Geometry & Topology

Volume 21 (2017)

Issue 5 (pages 2557–3190)

msp

GEOMETRY & TOPOLOGY

msp.org/gt

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

The subscription price for 2017 is US $495/year for the electronic version, and $730/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Geometry & Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Geometry & Topology (ISSN 1465-3060 printed, 1364-0380 electronic) at Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published 6 times per year and continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

GT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

Geometry & Topology 21 (2017) 2557–2600

A geometric construction of colored HOMFLYPT homology

BEN WEBSTER
GEORDIE WILLIAMSON

The aim of this paper is twofold. First, we give a fully geometric description of the HOMFLYPT homology of Khovanov and Rozansky. Our method is to construct this invariant in terms of the cohomology of various sheaves on certain algebraic groups, in the same spirit as the authors' previous work on Soergel bimodules. All the differentials and gradings which appear in the construction of HOMFLYPT homology are given a geometric interpretation.

In fact, with only minor modifications, we can extend this construction to give a categorification of the colored HOMFLYPT polynomial, *colored HOMFLYPT homology*. We show that it is in fact a knot invariant categorifying the colored HOMFLYPT polynomial and that it coincides with the categorification proposed by Mackaay, Stošić and Vaz.

17B10, 57T10

1 Introduction

The *colored HOMFLYPT polynomial* is an invariant of links together with a labeling or "coloring" of each component with a positive integer; in particular, for knots, there is an invariant for each positive integer. Its most important properties are that

- it reduces to the usual HOMFLYPT polynomial when all labels are 1, and
- colored HOMFLYPT encapsulates all Reshetikhin–Turaev invariants for the link labeled with wedge powers of the standard representation of $\mathfrak{sl}_n$, just as the HOMFLYPT polynomial does for the standard representation alone.

In this paper we give a geometric construction of a categorification of this invariant, *colored HOMFLYPT homology.* Like the HOMFLYPT homology of Khovanov and Rozansky [13], this associates a triply graded vector space to each colored link such that the bigraded Euler characteristic is the colored HOMFLYPT polynomial. In fact, we produce an infinite sequence of such invariants, one for each page of a spectral sequence, but only the first and second pages are connected via an Euler characteristic to a known classical invariant.

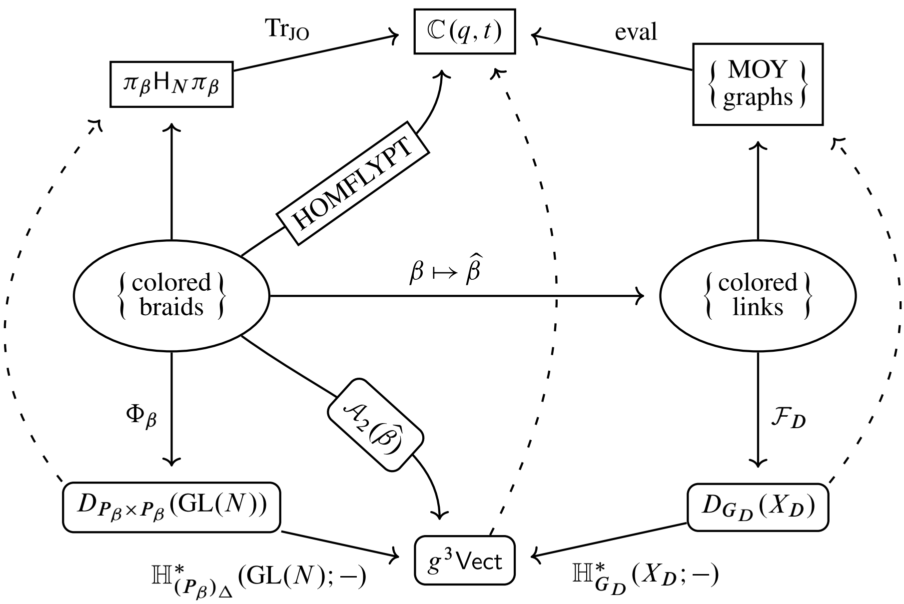
Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2557

Our construction and proofs of invariance and categorification are algebro-geometric in nature. As a special case we obtain a new and entirely geometric interpretation of Khovanov's Soergel bimodule construction of HOMFLYPT homology [12].

We also show that this invariant has a purely combinatorial description via the Hochschild homology of bimodules analogous to that of Khovanov. In fact, it coincides with the link homology proposed from an algebraic perspective by Mackaay, Stošić and Vaz [17]. Thus, the main result of our paper has an entirely algebraic statement:

Theorem 1.1 *The colored HOMFLYPT homology defined in [17] is a knot invariant, and its Euler characteristic is the colored HOMFLYPT polynomial.*

Our definition also has the advantage of categorifying essentially all algebraic objects involved in the definition of colored HOMFLYPT homology. Let us give a schematic diagram for the pieces here, with actual operations given by solid arrows, and (de)categorifications given by dashed ones:



The top half of the diagram shows two different definitions of the colored HOMFLYPT polynomial:

- The path through {MOY graphs} is the description of the colored HOMFLYPT polynomial by Murakami, Ohtsuki and Yamada [19]: one associates to a link diagram a sum of weighted trivalent graphs, and then defines an evaluation function on such graphs, which in turn gives a state sum interpretation of the colored HOMFLYPT polynomial. While the paper [19] only considers certain specializations of the HOMFLYPT polynomial, their technique is easily extended to the polynomial itself.

- The path through $\pi_\beta \mathsf{H}_N \pi_\beta$ is described by Lin and Zheng [16]: to each closable colored braid β, we have an associated element of the Hecke algebra H_N, where N is the colored braid index of β (the sum of the colorings of the strands). In fact, this element lies in a certain subalgebra $\pi_\beta \mathsf{H}_N \pi_\beta$, where π_β is a projection which depends on the coloring of β. The colored HOMFLY polynomial is obtained by applying a certain trace $\mathrm{Tr}_{\mathrm{JO}}$ defined by Ocneanu [11] on H_N.

In this paper, we show how to categorify both of these paths, as is schematically indicated in the bottom half of the diagram, and briefly summarized in Section 1.2. The final result of this construction is a knot invariant $\mathcal{A}_2(\widehat{\beta})$; we show that this invariant is well-defined in Theorem 1.2 and that it agrees with HOMFLYPT homology in Theorem 1.4.

- The leftmost dashed arrow is the isomorphism of $\pi_\beta \mathsf{H}_N \pi_\beta$ with the Grothendieck group of sheaves on $\mathrm{GL}(N)$ which are bi-equivariant for the left and right multiplication of a subgroup of block upper triangular matrices P_β.

- The rightmost dashed arrow can be described as follows: to each link diagram D, we associate a group G_D, a G_D–variety X_D and a G_D–equivariant perverse sheaf whose the composition factors are in bijection with the MOY graphs arising from this link diagram.

- The central dashed arrow simply indicates taking bigraded Euler characteristic of a trigraded vector space with respect to one of its gradings.

We must also show that this diagram, including the dashed arrows, "commutes". This follows from a result of the authors giving a similar construction of a Markov trace for the Hecke algebra of any semisimple Lie group, shown in the paper [27].

As should be clear from the above, the techniques we use are those of algebraic geometry and geometric representation theory. While these are not familiar to the average topologist, we have striven to make this paper accessible to the novice, at least if they are willing to accept a few deep results as black boxes. As a general rule, our actual calculations are simple and quite geometric in nature; however, we must cite rather serious machinery to show that these calculations are meaningful.

1.1 The geometric machinery

Let us briefly indicate the geometric setting in which we work. All material covered here is discussed in greater detail in Section 3.

Let X be an algebraic variety defined by equations with integer coefficients. (In this paper, our varieties are built from copies of the general linear group, so we can always describe them in terms of integral equations.) To X one may associate a derived category $D^b(X)$ of sheaves with constructible cohomology. There are numerous technicalities in the construction of this category, but we postpone discussion of these until Section 3.

The category $D^b(X)$ behaves similarly to the bounded derived category of constructible sheaves on the complex algebraic variety defined by these equations. However, since we used integral equations, we have an alternate perspective on these varieties; one can also reduce modulo a prime p, and work over the finite field $\mathbb{F}_p$. The objects in $D^b(X)$ can also be interpreted as sheaves on these varieties in characteristic p, and for technical reasons, this is the perspective we will take. In this situation, there is an extra structure which helps us to understand our complexes of sheaves: an action of the Frobenius Fr on our variety.

The category $D^b(X)$ contains a remarkable abelian subcategory $P(X)$ of "mixed perverse sheaves". For us the most important feature of $P(X)$ is that every object of $P(X)$ has a canonical "weight filtration" with semisimple subquotients, which is defined using the Frobenius.

As with any filtration, this leads to a spectral sequence

$$E_1^{p,q} = \mathbb{H}^{p+q}(\mathrm{gr}^W_{-p}\mathcal{F}) \implies \mathbb{H}^{p+q}(\mathcal{F}).$$

Each term on the left hand side also carries an action of Frobenius induced by that on the variety. Considering the norms of the eigenvalues of Frobenius may be used to give an additional grading to each page of the spectral sequence. It follows that each page of the spectral sequence is *triply* graded.

We will need to consider a generalization of this category, which is a version of equivariant sheaves for the action of an affine algebraic group on X. While in principle, the technical difficulties of understanding such a category could be resolved by working in the category of stacks, we have found it less burdensome to give a careful definition of the mixed equivariant derived category from a more elementary perspective. For the sake of brevity, this has been done in a separate note [25].

1.2 Application to knot theory

In order to apply the above machinery to knot theory, we must define a sheaf associated to a link. More precisely, as we discuss in Section 2, to any projection D of a colored link, we associate the natural graph with vertices given by crossings and edges by arcs.

To this graph, we associate a variety X_D together with the action of a reductive group G_D. Remembering the crossings in D allows us to construct a G_D–equivariant mixed shifted perverse sheaf $\mathcal{F}_D \in D^b_{G_D}(X_D)$. We then show that $\mathcal{F}_D$ may be used to construct a series of knot invariants.

Associated to any filtration on $\mathcal{F}_D$ (as a perverse sheaf), we have a canonical spectral sequence converging to $\mathbb{H}^*_{G_D}(X_D;\mathcal{F}_D)$. Furthermore, we can endow $\mathbb{H}^*_{G_D}(X_D;-)$ of any mixed sheaf with the weight grading. This is a grading which is preserved by all spectral sequence differentials, so we can think of any page of this spectral sequence as a triply graded vector space, where two gradings are given by the usual spectral sequence structure, and the third by weight.

The sheaf $\mathcal{F}_D$ carries a natural weight filtration. This is easily confused with, but distinct from, the weight grading discussed above.[1] We call the spectral sequence associated to this weight filtration *chromatographic*.

Theorem 1.2 *If D is the diagram of a closed braid, then every page E_i for $i \geq 2$ of the spectral sequence computing $\mathbb{H}^*_{G_D}(X_D;\mathcal{F}_D)$ associated to the weight filtration is an invariant of the underlying link L, up to an overall shift in the grading. We let $\mathcal{A}_i(\widehat{\beta})$ be the i^{th} page of this spectral sequence.*

If D is not a closed braid, then this theorem fails, since $\mathcal{A}_i(\widehat{\beta})$ can be changed by the Reidemeister IIb move; above we are using the fact that by the Markov theorem, any two braid closure diagrams for the same knot can be related without using this move.

This description has a similar flavor to that of Khovanov and Rozansky [13] or Bar-Natan [2]: it begins by assigning a simple object to a single crossing, and then an algebraic rule for gluing crossings together (this process can be formalized as an object called a *canopolis* as introduced by Bar-Natan [2]; we will discuss this perspective in Section 6.2). However, other papers, such as [12] or [17], have used a description which depended on the link diagram chosen being a closed braid. In order to show that our invariants coincide with those of [17], we must find a geometric description of this form.

Assume that β is a closable colored braid with coloring given by positive integers, $\widehat{\beta}$ its closure and let N be the colored braid index (the sum of the colorings over the strands of the braid). Let P_β be the block upper triangular matrices inside $G_N := \mathrm{GL}(N)$ with the sizes of the blocks given by the coloring of the strands of β. Using left and right multiplication, we obtain a natural $P_\beta \times P_\beta$ action on G_N. We let $(P_\beta)_\Delta$ be

[1] The weight grading mentioned above comes from the weight filtration of the pushforward of $\mathcal{F}_D$ to a point.

the diagonal subgroup, which acts on G_N by conjugation. By the classical theory of characteristic classes, we have a canonical isomorphism of $H^*(BP_\beta)$ to partially symmetric polynomials corresponding to the block sizes of P_β, which we will use freely from now on.

Theorem 1.3 *For each β, there is a $(P_\beta \times P_\beta)$–equivariant complex of sheaves Φ_β on G_N with a natural filtration, such that the associated spectral sequence computing $\mathbb{H}^*_{(P_\beta)_\Delta}(G_N; \Phi_\beta)$ is canonically isomorphic to the spectral sequence obtained from the weight filtration for $\mathbb{H}^*_{G_{\widehat{\beta}}}(X_{\widehat{\beta}}; \mathcal{F}_{\widehat{\beta}})$.*

Moreover, we have an isomorphism of the E_1 page $\mathcal{A}_1(\widehat{\beta})$ of the spectral sequence for the hypercohomology $\mathbb{H}^*_{P_\beta \times P_\beta}(G_N; \Phi_\beta)$, as a complex of bimodules over $H^*(BP_\beta)$, to the complex of singular Soergel bimodules considered by Mackaay, Stošić and Vaz in [17, Section 8].

Singular Soergel bimodules have been defined and classified in the thesis of the second author [28] and in the context of Harish-Chandra bimodules by Stroppel [23]. Since previous work of the authors [26] has related Hochschild homology to conjugation equivariant cohomology, we can identify our geometric knot invariant in terms of such bimodules.

Theorem 1.4 *If D is a closed braid, then the E_2 page of our spectral sequence is the categorification of the colored HOMFLYPT polynomial proposed in* [17].

If all the labels on the components of D are 1, then this agrees with the triply graded link homology as defined by Khovanov and Rozansky in [13].

The higher pages of this spectral sequence are not easy to compute, and it is not known what their Euler characteristics are. Whether they correspond to any classical link invariant is unknown.

Acknowledgements We would like to thank: Wolfgang Soergel for his observation that "Komplexe von Bimoduln sind die Gewichtsfiltrierung des armen Mannes" ("*Complexes of bimodules are the poor man's weight filtration*"), which formed a starting point for this work; Marco Mackaay for suggesting that it could be generalized to the colored case and explaining the constructions of [17]; Raphaël Rouquier and Olaf Schnürer for illuminating discussions; and Catharina Stroppel, Noah Snyder and Carl Mautner for comments on an earlier version of this paper. Part of this research was conducted whilst Williamson took part in the program "Algebraic Lie Theory" at the Isaac Newton Institute, Cambridge. Webster was supported by an NSF Postdoctoral Fellowship.

2 Description of the varieties

We start by recalling the steps involved in our categorification, beginning with a braidlike diagram D of an oriented colored link L:

- To D (with its coloring) we associate a reductive group G_D together with a G_D–variety X_D, which only depends on the graph obtained from the diagram D by forgetting the distinction between under and overcrossings.
- The crossing data allows us to define a G_D–equivariant sheaf $\mathcal{F}_D$ on X_D.
- This sheaf $\mathcal{F}_D$ has a chromatographic spectral sequence converging to the G_D–equivariant hypercohomology of $\mathcal{F}_D$.
- Each page E_i of this spectral sequence for $i \geq 2$ is a knot invariant (up to overall shift) and the E_2 page categorifies the colored HOMFLYPT polynomial.

In this section we discuss the first step.

First let us fix some notation. We fix a chain of vector spaces $0 \subset V_1 \subset V_2 \subset V_3 \subset \cdots$ over $\mathbb{F}_q$ such that $\dim V_i = i$ for all i. Let

$$G_{i_1,\dots,i_n} := \mathrm{GL}(i_1) \times \cdots \times \mathrm{GL}(i_n),$$

and let $P_{i_1,\dots,i_n}$ be the block upper triangular matrices with blocks $\{i_1, \dots, i_n\}$. We may identify $P_{i_1,i_2,\dots,i_n}$ with the stabilizer in $G_{i_1+\cdots+i_n}$ of the standard partial flag

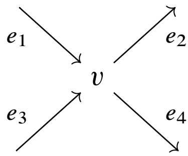
$$\{0 \subset V_{i_1} \subset V_{i_1+i_2} \subset \cdots \subset V_{i_1+\cdots+i_n}\}.$$

Let D be a diagram of an oriented tangle with marked points, with no marked points occurring at a crossing. Let Γ be the oriented graph obtained by the diagram's projection, with vertices corresponding to crossings and marked points in D. That is, we simply forget the over and undercrossings in D. We deal with the exterior ends of the tangle in a somewhat unconventional manner; we do not think of them as vertices in the graph, so we think of the arcs connecting to the edge as connecting to 1 or 0 vertices. By adding marked points to D if necessary, we may assume that every component of Γ contains at least one vertex.

Recall that to the diagram D we wish to associate a variety X_D acted on by an algebraic group G_D. Let us write $\mathcal{E}(D)$ and $\mathcal{V}(D)$ for the edges and vertices of Γ respectively. Given an edge $e \in \mathcal{E}(D)$ write G_e for G_i, where i is the label on e. Similarly, given $v \in \mathcal{V}(D)$ write G_v for G_i, where i is the sum of the labels on the incoming vertices at v (which necessarily equals the sum of the labels on the outgoing vertices). We define

$$X_D := \prod_{v \in \mathcal{V}(D)} G_v \quad \text{and} \quad G_D := \prod_{e \in \mathcal{E}(D)} G_e.$$

It remains to describe how G_D acts on X_D. Locally, near any crossing, Γ is isotopic to a graph of the form:



We will call e_1 and e_2 *upper* and e_3 and e_4 *lower* edges with respect to the vertex v. Whenever a vertex v lies on an edge e we define an inclusion map $i_e\colon G_e \to G_v$, which is the identity if v corresponds to a marked point, and is the composition of the canonical inclusions

$$\begin{cases} G_i \hookrightarrow G_{i,j} \hookrightarrow G_{i+j} & \text{if } e \text{ is upper,} \\ G_i \hookrightarrow G_{j,i} \hookrightarrow G_{i+j} & \text{if } e \text{ is lower.} \end{cases}$$

That is, G_e is included as the upper left or lower right block matrices in G_v, according to whether e is upper or lower.

We now describe how G_D acts on X_D by describing the action componentwise. Let $g \in G_e$ and $x \in G_v$. We have

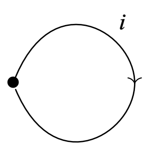
$$g \cdot x = i_e(g)^{\omega} x i_e(g)^{-\alpha},$$

where $\omega = 1$ if e is incoming at v, and 0 otherwise, and $\alpha = 1$ if e is outgoing at v and 0 otherwise. Thus, we have

$$g \cdot x = \begin{cases} x & \text{if } v \text{ does not lie on } e, \\ x i_e(g)^{-1} & \text{if } e \text{ is only outgoing at } v, \\ i_e(g) x & \text{if } e \text{ is only incoming at } v. \\ i_e(g) x i_e(g)^{-1} & \text{if } e \text{ forms a loop at the vertex } v. \end{cases}$$

Example 2.1 Here are two examples of X_D and G_D.

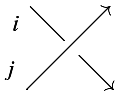
- Let D be the standard diagram of the unknot labeled i with one marked point:



Then we have $X_D = G_D = G_i$ and G_D acts on X_D by conjugation.

- If D is an oriented arc with a single marked point, then we have $X_D = G_i \times G_i$, and $G_D = G_i$, with the action $g \cdot (a, b) = (ag^{-1}, gb)$, where a corresponds to the arc leaving the marked point, and b to the arc entering it.

- Let D be the diagram of an (i,j)–crossing:



Here $X_D = G_{i+j}$ and $G_D = G_i \times G_j \times G_j \times G_i$ and (a,b,c,d) acts on $x \in G_{i+j}$ by

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} x \begin{pmatrix} ccc^{-1} & 0 \\ 0 & d^{-1} \end{pmatrix}.$$

Note that if we glue two open edges with label i of the tangle diagram D together to make a new diagram D' by adding a marked point, then the spaces $X_D = X_{D'}$ are isomorphic, but $G_{D'} = G_D \times G_i$, with the new factor acting on the factors in X_D corresponding to the glued edges.

The group G_i attached to a marked point acts freely if the two connected edges don't close into a loop, and removing this point simply quotients both X_D and G_D by this G_i, leaving the equivariant geometry unchanged. Combining these observations with the examples above is enough to construct X_D and G_D for any tangle diagram.

This is the variety and group that we shall use in our construction. But before defining our invariant, we must first cover some generalities on categories of sheaves on these varieties.

3 Mixed and equivariant sheaves

In the rest of this paper, we will be using the machinery of mixed equivariant sheaves. In this section we intend to summarize the essential features of the theory that are necessary for us, and to indicate to the reader where the details can be found.

3.1 Weight grading

An important point underlying what follows is that cohomology of a complex algebraic variety (as well as most variations, such as equivariant cohomology, or intersection cohomology) has an additional natural grading, the *weight grading*. This grading is difficult to describe explicitly without using methods over characteristic p (as we will later), but is best understood by two simple properties:

- The weight grading is preserved by cup products, pullback and all maps in long exact sequences (in fact, by all differentials in any Serre spectral sequence).
- This weight grading is equal to the cohomological grading on smooth projective varieties.

Example 3.1 (cohomology of $\mathbb{C}^*$) If we write $\mathbb{CP}^1$ as the union of $\mathbb{C}$ and $\mathbb{CP}^1-\{0\}$, then in the Mayer–Vietoris sequence we have an isomorphism $H^2(\mathbb{CP}^1)\cong H^1(\mathbb{C}^*)$. Thus, the cohomological and weight gradings do not agree on $H^1(\mathbb{C}^*)$.

We plan to describe homological knot invariants using the equivariant cohomology of varieties, and the weight grading will be necessary to give all the gradings we expect on our knot homology.

3.2 Sheaves and perverse sheaves

We must use a generalization of the weight grading, the weight filtration on a mixed perverse sheaf. References for this section include [1], [9], [5] and [14]. Although we will not use it below, we should also point out that there is a way to understand mixed perverse sheaves which only uses characteristic-0 methods (Saito's mixed Hodge modules [22]; see the book of Peter and Steenbrink [21]).

Let $q=p^e$ be a prime power. We consider throughout a finite field $\mathbb{F}_q$ with q elements, and an algebraic closure $\mathbb{F}$ of $\mathbb{F}_q$. Unless we state otherwise all varieties and morphisms will be defined over $\mathbb{F}_q$. Given a variety X we will write $X\otimes\mathbb{F}$ for its extension of scalars to $\mathbb{F}$.

We fix a prime number $\ell\neq p$ and let $\Bbbk$ denote the algebraic closure $\overline{\mathbb{Q}}_\ell$ of the field of ℓ–adic numbers. We should note that the choices of p, q and ℓ are purely auxiliary; the resulting knot homology will be independent of all of these choices. Throughout we fix a square root of q in $\Bbbk$ and denote it by $q^{1/2}$. Given a variety Y defined over $\mathbb{F}_q$ or $\mathbb{F}$ we denote by $D^b(Y)$ (resp. $D^+(Y)$) the bounded (resp. bounded below) derived category of constructible $\Bbbk$–sheaves on Y; see [9]. By abuse of language we also refer to objects in $D^b(X)$ or $D^+(X)$ as sheaves. Given a sheaf $\mathcal{F}$ on X we denote by $\mathcal{F}\otimes\mathbb{F}$ its extension of scalars to a sheaf on $X\otimes\mathbb{F}$. Given a sheaf $\mathcal{F}$ on X we abuse notation and write

$$\mathbb{H}^*(\mathcal{F}):=\mathbb{H}^*(X\otimes\mathbb{F},\mathcal{F}\otimes\mathbb{F})=\mathbb{H}^*(\mathcal{F}\otimes\mathbb{F}).$$

We *never* consider hypercohomology before extending scalars.

On the category $D^b(X)$, we have the Verdier duality functor $\mathbb{D}\colon D^b(X)\to D^b(X)^{\mathrm{op}}$, and for each map $f\colon X\to Y$, we have Verdier dual pushforward functors

$$f_*,\, f_!\colon D^b(X)\to D^b(Y)$$

(often denoted Rf_* and $Rf_!$), and Verdier dual pullback functors

$$f^*,\, f^!\colon D^b(Y)\to D^b(X).$$

In $D^b(X)$ we have the full abelian subcategory $P(X)$ of *perverse sheaves*; see [5]. We will call a sheaf $\mathcal{F}$ *shifted perverse* if $\mathcal{F}[n]$ is perverse for some $n\in\mathbb{Z}$.

3.3 The Frobenius and its action on sheaves

Given any variety X defined over $\mathbb{F}_q$ we have the Frobenius morphism

$$\mathrm{Fr}_q\colon X \to X,$$

which for affine $X \subset \mathbb{A}^m$ is given by $(x_1,\dots,x_m) \mapsto (x_1^q,\dots,x_m^q)$. The fixed points of $\mathrm{Fr}_{q^n} := (\mathrm{Fr}_q)^n$ are precisely $X(\mathbb{F}_{q^n})$, the points of X defined over $\mathbb{F}_{q^n}$.

Given any $\mathcal{F} \in D^b(X)$ we have an isomorphism (see [5, Chapter 5])

$$F_q^*\colon \mathrm{Fr}_q^*\mathcal{F} \xrightarrow{\sim} \mathcal{F},$$

and obtain an induced action of $F_{q^n}^* := (F_q^*)^n$ on the stalk of $\mathcal{F}$ at any point $x \in X(\mathbb{F}_{q^n})$. By considering the eigenvalues of the action of $F_{q^n}^*$ on the stalks of $\mathcal{F}$ at all points $x \in X(\mathbb{F}_{q^n})$ for all $n \geq 1$, one defines the subcategory of *mixed sheaves* $D_m^b(X)$ as well as the full subcategories of sheaves of *weight* $\leq w$ and *weight* $\geq w$ for $w \in \mathbb{Z}$, which we denote $D^b_{\leq w}(X)$ and $D^b_{\geq w}(X)$ respectively; see [5, Chapter 5], [10] or the first chapter of [14]. An object is called *pure of weight* i if it lies in both $D^b_{\leq i}(X)$ and $D^b_{\geq i}(X)$.

Given any mixed sheaf $\mathcal{F}$ on X, all eigenvalues $\alpha \in \Bbbk$ of Fr_q^* on $\mathbb{H}^*(\mathcal{F})$ are algebraic integers such that all complex numbers with the same minimal polynomial have the same complex norm, which by abuse of notation we denote by $|\alpha|$. As $\mathcal{F}$ is assumed mixed, all such norms will be $q^{i/2}$ for some i. Let $\mathbb{H}^*_\alpha(\mathcal{F}) \subset \mathbb{H}^*(\mathcal{F})$ be the generalized eigenspace of α, and let

$$\mathbb{H}^{*,i}(\mathcal{F}) := \bigoplus_{|\alpha|=q^{i/2}} \mathbb{H}^*_\alpha(\mathcal{F}). \tag{1}$$

If X is proper and $\mathcal{F}$ pure, then the decomposition (1) will agree with the cohomological grading of $\mathbb{H}^*(\mathcal{F})$ by the Riemann hypothesis for X; that is, $\mathbb{H}^{*,i}(\mathcal{F}) = \mathbb{H}^i(\mathcal{F})$. Since we are not assuming that X is proper, this can fail even if $\mathcal{F}$ is pure. For example, if $X = \mathbb{A}^1 \setminus \{0\} \cong G_1$ and $\mathcal{F} = \Bbbk_X$, then as in Example 3.1, we have $\mathbb{H}^{*,2}(\mathcal{F}) = \mathbb{H}^1(\mathcal{F})$.

Definition 3.2 The grading on $\mathbb{H}^*(\mathcal{F})$ where the elements of $\mathbb{H}^{*,i}(\mathcal{F})$ are defined to have degree i is called the *weight grading*.

Remark 1 The constant sheaf on X has a unique mixed structure for which the Frobenius acts trivially on all stalks, and its hypercohomology is the étale cohomology of X. The i^{th} graded component of $H^*(X;\Bbbk)$ for the weight grading is $H^{*;i}(X;\Bbbk)$. So our previous discussion was a reflection of some of the properties of the Frobenius action on the cohomology of algebraic varieties.

If $X = \operatorname{Spec} \mathbb{F}_q$, then a perverse sheaf on X is the same as a finite-dimensional $\Bbbk$–vector space together with a continuous action of the absolute Galois group of $\mathbb{F}_q$. In particular we have the *Tate sheaf* $\underline{\Bbbk}(1)$, which under the above equivalence corresponds to $\Bbbk$, with action of F_q^* given by q^{-1}. Recall that we have fixed a square root $q^{1/2}$ of q in $\Bbbk$, allowing us to define the *half Tate sheaf* $\underline{\Bbbk}\left(\frac{1}{2}\right)$, with F_q^* acting by $q^{-1/2}$.

Given any X with structure morphism $X \xrightarrow{a} \operatorname{Spec} \mathbb{F}_q$ and any sheaf $\mathcal{F}$ on X, we define

$$\mathcal{F}\left(\tfrac{m}{2}\right) := \mathcal{F} \otimes a^* \underline{\Bbbk}\left(\tfrac{1}{2}\right)^{\otimes m}.$$

The following notation will prove useful:

$$\mathcal{F}\langle d\rangle = \mathcal{F}[d]\left(\frac{d}{2}\right).$$

If $\mathcal{F}$ is pure of weight w, then $\mathcal{F}[d]$ is pure of weight $w+d$, and $\mathcal{F}(d/2)$ pure of weight $w-d$, so $\mathcal{F}\langle d\rangle$ is again pure of weight 0. The natural isomorphism $\mathbb{H}^*(F) \cong \mathbb{H}^*(F)\langle d\rangle$ has degree d for both the weight and cohomological gradings.

The most important fact about mixed sheaves for our purposes is that every mixed perverse sheaf $\mathcal{F}$ on X admits a unique increasing filtration W, called the *weight filtration*, such that, for all i,

$$\operatorname{gr}_i^W \mathcal{F} := W_i\mathcal{F}/W_{i-1}\mathcal{F}$$

is pure of weight i. Any morphism of mixed sheaves preserves this filtration. The decomposition (1) comes from the weight filtration applied to the pushforward sheaf $p_*\mathcal{F}$ to a point.

In fact, after extension of scalars to the algebraic closure, the extensions in this filtration are the only way that mixed perverse sheaves can fail to be semisimple.

Theorem 3.3 (Gabber; [5, Théorème 5.3.8]) *If $\mathcal{F}$ is a pure perverse sheaf on X then $\mathcal{F} \otimes \mathbb{F}$ is semisimple.*

3.4 The function–sheaf dictionary

The eigenvalues of the Frobenius on stalks are also valuable for analyzing the structure of a given perverse sheaf. To any mixed perverse sheaf $\mathcal{F}$ (or more generally, any mixed sheaf) one may associate a function on $X(\mathbb{F}_{q^n})$ for each n given by the supertrace of the Frobenius on the stalks of the cohomology sheaves at those points:

$$[\mathcal{F}]_n\colon X(\mathbb{F}_{q^n}) \to \Bbbk, \quad x \mapsto \operatorname{Tr}(F_{q^n}^*, \mathcal{F}_x) := \sum (-1)^j \operatorname{Tr}(F_{q^n}^*, \mathcal{H}^j(\mathcal{F}_x)).$$

Proposition 3.4 *These functions give a map from the Grothendieck group of the category of mixed perverse sheaves to the abelian group of functions on $X(\mathbb{F}_{q^n})$ for each n, and these maps are jointly injective. That is, if $\mathcal{F}$ and $\mathcal{G}$ are semisimple and $[\mathcal{F}]_n = [\mathcal{G}]_n$ for all n, then $\mathcal{F}$ and $\mathcal{G}$ are isomorphic.*

Proof The fact that these functions give a map of Grothendieck groups is just that all maps in the long exact sequence must respect the action of the Frobenius, so the supertrace is additive under extensions. The proof that this map is injective may be found in [15, Théorème 1.1.2]; see also [14, Theorem 12.1]. □

This reduces the calculation of the constituents of a weight filtration to a problem of computing $[\mathcal{F}]_n$ for simple perverse sheaves, followed by linear algebra. Indeed, suppose that $\mathcal{F}, \mathcal{G} \in D^b_m(X)$ are such that $[\mathcal{F}]_n$ and $[\mathcal{G}]_n$ agree for all n, with $\mathcal{G}$ semisimple. As $[\mathcal{F}]_n = \sum [\mathrm{gr}_i^W \mathcal{F}]_n$ for all n, we conclude that $\mathrm{gr}_i^W \mathcal{F}$ is isomorphic to the largest direct summand of $\mathcal{G}$ of weight i.

3.5 The chromatographic complex

We want to explain how to move between the weight filtration and a complex, which we term *the chromatographic complex*, composed of its pure constituents. For background, the reader is referred to [8, Section 1.4] and [5, Section 3.1].

Let $\mathcal{A}$ be an abelian category with enough injectives and let $D^+(\mathcal{A})$ denote its bounded below derived category. We may also consider the filtered derived category $DF^+(\mathcal{A})$ whose objects consist of $K \in D^+(\mathcal{A})$ together with a finite increasing filtration

$$\cdots \subset W_{i-1}K \subset W_i K \subset W_{i+1}K \subset \cdots,$$

where finite means that $W_i K = 0$ for $i \ll 0$ and $W_i K = W_{i+1} K$ for $i \gg 0$.

For all p we define

$$\mathrm{gr}_p^W K := W^p K / W^{p-1} K.$$

More generally, for $q \le p$, let

$$(W^p / W^q)(K) := W^p K / W^q K.$$

For all p we have a distinguished triangle

$$\mathrm{gr}_p^W K \to (W^{p+1}/W^{p-1})(K) \to \mathrm{gr}_{p+1}^W K \xrightarrow{[1]}$$

and in particular a "boundary" morphism $\mathrm{gr}_W^{p+1} \to \mathrm{gr}_W^p K[1]$. Shifting, we obtain a sequence

$$\cdots \to \mathrm{gr}_{p+1}^W K[-(p+1)] \to \mathrm{gr}_p^W K[-p] \to \mathrm{gr}_{p-1}^W K[-(p-1)] \to \cdots . \tag{2}$$

Lemma 3.5 *The morphisms in (2) define a complex.*

Proof After completing the (commuting) triangle

$$\begin{tikzcd} & (W^{p+1}/W^{p-1})(K) \arrow[dr] & \\ \mathrm{gr}_p^W K \arrow[ur] \arrow[rr] & & (W^{p+2}/W^{p-1})(K) \end{tikzcd}$$

to an octahedron, one sees that the morphism

$$\mathrm{gr}_{p+2}^W K \to \mathrm{gr}_{p+1}^W K[1] \to \mathrm{gr}_p^W K[2]$$

may be factored as

$$\mathrm{gr}_{p+2}^W \to W^{p+1}/W^{p-1}(K) \to \mathrm{gr}_{p+1}^W K[1] \to \mathrm{gr}_p^W K[2].$$

However, the second two morphisms form part of a distinguished triangle, and so their composition is zero. □

Given any left exact functor $T\colon \mathcal{A} \to \mathcal{B}$ between abelian categories we can consider the hypercohomology objects $R^i T(K) \in \mathcal{B}$, obtained by applying T to an injective resolution of K. One has a spectral sequence

$$E_1^{p,q} = R^{p+q} T(\mathrm{gr}_{-p}^W K) \Rightarrow R^{p+q} T(K) \tag{3}$$

(see [18, Theorem 2.6] or [8, Section 1.4.5]), and a diagram chase shows that the first differential of this spectral sequence (ie the differential on the E_1 page) is the same as the differential obtained by applying $R^q T(-)$ to the complex (2).

We now apply these ideas to $D_m^b(X)$, where X and $D_m^b(X)$ are as in Section 3.3.

Lemma 3.6 *Any $\mathcal{G} \in D_m^b(X)$ admits a "filtration" $\cdots \to \mathcal{G}_{\le i} \to \mathcal{G}_{\le i+1} \to \cdots$ such that:*

(1) *If we define $\mathrm{gr}_i(X)$ via the distinguished triangle*

$$\mathcal{G}_{\le i-1} \to \mathcal{G}_{\le i} \to \mathrm{gr}_i(\mathcal{G}) \xrightarrow{[1]},$$

then $\mathrm{gr}_i(\mathcal{G})$ is pure of weight i.

(2) $\mathrm{gr}_i(\mathcal{G}) = 0$ *for* $|i| \gg 0$.

We will refer to any sequence of maps satisfying the conditions of the lemma as a *weight filtration* on $\mathcal{G}$. As the choice of article emphasizes, this is *not* unique. (For example, the reader may convince themselves easily that the zero object admits many nonequivalent weight filtrations.)

Proof It is enough to show that for any $\mathcal{G} \in D^b_m(X)$ there exists a distinguished triangle

$$\mathcal{G}_{\leq 0} \to \mathcal{G} \to \mathcal{G}_{>0} \xrightarrow{[1]}$$

with $\mathcal{G}_{\leq 0}$ (resp. $G_{>0}$) of weight ≤ 0 (resp. > 0). If $\mathcal{G}$ is perverse then the statement is an immediate consequence of the existence of weight filtrations on perverse sheaves [5, Théorèm 5.3.5].

By induction on the perverse filtration it is enough to show the following claim: if

$$\mathcal{F} \to \mathcal{G} \to \mathcal{K} \xrightarrow{[1]}$$

is a distinguished triangle of sheaves, and there exist distinguished triangles

$$\mathcal{F}_{\leq 0} \to \mathcal{F} \to F_{>0} \xrightarrow{[1]} \quad \text{and} \quad \mathcal{K}_{\leq 0} \to \mathcal{K} \to \mathcal{K}_{>0} \xrightarrow{[1]}$$

with $\mathcal{F}_{\leq 0}$ and $\mathcal{K}_{\leq 0}$ (resp. $\mathcal{F}_{>0}$ and $\mathcal{K}_{>0}$) of weights ≤ 0 (resp. > 0), then there exists a filtration

$$\mathcal{G}_{\leq 0} \to \mathcal{G} \to G_{>0} \xrightarrow{[1]}$$

satisfying the same conditions. For the rest of the proof the following notation will be useful. Given a commutative triangle

$$\begin{array}{ccc} & \mathcal{B} & \\ \nearrow & & \searrow \\ \mathcal{A} & \longrightarrow & \mathcal{C} \end{array}$$

we denote by $O(\mathcal{A}, \mathcal{B}, \mathcal{C})$ the corresponding octahedron (the maps will be clear from the context).

By considering $O(\mathcal{G}, \mathcal{K}, \mathcal{K}_{>0})$ we deduce the existence of distinguished triangles

$$\mathcal{A} \to \mathcal{G} \to \mathcal{K}_{>0} \xrightarrow{[1]} \quad \text{and} \quad \mathcal{F} \to \mathcal{A} \to \mathcal{K}_{\leq 0} \xrightarrow{[1]} .$$

By considering $O(\mathcal{F}_{\leq 0}, \mathcal{F}, \mathcal{A})$ we deduce the existence of distinguished triangles

$$\mathcal{F}_{\leq 0} \to \mathcal{A} \to \mathcal{B} \xrightarrow{[1]} \quad \text{and} \quad \mathcal{F}_{>0} \to \mathcal{B} \to \mathcal{K}_{\leq 0} \xrightarrow{[1]} .$$

Because $\operatorname{Hom}(\mathcal{K}_{\leq 0}, \mathcal{F}_{>0}[1]) = 0$ by [5, Proposition 5.1.15(ii)], we deduce that the "connecting" map $\mathcal{K}_{\leq 0} \to F_{>0}[1]$ in the second triangle is zero, and hence that $\mathcal{B} \cong \mathcal{F}_{>0} \oplus \mathcal{K}_{\leq 0}$. (This decomposition is not canonical; we fix one.) By considering $O(\mathcal{A}, \mathcal{B}, \mathcal{F}_{>0})$ (the map $\mathcal{B} = \mathcal{F}_{>0} \oplus Z_{\leq 0} \to \mathcal{F}_{>0}$ is the projection) we deduce the existence of distinguished triangles

$$\mathcal{C} \to \mathcal{A} \to \mathcal{F}_{>0} \xrightarrow{[1]} \quad \text{and} \quad \mathcal{F}_{\leq 0} \to \mathcal{C} \to \mathcal{K}_{\leq 0} \xrightarrow{[1]} .$$

In particular, $\mathcal{C}$ has weight ≤ 0. Finally, by considering $O(\mathcal{C}, \mathcal{A}, \mathcal{G})$ we deduce the existence of distinguished triangles

$$\mathcal{C} \to \mathcal{G} \to \mathcal{D} \xrightarrow{[1]} \quad \text{and} \quad \mathcal{F}_{>0} \to \mathcal{D} \to \mathcal{K}_{>0} \xrightarrow{[1]} .$$

In particular, $\mathcal{D}$ has weight > 0. It follows that we can take $\mathcal{G}_{\leq 0} := \mathcal{C}$ and $\mathcal{G}_{>0} := \mathcal{D}$. □

Applying the above considerations to $\mathcal{F}$ together with its weight filtration we obtain:

Definition 3.7 The *local chromatographic complex* of a mixed sheaf $\mathcal{F} \in D^b_m(X)$ is the complex of objects in $D^b_m(X)$ given by

$$\cdots \to \mathrm{gr}^W_{p+1} \mathcal{F}[-(p+1)] \to \mathrm{gr}^W_p \mathcal{F}[-p] \to \mathrm{gr}^W_{p-1} \mathcal{F}[-(p-1)] \to \cdots .$$

Applying $T = \mathbb{H}^*(-)$ we obtain the *global chromatographic complex*,

$$\cdots \to \mathbb{H}^*(\mathrm{gr}^W_{i+1} \mathcal{F}[-(i+1)]) \to \mathbb{H}^*(\mathrm{gr}^W_i \mathcal{F}[-i]) \to \mathbb{H}^*(\mathrm{gr}^W_{i-1} \mathcal{F}[-(i-1)]) \to \cdots .$$

The spectral sequence (3) with $T = \mathbb{H}^*(-)$ is the *chromatographic spectral sequence*.

Note that the global chromatographic complex sends $\langle d \rangle$ to simultaneous grading shift on terms of the complex, and the Tate twist $(d/2)$ to homological shift of the complex. Unfortunately, this definition is not entirely an invariant of the object $\mathcal{G}$, but the dependence on choice of filtration is not very strong.

Proposition 3.8 *The chromatographic complexes associated to two different weight filtrations on a single object $\mathcal{G} \in D^b(X)$ are homotopy equivalent, after extending scalars to $\mathbb{F}$.*

In particular, this shows that all pages of the chromatographic spectral sequence after the first are independent of the choice of filtration.

Proof We note that if $\mathcal{G}$ is quasi-isomorphic to a complex $\cdots \to \mathcal{F}_i \to \cdots$, then we obtain a natural bicomplex by writing the chromatographic complexes of $\mathcal{F}_i$ vertically, and then the maps induced by the original differentials horizontally. By Gabber's theorem, we note that after passing to $\mathbb{F}$ every term in this bicomplex is semisimple, and the horizontal maps go between objects pure of the same degree, and thus split.

Now assume perverse sheaves $\mathcal{F}'_i$ form another complex isomorphic in the derived category to $\mathcal{G}$. For simplicity, we may assume there is a quasi-isomorphism $\phi_i\colon \mathcal{F}_i \to \mathcal{F}'_i$ between these complexes. This induces a map $\phi^\#$ between our bicomplexes, which is an isomorphism after taking horizontal cohomology, since this will give us the chromatographic complexes of the perverse cohomology of $\mathcal{G}$.

Consider the kernel of $\phi^\#$. This is itself a bicomplex, and each of its rows has trivial cohomology, and is split. Thus, each row is homotopic to 0. Furthermore, we can choose these homotopies so that they commute with the vertical differentials, and thus when applied to the total complex of the kernel, they show that this total complex is nullhomotopic.

We now use the result that any surjective chain map whose kernel is homotopic to the zero complex and is a summand of the chain complex *with the differentials forgotten* is a homotopy equivalence (this is a consequence of Gaussian elimination). Thus, the chromatographic complex from the $\mathcal{F}_i$ is homotopy equivalent to the total complex of the image of $\phi^\#$, and the dual result applied to the inclusion of the image shows that the chromatographic complex for $\mathcal{F}_i'$ is also equivalent to this image. □

Proposition 3.9 *The global chromatographic complex is preserved (up to homotopy) by proper pushforward.*

Proof Proper pushforward preserves purity, and thus sends weight filtrations to weight filtrations. Furthermore, pushforward always preserves hypercohomology. □

Corollary 3.10 *If we let $E_*^{*,*}$ be the chromatographic spectral sequence, then all differentials preserve the weight grading on hypercohomology. Furthermore, we have:*

- $E_1^{i,j} = \mathbb{H}^{i+j}(\mathrm{gr}_{-j}^W \mathcal{F})$ *is the global chromatographic complex.*
- E_2 *is the cohomology of the global chromatographic complex.*
- *The chromatographic spectral sequence converges to the hypercohomology* $\mathbb{H}^{i+j}(\mathcal{F})$.

Remark 2 It seems likely that it is possible to interpret the results of this section in terms of "weight structures", introduced by Bondarko [6] and Paukzsello [20]. In particular, Bondarko shows the existence of a functor from a derived category equipped with a suitable weight structure, to the homotopy category of pure complexes in a very general framework.

3.6 The equivariant derived category

We have thus far discussed the theory of perverse sheaves on schemes, but we will require a generalization of schemes that includes the quotient of a scheme X by the action of an algebraic group G, which can be understood as G–equivariant geometry on X.

This quotient can be understood as a stack, but the theory of perverse sheaves on stacks is not straightforward, and it proved more suitable to give a treatment of the equivariant derived category similar to that of Bernstein and Lunts [4], but with an eye to working over characteristic p with the action of the Frobenius (that is, "in the mixed setting"). We have done this in a separate note [25].

The result is the *bounded below equivariant derived category* $D^+_G(X)$ and its subcategory $D^b_G(X)$ of bounded complexes for a variety X acted on by an affine algebraic group G. The resulting formalism is essentially identical to that of Bernstein and Lunts. We now summarize the essential points.

We have a forgetful functor

$$\mathrm{For}\colon D^+_G(X) \to D^+(X)$$

which preserves the subcategories of bounded complexes and, given any $\mathcal{F} \in D^+_G(X)$, the cohomology sheaves of $\mathrm{For}(\mathcal{F})$ are locally constant along the G–orbits on X.

Given an equivariant map $f\colon X \to Y$ of G–varieties we have functors

$$f_*, f_!\colon D^+_G(X) \to D^+_G(Y) \quad \text{and} \quad f^*, f^!\colon D^+_G(Y) \to D^+_G(X)$$

for equivariant maps $f\colon X \to Y$ of G–varieties. These functors commute with the forgetful functor.

If $H \subset G$ is a closed subgroup and X is a G–space, we have an adjoint pair $(\mathrm{res}^G_H, \mathrm{ind}^G_H)$ of restriction and induction functors

$$\mathrm{res}^G_H\colon D^+_G(X) \to D^+_H(X) \quad \text{and} \quad \mathrm{ind}^G_H\colon D^+_H(X) \to D^+_G(X).$$

These functors preserve the subcategories of bounded complexes, and one has an isomorphism $\mathrm{res}^G_{\{1\}} \cong \mathrm{For}$.

More generally, given a map $\phi\colon H \to G$, a G–variety X, an H–variety Y and a ϕ–equivariant map $m\colon X \to Y$, we have an adjoint pair $({}^G_H m^*, {}^G_H m_*)$ of functors

$${}^G_H m^*\colon D^+_H(Y) \to D^+_G(X) \quad \text{and} \quad {}^G_H m_*\colon D^+_G(X) \to D^+_H(Y).$$

As a special case, we have ${}^G_H \mathrm{id}^* \cong \mathrm{res}^G_H$ and ${}^G_H \mathrm{id}_* \cong \mathrm{ind}^G_H$. The functor ${}^G_H m^*$ preserves the subcategory of bounded complexes, but this is not true in general for ${}^G_H m_*$. In fact, this is the reason that we are forced to consider complexes of sheaves which are not bounded above.

If $G = G_1 \times G_2$ and G_1 acts freely on X with quotient X/G_1, one has the *quotient equivalence*

$$D^+_G(X) \cong D^+_{G_2}(X/G_1), \tag{4}$$

which restricts to an equivalence between the subcategories of bounded complexes. If we let $\phi\colon G_1 \times G_2 \to G_2$ denote the projection, then the quotient map $X \to X/G_1$ is ϕ–equivariant and the above equivalence is realized by ${}^{G_2}_{G_1\times G_2}m^*$ and ${}^{G_2}_{G_1\times G_2}m_*$.

Using the forgetful functor For: $D^+_G(X) \to D^+(X)$ many notions carry over immediately. For example, we call an object $\mathcal{F}$ in $D^+_G(X)$ *perverse* if and only if For $\mathcal{F}$ is perverse.

Moreover, if X is defined over $\mathbb{F}_q$, then we can also incorporate the action of the Frobenius. In particular, perverse objects in $D^+_G(X)$ still have weight filtrations, which are preserved by the restriction functor and we can extend Proposition 3.4 to the equivariant setting.

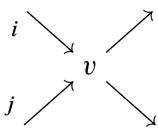
4 Description of the invariant

Equipped with these geometric tools, we continue the construction of our invariant.

4.1 The sheaf associated to a diagram

In this subsection we describe the sheaf $\mathcal{F}_D$ on X_D.

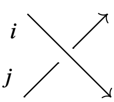
We first discuss the case of a single (i,j)–crossing:

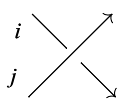


As we have seen, $X_D = G_{i+j}$. Consider the big Bruhat cell

$$U := \{g \in G_{i+j} \mid V_i \cap gV_j = 0\}, \tag{5}$$

and let $k\colon U \hookrightarrow G_{i+j}$ denote its inclusion. As U is an orbit under $P_{i,j} \times P_{j,i}$ it is certainly G_D–invariant. We now define $\mathcal{F}_v = \mathcal{F}_D \in D_{G_D}(X_D)$ as follows:



$\longmapsto k_*\underline{\Bbbk}_U\langle ij\rangle,$


$\longmapsto k_!\underline{\Bbbk}_U\langle ij\rangle.$

As U is the complement of a divisor in G_{i+j}, both these sheaves are shifted perverse.

We now consider the case of a general diagram D of an oriented colored tangle. After forgetting equivariance, $\mathcal{F}_D$ is simply the exterior product of the above sheaves associated to each crossing. To take care of the equivariant structure we need to proceed a little more carefully.

Let D be the diagram of an oriented colored tangle and Γ its underlying graph. Let D' be the diagram obtained from D by cutting each strand connecting two vertices in Γ (so that D' is a disjoint union of (i,j)–crossings). Let Γ' be the graph corresponding to D'. Obviously we have $X_D = X'_D$. Note also that for every e with two vertices in Γ, we have two edges, which we denote by e_1 and e_2, in Γ'. We have a natural map $G_D \to G'_D$, which is the identity on factors corresponding to external edges, and is the diagonal $G_e \to G_{e_1} \times G_{e_2}$ on the remaining factors.

We define

$$\mathcal{F}_D := \mathrm{res}^{G}_{G'}\Bigg(\mathop{\boxtimes}_{v \in \mathcal{V}(D')} \mathcal{F}_v\Bigg) \in D^b_{G_D}(X_D).$$

Of course, this sheaf depends on the link diagram used; different diagrams correspond to sheaves on different spaces. Instead, we will studying the hypercohomology of these sheaves, and the corresponding chromatographic spectral sequence.

Definition 4.1 We let $\mathcal{A}_i(L)$ denote the i^{th} page of the chromatographic spectral sequence (as given by Definition 3.7) for $\mathcal{F}_D$. This is triply graded, where by convention subquotients of $\mathbb{H}^{j-\ell;j-k}(\mathrm{gr}^W_\ell \mathcal{F}_D)$ lie in $\mathcal{A}_i^{j;k;\ell}(L)$.

Remark 3 These grading conventions may seem strange, but they are an attempt to match those already in use in the field. These conventions are almost those of [17], though we will not match perfectly since we have different grading shifts in our definition of the complex for a single crossing. We hope the reader finds these choices defensible on grounds of geometric naturality. This simply changes the shift we must apply to our invariant to assure it is a true knot invariant.

It is these spaces for $i > 1$ which we intend to show are knot invariants (up to shift).

4.2 Braids and sheaves on groups

As we mentioned in Section 1, in the special case of a braid β, there is a different perspective on this construction.

Let β be the diagram of a colored braid on n strands with labels $\boldsymbol{n} = (i_1, i_2, \ldots, i_n)$ and underlying labeled graph Γ. Let $N = \sum_{j=1}^n i_j$ denote the colored braid index. We assume our braid is in generic position, so reading from start to finish, we fix an order on the vertices $v_1, v_2, \ldots, v_p$ of Γ. This corresponds to an expression for β in the standard generators of the braid group.

In the previous section we described how to associate to β a group G_β and a G_β–variety X_β. We can decompose G_β as

$$G_\beta = G_\beta^+ \times G_\beta^\iota \times G_\beta^-,$$

where G_β^+, G_β^ι and G_β^- denote the factors of G_β corresponding to incoming, interior and outgoing edges of Γ respectively.

In what follows we will describe an action of $G_\beta^+ \times G_\beta^-$ on G_N and a map

$$m\colon X_\beta \to G_N$$

equivariant with respect to the natural projection $\phi\colon G_\beta \to G_\beta^+ \times G_\beta^-$. We will study our sheaf $\mathcal{F}_\beta$ by considering its equivariant pushforward under this map.

First we describe an embedding $\alpha_v\colon G_v \to G_N$ corresponding to each vertex $v \in \Gamma$. Let us fix a basis $e_1, \dots, e_N$ of V_N and let $W_1, W_2, \dots, W_n$ be vector spaces (again with fixed bases) of dimensions $i_1, i_2, \dots, i_n$ respectively.

Definition 4.2 Given any permutation $w \in S_n$, we let

$$h_w\colon W = \bigoplus_{j=1}^{n} W_j \xrightarrow{\sim} V$$

be the isomorphism defined by mapping the basis vectors of $W_{w^{-1}(1)}$ to the first $w^{-1}(1)$ basis vectors of V in their natural order, the basis vectors of $W_{w^{-1}(2)}$ to the next $w^{-1}(2)$ basis vectors, etc.

For any braid β, we have an induced permutation, and by abuse of notation, we let h_β be the map corresponding to this permutation.

In the obvious basis, this map is a permutation matrix. The corresponding permutation is a shortest coset representative for the Young subgroup preserving the partition of $[1, N]$ of sizes $i_1, \dots, i_n$, corresponding to the "cabling" of the permutation w.

Now choose a vertex v in Γ and let e' and e'' denote the two incoming edges which are in the strands connected to the incoming vertices labeled j' and j'' respectively, so $i_{j'}$ and $i_{j''}$ are the labels on e' and e''. Because we have ordered the vertices of Γ, we may factor β into braids $\alpha_v \cdot \beta_v \cdot \omega_v$, with β_v consisting of a simple crossing corresponding to v. The procedure described in the previous paragraph yields an embedding

$$W_{j'} \oplus W_{j''} \hookrightarrow W \xrightarrow{h_{\alpha_v}} V_N.$$

This induces an embedding

$$\iota_v\colon G_v \hookrightarrow G_N.$$

We let braids on n strands act on sequences of n elements on the right by the usual association of a permutation to each braid. We may then identify

$$G_\beta^+ \cong G_{\boldsymbol{n}}, \quad G_\beta^- \cong G_{\boldsymbol{n}\beta},$$

and therefore obtain an action of $G_\beta^+ \times G_\beta^-$ on G_N by left and right multiplication. We let $P_\beta^+ = P_{\boldsymbol{n}}$ and $P_\beta^- = P_{\boldsymbol{n}\beta}$. We denote by $\phi\colon G_\beta \to P_\beta^+ \times P_\beta^-$ the composition of the natural projection with the inclusion $G_\beta^\pm \hookrightarrow P_\beta^\pm$.

Consider the map

$$m\colon X_\beta \to G_N, \quad (g_{v_1}, \dots, g_{v_p}) \mapsto \iota_{v_1}(g_{v_1})\iota_{v_2}(g_{v_2}) \cdots \iota_{v_p}(g_{v_p}).$$

It is easy to see that this map is equivariant with respect to ϕ.

Definition 4.3 Let $\Phi_\beta = {}^{P_\beta^+ \times P_\beta^-}_{G_\beta} m_* \mathcal{F}_\beta$.

This definition is useful, since it is compatible with braid multiplication. We have a diagram of equivariant maps of spaces:

$$\begin{tikzcd} G_N & & \\ & G_N \times G_N \arrow[ul, "\pi_1"'] \arrow[dl, "\pi_2"] \arrow[r, "\mu"] & G_N \\ G_N & & \end{tikzcd}$$

As usual, this diagram can be used to construct the functor of sheaf convolution:

$$- \star - \colon D^b_{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta}}(G_N) \times D^b_{P_{\boldsymbol{n}\beta} \times P_{\boldsymbol{n}\beta\beta'}}(G_N) \to D^b_{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta\beta'}}(G_N),$$

$$\mathcal{F}_1 \star \mathcal{F}_2 \cong {}^{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta\beta'}}_{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta} \times P_{\boldsymbol{n}\beta\beta'}} \mu_* \Big(\mathrm{res}^{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta}^2 \times P_{\boldsymbol{n}\beta\beta'}}_{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta} \times P_{\boldsymbol{n}\beta\beta'}} (\mathcal{F}_1 \boxtimes \mathcal{F}_2) \Big).$$

Theorem 4.4 *We have a canonical isomorphism* $\Phi_\beta \star \Phi_{\beta'} \cong \Phi_{\beta\beta'}$.

We should note that here we are simply claiming that this holds for the composition of diagrams. We will prove in Sections 8 and 9 that the sheaf we associate to a braid doesn't depend on the choice of presentation.

Proof Immediate from the definition of Φ. □

As G_β^ι acts freely on X_β, we may factor m as

$$X_\beta \to X_\beta / G_\beta^\iota \to G_N.$$

One may verify that the second map is the composition of an affine bundle along which $\mathcal{F}_\beta$ is smooth, and a proper map. It follows that

$${}^{P_\beta^+ \times P_\beta^-}_{G_\beta} m_*$$

preserves the weight filtration on $\mathcal{F}_\beta$. Hence the chromatographic spectral sequences for $\mathcal{F}_\beta$ and Φ_β are isomorphic.

Note that if β is closable, then $\boldsymbol{n}\beta = \boldsymbol{n}$, and $P_\beta^\pm$ have the same image in the group. Thus these subgroups are canonically isomorphic. Let $(P_\beta)_\Delta \subset P_\beta^+ \times P_\beta^-$ be the diagonal and let $\widehat{\beta}$ be the colored link diagram given by the closure of β.

Theorem 4.5 *We have a canonical isomorphism between*

- *the chromatographic spectral sequence of $\mathcal{F}_{\widehat{\beta}}$ as a $G_{\widehat{\beta}}$–sheaf, and*
- *the chromatographic spectral sequence of Φ_β as a $(P_\beta)_\Delta$–sheaf.*

Proof Since P_* and G_* are homotopy equivalent, the functor $\mathrm{res}_{G_*}^{P_*}$ is fully faithful, so we may work with their restrictions instead. We have already observed that the weight filtration on Φ_β and the pushforward of the weight filtration on $\mathcal{F}_\beta$ agree. Thus the equivariant chromatographic spectral sequences of

$$\mathrm{res}_{\phi^{-1}(H)}^{G_\beta} \mathcal{F}_\beta \quad \text{and} \quad \mathrm{res}_H^{G_\beta^+ \times G_\beta^-} \Phi_\beta$$

are canonically isomorphic for any subgroup $H \subset G_\beta^+ \times G_\beta^-$.

On the other hand, we have a canonical identification $G_{\widehat{\beta}} \cong \phi^{-1}((G_\beta)_\Delta)$, and $X_\beta = X_{\widehat{\beta}}$, with

$$\mathcal{F}_{\widehat{\beta}} = \mathrm{res}_{G_{\widehat{\beta}}}^{G_\beta} \mathcal{F}_\beta.$$

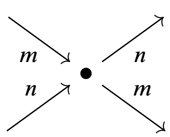
The result follows. □

5 Analyzing an (m,n)–crossing

5.1 Preliminary details

In this section we work out all the details for an (m,n)–crossing. This will be of use in expressing the invariant in terms of bimodules.

We consider an (m,n)–crossing. Its underlying graph is



and the variety in question is G_{m+n} acted on by $P_{m,n} \times P_{n,m}$ by left and right multiplication: $(p,q)\cdot g = pgq^{-1}$ for $g \in G_{n+m}$ and $(p,q) \in P_{m,n} \times P_{n,m}$. The orbits under this action are

$$\mathcal{O}_i = \{g \in G_{m+n} \mid \dim V_m \cap g V_n = i\},$$

for $0 \le i \le \min(n,m)$. Clearly $\mathcal{O}_j \subset \overline{\mathcal{O}}_i$ if and only if $j > i$. For all $0 \le i \le \min(n,m)$ we denote the inclusion of the orbit $\mathcal{O}_i$ by $f_i \colon \mathcal{O}_i \hookrightarrow G_{n+m}$.

For each orbit $\mathcal{O}_i$ we have the corresponding intersection cohomology complex. It will prove natural to normalize them by requiring

$$\mathrm{IC}(\overline{\mathcal{O}}_i)|_{\mathcal{O}_i} \cong \underline{\Bbbk}_{\mathcal{O}_i}\langle nm - i^2\rangle.$$

Under this normalization each $\mathrm{IC}(\overline{\mathcal{O}}_i)$ is pure of weight 0.

We first describe resolutions for the closures $\overline{\mathcal{O}}_i \subset G_{m+n}$. Consider the variety

$$\widetilde{\mathcal{O}}_i = \{(W,g) \in \mathrm{Gr}_i^m \times G_{m+n} \mid W \subset V_m \cap g V_n\}.$$

We have an action of $P_{m,n} \times P_{n,m}$ on $\widetilde{\mathcal{O}}_i$ given by $(p,q)\cdot(W,g) = (pW, pgq^{-1})$. The second projection induces an equivariant map:

$$\pi_i \colon \widetilde{\mathcal{O}}_i \to \overline{\mathcal{O}}_i.$$

Proposition 5.1 *This is a small resolution of singularities.*

Proof The morphism π_i is patently an isomorphism over $\mathcal{O}_i$. Since $\mathcal{O}_i$ is exactly the subset of G_{n+m} where the induced map $V_n \to V/V_m$ has rank $n-i$, we have that $\mathcal{O}_i$ has the same codimension in G_{m+n} as the space of rank $n-i$ matrices in G_n, which is i^2. Hence, for $j < i$, $\mathcal{O}_i$ is of codimension $i^2 - j^2$ in $\overline{\mathcal{O}}_j$. Over any $x \in \mathcal{O}_j$ the fiber is the Grassmannian Gr_i^j. Thus

$$2\dim \pi_i^{-1}(x) = 2i(j-i) < (j+i)(j-i) = \mathrm{codim}_{\overline{\mathcal{O}}_i} \mathcal{O}_j. \qquad \square$$

Corollary 5.2
$$\mathrm{IC}(\overline{\mathcal{O}}_i) \cong \pi_{i*}\underline{\Bbbk}_{\widetilde{\mathcal{O}}_i}\langle nm - i^2\rangle.$$

Proof Proposition 5.1 implies that $\pi_{i*}\underline{\Bbbk}_{\widetilde{\mathcal{O}}_i}$ is a shift and twist of $\mathrm{IC}(\overline{\mathcal{O}}_i)$, since pushforward by a small resolution sends the constant sheaf to a shift of the intersection cohomology sheaf on the target. The restriction of $\pi_{i*}\underline{\Bbbk}_{\widetilde{\mathcal{O}}_i}\langle nm-i^2\rangle$ to $\mathcal{O}_i$ is isomorphic to $\underline{\Bbbk}_{\mathcal{O}_i}\langle nm - i^2\rangle$, which is our choice of normalization. $\square$

Given sheaves $\mathcal{F}, \mathcal{G} \in D^b_G(X)$ let us write

$$\mathrm{Hom}^\bullet(\mathcal{F},\mathcal{G}) := \bigoplus_m \mathrm{Hom}(\mathcal{F}, \mathcal{G}[m]).$$

This is a graded vector space.

Proposition 5.3 *In $D^b_{P_{m,n}\times P_{n,m}}(G)$ we have an isomorphism*

$$\mathrm{Hom}^\bullet(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i'})) \cong \bigoplus_j \mathrm{Hom}^\bullet(f_j^*\,\mathrm{IC}(\mathcal{O}_i), f_j^!\,\mathrm{IC}(\mathcal{O}_{i'})).$$

Proof For flag varieties this is [3, Theorem 3.4.1]. One may reduce to this situation using the quotient equivalence. □

The space $\mathrm{Hom}^\bullet(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i'}))$ itself has a weight grading, when thought of as sections of the sheaf-Hom from $\mathcal{H}om^\bullet(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i'}))$, which has a natural mixed structure. The decomposition of Proposition 5.3 is compatible with the Frobenius structure, and so the purity of the cohomology of $\mathcal{O}_i$ (which is an affine bundle over a partial flag variety) and the pointwise purity of $\mathrm{IC}(\mathcal{O}_i)$ shows that the weight grading of $\mathrm{Hom}^\bullet(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i'}))$ agrees with the cohomological grading.

This shows that:

Proposition 5.4 *In the mixed equivariant derived category $D^b_{m,P_{m,n}\times P_{n,m}}(G)$, there are no higher Exts between $\mathrm{IC}(\mathcal{O}_i)$ and $\mathrm{IC}(\mathcal{O}_{i'})\langle d\rangle$.*

Proof By the purity discussed above, all of the eigenvalues of Frobenius on the space $\mathrm{Ext}^i(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i'})\langle d\rangle)$ have complex norm $p^{i/2}$, so they are not 1. Thus, there are no invariants of Frobenius in this space. □

This shows immediately that:

Corollary 5.5 *Any mixed $(P_{m,n}\times P_{n,m})$–equivariant sheaf $\mathcal{F}$ on G is the iterated cone of its local chromatographic complex (in any dg-refinement). In particular, $\mathcal{F}$ is indecomposable if and only if the same is true of its local chromatographic complex.*

5.2 Calculating the weight filtration

Our aim in this section is to calculate the weight filtration on the sheaves associated to positive and negative crossings. We set

$$[n]_q = 1 + q + \cdots + q^{n-1},$$
$$[n]_q! = [n]_q[n-1]_q \cdots [1]_q,$$
$$\begin{bmatrix} j \\ i \end{bmatrix}_q = \frac{[j]_q}{[j-i]_q![i]_q!}.$$

In order to understand the constituents via the function–sheaf correspondence discussed in Section 3.4, we must calculate the trace of the Frobenius on the stalks of $\mathrm{IC}(\overline{\mathcal{O}}_i)$. Base change combined with the Grothendieck–Lefschetz fixed point formula yields:

Corollary 5.6 *If $j > i$ and $x \in \mathcal{O}_j(\mathbb{F}_{q^a})$, we have*

$$\mathrm{Tr}(F^*_{q^a}, (\pi_{i*}\underline{\Bbbk}_{\widetilde{\mathcal{O}}_i})_x) = \#\,\mathrm{Gr}^j_i(\mathbb{F}_{q^a}) = \begin{bmatrix} j \\ i \end{bmatrix}_{q^a}.$$

In the following proposition W denotes the weight filtration.

Proposition 5.7 *One has isomorphisms*

$$\mathrm{gr}^W_{-i}\, j_!\underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle \cong \mathrm{IC}(\overline{\mathcal{O}}_i)\big(\tfrac{i}{2}\big), \quad \mathrm{gr}^W_i\, j_*\underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle \cong \mathrm{IC}(\overline{\mathcal{O}}_i)\big(-\tfrac{i}{2}\big).$$

Proof Because taking weight filtrations commutes with forgetting equivariance, it is enough to handle the nonequivariant case. Note also that $\mathrm{IC}(\mathcal{O}_i)(i/2)$ is pure of weight $-i$. Thus, by the remarks in Section 3.4, the first statement of the proposition follows from the equality of the functions

$$[j_!\underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle]_{q^a} = \sum_i \big[\mathrm{IC}(\mathcal{O}_i)\big(\tfrac{i}{2}\big)\big]_{q^a}$$

for all $a \geq 1$. Evaluating at a point $x \in \mathcal{O}_j(\mathbb{F}_{q^a})$ we need to verify that

$$(-1)^{nm/2}\delta_{0j}q^{-anm/2} = \sum_{0\leq i\leq j} (-1)^{nm-i^2} q^{a(i^2-nm-i)/2}\begin{bmatrix} j \\ i \end{bmatrix}_{q^a},$$

or equivalently,

$$\delta_{0j} = \sum_{0\leq i\leq j} (-1)^i q^{i(i-1)/2}\begin{bmatrix} j \\ i \end{bmatrix}_q,$$

which is a standard identity on q–binomial coefficients. The second statement follows from the first by Verdier duality. $\square$

Proposition 5.8 *We have equalities*

$$\dim \mathrm{Ext}^1(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i+1})) = \dim \mathrm{Ext}^1(\mathrm{IC}(\mathcal{O}_{i+1}), \mathrm{IC}(\mathcal{O}_i)) = 1.$$

Proof By the Verdier self-duality of IC sheaves, we have an equality of dimensions

$$\dim \mathrm{Ext}^1(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i+1})) = \dim \mathrm{Ext}^1(\mathrm{IC}(\mathcal{O}_{i+1}), \mathrm{IC}(\mathcal{O}_i)),$$

so we need only give a proof for one.

Using Proposition 5.3, we have that

$$\begin{aligned}\dim \mathrm{Ext}^1(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i+1})) &= \dim \mathrm{Hom}(\mathrm{IC}(\mathcal{O}_i), \mathrm{IC}(\mathcal{O}_{i+1})[1]) \\ &= \dim \mathrm{Hom}(f_{i+1}^* \mathrm{IC}(\mathcal{O}_i), f_{i+1}^! \mathrm{IC}(\mathcal{O}_{i+1})[1]),\end{aligned}$$

since no other terms that appear in Proposition 5.3 can contribute in this degree (by the conditions for being an IC sheaf).

Recall our small resolution $\pi_i\colon \widetilde{\mathcal{O}}_i \to \overline{\mathcal{O}}_i$ from earlier. We have

$$f_{i+1}^* \mathrm{IC}(\mathcal{O}_i) = f_{i+1}^* \pi_{i*}\underline{\Bbbk}_{\widetilde{\mathcal{O}}_i}[nm-i^2] = H^*(\mathbb{P}^i)\otimes \underline{\Bbbk}_{\widetilde{\mathcal{O}}_i}[nm-i^2]$$

by the proper base change theorem and the fact that π_i is a fiber bundle with fiber $\mathbb{P}^i$ over $\mathcal{O}_{i+1}$. Thus

$$\begin{aligned}\mathrm{Hom}(f_{i+1}^* \,\mathrm{IC}(\mathcal{O}_i),\, f_{i+1}^! \,\mathrm{IC}(\mathcal{O}_{i+1})[1]) \\ &= \mathrm{Hom}(H^*(\mathbb{P}^i) \otimes \underline{\Bbbk}_{\mathcal{O}_i}[nm-i^2], \underline{\Bbbk}_{\mathcal{O}_i}[nm-(i+1)^2+1]) \\ &= \mathrm{Hom}(H^*(\mathbb{P}^i) \otimes \underline{\Bbbk}_{\mathcal{O}_i}, \underline{\Bbbk}_{\mathcal{O}_i}[-2i]) = \Bbbk\end{aligned}$$

because $\mathrm{Hom}(\underline{\Bbbk}_{\mathcal{O}_i}, \underline{\Bbbk}_{\mathcal{O}_i}[a]) = 0$ for $a < 0$, and $H^{2i}(\mathbb{P}^i) = \Bbbk$. □

Corollary 5.9 *The local chromatographic complex of $j_! \underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle$ is the unique complex of the form*

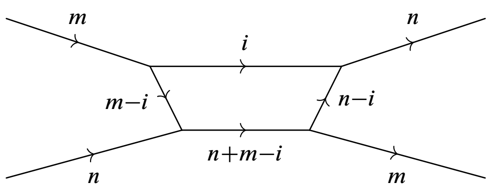
$$0 \to \mathrm{IC}(\mathcal{O}_0) \to \mathrm{IC}(\mathcal{O}_1)\langle 1\rangle \to \cdots \to \mathrm{IC}(\mathcal{O}_i)\langle i\rangle \to \cdots$$

where all differentials are nonzero. Similarly, that for $j_* \underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle$ is the unique complex of the form

$$\cdots \to \mathrm{IC}(\mathcal{O}_i)\langle -i\rangle \to \cdots \to \mathrm{IC}(\mathcal{O}_1)\langle -1\rangle \to \mathrm{IC}(\mathcal{O}_0) \to 0$$

also where all differentials are nonzero.

Remark 4 This corollary shows that this chromatographic complex categorifies the MOY expansion of a crossing in terms of trivalent graphs, with $\mathrm{IC}(\mathcal{O}_i)$ corresponding to the following MOY graph:



Proof The terms in the complex are determined by Proposition 5.7, and Proposition 5.8 implies that the isomorphism type of the complex is just determined by which maps are nonzero. Since $j_! \underline{\Bbbk}_{\mathcal{O}_0}$ and $j_* \underline{\Bbbk}_{\mathcal{O}_0}$ are indecomposable, all these maps must be nonzero by Corollary 5.5. □

6 The invariant via bimodules

6.1 The global chromatographic complex of a crossing

The following lemma gives a description of $\widetilde{\mathcal{O}}_i$ as a "Bott–Samelson" type space.

Lemma 6.1 *We have an isomorphism of $(P_{m,n}\times P_{n,m})$–equivariant varieties*

$$\widetilde{\mathcal{O}}_i \cong P_{m,n}\times_{P_{i,m-i,n}} P_{i,m+n-i}\times_{P_{i,n-i,m}} P_{n,m}.$$

Proof The map sending $[g,h,k]$ to (gV_i, ghV_n, ghk) defines a closed embedding

$$P_{m,n}\times_{P_{i,m-i,n}} P_{i,m+n-i}\times_{P_{i,n-i,m}} P_{n,m}\hookrightarrow \mathrm{Gr}_i^m\times\mathrm{Gr}_n^{n+m}\times G_{m+n}.$$

Its image is given by triples (W,V,g) satisfying $W\subset V$ and $V=gV_n$, which is isomorphic to $\widetilde{\mathcal{O}}_i$ under the map forgetting V. □

Definition 6.2 We let $R_{i_1,\dots,i_m}=\Bbbk[x_1,\dots,x_m]^{S_{i_1}\times\cdots\times S_{i_m}}$ be the rings of partially symmetric functions corresponding to Young subgroups. We will use without further mention the canonical isomorphism $R_{i_1,\dots,i_m}\cong H^*(BG_{i_1,\dots,i_m})$ sending Chern classes of tautological bundles to elementary symmetric functions.

Given a graded module or bimodule M over any ring R, we let $M(n)$ be the same module with the grading decreased by n.

Corollary 6.3 *As $(R_{m,n}\otimes R_{n,m})$–modules, we have natural isomorphisms*

$$H^*_{P_{m,n}\times P_{n,m}}(\widetilde{\mathcal{O}}_i)\cong M_i \overset{\mathrm{def}}{=} R_{i,m-i,n}\otimes_{R_{i,m+n-i}} R_{i,n-i,m},$$
$$\mathbb{H}^*_{P_{m,n}\times P_{n,m}}(\mathrm{IC}(\mathcal{O}_i))\cong M_i(nm-i^2).$$

Proof The first equality follows immediately from the main theorem of [4] (which we restated in the most convenient form for our work in our earlier paper [26, Theorem 3.3]) and Lemma 6.1. The second is a consequence of Corollary 5.2. □

Now have a global version of Proposition 5.8:

Proposition 6.4 *The spaces of bimodule maps*

$$\mathrm{Hom}_{R_{m,n}\otimes R_{n,m}}(M_i(-2i),M_{i-1})\quad\textit{and}\quad \mathrm{Hom}_{R_{m,n}\otimes R_{n,m}}(M_i(2i),M_{i+1})$$

are trivial in degrees <1, and one-dimensional in degree 1.

Proof This follows from [28, Theorem 5.4.1]. In fact, combined with Proposition 5.3, the theorem cited above implies that we have isomorphisms

$$\mathrm{Hom}_{R_{m,n}\otimes R_{n,m}}(M_i(-2i),M_{i-1})\cong \mathrm{Hom}^\bullet(\mathrm{IC}(\mathcal{O}_i),\mathrm{IC}(\mathcal{O}_{i-1})),$$
$$\mathrm{Hom}_{R_{m,n}\otimes R_{n,m}}(M_i(2i),M_{i+1})\cong \mathrm{Hom}^\bullet(\mathrm{IC}(\mathcal{O}_i),\mathrm{IC}(\mathcal{O}_{i+1})),$$

with grading degree on module maps matching the homological grading. Thus, this result is equivalent to Proposition 5.8. □

Corollary 6.5 *The global chromatographic complex of $j_!\underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle$ is the unique complex of the form*

$$\text{(6)}\qquad \boldsymbol{M}^- = \cdots \xrightarrow{\partial^-_{i+1}} M_{i+1}(nm-i(i+1)) \xrightarrow{\partial^-_i} M_i(nm-i(i-1)) \xrightarrow{\partial^-_{i-1}} \cdots$$

where all differentials are nonzero. Similarly, that for $j_*\underline{\Bbbk}_{\mathcal{O}_0}\langle nm\rangle$ is the unique complex of the form

$$\text{(7)}\qquad \boldsymbol{M}^+ = \cdots \xrightarrow{\partial^+_{i-1}} M_i(nm-i(1+i)) \xrightarrow{\partial^+_i} M_{i+1}(nm-(i+1)(i+2)) \xrightarrow{\partial^+_{i+1}} \cdots$$

also where all differentials are nonzero.

We note that these are the complexes defined in [17, Section 8], with slight change in grading shift, since they have the same modules, and there is only one such complex up to isomorphism.

We note that these maps have a geometric origin. Consider the correspondence

$$\widetilde{\mathcal{O}}_{i+1,i} = \{(U,W,g) \in \mathrm{Gr}^n_{i+1} \times \mathrm{Gr}^n_i \times G_{n+m} \mid gV_n \cap V_m \supset U \supset W\}.$$

Obviously, we have natural maps:

$$\begin{tikzcd} & \widetilde{\mathcal{O}}_{i+1,i} \arrow[dl, "p_i^1"'] \arrow[dr, "p_i^2"] & \\ \widetilde{\mathcal{O}}_{i+1} & & \widetilde{\mathcal{O}}_i \end{tikzcd}$$

Proposition 6.6 *Up to scaling, we have equalities*

$$\partial_i^- = (p_i^2)_*(p_i^1)^*, \quad \partial_i^+ = (p_i^1)_*(p_i^2)^*.$$

Proof We note that $(p_i^2)_*(p_i^1)^*$ has the expected degree and is nonzero. Thus it must be ∂_i^-. Similarly with $(p_i^1)_*(p_i^2)$. $\square$

6.2 Building the global chromatographic complex, I: via canopolises

Now, we are faced with the question of how to build the global chromatographic complex of an arbitrary braid fragment (by which we mean a tangle that can be completed to a closed braid by planar algebra operations).

While the operations we describe are nothing complicated or mysterious, it can be a bit difficult to both be precise and not pile on unnecessary notation. In an effort to give an understandable account for all readers, we give two similar, but slightly different, expositions of how to build the complex for a knot: one quite analogous to

Khovanov's exposition in [12] using braids and their closures, and one in the language of planar algebras and canopolises, in the vein of the work of Bar-Natan [2] and the first author [24].

This approach is based around planar diagrams in sense of planar algebra: a planar diagram is a crossingless tangle diagram in a planar disk with holes. A canopolis is a way of formalizing the process of building up a tangle by gluing smaller tangles into planar diagrams.

Our definition of our geometric invariant can be phrased in this language. Given a tangle T written as a union of smaller tangles T_i in a planar diagram D, the space X_T has a product decomposition $X_T \cong \prod_i X_{T_i}$, and G_T is a subgroup of $\prod_i G_{T_i}$, given by taking the diagonal inside the factors corresponding to the edges on T_i and T_j identified by D.

That is, the sheaf $\mathcal{F}_D$ can be built from the sheaves corresponding to crossings by successive applications of exterior product and restriction of groups. It is easy to understand how each of these affects chromatographic complexes, and our desired invariant can be built piece by piece.

Formally, to each oriented colored tangle diagram in a disk with boundary points $\{p_1, \dots, p_m\}$, we will associate a complex of modules over $R_\Pi = H^*\big(\prod_i BG_{p_i}\big)$, where we use Π to denote all the boundary data of the tangle (the points, their coloring, their orientation).

The association of the category $\mathcal{K}(R_\Pi\text{–mod})$ of complexes up to homotopy over R_Π to the boundary data Π (with their colorings) is a canopolis $\mathcal{K}$, where the functor associated to a planar diagram is an analogue of that used in the canopolis $\mathcal{M}_0$ in [24]. The canopolis functor

$$\tilde{\eta}\colon \mathcal{K}(R_{\Pi_1}\text{–mod}) \times \cdots \times \mathcal{K}(R_{\Pi_k}\text{–mod}) \to \mathcal{K}(R_{\Pi_0}\text{–mod}),$$

associated to a planar diagram with outer circle labeled with Π_0 and k inner circles labeled with $\Pi_1, \dots, \Pi_k$, will be given by tensoring with a complex of (R_{Π_0}, R_{Π_*})–bimodules, where $R_{\Pi_*} = R_{\Pi_1} \otimes \cdots \otimes R_{\Pi_k}$.

Let $\mathcal{A}(\eta)$ be the set of arcs in η, let α_a, ω_a be the tail and head of $a \in \mathcal{A}(\eta)$, and let n_a be the integer a is colored with. Associated to each arc is the sequence

$$(e_1(\omega_a) - e_1(\alpha_a), \dots, e_{n_a}(\omega_a) - e_{n_a}(\alpha_a)),$$

which identifies the classes $e_i \in H^*(BG_n)$ corresponding to the elementary symmetric polynomials (geometrically, these are the Chern classes of the tautological bundle on BG_n) for the endpoints connected by the arc. To our diagram, we associate the concatenation of these sequences.

Let $\kappa(\eta)$ be the Koszul complex over $R_{\Pi_0} \otimes \cdots \otimes R_{\Pi_k}$ of this concatenated sequence for our diagram η, which we think of as a bimodule with the R_{Π_0} action on the left and the R_{Π_*} on the right.

Definition 6.7 The canopolis functor $\tilde{\eta}$ associated to the diagram η is $\kappa(\eta) \otimes_{R_{\Pi_*}} -$.

Proposition 6.8 *The map sending a tangle T to the global chromatographic complex of $\mathcal{F}_T$ is a canopolis map.*

Proof We simply need to justify why tensoring with such a Koszul resolution (which is a free resolution of the diagonal bimodule for $H^*(BG_{p_i})$) is the same as changing G_T to only include the diagonal subgroup of $G_{\omega_a} \times G_{\alpha_a}$. This is one of the basic results of [4]; as we mentioned earlier, this is rephrased most conveniently for us in [26, Theorem 3.3]. □

Remark 5 We note that this construction at no point used the fact that our diagram should be a braid fragment; unfortunately, it is unclear whether our construction will be invariant under the oppositely oriented Reidemeister II move, as with Khovanov and Rozansky's original construction (see, for example, [24, Section 3]), though we will note that proving invariance under this move for the labeling with all labels 1 is sufficient to imply it for every labeling, by the same cabling arguments we will use later.

6.3 Building the global chromatographic complex, II: via bimodules

A less flexible, but perhaps more familiar, perspective is to associate to each braid a complex of bimodules, in a manner similar to [12] (though the same complex had previously appeared in other works on geometric representation theory). In the case where all labels are 1, our construction will coincide with Khovanov's.

As in Section 4.2, we let β be a braid with n strands, and $\boldsymbol{n} = (i_1, \dots, i_m)$ be the labels of the top end of the strands (so $\boldsymbol{n}\beta$ is the labeling of the bottom end). In that section, we showed that our invariant can also be described in terms of the chromatographic complex of a sheaf Φ_β on G_N.

This sheaf has the advantage that it can be built from the sheaves for smaller braids by convolution of sheaves. However, convolution of sheaves is a geometric operation which is not always easy to understand. Thus, we will give a description of it using the tensor product of bimodules. Let $F(\beta)$ be the $P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta}$–equivariant global chromatographic complex of Φ_β, considered as a complex of bimodules over $H^*(BP_{\boldsymbol{n}})$ and $H^*(BP_{\boldsymbol{n}\beta})$.

Proposition 6.9 *We have natural isomorphisms*

$$F(\beta\beta') \cong F(\beta) \otimes_{H^*(BP_{\boldsymbol{n}\beta})} F(\beta').$$

Proof Form the exterior product $\Phi_\beta \boxtimes \Phi_{\beta'}$ on $G_N \times G_N$. The $P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta} \times P_{\boldsymbol{n}\beta} \times P_{\boldsymbol{n}\beta\beta'}$–equivariant chromatographic complex of this is $F(\beta) \otimes_{\mathbb{C}} F(\beta')$. If we restrict to the diagonal $P_{\boldsymbol{n}\beta}$, then this complex is

$$F(\beta) \overset{L}{\otimes}_{H^*(BP_{\boldsymbol{n}\beta})} F(\beta').$$

By the equivariant formality of all simple, Schubert-smooth perverse sheaves on a partial flag variety, $F(\beta)$ is free as a right module, so it is not necessary to take derived tensor product.

By the convolution description, we have

$$\Phi_{\beta'\beta} \cong {}^{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta\beta'}}_{P_{\boldsymbol{n}} \times P_{\boldsymbol{n}\beta} \times P_{\boldsymbol{n}\beta\beta'}} \mu_*(\Phi_{\beta,\beta'}),$$

where $\mu\colon G_N \times G_N \to G_N$. Since $G/P_{\boldsymbol{n}\beta}$ is projective, this map simply has the effect of forgetting the $H^*(BP_{\boldsymbol{n}\beta})$ action on each page of the chromatographic spectral sequence. □

Thus, we can construct $F(\beta)$ just by knowing the complex $F(\sigma_i^{\pm 1})$ for the elementary twists $\sigma_i^{\pm 1}$. However, first we must compute the corresponding sheaves. Given $\boldsymbol{n}$, we let $Q_j = P_{i_1,\dots,i_j+i_{j+1},\dots,i_n}$ and $\mathring{Q}_j = Q_j - Q_0$.

Proposition 6.10 *We have isomorphisms*

$$\Phi_{\sigma_i} = j_* \underline{\Bbbk}_{\mathring{Q}_i} \langle i_i i_{i+1} \rangle, \quad \Phi_{\sigma_i^{-1}} = j_! \underline{\Bbbk}_{\mathring{Q}_i} \langle i_i i_{i+1} \rangle,$$

where $j\colon \mathring{Q}_i \hookrightarrow G_N$ is the obvious inclusion.

The global complex of this is very close to the complex $\boldsymbol{M}^+$ described in (6), considered as a complex of $(R_{i_i,i_{i+1}}, R_{i_{i+1},i_i})$–bimodules. However, we must extend scalars to get a complex of $(R_{\boldsymbol{n}}, R_{\sigma_i \boldsymbol{n}})$–bimodules:

Proposition 6.11 $F(\sigma_i^{\pm 1}) = R_{i_1,\dots,i_{i-1}} \otimes_{\mathbb{Q}} \boldsymbol{M}^{\pm} \otimes_{\mathbb{Q}} R_{i_{i+2},\dots,i_k}$.

Again, this is precisely the complex given in [17, Section 8] up to grading shift.

If $\boldsymbol{n}\beta = \boldsymbol{n}$, then we can close this braid to a link. Our definition of the knot invariant for this link is the equivariant chromatographic complex for the diagonal $P_{\boldsymbol{n}}$ action. By the authors' previous work [26, Theorem 1.2], this coincides with the Hochschild homology $\mathrm{HH}^*(F(\beta))$, applied termwise, of the complex $F(\beta)$.

Proposition 6.12 *The cohomology of the complex* $\mathrm{HH}^*_{R_n}(F(\beta))$ *coincides with the invariant* $\mathcal{A}_2(\widehat{\beta})$ *of the closure of the braid.*

In fact, the chromatographic spectral sequence is exactly the natural spectral sequence

$$\mathcal{H}^i(\mathrm{HH}^j(F(\beta))) \Rightarrow \mathcal{H}^{i+j}(K \otimes_{R_n \otimes R_n} F(\beta)),$$

where K *is a free resolution of* R_n *as a* $R_n \otimes R_n$*–module.*

Proof Let $\pi\colon G_N \to \mathrm{pt}$, and consider the object $\pi_* \Phi_\beta$ in the equivariant derived category $D_{P_n \times P_n}(\mathrm{pt})$. Under the equivalence to R_n–dg-bimodules given in [25, Theorem 7], this is sent to the complex $F(\sigma)$. Similarly, the weight filtration is sent to that induced by thinking of $F(\beta)$ as a complex. Thus, the spectral sequences match under this equivalence. □

Since $\mathcal{H}^*(\mathrm{HH}^*(F(\beta)))$ is precisely the invariant proposed by [17], Theorem 1.4 follows immediately.

7 Decategorification

We also wish to show that our knot invariant is, in fact, a categorification of the HOMFLYPT polynomial.

7.1 A categorification of the Hecke algebra

This requires a few basic results about the relationship between sheaves on G_n and the Hecke algebra H_n. As usual, $B = P_{1,\dots,1}$ is the standard Borel.

Definition 7.1 The Hecke algebra H_n is the algebra over $\mathbb{Z}[q^{1/2}, q^{-1/2}]$ given by the quotient of the group algebra of the braid group $\mathcal{B}_n$ by the quadratic relation

$$(\sigma_i + q^{1/2})(\sigma_i - q^{-1/2}) = 0$$

for each elementary twist σ_i.

Proposition 7.2 [14] *The Grothendieck group* $K^0(D^b_{B\times B}(G_n))$ *of the equivariant derived category* $D^b_{B\times B}(G_n)$ *is isomorphic to the Hecke algebra* H_n*, with the convolution product decategorifying to the algebra product in* H_n*.*

This map is fixed by the assignment

$$[j_* \underline{\Bbbk}_{Bs_iB}] \mapsto q^{1/2}\sigma_i,$$

where $j\colon Bs_iB \hookrightarrow G_n$ *is the obvious inclusion.*

Let $\mathcal{F}$ be a $B \times B$–equivariant sheaf on G_n. Then we have a map

$$\mathcal{E}_B(G;\mathcal{F}) = \sum_{i,j,k} (-1)^\ell q^{j/2} t^k \dim \mathbb{H}_{B_\Delta}^{j-\ell;j-k}(\mathrm{gr}_\ell^W \mathcal{F}),$$

sending the class of $\mathcal{F}$ in the Grothendieck group to the bigraded Euler characteristic of its global chromatographic complex, often called the *mixed Hodge polynomial*.

This map agrees with a previously known trace on the Hecke algebra, a fact that the authors have proven in a separate note, due to its independent interest and separate connection to the question of constructing Markov traces on general Hecke algebras.

Proposition 7.3 [27, Theorem 1] *The map $\mathcal{E}_B(G_n;-)$ is the Jones–Ocneanu trace* Tr *(see [11]) on H_n with appropriate normalization factors.*

Remark 6 This geometric definition applies equally well to any simple Lie group, and defines a canonical trace on the Hecke algebra for any type. In fact, our construction can be modified in a straightforward way to a "triply graded homology" invariant on all Artin braid groups. In type B, this can be interpreted as a homological knot invariant for knots in the complement of a solid torus.

7.2 Decategorification for colored HOMFLYPT

To apply this result, we must relate our construction to the categorification of the Hecke algebra above. Recall that if σ is a braid with all labels 1, then Φ_σ is an object of $D^b_{B\times B}(G_n)$.

Proposition 7.4 *The class $[\Phi_\sigma] \in \mathsf{H}_n$ is the image of σ under the natural map $\mathcal{B}_n \to \mathsf{H}_n$.*

This, combined with Proposition 7.3, gives a new proof of the result of Khovanov [12] that when all components are labeled with 1, the invariant

$$\mathcal{E}(L) = \mathcal{E}_{G_D}(X_D;\mathcal{F}_D) = \sum_{i,j,k} (-1)^\ell q^j t^k \dim \mathcal{A}_2^{j;k;\ell}(L)$$

is the appropriately normalized HOMFLYPT polynomial of the link L underlying the diagram D. We wish to extend this to the colored case. For this, we must use a "cabling/projection" formula.

Consider a closable colored braid σ, and let $P = P_{\boldsymbol{n}}$ and $G = G_N$. We have defined a $P\times P$–equivariant sheaf Φ_σ on G by the multiplication map $m\colon X_\sigma \to G$.

Theorem 7.5 *For any colored link L, the Euler characteristic $\mathcal{E}(D)$ is the (suitably normalized) colored HOMFLYPT polynomial for any diagram D of L.*

In order to prepare for the proof, we show a pair of lemmata. Let σ_{cab} denote the cabling of σ in the blackboard framing with multiplicities given by the colorings, thought of as colored with all labels 1.

Lemma 7.6 *We have an isomorphism of $P \times B$–equivariant sheaves,*

$$\mathrm{res}_{P\times B}^{P\times P} \Phi_\sigma \cong \mathrm{ind}_{B\times B}^{P\times B} \Phi_{\sigma_{\mathrm{cab}}}.$$

Proof The proof is a straightforward induction on the length of σ, left to the reader. □

Let $\lambda_{\boldsymbol{n}}$ be the partition given by arranging the parts of $\boldsymbol{n}$ in decreasing order, and let $\lambda_{\boldsymbol{n}}^t$ be its transpose. Let $\pi_{\boldsymbol{n}}$ be the projection in the Hecke algebra to the representations indexed by Young diagrams less than $\lambda_{\boldsymbol{n}}^t$ in dominance order. Alternatively, if we identify H_N with the endomorphisms of $V^{\otimes N}$, where V is the standard representation of $U_q(\mathfrak{sl}_m)$ for $m \geq n$, then this is the projection to $\bigwedge^{i_1} V \otimes \cdots \otimes \bigwedge^{i_n} V$.

Let $q_P = \sum_{W_P} q^{\ell(w)}$ be the Poincaré polynomial of the flag variety P/B.

Lemma 7.7 *For every complex Φ in $D^b_{B\times B}(G)$, we have*

$$[\mathrm{res}_{P\times B}^{B\times B} \mathrm{ind}_{B\times B}^{P\times B} \Phi] = q_P \pi_P[\Phi].$$

Proof First consider the case where $P = G$. In this case, the sheaf $\mathrm{res}_{G\times B}^{B\times B} \mathrm{ind}_{B\times B}^{G\times B} \Phi$ has a filtration whose successive quotients are of the form $\mathbb{H}^i(\Phi) \otimes \underline{\Bbbk}_G$. Thus we have

$$[\mathrm{res}_{G\times B}^{B\times B} \mathrm{ind}_{B\times B}^{G\times B} \Phi] = \dim_q \mathbb{H}^*(\Phi) \cdot [\underline{\Bbbk}_G].$$

It is a classical fact that $[\underline{\Bbbk}_G] = q_G \pi_G$; here π_G is just the projection to $\bigwedge^N V$. This computation immediately extends to the general case. □

Remark 7 This proposition shows why our approach works for colored HOMFLYPT polynomials, but would need to be modified to approach the HOMFLY polynomials for more general type A representations; we lack a good categorification of most of the projections in the Hecke algebra, but π_P has a beautiful geometric counterpart. This may be related to the fact that π_P is the projection not just to a subrepresentation, but in fact to a cellular ideal in $\mathsf{H}_{\boldsymbol{n}}$.

Proof of Theorem 7.5 Immediately from Lemmata 7.6 and 7.7, we have the equality of Grothendieck classes $[\mathrm{res}_{B\times B}^{P\times P} \Phi_\sigma] = q_P \pi_P [\Phi_{\sigma_{\mathrm{cab}}}]$. Thus

$$\begin{aligned}\mathcal{E}_P(G;\Phi_\sigma) &= q_P^{-1} \mathcal{E}_B(G; \mathrm{res}_{B\times B}^{P\times P} \Phi_\sigma) \\ &= \mathrm{Tr}(q_P^{-1}[\mathrm{res}_{B\times B}^{P\times P} \Phi_\sigma]) \\ &= \mathrm{Tr}(\pi_P[\Phi_{\sigma_{\mathrm{cab}}}]).\end{aligned}$$

By the "projection/cabling" formula (see, for example, [16, Lemma 3.3]), this is precisely the colored HOMFLYPT polynomial. □

8 The proof of invariance: the 1–colored case

We first concentrate on the simpler case of GL(2) before attacking the general case. In this case, we will obtain an invariant which matches the HOMFLYPT homology of Khovanov and Rozansky [13; 12], so the section below can be thought of as a geometric proof of the invariance of this homology theory.

Recall that if σ is a braidlike diagram on n strands, we described in Section 4.2 a map

$$m\colon X_\sigma \to G_n$$

which is equivariant with respect to $\phi\colon G_\sigma \to T\times T$, where $T\times T$ acts on G_n by left and right multiplication. This map gives rise to a functor

$$ {}^{B\times B}_{G_\sigma} m_*\colon D^+_{G_D}(X_D) \to D^+_{T\times T}(G_n),$$

and we denoted the image of $\mathcal{F}_\sigma$ by Φ_σ. We saw that this functor preserves weight filtrations.

Now suppose that w is an element of the symmetric group on n letters (which we regard as permutation matrices in G_n) and that $\sigma = \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p}$ is a (positive) braid in the standard generators corresponding to a reduced expression $s_{i_1}\cdots s_{i_p}$ for w.

It is straightforward to see that if we restrict m to the open set $\widetilde{U}$ in G_D consisting of tuples $(g_1,\dots,g_p)$ with each $g_i\in U$, where U denotes the open Bruhat cell in G_2, then we may factor m as

$$\widetilde{U} \to \widetilde{U}/\ker\phi \to G_n, \tag{8}$$

where the first map is a quotient by a free action, and the second map is an isomorphism.

Moreover, if we denote by B the subgroup of upper triangular matrices, then the image of the restriction of m to $\widetilde{U}$ is contained in Schubert cell BwB. It follows that

$$\Phi_\sigma = j_{w*}\Bbbk_{BwB}\langle \ell(w)\rangle, \tag{9}$$

where j_w denotes the inclusion of the Bruhat cell BwB into G_n.

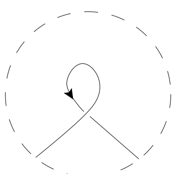
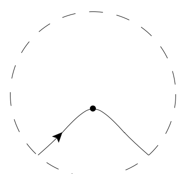
Proposition 8.1 *Theorem 1.2 holds in the case where all strands are labeled by* 1*.*

Proof As usual with proofs that knot invariants defined in terms of a projection are really invariants, we check that our description is unchanged by the Reidemeister moves. Since we only consider closed braids, we only need to check Reidemeister II and III in the braid-like case, when all strands are coherently oriented. Those who prefer to use the Markov theorem can consider the proof of Reidemeister I as a proof of the Markov 1 move, and the Reidemeister II and III calculations as proving the independence of the presentation of our braid in terms of elementary twists *and* of the Markov 2 move (which only uses Reidemeister IIa).

In each case, we will use the fact that while we wish to compare the pushforwards of sheaves corresponding to diagrams D and D' from X_D/G_D and $X_{D'}/G_{D'}$ to a point, we can accomplish this by showing that their pushforwards by any pair of maps to any common space coincide. Being able to use these techniques is one of the principal advantages of a geometric definition over a purely algebraic one.

In each case, the calculation we need to do is local in terms of diagrams. Proposition 6.8 implies that if we show that we have an isomorphism of global chromatographic complexes of two diagrams as modules over the polynomial rings attached to external edges, then "pasting" these into a fixed larger diagram again gives an isomorphism of global chromatographic complexes.

Reidemeister I Consider the following tangles:

(10) $D =$  $D' =$ 

To simplify notation we denote the associated varieties by X and X' and groups by G and G', respectively. We have $X = G_2$ and $X' = G_1$, $G = G_1^3$ and $G' = G_1^2$. The determinant gives a map

$$d\colon X \to X',$$

which is equivariant with respect to the map $\phi\colon G \to G'$ forgetting the factor corresponding to the internal edge. We wish to exhibit an isomorphism

$$\,^{G'}_{G} d_* \mathcal{F}_D \cong \mathcal{F}_{D'} \tag{11}$$

compatible with the weight filtrations on both sheaves. Note that the weight filtration on $\mathcal{F}_{D'}$ is trivial, whereas that on $\mathcal{F}_D$ is not.

Let $B \overset{a}{\hookrightarrow} X \overset{b}{\hookleftarrow} BsB$ be the decomposition of $X = G_2$ into its two Bruhat cells. We have a distinguished triangle

$$a_! a^! \underline{\Bbbk}_X \langle 1 \rangle \to \underline{\Bbbk}_X \langle 1 \rangle \to b_* b^* \underline{\Bbbk}_X \langle 1 \rangle \xrightarrow{[1]} .$$

Because a is the inclusion of a smooth divisor, $a^! \underline{\Bbbk}_X = \underline{\Bbbk}_X \langle -2 \rangle = \underline{\Bbbk}_X[-2](-1)$. Hence

$$a^! \underline{\Bbbk}_X \langle 1 \rangle = \underline{\Bbbk}_X[-1](-\tfrac{1}{2}).$$

Turning the triangle gives the weight filtration on $b_* \underline{\Bbbk}_{BsB} \langle 1 \rangle$:

$$\underline{\Bbbk}_X \langle 1 \rangle \to b_* \underline{\Bbbk}_{BsB} \langle 1 \rangle \to a_* \underline{\Bbbk}_B (-\tfrac{1}{2}) \xrightarrow{[1]} . \tag{12}$$

The left (resp. right) hand term is pure of weight 0 (resp. 1). In the following we analyze the effect of ${}^{G'}_{G}d_*$ on this triangle.

The restriction of d to $BsB \subset X$ is a trivial $G_1 \times \mathbb{A}^2$–bundle over X'. One may easily check that $\ker \phi$ acts freely on the multiplicative group in the fiber. It follows that

$${}^{G'}_{G}d_* b_* \underline{\Bbbk}_{BsB} \cong \underline{\Bbbk}_{X'}.$$

On the other hand, the restriction of d to $B \subset X$ yields a trivial $G_1 \times \mathbb{A}^1$–bundle, with $\ker \phi$ only acting on $\mathbb{A}^1$. It follows that

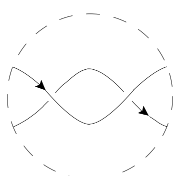
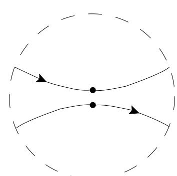
$${}^{G'}_{G}d_* a_* \underline{\Bbbk}_B = H^\bullet(\mathbb{P}^\infty) \otimes H^\bullet(G_1) \otimes \underline{\Bbbk}_{X'}.$$

Applying ${}^{G'}_{G}d_*$ to (12) and using the above isomorphisms, we obtain

$${}^{G'}_{G}d_* \underline{\Bbbk}_X \langle 1 \rangle \to \underline{\Bbbk}_{X'} \langle 1 \rangle \to H^\bullet(\mathbb{P}^\infty) \otimes H^\bullet(G_1) \otimes \underline{\Bbbk}_{X'} (-\tfrac{1}{2}) \xrightarrow{[1]} .$$

As $\mathrm{Hom}(\underline{\Bbbk}_{X'}, \underline{\Bbbk}_{X'}[i]) = H^i_{G'}(X')$ is zero for $i < 0$ we conclude that the second arrow above is zero. Thus, the induced weight filtration on $\underline{\Bbbk}_{X'}$ is trivial. Thus, we have the desired Equation (11). As discussed before, the general case follows from Section 6.2, where we think of adding the rest of the diagram as a canopolis operation.

Reidemeister IIa Here we are concerned with the following two tangles:

$D =$  $\qquad D' =$ 

We denote the associated varieties and groups X, X', G, G'. We denote by m the multiplication map $X \to G_2$ considered at the start of this section. We regard X' as the diagonal matrices inside G_2.

We have seen that ${}^{G'}_{G}m_*$ preserves weight filtrations, and hence we may ignore weight filtrations when comparing ${}^{G'}_{G}m_*\mathcal{F}_D$ and $\mathcal{F}_{D'}$. The map $B \to X'$ forgetting the off-diagonal entry is acyclic, and therefore it is enough to show that ${}^{G'}_{G}m_*\mathcal{F}_D \cong \underline{\Bbbk}_B$.

We decompose G_2 into its Bruhat cells $B \overset{a}{\hookrightarrow} G_2 \overset{b}{\hookleftarrow} BsB$ as before. We claim we have isomorphisms

$$ {}^{G'}_{G}m_*(a_*\underline{\Bbbk}_B \boxtimes b_!\underline{\Bbbk}_{BsB}) \cong b_!\underline{\Bbbk}_{BsB}, \tag{13}$$

$$ {}^{G'}_{G}m_*(\underline{\Bbbk}_G \boxtimes a_*\underline{\Bbbk}_B) \cong \underline{\Bbbk}_G, \tag{14}$$

$$ {}^{G'}_{G}m_*(\underline{\Bbbk}_G \boxtimes \underline{\Bbbk}_G) \cong \underline{\Bbbk}_G \oplus \underline{\Bbbk}_G\langle -2\rangle, \tag{15}$$

$$ {}^{G'}_{G}m_*(\underline{\Bbbk}_G \boxtimes b_!\underline{\Bbbk}_{BsB}) \cong \underline{\Bbbk}_G\langle -2\rangle. \tag{16}$$

(As always we regard the exterior tensor product of equivariant sheaves on G_2 as an equivariant sheaf on X via restriction.)

Indeed, (13) and (14) follow from the fact that the restriction of m to $B\times G$ or $G\times B$ is a trivial B–bundle, with $\ker\phi$ acting freely on the multiplicative groups in the fiber. The factorization (8) of m as "essentially a $\mathbb{P}^1$–bundle" implies (15). Then (16) follows from the others by taking the exterior tensor product of $\underline{\Bbbk}_G$ with the distinguished triangle $b_!\underline{\Bbbk}_{BsB} \to \underline{\Bbbk}_G \to a_*\underline{\Bbbk}_B \to$ and applying ${}^{G'}_{G}m_*$.

Now B is smooth of codimension 1 inside G_2 so $a^!\underline{\Bbbk}_G = \underline{\Bbbk}_B\langle -2\rangle$ and we have an exact triangle

$$ a_*\underline{\Bbbk}_B\langle -2\rangle \to \underline{\Bbbk}_G \to b_*\underline{\Bbbk}_{BsB} \xrightarrow{[1]} .$$

Taking the exterior tensor product with $b_!\underline{\Bbbk}_{BsB}$, applying ${}^{G'}_{G}m_*$ and using the above isomorphisms we obtain a distinguished triangle

$$ b_!\underline{\Bbbk}_{BsB}\langle -2\rangle \to \underline{\Bbbk}_G\langle -2\rangle \to {}^{G'}_{G}m_*(b_*\underline{\Bbbk}_{BsB} \boxtimes b_!\underline{\Bbbk}_{BsB}) \xrightarrow{[1]} . \tag{17}$$

Note that $\mathrm{Hom}(b_!\underline{\Bbbk}_{BsB}, \underline{\Bbbk}_G)$ is one-dimensional and contains the adjunction morphism $b_!b^!\underline{\Bbbk}_G \to \underline{\Bbbk}_G$. By considering its dual, one may show that the first arrow in (17) is nonzero. It follows that this arrow is the adjunction morphism (up to a nonzero scalar) and we have an isomorphism

$$ {}^{G'}_{G}m_*(b_*\underline{\Bbbk}_{BsB} \boxtimes b_!\underline{\Bbbk}_{BsB}) \cong \underline{\Bbbk}_B\langle -2\rangle. $$

Finally note that by definition $\mathcal{F}_D$ is $b_*\underline{\Bbbk}_{BsB} \boxtimes b_!\underline{\Bbbk}_{BsB}\langle 2\rangle$ and so

$$ {}^{G'}_{G}m_*\mathcal{F}_D \cong \underline{\Bbbk}_B, $$

which finishes the proof of invariance under Reidemeister II.

Reidemeister III This follows immediately from the considerations at the beginning of this section. Indeed, if σ and σ' are the diagrams corresponding to the words $\sigma_1\sigma_2\sigma_1$ and $\sigma_2\sigma_1\sigma_2$ we have maps

$$X_\sigma \xrightarrow{m} G_3 \xleftarrow{m'} X_{\sigma'}$$

and we have

$${}^{T\times T}_{G_\sigma} m_* \mathcal{F}_\sigma \cong j_{w_0} \underline{\Bbbk}_{Bw_0B} \cong {}^{T\times T}_{G_\sigma} m'_* \mathcal{F}_{\sigma'},$$

where w_0 indicates the longest element in S_3. □

9 The proof of invariance: arbitrary labels

Now, we expand to the full case of all possible positive integer labels.

Proof of Theorem 1.2 All of the Reidemeister moves can simply be reduced to the corresponding statement for the cabling with all labels 1. Interestingly, the same trick was used in [17] to prove invariance in a special case. Almost certainly our proof could be rephrased in a purely algebraic language like their paper, though at the moment it is unclear how.

Reidemeister IIa & III Here we need only establish the isomorphisms of $P\times P$–equivariant sheaves

$$\Phi_{\sigma_i} \star \Phi_{\sigma_i^{-1}} \cong \underline{\Bbbk}_P \quad \text{and} \quad \Phi_{\sigma_i} \star \Phi_{\sigma_{i+1}} \star \Phi_{\sigma_i} \cong \Phi_{\sigma_{i+1}} \star \Phi_{\sigma_i} \star \Phi_{\sigma_{i+1}}.$$

Lemma 7.6 implies that these hold as $P\times B$–equivariant sheaves, applying the invariance for the cabling with all labels 1.

In fact, both are the $*$–inclusion of a local system on a $P\times P$–orbit: P itself in the first case, and the $P\times P$ orbit of the permutation corresponding to the cabling of $\sigma_i\sigma_{i+1}\sigma_i$ in the second. Since the stabilizer of any point under $P\times P$ is connected, any $P\times B$–equivariant local system on an orbit has at most one $P\times P$–equivariant structure, and this equality holds as $P\times P$–equivariant sheaves.

Reidemeister I We again use the "cabling/projection" philosophy, but this argument requires a bit more subtlety. We are interested in the chromatographic complex of a single crossing with its right ends capped off; that is, the tangle projection denoted by D in (10). To construct the sheaf $\mathcal{F}_D$, we take $U \subset G_{2n}$, as defined in (5), and consider $j_*\underline{\Bbbk}_U\langle n^2\rangle$ or $j_!\underline{\Bbbk}_U\langle n^2\rangle$, depending on whether our crossing is positive or negative. These cases are Verdier dual, and the proofs of invariance are essentially identical, so we will treat the positive case, and only note where the negative differs. If we consider this sheaf equivariantly for the action of $G_{n,n}$ on the left and the right, then we obtain the sheaf attached to a single crossing with label n on both strands.

By convention, we let G^1 denote the first copy of $G_n \subset G_{n,n}$ and G^2 the second. As before, we let T_n be diagonal matrices in G_n, and we use T^1, T^2 for the inclusions into the two factors. We let $G^{1,1,2}$ denote $G^1 \times G^1 \times (G^2)_\Delta$; that is, the left and right action of G^1, and the conjugation action of G^2. This is the group G_D for the diagram labeled D in (10). The sheaf $\mathcal{F}_D$ for this diagram is thus $j_! \underline{\Bbbk}_U \langle n^2 \rangle$ (or $j_* \underline{\Bbbk}_U \langle n^2 \rangle$ if D is taken with a positive crossing) considered equivariantly for $G^{1,1,2}$.

Thus in order to prove the theorem, what we must do is consider the $G^{1,1,2}$–equivariant global chromatographic complex of $\mathcal{F}_D$ as a $H^*(BG^1)$–bimodule, and show that it matches that of an untwisted strand (the diagram denoted D' in (10)).

Note that for any G_n sheaf $\mathcal{F}$ on any G_n–space X, the inclusion of the symmetric group as permutation matrices normalizing T_n gives an action of S_n on $\mathbb{H}^*_{T_n}(X; \operatorname{res}^{G_n}_{T_n} \mathcal{F})$.

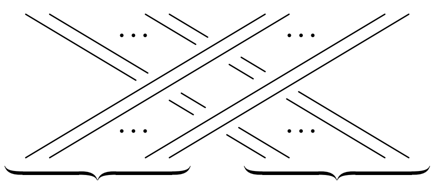
Lemma 9.1 *The natural transformation of functors*

$$\mathbb{H}^*_{G^{1,1,2}}(G_{2n}; -) \to \mathbb{H}^*_{G^{1,1} \times T^2}(G_{2n}; \operatorname{res}^{G^{1,1,2}}_{G^{1,1} \times T^2} -)$$

is the inclusion of the S_n–invariants for the permutation action on T^2.

Proof This is the abelianization theorem for equivariant cohomology; see, for example, [7, Proposition 1]. □

Let $\widehat{U}$ be the Bruhat cell $B w^{n,n}_{2n} B$, where $w^{n,n}_{2n}$ is the permutation which switches i and $i \pm n$, and let $\widehat{j}$ be its inclusion to G_{2n}. We note that $\widehat{j}_* \underline{\Bbbk}_{\widehat{U}}$ is Φ_σ where σ is the braid given by the n–cabling of a single crossing:



Lemma 9.2 *The $G^{1,1} \times T^2$–equivariant global chromatographic complex of $j_* \underline{\Bbbk}_U$ is isomorphic to the $T^{1,1} \times T^2$–equivariant one for $\widehat{j}_* \underline{\Bbbk}_{\widehat{U}}$, with the bimodule structure restricted to $H^*(BG^{1,1}) \subset H^*(BT^{1,1})$.*

Proof Let $Q = G^1 \cap B$ be the upper triangular matrices in G_n. Then

$$\operatorname{ind}^{G^{1,1} \times T^2}_{T^{1,1} \times T^2} j_* \underline{\Bbbk}_{\widehat{U}} \cong \operatorname{ind}^{G^{1,1} \times T^2}_{Q \times Q \times T^2} \operatorname{ind}^{Q \times Q \times T^2}_{T^{1,1} \times T^2} j_* \underline{\Bbbk}_{\widehat{U}} \cong \operatorname{res}^{G^{1,1,2}}_{T^{1,1} \times T^2} j_* \underline{\Bbbk}_U.$$

The first induction leaves chromatographic complexes unchanged, since Q and T^1 are homotopy equivalent, and $j_* \underline{\Bbbk}_{\widehat{U}}$ is smooth on $Q \times Q$–orbits.

For the second, we have a projective map

$$\mu\colon G_n \times_Q \overline{\widehat{U}} \times_Q G_n \to G_{2n},$$

which induces an isomorphism

$$G_n \times_Q \widehat{U} \times_Q G_n \cong U.$$

By [25, Theorem 5], under taking equivariant cohomology, induction of sheaves corresponds to the restriction of scalars, and since G_n/Q is projective, pushforward preserves purity and thus commutes with taking local chromatographic complex. This means that the result extends to all terms in the chromatographic spectral sequence. □

By definition, the $T^{1,1}\times T^2$–equivariant chromatographic complex for $\widehat{j}_*\Bbbk_{\widehat{U}}$ is just the complex of bimodules for the tangle diagram D_{cab} corresponding to closing the right half of the strands in the braid above. Applying the invariance result for labelings with all labels 1, this is the same as the complex corresponding to a full twist of n strands. Since $\widehat{j}_*\Bbbk_{\widehat{U}}$ is in fact equivariant for $T^{1,1}\times G^2$, this has an S_n action, which is compatible with its module structure over $H^*(BT^2) \cong \Bbbk[x_1,\dots,x_n]$. Doing this straightening one strand at a time, we see that the actions of $H^*(BT^2)$ and $H^*(BT^1) \cong \Bbbk[y_1,\dots,y_n]$ are intertwined by the map sending x_i to y_{n+1-i}. Thus, the S_n action discussed above is compatible with the standard S_n–module action on $H^*(BT_n)$ acting on the left and right after conjugation by the longest element w_0.

Note that if we consider a negative crossing, we will have to include n times the usual shift for removing a negative stabilization, but this is easily accounted for in the normalization.

Restricted to symmetric polynomials (that is, $H^*(BG_n)$), every Soergel bimodule is a number of copies of the regular bimodule, and every map in the complex for a single crossing splits, so restricted to $H^*(BG_n)$, the complex attached to a braid with all labels 1 is homotopic to a single copy of $H^*(BT_n)$ with the regular bimodule action and standard S_n action (conjugated by the longest element w_0).

By Lemma 9.1, to obtain the $G^{1,1,2}$–equivariant global chromatographic complex we simply take S_n–invariants and thus we obtain a single copy of the regular bimodule for $H^*(BG_n)$, as desired. □

References

[1] **M Artin**, **A Grothendieck**, **J L Verdier**, *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (SGA 4_3), Lecture Notes in Math. 305, Springer (1973) MR

[2] **D Bar-Natan**, *Khovanov's homology for tangles and cobordisms*, Geom. Topol. 9 (2005) 1443–1499 MR

[3] **A Beilinson**, **V Ginzburg**, **W Soergel**, *Koszul duality patterns in representation theory*, J. Amer. Math. Soc. 9 (1996) 473–527 MR

[4] **J Bernstein**, **V Lunts**, *Equivariant sheaves and functors*, Lecture Notes in Math. 1578, Springer (1994) MR

[5] **A A Beĭlinson**, **J Bernstein**, **P Deligne**, *Faisceaux pervers*, from "Analysis and topology on singular spaces, I", Astérisque 100, Soc. Math. France, Paris (1982) 5–171 MR

[6] **M V Bondarko**, *Weight structures and motives; comotives, coniveau and Chow-weight spectral sequences, and mixed complexes of sheaves: a survey*, preprint (2009) arXiv

[7] **M Brion**, *Equivariant cohomology and equivariant intersection theory*, from "Representation theories and algebraic geometry" (A Broer, G Sabidussi, editors), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 514, Kluwer Acad., Dordrecht (1998) 1–37 MR

[8] **P Deligne**, *Théorie de Hodge, II*, Inst. Hautes Études Sci. Publ. Math. (1971) 5–57

[9] **P Deligne**, *Cohomologie étale* (SGA 4 1/2), Lecture Notes in Math. 569, Springer (1977) MR

[10] **P Deligne**, *La conjecture de Weil, II*, Inst. Hautes Études Sci. Publ. Math. (1980) 137–252 MR

[11] **V F R Jones**, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. 126 (1987) 335–388 MR

[12] **M Khovanov**, *Triply-graded link homology and Hochschild homology of Soergel bimodules*, Internat. J. Math. 18 (2007) 869–885 MR

[13] **M Khovanov**, **L Rozansky**, *Matrix factorizations and link homology, II*, Geom. Topol. 12 (2008) 1387–1425 MR

[14] **R Kiehl**, **R Weissauer**, *Weil conjectures, perverse sheaves and l–adic Fourier transform*, Ergeb. Math. Grenzgeb. 42, Springer (2001) MR

[15] **G Laumon**, *Transformation de Fourier, constantes d'équations fonctionnelles et conjecture de Weil*, Inst. Hautes Études Sci. Publ. Math. (1987) 131–210 MR

[16] **X-S Lin**, **H Zheng**, *On the Hecke algebras and the colored HOMFLY polynomial*, Trans. Amer. Math. Soc. 362 (2010) 1–18 MR

[17] **M Mackaay**, **M Stošić**, **P Vaz**, *The $1,2$–coloured HOMFLY-PT link homology*, Trans. Amer. Math. Soc. 363 (2011) 2091–2124 MR

[18] **J McCleary**, *A user's guide to spectral sequences*, 2nd edition, Cambridge Studies in Advanced Math. 58, Cambridge Univ. Press (2001) MR

[19] **H Murakami**, **T Ohtsuki**, **S Yamada**, *Homfly polynomial via an invariant of colored plane graphs*, Enseign. Math. 44 (1998) 325–360 MR

[20] **D Pauksztello**, *Compact corigid objects in triangulated categories and co–t–structures*, Cent. Eur. J. Math. 6 (2008) 25–42 MR

[21] **C A M Peters**, **J H M Steenbrink**, *Mixed Hodge structures*, Ergeb. Math. Grenzgeb. 52, Springer (2008) MR

[22] **M Saito**, *Mixed Hodge modules*, Proc. Japan Acad. Ser. A Math. Sci. 62 (1986) 360–363 MR

[23] **C Stroppel**, *A structure theorem for Harish-Chandra bimodules via coinvariants and Golod rings*, J. Algebra 282 (2004) 349–367 MR

[24] **B Webster**, *Khovanov–Rozansky homology via a canopolis formalism*, Algebr. Geom. Topol. 7 (2007) 673–699 MR

[25] **B Webster**, **G Williamson**, *The bounded below equivariant derived category*, in preparation

[26] **B Webster**, **G Williamson**, *A geometric model for Hochschild homology of Soergel bimodules*, Geom. Topol. 12 (2008) 1243–1263 MR

[27] **B Webster**, **G Williamson**, *The geometry of Markov traces*, preprint (2009) arXiv

[28] **G Williamson**, *Singular Soergel bimodules*, Int. Math. Res. Not. 2011 (2011) 4555–4632 MR

Department of Pure Mathematics, University of Waterloo
and Perimeter Institute for Theoretical Physics
Waterloo ON, Canada

School of Mathematics and Statistics, University of Sydney
Sydney NSW, Australia

ben.webster@uwaterloo.ca, g.williamson@sydney.edu.au

Proposed: Peter S. Ozsváth
Seconded: Ciprian Manolescu, Haynes Miller

Received: 28 September 2010
Revised: 25 June 2016

msp

Geometry & Topology 21 (2017) 2601–2681

Categorical cell decomposition of quantized symplectic algebraic varieties

GWYN BELLAMY
CHRISTOPHER DODD
KEVIN MCGERTY
THOMAS NEVINS

We prove a new symplectic analogue of Kashiwara's equivalence from $\mathcal{D}$–module theory. As a consequence, we establish a structure theory for module categories over deformation-quantizations that mirrors, at a higher categorical level, the Białynicki-Birula stratification of a variety with an action of the multiplicative group $\mathbb{G}_m$. The resulting *categorical cell decomposition* provides an algebrogeometric parallel to the structure of Fukaya categories of Weinstein manifolds. From it, we derive concrete consequences for invariants such as K–theory and Hochschild homology of module categories of interest in geometric representation theory.

53D55; 14F05

1 Introduction

Since the 1970s, categories of (ordinary or twisted) $\mathcal{D}$–modules on algebraic varieties and stacks have become fundamental tools in geometric representation theory; see Beĭlinson and Bernstein [3]. More recently, an emerging body of important work in geometric representation theory relies on sheaves over deformation-quantizations of symplectic algebraic varieties more general than the cotangent bundles whose deformation-quantizations give rise to $\mathcal{D}$–modules; see Bellamy and Kuwabara [5], Bezrukavnikov and Kaledin [8], Bezrukavnikov and Losev [9], Braden, Proudfoot and Webster [13], Dodd and Kremnitzer [18], Gordon and Losev [26], Kaledin [35], Kashiwara and Rouquier [40] and McGerty and Nevins [53]. A sophisticated theory of such quantizations now exists thanks to the efforts of many (see Bezrukavnikov and Kaledin [7], D'Agnolo and Kashiwara [15], D'Agnolo and Polesello [16], D'Agnolo and Schapira [17], Kashiwara [39], Kashiwara and Schapira [41; 42] and Nest and Tsygan [60; 61] among many others).

The present paper establishes a structure theory for deformation-quantizations that mirrors, at a higher categorical level, the fundamental Białynicki-Birula stratification

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2601

of a variety with an action of the multiplicative group $\mathbb{G}_m$ and the corresponding decomposition of its cohomology. In the most prominent examples, the resulting *categorical cell decomposition* has many immediate and concrete consequences for invariants such as K–theory and Hochschild homology; it also makes possible the extension of powerful tools from $\mathcal{D}$–module theory, such as the Koszul duality relating $\mathcal{D}$–modules to dg modules over the de Rham complex, to a more general symplectic setting; see Bellamy, Dodd, McGerty and Nevins [4]. The structures that we identify parallel those described for Fukaya–type categories in real symplectic geometry by Nadler [55]. We derive these structures on module categories from a new symplectic analogue of Kashiwara's equivalence for $\mathcal{D}$–modules.

In Section 1.1 we describe an enhancement of the Białynicki-Birula decomposition for symplectic varieties with a nice $\mathbb{G}_m$–action. Section 1.2 explains our categorical cell decomposition for sheaves on the quantizations of such varieties; in Section 1.3, we lay out the symplectic Kashiwara equivalence that underlies categorical cell decomposition. Section 1.4 describes basic categorical consequences. Section 1.5 provides immediate applications of this structure theory for module categories of deformation-quantizations. Section 1.6 explores parallels with Fukaya categories.

1.1 Symplectic varieties with elliptic $\mathbb{G}_m$–action

We work throughout the paper over $\mathbb{C}$. Let $\mathfrak{X}$ be a smooth, connected symplectic algebraic variety with symplectic form ω.

Definition 1.1 A $\mathbb{G}_m$–action on $\mathfrak{X}$ is said to be *elliptic* if the following hold:

(1) $\mathbb{G}_m$ acts with positive weight on the symplectic form: $m_t^*\omega = t^l\omega$ for some $l > 0$.

(2) For every $x \in \mathfrak{X}$, the limit $\lim_{t\to\infty} t\cdot x$ exists in $\mathfrak{X}$.

We remark that if we assume that ω is rescaled by $\mathbb{G}_m$ with some weight $l \in \mathbb{Z}$, then the existence of limits already implies that $l \geq 0$.

Write $\mathfrak{X}^{\mathbb{G}_m} = \coprod Y_i$, a union of smooth connected components. For each i, let

$$C_i = \{p \in \mathfrak{X} \mid \lim_{t\to\infty} t\cdot p \in Y_i\};$$

these subsets are the Morse-theoretic attracting loci for the elliptic $\mathbb{G}_m$–action. Note that $\mathfrak{X} = \coprod C_i$ by Definition 1.1(2).

Recall that if $i\colon C \hookrightarrow \mathfrak{X}$ is a smooth coisotropic subvariety, a *coisotropic reduction* of C consists of a smooth symplectic variety (S, ω_S) and a morphism $\pi\colon C \to S$ for which $\omega|_C = \pi^*\omega_S$. We establish a basic structural result for the decomposition $\mathfrak{X} = \coprod C_i$ in Section 2:

Theorem 1.2 (see Theorem 2.1) (1) *Each C_i is a smooth, coisotropic subvariety of $\mathfrak{X}$ and a $\mathbb{G}_m$–equivariant affine bundle over the fixed point set Y_i.*

(2) *There exist symplectic manifolds (S_i, ω_i) with elliptic $\mathbb{G}_m$–action and $\mathbb{G}_m$–equivariant coisotropic reductions $\pi_i\colon C_i \to S_i$.*

Part (1) of the theorem is a symplectic refinement of the Białynicki-Birula stratification [10] arising from a $\mathbb{G}_m$–action. Our proof of assertion (2) relies on formal local normal forms for symplectic varieties in the neighborhood of a coisotropic subvariety, which we develop in Section 2.

We provide a refined description of the symplectic quotients S_i and corresponding $\mathbb{G}_m$–equivariant affine fibrations $S_i \to Y_i$ of Theorem 1.2. We need two definitions. First, let Y be a smooth connected variety. A *symplectic fibration* over Y is a tuple $(E, \eta, \{-,-\})$, where $\eta\colon E \to Y$ is an affine bundle and $\{-,-\}$ an $\mathcal{O}_Y$–linear Poisson bracket on E such that the restriction of $\{-,-\}$ to each fiber of η is nondegenerate. The symplectic fibration is said to be *elliptic* if $\mathbb{G}_m$ acts on E such that $\{-,-\}$ is homogeneous of negative weight, $Y = E^{\mathbb{G}_m}$ and all weights of $\mathbb{G}_m$ on the fibers of η are negative.

Second, note that T^*Y is naturally a group scheme over Y. Suppose $p\colon B \to Y$ is a smooth variety over Y equipped with a symplectic form ω_B. Suppose that B is equipped with an action $a\colon T^*Y \times_Y B \to B$ of the group scheme T^*Y over Y. We say $B \to Y$ is *symplectically automorphic* if, for any 1–form θ on Y, we have $a(\theta, -)^*\omega_B = \omega_B + p^* d\theta$. In the special case that B is a T^*Y–torsor, B is thus a *twisted cotangent bundle* in the sense of Beĭlinson and Bernstein [3].

Theorem 1.3 (see Theorem 2.21) *Keep the notation of Theorem 1.2(2). Then, for each i:*

(1) *The fibration $S_i \to Y_i$ comes equipped with a free T^*Y_i–action making S_i symplectically automorphic over Y_i.*

(2) *The quotient $E_i := S_i / T^*Y_i$ inherits a Poisson structure, making $E_i \to Y_i$ into an elliptic symplectic fibration.*

Locally in the Zariski topology on Y_i, we have $S_i \simeq T^*Y_i \times_{Y_i} E_i$ as smooth varieties with $\mathbb{G}_m$–actions.

1.2 Categorical cell decomposition

We next turn to deformation-quantizations. Suppose that $\mathfrak{X}$ is a smooth symplectic variety with elliptic $\mathbb{G}_m$–action. Let $\mathcal{A}$ be a $\mathbb{G}_m$–equivariant deformation-quantization of $\mathcal{O}_{\mathfrak{X}}$; this is a $\mathbb{G}_m$–equivariant sheaf of flat $\mathbb{C}[\![\hbar]\!]$–algebras, where $\mathbb{G}_m$ acts with weight l on $\hbar$, for which $\mathcal{A}/\hbar\mathcal{A}$ is isomorphic, as a sheaf of $\mathbb{G}_m$–equivariant Poisson algebras, to $\mathcal{O}_{\mathfrak{X}}$; see Section 3.2 for more details. Let $\mathcal{W} = \mathcal{A}[\hbar^{-1}]$. There is a natural analogue for $\mathcal{W}$ of coherent sheaves on $\mathfrak{X}$, the category of $\mathbb{G}_m$–equivariant *good* $\mathcal{W}$–modules, denoted by $\mathcal{W}$–good; see Section 3.3 for details.

Definition 1.4 The category of *quasicoherent $\mathcal{W}$–modules* is

$$\mathsf{Qcoh}(\mathcal{W}) := \mathrm{Ind}(\mathcal{W}\text{–good}).$$

For each subcollection of the coisotropic attracting loci C_i of Section 1.1 whose union C_K, with $K \subset \{1, \dots, k\}$, is a closed subset of $\mathfrak{X}$, we let $\mathsf{Qcoh}(\mathcal{W})_K$ denote the full subcategory of $\mathsf{Qcoh}(\mathcal{W})$ whose objects are supported on C_K. By Lemma 2.3, the loci C_i are naturally partially ordered. Refining to a total order, for each i there are closed subsets $C_{K_{\geq i}} = \bigcup_{j \geq i} C_j$ and $C_{K_{>i}} = \bigcup_{j > i} C_j$.

Theorem 1.5 (see Theorem 4.28 and Corollary 5.2)

1. *The category $\mathsf{Qcoh}(\mathcal{W})$ is filtered by localizing subcategories $\mathsf{Qcoh}(\mathcal{W})_{K_{\geq i}}$.*
2. *Each subquotient $\mathsf{Qcoh}(\mathcal{W})_{K_{\geq i}}/\mathsf{Qcoh}(\mathcal{W})_{K_{>i}}$ is equivalent to the category of quasicoherent modules over a deformation-quantization of the symplectic quotient S_i.*

Mirroring the structure of S_i in Theorem 1.3, the category of $\mathbb{G}_m$–equivariant quasicoherent modules over a deformation-quantization of S_i is equivalent to the category of modules for a specific type of algebra. We describe this relationship explicitly in Sections 4.10 and 4.11, and in Theorem 1.8 below. In particular, in the special case when Y_i is an isolated fixed point, we obtain:

Corollary 1.6 *Suppose the fixed point set $\mathfrak{X}^{\mathbb{G}_m}$ is finite. Then each subquotient*

$$\mathsf{Qcoh}(\mathcal{W})_{K_{\geq i}}/\mathsf{Qcoh}(\mathcal{W})_{K_{>i}}$$

is equivalent to the category of modules over the Weyl algebra $\mathcal{D}(\mathbb{A}^{t_i})$ for $t_i = \frac{1}{2}\dim S_i$.

The Weyl algebra $\mathcal{D}(\mathbb{A}^{t_i})$ is a quantization of the algebra of functions on an "algebraic cell" $\mathbb{A}^{2t_i}$. Moreover, this category, although complicated, looks contractible from the point of view of certain fundamental invariants, for example algebraic K–theory and Hochschild/cyclic homology. Thus, we view the category $\mathcal{D}(\mathbb{A}^{t_i})$–mod as a "categorical algebraic cell", parallel to the way that the category $\mathrm{Vect}_{\mathbb{C}}$ of finite-dimensional complex vector spaces is a categorical analogue of a topological cell. In particular, in the case when the fixed point set $\mathfrak{X}^{\mathbb{G}_m}$ is finite, we interpret the filtration of $\mathsf{Qcoh}(\mathcal{W})$ provided by Theorem 1.5 as providing a *categorical cell decomposition* of $\mathsf{Qcoh}(\mathcal{W})$. Building the category $\mathsf{Qcoh}(\mathcal{W})$ from algebraic cells is thus a "bulk analogue" of the process of building a quasihereditary category from "categorical topological cells", ie copies of $\mathrm{Vect}_{\mathbb{C}}$. In particular, to the extent that categories of the form $\mathsf{Qcoh}(\mathcal{W})$ undergird many representation-theoretic settings of intense recent interest, our categorical cell decompositions are a basic structural feature of the "big" geometric categories that arise in representation theory.

We expect Theorem 1.5 to have many consequences for $\mathsf{Qcoh}(\mathcal{W})$ and related algebraic categories from representation theory. One such application will appear in Bellamy, Dodd, McGerty and Nevins [4]: a symplectic analogue of the Koszul duality, sometimes called "$\mathcal{D}$–Ω–duality", between $\mathcal{D}$–modules on a smooth variety X and dg modules over the de Rham complex Ω_X of X; see Kapranov [38]. More precisely, the Koszul duality of [38] generalizes to arbitrary coherent $\mathcal{D}$–modules the Riemann–Hilbert correspondence between regular holonomic $\mathcal{D}$–modules and their associated de Rham complexes, which are constructible complexes on X. Since Ω_X sheafifies over X, embedded as the zero section of T^*X, one can view this correspondence as a categorical means of sheafifying the category of $\mathcal{D}_X$–modules over X. Such a sheafification is tautologous in the $\mathcal{D}$–module setting, but becomes less so in a general symplectic setting. Namely, in [4], starting from a *bionic symplectic variety* — a symplectic variety $\mathfrak{X}$ with *both* an elliptic $\mathbb{G}_m$–action and a commuting Hamiltonian $\mathbb{G}_m$–action defining a good *Lagrangian skeleton* Λ of $\mathfrak{X}$ — we will use Theorem 1.5 to establish a Koszul duality between $\mathcal{W}_{\mathfrak{X}}$–modules and dg modules over an analogue of Ω_X that "lives on" Λ. As a result, the bounded derived category $D^b(\mathcal{W}\text{–}\mathsf{good})$ naturally sheafifies over Λ.

We explain in Section 1.3 the main technical result that makes Theorem 1.5 possible, namely, an analogue of Kashiwara's equivalence. In Section 1.5 we derive several applications to the category $\mathsf{Qcoh}(\mathcal{W})$. In Section 1.6 we describe parallels to the structure of Fukaya categories in more detail, and indicate some future work in that direction.

1.3 Analogue of Kashiwara's equivalence for deformation-quantizations

Theorem 1.5 is a consequence of a symplectic version of a fundamental phenomenon of $\mathcal{D}$–module theory, classically encoded in the following topological invariance property:

Kashiwara's equivalence *Suppose that $C \subset X$ is a smooth closed subset of a smooth variety X. Then the category $\mathsf{Qcoh}(\mathcal{D}_X)_C$ of quasicoherent $\mathcal{D}_X$–modules supported set-theoretically on C is equivalent to the category $\mathsf{Qcoh}(\mathcal{D}_C)$ of quasicoherent $\mathcal{D}_C$–modules.*

Assume now that $(\mathfrak{X},\omega)$ is a smooth, connected symplectic variety with elliptic $\mathbb{G}_m$–action. Let Y be a connected component of the fixed point set of $\mathfrak{X}$ under $\mathbb{G}_m$–action and C the set of points in $\mathfrak{X}$ limiting to Y under $\mathbb{G}_m$. Assume that C is closed in $\mathfrak{X}$. Let $C \to S$ denote the symplectic quotient whose existence is assured by Theorem 1.2(2). The subcategory of $\mathcal{W}$–good consisting of objects supported on C is $\mathcal{W}$–good_C.

To the algebra $\mathcal{W}$ and the coisotropic subset C we associate a sheaf of algebras $\mathcal{W}_S$ on the symplectic quotient S. We define a natural *coisotropic reduction functor* $\mathbb{H}\colon \mathcal{W}_{\mathfrak{X}}$–$\mathsf{good}_C \to \mathcal{W}_S$–$\mathsf{good}$. The following provides an analogue of Kashiwara's equivalence for $\mathcal{W}$–modules:

Theorem 1.7 (see Theorem 4.28 and Corollary 4.30)

1. *The functor* $\mathbb{H}\colon \mathcal{W}_{\mathfrak{X}}$–$\mathsf{good}_C \to \mathcal{W}_S$–$\mathsf{good}$ *is an exact equivalence of categories. It induces an exact equivalence* $\mathbb{H}\colon \mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}})_C \to \mathsf{Qcoh}(\mathcal{W}_S)$.
2. *The functor* $\mathbb{H}$ *preserves both holonomicity and regular holonomicity.*

We also analyze the structure of $\mathcal{W}_S$ more fully. More precisely, we prove that, on Y, there exists a sheaf of "generalized twisted differential operators" $\mathcal{D}_S$, a filtered $\mathcal{O}_Y$–algebra, whose completed Rees algebra sheafifies over S and gives exactly $\mathcal{W}_S$. We obtain:

Theorem 1.8 (see Theorem 4.28 and Proposition 4.33)

1. *The functor* $\mathbb{H}$ *of coisotropic reduction, followed by taking* $\mathbb{G}_m$*–finite vectors, defines an equivalence*
$$\mathbb{H}\colon \mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}_C \xrightarrow{\sim} \mathsf{coh}(\mathcal{D}_S).$$
Passing to ind-categories defines an equivalence $\mathbb{H}\colon \mathsf{Qcoh}(\mathcal{W})_C \to \mathsf{Qcoh}(\mathcal{D}_S)$.
2. *In particular, if* Y *consists of a single isolated* $\mathbb{G}_m$*–fixed point,* $\mathbb{H}$ *defines an equivalence*
$$\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}})_C \cong \mathsf{Qcoh}(\mathcal{D}(\mathbb{A}^{t_i}))$$
for some t_i.

We emphasize that the elliptic $\mathbb{G}_m$–action is both essential to relate $\mathcal{W}$–modules to representation theory and an intrinsic part of the geometry behind Theorems 1.7 and 1.8: it is not simply a technical convenience to make the proofs work. As an illustration, consider $\mathcal{D}$–modules on $\mathbb{A}^1$ with singular support in the zero section, ie local systems. The zero section is a conical coisotropic subvariety in $T^*\mathbb{A}^1$ with symplectic quotient a point. However, the category of *all* algebraic local systems on $\mathbb{A}^1$ is *not* equivalent to the category $\mathrm{Vect}_{\mathbb{C}}$ of finite-dimensional vector spaces, ie coherent $\mathcal{D}$–modules on a point: such a statement is only true once one passes to the subcategory of local systems regular at infinity. One can define a good notion of regularity in the deformation-quantization setting; see D'Agnolo and Polesello [16]; and then a regular $\mathcal{W}$–module supported on a coisotropic subvariety C will be in the essential image of the corresponding functor quasi-inverse to $\mathbb{H}$. In particular, passing to regular objects yields a version of Theorem 1.7 as in [16], but at a cost too high for our intended applications: the subcategory obtained is no longer described in terms of support conditions, and correspondingly one loses control over what the subquotients look like.

Theorem 1.7, on the other hand, imposes a natural geometric condition on the coisotropic subvariety C: it must arise from the (algebrogeometric) Morse theory of the $\mathbb{G}_m$–action. With that condition satisfied, regularity can be replaced by the more natural geometric support condition, thus yielding a precise structural result on $\mathsf{Qcoh}(\mathcal{W})$.

1.4 Abelian and derived categories

Our main results and techniques also establish some basic properties of categories of $\mathcal{W}$–modules that are analogues of familiar assertions for categories of coherent or quasicoherent sheaves.

1.4.1 Abelian categories As one example, let $Z \subset \mathfrak{X}^{\mathbb{G}_m}$ be a closed, connected and smooth subvariety. Let $C = \{x \in \mathfrak{X} \mid \lim_{t\to\infty} t\cdot x \in Z\}$ be the attracting locus for Z; it is a smooth, locally closed subvariety of $\mathfrak{X}$. Assume that C is closed in $\mathfrak{X}$. The complement to C in $\mathfrak{X}$ is denoted by U and we write $j\colon U \hookrightarrow \mathfrak{X}$ for the embedding.

Theorem 1.9 (Theorem 3.27 and Corollary 3.28)

(1) *The inclusion functor* $i_*\colon \mathsf{Qcoh}(\mathcal{W})_C \to \mathsf{Qcoh}(\mathcal{W})$ *realizes* $\mathsf{Qcoh}(\mathcal{W})_C$ *as a localizing subcategory of* $\mathsf{Qcoh}(\mathcal{W})$*. In particular,* i_* *admits a right adjoint* $i^!$ *such that the adjunction* $\mathrm{id} \to i^! \circ i_*$ *is an isomorphism.*

(2) *The restriction functor* $j^*\colon \mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}}) \to \mathsf{Qcoh}(\mathcal{W}_U)$ *induces equivalences*

$$\mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}/\mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}_C \simeq \mathcal{W}_U\text{–}\mathsf{good} \quad \textit{and} \quad \mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}})/\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}})_C \simeq \mathsf{Qcoh}(\mathcal{W}_U).$$

In particular, j^* *admits a right adjoint* $j_*\colon \mathsf{Qcoh}(\mathcal{W}_U) \to \mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}})$ *with* $j^* \circ j_* \simeq \mathrm{id}$.

As corollaries of Theorem 1.9, we immediately get corresponding statements for both the bounded and unbounded derived categories. Unfortunately, it is clear that neither j^* nor i_* can admit *left* adjoints in general (see however Braden, Proudfoot and Webster [13, Theorem 5.20] in a special case). However, we show in the forthcoming paper [4] that the inclusions $C \xhookrightarrow{i} \mathfrak{X} \xhookleftarrow{j} U$ determine a full recollement pattern on the derived category of *holonomic* $\mathcal{W}$–modules.

1.4.2 Derived categories and compact generation The categorical cell decomposition of $\mathsf{Qcoh}(\mathcal{W})$ extends to the derived level: the (unbounded) derived category $D(\mathsf{Qcoh}(\mathcal{W}))$ is filtered by $D_{K_{\geq i}}(\mathsf{Qcoh}(\mathcal{W}))$, the full localizing triangulated subcategories of objects with cohomology supported on the closed subvarieties $C_{K_{\geq i}}$. The associated minimal subquotients are given by

$$D_{K_{\geq i}}(\mathsf{Qcoh}(\mathcal{W}))/D_{K_{>i}}(\mathsf{Qcoh}(\mathcal{W})) \simeq D(\mathsf{Qcoh}(\mathcal{D}_{S_i})).$$

Since $\mathcal{D}_{S_i}$ is a sheaf of $\mathcal{O}_{Y_i}$–algebras of finite homological dimension, standard arguments show that the triangulated category $D(\mathsf{Qcoh}(\mathcal{D}_{S_i}))$ enjoys strong generation properties. Namely, the category is compactly generated and the full subcategory of compact objects is precisely the category of perfect complexes. Our symplectic generalization of Kashiwara's equivalence allows one to inductively show that these properties lift to the categories $D_K(\mathsf{Qcoh}(\mathcal{W}))$. In particular, if $D(\mathsf{Qcoh}(\mathcal{W}))^c$ denotes the full subcategory of compact objects, then taking K to be all of $\{1,\dots,k\}$ we have:

Theorem 1.10 (see Corollaries 5.9 and 5.10) *The derived category $D(\mathsf{Qcoh}(\mathcal{W}))$ is compactly generated and there is an equality*

$$D(\mathsf{Qcoh}(\mathcal{W}))^c = \mathsf{perf}(\mathcal{W}) = D^b(\mathcal{W}\text{–}\mathsf{good})$$

of full, triangulated subcategories of $D(\mathsf{Qcoh}(\mathcal{W}))$.

An analogous compact generation result was shown by Petit [63], though the category of cohomologically complete deformation-quantization modules considered in [63] is (in a precise sense) orthogonal to $\mathsf{Qcoh}(\mathcal{W})$.

1.5 Applications to invariants

Theorem 1.5 yields immediate consequences for the structure of fundamental invariants associated to the category of sheaves over a deformation-quantization of $\mathfrak{X}$. For example:

Corollary 1.11 (see Section 5.3) *Suppose $\mathfrak{X}^{\mathbb{G}_m}$ is finite of cardinality k. Choose a refinement of the partial ordering of coisotropic attracting loci C_i of $\mathfrak{X}$ to a total ordering. Then the group $K_0(\mathsf{perf}(\mathcal{W}))$ comes equipped with a canonical k–step filtration each subquotient of which is isomorphic to $\mathbb{Z}$; in particular, $K_0(\mathsf{perf}(\mathcal{W}))$ is free abelian of rank k.*

In fact, using the fundamental properties of holonomic modules developed in the sequel [4], one can show that there are natural isomorphisms

$$\theta_n\colon K_n(\mathbb{C}) \otimes_{\mathbb{Z}} K_0(\mathsf{perf}(\mathcal{W})) \xrightarrow{\sim} K_n(\mathsf{perf}(\mathcal{W}))$$

for all $n \geq 0$. Similar results hold for the cyclic and Hochschild homology of the dg enhancement $\mathsf{Perf}(\mathcal{W})$.

Corollary 1.12 (see Section 5.2) *Suppose $\mathfrak{X}^{\mathbb{G}_m}$ is finite of cardinality k. Let $H_*(\mathfrak{X})$ denote the Borel–Moore homology of $\mathfrak{X}$, with coefficients in $\mathbb{C}$. There are isomorphisms of graded vector spaces*

$$HH_*(\mathsf{Perf}(\mathcal{W})) \simeq H_{*-\dim \mathfrak{X}}(\mathfrak{X}), \quad HC_*(\mathsf{Perf}(\mathcal{W})) \simeq H_{*-\dim \mathfrak{X}}(\mathfrak{X}) \otimes \mathbb{C}[\epsilon],$$

where ϵ is assumed to have degree two.

In most situations, such as those that appear in representation theory, it is also possible to calculate the Hochschild *co*homology of $\mathsf{Perf}(\mathcal{W})$. Namely, in Proposition 5.19 we show:

Corollary 1.13 (see Section 5.2) *Suppose $\mathfrak{X}^{\mathbb{G}_m}$ is finite of cardinality k. Then*

$$HH^*(\mathsf{Perf}(\mathcal{W})) = H^*(\mathfrak{X}, \mathbb{C}).$$

Via derived localization — see McGerty and Nevins [53] — the above results allows one to easily calculate the additive invariants HH_*, HC_* and HH^* of many quantizations of singular (affine) symplectic varieties that occur naturally in representation theory. See Section 5.4 for a discussion and applications. For example, let Γ be a cyclic group and $\mathfrak{S}_n \wr \Gamma$ the wreath product group that acts as a symplectic reflection group on $\mathbb{C}^{2n}$. The corresponding symplectic reflection algebra at $t = 1$ and parameter $\boldsymbol{c}$ is denoted by $\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)$. For the definition of the filtration F in the corollary below, see Example 5.20.

Corollary 1.14 (see Proposition 5.21) *Assume that $\boldsymbol{c}$ is spherical. Then*

$$HH^*(\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)) = HH_{2n-*}(\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)) = \mathrm{gr}_*^F(\mathsf{Z}\mathfrak{S}_n \wr \Gamma),$$

as graded vector spaces.

One can deduce similar results for finite W–algebras associated to nilpotent elements regular in a Levi, and quantizations of slices to Schubert varieties in affine Grassmannians. These examples are explained in more detail at the end of Section 5.3.

1.6 Relation to Fukaya categories of Weinstein manifolds

There are a close conceptual link and, conjecturally, a precise mathematical relationship between the categories $\mathsf{Qcoh}(\mathcal{W})$ that we study and the Fukaya categories of Weinstein manifolds in real symplectic geometry. More precisely, a growing body of important work in real symplectic geometry (by, among others, Abouzaid, Kontsevich, Nadler, Seidel, Soibelman, Tamarkin, Tsygan and Zaslow) establishes fundamental links between structures of microlocal sheaf theory and Fukaya categories. The exposition of Nadler [55] sets the Fukaya theory of Weinstein manifolds squarely in a Morse-theoretic context, by showing how to use integral transforms to realize brane categories as glued from the homotopically simpler categories of branes living on coisotropic cells. Theorem 1.5 provides an exact parallel to the structure described in [55]. One difference worth noting is that our categories include objects with arbitrary coisotropic support, not just Lagrangian support as in standard Fukaya theory: we provide such gluing structure for an algebrogeometric "bulk" category of all coisotropic branes rather just than the "thin" category of Lagrangian branes.

Expert opinion supports a direct relationship between the category $\mathsf{Qcoh}(\mathcal{W})$ (or more precisely the holonomic subcategory) and the structure of Fukaya categories described in [55]. Namely, many examples of hyperkähler manifolds with S^1–action fit our paradigm and have affine hyperkähler rotations possessing the requirements described in [55]. In such cases, it is natural to try to prove that the category of [55] is equivalent to $\mathsf{Qcoh}(\mathcal{W})$ for a particular choice of $\mathcal{W}$ by first proving cell-by-cell equivalences; next, describing a classifying object for categories built from cells as in Theorem 1.5 and [55]; and, finally and most difficult, isolating a collection of properties that distinguish the Fukaya category of [55] in the universal family and matching it to some $\mathsf{Qcoh}(\mathcal{W})$. We intend to return to this problem in future work.

1.7 Relation to other work on deformation-quantization

In recent years there has been much interest in the study of quantizations of certain classes of symplectic algebraic varieties, going back at least as far as the work of Kashiwara and Rouquier [40] on the Hilbert scheme of points in the plane mentioned above. The class of varieties which has attracted the most interest is that of conical symplectic resolutions. These are symplectic varieties Y with a $\mathbb{G}_m$–action such that the affinization map $f\colon Y \to X$ is birational and the resulting $\mathbb{G}_m$–action on X has a single attracting fixed point. Braden, Proudfoot and Webster [13] and Braden, Licata, Proudfoot and Webster [12] give a systematic study of quantizations of these varieties, and study in detail a class of holonomic modules in the spirit of the classical theory of category $\mathcal{O}$ (see also the subsequent work of Losev [51]).

Clearly any such conic symplectic resolution is an elliptic symplectic variety, but the class of elliptic symplectic varieties is strictly larger. For example, if Σ is a smooth complete curve, then $Y = \mathrm{Hilb}^n(T^*\Sigma)$, the Hilbert scheme of points of the cotangent bundle of Σ is naturally a symplectic variety (studied by Nakajima [56], for example) and it possess a natural elliptic $\mathbb{G}_m$–action induced by the scaling action on the fibers of $T^*\Sigma$. However, unless $\Sigma = \mathbb{P}^1$, the symplectic variety Y is not a symplectic resolution. Moreover, in this paper we seek to investigate the structure of the full category of $\mathbb{G}_m$–equivariant modules, rather than focusing on particular classes of holonomic modules. Our forthcoming work [4] was partly inspired by the question of which properties of the category of all (suitably equivariant) modules for a quantization of $\mathfrak{X}$ can be detected by small subcategories such as the geometric incarnations of category $\mathcal{O}$ studied by Braden, Licata, Proudfoot and Webster [12]. That paper however, as in the work of Braden, Proudfoot and Webster [13] mentioned above, requires Y to carry the action of a higher-dimensional torus.

1.8 Outline of the paper

Section 2 describes some of the basic geometric properties of symplectic manifolds equipped with an elliptic $\mathbb{G}_m$–action. The basic properties of modules over deformation-quantization algebras are recalled in Section 3. In Section 4 we describe a version of quantum coisotropic reduction for equivariant DQ algebras and prove a version of Kashiwara's equivalence. This equivalence is used in Section 5 to study the derived category $D(\mathsf{Qcoh}(\mathcal{W}))$ and also calculate the additive invariants of $\mathcal{W}$–good.

1.9 Conventions

Deformation-quantization algebras A, sheaves of DQ algebras $\mathcal{A}$, $\mathcal{W}$–algebras W (global section case) and $\mathcal{W}$ (sheaf case), and their modules and equivariant modules are defined in the body of the paper. For a (sheaf of) DQ algebra(s) A (resp. $\mathcal{A}$) with $\mathbb{G}_m$–action, we always write $(A, \mathbb{G}_m)$–mod for the category of finitely generated equivariant modules (resp. $(\mathcal{A}, \mathbb{G}_m)$–$\mathsf{coh}$ for the category of coherent modules).

For a $\mathcal{W}$–algebra $\mathcal{W}$, we write $\mathcal{W}$–good for the category of good $\mathcal{W}$–modules. If T is a torus acting on $\mathfrak{X}$ and $\mathcal{W}$ is a T–equivariant $\mathcal{W}$–algebra (where T acts on $\hbar$ via a character), we write $(\mathcal{W}, \mathsf{T})$–$\mathsf{good}$ for the category of good T–equivariant $\mathcal{W}$–modules, or, if T is clear from context, just $\mathcal{W}$–good. We write $\mathsf{Qcoh}(\mathcal{W})$ and $\mathsf{Qcoh}(\mathcal{W}, \mathsf{T})$ for the ind-categories of $\mathcal{W}$–good and $(\mathcal{W}, \mathsf{T})$–$\mathsf{good}$, respectively.

Acknowledgments The authors are grateful to David Ben-Zvi, Andrea D'Agnolo, Kobi Kremnitzer, David Nadler and Eric Vasserot for very helpful conversations.

Bellamy was supported by the EPSRC grant EP-H028153. McGerty was supported by a Royal Society research fellowship and also by the EPSRC grant EP/1033343/1. Nevins was supported by NSF grants DMS-0757987 and DMS-1159468 and NSA grant H98230-12-1-0216, and by an All Souls Visiting Fellowship. All four authors were supported by MSRI.

2 The geometry of symplectic varieties with elliptic $\mathbb{G}_m$–action

We assume throughout this section that $(\mathfrak{X}, \omega)$ is a smooth, connected symplectic, quasiprojective variety with elliptic $\mathbb{G}_m$–action. By *symplectic manifold* we mean a smooth quasiprojective variety over $\mathbb{C}$ equipped with an algebraic symplectic form. In this section, we describe some basic geometric consequences of the $\mathbb{G}_m$–action.

2.1 A symplectic Białynicki-Birula decomposition

The connected components of the fixed point set of $\mathfrak{X}$ under the $\mathbb{G}_m$–action will be denoted by $Y_1, \dots, Y_k$. Each Y_i is a smooth, closed subvariety of $\mathfrak{X}$. Recall that

$$C_i = \left\{x \in \mathfrak{X} \mid \lim_{t \to \infty} t \cdot x \in Y_i\right\}.$$

Then it follows that $\mathfrak{X} = \bigsqcup_{i=1}^k C_i$.

Let C a smooth, connected, locally closed coisotropic subvariety of $\mathfrak{X}$. A *coisotropic reduction* of C is a smooth symplectic variety (S, ω') together with a smooth morphism $\pi\colon C \to S$ such that $\omega|_C = \pi^*\omega'$. A classical example of a coisotropic reduction is given by $\mathfrak{X} = T^*X$, $Y \subset X$ a smooth, closed subvariety, $C = (T^*X)|_Y$ and $\pi\colon (T^*X)|_Y \to T^*Y$ the natural map.

Theorem 2.1 *Suppose $(\mathfrak{X}, \omega)$ is a symplectic manifold with elliptic $\mathbb{G}_m$–action. Then:*

1. *Each C_i is a smooth, coisotropic subvariety of $\mathfrak{X}$ and an $\mathbb{G}_m$–equivariant affine bundle over the fixed point set Y_i.*
2. *There exist symplectic manifolds (S_i, ω_i) with elliptic $\mathbb{G}_m$–action and $\mathbb{G}_m$–equivariant coisotropic reductions $\pi_i\colon C_i \to S_i$.*

The proof of the first statement of Theorem 2.1 is given in Section 2.2. The proof of the second statement of Theorem 2.1 is given in Section 2.3 after some preparatory work.

2.2 Proof of Theorem 2.1(1)

The proof of Theorem 2.1(1) is essentially a direct consequence of the Białynicki-Birula decomposition together with some elementary weight arguments. However, we provide details for completeness. The fact that each C_i is an $\mathbb{G}_m$–equivariant affine bundle over Y_i follows directly from [10, Theorem 4.1].

With regard to $\mathbb{G}_m$–representations, the following conventions will be used throughout the paper. If V is a graded vector space then V_i denotes the subspace of degree i. Let

V be a prorational $\mathbb{G}_m$–module; that is, V is the limit of its $\mathbb{G}_m$–equivariant rational quotients. Then $V^{\mathrm{rat}} = \bigoplus_{i\in\mathbb{Z}} V_i$, the subspace of $\mathbb{G}_m$–finite vectors, is a rational $\mathbb{G}_m$–module.

To show that each C_i is coisotropic, we first show that T_pC_i is a coisotropic subspace of $T_p\mathfrak{X}$ at each point $p \in Y_i$. Indeed we claim that

$$(T_pC_i)^{\perp} = \mathrm{rad}(\omega|_{C_i})_p = \bigoplus_{j<-l} (T_pC_i)_j. \tag{2-1}$$

To see this, let $V = T_p\mathfrak{X}$ and $W = T_pC_i$. Then V and W have weight space decompositions $V = \bigoplus_j V_j$ and $W = \bigoplus_j W_j$, where $W_j = V_j$ for $j \le 0$ and $W_j = 0$ for $j > 0$. If $v \in V_a$ and $w \in V_b$ then

$$t^l\omega(v,w) = (t\cdot\omega)(v,w) = \omega(t^{-1}\cdot v, t^{-1}\cdot w) = t^{-a-b}\omega(v,w).$$

This implies that $\omega(v,w) = 0$ if $l \ne -a-b$ and ω restricts to a nondegenerate pairing $V_j \times V_{-l-j} \to \mathbb{C}$. Therefore, if $v \in V_j \cap W^{\perp}$ then the equality $\omega(v,w) = 0$ for all $w \in W$ implies that $V_{-j-l} \cap W = 0$, ie $-j-l > 0$ and hence $j < -l$. Hence $V_j \subset W$. This implies that $W^{\perp} = W_{<-l}$ and (2-1) follows. Then the following lemma completes the proof of Theorem 2.1(1):

Lemma 2.2 *Let C be an attracting set in $\mathfrak{X}$ and $Y \subset C$ the set of fixed points. Then C is coisotropic if and only if $(T_pC)^{\perp} \subset T_pC$ for all $p \in Y$.*

Proof Fix $p \in Y$. There exists a $\mathbb{G}_m$–stable affine open neighborhood U of p on which the tangent bundle equivariantly trivializes, ie $T\mathfrak{X}|_U \simeq U \times T_p\mathfrak{X}$. To see this, let U_0 be a $\mathbb{G}_m$–stable affine open neighborhood of p and $\mathfrak{m} \lhd \mathbb{C}[U_0]$ the maximal ideal defining p. Choose a homogeneous lift $x_1,\dots,x_{2n}$ of a basis of $\mathfrak{m}/\mathfrak{m}^2$ in $\mathbb{C}[U_0]$. Then there exists some affine open $\mathbb{G}_m$–stable subset $U \subset U_0$ such that $\{dx_1,\dots,dx_{2n}\}$ is a basis of Ω^1_U as a $\mathbb{C}[U]$–module. Shrinking U if necessary, we may assume that $U^{\mathbb{G}_m} = Y \cap U$. Under the corresponding identification $T_x\mathfrak{X} \xrightarrow{\sim} T_p\mathfrak{X}$ of tangent spaces, T_xC is mapped to T_pC for all $x \in U \cap C$. Write $\omega = \sum_{i<j} f_{i,j}\,dx_i \wedge dx_j$, thought of as a family of skew-symmetric bilinear forms on the fixed vector space $T_p\mathfrak{X}$.

If ∂_i is dual to dx_i, then as shown above T_pC is spanned by all ∂_i of degree ≤ 0 and $(T_pC)^{\perp_p}$ is spanned by all ∂_i of degree less than $-l$. By definition, $U \cap C$ is the set of all points in U vanishing on all $f \in \mathbb{C}[U]$ of negative degree. Let $\partial_i \in (T_pC)^{\perp_p}$ and $\partial_j \in T_pC$. Then $\deg dx_i > l$ and $\deg x_j \ge 0$. Therefore $\deg f_{i,j} < 0$. This implies that $f_{i,j}(x) = 0$ for all $x \in U \cap C$ and hence $\omega_x(\partial_i,\partial_j) = 0$, so that $(T_pC)^{\perp_p} \subseteq (T_pC)^{\perp_x}$. Since $\dim(T_pC)^{\perp_p} = \dim\mathfrak{X} - \dim(T_pC) = \dim(T_pC)^{\perp_x}$ we must have $(T_pC)^{\perp_x} = (T_pC)^{\perp_p}$ and so the lemma follows. □

Define a relation on the coisotropic attracting loci C_i by $C_i \geq C_j$ if $\overline{C}_j \cap C_i \neq \varnothing$.

Lemma 2.3 *The relation $C_i \geq C_j$ is antisymmetric; in particular, it defines a partial order on the Białynicki-Birula strata.*

Proof Equivariantly compactify $\mathfrak{X}$ to a smooth projective $\mathbb{G}_m$–variety $\overline{\mathfrak{X}}$; see [14, Theorem 5.1.25]. Since we assume that every point in $\mathfrak{X}$ has a $\mathbb{G}_m$–limit in $\mathfrak{X}$, a point in $x \in \overline{\mathfrak{X}}$ has its $\mathbb{G}_m$–limit in the closed set $\partial\overline{\mathfrak{X}} = \overline{\mathfrak{X}} \smallsetminus \mathfrak{X}$ if and only if $x \in \partial\overline{\mathfrak{X}}$; in particular, $\partial\overline{\mathfrak{X}}$ is a union of BB strata of $\overline{\mathfrak{X}}$. The relation on BB strata defined above on $\mathfrak{X}$ is a subset of the relation defined using the BB stratification of $\overline{\mathfrak{X}}$ (some stratum closures may intersect at the boundary in $\overline{\partial}$ but not in $\mathfrak{X}$ itself). Since the BB stratification of $\overline{\mathfrak{X}}$ is *filterable* by [11], the conclusion follows. □

2.3 Proof of Theorem 2.1(2): from global to local

We fix C to be one of the coisotropic strata (one of the C_i) in $\mathfrak{X}$ and let $Y = C^{\mathbb{G}_m}$. Note that if $C = C_i$ then Lemma 2.3 shows that $C_{\geq i} = \bigcup_{j \geq i} C_j$ is open in $\mathfrak{X}$ and C_i is closed in $C_{\geq i}$. Since an open union of coisotropic cells in $\mathfrak{X}$ inherits the structure of an elliptic symplectic variety, we may thus assume without loss of generality that C is closed in $\mathfrak{X}$. We first describe a canonical global construction of a morphism $\pi\colon C \to S$. To show that this construction yields a coisotropic reduction is a local computation, which we carry out in the next section, giving a local normal form for the symplectic form on a formal neighborhood of C.

The global construction can be described as follows: Let $\rho\colon C \to Y$ be the projection map and $\mathcal{I}$ denote the sheaf of ideals in $\mathcal{O}_{\mathfrak{X}}$ defining C. The quotient $\mathcal{I}/\mathcal{I}^2$ is a locally free $\mathcal{O}_C$–module. By Lemma 2.4 below, $\mathcal{I}$ is involutive. Therefore the Lie algebroid $\mathcal{L} := \rho_\bullet(\mathcal{I}/\mathcal{I}^2)$ acts on $\rho_\bullet\mathcal{O}_C$ via Hamiltonian vector fields and we can consider the sheaf $\mathfrak{H} := H^0(\mathcal{L}, \rho_\bullet\mathcal{O}_C)$ of sections of $\rho_\bullet\mathcal{O}_C$ that are invariant under these Hamiltonian vector fields. The fact that the Poisson bracket has weight $-l$ and $\mathcal{O}_Y$ is concentrated in degree zero implies that $\mathcal{L}$ is actually an $\mathcal{O}_Y$–Lie algebra and $\mathfrak{H}$ a sheaf of $\mathcal{O}_Y$–algebras.

The embedding of $\mathfrak{H}$ into $\rho_\bullet\mathcal{O}_C$ defines a dominant map $\pi\colon C \to S := \operatorname{Spec}_Y \mathfrak{H}$ of schemes over Y. The final claim of Theorem 2.1 is that the map π is a coisotropic reduction and, in particular, S is a symplectic manifold. Since both C and S are affine over Y and the statement of the claim is local on S, it suffices to assume that we are in the local situation of Section 2.4 below. Then the claim is a consequence of Theorem 2.6; see the end of Section 2.4.

2.4 Proof of Theorem 2.1(2): affine local case

In this section we prove that the construction described in Section 2.3 does indeed give a coisotropic reduction. The idea is to show that, locally, there is a different construction of π using the $\mathbb{G}_m$–action on $\mathfrak{X}$. This construction clearly gives a coisotropic reduction. Unfortunately, this construction doesn't obviously lift to a global construction. Therefore the main thrust of this section is to show that this second construction agrees (locally) with the construction given in Section 2.3.

It follows from [10, Theorem 2.5] that, for each point $y \in Y$, there is some affine open neighborhood of y in Y such that the affine bundle $\rho\colon C \to Y$ trivializes equivariantly. Replacing Y by such an affine open subset, we assume that ρ is equivariantly trivial. Moreover, we may assume that $\mathfrak{X}$ is also affine (still equipped with a $\mathbb{G}_m$–action). Set $R = \mathbb{C}[\mathfrak{X}]$, a regular affine $\mathbb{C}$–algebra with nondegenerate Poisson bracket $\{-,-\}$ and $\mathbb{G}_m$–action such that $\{-,-\}$ has weight $-l$. Let I be the ideal in R defining C.

Let $\mathfrak{l} = I/I^2$. The Poisson bracket on R makes $\mathfrak{l}$ into a Lie algebra which acts on R/I. Let

$$H := H^0(\mathfrak{l}, R/I) = (R/I)^{\{I,-\}}$$

denote the "coisotropic reduction" of R with respect to I. The bracket $\{-,-\}$ descends to a bracket on H and R/I is a Poisson module for H. We set $S = \operatorname{Spec} H$ and let π be the dominant map coming from the inclusion $H \hookrightarrow R/I$. This is the local version of the construction described in Section 2.3.

