as $[\pi_1(H)]$ for some nontrivial core subgraph H of some marked graph $G$. An equivalent way of saying that a line or circuit $\gamma$ is carried by $\mathcal{F}$ is that for any marked graph $G$ and a subgraph $H \subset G$ with $[\pi_1(H)] = \mathcal{F}$, the realization of $\gamma$ in $G$ is contained in $H$.

We have the following fact:

**Lemma 2.1** [23, Fact 1.8]   *Given a subgroup system $\mathcal{K}$ the set of lines/circuits carried by $\mathcal{K}$ is a closed set in the weak topology.*

As a consequence, given a malnormal subgroup system $\mathcal{K}$ and a sequence of lines/circuits $\{\gamma_n\}$, if $\mathcal{K}$ carries every weak limit of every subsequence of $\{\gamma_n\}$, then $\gamma_n$ is carried by $\mathcal{K}$ for all sufficiently large $n$ [23, Lemma 1.11].

Given two free factor systems, an extension of the notion of "intersection" of subgroups, extended to free factor systems is called the "meet" of free factors:

**Lemma 2.2** [6, Section 2.6] *Every collection $\{\mathcal{F}_i\}$ of free factor systems has a well-defined meet $\bigwedge \{\mathcal{F}_i\}$, which is the unique maximal free factor system $\mathcal{F}$ such that $\mathcal{F} \sqsubset \mathcal{F}_i$ for all $i$. Moreover, for any free factor $F < \mathbb{F}$ we have $[F] \in \bigwedge \{\mathcal{F}_i\}$ if and only if there exists an indexed collection of subgroups $\{A_i\}_{i \in I}$ such that $[A_i] \in \mathcal{F}_i$ for each $i$ and $F = \bigcap_{i \in I} A_i$.*

The *free factor support* $\mathcal{F}_{\text{supp}}(B)$ of a set of lines $B$ in $\mathcal{B}$ is defined as the meet of all free factor systems that carry $B$ [6]. If $B$ is a single line then $\mathcal{F}_{\text{supp}}(B)$ is single free factor. We say that a set of lines $B$ is *filling* if $\mathcal{F}_{\text{supp}}(B) = \mathbb{F}$.

## 2.5 Attracting laminations and nonattracting subgroup systems

For any marked graph $G$, there is a natural identification $\mathcal{B} \approx \mathcal{B}(G)$ which induces a bijection between the closed subsets of $\mathcal{B}(\mathcal{G})$ of $\mathcal{B}$. A *lamination* $\Lambda$ is a closed subset of any of these two sets. Given a lamination $\Lambda \subset \mathcal{B}$ we look at the corresponding lamination in $\mathcal{B}(G)$ as the realization of $\Lambda$ in $G$. An element $\lambda \in \Lambda$ is called a *leaf* of the lamination.

A lamination $\Lambda$ is called an *attracting lamination* for $\phi$ if it is the weak closure of a line $\ell$ (called the *generic leaf of* $\Lambda$) satisfying:

- $\ell$ is birecurrent leaf of $\Lambda$: every finite subpath of $\ell$ occurs infinitely many times as subpath in both directions.
- There is a neighborhood $V^+$ such that every line in $V^+$ is weakly attracted to $\ell$ in the weak topology. $V^+$ is called an *attracting neighborhood* of $\ell$ (see [6, Definition 3.1.1]).
- no lift $\widetilde{\ell} \in \mathcal{B}$ of $\ell$ is the axis of a generator of a rank-1 free factor of $\mathbb{F}$.

Attracting neighborhoods of $\Lambda$ are defined by choosing sufficiently long segments of generic leaves of $\Lambda$. If $V^+$ is an attracting neighborhood, then $\phi(V^+) \subset V^+$.

Associated to each $\phi \in \text{Out}(\mathbb{F})$ is a finite set $\mathcal{L}^+(\phi)$ of laminations, called the set of *attracting laminations* of $\phi$ [6]. Similarly we define the set of attracting laminations (or the set of *repelling laminations*) $\mathcal{L}^-(\phi)$ of $\phi^{-1}$ (or of $\phi$). Both $\mathcal{L}^+(\phi)$ and $\mathcal{L}^-(\phi)$ are $\phi$-invariant [6, Lemmas 3.1.13 and 3.1.6]. Given a relative train-track map $f: G \to G$ representing $\phi$, then there is a bijection between the set $\mathcal{L}^+(\phi)$ and the set of exponentially growing strata in $G$ which is determined by the height of a generic leaf of $\Lambda^+ \in \mathcal{L}^+(\phi)$ [6, Section 3].

Free factor support of an element $\Lambda$ of $\mathcal{L}^+(\phi)$ or $\mathcal{L}^-(\phi)$ is the smallest free factor system that carries each leaf of $\Lambda$, denoted by $\mathcal{F}_{\text{supp}}(\Lambda)$ [6, Section 3.2]. Free factor support of $\Lambda^+ \in \mathcal{L}^+(\phi)$ is equal to free factor support of a generic leaf of $\Lambda^+$ [6, Corollaries 2.6.5 and 3.1.11]. If $\Lambda_1, \Lambda_2 \in \mathcal{L}^+(\phi)$ then $\Lambda_1 = \Lambda_2 \Leftrightarrow \mathcal{F}_{\text{supp}}(\Lambda_1) = \mathcal{F}_{\text{supp}}(\Lambda_2)$ [23, Fact 1.14].

An element of $\mathcal{L}^+(\phi)$ is dual to an element of $\mathcal{L}^-(\phi)$ if these two elements have the same free factor support. This imposes a bijection between these two sets [6, Lemma 3.2.4].

A line/circuit $\gamma$ is said to be *weakly attracted* to $\Lambda_1 \in \mathcal{L}^+(\phi)$ if $\gamma$ is weakly attracted to some (hence every) generic leaf of $\Lambda$ under action of $\phi$. No leaf of any element of $\mathcal{L}^-(\phi)$ is ever attracted to any leaf of any element of $\mathcal{L}^+(\phi)$. No Nielsen path is weakly attracted to either an element of $\mathcal{L}^+(\phi)$ or an element of $\mathcal{L}^-(\phi)$.

An attracting lamination $\Lambda^+$ and a repelling lamination $\Lambda^-$ of $\phi$ cannot have leaves which are asymptotic, for otherwise it would violate the fact that no leaf of $\Lambda^-$ is weakly attracted to $\Lambda^+$. This allows us to choose sufficiently long subpaths of generic leaves of $\Lambda^+$, $\Lambda^-$ and construct attracting and repelling neighborhoods $V^+$, $V^-$ so that $V^+ \cap V^- = \varnothing$.

An element of $\Lambda \in \mathcal{L}^+(\phi)$ is said to be *topmost* if there does not exists any $\Lambda_j \in \mathcal{L}^+(\phi)$ such that $\Lambda \subset \Lambda_j$.

Elements of $\mathcal{L}^+(\phi)$ can be divided into two distinct classes. If there exists a finite, $\phi$-invariant collection of distinct, nontrivial conjugacy classes $\mathcal{C} = \{[c_1], [c_2], \ldots, [c_s]\}$ such that $\mathcal{F}_{\mathrm{supp}}(\mathcal{C}) = \mathcal{F}_{\mathrm{supp}}(\Lambda^+)$, then $\Lambda^+ \in \mathcal{L}^+(\phi)$ is said to be *geometric* [23, Definition 2.19]. It is said to be *nongeometric* otherwise.

The *nonattracting subgroup system* of an attracting lamination was first introduced by Bestvina–Feighn–Handel in [6] and later explored in great details by Handel–Mosher in [23, Part III, pages 192–202]. Nonattracting subgroup system records the information about the lines and circuits which are not attracted to the lamination. We will list some of the properties which are central to our proofs.

**Lemma 2.3** [23, Lemmas 1.5 and 1.6]  *Let $\phi \in \mathrm{Out}(\mathbb{F})$ and $\Lambda^+ \in \mathcal{L}^+(\phi)$ be an attracting lamination such that $\phi(\Lambda^+) = \Lambda^+$. Then there exists a subgroup system $\mathcal{A}_{\mathrm{na}}(\Lambda^+)$ such that:*

(1) *$\mathcal{A}_{\mathrm{na}}(\Lambda^+)$ is a malnormal subgroup system and the set of lines carried by $\mathcal{A}_{\mathrm{na}}(\Lambda^+)$ is closed in the weak topology.*

(2) *A conjugacy class $[c]$ is not attracted to $\Lambda^+$ if and only if it is carried by $\mathcal{A}_{\mathrm{na}}(\Lambda^+)$.*

(3) *If each conjugacy class of a finite-rank subgroup $B < \mathbb{F}$ is not weakly attracted to $\Lambda_\phi^+$, then there exists some $A < \mathbb{F}$ such that $B < A$ and $[A] \in \mathcal{A}_{\mathrm{na}}(\Lambda_\phi^+)$.*

(4) *If $\Lambda^-$ and $\Lambda^+$ are dual to each other, we have $\mathcal{A}_{\mathrm{na}}(\Lambda^+) = \mathcal{A}_{\mathrm{na}}(\Lambda^-)$.*

(5) *$\mathcal{A}_{\mathrm{na}}(\Lambda^+)$ is a free factor system if and only if $\Lambda^+$ is not geometric.*

(6) *If $\{\gamma_n\}_{n \in \mathbb{N}}$ is a sequence of lines such that every weak limit of every subsequence of $\{\gamma_n\}$ is carried by $\mathcal{A}_{\mathrm{na}}(\Lambda^+)$ then for all sufficiently large $n$, $\{\gamma_n\}$ is carried by $\mathcal{A}_{\mathrm{na}}(\Lambda^+)$.*

A pseudo-Anosov on a surface with nonempty boundary induces a fully irreducible automorphism called *geometric* fully irreducible. Geometricity was relativized by Bestvina–Feighn–Handel [6] by extending it so that it is a property of an EG stratum of a relative train-track map. The definition of geometricity is expressed in terms of the existence of what we call a "geometric model" (see [23] for details of constructions of geometric models.)

When $\Lambda^+$ is geometric, the geometric model gives us the structure of the nonattracting subgroup system [23, Definition 2.1, Lemma 2.5]. In particular, there is a unique closed indivisible Nielsen path $\rho_r$ of height $r$ [23, Fact 1.42].

The facts listed above show that when $\Lambda^+$ is geometric, $\mathcal{A}_{\mathrm{na}}(\Lambda^+) = \mathcal{F} \cup \{[F_m]\}$, where $\mathcal{F}$ is some free-factor system such that $[\pi_1 G_{r-1}] \sqsubset \mathcal{F}$. Moreover, $\phi$ restricted to $F_m$ has polynomial growth and if $[c]$ is the conjugacy class determined by $\rho_r$, then $[c]$ is carried by $[F_m]$. We shall refer to the component $[F_m]$ as a *geometric component* of $\mathcal{A}_{\mathrm{na}}(\Lambda^+)$. In the case $H_r$ is a top stratum, $\mathcal{A}_{\mathrm{na}}(\Lambda^+) = \mathcal{F} \cup \{[\langle c \rangle]\}$.

## 2.6  Completely split improved relative train-track (CT) maps

A *splitting* of a line, a path or a circuit $\alpha$ is a concatenation $\ldots \alpha_0 \alpha_1 \ldots \alpha_k \ldots$ of subpaths of $\alpha$ in $G$ such that for all $i \geq 1$, $f_\#^i(\alpha) = \ldots f_\#^i(\alpha_0) f_\#^i(\alpha_1) \ldots f_\#^i(\alpha_k) \ldots$, for a relative train-track map $f : G \to G$. The subpath $\alpha_i$'s in a splitting are called *terms* or *components* of the splitting of $\alpha$. The notation $\alpha \cdot \beta$ will denote a splitting and $\alpha\beta$ will denote a concatenation of nontrivial paths $\alpha, \beta$.

A splitting $\alpha_1 \cdot \alpha_2 \cdots \alpha_k$ of a nontrivial path or circuit $\alpha$ is a *complete splitting* if each component $\alpha_i$ is either a single edge in an irreducible stratum, an indivisible Nielsen path, an exceptional path (see [16, Definition 4.1]) or a *taken* connecting path in a zero stratum (see [16, Definition 4.4]).

*Completely split improved relative train-track (CT) maps* are topological representatives with particularly nice properties. In particular, for each edge $E$ in each irreducible stratum, the path $f_\#(E)$ is completely split. Moreover, for each taken connecting path $\alpha$ in each zero stratum, the path $f_\#(\alpha)$ is completely split. CT maps are guaranteed to exist for *rotationless* (see [16, Definition 3.13]) outer automorphisms, as has been shown in the following.

**Lemma 2.4**  [16, Theorem 4.28]  *For each rotationless $\phi \in \mathrm{Out}(\mathbb{F})$ and each increasing sequence $\mathcal{F}$ of $\phi$-invariant free factor systems, there exists a CT map $f : G \to G$ that is a topological representative for $\phi$ and $f$ realizes $\mathcal{F}$.*

Feighn–Handel [16] showed that there exists some $k > 0$ such that, given any $\phi \in \mathrm{Out}(\mathbb{F})$, $\phi^k$ is rotationless. So given any outer automorphism $\phi$, some finite power of $\phi$ has a completely split improved relative train-track representative. In the body of our work, we will use CT's defined in [16]. In particular, we will use the properties given in [23, Definition 1.29] assume all exponentially growing outer automorphisms are rotationless.

## 2.7  Critical constant

Let $\phi \in \mathrm{Out}(\mathbb{F})$ be exponentially growing and $f : G \to G$ be a CT map representing $\phi$. Below we will adapt a *bounded cancellation* notion coming from Bestvina–Feighn–Handel's bounded cancellation lemma for train-track representatives of automorphisms of $\mathbb{F}$ [5]; which is inspired by Cooper's same

named lemma [9]. To summarize, if we have a path in $G$ which has some $r$-legal "central" subsegment of length greater than the critical constant given below, then this segment is protected by the bounded cancellation lemma and the length of this segment grows exponentially under iteration.

**Definition 2.5** Let $H_r$ be an exponentially growing stratum with associated Perron–Frobenius eigenvalue $\lambda_r$ and let $\mathrm{BCC}(f)$ denote the bounded cancellation constant for $f$. Then,

$$\frac{2\,\mathrm{BCC}(f)}{\lambda_r - 1}$$

is called the *critical constant* for $H_r$.

It is easy to see that for every number $C > 0$ that exceeds the critical constant, there is some $1 \geq \mu > 0$ such that if $\alpha\beta\gamma$ is a concatenation of $r$-legal paths where $\beta$ is some $r$-legal segment of length $\geq C$, then the $r$-legal leaf segment of $f_\#^k(\alpha\beta\gamma)$ corresponding to $\beta$ has length at least $\mu\lambda^k |\beta|_{H_r}$ (see [5, page 219]). Here $|\beta|_{H_r}$ is the length of an edge-path $\beta \subset G$ is the number of the edges of $\beta$ that remain only in $H_r$.

For the rest of this paper, let $C$ be a number larger than the maximum of all critical constants corresponding to EG strata of $f : G \to G$.

## 2.8 Admissible subgroup systems and relative hyperbolicity

Given a group $G$ and a collection $\{K_\alpha\}$ of subgroups $K_\alpha < G$, we obtain the *coned-off Cayley graph of $G$* or the *electrocuted $G$* relative to the collection $\{K_\alpha\}$ by assigning a vertex $v_\alpha$ for each left coset of $K_\alpha$ on the Cayley graph of $G$ such that each point of a left coset of $K_\alpha$ is joined to (or coned-off at) $v_\alpha$ by an edge of length $\frac{1}{2}$. The resulting metric space is denoted by $(\widehat{G}, |\cdot|_{\mathrm{el}})$.

A group $G$ is said to be (weakly) relatively hyperbolic relative to the collection of subgroups $\{K_\alpha\}$ if $\widehat{G}$ is a $\delta$-hyperbolic metric space. $G$ is said to be strongly hyperbolic relative to the collection $\{K_\alpha\}$ if the coned-off space $\widehat{G}$ is weakly hyperbolic relative to $\{K_\alpha\}$ and it satisfies a certain *bounded coset penetration* property (see [14]). In this paper, when we say "relative hyperbolicity" we always mean "strong relative hyperbolicity".

**Definition 2.6** Given $\phi \in \mathrm{Out}(\mathbb{F})$, let $\mathcal{K} = \{[K_1], [K_2], \ldots, [K_p]\}$ be a subgroup system and $\mathcal{L}_\mathcal{K}^+(\phi)$ (respectively, $\mathcal{L}_\mathcal{K}^-(\phi)$) denote the collection of attracting (respectively, repelling) laminations of $\phi$ whose generic leaves are not carried by $\mathcal{K}$. Assume:

(1) $\mathcal{L}_\mathcal{K}^+(\phi), \mathcal{L}_\mathcal{K}^-(\phi)$ are both nonempty.

(2) $\mathcal{K}$ is a malnormal subgroup system.

(3) $\phi(K_s) = K_s$ for each $s \in \{1, \ldots p\}$.

(4) Let $V^+$ denote the union of attracting neighborhoods of elements of $\mathcal{L}_\mathcal{K}^+(\phi)$ defined by generic leaf segments of length $\geq 2C$. Define $V^-$ similarly for $\mathcal{L}_\mathcal{K}^-(\phi)$. By increasing $C$ if necessary,

$$V^+ \cap V^- = \varnothing.$$

(5) Every conjugacy class which is not carried by $\mathcal{K}$ is weakly attracted to some element of $\mathcal{L}_\mathcal{K}^+(\phi)$.

We will call subgroup systems satisfying the properties above an *admissible subgroup system for $\phi$* (where $\phi \in \mathrm{Out}(\mathbb{F})$ is exponentially growing). If we are given some finitely generated group $Q = \langle \phi_1, \phi_2, \ldots, \phi_k \rangle$ we say that $\mathcal{K}$ is an *admissible subgroup system for $Q$* if $\mathcal{K}$ is an admissible subgroup system for each $\phi_i$. We will simply write "admissible subgroup system" when the context is clear.

In our previous work, we proved:

**Theorem 2.7** [20]   *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing outer automorphisms of $\mathbb{F}$ such that $\phi_i, \phi_j$ do not have a common power whenever $i \neq j$. Then the following are equivalent:*

(1) *There exists a subgroup system $\mathcal{K} = \{[K_1], \ldots, [K_p]\}$ which is admissible for each $\phi_i$ and for which $\mathcal{L}_{\mathcal{K}}^{\pm}(\phi_i) \cap \mathcal{L}_{\mathcal{K}}^{\pm}(\phi_j) = \varnothing$ whenever $i \neq j$.*

(2) *For every free group $Q$ generated by sufficiently high powers of $\phi_i$'s, $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic relative to a collection $\{K_s \rtimes \widehat{Q}_s\}_{s=1}^{p}$ where $\widehat{Q}_s$ is a lift of $Q$ that fixes $K_s$.*

# 3   Detecting obstructions to (relative) hyperbolicity on the complex of free factors

**Why high powers?**   Throughout this paper we have worked with a subgroup $Q$ which becomes a free group after taking high powers of outer automorphisms, which satisfy certain conditions. We wish to clarify the reasons for resorting to high powers. Firstly, there does not exist a good (for our purposes) hyperbolic metric space where hyperbolic outer automorphisms act loxodromically. The second necessity comes from a technicality, which we describe via the following example:

Let $\phi$ be the outer automorphism class of the automorphism corresponding to

$$\Phi \colon a \mapsto ad, b \mapsto a, c \mapsto b, d \mapsto c$$

and let $\psi$ be the outer automorphism class of the automorphism corresponding to

$$\Psi \colon a \mapsto ac, b \mapsto a, c \mapsto b, d \mapsto db.$$

Observe that both $\phi$ and $\psi$ are positive automorphisms and one can get a relative train-track map (in fact a CT map) on the standard rose with four petals. $\phi$ has only one strata (the full graph) and $\psi$ has an NEG strata (namely $d$) sitting on top of an EG strata (namely $\{a, b, c\}$), but the NEG edge $d$ grows exponentially under iteration. In both cases, we have exactly one EG strata and the corresponding attracting lamination has a trivial nonattracting subgroup system by design.

Thus $\phi, \psi$ are both atoroidal outer automorphisms (since nonattracting subgroup system for the unique attracting lamination is trivial in both cases). Also, they have disjoint laminations since by design the free factor support of the attracting lamination of $\psi$ is $[\langle a, b, c \rangle]$ whereas $\phi$ is fully irreducible, hence the associated lamination fills. These automorphisms are sufficiently different and have empty sinks.

If one looks at the extension group $E_Q$ associated to the group $Q = \langle \phi, \psi \rangle$, then it follows that $\Phi^{-1}\Psi$ fixes both $b, c$ and hence the group $E_Q$ has multiple copies of $\mathbb{Z} \oplus \mathbb{Z}$, so $E_Q$ cannot be hyperbolic. Now consider $H = \langle \Phi^{-1}\Psi, b, c \rangle$. If $E_Q$ were relatively hyperbolic, then $H$ would be conjugate into some peripheral subgroup (call it $P$) in this relatively hyperbolic structure. But $b \in \Phi H \Phi^{-1}$ together with malnormality implies that $H$ and $\Phi H \Phi^{-1}$ are both contained in $P$. Hence $\Phi b \Phi^{-1} \in P \implies a \in P$. Applying the same argument again, $\Phi a \Phi^{-1} \in P \implies ad \in P \implies d \in P$. As a result, $\mathbb{F} < P$. Hence, $E_Q$ cannot be relatively hyperbolic.

However, as we shall see later if the group $Q$, whose generators are sufficiently different with empty corresponding sinks is generated instead by *sufficiently high powers of $\phi$ and $\psi$*, its extension $E_Q$ is hyperbolic –which is something one expects. This example highlights a tractable but key issue due to which we need to pass to high powers throughout this paper.

## 3.1  Sufficiently different outer automorphisms and their dynamics on $\mathcal{FF}$

In this section we will let our exponentially growing automorphisms have fixed points, but require those fixed points to be sufficiently apart from each other in $\mathcal{FF}$. We begin our analysis on $\mathcal{FF}$ and combine dynamics of automorphisms on $\mathcal{FF}$ with the dynamics on the set of laminations via weak attraction theory.

**Definition 3.1**  A collection of outer automorphisms will be called *sufficiently different* if

- no pair of the automorphisms have a common power (hence they do not generate a virtually cyclic (elementary) subgroup of $\mathrm{Out}(\mathbb{F})$), and

- fixed vertices (if any) of these outer automorphisms on $\mathcal{FF}$ are distance at least 2 from each other in $\mathcal{FF}$.

Given that $\mathcal{FF}$ has infinite diameter, it is natural to expect that two randomly chosen exponentially growing outer automorphisms will be sufficiently different. Hence, we study their dynamics in relation to one another.

**Lemma 3.2**  *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing outer automorphisms of $\mathbb{F}$ that are sufficiently different. Then,*

(1)  *No generic leaf of any attracting or repelling lamination of a $\phi_i$ is carried by the nonattracting subgroup system of any attracting or repelling lamination of $\phi_j$, for $i \neq j$.*

(2)  *There exists some $M$ such that whenever $m_i \geq M$, for all $i$, $\langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group.*

**Proof**  If no $\phi_i$ fixes any vertex in the free factor complex, then they are all fully irreducible and the attracting and repelling laminations associated to them fill $\mathbb{F}$ (see Lemma 2.2 and the paragraph afterwards). The result then follows from an induction argument applied to [5, Proposition 3.7].

Suppose that some $\phi_i$ fixes some vertex in $\mathcal{FF}$. Let $\Lambda_j^+$ be an attracting lamination of $\phi_j$, $j \neq i$. If $\mathcal{F}_{\text{supp}}(\Lambda_j^+)$ fills then we are done as in this case no generic leaf of $\Lambda_j^+$ can be carried by the nonattracting subgroup system of any attracting or repelling lamination of $\phi_i$. Suppose then that $\mathcal{F}_{\text{supp}}(\Lambda_j^+)$ is proper and $\Lambda_j^+$ is carried by $\mathcal{A}_{\text{na}}(\Lambda_i^+)$ for some attracting lamination $\Lambda_i^+$ of $\phi_i$. This would imply that $\mathcal{F}_{\text{supp}}(\Lambda_j^+)$ is carried by a free factor component $[B] \in \mathcal{A}_{\text{na}}(\Lambda_i^+)$. This violates the distance requirement of being sufficiently different. So the result follows.

To prove (2), we claim that there exists an attracting (respectively, repelling) lamination $\Lambda_i^+$ (respectively, $\Lambda_i^-$) of $\phi_i$ and an attracting (respectively, repelling) lamination $\Lambda_j^+$ (respectively, $\Lambda_j^-$) of $\phi_j$ such that no generic leaf of $\Lambda_i^+$ or $\Lambda_i^-$ is asymptotic to any generic leaf of $\Lambda_j^+$ or $\Lambda_j^-$. Assuming the claim to be true we are done by [23, Corollary 2.17] and applying induction to [6, Lemma 3.4.2] (detecting $\mathbb{F}_2$ via laminations).

To prove the claim, suppose first that there exists some attracting lamination $\Lambda$ of $\phi_i$ such that $\mathcal{F}_{\text{supp}}(\Lambda)$ is proper. If some generic leaf of $\Lambda$ is asymptotic to some generic leaf of an attracting or repelling lamination of $\phi_j$ then birecurrency of generic leaves implies that both ends of a leaf would have the same height, and as a result these lines would have the same free factor support. Hence, $\mathcal{F}_{\text{supp}}(\Lambda)$ will be invariant under both $\phi_i$ and $\phi_j$, thus violating the distance requirement in $\mathcal{FF}$ between their fixed vertices. So the claim is proved in the case of existence of any attracting or repelling lamination for either $\phi_i$ or $\phi_j$ with a proper free factor support. Suppose then neither $\phi_i$ nor $\phi_j$ have an attracting lamination with a proper free factor support. This implies that there exist unique attracting laminations $\Lambda_i^+$, $\Lambda_j^+$ for $\phi_i$, $\phi_j$, respectively, that fill $\mathbb{F}$. The filling condition implies that $\Lambda_i^+$, $\Lambda_j^+$ are necessarily topmost. If $\Lambda_i^+$, $\Lambda_j^+$ have asymptotic generic leaves, say $\ell_i$, $\ell_j$, respectively, then $\ell_i$ cannot be carried by the nonattracting subgroup system of $\Lambda_j^+$. Also, asymptoticity will imply that $\ell_i$ cannot be weakly attracted to $\Lambda_j^-$. Hence by the weak attraction theorem [6, Theorem 6.0.1] $\ell_i$ must be a generic leaf of $\Lambda_j^+$. But by being a generic leaf, closure of $\ell_i$ is all of $\Lambda_j^+$ (and it is also all of $\Lambda_i^+$ by the nature of $\ell_i$). Hence $\Lambda_i^+ = \Lambda_j^+$. As a result, in this case the following proposition completes the proof of our claim by giving us a contradiction with the hypothesis that no pair $\phi_i$, $\phi_j$ have common power. $\qquad\square$

The following is a technical result which identifies dynamical conditions for detecting when a subgroup of $\text{Out}(\mathbb{F})$ is virtually cyclic.

**Proposition 3.3** *Let* $\phi_1, \ldots, \phi_k$ *be a collection of exponentially growing automorphisms of* $\text{Out}(\mathbb{F})$ *for which the following conditions are satisfied:*

(1) *No pair* $\phi_i$, $\phi_j$ *with* $i \neq j$ *have a common invariant free factor.*

(2) *There exists a common attracting lamination* $\Lambda^+$ *such that* $\Lambda^+$ *fills.*

*Then all* $\phi_i$ *are fully irreducibles with common powers.*

**Proof** We will prove the theorem for $k = 2$ to avoid extra notation as the general case is the same. Consider the group $\mathcal{H} = \langle \phi_1, \phi_2 \rangle < \text{Out}(\mathbb{F})$. Then $\mathcal{H} < \text{Stab}(\Lambda^+)$. If either $\phi_1$ or $\phi_2$ is fully irreducible, then $\text{Stab}(\Lambda^+)$ is virtually cyclic by [5, Theorem 2.14]. So $\mathcal{H}$ must be virtually cyclic and hence $\phi_1, \phi_2$ are both fully irreducible and must have common powers.

Suppose then that both $\phi_1, \phi_2$ are reducible. Then the group $\mathcal{H}$ is fully irreducible, since $\phi_1, \phi_2$ do not have any common invariant free factor. By applying [23, Theorem C', page 2] with $\mathcal{F} = \varnothing$, we see that there exists a fully irreducible $\xi \in \mathcal{H} < \text{Stab}(\Lambda^+)$. Let $\Lambda_\xi^+, \Lambda_\xi^-$ be the unique attracting and repelling laminations of $\xi$ and $\ell^+$ be a generic leaf of $\Lambda_\xi^+$. Then $\Lambda_\xi^+$ and $\Lambda_\xi^-$ both fill $\mathbb{F}$. If $\Lambda^+ = \Lambda_\xi^+$ or $\Lambda_\xi^-$, then we are reduced to the earlier case. So suppose $\Lambda^+$ is distinct from $\Lambda_\xi^+$ and $\Lambda_\xi^-$. Choose a train-track map $f : G \to G$ for $\xi$. Let $V_\xi^+$ be some attracting neighborhood of $\Lambda_\xi^+$, which is described by some sufficiently long generic leaf segment of $\ell^+$. Similarly choose a repelling neighborhood $V_\xi^-$. Now, by our assumption, if $\gamma^+$ is a generic leaf of $\Lambda^+$, then $\gamma^+$ fills, and it is not a generic leaf of $\Lambda_\xi^-$. Then by the weak attraction theorem [23, Corollary 2.17] together with [23, Theorem G], there exists some $k > 0$ such that $f_\#^k(\gamma^+) \in V_\xi^+$. Since $\xi \in \text{Stab}(\Lambda^+)$ and $f_\#^k(\gamma^+)$ is always birecurrent, $f_\#^k(\gamma^+) \in \Lambda^+$ is a birecurrent leaf. Hence we conclude that some birecurrent leaf of $\Lambda^+$ contains the subpath of $\ell^+$ which described $V_\xi^+$. As $\gamma^+$ was chosen to be generic leaf, we have $\overline{\gamma^+} = \Lambda^+$. Hence the chosen subpath of $\ell^+$ is also a subpath of $\gamma^+$. Since this subpath was chosen arbitrarily, we deduce that $\ell^+$ is in the closure of generic leaves of $\Lambda^+$, implying that $\Lambda_\xi^+ \subset \Lambda^+$.

Now consider the dual lamination $\Lambda^-$ of $\Lambda^+$. Since $\ell^+$ is a filling line, by the weak attraction theorem of [6, Theorem 6.0.1] either $\ell^+$ is weakly attracted to $\Lambda^-$ or $\ell^+$ is a generic leaf of $\Lambda^+$. Since being weakly attracted to a lamination is an open condition and every subpath of $\ell^+$ is a subpath of some generic leaf of $\Lambda^+$, the first option is not possible. Hence $\ell^+$ must be a generic leaf of $\Lambda^+$, implying that $\Lambda^+ = \Lambda_\xi^+$, which is a contradiction to our assumption that they are distinct. Therefore $\Lambda^+ = \Lambda_\xi^+$ or $\Lambda^+ = \Lambda_\xi^-$ and we are done by the first case. $\square$

**Remark 3.4** When $\Lambda_1^\pm$ and $\Lambda_2^\pm$ are both geometric and $\phi_1, \phi_2$ satisfy the hypothesis of Lemma 3.2, it may seem that no additional restriction is placed on the conjugacy class representing the unique indivisible Nielsen path associated to the laminations. However, we shall see in Lemma 3.9 that the hypothesis that the fixed vertex sets of $\phi_1$ and $\phi_2$ are of distance at least 2 in $\mathcal{FF}$ puts some very strong restrictions on them.

## 3.2 Cusp preserving relative hyperbolicity

Given a short exact sequence of finitely generated groups $1 \to \mathbb{F} \to E \to Q \to 1$, we say that $E$ has a *cusp-preserving* relatively hyperbolic structure if there exists a collection of finitely generated subgroups $\{H_i\}$ of $\mathbb{F}$ such that $Q$ preserves conjugacy class of each $H_i$ and $E$ is (strongly) hyperbolic relative to the collection $\{N_E(H_i)\}$ (where $N_E(H_i)$ is the normalizer of $H_i$ in $E$) (see [20, Theorem 5.1] for motivation behind this property).

Recall that an outer automorphism $\phi$ is said to be *polynomially growing* if for each automorphism $\Phi \in \text{Aut}(\mathbb{F})$ representing $\phi$ and each $c \in \mathbb{F}$ the cyclically reduced word length of $\phi^i(c)$ is bounded above by a polynomial of $i$.

A subgroup $Q < \text{Out}(\mathbb{F})$ is called *upper polynomially growing (UPG)* if each $\phi \in Q$ is polynomially growing and $Q$ has unipotent image in $\text{GL}_n(\mathbb{Z})$. If $Q < \text{Out}(\mathbb{F})$ is UPG we will call every $\phi \in Q$ also UPG.

Recall that if $\Lambda_1$, $\Lambda_2$ are attracting laminations for $\phi$ with nonattracting subgroup systems $\mathcal{K}_1$, $\mathcal{K}_2$, respectively, then we define the *meet* (Lemma 2.2) $\mathcal{K}_1 \wedge \mathcal{K}_2$ as follows:

$$\mathcal{K}_1 \wedge \mathcal{K}_2 = \{[A \cap B^w] : [A] \in \mathcal{K}_1, [B] \in \mathcal{K}_2, w \in \mathbb{F}, A \cap B^w \neq \{\text{id}\}\},$$

where $B^w$ denotes the group $w^{-1}Bw$, with $w \in \mathbb{F}$.

**Definition 3.5** (sink of an automorphism [20, Section 4.3])  Given any exponentially growing outer automorphism $\phi$ of $\mathbb{F}$, consider the full list of attracting laminations $\{\Lambda_q\}_{q=1}^r$ and let

$$\mathcal{K}_\phi^* = \mathcal{K}_1 \wedge \mathcal{K}_2 \wedge \cdots \wedge \mathcal{K}_r, \quad \text{where } \mathcal{K}_q = \mathcal{A}_{\text{na}}(\Lambda_q) \text{ for } q = 1, \dots, r.$$

If $\phi$ is not exponentially growing define $\mathcal{K}_\phi^* = \{[\mathbb{F}]\}$. We shall call $\mathcal{K}_\phi^*$ the *nonattracting sink* of $\phi$.