Lemma 2.4 *The ideal I equals the ideal of R generated by all homogeneous elements of negative degree and is involutive, ie $\{I, I\} \subset I$.*

Proof Let J be the ideal of R generated by all homogeneous elements of negative degree. The fact that $\{-,-\}$ has degree $-l$ implies that J is involutive. Moreover, the set of zeros of I and J clearly coincide. Therefore the lemma is really asserting that J is a radical ideal. If $\mathfrak{X}$ were not smooth then this need not be true. Let $D = \operatorname{Spec} R/J$. Since the nonreduced locus is a closed, $\mathbb{G}_m$–stable subscheme of D, it suffices to show that the local ring $\mathcal{O}_{D,y}$ is a domain for all $\mathbb{G}_m$–fixed closed points of D. But in this case, if $\mathfrak{m}$ is the maximal ideal in R defining y, then T_y^*D is precisely the subspace of $\mathfrak{m}/\mathfrak{m}^2$ of nonnegative weights and D is locally cut out by homogeneous lifts of the elements of $\mathfrak{m}/\mathfrak{m}^2$ of negative degree. Since $\mathfrak{X}$ is smooth at y, these elements form a regular sequence and hence $\mathcal{O}_{D,y}$ is reduced. □

Now we give our alternative, local construction of π which we will use to verify the morphism π is a coisotropic reduction. For the remainder of this section, let S denote the affine variety such that $\mathbb{C}[S] \subset \mathbb{C}[C]$ is the subalgebra generated by all homogeneous elements of degree at most l and let π denote the dominant morphism $C \to S$.

Lemma 2.5 *The variety S is symplectic and π is a coisotropic reduction of C.*

Proof Fix some $y \in Y$. First we establish that S is a smooth variety of dimension $\dim(T_y C)_{\geq -l}$. By definition, there is an equivariant trivialization $\phi\colon C \xrightarrow{\sim} Y \times V \times Z$, where $V \subset T_y C$ is the sum of all weight spaces with weight $-l \leq i < 0$ and Z the sum of all weight spaces of weight $< -l$. Since ϕ is equivariant and S defined in terms of weights, $\phi^*(\mathbb{C}[S]) = \mathbb{C}[Y \times V]$ and $\pi' \circ \phi^{-1}$ corresponds to the projection map onto $Y \times V$. Thus, S is smooth of the stated dimension and π is a smooth morphism.

Since C is only coisotropic, the bracket on $\mathfrak{X}$ does not restrict to a bracket on R/I. However, as explained above, it does induce a bracket on H. Since $\mathbb{C}[S]$ is generated by homogeneous elements of degree at most l, and I is generated by elements of negative degree, the algebra $\mathbb{C}[S]$ is contained in H. Again, weight considerations imply that it is Poisson closed. Thus, it inherits a bracket from $\mathfrak{X}$ making π a Poisson morphism.

Finally we need to show that the Poisson structure on $\mathbb{C}[S]$ is nondegenerate. Since $\mathbb{C}[S]$ is positively graded and smooth, it suffices to check the induced pairing

$$\{-,-\}_y\colon T_y^* S \times T_y^* S \to \mathbb{C}$$

is nondegenerate. As noted in the proof of Lemma 2.4, $T_y^* C = (T_y^* \mathfrak{X})/(T_y^* \mathfrak{X})_{<0}$. Since $\mathbb{C}[S]$ is generated by all elements of degree at most l in $\mathbb{C}[C]$, the space $(T_y^* \mathfrak{X})_{\leq l}/(T_y^* \mathfrak{X})_{<0}$ is contained in $T_y^* S$. But these two spaces have the same dimension. Therefore they are equal. Since $\{-,-\}$ has degree $-l$ and is nondegenerate on $T_y^* \mathfrak{X}$, the induced pairing

$$\{-,-\}_y\colon (T_y^* \mathfrak{X})_{\leq l}/(T_y^* \mathfrak{X})_{<0} \times (T_y^* \mathfrak{X})_{\leq l}/(T_y^* \mathfrak{X})_{<0} \to \mathbb{C}$$

is nondegenerate. □

As shown in the proof of Lemma 2.5, we have Poisson subalgebras $\mathbb{C}[S] \subset H$ of R/I and the Poisson structure on $\mathbb{C}[S]$ is nondegenerate. Therefore, to show that H is also a regular affine algebra with nondegenerate Poisson bracket it suffices to show that $\mathbb{C}[S] = H$. In order to do this, we shall need to investigate more closely the local structure of the symplectic form. In particular, we shall need a Darboux–Weinstein-type theorem to describe the behavior of the form near $C \subset \mathfrak{X}$. Since we are working in the Zariski topology, we shall take the formal completion along C.

Shrinking Y further if necessary, we may assume that the normal bundle $\mathsf{N}_{\mathfrak{X}/C}$ to C in $\mathfrak{X}$ is $\mathbb{G}_m$–equivariantly trivial. Thus, we have a $\mathbb{G}_m$–equivariant trivialization $\mathrm{Tot}(\mathsf{N}_{\mathfrak{X}/C}) \simeq C \times Z^*$. This implies that I/I^2 is a free R/I–module.

In these circumstances, [34, paragraphe III.1.1.10 et théorème III.1.2.3] imply that we can choose ($\mathbb{G}_m$–equivariantly) an identification of the formal neighborhood $\mathfrak{C}$

of C in $\mathfrak{X}$ with the formal neighborhood $\widehat{\mathsf{N}}_{\mathfrak{X}/C}$ of the zero section in $\mathsf{N}_{\mathfrak{X}/C}$. By Lemma 2.5, $C \simeq S \times Z$, where $Z \simeq \mathbb{A}^m$. It follows easily that there is a $\mathbb{G}_m$–equivariant isomorphism $\mathfrak{C} \simeq S \times \widehat{T^*Z}$, where $\widehat{T^*Z}$ denotes the completion of T^*Z along the zero section; the corresponding ideal in $\mathbb{C}[\mathfrak{C}] = \widehat{R}$ is therefore the completion of I, denoted by $\widehat{I}$. Then π extends to a projection map $S \times \widehat{T^*Z} \to S$. The inclusion $S \hookrightarrow \mathfrak{C}$ is denoted by ι.

For an arbitrary morphism $f\colon X \to X'$, there is a map $\Omega_f\colon f^*\Omega^\bullet_{X'} \to \Omega^\bullet_X$. Assume X and X' are affine. If ν is a closed k–form on X and $\overline{\nu}$ its image in $f^*\Omega^k_{X'}$, then $\Omega_f(\overline{\nu})$ is closed in Ω^k_X. From $\mathfrak{X}$, the space $\mathfrak{C}$ inherits a symplectic form ω of weight l. Set $\omega_S = \Omega_\pi(\Omega_\iota(\overline{\omega}))$, a closed 2–form on $\mathfrak{C}$. Our local normal form result states:

Theorem 2.6 *Under the identification $\mathfrak{C} \simeq S \times \widehat{T^*Z}$ there is a $\mathbb{G}_m$–equivariant automorphism ϕ of $\mathfrak{C}$, with $\phi(\widehat{I}) = \widehat{I}$, such that*

$$\phi^*\omega = \omega_{\mathrm{can}} := \omega_S + \sum_{i=1}^m dz_i \wedge dw_i \tag{2-2}$$

with respect to some homogeneous bases $\boldsymbol{z} = z_1, \dots, z_m$ and $\boldsymbol{w} = w_1, \dots, w_m$ of $Z^* \subset \mathbb{C}[Z]$ and $Z \subset \mathbb{C}[\![Z^*]\!]$, respectively.

The proof of Theorem 2.6 will be given in Section 2.5.

Remark 2.7 We have $\deg z_i > l$, $\deg w_i < 0$ and $\deg z_i + \deg w_i = l$ for all i.

Theorem 2.6 implies:

Corollary 2.8 *We have $\mathsf{N}_{\mathfrak{X}/C} \simeq (\pi^*Z^*) \otimes \chi^l$, where $\chi = \mathrm{id}\colon \mathbb{G}_m \to \mathbb{G}_m$ is the fundamental character of $\mathbb{G}_m$.*

We can now complete the proof of Theorem 2.1. Recall from Section 2.4 that our goal is to show that $\mathbb{C}[S] = H$ as subalgebras of R/I.

The ring $\widehat{R}$ can be identified with functions on $\mathfrak{C}$. Under this identification, the ideal $\widehat{I} = \widehat{R}I$ is the ideal of functions vanishing on the zero section C and $\widehat{R}/\widehat{I} = R/I$. Since I is involutive and $\widehat{I} = \widehat{R}I$, the ideal $\widehat{I}$ is involutive. By Theorem 2.6, there is an automorphism ϕ^* of $\widehat{R}$ such that $\phi^*(\{f, g\}) = \{\phi^*(f), \phi^*(g)\}_{\mathrm{can}}$, where $\{-,-\}_{\mathrm{can}}$ is the canonical Poisson bracket coming from the symplectic two-form (2-2). By construction, $\phi^*(\widehat{I}) = \widehat{I}$. Therefore, $(\widehat{R}/\widehat{I})^{\{\widehat{I},-\}} = (\widehat{R}/\widehat{I})^{\{\widehat{I},-\}_{\mathrm{can}}}$. Since

$$\widehat{R} = (R/I)[\![w_1, \dots, w_m]\!] = (\mathbb{C}[S] \otimes \mathbb{C}[z_1, \dots, z_m])[\![w_1, \dots, w_m]\!]$$

and $\widehat{I}$ is generated by $w_1, \dots, w_m$, the embedding $\mathbb{C}[S] \hookrightarrow \widehat{R}$ induces an isomorphism $\mathbb{C}[S] \simeq (\widehat{R}/\widehat{I})^{\{\widehat{I},-\}}$. Finally, the equality $\widehat{I} = \widehat{R}I$ implies that $(\widehat{R}/\widehat{I})^{\{\widehat{I},-\}} = (R/I)^{\{I,-\}} = H$.

The following observation will be useful later:

Lemma 2.9 *The projection $\rho\colon C \to Y$ factors through $\pi\colon C \to S$.*

Proof The algebra R/I is $\mathbb{N}$–graded such that the quotient by the ideal generated by all elements of strictly positive degree equals $\mathbb{C}[Y]$; the proof of this last claim is identical to the proof of Lemma 2.4, using the fact that C is smooth. On the other hand we can also identify $\mathbb{C}[Y]$ with the degree zero part of R/I. Therefore, it suffices to show that all sections of R/I of degree zero lie in T. To see this, notice that the Poisson bracket on R has degree $-l$. Therefore, given $f \in (R/I)_0$ and $g = \sum_i g_i h_i \in I$, where $\deg h_i < 0$ for all i, the element $\{f, g\} = \sum_i g_i\{f, h_i\} + h_i\{f, g_i\}$ lies in I because $\deg\{f, h_i\} < 0$. □

2.5 The proof of Theorem 2.6

The crucial tool in the proof of Theorem 2.6 is an algebraic version of the Darboux–Weinstein theorem which is due to Knop [45, Theorem 5.1]. In our setup, Knop's theorem can be stated as:

Theorem 2.10 *Let $\{-,-\}_{\omega}$ and $\{-,-\}_{\omega_{\mathrm{can}}}$ denote the Poisson brackets on $\widehat{R}$ associated to the symplectic forms ω and ω_{can}, respectively. Denote their difference by $\{-,-\}_{-}$. If $\{\widehat{R},\widehat{R}\}_{-} \subset \widehat{I}$, then there exists an automorphism ϕ of $\mathfrak{C}$ as described in Theorem 2.6.*

Knop's proof of Theorem 2.10 is based on the proof by Guillemin and Sternberg of the equivariant Darboux–Weinstein theorem [29], except that Knop works in the formal algebraic setting.

In our case, it is not necessarily true that $\{\widehat{R},\widehat{R}\}_{-} \subset \widehat{I}$. Instead, we construct an equivariant automorphism ψ of $\widehat{R}$, such that $\psi(\widehat{I}) = \widehat{I}$ and the difference of $\{-,-\}_{\psi^*(\omega)}$ and $\{-,-\}_{\omega_{\mathrm{can}}}$ has the desired properties. In fact we prove the following:

Proposition 2.11 *There exist homogeneous elements $w_1, \dots, w_m$ and $z_1, \dots, z_m$ in $\widehat{R}$ and a graded subalgebra T of $\widehat{R}$ such that:*

1. *$T \cap \widehat{I} = 0$ and the map $\widehat{R} \to R/I$ induces a graded algebra isomorphism $T \to \mathbb{C}[S]$.*
2. *The elements $w_1, \dots, w_m$ generate $\widehat{I}$.*
3. *There is a $\mathbb{G}_m$–equivariant isomorphism $R/I \simeq \mathbb{C}[S][z_1, \dots, z_m]$*
4. *With respect to the Poisson structure $\{-,-\} = \{-,-\}_{\omega}$, we have $\{w_i, z_j\} = \delta_{ij} \bmod \widehat{I}$, $\{z_i, T\}$, $\{w_i, T\} \subset \widehat{I}$ and $\{z_i, z_j\} \in \widehat{I}$ for all i and j.*

Since the choice of such elements $w_1, \dots, w_m$, $z_1, \dots, z_m$ clearly yield a $\mathbb{G}_m$–equivariant automorphism of $\widehat{R}$ (which fixes $\widehat{I}$ and S), one sees directly that this proposition implies the existence of ψ as described above.

So we turn to the proof of Proposition 2.11. As in the proof of Lemma 2.9, we make the identification $\mathbb{C}[Y] = (R/I)_0$. We have $S = Y \times V$ and $C = Y \times V \times Z$, where V and Z are as in Lemma 2.5. Let V_l^* be the l–weight subspace of V^*.

Lemma 2.12 *The Poisson bracket on $\mathbb{C}[S]$ defines an isomorphism of $\mathbb{C}[Y]$–modules*

$$\mathbb{C}[Y] \otimes V_l^* \xrightarrow{\sim} \mathrm{Der}(\mathbb{C}[Y]).$$

Proof That the map is well-defined and $\mathbb{C}[Y]$–linear follows from degree considerations. Lemma 2.5 implies that the Poisson bracket on $\mathbb{C}[S]$ is nondegenerate. This nondegeneracy implies that the above map is surjective. Since it is a surjective map between two projective $\mathbb{C}[Y]$–modules of the same rank, it is an isomorphism. □

Choose an arbitrary point $y \in Y$. Our strategy will be to complete at y and use the Darboux theorem for a symplectic formal disc to control the behavior of the Poisson bracket in some neighborhood of y. For the convenience of the reader we recall the statement of the Darboux theorem in the presence of an elliptic $\mathbb{G}_m$–action:

Theorem 2.13 *Let $\widehat{R}_y$ denote the completion of R at y. Then, in $\widehat{R}_y$, there is a regular sequence $\{u_1, \dots, u_n, v_1, \dots, v_n\}$ of homogenous elements such that $\{u_i, v_j\} = \delta_{ij}$ and $\{u_i, u_j\} = 0 = \{v_i, v_j\}$ for all i and j.*

The proof of the equivariant Darboux theorem is a slight modification of standard arguments and we shall omit it. Since $\mathbb{C}[Y] \subset H$, R/I is a Poisson $\mathbb{C}[Y]$–module, where $\mathbb{C}[Y]$ is equipped with the trivial Poisson bracket. Let $K \subset R/I$ denote the set of all elements k such that $\{\mathbb{C}[Y], k\} = 0$; it is a graded subalgebra of R/I.

Lemma 2.14 *Multiplication defines an isomorphism $\mathbb{C}[V_{-l}] \otimes K \xrightarrow{\sim} R/I$.*

Proof By Krull's intersection theorem, we may consider R/I as a subalgebra of $\widehat{R}_y/I\widehat{R}_y$, and hence of $(\widehat{R}_y/I\widehat{R}_y)^{\mathrm{rat}}$ too. The ring $\widehat{R}_y/I\widehat{R}_y$ is a Poisson module over the ring $\widehat{\mathbb{C}}[Y]_y$, where $\widehat{\mathbb{C}}[Y]_y$ is the completion of $\mathbb{C}[Y]$ at the point y. If $\widehat{K} \subset \widehat{R}_y/I\widehat{R}_y$ and $K^{\mathrm{rat}} \subset (\widehat{R}_y/I\widehat{R}_y)^{\mathrm{rat}}$ are defined analogously to K, then it suffices to show that $\mathbb{C}[V_{-l}] \otimes K^{\mathrm{rat}} \xrightarrow{\sim} (\widehat{R}_y/I\widehat{R}_y)^{\mathrm{rat}}$.

Reordering if necessary, we may suppose that $\{v_1, \dots, v_m\}$ are the elements of Theorem 2.13 that have negative degree. They generate the ideal $I\widehat{R}_y$. There exist

$\alpha_1,\dots,\alpha_r, \beta_1,\dots,\beta_s$ such that $\{\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s\} \subset \{u_1,\dots,u_n,v_{m+1},\dots,v_n\}$, $\mathbb{C}[\![\alpha_1,\dots,\alpha_r]\!] = \mathbb{C}[\![V_{-l}]\!]$ and $\mathbb{C}[\![\beta_1,\dots,\beta_s]\!] = \mathbb{C}[\![V_{>-l}\times Z]\!]$. In this case, it follows from Theorem 2.13 that

$$\widehat{R}_y/I\widehat{R}_y = \widehat{\mathbb{C}}[Y]_y \widehat{\otimes} \mathbb{C}[\![\alpha_1,\dots,\alpha_r,\beta_1,\dots,\beta_s]\!] \quad \text{with} \quad \widehat{K} = \widehat{\mathbb{C}}[Y]_y \widehat{\otimes} \mathbb{C}[\![\beta_1,\dots,\beta_s]\!].$$

Hence, since $\deg\alpha_i$ and $\deg\beta_j > 0$ for all i and j,

$$(\widehat{R}_y/I\widehat{R}_y)^{\mathrm{rat}} = \mathbb{C}[V_{-l}] \otimes \widehat{\mathbb{C}}[Y]_y \otimes \mathbb{C}[\beta_1,\dots,\beta_s]. \tag{2-3}$$

Lemma 2.12 implies that we may fix a basis $x_1,\dots,x_r$ of $V_l^* \subset \mathbb{C}[V_{-l}]$ and regular sequence $y_1,\dots,y_r$ in $\mathbb{C}[Y]$ such that $dy_1,\dots,dy_r$ are a basis of Ω^1_Y and $\{x_i,y_j\} = \delta_{ij}$ for all i and j. This, together with the identification (2-3), implies that $K^{\mathrm{rat}} = \widehat{\mathbb{C}}[Y]_y \otimes \mathbb{C}[\beta_1,\dots,\beta_s]$ and $\mathbb{C}[V_{-l}] \otimes K^{\mathrm{rat}} \xrightarrow{\sim} (\widehat{R}_y/I\widehat{R}_y)^{\mathrm{rat}}$. □

Now we begin constructing elements that satisfy the conditions of Proposition 2.11. Since the statement of Proposition 2.11 is local, it suffices to replace $\mathfrak{X}$ by some sufficiently small affine neighborhood of y where the statement holds. We prove:

Lemma 2.15 *There exists a $\mathbb{G}_m$–equivariant identification $C \simeq S \times Z$ such that $\{\mathbb{C}[Y],\mathbb{C}[Z]\} = 0$.*

Proof Let $\mathbb{C}[Z] = \mathbb{C}[z_1,\dots,z_m]$. We show that the elements z_i can be modified so that the lemma holds. If $x_1,\dots,x_r$ is the basis of V_l^* as in the proof of Lemma 2.14, then that lemma implies that we may uniquely decompose

$$z_i = \sum_{I\in\mathbb{N}^r} x^I \cdot p_I^{(i)}$$

for $p_I^{(i)} \in K$. Since the x_i have degree $l > 0$ it follows that $\deg p_I^{(i)} < \deg z_i$ for all $I \neq 0$. Using this fact, it is straightforward to show by induction on degree that

$$\mathbb{C}[Y\times V_{>-l}][z_1,\dots,z_m] = \mathbb{C}[Y\times V_{>-l}][p_0^{(1)},\dots,p_0^{(m)}],$$

which implies the lemma. Indeed if $z_1,\dots,z_k$ have the minimal possible degree then $\deg p_I^{(i)} < \deg z_i$ for $I\neq 0$ implies that $p_I^{(i)} \in \mathbb{C}[Y\times V_{>-l}]$ and hence

$$\mathbb{C}[Y\times V_{>-l}][z_1,\dots,z_k] = \mathbb{C}[Y\times V_{>-l}][p_0^{(1)},\dots,p_0^{(k)}].$$

The inductive step is entirely analogous. □

Recall that the Lie algebra $\mathfrak{l} = I/I^2$ acts on R/I as well. Next, we consider the action of $\mathfrak{l}$ on $\mathbb{C}[Z]$. Recall that we have assumed that Y is small enough so that we have a $\mathbb{G}_m$–equivariant trivialization $\mathrm{Tot}(\mathsf{N}_{\mathfrak{X}/C}) \cong C\times Z^*$. It follows that $\mathfrak{l}$ is a free R/I–module of rank equal to $\dim Z$. We now consider the action of $\mathfrak{l}$ on $\mathbb{C}[Z]$ viewed as a subalgebra of R/I via the isomorphism constructed in Lemma 2.15:

Lemma 2.16 *There exists a homogeneous subspace $W \subset \mathfrak{l}$ such that $\mathfrak{l} = R/I \otimes W$ as an R/I–module and the action of $\mathfrak{l}$ on R/I restricts to a nondegenerate, equivariant pairing $W \times Z^* \to \mathbb{C}$.*

Proof Let $\{z_1, \dots, z_m\}$ be the generators of $\mathbb{C}[Z] \subset R/I$ as in the proof of Lemma 2.15, viewed as elements of Z^*. We must show that there is a homogeneous R/I–basis of $\mathfrak{l}$ dual to the z_i with respect to the pairing induced by the Poisson bracket. Fix $\{w_1, \dots, w_m\}$ some homogeneous R/I–basis. We will modify this basis by increasing induction on degree in order to obtain the required dual basis.

For each integer $-m < 0$ the Poisson bracket induces a $\mathbb{C}[Y]$–linear pairing between the $\mathfrak{l}_{-m}$, the $(-m)^{\text{th}}$ graded piece of $\mathfrak{l}$, and $\mathbb{C}[Z]_{l+m}$. Let Z^*_{l+m} be $\mathbb{C}$–span of the z_i of degree $l+m$ and let N_{l+m} be the $\mathbb{C}[Y]$–module it generates. Similarly, let $\mathfrak{n}_{-m} \subseteq \mathfrak{l}_{-m}$ be the $\mathbb{C}[Y]$–module generated by the w_j of degree $-m$. Since the pairing induced by the Poisson bracket on the tangent space $T_y\mathfrak{X}$ at a closed point $y \in Y$ is homogeneous and nondegenerate, it follows that the $\mathbb{C}[Y]$–pairing between N_{l+m} and $\mathfrak{n}_{-m}$ is also nondegenerate. It follows that we may modify the w_j in $\mathfrak{n}_{-m}$ to be dual to the $z_j \in N_{l+m}$, that is, so that $\{w_j, z_k\} = \delta_{j,k}$.

Thus we suppose by induction that for all $d < -m$ the w_j of degree $-d$ are dual to the z_k of degree $l+d$ and the pairings $\{w_j, z_k\} = 0$ for w_j and z_k in degrees less than $-m$ and greater than $l+m$, respectively, which are not paired. Note that for sufficiently large m this holds vacuously. In degree $-m$, modify the w_i to be dual to the z_k of degree $l+m$ as above. To complete the inductive step we must ensure the new w_i have trivial Poisson bracket with the z_k of higher degree; modify the w_i according to the homogeneous substitution

$$w_i \mapsto w_i - \sum_{\deg w_j < -m} \{w_i, z_j\} w_j .$$

Since the $\mathfrak{l}$–module structure on R/I is left R/I–linear, it follows from the inductive hypothesis that for these new w_i we have $\{w_i, z_k\} = 0$ whenever $\deg z_k > l+m$. Moreover, these new w_i are still linearly independent, as they remain dual to the z_k of degree $l+m$, since $\{w_j, z_k\} = 0$ for all w_j of degree less than $-m$ by consideration of degree. □

With these two lemmas in hand we may now give:

Proof of Proposition 2.11 If we denote by the same letter a homogeneous lift of the space W of Lemma 2.16 to $\widehat{I}$, then $\widehat{R} = (R/I)[\![W]\!]$. In this way we regard W and

Z^* as subspaces of $\widehat{R}$. Now, by Lemma 2.16 we may choose bases $w_1, \dots, w_m$ and $z_1, \dots, z_m$ of W and Z^*, respectively, so that conditions (2), (3) and the first condition of (4) in Proposition 2.11 hold. It remains to ensure the conditions that $\{z_i, z_j\}$ belong to $\widehat{I}$ and to find a suitable subalgebra T. For this we repeatedly use the substitution strategy of Lemma 2.16: inductively define

$$z_i' = z_i - \sum_{j=1}^{i-1} \{z_i, z_j'\} w_j \quad \text{for all } i.$$

One checks directly that these elements Poisson commute with each other modulo $\widehat{I}$ and, since $z_i' - z_i \in \widehat{I}$, we have that $\{z_i', \mathbb{C}[Y]\} = 0 \bmod \widehat{I}$ and $\{z_i', w_j\} = \delta_{ij} \bmod \widehat{I}$. Finally, if V is as in Lemma 2.5, and we pick a basis $x_1, \dots, x_t$ of V^* so that $\mathbb{C}[S] = \mathbb{C}[Y][x_1, \dots, x_t]$, and set

$$x_i' = x_i - \sum_{j=1}^{m} \{z_j', x_i\} w_j \quad \text{for all } i$$

and let T be the algebra generated by $\mathbb{C}[Y]$ and the x_i'. Then it follows easily that $\{z_i', T\} = 0 \bmod \widehat{I}$ and $\{w_i, T\} = 0 \bmod \widehat{I}$ (the latter from elementary degree considerations) and thus the proposition is proved. □

2.6 Special case of Theorem 2.1: isolated fixed points

Assume now that each fixed point component Y_i of the elliptic $\mathbb{G}_m$–action is a single point $\{p_i\}$. In this case each S_i is isomorphic to $\mathbb{A}^{2t_i}$ and there exists globally a splitting $S_i \hookrightarrow C_i \xrightarrow{\pi_i} S_i$. Let $(\mathbb{A}^{2n}, \omega_{\mathrm{can}})$ be the $2n$–dimensional affine space equipped with the constant symplectic form. We now check that in this case the symplectic reduction is isomorphic to a linear symplectic form.

Proposition 2.17 *Let (S, ω_S) be an affine symplectic manifold, isomorphic to $\mathbb{A}^{2n}$, equipped with an elliptic $\mathbb{G}_m$–action with unique fixed point $o \in S$. Then there exists an isomorphism $\phi\colon S \to \mathbb{A}^{2n}$ and homogeneous algebraically independent generators $\boldsymbol{z} = z_1, \dots, z_n$ and $\boldsymbol{w} = w_1, \dots, w_n$ of $\mathbb{C}[S]$ such that*

$$\phi^*\omega = dz_1 \wedge dw_1 + \dots + dz_n \wedge dw_n.$$

Proof Take the completion at $o \in S$. Then by the formal Darboux theorem we may choose homogeneous generators $\{u_1, \dots, u_n, v_1, \dots, v_n\}$ for which the Poisson bracket associated to ω has standard form. We may view $\mathbb{C}[S] \subset \widehat{\mathbb{C}[S]}$ via the obvious map, and so we are reduced to showing that the generators u_i and v_j are actually elements of $\mathbb{C}[S]$. But the $\mathbb{G}_m$–action naturally extends to $\widehat{\mathbb{C}[S]}$, and since o is the unique fixed point, which is therefore the limit of every point in S, it follows that we may describe $\mathbb{C}[S] \subset \widehat{\mathbb{C}[S]}$ as $(\widehat{\mathbb{C}[S]})^{\mathrm{rat}}$ the elements for which the $\mathbb{G}_m$–action is locally finite. Clearly any homogeneous element is locally finite and so we are done. □

Corollary 2.18 *Suppose $\mathfrak{X}^{\mathbb{G}_m}$ is finite, of cardinality k. Then, for all i, there is a coisotropic reduction $\pi_i\colon C_i \to (\mathbb{A}^{2t_i}, \omega_{\mathrm{can}})$.*

2.7 Symplectic fibrations

In this subsection, we consider those symplectic manifolds equipped with an elliptic $\mathbb{G}_m$–action for which the set of $\mathbb{G}_m$–fixed points is a connected variety. At one extreme, we have cotangent bundles of a smooth variety, with $\mathbb{G}_m$ acting by rescaling the fibers; at the other extreme one has symplectic fibrations, which are affine bundles such that each fiber is a copy of affine symplectic space. We show that generally one gets a mix of these two extremes.

Let Y be a smooth connected variety.

Definition 2.19

- A *symplectic fibration* over Y is a tuple $(E, \eta, \{-,-\})$, where $\eta\colon E \to Y$ is an affine bundle and $\{-,-\}$ an $\mathcal{O}_Y$–linear Poisson bracket on $\eta_\bullet \mathcal{O}_E$ such that the restriction of $\{-,-\}$ to each fiber of η is nondegenerate.
- The symplectic fibration is said to be *elliptic* if $\mathbb{G}_m$ acts on E such that $\{-,-\}$ is homogeneous of negative weight, $Y = E^{\mathbb{G}_m}$ and all weights of $\mathbb{G}_m$ on the fibers of η are negative.

If (E, η, ω) is an elliptic symplectic fibration then $\left(\eta^{-1}(y), \{-,-\}|_{\eta^{-1}(y)}\right)$ is a symplectic manifold for each $y \in Y$.

Since T^*Y is a vector bundle, it is naturally an abelian group scheme over the base Y.

Definition 2.20 Suppose $p\colon B \to Y$ is a smooth variety over Y equipped with a symplectic form ω_B. Suppose that B is equipped with an action $a\colon T^*Y \times_Y B \to B$ of the group scheme T^*Y over Y. We say $B \to Y$ is *symplectically automorphic* if, for any 1–form θ on Y, we have

$$a(\theta, -)^* \omega_B = \omega_B + p^* d\theta.$$

As remarked in the introduction, if B is a T^*Y–torsor, this reduces to the notion of *twisted cotangent bundle* as in [3].

For the remainder of this section we assume that $\mathfrak{X}$ is equipped with an elliptic $\mathbb{G}_m$–action such that the fixed point set Y of $\mathfrak{X}$ is connected and hence every point of $\mathfrak{X}$

has a limit in Y. By [10, Theorem 4.1], there is a smooth map $\rho\colon \mathfrak{X} \to Y$, making $\mathfrak{X}$ an affine bundle over Y. We should like to show that such symplectic varieties are symplectically automorphic varieties built from cotangent bundles and elliptic symplectic fibrations.

Theorem 2.21 *Let $\mathfrak{X}$ be a smooth symplectic variety equipped with an elliptic $\mathbb{G}_m$–action such that the fixed point locus $Y = \mathfrak{X}^{\mathbb{G}_m}$ is connected. Then:*

(1) *The group scheme T^*Y acts freely on $\mathfrak{X}$, making $\mathfrak{X} \to Y$ symplectically automorphic.*
(2) *The cotangent bundle T^*Y embeds T^*Y–equivariantly in $\mathfrak{X}$. Moreover, the restriction of the symplectic form ω to $T^*Y \subset \mathfrak{X}$ equals the standard 2–form on T^*Y.*
(3) *The quotient $E := \mathfrak{X}/T^*Y$ inherits a Poisson structure making the projection $E \to Y$ an elliptic symplectic fibration.*

The proof of Theorem 2.21 is similar to the proof of the local normal form Theorem 2.6; however, we must also use the powerful Artin approximation theorem of [2]. For brevity, we write $\mathcal{O}_{\mathfrak{X}}$ for the algebra $\rho_\bullet \mathcal{O}_{\mathfrak{X}}$. Let $\mathcal{K}$ be the ideal in $\mathcal{O}_{\mathfrak{X}}$ generated by $(\mathcal{O}_{\mathfrak{X}})_{0<i<l}$.

Lemma 2.22 *The symplectic form on $\mathfrak{X}$ restricts to a symplectic form on $\operatorname{Spec}_Y \mathcal{O}_{\mathfrak{X}}/\mathcal{K}$. Moreover, $\operatorname{Spec}_Y \mathcal{O}_{\mathfrak{X}}/\mathcal{K} \simeq T^*Y$ as smooth symplectic varieties.*

Proof The $\mathcal{O}_Y$–submodule $(\mathcal{O}_{\mathfrak{X}})_{\le l}$ of $\mathcal{O}_{\mathfrak{X}}$ is preserved by the Poisson bracket. If $f \in (\mathcal{O}_{\mathfrak{X}})_{\le l}$ and $g \in \mathcal{O}_Y$, then $\{f, g\} \in \mathcal{O}_Y$. Thus, $\{f, -\}$ defines a derivation of $\mathcal{O}_Y$. If $f \in (\mathcal{O}_{\mathfrak{X}})_{<l}$ then this derivation is zero. Hence, the Poisson bracket defines a $\mathcal{O}_Y$–linear map $\phi\colon \mathcal{E} \to \Theta_Y$, where $\mathcal{E} = (\mathcal{O}_{\mathfrak{X}})_{\le l}/(\mathcal{O}_{\mathfrak{X}})_{<l}$. This is a map of Lie algebroids:

$$\begin{aligned}[]
[\phi(f), \phi(g)](h) &= \phi(f)(\phi(g)(h)) - \phi(g)(\phi(f)(h)) \\
&= \{f, \{g, h\}\} - \{g, \{f, h\}\} \\
&= -\{h, \{f, g\}\} - \{g, \{h, f\}\} - \{g, \{f, h\}\} \\
&= \phi(\{f, g\})(h).
\end{aligned}$$

Since the Poisson bracket on $\mathfrak{X}$ is nondegenerate, ϕ is surjective. Locally trivializing $\mathfrak{X} \simeq Y \times V$ shows that $\mathcal{E}$ is locally free of rank $\dim V_{-l} = \dim Y$ and $\mathcal{O}_{\mathfrak{X}}/\mathcal{K} = \operatorname{Sym}^\bullet \mathcal{E}$. Therefore ϕ is an isomorphism. □

Lemma 2.23 *The algebra $\mathcal{O}_{\mathfrak{X}}$ is a comodule for $\operatorname{Sym}^\bullet \Theta_Y$.*

Proof Recall that the comultiplication on the Hopf algebra $\operatorname{Sym}^\bullet \Theta_Y$ is defined by $\Delta(v) = v \otimes 1 + 1 \otimes v$ for $v \in \Theta_Y$. The algebra $\mathcal{O}_{\mathfrak{X}}$ is generated by $(\mathcal{O}_{\mathfrak{X}})_{\le l}$. We define $\Delta_{\mathfrak{X}}\colon \mathcal{O}_{\mathfrak{X}} \to \mathcal{O}_{\mathfrak{X}} \otimes_{\mathcal{O}_Y} \operatorname{Sym}^\bullet \Theta_Y$ by $\Delta_{\mathfrak{X}}(f) = f \otimes 1 + 1 \otimes \overline{f}$ for $f \in (\mathcal{O}_{\mathfrak{X}})_{\le l}$ and extending to $\mathcal{O}_{\mathfrak{X}}$ using the fact that $\Delta_{\mathfrak{X}}$ should be an algebra homomorphism. Here $\overline{f}$ denotes the image of f in $\mathcal{O}_{\mathfrak{X}}/\mathcal{K} \simeq \operatorname{Sym}^\bullet \Theta_Y$. Since $\mathcal{O}_{\mathfrak{X}}$ is not freely generated by $(\mathcal{O}_{\mathfrak{X}})_{\le l}$, we need to show that this is well-defined.

Choose a local set $x_1, \dots, x_k$, $z_1, \dots, z_r$ of homogeneous algebraically independent generators of $\mathcal{O}_{\mathfrak{X}}$ over $\mathcal{O}_Y$ such that $0 < \deg x_i < l$ and $\deg z_j = l$ for all i and j. The images $\overline{z}_i$ in $\operatorname{Sym}^\bullet \Theta_Y$ of the z_i form a basis of Θ_Y. Then we define $\Delta'_{\mathfrak{X}}$ by $\Delta'_{\mathfrak{X}}(f) = f \otimes 1 + 1 \otimes \overline{f}$ for $f \in \{x_1, \dots, x_k, z_1, \dots, z_r\}$. This clearly defines a local comodule structure on $\mathcal{O}_{\mathfrak{X}}$. We just need to show that it equals $\Delta_{\mathfrak{X}}$, ie it is independent of the choice of local generators. Take $f \in (\mathcal{O}_{\mathfrak{X}})_{\le l}$, a homogeneous element. There exist $a_i \in \mathcal{O}_Y$ and some g such that $f = \sum_{i=1}^r a_i z_i + g(x_1, \dots, x_k)$. Then

$$\begin{aligned}
\Delta'_{\mathfrak{X}}(f) &= \sum_{i=1}^r a_i \Delta'_{\mathfrak{X}}(z_i) + g(\Delta'_{\mathfrak{X}}(x_1), \dots, \Delta'_{\mathfrak{X}}(x_k)) \\
&= \sum_{i=1}^r (a_i z_i \otimes 1 + 1 \otimes a_i \overline{z}_i) + g(x_1 \otimes 1, \dots, x_k \otimes 1) \\
&= \Bigl(\sum_{i=1}^r a_i z_i\Bigr) \otimes 1 + 1 \otimes \overline{\Bigl(\sum_{i=1}^r a_i z_i\Bigr)} + g(x_1, \dots, x_k) \otimes 1 \\
&= f \otimes 1 + 1 \otimes \overline{f} = \Delta_{\mathfrak{X}}(f).
\end{aligned}$$

Finally, to check that $\Delta_{\mathfrak{X}}$ is compatible with the comultiplication on $\operatorname{Sym}^\bullet \Theta_Y$, it suffices to do so locally, as above, where it is clear. □

We denote by $\mathcal{F}$ the graded $\mathcal{O}_Y$–subalgebra of $\mathcal{O}_{\mathfrak{X}}$ generated by $(\mathcal{O}_{\mathfrak{X}})_{<l}$. Set $E := \operatorname{Spec}_Y \mathcal{F}$ and let $\mathfrak{X}/T^*Y$ denote the spectrum of $\mathcal{O}_{\mathfrak{X}}^{\Delta_{\mathfrak{X}}}$. From the definition of $\Delta_{\mathfrak{X}}$ given in the proof of Lemma 2.23, the algebra $\mathcal{F}$ is contained in $\mathcal{O}_{\mathfrak{X}}^{\Delta_{\mathfrak{X}}}$. Thus, we have a dominant map $\mathfrak{X}/T^*Y \to E$. Moreover, the local description of $\Delta_{\mathfrak{X}}$ given in the proof of Lemma 2.23 shows that $\mathfrak{X}/T^*Y$ is the space of T^*Y–orbits in $\mathfrak{X}$ and the map $\mathfrak{X}/T^*Y \to E$ is an isomorphism. Thus, parts (1) and (2) of Theorem 2.21, except for the "symplectically automorphic" assertion of (1), are a consequence of Lemmas 2.22 and 2.23. For part (3) we need to show that the Poisson bracket on $\mathcal{O}_{\mathfrak{X}}$, restricted to $\mathcal{F}$, makes E into a symplectic fibration. For this and the "symplectically automorphic" assertion, we need a local normal form of the Poisson bracket on $\mathcal{O}_{\mathfrak{X}}$.

Proposition 2.24 *Locally, in the étale topology on Y, there is a $\mathbb{G}_m$–equivariant isomorphism of Poisson algebras*

$$\mathcal{O}_{\mathfrak{X}} \simeq \mathcal{O}_{T^*Y} \otimes \mathbb{C}[\mathbb{A}^{2n}],$$

where the inherited Poisson bracket on $\mathbb{C}[\mathbb{A}^{2n}]$ is the standard one.

Proof Let $r = \dim Y$. Since the statement is local we assume that $\mathfrak{X} \simeq Y \times V$ is affine, where V is a $\mathbb{G}_m$–module with strictly negative weights. The algebra $R := \mathbb{C}[\mathfrak{X}]$ is isomorphic to $\mathbb{C}[Y] \otimes \mathbb{C}[V]$. If $x_1, \dots, x_{2n}$ is a homogeneous basis of $V^*_{<l} \subset \mathbb{C}[V]$ and $x_{2n+1}, \dots, x_{2n+r}$ a basis of $V^*_l \subset \mathbb{C}[V]$, then $\mathbb{C}[\mathfrak{X}] = \mathbb{C}[Y] \otimes \mathbb{C}[x_1, \dots, x_{2n+r}]$. Concretely, we wish to show that there exists an equivariant étale morphism $p \colon U \to Y$ such that $\Gamma(U, p^*\mathbb{C}[\mathfrak{X}])$ is isomorphic to $\mathbb{C}[U][x'_1, \dots, x'_{2n+r}]$, where the $\{x'_i\}$ are graded elements, with $\deg x_i = \deg x'_i$, and the Poisson bracket satisfies $\{x'_i, x'_j\} = \delta_{i+j,2n+1}$. Assuming we have done this, weight considerations imply that

$$\Gamma(U, p^*\mathbb{C}[\mathfrak{X}]) \simeq \mathbb{C}[U][x'_{2n+1}, \dots, x'_{2n+r}] \otimes \mathbb{C}[\mathbb{A}^{2n}]$$

as Poisson algebras and Lemma 2.12 implies that $\mathbb{C}[U][x'_{2n+1}, \dots, x'_{2n+r}] \simeq \mathbb{C}[T^*U]$ as Poisson algebras.

Remark 2.25 Since this identification is compatible with the action of T^*Y, the "symplectically automorphic" assertion of Theorem 2.21(1) follows immediately from the same assertion for T^*Y itself.

Choose $y \in Y$, and consider the algebra $\widehat{R}_y^{\mathrm{rat}} \simeq \widehat{\mathbb{C}}[Y]_y \otimes \mathbb{C}[V]$ of $\mathbb{G}_m$–locally finite sections of the completion of R at y. Let $\mathfrak{m}$ denote the maximal ideal of $(y, 0)$ in $\widehat{R}_y^{\mathrm{rat}}$. The formal Darboux theorem, Theorem 2.13, implies that there exist homogeneous elements $u_1, \dots, u_{2n+r}$ in $\widehat{R}_y^{\mathrm{rat}}$ such that $\{u_i, u_j\} = \delta_{i+j,2n+1}$ and their image in $\mathfrak{m}/\mathfrak{m}^2 = T_y^*Y \times V^*$ is a basis of V^*. With respect to our chosen basis, $u_i = \sum_j g_{ij} x_j$ for some $g_{ij} \in \widehat{\mathbb{C}}[Y]_y$. This implies the relations

$$\delta_{i+j,2n+1} = \{u_i, u_j\} = \sum_{k,l} g_{ik} g_{jl} \{x_k, x_l\}, \tag{2-4}$$

where the last equality follows from the fact that $\{x_k, g_{ij}\} = 0$ for all i, j and k for reasons of degree.

Since these relations are taking place in the finite, free $\mathbb{C}[Y]$–module $\mathbb{C}[Y] \otimes \mathbb{C}[V]_{\leq l}$ we may consider (2-4) as defining a system of polynomial equations in the variables g_{ij}. Then the formal Darboux theorem says that these equations have solutions in the completion of Y at y. The Artin approximation theorem [2, Corollary 2.1] now assures us of the existence of an étale neighborhood of y (ie an étale map $p \colon U \to Y$ with a chosen closed point x living above y) in which the equations have a solution $\{g'_{ij}\}$, whose difference with the solution in $\widehat{\mathbb{C}}[Y]_y = \widehat{\mathbb{C}}[U]_x$ lies in $\mathfrak{n}$, the maximal ideal of this local ring.

Now, we can base change the affine map $\mathfrak{X} \to Y$ to get an affine bundle $\mathfrak{X}_U \to U$; furthermore, the induced map $\mathfrak{X}_U \to \mathfrak{X}$ is étale since $U \to Y$ is. Therefore, the pullback of the Poisson bracket induces a Poisson bracket on $p^*\mathbb{C}[\mathfrak{X}]$, which is homogeneous of weight $-l$ by construction. Thus the given solutions of these equations yield elements

$$x_i' = \sum g_{ij}' x_j \in p^*\mathbb{C}[\mathfrak{X}]$$

which satisfy the Poisson relations as in the conclusion of the proposition. Furthermore, the choice of x_i' implies that the determinant of the $\mathbb{C}[U]$–linear transformation $x_i \mapsto x_i'$ is nonzero at $p^{-1}(y)$ (because this is true for $\widehat{\mathbb{C}}[Y]_y$–linear transformation $x_i \mapsto u_i$). Thus, restricting to smaller neighborhoods if necessary, we may assume that the map $x_i \mapsto x_i'$ is invertible; this implies that the x_i' are algebra generators of $p^*\mathbb{C}[\mathfrak{X}]$ over $\mathbb{C}[U]$, which proves the proposition □

Proposition 2.24 implies that E is a symplectic fibration. This proves Theorem 2.21.

Remark 2.26 It follows from [10, Theorem 2.2] that the equivariant closed embedding $T^*Y \hookrightarrow \mathfrak{X}$ of Theorem 2.21 is unique. We also note that Theorem 2.21 implies that if $\dim \mathfrak{X} = 2\dim Y$ then $\mathfrak{X} \simeq T^*Y$. Moreover, the proof of Theorem 2.21 shows that, locally in the Zariski topology,

$$\mathfrak{X} \simeq T^*Y \times_Y E$$

as smooth varieties with $\mathbb{G}_m$–actions, though not as Poisson varieties.

2.8 Reductions of coisotropic subvarieties

We can now give a generalization of Theorem 2.1. Fix a coisotropic stratum $\rho\colon C \to Y$ and let $\pi\colon C \to S$ be the coisotropic reduction of C given by Theorem 2.1. Let $Y' \subset Y$ be a smooth, closed subvariety and set $C' = \rho^{-1}(Y')$. The proof of Theorem 2.1(1) shows that $C' \subset C$ is coisotropic. Let $\mathcal{I}$ be the sheaf of ideals in $\mathcal{O}_S$ vanishing on C'. Since it is generated by elements in degree zero, it is an involutive ideal. Let $\rho'\colon S \to Y$ be the projection map. We perform coisotropic reduction as before and set

$$S' := \operatorname{Spec}_{Y'}(\rho'_\bullet \mathcal{O}_S/\mathcal{I})^{\{\mathcal{I},-\}}, \tag{2-5}$$

a Poisson variety. Just as in the proof of Theorem 2.1, to show that S' is a coisotropic reduction of C', it suffices to do so locally on Y. Instead of considering the formal neighborhood of Y in S, we pass to an étale local neighborhood. Thus, by Proposition 2.24, we may assume that $S \simeq T^*Y \times \mathbb{A}^{2n}$. Then $C' = (T^*Y)|_{Y'} \times \mathbb{A}^{2n}$ and S' becomes the classical coisotropic reduction $T^*Y' \times \mathbb{A}^{2n}$. Thus, we have shown:

Corollary 2.27 *The space C' is coisotropic in $\mathfrak{X}$ and there is a coisotropic reduction $\pi'\colon C' \to S'$.*

In Section 4.10 the above coisotropic reduction will be quantized.

3 Deformation-quantization modules

In this section, we recall the basic properties of DQ algebras and their modules. We also prove an extension result, Theorem 3.27, which will play a key role in Section 5.

3.1 DQ algebras: affine setting

We begin by recalling the definition of a deformation-quantization algebra. Let $(R, \{-,-\})$ be a regular Poisson $\mathbb{C}$–algebra.

Definition 3.1 A *deformation-quantization* of R is an $\hbar$–flat and $\hbar$–adically complete $\mathbb{C}[\![\hbar]\!]$–algebra A equipped with an isomorphism of Poisson algebras $A/\hbar A \cong R$. Here $A/\hbar A$ is equipped with a Poisson bracket via

$$\{\overline{a}, \overline{b}\} = \left(\frac{1}{\hbar}[a,b]\right) \mod \hbar A,$$

for arbitrary lifts a and b of $\overline{a}$ and $\overline{b}$ in A. The algebra A is a *deformation-quantization algebra* or *DQ algebra* if it is a deformation-quantization of some regular Poisson algebra R. An *isomorphism of DQ algebras* that are quantizations of the same Poisson algebra R is a $\mathbb{C}[\![\hbar]\!]$–algebra isomorphism such that the induced map on R is the identity.

Lemma 3.2 *Let A be a deformation-quantization algebra.*

(1) *A is a (left and right) Noetherian domain of finite global dimension.*
(2) *The Rees ring $\mathsf{Rees}_\hbar A = \bigoplus_{n\geq 0} \hbar^n A$ is Noetherian.*
(3) *If M is a finitely generated A–module, then M is $\hbar$–complete.*

Proof Part (1) follows from the fact that the associated graded of A with respect to the $\hbar$–adic filtration is $R[\hbar]$, which is a regular domain. Part (2) is shown in [48, Lemma 2.4.2]. Part (3) is a consequence of the Artin–Rees lemma. □

We also have the following well-known complete version of Nakayama's lemma:

Lemma 3.3 *Let M be a complete $\mathbb{C}[\![\hbar]\!]$–module. If $\hbar M = M$ then $M = 0$.*

3.2 Sheaves of DQ algebras

Let $(\mathfrak{X}, \{-,-\})$ be a smooth Poisson variety. A sheaf of $\mathbb{C}[\![\hbar]\!]$–modules $\mathcal{A}$ on $\mathfrak{X}$ is said to be $\hbar$–flat if each stalk $\mathcal{A}_p$ is a flat $\mathbb{C}[\![\hbar]\!]$–module. For each positive integer n, let $\mathcal{A}_n = \mathcal{A}/\hbar^n\mathcal{A}$. The $\hbar$–adic completion of $\mathcal{A}$ is $\widehat{\mathcal{A}} = \varprojlim_n \mathcal{A}_n$ and $\mathcal{A}$ is said to be $\hbar$–adically complete if the canonical morphism $\mathcal{A} \to \widehat{\mathcal{A}}$ is an isomorphism.

Definition 3.4 A sheaf of $\mathbb{C}[\![\hbar]\!]$–algebras $\mathcal{A}$ on X is said to be a *deformation-quantization* algebra if it is $\hbar$–flat and $\hbar$–adically complete, equipped with an isomorphism of Poisson algebras $\mathcal{A}_0 \cong \mathcal{O}_X$.

If, moreover, the algebra $\mathcal{A}$ is equipped with a $\mathbb{G}_m$–action that acts on $\hbar \in \mathcal{A}$ with weight l and the Poisson bracket on $\mathfrak{X}$ has degree $-l$, then we replace $\mathcal{A}$ by $\mathcal{A}[\hbar^{1/l}]$ and $\hbar$ by $\hbar^{1/l}$, so that, without loss of generality $t \cdot \hbar = t\hbar$ and the Poisson bracket on $\mathcal{O}_{\mathfrak{X}}$ coming from $\mathcal{A}$ is defined by

$$\{\overline{a}, \overline{b}\} := \frac{1}{\hbar^l}[a, b] \mod \hbar\mathcal{A}.$$

Remark 3.5 For a symplectic variety $\mathfrak{X}$ with $\mathbb{G}_m$–action and deformation-quantization $\mathcal{A}$ we always assume that $\mathcal{A}$ is equivariant in the above sense.

In the algebraic setting, the existence and classification of sheaves of deformation-quantization algebras is well understood. See [8; 49] for the equivariant setting. Assume that $\mathfrak{X}$ is affine and let $R = \mathbb{C}[\mathfrak{X}]$; let A be a deformation-quantization of R. For any multiplicatively closed subset S of R, there is an associated microlocalization $Q_S^{\mu}(A)$ of A; the algebra $Q_S^{\mu}(A)$ is, by definition, a deformation-quantization of R_S. Using Gabriel filters, one can extend the notion of microlocalization to define a presheaf $\mathcal{O}_A^{\mu}$ of algebras on $\mathfrak{X}$ such that $\Gamma(D(f), \mathcal{O}_A^{\mu}) = Q_f^{\mu}(A)$ for all $f \in R$. By [68, Theorem 4.2], the presheaf $\mathcal{O}_A^{\mu}$ is a sheaf and the following proposition holds:

Proposition 3.6 *Assume that $\mathfrak{X}$ is affine. Then microlocalization defines an equivalence between the category of DQ algebras quantizing R and sheaves of DQ algebras on $\mathfrak{X}$.*

3.3 Sheaves of DQ modules

In this section, we define those $\mathcal{A}$–modules that will be studied in Sections 4 and 5. First:

Lemma 3.7 *If $H^1(\mathfrak{X}, \mathcal{O}_X) = 0$, then $\Gamma(\mathfrak{X}, \mathcal{A}_n) = \Gamma(\mathfrak{X}, \mathcal{A})/\hbar^n\Gamma(\mathfrak{X}, \mathcal{A})$.*

Proof Consider the short exact sequence $0 \to \mathcal{A} \xrightarrow{\cdot \hbar^n} \mathcal{A} \to \mathcal{A}_n \to 0$. Taking derived global sections, it suffices to show that $H^1(\mathfrak{X}, \mathcal{A}) = 0$. Induction on n using the exact sequence $0 \to \hbar\mathcal{A}_{n-1} \to \mathcal{A}_n \to \mathcal{O}_{\mathfrak{X}} \to 0$ shows that $H^1(\mathfrak{X}, \mathcal{O}_X) = 0$ implies $H^1(\mathfrak{X}, \mathcal{A}_n) = 0$. Then [28, Proposition 13.3.1] implies that $H^1(\mathfrak{X}, \mathcal{A}) = 0$. □

Thus, write $A = \Gamma(\mathfrak{X}, \mathcal{A})$ and $A_n = \Gamma(\mathfrak{X}, \mathcal{A}_n)$. If M is an A_n–module then, thinking of A_n and M as constant sheaves on $\mathfrak{X}$, we define $M^{\Delta} = \mathcal{A}_n \otimes_{A_n} M$.

Definition 3.8 An $\mathcal{A}_n$–module $\mathscr{M}_n$ is *quasicoherent* if there exists an affine open covering $\{U_i\}$ of $\mathfrak{X}$ such that $\mathscr{M}_n|_{U_i} \simeq \Gamma(U_i, \mathscr{M}_n)^{\Delta}$. If, moreover, $\Gamma(U_i, \mathcal{A}_n)$ is a finitely generated $\Gamma(U_i, \mathscr{M}_n)$–module for all i then $\mathscr{M}_n$ is said to be *coherent*.

As in the commutative case, we have:

Proposition 3.9 *An $\mathcal{A}_n$–module $\mathscr{M}_n$ is quasicoherent if and only if*

$$\mathscr{M}_n|_U \simeq \Gamma(U, \mathscr{M}_n)^{\Delta}$$

for all affine open subsets U of $\mathfrak{X}$.

Assume that $\mathfrak{X}$ is affine and let $A = \Gamma(\mathfrak{X}, \mathcal{A})$. Let M be an A–module. We define $M^{\Delta} = \varprojlim_n (\mathcal{A}_n \otimes_{A_n} M/\hbar^n M)$. We remark that if $\{\mathcal{M}_n\}$ is an inverse system of sheaves, then $\varprojlim_n \mathcal{M}_n$ is defined to be the sheaf $U \mapsto \varprojlim_n \Gamma(U, \mathcal{M}_n)$; there is no need to sheafify. In particular, $\Gamma(\mathfrak{X}, \varprojlim_n \mathcal{M}_n) = \varprojlim_n \Gamma(\mathfrak{X}, \mathcal{M}_n)$.

We may now define the two classes of modules that play a role in this paper:

Definition 3.10 Let $\mathscr{M}$ be an $\mathcal{A}$–module.

(1) $\mathscr{M}$ is *coherent* if it is $\hbar$–complete and each $\mathscr{M}_n$ is a coherent $\mathcal{A}_n$–module.

(2) $\mathscr{M}$ is *quasicoherent* if it the union of its coherent $\mathcal{A}$–submodules.

The category of all coherent (resp. quasicoherent) $\mathcal{A}$–modules is denoted by $\mathcal{A}\text{–coh}$ (resp. $\mathcal{A}\text{–qcoh}$). The proof of the following is based on [1, Theorem 5.5]:

Proposition 3.11 *Assume that $\mathfrak{X}$ is affine. Then $\Gamma(\mathfrak{X}, -)$ defines an exact equivalence between $\mathcal{A}\text{–coh}$ and the category $A\text{–mod}$ of finitely generated A–modules, where $A = \Gamma(\mathfrak{X}, \mathcal{A})$. A quasi-inverse is given by $M \mapsto M^{\Delta}$. Moreover, $M^{\Delta} = \mathcal{A} \otimes_A M$.*

Proof As for any localization theorem, the proof has three parts. First, we show that Γ is exact on $\mathcal{A}\text{–coh}$. Then we show that $\Gamma(\mathfrak{X}, \mathscr{M})$ is a finitely generated A–module, for all $\mathscr{M} \in \mathcal{A}\text{–coh}$. Finally, we show that $\mathscr{M}$ is generated by its global sections.

Let $M_n = \Gamma(\mathfrak{X}, \mathscr{M}_n)$ and $M = \Gamma(\mathfrak{X}, \mathscr{M})$. Since the $\mathcal{A}_n$–module $\hbar^{n-1}\mathscr{M}/\hbar^n\mathscr{M}$ is a submodule of the coherent $\mathcal{A}_n$–module $\mathscr{M}_n$, it is coherent, and hence Proposition 3.9 implies that the cohomology groups $H^i(\mathfrak{X}, \hbar^{n-1}\mathscr{M}/\hbar^n\mathscr{M})$ are zero for all $i \neq 0$. Therefore, we have surjective maps $M_n \to M_{n-1} \to \cdots$. Therefore, the inverse system $\{M_n\}_n$ satisfies the Mittag-Leffler condition and hence [28, Proposition 13.3.1], together with the fact that $\mathscr{M}$ is assumed to be complete, implies that $H^i(\mathfrak{X}, \mathscr{M}) = 0$ for all $i \neq 0$.

Since $\mathscr{M}_n/\hbar^n\mathscr{M}_n \cong \mathscr{M}_{n-1}$, the fact that $H^1(\mathfrak{X}, \mathscr{M}_n) = 0$ implies $M_n/\hbar^n M_n = M_{n-1}$. Therefore, by [6, Lemma 3.2.2], the fact that each M_n is a finitely generated A_n–module implies that M is a finitely generated A–module.

The fact that Γ is exact implies that $M_n = M/\hbar^n M$. Therefore, by Proposition 3.9,

$$\mathscr{M} \simeq \varprojlim_n \mathscr{M}_n = \varprojlim_n \mathcal{A}_n \otimes_{A_n} M_n = \varprojlim_n \mathcal{A}_n \otimes_{A_n} M/\hbar^n M = M^\Delta.$$

Finally, to show that this is the same as $\mathcal{A} \otimes_A M$, let $A^n \to A^m \to M \to 0$ be a finite presentation of M. Since we have a natural map $\mathcal{A} \otimes_A M \to \mathcal{A}_n \otimes_{A_n} M/\hbar^n M$ for all n, there is a canonical morphism $\mathcal{A} \otimes_A M \to M^\Delta$. Then we get the usual commutative diagram

$$\begin{tikzcd}
\mathcal{A} \otimes_A A^n \arrow[r] \arrow[d] & \mathcal{A} \otimes A^m \arrow[r] \arrow[d] & \mathcal{A} \otimes_A M \arrow[r] \arrow[d] & 0 \\
(A^n)^\Delta \arrow[r] & (A^m)^\Delta \arrow[r] & M^\Delta \arrow[r] & 0
\end{tikzcd}$$

so the result follows from the five lemma and the fact that $(A^m)^\Delta = (\mathcal{A})^m$, which in turn is a consequence of the fact that microlocalization is an additive functor. □

Remark 3.12 It is clear that, under the identification of the proposition, if $U \subset \mathfrak{X}$ is an inclusion of affine opens, the restriction functor $\mathscr{M} \mapsto \mathscr{M}|_U$ is identified with $M \mapsto H^0(U, \mathcal{A}) \otimes_A M$.

Corollary 3.13 *Let $\mathscr{M} \in \mathcal{A}\text{–coh}$ and $U \subset \mathfrak{X}$ an affine open set. Then $H^i(U, \mathscr{M}) = 0$ for all $i \neq 0$ and $\mathscr{M}|_U \simeq \Gamma(U, \mathscr{M})^\Delta$, where M is a finitely generated $\Gamma(U, \mathcal{A})$–module.*

Lemma 3.14 *Let $\mathscr{M}$ be an $\mathcal{A}$–module. The following are equivalent:*

(1) *$\mathscr{M}$ is coherent.*

(2) *$\mathscr{M}$ is locally finitely presented.*

(3) *$\mathscr{M}$ is locally finitely generated.*

Proof (1) $\Longrightarrow$ (2) Let U be an affine open subset of $\mathfrak{X}$ and $M = \Gamma(U, \mathscr{M})$. Then M is finitely generated. Since A is Noetherian, it is actually finitely presented and hence there is sequence $A^n \to A^m \to M \to 0$. The functor Δ is an equivalence on $A\text{–mod}$, hence we have $\mathcal{A}^n|_U \to \mathcal{A}^m|_U \to \mathscr{M} \to 0$.

(2) $\Longrightarrow$ (3) This is clear.

(3) $\Longrightarrow$ (1) We have $\phi\colon \mathcal{A}^m|_U \twoheadrightarrow \mathscr{M}|_U$ and hence $\mathcal{A}_n^m|_U \twoheadrightarrow \mathscr{M}_n|_U$. Thus, each $\mathscr{M}_n|_U$ is coherent. The module $\ker\phi$ is a submodule of the coherent $\mathcal{A}$–module $\mathcal{A}^m|_U$. Therefore the Artin–Rees lemma implies that the filtrations $\{\hbar^n \ker\phi\}$ and $\{(\ker\phi) \cap (\hbar^n \mathcal{A}^m)\}$ are comparable. Hence $\varprojlim^{(1)} (\ker\phi)/[(\ker\phi) \cap \hbar^n(\mathcal{A}^m)] = 0$. This implies that $\mathscr{M}$ is complete. □

Proposition 3.15 *Suppose $\mathfrak{X}$ is affine and that $j\colon U \hookrightarrow \mathfrak{X}$ is an open subset. Suppose that $\mathscr{M}$ is a coherent $\mathcal{A}_U$–module. Then $\mathscr{M}$ is globally generated.*

Proof First, let $p \in U$ be any point. Write $\mathscr{M}_m = \mathscr{M}/\hbar^m\mathscr{M}$, and let $\mathscr{M}_0[p]$ denote the fiber of $\mathscr{M}_0$ at p. We have that $j_*\mathscr{M}_m$ is a quasicoherent $\mathcal{A}$–module; thus, by Proposition 3.11, $j_*\mathscr{M}_m$ is the union of its globally generated subsheaves, hence is itself globally generated. Thus $\Gamma(\mathscr{M}_m) \to \mathscr{M}_0[p]$ is surjective. Taking (inverse) limits and applying [30, Theorem 4.5], we get that $\Gamma(\mathscr{M}) \to \mathscr{M}_0[p]$ is surjective. It follows (by a standard argument) from Nakayama's lemma that $\Gamma(\mathscr{M}) \otimes \mathcal{A} \to \mathscr{M}_0$ is surjective. Writing ev: $\Gamma(\mathscr{M}) \otimes \mathcal{A} \to \mathscr{M}$ for the evaluation map, we get that $\mathscr{M} = \mathrm{Im}(\mathrm{ev}) + \hbar\mathscr{M}$, ie that $\mathscr{M}/\,\mathrm{Im}(\mathrm{ev}) = \hbar \cdot (\mathscr{M}/\,\mathrm{Im}(\mathrm{ev}))$, and thus by Lemma 3.3 that ev is surjective. □

We remark that, in the proof of Proposition 3.15, we use only quasicoherence of $j_*\mathscr{M}_m$ (and not of the naive sheaf-theoretic image $j_\bullet\mathscr{M}$, which we expect is not quasicoherent in general).

Corollary 3.16 *Let $\mathfrak{X}$ be an affine variety and $U \subset \mathfrak{X}$ an open subset with complement $C = \mathfrak{X} \smallsetminus U$. Let $\mathcal{A}\text{–coh}_C$ denote the subcategory of sheaves supported on C (that is, the kernel of the restriction-to-U functor). Suppose that the induced functor $\mathcal{A}\text{–coh}/\mathcal{A}\text{–coh}_C \to \mathcal{A}_U\text{–coh}$ is full. Then $\mathcal{A}\text{–coh} \to \mathcal{A}_U\text{–coh}$ is essentially surjective.*

Proof By Proposition 3.15, objects of $\mathcal{A}$–coh are globally generated; hence, given $\mathscr{M} \in \mathcal{A}$–coh, we may produce a presentation $\mathcal{A}_U^I \xrightarrow{\phi} \mathcal{A}_U^J \to \mathscr{M} \to 0$ with I and J finite index sets. It suffices to prove that there are objects F_1, $F_0 \in \mathcal{A}$–coh, a morphism $\tilde{\phi}\colon F_1 \to F_0$, and isomorphisms $F_1|_U \cong \mathcal{A}_U^I$ and $F_0|_U \cong \mathcal{A}_U^J$ that identify $\tilde{\phi}|_U$ with ϕ: then $\mathscr{M} \cong \mathrm{coker}(\tilde{\phi})|_U$. But since $\mathcal{A}_U^I$ and $\mathcal{A}_U^J$ are in the essential image of the functor $\mathcal{A}\text{–coh}/\mathcal{A}\text{–coh}_C \to \mathcal{A}_U\text{–coh}$, this follows immediately from the fullness hypothesis. □

Recall that a $\mathbb{C}[\![\hbar]\!]$–module is flat if and only if it is torsion-free. The following is a consequence of [71, Theorem 5.6]:

Lemma 3.17 *Let $\mathscr{M}$ be a $\hbar$–adically complete and $\hbar$–flat $\mathcal{A}$–module. Let $U \subset \mathfrak{X}$ be an affine open set. Then $\Gamma(U, \mathscr{M}) \simeq \Gamma(U, \mathscr{M}_0) \mathbin{\widehat{\otimes}} \mathbb{C}[\![\hbar]\!]$ as $\mathbb{C}[\![\hbar]\!]$–modules, ie $\Gamma(U, \mathscr{M})$ is $\hbar$–adically free.*

Based on Lemma 3.17, a $\mathcal{A}$–module $\mathscr{M}$ is said to be $\hbar$–adically free if it is $\hbar$–adically complete and $\hbar$–flat. At the other extreme, an $\mathcal{A}$–module $\mathscr{M}$ is said to be $\hbar$–torsion if, for each $p \in \mathfrak{X}$, there exists an affine open neighborhood U of p and $N \gg 0$ such that $\hbar^N \cdot \Gamma(U, \mathscr{M}) = 0$. Since $\mathfrak{X}$ is assumed to be quasicompact, this is equivalent to requiring that $\hbar^N \cdot \mathscr{M} = 0$ for $N \gg 0$. We define $\mathcal{A}\text{–coh}_{\mathrm{tor}}$ be the full subcategory of $\mathcal{A}$–coh consisting of all $\hbar$–torsion sheaves.

3.4 Equivariant algebras and modules

Terminology 3.18 Let T be a torus, ie T is isomorphic to $\mathbb{G}_m^n$ for some n. We recall that a representation M of T is *prorational* if it is the inverse limit of rational T–modules. Fix a character χ of T and let T act on $\mathbb{C}[\![\hbar]\!]$ by $t \cdot \hbar = \chi(t)\hbar$.

Let $(\mathfrak{X}, \omega)$ be an affine symplectic variety with $\mathbb{G}_m$–action. Assume that $m_t^* \omega = \chi(t)\omega$ for all $t \in \mathsf{T}$.

Definition 3.19 A complete $\mathbb{C}[\![\hbar]\!]$–algebra A is said to be T–*equivariant* if A is a prorational T–module such that $g \cdot (ab) = (g \cdot a)(g \cdot b)$, and $g \cdot \hbar = \chi(g)\hbar$. The algebra A is a T–equivariant deformation-quantization of $\mathbb{C}[\mathfrak{X}]$ if A comes equipped with a T–equivariant isomorphism $A/\hbar A \cong \mathbb{C}[\mathfrak{X}]$.

A finitely generated A–module is said to be T–*equivariant* (or just equivariant) if it is a prorational T–module such that the multiplication map $A \otimes M \to M$ is equivariant.

The category of all finitely generated, equivariant A–modules is denoted by (A, T)–mod and the corresponding ind-category is (A, T)–Mod. The morphisms in these categories are equivariant.

Proposition 3.20 *(A, T)–mod and (A, T)–Mod are abelian categories.*

Lemma 3.21 *Let $M \in (A, \mathsf{T})$–mod. Then there exists a finite-dimensional T–submodule V of M such that $M = A \cdot V$.*

Proof Nakayama's lemma implies that if V is any subspace of M whose image in $M/\hbar M$ generates $M/\hbar M$, then V generates M. As noted in [26, Section 5.2.1], each M_n is a rational T–module and M is the inverse of these T–modules. Since T is reductive, we may fix T–equivariant splittings $M_n = K_n \oplus M_{n-1}$ such that K_n is the kernel of $M_n \twoheadrightarrow M_{n-1}$. This implies that $M = \prod_n K_n$ as a T–module. Hence, if we choose a finite-dimensional T–submodule V' of $K_0 = M/\hbar M$ that generates $M/\hbar M$ as an A_0–module, then we can choose a T–module lift V of V' in M. $\square$

3.5 Equivariant DQ algebras and modules

We maintain Terminology 3.18.

Let $(\mathfrak{X}, \omega)$ be any smooth symplectic variety with T–action: assume $m_t^* \omega = \chi(t)\omega$ for all $t \in \mathsf{T}$.

Definition 3.22 A deformation-quantization $\mathcal{A}$ of $\mathfrak{X}$ is said to be T–*equivariant* if it is equipped with the structure of a T–equivariant sheaf of algebras, with T acting on $\mathbb{C}[\![\hbar]\!]$ as in Terminology 3.18, so that the T–action on each $\mathcal{A}_n$ is rational.

A coherent $\mathcal{A}$–module $\mathscr{M}$ is T–equivariant if it comes equipped with a T–equivariant structure making each $\mathscr{M}_n$ a T–rational $\mathcal{A}_n$–module. The category of T–equivariant coherent $\mathcal{A}$–modules is $(\mathcal{A},\mathsf{T})\text{–}\mathsf{coh}$.

Proposition 3.23 *Assume that $\mathfrak{X}$ is affine. Then $\Gamma(\mathfrak{X},-)$ defines an exact equivalence between $(\mathcal{A},\mathsf{T})\text{–}\mathsf{coh}$ and the category $(A,\mathsf{T})\text{–}\mathsf{mod}$ of finitely generated T–equivariant A–modules, where $A=\Gamma(\mathfrak{X},\mathcal{A})$. A quasi-inverse is given by $M\mapsto M^{\Delta}$. Moreover, $M^{\Delta}=\mathcal{A}\otimes_A M$.*

Proof This is immediate from Proposition 3.11. □

3.6 Support

Let $\mathscr{M}$ be an $\mathcal{A}$–module. Then $\operatorname{Supp}\mathscr{M}$ denotes the sheaf-theoretic support of $\mathscr{M}$, ie it is the set of all points $x\in\mathfrak{X}$ such that $\mathscr{M}_x\neq 0$.

Lemma 3.24 *Let $\mathscr{M}$ be a coherent $\mathcal{A}$–module. Then $\operatorname{Supp}\mathscr{M}=\operatorname{Supp}\mathscr{M}/\hbar\mathscr{M}$. In particular, it is a closed subvariety of $\mathfrak{X}$.*

Proof Since both notions of support are local, we may assume that $\mathfrak{X}$ is affine and set $M=\Gamma(\mathfrak{X},\mathscr{M})$, $A=\Gamma(\mathfrak{X},\mathcal{A})$ and $R=A/\hbar A$.

Claim 3.25 *Let $f\in R$. Then $\Gamma(D(f),\mathscr{M})=Q_f^{\mu}(M)$.*

Proof As noted in Section 3.2, it follows from [68, Theorem 4.2] that the claim is true when $\mathscr{M}=\mathcal{A}$. Since M is finitely generated we may, by Lemma 3.14, fix a finite presentation $A^n\to A^m\to M\to 0$ of M. Then the claim follows from the fact that $Q_f^{\mu}(-)$ is exact on finitely generated A–modules and the five lemma applied to the diagram:

$$\begin{tikzcd}
Q_f^{\mu}(A^n) \arrow[r] \arrow[d, "\wr"'] & Q_f^{\mu}(A^m) \arrow[r] \arrow[d, "\wr"'] & Q_f^{\mu}(M) \arrow[r] \arrow[d] & 0 \\
\Gamma(D(f),\mathcal{A}^n) \arrow[r] & \Gamma(D(f),\mathcal{A}^m) \arrow[r] & \Gamma(D(f),\mathscr{M}) \arrow[r] & 0
\end{tikzcd}$$

This completes the proof of the claim. □

For each $f\in R$, the short exact sequence $0\to\hbar M\to M\to M/\hbar M\to 0$ gives

$$0\to Q_f^{\mu}(\hbar M)\to Q_f^{\mu}(M)\to Q_f^{\mu}(M/\hbar M)\to 0.$$

Therefore, $Q_f^{\mu}(M/\hbar M)\neq 0$ implies that $Q_f^{\mu}(M)\neq 0$. On the other hand, if $Q_f^{\mu}(M)\neq 0$ then

$$Q_f^{\mu}(\hbar M)=Q_f^{\mu}(A)\otimes_A\hbar M=\hbar Q_f^{\mu}(A)\otimes_A M=\hbar Q_f^{\mu}(M).$$

Since $Q_f^{\mu}(M)$ is a finitely generated $Q_f^{\mu}(A)$–module and $Q_f^{\mu}(A)$ is $\hbar$–adically complete, Nakayama's lemma implies that $\hbar Q_f^{\mu}(M)$ is a proper submodule of $Q_f^{\mu}(M)$. Thus, $\Gamma(D(f), \mathscr{M}) \neq 0$ if and only if $\Gamma(D(f), \mathscr{M}/\hbar\mathscr{M}) \neq 0$. The lemma follows. □

3.7 $\mathcal{W}$–algebras and good modules

Let $\mathcal{A}$ be a T–equivariant DQ algebra on $\mathfrak{X}$. Then $\mathcal{W} := \mathcal{A}[\hbar^{-1}]$ is a sheaf of $\mathbb{C}((\hbar))$–algebras on $\mathfrak{X}$; it is the $\mathcal{W}$–algebra associated to $\mathcal{A}$. Base change defines a functor $(\mathcal{A}, \mathsf{T})\text{–}\mathsf{coh} \to (\mathcal{W}, \mathsf{T})\text{–}\mathsf{mod}$, $\mathscr{M} \mapsto \mathscr{M}[\hbar^{-1}]$. Let $U \subset \mathfrak{X}$ be a T–stable open subset and $\mathscr{M}$ be a $\mathcal{W}_U$–module. A *lattice* for $\mathscr{M}$ is a coherent $\mathcal{A}_U$–submodule $\mathscr{M}'$ of $\mathscr{M}$ such that $\mathscr{M}'[\hbar^{-1}] = \mathscr{M}$; it is a T–lattice if it is a T–equivariant coherent module. The category of all T–equivariant $\mathcal{W}_U$–modules that admit a (global) T–lattice is denoted by $(\mathcal{W}_U, \mathsf{T})\text{–}\mathsf{good}$, and we refer to a module in this category as a *good* (T–equivariant) $\mathcal{W}_U$–module. Recall that $(\mathcal{A}_U, \mathsf{T})\text{–}\mathsf{coh}_{\mathrm{tor}}$ denotes the full subcategory of $(\mathcal{A}_U, \mathsf{T})\text{–}\mathsf{coh}$ consisting of all $\hbar$–torsion sheaves.

Proposition 3.26 *The category $(\mathcal{A}_U, \mathsf{T})\text{–}\mathsf{coh}_{\mathrm{tor}}$ is a Serre subcategory of $(\mathcal{A}_U, \mathsf{T})\text{–}\mathsf{coh}$ and we have an equivalence of abelian categories*

$$(\mathcal{A}_U, \mathsf{T})\text{–}\mathsf{coh}/(\mathcal{A}_U, \mathsf{T})\text{–}\mathsf{coh}_{\mathrm{tor}} \simeq (\mathcal{W}_U, \mathsf{T})\text{–}\mathsf{good}.$$

3.8 Restriction and quotient categories

Let $\mathfrak{X}$ be a smooth symplectic manifold with $\mathbb{G}_m$–action of positive weight. Let $Z \subset \mathfrak{X}^{\mathbb{G}_m}$ be a closed, connected and smooth subvariety. Let $C = \{x \in \mathfrak{X} \mid \lim_{t \to \infty} t \cdot x \in Z\}$ be the attracting locus for Z; it is a smooth, locally closed subvariety of $\mathfrak{X}$.

Assume that C is closed in $\mathfrak{X}$. The complement to C in $\mathfrak{X}$ is denoted by U and we write $j \colon U \hookrightarrow \mathfrak{X}$ for the embedding. In this section we prove the following:

Theorem 3.27 *Suppose that $C \subset \mathfrak{X}$ is closed and let $U = \mathfrak{X} \smallsetminus C$. The functor j^* induces an equivalence*

$$\mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}/\mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}_C \xrightarrow{\sim} \mathcal{W}_U\text{–}\mathsf{good}. \tag{3-1}$$

Proof The remainder of this section is devoted to the proof of Theorem 3.27. First we note an immediate corollary. The Ind-category of $(\mathcal{W}, \mathsf{T})\text{–}\mathsf{good}$ is denoted by $\mathsf{Qcoh}(\mathcal{W}, \mathsf{T})$, or $\mathsf{Qcoh}(\mathcal{W})$ if T is understood from context; we call it (abusively) the *quasicoherent* category. We say a quasicoherent object has support in a closed subset $K \subset \mathfrak{X}$ if one can write $\mathscr{M} = \varinjlim \mathscr{M}_i$ where each $\mathscr{M}_i$ is good and has support in K. We write $\mathsf{Qcoh}(\mathcal{W})_K$ for the full subcategory whose objects have support in K.

Recall (for example, from [41]) that a full subcategory of a Grothendieck category is called *localizing* if it is closed under subobjects, quotients, extensions and small inductive limits.

Corollary 3.28 *Let $C \subset \mathfrak{X}$ be a closed subset as above. The functor*

$$j^*\colon \mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}}, \mathsf{T}) \to \mathsf{Qcoh}(\mathcal{W}_U, \mathsf{T})$$

is essentially surjective and induces an equivalence

$$\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}}, \mathsf{T})/\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}}, \mathsf{T})_C \simeq \mathsf{Qcoh}(\mathcal{W}_U).$$

Moreover, j^* admits a right adjoint.

Definition 3.29 We write

$$j_*\colon \mathsf{Qcoh}(\mathcal{W}_U) \to \mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}})$$

to denote the right adjoint of j^*, taking care to note that it *need not* be identified with the sheaf-theoretic direct image.

Proof of Corollary 3.28 Essential surjectivity and equivalence are immediate from the theorem. The existence of a right adjoint follows since the kernel of j^*, ie the subcategory of modules supported on U, is a localizing subcategory. □

We begin the proof of Theorem 3.27. The main part of the proof will show that the faithful functor $\mathcal{W}_{\mathfrak{X}}\text{–good}/\mathcal{W}_{\mathfrak{X}}\text{–good}_C \to \mathcal{W}_U\text{–good}$ is full.

We begin with a lemma that will allow us to reduce to affine statements.

Lemma 3.30 *Suppose that $Z \subseteq \mathfrak{X}^{\mathbb{G}_m}$ is connected and closed, and that*

$$C = \{x \in \mathfrak{X} \mid \lim_{t\to\infty} t\cdot x \in Z\}$$

is closed in $\mathfrak{X}$. If $U \subseteq \mathfrak{X}$ is an affine open subset of $\mathfrak{X}$ for which $Z \cap U \neq \varnothing$, then

$$C \cap U = \{x \in U \mid \lim_{t\to\infty} t\cdot x \in Z \cap U\}.$$

In particular, $\lim_{t\to\infty} t\cdot x$ exists in U for every $x \in C \cap U$.

Proof Note that $Z^\circ = Z \cap U$ is open in Z. Let

$$C^\circ = \{x \in U \mid \lim_{t\to\infty} t\cdot x \in Z^\circ\}.$$

Then C° is the preimage, under the projection morphism $C \to Z$, of the dense open set Z°; hence C° is dense in C. Suppose that $f \in \mathbb{C}[U]$ is a $\mathbb{G}_m$–semi-invariant, say $f(t\cdot x) = t^{-d} f(x)$ for all $t \in \mathbb{G}_m$, $x \in U$. If $x \in C^\circ$, then $\lim_{t\to 0} t^d f(x) = f(\lim_{t\to\infty} t\cdot x)$, so if $d < 0$ then $f(x) = 0$. Thus any $f \in \mathbb{C}[U]$ of negative weight vanishes on C° and consequently (by density) vanishes on $C \cap U$. It follows that $\mathbb{C}[C \cap U]$ has nonnegative $\mathbb{G}_m$–weights; since $C \cap U$ is closed in U, hence affine, we conclude that the $\mathbb{G}_m$–action on $C \cap U$ extends to an action of the monoid $\mathbb{A}^1$ on $C \cap U$, proving the lemma. □

Returning to the proof of the theorem, we claim that, if we assume fullness of (3-1), essential surjectivity follows from Corollary 3.16. Indeed, to prove essential surjectivity, it suffices to replace $\mathfrak{X}$ by any $\mathbb{G}_m$–stable open subset of $\mathfrak{X}$ that contains C. Thus, choose a collection $\{\mathfrak{X}_i\}$ of $\mathbb{G}_m$–stable affine open subsets of $\mathfrak{X}$ whose union contains C. Then, by Lemma 3.30, $C \cap \mathfrak{X}_i \subset \mathfrak{X}_i$ is a closed subset satisfying the hypotheses of the theorem, and so the fullness assertion holds for restriction from $\mathfrak{X}_i$ to $U_i = \mathfrak{X}_i \smallsetminus C$. Corollary 3.16 thus implies that for every coherent $\mathcal{A}_{U_i}$–module $\mathscr{M}_i$, there is a coherent $\mathcal{A}_{\mathfrak{X}_i}$–module $\overline{\mathscr{M}}_i$ and an isomorphism $\overline{\mathscr{M}}_i|_{U_i} \cong \mathscr{M}_i$. A standard gluing argument then shows that every coherent $\mathcal{A}_U$–module extends to a coherent $\mathcal{A}$–module, proving essential surjectivity.

Thus, we return to the proof of fullness of (3-1). We note that taking a covering of $\mathfrak{X}$ by affine open $\mathbb{G}_m$–stable sets, the sheaf property implies that the fullness statement is local. Therefore we may assume that $\mathfrak{X}$ is affine. Shrinking $\mathfrak{X}$ if necessary, we may assume that $C = Z(f_1, \dots, f_k)$ is a complete intersection in $\mathfrak{X}$ of codimension k, where each f_i is homogeneous with respect to $\mathbb{G}_m$. As in Lemma 3.30, if $f \in \mathcal{O}(\mathfrak{X})$ is homogeneous of negative weight with respect to $\mathbb{G}_m$, then $f \in I(C)$. The fact that $\mathfrak{X}$ is affine implies that we can (and will) fix an identification $\mathcal{A} = \mathcal{O}_{\mathfrak{X}}[\![\hbar]\!]$ of prorational sheaves of $\mathbb{C}[\![\hbar]\!]$–modules. Notice that for any $\mathbb{G}_m$–stable affine open subset V of $\mathfrak{X}$, the identification gives a canonical identification $\mathcal{A}(V) = \mathcal{O}(V)[\![\hbar]\!]$. Given $f \in \mathcal{O}(\mathfrak{X})$, let $\nu(f)$ denote the corresponding section of $\mathcal{A}(\mathfrak{X})$ under this identification.

Let W_{rat} denote the $\mathbb{C}[\hbar, \hbar^{-1}]$–subalgebra of $\mathbb{G}_m$–rational sections in $\Gamma(\mathfrak{X}, \mathcal{W}_{\mathfrak{X}})$. Given a $\mathbb{G}_m$–equivariant W–module M, let M_{rat} denote the W_{rat}–submodule of all rational sections. We say that a W_{rat}–module M is supported on C if, for each section $m \in M$, there exists $N \gg 0$ such that $\nu(f_i)^N \cdot m = 0$ for all i.

Lemma 3.31 *Suppose $\mathfrak{X}$ is affine.*

1. *The functor $\mathscr{M} \mapsto \Gamma(\mathfrak{X}, \mathscr{M})_{\mathrm{rat}}$ is an equivalence of categories $R\colon \mathcal{W}_{\mathfrak{X}}\text{–good} \xrightarrow{\sim} W_{\mathrm{rat}}\text{–mod}$.*
2. *Under the equivalence of* (1), *$\mathscr{M}$ is supported on C (in the usual sense) if and only if $\Gamma(\mathfrak{X}, \mathscr{M})_{\mathrm{rat}}$ is supported on C (in the above sense).*

Proof (1) This follows from Propositions 3.23 and 3.26 by a standard argument (see the proof of Proposition 4.33).

(2) If $\mathscr{M}_0$ is a coherent $\mathcal{A}$–submodule of $\mathscr{M}$ such that $\mathscr{M} = \mathscr{M}_0[\hbar^{-1}]$, then $M = M_0[\hbar^{-1}]$, where $M = \Gamma(\mathfrak{X}, \mathscr{M})_{\mathrm{rat}}$ and $M_0 = \Gamma(X, \mathscr{M}_0)_{\mathrm{rat}}$. Certainly, if $m \in M_0$ and $\nu(f_i)^N \cdot m = 0$ then $f_i^N \cdot \bar{m} = 0$ in $M_0/\hbar M_0$. Hence $\mathscr{M}$ is supported on C in the usual sense.

We need to check the converse. So assume that $\mathscr{M}$ is supported on C in the usual sense. Our assumptions on C imply that $\mathcal{O}(\mathfrak{X}) = \mathbb{C}[f_1, \dots, f_k, x_1, \dots, x_l]/I$, where $\deg x_i \geq 0$ and I is a homogeneous ideal. Hence $\mathbb{C}[C]$, a quotient of the algebra $\mathbb{C}[x_1, \dots, x_l]$, is nonnegatively graded. This implies that the finitely generated $\mathbb{C}[\mathfrak{X}]$–module $M_0/\hbar M_0$ has its grading bounded from below. Since $\hbar$ has positive weight, the same applies to M_0. Let $m \in M_0$ be a homogeneous section. If $f_i^N \cdot \bar{m} = 0$ in $M_0/\hbar M_0$, then $\nu(f_i)^N \cdot m \in \hbar M_0$ and hence $\nu(f_i)^{rN} \cdot m \in \hbar^r M_0$. On the other hand, $\deg f_i \leq 0$ and hence $\deg(\nu(f_i)^{rN} \cdot m) = rN \deg \nu(f_i) + \deg m \leq \deg(m)$. This implies that $\nu(f_i)^{rN} \cdot m = 0$ for $r,\ N \gg 0$, since the weights of all homogeneous elements in $\hbar^r M_0$ will be greater than $\deg m$ for $r \gg 0$. □

Write $U_\alpha = \mathfrak{X} \smallsetminus Z(f_\alpha)$ and $U_{\alpha_0,\dots,\alpha_i} = U_{\alpha_0} \cap \dots \cap U_{\alpha_i}$. Given $\mathscr{M} \in \mathcal{W}_U$–good, define the complex

$$\check{C}^i(\mathscr{M}) = \prod_{\alpha_0 < \dots < \alpha_i} \Gamma(U_{\alpha_0,\dots,\alpha_i}, \mathscr{M}),$$

with the usual differential

$$d^i \colon \check{C}^i(\mathscr{M}) \to \check{C}^{i+1}(\mathscr{M}), \quad d^i(f_{\alpha_0,\dots,\alpha_i})_{\alpha_0,\dots,\beta,\dots,\alpha_i} = (-1) f_{\alpha_0,\dots,\alpha_i}|_{U_{\alpha_0,\dots,\beta,\dots,\alpha_i}}.$$

For $\mathscr{M} \in \mathcal{W}_U$–good, we define $F(\mathscr{M}) = \check{C}^\bullet(\mathscr{M})_{\mathrm{rat}}$. There is a canonical transformation $R \to F(-|_U)$, where we identify W_{rat}–Mod with complexes concentrated in degree zero.

Lemma 3.32 $\mathrm{Cone}(R \to F(-|_U))$ *defines an exact functor from* $\mathcal{W}_U$–good *to complexes with terms in* W_{rat}–Mod.

Proof The cone $\mathrm{Cone}(R \to F(-|_U))$ is exact if and only if both R and $F(-|_U)$ are exact. The functor R is exact by Lemma 3.31. Therefore it suffices to show that the functor F defines an exact functor from $\mathcal{W}_U$–good to complexes with terms in W_{rat}–Mod. The exactness of $\mathscr{M} \mapsto \check{C}^\bullet(\mathscr{M})$ can be checked term by term. But it is clear that $\mathscr{M} \mapsto \check{C}^i(\mathscr{M})$ is exact because the open set $U_{\alpha_0,\dots,\alpha_i}$ is affine. Therefore, to show that F is exact, it suffices to show that the functor $\mathcal{W}_{U_{\alpha_0,\dots,\alpha_i}}$–good $\to W_{\mathrm{rat}}$–Mod, $\mathscr{M} \mapsto \Gamma(U_{\alpha_0,\dots,\alpha_i}, \mathscr{M})_{\mathrm{rat}}$ is an exact functor. Since $U_{\alpha_0,\dots,\alpha_i}$ is affine, this follows from Lemma 3.31(1). □

Lemma 3.33 *The cohomology of the cone* $D^\bullet = \mathrm{Cone}(W_{\mathrm{rat}} \to F(\mathcal{W}_U))$ *is zero outside degree* k. *The group* $H^k(D^\bullet)$ *is supported on* C.

Proof Notice that the differentials are $\mathbb{C}[\![\hbar]\!]$–linear and preserve the lattice defined by $\mathcal{A}$. Therefore $D^\bullet = D_0^\bullet[\hbar^{-1}]$ and flat base change implies that $H^i(D^\bullet) = H^i(D_0^\bullet)[\hbar^{-1}]$. Thus, it suffices to check that the corresponding statements hold for $D_0^\bullet$. Write $D_{\mathrm{com}}^\bullet$ for $\mathrm{Cone}(\mathcal{O}(X) \to \check{C}^\bullet(\mathcal{O}_U))$. Since we have an identification $\mathcal{A} = \mathcal{O}_X[\![\hbar]\!]$ of prorational sheaves, we have an identification of complexes $D_0^\bullet = D_{\mathrm{com}}^\bullet[\![\hbar]\!]$ and hence $H^i(D_0^\bullet) = H^i(D_{\mathrm{com}}^\bullet)[\![\hbar]\!]$, as prorational $\mathbb{C}[\![\hbar]\!]$–modules. Since $H^i(D_{\mathrm{com}}^\bullet) = 0$ for $i \neq 0$, the first claim follows.

Claim 3.34 *The space $H^k(D_{\mathrm{com}}^\bullet)$ is nonnegatively graded as a $\mathbb{G}_m$–module.*

Proof As in the proof of Lemma 3.31, $\mathbb{C}[\mathfrak{X}] = [f_1, \dots, f_k, x_1, \dots, x_l]/I$ for some homogeneous ideal I. Then

$$H^k(D_{\mathrm{com}}^\bullet) = \mathbb{C}[\mathfrak{X}]_{f_1 \cdots f_k} \Big/ \sum_i \mathbb{C}[\mathfrak{X}]_{f_1 \cdots \hat{f}_i \cdots f_k}$$

is a quotient of $\mathbb{C}[f_1, \dots, f_k, x_1, \dots, x_l]_{f_1 \cdots f_k} / \sum_i \mathbb{C}[f_1, \dots, f_k, x_1, \dots, x_l]_{f_1 \cdots \hat{f}_i \cdots f_k}$. The latter clearly has the desired properties. This proves the claim. □

We return to the proof of Lemma 3.33. To prove the second assertion of the lemma, we may assume that m belongs to $H^k(D_0^\bullet)$ and is rational. The image of m in $H^k(D_{\mathrm{com}}^\bullet) = H^k(D_0^\bullet)/\hbar H^k(D_0^\bullet)$ is torsion with respect to the f_i. Therefore there exists N such that $\nu(f_i)^N \cdot m \in \hbar H^k(D_0^\bullet)$ and hence $\nu(f_i)^{rN} \cdot m \in \hbar^r H^k(D_0^\bullet)$. Claim 3.34 implies that the weight of every homogeneous element in $H^k(D_0^\bullet)$ is nonnegative. Thus, every homogeneous element in $\hbar^r H^k(D_0^\bullet)$ has degree at least r. On the other hand $\deg f_i \leq 0$ and hence $\deg(\nu(f_i)^N \cdot m) = N \deg \nu(f_i) + \deg m \leq \deg m$. This implies that $\nu(f_i)^N \cdot m = 0$ for $N \gg 0$. □

Proposition 3.35 *The cohomology of* $\mathrm{Cone}(R(\mathscr{M}) \to F(\mathscr{M}|_U))$ *is supported on C for any $\mathscr{M} \in \mathcal{W}_{\mathfrak{X}}$–good.*

Proof Since $\mathcal{W}_{\mathfrak{X}}$–good has finite homological dimension, we prove the claim by induction on projective dimension. Certainly Lemma 3.33 implies that the claim holds for every summand of $\mathcal{W}_{\mathfrak{X}}^N$ (for N finite). By Lemma 3.32, the functor $\mathrm{Cone}(R \to F(-|_U))$ is exact. Therefore the long exact sequence in cohomology implies that if it holds for all modules of projective dimension i, then it also holds for all modules of projective dimension $i+1$. □

Finally, we show that the faithful functor $\mathcal{W}_{\mathfrak{X}}\text{–good}/\mathcal{W}_{\mathfrak{X}}\text{–good}_C \to \mathcal{W}_U\text{–good}$ is also full. This will complete the proof of Theorem 3.27. Thus, suppose $\mathscr{M}, \mathscr{N} \in \mathcal{W}_{\mathfrak{X}}$–good and $\phi \in \mathrm{Hom}_{\mathcal{W}_U\text{–good}}(\mathscr{M}|_U, \mathscr{N}|_U)$. Applying the functors R and F from above, we get a diagram:

$$\begin{array}{ccc} R(\mathscr{M}) & & R(\mathscr{N}) \\ \downarrow & & \downarrow \\ H^0(F(\mathscr{M}|_U)) & \xrightarrow{H^0(F(\phi))} & H^0(F(\mathscr{N}|_U)) \end{array}$$

By Proposition 3.35, the vertical arrows have kernel and cokernel supported on C. Letting N' be any finitely generated W_{rat}–submodule of $H^0(F(\mathscr{N}|_U))$ that contains the images of $R(\mathscr{M})$ and $R(\mathscr{N})$, we get a diagram

$$R(\mathscr{M}) \xrightarrow{\tilde{\phi}'} N' \leftarrow R(\mathscr{N}),$$

where $R(\mathscr{N}) \to N'$ has kernel and cokernel supported on C. Applying R^{-1}, we get a diagram of good $\mathcal{W}_{\mathfrak{X}}$–modules $\mathscr{M} \xrightarrow{\phi'} \mathscr{N}' \leftarrow \mathscr{N}$, where the support assertion guarantees that $\mathscr{N} \to \mathscr{N}'$ becomes an isomorphism in the quotient category. Thus, ϕ' defines a morphism $\mathscr{M} \to \mathscr{N}$ in the quotient category, whose restriction to U is, by construction, identified with ϕ. This proves Theorem 3.27. □

3.9 Holonomic modules

By Gabber's theorem, the support of any good $\mathcal{W}$–module has dimension at least $\frac{1}{2}\dim \mathfrak{X}$. A good $\mathcal{W}$–module is said to be holonomic if the dimension of its support is exactly $\frac{1}{2}\dim \mathfrak{X}$. The category of holonomic $\mathcal{W}$–modules is denoted by $\mathcal{W}\text{–}\mathsf{hol}$. The theory of characteristic cycles implies:

Lemma 3.36 *Let $\mathscr{M}$ be a holonomic $\mathcal{W}$–module. Then $\mathscr{M}$ has finite length.*

A $\mathcal{W}$–module $\mathscr{M}$ is said to be *regular holonomic* if there exists a lattice $\mathscr{M}'$ of $\mathscr{M}$ such that the support of $\mathscr{M}'/\hbar\mathscr{M}'$ is reduced. The category of regular holonomic $\mathcal{W}$–modules is denoted by $\mathcal{W}\text{–}\mathsf{reghol}$.

3.10 Equidimensionality of supports

In this subsection we note that the analogue of the Gabber–Kashiwara equidimensionality theorem holds for $\mathcal{W}$–algebras. First, given a coherent $\mathcal{A}$–module $\mathscr{M}$, the sheaf $\bigoplus_{n\geq 0} \hbar^n \mathscr{M}/\hbar^{n+1}\mathscr{M}$ is a coherent $\mathcal{O}_{\mathfrak{X}}[\hbar]$–module. Therefore, its support is a closed subvariety of $\mathfrak{X}\times\mathbb{A}^1$, which we will denote by $\widetilde{\mathrm{Supp}}\,\mathscr{M}$.

Lemma 3.37 *Let $p\colon \mathfrak{X}\times\mathbb{A}^1 \to \mathfrak{X}$ be the projection map. Then* $\mathrm{Supp}\,\mathscr{M} = p(\widetilde{\mathrm{Supp}}\,\mathscr{M})$.

Proof Since $\mathscr{M}$ is coherent and support is a local property, we may assume that $\mathfrak{X}$ is affine and set $M = \Gamma(\mathfrak{X}, \mathscr{M})$, an $A = \Gamma(\mathfrak{X}, \mathcal{A})$–module. Both Supp and $\widetilde{\mathrm{Supp}}$ are additive on short exact sequences, therefore the sequence $0 \to M_{\mathrm{tor}} \to M \to M_{\mathrm{tf}} \to 0$ implies that we may assume that M is either $\hbar$–torsion or $\hbar$–torsion-free. First, if M is $\hbar$–torsion-free, then $M \simeq (M/\hbar M)\,\widehat{\otimes}\,\mathbb{C}[\![\hbar]\!]$ as a $\mathbb{C}[\![\hbar]\!]$–module. This implies that $\widetilde{\mathrm{Supp}}\mathscr{M} = \mathrm{Supp}\mathscr{M} \times \mathbb{A}^1$. On the other hand, if M is $\hbar$–torsion, then $\mathrm{gr}\, M := \bigoplus_{n\geq 0} \hbar^n M/\hbar^{n+1} M$ equals $\bigoplus_{n=0}^{N} \hbar^n M/\hbar^{n+1} M$ for some N. If $I \cdot (M/\hbar M) = 0$ for some $I \lhd \mathbb{C}[\mathfrak{X}]$, then $I \cdot \mathrm{gr}\, M = 0$. Since $\hbar^{N+1}\,\mathrm{gr}\, M = 0$ too, this implies that $\widetilde{\mathrm{Supp}}\,\mathscr{M} = \mathrm{Supp}\mathscr{M} \times \{0\}$. □

Let $\dim \mathfrak{X} = 2m$. The analogue of the Gabber–Kashiwara theorem reads:

Theorem 3.38 *Let $\mathscr{M}$ be a good $\mathcal{W}$–module. Then there exists a unique filtration*

$$0 = D_{m-1}(\mathscr{M}) \subset D_m(\mathscr{M}) \subset D_{m+1}(\mathscr{M}) \subset \cdots \subset D_{2m}(\mathscr{M}) = \mathscr{M}$$

such that $\mathrm{Supp}\, D_i(\mathscr{M})/D_{i-1}(\mathscr{M})$ is pure i–dimensional.

Proof We fix a lattice $\mathscr{M}'$ of $\mathscr{M}$, it is a coherent torsion-free $\mathcal{A}$–module. We will show that the analogue of the above statement holds for $\mathscr{M}'$, then set $D_i(\mathscr{M}) = D_i(\mathscr{M}')$ and check that $D_i(\mathscr{M})$ is independent of the choice of lattice. The uniqueness property actually implies that the statement will hold globally on $\mathfrak{X}$ if it holds locally, therefore we may as well assume that $\mathfrak{X}$ is affine and set $M' = \Gamma(\mathfrak{X}, \mathscr{M}')$. Since M is $\hbar$–torsion-free, the proof of Lemma 3.37 shows that $\widetilde{\mathrm{Supp}}\mathscr{M} = \mathrm{Supp}\mathscr{M} \times \mathbb{C}$. Then, noting this fact, the required result is [24, Theorem V8, page 342].

Therefore, it suffices to show that $D_i(\mathscr{M})$ is independent of the choice of lattice. Let M'' be another choice of lattice and $N \gg 0$ such that $\hbar^N M'' \subset M'$. As explained in [24, Section 1], $D_i(M') := \{m \in M' \mid \dim \mathrm{gr}(A\cdot m) \leq i+1\}$. Therefore, if $m \in D_i(M'')$ then $\hbar^N m \in D_i(M')$, which implies that $\hbar^N D_i(M'') \subset D_i(M')$. By symmetry, we have $\hbar^{N'} D_i(M') \subset D_i(M'')$ and hence $D_i(M'')[\hbar^{-1}] = D_i(M')[\hbar^{-1}]$. □

As a corollary of the theorem, we can strengthen our extension result. Again, let T be a torus.

Corollary 3.39 *Let U be a T–stable open subset of $\mathfrak{X}$ whose complement is a union of coisotropic cells and $\mathscr{M}$ a holonomic, T–equivariant $\mathcal{W}|_U$–module. Then, there exists a holonomic, T–equivariant $\mathcal{W}$–module $\mathscr{M}'$ such that $\mathscr{M}'|_U \simeq \mathscr{M}$. Moreover, if $\mathscr{M}$ is simple then there exists a unique simple extension $\mathscr{M}'$.*

Proof Let $\mathscr{N}$ be a lattice of $\mathscr{M}$. By Theorem 3.27, there exists a coherent T–equivariant $\mathcal{A}$–module $\mathscr{N}'$ such that $\mathscr{N}'|_U = \mathscr{N}$. Replacing $\mathscr{N}'$ by its torsion-free quotient, we may assume that $\mathscr{N}'$ is torsion-free and set $\mathscr{M}'' = \mathscr{N}'[\hbar^{-1}]$. Then, let $\mathscr{M}' = D_m(\mathscr{M}'')$, a holonomic submodule of $\mathscr{M}''$. Since $\mathscr{M}$ is holonomic, $D_m(\mathscr{M}) = \mathscr{M}$. The uniqueness of $D_m(-)$ implies that $\mathscr{M}'|_U = D_m(\mathscr{M}'')|_U = D_m(\mathscr{M}'|_U) = \mathscr{M}$.

If $\mathscr{M}$ is assumed to be simple then there is some simple subquotient of $\mathscr{M}'$ whose restriction to U is isomorphic to $\mathscr{M}$. In order to show uniqueness of the extension, let

$\mathscr{M}^1$ and $\mathscr{M}^2$ be two simple extensions of $\mathscr{M}$. Denote by j the open embedding $U \hookrightarrow \mathfrak{X}$ and by j_* the right adjoint to j^* whose existence is established in Corollary 3.28. Then for each $i = 1, 2$ the canonical adjunction $\mathscr{M}^i \to j_*\mathscr{M}$ is an embedding because it is an isomorphism over U (and hence nonzero) and $\mathscr{M}^i$ is assumed to be simple. Therefore $\mathscr{M}^1 \cap \mathscr{M}^2$ is a $\mathcal{A}$–submodule of $\mathscr{M}^i$ whose restriction to U is $\mathscr{M}$. Thus, $\mathscr{M}^1 = \mathscr{M}^2$. □

Remark 3.40 Let $\mathscr{M}$ be a simple $\mathcal{W}$–module or a primitive quotient of $\mathcal{W}$. Then Theorem 3.38 implies that the support of $\mathscr{M}$ is equidimensional. A proof of the analogous result in the setting of localization via $\mathbb{Z}$–algebras was recently given by Gordon and Stafford [27].

4 Quantum coisotropic reduction

In this section we consider the process of quantum coisotropic reduction. Our main result is that quantum coisotropic reduction can be used to prove an analogue of Kashiwara's equivalence for DQ modules supported on a coisotropic stratum. At the end of the section we consider $\mathcal{W}$–algebras on a symplectic manifold $\mathfrak{X}$ with an elliptic $\mathbb{G}_m$–action such that $Y = \mathfrak{X}^{\mathbb{G}_m}$ is connected. We show that $\mathsf{Qcoh}(\mathcal{W})$ is equivalent to the category of quasicoherent sheaves for a sheaf of filtered $\mathcal{O}_Y$–algebras quantizing $\mathfrak{X}$. These filtered $\mathcal{O}_Y$–algebras behave much like the sheaf of differential operators on Y.

Notation 4.1 Throughout the remainder of the paper, $\mathcal{W}$–good and $\mathsf{Qcoh}(\mathcal{W})$ will denote the category of good, $\mathbb{G}_m$*–equivariant* $\mathcal{W}$–modules and the Ind-category of good, $\mathbb{G}_m$*–equivariant* $\mathcal{W}$–modules, respectively. Moreover, all $\mathcal{W}$–modules that we consider will be assumed to be $\mathbb{G}_m$–equivariant.

4.1 Quantum coisotropic reduction: local case

We maintain the notation and conventions of Sections 2 and 3.

Thus, let R be a regular affine $\mathbb{C}$–algebra equipped with a Poisson structure $\{-,-\}$ making $\mathfrak{X} = \operatorname{Spec}(R)$ into a smooth affine symplectic variety. We assume in addition that R comes equipped with a $\mathbb{G}_m$–action for which the Poisson structure has weight $-l$.

Let A be a deformation-quantization of R, equipped with an action of $\mathbb{G}_m$ that preserves the central subalgebra $\mathbb{C}[\![\hbar]\!] \subset A$ and has $\hbar$ as a weight vector. The quotient map $\rho\colon A \to R$ is equivariant. The fact that the Poisson bracket $\{-,-\}$ on R is graded of degree $-l$ implies that $\hbar$ has weight l. Let J be the left ideal generated by all homogeneous elements in A of negative degree.

As in Section 2.4 we have $I = \rho(J)$, the ideal in R generated by all homogeneous elements of strictly negative degree. We write $C = \operatorname{Spec}(R/I)$, a closed coisotropic subvariety of $\mathfrak{X}$. We write $Y = C^{\mathbb{G}_m}$, the $\mathbb{G}_m$–fixed locus of the coisotropic subset C. As in Section 2.4, we also assume that Y has been shrunk suitably so that the affine bundle $\rho\colon C \to Y$ of [10] splits equivariantly as

$$\phi\colon C \xrightarrow{\sim} Y \times V \times Z,$$

where V is isomorphic to a vector space with $\mathbb{G}_m$–weights lying in $[-l,-1]$ and Z is isomorphic to a vector space with $\mathbb{G}_m$–weights less than $-l$.

Notation 4.2 Let $\widehat{R}$ denote the completion of R with respect to the ideal I, and $\widehat{A}$ the completion of A with respect to the two-sided ideal $K := \rho^{-1}(I)$.

Lemma 4.3 *Let $\widehat{A}$ be as above.*

(1) *The algebra $\widehat{A}$ is flat over A and Noetherian.*

(2) *If M is a finitely generated A–module then $\varprojlim(M/K^n \cdot M) \simeq \widehat{A} \otimes_A M$.*

(3) *The algebra $\widehat{A}$ is $\hbar$–adically free and $\widehat{A}/\hbar\widehat{A} \simeq \widehat{R}$.*

(4) *The algebra $\widehat{A}$ is an equivariant quantization of $\widehat{R}$ with the Poisson structure on $\widehat{R}$ induced from the Poisson structure on R.*

Lemma 4.3(1)–(2) have also been shown by Losev [48] using a different argument.

Proof Take $u_1 = \hbar$ and let $u_2, \dots$ be arbitrary lifts in A of a set of generators of I; these form a normalizing sequence of generators as defined in [52, Theorem 4.2.7], and hence by that theorem the ideal K satisfies the Artin–Rees property in A. Then the proofs of [21, Lemma 7.15 and Theorem 7.2b] apply also in the noncommutative case to imply that $\widehat{A}$ is flat over A. The fact that $\widehat{A}$ is Noetherian follows from (3), which implies that $\operatorname{gr}\widehat{A} \simeq \widehat{R}[\hbar]$, a Noetherian ring.

Therefore we need to establish that $\widehat{A}/\hbar\widehat{A} \simeq \widehat{R}$ and that $\widehat{A}$ is a complete, flat $\hbar$–module. By Lemma 3.2, and using the fact that inverse limits commute, we have

$$\begin{aligned}\widehat{A} = \lim_{\infty \leftarrow n} A/K^n &= \lim_{\infty \leftarrow n}\big(\lim_{\infty \leftarrow s}(A/K^n)/\hbar^s(A/K^n)\big) \\ &= \lim_{\infty \leftarrow s}\big(\lim_{\infty \leftarrow n}(A/K^n)/\hbar^s(A/K^n)\big),\end{aligned} \tag{4-1}$$

which implies that $\widehat{A}$ is $\hbar$–adically complete. By (1), $\widehat{A}$ is A–flat, hence a fortiori it is $\hbar$–flat. Consider the short exact sequence

$$0 \to \left\{\frac{\hbar A + K^n}{K^n}\right\}_n \to \{A/K^n\}_n \to \left\{\frac{A}{\hbar A + K^n}\right\}_n \to 0$$

of inverse systems. Since $(\hbar A + K^n)/K^n \to (\hbar A + K^{n-1})/K^{n-1}$ is surjective, the inverse system $\{(\hbar A + K^n)/K^n\}_n$ satisfies the Mittag-Leffler condition. Therefore, we have

$$\begin{aligned}\big(\lim_{\infty \leftarrow n} A/K^n\big)/\big(\lim_{\infty \leftarrow n} \hbar(A/K^n)\big) &= \big(\lim_{\infty \leftarrow n} A/K^n\big)/\big(\lim_{\infty \leftarrow n} (\hbar A + K^n)/K^n\big) \\ &= \lim_{\infty \leftarrow n} A/(K^n + \hbar A) = \lim_{\infty \leftarrow n} R/I^n,\end{aligned}$$

where the final equality follows from the fact that $\hbar \in K$ and hence $(K^n + \hbar A)/\hbar A$ equals I^n.

The only thing left to show in order to conclude that $\widehat{A}$ is a quantization of $\widehat{R}$ is that the Poisson bracket on $\widehat{R}$ coming from $\widehat{A}$ equals the Poisson bracket on $\widehat{R}$ coming from the fact that it is a completion of R. To see this, we note the following well-known properties of algebras:

Sublemma 4.4 (1) *Suppose that $A \to B$ is a filtered homomorphism of almost-commutative filtered algebras. Then the induced map $\mathrm{gr}(A) \to \mathrm{gr}(B)$ is a Poisson homomorphism.*

(2) *Suppose that $\phi\colon R \to S$ is a continuous homomorphism of topological algebras and $\phi(R)$ is dense in S. Suppose R is equipped with a Poisson structure. Then there is at most one continuous Poisson structure on S making ϕ a Poisson homomorphism.*

Applying Sublemma 4.4(1) to $A \to \widehat{A}$ shows that the Poisson structure on $\widehat{R}$ induced from $\widehat{A}$ makes $R \to \widehat{R}$ a Poisson homomorphism. Sublemma 4.4(2) then implies that this Poisson structure must agree with the Poisson structure on $\widehat{R}$ induced from R. Lemma 4.3(4) follows. □

4.2 Quantizations of the formal neighborhood of C

The total space of the normal bundle $\mathsf{N}_{\mathfrak{X}/C}$ has a canonical symplectic structure of weight l. Choosing homogeneous bases $\boldsymbol{z}$ and $\boldsymbol{w}$ as in Theorem 2.6, the symplectic form on $\mathsf{N}_{\mathfrak{X}/C}$ is given by $\omega_S + \sum_{i=1}^m dz_i \wedge dw_i$. We denote by $\mathcal{D}$ the Moyal–Weyl quantization of T^*V, and by $\widehat{\mathcal{D}}$ the Moyal–Weyl quantization of the ring of functions $\widehat{F}$ on the formal neighborhood $\widehat{T^*V}$ of V in T^*V.

We write $\mathfrak{C}$ for the formal neighborhood of C in $\mathsf{N}_{\mathfrak{X}/C}$; recall that, as in Section 2.4, this is equivariantly isomorphic (though noncanonically) to the formal neighborhood of C in $\mathfrak{X}$. Let $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}} = \varprojlim \Omega^1_{\mathfrak{C}_n}$ denote the sheaf of continuous one-forms on $\mathfrak{C}$, where $\mathfrak{C}_n$ is the n^{th} infinitesimal neighborhood of C. Similarly, let $\Theta_{\mathfrak{C}}$ and $\Theta^{\mathrm{cts}}_{\mathfrak{C}}$ denote the sheaf of vector fields and continuous vector fields, respectively, on $\mathfrak{C}$. We denote by Π the bivector that defines the Poisson bracket on $\mathfrak{C}$. Then $d = [\Pi, -]$ defines a differential on $\bigwedge^* \Theta_{\mathfrak{C}}$, where $[-,-]$ is the Schouten bracket on polyvector fields. The cohomology of $\bigwedge^* \Theta_{\mathfrak{C}}$ is the Poisson cohomology $H^*_{\Pi}(\mathfrak{C})$ of $\mathfrak{C}$. The algebraic de Rham complex of $\mathfrak{C}$ is denoted by $\Omega^{*,\mathrm{cts}}_{\mathfrak{C}}$.

Lemma 4.5 *Let $i\colon \mathfrak{C} \to \mathsf{N}_{\mathfrak{X}/C}$ be the canonical morphism.*

(1) *Every derivation of $\widehat{R}$ is continuous, ie $\Theta_{\mathfrak{C}} = \Theta^{\mathrm{cts}}_{\mathfrak{C}}$.*

(2) $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}} \simeq \Theta^*_{\mathfrak{C}} \simeq i^*\Omega_{\mathsf{N}_{\mathfrak{X}/C}}$.

(3) *The Poisson structure defines an isomorphism of complexes $\bigwedge^* \Theta_{\mathfrak{C}} \simeq \Omega^{*,\mathrm{cts}}_{\mathfrak{C}}$.*

Proof Let $\delta \in \mathrm{Der}(\widehat{R})$. From the definition of a derivation, $\widehat{I}^{n+1} \subset \delta^{-1}(\widehat{I}^n)$. Since the translates of the powers of $\widehat{I}$ are a base of the topology, it follows that the preimage of $\widehat{I}^n$ is open in $\widehat{R}$. Hence δ is continuous and $\Theta_{\mathfrak{C}} = \Theta^{\mathrm{cts}}_{\mathfrak{C}}$. By [28, Proposition 20.7.15], the module $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}}$ is coherent. This implies that the dual of $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}}$ is the same as the continuous dual of $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}}$. Hence, the isomorphism $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}} \simeq \Theta^*_{\mathfrak{C}}$ follows from [28, Equation (20.7.14.4)].

Let $\mathcal{I}$ be the ideal defining the zero section in $\mathsf{N}_{\mathfrak{X}/C}$ and $i_n\colon \mathfrak{C}_n \to \mathsf{N}_{\mathfrak{X}/C}$ the canonical morphism. For each n, there is a short exact sequence

$$\mathcal{I}^n/\mathcal{I}^{2n} \xrightarrow{d_n} i_n^*\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}} \to \Omega^1_{\mathfrak{C}_n} \to 0, \tag{4-2}$$

where $d_n(\overline{f}) = 1 \otimes df$. Let $\mathscr{N}_n$ denote the image of d_n and notice that $\mathcal{I}\cdot\mathscr{N}_n = 0$ for all n. This implies that $\mathcal{I}\cdot\mathscr{N} = 0$, where $\mathscr{N} = \varprojlim_n \mathscr{N}_n$. But $\mathscr{N}$ is a submodule of the free $\mathcal{O}_{\mathfrak{C}}$–module $i^*\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}}$, implying that $\mathscr{N} = 0$. Similarly, the map $\mathscr{N}_{n+1} \to \mathscr{N}_n$ is zero for all n because $\mathscr{N}_{n+1}$ is a submodule of $\mathrm{ann}_{\mathcal{I}}(i^*_{n+1}\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}})$, which is mapped to zero under the map $i^*_{n+1}\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}} \to i^*_n\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}}$. Thus, $\{\mathscr{N}_n\}_{n\in\mathbb{N}}$ satisfies the Mittag-Leffler condition and hence $\varprojlim^{(1)}_n \mathscr{N}_n = 0$. Therefore, (4-2) induces an isomorphism $i^*\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}} \to \Omega^{1,\mathrm{cts}}_{\mathfrak{C}}$. Similarly, we have $\Theta_{\mathfrak{C}} = i^*\Theta_{\mathsf{N}_{\mathfrak{X}/C}}$. Thus the nondegenerate Poisson structure on $\mathsf{N}_{\mathfrak{X}/C}$ defines an isomorphism

$$\Omega^{1,\mathrm{cts}}_{\mathfrak{C}} \simeq i^*\Omega^1_{\mathsf{N}_{\mathfrak{X}/C}} \xrightarrow{\sim} i^*\Theta_{\mathsf{N}_{\mathfrak{X}/C}} \simeq \Theta_{\mathfrak{C}}.$$

The differential on the complex $\Omega^{*,\mathrm{cts}}_{\mathfrak{C}}$ is defined as in [30, Chapter I, Section 7]. Thus, the fact that we have an isomorphism of complexes $\bigwedge^*\Theta_{\mathfrak{C}} \simeq \Omega^{*,\mathrm{cts}}_{\mathfrak{C}}$ follows from the corresponding isomorphism $\bigwedge^*\Theta_{\mathsf{N}_{\mathfrak{X}/C}} \simeq \Omega^*_{\mathsf{N}_{\mathfrak{X}/C}}$ for $\mathsf{N}_{\mathfrak{X}/C}$ due to Lichnerowicz (see [20, Theorem 2.1.4]). □

Remark 4.6 Unlike for vector fields, we have $\Omega^{1,\mathrm{cts}}_{\mathfrak{C}} \neq \Omega^1_{\mathfrak{C}}$: indeed, the latter is not coherent over $\mathfrak{C}$.

As in Section 2.4 we let $C \to S$ denote the coisotropic reduction of C; as in loc. cit. this projection has a section $\nu\colon S \hookrightarrow C$. Let $z\colon C \hookrightarrow \mathfrak{C}$ be the embedding of the zero section and write $j = z \circ \nu \colon S \hookrightarrow \mathfrak{C}$ for the composite embedding. This is a closed immersion of formal Poisson schemes. Hence, restriction defines a morphism $j^{-1}(\bigwedge^* \Theta_{\mathfrak{C}}) \to \bigwedge^* \Theta_S$. Similarly, functoriality of the de Rham complex implies that we have a morphism $j^{-1}\Omega_{\mathfrak{C}}^{*,\mathrm{cts}} \to \Omega_S^*$. These form a commutative diagram:

$$\begin{tikzcd} j^{-1}\Omega_{\mathfrak{C}}^{*,\mathrm{cts}} \arrow[r, "\sim"] \arrow[d] & j^{-1}\bigwedge^* \Theta_{\mathfrak{C}} \arrow[d] \\ \Omega_S^* \arrow[r, "\sim"] & \bigwedge^* \Theta_S \end{tikzcd} \tag{4-3}$$

Lemma 4.7 *The morphism of complexes $j^{-1}\Omega_{\mathfrak{C}}^{*,\mathrm{cts}} \to \Omega_S^*$ is a quasi-isomorphism. Hence, the de Rham cohomology groups $H^2_{\mathrm{DR}}(\mathfrak{C})$ and $H^2_{\mathrm{DR}}(S)$ are isomorphic.*

Proof We factor $j^{-1}\Omega_{\mathfrak{C}}^{*,\mathrm{cts}} \to \Omega_S^*$ as

$$\nu^{-1}(z^{-1}\Omega_{\mathfrak{C}}^{*,\mathrm{cts}}) \to \nu^{-1}\Omega_C^* \to \Omega_S^*.$$

Thus, it suffices to show that each of $z^{-1}\Omega_{\mathfrak{C}}^{*,\mathrm{cts}} \to \Omega_C^*$ and $\nu^{-1}\Omega_C^* \to \Omega_S^*$ is a quasi-isomorphism. That the first is a quasi-isomorphism is [30, Chapter II, Proposition 1.1]. Since $C = S \times V$ and $V \simeq \mathbb{A}^m$ is contractible, the second morphism is a quasi-isomorphism. □

Now we would like to use the above results to relate quantizations of $\mathfrak{C}$ to quantizations of S. To accomplish this, we shall use some results of Bezrukavnikov and Kaledin [8] on period maps for quantizations. Their results are stated only for algebraic varieties, but they apply without essential change to smooth formal schemes as well. Since the results of this section are by now quite standard, and have also been summarized very well in [49], where the details of compatibility with $\mathbb{G}_m$–actions are also examined, we shall content ourselves with a very terse recollection.

To describe the results of [8], we first recall the notion of a Harish-Chandra torsor on $\mathfrak{X}$ (see also [49, page 1227] for more details): suppose that G is a (proalgebraic) group with Lie algebra $\mathfrak{g}$ and that $\mathfrak{h}$ is a Lie algebra such that $\mathfrak{g} \subset \mathfrak{h}$. Suppose further that $\mathfrak{h}$ is equipped with an action of G whose differential agrees with the adjoint action of $\mathfrak{g}$ on $\mathfrak{h}$. Then the pair $(G, \mathfrak{h})$ is known as a Harish-Chandra pair. A Harish-Chandra torsor for $(G, \mathfrak{h})$ is a pair (M, θ), where M is a G–torsor on $\mathfrak{X}$ and θ is an $\mathfrak{h}$–valued flat connection on M (the notions of torsor and flat connection are defined for (formal) schemes exactly as they are in usual differential geometry).

A symplectic variety comes equipped with a canonical Harish-Chandra torsor, defined as follows: Let $\mathfrak{A}$ denote the algebra of functions on a symplectic formal disc. Then the group of symplectomorphisms $\mathrm{Aut}(\mathfrak{A})$ of $\mathfrak{A}$ is naturally a proalgebraic group. Furthermore, the Lie algebra of Hamiltonian derivations of $\mathfrak{A}$, denoted by $\mathfrak{H}$, is a pro-Lie algebra and $(\mathrm{Aut}(\mathfrak{A}), \mathfrak{H})$ is a Harish-Chandra pair. Then the Harish-Chandra torsor $\mathcal{M}_{\mathrm{symp}}(\mathfrak{X})$ is defined to be the pro-scheme parametrizing all maps $\mathrm{Spec}(\mathfrak{A}) \to \mathfrak{X}$ which preserve the symplectic form.

Now, let D denote the (unique) quantization of a $2n$–dimensional formal disc over $\mathbb{C}$. Then the group of automorphisms $\mathrm{Aut}(D)$ comes with a natural proalgebraic group structure; similarly, the Lie algebra of derivations $\mathrm{Der}(D)$ is naturally a pro-Lie algebra making $(\mathrm{Aut}(D), \mathrm{Der}(D))$ a Harish-Chandra pair; we note that the map "reduction mod $\hbar$" gives a morphism of Harish-Chandra pairs $(\mathrm{Aut}(D), \mathrm{Der}(D)) \to (\mathrm{Aut}(\mathfrak{A}), \mathfrak{H})$.

Define

$$H^1_{\mathcal{M}_{\mathrm{symp}}}(\mathrm{Aut}(D), \mathrm{Der}(D))$$

to be the set of all isomorphism classes of $(\mathrm{Aut}(D), \mathrm{Der}(D))$–torsors on $\mathfrak{X}$ which are liftings of $\mathcal{M}_{\mathrm{symp}}$, ie those torsors equipped with a reduction of structure group to $(\mathrm{Aut}(\mathfrak{A}), \mathfrak{H})$ such that the resulting $(\mathrm{Aut}(\mathfrak{A}), \mathfrak{H})$–torsor is isomorphic to $\mathcal{M}_{\mathrm{symp}}$.

It is shown in [8, Section 3] that there is a natural bijection

$$\mathsf{Loc}\colon H^1_{\mathcal{M}_{\mathrm{symp}}}(\mathrm{Aut}(D), \mathrm{Der}(D)) \xrightarrow{\sim} Q(\mathfrak{X}),$$

where the right-hand side denotes the set of all isomorphism classes of quantizations of $\mathfrak{X}$. This bijection respects the $\mathbb{G}_m$–action on both sides, and hence it can be checked that it induces a bijection between equivariant quantizations and $H^1_{\mathcal{M}_{\mathrm{symp}}}(\mathrm{Aut}(D), \mathrm{Der}(D))^{\mathbb{G}_m}$.

Now the nonabelian cohomology group admits a natural "period map"

$$\mathsf{Per}\colon H^1_{\mathcal{M}_{\mathrm{symp}}}(\mathrm{Aut}(D), \mathrm{Der}(D)) \to H^2_{\mathrm{DR}}(\mathfrak{X})[\![\hbar]\!],$$

which moreover restricts to give a map $\mathsf{Per}\colon H^1_{\mathcal{M}_{\mathrm{symp}}}(\mathrm{Aut}(D), \mathrm{Der}(D))^{\mathbb{G}_m} \to H^2_{\mathrm{DR}}(\mathfrak{X})$. In good situations, such as when $H^i(\mathfrak{X}, \mathcal{O}_{\mathfrak{X}}) = 0$ for $i = 1, 2$, this equivariant period map is an isomorphism. We thus have the following classification:

Theorem 4.8 *Let $\mathfrak{X}$ be a smooth symplectic affine algebraic variety or formal scheme, with an elliptic action of $\mathbb{G}_m$ (assumed to be prorational if $\mathfrak{X}$ is formal). We let $Q^{\mathbb{G}_m}(\mathfrak{X})$ denote the set of isomorphism classes of $\mathbb{G}_m$–equivariant quantizations of $\mathfrak{X}$. Then there is a natural bijection*

$$Q^{\mathbb{G}_m}(\mathfrak{X}) \to H^2_{\mathrm{DR}}(\mathfrak{X}).$$

To relate quantizations of $\mathfrak{C}$ to quantizations of S we may thus relate the corresponding Harish-Chandra torsors: given a Harish-Chandra torsor on S and a Harish-Chandra torsor on $\widehat{T^*V}$, the external product defines a new Harish-Chandra torsor on $S \times \widehat{T^*V} \simeq \mathfrak{C}$. If we suppose that these Harish-Chandra torsors are liftings of $\mathcal{M}_{\mathrm{symp}}(S)$ and $\mathcal{M}_{\mathrm{symp}}(\widehat{T^*V})$, then under the map Loc, this external product of torsors corresponds to taking a quantization B of S and associating to it the quantization $B \otimes \widehat{\mathcal{D}}$ on $\mathfrak{C}$. On the other hand, the map Per is constructed by associating, to a Harish-Chandra torsor which is a lifting of $\mathcal{M}_{\mathrm{symp}}$, a certain canonical cohomology class associated to the lifting. From this is follows that the map $Q(S)^{\mathbb{G}_m} \to Q(\mathfrak{C})^{\mathbb{G}_m}$ (given by $B \to B \otimes \widehat{\mathcal{D}}$) corresponds, under $\mathsf{Per} \circ \mathsf{Loc}^{-1}$, to the map $H^2_{\mathrm{DR}}(S) \to H^2_{\mathrm{DR}}(\mathfrak{C})$ associated to the inclusion $S \to \mathfrak{C}$. Since this has been shown to be an isomorphism above, we have shown:

Proposition 4.9 *Let $\widehat{A}$ be a $\mathbb{G}_m$–equivariant quantization of $\mathfrak{C}$. Then there exist an equivariant quantization B of S and an equivariant isomorphism $\psi\colon \widehat{A} \xrightarrow{\sim} B \mathbin{\widehat{\otimes}} \widehat{\mathcal{D}}$ of deformation-quantization algebras.*

A quantization A of $\mathfrak{X}$ induces a quantization $\widehat{A}$ of $\mathfrak{C}$. In the following subsections we use the isomorphism ψ given by the above proposition to show that the quantum Hamiltonian reduction of A is isomorphic to B.

4.3 The completion of the ideal J

Recall the ideal $J \subset A$ from Section 4.1. Let $\widehat{J} = \widehat{A} \otimes_A J$. Lemma 4.3 implies that the natural map $\widehat{J} \to \widehat{A}$ is an embedding. Let $\mathfrak{u}$ be the m–dimensional subspace of $A_{<0}$ spanned by a collection of homogeneous lifts of the $w_i \in I$ and let $J' = A\mathfrak{u}$. Since the isomorphism ψ of Proposition 4.9 induces the identity on $\widehat{R}$, the image $\psi(\mathfrak{u})$ consists of lifts of the w_i to $\widehat{\mathcal{D}}$. We denote this space by $\mathfrak{u}$ as well.

Lemma 4.10 *We have $J = J'$ and hence $\psi(\widehat{J}) = B \mathbin{\widehat{\otimes}} (\widehat{\mathcal{D}}\mathfrak{u})$. Thus $(A/J)/\hbar \xrightarrow{\sim} R/I$.*

Proof Clearly we have $J' \subseteq J$. In order to get the opposite inclusion, let $a \in J$ be any element; since J is generated by homogeneous elements, it suffices to assume a is homogeneous. Recall that $\rho\colon A \to R$ is the projection. Since $\rho(J) = I = R\mathfrak{u}$, we may write

$$a_1' := a - \sum a_i y_i \in \hbar A$$

for some homogenous elements $\{a_i\}$, and we may select these elements so that $\deg \sum a_i y_i = \deg a < 0$. But since $\deg \hbar = l$, we have that $a_1' = \hbar a_1$ for some a_1 of strictly smaller degree than a; evidently a_1 is again in $J \smallsetminus J'$. We may repeat this process to find $a_2' = \hbar a_2$, etc. But now if we look at these equations in the finitely generated left A–module $N = J/J'$, we see that we have a sequence $a = \hbar a_1 = \hbar^2 a_2 = \cdots$. But this implies that $\bigcap_{n \geq 0} \hbar^n N$ is nonzero, which contradicts the fact that N is h–complete and hence separated.

To see the last statement, note that the first statement shows that the image of the natural map $J/\hbar \to A/\hbar = R$ is precisely I. Then the statement follows from applying the functor $M \to M/\hbar$ to the short exact sequence $J \to A \to A/J$. □

4.4 Identification of the formal quantum coisotropic reduction

We require the following result, which is [42, Lemmas 1.2.6 and 1.2.7]:

Lemma 4.11 *Let L be a finitely generated, free A–module and N a submodule. Then $\bigcap_{i=1}^{\infty}(N + \hbar^i L) = N$.*

Proposition 4.12 *The natural map $A/J \to \widehat{A/J}$ is an embedding. Hence, A/J is $\hbar$–flat.*

Proof Note that Lemma 4.3 implies that the sequence $0 \to \hat{J} \to \hat{A} \to \widehat{A/J} \to 0$ is exact, hence it suffices to show that $\hat{J} \cap A = J$. Since the image of J in A/K^n equals $(J + K^n)/K^n$, we have

$$\hat{J} \cap A = \bigcap_{n=1}^{\infty}(J + K^n).$$

Claim 4.13 *We have $\bigcap_{n=1}^{\infty}(J + K^n) \subset J$.*

Proof Notice that $K = JA + \hbar A = J + \hbar A$ because $K/\hbar A = (J + \hbar A)/\hbar A = I$. Consider the expansion of $(J + \hbar A)^n$. Since $\hbar$ is central, a term of this expansion containing i copies of $\hbar A$ equals $\hbar^i J^{n-i} A$. Multiplying $JA + \hbar A = J + \hbar A$ on the left by $\hbar^i J^{n-i-1}$ implies that

$$\hbar^i J^{n-i} A + \hbar^{i+1} J^{n-i-1} A = \hbar^i J^{n-i} + \hbar^{i+1} J^{n-i-1} A.$$

Thus,

$$(J + \hbar A)^n = \hbar^n A + \sum_{i=0}^{n-1} \hbar^i J^{n-i} \subset \hbar^n A + J.$$

Since $J + \bigcap_{n=1}^{\infty} \hbar^n A \subset J + \hbar^n A$ for all n, we have $J + \bigcap_{n=1}^{\infty} \hbar^n A \subset \bigcap_{n=1}^{\infty}(J + \hbar^n A)$. On the other hand, Lemma 4.11 says that $\bigcap_{n=1}^{\infty}(J + \hbar^n A) = J$. This completes the proof of Claim 4.13. □

Returning to the proof of Proposition 4.12, the second assertion will now follow from the fact that $\hat{A}/\hat{J}$ is $\hbar$–flat. By Lemma 4.10,

$$\hat{A}/\hat{J} = B \mathbin{\widehat{\otimes}} (\hat{\mathcal{D}}/\hat{\mathcal{D}}\mathfrak{u}). \tag{4-4}$$

The Poincaré–Birkhoff–Witt property of $\hat{\mathcal{D}}$ implies that the right-hand side of (4-4) is $\hbar$–flat. □

Recall again that we have constructed a coisotropic reduction $\pi\colon C \to S$. Let B be the quantization of the Poisson algebra $T = \mathbb{C}[S]$ given by Proposition 4.9. We have an identification $T = (R/I)^{\{I,-\}}$ such that the embedding $(R/I)^{\{I,-\}} \hookrightarrow R/I$ is just the comorphism π^*.

Proposition 4.14 (1) *The $\mathbb{C}[\![\hbar]\!]$–algebra $\mathrm{End}_{\widehat{\mathcal{D}}}(\widehat{\mathcal{D}}/\widehat{\mathcal{D}}\mathfrak{u})$ is isomorphic to $\mathbb{C}[\![\hbar]\!]$ and $\mathrm{Ext}^i_{\widehat{\mathcal{D}}}(\widehat{\mathcal{D}}/\widehat{\mathcal{D}}\mathfrak{u}, \widehat{\mathcal{D}}/\widehat{\mathcal{D}}\mathfrak{u})$ is a torsion $\mathbb{C}[\![\hbar]\!]$–module for $i > 0$.*

(2) *We have an isomorphism of $\mathbb{C}[\![\hbar]\!]$–algebras*

$$\mathrm{End}_{\widehat{A}}(\widehat{A}/\widehat{J})^{\mathrm{opp}} \xrightarrow{\sim} B$$

and $\mathrm{Ext}^i_{\widehat{A}}(\widehat{A}/\widehat{J}, \widehat{A}/\widehat{J})$ is $\hbar$–torsion for all $i > 0$.

Proof (1) Recall that, as above, $\widehat{\mathcal{D}}$ is the Moyal–Weyl quantization of the algebra $\widehat{F}$ of functions on $\widehat{T^*V}$, the formal neighborhood of the zero section; $\mathfrak{u}$ is a space spanned by homogeneous lifts of generators of I/I^2 to $\widehat{\mathcal{D}}$. Under the identification of $\widehat{\mathcal{D}}$ with $\widehat{F}[\![\hbar]\!]$, we may assume that the elements $w_1, \dots, w_m$ of $\mathfrak{u}$ are coordinate functions on V^*. Write $L := \widehat{\mathcal{D}}/\widehat{\mathcal{D}}\mathfrak{u}$.

It is standard, for a cyclic left module S/P over a ring S, that $\mathrm{End}_S(S/P) \cong \{q \in S/P \mid Pq \subseteq P\}$; it is also standard that $\{f \in \mathcal{D}/\mathcal{D}\mathfrak{u} \mid \mathfrak{u} \cdot f = 0\} \cong \mathbb{C}[[\hbar]]$ (and in any case this can be computed, by hand, inductively using the Moyal–Weyl product).

Letting z_i denote the dual coordinates on V as in Theorem 2.6, we get an identification

$$\widehat{\mathcal{D}} \cong \mathbb{C}[z_1, \dots, z_m][\![w_1, \dots, w_m, \hbar]\!], \tag{4-5}$$

with Moyal–Weyl product $*$ satisfying

$$\begin{aligned} w_i * w_j &= w_i w_j, & w_i * f(z,w) &= w_i \cdot f(z,w) + \frac{\hbar}{2}\frac{\partial f}{\partial z_i}, \\ z_i * z_j &= z_i z_j, & f(z,w) * w_i &= w_i \cdot f(z,w) - \frac{\hbar}{2}\frac{\partial f}{\partial z_i}. \end{aligned}$$

It follows from Lemma 4.10 that the natural composite $\mathbb{C}[z_1, \dots, z_m][\![\hbar]\!] \hookrightarrow \widehat{D} \to L$ is an isomorphism of vector spaces: via the vector space isomorphism of (4-5) and the formulas above, we can write any element of $\widehat{D}$ as $\sum_{I,j} f_{I,j}(z) * w^I \hbar^j$, and then those terms with nonconstant w^I vanish in L. Under this identification, for $f(z) \in L$, we have

$$w_i * f(z) = w_i * f(z) - f(z) * w_i = \hbar \frac{\partial f}{\partial z_i}. \tag{4-6}$$

Let $K(\mathfrak{u}) \cong \big(\bigwedge^{\bullet}(\mathfrak{u}) \otimes \mathrm{Sym}^{\bullet}(\mathfrak{u}), d\big)$ denote the Koszul complex of $\mathbb{C}$ as a $\mathrm{Sym}(\mathfrak{u})$–module, and let $K(\widehat{\mathcal{D}}, \mathfrak{u}) = \widehat{\mathcal{D}} \otimes_{\mathrm{Sym}^{\bullet}(\mathfrak{u})} K(\mathfrak{u})$; so $K(\widehat{\mathcal{D}}, \mathfrak{u})$ is a finite free resolution of L over $\widehat{\mathcal{D}}$. By adjunction,

$$\mathrm{Ext}^i_{\widehat{\mathcal{D}}}(L, L) \cong H^i\big(\mathrm{Hom}_{\mathrm{Sym}(\mathfrak{u})}(K(\mathfrak{u}), L)\big) \cong H^i\big(\textstyle\bigwedge^{\bullet}(\mathfrak{u}^*) \otimes \widehat{\mathcal{D}}/\widehat{\mathcal{D}}\mathfrak{u}\big).$$

It is then evident from (4-6) that the identification $\mathbb{C}[z_1, \dots, z_m][\![\hbar]\!] \to L$ intertwines the Koszul differential and $\hbar$ times the de Rham differential, thus yielding, when $\hbar$ is inverted,

$$H^i\big(\textstyle\bigwedge^{\bullet}(\mathfrak{u}^*) \otimes \widehat{\mathcal{D}}/\widehat{\mathcal{D}}\mathfrak{u}\big)[\hbar^{-1}] \cong H^i_{\mathrm{DR}}(\mathrm{Spec}\,\mathbb{C}[z_1, \dots, z_m])((\hbar)),$$

which proves the $i > 0$ part of (1).

(2) Again, we have

$$\mathrm{Ext}^i_{B\widehat{\otimes}\widehat{\mathcal{D}}}(B\widehat{\otimes}L, B\widehat{\otimes}L) \cong H^i\big(\mathrm{Hom}_{B\widehat{\otimes}\widehat{\mathcal{D}}}(B\widehat{\otimes}K(\widehat{\mathcal{D}}, \mathfrak{u}), B\widehat{\otimes}L)\big) \cong H^i\big(\textstyle\bigwedge^{\bullet}(\mathfrak{u}^*) \otimes B\widehat{\otimes}L\big),$$

where the last isomorphism follows by adjunction as before. Let d_B denote the Koszul differential on this completed tensor product and d the Koszul differential on $\bigwedge^{\bullet}(\mathfrak{u}^*) \otimes L$. The $\mathfrak{u}$–action commutes with all elements of B and $B \cong \mathbb{C}[S][\![\hbar]\!]$ as a free $\mathbb{C}[\![\hbar]\!]$–module. Thus, letting $\{s_i\}$ denote a vector space basis of $\mathbb{C}[S]$, for any element $\sum s_i \otimes l_i$ of $B \mathbin{\widehat{\otimes}} L$ we get $d_B\big(\sum s_i \otimes l_i\big) = \sum s_i d(l_i)$, and it follows that

$$\ker(d_B) = B \mathbin{\widehat{\otimes}} \ker(d), \quad \mathrm{Im}(d_B) = B \mathbin{\widehat{\otimes}} \mathrm{Im}(d).$$

Thus,

$$\mathrm{Ext}^i_{B\widehat{\otimes}\widehat{\mathcal{D}}}(B \mathbin{\widehat{\otimes}} L, B \mathbin{\widehat{\otimes}} L) \cong B \mathbin{\widehat{\otimes}}_{\mathbb{C}[\![\hbar]\!]} H^i\big(\textstyle\bigwedge^{\bullet}(\mathfrak{u}^*) \otimes L\big) \cong B \mathbin{\widehat{\otimes}}_{\mathbb{C}[\![\hbar]\!]} \mathrm{Ext}^i_{\widehat{\mathcal{D}}}(L, L),$$

reducing the assertions of (2) to (1). □

4.5 Identification of the quantum coisotropic reduction

For any $s > 0$, let $A_s = A/\hbar^s A$ and $\widehat{A}_J = \varprojlim_n A/J^n$. Even though J is only a left ideal of A, we can form the Rees algebra $\mathsf{Rees}_J(A) = \bigoplus_{n \geq 0} J^n$, with the obvious multiplication. We shall abuse notation and denote by J the left ideal generated by the image of J in A_s.

Lemma 4.15 (1) *The inclusion $J^n \subset K^n$ induces an isomorphism of complete topological algebras $\widehat{A}_J \xrightarrow{\sim} \widehat{A}$.*

(2) *The Rees algebra $\mathsf{Rees}_J(A_s)$ is (both left and right) Noetherian.*

(3) *Let M be a finitely generated A–module. Then $\widehat{A}_J \otimes_A M \simeq \varprojlim_n M/J^n M$.*

Proof (1) Since $[A, A] \subset \hbar A$, we have $K^n \subseteq J^n + \hbar J^n + \cdots + \hbar^s J^{n-s}$ for all $s, n > 0$. Therefore the filtrations $\{K^n\}_n$ and $\{J^n\}_n$ of A_s are comparable and the canonical morphism $\varprojlim A_s/J^n \to \varprojlim A_s/K^n$ is an isomorphism. Thus, commutativity of limits implies that

$$\varprojlim_n A/J^n = \varprojlim_s \big(\varprojlim_n A_s/J^n\big) \to \varprojlim_s \big(\varprojlim_n A_s/K^n\big) = \varprojlim_n A/K^n$$

is an isomorphism.

(2) Since A_s is finitely generated and nilpotent, we may choose a finite-dimensional vector subspace $\mathfrak{n}$ of A_s which is bracket closed and generates A_s as an algebra; enlarging $\mathfrak{n}$ if necessary, we may suppose $J \cap \mathfrak{n} = \mathfrak{n}_1$ is a Lie subalgebra which generates J as an ideal. Then $\mathfrak{n}$ is a nilpotent Lie algebra, and we have a surjection $U(\mathfrak{n}) \to A_s$. Further, we have the subalgebra $U(\mathfrak{n}_1)$; the image of its augmentation ideal in A_s is J. Thus the claim is reduced to showing the following: let $\mathfrak{n}_1 \subset \mathfrak{n}$ be nilpotent Lie algebras, and let $\mathfrak{a}$ be the left ideal of $U(\mathfrak{n})$ generated by $\mathfrak{n}_1$; then $\mathsf{Rees}_{\mathfrak{a}}(U(\mathfrak{n}))$ is Noetherian. But this is a standard argument; see for instance [64].

(3) This is a noncommutative analogue of [21, Theorem 7.2]. If M' is a submodule of M, then the argument given in the proof of [21, Lemma 7.15] shows that the claim reduces to showing that the morphism

$$\varprojlim_{s,n} M'/(J^n M' + \hbar^s M') \to \varprojlim_{s,n} M'/((J^n M) \cap M' + \hbar^s M')$$

is an isomorphism. This will be an isomorphism if, for each s, $n \geq 1$, there exist $N(s,n)$, $S(s,n) \gg 0$ such that

$$(J^N M) \cap M' + \hbar^S M' \subset J^n M' + \hbar^s M'.$$

By Lemma 3.2(2), the Rees algebra $\mathsf{Rees}_{\hbar A}(A)$ is Noetherian. Therefore, there exists some s_0 such that $\hbar^{i+s_0} M \cap M' \subset \hbar^i M'$ for all $i \geq 1$. The A_{s+s_0}–module $M'/(\hbar^{s+s_0} M \cap M')$ is a submodule of $M/\hbar^{s+s_0} M$. Since we have shown in (2) that the Rees algebra $\mathsf{Rees}_J(A_{s+s_0})$ is Noetherian, the usual Artin–Rees argument shows that there exists some $N \gg 0$ such that

$$(J^N \cdot M) \cap M' + (\hbar^{s+s_0} M \cap M') \subset J^n \cdot M' + (\hbar^{s+s_0} M \cap M'). \tag{4-7}$$

Since $\hbar^{s+s_0} M \cap M' \subset \hbar^s M'$, the inclusion (4-7) implies that

$$(J^N \cdot M) \cap M' + \hbar^s M' \subset J^n \cdot M' + \hbar^s M',$$

as required. □

Remark 4.16 One can check that the ring $\mathsf{Rees}_J(A)$ is *not* in general Noetherian.

Theorem 4.17 (1) *The $\mathbb{C}[\![\hbar]\!]$–algebras $\operatorname{End}_A(A/J)^{\mathrm{opp}}$ and B are isomorphic. Hence $\operatorname{End}_A(A/J)^{\mathrm{opp}}$ is a deformation-quantization of S.*

(2) *The ext groups* $\mathrm{Ext}^i_A(A/J, A/J)$ *are* $\hbar$*–torsion for all* $i > 0$*.*

(3) A/J *is a faithfully flat* B*–module.*

Proof (1)–(2) Since $\widehat{A}/\widehat{J} = \widehat{A} \otimes_A (A/J)$ and $\widehat{A}$ is flat over A, adjunction says that we have

$$\mathrm{Ext}^i_{\widehat{A}}(\widehat{A}/\widehat{J}, \widehat{A}/\widehat{J}) \simeq \mathrm{Ext}^i_A(A/J, \widehat{A}/\widehat{J}) \quad \text{for all } i \geq 0.$$

Since A/J is a finitely generated A–module, Lemmas 4.3 and 4.15 imply that $\widehat{A}/\widehat{J} = \widehat{A} \otimes_A (A/J)$ is isomorphic to $\widehat{A}_J \otimes_A (A/J)$. By Lemma 4.15, we have

$$\widehat{A}_J \otimes_A (A/J) = \varprojlim_n \frac{(A/J)}{J^n \cdot (A/J)} = A/J.$$

Therefore, (1) and (2) follows from Proposition 4.14.

(3) By Proposition 4.12, A/J is $\hbar$–flat, or equivalently $\hbar$–torsion-free. Since it is finitely generated over A, it is also $\hbar$–complete. Therefore, [42, Corollary 1.5.7] says that it is cohomologically complete. By 4.10, $(A/J)/\hbar(A/J) \simeq R/I$ and hence is a free T–module. Thus, [42, Theorem 1.6.6] implies that A/J is a faithfully flat B–module. □

Let $W = A[\hbar^{-1}]$, a $\mathcal{W}$–algebra. By base change, Theorem 4.17 implies:

Corollary 4.18 *The algebra*

$$\mathrm{End}_W(W/J[\hbar^{-1}], W/J[\hbar^{-1}])^{\mathrm{opp}} \xrightarrow{\sim} B[\hbar^{-1}] =: W_S$$

is a $\mathcal{W}$*–algebra, and* $\mathrm{Ext}^i_W(W/J[\hbar^{-1}], W/J[\hbar^{-1}]) = 0$ *for all* $i > 0$*.*

4.6 Equivariant modules

We maintain the notation and assumptions of the prior subsections of Section 4. In particular, $\mathfrak{X} = \mathrm{Spec}(R)$ is a smooth affine symplectic variety with $\mathbb{G}_m$–action and with coisotropic subvariety $C = \mathrm{Spec}(R/I)$, where I is generated by all homogeneous elements of negative degree. Moreover, A is a deformation-quantization of R, and $W = A[\hbar^{-1}]$. We have a symplectic quotient $C \to S$, also assumed affine, and B is a deformation-quantization of $\mathbb{C}[S]$.

Definition 4.19 The full abelian subcategory of $(A, \mathbb{G}_m)$–mod consisting of all modules supported on C is denoted by $(A, \mathbb{G}_m)$–mod_C. The full abelian subcategory of $(W, \mathbb{G}_m)$–good consisting of good $\mathcal{W}$–modules supported on C is denoted by $(W, \mathbb{G}_m)$–good_C.

We define a filtration $\mathcal{H}_i M$ on a finitely generated, equivariant A–module by letting $\mathcal{H}_i(M/\hbar^n M)$ be the direct sum $\bigoplus_{j\geq i}(M/\hbar^n M)_j$ and $\mathcal{H}_i M = \varprojlim_n \mathcal{H}_i(M/\hbar^n M)$. Then $\hbar\mathcal{H}_i M \subset \mathcal{H}_{i+1} M$. The filtration $\mathcal{H}_i M$ need not be exhaustive.

Lemma 4.20 *Let $M \in (A, \mathbb{G}_m)\text{–mod}$. Then $\operatorname{Supp} M \subset C$ if and only if $\mathcal{H}_N M = M$ for $N \ll 0$.*

Proof If $\mathcal{H}_N M = M$, then $\mathcal{H}_N(M/\hbar M) = M/\hbar M$. This means that $M/\hbar M$ is a graded R–module, with $(M/\hbar M)_i = 0$ for $i \ll 0$. This implies that $\operatorname{Supp}(M/\hbar M)$ is contained in C. Therefore, by Lemma 3.24, the support of M is contained in C. Conversely, if $\operatorname{Supp} M \subset C$ then clearly the support of $M/\hbar M$ is contained in C too. This implies that $\mathcal{H}_N(M/\hbar M) = M/\hbar M$ for some $N \ll 0$. By induction on n, the exact sequence

$$M/\hbar M \xrightarrow{\cdot\hbar^n} M/\hbar^{n+1} M \to M/\hbar^n M \to 0$$

implies that $\mathcal{H}_N(M/\hbar^n M) = M/\hbar^n M$ and hence $\mathcal{H}_N(M) = M$. □

4.7 Quantum coisotropic reduction: affine case

We maintain the notation of Section 4.6. For a module $M[\hbar^{-1}] \in (W, \mathbb{G}_m)\text{–good}_C$, we denote by M a choice of lattice in $(A, \mathbb{G}_m)\text{–mod}_C$. Recall that $W = A[\hbar^{-1}]$ and $W_S = B[\hbar^{-1}]$.

By Theorem 4.17, we can define an adjoint pair $(\mathbb{H}^{\perp}, \mathbb{H})$ of functors

$$\mathbb{H}\colon (A, \mathbb{G}_m)\text{–mod}_C \rightleftarrows (B, \mathbb{G}_m)\text{–mod} :\mathbb{H}^{\perp}$$

by

$$\mathbb{H}(M) = \operatorname{Hom}_A(A/J, M) \quad \text{and} \quad \mathbb{H}^{\perp}(N) = A/J \otimes_B N.$$

The functors $\mathbb{H}$ and $\mathbb{H}^{\perp}$ clearly preserve the subcategories of $\hbar$–torsion modules, and in particular by Proposition 3.26 they thus induce a well-defined adjoint pair of functors

$$\mathbb{H}\colon (W, \mathbb{G}_m)\text{–good}_C \rightleftarrows (W_S, \mathbb{G}_m)\text{–good} :\mathbb{H}^{\perp} \tag{4-8}$$

for which

$$\mathbb{H}(M[\hbar^{-1}]) = \operatorname{Hom}_A(A/J, M)[\hbar^{-1}] \quad \text{and} \quad \mathbb{H}^{\perp}(N[\hbar^{-1}]) = (A/J \otimes_B N)[\hbar^{-1}].$$

Theorem 4.21 *The functors $\mathbb{H}$ and $\mathbb{H}^{\perp}$ of (4-8) are exact, mutually quasi-inverse equivalences of abelian categories.*

Proof Suppose $M \in (A, \mathbb{G}_m)\text{–mod}_C$. Let $M^{\mathrm{rat}} = \bigoplus_i M_i$ be the subspace of $\mathbb{G}_m$–locally finite vectors. By Lemma 3.21, this space is nonzero if M is. Lemma 4.20

implies that there exists N such that $M_N \neq 0$ but $M_i = 0$ for all $i < N$. Now $\mathfrak{u} \cdot M_N = 0$ and hence:

$$\mathbb{H} \text{ is left-exact on } (A, \mathbb{G}_m)\text{–}\mathsf{mod}_C \text{ and } (W, \mathbb{G}_m)\text{–}\mathsf{good}_C, \\ \text{and conservative on } (A, \mathbb{G}_m)\text{–}\mathsf{mod}_C.$$

Remark 4.22 By Theorem 4.17(3), A/J is a faithfully flat B–module; thus $\mathbb{H}^{\perp}(N)$ is $\hbar$–torsion-free if N is $\hbar$–torsion-free. Similarly $\mathbb{H}(N)$ is $\hbar$–torsion-free if N is, by inspection.

Theorem 4.17(3) also implies:

$$\mathbb{H}^{\perp} \text{ is exact and conservative on } (B, \mathbb{G}_m)\text{–}\mathsf{mod} \text{ and } (W_S, \mathbb{G}_m)\text{–}\mathsf{good}.$$

Next, we show:

Claim 4.23 *The adjunction*

$$\mathrm{id} \to \mathbb{H} \circ \mathbb{H}^{\perp}$$

is an isomorphism of functors of $(W_S, \mathbb{G}_m)$–good.

To prove the claim, we observe that the global dimensions of B and W_S are finite, and therefore we prove, by induction on the projective dimension of $N \in (B, \mathbb{G}_m)$–mod, that:

(a) $N[\hbar^{-1}] \to \mathbb{H}(\mathbb{H}^{\perp}(N[\hbar^{-1}]))$ is an isomorphism.

(b) $\mathrm{Ext}^i_A(A/J, \mathbb{H}^{\perp}(N))[\hbar^{-1}] = 0$ for all $i > 0$.

When N is a finitely generated projective B–module, the assertions are immediate from Corollary 4.18. So we may assume that assertions (a) and (b) hold for all modules F with projective dimension less than $\mathrm{pd}_B\, N$. Fix a presentation $0 \to F \to B^k \to N \to 0$, so that the projective dimension of F is less than the projective dimension of N. Since $\mathbb{H}^{\perp}$ is exact we get an exact sequence

$$0 \to \mathbb{H} \circ \mathbb{H}^{\perp}(F) \to \mathbb{H} \circ \mathbb{H}^{\perp}(B^k) \to \mathbb{H} \circ \mathbb{H}^{\perp}(N) \to \mathrm{Ext}^1_A(A/J, \mathbb{H}^{\perp}(F)) \to \cdots. \tag{4-9}$$

Inverting $\hbar$ and using the inductive hypothesis that assertion (b) holds for F, we get a short exact sequence

$$0 \to \mathbb{H} \circ \mathbb{H}^{\perp}(F)[\hbar^{-1}] \to \mathbb{H} \circ \mathbb{H}^{\perp}(B^k)[\hbar^{-1}] \to \mathbb{H} \circ \mathbb{H}^{\perp}(N)[\hbar^{-1}] \to 0.$$

Assertion (a) for F and B^k then implies assertion (a) for N. Similarly, it follows, by continuing the exact sequence (4-9), from assertion (b) for B^k and F that $\mathrm{Ext}^i_A(A/J, \mathbb{H}^{\perp}(N))[\hbar^{-1}] = 0$ for $i \geq 1$, ie assertion (b) holds for N as well. This proves the inductive step, hence the claim.

Finally, we need to show:

Claim 4.24 *The adjunction*

$$\mathbb{H}^{\perp} \circ \mathbb{H} \to \mathrm{id}$$

is an isomorphism of functors on $(W_S, \mathbb{G}_m)$–good.

Suppose $M \subset M[\hbar^{-1}]$ is a lattice. Taking kernels and cokernels gives

$$0 \to E \to \mathbb{H}^{\perp} \circ \mathbb{H}(M) \to M \to E' \to 0.$$

Applying $\mathbb{H}$ and localizing gives

$$0 \to \mathbb{H}(E)[\hbar^{-1}] \to \mathbb{H} \circ \mathbb{H}^{\perp} \circ \mathbb{H}(M)[\hbar^{-1}] \xrightarrow{\mathrm{id}} \mathbb{H}(M)[\hbar^{-1}] \to \mathbb{H}(E')[\hbar^{-1}] \to \cdots.$$

This implies that $\mathbb{H}(E)[\hbar^{-1}] = 0$ and hence $\mathbb{H}(E)$ is $\hbar$–torsion. But by Remark 4.22, $\mathbb{H}^{\perp}(\mathbb{H}(M))$ is $\hbar$–torsion-free, hence so is E, hence again by the remark so is $\mathbb{H}(E)$; this implies that $\mathbb{H}(E) = 0$. But $\mathbb{H}$ is conservative, so $E = 0$. Thus, we have

$$0 \to \mathbb{H}^{\perp} \circ \mathbb{H}(M) \to M \to E' \to 0.$$

Again, applying $\mathbb{H}$, localizing and using assertion (b) above to obtain that the extension group $\mathrm{Ext}^1_A(A/J, \mathbb{H}^{\perp} \circ \mathbb{H}(M))[\hbar^{-1}]$ is zero, we get an exact sequence

$$0 \to \mathbb{H}(M)[\hbar^{-1}] \xrightarrow{\mathrm{id}} \mathbb{H}(M)[\hbar^{-1}] \to \mathbb{H}(E')[\hbar^{-1}] \to 0.$$

This implies that $\mathbb{H}(E') = \mathrm{Hom}_A(A/J, M)$ is $\hbar$–torsion. Let E'_{tf} denote the quotient of E' by its $\hbar$–torsion submodule. Using the exact sequence

$$0 \to \mathrm{Hom}_A(A/J, E'_{\mathrm{tor}}) \to \mathrm{Hom}_A(A/J, E') \to \mathrm{Hom}_A(A/J, E'_{tf}) \to \mathrm{Ext}^1_A(A/J, E'_{\mathrm{tor}})$$

and that the left-hand and right-hand terms are $\hbar$–torsion, together with Remark 4.22, we conclude that $\mathbb{H}(E'_{tf}) \subset \mathbb{H}(E'_{tf})[\hbar^{-1}] = \mathbb{H}(E')[\hbar^{-1}] = 0$. Since $\mathbb{H}$ is conservative, $E'_{tf} = 0$, and thus E' is a torsion module. The claim follows. This completes the proof of the theorem. □

4.8 Quantum coisotropic reduction: global case

In this section, we fix a connected component Y of the fixed point set of $\mathfrak{X}$ and let $C \subset \mathfrak{X}$ denote the set of points converging to Y under the elliptic $\mathbb{G}_m$–action. We assume that C is *closed* in $\mathfrak{X}$. Let $\mathcal{W}_{\mathfrak{X}}$–$\mathsf{good}_C$ denote the category of $\mathbb{G}_m$–equivariant, good $\mathcal{W}_{\mathfrak{X}}$–modules supported on C.

Lemma 4.25 *There exists an affine* $\mathbb{G}_m$*–stable open covering* $\{U_i\}_{i \in I}$ *of* $\mathfrak{X}$ *such that*

$$C \cap U_i = \left\{ x \in U_i \mid \lim_{t \to \infty} t \cdot x \in Y \cap U_i \right\}$$

for all i.

Proof First choose a $\mathbb{G}_m$–stable affine open covering $\{V_i\}$ of $\mathfrak{X} \smallsetminus C$. Replacing $\mathfrak{X}$ by $\mathfrak{X} \smallsetminus D$, where $D = \mathfrak{X}^{\mathbb{G}_m} \smallsetminus Y$, we may assume that if $\lim_{t \to \infty} t \cdot x$ exists then it belongs to Y. Now take any collection of affine $\mathbb{G}_m$–stable open subsets V'_j of $\mathfrak{X}$ such that (a) $V'_j \cap Y \neq \varnothing$ for all j, and (b) $\bigcup_j U'_j \cap Y = Y$. Then Lemma 3.30 implies that $\{V_i\} \cup \{V'_j\}$ is the desired covering. $\square$

Let $\mathcal{I} \subset \mathcal{O}_{\mathfrak{X}}$ denote the ideal of C.

Lemma 4.26 *There exists an $\hbar$–flat quantization $\mathcal{J}$ of $\mathcal{I}$.*

Proof Let $\{U_i\}_{i \in I}$ be a $\mathbb{G}_m$–stable open cover of $\mathfrak{X}$ satisfying the properties of Lemma 4.25. It suffices to construct a sheaf $\mathcal{J}_i$ on each U_i, together with a natural identification on over-laps $U_i \cap U_j$. If $U_i \cap C = \varnothing$ then we set $\mathcal{J}_i = \mathcal{A}$. If $U_i \cap C \neq \varnothing$, then $\mathcal{J}_i$ is defined to be the coherent sheaf associated to the left ideal of $\Gamma(U_i, \mathcal{A})$ generated by all homogeneous sections of negative degree. In each case, $\mathcal{J}_i$ is a subsheaf of $\mathcal{A}|_{U_i}$. Therefore, it suffices to show that $\mathcal{J}_i|_{U_i \cap U_j} = \mathcal{J}_j|_{U_i \cap U_j}$ as subsheaves of $\mathcal{A}|_{U_i \cap U_j}$.

If $U_i \cap C = U_j \cap C = \varnothing$ there is nothing to prove. Therefore, we assume that $U_i \cap C \neq \varnothing$. If $U_i \cap U_j \cap C = \varnothing$, let $D(f) \subset U_i \cap U_j$ be any $\mathbb{G}_m$–stable affine open subset that is the complement of the vanishing set of $f \in \Gamma(U_i, \mathcal{O}_{U_i})$. Then $C \cap D(f) = \varnothing$, ie if $I \subset \Gamma(U_i, \mathcal{O}_{U_i})$ is the ideal of $C \cap U_i$ then $I[f^{-1}] = \Gamma(U_i, \mathcal{O}_{U_i})[f^{-1}]$. It follows that $Q_f^{\mu}(\mathcal{J}) = Q_f^{\mu}(\mathcal{A}(U_i)) = \mathcal{A}(D(f))$. The argument being symmetric in i and j and applying to all $\mathbb{G}_m$–stable affines in an open cover of $U_i \cap U_j$, we get $\mathcal{J}_i|_{U_i \cap U_j} = \mathcal{A}|_{U_i \cap U_j} = \mathcal{J}_j|_{U_i \cap U_j}$.

Finally, assume $U_i \cap U_j \cap C \neq \varnothing$. Since $U_i \cap U_j$ is affine, it suffices to show that $\mathcal{J}_i|_{U_i \cap U_j} = \mathcal{J}_{i,j}$, where $\mathcal{J}_{i,j}$ is the left ideal in $\mathcal{A}_{i,j} := \mathcal{A}|_{U_i \cap U_j}$ generated by negative sections. Noting that $\mathcal{J}_i|_{U_i \cap U_j}$ is clearly contained in $\mathcal{J}_{i,j}$, we have

$$0 \to \mathcal{K} \to \mathcal{A}_{i,j}/(\mathcal{J}_i|_{U_i \cap U_j}) \to \mathcal{A}_{i,j}/\mathcal{J}_{i,j} \to 0.$$

By Proposition 4.12, $\mathcal{A}_{i,j}/\mathcal{J}_{i,j}$ is $\hbar$–flat, therefore tensoring by $\mathbb{C}[\![\hbar]\!]/(\hbar)$, and applying Lemma 2.4 we get

$$0 \to \mathcal{K}_0 \to \mathcal{O}_{U_i \cap U_j}/(\mathcal{I}_i|_{U_i \cap U_j}) \to \mathcal{O}_{U_i \cap U_j}/\mathcal{I}_{i,j} \to 0.$$

But this is just the sequence $0 \to \mathcal{K}_0 \to \mathcal{O}_{C_{i,j}} \to \mathcal{O}_{C_{i,j}} \to 0$, where $C_{i,j} = U_i \cap U_j \cap C$. By Nakayama's lemma, this implies that $\mathcal{K} = 0$. $\square$

Let us recall that $\pi \colon C \to S$ denotes the morphism of symplectic reduction. Since the sheaf $\mathcal{A}/\mathcal{J}$ is supported on C, we shall denote the sheaf restriction $i^{-1}(\mathcal{A}/\mathcal{J})$ simply by $\mathcal{A}/\mathcal{J}$. Under this convention, we have:

Proposition 4.27 *$\mathcal{B} = (\pi_\bullet \mathcal{E}nd_{\mathcal{A}}(\mathcal{A}/\mathcal{J}))^{\mathrm{opp}}$ is a deformation-quantization of S.*

Proof This is a local statement. Thus, the proposition follows from Corollary 4.18. □

Let $\mathcal{W}_S = \mathcal{B}[\hbar^{-1}]$; by Proposition 4.27, $\mathcal{W}_S$ is a $\mathcal{W}$–algebra on S. As we did in the affine setting in Section 4.7, define an adjoint pair $(\mathbb{H}^\perp, \mathbb{H})$ of functors of DQ modules by

$$\mathbb{H}(\mathscr{M}') = \pi_\bullet \mathcal{H}om_{\mathcal{A}}(\mathcal{A}/\mathcal{J}, \mathscr{M}'), \quad \mathbb{H}^\perp(\mathscr{N}') = \pi^{-1}(\mathcal{A}/\mathcal{J}) \otimes_{\pi^{-1}\mathcal{B}} \mathscr{N}'. \tag{4-10}$$

As in the affine setting, these functors preserve $\hbar$–torsion modules and thus descend to an adjoint pair on the $\mathcal{W}$–module categories defined by

$$\begin{aligned} \mathbb{H}(\mathscr{M}) &= \pi_\bullet \mathcal{H}om_{\mathcal{A}}(\mathcal{A}/\mathcal{J}, \mathscr{M}')[\hbar^{-1}], \\ \mathbb{H}^\perp(\mathscr{N}) &= (\pi^{-1}(\mathcal{A}/\mathcal{J}) \otimes_{\pi^{-1}\mathcal{B}} \mathscr{N}')[\hbar^{-1}], \end{aligned} \tag{4-11}$$

where $\mathscr{M}'$ and $\mathscr{N}'$ are choices of lattice in $\mathscr{M}$ and $\mathscr{N}$, respectively.

Theorem 4.28 *The adjoint functors $(\mathbb{H}^\perp, \mathbb{H})$ defined by (4-11) form a pair of exact, mutually quasi-inverse equivalences of abelian categories:*

$$\mathbb{H} \colon \mathcal{W}_{\mathfrak{X}}\text{–good}_C \rightleftarrows \mathcal{W}_S\text{–good} : \mathbb{H}^\perp$$

Proof It suffices to check locally that the canonical adjunctions $\mathrm{id} \to \mathbb{H} \circ \mathbb{H}^\perp$ and $\mathbb{H}^\perp \circ \mathbb{H} \to \mathrm{id}$ are exact isomorphisms. Therefore the theorem follows from Theorem 4.21. □

4.9 Support

The support of modules is well behaved under the functor of quantum coisotropic reduction.

Proposition 4.29 *Let $\mathscr{M} \in \mathcal{W}_{\mathfrak{X}}\text{–good}_C$ and $\mathscr{N} \in \mathcal{W}_S\text{–good}$. Then*

$$\operatorname{Supp} \mathbb{H}(\mathscr{M}) = \pi(\operatorname{Supp} \mathscr{M}), \quad \operatorname{Supp} \mathbb{H}^\perp(N) = \pi^{-1}(\operatorname{Supp} \mathscr{N}).$$

Proof Since support is a local property, we may assume that $S \hookrightarrow C = S \times V \twoheadrightarrow S$. Let N be the global sections of a lattice for $\mathscr{N}$. Since $\mathbb{H}$ is an equivalence it suffices to show that $\operatorname{Supp} \mathbb{H}^\perp(N) = \pi^{-1}(\operatorname{Supp} N)$ and $\pi(\pi^{-1}(\operatorname{Supp} N)) = \operatorname{Supp} N$. As noted in [42, Proposition 1.4.3],

$$\mathrm{gr}_\hbar(A/J \otimes_B N) = (\mathrm{gr}_\hbar A/J) \otimes^L_{B_0} (\mathrm{gr}_\hbar N).$$

Hence, using the fact that A/J is $\hbar$–flat and $H^0(\mathrm{gr}_\hbar A/J)$ is free over B_0,

$$\operatorname{Supp} \mathbb{H}^\perp(N) = \operatorname{Supp} \mathrm{gr}_\hbar(A/J \otimes_B N) = V \times \operatorname{Supp} \mathrm{gr}_\hbar(N) = V \times \operatorname{Supp} N.$$

From this, both claims are clear. □

We recall that a holonomic $\mathcal{W}_{\mathfrak{X}}$–module $\mathcal{M}$ is said to be *regular* if it admits a lattice $\mathcal{M}'$ such that the support of the $\mathcal{O}_{\mathfrak{X}}$–module $\mathcal{M}'/\hbar\mathcal{M}'$ is reduced. Proposition 4.29 implies that the functors $\mathbb{H}$ and $\mathbb{H}^{\perp}$ preserve both holonomicity and regular holonomicity:

Corollary 4.30 *The functor $\mathbb{H}$ restricts to equivalences*

$$\mathbb{H}\colon \mathcal{W}_{\mathfrak{X}}\text{–hol}_C \xrightarrow{\sim} \mathcal{W}_S\text{–hol}, \qquad \mathbb{H}\colon \mathcal{W}_{\mathfrak{X}}\text{–reghol}_C \xrightarrow{\sim} \mathcal{W}_S\text{–reghol}.$$

Proof The first claim follows directly from Proposition 4.29.

For the second claim, we may assume given a regular lattice. Then it suffices to check that applying either functor of DQ modules yields again a regular lattice; moreover, this can be checked locally. We thus revert to the affine setting of Section 4.7. Suppose first that M is an $\hbar$–torsion-free A–module for which $M/\hbar M$ has reduced support. We use the following variant of [53, Lemma 7.13], whose proof is identical:

Lemma 4.31 *Let $\mathcal{R}$ be a Noetherian, flat $\mathbb{C}[t]$–algebra (in particular, $\mathbb{C}[t]$ is central in $\mathcal{R}$). Suppose that M is an $\mathcal{R}$–module of finite type and $N^\bullet$ is a complex of $\mathbb{C}[t]$–flat $\mathcal{R}$–modules. Then, for any $a \in \mathbb{C}$,*

$$\mathbb{C}[t]/(t-a) \otimes_{\mathbb{C}[t]} \operatorname{Hom}_{\mathcal{R}}(M, N^\bullet) \cong \operatorname{Hom}_{\mathcal{R}/(t-a)\mathcal{R}}(M/(t-a)M, N^\bullet/(t-a)N^\bullet).$$

By the lemma, $\operatorname{Hom}_A(A/J, M) \otimes_{\mathbb{C}[\![\hbar]\!]} \mathbb{C}[\![\hbar]\!]/(\hbar) \cong \operatorname{Hom}_R(R/I, M/\hbar M)$. The support of the last module is the scheme-theoretic intersection of C with $\operatorname{Supp}(M/\hbar M)$, but since the latter is set-theoretically contained in C and is assumed to be reduced, this intersection is reduced. Since $\mathbb{C}[S] \subset R/I$, the annihilator of $\operatorname{Hom}_R(R/I, M/\hbar M)$ in $\mathbb{C}[S]$ is thus also a radical ideal, as required.

Suppose, on the other hand, that N is a B–module with $N/\hbar N$ reduced; we must show that $A/J \otimes N$ is also regular. For this we note the equivalence

$$(A/J \otimes_B N)/\hbar \xrightarrow{\sim} (A/J)/\hbar \otimes_{B/\hbar} (N/\hbar) \xrightarrow{\sim} (R/I) \otimes_{\mathbb{C}[S]} (N/\hbar)$$

which follows from standard base change identities. If we view the right-hand side as a module over R/I, then it follows that

$$\operatorname{ann}_{R/I}((R/I) \otimes_{\mathbb{C}[S]} (N/\hbar)) = R/I \cdot \operatorname{ann}_{\mathbb{C}[S]}(N/\hbar)$$

and since the map $T \to R/I$ is (locally) an inclusion of polynomial rings, we have that $\operatorname{ann}_{\mathbb{C}[S]}(N/\hbar)$ is reduced implies $\operatorname{ann}_{R/I}((R/I) \otimes_{\mathbb{C}[S]} (N/\hbar))$ is reduced. Finally, we have that the annihilator of $(R/I) \otimes_{\mathbb{C}[S]} (N/\hbar)$ as an R–module is the preimage under $R \to R/I$ of the annihilator as an R/I–module. The result follows because $I \subset R$ is a generated by a regular sequence. $\square$

4.10 Filtered quantizations

Using Theorem 4.28 we can reduce the study of equivariant modules supported on a smooth, closed, coisotropic subvariety $C \subseteq \mathfrak{X}$ to the situation where the set of fixed points Y of our symplectic manifold $\mathfrak{X}$ with elliptic $\mathbb{G}_m$–action is connected. In this section we study deformation-quantization algebras on such manifolds. Such algebras are equivalent to filtered quantizations, which we now define.

Suppose $\mathfrak{X}$ is a smooth symplectic variety with elliptic $\mathbb{G}_m$–action and connected fixed locus $Y = \mathfrak{X}^{\mathbb{G}_m}$. Let $\rho\colon \mathfrak{X} \to Y$ be the projection. Recall that, as in Theorem 2.21, the group scheme T^*Y acts on $\mathfrak{X}$ so that the quotient $E = \mathfrak{X}/T^*Y$ is an elliptic symplectic fibration. In this section we write $\mathcal{O}_{\mathfrak{X}}$ for the sheaf $\rho_\bullet \mathcal{O}_{\mathfrak{X}}$ of algebras on Y.

Definition 4.32 A *filtered quantization* of $\mathfrak{X}$ is a sheaf of quasicoherent $\mathcal{O}_Y$–algebras $\mathcal{D}_{\mathfrak{X}}$ equipped with an algebra filtration $\mathcal{D}_{\mathfrak{X}} = \bigcup_{i \geq 0} \mathcal{F}_i \mathcal{D}_{\mathfrak{X}}$ by coherent $\mathcal{O}_Y$–submodules and an isomorphism

$$\alpha\colon \operatorname{gr}_{\mathcal{F}} \mathcal{D}_{\mathfrak{X}} \xrightarrow{\sim} \mathcal{O}_{\mathfrak{X}} \quad \text{of } \mathcal{O}_Y\text{–algebras}$$

such that, for $D \in \mathcal{F}_i \mathcal{D}_{\mathfrak{X}}$ and $D' \in \mathcal{F}_j \mathcal{D}_{\mathfrak{X}}$, $[D, D'] \in \mathcal{F}_{i+j-l} \mathcal{D}_{\mathfrak{X}}$ defines a Poisson bracket on $\operatorname{gr}_{\mathcal{F}} \mathcal{D}_{\mathfrak{X}}$, making α an isomorphism of Poisson algebras.

Proposition 4.33 *The sheaf $\mathcal{D}_{\mathfrak{X}} := (\rho_\bullet \mathcal{W})^{\mathbb{G}_m}$ is a filtered quantization of $\mathfrak{X}$. The functor $\mathscr{M} \mapsto (\rho_\bullet \mathscr{M})^{\mathbb{G}_m}$ defines an equivalence $\mathcal{W}$–$\mathsf{good} \xrightarrow{\sim} \mathcal{D}_{\mathfrak{X}}$–$\mathsf{mod}$.*

Proof Let $\mathcal{A}$ be the DQ algebra on $\mathfrak{X}$ such that $\mathcal{W} = \mathcal{A}[\hbar^{-1}]$. We note that since $\mathbb{G}_m$ acts on $\mathcal{A}$ with positive weights, the sheaf of algebras $\widetilde{\mathcal{D}}_{\mathfrak{X}} := (\rho_\bullet \mathcal{A})^{\mathrm{rat}}$ is a polynomial quantization of $\mathfrak{X}$; ie $\widetilde{\mathcal{D}}_{\mathfrak{X}}$ is a flat sheaf of $\mathbb{C}[\hbar]$–algebras satisfying $\widetilde{\mathcal{D}}_{\mathfrak{X}}/\hbar\widetilde{\mathcal{D}}_{\mathfrak{X}} \simeq \mathcal{O}_{\mathfrak{X}}$; see [47] for a proof of this fact. From this it follows easily that $\widetilde{\mathcal{D}}_{\mathfrak{X}}/(\hbar - 1) \simeq (\widetilde{\mathcal{D}}_{\mathfrak{X}}[h^{-1}])^{\mathbb{G}_m} = \mathcal{D}_{\mathfrak{X}}$. Furthermore, via this isomorphism $\mathcal{D}_{\mathfrak{X}}$ inherits a filtration from the grading on $\widetilde{\mathcal{D}}_{\mathfrak{X}}$, inducing an isomorphism $\widetilde{\mathcal{D}}_{\mathfrak{X}} \simeq \mathsf{Rees}(\mathcal{D}_{\mathfrak{X}})$; see [47, Section 3.2] for a detailed discussion. Summing up, we see that $\mathcal{A}$ can be recovered as the $\hbar$–completion of the algebra $\mathsf{Rees}(\mathcal{D}_{\mathfrak{X}})$.

Now, for any finitely generated $\mathcal{D}_{\mathfrak{X}}$–module $\mathcal{N}$, we may choose a good filtration on $\mathcal{N}$ and obtain the $\widetilde{\mathcal{D}}_{\mathfrak{X}}$–module $\mathsf{Rees}(\mathcal{N})$. Completing at $\hbar$ gives a $\mathcal{A}$–module $\widehat{\mathsf{Rees}(\mathcal{N})}$, and it is easy to see that the $\mathcal{W}_{\mathfrak{X}}$–module $\widehat{\mathsf{Rees}(\mathcal{N})}[h^{-1}]$ doesn't depend on the choice of good filtration on $\mathcal{N}$. Now one checks directly that this is a well-defined functor which is quasi-inverse to the one above. □

When Y is a single point, $\mathfrak{X}$ is isomorphic to $\mathbb{A}^{2n}$ and a filtered quantization of $\mathfrak{X}$ is isomorphic to the Weyl algebra on $\mathbb{A}^n$, equipped with a filtration whose pieces are all finite-dimensional. In the other extreme, when $2 \dim Y = \dim \mathfrak{X}$, one gets a sheaf of twisted differential operators (we refer the reader to [3] for basics on twisted differential operators, Lie and Picard algebroids). If $0 \to \mathcal{O} \to \mathcal{P} \to \Theta_Y \to 0$ is a Picard algebroid on Y, then we denote by $U(\mathcal{P})$ the enveloping algebroid of the Picard algebroid [3]. Since $\mathcal{P}$ is locally free as an $\mathcal{O}_Y$–module, the algebra $U(\mathcal{P})$ is equipped with a canonical filtration such that $\operatorname{gr} U(\mathcal{P}) \simeq \operatorname{Sym}^\bullet \Theta_Y$.

Lemma 4.34 *When $\mathfrak{X} = T^*Y$, $\mathcal{D}_\mathfrak{X}$ is a sheaf of twisted differential operators on Y.*

Proof We construct a Picard algebroid $\mathcal{P}$ on Y and show that $\mathcal{D}_\mathfrak{X}$ is isomorphic, as a filtered algebra, to the sheaf of twisted differential operators $U(\mathcal{P})$. The assumption $\mathfrak{X} = T^*Y$ implies that $\mathcal{F}_i \mathcal{D}_Y = \mathcal{O}_Y$ for $0 \le i < l$ and we have a short exact sequence

$$0 \to \mathcal{O}_Y \to \mathcal{P} \xrightarrow{\sigma} \Theta_Y \to 0,$$

where $\mathcal{P} = \mathcal{F}_l \mathcal{D}_\mathfrak{X}$. Here $\sigma(D)(f) = [D, f]$ and $\mathcal{P}$ is closed under the commutator bracket on $\mathcal{D}_\mathfrak{X}$. Therefore we get a filtered morphism $U(\mathcal{P}) \to \mathcal{D}_\mathfrak{X}$ whose associated graded morphism is just the identity on $\operatorname{Sym}^\bullet \Theta_Y$. □

In fact, one can show that there is an equivalence of categories between twisted differential operators on Y and $\mathbb{G}_m$–equivariant deformation-quantizations of T^*Y.

4.11 Refining Kashiwara's equivalence

We continue to assume that $\mathfrak{X}$ is a smooth symplectic variety with elliptic $\mathbb{G}_m$–action and that the set of $\mathbb{G}_m$–fixed points Y in $\mathfrak{X}$ is connected. Let $Y' \subset Y$ be a closed, smooth subvariety, $i\colon Y' \hookrightarrow Y$ the embedding and $\mathcal{I}$ the sheaf of ideals in $\mathcal{O}_Y$ defining Y'. If $C' = \rho^{-1}(Y')$, then Corollary 2.27 says that C' is coisotropic and admits a reduction $C' \to S'$.

Proposition 4.35 *The sheaf $\mathcal{E}nd_{\mathcal{D}_\mathfrak{X}}(\mathcal{D}_\mathfrak{X}/\mathcal{D}_\mathfrak{X}\mathcal{I})^{\mathrm{opp}}$ is a filtered quantization of S'.*

In order to establish the above proposition we first consider the case where $\mathfrak{X} = T^*Y$. Let $\mathcal{D}_\mathfrak{X}$ be a filtered quantization of $\mathfrak{X}$. By Lemma 4.34, $\mathcal{D}_\mathfrak{X} \simeq U(\mathcal{P})$ is a sheaf of twisted differential operators on Y. The following is the analogue of Kashiwara's equivalence for twisted differential operators:

Lemma 4.36 *Let $\mathcal{P}$ be a Picard algebroid on Y and $\mathcal{D}_Y = U(\mathcal{P})$ the associated sheaf of twisted differential operators.*

(1) *The sheaf $\mathcal{P}'$ associated to the presheaf $\{s \in \mathcal{P}/\mathcal{P}\mathcal{I} \mid [\mathcal{I}, s] = 0\}$ is a Picard algebroid on Y'.*

(2) *The sheaf of twisted differential operators $U(\mathcal{P}')$ associated to the Picard algebroid $\mathcal{P}'$ is isomorphic to $\mathcal{E}nd_{\mathcal{D}_Y}(\mathcal{D}_Y/\mathcal{D}_Y\mathcal{I})^{\mathrm{opp}}$.*

(3) *The functor of Hamiltonian reduction*

$$\mathscr{M} \mapsto \mathcal{H}om_{\mathcal{D}_Y}(\mathcal{D}_Y/\mathcal{D}_Y\mathcal{I}, \mathscr{M})$$

defines an equivalence between the categories of quasicoherent $\mathcal{D}_Y$–modules supported on Y' and quasicoherent $U(\mathcal{P}')$–modules.

Proof The proof is analogous to the untwisted case; therefore we only sketch the proof. The main difference is that one must work in a formal neighborhood of each point of Y' since Picard algebroids do not trivialize in the étale topology; see [3]. One can check that $\mathcal{P}'$ is a sheaf of $\mathcal{O}_{Y'}$–modules, which is coherent since $\mathcal{P}$ is a coherent $\mathcal{O}_Y$–module. The anchor map $\sigma\colon \mathcal{P} \to \Theta_Y$ descends to a map $\sigma'\colon \mathcal{P}' \to \Theta_{Y'}$ such that $\mathcal{O}_{Y'}$ is in the kernel of this map. To show that the sequence

$$0 \to \mathcal{O}_{Y'} \to \mathcal{P}' \xrightarrow{\sigma} \Theta_{Y'} \to 0$$

is exact, it suffices to consider the sequence as a sequence of $\mathcal{O}_Y$–modules and check exactness in a formal neighborhood of each point of $Y' \subset Y$. But then $\mathcal{P}$ trivializes and the claim is clear. The other statements are analogously proved by reducing to the untwisted case in the formal neighborhood of each point of Y'. □

Proof of Proposition 4.35 Since each graded piece $(\mathcal{O}_{\mathfrak{X}})_i$ is a locally free, finite-rank $\mathcal{O}_Y$–module, it follows by induction on i that each piece $\mathcal{F}_i\mathcal{D}_{\mathfrak{X}}$ is locally free of finite rank over $\mathcal{O}_Y$. Therefore by vanishing of $\mathrm{Tor}_1^{\mathcal{O}_Y}(\mathcal{O}_{Y'}, -)$ the sequence

$$0 \to \mathcal{F}_{i-1}/\mathcal{F}_{i-1}\mathcal{I} \to \mathcal{F}_i/\mathcal{F}_i\mathcal{I} \to (\mathcal{O}_{\mathfrak{X}})_i/(\mathcal{O}_{\mathfrak{X}})_i\mathcal{I} \to 0$$

is exact. Consequently, $\mathrm{gr}_{\mathcal{F}}(\mathcal{D}_{\mathfrak{X}}/\mathcal{D}_{\mathfrak{X}}\mathcal{I}) \simeq \mathcal{O}_{\mathfrak{X}}/\langle\mathcal{I}\rangle$ and the fact that $\mathcal{D}_{\mathfrak{X}}$ quantizes the Poisson bracket on $\mathcal{O}_{\mathfrak{X}}$ implies that we have an embedding

$$\begin{aligned}\mathrm{gr}_{\mathcal{F}}\,\mathcal{E}nd_{\mathcal{D}_{\mathfrak{X}}}(\mathcal{D}_{\mathfrak{X}}/\mathcal{D}_{\mathfrak{X}}\mathcal{I})^{\mathrm{opp}} &= \mathrm{gr}_{\mathcal{F}}\big(\{s \in \mathcal{D}_{\mathfrak{X}}/\mathcal{D}_{\mathfrak{X}}\mathcal{I} \mid [\mathcal{I}, s] = 0\}\big)\\ &\hookrightarrow \{f \in \mathcal{O}_{\mathfrak{X}}/\langle\mathcal{I}\rangle \mid \{\mathcal{I}, f\} = 0\} = \mathcal{O}_{S'},\end{aligned}$$

where the final identification is (2-5). Therefore it suffices to show that the embedding is an isomorphism.

We will do this by étale base change. Let $\phi\colon U \to Y$ be an étale map and let $\mathfrak{X}' = U \times_Y \mathfrak{X}$. Assume now that U is affine, and replacing Y by the image of U we will assume that Y is too. Let $\mathcal{A}_{\mathfrak{X}}$ be the sheaf of DQ algebras on $\mathfrak{X}$ corresponding to $\mathcal{D}_{\mathfrak{X}}$ via Proposition 4.33. Then we have a $\mathbb{C}[\![\hbar]\!]$–module isomorphism $\mathcal{A}_{\mathfrak{X}} \simeq \prod_{n\ge 0}\mathcal{O}_{\mathfrak{X}}\hbar^n$. Since the multiplication in $\mathcal{A}_{\mathfrak{X}}$ is given by polydifferential operators, it uniquely extends to a multiplication structure on $\prod_{n\ge 0}\mathcal{O}_{\mathfrak{X}'}\hbar^n$, which by abuse of notation we write as $\phi^*\mathcal{A}_{\mathfrak{X}}$. This shows that filtered quantizations behave well under étale base change. Shrinking U if necessary, Proposition 2.24, together with the

Bezrukavnikov–Kaledin classification, implies that there is an equivariant isomorphism $\phi^*\mathcal{A}_{\mathfrak{X}} \simeq \mathcal{A}_{T^*U} \mathbin{\widehat{\boxtimes}} \mathcal{D}$ for some $\mathbb{G}_m$–equivariant quantization $\mathcal{A}_{T^*U}$ of T^*U. Inverting $\hbar$ and taking $\mathbb{G}_m$–invariance as in Proposition 4.33, we get an isomorphism of filtered algebras $\phi^*\mathcal{D}_{\mathfrak{X}} \simeq \mathcal{D}_{T^*U} \boxtimes \mathcal{D}(\mathbb{A}^{2n})$, where $\mathcal{D}(\mathbb{A}^{2n})$ is the usual Weyl algebra (but equipped with a particular filtration) and $\mathcal{D}_{T^*U}$ is a filtered quantization of T^*U. Let $\mathcal{I}_U$ be the ideal in $\mathcal{O}_U$ defining $\phi^{-1}(Y')$ in U. Then, since ϕ is flat,

$$\phi^*(\mathcal{D}_{\mathfrak{X}}/\mathcal{D}_{\mathfrak{X}}\mathcal{I}) \simeq (\mathcal{D}_{T^*U}/\mathcal{D}_{T^*U}\mathcal{I}_U) \boxtimes \mathcal{D}(\mathbb{A}^{n-j}).$$

Thus,

$$\phi^*\mathcal{E}nd_{\mathcal{D}_{\mathfrak{X}}}(\mathcal{D}_{\mathfrak{X}}/\mathcal{D}_{\mathfrak{X}}\mathcal{I})^{\mathrm{opp}} \simeq \mathcal{E}nd_{\mathcal{D}_{T^*U}}(\mathcal{D}_{T^*U}/\mathcal{D}_{T^*U}\mathcal{I}_U)^{\mathrm{opp}} \boxtimes \mathcal{D}(\mathbb{A}^{n-j}).$$

Lemma 4.34 says that $\mathcal{D}_{T^*U}$ is a sheaf of twisted differential operators on U. Hence, by Lemma 4.36, $\mathcal{E}nd_{\mathcal{D}_{T^*U}}(\mathcal{D}_{T^*U}/\mathcal{D}_{T^*U}\mathcal{I}_U)^{\mathrm{opp}}$ is a sheaf of twisted differential operators on $\phi^{-1}(Y')$. This completes the proof of Proposition 4.35. □

Now we may consider the category of coherent $\mathcal{D}_{\mathfrak{X}}$–modules supported on Y', or equivalently the category of good, $\mathbb{G}_m$–equivariant $\mathcal{W}_{\mathfrak{X}}$–modules whose support is contained in C'. Applying Lemma 4.36(3) and the étale local arguments of the proof of Proposition 4.35, one gets that

$$\mathbb{H}\colon \mathcal{D}_{\mathfrak{X}}\text{–}\mathsf{mod}_{Y'} \to \mathcal{D}_{S'}\text{–}\mathsf{mod}, \quad \mathscr{M} \mapsto \mathcal{H}om_{\mathcal{D}_{\mathfrak{X}}}(\mathcal{D}_{\mathfrak{X}}/\mathcal{D}_{\mathfrak{X}}\mathcal{I}, \mathscr{M}),$$

is an equivalence. The arguments involved are analogous to the proof of Theorem 4.28 and are omitted.

4.12 Generalizing Kashiwara's equivalence

Combining Theorem 4.28 with the above equivalence gives a direct generalization of Kashiwara's equivalence. Let $\mathfrak{X}$ be an arbitrary symplectic manifold with elliptic $\mathbb{G}_m$–action, equipped with a $\mathbb{G}_m$–equivariant DQ algebra $\mathcal{A}_{\mathfrak{X}}$. Fix a smooth, closed, coisotropic attracting locus $\rho\colon C \to Y$ and let $\pi\colon C \to S$ be the coisotropic reduction of C. Let $Y' \subset Y$ be a smooth, closed subvariety and set $C' = \rho^{-1}(Y')$. Proposition 4.29 implies that the equivalence $\mathbb{H}$ of Theorem 4.28 restricts to an equivalence between the category $\mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}_{C'}$ of good $\mathcal{W}_{\mathfrak{X}}$–modules supported on C' and the category $\mathcal{W}_S\text{–}\mathsf{good}_{\pi(C')}$ of good $\mathcal{W}_S$–modules supported on $\pi(C')$. By the $\mathbb{G}_m$–equivariance of π, we have $\pi(C') = \tilde{\rho}^{-1}(Y')$, where $\tilde{\rho}\colon S \to Y$ is the bundle map. By Corollary 2.27, there exists a coisotropic reduction $\pi'\colon \tilde{\rho}^{-1}(Y') \to S'$. As we have argued above in terms of filtered quantizations, the category $\mathcal{W}_S\text{–}\mathsf{good}_{\pi(C')}$ is equivalent to the category $\mathcal{W}_{S'}\text{–}\mathsf{good}$ of good $\mathcal{W}_{S'}$–modules. Thus, we have shown:

Theorem 4.37 *The category $\mathcal{W}_{\mathfrak{X}}\text{–}\mathsf{good}_{C'}$ of $\mathbb{G}_m$–equivariant good $\mathcal{W}_{\mathfrak{X}}$–modules supported on C' is equivalent to the category of $\mathbb{G}_m$–equivariant good $\mathcal{W}_{S'}$–modules.*

5 Categorical cell decomposition and applications

We use the generalized Kashiwara equivalence to show that $\mathsf{Qcoh}(\mathcal{W})$ admits a categorical cell decomposition. As a consequence we are able to calculate additive invariants of this category.

In this section, open subsets $U \subset \mathfrak{X}$ are always assumed to be $\mathbb{G}_m$–stable.

5.1 Categorical cell decomposition

Recall from Lemma 2.3 that the closure relation on coisotropic strata is a partial ordering. This provides a topology on the set $\mathcal{S}$ of indices of strata, so that a subset $K \subseteq \mathcal{S}$ is closed if and only if $i \in K$ implies that $j \in K$ for all $j \le i$.

Given a subset $K \subseteq \mathcal{S}$, we let $C_K := \bigcup_{i \in K} C_i$. When K is closed, we let $\mathsf{Qcoh}(\mathcal{W})_K$ denote the full subcategory of $\mathbb{G}_m$–equivariant objects whose support is contained in the closed set $C_K \subset \mathfrak{X}$. The open inclusion $\mathfrak{X} \smallsetminus C_K \hookrightarrow \mathfrak{X}$ is denoted by j_K.

The closed embedding $C_K \hookrightarrow \mathfrak{X}$ is denoted by i_K. For $K \subset L \subset \mathcal{S}$ closed, the inclusion functor $\mathsf{Qcoh}(\mathcal{W})_K \hookrightarrow \mathsf{Qcoh}(\mathcal{W})_L$ is denoted by $i_{K,L,*}$. We have:

Proposition 5.1 (1) *The functors $i_{K,L,*}$ have right adjoints $i^!_{K,L} = \Gamma_{K,L}$ of "submodule with support" such that the adjunction $\mathrm{id} \to i^!_{K,L} \circ i_{K,L,*}$ is an isomorphism.*

(2) *The categories $\mathsf{Qcoh}(\mathcal{W})_K$ provide a filtration of $\mathsf{Qcoh}(\mathcal{W})$, indexed by the collection of closed subsets K of $\mathcal{S}$, by localizing subcategories.*

(3) *The quotient $\mathsf{Qcoh}(\mathcal{W})_L/\mathsf{Qcoh}(\mathcal{W})_K$ is equivalent (via the canonical functor) to $\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X} \smallsetminus C_K})_{L \smallsetminus K}$.*

Proof Each module $\mathscr{M}$ in $\mathcal{W}$–good_L has a unique maximal submodule $\mathscr{M}_K$ supported on C_K. Then $i^!_{K,L}(\mathscr{M}) = \mathscr{M}_K$ defines a right adjoint to $i_{K,L,*}$ such that the adjunction $\mathrm{id} \to i^!_{K,L} \circ i_{K,L,*}$ is an isomorphism. Both functors are continuous and hence the functors extend to $\mathsf{Qcoh}(\mathcal{W})$.

Let $U = \mathfrak{X} \smallsetminus K$. It is clear that the subcategory $\mathcal{W}$–good_K of $\mathcal{W}$–good_L is a localizing subcategory, from which (2) follows.

Part (3) follows from Corollary 3.28, though this is not immediate since C_K is a union of cells, and not a single cell. The proof is an easy induction on $|K|$. If $i \in K$ is maximal, then let $K' = K \smallsetminus \{i\}$. By induction, $\mathsf{Qcoh}(\mathcal{W})_L/\mathsf{Qcoh}(\mathcal{W})_{K'}$ is equivalent to $\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X} \smallsetminus C_{K'}})_{L \smallsetminus K'}$. Under this equivalence, the full subcategory

$\mathsf{Qcoh}(\mathcal{W})_K/\mathsf{Qcoh}(\mathcal{W})_{K'}$ is sent to $\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}\smallsetminus C_{K'}})_{C_i}$. Corollary 3.28 now implies that the quotient of $\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}\smallsetminus C_{K'}})_{L\smallsetminus K'}$ by $\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}\smallsetminus C_{K'}})_{C_i}$ is canonically isomorphic to $\mathsf{Qcoh}(\mathcal{W}_{\mathfrak{X}\smallsetminus C_K})_{L\smallsetminus K}$. □

Corollary 5.2 *The category $\mathsf{Qcoh}(\mathcal{W})$ has a filtration by full, localizing subcategories whose subquotients are of the form $\mathsf{Qcoh}(\mathcal{W}_{S_i})$ for various $i \in \mathcal{S}$.*

Proof Fix $i \in \mathcal{S}$. Replacing $\mathcal{S}$ by $\{j \mid j \geq i\}$, we may assume that C_i is closed in $\mathfrak{X}$. Then the corollary follows from Proposition 5.1 and Theorem 1.8. □

Since the category $\mathsf{Qcoh}(\mathcal{W})_K$ is a Grothendieck category, it contains enough injectives. For $K \subset L \subset \mathcal{S}$ closed subsets, we let $D_K(\mathsf{Qcoh}(\mathcal{W})_L)$ denote the full subcategory of the unbounded derived category $D(\mathsf{Qcoh}(\mathcal{W})_L)$ consisting of those objects whose cohomology sheaves lie in $\mathsf{Qcoh}(\mathcal{W})_K$. It is a consequence of Proposition 5.1 and [41, Lemma 4.7] that $j_{K,L}$ defines an equivalence

$$j^*_{K,L}\colon D(\mathsf{Qcoh}(\mathcal{W})_L)/D_K(\mathsf{Qcoh}(\mathcal{W})_L) \simeq D(\mathsf{Qcoh}(\mathcal{W})_{L\smallsetminus K}). \tag{5-1}$$

Since $\mathsf{Qcoh}(\mathcal{W})_L$ has enough injectives the left exact functor $i^!_{K,L}$ can be derived to an exact functor $\mathbb{R}i^!_{K,L}\colon D(\mathsf{Qcoh}(\mathcal{W})_L) \to D(\mathsf{Qcoh}(\mathcal{W})_K)$ such that the adjunction $\mathrm{id} \to \mathbb{R}i^!_{K,L} \circ i_{K,L,*}$ is an isomorphism. We have:

Lemma 5.3 *The functor $i_{K,L,*}\colon D(\mathsf{Qcoh}(\mathcal{W})_K) \to D_K(\mathsf{Qcoh}(\mathcal{W})_L)$ is an equivalence.*

Proof The quasi-inverse to this functor is given by $\mathbb{R}i^!_{K,L}$. The already-noted isomorphism $\mathrm{id} \to \mathbb{R}i^!_{K,L} \circ i_{K,L,*}$ gives that $i_{K,L,*}$ is fully faithful. For the other direction, we note:

Claim 5.4 *The adjunction $i_{K,L,*} \circ \mathbb{R}i^!_{K,L}(\mathscr{M}) \to \mathscr{M}$ is an isomorphism for any $\mathscr{M} \in D^b_K(\mathsf{Qcoh}(\mathcal{W})_L)$.*

Proof Let $F = i_{K,L,*} \circ \mathbb{R}i^!_{K,L}$. The proof of the claim is essentially identical to the proof of [32, Corollary 1.6.2], but we provide details for the reader's convenience. As usual, let $l(\mathscr{M})$ denote the cohomological length of $\mathscr{M} \in D^b_K(\mathsf{Qcoh}(\mathcal{W})_L)$, ie it is the difference $\max\{i \mid H^i(\mathscr{M}) \neq 0\} - \min\{j \mid H^j(\mathscr{M}) \neq 0\}$. The claim will follow by induction. If $l(\mathscr{M}) = 0$, then we may assume without loss of generality that $\mathscr{M} \in \mathsf{Qcoh}(\mathcal{W})_K$. In this case the claim follows from Theorem 4.28.

In general, we choose $k \in \mathbb{Z}$ such that $l(\tau^{\leq k}\mathscr{M})$ and $l(\tau^{>k}\mathscr{M})$ are strictly less than $l(\mathscr{M})$. Applying F to the triangle $\tau^{\leq k}\mathscr{M} \to \mathscr{M} \to \tau^{>k}\mathscr{M} \xrightarrow{[1]}$ gives a commutative diagram:

$$\begin{tikzcd}
\tau^{\leq k}\mathscr{M} \arrow[r] \arrow[d, "\alpha"] & \mathscr{M} \arrow[r] \arrow[d, "\beta"] & \tau^{>k}\mathscr{M} \arrow[r, "{[1]}"] \arrow[d, "\gamma"] & {} \\
F(\tau^{\leq k}\mathscr{M}) \arrow[r] & F(\mathscr{M}) \arrow[r] & F(\tau^{>k}\mathscr{M}) \arrow[r, "{[1]}"] & {}
\end{tikzcd}$$

Since α and γ are isomorphisms by induction, so too is β. □

Since $i_{K,L,*}$ is fully faithful and t–exact with respect to the standard t–structure, it follows that $\mathbb{R}i^!_{K,L}$ has finite cohomological dimension on $D_K(\mathsf{Qcoh}(\mathcal{W})_L)$; and thus that $i_{K,L,*}\colon D^b(\mathsf{Qcoh}(\mathcal{W})_K) \to D^b_K(\mathsf{Qcoh}(\mathcal{W})_L)$ is an equivalence, with quasi-inverse $\mathbb{R}i^!_{K,L}$. Now the unbounded case follows from the bounded by noting that both functors are continuous. □

Let $\mathcal{T}^c$ be the full subcategory of compact objects in a triangulated or dg category $\mathcal{T}$.

The full subcategory of $D(\mathsf{Qcoh}(\mathcal{W}))$ consisting of all objects locally isomorphic to a bounded complex of projective objects inside $\mathcal{W}$–good (ie the perfect objects) is denoted by $\mathsf{perf}(\mathcal{W})$. To see that this is well behaved, we first note that $\mathsf{perf}(\mathcal{W})$ is contained in $D^b_{\mathcal{W}\text{–good}}(\mathsf{Qcoh}(\mathcal{W}))$; this follows from the fact that any object in $\mathsf{Qcoh}(\mathcal{W})$, being a limit of good modules, is in $\mathcal{W}$–good if and only if it is locally in $\mathcal{W}$–good. Next, we recall from [33, Lemma 2.6], that in fact $D^b(\mathcal{W}\text{–good}) \xrightarrow{\sim} D^b_{\mathcal{W}\text{–good}}(\mathsf{Qcoh}(\mathcal{W}))$. So we in fact have $\mathsf{perf}(\mathcal{W}) \subseteq D^b(\mathcal{W}\text{–good})$.

Lemma 5.5 *We have*

$$\mathsf{perf}(\mathcal{W}) = D^b(\mathcal{W}\text{–good}).$$

Moreover, when $\mathfrak{X}$ is affine,

$$\mathsf{perf}(\mathcal{W}) = D^b(\mathcal{W}\text{–good}) = D(\mathsf{Qcoh}(\mathcal{W}))^c.$$

Proof Since both $\mathsf{perf}(\mathcal{W})$ and $D^b(\mathcal{W}\text{–good})$ are locally defined full subcategories of $D(\mathsf{Qcoh}(\mathcal{W}))$, the first statement follows from the second. So we suppose $\mathfrak{X}$ is affine. In this case, Proposition 3.11 says that the category $\mathcal{A}$–mod is equivalent to A–mod. Under this equivalence the full subcategory of $\hbar$–torsion sheaves is sent to the subcategory of $\hbar$–torsion A–modules. The quotient of A–mod by this subcategory is equivalent to W–mod. Hence, since the global section functor commutes with colimits, $\mathsf{Qcoh}(\mathcal{W})$ is equivalent to $(W, \mathbb{G}_m)$–Mod. Hence $D(\mathsf{Qcoh}(\mathcal{W})) \simeq D((W, \mathbb{G}_m)\text{–Mod})$. As usual, the projective good objects in $(W, \mathbb{G}_m)$–Mod are precisely the summands of a finite graded free $\mathcal{W}$ modules. Since the category $(W, \mathbb{G}_m)$–Mod has finite global dimension and W is Noetherian, the claim is now standard. □

Lemma 5.6 *Let U be an open subset of $\mathfrak{X}$ whose complement is a union of coisotropic cells. Then, any perfect complex on U admits a perfect extension to $\mathfrak{X}$.*

Proof Write $j\colon U \hookrightarrow \mathfrak{X}$. As noted above, $\mathsf{perf}(\mathcal{W}) = D^b(\mathcal{W}\text{–good})$. Therefore it suffices to show that any bounded complex $\mathscr{M}^\bullet$ of good $\mathcal{W}_U$-modules can be extended to a bounded complex of good $\mathcal{W}$-modules. We shall construct this as a subcomplex of $j_*\mathscr{M}^\bullet$, where j_* is the right adjoint to j^* given by Corollary 3.28. Assume that $\mathscr{M}^0$ is the first nonzero term. By Theorem 3.27, there exists a good extension $\mathscr{N}^0$, which by adjunction maps naturally to $j_*\mathscr{M}^0$. Since its image is again a coherent extension of $\mathscr{M}^0$ we may assume $\mathscr{N}^0$ is a submodule of $j_*\mathscr{M}^0$. Similarly we can find some good submodule $\mathscr{E}^1$ of $j_*\mathscr{M}^1$ that extends $\mathscr{M}^1$. Let $\mathscr{N}^1$ be the sum inside $j_*\mathscr{M}^1$ of $\mathscr{E}^1$ and the image of $\mathscr{N}^0$ under the differential of $j_*\mathscr{M}^\bullet$. It is a good $\mathcal{W}$-module extending $\mathscr{M}^1$. Continuing in this fashion, the lemma is clear. □

Lemma 5.7 *Let $U \subset \mathfrak{X}$ be open. Then the perfect complexes in $D(\mathsf{Qcoh}(\mathcal{W}_U))$ are compact.*

Proof By Lemma 5.5, we already know this when U is affine. In general, the result follows from the argument given in [59, Example 1.13]. Namely, given a perfect complex $\mathscr{P}$, we consider the map of sheaves of $\mathbb{C}((\hbar))$–modules

$$\bigoplus_i \mathcal{RH}om(\mathscr{P}, \mathscr{M}_i) \to \mathcal{RH}om\Big(\mathscr{P}, \bigoplus_i \mathscr{M}_i\Big).$$

This is an isomorphism since Lemma 5.5 implies that its restriction to every $\mathbb{G}_m$–stable affine open subset of U is an isomorphism. Then, since U is a Noetherian topological space, Proposition III.2.9 of [31] implies

$$\begin{aligned}\bigoplus_i \mathrm{Hom}(\mathscr{P}, \mathscr{M}_i) &= H^0\Big(\mathfrak{X}, \bigoplus_i \mathcal{RH}om(\mathscr{P}, \mathscr{M}_i)\Big)\\ &\xrightarrow{\sim} H^0\Big(\mathfrak{X}, \mathcal{RH}om\Big(\mathscr{P}, \bigoplus_i \mathscr{M}_i\Big)\Big) = \mathrm{Hom}\Big(\mathscr{P}, \bigoplus_i \mathscr{M}_i\Big). \qquad \square\end{aligned}$$

In the following, we let $\mathcal{S}_i \subset \mathcal{S}$ be a collection of subsets such that $C_{\mathcal{S}_i} = U_i$ is open in $\mathfrak{X}$, $\mathcal{S}_i \subset \mathcal{S}_{i+1}$ and $C_{\mathcal{S}_{i+1}} \smallsetminus C_{\mathcal{S}_i}$ is a union of strata of the same dimension. We have that $C_{\mathcal{S}_0} = U_0$ is the open stratum, and that $\mathcal{S}_n = \mathcal{S}$ for $n \gg 0$, so that $U_n = \mathfrak{X}$.

Proposition 5.8 *The triangulated category $D(\mathsf{Qcoh}(\mathcal{W}_{U_i}))$ is compactly generated for all i.*

Proof Since every perfect complex is compact, it suffices to show that $D(\mathsf{Qcoh}(\mathcal{W}_{U_i}))$ is generated by its perfect complexes. The proof is by induction on i. When $i = 0$, U_0 is a single stratum and Proposition 4.33 implies that $D(\mathsf{Qcoh}(\mathcal{W})) \simeq D(\mathsf{Qcoh}(\mathcal{D}_{\mathfrak{X}}))$.

Then we have the restriction and induction functors

$$D(\mathsf{Qcoh}(\mathcal{D}_{\mathfrak{X}})) \leftrightarrows D(\mathsf{Qcoh}(\mathfrak{X}^{\mathbb{G}_m})).$$

We can argue in exactly the same way as in the case of $\mathcal{D}$–modules — see [59, Example 1.14] — to see that this category is compactly (indeed, perfectly) generated.

Next, we assume that the theorem is known for U_i, and we set $C = U_{i+1} \smallsetminus U_i$, a union of closed strata in $\mathfrak{X}$. Then Theorem 4.28 together with Proposition 4.33 implies that $D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}})_C)$ is equivalent to the direct sum of the $D(\mathsf{Qcoh}(\mathcal{D}_{S_k}))$, where the S_k are the coisotropic reductions of the strata in C. As above, these categories are generated by their subcategories of perfect objects. Since the natural functors $D(\mathsf{Qcoh}(\mathcal{D}_{S_k})) \to D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}})_C)$ take good modules to good modules, we see that $D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}})_C)$ is generated by perfect (and hence compact) objects in $D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}}))$.

Now we wish to show that $D(\mathsf{Qcoh}(\mathcal{W})_{U_{i+1}})$ is perfectly generated, given that both of the categories $D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}}))_C$ and $D(\mathsf{Qcoh}(\mathcal{W}_{U_i}))$ are generated by their perfect objects. Assume that $\mathscr{M} \in D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}}))$ is such that $\mathrm{Hom}(\mathscr{P}, \mathscr{M}) = 0$ for all perfect objects $\mathscr{P}$ in $D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}}))$. In particular, $\mathrm{Hom}(\mathscr{P}, \mathscr{M}) = 0$ for all perfect objects $\mathscr{P}$ in $D(\mathsf{Qcoh}(\mathcal{W}_{U_{i+1}})_C)$. Thus, Lemma 1.7 of [58] says that $\mathscr{M} \simeq j_{i,*} j_i^* \mathscr{M}$. Since $D(\mathsf{Qcoh}(\mathcal{W}_{U_i}))$ is perfectly generated, $j_i^* \mathscr{M} \neq 0$ implies that there is some perfect object $\mathscr{Q}$ in $D(\mathsf{Qcoh}(\mathcal{W}_{U_i}))$ such that $\mathrm{Hom}(\mathscr{Q}, j_i^* \mathscr{M}) \neq 0$. Take $\phi \neq 0$ in $\mathrm{Hom}(\mathscr{Q}, j_i^* \mathscr{M})$. By Lemma 5.6, there exists some perfect complex $\mathscr{Q}'$ on U_{i+1} whose restriction to U_i equals $\mathscr{Q}$. Then the composite $\mathscr{Q}' \to j_{i,*}\mathscr{Q} \to j_{i,*} j_i^* \mathscr{M}$ is nonzero since its restriction to U equals ϕ, and we have a contradiction. Thus, $D(\mathsf{Qcoh}(\mathcal{W})_{U_{i+1}})$ is perfectly, and hence compactly, generated, as claimed. □

Corollary 5.9 *Let $K \subset L \subset \mathcal{S}$ be closed subsets.*

(1) *The subcategory $D_K(\mathsf{Qcoh}(\mathcal{W}))$ of $D(\mathsf{Qcoh}(\mathcal{W}))$ is generated by*

$$D_K(\mathsf{Qcoh}(\mathcal{W})) \cap D(\mathsf{Qcoh}(\mathcal{W}))^c.$$

(2) *Let $U = \mathfrak{X} \smallsetminus C_K$. Then the exact functor*

$$j_{K,L}^*\colon D_L(\mathsf{Qcoh}(\mathcal{W})) \to D_{L\smallsetminus K}(\mathsf{Qcoh}(\mathcal{W}_U))$$

admits a right adjoint

$$j_{K,L,*}\colon D_{L\smallsetminus K}(\mathsf{Qcoh}(\mathcal{W}_U)) \to D_L(\mathsf{Qcoh}(\mathcal{W})).$$

Proof We first prove (1), by induction on $|K|$. We assume that $\mathcal{S}$ is totally ordered with $K = \{i \le k\}$. The case $|K| = 1$ has been done in Proposition 5.8. We define $\mathsf{perf}_K(\mathcal{W})$ to be the full subcategory of $\mathsf{perf}(\mathcal{W})$ consisting of all complexes whose cohomology is supported on C_K. Since $D_K(\mathsf{Qcoh}(\mathcal{W}))$ is a full subcategory of $D(\mathsf{Qcoh}(\mathcal{W}))$, the objects in $\mathsf{perf}_K(\mathcal{W})$ are compact in $D_K(\mathsf{Qcoh}(\mathcal{W}))$. Let $K' = \{i \le k-1\}$, so that, by induction, $D_{K'}(\mathsf{Qcoh}(\mathcal{W}))$ is generated by $\mathsf{perf}_{K'}(\mathcal{W})$. Since the objects in $\mathsf{perf}_{K'}(\mathcal{W})$ are compact in $D(\mathsf{Qcoh}(\mathcal{W}))$, Construction 1.6 of [58] says that there exist a right adjoint $j_{K',*}$ to $j^*_{K'}$. Let $\mathscr{M} \in D_K(\mathsf{Qcoh}(\mathcal{W}))$ such that $\mathrm{Hom}(\mathscr{P}, \mathscr{M}) = 0$ for all $\mathscr{P} \in \mathsf{perf}_K(\mathcal{W})$. In particular, $\mathrm{Hom}(\mathscr{P}, \mathscr{M}) = 0$ for all $\mathscr{P} \in \mathsf{perf}_{K'}(\mathcal{W})$ and hence $\mathscr{M} = j_{K,*} j^*_K \mathscr{M}$. Assume that there exists some $\mathscr{Q} \in \mathsf{perf}_{K \smallsetminus K'}(\mathcal{W}_U)$ and nonzero morphism $\phi\colon \mathscr{Q} \to j^*_{K'}\mathscr{M}$. Just as in the proof of Proposition 5.8, this implies that there is some $\mathscr{Q}' \in \mathsf{perf}(\mathcal{W})$ and nonzero $\phi'\colon \mathscr{Q}' \to \mathscr{M}$ extending ϕ. However, the fact that $\mathfrak{X} = U \sqcup C_{K'}$ and the cohomology of $\mathscr{Q}$ was assumed to be contained in $C_K \smallsetminus C_{K'} \subset U$ implies that the cohomology of $\mathscr{Q}'$ is supported on C_K, ie $\mathscr{Q}'$ belongs to $\mathsf{perf}_K(\mathcal{W})$. Thus, we conclude that $\mathrm{Hom}(\mathscr{Q}, j^*_{K'}\mathscr{M}) = 0$ for all $\mathscr{Q} \in \mathsf{perf}_{K \smallsetminus K'}(\mathcal{W}_U)$. Since $C_K \smallsetminus C_{K'}$ is a single closed stratum in U, Proposition 5.8 implies that $j^*_{K'}\mathscr{M} = 0$ and hence $\mathsf{perf}_K(\mathcal{W})$ generates $D_K(\mathsf{Qcoh}(\mathcal{W}))$.

Now we deduce part (2). By [58, Construction 1.6], a right adjoint $j_{K,*}$ to j^*_K exists. The image of $D_{L \smallsetminus K}(\mathsf{Qcoh}(\mathcal{W}_U))$ under $j_{K,*}$ is contained in $D_L(\mathsf{Qcoh}(\mathcal{W}))$ since $\mathfrak{X} = U \sqcup C_K$. Thus, $j_{K,*}$ restricted to $D_{L \smallsetminus K}(\mathsf{Qcoh}(\mathcal{W}_U))$ is a right adjoint to $j^*_{K,L}$. □

The proof of Proposition 5.8 shows that $\mathsf{perf}(\mathcal{W})$ generates $D(\mathsf{Qcoh}(\mathcal{W}))$. Since $\mathsf{perf}(\mathcal{W})$ equals $D^b(\mathcal{W}\text{–good})$ by Lemma 5.5, it is clear that $\mathsf{perf}(\mathcal{W})$ is closed under summands; the same is true of $\mathsf{perf}_K(\mathcal{W})$. Therefore [59, Theorem 2.1] implies:

Corollary 5.10 *For any closed $K \subseteq \mathcal{S}$, the compact objects in $D_K(\mathsf{Qcoh}(\mathcal{W}))$ are precisely the perfect complexes, ie $D_K(\mathsf{Qcoh}(\mathcal{W}))^c = \mathsf{perf}_K(\mathcal{W})$.*

5.2 Consequences: K–theory and Hochschild and cyclic cohomology

In this section we consider the case where $\mathfrak{X}$ has only finitely many $\mathbb{G}_m$–fixed points. The fact that $\mathcal{W}$–good admits an algebraic cell decomposition in this case (see Definition 5.11) allows us to inductively calculate K_0, and the additive invariants Hochschild and cyclic homology of $\mathsf{perf}(\mathcal{W})$.

When $\mathfrak{X}$ has isolated fixed points, the coisotropic strata C_i are affine spaces and their coisotropic reductions S_i are isomorphic as symplectic manifolds to $T^*\mathbb{A}^{t_i}$ for some t_i. Moreover, $\mathsf{Qcoh}(\mathcal{W}_{S_i}) \simeq \mathcal{D}(\mathbb{A}^{t_i})\text{–}\mathsf{Mod}$. For each $i \in \mathcal{S}$ we can form the open subsets $\geq i = \{j \mid j \ge i\}$ and $>i = \{j \mid j > i\}$.

Definition 5.11 Let $\mathcal{C}$ be an abelian category with a collection of Serre subcategories $\mathcal{C}_K$ indexed by closed subsets K in a finite poset. We say that the $\mathcal{C}_K$ form an *algebraic cell decomposition* of $\mathcal{C}$ if each subquotient $\mathcal{C}_{\le i}/\mathcal{C}_{<i}$ is equivalent to the category of modules over some Weyl algebra.

By Corollary 5.2, $\mathcal{W}$–good admits an algebraic cell decomposition. Let DG–cat$_{\mathbb{C}}$ denote the category of all small $\mathbb{C}$–linear dg categories and DG–vect$_{\mathbb{C}}$ the dg derived category of $\mathbb{C}$–vector spaces. Let L denote a $\mathbb{C}$–linear functor from DG–cat$_{\mathbb{C}}$ to DG–vect$_{\mathbb{C}}$. We are interested in the case when L admits a *localization formula*. This means that any short exact sequence of dg categories gives rise, in a natural way via L, to an exact triangle in DG–vect$_{\mathbb{C}}$.

For any quantization $\mathcal{W}$, we denote by $\mathsf{Perf}(\mathcal{W})$ the dg category of perfect complexes for $\mathcal{W}$; see [43, Section 4].

Moreover, we say that L is *even* if

$$H^i\big(\mathsf{L}(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^n)))\big) = 0$$

for all n and all odd i.

Proposition 5.12 *Suppose* L: DG–cat $\to$ DG–vect *is an even* $\mathbb{C}$*–linear functor that admits a localization formula. Then there is a (noncanonical) splitting*

$$\mathsf{L}(\mathsf{Perf}(\mathcal{W})) \cong \bigoplus_i \mathsf{L}(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^{t_i}))).$$

Proof By induction on $k := |\mathcal{S}|$, we may assume that the result is true for $U = \mathfrak{X} \smallsetminus C_K$, where $K = \{k\}$. Lemma 5.3 together with (5-1) implies that we have a short exact sequence

$$0 \to D(\mathsf{Qcoh}(\mathcal{W})_K) \to D(\mathsf{Qcoh}(\mathcal{W})) \to D(\mathsf{Qcoh}(\mathcal{W}_U)) \to 0.$$

By Theorem 4.28, we may identify $D(\mathsf{Qcoh}(\mathcal{W})_K)$ with $D(\mathsf{Qcoh}(\mathcal{W}_S))$, where S is the coisotropic reduction of C_K. This in turn can be identified with $D(\mathcal{D}(\mathbb{A}^{t_k})\text{–}\mathsf{Mod})$. By Lemma 5.6, the sequence

$$0 \to \mathsf{perf}(\mathcal{D}(\mathbb{A}^{n_k})) \to \mathsf{perf}(\mathcal{W}) \to \mathsf{perf}(\mathcal{W}_U) \to 0$$

is exact. Since all functors involved lift to the dg level, we obtain an exact sequence

$$0 \to \mathsf{Perf}(\mathcal{D}(\mathbb{A}^{n_k})) \to \mathsf{Perf}(\mathcal{W}) \to \mathsf{Perf}(\mathcal{W}_U) \to 0. \tag{5-2}$$

Applying L, we get a triangle

$$\mathsf{L}(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^{t_k}))) \to \mathsf{L}(\mathsf{Perf}(\mathcal{W})) \to \mathsf{L}(\mathsf{Perf}(\mathcal{W}_U)) \to \mathsf{L}(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^{t_k})))[1] \tag{5-3}$$

and hence a long exact sequence in cohomology. Therefore, the fact that L is even implies by induction that $H^i(\mathsf{L}(\mathsf{Perf}(\mathcal{W}))) = 0$ for all odd i, and we have short exact sequences of vector spaces

$$0 \to H^{2i}\big(\mathsf{L}(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^{t_k})))\big) \to H^{2i}(\mathsf{L}(\mathsf{Perf}(\mathcal{W}))) \to H^{2i}(\mathsf{L}(\mathsf{Perf}(\mathcal{W}_U))) \to 0$$

for all i. The claim of the proposition follows. $\square$

Next we prove Corollary 1.12. By [44, Theorem 1.5(c)], both Hochschild and cyclic homology are localizing functors. Let ϵ be a variable, given degree two. We recall that the main theorem of [70] implies $HH_*(\mathcal{D}(\mathbb{A}^n)) = \mathbb{C}\epsilon^{2n}$. Furthermore, we have $HC_*(\mathcal{D}(\mathbb{A}^n)) = \epsilon^{2n}\mathbb{C}[\epsilon]$. In particular, HH_* and HC_* are *even* localizing functors. Proposition 5.12 implies that the Hochschild and cyclic homology of $\mathsf{Perf}(\mathcal{W})$ are given by

$$HH_*(\mathsf{Perf}(\mathcal{W})) = \bigoplus_{i=1}^{k} \mathbb{C}\epsilon^{t_i} \quad \text{and} \quad HC_*(\mathsf{Perf}(\mathcal{W})) = \bigoplus_{i=1}^{k} \epsilon^{t_i}\mathbb{C}[\epsilon],$$

respectively. Therefore, in order to identify $HH_*(\mathsf{Perf}(\mathcal{W}))$ with $H_{*-\dim\mathfrak{X}}(\mathfrak{X})$ as graded vector spaces, it suffices to show that

$$H_*(\mathfrak{X}) = \epsilon^{\frac{1}{2}\dim\mathfrak{X}} \bigoplus_{i=1}^{k} \mathbb{C}\epsilon^{t_i}.$$

This follows from the BB decomposition of $\mathfrak{X}$, noting that $\dim C_i = \frac{1}{2}\dim\mathfrak{X} + 2t_i$.

5.3 The Grothendieck group of $\mathsf{Perf}(\mathcal{W})$

Finally, we turn to the proof of Corollary 1.11, which states that the Grothendieck group $K_0(\mathsf{Perf}(\mathcal{W}))$ is a free $\mathbb{Z}$–module of rank $|\mathcal{S}|$. Again, the proof is by induction on $k = |\mathcal{S}|$. Using the Bernstein filtration on $\mathcal{D}(\mathbb{A}^{t_i})$, Theorem 6.7 of [65] says that we have identifications

$$K_j(\mathcal{D}(\mathbb{A}^{t_i})) = K_j(\mathsf{Perf}\mathcal{D}(\mathbb{A}^{t_i})) \simeq K_j(\mathbb{C}) \quad \text{for all } j,$$

where $K_j(\mathcal{D}(\mathbb{A}^{t_i}))$ is the j^{th} K–group of the exact category of finitely generated projective $\mathcal{D}(\mathbb{A}^{t_i})$–modules. The higher K–groups $K_j(\mathbb{C})$ for $j \geq 1$ of $\mathbb{C}$ have been calculated by Suslin and are known to be divisible; see [62, Corollary 1.5]. Quillen's localization theorem [65, Theorem 5.5] says that the short exact sequence (5-2) induces long exact sequences

$$\cdots \to K_1(\mathsf{Perf}(\mathcal{W}_U)) \to K_0(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^{t_k})) \to K_0(\mathsf{Perf}(\mathcal{W})) \to K_0(\mathsf{Perf}(\mathcal{W}_U)) \to 0. \tag{5-4}$$

Since a divisible group is an injective $\mathbb{Z}$–module and the quotient of a divisible group is divisible, the subsequence

$$0 \to K_0(\mathsf{Perf}(\mathcal{D}(\mathbb{A}^{t_k})) \to K_0(\mathsf{Perf}(\mathcal{W})) \to K_0(\mathsf{Perf}(\mathcal{W}_U)) \to 0$$

is exact, and it follows by induction that $K_0(\mathsf{Perf}(\mathcal{W}))$ is free of rank k.

5.4 Hochschild cohomology

In many instances one can apply Van den Bergh's duality theorem to calculate the Hochschild cohomology of the category $\mathcal{W}$–good. In this subsection, we assume that

$\mathfrak{X}$ is a symplectic resolution $f\colon \mathfrak{X} \to X$, where X is an affine cone, ie X has a $\mathbb{G}_m$–action with a single attracting fixed point. Moreover, we assume that $\mathfrak{X}$ arises as the GIT quotient of a G–representation W, where G is a reductive algebraic group. Thus, $\mathfrak{X} = \mu^{-1}(0)/\!\!/_\chi G$ for an appropriate character χ of G, and $X = \operatorname{Spec}(\mathbb{C}[\mu^{-1}(0)])$. Note that in this case, Lemma 3.7 of [53] shows that Assumptions 3.2 of that paper apply once we assume (as we shall) that the moment map is flat. Let U denote the algebra of $\mathbb{G}_m$–invariant global sections of $\mathcal{W}_{\mathfrak{X}}$.

Assumption 5.13 Suppose that $f\colon \mathfrak{X} \to X$ is a symplectic resolution obtained via Hamiltonian reduction as above. Then, if $\mathfrak{g} = \operatorname{Lie}(G)$ and $\mathfrak{z} = (\mathfrak{g}/[\mathfrak{g},\mathfrak{g}])^*$, we have a natural "Duistermaat–Heckman" map $\mathfrak{z} \to H^2(\mathfrak{X})$. We assume this map is surjective.

Lemma 5.14 *If $\mathfrak{X}$ is a Nakajima quiver variety of type ADE, then Assumption 5.13 holds. Moreover, in general, for any Nakajima quiver variety, whenever the assumption holds, the natural map $\mathfrak{z} \to H^2(\mathfrak{X})$ is an isomorphism and hence the family of Hamiltonian reductions over $\mathfrak{z}$ realizes the universal deformation of $\mathfrak{X}$.*[1]

Proof In the ADE case, the Poincaré polynomial of the cohomology $H^*(\mathfrak{X})$ is known by the work of Kodera and Naoi [46]. It follows readily that $\dim H^2(\mathfrak{X}) = \dim \mathfrak{z}$, and hence it suffices to check that the Duistermaat–Heckman map is injective. While this can be checked directly in these cases, since the Duistermaat–Heckman map can be realized as (a graded component of) a natural map from the center of a quiver Hecke algebra to the cohomology of the quiver variety, it follows from a recent result of Webster [69] that this map is always injective. This completes the proof of the lemma. □

Recall that for any $c \in \mathfrak{z}$ we may define the quantum Hamiltonian reduction of the Weyl algebra associated to W. We denote this algebra by U_c.

Lemma 5.15 *Let $\mathfrak{X}$ be a conic symplectic resolution as above with a $\mathbb{G}_m$–equivariant quantization $\mathcal{W}_{\mathfrak{X}}$. Suppose that Assumption 5.13 holds for $\mathfrak{X}$. Then the filtered quantization U of X associated to $\mathcal{W}$ is of the form U_c for some $c \in \mathfrak{z}$.*

Proof In [49] it is shown that, via the Bezrukavnikov–Kaledin noncommutative period map, graded quantizations of $\mathfrak{X}$ are parametrized by $H^2(\mathfrak{X}, \mathbb{C})$. Moreover, one can show that, provided Assumption 5.13 holds, this period map for quantum Hamiltonian reductions U_c for $c \in \mathfrak{z}$ is realized by the Duistermaat–Heckman map, up to a shift corresponding to the canonical quantization. Since the Rees construction gives an equivalence between filtered and graded quantizations, it follows from this that any

[1] Added in proof: If $\mathfrak{X}$ is a Nakajima quiver variety, then the main theorem of the recent preprint [54] implies that the canonical map $\mathfrak{z} \to H^2(\mathfrak{X})$ is surjective.

filtered quantization is of the form U_c for some $c \in \mathfrak{z}$. As the $\mathbb{G}_m$–invariant global sections U of our quantization of $\mathfrak{X}$ gives such a filtered quantization, it is thus of the form U_c for some $c \in \mathfrak{z}$. □

For the rest of this section we fix $c \in \mathfrak{z}$ such that $U \cong U_c$.

Lemma 5.16 *Let W be a G–representation as above, and $\chi\colon G \to \mathbb{G}_m$. Then $W \oplus W$ is a $G \times G$–representation, and the (χ,χ)–unstable locus contains $W^{\chi\text{–un}} \times W^{\chi\text{–un}}$.*

Proof Let $W \times \mathbb{C}_\chi$ be the G–representation given by $g(w,z) = (g(w), \chi(g)\cdot z)$. By the Hilbert–Mumford criterion, a point $x \in W$ is unstable if there is a one-parameter subgroup $\lambda\colon \mathbb{G}_m \to G$ such that $\lim_{t\to 0} \lambda(t)(x,1) = (0,0)$ for $(x,1) \in W \times \mathbb{C}_\chi$, and similarly for $(x,y) \in W \oplus W$. Thus, if x and y are both in $W^{\chi\text{–un}}$, destabilized by one-parameter subgroups λ_1 and λ_2, respectively, it is clear that (λ_1,λ_2) gives a one-parameter subgroup of $G \times G$ which destabilizes (x,y), and we are done. □

Lemma 5.17 *Let $c\colon \mathfrak{g} \to \mathbb{C}$ be a character of $\mathfrak{g}$ for which the algebra $U = U_c$ has finite cohomological dimension. Then U_c is smooth; that is, U^e has finite global dimension.*

Proof First note that since U_c is (left or right) Noetherian, its global dimension is its Tor dimension, and hence it has finite global dimension if and only if U_c^{op} has finite global dimension. Moreover, it follows from the construction of the noncommutative period map in [8] and the Duistermaat–Heckman theorem (see for example [49]) that the algebra U_c^{op} is isomorphic to $U_{\rho-c}$, where $\frac{1}{2}\rho$ is the character corresponding to the canonical quantum moment map, ie the quantum moment map which yields the canonical quantization of $\mathfrak{X}$.

To see that $U^e = U_c \otimes U_c^{\mathrm{op}} \cong U_c \otimes U_{\rho-c}$ has finite global dimension, first note that

$$U_c \otimes U_{\rho-c} \cong \Gamma(\mathfrak{X}\times\mathfrak{X}, \mathcal{E}_{\mathfrak{X},c} \boxtimes \mathcal{E}_{\mathfrak{X},\rho-c})$$

or, in the notation of [53], $(M_c \boxtimes M_{\rho-c})^{G\times G}$. Now, since $\mathfrak{X}$ is smooth, Corollary 7.6 of [53] shows that $U_c \otimes U_{c-\rho}$ has finite global dimension if and only if the pullback functor $\mathbb{L} f^*$ is cohomologically bounded. Explicitly, $\mathbb{L} f^*$ is the functor from $D(U_c \otimes U_{c-\rho}\text{–mod})$ to $D_{G,(c,\rho-c)}(A\otimes A\text{–mod})$ given by

$$N \mapsto \pi((M_c \boxtimes M_{\rho-c}) \otimes^{\mathbb{L}}_{U_c\otimes U_{\rho-c}} N), \quad N \in D(U_c \boxtimes U_{\rho-c}\text{–mod}),$$

where π is the quotient functor given by the (χ,χ)–unstable locus.

Choose free U_c and $U_{\rho-c}$ resolutions $P^\bullet$ of M_c and $Q^\bullet$ of $M_{\rho-c}$, respectively. By [53, Corollary 7.6], each of $\pi(P^\bullet)$ and $\pi(Q^\bullet)$ has only finitely many cohomologies, ie for each of $P^\bullet$ and $Q^\bullet$, all but finitely many cohomologies have associated graded supported in $W^{\chi\text{–un}}$. Hence the resolution $\mathrm{Tot}(P^\bullet \boxtimes Q^\bullet)$ of $M_c \boxtimes M_{\rho-c}$ has all but finitely many cohomologies with associated graded supported in $W^{\chi\text{–un}} \times W^{\chi\text{–un}}$. Thus, using Lemma 5.16 it follows the cohomologies are (χ, χ)–unstably supported after finitely many terms also. Thus $\mathbb{L} f^*$ is bounded and hence U^e has finite global dimension, as required. □

Remark 5.18 In a number of examples, such as the Hilbert scheme of points in $\mathbb{C}^2$ or the minimal resolution of $\mathbb{C}^2/\Gamma_l$ the cyclic quotient singularity, it is known explicitly when the algebra U_c has finite global dimension; moreover it is shown in [13], building on work of Kaledin, that the algebra U_c has finite global dimension for sufficiently generic c.

Recall that an algebra U is said to have finite Hochschild dimension (or is *smooth*) if U has a finite resolution when considered as a $U^e = U \otimes U^{\mathrm{opp}}$–module.

Proposition 5.19 *Assume that U_c has finite global dimension. Then*

$$HH^*(\mathsf{Perf}(\mathcal{W})) = HH^*(U_c) = H^*(\mathfrak{X}, \mathbb{C}).$$

Proof By [53, Lemma 3.14], the algebra U_c is Auslander Gorenstein with rigid Auslander dualizing complex $\mathbb{D}^* = U_c$. By Lemma 5.17, the enveloping algebra U^e of U_c has finite global dimension, and hence U_c has finite Hochschild dimension, thus we are able to apply Van den Bergh's duality result [67, Theorem 1] to conclude that $HH^*(U_c) = HH_{\dim \mathfrak{X} - *}(U_c)$. Since U_c has finite global dimension, Theorem 1.1 of [53] and Corollary 1.12 show that

$$HH^*(\mathsf{Perf}(\mathcal{W})) = HH^*(U_c) = HH_{\dim \mathfrak{X} - *}(U_c) = H_{\dim \mathfrak{X} - *}(\mathfrak{X}),$$

where in the last equality we use the fact that the degrees in Borel–Moore homology are twice those in Hochschild homology. On the other hand, Poincaré duality [14, Equation (2.6.2)] says that $H_{\dim \mathfrak{X} - *}(\mathfrak{X}) = H^*(\mathfrak{X})$, and so the result follows. □

Presumably one can also apply the results of [19] and [42, Section 2.5] to the category $\mathsf{Perf}(\mathcal{W})$ in order to express directly the Hochschild cohomology of that category in terms of its Hochschild homology.

More generally, the proof of Proposition 5.19 shows that if $\mathfrak{X}$ is a symplectic resolution $f \colon \mathfrak{X} \to X$ of an affine cone X, the number of $\mathbb{G}_m$–fixed points on $\mathfrak{X}$ is finite, and U_c is a quantization of $\mathbb{C}[X]$ such that derived localization holds and the enveloping algebra U^e has finite global dimension, then $HH^*(U_c) = H^*(\mathfrak{X}, \mathbb{C})$. We conclude with a number of standard examples, arising from representation theory, where Proposition 5.19, or the above more general statement, is applicable.

Example 5.20 Let Γ be a cyclic group and $\mathfrak{S}_n \wr \Gamma$ the wreath product group that acts as a symplectic reflection group on $\mathbb{C}^{2n}$. The corresponding symplectic reflection algebra at $t = 1$ and parameter $\boldsymbol{c}$ is denoted by $\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)$. Define an increasing filtration $F_\bullet(\mathbb{C}[\mathfrak{S}_n \wr \Gamma])$ on the group algebra of $\mathfrak{S}_n \wr \Gamma$ by letting $F_i(\mathbb{C}[\mathfrak{S}_n \wr \Gamma])$ for $i \geq 0$ be the subspace spanned by all elements $g \in \mathfrak{S}_n \wr \Gamma$ such that $\operatorname{rk}(1-g) \leq k$, where $1-g$ is thought of as an endomorphism of $\mathbb{C}^{2n}$. This is an algebra filtration and restricts to a filtration $F_\bullet(\mathsf{Z}\mathfrak{S}_n \wr \Gamma)$ on the center of the group algebra. The following result, which was proved for generic $\boldsymbol{c}$ in [22, Theorem 1.8] for generic $\boldsymbol{c}$, follows easily from the results of this paper.

Proposition 5.21 *Assume that $\boldsymbol{c}$ is spherical. Then*

$$HH^*(\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)) = HH_{2n-*}(\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)) = \operatorname{gr}^F_*(\mathsf{Z}\mathfrak{S}_n \wr \Gamma),$$

as graded vector spaces.

In [22], it is shown that the identification $HH^*(\mathsf{H}_{\boldsymbol{c}}(\mathfrak{S}_n \wr \Gamma)) = \operatorname{gr}^F_*(\mathsf{Z}\mathfrak{S}_n \wr \Gamma)$ is as graded algebras.

Example 5.22 Let G be a connected, semisimple, complex Lie group and $\mathfrak{g}$ its Lie algebra. Fix a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ and let W be the Weyl group of G. Let $\mathcal{N}$ denote the nilpotent cone in $\mathfrak{g}$. The Springer resolution of $\mathcal{N}$ is $\pi\colon T^*\mathcal{B} \to \mathcal{N}$, where $\mathcal{B}$ is the flag variety. We fix $e \in \mathcal{N}$. Associated to e is a Slodowy slice $e \in S \subset \mathfrak{g}$. The intersection $S_0 := S \cap \mathcal{N}$ is a conic symplectic singularity, and the restriction of π defines a symplectic resolution $\widetilde{S}_0 := \pi^{-1}(S_0) \to S_0$. Quantizations of S_0 are given by finite W–algebras, which we denote by $A_\lambda(e)$ to avoid confusion with our notation for DQ algebras. Here $\lambda \in \mathfrak{h}^*$. Notice that $\mathcal{B} \cap \widetilde{S}_0$ is the Springer fiber $\mathcal{B}_e$ of e.

Let $\mathfrak{l} \subset \mathfrak{g}$ be a minimal Levi subalgebra containing e. Recall that the element e is said to be *of standard Levi type* if it is regular in $\mathfrak{l}$ (this is independent of the choice of $\mathfrak{l}$). In type A, every nilpotent element is of standard Levi type.

Proposition 5.23 *Let e be of standard Levi type and $\lambda \in \mathfrak{h}^*$ regular. Then*

$$HH^*(A_\lambda(e)) \simeq H^*(\mathcal{B}_e).$$

Proof If e is of standard Levi type then it follows from [23, Proposition 1] that there is a one-parameter subgroup $H \subset G$ acting on $\widetilde{S}_0$ such that $\widetilde{S}_0^H$ is finite. Since this action is Hamiltonian, we may assume by twisting that the elliptic action of $\mathbb{G}_m$ on $\widetilde{S}_0$ has only finitely many fixed points.

Therefore it suffices to check that, when λ is dominant regular, localization holds (ie there is a DQ algebra $\mathcal{W}_\lambda$ on $\widetilde{S}_0$ such that $A_\lambda(e)\text{–}\mathsf{mod} \simeq \mathcal{W}_\lambda\text{–}\mathsf{good}$) and the enveloping algebra of $A_\lambda(e)$ has finite global dimension. This is well known so we give the appropriate references. The statement about localization follows from [18, Theorem 6.5] or [25, Theorem 6.3.2], and the fact that the enveloping algebra of $A_\lambda(e)$ has finite global dimension is a consequence of [25, Proposition 6.5.1] and the fact that the abelian category $\mathcal{W}$–good has finite global dimension for any DQ algebra $\mathcal{W}$ on $\widetilde{S}_0 \times \widetilde{S}_0$. □

In particular, when $e = 0$, we have $S = \mathfrak{g}$ and hence $S \cap \mathcal{N} = \mathcal{N}$ and $\widetilde{S}_0 = T^*\mathcal{B}$. In this case $A_\lambda(e)$ is a primitive central quotient of the enveloping algebra $U(\mathfrak{g})$ and our result recovers a result of Soergel [66], who also used localization (but coupled with the Riemann–Hilbert correspondence).

Example 5.24 Let $M(r,n)$ be a framed moduli space of torsion-free sheaves on $\mathbb{P}^2$ with rank r and second Chern character $c_2 = n$. Specifically, $M(r,n)$ parametrizes isomorphism classes of pairs (E,ϕ) such that:

(1) The sheaf E is torsion-free of rank r and $\langle c_2(E), [\mathbb{P}^2]\rangle = n$.

(2) The sheaf E is locally free in a neighborhood of ℓ^∞, with fixed isomorphism $\phi\colon E|_{\ell^\infty} \xrightarrow{\sim} \mathcal{O}_{\ell^\infty}^{\oplus r}$.

Here $\ell^\infty = \{[0 : z_1 : z_2] \in \mathbb{P}^2\}$ is the line at infinity. The space $M(r,n)$ is isomorphic to the quiver variety associated to the framed Jordan quiver, with dimension vector (r,n); see [57]. Let $M_0^{\mathrm{reg}}(r,n)$ be the open subset of locally free sheaves. The space $M(r,n)$ is a symplectic resolution of $M_0(r,n)$, where the latter is the Uhlenbeck partial compactification of $M_0^{\mathrm{reg}}(r,n)$. Quantizations $\mathcal{A}_c(r,n)$, for $c \in \mathbb{C}$, of $M_0(r,n)$ have been studied in [50].

Proposition 5.25 *Assume that c is not of the form s/m, where $1 < m < n$ and $-rm < s < 0$. Then*

$$HH^*(\mathcal{A}_c(r,n)) = H^*(M(n,r),\mathbb{C}).$$

Proof This follows from Proposition 5.19 and [50, Theorem 1.1], once one knows that $M(r,n)$ has finitely many fixed points under $\mathbb{G}_m$. But this follows from [57, Theorem 3.7], which shows that there is a natural torus T acting by Hamiltonian automorphisms on $M(r,n)$ such that the set $M(r,n)^{\mathsf{T}}$ is finite. □

The above proposition implies that the graded dimension of $HH^*(\mathcal{A}_c(r,n))$ has a concise expression in terms of r–multipartitions of n; see [57, Theorem 3.8].

Example 5.26 Our final example is slices in affine Grassmannians. We follow [36; 37]. Let G be a complex semisimple Lie group and $\mathbf{Gr} = G((t^{-1}))/G[t]$ its thick affine Grassmannian. For any given pair of dominant coweights $\lambda \geq \mu$, we have Schubert subvarieties $\mathbf{Gr}^{\lambda}$ and $\mathbf{Gr}^{\mu}$ of $\mathbf{Gr}$ such that $\mathbf{Gr}^{\mu} \subset \overline{\mathbf{Gr}^{\lambda}}$. The intersection $\mathbf{Gr}^{\lambda}_{\mu} := \overline{\mathbf{Gr}^{\lambda}} \cap \mathbf{Gr}_{\mu}$, with $\mathbf{Gr}_{\mu}$ an orbit for the first congruence subgroup $G_1[\![t^{-1}]\!]$ of $G[\![t^{-1}]\!]$, is called the Lusztig slice. It is a finite-dimensional affine conic symplectic singularity. If λ is a sum of miniscule coweights, then it is shown in [37, Theorem 2.9] that $\mathbf{Gr}^{\lambda}_{\mu}$ admits a symplectic resolution $\widetilde{\mathbf{Gr}}^{\lambda}_{\mu}$, given by closed convolution of Schubert varieties associated to miniscule weights. If $\mathsf{T} \subset G$ is a maximal torus, then T acts Hamiltonian on both $\widetilde{\mathbf{Gr}}^{\lambda}_{\mu}$ and $\mathbf{Gr}^{\lambda}_{\mu}$ such that the resolution morphism is equivariant. It has been shown in [36, Lemma 4.4] that $(\widetilde{\mathbf{Gr}}^{\lambda}_{\mu})^{\mathsf{T}}$ is finite. Therefore, one can twist the elliptic action of $\mathbb{G}_m$ on $\widetilde{\mathbf{Gr}}^{\lambda}_{\mu}$ so that it has only finitely many fixed points. Thus, Proposition 5.19 implies that if U is a quantization of $\mathbb{C}[\mathbf{Gr}^{\lambda}_{\mu}]$ such that U^e has finite global dimension and derived localization hold, then

$$HH^*(U) \simeq H^*(\widetilde{\mathbf{Gr}}^{\lambda}_{\mu}, \mathbb{C}).$$

Conjecturally, any such quantization is given by a quotient $Y^{\mu}_{\lambda}(\boldsymbol{c})$ of a shifted Yangian Y^{μ}; see [37, Conjecture 4.11].

References

[1] **K Ardakov**, **S Wadsley**, *On irreducible representations of compact p–adic analytic groups*, Ann. of Math. 178 (2013) 453–557 MR

[2] **M Artin**, *Algebraic approximation of structures over complete local rings*, Inst. Hautes Études Sci. Publ. Math. 36 (1969) 23–58 MR

[3] **A Beĭlinson**, **J Bernstein**, *A proof of Jantzen conjectures*, from "I M Gel'fand seminar" (S Gindikin, editor), Adv. Soviet Math. 16, Amer. Math. Soc., Providence, RI (1993) 1–50 MR

[4] **G Bellamy**, **C Dodd**, **K McGerty**, **T Nevins**, *Lagrangian skeleta and Koszul duality on bionic symplectic varieties*, in preparation

[5] **G Bellamy**, **T Kuwabara**, *On deformation quantizations of hypertoric varieties*, Pacific J. Math. 260 (2012) 89–127 MR

[6] **P Berthelot**, *$\mathscr{D}$–modules arithmétiques, I: Opérateurs différentiels de niveau fini*, Ann. Sci. École Norm. Sup. 29 (1996) 185–272 MR

[7] **R Bezrukavnikov**, **D Kaledin**, *Fedosov quantization in algebraic context*, Mosc. Math. J. 4 (2004) 559–592 MR

[8] **R V Bezrukavnikov**, **D B Kaledin**, *McKay equivalence for symplectic resolutions of quotient singularities*, Tr. Mat. Inst. Steklova 246 (2004) 20–42 MR In Russian; translated in Proc. Steklov Inst. Math. 246 (2004) 13–33

[9] **R Bezrukavnikov**, **I Losev**, *Etingof conjecture for quantized quiver varieties*, preprint (2013) arXiv

[10] **A Białynicki-Birula**, *Some theorems on actions of algebraic groups*, Ann. of Math. 98 (1973) 480–497 MR

[11] **A Białynicki-Birula**, *Some properties of the decompositions of algebraic varieties determined by actions of a torus*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 24 (1976) 667–674 MR

[12] **T Braden**, **A Licata**, **N Proudfoot**, **B Webster**, *Quantizations of conical symplectic resolutions, II: Category $\mathcal{O}$ and symplectic duality*, preprint (2014) arXiv

[13] **T Braden**, **N Proudfoot**, **B Webster**, *Quantizations of conical symplectic resolutions, I: Local and global structure*, preprint (2012) arXiv

[14] **N Chriss**, **V Ginzburg**, *Representation theory and complex geometry*, Birkhäuser, Boston (1997) MR

[15] **A D'Agnolo**, **M Kashiwara**, *On quantization of complex symplectic manifolds*, Comm. Math. Phys. 308 (2011) 81–113 MR

[16] **A D'Agnolo**, **P Polesello**, *Deformation quantization of complex involutive submanifolds*, from "Noncommutative geometry and physics" (Y Maeda, N Tose, N Miyazaki, S Watamura, D Sternheimer, editors), World Sci., Hackensack, NJ (2005) 127–137 MR

[17] **A D'Agnolo**, **P Schapira**, *Quantization of complex Lagrangian submanifolds*, Adv. Math. 213 (2007) 358–379 MR

[18] **C Dodd**, **K Kremnizer**, *A localization theorem for finite W–algebras*, preprint (2009) arXiv

[19] **V Dolgushev**, *The Van den Bergh duality and the modular symmetry of a Poisson variety*, Selecta Math. 14 (2009) 199–228 MR

[20] **J-P Dufour**, **N T Zung**, *Poisson structures and their normal forms*, Progress in Mathematics 242, Birkhäuser, Basel (2005) MR

[21] **D Eisenbud**, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics 150, Springer, New York (1995) MR

[22] **P Etingof**, **V Ginzburg**, *Symplectic reflection algebras, Calogero–Moser space, and deformed Harish-Chandra homomorphism*, Invent. Math. 147 (2002) 243–348 MR

[23] **L Fresse**, *Borel subgroups adapted to nilpotent elements of standard Levi type*, Represent. Theory 18 (2014) 341–360 MR

[24] **V Ginsburg**, *Characteristic varieties and vanishing cycles*, Invent. Math. 84 (1986) 327–402 MR

[25] **V Ginzburg**, *Harish-Chandra bimodules for quantized Slodowy slices*, Represent. Theory 13 (2009) 236–271 MR

[26] **I G Gordon**, **I Losev**, *On category $\mathcal{O}$ for cyclotomic rational Cherednik algebras*, J. Eur. Math. Soc. 16 (2014) 1017–1079 MR

[27] **I G Gordon**, **J T Stafford**, *The Auslander–Gorenstein property for $\mathbb{Z}$–algebras*, J. Algebra 399 (2014) 102–130 MR

[28] **A Grothendieck**, *Eléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, I*, Inst. Hautes Études Sci. Publ. Math. 11 (1961) 5–167 MR

[29] **V Guillemin**, **S Sternberg**, *Symplectic techniques in physics*, 2nd edition, Cambridge University Press (1990) MR

[30] **R Hartshorne**, *On the De Rham cohomology of algebraic varieties*, Inst. Hautes Études Sci. Publ. Math. 45 (1975) 5–99 MR

[31] **R Hartshorne**, *Algebraic geometry*, Graduate Texts in Mathematics 52, Springer, New York (1977) MR

[32] **R Hotta**, **K Takeuchi**, **T Tanisaki**, *D–modules, perverse sheaves, and representation theory*, Progress in Mathematics 236, Birkhäuser, Boston, MA (2008) MR

[33] **D Huybrechts**, **E Macrì**, **P Stellari**, *Formal deformations and their categorical general fibre*, Comment. Math. Helv. 86 (2011) 41–71 MR

[34] **L Illusie**, *Complexe cotangent et déformations, I*, Lecture Notes in Mathematics 239, Springer, Berlin (1971) MR

[35] **D Kaledin**, *Derived equivalences by quantization*, Geom. Funct. Anal. 17 (2008) 1968–2004 MR

[36] **J Kamnitzer**, **P Tingley**, **B Webster**, **A Weekes**, **O Yacobi**, *Highest weights for truncated shifted Yangians and product monomial crystals*, preprint (2015) arXiv

[37] **J Kamnitzer**, **B Webster**, **A Weekes**, **O Yacobi**, *Yangians and quantizations of slices in the affine Grassmannian*, Algebra Number Theory 8 (2014) 857–893 MR

[38] **M M Kapranov**, *On DG-modules over the de Rham complex and the vanishing cycles functor*, from "Algebraic geometry" (S Bloch, I Dolgachev, W Fulton, editors), Lecture Notes in Math. 1479, Springer, Berlin (1991) 57–86 MR

[39] **M Kashiwara**, *Quantization of contact manifolds*, Publ. Res. Inst. Math. Sci. 32 (1996) 1–7 MR

[40] **M Kashiwara**, **R Rouquier**, *Microlocalization of rational Cherednik algebras*, Duke Math. J. 144 (2008) 525–573 MR

[41] **M Kashiwara**, **P Schapira**, *Categories and sheaves*, Grundl. Math. Wissen. 332, Springer, Berlin (2006) MR

[42] **M Kashiwara**, **P Schapira**, *Deformation quantization modules*, Astérisque 345, Soc. Math. France, Paris (2012) MR

[43] **B Keller**, *On the cyclic homology of ringed spaces and schemes*, Doc. Math. 3 (1998) 231–259 MR

[44] **B Keller**, *On the cyclic homology of exact categories*, J. Pure Appl. Algebra 136 (1999) 1–56 MR

[45] **F Knop**, *Weyl groups of Hamiltonian manifolds, I*, preprint (1997) arXiv

[46] **R Kodera**, **K Naoi**, *Loewy series of Weyl modules and the Poincaré polynomials of quiver varieties*, Publ. Res. Inst. Math. Sci. 48 (2012) 477–500 MR

[47] **I Losev**, *Quantized symplectic actions and W–algebras*, J. Amer. Math. Soc. 23 (2010) 35–59 MR

[48] **I Losev**, *Finite-dimensional representations of W–algebras*, Duke Math. J. 159 (2011) 99–143 MR

[49] **I Losev**, *Isomorphisms of quantizations via quantization of resolutions*, Adv. Math. 231 (2012) 1216–1270 MR

[50] **I Losev**, *Etingof conjecture for quantized quiver varieties, II: Affine quivers*, preprint (2014) arXiv

[51] **I Losev**, *On categories O for quantized symplectic resolutions*, preprint (2015) arXiv

[52] **J C McConnell**, **J C Robson**, *Noncommutative Noetherian rings*, revised edition, Graduate Studies in Mathematics 30, Amer. Math. Soc., Providence, RI (2001) MR

[53] **K McGerty**, **T Nevins**, *Derived equivalence for quantum symplectic resolutions*, Selecta Math. 20 (2014) 675–717 MR

[54] **K McGerty**, **T Nevins**, *Kirwan surjectivity for quiver varieties*, preprint (2016) arXiv

[55] **D Nadler**, *Fukaya categories as categorical Morse homology*, Symmetry Integrability Geom. Methods Appl. 10 (2014) art. id. 018 MR

[56] **H Nakajima**, *Lectures on Hilbert schemes of points on surfaces*, University Lecture Series 18, Amer. Math. Soc., Providence, RI (1999) MR

[57] **H Nakajima**, **K Yoshioka**, *Instanton counting on blowup, I: 4–dimensional pure gauge theory*, Invent. Math. 162 (2005) 313–355 MR

[58] **A Neeman**, *The connection between the K–theory localization theorem of Thomason, Trobaugh and Yao and the smashing subcategories of Bousfield and Ravenel*, Ann. Sci. École Norm. Sup. 25 (1992) 547–566 MR

[59] **A Neeman**, *The Grothendieck duality theorem via Bousfield's techniques and Brown representability*, J. Amer. Math. Soc. 9 (1996) 205–236 MR

[60] **R Nest**, **B Tsygan**, *Deformations of symplectic Lie algebroids, deformations of holomorphic symplectic structures, and index theorems*, Asian J. Math. 5 (2001) 599–635 MR

[61] **R Nest**, **B Tsygan**, *Remarks on modules over deformation quantization algebras*, Mosc. Math. J. 4 (2004) 911–940 MR

[62] **C Pedrini**, **C Weibel**, *The higher K–theory of a complex surface*, Compositio Math. 129 (2001) 239–271 MR

[63] **F Petit**, *DG affinity of DQ-modules*, Int. Math. Res. Not. 2012 (2012) 1414–1438 MR

[64] **A Premet**, *Special transverse slices and their enveloping algebras*, Adv. Math. 170 (2002) 1–55 MR

[65] **D Quillen**, *Higher algebraic K–theory, I*, from "Algebraic K–theory, I: Higher K–theories" (H Bass, editor), Lecture Notes in Math. 341, Springer, Berlin (1973) 85–147 MR

[66] **W Soergel**, *The Hochschild cohomology ring of regular maximal primitive quotients of enveloping algebras of semisimple Lie algebras*, Ann. Sci. École Norm. Sup. 29 (1996) 535–538 MR

[67] **M Van den Bergh**, *A relation between Hochschild homology and cohomology for Gorenstein rings*, Proc. Amer. Math. Soc. 126 (1998) 1345–1348 MR

[68] **F Van Oystaeyen**, **R Sallam**, *A microstructure sheaf and quantum sections over a projective scheme*, J. Algebra 158 (1993) 201–225 MR

[69] **B Webster**, *Centers of KLR algebras and cohomology rings of quiver varieties*, preprint (2015) arXiv

[70] **M Wodzicki**, *Cyclic homology of differential operators*, Duke Math. J. 54 (1987) 641–647 MR

[71] **A Yekutieli**, *On flatness and completion for infinitely generated modules over Noetherian rings*, Comm. Algebra 39 (2011) 4221–4245 MR

GB: *School of Mathematics and Statistics, University of Glasgow*
15 University Gardens, Glasgow, G12 8QW, United Kingdom

CD and TN: *Department of Mathematics, University of Illinois at Urbana-Champaign*
1409 W Green Street, MC-382, Urbana, IL 61801, United States

KM: *Mathematical Institute, University of Oxford*
Andrew Wiles Building, Woodstock Road, Oxford, OX2 6GG, United Kingdom

gwyn.bellamy@glasgow.ac.uk, csdodd2@illinois.edu,
mcgerty@maths.ox.ac.uk, nevins@illinois.edu

Proposed: Richard Thomas — Received: 7 May 2015
Seconded: Dan Abramovich, Frances Kirwan — Revised: 23 May 2016

Geometry & Topology 21 (2017) 2683–2723

The nonuniqueness of the tangent cones at infinity of Ricci-flat manifolds

KOTA HATTORI

Colding and Minicozzi established the uniqueness of the tangent cones at infinity of Ricci-flat manifolds with Euclidean volume growth where at least one tangent cone at infinity has a smooth cross section. In this paper, we raise an example of a Ricci-flat manifold implying that the assumption for the volume growth in the above result is essential. More precisely, we construct a complete Ricci-flat manifold of dimension 4 with non-Euclidean volume growth that has infinitely many tangent cones at infinity where one of them has a smooth cross section.

53C23

1 Introduction

For a complete Riemannian manifold (X, g) with nonnegative Ricci curvature, it is shown by Gromov's compactness theorem that if one takes a sequence

$$a_1 > a_2 > \cdots > a_i > \cdots > 0$$

such that $\lim_{i\to\infty} a_i = 0$, then there is a subsequence $\{a_{i(j)}\}_j$ such that $(X, a_{i(j)}g, p)$ converges to a pointed metric space (Y, d, q) as $j \to \infty$ in the sense of the pointed Gromov–Hausdorff topology; see Gromov [9; 10]. The limit (Y, d, q) is called the tangent cone at infinity of (X, g). In general, the pointed Gromov–Hausdorff limit might depend on the choice of $\{a_i\}_i$ or its subsequences.

The tangent cone at infinity is said to be unique if the isometry classes of the limits are independent of the choice of $\{a_i\}$ and its subsequences, and Colding and Minicozzi showed the next uniqueness theorem under the given assumptions.

Theorem 1.1 [6] *Let (X, g) be a Ricci-flat manifold with Euclidean volume growth, and suppose that one of the tangent cones at infinity has a smooth cross section. Then the tangent cone at infinity of (X, g) is unique.*

Among the assumptions in Theorem 1.1, the Ricci-flat condition is essential since there are several examples of complete Riemannian manifolds with nonnegative Ricci

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2683

curvature and Euclidean volume growth where one of the tangent cones at infinity has smooth cross section, but the tangent cone at infinity is not unique; see Perelman [12] and Colding and Naber [7].

Here, let $\mathcal{T}(X,g)$ be the set of all of the isometry classes of the tangent cones at infinity of (X,g). In this paper, an isometry between pointed metric spaces means a bijective map preserving the metrics and the base points. It is known that $\mathcal{T}(X,g)$ is closed with respect to the pointed Gromov–Hausdorff topology, and has the natural continuous $\mathbb{R}^+$–action defined by the rescaling of metrics. The uniqueness of the tangent cone at infinity means that $\mathcal{T}(X,g)$ consists of only one point.

In this paper, we show that the assumption for the volume growth in Theorem 1.1 is essential. More precisely, we obtain the next main result.

Theorem 1.2 *There is a complete Ricci-flat manifold (X,g) of dimension 4 such that $\mathcal{T}(X,g)$ is homeomorphic to S^1. Moreover, the $\mathbb{R}^+$–action on $\mathcal{T}(X,g)$ fixes $(\mathbb{R}^3,d_0^\infty,0)$, $(\mathbb{R}^3,h_0,0)$ and $(\mathbb{R}^3,h_1,0)$, where $h_0=\sum_{i=1}^3(d\zeta_i)^2$ is the Euclidean metric, d_0^∞ is the completion of the Riemannian metric*

$$\int_0^\infty \frac{dx}{|\zeta-(x^\alpha,0,0)|}\cdot h_0,$$

$h_1=(1/|\zeta|)h_0$, and $\mathbb{R}^+$ acts freely on

$$\mathcal{T}(X,g)\backslash\{(\mathbb{R}^3,d_0^\infty,0),(\mathbb{R}^3,h_0,0),(\mathbb{R}^3,h_1,0)\}.$$

Here, $\zeta=(\zeta_1,\zeta_2,\zeta_3)$ is the Cartesian coordinate on $\mathbb{R}^3$.

Here, we mention more about the metric spaces appearing in Theorem 1.2. For $0\le S<T\le\infty$, denote by d_S^T the metric on $\mathbb{R}^3$ induced by the Riemannian metric

$$\int_S^T \frac{dx}{|\zeta-(x^\alpha,0,0)|}\cdot h_0.$$

For (X,g) in Theorem 1.2, we show that $\mathcal{T}(X,g)$ contains $\{(\mathbb{R}^3,d_0^T,0):T\in\mathbb{R}^+\}$, $\{(\mathbb{R}^3,d_S^\infty,0):S\in\mathbb{R}^+\}$ and $\{(\mathbb{R}^3,h_0+\theta h_1,0):\theta\in\mathbb{R}^+\}$. Here, we can check easily that d_0^T and d_S^∞ are homothetic to d_0^1 and d_1^∞, respectively. We can show that

$$(\mathbb{R}^3,d_0^T,0)\xrightarrow[T\to\infty]{\mathrm{GH}}(\mathbb{R}^3,d_0^\infty,0),\qquad (\mathbb{R}^3,d_0^T,0)\xrightarrow[T\to 0]{\mathrm{GH}}(\mathbb{R}^3,h_1,0),$$

$$(\mathbb{R}^3,d_S^\infty,0)\xrightarrow[S\to\infty]{\mathrm{GH}}(\mathbb{R}^3,h_0,0),\qquad (\mathbb{R}^3,d_S^\infty,0)\xrightarrow[S\to 0]{\mathrm{GH}}(\mathbb{R}^3,d_0^\infty,0),$$

$$(\mathbb{R}^3,h_0+\theta h_1,0)\xrightarrow[\theta\to\infty]{\mathrm{GH}}(\mathbb{R}^3,h_1,0),\qquad (\mathbb{R}^3,h_0+\theta h_1,0)\xrightarrow[\theta\to 0]{\mathrm{GH}}(\mathbb{R}^3,h_0,0).$$

Both $(\mathbb{R}^3, h_0)$ and $(\mathbb{R}^3, h_1)$ can be regarded as the Riemannian cones with respect to the dilation $\zeta \mapsto \lambda\zeta$ on $\mathbb{R}^3$. Although the dilation also pulls back d_0^∞ to $\lambda^{(\alpha+1)/(2\alpha)} d_0^\infty$, $(\mathbb{R}^3, d_0^\infty)$ does not become the metric cone with respect to this dilation since $\boldsymbol{l} = \{(t,0,0) \in \mathbb{R}^3 : t \geq 0\}$ is not a ray. In fact, any open intervals contained in $\boldsymbol{l}$ have infinite length with respect to d_0^∞.

In general, tangent cones at infinity of complete Riemannian manifolds with nonnegative Ricci curvature and Euclidean volume growth are metric cones; see Cheeger and Colding [4]. In our case, it is shown in Section 9 that $(\mathbb{R}^3, d_0^\infty, 0)$ never becomes the metric cone of any metric space.

The Ricci-flat manifold (X, g) appearing in Theorem 1.2 is one of the hyper-Kähler manifolds of type A_∞, constructed by Anderson, Kronheimer and LeBrun in [1] applying Gibbons–Hawking ansatz, and by Goto in [8] as hyper-Kähler quotients. Combining Theorems 1.1 and 1.2, we can see that the volume growth of (X, g) should not be Euclidean. In fact, the author [11] has computed the volume growth of the hyper-Kähler manifolds of type A_∞ and showed that they are always greater than cubic growth and less than Euclidean growth. To construct (X, g), we "mix" the hyper-Kähler manifold of type A_∞ whose volume growth is r^a for some $3 < a < 4$, and $\mathbb{R}^4$ equipped with the standard hyper-Kähler structure. Unfortunately, the author could not compute the volume growth of (X, g) in Theorem 1.2 explicitly.

In this paper, we can show that many metric spaces may arise as the Gromov–Hausdorff limit of hyper-Kähler manifolds of type A_∞. Let

$$I \in \mathcal{B}_+(\mathbb{R}^+) := \{J \subset \mathbb{R}^+ : J \text{ is a Borel set of nonzero Lebesgue measure}\},$$

and denote by d_I the metric on $\mathbb{R}^3$ induced by the Riemannian metric

$$\int_I \frac{dx}{|\zeta - (x^\alpha, 0, 0)|} \cdot h_0.$$

Then we have the following result.

Theorem 1.3 *There is a complete Ricci-flat manifold (X, g) of dimension 4 such that $\mathcal{T}(X, g)$ contains*

$$\{(\mathbb{R}^3, d_I, 0) : I \in \mathcal{B}_+(\mathbb{R}^+)\}/\text{isometry}.$$

Since d_S^∞ and d_0^T are contained in $\mathcal{T}(X, g)$ in the above theorem, their limits h_0 and $(1/|\zeta|)h_0$ are also contained in $\mathcal{T}(X, g)$. The author does not know whether any other metric spaces are contained in $\mathcal{T}(X, g)$.

Theorems 1.2 and 1.3 are shown along the following process. The aforementioned hyper-Kähler manifolds are constructed from infinitely countable subsets Λ in $\mathbb{R}^3$

such that $\sum_{\lambda\in\Lambda} 1/(1+|\lambda|) < \infty$. We denote such a manifold by (X, g_Λ) and fix the base point $p \in X$. From the construction, (X, g_Λ) has a natural S^1–action preserving g_Λ and the hyper-Kähler structure; then we obtain a hyper-Kähler moment map $\mu_\Lambda\colon X \to \mathbb{R}^3$ such that $\mu_\Lambda(p) = 0$, which is a surjective map whose generic fibers are S^1. There is a unique distance function d_Λ on $\mathbb{R}^3$ such that μ_Λ is a submetry. Here, submetries are the generalizations of Riemannian submersions to the category of metric spaces. For $a > 0$, we can see $ag_\Lambda = g_{a\Lambda}$; hence by taking $a_i > 0$ such that $\lim_{i\to\infty} a_i = 0$, we obtain a sequence of submetries $\mu_{a_i\Lambda}\colon X \to \mathbb{R}^3$. Now, assume that $\{(\mathbb{R}^3, d_{a_i\Lambda}, 0)\}_i$ converges to a metric space $(\mathbb{R}^3, d_\infty, 0)$ for some d_∞ in the pointed Gromov–Hausdorff topology, and the diameters of fibers of $\mu_{a_i\Lambda}$ converge to 0 in some sense. Then we can show that $(\mathbb{R}^3, d_\infty, 0)$ is the Gromov–Hausdorff limit of $\{(X, g_{a_i\Lambda}, p)\}_i$. We raise a concrete example of Λ and sequences $\{a_i\}_i$, then obtain several limit spaces. Among them, it is shown in Section 9 that $(\mathbb{R}^3, d_0^\infty)$ is not a polar space in the sense of Cheeger and Colding [5].

This paper is organized as follows. We review the construction of hyper-Kähler manifolds of type A_∞ and the hyper-Kähler moment map μ_Λ in Section 2. Then we review the notion of submetry in Section 3, and the notion of Gromov–Hausdorff topology in Section 4. In Section 5, we construct a submetry μ_a from $(X, g_{a\Lambda})$ to $(\mathbb{R}^3, d_a)$ by using μ_Λ and dilation, where d_a is the metric induced by the Riemannian metric $\Phi_a(\zeta)h_0$. Here, Φ_a is a positive valued harmonic function determined by Λ and some constants. Then we see that the convergence of $\{(X, g_{a_i\Lambda})\}_i$ can be reduced to the convergence of $\{(\mathbb{R}^3, d_{a_i})\}_i$. In Sections 6 and 7, we raise concrete examples of Λ and fix $a > 0$, and then we estimate the difference of Φ_a and another positive valued harmonic function Φ_∞, which induces the metric d_∞ on $\mathbb{R}^3$. In Section 8, we observe some examples by applying the results in Sections 6 and 7, and then we show Theorems 1.2 and 1.3. In Section 9, we prove that $(\mathbb{R}^3, d_0^\infty)$ is not a polar space.

Acknowledgments The author would like to thank Professor Shouhei Honda who invited the author to this attractive topic, and also for advice on this paper. The author also would like to thank the referee for careful reading and several useful comments. Thanks to these, the author could make the main results much stronger. The author was supported by Grant-in-Aid for Young Scientists (B) Grant Number 16K17598. The author was partially supported by JSPS Core-to-Core Program, "Foundation of a Global Research Cooperative Center in Mathematics focused on Number Theory and Geometry".

2 Hyper-Kähler manifolds of type A_∞

Here we review briefly the construction of hyper-Kähler manifolds of type A_∞, along [1].

Let $\Lambda \subset \mathbb{R}^3$ be a countably infinite subset satisfying the convergence condition

$$\sum_{\lambda\in\Lambda} \frac{1}{1+|\lambda|} < \infty,$$

and take a positive valued harmonic function Φ_Λ over $\mathbb{R}^3 \backslash \Lambda$ defined by

$$\Phi_\Lambda(\zeta) := \sum_{\lambda\in\Lambda} \frac{1}{|\zeta-\lambda|}.$$

Then $*d\Phi_\Lambda \in \Omega^2(\mathbb{R}^3\backslash\Lambda)$ is a closed 2–form, where $*$ is the Hodge star operator of the Euclidean metric. We have an integrable cohomology class $[1/(4\pi) * d\Phi_\Lambda] \in H^2(\mathbb{R}^3\backslash\Lambda, \mathbb{Z})$, which is equal to the 1st Chern class of a principal S^1–bundle $\mu = \mu_\Lambda\colon X^* \to \mathbb{R}^3\backslash\Lambda$. For every $\lambda \in \Lambda$, we can take a sufficiently small open ball $B \subset \mathbb{R}^3$ centered at λ which does not contain any other elements in Λ. Then $\mu\colon \mu^{-1}(B\backslash\{\lambda\}) \to B\backslash\{\lambda\}$ is isomorphic to the Hopf fibration $\mu_0\colon \mathbb{R}^4\backslash\{0\} \to \mathbb{R}^3\backslash\{0\}$ as principal S^1–bundles; hence there exists a C^∞ 4–manifold X and an open embedding $X^* \subset X$, and μ can be extended to an S^1–fibration

$$\mu = (\mu_1, \mu_2, \mu_3)\colon X \to \mathbb{R}^3.$$

Moreover, we may write $X\backslash X^* = \{p_\lambda : \lambda \in \Lambda\}$ and $\mu(p_\lambda) = \lambda$. Next we take an S^1–connection $\Gamma \in \Omega^1(X^*)$ on $X^* \to \mathbb{R}^3\backslash\Lambda$, whose curvature form is given by $*d\Phi_\Lambda$. Then Γ is uniquely determined up to an exact 1–form on $\mathbb{R}^3\backslash\Lambda$. Now, we obtain a Riemannian metric

$$g_\Lambda := (\mu^*\Phi_\Lambda)^{-1}\Gamma^2 + \mu^*\Phi_\Lambda \sum_{i=1}^{3} (d\mu_i)^2$$

on X^*, which can be extended to a smooth Riemannian metric g_Λ over X by taking Γ appropriately.

Theorem 2.1 [1] *Let (X, g_Λ) be as above. Then it is a complete hyper-Kähler (hence Ricci-flat) metric of dimension 4.*

Since S^1 acts on (X, g_Λ) isometrically, it is easy to check that

$$\mu\colon (X^*, g_\Lambda) \to (\mathbb{R}^3\backslash\Lambda, \Phi_\Lambda \cdot h_0)$$

is a Riemannian submersion, where h_0 is the Euclidean metric on $\mathbb{R}^3$.

Next we consider the rescaling of (X, g_Λ). For $a > 0$, put $a\Lambda := \{a\lambda \in \mathbb{R}^3 : \lambda \in \Lambda\}$. Then it is easy to see

$$\Phi_{a\Lambda}(\zeta) = \sum_{\lambda\in\Lambda} \frac{1}{|\zeta - a\lambda|} = a^{-1} \sum_{\lambda\in\Lambda} \frac{1}{|a^{-1}\zeta - \lambda|} = a^{-1}\Phi_\Lambda(a^{-1}\zeta),$$

and $\mu_{a\Lambda} = a\mu_\Lambda$; hence $\mu_{a\Lambda}^*\Phi_{a\Lambda} = a^{-1}\mu_\Lambda^*\Phi_\Lambda$ holds. Thus we have

$$\begin{aligned} g_{a\Lambda} &= (\mu_{a\Lambda}^*\Phi_{a\Lambda})^{-1}\Gamma^2 + \mu_{a\Lambda}^*\Phi_{a\Lambda} \sum_{i=1}^3 (d\mu_{a\Lambda,i})^2 \\ &= a(\mu_\Lambda^*\Phi_\Lambda)^{-1}\Gamma^2 + a\mu_\Lambda^*\Phi_\Lambda \sum_{i=1}^3 (d\mu_{\Lambda,i})^2 = ag_\Lambda. \end{aligned}$$

3 Submetry

Throughout this paper, the distance between x and y in a metric space (X, d) is denoted by $d(x, y)$. If it is clear which metric is used, we often write $|xy| = d(x, y)$

The map $\mu\colon X \to \mathbb{R}^3$ appearing in the previous section is not a Riemannian submersion since $d\mu$ degenerates on $X \backslash X^*$ and $\Phi_\Lambda \cdot h_0$ is not defined on the whole of $\mathbb{R}^3$. However, we can regard μ as a submetry, which is a notion introduced in [3].

Definition 3.1 [3] Let X, Y be metric spaces and $\mu\colon X \to Y$ a map which is not necessarily continuous. Then μ is said to be a *submetry* if $\mu(D(p, r)) = D(\mu(p), r)$ for every $p \in X$ and $r > 0$, where $D(p, r)$ is the closed ball of radius r centered at p.

Any proper Riemannian submersions between smooth Riemannian manifolds are known to be submetries. Conversely, a submetry between smooth complete Riemannian manifolds becomes a $C^{1,1}$ Riemannian submersion [2].

Now we go back to the setting in Section 2. Denote by d_Λ the metric on $\mathbb{R}^3$ defined as the completion of the Riemannian distance induced from $\Phi_\Lambda \cdot h_0$. Since $\mu\colon (X^*, g_\Lambda) \to (\mathbb{R}^3 \backslash \Lambda, \Phi_\Lambda \cdot h_0)$ is a Riemannian submersion, we have the following proposition.

Proposition 3.2 *Let (X, g_Λ) be a hyper-Kähler manifold of type A_∞. The map $\mu\colon (X, d_{g_\Lambda}) \to (\mathbb{R}^3, d_\Lambda)$ is a submetry, where d_{g_Λ} is the Riemannian distance induced from g_Λ. Moreover, for any $p_0 \in \mu^{-1}(q_0)$, we have*

$$d_\Lambda(q_0, q_1) = \inf_{p_1 \in \mu^{-1}(q_1)} d_{g_\Lambda}(p_0, p_1).$$

4 The Gromov–Hausdorff convergence

In this section, we discuss the pointed Gromov–Hausdorff convergence of a sequence of pointed metric spaces equipped with submetries. First we review the definition of pointed Gromov–Hausdorff convergence of pointed metric spaces. Denote by $B(p,r)=B_X(p,r)$ the open ball of radius r centered at p in a metric space X.

Definition 4.1 Let (X,p) and (X',p') be pointed metric spaces, and let r and ε be positive real numbers. Then $f\colon B(p,r)\to X'$ is said to be an (r,ε)*–isometry* from (X,p) to (X',p') if

1. (1) $f(p)=p'$,
2. (2) $\big||xy|-|f(x)f(y)|\big|<\varepsilon$ holds for any $x,y\in B(p,r)$, and
3. (3) $B(f(B(p,r)),\varepsilon)$ contains $B(p',r-\varepsilon)$.

Definition 4.2 Let $\{(X_i,p_i)\}_i$ be a sequence of pointed metric spaces. Then we say $\{(X_i,p_i)\}_i$ *converges to a metric space* (X,p) *in the pointed Gromov–Hausdorff topology*, or $\{(X_i,p_i)\}_i\xrightarrow{\mathrm{GH}}(X,p)$, if for any $r,\varepsilon>0$ there exists a positive integer $N_{(r,\varepsilon)}$ such that an (r,ε)–isometry from (X_i,p_i) to (X,p) exists for every $l\geq N_{(r,\varepsilon)}$.

For metric spaces X, Y, a map $\mu\colon X\to Y$ and $q\in Y$, define $\delta_{q,\mu}(r)\in\mathbb{R}_{\geq 0}\cup\{\infty\}$ by

$$\delta_{q,\mu}(r):=\sup_{y\in B(q,r)}\operatorname{diam}(\mu^{-1}(y))=\sup_{\substack{y\in B(q,r)\\ x,x'\in\mu^{-1}(y)}}|xx'|.$$

Proposition 4.3 *Let (X,p) and (Y,q) be pointed metric spaces equipped with submetries $\mu\colon X\to Y$ satisfying $\mu(p)=q$, and let (Y_∞,q_∞) be another pointed metric space. Assume that $\delta_{q,\mu}(r)<\infty$ and we have an (r,δ)–isometry from (Y,q) to (Y_∞,q_∞). Then there exists an $(r,\delta+\delta_{q,\mu})$–isometry from (X,p) to (Y_∞,q_∞).*

Proof There is an (r,δ)–isometry f from (Y,q) to (Y_∞,q_∞). It is easy to check that the composition $\hat{f}:=f\circ\mu$ is an $(r,\delta+\delta_{q,\mu})$–isometry from (X,p) to (Y_∞,q_∞). □

5 Tangent cones at infinity

Let (X,d) be a metric space and $\{a_i\}_i$ a decreasing sequence of positive numbers converging to 0. If (Y,q) is the pointed Gromov–Hausdorff limit of $\{(X,a_id,p)\}_i$, then it is called an tangent cone at infinity of X. It is clear that the limit does not depend on $p\in X$, but may depend on the choice of the sequence $\{a_i\}_i$.

In this paper, we are considering the tangent cones at infinity of (X,d_{g_Λ}). In Section 2, we have seen that $\sqrt{a}d_{g_\Lambda}=d_{g_{a\Lambda}}$ for $a>0$; hence $\mu_{a\Lambda}\colon(X,\sqrt{a}d_{g_\Lambda})\to(\mathbb{R}^3,d_{a\Lambda})$

is a submetry. By taking $N \in \mathbb{R}^+$ and the dilation $I_N\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by $I_N(\zeta) := (1/N)\zeta$, we have another submetry

$$\mu_a := I_N^{-1} \circ \mu_{a\Lambda}\colon (X, \sqrt{a} d_{g_\Lambda}) \to (\mathbb{R}^3, d_a := I_N{}^* d_{a\Lambda}).$$

Here, $I_N{}^* d_{a\Lambda}$ is the completion of the Riemannian distance of

$$I_N{}^*(\Phi_{a\Lambda} \cdot h_0) = I_N{}^* \Phi_{a\Lambda} \cdot \frac{1}{N^2} h_0 = N \Phi_{Na\Lambda} \cdot \frac{1}{N^2} h_0 = \frac{1}{N} \Phi_{Na\Lambda} \cdot h_0.$$

Thus we obtain d_a, which is the completion of the Riemannian metric $\Phi_a \cdot h_0$, where

$$\Phi_a := \frac{1}{N} \Phi_{Na\Lambda}.$$

In other words, d_a is given by

(1) $$d_a(x, y) = \inf_{\gamma \in \mathrm{Path}(x,y)} l_a(\gamma),$$

where $\mathrm{Path}(x, y)$ is the set of smooth paths in $\mathbb{R}^3$ joining $x, y \in \mathbb{R}^3$, and

(2) $$l_a(\gamma) = \int_{t_0}^{t_1} \sqrt{\Phi_a(\gamma(t))}\, |\gamma'(t)|\, dt.$$

By the definition of g_Λ, one can see that the diameter of the fiber $\mu_\Lambda^{-1}(\zeta)$ is given by $\pi/\sqrt{\Phi_\Lambda(\zeta)}$. Accordingly, the diameter of $\mu_a^{-1}(\zeta)$ is given by $\pi/(N\sqrt{\Phi_a(\zeta)})$.

For a metric d_∞ on $\mathbb{R}^3$ and constants $r, \delta, \delta' > 0$, we introduce the next assumptions.

(A1) The identity map $\mathrm{id}_{\mathbb{R}^3}\colon (\mathbb{R}^3, d_a, 0) \to (\mathbb{R}^3, d_\infty, 0)$ is an (r, δ)–isometry.

(A2) $$\sup_{\zeta \in B_{d_a}(0,r)} \frac{\pi}{N\sqrt{\Phi_a(\zeta)}} < \delta'.$$

Then we obtain the next proposition by Proposition 4.3.

Proposition 5.1 *Let (X, g_Λ) and μ_a be as above, $p \in X$ satisfy $\mu_\Lambda(p) = 0$ and d_∞ be a metric on $\mathbb{R}^3$. If* **(A1)** *and* **(A2)** *are satisfied for given constants $r, \delta, \delta' > 0$, then μ_a is an $(r, \delta+\delta')$–isometry from (X, ag_Λ, p) to $(\mathbb{R}^3, d_\infty, 0)$.*

6 Construction

Fix $\alpha > 1$, and let

$$\Lambda^\alpha := \{(k^\alpha, 0, 0) : k \in \mathbb{Z}_{\geq 0}\}.$$

Take an increasing sequence of integers $0 < K_0 < K_1 < K_2 < \cdots$.

In this paper, many constants will appear, and they may depend on α or $\{K_n\}$. However, we do not mind the dependence on these parameters. Put

$$\Lambda_{2n} := \{(k^\alpha, 0, 0) \in \Lambda^\alpha : K_{2n} \le k < K_{2n+1}\}, \quad \Lambda := \bigcup_{n=0}^{\infty} \Lambda_{2n}.$$

Since $\Lambda \subset \Lambda^\alpha$, we can see that $\sum_{\lambda\in\Lambda} 1/(1+|\lambda|) < \infty$; accordingly, we obtain a hyper-Kähler manifold (X, g_Λ).

From now on, we fix $a > 0$, $n \in \mathbb{N}$ and $P > 0$, then put $N := a^{-1/(1+\alpha)} P^{1/(1+\alpha)}$ and

$$\Phi_a(\zeta) := \frac{1}{N}\Phi_{Na\Lambda}(\zeta) = \sum_{\lambda\in\Lambda} \frac{1}{N|\zeta - PN^{-\alpha}\lambda|}.$$

Let $\boldsymbol{l} := \{(t, 0, 0) \in \mathbb{R}^3 : t \ge 0\}$, and put

$$K(R, D) := \Big\{\zeta \in \mathbb{R}^3 : |\zeta| \le R,\ \inf_{y\in\boldsymbol{l}} |\zeta - y| \ge D\Big\}.$$

Here, $\inf_{y\in\boldsymbol{l}} |\zeta - y|$ is given by

$$\inf_{y\in\boldsymbol{l}} |\zeta - y| = \begin{cases} \sqrt{|\zeta_{\mathbb{C}}|^2} & \text{if } \zeta_{\mathbb{R}} \ge 0, \\ |\zeta| & \text{if } \zeta_1 < 0 \end{cases}$$

for $\zeta = (\zeta_{\mathbb{R}}, \zeta_{\mathbb{C}}) \in \mathbb{R}^3 = \mathbb{R} \oplus \mathbb{C}$. For $0 \le S < T \le \infty$, define a positive valued function $\Phi_{S,P}^T : \mathbb{R}^3 \backslash \boldsymbol{l} \to \mathbb{R}$ by

$$\Phi_{S,P}^T(\zeta) := \int_S^T \frac{dx}{|\zeta - P(x^\alpha, 0, 0)|}.$$

Throughout this section, we put

$$S_n := \frac{K_{2n}}{N} = a^{\frac{1}{1+\alpha}} P^{\frac{-1}{1+\alpha}} K_{2n}, \quad T_n := \frac{K_{2n+1}}{N} = a^{\frac{1}{1+\alpha}} P^{\frac{-1}{1+\alpha}} K_{2n+1}.$$

Proposition 6.1 *For any $\zeta \in K(R, D)$, we have*

$$\left|\Phi_a(\zeta) - \sum_{n=0}^{\infty} \Phi_{S_n,P}^{T_n}(\zeta)\right| \le \frac{2}{ND} = \frac{2}{D}\left(\frac{a}{P}\right)^{\frac{1}{1+\alpha}}.$$

Proof Since

$$\Lambda_{2n} = \{(k^\alpha, 0, 0) : K_{2n} \le k < K_{2n+1}\},$$

we have

$$\sum_{\lambda\in\Lambda_{2n}} \frac{1}{N|\zeta - PN^{-\alpha}\lambda|} = \sum_{k=K_{2n}}^{K_{2n+1}-1} \frac{1}{N|\zeta - PN^{-\alpha}(k^\alpha, 0, 0)|}.$$

Then we obtain

$$(3)\qquad \left|\sum_{n=0}^{\infty}\left(\sum_{\lambda\in\Lambda_{2n}}\frac{1}{N|\zeta-PN^{-\alpha}\lambda|}-\int_{K_{2n}/N}^{K_{2n+1}/N}\frac{dx}{|\zeta-P(x^{\alpha},0,0)|}\right)\right|\leq\frac{2}{ND}.$$

The above inequality holds since the function $x\mapsto 1/|\zeta-P(x^{\alpha},0,0)|$ has at most one critical point, and for all $\zeta\in K(R,D)$,

$$\sup_{x\in\mathbb{R}}\frac{1}{|\zeta-P(x^{\alpha},0,0)|}-\inf_{x\in\mathbb{R}}\frac{1}{|\zeta-P(x^{\alpha},0,0)|}\leq\frac{1}{D}.\qquad\square$$

Next we obtain the lower estimate of Φ_a as follows.

Proposition 6.2 *We have*

$$(4)\qquad \Phi_{S_n,P}^{T_n}(\zeta)\geq\left(\int_{S_n}^{T_n}\frac{dx}{1+Px^{\alpha}}\right)\min\left\{\frac{1}{|\zeta|},1\right\},$$

$$(5)\qquad \Phi_a(\zeta)\geq\left(\sum_{n=0}^{\infty}\int_{S_n}^{T_n}\frac{dx}{1+Px^{\alpha}}-2(aP^{-1})^{\frac{1}{1+\alpha}}\right)\min\left\{\frac{1}{|\zeta|},1\right\},$$

$$(6)\qquad \sum_{n=0}^{\infty}\Phi_{S_n,P}^{T_n}(\zeta)\leq P^{-\frac{1}{\alpha}}\frac{\alpha 2^{\frac{1}{\alpha}}}{\alpha-1}\frac{|\zeta|^{\frac{1}{\alpha}}}{|\zeta_{\mathbb{C}}|},$$

$$(7)\qquad \sum_{n=n_0}^{\infty}\Phi_{S_n,P}^{T_n}(\zeta)\leq\frac{2S_{n_0}^{-\alpha+1}}{P(\alpha-1)}\quad \textit{if } S_{n_0}\geq\left(\frac{2|\zeta|}{P}\right)^{\frac{1}{\alpha}},$$

$$(8)\qquad \sum_{n=0}^{n_0}\Phi_{S_n,P}^{T_n}(\zeta)\leq\frac{T_{n_0}}{D}\quad \textit{if } \zeta\in K(R,D).$$

Proof First of all, one can see

$$\Phi_{S_n,P}^{T_n}(\zeta)\geq\int_{S_n}^{T_n}\frac{dx}{|\zeta|+Px^{\alpha}}\geq\frac{1}{|\zeta|}\int_{S_n}^{T_n}\frac{dx}{1+Px^{\alpha}}$$

if $|\zeta|\geq 1$, and

$$\Phi_{S_n,P}^{T_n}(\zeta)\geq\int_{S_n}^{T_n}\frac{dx}{|\zeta|+Px^{\alpha}}\geq\int_{S_n}^{T_n}\frac{dx}{1+Px^{\alpha}}$$

if $|\zeta|\leq 1$. Next we have

$$\Phi_a(\zeta)\geq\sum_{n=0}^{\infty}\sum_{k=K_{2n}}^{K_{2n+1}-1}\frac{1}{N(|\zeta|+PN^{-\alpha}k^{\alpha})},$$

and a similar argument to the proof of Proposition 6.1 gives

$$\left|\sum_{n=0}^{\infty}\left(\sum_{k=K_{2n}}^{K_{2n+1}-1}\frac{1}{N(|\zeta|+PN^{-\alpha}k^{\alpha})}-\int_{S_n}^{T_n}\frac{dx}{|\zeta|+Px^{\alpha}}\right)\right|\le\frac{2}{N|\zeta|}.$$

Combining these inequalities, one has the second assertion if $|\zeta|\ge 1$. If $|\zeta|\le 1$, then

$$\Phi_a(\zeta)\ge\sum_{n=0}^{\infty}\sum_{k=K_{2n}}^{K_{2n+1}-1}\frac{1}{N(1+PN^{-\alpha}k^{\alpha})},$$

and by a similar argument, we obtain the assertion.

Next we consider (6). If $t\ge(2|\zeta|/P)^{1/\alpha}$, then

$$\int_t^{\infty}\frac{dx}{|\zeta-P(x^{\alpha},0,0)|}\le\int_t^{\infty}\frac{2dx}{Px^{\alpha}}=\frac{2}{P(\alpha-1)}t^{-\alpha+1}\tag{9}$$

holds. Hence one can see

$$\begin{aligned}\sum_{n=0}^{\infty}\Phi_{S_n,P}^{T_n}(\zeta_{\mathbb{R}},\zeta_{\mathbb{C}})&\le\int_0^{\infty}\frac{dx}{|\zeta-P(x^{\alpha},0,0)|}\\&=\int_0^{\left(\frac{2|\zeta|}{P}\right)^{\frac{1}{\alpha}}}\frac{dx}{|\zeta-P(x^{\alpha},0,0)|}+\int_{\left(\frac{2|\zeta|}{P}\right)^{\frac{1}{\alpha}}}^{\infty}\frac{dx}{|\zeta-P(x^{\alpha},0,0)|}\\&\le\frac{(2|\zeta|/P)^{\frac{1}{\alpha}}}{|\zeta_{\mathbb{C}}|}+\frac{2}{P(\alpha-1)}\left(\frac{2|\zeta|}{P}\right)^{\frac{1}{\alpha}(-\alpha+1)}\\&=P^{-\frac{1}{\alpha}}\left(\frac{(2|\zeta|)^{\frac{1}{\alpha}}}{|\zeta_{\mathbb{C}}|}+\frac{(2|\zeta|)^{\frac{1}{\alpha}}}{\alpha-1}\frac{1}{|\zeta_{\mathbb{C}}|}\right).\end{aligned}$$

The statement (7) follows from (9), and (8) is obvious. □

Put

$$A_{S,P}^{T}:=\int_S^T\frac{dx}{1+Px^{\alpha}}.$$

By Proposition 6.2, we have the following.

Proposition 6.3 *Let Φ_a be as above. Then for every $R\ge 1$,*

$$\sup_{|\zeta|\le R}\frac{1}{N\sqrt{\Phi_a(\zeta)}}\le\left(\frac{a}{P}\right)^{\frac{1}{1+\alpha}}\left(\sum_{n=0}^{\infty}A_{S_n,P}^{T_n}-2\left(\frac{a}{P}\right)^{\frac{1}{1+\alpha}}\right)^{-\frac{1}{2}}\sqrt{R}.$$

7 Distance

In the previous section, we estimated $\left|\Phi_a - \sum_n \Phi^{T_n}_{S_n,P}\right|$ on $K(R,D)$.

In this section, we introduce more general positive functions Φ and Φ_∞, and induced metrics d and d_∞ on $\mathbb{R}^3$, respectively. What we hope to show in this section is that if we fix a very large $R \geq 1$ and assume that $\sup_{K(R,D)}|\Phi - \Phi_\infty| \leq \varepsilon/D$ holds for a very small ε and every $D \leq 1$, then the identity map of $\mathbb{R}^3$ becomes an (r,δ)–isometry from $(\mathbb{R}^3, d, 0)$ to $(\mathbb{R}^3, d_\infty, 0)$ for a large r and a small δ. Here, we explain the difficulty in showing it.

We hope to show that $\sup_{K(R,D)}|d - d_\infty|$ is small for every $R \geq 1$ and $0 < D \leq 1$. By the estimate of $\sup_{K(R,D)}|\Phi - \Phi_\infty|$, it is easy to see that $\sup_{K(R,D)}|d_{R,D} - d_{\infty,R,D}|$ is small, where $d_{R,D}$ (resp. $d_{\infty,R,D}$) is the Riemannian distance of the Riemannian metric $\Phi h_0|_{K(R,D)}$ (resp. $\Phi_\infty h_0|_{K(R,D)}$). However, $d_{R,D}$ may not equal d in general since the geodesic of Φh_0 joining two points in $K(R,D)$ might leave $K(R,D)$. To see that $\sup_{K(R,D)}|d_{R,D} - d|$ is sufficiently small, we have to observe that a path joining two points in $K(R,D)$ which leaves $K(R,D)$ can be replaced by a shorter path included in $K(R,D)$.

In this section, we consider positive valued functions $\Phi, \Phi_\infty \in C^\infty(\mathbb{R}^3 \backslash \boldsymbol{l})$ satisfying the following conditions for given constants $R \geq 1$, $m, \varepsilon, C_0, C_1 > 0$ and $\kappa \geq 0$:

(A3)
$$|\Phi(\zeta) - \Phi_\infty(\zeta)| \leq \frac{\varepsilon}{D^m} \quad \text{and} \quad |\Phi(\zeta) - \Phi_\infty(\zeta)| \leq \frac{C_1}{D}$$

hold for any $D \leq 1$ and $\zeta \in K(R,D)$.

(A4) Along the decomposition $\mathbb{R}^3 = \mathbb{R} \oplus \mathbb{C}$, put $\zeta = (\zeta_\mathbb{R}, \zeta_\mathbb{C}) \in \mathbb{R} \oplus \mathbb{C}$. Then

$$\begin{aligned} \Phi(\zeta_\mathbb{R}, e^{i\theta}\zeta_\mathbb{C}) &= \Phi(\zeta_\mathbb{R}, \zeta_\mathbb{C}), & \Phi(\zeta_\mathbb{R}, \zeta_\mathbb{C}) &\leq \Phi(\zeta_\mathbb{R}, \zeta'_\mathbb{C}), \\ \Phi_\infty(\zeta_\mathbb{R}, e^{i\theta}\zeta_\mathbb{C}) &= \Phi_\infty(\zeta_\mathbb{R}, \zeta_\mathbb{C}), & \Phi_\infty(\zeta_\mathbb{R}, \zeta_\mathbb{C}) &\leq \Phi_\infty(\zeta_\mathbb{R}, \zeta'_\mathbb{C}) \end{aligned}$$

hold for any $e^{i\theta} \in S^1$, if $|\zeta_\mathbb{C}| \geq |\zeta'_\mathbb{C}|$.

(A5)
$$\min\{\Phi(\zeta), \Phi_\infty(\zeta)\} \geq \begin{cases} C_0/|\zeta| & \text{if } |\zeta| \geq 1, \\ C_0 & \text{if } |\zeta| \leq 1. \end{cases}$$

(A6) For any $u \geq 1$ and $\zeta \in \mathbb{R}^3 \backslash \boldsymbol{l}$ with $|\zeta| \leq u$,

$$\Phi_\infty(\zeta) \leq \frac{C_1 u^\kappa}{|\zeta_\mathbb{C}|}.$$

Remark 7.1 Let $\Phi = \Phi_a$ and $\Phi_\infty = \Phi^T_{S,P}$ be as in Section 6. Then they satisfy **(A4)**, and also satisfy **(A3)**, **(A5)** and **(A6)** for appropriate constants ε, C_0 and C_1 given by Propositions 6.1 and 6.2.

From now on, let Φ, Φ_∞ satisfy **(A3)–(A6)** for constants $R, \varepsilon, C_0, C_1, \kappa$. Denote by d, d_∞ the metrics on $\mathbb{R}^3$ induced by $\Phi \cdot h$, $\Phi_\infty \cdot h$, and by l, l_∞ the lengths of paths with respect to d, d_∞, respectively.

7.1 Estimates (1)

Let $\boldsymbol{B}(u) := \{\zeta \in \mathbb{R}^3 : |\zeta| < u\}$ and $\mathrm{Path}(u, x, y)$ be the set of smooth paths in $\overline{\boldsymbol{B}(u)}$ joining $x, y \in \overline{\boldsymbol{B}(u)}$; then put

$$d_u(x, y) = \inf_{\gamma \in \mathrm{Path}(u,x,y)} l(\gamma), \qquad d_{\infty,u}(x, y) = \inf_{\gamma \in \mathrm{Path}(u,x,y)} l_\infty(\gamma)$$

By the definition, $d(x, y) \le d_u(x, y)$ and $d_\infty(x, y) \le d_{\infty,u}(x, y)$ always hold. However, the opposite inequality may not hold since the minimizing geodesic γ joining $x, y \in \overline{\boldsymbol{B}(u)}$ may leave $\overline{\boldsymbol{B}(u)}$. The goal of this subsection is to show $d_{\rho(u)}(x, y) \le d(x, y)$ and $d_{\infty,\rho(u)}(x, y) \le d_\infty(x, y)$ for a sufficiently large $\rho(u)$.

Proposition 7.2 *Suppose Φ and Φ_∞ satisfy* **(A3)–(A6)***. Let D_u and $D_{u,u'}$ be the diameters of $\overline{\boldsymbol{B}(t)}$ with respect to d and $d_{u'}$, respectively, where $0 < u \le u'$. Define $D_{\infty,u}$ and $D_{\infty,u,u'}$ in the same way. Then the inequality*

$$2\sqrt{C_0}(\sqrt{|\zeta|} - 1) \le \min\{d(0, \zeta), d_\infty(0, \zeta)\}$$

holds for all $\zeta \in \mathbb{R}^3$, and

$$d(0, \zeta) \le D_u \le D_{u,u} \le C_2 u^{\kappa'}, \quad d_\infty(0, \zeta) \le D_{\infty,u} \le D_{\infty,u,u} \le C_2 u^{\kappa'}$$

hold for all $\zeta \in \mathbb{R}^3$ and $u \ge 1$ with $|\zeta| \le u \le R$, where C_2 is the constant depending only on C_1 and $\kappa' = \frac{1}{2}(1 + \kappa)$.

Proof First we show the first inequality. Let $\gamma\colon [a, b] \to \mathbb{R}^3$ be a smooth path such that $\gamma(a) = 0$ and $\gamma(b) = \zeta$. We may suppose $|\zeta| \ge 1$ since it is obviously satisfied when $|\zeta| < 1$. Then there is $s \in [a, b]$ such that $|\gamma(s)| = 1$ and $|\gamma(t)| \ge 1$ for any $t \in [s, b]$. Then by the assumption **(A5)**, one can see

$$l(\gamma) = \int_a^b \sqrt{\Phi(\gamma(t))}\, |\gamma'(t)|\, dt \ge \int_s^b \sqrt{\Phi(\gamma)}\, |\gamma'|\, dt \ge \int_s^b \sqrt{\frac{C_0}{|\gamma|}}\, |\gamma'|\, dt.$$

Since we have $|\gamma'| \ge |\gamma|'$, we obtain, for all $\zeta \in \mathbb{R}^3$ with $|\zeta| \ge 1$,

$$l(\gamma) \ge \int_s^b \sqrt{\frac{C_0}{|\gamma|}}\, |\gamma|'\, dt \ge 2\sqrt{C_0} \int_s^b \frac{d}{dt}\sqrt{|\gamma|}\, dt \ge 2\sqrt{C_0}(\sqrt{|\zeta|} - 1).$$

By the definition, $d(0,\zeta) \le D_u \le D_{u,R_1} \le D_{u,R_0}$ always hold for any $u \le R_0 \le R_1$ and $\zeta \in \mathbb{R}^3$ with $|\zeta| \le u$. Next we estimate $D_{u,u}$ from the above. For every ζ, we will prepare the piecewise smooth paths γ_ζ in $\overline{\boldsymbol{B}(u)}$ joining 0 and ζ as described below. Then we will have an upper bound

$$D_{u,u} \le 2 \sup_{\zeta \in \overline{\boldsymbol{B}(u)}} l(\gamma_\zeta).$$

Here we define γ_ζ as follows. We have the isometric S^1–action on $\mathbb{R}^3$ with respect to d and d_∞ by **(A4)**. By supposing $\gamma_{e^{i\theta}\zeta} = e^{i\theta}\gamma_\zeta$, it suffices to consider γ_ζ in the case of $\zeta = r(\sin s, -\cos s, 0)$, where $r > 0$ and $-\pi < s \le \pi$. Let

$$\gamma_\zeta|_{[0,1]}(t) := (0, -rt, 0),$$
$$\gamma_\zeta|_{[1,2]}(t) := r\big(\sin(s(t-1)), -\cos(s(t-1)), 0\big).$$

Since $\zeta \in K(R, |\zeta_{\mathbb{C}}|)$ holds, **(A3)** gives $|\Phi(\zeta) - \Phi_\infty(\zeta)| \le C_1/|\zeta_{\mathbb{C}}|$, and **(A6)** gives $\Phi_\infty(\zeta) \le C_1 u^\kappa/|\zeta_{\mathbb{C}}|$. Then we can see

$$\begin{aligned}
l(\gamma_\zeta|_{[0,1]}) &= \int_0^1 \sqrt{\Phi(\gamma_\zeta)}\, |\gamma_\zeta'|\, dt \\
&\le \int_0^1 r\sqrt{|\Phi(\gamma_\zeta) - \Phi_\infty(\gamma_\zeta)|}\, dt + \int_0^1 r\sqrt{|\Phi_\infty(\gamma_\zeta)|}\, dt \\
&\le \int_0^1 r\sqrt{\frac{C_1}{rt}}\, dt + \int_0^1 r\sqrt{\frac{C_1 u^\kappa}{rt}}\, dt \\
&\le 2\sqrt{C_1 u} + 2\sqrt{C_1}\, u^{\frac{\kappa+1}{2}}.
\end{aligned}$$

Simultaneously, we also have

$$\begin{aligned}
l(\gamma_\zeta|_{[1,2]}) &\le \int_1^2 |\gamma_\zeta'| \sqrt{|\Phi(\gamma_\zeta) - \Phi_\infty(\gamma_\zeta)|}\, dt + \int_1^2 |\gamma_\zeta'|\sqrt{|\Phi_\infty(\gamma_\zeta)|}\, dt \\
&\le \int_0^1 r|s|\sqrt{\frac{C_1}{r|\cos st|}}\, dt + \int_0^1 r|s|\sqrt{\frac{C_1 u^\kappa}{r|\cos st|}}\, dt \\
&\le \sqrt{C_1 u} + \sqrt{C_1 u^{1+\kappa}} \int_0^{|s|} \sqrt{\frac{1}{\cos t}}\, dt.
\end{aligned}$$

Here, $\int_0^\pi \sqrt{1/\cos t}\, dt$ is finite. Since $u \ge 1$ and $\kappa \ge 0$, we may suppose that $\max\{\sqrt{u}, \sqrt{u^{1+\kappa}}\} = u^{1+\kappa}$. By combining these estimates and putting

$$C_2 = \left(2 + \int_0^\pi \sqrt{\frac{1}{\cos t}}\, dt\right)\sqrt{C_1},$$

we have the assertion. We also obtain the estimate of $D_{\infty,u,u}$ by the above argument. □

Proposition 7.3 *Suppose* Φ, Φ_∞ *satisfy* **(A3)–(A6)**, *and for* $t > 0$, *let*

$$\rho(t) := \max\{t-1, (1+C_3 t^{\kappa'})^2\},$$

where $C_3 = 3C_2/(2\sqrt{C_0})$ *and* C_2 *is the constant in* Proposition 7.2. *Then* $d_{\rho(u)}(x,y) = d(x,y)$ *and* $d_{\infty,\rho(u)}(x,y) = d_\infty(x,y)$ *for any* $x, y \in \boldsymbol{B}(u)$ *and* $1 \le u \le R$.

Proof By the definition, $d(x,y) \le d_{\rho(u)}(x,y)$ always holds. We assume $d(x,y) < d_{\rho(u)}(x,y)$ for some $x, y \in \boldsymbol{B}(u)$. Then there is a smooth $\gamma\colon [a,b] \to \mathbb{R}^3$ joining x and y such that $d(x,y) \le l(\gamma) < d_{\rho(u)}(x,y)$, which implies the existence of $c \in [a,b]$ satisfying $|\gamma(c)| = \rho(u)$. Then one can see

$$\begin{aligned} l(\gamma) \ge l(\gamma|_{[a,c]}) &\ge d(0,\gamma(c)) - d(0,\gamma(a)) \\ &\ge 2\sqrt{C_0}(\sqrt{\rho(u)}-1) - D_{u,u} \\ &\ge 2\sqrt{C_0}\Big(\sqrt{(1+C_3u^{\kappa'})^2}-1\Big) - C_2u^{\kappa'} \\ &\ge 2C_2u^{\kappa'} \end{aligned}$$

by Proposition 7.2. On the other hand, we have

$$d_{\rho(u)}(x,y) \le D_{u,\rho(u)} \le D_{u,u} \le C_2u^{\kappa'}$$

by Proposition 7.2. Therefore, we obtain

$$2C_2u^{\kappa'} \le l(\gamma) < d_{\rho(u)}(x,y) \le C_2u^{\kappa'},$$

a contradiction. We can show $d_\infty(x,y) = d_{\infty,\rho(u)}(x,y)$ in the same way. □

7.2 Estimates (2)

In this subsection, let $\gamma\colon [a,b] \to \boldsymbol{B}(u)$ be a smooth path joining $x, y \in \mathbb{R}^3 \setminus L(D)$, where

$$L(D) := \{\zeta \in \mathbb{R}^3 : |\zeta_{\mathbb{C}}| < D\}.$$

Now, we are going to show that if γ is a minimizing geodesic joining x and y, then it never approaches the axis $\{(t,0,0) \in \mathbb{R}^3 : t \in \mathbb{R}\}$. To show this, if the given γ invades $L(D)$, then we modify γ and construct the new path c_γ to not invade $L(D)$.

Lemma 7.4 *Suppose* Φ, Φ_∞ *satisfy* **(A4)**. *Let* $\gamma = (\gamma_{\mathbb{R}}, \gamma_{\mathbb{C}})\colon [a,b] \to \mathbb{R}^3 = \mathbb{R} \oplus \mathbb{C}$ *be a smooth path satisfying* $|\gamma_{\mathbb{C}}(a)| = |\gamma_{\mathbb{C}}(b)| = D$ *and* $|\gamma_{\mathbb{C}}(t)| \le D$ *for any* $t \in [a,b]$. *Define* $P_\gamma\colon [a,b] \to \mathbb{R}^3$ *by*

$$P_\gamma(t) := (\gamma_{\mathbb{R}}(t), \gamma_{\mathbb{C}}(a)).$$

Then $l(P_\gamma) \le l(\gamma)$ *and* $l_\infty(P_\gamma) \le l_\infty(\gamma)$ *hold.*

Proof Since $\Phi(\gamma(t)) \geq \Phi(P_\gamma(t))$ holds by the second inequality of **(A4)**, and

$$|\gamma'|^2 = |\gamma'_{\mathbb{R}}|^2 + |\gamma'_{\mathbb{C}}|^2 \geq |\gamma'_{\mathbb{R}}|^2 = |P'_\gamma|^2$$

holds, we can deduce

$$l(\gamma) = \int_a^b \sqrt{\Phi(\gamma(t))}\,|\gamma'(t)|\,dt \geq \int_a^b \sqrt{\Phi(P_\gamma(t))}\,|P'_\gamma(t)|\,dt \geq l(P_\gamma). \qquad \square$$

Let $\gamma\colon [a,b] \to \mathbb{R}^3$ be a smooth path joining $x, y \in \mathbb{R}^3 \backslash L(D)$, and assume that $|\gamma_{\mathbb{C}}(a')| = |\gamma_{\mathbb{C}}(b')| = D$ and that $\gamma((a',b'))$ is contained in $\overline{L(D)}$ for some $a \leq a' < b' \leq b$. Then define a new path $\Gamma(\gamma,[a',b'])\colon [a,b] \to \mathbb{R}^3$ by connecting

$$\gamma|_{[a,a']}, \quad P_{\gamma|_{[a',b']}}, \quad e^{i\theta}\gamma|_{[b',b]}.$$

Here, by choosing $e^{i\theta}$ appropriately, $\Gamma(\gamma,[a',b'])$ is continuous and piecewise smooth. By Lemma 7.4, the length of $\Gamma(\gamma,[a',b'])$ is not longer than that of γ since S^1–rotation preserves d and d_∞.

Put $J := \gamma^{-1}(L(D)) \cap (a,b)$. Since J is open in (a,b), it is decomposed into disjoint open intervals

$$J = \bigsqcup_{q \in \mathcal{Q}} (a_q, b_q)$$

for some $a_q, b_q \in [a,b]$ and countable set $\mathcal{Q}$. If $q \in \mathcal{Q}$, then $|\gamma_{\mathbb{C}}(a_q)| = |\gamma_{\mathbb{C}}(b_q)| = D$ holds. Then we have $\gamma_1 := \Gamma(\gamma,[a_q,b_q])$ for a fixed $q \in \mathcal{Q}$; moreover, we obtain $\gamma_2 := \Gamma(\gamma_1,[a_{q'},b_q])$ for another $q' \in \mathcal{Q}$, and repeating this process for all $q \in \mathcal{Q}$ we finally obtain the piecewise smooth path $c\colon [a,b] \to \mathbb{R}^3$ such that $c(a) = \gamma(a)$, $c(b) = e^{i\theta}\gamma(b)$ for some $e^{i\theta_0}$ and

$$l(c) \leq l(\gamma), \quad l_\infty(c) \leq l_\infty(\gamma).$$

Here, we have to modify c so that the terminal points of both paths coincide. Put $\bar{b} := \sup\{t \in [a,b] : |\gamma_{\mathbb{C}}(t)| = D\}$. Then define a path $\hat{\gamma}$ by connecting $c|_{[a,\bar{b}]}$ and $\gamma|_{[\bar{b},b]}$. Here, to connect $c(\bar{b})$ and $\gamma(\bar{b})$, we add the path $c_{\theta_0}\colon [0,\theta_0] \to \partial L(D)$ defined by $c_{\theta_0}(t) = e^{it}\gamma(\bar{b})$. Then by **(A6)**, we obtain

$$l(c_{\theta_0}) \leq \sqrt{C_1(1+u^\kappa)}\sqrt{D} \quad \text{and} \quad l_\infty(c_{\theta_0}) \leq \sqrt{C_1 u^\kappa}\sqrt{D}$$

if $|\gamma(\bar{b})| \leq u \leq R$. Hence we have the following proposition.

Proposition 7.5 *Let $D \leq 1$ and $1 \leq u \leq R$, and let $x, y, \gamma, \hat{\gamma}$ be as above. If the image of γ is contained in $\boldsymbol{B}(u)$, then we have*

$$l(\hat{\gamma}) - l(\gamma) \leq \sqrt{C_1(1+u^\kappa)}\sqrt{D}, \quad l_\infty(\hat{\gamma}) - l_\infty(\gamma) \leq \sqrt{C_1 u^\kappa}\sqrt{D}.$$

Proposition 7.6 *Let $x, y, \gamma, \hat{\gamma}$ be as above. If the image of γ is contained in $\boldsymbol{B}(u)$, then the image of $\hat{\gamma}$ is contained in $\boldsymbol{B}(u+D)\backslash L(D)$.*

Proof It is obvious by the construction that the image of $\hat{\gamma}$ is contained in $\mathbb{R}^3\backslash L(D)$. Since S^1–action preserves $\boldsymbol{B}(u)$, and

$$|P_\gamma|^2 \le |\gamma|^2 + D^2, \qquad \left|\left(\gamma_{\mathbb{R}}(t), \frac{D\gamma_{\mathbb{C}}(t)}{|\gamma_{\mathbb{C}}(t)|}\right)\right|^2 \le |\gamma|^2 + D^2$$

hold, we have the assertion. □

7.3 Estimates (3)

Let

$$\begin{aligned} \mathrm{Path}(u, D, x, y) &:= \{\gamma \in \mathrm{Path}(x, y) : \mathrm{Im}(\gamma) \subset K(u, D)\}, \\ d_{u,D}(x, y) &:= \inf_{\gamma \in \mathrm{Path}(u,D,x,y)} l(\gamma), \\ d_{\infty,u,D}(x, y) &:= \inf_{\gamma \in \mathrm{Path}(u,D,x,y)} l_\infty(\gamma). \end{aligned}$$

for $x, y \in K(u, D)$. By the definition, $d(x, y) \le d_{u,D}(x, y)$ always holds. In this subsection, we consider the opposite estimate.

Lemma 7.7 *Let $\hat{\zeta} := (\zeta_{\mathbb{R}}, D\zeta_{\mathbb{C}}/|\zeta_{\mathbb{C}}|)$ if $\zeta_{\mathbb{C}} \ne 0$, and $\hat{\zeta} := (\zeta_{\mathbb{R}}, D)$ if $\zeta_{\mathbb{C}} = 0$. Suppose Φ, Φ_∞ satisfy* **(A3)–(A6)**, *and $1 \le u \le R$.*

(1) *If $\zeta \in L(D) \cap \boldsymbol{B}(u-1)$ and $0 < D \le 1$, then*

$$d_u(\zeta, \hat{\zeta}) \le 2\sqrt{C_1(1+u^\kappa)D}, \quad d_{\infty,u}(\zeta, \hat{\zeta}) \le 2\sqrt{C_1 u^\kappa D}.$$

(2) *If $\zeta \in L(D) \cap K(u-1, D)$ and $0 < D \le 1$, then*

$$d_{u,D}(\zeta, \hat{\zeta}) \le 2\sqrt{C_1(1+u^\kappa)D}, \quad d_{\infty,u,D}(\zeta, \hat{\zeta}) \le 2\sqrt{C_1 u^\kappa D}.$$

Proof Let $\gamma(t) = (\zeta_{\mathbb{R}}, t\hat{\zeta}_{\mathbb{C}})$ for $t \in [|\zeta_{\mathbb{C}}|/D, 1]$. Then γ is joining ζ and $\hat{\zeta}$, and the image of γ is contained in $\boldsymbol{B}(u-1+D) \subset \boldsymbol{B}(u)$. Then by **(A3)** and **(A6)**, we have $\Phi(\gamma(t)) \le C_1(1+u^\kappa)/(tD)$. Then we have

$$d_u(\zeta, \hat{\zeta}) \le l(\gamma) \le 2\sqrt{C_1(1+u^\kappa)D}.$$

Moreover, if $\zeta \in K(u-1, D)$, then the image of γ is contained in $K(u, D)$; therefore,

$$d_{u,D}(\zeta, \hat{\zeta}) \le l(\gamma) \le 2\sqrt{C_1(1+u^\kappa)D}.$$

The estimates for $d_{\infty,u}(\zeta, \hat{\zeta})$ and $d_{\infty,u,D}(\zeta, \hat{\zeta})$ follow in the same way. □

Proposition 7.8 *Suppose Φ, Φ_∞ satisfy* **(A3)–(A6)** *and let ρ be as in Proposition 7.3. If $\rho(u+1)+1 \le R$, then*

$$|d_{\rho(u+1)+1,D}(x,y) - d(x,y)| \le \xi(u)\sqrt{D},$$
$$|d_{\infty,\rho(u+1)+1,D}(x,y) - d_\infty(x,y)| \le \xi_\infty(u)\sqrt{D}$$

hold for any $x, y \in K(u,D)$ and $0 < D \le 1$, where

$$\xi(u) := \sqrt{C_1(1+(\rho(u+1)+1)^\kappa)} + 8\sqrt{C_1(1+(u+1)^\kappa)} + 2,$$
$$\xi_\infty(u) := \sqrt{C_1(\rho(u+1)+1)^\kappa} + 8\sqrt{C_1(u+1)^\kappa} + 2.$$

Proof Since $d(x,y) \le d_{\rho(u+1)+1,D}(x,y)$ always holds, it suffices to show that $d_{\rho(u+1)+1,D}(x,y) - d(x,y) \le \xi(u)\sqrt{D}$. Let $x, y \in K(u,D)$ and $0 < D \le 1$. By the assumption $\rho(u+1)+1 \le R$ and the definition of ρ, we have that $u+1 \le R$. Define $\hat{x} \in \mathbb{R}^3$ as in Lemma 7.7 if $x \in L(D)$, and $\hat{x} := x$ if $x \notin L(D)$. Define $\hat{y}$ in the same way. Then we can see $\hat{x}, \hat{y} \in \boldsymbol{B}(u+1) \setminus L(D)$ and $d_{u+1,D}(x,\hat{x}) \le 2\sqrt{C_1(1+(u+1)^\kappa)D}$ by Lemma 7.7; consequently, we obtain

$$d_{u+1,D}(x,\hat{x}) + d_{u+1,D}(y,\hat{y}) \le 4\sqrt{C_1(1+(u+1)^\kappa)D}. \tag{10}$$

For any $\gamma \in \mathrm{Path}(\hat{x},\hat{y})$, we construct $F(\gamma) \in \mathrm{Path}(\rho(u+1)+1, D, \hat{x}, \hat{y})$ as follows. By Proposition 7.3, we can see

$$l(\gamma) \ge d(\hat{x},\hat{y}) = d_{\rho(u+1)}(\hat{x},\hat{y}) = \inf_{c \in \mathrm{Path}(\rho(u+1),x,y)} l(c).$$

Accordingly, we can take $c \in \mathrm{Path}(\rho(u+1),x,y)$ such that $l(c) \le l(\gamma) + \sqrt{D}$. Then we can apply the argument in Section 7.2 to $\hat{x}, \hat{y}$ and c so that we obtain a piecewise smooth path $\hat{c}$ whose image is contained in $\boldsymbol{B}(\rho(u+1)+1) \setminus L(D)$, hence in $K(\rho(u+1)+1, D)$. Then we have

$$\liminf_{k\to\infty} l(\hat{c}) - l(c) \le \sqrt{C_1(1+(\rho(u+1)+1)^\kappa)D}$$

by Proposition 7.5. Therefore, there is a sufficiently large k, which may depend on n and D, such that $l(\hat{c}) - l(c) \le \sqrt{C_1(1+(\rho(u+1)+1)^\kappa)D} + \sqrt{D}$. Put $F(\gamma) = \hat{c}$. Then we can see

$$\begin{aligned} l(F(\gamma)) - l(\gamma) &\le l(F(\gamma)) - l(c) + l(c) - l(\gamma) \\ &\le \sqrt{C_1(1+(\rho(u+1)+1)^\kappa)D} + \sqrt{D} + \sqrt{D} \\ &= \big(\sqrt{C_1(1+(\rho(u+1)+1)^\kappa)} + 2\big)\sqrt{D}. \end{aligned}$$

Thus we obtain $F(\gamma) \in \mathrm{Path}(\rho(u+1)+1, D, \hat{x}, \hat{y})$ for every $\gamma \in \mathrm{Path}(\hat{x},\hat{y})$ such that

$$l(F(\gamma)) - l(\gamma) \le \big(\sqrt{C_1(1+(\rho(u+1)+1)^\kappa)} + 2\big)\sqrt{D}. \tag{11}$$

By taking the infimum of (11) for all $\gamma \in \mathrm{Path}(\hat{x}, \hat{y})$, we obtain

$$d_{\rho(u+1)+1,D}(\hat{x}, \hat{y}) \le d(\hat{x}, \hat{y}) + \left(\sqrt{C_1(1+(\rho(u+1)+1)^\kappa)} + 2\right)\sqrt{D}. \tag{12}$$

Since $\rho(u+1) \ge u+1$, we have

$$\begin{aligned} |d_{\rho(u+1)+1,D}(\hat{x}, \hat{y}) - d_{\rho(u+1)+1,D}(x, y)| &\le d_{\rho(u+1)+1,D}(\hat{x}, x) + d_{\rho(u+1)+1,D}(\hat{y}, y) \\ &\le d_{u+1,D}(\hat{x}, x) + d_{u+1,D}(\hat{y}, y) \\ &\le 4\sqrt{C_1(1+(u+1)^\kappa)D}, \\ |d(\hat{x}, \hat{y}) - d(x, y)| &\le d(\hat{x}, x) + d(\hat{y}, y) \\ &\le d_{u+1,D}(\hat{x}, x) + d_{u+1,D}(\hat{y}, y) \\ &\le 4\sqrt{C_1(1+(u+1)^\kappa)D} \end{aligned}$$

by (10); hence

$$\begin{aligned} d_{\rho(u+1)+1,D}(x, y) &\le d_{\rho(u+1)+1,D}(\hat{x}, \hat{y}) + 4\sqrt{C_1(1+(u+1)^\kappa)D}, \\ d(\hat{x}, \hat{y}) &\le d(x, y) + 4\sqrt{C_1(1+(u+1)^\kappa)D} \end{aligned}$$

hold. By combining these inequalities with (12), we obtain

$$d_{\rho(u+1)+1,D}(x, y) \le d(x, y) + \xi(u)\sqrt{D}.$$

The second inequality can be shown in the same way. □

7.4 From (A3)–(A6) to (A1) and (A2)

Proposition 7.9 *Suppose that Φ, Φ_∞ satisfy (A3)–(A6), and let $\gamma\colon [a,b] \to K(u, D)$ and $1 \le u \le R$. Then*

$$|l(\gamma) - l_\infty(\gamma)| \le \sqrt{\frac{\varepsilon u}{C_0 D^m}}\, l_\infty(\gamma).$$

Proof Since $l(\gamma) = \int_a^b \sqrt{\Phi(\gamma(t))}\,|\gamma'(t)|\,dt$, one can see

$$\begin{aligned} |l(\gamma) - l_\infty(\gamma)| &\le \int_a^b \sqrt{|\Phi(\gamma) - \Phi_\infty(\gamma)|}\,|\gamma'|\,dt \\ &\le \int_a^b \sqrt{\frac{|\Phi(\gamma) - \Phi_\infty(\gamma)|}{\Phi_\infty(\gamma)}}\sqrt{\Phi_\infty(\gamma)}\,|\gamma'|\,dt \\ &\le \int_a^b \sqrt{\frac{\varepsilon \max\{|\gamma|, 1\}}{C_0 D^m}}\sqrt{\Phi_\infty(\gamma(t))}\,|\gamma'(t)|\,dt \end{aligned}$$

by (A3) and (A5). Since we have assumed $|\gamma| \le u$ and $u \ge 1$, we have

$$|l(\gamma) - l_\infty(\gamma)| \le \sqrt{\frac{\varepsilon u}{C_0 D^m}}\, l_\infty(\gamma). \qquad \square$$

Proposition 7.10 *Suppose that Φ and Φ_∞ satisfy* (A3), (A5) *and* (A6). *Then*

$$|d_{u,D}(x,y) - d_{\infty,u,D}(x,y)| \le \sqrt{\frac{\varepsilon u}{C_0 D^m}}\, d_{\infty,u,D}(x,y)$$

holds for all $1 \le u \le R$.

Proof Put $\delta = \sqrt{\varepsilon u/(C_0 D^m)}$. Then Proposition 7.9 gives

$$(1-\delta) l_\infty(\gamma) \le l(\gamma) \le (1+\delta) l_\infty(\gamma). \tag{13}$$

Then by taking the infimum of (13) for all $\gamma \in \mathrm{Path}(u, D, x, y)$, we can see

$$(1-\delta) d_{\infty,u,D}(x,y) \le d_{n,u,D}(x,y) \le (1+\delta) d_{\infty,u,D}(x,y)$$

for all $u \ge 0$. $\square$

Proposition 7.11 *Suppose that Φ, Φ_∞ satisfy* (A3)–(A6) *and $u \le 1$. Let $u^{(2)} := \rho(u+2)+1 \le R$. Then for all $x, y \in \boldsymbol{B}(u)$, we have*

$$\begin{aligned}|d(x,y) - d_\infty(x,y)|& \\ &\le 26\sqrt{C_1(1+R^\kappa)D} + 4\sqrt{D} + \sqrt{\frac{\varepsilon R}{C_0 D^m}}\big(C_2 R^{\kappa'} + (9\sqrt{C_1 R^\kappa} + 2)\sqrt{D}\big).\end{aligned}$$

Proof Put $u^{(1)} = \rho(u+1)+1$ and let $x, y \in K(u, D)$. Then $u^{(1)} \le R$. By combining Propositions 7.8 and 7.10, we have

$$\begin{aligned}|d(x,y) - d_\infty(x,y)| &\le |d(x,y) - d_{u^{(1)},D}(x,y)| + |d_{u^{(1)},D}(x,y) - d_{\infty,u^{(1)},D}(x,y)| \\ &\qquad + |d_\infty(x,y) - d_{\infty,u^{(1)},D}(x,y)| \\ &\le \xi(u)\sqrt{D} + \xi_\infty(u)\sqrt{D} + \sqrt{\frac{\varepsilon u^{(1)}}{C_0 D^m}}\, d_{\infty,u^{(1)},D}(x,y) \\ &\le 2\xi(u)\sqrt{D} + \sqrt{\frac{\varepsilon u^{(1)}}{C_0 D^m}}\big(d_\infty(x,y) + \xi_\infty(u)\sqrt{D}\big).\end{aligned}$$

By Proposition 7.2, $D_{\infty,u} < C_2 u^{\kappa'}$ holds if $u \ge 1$; consequently, $d_\infty(x,y)$ is not more than $C_2 u^{\kappa'}$. Therefore, for all $x, y \in K(u, D)$, we obtain

$$|d(x,y) - d_\infty(x,y)| \le 2\xi(u)\sqrt{D} + \sqrt{\frac{\varepsilon u^{(1)}}{C_0 D^m}}\big(C_2 u^{\kappa'} + \xi_\infty(u)\sqrt{D}\big).$$

Next we consider the case of $x \in \boldsymbol{B}(u)$, but $x \notin K(u, D)$. In this case, $x \in \boldsymbol{B}(u) \cap L(D)$ holds, hence we can apply Lemma 7.7. Let $\hat{x}$ be as in Lemma 7.7. Then we can see

$$d(x, \hat{x}) \le 2\sqrt{C_1(1 + (u+1)^\kappa)D},$$

and $\hat{x}$ is contained in $K(u+1, D)$. Here we suppose that y is also contained in $\boldsymbol{B}(u) \cap L(D)$ and follow the same procedure. If y is in $K(u, D)$, then suppose $y = \hat{y}$ in the following discussion. Now we have

$$|d(x, y) - d(\hat{x}, \hat{y})| \le d(x, \hat{x}) + d(y, \hat{y}) \le 4\sqrt{C_1(1 + (u+1)^\kappa)D}.$$

Hence we can see

$$\begin{aligned}
&|d(x,y) - d_\infty(x,y)| \\
&\le 8\sqrt{C_1(1+(u+1)^\kappa)D} + |d(\hat{x},\hat{y}) - d_\infty(\hat{x},\hat{y})| \\
&\le 8\sqrt{C_1(1+(u+1)^\kappa)D} + 2\xi(u+1)\sqrt{D} + \sqrt{\frac{\varepsilon u^{(2)}}{C_0 D^m}}\big(C_2(u+1)^{\kappa'} + \xi_\infty(u+1)\sqrt{D}\big).
\end{aligned}$$

Since $\xi(u)$ is monotonically increasing and $u + 2 \le u^{(2)} \le R$ holds, we have

$$\xi(u+1) \le 9\sqrt{C_1(1 + R^\kappa)} + 2, \quad \xi_\infty(u+1) \le 9\sqrt{C_1 R^\kappa} + 2. \qquad \square$$

Corollary 7.12 *Suppose that Φ and Φ_∞ satisfy* **(A3)–(A6)** *and that $\varepsilon \le 1$, and let $u^{(2)} := \rho(u+2) + 1 \le R$. Then there exists a constant C independent of any other constants such that, for all $x, y \in \boldsymbol{B}(u)$,*

$$|d(x, y) - d_\infty(x, y)| < C(1 + \sqrt{C_1})(1 + C_0^{-\frac{1}{2}}) R^{1+\frac{\kappa}{2}} \varepsilon^{\frac{1}{2(1+m)}}.$$

Proof In Proposition 7.11, let $D = \varepsilon^{1/(1+m)} \le 1$. As described in the proof of Proposition 7.2, C_2 is linearly dependent on $\sqrt{C_1}$. Then assertion follows by using $R \ge 1$, $\varepsilon \le 1$ and unifying constants. $\square$

Proposition 7.13 *Suppose that $\Phi(\zeta) \ge A/|\zeta|$ holds for some $A > 0$ and all ζ with $|\zeta| \le 1$, and let $u(r) := \big(1 + \frac{1}{2}A^{-1/2}r\big)^2$. Then $B(0, r) \subset \boldsymbol{B}(u(r))$ holds for all $r > 0$, where $B(0, r)$ is the metric ball with respect to d.*

Proof Let $\zeta \in B(0, r)$. Then by the same argument as in the proof of the first inequality of Proposition 7.2, we have

$$2\sqrt{A}(\sqrt{|\zeta|} - 1) \le d(0, \zeta) < r,$$

which gives $|\zeta| < \big(1 + \frac{1}{2}A^{-1/2}r\big)^2 = u(r)$. $\square$

Proposition 7.14 *Suppose that Φ, Φ_∞ satisfy* **(A3)–(A6)**, *and suppose $\varepsilon \le 1$. Then the identity map of $\mathbb{R}^3$ is an (r,δ)–isometry from $(\mathbb{R}^3, d, 0)$ to $(\mathbb{R}^3, d_\infty, 0)$, where $r, \delta > 0$ are defined by*

$$\rho(u(r)+2)+1 = R, \quad \delta = C(1+\sqrt{C_1})(1+C_0^{-\frac{1}{2}})R^{1+\frac{\kappa}{2}}\varepsilon^{\frac{1}{2(1+m)}}.$$

Proof Let $x, y \in B(0,r)$. Then $x, y \in \boldsymbol{B}(u(r))$; hence

$$|d(x,y) - d_\infty(x,y)| < C(1+\sqrt{C_1})(1+C_0^{-\frac{1}{2}})R^{1+\frac{\kappa}{2}}\varepsilon^{\frac{1}{2(1+m)}} \tag{14}$$

holds. Next we show $B_\infty(0, r-\delta) \subset B(B(0,r),\delta)$. If $x \in B_\infty(0, r-\delta)$, then $x \in \boldsymbol{B}(u(r))$ holds; therefore, (14) gives

$$\begin{aligned} d(0,x) &< d_\infty(0,x) + C(1+\sqrt{C_1})(1+C_0^{-\frac{1}{2}})R^{1+\frac{\kappa}{2}}\varepsilon^{\frac{1}{2(1+m)}} \\ &< r - \delta + C(1+\sqrt{C_1})(1+C_0^{-\frac{1}{2}})R^{1+\frac{\kappa}{2}}\varepsilon^{\frac{1}{2(1+m)}} = r, \end{aligned}$$

which implies $B_\infty(0, r-\delta) \subset B(0,r)$. □

By Propositions 7.13 and 6.3, the following estimate is obtained.

Proposition 7.15 *Let Φ_a be as in Section 6 and assume $\sum_{n=0}^\infty A_{S_n,P}^{T_n} - 2a^{1/(1+\alpha)} > 0$. Then $\sup_{\zeta \in B(0,r)} 1/(N\sqrt{\Phi_a(\zeta)})$ is not more than*

$$\frac{(a/P)^{\frac{1}{1+\alpha}}}{\sqrt{\sum_{n=0}^\infty A_{S_n,P}^{T_n} - 2(a/P)^{\frac{1}{1+\alpha}}}}\left(1 + \frac{r}{2\sqrt{\sum_{n=0}^\infty A_{S_n,P}^{T_n} - 2(a/P)^{\frac{1}{1+\alpha}}}}\right).$$

Combining Propositions 7.14 and 7.15, we obtain the following theorem.

Theorem 7.16 *Let $a_i, P_i, n_i > 0$, $\lim_{i\to\infty} a_i = 0$ and $\lim_{i\to\infty} n_i \to \infty$. Put S_{i,n_i} and T_{i,n_i} as in Section 6. Suppose that there are constants $\varepsilon = \varepsilon_i(R)$, C_0, C_1, κ, m for all $R \ge 1$ such that $\Phi = \Phi_{a_i}$ and Φ_∞ satisfy* **(A3)–(A6)**. *If*

$$\lim_{i\to\infty} \varepsilon_i(R) = \lim_{i\to\infty} \frac{a_i}{P_i} = 0, \quad \liminf_{i\to\infty} \sum_{l=0}^{\infty} A_{S_{i,n_i},P_i}^{T_{i,n_i}} > 0$$

and C_0, C_1, κ, m are independent of i, R, then

$$\{(X, a_i g_\Lambda, p)\}_i \xrightarrow[i\to\infty]{\mathrm{GH}} (\mathbb{R}^3, d_\infty, 0).$$

Proof Fix $r>0$ and $\delta>0$ arbitrarily. Put $R(r)=\rho(u(r)+2)+1$, and let $C>0$ be the constant in Corollary 7.12. By the assumption, there exists $i(r,\delta)>0$ such that

$$C(1+\sqrt{C_1})(1+C_0^{-\frac{1}{2}})R(r)^{1+\frac{\kappa}{2}}\varepsilon_i(R(r))^{\frac{1}{2(1+m)}}<\tfrac{1}{2}\delta$$

holds for all $i\geq i(r,\delta)$. Then by Proposition 7.14, $\mathrm{id}_{\mathbb{R}^3}$ is an $\left(r,\frac{1}{2}\delta\right)$–isometry from $(\mathbb{R}^3,d_{a_i},0)$ to $(\mathbb{R}^3,d_{a_\infty},0)$. By Proposition 7.15, we can take $i'(r,\delta)\geq i(r,\delta)$ such that $\sup_{\zeta\in B(0,r)}1/(N\sqrt{\Phi_a(\zeta)})<\frac{1}{2}\delta$ for all $i\geq i'(r,\delta)$. Then Proposition 5.1 gives the assertion. □

8 Convergence

In this section, we consider the convergence of $\{(X,a_ig_\Lambda)\}_i$, where Λ is as defined in Section 6, and $\{a_i\}_i$ is a sequence with $a_i>0$ and $\lim_{i\to\infty}a_i$, applying Theorem 7.16. To apply them, we have to estimate constants ε, C_0, C_1 in **(A3)–(A6)** uniformly with respect to $i\in\mathbb{N}$, and show that $\varepsilon\to 0$ as $i\to\infty$. In Section 8.1, we consider the uniform estimate for the case of $P=1$, which is the simplest case. In Sections 8.2 and 8.3, we suppose P is depending on some parameters. Then we apply them to show Theorems 1.2 and 1.3 in Sections 8.4 and 8.5.

Put $S_{a,n}:=a^{1/(1+\alpha)}K_{2n}$ and $T_{a,n}:=a^{1/(1+\alpha)}K_{2n+1}$. We take a subsequence

$$\{K_{n_0}<K_{n_1}<K_{n_2}<\cdots\}\subset\{K_0<K_1<K_2<\cdots\}.$$

We are now going to consider the convergence (in several cases according to the rate of the convergence of $\{a_i\}_i$) or the divergence of $\{K_n\}_n$.

From now on, we put

$$\Phi_S^T(\zeta):=\Phi_{S,1}^T(\zeta)=\int_S^T\frac{dx}{|\zeta-(x^\alpha,0,0)|},\quad A_S^T:=A_{S,1}^T=\int_S^T\frac{dx}{1+x^\alpha}.$$

8.1 Convergence (1)

Fix $a>0$, n and $0\leq S<T\leq\infty$, and put $P=1$.