The nonattracting sink $\mathcal{K}_\phi^*$ of $\phi$ is thus a $\phi$-invariant malnormal subgroup system which carries all conjugacy classes which are either fixed or grow polynomially under iteration by $\phi$. Thus, by construction, the nonattracting sink is designed to trap obstructions to hyperbolicity properties of the extension groups.

**Lemma 3.6**  *Given an exponentially growing $\phi \in \text{Out}(\mathbb{F})$ and its nonattracting sink $\mathcal{K}_\phi^*$, we have:*

(1)  *The nonattracting sink of $\phi$ is an admissible subgroup system for $\phi$.*

(2)  *A conjugacy class is carried by $\mathcal{K}_\phi^*$ if and only if it is not weakly attracted to any attracting lamination of $\phi$.*

(3)  *A conjugacy class is carried by $\mathcal{K}_\phi^*$ if and only if it is either fixed by $\phi$ or grows polynomially under iteration by $\phi$.*

(4)  $K_\phi^* = \mathcal{F} \cup \{[F_1], \dots, [F_p]\}$, *for some free factor system $\mathcal{F}$ and finite-rank $\phi$-invariant subgroups $F_i \in \mathbb{F}$, each of which is a geometric component of the nonattracting subgroup system of some geometric lamination of $\phi$.*

(5)  *If $\mathcal{K}_\phi^*$ has only infinite cyclic components, then no conjugacy class has nontrivial polynomial growth under iteration by $\phi$.*

**Proof**  (1) and (2) follows from [20, Corollary 4.13]. (3) follows from the fact that any conjugacy class grows exponentially under iteration by $\phi$ if and only if it is weakly attracted to some attracting lamination of $\phi$ under iteration.

(4) follows from the proof of part (1) of [20, Proposition 4.12].

To prove (5) note that if $[w]$ is a conjugacy class that grows polynomially under iteration by $\phi$, then there must be at least a rank-2 subgroup which carries both $[w]$ and a twistor to generate at least linear-order growth. This proves that if $\mathcal{K}_\phi^*$ has only infinite cyclic components, then no conjugacy class can have nontrivial polynomial growth. $\qquad\square$

**Remark**  If $K_\phi^* = \mathcal{F} \cup \{[F_1], \ldots, [F_p]\}$ then each nontrivial $[F_i]$ arises due to presence of a geometric strata and restriction of $\phi$ to each nontrivial $F_i$ is polynomially growing. Moreover, each nontrivial $[F_i]$ carries the conjugacy of the unique closed indivisible Nielsen path associated to a geometric strata. This is an essential ingredient in the proof of [20, Proposition 4.12]

We now proceed with the following proposition which identifies a sufficient condition for a collection of outer automorphisms to not have a common admissible subgroup system.

**Proposition 3.7**  *Let* $\phi_1, \ldots, \phi_k$ *be a collection of exponentially growing automorphisms so that no pair of elements have a common power. Let* $Q$ *be any free group generated by some powers of* $\phi_i$*'s. Let* $\mathcal{K}_i^*$ *be the nonattracting sink of* $\phi_i$ *and assume that the following conditions hold:*

(1)  $\mathcal{K}_i^* \neq \varnothing$ *for some* $i \in \{1, \ldots, k\}$.

(2)  *If* $\mathcal{K}_i^* \neq \varnothing$, *then* $\mathcal{K}_i^*$ *is not carried by* $\mathcal{A}_{\mathrm{na}}(\Lambda_j^+)$ *for any attracting lamination* $\Lambda_j^+$ *of* $\phi_j$ *for some* $j \neq i$.

*Then we have the following:*

(a)  *There is no admissible subgroup system for* $Q$.

(b)  *The extension group* $\mathbb{F} \rtimes \widehat{Q}$ *is not hyperbolic relative to any collection of subgroups so that the cusps are preserved.*

**Proof**  Suppose that (1) and (2) hold. Without loss of generality assume that $\mathcal{K}_1^* \neq \varnothing$. Using item (3) of Lemma 2.3, condition (2) implies that $\mathcal{K}_1^*$ cannot be carried by any admissible subgroup system for $\phi_j$, for some $j \neq 1$. If $\mathcal{K}$ were an admissible subgroup system for $Q$, then every conjugacy class not carried by $\mathcal{K}$ would get weakly attracted to some element of $\mathcal{L}_\mathcal{K}^+(\phi_s)$ (hence would grow exponentially under iteration by $\phi_s$), for each $1 \leq s \leq k$. Hence (a) holds.

Now suppose that $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic relative to some collection of peripheral subgroups. Let $[g]$ be carried by $\mathcal{K}_i^*$. Then $[g]$ is not weakly attracted to any attracting lamination of $\phi_i$ and hence by definition either $[g]$ is fixed by $\phi_i$ or $|\phi_{i\#}^n([g])|$ grows polynomially. So $g$ must be conjugate to some element in a peripheral subgroup. As a consequence, the peripheral subgroup system in Theorem 2.7 is nonempty. Moreover, whenever $\mathbb{F} \rtimes \widehat{Q}$ preserves cusps, the same theorem ensures the existence of an admissible subgroup system, contradicting (a) above. This completes proof of (b). $\qquad\square$

**Corollary 3.8**  *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing and sufficiently different outer automorphisms of $\mathbb{F}$. Then if some $\phi_i$ fixes the conjugacy class of some rank-1 free factor $[\langle g \rangle]$, then $[g]$ can not be carried by a free factor component of the nonattracting subgroup system of any attracting lamination of $\phi_j$, for any $j \neq i$.*

*Moreover, if the distance between the fixed points sets in $\mathcal{FF}$ of any distinct pair of $\phi_i, \phi_j$ is at least 5, then $[g]$ is weakly attracted to every attracting lamination of $\phi_j$ for every $j \neq i$.*

**Proof**  For concreteness, suppose $\phi_1$ fixes the conjugacy class of a rank-1 free factor $[\langle g \rangle]$. Let $\Lambda_j^+$ be any attracting lamination of $\phi_j$, $1 \neq j$. If $[g]$ is carried by $[B] \in \mathcal{A}_{\mathrm{na}}(\Lambda_j^+) = \mathcal{F}_j \cup [F_{m_j}]$, then either $[B]$ is a free factor component of $\mathcal{F}_j$ or $[B] = [F_{m_j}]$. The former case directly violates the sufficiently different hypothesis.

To complete the proof, it remains to show that if the distance between the fixed points sets in $\mathcal{FF}$ of any distinct pair of $\phi_i, \phi_j$ is at least 5, then $[g]$ is weakly attracted to every attracting lamination of $\phi_j$ for every $j \neq i$. To see this, observe that if $[g]$ is carried by $[F_{m_j}]$, then the free factor conjugacy class $[\langle g \rangle]$ is disjoint from $\mathcal{F}_{\mathrm{supp}}(\Lambda_j^+)$ and hence their distance in $\mathcal{FF}$ is at most 4 — violating our hypothesis.   □

In Proposition 3.7 we identified sufficient conditions for a collection of exponentially growing outer automorphisms to not have a common admissible subgroup system. In the following lemma, we have a hypothesis which negates that sufficient condition and examine the impact of our hypothesis on the nonattracting sinks for a collection of sufficiently different outer automorphisms. As it turns out, the nonattracting sinks in such cases will have a very rigid structure.

For an automorphism $\phi$, closed Nielsen paths are circuits which are in bijective correspondence with conjugacy classes which are fixed by $\phi$. There are three primary cases in our analysis where Nielsen paths appear. First one (and the most important for us) as an indivisible Nielsen path associated to a geometric EG strata. The second one is from the "twistor" in Dehn twist parts. The third case is when there is an isolated invariant rank-1 free factor, not covered by the previous two cases. All Nielsen paths are carried by the nonattracting sink. In case of absence of a free factor component in the nonattracting sink, which is essentially what the hypothesis of Lemma 3.9 ensures, the second and third cases of Nielsen paths are ruled out.

**Lemma 3.9**  *Let $\phi_1, \ldots, \phi_k$ be a collection of sufficiently different exponentially growing outer automorphisms of $\mathrm{Out}(\mathbb{F})$. Let $\mathcal{K}_i^*$ be the nonattracting sink of $\phi_i$ for each $i$ and suppose that the following condition holds*:

- *If $\mathcal{K}_i^* \neq \varnothing$ then for each $j \neq i$, there exists an attracting lamination $\Lambda_j^+$ of $\phi_j$ such that $\mathcal{K}_i^*$ is carried by $\mathcal{A}_{\mathrm{na}}(\Lambda_j^+)$.*

*Then the following are true*:

(a)  *If $\mathcal{K}_i^* \neq \varnothing$ for some $i$, then $\mathcal{K}_i^* \neq \varnothing$ for every $i$. Moreover, each $\mathcal{K}_i^*$ has exactly one component.*

(b)  *If $\mathcal{K}_i^* \neq \varnothing$ for some $i$, then there exists a finite-rank subgroup $F$ such that $\mathcal{K}_i^* = [F]$ for each $i$.*

(c)  *There exists an admissible subgroup system for the collection $\phi_1, \ldots, \phi_k$.*

**Proof** For concreteness assume that $\mathcal{K}_i^* = \mathcal{F}_i \cup [F_{i_1}] \cup \cdots \cup [F_{i_n}]$ is nonempty and let $j \neq i$. Let $\Lambda_j^+$ be an attracting lamination of $\phi_j$ such that $\mathcal{A}_{\mathrm{na}}(\Lambda_j^+) = \mathcal{F}_j \cup [F_j]$ carries $\mathcal{K}_i^*$. Sufficiently different condition implies that each component of the free factor system $\mathcal{F}_i$ must be carried by $[F_j]$. Moreover, if the conjugacy class $[g_{i_k}]$ carried by the geometric component $[F_{i_k}]$, which represents the indivisible Nielsen path associated to the corresponding geometric strata, is a free factor conjugacy class, then $[F_{i_k}]$ must be carried by $[F_j]$ due to sufficiently different condition. On the other hand if $[g_{i_k}]$ is not a free factor conjugacy class and is carried by some component of $\mathcal{F}_j$, then the same component will also carry the free factor support of $[g_{i_k}]$, hence the free factor support of the corresponding geometric lamination (which is invariant under $\phi_i$) — violating sufficiently different condition. Thus each $[F_{i_k}]$ (and hence entire $\mathcal{K}_i^*$) must be carried by $[F_j]$. This also proves that $\mathcal{K}_j^*$ is nonempty as $[F_j]$ is one of the components of the sink of $\phi_j$.

Now, switching the roles of $i$, $j$ in the above argument, we see that there exists an attracting lamination $\Lambda_i^+$ with a nonattracting subgroup system $\mathcal{F}_i' \cup [F_{i_l}]$ for some $1 \leq l \leq n$ and entire $\mathcal{K}_j^*$ is carried by $[F_{i_l}]$. But we know that $[F_j]$ carries each $[F_{i_k}]$, thus proving that $[F_j] = [F_{i_l}]$ and $[F_{i_k}]$ is trivial if $k \neq l$. For the final step, observe that we have reduced to the case $\mathcal{K}_i^* = \mathcal{F}_i \cup [F_j]$ is carried by $\mathcal{A}_{\mathrm{na}}(\Lambda_j^+) = \mathcal{F}_j \cup [F_j]$. If $[B] \in \mathcal{F}_i$ is nontrivial free factor component, as seen earlier — it must be carried by $[F_j]$, which violates malnormality of $\mathcal{K}_i^*$. Hence $\mathcal{K}_i^* = [F_j]$ has a single component. Since $i$ was chosen arbitrarily, this proves (a).

To prove (b), observe that in the paragraph above, we have actually shown that $\mathcal{K}_i^* = [F_j] = [F_{i_l}] = \mathcal{K}_j^*$. Since $i$, $j$ are arbitrary here, we may let $[F] = [F_j]$ and complete the proof of (b).

For the final part, we note that if all of the nonattracting sinks are empty, then we take $\mathcal{K} = \varnothing$ to be our admissible subgroup system, otherwise we may take $\mathcal{K} = \{[F]\}$ as our admissible subgroup system. $\square$

Having identified the structure of the nonattracting sinks for a sufficiently different collection of exponentially growing outer automorphisms, we proceed to answering the question about when we can expect extension groups to have a cusp-preserving relative hyperbolic structure. The answer breaks down into three cases. The first case that we deal with is the following proposition, when all the nonattracting sinks are nonempty. The remaining two cases when either all nonattracting are empty (Theorem 1.3) or the mixed case with some being empty (Theorem 5.1) are dealt with later.

**Proposition 3.10** *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing, sufficiently different outer automorphisms. If all the nonattracting sinks are nonempty then exactly one of following are true:*

(1) *$\phi_1, \ldots, \phi_k$ is a collection of fully irreducible geometric outer automorphisms induced by some pseudo-Anosov homemorphisms of the same compact surface with one boundary component. Moreover, all sufficiently high powers of $\phi_i$'s generate a free group $Q$ such that $E_Q$ has a cusp-preserving relatively hyperbolic structure.*

(2) *$\mathcal{K}_i^* \neq \mathcal{K}_j^*$ for some $i \neq j$ and the extension $E_Q$ of a free group $Q$, which is generated by sufficiently high powers of $\phi_i$'s, cannot have a cusp preserving relatively hyperbolic structure.*

**Proof** When (1) holds, the work of Bowditch [7] and Mj–Reeves [27] for mapping classes of once punctured surfaces shows that sufficiently high powers of $\phi_i$'s generate a free group $Q$ of rank $k$ and the corresponding extension group $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic relative to the collection $\{\langle\sigma\rangle \oplus \langle\phi_i\rangle\}$, where $\sigma$ is the conjugacy class corresponding to the boundary of the surface. This gives us a cusp-preserving relatively hyperbolic structure and an admissible subgroup system $\{[\langle\sigma\rangle]\}$ for $Q$, implying that (2) fails.

Suppose that (1) is false. We argue by contradiction to show that (2) must be true. Our argument involves showing that we cannot satisfy the hypothesis of Lemma 3.9, hence we must satisfy the hypothesis of Proposition 3.7. This will allow us to conclude that admissible subgroup system does not exist. Suppose on the contrary that we do indeed satisfy the hypothesis of Lemma 3.9. Conclusion (2) of Lemma 3.9 proves the existence of a finite-rank subgroup $[F]$ such that $\mathcal{K}_i^* = [F]$ for each $i$. Moreover, the proof shows that this $[F]$ is actually a geometric component of $\mathcal{K}_i^*$, for each $i$, ie, for each $\phi_i$ there exists some (geometric) attracting lamination $\Lambda_i^+$ such that $\mathcal{A}_{\mathrm{na}}(\Lambda_i^+) = \mathcal{F}_i \cup [F]$ for some free factor system $\mathcal{F}_i$.

If $F$ is not infinite cyclic, then this means there are NEG edges attached to the boundary circle of the surface which supports the lamination $\Lambda_i^+$ for each $i$. Using the fact that $\mathcal{A}_{\mathrm{na}}(\Lambda_i^+)$ cannot carry the free factor support of $\Lambda_i^+$, we may conclude that if $[F]$ is not cyclic, then there is a common free factor conjugacy class (namely, the free factor support of the surface noted earlier) which is left invariant by each $\phi_i$ — thus violating sufficiently different hypothesis. Thus $F$ must be infinite cyclic, say $\langle c\rangle$, where $[c]$ represents the conjugacy class of the word representing the boundary of the surface which supports all the laminations $\Lambda_i^+$'s. If the free factor support of this surface was proper, then we would violate sufficiently different hypothesis. Hence the surface and therefore $[c]$ must fill, ie $\mathcal{F}_{\mathrm{supp}}([c]) = \mathbb{F}$. This implies that the geometric stratum corresponding to the $\Lambda_i^+$ is a top stratum for each $i$. Moreover, a geometric lamination being bottommost [23, Proposition 2.15] implies that any of the lower strata cannot be EG, since the lamination fills. But since the sink of each $\phi_i$ is just $[\langle c\rangle]$, the lower filtration is just trivial (contractible) in each case. Hence $\mathcal{A}_{\mathrm{na}}(\Lambda_i^+) = [F] = [\langle c\rangle]$ for each $i$.

Using [17, Lemma 5.3], this implies that each $\phi_i$ is a geometric fully irreducible map which is induced by a pseudo-Anosov homeomorphism on the same compact surface (with one boundary component), bringing us back to (1) — which we had assumed to fail — hence giving us a contradiction. In conclusion, if (1) fails, then we must satisfy the hypothesis of Proposition 3.7, and hence (2) follows using Theorem 2.7. $\square$

# 4 Small displacement sets, weak attraction theory and further obstructions to (relative) hyperbolicity

To further our analysis and suggested by the previous section, we will consider a *small displacement sets* in $\mathcal{FF}$ for exponentially growing outer automorphisms. We will then investigate interactions between attracting/repelling laminations, their corresponding nonattracting subgroup systems and the small displacements sets.

## 4.1 Small displacement sets of partial fully irreducibles

**Definition 4.1** Let $F$ be some free-factor of $\mathbb{F}$ which is invariant under $\psi \in \mathrm{Out}(\mathbb{F})$. If the restriction of $\psi$ to $F$ is fully irreducible element of $\mathrm{Out}(F)$, then we say that $\psi$ is partial fully irreducible on $F$.

Consider the isometric action of $\mathrm{Out}(\mathbb{F})$ on the free factor complex, $\mathcal{FF}$. For an automorphism $\phi$ and constant $C > 0$, we define the *small displacement set* of $\langle \phi \rangle$ on $\mathcal{FF}$ to be

$$\mathcal{S}_C = \{x \in \mathcal{FF} : \exists k \neq 0 \text{ such that } d(x, \phi^k(x)) \leq C\}.$$

Below we will prove that a small displacement set of a partial fully irreducible outer automorphism has a diameter bounded from above. For that we will need a few definitions regarding $\mathcal{FF}$.

A *marking* of a graph $G$ is a homotopy equivalence $\mathcal{R}_n \to G$ from a rose $\mathcal{R}_n$. A metric on $G$ is a function that assigns a positive number (length) to each edge of $G$. The (unprojectivized) *outer space* is a space of marked metric graphs which is introduced by Culler and Vogtmann in [10] as an analog of Teichmüller space. We will denote by CV the *projectivized* outer space, in which the graphs will all have total volume 1. For the details we refer the reader to [10; 31].

We will use the coarse projection $\pi : \mathrm{CV} \to \mathcal{FF}$ defined as follows. For each proper subgraph $\Gamma_0$ of a marked graph $G$ that contains a circle, its image in $\mathcal{FF}$ is the conjugacy class of the smallest free factor containing $\Gamma_0$. Now by [3], for two such proper subgraphs $\Gamma_1$ and $\Gamma_2$, $d_{\mathcal{FF}}(\pi(\Gamma_1), \pi(\Gamma_2)) \leq 4$ [3, Lemma 3.1]. Then for $G \in \mathrm{CV}$ we define

$$\pi(G) := \{\pi(\Gamma) \mid \Gamma \text{ is a proper, connected, noncontractible subgraph of } G\}$$

We will call the induced map $\mathrm{CV} \to \mathcal{FF}$ also $\pi$ which is clearly a *coarse* projection in that the diameter of each $\pi(G)$ is bounded by 4.

For a point $G \in \mathrm{CV}$ and $A < \mathbb{F}$ we consider the core graph $A|G$ corresponding to the conjugacy class of $A$. We will denote the projection of this core subgraph to $\mathcal{FF}(A)$ by $\pi_A(G)$.

In other words, $\pi_A(G) = \pi(A|G)$ where $\pi : \mathrm{CV} \to \mathcal{FF}$ is the coarse Lipschitz map that takes a graph to the collection of all the proper subgraphs it contains. The image of this projection is of diameter at most 4, by Bestvina–Feighn [3]. The pulled back metric via the immersion $p : A|G \to G$ gives that $A|G \in \mathrm{CV}(A)$, the outer space of $A$.

Let $A$ and $B$ be conjugacy classes of two free factors. Assume that $A$ and $B$ are not disjoint, in other words they are not free factors of a common free splitting of $\mathbb{F}$ and one of them is not included in the other. In this case $A$ and $B$ said to *overlap*. For two overlapping conjugacy classes of free factors, we have the following definition.

**Definition 4.2** [4; 28; 29] Let $A$, $B$ be two overlapping free factors with rank of $A$ at least 2. Then the *subfactor projection* $\pi_A(B)$ of $B$ to $A$ is defined to be

$$\bigcup \{\pi_A(G) : G \in \mathrm{CV} \text{ and } B|G \subset G\},$$

where $B|G \subset G$ will mean embedding of the core graph $B|G$ in $G$.

In other words, given any tree $T$ with a vertex stabilizer $B$ and $\mathbb{F}$ action, $A$ fixes a tree $T^A$. If $T^A$ is not degenerate, which is guaranteed by the overlapping condition, the induced action of $A$ on $T^A$ gives a set of vertex stabilizers by the Bass–Serre theory. We define $\pi_A(B)$ to be this set of vertex stabilizers.

**Remark** Any two rank-1 free factors of $\mathbb{F}$ are either disjoint or equal. There is no possibility of overlap. Hence the definition of subfactor projections uses rank $\geq 2$.

**Lemma 4.3** (bounded geodesic image theorem [4; 28]) *For a free group of rank $n \geq 3$ there is a number $M_0 \geq 0$ such that if $A$ is a free factor and $\gamma$ is a geodesic in $\mathcal{FF}$ such that every vertex of $\gamma$ meets $A$, then*

$$\mathrm{diam}\{\pi_A(\gamma)\} \leq M_0.$$

**Proposition 4.4** *Let $\phi$ be partially fully irreducible automorphism with respect to a free factor $F^1$ with minimum translation length $\lambda$ on $F^1$. Then, whenever $\lambda \geq M_0$ and for every $C > 0$, there exists a sufficiently large constant $D = D(C, \lambda)$ such that the diameter of the small displacement set*

$$\mathcal{S}_{\phi, C} = \{x \in \mathcal{FF} : \exists k \neq 0 \text{ such that } d_{\mathcal{FF}}(x, \phi^k(x)) \leq C\}$$

*corresponding to $\langle \phi \rangle$ is bounded above by $D$.*

**Proof** Let $x \in S_C$. Then, by definition of this set, for some $k > 0$ we have $d_{\mathcal{FF}}(x, \phi^k(x)) \leq C$. Let also $G \in \mathrm{CV}$ such that $\pi(G) = x$. Now, let $\{G_t\}$ be a folding path in CV between $G$ and $\phi^k(G)$. The projection path $\pi(\{G_t\})$ in $\mathcal{FF}$ is an unparametrized quasigeodesic between $x$ and $\phi^k(x)$ and it is Hausdorff close to a geodesic [3].

Now, since $F^1$ is of rank at least 2, the subfactor projection to the free factor complex $\mathcal{FF}(F^1)$ is coarsely defined. Moreover, by the hypothesis (or for a suitable power $> k$ of $\phi$), $d_{F_1}(\tilde{x}, \phi^k_{|F1}(\tilde{x})) > M_0$ where $\tilde{x}$ is the subfactor projection of $x$ to $F^1$ and $M_0$ is the constant from the Lemma 4.3. By Lemma 4.3 again, there is a vertex $F^2$ along $\pi(\{G_t\})$ which does not project to $\mathcal{FF}(F^1)$. By [28] this means that either $F^1$ and $F^2$ are disjoint or one is included in the other as subgroups. Hence we have $d_{\mathcal{FF}}(F^1, F^2) \leq 4$.

Use triangle inequality to deduce that $d_{\mathcal{FF}}(x, F^1) \leq 4 + C + C_1$ where $C_1$ is the Hausdorff distance between $\pi(\{G_t\})$ and the geodesic between $x$ and $\phi^k(x)$. Hence, $D = 2(4 + C + C_1)$. $\square$

The following proposition captures the impact of the separation of small displacement sets of automorphisms $\phi_i$, $\phi_j$, $i \neq j$, on the weak attraction property of their corresponding lamination sets. We continue using the same notation we set up at the beginning of this subsection (see the paragraph before Proposition 5.3).

Thus, let $\mathcal{S}_\phi$ denote the small displacement set of $\langle \phi \rangle$ in $\mathcal{FF}$. We define the distance between the small displacement sets of $\phi_i$ and $\phi_j$ in $\mathcal{FF}$ to be

$$d_{\mathcal{FF}}(\mathcal{S}_{\phi_i}, \mathcal{S}_{\phi_j}) := \min\{d_{\mathcal{FF}}([A], [B]) : A \in \mathcal{S}_{\phi_i}, B \in \mathcal{S}_{\phi_j}\}$$

**Proposition 4.5** *Let $\phi_i$ be partially fully irreducible on $F^i$ for all $i \in \{1, \ldots, k\}$. Assume that $d_{\mathcal{FF}}(\mathcal{S}_{\phi_i}, \mathcal{S}_{\phi_j}) \geq 5$ for all $i \neq j$.*

(1) *There exists some $M$ such that whenever $m_i \geq M$, for all $i$, $\langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group.*

(2) *For $i \neq j$, $\phi_i$ and $\phi_j$ do not have any common invariant free factors.*

(3) *For all $i \in \{1, \ldots, k\}$, let $\Lambda_i^+$ be any attracting lamination of $\phi_i$ and $\mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$ be nontrivial. For a pair $\phi_i, \phi_j, i \neq j$, let $\{[A_k^i]\}_{k=1}^p = \mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$ and $\{[B_s^j]\}_{s=1}^q = \mathcal{A}_{\mathrm{na}}(\Lambda_j^{\pm})$. Then, $\{[A_k^i \cap B_s^j]\}_{k,s}$ is a malnormal subgroup system.*

(4) *No generic leaf of any attracting or repelling lamination of a $\phi_i$ is carried by a nonattracting subgroup system for **any** nongeometric dual lamination pair of $\phi_j$, for $j \neq i$.*

(5) *If $[g]$ is a nontrivial conjugacy class of a free factor which is fixed by $\phi_i$, then $[g]$ is cannot be carried by the nonattracting subgroup system of any attracting lamination of any $\phi_j$ for $j \neq i$.*

**Proof** (1) is straightforward using Lemma 3.2.

Any invariant free factor of $\phi_i$ is contained in the respective small displacement set. When the respective small displacement sets are disjoint, the distance between any invariant free factor of $\phi_i$ and any invariant free factor of $\phi_j$ is greater than 0. Hence we cannot have a common invariant free factor. This proves (2).

Every free factor component of $\mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$ is a vertex in the small displacement set of $\phi_i$. When the small displacement sets are of distance at least 5, every free factor component of $\mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$ of rank $\geq 2$ meets every free factor component of $\mathcal{A}_{\mathrm{na}}(\Lambda_j^{\pm})$ of rank $\geq 2$. Malnormality of the collection of subgroups $\{[A_k^i \cap B_s^j]\}_{k,s}$ follows directly from the malnormality of $\mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$, for all $i$. This proves (3).

Let $\Lambda_j^+$ be any attracting lamination for $\phi_j$ with nontrivial nonattracting subgroup system $\mathcal{A}_{\mathrm{na}}(\Lambda_j^{\pm})$. By definition every free factor component of $\mathcal{A}_{\mathrm{na}}(\Lambda_j^{\pm})$ is invariant under $\phi_j$. Hence every free factor component of $\mathcal{A}_{\mathrm{na}}(\Lambda_j^{\pm})$ is an element of the small displacement set of $\phi_j$. Let $\ell$ be a generic leaf of some attracting lamination $\Lambda_i^+$, $i \neq j$, and let $[B] \in \mathcal{A}_{\mathrm{na}}(\Lambda_j^{\pm})$ be some free factor component of rank $\geq 2$. Assume $\ell$ is carried by $[B]$. Then $\Lambda_i^+$ is carried by $[B]$ which implies that $\mathcal{F}_{\mathrm{supp}}(\Lambda_i^+) < [B]$, since $\mathcal{F}_{\mathrm{supp}}(\Lambda_i^+)$ is the smallest free factor system that contains $\Lambda_i^+$. But $\mathcal{F}_{\mathrm{supp}}(\Lambda_i^+)$ is invariant under $\phi_i$ and hence $\mathcal{F}_{\mathrm{supp}}(\Lambda_i^+) \in \mathcal{S}_{\phi_i}$ which violates $d_{\mathcal{FF}}(\mathcal{S}_{\phi_i}, \mathcal{S}_{\phi_j}) \geq 5$ when $i \neq j$. This completes the proof of (4).

Let $[\langle g \rangle]$ be the conjugacy class of a rank-1 free factor fixed by $\phi_1$. If $j \neq 1$ and $\Lambda_j^+$ is an attracting lamination of $\phi_j$, with $\mathcal{A}_{\mathrm{na}}(\Lambda_j^+) = \mathcal{F}_j \cup [F_{m_j}]$. If $[g]$ is carried by $\mathcal{F}_j$, we would violate the distance between small displacement sets $\geq 5$ condition. So suppose that $[g]$ is carried by $[F_{m_j}]$. In this case, note that the free factor support of $\Lambda_j^+$ is not carried by $[F_{m_j}]$ and hence the free factor conjugacy classes $[\langle g \rangle]$ and $\mathcal{F}_{\mathrm{supp}}(\Lambda_j^+)$ are disjoint and therefore violate that the small displacement sets have distance $\geq 5$. Therefore, $[g]$ cannot be carried by $\mathcal{A}_{\mathrm{na}}(\Lambda_j^+)$.                                                                                                      $\square$

The following proposition produces examples similar to Example 6.2 when the small displacement sets are sufficiently separated.

**Proposition 4.6**   *Let $\phi_i$ be partially fully irreducible on $F^i$ and assume $d_{\mathcal{FF}}(\mathcal{S}_{\phi_i}, \mathcal{S}_{\phi_j}) \geq 5$ for all $j \neq i$. Suppose that some $\phi_i$ fixes a rank-1 free factor $[g]$. Then $\mathbb{F} \rtimes \widehat{Q}$ is not hyperbolic relative to any collection of subgroups so that cusps are preserved.*

**Proof**   By Proposition 4.5(5), $[g]$ is not carried by the nonattracting subgroup system of any attracting lamination of any $\phi_j$ for $j \neq i$. This tells us that $[g]$ is weakly attracted to every attracting lamination of $\phi_j$, $j \neq i$. Using the fact that $[g]$ is carried by $\mathcal{K}_i^*$, Proposition 3.7 now completes the proof.   $\square$

# 5   Geometry of free by free extensions

## 5.1   Hyperbolicity of the extension from independent atoroidal automorphisms

Now we are ready to characterize hyperbolicity of the extension group when automorphisms are atoroidal.

**Theorem 1.1**   *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing outer automorphisms such that no pair of automorphisms have a common power. Then the following are equivalent:*

(1)   *Each $\phi_i$ is an atoroidal outer automorphism and $\mathcal{L}^{\pm}(\phi_i) \cap \mathcal{L}^{\pm}(\phi_j) = \varnothing$ for all $i \neq j$.*

(2)   *There exists $M > 0$ such that for all $m_i \geq M$, $Q = \langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group and the extension group $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic (where $\widehat{Q}$ is any lift of $Q$).*

**Proof**   We prove the result for $k = 2$ for the sake of simplicity of the notation. The general case is an exact replica of the argument.

**(1) $\implies$ (2)**   Since $\phi_1$, $\phi_2$ are both hyperbolic, by [18, Lemma 3.1] every conjugacy class grows exponentially under iteration by both of them. So $\mathcal{K}_1^* = \mathcal{K}_2^* = \varnothing$. We will use Theorem 2.7 for $\mathcal{K} = \mathcal{K}_1^* = \mathcal{K}_2^* = \varnothing$ (there is no restriction on $\mathcal{K}$ being nonempty in that theorem). To this end, all that remains is to show that the set of laminations are pairwise disjoint. Let $f \colon G \to G$ be a CT–map representing $\phi_1$. If a generic leaf $\ell_1 \in \Lambda_1^+ \in \mathcal{L}^{\pm}(\phi_1)$ has a common end with a generic leaf $\ell_2 \in \Lambda_2^+ \in \mathcal{L}^{\pm}(\phi_2)$, then birecurrence implies that both ends of $\ell_i$ have height $s$, where $H_s$ is the exponentially growing strata corresponding to $\Lambda_1^+$ and $\ell_1, \ell_2 \subset G_s$. Since $\ell_2$ is asymptotic to $\ell_1$ and being weakly attracted to the dual lamination $\Lambda_1^-$ is an open condition, $\ell_2$ cannot be weakly attracted to $\Lambda_1^-$ under iteration by $\phi_1^{-1}$. Using [6, Proposition 6.0.8], we get that $\ell_2$ is a generic leaf of $\Lambda_1^+$. Hence $\Lambda_1^+ = \bar{\ell}_2 = \Lambda_2^+$, violating our hypothesis. Hence (2) follows from Theorem 2.7 as $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic relative to empty sets by this theorem.

**(2) $\implies$ (1)**   Hyperbolicity of $\mathbb{F} \rtimes \widehat{Q}$ implies that the Cannon–Thurston map for the inclusion $\iota \colon \mathbb{F} \to \mathbb{F} \rtimes \widehat{Q}$ exists. Since $\phi_1$, $\phi_2$ do not have a common power, $\phi_1^{\infty}$, $\phi_2^{\infty}$ represent two distinct points in the boundary of $\mathbb{F} \rtimes \widehat{Q}$ corresponding the forward end of the axis generated by the elements $\phi_1$, $\phi_2$, respectively. Using [26, Proposition 5.1], we see that the ending lamination sets of corresponding to $\phi_1^{\infty}$ and $\phi_2^{\infty}$

must be disjoint. Since [18, Theorem 3.1, Lemma 4.4] guarantees that generic leaves of all attracting laminations for $\phi_i$ is contained in the ending lamination set corresponding to $\phi_i^\infty$ and we know that the ending lamination set corresponding to any boundary point is a closed set, we get that $\phi_1$ and $\phi_2$ have no common attracting lamination. Similar arguments, while working with pairs of distinct boundary points $\{\phi_1^\infty, \phi_2^{-\infty}\}$, $\{\phi_1^{-\infty}, \phi_2^{-\infty}\}$, $\{\phi_1^{-\infty}, \phi_2^\infty\}$, gives us $\mathcal{L}^\pm(\phi_1) \cap \mathcal{L}^\pm(\phi_2) = \varnothing$. $\square$

A simple application of Theorem 2.7 given below demonstrates a way to construct free-by-free hyperbolic groups. We neither need the automorphisms to be fully irreducible, nor do we need quasi-isometrically embedded orbits for the quotient group for the hypothesis. The converse to the corollary is obviously false.