Proposition 8.1 *Let $R\geq 1$ and $D\leq 1$. There exists a constant $C_\alpha>0$ depending only on α such that*

$$|\Phi_a(\zeta)-\Phi_S^T(\zeta)|\leq\frac{C_\alpha\varepsilon_n}{D}$$

holds for any $\zeta\in K(R,D)$, where

$$\varepsilon_{a,n}=a^{\frac{1}{1+\alpha}}+\frac{K_{2n-1}}{K_{2n}}S_{a,n}+\left(\frac{K_{2n+2}}{K_{2n+1}}T_{a,n}\right)^{-\alpha+1}+|S_{a,n}-S|+|T_{a,n}^{-\alpha+1}-T^{-\alpha+1}|.$$

Proof By combining Proposition 6.1, (7) and (8), we have

$$\left|\Phi_a(\zeta) - \Phi_{S_{a,n}}^{T_{a,n}}(\zeta)\right| \le \frac{2a^{\frac{1}{1+\alpha}}}{D} + \frac{T_{a,n-1}}{D} + \frac{2S_{a,n+1}^{-\alpha+1}}{\alpha-1}$$

if $S_{a,n+1} \ge (2|\zeta|)^{1/\alpha}$. Here $|\zeta| \le R$, and

$$S_{a,n+1} = a^{\frac{1}{1+\alpha}} K_{2n+2} = \frac{K_{2n+2}}{K_{2n+1}} T_{a,n},$$
$$T_{a,n-1} = a^{\frac{1}{1+\alpha}} K_{2n-1} = \frac{K_{2n-1}}{K_{2n}} S_{a,n}.$$

On the other hand, we can see

$$\left|\Phi_{S_{a,n}}^{T_{a,n}}(\zeta) - \Phi_S^T(\zeta)\right| \le \frac{|S_{a,n} - S|}{D} + \frac{2|T_{a,n}^{-\alpha+1} - T^{-\alpha+1}|}{\alpha-1}.$$

Thus we obtain the assertion. □

Now, we put $\Phi = \Phi_a$, $\Phi_\infty = \Phi_S^T$, and suppose a, $|S_{a,n} - S|$ and $|T_{a,n}^{-\alpha+1} - T^{-\alpha+1}|$ are sufficiently small. Then the constants in **(A3)–(A6)** can be taken uniformly as

$$C_0 = \frac{1}{2} A_S^T, \quad C_1 = \frac{\alpha 2^{\frac{1}{\alpha}}}{\alpha - 1}, \quad m = 1, \quad \kappa = \frac{1}{\alpha}.$$

Then by Proposition 8.1, if $\lim_{n\to\infty} K_{2n+1}/(2n) = \infty$, then we have $\varepsilon_{a,n} \to 0$ as $a \to 0$, $n \to \infty$, $|S_{a,n} - S| \to 0$ and $|T_{a,n}^{-\alpha+1} - T^{-\alpha+1}| \to 0$. Hence by Theorem 7.16, we have the next results.

Theorem 8.2 *Let (X, g_Λ) be as in Section 6 and suppose $\lim_{n\to\infty} K_{2n+1}/K_{2n} = \infty$. Assume that $\{a_i\}_i \subset \mathbb{R}^+$ and*

$$\{K_{n_0} < K_{n_1} < K_{n_2} < \cdots\} \subset \{K_0 < K_1 < K_2 < \cdots\}$$

satisfy

$$\lim_{i\to\infty} a_i^{\frac{1}{1+\alpha}} K_{2n_i} = S \ge 0, \quad \lim_{i\to\infty} a_i^{\frac{1}{1+\alpha}} K_{2n_i+1} = T \le \infty, \quad S < T.$$

Then $\{(X, a_n g_\Lambda, p)\}_n \xrightarrow{\mathrm{GH}} (\mathbb{R}^3, d_S^T, 0)$, where d_S^T is the metric induced by $\Phi_S^T \cdot h_0$.

8.2 Convergence (2)

Let (X, d_X, p), (Y, d_Y, q) be pointed metric spaces and suppose $\lim_{n\to\infty} a_n = 0$. Assume that $\{(X, a_n d_X, p)\}_n \xrightarrow{\mathrm{GH}} (Y, d_Y, q)$. It is easy to check that we have

$\{(X, sa_n d_X, p)\}_n \xrightarrow{\mathrm{GH}} (Y, sd_Y, q)$ for any $s > 0$. Moreover, if $\{a_{n,m}\}_{n,m\in\mathbb{N}}$ satisfies $\lim_{n\to\infty} a_{n,m} = 0$ for every m, and

$$\{(X, a_{n,m} d_X, p)\}_n \xrightarrow{\mathrm{GH}} (Y_m, d_{Y_m}, q_m), \quad \{(Y_m, d_{Y_m}, q_m)\}_m \xrightarrow{\mathrm{GH}} (Y, d_Y, q)$$

hold for every m, then by the diagonal argument one can show there exists a subset $\{a_{n,m(n)}\}_n \subset \{a_{n,m}\}_{n,m}$ such that $\lim_{n\to\infty} a_{n,m(n)} = 0$ and

$$\{(X, a_{n,m(n)} d_X, p)\}_n \xrightarrow{\mathrm{GH}} (Y, d_Y, q).$$

Now, let $\mathcal{T}(X,d)$ be the set of isometry classes of tangent cones at infinity of (X,d). From the above argument, one can see that $\mathcal{T}(X,d)$ is closed with respect to the pointed Gromov–Hausdorff topology, and if $(Y,d') \in \mathcal{T}(X,d)$, then its rescaling (Y, ad') is also contained in $\mathcal{T}(X,d)$.

From Section 8.1, $(\mathbb{R}^3, d_S^T, 0)$ may appear as the tangent cone at infinity of some (X, g_Λ), where Λ is as in Section 6.

Let $\sigma > 0$, $0 \le S < T \le \infty$ and $I_\sigma\colon \zeta \mapsto \sigma^{-1}\zeta$ be the dilation. Then we have

$$I^*_{P^{\frac{1}{1+\alpha}}}(\Phi_S^T h_0) = P^{\frac{-1}{\alpha}}\,\Phi_{S''}^{T''}(\zeta) h_0 = \Phi_{S',P}^{T'}(\zeta) h_0,$$

where

$$S' = P^{\frac{-1}{1+\alpha}} S, \quad T' = P^{\frac{-1}{1+\alpha}} T, \quad S'' = P^{\frac{1}{\alpha(1+\alpha)}} S, \quad T'' = P^{\frac{1}{\alpha(1+\alpha)}} T. \tag{15}$$

Hence if $(\mathbb{R}^3, d_S^T, 0) \in \mathcal{T}(X, g_\Lambda)$, then $\{(\mathbb{R}^3, d_{\sigma S}^{\sigma T}, 0)\}_{\sigma\in\mathbb{R}^+}$ is also contained in $\mathcal{T}(X, g_\Lambda)$.

Fix a constant $\theta > 0$, put $P^{1/(1+\alpha)} = \theta\sqrt{S^{-\alpha+1} - T^{-\alpha+1}} > 0$, and let S', T' be defined by (15).

Proposition 8.3 *Let $R \ge 1$. There is a constant $C > 0$ depending only on α such that*

$$\left|\Phi_{S',P}^{T'}(\zeta) - \frac{1}{\theta^2(\alpha-1)}\right| \le \frac{CR}{\theta^3 S^\alpha \sqrt{S^{-\alpha+1} - T^{-\alpha+1}}}$$

holds for any $\zeta \in K(R,D)$ if $\theta S^\alpha \sqrt{S^{-\alpha+1} - T^{-\alpha+1}} \ge 2R$.

Proof Let S'' and T'' be defined by (15). Note that

$$\Phi_{S',P}^{T'}(\zeta) = P^{\frac{-1}{\alpha}} \int_{S''}^{T''} \frac{dx}{|\zeta - (x^\alpha, 0, 0)|}.$$

By the assumption, we have $P^{1/(1+\alpha)}S^{\alpha} = \theta S^{\alpha}\sqrt{S^{-\alpha+1} - T^{-\alpha+1}} \geq 2R$; then we see

$$\begin{aligned}\left|\int_{S''}^{T''} \frac{dx}{|\zeta - (x^{\alpha},0,0)|} - \int_{S''}^{T''} \frac{dx}{x^{\alpha}}\right| &\\ &\leq \int_{S''}^{T''} \left|\frac{1}{|\zeta-(x^{\alpha},0,0)|} - \frac{1}{x^{\alpha}}\right| dx \\ &\leq \int_{S''}^{T''} \frac{2x^{\alpha}|\zeta| + |\zeta|^2}{|\zeta - (x^{\alpha},0,0)|x^{\alpha}(|\zeta - (x^{\alpha},0,0)| + x^{\alpha})}\, dx \\ &\leq \int_{S''}^{T''} \frac{8R}{x^{2\alpha}}\, dx + \int_{S''}^{T''} \frac{4R^2}{x^{3\alpha}}\, dx \\ &\leq \frac{8RP^{\frac{-2\alpha+1}{\alpha(1+\alpha)}}}{2\alpha-1}(S^{-2\alpha+1} - T^{-2\alpha+1}) + \frac{4R^2P^{\frac{-3\alpha+1}{\alpha(1+\alpha)}}}{3\alpha-1}(S^{-3\alpha+1} - T^{-3\alpha+1}).\end{aligned}$$

Since we have

$$\int_{S''}^{T''} \frac{dx}{x^{\alpha}} = \frac{P^{\frac{-\alpha+1}{\alpha(1+\alpha)}}}{\alpha-1}(S^{-\alpha+1} - T^{-\alpha+1}) = \frac{P^{\frac{1}{\alpha}}}{\theta^2(\alpha-1)},$$

we obtain

$$\begin{aligned}\left|\Phi_{S',P}^{T'}(\zeta) - \frac{1}{\theta^2(\alpha-1)}\right| &\\ &\leq \frac{8RP^{\frac{-3}{1+\alpha}}}{2\alpha-1}(S^{-2\alpha+1} - T^{-2\alpha+1}) + \frac{4R^2P^{\frac{-4}{1+\alpha}}}{3\alpha-1}(S^{-3\alpha+1} - T^{-3\alpha+1}).\end{aligned}$$

Using the assumption $2R \leq P^{1/(1+\alpha)}S^{\alpha}$ once more, we have

$$\begin{aligned}\left|\Phi_{S',P}^{T'}(\zeta) - \frac{1}{\theta^2(\alpha-1)}\right| &\\ &\leq \frac{8RP^{\frac{-3}{1+\alpha}}}{2\alpha-1}(S^{-2\alpha+1} - T^{-2\alpha+1}) + \frac{2RP^{\frac{-3}{1+\alpha}}}{3\alpha-1}(S^{-2\alpha+1} - S^{\alpha}T^{-3\alpha+1}) \\ &\leq \theta^{-3}C_{\alpha}RS^{-\frac{1+\alpha}{2}}\frac{1-(S/T)^{3\alpha-1}}{(1-(S/T)^{\alpha-1})^{\frac{3}{2}}}.\end{aligned}$$

Now, put $f(x) := (1 - x^{3\alpha-1})/((1 - x^{\alpha-1})^{3/2})$ for $0 \leq x < 1$. Then there exists a constant $C'_{\alpha} > 0$ such that $f(x) \leq C'_{\alpha}(1 - x^{\alpha-1})^{-1/2}$ holds for all $0 \leq x < 1$. Consequently, by replacing C_{α} larger if necessary, we can see

$$\left|\Phi_{S',P}^{T'}(\zeta) - \frac{1}{\theta^2(\alpha-1)}\right| \leq \frac{C_{\alpha}R}{\theta^3S^{\alpha}\sqrt{S^{-\alpha+1} - T^{-\alpha+1}}}. \qquad \square$$

Proposition 8.4 *Suppose $\theta S^{\alpha}\sqrt{S^{-\alpha+1}-T^{-\alpha+1}}\geq 2R$ for $R\geq 1$. Then*

$$A_{S',P}^{T'}\geq\frac{1}{2\theta^2(\alpha-1)},\quad \Phi_{S',P}^{T'}(\zeta_{\mathbb{R}},\zeta_{\mathbb{C}})\leq\frac{2|\zeta|}{\theta^2(\alpha-1)|\zeta_{\mathbb{C}}|}$$

holds for any $\zeta=(\zeta_{\mathbb{R}},\zeta_{\mathbb{C}})\in\mathbb{R}^3=\mathbb{R}\oplus\mathbb{C}$ with $|\zeta|\leq R$.

Proof We have $1\leq S^{-\alpha}x^{\alpha}$ for all $x\geq S$, then we can see

$$\begin{aligned}A_{S',P}^{T'}&\geq P^{-\frac{1}{1+\alpha}}\int_S^T\frac{dx}{S^{-\alpha}x^{\alpha}+P^{\frac{1}{1+\alpha}}x^{\alpha}}\\&=\frac{1}{P^{\frac{1}{1+\alpha}}(S^{-\alpha}+P^{\frac{1}{1+\alpha}})}\int_S^T\frac{dx}{x^{\alpha}}\\&=\frac{1}{P^{\frac{1}{1+\alpha}}S^{-\alpha}(1+S^{\alpha}P^{\frac{1}{1+\alpha}})}\frac{S^{-\alpha+1}-T^{-\alpha+1}}{\alpha-1}.\end{aligned}$$

Since we have

$$S^{\alpha}P^{\frac{1}{1+\alpha}}=\theta\sqrt{S^{-\alpha+1}-T^{-\alpha+1}}\geq 2R\geq 1,$$

we obtain

$$A_{S',P}^{T'}\geq\frac{S^{-\alpha+1}-T^{-\alpha+1}}{2(\alpha-1)P^{\frac{1}{1+\alpha}}S^{-\alpha}\cdot S^{\alpha}P^{\frac{1}{1+\alpha}}}=\frac{1}{2\theta^2(\alpha-1)}.$$

Next we consider the upper estimate of $\Phi_{S',P}^{T'}(\zeta)$. Take ζ such that $|\zeta|\leq R$; then we have $2|\zeta|\leq P^{1/(1+\alpha)}S^{\alpha}$ by the assumption. Then one can see

$$\begin{aligned}\Phi_{S',P}^{T'}(\zeta)&\leq P^{-\frac{1}{1+\alpha}}\int_S^T\frac{2dx}{P^{\frac{1}{1+\alpha}}x^{\alpha}}=P^{\frac{-2}{1+\alpha}}\frac{2(S^{-\alpha+1}-T^{-\alpha+1})}{\alpha-1}\\&\leq\frac{2}{\theta^2(\alpha-1)}\leq\frac{2|\zeta|}{\theta^2(\alpha-1)|\zeta_{\mathbb{C}}|}.\end{aligned}$$

□

Proposition 8.5 *Let $\Phi=\Phi_{S',P}^{T'}$ and $\Phi_\infty\equiv 1/(\theta^2(\alpha-1))$. Then there exists $C>0$ such that Φ, Φ_∞ satisfy* **(A3)–(A6)** *for $R\geq 1$, and*

$$\varepsilon=\frac{CR}{\theta^3S^{\alpha}\sqrt{S^{-\alpha+1}-T^{-\alpha+1}}},\qquad C_0=\frac{1}{2\theta^2(\alpha-1)},$$

$$C_1=\frac{1}{\theta^2}\max\left\{\frac{1}{\alpha-1},\frac{C}{2}\right\},\quad m=1,\qquad \kappa=1$$

if $\theta S^{\alpha}\sqrt{S^{-\alpha+1}-T^{-\alpha+1}}\geq 2R$.

Proof It is obvious that **(A4)** holds. Proposition 8.4 gives **(A5)** for $C_0 = 1/(2\theta^2(\alpha - 1))$ if we take $\theta S^\alpha \sqrt{S^{-\alpha+1} - T^{-\alpha+1}} \geq 2R$. **(A6)** holds for $C_1 = 1/(\theta^2(\alpha-1))$ since

$$\frac{1}{\alpha-1} = \frac{1}{\alpha-1}\frac{|\zeta|}{|\zeta|} \leq \frac{1}{\alpha-1}\frac{|\zeta|}{|\zeta_{\mathbb{C}}|}.$$

Combining $\theta S^\alpha \sqrt{S^{-\alpha+1} - T^{-\alpha+1}} \geq 2R$ and Proposition 8.3, we can see that $\varepsilon \leq C/(2\theta^2)$. □

Now, Propositions 7.14 and 8.5 with $\theta = 1$ give the following theorem.

Theorem 8.6 *Let $\{S_i\}_i$ and $\{T_i\}_i$ be sequences such that $0 \leq S_i < T_i \leq \infty$ and $\lim_{i\to\infty} S_i^\alpha \sqrt{S_i^{-\alpha+1} - T_i^{-\alpha+1}} = \infty$. Then $\{(\mathbb{R}^3, d_{S_i}^{T_i}, 0)\}_i$ converges to $(\mathbb{R}^3, h_0, 0)$ in the pointed Gromov–Hausdorff topology.*

Next put $P^{1/(1+\alpha)} = \theta|T - S|$ for $0 \leq S < T$ and $\theta > 0$, and let S', T' be as in (15). Then we can show the following similarly to Proposition 8.9.

Proposition 8.7 *Let $D \geq 1$. For all $\zeta \in K(R, D)$, we have*

$$\left|\Phi_{S',P}^{T'}(\zeta) - \frac{1}{\theta|\zeta|}\right| \leq \frac{2}{\theta D}, \qquad \left|\Phi_{S',P}^{T'}(\zeta) - \frac{1}{\theta|\zeta|}\right| \leq \frac{1 + \theta T^\alpha (T-S)}{D^3} T^\alpha (T-S).$$

Proof Let S'' and T'' be defined by (15). The first inequality is obviously shown by $\Phi_{S',P}^{T'}(\zeta) \leq 1/(\theta D)$ and $1/|\zeta| \leq 1/D$. The second inequality follows from

$$\begin{aligned}
&\left|\int_{S''}^{T''} \frac{dx}{|\zeta - (x^\alpha, 0, 0)|} - \int_{S''}^{T''} \frac{dx}{|\zeta|}\right| \\
&\quad\leq \int_{S''}^{T''} \frac{2x^\alpha|\zeta| + x^{2\alpha}}{|\zeta - (x^\alpha,0,0)|\,|\zeta|\,(|\zeta - (x^\alpha,0,0)| + |\zeta|)}\,dx \\
&\quad\leq \int_{S''}^{T''} \frac{2x^\alpha}{D^2}\,dx + \int_{S''}^{T''} \frac{x^{2\alpha}}{D^3}\,dx \\
&\quad\leq C_\alpha \frac{P^{\frac{1}{\alpha}}(T^{\alpha+1} - S^{\alpha+1}) + P^{\frac{2\alpha+1}{\alpha(1+\alpha)}}(T^{2\alpha+1} - S^{2\alpha+1})}{D^3} \\
&\quad= C_\alpha \theta^{1+\frac{1}{\alpha}} T^{\alpha+1}(T-S)^{1+\frac{1}{\alpha}} \frac{1 - (S/T)^{\alpha+1} + \theta(T-S)T^\alpha(1 - (S/T)^{2\alpha+1})}{D^3}.
\end{aligned}$$

By the similar argument to Proposition 8.3, we can replace either $1 - (S/T)^{\alpha+1}$ or $1 - (S/T)^{2\alpha+1}$ by $1 - S/T$; hence we obtain the assertion. □

Proposition 8.8 *For any* $\zeta = (\zeta_{\mathbb{R}}, \zeta_{\mathbb{C}}) \in \mathbb{R}^3 = \mathbb{R} \oplus \mathbb{C}$,

$$A^{T'}_{S',P} \geq \frac{1}{\theta(1+\theta T^\alpha(T-S))}, \quad \Phi^{T'}_{S',P}(\zeta_{\mathbb{R}}, \zeta_{\mathbb{C}}) \leq \frac{1}{\theta|\zeta_{\mathbb{C}}|}.$$

Proof One can see

$$\begin{aligned} A^{T'}_{S',P} = P^{-\frac{1}{1+\alpha}} \int_S^T \frac{dx}{1+P^{\frac{1}{1+\alpha}} x^\alpha} &\geq P^{-\frac{1}{1+\alpha}} \int_S^T \frac{dx}{1+P^{\frac{1}{1+\alpha}} T^\alpha} \\ &\geq \frac{T-S}{P^{\frac{1}{1+\alpha}}(1+P^{\frac{1}{1+\alpha}} T^\alpha)} \\ &= \frac{1}{\theta(1+\theta T^\alpha(T-S))}. \end{aligned}$$

We can also obtain

$$\Phi^{T'}_{S',P}(\zeta) \leq \frac{T-S}{P^{\frac{1}{1+\alpha}}|\zeta_{\mathbb{C}}|} = \frac{1}{\theta|\zeta_{\mathbb{C}}|}. \qquad \square$$

Combining Propositions 8.7 and 8.8, the next proposition is obtained.

Proposition 8.9 *Let* $\Phi = \Phi^{T'}_{S',P}$ *and* $\Phi_\infty(\zeta) = 1/(\theta|\zeta|)$. *Then* Φ, Φ_∞ *satisfy* **(A3)**–**(A6)** *for* $R \geq 1$, *and*

$$\varepsilon = (1+\theta T^\alpha(T-S))T^\alpha(T-S), \quad C_0 = \frac{1}{\theta(1+\theta T^\alpha(T-S))},$$
$$C_1 = \frac{2}{\theta}, \qquad m = 3, \qquad \kappa = 0$$

for any $0 \leq S < T$.

By Propositions 7.14 and 8.9 for $\theta = 1$, we have the next result.

Theorem 8.10 *Let* $\{S_i\}_i$ *and* $\{T_i\}_i$ *be a sequence such that* $0 \leq S_i < T_i$ *and* $\lim_{i\to\infty} T_i^\alpha |T_i - S_i| = 0$. *Then* $\{(\mathbb{R}^3, d^{T_i}_{S_i}, 0)\}_i$ *converges to* $(\mathbb{R}^3, (1/|\zeta|)h_0, 0)$ *in the pointed Gromov–Hausdorff topology.*

8.3 Convergence (3)

Here, we fix $a > 0$ and n, and suppose that $T_{a,n} = a^{1/(1+\alpha)} K_{2n+1}$ is sufficiently small and that $S_{a,n+1} = a^{1/(1+\alpha)} K_{2n+2}$ is sufficiently large. Fix P and θ such that

$$P^{\frac{1}{1+\alpha}} = \theta(T_{a,n} - S_{a,n}) = \sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}}.$$

Put $S'_l = P^{-1/(1+\alpha)} S_{a,l}$ and $T'_l = P^{-1/(1+\alpha)} T_{a,l}$.

Proposition 8.11 *Let $R \geq 1$ and $D \leq 1$, and let P be as above. Assume that $P^{1/(\alpha(1+\alpha))} S_{a,n+2} \geq (2R)^{1/\alpha}$. Then there exists a constant $C_\alpha > 0$ depending only on α such that*

$$\left|\Phi_a(\zeta) - \Phi^{T'_n}_{S'_n,P}(\zeta) - \Phi^{T'_{n+1}}_{S'_{n+1},P}(\zeta)\right| \leq \frac{C_\alpha \varepsilon_{a,n}}{D},$$

for any $\zeta \in K(R,D)$, where $\varepsilon_{a,n}$ is the constant defined by

$$\varepsilon_{a,n} = \frac{1+K_{2n-1}}{\theta(K_{2n+1}-K_{2n})} + \frac{K_{2n+4}^{-\alpha+1}}{K_{2n+2}^{-\alpha+1} - K_{2n+3}^{-\alpha+1}}.$$

Proof By Proposition 6.1, (7) and (8), we have

$$|\Phi_a - \Phi^{T'_n}_{S'_n,P} - \Phi^{T'_{n+1}}_{S'_{n+1},P}| \leq \frac{2(a/P)^{\frac{1}{1+\alpha}} + P^{\frac{-1}{1+\alpha}} T_{a,n-1}}{D} + \frac{2S_{a,n+2}^{-\alpha+1}}{P^{\frac{2}{1+\alpha}}(\alpha-1)}$$

if $P^{1/(\alpha(1+\alpha))} S_{a,n+2} \geq (2R)^{1/\alpha}$. Since we have

$$\left(\frac{a}{P}\right)^{\frac{1}{1+\alpha}} = \frac{1}{\theta(K_{2n+1}-K_{2n})},$$

$$P^{\frac{-1}{1+\alpha}} T_{a,n-1} = \frac{K_{2n-1}}{\theta(K_{2n+1}-K_{2n})},$$

$$\frac{S_{a,n+2}^{-\alpha+1}}{P^{\frac{2}{1+\alpha}}} = \frac{K_{2n+4}^{-\alpha+1}}{K_{2n+2}^{-\alpha+1} - K_{2n+3}^{-\alpha+1}},$$

we have the assertion. □

Here, the assumption $P^{1/(\alpha(1+\alpha))} S_{a,n+2} \geq (2R)^{1/\alpha}$ can be replaced by

$$\left(\frac{K_{2n+4}}{K_{2n+2}}\right)^\alpha S_{a,n+1}^\alpha \sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}} \geq 2R.$$

We can apply Propositions 8.3 and 8.7 to $\Phi^{T'_n}_{S'_n,P}$ and $\Phi^{T'_{n+1}}_{S'_{n+1},P}$. If we put

$$S = S_{a,n+1}, \quad T = T_{a,n+1}, \quad \theta = 1, \quad P^{\frac{1}{1+\alpha}} = \sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}}$$

in Proposition 8.3, then we have

$$\left|\Phi^{T'_{n+1}}_{S'_{n+1},P} - \frac{1}{\alpha-1}\right| \leq \frac{CR}{S_{a,n+1}^\alpha \sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}}}$$

for any $\zeta \in K(R,D)$ if $S_{a,n+1}^{\alpha}\sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}} \geq 2R$. If we put

$$S = S_{a,n}, \quad T = T_{a,n}, \quad P^{\frac{1}{1+\alpha}} = \theta(T_{a,n} - S_{a,n})$$

in Proposition 8.7, then we have

$$\left|\Phi_{S_n',P}^{T_n'} - \frac{1}{\theta|\zeta|}\right| \leq \frac{2}{\theta D},$$
$$\left|\Phi_{S_n',P}^{T_n'} - \frac{1}{\theta|\zeta|}\right| \leq \frac{1+\theta T_{a,n}^{\alpha}(T_{a,n} - S_{a,n})}{D^3} T_{a,n}^{\alpha}(T_{a,n} - S_{a,n}).$$

Now, we put $\Phi = \Phi_a$, $\Phi_\infty = 1/(\alpha-1) + 1/(\theta|\zeta|)$. Combining the above arguments and Proposition 8.11, we can describe ε, C_1 in **(A3)** explicitly with $m=3$. Moreover, by Propositions 8.3, 8.7, 8.4 and 8.8, we obtain C_0, C_1 in **(A5)** and **(A6)**, and $\kappa = 1$. Fix a constant $A>0$, suppose that

$$A^{-1} \leq \theta \leq A, \quad S_{a,n+1}^{\alpha}\sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}} \geq 2R,$$

and take P as above. Then we can take these constants in **(A3)**–**(A6)** being only depending on α, A and R, if $\varepsilon_{a,n}$, $S_{a,n+1}^{-\alpha}(S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1})^{-1/2}$ and $T_{a,n}^{\alpha}(T_{a,n} - S_{a,n})$ are sufficiently small. Therefore, we obtain the following result.

Theorem 8.12 *Let (X, g_Λ) be as in Section 6, take a subsequence*

$$\{K_{n_0} < K_{n_1} < K_{n_2} < \cdots\} \subset \{K_0 < K_1 < K_2 < \cdots\},$$

and suppose

$$\lim_{i\to\infty}\left(\frac{K_{2n_i-1}}{K_{2n_i+1} - K_{2n_i}} + \frac{K_{2n_i+4}^{-\alpha+1}}{K_{2n_i+2}^{-\alpha+1} - K_{2n_i+3}^{-\alpha+1}}\right) = 0. \tag{16}$$

If a sequence $\{a_i\}_i \subset \mathbb{R}^+$ satisfies

$$\lim_{i\to\infty}\frac{\sqrt{S_{a_i,n_i+1}^{-\alpha+1} - T_{a_i,n_i+1}^{-\alpha+1}}}{T_{a_i,n_i} - S_{a_i,n_i}} = \theta > 0,$$
$$\lim_{i\to\infty} S_{a_i,n_i+1}^{-\alpha}(S_{a_i,n_i+1}^{-\alpha+1} - T_{a_i,n_i+1}^{-\alpha+1})^{\frac{-1}{2}} = \lim_{i\to\infty} T_{a_i,n_i}^{\alpha}(T_{a_i,n_i} - S_{a_i,n_i}) = 0,$$

then $\{(X, a_i g_\Lambda, p)\}_n \xrightarrow{\mathrm{GH}} \big(\mathbb{R}^3, \big(1/(\alpha-1) + 1/(\theta|\zeta|)\big)h_0, 0\big)$.

Next we estimate $\Phi_a - 1/(\alpha-1)$ in the same situation, as $\theta \to \infty$. We have

$$\left|\Phi_a - \Phi^{T'_{n+1}}_{S'_{n+1},P}\right| \le \frac{2(a/P)^{\frac{1}{1+\alpha}} + P^{\frac{-1}{1+\alpha}} T_{a,n-1} + P^{\frac{-1}{1+\alpha}}(T_{a,n} - S_{a,n})}{D} + \frac{2S^{-\alpha+1}_{a,n+2}}{P^{\frac{2}{1+\alpha}}(\alpha-1)}$$
$$\le \frac{C_\alpha}{D}\left(\frac{1+K_{2n-1}}{\theta(K_{2n+1}-K_{2n})} + \frac{1}{\theta} + \frac{K^{-\alpha+1}_{2n+4}}{K^{-\alpha+1}_{2n+2} - K^{-\alpha+1}_{2n+3}}\right).$$

Applying Propositions 8.3 and 8.4 with $\theta = 1$ and (5), we have

$$\left|\Phi_a - \frac{1}{\alpha-1}\right|$$
$$\le \frac{C_\alpha}{D}\left(\frac{1+K_{2n-1}}{\theta(K_{2n+1}-K_{2n})} + \frac{1}{\theta} + \frac{K^{-\alpha+1}_{2n+4}}{K^{-\alpha+1}_{2n+2} - K^{-\alpha+1}_{2n+3}} + \frac{R}{S^{\alpha}_{a,n+1}\sqrt{S^{-\alpha+1}_{a,n+1} - T^{-\alpha+1}_{a,n+1}}}\right),$$

$$\Phi_a \ge \left(A^{T'_{n+1}}_{S'_{n+1},P} - \frac{2}{\theta(K_{2n+1}-K_{2n})}\right)\min\left\{\frac{1}{|\zeta|}, 1\right\}$$
$$\ge \left(\frac{1}{2(\alpha-1)} - \frac{2}{\theta(K_{2n+1}-K_{2n})}\right)\min\left\{\frac{1}{|\zeta|}, 1\right\}$$

if $D \le 1$, $R \ge 1$ and $|\zeta| \le R$. Therefore, we can take C_0, C_1, κ and m in **(A3)–(A6)** depending only on α and R if $\varepsilon \to 0$, where $\Phi = \Phi_a$ and $\Phi_\infty = 1/(\alpha-1)$. Hence we have the following theorem.

Theorem 8.13 *Let (X, g_Λ) be as in Section 6 and suppose $\{K_{n_i}\}_i$ satisfies (16). If a sequence $\{a_i\}_i \subset \mathbb{R}^+$ satisfies*

$$\lim_{i\to\infty} \frac{\sqrt{S^{-\alpha+1}_{a_i,n_i+1} - T^{-\alpha+1}_{a_i,n_i+1}}}{T_{a_i,n_i} - S_{a_i,n_i}} = \infty, \qquad \lim_{i\to\infty} S^{-\alpha}_{a_i,n_i+1}(S^{-\alpha+1}_{a_i,n_i+1} - T^{-\alpha+1}_{a_i,n_i+1})^{\frac{-1}{2}} = 0,$$

then $\{(X, a_i g_\Lambda, p)\}_n \xrightarrow{\mathrm{GH}} (\mathbb{R}^3, h_0, 0)$.

By the similar argument, we have the following.

Theorem 8.14 *Let (X, g_Λ) be as in Section 6 and suppose $\{K_{n_i}\}_i$ satisfies (16). If a sequence $\{a_i\}_i \subset \mathbb{R}^+$ satisfies*

$$\lim_{i\to\infty} \frac{\sqrt{S^{-\alpha+1}_{a_i,n_i+1} - T^{-\alpha+1}_{a_i,n_i+1}}}{T_{a_i,n_i} - S_{a_i,n_i}} = 0,$$
$$\lim_{i\to\infty} S^{-\alpha}_{a_i,n_i+1}(S^{-\alpha+1}_{a_i,n_i+1} - T^{-\alpha+1}_{a_i,n_i+1})^{\frac{-1}{2}} = \lim_{i\to\infty} T^{\alpha}_{a_i,n_i}(T_{a_i,n_i} - S_{a_i,n_i}) = 0,$$

then $\{(X, a_i g_\Lambda, p)\}_n \xrightarrow{\mathrm{GH}} (\mathbb{R}^3, (1/|\zeta|)h_0, 0)$.

Proof Put

$$P^{\frac{1}{1+\alpha}} := (T_{a,n} - S_{a,n}) = \theta \sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}},$$

$$S'_l = P^{\frac{-1}{1+\alpha}} S_{a,l}, \quad T'_l = P^{\frac{-1}{1+\alpha}} T_{a,l}.$$

An argument similar to (7) gives

$$\Phi_{S'_{n+1},P}^{T'_{n+1}}(\zeta) \le \frac{2((S'_{n+1})^{-\alpha+1} - (T'_{n+1})^{-\alpha+1})}{P(\alpha-1)}$$

if $P(S'_{n+1})^{\alpha} \ge 2R$, which is equivalent to $\theta S_{a,n+1}^{\alpha}\sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}} \ge 2R$. Then an argument similar to Proposition 8.11 gives

$$\left|\Phi_a(\zeta) - \Phi_{S'_n,P}^{T'_n}(\zeta)\right| \le \frac{C_\alpha \varepsilon_{a,n}}{D} + \frac{2}{(\alpha-1)\theta}$$

for any $\zeta \in K(R,D)$, where $\varepsilon_{a,n}$ is the constant defined by

$$\varepsilon_{a,n} = \frac{1+K_{2n-1}}{K_{2n+1} - K_{2n}} + \frac{K_{2n+4}^{-\alpha+1}}{\theta^2(K_{2n+2}^{-\alpha+1} - K_{2n+3}^{-\alpha+1})}.$$

Moreover, Proposition 8.7 with $\theta = 1$ gives

$$\left|\Phi_{S'_n,P}^{T'_n}(\zeta) - \frac{1}{|\zeta|}\right| \le \frac{2}{D},$$

$$\left|\Phi_{S'_n,P}^{T'_n}(\zeta) - \frac{1}{|\zeta|}\right| \le \frac{1+T_{a,n}^{\alpha}(T_{a,n}-S_{a,n})}{D^3} T_{a,n}^{\alpha}(T_{a,n}-S_{a,n}).$$

Then we can see $|\Phi_a - 1/|\zeta|| \le \varepsilon/D^3$ for some $\varepsilon > 0$ if $D \le 1$ and $\zeta \in K(R,D)$. Here, ε goes to 0 as

$$\varepsilon_{a,n} \to 0, \quad \theta \to \infty, \quad S_{a,n+1}^{\alpha}\sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}} \to \infty \quad \text{and} \quad T_{a,n}^{\alpha}(T_{a,n} - S_{a,n}) \to 0.$$

Since one can take C_0, C_1, m, κ in **(A3)**–**(A6)** depending only on α if ε is sufficiently small, by Proposition 8.8 with $\theta = 1$ and (5), we obtain the result. □

8.4 Example (1)

Let Λ be as in Section 6. Moreover, we take an increasing sequence $\{K_n\}_n$ such that

$$\lim_{n\to\infty} \frac{K_n}{K_{n-1}} = \infty.$$

In this situation, we observe which pointed metric spaces can be contained in $\mathcal{T}(X, g_\Lambda)$ and prove Theorem 1.2.

Take $S>0$ and put $a_i := K_{2i}^{-1-\alpha}S^{1+\alpha}$. Then we have $a_i^{1/(1+\alpha)}K_{2i}=S$ and $\lim_i a_i^{1/(1+\alpha)}K_{2i+1}=\infty$. Hence Theorem 8.2 implies $(X,a_i g_\Lambda,p)\xrightarrow{\mathrm{GH}}(\mathbb{R}^3,d_S^\infty,0)$. Similarly, if we take $a_i := K_{2i+1}^{-1-\alpha}T^{1+\alpha}$ for $T>0$, then we obtain $(\mathbb{R}^3,d_0^T,0)$ as the pointed Gromov–Hausdorff limit.

Next we fix $\theta>0$ and put $a_i=\theta^{-1}K_{2i+1}^{-2}K_{2i+2}^{-\alpha+1}$. Then one can check that the assumptions of Theorem 8.12 are satisfied; hence one obtains $\big(\mathbb{R}^3,\ \big(1/(\alpha-1)+1/(\theta|\zeta|)\big)h_0,\ 0\big)$ as the pointed Gromov–Hausdorff limit. By taking the limit $\theta\to 0$ or $\theta\to\infty$, we obtain $(\mathbb{R}^3,h_0,0)$ or $(\mathbb{R}^3,(1/|\zeta|)h_0,0)$ as the pointed Gromov–Hausdorff limit. In fact, we obtain the next result.

Theorem 8.15 *Let $\Lambda,\{K_n\}_n$ satisfy $\lim_{n\to\infty}K_n/K_{n-1}=\infty$. Then $\mathcal{T}(X,g_\Lambda)$ is equal to the closure of*

$$\{(\mathbb{R}^3,sd_1^\infty,0):s>0\}\cup\{(\mathbb{R}^3,sd_0^1,0):s>0\}\cup\Big\{\Big(\mathbb{R}^3,s\Big(1+\frac{1}{|\zeta|}\Big)h_0,0\Big):s>0\Big\}$$

with respect to the Gromov–Hausdorff topology. Moreover, we have

$$\begin{aligned}\lim_{s\to\infty}(\mathbb{R}^3,sd_1^\infty,0)&=\lim_{s\to 0}\Big(\mathbb{R}^3,s\Big(1+\frac{1}{|\zeta|}\Big)h_0,0\Big)=(\mathbb{R}^3,h_0,0),\\ \lim_{s\to 0}(\mathbb{R}^3,sd_0^1,0)&=\lim_{s\to\infty}\Big(\mathbb{R}^3,s\Big(1+\frac{1}{|\zeta|}\Big),0\Big)=\Big(\mathbb{R}^3,\frac{1}{|\zeta|}h_0,0\Big),\\ \lim_{s\to 0}(\mathbb{R}^3,sd_1^\infty,0)&=\lim_{s\to\infty}(\mathbb{R}^3,sd_0^1,0)=(\mathbb{R}^3,d_0^\infty,0).\end{aligned}$$

Proof We have already shown that the pointed metric spaces in the above list are contained in $\mathcal{T}(X,g_\Lambda)$. Accordingly, what we have to show is that any other pointed metric spaces may not arise as the tangent cone at infinity of (X,g_Λ).

Suppose that a sequence $\{a_i\}_i\subset\mathbb{R}^+$ is given such that $(X,a_i g_\Lambda,p)\xrightarrow{\mathrm{GH}}(Y,d,q)$ as $i\to\infty$. It suffices to show that (Y,d,q) is one of the metric spaces in the list.

First of all, we may assume that for any large $M>0$, there exists $i(M)$ such that

$$\big\{a_i^{\frac{1}{1+\alpha}}K_n\in\mathbb{R}^+:n\in\mathbb{N}\big\}\cap[M^{-1},M]$$

is empty for any $i\ge i(M)$. If not, there is an $M>0$ and a map $i\mapsto n_i$ such that $M^{-1}\le a_i^{1/(1+\alpha)}K_{n_i}\le M$ holds for infinitely many i. Then by taking a subsequence $\{a_{i_j}\}\subset\{a_i\}_i$, we may suppose that $M^{-1}\le a_{i_j}^{1/(1+\alpha)}K_{2n_{i_j}}\le M$ holds for any j or that $M^{-1}\le a_{i_j}^{1/(1+\alpha)}K_{2n_{i_j}+1}\le M$ holds for any j. If the former case holds, then by replacing with a subsequence, we may suppose

$$\lim_{i\to\infty} a_i^{\frac{1}{1+\alpha}} K_{2n_i} = S \in [M^{-1}, M],$$
$$\lim_{i\to\infty} a_i^{\frac{1}{1+\alpha}} K_{2n_i+1} = S \lim_{i\to\infty} \frac{K_{2n_i+1}}{K_{2n_i}} = \infty,$$

and we can apply Theorem 8.2; hence we obtain $(Y,d,q)=(\mathbb{R}^3, d_S^{\infty}, 0)$. If the latter case holds, then we have $(Y,d,q)=(\mathbb{R}^3, d_0^T, 0)$ for some $T>0$.

We may suppose that there exists $l_i \in \mathbb{N}$ for each i such that $\lim_{i\to\infty} a_i^{1/(1+\alpha)} K_{l_i} = 0$ and $\lim_{i\to\infty} a_i^{1/(1+\alpha)} K_{l_i+1} = \infty$ hold. If $\{i\in\mathbb{N} : l_i \text{ is even}\}$ is an infinite set, then we can apply Theorem 8.2 again and obtain $(Y,d,q)=(\mathbb{R}^3, d_0^{\infty}, 0)$. Therefore, replacing with a subsequence, we may suppose

$$\lim_{i\to\infty} a_i^{\frac{1}{1+\alpha}} K_{2n_i+1} = 0, \qquad \lim_{i\to\infty} a_i^{\frac{1}{1+\alpha}} K_{2n_i+2} = \infty.$$

Now, we have

$$\sqrt{S_{a,n+1}^{-\alpha+1} - T_{a,n+1}^{-\alpha+1}} \geq \tfrac{1}{2} S_{a,n+1}^{\frac{-\alpha+1}{2}}, \qquad T_{a,n} - S_{a,n} \geq \tfrac{1}{2} T_{a,n}$$

holds for sufficiently large n. Hence if

$$0 < \liminf_{i\to\infty} \frac{S_{a_i,n_i+1}^{\frac{1-\alpha}{2}}}{T_{a_i,n_i}} \leq \limsup_{i\to\infty} \frac{S_{a_i,n_i+1}^{\frac{1-\alpha}{2}}}{T_{a_i,n_i}} < \infty,$$

then Theorem 8.12 can be applied to this situation by taking a subsequence. Then we obtain $(Y,d,q)=(\mathbb{R}^3, (1+\theta/|\zeta|)h_0, 0)$ for some $\theta>0$. Hence the remaining cases are

$$\lim_{i\to\infty} \frac{S_{a_i,n_i+1}^{\frac{1-\alpha}{2}}}{T_{a_i,n_i}} = 0 \quad \text{or} \quad \lim_{i\to\infty} \frac{S_{a_i,n_i+1}^{\frac{1-\alpha}{2}}}{T_{a_i,n_i}} = \infty.$$

In both of the cases, we can apply Theorems 8.13 or 8.14, and then obtain $(Y,d,q)=(\mathbb{R}^3, h_0, 0)$ or $(\mathbb{R}^3, (1/|\zeta|)h_0, 0)$. ◻

One can also see that there are no nontrivial isometries between two pointed metric spaces appearing in the list of Theorem 8.15. Here, an isometry of pointed metric spaces means a bijective morphism preserving the metrics and the base points.

Obviously, there is no isometry between $(\mathbb{R}^3, h_0, 0)$ and $(\mathbb{R}^3, (1/|\zeta|)h_0, 0)$. In the next section, we will show that $(\mathbb{R}^3, d_0^{\infty}, 0)$ is isometric to neither $(\mathbb{R}^3, h_0, 0)$ nor $(\mathbb{R}^3, (1/|\zeta|)h_0, 0)$.

metric	tangent cone at 0	tangent cone at ∞
d_S^T $(S<T)$	h_0	$\frac{1}{\lvert\zeta\rvert}h_0$
d_S^∞	h_0	d_0^∞
d_0^T	d_0^∞	$\frac{1}{\lvert\zeta\rvert}h_0$
d_0^∞	d_0^∞	d_0^∞
h_0	h_0	h_0
$\frac{1}{\lvert\zeta\rvert}h_0$	$\frac{1}{\lvert\zeta\rvert}h_0$	$\frac{1}{\lvert\zeta\rvert}h_0$
$(1+\frac{\theta}{\lvert\zeta\rvert})h_0$	$\frac{1}{\lvert\zeta\rvert}h_0$	h_0

Table 1: Tangent cones $(0<S,T,\theta<\infty)$

Then Table 1 implies that nontrivial isometries may exist between

$$\begin{aligned} &(\mathbb{R}^3, d_S^\infty, 0) &&\text{and}&& (\mathbb{R}^3, d_{S'}^\infty, 0) &&\text{for}\quad S\neq S',\\ &(\mathbb{R}^3, d_0^T, 0) &&\text{and}&& (\mathbb{R}^3, d_0^{T'}, 0) &&\text{for}\quad T\neq T',\\ &\Big(\mathbb{R}^3, \Big(1+\frac{\theta}{|\zeta|}\Big)h_0, 0\Big) &&\text{and}&& \Big(\mathbb{R}^3, \Big(1+\frac{\theta'}{|\zeta|}\Big)h_0, 0\Big) &&\text{for}\quad \theta\neq\theta'. \end{aligned}$$

Suppose $(\mathbb{R}^3, d_S^\infty, 0)$ is isometric to $(\mathbb{R}^3, d_{S'}^\infty, 0)$ for some $S\neq S'$. Then the topological space

$$\{(\mathbb{R}^3, d_S^\infty, 0) : S\in\mathbb{R}^+\}$$

with respect to pointed Gromov–Hausdorff topology is homeomorphic to S^1 or 1–point; hence it is compact. Then its closure is itself; therefore $(\mathbb{R}^3, h_0, 0)$ is isometric to some $(\mathbb{R}^3, d_S^\infty, 0)$, which is a contradiction by Table 1. Similarly, we can show that there are no isometries between $(\mathbb{R}^3, d_0^T, 0)$ and $(\mathbb{R}^3, d_0^{T'}, 0)$, or between $(\mathbb{R}^3, (1+\theta/|\zeta|)h_0, 0)$ and $(\mathbb{R}^3, (1+\theta'/|\zeta|)h_0, 0)$.

8.5 Example (2)

Next we suppose that $\{K_n\}_n$ satisfies

$$\lim_{n\to\infty}\frac{K_{2n}}{K_{2n-1}}=\infty,\qquad \frac{K_{2n+1}}{K_{2n}}=\beta>1.$$

Take $S>0$ and put $a_n := K_{2n}^{-1-\alpha}S^{1+\alpha}$. Then we have $a_n^{1/(1+\alpha)}K_{2n}=S$ and $a_n^{1/(1+\alpha)}K_{2n+1}=\beta S$. Hence Theorem 8.2 implies that $(X, a_n g_\Lambda, p)\xrightarrow{\mathrm{GH}}(\mathbb{R}^3, d_S^{\beta S})$. By arguing similarly to the proof of Theorem 8.15, we obtain the following.

Theorem 8.16 *Let* $\Lambda, \{K_n\}_n$ *satisfy*

$$\lim_{n\to\infty} \frac{K_{2n}}{K_{2n-1}} = \infty, \qquad \lim_{n\to\infty} \frac{K_{2n+1}}{K_{2n}} = \beta > 1.$$

Then $\mathcal{T}(X, g_\Lambda)$ *is equal to the closure of*

$$\{(\mathbb{R}^3, sd_1^\beta, 0) : s > 0\} \cup \left\{\left(\mathbb{R}^3, s\left(1 + \frac{1}{|\zeta|}\right)h_0, 0\right) : s > 0\right\}$$

with respect to the Gromov–Hausdorff topology. Moreover, we have

$$\lim_{s\to\infty} (\mathbb{R}^3, sd_1^\beta, 0) = \lim_{s\to 0}\left(\mathbb{R}^3, s\left(1 + \frac{1}{|\zeta|}\right)h_0, 0\right) = (\mathbb{R}^3, h_0, 0),$$
$$\lim_{s\to 0} (\mathbb{R}^3, sd_1^\beta, 0) = \lim_{s\to\infty}\left(\mathbb{R}^3, s\left(1 + \frac{1}{|\zeta|}\right), 0\right) = \left(\mathbb{R}^3, \frac{1}{|\zeta|}h_0, 0\right).$$

By a similar argument to Section 8.4, we can see that $(\mathbb{R}^3, d_S^{\beta S}, 0)$ is isometric to neither $(\mathbb{R}^3, h_0, 0)$, $(\mathbb{R}^3, (1/|\zeta|)h_0, 0)$ nor $(\mathbb{R}^3, d_{S'}^{\beta S'}, 0)$ for $S' \neq S$.

8.6 Example (3)

For $I \subset \mathbb{R}^+$, denote by d_I the metric on $\mathbb{R}^3$ induced by

$$\int_{x\in I} \frac{dx}{|\zeta - (x^\alpha, 0, 0)|} \cdot h_0.$$

Denote by $\mathcal{B}_+(\mathbb{R}^+)$ the set consisting of all Borel subsets of $\mathbb{R}^+$ of nonzero Lebesgue measure. In this subsection, we show the next theorem.

Theorem 8.17 *There is a sequence* $\{K_n\}_n$ *such that* $\mathcal{T}(X, g_\Lambda)$ *contains*

$$\{(\mathbb{R}^3, d_I, 0) : I \in \mathcal{B}_+(\mathbb{R}^+)\}/\text{isometry}.$$

Proof Put

$$\mathcal{O}_0 := \{I \subset \mathbb{R}^+ : I \text{ is nonempty and open}\},$$
$$\mathcal{O}_1 := \left\{\bigcup_{i=1}^k (S_l, T_l) \subset \mathbb{R}^+ : \begin{array}{l} S_l, T_l \in \mathbb{Q},\ 1 \le k < \infty, \\ 0 < S_l < T_l < S_{l+1} < \infty \end{array}\right\}.$$

Then one can see $\mathcal{O}_1 \subset \mathcal{O}_0 \subset \mathcal{B}_+(\mathbb{R}^+)$. Since $\mathcal{O}_1$ is countable, we can label the open sets in $\mathcal{O}_1$ as follows:

$$\mathcal{O}_1 = \{I_1, I_2, I_3, \dots\}, \qquad I_m = \bigcup_{l=1}^{k_m} (S_{m,l}, T_{m,l}).$$

Now we fix a bijection $F\colon \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ and write $F(q) = (i(q), m(q))$. Define $L_q > 0$ inductively by

$$L_{q+1} := 2^{i(q)+i(q+1)} L_q \cdot \frac{T_{m(q),k_{m(q)}}}{S_{m(q),1}}, \quad L_0 := 1.$$

Then we can define $0 < K_0 < K_1 < \cdots$ such that

$$\{K_0 < K_1 < \cdots\} = \left\{ L_q \frac{S_{m(q),l}}{S_{m(q),1}}, L_q \frac{T_{m(q),l}}{S_{m(q),1}} : 1 \le l \le k_{m(q)},\ q = 0, 1, \dots \right\}.$$

First we show that $(\mathbb{R}^3, d_{I_m}, 0) \in \mathcal{T}(X, g_\Lambda)$ for every $I_m \in \mathcal{O}_1$. Fix m. For any $i \in \mathbb{N}$, we can take a unique q such that $i(q) = i$ and $m(q) = m$. Put $a_i^{1/(1+\alpha)} := L_q^{-1} S_{m,1}$; then we have

$$a_i^{\frac{1}{1+\alpha}} L_q \frac{S_{m,l}}{S_{m,1}} = S_{m,l}, \quad a_i^{\frac{1}{1+\alpha}} L_q \frac{T_{m,l}}{S_{m,1}} = T_{m,l}.$$

Note that $L_{q+1} \ge 2^{i(q)+i(q+1)} L_q$ implies $L_q \to \infty$ as $i \to \infty$, hence $a_i \to 0$ as $i \to \infty$. Here, we put $\Phi = \Phi_{a_i}$ and $\Phi_\infty = \sum_{l=1}^{k_m} \Phi_{S_{m,l}}^{T_{m,l}}$. By applying Proposition 6.1 and (4)–(8) with $P = 1$, the constants appearing in **(A3)**–**(A6)** are given by

$$\varepsilon = 2a_i^{\frac{1}{1+\alpha}} + 2^{-i} S_{m,1} + \frac{2^{1-(\alpha-1)i} T_{m,k_m}^{-\alpha+1}}{\alpha - 1}, \quad C_0 = \frac{1}{2}\sum_{l=1}^{k_m} A_{S_{m,l}}^{T_{m,l}},$$

$$C_1 = \frac{\alpha 2^{\frac{1}{\alpha}}}{\alpha - 1}, \qquad m = 1, \qquad \kappa = \frac{1}{\alpha}$$

if we suppose ε is sufficiently small. One can see $\varepsilon \to 0$ as $i \to \infty$, so we obtain $\{(X, a_i g_\Lambda, p)\}_i \xrightarrow{\mathrm{GH}} (\mathbb{R}^3, d_{I_m}, 0)$.

Next we show that $(\mathbb{R}^3, d_I, 0) \in \mathcal{T}(X, g_\Lambda)$ for any $I \in \mathcal{O}_0$. To show it, we apply Vitali's covering theorem. Fix $I \in \mathcal{O}_0$ and put $\mathcal{I} := \{(a,b) \in \mathcal{O}_0 : [a,b] \subset I\}$. Then $\mathcal{I}$ is a Vitali cover of I; hence there exists $\{J_n\}_{n\in\mathbb{N}} \subset \mathcal{I}$ such that

$$J_n \ne J_{n'} \text{ if } n \ne n', \quad m\Big(I \setminus \bigsqcup_{n\in\mathbb{N}} J_n\Big) = 0,$$

where m is the Lebesgue measure. Put $\hat{J}_n := \bigsqcup_{k=1}^n J_k$. Since $\hat{J}_n \in \mathcal{O}_1$ holds, we have $(\mathbb{R}^3, d_{\hat{J}_n}, 0) \in \mathcal{T}(X, g_\Lambda)$. If we put

$$\Phi_J(\zeta) := \int_{x\in J} \frac{dx}{|\zeta - (x^\alpha, 0, 0)|},$$

then we can see

$$|\Phi_{\hat{J}_n}(\zeta) - \Phi_I(\zeta)| \le \frac{m(I \setminus \hat{J}_n)}{D} \to 0 \quad \text{as } n \to \infty,$$

and we can take the constants in **(A3)**–**(A6)** independent of n by using Proposition 6.2. Therefore, we obtain $\{(\mathbb{R}^3, d_{\hat{J}_n}, 0)\}_n \xrightarrow{\mathrm{GH}} (\mathbb{R}^3, d_I, 0)$.

Finally, let $I \in \mathcal{B}_+(\mathbb{R}^+)$. Since the Lebesgue measure is the Radon measure, there exists $U_n \subset \mathcal{O}_1$ for any n such that $I \subset U$ and $m(U) \le m(I) + 1/n$. Then we have $|\Phi_I(\zeta) - \Phi_{U_n}(\zeta)| \le 1/(nD)$, and thus $\{(\mathbb{R}^3, d_{U_n}, 0)\}_n \xrightarrow{\mathrm{GH}} (\mathbb{R}^3, d_I, 0)$ by a similar argument. Here, the positivity of $m(I)$ is necessary since, by (4), C_0 in **(A5)** is given by

$$\int_I \frac{dx}{1 + x^\alpha} > 0. \qquad \square$$

By Theorem 8.17, we can see that $(\mathbb{R}^3, h_0, 0)$ and $(\mathbb{R}^3, (1/|\zeta|)h_0, 0)$ are also contained in $\mathcal{B}_+(\mathbb{R}^+)$. The author does not know whether any other metric spaces may appear as the tangent cone at infinity of (X, g_Λ) or not.

9 On the geometry of the limit spaces

In this section, we study the geometry of $(\mathbb{R}^3, d_0^\infty)$ and conclude that there is no isometry between $(\mathbb{R}^3, d_0^\infty)$ and $(\mathbb{R}^3, h_0)$, nor between $(\mathbb{R}^3, d_0^\infty)$ and $(\mathbb{R}^3, (1/|\zeta|)h_0)$.

Proposition 9.1 $(\mathbb{R}^3, (1/|\zeta|)h_0)$ *is the Riemannian cone* $S^2 \times \mathbb{R}^+$*, where the Riemannian metric on* S^2 *is the homogeneous one whose area is equal to* π*.*

Proof Put $\zeta = (\zeta_1, \zeta_2, \zeta_3) \neq 0$ and $r = \sqrt{\zeta_1^2 + \zeta_2^2 + \zeta_3^2}$, and let g_{S^2} be the standard Riemannian metric on S^2 with constant curvature and volume 4π. Then by putting $R := 2\sqrt{r}$, we have

$$\frac{1}{|\zeta|} h_0 = \frac{1}{r}\big((dr)^2 + r^2 g_{S^2}\big) = (dR)^2 + R^2 \cdot \frac{g_{S^2}}{4}. \qquad \square$$

Next we review the notion of polar spaces, introduced by Cheeger and Colding in [5], and then we show that the metric space $(\mathbb{R}^3, d_0^\infty)$ is never a polar space.

Let Y be a metric space, and suppose that there is a tangent cone Y_y at $y \in Y$. Then we can consider tangent cones at any points in Y_y. The tangent cones obtained by repeating this process are called *iterated tangent cones* of Y. A point x in a length-space X is called a *pole* if there is a ray $\gamma\colon [0, \infty) \to X$ and $t \ge 0$ for any $\underline{x} \neq x$ such that $\gamma(0) = x$ and $\gamma(t) = \underline{x}$. Here, the ray $\gamma\colon [0, \infty) \to X$ is a continuous curve such that the length of $\gamma|_{[t_0, t_1]}$ is equal to $|\gamma(t_0)\gamma(t_1)|$.

Definition 9.2 [5] The metric space Y is called a *polar space* if all of the base points of the iterated tangent cones of Y are poles.

For example, let $C(X)$ be a metric cone of a metric space X. Then every γ defined by $\gamma(t) := (x,t) \in X \times \mathbb{R}^+ = C(X)$ is a ray; hence the base points of any metric cones are poles. Now, since $(\mathbb{R}^3, (1/|\zeta|)h_0)$ is a Riemannian cone of a smooth compact Riemannian manifold, then all of the iterated tangent cones are $(\mathbb{R}^3, (1/|\zeta|)h_0)$ itself or $(\mathbb{R}^3, h_0)$. Consequently, we can conclude that $(\mathbb{R}^3, (1/|\zeta|)h_0)$ is polar. Obviously, $(\mathbb{R}^3, h_0)$ is also polar. We can also see in the similar way that $(\mathbb{R}^3, (1+\theta/|\zeta|)h_0)$ is polar. On the other hand, we can show the next proposition.

Proposition 9.3 *The origin $0 \in \mathbb{R}^3$ is not a pole of the metric space $(\mathbb{R}^3, d_0^\infty)$. In particular, $(\mathbb{R}^3, d_0^\infty)$ is neither a polar space nor a metric cone of any metric spaces.*

Proof First of all we show that $0 \in \mathbb{R}^3$ is not a pole with respect to d_0^∞. Put $p := (1,0,0) \in \mathbb{R}^3$, and suppose that there is a ray $\gamma\colon [0,\infty) \to \mathbb{R}^3$ such that $\gamma(0) = 0$ and $\gamma(t_0) = p$ for some $t_0 > 0$. Then we have

$$d_0^\infty(\gamma(s_0), \gamma(s_1)) = \int_{s_0}^{s_1} \sqrt{\Phi_0^\infty(\gamma(t))}\,|\gamma'(t)|\,dt$$

for any $0 \le s_0 < s_1$. For $\delta > 0$, let

$$A_\delta := \{t \in \mathbb{R} : |\gamma_{\mathbb{C}}(t)| \ge \delta\}.$$

Then there is a sufficiently small δ such that $A_\delta \cap (0,t_0) \ne \varnothing$ and $A_\delta \cap (t_0,\infty) \ne \varnothing$. This is because the length of $\gamma|_I$ becomes infinity for any small interval $I \subset \mathbb{R}$ if not. Since A_δ is closed and does not contain t_0, we can take a connected component (a_0,a_1) of $\mathbb{R}\backslash A_\delta$ containing t_0. Then we can see that $|\gamma_{\mathbb{C}}(a_0)| = |\gamma_{\mathbb{C}}(a_1)| = \delta$ and $|\gamma_{\mathbb{C}}(t)| < \delta$ for any $t \in (a_0,a_1)$. Now define $\tilde{\gamma}\colon [0,a_1] \to X$ by

$$\tilde{\gamma}(t) := \begin{cases} (\gamma_{\mathbb{R}}(t), e^{i\theta}\gamma_{\mathbb{C}}(t)), & 0 \le t \le a_0, \\ e^{i\theta} P_{\gamma|_{[a_0,a_1]}}(t), & a_0 \le t \le a_1, \end{cases}$$

where θ is defined by $e^{i\theta}\gamma_{\mathbb{C}}(a_0) = \gamma_{\mathbb{C}}(a_1)$. Recall that $P_{\gamma|_{[a_0,a_1]}}$ is already defined in Lemma 7.4. Then by applying Lemma 7.4, we can see that the length of $\tilde{\gamma}$ is strictly less than the length of $\gamma|_{[0,a_1]}$; therefore, γ is not a ray, which is a contradiction. Hence $0 \in \mathbb{R}^3$ is not a pole.

Now we can check that the $\mathbb{R}^+$–action on $\mathbb{R}^3$ defined by scalar multiplication is homothetic with respect to d_0^∞; thus the tangent cone of $(\mathbb{R}^3, d_0^\infty)$ at 0 is itself. Consequently, $(\mathbb{R}^3, d_0^\infty)$ is not a polar space.

Suppose that $(\mathbb{R}^3, d_0^\infty)$ is the metric cone of some metric space X; then the origin 0 is nothing but the base point of the metric cone. Since the base point of the metric cone is always a pole, we have a contradiction. □

Corollary 9.4 *There is no isometry between* $(\mathbb{R}^3, d_0^\infty)$ *and* $(\mathbb{R}^3, h_0)$*, nor between* $(\mathbb{R}^3, d_0^\infty)$ *and* $(\mathbb{R}^3, (1/|\zeta|)h_0)$*.*

References

[1] **M T Anderson**, **P B Kronheimer**, **C LeBrun**, *Complete Ricci-flat Kähler manifolds of infinite topological type*, Comm. Math. Phys. 125 (1989) 637–642 MR

[2] **V N Berestovskii**, **L Guijarro**, *A metric characterization of Riemannian submersions*, Ann. Global Anal. Geom. 18 (2000) 577–588 MR

[3] **V N Berestovskiĭ**, *"Submetries" of three-dimensional forms of nonnegative curvature*, Sibirsk. Mat. Zh. 28 (1987) 44–56, 224 MR In Russian; translated in Siberian Math. J. 28 (1987) 552–562

[4] **J Cheeger**, **T H Colding**, *Lower bounds on Ricci curvature and the almost rigidity of warped products*, Ann. of Math. 144 (1996) 189–237 MR

[5] **J Cheeger**, **T H Colding**, *On the structure of spaces with Ricci curvature bounded below, I*, J. Differential Geom. 46 (1997) 406–480 MR

[6] **T H Colding**, **W P Minicozzi, II**, *On uniqueness of tangent cones for Einstein manifolds*, Invent. Math. 196 (2014) 515–588 MR

[7] **T H Colding**, **A Naber**, *Characterization of tangent cones of noncollapsed limits with lower Ricci bounds and applications*, Geom. Funct. Anal. 23 (2013) 134–148 MR

[8] **R Goto**, *On hyper-Kähler manifolds of type* A_∞, Geom. Funct. Anal. 4 (1994) 424–454 MR

[9] **M Gromov**, *Structures métriques pour les variétés riemanniennes*, Textes Mathématiques 1, CEDIC, Paris (1981) MR

[10] **M Gromov**, *Metric structures for Riemannian and non-Riemannian spaces*, Birkhäuser, Boston (2007) MR

[11] **K Hattori**, *The volume growth of hyper-Kähler manifolds of type* A_∞, J. Geom. Anal. 21 (2011) 920–949 MR

[12] **G Perelman**, *A complete Riemannian manifold of positive Ricci curvature with Euclidean volume growth and nonunique asymptotic cone*, from "Comparison geometry" (K Grove, P Petersen, editors), Math. Sci. Res. Inst. Publ. 30, Cambridge Univ. Press (1997) 165–166 MR

Department of Mathematics, Keio University
3-14-1 Hiyoshi, Kohoku, Yokohama 223-8522, Japan

hattori@math.keio.ac.jp

Proposed: Tobias H. Colding
Seconded: Dmitri Burago, Bruce Kleiner

Received: 10 May 2015
Revised: 12 October 2016

Geometry & Topology 21 (2017) 2725–2809

Smooth Kuranishi atlases with isotropy

DUSA MCDUFF
KATRIN WEHRHEIM

Kuranishi structures were introduced in the 1990s by Fukaya and Ono for the purpose of assigning a virtual cycle to moduli spaces of pseudoholomorphic curves that cannot be regularized by geometric methods. Their core idea was to build such a cycle by patching local finite-dimensional reductions, given by smooth sections that are equivariant under a finite isotropy group.

Building on our notions of topological Kuranishi atlases and perturbation constructions in the case of trivial isotropy, we develop a theory of Kuranishi atlases and cobordisms that transparently resolves the challenges posed by nontrivial isotropy. We assign to a cobordism class of weak Kuranishi atlases both a virtual moduli cycle (a cobordism class of weighted branched manifolds) and a virtual fundamental class (a Čech homology class).

53D35, 53D45, 54B15, 57R17, 57R95

1 Introduction

1.1 Overview

This is the third in a series [13; 14] of papers that construct a fundamental class for compact spaces X that are modeled locally by the zero sets of smooth sections $s_i\colon U_i \to E_i$ in finite rank bundles over finite-dimensional manifolds. While these obstruction bundles have fixed index, they may have varying rank, and thus an ambient space $\bigcup U_i/{\sim}$ naively constructed from the ambient manifolds of the local zero sets $s_i^{-1}(0)$ modulo transition data is lacking all topological controls (Hausdorffness, local compactness, in fact existence) that are needed for a perturbative construction $[X] := \bigcup (s_i + \nu_i)^{-1}(0)/{\sim}$ of the fundamental class. Moreover, most interesting cases involve nontrivial isotropy groups that are captured in the local charts as finite symmetry groups Γ_i of the sections s_i, so that X is locally modeled by the quotients $s_i^{-1}(0)/\Gamma_i$.

Pioneered by Fukaya et al [6; 3], this problem has been considered by symplectic topologists since the 1990s as a tool for "counting curves", ie assigning homological information to moduli spaces of pseudoholomorphic curves, such as the Gromov–Witten

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2725

moduli spaces (in which isotropy arises from components that are multiply covered). In the case of trivial isotropy, a comprehensive solution was developed in [13; 14] by introducing notions of Kuranishi atlases, which on the one hand can in practice be constructed from moduli spaces, and on the other hand have sufficient compatibility between the local models for the construction of a virtual fundamental class. This paper extends these techniques to the case of nontrivial isotropy, proving the following result.

Theorem A *Let $\mathcal{K}$ be an oriented, d–dimensional, additive, smooth weak Kuranishi atlas on a compact metrizable space X. Then $\mathcal{K}$ determines*

- *a* ***virtual moduli cycle (VMC)*** *as cobordism class of weighted branched manifolds,*
- *a* ***virtual fundamental class (VFC)*** $[X]^{\mathrm{vir}}_{\mathcal{K}} \in \check{H}_d(X;\mathbb{Q})$ *in Čech homology,*

both of which depend only on the cobordism class of $\mathcal{K}$.

A more precise statement that also applies when $\mathcal{K}$ is a cobordism from an atlas $\mathcal{K}^0$ on X^0 to an atlas $\mathcal{K}^1$ on X^1 is given in Theorem 3.3.5. Notice further that the VMC contains more information than the VFC since cobordism classes of weighted branched manifolds contain more information than just their fundamental class; for example, Pontryagin numbers are invariants of weighted branched cobordism by [11, Remark 4.7].

The guiding idea of a Kuranishi atlas $\mathcal{K}$ is to start with a family of basic charts $(\boldsymbol{K}_i)_{i=1,\dots,N}$, where each basic chart

$$\boldsymbol{K}_i = (U_i, E_i, \Gamma_i, s_i, \psi_i)$$

is a tuple consisting of a domain U_i, an obstruction space E_i, a group Γ_i, a section $s_i\colon U_i \to E_i$, and a footprint map $\psi_i\colon s_i^{-1}(0) \to X$ inducing a homeomorphism from $s_i^{-1}(0)/\Gamma_i$ onto the "footprint", an open subset $F_i \subset X$ such that $(F_i)_{i=1,\dots,N}$ covers X. The compatibility of these charts then involves transition charts $\boldsymbol{K}_I = (U_I, E_I, \Gamma_I, s_I, \psi_I)$ of the same type as the basic charts, but with $I \subset \{1,\dots,N\}$ such that $F_I := \bigcap_{i\in I} F_i \neq \varnothing$. Finally, the basic and transition charts are related by coordinate changes from $\boldsymbol{K}_I$ to $\boldsymbol{K}_J$ whenever $I \subset J$. This gives rise to an "étale-like" category $\boldsymbol{B}_{\mathcal{K}}$ whose space of objects is $\bigsqcup_I U_I$, and whose morphisms are determined by the local group actions and the coordinate changes. The category $\boldsymbol{B}_{\mathcal{K}}$ is not a groupoid since some morphisms (those relating the different charts) are not invertible. On the other hand, its spaces of objects and morphisms are very closely controlled, which enables us to carry out various geometric constructions, in particular the construction of perturbations, very explicitly. The realization $|\mathcal{K}|$ of $\boldsymbol{B}_{\mathcal{K}}$ (the space of objects modulo the equivalence relation generated by the morphisms) is much larger than X, though it does contain a homeomorphic image of X formed from the zero sets of the local sections s_I. As in [14], the class $[X]^{\mathrm{vir}}_{\mathcal{K}}$ is constructed from the zero sets of suitable perturbations $\mathfrak{s}_{\mathcal{K}}+\nu$ of the basic section $\mathfrak{s}_{\mathcal{K}}=(s_I)$ of $\mathcal{K}$.

Even if X is an orbifold so that no obstruction spaces are needed, our formulations are new.[1] Rather than being given by inclusions $U_I \supset U_{IJ} \hookrightarrow U_J$ as in the case with trivial isotropy, our notion of coordinate changes in the presence of isotropy involves equivariant covering maps $\widetilde{\rho}_{IJ}\colon (\widetilde{U}_{IJ},\Gamma_J)\to(U_{IJ},\Gamma_I)\subset(U_I,\Gamma_I)$, where $\widetilde{U}_{IJ}$ is a suitable subset of the domain U_J and $\widetilde{U}_{IJ}\to U_{IJ}$ is a principal Γ_J/Γ_I-bundle. As the following result from [11, Proposition 3.3] shows, every orbifold has a structure of this kind.

Proposition *Every compact orbifold Y has an orbifold atlas $\mathcal{K}$ with trivial obstruction spaces whose associated groupoid $\boldsymbol{G}_{\mathcal{K}}$ is an orbifold structure on Y. Moreover, there is a bijective correspondence between commensurability classes of such Kuranishi atlases and Morita equivalence classes of ep groupoids.*

To apply the above theory to moduli spaces X that arise in geometric examples, one needs to develop methods for constructing Kuranishi atlases on such X. Some parts of this construction were detailed in the 2012 preprint [12], and now appear in [14]. They will be extended in [10] to include multiply covered curves (and hence nontrivial isotropy) as well as nodal curves. Both McDuff [10] and Pardon [16] outline the needed construction for moduli spaces of closed stable maps, though neither approach is sufficient to give the smooth charts whose existence is assumed in the current paper. In [10] we will combine the same setup with an implicit function theorem from polyfold theory (see Hofer, Wysocki and Zehnder [7]) to obtain compatible choices of smooth structures near nodal curves. An alternative approach is to extend the VMC/VFC construction to less smooth sections. In fact, Castellano [2] proves a gluing theorem for Gromov–Witten moduli spaces that allows the construction of stratified smooth Kuranishi atlases with $\mathcal{C}^1$–differentiability across strata, to which our construction applies with minor modifications. He moreover shows that the resulting genus zero Gromov–Witten invariants satisfy the standard axioms.

1.2 Outline of the construction

This paper contains all relevant definitions and a fair amount of review so that it can be read independently of the previous papers in this series. This outline will also be

[1] Our construction was outlined in [9]. In [16], Pardon independently takes a similar approach to handling the isotropy groups.

rather brief since the earlier papers give extensive explanation and justification for our approach:

- The first part of [14] is a general discussion of different approaches to regularizing moduli spaces — eg as VMC/VFC — and explains important analytic background.

- The paper [13] starts with an overview of the topological challenges that need to be addressed in constructing a VMC/VFC, and then proves the basic topological results needed to show that a filtered weak Kuranishi atlas determines a tame Kuranishi atlas $\mathcal{K}$, well-defined up to cobordism, whose realization $|\mathcal{K}|$ is Hausdorff, contains a homeomorphic copy of the moduli space X, and can be equipped with a metric that is compatible with local charts (but generally induces a different topology on $|\mathcal{K}|$).

- The second part of [14] carries out the full construction of the VMC as the zero set of a suitable perturbation of the canonical section $\mathfrak{s}_{\mathcal{K}}$ in the case of trivial isotropy.

We now discuss the main steps in the construction below in more detail, highlighting the new features needed to deal with nontrivial isotropy.

- In order to simplify the abstract discussion, we decided to give a rather narrow definition of a Kuranishi atlas $\mathcal{K}$. Thus the domains of both the basic and transition charts are group quotients (U_I, Γ_I), and the coordinate changes are determined by rather special equivariant covering maps $(\widetilde{U}_{IJ}, \Gamma_J) \to (U_{IJ}, \Gamma_I)$. The basic theory is set up in Section 2.1; see in particular Definition 2.1.4 and Lemma 2.1.5. If there were a need, one could no doubt replace these group quotients by more general étale groupoids and use more general covering maps and obstruction bundles, at the expense of revisiting the construction of perturbations.

- Smooth atlases and coordinate changes are defined in Sections 2.2 and 2.3. Though in general the definitions are similar to those in the case with trivial isotropy, there is an important difference in the notion of coordinate change: when $I \subset J$ this is now given by a covering map from an appropriate submanifold $\widetilde{U}_{IJ}$ of the domain of the higher dimensional domain U_J onto an open subset U_{IJ} of the lower dimensional domain U_I. If the isotropy groups are trivial, this map is a diffeomorphism with inverse equal to the coordinate changes $\phi_{IJ}\colon U_{IJ} \to U_J$ considered in [13; 14]. Another small difference is that we build in the notion of additivity since at least some version of this is needed for the taming construction discussed below. (In some situations, for example when considering products, this formulation is too rigid; for appropriate generalizations see [10].)

- An important feature of our definitions is that the quotients $\underline{U}_I := U_I/\Gamma_I$ fit together to form an *intermediate atlas*, which Lemma 2.3.4 shows to be a filtered topological atlas in the sense of [13]. In particular it has an associated category $\boldsymbol{B}_{\underline{\mathcal{K}}}$ with space of objects the orbifold $\mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} := \bigsqcup_I \underline{U}_I$, and identical realization $|\underline{\mathcal{K}}| = |\mathcal{K}|$.

- One difficulty in constructing a VFC for a given moduli space X is that in practice one cannot usually construct an atlas on X. Instead one constructs a weak atlas, which is like an atlas except that one has less control of the domains of the charts and coordinate changes; cf the various *cocycle conditions* discussed in Definition 2.2.12 and Lemma 2.2.13. But a weak atlas does not even define a category, let alone one whose realization $|\boldsymbol{B}_{\mathcal{K}}| =: |\mathcal{K}|$ has good topological properties. For example, we would like $|\mathcal{K}|$ to be Hausdorff and (in order to make local constructions possible) for the projection $\pi_{\mathcal{K}}\colon U_I \to |\mathcal{K}|$ to be a homeomorphism to its image. Theorem 2.5.3 summarizes the main topological facts about $\mathcal{K}$ that are needed for subsequent constructions. We achieve these via *shrinking* and *taming*. Our definitions were designed so that all the topological constructions of [13], such as the taming, cobordism and reduction constructions, apply to the intermediate atlas $\underline{\mathcal{K}}$ and then lift to $\mathcal{K}$ because the quotient maps $U_I \to \underline{U}_I$ are proper. However, we do need to take some care with the proof of the linearity properties of the projection $\mathrm{pr}\colon |\boldsymbol{E}_{\mathcal{K}}| \to |\boldsymbol{B}_{\mathcal{K}}|$.

- Another important part of Theorem 2.5.3 is the claim that any two tame shrinkings of a weak atlas $\mathcal{K}$ are *concordant*, ie cobordant over $[0,1]\times X$, which is required to show independence of the VMC/VFC from the choice of shrinking. In Section 2.4 we give the precise definition of a cobordism atlas. This is an immediate generalization of the notion of cobordism in [13; 14], and the relevant proofs generalize easily.

- Given a weak atlas, the taming procedure gives us two categories $\boldsymbol{B}_{\mathcal{K}}$ and $\boldsymbol{E}_{\mathcal{K}}$ with a projection functor $\mathrm{pr}\colon \boldsymbol{E}_{\mathcal{K}} \to \boldsymbol{B}_{\mathcal{K}}$ and section functor $\mathfrak{s}_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$. However, even when the isotropy is trivial, the category has too many morphisms (ie the chart domains overlap too much) for us to be able to construct a perturbation $\nu\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$ that is transverse to 0 (written $\mathfrak{s}_{\mathcal{K}}+\nu \pitchfork 0$). We therefore pass to a full subcategory $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}$ of $\boldsymbol{B}_{\mathcal{K}}$ with objects $\mathcal{V} := \bigsqcup V_I$ that does support suitable perturbations $\nu\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}$. This subcategory $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}$ is called a *reduction* of $\mathcal{K}$; cf Definition 3.2.1. Constructing it is akin to passing from the covering of a triangulated space by the stars of its vertices to the covering by the stars of its first barycentric subdivision. Again this construction can be done at the level of the intermediate category, so that the methods of [13] immediately give us the required reductions.

- In the presence of nontrivial isotropy, we may still not be able to construct a transverse perturbation $\nu\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}$ as a functor, since local perturbations ν_I are required to be Γ_I–equivariant. In general, this can be resolved by using multivalued perturbations. Our setup allows for a simplified approach: we define perturbations $\nu = (\nu_I)_{I\in\mathcal{I}_{\mathcal{K}}}$ to be families of maps that are compatible with the covering maps ρ_{IJ}

but need not be Γ_I–equivariant. We show in Section 3.2 that this construction inherits enough equivariance to yield an étale category that represents the zero set of the perturbed section $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}} + \nu$, assuming that this is transverse to 0. The remaining morphisms are then added back in at the expense of weighting functions, which give the perturbed zero set the structure of a *weighted branched manifold*. More precisely, we construct the perturbed zero set in Theorem 3.2.8 as the Hausdorff realization $|\boldsymbol{Z}^\nu|_{\mathcal{H}}$ of an étale (but nonproper) category $\boldsymbol{Z}^\nu$ whose space of objects has one component $Z_I = (s_I|_{V_I} + \nu)^{-1}(0)$ for each $I \in \mathcal{I}_{\mathcal{K}}$, and whose branching locus and weighting function are explicitly determined by the reduction $\mathcal{V}$ and the isotropy groups.

- For the convenience of the reader we prove the needed results about weighted branched manifolds and cobordisms in the appendix. Moreover, the short paper [11] explains the construction of $\boldsymbol{Z}^\nu$ in the orbifold case. This is much simpler, since the obstruction spaces, and hence also the sections $\mathfrak{s}_{\mathcal{K}}$, ν are zero.

- Moreover, we must ensure that the perturbed zero sets are compact and unique up to cobordism. As we show in Proposition 3.3.3 the rather intricate construction in [14] carries through in the current situation without essential change.

- In Section 3.1 we extend the notion of orientation to atlases with nontrivial isotropy. As in [14], we define the orientation line bundle of $\mathcal{K}$ in two equivalent ways, showing in Proposition 3.1.13 that the bundle $\det \mathfrak{s}_{\mathcal{K}}$ (with local bundles $(\det s_I)_{I \in \mathcal{I}_{\mathcal{K}}}$) is isomorphic to $\Lambda_{\mathcal{K}}$ (with local bundles $(\Lambda^{\max} U_I \otimes (\Lambda^{\max} E_I)^*)_{I \in \mathcal{I}_{\mathcal{K}}}$). Most of the needed proofs can again be quoted directly from [14]. Lemma 3.1.14 explains how these bundles are used to orient local zero sets of sections.

- The final step is to build the homology class $[X]^{\text{vir}}_{\mathcal{K}} \in \check{H}_d(X;\mathbb{Q})$ from the zero set $(\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}} + \nu)^{-1}(0)$. Many of the details here are again the same as in [14]. In particular, we build a geometric representative $|\boldsymbol{Z}^\nu|_{\mathcal{H}}$ for this class that maps to the precompact "neighborhood"[2] $|\mathcal{V}| = \bigcup_I \pi_{\mathcal{K}}(V_I) \subset |\mathcal{K}|$ of $\iota_{\mathcal{K}}(X) = |\mathfrak{s}_{\mathcal{K}}^{-1}(0)|$, and then define $[X]^{\text{vir}}_{\mathcal{K}}$ by taking an appropriate inverse limit in rational Čech homology.

Acknowledgements We would like to thank Mohammed Abouzaid, Kenji Fukaya, Tom Mrowka, Kaoru Ono, Yongbin Ruan, Dietmar Salamon, Bernd Siebert, Cliff Taubes, Gang Tian, and Aleksey Zinger for encouragement and enlightening discussions about this project. We moreover thank MSRI, IAS, BIRS and SCGP for hospitality, and acknowledge support by NSF grants DMS 1308669, 0844188, 0905191.

[2] In fact, $\iota_{\mathcal{K}}(X)$ does *not* have a compact neighborhood in $|\mathcal{K}|$; we should think of $|\mathcal{V}|$ as the closest we can come to such a neighborhood.

2 Smooth Kuranishi atlases with isotropy

In this section we extend the notions of smooth Kuranishi charts and transition data introduced in [14] to nontrivial isotropy and then discuss cobordisms and taming. The main result is Theorem 2.5.3.

Throughout this section we fix X to be a compact metrizable space. The main change from [14] is that the domains of the charts are no longer smooth manifolds, but rather group quotients. We begin by setting up notation for the latter. As in [14, Remark 5.1.2] we assume all manifolds are smooth and second countable.

2.1 Group quotients

Definition 2.1.1 A *group quotient* is a pair (U,Γ) consisting of a smooth manifold U and a finite group Γ together with a smooth action $\Gamma\times U\to U$. We will denote the quotient space by

$$\underline{U} := U/\Gamma,$$

giving it the quotient topology, and write $\pi\colon U\to \underline{U}$ for the associated projection. Moreover, we denote the *stabilizer* of each $x\in U$ by

$$\Gamma^x := \{\gamma\in\Gamma \mid \gamma x = x\}\subset\Gamma.$$

We could consider a group quotient as a topological category with space of objects U and morphisms $U\times\Gamma$, but in the interest of simplicity will often avoid doing this.

Both the basic and transition charts of Kuranishi atlases will be group quotients, related by coordinate changes that are composites of the following kinds of maps.

Definition 2.1.2 Let (U,Γ), (U',Γ') be group quotients. A *group embedding*

$$(\phi,\phi^\Gamma)\colon (U,\Gamma)\to(U',\Gamma')$$

is a smooth embedding $\phi\colon U\to U'$ that is equivariant with respect to an injective group homomorphism $\phi^\Gamma\colon\Gamma\to\Gamma'$ and induces an injection $\underline{\phi}\colon \underline{U}\to\underline{U}'$ on the quotient spaces. We call a group embedding *equidimensional* if $\dim U=\dim U'$.

In a Kuranishi atlas we often consider embeddings $(\phi,\phi^\Gamma)\colon(U,\Gamma)\to(U',\Gamma)$ where $\dim U<\dim U'$ and $\phi^\Gamma\colon\Gamma\to\Gamma' := \Gamma$ is the identity map. On the other hand, group quotients of the same dimension are usually related either by restriction or by coverings as follows.

Definition 2.1.3 Let (U,Γ) be a group quotient and $\underline{V}\subset\underline{U}$ an open subset. Then the *restriction of* (U,Γ) *to* $\underline{V}$ is the group quotient $(\pi^{-1}(\underline{V}),\Gamma)$.

Note that the inclusion $\pi^{-1}(\underline{V})\to U$ induces an equidimensional group embedding $(\pi^{-1}(\underline{V}),\Gamma)\to(U,\Gamma)$ that covers the inclusion $\underline{V}\to\underline{U}$. The third kind of map that occurs in a coordinate change is a group covering. This notion is less routine; notice in particular the requirement in (ii) that $\ker\rho^\Gamma$ act freely. Further, the two domains $\widetilde{U}, U$ will necessarily have the same dimension since they are related by a regular covering ρ.

Definition 2.1.4 Let (U,Γ) be a group quotient. A *group covering* of (U,Γ) is a tuple $(\widetilde{U},\widetilde{\Gamma},\rho,\rho^\Gamma)$ consisting of

(i) a surjective group homomorphism $\rho^\Gamma\colon\widetilde{\Gamma}\to\Gamma$,

(ii) a group quotient $(\widetilde{U},\widetilde{\Gamma})$, where $\ker\rho^\Gamma$ acts freely,

(iii) a regular covering $\rho\colon\widetilde{U}\to U$ that is the quotient map $\widetilde{U}\to\widetilde{U}/\ker\rho^\Gamma$ composed with a diffeomorphism $\widetilde{U}/\ker\rho^\Gamma\cong U$ that is equivariant with respect to the induced $\Gamma=\mathrm{im}(\rho^\Gamma)$ action on both spaces.

Thus $\rho\colon\widetilde{U}\to U$ is equivariant with respect to $\rho^\Gamma\colon\widetilde{\Gamma}\to\Gamma$ and ρ^Γ acts transitively on the fibers of ρ. We denote by $\underline{\rho}\colon\underline{\widetilde{U}}\to\underline{U}$ the induced map on quotients.

Next, we establish some basic properties of group quotients, in particular the fact that coverings induce homeomorphisms between the quotients. Here and subsequently we denote a precompact inclusion by $V\sqsubset U$.

Lemma 2.1.5 *Let (U,Γ) be a group quotient.*

(i) *The projection $\pi\colon U\to\underline{U}$ is open, closed and proper. In particular, any precompact set $P\sqsubset\underline{U}$ has precompact preimage $\pi^{-1}(P)\sqsubset U$, Moreover, $\underline{U}$ is a separable, locally compact metric space.*

(ii) *Every point $x\in U$ has a neighborhood U_x that is invariant under Γ^x and is such that inclusion $U_x\hookrightarrow U$ induces a homeomorphism from U_x/Γ^x to $\pi(U_x)$. In particular, the inclusion $(U_x,\Gamma^x)\to(\pi^{-1}(\pi(U_x)),\Gamma)$ is a group embedding.*

(iii) *If $(\widetilde{U},\widetilde{\Gamma},\rho,\rho^\Gamma)$ is a group covering of (U,Γ), then $\underline{\rho}\colon\underline{\widetilde{U}}\to\underline{U}$ is a homeomorphism and ρ^Γ induces isomorphisms between the stabilizers $\widetilde{\Gamma}^y\to\Gamma^{\rho(y)}$ for all $y\in\widetilde{U}$.*

Proof Let $W\subset U$ be open. Then $\pi^{-1}(\pi(W))=\bigcup_{\gamma\in\Gamma}\gamma W$ is open since each γW is the preimage, under the continuous action of γ^{-1}, of the open set W. Hence, by definition of the quotient topology, $\pi(W)$ is open. This shows that π is open. The same argument applied to the complement of a closed set shows that π is closed.

To see that π is proper, consider a compact set $\underline{V} \subset \underline{U}$. Given any open cover $(U_\alpha)_{\alpha\in A}$ of $\pi^{-1}(\underline{V})$, choose for each $x \in \pi^{-1}(\underline{V})$ an element $\alpha_x \in A$ such that $x \in U_{\alpha_x}$. Then for each $\underline{x} \in \underline{V}$ define

$$\underline{W}_{\underline{x}} := \bigcap_{x\in\pi^{-1}(\underline{x})} \pi(U_{\alpha_x}) \subset \underline{U}.$$

These are open sets since $\pi^{-1}(\underline{x})$ is finite and the map π is open, and they cover the compact set $\underline{V}$. So we may choose a finite subcover $(\underline{W}_{\underline{x_i}})_{i=1,\dots,n}$ of $\underline{V}$. Then $(U_{\alpha_x})_{x\in\pi^{-1}\{\underline{x_1},\dots,\underline{x_n}\}}$ is a finite subcover of $\pi^{-1}(\underline{V})$. This shows that preimages of compact sets are compact, ie π is proper.

To see that preimages of precompact sets $P \sqsubset \underline{U}$ are precompact, it suffices to note that the continuity of π gives $\overline{\pi^{-1}(P)} \subset \pi^{-1}(\overline{P})$, so that $\overline{\pi^{-1}(P)}$ is compact because it is a closed subset of $\pi^{-1}(\overline{P})$, which is compact as preimage of the compact set $\overline{P} \subset \underline{U}$.

To finish the proof of (i) we must show that $\underline{U}$ is a separable, locally compact metric space. But $\underline{U}$ inherits these properties from U by [15, Exercise 31.7] which applies to closed continuous surjective maps $\pi\colon X \to Y$ such that $\pi^{-1}(y)$ is compact for all $y \in Y$.

To prove (ii), first choose any open neighborhood $V_x \subset U$ of x that is disjoint from its images under the elements of $\Gamma \setminus \Gamma^x$, and then set

$$U_x := \bigcap_{\gamma\in\Gamma^x} \gamma V_x.$$

Then U_x is open since Γ^x is finite and each γV_x is open. Moreover, U_x is invariant under Γ^x, and has the property that its intersection with each Γ–orbit is either empty or is a Γ^x–orbit. Thus the restriction of π to U_x is simply the quotient by the Γ^x action, so that $U_x/\Gamma^x \to \pi(U_x)$ is the identity.

To prove the first claim in (iii), note that Γ acts on the partial quotient $\widetilde{U}/\ker\rho^\Gamma$ via its identification with $\operatorname{im}\rho^\Gamma = \widetilde{\Gamma}/\ker\rho^\Gamma$ to induce a homeomorphism $\widetilde{U}/\widetilde{\Gamma} \cong (\widetilde{U}/\ker\rho^\Gamma)/\Gamma$. Now $\underline{\rho}$ is this identification composed with the homeomorphism $(\widetilde{U}/\ker\rho^\Gamma)/\Gamma \to U/\Gamma$ induced by the Γ–equivariant diffeomorphism $\widetilde{U}/\ker\rho^\Gamma \cong U$.

As for the statement about stabilizers, notice that we have $\widetilde{\Gamma}^y \cap (\ker\rho^\Gamma) = \mathrm{id}$, because $\ker\rho^\Gamma$ acts freely. Thus $\rho^\Gamma|_{\widetilde{\Gamma}^y}$ is injective. It takes values in Γ^x for $x := \rho(y)$ by the equivariance of ρ with respect to ρ^Γ. To see that $\rho^\Gamma|_{\widetilde{\Gamma}^y}\colon \widetilde{\Gamma}^y \to \Gamma^x$ is surjective, fix an element $\delta \in \Gamma^x$. By surjectivity of $\rho^\Gamma\colon \widetilde{\Gamma} \to \Gamma$ we can choose a lift $\widetilde{\delta} \in (\rho^\Gamma)^{-1}(\delta)$. Since $\rho(\widetilde{\delta}y) = \rho^\Gamma(\widetilde{\delta})\rho(y) = \delta x = \rho(y)$ and the fibers of ρ are $\ker\rho^\Gamma$ orbits, there is a unique $\gamma \in \ker \rho^\Gamma$ such that $\gamma\widetilde{\delta} y = y$, and hence $\gamma\widetilde{\delta} \in \widetilde{\Gamma}^y$. Since $\rho^\Gamma(\gamma\widetilde{\delta}) = \rho^\Gamma(\widetilde{\delta}) = \delta$, this shows that the induced map on stabilizers $\widetilde{\Gamma}^y \to \Gamma^x$ is surjective and hence an isomorphism. □

Remark 2.1.6 In order to make our presentation more accessible we have chosen to require that the domains of our Kuranishi charts are explicit group quotients (U, Γ). Instead we could have worked with étale proper groupoids $\mathcal{G}$ with the additional property that the realization map $\mathrm{Obj}_{\mathcal{G}} \to \mathrm{Obj}_{\mathcal{G}}/{\sim}$, that identifies two objects if and only if there is a morphism between them, is proper. This extra properness assumption is proved for group quotients in Lemma 2.1.5(i). We will see below that this properness allows us to deduce results about a Kuranishi atlas $\mathcal{K}$ from results of [13] applied to the intermediate atlas $\underline{\mathcal{K}}$ in which the charts have domains $\underline{U} = U/\Gamma$. ◇

2.2 Kuranishi charts and coordinate changes

We begin by generalizing the notion of smooth Kuranishi chart (with trivial isotropy) from [14] to the case of nontrivial finite isotropy.

Remark 2.2.1 To simplify language, we will not add the specifications "smooth", "nontrivial isotropy" or "additive" to Kuranishi charts, coordinate changes, and atlases in this paper. Hence a Kuranishi atlas in this paper is a generalization (allowing nontrivial isotropy) of the notion of smooth additive Kuranishi atlas in [14]. We will see that it induces a filtered topological Kuranishi atlas in the sense of [13], given by the "intermediate charts and coordinate changes" introduced in Definition 2.2.3 and Remark 2.2.11 below. So in this paper we will take "intermediate" to include the specification "topological". ◇

Definition 2.2.2 A Kuranishi chart for X is a tuple $\boldsymbol{K} = (U, E, \Gamma, s, \psi)$ consisting of

- the *domain* U, which is a smooth finite-dimensional manifold;
- a finite-dimensional vector space E called the *obstruction space*;
- a finite *isotropy group* Γ with a smooth action on U and a linear action on E;
- a smooth Γ–equivariant function $s\colon U \to E$, called the *section*;
- a continuous map $\psi\colon s^{-1}(0) \to X$ that induces a homeomorphism

$$\underline{\psi}\colon \underline{s^{-1}(0)} := s^{-1}(0)/\Gamma \;\to\; F$$

with open image $F \subset X$, called the *footprint* of the chart.

The *dimension* of $\boldsymbol{K}$ is $\dim(\boldsymbol{K}) := \dim U - \dim E$.

In order to extend topological constructions from [13] to the case of nontrivial isotropy, we will also consider the following notion of intermediate Kuranishi charts which have trivial isotropy but less smooth structure.

Definition 2.2.3 We associate to each Kuranishi chart $\boldsymbol{K} = (U, E, \Gamma, s, \psi)$ the *intermediate chart* $\underline{\boldsymbol{K}} := (\underline{U}, \underline{\mathbb{E}}, \underline{\mathfrak{s}}, \underline{\psi})$ consisting of

- the *intermediate domain* $\underline{U} := U/\Gamma$;
- the *intermediate obstruction "bundle"*, whose total space $\underline{\mathbb{E}} := \underline{U \times E}$ is the quotient by the diagonal action of Γ, with the projection map $\mathrm{pr}\colon \underline{\mathbb{E}} \to \underline{U}$, $\Gamma(u,e) \mapsto \Gamma u$, and zero section $0\colon \underline{U} \to \underline{\mathbb{E}}$, $\Gamma u \mapsto \Gamma(u,0)$;
- the *intermediate section* $\underline{\mathfrak{s}}\colon \underline{U} \to \underline{\mathbb{E}}$ induced by $\mathfrak{s} = \mathrm{id}_U \times s\colon U \to U \times E$;
- the *intermediate footprint map* $\underline{\psi}\colon \underline{\mathfrak{s}}^{-1}(\mathrm{im}\, 0) \to X$ induced by $\psi\colon s^{-1}(0) \to X$.

We write $\pi\colon U \to \underline{U}$ for the projection from the Kuranishi domain. Moreover if a chart $\boldsymbol{K}_I = (U_I, E_I, \Gamma_I, s_I, \psi_I)$ has the label I, then $\underline{\boldsymbol{K}}_I = (\underline{U}_I, \underline{\mathbb{E}}_I, \underline{\mathfrak{s}}_I, \underline{\psi}_I)$ and $\pi_I\colon U_I \to \underline{U}_I$ denote the corresponding intermediate chart and projection.

The intermediate charts and coordinate changes of a Kuranishi atlas (with isotropy) will form a topological Kuranishi atlas (without isotropy). For the charts, the following is a direct consequence of Lemma 2.1.5.

Lemma 2.2.4 *The intermediate chart* $\underline{\boldsymbol{K}}$ *is a topological chart in the sense of* [13, Definition 2.1.3]. *In other words,*

- *the intermediate domain* $\underline{U}$ *is a separable, locally compact metric space;*
- *the intermediate obstruction "bundle"* $\mathrm{pr}\colon \underline{\mathbb{E}} \to \underline{U}$ *is a continuous map between separable, locally compact metric spaces, so that the zero section* $0\colon \underline{U} \to \underline{\mathbb{E}}$ *is a continuous map with* $\mathrm{pr} \circ 0 = \mathrm{id}_{\underline{U}}$*;*
- *the intermediate section* $\underline{\mathfrak{s}}\colon \underline{U} \to \underline{\mathbb{E}}$ *is a continuous map with* $\mathrm{pr} \circ \underline{\mathfrak{s}} = \mathrm{id}_{\underline{U}}$*;*
- *the intermediate footprint map* $\underline{\psi}\colon \underline{\mathfrak{s}}^{-1}(0) \to X$ *is a homeomorphism onto the footprint* $\underline{\psi}(\underline{\mathfrak{s}}^{-1}(0)) = F$*, which is an open subset of* X*.*

Remark 2.2.5 (i) The intermediate bundle $\mathrm{pr}\colon \underline{\mathbb{E}} \to \underline{U}$ is an orbibundle and hence has more structure than a general topological chart. In particular, it has a natural zero section $0\colon \underline{U} \to \underline{\mathbb{E}}$. Hence, when working with labeled charts $\underline{\boldsymbol{K}}_I$, we will usually simply denote the projection and zero section by pr and 0 rather than $\mathrm{pr}_I, 0_I$.

(ii) We will find that many results from [13], in particular the taming constructions, carry over to nontrivial isotropy via the intermediate charts, since precompact subsets of $\underline{U}$ lift to precompact subsets of U by Lemma 2.1.5(i). An important exception is the construction of perturbations which must be done on the smooth spaces U. $\diamond$

Next, as in [13; 14], compatibility of Kuranishi charts will require restrictions and embeddings to common transition charts.

Definition 2.2.6 Let $\boldsymbol{K} = (U, E, \Gamma, s, \psi)$ be a Kuranishi chart and $F' \subset F$ an open subset of its footprint. A *restriction of* $\boldsymbol{K}$ *to* F' is a Kuranishi chart of the form

$$\boldsymbol{K}' = \boldsymbol{K}|_{\underline{U}'} := (U', E, \Gamma, s' = s|_{U'}, \psi' = \psi|_{s'^{-1}(0)}) \quad \text{with } U' := \pi^{-1}(\underline{U}')$$

given by a choice of open subset $\underline{U}' \subset \underline{U}$ such that $\underline{U}' \cap \underline{\psi}^{-1}(F) = \underline{\psi}^{-1}(F')$. We call $\underline{U}'$ the *domain* of the restriction.

Note that the restriction $\boldsymbol{K}'$ in the above definition has footprint $\psi'(s'^{-1}(0)) = F'$, and its domain group quotient (U', Γ) is the restriction of (U, Γ) to $\underline{U}'$ in the sense of Definition 2.1.3. Moreover, because the restriction of a chart is determined by a subset of the intermediate domain $\underline{U}$, we can in the following use the existence result in [13] for restrictions of topological charts to obtain restrictions of charts with isotropy. Here we use the notation $\sqsubset$ to denote a precompact inclusion and we write $\mathrm{cl}_V(V')$ for the closure of a subset $V' \subset V$ in the relative topology of V.

Lemma 2.2.7 *Let* $\boldsymbol{K}$ *be a Kuranishi chart. Then for any open subset* $F' \subset F$ *there is a restriction* $\boldsymbol{K}'$ *to* F' *with domain* $\underline{U}'$ *such that* $U' := \pi^{-1}(\underline{U}')$ *satisfies* $\mathrm{cl}_U(U') \cap s^{-1}(0) = \psi^{-1}(\mathrm{cl}_X(F'))$. *Moreover, if* $F' \sqsubset F$ *is precompact, then* $\underline{U}' \sqsubset \underline{U}$ *can be chosen precompact so that* $U' \sqsubset U$.

Proof By [13, Lemma 2.1.6] applied to the intermediate chart $\underline{\boldsymbol{K}}$, there is a subset $\underline{U}' \subset \underline{U}$ that defines a restriction of this topological chart, and in particular satisfies $\underline{U}' \cap \underline{\mathfrak{s}}^{-1}(0) = \underline{\psi}^{-1}(F')$, with the additional property $\mathrm{cl}_{\underline{U}}(\underline{U}') \cap \underline{\mathfrak{s}}^{-1}(0) = \underline{\psi}^{-1}(\mathrm{cl}_X(F'))$. Further, we may assume that $\underline{U}'$ is precompact in $\underline{U}$ if $F' \sqsubset F$. Then $U' = \pi^{-1}(\underline{U}')$ is the required domain. It inherits precompactness by Lemma 2.1.5(i). Further, the same lemma shows that $\pi^{-1}(\mathrm{cl}_{\underline{U}}(\underline{U}')) = \mathrm{cl}_U(U')$. Hence applying π^{-1} to the identity $\mathrm{cl}_{\underline{U}}(\underline{U}') \cap \underline{\mathfrak{s}}^{-1}(0) = \underline{\psi}^{-1}(\mathrm{cl}_X(F'))$ implies that $\mathrm{cl}_U(U') \cap s^{-1}(0) = \psi^{-1}(\mathrm{cl}_X(F'))$. □

Most definitions in [14] extend, as the previous ones, with only minor changes to the case of nontrivial isotropy. However, the notion of smooth coordinate change [14, Definition 5.2.2] needs to be generalized significantly to include a covering map. For simplicity we will formulate the definition in the situation that is relevant to additive Kuranishi atlases.[3] That is, we suppose that a finite set of basic Kuranishi charts $(\boldsymbol{K}_i)_{i \in \{1,\dots,N\}}$ is given so that for each $I \subset \{1, \dots, N\}$ with $F_I := \bigcap_{i \in I} F_i \neq \varnothing$ we have another Kuranishi chart $\boldsymbol{K}_I$ with

[3] While additivity was introduced as separate property in [12], it is both so crucial and natural that below in Section 2.3 we will define the notion of Kuranishi atlas to be automatically additive.

- isotropy group $\Gamma_I := \prod_{i\in I} \Gamma_i$,
- obstruction space $E_I := \prod_{i\in I} E_i$ on which Γ_I acts with the product action,
- footprint $F_I := \bigcap_{i\in I} F_i$.

Then for $I \subset J$ we have the natural splitting $\Gamma_J \cong \Gamma_I \times \Gamma_{J\setminus I}$ with induced inclusion $\Gamma_I \hookrightarrow \Gamma_I \times \{\mathrm{id}\} \subset \Gamma_J$ and projection $\rho^{\Gamma}_{IJ}\colon \Gamma_J \to \Gamma_I$ with kernel $\Gamma_{J\setminus I}$. (Here we include the case $I = J$, interpreting $\Gamma_\varnothing := \{\mathrm{id}\}$.) Moreover, we have the natural inclusion $\widehat{\phi}_{IJ}\colon E_I \to E_J$, which is equivariant with respect to the inclusion $\Gamma_I \hookrightarrow \Gamma_J$ and such that the complement of this inclusion $\Gamma_{J\setminus I}$ acts trivially on the image $\widehat{\phi}_{IJ}(E_I) \subset E_J$.

Definition 2.2.8 Given $I \subset J \subset \{1, \dots, N\}$ let $\boldsymbol{K}_I$ and $\boldsymbol{K}_J$ be Kuranishi charts as above with $F_I \supset F_J$. A *smooth coordinate change* $\widehat{\Phi}_{IJ}$ from $\boldsymbol{K}_I$ to $\boldsymbol{K}_J$ consists of

- a choice of domain $\underline{U}_{IJ} \subset \underline{U}_I$ such that $\boldsymbol{K}_I|_{\underline{U}_{IJ}}$ is a restriction of $\boldsymbol{K}_I$ to F_J,
- the splitting $\Gamma_J \cong \Gamma_I \times \Gamma_{J\setminus I}$ as above, and the induced inclusion $\Gamma_I \hookrightarrow \Gamma_J$ and projection $\rho^{\Gamma}_{IJ}\colon \Gamma_J \to \Gamma_I$,
- a Γ_J–invariant submanifold $\widetilde{U}_{IJ} \subset U_J$ on which $\Gamma_{J\setminus I}$ acts freely, and the induced Γ_J–equivariant inclusion $\widetilde{\phi}_{IJ}\colon \widetilde{U}_{IJ} \hookrightarrow U_J$,
- a group covering $(\widetilde{U}_{IJ}, \Gamma_J, \rho_{IJ}, \rho^{\Gamma}_{IJ})$ of the group quotient (U_{IJ}, Γ_I), where $U_{IJ} := \pi_I^{-1}(\underline{U}_{IJ}) \subset U_I$,
- the linear equivariant injection $\widehat{\phi}_{IJ}\colon E_I \to E_J$ as above,

such that the inclusions $\widetilde{\phi}_{IJ}, \widehat{\phi}_{IJ}$ and covering ρ_{IJ} intertwine the sections and footprint maps,

$$
\begin{aligned}
&s_J \circ \widetilde{\phi}_{IJ} = \widehat{\phi}_{IJ} \circ s_I \circ \rho_{IJ} && \text{on } \widetilde{U}_{IJ}, \\
&\psi_J \circ \widetilde{\phi}_{IJ} = \psi_I \circ \rho_{IJ} && \text{on } s_J^{-1}(0) \cap \widetilde{U}_{IJ} = \rho_{IJ}^{-1}(s_I^{-1}(0)).
\end{aligned}
\tag{2.2.1}
$$

Moreover, we denote $s_{IJ} := s_I \circ \rho_{IJ}\colon \widetilde{U}_{IJ} \to E_I$ and require the *index condition*:

(i) The embedding $\widetilde{\phi}_{IJ}\colon \widetilde{U}_{IJ} \hookrightarrow U_J$ identifies the kernels:

$$
\mathrm{d}_u \widetilde{\phi}_{IJ}(\ker \mathrm{d}_u s_{IJ}) = \ker \mathrm{d}_{\widetilde{\phi}_{IJ}(u)} s_J \quad \forall u \in \widetilde{U}_{IJ}.
$$

(ii) The linear embedding $\widehat{\phi}_{IJ}\colon E_I \to E_J$ identifies the cokernels:

$$
E_I = \mathrm{im}(\mathrm{d}_u s_{IJ}) \oplus C_{u,I} \implies E_J = \mathrm{im}(\mathrm{d}_{\widetilde{\phi}_{IJ}(u)} s_J) \oplus \widehat{\phi}_{IJ}(C_{u,I}) \quad \forall u \in \widetilde{U}_{IJ}.
$$

The subset $\underline{U}_{IJ} \subset \underline{U}_I$ is called the *domain* of the coordinate change, while $\widetilde{U}_{IJ} \subset U_J$ is its *lifted domain*.

Recall that we have $\dim \widetilde{U}_{IJ} = \dim U_I$ since $\rho_{IJ}\colon \widetilde{U}_{IJ} \to U_{IJ}$ is a regular covering. Moreover, ρ_{IJ} identifies the kernels and images of $\mathrm{d}s_{IJ}$ and $\mathrm{d}s_I$, in other words

(2.2.2) $$\mathrm{d}_u\rho_{IJ}(\ker \mathrm{d}_u s_{IJ}) = \ker \mathrm{d}_{\rho_{IJ}(u)} s_I, \quad \operatorname{im}(\mathrm{d}_u s_{IJ}) = \operatorname{im}(\mathrm{d}_{\rho_{IJ}(u)} s_I) \subset E_I.$$

Hence the index condition is equivalent to kernels and cokernels of $\mathrm{d}_{\rho_{IJ}(u)} s_I$ and $\mathrm{d}_u s_J$ being identified by the coordinate change. As in [14, Lemma 5.2.5] it is also equivalent to the *tangent bundle condition*

(2.2.3) $$\mathrm{d}_{\widetilde{\phi}_{IJ}(u)} s_J\colon \mathrm{T}_{\widetilde{\phi}_{IJ}(u)} U_J \big/ \mathrm{d}_u\widetilde{\phi}_{IJ}(\mathrm{T}_u \widetilde{U}_{IJ}) \xrightarrow{\cong} E_J \big/ \widehat{\phi}_{IJ}(E_I) \quad \forall u \in \widetilde{U}_{IJ}.$$

This also shows that any two charts that are related by a coordinate change have the same dimension. To keep our language similar to that in [14], we denote a coordinate change by $\widehat{\Phi}_{IJ} = (\widetilde{\phi}_{IJ}, \widehat{\phi}_{IJ}, \rho_{IJ})\colon \boldsymbol{K}_I|_{\underline{U}_{IJ}} \to \boldsymbol{K}_J$. However, since the linear map $\widehat{\phi}_{IJ}$ is fixed by our conventions, the coordinate change $\widehat{\Phi}_{IJ}$ is in fact determined by a group covering $(\widetilde{U}_{IJ}, \Gamma_J, \rho_{IJ}, \rho^\Gamma_{IJ})$ of $(\pi_I^{-1}(\underline{U}_{IJ}), \Gamma_I)$, where $\underline{U}_{IJ} \subset \underline{U}_I$ is a choice of domain for which $\underline{U}_{IJ} \cap \underline{\psi}_I^{-1}(F_I) = \underline{\psi}_I^{-1}(F_J)$.

Remark 2.2.9 (i) In the case of trivial isotropy and with trivial covering $\rho_{IJ} =: \phi_{IJ}^{-1}$, this definition is the notion of coordinate change in [14] with $\widetilde{U}_{IJ} = \phi_{IJ}(U_{IJ})$. Because $U_{IJ} \subset U_I$ is open, the index condition together with the condition that $\widetilde{U}_{IJ}$ is a submanifold of U_J implies that $\widetilde{U}_{IJ}$ is an open subset of $s_J^{-1}(E_I)$.

(ii) The following diagram of group embeddings and group coverings is associated to each coordinate change:

(2.2.4) $$\begin{tikzcd}
 & (\widetilde{U}_{IJ}, \Gamma_J) \arrow[r, hook, "(\widetilde{\phi}_{IJ}, \mathrm{id})"] \arrow[d, "(\rho_{IJ}, \rho^\Gamma_{IJ})"] & (U_J, \Gamma_J) \\
(U_I, \Gamma_I) & (U_{IJ}, \Gamma_I) \arrow[l, hook'] &
\end{tikzcd}$$

(iii) Since $\underline{\rho}_{IJ}\colon \underline{\widetilde{U}}_{IJ} \to \underline{U}_{IJ}$ is a homeomorphism by Lemma 2.1.5(iii), each coordinate change $(\phi_{IJ}, \widehat{\phi}_{IJ}, \rho_{IJ})\colon \boldsymbol{K}_I|_{\underline{U}_{IJ}} \to \boldsymbol{K}_J$ induces an injective map

$$\underline{\phi}_{IJ} := \underline{\widetilde{\phi}}_{IJ} \circ \underline{\rho}_{IJ}^{-1}\colon \underline{U}_{IJ} \to \underline{U}_J$$

on the domain of the intermediate chart. In fact there is an induced coordinate change $\underline{\widehat{\Phi}}_{IJ}\colon \underline{\boldsymbol{K}}_I|_{\underline{U}_{IJ}} \to \underline{\boldsymbol{K}}_J$ between the intermediate charts, given by the bundle map $\underline{\widehat{\Phi}}_{IJ}\colon \underline{U_{IJ} \times E_I} \to \underline{U_J \times E_J}$ which is induced by the multivalued map $(\widetilde{\phi}_{IJ} \circ \rho_{IJ}^{-1}) \times \widehat{\phi}_{IJ}$ and thus covers $\underline{\widetilde{\phi}}_{IJ} \circ \underline{\rho}_{IJ}^{-1} =: \underline{\phi}_{IJ}$. This is a topological coordinate change in the sense of [13, Definition 2.2.1]. This means in particular that the map

$$\underline{\widehat{\Phi}}_{IJ}\colon \underline{U_{IJ} \times E_I} =: \underline{\mathbb{E}}_I|_{\underline{U}_{IJ}} := \mathrm{pr}_I^{-1}(\underline{U}_{IJ}) \to \underline{\mathbb{E}}_J$$

is a topological embedding (ie homeomorphism to its image) that satisfies the following:

- It is a bundle map, ie we have $\mathrm{pr}_J \circ \underline{\widehat{\Phi}}_{IJ} = \underline{\phi}_{IJ} \circ \mathrm{pr}_I|_{\mathrm{pr}_I^{-1}(\underline{U}_{IJ})}$ for a topological embedding $\underline{\phi}_{IJ}\colon \underline{U}_{IJ} \to \underline{U}_J$, and it is linear in the sense that $0_J \circ \underline{\phi}_{IJ} = \underline{\widehat{\Phi}}_{IJ} \circ 0_I|_{U_{IJ}}$, where 0_I denotes the zero section $0_I\colon \underline{U}_I \to \underline{\mathbb{E}}_I$ in the chart $\underline{\boldsymbol{K}}_I$.
- It intertwines the sections and footprints maps, ie

$$\underline{\mathfrak{s}}_J \circ \underline{\phi}_{IJ} = \underline{\widehat{\Phi}}_{IJ} \circ \underline{\mathfrak{s}}_I|_{U_{IJ}}, \quad \underline{\phi}_{IJ}|_{\underline{\psi}_I^{-1}(F_I \cap F_J)} = \underline{\psi}_J^{-1} \circ \underline{\psi}_I.$$

However, $\underline{\widehat{\Phi}}_{IJ}$ has more smooth structure than a general topological coordinate change since $\underline{\phi}_{IJ}\colon \underline{U}_{IJ} \to \underline{U}_J$ preserves the orbifold structure and $\underline{\widehat{\Phi}}_{IJ}$ is a map of orbibundles.

(iv) Conversely, suppose we are given a topological coordinate change $\underline{\widehat{\Phi}}_{IJ}\colon \underline{\boldsymbol{K}}_I \to \underline{\boldsymbol{K}}_J$ with domain $\underline{U}_{IJ}$. Then any coordinate change from $\boldsymbol{K}_I$ to $\boldsymbol{K}_J$ that induces $\underline{\widehat{\Phi}}_{IJ}$ is determined by the Γ_J–invariant set $\widetilde{U}_{IJ} := \pi_J^{-1}(\underline{\phi}_{IJ}(\underline{U}_{IJ}))$ and a choice of Γ_I–equivariant homeomorphism between $\widetilde{U}_{IJ}/\Gamma_{J\setminus I}$ and $U_{IJ} := \pi_I^{-1}(\underline{U}_{IJ})$. If we can choose this homeomorphism to be smooth, then we obtain a smooth coordinate change $\boldsymbol{K}_I \to \boldsymbol{K}_J$ with domain $\underline{U}_{IJ}$ provided that the index condition is satisfied, which is a condition on the relation between the set $\widetilde{U}_{IJ}$ and the section s_J. When constructing coordinate changes in the Gromov–Witten setting in [10], we will see that there is a natural choice of this diffeomorphism since the covering maps ρ_{IJ} are given by forgetting certain added marked points. Further, the index condition is automatically satisfied in this setting.

(v) Because $\widetilde{U}_{IJ}$ is defined to be a subset of U_J, it is sometimes convenient to think of an element $\tilde{x} \in \widetilde{U}_{IJ}$ as an element in U_J, omitting the notation for the inclusion map $\tilde{\phi}_{IJ}\colon \widetilde{U}_{IJ} \to U_J$. ◇

The next step is to consider restrictions and composites of coordinate changes. Restrictions exist analogously to [14, Lemma 5.2.6]: for $I \subset J$, given a coordinate change $\widehat{\Phi}_{IJ}\colon \boldsymbol{K}_I|_{\underline{U}_{IJ}} \to \boldsymbol{K}_J$ and restrictions $\boldsymbol{K}'_I := \boldsymbol{K}_I|_{\underline{U}'_I}$ and $\boldsymbol{K}'_J := \boldsymbol{K}_J|_{\underline{U}'_J}$ whose footprints $F'_I \cap F'_J$ have nonempty intersection, there is an induced *restricted coordinate change* $\widehat{\Phi}_{IJ}|_{\underline{U}'_{IJ}}\colon \boldsymbol{K}'_I|_{\underline{U}'_{IJ}} \to \boldsymbol{K}'_J$ for any open subset $\underline{U}'_{IJ} \subset \underline{U}_{IJ}$ satisfying the conditions

$$\underline{U}'_{IJ} \subset \underline{U}'_I \cap \underline{\phi}_{IJ}^{-1}(\underline{U}'_J), \quad \underline{U}'_{IJ} \cap \underline{\mathfrak{s}}_I^{-1}(0) = \underline{\psi}_I^{-1}(F'_I \cap F'_J). \tag{2.2.5}$$

However, coordinate changes now do not directly compose due to the coverings involved. The induced coordinate changes on the intermediate charts still compose directly, but the analog of [14, Lemma 5.2.7] is the following.

Lemma 2.2.10 *Let $I \subset J \subset K$ (so that automatically $F_I \supset F_J \supset F_K$) and suppose that $\widehat{\Phi}_{IJ}\colon \boldsymbol{K}_I \to \boldsymbol{K}_J$ and $\widehat{\Phi}_{JK}\colon \boldsymbol{K}_J \to \boldsymbol{K}_K$ are coordinate changes with domains $\underline{U}_{IJ}$ and $\underline{U}_{JK}$ respectively. Then:*

(i) *The domain $\underline{U}_{IJK} := \underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK}) \subset \underline{U}_I$ defines a restriction $\boldsymbol{K}_I|_{\underline{U}_{IJK}}$ of $\boldsymbol{K}_I$ to F_K.*

(ii) *The composite $\rho_{IJK} := \rho_{IJ} \circ \rho_{JK}\colon \widetilde{U}_{IJK} \to U_{IJK} := \pi_I^{-1}(\underline{U}_{IJK})$ is defined on $\widetilde{U}_{IJK} := \pi_K^{-1}((\underline{\phi}_{JK} \circ \underline{\phi}_{IJ})(\underline{U}_{IJK}))$ via the natural identification of $\rho_{JK}(\widetilde{U}_{IJK}) \subset U_J$ with a subset of $\widetilde{U}_{IJ}$. Together with the natural projection $\rho^{\Gamma}_{IK}\colon \Gamma_K \to \Gamma_I$ with kernel $\Gamma_{K\setminus I}$, which factors $\rho^{\Gamma}_{IK} = \rho^{\Gamma}_{IJ} \circ \rho^{\Gamma}_{JK}$, this forms a group covering $(\widetilde{U}_{IJK}, \Gamma_K, \rho_{IJK}, \rho^{\Gamma}_{IK})$ of (U_{IJK}, Γ_I).*

(iii) *The inclusion $\widetilde{\phi}_{IJK}\colon \widetilde{U}_{IJK} \hookrightarrow U_K$, taken together with the natural inclusion $\widehat{\phi}_{IK}\colon E_I \to E_K$ (which factors $\widehat{\phi}_{IK} = \widehat{\phi}_{JK} \circ \widehat{\phi}_{IJ}$) and ρ_{IJK}, satisfies (2.2.1) and the index condition with respect to the indices I, K.*

Hence this defines a ***composite coordinate change***

$$\widehat{\Phi}_{JK} \circ \widehat{\Phi}_{IJ} := \widehat{\Phi}_{IJK} = (\widetilde{\phi}_{IJK}, \widehat{\phi}_{IK}, \rho_{IJK})$$

from $\boldsymbol{K}_I$ to $\boldsymbol{K}_K$ with domain $\underline{U}_{IJK}$.

Proof The corresponding statement for the induced coordinate changes for the intermediate charts is proved in [13, Lemma 2.2.5]. Thus claim (i) follows from part (i) of [13, Lemma 2.2.5].

To see that ρ_{IJK} in (ii) is well defined, we need to verify that $\rho_{JK}(\widetilde{U}_{IJK}) \subset \widetilde{U}_{IJ}$, or (due to equivariance) equivalently $\underline{\rho}_{JK}(\underline{\widetilde{U}}_{IJK}) \subset \underline{\widetilde{U}}_{IJ}$. For that purpose we drop the natural identifications $\underline{\widetilde{\phi}}_{IJ}\colon \underline{\widetilde{U}}_{IJ} \to \underline{U}_J$ from the notation so that the intermediate coordinate changes are $\underline{\phi}_{IJ} = \underline{\rho}_{IJ}^{-1}\colon \underline{U}_{IJ} \to \underline{\widetilde{U}}_{IJ} \subset \underline{U}_J$ and the inclusion follows from

$$\begin{aligned}\underline{\rho}_{JK}(\underline{\widetilde{U}}_{IJK}) &= \underline{\rho}_{JK}((\underline{\phi}_{JK} \circ \underline{\phi}_{IJ})(\underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK}))) \\ &= (\underline{\rho}_{JK} \circ \underline{\phi}_{JK})(\underline{\widetilde{U}}_{IJ} \cap \underline{U}_{JK}) \\ &= \underline{\widetilde{U}}_{IJ} \cap \underline{U}_{JK}.\end{aligned}$$

Next, observe that composites of group covering maps are also group covering maps. In particular, since $\Gamma_{K\setminus J}$ acts freely on $\widetilde{U}_{IJK} \subset \widetilde{U}_{JK}$ and $\Gamma_{J\setminus I}$ acts freely on the quotient $\widetilde{U}_{IJK}/\Gamma_{K\setminus J}$ (because it is identified Γ_J–equivariantly with a subset of $\widetilde{U}_{IJ}$), the group $\Gamma_{K\setminus I} \cong \Gamma_{K\setminus J} \times \Gamma_{J\setminus I}$ acts freely on $\widetilde{U}_{IJK}$.

To prove (iii), first observe that (2.2.1) holds for the index pair IK because it holds for IJ and JK:

$$\begin{aligned} s_K \circ \widetilde{\phi}_{IJK} &= \widehat{\phi}_{JK} \circ s_J \circ \rho_{JK}|_{\widetilde{U}_{IJK}} \\ &= \widehat{\phi}_{JK} \circ (\widehat{\phi}_{IJ} \circ s_I \circ \rho_{IJ}) \circ \rho_{JK}|_{\widetilde{U}_{IJK}} \\ &= \widehat{\phi}_{IK} \circ s_I \circ \rho_{IJK} && \text{on } \widetilde{U}_{IJK}, \\ \psi_K \circ \widetilde{\phi}_{IJK} &= \psi_J \circ \rho_{JK} \\ &= \psi_I \circ \rho_{IJ} \circ \rho_{JK} \\ &= \psi_I \circ \rho_{IJK} && \text{on } s_K^{-1}(0) \cap \widetilde{U}_{IJK}. \end{aligned}$$

Finally, it is easiest to check the index condition in the form given in (2.2.3), ie we need to establish isomorphisms for all $u \in \widetilde{U}_{IJK}$,

$$\mathrm{d}_{\widetilde{\phi}_{IJK}(u)} s_K\colon\ \mathrm{T}_{\widetilde{\phi}_{IJK}(u)} U_K / \mathrm{d}_u \widetilde{\phi}_{IJK}(\mathrm{T}_u \widetilde{U}_{IJK}) \xrightarrow{\cong} E_K / \widehat{\phi}_{IK}(E_I). \tag{2.2.6}$$

Here and below we will suppress the natural embedding $\widetilde{\phi}_{IJK}\colon \widetilde{U}_{IJK} \to U_K$ from the notation, hence identifying eg $u \in \widetilde{U}_{IJK}$ with $\widetilde{\phi}_{IJK}(u) \in U_K$. With that, the quotient on the left is naturally identified with the normal fiber $\mathrm{T}_u U_K / \mathrm{T}_u \widetilde{U}_{IJK}$ to the submanifold $\widetilde{U}_{IJK}$ of U_K. Next, $\widetilde{U}_{IJK}$ is by construction a submanifold of $\widetilde{U}_{JK}$, which in turn is a submanifold of U_K, hence this normal fiber is isomorphic to the direct sum of the normal fiber of $\widetilde{U}_{IJK}$ in $\widetilde{U}_{JK}$ together with that of $\widetilde{U}_{JK}$ in U_K,

$$\mathrm{T}_u U_K / \mathrm{T}_u \widetilde{U}_{IJK} \cong \mathrm{T}_u U_K / \mathrm{T}_u \widetilde{U}_{JK} \oplus \mathrm{T}_u \widetilde{U}_{JK} / \mathrm{T}_u \widetilde{U}_{IJK}.$$

By the index condition for $\widehat{\Phi}_{JK}$, the map $\mathrm{d}_u s_K$ restricted to the first summand induces an isomorphism $\mathrm{T}_u U_K / \mathrm{T}_u \widetilde{U}_{JK} \xrightarrow{\cong} E_K / \widehat{\Phi}_{JK}(E_J)$. Considering the second summand, recall that on $\widetilde{U}_{JK}$ we have $s_K = s_J \circ \rho_{JK}$, where $\rho_{JK}\colon \widetilde{U}_{JK} \to U_{JK}$ is a local diffeomorphism onto an open subset of U_J. It maps $\widetilde{U}_{IJK}$ to $\rho_{JK}(\widetilde{U}_{IJK}) = \widetilde{U}_{IJ} \cap U_{JK}$ so that, with $v := \rho_{JK}(u)$, the map $\mathrm{d}_u \rho_{JK}$ induces an isomorphism $\mathrm{T}_u \widetilde{U}_{JK} / \mathrm{T}_u \widetilde{U}_{IJK} \xrightarrow{\cong} \mathrm{T}_v U_J / \mathrm{T}_v \widetilde{U}_{IJ}$. Thus the restriction of $\mathrm{d}_u s_K$ to the second summand induces the isomorphism

$$\mathrm{d}_v s_J \circ \mathrm{d}_u \rho_{JK}\colon \mathrm{T}_u \widetilde{U}_{JK} / \mathrm{T}_u \widetilde{U}_{IJK} \xrightarrow{\cong} \mathrm{T}_v U_J / \mathrm{T}_v \widetilde{U}_{IJ} \xrightarrow{\cong} E_J / \widehat{\Phi}_{IJ}(E_I),$$

where the second isomorphism results from the index condition for $\widehat{\Phi}_{IJ}$. Putting this all together, $\mathrm{d}_u s_K$ induces an isomorphism from $\mathrm{T}_u U_K / \mathrm{T}_u \widetilde{U}_{IJK}$ to

$$E_K / \widehat{\Phi}_{JK}(E_J) \oplus E_J / \widehat{\Phi}_{IJ}(E_I) \cong E_K / \widehat{\Phi}_{IK}(E_I),$$

where in the last step we used the fact that $\widehat{\Phi}_{JK}\colon E_J \to E_K$ is the natural inclusion. This establishes the isomorphism (2.2.6) and thus completes the proof. □

Remark 2.2.11 The composition $\widehat{\Phi}_{IJK}\colon \boldsymbol{K}_I \to \boldsymbol{K}_K$ induces a coordinate change $\underline{\widehat{\Phi}}_{IJK}\colon \underline{\boldsymbol{K}}_I \to \underline{\boldsymbol{K}}_K$ on the intermediate charts. This agrees with the composition of the intermediate coordinate changes $\underline{\widehat{\Phi}}_{IJ}, \underline{\widehat{\Phi}}_{JK}$ as defined for topological charts in [13, Lemma 2.2.5]. ◇

Next, the cocycle conditions from [13, Definition 2.3.2] have direct generalizations.

Definition 2.2.12 Let $\boldsymbol{K}_\alpha$ for $\alpha = I, J, K$ be Kuranishi charts with $I \subset J \subset K$, and let $\widehat{\Phi}_{\alpha\beta}\colon \boldsymbol{K}_\alpha|_{\underline{U}_{\alpha\beta}} \to \boldsymbol{K}_\beta$ for $(\alpha,\beta) \in \{(I,J),(J,K),(I,K)\}$ be coordinate changes. We say that this triple $\widehat{\Phi}_{IJ}, \widehat{\Phi}_{JK}, \widehat{\Phi}_{IK}$ satisfies the

- *weak cocycle condition* if $\widehat{\Phi}_{JK} \circ \widehat{\Phi}_{IJ} \approx \widehat{\Phi}_{IK}$ are equal on the overlap, in the sense that

$$\rho_{IK} = \rho_{IJ} \circ \rho_{JK} \quad \text{on } \widetilde{U}_{IK} \cap \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}); \tag{2.2.7}$$

- *cocycle condition* if $\widehat{\Phi}_{JK} \circ \widehat{\Phi}_{IJ} \subset \widehat{\Phi}_{IK}$, ie $\widehat{\Phi}_{IK}$ extends the composed coordinate change in the sense that (2.2.7) holds and

$$\underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK}) \subset \underline{U}_{IK}; \tag{2.2.8}$$

- *strong cocycle condition* if $\widehat{\Phi}_{JK} \circ \widehat{\Phi}_{IJ} = \widehat{\Phi}_{IK}$ are equal as coordinate changes, that is if (2.2.7) holds and

$$\underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK}) = \underline{U}_{IK}. \tag{2.2.9}$$

We stated these last two conditions on the level of the intermediate category because, as we now show, they imply corresponding identities on the level of the Kuranishi atlas.

Lemma 2.2.13 (i) *Condition (2.2.7) implies*

$$\underline{\phi}_{IK} = \underline{\phi}_{JK} \circ \underline{\phi}_{IJ} \quad \text{on } \underline{U}_{IK} \cap (\underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK}));$$

(ii) *The cocycle condition (2.2.8) implies that*

$$\rho_{IK} = \rho_{IJ} \circ \rho_{JK} \quad \text{on } \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}) \subset \widetilde{U}_{IK}.$$

(iii) *The strong cocycle condition (2.2.9) implies that*

$$\rho_{IK} = \rho_{IJ} \circ \rho_{JK} \quad \text{on } \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}) = \widetilde{U}_{IK}.$$

Proof By definition, $\underline{\rho}_{\alpha\beta} \circ \pi_\beta = \pi_\alpha \circ \rho_{\alpha\beta}$ when $\alpha \subset \beta$, so condition (2.2.7) implies

$$\underline{\rho}_{IK} = \underline{\rho}_{IJ} \circ \underline{\rho}_{JK} \quad \text{on } \pi_K(\widetilde{U}_{IK} \cap \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK})).$$

The identity $\underline{\phi}_{\alpha\beta} = \underline{\rho}_{\alpha\beta}^{-1}$ from Remark 2.2.9(iii) then implies $\underline{\phi}_{IK} = \underline{\phi}_{JK} \circ \underline{\phi}_{IJ}$ on

$$\begin{aligned}
\underline{\rho}_{IK}(\pi_K(\widetilde{U}_{IK} \cap \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}))) &= \pi_I(\rho_{IK}(\widetilde{U}_{IK} \cap \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}))) \\
&= \pi_I(\rho_{IK}(\widetilde{U}_{IK}) \cap \rho_{IJ} \circ \rho_{JK}(\rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}))) \\
&= \pi_I(U_{IK} \cap \rho_{IJ}(\widetilde{U}_{IJ} \cap U_{JK})) \\
&= \underline{U}_{IK} \cap (\underline{U}_{IJ} \cap \underline{\rho}_{IJ}(\underline{U}_{JK})) \\
&= \underline{U}_{IK} \cap (\underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK})),
\end{aligned}$$

where the second equality uses $\rho_{IK} = \rho_{IJ} \circ \rho_{JK}$ on $\widetilde{U}_{IK} \cap \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK})$, and the last uses $\underline{\rho}_{IJ} = \underline{\phi}_{IJ}^{-1}$. This proves (i).

Using in addition the identities $U_{\alpha\beta} = \pi_\alpha^{-1}(\underline{U}_{\alpha\beta})$ and $\widetilde{U}_{\alpha\beta} = \pi_\beta^{-1}(\underline{\phi}_{\alpha\beta}(\underline{U}_{\alpha\beta}))$, the cocycle condition (2.2.8) implies the inclusion claimed in (ii),

$$\begin{aligned}
\rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK}) &= (\pi_J \circ \rho_{JK})^{-1}(\underline{\phi}_{IJ}(\underline{U}_{IJ}) \cap \underline{U}_{JK}) \\
&= (\underline{\phi}_{IJ} \circ \underline{\rho}_{IK} \circ \pi_K)^{-1}(\underline{\phi}_{IJ}(\underline{U}_{IJ}) \cap \underline{U}_{JK}) \\
&= (\underline{\rho}_{IK} \circ \pi_K)^{-1}(\underline{U}_{IJ} \cap \underline{\phi}_{IJ}^{-1}(\underline{U}_{JK})) \subset \pi_K^{-1}(\underline{\rho}_{IK}^{-1}(\underline{U}_{IK})) \\
&= \widetilde{U}_{IK}.
\end{aligned}$$

The proof of (iii) is the same, with the strong cocycle condition implying equality in the second to last step. □

2.3 Kuranishi atlases

With the notions of Kuranishi charts and coordinate changes with nontrivial isotropy in place, we can now directly extend the notion of smooth Kuranishi atlas from [14, Definition 6.1.3]. For comparison with the notions of smooth and topological Kuranishi atlas from [13; 14], see Remark 2.2.1.

Definition 2.3.1 A *(weak) Kuranishi atlas of dimension d* on a compact metrizable space X is a tuple

$$\mathcal{K} = (\boldsymbol{K}_I, \widehat{\Phi}_{IJ})_{I,J \in \mathcal{I}_\mathcal{K}, I \subsetneq J}$$

consisting of a covering family of basic charts $(\boldsymbol{K}_i)_{i=1,\dots,N}$ of dimension d and transition data $(\boldsymbol{K}_J)_{|J| \geq 2}$, $(\widehat{\Phi}_{IJ})_{I \subsetneq J}$ for $(\boldsymbol{K}_i)_{i=1,\dots,N}$, where:

- A *covering family of basic charts* for X is a finite collection $(\boldsymbol{K}_i)_{i=1,\dots,N}$ of Kuranishi charts for X whose footprints cover $X = \bigcup_{i=1}^N F_i$.
- *Transition data* for a covering family $(\boldsymbol{K}_i)_{i=1,\dots,N}$ is a collection of Kuranishi charts $(\boldsymbol{K}_J)_{J \in \mathcal{I}_\mathcal{K}, |J| \geq 2}$ and coordinate changes $(\widehat{\Phi}_{IJ})_{I,J \in \mathcal{I}_\mathcal{K}, I \subsetneq J}$ as follows:

(i) $\mathcal{I}_{\mathcal{K}}$ denotes the set of subsets $I \subset \{1, \dots, N\}$ for which the intersection of footprints is nonempty,

$$F_I := \bigcap_{i \in I} F_i \neq \varnothing.$$

(ii) For each $J \in \mathcal{I}_{\mathcal{K}}$ with $|J| \geq 2$, $\boldsymbol{K}_J$ is a Kuranishi chart for X with footprint $F_J = \bigcap_{i \in J} F_i$, group $\Gamma_J = \prod_{j \in J} \Gamma_j$, and obstruction space $E_J = \prod_{j \in J} E_j$. Further, for one element sets $J = \{i\}$ we denote $\boldsymbol{K}_{\{i\}} := \boldsymbol{K}_i$.

(iii) $\widehat{\Phi}_{IJ} = (\rho_{IJ}, \rho^{\Gamma}_{IJ}, \widehat{\phi}_{IJ})$ is a coordinate change $\boldsymbol{K}_I \to \boldsymbol{K}_J$ for every $I, J \in \mathcal{I}_{\mathcal{K}}$ with $I \subsetneq J$, where $\rho^{\Gamma}_{IJ} \colon \Gamma_J \to \Gamma_I$ is the natural projection $\prod_{j \in J} \Gamma_j \to \prod_{i \in I} \Gamma_i$ and $\widehat{\phi}_{IJ} \colon E_I \to E_J$ is the natural inclusion $\prod_{i \in I} E_j \to \prod_{j \in J} E_j$.

For a weak atlas we require that the weak cocycle condition in Definition 2.2.12 hold for every triple $I, J, K \in \mathcal{I}_K$ with $I \subsetneq J \subsetneq K$, while for an atlas the cocycle condition must hold for all such triples.

Remark 2.3.2 Note that we have built *additivity* in the sense of [14, Definition 6.1.5] into the above definitions. Namely, the natural embeddings $\widehat{\phi}_{iI} \colon E_i \to E_I = \prod_{\ell \in I} E_\ell$ for each $I \in \mathcal{I}_{\mathcal{K}}$ induce the identity isomorphism

$$\prod_{i \in I} \widehat{\phi}_{iI} \colon \prod_{i \in I} E_i \xrightarrow{\cong} E_I = \prod_{\ell \in I} E_\ell, \tag{2.3.1}$$

and for $I \subset J$ the linear map $\widehat{\phi}_{IJ} \colon E_I \to E_J$ is the induced inclusion $\prod_{i \in I} E_i \to \prod_{i \in J} E_i$. Further, each group Γ_I is the product $\prod_{i \in I} \Gamma_i$ and we use the natural projections $\rho^{\Gamma}_{IJ} \colon \Gamma_J \to \Gamma_I$ in the group covering maps of the coordinate changes. Hence, when $I \subset J \subset K$ the projections $\rho^{\Gamma}_{\bullet\bullet}$ and linear inclusions $\widehat{\phi}_{\bullet\bullet}$ are automatically compatible:

$$\rho^{\Gamma}_{IK} = \rho^{\Gamma}_{IJ} \circ \rho^{\Gamma}_{JK}, \quad \widehat{\phi}_{IK} = \widehat{\phi}_{JK} \circ \widehat{\phi}_{IJ}.$$

So when $I \subset J$ we will almost always write $E_I \subset E_J$ for the subspace $\widehat{\phi}_{IJ}(E_I) \subset E_J$, and similarly we have a natural identification of Γ_J with $\Gamma_I \times \Gamma_{J \setminus I}$. ◇

Remark 2.3.3 Although it seems that many interdependent choices are needed in order to construct a Kuranishi atlas, this is somewhat deceptive. For example, in the Gromov–Witten case considered in [10] (see also [10]), the geometric choices involved in the construction of a family of basic charts $(\boldsymbol{K}_i)_{i=1,\dots,N}$ essentially induce the transition data as well. Namely, each basic chart $\boldsymbol{K}_i$ is constructed by adding a certain tuple $\vec{w}_i$ of marked points to the domains of the stable maps (f, z), given by the preimages of a fixed hypersurface of M in a fixed set of disjoint disks. The group Γ_i acts by permuting these disks, which has a rather nontrivial effect when viewing the chart in a local slice, in which the first three marked points are fixed. However, the transition charts $\boldsymbol{K}_J$ are constructed very similarly: Elements of the domain U_J consist of stable maps $(f, \mathbf{z})$ together with $|J|$ sets of added tuples of marked points $(\vec{w}_j)_{j\in J}$, each lying in an appropriate set of disks and mapping to certain hypersurfaces. Each factor Γ_j of the group Γ_J acts by permuting the components of the j^{th} set of disks, leaving the others alone. Moreover, the covering map $\widetilde{U}_{IJ} \to U_I$ simply forgets the extra tuples $(\vec{w}_j)_{j\in J\setminus I}$. Thus it is immediate from the construction that the group $\Gamma_{J\setminus I}$ acts freely on the subset $\widetilde{U}_{IJ}$ of U_J, and that the covering map is equivariant in the appropriate sense. Further, when $I \subset J \subset K$ the compatibility condition $\rho_{IK} = \rho_{IJ} \circ \rho_{JK}$ holds whenever both sides are defined.

Furthermore, the stabilization process explained in [10] (see also [13, Remark 6.1.6]) allows us to directly work with products of obstruction spaces $E_I := \prod_{i\in I} E_i$; there is no need for a transversality requirement such as Sum Condition II$'$ in [13, Section 4.3]. In fact, already each E_i is a product of the form $E_i = \prod_{\gamma\in\Gamma_i} (E_{0i})_\gamma$, on which Γ_i acts by permutation of the $|\Gamma_i|$ copies of a vector space E_{0i}. Therefore, just as in the case with no isotropy, once given the geometric choices that determine the basic charts, we naturally obtain an additive weak Kuranishi atlas in which the only new choices are those of the domains $\underline{U}_I = \underline{U}_{II}$ and $\underline{U}_{IJ}$ of the transition charts and coordinate changes which are required to intersect the zero set $\underline{\mathfrak{s}}_I^{-1}(0)$ in $\underline{\psi}_I^{-1}(F_J)$. Note that there is no simple hierarchy by which one could organize these choices to automatically fulfill the cocycle condition. Hence concrete constructions will usually only satisfy a weak cocycle condition. However, we show below that any weak (automatically additive) atlas can be "tamed" so that it satisfies the strong cocycle condition, and hence in particular gives a Kuranishi atlas. $\diamond$

Given a (weak) atlas $\mathcal{K} = (\boldsymbol{K}_I, \widehat{\Phi}_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}},\, I\subsetneq J}$, we define the associated *intermediate atlas* $\underline{\mathcal{K}} := (\underline{\boldsymbol{K}}_I, \underline{\widehat{\Phi}}_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}},\, I\subsetneq J}$ to consist of the intermediate charts and coordinate changes. The next lemma shows that the intermediate atlas is a (weak) topological atlas in the sense of [13, Definition 3.1.1], and that it is *filtered* in the sense that there are closed sets $\underline{\mathbb{E}}_{IJ} \subset \underline{\mathbb{E}}_J := \underline{U_J \times E_J}$ for each $I \subset J$ that satisfy the following conditions (cf [13, Definition 3.1.3]):

(i) $\underline{\mathbb{E}}_{JJ} = \underline{\mathbb{E}}_J$ and $\underline{\mathbb{E}}_{\varnothing J} = \operatorname{im} 0_J$ for all $J \in \mathcal{I}_{\mathcal{K}}$;

(ii) $\underline{\widehat{\Phi}}_{JK}(\mathrm{pr}_J^{-1}(\underline{U}_{JK}) \cap \underline{\mathbb{E}}_{IJ}) = \underline{\mathbb{E}}_{IK} \cap \mathrm{pr}_K^{-1}(\operatorname{im} \underline{\phi}_{JK})$ for all $I, J, K \in \mathcal{I}_{\mathcal{K}}$ with $I \subset J \subsetneq K$;

(iii) $\underline{\mathbb{E}}_{IJ} \cap \underline{\mathbb{E}}_{HJ} = \underline{\mathbb{E}}_{(I\cap H)J}$ for all $I, H, J \in \mathcal{I}_{\mathcal{K}}$ with $I, H \subset J$;

(iv) $\operatorname{im} \underline{\phi}_{IJ}$ is an open subset of $\underline{\mathfrak{s}}_J^{-1}(\underline{\mathbb{E}}_{IJ})$ for all $I, J \in \mathcal{I}_{\mathcal{K}}$ with $I \subsetneq J$.

Lemma 2.3.4 *Let $\mathcal{K}$ be a weak Kuranishi atlas. Then the intermediate atlas $\underline{\mathcal{K}}$ is a filtered weak topological Kuranishi atlas, with filtration $\underline{\mathbb{E}}_{IJ} := \underline{U_J \times \widehat{\phi}_{IJ}(E_I)}$, using the conventions $E_\varnothing := \{0\}$ and $\widehat{\phi}_{JJ} := \mathrm{id}_{E_J}$.*

Proof Lemma 2.2.4 and Remark 2.2.9(iii) assert that $\underline{\mathcal{K}}$ consists of topological Kuranishi charts and coordinate changes. The intermediate basic charts cover X since they have the same footprints as the basic charts of $\mathcal{K}$, and this also implies that the intermediate transition charts have the prescribed footprints. Moreover, the weak cocycle condition for $\mathcal{K}$ transfers to $\underline{\mathcal{K}}$ by Lemma 2.2.13(i), and the same holds for the cocycle condition since its definition (2.2.8) is in terms of the intermediate domains.

Next, to see that $\underline{\mathbb{E}}_{IJ}$ defines a filtration on $\underline{\mathcal{K}}$, we need a mild generalization of [14, Lemma 6.3.1]. First note that $\underline{U_J \times \widehat{\phi}_{IJ}(E_I)} \subset \underline{U_J \times E_J}$ is closed since $U_J \times \widehat{\phi}_{IJ}(E_I) \subset U_J \times E_J$ is closed and the projection $U_J \times E_J \to \underline{U_J \times E_J}$ is a closed map by Lemma 2.1.5(i). The filtration property (i) above holds by definition, and property (iii) holds because additivity implies

$$\widehat{\phi}_{IJ}(E_I) \cap \widehat{\phi}_{HJ}(E_H) = \widehat{\phi}_{(I\cap H)J}(E_{I\cap H}).$$

Moreover, because $\underline{\widehat{\Phi}}_{JK} = \underline{\phi_{JK} \times \widehat{\phi}_{JK}}$, property (ii) follows by quotienting the next identity by the group Γ_K,

$$\begin{aligned}\widehat{\Phi}_{JK}(U_{JK} \times \widehat{\phi}_{IJ}(E_I)) &= \operatorname{im}\phi_{JK} \times \widehat{\phi}_{JK}(\widehat{\phi}_{IJ}(E_I)) \\ &= \operatorname{im}\phi_{JK} \times \widehat{\phi}_{IK}(E_I) \\ &= (U_K \times \widehat{\phi}_{IK}(E_I)) \cap (\operatorname{im}\phi_{JK} \times E_K).\end{aligned}$$

Finally, to check property (iv) we first apply [14, Lemma 5.2.5] to the embedding $\widetilde{\phi}_{IJ}\colon \widetilde{U}_{IJ} \to U_J$, which satisfies the index condition, ie identifies kernel and cokernel of $\mathrm{d}s_J$ and $\mathrm{d}s_I$ (the latter being pulled back with the covering ρ_{IJ}). It implies that $\operatorname{im}\widetilde{\phi}_{IJ}$ is an open subset of $s_J^{-1}(E_I)$. This openness is preserved in the Γ_J quotient, since Lemma 2.1.5 applies to the projection

$$s_J^{-1}(E_I) \to s_J^{-1}(E_I)/\Gamma_J = \mathfrak{s}_J^{-1}(\underline{U_J \times E_I}) = \mathfrak{s}_J^{-1}(\underline{\mathbb{E}}_{IJ}),$$

which maps $\operatorname{im}\phi_{IJ}$ to $\operatorname{im}\underline{\phi}_{IJ}$. □

If $\mathcal{K}$ is a Kuranishi atlas, then the topological atlas $\underline{\mathcal{K}}$ also satisfies the cocycle conditions, and hence by [13, Lemma 2.3.7] there is an *intermediate domain category* $\boldsymbol{B}_{\underline{\mathcal{K}}}$ with objects $\mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} := \bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}} \underline{U}_I$ equal to the disjoint union of the intermediate domains, and morphisms

$$\mathrm{Mor}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} := \bigsqcup_{I\subset J} \underline{U}_{IJ}$$

given by the intermediate coordinate changes $\underline{\phi}_{IJ}\colon \underline{U}_{IJ} \to \underline{U}_J$, where the identity maps $\underline{\phi}_{II}$ on $\underline{U}_{II} = \underline{U}_I$ are included. Thus the source and target maps are

$$s \times t\colon\ \underline{U}_{IJ} \to \underline{U}_I \times \underline{U}_J \subset \mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} \times \mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}}, \quad (I,x) \mapsto ((I,x),(J,\underline{\phi}_{IJ}(x))).$$

The following gives the analogous categorical interpretation for the Kuranishi atlas itself.

Definition 2.3.5 Given a Kuranishi atlas $\mathcal{K}$ we define its *domain category* $\boldsymbol{B}_{\mathcal{K}}$ to consist of the space of objects

$$\mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}} := \bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} U_I = \{(I,x) \mid I \in \mathcal{I}_{\mathcal{K}}, x \in U_I\}$$

and the space of morphisms

$$\mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}} := \bigsqcup_{I,J \in \mathcal{I}_{\mathcal{K}},\, I \subset J} \widetilde{U}_{IJ} \times \Gamma_I = \{(I,J,y,\gamma) \mid I \subset J, y \in \widetilde{U}_{IJ}, \gamma \in \Gamma_I\}.$$

Here we denote $\widetilde{U}_{II} := U_I$ for $I = J$, and for $I \subsetneq J$ use the lifted domain $\widetilde{U}_{IJ} \subset U_J$ of the restriction $\boldsymbol{K}_I|_{\underline{U}_{IJ}}$ to F_J that is part of the coordinate change $\widehat{\Phi}_{IJ}\colon \boldsymbol{K}_I|_{\underline{U}_{IJ}} \to \boldsymbol{K}_J$. Source and target of these morphisms are given by

$$(I,J,y,\gamma) \in \mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}}\big((I,\gamma^{-1}\rho_{IJ}(y)),(J,\widetilde{\phi}_{IJ}(y))\big), \tag{2.3.2}$$

where we denote $\widetilde{\phi}_{II} = \mathrm{id}$. Composition[4] is defined by

$$(I,J,y,\gamma) \circ (J,K,z,\delta) := (I,K,z{=}\widetilde{\phi}_{IK}^{-1}(\widetilde{\phi}_{JK}(z)), \rho_{IJ}^{\Gamma}(\delta)\gamma)$$

whenever $\delta^{-1}\rho_{JK}(z) = \widetilde{\phi}_{IJ}(y)$.

The *obstruction category* $\boldsymbol{E}_{\mathcal{K}}$ is defined in complete analogy to $\boldsymbol{B}_{\mathcal{K}}$ to consist of the spaces of objects $\mathrm{Obj}_{\boldsymbol{E}_{\mathcal{K}}} := \bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} U_I \times E_I$ and morphisms

$$\mathrm{Mor}_{\boldsymbol{E}_{\mathcal{K}}} := \bigsqcup_{I \subset J,\, I,J \in \mathcal{I}_{\mathcal{K}}} \widetilde{U}_{IJ} \times E_I \times \Gamma_I,$$

with source and target maps

$$(I,J,y,e,\gamma) \mapsto (I,\gamma^{-1}\rho_{IJ}(y),\gamma^{-1}e), \quad (I,J,y,e,\gamma) \mapsto (J,\widetilde{\phi}_{IJ}(y),\widehat{\phi}_{IJ}(e)),$$

[4]Note that we write compositions in the categorical ordering here. Recall that $\widetilde{\phi}_{JK}\colon \widetilde{U}_{JK} \to U_K$ is the canonical inclusion of the subset $\widetilde{U}_{JK} \subset U_K$. We then identify $z = \widetilde{\phi}_{IK}^{-1}(\widetilde{\phi}_{JK}(z))$, since composability of the morphisms implies $z \in \rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK})$ and the cocycle condition ensures that $\rho_{JK}^{-1}(\widetilde{U}_{IJ} \cap U_{JK})$ is contained in $\widetilde{U}_{IK}$, where both are considered as subsets of U_K.

and composition defined by

$$(I,J,y,e,\gamma)\circ(J,K,z,f,\delta) := \bigl(I,K,\widetilde{\phi}_{IK}^{-1}(\widetilde{\phi}_{JK}(z)), f, \rho^{\Gamma}_{IJ}(\delta)\gamma\bigr)$$

for any $I\subset J\subset K$ and $(y,e,\gamma)\in\widetilde{U}_{IJ}\times E_I\times\Gamma_I$, $(z,f,\delta)\in\widetilde{U}_{JK}\times E_J\times\Gamma_J$ such that $\rho^{\Gamma}_{IJ}(\delta^{-1})\rho_{JK}(z)=\widetilde{\phi}_{IJ}(y)$ and $\delta^{-1}f=e$.

Lemma 2.3.6 *If $\mathcal{K}$ is a Kuranishi atlas, then the categories $\boldsymbol{B}_{\mathcal{K}}$, $\boldsymbol{E}_{\mathcal{K}}$ are well defined.*

Proof We must check that the composition of morphisms in $\boldsymbol{B}_{\mathcal{K}}$ is well defined, has identities, and is associative; the proof for $\boldsymbol{E}_{\mathcal{K}}$ is analogous. We begin by checking that $z=\widetilde{\phi}_{IK}^{-1}(\widetilde{\phi}_{JK}(z))$ lies in the lifted domain $\widetilde{U}_{IK}$ of $\widehat{\Phi}_{IK}$. For that purpose we drop the natural inclusions $\widetilde{\phi}_{**}$ from the notation and note that the composition $(I,J,y,\gamma)\circ(J,K,z,\delta)$ is defined only when the target of (I,J,y,γ) equals the source of (J,K,z,δ); ie when $y=\delta^{-1}\rho_{JK}(z)$. So the cocycle condition in Lemma 2.2.13(ii) implies that $z\in\rho_{JK}^{-1}(\delta y)$ is contained in $\rho_{JK}^{-1}(\widetilde{U}_{IJ}\cap U_{JK})\subset\widetilde{U}_{IK}$, as claimed. This means that $(I,K,z,\rho^{\Gamma}_{IJ}(\delta)\gamma)$ is a well-defined morphism of $\boldsymbol{B}_{\mathcal{K}}$. Its source is

$$(\rho^{\Gamma}_{IJ}(\delta)\gamma)^{-1}\rho_{IK}(z)=\gamma^{-1}\rho^{\Gamma}_{IJ}(\delta)^{-1}\rho_{IJ}(\delta y)=\gamma^{-1}\rho_{IJ}(y),$$

which coincides with the source of (I,J,y,γ) as required. Finally, the target of the composed morphism, $z=\widetilde{\phi}_{IK}(\widetilde{\phi}_{IK}^{-1}(\widetilde{\phi}_{JK}(z))$ coincides with the target $\widetilde{\phi}_{JK}(z)$ of (J,K,z,δ). This shows that composition is well defined. The identity morphisms are given by (I,I,x,id) for all $x\in U_{II}:=U_I$. To check associativity we consider $I\subset J\subset K\subset L$ and suppose that the three morphisms (I,J,y,γ), (J,K,z,δ), (K,L,w,σ) are composable. Then we have

$$\begin{aligned}(I,J,y,\gamma)\circ\bigl((J,K,z,\delta)\circ(K,L,w,\sigma)\bigr)&=(I,J,y,\gamma)\circ(J,L,w,\rho^{\Gamma}_{JK}(\sigma)\delta)\\&=(I,L,w,\rho^{\Gamma}_{IJ}(\rho^{\Gamma}_{JK}(\sigma)\delta)\gamma),\end{aligned}$$

and associativity follows from comparing this expression with

$$\begin{aligned}\bigl((I,J,y,\gamma)\circ(J,K,z,\delta)\bigr)\circ(K,L,w,\sigma)&=(I,K,z,\rho^{\Gamma}_{IJ}(\delta)\gamma)\circ(K,L,w,\sigma)\\&=(I,L,w,\rho^{\Gamma}_{IK}(\sigma)\rho^{\Gamma}_{IJ}(\delta)\gamma).\end{aligned}$$

This completes the proof. □

For the rest of this subsection we will make the standing assumption that $\mathcal{K}$ is a Kuranishi atlas, ie satisfies the cocycle condition (not just the weak cocycle condition). Given the categorical interpretation of domains and obstruction spaces of Kuranishi charts, we can now express the bundles, sections and footprint maps as functors:

- The obstruction category $\boldsymbol{E}_{\mathcal{K}}$ is a bundle over $\boldsymbol{B}_{\mathcal{K}}$ in the sense that there is a functor $\mathrm{pr}_{\mathcal{K}}\colon \boldsymbol{E}_{\mathcal{K}} \to \boldsymbol{B}_{\mathcal{K}}$ that is given on objects and morphisms by projection $(I,x,e) \mapsto (I,x)$ and $(I,J,y,e,\gamma) \mapsto (I,J,y,\gamma)$.
- The sections s_I induce a smooth section of this bundle, ie a functor $\mathfrak{s}_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$ which acts smoothly on the spaces of objects and morphisms, and whose composite with the projection $\mathrm{pr}_{\mathcal{K}}\colon \boldsymbol{E}_{\mathcal{K}} \to \boldsymbol{B}_{\mathcal{K}}$ is the identity. More precisely, $\mathfrak{s}_{\mathcal{K}}$ is given by $(I,x) \mapsto (I,x,s_I(x))$ on objects and by $(I,J,y,\gamma) \mapsto (I,J,y,s_I(y),\gamma)$ on morphisms.
- The zero sections also fit together to give a functor $0_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$ given by $(I,x) \mapsto (I,x,0)$ on objects and by $(I,J,y,\gamma) \mapsto (I,J,y,0,\gamma)$ on morphisms.
- The footprint maps ψ_I induce a surjective functor

  $$\psi_{\mathcal{K}}\colon \mathfrak{s}_{\mathcal{K}}^{-1}(0) := \bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} s_I^{-1}(0) \to \boldsymbol{X}$$

  to the category $\boldsymbol{X}$ with object space X and trivial morphism spaces. It is given by $(I,x) \mapsto \psi_I(x)$ on objects and by $(I,J,y,\gamma) \mapsto \mathrm{id}_{\psi_J(\tilde{\phi}_{IJ}(y))} = \mathrm{id}_{\psi_I(\gamma^{-1}\rho_{IJ}(y))}$ on morphisms.

As in [13] we denote by $|\mathcal{K}|$ (resp. $|\underline{\mathcal{K}}|$) the *realization* of the category $\boldsymbol{B}_{\mathcal{K}}$ (resp. $\boldsymbol{B}_{\underline{\mathcal{K}}}$). This is the topological space obtained as the quotient of the object space by the equivalence relation generated by the morphisms. The next lemma fits the quotient maps $\pi_{\mathcal{K}}\colon \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}} \to |\mathcal{K}|$, $(I,x) \mapsto [I,x]$ and $\pi_{\underline{\mathcal{K}}}\colon \mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} \to |\underline{\mathcal{K}}|$, $(I,\underline{x}) \mapsto [I,\underline{x}]$ into a commutative diagram that will allow us to identify the realizations $|\mathcal{K}| \cong |\underline{\mathcal{K}}|$ as topological spaces.

Lemma 2.3.7 *If $\mathcal{K}$ is a Kuranishi atlas, then there is a functor $\rho_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{B}_{\underline{\mathcal{K}}}$ that is given on objects by the quotient maps $U_I \to \underline{U}_I$, $x \mapsto \underline{x}$, and on morphisms by the group coverings ρ_{IJ} together with a quotient,*

$$\tilde{U}_{IJ} \times \Gamma_I \to \underline{U}_{IJ}, \quad (I,J,y,\gamma) \mapsto (I,J,\underline{\rho_{IJ}(y)}).$$

It induces a homeomorphism $|\rho_{\mathcal{K}}|\colon |\mathcal{K}| \to |\underline{\mathcal{K}}|$ between the realizations that fits into a commutative diagram:

$$\begin{array}{ccc} \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}} & \xrightarrow{\rho_{\mathcal{K}}} & \mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} \\ \downarrow{\scriptstyle \pi_{\mathcal{K}}} & & \downarrow{\scriptstyle \pi_{\underline{\mathcal{K}}}} \\ |\mathcal{K}| & \xrightarrow{|\rho_{\mathcal{K}}|} & |\underline{\mathcal{K}}| \end{array}$$

Proof To see that $\rho_{\mathcal{K}}$ is a functor, recall that $(y,\gamma) \in \tilde{U}_{IJ} \times \Gamma_I$ represents a morphism from $\gamma^{-1}\rho_{IJ}(y)$ to $y \in U_J$. On the other hand, $\underline{\rho_{IJ}(y)} = \underline{\rho}_{IJ}(\underline{y}) \in \underline{U}_{IJ}$ represents a morphism from $\underline{\rho_{IJ}(y)} = \underline{\gamma^{-1}\rho_{IJ}(y)}$ to $\underline{\phi}_{IJ}(\underline{\rho_{IJ}(y)}) = \underline{y}$, which shows compatibility of $\rho_{\mathcal{K}}$ with source and target maps. Compatibility with composition as in (2.3.2) follows from $\underline{\rho}_{IK}(\underline{z}) = \underline{\rho}_{IJ}(\underline{y})$ when $\underline{y} = \underline{\rho}_{JK}(\underline{z})$.

Next, any functor such as $\rho_{\mathcal{K}}$ induces a map $|\rho_{\mathcal{K}}|$ between the realizations that is defined exactly by the above commutative diagram. The map $|\rho_{\mathcal{K}}|$ is surjective because the functor $\rho_{\mathcal{K}}$ is surjective on the level of objects. It is injective because $\rho_{\mathcal{K}}$ is surjective on the level of morphisms.

To check that $|\rho_{\mathcal{K}}|$ is open and continuous note that $|\rho_{\mathcal{K}}|(U) = V$ is equivalent to $\rho_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}^{-1}(U)) = \pi_{\underline{\mathcal{K}}}^{-1}(V)$. Since $\rho_{\mathcal{K}}$ is continuous and open by Lemma 2.1.5(i), and $|\mathcal{K}|$, $|\underline{\mathcal{K}}|$ are equipped with the quotient topologies, the openness of $U \subset |\mathcal{K}|$, $\pi_{\mathcal{K}}^{-1}(U)$, $\pi_{\underline{\mathcal{K}}}^{-1}(V)$ and $V \subset |\underline{\mathcal{K}}|$ are all equivalent. This proves that $|\rho_{\mathcal{K}}|$ is a homeomorphism. $\square$

Remark 2.3.8 (i) If $\mathcal{K}$ is a Kuranishi atlas with trivial isotropy groups $\Gamma_I = \{\mathrm{id}\}$, then the intermediate atlas $\underline{\mathcal{K}}$ has the exact same object space and naturally diffeomorphic morphism spaces, only the direction of the maps in the coordinate changes are reversed from $\rho_{IJ}\colon \widetilde{U}_{IJ} \to U_{IJ} \subset U_I$ to $\underline{\phi}_{IJ} = \rho_{IJ}^{-1}\colon U_{IJ} \to \widetilde{U}_{IJ} \subset U_J$. In this special case, $\underline{\mathcal{K}}$ is a Kuranishi atlas in the sense of [14], and Lemma 2.3.7 identifies the atlases and their realizations.

(ii) In general, the spaces of objects and morphisms of the intermediate category are orbifolds, and there is at most one morphism between any pair of objects. However, just as in the case of trivial isotropy, we do not attempt to make this category into a groupoid by formally inverting the morphisms and then adding all resulting composites, since doing so would in general give components of the morphism space without orbifold structure; cf [14, Remark 6.1.8]. This objection does not apply if all the obstruction spaces are trivial. It is shown in [10; 11] that every such atlas can be completed to a groupoid without changing its realization. $\diamond$

In complete analogy to Lemma 2.3.7, the obstruction categories $\boldsymbol{E}_{\mathcal{K}}$ and $\boldsymbol{E}_{\underline{\mathcal{K}}}$ of the Kuranishi atlas $\mathcal{K}$ and the intermediate atlas $\underline{\mathcal{K}}$ also fit into a commutative diagram that identifies their realizations $|\boldsymbol{E}_{\mathcal{K}}| \cong |\boldsymbol{E}_{\underline{\mathcal{K}}}|$. Moreover, these two diagrams also intertwine the section functors $\mathfrak{s}_{\mathcal{K}}$, $\mathfrak{s}_{\underline{\mathcal{K}}}$ and their realizations:

(2.3.3)
$$
\begin{tikzcd}
{} & \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}} \arrow[l, "\rho_{\mathcal{K}}"'] \arrow[r, "\mathfrak{s}_{\mathcal{K}}"] \arrow[d, "\pi_{\mathcal{K}}"] & \mathrm{Obj}_{\boldsymbol{E}_{\mathcal{K}}} \arrow[r] \arrow[d, "\pi_{\boldsymbol{E}_{\mathcal{K}}}"] & \mathrm{Obj}_{\boldsymbol{E}_{\underline{\mathcal{K}}}} \arrow[d, "\pi_{\boldsymbol{E}_{\underline{\mathcal{K}}}}"] & \mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}} \arrow[l, "\mathfrak{s}_{\underline{\mathcal{K}}}"'] \arrow[d, "\pi_{\underline{\mathcal{K}}}"] & {} \arrow[l, "\rho_{\mathcal{K}}"'] \\
{} & |\mathcal{K}| \arrow[l, "|\rho_{\mathcal{K}}|"'] \arrow[r, "|\mathfrak{s}_{\mathcal{K}}|"] & |\boldsymbol{E}_{\mathcal{K}}| \arrow[r] & |\boldsymbol{E}_{\underline{\mathcal{K}}}| & |\underline{\mathcal{K}}| \arrow[l, "|\mathfrak{s}_{\underline{\mathcal{K}}}|"'] & {} \arrow[l, "|\rho_{\mathcal{K}}|"']
\end{tikzcd}
$$

There are analogous diagrams for the projection functors $\mathrm{pr}_{\mathcal{K}}$, $\mathrm{pr}_{\underline{\mathcal{K}}}$ and zero sections $0_{\mathcal{K}}$ and $0_{\underline{\mathcal{K}}}$, which identify the induced maps between the realizations as stated below.

Lemma 2.3.9 *Let $\mathcal{K}$ be a Kuranishi atlas.*

(i) *The functors $\mathrm{pr}_{\mathcal{K}}\colon \boldsymbol{E}_{\mathcal{K}} \to \boldsymbol{B}_{\mathcal{K}}$, $\mathrm{pr}_{\underline{\mathcal{K}}}\colon \underline{\boldsymbol{E}}_{\mathcal{K}} \to \underline{\boldsymbol{B}}_{\mathcal{K}}$ induce the same continuous map*

$$|\mathrm{pr}_{\mathcal{K}}|\colon |\boldsymbol{E}_{\mathcal{K}}| \to |\mathcal{K}|,$$

which we call the ***obstruction bundle*** *of $\mathcal{K}$, although its fibers generally do not have the structure of a vector space.*

(ii) *The zero sections $0_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$, $0_{\underline{\mathcal{K}}}\colon \boldsymbol{B}_{\underline{\mathcal{K}}} \to \boldsymbol{E}_{\underline{\mathcal{K}}}$ as well as the section functors $\mathfrak{s}_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$, $\mathfrak{s}_{\underline{\mathcal{K}}}\colon \boldsymbol{B}_{\underline{\mathcal{K}}} \to \boldsymbol{E}_{\underline{\mathcal{K}}}$ induce the same continuous maps*

$$|0_{\mathcal{K}}| \cong |0_{\underline{\mathcal{K}}}|\colon |\mathcal{K}| \to |\boldsymbol{E}_{\mathcal{K}}|, \quad |\mathfrak{s}_{\mathcal{K}}| \cong |\mathfrak{s}_{\underline{\mathcal{K}}}|\colon |\mathcal{K}| \to |\boldsymbol{E}_{\mathcal{K}}|,$$

which are sections in the sense that $|\mathrm{pr}_{\mathcal{K}}| \circ |0_{\mathcal{K}}| = \mathrm{id}_{|\mathcal{K}|} = |\mathrm{pr}_{\mathcal{K}}| \circ |\mathfrak{s}_{\mathcal{K}}|$.

(iii) *There is a natural homeomorphism from the realization of the subcategory $\mathfrak{s}_{\mathcal{K}}^{-1}(0)$ to the zero set of $|\mathfrak{s}_{\mathcal{K}}|$, with the relative topology induced from $|\mathcal{K}|$,*

$$|\mathfrak{s}_{\mathcal{K}}^{-1}(0)| = \mathfrak{s}_{\mathcal{K}}^{-1}(0)/\!\sim \;\xrightarrow{\cong}\; |\mathfrak{s}_{\mathcal{K}}|^{-1}(|0_{\mathcal{K}}|) := \bigl\{[I,x] \bigm| |\mathfrak{s}_{\mathcal{K}}|([I,x]) = |0_{\mathcal{K}}|([I,x])\bigr\} \subset |\mathcal{K}|.$$

Proof The induced maps on the realizations are identified by commutative diagrams such as (2.3.3). The continuity and other identities are proven exactly as in [14, Lemma 6.1.10] for the case of trivial isotropy. □

Next, we extend the notion of metrizability to Kuranishi atlases with nontrivial isotropy. In the case of trivial isotropy, recall from [14, Definition 6.1.14] that an admissible metric is a bounded metric d on the set $|\mathcal{K}|$ such that for each $I \in \mathcal{I}_{\mathcal{K}}$ the pullback metric $d_I := (\pi_{\mathcal{K}}|_{U_I})^* d$ on U_I induces the given topology on the manifold U_I. However, in the presence of isotropy, it makes no sense to try to pull this metric back to U_I since the pullback of a metric by a noninjective map is no longer a metric. Instead, we use the fact that the realizations $|\mathcal{K}| \cong |\underline{\mathcal{K}}|$ of the Kuranishi atlas and its intermediate atlas are canonically identified, which allows us to work with admissible metrics on $|\underline{\mathcal{K}}|$, which is the realization of a topological Kuranishi atlas $\underline{\mathcal{K}}$ with trivial isotropy and given metrizable topologies on the domains $\underline{U}_I = U_I/\Gamma_I$.