**Theorem 1.2** *Let $\phi_1, \ldots, \phi_k$ be a collection of atoroidal elements which do not have a common power. If no pair $\phi_i, \phi_j$, $i \neq j$, fixes a common vertex in the free factor complex $\mathcal{FF}$, then*

(1) *there exists $M > 0$ such that for all $m_i \geq M$, $Q = \langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group,*

(2) *the extension group $\mathbb{F} \rtimes \widehat{Q}$ is a hyperbolic group (where $\widehat{Q}$ is any lift of $Q$).*

**Proof** Let $\Lambda^+ \in \mathcal{L}^\pm(\phi_i) \cap \mathcal{L}^\pm(\phi_j)$ be an attracting lamination for some $i \neq j$. If $\mathcal{F}_{\text{supp}}(\Lambda)$ is proper, then $\phi_i, \phi_j$ both fix the vertex corresponding to $\mathcal{F}_{\text{supp}}(\Lambda)$, giving us a contradiction. If $\Lambda$ fills $\mathbb{F}$, then Proposition 3.3 gives us a contradiction. Hence $\mathcal{L}^\pm(\phi_i) \cap \mathcal{L}^\pm(\phi_j) = \varnothing$ and the conclusions follow from Theorem 1.1. $\square$

## 5.2 Hyperbolicity and relative hyperbolicity from sufficiently different automorphisms

In this section we will focus more on the action on $\mathcal{FF}$ of our outer automorphisms; fixing some free factors.

**Theorem 1.3** *Let $\phi_1, \ldots, \phi_k$ be a collection of exponentially growing and sufficiently different outer automorphisms. Let $Q$ be a free group generated by sufficiently high powers of $\phi_i$'s. Then, $\mathbb{F} \rtimes \widehat{Q}$ is a hyperbolic group if and only if $\mathcal{K}_i^* = \varnothing$ for all $i \in \{1, \ldots, k\}$.*

**Proof** Suppose that $\mathcal{K}_i^* = \varnothing$ for all $i \in \{1, \ldots, k\}$. If none of the $\phi_i$ fixes a vertex in $\mathcal{FF}$, then they are all fully irreducibles which do not have a common power, and we are done by [5, Theorem 5.2]. Suppose some $\phi_i$ fixes a vertex in $\mathcal{FF}$ and $\mathcal{K}_i^* = \varnothing$ for all $i \in \{1, \ldots, k\}$. To apply Theorem 1.1 it is enough to show that $\mathcal{L}^\pm(\phi_i) \cap \mathcal{L}^\pm(\phi_j) = \varnothing$ for all $j \neq i$.

Since $\mathcal{K}_i^*$ is trivial every conjugacy class in $\mathbb{F}$ is weakly attracted to some attracting lamination of $\phi_i$. Hence $\phi_i$ is hyperbolic (atoroidal) outer automorphism for all $i \in \{1, \ldots, k\}$.

**Case 1** Let $\Lambda_i^+$ be an attracting lamination of $\phi_i$ such that $\mathcal{F}_{\text{supp}}(\Lambda_i^+)$ is a proper free factor. Let $j \neq i$. If $\phi_j$ is fully irreducible, then $\Lambda_i^+$ cannot be an attracting or repelling lamination of $\phi_j$ since $\phi_j$ has unique attracting and repelling laminations which fill. If $\phi_j$ is reducible, then $\Lambda_i^+$ cannot be an attracting or repelling lamination of $\phi_j$ for otherwise $\mathcal{F}_{\text{supp}}(\Lambda_i^+)$ would be a common invariant free factor and this contradicts the fact that any vertex fixed by $\phi_i$ is distance at least 2 from any vertex fixed by $\phi_j$.

**Case 2**  Suppose that $\Lambda_i^+$ fills. Pick some $j \neq i$. Then $\Lambda_i^+$ is not an attracting lamination of $\phi_j$ for otherwise hypotheses of Proposition 3.3 are satisfied, which contradicts with the fact that $\phi_i$, $\phi_j$ do not have a common power. Same contradiction occurs if $\Lambda_i^+$ is a repelling lamination of $\phi_j$ by replacing $\phi_j$ with $\phi_j^{-1}$ in Proposition 3.3. Therefore $\phi_i$ and $\phi_j$ do not have any common attracting or repelling laminations. Now apply Theorem 1.1 to get hyperbolicity of $\mathbb{F} \rtimes \widehat{Q}$.

The converse part easily follows from the observation that hyperbolicity of $\mathbb{F} \rtimes \widehat{Q}$ implies $\phi_i$, $\phi_j$ are both hyperbolic outer automorphisms for all $i \neq j$, $i, j \in \{1, \dots, k\}$, and using [18, Lemma 3.1].     $\square$

The following theorem gives a necessary and sufficient condition for $\mathbb{F} \rtimes \widehat{Q}$ to have a cusp preserving relative hyperbolic structure when $Q$ is generated by a collection of sufficiently different outer automorphisms. One should read the following result as a relative hyperbolic analog of Theorem 1.3.

**Theorem 5.1**  *Let $\phi_1, \dots, \phi_k$ be a collection of pairwise sufficiently different and exponentially growing outer automorphisms. Let $\mathcal{K}_i^*$ be the nonattracting sink of $\phi_i$, $i \in \{1, \dots, k\}$, and assume that $\mathcal{K}_j^* \neq \varnothing$ for some fixed $j$. Then the following are equivalent:*

(1)  *$\mathbb{F} \rtimes \widehat{Q}$ has a cusp preserving relatively hyperbolic structure where $Q$ is a free group generated by sufficiently large exponents of $\phi_i$'s.*

(2)  *$\mathcal{K}_i^* = \mathcal{K}_j^*$ for all $i$.*

**Proof**  Lemma 3.2 ensures high enough $m_i$'s so that $Q = \langle \phi_1^{m_1}, \dots, \phi_k^{m_k} \rangle$ is a free group. If (2) holds, then (1) immediately follows from Theorem 2.7.

Suppose next that for a free group $Q$ generated by sufficiently large exponents of $\phi_i$'s, $\mathbb{F} \rtimes \widehat{Q}$ has a relatively hyperbolic structure where cusps are preserved. By Proposition 3.7 $\mathcal{K}_j^*$ must be carried by nonattracting subgroup system of some $\Lambda_i \in \mathcal{L}^+(\phi_i)$ for each $i \in \{1, \dots, k\}$, $i \neq j$. Lemma 3.9(b) completes the proof.     $\square$

## 5.3  Relatively hyperbolic extensions using partially fully irreducibles

In this section we will give examples of relatively hyperbolic extensions using specific types of exponentially growing outer automorphisms. Our first example is for an exponentially growing outer automorphism which is *partially* fully irreducible.

Let $\phi_1, \dots, \phi_k$ be a collection of partially fully irreducibles, each with respect to a free factor $F^i$ with $[F^i] \neq [F^j]$ whenever $i \neq j$. We denote the corresponding dual attracting and repelling lamination pair of each $\phi_i$ by $\Lambda_i^+$ and $\Lambda_i^-$. By partial fully irreducibility we have $\mathcal{F}_{\text{supp}}(\Lambda_i^\pm) = [F^i]$. We use these conventions below.

**Lemma 5.2**  *Let $\phi_i$ be partially fully irreducible on $F^i$, $i \in \{1, \dots, k\}$. Then $\Lambda_i^\pm$ and $\Lambda_j^\pm$ have asymptotic leaves if and only if $[F^i] = [F^j]$.*

**Proof** Since the restriction of $\phi_i$ to $[F^i]$ is fully irreducible, every leaf of $\Lambda_i^+$ and of $\Lambda_i^-$ is generic. If $\ell_i \in \Lambda_i^+$ and $\ell_j \in \Lambda_j^+$ have a common end, birecurrency of generic leaves implies that $\ell_i$ is carried by $[F^j] = \mathcal{F}_{\text{supp}}(\Lambda_j^+)$. But since $\mathcal{F}_{\text{supp}}(\Lambda_i^+) = \mathcal{F}_{\text{supp}}(\ell_i) = [F^i]$, we must have $[F^i] = [F^j]$. The converse direction is obvious. $\qquad\square$

**Proposition 5.3** *Let $\phi_i$ be partially fully irreducible on $F^i$, where $[F^i] \neq [F^j]$ whenever $i \neq j$, $i, j \in \{1, \ldots, k\}$.*

(1) *For all $i \in \{1, \ldots, k\}$, $[F^i]$ and free factor components of $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$ are close in the free factor complex $\mathcal{FF}$.*

(2) *For $i \neq j$, $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$ carries a leaf of $\Lambda_j^+$ if and only if $[F^j]$ is carried by $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$. Equivalently, $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$ carries a leaf of $\Lambda_j^+$ if and only if $[F^j]$ is distance $\leq 1$ from some free factor component of $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$.*

(3) *If $[F^i]$ is at distance at least 2 from every free factor component of $\mathcal{A}_{\text{na}}(\Lambda_j^\pm)$, then any generic leaf of $\Lambda_i^+$ is weakly attracted to $\Lambda_j^+$ under the action of $\phi_i$ $(\phi_i^{-1})$, for all $i, j \in \{1, \ldots, k\}, i \neq j$.*

**Proof** Fix $i$. To prove (1), we show that any representative of $[F^i]$ and any representative of any free factor component of $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$ intersect at identity only. If they are not, there is some nontrivial conjugacy class $[g]$ of $F^i$ which is carried also by $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$. Since $\phi_i$ restricted to $F^i$ is fully irreducible, this implies that $\Lambda_i^+$ must be geometric and $[g]$ must be a representative of the unique closed indivisible Nielsen path corresponding to $\Lambda_i^+$. But this implies that $[g]$ cannot be a free factor of $\mathbb{F}$ and hence cannot be carried by any free factor component of $\mathcal{A}_{\text{na}}(\Lambda_i^\pm)$. The contradiction proves our claim.

Since the restriction of each $\phi_i$ to $F^i$ is fully irreducible, every leaf of $\Lambda_i^+$ is generic and hence not a circuit. If $\ell \in \Lambda_i^+$ is any leaf which is carried by $\mathcal{A}_{\text{na}}(\Lambda_j^\pm)$, $i \neq j$, malnormality of $\mathcal{A}_{\text{na}}(\Lambda_j^\pm)$ implies that there must be some free factor conjugacy class $[H] \in \mathcal{A}_{\text{na}}(\Lambda_j^\pm)$ which carries $\Lambda_i^+$ (since the set of lines carried by $[H]$ is closed and $\bar{\ell} = \Lambda_i^+$). By definition of free factor support we then get $F^i \leq H$ up to conjugation, which is equivalent to saying that the distance between $[F^i]$ and $[H]$ in the free factor complex of $\mathbb{F}$ is $\leq 1$. The converse follows directly from definitions.

The proof of (3) is exactly the same as the proof of [19, Lemma 5.2] (the cited proof does not anywhere use the assumptions of being atoroidal on the automorphisms) or one can see [23, Remark 2.8, page 208]. $\qquad\square$

**Lemma 5.4** *Let $\psi \in \text{Out}(\mathbb{F})$ be partially fully irreducible with respect to $F$. Then there exists a $\psi$-invariant, malnormal subgroup system $\mathcal{K} = \{[K_1], \ldots, [K_p]\}$ such that the extension group $\mathbb{F} \rtimes \langle \psi \rangle$ is strongly hyperbolic relative to the collection of subgroups $\{K_s \rtimes_{\Psi_s} \mathbb{Z}\}_{s=1}^p$, where $\Psi_s$ is a chosen lift of $\phi$ such that $\Psi_s(K_s) = K_s$.*

**Proof** If $\psi$ is partially fully irreducible with respect to $F$, then restriction of $\psi$ to $F$ is fully irreducible and hence there exists an attracting lamination $\Lambda^+$ supported by $[F]$.

Let $\mathcal{K}$ be the nonattracting subgroup system associated to $\Lambda^+$. Now apply the main result of [19] with $\mathcal{K}$ and $\psi$. $\qquad\square$

**Theorem 5.5** *Let $\phi_i$ be partially fully irreducible on $F^i$, and $[F^i] \neq [F^j]$, $i, j \in \{1, \ldots, k\}$. Suppose that $[F^i]$ is distance $\geq 2$ from every free-factor component of $\mathcal{A}_{\mathrm{na}}(\Lambda_j^\pm)$ for all $i = 1, \ldots, k$ and $i \neq j$; if such free factor components exist. Then there exists $M > 0$ such that for every $m_i \geq M$, $Q = \langle \phi_1^{m_1}, \ldots, \phi_k^{m_k} \rangle$ is a free group of rank $k$.*

*Moreover, if $\mathcal{A}_{\mathrm{na}}(\Lambda_1^\pm) = \mathcal{A}_{\mathrm{na}}(\Lambda_2^\pm) = \cdots = \mathcal{A}_{\mathrm{na}}(\Lambda_k^\pm)$, then for any lift $\widehat{Q}$ of $Q$, $\mathbb{F} \rtimes \widehat{Q}$ is relatively hyperbolic group.*

**Proof** Lemma 5.2 tells us that $\Lambda_i^\pm$ and $\Lambda_j^\pm$ do not have any asymptotic leaves. This implies that we can choose long generic-leaf segments of $\Lambda_i^+$ and $\Lambda_i^-$ and construct attracting and repelling neighborhoods $V_i^+$ and $V_i^-$ for each $i = 1, \ldots, k$ so that $V_i^+, V_i^-, V_j^+, V_j^-$ are pairwise disjoint for $i \neq j$.

Since generic leaves of $\Lambda_i^+$ are not carried by $\mathcal{A}_{\mathrm{na}}(\Lambda_j^\pm)$ for $i \neq j$, by Proposition 5.3 generic leaves of $\Lambda_i^+$ are weakly attracted to $\Lambda_j^+$ (respectively, to $\Lambda_j^-$) under the action of $\phi_j$ (respectively, of $\phi_j^{-1}$).

The conclusion about $Q$ being free group now follows directly using [6, Lemma 3.4.2, page 551]. For the relative hyperbolicity of $\mathbb{F} \rtimes \widehat{Q}$, let $\mathcal{K} = \mathcal{A}_{\mathrm{na}}(\Lambda_1^\pm)$ and apply of Theorem 2.7. □

The above theorem remains true if we relax the requirement of $\mathcal{A}_{\mathrm{na}}(\Lambda_1^\pm) = \mathcal{A}_{\mathrm{na}}(\Lambda_2^\pm) = \cdots = \mathcal{A}_{\mathrm{na}}(\Lambda_k^\pm)$ and replace it by the requirement of $Q$ having an admissible subgroup system.

## 5.4 Relatively hyperbolic extensions using relative fully irreducibles

Given a collection of free factors $F^1, F^2, \ldots, F^p$ of $\mathbb{F}$ such that $\mathbb{F} = F^1 * F^2 * \cdots * F^p * B$ with $B$ possibly trivial, we say that the collection forms a *free factor system*, written as $\mathcal{F} := \{[F^1], [F^2], \ldots, [F^p]\}$ and we say that a conjugacy class $[c]$ of a word $c \in \mathbb{F}$ is carried by $\mathcal{F}$ if there exists some $1 \leq s \leq p$ and a representative $H^s$ of $F^s$ such that $c \in H^s$.

**Definition 5.6** We say that an automorphism $\phi \in \mathrm{Out}(\mathbb{F})$ is *fully irreducible relative to* a free factor system $\mathcal{F}$ if $\mathcal{F}$ is $\phi$-invariant and there is no proper $\phi$-invariant free factor system $\mathcal{F}'$ such that $\mathcal{F} \sqsubset \mathcal{F}'$.

Relative fully irreducibles were developed by Handel and Mosher [23] to find a better analog of Ivanov's theorem on subgroups of mapping class groups of a surface so that the analogous theorem for subgroups of $\mathrm{Out}(\mathbb{F})$ works "inductively" on free factors that are fixed and that the behavior of the automorphism in between free factors is better understood. Below we will describe how subgroups made of relative fully irreducibles give us easy ways to construct relatively hyperbolic extensions, under some mild restrictions on the generators.

Let $\mathcal{F} \sqsubset \mathcal{F}'$ be free factor systems, where $\mathcal{F}$ is a proper free factor system. If there exists a marked graph $G$ and realizations $H \subset H' \subset G$ of these free factor systems such that $H' \smallsetminus H$ is a single edge, then $\mathcal{F}'$ is said to be a one-edge extension of $\mathcal{F}$. If no such realization exist, then $\mathcal{F}'$ is said to be a *multiedge extension* of $\mathcal{F}$.

We record the following lemma which follows from various nontrivial results in [23] and use the conclusion as our working definition of relative fully irreducible outer automorphism.

**Lemma 5.7** *Suppose $\mathcal{F} \sqsubset \{[\mathbb{F}]\}$ is a multiedge extension invariant under $\phi$ and every component of $\mathcal{F}$ is $\phi$-invariant. If $\phi$ is fully irreducible rel $\mathcal{F}$ then there exists $\phi$-invariant dual lamination pair $\Lambda_\phi^\pm$ such that the following hold:*

(1) $\mathcal{F}_{\text{supp}}(\mathcal{F}, \Lambda^\pm) = [\mathbb{F}]$.

(2) *If $\Lambda_\phi^\pm$ is nongeometric then $\mathcal{A}_{\text{na}}(\Lambda_\phi^\pm) = \mathcal{F}$.*

(3) *If $\Lambda_\phi^\pm$ is geometric then there exists a root free $\sigma \in \mathbb{F}$ such that $\mathcal{A}_{\text{na}}(\Lambda_\phi^\pm) = \mathcal{F} \cup \{[\langle \sigma \rangle]\}$.*

*Conversely, there exists a $\phi$-invariant dual lamination pair such that if (1) and (2) hold or if (1) and (3) hold then $\phi$ is fully irreducible rel $\mathcal{F}$.*

**Proof** Let $f : G \to G$ be an improved relative train-track map representing $\phi$ and $G_{r-1}$ be the filtration element realizing $\mathcal{F}$. Apply [23, Proposition 2.2] to get all the conclusions for some iterate $\phi^k$ of $\phi$.

We claim that $\Lambda_\phi^\pm$ obtained by applying [23, Proposition 2.2] must be $\phi$-invariant. Otherwise, by the definition of being fully irreducible relative to $\mathcal{F}$ $\phi(\Lambda_\phi^+)$ will be an attracting lamination which is properly contained in $G_{r-1}$ and hence is carried by $\mathcal{F}$ which in turn is carried by the nonattracting subgroup system for $\Lambda_\phi^+$. This is a contradiction, hence $\Lambda_\phi^+$ is $\phi$-invariant. Similar arguments work for $\Lambda_\phi^-$.

The converse part follows from the case analysis in the proof of [23, Theorem I, pages 18–19]. $\qquad\square$

**Theorem 5.8** *Let $\phi, \psi$ be fully irreducible relative to the multiedge extension $\mathcal{F}$, with corresponding invariant lamination pairs $\Lambda_\phi^\pm, \Lambda_\psi^\pm$ (as in the equivalence Lemma 5.7). If no leaf of $\Lambda_\phi^+ \cup \Lambda_\phi^-$ is asymptotic to any leaf of $\Lambda_\psi^+ \cup \Lambda_\psi^-$, then there exists an integer $M \geq 1$ such that:*

(1) $Q = \langle \phi^m, \psi^n \rangle$ *is free group of rank 2 for all $m, n \geq M$.*

(2) *If $\Lambda_\phi^\pm, \Lambda_\psi^\pm$ are both nongeometric then the extension group $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic relative to the finite collection of subgroups $\{F_s \rtimes \widehat{Q}_s\}_{s=1}^p$, where $\widehat{Q}_s$ is a lift that preserves $F_s$.*

(3) *If both $\Lambda_\phi^\pm$ and $\Lambda_\psi^\pm$ are geometric laminations which come from the same surface, both fixing the conjugacy class $[\sigma]$ representing the surface boundary, then $\mathbb{F} \rtimes \widehat{Q}$ is hyperbolic relative to the finite collection of subgroups $\{\langle \sigma \rangle \rtimes \widehat{Q}_\sigma\} \cup \{F_s \rtimes \widehat{Q}_s\}_{s=1}^p$, where $\widehat{Q}_s$ is a lift that preserves $F_s$ and $\widehat{Q}_\sigma$ is a lift that fixes $\sigma$.*

**Proof** To prove (1) we apply Theorem 2.7 with $\mathcal{K} = \mathcal{F}$ and to prove (2) we apply Theorem 2.7 with $\mathcal{K} = \mathcal{F} \cup [\langle \sigma \rangle]$. $\qquad\square$

As a corollary of the theorem for $\mathcal{F} = \varnothing$, we recover the case for surface group with punctures, which was proved in [27, Theorem 4.9].

# 6 Further discussion: nonrelative hyperbolicity

From the results we have so far we gather a necessary condition of nonrelative hyperbolicity:

**Proposition 6.1** *Let* $\mathbb{F} \rtimes \widehat{Q}$ *be NRH for every free group* $Q$ *which is generated by nonzero powers of* $\phi_i$*'s which are sufficiently different. Then, there exists* $i \in \{1, \ldots, k\}$ *such that* $\mathcal{K}_i^* \neq \varnothing$ *and* $\mathcal{K}_i^* \neq \mathcal{K}_j^*$ *for some* $j \neq i$.

A sufficient condition for NRH remains to be unknown. However, we investigate the following example.

**Example 6.2** Let $\mathbb{F} = \langle a, b, c, d, e \rangle$, $F^1 = \langle a, b, c \rangle$, $F^2 = \langle a, b, c, d \rangle$. Define $\phi_1$ as the outer automorphism class of the map $\Phi_1 \colon a \mapsto ac, b \mapsto a, c \mapsto b, d \mapsto dc, e \mapsto ec$ and $\phi_2$ to be the outer automorphism class of the map $\Phi_2 \colon a \mapsto ad, b \mapsto a, c \mapsto b, d \mapsto c, e \mapsto e$. Then $\mathcal{A}_{\mathrm{na}}(\Lambda_1^{\pm}) = \varnothing$ and $\mathcal{A}_{\mathrm{na}}(\Lambda_2^{\pm}) = \{[\langle e \rangle]\}$. In any relatively hyperbolic structure, the subgroup $\mathbb{Z} \oplus \mathbb{Z} \cong \langle e, \Phi_2 \rangle$ must be contained in some peripheral subgroup. A simple computation shows that $\mathbb{F} \leq \langle \Phi_1^{-1}\Phi_2, b, c \rangle$ and hence the extension group $1 \to \mathbb{F} \to E \to \langle \phi_1, \phi_2 \rangle \to 1$ is not relatively hyperbolic (NRH as in [1]).

The hypothesis for freeness in the above theorem is easily checked, and so there exists $M$ such that for every $m, n \geq M$, the group $Q = \langle \phi_1^m, \phi_2^n \rangle$ is a free group of rank 2. Once we raise the automorphisms to sufficiently high powers, the abnormality with peripheral subgroups goes away. We ask the question: is $\mathbb{F} \rtimes \widehat{Q}$ hyperbolic relative to $\langle e, \Phi_2 \rangle$ when $Q$ is generated by high enough powers of $\phi_1, \phi_2$?

**Remark 6.3** Note that $[F^i]$ is a fixed vertex for the action of $\phi_i$ on the free-factor complex $\mathcal{FF}$, for $i = 1, 2$. Also the free-factors in the set $\mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$ are fixed vertices for the same action. If we translate the hypothesis of Theorem 5.5 in the framework of the free factor complex $\mathcal{FF}$ of $\mathbb{F}$, then what we require is that $[F^1]$ and $[F^2]$ are distinct points in $\mathcal{FF}$ and $\mathcal{A}_{\mathrm{na}}(\Lambda_1^{\pm}) = \mathcal{A}_{\mathrm{na}}(\Lambda_2^{\pm})$. Using Proposition 5.3, the second hypothesis in turn implies that $[F^1]$ and the free-factor components of $\mathcal{A}_{\mathrm{na}}(\Lambda_2^{\pm})$ are close in $\mathcal{FF}$. Similarly free-factor components of $\mathcal{A}_{\mathrm{na}}(\Lambda_1^{\pm})$ and $[F^2]$ are also close in $\mathcal{FF}$. By partial fully irreducibility $[F^i]$ and free-factor components of $\mathcal{A}_{\mathrm{na}}(\Lambda_i^{\pm})$ are also close. Therefore the entire setup of Theorem 5.5 is in a small (bounded) region of $\mathcal{FF}$. The obvious question that comes to mind is: if we assume that $[F^1]$ and $[F^2]$ are far apart in Theorem 5.5 can we still expect relative hyperbolicity of $\mathbb{F} \rtimes \widehat{Q}$?

# References

[1] **J Behrstock**, **C Druţu**, **L Mosher**, *Thick metric spaces, relative hyperbolicity, and quasi-isometric rigidity*, Math. Ann. 344 (2009) 543–595 MR

[2] **M Bestvina**, **M Feighn**, *A combination theorem for negatively curved groups*, J. Differential Geom. 35 (1992) 85–101 MR

[3] **M Bestvina**, **M Feighn**, *Hyperbolicity of the complex of free factors*, Adv. Math. 256 (2014) 104–155 MR

[4] **M Bestvina**, **M Feighn**, *Subfactor projections*, J. Topol. 7 (2014) 771–804 MR

[5] **M Bestvina**, **M Feighn**, **M Handel**, *Laminations, trees, and irreducible automorphisms of free groups*, Geom. Funct. Anal. 7 (1997) 215–244 MR

[6] **M Bestvina**, **M Feighn**, **M Handel**, *The Tits alternative for* Out($F_n$), *I*: *Dynamics of exponentially-growing automorphisms*, Ann. of Math. 151 (2000) 517–623 MR

[7] **B H Bowditch**, *The Cannon–Thurston map for punctured-surface groups*, Math. Z. 255 (2007) 35–76 MR

[8] **P Brinkmann**, *Hyperbolic automorphisms of free groups*, Geom. Funct. Anal. 10 (2000) 1071–1089 MR

[9] **D Cooper**, *Automorphisms of free groups have finitely generated fixed point sets*, J. Algebra 111 (1987) 453–456 MR

[10] **M Culler**, **K Vogtmann**, *Moduli of graphs and automorphisms of free groups*, Invent. Math. 84 (1986) 91–119 MR

[11] **S Dowdall**, **I Kapovich**, **S J Taylor**, *Cannon–Thurston maps for hyperbolic free group extensions*, Israel J. Math. 216 (2016) 753–797 MR

[12] **S Dowdall**, **S J Taylor**, *The co-surface graph and the geometry of hyperbolic free group extensions*, J. Topol. 10 (2017) 447–482 MR

[13] **S Dowdall**, **S J Taylor**, *Hyperbolic extensions of free groups*, Geom. Topol. 22 (2018) 517–570 MR

[14] **B Farb**, *Relatively hyperbolic groups*, Geom. Funct. Anal. 8 (1998) 810–840 MR

[15] **B Farb**, **L Mosher**, *Convex cocompact subgroups of mapping class groups*, Geom. Topol. 6 (2002) 91–152 MR

[16] **M Feighn**, **M Handel**, *The recognition theorem for* Out($F_n$), Groups Geom. Dyn. 5 (2011) 39–106 MR

[17] **P Ghosh**, *Applications of weak attraction theory in* Out($\mathbb{F}_N$), Geom. Dedicata 181 (2016) 1–22 MR

[18] **P Ghosh**, *Limits of conjugacy classes under iterates of hyperbolic elements of* Out($\mathbb{F}$), Groups Geom. Dyn. 14 (2020) 177–211 MR

[19] **P Ghosh**, *Relative hyperbolicity of free-by-cyclic extensions*, Compos. Math. 159 (2023) 153–183 MR

[20] **P Ghosh**, **F Gültepe**, *Relative hyperbolicity of free extensions of free groups*, preprint (2023) arXiv 2307.09674 To appear in Trans. Amer. Math. Soc.

[21] **U Hamenstädt**, **S Hensel**, *Stability in outer space*, Groups Geom. Dyn. 12 (2018) 359–398 MR

[22] **U Hamenstaedt**, *Word hyperbolic extensions of surface groups*, preprint (2005) arXiv math/0505244

[23] **M Handel**, **L Mosher**, *Subgroup decomposition in* Out($F_n$), Mem. Amer. Math. Soc. 1280, Amer. Math. Soc., Providence, RI (2020) MR

[24] **A Hatcher**, **K Vogtmann**, *The complex of free factors of a free group*, Quart. J. Math. Oxford Ser. 49 (1998) 459–468 MR

[25] **R P Kent, IV**, **C J Leininger**, *Shadows of mapping class groups*: *capturing convex cocompactness*, Geom. Funct. Anal. 18 (2008) 1270–1325 MR

[26] **M Mitra**, *Ending laminations for hyperbolic group extensions*, Geom. Funct. Anal. 7 (1997) 379–402 MR

[27] **M Mj**, **L Reeves**, *A combination theorem for strong relative hyperbolicity*, Geom. Topol. 12 (2008) 1777–1798 MR

[28] **S J Taylor**, *A note on subfactor projections*, Algebr. Geom. Topol. 14 (2014) 805–821 MR

[29]   **S J Taylor**, *Right-angled Artin groups and* Out($\mathbb{F}_n$)*, I*: *Quasi-isometric embeddings*, Groups Geom. Dyn. 9 (2015) 275–316  MR

[30]   **W P Thurston**, *Hyperbolic structures on* 3-*manifolds*, *II*: *Surface groups and* 3-*manifolds which fiber over the circle*, preprint (1986)  arXiv math/9801045  Reprinted as pages 79–110 in his *Collected works with commentary*, *II*, Amer. Math. Soc., Providence, RI (2022)

[31]   **K Vogtmann**, *Automorphisms of free groups and outer space*, from "Proceedings of the Conference on Geometric and Combinatorial Group Theory, I" (L Mosher, M Sageev, editors), volume 94 (2002) 1–31  MR

*Department of Mathematics, Ashoka University*
*Sonipat, India*

*Department of Mathematics and Statistics, University of Toledo*
*Toledo, OH, United States*

`pritam.ghosh@ashoka.edu.in`,   `funda.gultepe@utoledo.edu`

# Coarse cohomology of configuration space and coarse embedding

ARKA BANERJEE

We introduce a notion of equivariant coarse cohomology of the complement of a subspace in a metric space. We use this cohomology to define a notion of coarse cohomology of the two-points configuration space of a metric space and develop tools to compute this cohomology under various conditions. As an application of this theory, we show that certain classes in the coarse cohomology of two-points configuration space obstruct coarse embedding between two metric spaces.

## 1 Introduction

Van Kampen [10] developed an obstruction theory for embeddings of $n$-dimensional simplicial complexes into $\mathbb{R}^{2n}$. A modern approach to his theory uses (co)homology of the two-points configuration space. In this article, we develop an analogous obstruction theory for coarse embedding by introducing a notion of coarse cohomology of two-points configuration space. For simplicity, throughout this article, we will say configuration space to mean two-points configuration space.

Let us first briefly describe the classical van Kampen obstruction for a topological space $X$. Let $\delta(X)$ be the diagonal set $\{(x, x) \mid x \in X\} \subset X \times X$. Consider the deleted product

$$\tilde{X} := X \times X - \delta(X) = \{(x, y) \in X \times X \mid x \neq y\},$$

where $\tilde{X}$ has a natural free action by $\mathbb{Z}_2$ by switching the coordinates. Consider the corresponding $\mathbb{Z}_2$ covering map $q: \tilde{X} \to \tilde{X}/\mathbb{Z}_2$. The space $\tilde{X}/\mathbb{Z}_2$, which we denote by $\mathrm{Conf}(X)$, is the unordered configuration space of two points in $X$. There exists a classifying map from the $\mathbb{Z}_2$-bundle $q: \tilde{X} \to \tilde{X}/\mathbb{Z}_2$ to the universal $\mathbb{Z}_2$-bundle $S^\infty \to \mathbb{R}P^\infty$ as follows:

$$
\begin{array}{ccc}
\tilde{X} & \xrightarrow{\tilde{\phi}} & S^\infty \\
\downarrow{\scriptstyle q} & & \downarrow \\
\mathrm{Conf}(X) & \xrightarrow{\phi} & \mathbb{R}P^\infty
\end{array}
$$

If there is an embedding $g: X \hookrightarrow \mathbb{R}^n$, then we can choose $\tilde{\phi}$ so that it factors through $S^{n-1}$. More precisely, we can choose $\tilde{\phi}$ to be the following map $\tilde{X} \to S^{n-1} \subset S^\infty$:

$$(x, y) \mapsto \frac{g(x) - g(y)}{|g(x) - g(y)|}.$$

In this case, $\phi$ maps $\mathrm{Conf}(X)$ to $\mathbb{R}P^{n-1} \subset \mathbb{R}P^\infty$.

So the induced map $\phi^*: \mathrm{H}^n(\mathbb{R}P^\infty) \to \mathrm{H}^n(\mathrm{Conf}(X))$ is trivial as it factors through $\mathrm{H}^n(\mathbb{R}P^{n-1})$. In particular, if $\eta^n \in \mathrm{H}^n(\mathbb{R}P^\infty; \mathbb{Z}_2)$ denotes the nonzero class, then $\phi^*(\eta^n)$ would be trivial. In other words, the cohomology class $\phi^*(\eta^n)$ gives an obstruction for the embedding of $X$ into $\mathbb{R}^n$. We will call the class $\phi^*(\eta^n)$ the van Kampen obstruction class of degree $n$ and denote it by $vk^n(X)$.

Let us now turn our attention to the coarse world. A map $f: X \to Y$ between two metric spaces is said to be a *coarse embedding* if there exist two proper nondecreasing maps $\rho_-, \rho_+: [0, \infty) \to [0, \infty)$ such that

$$\rho_-(d(x, y)) \leq d(f(x), f(y)) \leq \rho_+(d(x, y)) \quad \text{for all } x, y \in X.$$

Roe [13] defined the notion of coarse cohomology of a metric space which can be thought of as a coarse analog of Alexander–Spanier cohomology in the topological setting. The main motivation of this paper is to define a coarse version of $vk^*$ that obstructs coarse embedding between metric spaces.

The study of obstruction to the coarse embedding of finitely generated groups into proper contractible $n$-manifold was initiated by Bestvina, Kapovich, and Kleiner [3]. For several interesting classes of metric spaces (for example CAT(0) space, Gromov-hyperbolic space) one can attach a boundary at infinity to the space. Suppose $\partial_\infty X$ denotes such a boundary of a space $X$. A popular theme in the study of large scale geometry is to find topological properties of $\partial_\infty X$ that provide information about the coarse geometry of $X$. A special case of [3] roughly says the following: if $\partial_\infty X$ does not embed into $\mathbb{R}^n$ due to the van Kampen embedding obstruction, then $X$ does not coarsely embed into $\mathbb{R}^{n+1}$ with the Euclidean metric. One of the motivations behind the present work was to understand their obstruction in the language of Roe's coarse cohomology.

To this end, recall that the van Kampen obstruction class $vk^*(X)$ lives in the cohomology of the configuration space of $X$, which is the same as the $\mathbb{Z}_2$-equivariant cohomology of $X \times X - \delta(X)$ where the coefficients are $\mathbb{Z}_2$ with the trivial $\mathbb{Z}_2$ action. This suggests that a coarse version of the van Kampen obstruction class should live in some coarse version of the $\mathbb{Z}_2$-equivariant cohomology of the complement of $\delta(X)$ in $X \times X$. This motivates us to define a notion of equivariant coarse cohomology of the complement of a subspace in a metric space.

Building on Roe's theory, Banerjee and Okun [2] defined a notion of coarse cohomology of the complement of a subspace in a metric space. In this paper, we extend [2] to the equivariant setting. We then define the coarse cohomology of the configuration space of $X$ simply to be the $\mathbb{Z}_2$-equivariant coarse cohomology of the complement of $\delta(X)$ in $X \times X$, where the action of $\mathbb{Z}_2$ on $X \times X$ is by switching coordinate.

Once we have a proper notion of the coarse cohomology of the configuration space, we can search for a coarse $vk^*$ in that cohomology that obstructs coarse embeddings. Indeed, when $X$ is a separable metric space, we find a class in the $n^{\mathrm{th}}$ degree of the coarse cohomology of the configuration space of $X$, which we denote by $\mathrm{cvk}^n(X)$, that obstructs coarse embedding of $X$ into $\mathbb{R}^{n-1}$.[1] In general, the class $\mathrm{cvk}^n(X)$

---

[1]While the van Kampen obstruction $vk^n$ is associated to $\mathbb{R}^n$, note that the coarse van Kampen obstruction $\mathrm{cvk}^n$ is related to $\mathbb{R}^{n-1}$. The reason for that is coarse cohomology of the configuration space of $\mathbb{R}^n$ is the same (except in degree 0 and 1) as the cohomology of the configuration space of $\mathbb{R}^{n-1}$ (cf. Example 5.2).

can be used to get obstruction to coarse embedding into any other metric space. We define the coarse obstruction dimension $\mathrm{cobdim}(X)$ of a separable metric space $X$ to be the largest $n$ such that $\mathrm{cvk}^n(X)$ is nonzero. One of our main theorems is the following:

**Theorem 1.1** *If $X$ admits a coarse embedding into $Y$, then* $\mathrm{cobdim}(X) \le \mathrm{cobdim}(Y)$.

So one way to determine when a given space does not admit coarse embedding into another space is to compare their cobdim. However, the equivariant coarse cohomology of the complement where $\mathrm{cvk}^*$ lives is hard to compute in general. A big part of this paper is devoted to the computation of equivariant coarse cohomology of the complement for certain spaces which may be of independent interest. We use these computations to estimate cobdim of certain spaces and obtain coarse embedding obstructions between certain classes of spaces. Below we highlight some of our results in this direction.

• We show that $\mathrm{cobdim}(\mathbb{R}^n) = n$. Hence Theorem 1.1 implies that $X$ does not admit coarse embedding into $\mathbb{R}^{n-1}$ if $\mathrm{cobdim}(X) \ge n$.

• Suppose $X = K \times [0, \infty)/K \times \{0\}$ is the open cone on a finite simplicial complex $K$. A metric $d$ on $X$ is called expanding if for any two disjoint simplices $\sigma, \tau \in K$ and $S \ge 0$, there exists $r \ge 0$ such that $d(\sigma \times [r, \infty), \tau \times [r, \infty)) \ge S$. If there is a class $c \in \mathrm{H}_n(\mathrm{Conf}(X))$ such that $vk^n(X)(c) \ne 0$, then we show that $\mathrm{cobdim}(X) \ge n + 1$ whenever $X$ is equipped with a proper, expanding metric (Example 7.3). Hence such $X$ does not admit coarse embedding into $\mathbb{R}^n$ by Theorem 1.1 and the previous example. This was initially proved in [3].

• If $X$ is a proper, uniformly acyclic $n$-manifold with uniformly locally acyclic boundary, $\mathrm{cobdim}(X) \le n$ (Theorem 9.14). Examples of such spaces include universal cover of compact aspherical $n$-manifolds. Hence Theorem 1.1 implies that, if $\mathrm{cobdim}(X) \ge n$, then $X$ does not admit coarse embedding into the universal cover of any compact aspherical $(n-1)$-manifold.

• If $\mathrm{H X}^*(X^2 - \delta(X)) = 0$ for $* \le n-1$, then $\mathrm{cobdim}(X) \ge n$ (Theorem 10.1). This implies, in particular, that any proper, uniformly contractible $n$-manifold has $\mathrm{cobdim} \ge n$. Hence it follows from the last example that any proper, uniformly contractible $n$-manifold does not admit coarse embedding into the universal cover of any compact aspherical $(n-1)$-manifold (see Corollary 10.2 for a more general version). This recovers a result of Yoon [16].

• If a finitely generated group $G$ acts properly on $X$ by isometries, then there exists a coarse embedding of $G$ (equipped with a word metric coming from a finite generating set) into $X$ by mapping $G$ into an orbit of the action. That means, from the coarse point of view, any space $X$ with a proper $G$-action has to be at least as large as $G$. More precisely, using Theorem 1.1 we show that if $G$ acts properly, cocompactly on a contractible manifold $M$ (possibly with boundary), then $\dim(M) \ge \mathrm{cobdim}(G)$ (Corollary 9.17).

Equivariant coarse (co)homology has been previously studied in [5; 15]. While our approach focuses specifically on metric spaces to address the obstruction theory and computations relevant to this article,

the theories developed in [5; 15] apply to more general spaces (referred to as coarse spaces) and are motivated by certain coarse K-theoretic considerations (as discussed in the introduction of [15]). It would be intriguing to explore the extent to which our theory and computations relate to those in [5; 15].

**Overview** In Section 2 we describe several variations of Alexander–Spanier cochains and the corresponding cohomology theories. We also introduce coarse language and define coarse (co)homology. In Section 3, we define the equivariant coarse cohomology of the complement and give some examples using Theorem 3.8 that relate, for certain cases, equivariant coarse cohomology to the Alexander–Spanier cohomology of the quotient. In Section 4 we prove Theorem 3.8. In Section 5, we define the coarse cohomology of the configuration space. In Section 6, we give a class in the coarse cohomology of the configuration space that obstructs coarse embedding maps. Then we define the coarse obstruction dimension of a space and prove Theorem 6.4 which is a slightly stronger version of Theorem 1.1. In Section 7, we give a relation between classical van Kampen obstruction and coarse van Kampen obstruction. We use this relation to compute coarse van Kampen obstruction for certain Euclidean cones on simplicial complexes. In Section 8, we produce a coarse version of the Gysin sequence for the $\mathbb{Z}_2$-bundle to compute the coarse cohomology of configuration space. In Sections 9 and 10, we use the coarse Gysin sequence to estimate cobdim of certain spaces.

## 2 Preliminaries

### Alexander–Spanier complexes

We will refer to points in $X^{n+1}$ as *n-simplices*. Let $R$ be an abelian group. We will think of functions $X^{n+1} \to R$ as $n$-cochains or $n$-chains on $X$, depending on the context, and in the latter case will use additive notation $c = \sum_{\sigma \in X^{*+1}} c_\sigma \sigma$.

The basic complex is the complex of finitely supported chains

$$\mathrm{C}_*(X; R) := \left\{ c = \sum_{i=0}^{n} c_i \sigma_i \mid c_i \in R \text{ and } \sigma_i \in X^{*+1} \right\}$$

equipped with the usual boundary map, defined on the basis by

$$\partial(x_0, \ldots, x_n) := \sum_{i=0}^{n} (-1)^i (x_0, \ldots, \widehat{x}_i, \ldots, x_n).$$

The boundary map $\partial$ is well defined on a larger complex of *locally finite* chains $C_*^{lf}(X; R)$ which consists of chains $c$ satisfying the following property: for any bounded $B \subset X$ only finitely many simplices in $c$ have vertices in $B$.

The algebraic dual of $C_*(X)$ is the complex of all Alexander–Spanier cochains

$$C^*(X; R) = \{\phi : X^{*+1} \to R\}$$

with the coboundary operator

$$d(\phi)(x_0, \dots, x_n) = \sum_{i=0}^{n} (-1)^i \phi(x_0, \dots, \hat{x}_i, \dots, x_n).$$

**Lemma 2.1** $C^*(X; R)$ *is an acyclic complex, ie the homology of the complex is trivial in each degree except at degree zero where it is isomorphic to* $R$.

**Proof** In degree zero, the cohomology consists of all the constant functions $X \to R$ which are isomorphic to $R$.

To prove that the cohomology is trivial in degree $\geq 1$, choose $a \in X$. Consider the following cone operator $D_* : C_*(X) \to C_{*+1}(X)$:

$$D_n : (x_0, x_1, \dots, x_n) \mapsto (a, x_0, \dots, x_n).$$

For any $n$-simplex $\sigma$ with $n \geq 1$, we observe that $D_* \partial(\sigma) + \partial D_*(\sigma) = \sigma$. By considering the dual, we have that for any cochain $\phi \in C^*(X; R)$ in degree $\geq 1$,

$$dD^*(\phi) + D^* d(\phi) = \phi.$$

In particular, $dD^*(\phi) = \phi$ when $\phi$ is a cocycle. $\qquad\square$

All the cochain complexes that we will consider in this paper are subcomplexes of the complex $(C^*, d)$.

For a cochain $\phi \in C^*(X)$, let $\|\phi\|$ be the intersection of the diagonal $\Delta = \{(x, x, \dots, x) \mid x \in X\} \subset X^{*+1}$ and the closure of the support of the function $\phi : X^{*+1} \to R$. Let $C_0^*(X; R)$ be the complex of locally zero cochains:

$$C_0^*(X; R) = \{\phi \in C^*(X; R) \mid \|\phi\| = \varnothing\}.$$

The restriction of $d$ gives a well-defined map $C_0^*(X) \to C_0^{*+1}(X)$. Consequently, $d$ induces a well-defined map $C_{as}^*(X; R) \to C_{as}^{*+1}(X; R)$, where

$$C_{as}^*(X; R) = C^*(X; R) / C_0^*(X; R).$$

Alexander–Spanier cohomology, denoted by $H^*(X; R)$, is the cohomology of the complex $(C_{as}^*(X; R), d)$. We will denote by $\tilde{H}^*(X; R)$ the reduced Alexander–Spanier cohomology.

## Coarse inclusion

We adopt the notation from [12]. Let $(X, d)$ be a metric space. For $A \subset X$ and $r \geq 0$, we define $N_r(A) = \{x \in X \mid d(x, A) \leq r\}$. We will call such neighborhoods *metric neighborhoods* of $A$. We will say that $A$ is *$r$-contained* in $B$, $A \overset{r}{\subset} B$, if $A \subset N_r(B)$. We will say that $A$ is *coarsely contained* in $B$, $A \overset{c}{\subset} B$, if $A \overset{r}{\subset} B$ for some $r \geq 0$. Two subsets are *coarsely equal*, $A \overset{c}{=} B$, if $A \overset{c}{\subset} B$ and $B \overset{c}{\subset} A$.

## Coarse intersection

Now we recall from [12] the notion of *coarse intersection*: $A \overset{c}{\cap} B \overset{c}{=} C$ if for all sufficiently large $r \geq 0$, $N_r(A) \cap N_r(B) \overset{c}{=} C$. The coarse intersection is not always well defined, it may happen that the coarse type of $N_r(A) \cap N_r(B)$ does not stabilize as $r$ goes to infinity. However the notion "coarse intersection is coarsely contained in" is well defined. $A \overset{c}{\cap} B \overset{c}{\subset} C$ means that for any $r \geq 0$, $N_r(A) \cap N_r(B) \overset{c}{\subset} C$. It is not hard to see that this condition is equivalent to the condition that for any $r \geq 0$, $A \cap N_r(B) \overset{c}{\subset} C$.

**Notation** From now on, all spaces will be assumed to be metric spaces unless stated otherwise. We will use the letter $r$ to represent a nonnegative real number, while $R$ will denote an abelian group unless specified otherwise.

## Coarse (co)homology

In what follows, we will need to measure distances between simplices of different dimensions. A convenient way to do this is to stabilize simplices by repeating the last coordinate, as follows. Denote by $X^\infty$ the subset of the product of countably many copies of $X$, consisting of eventually constant sequences. Equip $X^\infty$ with the sup metric. Let $i \colon X^{n+1} \to X^\infty$ denote the map $(x_0, \ldots, x_n) \mapsto (x_0, \ldots, x_n, x_n, x_n, \ldots)$. For a function $\phi \colon X^{n+1} \to R$ define its stabilized support

$$|\phi| = \{i(\sigma) \mid \sigma \in X^{n+1} \text{ and } \phi(\sigma) \neq 0\} \subset X^\infty.$$

Let $\Delta = i(X)$ denote the diagonal of $X^\infty$. Define the support of $\phi$ at scale $r$ to be $|\phi|_r = |\phi| \cap N_r(\Delta)$.

We now define coarse (co)homology theories, following Roe [13] and Hair [8] using our language. The coarse cochain complex is

$$\mathrm{C}X^*(X; R) := \{\phi \in \mathrm{C}^*(X; R) \mid |\phi|_r \text{ is bounded for all } r\}.$$

An equivalent way to define coarse cochains is to require the support to be coarsely disjoint from the diagonal:

$$\mathrm{C}X^*(X; R) = \{\phi \in \mathrm{C}^*(X; R) \mid |\phi| \overset{c}{\cap} \Delta \overset{c}{=} *\}.$$

The coboundary operator $d$ preserves this property, and the coarse cohomology $\mathrm{H}\,\mathrm{X}^*(X; R)$ is defined to be the cohomology of the complex $(\mathrm{C}\,\mathrm{X}^*(X; R), d)$. The coarse homology[2] $\mathrm{H}\,\mathrm{X}_*(X; G)$ is the homology of the subcomplex of $\mathrm{C}^{\mathrm{lf}}_*(X; G)$,

$$\mathrm{C}\,\mathrm{X}_*(X; R) := \{c \in \mathrm{C}^{\mathrm{lf}}_*(X; R) \mid |c| \overset{c}{\subset} \Delta\},$$

equipped with the restriction of the boundary operator $\partial$. In the presence of the support condition $|c| \overset{c}{\subset} \Delta$ local finiteness is equivalent to $|c|$ having finite intersections with bounded subsets of $X^\infty$.

**Example 2.2** Roe [13] showed that the coarse cohomology is isomorphic to the compactly supported Alexander–Spanier cohomology if the space is uniformly contractible. In particular, this applies to the universal cover of finite aspherical complexes. In this case, the coarse homology is isomorphic to the locally finite homology [13, Chapter 2]. For example,

$$\mathrm{H}\,\mathrm{X}_*(\mathbb{R}^n; R) = \mathrm{H}\,\mathrm{X}^*(\mathbb{R}^n; R) = \begin{cases} R & \text{if } * = n, \\ 0 & \text{otherwise.} \end{cases}$$

# 3 Equivariant coarse cohomology of the complement

We start by briefly reviewing the notion of coarse cohomology of the complement from [2]. Roughly the idea is that the coarse complement of a subset $A$ in $X$ is determined by the collection of subsets $\mathcal{S}$ of $X$ which are *coarsely disjoint* from $A$:

$$\mathcal{S} := \{B \subset X \mid N_r(B) \cap N_r(A) \text{ is bounded for any } r \geq 0\}.$$

Coarse cohomology of the complement of $A$, denoted by $\mathrm{H}\,\mathrm{X}^*(X - A)$, is defined to be the cohomology of the following complex with the usual coboundary operator $d$ mentioned in the previous section:

$$\mathrm{C}\,\mathrm{X}^n(X - A; R) = \{\phi \in \mathrm{C}^n(X; R) \mid \phi|_B \in \mathrm{C}\,\mathrm{X}^n(B; R) \text{ for all } B \in \mathcal{S}\}.$$

Recall that $i: X \to X^\infty$ denotes the map $x \mapsto (x, \dots, x, \dots)$ and $\Delta = i(X)$. For a subset $A \subset X$, we denote the set $i(A) \subset \Delta$ by $\Delta_A$. For our purpose, we will work with the following equivalent description of $\mathrm{C}\,\mathrm{X}^*(X - A)$ (see [2], Lemma 3.2):

$$\mathrm{C}\,\mathrm{X}^n(X - A; R) = \{\phi \in \mathrm{C}^n(X; R) \mid |\phi| \overset{c}{\cap} \Delta \overset{c}{\subset} \Delta_A\}$$

This description of coarse cochains of the complement is closer to the spirit of Roe's original definition of coarse cochains discussed in the previous section.

---

[2]We will need coarse homology only in Section 9 for a brief discussion of coarse PD($n$) spaces. For the rest of the paper, we will work with coarse cohomology.

**Example 3.1** When $X \overset{c}{=} A$, we can see that $\mathrm{C\,X}^*(X - A; R)$ coincides with $\mathrm{C}^*(X; R)$ because in this case $\Delta \overset{c}{\subset} \Delta_A$ and therefore the support condition $|\phi| \overset{c}{\cap} \Delta \overset{c}{\subset} \Delta_A$ holds for any $\phi \in \mathrm{C}^*(X; R)$. Since $\mathrm{C}^*(X; R)$ is an acyclic complex, in this case $\mathrm{H\,X}^*(X - A; R)$ is trivial in all degrees except at degree $0$ where it is isomorphic to $R$.

**Example 3.2** If $A$ is bounded, then $\mathrm{C\,X}^*(X - A; R)$ coincides with $\mathrm{C\,X}^*(X; R)$. Hence,

$$\mathrm{H\,X}^*(X - A, R) = \mathrm{H\,X}^*(X; R)$$

whenever $A$ is bounded.

**Example 3.3** As explained in [2, Theorem 5.9], the elements of $\mathrm{H\,X}^1(X - A; \mathbb{Z}_2)$ correspond to the "coarse complementary components" of $A$ inside $X$. More precisely, if $X$ is a geodesic space and $A \subset X$ and $k = \dim_{\mathbb{Z}/2} \mathrm{H\,X}^1(X - A; \mathbb{Z}/2)$ is finite, then there exists $r \geq 0$ such that for any $L \geq r$, $X - N_L(A)$ has exactly $k + 1$ deep path components, where deep path components are those path components which are not coarsely contained in $A$.

### Equivariant coarse cohomology of the complement

We now define an equivariant version of the coarse cohomology of the complement. Let $G$ be a group acting on $X$ by isometries and $R$ be an abelian group with a $G$-action. Suppose $G$ acts on $X^{n+1}$ by the diagonal action, $g(x_0, \dots, x_n) := (gx_0, \dots, gx_n)$. $G$-equivariant coarse cohomology of the complement of $A$ in $X$, denoted by $\mathrm{H\,X}^*_G(X - A; R)$, is defined to be the cohomology of the cochain complex

$$\mathrm{C\,X}^*_G(X - A; R) := \{\phi \in \mathrm{C\,X}^*(X - A; R) \mid \phi \text{ is } G\text{-equivariant}\}$$

with the usual coboundary operator $d$.

In particular, if $G$ is acting trivially on $X$ and $R$, then $\mathrm{C\,X}^*_G(X - A; R)$ coincides with $\mathrm{C\,X}^*(X - A; R)$ and therefore $\mathrm{H\,X}^*_G(X - A; R) = \mathrm{H\,X}^*(X - A; R)$. While $\mathrm{H\,X}^*_G$ is hard to compute in general, we can relate $\mathrm{H\,X}^*_G$ to a more computable cohomology under certain acyclicity conditions which we define below.

**Definition 3.4** $X$ is locally acyclic with coefficients in $R$ if for any $x \in X$ and a neighborhood $U$ of $x$, there exists another open neighborhood $V \subset U$ of $x$ such that the inclusion $V \hookrightarrow U$ induces the trivial map in the singular homology with coefficients in $R$.

$X$ is called locally acyclic away from $A$ with coefficients in $R$ if for some $r$, $X - N_r(A)$ is locally acyclic with coefficients in $R$.

**Example 3.5** Any locally finite simplicial complex is locally acyclic. If a space is locally acyclic, then it is locally acyclic away from any of its subset. However, the converse is not true. For instance, any bounded metric space is locally acyclic away from any subset.
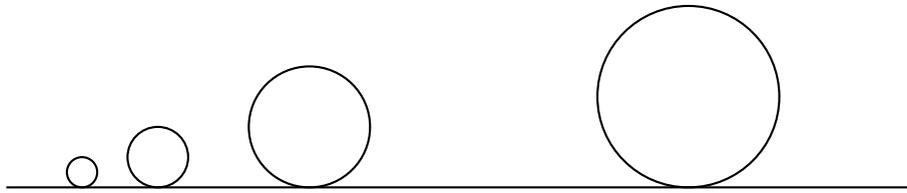
Figure 1: A subspace of $\mathbb{R}^2$ that consists of a countable union of circles $\{C_i\}_{i \in \mathbb{N}}$ and the ray $\mathcal{R} := [0, \infty) \times \{0\}$ such that the $i^{\text{th}}$ circle has radius $i$ and touches $\mathcal{R}$ at $(i^2, 0)$ so that the distance between two consecutive circles grows to infinity. This space is uniformly acyclic away from any bounded set.

**Definition 3.6** $X$ is uniformly acyclic with coefficients in $R$ if there exists a nondecreasing function $\rho \colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that for any $B \subset X$ the inclusion map $B \hookrightarrow N_{\rho(\mathrm{diam}(B))}(B)$ induces the trivial map in the singular homology with coefficients in $R$.

$X$ is called uniformly acyclic away from $A \subset X$ with coefficients in $R$ if there exist two nondecreasing functions $\rho, \mu \colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that for any $B \subset X$ the inclusion map $B \hookrightarrow N_{\rho(\mathrm{diam}(B))}(B)$ induces the trivial map in the singular homology with coefficients in $R$ if $d(B, A) \geq \mu(r)$.

**Example 3.7** Universal covers of compact aspherical complexes are uniformly acyclic. Any uniformly acyclic space is uniformly acyclic away from any of its subset. Moreover, if we remove a bounded set from a uniformly acyclic space, then the resulting space is uniformly acyclic away from any point. In general, uniform acyclicity away from a point can be very far from being uniformly acyclic. Figure 1 describes such an example.

For the rest of the paper $G \curvearrowright X$ will mean that $G$ is acting on $X$ by isometries and $G \curvearrowright (X, A)$ will mean $G \curvearrowright X$ and $GA = A$. We let

$$\mathrm{Fix}_G(X) := \{x \in X \mid gx = x \text{ for all } g \in G\}.$$

We now state a theorem that relates the equivariant coarse cohomology of the complement and the reduced Alexander–Spanier cohomology for certain spaces.

**Theorem 3.8** *Suppose* $G \curvearrowright (X, A)$ *and* $\mathrm{Fix}_G(X) \neq \varnothing$. *Let* $R$ *be an abelian group with trivial* $G$-action. *Suppose* $X$ *is uniformly acyclic away from* $A \subset X$ *with coefficients in* $R$ *and locally acyclic away from* $A$ *with coefficients in* $R$.

(1) *If* $X \overset{c}{\subset} A$, *then*
$$\mathrm{HX}_G^*(X - A; R) = \begin{cases} R & \text{if } * = 0, \\ 0 & \text{otherwise.} \end{cases}$$

(2) *If* $X$ *is not coarsely contained in* $A$, *then*
$$\mathrm{HX}_G^*(X - A; R) = \begin{cases} 0 & \text{if } * = 0, \\ \varinjlim \tilde{\mathrm{H}}^{*-1}\big((X - N_r(A))/G; R\big) & \text{otherwise,} \end{cases}$$
*where* $\tilde{\mathrm{H}}^*(-)$ *is the reduced Alexander–Spanier cohomology.*

We will postpone the proof of Theorem 3.8 until the next section. We conclude this section with an example.

**Example 3.9** Consider the action of $\mathbb{Z}_2$ on $\mathbb{R}^n$ by the antipodal map. Let $M$ be a codimension-$k$ vector subspace. Then for any abelian group $R$ with a trivial action of $\mathbb{Z}_2$ and $i \geq 1$, we have

$$
\begin{aligned}
\mathrm{H X}_G^i(\mathbb{R}^n - M; R) &= \varinjlim \tilde{\mathrm{H}}^{i-1}\big((\mathbb{R}^n - N_r(M))/\mathbb{Z}_2; R\big) \quad \text{(by Theorem 3.8)} \\
&= \tilde{\mathrm{H}}^{i-1}\big((\mathbb{R}^n - M)/\mathbb{Z}_2; R\big) \\
&= \tilde{\mathrm{H}}^{i-1}(\mathbb{R}P^{k-1}; R).
\end{aligned}
$$

The second equality holds because $\mathbb{R}^n - N_r(M)$ is $\mathbb{Z}_2$-equivariantly homotopic to $\mathbb{R}^n - M$ and the last equality follows because $\mathbb{R}^n - M$ is $\mathbb{Z}_2$-equivariantly homotopic to $S^{k-1}$ with the antipodal $\mathbb{Z}_2$-action.

# 4 Computation of equivariant coarse cohomology

In this section, our main goal is to prove Theorem 3.8. The proof is similar to the proof of a nonequivariant version of Theorem 3.8 proved in [1, Corollary 3.5]. The key is to relate the coarse cohomology of the complement to the boundedly supported cohomology of the complement which we introduce next.

**Boundedly supported cohomology of the complement**

Recall that $i \colon X \to X^\infty$ denotes the map $x \mapsto (x, \ldots, x, \ldots)$ and $\Delta = i(X)$, For a subset $A \subset X$, we denote the set $i(A) \subset \Delta$ by $\Delta_A$. Also recall that for a cochain $\phi \in \mathrm{C}^*(X)$, $\|\phi\|$ denotes the intersection of the diagonal $\Delta = \{(x, x, \ldots, x) \mid x \in X\} \subset X^{*+1}$ and the closure of the support of the function $\phi \colon X^{*+1} \to R$. Boundedly supported cohomology of the complement of $A \subset X$, denoted by $\mathrm{H B}^*(X - A)$, is the cohomology of the following cochain complex with $d$ being the coboundary operator:

$$
\mathrm{C B}^*(X - A; R) := \{\phi \in \mathrm{C}^*(X; R) \mid \|\phi\| \overset{c}{\subset} \Delta_A\}
$$

Suppose $R$ is an abelian group with a $G$-action. We define the equivariant boundedly supported cohomology of the complement with coefficients in $R$ to be the cohomology of the cochain complex

$$
\mathrm{C B}_G^*(X - A; R) := \{\phi \in \mathrm{C B}^*(X - A; R) \mid \phi \text{ is } G\text{-equivariant}\}.
$$

We denote the cohomology of the above complex by $\mathrm{H B}_G^*(X - A; R)$.

Our next goal is to relate $\mathrm{H B}_G^*$ to the equivariant Alexander–Spanier cohomology. We start by briefly recalling the notion of equivariant Alexander–Spanier cohomology.

**Equivariant Alexander–Spanier cohomology**

Honkasalo defined a notion of equivariant Alexander–Spanier cohomology in [9]. Let $R$ be an abelian group with a $G$-action. Consider the cochain complex

$$(\dagger) \qquad \mathcal{C}_G^*(X; R) := \{\phi \in \mathrm{C}^*(X; R) \mid \phi \text{ is } G \text{ equivariant}\}.$$

The equivariant Alexander–Spanier cochain complex is defined as

$$\mathrm{C}_G^*(X; R) := \mathcal{C}_G^*(X; R)/\mathcal{C}_G^*(X; R) \cap \mathrm{C}_0^*(X; R)$$

with the usual coboundary operator $d$. We will denote the cohomology of this complex by $\mathrm{H}_G^*(X; R)$.

The following theorem of Honkasalo relates the $G$-equivariant Alexander–Spanier cohomology with the Alexander–Spanier cohomology of the quotient by the $G$-action.

**Theorem 4.1** (see [9, Corollary 6.8]) *Let $R$ be an abelian group with the trivial $G$-action. There is a natural isomorphism*

$$\mathrm{H}_G^*(X; R) \cong \mathrm{H}^*(X/G; R),$$

*where the right-hand side is the ordinary Alexander–Spanier cohomology of $X/G$.*

**Remark 4.2** To define $\mathrm{C}_G^*$ in general, one needs a contravariant coefficient system — a contravariant functor from the category of $G$-spaces $G/H$ ($H \leq G$) and $G$-maps between them to the category of abelian groups. If $R$ is an abelian group with a $G$-action, it defines a contravariant coefficient system $G/H \mapsto R$, each $G$-map $G/H \to G/K$ inducing the identity $R \to R$. With this coefficient system, the equivariant Alexander–Spanier cochain complex takes the form of $\mathrm{C}_G^*(X; R)$ discussed above.

## Relation between $\mathrm{H}\,\mathrm{B}_G^*$ and $\mathrm{H}_G^*$

To relate the equivariant boundedly supported cohomology with the equivariant Alexander–Spanier cohomology, we will use the cohomology of the cochain complex $\mathcal{C}_G^*(X; R)$ as in †. We will denote this cohomology by $\mathcal{H}_G^*(X; R)$. We start with the following lemma.

**Lemma 4.3** *Suppose $G \curvearrowright X$ and $\mathrm{Fix}_G(X) \neq \varnothing$. Then*

$$\mathcal{H}_G^*(X; R) = \begin{cases} R & \text{if } * = 0, \\ 0 & \text{otherwise.} \end{cases}$$

**Proof** The proof is basically the equivariant version of Lemma 2.1. $\mathcal{H}_G^0(X; R)$ consists of all the constant functions from $X$ to $R$, and hence $\mathcal{H}_G^0(X; R) = R$.

To prove that the cohomology is trivial in degree $\geq 1$, choose $a \in \mathrm{Fix}_G(X)$. Consider the following $G$-equivariant cone operator $D_*: \mathrm{C}_*(X) \to \mathrm{C}_{*+1}(X)$:

$$D_n: (x_0, x_1, \ldots, x_n) \mapsto (a, x_0, \ldots, x_n).$$

Proceeding similar to the proof of Lemma 2.1, we obtain that $dD^*(\phi) = \phi$ when $\phi$ is a cocycle where $D^*$ is the dual of $D_*$. Since $D_*$ is $G$-equivariant, we have $D^*(\phi) \in \mathcal{C}_G^{*+1}(X; R)$. Hence, $\mathcal{H}_G^*(X) = 0$ for $* \geq 1$. $\qquad\square$

We now state our main proposition that relates $HB_G^*$ and $\tilde{H}_G^*$ under the assumption that $\mathrm{Fix}_G(X) \neq \varnothing$.

**Proposition 4.4** *Suppose $G \curvearrowright (X, A)$ and $\mathrm{Fix}_G(X) \neq \varnothing$.*

(1) *If $X \overset{c}{\subset} A$, then*
$$HB_G^*(X - A; R) = \begin{cases} R & \text{if } * = 0, \\ 0 & \text{otherwise.} \end{cases}$$

(2) *If $X$ is not coarsely contained in $A$, then*
$$HB_G^*(X - A; R) = \begin{cases} 0 & \text{if } * = 0, \\ \varinjlim \tilde{H}_G^{*-1}(X - N_r(A); R) & \text{otherwise.} \end{cases}$$

**Proof** (1) If $X \overset{c}{\subset} A$, then $CB_G^*(X - A; R)$ contain all $G$-equivariant cochains on $X$, in other words, $CB_G^*(X - A; R) = \mathcal{C}_G^*(X; R)$. The claim now follows from Lemma 4.3.

(2) Elements in $HB_G^0(X - A; R)$ are constant functions on $X$ with support contained in $N_r(A)$ for some $r$. Therefore, $HB_G^0(X - A; R) = 0$ if $X$ is not coarsely contained in $A$.

To calculate $HB_G^*(X - A; R)$ for $* \geq 1$, we first observe that we have a short exact sequence of $G$-equivariant cochain complexes,

$(\star)$ $\qquad 0 \to CB_G^*(X - A; R) \overset{j}{\longrightarrow} \mathcal{C}_G^*(X; R) \overset{i}{\longrightarrow} \varinjlim C_G^*(X - N_r(A); R) \to 0,$

where $j$ is the inclusion map and $i$ is induced by the composition of, for each $r$, the canonical maps

$$\mathcal{C}_G^*(X; R) \xrightarrow{\text{restriction}} \mathcal{C}_G^*(X - N_r(A); R) \xrightarrow{\text{quotient}} C_G^*(X - N_r(A); R).$$

To prove $(\star)$ is a short exact sequence, we observe

$$\begin{aligned} \ker(i) &= \{\phi \in \mathcal{C}_G^*(X; R) \mid \phi \in C_0^*(X - N_r(A); R) \text{ for some r}\} \\ &= \{\phi \in \mathcal{C}_G^*(X; R) \mid |\phi| \cap \Delta \subset N_r(\Delta_A) \text{ for some } r\} \\ &= \{\phi \in \mathcal{C}_G^*(X; R) \mid \|\phi\| \overset{c}{\subset} \Delta_A\} \\ &= \mathrm{im}(j). \end{aligned}$$

The short exact sequence $(\star)$ induces a long exact sequence of corresponding reduced cohomologies. The reduced cohomology of the middle cochain complex $\mathcal{C}_G^*(X; R)$ is trivial in all degrees by Lemma 4.3. Hence, the long exact sequence implies that

$$HB_G^*(X - A; R) \cong \varinjlim \tilde{H}_G^{*-1}(X - N_r(A); R) \quad \text{for } * \geq 1. \qquad \square$$

Combining Proposition 4.4 and Theorem 4.1 we get the following.

**Corollary 4.5** *Suppose $X$ is not coarsely contained in $A$ for some $A \subset X$, $G \curvearrowright (X, A)$ and $\mathrm{Fix}_G(X) \neq \varnothing$. Then*
$$HB_G^*(X - A; R) = \begin{cases} 0 & \text{if } * = 0, \\ \varinjlim \tilde{H}^{*-1}\big((X - N_r(A))/G; R\big) & \text{otherwise.} \end{cases}$$

## Relation between $HX_G^*$ and $HB_G^*$

The main reason for defining $HB_G^*(X - A; R)$ is the following theorem.