Definition 2.3.10 Let $\mathcal{K}$ be a Kuranishi atlas. Then an *admissible metric* on $|\mathcal{K}| \cong |\underline{\mathcal{K}}|$ is a bounded metric on this set (not necessarily compatible with the topology of the realization) such that for each $I \in \mathcal{I}_{\mathcal{K}}$ the pullback metric $\underline{d}_I := (\pi_{\underline{\mathcal{K}}}|_{\underline{U}_I})^* d$ on $\underline{U}_I$ induces the given quotient topology on $\underline{U}_I = U_I/\Gamma_I$.

A *metric Kuranishi atlas* is a pair $(\mathcal{K}, d)$ consisting of a Kuranishi atlas $\mathcal{K}$ together with a choice of admissible metric d on $|\mathcal{K}|$.

We finish this subsection with two comparisons of our notion of Kuranishi atlas: on the one hand with orbifolds, and on the other hand with Kuranishi structures.

Example 2.3.11 If the obstruction spaces are trivial, ie $E_I = \{0\}$ for all I, then the two categories $\boldsymbol{B}_{\mathcal{K}}$, $\boldsymbol{E}_{\mathcal{K}}$ are equal, and their realization is an orbifold. A first nontrivial example is a "football" $X = S^2$ with two basic Kuranishi charts

$$(U_1, \Gamma_1 = \mathbb{Z}_2, \psi_1), \quad (U_2, \Gamma_2 = \mathbb{Z}_3, \psi_2),$$

covering neighborhoods $\underline{\psi}_i(\underline{U}_i) \subset S^2$ of the northern (resp. southern) hemisphere with isotropy of order 2 (resp. 3) at the north (resp. south) pole. We may moreover assume that the overlap $\underline{\psi}_1(\underline{U}_1) \cap \underline{\psi}_2(\underline{U}_2) = \underline{A}$ is an annulus around the equator. The restrictions of the basic charts to $\underline{A} \subset X$ are $(A_1, \mathbb{Z}_2)$ and $(A_2, \mathbb{Z}_3)$, where both $A_i = \psi_i^{-1}(\underline{A})$ are annuli, but the freely acting isotropy groups are different. There is no functor between these restrictions because the coverings $A_1 \to \underline{A}$ and $A_2 \to \underline{A}$ are incompatible. However, they both have functors (ie coordinate changes) to a common free covering, namely the pullback defined by the diagram

$$\begin{array}{ccc} U_{12} & \longrightarrow & A_1 \\ \downarrow & & \downarrow \pi_1 \\ A_2 & \xrightarrow{\pi_2} & \underline{A} \subset X \end{array}$$

ie $U_{12} := \{(x, y) \in A_1 \times A_2 \mid \pi_1(x) = \pi_2(y)\}$ with group $\Gamma_{12} := \Gamma_1 \times \Gamma_2 = \mathbb{Z}_2 \times \mathbb{Z}_3$. The corresponding footprint map $\psi_{12} \colon U_{12} \to \underline{A}$ is the 6–fold covering of the annulus, and the coordinate changes from $(U_i, \Gamma_i, \psi_i)|_{\underline{A}}$ to $(U_{12}, \Gamma_{12}, \psi_{12})$ are the coverings $\widetilde{U}_{i,12} := U_{12} \to A_i =: U_{i,12}$ in the diagram. Therefore the category $\boldsymbol{B}_{\mathcal{K}}$ in this example has index set $\mathcal{I}_{\mathcal{K}} = \{1, 2, 12\}$, objects the disjoint union $\bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} U_I$, and morphisms

$$\Bigl(\bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} U_I \times \Gamma_I\Bigr) \cup \Bigl(\bigsqcup_{i=1,2} U_{12} \times \Gamma_i\Bigr),$$

where for $i = 1, 2$ the elements in $U_{12} \times \Gamma_i$ represent the morphisms from U_i to U_{12}.

This simple construction does not work for arbitrary orbifolds since the (set-theoretic) pullback U_{12} considered above will not be a smooth manifold if any point in $\psi_1(U_1) \cap \psi_2(U_2)$ has nontrivial stabilizer. However, we show in [11, Proposition 3.3] that the construction can be generalized to show that every orbifold has a Kuranishi atlas with trivial obstruction spaces. ◇

Remark 2.3.12 (relation to Kuranishi structures) A Kuranishi structure in the sense of [3, Appendix A] and [4] consists of a Kuranishi chart $\boldsymbol{K}_p$ at every point $p \in X$ and coordinate changes $\boldsymbol{K}_q|_{U_{qp}} \to \boldsymbol{K}_p$ whenever $q \in F_p$, that satisfy a suitable weak cocycle condition. Much as in the case of Kuranishi atlases with trivial isotropy (see [14, Remark 6.1.16]), a weak Kuranishi atlas in the sense of Definition 2.3.1 induces a Kuranishi structure. Indeed, given a covering family of basic charts $(\boldsymbol{K}_i)_{i=1,\dots,N}$ with footprints F_i, we may choose a family of compact subsets $C_i \subset F_i$ that also cover X. Then we use the transition data $(\boldsymbol{K}_I, \widehat{\Phi}_{IJ})$ and weak cocycle conditions to obtain a Kuranishi structure as follows:

- For any $p \in X$, we define $\boldsymbol{K}_p := \boldsymbol{K}_{I_p}|_{U_p}$ to be a restriction of $\boldsymbol{K}_{I_p}$, where $I_p := \{i \mid p \in C_i\}$ and $U_p \subset U_{I_p}$ is an open subset such that the footprint
$$F_p := \psi_{I_p}(s_{I_p}^{-1}(0) \cap U_p)$$
is a neighborhood of p and contained in $\bigcap_{i \in I_p} F_i \setminus \bigcup_{i \notin I_p} C_i$. Here we use a more general notion of restriction than Definition 2.2.6, in that we allow for a domain U_p that is invariant only under a subgroup $\Gamma_p \subset \Gamma_{I_p}$ such that the induced map $U_p/\Gamma_p \to U_{I_p}/\Gamma_{I_p}$ is a homeomorphism to its image. More precisely, to satisfy the minimality requirements of [3, Appendix A1.1], we choose a lift $x_p \in \pi^{-1}(p) \cap U_{I_p}$, set $\Gamma_p := \Gamma_{I_p}^{x_p}$ to be its stabilizer in Γ_{I_p}, and take the domain $U_p \subset U_{I_p}$ to be a $\Gamma_{I_p}^{x_p}$–invariant neighborhood of x_p, which exists with the required topological properties by Lemma 2.1.5(ii).

- For $q \in F_p$ we have $I_q \subset I_p$, since by construction $F_p \cap C_i = \varnothing$ for $i \notin I_p$. So we obtain a coordinate change[5] $\widehat{\Phi}_{qp}\colon \boldsymbol{K}_q \to \boldsymbol{K}_p$ from a suitable restriction of $\widehat{\Phi}_{I_q I_p}$ to a $\Gamma_q^{x_q}$–invariant neighborhood $U_{qp} \subset U_q$ of x_q. More precisely, we choose $U_{qp} \subset U_q$ small enough so that the projection $\rho_{I_q I_p}\colon U_p \cap \widetilde{U}_{I_q I_p} \to U_{I_q I_p}$ has a continuous section over U_{qp}. Writing $\widetilde{U}_{qp}$ for its image we thus obtain an embedding $\phi_{qp} := \rho_{I_q I_p}^{-1}\colon U_{qp} \to \widetilde{U}_{qp} \subset U_p \cap \widetilde{U}_{I_q I_p}$. Since the projection $\rho_{I_q I_p}$ induces an isomorphism on stabilizer subgroups by Lemma 2.1.5(iii), this is equivariant with respect to a suitable injective homomorphism $h_{qp}\colon \Gamma_q \to \Gamma_p$ and induces an injection
$$\underline{\phi}_{qp}\colon \underline{U}_{qp} := U_{qp}/\Gamma_q \to \underline{U}_p := U_p/\Gamma_p.$$
By construction of $\underline{U}_q \to \underline{U}_{I_q}$ above, the map $\underline{U}_{qp} = U_{qp}/\Gamma_q \to \underline{U}_{I_q} = U_{I_q}/\Gamma_{I_q}$ is a homeomorphism to its image, and similarly for p. Thus we can identify $\underline{\phi}_{qp}$ with a suitable restriction of the map $\underline{\phi}_{I_q I_p}$ underlying the coordinate change $\widehat{\Phi}_{I_q I_p}$ in the given Kuranishi atlas. The coordinate change $\widehat{\Phi}_{qp} = (U_{qp}, \underline{\phi}_{qp})$ is then given by the domain U_{qp} and the restriction of $\underline{\phi}_{I_q I_p}$ to $\underline{U}_{qp} \subset \underline{U}_q$.

Further, the weak cocycle condition for $\mathcal{K}$ implies the compatibility condition required by [3], namely for all triples $p, q, r \in X$ with $q \in F_p$ and
$$r \in \psi_q(U_{qp} \cap s_q^{-1}(0)) \subset F_q \cap F_p,$$

[5] While [3] denotes this coordinate change by ϕ_{pq}, we will write $\widehat{\Phi}_{qp}$ for consistency with our notation $\Phi_{IJ}\colon \boldsymbol{K}_I \to \boldsymbol{K}_J$.

the equality $\underline{\phi}_{qp} \circ \underline{\phi}_{rq} = \underline{\phi}_{rp}$ holds on the common domain $\underline{\phi}_{rq}^{-1}(\underline{U}_{qp}) \cap \underline{U}_{rp}$ of the maps in this equation.

- This atlas satisfies the effectivity condition required by [3] only if the action of Γ_p on U_p is locally effective in the sense that $s_p^{-1}(0)$ has a Γ_p–invariant open neighborhood that is disjoint from the interior of the fixed point set $\mathrm{Fix}(\gamma) \subset U_p$ for each $\gamma \in \Gamma_p \setminus \{\mathrm{id}\}$.

With this construction, we lose the distinction between basic charts and transition charts, and also in general can no longer recover the original transition charts with their group actions from the Kuranishi structure. Indeed, [4] works with a "good coordinate system" (an analog of our notion of reduction in Definition 3.2.1) that is defined on the orbifold level, ie on the level of the intermediate category. Notice also that the construction of a Kuranishi structure given for example in [4] essentially follows the above outline, and in particular starts with a finite covering family of basic charts and uses transition charts much like ours, though they are more localized and are not required to cover the full footprint F_I. However, the properties of these charts are never explicitly formulated. Indeed our work started by trying to understand precisely this point in their construction. Though it is not clear how relevant the extra information contained in a Kuranishi atlas is to the question of how to define Gromov–Witten invariants for closed curves, it might prove useful in other situations, for example in the case of orbifold Gromov–Witten invariants, or in the recent work of Fukaya et al [5], where the authors consider a process that rebuilds a Kuranishi structure from a coordinate system. Further, our categorical formulation makes it very easy to give an explicit description and construction for sections as in Definition 3.2.4. ◇

2.4 Kuranishi cobordisms and concordance

This section extends the notions of cobordism and concordance developed in [13, Section 4] and [14, Section 6.2] to the case of smooth Kuranishi atlases with nontrivial isotropy. It is a straightforward generalization that can be skipped until precise concordance notions are needed in the proof of Theorem 2.5.3. We begin by summarizing the topological cobordism notions from [13, Section 4.1].

A *collared cobordism* $(Y, \iota_Y^0, \iota_Y^1)$ is a separable, locally compact, metrizable space Y together with disjoint (possibly empty) closed subsets $\partial^0 Y$, $\partial^1 Y \subset Y$ and equipped with *collared neighborhoods*

$$\iota_Y^0\colon [0,\varepsilon) \times \partial Y^0 \to Y, \quad \iota_Y^1\colon (1-\varepsilon, 1] \times \partial Y^1 \to Y,$$

for some $\varepsilon > 0$. The latter are homeomorphisms onto disjoint open neighborhoods of $\partial^\alpha Y \subset Y$, extending the inclusions $\iota_Y^\alpha(\alpha, \cdot)\colon \partial^\alpha Y \hookrightarrow Y$ for $\alpha = 0, 1$. We call $\partial^0 Y$ and $\partial^1 Y$ the *boundary components* of $(Y, \iota_Y^0, \iota_Y^1)$. The main example is the *trivial* *cobordism* $Y = [0,1] \times X$ with the natural inclusions $\iota_Y^\alpha \colon A_\varepsilon^\alpha \times X \to [0,1] \times X$, where we denote

$$A_\varepsilon^0 := [0,\varepsilon) \quad \text{and} \quad A_\varepsilon^1 := (1-\varepsilon, 1], \quad \text{for } 0 < \varepsilon < \tfrac{1}{2}.$$

Next, a subset $F \subset Y$ is *collared* if there is $0 < \delta \le \varepsilon$ such that for $\alpha = 0, 1$ we have

$$F \cap \operatorname{im}(\iota_Y^\alpha) \ne \varnothing \iff F \cap \iota_Y^\alpha(A_\delta^\alpha \times \partial^\alpha Y) = \iota_Y^\alpha(A_\delta^\alpha \times \partial^\alpha F), \tag{2.4.1}$$

where the intersections with the boundary components $\partial^\alpha F := F \cap \partial^\alpha Y$ may be empty.

In the notion of Kuranishi cobordism, we will require all charts and coordinate changes to be of product form in sufficiently small collars, as follows.

Definition 2.4.1 Let $(Y, \iota_Y^0, \iota_Y^1)$ be a compact collared cobordism.

- Given a Kuranishi chart $\boldsymbol{K}^\alpha = (U^\alpha, E^\alpha, \Gamma^\alpha, s^\alpha, \psi^\alpha)$ for $\partial^\alpha Y$ and an open subset $A \subset [0,1]$, the *product chart* for $[0,1] \times \partial^\alpha Y$ with footprint $A \times F^\alpha$ is

$$A \times \boldsymbol{K}^\alpha := (A \times U^\alpha, E^\alpha, \Gamma^\alpha, s^\alpha \circ \mathrm{pr}_{U^\alpha}, \mathrm{id}_A \times \psi^\alpha),$$

where Γ^α acts trivially on the first factor of $A \times U^\alpha$ and $\mathrm{pr}_{U^\alpha} \colon A \times U^\alpha \to U^\alpha$ is the evident projection.

- Given a coordinate change $\widehat{\Phi}_{IJ}^\alpha = (\widetilde{\phi}_{IJ}^\alpha, \widehat{\phi}_{IJ}^\alpha, \rho_{IJ}^\alpha) \colon \boldsymbol{K}_I^\alpha \to \boldsymbol{K}_J^\alpha$ between Kuranishi charts for $\partial^\alpha Y$ with lifted domain $\widetilde{U}_{IJ}^\alpha$, and open subsets $A_I, A_J \subset [0,1]$, the *product coordinate change* $(A_I \cap A_J) \times \boldsymbol{K}_I^\alpha \to A_J \times \boldsymbol{K}_J^\alpha$ is

$$\mathrm{id}_{A_I \cap A_J} \times \widehat{\Phi}_{IJ}^\alpha : (\mathrm{id}_{A_I \cap A_J} \times \widetilde{\phi}_{IJ}^\alpha, \widehat{\phi}_{IJ} := \widehat{\phi}_{IJ}^\alpha, \ \mathrm{id}_{A_I \cap A_J} \times \rho_{IJ}^\alpha)$$

with the lifted domain $(A_I \cap A_J) \times \widetilde{U}_{IJ}^\alpha$.

- A *Kuranishi chart with collared boundary* for $(Y, \iota_Y^0, \iota_Y^1)$ is given by a tuple $\boldsymbol{K} = (U, E, \Gamma, s, \psi)$ as in Definition 2.2.2, with the following collar form requirements:

  (i) The footprint $F \subset Y$ is collared with at least one nonempty boundary $\partial^\alpha F$.

  (ii) The domain is a collared cobordism $(U, \iota_U^0, \iota_U^1)$ whose boundary components $\partial^\alpha U$ are nonempty if and only if $\partial^\alpha F \ne \varnothing$. It is smooth in the sense that U is a manifold with boundary $\partial U = \partial^0 U \sqcup \partial^1 U$ and the ι_U^α are tubular neighborhood diffeomorphisms.

  (iii) If $\partial^\alpha F \ne \varnothing$ then there is a *restriction of* $\boldsymbol{K}$ *to the boundary* $\partial^\alpha Y$; that is, a Kuranishi chart $\partial^\alpha \boldsymbol{K} = (\partial^\alpha U^\alpha, E, \Gamma, s^\alpha, \psi^\alpha)$ for $\partial^\alpha Y$, with the isotropy group Γ and obstruction space E of $\boldsymbol{K}$ and footprint $\partial^\alpha F$, and an embedding of the product chart $A_\varepsilon^\alpha \times \partial^\alpha \boldsymbol{K}$ into $\boldsymbol{K}$ for some $\varepsilon > 0$, in the sense that the boundary embedding ι^α_U is Γ–equivariant and the following diagrams commute:

$$
\begin{tikzcd}
A^\alpha_\varepsilon \times \partial^\alpha U \arrow[r, "\iota^\alpha_U"] \arrow[d, "s^\alpha \circ \mathrm{pr}_{\partial^\alpha U}"'] & U \arrow[d, "s"] \\
E \arrow[r, "\mathrm{id}_E"] & E
\end{tikzcd}
\qquad
\begin{tikzcd}
(\mathrm{id}_{A^\alpha_\varepsilon} \times s^\alpha)^{-1}(0) \arrow[r, "\iota^\alpha_U"] \arrow[d, "\mathrm{id}_{A^\alpha_\varepsilon} \times \psi^\alpha"'] & s^{-1}(0) \arrow[d, "\psi"] \\
A^\alpha_\varepsilon \times \partial^\alpha Y \arrow[r, "\iota^\alpha_Y"] & Y
\end{tikzcd}
$$

- Let $\boldsymbol{K}_I, \boldsymbol{K}_J$ be Kuranishi charts for $(Y, \iota^0_Y, \iota^1_Y)$ such that only $\boldsymbol{K}_I$ or both $\boldsymbol{K}_I, \boldsymbol{K}_J$ have collared boundary. A *coordinate change with collared boundary* $\widehat{\Phi}_{IJ}\colon \boldsymbol{K}_I \to \boldsymbol{K}_J$ with domain $\underline{U}_{IJ}$ satisfies the conditions in Definition 2.2.8, with the following collar form requirements:

(i) The lifted domain $\widetilde{U}_{IJ} \subset U_J$, as well as $U_{IJ} \subset U_I$, are collared subsets.

(ii) If $F_J \cap \partial^\alpha Y \neq \varnothing$ then $F_I \cap \partial^\alpha Y \neq \varnothing$ and there is a *restriction of* $\widehat{\Phi}_{IJ}$ *to the boundary* $\partial^\alpha Y$; that is, a coordinate change $\partial^\alpha \widehat{\Phi}_{IJ}\colon \partial^\alpha \boldsymbol{K}_I \to \partial^\alpha \boldsymbol{K}_J$ such that the restriction of $\widehat{\Phi}_{IJ}$ to

$$\underline{U_{IJ} \cap \iota^\alpha_{U_I}(A^\alpha_\varepsilon \times \partial^\alpha U_I)}$$

pulls back via the collar inclusions $\iota^\alpha_{U_I}$, $\iota^\alpha_{U_J}$ to the product coordinate change $\mathrm{id}_{A^\alpha_\varepsilon} \times \partial^\alpha \widehat{\Phi}_{IJ}$ for some $\varepsilon > 0$. In particular we have

$$(\iota^\alpha_{U_J})^{-1}(\widetilde{U}_{IJ}) \cap (A^\alpha_\varepsilon \times \partial^\alpha U_J) = A^\alpha_\varepsilon \times \partial^\alpha \widetilde{U}_{IJ},$$
$$(\iota^\alpha_{U_I})^{-1}(U_{IJ}) \cap (A^\alpha_\varepsilon \times \partial^\alpha U_I) = A^\alpha_\varepsilon \times \partial^\alpha U_{IJ}.$$

(iii) If $F_J \cap \partial^\alpha Y = \varnothing$ but $F_I \cap \partial^\alpha Y \neq \varnothing$, then $U_{IJ} \subset U_I$ is collared with $\partial^\alpha U_{IJ} = \varnothing$. As a consequence we have $U_{IJ} \cap \iota^\alpha_{U_I}(A^\alpha_\varepsilon \times \partial^\alpha \widetilde{U}_I) = \varnothing$ for some $\varepsilon > 0$.

Definition 2.4.2 A *(weak) Kuranishi cobordism* on a compact collared cobordism $(Y, \iota^0_Y, \iota^1_Y)$ is a tuple $\mathcal{K} = (\boldsymbol{K}_I, \widehat{\Phi}_{IJ})_{I,J \in \mathcal{I}_\mathcal{K}}$ of basic charts and transition data as in Definition 2.3.1, with the following collar form requirements:

- The charts of $\mathcal{K}$ are either Kuranishi charts with collared boundary or standard Kuranishi charts whose footprints are precompactly contained in $Y \setminus (\partial^0 Y \cup \partial^1 Y)$.
- The coordinate changes $\widehat{\Phi}_{IJ}\colon \boldsymbol{K}_I \to \boldsymbol{K}_J$ are either standard coordinate changes on $Y \setminus (\partial^0 Y \cup \partial^1 Y)$ between pairs of standard charts, or coordinate changes with collared boundary between pairs of charts, of which at least the first has collared boundary.

We say that $\mathcal{K}$ has *uniform collar width* $\delta > 0$ if all domains and coordinate changes have the required collar form over intervals A^α_ε of length $\varepsilon > \delta$.

Remark 2.4.3 Let $\mathcal{K}$ be a (weak) Kuranishi cobordism on $(Y, \iota_Y^0, \iota_Y^1)$.

(i) $\mathcal{K}$ induces by restriction (weak) Kuranishi atlases $\partial^\alpha \mathcal{K}$ on the boundary components $\partial^\alpha Y$ for $\alpha = 0, 1$ with

- basic charts $\partial^\alpha \boldsymbol{K}_i$ given by restriction of basic charts of $\mathcal{K}$ with $F_i \cap \partial^\alpha Y \neq \varnothing$;
- index set $\mathcal{I}_{\partial^\alpha \mathcal{K}} = \{I \in \mathcal{I}_{\mathcal{K}} \mid F_I \cap \partial^\alpha Y \neq \varnothing\}$;
- transition charts $\partial^\alpha \boldsymbol{K}_I$ given by restriction of transition charts of $\mathcal{K}$;
- coordinate changes $\partial^\alpha \widehat{\Phi}_{IJ}$ given by restriction of coordinate changes of $\mathcal{K}$.

(ii) The charts and coordinate changes of $\mathcal{K}$ induce intermediate charts and coordinate changes as in Definition 2.2.3 and Remark 2.2.9(iii). These fit together to form a filtered (weak) topological cobordism $\underline{\mathcal{K}}$ in the sense of [13, Definitions 4.1.12] by a direct generalization of Lemma 2.3.4. Its boundary restrictions are the intermediate Kuranishi atlases $\partial^\alpha \underline{\mathcal{K}} = \underline{\partial^\alpha \mathcal{K}}$ induced by the boundary restrictions $\partial^\alpha \mathcal{K}$.

(iii) As in [13, Remark 4.1.11] we can think of the virtual neighborhood $|\mathcal{K}|$ as a collared cobordism with boundary components $\partial^0|\mathcal{K}| \cong |\partial^0 \mathcal{K}|$ and $\partial^1|\mathcal{K}| \cong |\partial^1 \mathcal{K}|$, with the exception that $|\mathcal{K}|$ is usually not locally compact or metrizable. More precisely, if $\mathcal{K}$ has collar width $\varepsilon > 0$, then the inclusions $\iota_{U_I}^\alpha \colon A_\varepsilon^\alpha \times U_I^\alpha \hookrightarrow U_I$ induce topological embeddings

$$\iota_{|\mathcal{K}|}^0 \colon [0, \varepsilon) \times |\partial^0 \mathcal{K}| \hookrightarrow |\mathcal{K}|, \quad \iota_{|\mathcal{K}|}^1 \colon (1-\varepsilon, 1] \times |\partial^1 \mathcal{K}| \hookrightarrow |\mathcal{K}|$$

to open neighborhoods of the closed subsets

$$\partial^\alpha |\mathcal{K}| := \bigsqcup_{I \in \mathcal{I}_{\partial^\alpha \mathcal{K}}} \iota_{U_I}^\alpha(\{\alpha\} \times U_I^\alpha)/\sim \; \subset |\mathcal{K}|. \qquad \diamond$$

With this language in hand, one obtains cobordism relations between (weak) Kuranishi atlases in complete analogy with [13, Definition 4.1.8] and [14, Definition 6.2.10]. For the uniqueness results in this paper, the more important notion is the following. Here we use the notion of tameness, a refinement of the strong cocycle condition that is formalized in Definition 2.5.1 below.

Definition 2.4.4 Two (weak/tame) Kuranishi atlases $\mathcal{K}^0$, $\mathcal{K}^1$ on the same compact metrizable space X are said to be *(weakly/tamely) concordant* if there exists a (weak/tame) Kuranishi cobordism $\mathcal{K}$ on the trivial cobordism $Y = [0, 1] \times X$ whose boundary restrictions are $\partial^0 \mathcal{K} = \mathcal{K}^0$ and $\partial^1 \mathcal{K} = \mathcal{K}^1$. More precisely, there are injections $\iota^\alpha \colon \mathcal{I}_{\mathcal{K}^\alpha} \hookrightarrow \mathcal{I}_{\mathcal{K}}$ for $\alpha = 0, 1$ such that $\operatorname{im} \iota^\alpha = \mathcal{I}_{\partial^\alpha \mathcal{K}}$ and we have

$$\boldsymbol{K}_I^\alpha = \partial^\alpha \boldsymbol{K}_{\iota^\alpha(I)}, \quad \widehat{\Phi}_{IJ}^\alpha = \partial^\alpha \widehat{\Phi}_{\iota^\alpha(I)\iota^\alpha(J)} \quad \forall I, J \in \mathcal{I}_{\mathcal{K}^\alpha}.$$

Moreover, two metric Kuranishi atlases $(\mathcal{K}^0, d_0)$, $(\mathcal{K}^1, d_1)$ are called *metric concordant* if they are concordant as above with $\mathcal{K}$ a Kuranishi cobordism whose realization $|\mathcal{K}| \cong |\underline{\mathcal{K}}|$ supports an admissible, ε–collared metric d in the sense of [13, Definition 4.2.1] for the intermediate cobordism atlas $\underline{\mathcal{K}}$ such that $d|_{\partial^\alpha|\mathcal{K}|} = d_\alpha$ for $\alpha = 0, 1$.

2.5 Tameness and shrinkings

As in the case of trivial isotropy, we must adjust the Kuranishi atlas in order for its realization $|\mathcal{K}|$ to have good topological properties; for example, so that it is Hausdorff and has "enough" compact subsets. We essentially already dealt with these problems in [13] by

- introducing notions of tameness and preshrunk shrinking for topological Kuranishi atlases, which ensure the desired topological properties of the realization;
- constructing tame shrinkings of filtered weak topological Kuranishi atlases;
- proving that tame shrinkings are unique up to tame concordance.

In order to apply these results to smooth Kuranishi atlases with nontrivial isotropy, recall first that we built additivity into the notion of Kuranishi atlas, and showed in Lemma 2.3.4 that the resulting intermediate atlases are naturally filtered by

$$(\mathbb{E}_{IJ} := \underline{U_J \times \widehat{\phi}_{IJ}(E_I)})_{I \subset J}.$$

The same holds for Kuranishi cobordisms by Remark 2.4.3(ii). We can thus extend the notions of tameness to the case of nontrivial isotropy by working at the level of the intermediate category.

Definition 2.5.1 A weak Kuranishi atlas or cobordism is *tame* if its intermediate atlas is tame in the sense of [13, Definition 3.1.10]; that is, for all $I, J, K \in \mathcal{I}_{\mathcal{K}}$ we have

$$\underline{U}_{IJ} \cap \underline{U}_{IK} = \underline{U}_{I(J \cup K)} \qquad \forall I \subset J, K, \tag{2.5.1}$$

$$\underline{\phi}_{IJ}(\underline{U}_{IK}) = \underline{U}_{JK} \cap \underline{\mathfrak{s}}_J^{-1}(\mathbb{E}_{IK}) \qquad \forall I \subset J \subset K. \tag{2.5.2}$$

Here we allow equalities between I, J and K using the notation $\underline{U}_{II} := \underline{U_I}$ and $\underline{\phi}_{II} := \mathrm{Id}_{\underline{U}_I}$.

Similarly, a shrinking of a Kuranishi atlas or cobordism will arise exactly from a shrinking $(\underline{U}'_I \sqsubset \underline{U}_I)_{I \in \mathcal{I}_{\mathcal{K}}}$ of the intermediate atlas in the sense of [13, Definition 3.3.2]. Recall that shrinkings of cobordisms are necessarily given by collared subsets $\underline{U}'_I \sqsubset \underline{U}_I$.

Definition 2.5.2 Let $\mathcal{K} = (\boldsymbol{K}_I, \widehat{\Phi}_{IJ})_{I,J \in \mathcal{I}_{\mathcal{K}}, I \subsetneq J}$ be a weak Kuranishi atlas or cobordism. Then a weak Kuranishi atlas or cobordism $\mathcal{K}' = (\boldsymbol{K}'_I, \widehat{\Phi}'_{IJ})_{I,J \in \mathcal{I}_{\mathcal{K}'}, I \subsetneq J}$ is a *shrinking* of $\mathcal{K}$ if:

(i) The footprint cover $(F'_i)_{i=1,\dots,N}$ is a shrinking of the cover $(F_i)_{i=1,\dots,N}$; that is, $F'_i \sqsubset F_i$ are precompact open subsets such that $X = \bigcup_{i=1,\dots,N} F'_i$ and $F'_I := \bigcap_{i\in I} F'_i$ is nonempty whenever F_I is, so that the index sets $\mathcal{I}_{\mathcal{K}'} = \mathcal{I}_{\mathcal{K}}$ agree.

(ii) For each $I \in \mathcal{I}_{\mathcal{K}}$ the chart $\boldsymbol{K}'_I$ is the restriction of $\boldsymbol{K}_I$ to a precompact domain $\underline{U}'_I \sqsubset \underline{U}_I$ as in Definition 2.2.6.

(iii) For each $I, J \in \mathcal{I}_{\mathcal{K}}$ with $I \subsetneq J$ the coordinate change $\widehat{\Phi}'_{IJ}$ is the restriction of $\widehat{\Phi}_{IJ}$ to the open subset $\underline{U}'_{IJ} := \underline{\phi}_{IJ}^{-1}(\underline{U}'_J) \cap \underline{U}'_I$ as in Equation (2.2.5).

A *tame shrinking* of $\mathcal{K}$ is a shrinking that is tame in the sense of Definition 2.5.1. Finally, a *preshrunk tame shrinking* of $\mathcal{K}$ is a tame shrinking $\mathcal{K}''$ that is obtained as a shrinking of a tame shrinking $\mathcal{K}'$ of $\mathcal{K}$.

With this language in place, we can directly generalize [14, Theorem 6.3.9]. Recall here that by [13, Example 2.4.5] the quotient topology on $|\mathcal{K}|$ is never metrizable except in the most trivial cases. In fact, for any point $x \in \overline{U_{IJ}} \setminus U_{IJ}$ where $\dim U_I < \dim U_J$, the projection $\pi_{\mathcal{K}}(x)$ does not have a countable neighborhood basis in $|\mathcal{K}|$ with respect to the quotient topology. So an admissible metric will almost always induce a different topology on $|\mathcal{K}|$, which we will make no use of in the following statement.

Theorem 2.5.3 (i) *Any weak Kuranishi atlas or cobordism $\mathcal{K}$ has a preshrunk tame shrinking $\mathcal{K}'$.*

(ii) *For any tame Kuranishi atlas or cobordism $\mathcal{K}'$, the realizations $|\mathcal{K}'|$ and $|\boldsymbol{E}_{\mathcal{K}'}|$ are Hausdorff in the quotient topology, and for each $I \in \mathcal{I}_{\mathcal{K}'}$ the projection maps $\pi_{\underline{\mathcal{K}'}}\colon \underline{U}'_I \to |\mathcal{K}'|$ and $\pi_{\boldsymbol{E}_{\underline{\mathcal{K}'}}}\colon \underline{U'_I \times E_I} \to |\boldsymbol{E}_{\mathcal{K}'}|$ are homeomorphisms onto their images.*

(iii) *For any preshrunk tame shrinking $\mathcal{K}'$ as in (i), there exists an admissible metric on the set $|\mathcal{K}'|$. If $\mathcal{K}$ is a cobordism, then the metric can also be taken to be collared.*

(iv) *Any two metric preshrunk tame shrinkings of a weak Kuranishi atlas are metric tame concordant.*

Proof Since tameness, shrinking and admissible metrics are all defined on the level of intermediate atlases, and we are only concerned with homeomorphism properties of the intermediate projections, in the case of Kuranishi atlases we can simply quote [13, Proposition 3.3.5] for (i), [13, Proposition 3.1.13] for (ii), and [13, Proposition 3.3.8] for (iii). Moreover, [13, Proposition 4.2.3] proves (iv), as well as (i) and (iii) for Kuranishi cobordisms, and (ii) for cobordisms is established in [13, Lemma 4.1.15]. □

3 From Kuranishi atlases to the virtual fundamental class

In this section, Section 3.1 discusses orientations, Section 3.2 establishes the notions of reductions and perturbations. The main result here is Theorem 3.2.8, which shows that the zero set of a suitable perturbation $\mathfrak{s}_{\mathcal{K}} + \nu$ of the canonical section $\mathfrak{s}_{\mathcal{K}}$ has the structure of a compact weighted branched manifold. The construction of such perturbations is deferred to Proposition 3.3.3, and is followed by the construction of the VMC and VFC in Theorem 3.3.5.

3.1 Orientations

This section extends the theory of orientations of weak Kuranishi atlases from [14, Section 8.1] to the case with nontrivial isotropy. Since we use the method of determinant bundles, we first need to generalize the notions of vector bundles and isomorphisms.

Definition 3.1.1 A *vector bundle* $\Lambda = (\Lambda_I, \widetilde{\phi}^{\Lambda}_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}}}$ *over a weak Kuranishi atlas* $\mathcal{K}$ consists of local bundles and compatible transition maps as follows:

- For each $I \in \mathcal{I}_{\mathcal{K}}$, a vector bundle $\Lambda_I \to U_I$ with an action of Γ_I on Λ_I that covers the given action on U_I.
- For each $I \subsetneq J$, a Γ_J–equivariant map $\widetilde{\phi}^{\Lambda}_{IJ}\colon \rho^*_{IJ}(\Lambda_I|_{U_{IJ}}) \to \Lambda_J$ that is a linear isomorphism on each fiber and covers the embedding $\widetilde{\phi}_{IJ}\colon \widetilde{U}_{IJ} \to U_J$. Here $\Gamma_J \cong \Gamma_I \times \Gamma_{J\setminus I}$ acts on $\rho^*_{IJ}(\Lambda_I|_{U_{IJ}}) \to \widetilde{U}_{IJ}$ by the pullback action of Γ_I together with the natural identification of the fibers of $\rho^*_{IJ}(\Lambda_I|_{U_{IJ}})$ along $\Gamma_{J\setminus I}$–orbits in $\widetilde{U}_{IJ}$.
- For each $I \subsetneq J \subsetneq K$, we have the weak cocycle condition
$$\widetilde{\phi}^{\Lambda}_{IK} = \widetilde{\phi}^{\Lambda}_{JK} \circ \rho^*_{JK}(\widetilde{\phi}^{\Lambda}_{IJ}) \quad \text{on } \rho^{-1}_{JK}(\widetilde{\phi}_{IJ}(\widetilde{U}_{IJ})) \cap \widetilde{U}_{IK}.$$

A *section* of a vector bundle Λ over $\mathcal{K}$ is a collection of smooth Γ_I–equivariant sections $\sigma = (\sigma_I\colon U_I \to \Lambda_I)_{I\in\mathcal{I}_{\mathcal{K}}}$ that are compatible with the pullbacks ρ^*_{IJ} and bundle maps $\widetilde{\phi}^{\Lambda}_{IJ}$ in the sense that there are commutative diagrams for each $I \subsetneq J$:

$$\begin{tikzcd}
\Lambda_I|_{U_{IJ}} & \rho^*_{IJ}(\Lambda_I|_{U_{IJ}}) \arrow[l, "\rho_{IJ}"'] \arrow[r, "\widetilde{\phi}^{\Lambda}_{IJ}"] & \Lambda_J \\
U_{IJ} \arrow[u, "\sigma_I"] & \widetilde{U}_{IJ} \arrow[l, "\rho_{IJ}"'] \arrow[u, "\rho^*_{IJ}(\sigma_I)"'] \arrow[r, "\widetilde{\phi}_{IJ}"] & U_J. \arrow[u, "\sigma_J"']
\end{tikzcd}$$

Definition 3.1.2 If $\Lambda = (\Lambda_I, \widetilde{\phi}^{\Lambda}_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}}}$ is a bundle over $\mathcal{K}$ and $A \subset [0,1]$ an interval, then the *product bundle* $A \times \Lambda$ over $A \times \mathcal{K}$ is the tuple $(A \times \Lambda_I, \mathrm{id}_A \times \widetilde{\phi}^{\Lambda}_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}}}$. Here and in the following we denote by $A \times \Lambda_I \to A \times U_I$ the pullback bundle of $\Lambda_I \to U_I$ under the projection $\mathrm{pr}_{U_I}\colon A \times U_I \to U_I$.

Definition 3.1.3 A *vector bundle over a weak Kuranishi cobordism* $\mathcal{K}$ is a collection $\Lambda = (\Lambda_I, \widetilde{\phi}^{\Lambda}_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}}}$ of vector bundles and bundle maps as in Definition 3.1.1, together with a choice of isomorphism from its collar restriction to a product bundle. More precisely, this requires for $\alpha = 0, 1$ the choice of a *restricted vector bundle*

$$\Lambda|_{\partial^\alpha\mathcal{K}} = (\Lambda^\alpha_I \to \partial^\alpha U_I, \widetilde{\phi}^{\Lambda,\alpha}_{IJ})_{I,J\in\mathcal{I}_{\partial^\alpha\mathcal{K}}}$$

over $\partial^\alpha\mathcal{K}$, and, for some $\varepsilon > 0$ less than the collar width of $\mathcal{K}$, a choice of lifts of the embeddings ι^α_I for $I \in \mathcal{I}_{\partial^\alpha\mathcal{K}}$ to Γ_I–equivariant bundle isomorphisms

$$\widetilde{\iota}^{\Lambda,\alpha}_I \colon A^\alpha_\varepsilon \times \Lambda^\alpha_I \to \Lambda_I|_{\operatorname{im}\iota^\alpha_I}$$

such that, with $A := A^\alpha_\varepsilon$ and $\rho^*\widetilde{\iota}^{\Lambda,\alpha}_I := \rho^*_{IJ} \circ \widetilde{\iota}^{\Lambda,\alpha}_I \circ (\mathrm{id}_A \times (\rho^\alpha_{IJ})_*)$, the following diagrams commute:

$$\begin{tikzcd}
A \times \Lambda^\alpha_I \arrow[r, "\widetilde{\iota}^{\Lambda,\alpha}_I"] \arrow[d] & \Lambda_I|_{\operatorname{im}\iota^\alpha_I} \arrow[d] \\
A \times \partial^\alpha U_I \arrow[r, "\iota^\alpha_I"] & \operatorname{im}\iota^\alpha_I \subset U_I
\end{tikzcd}
\qquad
\begin{tikzcd}
A \times (\rho^\alpha_{IJ})^*(\Lambda^\alpha_I|_{\partial^\alpha U_{IJ}}) \arrow[r, "\rho^*\widetilde{\iota}^{\Lambda,\alpha}_I"] \arrow[d, "\mathrm{id}_A \times \widetilde{\phi}^{\Lambda,\alpha}_{IJ}"'] & \rho^*_{IJ}(\Lambda_I|_{\iota^\alpha_I(A\times\partial^\alpha U_{IJ})}) \arrow[d, "\widetilde{\phi}^{\Lambda}_{IJ}"] \\
A \times \Lambda^\alpha_J \arrow[r, "\widetilde{\iota}^{\Lambda,\alpha}_J"] & \Lambda_J|_{\operatorname{im}\iota^\alpha_J}
\end{tikzcd}$$

A *section* of a vector bundle Λ over a Kuranishi cobordism as above is a compatible collection $(\sigma_I \colon U_I \to \Lambda_I)_{I\in\mathcal{I}_{\mathcal{K}}}$ of equivariant sections as in Definition 3.1.1 that in addition are of product form in the collar. That is, we require that for each $\alpha = 0, 1$ there is a *restricted section* $\sigma|_{\partial^\alpha\mathcal{K}} = (\sigma^\alpha_I \colon \partial_\alpha U_I \to \Lambda^\alpha_I)_{I\in\mathcal{I}_{\partial^\alpha\mathcal{K}}}$ of $\Lambda|_{\partial^\alpha\mathcal{K}}$ such that for $\varepsilon > 0$ sufficiently small, $(\widetilde{\iota}^{\Lambda,\alpha}_I)^*\sigma_I = \mathrm{id}_{A^\alpha_\varepsilon} \times \sigma^\alpha_I$.

In the above definition we implicitly work with an isomorphism $(\widetilde{\iota}^{\Lambda,\alpha}_I)_{I\in\mathcal{I}_{\partial^\alpha\mathcal{K}}}$ that satisfies all but the product structure requirements of the following notion of isomorphism on Kuranishi cobordisms.

Definition 3.1.4 An *isomorphism* $\Psi \colon \Lambda \to \Lambda'$ between vector bundles over $\mathcal{K}$ is a collection $(\Psi_I \colon \Lambda_I \to \Lambda'_I)_{I\in\mathcal{I}_{\mathcal{K}}}$ of Γ_I–equivariant bundle isomorphisms covering the identity on U_I that intertwine the transition maps, ie $\widetilde{\phi}^{\Lambda'}_{IJ} \circ \rho^*_{IJ}(\Psi_I) = \Psi_J \circ \widetilde{\phi}^{\Lambda}_{IJ}|_{\widetilde{U}_{IJ}}$ for all $I \subsetneq J$.

If $\mathcal{K}$ is a Kuranishi cobordism then we additionally require Ψ to be of product form in the collar. That is, we require that for each $\alpha = 0, 1$ there is a restricted isomorphism $\Psi|_{\partial^\alpha\mathcal{K}} = (\Psi^\alpha_I \colon \Lambda^\alpha_I \to \Lambda'^\alpha_I)_{I\in\mathcal{I}_{\partial^\alpha\mathcal{K}}}$ from $\Lambda|_{\partial^\alpha\mathcal{K}}$ to $\Lambda'|_{\partial^\alpha\mathcal{K}}$ such that for $\varepsilon > 0$ sufficiently small we have $(\widetilde{\iota}'_I)^{\Lambda,\alpha} \circ (\mathrm{id}_A \times \Psi^\alpha_I) = \Psi_I \circ \widetilde{\iota}^{\Lambda,\alpha}_I$ on $A^\alpha_\varepsilon \times \partial^\alpha U_I$.

Note that although the compatibility conditions are the same, the canonical section $\mathfrak{s}_{\mathcal{K}} = (s_I \colon U_I \to E_I)_{I\in\mathcal{I}_{\mathcal{K}}}$ of a Kuranishi atlas does not form a section of a vector bundle since the obstruction spaces E_I are in general not of the same dimension, hence no bundle isomorphisms $\widetilde{\phi}_{IJ}^{\Lambda}$ as above exist. Nevertheless, we will see that there is a natural bundle associated with the section $\mathfrak{s}_{\mathcal{K}}$, namely its determinant line bundle, and that this line bundle is isomorphic to a bundle constructed by combining the determinant lines of the obstruction spaces E_I and the domains U_I.

Remark 3.1.5 If Λ is a bundle over a Kuranishi atlas $\mathcal{K}$ (rather than a weak atlas), then it is straightforward to verify that the union $\bigsqcup_I \Lambda_I$ of the local bundles form the objects of a category with projection to the Kuranishi category $\boldsymbol{B}_{\mathcal{K}}$. We did not formulate the above definitions in this language since orientations in applications to moduli spaces (eg Gromov–Witten as in [10]) will usually be constructed on a weak Kuranishi atlas, which does not form a category. $\diamond$

Here and in the following we will exclusively work with finite-dimensional vector spaces. First recall that the determinant line of a vector space V is its maximal exterior power $\Lambda^{\max} V := \bigwedge^{\dim V} V$, with $\bigwedge^0 \{0\} := \mathbb{R}$. More generally, the *determinant line of a linear map* $D\colon V \to W$ is defined to be

$$\det(D) := \Lambda^{\max} \ker D \otimes (\Lambda^{\max}(W/\operatorname{im} D))^*. \tag{3.1.1}$$

In order to construct isomorphisms between determinant lines, we will need to fix various conventions, in particular pertaining to the ordering of factors in their domains and targets. We begin by noting that every isomorphism $F\colon Y \to Z$ between finite-dimensional vector spaces induces an isomorphism

$$\Lambda_F\colon \Lambda^{\max} Y \xrightarrow{\cong} \Lambda^{\max} Z, \quad y_1 \wedge \cdots \wedge y_k \mapsto F(y_1) \wedge \cdots \wedge F(y_k). \tag{3.1.2}$$

For example, the fact that $\gamma \circ s_I := s_I \circ \gamma\colon U_I \to E_I$ for all $\gamma \in \Gamma_I$, implies that

$$\gamma_\Lambda := \Lambda_{\mathrm{d}_x \gamma} \otimes (\Lambda_{[\gamma]^{-1}})^*\colon \det(\mathrm{d}_x s_I) \to \det(\mathrm{d}_{\gamma x} s_I) \tag{3.1.3}$$

is an isomorphism, where $[\gamma]\colon E_I/\operatorname{im} \mathrm{d}_x s_I \to E_I/\operatorname{im} \mathrm{d}_{\gamma x} s_I$ is the induced map. Further, if $I \subsetneq J$ and $\widetilde{x} \in \widetilde{U}_{IJ}$ is such that $\rho_{IJ}(\widetilde{x}) = x$, then because

$$s_I \circ \rho_{IJ} =: s_{IJ}\colon \widetilde{U}_{IJ} \to E_I,$$

the derivative $\mathrm{d}_{\widetilde{x}} \rho_{IJ}\colon \ker \mathrm{d}s_{IJ} \to \ker \mathrm{d}s_I$ induces an isomorphism

$$\Lambda_{\mathrm{d}_{\widetilde{x}} \rho_{IJ}} \otimes \Lambda_{\mathrm{Id}}\colon \det(\mathrm{d}_{\widetilde{x}} s_{IJ}) \to \det(\mathrm{d}_x s_I)$$

and composition with pullback by ρ_{IJ} defines an isomorphism

$$P_{IJ}(\widetilde{x})\colon \det(\mathrm{d}_{\widetilde{x}} s_{IJ}) \to \rho_{IJ}^*(\det \mathrm{d}_x s_I). \tag{3.1.4}$$

Further, it follows from the index condition in Definition 2.2.8 that with $y := \widetilde{\phi}_{IJ}(\widetilde{x})$, the map

$$\widetilde{\Lambda}_{IJ}(\widetilde{x}) := \Lambda_{\mathrm{d}_{\widetilde{x}}\widetilde{\phi}_{IJ}} \otimes (\Lambda_{[\widehat{\phi}_{IJ}]^{-1}})^* \colon \det(\mathrm{d}_{\widetilde{x}} s_{IJ}) \to \det(\mathrm{d}_y s_J) \tag{3.1.5}$$

is an isomorphism, induced by the isomorphisms $\mathrm{d}\widetilde{\phi}_{IJ} \colon \ker \mathrm{d}s_{IJ} \to \ker \mathrm{d}s_J$ and $[\widehat{\phi}_{IJ}] \colon E_I / \operatorname{im} \mathrm{d}s_I \to E_J / \operatorname{im} \mathrm{d}s_J$. We can therefore define the determinant bundle $\det(\mathfrak{s}_{\mathcal{K}})$ of a Kuranishi atlas. A second, isomorphic, determinant line bundle $\det(\mathcal{K})$ with fibers $\Lambda^{\max}\mathrm{T}_x U_I \otimes (\Lambda^{\max} E_I)^*$ will be constructed in Proposition 3.1.13.

Definition 3.1.6 The *determinant line bundle* of a weak Kuranishi atlas (or cobordism) $\mathcal{K}$ is the vector bundle $\det(\mathfrak{s}_{\mathcal{K}})$ given by the line bundles

$$\det(\mathrm{d}s_I) := \bigcup_{x \in U_I} \det(\mathrm{d}_x s_I) \to U_I \qquad \text{for all } I \in \mathcal{I}_{\mathcal{K}},$$

with Γ_I actions given by the isomorphisms γ_Λ of (3.1.3), and the isomorphisms $\widetilde{\phi}^{\Lambda}_{IJ}(\widetilde{x}) := \Lambda_{IJ}(\widetilde{x}) \circ P_{IJ}(\widetilde{x})^{-1}$ in (3.1.4) and (3.1.5) for $I \subsetneq J$ and $\widetilde{x} \in U_{IJ}$.

To show that $\det(\mathfrak{s}_{\mathcal{K}})$ is well defined, in particular that $\widetilde{x} \mapsto \Lambda_{IJ}(\widetilde{x})$ is smooth, we introduce some further natural[6] isomorphisms and fix various ordering conventions.

- For any subspace $V' \subset V$ the *splitting isomorphism*

$$\Lambda^{\max} V \cong \Lambda^{\max} V' \otimes \Lambda^{\max}(V/V') \tag{3.1.6}$$

is given by completing a basis $v_1, \dots, v_k$ of V' to a basis $v_1, \dots, v_n$ of V and mapping $v_1 \wedge \cdots \wedge v_n \mapsto (v_1 \wedge \cdots \wedge v_k) \otimes ([v_{k+1}] \wedge \cdots \wedge [v_n])$.

- For each isomorphism $F \colon Y \xrightarrow{\cong} Z$ the *contraction isomorphism*

$$\mathfrak{c}_F \colon \Lambda^{\max} Y \otimes (\Lambda^{\max} Z)^* \xrightarrow{\cong} \mathbb{R}, \tag{3.1.7}$$

is given by the map $(y_1 \wedge \cdots \wedge y_k) \otimes \eta \mapsto \eta(F(y_1) \wedge \cdots \wedge F(y_k))$.

- For any space V we use the *duality isomorphism*

$$\Lambda^{\max} V^* \xrightarrow{\cong} (\Lambda^{\max} V)^*, \quad v_1^* \wedge \cdots \wedge v_n^* \mapsto (v_1 \wedge \cdots \wedge v_n)^*, \tag{3.1.8}$$

which corresponds to the natural pairing

$$\Lambda^{\max} V \otimes \Lambda^{\max} V^* \xrightarrow{\cong} \mathbb{R}, \quad (v_1 \wedge \cdots \wedge v_n) \otimes (\eta_1 \wedge \cdots \wedge \eta_n) \mapsto \prod_{i=1}^{n} \eta_i(v_i)$$

[6] Here a "natural" isomorphism is one that is functorial, ie it commutes with the action on both sides induced by a vector space isomorphism.

via the general identification

$$\text{(3.1.9)} \qquad \operatorname{Hom}(A\otimes B,\mathbb{R}) \xrightarrow{\cong} \operatorname{Hom}(B,A^*), \quad H\mapsto (b\mapsto H(\cdot\otimes b)),$$

which in the case of line bundles A, B maps $\eta\neq 0$ to a nonzero homomorphism, ie an isomorphism. Next, we combine the above isomorphisms to obtain a more elaborate contraction isomorphism.

Lemma 3.1.7 [14, Lemma 8.1.7] *Every linear map $F\colon V\to W$ together with an isomorphism $\phi\colon K\to\ker F$ induces an isomorphism*

$$\text{(3.1.10)} \qquad \mathfrak{C}^{\phi}_F\colon \Lambda^{\max}V\otimes(\Lambda^{\max}W)^* \xrightarrow{\cong} \Lambda^{\max}K\otimes(\Lambda^{\max}(W/F(V)))^*$$

given by

$$(v_1\wedge\cdots\wedge v_n)\otimes(w_1\wedge\cdots\wedge w_m)^*\mapsto(\phi^{-1}(v_1)\wedge\cdots\wedge\phi^{-1}(v_k))\otimes([w_1]\wedge\cdots\wedge[w_{m-n+k}])^*,$$

where $v_1,\dots,v_n$ is a basis for V with $\operatorname{span}(v_1,\dots,v_k)=\ker F$, and $w_1,\dots,w_m$ is a basis for W whose last $n-k$ vectors are $w_{m-n+i}=F(v_i)$ for $i=k+1,\dots,n$.

In particular, for every linear map $D\colon V\to W$ we may pick ϕ as the inclusion $K=\ker D\hookrightarrow V$ to obtain an isomorphism

$$\mathfrak{C}_D\colon \Lambda^{\max}V\otimes(\Lambda^{\max}W)^* \xrightarrow{\cong} \det(D).$$

Remark 3.1.8 If F is equivariant with respect to actions of the group Γ on V and W, and we equip K with the induced Γ action so that ϕ is also equivariant, then the above isomorphism $\mathfrak{C}^{\phi}_F$ is equivariant with respect to the action of Γ on $\Lambda^{\max}V\otimes(\Lambda^{\max}W)^*$ given by the maps $\Lambda_\gamma\otimes(\Lambda_{\gamma^{-1}})^*$ on $\Lambda^{\max}V\otimes(\Lambda^{\max}W)^*$ and by the corresponding maps $\Lambda_\gamma\otimes(\Lambda_{[\gamma]^{-1}})^*$ on $\Lambda^{\max}K\otimes(\Lambda^{\max}(W/F(V)))^*$, with $\Lambda_{[\gamma]}$ as in (3.1.3). ◇

With this notation in hand, we can now prove one of the main results of this section.

Proposition 3.1.9 *For any weak Kuranishi atlas, $\det(\mathfrak{s}_{\mathcal{K}})$ is a well-defined line bundle over $\mathcal{K}$. Further, if $\mathcal{K}$ is a weak Kuranishi cobordism, then $\det(\mathfrak{s}_{\mathcal{K}})$ can be given product form on the collar of $\mathcal{K}$ with restrictions $\det(\mathfrak{s}_{\mathcal{K}})|_{\partial^\alpha\mathcal{K}}=\det(\mathfrak{s}_{\partial^\alpha\mathcal{K}})$ for $\alpha=0,1$. The required bundle isomorphisms from the product $A^\alpha_\varepsilon\times\det(\mathfrak{s}_{\partial^\alpha\mathcal{K}})$ to the collar restriction $(\iota^\alpha_\varepsilon)^*\det(\mathfrak{s}_{\mathcal{K}})$ are given in (3.1.12).*

Proof We use the same local trivializations of $\det(\mathrm{d}s_I)$ as in the proof of the analogous result [14, Proposition 8.1.8] for trivial isotropy, and must check that these are compatible with the isotropy group actions and coordinate changes. We will begin by defining these trivializations, referring to [14] for many details of proofs.

Let $x_0 \in U_I$, and denote its stabilizer subgroup in Γ_I by $\Gamma_I^{x_0}$. Take a subspace of E_I that covers the cokernel of $\mathrm{d}_{x_0} s_I$, sweep it out to obtain a $\Gamma_I^{x_0}$–invariant subspace $E' \subset E_I$, then choose an isomorphism $\mathbb{R}^N \cong E'$ and equip $\mathbb{R}^N$ with the pullback action of $\Gamma_I^{x_0}$ denoted $(\gamma, v) \mapsto \gamma \cdot_{x_0} v$. The resulting equivariant map $R_I \colon (\mathbb{R}^N, \Gamma_I^{x_0}) \to (E_I, \Gamma_I^{x_0})$ covers the cokernel of $\mathrm{d}_x s_I$ for all x in some neighborhood O of x_0.

Thus $\mathrm{d}_x s_I \oplus R_I$ is surjective for $x \in O$, and as in [14, Equation 8.1.9] we may define a trivialization of $\det(\mathrm{d}s_I)|_O$ by

$$(3.1.11)\quad \begin{aligned} &\widehat{T}_{I,x} \colon \Lambda^{\max} \ker(\mathrm{d}_x s_I \oplus R_I) \xrightarrow{\cong} \det(\mathrm{d}_x s_I),\\ &\overline{v}_1 \wedge \cdots \wedge \overline{v}_n \mapsto (v_1 \wedge \cdots \wedge v_k) \otimes ([R_I(e_1)] \wedge \cdots \wedge [R_I(e_{N-n+k})])^*, \end{aligned}$$

where $\overline{v}_i = (v_i, r_i)$ is a basis of $\ker(\mathrm{d}_x s_I \oplus R_I) \subset \mathrm{T}_x U_I \times \mathbb{R}^N$ such that $v_1, \dots, v_k$ span $\ker \mathrm{d}_x s_I$ (and hence $r_1 = \cdots = r_k = 0$), and $e_1, \dots, e_N$ is a positively ordered normalized basis of $\mathbb{R}^N$ (that is, $e_1 \wedge \cdots \wedge e_N = 1 \in \mathbb{R} \cong \Lambda^{\max}\mathbb{R}^N$) such that $R_I(e_{N-n+i}) = \mathrm{d}_x s_I(v_i)$ for $i = k+1, \dots, n$. In particular, the last $n-k$ vectors span $\operatorname{im} \mathrm{d}_x s_I \cap \operatorname{im} R_I \subset E_I$, and thus the first $N-n+k$ vectors $[R_I(e_1)], \dots, [R_I(e_{N-n+k})]$ span the cokernel $E_I / \operatorname{im} \mathrm{d}_x s_I \cong \operatorname{im} R_I / \operatorname{im} \mathrm{d}_x s_I \cap \operatorname{im} R_I$. In [14, Proposition 8.1.8] we prove that these trivializations do not depend on the choice of injection $R_I \colon \mathbb{R}^N \to E_I$. In other words, if $R'_I \colon \mathbb{R}^{N'} \to E_I$ is another Γ_I–equivariant injection that also maps onto the cokernel of $\mathrm{d}_{x_0} s_I$, then there is a bundle isomorphism

$$\Psi \colon \Lambda^{\max} \ker(\mathrm{d}s_I \oplus R_I)|_O \to \Lambda^{\max} \ker(\mathrm{d}s_I \oplus R'_I)|_O$$

which is necessarily Γ_I–equivariant and such that $\widehat{T}_I = \widehat{T}'_I \circ \Psi$. Thus $\det(\mathrm{d}s_I)$ is a smooth line bundle over U_I for each $I \in \mathcal{I}_{\mathcal{K}}$.

It remains to check that the action $\gamma \in \Gamma_I$ on

$$\det(\mathrm{d}s_I) = \Lambda^{\max}(\ker \mathrm{d}s_I) \otimes \Lambda^{\max}(E_I / \operatorname{im} \mathrm{d}s_I)^*$$

is smooth. We prove this by choosing suitable trivializations near x_0 and γx_0 and then lifting the action of γ to a smooth action on the domains $\ker(\mathrm{d}s_I \oplus R_I)$ of the trivializations. To this end, first consider the trivialization $T_{I,x}$ defined near $x_0 \in U_I$ by a $\Gamma_I^{x_0}$–equivariant injection $R_I \colon (\mathbb{R}^N, \Gamma_I^{x_0}) \to (E_I, \Gamma_I^{x_0})$, and for $\gamma \in \Gamma_I$ the associated trivialization $T'_{I,\gamma x}$ defined near $\gamma x_0 \in U_I$ by

$$R'_I := \gamma \circ R_I \colon (\mathbb{R}^N, \Gamma_I^{\gamma x_0}) \to (E_I, \Gamma_I^{\gamma x_0}),$$

where $(\mathbb{R}^N, \Gamma_I^{x_0})$ denotes $\mathbb{R}^N$ with the $\Gamma_I^{x_0}$–action $\delta \colon v \mapsto \delta \cdot_{x_0} v$, and $(\mathbb{R}^N, \Gamma_I^{\gamma x_0})$ denotes $\mathbb{R}^N$ with the $\Gamma_I^{\gamma x_0}$–action

$$\delta' \colon v \mapsto \delta' \cdot_{\gamma x_0} v := \gamma^{-1} \delta' \gamma \cdot_{x_0} v,$$

which is well defined since conjugation by γ defines an isomorphism $c_\gamma\colon \Gamma_I^{\gamma x_0} \to \Gamma_I^{x_0}$, $\delta' \mapsto \gamma^{-1}\delta'\gamma$. Then $R_I' = \gamma \circ R_I$ is $\Gamma_I^{\gamma x_0}$–equivariant because when $\delta' \in \Gamma_I^{\gamma x_0}$,

$$\begin{aligned} R_I'(\delta' \cdot_{\gamma x_0} v) = R_I'(\gamma^{-1}\delta'\gamma \cdot_{x_0} v) &= \gamma R_I(\gamma^{-1}\delta'\gamma \cdot_{x_0} v) \\ &= \gamma(\gamma^{-1}\delta'\gamma) R_I(v) \\ &= \gamma\gamma^{-1}\delta'\gamma R_I(v) = \delta'\gamma R_I(v) = \delta' \circ R_I'(v), \end{aligned}$$

where the fourth equality holds because the full group Γ_I acts on E_I. Thus the diagram

$$\begin{tikzcd} (\mathbb{R}^N, \Gamma_I^{x_0}) \arrow[r, "(R_I{,}\mathrm{id})"] \arrow[d, "(\mathrm{id}{,}c_\gamma^{-1})"] & (E_I, \Gamma_I^{x_0}) \arrow[d, "(\gamma{,}c_\gamma^{-1})"] \\ (\mathbb{R}^N, \Gamma_I^{\gamma x_0}) \arrow[r, "(R_I'{,}\mathrm{id})"] & (E_I, \Gamma_I^{\gamma x_0}) \end{tikzcd}$$

commutes; in other words, the action of the element $\gamma \in \Gamma_I$ on E_I lifts to the identity map of $\mathbb{R}^N$. Hence the definition (3.1.11) of the maps $\widehat{T}_{I,x}$ implies that the following diagram commutes:

$$\begin{tikzcd} \Lambda^{\max}(\ker(\mathrm{d}_x s_I \oplus R_I)) \arrow[r, "\widehat{T}_{I,x}"] \arrow[d, "\Lambda_{\mathrm{d}_x\gamma \oplus \mathrm{id}_{\mathbb{R}^N}}"] & \det \mathrm{d}_x s_I \arrow[d, "\Lambda_{\mathrm{d}_x\gamma} \otimes (\Lambda_{[\gamma^{-1}]})^*"] \\ \Lambda^{\max}(\ker(\mathrm{d}_{\gamma x} s_I \oplus R_I')) \arrow[r, "\widehat{T}_{I,\gamma x}"] & \det \mathrm{d}_{\gamma x} s_I \end{tikzcd}$$

Since the map $\Lambda_{\mathrm{d}_x\gamma \oplus \mathrm{id}_{\mathbb{R}^N}}$ is smooth, so is $\gamma_\Lambda := \Lambda_{\mathrm{d}_x\gamma} \otimes (\Lambda_{[\gamma^{-1}]})^*$. Thus $\det(\mathrm{d}s_I)$ is a Γ_I–equivariant smooth line bundle over U_I for each $I \in \mathcal{I}_{\mathcal{K}}$.

Next note that because $\Gamma_{J \setminus I}$ acts freely on $\widetilde{U}_{IJ}$, the stabilizer subgroup $\Gamma_J^{\widetilde{x}_0}$ of a point $\widetilde{x}_0 \in \rho_{IJ}^{-1}(x_0)$ is taken isomorphically to $\Gamma_I^{x_0}$ by the projection $\rho_{IJ}^\Gamma\colon \Gamma_J \to \Gamma_I$. For simplicity we will identify these groups. Since $s_{IJ}\colon \widetilde{U}_{IJ} \to E_I$ is the composite $s_I \circ \rho_{IJ}$, we may therefore trivialize the bundle $\det(\mathrm{d}_{\widetilde{x}} s_{IJ})$ near $\widetilde{x}_0 \in \rho_{IJ}^{-1}(x_0)$ by using the same injection $R_I\colon \mathbb{R}^N \to E_I$, now considered as a $\Gamma_J^{\widetilde{x}_0}$–equivariant map. Since the diagram

$$\begin{tikzcd} \Lambda^{\max}(\ker(\mathrm{d}_{\widetilde{x}} s_{IJ} \oplus R_I)) \arrow[r, "\widehat{T}_{I,\widetilde{x}}"] \arrow[d, "\Lambda_{\rho_{IJ} \oplus \mathrm{id}_{\mathbb{R}^N}}"] & \det \mathrm{d}_{\widetilde{x}} s_{IJ} \arrow[d, "\Lambda_{\rho_{IJ}} \otimes (\Lambda_{\mathrm{id}})^*"] \\ \Lambda^{\max}(\ker(\mathrm{d}_x s_I \oplus R_I)) \arrow[r, "\widehat{T}_{I,x}"] & \det \mathrm{d}_x s_I \end{tikzcd}$$

commutes, the isomorphism $P_{IJ}(\widetilde{x})\colon \det \mathrm{d}_{\widetilde{x}} s_{IJ} \to \det \mathrm{d}_x s_I$ of (3.1.4) is smooth. Moreover the equivariance of the covering map $\rho_{IJ}\colon (\widetilde{U}_{IJ}, \Gamma_J) \to (U_{IJ}, \Gamma_I)$ and the identity

$s_I \circ \rho_{IJ} = s_{IJ}\colon \widetilde{U}_{IJ} \to E_I$ imply that it is equivariant. Therefore, to complete the proof that the transition maps $\widetilde{\phi}^{\Lambda}_{IJ}$ are smooth, we must check that the map

$$\widetilde{\Lambda}_{IJ}(\widetilde{x}) := \Lambda_{\mathrm{d}_{\widetilde{x}}\widetilde{\phi}_{IJ}} \otimes (\Lambda_{[\widehat{\phi}_{IJ}]^{-1}})^*\colon \det(\mathrm{d}_{\widetilde{x}} s_{IJ}) \to \det(\mathrm{d}_y s_J)$$

in (3.1.5) is equivariant and smooth. Its equivariance follows from the equivariance of its constituent maps $\widetilde{\phi}_{IJ}$ and $\widehat{\phi}_{IJ}$. To see that it is smooth, it suffices to show that the composite $\Lambda_{IJ}(x) := \widetilde{\Lambda}_{IJ}(\rho_{IJ}^{-1}(x))$ is smooth in some neighborhood O of $x_0 \in U_{IJ}$, where $\rho_{IJ}^{-1}\colon O \to \widetilde{U}_{IJ}$ is a local inverse for the covering map ρ_{IJ}. But if we define $\phi_{IJ}(x) := \widetilde{\phi}_{IJ}(\rho_{IJ}^{-1}(x))\colon O \to U_J$, then $\Lambda_{IJ}(x) = \Lambda_{\mathrm{d}_x\phi_{IJ}} \otimes (\Lambda_{[\widehat{\phi}_{IJ}]^{-1}})^*$ is identical to the map of the same name in [14, Equation 8.1.11], so that smoothness follows by the Claim proved as part of [14, Proposition 5.1.8]. This completes the proof that $\det(s_{\mathcal{K}})$ is a vector bundle over $\mathcal{K}$.

In the case of a weak Kuranishi cobordism $\mathcal{K}$, Proposition 8.1.8 in [14] also constructs smooth bundle isomorphisms from the collar restrictions to the product bundles $A^\alpha_\varepsilon \times \det(\mathfrak{s}_{\partial^\alpha\mathcal{K}})$ of the form

$$\widetilde{\iota}^{\Lambda,\alpha}_I(t,x) := (\Lambda_{\mathrm{d}_{(t,x)}\iota^\alpha_I} \circ \wedge_1) \otimes (\Lambda_{\mathrm{id}_{E_I}})^*\colon A^\alpha_\varepsilon \times \det(\mathrm{d}_x s^\alpha_I) \to \det(\mathrm{d}_{\iota^\alpha_I(x,t)} s_I), \tag{3.1.12}$$

where $\wedge_1\colon \Lambda^{\max}\ker \mathrm{d}_x s^\alpha_I \to \Lambda^{\max}(\mathbb{R} \times \ker \mathrm{d}_x s^\alpha_I)$ is given by $\eta \mapsto 1 \wedge \eta$. These are equivariant because they are induced by the equivariant map ι^α_I, and are compatible with the coordinate changes because the collar embeddings ι^α_I are. □

We next use the determinant bundle $\det(s_{\mathcal{K}})$ to define the notion of an orientation of a Kuranishi atlas.

Definition 3.1.10 A weak Kuranishi atlas or Kuranishi cobordism $\mathcal{K}$ is *orientable* if there exists a nonvanishing section σ of the bundle $\det(\mathfrak{s}_{\mathcal{K}})$, ie with $\sigma_I^{-1}(0) = \varnothing$ for all $I \in \mathcal{I}_{\mathcal{K}}$. An *orientation* of $\mathcal{K}$ is a choice of nonvanishing section σ of $\det(\mathfrak{s}_{\mathcal{K}})$. An *oriented Kuranishi atlas or cobordism* is a pair $(\mathcal{K}, \sigma)$ consisting of a Kuranishi atlas or cobordism and an orientation σ of $\mathcal{K}$.

For an oriented Kuranishi cobordism $(\mathcal{K}, \sigma)$ the *induced orientation of the boundary* $\partial^\alpha\mathcal{K}$ for $\alpha = 0, 1$ is the orientation of $\partial^\alpha\mathcal{K}$

$$\partial^\alpha\sigma := \big(((\widetilde{\iota}^{\Lambda,\alpha}_I)^{-1} \circ \sigma_I \circ \iota^\alpha_I)\big|_{\{\alpha\}\times\partial^\alpha U_I}\big)_{I \in \mathcal{I}_{\partial^\alpha\mathcal{K}}}$$

given by the isomorphism $(\widetilde{\iota}^{\Lambda,\alpha}_I)_{I \in \mathcal{I}_{\partial^\alpha\mathcal{K}}}$ in (3.1.12) between a collar neighborhood of the boundary in $\mathcal{K}$ and the product Kuranishi atlas $A^\alpha_\varepsilon \times \partial^\alpha\mathcal{K}$, followed by restriction to the boundary $\partial^\alpha\mathcal{K} = \partial^\alpha(A^\alpha_\varepsilon \times \partial^\alpha\mathcal{K})$, where we identify $\{\alpha\} \times \partial^\alpha U_I \cong \partial^\alpha U_I$.

With that, we say that two oriented weak Kuranishi atlases $(\mathcal{K}^0,\sigma^0)$ and $(\mathcal{K}^1,\sigma^1)$ are *oriented cobordant* if there exists a weak Kuranishi cobordism $\mathcal{K}$ from $\mathcal{K}^0$ to $\mathcal{K}^1$ and a section σ of $\det(\mathfrak{s}_{\mathcal{K}})$ such that $\partial^\alpha\sigma=\sigma^\alpha$ for $\alpha=0,1$.

Remark 3.1.11 Here we have defined the induced orientation on the boundary $\partial^\alpha\mathcal{K}$ of a cobordism so that it is completed to an orientation of the collar by adding the positive unit vector 1 along $A^\alpha_\varepsilon\subset\mathbb{R}$ rather than the more usual outward normal vector. In particular, by [14, Equation (8.1.12)], $\eta_1,\dots,\eta_n$ is a positively ordered basis for $\mathrm{T}_xU^\alpha_I$ exactly if $1,\eta_1,\dots,\eta_n$ is a positively ordered basis for $\mathrm{T}_x(A^\alpha_\varepsilon\times U^\alpha_I)$. ◇

Lemma 3.1.12 *Let $(\mathcal{K},\sigma)$ be an oriented weak Kuranishi atlas or cobordism.*

(i) *The orientation σ induces a canonical orientation $\sigma|_{\mathcal{K}'}:=(\sigma_I|_{U'_I})_{I\in\mathcal{I}_{\mathcal{K}'}}$ on each shrinking $\mathcal{K}'$ of $\mathcal{K}$ with domains $(U'_I\subset U_I)_{I\in\mathcal{I}_{\mathcal{K}'}}$.*

(ii) *In the case of a Kuranishi cobordism $\mathcal{K}$, the restrictions to boundary and shrinking commute; that is, $(\sigma|_{\mathcal{K}'})|_{\partial^\alpha\mathcal{K}'}=(\sigma|_{\partial^\alpha\mathcal{K}})|_{\partial^\alpha\mathcal{K}'}$.*

(iii) *In the case of a weak Kuranishi atlas $\mathcal{K}$, the orientation σ on $\mathcal{K}$ induces an orientation $\sigma^{[0,1]}$ on $[0,1]\times\mathcal{K}$, which induces the given orientation $\partial^\alpha\sigma^{[0,1]}=\sigma$ of the boundaries $\partial^\alpha([0,1]\times\mathcal{K})=\mathcal{K}$ for $\alpha=0,1$.*

Proof See the proof of [14, Lemma 8.1.11]. □

As in [14], in order to orient the zero sets of a perturbed section $\mathfrak{s}_{\mathcal{K}}+\nu$ we will work with a "more universal" determinant bundle $\det(\mathcal{K})$ over $\mathcal{K}$ that is constructed from the determinant bundles of the zero sections in each chart. Since the zero section $0_{\mathcal{K}}$ does not satisfy the index condition, we need to construct different transition maps for $\det(\mathcal{K})$, which will now depend on the section $\mathfrak{s}_{\mathcal{K}}$. For this purpose, we again use contraction isomorphisms from Lemma 3.1.7.

On the one hand, this provides families of isomorphisms

$$\mathfrak{C}_{\mathrm{d}_x s_I}\colon \Lambda^{\max}\mathrm{T}_xU_I\otimes(\Lambda^{\max}E_I)^*\xrightarrow{\cong}\det(\mathrm{d}_x s_I)\quad\text{for } x\in U_I, \tag{3.1.13}$$

which, by Remark 3.1.8, are equivariant with the respect to the action of $\gamma\in\Gamma_I$ on $\Lambda^{\max}\mathrm{T}U_I\otimes(\Lambda^{\max}E_I)^*$ given by

$$\begin{aligned}\widehat{\gamma}_\Lambda:=\Lambda_{\mathrm{d}_x\gamma}\otimes(\Lambda_{\gamma^{-1}})^*\colon \Lambda^{\max}\mathrm{T}_xU_I\otimes(\Lambda^{\max}E_I)^*&\\ \to\Lambda^{\max}\mathrm{T}_{\gamma x}U_I\otimes(\Lambda^{\max}E_I)^*&\end{aligned} \tag{3.1.14}$$

and the corresponding action on $\det(\mathrm{d}_x s_I)$ in Equation (3.1.3).

On the other hand, recall that the tangent bundle condition (2.2.3) implies that $\mathrm{d}s_J$ restricts to an isomorphism

$$\mathrm{T}_y U_J / \mathrm{d}_{\widetilde{x}} \widetilde{\phi}_{IJ}(\mathrm{T}_{\widetilde{x}} \widetilde{U}_{IJ}) \xrightarrow{\cong} E_J / \widehat{\phi}_{IJ}(E_I)$$

for $y = \widetilde{\phi}_{IJ}(\widetilde{x})$.[7] Therefore, if we choose a Γ_J–equivariant smooth normal bundle

$$N_{IJ} = \bigcup_{y \in \mathrm{im}\, \widetilde{\phi}_{IJ}} N_{IJ,y} \subset \mathrm{T}_y U_J$$

to the submanifold $\mathrm{im}\, \widetilde{\phi}_{IJ} \subset U_J$, then the subspaces $\mathrm{d}_y s_J(N_{IJ,y})$ form a smooth family of subspaces of E_J that are complements to $\widehat{\phi}_{IJ}(E_I)$. Hence, if we write $\mathrm{pr}_{N_{IJ}}(y)\colon E_J \to \mathrm{d}_y s_J(N_{IJ,y}) \subset E_J$ for the smooth family of projections with kernel $\widehat{\phi}_{IJ}(E_I)$, we obtain a smooth family of linear maps

$$F_{\widetilde{x}} := \mathrm{pr}_{N_{IJ}}(y) \circ \mathrm{d}_y s_J \colon \mathrm{T}_y U_J \longrightarrow E_J \quad \text{for } y = \widetilde{\phi}_{IJ}(\widetilde{x}),$$

with images $\mathrm{im}\, F_{\widetilde{x}} = \mathrm{d}_y s_J(N_{IJ,y})$, and also isomorphisms to their kernel

$$\phi_{\widetilde{x}} := \mathrm{d}_{\widetilde{x}} \widetilde{\phi}_{IJ} \colon \mathrm{T}_{\widetilde{x}} \widetilde{U}_{IJ} \xrightarrow{\cong} \ker F_{\widetilde{x}} = \mathrm{T}_y(\mathrm{im}\, \widetilde{\phi}_{IJ}) \subset \mathrm{T}_y U_J.$$

By Lemma 3.1.7 these induce isomorphisms

$$\mathfrak{C}^{\phi_{\widetilde{x}}}_{F_{\widetilde{x}}} \colon \Lambda^{\max} \mathrm{T}_{\widetilde{\phi}_{IJ}(\widetilde{x})} U_J \otimes (\Lambda^{\max} E_J)^* \xrightarrow{\cong} \Lambda^{\max} \mathrm{T}_{\widetilde{x}} \widetilde{U}_{IJ} \otimes (\Lambda^{\max}(E_J / \mathrm{im}\, F_{\widetilde{x}}))^*.$$

We may combine this with the isomorphism $\Lambda^{\max} \mathrm{T}_{\widetilde{x}} \widetilde{U}_{IJ} \to \rho_{IJ}^*(\Lambda^{\max} \mathrm{T}_x U_I)$ induced by $\mathrm{d}_{\widetilde{x}} \rho_{IJ}$, where $x := \rho_{IJ}(\widetilde{x})$, and the dual of the isomorphism

$$\Lambda^{\max}(E_J / \mathrm{d}_y s_J(N_{IJ,y})) \cong \Lambda^{\max} E_I$$

induced via (3.1.2) by $\mathrm{pr}^{\perp}_{N_{IJ}}(y) \circ \widehat{\phi}_{IJ} \colon E_I \to E_J / \mathrm{d}_y s_J(N_{IJ,y})$, to obtain for each $\widetilde{x} \in \widetilde{U}_{IJ}$ an isomorphism

$$\widetilde{\mathfrak{C}}_{IJ}(\widetilde{x}) \colon \Lambda^{\max} \mathrm{T}_y U_J \otimes (\Lambda^{\max} E_J)^* \xrightarrow{\cong} \rho_{IJ}^*(\Lambda^{\max} \mathrm{T}_x U_I) \otimes (\Lambda^{\max} E_I)^* \tag{3.1.15}$$

with

$$y := \widetilde{\phi}_{IJ}(\widetilde{x}), \quad x := \rho_{IJ}(\widetilde{x}),$$

given by the composite of $\mathfrak{C}^{\phi_{\widetilde{x}}}_{F_{\widetilde{x}}}$ with the map

$$(\Lambda_{\mathrm{d}_{\widetilde{x}} \rho_{IJ}}) \otimes \big(\Lambda_{(\mathrm{pr}^{\perp}_{N_{IJ}}(y) \circ \widehat{\phi}_{IJ})^{-1}}\big)^* \colon \Lambda^{\max} \mathrm{T}_{\widetilde{x}} \widetilde{U}_{IJ} \otimes (\Lambda^{\max}(E_J / \mathrm{im}\, F_{\widetilde{x}}))^* \\ \to \rho_{IJ}^*(\Lambda^{\max} \mathrm{T}_x U_I) \otimes (\Lambda^{\max} E_I)^*.$$

[7] Here and subsequently, we will distinguish between the manifold $\widetilde{U}_{IJ}$ and its image $\mathrm{im}\, \widetilde{\phi}_{IJ}$ in U_J, denoting points of $\widetilde{U}_{IJ}$ by $\widetilde{x}$, with $y = \widetilde{\phi}_{IJ}(x) \in U_J$ and $x = \rho_{IJ}(\widetilde{x}) \in U_{IJ}$.

Proposition 3.1.13 (i) *Let $\mathcal{K}$ be a weak Kuranishi atlas. Then there is a well-defined line bundle* $\det(\mathcal{K})$ *over* $\mathcal{K}$ *given by the line bundles*

$$\Lambda_I^{\mathcal{K}} := \Lambda^{\max}\mathrm{T}U_I \otimes (\Lambda^{\max} E_I)^* \to U_I \quad \text{for } I \in \mathcal{I}_{\mathcal{K}},$$

with group actions as in (3.1.14) and the transition maps

$$\widetilde{\mathfrak{C}}_{IJ}^{-1}\colon \rho_{IJ}^*(\Lambda_I^{\mathcal{K}}|_{U_{IJ}}) \to \Lambda_J^{\mathcal{K}}|_{\mathrm{im}\,\widetilde{\phi}_{IJ}}$$

from (3.1.15) for $I \subsetneq J$. In particular, the latter isomorphisms are independent of the choice of normal bundle N_{IJ}.
Furthermore, the contractions $\mathfrak{C}_{\mathrm{d}s_I}\colon \Lambda_I^{\mathcal{K}} \to \det(\mathrm{d}s_I)$ from (3.1.13) define an isomorphism $\Psi^{s_{\mathcal{K}}} := (\mathfrak{C}_{\mathrm{d}s_I})_{I\in\mathcal{I}_{\mathcal{K}}}$ from $\det(\mathcal{K})$ to $\det(\mathfrak{s}_{\mathcal{K}})$.

(ii) *If $\mathcal{K}$ is a weak Kuranishi cobordism, then the determinant bundle $\det(\mathcal{K})$ defined as in (i) can be given a product structure on the collar so that its boundary restrictions are $\det(\mathcal{K})|_{\partial^\alpha\mathcal{K}} = \det(\partial^\alpha\mathcal{K})$ for $\alpha = 0, 1$.*
Further, the isomorphism $\Psi^{s_{\mathcal{K}}}\colon \det(\mathcal{K}) \to \det(s_{\mathcal{K}})$ defined as in (i) has product structure on the collar with restrictions $\Psi^{s_{\mathcal{K}}}|_{\partial^\alpha\mathcal{K}} = \Psi^{s_{\partial^\alpha\mathcal{K}}}$ for $\alpha = 0, 1$.

Proof To begin, note that each $\Lambda_I^{\mathcal{K}} = \Lambda^{\max}\mathrm{T}U_I \otimes (\Lambda^{\max} E_I)^*$ is a smooth line bundle over U_I, since it inherits local trivializations from the tangent bundle $\mathrm{T}U_I \to U_I$. Moreover the action of Γ_I on $U_I \times E_I$ induces a smooth action on $\Lambda_I^{\mathcal{K}}$ given by (3.1.14) that covers its action on U_I. Thus $\Lambda_I^{\mathcal{K}} \to U_I$ is a smooth Γ_I–equivariant bundle. We showed in [14, Proposition 8.1.12] that the isomorphisms $\mathfrak{C}_{\mathrm{d}_x s_I}$ from (3.1.13) are smooth in this trivialization, where $\det(\mathrm{d}s_I)$ is trivialized via the maps $\widehat{T}_{I,x}$ as in Proposition 3.1.9. Since $\mathfrak{C}_{\mathrm{d}s_I}$ is equivariant, we can define preliminary transition maps

$$\widetilde{\phi}_{IJ}^{\Lambda} := \mathfrak{C}_{\mathrm{d}s_J}^{-1} \circ \widetilde{\Lambda}_{IJ} \circ \rho_{IJ}^*(\mathfrak{C}_{\mathrm{d}s_I})\colon \rho_{IJ}^*(\Lambda_I^{\mathcal{K}}|_{U_{IJ}}) \to \Lambda_J^{\mathcal{K}} \quad \text{for } I \subsetneq J \in \mathcal{I}_{\mathcal{K}} \tag{3.1.16}$$

by the transition maps (3.1.5) of $\det(\mathfrak{s}_{\mathcal{K}})$, the isomorphisms (3.1.13) and the pullback by ρ_{IJ}. These define a line bundle

$$\Lambda^{\mathcal{K}} := (\Lambda_I^{\mathcal{K}}, \widetilde{\phi}_{IJ}^{\Lambda})_{I,J\in\mathcal{I}_{\mathcal{K}}}$$

since the weak cocycle condition follows directly from that for the $\widetilde{\Lambda}_{IJ}$. Moreover, this automatically makes the family of bundle isomorphisms $\Psi^{\mathcal{K}} := (\widetilde{\mathfrak{C}}_{\mathrm{d}s_I})_{I\in\mathcal{I}_{\mathcal{K}}}$ an isomorphism from $\Lambda^{\mathcal{K}}$ to $\det(\mathfrak{s}_{\mathcal{K}})$. It remains to see that $\Lambda^{\mathcal{K}} = \det(\mathcal{K})$ and $\Psi^{\mathcal{K}} = \Psi^{\mathfrak{s}_{\mathcal{K}}}$, ie we claim equality of transition maps $\widetilde{\phi}_{IJ}^{\Lambda} = \widetilde{\mathfrak{C}}_{IJ}^{-1}$. This also shows that $\widetilde{\mathfrak{C}}_{IJ}^{-1}$ and thus $\det(\mathcal{K})$ is independent of the choice of normal bundle N_{IJ} in (3.1.15).

So to finish the proof of (i), it suffices to establish the following commuting diagram at a fixed $\widetilde{x} \in U_{IJ}$ with $x = \rho_{IJ}(\widetilde{x})$, $y = \widetilde{\phi}_{IJ}(\widetilde{x})$:

$$\begin{tikzcd}
\Lambda^{\max}\mathrm{T}_x U_I \otimes (\Lambda^{\max} E_I)^* \arrow[rr, "\mathfrak{C}_{\mathrm{d}_x s_I}"] & & \det(\mathrm{d}_x s_I) \\
\rho_{IJ}^*(\Lambda^{\max}\mathrm{T}_x U_I) \otimes (\Lambda^{\max} E_I)^* \arrow[u, "\rho_{IJ}"] \arrow[rr, "\rho_{IJ}^*(\mathfrak{C}_{\mathrm{d}_x s_I})"] & & \rho_{IJ}^*(\det(\mathrm{d}_x s_I)) \arrow[u, "\rho_{IJ}"'] \arrow[d, "\widetilde{\Lambda}_{IJ}(\widetilde{x})"] \\
\Lambda^{\max}\mathrm{T}_y U_J \otimes (\Lambda^{\max} E_J)^* \arrow[u, "\widetilde{\mathfrak{C}}_{IJ}(\widetilde{x})"] \arrow[rr, "\mathfrak{C}_{\mathrm{d}_y s_J}"] & & \det(\mathrm{d}_y s_J)
\end{tikzcd} \tag{3.1.17}$$

However, the composition $y \mapsto \rho_{IJ} \circ \widetilde{\mathfrak{C}}_{IJ}(\widetilde{\phi}_{IJ}^{-1}(\widetilde{x}))$ of the left-hand vertical maps is precisely the map denoted by $y \mapsto \mathfrak{C}_{IJ}(x)$ in [14, Equation (8.1.15)], while, as in the proof of Proposition 3.1.9 above, the right-hand vertical maps combine to $\Lambda_{IJ}(x) = \widetilde{\Lambda}(\rho_{IJ}^{-1}(x))\colon \det(\mathrm{d}_x s_I) \to \det(\mathrm{d}_y s_J)$, where ρ_{IJ}^{-1} is a local inverse to ρ_{IJ}. Therefore the desired result follows from the commutativity of the diagram

$$\begin{tikzcd}
\Lambda^{\max}\mathrm{T}_x U_I \otimes (\Lambda^{\max} E_I)^* \arrow[r, "\mathfrak{C}_{\mathrm{d}_x s_I}"] & \det(\mathrm{d}_x s_I) \arrow[d, "\Lambda_{IJ}(x)"] \\
\Lambda^{\max}\mathrm{T}_y U_J \otimes (\Lambda^{\max} E_J)^* \arrow[u, "\mathfrak{C}_{IJ}(x)"] \arrow[r, "\mathfrak{C}_{\mathrm{d}_y s_J}"] & \det(\mathrm{d}_y s_J)
\end{tikzcd}$$

which is established in [14, Proposition 8.1.12].

For part (ii) the same arguments apply to define a bundle $\det(\mathcal{K})$ and isomorphism $\Psi^{\mathfrak{s}_{\mathcal{K}}}$. The required product structure on a collar follows as in [14]. □

We end this section by explaining how orientations of a Kuranishi atlas induce compatible orientations on local zero sets of transverse sections.

Lemma 3.1.14 *Let $(\mathcal{K}, \sigma)$ be a d–dimensional oriented, tame Kuranishi atlas or cobordism, and for some $I \in \mathcal{I}_{\mathcal{K}}$ let $f\colon W \to E_I$ be a smooth section over an open subset $W \subset U_I$ that is transverse to 0.*

(i) *The zero set $Z_f := f^{-1}(0) \subset U_I$ inherits the structure of a smooth oriented d–dimensional submanifold.*
(ii) *The action of any $\gamma \in \Gamma_I$ on U_I induces an orientation-preserving diffeomorphism $Z_f \to Z_{\gamma * f}$ to the zero set of $\gamma * f\colon \gamma(W) \to E_I$, $x \mapsto \gamma f(\gamma^{-1}(x))$.*
(iii) *Suppose further that $f(W) \subset \widehat{\phi}_{HI}(E_H)$, $\widetilde{W}_{HI} := W \cap \widetilde{U}_{HI} \neq \varnothing$ and $\rho_{HI}|_{\widetilde{W}_{HI}}$ is injective for some $H \subset I$. Then ρ_{HI} induces an orientation-preserving diffeomorphism $Z_f \to Z_{\rho_{HI} * f}$ to the zero set of $\rho_{HI} * f\colon \rho_{HI}(W) \to E_H$, $x \mapsto \widehat{\phi}_{HI}^{-1}(f(\rho_{HI}^{-1}(x)))$.*

(iv) *If $\mathcal{K}$ is a cobordism, suppose in addition that $\boldsymbol{K}_I$ is a chart that intersects the boundary $\partial^\alpha\mathcal{K}$, with $W = \iota_I^\alpha(A_\varepsilon^\alpha \times W^\alpha)$ for some $W^\alpha \subset \partial^\alpha W$, and $f(\iota_I^\alpha(t,x)) = f^\alpha(x)$ for some transverse section $f^\alpha\colon W^\alpha \to E_I$. Then the atlas $(\partial^\alpha\mathcal{K}, \partial^\alpha\sigma)$ induces an oriented smooth structure on $Z_{f^\alpha} \subset W^\alpha$ by* (i)*, $Z_f \subset U_I$ is a submanifold with boundary and $j_I^\alpha := \iota_I(\alpha,\cdot)$ is a diffeomorphism $Z_{f^\alpha} \to \partial Z_f$ that preserves (resp. reverses) orientations when $\alpha = 1$ (resp. $\alpha = 0$).*

Proof Except for (ii) these local claims follow directly from the corresponding parts of the proof of [14, Proposition 8.1.13]. For (iii) note that the injectivity assumption allows us to write $\rho_{HI}{*}f = \phi_{HI}^* f$ for an embedding $\phi_{HI}\colon \rho_{HI}(\widetilde{W}_{HI}) \to W$. Before we can prove (ii), recall that the orientation on Z_f is induced from the orientation of the Kuranishi atlas/cobordism $\sigma_I\colon U_I \to \det(\mathrm{d}s_I)$ via the isomorphisms for $z \in Z_f$,

$$\Lambda^{\max}\mathrm{T}_z Z_f = \Lambda^{\max}\ker \mathrm{d}_z f \cong \Lambda^{\max}\ker \mathrm{d}_z f \otimes \mathbb{R} = \det(\mathrm{d}_z f),$$
$$\mathfrak{C}_{\mathrm{d}_z f}\colon \det(\mathrm{d}_z f) \to \Lambda^{\max}\mathrm{T}_z U_I \otimes (\Lambda^{\max} E_I)^*,$$
$$\mathfrak{C}_{\mathrm{d}_z s_I}\colon \det(\mathrm{d}_z s_I) \to \Lambda^{\max}\mathrm{T}_z U_I \otimes (\Lambda^{\max} E_I)^*.$$

Now to prove that $\gamma \in \Gamma_I$ acts by an orientation-preserving diffeomorphism, note that a smooth group action always acts by diffeomorphisms. Restriction to Z_f of the action by $\gamma \in \Gamma_I$ thus yields a diffeomorphism to its image, which is easily seen to be the zero set of $\gamma{*}f$. To show that this diffeomorphism is compatible with the induced orientations at $z \in Z_f$ and $\gamma z \in Z_{\gamma*f}$, we begin by noting that the action of γ is $\Lambda_{\mathrm{d}_z\gamma}\colon \Lambda^{\max}\mathrm{T}_z Z_f \to \Lambda^{\max}\mathrm{T}_{\gamma z} Z_{\gamma*f}$. On the other hand, the orientations $\sigma_I(z)$ and $\sigma_I(\gamma z)$ are by assumption intertwined by the isomorphism $\Lambda_{\mathrm{d}_z\gamma} \otimes (\Lambda_{[\gamma^{-1}]})^*\colon \det \mathrm{d}_z s_I \to \det \mathrm{d}_{\gamma z} s_I$, and by Proposition 3.1.13(i) this implies that their pullbacks to $\Lambda^{\max}\mathrm{T}_x U_I \otimes (\Lambda^{\max} E_I)^*$ for $x = z, \gamma z$ are intertwined by $\Lambda_{\mathrm{d}_z\gamma} \otimes (\Lambda_{\gamma^{-1}})^*$. Thus it remains to prove that the following diagram commutes:

$$\begin{tikzcd}
\Lambda^{\max}\mathrm{T}_z U_I \otimes (\Lambda^{\max} E_I)^* \arrow[r, "\mathfrak{C}_{\mathrm{d}_z f}"] \arrow[d, "\Lambda_{\mathrm{d}_z\gamma}\otimes(\Lambda_{\gamma^{-1}})^*"'] & \Lambda^{\max}\ker(\mathrm{d}_z f) \otimes \mathbb{R} \cong \Lambda^{\max}\mathrm{T}_z Z_f \arrow[d, "\Lambda_{\mathrm{d}_z\gamma}\otimes\mathrm{id}_{\mathbb{R}}"] \\
\Lambda^{\max}\mathrm{T}_{\gamma z} U_I \otimes (\Lambda^{\max} E_I)^* \arrow[r, "\mathfrak{C}_{\mathrm{d}_{\gamma z}(\gamma*f)}"] & \Lambda^{\max}\ker(\mathrm{d}_{\gamma z}(\gamma * f)) \otimes \mathbb{R} \cong \Lambda^{\max}\mathrm{T}_{\gamma z} Z_{\gamma*f}
\end{tikzcd}$$

By Lemma 3.1.7, $\mathfrak{C}_{\mathrm{d}_z f}$ is given by $(v_1\wedge\cdots\wedge v_n)\otimes(w_1\wedge\cdots\wedge w_m)^* \mapsto v_1\wedge\cdots\wedge v_k$, where $v_1,\dots,v_n$ is any basis for $\mathrm{T}_z U_I$ whose first k elements span $\ker \mathrm{d}_z f$, and $w_1,\dots,w_m$ is a basis for E_I, and similarly for $\mathfrak{C}_{\mathrm{d}_{\gamma z}(\gamma*f)}$. Therefore, if we denote $v_i' := \mathrm{d}_z\gamma(v_i)$ and $w_j' := \gamma w_j$, we find that $(\Lambda_{\gamma^{-1}})^*(w_1\wedge\cdots\wedge w_m)^* = (w_1'\wedge\cdots\wedge w_m')^*$

and thus the diagram commutes as required:

$$\begin{aligned}
&\mathfrak{C}_{\mathrm{d}_{\gamma z}(\gamma * f)}\big(\Lambda_{\mathrm{d}_z\gamma}\otimes(\Lambda_{\gamma^{-1}})^*((v_1\wedge\cdots\wedge v_n)\otimes(w_1\wedge\cdots\wedge w_m)^*)\big)\\
&\qquad\qquad = \mathfrak{C}_{\mathrm{d}_{\gamma z}(\gamma * f)}((v_1'\wedge\cdots\wedge v_n')\otimes(w_1'\wedge\cdots\wedge w_m')^*)\\
&\qquad\qquad = v_1'\wedge\cdots\wedge v_k'\\
&\qquad\qquad = \Lambda_{\mathrm{d}_z\gamma}(v_1\wedge\cdots\wedge v_k)\\
&\qquad\qquad = \Lambda_{\mathrm{d}_z\gamma}\big(\mathfrak{C}_{\mathrm{d}_z f}((v_1\wedge\cdots\wedge v_n)\otimes(w_1\wedge\cdots\wedge w_m)^*)\big).
\end{aligned}$$

□

3.2 Perturbed zero sets

With Theorem 2.5.3 providing existence and uniqueness of tame shrinkings, the second part of the proof of Theorem A is the construction of the VMC/VFC from the zero sets of suitable perturbations $\mathfrak{s}_{\mathcal{K}}+\nu$ of the canonical section $\mathfrak{s}_{\mathcal{K}}$ of a tame Kuranishi atlas or cobordism. In this section, we describe a suitable class of perturbations ν, and prove that the corresponding perturbed zero sets are compact weighted branched manifolds, a notion from [8] that we review in the appendix. The existence and uniqueness of such perturbations will be established in Section 3.3, as part of the perturbative construction of VMC and VFC. The main work is done by the setup in this section, which will put us into a situation in which the construction of perturbations and the resulting VMC/VFC can essentially be copied from [14]. Since the construction of perturbations requires tameness and the notion of weighted branched manifolds requires an orientation in [8], we will — unless otherwise stated — work with an oriented tame Kuranishi atlas or cobordism $\mathcal{K}$.

As in the case of trivial isotropy, one cannot in general find transverse perturbations $s_I+\nu_I \pitchfork 0$ that are also compatible with the coordinate changes $\widehat{\Phi}_{IJ}$. Instead, we will construct perturbations over the following notion of a reduced atlas that still covers X but generally does not form a Kuranishi atlas.

Definition 3.2.1 [13, Definition 5.1.2] A *(cobordism) reduction* of a tame Kuranishi atlas or cobordism $\mathcal{K}$ is an open subset $\mathcal{V}=\bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}}V_I\subset \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}}$, ie a tuple of (possibly empty) open subsets $V_I\subset U_I$ satisfying the following conditions:

(i) $V_I=\pi_I^{-1}(\underline{V}_I)$ for each $I\in\mathcal{I}_{\mathcal{K}}$, ie V_I is pulled back from the intermediate category and so is Γ_I–invariant.

(ii) $V_I\sqsubset U_I$ for all $I\in\mathcal{I}_{\mathcal{K}}$, and if $V_I\neq\varnothing$ then $V_I\cap s_I^{-1}(0)\neq\varnothing$.

(iii) If $\pi_{\mathcal{K}}(\overline{V_I})\cap\pi_{\mathcal{K}}(\overline{V_J})\neq\varnothing$ then $I\subset J$ or $J\subset I$.

(iv) The zero set $\iota_{\mathcal{K}}(X)=|s_{\mathcal{K}}|^{-1}(0)$ is contained in $\pi_{\mathcal{K}}(\mathcal{V})=\bigcup_{I\in\mathcal{I}_{\mathcal{K}}}\pi_{\mathcal{K}}(V_I)$.

If $\mathcal{K}$ is a cobordism, we require in addition that $\mathcal{V}$ is collared in the following sense:

(v) For each $\alpha \in \{0,1\}$ and $I \in \mathcal{I}_{\partial^\alpha \mathcal{K}} \subset \mathcal{I}_{\mathcal{K}}$, there exists an $\varepsilon > 0$ and a subset $\partial^\alpha V_I \subset \partial^\alpha U_I$ such that $\partial^\alpha V_I \neq \varnothing$ if and only if $V_I \cap \psi_I^{-1}(\partial^\alpha F_I) \neq \varnothing$, and

$$(\iota_I^\alpha)^{-1}(V_I) \cap (A_\varepsilon^\alpha \times \partial^\alpha U_I) = A_\varepsilon^\alpha \times \partial^\alpha V_I.$$

We call $\partial^\alpha \mathcal{V} := \bigsqcup_{I \in \mathcal{I}_{\partial 0_{\mathcal{K}}}} \partial^\alpha V_I \subset \mathrm{Obj}_{\boldsymbol{B}_{\partial^\alpha \mathcal{K}}}$ the *boundary restriction* of $\mathcal{V}$ to $\partial^\alpha \mathcal{K}$.

Remark 3.2.2 (i) The notion of (cobordism) reduction is equivalent to saying that $\mathcal{V} := \bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} V_I \subset \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}}$ is the lift $V_I := \pi_I^{-1}(\underline{V}_I)$ of a (cobordism) reduction $\underline{\mathcal{V}} = \bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} \underline{V}_I \subset \mathrm{Obj}_{\boldsymbol{B}_{\underline{\mathcal{K}}}}$ of the intermediate Kuranishi atlas/cobordism. Thus existence and uniqueness of reductions is proven in [13, Theorem 5.1.6].

(ii) The restrictions $\partial^\alpha \mathcal{V}$ of a cobordism reduction $\mathcal{V}$ of a Kuranishi cobordism $\mathcal{K}$ are reductions of the restricted Kuranishi atlases $\partial^\alpha \mathcal{K}$ for $\alpha = 0, 1$. In particular, condition (ii) holds because part (v) of Definition 3.2.1 implies that if $\partial^\alpha V_I \neq \varnothing$ then $\partial^\alpha V_I \cap \psi_I^{-1}(\partial^\alpha F_I) \neq \varnothing$. Note that condition (v) also implies that $V_I \subset U_I$ is a collared subset in the sense of (2.4.1). ◇

Given a reduction $\mathcal{V}$, we define the *reduced domain category* $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}$ and the *reduced obstruction category* $\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}$ to be the full subcategories of $\boldsymbol{B}_{\mathcal{K}}$ and $\boldsymbol{E}_{\mathcal{K}}$ with objects $\bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} V_I$ and $\bigsqcup_{I \in \mathcal{I}_{\mathcal{K}}} V_I \times E_I$ respectively, and denote by $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}} \colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}$ the section given by restriction of $\mathfrak{s}_{\mathcal{K}}$. Now one might hope to find transverse perturbation functors $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}} + \nu \colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}$ by iteratively constructing $\nu_I \colon V_I \to E_I$ as in [14], where compatibility with the morphisms can be ensured by working along the partial order $\subsetneq$ on $\mathcal{I}_{\mathcal{K}}$, using the separation property (iii) of a reduction. However, we also have to ensure compatibility with the morphisms given by the action of nontrivial isotropy groups Γ_I. Depending on their action, we might not even be able to even find a Γ_I–equivariant perturbation ν_I in a single chart such that $s_I + \nu_I \pitchfork 0$. In general, this can be resolved by using multivalued perturbations such as in the perturbative construction of the Euler class of an orbibundle, explained for example in [6] as motivation for perturbations in Kuranishi structures. We could also formulate our perturbation scheme in these terms, but due to the particularly simple setup — notably additivity $\Gamma_I = \prod_{i \in I} \Gamma_i$ of the isotropy groups — we can construct the "multivalued perturbations" as single-valued section functors $\nu \colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ over a *pruned domain category* $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$, which is obtained in Lemma 3.2.3 from the reduced domain category $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}$ by forgetting sufficiently many morphisms to obtain trivial isotropy. It is to this category that the iterative perturbation scheme of [14] will be applied in Section 3.3 to obtain a suitable class of transverse perturbations ν. Once a zero set is cut out transversely from $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$, we will then show in Theorem 3.2.8 that adding some of the isotropy morphisms back in — at the expense of adding weights to corresponding branches of the solution set — yields the structure of a weighted branched manifold on the Hausdorff quotient of the perturbed solution set $\big|(\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}+\nu)^{-1}(0)\big| \subset \big|\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}\big|$. This perturbed solution set is not a subset of the virtual neighborhood $|\mathcal{K}|$, but its Hausdorff quotient supports a fundamental class by Proposition A.7, and the inclusion $\iota^\nu\colon (\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}+\nu)^{-1}(0) \to \mathrm{Obj}_{\mathcal{B}_{\mathcal{K}}}$ induces a continuous map $|\iota^\nu|_{\mathcal{H}}\colon \big|(\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}+\nu)^{-1}(0)\big|_{\mathcal{H}} \to |\mathcal{K}|$ that will represent the virtual fundamental cycle of $\mathcal{K}$.

We will describe the pruned categories in terms of the sets

$$\widetilde{V}_{IJ} := V_J \cap \rho_{IJ}^{-1}(V_I) = V_J \cap \pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(V_I)) \subset \widetilde{U}_{IJ}.$$

Note that $\widetilde{V}_{IJ}$ is invariant under the action of Γ_J, and is an open subset of the closed submanifold $\widetilde{U}_{IJ} = s_J^{-1}(E_I)$ of V_J, where the last equality holds by the tameness condition (2.5.2). Further if $F \subset I \subset J$,

$$V_J \cap \rho_{IJ}^{-1}(\widetilde{V}_{FI}) = \widetilde{V}_{IJ} \cap \widetilde{V}_{FJ} = V_J \cap \pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(V_I) \cap \pi_{\mathcal{K}}(V_F)) \subset \widetilde{U}_{FJ}. \tag{3.2.1}$$

In fact, the second equality above holds for any pair of subsets $F, I \subset J$. However, because $\mathcal{V}$ is a reduction, the intersection is empty unless F and I are nested, ie either $F \subset I$ or $I \subset F$. Finally, the group $\Gamma_{I\setminus F}$ acts freely on $\widetilde{U}_{FI}$ (by Definition 2.2.8 for a coordinate change), and hence also on $\widetilde{V}_{FI}$. If $I = F$ we define $\Gamma_{I\setminus F} := \Gamma_\varnothing := \{\mathrm{id}\}$.

Lemma 3.2.3 *Let $\mathcal{V}$ be a (cobordism) reduction of a tame Kuranishi atlas or cobordism $\mathcal{K}$. Then there are well-defined categories — the* ***pruned domain category*** $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ *and the* ***pruned obstruction category*** $\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ *— obtained from $\boldsymbol{B}_{\mathcal{K}}$ and $\boldsymbol{E}_{\mathcal{K}}$ as follows:*

- *Object spaces are given by restriction to the reduction* $\mathcal{V} = \bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}} V_I \subset \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}}$:

$$\mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}} := \bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}} V_I \subset \mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}}, \quad \mathrm{Obj}_{\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}} := \bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}} V_I \times E_I \subset \mathrm{Obj}_{\boldsymbol{E}_{\mathcal{K}}}.$$

- *Morphism spaces are open subsets of* $\mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}}$ *and* $\mathrm{Mor}_{\boldsymbol{E}_{\mathcal{K}}}$ *respectively, with components*

$$\mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}} := \bigsqcup_{I,J\in\mathcal{I}_{\mathcal{K}}} \mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}(V_I, V_J), \quad \mathrm{Mor}_{\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}} := \bigsqcup_{I,J\in\mathcal{I}_{\mathcal{K}}} \mathrm{Mor}_{\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}(V_I, V_J),$$

  given by $\mathrm{Mor}_{\dots}(V_I, V_J) = \varnothing$ *unless $I \subset J$, in which case the morphisms are given in terms of the open subsets* $\widetilde{V}_{IJ} := V_J \cap \rho_{IJ}^{-1}(V_I) \subset \widetilde{U}_{IJ}$ *as*

$$\mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}(V_I, V_J) := \widetilde{V}_{IJ} \times \{\mathrm{id}\} \subset \widetilde{U}_{IJ} \times \Gamma_I = \mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}}(U_I, U_J),$$
$$\mathrm{Mor}_{\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}(V_I, V_J) := \widetilde{V}_{IJ} \times E_I \times \{\mathrm{id}\} \subset \widetilde{U}_{IJ} \times E_I \times \Gamma_I = \mathrm{Mor}_{\boldsymbol{E}_{\mathcal{K}}}(U_I, U_J).$$

- *All structure maps (source, target, identity, and composition) are given by restriction of the respective structure maps of $\boldsymbol{B}_{\mathcal{K}}$ and $\boldsymbol{E}_{\mathcal{K}}$ in Definition 2.3.5.*

These pruned categories are nonsingular in the sense that there is at most one morphism between any two objects. Moreover, the projection and section functors $\mathrm{pr}_{\mathcal{K}}\colon \boldsymbol{E}_{\mathcal{K}} \to \boldsymbol{B}_{\mathcal{K}}$ *and* $\mathfrak{s}_{\mathcal{K}}\colon \boldsymbol{B}_{\mathcal{K}} \to \boldsymbol{E}_{\mathcal{K}}$ *restrict to well-defined functors* $\mathrm{pr}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ *and* $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ *with* $\mathrm{pr}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} \circ \mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} = \mathrm{id}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}$.

Proof Recall that $(I, J, y, \mathrm{id}) \in \mathrm{Mor}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}$ has source $(I, \rho_{IJ}(y))$ and target (J, y) (where, as in Lemma 2.3.6, we suppress mention of the inclusion $\tilde{\phi}_{IJ}$). Now morphisms are closed under composition because the strong cocycle condition guarantees that $\rho_{IJ} \circ \rho_{JK} = \rho_{IK}$, with identical domains whenever $I \subset J \subset K$. Moreover, the category is nonsingular because source and target determine the morphism uniquely. Similar arguments apply to $\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$. Finally, the projection and section functors of $\mathcal{K}$ act trivially on the isotropy groups Γ_I, and thus restrict to well-defined functors when we drop these. □

The following combines Definitions 7.2.1, 7.2.5, 7.2.6 and 7.2.9 from [14].

Definition 3.2.4 A *(cobordism) perturbation* of $\mathcal{K}$ is a smooth functor

$$\nu\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$$

between the pruned domain and obstruction categories of some (cobordism) reduction $\mathcal{V}$ of $\mathcal{K}$, such that $\mathrm{pr}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} \circ \nu = \mathrm{id}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}$.

That is, $\nu = (\nu_I)_{I \in \mathcal{I}_{\mathcal{K}}}$ is given by a family of smooth maps $\nu_I\colon V_I \to E_I$ that are compatible with coordinate changes in the sense that for all $I \subsetneq J$ we have

$$\nu_J|_{\widetilde{V}_{IJ}} = \hat{\phi}_{IJ} \circ \nu_I \circ \rho_{IJ}|_{\widetilde{V}_{IJ}} \quad \text{on} \quad \widetilde{V}_{IJ} = V_J \cap \rho_{IJ}^{-1}(V_I). \tag{3.2.2}$$

If $\mathcal{K}$ is a Kuranishi cobordism we require in addition that ν has product form in a collar neighborhood of the boundary. That is, for $\alpha = 0, 1$ and $I \in \mathcal{I}_{\mathcal{K}^\alpha} \subset \mathcal{I}_{\mathcal{K}}$ there is an $\varepsilon > 0$ and a map $\nu_I^\alpha\colon \partial^\alpha V_I \to E_I$ such that

$$\nu_I(\iota_I^\alpha(t, x)) = \nu_I^\alpha(x) \quad \forall x \in \partial^\alpha V_I,\ t \in A_\varepsilon^\alpha.$$

We say that a (cobordism) perturbation ν is

- *admissible* if we have $\mathrm{d}_y \nu_J(\mathrm{T}_y V_J) \subset \mathrm{im}\, \hat{\phi}_{IJ}$ for all $I \subsetneq J$ and $y \in \widetilde{V}_{IJ}$;
- *transverse* if $s_I|_{V_I} + \nu_I\colon V_I \to E_I$ is transverse to 0 for each $I \in \mathcal{I}_{\mathcal{K}}$;
- *precompact* if there is a precompact open subset $\mathcal{C} \sqsubset \mathcal{V}$ which itself is a (cobordism) reduction, such that

$$\pi_{\mathcal{K}}\Big(\bigcup_{I \in \mathcal{I}_{\mathcal{K}}} (s_I|_{V_I} + \nu_I)^{-1}(0)\Big) \subset \pi_{\mathcal{K}}(\mathcal{C}). \tag{3.2.3}$$

Remark 3.2.5 Although $\pi_{\mathcal{K}}\colon \mathrm{Obj}_{\mathcal{B}_{\mathcal{K}}} \to |\mathcal{K}|$ is not induced by a functor on $\mathcal{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$, we will work with $\pi_{\mathcal{K}}\colon \bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}} U_I \to |\mathcal{K}|$ as continuous map, in particular for the notion of precompactness. As in the case of trivial isotropy, we do not have a nicely controlled cover of sets $U_J \cap \pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(\mathcal{C}))$ for $\mathcal{C}\subset\bigsqcup U_I$. However, because $\mathcal{C}=\bigsqcup C_I\subset\mathcal{V}=\bigsqcup V_I\subset\bigsqcup U_I$ are lifts of reductions of $|\underline{\mathcal{K}}|$ as in Remark 3.2.2, the morphisms between V_J and $\mathcal{C}$ are better understood, yielding

$$V_J\cap\pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(\mathcal{C})) = V_J\cap\Bigl(\bigcup_{H\supset J}\rho_{JH}(C_H)\cup\bigcup_{H\subsetneq J}\rho_{HJ}^{-1}(C_H)\Bigr). \tag{3.2.4}$$

Indeed, by the reduction property, $\pi_{\mathcal{K}}(V_J)$ only intersects $\pi_{\mathcal{K}}(C_H)$ for $H\supset J$ or $H\subset J$. The morphisms between V_J and C_H are then given by ρ_{JH} and Γ_J in the first case, or ρ_{HJ} and Γ_H in the second, and the isotropy groups are absorbed by the equivariance $\Gamma_J\rho_{JH}(C_H)=\rho_{JH}(\Gamma_H C_H)$ and fact that $\Gamma_H C_H = C_H = \pi_H^{-1}(\underline{C}_H)$. As a result, we can write (3.2.3) in terms of the covering maps $(\rho_{IJ})_{I,J\in\mathcal{I}_{\mathcal{K}}}$, without explicit reference to the isotropy groups Γ_I, as

$$(s_J|_{V_J}+\nu_J)^{-1}(0)\subset\bigcup_{H\supset J}\rho_{JH}(C_H)\cup\bigcup_{H\subsetneq J}\rho_{HJ}^{-1}(C_H)\quad\forall J\in\mathcal{I}_{\mathcal{K}}. \tag{3.2.5}$$ ◇

Definition 3.2.6 Given a (cobordism) perturbation ν, the *perturbed zero set* $|\boldsymbol{Z}^\nu|$ is defined to be the realization of the full subcategory $\boldsymbol{Z}^\nu$ of $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ with object space

$$\bigl(\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}+\nu\bigr)^{-1}(0) := \bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}}(s_I|_{V_I}+\nu_I)^{-1}(0)\subset\mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}}$$

given by the local zero sets $Z_I := (s_I|_{V_I}+\nu_I)^{-1}(0)$. That is, we equip

$$|\boldsymbol{Z}^\nu| := \bigl|(\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}+\nu)^{-1}(0)\bigr| = \Bigl(\bigsqcup_{I\in\mathcal{I}_{\mathcal{K}}}Z_I\Bigr)/\!\sim$$

with the quotient topology generated by the morphisms of $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$. Moreover, we denote by $\iota^\nu\colon \boldsymbol{Z}^\nu\to\boldsymbol{B}_{\mathcal{K}}$ the functor induced by the inclusion $(\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}+\nu)^{-1}(0)\to\mathrm{Obj}_{\boldsymbol{B}_{\mathcal{K}}}$ and corresponding inclusion of morphism spaces (to a generally not full subcategory), with resulting continuous map

$$|\iota^\nu|\colon |\boldsymbol{Z}^\nu|\to|\mathcal{K}|. \tag{3.2.6}$$

Remark 3.2.7 If $\nu\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}\to\boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ is a cobordism perturbation of a tame Kuranishi cobordism $\mathcal{K}$, then each *restriction* $\nu|_{\partial^\alpha\mathcal{V}} := (\nu_I^\alpha)_{I\in\mathcal{I}_{\mathcal{K}^\alpha}}$ for $\alpha=0,1$ forms a perturbation of the Kuranishi atlas $\partial^\alpha\mathcal{K}$ with respect to the boundary restriction $\partial^\alpha\mathcal{V}$ of the reduction.

If in addition ν is admissible/transverse/precompact, then so are the restrictions $\nu|_{\partial^\alpha\mathcal{V}}$. Moreover, in the case of transversality each perturbed section $s_I|_{V_I}+\nu_I\colon V_I \to E_I$ for $I \in \mathcal{I}_{\partial^0\mathcal{K}} \cup \mathcal{I}_{\partial^1\mathcal{K}} \subset \mathcal{I}_{\mathcal{K}}$ is transverse to 0 as a map on a domain with boundary; ie the kernel of its differential is transverse to the boundary

$$\partial V_I = \bigsqcup_{\alpha=0,1} \iota_I^\alpha(\{\alpha\}\times\partial^\alpha V_I). \qquad \diamond$$

Theorem 3.2.8 *Let $(\mathcal{K},\sigma)$ be an oriented tame Kuranishi atlas/cobordism of dimension d and let ν be an admissible, transverse, precompact (cobordism) perturbation of $\mathcal{K}$ with respect to nested (cobordism) reductions $\mathcal{C} \sqsubset \mathcal{V} \sqsubset \mathrm{Obj}_{\mathcal{B}_\mathcal{K}}$. Then $\boldsymbol{Z}^\nu$ can be completed to a compact, d–dimensional wnb (cobordism) groupoid $\widehat{\boldsymbol{Z}}^\nu$, in the sense of Definition A.4, with the same realization $|\widehat{\boldsymbol{Z}}^\nu| = |\boldsymbol{Z}^\nu|$. In addition,*

$$\Lambda^\nu(p) := |\Gamma_I|^{-1}\#\{z \in Z_I \mid \pi_H(|z|) = p\}, \quad \text{for } p \in |Z_I|_{\mathcal{H}},$$

defines a weighting function $\Lambda^\nu\colon |\boldsymbol{Z}^\nu|_{\mathcal{H}} \to \mathbb{Q}^+$ on the Hausdorff quotient of the perturbed zero set $|\boldsymbol{Z}^\nu|_{\mathcal{H}}$. Together, these give $(|\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}}, \Lambda^\nu)$ the structure of a compact, d–dimensional weighted branched manifold/cobordism, in the sense of Definition A.5. It defines a cycle in $|\mathcal{C}|$ in the sense that the map $|\iota^\nu|_{\mathcal{H}}\colon |\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}} \to |\mathcal{K}|$ induced by (3.2.6) has image in $|\mathcal{C}|$.

Moreover, if $\mathcal{K}$ is a Kuranishi cobordism and the boundary restrictions of ν are denoted $\nu^\alpha := \nu|_{\partial^\alpha\mathcal{V}}$, then $(\widehat{\boldsymbol{Z}}^\nu, \Lambda^\nu)$ has oriented boundaries $(\widehat{\boldsymbol{Z}}^{\nu^0}, \Lambda^{\nu^0})$ and $(\widehat{\boldsymbol{Z}}^{\nu^1}, \Lambda^{\nu^1})$, and the cycle $|\iota^\nu|_{\mathcal{H}}\colon |\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}} \to |\mathcal{C}|$ restricts on the boundaries to $|\iota^{\nu^\alpha}|_{\mathcal{H}}\colon |\widehat{\boldsymbol{Z}}^{\nu^\alpha}|_{\mathcal{H}} \to |\partial^\alpha\mathcal{C}|$.

We begin the proof of Theorem 3.2.8 by explaining the structure of the groupoid completion $\widehat{\boldsymbol{Z}}^\nu$. Note that the compatibility condition (3.2.2) implies partial equivariance of the perturbation: $\nu_J(\alpha y) = \nu_J(y)$ for $y \in \widetilde{V}_{IJ}, \alpha \in \Gamma_{J\setminus I}$. This fact is reflected in the structure of the morphisms in the groupoid $\widehat{\boldsymbol{Z}}^\nu$, which contain this action of $\Gamma_{J\setminus I}$ on $\widetilde{V}_{IJ} \cap Z_J$ as part of the morphism space $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_J, Z_J)$.

Lemma 3.2.9 *Let ν be any (cobordism) perturbation of a tame d–dimensional Kuranishi atlas/cobordism $\mathcal{K}$.*

(i) *There is a unique nonsingular groupoid $\widehat{\boldsymbol{Z}}^\nu$ with the same objects and realization as $\boldsymbol{Z}^\nu$. Its morphism space for $I \subset J$ is given by*

$$\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I, Z_J) := \bigcup_{\varnothing\neq F\subset I} (Z_J\cap\widetilde{V}_{IJ}\cap\widetilde{V}_{FJ})\times\Gamma_{I\setminus F} \subset \widetilde{U}_{IJ}\times\Gamma_I = \mathrm{Mor}_{\boldsymbol{B}_\mathcal{K}}(U_I, U_J).$$

(ii) *If ν is admissible and transverse, then the subsets $Z_J \cap \widetilde{V}_{IJ} \subset Z_J$ are open for all $I \subset J$ and the groupoid $\widehat{\boldsymbol{Z}}^\nu$ is étale and has dimension d. Further, $\widehat{\boldsymbol{Z}}^\nu$ is oriented if in addition $\mathcal{K}$ is oriented.*

(iii) *If $\mathcal{K}$ is an oriented tame Kuranishi cobordism and ν is admissible and transverse, then $\widehat{\boldsymbol{Z}}^\nu$ satisfies all conditions given in the appendix for being an étale, oriented, cobordism groupoid, except possibly that of compactness.*

Proof First note that there is at most one nonsingular groupoid with the same objects and realization as $\boldsymbol{Z}^\nu$, since any such groupoid has a unique morphism $(I,x)\mapsto(J,y)$ whenever $(I,x)\sim(J,y)$, where $\sim$ is the equivalence relation on $\mathrm{Obj}_{\boldsymbol{Z}^\nu}$ generated by $\mathrm{Mor}_{\boldsymbol{Z}^\nu}$. To prove existence of such a groupoid, we show below that when $I\subset J$,

(a) each element in $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I,Z_J)$ is uniquely determined by its source and target;

(b) if there is a morphism $(I,J,y,\alpha)\in\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I,Z_J)$ with source (I,x) and target (J,y), then $(I,x)\sim(J,y)$;

(c) the set of morphisms $\bigcup_{I\subset J}\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I,Z_J)$ together with their inverses (which are uniquely defined by (a)) is closed under composition.

Parts (a) and (c) show that there is a nonsingular groupoid $\widehat{\boldsymbol{Z}}^\nu$ with the given morphisms. Moreover, since the equivalence relation $\sim$ is generated by the morphisms $(I,J,y,\mathrm{id})\in\mathrm{Mor}_{\boldsymbol{Z}^\nu}(Z_I,Z_J)\subset\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I,Z_J)$, (c) shows that if $(I,x)\sim(J,y)$, where $I\subset J$, then $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}((I,x),(J,y))\neq\varnothing$. Together with (b) this implies that $\widehat{\boldsymbol{Z}}^\nu$ has realization $|\boldsymbol{Z}^\nu|$.

To prove (a) we must check that given $x\in U_I$, $y\in Z_J\cap\widetilde{V}_{IJ}$, where $I\subset J$, there is at most one element $\alpha\in\Gamma_I$ such that

- $x=\alpha^{-1}\rho_{IJ}(y)$;
- there is $F\subset I$ such that $\alpha\in\Gamma_{I\setminus F}$ and $y\in\widetilde{V}_{FJ}\cap\widetilde{V}_{IJ}$.

But if α_1 and α_2 are two such elements, corresponding to F_1 and F_2, then $\alpha_1^{-1}\alpha_2$ fixes the point $\rho_{IJ}(y)$. On the other hand, because the set of F such that $y\in\widetilde{V}_{FJ}$ is nested, we can suppose that $F_1\subset F_2$. Then $\rho_{IJ}(y)\in\widetilde{V}_{F_1I}$ and $\alpha_1^{-1}\alpha_2\in\Gamma_{I\setminus F_1}$. Since $\Gamma_{I\setminus F_1}$ acts freely on $\widetilde{V}_{F_1I}$, this implies that $\alpha_1=\alpha_2$ as required.

To prove (b), observe that if $I\subset J$ and $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}((I,x),(J,y))\neq\varnothing$ then there is $F\subset I$ and $\alpha\in\Gamma_{I\setminus F}$ such that $x=\alpha^{-1}\rho_{IJ}(y)$, which implies that

$$\rho_{FI}(x)=\rho_{FI}(\alpha^{-1}\rho_{IJ}(y))=\rho_{FI}(\rho_{IJ}(y))=\rho_{FJ}(y).$$

Hence, the composite $(F,I,x,\mathrm{id})\circ(I,J,y,\alpha)$ is well defined and equal to (F,J,y,id). Therefore $(F,\rho_{FI}(x))\sim(I,x)$ and $(F,\rho_{FI}(x))=(F,\rho_{FJ}(y))\sim(J,y)$, which gives $(I,x)\sim(J,y)$ since $\sim$ is an equivalence relation.

Finally, to prove (c), it is convenient to consider two special kinds of morphisms: morphisms denoted μ^A with $I=J$, and morphisms denoted μ^B with $I\subsetneq J$ and $\alpha=\mathrm{id}$

that therefore belong to $\mathrm{Mor}_{\boldsymbol{Z}^\nu}$. We first observe that every morphism (I, J, y, α) in $\mathrm{Mor}_{\hat{\boldsymbol{Z}}^\nu}(Z_I, Z_J)$ can be written in two ways as a composite of morphisms of types (A) and (B). More precisely, the identity $\mu_1^A \circ \mu_1^B = \mu_2^B \circ \mu_2^A$ holds, where μ_i^A, μ_j^B are the morphisms in the following diagram:

$$\begin{tikzcd}
(I, \alpha^{-1}\rho_{IJ}(y)) \arrow[r, "\mu_2^B"] \arrow[d, "\mu_1^A"] & (J, \alpha^{-1} y) \arrow[d, "\mu_2^A"] \\
(I, \rho_{IJ}(y)) \arrow[r, "\mu_1^B"] & (J, y)
\end{tikzcd} \tag{3.2.7}$$

Therefore these morphisms μ^A, μ^B and their inverses generate $\mathrm{Mor}_{\hat{\boldsymbol{Z}}^\nu}$. The commutativity of the above diagram also shows that we can interchange their order: ie every morphism of the form $\mu_1^A \circ \mu_1^B$ can also be written as $\mu_2^B \circ \mu_2^A$, which we abbreviate below as the identity $\mu_1^A \circ \mu_1^B = \mu_2^B \circ \mu_2^A$.

Next let us consider the other composites. Morphisms of type (A) with fixed $F \subset I$ are closed under composition since they are given by the action of $\Gamma_{I \setminus F}$. Moreover, two morphisms of this type corresponding to different subgroups F_1, F_2 can be composed only if the sets $\pi_{\mathcal{K}}(V_I), \pi_{\mathcal{K}}(V_{F_1}), \pi_{\mathcal{K}}(V_{F_2})$ intersect. Hence the sets F_1, F_2 are nested, either $F_1 \subset F_2$ or $F_2 \subset F_1$, and in either case the composite is another morphism of this type. The situation for morphisms of type (B) is more complicated (which is precisely why we needed to add the morphisms of type (A) to obtain a groupoid). We have:

- $\mu_1^B \circ \mu_2^B = \mu_3^B$: ie if $I \subset J \subset K$ and $y = \rho_{JK}(z)$ then the following identity holds (this statement includes the claim that the left-hand composite is well defined):

$$(I, J, y, \mathrm{id}) \circ (J, K, z, \mathrm{id}) = (I, K, z, \mathrm{id}) \in \mathrm{Mor}_{\hat{\boldsymbol{Z}}^\nu}((I, \rho_{IK}(z)),\ (K, z)).$$

- $(\mu_1^B)^{-1} \circ \mu_2^B = \mu^A \circ \mu_3^B$ or $= \mu^A \circ (\mu_3^B)^{-1}$:
  - If $I \subset J \subset K$ and $\rho_{IJ}(y') = \rho_{IK}(y) = \rho_{IJ} \circ \rho_{JK}(y)$, then $\rho_{JK}(y)$ and y' lie in the same $\Gamma_{J \setminus I}$–orbit so that $y' = \alpha^{-1}\rho_{JK}(y)$ for some $\alpha \in \Gamma_{J \setminus I}$, and

$$\begin{aligned}(I, J, y', \mathrm{id})^{-1} \circ (I, K, y, \mathrm{id}) &= (J, J, \rho_{JK}(y), \alpha) \circ (J, K, y, \mathrm{id}) \\ &\in \mathrm{Mor}_{\hat{\boldsymbol{Z}}^\nu}((J, \alpha^{-1}\rho_{JK}(y)), (K, y)).\end{aligned}$$

  - If $I \subset K \subset J$ and there are $y' \in \widetilde{V}_{IJ} \cap Z_J,\ y \in \widetilde{V}_{IK} \cap Z_K$ with

$$\rho_{IJ}(y') = \rho_{IK}(\rho_{KJ}(y')) = \rho_{IK}(y) \in Z_I,$$

  then there is $\beta \in \Gamma_{K \setminus I}$ such that $y = \beta\, \rho_{KJ}(y') = \rho_{KJ}(\beta y') \in Z_K$, and

$$\begin{aligned}(I, J, y', \mathrm{id})^{-1} \circ (I, K, y, \mathrm{id}) &= (J, J, \beta\, y', \beta) \circ (K, J, \beta\, y', \mathrm{id})^{-1} \\ &\in \mathrm{Mor}_{\hat{\boldsymbol{Z}}^\nu}((J, y'), (K, y)).\end{aligned}$$

- One can check similarly that if $\mu_1^B = (I, J, y, \mathrm{id})$ and $\mu_2^B = (K, J, y, \mathrm{id})$ then

$$\mu_1^B \circ (\mu_2^B)^{-1} = \begin{cases} \mu^A \circ \mu_3^B & \text{if } I \subset K, \\ \mu^A \circ (\mu_3^B)^{-1} & \text{if } K \subset I. \end{cases}$$

Combining these identities with $\mu_1^A \circ \mu_1^B = \mu_2^B \circ \mu_2^A$ and its inverse, we see that if $I \subset J$, every composite morphism $Z_I \to Z_J \to Z_K$ can be written in the form $\mu^B \circ \mu^A$ if $I \subset K$, and in the form $\mu_1^A \circ (\mu_1^B)^{-1} = (\mu_2^B)^{-1} \circ \mu_2^A$ if $K \subset I$. This completes the proof of (c) and hence of part (i) of the lemma.

The claims in (ii) are proved by applying Lemma 3.1.14 with $f\colon W \to E_I$ given by $s_I + \nu_I\colon V_I \to E_I$. Since $s_I + \nu_I \pitchfork 0$, Lemma 3.1.14(i) shows that Z_I is a manifold, while the admissibility of ν implies that the hypothesis of Lemma 3.1.14(iii) holds on $\widetilde{V}_{HI}$ so that the subset $Z_I \cap \widetilde{V}_{HI}$ of Z_I is open and ρ_{HI} induces a local diffeomorphism from $Z_I \cap \widetilde{V}_{HI}$ to $Z_H \cap \rho_{HI}(\widetilde{V}_{HI})$. Further, by the compatibility condition (3.2.2) we can identify with the zero set of $\rho_{HI} * (s_I + \nu_I) = s_H + \rho_{HI} * (s_I)$. Since the maps ρ_{IJ} together with their inverses generate the structure maps in $\widehat{\boldsymbol{Z}}^\nu$, this shows that this groupoid is étale. Moreover, if $\mathcal{K}$ is oriented, then Lemma 3.1.14(ii)–(iii) also implies that the structure maps in $\widehat{\boldsymbol{Z}}^\nu$ are orientation-preserving.

Finally, (iii) holds by Lemma 3.1.14(iv). □

In order to show that $\widehat{\boldsymbol{Z}}^\nu$ represents a weighted branched manifold, we must understand its maximal Hausdorff quotient $|\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}}$ as defined in Lemma A.2. The morphisms in a nonsingular groupoid $\boldsymbol{G}$ correspond bijectively to the equivalence relation $\sim_{\boldsymbol{G}}$ on $\mathrm{Obj}_{\boldsymbol{G}}$ where $x \sim_{\boldsymbol{G}} y$ if and only if $\mathrm{Mor}_{\boldsymbol{G}}(x, y) \neq \varnothing$. A necessary condition for the quotient $|\boldsymbol{G}| := \mathrm{Obj}_{\boldsymbol{G}} / {\sim_{\boldsymbol{G}}}$ to be Hausdorff is that this equivalence relation be given by a closed subset of $\mathrm{Obj}_{\boldsymbol{G}} \times \mathrm{Obj}_{\boldsymbol{G}}$; in other words, we need the map $s \times t\colon \mathrm{Mor}_{\boldsymbol{G}} \to \mathrm{Obj}_{\boldsymbol{G}} \times \mathrm{Obj}_{\boldsymbol{G}}$ that takes a morphism to its source and target to have closed image. The following lemma shows that in the special case of the groupoid $\widehat{\boldsymbol{Z}}^\nu$ this necessary condition is also sufficient.

Lemma 3.2.10 *Let ν be an admissible, transverse, (cobordism) perturbation of a tame Kuranishi atlas/cobordism $\mathcal{K}$. Then:*

(i) *Let $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$ be the groupoid obtained from $\widehat{\boldsymbol{Z}}^\nu$ by closing the relation $\sim$ on $\mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu}$. Then we have that $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$ is nonsingular and $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is Hausdorff. Further, we can identify $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ with the maximal Hausdorff quotient $|\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}}$ in such a way that the canonical quotient map $|\widehat{\boldsymbol{Z}}^\nu| \to |\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}} = |\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is induced by the functor $\iota_H\colon \widehat{\boldsymbol{Z}}^\nu \to \widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$.*

(ii) *For each* $I \in \mathcal{I}_{\mathcal{K}}$, *the projection*

$$\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}\colon \mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}} \to |\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$$

takes Z_I *onto a subset of* $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ *that is open with respect to the quotient topology. This topology on* $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ *is metrizable.*

(iii) *If* $x \in Z_I$ *and* $p = \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x) \in |\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$, *then* $\{x' \in Z_I \mid \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(I,x')=p\}$ *is the* $(\Gamma_{I\setminus F_x})$*–orbit of* x, *so*

$$\#\{x \in Z_I \mid \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(I,x)=p\} = |\Gamma_{I\setminus F_x}|,$$

where $F_x = \min\{F : Z_I \cap \mathrm{cl}(\widetilde{V}_{FI}) \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(p) \neq \varnothing\} = \min\{F : p \in \overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)}\}$.

Proof We use the notation in Lemma 3.2.9. The components of $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I, Z_J)$ consisting of morphisms of type (B) are taken by $s \times t\colon \mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I, Z_J) \to Z_I \times Z_J \subset \mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu} \times \mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu}$ to the set of pairs

$$\bigl\{(\rho_{IJ}(y), y) \mid y \in Z_J \cap \widetilde{V}_{IJ} \cap \pi_{\mathcal{K}}^{-1}(V_I)\bigr\} \subset Z_I \times Z_J,$$

where we simplify notation by writing y instead of (J, y), and similarly for the source. If $(\rho_{IJ}(y_n), y_n) \to (x_\infty, y_\infty) \in Z_I \times Z_J$ is a convergent sequence of such points with limit $(x_\infty, y_\infty) \in Z_I \times Z_J$, then $y_\infty \in Z_J \cap \widetilde{U}_{IJ}$ since $y_n \in Z_J \cap \widetilde{V}_{IJ} \subset \widetilde{U}_{IJ}$ and $\widetilde{U}_{IJ}$ is closed in U_J, which implies that $\rho_{IJ}(y_\infty)$ is defined. We then must have $x_\infty = \rho_{IJ}(y_\infty)$ by the continuity of ρ_{IJ}. Thus

$$y_\infty \in \rho_{IJ}^{-1}(Z_I) \cap Z_J \subset \rho_{IJ}^{-1}(V_I) \cap V_J = \widetilde{V}_{IJ}.$$

Hence $y_\infty \in Z_J \cap \widetilde{V}_{IJ}$, so that $(I, J, y_\infty, \mathrm{id}) \in \mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I, Z_J)$. Therefore the graph of this set of morphisms is closed in $Z_I \times Z_J$.

However the set of morphisms of type (A) from $Z_I \to Z_I$ is not closed in general; instead it has closure[8]

$$\overline{\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I, Z_I)} := \bigcup_{F \subsetneq I} \bigl\{(I, I, y, \alpha) \mid y \in \mathrm{cl}(\widetilde{V}_{FI}) \cap Z_I,\ \alpha \in \Gamma_{I\setminus F}\bigr\}.$$

Notice that, as in the proof of Lemma 3.2.9(i), this set $\overline{\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I, Z_I)}$ is invariant under compositions (and inverses) because the intersection properties of the sets in a reduction apply to their closures: $\pi_{\mathcal{K}}(\overline{V_{F_1}}) \cap \pi_{\mathcal{K}}(\overline{V_{F_2}}) \neq \varnothing \implies F_1 \subset F_2$ or $F_2 \subset F_1$. Next, observe that because $\rho_{IJ}\colon \widetilde{U}_{IJ} \to U_{IJ}$ is a local diffeomorphism, the map ρ_{IJ} induces a local diffeomorphism from $\widetilde{V}_{IJ} \cap \mathrm{cl}(\widetilde{V}_{FJ}) \cap Z_J$ into $U_{IJ} \cap \mathrm{cl}(\widetilde{V}_{FI}) \cap Z_I$.

[8] While we usually denote the closure of a set A by $\bar{A}$, for sets such as $\widetilde{V}_{IJ}$ that involve a tilde we will write $\mathrm{cl}(\widetilde{V}_{IJ})$.

Similarly, because $\mathrm{cl}(\widetilde{V}_{FJ}) \subset \widetilde{U}_{FJ}$ whenever $F \subset J$, the group $\Gamma_{I \setminus F}$ acts freely on $\mathrm{cl}(\widetilde{V}_{FJ})$, and, if $F \subset I$, commutes with the action of ρ_{IJ} as in diagram (3.2.7). Therefore the closure of $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^{\nu}}(Z_I, Z_J)$ when $I \subset J$ is given as follows:

$$\begin{aligned}(3.2.8)\quad \mathrm{Mor}_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I, Z_J) &= \bigcup_{F \subset I} \big(Z_J \cap \widetilde{V}_{IJ} \cap \mathrm{cl}(\widetilde{V}_{FJ})\big) \times \Gamma_{I \setminus F} \\ &= \big\{(I, J, y, \alpha) \,\big|\, \exists F \subset I,\ \alpha \in \Gamma_{I \setminus F} \\ &\qquad\qquad \text{such that } y \in \mathrm{cl}(\widetilde{V}_{FJ}) \cap \widetilde{V}_{IJ} \cap Z_J\big\}.\end{aligned}$$

The arguments in Lemma 3.2.10 apply to show that this set of morphisms, together with inverses, are closed under composition and are uniquely determined by their source and target. Thus $\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}$ is a nonsingular groupoid. Its realization $|\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ is Hausdorff as it is the quotient of the separable, locally compact metric space $\mathrm{Obj}_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}$ by a relation with closed graph; see [1, Chapter I, Section 10, Example 19] or [13, Lemma 3.2.4]. Moreover, the space $|\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ can be identified with the maximal Hausdorff quotient of $|\widehat{\boldsymbol{Z}}^{\nu}|$ because any continuous map from $\mathrm{Obj}_{\widehat{\boldsymbol{Z}}^{\nu}}/\!\sim$ to a Hausdorff space Y must factor through the closure of the relation $\sim$ induced by the morphisms in $\widehat{\boldsymbol{Z}}^{\nu}$, and hence descends to $|\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$. This proves (i).

To see that $\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I)$ is open in $|\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ we must show that each intersection

$$Z_J \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I))$$

is open. Since $\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I) \cap \pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_J) \neq \varnothing$ only if $I \subset J$ or $J \subset I$, it suffices to consider these two cases. Now

$$Z_J \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I))$$

consists of all elements in Z_J that are targets of morphisms with source in Z_I. Therefore if $I \subsetneq J$, then

$$Z_J \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I)) = Z_J \cap \widetilde{V}_{IJ},$$

which is open by Lemma 3.2.9(i). On the other hand, if $J \subset I$ then because the set $\rho_{JI}(\widetilde{V}_{JI})$ is Γ_J–invariant, we have

$$Z_J \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I)) = Z_J \cap \rho_{JI}(\widetilde{V}_{JI}),$$

which is open by Lemma 3.2.9(ii). Thus $\pi_{\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}}(Z_I)$ is open. It follows that the quotient topology on $|\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ has a countable basis because each Z_I does. We also have that $|\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ is regular. Indeed, by [15, Lemma 31.1], we only need to check that each point $p \in |\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ with neighborhood $W \subset |\widehat{\boldsymbol{Z}}^{\nu}_{\mathcal{H}}|$ has a smaller neighborhood $W_1 \subset W$ such that $\overline{W_1} \subset W$, and this is an immediate consequence of the regularity and local compactness of the sets Z_I and the openness of the sets $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I)$. Therefore $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is metrizable by the Urysohn metrization theorem. This proves (ii).

To prove (iii), note first that for each $x \in Z_I$ the subsets $F \in \mathcal{I}_{\mathcal{K}}$ such that $x \in \mathrm{cl}(\widetilde{V}_{FI})$ are nested and hence have a minimal element F_x. The precompactness of V_I in U_I implies that $x \in \mathrm{cl}(\widetilde{V}_{FI}) \subset \widetilde{U}_{F_x I}$ so that its orbit under $\Gamma_{I \setminus F_x}$ is free. Moreover, because $F_x \subset F$ for every F for which $x \in \mathrm{cl}(\widetilde{V}_{FI})$, this orbit $\Gamma_{I \setminus F_x}(x)$ contains the targets of all the morphisms in $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}$ with source (I,x). This proves the formula $|\Gamma_{I \setminus F_x}| = \#\{x \in Z_I \mid \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(I,x) = p\}$.

It remains to check that F_x, which we defined as

$$\min\{F : Z_I \cap \mathrm{cl}(\widetilde{V}_{FI}) \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(p) \neq \varnothing\},$$

also equals $F'_x := \min\{F : p \in \overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)}\}$. But if

$$Z_I \cap \mathrm{cl}(\widetilde{V}_{FI}) \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(p) \neq \varnothing$$

there is a sequence of elements $x_k \in Z_I \cap \widetilde{V}_{FI}$ with limit $x_\infty \in \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(p)$, implying by the continuity of $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}$ that, with $x'_k := \rho_{FI}(x_k)$, the sequence

$$\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x'_k) = \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x_k)$$

converges to p. Hence $p \in \overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)}$, which implies $F'_x \subset F_x$. Conversely, if

$$p \in \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I) \cap \overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)},$$

then since $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I)$ is open in $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ there is a sequence p_k of elements in

$$\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I) \cap \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)$$

that converges to $p \in \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I)$. By (3.2.1), this lifts to a sequence $x_k \in \widetilde{V}_{FI} \subset Z_I$, and the sequence of images $\pi_{\mathcal{K}}(\iota_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x_k))$ in $|\mathcal{V}| \subset |\mathcal{K}|$ converges to $|\iota_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}|(p)$, where $\iota_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}$ is as in (ii). But the composite

$$\pi_{\mathcal{K}} \circ \iota_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}} \colon V_I \to \pi_{\mathcal{K}}(V_I) \cong V_I / \Gamma_I$$

simply quotients out by the action of Γ_I on V_I. Since the projection $V \to V_I/\Gamma_I$ is proper by Lemma 2.1.5(i), the sequence (x_k) must have a convergent subsequence with limit $x_\infty \in V_I$. But then by uniqueness of limits in the Hausdorff space $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$, $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x_\infty) = \lim_{k\to\infty} \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x_k) = p$. Therefore

$$x_\infty \in \mathrm{cl}(\widetilde{V}_{FI}) \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(p).$$

Hence by the minimality of F_x we must have $F_x \subset F'_x$. This completes the proof. □

Proof of Theorem 3.2.8 Let us first consider the case when ν is an admissible, transverse, precompact perturbation of an oriented tame Kuranishi atlas $\mathcal{K}$ with respect to nested reductions $\mathcal{C} \sqsubset \mathcal{V} \sqsubset \mathrm{Obj}_{\mathcal{B}_{\mathcal{K}}}$. Then Lemma 3.2.9 shows that $\boldsymbol{Z}^\nu$ can be completed to an oriented, nonsingular étale groupoid $\widehat{\boldsymbol{Z}}^\nu$. Moreover, by Lemma 3.2.10 the maximal Hausdorff quotient $|\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}}$ can be identified with the realization $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ of the groupoid $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$. To complete the proof of the first part of the theorem it remains to show that $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is compact, and that $(\widehat{\boldsymbol{Z}}^\nu, \Lambda^\nu)$ has the structure of a wnb groupoid as in Definition A.4.

Because $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is metrizable by Lemma 3.2.9(ii), it suffices to prove that $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is sequentially compact. Further, we saw in (3.2.5) that the precompactness condition for ν can be written without explicit mention of the isotropy groups Γ_I. Hence the proof of the sequential compactness of the zero set given in [13, Theorem 5.2.2] carries through, without change, to the current situation.

We next check that the weighting function Λ^ν is well defined, and compatible with a local branching structure as required by Definition A.4. To see that it is well defined, suppose that $p \in \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I) \cap \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_J)$. As usual we may suppose that $I \subset J$, so that $p = \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(y)$ for some $y \in \widetilde{V}_{IJ} \subset Z_J$. Let the minimal set F such that $p \in \overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)}$ be denoted by F_p. Then there are $|\Gamma_{J\setminus F_p}|$ distinct elements in Z_J that map to p. Hence $\Lambda(p) = |\Gamma_{J\setminus F_p}|/|\Gamma_J|$, and we must check that this agrees with the calculation provided by replacing J by I. But if $x = \rho_{IJ}(y)$, then because $F_p \subset I$ does not depend on I, J we have $\Gamma_{J\setminus F_p} = \Gamma_{I\setminus F_p} \times \Gamma_{J\setminus I}$. Hence $|\Gamma_{J\setminus F_p}|/|\Gamma_J| = |\Gamma_{I\setminus F_p}|/|\Gamma_I|$. Thus Λ^ν is well defined.

Finally we describe the local branches at $p \in |\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$. Given $p \in |\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$, choose a minimal I such that $p \in \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I)$, and a minimal $F_p \subset I$ such that $p \in \overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_{F_p})}$. Then $F_p \subset I$, and there is $x \in Z_I \cap \mathrm{cl}(\widetilde{V}_{F_p I})$ such that $p = \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(x)$. As $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I)$ is open in $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$, we may choose an open neighborhood $N \subset \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_I)$ of p whose closure $\overline{N}$ is disjoint from all sets $\overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)}$ with $F \subsetneq F_p$. We saw in Lemma 3.2.10(iii) that $Z_I \cap (\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}})^{-1}(p) = \Gamma_{I\setminus F_p}(x)$. Hence, by shrinking N further if necessary, we may suppose that there is an precompact open neighborhood B_x of x in Z_I such that

- $\bigcup_{\gamma \in \Gamma_{I\setminus F_p}} \pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(\gamma B_x) = N$;
- the closure $\overline{B_x}$ in Z_J is disjoint from its images under the action of $\Gamma_{I\setminus F_p}$.

Then choose the local branches to be the disjoint subsets $(\gamma B_x)_{\gamma \in \Gamma_{I\setminus F_p}}$ of Z_I, each with weight $1/|\Gamma_I|$. Notice that

$$(3.2.9) \qquad \bigcup_{\gamma \in \Gamma_{I\setminus F_p}} \gamma \overline{B_x} = Z_I \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(\overline{N}) \quad \text{and} \quad \bigcup_{\gamma \in \Gamma_I \setminus F_p} \gamma B_x = Z_I \cap \pi^{-1}_{\widehat{\boldsymbol{Z}}^\nu}(N).$$

Here, the first claim holds because, by the minimality of F_p and the choice of B_x,

$$\overline{B_x} \cap \mathrm{cl}(\widetilde{V}_{FI}) \neq \varnothing \implies F_p \subset F,$$

so that the only morphisms in $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$ with source in $\bigcup_{\gamma\in\Gamma_{I\setminus F_p}} \gamma\overline{B_x}$ and target in Z_I are given by the action of an element in $\Gamma_{I\setminus F_p}$ and hence also have target in this set. The second claim holds similarly.

We must check that the three conditions in Definition A.4 hold.

- The covering property states that

$$\big(\pi^{\mathcal{H}}_{|\widehat{\boldsymbol{Z}}^\nu|}\big)^{-1}(N) \;=\; \bigcup_{\gamma\in\Gamma_{I\setminus F_p}} |\gamma B_x| \;\subset\; |\boldsymbol{Z}|.$$

If this were false there would be a point $y\in Z_J$ for some J such that there is a morphism in $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$ from (J,y) to a point $(I,x')\in\gamma B_x$ for some $\gamma\in\Gamma_{I\setminus F_p}$, but there is no such morphism in $\widehat{\boldsymbol{Z}}^\nu$. By construction, the morphisms in $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$ from Z_J to Z_I are composites of morphisms of type (B) from Z_J to Z_I (which lie in $\widehat{\boldsymbol{Z}}^\nu$) with morphisms in the closure of $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I,Z_I)$. Therefore it suffices to consider the case $J=I$, and $y\notin\bigcup_{\gamma\in\Gamma_{I\setminus F_p}}\gamma B_x$. But (3.2.9) implies that the only elements of $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu}(Z_I,Z_I)$ with target in γB_x must have source in some set αB_x. Therefore such y does not exist.

- For local regularity, we must check that for each γ the projection $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}\colon \gamma B_x\to|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is a homeomorphism onto a relatively closed subset of N. But (3.2.9) implies that this map extends to an injective, continuous map f with compact domain $\gamma\overline{B_x}$. Hence f is a homeomorphism onto its image because compact subsets of the Hausdorff space $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ are closed. Further, $\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(\gamma B_x) = N\cap\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(\gamma\overline{B_x})$ is closed in N because it is the intersection of a compact set with N.

- Finally, note that Λ^ν equals the branching function specified in Definition A.4; indeed, the number of branches through $q\in N$ is just the number of preimages of $q\in N$ in $\bigcup_{\gamma\in\Gamma_{I\setminus F_p}}\gamma B_x$, and we saw in Lemma 3.2.10(iii) that this is $|\Gamma_{I\setminus F_q}|$, where $F_q\supset F_p$ is the minimal set F such that $q\in\overline{\pi_{\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}}(Z_F)}$.

This completes the proof that $(\widehat{\boldsymbol{Z}}^\nu,\Lambda^\nu)$ is a compact wnb groupoid. It has a fundamental class by Proposition A.7, and hence defines a cycle in $\mathcal{C}$ as claimed.

The same arguments apply when $\mathcal{K}$ is a Kuranishi cobordism. In particular, $|\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ is compact so that, by Lemma 3.2.9(iii), $(\widehat{\boldsymbol{Z}}^\nu,\Lambda^\nu)$ is a wnb cobordism groupoid, and the boundary restrictions have the required properties by Lemma 3.1.14(iv). □

We end this section by some elementary examples of this construction: the fundamental class and the Euler class of an orbifold represented by Kuranishi atlases.

Example 3.2.11 Consider the orbifold case, ie a Kuranishi atlas $\mathcal{K}$ on X with trivial obstruction spaces so that $\mathfrak{s}_{\mathcal{K}}$ and ν are identically zero and $\iota_{\mathcal{K}}\colon X \to |\mathcal{K}|$ is a homeomorphism. In this case the zero set $\boldsymbol{Z}$ should represent the fundamental class of the oriented orbifold. We suppose that $X = M/\Gamma$ is the quotient of a compact oriented smooth manifold M by the action of a finite group Γ, and that $\mathcal{K}$ is the atlas with a single chart with domain M and $E = \{0\}$. Then $\boldsymbol{Z} = |\boldsymbol{B}_{\mathcal{K}}|^{\setminus\Gamma}$ is the category with objects M and only identity morphisms, because there are no pairs $I, J \in \mathcal{I}_{\mathcal{K}}$ such that $\varnothing \neq I \subsetneq J$. Therefore $\boldsymbol{Z} = \boldsymbol{Z}_{\mathcal{H}}$ has realization $|\boldsymbol{Z}_{\mathcal{H}}| = M$ and the weighting function $\Lambda\colon M \to \mathbb{Q}$ is given by $\Lambda(x) = 1/|\Gamma|$. If the action of Γ is effective on every open subset of M, then the pushforward of Λ by $\iota_{\boldsymbol{Z}_{\mathcal{H}}}\colon M \to X$, which is defined by

$$(\iota_{\boldsymbol{Z}_{\mathcal{H}}})_*\Lambda(p) \;:=\; \sum_{x \in (\iota_{\boldsymbol{Z}_{\mathcal{H}}})^{-1}(p)} \Lambda(x),$$

takes the value 1 at every smooth point (ie point with trivial stabilizer) of the orbifold M/Γ. On the other hand, if Γ acts by the identity so that the action is totally noneffective, then $\iota_{\boldsymbol{Z}_{\mathcal{H}}}\colon M \xrightarrow{\cong} X$ is the identity map and the weighting function $X \to \mathbb{Q}^+$ takes the constant value $1/|\Gamma|$.

Note that if we construct a fundamental class on $|\boldsymbol{Z}|_{\mathcal{H}}$ by the method of Proposition A.7 then our choice of weights gives a class that is consistent with standard conventions. For example, in dimension $d = 0$ the branched manifold $Z = |\boldsymbol{Z}|_{\mathcal{H}}$ is a finite collection of points $\{p_1, \dots, p_k\}$, one for each equivalence class in $\mathrm{Obj}_{\boldsymbol{Z}}$, where the point p_i corresponding to an equivalence class with stabilizer Γ^i has weight $1/|\Gamma^i|$. If each point is positively oriented, then the "number of elements" in $|\boldsymbol{Z}|_{\mathcal{H}}$ is the sum $\sum_{i=1}^k 1/|\Gamma^i|$, which gives the Euler characteristic of the groupoid; cf [17]. Other more substantive examples such as that of the football of Example 2.3.11 are discussed in [11, Example 4.6].

Example 3.2.12 Examples of Kuranishi atlases with nontrivial obstruction spaces can be seen in the calculation of the Euler class of the tangent bundle of S^2 and of the football orbifold using Kuranishi atlases.

(i) To build a Kuranishi atlas that models $\mathrm{T}S^2$, cover S^2 by two discs D_1, D_2 whose intersection $D_1 \cap D_2 =: D_{12} =: A$ is an annulus, and for $i = 1, 2$ define $\boldsymbol{K}_i := (U_i := D_i,\ E_i := \mathbb{C},\ s_i := 0,\ \psi_i := \mathrm{id})$. For $i = 1, 2$ choose trivializations $\tau_i\colon D_i \times \mathbb{C} \to \mathrm{T}S^2|_{D_i}$, $(x, e) \mapsto \tau_{i,x}(e)$ and then define the transition chart

$$\boldsymbol{K}_{12} := (U_{12} \subset E_1 \times E_2 \times A,\ E_1 \times E_2,\ s_{12} = \mathrm{pr}_{E_1 \times E_2},\ \psi_{12} = \mathrm{pr}_A|_{0 \times 0 \times A}),$$

where

$$U_{12} := \{(e_1, e_2, x) \mid x \in A,\ \tau_{1,x}(e_1) + \tau_{2,x}(e_2) = 0\}.$$

The coordinate changes $\widehat{\Psi}_{i,12}$ are given by taking $U_{i,12} = A$ and $\psi_{i,12}(x) = (0,0,x)$. To justify this choice of Kuranishi atlas note that one can construct a commutative diagram

$$\begin{tikzcd} |\boldsymbol{E}_{\mathcal{K}}| \arrow[r] \arrow[d] & \mathrm{T}S^2 \arrow[d] \\ |\boldsymbol{B}_{\mathcal{K}}| \arrow[r] & S^2 \end{tikzcd}$$

that restricts over $U_{12} \times E_{12}$ to the map

$$((e_1, e_2, x), e_1', e_2') \mapsto \tau_{1,x}(e_1 + e_1') + \tau_{2,x}(e_2 + e_2') \in \mathrm{T}S^2|_A.$$

This construction is generalized to other (orbi)bundles in [10].

Next, in order to calculate the Euler class we identify A with $[0,1] \times S^1$ and consider the corresponding trivialization $\mathrm{T}S^2|_A = A \times \mathbb{R}_t \times \mathbb{R}_\theta$, where $t \in [0,1]$ and $\theta \in S^1$ are coordinates. Then for $i = 1, 2$ there is a section $\nu_i \colon U_i \to E_i$ with one transverse zero such that $\tau_{i,x}(\nu_i(x)) = (x, 1, 0) \in \mathrm{T}S^2|_A$ for $x \in A$. (Take suitably modified versions of the sections $\nu_1(z) = z$ and $\nu_2(z) = -z$, where $D_i \subset \mathbb{C}$.)

Choose a reduction of the footprint covering with $V_{12} = (\varepsilon, 1-\varepsilon) \times S^1$ for some $\varepsilon \in \left(0, \frac{1}{4}\right)$ and so that $\widetilde{V}_{1,12} = (0,0) \times \left(\varepsilon, \frac{1}{4}\right] \times S^1 \subset U_{12}$ and $\widetilde{V}_{2,12} = (0,0) \times \left(\frac{3}{4}, 1-\varepsilon\right) \times S^1$, and choose a cutoff function $\beta \colon [0,1] \to [0,1]$ that equals 1 in $\left[0, \frac{1}{4}\right]$ and 0 in $\left[\frac{3}{4}, 1\right]$. Then the map $\nu_{12} \colon V_{12} \to E_1 \times E_2$ given by

$$\nu_{12}(e_1, e_2, x) = \big(\beta(x)\nu_1(x), (1-\beta(x))\nu_2(x)\big) \in E_1 \times E_2$$

defines an admissible perturbation section that restricts to ν_i on $V_{i,12} \subset (0,0) \times A$ for $i = 1, 2$. Moreover $s_{12} + \nu_{12}$ does not vanish at any point $(e_1, e_2, x) \in V_{12}$ because the equation $\tau_{1,x}(e_1) + \tau_{2,x}(e_2) = 0$ together with

$$0 = \tau_{1,x}(e_1) + \beta(x)(1,0) = \tau_{2,x}(e_2) + (1-\beta(x))(1,0) \in x \times \mathbb{R}_t \times \mathbb{R}_\theta \in \mathrm{T}S^2|_A$$

imply that the vector $(1,0)$ is zero, a contradiction. Hence the perturbed zero set $\boldsymbol{Z}^\nu$ consists of two points, each with weight one.

(ii) It is easy to adjust this example to the tangent bundle of the "football" discussed in Example 2.3.11. In this case, the zero of the section $s_i + \nu_i$ would count with weight $1/|\Gamma_i|$ so that the Euler class is $\frac{1}{2} + \frac{1}{3}$. For further details of this and other related examples see [10, Section 5].

3.3 Construction of the virtual moduli cycle and fundamental class

The next step in the Kuranishi regularization Theorem A is to construct admissible, transverse, precompact perturbations ν that are unique up to interpolation by admissible, transverse, precompact cobordism perturbations. This — quite complicated — construction is developed in complete detail in [14] in such a way that it applies directly to our present setting, in which the Kuranishi atlas $\mathcal{K}$ has nontrivial isotropy groups, but the reduced and pruned category $\mathcal{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ is nonsingular, ie the remaining isotropy groups act freely. While we defer most of the proofs to [14], we will give full technical statements of the existence and uniqueness of perturbations, so that our constructions of VMC/VFC can be compared directly to other approaches, without reference to [14]. Based on this, Definition 3.3.4 and Theorem 3.3.5 then define the virtual moduli cycle (VMC) as a cobordism class of closed oriented weighted branched manifolds and construct the virtual fundamental class (VFC) as Čech homology class.

For the construction of (cobordism) perturbations we will consider a metric tame Kuranishi atlas (or cobordism) $(\mathcal{K}, d)$. That is, we fix the following data:

- $\mathcal{K}$ is a tame Kuranishi atlas on a compact metrizable space X in the sense of Definitions 2.3.1 and 2.5.1, or it is a tame Kuranishi cobordism on a compact collared cobordism Y in the sense of Definitions 2.4.2 and 2.5.1.
- d is an admissible metric on $|\mathcal{K}|$ in the sense of Definition 2.3.10.
- If $(\mathcal{K}, d)$ is a metric, tame Kuranishi cobordism on Y, then the boundary restrictions $(\mathcal{K}^\alpha, d^\alpha) := (\partial^\alpha \mathcal{K}, d|_{|\partial^\alpha \mathcal{K}|})$ are metric, tame Kuranishi atlases on $\partial^\alpha Y$ for $\alpha = 0, 1$.

For easy reference we list some consequences of this setting and notation conventions.

- The associated intermediate Kuranishi atlas $\underline{\mathcal{K}}$ is a tame topological Kuranishi atlas (resp. cobordism) by Lemma 2.3.4 (resp. Remark 2.4.3(ii)), which has the same realization $|\underline{\mathcal{K}}| = |\mathcal{K}|$, equipped with the quotient topology.
- d is a bounded metric on the set $|\underline{\mathcal{K}}|$ such that for each $I \in \mathcal{I}_{\mathcal{K}}$ the pullback metric $\underline{d}_I := (\pi_{\underline{\mathcal{K}}}|_{\underline{U}_I})^* d$ on $\underline{U}_I$ induces the quotient topology on the intermediate domain $\underline{U}_I = U_I / \Gamma_I$. By construction, these also induce Γ_I–invariant pseudometrics $d_I := (\pi_{\underline{\mathcal{K}}}|_{U_I})^* d = \pi_I^* \underline{d}_I$ on the Kuranishi domains U_I of $\mathcal{K}$. Moreover, [13, Lemma 3.1.8] shows that these (pseudo)metrics are compatible with coordinate changes. We denote the δ–balls around subsets $Q \subset |\underline{\mathcal{K}}|$, $R \subset \underline{U}_I$ and $S \subset U_I$ for $\delta > 0$ by, respectively,

$$B_\delta(Q) := \{w \in |\mathcal{K}| \mid \exists q \in Q \text{ such that } d(w, q) < \delta\},$$
$$B^I_\delta(R) := \{x \in \underline{U}_I \mid \exists r \in R \text{ such that } \underline{d}_I(x, r) < \delta\},$$
$$\widehat{B}^I_\delta(S) := \{y \in U_I \mid \exists s \in S \text{ such that } d_I(y, s) < \delta\},$$

and note that balls in the pseudometric are Γ_I–invariant preimages of balls in $\underline{U}_I$,

$$\widehat{B}^I_\delta(S) = \pi_I^{-1}(B^I_\delta(\underline{S})) \quad \text{and} \quad B_\delta(\pi_{\mathcal{K}}(S)) = B_\delta(\pi_{\underline{\mathcal{K}}}(\underline{S})). \tag{3.3.1}$$

- While the metric topology on $|\underline{\mathcal{K}}|$ is generally not compatible with the quotient topology, we know from [13, Lemma 3.1.8] that the identity map $|\mathcal{K}| \to (|\mathcal{K}|, d)$ is continuous, and thus $|\mathcal{K}|$ is a Hausdorff topology in which the metric δ–balls are open, and thus neighborhoods.

Given this setting, our goal is to construct admissible, precompact, transverse (cobordism) perturbations of the section $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ over a pruned domain category $\boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$; see Definition 3.2.4 and Lemma 3.2.3. For that purpose we will also need to fix nested (cobordism) reductions $\mathcal{C} \sqsubset \mathcal{V}$ of $\mathcal{K}$. These induce the following crucial data, on which the iterative construction of perturbations depends. The claims here are all consequences of [13, Theorem 5.1.6(iii)] and [14, Lemma 7.3.4, Proposition 7.3.10] applied to $\underline{\mathcal{K}}$ together with (3.3.1) and properness of the projections $\pi_I\colon U_I \to \underline{U}_I$ established in Lemma 2.1.5(i).

- Given a reduction $\mathcal{V}$ of $\mathcal{K}$, there exists $\delta_{\mathcal{V}} \in \left(0, \frac{1}{4}\right]$ such that for any $\delta < \delta_{\mathcal{V}}$,

$$\widehat{B}^I_{2\delta}(V_I) \sqsubset U_I \quad \forall I \in \mathcal{I}_{\mathcal{K}},$$
$$B_{2\delta}(\pi_{\mathcal{K}}(V_I)) \cap B_{2\delta}(\pi_{\mathcal{K}}(V_J)) \neq \varnothing \quad \Longrightarrow \quad I \subset J \text{ or } J \subset I.$$

  This gives rise to a continuum of nested reductions $V_I \sqsubset \cdots V_I^{k''} \sqsubset V_I^{k'} \cdots \sqsubset V_I^0$ for $k'' > k' > 0$, which for $k \geq 0$ are given by

$$V_I^k := \widehat{B}^I_{2^{-k}\delta}(V_I) = \pi_I^{-1}(\underline{V}_I^k) \sqsubset U_I \quad \text{with} \quad \underline{V}_I^k := B^I_{2^{-k}\delta}(\underline{V}_I).$$

- For suitable $k \geq 0$, the iteration will construct ν_J by extension of the pullbacks $\rho_{IJ}^*\nu_I$, which are defined for $I \subsetneq J$ on $N_{JI}^k := V_J^k \cap \pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(V_I^k))$, also given as

$$N_{JI}^k = V_J^k \cap \rho_{IJ}^{-1}(V_I^k) = \pi_J^{-1}(\underline{N}_{JI}^k) \quad \text{with} \quad \underline{N}_{JI}^k := \underline{V}_J^k \cap \underline{\phi}_{IJ}(\underline{V}_I^k \cap \underline{U}_{IJ}).$$

- We need to make a choice of *equivariant norms on the obstruction spaces* as follows. For each basic chart $i \in \{1, \dots, N\}$ we choose a Γ_i–invariant norm $\|\cdot\|$ on E_i. Then the Γ_J–invariant norm on E_J for each $J \in \mathcal{I}_{\mathcal{K}}$ is given by

$$\|e\| := \left\| \sum_{i \in J} \widehat{\phi}_{iJ}(e_i) \right\| := \max_{i \in J} \|e_i\| \quad \forall e = \sum_{i \in J} \widehat{\phi}_{iJ}(e_i) \in E_J.$$

- While the sections $s_I\colon U_I \to E_I$ only induce continuous maps $\underline{s}_I\colon \underline{U}_I \to E_I/\Gamma_I$ to the quotient of obstruction spaces, equivariance of the norms guarantees that the norm of sections descends to a continuous function $\|\underline{s}_I\|\colon \underline{U}_I \to [0,\infty)$ given by $x \mapsto \|s_I(y)\|$ for any $y \in \pi_I^{-1}(x)$. These functions provide (rather nontransverse) topological Kuranishi charts over the intermediate domain with the same footprint: $\underline{\psi}_I$ maps $\|\underline{s}_I\|^{-1}(0) = s_I^{-1}(0)/\Gamma_I$ homeomorphically to F_I.
- Given equivariant norms $\|\cdot\|$, nested reductions $\mathcal{C} \sqsubset \mathcal{V}$ and $0<\delta<\delta_{\mathcal{V}}$, we have

$$\begin{aligned}\sigma(\mathcal{V},\mathcal{C},\|\cdot\|,\delta) &:= \min_{J\in\mathcal{I}_{\mathcal{K}}} \inf\Big\{ \|s_J(x)\| \;\Big|\; x \in \overline{V_J^{|J|}} \setminus \Big(\widetilde{C}_J \cup \bigcup_{I\subsetneq J} \widehat{B}^J_{\eta_{|J|-\frac12}}\big(N_{JI}^{|J|-\frac14}\big)\Big)\Big\} \\ &= \min_{J\in\mathcal{I}_{\mathcal{K}}} \inf\Big\{ \|\underline{s}_J\|(y) \;\Big|\; y \in \overline{\underline{V}_J^{|J|}} \setminus \Big(\underline{\widetilde{C}}_J \cup \bigcup_{I\subsetneq J} B^J_{\eta_{|J|-\frac12}}\big(\underline{N}_{JI}^{|J|-\frac14}\big)\Big)\Big\} \\ &> 0,\end{aligned}$$

  where $\eta_{k-\frac12} := 2^{-k+\frac12}\big(1-2^{-\frac14}\big)\delta$ and

$$\widetilde{C}_J := \bigcup_{K\supset J} \rho_{JK}(C_K) \;=\; \pi_J^{-1}(\underline{\widetilde{C}}_J), \quad \text{with} \quad \underline{\widetilde{C}}_J := \bigcup_{K\supset J} \underline{\phi}_{JK}^{-1}(\underline{C}_K),$$

  is a set containing $s_J^{-1}(0) = \pi_J^{-1}(\|\underline{s}_J\|^{-1}(0))$.
- In the case of a metric tame Kuranishi cobordism $(\mathcal{K},d)$ with equivariant norms $\|\cdot\|$ and nested cobordism reductions $\mathcal{C}\sqsubset\mathcal{V}$, let $\varepsilon>0$ be the smallest of the collar widths of $\mathcal{K}$, d, $\mathcal{C}$ and $\mathcal{V}$. Then for $0<\delta<\min\{\varepsilon,\delta_{\mathcal{V}}\}$, we obtain positive numbers

$$\sigma'(\mathcal{V},\mathcal{C},\|\cdot\|,\delta) := \min_{J\in\mathcal{I}_{\mathcal{K}}} \inf\Big\{\|s_J(x)\| \;\Big|\; x\in \overline{V_J^{|J|+1}} \setminus \Big(\widetilde{C}_J \cup \bigcup_{I\subsetneq J} \widehat{B}^J_{\eta_{|J|+\frac12}}\big(N_{JI}^{|J|+\frac34}\big)\Big)\Big\},$$

$$\sigma_{\mathrm{rel}}(\mathcal{V},\mathcal{C},\|\cdot\|,\delta) := \min\big(\{\sigma'(\mathcal{V},\mathcal{C},\|\cdot\|,\delta)\} \cup \{\sigma(\partial^\alpha\mathcal{V},\partial^\alpha\mathcal{C},\partial^\alpha\|\cdot\|,\delta) \mid \alpha=0,1\}\big).$$

  Here $\partial^\alpha\|\cdot\|$ denotes the collection of equivariant norms on E_I for $I\in\mathcal{I}_{\partial^\alpha\mathcal{K}}\subset\mathcal{I}_{\mathcal{K}}$.