**Theorem 4.6** *Suppose $G \curvearrowright (X, A)$ and $R$ is an abelian group with a $G$-action. Suppose $X$ is uniformly acyclic away from $A$, and is locally acyclic away from $A$ with coefficients in $R$. Then the inclusion $CX_G^*(X - A; R) \hookrightarrow CB_G^*(X - A; R)$ induces an isomorphism on the cohomology*

$$HX_G^*(X - A; R) \cong HB_G^*(X - A; R).$$

The proof of the above theorem follows a sketch provided by Roe in [14], which outlines a proof of a related result connecting coarse cohomology with compactly supported cohomology. The core idea of the proof involves a standard "connect the dots" construction, which allows us to attach a singular chain to each simplex while remaining close to it, thanks to the uniform acyclicity of the space. We will state this more precisely in Lemma 4.8.

Let us first fix some notation. For the rest of the section, we suppress the coefficient $R$ from the notation unless it is important. Let $C_*^s(X)$ be the singular chain complex on $X$.

For the upcoming lemmas, we need to measure the distance between the support of a singular simplex and an $n$-simplex. Support of a singular simplex is in $X$ while an $n$-simplex is in $X^{n+1}$. However, recall from Section 2 that $X^n$ can be realized as a subset of $X^\infty$ for any $n \in \mathbb{N}$ by stabilizing the last coordinate. This way both the support of a singular chain, and more generally any subset of $X$ and $n$-simplices live in $X^\infty$. Consequently, we can measure the distance between the support of a singular simplex and an $n$-simplex in $X^\infty$.

In what follows, we will need a bound on the distance between a simplex and its boundary. This is the purpose of the following lemma.

**Lemma 4.7** *We have $\max_{\tau \in |\partial\sigma|} d(\sigma, \tau) \leq \mathrm{diam}(\sigma)$ for any simplex $\sigma$.*

**Proof** Let $\sigma = (x_0, x_1, \ldots, x_n)$ and $\tau \in |\partial\sigma|$. The stabilization of both $\sigma$ and $\tau$ in $X^\infty$ has $x_i$ in each coordinate for some $i$. It follows that

$$d(\sigma, \tau) \leq \max_{i \neq j}\{|x_i, x_j|\} = \mathrm{diam}(\sigma). \qquad \square$$

**Lemma 4.8** *Suppose $X$ is uniformly acyclic away from $A$ and locally acyclic away from $A$ and $G \curvearrowright (X, A)$. Then there exist two nondecreasing sequences of functions $\mu_n, \rho_n \colon [0, \infty) \to [0, \infty)$ and a $G$-equivariant chain map $M \colon C_*^F(X) \to C_*^s(X)$ where*

$$C_n^F(X) = \langle \sigma^n \mid d(\sigma^n, A) \geq \mu_n(\mathrm{diam}\,\sigma^n) \rangle \subset C_n(X)$$

*such that*:

(1) $|M(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$.

(2) *There exists an $L \geq 0$ such that for every $x \in X - N_L(A)$, and a neighborhood $U$ of $x$, there is a neighborhood $W \subset U$ of $x$ such that $|M(\sigma^n)| \subset U$ for all $\sigma^n \in W^{n+1}$.*

**Proof** We will define $\mu_i$, $\rho_i$, and the chain map $M : C_i^F \to C_i^s$ by induction on $i$.

For $i = 0$, define $\mu_0$ and $\rho_0$ to be the constant functions $[0, \infty) \to [0, \infty)$ with image $\{0\}$. That means, $C_0^F(X) = C_0(X)$ and we define $M : C_0^F(X) \to C_0^s(X)$ to be the identity map.

Since $X$ is uniformly acyclic away from $A$, there exist $\rho, \mu : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that the map $i_* : H_*(B) \to H_*(N_{\rho(\mathrm{diam}(B))}(B))$ is trivial if $d(B, A) \geq \mu(\mathrm{diam}(B))$. By the induction hypothesis, suppose $\mu_i$, $\rho_i$, and $M : C_i^F \to C_i^s$ are already defined with the desired properties for all $i \leq n$. For $i = n + 1$, define

$$(4\text{-}1) \qquad \mu_{n+1}(x) = \max\{\mu[2\rho_n(x) + x] + \rho_n(x) + x, \mu_n(x)\}.$$

By construction, $\mu_{n+1} \geq \mu_n$.

Next, we define $M$ on $C_{n+1}^F(X)$. Take $\sigma \in C_{n+1}^F(X)$ with $\mathrm{diam}(\sigma) = r$. By the induction hypothesis, $M$ satisfies property (1) when applied to $\partial \sigma$ and therefore

$$(4\text{-}2) \qquad |M(\partial\sigma)|) \leq N_{\rho_n(r)}(|\partial\sigma|).$$

Consequently we obtain

$$(4\text{-}3) \qquad d\big(\sigma, |M(\partial\sigma)|\big) \leq d\big(|\partial\sigma|, |M(\partial\sigma)|\big) + d(\sigma, |\partial\sigma|) \leq \rho_n(r) + r,$$

where the first inequality is due to the triangle inequality and the second inequality follows from (4-2) and Lemma 4.7.

Since $\sigma \in C_{n+1}^F(X)$, we have $d(\sigma, A) \geq \mu_{n+1}(r)$. It follows that

$$\begin{aligned}
d\big(|M(\partial\sigma)|, A\big) &\geq d(\sigma, A) - d\big(\sigma, |M(\partial\sigma)|\big) &&\text{(by the triangle inequality)} \\
&\geq \mu_{n+1}(r) - \rho_n(r) - r &&\text{(by (4-3))} \\
&\geq \mu(2\rho_n(r) + r) + \rho_n(r) + r - \rho_n(r) - r &&\text{(by (4-1))} \\
&= \mu(2\rho_n(r) + r).
\end{aligned}$$

Note that, $|M(\partial\sigma)| \subset N_{\rho_n(r)}(|\partial\sigma|)$ and $\mathrm{diam}(|\partial d\sigma|) \leq r$ implies that

$$\mathrm{diam}\big(|M(\partial\sigma)|\big) \leq 2\rho_n(r) + r.$$

Since $d\big(|M(\partial\sigma)|, A\big) \geq \mu(2\rho_n(r) + r)$, and $X$ is $(\mu, \rho)$-uniformly acyclic away from $A$, it follows that $M(\partial\sigma)$ is a boundary of some singular chain of diameter at most $\rho[2\mu_n(r) + r]$. Let

$$k(\sigma) := \inf\{\mathrm{diam}(c) \mid \partial c = M(\partial\sigma)\}.$$

We define $M(\sigma)$ to be a singular chain whose boundary is $M(\partial\sigma)$ and has diameter at most $2k(\sigma)$. To make $M$ a $G$-equivariant chain, we can first define $M$ on a set of simplices from $C_{n+1}^F(M)$ that contains one element from each orbit of simplices under the action of $G$ and then extend the map $G$-equivariantly.

Next we define $\rho_{n+1}$, so that $M$ satisfies (1). Let $\sigma$ be as before. By construction, $\partial M(\sigma) = M(\partial\sigma)$ and $k(\sigma) \leq \rho[2\mu_n(r) + r]$. Therefore we have

$$|M(\sigma)| \subset N_{2k(\sigma)}|M(\partial\sigma)| \subset N_{2k(\sigma)+\rho_n(r)}(|\partial\sigma|) \qquad \text{(by (4-2))}$$
$$\subset N_{2k(\sigma)+\rho_n(r)+r}(\sigma) \qquad \text{(by Lemma 4.7)}$$
$$\subset N_{2\rho[(2\mu_n(r)+r)]+\rho_n(r)+r}(\sigma).$$

Finally if we define

$$\rho_{n+1}(x) := 2\rho[2\mu_n(x) + x] + \rho_n(x) + x,$$

then by construction

$$M(\sigma) \subset N_{\rho_{n+1}(r)}(\sigma) = N_{\rho_{n+1}(\mathrm{diam}(\sigma))}(\sigma),$$

which is the desired property (1).

To prove (2), recall that $X$ is locally acyclic away from $A$ and hence there exists some $L > 0$ such that $X - N_L(A)$ is locally acyclic. Recall that $k(\sigma)$ is the infimum of the diameter of chains bounding $M(\partial\sigma)$. By induction on the dimension of $\sigma$, we observe that $k(\sigma)$ goes to zero as $\mathrm{diam}(\sigma)$ goes to 0, given that $d(\sigma, A) \geq L$ because of the local acyclicity of $X - N_L(A)$. By construction, $M(\sigma)$ is of diameter at most $2k(\sigma)$. Therefore $\mathrm{diam}(|M(\sigma)|)$ goes to zero as $\mathrm{diam}(\sigma)$ goes to 0, given that $d(\sigma, A) \geq L$. This gives us (2). $\qquad\square$

**Lemma 4.9** *Assume that $A \subset X$ and $G \curvearrowright (X, A)$. Let $f : \mathrm{C}_*(X) \to \mathrm{C}_*(X)$ be a $G$-equivariant chain map and $\rho_n : [0, \infty) \to [0, \infty)$ is some nondecreasing sequence of nondecreasing functions such that:*

- *$|f(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$ for any $n$-simplex $\sigma^n$.*

- *There exists an $L \geq 0$ such that for every $x \in X - N_L(A)$, and a neighborhood $U$ of $x$, there is a neighborhood $W \subset U$ of $x$ such that $|f(\sigma^n)| \subset U^{n+1}$ whenever $\sigma^n \in W^{n+1}$.*

*Then $f$ and the identity map on $\mathrm{C}_*(X)$ are chain homotopic via a $G$-equivariant chain homotopy*

$$H_n : \mathrm{C}_n(X) \to \mathrm{C}_{n+1}(X)$$

*with an associated nondecreasing sequence of nondecreasing functions $\rho'_n : [0, \infty) \to [0, \infty)$ such that:*

(1) *$|H_n(\sigma)| \subset N_{\rho'_n(\mathrm{diam}(\sigma))}(\sigma)$ for any $n$-simplex $\sigma^n$.*

(2) *There exists an $L' \geq 0$ such that for every $x \in X - N_{L'}(A)$, and a neighborhood $U$ of $x$, there is a neighborhood $W \subset U$ of $x$ such that $|H_n(\sigma^n)| \subset U^{n+2}$ for all $\sigma^n \in W^{n+1}$.*

**Proof** We can define $H_n$ by induction on the dimension $n$. Define $H_0(x) := (x, f(x))$. Note that $H_0$ is $G$-equivariant and $f(x) - x = \partial H_0(x)$. Since $f(x) \subset N_{\rho_0(\mathrm{diam}(x))}(x)$, we get $H_0(x) \subset N_{\rho_0(\mathrm{diam}\{x\})}(x)$. We define $\rho'_0 := \rho_0$.

Suppose we have already defined $H_m \colon C_m(X) \to C_{m+1}(X)$ and the nondecreasing map $\rho'_m$ such that $H_m$ is $G$-equivariant, $H_m(\sigma) \subset N_{\rho'_m(\operatorname{diam}(\sigma))}(\sigma)$ and $\partial H_m(\sigma) + D_{m-1}\partial(\sigma) = i(\sigma) - f(\sigma)$ for any $m \le n$. To define $H_{n+1}(\sigma)$ for an $(n+1)$-simplex $\sigma$, consider the chain

$$c_\sigma := \sigma - f(\sigma) - H_n\partial(\sigma).$$

By induction hypothesis on $H_n$, $c$ is a cycle. Take a vertex $b$ from the chain $c$, and consider the cone operator

$$T_b \colon C_{n+1}(X) \to C_{n+2}(X), \quad (x_0, \ldots, x_{n+1}) \mapsto (b, x_0, \ldots, x_{n+1}).$$

Define $H_{n+1}(\sigma) := T_b(c_\sigma)$. It follows that

$$\partial H_{n+1}(\sigma) = c_\sigma - T_b(\partial c_\sigma) = c_\sigma = \sigma - f(\sigma) - H_n\partial(\sigma).$$

To make $H_{n+1}$ equivariant, we first define it on elements from each orbit class of $(n+1)$-simplices and then extend it $G$-equivariantly.

We now focus on the support of $H_{n+1}(\sigma)$. Suppose $\operatorname{diam}(\sigma) = r$. By the induction hypothesis, $H_n(\tau) \subset N_{\rho'_n}(\tau)$ for any $n$-simplex $\tau$. Since $\rho'_n$ is a nondecreasing function, $\rho'_n(r) \ge \rho'_n(\operatorname{diam}(\tau))$ for any $\tau \in |\partial\sigma|$. We have

$$|H_n(\partial\sigma)| \subset \bigcup_{\tau \in |\partial\sigma|} N_{\rho'_n(\operatorname{diam}(\tau))}(\tau) \subset N_{\rho'_n(r)}(|\partial\sigma|) \subset N_{[\rho'_n(r)+r]}(\sigma),$$

where the last containment is by Lemma 4.7. Also, $\sigma$ and $|f(\sigma)|$ are subsets of $N_{\rho_{n+1}(r)}(\sigma)$. If we define

$$q(x) := \max\{\rho'_n(x) + x, \rho_{n+1}(x)\}$$

then $|c_\sigma| \subset N_{q(r)}(\sigma)$ and consequently

$$\operatorname{diam}(|c_\sigma|) \le \operatorname{diam}(\sigma) + 2q(r) \le r + 2L(r).$$

Since $|T_b(c_\sigma)| \subset N_{\operatorname{diam}(|c_\sigma|)}(|c_\sigma|)$, we obtain

$$
\begin{aligned}
|H_{n+1}(\sigma)| = |T_b(c_\sigma)| &\subset N_{\operatorname{diam}(|c_\sigma|)}(|c_\sigma|) \\
&\subset N_{\operatorname{diam}(|c_\sigma|)}[N_{q(r)}(\sigma)] \\
&\subset N_{[\operatorname{diam}(|c_\sigma|)+q(r)]}(\sigma) \subset N_{[r+3q(r)]}(\sigma).
\end{aligned}
$$

Letting $\rho'_{n+1}(r) := r + 3q(r)$, we get property (1).

To get property (2), note that vertex of any simplex in $|H_n(\sigma)|$ is either a vertex of $\sigma$ or a vertex of some simplices in $|f(\tau)|$ where $\tau$ is some subsimplex of $\sigma$. The claim then follows from the analogous property of the map $f$. $\qquad\square$

Let $\mathcal{U}$ denotes an open cover $X$. We say $\mathcal{U}$ is $G$-invariant, if for any $U \in \mathcal{U}$, $gU \in \mathcal{U}$ for all $g \in G$. Let $C_*^{\mathcal{U}}(X)$ be the chain complex generated by singular simplices supported in some open set in $\mathcal{U}$. Let $V \colon C_*^s(X) \to C_*(X)$ be the forgetful map, which maps a singular simplex to its vertices.

To prove Theorem 4.6, we will need to fill in simplices by singular chains in $C_*^{\mathcal{U}}(X)$ for some $G$-invariant cover $\mathcal{U}$. More precisely, we need the following.

**Proposition 4.10** *Suppose $X$ is uniformly acyclic away from $A$ and locally acyclic away from $A$ and $G \curvearrowright (X, A)$. Let $\mathcal{U}$ be a $G$-invariant open cover of $X$. Then there exist two nondecreasing sequences of functions $\mu_n, \rho_n \colon [0, \infty) \to [0, \infty)$, a $G$-equivariant chain map $S \colon C_*^F(X) \to C_*^{\mathcal{U}}(X)$ where*

$$C_n^F(X) = \langle \sigma^n \mid d(\sigma^n, A) \geq \mu_n(\operatorname{diam} \sigma^n) \rangle \subset C_n(X)$$

*and a $G$-equivariant chain homotopy $H \colon C_*^F(X) \to C_{*+1}(X)$ between $VS \colon C_*^F(X) \to C_*(X)$ and the inclusion map so that*:

(1) $|VS(\sigma^n)| \subset N_{\rho_n(\operatorname{diam}(\sigma^n))}(\sigma^n)$ *for any $n$-simplex $\sigma^n$.*

(2) $|H(\sigma^n)| \subset N_{\rho_n(\operatorname{diam}(\sigma^n))}(\sigma^n)$ *for any $n$-simplex $\sigma^n$.*

(3) *There exists $r > 0$ so that for every $x \notin N_r(A)$, there is a neighborhood $W$ of $x$ such that for all $\sigma^n \in W^{n+1}$, we have $H(\sigma^n) \in C_{n+1}^{\mathcal{U}}(X)$.*

**Proof** Because of the assumptions on $X$, we can invoke Lemma 4.8, which outputs a $G$-equivariant map $M \colon C_*^F(X) \to C_*^s(X)$ satisfying (1) and (2) from Lemma 4.8.

Next, we choose a barycentric subdivision map $P \colon C_*^s(X) \to C_*^{\mathcal{U}}(X)$. We can choose $P$ in a $G$-equivariant way by the same trick as before: in each dimension, first define it on an element from each orbit class of simplices and then extend $G$-equivariantly. Note that $P$ satisfies

$$(*) \qquad\qquad |P(\sigma)| \subset N_{\operatorname{diam}(\sigma)}(|\sigma|) \quad \text{for any } \sigma.$$

We define $S := PM$. Since both $P$ and $M$ are $G$-equivariant, so is $S$.

Since $M$ satisfies Lemma 4.8(1) and $P$ satisfies $(*)$, we obtain that

$$|VS(\sigma^n)| \subset N_{\rho_n(\operatorname{diam}(\sigma^n))}(\sigma^n)$$

for some function nondecreasing sequence of function $\rho_n' \colon [0, \infty) \to [0, \infty)$.

Moreover, since $M$ satisfies (2) from Lemma 4.8 and $P$ satisfies $(*)$, there exists an $L \geq 0$ such that for every $x \in X - N_L(A)$, and a neighborhood $U$ of $x$, there is a neighborhood $W \subset U$ of $x$ such that $|VS(\sigma^n)| \subset U^{n+1}$ whenever $\sigma^n \in W^{n+1}$.

In conclusion, $VS$ satisfies the hypothesis of Lemma 4.9. Applying Lemma 4.9 to the map $VS$ we obtain a chain homotopy $H_*$ between $VS$ and the $id$ that satisfies properties (2) and (3). $\qquad\square$

**Remark 4.11** If $X$ is uniformly acyclic, then we can take $C_*^F(X)$ to be $C_*(X)$ in Proposition 4.10. Also, the local acyclicity away from $A$ was only used to ensure property (3). So, if the space is just uniformly acyclic away from $A$, we still get $H$ and $S$ satisfying the properties (1) and (2).

**Proof of Theorem 4.6** Consider the long exact sequence

$$\cdots \to \mathrm{H}\,\mathrm{B}_G^{*-1}(X-A) \to \mathrm{H}^*(\mathrm{C}\,\mathrm{B}_G^*(X-A)/\mathrm{C}\,\mathrm{X}_G^*(X-A)) \to \mathrm{H}\,\mathrm{X}_G^*(X-A) \to \mathrm{H}\,\mathrm{B}_G^*(X-A) \to .$$

Therefore it is enough to show that

$$\mathrm{H}^*(\mathrm{C}\,\mathrm{B}_G^*(X-A)/\mathrm{C}\,\mathrm{X}_G^*(X-A)) = 0.$$

In other words, for $\phi \in \mathrm{C}\,\mathrm{B}_G^n(X-A)$ with $d\phi \in \mathrm{C}\,\mathrm{X}_G^{n+1}(X-A)$ we need to find $\psi \in \mathrm{C}\,\mathrm{B}_G^{n-1}(X-A)$ so that $\phi - d\psi \in \mathrm{C}\,\mathrm{X}_G^n(X-A)$.

Our goal is to apply Proposition 4.10. In order to do that, we first need to choose a $G$-invariant cover of $X$. Let $r > 0$ and let $U$ be union of all the balls of radius $r$ that are centered at points in $A \cap \|\phi\|$. Since both $A$ and $\|\phi\|$ are $G$-invariant, $U$ is $G$-invariant: $gU = U$ for all $g \in G$. For each $x \in X - \|\phi\|$, choose a metric neighborhood $U_x$ with diameter $\leq 1$ such that $U_x^{n+1} \cap |\phi| = \varnothing$. Since $\phi$ is $G$-equivariant, we can choose the association $x \mapsto U_x$ so that $U_{gx} = gU_x$ for all $g \in G$. Let $\mathcal{U}$ denote the collection of $U_x$ together with $U$. By construction, this is a $G$-invariant cover.

Now we apply Proposition 4.10 to this setup which outputs a complex $\mathrm{C}_*^F(X)$, a $G$-equivariant chain map $S : \mathrm{C}_*^F(X) \to \mathrm{C}_*^{\mathcal{U}}(X)$ and a $G$-equivariant chain homotopy $H : \mathrm{C}_*^F(X) \to \mathrm{C}_{*+1}(X)$ between $VS$ and the inclusion map that satisfy properties (1), (2), and (3) from Proposition 4.10.

We define a linear map $D : \mathrm{C}_*(X) \to \mathrm{C}_{*+1}(X)$ by setting

$$D(\sigma^n) = \begin{cases} H(\sigma^n) & \text{if } \sigma^n \in \mathrm{C}_n^F(X), \\ 0 & \text{otherwise.} \end{cases}$$

We define $\tau : \mathrm{C}_*(X) \to \mathrm{C}_{*+1}(X)$ as

$$\tau = id + \partial D + D\partial.$$

If $\sigma \in \mathrm{C}_*^F(X)$, then by construction $D(\sigma) = H(\sigma)$ and therefore $\tau(\sigma) = VS(\sigma)$. Suppose $\tau^*$ denote the dual of $\tau$. By applying $\tau^*$ on $\phi$, we get

$$\tau^*\phi = \phi + dD^*\phi + D^*d\phi.$$

We claim that $\tau^*\phi \in \mathrm{C}\,\mathrm{X}_G^n(X-A)$. If $\mathrm{diam}(\sigma^n) \leq k$ and $d(\sigma^n, A) > \mu_n(k)$, then $\sigma^n \in \mathrm{C}_n^F(X)$ and hence $\tau(\sigma^n) = VS(\sigma^n)$. Moreover, if $\sigma^n$ is outside of the $\rho_n(k)$-neighborhood of $U^{n+1}$, then $|\tau(\sigma^n)|$ does not touch $U^{n+1}$ because $|\tau(\sigma^n)| = |VS\sigma^n| \subset N_{\rho_n(k)}(\sigma^n)$ by property (1) from Proposition 4.10. This implies $(\tau^*\phi)(\sigma^n) = \phi(\tau(\sigma^n)) = 0$. In other words, $\sigma^n \notin |\tau^*\phi|$. It follows that

$$|\tau^*\phi| \cap N_k(\Delta) \subset N_{\mu_n(k)}(\Delta_A) \cup N_{\rho_n(k)}(U^{n+1}).$$

Since $U \overset{c}{\subset} A$, we have

$$|\tau^*\phi| \cap N_k(\Delta) \overset{c}{\subset} \Delta_A. \qquad \square$$

Next, we claim that $D^*(\phi) \in \mathrm{C}\,\mathrm{B}_G^{n-1}(X-A)$. Since $H$ satisfies property (3) from Proposition 4.10, we can choose a set $N_r(A)$ containing $U$ such that for any $x \notin N_r(A)$, there is a neighborhood $W_x$ of $x$

so that $|H(\sigma^n)| \subset U_x^{n+2}$ for any $\sigma^n \in W_x^{n+1}$. Hence for any $x \notin N_r(A) \cup \|\phi\|$, we have $|H(\sigma^n)| \notin |\phi|$ and hence $|D(\sigma^n)| \notin |\phi|$ for all $\sigma^n \in W_x^{n+1}$. Therefore, $\|D^*(\phi)\| \subset \|\phi\| \cup N_r(A)$. The claim follows since $\phi \in C B^*(X)$.

Finally, we claim that $D^* d(\phi) \in C X_G^n(X - A)$. By construction of $D$, if $d(\sigma^n, A) \geq \mu_n(k)$ then $|D\sigma^n| \subset N_{\rho_n(k)}(\sigma^n)$ where $k$ is the diameter of $\sigma^n$. Since $d(\phi) \in C X_G^*(X - A)$, the claim follows.

Now setting $\psi = -D^*(\phi)$ we get what we want.

We can now prove Theorem 3.8.

**Proof of Theorem 3.8** We observe that $C X_G^*(X - A; R) = \mathcal{C}_G^*(X; R)$ when $X$ is coarsely contained in $A$. Hence Theorem 3.8(1) follows from Lemma 4.3.

Theorem 3.8(2) follows immediately from Proposition 4.4 and Theorem 4.6. $\qquad\square$

# 5 Coarse cohomology of the configuration space

Let $(X, d)$ be a metric space. Equip $X^2 = X \times X$ with the sup metric. Consider the $\mathbb{Z}_2$-action on $X^2$ that flips the coordinates. Since the fixed point set for this action is the diagonal subspace

$$\delta(X) = \{(x, x) \mid x \in X\} \subset X^2,$$

we have $\mathbb{Z}_2 \curvearrowright (X^2, \delta(X))$. Let $R$ be an abelian group with a $\mathbb{Z}_2$-action. We define *coarse cohomology of the two-points configuration space of $X$* with coefficients in $R$ to be the cohomology of the complex $C X_{\mathbb{Z}_2}^*(X^2 - \delta(X); R)$. From now on, for the sake of simplicity, we will omit the term "two-points" from our terminology and refer to it simply as the coarse cohomology of the configuration space.

Theorem 3.8 immediately gives us the following.

**Proposition 5.1** *If $X$ is unbounded, uniformly acyclic, and locally acyclic with coefficients in $R$ where $R$ is an abelian group with trivial $\mathbb{Z}_2$-action, then*

$$H X_{\mathbb{Z}_2}^*(X^2 - \delta(X); R) = \begin{cases} 0 & \text{if } * = 0, \\ \varinjlim \tilde{H}^{*-1}\big((X^2 - N_r(\delta(X)))/\mathbb{Z}_2; R\big) & \text{otherwise.} \end{cases}$$

**Example 5.2** $\mathbb{R}^n$ satisfies the hypothesis of Proposition 5.1. Moreover, for any $r$, there is a $\mathbb{Z}_2$-equivariant deformation retraction of $(\mathbb{R}^n)^2 - \delta(\mathbb{R}^n)$ to $(\mathbb{R}^n)^2 - N_r(\delta(\mathbb{R}^n))$. Therefore, applying Proposition 5.1 we obtain the following where the coefficients group is $\mathbb{Z}_2$ with the trivial $\mathbb{Z}_2$-action:

$$\begin{aligned} H X_{\mathbb{Z}_2}^*((\mathbb{R}^n)^2 - \delta(\mathbb{R}^n); \mathbb{Z}_2) &= \begin{cases} 0 & \text{if } * = 0, \\ \tilde{H}^{*-1}\big(((\mathbb{R}^n)^2 - \delta(\mathbb{R}^n))/\mathbb{Z}_2; \mathbb{Z}_2\big) & \text{otherwise,} \end{cases} \\ &= \begin{cases} 0 & \text{if } * = 0, \\ \tilde{H}^{*-1}(\mathbb{R}P^{n-1}; \mathbb{Z}_2) & \text{otherwise,} \end{cases} \\ &= \begin{cases} \mathbb{Z}_2 & \text{if } 2 \leq * \leq n, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

**Example 5.3** Recall that $\ell^\infty$ is the space of bounded sequences of real numbers with the sup-norm metric. The space $(\ell^\infty)^2 - \delta(\ell^\infty)$ $\mathbb{Z}_2$-equivariantly deformation retracts to $(\ell^\infty)^2 - N_r(\delta(\ell^\infty))$. Since $(\ell^\infty)^2 - \delta(\ell^\infty)$ is acyclic, $((\ell^\infty)^2 - \delta(\ell^\infty))/\mathbb{Z}_2$ is a classifying space for $\mathbb{Z}_2$ and hence is homotopy equivalent to $\mathbb{R}P^\infty$. Hence arguing as in Example 5.2 we obtain

$$\mathrm{H\,X}^*_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty); \mathbb{Z}_2) = \begin{cases} \mathbb{Z}_2 & \text{for } * \geq 2, \\ 0 & \text{otherwise.} \end{cases}$$

Next we describe the maps between two metric spaces that induce map between the corresponding coarse cohomology of the configuration spaces. In the topological setting, an injective continuous map induces a map between the cohomology of the corresponding configuration spaces. In the coarse setting, the role of injective continuous maps are played by *coarse expanding* maps which we define next.

**Definition 5.4** A map $f : (X, A) \to (Y, B)$ between pairs is called relatively proper if $f^{-1}(N_r(B)) \overset{\mathrm{c}}{\subset} A$ for any $r$.

A map $f : (X, A) \to (Y, B)$ between pairs is called relatively coarse if $f$ is relatively proper and there exists a nondecreasing function $\mu : [0, \infty) \to [0, \infty)$ such that $d(f(x), f(y)) \leq \mu(d(x, y))$ for all $x, y \in X$.

A map $f : X \to Y$ is a coarse expanding map if the induced map $(x, y) \mapsto (f(x), f(y))$ from $(X^2, \delta(X))$ to $(Y^2, \delta(Y))$ is a relatively coarse map.

**Example 5.5** Recall that a map $f : X \to Y$ between two metric spaces is said to be a *coarse embedding* if there exist two proper nondecreasing maps $\rho_-, \rho_+ : [0, \infty) \to [0, \infty)$ such that

$$\rho_-(d(x, y)) \leq d(f(x), f(y)) \leq \rho_+(d(x, y)) \quad \text{for all } x, y \in X.$$

One can see that any coarse embedding map is a coarse expanding map.

**Example 5.6** A map that is not coarse expanding is the map $x \mapsto |x|$ between real numbers. The reason is that the map $(x, y) \mapsto (|x|, |y|)$ from $(\mathbb{R}^2, \delta(\mathbb{R}))$ to itself is not relatively proper.

We now recall the following from [2].

**Lemma 5.7** *For any abelian group $R$, a relatively coarse map $f : (X, A) \to (Y, B)$ between pairs induces the chain map $f^* : \mathrm{C\,X}^*(Y - B; R) \to \mathrm{C\,X}^*(X - A; R)$ by the canonical formula*

$$(f^*\phi)(x_0, x_1, \ldots, x_n) = \phi(f(x_0), f(x_1), \ldots, f(x_n)).$$

The following is immediate from Lemma 5.7.

**Lemma 5.8** *If $f : X \to Y$ is a coarse expanding map and $R$ is an abelian group, then the map $(x, y) \mapsto (f(x), f(y))$ induces a map $f^* : \mathrm{H\,X}^*_{\mathbb{Z}_2}(Y^2 - \delta(Y); R) \to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X); R)$.*

# 6 Coarse van Kampen obstruction

In this section, we find an obstruction to the existence of coarse expanding maps between two metric spaces. Our key observation is the next proposition.

**Proposition 6.1** *Any two coarse expanding maps from the space $X$ to $\ell^\infty$ induce the same map from* $\mathrm{H\,X}^*_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty); R)$ *to* $\mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X); R)$.

**Proof**  For convenience, we will suppress the coefficient $R$ from the notation in the proof.

Suppose $f, g: X^2 \to (\ell^\infty)^2$ are two maps induced by two coarse expanding maps from $X$ to $\ell^\infty$. For convenience, we will denote the induced map between $\mathrm{C}_*(X^2)$ and $\mathrm{C}_*((\ell^\infty)^2)$ by $f$ and $g$ as well. Let $f^*$ and $g^*$ be the corresponding map from $\mathrm{C\,X}^*_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty))$ to $\mathrm{C\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X))$. We will show that there is a chain homotopy $\mathrm{C\,X}^*_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty)) \to \mathrm{C\,X}^{*+1}_{\mathbb{Z}_2}(X^2 - \delta(X))$ between $f^*$ and $g^*$.

Since $(\ell^\infty)^2$ is uniformly contractible, the proof of Proposition 4.10 (see Remark 4.11) gives us a $\mathbb{Z}_2$-equivariant chain map $S: \mathrm{C}_*((\ell^\infty)^2) \to \mathrm{C}^s_*((\ell^\infty)^2)$ such that $|S(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$ for some nondecreasing sequence of functions $\rho_n: [0, \infty) \to [0, \infty)$. Moreover, by Proposition 4.10, the composition $VS$ is chain homotopic to the identity map by a $\mathbb{Z}_2$-equivariant chain homotopy $H$ such that $|H(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$.

We now construct a chain homotopy $D_*: \mathrm{C}_*(X^2) \to \mathrm{C}^s_{*+1}((\ell^\infty)^2)$ between the two maps

$$S \circ f, S \circ g: \mathrm{C}_*(X^2) \to \mathrm{C}^s_*((\ell^\infty)^2)$$

with certain properties: in particular we want the homotopy to avoid $\delta(\ell^\infty)$.

For convenience let us fix some notation. For a singular chain $c \in \mathrm{C}^s_*(X)$, let

$$\mathrm{size}(c) := \sup_{\tau \in |c|} \{\mathrm{diam}(\tau)\}.$$

For a simplex $\sigma \in (X^2)^*$, let

$$r_\sigma := \min\{d(|S(f(\sigma))|, \delta(\ell^\infty)), d(|S(g(\sigma))|, \delta(\ell^\infty))\}$$

and

$$R_\sigma := \max\{d(|S(f(\sigma))|, \delta(\ell^\infty)), d(|S(g(\sigma))|, \delta(\ell^\infty))\}.$$

Let $C_\sigma \subset (\ell^\infty)^2$ be the following annulus around $\delta(\ell^\infty)$:

$$\{z \mid r_\sigma \le d(z, \delta(\ell^\infty)) \le R_\sigma\}.$$

We note that $C_\sigma$ is acyclic because it is homotopy equivalent to $S^\infty$.

We claim that there exists a $D_i: \mathrm{C}_i(X^2) \to \mathrm{C}^s_{i+1}((\ell^\infty)^2)$ such that for any $i$-simplex $\sigma$ the following hold.

(1)  $S \circ f(\sigma) - S \circ g(\sigma) = \partial D_i(\sigma) + D_{i-1}\partial(\sigma)$.

(2)  $|D_i(\sigma)| \subset C_\sigma^{i+2}$.

(3)  $\mathrm{size}(D_i(\sigma)) \leq \mathrm{size}\big(S(f(\sigma)) - S(g(\sigma)) - D_{i-1}\partial(\sigma)\big)$.

(4)  $D_i$ is $\mathbb{Z}_2$-equivariant.

We will defer the construction of $D_i$ to Lemma 6.2. Assuming the existence of such $D_i$, we now complete the proof of the proposition.

Let $V \colon C_*^s \to C_*$ be the map that sends a singular simplex to its vertices. Since $V\partial = \partial V$, applying $V$ on both sides of (1), we have

$$(**) \qquad\qquad VS(f - g) = \partial V D_i + V D_{i-1}\partial.$$

Recall that $VS$ is homotopic to the identity map with a $\mathbb{Z}_2$-equivariant chain homotopy $H$ such that $|H(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$. So, we have

$$(***) \qquad\qquad VS = id + \partial H_i + H_{i-1}\partial.$$

Combining $(**)$ and $(***)$, we obtain

$$f - g = \partial[V D_i + H_i(f - g)] + [V D_{i-1} + H_{i-1}(f - g)]\partial.$$

Dualizing the above we get $f^*$ and $g^*$ are chain homotopic via the cochain homotopy $(VD)^* + (f-g)^* H^*$. To complete the proof, we need to show that this cochain homotopy maps $C X_{\mathbb{Z}_2}^*((\ell^\infty)^2 - \delta(\ell^\infty))$ to $C X_{\mathbb{Z}_2}^{*-1}((\ell^\infty)^2 - \delta(\ell^\infty))$.

Since $|H(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$ and $H$ is $\mathbb{Z}_2$-equivariant, the dual $H^*$ maps $C X_{\mathbb{Z}_2}^*((\ell^\infty)^2 - \delta(\ell^\infty))$ to $C X_{\mathbb{Z}_2}^{*-1}((\ell^\infty)^2 - \delta(\ell^\infty))$. Consequently, $(f-g)^* H^*$ maps $C X_{\mathbb{Z}_2}^*((\ell^\infty)^2 - \delta(\ell^\infty))$ to $C X_{\mathbb{Z}_2}^{*-1}(X^2 - \delta(X))$. It is therefore enough to prove that $(VD)^*$ maps $C X_{\mathbb{Z}_2}^*((\ell^\infty)^2 - \delta(\ell^\infty))$ to $C X_{\mathbb{Z}_2}^{*-1}(X^2 - \delta(X))$. Let $\phi \in C X_{\mathbb{Z}_2}^*((\ell^\infty)^2 - \delta(\ell^\infty))$ and $\sigma \in |(VD)^*(\phi)| \cap N_r(\Delta)$ for some $r \geq 0$. Since $\sigma \in N_r(\Delta)$, and $f$ and $g$ are coarse expanding maps, we have $|S(f(\sigma))| \subset N_s(\Delta)$ and $|S(g(\sigma))| \subset N_s(\Delta)$ for some $s$ that depends only on $r$. Property (3) then implies that $|VD(\sigma)| \subset N_s(\Delta)$ for some $s$ that depends only on $r$. Since $\phi \in C X_{\mathbb{Z}_2}^*((\ell^\infty)^2 - \delta(\ell^\infty))$ and $\phi(VD(\sigma)) \neq 0$, it then follows that $|VD(\sigma)| \subset N_t(\Delta_{\delta(\ell^\infty)})$ where $t$ depends only on $s$ and hence depends only on $r$. It now follows from property (2) that $\sigma \in N_p(\Delta_{\delta(X)})$ for some $p$ that depends only on $t$ and hence only on $r$. Hence, we proved that for each $r \geq 0$, there exists $p \geq 0$ such that $|(VD)^*(\phi)| \cap N_r(\Delta) \subset N_p(\Delta_{\delta(X)})$. Finally, $(VD)^*(\phi)$ is $\mathbb{Z}_2$-equivariant by property (4). Hence, $(VD)^*(\phi) \in C X_{\mathbb{Z}_2}^{*+1}(X^2 - \delta(X))$. $\qquad\square$

**Lemma 6.2**  *Let $f$ and $g$ are two coarse expanding maps from $X$ to $\ell^\infty$. Let $S \colon C_*((\ell^\infty)^2) \to C_*^s((\ell^\infty)^2)$ be an $\mathbb{Z}_2$-equivariant chain map such that $|S(\sigma^n)| \subset N_{\rho_n(\mathrm{diam}(\sigma^n))}(\sigma^n)$ for some sequence of functions $\rho_n \colon [0, \infty) \to [0, \infty)$. Then, for each $i$, there exists $D_i \colon C_i(X^2) \to C_{i+1}^s((\ell^\infty)^2)$ such that for any $i$-simplex $\sigma$ the following hold.*

(1)  $S \circ f(\sigma) - S \circ g(\sigma) = \partial D_i(\sigma) + D_{i-1}\partial(\sigma).$

(2)  $|D_i(\sigma)| \subset C_\sigma^{i+2}.$

(3)  $\text{size}(D_i(\sigma)) \leq \text{size}\big(S(f(\sigma)) - S(g(\sigma)) - D_{i-1}\partial(\sigma)\big).$

(4)  $D_i$ is $\mathbb{Z}_2$-equivariant.

**Proof**  We first define $D_0 \colon C_0(X^2) \to C_1^s((\ell^\infty)^2)$. For a 0-simplex $\sigma \in X^2$, join $f(\sigma)$ and $g(\sigma)$ by a path $\gamma$ in $C_\sigma$. Such a path exists because the annulus $C_\sigma$ is contractible. Next, we subdivide $\gamma$ so that each subarc is of diameter $\leq 1$. We define $D_0(\sigma)$ to be this subdivided path. By construction, $D_0$ satisfies the first three desired properties. To make it $G$-equivariant, we first define $D_0$ on a simplex from each $G$-orbit of simplices and then move it $\mathbb{Z}_2$-equivariantly. It is straightforward to see that $D_0$ still satisfies the first property because $\partial$ commutes with the $G$-action. It satisfies the second property because $gC_\sigma = C_{g\sigma}$ for any $g \in \mathbb{Z}_2$. $D_0$ satisfies the third property because isometric action preserves size of chains.

Inductively, let us assume that $D_i \colon C_i(X^2) \to C_{i+1}^s((\ell^\infty)^2)$ is already defined for all $i \leq n$ with the desired properties. To define $D_{n+1}(\sigma)$, let $K = S(f(\sigma)) - S(g(\sigma)) - D_n(\partial(\sigma))$. By the induction hypothesis (1), $\partial K = D_{n-1}(\partial^2(\sigma)) = 0$ and hence $K$ is a cycle. By (2), $|K| \subset C_\sigma$. Since $C_\sigma$ is acyclic, there exists a singular chain $c$ supported in $C_\sigma$ such that $\partial c = K$. After applying the appropriate subdivision, we can make $c$ to satisfy $\text{size}(c) \leq \text{size}(K)$ without changing its boundary. Define $D_{n+1}(\sigma)$ to be that $c$. By construction, $D_{n+1}$ satisfies conditions (1), (2), and (3). To make $D_{n+1}$ satisfy (4), we can use the same trick as before: first define it on a simplex from each $\mathbb{Z}_2$-orbit of $(n+1)$-simplices and then extend equivariantly. For the similar reason as $D_0$, this $D_{n+1}$ has all the desired four properties.  □

**Coarse van Kampen obstruction class**  Let $X$ be a separable metric space and $g \colon X \to \ell^\infty$ be a coarse expanding map. Such a map exists because any separable metric space admits an isometric embedding into $\ell^\infty$ by the work of Fréchet ([7]). We consider the induced map

$$g^* \colon H X^n_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty); \mathbb{Z}_2) \to H X^n_{\mathbb{Z}_2}(X^2 - \delta(X); \mathbb{Z}_2).$$

Recall from Example 5.3 that $H X^n_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty); \mathbb{Z}_2) = \mathbb{Z}_2$ if $n \geq 2$. Let $e^n$ be the nontrivial element in $H X^n_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty))$ for $n \geq 2$. Proposition 6.1 implies that $g^*(e^n)$ depends only on the space $X$, not on $g$. We call the class $g^*(e^n)$ to be the $n^{th}$-*degree coarse van Kampen obstruction class* of $X$ and denote it by $\text{cvk}^n(X)$ where $n \geq 2$.

**Assumption**  From now on all our metric spaces will be separable. In the nonequivariant setting, our coefficient group will be $\mathbb{Z}_2$ and in the equivariant setting, the coefficient group will be $\mathbb{Z}_2$ with the trivial $\mathbb{Z}_2$-action unless stated otherwise. In those cases, we will omit the coefficient from the notation.

**Definition 6.3**  (coarse obstruction dimension)  The *coarse obstruction dimension* of a space $X$, denoted by $\text{cobdim}(X)$, is 0 if $X$ is bounded, is 1 if $\text{cvk}^n(X) = 0$ for all $n$, and otherwise it is the largest $n$ such that $\text{cvk}^n(X) \neq 0$.

Now we prove the main theorem of this section.

**Theorem 6.4** *If $X$ admits a coarse expanding map into $Y$, then* $\operatorname{cobdim}(X) \leq \operatorname{cobdim}(Y)$.

**Proof** If $\operatorname{cobdim}(X) = 0$, then there is nothing to prove.

If $\operatorname{cobdim}(X) = 1$, then $X$ is unbounded by definition. This means $Y$ is also unbounded and hence $\operatorname{cobdim}(Y) \geq 1$ by definition.

Suppose $\operatorname{cobdim}(X) = n \geq 2$. Let $g \colon Y \to \mathbb{R}^\infty$ be a coarse expanding map. Consider the composition

$$X \times X \xrightarrow{f} Y \times Y \xrightarrow{g} \ell^\infty \times \ell^\infty$$

By Lemma 5.8, the above maps induce, between coarse cohomology of the configuration spaces, the maps

$$\operatorname{H}\!\operatorname{X}^n_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty)) \xrightarrow{g^*} \operatorname{H}\!\operatorname{X}^n_{\mathbb{Z}_2}(Y^2 - \delta(Y)) \xrightarrow{f^*} \operatorname{H}\!\operatorname{X}^n_{\mathbb{Z}_2}(X^2 - \delta(X)).$$

Let $e^n \in \operatorname{H}\!\operatorname{X}^n_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty))$ be the generator. Then

$$\operatorname{cvk}^n(Y) = g^*(e^n) \quad \text{and} \quad \operatorname{cvk}^n(X) = f^* g^*(e^n)) = f^*(\operatorname{cvk}^n(Y)).$$

By assumption $\operatorname{cvk}^n(X) \neq 0$ and hence $\operatorname{cvk}^n(Y) \neq 0$. So, we get $\operatorname{cobdim}(Y) \geq n$. $\qquad\square$

Recall that $X$ and $Y$ are said to be *coarsely equivalent* if there exists a coarse embedding map $f : X \to Y$ such that $Y \overset{c}{\subset} f(X)$. One can observe that two coarsely equivalent spaces coarsely embed into each other. As a consequence, the above theorem immediately yields the following.

**Corollary 6.5** *If $X$ and $Y$ are coarsely equivalent, then* $\operatorname{cobdim}(X) = \operatorname{cobdim}(Y)$.

In Example 5.2, we saw that $\operatorname{H}\!\operatorname{X}^*_{\mathbb{Z}_2}((\mathbb{R}^n)^2 - \delta(\mathbb{R}^n)) = 0$ for all $* > n$. Hence $\operatorname{cobdim}(\mathbb{R}^n) \leq n$. Using Theorem 6.4, we obtain:

**Corollary 6.6** *If* $\operatorname{cobdim}(X) \geq n$, *then $X$ does not admit a coarse expanding map into* $\mathbb{R}^{n-1}$.

# 7 Relation to the classical van Kampen obstruction

Let us recall the classical van Kampen obstruction class. Let $X$ be any topological space. Any continuous embedding $f : X \hookrightarrow \mathbb{R}^\infty$ induces a map

$$f : \operatorname{H}^*((\mathbb{R}^\infty)^2 - \delta(\mathbb{R}^\infty)/\mathbb{Z}_2) \to \operatorname{H}^*((X^2 - \delta(X))/\mathbb{Z}_2).$$

This map depends only on $X$ because the quotient map

$$q : (\mathbb{R}^\infty)^2 - \delta(\mathbb{R}^\infty) \to ((\mathbb{R}^\infty)^2 - \delta(\mathbb{R}^\infty))/\mathbb{Z}_2$$

is a universal $\mathbb{Z}_2$-bundle. Let $\eta^* \in H^*((\mathbb{R}^\infty)^2 - \delta(\mathbb{R}^\infty)/\mathbb{Z}_2)$ be the generator. The cohomology class $f(\eta^*)$ is called the van Kampen obstruction class in degree $*$ and will be denoted by $vk^*(X)$. The quotient map $(\ell^\infty)^2 - \delta(\ell^\infty) \to ((\ell^\infty)^2 - \delta(\ell^\infty))/\mathbb{Z}_2$ can also be considered as a universal $\mathbb{Z}_2$-bundle because $(\ell^\infty)^2 - \delta(\ell^\infty)$ is contractible. Hence we can use $\ell^\infty$ instead of $\mathbb{R}^\infty$ to define $vk^*(X)$. We use this viewpoint in the next proposition.

**Proposition 7.1** *Let $X$ be a metric space. Suppose*

$$i : H^*((X^2 - \delta(X))/\mathbb{Z}_2) \to \varinjlim H^*((X^2 - N_r(\delta(X)))/\mathbb{Z}_2)$$

*is the map induced by inclusions $X^2 - N_r(\delta(X)) \hookrightarrow X^2 - \delta(X)$ for each $r > 0$. If $i(vk^{n-1}(X))$ is nontrivial for some $n \geq 2$, then $\mathrm{cvk}^n(X)$ is nontrivial.*

**Proof** Let $f : X \to \ell^\infty$ be an isometry. Fix $n \geq 2$. Suppose $i(vk^{n-1}(X)) \neq 0$. We consider the following commutative diagram where all the horizontal maps are canonically induced by $f$. We want to show that the top horizontal map is nontrivial:

$$
\begin{array}{ccc}
H\,X^n_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty)) & \xrightarrow{\quad f^* \quad} & H\,X^n_{\mathbb{Z}_2}(X^2 - \delta(X)) \\
\Big\downarrow{\scriptstyle\cong} & & \Big\downarrow \\
\varinjlim H^{n-1}(((\ell^\infty)^2 - N_r(\delta(\ell^\infty)))/\mathbb{Z}_2) & \xrightarrow{\ f^*\ } & \varinjlim H^{n-1}((X^2 - N_r(\delta(X)))/\mathbb{Z}_2) \\
\Big\uparrow & & {\scriptstyle i}\Big\uparrow \\
H^{n-1}(((\ell^\infty)^2 - \delta(\ell^\infty))/\mathbb{Z}_2) & \xrightarrow{\ f^*\ } & H^{n-1}((X^2 - \delta(X))/\mathbb{Z}_2)
\end{array}
$$

The top left isomorphism map is due to the Proposition 5.1.

The image of the generator of $H^{n-1}(((\ell^\infty)^2 - \delta(\ell^\infty))/\mathbb{Z}_2)$ under the bottom horizontal map is $vk^{n-1}(X)$ if $n \geq 2$. Since $i(vk^{n-1}(X)) \neq 0$, the commutativity of the above diagram implies that the middle horizontal map is nontrivial. Again using commutativity of the diagram, we conclude that the top horizontal map is nontrivial. $\qquad\square$

Suppose $X = K \times [0, \infty)/K \times \{0\}$ is the open cone on a finite simplicial complex $K$. A metric $d$ on $X$ is called *expanding* if for any two disjoint simplices $\sigma, \tau \in K$ and $S \geq 0$, there exists $r \geq 0$ such that $d(\sigma \times [r, \infty), \tau \times [r, \infty)) \geq S$.

**Proposition 7.2** *Let $X$ be an open cone on a finite simplicial complex $K$ and $X$ is equipped with an expanding metric. If $vk^{n-1}(X) \neq 0$ for some $n \geq 2$, then $\mathrm{cvk}^n(X) \neq 0$.*

**Proof** Since the metric on $X$ is expanding, any representative of $c \in H_*(\mathrm{Conf}(X))$ can be homotoped to a cycle that lives in $H^*((X^2 - N_r(\delta(X)))/\mathbb{Z}_2)$ for any $r$. Considering the dual, this implies that the restriction map $\tilde{H}^*((X^2 - \delta(X))/\mathbb{Z}_2) \to \varinjlim \tilde{H}^*((X^2 - N_r(\delta(X)))/\mathbb{Z}_2)$ sends $vk^{n-1}(X)$ to a nontrivial element for any $r \geq 0$. Hence, $i(vk^{n-1}(X))$ is nontrivial where $i$ is as in Proposition 7.1. Hence Proposition 7.1 implies $\mathrm{cvk}^{n+1}(X) \neq 0$. $\qquad\square$

**Example 7.3** (Bestvina–Kapovich–Kleiner obstruction)  One of the results in [3] can be stated as follows: if there is class $c \in H_n(\mathrm{Conf}(X))$ such that $vk^n(X)(c) \neq 0$, then $X$ with a proper, expanding metric cannot be coarsely embedded inside $\mathbb{R}^n$. One can see that Proposition 7.2 combined with Corollary 6.6 recovers Bestvina, Kapovich, and Kleiner's result.

**Remark 7.4**  Note that $vk^{n-1}(X) \neq 0$ does not necessarily imply $cvk^n(X) \neq 0$. For example, take $X$ to be a unit disk in $\mathbb{R}^2$, then $vk^1(X) \neq 0$. However, since $X$ is bounded, Proposition 4.4 implies $H X_{\mathbb{Z}_2}^2(X^2 - \delta(X)) = 0$, and therefore $cvk^2(X) = 0$.

# 8  A coarse Gysin sequence

Recall that coarse van Kampen obstruction class lives in $H X_{\mathbb{Z}_2}^*(X^2 - \delta(X))$. In this section, we relate $H X_{\mathbb{Z}_2}^*$ to $H X^*$ and apply that to compute $H X_{\mathbb{Z}_2}^*(X^2 - \delta(X))$ for certain $X$.

Suppose $\mathbb{Z}_2$ is acting on some metric space $X$ by isometries and $A$ is the subset of $X$ that is fixed by the action. We consider the exact sequence

$$0 \to C X_{\mathbb{Z}_2}^*(X - A) \xrightarrow{i} C X^*(X - A) \xrightarrow{p} C X_{\mathbb{Z}_2}^*(X - A) \xrightarrow{r} C^*(A) \to 0,$$

where $i$ is the inclusion map and $p(\phi): \sigma \mapsto \phi(\sigma) + \phi(g\sigma)$ where $g$ is the generator of $\mathbb{Z}_2$ and $r$ is the restriction map. The image of $p$ consists of those cochains in $C X_{\mathbb{Z}_2}^*(X - A)$ that send any simplex supported on $A$ to zero. It is easy to see that collection of such cochains gives a subcomplex of $C X_{\mathbb{Z}_2}^*(X - A)$. We denote this complex by $C X_{\mathbb{Z}_2}^*(X - A, A)$, and the corresponding cohomology by $H X_{\mathbb{Z}_2}^*(X - A, A)$. Hence, the above four-term short exact sequence splits into the two short exact sequences

$$0 \to C X_{\mathbb{Z}_2}^*(X - A) \xrightarrow{i} C X^*(X - A) \xrightarrow{p} C X_{\mathbb{Z}_2}^*(X - A, A) \to 0,$$

$$0 \to C X_{\mathbb{Z}_2}^*(X - A, A) \xrightarrow{i} C X_{\mathbb{Z}_2}^*(X - A) \xrightarrow{r} C^*(A) \to 0.$$

These two short exact sequences give us two long exact sequences in the cohomology which we record as our next lemma.

**Lemma 8.1**  *Let $X$ be a metric space and $\mathbb{Z}_2 \curvearrowright (X, A)$ such that $A$ is the fixed point set of the action. Then we have the two long exact sequences*

(8-1)    $\cdots \to H X_{\mathbb{Z}_2}^*(X - A) \to H X^*(X - A) \to H X_{\mathbb{Z}_2}^*(X - A, A) \to H X_{\mathbb{Z}_2}^{*+1}(X - A) \to \cdots,$

(8-2)    $\cdots \to H X_{\mathbb{Z}_2}^*(X - A, A) \to H X_{\mathbb{Z}_2}^*(X - A) \to H^*(C^*(A)) \to H X_{\mathbb{Z}_2}^{*+1}(X - A, A) \to \cdots.$

**Lemma 8.2**  *Suppose $A \subset X$ and $X$ is not coarsely contained in $A$. If $\mathbb{Z}_2 \curvearrowright (X, A)$ where $A$ is the fixed point set of the action, then*

$$H X_{\mathbb{Z}_2}^*(X - A, A) = \begin{cases} 0 & \text{if } * = 0, \\ H X_{\mathbb{Z}_2}^*(X - A) \oplus \mathbb{Z}_2 & \text{if } * = 1, \\ H X_{\mathbb{Z}_2}^*(X - A) & \text{if } * \geq 2. \end{cases}$$

**Proof** Since $X$ is not coarsely contained in $A$, we have

$$\mathrm{H\,X}^0_{\mathbb{Z}_2}(X - A, A) = \mathrm{H\,X}^0_{\mathbb{Z}_2}(X - A) = 0.$$

By Lemma 2.1, the homology of the complex $\mathrm{C}^*(A)$ vanishes everywhere except in the degree 0 where it is $\mathbb{Z}_2$. Combining this with the second long exact sequence of the Lemma 8.1, we obtain

$$(8\text{-}3) \quad 0 \to \mathbb{Z}_2 \to \mathrm{H\,X}^1_{\mathbb{Z}_2}(X - A, A) \to \mathrm{H\,X}^1_{\mathbb{Z}_2}(X - A) \to 0 \to$$
$$\cdots \to 0 \to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X - A, A) \to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X - A) \to 0 \cdots,$$

where $* \geq 2$.

Since the coefficient group is $\mathbb{Z}_2$, the first five terms give us a split short exact sequence. Hence $\mathrm{H\,X}^1_{\mathbb{Z}_2}(X - A, A) = \mathrm{H\,X}^1_{\mathbb{Z}_2}(X - A) \oplus \mathbb{Z}_2$. It also follows from (8-3) that $\mathrm{H\,X}^*(X - A, A) = \mathrm{H\,X}^*(X - A)$ for all $* \geq 2$. $\qquad\square$

Hence under the hypothesis of Lemma 8.2, we can rewrite the first long exact sequence of Lemma 8.1 as follows.

**Lemma 8.3** (coarse Gysin sequence) *Suppose* $\mathbb{Z}_2 \curvearrowright (X, A)$ *where* $A \subset X$ *is the fixed point set of the action. Moreover, assume that $A$ is unbounded and $X$ is not coarsely contained in $A$. Then there is a long exact sequence of the form*

$$0 \to \mathrm{H\,X}^0_{\mathbb{Z}_2}(X - A) \to \mathrm{H\,X}^0(X - A) \to 0$$
$$\to \mathrm{H\,X}^1_{\mathbb{Z}_2}(X - A) \to \mathrm{H\,X}^1(X - A) \to \mathrm{H\,X}^1_{\mathbb{Z}_2}(X - A) \oplus \mathbb{Z}_2 \to \cdots$$
$$\to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X - A) \to \mathrm{H\,X}^*(X - A) \to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X - A) \to \cdots,$$

*where* $* \geq 2$.

We are now ready to state our main result of this section.

**Proposition 8.4** *Suppose $X$ is unbounded and $\mathbb{Z}_2$ is acting on $X^2$ by permuting the coordinates. Then there is a long exact sequence of the form*

$$(8\text{-}4) \quad 0 \to \mathrm{H\,X}^0_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathrm{H\,X}^0(X^2 - \delta(X)) \to 0$$
$$\to \mathrm{H\,X}^1_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathrm{H\,X}^1(X^2 - \delta(X)) \to \mathrm{H\,X}^1_{\mathbb{Z}_2}(X^2 - \delta(X)) \oplus \mathbb{Z}_2 \to \cdots$$
$$\to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathrm{H\,X}^*(X^2 - \delta(X)) \to \mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \cdots,$$

*where* $* \geq 2$.

**Proof** By hypothesis, $\mathbb{Z}_2 \curvearrowright (X^2, \delta(X))$ and $\delta(X)$ is the set of fixed points of the action. Since $X$ is unbounded $X^2$ is not coarsely contained in $\delta(X)$. So, we can apply Lemma 8.3 on $(X^2, \delta(X))$, and the claim follows. $\qquad\square$

**Remark 8.5** If $X$ is unbounded and $\mathrm{H}X^1(X^2 - \delta(X)) = 0$, then it follows from Proposition 8.4 that $\mathrm{H}X^1_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$. As a consequence, we can write the beginning part of the coarse Gysin sequence (8-4) as

$$(8\text{-}5) \qquad 0 \to \mathbb{Z}_2 \to \mathrm{H}X^2_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathrm{H}X^2(X^2 - \delta(X)) \to \mathrm{H}X^2_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \cdots.$$

This observation will be useful for us in our next theorem where we compute $\mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X))$ when $\mathrm{H}X^*(X^2 - \delta(X))$ is concentrated in some degree.

**Theorem 8.6** *Suppose $X$ is a metric space such that, for some $n \geq 1$,*

$$\mathrm{H}X^*(X^2 - \delta(X)) = \begin{cases} \mathbb{Z}_2 & \text{if } * = n, \\ 0 & \text{otherwise.} \end{cases}$$

*Moreover, suppose that $\mathrm{H}X^i_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for some $i \geq n+1$. Then*

$$\mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = \begin{cases} \mathbb{Z}_2 & \text{if } n \geq 2 \text{ and } 2 \leq * \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

**Proof** Elements of $\mathrm{H}X^0(X^2 - \delta(X))$ and $\mathrm{H}X^0_{\mathbb{Z}_2}(X^2 - \delta(X))$ are constant functions on $X^2$ with support contained in a neighborhood of $\delta(X)$. Since $\mathrm{H}X^0(X^2 - \delta(X)) = 0$, we have $\mathrm{H}X^0_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$.

Next, we will show that $\mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ if $* \geq n+1$. Consider the following part of the coarse Gysin sequence where $* \geq 2$:

$$(8\text{-}6) \quad \to \mathrm{H}X^*(X^2 - \delta(X)) \to \mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathrm{H}X^{*+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathrm{H}X^{*+1}(X^2 - \delta(X)) \to.$$

Since $\mathrm{H}X^*(X^2 - \delta(X)) = 0$ for $* \geq n+1$, the middle map is an isomorphism for $* \geq n+1$. Hence,

$$\mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = \mathrm{H}X^{n+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) \quad \text{for all } * \geq n+1.$$

That means if $\mathrm{H}X^{n+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) \neq 0$ then $\mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X)) \neq 0$ for all $* \geq n+1$. By hypothesis, $\mathrm{H}X^i_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for some $i \geq n+1$. Therefore, $\mathrm{H}X^{n+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ and consequently,

$$\mathrm{H}X^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0 \quad \text{for all } * \geq n+1.$$

We divide the rest of the calculations into three cases.

**Case 1 ($n = 1$)** We only have to show that $\mathrm{H}X^1_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$. By the hypothesis,

$$\mathrm{H}X^1(X^2 - \delta(X)) = \mathbb{Z}_2 \quad \text{and} \quad \mathrm{H}X^2_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0.$$

Hence, we get the following from the coarse Gysin sequence:

$$0 \to \mathrm{H}X^1_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathbb{Z}_2 \to \mathrm{H}X^1_{\mathbb{Z}_2}(X^2 - \delta(X)) \oplus \mathbb{Z}_2 \to 0.$$

Since the third map is surjective, $\mathrm{H}X^1_{\mathbb{Z}_2}(X^2 - \delta(X))$ cannot have more than one element and hence it is trivial.

**Case 2 ($n = 2$)** By hypothesis, $HX^1(X^2 - \delta(X)) = 0$, Therefore, $HX^1_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ (see Remark 8.5). Furthermore, $HX^2(X^2 - \delta(X)) = \mathbb{Z}_2$ by hypothesis and we already showed $HX^3_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$. So a part of the sequence (8-5) takes the form

$$0 \to \mathbb{Z}_2 \to HX^2_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \mathbb{Z}_2 \to HX^2_{\mathbb{Z}_2}(X^2 - \delta(X)) \to 0.$$

So there is an injective map and a surjective map from $\mathbb{Z}_2$ into $HX^2_{\mathbb{Z}_2}(X^2 - \delta(X))$. It follows that $HX^2_{\mathbb{Z}_2}(X^2 - \delta(X)) = \mathbb{Z}_2$.

**Case 3 ($n > 2$)** In this case, $HX^1(X^2 - \delta(X)) = 0$. Remark 8.5 gives us $HX^1_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$.

Since $HX^2(X^2 - \delta(X)) = 0$, the beginning part of the sequence (8-5) takes the form

$$0 \to \mathbb{Z}_2 \to HX^2_{\mathbb{Z}_2}(X^2 - \delta(X)) \to 0 \to \cdots.$$

Hence $HX^2_{\mathbb{Z}_2}(X^2 - \delta(X)) = \mathbb{Z}_2$.

Since $HX^*(X^2 - \delta(X)) = 0$ for $* \leq n - 1$, it follows from (8-6) that

$$HX^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = HX^{*+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) \quad \text{when } 2 \leq * \leq n - 2.$$

That implies

$$HX^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = HX^2_{\mathbb{Z}_2}(X^2 - \delta(X)) = \mathbb{Z}_2 \quad \text{when } 2 \leq * \leq n - 1.$$

Finally to compute $HX^n_{\mathbb{Z}_2}(X^2 - \delta(X))$, consider the following part of the coarse Gysin sequence (Lemma 8.3):

$$\cdots \to HX^n_{\mathbb{Z}_2}(X^2 - \delta(X)) \to HX^n(X^2 - \delta(X)) \to HX^n_{\mathbb{Z}_2}(X^2 - \delta(X)) \to HX^{n+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) \to \cdots.$$

Since $HX^n(X^2 - \delta(X)) \neq 0$ by assumption, it follows from the above sequence that $HX^n_{\mathbb{Z}_2}(X^2 - \delta(X)) \neq 0$. We already know that the fourth term is trivial in the above sequence. That means the third map is surjective. Since $HX^n(X^2 - \delta(X)) = \mathbb{Z}_2$, we can conclude that $HX^n_{\mathbb{Z}_2}(X^2 - \delta(X)) = \mathbb{Z}_2$. $\qquad \square$

**Remark 8.7** It follows from the first part of the proof of the above theorem that if $HX^*(X^2 - \delta(X)) = 0$ for all $* \geq n + 1$ and $HX^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for some $* \geq n + 1$, then $HX^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for all $* \geq n + 1$. In particular, for such $X$, we have $\text{cvk}^*(X)$ is trivial for $* \geq n + 1$ and hence $\text{cobdim}(X) \leq n$.

# 9 Upper bound of cobdim

In this section, our goal is to apply Theorem 8.6 to estimate cobdim for proper, uniformly acyclic $n$-manifolds. The key tool that we are going to exploit here is a coarse Alexander duality theorem that holds for such spaces. In fact, such duality holds for a more general class of metric spaces called coarse PD($n$) spaces, first introduced by Kapovich and Kleiner in [11] where they proved a coarse Alexander duality theorem for these spaces. A different treatment of coarse PD($n$) space and coarse Alexander duality using coarse cohomology is given in [2]. Let us now recall the definition of a coarse PD($n$) space from [2].

**Definition 9.1** A metric space $X$ is a *coarse* $\mathrm{PD}(n)$ *space* if there exist chain maps

$$p \colon \mathrm{C}^*(X;\mathbb{Z}) \to \mathrm{C}\,\mathrm{X}_{n-*}(X;\mathbb{Z}) \quad \text{and} \quad q \colon \mathrm{C}\,\mathrm{X}_{n-*}(X;\mathbb{Z}) \to \mathrm{C}^*(X;\mathbb{Z}),$$

so that $pq$ and $qp$ are chain homotopic to identities via chain homotopies $G \colon \mathrm{C}\,\mathrm{X}_*(X;\mathbb{Z}) \to \mathrm{C}\,\mathrm{X}_{*+1}(X;\mathbb{Z})$ and $F \colon \mathrm{C}^*(X;\mathbb{Z}) \to \mathrm{C}^{*-1}(X;\mathbb{Z})$ which are controlled:

$$\forall \phi \in \mathrm{C}^*(X;\mathbb{Z}) \ |p(\phi)| \stackrel{\mathrm{c}}{\subset} |\phi|,$$

$$\forall \phi \in \mathrm{C}^*(X;\mathbb{Z}) \ |F(\phi)| \stackrel{\mathrm{c}}{\cap} \Delta \stackrel{\mathrm{c}}{\subset} |\phi|,$$

$$\forall c \in \mathrm{C}\,\mathrm{X}_*(X;\mathbb{Z}) \ |q(c)| \stackrel{\mathrm{c}}{\cap} \Delta \stackrel{\mathrm{c}}{\subset} |c|,$$

$$\forall c \in \mathrm{C}\,\mathrm{X}_*(X;\mathbb{Z}) \ |G(c)| \stackrel{\mathrm{c}}{\subset} |c|.$$

**Example 9.2** Any proper, uniformly acyclic $n$-manifold is a coarse $\mathrm{PD}(n)$ space [2, Corollary 8.3]. In particular, the universal cover of a closed, aspherical $n$-manifold is a coarse $\mathrm{PD}(n)$ space.

We now recall the coarse Alexander duality theorem from [2].

**Theorem 9.3** (coarse Alexander duality [2]) *If $X$ is a coarse $\mathrm{PD}(n)$ space, then, for any $A \subset X$ and finitely generated abelian group $G$,*

$$\mathrm{H}\,\mathrm{X}^*(X - A; G) \cong \mathrm{H}\,\mathrm{X}_{n-*}(A; G).$$

As a consequence we have the following.

**Lemma 9.4** *Let $X$ be a coarse $\mathrm{PD}(n)$ space. Then*

$$\mathrm{H}\,\mathrm{X}_*(X;\mathbb{Z}_2) = \begin{cases} \mathbb{Z}_2 & \text{if } * = n, \\ 0 & \text{otherwise.} \end{cases}$$

**Proof** By Theorem 9.3, $\mathrm{H}\,\mathrm{X}_{n-*}(X;G) = \mathrm{H}\,\mathrm{X}^*(X-X;G)$. Observe that $\mathrm{C}\,\mathrm{X}^*(X-X;\mathbb{Z}_2) = \mathrm{C}^*(X;\mathbb{Z}_2)$. Since $\mathrm{C}^*(X;\mathbb{Z}_2)$ is acyclic by Lemma 2.1, we obtain

$$\mathrm{H}\,\mathrm{X}_*(X;\mathbb{Z}_2) = \mathrm{H}\,\mathrm{X}^{n-*}(X - X;\mathbb{Z}_2) = \begin{cases} \mathbb{Z}_2 & \text{if } * = n, \\ 0 & \text{otherwise.} \end{cases} \qquad \square$$

For the rest of the paper, the omitted coefficient will mean $\mathbb{Z}_2$.