The constants $\delta_{\mathcal{V}}$ and $\sigma(\mathcal{V},\mathcal{C},\|\cdot\|,\delta)$ defined here will control the permitted support and norm of the perturbation ν for a Kuranishi atlas. In particular, $\delta_{\mathcal{V}}$ measures the separation between the components $V_I \neq V_J$ of the reduction $\mathcal{V}$, while $\sigma(\mathcal{V},\mathcal{C},\|\cdot\|,\delta)$ measures the minimal norm of $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ on the complement of an open neighborhood of the set $\pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(\mathcal{C}))$, in which all perturbed zero sets will need to be contained. We will construct perturbations $\nu = (\nu_I\colon V_I \to E_I)_{I\in\mathcal{I}_{\mathcal{K}}}$ by an iteration which constructs and controls each ν_I over the larger set $V_I^{|I|}$. Here the domains are determined by a choice of $0<\delta<\delta_{\mathcal{V}}$, and we ensure that the perturbed zero sets are contained in

$\pi_{\mathcal{K}}^{-1}(\pi_{\mathcal{K}}(\mathcal{C}))$ by bounding the perturbations $\|\nu_I\| < \sigma$ by some $0 < \sigma < \sigma(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta)$. In order to prove uniqueness of the VMC, we moreover have to interpolate between any two such perturbations. This requires the adjusted bound $\sigma_{\mathrm{rel}}(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta)$ on the norm of cobordism perturbations for the following reason. The construction of a cobordism perturbation with prescribed boundary values is achieved by an iteration on the domains $V_I^{|I|+1}$ instead of $V_I^{|I|}$, which guarantees that the boundary values — which got constructed in iterations over $\partial^\alpha V_I^{|I|}$ — are given on sufficiently large boundary collars. In view of this, it is also necessary to keep track of the refined properties arising from the iterative construction of a perturbation by the following notion of $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted, as well as a stronger notion which guarantees extensions to Kuranishi concordances.

Definition 3.3.1 Given nested reductions $\mathcal{C} \sqsubset \mathcal{V}$ of a metric tame Kuranishi atlas $(\mathcal{K}, d)$, a choice of equivariant norms $\|\cdot\|$ on the obstruction spaces, and constants $0 < \delta < \delta_{\mathcal{V}}$ and $0 < \sigma \le \sigma(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta)$, we say that a perturbation ν of $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ is $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$*–adapted* if the sections $\nu_I \colon V_I \to E_I$ extend to sections over $V_I^{|I|}$ (also denoted ν_I) so that the following conditions hold for every $k = 1, \dots, M_{\mathcal{K}} := \max_{I \in \mathcal{I}_{\mathcal{K}}} |I|$:

(a) The perturbations are compatible in the sense that for $H \subsetneq I$ with $|I| \le k$,
$$\nu_I|_{\rho_{HI}^{-1}(V_H^k) \cap V_I^k} = \widehat{\phi}_{HI} \circ \nu_H \circ \rho_{HI}|_{\rho_{HI}^{-1}(V_H^k) \cap V_I^k}.$$

(b) The perturbed sections are transverse; that is, $(s_I|_{V_I^k} + \nu_I) \pitchfork 0$ for each $|I| \le k$.

(c) The perturbations are strongly admissible; that is, for all $H \subsetneq I$ and $|I| \le k$ we have $\nu_I(\widehat{B}^I_{\eta_k}(N^k_{IH})) \subset \widehat{\phi}_{HI}(E_H)$.

(d) The perturbed zero sets are controlled by $\pi_{\mathcal{K}}\big((s_I|_{V_I^k} + \nu_I)^{-1}(0)\big) \subset \pi_{\mathcal{K}}(\mathcal{C})$ for $|I| \le k$.

e) The perturbations are small; that is, $\sup_{x \in V_I^k} \|\nu_I(x)\| < \sigma$ for $|I| \le k$.

Also, we say that a perturbation ν is *strongly* $(\mathcal{V}, \mathcal{C})$*–adapted* if it is a $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted perturbation of $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ for some choice of equivariant norms $\|\cdot\|$ and constants $0 < \delta < \delta_{\mathcal{V}}$, and using the product metric on $[0,1] \times |\mathcal{K}|$ we have
$$\begin{aligned} 0 < \sigma &\le \sigma_{\mathrm{rel}}([0,1] \times \mathcal{V}, [0,1] \times \mathcal{C}, \|\cdot\|, \delta) \\ &= \min\big\{\sigma(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta), \sigma'([0,1] \times \mathcal{V}, [0,1] \times \mathcal{C}, \|\cdot\|, \delta)\big\}, \end{aligned}$$

Remark 3.3.2 (i) Adapted perturbations are automatically admissible, precompact and transverse in the sense of Definition 3.2.4. Indeed, these properties are guaranteed by the inclusions $V_I \subset V_I^k$ and the fact that strong admissibility $\nu_I(x) \in \operatorname{im} \widehat{\phi}_{HI}$ for

$x \in \widehat{B}^I_{\eta_k}(N^k_{IH})$ for $H \subsetneq I$ implies admissibility $\operatorname{im} \mathrm{d}_y \nu_I \subset \operatorname{im} \widehat{\phi}_{HI}$ for $y \in \widetilde{V}_{HI} = V_I \cap \rho_{HI}^{-1}(V_H) \subset V_I^k \cap \rho_{IJ}^{-1}(V_I^k) = N^k_{IH}$.

(ii) The admissibility condition is crucial for the transfer of transversality as follows. Let ν be an admissible perturbation, and let $z \in V_I$ and $w \in V_J$ so that $\pi_{\mathcal{K}}(z) = \pi_{\mathcal{K}}(w) \in |\mathcal{K}|$. Then z is a transverse zero of $s_I|_{V_I} + \nu_I$ if and only if w is a transverse zero of $s_J|_{V_J} + \nu_J$.

Indeed, by the reduction property we can assume without loss of generality that $I \subset J$ and thus $z = \rho_{IJ}(w)$. Since ρ_{IJ} is a regular covering, we can pick a local inverse $\phi_I J$ so that $w = \phi_{IJ}(z)$. Then the proof of [14, Lemma 7.2.4] directly applies, using the index condition in terms of ϕ_{IJ}.

(iii) Any $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted perturbation for fixed $\mathcal{V}$, $\mathcal{C}$, $\|\cdot\|$, δ and sufficiently small $\sigma > 0$ is in fact strongly adapted. Indeed, given the product structure of all sets and maps involved in the definition of σ', we can rewrite the condition on $\sigma > 0$ in the definition of strong adaptivity as $\sigma < \|s_J(x)\|$ for all $x \in \overline{V_J^k} \setminus \big(\widetilde{C}_J \cup \bigcup_{I \subsetneq J} \widehat{B}^J_{\eta_{k-\frac{1}{2}}}(N_{JI}^{k-\frac{1}{4}})\big)$, $J \in \mathcal{I}_{\mathcal{K}}$ and $k \in \{|J|, |J|+1\}$. $\diamond$

By the above remark, the following in particular proves the existence of admissible, precompact, transverse perturbations as well as strongly adapted perturbations.

Proposition 3.3.3 (i) *Let $(\mathcal{K}, d)$ be a metric tame Kuranishi atlas with nested reductions $\mathcal{C} \sqsubset \mathcal{V}$ and equivariant norms $\|\cdot\|$ on the obstruction spaces. Then for any $0 < \delta < \delta_{\mathcal{V}}$ and $0 < \sigma \le \sigma(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta)$, there exists a $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted perturbation ν of $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$.*

(ii) *Let $(\mathcal{K}, d)$ be a metric tame Kuranishi cobordism with nested cobordism reductions $\mathcal{C} \sqsubset \mathcal{V}$, equivariant norms $\|\cdot\|$ on the obstruction spaces, and minimal collar width $\varepsilon > 0$ of $(\mathcal{K}, d)$ and the reductions $\mathcal{C}, \mathcal{V}$. Then, given $0 < \delta < \min\{\varepsilon, \delta_{\mathcal{V}}\}$, $0 < \sigma \le \sigma_{\mathrm{rel}}(\delta, \mathcal{V}, \mathcal{C})$, and perturbations ν^α of $\mathfrak{s}_{\partial^\alpha \mathcal{K}}|_{\partial^\alpha \mathcal{V}}^{\setminus\Gamma}$ for $\alpha = 0, 1$ that are $(\partial^\alpha \mathcal{V}, \partial^\alpha \mathcal{C}, \delta, \sigma)$–adapted, there exists an admissible, precompact, transverse cobordism perturbation ν of $\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ with $\pi_{\mathcal{K}}\big((\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma} + \nu)^{-1}(0)\big) \subset \pi_{\mathcal{K}}(\mathcal{C})$ and $\nu|_{\partial^\alpha \mathcal{V}} = \nu^\alpha$ for $\alpha = 0, 1$.*

(iii) *In the case of a product cobordism $[0,1] \times \mathcal{K}$ with product metric and nested product reductions $[0,1] \times \mathcal{C} \sqsubset [0,1] \times \mathcal{V}$, (ii) holds for $0 < \delta < \delta_{[0,1] \times \mathcal{V}}$ without restriction from the collar width.*

Proof As explained in [14, Remark 7.3.2], the iterative constructions in [14, Propositions 7.3.7 and 7.3.10] generalize directly to our setup based on the pruned domain category $\mathcal{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$. We indicated the necessary adjustments in a series of footnotes in the proofs of [14]. Beyond the above setting and notations, this requires the following two systematic changes.

Firstly, all relationships between (or definitions/constructions of) subsets of $\mathrm{Obj}_{\mathcal{B}_{\mathcal{K}}} = \bigcup_{I\in\mathcal{I}_{\mathcal{K}}} U_I$ in [14] should be replaced by two statements: one for subsets of $\mathrm{Obj}_{\mathcal{B}_{\underline{\mathcal{K}}}} = \bigcup_{I\in\mathcal{I}_{\mathcal{K}}} \underline{U}_I$ in the intermediate atlas $\underline{\mathcal{K}}$, and one for subsets in the pruned domain category $\mathcal{B}_{\mathcal{K}}|_{\mathcal{V}}^{\setminus\Gamma}$ with B_δ replaced by $\widehat{B}_\delta$. These two statements will always be equivalent via the projection π_I. Statements can then be checked by working in the intermediate category, but they will be applied on the level of the pruned domain category. Here it is crucial to know that the projections $\pi_I\colon U_I \to \underline{U}_I$ are continuous (by definition of the quotient topology) and proper by Lemma 2.1.5(i).

Secondly, our goal of constructing a precompact, transverse, admissible (cobordism) perturbation $\nu\colon \boldsymbol{B}_{\mathcal{K}}|_{\mathcal{V}} \to \boldsymbol{E}_{\mathcal{K}}|_{\mathcal{V}}$ is essentially the same as that of Definitions 7.2.1, 7.2.5 and 7.2.6 in [14]. Writing it in terms of the maps $\nu = (\nu_I\colon V_I \to E_I)_{I\in\mathcal{I}_{\mathcal{K}}}$, the only difference is that the compatibility conditions in [14, Equation (7.2.1)],

$$\nu_J\big|_{N_{JI}} = \widehat{\phi}_{IJ}\circ\nu_I\circ\phi_{IJ}^{-1}\big|_{N_{JI}} \quad \text{on} \quad N_{JI} := V_J\cap\phi_{IJ}(V_I\cap U_{IJ})$$

for all $I \subsetneq J$, are replaced by

$$\nu_J\big|_{\widetilde{V}_{IJ}} = \widehat{\phi}_{IJ}\circ\nu_I\circ\rho_{IJ}\big|_{\widetilde{V}_{IJ}} \quad \text{on} \quad \widetilde{V}_{IJ} := V_J\cap\rho_{IJ}^{-1}(V_I),$$

and the precompactness conditions in [14, Equation (7.2.5)],

$$(s_J|_{V_J}+\nu_J)^{-1}(0) \subset \bigcup_{H\supset J}\phi_{JH}^{-1}(C_H)\cup\bigcup_{H\subsetneq J}\phi_{HJ}(C_H)$$

for all $J\in\mathcal{I}_{\mathcal{K}}$, are replaced by (3.2.5) above,

$$(s_J|_{V_J}+\nu_J)^{-1}(0) \subset \bigcup_{H\supset J}\rho_{JH}(C_H)\cup\bigcup_{H\subsetneq J}\rho_{HJ}^{-1}(C_H).$$

Here our setup guarantees that $\rho_{IJ}\colon \widetilde{V}_{IJ}\to V_I\cap\rho_{IJ}(V_J)\subset U_{IJ}$ is a regular covering (ie local diffeomorphism with fibers given by the free action of a finite group $\Gamma_{J\setminus I}\cong\Gamma_J/\Gamma_I$) analogous to $\phi_{IJ}^{-1}\colon N_{IJ}\to V_I\cap\phi_{IJ}^{-1}(V_J)\subset U_{IJ}$ in [14], which is a regular covering with trivial fibers. Thus to adapt the proofs of [14] one should replace ϕ_{IJ} with ρ_{IJ}^{-1} and identify $N_{IJ} = \widetilde{V}_{IJ}$. □

Finally, we make the additional choice of an orientation of the Kuranishi atlases or cobordisms in the sense of Definition 3.1.10 to prove Theorem A from the introduction.

Definition 3.3.4 Let $(\mathcal{K},\sigma)$ be an oriented weak Kuranishi atlas of dimension D on a compact, metrizable space X. Then its *virtual moduli cycle* $\mathcal{Z}^{\mathcal{K}} := [(|\boldsymbol{Z}^\nu|_{\mathcal{H}}, \Lambda^\nu)]$ is the cobordism class of weighted branched manifolds (without boundary) of dimension D given by the choices of a preshrunk tame shrinking $\mathcal{K}_{\rm sh}$ of $\mathcal{K}$, an admissible metric on $|\mathcal{K}_{\rm sh}|$, nested reductions $\mathcal{C} \sqsubset \mathcal{V}$ of $\mathcal{K}_{\rm sh}$ and a strongly $(\mathcal{V},\mathcal{C})$–adapted perturbation ν.

Moreover, the *virtual fundamental class*

$$[X]^{\rm vir}_{\mathcal{K}} := |\psi_{\mathcal{K}_{\rm sh}}|_*(\varprojlim[\iota^{\nu_k}]) \in \check{H}_D(X;\mathbb{Q})$$

is constructed as follows:

- Choose a preshrunk tame shrinking $\mathcal{K}_{\rm sh}$ of $\mathcal{K}$, an admissible metric on $|\mathcal{K}_{\rm sh}|$ and a nested sequence of open sets $\mathcal{W}_{k+1} \subset \mathcal{W}_k \subset (|\mathcal{K}_{\rm sh}|, d)$ with $\bigcap_{k\in\mathbb{N}} \mathcal{W}_k = |\mathfrak{s}_{\mathcal{K}_{\rm sh}}^{-1}(0)|$. (These exist by Theorem 2.5.3, and taking for instance $\mathcal{W}_k = B_{\frac{1}{k}}(|\mathfrak{s}_{\mathcal{K}_{\rm sh}}^{-1}(0)|$.) Then equip $\mathcal{K}_{\rm sh}$ with the orientation induced from $\mathcal{K}$ by Lemma 3.1.12.
- For each $k \in \mathbb{N}$ choose a $(\mathcal{V}_k,\mathcal{C}_k)$–adapted perturbation ν_k of $\mathfrak{s}_{\mathcal{K}_{\rm sh}}|_{\mathcal{V}_k}^{\setminus\Gamma}$ for some nested reductions $\mathcal{C}_k \sqsubset \mathcal{V}_k$ with $\pi_{\mathcal{K}_{\rm sh}}(\mathcal{C}_k) \subset \mathcal{W}_k$. (These exist by Remark 3.2.2 and Proposition 3.3.3.)
- Denote by $[|\iota^{\nu_k}|_{\mathcal{H}}] \in \check{H}_D(\mathcal{W}_k;\mathbb{Q})$ the Čech homology classes induced by the maps

$$|\iota^{\nu_k}|_{\mathcal{H}}\colon (|\boldsymbol{Z}^{\nu_k}|_{\mathcal{H}}, \Lambda^{\nu_k}) \hookrightarrow \mathcal{W}_k \subset (|\mathcal{K}_{\rm sh}|, d),$$

take their inverse limit under pushforward with the inclusions $\mathcal{W}_{k+1} \hookrightarrow \mathcal{W}_k$, and finally take the pushforward under the homeomorphism $|\psi_{\mathcal{K}_{\rm sh}}| = \iota_{\mathcal{K}_{\rm sh}}^{-1}\colon |\mathfrak{s}_{\mathcal{K}_{\rm sh}}^{-1}(0)| \to X$ from Lemma 2.3.9(iii).

Note here that every weighted branched manifold (Y,Λ_Y) has a fundamental class $[Y] \in H_d(Y);\mathbb{Q})$ by Proposition A.7. This was constructed in [8] as an element of rational singular homology, and by the discussion after [13, Remark 8.2.4] gives a well-defined element in rational Čech homology. Thus the above construction makes sense. Further, Lemma 2.3.9(iii) identifies the quotient topology on $|\mathfrak{s}_{\mathcal{K}_{\rm sh}}^{-1}(0)|$ with the relative topology induced by the embedding $|\mathfrak{s}_{\mathcal{K}_{\rm sh}}^{-1}(0)| \hookrightarrow |\mathcal{K}_{\rm sh}|$. The latter is also identified with the metric topology given by restriction of d, due to the nesting uniqueness of Hausdorff topologies and the fact that the identity map $|\mathcal{K}| \to (|\mathcal{K}|, d)$ is continuous; see [13, Lemma 3.1.8, Remark 3.1.15]. Hence there is no ambiguity of topologies in the isomorphism explained in [14, Remark 8.2.4] and used in the definition of $[X]^{\rm vir}_{\mathcal{K}}$,

$$\check{H}_D(|\mathfrak{s}_{\mathcal{K}}^{-1}(0)|;\mathbb{Q}) \xrightarrow{\cong} \varprojlim \check{H}_D(\mathcal{W}_k;\mathbb{Q}).$$

Finally, we can prove our main theorem: the VMC/VFC are well defined and are invariants of the oriented weak Kuranishi cobordism class. The proof uses the same line of argument as [14, Theorems 8.2.2 and 8.2.5], just replacing manifolds with weighted branched manifolds. We summarize and unify these arguments here for ease of reference.

Theorem 3.3.5 (i) *The virtual moduli cycle $\mathcal{Z}^{\mathcal{K}}$ and the virtual fundamental class $[X]^{\mathrm{vir}}_{\mathcal{K}}$ are well defined and independent of the cobordism class of oriented weak Kuranishi atlases on a fixed compact, metrizable space X.*

(ii) *Let $\mathcal{K}$ be an oriented weak Kuranishi cobordism, and choose strongly adapted perturbations ν^α in the definition of $\mathcal{Z}^{\partial^\alpha\mathcal{K}} = [(|\boldsymbol{Z}^{\nu^\alpha}|_{\mathcal{H}}, \Lambda^{\nu^\alpha})]$ for $\alpha = 0, 1$. Then the perturbed zero sets $(|\boldsymbol{Z}^{\nu^0}|_{\mathcal{H}}, \Lambda^{\nu^0}) \sim (|\boldsymbol{Z}^{\nu^1}|_{\mathcal{H}}, \Lambda^{\nu^1})$ are cobordant as weighted branched manifolds, and thus $\mathcal{Z}^{\partial^0\mathcal{K}} = \mathcal{Z}^{\partial^1\mathcal{K}}$.*

(iii) *Let $\mathcal{K}$ be an oriented weak Kuranishi cobordism of dimension $D+1$ on a compact, metrizable collared cobordism $(Y, \iota^0_Y, \iota^1_Y)$. Then the virtual fundamental classes $[\partial^\alpha Y]^{\mathrm{vir}}_{\partial^\alpha\mathcal{K}} \simeq [\partial^1 Y]^{\mathrm{vir}}_{\partial^1\mathcal{K}}$ of the boundary restrictions are homologous in Y,*

$$(\iota^0_Y)_*\big([\partial^0 Y]^{\mathrm{vir}}_{\partial^0\mathcal{K}}\big) = (\iota^1_Y)_*\big([\partial^1 Y]^{\mathrm{vir}}_{\partial^1\mathcal{K}}\big) \in \check{H}_D(Y;\mathbb{Q}).$$

Proof First note that all the necessary choices of data exist, as noted in Definition 3.3.4. Given such choices, Step 1 below constructs a representative of the virtual moduli cycle, and Step 5 constructs the virtual fundamental class. To prove independence of those choices in (i), we use transitivity of the cobordism relation for compact weighted branched manifolds to prove increasing independence of choices in Steps 1–5. Parts (ii) and (iii) are then proven in Step 6. In the following, all Kuranishi atlases will be of dimension D, and all cobordisms of dimension $D+1$.

Step 1 *Fix an oriented, metric, tame Kuranishi atlas $(\mathcal{K}, d)$, nested reductions $\mathcal{C} \sqsubset \mathcal{V}$, equivariant norms $\|\cdot\|$, and constants δ, σ such that $0 < \delta < \delta_{\mathcal{V}}$ and $0 < \sigma \le \sigma_{\mathrm{rel}}([0,1]\times\mathcal{V}, [0,1]\times\mathcal{C}, \|\cdot\|, \delta)$. Then each $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted perturbation ν induces a D–dimensional weighted branched manifold $\mathcal{Z}^\nu := (|\boldsymbol{Z}^\nu|_{\mathcal{H}}, \Lambda^\nu)$ and a cycle $|\iota^\nu|_{\mathcal{H}}\colon \mathcal{Z}^\nu \to |\mathcal{C}|$, whose respective cobordism class and Čech homology class $[|\iota^\nu|_{\mathcal{H}}] \in \check{H}_D(|\mathcal{C}|;\mathbb{Q})$ are independent of the choice of ν.*

The regularity of the perturbed zero sets is proven in Theorem 3.2.8. To prove independence of the choice of ν, we consider two $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted perturbations ν^0 and ν^1. Then Proposition 3.3.3(iii) provides an admissible, precompact, transverse cobordism perturbation ν^{01} of $\mathfrak{s}_{[0,1]\times\mathcal{K}}|^{\setminus\Gamma}_{[0,1]\times\mathcal{V}}$ with boundary restrictions $\nu^{01}|_{\{\alpha\}\times\mathcal{V}} = \nu^\alpha$ for $\alpha = 0, 1$. Moreover, by Lemma 3.1.12(iii) the orientation of $\mathcal{K}$ induces an orientation of $[0,1]\times\mathcal{K}$, whose restriction to the boundaries $\partial^\alpha([0,1]\times\mathcal{K}) = \mathcal{K}$ equals the given orientation on $\mathcal{K}$. Now Theorem 3.2.8 implies that $\mathcal{Z} := (|\boldsymbol{Z}^{\nu^{01}}|, \Lambda^{\nu^{01}})$ is a cobordism from $\partial^0\mathcal{Z} = (|\boldsymbol{Z}^{\nu^0}|, \Lambda^{\nu^0})$ to $\partial^1\mathcal{Z} = (|\boldsymbol{Z}^{\nu^1}|, \Lambda^{\nu^1})$ and induces a cycle $|\iota^{\nu^{01}}|_{\mathcal{H}}\colon \mathcal{Z} \to [0,1]\times|\mathcal{C}|$. Finally, the boundary restrictions of this cycle prove the equality $[|\iota^{\nu^0}|_{\mathcal{H}}] = [|\iota^{\nu^1}|_{\mathcal{H}}]$ in $\check{H}_D(|\mathcal{C}|;\mathbb{Q})$; see [14, Equation (8.2.6)] for the detailed homological argument.

Step 2 *Fix an oriented, metric, tame Kuranishi atlas $(\mathcal{K}, d)$ and nested reductions $\mathcal{C} \sqsubset \mathcal{V}$. Then the cobordism class of $\mathcal{Z}^\nu$, as well as $[|\iota^\nu|_{\mathcal{H}}] \in \check{H}_D(|\mathcal{C}|;\mathbb{Q})$, are independent of the choice of strongly $(\mathcal{V}, \mathcal{C})$–adapted perturbation ν.*

To prove this we consider two strongly $(\mathcal{V}, \mathcal{C})$–adapted perturbations ν^α for $\alpha = 0, 1$. Thus ν^α is $(\mathcal{V}, \mathcal{C}, \|\cdot\|^\alpha, \delta^\alpha, \sigma^\alpha)$–adapted for some choices of equivariant norms $\|\cdot\|^\alpha$ and constants $0 < \delta^\alpha < \delta_{\mathcal{V}}$ and $0 < \sigma^\alpha \le \sigma_{\mathrm{rel}}([0,1]\times\mathcal{V}, [0,1]\times\mathcal{C}, \|\cdot\|^\alpha, \delta^\alpha)$. We note that $\delta := \max(\delta^0, \delta^1) < \delta_{\mathcal{V}} = \delta_{[0,1]\times\mathcal{V}}$, pick equivariant norms $\|\cdot\|$ on $\mathcal{K}$ such that $\|\cdot\|^\alpha \le \|\cdot\|$ for $\alpha = 0, 1$, and choose

$$\sigma \le \min\{\sigma^0, \sigma^1, \sigma_{\mathrm{rel}}([0,1]\times\mathcal{V}, [0,1]\times\mathcal{C}, \|\cdot\|, \delta)\}.$$

Then Proposition 3.3.3(iii) provides an admissible, precompact, transverse cobordism perturbation ν^{01} of $\mathfrak{s}_{[0,1]\times\mathcal{K}}|^{\setminus\Gamma}_{[0,1]\times\mathcal{V}}$, whose restrictions $\widetilde{\nu}^\alpha := \nu^{01}|_{\{\alpha\}\times\mathcal{V}}$ for $\alpha = 0, 1$ are $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta, \sigma)$–adapted perturbations of $\mathfrak{s}_{\mathcal{K}}|^{\setminus\Gamma}_{\mathcal{V}}$. Since we have that $\delta^\alpha \le \delta$, $\|\nu^{01}|_{\{\alpha\}\times\mathcal{V}}\|^\alpha \le \|\nu^{01}|_{\{\alpha\}\times\mathcal{V}}\| < \sigma$ and $\sigma \le \sigma^\alpha \le \sigma_{\mathrm{rel}}([0,1]\times\mathcal{V}, [0,1]\times\mathcal{C}, \|\cdot\|, \delta^\alpha)$, they are also $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta^\alpha, \sigma^\alpha)$–adapted. Then, as in Step 1, the perturbed zero set of ν^{01} is a cobordism from $\mathcal{Z}^{\widetilde{\nu}^0}$ to $\mathcal{Z}^{\widetilde{\nu}^1}$, and the induced cycle in $[0,1]\times|\mathcal{C}|$ shows $[|\iota^{\widetilde{\nu}^0}|_{\mathcal{H}}] = [|\iota^{\widetilde{\nu}^1}|_{\mathcal{H}}]$ in $\check{H}_D(|\mathcal{C}|;\mathbb{Q})$.

Moreover, for fixed $\alpha \in \{0,1\}$ both the restriction $\widetilde{\nu}^\alpha = \nu^{01}|_{\{\alpha\}\times\mathcal{V}}$ and the given perturbation ν^α are $(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta^\alpha, \sigma^\alpha)$–adapted, so that Step 1 provides cobordisms $\mathcal{Z}^{\nu^\alpha} \sim \mathcal{Z}^{\widetilde{\nu}^\alpha}$ and identities $[|\iota^{\nu^\alpha}|_{\mathcal{H}}] = [|\iota^{\widetilde{\nu}^\alpha}|_{\mathcal{H}}]$ in $\check{H}_D(|\mathcal{C}|;\mathbb{Q})$. By transitivity of the cobordism relation this proves $\mathcal{Z}^{\nu^0} \sim \mathcal{Z}^{\nu^1}$ as claimed, and also $[|\iota^{\nu^0}|_{\mathcal{H}}] = [|\iota^{\nu^1}|_{\mathcal{H}}] \in \check{H}_D(|\mathcal{C}|;\mathbb{Q})$.

Step 3 *For a fixed oriented, metric, tame Kuranishi atlas $(\mathcal{K}, d)$, the oriented cobordism class $\mathcal{A}^{(\mathcal{K},d)}$ of weighted branched manifolds $\mathcal{Z}^\nu$ is independent of the choice of strongly adapted perturbation ν. Moreover, given any open neighborhood $\mathcal{W} \subset (|\mathcal{K}|, d)$ of $|\mathfrak{s}_{\mathcal{K}}^{-1}(0)|$, the class $A^{(\mathcal{K},d)}_{\mathcal{W}} := [|\iota^\nu|_{\mathcal{H}} \colon \mathcal{Z}^\nu \to \mathcal{W}] \in \check{H}_D(\mathcal{W};\mathbb{Q})$ is independent of the choice of strongly $(\mathcal{V}, \mathcal{C})$–adapted perturbation ν for nested reductions $\mathcal{C} \sqsubset \mathcal{V}$ with $\pi_{\mathcal{K}}(\mathcal{C}) \subset \mathcal{W}$.*

To prove this we consider two strongly $(\mathcal{V}^\alpha, \mathcal{C}^\alpha)$–adapted perturbations ν^α with respect to nested reductions $\mathcal{C}^\alpha \sqsubset \mathcal{V}^\alpha$ with $\pi_{\mathcal{K}}(\mathcal{C}) \subset \mathcal{W}$, equivariant norms $\|\cdot\|^\alpha$ and admissible metrics d^α for $\alpha = 0, 1$. Remark 3.2.2 provides a nested cobordism reduction $\mathcal{C} \sqsubset \mathcal{V}$ of $[0,1]\times\mathcal{K}$ with $\partial^\alpha\mathcal{C} = \mathcal{C}^\alpha$, $\partial^\alpha\mathcal{V} = \mathcal{V}^\alpha$ and $\pi_{[0,1]\times\mathcal{C}} \subset [0,1]\times\mathcal{W}$. Now pick equivariant norms $\|\cdot\|$ on $\mathcal{K}$ such that $\|\cdot\|^\alpha \le \|\cdot\|$ for $\alpha = 0, 1$, and choose $0 < \delta < \delta_{\mathcal{V}}$ smaller than the collar width of d, $\mathcal{V}$, and $\mathcal{C}$. Then, for any $0 < \sigma \le \sigma_{\mathrm{rel}}(\mathcal{V}, \mathcal{C}, \|\cdot\|, \delta)$, Proposition 3.3.3(ii) provides an admissible, precompact, transverse cobordism perturbation ν^{01} of $\mathfrak{s}_{[0,1]\times\mathcal{K}}|^{\setminus\Gamma}_{\mathcal{V}}$ whose boundary restrictions $\widetilde{\nu}^\alpha := \nu^{01}|_{\partial^\alpha\mathcal{V}}$ for $\alpha = 0, 1$ are $(\mathcal{V}^\alpha, \mathcal{C}^\alpha, \|\cdot\|, \delta, \sigma)$–adapted perturbations of $\mathfrak{s}_{\mathcal{K}}|^{\setminus\Gamma}_{\mathcal{V}^\alpha}$.

As before, $\mathcal{Z}^{\nu^{01}}$ is an oriented cobordism from $\mathcal{Z}^{\widetilde{\nu}^0}$ to $\mathcal{Z}^{\widetilde{\nu}^1}$ and induces a cycle in $[0,1]\times\mathcal{W}$ that shows $[|\iota^{\widetilde{\nu}^0}|_{\mathcal{H}}] = [|\iota^{\widetilde{\nu}^1}|_{\mathcal{H}}]$ in $\check{H}_D(\mathcal{W};\mathbb{Q})$. Moreover, we can pick $\sigma \le \sigma_{\rm rel}([0,1]\times\mathcal{V}^\alpha, [0,1]\times\mathcal{C}^\alpha, \|\cdot\|^\alpha, \delta)$ for $\alpha = 0,1$, so that each $\nu^{01}|_{\partial^\alpha\mathcal{V}}$ is also strongly $(\mathcal{V}^\alpha,\mathcal{C}^\alpha)$–adapted. Then the claim follows by transitivity as in Step 2.

Step 4 *Let $(\mathcal{K},d)$ be an oriented, metric, tame Kuranishi atlas, and let $\mathcal{W}_k \subset (|\mathcal{K}|,d)$ be a nested sequence of open sets with $\bigcap_{k\in\mathbb{N}}\mathcal{W}_k = |\mathfrak{s}_{\mathcal{K}}^{-1}(0)|$ as in Definition 3.3.4. Then the Čech homology class*

$$A^{(\mathcal{K},d)} := \varprojlim A^{(\mathcal{K},d)}_{\mathcal{W}_k} \in \check{H}_D(|\mathfrak{s}_{\mathcal{K}}^{-1}(0)|;\mathbb{Q})$$

is well defined and independent of the choice of nested sequence $(\mathcal{W}_k)_{k\in\mathbb{N}}$.

The pushforward $\check{H}_D(\mathcal{W}_{k+1};\mathbb{Q}) \to \check{H}_D(\mathcal{W}_k;\mathbb{Q})$ by the inclusion $\mathcal{I}_{k+1}\colon \mathcal{W}_{k+1}\to\mathcal{W}_k$ maps $A^{(\mathcal{K},d)}_{\mathcal{W}_{k+1}} = [|\iota^{\nu_{k+1}}|_{\mathcal{H}}]$ to $A^{(\mathcal{K},d)}_{\mathcal{W}_k}$ since any strongly adapted perturbation ν_{k+1} with respect to nested reductions $\mathcal{C}_{k+1}\sqsubset\mathcal{V}_{k+1}$ with $\pi_{\mathcal{K}}(\mathcal{C}_{k+1})\subset\mathcal{W}_{k+1}$ can also be used as strongly adapted perturbation for $A^{(\mathcal{K},d)}_{\mathcal{W}_k}$. This shows that the homology classes $A^{(\mathcal{K},d)}_{\mathcal{W}_k}$ form an inverse system and thus have a well-defined inverse limit. To see that this limit is independent of the choice of nested sequence, note that the intersection $\mathcal{W}_k := \mathcal{W}^0_k\cap\mathcal{W}^1_k$ of any two such sequences $(\mathcal{W}^\alpha_k)_{k\in\mathbb{N}}$ is another nested sequence of open sets with $\bigcap_{k\in\mathbb{N}}\mathcal{W}_k = |\mathfrak{s}_{\mathcal{K}}^{-1}(0)|$. Now choose a sequence of strongly adapted perturbations ν_k with respect to nested reductions $\mathcal{C}_k\sqsubset\mathcal{V}_k$ with $\pi_{\mathcal{K}}(\mathcal{C}_k)\subset\mathcal{W}_k$, then these also fit the requirements for the larger open sets $\mathcal{W}^\alpha_k$ and hence the inclusions $\mathcal{W}_k\hookrightarrow\mathcal{W}^\alpha_k$ push $[|\iota^{\nu_k}|_{\mathcal{H}}]\in H_D(\mathcal{W}_k;\mathbb{Q})$ forward to $[|\iota^{\nu_k}|_{\mathcal{H}}]\in H_D(\mathcal{W}^\alpha_k;\mathbb{Q})$. Hence, by the definition of the inverse limit, we have equality

$$\varprojlim A^{(\mathcal{K},d)}_{\mathcal{W}^0_k} = \varprojlim A^{(\mathcal{K},d)}_{\mathcal{W}_k} = \varprojlim A^{(\mathcal{K},d)}_{\mathcal{W}^0_k} \in \check{H}_D(|\mathfrak{s}_{\mathcal{K}}^{-1}(0)|;\mathbb{Q}).$$

Step 5 *Given an oriented weak Kuranishi atlas $\mathcal{K}$, the cobordism class $\mathcal{Z}^{\mathcal{K}} := \mathcal{A}^{(\mathcal{K}_{\rm sh},d)}$ of weighted branched manifolds in Step 3 and the pullback $[X]^{\rm vir}_{\mathcal{K}} := |\psi_{\mathcal{K}_{\rm sh}}|_* A^{(\mathcal{K},d)} \in \check{H}_D\big(X;\mathbb{Q}\big)$ of the Čech homology classes in Step 4 are independent of the choice of tame shrinking $\mathcal{K}_{\rm sh}$ of $\mathcal{K}$ and admissible metric d on $|\mathcal{K}_{\rm sh}|$.*

Here the pushforward under $|\psi_{\mathcal{K}_{\rm sh}}|$ is well defined since this is a homeomorphism by Lemma 2.3.9(iii). Given different choices $(\mathcal{K}^\alpha_{\rm sh}, d^\alpha)$ of metric tame shrinkings of $\mathcal{K}$ and strongly adapted perturbations ν^α and $(\nu^\alpha_k)_{k\in\mathbb{N}}$) that define $\mathcal{A}^{(\mathcal{K}^\alpha_{\rm sh},d^\alpha)}\sim\mathcal{Z}^{\nu^\alpha}$ and

$$A^{(\mathcal{K}^\alpha_{\rm sh},d^\alpha)} = \varprojlim [|\iota^{\nu^\alpha_k}|_{\mathcal{H}}] \in \check{H}_D(|\mathfrak{s}_{\mathcal{K}^\alpha_{\rm sh}}^{-1}(0)|;\mathbb{Q})$$

respectively, we can apply Step 6 below to the cobordism $[0,1]\times\mathcal{K}$ to obtain a weighted branched cobordism from $\mathcal{Z}^{\nu^0}$ to $\mathcal{Z}^{\nu^1}$ and the identity

$$I^0_*([X]^{\rm vir}_{\mathcal{K}^0_{\rm sh}}) = I^1_*([X]^{\rm vir}_{\mathcal{K}^1_{\rm sh}}) \in \check{H}_D([0,1]\times X;\mathbb{Q})$$

with the natural boundary embeddings $I^\alpha\colon X \to \{\alpha\}\times X \subset [0,1]\times X = Y$. Further, $I^0_* = I^1_*\colon \check H_D(X;\mathbb{Q}) \to \check H_D([0,1]\times X;\mathbb{Q})$ are the same isomorphisms, because the two maps I^0, I^1 are both homotopy equivalences and homotopic to each other. Hence we obtain the identity $[X]^{\mathrm{vir}}_{\mathcal{K}^0_{\mathrm{sh}}} = [X]^{\mathrm{vir}}_{\mathcal{K}^1_{\mathrm{sh}}}$ in $\check H_D(X;\mathbb{Q})$, which proves Step 5.

Step 6 *Let $\mathcal{K}$ be an oriented weak Kuranishi cobordism over a compact collared cobordism Y. For $\alpha = 0,1$ fix choices of preshrunk tame shrinkings $\mathcal{K}^\alpha_{\mathrm{sh}}$ of $\partial^\alpha\mathcal{K}$, and admissible metrics d^α on $|\partial^\alpha\mathcal{K}|$. Then, for any choice of strongly adapted perturbations ν^α on $\mathcal{K}^\alpha_{\mathrm{sh}}$, there is a weighted branched cobordism $\mathcal{Z}^{\nu^{01}}$ from $\mathcal{Z}^{\nu^0}$ to $\mathcal{Z}^{\nu^1}$. Moreover, the VFCs of the boundary components push forward by the embeddings $\iota^\alpha_Y\colon \{\alpha\}\times\partial^\alpha Y \to Y$ to the same Čech homology class in Y,*

$$(\iota^0_Y)_*([\partial^0 Y]^{\mathrm{vir}}_{\partial^0\mathcal{K}}) = (\iota^1_Y)_*([\partial^1 Y]^{\mathrm{vir}}_{\partial^1\mathcal{K}}) \in \check H_D(Y;\mathbb{Q}).$$

First, use Theorem 2.5.3 to find a preshrunk tame shrinking $\mathcal{K}_{\mathrm{sh}}$ of $\mathcal{K}$ with $\partial^\alpha\mathcal{K}_{\mathrm{sh}} = \mathcal{K}^\alpha_{\mathrm{sh}}$, and an admissible metric d on $|\mathcal{K}_{\mathrm{sh}}|$ with boundary restrictions $d|_{|\partial^\alpha\mathcal{K}_{\mathrm{sh}}|} = d^\alpha$. If we equip $\mathcal{K}_{\mathrm{sh}}$ with the orientation induced by $\mathcal{K}$, then by Lemma 3.1.12 the induced boundary orientation on $\partial^\alpha\mathcal{K}_{\mathrm{sh}} = \mathcal{K}^\alpha_{\mathrm{sh}}$ agrees with that induced by shrinking from $\partial^\alpha\mathcal{K}$. Next, Remark 3.2.2 provides nested cobordism reductions $\mathcal{C}\sqsubset\mathcal{V}$ of $\mathcal{K}_{\mathrm{sh}}$ and we may choose equivariant norms $\|\cdot\|$ on $\mathcal{K}_{\mathrm{sh}}$. Then Proposition 3.3.3 with

$$\sigma = \min\Bigl\{\sigma_{\mathrm{rel}}(\mathcal{V},\mathcal{C},\|\cdot\|,\delta),\ \min_{\alpha=0,1}\sigma_{\mathrm{rel}}([0,1]\times\partial^\alpha\mathcal{V},[0,1]\times\partial^\alpha\mathcal{C},\partial^\alpha\|\cdot\|,\delta)\Bigr\}$$

yields an admissible, precompact, transverse cobordism perturbation ν^{01} of $\mathfrak{s}_{\mathcal{K}_{\mathrm{sh}}}|^{\setminus\Gamma}_{\mathcal{V}}$, whose restrictions $\widetilde\nu^\alpha := \nu^{01}|_{\partial^\alpha\mathcal{V}}$ for $\alpha = 0,1$ are $(\partial^\alpha\mathcal{V},\partial^\alpha\mathcal{C},\partial^\alpha\|\cdot\|,\delta,\sigma)$–adapted perturbations of $\mathfrak{s}_{\mathcal{K}^\alpha_{\mathrm{sh}}}|^{\setminus\Gamma}_{\partial^\alpha\mathcal{V}}$. In particular, these are strongly adapted by the choice of σ, and $\mathcal{Z}^{\nu^{01}}$ is a cobordism from $\mathcal{Z}_{\widetilde\nu^0}$ to $\mathcal{Z}_{\widetilde\nu^1}$. Invariance of the VMC under oriented weak Kuranishi cobordism then follows from Step 3 by transitivity of weighted branched cobordism.

To prove the identity between VFCs, we first construct a sequence of nested cobordism reductions $\mathcal{C}_k\sqsubset\mathcal{V}$ of $\mathcal{K}_{\mathrm{sh}}$ by

$$\mathcal{C}_k := \mathcal{C}\cap\pi^{-1}_{\mathcal{K}_{\mathrm{sh}}}(\mathcal{W}_k)\sqsubset\mathcal{V} \quad\text{with}\quad \mathcal{W}_k := B_{\frac{1}{k}}(\iota_{\mathcal{K}_{\mathrm{sh}}}(Y))\subset|\mathcal{K}_{\mathrm{sh}}|,$$

in addition discarding components $C_k\cap V_I$ that have empty intersection with $s_I^{-1}(0)$. With that, Proposition 3.3.3 provides admissible, precompact, transverse cobordism perturbations ν_k with $|(\mathfrak{s}_{\mathcal{K}_{sh}}|^{\setminus\Gamma}_{\mathcal{V}}+\nu_k)^{-1}(0)|\subset\mathcal{W}_k$, and with boundary restrictions $\nu^\alpha_k := \nu_k|_{\partial^\alpha\mathcal{V}}$ that are strongly adapted perturbations of $(\mathcal{K}^\alpha_{\mathrm{sh}},d^\alpha)$ for $\alpha = 0,1$. Since these boundary restrictions satisfy the requirements of Step 4, they define the Čech homology classes $A^{(\mathcal{K}^\alpha_{\mathrm{sh}},d^\alpha)} = \varprojlim[|\iota^{\nu^\alpha_k}|_{\mathcal{H}}]\in\check H_D(|\mathfrak{s}^{-1}_{\mathcal{K}^\alpha_{\mathrm{sh}}}(0)|;\mathbb{Q})$.

On the other hand, pushforward with the topological embeddings $J^\alpha\colon(|\mathcal{K}^\alpha_{\mathrm{sh}}|,d^\alpha)\to(|\mathcal{K}_{\mathrm{sh}}|,d)$ also yields Čech homology classes $J^\alpha_*[|\iota^{\nu^\alpha_k}|_{\mathcal{H}}]$ that form two inverse systems in $H_D(|\mathcal{K}_{\rm sh}|;\mathbb{Q})$. Now the cycles $\iota^{\nu_k}\colon |\mathcal{Z}^{\nu_k}| \to \mathcal{W}_k$ given by Theorem 3.2.8 give rise to identities $J^0_*[|\iota^{\nu^0_k}|_{\mathcal{H}}] = J^1_*[|\iota^{\nu^1_k}|_{\mathcal{H}}]$ in $\check{H}_D(\mathcal{W}_k;\mathbb{Q})$, and taking the inverse limit, which commutes with pushforward, we obtain $J^0_*(\varprojlim[|\iota^{\nu^0_k}|_{\mathcal{H}}]) = J^1_*(\varprojlim[|\iota^{\nu^1_k}|_{\mathcal{H}}])$ in $\check{H}_D(|\mathfrak{s}^{-1}_{\mathcal{K}_{\rm sh}}(0)|;\mathbb{Q})$. Further pushforward with $|\psi_{\mathcal{K}_{\rm sh}}|$ turns this into an equality in $\check{H}_D(Y;\mathbb{Q})$. Finally, we use the identities

$$|\psi_{\mathcal{K}_{\rm sh}}| \circ J^\alpha\big|_{\iota_{\mathcal{K}^\alpha_{\rm sh}}(\partial^\alpha Y)} = \iota^\alpha_Y \circ |\psi_{\mathcal{K}^\alpha_{\rm sh}}|$$

to obtain, in $\check{H}_D(Y;\mathbb{Q})$,

$$(|\psi_{\mathcal{K}_{\rm sh}}| \circ J^\alpha)_*(\varprojlim[i^{\nu^\alpha_k}]) = (\iota^\alpha_Y)_*\big(|\psi_{\mathcal{K}^\alpha_{\rm sh}}|_*(\varprojlim[i^{\nu^\alpha_k}])\big) = (\iota^\alpha_Y)_*[\partial^\alpha Y]^{\rm vir}_{\mathcal{K}^\alpha_{\rm sh}}.$$

This proves Step 6 since the left-hand side was shown to be independent of $\alpha = 0, 1$. □

Appendix: Groupoids and weighted branched manifolds

The purpose of this appendix is to review the definition and properties of weighted branched manifolds from [8], and slightly generalize these notions to a cobordism theory. This will be based on the following language of groupoids.

An *étale groupoid* $\boldsymbol{G}$ is a small category whose sets of objects $\mathrm{Obj}_{\boldsymbol{G}}$ and morphisms $\mathrm{Mor}_{\boldsymbol{G}}$ are equipped with the structure of a smooth manifold of a fixed finite dimension such that

- all morphisms are invertible;
- all structural maps[9] are local diffeomorphisms.

All groupoids considered in this appendix are étale. Moreover, a groupoid is called

- *proper* if the source and target map $s \times t\colon \mathrm{Mor}_{\boldsymbol{G}} \to \mathrm{Obj}_{\boldsymbol{G}} \times \mathrm{Obj}_{\boldsymbol{G}}$ is proper (ie preimages of compact sets are compact);
- *nonsingular* if there is at most one morphism between any two of its objects;
- *oriented* if its spaces of objects and morphisms are oriented manifolds and if all structural maps preserve these orientations;
- *d–dimensional* if $\mathrm{Obj}_{\boldsymbol{G}}$ and $\mathrm{Mor}_{\boldsymbol{G}}$ are d–dimensional manifolds;
- *compact* if its realization $|\boldsymbol{G}|$ is compact.

[9] The structure maps of a category are source and target maps $s,t\colon \mathrm{Mor}_{\boldsymbol{G}} \to \mathrm{Obj}_{\boldsymbol{G}}$, identity map $\mathrm{id}\colon \mathrm{Obj}_{\boldsymbol{G}} \to \mathrm{Mor}_{\boldsymbol{G}}$, and composition map $\mathrm{comp}\colon \mathrm{Mor}_{\boldsymbol{G}}\, {}_t{\times}_s\, \mathrm{Mor}_{\boldsymbol{G}} \to \mathrm{Mor}_{\boldsymbol{G}}$. If source and target are local diffeomorphisms, then the fiber product in the domain of composition is transverse and hence inherits a smooth structure. A groupoid has the additional structure map $\mathrm{inv}\colon \mathrm{Mor}_{\boldsymbol{G}} \to \mathrm{Mor}_{\boldsymbol{G}}$ given by the unique inverses.

Étale proper groupoids are often called *ep groupoids*. It is well known that in the current finite-dimensional context the properness assumption is equivalent to the condition that the realization $|\boldsymbol{G}|$ is Hausdorff.[10] Here the realization $|\boldsymbol{G}|$ of $\boldsymbol{G}$ is the quotient of the space of objects by the equivalence relation given by the morphisms, ie $x \sim y \iff \mathrm{Mor}_{\boldsymbol{G}}(x,y) \neq \varnothing$. It is equipped with the quotient topology, and the natural projection is denoted $\pi_{\boldsymbol{G}}\colon \mathrm{Obj}_{\boldsymbol{G}} \to |\boldsymbol{G}|$. In general, the realization $|\boldsymbol{G}|$ of an ep groupoid is an orbifold. It is a manifold if the groupoid is nonsingular, and an orientation of the groupoid induces an orientation of $|\boldsymbol{G}|$.

Two kinds of groupoids appear in this paper: Theorem 3.2.8 shows that the zero set of a transverse section defines a wnb groupoid (which is étale but generally not proper, and equipped with an additional weighting function; see Definition A.4). On the other hand, each Kuranishi chart $\boldsymbol{K}_I$ comprises two ep groupoids $\boldsymbol{G}_{(U_I,\Gamma_I)}$ and $\boldsymbol{G}_{(U_I\times E_I,\Gamma_I)}$, which arise from group quotients as follows.

Example A.1 (i) A group quotient (U,Γ) in the sense of Definition 2.1.1 defines an ep groupoid $\boldsymbol{G}_{(U,\Gamma)}$ with $\mathrm{Obj}_{\boldsymbol{G}} = U$, $\mathrm{Mor}_{\boldsymbol{G}} = U \times \Gamma$, $(s\times t)(u,\gamma) = (u,\gamma u)$, $\mathrm{id}(u) = (u,\mathrm{id})$, $\mathrm{comp}((u,\gamma),(\gamma u,\delta)) = (u,\delta\gamma)$, $\mathrm{inv}(u,\gamma) = (u,\gamma^{-1})$, and realization $|\boldsymbol{G}| = U/\Gamma = \underline{U}$. In particular, properness is proven in Lemma 2.1.5(i). This groupoid is nonsingular if and only if the action of Γ is free. It is oriented if U is oriented and the action of each $\gamma\in\Gamma$ preserves the orientation.

(ii) The category $\boldsymbol{B}_{\mathcal{K}}$ defined by a Kuranishi atlas with trivial obstruction spaces on a compact space X is not a groupoid, because when $I \subsetneq J$ the morphisms from U_I to U_J are not invertible. However, it is shown in [10] that $\boldsymbol{B}_{\mathcal{K}}$ may be completed to an ep groupoid with the same realization (namely, X itself) by adding appropriate inverses and composites to its set of morphisms. ◇

When we take restrictions of Kuranishi charts in the sense of Definition 2.2.6, this is reflected in the associated groupoids by an analogous notion:

- If $\boldsymbol{G}$ is an étale groupoid and $\underline{V} \subset |\boldsymbol{G}|$ is open, we define the *restriction* $\boldsymbol{G}|_{\underline{V}}$ to be the full subcategory of $\boldsymbol{G}$ with objects $\pi_{\boldsymbol{G}}^{-1}(\underline{V})$.

To discuss the theory of Kuranishi cobordisms in terms of groupoids, we need the following notions. Here we use the notation $A^0_\varepsilon := [0,\varepsilon)$ and $A^1_\varepsilon := (1-\varepsilon,1]$ for neighborhoods of $0,1\in[0,1]$ of size $\varepsilon>0$ as in [13].

- If $\boldsymbol{G}$ is a groupoid and $A\subset\mathbb{R}$ is an interval we define the *product groupoid* $A\times\boldsymbol{G}$ to be the groupoid with objects $A\times\mathrm{Obj}_{\boldsymbol{G}}$ and morphisms $A\times\mathrm{Mor}_{\boldsymbol{G}}$, and with all structural maps given by products with id_A.

[10] To see that proper groupoids have Hausdorff realization one can argue that the equivalence relation has closed graph and then use [1, Chapter I, Section 10, Exercise 19] or [13, Lemma 3.2.4].

- A *cobordism groupoid* is a triple $(\boldsymbol{G}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}})$ consisting of a compact proper groupoid $\boldsymbol{G}$ and collaring functors $\iota^\alpha_{\boldsymbol{G}}\colon A^\alpha_\varepsilon \times \partial^\alpha \boldsymbol{G} \to \boldsymbol{G}$ for $\alpha = 0, 1$. Here $\boldsymbol{G}$ is required to be "étale with boundary" in the sense that its object and morphism spaces are manifolds with boundary. Moreover, these boundaries form a strictly full[11] subcategory $\partial \boldsymbol{G}$ of $\boldsymbol{G}$ that splits, $\partial(\mathrm{Obj}_{\boldsymbol{G}}) = \mathrm{Obj}_{\partial^0 \boldsymbol{G}} \sqcup \mathrm{Obj}_{\partial^1 \boldsymbol{G}}$, $\partial(\mathrm{Mor}_{\boldsymbol{G}}) = \mathrm{Mor}_{\partial^0 \boldsymbol{G}} \sqcup \mathrm{Mor}_{\partial^1 \boldsymbol{G}}$, into the disjoint union of two ep groupoids $\partial^0 \boldsymbol{G}$ and $\partial^1 \boldsymbol{G}$. Finally, the functors $\iota^\alpha_{\boldsymbol{G}}\colon A^\alpha_\varepsilon \times \partial^\alpha \boldsymbol{G} \to \boldsymbol{G}$ are defined for some $\varepsilon > 0$ and required to be tubular neighborhood diffeomorphisms on both the sets of objects and morphisms. In particular, $\iota^\alpha_{\boldsymbol{G}}(\alpha, \cdot)$ is the identification between $\partial^\alpha \boldsymbol{G}$ and the full subcategories formed by the boundary components of $\boldsymbol{G}$.
- An *oriented cobordism groupoid* is a cobordism groupoid $(\boldsymbol{G}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}})$ such that both $\boldsymbol{G}$ and its boundary groupoids $\partial^0 \boldsymbol{G}, \partial^1 \boldsymbol{G}$ are oriented. Moreover the collaring functors are required to consist of orientation-preserving maps $\iota^\alpha_{\boldsymbol{G}}\colon A^\alpha_\varepsilon \times \mathrm{Obj}_{\partial^\alpha \boldsymbol{G}} \to \mathrm{Obj}_{\boldsymbol{G}}$ and $\iota^\alpha_{\boldsymbol{G}}\colon A^\alpha_\varepsilon \times \mathrm{Mor}_{\partial^\alpha \boldsymbol{G}} \to \mathrm{Mor}_{\boldsymbol{G}}$ for $\alpha = 0, 1$, where products are oriented as in Remark 3.1.11.

Lemma A.2 *Any topological space Y has a unique* ***maximal Hausdorff quotient*** *$Y_{\mathcal{H}}$, that is, a quotient of Y which is Hausdorff and satisfies the universal property: any continuous map from Y to a Hausdorff space factors through the quotient map $\pi_{\mathcal{H}}\colon Y \to Y_{\mathcal{H}}$.*

Proof To construct the maximal Hausdorff quotient let A be the set of all equivalence relations $\sim$ on Y for which the quotient topology on $Y/\sim$ is Hausdorff. This is a set since every relation $\sim$ on Y is represented by a subset of $Y \times Y$. Then the space $Y_A := \prod_{\sim \in A} Y/\sim$ is a product of Hausdorff spaces, hence Hausdorff. The map $\pi\colon Y \to Y_A$, $y \mapsto \prod_{\sim \in A} [y]_\sim$ is continuous by the definition of quotient topologies. Now the image $Y_{\mathcal{H}} := \pi(Y) \subset Y_A$ with the relative topology is Hausdorff, and π induces a continuous surjection $\pi_{\mathcal{H}}\colon Y \to Y_{\mathcal{H}}$.

To check that $\pi_{\mathcal{H}}\colon Y \to Y_{\mathcal{H}}$ satisfies the universal property, consider a continuous map $f\colon Y \to Z$ to a Hausdorff space Z. This induces an equivalence relation $\sim_f$ on Y given by $x \sim_f y \iff f(x) = f(y)$, whose quotient space $Y/\sim_f$ we equip with the quotient topology. Then $f\colon Y \to Z$ factors as

$$Y \xrightarrow{\pi_f} Y/\sim_f \xrightarrow{\iota_f} Z,$$

where $\iota_f\colon [y] \mapsto f(y)$ is continuous by definition of the quotient topology. Since ι_f is also injective, this implies that $Y/\sim_f$ is Hausdorff. Therefore, $Y/\sim_f$ is one of the

[11] A subcategory is strictly full if it contains all morphisms that have source or target in its objects.

factors of Y_A, so that $f\colon Y \to Z$ factors as the following sequence of continuous maps

$$Y \xrightarrow{\pi_{\mathcal{H}}} Y_{\mathcal{H}} \xrightarrow{\mathrm{pr}_f} Y/{\sim_f} \xrightarrow{\iota_f} Z,$$

where pr_f denotes the restriction to $Y_{\mathcal{H}}$ of the projection from Y_A to its factor $Y/{\sim_f}$.

To see that $Y_{\mathcal{H}}$ is in fact a quotient of Y, we will identify $Y_{\mathcal{H}} = \pi(Y)$ with the quotient $Y/\sim_\pi$ that is induced by the surjection $\pi_{\mathcal{H}}\colon Y \to Y_{\mathcal{H}}$. In this case the injection $\iota_\pi\colon Y/{\sim_\pi} \to Y_{\mathcal{H}}$ is in fact a continuous bijection by continuity and surjectivity of $\pi_{\mathcal{H}}$. In particular, this implies that $Y/{\sim_\pi}$ is Hausdorff, so that we have a continuous map $\mathrm{pr}_\pi\colon Y_{\mathcal{H}} \to Y/{\sim_\pi}$ by restriction of the projection $Y_A \to Y/{\sim_\pi}$ as above. It is inverse to ι_π because for $[y] \in Y/{\sim_\pi}$, we have

$$\mathrm{pr}_\pi(\iota_\pi([y])) = \mathrm{pr}_\pi(\pi_{\mathcal{H}}(y)) = \mathrm{pr}_\pi(\cdots \times [y] \times \cdots) = [y].$$

This identifies $Y_{\mathcal{H}} \cong Y/{\sim_\pi}$ as topological spaces and thus finishes the proof that a topological space $Y_{\mathcal{H}}$ with the above properties exists.

To prove uniqueness, consider another Hausdorff quotient $\mathrm{pr}\colon Y \to Y/{\sim}$ that satisfies the universal property. Then pr factors,

$$Y \xrightarrow{\pi_{\mathcal{H}}} Y_{\mathcal{H}} \xrightarrow{a} Y/{\sim},$$

and by the universal property $\pi_{\mathcal{H}}\colon Y \to Y_{\mathcal{H}}$ factors,

$$Y \xrightarrow{\mathrm{pr}} Y/{\sim} \xrightarrow{b} Y_{\mathcal{H}}.$$

Then a is surjective since pr is. Moreover, a is injective, because otherwise there would be two points $y_1, y_2 \in Y$ with $\pi_{\mathcal{H}}(y_1) \neq \pi_{\mathcal{H}}(y_2)$ but $\mathrm{pr}(y_1) = a(\pi(y_1)) = a(\pi(y_2)) = \mathrm{pr}(y_2)$, so that $\pi(y_1) = b(\mathrm{pr}(y_1)) = b(\mathrm{pr}(y_2)) = \pi(y_2)$, a contradiction. A similar argument shows that b is bijective. Moreover, the composite $b^{-1}a\colon Y_{\mathcal{H}} \to Y_{\mathcal{H}}$ has the property that $b^{-1}a \circ \pi_{\mathcal{H}} = \pi_{\mathcal{H}}$. Since $\pi_{\mathcal{H}}$ is surjective this implies $b^{-1}a = \mathrm{id}$, and similarly $a^{-1}b = \mathrm{id}$. Finally, note that because both $Y_{\mathcal{H}}$ and $Y/{\sim}$ have the quotient topology, a and b are continuous, and hence homeomorphisms. □

In the following we write $|\boldsymbol{G}|$ for the realization $\mathrm{Obj}_{\boldsymbol{G}}/{\sim}$ of an étale groupoid $\boldsymbol{G}$, and $|\boldsymbol{G}|_{\mathcal{H}}$ for its maximal Hausdorff quotient. We denote the natural maps by

$$\pi_{\boldsymbol{G}}\colon \mathrm{Obj}_{\boldsymbol{G}} \to |\boldsymbol{G}|, \quad \pi^{\mathcal{H}}_{|\boldsymbol{G}|}\colon |\boldsymbol{G}| \longrightarrow |\boldsymbol{G}|_{\mathcal{H}}, \quad \pi^{\mathcal{H}}_{\boldsymbol{G}} := \pi^{\mathcal{H}}_{|\boldsymbol{G}|} \circ \pi_{\boldsymbol{G}}\colon \mathrm{Obj}_{\boldsymbol{G}} \to |\boldsymbol{G}|_{\mathcal{H}}.$$

Moreover, for $U \subset \mathrm{Obj}_{\boldsymbol{G}}$ we write $|U| := \pi_{\boldsymbol{G}}(U) \subset |\boldsymbol{G}|$ and $|U|_{\mathcal{H}} := \pi_{\mathcal{H}}(U) \subset |\boldsymbol{G}|_{\mathcal{H}}$.

Lemma A.3 *Let $\boldsymbol{G}$ be an étale groupoid.*

(i) *Any smooth functor $F\colon \boldsymbol{G}\to\boldsymbol{G}'$ induces a continuous map $|F|_{\mathcal{H}}\colon |\boldsymbol{G}|_{\mathcal{H}}\to|\boldsymbol{G}'|_{\mathcal{H}}$.*

(ii) *If $A\subset\mathbb{R}$ is any interval, we may identify $|A\times\boldsymbol{G}|$ with $A\times|\boldsymbol{G}|$, and $|A\times\boldsymbol{G}|_{\mathcal{H}}$ with $A\times|\boldsymbol{G}|_{\mathcal{H}}$. More precisely, there are commutative diagrams*

$$\begin{tikzcd}
\mathrm{Obj}_{A\times\boldsymbol{G}} \arrow[r, "\mathrm{pr}_A\times\mathrm{pr}_{\boldsymbol{G}}"] \arrow[d, "\pi_{A\times\boldsymbol{G}}"'] & A\times\mathrm{Obj}_{\boldsymbol{G}} \arrow[d, "\mathrm{id}_A\times\pi_{\boldsymbol{G}}"] \\
|A\times\boldsymbol{G}| \arrow[r, "|\mathrm{pr}_A|\times|\mathrm{pr}_{\boldsymbol{G}}|"] & A\times|\boldsymbol{G}|
\end{tikzcd}
\qquad
\begin{tikzcd}
|A\times\boldsymbol{G}| \arrow[r, "|\mathrm{pr}_A|\times|\mathrm{pr}_{\boldsymbol{G}}|"] \arrow[d, "\pi^{\mathcal{H}}_{|A\times\boldsymbol{G}|}"'] & A\times|\boldsymbol{G}| \arrow[d, "\mathrm{id}_A\times\pi^{\mathcal{H}}_{|\boldsymbol{G}|}"] \\
|A\times\boldsymbol{G}|_{\mathcal{H}} \arrow[r, "|\mathrm{pr}_A|_{\mathcal{H}}\times|\mathrm{pr}_{\boldsymbol{G}}|_{\mathcal{H}}"] & A\times|\boldsymbol{G}|_{\mathcal{H}}
\end{tikzcd}$$

where the horizontal maps are homeomorphisms. Here $\mathrm{pr}_A\colon A\times\boldsymbol{G}\to\boldsymbol{A}$ and $\mathrm{pr}_{\boldsymbol{G}}\colon A\times\boldsymbol{G}\to\boldsymbol{G}$ are the two projection functors from the product groupoid to its factors and $\boldsymbol{A}$ is the groupoid with objects A and only identity morphisms so that $A=|\boldsymbol{A}|=|\boldsymbol{A}|_{\mathcal{H}}$.

Proof Any smooth functor $F\colon\boldsymbol{G}\to\boldsymbol{G}'$ induces a continuous map $|\boldsymbol{G}|\xrightarrow{|F|}|\boldsymbol{G}'|$. Then by Lemma A.2 applied to $|\boldsymbol{G}|$, the composite

$$|\boldsymbol{G}|\xrightarrow{|F|}|\boldsymbol{G}'|\xrightarrow{\pi^{\mathcal{H}}_{|\boldsymbol{G}'|}}|\boldsymbol{G}'|_{\mathcal{H}}$$

factors uniquely through the quotient map $|\boldsymbol{G}|\xrightarrow{\pi^{\mathcal{H}}_{|\boldsymbol{G}|}}|\boldsymbol{G}|_{\mathcal{H}}$. The resulting continuous map $|F|_{\mathcal{H}}\colon|\boldsymbol{G}|_{\mathcal{H}}\to|\boldsymbol{G}'|_{\mathcal{H}}$ is uniquely determined by $\pi^{\mathcal{H}}_{|\boldsymbol{G}'|}\circ|F|=|F|_{\mathcal{H}}\circ\pi^{\mathcal{H}}_{|\boldsymbol{G}|}$. This proves (i).

To prove (ii), first consider the diagram on the left. The bottom horizontal map is bijective because $\mathrm{Mor}_{A\times\boldsymbol{G}}=A\times\mathrm{Mor}_{\boldsymbol{G}}$, and continuous by definition of the product topology. Finally, it is a homeomorphism because A is locally compact; cf [15, Exercise 29.11]. In the diagram on the right we define the bottom horizontal arrow using the product of the maps induced as in (i) by the two functors pr_A and $\mathrm{pr}_{\boldsymbol{G}}$. Hence it is continuous. Since the diagram commutes and we have already seen that the top horizontal map is a homeomorphism, it remains to check this for the bottom map. But this holds because the uniqueness property of the maximal Hausdorff quotient implies that for any homeomorphism $\phi\colon Y\to Y'$, the unique continuous map $\phi_{\mathcal{H}}\colon Y_{\mathcal{H}}\to Y'_{\mathcal{H}}$ such that $Y\xrightarrow{\phi}Y'\xrightarrow{\pi_{Y'}}Y'_{\mathcal{H}}$ equals $Y\xrightarrow{\pi_{\mathcal{H}}}Y_{\mathcal{H}}\xrightarrow{\phi_{\mathcal{H}}}Y'_{\mathcal{H}}$ must be a homeomorphism. □

The smooth structure on a weighted branched manifold will be given by a homeomorphism to the realization of an étale groupoid with the following weighting structure.

Definition A.4 [8, Definition 3.2] A *weighted nonsingular branched groupoid* (or *wnb groupoid* for short) of dimension d is a pair $(\boldsymbol{G},\Lambda)$ consisting of an oriented, nonsingular, étale groupoid $\boldsymbol{G}$ of dimension d, together with a rational weighting function $\Lambda\colon |\boldsymbol{G}|_{\mathcal{H}} \to \mathbb{Q}^+ := \mathbb{Q} \cap (0,\infty)$ that satisfies the following compatibility conditions. For each $p \in |\boldsymbol{G}|_{\mathcal{H}}$ there is an open neighborhood $N \subset |\boldsymbol{G}|_{\mathcal{H}}$ of p, a collection $U_1,\dots,U_\ell$ of disjoint open subsets of $(\pi^{\mathcal{H}}_{\boldsymbol{G}})^{-1}(N) \subset \mathrm{Obj}_{\boldsymbol{G}}$ (called *local branches*), and a set of positive rational weights $m_1,\dots,m_\ell$ such that the following properties hold:

Covering $\quad (\pi^{\mathcal{H}}_{|\boldsymbol{G}|})^{-1}(N) = |U_1| \cup \cdots \cup |U_\ell| \subset |\boldsymbol{G}|$.

Local regularity For each $i = 1,\dots,\ell$ the projection $\pi^{\mathcal{H}}_{\boldsymbol{G}}|_{U_i}\colon U_i \to |\boldsymbol{G}|_{\mathcal{H}}$ is a homeomorphism onto a relatively closed subset of N.

Weighting For all $q \in N$, the number $\Lambda(q)$ is the sum of the weights of the local branches whose image contains q:

$$\Lambda(q) = \sum_{i\colon q \in |U_i|_{\mathcal{H}}} m_i.$$

A *wnb cobordism groupoid* is a tuple $(\boldsymbol{G}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}}, \Lambda)$ in which $(\boldsymbol{G}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}})$ is an oriented, nonsingular, étale cobordism groupoid of dimension d, and $\Lambda\colon |\boldsymbol{G}|_{\mathcal{H}} \to \mathbb{Q}^+$ is a weighting function as above with the additional property that Λ and the local branches $U_1,\dots,U_\ell$ are of product form in the collars.

In particular, this means that each boundary groupoid $\partial^\alpha \boldsymbol{G}$ is equipped with a weighting function Λ^α as above such that the following diagram commutes:

$$\begin{tikzcd}
A^\alpha_\varepsilon \times |\partial^\alpha \boldsymbol{G}|_{\mathcal{H}} \arrow[r, "|\iota^\alpha_{\boldsymbol{G}}|_{\mathcal{H}}"] \arrow[d, "\mathrm{id}_{A^\alpha_\varepsilon} \times \Lambda^\alpha"'] & |\boldsymbol{G}|_{\mathcal{H}} \arrow[d, "\Lambda"] \\
\mathbb{Q}^+ \arrow[r, "\mathrm{id}"] & \mathbb{Q}^+
\end{tikzcd}$$

where $|\iota^\alpha_{\boldsymbol{G}}|_{\mathcal{H}}$ is induced by the collaring functor $\iota^\alpha_{\boldsymbol{G}}\colon A^\alpha_\varepsilon \times \partial^\alpha \boldsymbol{G} \to \boldsymbol{G}$ and we identify $|A^\alpha_\varepsilon \times \partial^\alpha \boldsymbol{G}|_{\mathcal{H}}$ with $A^\alpha_\varepsilon \times |\partial^\alpha \boldsymbol{G}|_{\mathcal{H}}$ as in Lemma A.3 with orientation as specified in Definition 3.1.10.

Now we can formulate the notions of weighted branched manifold and cobordism.

Definition A.5 A *weighted branched manifold/cobordism* of dimension d is a pair (Z, Λ_Z) consisting of a topological space Z together with a function $\Lambda_Z\colon Z \to \mathbb{Q}^+$ and an equivalence class[12] of wnb (cobordism) d–dimensional groupoids $(\boldsymbol{G}, \Lambda_{\boldsymbol{G}})$ and homeomorphisms $f\colon |\boldsymbol{G}|_{\mathcal{H}} \to Z$ that induce the function $\Lambda_Z = \Lambda_{\boldsymbol{G}} \circ f^{-1}$.

[12] The precise notion of equivalence is given in [8, Definition 3.12]. In particular it ensures that the induced function $\Lambda_Z := \Lambda_{\boldsymbol{G}} \circ f^{-1}$ and the dimension of $\mathrm{Obj}_{\boldsymbol{G}}$ is the same for equivalent structures $(\boldsymbol{G}, \Lambda_{\boldsymbol{G}}, f)$. Moreover, if $(\boldsymbol{G}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}})$ is a cobordism groupoid, then the images $f(|\partial^\alpha \boldsymbol{G}|_{\mathcal{H}}) := \partial^\alpha Z \subset Z$ of the two boundary components are well defined.

For a weighted branched cobordism $(Z, \Lambda_Z, [\boldsymbol{G}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}}, \Lambda_{\boldsymbol{G}}, f])$, the induced *boundary components* $\partial^\alpha Z := f\big(|\iota^\alpha_{\boldsymbol{G}}|_{\mathcal{H}}(|\partial^\alpha \boldsymbol{G}|_{\mathcal{H}})\big) \subset Z$ for $\alpha = 0, 1$ are equipped with the weighted branched manifold structures $[(\partial^\alpha \boldsymbol{G}, \Lambda^\alpha_{\boldsymbol{G}}), f|_{|\partial^\alpha \boldsymbol{G}|_{\mathcal{H}}}]$.

The underlying space Z of a weighted branched manifold or cobordism is always Hausdorff due to the homeomorphism $Z \cong |\boldsymbol{G}|_{\mathcal{H}}$ to a Hausdorff quotient. Moreover, since cobordism groupoids are compact by definition, the underlying space Z of a weighted branched cobordism is always compact.

It is shown in [8, Proposition 3.5] that the weighting function $\Lambda\colon |\boldsymbol{G}|_{\mathcal{H}} \to (0, \infty)$ is locally constant on the complement of the *branch locus* $\mathrm{Br}(\boldsymbol{G}) \subset |\boldsymbol{G}|_{\mathcal{H}}$. (This is defined to be the set of points in $|\boldsymbol{G}|_{\mathcal{H}}$ over which $|\pi|^{\mathcal{H}}_{|\boldsymbol{G}|}\colon |\boldsymbol{G}| \to |\boldsymbol{G}|_{\mathcal{H}}$ is not injective, and is closed and nowhere dense.) Further, every point in $|\boldsymbol{G}|_{\mathcal{H}} \setminus \mathrm{Br}(\boldsymbol{G})$ has a neighborhood that is homeomorphic via $\pi^{\mathcal{H}}_{|\boldsymbol{G}|}$ to an open subset in a local branch and so has the structure of a smooth oriented manifold.

Example A.6 (i) Any compact oriented smooth manifold/cobordism may be considered as a weighted branched manifold/cobordism with weighting function $\Lambda_Z \equiv 1$ and empty branch locus.

(ii) A compact weighted branched manifold of dimension 0 also necessarily has empty branch locus and consists of a finite set of points $\{p_1, \dots, p_k\}$, each with a positive rational weight $m(p_i) \in \mathbb{Q}^+$ and orientation $\mathfrak{o}(p_i) \in \{\pm\}$. Any representing groupoid $\boldsymbol{G}$ has as object space $\mathrm{Obj}_{\boldsymbol{G}}$ a set with the discrete topology, which is equipped with an orientation function $\mathfrak{o}\colon \mathrm{Obj}_{\boldsymbol{G}} \to \{\pm\}$. The morphism space $\mathrm{Mor}_{\boldsymbol{G}}$ is also a discrete set and, because we assume that $\boldsymbol{G}$ is oriented, defines an equivalence relation on $\mathrm{Obj}_{\boldsymbol{G}}$ such that $x \sim y \Longrightarrow \mathfrak{o}(x) = \mathfrak{o}(y)$. Moreover, because $|\boldsymbol{G}|$ is Hausdorff, we can identify $|\boldsymbol{G}| = |\boldsymbol{G}|_{\mathcal{H}}$ and hence conclude that $\mathrm{Obj}_{\boldsymbol{G}}$ consists of precisely k classes of points that are equivalent under $\mathrm{Mor}_{\boldsymbol{G}}$ and project to $p_1, \dots, p_k$ in $Z \cong |\boldsymbol{G}|_{\mathcal{H}}$.

(iii) For the prototypical example of a 1–dimensional weighted branched cobordism $(|\boldsymbol{G}|_{\mathcal{H}}, \Lambda)$, take $\mathrm{Obj}(\boldsymbol{G}) = I \sqcup I'$ equal to two copies of the interval $I = I' = [0, 1]$, with nonidentity morphisms from $x \in I$ to $x \in I'$ for $x \in \big[0, \tfrac{1}{2}\big)$ and their inverses, where we suppose that I is oriented in the standard way. Then the realization and its Hausdorff quotient are

$$
\begin{aligned}
|\boldsymbol{G}| &= I \sqcup I' / \big\{(I, x) \sim (I', x) \text{ if and only if } x \in \big[0, \tfrac{1}{2}\big)\big\}, \\
|\boldsymbol{G}|_{\mathcal{H}} &= I \sqcup I' / \big\{(I, x) \sim (I', x) \text{ if and only if } x \in \big[0, \tfrac{1}{2}\big]\big\},
\end{aligned}
$$

and the branch locus is a single point $\mathrm{Br}(\boldsymbol{G}) = \big\{\big[I, \tfrac{1}{2}\big] = \big[I', \tfrac{1}{2}\big]\big\} \subset |\boldsymbol{G}|_{\mathcal{H}}$. The choice of weights $m, m' > 0$ on the two local branches I and I' determines the weighting function $\Lambda\colon |\boldsymbol{G}|_{\mathcal{H}} \to (0,\infty)$ as

$$\Lambda([I,x]) = \begin{cases} m+m' & \text{if } x \in \left[0,\tfrac{1}{2}\right], \\ m & \text{if } \ x \in \left(\tfrac{1}{2},1\right], \end{cases}$$

$$\Lambda([I',x]) = \begin{cases} m+m' & \text{if } x \in \left[0,\tfrac{1}{2}\right], \\ m' & \text{if } x \in \left(\tfrac{1}{2},1\right]. \end{cases}$$

For example, giving each branch I, I' the weight $m = m' = \frac{1}{2}$, together with an appropriate choice of collar functors $\iota^\alpha_{\boldsymbol{G}}$, yields a weighted branched cobordism $(|\boldsymbol{G}|_{\mathcal{H}}, \iota^0_{\boldsymbol{G}}, \iota^1_{\boldsymbol{G}}, \Lambda)$ with $|\partial^0 \boldsymbol{G}|_{\mathcal{H}} = \{[I,0] = [I',0]\}$, which is a single point with weight 1, and $|\partial^1 \boldsymbol{G}|_{\mathcal{H}} = \{[I,1],[I',1]\}$, which consists of two points with weight $\frac{1}{2}$, all with positive orientation because as explained in Remark 3.1.11, the induced orientation on the boundary $\partial^\alpha \boldsymbol{G}$ of a cobordism is completed to an orientation of the collar by adding as the first component the positive unit vector along A^α_ε.

Another choice of collar functors for the same weighted groupoid $(\boldsymbol{G},\Lambda)$ might give rise to a different partition of the boundary into incoming $\partial^0 \boldsymbol{G}$ and outgoing $\partial^1 \boldsymbol{G}$, for example yielding a weighted branched cobordism with $|\partial^0 \boldsymbol{G}|_{\mathcal{H}} = \{[I,0]{=}[I',0],[I,1]\}$ consisting of two points with weights and orientations $(1,+)$ and $\left(\frac{1}{2},-\right)$, and with $|\partial^1 \boldsymbol{G}|_{\mathcal{H}} = \{[I',1]\}$ consisting of one point with weight $\left(\frac{1}{2},+\right)$.

(iv) In the situation of Theorem 3.2.8, the nonsingular étale groupoid $\widehat{\boldsymbol{Z}}^\nu$ with $\mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu} = (\mathfrak{s}_{\mathcal{K}}|_{\mathcal{V}} + \nu)^{-1}(0)$ has a maximal Hausdorff quotient $|\widehat{\boldsymbol{Z}}^\nu|_{\mathcal{H}} = |\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}|$ that, as we show in Lemma 3.2.10, is given by the realization of the groupoid $\widehat{\boldsymbol{Z}}^\nu_{\mathcal{H}}$ obtained as in (iii) above by closing the set of morphisms $\mathrm{Mor}_{\widehat{\boldsymbol{Z}}^\nu} \subset \mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu} \times \mathrm{Obj}_{\widehat{\boldsymbol{Z}}^\nu}$. Therefore, in this case we can give a completely explicit description of $|\boldsymbol{Z}|_{\mathcal{H}}$ and its weighting function $\Lambda_{\boldsymbol{Z}}$; see the proof of Theorem 3.2.8. $\diamond$

The following is a version of some parts of [8, Proposition 3.25], which more generally defines a notion of integration over weighted branched manifolds and cobordisms.

Proposition A.7 *Any compact d–dimensional weighted branched manifold (Y,Λ_Y) induces a **fundamental class** $[Y] \in H_d(Y;\mathbb{Q})$, and any d–dimensional weighted branched cobordism (Z,Λ_Z) with boundary $\partial Z := \partial^0 Z \cup \partial^1 Z$ induces a **fundamental class** $[Z] \in H_d(Z,\partial Z;\mathbb{Q})$, whose image under the boundary map*

$$\partial\colon H_d(Z,\partial Z;\mathbb{Q}) \to H_{d-1}(\partial Z;\mathbb{Q}) \cong H_{d-1}(\partial^0 Z;\mathbb{Q}) + H_{d-1}(\partial^1 Z;\mathbb{Q})$$

is $\partial[Z] = [\partial^1 Z] - [\partial^0 Z]$.

Proof If (Y,Λ_Y) has a weighted branched manifold structure $(\boldsymbol{G},\Lambda_{\boldsymbol{G}})$ with well-behaved (eg piecewise smooth) branch locus, then one can triangulate $|\boldsymbol{G}|_{\mathcal{H}} \cong Y$

so that the branch locus lies in the codimension 1 skeleton. We may then define a singular cycle on Y by using the local weights m_i to assign a rational weight to each top-dimensional simplex. As explained in Remark 3.1.11, in the case of a cobordism Z the induced orientation on the boundary component $\partial^\alpha Z$ is completed to the orientation of the collar by adding the unit positive vector along the collar as the first component. In the case of $\partial^0 Z$ this yields an orientation of $\partial^0 Z$ that is the opposite of the standard way of orienting a boundary component by adding the outward pointing normal, a fact that is reflected in the minus sign in the formula $\partial[Z] = [\partial^1 Z] - [\partial^0 Z]$. For more details and the general case, see [8]. □

References

[1] **N Bourbaki**, *General topology: Chapters 1–4*, Springer (1989) MR

[2] **R Castellano**, *Genus zero Gromov–Witten axioms via Kuranishi atlases*, preprint (2016) arXiv

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

[4] **K Fukaya**, **Y-G Oh**, **H Ohta**, **K Ono**, *Technical details on Kuranishi structure and virtual fundamental chain*, preprint (2012) arXiv

[5] **K Fukaya**, **Y-G Oh**, **H Ohta**, **K Ono**, *Kuranishi structure, pseudoholomorphic curve, and virtual fundamental chain, I*, preprint (2015) arXiv

[6] **K Fukaya**, **K Ono**, *Arnold conjecture and Gromov–Witten invariant*, Topology 38 (1999) 933–1048 MR

[7] **H Hofer**, **K Wysocki**, **E Zehnder**, *A general Fredholm theory, II: Implicit function theorems*, Geom. Funct. Anal. 19 (2009) 206–293 MR

[8] **D McDuff**, *Groupoids, branched manifolds and multisections*, J. Symplectic Geom. 4 (2006) 259–315 MR

[9] **D McDuff**, *Smooth Kuranishi atlases*, lecture slides (2013) Available at `https://math.barnard.edu/profiles/dusa-mcduff`

[10] **D McDuff**, *Notes on Kuranishi atlases*, preprint (2014) arXiv

[11] **D McDuff**, *Strict orbifold atlases and weighted branched manifolds*, preprint (2015) arXiv To appear in J. Symplectic Geom.

[12] **D McDuff**, **K Wehrheim**, *Smooth Kuranishi atlases with trivial isotropy*, preprint (2012) arXiv

[13] **D McDuff**, **K Wehrheim**, *The topology of Kuranishi atlases*, preprint (2015) arXiv To appear in Proc. London Math. Soc.

[14] **D McDuff**, **K Wehrheim**, *The fundamental class of smooth Kuranishi atlases with trivial isotropy*, J. Topol. Anal. (online publication November 2016)

[15] **J R Munkres**, *Topology*, 2nd edition, Prentice Hall, Upper Saddle River, NJ (2000)

[16] **J Pardon**, *An algebraic approach to virtual fundamental cycles on moduli spaces of pseudoholomorphic curves*, Geom. Topol. 20 (2016) 779–1034 MR

[17] **A Weinstein**, *The volume of a differentiable stack*, Lett. Math. Phys. 90 (2009) 353–371 MR

Mathematics Department, Barnard College, Columbia University
MC4410, 2990 Broadway, New York, NY 10027-6840, United States

Department of Mathematics, University of California, Berkeley
907 Evans Hall #3840, Berkeley, CA 94705-3840, United States

dusa@math.columbia.edu, katrin@math.berkeley.edu

Proposed: Leonid Polterovich Received: 28 August 2015
Seconded: Yasha Eliashberg, Peter Teichner Revised: 8 September 2016

Geometry & Topology 21 (2017) 2811–2850

Brown's moduli spaces of curves and the gravity operad

CLÉMENT DUPONT
BRUNO VALLETTE

This paper is built on the following observation: the purity of the mixed Hodge structure on the cohomology of Brown's moduli spaces is essentially equivalent to the freeness of the dihedral operad underlying the gravity operad. We prove these two facts by relying on both the geometric and the algebraic aspects of the problem: the complete geometric description of the cohomology of Brown's moduli spaces and the coradical filtration of cofree cooperads. This gives a conceptual proof of an identity of Bergström and Brown which expresses the Betti numbers of Brown's moduli spaces via the inversion of a generating series. This also generalizes the Salvatore–Tauraso theorem on the nonsymmetric Lie operad.

14H10; 14C30, 18D50

Introduction

The moduli space of genus zero smooth curves with n marked points, denoted by $\mathcal{M}_{0,n}$, is a classical object in algebraic geometry, as well as its Deligne–Mumford–Knudsen compactification $\overline{\mathcal{M}}_{0,n}$, which parametrizes stable genus zero curves with n marked points. In [5], Brown introduced a "partial compactification"

$$\mathcal{M}_{0,n} \subset \mathcal{M}_{0,n}^{\delta} \subset \overline{\mathcal{M}}_{0,n}$$

in order to prove a conjecture of Goncharov and Manin [17] on the relation between certain period integrals on $\overline{\mathcal{M}}_{0,n}$ and multiple zeta values.

The homology groups of the moduli spaces $\mathcal{M}_{0,n}$, as well as those of the compactified moduli spaces $\overline{\mathcal{M}}_{0,n}$, assemble to form two operads, respectively called the gravity and hypercommutative operads by Getzler. These two operads are Koszul dual in the sense of the Koszul duality of operads; see Getzler [14] and Ginzburg and Kapranov [16]. As pointed out by Getzler, this is very much related to the purity of the mixed Hodge structures on the cohomology groups under consideration. This implies that the exponential generating series encoding the Betti numbers of $\mathcal{M}_{0,n}$ and $\overline{\mathcal{M}}_{0,n}$ are inverse to one another.

A similar identity was conjectured by Bergström and Brown in [4]: the ordinary generating series encoding the Betti numbers of the moduli spaces $\mathcal{M}_{0,n}$ and $\mathcal{M}_{0,n}^{\delta}$

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2811

should be inverse to one another. More precisely, it is showed how such a relation can be derived from a more conceptual fact: the purity of the mixed Hodge structure on the cohomology groups of Brown's moduli spaces. This is the first result of the present paper.

Theorem A *For every integers k and n, the mixed Hodge structure on the cohomology group $H^k(\mathcal{M}^\delta_{0,n})$ is pure Tate of weight $2k$.*

This theorem has the following straightforward consequences:

- the cohomology algebra of Brown's moduli space $\mathcal{M}^\delta_{0,n}$ embeds into that of the moduli space $\mathcal{M}_{0,n}$ (Corollary 4.18);
- there is a recursive formula for the Betti numbers of $\mathcal{M}^\delta_{0,n}$, conjectured in Bergström and Brown [4] (Corollary 4.19);
- Brown's moduli spaces $\mathcal{M}^\delta_{0,n}$ are formal topological spaces in the sense of rational homotopy theory (Corollary 4.20).

It turns out that the purity of the mixed Hodge structure of Theorem A can be equivalently interpreted in the following operadic terms.

Theorem B *The dihedral gravity operad is free. Its space of generators in arity n and degree k is (noncanonically) isomorphic to the homology group $H_{k+n-3}(\mathcal{M}^\delta_{0,n})$.*

We introduce here the new notion of a dihedral operad, which faithfully takes into account the dihedral symmetry of Brown's moduli spaces. Such a notion forgets almost all the symmetry properties of a cyclic operad, except for the dihedral structure. Theorem B can also be viewed as a kind of nonsymmetric analog of the Koszul duality between the gravity and the hypercommutative operad, since a free operad is Koszul, its dual being a nilpotent operad. We prove it by introducing a combinatorial filtration on the cohomology groups of the spaces $\mathcal{M}_{0,n}$, and identifying it with the coradical filtration of the dihedral gravity cooperad.

The problem of studying whether the nonsymmetric operad underlying a given operad is free is not new. In [26], Salvatore and Tauraso proved that the nonsymmetric operad underlying the operad of Lie algebras is free. This result is actually the top dimensional part of Theorem B. Thus, the geometric methods developed throughout this paper provide us with a new proof of (a dihedral enhancement of) the theorem of Salvatore and Tauraso.

Note that in the preprint [2] (which appeared on the arXiv one day after the present article), Alm and Petersen give independent proofs of Theorem A and Theorem B. Their proofs rely on an explicit basis for the gravity cooperad, and a construction of Brown's moduli spaces in terms of blow-ups and deletions. The freeness of the (nonsymmetric) gravity operad has been used by Alm in [1] to study an exotic A_∞–structure on Batalin–Vilkovisky algebras.

Layout The first section deals with the various combinatorial objects and notions of operads used in this text. In the second section, we introduce the moduli spaces of curves $\mathcal{M}_{0,n}$ and $\overline{\mathcal{M}}_{0,n}$, as well as the notion of mixed Hodge structure. The study of Brown's moduli spaces $\mathcal{M}_{0,n}^{\delta}$ and the dihedral gravity cooperad fills the third section. The fourth section contains the proofs of Theorems A and B and their corollaries.

Conventions Throughout the paper, the field of coefficients is the field $\mathbb{Q}$ of rational numbers. For a topological space X, we simply denote by $H_\bullet(X)$ and $H^\bullet(X)$ the (co)homology groups of X with rational coefficients. We work with graded vector spaces and switch between the homological convention (with degrees as subscripts) and the cohomological convention (with degrees as superscripts), the two conventions being linear dual to one another.

Acknowledgements We would like to express our sincere appreciation to Johan Alm, Francis Brown and Dan Petersen for useful discussions, and to the anonymous referee for their suggestions, which helped improve the clarity of this article. We would like to thank the Max-Planck-Institut für Mathematik (where Dupont was holding a position and where Vallette came during several visits) and the University Nice Sophia Antipolis (vice versa) for the excellent working conditions. Vallette was supported by the ANR SAT grant.

1 Freeness criteria for dihedral cooperads

The purpose of this first section is to recall the various notions of operads (classical, cyclic, nonsymmetric, cyclic nonsymmetric) and to introduce a new one (dihedral operad) which suits the geometry of Brown's moduli spaces. We first describe the combinatorial objects (trees and polygon dissections) involved in the proof of the results of the paper. In the end of this section, we prove two freeness criteria for dihedral cooperads: one based on their cobar construction and the other based on their coradical filtration.

1.1 Dissections of polygons and trees

Definition 1.1 (structured sets) Let S be a finite set of cardinality n.

- A *basepoint* ρ on S is a map $\rho\colon \{*\} \to S$. A pair (S, ρ) is called a *pointed set*.
- A *total order* ω on S is a bijection between S and the set $\{1, \dots, n\}$. There are $n!$ total orders on S. A pair (S, ω) is called a *totally ordered set*. By

convention, we view a totally ordered set as a pointed set, the basepoint being the maximal element.

- A *cyclic structure* γ on S is an identification of S with the edges of an oriented n–gon, modulo rotations. There are $n!/n = (n-1)!$ cyclic structures on S. A pair (S,γ) is called a *cyclic set*.
- A *dihedral structure* δ on S is an identification of S with the edges of an unoriented n–gon, modulo dihedral symmetries. There are $n!/(2n) = \frac{1}{2}(n-1)!$ dihedral structures on S. A pair (S,δ) is called a *dihedral set*.

In the sequel, we will identify a dihedral set (S,δ) with an unoriented polygon with its edges decorated by S in the dihedral order prescribed by δ.

Definition 1.2 (chords and dissections) Let (S,δ) be a dihedral set.

- A *chord* of (S,δ) is an unordered pair of nonconsecutive vertices of the underlying unoriented polygon.
- A *dissection* $\mathfrak{d}$ of (S,δ) is a (possibly empty) set of noncrossing chords. The refinement of dissections endows them with a poset structure:

$$\mathfrak{d} \leqslant \mathfrak{d}' \quad \text{if } \mathfrak{d} \subset \mathfrak{d}',$$

  in which the smallest element is the empty dissection. We denote by $\mathsf{Diss}(S,\delta)$ the poset of dissections of (S,δ), and by $\mathsf{Diss}_k(S,\delta)$ the subset consisting of dissections with k chords.

For a dissection $\mathfrak{d} \in \mathsf{Diss}(S,\delta)$, we denote by $P(\mathfrak{d})$ the set of subpolygons that it defines; see Figure 1. If $\mathfrak{d}$ is in $\mathsf{Diss}_k(S,\delta)$, then $P(\mathfrak{d})$ has cardinality $k+1$. A subpolygon $p \in P(\mathfrak{d})$ corresponds to a dihedral set that we denote by $(E(p),\delta(p))$, where $E(p)$ consists of edges and chords of the polygon (S,δ).

Definition 1.3 (trees) A *tree* is a finite graph with no cycle. The contraction of internal edges endows trees with a poset structure: we set $\mathfrak{t} \leqslant \mathfrak{t}'$ if the tree $\mathfrak{t}$ can be obtained from the tree $\mathfrak{t}'$ by contracting some internal edges. If the number of external vertices is fixed, the minimal element of this poset is the only tree with zero internal edge, called a *corolla*. By looking at the possible structures on the set of external vertices of a tree, we get different posets:

- the poset $\mathsf{Tree}(S)$ of trees with external vertices labeled by S;
- the poset $\mathsf{RTree}(S,\rho)$ of rooted trees with external vertices labeled by S, the root being labeled by the basepoint ρ;
- the poset $\mathsf{PRTree}(S,\omega)$ of planar rooted trees with external vertices labeled by S in the total order ω, the root being labeled by the maximal element;

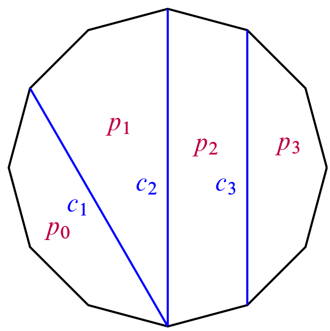


Figure 1: A dissection $\mathfrak{d} = \{c_1, c_2, c_3\}$, with the set of subpolygons $P(\mathfrak{d}) = \{p_0, p_1, p_2, p_3\}$

- the poset $\mathsf{PTree}(S, \gamma)$ of planar trees with external vertices labeled by S in the cyclic order γ;
- the poset $\mathsf{DTree}(S, \delta)$ of dihedral trees (trees embedded in an unoriented plane) with external vertices labeled by S in the dihedral order δ.

All these posets are graded by the number of internal edges of the trees.

For a tree $\mathfrak{t}$, we denote its set of vertices by $V(\mathfrak{t})$. For each vertex $v \in V(\mathfrak{t})$, we denote its set of adjacent edges by $E(v)$. Notice that if $\mathfrak{t}$ is a rooted tree then we get a pointed set $(E(v), \rho(v))$; if $\mathfrak{t}$ is a planar rooted tree then we get a totally ordered set $(E(v), \omega(v))$; if $\mathfrak{t}$ is a planar tree then we get a cyclic set $(E(v), \gamma(v))$; if $\mathfrak{t}$ is a dihedral tree then we get a dihedral set $(E(v), \delta(v))$. We refer the reader to [22, Section C.4] for more details on the notions related to trees.

Lemma 1.4 *The graded poset* $\mathsf{Diss}(S, \delta)$ *of dissections of a polygon* (S, δ) *and the graded poset* $\mathsf{DTree}(S, \delta)$ *of dihedral trees labeled by the dihedral set* (S, δ) *are isomorphic.*

Proof Let us describe the isomorphism $\mathsf{Diss}(S, \delta) \to \mathsf{DTree}(S, \delta)$. Given a dissection $\mathfrak{d} \in \mathsf{Diss}(S, \delta)$, one considers its "dual graph" $\mathfrak{t}$: each subpolygon $p \in P(\mathfrak{d})$ gives rise to a vertex $v \in V(\mathfrak{t})$ of the tree $\mathfrak{t}$ and each edge of this polygon gives rise to an edge of the tree; see Figure 2. The tree $\mathfrak{t}$ is naturally a dihedral tree, and it is straightforward to check that this defines a bijection between $\mathsf{Diss}(S, \delta)$ and $\mathsf{DTree}(S, \delta)$. Under this bijection, removing a chord from the dissection corresponds to contracting internal edges of trees; hence we get an isomorphism of posets, which respects the grading by construction. □

1.2 Dihedral operads

In this section, we recall the classical notions of operads, and we introduce a new one, the notion of dihedral operad, which suits the geometrical problem studied here. We

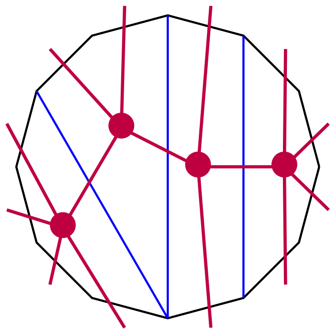


Figure 2: The isomorphism between polygon dissections and dihedral trees

work in the general setting of an abelian symmetric monoidal category $(\mathsf{A}, \otimes)$ such that the monoidal product preserves coproducts. In the next section and later on, we will specify the category A to be the category of graded mixed Hodge structures.

Definition 1.5 (categories of structured sets) We consider the following categories of structured sets:

- The category Bij of finite sets S and bijections.
- The category Bij_* of pointed sets (S, ρ) and bijections respecting the basepoint.
- The category Ord_* of totally ordered sets (S, ω) and bijections respecting the total order.
- The category Cyc of cyclic sets (S, γ) and bijections respecting the cyclic order.
- The category Dih of dihedral sets (S, δ) and bijections respecting the dihedral structure.

The forgetful functors between the various categories of structured sets assemble as a commutative diagram

$$\begin{tikzcd}
\mathsf{Dih} \arrow[d] & \mathsf{Cyc} \arrow[l] & \mathsf{Ord}_* \arrow[l] \arrow[d] \\
\mathsf{Bij} & & \mathsf{Bij}_* \arrow[ll]
\end{tikzcd}$$

where the functor $\mathsf{Ord}_* \to \mathsf{Bij}_*$ picks the maximal element as basepoint.

In each case, we consider the category of functors from these categories to the category A, for instance $\mathcal{M}\colon \mathsf{Bij} \to \mathsf{A}$, that we respectively call the category of Bij–modules, Bij_*–modules, Ord_*–modules, Cyc–modules and Dih–modules. We denote them respectively by $\mathsf{Bij}\text{-}\mathsf{Mod}$, $\mathsf{Bij}_*\text{-}\mathsf{Mod}$, $\mathsf{Ord}_*\text{-}\mathsf{Mod}$, $\mathsf{Cyc}\text{-}\mathsf{Mod}$ and $\mathsf{Dih}\text{-}\mathsf{Mod}$. We then get a commutative diagram of forgetful functors:

$$\begin{tikzcd}
\mathsf{Dih}\text{-}\mathsf{Mod} \arrow[r] & \mathsf{Cyc}\text{-}\mathsf{Mod} \arrow[r] & \mathsf{Ord}_*\text{-}\mathsf{Mod} \\
\mathsf{Bij}\text{-}\mathsf{Mod} \arrow[u] \arrow[rr] & & \mathsf{Bij}_*\text{-}\mathsf{Mod} \arrow[u]
\end{tikzcd}$$

In the next definition we are using tensor products labeled by sets; see [22, Section 5.1.14] for more details on this notion.

Definition 1.6 (monads of trees) We consider the following monads of trees.

- The monad $\mathbb{T}\colon \mathsf{Bij}\text{-}\mathsf{Mod} \to \mathsf{Bij}\text{-}\mathsf{Mod}$ is defined via trees:

$$\mathbb{T}\mathcal{M}(S) := \bigoplus_{\mathsf{t}\in\mathsf{Tree}(S)} \Big(\bigotimes_{v\in V(\mathsf{t})} \mathcal{M}(E(v))\Big).$$

- The monad $\mathbb{RT}\colon \mathsf{Bij}_*\text{-}\mathsf{Mod} \to \mathsf{Bij}_*\text{-}\mathsf{Mod}$ is defined via rooted trees:

$$\mathbb{RT}\mathcal{M}(S,\rho) := \bigoplus_{\mathsf{t}\in\mathsf{RTree}(S,\rho)} \Big(\bigotimes_{v\in V(\mathsf{t})} \mathcal{M}(E(v),\rho(v))\Big).$$

- The monad $\mathbb{PRT}\colon \mathsf{Ord}_*\text{-}\mathsf{Mod} \to \mathsf{Ord}_*\text{-}\mathsf{Mod}$ is defined via planar rooted trees:

$$\mathbb{PRT}\mathcal{M}(S,\omega) := \bigoplus_{\mathsf{t}\in\mathsf{PRTree}(S,\omega)} \Big(\bigotimes_{v\in V(\mathsf{t})} \mathcal{M}(E(v),\omega(v))\Big).$$

- The monad $\mathbb{PT}\colon \mathsf{Cyc}\text{-}\mathsf{Mod} \to \mathsf{Cyc}\text{-}\mathsf{Mod}$ is defined via planar trees:

$$\mathbb{PT}\mathcal{M}(S,\gamma) := \bigoplus_{\mathsf{t}\in\mathsf{PTree}(S,\gamma)} \Big(\bigotimes_{v\in V(\mathsf{t})} \mathcal{M}(E(v),\gamma(v))\Big).$$

- The monad $\mathbb{DT}\colon \mathsf{Dih}\text{-}\mathsf{Mod} \to \mathsf{Dih}\text{-}\mathsf{Mod}$ is defined via dihedral trees:

$$\mathbb{DT}\mathcal{M}(S,\delta) := \bigoplus_{\mathsf{t}\in\mathsf{DTree}(S,\delta)} \Big(\bigotimes_{v\in V(\mathsf{t})} \mathcal{M}(E(v),\delta(v))\Big).$$

The composition law of these monads, eg $\mathbb{T}\circ\mathbb{T}\to\mathbb{T}$, is given by substitution of trees, and the unit, eg $1\to\mathbb{T}$, is given by the inclusion into the direct summand indexed by corollas. See [22, Section 5.6.1] for more details.

Remark 1.7 In the above commutative diagram, the horizontal forgetful functors commute with the respective monads: the forgetful functor $\mathsf{Dih}\text{-}\mathsf{Mod} \to \mathsf{Cyc}\text{-}\mathsf{Mod}$ commutes with $\mathbb{DT}$ and $\mathbb{PT}$; the forgetful functor $\mathsf{Cyc}\text{-}\mathsf{Mod} \to \mathsf{Ord}_*\text{-}\mathsf{Mod}$ commutes with $\mathbb{PT}$ and $\mathbb{PRT}$; the forgetful functor $\mathsf{Bij}\text{-}\mathsf{Mod} \to \mathsf{Bij}_*\text{-}\mathsf{Mod}$ commutes with $\mathbb{T}$ and $\mathbb{RT}$. There is no corresponding statement for the vertical forgetful functors.

Definition 1.8 (types of operads) An *operad* (resp. a *cyclic operad*, a *nonsymmetric operad*, a *nonsymmetric cyclic operad* and a *dihedral operad*) is an algebra over the monad $\mathbb{RT}$ of rooted trees (resp. the monad $\mathbb{T}$ of trees, the monad $\mathbb{PRT}$ of planar rooted trees, the monad $\mathbb{PT}$ of planar trees and the monad $\mathbb{DT}$ of dihedral trees).

Remark 1.9 In the rest of this article, we will always assume that all finite sets S have cardinality $n \geqslant 3$. This is more convenient for our geometric purposes, since the moduli spaces $\mathcal{M}_{0,S}$ and $\overline{\mathcal{M}}_{0,S}$ are only defined for those sets, and also to avoid speaking of polygons with two sides. The operads that we manipulate are then *nonunital operads*.

The aforementioned diagram of categories gives rise to the following forgetful functors between the categories of operads:

$$\begin{tikzcd}
\mathsf{Dih}\text{-}\mathsf{Op} \arrow[r] & \mathsf{ns}\text{-}\mathsf{Cyc}\text{-}\mathsf{Op} \arrow[r] & \mathsf{ns}\text{-}\mathsf{Op} \\
\mathsf{Cyc}\text{-}\mathsf{Op} \arrow[u] \arrow[rr] & & \mathsf{Op} \arrow[u]
\end{tikzcd}$$

Remark 1.10 In view of Remark 1.7, the free dihedral operad, the free nonsymmetric cyclic operad and the free nonsymmetric operad on a given Dih–module have the same underlying nonsymmetric operad.

1.3 Dihedral cooperads

By dualizing Definitions 1.6 and 1.8, one defines comonads of trees and the corresponding notions of cooperads. For more details, we refer the reader to [22, Section 5.8.8]. For instance, the comonad of trees is defined by the endofunctor $\mathbb{T}^c\colon \mathsf{Bij}\text{-}\mathsf{Mod} \to \mathsf{Bij}\text{-}\mathsf{Mod}$ defined by

$$\mathbb{T}^c\mathcal{M}(S) := \bigoplus_{\mathsf{t}\in\mathsf{Tree}(S)} \mathcal{M}(\mathsf{t}),$$

where we have set

$$\mathcal{M}(\mathsf{t}) := \bigotimes_{v\in V(\mathsf{t})} \mathcal{M}(E(v)).$$

A cyclic cooperad consists of a Bij–module $\mathcal{C}$ along with decomposition morphisms

$$\Delta_{\mathsf{t}}\colon \mathcal{C}(S) \to \mathcal{C}(\mathsf{t}),$$

for any tree $\mathsf{t} \in \mathsf{Tree}(S)$, satisfying some coassociativity conditions. For the convenience of the reader, we make the definition explicit in the case of dihedral cooperads, switching from dihedral trees to polygon dissections; see Lemma 1.4.

A Dih–module $\mathcal{M}$ assigns to every dihedral set (S,δ) an object $\mathcal{M}(S,\delta)$, and to every dihedral bijection $(S,\delta) \simeq (S',\delta')$ an isomorphism $\mathcal{M}(S,\delta) \simeq \mathcal{M}(S',\delta')$. We introduce the notation, for a dissection $\mathfrak{d} \in \mathsf{Diss}(S,\delta)$,

$$\mathcal{M}(\mathfrak{d}) := \bigotimes_{p \in P(\mathfrak{d})} \mathcal{M}(E(p), \delta(p)).$$

Definition 1.11 (comonad of dissections) The *comonad of dissections*, denoted by $\mathbb{DT}^c$, consists of the endofunctor $\mathbb{DT}^c \colon \mathsf{Dih\text{-}Mod} \to \mathsf{Dih\text{-}Mod}$ defined by

$$\mathbb{DT}^c\mathcal{M}(S,\delta) := \bigoplus_{\mathfrak{d} \in \mathsf{Diss}(S,\delta)} \mathcal{M}(\mathfrak{d}).$$

Its law $\mathbb{DT}^c \to \mathbb{DT}^c \circ \mathbb{DT}^c$ sends the direct summand indexed by a dissection $\mathfrak{d}$ to the direct summands indexed by all subdissections of $\mathfrak{d}$. The counit $\mathbb{DT}^c \to 1$ is the projection on the direct summand indexed by empty dissections.

Definition 1.12 (dihedral cooperad) A *dihedral cooperad* is a coalgebra over the comonad of dissections.

The data of a dihedral cooperad is equivalent to a collection of *decomposition morphisms*

$$\Delta_{\mathfrak{d}} \colon \mathcal{C}(S,\delta) \to \mathcal{C}(\mathfrak{d}),$$

for any dihedral tree $\mathfrak{d} \in \mathsf{Diss}(S,\delta)$, satisfying some coassociativity conditions. The first nontrivial decomposition morphisms correspond to dissections with one chord; such decomposition morphisms are called *infinitesimal* and their iterations can generate any decomposition morphism.

1.4 Cobar construction and cofree dihedral cooperads

In this subsection and in the next one, we assume that the underlying symmetric monoidal category A consists of graded objects, like chain complexes for instance. We use the cohomological convention for cooperads. In this case, one can consider the *desuspension* $s^{-1}\mathcal{C}$ of any dihedral module $\mathcal{C}$ defined by the formula $s^{-1}\mathcal{C}(S,\delta)^\bullet := \mathcal{C}(S,\delta)^{\bullet+1}$. (Alternatively, one can view the element s^{-1} as a dimension-one element of the category A concentrated in cohomological degree -1. In this case, the desuspension coincide with the tensor product with the element s^{-1}.)

Definition 1.13 (cobar construction) The *cobar construction* $\Omega\mathcal{C} := (\mathbb{DT}(s^{-1}\mathcal{C}), \mathrm{d})$ of a dihedral cooperad $\mathcal{C}$ is the free dihedral operad generated by $s^{-1}\mathcal{C}$ equipped with the unique derivation d which extends the infinitesimal decomposition morphisms of $\mathcal{C}$. The signs induced by the desuspension force the derivation d to square to zero, which makes the cobar construction into a differential graded dihedral operad.

Remark 1.14 As usual [22, Section 6.5.2], if the underlying dihedral module $\mathcal{C}$ carries an internal differential, one takes it into account in the definition of the cobar construction. This will not be the case in the sequel.

The underlying cochain complex of the cobar construction looks like

$$0 \to s^{-1}\mathcal{C}(S,\delta) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_1(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_2(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \cdots .$$

One can read whether a dihedral cooperad is cofree on its cobar construction as follows.

Proposition 1.15 *Let $\mathcal{C}$ be a dihedral cooperad. The following are equivalent:*

(i) *the dihedral cooperad $\mathcal{C}$ is cofree;*

(ii) *for every dihedral set (S,δ), the cobar construction of $\mathcal{C}$ induces a long exact sequence*

$$s^{-1}\mathcal{C}(S,\delta) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_1(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_2(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \cdots . \tag{1}$$

In such a situation, the space of cogenerators of $\mathcal{C}$ is (noncanonically) isomorphic to the space of indecomposables

$$\mathcal{X}(S,\delta) = \ker\Bigg(\mathcal{C}(S,\delta) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_1(S,\delta)} \mathcal{C}(\mathfrak{d})\Bigg).$$

More precisely, any choice of splitting for the inclusion of Dih*–modules $\mathcal{X} \hookrightarrow \mathcal{C}$ leads to an isomorphism*

$$\mathcal{C} \xrightarrow{\cong} \mathbb{DT}^c(\mathcal{X}).$$

Proof The long sequence (1) is exact if and only if the long sequence

$$0 \to s^{-1}\mathcal{X}(S,\delta) \to s^{-1}\mathcal{C}(S,\delta) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_1(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_2(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \cdots \tag{2}$$

is exact.

(i) $\Longrightarrow$ (ii) Suppose that the dihedral cooperad $\mathcal{C} \cong \mathbb{DT}^c(\mathcal{X})$ is cofree on a dihedral module $\mathcal{X}$. Since the sequence (2) is the analog of the bar-cobar resolution [22, Theorem 6.6.5] for the nilpotent dihedral operad $s^{-1}\mathcal{X}$, one proves that this sequence is exact by the same kind of arguments.

(ii) $\Longrightarrow$ (i) Let us assume that the long sequence (2) is exact. We choose a splitting $\mathcal{C} \twoheadrightarrow \mathcal{X}$ for the inclusions $\mathcal{X} \hookrightarrow \mathcal{C}$ in the category of Dih–modules. This defines a morphism of dihedral cooperads $\mathcal{C} \to \mathbb{DT}^c(\mathcal{X})$. Let us prove, by induction on the arity $n \geqslant 3$ of a dihedral set (S,δ), that the morphism $\mathcal{C}(S,\delta) \to \mathbb{DT}^c(\mathcal{X})(S,\delta)$ is an isomorphism. The case $n=3$ is obvious and initiates the induction. Suppose that the property holds up to $n-1$. We prove that it holds for n as follows. The preceding point shows that the long sequence (2) associated to the dihedral cooperad $\mathbb{DT}^c(\mathcal{X})$ is exact. The induction hypothesis provides us with the commutative diagram

$$\begin{tikzcd}
0 \arrow[d] & 0 \arrow[d] \\
s^{-1}\mathcal{X}(S,\delta) \arrow[r, "\cong"] \arrow[d] & s^{-1}\mathcal{X}(S,\delta) \arrow[d] \\
s^{-1}\mathcal{C}(S,\delta) \arrow[r] \arrow[d] & s^{-1}\mathbb{DT}^c(\mathcal{X})(S,\delta) \arrow[d] \\
\displaystyle\bigoplus_{\mathfrak{d}\in\mathsf{Diss}_1(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \arrow[r, "\cong"] \arrow[d] & \displaystyle\bigoplus_{\mathfrak{d}\in\mathsf{Diss}_1(S,\delta)} s^{-1}\mathbb{DT}^c(\mathcal{X})(\mathfrak{d}) \arrow[d] \\
\displaystyle\bigoplus_{\mathfrak{d}\in\mathsf{Diss}_2(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \arrow[r, "\cong"] \arrow[d] & \displaystyle\bigoplus_{\mathfrak{d}\in\mathsf{Diss}_2(S,\delta)} s^{-1}\mathbb{DT}^c(\mathcal{X})(\mathfrak{d}) \arrow[d] \\
\vdots & \vdots
\end{tikzcd}$$

where the columns are exact and nearly all the horizontal maps are isomorphisms. A diagram chase (the 5–lemma) completes the proof. □

Proposition 1.16 *Let $\mathcal{C}$ be a dihedral cooperad. Then the following statements are equivalent:*

(i) *The dihedral cooperad $\mathcal{C}$ is cofree.*

(ii) *The nonsymmetric cyclic cooperad underlying $\mathcal{C}$ is cofree.*

(iii) *The nonsymmetric cooperad underlying $\mathcal{C}$ is cofree.*

Proof The same proof shows that Proposition 1.15 is valid in the category of nonsymmetric cyclic cooperads (resp. nonsymmetric cooperads), replacing the dihedral cobar construction by the nonsymmetric cyclic cobar construction (resp. the nonsymmetric cobar construction). By Remark 1.10, these three cobar constructions have the same underlying nonsymmetric operad, which is the nonsymmetric cobar construction of the nonsymmetric cooperad underlying $\mathcal{C}$. In particular, they have the same underlying chain complex, and the claim follows. □

1.5 The coradical filtration and a freeness criterion

To understand the behavior of a dihedral cooperad with respect to the freeness property, one can consider its coradical filtration. This is the direct generalization of the same notion on the level of coalgebras [24, Appendix B] and on the level of cooperads [22, Section 5.8.4].

Definition 1.17 (coradical filtration) Let $\mathcal{C}$ be a dihedral cooperad. The *coradical filtration*, defined by

$$F_k\mathcal{C}(S,\delta) := \bigcap_{\mathfrak{d}\in\mathsf{Diss}_{k+1}(S,\delta)} \ker(\Delta_{\mathfrak{d}}),$$

for $k \geqslant 0$, is an increasing filtration of the Dih–module $\mathcal{C}$:

$$0 = F_{-1}\mathcal{C} \subset F_0\mathcal{C} \subset F_1\mathcal{C} \subset \cdots \subset \mathcal{C}.$$

The next proposition gives a way to recognize coradical filtrations of cofree dihedral cooperads. Let us make the following convention: if we are given an increasing filtration $\cdots \subset R_{k-1}\mathcal{C} \subset R_k\mathcal{C} \subset \cdots$ of a Dih–module $\mathcal{C}$, then we extend this filtration, in the natural way, to all objects $\mathcal{C}(\mathfrak{d})$ for a dissection $\mathfrak{d}$ as follows. If the dissection $\mathfrak{d}$ dissects (S,δ) into polygons $p_0, p_1, \ldots, p_k$, then we set

$$R_r\mathcal{C}(\mathfrak{d}) := \sum_{i_0+\cdots+i_k=r} R_{i_0}\mathcal{C}(p_0)\otimes\cdots\otimes R_{i_k}\mathcal{C}(p_k).$$

Proposition 1.18 *Let $\mathcal{C}$ be a dihedral cooperad. Assume that the underlying Dih–module is equipped with an increasing filtration*

$$0 = R_{-1}\mathcal{C} \subset R_0\mathcal{C} \subset R_1\mathcal{C} \subset \cdots \subset \mathcal{C}$$

which is finite in every arity n and such that the following properties are satisfied:

(a) *for every dissection $\mathfrak{d} \in \mathsf{Diss}_k(S,\delta)$ of cardinality k and every integer r, the decomposition map $\Delta_{\mathfrak{d}}$ sends $R_r\mathcal{C}(S,\delta)$ to $R_{r-k}\mathcal{C}(S,\delta)$;*

(b) *for every integer r, the iterated decomposition map*

$$\mathrm{gr}_r^R\,\mathcal{C}(S,\delta) \xrightarrow{\bigoplus\Delta_{\mathfrak{d}}} \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_r(S,\delta)} R_0\mathcal{C}(\mathfrak{d}) \tag{3}$$

is an isomorphism.

Then the dihedral cooperad $\mathcal{C}$ is cofree and the filtration R is its coradical filtration. More precisely, any choice of splitting of the inclusion $R_0\mathcal{C} \hookrightarrow \mathcal{C}$ induces an isomorphism

$$\mathcal{C} \xrightarrow{\cong} \mathbb{DT}^c(R_0\mathcal{C}).$$

Proof Let us choose a splitting $\theta\colon \mathcal{C} \twoheadrightarrow R_0\mathcal{C}$ for the inclusions $R_0\mathcal{C} \hookrightarrow \mathcal{C}$ in the category of Dih–modules. By the universal property of the cofree dihedral cooperads, this induces a morphism of dihedral cooperads $\Theta\colon \mathcal{C} \to \mathbb{DT}^c(R_0\mathcal{C})$. The coradical filtration on the cofree dihedral cooperad $\mathbb{DT}^c(R_0\mathcal{C})$ is given by

$$F_k\mathbb{DT}^c(R_0\mathcal{C})(S,\delta) = \bigoplus_{\substack{r\leqslant k \\ \mathfrak{d}\in \mathsf{Diss}_r(S,\delta)}} R_0\mathcal{C}(\mathfrak{d}).$$

For any dissection $\mathfrak{d} \in \mathsf{Diss}_k(S,\delta)$ and $k > r$, the first assumption implies that we have $\Delta_{\mathfrak{d}}(R_r\mathcal{C}(S,\delta)) = 0$. Therefore, the morphism Θ is compatible with the filtrations R and it induces a morphism of graded dihedral modules

$$\mathrm{gr}_r^R\,\Theta\colon \mathrm{gr}_r^R\,\mathcal{C}(S,\delta) \to \mathrm{gr}_r^R\,\mathbb{DT}^c(R_0\mathcal{C}) = \bigoplus_{\mathfrak{d}\in \mathsf{Diss}_r(S,\delta)} R_0\mathcal{C}(\mathfrak{d}),$$

which is nothing but the iterated decomposition map (3). So it is an isomorphism by the second assumption. Finally, the morphism of dihedral cooperads Θ is an isomorphism and the proposition is proved. □

2 Moduli spaces of genus zero curves and the cyclic gravity operad

In this section, we begin by recalling the definitions of the moduli space of genus zero curves with marked points and its Deligne–Mumford–Knudsen compactification. We recall the definition of residues along normal crossing divisors in the context of mixed Hodge theory. This produces the cyclic gravity operad structure on the cohomology of the moduli spaces of curves.

2.1 Normal crossing divisors and stratifications

We introduce some vocabulary and notations on normal crossing divisors and the stratifications that they induce on complex algebraic varieties.

2.1.1 The local setting Let $\overline{X}$ be a small neighborhood of 0 in $\mathbb{C}^n$ and let us define a divisor $\partial\overline{X} = \{z_1 \cdots z_r = 0\}$ in $\overline{X}$ for some fixed integer r. Its irreducible components are the (intersections with $\bar{X}$ of the) coordinate hyperplanes $\{z_i = 0\}$ for $i = 1, \dots, r$. This induces a stratification

$$\bar{X} = \bigsqcup_{I \subset \{1,\dots,r\}} X(I), \tag{4}$$

where $X(I)$ is the locally closed subset of $\bar{X}$ defined by the conditions: $z_i = 0$ for $i \in I$ and $z_i \neq 0$ for $i \in \{1, \dots, r\} \setminus I$. Notice that

$$I \subset I' \iff \bar{X}(I) \supset \bar{X}(I').$$

The codimension of $X(I)$ is equal to the cardinality of I, and its closure $\bar{X}(I)$ is defined by the vanishing of the coordinates z_i, $i \in I$. In other words, the closure $\bar{X}(I)$ is the union of the strata $X(I')$ for $I' \supset I$:

$$\bar{X}(I) = \bigsqcup_{I' \supset I} X(I').$$

For a given set $I \subset \{1, \dots, r\}$, the complement $\partial \bar{X}(I) := \bar{X}(I) \setminus X(I)$ is defined by the equation $\prod_{i \in \{1,\dots,r\} \setminus I} z_i = 0$.

2.1.2 The global setting Let $\bar{X}$ be a smooth (not necessarily compact) complex algebraic variety and let $\partial \bar{X}$ be a normal crossing divisor inside $\bar{X}$. This means that around every point of $\bar{X}$, there is a system of coordinates $(z_1, \dots, z_n)$, where n is the complex dimension of $\bar{X}$, such that $\partial \bar{X}$ is defined by an equation of the form $z_1 \cdots z_r = 0$ for some integer r that depends on the point.

This induces a global stratification

$$\bar{X} = \bigsqcup_{\mathfrak{s} \in \mathsf{Strat}} X(\mathfrak{s}) \tag{5}$$

that is constructed as (4) in every local chart. For every $\mathfrak{s}$ in the indexing set Strat, the stratum $X(\mathfrak{s})$ is a connected locally closed subset of $\bar{X}$. Let $\bar{X}(\mathfrak{s})$ denote its closure. The indexing set Strat for the strata is actually endowed with a poset structure defined by

$$\mathfrak{s} \leqslant \mathfrak{s}' \iff \bar{X}(\mathfrak{s}) \supset \bar{X}(\mathfrak{s}').$$

In other words, the closure $\bar{X}(\mathfrak{s})$ of $X(\mathfrak{s})$ is the union of the strata $X(\mathfrak{s}')$ for $\mathfrak{s}' \geqslant \mathfrak{s}$:

$$\bar{X}(\mathfrak{s}) = \bigsqcup_{\mathfrak{s}' \geqslant \mathfrak{s}} X(\mathfrak{s}').$$

For an integer k, we write Strat_k for the indexing set of strata of codimension k, making Strat into a graded poset. The set Strat_0 only has one element corresponding to the open stratum $X = \bar{X} \setminus \partial \bar{X}$. The closures $\bar{X}(\mathfrak{s})$, for $\mathfrak{s} \in \mathsf{Strat}_1$, are the irreducible components of the normal crossing divisor $\partial \bar{X}$.

For a given stratum $X(\mathfrak{s})$, the complement $\partial\overline{X}(\mathfrak{s}) := \overline{X}(\mathfrak{s}) \setminus X(\mathfrak{s})$ is a normal crossing divisor inside $\overline{X}(\mathfrak{s})$.

2.2 The moduli spaces $\mathcal{M}_{0,S}$ and $\overline{\mathcal{M}}_{0,S}$

We introduce the moduli spaces of genus zero curves $\mathcal{M}_{0,S}$ and $\overline{\mathcal{M}}_{0,S}$. We refer the reader to [20; 18; 14; 17] for more details.

2.2.1 The open moduli spaces $\mathcal{M}_{0,S}$ Let S be a finite set of cardinality $n \geqslant 3$. The *moduli space of genus zero curves with S–marked points* is the quotient of the configuration space of points labeled by S on the Riemann sphere $\mathbb{P}^1(\mathbb{C})$ by the automorphisms of $\mathbb{P}^1(\mathbb{C})$. It is denoted by

$$\mathcal{M}_{0,S} := \big\{(z_s)_{s\in S} \in \mathbb{P}^1(\mathbb{C})^S \mid z_s \neq z_{s'} \text{ for all } s \neq s'\big\}/\mathrm{PGL}_2(\mathbb{C}),$$

where an element $g \in \mathrm{PGL}_2(\mathbb{C})$ acts diagonally by $g.(z_s)_{s\in S} = (g.z_s)_{s\in S}$.

Every bijection $S \simeq S'$ induces an isomorphism $\mathcal{M}_{0,S} \simeq \mathcal{M}_{0,S'}$. If $S = \{1,\dots,n\}$ then $\mathcal{M}_{0,S}$ is simply denoted by $\mathcal{M}_{0,n}$.

The action of $\mathrm{PGL}_2(\mathbb{C})$ on $\mathbb{P}^1(\mathbb{C})$ is strictly tritransitive: for every triple (a,b,c) of pairwise distinct points on $\mathbb{P}^1(\mathbb{C})$, there exists a unique element $g \in \mathrm{PGL}_2(\mathbb{C})$ such that $(g.a, g.b, g.c) = (\infty, 0, 1)$. By fixing an identification

$$(z_1,\dots,z_n) = (\infty, 0, t_1, \dots, t_{n-3}, 1),$$

we can thus get rid of the quotient by $\mathrm{PGL}_2(\mathbb{C})$ and obtain an isomorphism

$$\text{(6)} \quad \mathcal{M}_{0,n} \simeq \big\{(t_1,\dots,t_{n-3}) \in \mathbb{C}^{n-3} \mid t_i \neq 0,1 \text{ for all } i, \text{ and } t_i \neq t_j \text{ for all } i \neq j\big\}.$$

This description makes it clear that $\mathcal{M}_{0,S}$ is a smooth and affine complex algebraic variety of dimension $n-3$.

2.2.2 The compactified moduli spaces $\overline{\mathcal{M}}_{0,S}$ Let S be a finite set of cardinality $n \geqslant 3$, and let

$$\mathcal{M}_{0,S} \subset \overline{\mathcal{M}}_{0,S}$$

be the Deligne–Mumford–Knudsen compactification of $\mathcal{M}_{0,S}$. Every bijection $S \simeq S'$ induces an isomorphism $\overline{\mathcal{M}}_{0,S} \simeq \overline{\mathcal{M}}_{0,S'}$. If $S = \{1,\dots,n\}$, then $\overline{\mathcal{M}}_{0,S}$ is simply denoted by $\overline{\mathcal{M}}_{0,n}$.

The compactified moduli space $\overline{\mathcal{M}}_{0,S}$ is a smooth projective complex algebraic variety, and the complement $\partial\overline{\mathcal{M}}_{0,S} := \overline{\mathcal{M}}_{0,S} \setminus \mathcal{M}_{0,S}$ is a simple normal crossing divisor. The corresponding stratification (5) is indexed by the graded poset of S–trees:

(7)
$$\overline{\mathcal{M}}_{0,S} = \bigsqcup_{\mathrm{t}\in\mathsf{Tree}(S)} \mathcal{M}(\mathrm{t}).$$

The codimension of a stratum $\mathcal{M}(\mathrm{t})$ is equal to the number of internal edges of the tree t. If we denote by $\overline{\mathcal{M}}(\mathrm{t})$ the closure of a stratum $\mathcal{M}(\mathrm{t})$ in $\overline{\mathcal{M}}_{0,S}$, then we have

$$\overline{\mathcal{M}}(\mathrm{t}) \supset \overline{\mathcal{M}}(\mathrm{t}') \iff \mathrm{t} \leqslant \mathrm{t}',$$

where the order $\leqslant$ on trees is the one defined in Definition 1.3. The closure $\overline{\mathcal{M}}(\mathrm{t})$ is thus the union of the strata $\mathcal{M}(\mathrm{t}')$ for $\mathrm{t}' \geqslant \mathrm{t}$.

For a tree $\mathrm{t} \in \mathsf{Tree}(S)$, we have compatible product decompositions

(8)
$$\mathcal{M}(\mathrm{t}) \cong \prod_{v\in V(\mathrm{t})} \mathcal{M}_{0,E(v)} \quad \text{and} \quad \overline{\mathcal{M}}(\mathrm{t}) \cong \prod_{v\in V(\mathrm{t})} \overline{\mathcal{M}}_{0,E(v)}.$$

The stratum corresponding to the corolla is the open stratum $\mathcal{M}_{0,S}$. For $\mathrm{t} \in \mathsf{Tree}_1(S)$ a tree with only one internal edge, we get a divisor

$$\overline{\mathcal{M}}(\mathrm{t}) \cong \overline{\mathcal{M}}_{0,E_0} \times \overline{\mathcal{M}}_{0,E_1}$$

inside $\overline{\mathcal{M}}_{0,S}$. These divisors are the irreducible components of $\partial\overline{\mathcal{M}}_{0,S}$.

Example 2.1 (1) We have $\mathcal{M}_{0,3} = \overline{\mathcal{M}}_{0,3} = \{*\}$.

(2) If we write $\mathcal{M}_{0,4} = \mathbb{P}^1(\mathbb{C})\setminus\{\infty, 0, 1\}$ as in (6), then we have $\overline{\mathcal{M}}_{0,4} = \mathbb{P}^1(\mathbb{C})$. The divisor at infinity $\partial\overline{\mathcal{M}}_{0,4} = \{\infty, 0, 1\}$ has three irreducible components, all isomorphic to a product $\overline{\mathcal{M}}_{0,3} \times \overline{\mathcal{M}}_{0,3}$, indexed by the three 4–trees with one internal edge.

(3) If we write $\mathcal{M}_{0,5} = (\mathbb{P}^1(\mathbb{C}) \setminus \{\infty, 0, 1\})^2 \setminus \{t_1 = t_2\}$ as in (6), then $\overline{\mathcal{M}}_{0,5}$ can be realized as the blow-up of $\mathbb{P}^1(\mathbb{C})^2$ along the three points $(0,0)$, $(1,1)$ and (∞,∞); see Figure 3. The divisor at infinity $\partial\overline{\mathcal{M}}_{0,5}$ has ten irreducible components: the three exceptional divisors and the strict transforms of the lines $t_1 = 0, 1, \infty$, $t_2 = 0, 1, \infty$ and $t_1 = t_2$. They are all isomorphic to a product $\overline{\mathcal{M}}_{0,3} \times \overline{\mathcal{M}}_{0,4}$ and are indexed by the ten 5–trees with one internal edge. The fifteen different intersection points of these components are indexed by the fifteen 5–trees with two internal edges.

2.3 The category of mixed Hodge structures

We recall some useful facts on the category of mixed Hodge structures. The main references are the original articles by Deligne [6; 7; 8] and the book [23].

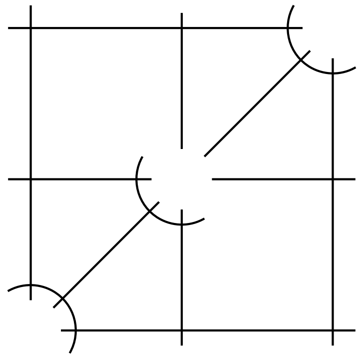


Figure 3: The combinatorial structure of $\overline{\mathcal{M}}_{0,5}$

Definition 2.2 (pure Hodge structures) A *pure Hodge structure* of weight w is the data of

- a finite-dimensional $\mathbb{Q}$–vector space H;
- a finite decreasing filtration, the *Hodge filtration*, $F^\bullet H_{\mathbb{C}}$ of the complexification $H_{\mathbb{C}} := H \otimes_{\mathbb{Q}} \mathbb{C}$,

such that for every integer p, we have

$$H_{\mathbb{C}} = F^p H_{\mathbb{C}} \oplus \overline{F^{w-p+1} H_{\mathbb{C}}}.$$

A morphism of pure Hodge structures is a morphism of $\mathbb{Q}$–vector spaces that is compatible with the Hodge filtration.

Definition 2.3 (mixed Hodge structures) A *mixed Hodge structure* is the data of

- a finite-dimensional $\mathbb{Q}$–vector space H;
- a finite increasing filtration, the *weight filtration*, $W_\bullet H$ of H;
- a finite decreasing filtration, the *Hodge filtration*, $F^\bullet H_{\mathbb{C}}$ of the complexification $H_{\mathbb{C}}$,

such that for every integer w, the Hodge filtration induces a pure Hodge structure of weight w on $\mathrm{gr}_w^W H := W_w H / W_{w-1} H$. A morphism of mixed Hodge structures is a morphism of $\mathbb{Q}$–vector spaces that is compatible with the weight and Hodge filtrations.

A pure Hodge structure of weight w is thus nothing but a mixed Hodge structure whose weight filtration is concentrated in weight w.

A very important remark is that morphisms of mixed Hodge structures are strictly compatible with the weight and Hodge filtrations. This implies that mixed Hodge structures form an abelian category. One easily defines on it a compatible structure of a symmetric monoidal category.

Another consequence of this strictness property is the following lemma, used in practice to prove degeneration of spectral sequences, like in Proposition 3.10.

Lemma 2.4 *Let $f\colon H \to H'$ be a morphism of mixed Hodge structures. If H is pure of weight w and H' is pure of weight w' with $w \neq w'$, then $f = 0$.*

The *pure Tate structure of weight $2k$*, denoted by $\mathbb{Q}(-k)$, is the only pure Hodge structure of weight $2k$ and dimension 1; its Hodge filtration is concentrated in degree k. They satisfy $\mathbb{Q}(-k) \otimes \mathbb{Q}(-l) \cong \mathbb{Q}(-k-l)$ and $\mathbb{Q}(-k)^\vee \cong \mathbb{Q}(k)$. A mixed Hodge structure is said to be *pure Tate of weight $2k$* if it is isomorphic to a direct sum $\mathbb{Q}(-k)^{\oplus d}$ for a certain integer d.

If H is a mixed Hodge structure and k is an integer, we denote by $H(-k)$ the *Tate twist* of H consisting in shifting the weight filtration by $2k$ and the Hodge filtration by k. It is equal to the tensor product of H by $\mathbb{Q}(-k)$.

The importance of mixed Hodge structures in the study of the topology of complex algebraic varieties is explained by the following fundamental theorem of Deligne.

Theorem 2.5 [8, Proposition 8.2.2] *Let X be a complex algebraic variety. For every integer k, the cohomology group $H^k(X)$ is endowed with a functorial mixed Hodge structure.*

2.4 Logarithmic forms and residues

We recall the notion of logarithmic form along a normal crossing divisor and that of a residue. We refer the reader to [7, 3.1] for more details.

2.4.1 The local setting We work in the local setting of Section 2.1.1. We say that a meromorphic differential form on $\bar{X}$ has *logarithmic poles* along $\partial\bar{X}$, or that it is a *logarithmic form* on $(\bar{X}, \partial\bar{X})$, if it can be written as a linear combination of forms of the type

$$\frac{dz_{i_1}}{z_{i_1}} \wedge \cdots \wedge \frac{dz_{i_s}}{z_{i_s}} \wedge \eta,$$

with $1 \leqslant i_1 < \cdots < i_s \leqslant r$ and where η a holomorphic form on $\bar{X}$. Logarithmic forms are closed under the exterior derivative on forms.

Any logarithmic form on $(\bar{X}, \partial\bar{X})$ can be written as

$$\omega = \frac{dz_1}{z_1} \wedge \alpha + \beta,$$

where α and β are forms with logarithmic poles along $\{z_2 \cdots z_r = 0\}$. We define the *residue* of ω on $\overline{X}(1) = \{z_1 = 0\}$ to be the restriction

$$\mathrm{Res}(\omega) := 2\pi i \, \alpha_{|\overline{X}(1)}. \tag{9}$$

It is a well-defined logarithmic form on $(\overline{X}(1), \partial\overline{X}(1))$. The residue operation lowers the degree of the forms by 1 and anticommutes with the exterior derivative: $d \circ \mathrm{Res} + \mathrm{Res} \circ d = 0$.

More generally, for sets $I \subset I' \subset \{1, \dots, r\}$ with $|I'| = |I| + 1$, we get residue operations $\mathrm{Res}^{I}_{I'}$ from logarithmic forms on $(\overline{X}(I), \partial\overline{X}(I))$ to logarithmic forms on $(\overline{X}(I'), \partial\overline{X}(I'))$.

2.4.2 The global setting We work in the global setting of Section 2.1.2. By gluing together the local definitions of the previous paragraph, one defines on each closure $\overline{X}(\mathfrak{s})$ a complex of sheaves of logarithmic forms on $(\overline{X}(\mathfrak{s}), \partial\overline{X}(\mathfrak{s}))$:

$$\Omega^{\bullet}_{\overline{X}(\mathfrak{s})}(\log \partial\overline{X}(\mathfrak{s})).$$

If $j_{\mathfrak{s}}\colon X(\mathfrak{s}) \hookrightarrow \overline{X}(\mathfrak{s})$ denotes the natural open immersion, we have a quasi-isomorphism $(j_{\mathfrak{s}})_* \mathbb{C}_{X(\mathfrak{s})} \simeq \Omega^{\bullet}_{\overline{X}(\mathfrak{s})}(\log \partial\overline{X}(\mathfrak{s}))$, which induces isomorphisms between cohomology groups:

$$H^k(X(\mathfrak{s}), \mathbb{C}) \cong \mathbb{H}^k\big(\overline{X}(\mathfrak{s}), \Omega^{\bullet}_{\overline{X}(\mathfrak{s})}(\log \partial\overline{X}(\mathfrak{s}))\big). \tag{10}$$

For elements $\mathfrak{s} \leqslant \mathfrak{s}'$ in Strat with $|\mathfrak{s}'| = |\mathfrak{s}| + 1$, we denote the corresponding closed immersion by $i^{\mathfrak{s}}_{\mathfrak{s}'}\colon \overline{X}(\mathfrak{s}') \hookrightarrow \overline{X}(\mathfrak{s})$. By applying the local construction of the previous paragraph in every local chart, we get a residue morphism

$$\mathrm{Res}^{\mathfrak{s}}_{\mathfrak{s}'}\colon \Omega^{\bullet}_{\overline{X}(\mathfrak{s})}(\log \partial\overline{X}(\mathfrak{s})) \to (i^{\mathfrak{s}}_{\mathfrak{s}'})_* \Omega^{\bullet-1}_{\overline{X}(\mathfrak{s}')}(\log \partial\overline{X}(\mathfrak{s}')), \tag{11}$$

which anticommutes with the exterior derivative on forms. In view of (10), this induces a residue morphism between cohomology groups:

$$\mathrm{Res}^{\mathfrak{s}}_{\mathfrak{s}'}\colon H^{\bullet}(X(\mathfrak{s}), \mathbb{C}) \to H^{\bullet-1}(X(\mathfrak{s}'), \mathbb{C}).$$

This residue morphism is actually defined over $\mathbb{Q}$ and it is compatible with the mixed Hodge structures if we add the right Tate twist, giving rise to residue morphisms

$$\mathrm{Res}^{\mathfrak{s}}_{\mathfrak{s}'}\colon H^{\bullet}(X(\mathfrak{s})) \to H^{\bullet-1}(X(\mathfrak{s}'))(-1). \tag{12}$$

2.5 The cyclic gravity cooperad

Following Getzler, we use the residue morphisms of the previous paragraph to define the cyclic gravity cooperad in the category of graded mixed Hodge structures. Let S

be a finite set of cardinality $n \geqslant 3$, and let us choose an S–tree $\mathfrak{t} \in \mathsf{Tree}_1(S)$ with one internal edge. Let us denote by v_0 and v_1 its two vertices, and by $E_0 := E(v_0)$ and $E_1 := E(v_1)$ the corresponding sets of adjacent edges. The stratum indexed by $\mathfrak{t}$ in the moduli space $\mathcal{M}_{0,S}$ is

$$\mathcal{M}(\mathfrak{t}) \cong \mathcal{M}_{0,E_0} \times \mathcal{M}_{0,E_1}.$$

For integers a and b, we thus get residue morphisms

$$\Delta_{\mathfrak{t}}\colon H^{a+b-1}(\mathcal{M}_{0,S})(-1) \to H^{a-1}(\mathcal{M}_{0,E_0})(-1) \otimes H^{b-1}(\mathcal{M}_{0,E_1})(-1). \tag{13}$$

They are obtained from (12) by using the Künneth formula, adding a Tate twist (-1) and multiplying by the Koszul sign $(-1)^{a-1}$, which reflects the cohomological degree shift. Let us define the Bij–module $\mathcal{C}$ in the category of graded mixed Hodge structures by

$$\mathcal{C}(S) := H^{\bullet-1}(\mathcal{M}_{0,S})(-1).$$

Associated to any set V, one considers the one-dimensional vector space $\det(V) := \bigwedge_{v\in V} \mathbb{Q}v$. The signed residue morphisms (13) are not quite the decomposition morphisms of a cyclic cooperad. Instead, they give rise to decomposition morphisms

$$\Delta_{\mathfrak{t}}\colon \mathcal{C}(S) \to \det(V(\mathfrak{t})) \otimes \mathcal{C}(\mathfrak{t}) \tag{14}$$

for any S–tree $\mathfrak{t} \in \mathsf{Tree}(S)$, that satisfy analogs of the axioms a cyclic cooperad, but with different signs. Such an algebraic structure on $\mathcal{C}$ is actually called an *anticyclic cooperad*; see [15, 2.10]. Note that in (13) the choice of an ordering $V(\mathfrak{t}) = \{v_0, v_1\}$ gives a trivialization $\det(V(\mathfrak{t})) \simeq \mathbb{Q}$ of the determinant. The following definition was introduced by Getzler [13; 14].

Definition 2.6 (cyclic gravity cooperad) The *cyclic gravity cooperad* is the cyclic suspension [15, 2.10] of the anticyclic cooperad $\mathcal{C}$:

$$\mathcal{G}rav(S) := \det(S) \otimes H^{\bullet+n-3}(\mathcal{M}_{0,S})(-1)$$

for any finite set S of cardinality $n \geqslant 3$. It forms a cyclic cooperad in the category of graded mixed Hodge structures, which is concentrated in nonpositive cohomological degree $-(n-3) \leqslant \bullet \leqslant 0$. The decomposition morphisms

$$\Delta_{\mathfrak{t}}\colon \mathcal{G}rav(S) \to \mathcal{G}rav(\mathfrak{t})$$

for the cyclic gravity cooperad are given by signed residues.

Getzler showed [13, Theorem 4.5] that the cyclic gravity operad, linear dual to the cyclic gravity cooperad, is generated by one element in each cyclic arity $n \geqslant 3$, and he also gave a presentation for the operadic ideal of relations. More specifically, the generator in cyclic arity n is the natural generator of the space $H_0(\mathcal{M}_{0,n})(1)$, which lies in homological degree $-(n-3)$, and the relations are generalizations of the Jacobi identity for Lie algebras. In particular, the generator of cyclic arity 3 satisfies the Jacobi identity, and one gets the following theorem.

Theorem 2.7 [14, 3.8] *The degree-zero suboperad of the cyclic gravity operad is isomorphic to the cyclic Lie operad. In particular, we get an isomorphism of* Bij*–modules*

$$\mathcal{L}ie(S) \cong \det(S) \otimes H_{n-3}(\mathcal{M}_{0,S})(1).$$

3 Brown's moduli spaces and the dihedral gravity cooperad

In this section, we introduce Brown's moduli spaces as a partial compactification of the moduli spaces of genus zero curves. Forgetting many of the symmetries of the gravity operad, one obtains the dihedral gravity operad. We conclude with the proof of the equivalence between the purity of the mixed Hodge structure on the cohomology of Brown's moduli spaces and the cofreeness of the dihedral gravity cooperad.

3.1 Brown's moduli spaces $\mathcal{M}^{\delta}_{0,S}$

Let S be a finite set of cardinality $n \geqslant 3$ and let δ be a dihedral structure on S. Brown defined [5, Section 2] a space $\mathcal{M}^{\delta}_{0,S}$ that fits between the moduli space $\mathcal{M}_{0,S}$ and its compactification $\overline{\mathcal{M}}_{0,S}$ with open immersions:

$$\mathcal{M}_{0,S} \subset \mathcal{M}^{\delta}_{0,S} \subset \overline{\mathcal{M}}_{0,S}.$$

Recall that $\mathsf{DTree}(S,\delta) \subset \mathsf{Tree}(S)$ denotes the set of S–trees that have a dihedral embedding compatible with δ.

Definition 3.1 (Brown's moduli space $\mathcal{M}^{\delta}_{0,S}$) Brown's moduli space $\mathcal{M}^{\delta}_{0,S}$ is the subspace of $\overline{\mathcal{M}}_{0,S}$ defined as the union of strata indexed by the trees underlying dihedral trees:

$$\mathcal{M}^{\delta}_{0,S} := \bigsqcup_{\mathfrak{t}\in\mathsf{DTree}(S,\delta)} \mathcal{M}(\mathfrak{t}).$$

For $\mathfrak{t}$ and $\mathfrak{t}'$ two S–trees such that $\mathfrak{t} \leqslant \mathfrak{t}'$, we have $\mathfrak{t}' \in \mathsf{DTree}(S,\delta) \Longrightarrow \mathfrak{t} \in \mathsf{DTree}(S,\delta)$; thus, Brown's moduli space $\mathcal{M}^{\delta}_{0,S}$ is an open subvariety of $\overline{\mathcal{M}}_{0,S}$. In other words, it is the complement in $\overline{\mathcal{M}}_{0,S}$ of the union of the closed subvarieties $\overline{\mathcal{M}}(\mathfrak{t})$ for $\mathfrak{t} \in \mathsf{Tree}(S) \setminus \mathsf{DTree}(S,\delta)$; in this description, it is actually enough to delete the divisors $\overline{\mathcal{M}}(\mathfrak{t})$ for trees $\mathfrak{t}$ with one internal edge.

Every dihedral bijection $(S,\delta) \simeq (S',\delta')$ induces an isomorphism $\mathcal{M}^{\delta}_{0,S} \simeq \mathcal{M}^{\delta'}_{0,S'}$. If we consider $S = \{1,\ldots,n\}$ with its standard dihedral structure δ, then $\mathcal{M}^{\delta}_{0,S}$ is simply denoted by $\mathcal{M}^{\delta}_{0,n}$.

Theorem 3.2 [5, Theorem 2.21] *Brown's moduli space $\mathcal{M}^{\delta}_{0,S}$ is a smooth and affine complex algebraic variety, and the complement $\partial\mathcal{M}^{\delta}_{0,S} := \mathcal{M}^{\delta}_{0,S} \setminus \mathcal{M}_{0,S}$ is a normal crossing divisor.*

With our definition of Brown's moduli spaces, the only nontrivial statement in the above theorem is the fact that $\mathcal{M}^{\delta}_{0,S}$ is affine. Brown's original definition is via an explicit presentation of the ring of functions of $\mathcal{M}^{\delta}_{0,S}$. The equivalence of the two definitions can be found in [5, Section 2.6].

3.2 The dihedral gravity cooperad

Definition 3.3 (the dihedral gravity operad) The *dihedral gravity cooperad*, still denoted by $\mathcal{G}rav$, is the dihedral cooperad in the category of graded mixed Hodge structures underlying the cyclic gravity cooperad. In other words, it is obtained by applying the forgetful functor $\mathsf{Cyc\text{-}Op} \to \mathsf{Dih\text{-}Op}$ of Section 1.2 to the cyclic gravity operad. Recall that its underling dihedral module is given by

$$\mathcal{G}rav(S,\delta) := \det(S) \otimes H^{\bullet+n-3}(\mathcal{M}_{0,S})(-1).$$

For the convenience of the reader and for future use, we restate its definition in the dihedral setting by using the bijection between graded posets of Lemma 1.4:

$$\mathsf{DTree}(S,\delta) \cong \mathsf{Diss}(S,\delta), \quad \mathfrak{t} \leftrightarrow \mathfrak{d}.$$

We may then write

$$\mathcal{M}^{\delta}_{0,S} = \bigsqcup_{\mathfrak{d}\in\mathsf{Diss}(S,\delta)} \mathcal{M}(\mathfrak{d}). \tag{15}$$

The codimension of a stratum $\mathcal{M}(\mathfrak{d})$ is the number of chords in the dissection $\mathfrak{d}$. If we denote by $\mathcal{M}^{\delta}(\mathfrak{d})$ the closure of a stratum $\mathcal{M}(\mathfrak{d})$ in $\mathcal{M}^{\delta}_{0,S}$, then we have

$$\mathcal{M}^{\delta}(\mathfrak{d}) \supset \mathcal{M}^{\delta}(\mathfrak{d}') \iff \mathfrak{d} \leqslant \mathfrak{d}',$$

where the order $\leqslant$ on dissections is the one defined in Section 1.1. The closure $\mathcal{M}^{\delta}(\mathfrak{d})$ is thus the union of the strata $\mathcal{M}(\mathfrak{d}')$ for $\mathfrak{d}' \geqslant \mathfrak{d}$.

For a dissection $\mathfrak{d} \in \mathsf{Diss}(S,\delta)$, we have the product decompositions

$$\mathcal{M}(\mathfrak{d}) \cong \prod_{p\in P(\mathfrak{d})} \mathcal{M}_{0,E(p)} \quad \text{and} \quad \mathcal{M}^{\delta}(\mathfrak{d}) \cong \prod_{p\in P(\mathfrak{d})} \mathcal{M}^{\delta(p)}_{0,E(p)},$$

which are compatible with the product decompositions (8).

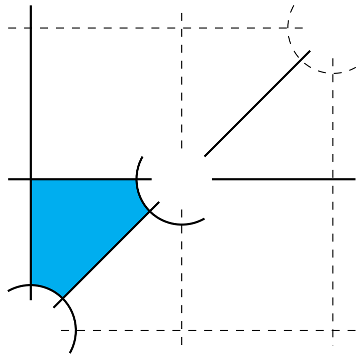


Figure 4: The combinatorial structure of $\mathcal{M}^{\delta}_{0,5}$

The stratum corresponding to the corolla is the open stratum $\mathcal{M}_{0,S}$. For $\mathfrak{d} = \{c\} \in \mathsf{Diss}_1(S,\delta)$ a dissection consisting of only one chord, we get a divisor

$$\mathcal{M}^{\delta}(\{c\}) \cong \mathcal{M}^{\delta_0}_{0,E_0} \times \mathcal{M}^{\delta_1}_{0,E_1}$$

inside $\mathcal{M}^{\delta}_{0,S}$. These divisors are the irreducible components of $\partial\mathcal{M}^{\delta}_{0,S}$.

Example 3.4 (1) We have $\mathcal{M}^{\delta}_{0,3} = \{*\}$.

(2) If we write $\mathcal{M}_{0,4} = \mathbb{P}^1(\mathbb{C}) \setminus \{\infty, 0, 1\}$ and $\overline{\mathcal{M}}_{0,4} = \mathbb{P}^1(\mathbb{C})$, then we have $\mathcal{M}^{\delta}_{0,4} = \mathbb{P}^1(\mathbb{C}) \setminus \{\infty\}$. The divisor at infinity $\partial\mathcal{M}^{\delta}_{0,4} = \{0,1\}$ has two irreducible components, all isomorphic to a product $\mathcal{M}^{\delta}_{0,3} \times \mathcal{M}^{\delta}_{0,3}$, indexed by the two dissection of a 4–gon with one chord.

(3) Figure 4 shows the combinatorial structure of $\mathcal{M}^{\delta}_{0,5}$ inside $\overline{\mathcal{M}}_{0,5}$. The curves in dashed lines are the complement $\overline{\mathcal{M}}_{0,5} \setminus \mathcal{M}^{\delta}_{0,5}$. The five curves in straight lines are the five irreducible components of the divisor at infinity $\partial\mathcal{M}^{\delta}_{0,5}$, indexed by the five dissection of a 5–gon with one chord. They bound a pentagon (shaded). The five different intersection points of these components are indexed by the five dissections of a 5–gon with two chords.

Remark 3.5 The stratification of $\mathcal{M}^{\delta}_{0,n}$ has the same combinatorial structure as the natural stratification of an associahedron K_n of dimension $n-3$. More precisely, there is a natural smooth embedding of K_n inside $\mathcal{M}^{\delta}_{0,n}$ which is compatible with these stratifications (the shaded pentagon in Figure 4). This is the same as Devadoss's realization of the associahedron [9, Definition 3.2.1]. In that sense, Brown's moduli spaces $\mathcal{M}^{\delta}_{0,n}$ are algebro-geometric analogs of associahedra.

The dihedral decomposition morphisms

$$\Delta_{\mathfrak{d}} \colon \mathcal{G}rav(S,\delta) \to \mathcal{G}rav(\mathfrak{d})$$

may be computed as (signed) residues of logarithmic forms on $(\mathcal{M}_{0,S}^{\delta}, \partial\mathcal{M}_{0,S}^{\delta})$. This is particularly interesting since $\mathcal{M}_{0,S}^{\delta}$ is affine (Theorem 3.2) and we can thus use global logarithmic forms. We will give explicit formulas for these dihedral decomposition morphisms in Proposition 4.4.

3.3 The residue spectral sequence

In the global setting of Section 2.1.2, we prove the existence of a *residue spectral sequence* which computes the cohomology of the ambient space $\bar{X}$ in terms of the cohomology of the strata $X(\mathfrak{s})$ and the residue morphisms. In the next paragraph, we will apply this spectral sequence to the dihedral gravity cooperad.

Proposition 3.6 *Let $\bar{X}$ be a smooth (not necessarily compact) complex algebraic variety and let $\partial\bar{X}$ be a normal crossing divisor inside $\bar{X}$, inducing a stratification*

$$\bar{X} = \bigsqcup_{\mathfrak{s}\in\mathsf{Strat}} X(\mathfrak{s}).$$

There exists a first quadrant spectral sequence in the category of mixed Hodge structures:

$$E_1^{p,q} = \bigoplus_{\mathfrak{s}\in\mathsf{Strat}_p} H^{q-p}(X(\mathfrak{s}))(-p) \Longrightarrow H^{p+q}(\bar{X}),$$

where the differential $d_1\colon E_1^{p,q} \to E_1^{p+1,q}$ is the sum of the residue morphisms (12)

$$\mathrm{Res}^{\mathfrak{s}}_{\mathfrak{s}'}\colon H^{q-p}(X(\mathfrak{s}))(-p) \to H^{q-p-1}(X(\mathfrak{s}'))(-p-1)$$

for $\mathfrak{s}\in\mathsf{Strat}_p$ and $\mathfrak{s}'\in\mathsf{Strat}_{p+1}$ such that $\mathfrak{s}\leqslant\mathfrak{s}'$.

Proof We first forget about mixed Hodge structures and prove the existence of the spectral sequence for the cohomology over $\mathbb{C}$. Let us denote by $i_{\mathfrak{s}}\colon \bar{X}(\mathfrak{s}) \hookrightarrow \bar{X}$ the natural closed immersions. Let us write

$$\mathcal{K}^{p,q} = \bigoplus_{\mathfrak{s}\in\mathsf{Strat}_p} (i_{\mathfrak{s}})_*\Omega^{q-p}_{\bar{X}(\mathfrak{s})}(\log\partial\bar{X}(\mathfrak{s})).$$

We give the collection of the $\mathcal{K}^{p,q}$ the structure of a double complex of sheaves on $\bar{X}$. The horizontal differential $d'\colon \mathcal{K}^{p,q}\to\mathcal{K}^{p+1,q}$ is induced by the residues

$$(i_{\mathfrak{s}})_*\Omega^{q-p}_{\bar{X}(\mathfrak{s})}(\log\partial\bar{X}(\mathfrak{s})) \to (i_{\mathfrak{s}'})_*\Omega^{q-p-1}_{\bar{X}(\mathfrak{s}')}(\log\partial\bar{X}(\mathfrak{s}'))$$

for $\mathfrak{s}\in\mathsf{Strat}_p$ and $\mathfrak{s}'\in\mathsf{Strat}_{p+1}$ such that $\mathfrak{s}\leqslant\mathfrak{s}'$. The vertical differential $d''\colon \mathcal{K}^{p,q}\to\mathcal{K}^{p,q+1}$ is induced by the exterior derivative on differential forms. One checks that we have $d' \circ d' = 0$, $d'' \circ d'' = 0$ and $d' \circ d'' + d'' \circ d' = 0$. We denote the corresponding total complex by

$$\mathcal{K}^n = \bigoplus_{p+q=n} \mathcal{K}^{p,q}.$$

Using local coordinates on $\bar{X}$, it is easy to check that we have a long exact sequence

$$0 \to \Omega^\bullet_{\bar{X}} \to \mathcal{K}^{0,\bullet} \to \mathcal{K}^{1,\bullet} \to \mathcal{K}^{2,\bullet} \to \cdots, \tag{16}$$

which induces a quasi-isomorphism $\Omega^\bullet_{\bar{X}} \simeq \mathcal{K}^\bullet$. The holomorphic Poincaré lemma implies that we have a quasi-isomorphism $\mathbb{C}_{\bar{X}} \simeq \Omega^\bullet_{\bar{X}}$; hence we get an isomorphism

$$H^q(\bar{X}, \mathbb{C}) \cong \mathbb{H}^q(\bar{X}, \mathcal{K}^\bullet).$$

Now, the hypercohomology spectral sequence for the double complex $\mathcal{K}^{\bullet,\bullet}$ filtered by the columns is exactly

$$E_1^{p,q} = \bigoplus_{\mathfrak{s} \in \mathsf{Strat}_p} \mathbb{H}^q\big(\bar{X}(\mathfrak{s}), \Omega^{\bullet - p}_{\bar{X}(\mathfrak{s})}(\log \partial \bar{X}(\mathfrak{s}))\big) \Longrightarrow H^{p+q}(\bar{X}, \mathbb{C}).$$

Taking into account the isomorphisms $\mathbb{H}^q\big(\bar{X}(\mathfrak{s}), \Omega^{\bullet - p}_{\bar{X}(\mathfrak{s})}(\log \partial \bar{X}(\mathfrak{s}))\big) \simeq H^{q-p}(X(\mathfrak{s}), \mathbb{C})$, one gets the desired spectral sequence.

In order to prove that this spectral sequence is defined over $\mathbb{Q}$, it is convenient to work in the category of perverse sheaves. We let $u_\mathfrak{s}\colon X(\mathfrak{s}) \hookrightarrow \bar{X}$ denote the natural locally closed immersions. We replace (16) by the following long exact sequence in the category of perverse sheaves on $\bar{X}$, where d denotes the complex dimension of $\bar{X}$:

$$0 \to \mathbb{Q}_{\bar{X}}[d] \to u_* \mathbb{Q}_X[d] \to \bigoplus_{\mathfrak{s} \in \mathsf{Strat}_1} (u_\mathfrak{s})_* \mathbb{Q}_{X(\mathfrak{s})}[d-1] \to \bigoplus_{\mathfrak{s} \in \mathsf{Strat}_2} (u_\mathfrak{s})_* \mathbb{Q}_{X(\mathfrak{s})}[d-2] \to \cdots.$$

Taking the hypercohomology spectral sequence and shifting all the degrees by d gives the result.

The proof via perverse sheaves can be copied in the category of mixed Hodge modules [25] (see [23, Section 14]), which proves the compatibility with mixed Hodge structures. □

3.4 Purity and freeness

We start with a classical theorem on the cohomology of the moduli spaces $\mathcal{M}_{0,S}$.

Theorem 3.7 *For every integer k and every set S, the cohomology group $H^k(\mathcal{M}_{0,S})$ is pure Tate of weight $2k$.*

Proof Since the moduli space $\mathcal{M}_{0,S}$ is a complement of a union of hyperplanes in the affine space $\mathbb{C}^{n-3}$ by (6), this is a consequence of a general result on complements of hyperplane arrangements [21; 27; 19]. See also Getzler's proof [14, Lemma 3.12] which only uses Arnol'd's result [3]. $\square$

The residue spectral sequence of the previous paragraph now allows us to compute the cohomology of Brown's moduli spaces $\mathcal{M}^{\delta}_{0,S}$ in term of the cohomology of the spaces $\mathcal{M}_{0,S}$.

Proposition 3.8 *There exists a first quadrant spectral sequence in the category of mixed Hodge structures:*

$$E_1^{p,q} = \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_p(S,\delta)} H^{q-p}(\mathcal{M}(\mathfrak{d}))(-p) \Longrightarrow H^{p+q}(\mathcal{M}^{\delta}_{0,S}), \tag{17}$$

where the differential $d_1\colon E_1^{p,q} \to E_1^{p+1,q}$ *is the sum of the residue morphisms*

$$\mathrm{Res}^{\mathfrak{d}}_{\mathfrak{d}'}\colon H^{q-p}(\mathcal{M}(\mathfrak{d}))(-p) \to H^{q-p-1}(\mathcal{M}(\mathfrak{d}'))(-p-1)$$

for $\mathfrak{d}\in\mathsf{Diss}_p(S,\delta)$ *and* $\mathfrak{d}'\in\mathsf{Diss}_{p+1}(S,\delta)$ *such that* $\mathfrak{d}\leqslant\mathfrak{d}'$.

Proof This is a direct application of Proposition 3.6 to the case $\bar{X} = \mathcal{M}^{\delta}_{0,S}$ with the stratification (15). $\square$

The q^{th} row of the first page E_1 of the spectral sequence (17) looks like

$$\begin{aligned} 0 \to H^q(\mathcal{M}_{0,S}) \to \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_1(S,\delta)} H^{q-1}(\mathcal{M}(\mathfrak{d}))(-1) & \\ \to \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_2(S,\delta)} H^{q-2}(\mathcal{M}(\mathfrak{d}))(-2) \to \cdots . & \end{aligned} \tag{18}$$

Proposition 3.9 *The direct sum of the rows* $E_1^{\bullet,q}$ *of the first page of the spectral sequence (17) is, up to a Tate twist* (-1), *the dihedral cobar construction of the* (*desuspension of the*) *dihedral gravity cooperad.*

Proof After twisting by (-1), the direct sum of the complexes (18) can be written as

$$0 \to s^{-1}\mathcal{C}(S,\delta) \to \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_1(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_2(S,\delta)} s^{-1}\mathcal{C}(\mathfrak{d}) \to \cdots ,$$

where the arrows are (signed) infinitesimal decomposition morphisms. We leave it to the reader to check that the sign conventions are consistent. $\square$

We now turn to the degeneration of this spectral sequence.

Proposition 3.10 *The spectral sequence (17) degenerates at the second page E_2; that is, $E_\infty = E_2$.*

Proof As a consequence of Theorem 3.7 and the Künneth formula, $H^{q-p}(\mathcal{M}(\mathfrak{d}))$ is pure Tate of weight $2(q-p)$ for every dissection $\mathfrak{d} \in \mathsf{Diss}_p(S,\delta)$, and hence $H^{q-p}(\mathcal{M}(\mathfrak{d}))(-p)$ is pure Tate of weight $2(q-p)+2p = 2q$. The differential $d_r\colon E_r^{p,q} \to E_r^{p+r,q-r+1}$ thus maps a pure Hodge structure of weight $2q$ to a pure Hodge structure of weight $2(q-r+1)$, and is zero for $r \geqslant 2$ by Lemma 2.4. □

In the next proposition, we prove the equivalence between two statements: a geometric statement (i), namely the purity of the Hodge structure on the cohomology of Brown's moduli spaces, and an algebraic statement (ii), namely the freeness of the dihedral gravity cooperad. In the next section, we will prove the algebraic statement (ii) and derive the geometric statement (i). We nevertheless state this proposition as an equivalence to convince the reader that the mathematical content of the two statements is *essentially the same*.

Theorem 3.11 *The following statements are equivalent:*

(i) *for every integer k and every dihedral set (S,δ), the cohomology group $H^k(\mathcal{M}_{0,S}^{\delta})$ is pure Tate of weight $2k$;*

(ii) *the dihedral gravity cooperad is cofree.*

When they are true, there is a (noncanonical) isomorphism between the dihedral gravity cooperad and the cofree dihedral cooperad on the dihedral module:

$$(S,\delta) \mapsto \det(S) \otimes H^{\bullet+n-3}(\mathcal{M}_{0,S}^{\delta})(-1).$$

Proof Let us denote by A the filtration on the cohomology of $\mathcal{M}_{0,S}^{\delta}$ that is induced by the spectral sequence (17). It is a filtration by mixed Hodge substructures. By Proposition 3.10, we get at the second page:

$$E_2^{p,q} = \mathrm{gr}_A^p\, H^{p+q}(\mathcal{M}_{0,S}^{\delta}).$$

By the proof of Proposition 3.10, the space $E_2^{p,q}$ is pure Tate of weight $2q$. Thus, (i) is equivalent to the fact that for every (S,δ), the spectral sequence (17) satisfies $E_2^{p,q} = 0$ for $p > 0$. This is the same as requesting that each row $E_1^{\bullet,q}$ is exact except possibly at $\bullet = 0$. According to Proposition 3.9 and Proposition 1.15, this is equivalent to (ii), and we have proved the equivalence between statements (i) and (ii). Assuming them, we see that $H^k(\mathcal{M}_{0,S}^{\delta}) = E_2^{0,k}$ is the kernel of the map

$$H^k(\mathcal{M}_{0,S}) \xrightarrow{\oplus \Delta_{\mathfrak{d}}} \bigoplus_{\mathfrak{d} \in \mathsf{Diss}_1(S,\delta)} H^{k-1}(\mathcal{M}(\mathfrak{d}))(-1),$$

hence the result about the cogenerators of the dihedral gravity cooperad, after a degree shift and an operadic suspension. □

Remark 3.12 We can also apply the residue spectral sequence to the case $\bar{X} = \overline{\mathcal{M}}_{0,S}$, with the stratification (7). We then get a spectral sequence in the category of mixed Hodge structures:

$$E_1^{p,q} = \bigoplus_{\mathfrak{t} \in \mathsf{Tree}_p(S)} H^{q-p}(\mathcal{M}(\mathfrak{t}))(-p) \Longrightarrow H^{p+q}(\overline{\mathcal{M}}_{0,S}),$$

which degenerates at the second page E_2. It is a classical fact that the odd cohomology groups of $\overline{\mathcal{M}}_{0,S}$ are zero, and that for every k, $H^{2k}(\overline{\mathcal{M}}_{0,S})$ is pure Tate of weight $2k$. Thus, the degeneration of the spectral sequence gives rise to a long exact sequence

$$0 \to H^k(\mathcal{M}_{0,S}) \to \bigoplus_{\mathfrak{t} \in \mathsf{Tree}_1(S)} H^{k-1}(\mathcal{M}(\mathfrak{t}))(-1) \to \cdots$$
$$\to \bigoplus_{\mathfrak{t} \in \mathsf{Tree}_k(S)} H^0(\mathcal{M}(\mathfrak{t}))(-k) \to H^{2k}(\overline{\mathcal{M}}_{0,S}) \to 0.$$

After dualizing and performing an operadic suspension, this long exact sequence gives a quasi-isomorphism from the cyclic hypercommutative operad $S \mapsto H_\bullet(\overline{\mathcal{M}}_{0,S})$ to the cyclic bar construction of the cyclic gravity operad. Under the bar-cobar adjunction, this corresponds to Getzler's quasi-isomorphism [14, Theorem 4.6], which proves the Koszul duality between the cyclic hypercommutative operad and the cyclic gravity operad.