**Lemma 9.5** *If $X$ is a coarse $\mathrm{PD}(n)$ space, then*

$$\mathrm{H}\,\mathrm{X}^*(X^2 - \delta(X)) = \begin{cases} \mathbb{Z}_2 & \text{if } * = n, \\ 0 & \text{otherwise.} \end{cases}$$

**Proof** If $X$ is a coarse $\mathrm{PD}(n)$ space, then $X^2$ is a coarse $\mathrm{PD}(2n)$ space. We obtain

$$\mathrm{H}\,\mathrm{X}^*(X^2 - \delta(X)) = \mathrm{H}\,\mathrm{X}_{2n-*}(\delta(X)) \quad \text{(by Theorem 9.3)}$$

$$= \begin{cases} \mathbb{Z}_2 & \text{if } * = n, \\ 0 & \text{otherwise} \end{cases} \quad \text{(by Lemma 9.4)}. \qquad \square$$

**Lemma 9.6** *If $X$ is a proper, uniformly acyclic $n$-manifold where $n \geq 1$, then $\mathrm{H\,X}^i_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for $i \geq 2n + 2$.*

**Proof** If $X$ is bounded, then $X^2 \overset{\mathrm{c}}{=} \delta(X)$ and the claim follows from Theorem 3.8(1). So, we assume that $X$ is unbounded. This implies that $X^2$ is not coarsely contained in $\delta(X)$. By hypothesis, $X^2$ is uniformly acyclic and locally acyclic. Hence we can apply Theorem 3.8 to the pair $(X^2, \delta(X))$ and obtain

$$\mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = \varinjlim \tilde{\mathrm{H}}^{*-1}\big(X^2 - N_r(\delta(X))/\mathbb{Z}_2\big).$$

Since $\big(X^2 - N_r(\delta(X))\big)/\mathbb{Z}_2$ is a $2n$-manifold, $\mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for all $* \geq 2n + 2$. $\qquad \square$

**Theorem 9.7** *If $X$ is a proper, uniformly acyclic $n$-manifold, then $\mathrm{cobdim}(X) \leq n$.*

**Proof** For $n = 0$, the claim is trivial. For $n \geq 1$, Lemmas 9.5 and 9.6 imply that $X$ satisfies the hypothesis of Theorem 8.6. Hence by Theorem 8.6, we have $\mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for $* \geq n + 1$. This implies $\mathrm{cobdim}(X) \leq n$. $\qquad \square$

**Remark 9.8** To prove Theorem 9.7, we did not need the full strength of the Theorem 8.6. We needed to show $\mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X)) = 0$ for $* \geq n + 1$ which only requires vanishing of $\mathrm{H\,X}^*(X^2 - \delta(X))$ for $* \geq n + 1$ and vanishing of $\mathrm{H\,X}^*_{\mathbb{Z}_2}(X^2 - \delta(X))$ for some $* \geq n + 1$ (see Remark 8.7).

Our next goal is to improve Theorem 9.7 to include manifolds with boundaries. For that, we need to impose a condition on the metric of the boundary. This is the purpose of the following definition which is inspired by the uniformly locally $k$-connected space defined in [6].

**Definition 9.9** A metric space $(X, d)$ is uniformly locally acyclic if for every $\epsilon > 0$, there is a $\delta > 0$ such that any ball of radius $\delta$ is acyclic inside a ball of radius $\epsilon$.

**Example 9.10** Any compact, locally acyclic space is uniformly locally acyclic. Similarly, any locally acyclic space that admits a cocompact group action by homeomorphisms is uniformly locally acyclic. In particular, the universal cover of a compact manifold is uniformly locally acyclic.

Any uniformly locally acyclic space is also locally acyclic. However, the converse is not true. For example, the set $\big\{\frac{1}{n}\big\}$ with the subspace metric from $\mathbb{R}$ is locally acyclic, but it is not uniformly locally acyclic.

**Lemma 9.11** *Let $(X, d)$ be a uniformly locally acyclic metric space. Then the space $X \times [1, \infty)$ has a metric that makes the space uniformly acyclic away from $X \times \{1\}$ and the map $x \mapsto (x, 1)$ is an isometric embedding of $X$ into $X \times [1, \infty)$.*

**Proof** The construction of the metric follows the one appearing in Lemma 2.2 of [6]. Choose a continuous strictly increasing function $\phi \colon [1, \infty) \to [1, \infty)$ with $\phi(1) = 1$. Let $d$ be the original metric on $X$ and define a function $\rho'$ by

  (1)  $\rho'((x, t), (x', t)) = \phi(t)d(x, x')$,

  (2)  $\rho'((x, t), (x, t')) = |t - t'|$.

We then define $\rho \colon (X \times [1, \infty))^2 \to [0, \infty)$ to be

$$\rho((x, t), (x', t')) = \inf \sum_{i=1}^{l} \rho'((x_i, t_i), (x_{i-1}, t_{i-1})),$$

where the sum is over all chains

$$(x, t) = (x_0, t_0), (x_1, t_1), \dots, (x_l, t_l) = (x', t')$$

and each segment is either horizontal or vertical. Also $\phi(1) = 1$ implies that $X \times \{1\}$ with the subspace metric is isometric to $X$ via the map $(x, 1) \mapsto x$. Now we will describe a $\phi$, so that the corresponding metric $\rho$ makes $X \times [1, \infty)$ uniformly acyclic away from $X \times \{1\}$. Since $X$ is uniformly locally acyclic, we have an infinite positive decreasing sequence $\{r_i\}$ with $r_1 = 1$ such that for every $x \in X$, the inclusions $\cdots \subset B_{r_i}^d(x) \subset B_{r_{i-1}}^d(x) \subset$ are nullhomotopic maps. Set $\phi(t) = 1/r_t$ for $t \in \mathbb{N}$. For nonintegral values of $t$, we set

$$\phi(t) = \phi([t]) + (t - [t])\phi([t] + 1).$$

Suppose, $N_i = \phi(i)/\phi(i - 1)$. Now we consider the ball $B_k^\rho(x, i) \subset X \times [1, \infty)$. Note that

$$B_k^\rho(x, i) \subset B_{k/N_{i-k}}^d(x) \times [i - k, i + k]$$

and that $B_k^\rho(x, i)$ contracts in itself to $B_k^\rho(x, i) \cap (X \times [i - k, i]) \subset B_{k/N_{i-k}}^d(x) \times [i - k, i]$. Also, $B_{k/N_{i-k-1}}^d(x) \times \{i - k - 1\} \subset B_{k+2}^\rho(x, i)$. So, $B_k^\rho(x, i)$ can be contracted inside $B_{k+2}^\rho(x, i)$ by pushing it down to the $(i-k-1)$-level and contracting it there. $\qquad \square$

The following gluing lemma in a slightly different form can be found in [4].

**Lemma 9.12** [4, Lemma I.5.24]  *Let $X_1$ and $X_2$ be two proper metric spaces. Let $A_i \subset X_i$ be closed subsets and $f \colon A_1 \to A_2$ be an isometry. Let $Y$ be the space obtained by gluing $(X_i, A_i)$ along $A_i$ via the map $f$. Define $d \colon Y \times Y \to \mathbb{R}$ as*

$$d(x, y) = \begin{cases} d_i(x, y) & \text{if } x, y \in X_i, \\ \inf_{a \in A_1}\{d_1(x, a) + d_2(f(a), y)\} & \text{if } x \in X_1, \ y \in X_2. \end{cases}$$

*Then*:

(1)  *$d$ is a proper metric on $Y$.*

(2)  *The canonical inclusions $X_i \hookrightarrow Y$ are isometric embedding.*

**Proposition 9.13**  *Any uniformly acyclic, proper $n$-manifold with uniformly locally acyclic boundary admits an isometric embedding into a uniformly acyclic, proper $n$-manifold.*

**Proof** Since $\partial M$ is uniformly locally acyclic, Lemma 9.11 allows us to equip $\partial M \times [1, \infty)$ with a metric so that it is uniformly acyclic away from $\partial M \times \{1\}$. Let $\rho, \mu \colon \mathbb{R}_+ \to \mathbb{R}_+$ be two functions such that any ball $B(x, r)$ in $\partial M \times [1, \infty)$ is acyclic inside $B(x, \rho(r))$ whenever $d(x, \partial M \times \{1\}) \geq \mu(r)$. We glue $\partial M \times [1, \infty)$ to $M$ along $\partial M$ by the attaching map $(x, 1) \mapsto x$. Let $Y$ be the resulting space. By Lemma 9.12, there is a proper metric on $Y$ such that the canonical maps $M \hookrightarrow Y$ and $\partial M \times [1, \infty) \hookrightarrow Y$ are isometric embedding. For the rest of the proof we will regard $M$ and $\partial M \times [1, \infty)$ as subspaces of $Y$.

We claim that $Y$ is uniformly acyclic. Since $M$ is uniformly acyclic, there is a function $\tau \colon [0, \infty) \to [0, \infty)$ such that, for any $r \geq 0$, any ball of radius $r$ in $M$ is acyclic inside a concentric ball of radius $\tau(r)$. Take a ball $B_r(x)$ of radius $r$ in $Y$. If $x \in M$ and $d(x, \partial M) \geq r$, then $B_r(x)$ is contained inside $M$ because points in $\partial M \times [1, \infty)$ are at least as far from $x$ as points in $\partial M$ by the construction of the metric on $Y$. Hence, $B_r(x)$ is acyclic inside $B_{\tau(r)}(x)$. If $x \in \partial M \times [1, \infty)$ and $d(x, \partial M) \geq \mu(r) + r$, then $B(x, r) \subset \partial M \times [1, \infty)$ and hence is acyclic in $B(x, \rho(r))$. If $d(x, \partial M) \leq \mu(r) + r$, we can deformation retract $B_r(x) \cap (\partial M \times [1, \infty))$ inside $\partial M \times [1, \infty)$ by sliding it along $[1, \infty)$ until it lands on $\partial M \times \{1\}$. Note that this deformation takes place inside $\partial M \times [1, \mu(r) + 2r]$, and the diameter shrinks as one approaches $\partial M \times \{1\}$. Hence the deformation takes place in a set of diameter at most $\mu(r) + 2r + 2r$. Also, the deformed ball is now contained in $M$ and has diameter at most $2r$ and hence it is acyclic inside a set of diameter at most $\tau(2r)$ by uniform acyclicity of $M$. Hence, we conclude that $B_r(x)$ is acyclic inside a set of diameter at most $\mu(r) + 4r + \tau(2r)$. Hence $Y$ is uniformly acyclic. $\square$

As a consequence of the above proposition, we get the following.

**Theorem 9.14** *If $X$ is a proper, uniformly acyclic $n$-manifold with uniformly locally acyclic boundary, then* $\mathrm{cobdim}(X) \leq n$.

**Proof** By Proposition 9.13, we have a uniformly acyclic proper $n$-manifold $Y$ such that $X$ embeds isometrically in $Y$. By Theorem 6.4, it follows that $\mathrm{cobdim}(X) \leq \mathrm{cobdim}(Y)$. By Theorem 9.7, we know $\mathrm{cobdim}(Y) \leq n$ and hence $\mathrm{cobdim}(X) \leq n$. $\square$

Now we can state the following improvement of Corollary 6.6. The proof is immediate from Theorems 6.4 and 9.14.

**Corollary 9.15** *If $\mathrm{cobdim}(X) \geq n$, then $X$ cannot be coarsely embedded into a proper, uniformly acyclic $(n-1)$-manifold with a uniformly locally acyclic boundary.*

**Definition 9.16** (cocompact action dimension) The *cocompact action dimension* $\mathrm{cadim}(G)$ of a group $G$ is the least dimension of a contractible manifold (possibly with boundary) that admits a proper cocompact $G$-action.

**Corollary 9.17** $$\mathrm{cadim}(G) \geq \mathrm{cobdim}(G).$$

**Proof** Suppose $\mathrm{cadim}(G) = n$. Then $G$ admits a proper, cocompact action on an acyclic $n$-manifold $M$. Choose a point $x_0 \in M$. By the Milnor–Schwarz lemma, the map $g \mapsto g.x_0$ gives a coarse equivalence $f: G \to M$. Since $M$ is acyclic and it admits a cocompact action, $M$ is uniformly acyclic. Similarly, since $\partial M$ is locally acyclic and it admits a cocompact action, it is uniformly locally acyclic. Since $M$ is proper, by Corollary 9.15, we get $\mathrm{cobdim}(G) \leq n = \mathrm{cadim}(G)$. □

**Remark 9.18** The action dimension $\mathrm{actdim}(G)$ of a group $G$ is the least dimension of a contractible manifold that admits a proper $G$-action. Note that $\mathrm{cadim}(G) \geq \mathrm{actdim}(G)$, however, we do not know of any groups where $\mathrm{cadim} > \mathrm{actdim}$. Nonetheless, we believe that $\mathrm{cobdim}(G)$ gives a lower bound to $\mathrm{actdim}(G)$. A naive approach to prove this might be as follows: Suppose, $G$ admits a proper action on a contractible $n$-manifold $M$. Then the map $f: G \to M$, sending $G$ to one of its orbit gives a coarse embedding. Furthermore, the image of $f$ is uniformly contractible in $M$: for any $r$ there exists $s$ such that any ball of radius $r$ in $M$ centered at a point in $f(G)$ is uniformly contractible inside a concentric ball of radius $s$. This should imply that there is a bounded function $p: f(G) \to (0, \infty)$ such that the space $X = \bigcup_{x \in f(G)} N_{p(x)}(x)$ is a uniformly contractible manifold with boundary. Since $G$ acts on $X$ cocompactly, $X$ has uniformly locally acyclic boundary. Theorem 9.14 then implies that $\mathrm{cobdim}(X) \leq n$. Since $G$ coarsely embed into $X$ by the map $f$, we get $\mathrm{cobdim}(G) \leq \mathrm{cobdim}(X) \leq n$, proving the desired claim. However, the author does not know how to prove that there exists such $X$.

**Question** Is $\mathrm{cobdim}(G) \geq \mathrm{actdim}(G)$ true?

# 10 Lower bound of cobdim

**Theorem 10.1** If $\mathrm{H\,X}^*(X^2 - \delta(X)) = 0$ for $* \leq n - 1$, then $\mathrm{cobdim}(X) \geq n$.

**Proof** For $n = 0$, the claim is trivial. If $n = 1$, the assumption says $\mathrm{H\,X}^0(X^2 - \delta(X)) = 0$. This means $X$ is unbounded, otherwise, nonzero constant functions from $X^2$ to $\mathbb{Z}_2$ give nontrivial elements in $\mathrm{H\,X}^0(X^2 - \delta(X))$. Hence, $\mathrm{cobdim}(X) \geq 1$ in this case.

Suppose $n \geq 2$ and $f: X \to \ell^\infty$ is an isometry. We first show that

$$f^*: \mathrm{H\,X}^2_{\mathbb{Z}_2}(\ell^\infty)^2 - \delta(\ell^\infty)) \xrightarrow{f^*} \mathrm{H\,X}^2_{\mathbb{Z}_2}(X^2 - \delta(X))$$

is a nontrivial map. To see that, consider the following part of the maps between the concerned coarse Gysin sequences. Our goal is to show that the second vertical map is nontrivial:

$$
\begin{array}{ccccccc}
\mathrm{H\,X}^1((\ell^\infty)^2 - \delta(\ell^\infty)) & \to & \mathrm{H\,X}^1_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty)) \oplus \mathbb{Z}_2 & \to & \mathrm{H\,X}^2_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty)) & \to & \cdots \\
 & & \downarrow{f^*} & & \downarrow{f^*} & & \\
\mathrm{H\,X}^1(X^2 - \delta(X)) & \longrightarrow & \mathrm{H\,X}^1_{\mathbb{Z}_2}(X^2 - \delta(X)) \oplus \mathbb{Z}_2 & \xrightarrow{j} & \mathrm{H\,X}^2_{\mathbb{Z}_2}(X^2 - \delta(X)) & \longrightarrow & \cdots
\end{array}
$$

By the commutativity of the diagram, our claim follows if we can show that $j$ is injective and the first vertical map is nontrivial.

Since $HX^1(X^2 - \delta(X)) = 0$, it follows from the long exact sequence (8-5) that $j$ is injective.

Next, we show that the first vertical map in the above commutative diagram is nontrivial. It is equivalent to showing that the following map is nontrivial:

$$f^*: HX^1_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty), \delta(\ell^\infty)) \to HX^1_{\mathbb{Z}_2}((X^2 - \delta(X), \delta(X))).$$

We can show that using the diagram

$$
\begin{array}{ccccccccc}
0 & \to & H^0(C^*(\delta(\ell^\infty))) & \to & HX^1_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty), \delta(\ell^\infty)) & \to & HX^1((\ell^\infty)^2 - \delta(\ell^\infty)) & \to & 0 \\
& & \downarrow f^* & & \downarrow f^* & & \downarrow f^* & & \downarrow f^* \\
0 & \to & H^0(C^*(\delta(X))) & \longrightarrow & HX^1_{\mathbb{Z}_2}(X^2 - \delta(X), \delta(X)) & \xrightarrow{\ j\ } & HX^1(X^2 - \delta(X)) & \longrightarrow & 0
\end{array}
$$

The long exact sequence is due to the second long exact sequence of Lemma 8.1 combined with the fact that $HX^1((\ell^\infty)^2 - \delta(\ell^\infty))$, $H^1(C^*(\delta(\ell^\infty)))$, $HX^1(X^2 - \delta(X))$ and $H^1(C^*(\delta(X)))$ all are trivial. The second vertical map is an isomorphism as both the domain and the range are constant functions defined on the respective spaces. The fourth vertical map is an isomorphism because both the domain and the range are trivial. So, applying the five lemma on the above diagram, we conclude that the middle vertical map is an isomorphism, as desired.

Hence we obtain that the map $f^*: HX^2_{\mathbb{Z}_2}(\ell^\infty)^2 - \delta(\ell^\infty)) \to HX^2_{\mathbb{Z}_2}(X^2 - \delta(X))$ is nontrivial.

Let us now consider the maps between the following parts of the coarse Gysin sequences where $* \geq 2$:

$$
\begin{array}{ccccccc}
\to & HX^*((\ell^\infty)^2 - \delta(\ell^\infty)) & \to & HX^*_{\mathbb{Z}_2}(\ell^\infty)^2 - \delta(\ell^\infty)) & \to & HX^{*+1}_{\mathbb{Z}_2}((\ell^\infty)^2 - \delta(\ell^\infty)) & \to \\
& & & \downarrow f^* & & \downarrow f^* & \\
\longrightarrow & HX^*(X^2 - \delta(X)) & \longrightarrow & HX^*_{\mathbb{Z}_2}(X^2 - \delta(X)) & \longrightarrow & HX^{*+1}_{\mathbb{Z}_2}(X^2 - \delta(X)) & \longrightarrow
\end{array}
$$

The first terms of both sequences above are trivial for $* \leq n - 1$. That implies that the third horizontal maps in the above diagram are injective. Hence by commutativity of the diagram, if the second vertical map is nontrivial then so is the third vertical map when $* \leq n - 1$. We saw previously that, when $* = 2$, the second vertical map is injective. It now follows by induction that the third vertical map is injective when $* \leq n - 1$. In particular, when $* = n - 1$, injectivity of the third vertical map means $cvk^n(X) \neq 0$. Hence, $cobdim(X) \geq n$. $\qquad\square$

As a consequence we get the following, which also follows from [16, Lemma 5.3].

**Corollary 10.2** *If $X$ is a coarse $PD(n)$ space, then $X$ does not coarsely embed into a proper, uniformly acyclic $(n-1)$-manifold with a uniformly locally acyclic boundary.*

**Proof** If $X$ is a coarse $PD(n)$ space, it satisfies the hypotheses of Theorem 10.1. Hence, $cobdim(X) \geq n$. The claim now follows from Corollary 9.15. $\qquad\square$

**Example 10.3** If $X$ is a proper, uniformly acyclic $n$-manifold, then it satisfies all the hypotheses of the above corollary. Hence $\mathrm{cobdim}(X) \geq n$. Theorem 9.14 implies that $\mathrm{cobdim}(X) \leq n$. Hence $\mathrm{cobdim}(X) = n$ whenever $X$ is a proper, uniformly acyclic $n$-manifold.

# References

[1] **A Banerjee**, *On the computation of coarse cohomology*, preprint (2024)  arXiv 2401.02231

[2] **A Banerjee**, **B Okun**, *Coarse cohomology of the complement*, Groups Geom. Dyn. (online publication June 2025)

[3] **M Bestvina**, **M Kapovich**, **B Kleiner**, *Van Kampen's embedding obstruction for discrete groups*, Invent. Math. 150 (2002) 219–235  MR

[4] **M R Bridson**, **A Haefliger**, *Metric spaces of non-positive curvature*, Grundl. Math. Wissen. 319, Springer (1999)  MR

[5] **U Bunke**, **A Engel**, **D Kasprowski**, **C Winges**, *Equivariant coarse homotopy theory and coarse algebraic K-homology*, from "$K$-theory in algebra, analysis and topology" (G Cortiñas, C A Weibel, editors), Contemp. Math. 749, Amer. Math. Soc., Providence, RI (2020) 13–104  MR

[6] **A N Dranishnikov**, **S C Ferry**, **S Weinberger**, *Large Riemannian manifolds which are flexible*, Ann. of Math. 157 (2003) 919–938  MR

[7] **M Fréchet**, *Les dimensions d'un ensemble abstrait*, Math. Ann. 68 (1910) 145–168

[8] **S Hair**, *Homological methods in coarse geometry*, PhD thesis, The Pennsylvania State University (2010)  MR  Available at `https://www.proquest.com/docview/853739344`

[9] **H Honkasalo**, *Equivariant Alexander–Spanier cohomology*, Math. Scand. 63 (1988) 179–195  MR

[10] **E R van Kampen**, *Komplexe in euklidischen Räumen*, Abh. Math. Sem. Univ. Hamburg 9 (1933) 72–78  MR

[11] **M Kapovich**, **B Kleiner**, *Coarse Alexander duality and duality groups*, J. Differential Geom. 69 (2005) 279–352  MR

[12] **L Mosher**, **M Sageev**, **K Whyte**, *Quasi-actions on trees, II*: *Finite depth Bass–Serre trees*, Mem. Amer. Math. Soc. 1008, Amer. Math. Soc. (2011)  MR

[13] **J Roe**, *Coarse cohomology and index theory on complete Riemannian manifolds*, Mem. Amer. Math. Soc. 497, Amer. Math. Soc. (1993)  MR

[14] **J Roe**, *Lectures on coarse geometry*, University Lecture Series 31, Amer. Math. Soc., Providence, RI (2003)  MR

[15] **C Wulff**, *Equivariant coarse (co-)homology theories*, Symmetry Integrability Geom. Methods Appl. 18 (2022) art. id. 057  MR

[16] **S Y Yoon**, *A lower bound to the action dimension of a group*, Algebr. Geom. Topol. 4 (2004) 273–296  MR

*Department of Mathematics and Statistics, Auburn University*
*Auburn, AL, United States*

`azb0263@auburn.edu`

# Correction to the article
# An algebraic model for finite loop spaces

CARLES BROTO

RAN LEVI

BOB OLIVER

We correct here two errors in our earlier paper *An algebraic model for finite loop spaces*.

55R35; 20D20, 20E22

We correct two erroneous arguments found in [3].

In the proof of [3, Lemma A.8], we applied [2, Proposition 5.4] to orbit categories of transporter systems, while that proposition is stated only for orbit categories of fusion systems. After replacing that by Proposition 2.3 below, Lemma A.8 can be proven as stated in [3].

In the proof of [3, Corollary A.10], we applied [3, Proposition A.9(b)] in a way that is not valid unless all objects in $\mathcal{L}$ are $\mathcal{F}$-centric. That corollary is a special case of Proposition 1.6 here.

# 1   Inclusions of linking systems for the same fusion system

We refer to [3, Definition 1.9] for the complete definition of a linking system associated to a fusion system $\mathcal{F}$ over a discrete $p$-toral group $S$. Very briefly, it consists of a category $\mathcal{L}$ whose objects are subgroups of $S$, together with a pair of functors

$$\mathcal{T}_{\mathrm{Ob}(\mathcal{L})}(S) \xrightarrow{\delta} \mathcal{L} \xrightarrow{\pi} \mathcal{F}$$

satisfying certain conditions. Here, $\mathcal{T}_{\mathrm{Ob}(\mathcal{L})}(S)$ is the *transporter category*: the category with the same objects as $\mathcal{L}$, and where $\mathrm{Mor}_{\mathcal{T}_{\mathrm{Ob}(\mathcal{L})}(S)}(P, Q)$ is the set of all $g \in S$ such that ${}^{g}P \leq Q$. Also, the set $\mathrm{Ob}(\mathcal{L})$ is required to be closed under $\mathcal{F}$-conjugacy and overgroups, and must include all subgroups of $S$ that are $\mathcal{F}$-centric and $\mathcal{F}$-radical.

As usual, when $P \leq Q$ are objects in a linking system $\mathcal{L}$, we write $\iota_{P,Q} = \delta_{P,Q}(1) \in \mathrm{Mor}_{\mathcal{L}}(P, Q)$ for the inclusion morphism. By "extensions" and "restrictions" of morphisms we mean extensions and restrictions of source and target both, with respect to these inclusions.

**Proposition 1.1** *The following hold for each linking system $\mathcal{L}$ associated to a saturated fusion system $\mathcal{F}$ over a discrete $p$-toral group $S$.*

(a) *For each $P, Q \in \mathrm{Ob}(\mathcal{L})$, the homomorphism $\pi_{P,Q} \colon \mathrm{Mor}_{\mathcal{L}}(P, Q) \to \mathrm{Hom}_{\mathcal{F}}(P, Q)$ is surjective. The group $\mathrm{Ker}(\pi_P)$ acts freely on $\mathrm{Mor}_{\mathcal{L}}(P, Q)$, and $\pi_{P,Q}$ induces a bijection*

$$\mathrm{Mor}_{\mathcal{L}}(P, Q)/\mathrm{Ker}(\pi_P) \xrightarrow{\cong} \mathrm{Hom}_{\mathcal{F}}(P, Q).$$

*If $P$ is fully centralized in $\mathcal{F}$, then $\mathrm{Ker}(\pi_P) = \delta_P(C_S(P))$.*

(b) *For every morphism $\psi \in \mathrm{Mor}_{\mathcal{L}}(P, Q)$, and every $P_*, Q_* \in \mathrm{Ob}(\mathcal{L})$ such that $P_* \leq P$, $Q_* \leq Q$, and $\pi(\psi)(P_*) \leq Q_*$, there is a unique morphism $\psi|_{P_*,Q_*} \in \mathrm{Mor}_{\mathcal{L}}(P_*, Q_*)$ (the "restriction" of $\psi$) such that $\psi \circ \iota_{P_*,P} = \iota_{Q_*,Q} \circ \psi|_{P_*,Q_*}$.*

(c) *Let $P \trianglelefteq \overline{P} \leq S$ and $Q \leq \overline{Q} \leq S$ be objects in $\mathcal{L}$. Let $\psi \in \mathrm{Mor}_{\mathcal{L}}(P, Q)$ be such that for each $g \in \overline{P}$, there is $h \in \overline{Q}$ satisfying $\iota_{Q,\overline{Q}} \circ \psi \circ \delta_P(g) = \delta_{Q,\overline{Q}}(h) \circ \psi$. Then there is a unique morphism $\overline{\psi} \in \mathrm{Mor}_{\mathcal{L}}(\overline{P}, \overline{Q})$ such that $\overline{\psi}|_{P,Q} = \psi$.*

(d) *All morphisms in $\mathcal{L}$ are monomorphisms and epimorphisms in the categorical sense.*

(e) *A morphism $\psi$ in $\mathcal{L}$ is an isomorphism in $\mathcal{L}$ if $\pi(\psi)$ is an isomorphism in $\mathcal{F}$.*

**Proof** Points (a)–(d) are shown in [3, Proposition A.4], while (e) follows from [3, Propositions A.2(c) and A.5]. $\qquad\square$

We want to show that the geometric realizations of two linking systems associated to the same fusion system are homotopy equivalent. The next lemma is a first step towards doing that.

**Lemma 1.2** *Let $\mathcal{F}$ be a saturated fusion system over a discrete $p$-toral group $S$. Let $\mathcal{L}_0 \subseteq \mathcal{L}$ be linking systems associated to $\mathcal{F}$ such that $\mathrm{Ob}(\mathcal{L}) \smallsetminus \mathrm{Ob}(\mathcal{L}_0) = \mathcal{P}$, where $\mathcal{P}$ is an $\mathcal{F}$-conjugacy class of subgroups of $S$. Then the inclusion of nerves $|\mathcal{L}_0| \subseteq |\mathcal{L}|$ is a homotopy equivalence.*

**Proof** The following proof is essentially that given in [1, Proposition 3.11], modified for fusion systems over discrete $p$-toral groups. To simplify notation, for $\varphi \in \mathrm{Mor}_{\mathcal{L}}(Q, R)$, we write $\mathrm{Im}(\varphi) = \mathrm{Im}(\pi(\varphi)) \leq R$, and $\varphi(Q_0) = \pi(\varphi)(Q_0) \leq R$ if $Q_0 \leq Q$.

We must show that the inclusion functor $\mathcal{I} \colon \mathcal{L}_0 \to \mathcal{L}$ induces a homotopy equivalence $|\mathcal{L}_0| \simeq |\mathcal{L}|$. By Quillen's Theorem A (see [5]), it will be enough to prove that the undercategory $P \downarrow \mathcal{I}$ is contractible (ie $|P \downarrow \mathcal{I}| \simeq *$) for each $P$ in $\mathcal{L}$. This is clear when $P \notin \mathcal{P}$ (since $P \downarrow \mathcal{I}$ has initial object $(P, \mathrm{Id})$ in that case), so it suffices to consider the case $P \in \mathcal{P}$. Since all subgroups in $\mathcal{P}$ are isomorphic in the category $\mathcal{L}$, we can assume that $P$ is fully normalized.

Set

$$\widehat{P} = \{x \in N_S(P) \mid c_x \in O_p(\mathrm{Aut}_{\mathcal{F}}(P))\}.$$

Recall that $P \notin \mathrm{Ob}(\mathcal{L}_0)$, and hence by definition of a linking system cannot be both $\mathcal{F}$-centric and $\mathcal{F}$-radical. If $P$ is not $\mathcal{F}$-centric, then $\widehat{P} \geq P \cdot C_S(P) > P$. If $P$ is not $\mathcal{F}$-radical, then $\mathrm{Inn}(P) < O_p(\mathrm{Aut}_{\mathcal{F}}(P)) \leq$

$\text{Aut}_S(P)$, the last inclusion since $P$ is fully normalized, and so $\widehat{P} > P$. Thus $\widehat{P} \in \text{Ob}(\mathcal{L}_0)$ (and $P \trianglelefteq \widehat{P}$) in either case.

Set $N_{\mathcal{L}_0}(P) = \mathcal{L}_0 \cap N_{\mathcal{L}}(P)$, and let

$$\mathcal{I}_N \colon N_{\mathcal{L}_0}(P) \to N_{\mathcal{L}}(P)$$

be the inclusion (thus the restriction of $\mathcal{I}$). Consider the functors

$$\widehat{P} {\downarrow} \mathcal{I}_N \xrightarrow{j_2} P {\downarrow} \mathcal{I}_N \xrightarrow{j_1} P {\downarrow} \mathcal{I},$$

where $j_1$ is the inclusion of undercategories, induced by the inclusions $N_{\mathcal{L}_0}(P) \to \mathcal{L}_0$ and $N_{\mathcal{L}}(P) \to \mathcal{L}$, and where $j_2$ sends an object $(Q, \alpha)$ to $(Q, \alpha \circ \iota_{P,\widehat{P}})$ (for $\alpha \in \text{Mor}_{N_{\mathcal{L}}(P)}(\widehat{P}, Q)$). For each $i = 1, 2$, we will construct a retraction $r_i$ such that $r_i \circ j_i = \text{Id}$, together with a natural transformation of functors between $j_i \circ r_i$ and the identity. It then follows that $|P {\downarrow} \mathcal{I}| \simeq |P {\downarrow} \mathcal{I}_N| \simeq |\widehat{P} {\downarrow} \mathcal{I}_N|$, and the last space is contractible since $(\widehat{P}, \text{Id}_{\widehat{P}})$ is an initial object in $\widehat{P} {\downarrow} \mathcal{I}_N$ (recall $\widehat{P} \in \text{Ob}(N_{\mathcal{L}_0}(P))$).

**Step 1** We first construct the retraction functor $r_1 \colon P {\downarrow} \mathcal{I} \to P {\downarrow} \mathcal{I}_N$, together with a natural transformation $(j_1 \circ r_1 \xrightarrow{\eta} \text{Id}_{P {\downarrow} \mathcal{I}})$. By [3, Lemma 1.7(c)], for each $R \in \mathcal{P}$ (the $\mathcal{F}$-conjugacy class of $P$), there is a morphism in $\mathcal{F}$ from $N_S(R)$ to $N_S(P)$ which sends $R$ isomorphically to $P$. Since $\pi$ is surjective on morphism sets by Proposition 1.1(a), we can choose a morphism

$$\Phi_R \in \text{Mor}_{\mathcal{L}}(N_S(R), N_S(P)),$$

for each $R \in \mathcal{P}$ such that $\Phi_R(R) = P$. We also require that $\Phi_P = \text{Id}_{N_S(P)}$.