4 The dihedral gravity cooperad is cofree

We prove that the dihedral gravity cooperad is cofree by using explicit formulas describing the cohomology of the moduli spaces $\mathcal{M}_{0,S}$. The main point consist in showing that the filtration given by residual chords is the coradical filtration of the dihedral gravity cooperad. We then derive geometric consequences for Brown's moduli spaces $\mathcal{M}_{0,S}^{\delta}$ and a new proof of a theorem of Salvatore–Tauraso.

4.1 Conventions

In this section, we will work with explicit formulas for the decomposition morphisms in the dihedral gravity cooperad. For reasons of signs, it is easier to work with its desuspension $\mathcal{C}$, whose underlying Dih–module is given by

$$\mathcal{C}(S,\delta) = H^{\bullet-1}(\mathcal{M}_{0,S})(-1).$$

We use the notation $\mathcal{C}(S,\delta)$ instead of $\mathcal{C}(S)$ because we will use a spanning set and a filtration for this space that depend on the choice of a dihedral structure.

For $\mathfrak{d} \in \mathsf{Diss}_k(S,\delta)$ a dissection of cardinality r, we will always choose an ordering $P(\mathfrak{d}) = \{p_0, \dots, p_k\}$ and write $E_i := E(p_i)$ for the set of edges of the subpolygons p_i, $\delta_i := \delta(p_i)$ for the induced dihedral orders. The ordering of $P(\mathfrak{d})$ gives a trivialization $\det(P(\mathfrak{d})) \simeq \mathbb{Q}$ and hence we can simply write

$$\Delta_{\mathfrak{d}} \colon \mathcal{C}(S,\delta) \to \mathcal{C}(\mathfrak{d}) = \mathcal{C}(E_0,\delta_0) \otimes \cdots \otimes \mathcal{C}(E_k,\delta_k)$$

for the dihedral decompositions (14).

4.2 Cohomology of the moduli spaces $\mathcal{M}_{0,S}$

Let S be a finite set of cardinality $n \geqslant 3$ and let δ be a dihedral structure on S. We first recall Brown's presentation of the cohomology algebra of the moduli space $\mathcal{M}_{0,S}$, which is well suited for computing residues on $\mathcal{M}^{\delta}_{0,S}$. For any chord c of (S,δ), there exists a global holomorphic function $u_c \in \mathcal{O}(\mathcal{M}^{\delta}_{0,S})$ such that the divisor $\mathcal{M}^{\delta}(\{c\})$ is defined by the vanishing of u_c:

$$\mathcal{M}^{\delta}(\{c\}) = \{u_c = 0\}.$$

We then define the following closed logarithmic differential 1–form on $\mathcal{M}_{0,S}$:

$$\omega_c := \frac{1}{2\pi i}\frac{du_c}{u_c}.$$

We denote by the same symbol ω_c its class in $H^1(\mathcal{M}_{0,S})$.

Proposition 4.1 [5, Proposition 6.2] *The cohomology algebra $H^\bullet(\mathcal{M}_{0,S})$ is generated by the classes ω_c. In other words, $\mathcal{C}(S,\delta)$ is spanned by monomials $\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}$ for some chords $c_1, \dots, c_k$ of (S,δ).*

Note that every differential form $\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}$ is a logarithmic form on $(\mathcal{M}^{\delta}_{0,S}, \partial\mathcal{M}^{\delta}_{0,S})$.

Remark 4.2 It is convenient to represent a monomial $\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}$, up to a sign, by the picture of the set of chords $\{c_1, \dots, c_k\}$, as in Figure 5, where the chords are pictured in dashed lines.

Remark 4.3 The ideal of relations between the classes ω_c in $H^\bullet(\mathcal{M}_{0,S})$ can be described in pure combinatorial terms with sets of chords that cross completely; see [5, Proposition 6.2]. Surprisingly enough, this will not play any role in the sequel.

The decomposition morphisms of the dihedral gravity cooperad are easily computed in terms of the symbols ω_c. They are completely determined by the infinitesimal ones which correspond to dissections made up of one chord.

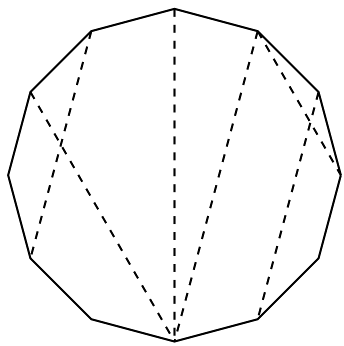


Figure 5: A monomial (up to a sign) in $H^6(\mathcal{M}_{0,10})$

Proposition 4.4 *Let c be a chord which dissects (S,δ) into two polygons p_0 and p_1. The corresponding dihedral decomposition morphism*

$$\Delta_{\{c\}}\colon \mathcal{C}(S,\delta) \to \mathcal{C}(E_0,\delta_0) \otimes \mathcal{C}(E_1,\delta_1)$$

is given by

(1) $\Delta_{\{c\}}(\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}) = 0$ *if* $c \notin \{c_1,\ldots,c_k\}$*;*

(2) $\Delta_{\{c\}}(\omega_c \wedge \omega_{c_1} \wedge \cdots \wedge \omega_{c_k}) = 0$ *if c crosses some chord c_i for $i = 1,\ldots,k$;*

(3) $\Delta_{\{c\}}(X_0 \wedge \omega_c \wedge X_1) = X_0 \otimes X_1$ *if X_i is a monomial formed with chords in p_i, $i = 0, 1$.*

Proof (1) This is because the differential form $\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}$ has no pole along $\mathcal{M}^\delta(\{c\})$ if $c \notin \{c_1,\ldots,c_k\}$.

(2) By definition of the residue morphisms, $\Delta_{\{c\}}(\omega_c \wedge \omega_{c_1} \wedge \cdots \wedge \omega_{c_k})$ is, up to a sign, the restriction of the differential form $\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}$ on $\mathcal{M}^\delta(\{c\})$. If c crosses some chord c_i for $i = 1,\ldots,k$, then the proof of [5, Lemma 2.6] implies that ω_{c_i} is zero when restricted to $\mathcal{M}^\delta(\{c\})$, hence the result.

(3) Let us denote by $a-1$ and $b-1$ the respective degrees of X_0 and X_1, so that they respectively live in degree a and b in $\mathcal{C}$. Then we get $X_0 \wedge \omega_c \wedge X_1 = (-1)^{a-1}\omega_c \wedge X_0 \wedge X_1$, whose residue on $\mathcal{M}^\delta(\{c\})$ is the restriction of $(-1)^{a-1}X_0 \wedge X_1$ on $\mathcal{M}^\delta(\{c\})$. Note that the sign $(-1)^{a-1}$ is canceled by the Koszul sign in the definition (13) of $\Delta_{\{c\}}$. By the proof of [5, Lemma 2.6], the pullback morphism $\mathcal{O}(\mathcal{M}^\delta_{0,S}) \to \mathcal{O}(\mathcal{M}^{\delta_0}_{0,E_0}) \otimes \mathcal{O}(\mathcal{M}^{\delta_1}_{0,E_1})$ is given by $u_{c_0} \mapsto u_{c_0} \otimes 1$ and $u_{c_1} \mapsto 1 \otimes u_{c_1}$ for c_i a chord in p_i for $i = 0, 1$. The result follows. □

Remark 4.5 The formula of Proposition 4.4 (3), is easy to represent pictorially: if c is a chord that is not crossed by any other, applying $\Delta_{\{c\}}$ has the effect of cutting the polygon along c into two parts; see Figure 6.

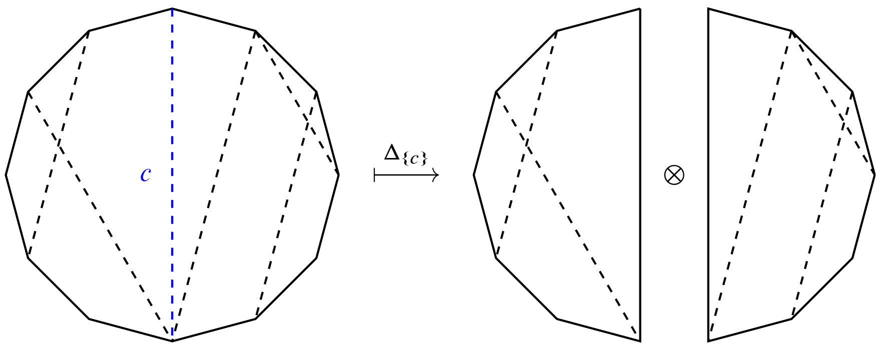


Figure 6: The dihedral decomposition $\Delta_{\{c\}}\colon \mathcal{C}(10,\delta) \to \mathcal{C}(6,\delta) \otimes \mathcal{C}(6,\delta)$ applied to a monomial

4.3 The residual filtration

Definition 4.6 (residual chord) Let $\{c_1,\dots,c_k\}$ be a set of chords of a polygon (S,δ). We say that c_i is a *residual chord* in $\{c_1,\dots,c_k\}$ if c_i is not crossed by any c_j for $j \neq i$.

Definition 4.7 (residual filtration) For every integer r, we denote by

$$R_r\mathcal{C}(S,\delta) \subset \mathcal{C}(S,\delta)$$

the subspace spanned by monomials $\omega_{c_1} \wedge \cdots \wedge \omega_{c_k}$ with at most r residual chords in $\{c_1,\dots,c_k\}$. This gives a finite filtration

$$0 = R_{-1}\mathcal{C}(S,\delta) \subset R_0\mathcal{C}(S,\delta) \subset R_1\mathcal{C}(S,\delta) \subset \cdots \subset \mathcal{C}(S,\delta)$$

called the *residual filtration*.

Lemma 4.8 *For a dissection $\mathfrak{d} \in \mathrm{Diss}_k(S,\delta)$ of (S,δ) of cardinality k, the dihedral decomposition*

$$\Delta_{\mathfrak{d}}\colon \mathcal{C}(S,\delta) \to \mathcal{C}(\mathfrak{d}) = \mathcal{C}(E_0,\delta_0) \otimes \cdots \otimes \mathcal{C}(E_k,\delta_k)$$

sends $R_r\mathcal{C}(S,\delta)$ to $R_{r-k}\mathcal{C}(\mathfrak{d})$.

Proof Since any decomposition map can be obtained by iterating infinitesimal decomposition maps, it is enough to do the case $k=1$, which follows from Proposition 4.4: applying $\Delta_{\{c\}}$ to a monomial either gives zero or erases a residual chord from the monomial. □

Example 4.9 In Figure 6, the left-hand side lives in $R_2\mathcal{C}(10,\delta)$ and the right-hand side lives in $R_0\mathcal{C}(6,\delta) \otimes R_1\mathcal{C}(6,\delta)$.

Theorem 4.10 *For every integer r and every dihedral set (S,δ), the morphism*

$$\Phi\colon \mathrm{gr}^R_r\, \mathcal{C}(S,\delta) \xrightarrow{\bigoplus \Delta_{\mathfrak{d}}} \bigoplus_{\mathfrak{d}\in \mathrm{Diss}_r(S,\delta)} R_0\mathcal{C}(\mathfrak{d})$$

is an isomorphism.

We postpone the proof of this theorem to Section 4.6, after we have introduced a technical tool.

4.4 The forgetful maps

Let S be a finite set and $S' \subset S$ be a subset. This inclusion gives rises to a forgetful morphism

$$f\colon \mathcal{M}_{0,S} \to \mathcal{M}_{0,S'}$$

and hence a pullback in cohomology

$$f^*\colon H^\bullet(\mathcal{M}_{0,S'}) \to H^\bullet(\mathcal{M}_{0,S}), \tag{19}$$

which is a map of graded algebras. Now suppose that we are given a dihedral structure δ on S and let δ' be the induced dihedral structure on S'. We view (S',δ') as the decorated polygon obtained by contracting the sides of (S,δ) that are not in S'. For a chord c of (S,δ) and a chord c' of (S',δ'), we write $c \rightsquigarrow c'$ if this contraction transforms c into c'.

Lemma 4.11 (1) *The pullback morphism f^* is given, for c' a chord of (S',δ'), by*

$$f^*(\omega_{c'}) = \sum_{c\rightsquigarrow c'} \omega_c.$$

(2) *The pullback morphism f^* is compatible with the residual filtration R.*

Proof (1) At the level of global functions, the pullback $\mathcal{O}(\mathcal{M}_{0,S'}) \to \mathcal{O}(\mathcal{M}_{0,S})$ is computed in [5, Lemma 2.9], and is given by

$$u_{c'} \mapsto \prod_{c\rightsquigarrow c'} u_c.$$

The result then follows from taking the logarithmic derivative.

(2) According to (1), the pullback of a monomial is given by

$$f^*(\omega_{c'_1}\wedge\cdots\wedge\omega_{c'_k}) = \sum_{\{c_i\rightsquigarrow c'_i\}} \omega_{c_1}\wedge\cdots\wedge\omega_{c_k}.$$

By construction, every set $\{c_1,\dots,c_k\}$ contains at most as many residual chords as $\{c'_1,\dots,c'_k\}$, hence the result. □

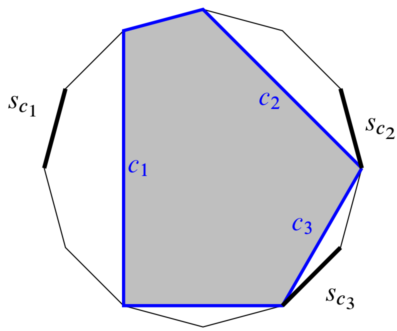


Figure 7: An inscribed polygon and a possible choice of matching sides

4.5 A technical lemma

Let us fix a polygon (S,δ). Let (E,δ_E) be an inscribed polygon inside (S,δ), that is, a polygon whose sides are either sides of (S,δ) or chords of (S,δ); see Figure 7. We let $E_{\mathrm{sides}} \subset E$ and $E_{\mathrm{chords}} \subset E$ denote the set of sides of (E,δ_E) which are respectively sides of (S,δ) and chords of (S,δ). In such a situation, we have a partition

$$S \setminus E_{\mathrm{sides}} = \bigsqcup_{c\in E_{\mathrm{chords}}} S_c$$

into components S_c delimited by c, that are outside of the inscribed polygon (E,δ_E), and connected with respect to the dihedral order δ.

For every chord $c \in E_{\mathrm{chords}}$, let us choose a *matching side* $s_c \in S_c$, and write

$$S' := E_{\mathrm{sides}} \sqcup \{s_c,\ c \in E_{\mathrm{chords}}\} \subset S.$$

We let δ' be the dihedral structure on S' induced by δ. Identifying a chord c and the matching side s_c gives rise to natural dihedral isomorphism $(E,\delta_E) \cong (S',\delta')$.

Example 4.12 In Figure 7, the inscribed polygon is shaded with $E_{\mathrm{chords}} = \{c_1,c_2,c_3\}$ and a possible choice of matching sides s_{c_1}, s_{c_2}, s_{c_3}.

The construction of the previous paragraph gives rise to a pullback morphism (19) that we denote by

$$\psi\colon \mathcal{C}(E,\delta_E) \cong \mathcal{C}(S',\delta') \to \mathcal{C}(S,\delta). \tag{20}$$

Lemma 4.13 *Let $X \in \mathcal{C}(E,\delta_E)$ be a monomial formed with chords of (E,δ_E), and let us denote by the same letter X the corresponding monomial viewed in $\mathcal{C}(S,\delta)$. Then $\psi(X)-X$ can be written as a sum of monomials $\omega_{c_1}\wedge\dots\wedge\omega_{c_k}$ for which some chord c_i crosses a chord in E_{chords}.*

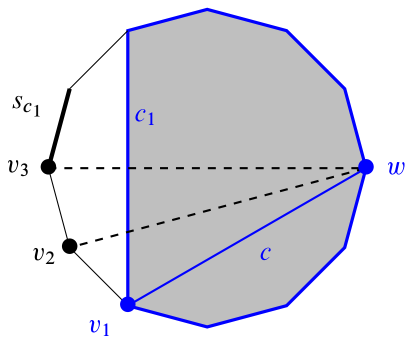


Figure 8: Illustration of the proof of Lemma 4.13

Proof It is enough to do the proof for a monomial $X = \omega_c$. We do the proof in the case where E_{chords} only contains one element c_1 corresponding to a side $s_{c_1} \in S$, the general case being similar. The formula for $\psi(\omega_c)$ is given in Lemma 4.11. If c and c_1 do not have a vertex in common, then $\psi(\omega_c) = \omega_c$. Else, let us denote by v_1 the common vertex of c and w the other vertex. We use the notation $c = v_1 w$. We then have

$$\psi(\omega_c) = \sum_v \omega_{vw},$$

where the sum ranges over the vertices $v \in S_{c_1}$ that are between v_1 and the first vertex of s_{c_1}. For such vertices v, the chord vw crosses c_1 except if $v = v_1$. The claim follows. □

Example 4.14 Figure 8 illustrates the proof of Lemma 4.13: the inscribed polygon (E, δ_E) is shaded. We have $\psi(\omega_{v_1 w}) = \omega_{v_1 w} + \omega_{v_2 w} + \omega_{v_3 w}$.

4.6 Proof of the main result

We now have all the tools to prove Theorem 4.10.

Proof of Theorem 4.10 To prove this theorem, we will construct the inverse morphism Ψ. To this aim, let us make some ordering conventions to make the signs explicit. For a dissection $\mathfrak{d} \in \mathsf{Diss}_r(S, \delta)$, we will choose compatible orderings

$$\mathfrak{d} = \{c_1, \dots, c_r\} \quad \text{and} \quad P(\mathfrak{d}) = \{p_0, \dots, p_r\} \tag{21}$$

that obey the following constraint. Let $\tilde{\mathfrak{t}}$ be the tree obtained by removing the leaves (external vertices) of the tree $\mathfrak{t}$ corresponding to $\mathfrak{d}$. The chords c_i label the edges of $\tilde{\mathfrak{t}}$, and the polygons p_i label the vertices of $\tilde{\mathfrak{t}}$. We choose the orderings (21) such that for every $j = 1, \dots, r-1$, deleting the edges labeled by c_i for $i = 1, \dots, j$ only disconnects the vertices p_i for $i = 0, \dots, j-1$.

An element of $\mathrm{gr}_r^R\,\mathcal{C}(S,\delta)$ can be represented as a sum of elements

$$X_0 \wedge \omega_{c_1} \wedge X_1 \wedge \omega_{c_2} \wedge \cdots \wedge \omega_{c_r} \wedge X_r,$$

for some dissection $\mathfrak{d} = \{c_1, \ldots, c_r\} \in \mathsf{Diss}_r(S,\delta)$, with $X_i \in R_0\mathcal{C}(E_i,\delta_i)$. According to the constraint we put on the orderings (21), the image of such an element by Φ is

$$\Delta_{\mathfrak{d}}(X_0 \wedge \omega_{c_1} \wedge X_1 \wedge \omega_{c_2} \wedge \cdots \wedge \omega_{c_r} \wedge X_r) = X_0 \otimes \cdots \otimes X_r \tag{22}$$

by repeated applications of Proposition 4.4.

For every $i = 0, \ldots, r$, we let

$$\psi_i \colon \mathcal{C}(E_i,\delta_i) \to \mathcal{C}(S,\delta)$$

denote the pullback map (20) defined in the previous paragraph, corresponding to the inscribed polygon $p_i = (E_i,\delta_i)$ and any choice of matching sides s_c for $c \in (E_i)_{\text{chords}}$.

Let us recall that we have

$$R_0\mathcal{C}(\mathfrak{d}) = R_0\mathcal{C}(E_0,\delta_0) \otimes \cdots \otimes R_0\mathcal{C}(E_r,\delta_r).$$

We then define

$$\Psi_{\mathfrak{d}} \colon R_0\mathcal{C}(\mathfrak{d}) \to \mathrm{gr}_r^R\,\mathcal{C}(S,\delta)$$

by the formula

$$\Psi_{\mathfrak{d}}(X_0 \otimes \cdots \otimes X_r) := \psi_0(X_0) \wedge \omega_{c_1} \wedge \psi_1(X_1) \wedge \omega_{c_2} \wedge \cdots \wedge \omega_{c_r} \wedge \psi_r(X_r).$$

Let us first prove that $\Psi_{\mathfrak{d}}$ is well defined. According to Lemma 4.11, each map ψ_i sends $R_0\mathcal{C}(E_i,\delta_i)$ to $R_0\mathcal{C}(S,\delta)$; hence the term $\psi_0(X_0) \wedge \cdots \wedge \psi_r(X_r)$ is in $R_0\mathcal{C}(S,\delta)$. Since the cardinality of $\mathfrak{d}$ is r, multiplying by $\omega_{c_1} \wedge \cdots \wedge \omega_{c_r}$ gives an element of $R_r\mathcal{C}(S,\delta)$.

With the same abuse of notation as in Lemma 4.13, we claim that we have

$$\Psi_{\mathfrak{d}}(X_0 \otimes \cdots \otimes X_r) = X_0 \wedge \omega_{c_1} \wedge X_1 \wedge \omega_{c_2} \wedge \cdots \wedge \omega_{c_r} \wedge X_r \mod R_{r-1}\mathcal{C}(S,\delta). \tag{23}$$

We do the proof of this equality in the case $r = 1$ and $\mathfrak{d} = \{c\}$ a chord, the general case being similar and left to the reader. Let us choose monomials $X_0 \in R_0\mathcal{C}(E_0,\delta_0)$ and $X_1 \in R_0\mathcal{C}(E_1,\delta_1)$ with zero residual chord. We want to prove the equality

$$\Psi_{\{c\}}(X_0 \otimes X_1) = X_0 \wedge \omega_c \wedge X_1 \mod R_0\mathcal{C}(S,\delta).$$

According to Lemma 4.13, we may write

$$\psi_1(X_0) = X_0 + \sum_{i_0} X_0^{(i_0)} \quad \text{and} \quad \psi_1(X_1) = X_1 + \sum_{i_1} X_1^{(i_1)},$$

where each monomial $X_0^{(i_0)}$ and $X_1^{(i_1)}$ has zero residual chord and contains a symbol $\omega_{c'}$ with c' crossing c. We can then write the difference $\Psi_{\{c\}}(X_0\otimes X_1)-X_0\wedge\omega_c\wedge X_1$ as

$$\sum_{i_0} X_0^{(i_0)}\wedge\omega_c\wedge X_1+\sum_{i_1} X_0\wedge\omega_c\wedge X_1^{(i_1)}+\sum_{i_0,i_1} X_0^{(i_0)}\wedge\omega_c\wedge X_1^{(i_1)}.$$

All the monomials appearing in the above expression have zero residual chord, hence the result. Equations (22) and (23) imply that Ψ is the inverse for Φ. □

Theorem 4.15 *The dihedral gravity cooperad is cofree. More precisely, it is (non-canonically) isomorphic to the cofree dihedral cooperad on the dihedral module:*

$$(S,\delta)\mapsto \det(S)\otimes H^{\bullet+n-3}(\mathcal{M}_{0,S}^{\delta})(-1).$$

Proof It is a consequence of Proposition 1.18, using Lemma 4.8 and Theorem 4.10, which imply, after operadic suspension, the corresponding statements for the dihedral gravity cooperad. The last statement follows from the last statement of Theorem 3.11. □

Remark 4.16 In [10], Dotsenko built a general a criterion to prove the freeness of the nonsymmetric operad underlying an operad in terms of Gröbner bases [11]. It would be interesting to know whether this criterion can give an alternate proof of Theorem 4.15.

4.7 Consequences for Brown's moduli spaces

We gather here some consequences of Theorem 4.15 on the geometry of the moduli spaces $\mathcal{M}_{0,S}^{\delta}$.

Corollary 4.17 *For every integer k and every dihedral set (S,δ), the cohomology group $H^k(\mathcal{M}_{0,S}^{\delta})$ is pure Tate of weight $2k$.*

Proof This follows from Theorem 4.15 and Theorem 3.11. □

Corollary 4.18 *For every integer k and every dihedral set (S,δ), the natural map $H^k(\mathcal{M}_{0,S}^{\delta})\to H^k(\mathcal{M}_{0,S})$ is injective and fits into a long exact sequence*

$$0\to H^k(\mathcal{M}_{0,S}^{\delta})\to H^k(\mathcal{M}_{0,S}) \to \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_1(S,\delta)} H^{k-1}(\mathcal{M}(\mathfrak{d}))(-1)\to \bigoplus_{\mathfrak{d}\in\mathsf{Diss}_2(S,\delta)} H^{k-2}(\mathcal{M}(\mathfrak{d}))(-1)\to\cdots. \tag{24}$$

Proof By Theorem 4.15 and the proof of Theorem 3.11, we get an injective map $H^k(\mathcal{M}_{0,S}^{\delta})\to E_1^{0,k}=H^k(\mathcal{M}_{0,S})$. By the construction of the residue spectral sequence, this map is indeed the one induced in cohomology by the inclusion $\mathcal{M}_{0,S}\hookrightarrow\mathcal{M}_{0,S}^{\delta}$. □

We note that the image of the natural map $H^{\bullet-1}(\mathcal{M}^{\delta}_{0,S})(-1) \hookrightarrow \mathcal{C}(S,\delta)$ is exactly the subspace $R_0\mathcal{C}(S,\delta)$.

Let us recall that the Betti numbers of the spaces $\mathcal{M}_{0,n}$ are given by the Poincaré polynomials

$$\sum_{k=0}^{n-3} b_k(\mathcal{M}_{0,n})x^k = \prod_{j=2}^{n-2}(x-j).$$

By taking the Euler characteristic of the exact sequence (24), one may thus derive a formula for the Betti numbers of the spaces $\mathcal{M}^{\delta}_{0,n}$ as follows.

Corollary 4.19 [4] *The generating series*

$$f(x,t) = x - \sum_{n\geqslant 3}\left(\sum_{k=0}^{n-3}(-1)^k b_k(\mathcal{M}_{0,n})t^{n-3-k}\right)x^{n-1},$$

$$f^{\delta}(x,t) = x + \sum_{n\geqslant 3}\left(\sum_{k=0}^{n-3}(-1)^k b_k(\mathcal{M}^{\delta}_{0,n})t^{n-3-k}\right)x^{n-1}$$

are inverse one to another: $f(f^{\delta}(x,t),t) = f^{\delta}(f(x,t),t) = x$.

We note that in [4, Section 3], the injectivity statement of Corollary 4.18 is used but not proved.

Corollary 4.20 *For every dihedral set (S,δ), Brown's moduli space $\mathcal{M}^{\delta}_{0,S}$ is a formal topological space.*

Proof It is a consequence of Corollary 4.17 and [12, Theorem 2.5]. A more direct proof goes as follows. Recall that $\mathcal{M}^{\delta}_{0,S}$ is a smooth affine complex variety. We denote by $\Omega^{\bullet}(\mathcal{M}^{\delta}_{0,S})$ the complex of global holomorphic differential forms on $\mathcal{M}^{\delta}_{0,S}$, and by $\Omega^{\bullet}(\mathcal{M}^{\delta}_{0,S}, \log\partial\mathcal{M}^{\delta}_{0,S})$ the complex of global holomorphic logarithmic differential forms on $\mathcal{M}^{\delta}_{0,S}$ along $\partial\mathcal{M}^{\delta}_{0,S}$. Let us recall that the morphism $H^{\bullet}(\mathcal{M}_{0,S})\to\Omega^{\bullet}(\mathcal{M}^{\delta}_{0,S}, \log\partial\mathcal{M}^{\delta}_{0,S})$ which maps the class of ω_c to ω_c is well defined and is a quasi-isomorphism. We consider the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & H^{\bullet}(\mathcal{M}^{\delta}_{0,S}) \arrow[r] & H^{\bullet}(\mathcal{M}_{0,S}) \arrow[r] \arrow[d] & \bigoplus_{\mathfrak{d}\in\mathrm{Diss}(S,\delta)} H^{\bullet-1}(\mathcal{M}(\mathfrak{d}))(-1) \arrow[d] \\
0 \arrow[r] & \Omega^{\bullet}(\mathcal{M}^{\delta}_{0,S}) \arrow[r] & \Omega^{\bullet}(\mathcal{M}^{\delta}_{0,S}, \log\partial\mathcal{M}^{\delta}_{0,S}) \arrow[r] & \bigoplus_{\mathfrak{d}\in\mathrm{Diss}(S,\delta)} \Omega^{\bullet-1}(\mathcal{M}^{\delta}(\mathfrak{d}), \log\partial\mathcal{M}^{\delta}(\mathfrak{d}))
\end{tikzcd}$$

where all arrows are morphisms of cochain complexes and where the vertical arrows are quasi-isomorphisms. The first row is exact by Corollary 4.18; the exactness of the second row follows from the fact that a logarithmic differential form on $\mathcal{M}^{\delta}_{0,S}$ along $\partial\mathcal{M}^{\delta}_{0,S}$ is regular on $\mathcal{M}^{\delta}_{0,S}$ if and only if its residue along each $\mathcal{M}^{\delta}(\mathfrak{d})$ is zero. Completing the diagram gives the following quasi-isomorphism, hence the result:

$$H^\bullet(\mathcal{M}^{\delta}_{0,S}) \to \Omega^\bullet(\mathcal{M}^{\delta}_{0,S}). \qquad \square$$

4.8 The dihedral Lie operad is free

As a corollary of Theorem 4.15 and in view of Theorem 2.7, we get a geometric proof of a dihedral enhancement of the theorem of Salvatore and Tauraso about the nonsymmetric Lie operad [26].

Corollary 4.21 *The dihedral Lie operad is free. More precisely, it is* (*noncanonically*) *isomorphic to the free dihedral operad on the dihedral module:*

$$(S,\delta) \mapsto \det(S) \otimes H_{n-3}(\mathcal{M}^{\delta}_{0,S})(1).$$

Remark 4.22 The equality between the top Betti number of $\mathcal{M}^{\delta}_{0,n}$ and the number of generators of the nonsymmetric Lie operad in arity n in [26] was already noticed in [4].

References

[1] **J Alm**, *A universal A-infinity structure on Batalin–Vilkovisky algebras with multiple zeta value coefficients*, preprint (2015) arXiv

[2] **J Alm**, **D Petersen**, *Brown's dihedral moduli space and freedom of the gravity operad*, preprint (2015) arXiv To appear in Ann. Sci. Éc. Norm. Supér

[3] **V I Arnol'd**, *The cohomology ring of the group of dyed braids*, Mat. Zametki 5 (1969) 227–231 MR In Russian; translated in "Vladimir I Arnold — Collected Works" II, Springer (2014) 183–186

[4] **J Bergström**, **F Brown**, *Inversion of series and the cohomology of the moduli spaces* $\mathcal{M}^{\delta}_{0,n}$, from "Motives, quantum field theory, and pseudodifferential operators" (A Carey, D Ellwood, S Paycha, S Rosenberg, editors), Clay Math. Proc. 12, Amer. Math. Soc., Providence, RI (2010) 119–126 MR

[5] **F C S Brown**, *Multiple zeta values and periods of moduli spaces* $\overline{\mathfrak{M}}_{0,n}$, Ann. Sci. Éc. Norm. Supér. 42 (2009) 371–489 MR

[6] **P Deligne**, *Théorie de Hodge, I*, from "Actes du Congrès International des Mathématiciens", volume 1, Gauthier-Villars, Paris (1971) 425–430 MR

[7] **P Deligne**, *Théorie de Hodge, II*, Inst. Hautes Études Sci. Publ. Math. 40 (1971) 5–57 MR

[8] **P Deligne**, *Théorie de Hodge, III*, Inst. Hautes Études Sci. Publ. Math. 44 (1974) 5–77 MR

[9] **S L Devadoss**, *Tessellations of moduli spaces and the mosaic operad*, from "Homotopy invariant algebraic structures" (J-P Meyer, J Morava, W S Wilson, editors), Contemp. Math. 239, Amer. Math. Soc., Providence, RI (1999) 91–114 MR

[10] **V Dotsenko**, *Freeness theorems for operads via Gröbner bases*, from "Operads 2009" (J-L Loday, B Vallette, editors), Sémin. Congr. 26, Soc. Math. France, Paris (2013) 61–76 MR

[11] **V Dotsenko**, **A Khoroshkin**, *Gröbner bases for operads*, Duke Math. J. 153 (2010) 363–396 MR

[12] **C Dupont**, *Purity, formality, and arrangement complements*, Int. Math. Res. Not. 2016 (2016) 4132–4144 MR

[13] **E Getzler**, *Two-dimensional topological gravity and equivariant cohomology*, Comm. Math. Phys. 163 (1994) 473–489 MR

[14] **E Getzler**, *Operads and moduli spaces of genus 0 Riemann surfaces*, from "The moduli space of curves" (R Dijkgraaf, C Faber, G van der Geer, editors), Progr. Math. 129, Birkhäuser, Boston (1995) 199–230 MR

[15] **E Getzler**, **M M Kapranov**, *Cyclic operads and cyclic homology*, from "Geometry, topology, & physics" (S-T Yau, editor), Conf. Proc. Lecture Notes Geom. Topology IV, Int. Press, Cambridge, MA (1995) 167–201 MR

[16] **V Ginzburg**, **M Kapranov**, *Koszul duality for operads*, Duke Math. J. 76 (1994) 203–272 MR

[17] **A B Goncharov**, **Y I Manin**, *Multiple ζ–motives and moduli spaces $\overline{\mathcal{M}}_{0,n}$*, Compos. Math. 140 (2004) 1–14 MR

[18] **S Keel**, *Intersection theory of moduli space of stable n–pointed curves of genus zero*, Trans. Amer. Math. Soc. 330 (1992) 545–574 MR

[19] **M Kim**, *Weights in cohomology groups arising from hyperplane arrangements*, Proc. Amer. Math. Soc. 120 (1994) 697–703 MR

[20] **F F Knudsen**, *The projectivity of the moduli space of stable curves, II: The stacks $M_{g,n}$*, Math. Scand. 52 (1983) 161–199 MR

[21] **G I Lehrer**, *The l–adic cohomology of hyperplane complements*, Bull. London Math. Soc. 24 (1992) 76–82 MR

[22] **J-L Loday**, **B Vallette**, *Algebraic operads*, Grundl. Math. Wissen. 346, Springer (2012) MR

[23] **C A M Peters**, **J H M Steenbrink**, *Mixed Hodge structures*, Ergeb. Math. Grenzgeb. 52, Springer (2008) MR

[24] **D Quillen**, *Rational homotopy theory*, Ann. of Math. 90 (1969) 205–295 MR

[25] **M Saito**, *Modules de Hodge polarisables*, Publ. Res. Inst. Math. Sci. 24 (1988) 849–995 MR

[26] **P Salvatore**, **R Tauraso**, *The operad Lie is free*, J. Pure Appl. Algebra 213 (2009) 224–230 MR

[27] **B Z Shapiro**, *The mixed Hodge structure of the complement to an arbitrary arrangement of affine complex hyperplanes is pure*, Proc. Amer. Math. Soc. 117 (1993) 931–933 MR

Institut Montpelliérain Alexander Grothendieck, Université de Montpellier
Place Eugène Bataillon, 34090 Montpellier, France

Laboratoire Analyse, Géométrie et Applications, Université Paris 13, Sorbonne Paris Cité
CNRS, UMR 7539, 93430 Villetaneuse, France

clement.dupont@umontpellier.fr, vallette@math.univ-paris13.fr

Proposed: Peter Teichner
Seconded: Dan Abramovich, Frances Kirwan

Received: 12 October 2015
Revised: 26 September 2016

msp

Geometry & Topology 21 (2017) 2851–2896

On the second homology group of the Torelli subgroup of $\mathrm{Aut}(F_n)$

MATTHEW DAY
ANDREW PUTMAN

Let IA_n be the Torelli subgroup of $\mathrm{Aut}(F_n)$. We give an explicit finite set of generators for $\mathrm{H}_2(\mathrm{IA}_n)$ as a $\mathrm{GL}_n(\mathbb{Z})$–module. Corollaries include a version of surjective representation stability for $\mathrm{H}_2(\mathrm{IA}_n)$, the vanishing of the $\mathrm{GL}_n(\mathbb{Z})$–coinvariants of $\mathrm{H}_2(\mathrm{IA}_n)$, and the vanishing of the second rational homology group of the level ℓ congruence subgroup of $\mathrm{Aut}(F_n)$. Our generating set is derived from a new group presentation for IA_n which is infinite but which has a simple recursive form.

20E05, 20E36, 20F05, 20J06

1 Introduction

The *Torelli subgroup* of the automorphism group of a free group F_n on n letters, denoted by IA_n, is the kernel of the action of $\mathrm{Aut}(F_n)$ on $F_n^{\mathrm{ab}} \cong \mathbb{Z}^n$. The group of automorphisms of $\mathbb{Z}^n$ is $\mathrm{GL}_n(\mathbb{Z})$ and the resulting map $\mathrm{Aut}(F_n) \to \mathrm{GL}_n(\mathbb{Z})$ is easily seen to be surjective, so we have a short exact sequence

$$1 \to \mathrm{IA}_n \to \mathrm{Aut}(F_n) \to \mathrm{GL}_n(\mathbb{Z}) \to 1.$$

Though it has a large literature, the cohomology and combinatorial group theory of IA_n remain quite mysterious. Magnus [26] proved that IA_n is finitely generated, and thus that $\mathrm{H}_1(\mathrm{IA}_n)$ has finite rank. Krstić and McCool [24] later showed that IA_3 is not finitely presentable. This was improved by Bestvina, Bux and Margalit [4], who showed that $\mathrm{H}_2(\mathrm{IA}_3)$ has infinite rank. However, for $n \geq 4$ it is not known whether or not IA_n is finitely presentable, or whether or not $\mathrm{H}_2(\mathrm{IA}_n)$ has finite rank.

Representation-theoretic finiteness It seems to be very difficult to determine whether or not $\mathrm{H}_2(\mathrm{IA}_n)$ has finite rank, so it is natural to investigate weaker finiteness properties. Since inner automorphisms act trivially on homology, the conjugation action of $\mathrm{Aut}(F_n)$ on IA_n induces an action of $\mathrm{GL}_n(\mathbb{Z})$ on $\mathrm{H}_k(\mathrm{IA}_n)$. Church and Farb [12, Conjecture 6.7] conjectured that $\mathrm{H}_k(\mathrm{IA}_n)$ is finitely generated as a $\mathrm{GL}_n(\mathbb{Z})$–module. In other words, they conjectured that there exists a finite subset of $\mathrm{H}_k(\mathrm{IA}_n)$ whose $\mathrm{GL}_n(\mathbb{Z})$–orbit spans $\mathrm{H}_k(\mathrm{IA}_n)$. Our first main theorem verifies their conjecture for $k = 2$.

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2851

Theorem A (generators for $\mathrm{H}_2(\mathrm{IA}_n)$) *For all $n \geq 2$, there exists a finite subset of $\mathrm{H}_2(\mathrm{IA}_n)$ whose $\mathrm{GL}_n(\mathbb{Z})$–orbit spans $\mathrm{H}_2(\mathrm{IA}_n)$.*

Each element of our finite subset corresponds to a map of a surface into a classifying space for IA_n; the genera of these surfaces range from 1 to 3. Table 1 below lists our finite set of $\mathrm{GL}_n(\mathbb{Z})$–generators for $\mathrm{H}_2(\mathrm{IA}_n)$. This table expresses these generators using specific "commutator relators" in IA_n; see below for how to translate these into elements of $\mathrm{H}_2(\mathrm{IA}_n)$.

Remark 1.1 The special case $n = 3$ of Theorem A was proven in the unpublished thesis of Owen Baker [2]. His proof uses a "Jacobian" map on outer space and is quite different from our proof. It seems difficult to generalize his proof to higher n.

Surjective representation stability The generators for $\mathrm{H}_2(\mathrm{IA}_n)$ given in Theorem A are explicit enough that they can be used to perform a number of interesting calculations. The first verifies part of a conjecture of Church and Farb, which asserts that the homology groups of IA_n are "representation stable". We begin with some background. An increasing sequence

$$G_1 \subset G_2 \subset G_3 \subset \cdots$$

of groups is *homologically stable* if for all $k \geq 1$, the k^{th} homology group of G_n is independent of n for $n \gg 0$. Many sequences of groups are homologically stable; see Hatcher and Wahl [21] for a bibliography. In particular, Hatcher and Vogtmann [20] proved this for $\mathrm{Aut}(F_n)$. However, it is known that IA_n is *not* homologically stable; indeed, even $\mathrm{H}_1(\mathrm{IA}_n)$ does not stabilize (see below).

Church and Farb [12] introduced a new form of homological stability for groups like IA_n whose homology groups possess natural group actions. For IA_n, they conjectured that for all $k \geq 1$, there exists some $n_k \geq 1$ such that the following two properties hold for all $n \geq n_k$:

- **Injective stability** The map $\mathrm{H}_k(\mathrm{IA}_n) \to \mathrm{H}_k(\mathrm{IA}_{n+1})$ is injective.
- **Surjective representation stability** The map $\mathrm{H}_k(\mathrm{IA}_n) \to \mathrm{H}_k(\mathrm{IA}_{n+1})$ is surjective "up to the action of $\mathrm{GL}_{n+1}(\mathbb{Z})$"; more precisely, the $\mathrm{GL}_{n+1}(\mathbb{Z})$–orbit of its image spans $\mathrm{H}_k(\mathrm{IA}_{n+1})$.

Remark 1.2 In fact, they made this conjecture in [11] for the Torelli subgroup of the mapping class group; however, they have informed us that they also conjecture it for IA_n.

Our generators for $\mathrm{H}_2(\mathrm{IA}_n)$ are "the same in each dimension" starting at $n = 6$, so we are able to derive the following special case of Church and Farb's conjecture:

Theorem B (surjective representation stability for $\mathrm{H}_2(\mathrm{IA}_n)$) *The* $\mathrm{GL}_{n+1}(\mathbb{Z})$*–orbit of the image of the natural map* $\mathrm{H}_2(\mathrm{IA}_n) \to \mathrm{H}_2(\mathrm{IA}_{n+1})$ *spans* $\mathrm{H}_2(\mathrm{IA}_{n+1})$ *for* $n \geq 6$.

Remark 1.3 Boldsen and Hauge Dollerup [5] proved a theorem similar to Theorem B for the *rational* second homology group of the Torelli subgroup of the mapping class group. Their proof is different from ours; in particular, they were not able to prove an analogue of Theorem A. It seems hard to use their techniques to prove Theorem B. Similarly, our proof uses special properties of IA_n and does not work for the Torelli subgroup of the mapping class group.

Coinvariants Our tools do not allow us to easily distinguish different homology classes; indeed, for all we know our generators for $\mathrm{H}_2(\mathrm{IA}_n)$ might be redundant. This prevents us from proving injective stability for $\mathrm{H}_2(\mathrm{IA}_n)$. However, we still can prove some interesting vanishing results. If G is a group and M is a G–module, then the *coinvariants* of G acting on M, denoted by M_G, are the largest quotient of M on which G acts trivially. More precisely, $M_G = M/K$ with $K = \langle m - g\cdot m \mid g \in G,\ m \in M\rangle$. We then have the following.

Theorem C (vanishing coinvariants) *For* $n \geq 6$, *we have* $(\mathrm{H}_2(\mathrm{IA}_n))_{\mathrm{GL}_n(\mathbb{Z})} = 0$.

Remark 1.4 Church and Farb [12, Conjecture 6.5] conjectured that the $\mathrm{GL}_n(\mathbb{Z})$–invariants in $\mathrm{H}^k(\mathrm{IA}_n;\mathbb{Q})$ are 0. For $k = 1$, this follows from the known computation of $\mathrm{H}^1(\mathrm{IA}_n;\mathbb{Q})$; see below. Theorem C implies that this also holds for $k = 2$.

Linear congruence subgroups For $\ell \geq 2$, the *level* ℓ *congruence subgroup* of $\mathrm{Aut}(F_n)$, denoted by $\mathrm{Aut}(F_n,\ell)$, is the kernel of the natural map $\mathrm{Aut}(F_n) \to \mathrm{GL}_n(\mathbb{Z}/\ell)$; one should think of it as a "mod-ℓ" version of IA_n. It is natural to conjecture that for all $k \geq 1$, there exists some $n_k \geq 1$ such that $\mathrm{H}_k(\mathrm{Aut}(F_n,\ell);\mathbb{Q}) \cong \mathrm{H}_k(\mathrm{Aut}(F_n);\mathbb{Q})$ for $n \geq n_k$; an analogous theorem for congruence subgroups of $\mathrm{GL}_n(\mathbb{Z})$ is due to Borel [7]. Galatius [19] proved that $\mathrm{H}_k(\mathrm{Aut}(F_n);\mathbb{Q}) = 0$ for $n \gg 0$, so this conjecture really asserts that $\mathrm{H}_k(\mathrm{Aut}(F_n,\ell);\mathbb{Q}) = 0$ for $n \gg 0$. The case $k = 1$ of this is known. Indeed, Satoh [31] calculated the abelianization of $\mathrm{Aut}(F_n,\ell)$ for $n \geq 3$ and the answer consisted entirely of torsion, so $\mathrm{H}_1(\mathrm{Aut}(F_n,\ell);\mathbb{Q}) = 0$ for $n \geq 3$. Using Theorem A, we will prove the case $k = 2$.

Theorem D (second homology of congruence subgroups) *For* $\ell \geq 2$ *and* $n \geq 6$, *we have* $\mathrm{H}_2(\mathrm{Aut}(F_n,\ell);\mathbb{Q}) = 0$.

The key to our proof is that Theorem A allows us to show that the image of $\mathrm{H}_2(\mathrm{IA}_n;\mathbb{Q})$ in $\mathrm{H}_2(\mathrm{Aut}(F_n,\ell);\mathbb{Q})$ vanishes; this allows us to derive Theorem D using standard techniques.

Remark 1.5 The second author proved an analogue of Theorem D for congruence subgroups of the mapping class group in [30]. The techniques in [30] are different from those in the present paper and it seems difficult to prove Theorem D via those techniques.

Basic elements of Torelli We now wish to describe our generating set for $\mathrm{H}_2(\mathrm{IA}_n)$. This requires introducing some basic elements of IA_n. Let $\{x_1, \dots, x_n\}$ be a free basis for F_n. We then make the following definitions:

- For distinct $1 \le i, j \le n$, let $C_{x_i, x_j} \in \mathrm{IA}_n$ be defined via the formulas
$$C_{x_i,x_j}(x_i) = x_j x_i x_j^{-1} \qquad \text{and} \qquad C_{x_i,x_j}(x_\ell) = x_\ell \quad \text{if } \ell \neq i.$$
- For $\alpha, \beta, \gamma \in \{\pm 1\}$ and distinct $1 \le i, j, k \le n$, let $M_{x_i^\alpha, [x_j^\beta, x_k^\gamma]} \in \mathrm{IA}_n$ be defined via the formulas
$$M_{x_i^\alpha,[x_j^\beta,x_k^\gamma]}(x_i^\alpha) = [x_j^\beta, x_k^\gamma] x_i^\alpha \qquad \text{and} \qquad M_{x_i^\alpha,[x_j^\beta,x_k^\gamma]}(x_\ell) = x_\ell \quad \text{if } \ell \neq i.$$

Observe that by definition
$$M_{x_i^{-1},[x_j^\beta,x_k^\gamma]}(x_i^{-1}) = [x_j^\beta, x_k^\gamma] x_i^{-1} \quad \text{and} \quad M_{x_i^{-1},[x_j^\beta,x_k^\gamma]}(x_i) = x_i [x_j^\beta, x_k^\gamma]^{-1}.$$
We call C_{x_i,x_j} a *conjugation move* and $M_{x_i^\alpha,[x_j^\beta,x_k^\gamma]}$ a *commutator transvection.*

Surfaces in a classifying space: our generators A *commutator relator* in IA_n is a formula of the form $[a_1, b_1] \cdots [a_g, b_g] = 1$ with $a_i, b_i \in \mathrm{IA}_n$. Given such a commutator relator r, let Σ_g be a genus g surface. There is a continuous map $\zeta\colon \Sigma_g \to K(\mathrm{IA}_n, 1)$ that takes the standard basis for $\pi_1(\Sigma_g)$ to $a_1, b_1, \dots, a_g, b_g \in \mathrm{IA}_n$. We obtain an element $\mathfrak{h}_r = \zeta_*([\Sigma_g]) \in \mathrm{H}_2(\mathrm{IA}_n)$. With this notation, the generators for $\mathrm{H}_2(\mathrm{IA}_n)$ given by Theorem A are the elements $\mathfrak{h}_r$ where r is one of the relators in Table 1.

The Johnson homomorphism To motivate our proof of Theorem A, we first recall the computation of $\mathrm{H}_1(\mathrm{IA}_n)$, which is due independently to Farb [18], Kawazumi [23] and Cohen and Pakianathan [14]. The basic tool is the Johnson homomorphism [22], which was introduced in the context of the Torelli subgroup of the mapping class group (though it also appears in earlier work of Andreadakis [1]). See Satoh [32] for a survey of the IA_n version of it. The Johnson homomorphism is a homomorphism
$$\tau\colon \mathrm{IA}_n \to \mathrm{Hom}\big(\mathbb{Z}^n, \textstyle\bigwedge^2 \mathbb{Z}^n\big)$$
that arises from studying the action of IA_n on the second nilpotent truncation of F_n. It can be defined as follows. For $z \in F_n$, let $[z] \in \mathbb{Z}^n$ be the associated element of the abelianization of F_n. Consider $f \in \mathrm{IA}_n$. For $x \in F_n$, we have $f(x) \cdot x^{-1} \in [F_n, F_n]$.

(H1) $[C_{x_a,x_b}, C_{x_c,x_d}] = 1$, possibly with $b = d$.

(H2) $[M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}, M_{x_d^\delta,[x_e^\epsilon,x_f^\zeta]}] = 1$, possibly with $\{b,c\} \cap \{e,f\} \neq \varnothing$, or with $x_a^\alpha = x_d^{-\delta}$ as long as $x_a^\alpha \neq x_d^\delta$, $a \notin \{e,f\}$ and $d \notin \{b,c\}$.

(H3) $[C_{x_a,x_b}, M_{x_c^\gamma,[x_d^\delta,x_e^\epsilon]}] = 1$,
possibly with $b \in \{d,e\}$ if $c \notin \{a,b\}$ and $a \notin \{c,d,e\}$.

(H4) $[C_{x_c,x_b}^\beta C_{x_a,x_b}^\beta, C_{x_c,x_a}^\alpha] = 1$.

(H5) $[C_{x_a,x_c}^{-\gamma}, C_{x_a,x_d}^{-\delta}][C_{x_a,x_b}^{-\beta}, M_{x_b^\beta,[x_c^\gamma,x_d^\delta]}] = 1$.

(H6) $[M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}, M_{x_d^\delta,[x_a^\alpha,x_e^\epsilon]}][M_{x_d^\delta,[x_a^\alpha,x_e^\epsilon]}, M_{x_d^\delta,[x_c^\gamma,x_b^\beta]}] \times$
$[M_{x_d^\delta,[x_c^\gamma,x_b^\beta]}, C_{x_d,x_e}^{-\epsilon}] = 1$, possibly with $b = e$ or $c = e$.

(H7) $[M_{x_c^\gamma,[x_a^\alpha,x_d^\delta]}, C_{x_a,x_b}^\beta][C_{x_c,x_d}^{-\delta}, M_{x_c^\gamma,[x_a^\alpha,x_b^\beta]}][M_{x_c^\gamma,[x_a^\alpha,x_b^\beta]}, M_{x_c^\gamma,[x_a^\alpha,x_d^\delta]}] = 1$,
possibly with $b = d$.

(H8) $[M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}, C_{x_a,x_d}^\delta C_{x_b,x_d}^\delta C_{x_c,x_d}^\delta] = 1$.

(H9) $[C_{x_a,x_c}^\gamma C_{x_b,x_c}^\gamma, C_{x_a,x_b}^\beta C_{x_c,x_b}^\beta][M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}, C_{x_b,x_a}^\alpha C_{x_c,x_a}^\alpha] = 1$.

Table 1: The set of commutator relators whose associated elements of $\mathrm{H}_2(\mathrm{IA}_n)$ generate it as a $\mathrm{GL}_n(\mathbb{Z})$–module. Distinct letters represent distinct indices unless stated otherwise.

There is a natural surjection $\rho\colon [F_n, F_n] \to \bigwedge^2 \mathbb{Z}^n$ satisfying $\rho([a,b]) = [a] \wedge [b]$; the kernel of ρ is $[F_n,[F_n,F_n]]$. We can then define a map $\widetilde{\tau}_f\colon F_n \to \bigwedge^2 \mathbb{Z}^n$ via the formula $\widetilde{\tau}_f(x) = \rho(f(x)\cdot x^{-1})$. One can check that $\widetilde{\tau}_f$ is a homomorphism. It factors through a homomorphism $\tau_f\colon \mathbb{Z}^n \to \bigwedge^2 \mathbb{Z}^n$. We can then define $\tau\colon \mathrm{IA}_n \to \mathrm{Hom}\big(\mathbb{Z}^n, \bigwedge^2 \mathbb{Z}^n\big)$ via the formula $\tau(f) = \tau_f$. One can check that τ is a homomorphism.

Generators and their images Define

$$S_{\mathrm{MA}}(n) = \{C_{x_i,x_j} \mid 1 \le i,j \le n \text{ distinct}\} \cup \{M_{x_i,[x_j,x_k]} \mid 1 \le i,j,k \le n \text{ distinct}, \ j < k\}.$$

Magnus [26] proved that IA_n is generated by $S_{\mathrm{MA}}(n)$; see Day and Putman [16] and Bestvina, Bux and Margalit [4] for modern proofs. For distinct $1 \le i,j \le n$, the image $\tau(C_{x_i,x_j}) \in \mathrm{Hom}(\mathbb{Z}^n, \bigwedge^2 \mathbb{Z}^n)$ is the homomorphism defined via the formulas

$$[x_i] \mapsto [x_j] \wedge [x_i] \qquad \text{and} \qquad [x_\ell] \mapsto 0 \quad \text{if } \ell \neq i.$$

Similarly, for distinct $1 \le i,j,k \le n$ with $j < k$, the image $\tau(M_{x_i,[x_j,x_k]})$ in $\mathrm{Hom}\big(\mathbb{Z}^n, \bigwedge^2 \mathbb{Z}^n\big)$ is the homomorphism defined via the formulas

$$[x_i] \mapsto [x_j] \wedge [x_k] \qquad \text{and} \qquad [x_\ell] \mapsto 0 \quad \text{if } \ell \neq i.$$

The key observation is that these form a *basis* for $\mathrm{Hom}\big(\mathbb{Z}^n, \bigwedge^2 \mathbb{Z}^n\big)$.

The abelianization Let $F(S_{\mathrm{MA}}(n))$ be the free group on $S_{\mathrm{MA}}(n)$ and let $R_{\mathrm{MA}}(n) \subset F(S_{\mathrm{MA}}(n))$ be a set of relations for IA_n, so $\mathrm{IA}_n = \langle S_{\mathrm{MA}}(n) \mid R_{\mathrm{MA}}(n) \rangle$. Since τ takes $S_{\mathrm{MA}}(n)$ bijectively to a basis for the free abelian group $\mathrm{Hom}\big(\mathbb{Z}^n, \bigwedge^2 \mathbb{Z}^n\big)$, we must have $R_{\mathrm{MA}}(n) \subset [F(S_{\mathrm{MA}}(n)), F(S_{\mathrm{MA}}(n))]$. This immediately implies that $\mathrm{H}_1(\mathrm{IA}_n) \cong \mathrm{Hom}\big(\mathbb{Z}^n, \bigwedge^2 \mathbb{Z}^n\big)$.

Hopf's formula But even more is true. Recall that Hopf's formula (see Brown [10]) says that if G is a group with a presentation $G = \langle S \mid R \rangle$, then

$$\mathrm{H}_2(G) \cong \frac{\langle\!\langle R \rangle\!\rangle \cap [F(S), F(S)]}{[F(S), \langle\!\langle R \rangle\!\rangle]};$$

here $\langle\!\langle R \rangle\!\rangle$ is the normal closure of R. The intersection in the numerator of this is usually hard to calculate, so Hopf's formula is not often useful for computation. However, by what we have said it simplifies for IA_n to

$$\mathrm{H}_2(\mathrm{IA}_n) \cong \frac{\langle\!\langle R_{\mathrm{MA}}(n) \rangle\!\rangle}{[F(S_{\mathrm{MA}}(n)), \langle\!\langle R_{\mathrm{MA}}(n) \rangle\!\rangle]}. \tag{1}$$

This isomorphism is very concrete: an element $r \in \langle\!\langle R_{\mathrm{MA}}(n) \rangle\!\rangle$ is a commutator relator, and the associated element of $\mathrm{H}_2(\mathrm{IA}_n)$ is the homology class $\mathfrak{h}_r$ discussed above.

Summary and trouble For $r \in \langle\!\langle R_{\mathrm{MA}}(n) \rangle\!\rangle$ and $z \in F(S_{\mathrm{MA}}(n))$, the element $zrz^{-1}r^{-1}$ lies in the denominator of (1), $[F(S_{\mathrm{MA}}(n)), \langle\!\langle R_{\mathrm{MA}}(n) \rangle\!\rangle]$. Hence $\mathfrak{h}_{zrz^{-1}} = \mathfrak{h}_r$. It follows that $\mathrm{H}_2(\mathrm{IA}_n)$ is generated by the set $\{\mathfrak{h}_r \mid r \in R_{\mathrm{MA}}(n)\}$. In other words, to calculate generators for $\mathrm{H}_2(\mathrm{IA}_n)$, it is enough to find a presentation for IA_n with $S_{\mathrm{MA}}(n)$ as its generating set. However, this seems like a difficult problem (especially if, as we suspect, IA_n is not finitely presentable). Moreover, the $\mathrm{GL}_n(\mathbb{Z})$–action on $\mathrm{H}_2(\mathrm{IA}_n)$ has not yet appeared.

L–presentations To incorporate the $\mathrm{GL}_n(\mathbb{Z})$–action on $\mathrm{H}_2(\mathrm{IA}_n)$ into our presentation for IA_n, we use the notion of an L–presentation, which was introduced by Bartholdi [3] (we use a slight simplification of his definition). An *L–presentation* for a group G is a triple $\langle S \mid R^0 \mid E \rangle$, where S and R^0 and E are as follows:

- S is a generating set for G.
- $R^0 \subset F(S)$ is a set consisting of relations for G (not necessarily complete).
- E is a subset of $\mathrm{End}(F(S))$.

This data must satisfy the following condition: Let $M \subset \mathrm{End}(F(S))$ be the monoid generated by E. Define $R = \{f(r) \mid f \in M,\ r \in R^0\}$. Then we require $G = \langle S \mid R \rangle$. Each element of E descends to an element of $\mathrm{End}(G)$; we call the resulting subset $\widetilde{E} \subset \mathrm{End}(G)$ the *induced endomorphisms* of our L–presentation. We say that our L–presentation is *finite* if the sets S and R^0 and E are all finite.

In our examples, the induced endomorphisms of our L–presentations will actually be automorphisms. Thus in our context one should think of an L–presentation as a group presentation incorporating certain symmetries of a group. Here is an easy example:

Example Let $S = \{z_i \mid i \in \mathbb{Z}/p\}$ and $R^0 = \{z_0^2\}$. Let $\psi\colon F(S) \to F(S)$ be the homomorphism defined via the formula $\psi(z_i) = z_{i+1}$. Then $\langle S \mid R^0 \mid \{\psi\}\rangle$ is an L–presentation for the free product of p copies of $\mathbb{Z}/2$.

A finite L–presentation for Torelli The conjugation action of $\mathrm{Aut}(F_n)$ on IA_n gives an injection $\mathrm{Aut}(F_n) \hookrightarrow \mathrm{Aut}(\mathrm{IA}_n)$. If we could somehow construct a finite L–presentation $\langle S_{\mathrm{MA}}(n) \mid R^0_{\mathrm{MA}}(n) \mid E_{\mathrm{MA}}(n)\rangle$ for IA_n whose set of induced endomorphisms generated

$$\mathrm{Aut}(F_n) \subset \mathrm{Aut}(\mathrm{IA}_n) \subset \mathrm{End}(\mathrm{IA}_n),$$

then Theorem A would immediately follow. Indeed, since the $\mathrm{GL}_n(\mathbb{Z})$–action on $\mathrm{H}_2(\mathrm{IA}_n)$ is induced by the conjugation action of $\mathrm{Aut}(F_n)$ on IA_n, it would follow that the $\mathrm{GL}_n(\mathbb{Z})$–orbit of the set $\{\mathfrak{h}_r \mid r \in R^0_{\mathrm{MA}}(n)\} \subset \mathrm{H}_2(\mathrm{IA}_n)$ spanned $\mathrm{H}_2(\mathrm{IA}_n)$.

Although we find the idea in the previous paragraph illuminating, we do not follow it strictly. To make our L–presentation for IA_n easier to comprehend, we will use the following generating set, which is larger than S_{MA}:

$$S_{\mathrm{IA}}(n) = \{C_{x_i,x_j} \mid 1 \le i, j \le n \text{ distinct}\} \\ \cup \{M_{x_i^{\alpha},[x_j^{\beta},x_k^{\gamma}]} \mid 1 \le i, j, k \le n \text{ distinct}, \alpha, \beta, \gamma \in \{\pm 1\}\}.$$

This has the advantage of making our relations and rewriting rules shorter, and making their meaning easier to understand. It has the disadvantage of making the proof of Theorem A less direct. Our theorem giving an L–presentation for IA_n is as follows:

Theorem E (finite L–presentation for Torelli) *For all $n \ge 2$, there exists a finite L–presentation $\mathrm{IA}_n = \langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n)\rangle$ whose set of induced endomorphisms generates $\mathrm{Aut}(F_n) \subset \mathrm{Aut}(\mathrm{IA}_n) \subset \mathrm{End}(\mathrm{IA}_n)$.*

We note that our presentation is not a presentation in which all relators are commutators. The formulas for the $R^0_{\mathrm{IA}}(n)$ and $E_{\mathrm{IA}}(n)$ in our finite L–presentation are a little complicated, so we postpone them until Section 2. The formulas in that section make it clear that $R_{\mathrm{IA}}(n)$ does not lie in $[F(S_{\mathrm{IA}}(n)), F(S_{\mathrm{IA}}(n))]$. Therefore we cannot prove Theorem A simply by interpreting the relators as homology classes. We must do something more complicated to deduce that theorem from our presentation.

Remark 1.6 The relations in Table 1 are not sufficient for our L–presentation. Indeed, they all lie in the commutator subgroup, but the generators $S_{\mathrm{IA}}(n)$ do not map to linearly independent elements of the abelianization.

Sketch of proof We close this introduction by briefly discussing how we prove Theorem E. In particular, we explain why it is easier to verify an L–presentation than a standard presentation. We remark that our proof is inspired by a recent paper [9] of the second author together with Brendle and Margalit, which constructed generators for the kernel of the Burau representation evaluated at -1.

Assume that we have guessed a finite L–presentation $\langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n)\rangle$ for IA_n as in Theorem E (we found the one that we use by first throwing in all the relations we could think of and then attempting the proof below; each time it failed it revealed a relation that we had missed). Let $\mathcal{Q}_n$ be the group presented by the purported L–presentation. There is thus a surjection $\pi\colon \mathcal{Q}_n \to \mathrm{IA}_n$, and the goal of our proof will be to construct an inverse map $\phi\colon \mathrm{IA}_n \to \mathcal{Q}_n$ satisfying $\phi \circ \pi = \mathrm{id}$. This will involve several steps.

Step 1 *We decompose* IA_n *in terms of stabilizers of conjugacy classes of primitive elements of* F_n*.*

For $z \in F_n$, let $[\![z]\!]$ denote the union of the conjugacy classes of z and z^{-1}. A *primitive element* of F_n is an element that forms part of a free basis. Let $\mathcal{C} = \{[\![z]\!] \mid z \in F_n \text{ primitive}\}$. The set $\mathcal{C}$ forms the set of vertices of a simplicial complex called the *complex of partial bases*, which is analogous to the complex of curves for the mapping class group. Applying a theorem of the second author [29] to the action of IA_n on the complex of partial bases, we will obtain a decomposition

$$\mathrm{IA}_n = \mathop{\ast}_{c \in \mathcal{C}} (\mathrm{IA}_n)_c / (\text{some relations}); \tag{2}$$

here $(\mathrm{IA}_n)_c$ denotes the stabilizer in IA_n of c. The unlisted relations play only a small role in our proof and can be ignored at this point.

Step 2 *We use induction to construct a partial inverse.*

Fix some $c_0 \in \mathcal{C}$. The stabilizer $(\mathrm{IA}_n)_{c_0}$ is very similar to IA_{n-1}; in fact, it is connected to IA_{n-1} by an exact sequence that is analogous to the Birman exact sequence for the mapping class group. We construct this exact sequence in the companion paper [17], which builds on our previous paper [15]. By analyzing this exact sequence and using induction, we will construct a "partial inverse" $\phi_{c_0}\colon (\mathrm{IA}_n)_{c_0} \to \mathcal{Q}_n$. We remark that this step is where most of our relations arise — they actually are relations in the kernel of the Birman exact sequence we construct in [15].

Step 3 *We use the L–presentation to lift the conjugation action of* $\mathrm{Aut}(F_n)$ *on* IA_n *to* $\mathcal{Q}_n$*.*

Let $\widetilde{E}$ be the induced endomorphisms of our L–presentation. We directly prove that these endomorphisms actually give an action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ such that the projection map $\pi\colon \mathcal{Q}_n \to \mathrm{IA}_n$ is equivariant. This is the key place where we use properties of L–presentations; in general, it is difficult to construct group actions on groups given by generators and relations.

Step 4 *We use our group action to construct the inverse.*

The conjugation action of $\mathrm{Aut}(F_n)$ on IA_n transitively permutes the terms of (2). Using our lifted action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ as a "guide", we then "move" the partially defined inverse ϕ_{c_0} around and construct ϕ on the rest of IA_n, completing the proof.

Remark 1.7 In [28], the second author constructed an infinite presentation of the Torelli subgroup of the mapping class group. Though this used the same result [29] that we quoted above, the details are quite different. One source of this difference is that instead of an L–presentation with a finite generating set, the paper [28] constructed an ordinary presentation with an infinite generating set.

Computer calculations At several places in this paper, we will need to verify large numbers of equations in group presentations. Rather than displaying these equations in the paper or leaving them as exercises, we use the GAP System to store and check our equations mechanically. The code to verify these equations is in the file `h2ia.g`, which is in an online supplement. We found the equations in this file by hand, and our proof does not rely on a computer search. We will say more about this in Section 5, where said calculations begin.

This is a good place to note that our results rely strongly on the authors' earlier paper [17] and the computer calculations from that paper. There we use a similar approach to automatically verify identities and prove the existence of certain homomorphisms between groups given by presentations. This is used in more than one place in the present paper, but most crucially in Proposition 3.14. The computations from our earlier paper are in a file `iabes.g`, which is available on the authors' websites and on arXiv.

Outline We begin in Section 2 by giving a precise statement of the L–presentation whose existence is asserted in Theorem E. Next, in Section 3 we discuss several tools that are needed for the proof of Theorem E. The proof of Theorem E is in Section 4. This proof depends on some combinatorial group theory calculations that are stated in Sections 2 and 3 but whose proofs are postponed until Section 5. In Section 6, we prove Theorem A. That section also shows how to derive Theorem B from Theorem A. Finally, Theorems C and D are proven in Section 7.

Acknowledgements Day was supported in part by NSF grant DMS-1206981 and Putman was supported in part by NSF grant DMS-1255350 and the Alfred P Sloan Foundation.

(R0) $M^{-1}_{x_a^\alpha,[x_b^\beta,x_c^\gamma]} = M_{x_a^\alpha,[x_c^\gamma,x_b^\beta]}$.

(R1) $[C_{x_a,x_b}, C_{x_c,x_d}] = 1$, possibly with $b = d$.

(R2) $[M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}, M_{x_d^\delta,[x_e^\epsilon,x_f^\zeta]}] = 1$, possibly with $\{b,c\} \cap \{e,f\} \neq \varnothing$, or with $x_a^\alpha = x_d^{-\delta}$ as long as $x_a^\alpha \neq x_d^\delta$, $a \notin \{e,f\}$ and $d \notin \{b,c\}$.

(R3) $[C_{x_a,x_b}, M_{x_c^\gamma,[x_d^\delta,x_e^\epsilon]}] = 1$,
possibly with $b \in \{d,e\}$ if $c \notin \{a,b\}$ and $a \notin \{c,d,e\}$.

(R4) $[C_{x_a,x_b} C_{x_c,x_b}, C_{x_c,x_a}] = 1$.

(R5) $C^\beta_{x_a,x_b} M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]} C^{-\beta}_{x_a,x_b} = M_{x_a^\alpha,[x_c^\gamma,x_b^{-\beta}]}$.

(R6) $M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]} M_{x_a^{-\alpha},[x_b^\beta,x_c^\gamma]} = [C^{-\gamma}_{x_a,x_c}, C^{-\beta}_{x_a,x_b}]$.

(R7) $[C^{-\beta}_{x_a,x_b}, M_{x_b^\beta,[x_c^\gamma,x_d^\delta]}] = [C^{-\delta}_{x_a,x_d}, C^{-\gamma}_{x_a,x_c}]$.

(R8) $M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]} M_{x_d^\delta,[x_a^\alpha,x_e^\epsilon]} M_{x_a^\alpha,[x_c^\gamma,x_b^\beta]} =$
$C^{-\epsilon}_{x_d,x_e} M_{x_d^\delta,[x_c^\gamma,x_b^\beta]} C^{\epsilon}_{x_d^\delta,x_e} M_{x_d^\delta,[x_a^\alpha,x_e^\epsilon]} M_{x_d^\delta,[x_b^\beta,x_c^\gamma]}$, possibly with $e \in \{b,c\}$.

(R9) $C^\beta_{x_a,x_b} M_{x_c^\gamma,[x_a^\alpha,x_d^\delta]} C^{-\beta}_{x_a,x_b} =$
$C^{-\delta}_{x_c,x_d} M_{x_c^\gamma,[x_a^\alpha,x_b^\beta]} C^{\delta}_{x_c,x_d} M_{x_c^\gamma,[x_a^\alpha,x_d^\delta]} M_{x_c^\gamma,[x_b^\beta,x_a^\alpha]}$, possibly with $b = d$.

Table 2: Basic relations for the L–presentation of IA_n. Distinct letters are assumed to represent distinct indices unless stated otherwise. Let $R_{\mathrm{IA}}(n)$ denote the finite set of all relations from the above ten classes.

2 Our finite L–presentation

We now discuss the relations $R^0_{\mathrm{IA}}(n)$ and endomorphisms $E_{\mathrm{IA}}(n)$ of our L–presentation. Two calculations (Propositions 2.1 and 2.3) are postponed until Section 5.

Relations Our set $R^0_{\mathrm{IA}}(n)$ of relations consists of the relations in Table 2. It is easy to verify that these are all indeed relations:

Proposition 2.1 *The relations $R^0_{\mathrm{IA}}(n)$ all hold when interpreted in* IA_n.

The proof is computational and is postponed until Section 5. We remark that unlike many of our computer calculations, it is not particularly difficult to verify by hand.

Since the relations are rather complicated we suggest to the reader that they not pay too close attention to them on their first pass through the paper. The overall structure of our proof (and, in fact, the majority of its details) can be understood without much knowledge of our relations.

Remark 2.2 The relations in $R^0_{\mathrm{IA}}(n)$ have reasonable intuitive interpretations. Relations (R1)–(R3) state that generators acting only in different places commute with each other. Relation (R4) is a generalization of the fact that for $n = 3$, the conjugation move C_{x_3,x_1} conjugates the inner automorphism $C_{x_1,x_2}C_{x_3,x_2}$ back to itself (since it fixes the conjugating element x_2). Relation (R5) makes sense by looking at either side of x_a: on the right of x_a^{α}, instances of $x_b^{\pm\beta}$ cancel, but on the left side of x_a^{α}, we get a conjugate of a basic commutator that is itself a basic commutator. Relation (R6) states that conjugation by a commutator is the same as acting by a commutator of conjugation moves. Relations (R7)–(R9) allow us to rewrite a conjugate of a generator acting on a given element as a product of generators acting only on that same element (x_a, x_d or x_c as stated here, respectively). In this sense, these relations are like the Steinberg relations from the presentation of $\mathrm{GL}_n(\mathbb{Z})$ in algebraic K–theory.

Generators for the automorphism group of a free group Before discussing our endomorphisms $E_{\mathrm{IA}}(n)$, we first introduce a generating set for $\mathrm{Aut}(F_n)$ that goes back to work of Nielsen. For $\alpha = \pm 1$ and distinct $1 \le i, j \le n$, let $M_{x_i^{\alpha},x_j} \in \mathrm{Aut}(F_n)$ be the *transvection* that takes x_i^{α} to $x_j x_i^{\alpha}$ and fixes x_ℓ for $\ell \ne i$. Just like before, we have

$$M_{x_i^{-1},x_j}(x_i^{-1}) = x_j x_i^{-1} \quad \text{and} \quad M_{x_i^{-1},x_j}(x_i) = x_i x_j^{-1}.$$

Next, for distinct $1 \le i, j \le n$ let $P_{i,j} \in \mathrm{Aut}(F_n)$ be the *swap automorphism* that exchanges x_i and x_j while fixing x_ℓ for $\ell \ne i, j$. Finally, for $1 \le i \le n$ let $I_i \in \mathrm{Aut}(F_n)$ be the *inversion automorphism* that takes x_i to x_i^{-1} and fixes x_ℓ for $\ell \ne i$. Define

$$\begin{aligned} S_{\mathrm{Aut}}(n) = \{M_{x_i^{\alpha},x_j}^{\beta} \mid 1 \le i, j \le n \text{ distinct}, \alpha, \beta \in \{\pm 1\}\} \\ \cup \{P_{i,j} \mid 1 \le i, j \le n \text{ distinct}\} \cup \{I_i \mid 1 \le i \le n\}. \end{aligned}$$

Observe that the set $S_{\mathrm{Aut}}(n) \subset \mathrm{Aut}(F_n)$ is closed under inversion.

Endomorphisms Below we will define a function $\theta\colon S_{\mathrm{Aut}}(n) \to \mathrm{End}(F(S_{\mathrm{IA}}(n)))$ with this key property: Let $\pi\colon F(S_{\mathrm{IA}}(n)) \to \mathrm{IA}_n$ and $\rho\colon F(S_{\mathrm{Aut}}(n)) \to \mathrm{Aut}(F_n)$ be the projections. Then for $s \in S_{\mathrm{Aut}}(n)$ and $w \in F(S_{\mathrm{IA}}(n))$, we have

$$\pi(\theta(s)(w)) = \rho(s)\pi(w)\rho(s)^{-1} \in \mathrm{IA}_n . \tag{3}$$

Our set of endomorphisms will then be

$$E_{\mathrm{IA}}(n) = \{\theta(s) \mid s \in S_{\mathrm{Aut}}(n)\}.$$

The relevance of the formula (3) is that we want the induced endomorphisms of our IA–presentation of IA_n to generate the image of $\mathrm{Aut}(F_n)$ in $\mathrm{Aut}(\mathrm{IA}_n) \subset \mathrm{End}(\mathrm{IA}_n)$ arising from the conjugation action of $\mathrm{Aut}(F_n)$ on IA_n.

Defining θ To define an endomorphism $\theta(s)\colon F(S_{\mathrm{IA}}(n)) \to F(S_{\mathrm{IA}}(n))$ for $s \in S_{\mathrm{Aut}}(n)$, it is enough to say what $\theta(s)$ does to each element of $S_{\mathrm{IA}}(n)$. There are two cases:

$s \in S_I$	$\theta(M^{\beta}_{x_a^{\alpha}, x_b})(s)$
C_{x_c, x_a}	$(C^{\alpha}_{x_c, x_a} C^{\beta}_{x_c, x_b})^{\alpha}$
C_{x_a, x_c}	$C_{x_a, x_c} M_{x_a^{\alpha}, [x_b^{-\beta}, x_c]}$
C_{x_b, x_c}	$C_{x_b, x_c} M_{x_a^{\alpha}, [x_b^{-\beta}, x_c^{-1}]}$
C_{x_b, x_a}	$(C^{\beta}_{x_a, x_b} C^{\alpha}_{x_b, x_a})^{\alpha}$
$M_{x_a^{\alpha}, [x_c^{\gamma}, x_d^{\delta}]}$	$C^{\beta}_{x_a, x_b} M_{x_a^{\alpha}, [x_c^{\gamma}, x_d^{\delta}]} C^{-\beta}_{x_a, x_b}$
$M_{x_c^{\gamma}, [x_a^{\alpha}, x_d^{\delta}]}$	$M_{x_c^{\gamma}, [x_b^{\beta}, x_d^{\delta}]} C^{-\beta}_{x_c, x_b} M_{x_c^{\gamma}, [x_a^{\alpha}, x_d^{\delta}]} C^{\beta}_{x_c, x_b}$
$M_{x_c^{\gamma}, [x_a^{-\alpha}, x_d^{\delta}]}$	$M_{x_c^{\gamma}, [x_a^{-\alpha}, x_d^{\delta}]} C^{\alpha}_{x_c, x_a} M_{x_c^{\gamma}, [x_b^{-\beta}, x_d^{\delta}]} C^{-\alpha}_{x_c, x_a}$
$M_{x_b^{\beta}, [x_c^{\gamma}, x_d^{\delta}]}$	$C^{\beta}_{x_a, x_b} M_{x_a^{-\alpha}, [x_c^{\gamma}, x_d^{\delta}]} M_{x_b^{\beta}, [x_c^{\gamma}, x_d^{\delta}]} C^{-\beta}_{x_a, x_b}$
$M_{x_b^{-\beta}, [x_c^{\gamma}, x_d^{\delta}]}$	$M_{x_a^{\alpha}, [x_c^{\gamma}, x_d^{\delta}]} M_{x_b^{-\beta}, [x_c^{\gamma}, x_d^{\delta}]}$
$M_{x_a^{\alpha}, [x_b^{\beta}, x_c^{\gamma}]}$	$C^{\beta}_{x_a, x_b} M_{x_a^{\alpha}, [x_b^{\beta}, x_c^{\gamma}]} C^{-\beta}_{x_a, x_b}$
$M_{x_a^{\alpha}, [x_b^{-\beta}, x_c^{\gamma}]}$	$C^{\beta}_{x_a, x_b} M_{x_a^{\alpha}, [x_b^{-\beta}, x_c^{\gamma}]} C^{-\beta}_{x_a, x_b}$
$M_{x_b^{\beta}, [x_a^{\alpha}, x_c^{\gamma}]}$	$C^{\gamma}_{x_a, x_c} M_{x_a^{\alpha}, [x_b^{-\beta}, x_c^{\gamma}]} C^{-\gamma}_{x_b, x_c} M_{x_b^{-\beta}, [x_a^{\alpha}, x_c^{-\gamma}]}$
$M_{x_b^{\beta}, [x_a^{-\alpha}, x_c^{\gamma}]}$	$C^{-\beta}_{x_c, x_b} C^{-\alpha}_{x_c, x_a} M_{x_b^{-\beta}, [x_c^{-\gamma}, x_a^{\alpha}]} C^{\gamma}_{x_b, x_c} M_{x_a^{\alpha}, [x_c^{\gamma}, x_b^{-\beta}]} C^{-\gamma}_{x_a, x_c} C^{\alpha}_{x_c, x_a} C^{\beta}_{x_c, x_b}$
$M_{x_b^{-\beta}, [x_a^{\alpha}, x_c^{\gamma}]}$	$C^{-\gamma}_{x_a, x_c} C^{\alpha}_{x_c, x_a} M_{x_b^{-\beta}, [x_c^{\gamma}, x_a^{-\alpha}]} C^{\beta}_{x_c^{\gamma}, x_b} M_{x_a^{\alpha}, [x_c^{\gamma}, x_b^{-\beta}]} C^{\gamma}_{x_b, x_c} C^{-\beta}_{x_c, x_b} C^{-\alpha}_{x_c, x_a}$
$M_{x_b^{-\beta}, [x_a^{-\alpha}, x_c^{\gamma}]}$	$C^{-\gamma}_{x_b, x_c} M_{x_a^{\alpha}, [x_b^{-\beta}, x_c^{\gamma}]} C^{-\beta}_{x_c, x_b} M_{x_b^{-\beta}, [x_a^{-\alpha}, x_c^{\gamma}]} C^{-\alpha}_{x_c, x_a} C^{\gamma}_{x_a, x_c} C^{\alpha}_{x_c, x_a} C^{\beta}_{x_c, x_b}$
$M_{x_c^{\gamma}, [x_a^{\alpha}, x_b^{\beta}]}$	$C^{\beta}_{x_a, x_b} M_{x_c^{\gamma}, [x_a^{\alpha}, x_b^{\beta}]} C^{-\beta}_{x_a, x_b}$
$M_{x_c^{\gamma}, [x_a^{\alpha}, x_b^{-\beta}]}$	$M_{x_c^{\gamma}, [x_b^{\beta}, x_a^{\alpha}]}$

Table 3: Definition of $\theta(M^{\beta}_{x_a^{\alpha}, x_b})$ on the generators $S_{\mathrm{IA}}(n)$. All indices in each entry are assumed to be distinct. If no entry is listed for $t \in S_{\mathrm{IA}}(n)$ or for the generator representing t^{-1} (as in relation (R0)) then $\theta(M^{\beta}_{x_a^{\alpha}, x_b})(t) = t$.

- $\boldsymbol{s = P_{i,j}}$ **or** $\boldsymbol{s = I_i}$ We then define $\theta(s)$ using the action of s on F_n via
$$\theta(s)(C_{x_a, x_b}) = C_{s(x_a), s(x_b)} \quad \text{and} \quad \theta(s)(M_{x_a^{\alpha}, [x_b^{\beta}, x_c^{\gamma}]}) = M_{s(x_a^{\alpha}), [s(x_b^{\beta}), s(x_c^{\gamma})]}.$$
Here one should interpret $C_{x_e^{-1}, x_f}$ as C_{x_e, x_f}, $C_{x_e, x_f^{-1}}$ as $C^{-1}_{x_e, x_f}$ and $C_{x_e^{-1}, x_f^{-1}}$ as $C^{-1}_{x_e, x_f}$.
- $\boldsymbol{s = M_{x_a^{\alpha}, x_b}}$ In this case, we define $\theta(s)$ via the formulas in Table 3. These list the cases where $\theta(s)$ does not fix a generator, except that, to avoid redundancy, we do not always list both a commutator transvection and its inverse. Specifically, if $t = M_{x_c^{\gamma}, [x_d^{\delta}, x_e^{\epsilon}]}$, possibly with $\{c, d, e\} \cap \{a, b\} \neq \varnothing$, and a formula is listed for $t' = M_{x_c^{\gamma}, [x_e^{\epsilon}, x_d^{\delta}]}$ but not for t, then we define
$$\theta(s)(t) = \theta(s)(t')^{-1}.$$

If Table 3 lists no entry for t or t', or the table lists no entry for t and t is a conjugation move, then we define $\theta(s)(t) = t$.

These formulas were chosen to be as simple as possible, among formulas realizing (3). Just like for the relations, we recommend not dwelling on these formulas during one's first read through this paper.

Proposition 2.3 *The definition of θ satisfies (3).*

This proof uses a computer verification and is postponed until Section 5. Propositions 2.1 and 2.3 together imply that all of the extended relations from our L–presentation are trivial in IA_n. This means that the obvious map on generators (sending each generator to the automorphism it names) extends to a well-defined homomorphism

$$\langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n) \rangle \to \mathrm{IA}_n \,.$$

3 Tools for the proof

In this section, we assemble the tools we will need to prove Theorem E. In Section 3.1, we discuss a theorem of the second author that gives a sort of infinite presentation for a group acting on a simplicial complex. In Section 3.2, we introduce the complex of partial bases. In Section 3.3, we give generators for the IA_n–stabilizers of simplices in the complex of partial bases. In Section 3.4, we introduce an action of $\mathrm{Aut}(F_n)$ on the group given by our purported L–presentation for IA_n. Finally, in Section 3.5 we introduce a certain morphism between groups given by L–presentations.

Two results in this sections have computer-aided proofs which are postponed until Section 5: Proposition 3.13 from Section 3.4 and Proposition 3.14 from Section 3.5.

3.1 Presentations from group actions

Consider a group G acting on a simplicial complex X. We say that G acts *without rotations* if for all simplices σ of X, the setwise and pointwise stabilizers of σ coincide. For a simplex σ, denote by G_σ the stabilizer of σ. Letting $X^{(0)}$ denote the vertex set of X, there is a homomorphism from the free product of vertex stabilizers

$$\psi\colon \mathop{\ast}_{v \in X^{(0)}} G_v \to G.$$

As notation, if $g \in G$ stabilizes a vertex v of X, then denote by g_v the associated element of

$$G_v < \mathop{\ast}_{v \in X^{(0)}} G_v.$$

The map ψ is rarely injective. Two families of elements in its kernel are as follows:

- If e is an edge of X joining vertices v and v' and if $g \in G_e$, then $g_v g_{v'}^{-1} \in \ker(\psi)$. We call these the *edge relators*.
- If $v, w \in X^{(0)}$ and $g \in G_v$ and $h \in G_w$, then $h_w g_v h_w^{-1} (hgh^{-1})_{h(v)}^{-1} \in \ker(\psi)$. We call these the *conjugation relators*.

The second author gave hypotheses under which these generate $\ker(\psi)$:

Theorem 3.1 [29] *Consider a group G acting without rotations on a 1–connected simplicial complex X. Assume that X/G is 2–connected. Then the kernel of the map ψ described above is normally generated by the edge and conjugation relators.*

3.2 The complex of partial bases

We now introduce the simplicial complex to which we will apply Theorem 3.1. For $z \in F_n$, let $[\![z]\!]$ denote the union of the conjugacy classes of z and z^{-1}.

Definition 3.2 A *partial basis* for F_n is a set $\{z_1, \ldots, z_k\} \subset F_n$ such that there exist $z_{k+1}, \ldots, z_n \in F_n$ with $\{z_1, \ldots, z_n\}$ a free basis for F_n. The *complex of partial bases for F_n*, denoted by $\mathcal{B}_n$, is the simplicial complex whose $(k-1)$–simplices are sets $\{[\![z_1]\!], \ldots, [\![z_k]\!]\}$, where $\{z_1, \ldots, z_k\}$ is a partial basis for F_n.

The group $\mathrm{Aut}(F_n)$ acts on $\mathcal{B}_n$, and we wish to apply Theorem 3.1 to the restriction of this action to IA_n. It is clear that IA_n acts on $\mathcal{B}_n$ without rotations, so we must check that $\mathcal{B}_n$ is 1–connected and that $\mathcal{B}_n/\mathrm{IA}_n$ is 2–connected.

We start by verifying that $\mathcal{B}_n$ is 1–connected.

Proposition 3.3 *The simplicial complex $\mathcal{B}_n$ is 1–connected for $n \geq 3$.*

Proof For $z \in F_n$, let $[\![z]\!]'$ be the conjugacy class of z. Define $\mathcal{B}_n'$ to be the simplicial complex whose $(k-1)$–simplices are sets $\{[\![z_1]\!]', \ldots, [\![z_k]\!]'\}$, where $\{z_1, \ldots, z_k\}$ is a partial basis for F_n. In [16], the authors proved that $\mathcal{B}_n'$ is 1–connected for $n \geq 3$. There is a natural simplicial map $\rho\colon \mathcal{B}_n' \to \mathcal{B}_n$. Letting $\psi^0\colon (\mathcal{B}_n)^{(0)} \to (\mathcal{B}_n')^{(0)}$ be an arbitrary map satisfying $\rho \circ \psi^0 = \mathrm{id}$, it is clear that ψ^0 extends to a simplicial map $\psi\colon \mathcal{B}_n \to \mathcal{B}_n'$ satisfying $\rho \circ \psi = \mathrm{id}$. This implies that ρ induces a surjection on all homotopy groups, so $\mathcal{B}_n$ is 1–connected for $n \geq 3$. □

It also follows from [16] that $\mathcal{B}_n/\mathrm{IA}_n$ is $(n-2)$–connected. In particular, it is 2–connected for $n \geq 4$, and thus satisfies the conditions of Theorem 3.1 for $n \geq 4$. However, we will need a complex that satisfies the conditions of Theorem 3.1 for $n = 3$ as well. We therefore attach cells to increase the connectivity.

Definition 3.4 The *augmented complex of partial bases for* F_n, denoted by $\widehat{\mathcal{B}}_n$, is the simplicial complex whose $(k-1)$–simplices are as follows:

- Sets of the form $\{[\![z_1]\!], \dots, [\![z_k]\!]\}$, where $\{z_1, \dots, z_k\}$ is a partial basis for F_n. These will be called the *standard simplices*.
- Sets of the form $\{[\![z_1 z_2]\!], [\![z_1]\!], [\![z_2]\!], \dots, [\![z_{k-1}]\!]\}$, where $\{z_1, \dots, z_{k-1}\}$ is a partial basis for F_n. These will be called the *additive simplices*.

Remark 3.5 Since $z_1 z_2$ and $z_2 z_1$ are conjugate, the two additive simplices

$$\{[\![z_1 z_2]\!], [\![z_1]\!], [\![z_2]\!], \dots, [\![z_{k-1}]\!]\} \quad \text{and} \quad \{[\![z_2 z_1]\!], [\![z_1]\!], [\![z_2]\!], \dots, [\![z_{k-1}]\!]\}$$

of $\widehat{\mathcal{B}}_n$ are the same.

The group $\mathrm{Aut}(F_n)$ (and hence IA_n) still acts on $\widehat{\mathcal{B}}_n$. Since $\widehat{\mathcal{B}}_n$ is obtained from $\mathcal{B}_n$ by adding simplices of dimension at least 2, it inherits the 1–connectivity of $\mathcal{B}_n$ for $n \geq 3$ asserted in Proposition 3.3.

Proposition 3.6 *The complex* $\widehat{\mathcal{B}}_n$ *is* 1*–connected for* $n \geq 3$.

To help us understand the connectivity of $\widehat{\mathcal{B}}_n / \mathrm{IA}_n$, we introduce the following complex. For $\vec{v} \in \mathbb{Z}^n$, let $(\vec{v})_\pm$ denote the set $\{\vec{v}, -\vec{v}\}$.

Definition 3.7 A *partial basis* for $\mathbb{Z}^n$ is a set $\{\vec{v}_1, \dots, \vec{v}_k\} \subset \mathbb{Z}^n$ such that there exist $\vec{v}_{k+1}, \dots, \vec{v}_n \in \mathbb{Z}^n$ with $\{\vec{v}_1, \dots, \vec{v}_n\}$ a basis for $\mathbb{Z}^n$. The *augmented complex of lax partial bases for* $\mathbb{Z}^n$, denoted by $\widehat{\mathcal{B}}_n(\mathbb{Z})$, is the simplicial complex whose $(k-1)$–simplices are as follows:

- Sets of the form $\{(\vec{v}_1)_\pm, \dots, (\vec{v}_k)_\pm\}$, where $\{\vec{v}_1, \dots, \vec{v}_k\}$ is a partial basis for $\mathbb{Z}^n$. These will be called the *standard simplices*.
- Sets of the form $\{(\vec{v}_1 + \vec{v}_2)_\pm, (\vec{v}_1)_\pm, (\vec{v}_2)_\pm, \dots, (\vec{v}_{k-1})_\pm\}$, where $\{\vec{v}_1, \dots, \vec{v}_{k-1}\}$ is a partial basis for $\mathbb{Z}^n$. These will be called the *additive simplices*.

We then have the following lemma:

Lemma 3.8 *We have* $\widehat{\mathcal{B}}_n / \mathrm{IA}_n \cong \widehat{\mathcal{B}}_n(\mathbb{Z})$ *for* $n \geq 1$.

For the proof of Lemma 3.8, we will need the following result of the authors. For $z \in F_n$, let $[z] \in \mathbb{Z}^n$ be the associated element of the abelianization of F_n.

Lemma 3.9 [16, Lemma 5.3] *Let* $\{\vec{v}_1, \dots, \vec{v}_n\}$ *be a basis for* $\mathbb{Z}^n$ *and let* $\{z_1, \dots, z_k\}$ *be a partial basis for* F_n *such that* $[z_i] = \vec{v}_i$ *for* $1 \leq i \leq k$. *Then there exist* $z_{k+1}, \dots, z_n \in F_n$ *with* $[z_i] = \vec{v}_i$ *for* $k+1 \leq i \leq n$ *such that* $\{z_1, \dots, z_n\}$ *is a basis for* F_n.

Proof of Lemma 3.8 The map $(\widehat{\mathcal{B}}_n)^{(0)} \to (\widehat{\mathcal{B}}_n(\mathbb{Z}))^{(0)}$ that takes $[\![z]\!]$ to $[z]$ extends to a simplicial map $\rho\colon \widehat{\mathcal{B}}_n \to \widehat{\mathcal{B}}_n(\mathbb{Z})$. Since IA_n acts without rotations on $\widehat{\mathcal{B}}_n$, the quotient $\widehat{\mathcal{B}}_n/\,\mathrm{IA}_n$ has a natural CW–complex structure whose k–cells are the IA_n–orbits of the k–cells of $\widehat{\mathcal{B}}_n$ (warning: though it will turn out that in this case it is, this CW–complex structure need not be a simplicial complex structure; consider, for example, the action of $\mathbb{Z}$ by translations on the standard triangulation of $\mathbb{R}$ whose vertices are $\mathbb{Z}$). Since ρ is IA_n–invariant, it factors through a map $\overline{\rho}\colon \widehat{\mathcal{B}}_n/\,\mathrm{IA}_n \to \widehat{\mathcal{B}}_n(\mathbb{Z})$. We will prove that $\overline{\rho}$ is an isomorphism of CW–complexes.

This requires checking two things. The first is that every simplex of $\widehat{\mathcal{B}}_n(\mathbb{Z})$ is in the image of ρ, which is an immediate consequence of Lemma 3.9. The second is that if σ and σ' are simplices of $\widehat{\mathcal{B}}_n$ such that $\rho(\sigma) = \rho(\sigma')$, then there exists some $f \in \mathrm{IA}_n$ such that $f(\sigma) = \sigma'$. It is clear that σ and σ' are either both standard simplices or both additive simplices. Assume first that they are both standard simplices. We can then write

$$\sigma = \{[\![z_1]\!], \dots, [\![z_k]\!]\} \quad \text{and} \quad \sigma' = \{[\![z'_1]\!], \dots, [\![z'_k]\!]\}$$

as in the definition of standard simplices, with $[z_i] = [z'_i]$ for $1 \le i \le k$. Set $\vec{v}_i = [z_i] = [z'_i]$ for $1 \le i \le k$. The set $\{\vec{v}_1, \dots, \vec{v}_k\}$ is a partial basis for $\mathbb{Z}^n$, so we can extend it to a basis $\{\vec{v}_1, \dots, \vec{v}_n\}$. Applying Lemma 3.9 twice, we can find $z_{k+1}, \dots, z_n \in F_n$ and $z'_{k+1}, \dots, z'_n \in F_n$ such that $[z_i] = [z'_i] = \vec{v}_i$ for $k+1 \le i \le n$ and such that both $\{z_1, \dots, z_n\}$ and $\{z'_1, \dots, z'_n\}$ are free bases for F_n. There then exists $f \in \mathrm{Aut}(F_n)$ such that $f(z_i) = z'_i$ for $1 \le i \le n$. By construction, we have $f \in \mathrm{IA}_n$ and $f(\sigma) = \sigma'$.

It remains to deal with the case where σ and σ' are both simplices of additive type. Write

$$\sigma = \{[\![z_1 z_2]\!], [\![z_1]\!], [\![z_2]\!], \dots, [\![z_{k-1}]\!]\} \quad \text{and} \quad \sigma' = \{[\![z'_1 z'_2]\!], [\![z'_1]\!], \dots, [\![z'_{k-1}]\!]\}$$

as in the definition of additive simplices. The unordered sets

$$\bigl\{([z_1]+[z_2])_\pm, ([z_1])_\pm, ([z_2])_\pm\bigr\} \quad \text{and} \quad \bigl\{([z'_1]+[z'_2])_\pm, ([z'_1])_\pm, ([z'_2])_\pm\bigr\}$$

are minimal nonempty subsets of $\rho(\sigma) = \rho(\sigma')$ such that the defining elements of $\mathbb{Z}^n$ are not linearly independent. It follows that as unordered sets we have

$$\rho\bigl(\{[\![z_1 z_2]\!], [\![z_1]\!], [\![z_2]\!]\}\bigr) = \rho\bigl(\{[\![z'_1 z'_2]\!], [\![z'_1]\!], [\![z'_2]\!]\}\bigr),$$
$$\rho\bigl(\{[\![z_3]\!], \dots, [\![z_{k-1}]\!]\}\bigr) = \rho\bigl(\{[\![z'_3]\!], \dots, [\![z'_{k-1}]\!]\}\bigr).$$

Reordering the z_i and possibly replacing some of the z_i by z_i^{-1} (which does not change $[\![z_i]\!]$), we can assume that $[z_i] = [z'_i]$ for $3 \le i \le k-1$.

The next observation is that all of the following sets define the same additive simplex (but with the vertices in a different order; all six possible orderings occur):

$$\{[\![z_1z_2]\!],[\![z_1]\!],[\![z_2]\!]\},\quad \{[\![z_2z_1]\!],[\![z_2]\!],[\![z_1]\!]\},\quad \{[\![z_1]\!],[\![z_1z_2]\!],[\![z_2^{-1}]\!]\},$$
$$\{[\![z_1^{-1}]\!],[\![z_2]\!],[\![z_2^{-1}z_1^{-1}]\!]\},\quad \{[\![z_2]\!],[\![z_2z_1]\!],[\![z_1^{-1}]\!]\},\quad \{[\![z_2^{-1}]\!],[\![z_1]\!],[\![z_1^{-1}z_2^{-1}]\!]\}.$$

By reordering σ and possibly changing some of our expressions for the elements in it again, we can assume that

$$([z_1]+[z_2])_\pm = ([z_1']+[z_2'])_\pm,\quad ([z_1])_\pm = ([z_1'])_\pm,\quad ([z_2])_\pm = ([z_2'])_\pm$$

and that $[z_i]=[z_i']$ for $3\le i\le k-1$.

The final observation is that either

$$([z_1],[z_2]) = ([z_1'],[z_2']) \quad\text{or}\quad ([z_1],[z_2]) = (-[z_1'],-[z_2']);$$

the key point here is that changing the sign of one of $\{[z_1],[z_2]\}$ but not the other changes $([z_1]+[z_2])_\pm$. If the second possibility occurs, then replace z_1 and z_2 with z_1^{-1} and z_2^{-1}, respectively; this does not change σ. The upshot is that we now have arranged for $[z_i]=[z_i']$ for all $1\le i\le k-1$. By the same argument we used to deal with standard simplices, there exists some $f\in \mathrm{IA}_n$ such that $f(z_i)=z_i'$ for $1\le i\le k-1$. Since $f(z_1z_2)=z_1'z_2'$, we see that $f(\sigma)=\sigma'$, as desired. □

The second author together with Church proved in [13] that $\widehat{\mathcal{B}}_n(\mathbb{Z})$ is $(n-1)$–connected for $n\ge 1$. We therefore deduce the following:

Proposition 3.10 *The complex $\widehat{\mathcal{B}}_n/\mathrm{IA}_n$ is $(n-1)$–connected for $n\ge 1$.*

3.3 Generators for simplex stabilizers

This section is devoted to the following proposition, which gives generators for the stabilizers in IA_n of simplices of $\mathcal{B}_n$. Recall that $S_{\mathrm{MA}}(n)$ is Magnus's generating set for IA_n discussed in the introduction.

Proposition 3.11 *Fix $1\le k\le n$ and define $\Gamma = (\mathrm{IA}_n)_{[\![x_{n-k+1}]\!],[\![x_{n-k+2}]\!],\dots,[\![x_n]\!]}$. Then Γ is generated by*

$$S_{\mathrm{MA}}(n)\cap\Gamma = \{C_{x_a,x_b} \mid 1\le a,b\le n \text{ distinct}\}$$
$$\cup\{M_{x_a,[x_b,x_c]} \mid 1\le a\le n-k,\ 1\le b,c\le n \text{ distinct}\}.$$

Proof The map $F_n\to F_{n-k}$ that quotients by the normal closure of $\{x_{n-k+1},\dots,x_n\}$ induces a split surjection $\rho\colon \Gamma\to \mathrm{IA}_{n-k}$. Define $\mathcal{K}_{n-k,k}=\ker(\rho)$, so we have

$\Gamma = \mathcal{K}_{n-k,k} \rtimes \mathrm{IA}_{n-k}$. As we said in the introduction, Magnus [26] proved that IA_{n-k} is generated by

$$(4)\quad \{C_{x_a,x_b} \mid 1 \le a,b \le n-k \text{ distinct}\} \cup \{M_{x_a,[x_b,x_c]} \mid 1 \le a,b,c \le n-k \text{ distinct}\}.$$

The authors proved in [17, Theorem A] that $\mathcal{K}_{n-k,k}$ is generated by

$$\begin{aligned}(5)\quad & \{C_{x_a,x_b} \mid n-k+1 \le a \le n,\ 1 \le b \le n \text{ distinct}\} \\ & \cup \{C_{x_a,x_b} \mid 1 \le a \le n,\ n-k+1 \le b \le n \text{ distinct}\} \\ & \quad \cup \{M_{x_a,[x_b,x_c]} \mid 1 \le a \le n-k,\ n-k-1 \le b \le n,\ 1 \le c \le n \text{ distinct}\}.\end{aligned}$$

The union of (4) and (5) is the claimed generating set for Γ. □

Remark 3.12 For $z \in F_n$, define $[\![z]\!]'$ to be the conjugacy class of z. In [17] we deal with $(\mathrm{IA}_n)_{[\![x_{n-k+1}]\!]', [\![x_{n-k+2}]\!]', \dots, [\![x_n]\!]'}$ instead of $(\mathrm{IA}_n)_{[\![x_{n-k+1}]\!], [\![x_{n-k+2}]\!], \dots, [\![x_n]\!]}$; however, since x_i and x_i^{-1} have different images in F_n^{ab}, these two stabilizer subgroups are actually equal. There are also notational differences: the group denoted by $\mathcal{K}_{n-k,k}$ here is denoted by $\mathcal{K}_{n-k,k}^{\mathrm{IA}}$ in that paper.

3.4 The action of $\mathrm{Aut}(F_n)$

Let $\mathcal{Q}_n$ be the group with the L–presentation $\langle S_{\mathrm{IA}}(n) \mid R_{\mathrm{IA}}^0(n) \mid E_{\mathrm{IA}}(n)\rangle$ discussed in Section 2. By Propositions 2.1 and 2.3, there is a map $\pi\colon \mathcal{Q}_n \to \mathrm{IA}_n$. The group $\mathrm{Aut}(F_n)$ acts on IA_n by conjugation. The goal of this section is to state Proposition 3.13 below, which asserts that this action can be lifted to $\mathcal{Q}_n$.

To state some important properties of this lifted action, we must introduce some notation. First, let $S_{\mathrm{Aut}}(n) \subset \mathrm{Aut}(F_n)$ be the generating set discussed in Section 2. Recall that $E_{\mathrm{IA}}(n) \subset \mathrm{End}(F(S_{\mathrm{IA}}(n)))$ is the image of a map $\theta\colon S_{\mathrm{Aut}}(n) \to \mathrm{End}(F(S_{\mathrm{IA}}(n)))$. There is thus a map $e\colon S_{\mathrm{Aut}}(n) \to \mathrm{End}(\mathcal{Q}_n)$ whose image is the set of induced endomorphisms of our L–presentation. It will turn out that the image of e consists of automorphisms, and these automorphisms generate the action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$.

Second, recall that $\{x_1, \dots, x_n\}$ is a fixed free basis for F_n. Let

$$\begin{aligned}(S_{\mathrm{IA}}(n))_{[\![x_n]\!]} = & \{C_{x_a,x_b} \mid 1 \le a,b \le n \text{ distinct}\} \\ & \cup \{M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]} \mid 1 \le a,b,c \le n \text{ distinct},\ \alpha,\beta,\gamma \in \{\pm 1\},\ a \ne n\}.\end{aligned}$$

This is exactly the subset of $S_{\mathrm{IA}}(n) \subset \mathrm{IA}_n$ consisting of automorphisms that fix $[\![x_n]\!]$; Proposition 3.11 (with $k=1$) implies that it generates the stabilizer subgroup $(\mathrm{IA}_n)_{[\![x_n]\!]}$. Define $(\mathcal{Q}_n)_{[\![x_n]\!]}$ be the subgroup of $\mathcal{Q}_n$ generated by $(S_{\mathrm{IA}}(n))_{[\![x_n]\!]}$. We will then require the stabilizer subgroup $(\mathrm{Aut}(F_n))_{[\![x]\!]}$ to preserve the subgroup $(\mathcal{Q}_n)_{[\![x_n]\!]}$.

Our proposition is as follows:

Proposition 3.13 *For all $n \geq 2$, there is an action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ that satisfies the following three properties:*

(1) *The action comes from the induced endomorphisms in the sense that, for $s \in S_{\mathrm{IA}}(n) \subset \mathrm{Aut}(F_n)$ and $q \in \mathcal{Q}_n$, we have $s \cdot q = e(s) \cdot q$.*
(2) *The restriction of the action to IA_n induces the conjugation action of $\mathcal{Q}_n$ on itself in the sense that, for $q, r \in \mathcal{Q}_n$, we have $\pi(r) \cdot q = rqr^{-1}$.*
(3) *For $\eta \in (\mathrm{Aut}(F_n))_{[\![x_n]\!]}$ and $q \in (\mathcal{Q}_n)_{[\![x_n]\!]}$, we have $\eta \cdot q \in (\mathcal{Q}_n)_{[\![x_n]\!]}$.*

The proof of Proposition 3.13 is a computation with generators and relations (mostly done by computer), so we have postponed it until Section 5.

3.5 A homomorphism between L–presentations

There is a natural split surjection $\rho\colon (\mathrm{IA}_n)_{[\![x_n]\!]} \to \mathrm{IA}_{n-1}$ arising from the quotient map $F_n \to F_{n-1}$ whoe kernel is the normal closure of x_n. Let $\mathcal{K}_{n-1,1} = \ker(\rho)$; so we have a decomposition $(\mathrm{IA}_n)_{[\![x_n]\!]} = \mathcal{K}_{n-1,1} \rtimes \mathrm{IA}_{n-1}$. Building on the Birman exact sequence for $\mathrm{Aut}(F_n)$ we constructed in [15], we constructed an L–presentation for $\mathcal{K}_{n-1,1}$ in [17, Theorem D] ($\mathcal{K}_{n-1,1}$ is denoted by $\mathcal{K}^{\mathrm{IA}}_{n-1,1}$ in that paper). This L–presentation plays a crucial role in the inductive step of our proof, because it allows us to obtain the following proposition:

Proposition 3.14 *There is a homomorphism $\mathcal{K}_{n-1,1} \to \langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n) \rangle$ fitting into the following commuting triangle:*

$$\begin{tikzcd}
\mathcal{K}_{n-1,1} \arrow[r] \arrow[dr, hook] & \langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n) \rangle \arrow[d, two heads] \\
& \mathrm{IA}_n
\end{tikzcd}$$

Usually finding a homomorphism between groups given by presentations is simple: one checks that the relations map to products of conjugates of relations. This is the spirit of the proof of Proposition 3.14, but the substitution rules and extended relations complicate the picture. Our proof of Proposition 3.14 is computer-assisted and is postponed until Section 5.

4 Verification of our L–presentation

In this section, we prove Theorem E, which says that IA_n has the finite L–presentation $\langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n) \rangle$ discussed in Section 2. Our proof is inspired by the proof of the main theorem of [9]. We will make use of Propositions 2.1, 2.3, 3.13 and 3.14, which were all stated in previous sections and which will be proved (with the aid of a computer) in Section 5.

Proof of Theorem E Let $\mathcal{Q}_n$ be the group given by $\langle S_{\mathrm{IA}}(n) \mid R^0_{\mathrm{IA}}(n) \mid E_{\mathrm{IA}}(n)\rangle$. Elements of $S_{\mathrm{IA}}(n)$ play dual roles as elements of $\mathcal{Q}_n$ and as elements of IA_n, and during our proof it will be important to distinguish them. Therefore, throughout this proof elements C_{x_a,x_b} and $M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}$ will always lie in IA_n; the associated elements of $\mathcal{Q}_n$ will be denoted by $\mathfrak{C}_{x_a,x_b}$ and $\mathfrak{M}_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}$.

There is a natural projection map $\pi\colon \mathcal{Q}_n \to \mathrm{IA}_n$. We will prove that π is an isomorphism by induction on n. The base cases are $n=1$ and $n=2$. For $n=1$, both IA_n and $\mathcal{Q}_n$ are the trivial group, so there is nothing to prove. For $n=2$, it is a classical theorem of Nielsen [27] (see also [25, Proposition 4.5]) that IA_2 is the group of inner automorphisms of F_2, so IA_2 is a free group on the generators C_{x_1,x_2} and C_{x_2,x_1}. Our generating set for $\mathcal{Q}_2$ is $\{\mathfrak{C}_{x_1,x_2}, \mathfrak{C}_{x_2,x_1}\}$, and for $n=2$ the set of basic relations $R^0_{\mathrm{IA}}(2)$ is empty. Even though our set of substitution rules $E_{\mathrm{IA}}(2)$ is nonempty, it follows that our full set of relations for $\mathcal{Q}_2$ is empty. So our presentation for $\mathcal{Q}_2$ is $\langle \mathfrak{C}_{x_1,x_2}, \mathfrak{C}_{x_2,x_1} \mid \varnothing\rangle$, and the result is also true in this case.

Assume now that $n \ge 3$ and that the projection map $\mathcal{Q}_{n'} \to \mathrm{IA}_{n'}$ is an isomorphism for all $1 \le n' < n$. Since π is a surjection, to prove that π is an isomorphism it is enough to construct a homomorphism $\phi\colon \mathrm{IA}_n \to \mathcal{Q}_n$ such that $\phi \circ \pi = \mathrm{id}$. Propositions 3.6 and 3.10 show that the action of IA_n on $\widehat{\mathcal{B}}_n$ satisfies the conditions of Theorem 3.1, so

$$\mathrm{IA}_n \cong \Big(\mathop{\ast}_{[\![z]\!] \in (\widehat{\mathcal{B}}_n)^{(0)}} (\mathrm{IA}_n)_{[\![z]\!]} \Big)/R,$$

where R is the normal closure of the edge and conjugation relators. The construction of ϕ will have two steps. First, we will use the action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ provided by Proposition 3.13 to construct a map

$$\widetilde{\phi}\colon \mathop{\ast}_{[\![z]\!] \in (\widehat{\mathcal{B}}_n)^{(0)}} (\mathrm{IA}_n)_{[\![z]\!]} \to \mathcal{Q}_n.$$

Second, we will show that $\widetilde{\phi}$ takes the edge and conjugation relators to 1, and thus induces a map $\phi\colon \mathrm{IA}_n \to \mathcal{Q}_n$. We will close by verifying that $\phi \circ \pi = \mathrm{id}$.

Construction of $\widetilde{\phi}$ To construct $\widetilde{\phi}$, we must construct a map

$$\widetilde{\phi}_{[\![z]\!]}\colon (\mathrm{IA}_n)_{[\![z]\!]} \to \mathcal{Q}_n$$

for each vertex $[\![z]\!]$ of $\widehat{\mathcal{B}}_n$. Recalling that $\{x_1, \dots, x_n\}$ is our fixed free basis for F_n, we begin with the vertex $[\![x_n]\!]$. In the following claim, we will use the notation $(S_{\mathrm{IA}}(n))_{[\![x_n]\!]}$ and $(\mathcal{Q}_n)_{[\![x_n]\!]}$ introduced in Section 3.4.

Claim 1 *The restriction of π to $(\mathcal{Q}_n)_{[\![x_n]\!]}$ is an isomorphism onto $(\mathrm{IA}_n)_{[\![x_n]\!]}$.*

Proof Proposition 3.11 implies that natural map $\pi|_{(\mathcal{Q}_n)_{[\![x_n]\!]}}\colon (\mathcal{Q}_n)_{[\![x_n]\!]} \to (\mathrm{IA}_n)_{[\![x_n]\!]}$ is surjective, since the generators from that proposition (with $k=1$) are in the image.

Our inductive hypothesis says that the map $\pi|_{\mathcal{Q}_{n-1}}\colon \mathcal{Q}_{n-1} \to \mathrm{IA}_{n-1}$ is an isomorphism. Recall from Section 3.5 that $(\mathrm{IA}_n)_{[\![x_n]\!]} = \mathcal{K}_{n-1,1} \rtimes \mathrm{IA}_{n-1}$, where the projection $(\mathrm{IA}_n)_{[\![x_n]\!]} \to \mathrm{IA}_{n-1}$ is the one induced by the map $F_n \to F_{n-1}$ that quotients out by the normal closure of x_n, and $\mathcal{K}_{n-1,1}$ is the kernel of this projection. The composition

$$(\mathcal{Q}_n)_{[\![x_n]\!]} \to (\mathrm{IA}_n)_{[\![x_n]\!]} \to \mathrm{IA}_{n-1} \xrightarrow[\cong]{\pi|^{-1}_{\mathcal{Q}_{n-1}}} \mathcal{Q}_{n-1}$$

is a well-defined homomorphism. It is a composition of surjective maps, and is therefore surjective. We define $\mathfrak{K}_{n-1,1}$ to be the kernel of this composition of maps.

The restriction of π to $\mathfrak{K}_{n-1,1}$ has image in $\mathcal{K}_{n-1,1}$ since the map $(\mathcal{Q}_n)_{[\![x_n]\!]} \to \mathcal{Q}_{n-1}$ factors through $(\mathrm{IA}_n)_{[\![x_n]\!]} \to \mathrm{IA}_{n-1}$. Proposition 3.11 says that $\mathcal{K}_{n-1,1}$ is generated by the set

$$\begin{aligned} S_{\mathcal{K}}(n) := \{&C_{x_n,x_a}, C_{x_a,x_n} \mid 1 \le a < b\} \\ &\cup \{M_{x_a^\alpha,[x_b^\beta,x_n^\gamma]}, M_{x_a^\alpha,[x_n^\gamma,x_b^\beta]} \mid 1 \le a,b < n \text{ distinct}, \alpha,\beta,\gamma \in \{\pm 1\}\}. \end{aligned}$$

Since these generators are contained in $\mathfrak{K}_{n-1,1}$, the map $\mathfrak{K}_{n-1,1} \to \mathcal{K}_{n-1,1}$ is surjective. Further, Proposition 3.14 gives us a left inverse to $\pi|_{\mathfrak{K}_{n-1,1}}$. We conclude that $\pi|_{\mathfrak{K}_{n-1,1}}$ is an isomorphism $\mathfrak{K}_{n-1,1} \cong \mathcal{K}_{n-1,1}$. We note that existence of this isomorphism is a deceptively difficult part of the proof, and it is the main consequence that we draw from [17].

Summing up, we have a commutative diagram of short exact sequences as follows:

$$\begin{tikzcd} 1 \arrow[r] & \mathfrak{K}_{n-1,1} \arrow[r] \arrow[d, "\cong"] & (\mathcal{Q}_n)_{[\![x_n]\!]} \arrow[r] \arrow[d] & \mathcal{Q}_{n-1} \arrow[r] \arrow[d, "\cong"] & 1 \\ 1 \arrow[r] & \mathcal{K}_{n-1,1} \arrow[r] & (\mathrm{IA}_n)_{[\![x_n]\!]} \arrow[r] & \mathrm{IA}_{n-1} \arrow[r] & 1 \end{tikzcd}$$

The five lemma therefore says that the projection map $(\mathcal{Q}_n)_{[\![x_n]\!]} \to (\mathrm{IA}_n)_{[\![x_n]\!]}$ is an isomorphism, as desired. □

Claim 1 implies that we can define a map $\widetilde{\phi}_{[\![x_n]\!]}\colon (\mathrm{IA}_n)_{[\![x_n]\!]} \to \mathcal{Q}_n$ via the formula $\widetilde{\phi}_{[\![x_n]\!]} = (\pi|_{(\mathcal{Q}_n)_{[\![x_n]\!]}})^{-1}$.

Now consider a general vertex $[\![z]\!]$ of $\widehat{\mathcal{B}}_n$. Here we will use the action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ provided by Proposition 3.13. The group $\mathrm{Aut}(F_n)$ acts transitively on the set of primitive elements of F_n, so there exists some $\nu \in \mathrm{Aut}(F_n)$ such that $\nu(x_n) = z$. We then define a map $\widetilde{\phi}_{[\![z]\!]}\colon (\mathrm{IA}_n)_{[\![z]\!]} \to \mathcal{Q}_n$ via the formula

$$\widetilde{\phi}_{[\![z]\!]}(\eta) = \nu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu^{-1}\eta\nu) \quad (\eta \in (\mathrm{IA}_n)_{[\![z]\!]}).$$

This appears to depend on the choice of ν, but the following claim says that this choice does not matter.

Claim 2 *The map $\widetilde{\phi}_{[\![z]\!]}(\eta)$ does not depend on the choice of ν.*

Proof Assume that $\nu_1, \nu_2 \in \mathrm{Aut}(F_n)$ both satisfy $\nu_i(x_n) = z$, and consider some $\eta \in (\mathrm{IA}_n)_{[\![z]\!]}$. Our goal is to prove that

$$\nu_1 \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu_1^{-1}\eta\nu_1) = \nu_2 \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu_2^{-1}\eta\nu_2). \tag{6}$$

Define $\mu = \nu_1^{-1}\nu_2$ and $\omega = \nu_2^{-1}\eta\nu_2$, so $\mu \in (\mathrm{Aut}(F_n))_{[\![x_n]\!]}$ and $\omega \in (\mathrm{IA}_n)_{[\![x_n]\!]}$. We will first prove that

$$\widetilde{\phi}_{[\![x_n]\!]}(\mu\omega\mu^{-1}) = \mu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\omega). \tag{7}$$

To see this, observe first that by construction both $\widetilde{\phi}_{[\![x_n]\!]}(\mu\omega\mu^{-1})$ and $\widetilde{\phi}_{[\![x_n]\!]}(\omega)$ lie in $(\mathcal{Q}_n)_{[\![x_n]\!]}$. The third part of Proposition 3.13 implies that $\mu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\omega)$ also lies in $(\mathcal{Q}_n)_{[\![x_n]\!]}$. Claim 1 says that $\pi|_{(\mathcal{Q}_n)_{[\![x_n]\!]}}$ is injective, so to prove (7), it is thus enough to prove that $\widetilde{\phi}_{[\![x_n]\!]}(\mu\omega\mu^{-1})$ and $\mu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\omega)$ have the same image under π. This follows from the calculation

$$\pi(\widetilde{\phi}_{[\![x_n]\!]}(\mu\omega\mu^{-1})) = \mu\omega\mu^{-1} = \mu\pi(\widetilde{\phi}_{[\![x_n]\!]}(\omega))\mu^{-1} = \pi(\mu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\omega)),$$

where the first two equalities follow from the fact that $\pi \circ \widetilde{\phi}_{[\![x_n]\!]} = \mathrm{id}$ and the third follows from the first conclusion of Proposition 3.13.

We now verify (6) as follows:

$$\nu_1 \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu_1^{-1}\eta\nu_1) = \nu_1 \cdot \widetilde{\phi}_{[\![x_n]\!]}(\mu\omega\mu^{-1}) = \nu_1\mu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\omega) = \nu_2 \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu_2^{-1}\eta\nu_2). \qquad \square$$

This completes the construction of $\widetilde{\phi}$.

Some naturality properties Before we study the edge and conjugation relators, we first need to verify the following two naturality properties of $\widetilde{\phi}$. Starting now we will use the notation which was introduced in Section 2: for a vertex $[\![z]\!]$ of $\widehat{\mathcal{B}}_n$ and $\eta \in \mathrm{IA}_n$ satisfying $\eta([\![z]\!]) = [\![z]\!]$, we will denote η, considered as an element of

$$(\mathrm{IA}_n)_{[\![z]\!]} < \mathop{\ast}_{[\![z]\!] \in (\widehat{\mathcal{B}}_n)^{(0)}} (\mathrm{IA}_n)_{[\![z]\!]},$$

by $\eta_{[\![z]\!]}$.

Claim 3 *The following two identities hold:*

- *Let* $1 \le a, b \le n$ *be distinct and* $1 \le i \le n$ *be arbitrary. Then* $\widetilde{\phi}((C_{x_a,x_b})_{[\![x_i]\!]}) = \mathfrak{C}_{x_a,x_b}$.
- *Let* $1 \le a, b, c \le n$ *be distinct, let* $\alpha, \beta, \gamma \in \{\pm 1\}$ *be arbitrary, and let* $1 \le i \le n$ *be such that* $i \ne a$. *Then* $\widetilde{\phi}((M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]})_{[\![x_i]\!]}) = \mathfrak{M}_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}$.

Proof The proofs of the two identities are similar; we will deal with the first and leave the second to the reader. It is clear from the construction that $\widetilde{\phi}((C_{x_a,x_b})_{[\![x_n]\!]}) = \mathfrak{C}_{x_a,x_b}$. For $1 \le i < n$, we have $P_{i,n}(x_n) = x_i$, and thus by definition we have

$$\begin{aligned}\widetilde{\phi}((C_{x_a,x_b})_{[\![x_i]\!]}) &= P_{i,n} \cdot \widetilde{\phi}_{[\![x_n]\!]}(P_{i,n}^{-1} C_{x_a,x_b} P_{i,n}) = P_{i,n} \cdot \widetilde{\phi}_{[\![x_n]\!]}(C_{P_{i,n}^{-1}(x_a),\, P_{i,n}^{-1}(x_b)}) \\ &= P_{i,n} \cdot \mathfrak{C}_{P_{i,n}^{-1}(x_a),\, P_{i,n}^{-1}(x_b)} = \mathfrak{C}_{x_a,x_b};\end{aligned}$$

here the last equality follows from the first part of Proposition 3.13 and the definition of the endomorphisms in Section 2. □

Claim 4 *Let* $[\![z]\!]$ *be a vertex of* $\widehat{\mathcal{B}}_n$. *Then for* $\eta \in (\mathrm{IA}_n)_{[\![z]\!]}$ *we have* $\pi(\widetilde{\phi}_{[\![z]\!]}(\eta)) = \eta$.

Proof Pick $\nu \in \mathrm{Aut}(F_n)$ such that $\nu(x_n) = z$. Then

$$\pi(\widetilde{\phi}_{[\![z]\!]}(\eta)) = \pi(\nu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu^{-1}\eta\nu)) = \nu\pi(\widetilde{\phi}_{[\![x_n]\!]}(\nu^{-1}\eta\nu))\nu^{-1} = \nu\nu^{-1}\eta\nu\nu^{-1} = \eta;$$

here the second equality uses the first part of Proposition 3.13. □

The edge and conjugation relators We now check that $\widetilde{\phi}$ takes the edge and conjugation relators to 1.

Claim 5 (edge relators) *If* e *is an edge of* $\widehat{\mathcal{B}}_n$ *with endpoints* $[\![z]\!]$ *and* $[\![z']\!]$ *and* $\eta \in (\mathrm{IA}_n)_e$, *then* $\widetilde{\phi}(\eta_{[\![z]\!]}\eta_{[\![z']\!]}^{-1}) = 1$.

Proof We first consider the special case where $z = x_n$ and $z' = x_{n-1}$. Proposition 3.11, with $k = 2$, states that $(\mathrm{IA}_n)_{[\![x_{n-1}]\!],[\![x_n]\!]}$ is generated by

$$(8)\quad \{C_{x_a,x_b} \mid 1 \le a, b \le n \text{ distinct}\} \cup \{M_{x_a,[x_b,x_c]} \mid 1 \le a, b, c \le n \text{ distinct}, a \ne n-1, n\}.$$

Claim 3 implies that for all elements ω in (8), we have $\widetilde{\phi}(\omega_{[\![x_{n-1}]\!]}) = \widetilde{\phi}(\omega_{[\![x_n]\!]})$. It follows that for all $\eta \in (\mathrm{IA}_n)_{[\![x_{n-1}]\!],[\![x_n]\!]}$ we have $\widetilde{\phi}(\eta_{[\![x_{n-1}]\!]}) = \widetilde{\phi}(\eta_{[\![x_n]\!]})$, as desired.

We now turn to general edges e with endpoints $[\![z]\!]$ and $[\![z']\!]$ and $\eta \in (\mathrm{IA}_n)_e$. There exists some $\nu \in \mathrm{Aut}(F_n)$ such that $\nu(x_n) = z$ and $\nu(x_{n-1}) = z'$, and hence $\nu P_{n-1,n}(x_n) = z'$. Setting $\eta' = \nu^{-1}\eta\nu \in (\mathrm{IA}_n)_{[\![x_{n-1}]\!],[\![x_n]\!]}$, we have

$$\begin{aligned}\widetilde{\phi}(\eta_{[\![z]\!]}) &= \nu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu^{-1}\eta\nu) = \nu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\eta'), \\ \widetilde{\phi}(\eta_{[\![z']\!]}) &= \nu P_{n-1,n}\widetilde{\phi}_{[\![x_n]\!]}(P_{n-1,n}^{-1}\nu^{-1}\eta\nu P_{n-1,n}) = \nu \cdot \widetilde{\phi}_{[\![x_{n-1}]\!]}(\eta').\end{aligned}$$

By the previous paragraph, we have $\widetilde{\phi}_{[\![x_n]\!]}(\eta') = \widetilde{\phi}_{[\![x_{n-1}]\!]}(\eta')$, so we conclude that $\widetilde{\phi}(\eta_{[\![z]\!]}) = \widetilde{\phi}(\eta_{[\![z']\!]})$, as desired. □

Claim 6 (conjugation relators) *If $[\![z]\!]$ and $[\![z']\!]$ are vertices of $\widehat{\mathcal{B}}_n$ and $\eta \in (\mathrm{IA}_n)_{[\![z]\!]}$ and $\omega \in (\mathrm{IA}_n)_{[\![z']\!]}$, then $\widetilde{\phi}(\omega_{[\![z']\!]}\eta_{[\![z]\!]}\omega^{-1}_{[\![z']\!]}(\omega\eta\omega^{-1})_{[\![\omega(z)]\!]}) = 1$.*

Proof Choose $\nu \in \mathrm{Aut}(F_n)$ such that $\nu(x_n) = z$. We then have

$$\begin{aligned}\widetilde{\phi}(\omega_{[\![z']\!]}\eta_{[\![z]\!]}\omega^{-1}_{[\![z']\!]}) &= \widetilde{\phi}_{[\![z']\!]}(\omega)\widetilde{\phi}_{[\![z]\!]}(\eta)\widetilde{\phi}_{[\![z']\!]}(\omega)^{-1} = \pi(\widetilde{\phi}_{[\![z']\!]}(\omega)) \cdot \widetilde{\phi}_{[\![z]\!]}(\eta) \\ &= \omega \cdot \widetilde{\phi}_{[\![z]\!]}(\eta) = \omega\nu \cdot \widetilde{\phi}_{[\![x_n]\!]}(\nu^{-1}\eta\nu) \\ &= \omega\nu \cdot \widetilde{\phi}_{[\![x_n]\!]}((\omega\nu)^{-1}\omega\eta\omega^{-1}(\omega\nu)) = \widetilde{\phi}((\omega\eta\omega^{-1})_{[\![\omega(z)]\!]}),\end{aligned}$$

as desired. The second equality follows from the third part of Proposition 3.13, the third equality follows from Claim 4, and the remainder of the equalities are straightforward applications of the definitions. □

Claims 5 and 6 imply that $\widetilde{\phi}$ descends to a homomorphism $\phi\colon \mathrm{IA}_n \to \mathcal{Q}_n$.

We have an inverse To complete the proof, it remains to prove the following:

Claim 7 *We have $\phi \circ \pi = \mathrm{id}$.*

Proof Claim 3 implies that this holds for the generators of $\mathcal{Q}_n$. □

This completes the proof of Theorem E. □

5 Computations for the L–presentation

This section contains the postponed proofs of Propositions 2.1, 2.3, 3.13 and 3.14. These proofs are done with the aid of a computer. We will discuss our computational framework in Section 5.1 and then prove the propositions in Sections 5.2–5.4.

We will use the following notation throughout the rest of the paper. Let $S_{\mathrm{Aut}}(n)^*$ denote the free monoid on the set $S_{\mathrm{Aut}}(n)$. In Section 2, we defined a function $\theta\colon S_{\mathrm{Aut}}(n) \to \mathrm{End}(F(S_{\mathrm{IA}}(n)))$. This naturally extends to a function

$$\theta\colon S_{\mathrm{Aut}}(n)^* \to \mathrm{End}(F(S_{\mathrm{IA}}(n))).$$

5.1 Computational framework

As we discussed in the introduction, we use the GAP system to mechanically verify the large number of equations we have to check. These verifications are in the file `h2ia.g`, available as an online supplement.

We use GAP's built-in functionality to model F_n as a free group on the eight generators `xa`, `xb`, `xc`, `xd`, `xe`, `xf`, `xg` and `y`. Since our computations never involve more than 8 variables, computations in this group suffice to show that our computations hold in general.

Elements of the sets $S_{\mathrm{Aut}}(n)$ and $S_{\mathrm{IA}}(n)$ are parametrized over basis elements from F_n and their inverses, so we model these sets using lists. For example, we model the generator M_{x_a,x_b} as the list `["M",xa,xb]`, C_{y,x_a} as `["C",y,xa]`, and $M_{x_a^{-1},[y,x_c]}$ as `["Mc",xa^-1,y,xc]`. We model $P_{a,b}$ as `["P",xa,xb]` and I_a as `["I",xa]`. The examples should make clear: the first entry in the list is a string key `"M"`, `"C"`, `"Mc"`, `"P"` or `"I"`, indicating whether the list represents a transvection, conjugation move, commutator transvection, swap or inversion. The parameters given as subscripts in the generator are then the remaining elements of the list, in the same order.

GAP's built-in free group functionality expects the basis elements to be variables, not lists, so we do not use it to model $S_{\mathrm{Aut}}(n)^*$ and $F(S_{\mathrm{IA}}(n))$. We model inverses of generators as follows: the inverse of `["M",xa,xb]` is `["M",xa,xb^-1]` and the inverse of `["C",xa,xb]` is `["C",xa,xb^-1]`, but the inverse of `["Mc",xa,xb,xc]` is `["Mc",xa,xc,xb]`. Swaps and inversions are their own inverses. Technically, this means that we are not really modeling $S_{\mathrm{Aut}}(n)^*$ and $F(S_{\mathrm{IA}}(n))$; instead we model structures where the order relations for swaps and inversions and the relation (R0) for inverting commutator transvections are built in. This is not a problem because our verifications always show that certain formulas are trivial modulo our relations, and we can always apply the (R0) and order relations as needed.

We model words in $S_{\mathrm{Aut}}(n)^*$ and $F(S_{\mathrm{IA}}(n))$ as lists of generators and inverse generators. The empty word `[]` represents the trivial element. We wrote several functions in `h2ia.g` that perform common tasks on words. The function `pw` takes any number of words (reduced or not) as arguments and returns the freely reduced product of those words in the given order, as a single word. The function `iw` inverts its input word and the function `cyw` cyclically permutes its input word.

The function `iarel` outputs the relations $R^0_{\mathrm{IA}}(n)$. We introduce some extra relations for convenience. The function `exiarel` outputs these extra relations and the code generating the list `exiarelchecklist` derives the extra relations from the basic relations. The function `theta` takes in a word w in $S_{\mathrm{Aut}}(n)$ and a word v in $S_{\mathrm{IA}}(n)$, and returns $\theta(w)(v)$. In addition to the functions described here, we often define simple macros for carrying out the verifications.

The function `applyrels` is particularly useful, because it inserts multiple relations into a word. It takes two inputs: a starting word and a list of words with placement indicators. The function recursively inserts the first word from the list in the starting word at the given position, reduces the word, and then calls itself with the new word as the starting word and with the same list of insertions, with the first dropped.

For example, the following command appears in the justification of Proposition 3.13:

```
applyrels(
     pw(
          theta([["M",xa,xb],["M",xa,xb^-1]], [["Mc",xb,xa,xe]]),
          [["Mc",xb,xe,xa]]
     ),
     [
          [6,iarel(5,[xa,xb,xe])],
          [6,iw(exiarel(1,[xb^-1,xe,xa]))],
          [2,iw(iarel(4,[xe,xa,xb]))],
          [2,iarel(6,[xb^-1,xe^-1,xa])],
          [2,iw(iarel(5,[xb,xe,xa]))]
     ]
)
```

This tells the GAP system to compute the effect of $\theta(M_{x_a,x_b}M_{x_a,x_b}^{-1})$ on $M_{x_b,[x_a,x_e]}$. Then it multiplies this by $M_{x_b,[x_e,x_a]}$, the inverse of $M_{x_b,[x_a,x_e]}$. It then freely reduces this word. The system inserts a version of (R5) after the sixth letter in this word, and reduces the result to a new word. Then it inserts the inverse of one of the extra relations after the sixth letter in the new word and reduces it. It continues with inserting relations and reducing the resulting expressions, inserting instances of (R4), (R5) and (R6). Since the entire expression evaluates to `[]`, we have expressed

$$\theta(M_{x_a,x_b}M_{x_a,x_b}^{-1})(M_{x_b,[x_a,x_e]})\cdot M_{x_b,[x_e,x_a]}$$

as a product of relations in $\mathcal{Q}_n$. In any example like this, an interested reader can reproduce our reduction process by removing all the list entries from the second input of the `applyrels` call, and then adding them back in one at a time, evaluating after each one.

5.2 Verifying the map to IA_n

First we prove Proposition 2.1, which states that our relations $R_{\mathrm{IA}}^0(n)$ hold in IA_n.

Proof of Proposition 2.1 The code generating the list `verifyiarel` generates examples of all the relations in $R_{\mathrm{IA}}^0(n)$, with all allowable configurations of coincidences between the subscripts on the generators. It converts each of these relations into automorphisms of F_n and evaluates them on a basis for F_n, returning `true` if all basis elements are unchanged. We evaluate on a fixed finite-rank free group, but since the basic relations involve at most six generators, those evaluations suffice to show the result in general. Since `verifyiarel` evaluates to a list of `true`, this means that all these relations are true. □

Next we prove Proposition 2.3, which states that θ acts by conjugation when evaluated on generators (see (3)).

Proof of Proposition 2.3 The code generating the list `thetavsconjaut` goes through all possible configurations for a pair of generators s from $S_{\mathrm{Aut}}(n)$ and t from $S_{\mathrm{IA}}(n)$, evaluates $\theta(s)(t)$ as a product of generators, and then evaluates both $\theta(s)(t)$ and sts^{-1} on a basis for F_n. It returns `true` when both have the same effect on all basis elements. Since `thetavsconjaut` evaluates to a list of `true`, the proposition holds. □

5.3 Verifying Proposition 3.13

Proof of Proposition 3.13 The action of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ is given by our substitution rule endomorphism map

$$\theta\colon S_{\mathrm{Aut}}(n) \to \mathrm{End}(F(S_{\mathrm{IA}}(n))).$$

First of all, it is clear that for each $s \in S_{\mathrm{Aut}}(n)$, the element $\theta(s)$ defines an endomorphism of $\mathcal{Q}_n$. This is because the subgroup of $F(S_{\mathrm{IA}}(n))$ normally generated by the relations of $\mathcal{Q}_n$ is invariant under $\theta(s)$ by the definition of $\mathcal{Q}_n$.

Next, we verify that $\theta(s)$ is an automorphism of $\mathcal{Q}_n$. If s is a swap or an inversion, then it is clear from the definition of θ that this is the case. In the code generating the list `thetainverselist`, we compute $\theta(s)(\theta(s^{-1})(t))t^{-1}$ for $s = M_{x_a,x_b}$ and for all possible configurations of t relative to s. In each case, we reduce it to the trivial word using relations for $\mathcal{Q}_n$. It is not hard to deduce that $\theta(s)(\theta(s^{-1})(t)) = t$ in $\mathcal{Q}_n$ for the remaining choices of $s = M_{x_a,x_b}^{-1}$, $M_{x_a^{-1},x_b}$ and $M_{x_a^{-1},x_b}^{-1}$, using the fact that it is true for $s = M_{x_a,x_b}$.

So this shows that θ defines an action

$$F(S_{\mathrm{Aut}}(n)) \to \mathrm{Aut}(\mathcal{Q}_n).$$

Now we need to verify that this action descends to an action of $\mathrm{Aut}(F_n)$. To show this, it is enough to show that for every relation r in a presentation for $\mathrm{Aut}(F_n)$, we have

$$\theta(r)(t) = t \quad \text{in } \mathcal{Q}_n, \tag{9}$$

for t taken from a generating set for $\mathcal{Q}_n$. To check this, we use the same version of Nielsen's presentation for $\mathrm{Aut}(F_n)$ that we used in [17, Theorem 5.5]. The generators are the same set $S_{\mathrm{Aut}}(n)$ we use here, and the relations fall into five classes (N1)–(N5). Relations (N1) are sufficient for the subgroup generated by swaps and inversions, and (N2) are relations indicating how to conjugate transvections by swaps and inversions. It is an exercise to see that (9) holds for relations of class (N1) and (N2). Relations (N3)–(N5) are more complicated relations. For each of these, we compute $\theta(r)(t)t^{-1}$ on generators t (for t with enough configurations of subscripts to include a generating set) and reduce the resulting expressions to 1 using relations from $\mathcal{Q}_n$. These computations are given in the code generating the lists `thetaN3list`, `thetaN4list` and `thetaN5list`. Since these evaluate to lists of the trivial word, this verifies (9). We have shown that the action of an element of $\mathrm{Aut}(F_n)$ on $\mathcal{Q}_n$ does not depend on the word in $F(S_{\mathrm{Aut}}(n))$ we use to represent it.

Since we have shown that θ defines an action, now we can check the three properties asserted in Proposition 3.13. We have already verified the first point (we took the definition of the action to agree with it). To verify the second point, we need to check that for $w, s \in S_{\mathrm{IA}}(n)$, there is $\widetilde{w} \in F(S_{\mathrm{Aut}}(n))$ representing w with

$$\theta(\widetilde{w})(s) = wsw^{-1} \quad \text{in } \mathcal{Q}_n. \tag{10}$$

In fact, it is enough to verify this for w and s in a smaller generating set, and the generating set that s is taken from may depend on w. In the code generating `thetaconjrellist`, for each choice of w from $S_{\mathrm{MA}}(n)$, we lift w to $\widetilde{w} \in F(S_{\mathrm{Aut}}(n))$, and for several configurations of subscripts in the generator s, we reduce the element $\theta(\widetilde{w})(s)ws^{-1}w^{-1}$ to the identity using relations from $\mathcal{Q}_n$. We use enough configurations of subscripts in s to cover all cases for s in a generating set (a conjugate of $S_{\mathrm{MA}}(n)$).

To check the third point, we use the generating set $(S_{\mathrm{Aut}}(n))_{[\![x_n]\!]}$ for $(\mathrm{Aut}(F_n))_{[\![x_n]\!]}$ mentioned in the proof of Proposition 3.11 above, namely

$$\begin{aligned}\{M_{x_a^\alpha, x_b} \mid 1 \le a \le n-1,\ 1 \le b \le n,\ \alpha = \pm 1,\ a \ne b\}\\ \cup \{P_{a,b} \mid 1 \le a, b \le n,\ a \ne b,\ a \ne n\}\\ \cup \{I_a \mid 1 \le a \le n\} \cup \{C_{x_n, x_a} \mid 1 \le a \le n-1\}.\end{aligned}$$

We need to check that for each of these generators, there is $w \in F(S_{\mathrm{Aut}}(n))$ representing it with $\theta(w)(s)$ in $(\mathcal{Q}_n)_{[\![x_n]\!]}$ (really, that $\theta(w)(s)$ is equal in $\mathcal{Q}_n$ to an element of $(\mathcal{Q}_n)_{[\![x_n]\!]}$). This is clear from the definition of θ for w a swap or inversion. It can be verified for $w = M_{x_a^\alpha, x_b}$ by inspecting Table 3. For w representing C_{x_n, x_a}, the fact that $\theta(w)(s) \in (\mathcal{Q}_n)_{[\![x_n]\!]}$ follows from the second point in this proposition, since $C_{x_n, x_a} \in \mathrm{IA}_n$. □

5.4 Verifying Proposition 3.14

Here we prove Proposition 3.14. We recall the statement: $\mathcal{K}_{n-1,1}$ is the kernel of the natural map $(\mathrm{IA}_n)_{[\![x_n]\!]} \to \mathrm{IA}_{n-1}$, and the proposition asserts that the inclusion $\mathcal{K}_{n-1,1} \hookrightarrow \mathrm{IA}_n$ factors as the composition of a map $\mathcal{K}_{n-1,1} \to \mathcal{Q}_n$ with the projection $\mathcal{Q}_n \to \mathrm{IA}_n$.

The proof uses the finite L–presentation for $\mathcal{K}_{n-1,1}$ from [17]. We note that [17, Theorem D] asserts the existence of such a presentation, and [17, Theorem 6.2] gives the precise statement that we use in the computations. Since this L–presentation is in fact a presentation for $\mathcal{K}_{n-1,1}$, we use the same notation for $\mathcal{K}_{n-1,1}$ as a subset of IA_n and $\mathcal{K}_{n-1,1}$ as the group given by this presentation.

We do not reproduce the L–presentation here, but instead we describe some of its features. Its finite generating set is

$$\begin{aligned} S_{\mathcal{K}}(n) = \{M_{x_a^\alpha,[x_n^\epsilon,x_b^\beta]} \mid 1 \le a,b \le n-1,\ a \ne b,\ \alpha,\beta,\epsilon \in \{1,-1\}\} \\ \cup \{C_{x_n,x_a} \mid 1 \le a \le n-1\} \cup \{C_{x_a,x_n} \mid 1 \le a \le n-1\}. \end{aligned}$$

The substitution endomorphisms of the L–presentation for $\mathcal{K}_{n-1,1}$ are indexed by a finite generating set $(S_{\mathrm{Aut}}(n))_{[\![x_n]\!]}$ for $(\mathrm{Aut}(F_n))_{[\![x_n]\!]}$. The endomorphisms themselves are the image of a map

$$\phi\colon (S_{\mathrm{Aut}}(n))_{[\![x_n]\!]} \to \mathrm{End}(F(S_{\mathcal{K}}(n))).$$

Proof of Proposition 3.14 Since $S_{\mathcal{K}}(n)$ is a subset of $S_{\mathrm{IA}}(n)$, we map $\mathcal{K}_{n-1,1}$ to $\mathcal{Q}_n$ by sending each generator to the generator of the same name. To verify that this map on generators extends to a well-defined map of groups, we need to check that each defining relation from $\mathcal{K}_{n-1,1}$ maps to the trivial element of $\mathcal{Q}_n$. Since $\mathcal{K}_{n-1,1}$ is given by a L–presentation, we proceed as follows:

(1) We check that each of the basic relations from $\mathcal{K}_{n-1,1}$ maps to the trivial element of $\mathcal{Q}_n$.

(2) We check that for $s \in (S_{\mathrm{Aut}}(n))_{[\![x_n]\!]}$ and $t \in S_{\mathcal{K}}(n)$, we have

$$\phi(s)(t) = \theta(s)(t) \quad \text{in } \mathcal{Q}_n,$$

where we use $(S_{\mathrm{Aut}}(n))_{[\![x_n]\!]} \subset S_{\mathrm{Aut}}(n)$ to plug s into θ, and we interpret both expressions in $\mathcal{Q}_n$ using $F(S_{\mathcal{K}}(n)) \subset F(S_{\mathrm{IA}}(n))$.

The first point is verified in the code generating the list `kfromialist`. The function `krel` produces the basic relations from $\mathcal{K}_{n-1,1}$, and we reduce each relation to the identity by applying relations from $\mathcal{Q}_n$. The second point is verified in the code generating the list `thetavsphlist`. For each choice of pairs of generators, we reduce the difference of θ and ϕ using relations from $\mathcal{Q}_n$.

With these two points verified, one can easily check by induction that every extended relation (starting with a basic relation, and applying any sequence of rewriting rules) maps to the identity element in $\mathcal{Q}_n$. □

6 Generators for $\mathrm{H}_2(\mathrm{IA}_n)$

In this section, we prove Theorem A, which asserts that there exists a finite subset of $\mathrm{H}_2(\mathrm{IA}_n)$ whose $\mathrm{GL}_n(\mathbb{Z})$–orbit spans $\mathrm{H}_2(\mathrm{IA}_n)$. In fact, we gave an explicit list of generators in Table 1; each generator is of the form $\mathfrak{h}_r$ for a commutator relation r. This list is reproduced in Table 4, which also introduces the notation $h_i(\cdot) \in \mathrm{H}_2(\mathrm{IA}_n)$ for the associated elements of homology (this notation will be used during the calculations in Section 7, though we will not use it in this section). The following theorem asserts that this list is complete; it is a more precise form of Theorem A and will be the main result of this section.

Theorem 6.1 *Fix $n \geq 2$. Let $S_H(n)$ be the set of commutator relators in Table 4. Then the $\mathrm{GL}_n(\mathbb{Z})$–orbit of the set $\{\mathfrak{h}_r \mid r \in S_H(n)\}$ spans $\mathrm{H}_2(\mathrm{IA}_n)$.*

Before proving Theorem 6.1, we will use it to derive Theorem B.

Proof of Theorem B Recall that this theorem asserts that for $n \geq 6$, the $\mathrm{GL}_{n+1}(\mathbb{Z})$–orbit of the image of the natural map $\mathrm{H}_2(\mathrm{IA}_n) \to \mathrm{H}_2(\mathrm{IA}_{n+1})$ spans $\mathrm{H}_2(\mathrm{IA}_{n+1})$. Let $S_{n+1} \subset \mathrm{GL}_{n+1}(\mathbb{Z})$ be the subgroup consisting of permutation matrices. By inspecting Table 4, it is clear that the S_{n+1}–orbit of the image of $\{\mathfrak{h}_r \mid r \in S_H(n)\} \subset \mathrm{H}_2(\mathrm{IA}_n)$ in $\mathrm{H}_2(\mathrm{IA}_{n+1})$ is $\{\mathfrak{h}_r \mid r \in S_H(n+1)\}$. This uses the fact that $n \geq 6$, since the commutator relations in $S_H(n)$ use generators involving at most six basis elements. □

We now turn to the proof of Theorem 6.1. We start by introducing some notation. Let $F = F(S_{\mathrm{IA}}(n))$ and let $R \subset F$ denote the full set of relations of IA_n, so $\mathrm{IA}_n = F/R$. Define $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)} = R/[F,R]$, and for $r \in R$ denote by $\|r\|$ the associated element of $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$. There is a natural map $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)} \to F^{\mathrm{ab}}$, and the starting point for our proof is the following lemma. In it, recall from the beginning of Section 5 that $S_{\mathrm{Aut}}(n)^*$ is the free monoid on the set $S_{\mathrm{Aut}}(n)$.

Lemma 6.2 *The group $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$ is an abelian group which is generated by*

$$\{\|\theta(w)(r)\| \mid w \in S_{\mathrm{Aut}}(n)^* \text{ and } r \text{ is one of the relations (R0)–(R9) from Table 2}\}.$$

Also, we have $\mathrm{H}_2(\mathrm{IA}_n) = \ker(\widetilde{\mathrm{H}_2(\mathrm{IA}_n)} \to F^{ab})$.

Proof The group $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$ is abelian since $[R,R] \subset [F,R]$. For $v \in F$ and $r \in R$, we have $[v,r] \in [F,R]$, so $\|vrv^{-1}\| = \|r\|$. The indicated generating set for $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$ thus follows from Theorem E. As for the final statement of the lemma, we follow one of the standard proofs of Hopf's formula [10]. The 5–term exact sequence in group homology associated to the short exact sequence

$$0 \to R \to F \to \mathrm{IA}_n \to 0$$

is

$$\mathrm{H}_2(F) \to \mathrm{H}_2(\mathrm{IA}_n) \to R/[F,R] \to \mathrm{H}_1(F) \to \mathrm{H}_1(\mathrm{IA}_n) \to 0.$$

Since F is free, we have $\mathrm{H}_2(F) = 0$, and the claim follows. □

Our goal will be to take an element of $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$ that happens to lie in $\mathrm{H}_2(\mathrm{IA}_n)$ and rewrite it as a sum of elements of the form

(11) $\{\|\theta(w)(r)\| \mid w \in S_{\mathrm{Aut}}(n)^*$ and r is one of the relations (H1)–(H9) of Table 4$\}$.

The relations (H1)–(H4) are the same as (R1)–(R4), and the relations (H5)–(H7) are the same as (R6)–(R9). The troublesome relations are (R0), (R5) and (R6), none of which lie in $\mathrm{H}_2(\mathrm{IA}_n)$. For $r \in R$, we have $\|r\| \in \mathrm{H}_2(\mathrm{IA}_n)$ if and only if the exponent-sum of each generator in $S_{\mathrm{IA}}(n)$ appearing in it is 0. For our problematic relations (R0), (R5) and (R6), the exponent-sum of all the conjugations moves is already 0, so we will only need to study the exponent-sums of the commutator transvections.

We begin with the following lemma, which will allow us to mostly ignore our rewriting rules $\theta(\cdot)$.

Lemma 6.3 *Consider $w \in S_{\mathrm{Aut}}(n)^*$, and let r be a relation of the form* (R0), (R5) *or* (R6). *Then $\|\theta(w)(r)\| = h + h'$, where h and h' are as follows:*

- *$h \in \mathrm{H}_2(\mathrm{IA}_n)$ is a sum of elements from* (11).
- *h' is a sum of elements of the form*

$$\{\|r\| \mid r \text{ is one of the relations (R0), (R5) and (R6) from Table 2}\}.$$

Proof We can use induction to reduce to the case where $w = s \in S_{\mathrm{Aut}}(n)$. The proof now is a combinatorial group-theoretic calculation: we will show how to rewrite $\theta(s)(r)$ as a product of relations of the desired form.

We start by dealing with the case where r is of the form (R0). Observe that

$$\theta(s)(M_{x_a^\alpha,[x_b^\beta,x_c^\gamma]}) \quad \text{and} \quad \theta(s)(M_{x_a^\alpha,[x_c^\gamma,x_b^\beta]})^{-1}$$

agree up to (R0), except in two cases. These are

$$\theta(M^{\beta}_{x_a^\alpha,x_b})(M_{x_b^{-\beta},[x_c^\gamma,x_d^\delta]}) = M_{x_a^\alpha,[x_c^\gamma,x_d^\delta]} M_{x_b^{-\beta},[x_c^\gamma,x_d^\delta]}$$

$$\theta(M^{\beta}_{x_a^\alpha,x_b})(M_{x_b^{\beta},[x_c^\gamma,x_d^\delta]}) = C^{\beta}_{x_a,x_b} M_{x_a^{-\alpha},[x_c^\gamma,x_d^\delta]} M_{x_b^{\beta},[x_c^\gamma,x_d^\delta]} C^{-\beta}_{x_a,x_b}.$$

(H1) $[C^{\beta}_{x_a,x_b}, C^{\delta}_{x_c,x_d}] = 1$, possibly with $b = d$.

(H2) $[M_{x_a^{\alpha},[x_b^{\beta},x_c^{\gamma}]}, M_{x_d^{\delta},[x_e^{\epsilon},x_f^{\zeta}]}] = 1$, possibly with $\{b,c\} \cap \{e,f\} \neq \varnothing$ or $x_a^{\alpha} = x_d^{-\delta}$ as long as $x_a^{\alpha} \neq x_d^{\delta}$, $a \notin \{e,f\}$ and $d \notin \{b,c\}$.

(H3) $[C^{\beta}_{x_a,x_b}, M_{x_c^{\gamma},[x_d^{\delta},x_e^{\epsilon}]}] = 1$,
possibly with $b \in \{d,e\}$ if $c \notin \{a,b\}$ and $a \notin \{c,d,e\}$.

(H4) $[C^{\beta}_{x_c,x_b} C^{\beta}_{x_a,x_b}, C^{\alpha}_{x_c,x_a}] = 1$.

(H5) $[C^{-\gamma}_{x_a,x_c}, C^{-\delta}_{x_a,x_d}][C^{-\beta}_{x_a,x_b}, M_{x_b^{\beta},[x_c^{\gamma},x_d^{\delta}]}] = 1$.

(H6) $[M_{x_a^{\alpha},[x_b^{\beta},x_c^{\gamma}]}, M_{x_d^{\delta},[x_a^{\alpha},x_e^{\epsilon}]}][M_{x_d^{\delta},[x_a^{\alpha},x_e^{\epsilon}]}, M_{x_d^{\delta},[x_c^{\gamma},x_b^{\beta}]}] \times$
$[M_{x_d^{\delta},[x_c^{\gamma},x_b^{\beta}]}, C^{-\epsilon}_{x_d,x_e}] = 1$, possibly with $b = e$ or $c = e$.

(H7) $[M_{x_c^{\gamma},[x_a^{\alpha},x_d^{\delta}]}, C^{\beta}_{x_a,x_b}][C^{-\delta}_{x_c,x_d}, M_{x_c^{\gamma},[x_a^{\alpha},x_b^{\beta}]}][M_{x_c^{\gamma},[x_a^{\alpha},x_b^{\beta}]}, M_{x_c^{\gamma},[x_a^{\alpha},x_d^{\delta}]}] = 1$, possibly with $b = d$.

(H8) $[M_{x_a^{\alpha},[x_b^{\beta},x_c^{\gamma}]}, C^{\delta}_{x_a,x_d} C^{\delta}_{x_b,x_d} C^{\delta}_{x_c,x_d}] = 1$.

(H9) $[C^{\gamma}_{x_a,x_c} C^{\gamma}_{x_b,x_c}, C^{\beta}_{x_a,x_b} C^{\beta}_{x_c,x_b}][M_{x_a^{\alpha},[x_b^{\beta},x_c^{\gamma}]}, C^{\alpha}_{x_b,x_a} C^{\alpha}_{x_c,x_a}] = 1$.

Table 4: The set $S_H(n)$ of basic commutator relators such that the $\mathrm{GL}_n(\mathbb{Z})$–orbit of $\{\mathfrak{h}_r \mid r \in S_H(n)\}$ spans $\mathrm{H}_2(\mathrm{IA}_n)$. Distinct letters are assumed to represent distinct indices unless stated otherwise. We use the notation $h_i(\cdot)$, with inputs the appropriate x_a^{α}, x_b^{β}, x_c^{γ}, x_d^{δ}, x_e^{ϵ}, x_f^{ζ}, for the elements in $\mathrm{H}_2(\mathrm{IA}_n)$, which we use later.

In the first case,

$$\theta(M^{\beta}_{x_a^{\alpha},x_b})(M_{x_b^{-\beta},[x_d^{\delta},x_c^{\gamma}]})^{-1} = M_{x_b^{-\beta},[x_c^{\gamma},x_d^{\delta}]} M_{x_a^{\alpha},[x_c^{\gamma},x_d^{\delta}]}.$$

In the second case,

$$\theta(M^{\beta}_{x_a^{\alpha},x_b})(M_{x_b^{\beta},[x_d^{\delta},x_c^{\gamma}]})^{-1} = C^{\beta}_{x_a,x_b} M_{x_b^{\beta},[x_c^{\gamma},x_d^{\delta}]} M_{x_a^{-\alpha},[x_c^{\gamma},x_d^{\delta}]} C^{-\beta}_{x_a,x_b}.$$

In both cases, the two expressions differ by an application of (R2). This means that $\theta(s)$ of an (R0) relation can always be written using (R0) and (R2) relations.

Next we explain the computations that prove the lemma for (R5) and (R6) relations. These are in the list `rewritetheta(R5)(R6)`. In these computations we reduce $\theta(s)(r)$ to the trivial word where $s \in S_A^{\pm 1}$ and r is an (R5) or (R6) relation. These reductions may use any of the basic relations for IA_n, including (R5) and (R6) themselves, but notably may not use the images of (R5) and (R6) relations under θ. We may use the images of (R1)–(R4), (R7)–(R9), (H8) and (H9) under θ.

Despite these restrictions, we may use the extra relations from `exiarel` in these computations. Our relation `exiarel(3,[xa,xb,xc,xd])` is (H8), and the relation `exiarel(4,[xa,xb,xc,xd])` is equivalent to (H8) modulo the basic relations. Relation `exiarel(7,[xa,xb,xc])` is (H9), and relation `exiarel(6,[xa,xb,xc])` and relation `exiarel(8,[xa,xb,xc,xd])` are equivalent to (H9) modulo the basic relations. All the other `exiarel` relations can be derived without using images of (R5) or (R6) under θ. These facts can be verified by inspecting `exiarelchecklist`.

If s is a swap or an inversion move, then acting by $\theta(s)$ is always the same as acting on the parameters of the relation by s in the obvious way. Therefore the list `rewritetheta(R5)(R6)` only contains cases where s is a transvection.

We use several redundancies between different forms of the relations (R5) and (R6) to reduce the number of computations. Inverting the parameter x_b^β in (R5) (as it appears in Table 2) is the same as cyclically permuting the relation. Inverting the parameter x_a^α in (R6) is the same as applying a relation from (R2) to the original (R6) relation. Swapping the roles of x_b^β and x_c^γ in (R6) is the same as inverting and cyclically permuting the original (R6) relation and applying a relation from (R2).

We use the identity

$$C_{x_a,x_b}^{-\beta}\theta(M_{x_a^\alpha,x_b}^{\beta})(t)C_{x_a,x_b}^{\beta}=\theta(M_{x_a^{-\alpha},x_b}^{-\beta})(t),$$

which holds in IA_n for any $t\in\mathrm{IA}_n$. This is a consequence of Proposition 3.13. In particular, this means that we only need to consider one of $\theta(M_{x_a^\alpha,x_b^\beta})(r)$ and $\theta(M_{x_a^{-\alpha},x_b^{-\beta}})(r)$; if one is trivial then so is the other.

Since the computations in `rewritetheta(R5)(R6)` rewrite all configurations of $\theta(s)(r)$ for r an (R5) or (R6) relation, up to these reductions, this proves the lemma. □

The next lemma allows us to deal with certain combinations of (R5) and (R6) relations. The ordered triple of generators of F_n *involved* in a commutator transvection $M_{x_i^\alpha,[x_j^\beta,x_k^\gamma]}$ is (x_i,x_j,x_j). There are eight commutator transvections involving a given triple of generators.

Lemma 6.4 *Fix distinct* $1\le a,b,c\le n$*, and let* $w\in R\cap[F,F]$ *be a product of* (R5) *and* (R6) *relations whose commutator transvections involve only* (x_a,x_b,x_c)*, in order. Then* $\|w\|$ *can be written as a sum of elements of the form* $\|v\|$ *with* v *an* (H2) *relation.*

Proof Let F' be the subgroup of F generated by the eight commutator transvections involving (x_a,x_b,x_c) and the two conjugation moves $\{C_{x_a,x_b},C_{x_a,x_c}\}$, and let $R'\subset F'$ be the normal closure in F' of the (R5) and (R6) relations that can be written as products of elements of F'. We thus have $w\in R'\cap[F',F']$. Our first goal is to better understand $(R'\cap[F',F'])/[F',R']$ and $R'/[F',R']$. Consider the exact sequence of abelian groups

$$0 \to \frac{R' \cap [F', F']}{[F', R']} \to \frac{R'}{[F', R']} \to \frac{R'}{R' \cap [F', F']} \to 0.$$

We find generators for the first group in the sequence by considering a related exact sequence of free abelian groups.

Let $v_1, \dots, v_8$ be the eight commutator transvections in $S_{\mathrm{IA}}(n)$ that only involve (x_a, x_b, x_c), enumerated as in Table 5. Similarly, let $r_1, \dots, r_8$ be the eight (R5) relations in F' and let $r_9, \dots, r_{16}$ denote the eight (R6) relations lying in F', enumerated as in Table 5. Let A be the free abelian group freely generated by $r_1, \dots, r_{16}$, and let B be the free abelian group freely generated by $v_1, \dots, v_8$. We consider the map $A \to B$ that counts the exponent-sum of each commutator transvection generator. Let C denote the kernel of this map and let B' denote the image.

name	generator
v_1	$M_{x_a, [x_b, x_c]}$
v_2	$M_{x_a^{-1}, [x_b, x_c]}$
v_3	$M_{x_a, [x_b^{-1}, x_c]}$
v_4	$M_{x_a^{-1}, [x_b^{-1}, x_c]}$
v_5	$M_{x_a, [x_b, x_c^{-1}]}$
v_6	$M_{x_a^{-1}, [x_b, x_c^{-1}]}$
v_7	$M_{x_a, [x_b^{-1}, x_c^{-1}]}$
v_8	$M_{x_a^{-1}, [x_b^{-1}, x_c^{-1}]}$

name	relation
r_1	$C_{x_a, x_b} M_{x_a, [x_b, x_c]} C_{x_a, x_b}^{-1} M_{x_a, [x_b^{-1}, x_c]}$
r_2	$C_{x_a, x_b} M_{x_a^{-1}, [x_b, x_c]} C_{x_a, x_b}^{-1} M_{x_a^{-1}, [x_b^{-1}, x_c]}$
r_3	$C_{x_a, x_c}^{-1} M_{x_a, [x_b, x_c]} C_{x_a, x_c} M_{x_a, [x_b, x_c^{-1}]}$
r_4	$C_{x_a, x_c}^{-1} M_{x_a^{-1}, [x_b, x_c]} C_{x_a, x_c} M_{x_a^{-1}, [x_b, x_c^{-1}]}$
r_5	$C_{x_a, x_c}^{-1} M_{x_a, [x_b^{-1}, x_c]} C_{x_a, x_c} M_{x_a, [x_b^{-1}, x_c^{-1}]}$
r_6	$C_{x_a, x_c}^{-1} M_{x_a^{-1}, [x_b^{-1}, x_c]} C_{x_a, x_c} M_{x_a^{-1}, [x_b^{-1}, x_c^{-1}]}$
r_7	$C_{x_a, x_b} M_{x_a, [x_b, x_c^{-1}]} C_{x_a, x_b}^{-1} M_{x_a, [x_b^{-1}, x_c^{-1}]}$
r_8	$C_{x_a, x_b} M_{x_a^{-1}, [x_b, x_c^{-1}]} C_{x_a, x_b}^{-1} M_{x_a^{-1}, [x_b^{-1}, x_c^{-1}]}$
r_9	$M_{x_a, [x_b, x_c]} M_{x_a^{-1}, [x_b, x_c]} [C_{x_a, x_b}^{-1}, C_{x_a, x_c}^{-1}]$
r_{10}	$M_{x_a, [x_b^{-1}, x_c]} M_{x_a^{-1}, [x_b^{-1}, x_c]} [C_{x_a, x_b}, C_{x_a, x_c}^{-1}]$
r_{11}	$M_{x_a, [x_b, x_c^{-1}]} M_{x_a^{-1}, [x_b, x_c^{-1}]} [C_{x_a, x_b}^{-1}, C_{x_a, x_c}]$
r_{12}	$M_{x_a, [x_b^{-1}, x_c^{-1}]} M_{x_a^{-1}, [x_b^{-1}, x_c^{-1}]} [C_{x_a, x_b}, C_{x_a, x_c}]$
r_{13}	$M_{x_a^{-1}, [x_b, x_c]} M_{x_a, [x_b, x_c]} [C_{x_a, x_b}^{-1}, C_{x_a, x_c}^{-1}]$
r_{14}	$M_{x_a^{-1}, [x_b^{-1}, x_c]} M_{x_a, [x_b^{-1}, x_c]} [C_{x_a, x_b}, C_{x_a, x_c}^{-1}]$
r_{15}	$M_{x_a^{-1}, [x_b, x_c^{-1}]} M_{x_a, [x_b, x_c^{-1}]} [C_{x_a, x_b}^{-1}, C_{x_a, x_c}]$
r_{16}	$M_{x_a^{-1}, [x_b^{-1}, x_c^{-1}]} M_{x_a, [x_b^{-1}, x_c^{-1}]} [C_{x_a, x_b}, C_{x_a, x_c}]$

Table 5: Labels for the eight commutator transvections using x_a, x_b and x_c in order, and for the sixteen (R5) and (R6) relations using these commutator transvections.

Since R' is normally generated by relations $r_1, \dots, r_{16}$, we know that $R'/[F', R']$ is generated by the images of these relations. Thus there is a surjection $A \to R'/[F', R']$ that sends each basis element to the image of the relation with the same name. The group B is a subgroup of $F'/[F', F']$. The natural map $R'/[F', R'] \to F'/[F', F']$ counts exponent-sums of generators. Since the generators $r_1, \dots, r_{16}$ all have zero exponent-sum for conjugation move generators, we do not lose any information by counting only commutator transvection generators in B. This means we have a commuting square

$$\begin{tikzcd} A \arrow[r] \arrow[d] & B \arrow[d] \\ R'/[F', R'] \arrow[r] & F'/[F', F'] \end{tikzcd}$$

The subgroup B' thus maps surjectively onto $(R'[F', F'])/[F', F']$, which is isomorphic to $R'/(R' \cap [F', F'])$. Therefore we have a commuting diagram with exact rows

$$\begin{tikzcd} 0 \arrow[r] & C \arrow[r] \arrow[d] & A \arrow[r] \arrow[d] & B' \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & (R' \cap [F', F'])/[F', R'] \arrow[r] & R'/[F', R'] \arrow[r] & R'/(R' \cap [F', F']) \arrow[r] & 0 \end{tikzcd}$$

The map $B' \to F'/[F', F']$ is injective, so $B' \to R'/(R' \cap [F', F'])$ is also injective. By construction, $A \to R'/[F', R']$ is surjective. It follows from a simple diagram chase that $C \to (R' \cap [F', F'])/[F', R']$ is surjective.

The map $A \to B$ is given by the 8×16 matrix

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

A straightforward linear algebra computation shows that C, the kernel of this map, is generated by the nine vectors

$$r_1 - r_3 - r_5 + r_7, \quad r_2 - r_4 - r_6 + r_8, \quad -r_1 - r_2 + r_{13} + r_{14},$$
$$-r_3 - r_4 + r_{13} + r_{15}, \quad r_1 + r_2 - r_5 - r_6 - r_{13} + r_{12},$$
$$r_9 - r_{13}, \quad r_{10} - r_{14}, \quad r_{11} - r_{15}, r_{12} - r_{16}.$$

Since C surjects on $(R' \cap [F', F'])/[F', R']$, we know that $(R' \cap [F', F'])/[F', R']$ is generated by the images of these nine elements.

We will now describe calculations that show that each of the generators above is equivalent modulo $[F', R']$ to an (H2) relation. In each case, we find representatives in $R' \cap [F', F']$ of the image of the given element of C. Since we are working modulo $[F', R']$ we may conjugate any r_i in computing the representative. In reducing to an (H2) relation, we may also apply any relation at all, as long as we apply its inverse somewhere else.

The last four generators are easily equivalent to (H2) relations. We skip the second kernel generator because its image is equal to the first after inverting x_a, and we skip the fourth because its image is equal to the third after swapping x_b and x_c. The three computations in the list `kernellist` finish the lemma by showing that the first, third and fifth generators are equivalent to (H2) relations. □

Proof of Theorem 6.1 We must show that every element of $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$ that happens to lie in $\mathrm{H}_2(\mathrm{IA}_n)$ can be written as a sum of elements of

$$\{\|\theta(w)(r)\| \mid w \in S_{\mathrm{Aut}}(n)^* \text{ and } r \text{ is one of the relations (H1)–(H9) from Table 4}\}.$$

Combining Lemmas 6.2 and 6.3 with the fact that (R0) and (R5) and (R6) are the only relations in our L–presentation for IA_n that do not appear as one of the commutator relations in Table 4, we see that it enough to deal with sums of elements of the set

$$\{\|r\| \mid r \text{ is one of the relations (R0), (R5) and (R6)}\}.$$

So consider $\|w\| \in \mathrm{H}_2(\mathrm{IA}_n)$ that can be written as

$$\|w\| = \sum_{i=1}^{m} \|r_i\|$$

with each r_i either an (R0), (R5) or (R6) relation.