For each $(Q, \varphi)$ in $P {\downarrow} \mathcal{I}$, set $N_{(Q,\varphi)}(P) = \Phi_{\varphi(P)}(N_Q(\varphi(P)))$. Thus $P < N_{(Q,\varphi)}(P) \le N_S(P)$, since $N_Q(\varphi(P)) > \varphi(P)$ by [2, Lemma 1.8]. Consider the diagram

(1-1)

$$P \xrightarrow{\varphi_*} N_Q(\varphi(P)) \xrightarrow{\iota} Q$$

with $\varphi'$ and $\varphi''$ to $N_{(Q,\varphi)}(P)$ via $\cong \big\downarrow (\Phi_{\varphi(P)})_*$

where $\iota$ is the inclusion (in $\mathcal{L}$), $\varphi_*$ and $(\Phi_{\varphi(P)})_*$ denote the restriction morphisms of Proposition 1.1(b) (restricting source and/or target as indicated), and $\varphi' = (\Phi_{\varphi(P)})_* \circ \varphi_*$ and $\varphi'' = \iota \circ (\Phi_{\varphi(P)})_*^{-1}$ (so the triangles commute). Define $r_1 \colon P {\downarrow} \mathcal{I} \to P {\downarrow} \mathcal{I}_N$ on objects by setting

$$r_1(Q, \varphi) = (N_{(Q,\varphi)}(P), \varphi').$$

We still need to define $r_1$ on morphisms. For each morphism $\beta \in \text{Mor}_{P {\downarrow} \mathcal{I}}((Q, \varphi), (R, \psi))$, ie for each commutative square of the form

(1-2)

$$\begin{array}{ccc} P & \xrightarrow{\varphi} & Q \\ \text{Id} \downarrow & & \downarrow \beta \\ P & \xrightarrow{\psi} & R \end{array}$$

we get the following commutative diagrams:

$$(1\text{-}3) \quad \begin{array}{ccccc}
P & \xrightarrow{\varphi_*} & N_Q(\varphi(P)) & \xrightarrow{\iota} & Q \\
{\scriptstyle \mathrm{Id}}\downarrow & & {\scriptstyle \beta_*}\downarrow & & {\scriptstyle \beta}\downarrow \\
P & \xrightarrow{\psi_*} & N_R(\psi(P)) & \xrightarrow{\iota} & R
\end{array}
\qquad
\begin{array}{ccc}
N_Q(\varphi(P)) & \xrightarrow[\cong]{(\Phi_{\varphi(P)})_*} & N_{(Q,\varphi)}(P) \\
{\scriptstyle \beta_*}\downarrow & & {\scriptstyle \hat\beta}\downarrow \\
N_R(\psi(P)) & \xrightarrow[\cong]{(\Phi_{\psi(P)})_*} & N_{(R,\psi)}(P)
\end{array}$$

Again, $\beta_*$, $\varphi_*$, and $(\Phi_{\varphi(P)})_*$ denote the morphisms after restricting source and target as shown, while $\hat\beta = (\Phi_{\psi(P)})_* \circ \beta_* \circ (\Phi_{\varphi(P)})_*^{-1}$.

Since each of the three squares in (1-3) commutes, it follows that $\psi' = \hat\beta \circ \varphi'$ and $\psi'' \circ \hat\beta = \varphi''$, where $\varphi'$, $\psi'$, $\varphi''$, and $\psi''$ are as in (1-1). So we can define $r_1$ on morphisms by setting

$$r_1(\beta) = \hat\beta \colon (N_{(Q,\varphi)}(P), \varphi') \to (N_{(R,\psi)}(P), \psi').$$

Define a natural transformation $\eta \colon j_1 \circ r_1 \to \mathrm{Id}_{P\downarrow\mathcal{I}}$ by sending an object $(Q,\varphi)$ to the morphism

$$\varphi'' \colon (j_1 \circ r_1)(Q, \varphi) = (N_{(Q,\varphi)}(P), \varphi') \to (Q, \varphi)$$

(a natural transformation since $\psi'' \circ \hat\beta = \varphi''$). Since $r_1 \circ j_1 = \mathrm{Id}_{P\downarrow\mathcal{I}_N}$, this finishes the proof that $|P\downarrow\mathcal{I}| \simeq |P\downarrow\mathcal{I}_N|$.

**Step 2**  Recall that $j_2 \colon \widehat{P}\downarrow\mathcal{I}_N \to P\downarrow\mathcal{I}_N$ is induced by precomposing with the inclusion $\iota_{P,\widehat{P}} \in \mathrm{Mor}_{\mathcal{L}}(P, \widehat{P})$. We now construct a retraction functor $r_2 \colon P\downarrow\mathcal{I}_N \to \widehat{P}\downarrow\mathcal{I}_N$, together with a natural transformation of functors from $\mathrm{Id}_{P\downarrow\mathcal{I}_N}$ to $j_2 \circ r_2$.

Since $P$ is fully normalized in $\mathcal{F}$ by assumption, it is also fully centralized by the Sylow axiom for a saturated fusion system [3, Definition 1.4(I)]. So by Proposition 1.1(a), the structure homomorphism $\pi_P \colon \mathrm{Aut}_{\mathcal{L}}(P) \to \mathrm{Aut}_{\mathcal{F}}(P)$ is surjective and $\mathrm{Ker}(\pi_P) = \delta_P(C_S(P))$. Since $\pi_P \circ \delta_P \colon N_S(P) \to \mathrm{Aut}_{\mathcal{F}}(P)$ sends $g$ to $c_g$, we have $\widehat{P} = \delta_P^{-1}(\pi_P^{-1}(O_p(\mathrm{Aut}_{\mathcal{F}}(P))))$, and so $\delta_P(\widehat{P}) = \pi_P^{-1}(O_p(\mathrm{Aut}_{\mathcal{F}}(P)))$. Since $\pi_P$ is surjective, this proves that $\delta_P(\widehat{P}) \trianglelefteq \mathrm{Aut}_{\mathcal{L}}(P)$.

Fix subgroups $Q, R \leq N_S(P)$ containing $P$, and let $\varphi \in \mathrm{Mor}_{N_{\mathcal{L}}(P)}(Q, R)$ be a morphism. For $g \in Q\widehat{P}$, write $g = g'x$ for $g' \in Q$ and $x \in \widehat{P}$, and let $y \in \widehat{P}$ be such that $\delta_P(y) = (\varphi|_P)\delta_P(x)(\varphi|_P)^{-1}$. Then $\varphi\delta_Q(g') = \delta_R(\varphi(g'))\varphi$ by condition (C) in the definition of a linking system ([3, Definition 1.9]), so after restriction of source and target to $P$, we get $(\varphi|_P)\delta_P(g')(\varphi|_P)^{-1} = \delta_P(\varphi(g'))$. Set $h = \varphi(g')y$. Then

$$(\varphi|_P)\delta_P(g)(\varphi|_P)^{-1} = (\varphi|_P)\delta_P(g')(\varphi|_P)^{-1} \circ (\varphi|_P)\delta_P(x)(\varphi|_P)^{-1} = \delta_P(\varphi(g')y) = \delta_P(h)$$

in $\mathrm{Aut}_{\mathcal{L}}(P)$, and hence

$$\begin{aligned}
\iota_{R,R\widehat{P}} \circ \varphi \circ \delta_Q(g) \circ \iota_{P,Q} &= \iota_{P,R\widehat{P}} \circ (\varphi|_P) \circ \delta_P(g) \\
&= \iota_{P,R\widehat{P}} \circ \delta_P(h) \circ (\varphi|_P) = \delta_{R,R\widehat{P}}(h) \circ \varphi \circ \iota_{P,Q} \in \mathrm{Mor}_{N_{\mathcal{L}_0}(P)}(P, R\widehat{P}).
\end{aligned}$$

So $\iota_{R,R\widehat{P}} \circ \varphi \circ \delta_Q(g) = \delta_{R,R\widehat{P}}(h) \circ \varphi$ since $\iota_{P,Q}$ is an epimorphism by Proposition 1.1(d). Hence by Proposition 1.1(c), there is a unique morphism $\hat{\varphi} \in \mathrm{Mor}_{\mathcal{L}_0}(Q\widehat{P}, R\widehat{P})$ (the "extension" of $\varphi$) such that the following diagram commutes in $\mathcal{L}$:

$$
\begin{array}{ccc}
Q & \xrightarrow{\ \varphi\ } & R \\
{\scriptstyle \iota_{Q,Q\widehat{P}}}\big\downarrow & {\scriptstyle \iota_{R,R\widehat{P}}} & \big\downarrow \\
Q\widehat{P} & \xrightarrow{\ \hat{\varphi}\ } & R\widehat{P}
\end{array}
$$

Note that $Q\widehat{P}, R\widehat{P} \in \mathrm{Ob}(\mathcal{L}_0)$ (and hence $\hat{\varphi}$ lies in $N_{\mathcal{L}_0}(P)$) since $\widehat{P} > P$.

The functor $\boldsymbol{r}_2 \colon P \downarrow \mathcal{I}_N \to \widehat{P}\downarrow\mathcal{I}_N$ is defined on objects by setting

$$
\boldsymbol{r}_2(Q,\alpha) = (Q\widehat{P}, \hat{\alpha}).
$$

If $\beta \colon (Q,\alpha) \to (R,\gamma)$ is a morphism in $P\downarrow\mathcal{I}_N$, that is $\beta \in \mathrm{Mor}_{\mathcal{L}}(Q, R)$ is such that $\beta \circ \alpha = \gamma$, then we define $\boldsymbol{r}_2(\beta) = \hat{\beta}$. Because of the uniqueness of the extension $\hat{\beta}$, this construction defines a functor from $P\downarrow\mathcal{I}_N$ to $\widehat{P}\downarrow\mathcal{I}_N$. Moreover, $\boldsymbol{r}_2 \circ \boldsymbol{j}_2 = \mathrm{Id}_{\widehat{P}\downarrow\mathcal{I}_N}$, and $\boldsymbol{j}_2 \circ \boldsymbol{r}_2 \simeq \mathrm{Id}_{P\downarrow\mathcal{I}_N}$, where the homotopy is induced by the natural transformation given by the inclusions $\iota_{Q,Q\widehat{P}}$. $\qquad\square$

Lemma 1.2 says that under certain conditions, we can remove one class of objects from a linking system without changing the homotopy type of its nerve. We need to combine this with the "bullet construction", defined in [2], to show that we can remove infinitely many classes of objects without changing the homotopy type.

**Definition 1.3** [2, Definition 3.1] Let $\mathcal{F}$ be a fusion system over a discrete $p$-toral group $S$, and let $T \trianglelefteq S$ be its identity component. Let $m \geq 0$ be such that $S/T$ has exponent $p^m$; thus $g^{p^m} \in T$ for all $g \in S$. Set $W = \mathrm{Aut}_{\mathcal{F}}(T)$.

    (a) For each $A \leq T$, set $I(A) = C_T(C_W(A))$, and let $I(A)_0$ be its identity component.

    (b) For each $P \leq S$, set $P^{[m]} = \langle g^{p^m} \mid g \in P \rangle \leq T$, and set $P^{\bullet} = P \cdot I(P^{[m]})_0$.

    (c) Let $\mathcal{F}^{\bullet} \subseteq \mathcal{F}$ be the full subcategory with $\mathrm{Ob}(\mathcal{F}^{\bullet}) = \{P^{\bullet} \mid P \leq S\}$.

Some of the key properties of the bullet construction are listed in the following proposition:

**Proposition 1.4** [2] *Let $\mathcal{F}$ be a saturated fusion system over a discrete $p$-toral group $S$.*

    (a) *The set $\mathrm{Ob}(\mathcal{F}^{\bullet}) = \{P^{\bullet} \mid P \leq S\}$ contains only finitely many $S$-conjugacy classes of subgroups.*

    (b) *For each $P \leq S$, we have $(P^{\bullet})^{\bullet} = P^{\bullet}$.*

    (c) *For each $P \leq Q \leq S$, we have $P^{\bullet} \leq Q^{\bullet}$.*

    (d) *For each $P, Q \leq S$ and each $\varphi \in \mathrm{Hom}_{\mathcal{F}}(P, Q)$, there is a unique map $\varphi^{\bullet} \in \mathrm{Hom}_{\mathcal{F}}(P^{\bullet}, Q^{\bullet})$ such that $\varphi^{\bullet}|_P = \varphi$.*

**Proof** Points (a)–(c) are stated as Lemma 3.2 in [2], and (d) is stated as Proposition 3.3. □

In particular, points (b)–(d) in Proposition 1.4 imply that there is a well defined retraction functor $\mathcal{F} \to \mathcal{F}^{\bullet}$ that sends an object $P$ to $P^{\bullet}$ and a morphism $\varphi$ to $\varphi^{\bullet}$.

The following lemma is shown in [2, Proposition 4.5] when $\mathcal{L}$ is a centric linking system, but we need it in a more general situation.

**Lemma 1.5** *Let $\mathcal{F}$ be a saturated fusion system over a discrete $p$-toral group $S$, and let $\mathcal{L}$ be a linking system associated to $\mathcal{F}$.*

   (a) *For each $P, Q \in \mathrm{Ob}(\mathcal{L})$ and each $\psi \in \mathrm{Mor}_{\mathcal{L}}(P, Q)$, there is a unique morphism $\psi^{\bullet} \in \mathrm{Mor}_{\mathcal{L}}(P^{\bullet}, Q^{\bullet})$ that restricts to $\psi$, and is such that $\pi(\psi^{\bullet}) = (\pi(\psi))^{\bullet}$.*

   (b) *The space $|\mathcal{L}^{\bullet}|$ is a deformation retract of $|\mathcal{L}|$, with retraction $|\mathcal{L}| \to |\mathcal{L}^{\bullet}|$ induced by the functor that sends $P$ to $P^{\bullet}$ and $\psi$ to $\psi^{\bullet}$.*

**Proof** (a) It suffices to show this when $P$ is fully normalized in $\mathcal{F}$. Set $\varphi = \pi(\psi) \in \mathrm{Hom}_{\mathcal{F}}(P, Q)$. By Proposition 1.4(d), there is a unique $\varphi^{\bullet} \in \mathrm{Hom}_{\mathcal{F}}(P^{\bullet}, Q^{\bullet})$ that extends $\varphi$. By [4, Proposition 1.13(v)], we have $C_S(P) = C_S(P^{\bullet})$, and by Proposition 1.1(a), this group acts freely and transitively on the sets $\pi_{P,Q}^{-1}(\varphi)$ and $\pi_{P^{\bullet},Q^{\bullet}}^{-1}(\varphi^{\bullet})$. The restriction map from $\pi_{P^{\bullet},Q^{\bullet}}^{-1}(\varphi^{\bullet})$ to $\pi_{P,Q}^{-1}(\varphi)$ commutes with this action, and hence is a bijection.

(b) Point (a) implies that there is a well defined functor $(-)^{\bullet} \colon \mathcal{L} \to \mathcal{L}^{\bullet}$ that sends $P$ to $P^{\bullet}$ and $\psi$ to $\psi^{\bullet}$. This defines a map $r \colon |\mathcal{L}| \to |\mathcal{L}^{\bullet}|$, which is a retraction by Proposition 1.4(b). Let $i$ denote the inclusion; then $i \circ r$ homotopic to the identity on $|\mathcal{L}|$ since there is a natural transformation of functors from $(-)^{\bullet}$ to the identity that sends each object $P$ to the inclusion $\iota_{P,P^{\bullet}}$ from $P$ to $P^{\bullet}$. □

We now combine Lemmas 1.2 and 1.5 to get the result we need.

**Proposition 1.6** *Let $\mathcal{F}$ be a saturated fusion system over a discrete $p$-toral group $S$, and let $\mathcal{L}_0 \subseteq \mathcal{L}$ be a pair of linking systems associated to $\mathcal{F}$. Then the inclusion of $|\mathcal{L}_0|$ in $|\mathcal{L}|$ is a homotopy equivalence.*

**Proof** Let $\boldsymbol{L}$ be the set of all linking subsystems $\mathcal{L}' \subseteq \mathcal{L}$ containing $\mathcal{L}_0$ such that the inclusion $|\mathcal{L}_0| \subseteq |\mathcal{L}'|$ is a homotopy equivalence. We must show that $\mathcal{L} \in \boldsymbol{L}$.

Assume otherwise, and choose $\mathcal{L}_1 \in \boldsymbol{L}$ for which $\mathrm{Ob}(\mathcal{L}_1^{\bullet})$ contains the largest possible number of $\mathcal{F}$-conjugacy classes. Let $\mathcal{L}_2 \subseteq \mathcal{L}$ be the full subcategory with $\mathrm{Ob}(\mathcal{L}_2) = \{P \in \mathrm{Ob}(\mathcal{L}) \mid P^{\bullet} \in \mathrm{Ob}(\mathcal{L}_1)\}$. By Lemma 1.5(b), $|\mathcal{L}_1^{\bullet}| = |\mathcal{L}_2^{\bullet}|$ is a deformation retract of $|\mathcal{L}_1|$ and of $|\mathcal{L}_2|$, so $|\mathcal{L}_1| \simeq |\mathcal{L}_2|$, and $\mathcal{L}_2 \in \boldsymbol{L}$. Since $\mathcal{L}_2 \subsetneqq \mathcal{L}$ by the assumption that $\mathcal{L} \notin \boldsymbol{L}$, we have $\mathrm{Ob}(\mathcal{L}^{\bullet}) \not\subseteq \mathrm{Ob}(\mathcal{L}_2)$. Let $P$ be maximal among objects in $\mathcal{L}^{\bullet}$ not in $\mathcal{L}_2$. By definition of $\mathrm{Ob}(\mathcal{L}_2)$, $P$ is maximal among all objects of $\mathcal{L}$ not in $\mathcal{L}_2$. Let $\mathcal{L}_3 \subseteq \mathcal{L}$ be the full subcategory with $\mathrm{Ob}(\mathcal{L}_3) = \mathrm{Ob}(\mathcal{L}_2) \cup P^{\mathcal{F}}$.

By Lemma 1.2, the inclusion of $|\mathcal{L}_2|$ into $|\mathcal{L}_3|$ is a homotopy equivalence. Hence $\mathcal{L}_3 \in \boldsymbol{L}$, contradicting our maximality assumption on $\mathcal{L}_1^{\bullet}$. □

## 2 Λ-functors

For any group $G$, let $\mathcal{O}_p(G)$ be the *p-subgroup orbit category* of $G$: the category whose objects are the $p$-subgroups of $G$, and where

$$\mathrm{Mor}_{\mathcal{O}_p(G)}(P, Q) = \{Qg \in Q \backslash G \mid {}^g P \le Q\} = Q \backslash \{g \in G \mid {}^g P \le Q\}.$$

Note that there is a bijection

$$\mathrm{Mor}_{\mathcal{O}_p(G)}(P, Q) \xrightarrow{\cong} \mathrm{map}_G(G/P, G/Q)$$

that sends a coset $Qg$ to the $G$-map $(xP \mapsto xg^{-1}Q)$.

**Definition 2.1** Let $G$ be a locally finite group, and let $M$ be a $\mathbb{Z}G$-module. Define a functor

$$F_M : \mathcal{O}_p(G)^{\mathrm{op}} \to \mathrm{Ab}$$

by setting

$$F_M(P) = \begin{cases} M & \text{if } P = 1, \\ 0 & \text{if } P \ne 1, \end{cases}$$

for each $p$-subgroup $P \le G$. Here, $\mathrm{Aut}_{\mathcal{O}_p(G)}(1) \cong G$ has the given action on $M$. Set

$$\Lambda^*(G; M) = \varprojlim_{\mathcal{O}_p(G)}{}^*(F_M).$$

We refer to [3, Definition A.1] for the definition of a transporter system. As in Section 1, when $S$ is a group and $\mathcal{H}$ is a set of subgroups of $S$, we let $\mathcal{T}_{\mathcal{H}}(S)$ be the transporter category of $S$ with object set $\mathcal{H}$, where $\mathrm{Mor}_{\mathcal{T}_{\mathcal{H}}(S)}(P, Q)$ is the set of all $g \in S$ such that ${}^g P \le Q$. A transporter system associated to a fusion system $\mathcal{F}$ over a discrete $p$-toral group $S$ consists of a category $\mathcal{T}$ whose objects are subgroups of $S$, together with a pair of functors

$$\mathcal{T}_{\mathrm{Ob}(\mathcal{T})}(S) \xrightarrow{\varepsilon} \mathcal{T} \xrightarrow{\rho} \mathcal{F},$$

such that $\varepsilon$ is the identity on objects and injective on morphism sets, $\rho$ is the inclusion on objects and surjective on morphism sets, and several other conditions are satisfied. The only requirements on the set $\mathrm{Ob}(\mathcal{T})$ are that it be nonempty and closed under $\mathcal{F}$-conjugacy and overgroups.

As with linking systems, when $P \le Q$ are objects in a transporter system $\mathcal{T}$, we write $\iota_{P,Q} = \varepsilon_{P,Q}(1) \in \mathrm{Mor}_{\mathcal{T}}(P, Q)$ for the inclusion morphism. By "extensions" and "restrictions" of morphisms we mean extensions and restrictions of source and target both, with respect to these inclusions. We will only need to refer to the following properties:

**Proposition 2.2** *The following hold for any transporter system $\mathcal{T}$ associated to a saturated fusion system $\mathcal{F}$ over a discrete $p$-toral group $S$.*

(a) *For each $P, Q \in \mathrm{Ob}(\mathcal{T})$, the composite $\rho_{P,Q} \circ \varepsilon_{P,Q}$ sends $g \in N_S(P, Q)$ to the homomorphism $c_g = (x \mapsto {}^g x) \in \mathrm{Hom}_{\mathcal{F}}(P, Q)$.*

(b) *Let $P \trianglelefteq \overline{P} \leq S$ and $Q \leq \overline{Q} \leq S$ be objects in $\mathcal{T}$. Let $\psi \in \mathrm{Iso}_{\mathcal{T}}(P, Q)$ be such that for each $g \in \overline{P}$, there is $h \in \overline{Q}$ satisfying $\iota_{Q,\overline{Q}} \circ \psi \circ \varepsilon_P(g) = \varepsilon_{Q,\overline{Q}}(h) \circ \psi$. Then $\psi$ extends to a unique morphism $\overline{\psi} \in \mathrm{Mor}_{\mathcal{T}}(\overline{P}, \overline{Q})$ such that $\overline{\psi}|_{P,Q} = \psi$.*

(c) *All morphisms in $\mathcal{T}$ are monomorphisms and epimorphisms in the categorical sense.*

**Proof** Point (a) is axiom (B) in [3, Definition A.1], and point (c) is shown in [3, Proposition A.2(d)].

The existence of an extension in point (b) is axiom (II) in [3, Definition A.1], except that it is stated there under the additional assumption that $Q$ is normal in $\overline{Q}$. But if $Q \leq \overline{Q}$ is not normal, and $\psi \in \mathrm{Iso}_{\mathcal{T}}(P, Q)$, $g \in \overline{P}$, and $h \in \overline{Q}$ are as above, then

$$Q = \mathrm{Im}(\rho(\iota_{Q,\overline{Q}} \circ \psi \circ \varepsilon_P(g))) = \mathrm{Im}(\rho(\varepsilon_{Q,\overline{Q}}(h) \circ \psi)) = {}^h Q,$$

the last equality by (a), and so $h \in N_{\overline{Q}}(Q)$. Hence $\psi$ extends to $\widehat{\psi} \in \mathrm{Mor}_{\mathcal{T}}(\overline{P}, N_{\overline{Q}}(Q))$ by axiom (II) as stated in [3], and we can take $\overline{\psi} = \iota_{N_{\overline{Q}}(Q),\overline{Q}} \circ \widehat{\psi}$.

The extension in (b) is unique since inclusion morphisms in $\mathcal{T}$ are epimorphisms by (c). $\square$

If $\mathcal{T}$ is a transporter system over a discrete $p$-toral group $S$, then its orbit category $\mathcal{O}(\mathcal{T})$ is the category with the same objects, and where for each $P, Q \in \mathrm{Ob}(\mathcal{T})$,

$$\mathrm{Mor}_{\mathcal{O}(\mathcal{T})}(P, Q) = \varepsilon_Q(Q)\backslash\mathrm{Mor}_{\mathcal{T}}(P, Q).$$

Thus, for example, if $\mathcal{T}$ is the transporter system of a group $G$ (with some set of objects), then $\mathcal{O}(\mathcal{T})$ is the orbit category of $G$ in the above sense.

**Proposition 2.3** *Let $\mathcal{T}$ be a transporter system associated to a saturated fusion system $\mathcal{F}$ over a discrete $p$-toral group $S$. Fix $P \in \mathrm{Ob}(\mathcal{T})$, and let*

$$\Phi \colon \mathcal{O}(\mathcal{T})^{\mathrm{op}} \to \mathrm{Ab}$$

*be a functor such that $\Phi(Q) = 0$ for each $Q \notin P^{\mathcal{F}}$. Then*

$$\varprojlim_{\mathcal{O}(\mathcal{T})}{}^*(\Phi) \cong \Lambda^*(\mathrm{Aut}_{\mathcal{O}(\mathcal{T})}(P); \Phi(P)).$$

**Proof** By axiom (I) for a transporter system (see [3, Definition A.1]), there is $Q \in P^{\mathcal{F}}$ such that the index of $\varepsilon_Q(N_S(Q))$ in $\mathrm{Aut}_{\mathcal{T}}(Q)$ is finite and prime to $p$. Thus $\mathrm{Aut}_{\mathcal{T}}(Q)$ is an extension of a discrete $p$-toral group by a finite group, and we write $\varepsilon_Q(N_S(Q)) \in \mathrm{Syl}_p(\mathrm{Aut}_{\mathcal{T}}(Q))$ for short. Since $\Lambda^*(\mathrm{Aut}_{\mathcal{O}(\mathcal{T})}(Q); \Phi(Q)) \cong \Lambda^*(\mathrm{Aut}_{\mathcal{O}(\mathcal{T})}(P); \Phi(P))$ for each $Q \in P^{\mathcal{F}}$, it suffices to prove the proposition when $\varepsilon_P(N_S(P)) \in \mathrm{Syl}_p(\mathrm{Aut}_{\mathcal{T}}(P))$. We will show that this is a special case of [2, Proposition 5.3].

Set $\Gamma = \mathrm{Aut}_{\mathcal{O}(\mathcal{T})}(P) = \varepsilon_P(P)\backslash\mathrm{Aut}_{\mathcal{T}}(P)$. To simplify the notation, we identify $N_S(P)$ with its image $\varepsilon_P(N_S(P)) \leq \mathrm{Aut}_{\mathcal{T}}(P)$, and identify $N_S(P)/P$ with $\varepsilon_P(N_S(P))/\varepsilon_P(P)$.

Let $\mathcal{H}$ be the set of all subgroups of $N_S(P)/P \in \mathrm{Syl}_p(\Gamma)$. Define

$$\alpha \colon \mathcal{O}_{\mathcal{H}}(\Gamma) \to \mathcal{O}(\mathcal{T})$$

by setting $\alpha(Q/P) = Q$, and

$$\alpha(Q/P \xrightarrow{R\gamma} R/P) = [\bar{\gamma}] \in \mathrm{Mor}_{\mathcal{O}(\mathcal{T})}(Q, R)$$

where $\gamma \in \mathrm{Aut}_{\mathcal{T}}(P)$ extends to a unique morphism $\bar{\gamma} \in \mathrm{Mor}_{\mathcal{T}}(Q, R)$ by Proposition 2.2(b).

The proposition will follow immediately from [2, Proposition 5.3] once we have shown that conditions (a)–(d) in the proposition all hold. Set $c_0 = \alpha(1)$, following the notation used in that proposition, and note that $c_0 = P$.

(a)   For each $P \in \mathrm{Ob}(\mathcal{T})$ and each $\psi \in \mathrm{End}_{\mathcal{T}}(P)$, the homomorphism $\rho(\psi) \in \mathrm{End}_{\mathcal{F}}(P)$ is an isomorphism of groups, and hence $\psi \in \mathrm{Aut}_{\mathcal{T}}(P)$ by [3, Proposition A.2(c)]. So by construction, $\alpha$ sends $\Gamma = \mathrm{Aut}_{\mathcal{O}_{\mathcal{H}}(\Gamma)}(1)$ isomorphically to $\mathrm{End}_{\mathcal{O}(\mathcal{T})}(P) = \mathrm{Aut}_{\mathcal{O}(\mathcal{T})}(P)$.

(b)   Let $U \in \mathrm{Ob}(\mathcal{T})$ be such that $U \notin P^{\mathcal{F}}$. We must show that all isotropy subgroups of the $\Gamma$-action on $\mathrm{Mor}_{\mathcal{O}(\mathcal{T})}(P, U)$ are nontrivial and conjugate in $\Gamma$ to subgroups in $\mathcal{H}$. Since $\mathcal{H}$ is the set of all subgroups of a Sylow $p$-subgroup of $\Gamma$, where $\Gamma$ has a discrete $p$-toral subgroup of finite index, this means showing that each isotropy subgroup is a nontrivial discrete $p$-toral subgroup.

Fix $\psi \in \mathrm{Mor}_{\mathcal{T}}(P, U)$, and let $[\psi]$ be its class in $\mathrm{Mor}(\mathcal{O}(\mathcal{T}))$. Set $Q = \rho(\psi)(P) < U$. Then $\psi = \iota_{Q,U} \circ (\psi|_{P,Q})$, where $\psi|_{P,Q} \in \mathrm{Iso}_{\mathcal{T}}(P, Q)$. So the isotropy subgroup for the action of $\Gamma$ on $\psi$ is isomorphic to the isotropy subgroup of $\mathrm{Aut}_{\mathcal{O}(\mathcal{T})}(Q)$ on $\iota_{Q,U}$.

For each $\gamma \in \mathrm{Aut}_{\mathcal{T}}(Q)$, we have $[\iota_{Q,U}\gamma] = [\iota_{P,U}]$ if and only if there is $g \in U$ such that $^{g}Q = Q$ and $\gamma = \varepsilon_U(g)|_{Q,Q} = \varepsilon_Q(g)$. Thus the isotropy subgroup is the group of all $[\varepsilon_Q(g)]$ for $g \in N_U(Q)$, and hence is isomorphic to $N_U(Q)/Q$. This last group is a nontrivial discrete $p$-toral group by [2, Lemma 1.8] and since $Q < U$ are both discrete $p$-toral groups.

(c)   We claim that all morphisms in $\mathcal{O}(\mathcal{T})$ are epimorphisms in the categorical sense. To see this, fix subgroups $P, Q, R \in \mathrm{Ob}(\mathcal{T})$ and morphisms $\varphi \in \mathrm{Mor}_{\mathcal{T}}(P, Q)$ and $\alpha, \beta \in \mathrm{Mor}_{\mathcal{T}}(Q, R)$ such that $[\alpha][\varphi] = [\beta][\varphi]$, where $[-]$ denotes the class in $\mathcal{O}(\mathcal{T})$ of a morphism in $\mathcal{T}$. Thus there is $g \in R$ such that $\alpha\varphi = \varepsilon_R(g)\beta\varphi$ in $\mathcal{T}$.

By Proposition 2.2(c), all morphisms in $\mathcal{T}$ are epimorphisms. So $\alpha = \varepsilon_R(g)\beta$, and hence $[\alpha] = [\beta]$, proving that $\varphi$ is an epimorphism in $\mathcal{O}(\mathcal{T})$.

(d)   Fix $Q/P \leq N_S(P)/P$, $U \in \mathrm{Ob}(\mathcal{T})$, and $\varphi \in \mathrm{Mor}_{\mathcal{T}}(P, U)$ such that $[\varphi\varepsilon_P(g)] = [\varphi]$ in $\mathrm{Mor}_{\mathcal{O}(\mathcal{T})}(P, U)$ for each $g \in Q$. We must show that there is $\bar{\varphi} \in \mathrm{Mor}_{\mathcal{T}}(Q, U)$ such that $[\bar{\varphi}|_{P,U}] = [\varphi]$ in $\mathrm{Mor}_{\mathcal{O}(\mathcal{T})}(P, U)$.

By assumption, for each $g \in Q$, there is $u \in U$ such that $\varepsilon_U(u)\varphi = \varphi\varepsilon_P(g)$. Set $U_0 = \rho(\varphi)(P) \leq U$. Thus $\varphi = \iota_{U_0,U}\varphi_0$, where $\varphi_0 = \varphi|_{P,U_0} \in \mathrm{Iso}_{\mathcal{T}}(P, U_0)$. So

$$\varepsilon_U(u)\iota_{U_0,U}\varphi_0 = \iota_{U_0,U}\varphi_0\varepsilon_P(g).$$

Hence $U_0 = \mathrm{Im}(\rho(\varphi_0\varepsilon_P(g))) = \mathrm{Im}(\rho(\varepsilon_U(u)\iota_{U_0,U}\varphi_0)) = {}^{u}U_0$ (the last equality by Proposition 2.2(a)), and so $u \in N_U(U_0)$. Thus $\iota_{U_0,U}\varphi_0\varepsilon_P(g)\varphi_0^{-1} = \varepsilon_U(u)\iota_{U_0,U} = \iota_{U_0,U}\varepsilon_{U_0}(u)$. Since morphisms in a transporter category are monomorphisms by Proposition 2.2(c), this implies that $\varphi_0\varepsilon_P(g)\varphi_0^{-1} \in \varepsilon_{U_0}(N_U(U_0))$. So by Proposition 2.2(b), there is $\bar\varphi \in \mathrm{Mor}_{\mathcal{T}}(Q,U)$ such that $\bar\varphi|_{P,U_0} = \varphi_0$, and hence $\bar\varphi|_{P,U} = \varphi$. $\square$

## Acknowledgements

## References

[1] **C Broto**, **N Castellana**, **J Grodal**, **R Levi**, **B Oliver**, *Subgroup families controlling p-local finite groups*, Proc. London Math. Soc. 91 (2005) 325–354 MR

[2] **C Broto**, **R Levi**, **B Oliver**, *Discrete models for the p-local homotopy theory of compact Lie groups and p-compact groups*, Geom. Topol. 11 (2007) 315–427 MR

[3] **C Broto**, **R Levi**, **B Oliver**, *An algebraic model for finite loop spaces*, Algebr. Geom. Topol. 14 (2014) 2915–2981 MR

[4] **A Gonzalez**, *Finite approximations of p-local compact groups*, Geom. Topol. 20 (2016) 2923–2995 MR

[5] **D Quillen**, *Higher algebraic K-theory, I*, from "Algebraic K-theory, I: Higher K-theories" (H Bass, editor), Lecture Notes in Math. 341, Springer (1973) 85–147 MR

*Departament de Matemàtiques, Universitat Autònoma de Barcelona*
*Bellaterra, Spain*

*Institute of Mathematics, University of Aberdeen*
*Aberdeen, United Kingdom*

*LAGA, Institut Galilée, Université Sorbonne Paris Nord*
*Villetaneuse, France*

carles.broto@uab.cat,  r.levi@abdn.ac.uk,  bobol@math.univ-paris13.fr

# 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.