For any choice of distinct $1 \leq a, b, c \leq n$, we consider the commutator transvection generators involving (x_a, x_b, x_c) or (x_a, x_c, x_b), and the (R5), (R6) and (R0) relations involving only these commutator transvections. We write

$$\|w\| = \sum_{i=1}^{n+\binom{n-1}{2}} \|w_i\|,$$

where each $\|w_i\|$ is a sum of (R5), (R6) and (R0) relations involving only a single choice of $(x_a, \{x_b, x_c\})$. To prove the theorem, it is enough to show that we can write each of the $\|w_i\|$ as a sum of our generators for $\mathrm{H}_2(\mathrm{IA}_n)$. So we assume $\|w\| = \|w_i\|$ for some i; this amounts to fixing a choice of $(x_a, \{x_b, x_c\})$ and assuming $\|w\|$ is a sum of (R0), (R5) and (R6) generators using only commutator transvections involving this triple.

We call a commutator transvection $M_{x_i^\alpha, [x_j^\beta, x_k^\gamma]}$ *positive* if $j < k$, and *negative* otherwise. Each (R5) or (R6) relation contains two positive commutator transvections (and no negative ones), or two negative ones (and no positive ones). Suppose $\|r_i\|$ is an (R5) or (R6) relation with negative generators, appearing in the sum defining $\|w\|$. By inserting an (R0) relation and its inverse into r_i, we replace both of the negative generators with positive ones. Let r_i' denote the word we get by doing this to r_i. Since we have added and subtracted the same element in $\widetilde{\mathrm{H}_2(\mathrm{IA}_n)}$, we have $\|r_i'\| = \|r_i\|$. Modifying an (R5) or (R6) relation in this way gives us the inverse of an (R5) or (R6) relation involving the same $(x_a, \{x_b, x_c\})$, up to cyclic permutation of the relation. So we interpret this move as rewriting the sum defining $\|w\|$: we replace the relation $\|r_i\|$ with the new relation $\|r_i'\|$, which is an (R5) or (R6) relation without negative commutator transvections. We proceed to eliminate all the negative commutator transvections in (R5) and (R6) relations in $\|w\|$ this way.

Having done this, the only negative commutator transvections the sum defining $\|w\|$ appear in (R0) relations. Since $\|w\| \in \mathrm{H}_2(\mathrm{IA}_n)$, the negative generators appear with exponent-sum zero; so the (R0) relations appear in inverse pairs. This means that we can simply rewrite the sum without any (R0) relations. So $\|w\|$ is a sum of (R5) and (R6) relations whose only commutator transvections are positive ones involving $(x_a, \{x_b, x_c\})$. Then $\|w\|$ satisfies the hypotheses of Lemma 6.4 and therefore is a sum of (H2) generators. □

7 Coinvariants and congruence subgroups

This section contains the proofs of Theorems C and D, which can be found in Section 7.2 and 7.3, respectively. Both of these proofs depend on calculations that are contained in Section 7.1.

7.1 The action of $\mathrm{GL}_n(\mathbb{Z})$ on $\mathrm{H}_2(\mathrm{IA}_n)$

This section is devoted to understanding the action of $\mathrm{GL}_n(\mathbb{Z})$ on our generators for $\mathrm{H}_2(\mathrm{IA}_n)$. The results in this section consist of long lists of equations that are verified by a computer, so on their first pass a reader might want to skip to the next two sections to see how they are used. For $i = 1, \dots, 9$, we use the notation $h_i(x_{a_1}^{\alpha_1}, \dots, x_{a_{k_i}}^{\alpha_{k_i}}) \in \mathrm{H}_2(\mathrm{IA}_n)$ for the image in $\mathrm{H}_2(\mathrm{IA}_n)$ of the i^{th} relation from $S_H(n)$, with the given parameters, as specified in Table 4. Since the action of $\mathrm{GL}_n(\mathbb{Z})$ is induced from the action of $\mathrm{Aut}(F_n)$, we record the action of various $\mathrm{Aut}(F_n)$ generators on these generators.

The computations justifying Lemmas 7.2–7.6 are in the file `h2ia.g`. We use the Hopf isomorphism $\mathrm{H}_2(\mathrm{IA}_n) \cong (R \cap [F, F])/[F, R]$, where $F = F(S_{\mathrm{IA}}(n))$ and $R < F$ is the group of relations of IA_n. We justify these equations by performing computations in $R \cap [F, F] \subset F$. In each computation, we start with a word representing one side of the equation and reduce to the trivial word using words representing the other side. Since $[F, R]$ is trivial, we may use any relations in inverse pairs, we may apply relations from in any order, and we may cyclically permute relations.

We note some identities, which we leave as an exercise.

Lemma 7.1 *The following identities hold in* $\mathrm{H}_2(\mathrm{IA}_n)$. *The letters in subscripts are assumed distinct unless otherwise noted.*

(a) $h_1(x_a^\alpha, x_b^\beta, x_c^\gamma, x_b^\delta) = -h_1(x_a^\alpha, x_b^\beta, x_c^\gamma, x_b^{-\delta})$, *even if* $b = d$.

(b) $h_3(x_a^\alpha, x_b^\beta, x_c^\gamma, x_b^\delta, x_e^\epsilon) = -h_3(x_a^\alpha, x_b^\beta, x_c^\gamma, x_b^\delta, x_e^{-\epsilon})$, *even if* $b = e$ *or* $c = e$.

(c) $h_3(x_a^\alpha, x_b^\beta, x_c^\gamma, x_b^\delta, x_e^\epsilon) = -h_3(x_a^\alpha, x_c^\gamma, x_b^\beta, x_b^\delta, x_e^\epsilon)$, *even if* $b = e$ *or* $c = e$.

We also need the following, which is not obvious.

Lemma 7.2 *The following identities hold in* $\mathrm{H}_2(\mathrm{IA}_n)$. *The letters in subscripts are assumed distinct unless otherwise noted.*

(a) $h_6(x_a^\alpha, x_e^\epsilon, x_c^\gamma, x_d^\delta, x_e^\epsilon) =$
$h_6(x_a^\alpha, x_c^\gamma, x_e^{-\epsilon}, x_d^\delta, x_e^\epsilon) - h_7(x_a^\alpha, x_e^\epsilon, x_d^\delta, x_e^\epsilon) - h_7(x_a^\alpha, x_e^{-\epsilon}, x_d^\delta, x_e^\epsilon)$.

(b) $h_6(x_a^\alpha, x_b^\beta, x_c^\gamma, x_d^\delta, x_e^\epsilon) = -h_6(x_a^\alpha, x_c^\gamma, x_b^\beta, x_d^\delta, x_e^\epsilon)$, *even if* $b = e$ *or* $c = e$.

Proof Computations justifying these equations appear in the list `lemma7pt2`. □

We proceed by expressing the action of many elementary matrices from $\mathrm{GL}_n(\mathbb{Z})$ on our generators.

Lemma 7.3 *The following identities hold in* $\mathrm{H}_2(\mathrm{IA}_n)$. *The letters in subscripts are assumed distinct unless otherwise noted.*

(a) $M^\epsilon_{x_b^\beta, x_e} \cdot h_1(x_a^\alpha, x_b^\beta, x_c^\gamma, x_d^\delta) = h_1(x_a^\alpha, x_b^\beta, x_c^\gamma, x_d^\delta) + h_1(x_a^\alpha, x_e^\epsilon, x_c^\gamma, x_d^\delta)$.

(b) $M^\epsilon_{x_b^\beta, x_e} \cdot h_1(x_a^\alpha, x_e^\epsilon, x_c^\gamma, x_d^\delta) = h_1(x_a^\alpha, x_e^\epsilon, x_c^\gamma, x_d^\delta)$.

(c) $M^{\delta}_{x_b^{\beta}, x_d} \cdot h_1(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) = h_1(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) + h_1(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}, x_d^{\delta})$.

(d) $M^{-\epsilon}_{x_a^{\alpha}, x_e} \cdot h_1(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) = h_1(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) - h_3(x_a^{\alpha}, x_e^{\epsilon}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta})$, *even if* $x_e^{\epsilon} = x_d^{\delta}$.

(e) $M^{-\epsilon}_{x_d^{\delta}, x_e} \cdot h_3(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) =$
$h_3(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) - h_2(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon}, x_f^{\zeta})$, *even if* $\{b,c\} \cap \{e,f\} \neq \varnothing$.

(f) $M^{\delta}_{x_a^{-\alpha}, x_d} \cdot h_2(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{-\delta}, x_f^{\zeta}, x_e^{\epsilon}) =$
$h_2(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{-\delta}, x_f^{\zeta}, x_e^{\epsilon}) - h_2(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_a^{-\alpha}, x_e^{\epsilon}, x_f^{\zeta})$,
even if $\{b,c\} \cap \{e,f\} \neq \varnothing$.

Proof These computations appear in `lemma7pt3`. The equations where coincidences are allowed are justified in several different computations. □

Lemma 7.4 *The following identities hold in* $\mathrm{H}_2(\mathrm{IA}_n)$. *The letters in subscripts are assumed distinct.*

(a) $M^{\beta}_{x_d^{\delta}, x_b} \cdot h_4(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}) = h_4(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}) + h_4(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma})$.

(b) $M^{\beta}_{x_d^{\delta}, x_b} \cdot h_4(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}) = h_4(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma})$.

(c) $M^{-\gamma}_{x_b^{\beta}, x_c} \cdot h_1(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) =$
$h_1(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) - h_3(x_b^{\beta}, x_d^{\delta}, x_c^{\gamma}, x_a^{\alpha}, x_c^{\gamma}) +$
$h_5(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) + h_4(x_c^{\gamma}, x_d^{\delta}, x_a^{\alpha})$.

(d) $M^{\epsilon}_{x_d^{\delta}, x_e} \cdot h_8(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) =$
$h_8(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) + h_8(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_e^{\epsilon}) + ((\mathrm{H1})\ \text{generators})$.

(e) $M^{\epsilon}_{x_d^{\delta}, x_e} \cdot h_8(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_e^{\epsilon}) = h_8(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_e^{\epsilon})$.

Proof These computations appear in `lemma7pt4`. □

Lemma 7.5 *The following identities hold in* $\mathrm{H}_2(\mathrm{IA}_n)$. *The letters in subscripts are assumed distinct.*

(a) $M^{\epsilon}_{x_f^{\zeta}, x_e} \cdot h_6(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) =$
$h_6(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) + h_6(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon})$.

(b) $M^{\epsilon}_{x_f^{\zeta}, x_e} \cdot h_6(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon}) = h_6(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon})$.

(c) $M^{-\epsilon}_{x_c^{\gamma},x_e} \cdot h_7(x_a^{\alpha}, x_b^{\beta}, x_e^{\epsilon}, x_d^{\delta}) =$
$h_7(x_a^{\alpha}, x_b^{\beta}, x_e^{\epsilon}, x_d^{\delta}) + h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) + h_6(x_e^{\epsilon}, x_b^{\beta}, x_a^{\alpha}, x_c^{\gamma}, x_d^{\delta}) +$ ((H1)–(H3) generators).

(d) $M^{-\epsilon}_{x_c^{\gamma},x_e} \cdot h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) =$
$h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) + h_2(x_c^{-\gamma}, x_e^{-\epsilon}, x_d^{-\delta}, x_c^{\gamma}, x_b^{\beta}, x_a^{\alpha})$.

(e) $M^{\beta}_{x_d^{\delta},x_b} \cdot h_7(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}, x_d^{\delta}) =$
$h_7(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}, x_d^{\delta}) + h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_b^{\beta} +$
$h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_d^{\delta}) + h_7(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}, x_b^{\beta}) +$ ((H1) generators).

(f) $M^{\beta}_{x_d^{\delta},x_b} \cdot h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_b^{\beta}) = h_7(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}, x_b^{\beta})$.

(g) $M^{\beta}_{x_d^{\delta},x_b} \cdot h_7(x_a^{\alpha}, x_d^{-\delta}, x_c^{\gamma}, x_d^{\delta}) =$
$h_7(x_a^{\alpha}, x_d^{-\delta}, x_c^{\gamma}, x_d^{\delta}) + h_7(x_a^{\alpha}, x_b^{-\beta}, x_c^{\gamma}, x_b^{\beta}) +$
$h_7(x_a^{\alpha}, x_d^{-\delta}, x_c^{\gamma}, x_b^{\beta}) - h_7(x_a^{\alpha}, x_b^{-\beta}, x_c^{\gamma}, x_d^{-\delta})$.

(h) $M^{\beta}_{x_d^{\delta},x_b} \cdot h_7(x_a^{\alpha}, x_b^{-\beta}, x_c^{\gamma}, x_b^{\beta}) = h_7(x_a^{\alpha}, x_b^{-\beta}, x_c^{\gamma}, x_b^{\beta})$.

(i) $M^{\epsilon}_{x_f^{\zeta},x_e} \cdot h_6(x_a^{\alpha}, x_f^{\zeta}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) =$
$h_6(x_a^{\alpha}, x_f^{\zeta}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) - h_7(x_a^{\alpha}, x_e^{\epsilon}, x_d^{\delta}, x_f^{\zeta}) - h_7(x_a^{\alpha}, x_e^{\epsilon}, x_d^{\delta}, x_e^{\epsilon}) +$
$h_6(x_a^{\alpha}, x_f^{\zeta}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon}) - h_7(x_a^{\alpha}, x_e^{-\epsilon}, x_d^{\delta}, x_f^{\zeta}) - h_7(x_a^{\alpha}, x_e^{-\epsilon}, x_d^{\delta}, x_e^{\epsilon}) +$
$h_6(x_a^{\alpha}, x_e^{\epsilon}, x_c^{\gamma}, x_d^{\delta}, x_f^{\zeta}) + h_6(x_a^{\alpha}, x_e^{\epsilon}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon})$.

(j) $M^{\epsilon}_{x_f^{\zeta},x_e} \cdot h_6(x_a^{\alpha}, x_e^{\epsilon}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon}) = h_6(x_a^{\alpha}, x_e^{\epsilon}, x_c^{\gamma}, x_d^{\delta}, x_e^{\epsilon})$.

Proof These computations appear in `lemma7pt5`. □

Lemma 7.6 *The following identities hold in* $\mathrm{H}_2(\mathrm{IA}_n)$. *The letters in subscripts are assumed distinct.*

(a) $M^{\delta}_{x_b^{\beta},x_d} \cdot h_9(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}) =$
$h_9(x_a^{\alpha}, x_b^{\beta}, x_c^{\gamma}) + h_9(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}) +$ ((H1)–(H5) generators).

(b) $M^{\delta}_{x_b^{\beta},x_d} \cdot h_9(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}) = h_9(x_a^{\alpha}, x_d^{\delta}, x_c^{\gamma}) +$ ((H1)–(H6) generators).

Proof These computations are contained in the list `lemma7pt6`, but they take some explanation. We use the relations `exiarel(5,[...])` frequently in these computations; these relations are always combinations of (H1) and (H4) relations. We use `exiarel(7,[xa,xb,xc])` to represent $h_9(x_a^\alpha, x_b^\beta, x_c^\gamma)$, but we also use other relations in this computation. The relation `exiarel(8,[xa,xb,xc,xd])` is an expanded version of this relation that behaves better under this action. Its derivation in `exiarelchecklist` shows that it differs from $-h_9(x_a^\alpha, x_b^\beta, x_c^\gamma)$ only by (H1), (H4) and (H5) relations. We also use `exiarel(6,[xa,xb,xc,xd])`; this differs from $h_9(x_a^\alpha, x_b^\beta, x_c^\gamma)$ by (H1) and (H4) relations, and an (R6) relation. This is apparent in the derivation of `exiarel(7,[xa,xb,xc])` in `exiarelchecklist`.

The computation justifying Lemma 7.6(a) shows directly that the image of the relation `exiarel(8,[xa,xb,xc,xd])` under `[["M",xb,xd]]` can be reduced to the trivial word by applying

- `exiarel(6,[xa,xd,xc])`,
- `iw(exiarel(7,[xa,xb,xc]))`,
- (H1)–(H5) relations (some in `exiarel(5,[...])` relations),
- (R5) and (R6) relations, and
- elements from $[F, R]$, including inverse pairs of instances of `exiarel(1,[...])` and `exiarel(3,[...])`.

Since we start and end with elements of $R \cap [F, F]$ in this computation, the use of (R5) and (R6) (in one case inside `exiarel(6,[xa,xb,xc,xd])`) is inconsequential; by Lemma 6.4 this can only change the outcome by (H2) relations. So this proves Lemma 7.6(a).

The computation justifying Lemma 7.6(b) is similar, but uses an instance of the relation `exiarel(2,[...])`. The derivation in `exiarelchecklist` shows that this relation is a combination of (H2), (H5) and (H6) relations, and elements of $[F, R]$. $\square$

7.2 Coinvariants of $\mathrm{H}_2(\mathrm{IA}_n)$

In this section, we prove Theorem C, which asserts that for the $\mathrm{GL}_n(\mathbb{Z})$–coinvariants of $\mathrm{H}_2(\mathrm{IA}_n)$ vanish for $n \geq 6$.

Proof of Theorem C We use the generators (H1)–(H9) from Table 4. To show that the coinvariants $\mathrm{H}_2(\mathrm{IA}_n)_{\mathrm{GL}_n(\mathbb{Z})}$ are trivial, we show that the coinvariance class of each of these generators is trivial. The coinvariants are defined to be the largest quotient of $\mathrm{H}_2(\mathrm{IA}_n)$ with a trivial induced action of $\mathrm{GL}(n,\mathbb{Z})$. Since the action of $\mathrm{GL}(n,\mathbb{Z})$ on $\mathrm{H}_*(\mathrm{IA}_n)$ is induced by the action of $\mathrm{Aut}(F_n)$ on IA_n, this means that $\mathrm{H}_2(\mathrm{IA}_n)_{\mathrm{GL}_n(\mathbb{Z})}$ is the quotient of $\mathrm{H}_2(\mathrm{IA}_n)$ by the subgroup generated by classes of the form $f \cdot r - r$, where $f \in \mathrm{Aut}(F_n)$ and $r \in \mathrm{H}_2(\mathrm{IA}_n)$. Elements of the form $f \cdot r - r$ are called *coboundaries*.

In fact, we have already shown in Lemmas 7.1–7.6 that the subgroup of $\mathrm{H}_2(\mathrm{IA}_n)$ generated by coboundaries contains our generators from Theorem 6.1. We show this for each generator in the equations above as follows:

- (H1) in equations (a) and (c) of Lemma 7.3, also using an observation from Lemma 7.1,
- (H2) in equations (e) and (f) of Lemma 7.3,
- (H3) in Lemma 7.3(d), also using an observation from Lemma 7.1,
- (H4) in Lemma 7.4(a),
- (H5) in Lemma 7.4(c),
- generic (H6) in Lemma 7.5(a),
- (H7) in equations (c), (e) and (g) of Lemma 7.5,
- the special cases of (H6) in Lemma 7.5(i), also using equations (a) and (b) of Lemma 7.2,
- (H8) in Lemma 7.4(d) and
- (H9) in Lemma 7.6(a).

Each equation shows how to express the given generator as a sum of coboundaries and generators previously expressed in terms of coboundaries: □

Remark 7.7 The equations above assume that distinct subscripts label distinct elements. This means that Lemma 7.5(a) requires six basis elements. We do not know if the generic (H6) generator (a five-parameter generator) can be expressed as a sum of coboundaries without using a sixth basis element. Therefore we require $n \geq 6$ in the statement and we do not know if the theorem holds for smaller n.

7.3 Second homology of congruence subgroups

In this section, we prove Theorem D, which asserts that $\mathrm{H}_2(\mathrm{Aut}(F_n,\ell);\mathbb{Q}) = 0$ for $n \geq 6$ and $\ell \geq 2$. The key to this is the following lemma. Let $\mathrm{GL}_n(\mathbb{Z},\ell)$ be the level-ℓ congruence subgroup of $\mathrm{GL}_n(\mathbb{Z})$, that is, the kernel of the natural map $\mathrm{GL}_n(\mathbb{Z}) \to \mathrm{GL}_n(\mathbb{Z}/\ell)$.

Like in Theorem C, we require $n \geq 6$ because of Lemma 7.5(a). We do not know if the result holds for smaller n.

Lemma 7.8 *For $n \geq 6$ and $\ell \geq 2$ we have $(\mathrm{H}_2(\mathrm{IA}_n;\mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)} = 0$.*

Proof Again we use our generators from Theorem 6.1. The universal coefficient theorem implies that $\mathrm{H}_2(\mathrm{IA}_n;\mathbb{Q})$ is generated by the images of our generators from $\mathrm{H}_2(\mathrm{IA}_n)$.

We have two approaches for showing that a generator has trivial image. The first is the following: if $f \in \mathrm{Aut}(F_n)$ and $r, s \in \mathrm{H}_2(\mathrm{IA}_n)$ with $f \cdot r - r = s$ and $f \cdot s = s$ in $(\mathrm{H}_2(\mathrm{IA}_n; \mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)}$, then

$$f^{\ell} \cdot r - r = \ell s \quad \text{in } (\mathrm{H}_2(\mathrm{IA}_n; \mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)}.$$

If further f^{ℓ} lies in $\mathrm{Aut}(F_n, \ell)$ then this shows that s is trivial in $(\mathrm{H}_2(\mathrm{IA}_n; \mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)}$.

The second approach is simpler: if $f \in \mathrm{Aut}(F_n)$ and $r, s \in \mathrm{H}_2(\mathrm{IA}_n)$ with $f \cdot r - r = s$ and $r = 0$ in $(\mathrm{H}_2(\mathrm{IA}_n; \mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)}$, then of course, $s = 0$ in $(\mathrm{H}_2(\mathrm{IA}_n; \mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)}$.

We show the generators have trivial images as follows:

- generic (H1) using equations (a) and (b) of Lemma 7.3 by the first approach,
- special cases of (H1) using Lemma 7.3(c) by the second approach, and using Lemma 7.1,
- (H3) by Lemma 7.3(d) using the second approach, and using Lemma 7.1,
- (H2) by equations (e) and (f) of Lemma 7.3, using the second approach,
- (H4) by equations (a) and (b) of Lemma 7.4, using the first approach,
- (H5) by Lemma 7.4(c), by the second approach,
- generic (H6) by equations (a) and (b) of Lemma 7.5, using the first approach,
- generic (H7) by equations (c) and (d) of Lemma 7.5, by the first approach,
- special cases of (H7) by equations (e) and (f), and by equations (g) and (h), of Lemma 7.5, both by the first approach,
- one special case of (H6) by equations (i) and (j) of Lemma 7.5, by the first approach,
- other special cases of (H6) using the first case and equations (a) and (b) of Lemma 7.2,
- (H8) by equations (d) and (e) of Lemma 7.4 by the first approach and
- (H9) by equations (a) and (b) of Lemma 7.6 by the first approach. □

Proof of Theorem D We examine the Hochschild–Serre spectral sequence associated to the short exact sequence

$$1 \to \mathrm{IA}_n \to \mathrm{Aut}(F_n, \ell) \to \mathrm{GL}_n(\mathbb{Z}, \ell) \to 1. \tag{12}$$

First, the Borel stability theorem [7] implies that $\mathrm{H}_2(\mathrm{GL}_n(\mathbb{Z}, \ell); \mathbb{Q}) = 0$. Next, recall from the introduction that $\mathrm{H}_1(\mathrm{IA}_n; \mathbb{Q}) \cong \mathrm{Hom}(\mathbb{Q}^n, \bigwedge^2 \mathbb{Q}^n)$; the group $\mathrm{GL}_n(\mathbb{Z})$ acts on this in the obvious way. Since $\mathrm{Hom}(\mathbb{R}^n, \bigwedge^2 \mathbb{R}^n)$ is an irreducible representation of the algebraic group $\mathrm{SL}_n(\mathbb{R})$, it follows from work of Borel [6, Proposition 3.2] that $\mathrm{H}_1(\mathrm{IA}_n;\mathbb{Q})$ is an irreducible representation of $\mathrm{GL}_n(\mathbb{Z},\ell)$ (we remark that the above reference is one of the steps in the original proof of the Borel density theorem; the result can also be derived directly from the Borel density theorem). It thus follows from the extension of the Borel stability theorem to nontrivial coefficient systems [8] that

$$\mathrm{H}_1(\mathrm{GL}_n(\mathbb{Z},\ell);\mathrm{H}_1(\mathrm{IA}_n;\mathbb{Q}))=0.$$

Lemma 7.8 says that

$$\mathrm{H}_0(\mathrm{GL}_n(\mathbb{Z},\ell);\mathrm{H}_2(\mathrm{IA}_n;\mathbb{Q}))\cong(\mathrm{H}_2(\mathrm{IA}_n;\mathbb{Q}))_{\mathrm{GL}_n(\mathbb{Z},\ell)}=0.$$

The $p+q=2$ terms of the Hochschild–Serre spectral sequence associated to (12) thus all vanish, so $\mathrm{H}_2(\mathrm{Aut}(F_n,\ell);\mathbb{Q})=0$, as desired. □

References

[1] **S Andreadakis**, *On the automorphisms of free groups and free nilpotent groups*, Proc. London Math. Soc. 15 (1965) 239–268 MR

[2] **O Baker**, *The Jacobian map on outer space*, PhD thesis, Cornell University (2011) Available at https://ecommons.cornell.edu/handle/1813/30769

[3] **L Bartholdi**, *Endomorphic presentations of branch groups*, J. Algebra 268 (2003) 419–443 MR

[4] **M Bestvina**, **K-U Bux**, **D Margalit**, *Dimension of the Torelli group for* $\mathrm{Out}(F_n)$, Invent. Math. 170 (2007) 1–32 MR

[5] **S K Boldsen**, **M Hauge Dollerup**, *Towards representation stability for the second homology of the Torelli group*, Geom. Topol. 16 (2012) 1725–1765 MR

[6] **A Borel**, *Density properties for certain subgroups of semi-simple groups without compact components*, Ann. of Math. 72 (1960) 179–188 MR

[7] **A Borel**, *Stable real cohomology of arithmetic groups*, Ann. Sci. École Norm. Sup. 7 (1974) 235–272 MR

[8] **A Borel**, *Stable real cohomology of arithmetic groups, II*, from "Manifolds and Lie groups" (S Murakami, J Hano, K Okamoto, A Morimoto, H Ozeki, editors), Progr. Math. 14, Birkhäuser, Boston (1981) 21–55 MR

[9] **T Brendle**, **D Margalit**, **A Putman**, *Generators for the hyperelliptic Torelli group and the kernel of the Burau representation at* $t=-1$, Invent. Math. 200 (2015) 263–310 MR

[10] **K S Brown**, *Cohomology of groups*, Graduate Texts in Mathematics 87, Springer (1994) MR

[11] **T Church**, **B Farb**, *Parameterized Abel–Jacobi maps and abelian cycles in the Torelli group*, J. Topol. 5 (2012) 15–38 MR

[12] **T Church**, **B Farb**, *Representation theory and homological stability*, Adv. Math. 245 (2013) 250–314 MR

[13] **T Church**, **A Putman**, *The codimension-one cohomology of* $\mathrm{SL}_n\mathbb{Z}$, Geom. Topol. 21 (2017) 999–1032 MR

[14] **F Cohen**, **J Pakianathan**, *On automorphism groups of free groups and their nilpotent quotients*, in preparation

[15] **M Day**, **A Putman**, *A Birman exact sequence for* $\mathrm{Aut}(F_n)$, Adv. Math. 231 (2012) 243–275 MR

[16] **M Day**, **A Putman**, *The complex of partial bases for* F_n *and finite generation of the Torelli subgroup of* $\mathrm{Aut}(F_n)$, Geom. Dedicata 164 (2013) 139–153 MR

[17] **M Day**, **A Putman**, *A Birman exact sequence for the Torelli subgroup of* $\mathrm{Aut}(F_n)$, Internat. J. Algebra Comput. 26 (2016) 585–617 MR

[18] **B Farb**, *Automorphisms of* F_n *which act trivially on homology*, in preparation

[19] **S Galatius**, *Stable homology of automorphism groups of free groups*, Ann. of Math. 173 (2011) 705–768 MR

[20] **A Hatcher**, **K Vogtmann**, *Cerf theory for graphs*, J. London Math. Soc. 58 (1998) 633–655 MR

[21] **A Hatcher**, **N Wahl**, *Stabilization for mapping class groups of 3–manifolds*, Duke Math. J. 155 (2010) 205–269 MR

[22] **D Johnson**, *An abelian quotient of the mapping class group* $\mathcal{I}_g$, Math. Ann. 249 (1980) 225–242 MR

[23] **N Kawazumi**, *Cohomological aspects of Magnus expansions*, preprint (2005) arXiv

[24] **S Krstić**, **J McCool**, *The non-finite presentability of* $\mathrm{IA}(F_3)$ *and* $\mathrm{GL}_2(\mathbb{Z}[t,t^{-1}])$, Invent. Math. 129 (1997) 595–606 MR

[25] **R C Lyndon**, **P E Schupp**, *Combinatorial group theory*, Ergeb. Math. Grenzgeb. 89, Springer (1977) MR

[26] **W Magnus**, *Über* n*–dimensionale Gittertransformationen*, Acta Math. 64 (1935) 353–367 MR

[27] **J Nielsen**, *Die Isomorphismen der allgemeinen, unendlichen Gruppe mit zwei Erzeugenden*, Math. Ann. 78 (1917) 385–397 MR

[28] **A Putman**, *An infinite presentation of the Torelli group*, Geom. Funct. Anal. 19 (2009) 591–643 MR

[29] **A Putman**, *Obtaining presentations from group actions without making choices*, Algebr. Geom. Topol. 11 (2011) 1737–1766 MR

[30] **A Putman**, *The second rational homology group of the moduli space of curves with level structures*, Adv. Math. 229 (2012) 1205–1234 MR

[31] **T Satoh**, *The abelianization of the congruence* IA*–automorphism group of a free group*, Math. Proc. Cambridge Philos. Soc. 142 (2007) 239–248 MR

[32] **T Satoh**, *A survey of the Johnson homomorphisms of the automorphism groups of free groups and related topics*, preprint (2012) arXiv

Department of Mathematical Sciences, University of Arkansas
Fayetteville, AR, United States

Department of Mathematics, University of Notre Dame
Notre Dame, IN, United States

matthewd@uark.edu, andyp@nd.edu

Proposed: Shigeyuki Morita Received: 22 October 2015
Seconded: Benson Farb, Mark Behrens Revised: 8 November 2016

Geometry & Topology 21 (2017) 2897–2944

Tautological integrals on curvilinear Hilbert schemes

GERGELY BÉRCZI

We take a new look at the curvilinear Hilbert scheme of points on a smooth projective variety X as a projective completion of the nonreductive quotient of holomorphic map germs from the complex line into X by polynomial reparametrisations. Using an algebraic model of this quotient coming from global singularity theory we develop an iterated residue formula for tautological integrals over curvilinear Hilbert schemes.

14C05, 14N10, 55N91

1 Introduction

Let X be a smooth projective variety of dimension n and let F be a rank-r algebraic vector bundle on X. Let $X^{[k]}$ denote the Hilbert scheme of length-k subschemes of X and let $F^{[k]}$ be the corresponding tautological rank-rk bundle on $X^{[k]}$ whose fibre at $\xi \in X^{[k]}$ is $H^0(\xi, F|_\xi)$.

Let $\mathrm{Hilb}_0^k(\mathbb{C}^n)$ be the punctual Hilbert scheme defined as the closed subset of $(\mathbb{C}^n)^{[k]} = \mathrm{Hilb}^k(\mathbb{C}^n)$ parametrising subschemes supported at the origin. Following Rennemo [33] we define punctual geometric subsets as constructible subsets $Q \subseteq \mathrm{Hilb}_0^k(\mathbb{C}^n)$ which are unions of isomorphism classes of schemes, that is, if $\xi \in Q$ and $\xi' \in \mathrm{Hilb}_0^k(\mathbb{C}^n)$ are isomorphic (ie they have isomorphic coordinate rings), then $\xi' \in Q$. Geometric subsets of $X^{[k]}$ are those generated by finite unions, intersections and complements from sets of the form

$$P(Q_1, \dots, Q_s) = \{\xi \in X^{[k]} : \xi = \xi_1 \sqcup \cdots \sqcup \xi_s \text{ for some } \xi_i \in Q_i\}.$$

For a geometric subset $\mathcal{Z}$ let $\overline{\mathcal{Z}}$ denote its Zariski closure in $X^{[k]}$. Let $M(c_1, \dots, c_{rk})$ be a monomial in the Chern classes $c_i = c_i(F^{[k]})$ of weighted degree equal to $\dim \overline{\mathcal{Z}}$, where the weight of c_i is i. Let $\Omega^*(\overline{\mathcal{Z}})$ denote the algebra of differential forms supported on the smooth part of $\overline{\mathcal{Z}}$ (see eg Remark 2.3). If $\alpha_M \in \Omega^*(\overline{\mathcal{Z}})$ is a closed differential form representing the cohomology class of $M(c_1, \dots, c_{rk})$ then the Chern numbers

$$[\overline{\mathcal{Z}}] \cap M(c_1, \dots, c_{rk}) = \int_{\overline{\mathcal{Z}}} \alpha_M$$

are called tautological integrals of $F^{[k]}$. Rennemo [33] shows that these integrals can be expressed in terms of the Chern numbers of X and F.

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2897

Theorem 1.1 (Rennemo [33]) *Let $\mathcal{M}_{r,n}$ denote the set of weighted-degree-n monomials in the Chern classes $c_1(F),\dots,c_r(F)$ and $c_1(X),\dots,c_n(X)$. For $S\in\mathcal{M}_{r,n}$ let $\alpha_S\in\Omega^{\mathrm{top}}(X)$ be a closed differential form representing the cohomology class of S and let $y_S=\int_X\alpha_S$ denote the corresponding intersection number. Let $\mathcal{Z}\subset X^{[k]}$ be a geometric subset. Then for any Chern monomial $M=M(c_1,\dots,c_{rk})$ of weighted degree $\dim\overline{\mathcal{Z}}$ there is a polynomial R_M in $|\mathcal{M}_{r,n}|$ variables depending only on M such that*

$$[\overline{\mathcal{Z}}]\cap M(c_1,\dots,c_{rk})=R_M(y_S:S\in\mathcal{M}_{r,n}).$$

The proof in [33] is nonconstructive and based on constructing homology classes supported on certain diagonals of X^n (see also Li [27]) and the fact that an element in the cohomology ring of a Grassmannian is a polynomial in the Chern classes of the universal bundle. Lacking a method of obtaining information about this polynomial, there is no apparent way of turning this proof into an algorithm. Explicit expressions for tautological integrals are not known in general. On surfaces the method of Ellingsrud, Göttsche and Lehn [15] yields a recursion which in principle computes the universal polynomial explicitly. The top Segre classes of tautological bundles over surfaces provides an example of this problem and the conjecture of Lehn [26] has been recently proved by Marian, Oprea and Pandharipande [29] for K3 surfaces using virtual localisation. However, [15] and [29] deal only with surfaces and their authors integrate over the whole Hilbert scheme rather than over geometric subsets. Our method works in any dimension for integration over a geometric subset called the curvilinear component.

Let X be a smooth projective variety of dimension n. This paper provides a closed iterated residue formula for tautological integrals over the simplest geometric subsets $P(Q)$ where $s=1$ and the punctual geometric subset Q is defined as

$$Q=\{\xi\in\mathrm{Hilb}_0^k(\mathbb{C}^n):\mathcal{O}_\xi\simeq\mathbb{C}[z]/z^k\}.$$

We will see that $\overline{Q}$ is an irreducible component of the punctual Hilbert scheme. Points of $P(\overline{Q})$ correspond to curvilinear subschemes on X, ie subschemes contained in the germ of some smooth curve on X. In other words, these are the limit points on $X^{[k]}$ where k distinct points come together along a smooth curve. We denote this curvilinear locus by CX^k and its closure by $\overline{CX}^{[k]}$, which we call the curvilinear Hilbert scheme.

The main result of the present paper is the following explicit formula for tautological integrals over curvilinear Hilbert schemes.

Theorem 1.2 *Let $k\geq 1$ and $P(\mathbf{x})=P(x_1,\dots,x_{r(k+1)})$ be a polynomial of weighted degree $\dim\overline{CX}^{[k+1]}=n+(n-1)k$ in the variables x_l of weight l for $1\leq l\leq r(k+1)$.*

Let $c_l = c_l(F^{[k+1]})$ denote the l^{th} Chern class of the tautological rank-$r(k+1)$ bundle on $X^{[k+1]}$. Then

$$\int_{\overline{CX}^{[k+1]}} P(\boldsymbol{c}) = \int_X \operatorname*{Res}_{\boldsymbol{z}=\infty} \frac{(-1)^k \prod_{1\le i<j\le k}(z_i-z_j)Q_k(\boldsymbol{z})P(\boldsymbol{c}(\theta-\boldsymbol{z},\theta))d\boldsymbol{z}}{\prod_{i+j\le l\le k}(z_i+z_j-z_l)(z_1\cdots z_k)^n}\prod_{i=1}^k s_X\Big(\frac{1}{z_i}\Big),$$

where $\theta_1,\dots,\theta_r$ are the Chern roots of F and $c_l(\theta-\boldsymbol{z},\theta)$ denotes the l^{th} symmetric polynomial in the formal Chern roots $\{\theta_j-z_i,\theta_j : 1\le i\le k, 1\le j\le r\}$. The iterated residue is $(-1)^k$ times the coefficient of $(z_1\cdots z_k)^{-1}$ in the expansion of the rational expression in the domain $z_1 \ll \cdots \ll z_k$ and

$$s_X\Big(\frac{1}{z_i}\Big) = 1 + \frac{s_1(X)}{z_i} + \frac{s_2(X)}{z_i^2} + \cdots + \frac{s_n(X)}{z_i^n}$$

is the total Segre class of X at $1/z_i$. Finally $Q_k(\boldsymbol{z})$ is a homogeneous polynomial invariant of Morin singularities given as the equivariant Poincaré dual of a Borel orbit defined in the following Remark.

Remark (explanation and features of the residue formula) • The iterated residue gives a degree-n symmetric polynomial in Chern roots of F and Segre classes of X reproving Theorem 1.1 This shows that the dependence on Chern classes of X in fact can be expressed via the Segre classes of X. In particular, in Example 7.2 we give a formula for the top Segre classes of tautological bundles over curvilinear Hilbert schemes.

- For fixed k the rational expression

$$\mathcal{R}_k = \frac{\prod_{1\le i<j\le k}(z_i-z_j)Q_k(\boldsymbol{z})P(c_l(\theta-\boldsymbol{z},\theta))d\boldsymbol{z}}{\prod_{i+j\le l\le k}(z_i+z_j-z_l)}$$

in the formula is independent of the dimension n and the iterated residue depends on n only through the total Segre class s_X of X. The iterated residue is then some linear combination of the coefficients of (the expansion of) $\mathcal{R}_k$ multiplied by Segre classes of X. By increasing the dimension, the iterated residue involves new terms of the expansion of $\mathcal{R}_k$, and we can think of $\mathcal{R}_k$ as a universal rational expression encoding the integrals for fixed k but varying n.

- The Chern class $c_l(\theta-\boldsymbol{z},\theta)$ is the coefficient of t^l in

$$\prod_{j=1}^r(1+\theta_j t)\prod_{i=1}^k\prod_{j=1}^r(1-z_i t+\theta_j t),$$

that is, the l^{th} Chern class of the bundle with formal Chern roots θ_j, $\theta_j - z_i$.

- The quick description of Q_k is the following; see Remark 6.2 for details. The $\mathrm{GL}(k)$–module of 3–tensors $\mathrm{Hom}(\mathbb{C}^k, \mathrm{Sym}^2\mathbb{C}^k)$ has a diagonal decomposition
$$\mathrm{Hom}(\mathbb{C}^k, \mathrm{Sym}^2\mathbb{C}^k) = \bigoplus_{1\le m,r,l\le k} \mathbb{C}q_l^{mr},$$
where the T_k–weight of q_l^{mr} is $(z_m + z_r - z_l)$. Define $\epsilon = \sum_{m=1}^{k}\sum_{r=1}^{k-m} q_{m+r}^{mr}$ as a point in the B_k–invariant subspace
$$W_k = \bigoplus_{1\le m+r\le l\le k} \mathbb{C}q_l^{mr} \subset \mathrm{Hom}(\mathbb{C}^k, \mathrm{Sym}^2\mathbb{C}^k).$$
Then $Q_k(\mathbf{z}) = \mathrm{eP}[\overline{B_k\epsilon}, W_k]$ is the equivariant Poincaré dual of the Borel orbit $\overline{B_k\epsilon}$ in W_k. The list of these polynomials begins as follows: $Q_1 = Q_2 = Q_3 = 1$, $Q_4 = 2z_1 + z_2 - z_4$. In principle, Q_k may be calculated for each concrete k using a computer algebra program, but at the moment, we do not have an efficient algorithm for performing such calculations for large k and Q_k is only known for $k \le 6$.

The main motivation for studying tautological integrals is their immediate applications in enumerative geometry and in particular in counting hypersurfaces in sufficiently ample linear systems on X with a prescribed set of singularities. Let X be a smooth, projective, connected variety, L a sufficiently ample line bundle on X and let $T_1, \dots, T_s$ be analytic singularity types. There are expected codimensions d_i associated with each T_i, and we let $d = \sum d_i$. Rennemo [33] shows (see also Göttsche [20] and Kleiman and Piene [24]) that there is an m and a geometric set $W = W(T_1, \dots, T_s) \subset X^{[m]}$ such that a generic hypersurface containing a $Z \in W$ has the specified singularities. Therefore in a general $\mathbb{P}^d \subseteq |L|$ the number of hypersurfaces containing a subscheme $Z \in W$ is equal to $\int_W c_{\dim(W)}(L^{[m]})$; hence this tautological integral gives the number of hypersurfaces in $\mathbb{P}^d$ with singularities $T_1, \dots, T_s$.

In the forthcoming paper Bérczi and Szenes [6], we extend the methods of the present paper to study tautological integrals over more general geometric subsets supported at more than one point on X and develop residue formulae for counts of hypersurfaces with given sets of singularities. As a special case we present a new formula for the number of δ–nodal curves on surfaces (and more generally δ–nodal hypersurfaces on projective varieties) different from the well-known Göttsche conjecture [20], which by now has several proofs; see Kazarian [23], Kool, Shende and Thomas [25], Liu [28] and Tzeng [36].

The intersection theory of the Hilbert scheme of points on surfaces has been extensively studied and it can be approached from different directions. One is the inductive recursions set up by Ellingsrud, Göttsche and Lehn [15], an other possibility is using

Nakajima calculus (see Lehn [26] and Nakajima [32]). By these methods, the integration of tautological classes is reduced to a combinatorial problem. Another strategy is to prove an equivariant version of Lehn's conjecture for the Hilbert scheme of points of $\mathbb{C}^2$ via appropriately weighted sums over partitions. More recently Marian, Oprea and Pandharipande [29] proved a conjecture of Lehn [26] on integrals of top Segre classes of tautological bundles over the Hilbert schemes of points over surfaces in the K3 case via virtual localisation on the Quot schemes of the surface.

In this paper we suggest a new approach by taking a look at Hilbert schemes of points from a different perspective. We work in arbitrary dimension, not just over surfaces. For $n \geq 3$ not much is known about the irreducible components and singularities of the punctual Hilbert scheme $\mathrm{Hilb}_0^k(\mathbb{C}^n)$ so we only focus on the curvilinear component. The crucial observation is that for $k \geq 1$ the punctual curvilinear locus $CX_p^{[k+1]}$ at $p \in X$ can be described as the nonreductive quotient of k–jets of holomorphic map germs $(\mathbb{C}, 0) \to (X, p)$ by polynomial reparametrisations of $\mathbb{C}$ at the origin.

Let $J_k^{\mathrm{reg}} X$ denote the regular k–jet bundle over X whose elements are equivalence classes of germs of holomorphic maps $f\colon (\mathbb{C}, 0) \to (X, p)$ with the equivalence relation $f \sim g$ if and only if the derivatives satisfy $f^{(j)}(0) = g^{(j)}(0)$ for $0 \leq j \leq k$ when computed in some local coordinate system of X near $p \in X$ and $f'(0) \neq 0$. The reparametrisation group $\mathrm{Diff}_k(1)$ formed by k–jets of regular reparametrisations of $\mathbb{C}$ at the origin acts fibrewise on $J_k^{\mathrm{reg}} X$ and the curvilinear locus (as a set) can be identified with the quasiprojective quotient

$$CX^{[k+1]} \simeq J_k^{\mathrm{reg}} X / \mathrm{Diff}_k(1).$$

Using an algebraic model coming from global singularity theory (we call this the test-curve model) we reinterpret the natural embedding of the punctual curvilinear locus $CX_p^{[k+1]}$ into the Grassmannian of codimension-k subspaces in the maximal ideal $\mathfrak{m} = (x_1, \ldots, x_n)$ as a parametrised map $CX_p^{[k+1]} \hookrightarrow \mathrm{Grass}_k(\mathcal{D}_{X,p}^k)$, where $\mathcal{D}_X^k = \mathcal{D}_X^{\leq k} / \mathcal{O}_X$ is the bundle of order-k differential operators over X. The punctual curvilinear Hilbert scheme $\overline{CX}_p^{[k+1]}$ is the closure of the image of this map in $\mathrm{Grass}_k(\mathcal{D}_{X,p}^k)$, and moving the point p on X, this gives an embedding of the curvilinear component

$$\phi^{\mathrm{Grass}}\colon \overline{CX}^{[k+1]} \hookrightarrow \mathrm{Grass}_k(\mathcal{D}_X^k).$$

Integration on $\overline{CX}^{[k+1]}$ can be reduced to integration along the fibre $\overline{CX}_p^{[k+1]}$; see Section 7. We use equivariant localisation on $\overline{CX}_p^{[k+1]}$ following the strategy of Bérczi and Szenes [7]. However, for tautological integrals we need to modify the proof in [7] in two crucial points:

- First, the main obstacle to applying localisation directly is that we don't know which fixed points of the ambient Grassmannian sit in the image $\overline{CX}_p^{[k+1]}$. However, for $k+1 \le n$ we prove in [7] a residue vanishing theorem which tells that after transforming the localisation formula into an iterated residue only one distinguished fixed point of the torus action contributes to the sum. This mysterious property remains valid for tautological integrals but its proof needs a more detailed study of the rational differential form.

- Second, we need to extend the formula to the domain where $k+1 > n$, that is, the number of points is not smaller than the dimension of X. The trick here is to increase the dimension of the variety and study $\mathrm{Hilb}_0^{k+1}(\mathbb{C}^n)$ as a subvariety of $\mathrm{Hilb}_0^{k+1}(\mathbb{C}^{k+1})$.

The developed method reflects a surprising feature of curvilinear Hilbert schemes: in order to evaluate tautological integrals and make the residue vanishing principle work we need to increase the dimension of the variety first and work in the range where the number of points does not exceed the dimension.

Acknowledgements I warmly thank Frances Kirwan and Jørgen Vold Rennemo for the valuable discussions. I also thank the referee for the insightful comments and suggestions. This paper has outgrown from [7] and my special thanks go to András Szenes.

2 Tautological integrals

Let X be a smooth projective variety of dimension n and let F be a rank-r bundle (locally free sheaf) on X. Let

$$X^{[k]} = \{\xi \subset X : \dim(\xi) = 0 \text{ and } \mathrm{length}(\xi) = \dim H^0(\xi, \mathcal{O}_\xi) = k\}$$

denote the Hilbert scheme of k points on X parametrising length-k subschemes of X and $F^{[k]}$ the corresponding rank-rk bundle on $X^{[k]}$ whose fibre over $\xi \in X^{[k]}$ is $F \otimes \mathcal{O}_\xi = H^0(\xi, F|_\xi)$.

Equivalently, $F^{[k]} = q_* p^*(F)$, where p and q denote the projections from the universal family of subschemes $\mathcal{U}$ to X and $X^{[k]}$ respectively:

$$\begin{array}{ccc} X^{[k]} \times X \supset \mathcal{U} & \xrightarrow{q} & X^{[k]} \\ \big\downarrow p & & \\ X & & \end{array}$$

For simplicity let $\mathrm{Hilb}_0^k(\mathbb{C}^n)$ denote the punctual Hilbert scheme of k points on $\mathbb{C}^n$ defined as the closed subset of $\mathrm{Hilb}^k(\mathbb{C}^n)$ parametrising subschemes supported at the origin. Following Rennemo [33] we define punctual geometric subsets to be the constructible subsets of the punctual Hilbert scheme containing all 0–dimensional schemes of given isomorphism types.

Definition 2.1 A punctual geometric set is a constructible subset $Q \subseteq \mathrm{Hilb}_0^k(\mathbb{C}^n)$ which is the union of isomorphism classes of subschemes, that is, if $\xi \in Q$ and $\xi' \in \mathrm{Hilb}_0^k(\mathbb{C}^n)$ are isomorphic schemes then $\xi' \in Q$.

Definition 2.2 For an s–tuple $\boldsymbol{Q} = (Q_1, \dots, Q_s)$ of punctual geometric sets such that $Q_i \subseteq \mathrm{Hilb}_0^{k_i}(\mathbb{C}^n)$ and $k = \sum k_i$ define

$$P(\boldsymbol{Q}) = \{\xi \in X^{[k]} : \xi = \xi_1 \sqcup \dots \sqcup \xi_s \text{ where } \xi_i \in X_{p_i}^{[k_i]} \cap Q_i \text{ for distinct } p_1, \dots, p_s\} \subseteq X^{[k]}.$$

A subset $\mathcal{Z} \subseteq X^{[k]}$ is geometric if it can be expressed as finite union, intersection and complement of sets of the form $P(\boldsymbol{Q})$.

A straightforward way to produce punctual geometric subsets is by taking a complex algebra A of complex dimension k and making the corresponding definition

$$Q_A = \{\xi \in X^{[k]} : \mathcal{O}_\xi \simeq A\}.$$

When $A = \mathbb{C}[z]/z^k$ then $Q_A = CX_p^{[k]}$ is the punctual curvilinear locus defined in the next section and

$$\overline{CX}^{[k]} = \bigcup_{p \in X} \overline{CX}_p^{[k]}$$

is the curvilinear Hilbert scheme, the central object of this paper.

In this paper we work with singular homology and cohomology with rational coefficients. For a smooth manifold X the degree of a class $\eta \in H_*(X)$ means its push-forward to $H_*(\mathrm{pt}) = \mathbb{Q}$. By choosing $\alpha_\eta \in \Omega^{\mathrm{top}}(X)$, a closed compactly supported differential form representing the cohomology class η, this degree is equal to the integral

$$\eta \cap [X] = \int_X \alpha_\eta.$$

Let $\mathcal{Z} \subset X^{[k]}$ be a geometric subset with closure $\overline{\mathcal{Z}}$ and $M(c_1, \dots, c_{rk})$ be a monomial in the Chern classes $c_i = c_i(F^{[k]})$ of weighted degree equal to $\dim \overline{\mathcal{Z}}$, where the weight of c_i is i. By choosing $\alpha_M \in \Omega^*(X^{[k]})$, a closed compactly supported differential form representing the cohomology class of $M(c_1, \dots, c_{rk})$, the degree

$$[\overline{\mathcal{Z}}] \cap M(c_1, \dots, c_{rk}) = \int_{\overline{\mathcal{Z}}} \alpha_M \tag{1}$$

is called a tautological integral of $F^{[k]}$.

Remark 2.3 (1) In (1) the integral of α_M on the smooth part of $\overline{\mathcal{Z}}$ is absolutely convergent and by definition we denote this by $\int_{\overline{\mathcal{Z}}} \alpha_M$.

(2) Recall (see eg Bott and Tu [10]) that if $f\colon X \to Y$ is a smooth proper map between connected oriented manifolds such that f restricted to some open subset of X is a diffeomorphism, then for a compactly supported form μ on Y, we have $\int_X f^*\mu = \int_Y \mu$. The analogous statement for singular varieties is the following. Let $f\colon M \to N$ be a smooth proper map between smooth quasiprojective varieties and assume that $X \subset M$ and $Y \subset N$ are possibly singular closed subvarieties, such that f restricted to X is a birational map from X to Y. If μ is a closed differential form on N then the integral of μ on the smooth part of Y is absolutely convergent; we denote this by $\int_Y \mu$. With this convention we again have $\int_X f^*\mu = \int_Y \mu$.

In particular this means that the integral $\int_Y \mu$ of the compactly supported form μ on N is the same as the integral $\int_{\widetilde{Y}} f^*\mu$ of the pull-back form $f^*\mu$ over any (partial) resolution $f\colon (\widetilde{Y}, \widetilde{M}) \to (Y, M)$.

3 Curvilinear Hilbert schemes

In this section we describe a geometric model for curvilinear Hilbert schemes. Let X be a smooth projective variety of dimension n and let

$$X^{[k]} = \{\xi \subset X : \dim(\xi) = 0 \text{ and } \operatorname{length}(\xi) = \dim H^0(\xi, \mathcal{O}_\xi) = k\}$$

denote the Hilbert scheme of k points on X parametrising all length-k subschemes of X. For $p \in X$ let

$$X_p^{[k]} = \{\xi \in X^{[k]} : \operatorname{supp}(\xi) = p\}$$

denote the punctual Hilbert scheme consisting of subschemes supported at p. If $\rho\colon X^{[k]} \to S^k X$ given by $\xi \mapsto \sum_{p \in X} \operatorname{length}(\mathcal{O}_{\xi,p}) p$ denotes the Hilbert–Chow morphism then $X_p^{[k]} = \rho^{-1}(kp)$.

Definition 3.1 A subscheme $\xi \in X_p^{[k]}$ is called curvilinear if ξ is contained in some smooth curve $C \subset X$. Equivalently, ξ is curvilinear if $\mathcal{O}_\xi$ is isomorphic to the $\mathbb{C}$–algebra $\mathbb{C}[z]/z^k$. The punctual curvilinear locus at $p \in X$ is the set of curvilinear subschemes supported at p:

$$\begin{aligned} CX_p^{[k]} &= \{\xi \in X_p^{[k]} : \xi \subset \mathcal{C}_p \text{ for some smooth curve } \mathcal{C} \subset X\} \\ &= \{\xi \in X_p^{[k]} : \mathcal{O}_\xi \simeq \mathbb{C}[z]/z^k\}. \end{aligned}$$

If X is a surface (ie $\dim X = 2$), $CX_p^{[k]}$ is an irreducible quasiprojective variety of dimension $k-1$ which is an open dense subset in $X_p^{[k]}$ and therefore its closure is the full punctual Hilbert scheme at p, that is, $\overline{CX}_p^{[k]} = X_p^{[k]}$. When $n \geq 3$ the punctual Hilbert scheme $X_p^{[k]}$ is not necessarily irreducible or reduced, but the closure of the curvilinear locus is one of its irreducible components:

Lemma 3.2 *$\overline{CX}_p^{[k]}$ is an irreducible component of the punctual Hilbert scheme $X_p^{[k]}$ of dimension $(n-1)(k-1)$.*

Proof Note that $\xi \in \mathrm{Hilb}_0^{[k]}(\mathbb{C}^n)$ is not curvilinear if and only if $\mathcal{O}_\xi$ does not contain elements of degree $k-1$, that is, after fixing some local coordinates $x_1, \dots, x_n$ of $\mathbb{C}^n$ at the origin we have

$$\mathcal{O}_\xi \simeq \mathbb{C}[x_1, \dots, x_n]/I \quad \text{for some} \quad I \supseteq (x_1, \dots, x_n)^{k-1}.$$

This is a closed condition and hence curvilinear subschemes can't be approximated by noncurvilinear subschemes in $\mathrm{Hilb}_0^{[k]}(\mathbb{C}^n)$. The dimension of $\overline{CX}_p^{[k]}$ will come from the description of it as a nonreductive quotient in the next subsection. □

Note that any curvilinear subscheme contains only one subscheme for any given smaller length and any small deformation of a curvilinear subscheme is again locally curvilinear.

Remark 3.3 Fix coordinates $x_1, \dots, x_n$ on $\mathbb{C}^n$. Recall that the defining ideal I_ξ of any subscheme $\xi \in \mathrm{Hilb}_0^{k+1}(\mathbb{C}^n)$ is a codimension-k subspace in the maximal ideal $\mathfrak{m} = (x_1, \dots, x_n)$. The dual of this is a k–dimensional subspace S_ξ in $\mathfrak{m}^* \simeq \mathrm{Sym}^{\le k}\mathbb{C}^n$ giving us a natural embedding $\varphi\colon X_p^{[k+1]} \hookrightarrow \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$. In what follows, we give an explicit parametrisation of this embedding using an algebraic model coming from global singularity theory.

3.1 Test-curve model for $\overline{CX}^{[k]}$

3.1.1 Jets of holomorphic maps If u and v are positive integers let $J_k(u,v)$ denote the vector space of k–jets of holomorphic maps $(\mathbb{C}^u, 0) \to (\mathbb{C}^v, 0)$ at the origin, that is, the set of equivalence classes of maps $f\colon (\mathbb{C}^u, 0) \to (\mathbb{C}^v, 0)$, where $f \sim g$ if and only if $f^{(j)}(0) = g^{(j)}(0)$ for all $j = 1, \dots, k$. This is a finite-dimensional complex vector space, which one can identify with $J_k(u,1) \otimes \mathbb{C}^v$; hence $\dim J_k(u,v) = v\binom{u+k}{k} - v$. We will call the elements of $J_k(u,v)$ *map-jets of order* k, or simply map-jets.

Eliminating the terms of degree $k+1$ results in an algebra homomorphism $J_k(u,1) \twoheadrightarrow J_{k-1}(u,1)$, and the chain $J_k(u,1) \twoheadrightarrow J_{k-1}(u,1) \twoheadrightarrow \cdots \twoheadrightarrow J_1(u,1)$ induces the following increasing filtration on $J_k(u,1)^*$:

(2) $$J_1(u,1)^* \subset J_2(u,1)^* \subset \cdots \subset J_k(u,1)^*.$$

Remark 3.4 The space $J_i(u,1)^*$ may be interpreted as set of differential operators on $\mathbb{C}^u$ of degree at most i, and in particular, by taking symbols, we have

(3) $$J_k(u,1)^* \simeq \mathrm{Sym}^{\le k}\mathbb{C}^u \stackrel{\mathrm{def}}{=} \bigoplus_{l=1}^{k} \mathrm{Sym}^l \mathbb{C}^u,$$

where Sym^l stands for the symmetric tensor product and the isomorphism is that of filtered $\mathrm{GL}(n)$–modules. Given a regular k–jet $f\colon (\mathbb{C},0)\to(\mathbb{C}^n,0)$ in $J_k^{\mathrm{reg}}(1,n)$ we may push forward the differential operators of order k on $\mathbb{C}$ (with constant coefficients) to $\mathbb{C}^n$ along f which gives us a map

$$\tilde{f}\colon J_k(1,1)^*\to \mathrm{Grass}(k,J_k(n,1)^*).$$

In Section 3.1.3 we describe a parametrisation of this map, and identify the image in the Grassmannian with the punctual curvilinear locus $CX_p^{[k+1]}$ using local coordinates on X near p.

Choosing coordinates on $\mathbb{C}^u$ and $\mathbb{C}^v$ a k–jet $f\in J_k(u,v)$ can be identified with the set of derivatives at the origin, that is, the vector $(f'(0),f''(0),\dots,f^{(k)}(0))$, where $f^{(j)}(0)\in\mathrm{Hom}(\mathrm{Sym}^j\mathbb{C}^u,\mathbb{C}^v)$. This way we get the identification

$$J_k(u,v)\simeq J_k(u,1)\otimes\mathbb{C}^v\simeq\bigoplus_{j=1}^{k}\mathrm{Hom}(\mathrm{Sym}^j\mathbb{C}^u,\mathbb{C}^v). \tag{4}$$

One can compose map-jets via substitution and elimination of terms of degree greater than k; this leads to the composition map

$$\begin{aligned} &J_k(u,v)\times J_k(v,w)\to J_k(u,w),\\ &\qquad\qquad(\Psi_1,\Psi_2)\mapsto\Psi_2\circ\Psi_1 \text{ modulo terms of degree} > k. \end{aligned} \tag{5}$$

When $k=1$, the map-jets in $J_1(u,v)$ may be identified with u-by-v matrices, and (5) reduces to multiplication of matrices.

The k–jet of a curve $(\mathbb{C},0)\to(\mathbb{C}^n,0)$ is simply an element of $J_k(1,n)$. We call such a curve γ *regular* if $\gamma'(0)\neq 0$; introduce the notation $J_k^{\mathrm{reg}}(1,n)$ for the set of regular curves:

$$J_k^{\mathrm{reg}}(1,n)=\{\gamma\in J_k(1,n):\gamma'(0)\neq 0\}.$$

Note that $J_k^{\mathrm{reg}}(u,u)$ with the composition map (5) has a natural group structure and we will often use the notation

$$\mathrm{Diff}_k(u)=J_k^{\mathrm{reg}}(u,u)$$

and refer to this set as the *k–jet diffeomorphism group* to underline this property.

3.1.2 Jet bundles and differential operators Let X be a smooth projective variety. Following Green and Griffiths [21] we let $J_kX\to X$ be the bundle of k–jets of germs of parametrised curves in X; that is, J_kX is the of equivalence classes of germs of holomorphic maps $f\colon(\mathbb{C},0)\to(X,p)$, with the equivalence relation $f\sim g$ if and only if the derivatives satisfy $f^{(j)}(0)=g^{(j)}(0)$ for $0\le j\le k$ when computed in some local coordinate system of X near $p \in X$. The projection map $J_k X \to X$ is simply $f \mapsto f(0)$. If we choose local holomorphic coordinates on an open neighbourhood $\Omega \subset X$ around p, the elements of the fibre $J_k X_p$ can be represented by Taylor expansions

$$f(t) = p + tf'(0) + \frac{t^2}{2!} f''(0) + \cdots + \frac{t^k}{k!} f^{(k)}(0) + O(t^{k+1})$$

up to order k at $t = 0$ of $\mathbb{C}^n$–valued maps $f = (f_1, f_2, \dots, f_n)$ on open neighbourhoods of 0 in $\mathbb{C}$. Locally in these coordinates the fibre $J_k X_p$ can be identified with the set of k–tuples of vectors $(f'(0), \dots, f^{(k)}(0)/k!) = (\mathbb{C}^n)^k$ which further can be identified with $J_k(1,n)$. These jet bundles and the corresponding jet differential bundles play central role in the study of hyperbolic varieties and the Green–Griffiths–Lang conjecture; see Demailly [11] and Green and Griffiths [21].

Remark 3.5 Note that $J_k X$ is not a vector bundle over X since the transition functions are polynomial but not linear; see Section 5 of Demailly [11]. In fact, let Diff_X denote the principal $\mathrm{Diff}_k(n)$–bundle over X formed by all local polynomial coordinate systems on X. Then

$$J_k X = \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} J_k(1,n)$$

is the associated bundle whose structure group is $\mathrm{Diff}_k(n)$.

Let $J_k^{\mathrm{reg}} X$ denote the bundle of k–jets of germs of parametrised regular curves in X, that is, where the first derivative satisfies $f' \neq 0$. After fixing local coordinates near $p \in X$ the fibre $J_k^{\mathrm{reg}} X_p$ can be identified with $J_k^{\mathrm{reg}}(1,n)$ and

$$J_k^{\mathrm{reg}} X = \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} J_k^{\mathrm{reg}}(1,n).$$

Let $\mathcal{D}_X^{\leq k}$ denote the bundle of k^{th}–order differential operators over X. Then we have $\mathcal{D}_X^{\leq 0} = \mathcal{O}_X$, and we let $\mathcal{D}_X^k = \mathcal{D}_X^{\leq k}/\mathcal{D}_X^{\leq 0}$. We have a filtration

$$\mathcal{O}_X = \mathcal{D}_X^{\leq 0} \subset \mathcal{D}_X^{\leq 1} \subset \cdots \subset \mathcal{D}_X^{\leq k}, \tag{6}$$

where the graded component $\mathcal{D}_X^{\leq i}/\mathcal{D}_X^{\leq i-1} \simeq \mathrm{Sym}^i T_X$ but this filtration is not split in general, so $\mathcal{D}_X^k \neq \mathrm{Sym}^{\leq k} T_X$; see Section 4.1 for details. Recall from Remark 3.4 that after choosing local coordinates on X near p the fibre $\mathcal{D}_{X,p}^k$ can be identified with the space $J_k(n,1)^* \simeq \mathrm{Sym}^{\leq k}\mathbb{C}^n$ of k^{th}–order differential operators on $\mathbb{C}^n$ and the filtration (6) restricted to this fibre is the one given in (2).

Remark 3.6 We have a description of $\mathcal{D}_X^k$ as an associated bundle similar to that of $J_k X$ in Remark 3.5 , namely

$$\mathcal{D}_X^k = \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} \mathrm{Sym}^{\leq k}\mathbb{C}^n.$$

Remark 3.7 Given a regular k–jet $(\mathbb{C},0)\to(X,p)$ we may push forward the differential operators of order k on $\mathbb{C}$ to X and obtain a k–dimensional subspace of $\mathcal{D}_{X,p}^{\le k}$. This gives the bundle map

$$J_k^{\mathrm{reg}}X\to \mathrm{Grass}_k(\mathcal{D}_X^k), \tag{7}$$

which is the fibred version of the map in Remark 3.4. Note that $\mathrm{Diff}_k(1)=J_k^{\mathrm{reg}}(1,1)$ acts fibrewise on the jet bundle $J_k^{\mathrm{reg}}X$ via the composition map (5) and the map (7) is $\mathrm{Diff}_k(1)$–invariant, resulting in an embedding

$$J_k^{\mathrm{reg}}X/\mathrm{Diff}_k(1)\hookrightarrow \mathrm{Grass}_k(\mathcal{D}_X^k). \tag{8}$$

In Section 3.1.3 we show that the set $CX^{[k+1]}$ of curvilinear subschemes on X can be identified with the nonreductive fibrewise quotient of $J_k^{\mathrm{reg}}X$ by $\mathrm{Diff}_k(1)$:

$$CX^{[k+1]}=J_k^{\mathrm{reg}}X/\mathrm{Diff}_k(1).$$

This, together with (8) gives an embedding

$$CX^{[k+1]}\hookrightarrow \mathrm{Grass}_k(\mathcal{D}_X^k).$$

In the next subsection we describe a parametrisation of this embedding which turns out to be crucial to control equivariant localisation on $\overline{CX}_p^{[k+1]}$.

3.1.3 The test-curve model of $\overline{CX}^{[k]}$ Let $\xi\in CX_p^{[k+1]}$ be a curvilinear subscheme supported at $p\in X$. Then ξ is (scheme-theoretically) contained in a smooth curve germ $\mathcal{C}_p$ in X:

$$\xi\subset\mathcal{C}_p\subset X.$$

Let $f_\xi\colon(\mathbb{C},0)\to(X,p)$ be a k–jet of a germ parametrising $\mathcal{C}_p$. Then $f_\xi\in J_k^{\mathrm{reg}}X_p$ is determined only up to polynomial reparametrisation germs $\phi\colon(\mathbb{C},0)\to(\mathbb{C},0)$ and therefore we get the following lemma.

Lemma 3.8 *The punctual curvilinear locus $CX_p^{[k+1]}$ is equal (as a set) to the set of k–jets of regular germs at $p\in X$ modulo polynomial reparametrisations:*

$$\begin{aligned}CX_p^{[k+1]}&=\{\text{regular } k\text{–jets } (\mathbb{C},0)\to(X,p)\}/\{\text{regular } k\text{–jets } (\mathbb{C},0)\to(\mathbb{C},0)\}\\&=J_k^{\mathrm{reg}}X_p/\mathrm{Diff}_k(1).\end{aligned}$$

Therefore the curvilinear locus $CX^{[k+1]}$ is the fibrewise quotient

$$CX^{[k+1]}=J_k^{\mathrm{reg}}X/\mathrm{Diff}_k(1).$$

Recall that after choosing local coordinates on X near p we can identify $J_k^{\mathrm{reg}}X_p$ with $J_k^{\mathrm{reg}}(1,n)$. We can explicitly write out the reparametrisation action (defined in (5)) of $\mathrm{Diff}_k(1)$ on $J_k^{\mathrm{reg}}(1,n)$ as follows:

Let
$$f_\xi(z)=zf'(0)+\frac{z^2}{2!}f''(0)+\cdots+\frac{z^k}{k!}f^{(k)}(0)\in J_k^{\mathrm{reg}}(1,n)$$
be the k–jet of a germ at the origin (ie it has no constant term) in $\mathbb{C}^n$ with $f^{(i)}\in\mathbb{C}^n$ such that $f'\neq 0$ and let $\varphi(z)=\alpha_1z+\alpha_2z^2+\cdots+\alpha_kz^k\in J_k^{\mathrm{reg}}(1,1)$ with $\alpha_i\in\mathbb{C}$ and $\alpha_1\neq 0$. Then
$$\begin{aligned}&f\circ\varphi(z)\\&=(f'(0)\alpha_1)z+\left(f'(0)\alpha_2+\frac{f''(0)}{2!}\alpha_1^2\right)z^2+\cdots+\left(\sum_{i_1+\cdots+i_l=k}\frac{f^{(l)}(0)}{l!}\alpha_{i_1}\cdots\alpha_{i_l}\right)z^k,\end{aligned}$$
which equals
$$(f'(0),\ldots,f^{(k)}(0)/k!)\cdot\begin{pmatrix}\alpha_1&\alpha_2&\alpha_3&\cdots&\alpha_k\\0&\alpha_1^2&2\alpha_1\alpha_2&\cdots&2\alpha_1\alpha_{k-1}+\cdots\\0&0&\alpha_1^3&\cdots&3\alpha_1^2\alpha_{k-2}+\cdots\\0&0&0&\cdots&\cdot\\\vdots&\vdots&\vdots&\cdots&\vdots\\0&0&0&\cdots&\alpha_1^k\end{pmatrix},\tag{9}$$
where the (i,j) entry is $p_{i,j}(\bar{\alpha})=\sum_{a_1+a_2+\cdots+a_i=j}\alpha_{a_1}\alpha_{a_2}\cdots\alpha_{a_i}$.

Remark 3.9 The linearisation of the action of $\mathrm{Diff}_k(1)$ on $J_k^{\mathrm{reg}}(1,n)$ given as the matrix multiplication in (9) represents $\mathrm{Diff}_k(1)$ as a group of upper triangular matrices in $\mathrm{GL}(n)$. This is a nonreductive group so Mumford's reductive GIT is not applicable to study the geometry of the quotient $J_k^{\mathrm{reg}}(1,n)/\mathrm{Diff}_k(1)$; see Bérczi, Doran, Hawes and Kirwan [3; 4] for details. Note that our matrix group is parametrised along its first row with the free parameters $\alpha_1,\ldots,\alpha_k$ and the other entries are certain (weighted homogeneous) polynomials in these free parameters. It is a $\mathbb{C}^*$ extension of its maximal unipotent radical
$$\mathrm{Diff}_k(1)=U\rtimes\mathbb{C}^*,$$
where U is the subgroup we get via substituting $\alpha_1=1$ and the diagonal $\mathbb{C}^*$ acts with weights $0,1,\ldots,n-1$ on the Lie algebra $\mathrm{Lie}(U)$. In Bérczi and Kirwan [5] and Bérczi, Doran, Hawes and Kirwan [3; 4] we study actions of groups of this type in a more general context.

Fix an integer $N\geq 1$ and define
$$\Theta_k=\{\Psi\in J_k(n,N):\exists\gamma\in J_k^{\mathrm{reg}}(1,n)\text{ such that }\Psi\circ\gamma=0\},$$
that is, Θ_k is the set of those k–jets of germs on $\mathbb{C}^n$ at the origin which vanish on some regular curve. By definition, Θ_k is the image of the closed subvariety

of $J_k(n,N) \times J_k^{\mathrm{reg}}(1,n)$ defined by the algebraic equations $\Psi \circ \gamma = 0$, under the projection to the first factor. If $\Psi \circ \gamma = 0$, we call γ a *test curve* of Θ.

Remark 3.10 The subset Θ_k is the closure of an important singularity class in the jet space $J_k(n,N)$. These are called Morin singularities and the equivariant dual of Θ_k in $J_k(n,N)$ is called the Thom polynomial of Morin singularities; see Bérczi and Szenes [7] and Fehér and Rimányi [16] for details.

Test curves of germs are generally not unique. A basic but crucial observation is the following. If γ is a test curve of $\Psi \in \Theta_k$, and $\varphi \in \mathrm{Diff}_k(1)$ is a holomorphic reparametrisation of $\mathbb{C}$, then $\gamma \circ \varphi$ is, again, a test curve of Ψ:

$$\mathbb{C} \xrightarrow{\varphi} \mathbb{C} \xrightarrow{\gamma} \mathbb{C}^n \xrightarrow{\Psi} \mathbb{C}^N \text{ with } \Psi \circ \gamma = 0 \quad \Longrightarrow \quad \Psi \circ (\gamma \circ \varphi) = 0.$$

In fact, we get all test curves of Ψ in this way if the following property, open and dense in Θ_k, holds: the linear part of Ψ has 1–dimensional kernel. Before stating this in Theorem 3.12, let us write down the equation $\Psi \circ \gamma = 0$ in coordinates in an illustrative case. Let $\gamma = (\gamma', \gamma'', \dots, \gamma^{(k)}) \in J_k^{\mathrm{reg}}(1,n)$ and $\Psi = (\Psi', \Psi'', \dots, \Psi^{(k)}) \in J_k(n,N)$ be the k–jets of the test curve γ and the map Ψ respectively. Using the chain rule and the notation $v_i = \gamma^{(i)}/i!$, the equation $\Psi \circ \gamma = 0$ reads as follows for $k = 4$:

$$\begin{gathered}\Psi'(v_1) = 0,\\ \Psi'(v_2) + \Psi''(v_1, v_1) = 0,\\ \Psi'(v_3) + 2\Psi''(v_1, v_2) + \Psi'''(v_1, v_1, v_1) = 0,\\ \Psi'(v_4)+2\Psi''(v_1,v_3)+\Psi''(v_2,v_2)+3\Psi'''(v_1,v_1,v_2)+\Psi''''(v_1,v_1,v_1,v_1)=0.\end{gathered} \tag{10}$$

Lemma 3.11 (Gaffney [19]; Bérczi and Szenes [7]) *Let $\gamma = (\gamma', \gamma'', \dots, \gamma^{(k)}) \in J_k^{\mathrm{reg}}(1,n)$ and $\Psi = (\Psi', \Psi'', \dots, \Psi^{(k)}) \in J_k(n,N)$ be k–jets. Then substituting $v_i = \gamma^{(i)}/i!$, the equation $\Psi \circ \gamma = 0$ is equivalent to the following system of k linear equations with values in $\mathbb{C}^N$:*

$$\sum_{\tau \in \mathcal{P}(m)} \Psi(\boldsymbol{v}_\tau) = 0 \quad \text{for } m = 1, 2, \dots, k. \tag{11}$$

Here $\mathcal{P}(m)$ denotes the set of partitions $\tau = 1^{\tau_1} \dots m^{\tau_m}$ of m into nonnegative integers and $\boldsymbol{v}_\tau = v_1^{\tau_1} \cdots v_m^{\tau_m}$.

For a given $\gamma \in J_k^{\mathrm{reg}}(1,n)$ and $1 \le i \le k$ let $\mathcal{S}_\gamma^{i,N}$ denote the set of solutions of the first i equations in (11), that is,

$$\mathcal{S}_\gamma^{i,N} = \{\Psi \in J_k(n,N) : \Psi \circ \gamma = 0 \text{ up to order } i\}. \tag{12}$$

The equations (11) are linear in Ψ, and hence

$$\mathcal{S}_\gamma^{i,N} \subset J_k(n,N)$$

is a linear subspace of codimension iN, ie a point of $\mathrm{Grass}_{\mathrm{codim}=iN}(J_k(n,N))$, whose orthogonal, $(\mathcal{S}_\gamma^{i,N})^\perp$, is an iN–dimensional subspace of $J_k(n,N)^*$. These subspaces are invariant under the reparametrisation of γ. In fact, $\Psi \circ \gamma$ has N vanishing coordinates and therefore

$$(\mathcal{S}_\gamma^{i,N})^\perp = (\mathcal{S}_\gamma^{i,1})^\perp \otimes \mathbb{C}^N.$$

For $\Psi \in J_k(n,N)$ let $\Psi^1 \in \mathrm{Hom}(\mathbb{C}^n,\mathbb{C}^N)$ denote the linear part. When $N \geq n$ then the subset

$$\widetilde{\mathcal{S}}_\gamma^{i,N} = \{\Psi \in \mathcal{S}_\gamma^{i,N} : \dim\ker\Psi^1 = 1\}$$

is an open dense subset of the subspace $\mathcal{S}_\gamma^{i,N}$. In fact it is not hard to see that the complement $\widetilde{\mathcal{S}}_\gamma^{i,N} \setminus \mathcal{S}_\gamma^{i,N}$ where the kernel of Ψ^1 has dimension at least two is a closed subvariety of codimension $N-n+2$.

Theorem 3.12 (1) *The map*

$$\phi\colon J_k^{\mathrm{reg}}(1,n) \to \mathrm{Grass}_k(J_k(n,1)^*)$$

defined as $\gamma \mapsto (\mathcal{S}_\gamma^{k,1})^\perp$ *is* $\mathrm{Diff}_k(1)$*–invariant and induces an injective map on the* $\mathrm{Diff}_k(1)$*–orbits into the Grassmannian*

$$\phi^{\mathrm{Grass}}\colon J_k^{\mathrm{reg}}(1,n)/\mathrm{Diff}_k(1) \hookrightarrow \mathrm{Grass}_k(J_k(n,1)^*).$$

Moreover, ϕ *and* ϕ^{Grass} *are* $\mathrm{GL}(n)$*–equivariant with respect to the standard action of* $\mathrm{GL}(n)$ *on* $J_k^{\mathrm{reg}}(1,n) \subset \mathrm{Hom}(\mathbb{C}^k,\mathbb{C}^n)$ *and the induced action on* $\mathrm{Grass}_k(J_k(n,1)^*)$.

(2) *Recall form Remark 3.4 that* $J_k(n,1)^* = \mathrm{Sym}^{\leq k}\mathbb{C}^n$. *The image of* ϕ *and the image of* φ *defined in Remark 3.3 coincide in* $\mathrm{Grass}_k(\mathrm{Sym}^{\leq k}\mathbb{C}^n)$*:*

$$\mathrm{Im}(\phi) = \mathrm{Im}(\varphi) \subset \mathrm{Grass}_k(\mathrm{Sym}^{\leq k}\mathbb{C}^n).$$

Proof For the first part it is enough to prove that for $\Psi \in \Theta_k$ with $\dim\ker\Psi^1 = 1$ and $\gamma,\delta \in J_k^{\mathrm{reg}}(1,n)$,

$$\Psi\circ\gamma = \Psi\circ\delta = 0 \iff \exists \Delta \in J_k^{\mathrm{reg}}(1,1) \text{ such that } \gamma = \delta\circ\Delta.$$

We prove this statement by induction. Let $\gamma = v_1t+\cdots+v_kt^k$ and $\delta = w_1t+\cdots+w_kt^k$. Since $\dim\ker\Psi^1 = 1$, we have $v_1 = \lambda w_1$, for some $\lambda \neq 0$. This proves the $k=1$ case.

Suppose the statement is true for $k-1$. Then, using the appropriate order-$(k{-}1)$ diffeomorphism, we can assume that $v_m = w_m$ for $m = 1,\dots,k-1$. It is clear then from the explicit form (11) (see (10)) of the equation $\Psi\circ\gamma=0$, that $\Psi^1(v_k)=\Psi^1(w_k)$, and hence $w_k=v_k-\lambda v_1$ for some $\lambda\in\mathbb{C}$. Then $\gamma=\Delta\circ\delta$ for $\Delta=t+\lambda t^k$, and the proof is complete.

The second part immediately follows from the definition of φ and ϕ. □

Remark 3.13 (1) For a point $\gamma\in J_k^{\mathrm{reg}}(1,n)$ let $v_i=\gamma^{(i)}/i!\in\mathbb{C}^n$ denote the normed i^{th} derivative. Then from Lemma 3.11 it immediately follows that for $1\le i\le k$ (see [7]),

(13) $$(\mathcal{S}_\gamma^{i,1})^\perp=\mathrm{Span}_{\mathbb{C}}\Big(v_1,v_2+v_1^2,\dots,\sum_{\tau\in\mathcal{P}(i)}\boldsymbol{v}_\tau\Big)\subset\mathrm{Sym}^{\le k}\mathbb{C}^n.$$

This explicit parametrisation of the curvilinear component is crucial in building our localisation process in the next section.

(2) Let $\{e_1,\dots,e_n\}$ be a basis of $\mathbb{C}^n$. Since ϕ is $\mathrm{GL}(n)$–equivariant, for $k\le n$ the $\mathrm{GL}(n)$–orbit of $p_{k,n}$ satisfies

$$p_{k,n}=\phi(e_1,\dots,e_k)=\mathrm{Span}_{\mathbb{C}}\Big(e_1,e_2+e_1^2,\dots,\sum_{\tau\in\mathcal{P}(k)}e_\tau\Big)$$

and forms a dense subset of the image $J_k^{\mathrm{reg}}(1,n)$ and therefore

$$\overline{\phi(J_k^{\mathrm{reg}}(1,n))}=\overline{\mathrm{GL}(n)\cdot p_{k,n}}.$$

Recall that after choosing local coordinates on X near p we can identify the fibre $J_k^{\mathrm{reg}}X_p$ with $J_k^{\mathrm{reg}}(1,n)$ and the fibre $\mathcal{D}_{X,p}^k$ with $J_k(n,1)^*$. Lemma 3.8 and Theorem 3.12 therefore give us the following.

Corollary 3.14 *We have an embedding of the punctual curvilinear locus $CX_p^{[k+1]}$,*

$$\phi_p^{\mathrm{Grass}}\colon CX_p^{[k+1]}=J_k^{\mathrm{reg}}X_p/\mathrm{Diff}_k(1)\hookrightarrow\mathrm{Grass}_k(\mathcal{D}_{X,p}^k),$$

into the Grassmannian bundle of k–dimensional subspaces of the fibre $\mathcal{D}_{X,p}^k$. The quotient $J_k^{\mathrm{reg}}X/\mathrm{Diff}_k(1)$ has the structure of a locally trivial bundle over X which has a holomorphic embedding

$$\phi^{\mathrm{Grass}}\colon CX^{[k+1]}=J_k^{\mathrm{reg}}X/\mathrm{Diff}_k(1)\hookrightarrow\mathrm{Grass}_k(\mathcal{D}_X^k)$$

into the Grassmannian bundle of k–dimensional subspaces of $\mathcal{D}_X^k$ over X. The closure of the image

$$\overline{CX}^{[k+1]}=\overline{\phi^{\mathrm{Grass}}(J_k^{\mathrm{reg}}X)}$$

is the curvilinear component of the Hilbert scheme of $k+1$ points on X.

3.2 Tautological bundles over $\overline{CX}^{[k]}$

Let F be a rank-r vector bundle over X. The fibre of the corresponding rank-rk tautological bundle $F^{[k]}$ on $\overline{CX}^{[k]}$ at the point ξ is

$$F_\xi^{[k]} = H^0(\xi, F|_\xi) = H^0(\mathcal{O}_\xi \otimes F).$$

On the level of bundles we have the following.

Lemma 3.15 *There is an isomorphism of topological vector bundles $F^{[k]}|_{\overline{CX}^{[k]}} \simeq \mathcal{O}^{[k]}_{\overline{CX}^{[k]}} \otimes \pi^* F$ where $\pi\colon \overline{CX}^{[k]} \to X$ is the projection.*

Proof This comes from Rennemo [33, Lemma 5.2] as follows. Let us adopt the notations of [33] and denote by $X^{[\![k]\!]}$ the Hilbert scheme of k ordered points on X and let $X_0^{[\![k]\!]} \subset X^{[\![k]\!]}$ be the set of pairs $(\xi, (x_i))$ such that ξ is supported at a single point, that is, $x_1 = \cdots = x_k$. Let $\overline{TX} = \mathbb{P}(\mathcal{O}_X \oplus TX)$ denote the natural fibrewise compactification of the tangent bundle. Let $\overline{TX}_0^{[\![k]\!]} \subset \overline{TX}^{[\![k]\!]}$ denote the set of pairs $(\xi, (x_i))$ such that ξ is supported at the 0–section of $\overline{TX}$. Let $q\colon X^{[\![k]\!]} \to X$ be defined by $q(\xi, (x_i)) = x_1$ and let $r\colon \overline{TX}^{[\![k]\!]} \to X$ be defined by $r((\xi, (x_i))) = \pi(x_1)$. Let W be the set of pairs $(\xi, (x_i)) \in \overline{TX}^{[\![k]\!]}$ such that x_1 lies in the 0–section of $\overline{TX}$.

Rennemo [33] shows that there is an open neighbourhood U of $X_0^{[\![k]\!]}$ in $X^{[\![k]\!]}$, an open neighbourhood $\mathcal{U}$ of $\overline{TX}_0^{[\![k]\!]}$ in W, and a homeomorphism $f\colon U \to \mathcal{U}$ such that $q = r \circ f$ and $f|_{q^{-1}(x)}$ is holomorphic for all $x \in X$. Furthermore, there is an isomorphism of topological vector bundles $f^*(\mathcal{F}^{[\![k]\!]}) \simeq F^{[k]}$.

In particular, f is constructed using a similar but simpler statement about the neighbourhood of the diagonal in $X \times X$. Let $p_1, p_2\colon X \times X \to X$ be the projections to the first and second factors, and let $\pi\colon TX \to X$ be the tangent bundle. There is an open neighbourhood U_1 of the diagonal $\delta \subset X \times X$, an open neighbourhood $\mathcal{U}_1$ of the 0–section $X \subset TX$ and a homeomorphism $f_1\colon U_1 \to \mathcal{U}_1$, such that $\pi \circ f_1 = p_1$ and such that $f_1|_\Delta$ is the identification between Δ and the 0–section of TX. Furthermore, the restriction of f_1 to each fibre $p^{-1}(x)$ is holomorphic. There is an isomorphism of topological vector bundles $p_1^*(E)|_U \to p_2^*(E)|_U$, which is an isomorphism of holomorphic bundles on the restriction to each fibre $p^{-1}(x)$.

Then f is given by

$$f\big((\xi, (x_i))\big) = \big((f_1)_*(\{q(x)\} \times \xi), f_1(q(x), x_i)\big)$$

on a small neighbourhood U of $X_0^{[\![k]\!]}$ and over a point $(\xi, (x_i)) \in X^{[\![k]\!]}$ we have

$$\begin{aligned} f^*(\mathcal{F}^{[\![k]\!]})_{(\xi,(x_i))} &= H^0(\{x_1\} \times \xi, p_1^*(F)|_{\{x_1\} \times \xi}) \\ &\simeq H^0(\{x_1\} \times \xi, p_2^*(F)|_{\{x_1\} \times \xi}) = F^{[\![k]\!]}_{(\xi,(x_i))}. \end{aligned}$$

For $(\xi,(x_i)) \in \overline{CX}_p^{[k]} \subset X_p^{[\![k]\!]}$ we have $p = x_1 = \cdots = x_k$ and therefore

$$F_\xi^{[\![k]\!]} = H^0(\{x_1\}\times\xi, p_1^*(F)|_{\{x_1\}\times\xi}) = F_p \otimes H^0(\mathcal{O}_\xi),$$

which gives the isomorphism of the lemma. □

Our embedding $\phi^{\mathrm{Grass}}\colon \overline{CX}^{[k+1]} \hookrightarrow \mathrm{Grass}_k(\mathcal{D}^k)$ then identifies the fibres of $\mathcal{O}^{[k+1]}_{\overline{CX}^{[k+1]}}$ over $\xi \in \overline{CX}_p^{[k+1]}$ with

$$H^0(\mathcal{O}_\xi) \simeq \mathcal{O}_p \otimes \mathcal{E}_{\phi^{\mathrm{Grass}}(\xi)},$$

where $\mathcal{E}$ is the tautological rank-k bundle over $\mathrm{Grass}_k(\mathcal{D}^k)$. Hence the total Chern class of $F^{[k+1]}$ can be written as

$$c(F^{[k+1]}) = \prod_{j=1}^{r}(1+\theta_j)\prod_{i=1}^{k}\prod_{j=1}^{r}(1+\eta_i+\theta_j), \tag{14}$$

where $c(F) = \prod_{j=1}^r(1+\theta_j)$ and $c(\mathcal{E}) = \prod_{i=1}^k(1+\eta_i)$ are the Chern classes for the corresponding bundles. In particular the Chern class

$$c_i(F^{[k+1]}) = \mathcal{C}_i(c_1(\mathcal{E}),\dots c_k(\mathcal{E}), c_1(F),\dots,c_r(F)) \tag{15}$$

can be expressed as a polynomial function $\mathcal{C}_i$ in Chern classes of $\mathcal{E}$ and F.

4 Partial resolutions of $\overline{CX}^{[k+1]}$

The structure group of the bundles $\overline{CX}^{[k+1]}$ and $\mathrm{Grass}_k(\mathcal{D}_X^k)$ is the polynomial reparametrisation group $\mathrm{Diff}_k(n)$. The subgroup $\mathrm{GL}(n)$ of linear coordinate changes sit in $\mathrm{Diff}_k(n)$ and using this in Section 4.1 we define the corresponding linearised bundles $\overline{CX}_{\mathrm{GL}}^{[k+1]} \subset \mathrm{Grass}_k(\mathrm{Sym}^{\le k} T_X)$. In Section 4.2 we construct a fibrewise partial resolution of the (highly singular) curvilinear component $\overline{CX}^{[k+1]} \subset \mathrm{Grass}_k(\mathcal{D}_X^k)$ and also its linearised bundle $\overline{CX}_{\mathrm{GL}}^{[k+1]}$. This resolution is defined for any choice of parameters n,k and it uses nested Hilbert schemes. In Section 4.3 we construct a second partial resolution of $\overline{CX}_{\mathrm{GL}}^{[k+1]}$ under the very restrictive condition $k \le n$, that is, the number of points can't exceed the dimension of the variety plus one. We will see how to dispose this condition in Section 6.2.

4.1 Linearisation of $\overline{CX}^{[k+1]}$

Recall from Section 3.1.1 the notation $\mathrm{Diff}_k(n) = J_k^{\mathrm{reg}}(n,n)$ for the group of k^{th}–order diffeomorphism germs of $\mathbb{C}^n$ at the origin. Then $\mathrm{Diff}_k(n)$ is the set of local (polynomial) coordinate changes on $\mathbb{C}^n$ at the origin. The set $\mathrm{GL}(n)$ of linear coordinate changes forms a subgroup of $\mathrm{Diff}_k(n)$.

We have seen that after choosing local coordinates on X near p we can identify the fibre $\mathcal{D}^k_{X,p}$ of the bundle $\mathcal{D}^k_X$ with $J_k(n,1)^* = \mathrm{Sym}^{\le k}\mathbb{C}^n$. Then Diff_k (and therefore its subgroup $\mathrm{GL}(n)$) acts on $\mathcal{D}^k_{X,p} \simeq \mathrm{Sym}^{\le k}\mathbb{C}^n$. Let Diff_X denote the principal $\mathrm{Diff}_k(n)$–bundle over X formed by all local polynomial coordinate systems on X. Then $\mathcal{D}^k_X$ can be described as the associated bundle (see also Remark 3.4)

$$\mathcal{D}^k_X = \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} \mathrm{Sym}^{\le k}\mathbb{C}^n.$$

On the other hand, if GL_X denotes the principal $\mathrm{GL}(n)$–bundle over X formed by all local linear coordinate systems on X then the vector bundle $\mathrm{Sym}^{\le k} T_X = \bigoplus_{i=1}^k \mathrm{Sym}^i T_X$ is associated to the same $\mathrm{Sym}^{\le k}\mathbb{C}^n$ considered as a $\mathrm{GL}(n)$–module:

$$\mathrm{Sym}^{\le k} T_X = \mathrm{GL}_X \times_{\mathrm{GL}(n)} \mathrm{Sym}^{\le k}\mathbb{C}^n.$$

Therefore $\mathcal{D}^k_X$ and $\mathrm{Sym}^{\le k} T_X$ are not isomorphic bundles and in particular the filtration defined in (6) does not split. Hence there is no projection map $\mathcal{D}^k_X \to T_X$ but there is a natural projection $\mathrm{Sym}^{\le k} T_X \to T_X$.

There is an induced $\mathrm{Diff}_k(n)$–action on $\mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ and the image $\mathrm{Im}(\phi^{\mathrm{Grass}})$ in Theorem 3.12 is $\mathrm{Diff}_k(n)$–invariant subvariety of $\mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$. The curvilinear locus $\overline{CX}^{[k+1]}$ is the associated bundle

$$\begin{aligned}\overline{CX}^{[k+1]} &= \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} \overline{\mathrm{Im}(\phi^{\mathrm{Grass}})} \\ &\subset \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \mathrm{Grass}_k(\mathcal{D}^k_X).\end{aligned}$$

We can form the corresponding linearised bundle

$$\begin{aligned}\overline{CX}^{[k+1]}_{\mathrm{GL}} &= \mathrm{GL}_X \times_{\mathrm{GL}(n)} \overline{\mathrm{Im}(\phi^{\mathrm{Grass}})} \\ &\subset \mathrm{GL}_X \times_{\mathrm{GL}(n)} \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \mathrm{Grass}_k(\mathrm{Sym}^{\le k} T_X),\end{aligned}$$

which is the linearised version of $\overline{CX}^{[k+1]}$ remembering the linear action on the fibres. We will explain in Section 7 that for torus localisation purposes we can replace $\overline{CX}^{[k+1]}$ with its linearised version $\overline{CX}^{[k+1]}_{\mathrm{GL}}$.

4.2 Completion in nested Hilbert schemes

Let

$$X^{[k_1,\dots,k_t]} = \{(\xi_1 \subset \xi_2 \subset \cdots \subset \xi_t) : \xi_i \in X^{[k_i]}\} \subset X^{[k_1]} \times \cdots \times X^{[k_t]}$$

denote the nested Hilbert scheme defining flags of subschemes of length vector $(k_1,\dots,k_t)$.

Curvilinear subschemes contain only one subscheme for any given smaller length. Therefore $\xi \in CX_p^{[k+1]}$ defines a unique flag

$$\mathcal{F}(\xi) = (\xi_1 \subset \xi_2 \subset \cdots \subset \xi_k) \in CX_p^{[2]} \times \cdots \times CX_p^{[k+1]} \subset X^{[2,\dots,k+1]},$$

where ξ_i is the unique subscheme of ξ satisfying

$$\mathcal{O}_{\xi_i} = \mathcal{O}_\xi / \mathcal{O}_{X,p}^{i+1} \simeq \mathbb{C}[z]/z^{i+1},$$

and therefore $\xi_i \in CX_p^{[i+1]}$. This defines an embedding

$$\widetilde{\phi}\colon CX_p^{[k+1]} \hookrightarrow X^{[2,\dots,k+1]}, \quad \xi \mapsto (\xi_1 \subset \cdots \subset \xi_k).$$

Fix local coordinates on X near p such that $J_k X_p$ is identified with $J_k(1,n)$ and $\mathcal{D}^k_{X,p}$ is identified with $J_k(n,1)^*$. Let $f_\xi \in J_k^{\mathrm{reg}}(1,n)$ denote the k–jet corresponding to $\xi \in CX_p^{[k+1]}$ and let $\mathcal{S}^i_\xi = \mathcal{S}^{i,1}_{f_\xi} \subset J_k(n,1)$ be the solution space defined in (12) where $N = 1$. Then $\widetilde{\phi}$ can be equivalently written as

$$f_\xi \mapsto ((\mathcal{S}^1_\xi)^\perp \subset (\mathcal{S}^2_\xi)^\perp \subset \cdots \subset (\mathcal{S}^k_\xi)^\perp) \in \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$$

or, using coordinates, as

$$f_\xi \mapsto \Big(\mathrm{Span}(f') \subset \mathrm{Span}(f', f''+(f')^2) \subset \\ \cdots \subset \mathrm{Span}\Big(f', f''+(f')^2, \dots, \sum_{\sum a_i = k} (f^{[i]})^{a_i}\Big)\Big).$$

Theorem 3.12 has the following immediate corollary:

Corollary 4.1 *The map*

$$\widetilde{\phi}\colon J_k^{\mathrm{reg}}(1,n) \to \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n), \quad \gamma \mapsto \mathcal{F}_\gamma = ((\mathcal{S}^1_\gamma)^\perp \subset \cdots \subset (\mathcal{S}^k_\gamma)^\perp)$$

is $\mathrm{Diff}_k(1)$*–invariant and induces an injective map on the* $\mathrm{Diff}_k(1)$*–orbits into the flag manifold*

$$\phi^{\mathrm{Flag}}\colon J_k^{\mathrm{reg}}(1,n)/\mathrm{Diff}_k(1) \hookrightarrow \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n).$$

Moreover, all these maps are $\mathrm{GL}(n)$*–equivariant with respect to the standard action of* $\mathrm{GL}(n)$ *on* $J_k^{\mathrm{reg}}(1,n) \subset \mathrm{Hom}(\mathbb{C}^k, \mathbb{C}^n)$ *and the induced action on* $\mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$.

This implies that similarly to Remark 3.13, for any basis $\{e_1, \dots, e_n\}$ *of* $\mathbb{C}^n$ *and* $k \le n$ *the* $\mathrm{GL}(n)$*–orbit of* $\mathfrak{p}_{k,n} = \widetilde{\phi}(e_1, \dots, e_k)$*, that is, in coordinates,*

$$\mathfrak{p}_{k,n} = \Big(\mathrm{Span}(e_1) \subset \mathrm{Span}(e_1, e_2 + e_1^2) \subset \cdots \subset \mathrm{Span}\Big(e_1, e_2 + e_1^2, \dots, \sum_{\tau \in \mathcal{P}(k)} e_\tau\Big)\Big),$$

forms a dense subset of the image $J_k^{\mathrm{reg}}(1,n)$ *and therefore*

$$\overline{\widetilde{\phi}(J_k^{\mathrm{reg}}(1,n))} = \overline{\mathrm{GL}(n) \cdot \mathfrak{p}_{k,n}}.$$

Definition 4.2 We define the bundle

$$\begin{aligned}\widehat{CX}^{[k+1]} &= \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} \overline{\widetilde{\phi}(J_k^{\mathrm{reg}}(1,n))} \\ &\subset \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \mathrm{Flag}_k(\mathcal{D}_X^k),\end{aligned}$$

which is a fibrewise partial resolution of $\overline{CX}^{[k+1]}$. The corresponding linearised bundle is defined as

$$\begin{aligned}\widehat{CX}_{\mathrm{GL}}^{[k+1]} &= \mathrm{GL}_X \times_{\mathrm{GL}(n)} \overline{\widetilde{\phi}(J_k^{\mathrm{reg}}(1,n))} \\ &\subset \mathrm{GL}_X \times_{\mathrm{GL}(n)} \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \mathrm{Flag}_k(\mathrm{Sym}^{\le k} T_X).\end{aligned}$$

4.3 Blowing up along the linear part

Assume $k \le n$. Let $J_k^{\mathrm{nondeg}}(1,n) \subset J_k^{\mathrm{reg}}(1,n)$ denote the Zariski open set of jets $(\gamma', \gamma'', \dots, \gamma^{[k]})$ with $\gamma', \dots, \gamma^{(k)}$ linearly independent. These correspond to the regular $n \times k$ matrices in $\mathrm{Hom}(\mathbb{C}^k, \mathbb{C}^n)$, and they fibre over the set of complete flags in $\mathbb{C}^n$:

$$J_k^{\mathrm{nondeg}}(1,n) \to \mathrm{Hom}(\mathbb{C}^k, \mathbb{C}^n)/B_k = \mathrm{Flag}_k(\mathbb{C}^n),$$

where $B_k \subset \mathrm{GL}(k)$ is the upper Borel subgroup. Since $J_k^{\mathrm{reg}}(1,1) \subset B_k$ this induces a surjective fibration

$$\pi\colon J_k^{\mathrm{nondeg}}(1,n)/\mathrm{Diff}_k(1) \to \mathrm{Flag}_k(\mathbb{C}^n) \tag{16}$$

which factors through ϕ^{Grass}:

$$\begin{tikzcd} J_k^{\mathrm{nondeg}}(1,n)/\mathrm{Diff}_k(1) \arrow[r, "\phi^{\mathrm{Flag}}"] \arrow[dr, "\pi"] & \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) \arrow[d, dashed] \\ & \mathrm{Flag}_k(\mathbb{C}^n) \end{tikzcd} \tag{17}$$

Here the vertical rational map is induced by the projection $\mathrm{Sym}^{\le k}\mathbb{C}^n \to \mathbb{C}^n$ and the image of ϕ^{Flag} sits in its domain.

Since $J_k^{\mathrm{nondeg}}(1,n) \subset J_k(1,n)$ is $\mathrm{GL}(n)$–invariant, we can form the associated bundle

$$J_k^{\mathrm{nondeg}} X = \mathrm{GL}_X \times_{\mathrm{GL}(n)} J_k^{\mathrm{nondeg}}(1,n).$$

Note, however, that $J_k^{\mathrm{nondeg}} X$ is not a subbundle of $J_k X = \mathrm{Diff}_X \times_{\mathrm{Diff}_k(n)} J_k(1,n)$. Similarly, we can form the bundle

$$\begin{aligned}CX_{\mathrm{nondeg}}^{[k+1]} &= \mathrm{GL}_X \times_{\mathrm{GL}(n)} (\widetilde{\phi}(J_k^{\mathrm{nondeg}}(1,n))) \subset \mathrm{GL}_X \times_{\mathrm{GL}(n)} \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) \\ &= \mathrm{Flag}_k(\mathrm{Sym}^{\le k} T_X),\end{aligned}$$

which is a dense subbundle of $\widehat{CX}^{[k+1]}_{\mathrm{GL}}$ but not a subbundle of $\widehat{CX}^{[k+1]}$. The projection $\mathrm{Sym}^{\le k} T_X \to T_X$ then induces the following diagram whose restriction to the fibres over X was given in (17):

$$\begin{tikzcd} J_k^{\mathrm{nondeg}} X/\mathrm{Diff}_k(1) \arrow[r, "\phi^{\mathrm{Flag}}"] \arrow[dr, "\pi"'] & \mathrm{Flag}_k(\mathrm{Sym}^{\le k} T_X) \arrow[d, dashed] \\ & \mathrm{Flag}_k(T_X) \end{tikzcd} \tag{18}$$

The image of ϕ^{Flag} sits in the domain of the vertical rational map and therefore we have a fibration

$$\mu\colon CX^{[k+1]}_{\mathrm{nondeg}} \to \mathrm{Flag}_k(T_X)$$

of the bundles.

Definition 4.3 Let $\widetilde{CX}^{[k+1]} \to \mathrm{Flag}_k(T_X)$ denote the fibrewise compactification of the bundle $\pi\colon CX^{[k+1]}_{\mathrm{nondeg}} \to \mathrm{Flag}_k(T_X)$. In other words, if $P_{k,n} \subset \mathrm{GL}(n)$ denotes the parabolic subgroup which preserves the flag

$$\boldsymbol{f} = (\mathrm{Span}(e_1) \subset \mathrm{Span}(e_1, e_2) \subset \cdots \subset \mathrm{Span}(e_1, \ldots, e_k) \subset \mathbb{C}^n).$$

and $\mathfrak{p}_{k,n} = \widetilde{\phi}(e_1, \ldots, e_k)$ denotes the base point in $\mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$, then

$$\widetilde{CX}^{[k+1]} = \mathrm{GL}_X \times_{\mathrm{GL}(n)} (\mathrm{GL}(n) \times_{P_{k,n}} \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}),$$

and we have a partial resolution map

$$\begin{aligned} \rho\colon \widetilde{CX}^{[k+1]} &= \mathrm{GL}_X \times_{\mathrm{GL}(n)} (\mathrm{GL}(n) \times_{P_{k,n}} \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}) \\ &\to \mathrm{GL}_X \times_{\mathrm{GL}(n)} (\overline{\mathrm{GL}(n) \cdot \mathfrak{p}_{k,n}}) = \widehat{CX}^{[k+1]}_{\mathrm{GL}}. \end{aligned}$$

Remark 4.4 Equivalently, let $\pi\colon J_k(n,1)^* \simeq \mathrm{Sym}^{\le k}\mathbb{C}^n = \bigoplus_{i=1}^k \mathrm{Sym}^i\mathbb{C}^n \to \mathbb{C}^n$ denote the projection to the first (linear) factor. Then

$$\begin{aligned} &\mathrm{GL}(n) \times_{P_{k,n}} \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}} \\ &= \big\{\big(((\mathcal{S}^1_\gamma)^\perp \subset \cdots \subset (\mathcal{S}^k_\gamma)^\perp), (V_1 \subset \cdots \subset V_k)\big) : \pi(\mathcal{S}^i_\gamma)^\perp \subset V_i\big\} \subset \overline{\mathrm{Im}(\widetilde{\phi})} \times \mathrm{Flag}_k(\mathbb{C}^n). \end{aligned}$$

5 Equivariant localisation on $\widetilde{CX}^{[k+1]}_p$

In this section we fix a point $p \in X$ and a holomorphic coordinate system on X near p. We identify the fibre $J_k X_p$ with $J_k(1,n)$ and $\mathcal{D}^k_{X,p}$ with $J_k(n,1)^* = \mathrm{Sym}^{\le k}\mathbb{C}^n$.

With these identifications we can use Theorem 3.12, Remark 3.13, Definition 4.2, Definition 4.3 and the maps

$$\phi^{\mathrm{Grass}}\colon J_k^{\mathrm{reg}}(1,n)/\mathrm{Diff}_k(1) \hookrightarrow \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$$

and

$$\phi^{\mathrm{Flag}}\colon J_k^{\mathrm{reg}}(1,n)/\mathrm{Diff}_k(1) \hookrightarrow \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$$

to describe the punctual curvilinear Hilbert scheme and its partial resolutions at $p \in X$ as

$$\begin{aligned}
\overline{CX}_p^{[k+1]} &= \overline{\mathrm{Im}(\phi^{\mathrm{Grass}})} = \overline{\mathrm{GL}(n)\cdot \mathfrak{p}_{k,n}} \subset \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n),\\
\widehat{CX}_p^{[k+1]} &= \overline{\mathrm{Im}(\phi^{\mathrm{Flag}})} = \overline{\mathrm{GL}(n)\cdot \mathfrak{p}_{k,n}} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n),\\
\widetilde{CX}_p^{[k+1]} &= \mathrm{GL}(n) \times_{P_{k,n}} \overline{P_{k,n}\cdot \mathfrak{p}_{k,n}} \to \mathrm{Flag}_k(\mathbb{C}^n).
\end{aligned}$$

Let F be a rank-r vector bundle over X and let $F^{[k+1]}$ denote the corresponding rank-$(k+1)r$ tautological bundle over $X^{[k+1]}$. We use the same notation $F^{[k+1]}$ for its pull-back along the partial resolution map $\rho\colon \widetilde{CX}_p^{[k+1]} \to \overline{CX}_p^{[k+1]}$. Then $\widetilde{CX}_p^{[k+1]} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ is endowed with a natural $\mathrm{GL}(n)$–action. In this section we start developing an iterated residue formula for $\int_{\widetilde{CX}_p^{[k]}} \alpha$ for closed torus equivariant forms α. This formula is attained via equivariant localisation process using the fibration $\pi\colon \widetilde{CX}_p^{[k+1]} \to \mathrm{Flag}_k(\mathbb{C}^n)$ and it is crucially based on a vanishing theorem of residues.

5.1 Equivariant de Rham model and the Atiyah–Bott formula

This section is a short introduction to equivariant cohomology and localisation. For more details, we refer the reader to Section 2 of Bérczi and Szenes [7] and Berline, Getzler and Vergne [8].

Let G be a compact Lie group with Lie algebra $\mathfrak{g}$ and let M be a C^∞ manifold endowed with the action of G. The G–equivariant differential forms are defined as differential-form-valued polynomial functions on $\mathfrak{g}$:

$$\Omega_G^\bullet(M) = \{\alpha\colon \mathfrak{g} \to \Omega^\bullet(M) : \alpha(gX) = g\alpha(X) \text{ for } g \in G,\ X \in \mathfrak{g}\} = (S^\bullet\mathfrak{g}^* \otimes \Omega^\bullet(M))^G,$$

where $(g\cdot\alpha)(X) = g\cdot(\alpha(g^{-1}\cdot X))$. One can define equivariant the exterior differential d_G on $(S^\bullet\mathfrak{g}^*\otimes\Omega^\bullet(M))^G$ by the formula

$$(d_G\alpha)(X) = (d - \iota(X_M))\alpha(X),$$

where $\iota(X_M)$ denotes the contraction by the vector field X_M. This increases the degree of an equivariant form by one if the $\mathbb{Z}$–grading is given on $(S^\bullet\mathfrak{g}^*\otimes\Omega^\bullet(M))^G$ by $\deg(P\otimes\alpha) = 2\deg(P) + \deg(\alpha)$ for $P \in S^\bullet\mathfrak{g}^*$ and $\alpha \in \Omega^\bullet(M)$. The homotopy formula $\iota(X)d + d\iota(X) = \mathcal{L}(X)$ implies that $d_G^2(\alpha)(X) = -\mathcal{L}(X)\alpha(X) = 0$

for any $\alpha \in (S^\bullet\mathfrak{g}^* \otimes \Omega^\bullet(M))^G$, and therefore $(d_G, \Omega^\bullet_G(M))$ is a complex. The equivariant cohomology $H^*_G(M)$ of the G–manifold M is the cohomology of the complex $(d_G, \Omega^\bullet_G(M))$. Note that $\alpha \in \Omega^\bullet_G(M)$ is equivariantly closed if and only if

$$\alpha(X) = \alpha(X)^{[0]} + \cdots + \alpha(X)^{[n]} \quad \text{such that } \iota(X_M)\alpha(X)^{[i]} = d\alpha(X)^{[i-2]}.$$

Here $\alpha(X)^{[i]} \in \Omega^i(M)$ is the degree-i part of $\alpha(X) \in \Omega^\bullet(M)$ and $\alpha^{[i]}\colon \mathfrak{g} \to \Omega^i(M)$ is a polynomial function.

The equivariant push-forward map $\int_M\colon \Omega_G(M) \to (S^\bullet\mathfrak{g}^*)^G$ is defined by the formula

$$\left(\int_M \alpha\right)(X) = \int_M \alpha(X) = \int_M \alpha^{[n]}(X). \tag{19}$$

When the n–dimensional complex torus $T = (\mathbb{C}^*)^n$ acts on M let $K = U(1)^n$ be its maximal unipotent subgroup and $\mathfrak{t} = \mathrm{Lie}(K)$ its Lie algebra. We define the T–equivariant cohomology $H^\bullet_T(M)$ to be $H^\bullet_K(M)$, the equivariant de Rham cohomology defined by the action of K. If $M_0(X)$ is the zero locus of the vector field X_M, then the form $\alpha(X)^{[n]}$ is exact outside $M_0(X)$ (see Proposition 7.10 in [8]), and this suggests that the integral $\int_M \alpha(X)$ depends only on the restriction $\alpha(X)|_{M_0(X)}$.

Theorem 5.1 (Atiyah and Bott [1]; Berline and Vergne [9]) *Suppose that M is a compact manifold and T is a complex torus acting smoothly on M, and the fixed-point set M^T of the T–action on M is finite. Then for any cohomology class $\alpha \in H^\bullet_T(M)$,*

$$\int_M \alpha = \sum_{f \in M^T} \frac{\alpha^{[0]}(f)}{\mathrm{Euler}^T(T_f M)}.$$

Here $\mathrm{Euler}^T(T_f M)$ is the T–equivariant Euler class of the tangent space $T_f M$, and $\alpha^{[0]}$ is the differential-form-degree-0 part of α.

The right-hand side in the localisation formula sits in the fraction field of the polynomial ring $H^\bullet_T(\text{point}) = H^\bullet(BT) = S^\bullet\mathfrak{t}^*$. Part of the statement is that the denominators cancel when the sum is simplified.

5.2 Equivariant Poincaré duals and multidegrees

The Atiyah–Bott formula works for holomorphic actions of tori on nonsingular projective varieties. In our case, however, the punctual curvilinear component $\overline{CX}_p^{[k+1]}$ is highly singular at the fixed points so the AB localisation does not apply directly as the equivariant Euler class of the tangent space at a singular fixed point is not well defined. But $\overline{CX}_p^{[k+1]}$ sits in the nonsingular ambient space $\mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ and an intuitive idea would be to put $\mathrm{Euler}^T(T_f \mathrm{Grass}_k(\mathrm{Sym}^{\leq k}\mathbb{C}^n))$ into the denominator on the right-hand side of the equation in Theorem 5.1 which we then compensate in the numerator with some sort of dual of the tangent cone of $\overline{CX}_p^{[k+1]}$ at f sitting in the tangent space of $\mathrm{Grass}_k(\mathrm{Sym}^{\leq k}\mathbb{C}^n)$ at f. This idea indeed works and it becomes incarnate in the Rossmann formula in Section 5.3.

Let $T=(\mathbb{C}^*)^n$ be a complex torus with $K=U(1)^n$ its maximal compact subgroup and $\mathfrak{t}=\mathrm{Lie}(K)$ its Lie algebra. Let M be a manifold endowed with a T–action. The compactly supported equivariant cohomology groups $H^\bullet_{K,\mathrm{cpt}}(M)$ are obtained by restricting the equivariant de Rham complex to compactly supported (or quickly decreasing at infinity) differential forms. Clearly $H^\bullet_{K,\mathrm{cpt}}(M)$ is a module over $H^\bullet_K(M)$. When $M=W$ is an r–dimensional complex vector space, and the action is linear, one has $H^\bullet_K(W)=S^\bullet\mathfrak{t}^*$ and $H^\bullet_{K,\mathrm{cpt}}(W)$ is a free module over $H^\bullet_K(W)$ generated by a single element of degree $2r$ called the Thom class of W:

$$H^\bullet_{K,\mathrm{cpt}}(W)=H^\bullet_K(W)\cdot\mathrm{Thom}_K(W). \tag{20}$$

A T–invariant algebraic subvariety Σ of dimension d in W represents a T–equivariant $2d$–cycle in the sense that

- a compactly supported equivariant form μ is absolutely integrable over the components of maximal dimension of Σ, and $\int_\Sigma \mu \in S^\bullet\mathfrak{t}^*$,
- if $d_K\mu=0$, then $\int_\Sigma\mu$ depends only on the class of μ in $H^\bullet_{K,\mathrm{cpt}}(W)$, and
- $\int_\Sigma\mu=0$ if $\mu=d_K\nu$ for a compactly supported equivariant form ν.

Definition 5.2 Let Σ be an T–invariant algebraic subvariety of dimension d in the vector space W. Then the equivariant Poincaré dual of Σ is the polynomial on $\mathfrak{t}$ of degree $2r-2d$ defined by the integral

$$\mathrm{eP}[\Sigma,W]=\frac{1}{(2\pi)^d}\int_\Sigma \mathrm{Thom}_K(W). \tag{21}$$

An immediate consequence of this definition is that for an equivariantly closed differential form μ with compact support, we have

$$\int_\Sigma\mu=\int_W \mathrm{eP}[\Sigma,W]\cdot\mu.$$

This formula serves as the motivation for the term *equivariant Poincaré dual.* This definition naturally extends to the case of an analytic subvariety of $\mathbb{C}^n$ defined in the neighbourhood of the origin, or more generally, to any T–invariant cycle in $\mathbb{C}^n$. Note that we do not require for Σ to be smooth, and for singular Σ integration on the right-hand side means integration over the smooth part.

The fibred version of Thom classes of vector spaces are the so-called equivariant Thom classes of vector bundles. We recall the definition and basic properties from Section 2.3 of Duflo and Vergne [12] (see also Mathai and Quillen [30]). Let $\pi\colon E \to M$ be a K–equivariant rank-r complex vector bundle and assume M is compact. Then according to [12, Proposition 16], $H^\bullet_{K,\mathrm{cpt}}(E)$ is a free module over $H^\bullet_K(M)$ generated by a single element of degree $2r$ called the equivariant Thom class of E:

$$H^\bullet_{K,\mathrm{cpt}}(E) = H^\bullet_K(M)\cdot \mathrm{Thom}_K(E). \tag{22}$$

The multiplication map $\alpha \mapsto \pi^*(\alpha)\cdot \mathrm{Thom}_K(E)$ establishes an $H^\bullet_K(M)$–module isomorphism from $H^\bullet_K(M)$ to $H^\bullet_{K,\mathrm{cpt}}(E)$. In particular

$$\int_{E/M} \mathrm{Thom}_K(E) = 1$$

holds for the equivariant push-forward map $\int_{E/M}\colon H^\bullet_K(E) \to H^{\bullet-r}_K(M)$. In fact, there is an equivariantly closed form with compact support on E representing $\mathrm{Thom}_K(E)$. By an abuse of notation let $\mathrm{Thom}_K(E) \in \Omega_{K,\mathrm{cpt}}(E) \subset (S^\bullet\mathfrak{t}^* \otimes \Omega^\bullet_{K,\mathrm{cpt}}(E))^K$ denote this compactly supported form.

Note that for nonsingular Σ the definition (21) can be rewritten using the equivariant normal bundle $\mathcal{N}_\Sigma$ of Σ in W as

$$\mathrm{eP}[\Sigma, W] = \mathrm{Thom}_K(\mathcal{N}_\Sigma) \in S^\bullet\mathfrak{t}^*.$$

More generally, let $Z \subset M$ be a T–invariant complex submanifold of codimension r in the complex manifold M. Let $\mathcal{N}_Z$ denote the normal bundle of Z in M. By the equivariant tubular neighbourhood theorem there exists a K–invariant tubular neighbourhood U of Z in M and a K–invariant diffeomorphism $\gamma\colon Z \to U$ such that $\gamma \circ i_0 = i$, where $i_0\colon Z \hookrightarrow \mathcal{N}_Z$ is the embedding of Z into $\mathcal{N}_Z$ as the zero section. Let $\mathrm{Thom}_K(\mathcal{N}_Z) \in \Omega^\bullet_{K,\mathrm{cpt}}(M)$ denote the extension by zero of the equivariant Thom form of $\mathcal{N}_Z$ to M.

Definition 5.3 Let $Z \subset M$ be a T–invariant complex submanifold of codimension r in the (not necessary compact) complex manifold M. Let $\mathcal{N}_Z$ denote the normal bundle of Z in M. The equivariant Poincaré dual of Z is defined as

$$\mathrm{eP}[Z, M] = \mathrm{Thom}_K(\mathcal{N}_Z) \in \Omega^{2r}_{K,\mathrm{cpt}}(M).$$

Then for any closed (not necessarily compactly supported) form $\mu \in \Omega^\bullet_{K,\mathrm{cpt}}(M)$ we have

$$\int_Z \mu = \int_M \mathrm{eP}[Z, M]\cdot \mu.$$

More generally, for a vector bundle $\pi\colon E \to M$ over the compact variety M and $\mu \in \Omega^{\bullet}_K(E)$ we have

$$\int_M \mu = \int_E \mathrm{Thom}_K(E) \cdot \mu. \tag{23}$$

The following lemma is a special case of Proposition 2.8 in [7].

Lemma 5.4 *Let $\pi\colon E \to M$ be a complex vector bundle and $s\colon M \to E$ a smooth section generically transversal to the zero section $\iota\colon M \hookrightarrow E$. Then the zero locus $s^{-1}(M) \subset M$ of s defines a cycle and it is Poincaré dual to the K–equivariant Euler class* $\mathrm{Euler}_K(E) = \iota^* \mathrm{Thom}_K(E)$ *of E.*

Let W be again a complex N–dimensional vector space. Note that $\mathrm{eP}[\Sigma, W]$ is determined by the maximal dimensional components of Σ and in fact it can be characterised and axiomatised by some of its basic properties. These are carefully stated in Bérczi and Szenes [7, Proposition 2.3] and proofs can be found in Rossmann [34], Vergne [37] and Miller and Sturmfels [31]. The list of these properties is the following: positivity, additivity on maximal dimensional component, deformation invariance, symmetry and finally a formula for complete intersections of hypersurfaces. These properties provide an algorithm for computing $\mathrm{eP}[\Sigma, W]$ as follows (see Miller and Sturmfels [31, Chapter 8.5], Bérczi and Szenes [7] and Bérczi [2] for details). We pick any monomial order on the coordinates of W and apply Gröbner deformation on the ideal of Σ to deform it onto its initial monomial ideal (see Eisenbud [14]). The spectrum of this monomial ideal is the union of some coordinate subspaces in W with multiplicities whose equivariant dual is then given as the sum of the duals of the maximal dimensional subspaces by the additivity property. For these linear subspaces the formula for complete intersections has the following special form. Let $W = \mathrm{Spec}(\mathbb{C}[y_1, \dots, y_N])$ acted on by the n–dimensional torus T diagonally where the weight of y_i is η_i. Then for every subset $\boldsymbol{i} \subset \{1, \dots, N\}$ we have

$$\mathrm{eP}[\{y_i = 0,\ i \in \boldsymbol{i}\}, W] = \prod_{i \in \boldsymbol{i}} \eta_i. \tag{24}$$

The weights $\eta_1, \dots, \eta_N$ are linear forms of the basis elements $\lambda_1, \dots, \lambda_n$ of $\mathfrak{t}^*$. Let $\mathrm{coeff}(\eta_i, j)$ denote the coefficient of λ_j in η_i (for $1 \le i \le N$ and $1 \le j \le n$) and introduce the notation

$$\deg(\eta_1, \dots, \eta_N; m) = \#\{i : \mathrm{coeff}(\eta_i, m) \neq 0\}.$$

Let $\Sigma \subset W$ be a T–invariant subvariety. It is clear from the formula (24) that for any $1 \le m \le n$, the λ_m–degree of $\mathrm{eP}[\Sigma, W]$ satisfies

$$\deg_{\lambda_m} \mathrm{eP}[\Sigma, W] \le \deg(\eta_1, \dots, \eta_N; m). \tag{25}$$

Example 5.5 Let W be $\mathbb{C}^4$ endowed with a $T=(\mathbb{C}^*)^3$–action, whose weights η_1,η_2,η_3 and η_4 span $\mathfrak{t}^*$, and satisfy $\eta_1+\eta_3=\eta_2+\eta_4$. Choose $p=(1,1,1,1)\in W$; then the affine toric variety

$$\overline{T\cdot p}=\{(y_1,y_2,y_3,y_4)\in\mathbb{C}^4 : y_1y_3=y_2y_4\}$$

is a hypersurface and its equivariant dual is given by the weight of the equation

$$\mathrm{eP}[\overline{T\cdot p},W]=\eta_1+\eta_3=\eta_2+\eta_4.$$

Another way to see this is to fix the monomial order $>$ induced by $y_1>y_2>y_3>y_4$; then the ideal $I=(y_1y_3-y_2y_4)$ has initial ideal $\mathrm{in}_I=(y_1y_3)$ whose spectrum is the union of the hyperplanes $\{y_1=0\}$ and $\{y_3=0\}$ with duals η_1, η_3 respectively.

Remark 5.6 An alternative and slightly more general topological definition of the equivariant dual is the following; see the notes of Fulton [18], Kazarian [22] and Edidin and Graham [13] for details. For a Lie group G let $EG\to BG$ be a right principal G–bundle with EG contractible. Such a bundle is universal in the topological setting: if $E\to B$ is any principal G–bundle, then there is a map $B\to BG$, unique up to homotopy, such that E is isomorphic to the pull-back of EG. If X is a smooth algebraic G–variety then the topological definition of the G–equivariant cohomology of X is

$$H_G^*(X)=H^*(EG\times_G X).$$

If Y is a G–invariant subvariety then Y represents a G–equivariant cohomology class in the equivariant cohomology of X, namely the ordinary Poincaré dual of $EG\times_G Y$ in $EG\times_G X$. This is the equivariant dual of Y in X:

$$\mathrm{eP}[Y,X]=\mathrm{PD}(EG\times_G Y, EG\times_G X).$$

5.3 The Rossmann formula

Let Z be a complex manifold with a holomorphic T–action, and let $M\subset Z$ be a T–invariant analytic subvariety with an isolated fixed point $p\in M^T$. Then one can find local analytic coordinates near p, in which the action is linear and diagonal. Using these coordinates, one can identify a neighbourhood of the origin in T_pZ with a neighbourhood of p in Z. We denote by $\widehat{\mathrm{T}}_pM$ the part of T_pZ which corresponds to M under this identification; informally, we will call $\widehat{\mathrm{T}}_pM$ the T–invariant *tangent cone* of M at p. This tangent cone is not quite canonical: it depends on the choice of coordinates; the equivariant dual of $\Sigma=\widehat{\mathrm{T}}_pM$ in $W=\mathrm{T}_pZ$, however, does not. Rossmann named this the *equivariant multiplicity of M in Z at p*:

$$\mathrm{emult}_p[M,Z]\overset{\mathrm{def}}{=}\mathrm{eP}[\widehat{\mathrm{T}}_pM,\mathrm{T}_pZ]. \tag{26}$$

Remark 5.7 In the algebraic framework one might need to pass to the *tangent scheme* of M at p (see Fulton [17]). This is canonically defined, but we will not use this notion.

The analogue of the Atiyah–Bott formula for singular subvarieties of smooth ambient manifolds is the following statement.

Proposition 5.8 (Rossmann's localisation formula [34]) *Let $\mu \in H_T^*(Z)$ be an equivariant class represented by a holomorphic equivariant map $\mathfrak{t} \to \Omega^\bullet(Z)$. Then*

$$\int_M \mu = \sum_{p \in M^T} \frac{\mathrm{emult}_p[M,Z]}{\mathrm{Euler}^T(\mathrm{T}_p Z)} \cdot \mu^{[0]}(p), \tag{27}$$

where $\mu^{[0]}(p)$ is the differential-form-degree-0 component of μ evaluated at p.

5.4 Equivariant localisation on $\widetilde{CX}_p^{[k+1]}$ for $k \le n$.

In this subsection we start to develop a two step equivariant localisation method on $\widetilde{CX}_p^{[k+1]}$ using the Rossmann formula. As the partial resolution $\widetilde{CX}_p^{[k+1]}$ described in Section 4.3 is defined only for $k \le n$ we impose this condition in this section.

Recall from the introduction to Section 5 that we fix a holomorphic coordinate system on X near p and using this we identify the fibre $J_k X_p$ with $J_k(1,n)$ and $\mathcal{D}_{X,p}^k$ with $J_k(n,1)^* = \mathrm{Sym}^{\le k}\mathbb{C}^n$. With these identifications the partial resolution map

$$\rho\colon \widetilde{CX}_p^{[k+1]} = \mathrm{GL}(n) \times_{P_{k,n}} \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}} \to \overline{\mathrm{GL}(n) \cdot \mathfrak{p}_{k,n}} = \widehat{CX}_p^{[k+1]}$$

fits into the following diagram:

$$\begin{array}{ccl} \widetilde{CX}_p^{[k+1]} & \xrightarrow{\rho} & \widehat{CX}_p^{[k+1]} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) \\ \Big\downarrow \mu & & \\ \mathrm{Flag}_k(\mathbb{C}^n) & & \end{array}$$

The fibres of μ are isomorphic to $\overline{P_{k,n} \cdot \mathfrak{p}_{k,n}} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ and μ is $\mathrm{GL}(n)$–equivariant.

Let $e_1, \dots, e_n \in \mathbb{C}^n$ be elements of an eigenbasis for the $T \subset \mathrm{GL}(n)$–action with weights $\lambda_1, \dots, \lambda_n \in \mathfrak{t}^*$ and let

$$f = (\langle e_1 \rangle \subset \langle e_1, e_2 \rangle \subset \cdots \subset \langle e_1, \dots, e_k \rangle \subset \mathbb{C}^n)$$

denote the standard flag in $\mathbb{C}^n$ fixed by the parabolic subgroup $P_{k,n} \subset \mathrm{GL}(n)$. Since the torus action on $\widetilde{CX}_p^{[k+1]}$ is obtained by the restriction of a $\mathrm{GL}(n)$–action to its subgroup of diagonal matrices T_n, the Weyl group of permutation matrices S_n acts transitively on the fixed-point set $\mathrm{Flag}_k(\mathbb{C}^n)^{T_n}$ taking the standard flag $\boldsymbol{f}$ to $\sigma(\boldsymbol{f})$ and for any closed equivariant form $\alpha \in \Omega_T^*(\widetilde{CX}_p^{[k+1]})$ Theorem 5.1 gives us

$$\int_{\widetilde{CX}_p^{[k+1]}} \alpha = \sum_{\sigma \in \mathcal{S}_n/\mathcal{S}_{n-k}} \frac{\alpha_{\sigma(\boldsymbol{f})}}{\prod_{1 \le m \le k} \prod_{i=m+1}^n (\lambda_{\sigma \cdot i} - \lambda_{\sigma \cdot m})}, \tag{28}$$

where:

- σ runs over the ordered k–element subsets of $\{1, \dots, n\}$ labelling the fixed flags $\sigma(\boldsymbol{f}) = (\langle e_{\sigma(1)} \rangle \subset \cdots \subset \langle e_{\sigma(1)}, \dots, e_{\sigma(k)} \rangle \subset \mathbb{C}^n)$ in $\mathbb{C}^n$.
- $\prod_{1 \le m \le k} \prod_{i=m+1}^n (\lambda_{\sigma(i)} - \lambda_{\sigma(m)})$ is the equivariant Euler class of the tangent space of $\mathrm{Flag}_k(\mathbb{C}^n)$ at $\sigma(\boldsymbol{f})$. Note that $\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(k)} \in S^\bullet \mathfrak{t}^*$ can be identified with the Chern roots of the tautological rank-k bundle $\mathcal{E}$ at $\sigma(\boldsymbol{f})$.
- If $\widetilde{CX}_{\sigma(\boldsymbol{f})}^{[k+1]} = \mu^{-1}(\sigma(\boldsymbol{f}))$ denotes the fibre then
$$\alpha_{\sigma(\boldsymbol{f})} = \left(\int_{\widetilde{CX}_{\sigma(\boldsymbol{f})}^{[k+1]}} \alpha \right)^{[0]} (\sigma(\boldsymbol{f})) \in S^\bullet \mathfrak{t}^* \otimes H^*(X)$$
is the differential-form-degree-0 part with coefficients in $H^*(X)$ evaluated at $\sigma(\boldsymbol{f})$ and $\alpha_{\sigma(\boldsymbol{f})} = \sigma \cdot \alpha_{\boldsymbol{f}}$ with respect to the natural Weyl group action on $S^\bullet \mathfrak{t}^*$.

In particular, when $P = P(c_1, \dots, c_{r(k+1)})$ is a polynomial in the Chern classes $c_i = c_i(F^{[k+1]})$ of the tautological rank-$r(k+1)$ bundle on the curvilinear Hilbert scheme then, according to Section 3.2, P is represented by a closed form $\alpha = \alpha(\theta_1, \dots, \theta_r, \eta_1, \dots, \eta_k)$ which is a bisymmetric polynomial in the Chern roots θ_i of the pull-back of F over $\widetilde{CX}_p^{[k+1]} \subset \mathrm{Flag}_k(\mathbb{C}^n)$ and the Chern roots η_j of the tautological rank-k bundle $\mathcal{E}$. Then $\alpha_{\boldsymbol{f}}$ is a polynomial in two sets of variables: the basic weights $\lambda = (\lambda_1, \dots, \lambda_n)$ of T on $\mathbb{C}^n$ and $\theta = (\theta_1, \dots, \theta_r)$. More precisely, the Chern roots of the tautological rank-k bundle $\mathcal{E}$ over $\alpha_{\boldsymbol{f}}$ correspond to the weights $\lambda_1, \dots, \lambda_k$ and therefore

$$\alpha_{\boldsymbol{f}} = \alpha_{\boldsymbol{f}}(\theta_1, \dots, \theta_r, \lambda_1, \dots, \lambda_k) \in S^\bullet \mathfrak{t}^* \otimes H^*(X)$$

is a bisymmetric polynomial of these $r + k$ variables. Then

$$\alpha_{\sigma(\boldsymbol{f})} = \sigma \cdot \alpha_{\boldsymbol{f}} = \alpha_{\boldsymbol{f}}(\theta_1, \dots, \theta_r, \lambda_{\sigma(1)}, \dots, \lambda_{\sigma(k)}) \in S^\bullet \mathfrak{t}^* \otimes H^*(X) \tag{29}$$

is the σ–shift of the polynomial $\alpha_{\boldsymbol{f}}$ corresponding to the distinguished fixed flag $\boldsymbol{f}$.

5.5 Transforming the localisation formula into iterated residue

In this section we transform the right-hand side of (28) into an iterated residue. This step turns out to be crucial in handling the combinatorial complexity of the Atiyah–Bott localisation formula and captures the symmetry of the fixed-point data in an efficient way which enables us to prove the vanishing of the contribution of all but one of the fixed points.

To describe this formula, we will need the notion of an *iterated residue* (see Szenes [35]) at infinity. Let $\omega_1,\dots,\omega_N$ be affine linear forms on $\mathbb{C}^k$; denoting the coordinates by $z_1,\dots,z_k$, this means that we can write $\omega_i = a_i^0 + a_i^1 z_1 + \cdots + a_i^k z_k$. We will use the shorthand $h(\boldsymbol{z})$ for a function $h(z_1,\dots,z_k)$, and $d\boldsymbol{z}$ for the holomorphic n–form $dz_1\wedge\cdots\wedge dz_k$. Now, let $h(\boldsymbol{z})$ be an entire function, and define the *iterated residue at infinity* as follows:

$$\operatorname*{Res}_{z_1=\infty}\operatorname*{Res}_{z_2=\infty}\cdots\operatorname*{Res}_{z_k=\infty}\frac{h(\boldsymbol{z})\,d\boldsymbol{z}}{\prod_{i=1}^N\omega_i} \overset{\text{def}}{=} \left(\frac{1}{2\pi i}\right)^k\int_{|z_1|=R_1}\cdots\int_{|z_k|=R_k}\frac{h(\boldsymbol{z})\,d\boldsymbol{z}}{\prod_{i=1}^N\omega_i}, \tag{30}$$

where $1\ll R_1\ll\cdots\ll R_k$. The torus $\{|z_m|=R_m : m=1,\dots,k\}$ is oriented in such a way that $\operatorname{Res}_{z_1=\infty}\cdots\operatorname{Res}_{z_k=\infty} d\boldsymbol{z}/(z_1\cdots z_k)=(-1)^k$. We will also use the simplified notation $\operatorname{Res}_{\boldsymbol{z}=\infty}\overset{\text{def}}{=}\operatorname{Res}_{z_1=\infty}\operatorname{Res}_{z_2=\infty}\cdots\operatorname{Res}_{z_k=\infty}$.

In practice, one way to compute the iterated residue (30) is the following algorithm: for each i, use the expansion

$$\frac{1}{\omega_i}=\sum_{j=0}^{\infty}(-1)^j\frac{(a_i^0+a_i^1z_1+\cdots+a_i^{q(i)-1}z_{q(i)-1})^j}{(a_i^{q(i)}z_{q(i)})^{j+1}}, \tag{31}$$

where $q(i)$ is the largest value of m for which $a_i^m\neq 0$, then multiply the product of these expressions with $(-1)^k h(z_1,\dots,z_k)$, and then take the coefficient of $z_1^{-1}\cdots z_k^{-1}$ in the resulting Laurent series.

Proposition 5.9 [7, Proposition 5.4] *For any homogeneous polynomial $Q(\boldsymbol{z})$ on $\mathbb{C}^k$ we have*

$$\sum_{\sigma\in\mathcal{S}_n/\mathcal{S}_{n-k}}\frac{Q(\lambda_{\sigma(1)},\dots,\lambda_{\sigma(k)})}{\prod_{1\le m\le k}\prod_{i=m+1}^n(\lambda_{\sigma\cdot i}-\lambda_{\sigma\cdot m})}=\operatorname*{Res}_{\boldsymbol{z}=\infty}\frac{\prod_{1\le m<l\le k}(z_m-z_l)Q(\boldsymbol{z})d\boldsymbol{z}}{\prod_{l=1}^k\prod_{i=1}^n(\lambda_i-z_l)}. \tag{32}$$

Remark 5.10 Changing the order of the variables in iterated residues, usually, changes the result. In this case, however, because all the poles are normal crossing, formula (32) remains true no matter in what order we take the iterated residues.

Proposition 5.9 together with (28) and (29) give us the following proposition.

Proposition 5.11 *Let $k \le n$ and $\alpha = \alpha(\theta_1, \dots, \theta_r, \eta_1, \dots, \eta_k)$ be a bisymmetric polynomial in the Chern roots θ_i of the pull-back of F over $\widetilde{CX}_p^{[k+1]} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ and the Chern roots η_j of the tautological rank-k bundle $\mathcal{E}$. Then*

$$\int_{\widetilde{CX}_p^{[k+1]}} \alpha = \operatorname*{Res}_{z=\infty} \frac{\prod_{1\le m<l\le k}(z_m - z_l)\,\alpha_f(\theta_1, \dots, \theta_r, z_1, \dots, z_k)\,dz}{\prod_{l=1}^k \prod_{i=1}^n (\lambda_i - z_l)}.$$

Next, we fix the T–eigenbasis $\{e_1, \dots, e_n\}$ of $\mathbb{C}^n$ and proceed a second localisation on the fibre

$$\widetilde{CX}_f^{[k+1]} = \mu^{-1}(f) \simeq \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$$

to compute $\alpha_f(\boldsymbol{\theta}, \boldsymbol{z})$. Recall from Corollary 4.1 that for $k \le n$

$$\mathfrak{p}_{k,n} = \Big(\mathrm{Span}(e_1) \subset \mathrm{Span}(e_1, e_2 + e_1^2) \subset \cdots \subset \mathrm{Span}\Big(e_1, e_2 + e_1^2, \dots, \sum_{\tau \in \mathcal{P}(k)} e_\tau\Big)\Big)$$

and $P_{k,n} \subset \mathrm{GL}(n)$ is the parabolic subgroup which preserves the flag

$$f = (\mathrm{Span}(e_1) \subset \mathrm{Span}(e_1, e_2) \subset \cdots \subset \mathrm{Span}(e_1, \dots, e_k) \subset T_pX).$$

Here for the partition $\tau = \{\tau_1 \le \tau_2 \le \cdots \le \tau_s\} \in \mathcal{P}(k)$ we use the notation

(1) k for $\mathrm{sum}(\tau) = \tau_1 + \cdots + \tau_s$,

(2) s for the length $|\tau|$,

(3) e_τ for $e_{\tau_1} e_{\tau_2} \cdots e_{\tau_s} \in \mathrm{Sym}^s(\mathbb{C}^n)$.

We define the subspaces

$$W_i = \mathrm{Span}_{\mathbb{C}}(e_\tau : \mathrm{sum}(\tau) \le i) \subset \mathrm{Sym}^{\le k}\mathbb{C}^n \quad \text{for } 1 \le i \le k.$$

These are invariant under the parabolic subgroup $P_{k,n} \subset \mathrm{GL}(n)$ which fixes the flag f. Note that the fibre $\widetilde{CX}_f^{[k+1]} = \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}$ sits in the submanifold

$$\mathrm{Flag}_k^*(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \{V_1 \subset \cdots \subset V_k \subset \mathrm{Sym}^{\le k}\mathbb{C}^n : \dim(V_i) = i,\ V_i \subset W_i\}$$

of $\mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$. Moreover, $\mathrm{Flag}_k^*(\mathrm{Sym}^{\le k}\mathbb{C}^n) \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ is a $P_{k,n}$–invariant subvariety.

As $\widetilde{CX}_f^{[k+1]}$ is invariant under the T–action on $\mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$, we can apply Rossmann's integration formula; see Proposition 5.8. More precisely, we apply the Rossmann formula for $M = X_f$, $Z = \mathrm{Flag}_k^*(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ and $\mu = \alpha_f$. The fixed points on

$$Z = \mathrm{Flag}_k^*(\mathrm{Sym}^{\le k}\mathbb{C}^n) \subset \bigoplus_{i=1}^k W_1 \wedge \cdots \wedge W_i$$

are parametrised by *admissible* sequences of partitions $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$. We call a sequence of partitions $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$ admissible if

(1) $\operatorname{sum}(\pi_l) \le l$ for $1 \le l \le k$, and

(2) $\pi_l \ne \pi_m$ for $1 \le l \ne m \le k$.

We will denote the set of admissible sequences of length k by $\mathbf{\Pi}_k$. The corresponding fixed point is then

$$\bigoplus_{i=1}^{k} e_{\pi_1} \wedge \dots \wedge e_{\pi_i} \in \bigoplus_{i=1}^{k} W_1 \wedge \dots \wedge W_i,$$

where $e_\pi = \prod_{j \in \pi} e \in \operatorname{Sym}^{|\pi|}\mathbb{C}^n$.

Then the Rossmann formula (27) and Proposition 5.11 give us the following proposition.

Proposition 5.12 *Let $k \le n$. Let $\alpha = \alpha(\theta_1, \dots, \theta_r, \eta_1, \dots, \eta_k)$ be a bisymmetric polynomial in the Chern roots θ_i of the pull-back of F over $\widetilde{CX}_p^{[k+1]} \subset \operatorname{Flag}_k(\operatorname{Sym}^{\le k}\mathbb{C}^n)$ and the Chern roots η_j of the tautological rank-k bundle $\mathcal{E}$. Then*

$$\int_{\widetilde{CX}_p^{[k+1]}} \alpha = \sum_{\boldsymbol{\pi} \in \mathbf{\Pi}_k \cap \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}} \operatorname*{Res}_{\boldsymbol{z}=\infty} \frac{Q_{\boldsymbol{\pi}}(\boldsymbol{z}) \prod_{m<l}(z_m - z_l)\alpha(\boldsymbol{\theta}, z_{\pi_1}, \dots, z_{\pi_k})}{\prod_{l=1}^{k} \prod_{\operatorname{sum}(\tau) \le l}^{\tau \ne \pi_1, \dots, \pi_l} (z_\tau - z_{\pi_l}) \prod_{l=1}^{k} \prod_{i=1}^{n} (\lambda_i - z_l)} \, d\boldsymbol{z}, \tag{33}$$

where

$$Q_{\boldsymbol{\pi}}(\boldsymbol{z}) = \operatorname{emult}_{\boldsymbol{\pi}}[X_{\boldsymbol{f}}, \operatorname{Flag}_k^*] \quad \textit{and} \quad z_\pi = \sum_{i \in \pi} z_i.$$

This formula reduces the computation of the tautological integrals $\int_{\widetilde{CX}_p^{[k+1]}} \alpha$ to determining the fixed-point set $\mathbf{\Pi}_k \cap \widetilde{CX}_{\boldsymbol{f}}^{[k+1]}$ and determining the multidegree $Q_{\boldsymbol{\pi}}(\boldsymbol{z}) = \operatorname{emult}_{\boldsymbol{\pi}}[X_{\boldsymbol{f}}, \operatorname{Flag}_k^*]$ of the tangent cone of $\widetilde{CX}_{\boldsymbol{f}}^{[k+1]}$ in $\operatorname{Flag}_k^*(\operatorname{Sym}^{\le k}\mathbb{C}^n)$.

6 The residue vanishing theorem

The first immediate problem arising with our formula (33) is our not having a complete description of the fixed-point set $\mathbf{\Pi}_k \cap \widetilde{CX}_{\boldsymbol{f}}^{[k+1]}$, and in fact deciding which torus fixed points on $\operatorname{Flag}_k^*(\operatorname{Sym}^{\le k}\mathbb{C}^n)$ sit in the orbit closure $\widetilde{CX}_{\boldsymbol{f}}^{[k+1]} = \overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}$ seems to be a hard question. The second problem we face is how to compute the multidegrees $Q_{\boldsymbol{\pi}}(\boldsymbol{z}) = \operatorname{emult}_{\boldsymbol{\pi}}[X_{\boldsymbol{f}}, \operatorname{Flag}_k^*]$ for those admissible sequences representing fixed points in $\overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}$. We postpone this second problem to the next section and here we focus on the first question which has a particularly nice and surprising answer. Namely, we do not need to know which fixed points sit in $\overline{P_{k,n} \cdot \mathfrak{p}_{k,n}}$ because our limited knowledge on the equations of the $P_{k,n}$–orbit is enough to show that all but one term on the right-hand side of (33) vanish. This key feature of the iterated residue has already appeared in Bérczi and Szenes [7] but here we need to prove a stronger version where the total degree of the rational forms are zero. We devote the rest of this section to the proof of the following theorem.

Theorem 6.1 (residue vanishing theorem) *Let $k+1 \leq n$ and let*

$$\alpha = \alpha(\theta_1, \dots, \theta_r, \eta_1, \dots, \eta_k)$$

be a bisymmetric polynomial in the Chern roots θ_i of the pull-back of F over $\widetilde{CX}_p^{[k+1]} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\leq k}\mathbb{C}^n)$ and the Chern roots η_j of the tautological rank-k bundle $\mathcal{E}$. Then:

(1) *All terms but the one corresponding to $\boldsymbol{\pi}_{\mathrm{dst}} = ([1],[2],\dots,[k])$ vanish in (33) leaving us with*

$$\int_{\widetilde{CX}_p^{[k+1]}} \alpha = \operatorname*{Res}_{z=\infty} \frac{Q_{([1],\dots,[k])}(\boldsymbol{z}) \prod_{m<l}(z_m - z_l)\alpha(\boldsymbol{\theta}, \boldsymbol{z})}{\prod_{\mathrm{sum}(\tau)\leq l\leq k}(z_\tau - z_l) \prod_{l=1}^k \prod_{i=1}^n (\lambda_i - z_l)}\, d\boldsymbol{z}. \tag{34}$$

(2) *If $|\tau| \geq 3$ then $Q_{([1],\dots,[k])}(\boldsymbol{z})$ is divisible by $z_\tau - z_l$ for all $l \geq \mathrm{sum}(\tau)$. Let*

$$Q_k(\boldsymbol{z}) = \frac{Q_{([1],\dots,[k])}(\boldsymbol{z})}{\prod_{|\tau|\geq 3,\, \mathrm{sum}(\tau)\leq l\leq k}(z_\tau - z_l)}$$

denote the quotient polynomial, and then we get the simplified formula

$$\int_{\widetilde{CX}_p^{[k+1]}} \alpha = \operatorname*{Res}_{z=\infty} \frac{Q_k(\boldsymbol{z}) \prod_{m<l}(z_m - z_l)\alpha(\theta, \boldsymbol{z})}{\prod_{m+r\leq l\leq k}(z_m + z_r - z_l) \prod_{l=1}^k \prod_{i=1}^n (\lambda_i - z_l)}\, d\boldsymbol{z}. \tag{35}$$

Remark 6.2 (1) Here we describe the geometric meaning of $Q_k(\boldsymbol{z})$ in (35); see also [7, Theorem 6.16]. Let $T_k \subset B_k \subset \mathrm{GL}(k)$ be the subgroups of invertible diagonal and upper triangular matrices, respectively; denote the diagonal weights of T_k by $z_1, \dots, z_k$. Consider the $\mathrm{GL}(k)$–module of 3–tensors $\mathrm{Hom}(\mathbb{C}^k, \mathrm{Sym}^2\mathbb{C}^k)$; identifying the weight-$(z_m+z_r-z_l)$ symbols q_l^{mr} and q_l^{rm}, we can write this space in terms of a basis as follows:

$$\mathrm{Hom}(\mathbb{C}^k, \mathrm{Sym}^2\mathbb{C}^k) = \bigoplus_{1\leq m,r,l\leq k} \mathbb{C}q_l^{mr}.$$

Consider the point $\epsilon = \sum_{m=1}^{k} \sum_{r=1}^{k-m} q_{mr}^{m+r}$ in the B_k–invariant subspace

$$W_k = \bigoplus_{1 \le m+r \le l \le k} \mathbb{C} q_l^{mr} \subset \mathrm{Hom}(\mathbb{C}^k, \mathrm{Sym}^2 \mathbb{C}^k).$$

Set the notation $\mathcal{O}_k$ for the orbit closure $\overline{B_k \epsilon} \subset W_k$; then $Q_k(\boldsymbol{z})$ is the T_k–equivariant Poincaré dual $Q_k(\boldsymbol{z}) = \mathrm{eP}[\mathcal{O}_k, W_k]_{T_k}$, which is a homogeneous polynomial of degree $\dim(W_k) - \dim(\mathcal{O}_k)$. For small k these polynomials are (see [7, Section 7])

$$Q_2 = Q_3 = 1, \quad Q_4 = 2z_1 + z_2 - z_4,$$
$$Q_5 = (2z_1 + z_2 - z_5)(2z_1^2 + 3z_1 z_2 - 2z_1 z_5 + 2z_2 z_3 - z_2 z_4 - z_2 z_5 - z_3 z_4 + z_4 z_5).$$

(2) To understand the significance of this vanishing theorem we note that while the fixed-point set $\mathbf{\Pi}_k$ on $\mathrm{Flag}_k^*(\mathrm{Sym}^{\le k} \mathbb{C}^n)$ is well understood, it is not clear which of these fixed points sit in $X_{\boldsymbol{f}}$. But we have enough information to prove that none of those fixed points in $X_{\boldsymbol{f}}$ contribute to the iterated residue except for the distinguished fixed point $\boldsymbol{\pi}_{\mathrm{dst}} = ([1], [2], \dots, [k])$. This simplification is dramatic: the number of terms in (34) grows exponentially with k, and of this sum now a single term survives.

(3) The residue vanishing theorem is valid under the condition $k+1 \le n$ which is slightly stronger than the condition $k \le n$ we worked with so far and which guaranteed the existence of $\widetilde{CX}_p^{[k+1]}$. We will remedy this condition in Section 6.2.

Remark 6.3 Remark 2.3 for singular varieties and ordinary compactly supported differential forms holds for compactly supported equivariant forms as follows. Let T be a complex torus and $f \colon M \to N$ be a smooth proper T–equivariant map between smooth quasiprojective varieties. Now assume that $X \subset M$ and $Y \subset N$ are possibly singular T–invariant closed subvarieties, such that f restricted to X is a birational map from X to Y. Next, let μ be an equivariantly closed differential form on N with values in polynomials on $\mathfrak{t}$. Then the integral of μ on the smooth part of Y is absolutely convergent; we denote this by $\int_Y \mu$. With this convention we again have

$$\int_X f^* \mu = \int_Y \mu, \tag{36}$$

and we can define integrals of equivariant forms on singular quasiprojective varieties simply by passing to any partial equivariant resolution or equivalently to integration over the smooth locus. In particular, applying this for the partial resolution $\rho \colon \widetilde{CX}_p^{[k+1]} \to \overline{CX}_p^{[k+1]}$ we get

$$\int_{\overline{CX}_p^{[k+1]}} \alpha = \int_{\widetilde{CX}_p^{[k+1]}} \rho^* \alpha$$

for any closed compactly supported differential form $\alpha \in \Omega^*(\overline{CX}_p^{[k+1]})$.

6.1 The vanishing of residues

In this subsection, following Bérczi and Szenes [7, Section 6.2], we describe the conditions under which iterated residues of the type appearing in the sum in (33) vanish and we prove Theorem 6.1.

We start with the 1–dimensional case, where the residue at infinity is defined by (30) with $d=1$. By bounding the integral representation along a contour $|z|=R$ with R large, one can easily prove the following lemma.

Lemma 6.4 *Let $p(z)$ and $q(z)$ be polynomials of one variable. Then*

$$\operatorname*{Res}_{z=\infty} \frac{p(z)\,dz}{q(z)} = 0 \quad \text{if } \deg(p(z))+1 < \deg(q).$$

Consider now the multidimensional situation. Let $p(\mathbf{z})$ and $q(\mathbf{z})$ be polynomials in the k variables $z_1,\dots,z_k$, and assume that $q(\mathbf{z})$ is the product of linear factors $q=\prod_{i=1}^N L_i$, as in (33). We continue to use the notation $d\mathbf{z}=dz_1\cdots dz_k$. We would like to formulate conditions under which the iterated residue

$$\operatorname*{Res}_{z_1=\infty}\operatorname*{Res}_{z_2=\infty}\cdots\operatorname*{Res}_{z_k=\infty}\frac{p(\mathbf{z})\,d\mathbf{z}}{q(\mathbf{z})} \tag{37}$$

vanishes. Introduce the following notation:

- When $p(\mathbf{z})$ is the product of linear forms and $1\le m\le k$ let $\deg(p(\mathbf{z});m)$ denote the number of terms in $p(\mathbf{z})$ with nonzero coefficients in front of z_m.
- For a nonzero linear form $L=a_0+a_1z_1+\dots+a_kz_k$, denote by $\operatorname{coeff}(L,z_l)=a_i$ the coefficient in front of z_i.
- Finally, for $1\le m\le k$, set
$$\operatorname{lead}(q(\mathbf{z});m)=\#\{i:\max\{l:\operatorname{coeff}(L_i,z_l)\neq 0\}=m\},$$
which is the number of those factors L_i in which the coefficient of z_m does not vanish, but the coefficients of $z_{m+1},\dots,z_k$ are 0.

We can group the N linear factors of $q(\mathbf{z})$ according to the nonvanishing coefficient with the largest index; in particular, for $1\le m\le k$ we have

$$\deg(q(\mathbf{z});m)\ge \operatorname{lead}(q(\mathbf{z});m) \quad \text{and} \quad \sum_{m=1}^k \operatorname{lead}(q(\mathbf{z});m)=N.$$

Proposition 6.5 [7, Proposition 6.3] *Let $p(z)$ and $q(z)$ be polynomials in the variables $z_1,\dots,z_k$, and assume that $q(z)$ is a product of linear factors: $q(z)=\prod_{i=1}^N L_i$; set $dz=dz_1\cdots dz_k$. Then*

$$\operatorname*{Res}_{z_1=\infty}\operatorname*{Res}_{z_2=\infty}\cdots\operatorname*{Res}_{z_k=\infty}\frac{p(z)\,dz}{q(z)}=0$$

if for some $l\le k$, the following holds:

$$\deg(p(z);l)+1<\deg(q(z);l)=\operatorname{lead}(q(z);l).$$

Note that the equality $\deg(q(z);l)=\operatorname{lead}(q(z);l)$ means that

(38) for each $i=1,\dots,N$ and $m>l$, $\operatorname{coeff}(L_i,z_l)\neq 0$ implies $\operatorname{coeff}(L_i,z_m)=0$.

We are ready to prove the residue vanishing theorem. Recall that our goal is to show that all the terms of the sum in (33) vanish except for the one corresponding to $\boldsymbol{\pi}_{\mathrm{dst}}=([1],\dots,[k])$. The plan is to apply Proposition 6.5 in stages to show that the iterated residue vanishes unless $z_i=[i]$ holds, starting with $i=k$ and going backwards.

Fix a sequence $\boldsymbol{\pi}=(\pi_1,\dots,\pi_k)\in\boldsymbol{\Pi}_k$, and consider the iterated residue corresponding to it on the right-hand side of (33). The expression under the residue is the product of two fractions:

$$\frac{p(z)}{q(z)}=\frac{p_1(z)}{q_1(z)}\cdot\frac{p_2(z)}{q_2(z)},$$

where

$$(39)\quad \frac{p_1(z)}{q_1(z)}=\frac{Q_{\boldsymbol{\pi}}(z)\prod_{m<l}(z_m-z_l)}{\prod_{l=1}^k\prod_{\mathrm{sum}(\tau)\le l}^{\tau\neq\pi_1,\dots,\pi_l}(z_\tau-z_{\pi_l})}\quad\text{and}\quad\frac{p_2(z)}{q_2(z)}=\frac{\alpha(\theta_1,\dots,\theta_r,z_{\pi_1},\dots,z_{\pi_k})}{\prod_{l=1}^k\prod_{i=1}^n(\lambda_i-z_l)}.$$

Note that $p(z)$ is a polynomial, while $q(z)$ is a product of linear forms. As a first step we show that if $\pi_k\neq[k]$, then already the first residue in the corresponding term on the right-hand side of (33) — the one with respect to z_k — vanishes. Indeed, if $\pi_k\neq[k]$, then $\deg(q_2(z);k)=n$, while z_k does not appear in $p_2(z)$. On the other hand, $\deg(q_1(z);k)=1$, because the only term which contains z_k is the one corresponding to $l=k$ and $\tau=[k]\neq\pi_k$. This also means that the only coordinate on $T_\pi\mathrm{Flag}_k^*$ which contains the z_k coordinate of the torus is $z_k-z_{\pi_k}$, and since $Q_{\boldsymbol{\pi}}(z)=\operatorname{emult}_{\boldsymbol{\pi}}[X_{\boldsymbol{f}},\mathrm{Flag}_k^*]$, (25) tells us that $\deg(Q_{\boldsymbol{\pi}}(z);k)\le 1$ holds. Collecting this data gives

$$(40)\qquad \deg(p_1(z)p_2(z);k)=k\quad\text{and}\quad\deg(q_1(z)q_2(z);k)=n+1,$$

and $k\le n-1$, so $\deg(p(z))\le\deg(q(z))-2$ holds and we can apply Lemma 6.4.

We can thus assume that $\pi_k=[k]$, and proceed to the next step and take the residue with respect to z_{k-1}. If $\pi_{k-1}\neq[k-1]$ then

$$\deg(q_2(z);k-1)=\operatorname{lead}(q_2(z);k-1)=n \quad\text{and}\quad \deg(p_2(z);k-1)=0. \tag{41}$$

In $q_1(z)$ the linear terms containing z_{k-1} are

$$z_{k-1}-z_k,\quad z_1+z_{k-1}-z_k \quad\text{and}\quad z_{k-1}-z_{\pi_{k-1}}. \tag{42}$$

The first term here cancels with the identical term in the Vandermonde in p_1. The second term divides $Q_{\boldsymbol{\pi}}$, according to the following proposition from [7] applied for $l=k-1$.

Proposition 6.6 [7, Proposition 6.4] *Let $l\geq 1$, and let $\boldsymbol{\pi}$ be an admissible sequence of partitions of the form $\boldsymbol{\pi}=(\pi_1,\dots,\pi_l,[l+1],\dots,[k])$, where $\pi_l\neq[l]$. Then for $m>l$, and every partition τ such that $l\in\tau$, $\operatorname{sum}(\tau)\leq m$, and $|\tau|>1$, we have*

$$(z_\tau-z_m)\mid Q_{\boldsymbol{\pi}}. \tag{43}$$

Therefore, after cancellation, all linear factors from $q_1(z)$ which have nonzero coefficients in front of both z_{k-1} and z_k vanish, and for the new fraction $p_1'(z)/q_1'(z)$,

$$\deg(q_1'(z);k-1)=\operatorname{lead}(q_1'(z);k-1)=1.$$

By (42) and (25), $\deg(Q_\pi;k-1)\leq 3$ and therefore after cancellation we have

$$\deg(p_1'(z);k-1)\leq k-2+2=k.$$

Using (41) we get

$$\deg(p_1'(z)p_2(z);k-1)\leq k$$

and

$$\deg(q_1'(z)q_2(z);k-1)=\operatorname{lead}(q_1'(z)q_2(z);k-1)=n+1,$$

so we can apply Proposition 6.5 with $l=k-1$ to deduce the vanishing of the residue with respect to $k-1$.

In general, assume that

$$\boldsymbol{\pi}=(\pi_1,\pi_2,\dots,\pi_l,[l+1],\dots,[k]) \quad\text{where } \pi_l\neq[l]$$

and proceed to the study of the residue with respect to z_l. For the second fraction we have again

$$\deg(q_2(z);l)=\operatorname{lead}(q_2(z);l)=n \quad\text{and}\quad \deg(p_2(z);l)=0. \tag{44}$$

The linear terms containing z_l in $q_1(\boldsymbol{z})$ are

$$z_l - z_k, \quad z_l - z_{k-1}, \dots, \quad z_l - z_{l+1}, \tag{45}$$

$$z_\tau - z_s \qquad \text{with } l \in \tau, \quad \tau \neq l, \quad l+1 \le s \le k \quad \text{and} \quad \mathrm{sum}(\tau) \le s, \tag{46}$$

$$z_l - z_{\pi_l}. \tag{47}$$

The weights in (45) cancel out with the identical terms of the Vandermonde in $p_1(\boldsymbol{z})$, and by Proposition 6.6, $Q_{\boldsymbol{\pi}}(\boldsymbol{z})$ is divisible by the weights in (46). Hence all linear factors with nonzero coefficient in front of z_l and at least one of $z_{l+1}, \dots, z_k$ vanish from $q_1(z)$. Let again $p_1'(\boldsymbol{z})/q_1'(\boldsymbol{z})$ denote the new fraction arising from $p_1(\boldsymbol{z})/q_1(\boldsymbol{z})$ after these cancellations. Then in $q_1'(\boldsymbol{z})$ only the term (47) contains z_l, and

$$\deg(q_1'(\boldsymbol{z}), l) = \mathrm{lead}(q_1'(\boldsymbol{z}), l) = 1. \tag{48}$$

In $p_1'(\boldsymbol{z})$ the linear terms which are left from the Vandermonde after cancellation and contain z_l are $z_{l-1} - z_l, \dots, z_1 - z_l$. The reduced $Q_\pi'(\boldsymbol{z})$ which we get after dividing by the terms in (46) is then a polynomial of the remaining weights, and the only remaining weights which contain z_l are

$$z_l - z_{\pi_l} \qquad \text{and} \qquad z_l - z_k, \quad z_l - z_{k-1}, \dots, \quad z_l - z_{l+1}.$$

This is because Q_π is the sum of monomials of the form $\Pi_{\omega \in I}\omega$ where

$$I \subset \{z_\tau - z_{\pi_l} : \mathrm{sum}(\tau) \le l, \, \tau \neq \pi_1, \dots, \pi_l, \, 1 \le l \le k\}$$

is a subset of the weights in $q_1(\boldsymbol{z})$ and therefore Q_π does not contain repeated weights. Then (25) tells us that $\deg(Q_\pi(\boldsymbol{z}); l) \le k - l + 1$. Therefore

$$\deg(p_1'(\boldsymbol{z}); l) \le (l-1) + (k-l+1) = k. \tag{49}$$

Putting (48) and (49) together we get

$$\deg(p_1'(\boldsymbol{z})p_2(\boldsymbol{z}), l) \le k \quad \text{and} \quad \deg(q_1'(\boldsymbol{z})q_2(\boldsymbol{z}), l) = \mathrm{lead}(q_1'(\boldsymbol{z})q_2(\boldsymbol{z}), l) = n+1.$$

Since $k \le n-1$, by applying Proposition 6.5 we arrive at the vanishing of the residue, forcing π_l to be $[l]$. This proves (1) of Theorem 6.1

The second part of the residue vanishing theorem is proved in Section 6.5 of [7] where we show that for $|\tau| > 1$,

$$(z_\tau - z_l) \mid Q_{([1],\dots,[k])} \quad \text{if } ([1],[2],\dots,[l-1],\tau,[l+1],\dots,[k-1],[k]) \text{ is not complete.}$$

We call an admissible sequence of partitions $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k)$ *complete* if for every $l \in \{1, \dots, k\}$ and every nontrivial subpartition $\tau \subset \pi_l$, there is an $m \in \{1, \dots, k\}$ such that $\pi_m = \tau$. Clearly, a sequence $([1],[2],\dots,[l-1],\tau,[l+1],\dots,[k-1],[k])$ is complete if and only if $|\tau| = 2$.

6.2 Increasing the number of points

The residue vanishing theorem provides a closed iterated residue formula for tautological integrals on $\widetilde{CX}_p^{[k+1]}$ in the case when $k+1 \le n$, that is, the number of points does not exceed the dimension of X. In this section we show how one can drop this very restrictive condition.

Recall that after fixing local coordinates on X near p, the test-curve model in Section 3.1 establishes a $\mathrm{GL}(n)$–equivariant isomorphism of quasiprojective varieties

$$J_k^{\mathrm{reg}}(1,n)/J_k^{\mathrm{reg}}(1,1) \simeq CX_p^{[k+1]} \subset \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\,\mathbb{C}^n)$$

between the moduli of k–jets of regular germs and the curvilinear locus of the punctual Hilbert scheme sitting in the Grassmannian of k–dimensional subspaces in $\mathrm{Sym}^{\le k}\,\mathbb{C}^n$.

Assume that $k+1 > \dim(X) = n$. Fix a basis $\{e_1,\dots,e_{k+1}\}$ of $\mathbb{C}^{k+1}$ and let

$$\mathbb{C}_{[n]} = \mathrm{Span}(e_1,\dots,e_n) \hookrightarrow \mathbb{C}^{k+1} \quad \text{and} \quad \mathbb{C}_{[k+1-n]} = \mathrm{Span}(e_{n+1},\dots,e_{k+1}) \hookrightarrow \mathbb{C}^{k+1}$$

denote the subspaces spanned by the first n and last $k+1-n$ basis vectors respectively. These are T_{k+1}–equivariant embeddings under the diagonal action of the maximal torus $T_{k+1} \subset \mathrm{GL}(k+1)$.

We can write $J_k^{\mathrm{reg}}(1,n) \subset J_k^{\mathrm{reg}}(1,k+1)$ as a zero locus of a smooth section of a vector bundle over $J_k^{\mathrm{reg}}(1,k+1)$. Indeed, consider the projection

$$J_k^{\mathrm{reg}}(1,k+1) \to J_k(1,k+1-n) \simeq \mathrm{Hom}(\mathbb{C}^k, \mathbb{C}_{[k+1-n]})$$

which sends a regular k–jet $f\colon \mathbb{C} \to \mathbb{C}^{k+1}$ to the composition $\mathbb{C} \to \mathbb{C}^{k+1} \to \mathbb{C}^{k+1-n}/\mathbb{C}_{[n]} = \mathbb{C}_{[k+1-n]}$. This map is $T_{k+1} \times J_k^{\mathrm{reg}}(1,1)$–equivariant and therefore it defines a T_{k+1}–equivariant section π of the bundle

$$E = J_k^{\mathrm{reg}}(1,k+1) \times_{J_k^{\mathrm{reg}}(1,1)} J_k(1,k+1-n).$$

This is a bundle over the quasiprojective base space $J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)$ with fibres isomorphic to $J_k(1,k+1-n)$ and the zero locus of the section π is

$$\pi^{-1}(0) = J_k^{\mathrm{reg}}(1,n)/J_k^{\mathrm{reg}}(1,1) \subset J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1).$$

Lemma 5.4 then suggests that for any T_{k+1}–equivariantly closed form on the quasiprojective quotient μ on $J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)$ we have

$$\int_{J_k^{\mathrm{reg}}(1,n)/J_k^{\mathrm{reg}}(1,1)} \mu = \int_{J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)} \mu \cdot \mathrm{Euler}^{T_{k+1}}(E),$$

but our base space $J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)$ is quasiprojective and not compact so

Lemma 5.4 does not apply directly. Note, however, that E extends T_{k+1}–equivariantly over the closure

$$\overline{J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)} \subset \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^{k+1})$$

in the Grassmannian, namely E is the restriction of the bundle

$$\widetilde{E} = \mathrm{Hom}^{\mathrm{reg}}(\mathbb{C}^k, \mathrm{Sym}^{\le k}\mathbb{C}^{k+1}) \times_{\mathrm{GL}(k)} \mathrm{Hom}(\mathbb{C}^k, \mathbb{C}_{[k+1-n]})$$

over the Grassmannian $\mathrm{Hom}^{\mathrm{reg}}(\mathbb{C}^k, \mathrm{Sym}^{\le k}\mathbb{C}^{k+1})/\mathrm{GL}(k) = \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^{k+1})$. That is, we have a T_{k+1}–equivariant embedding

$$\begin{array}{ccc} E & \hookrightarrow & \widetilde{E} \\ \downarrow{\scriptstyle \pi} & & \downarrow \\ J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1) & \xhookrightarrow{\phi^{\mathrm{Grass}}} & \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^{k+1}) \end{array}$$

of E into $\widetilde{E}$ and Lemma 5.4 gives us the following:

Proposition 6.7 *Let $\overline{E} = (\phi^{\mathrm{Grass}})^*\widetilde{E}$ denote the restriction of $\widetilde{E}$ to the closure $\overline{J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)}$. Then for any T_{k+1}–equivariantly closed form μ on this closure we have*

$$\int_{\overline{J_k^{\mathrm{reg}}(1,n)/J_k^{\mathrm{reg}}(1,1)}} \mu = \int_{\overline{J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1)}} \mu \cdot \mathrm{Euler}^{T_{k+1}}(\overline{E}).$$

We are ready to prove the iterated residue formula on the domain $n \le k+1$.

Theorem 6.8 (extended residue vanishing theorem) *Formula (35) remains valid for any $2 \le n < k+1$.*

Proof The embedding $\phi^{\mathrm{Flag}}\colon J_k^{\mathrm{reg}}(1,k+1)/J_k^{\mathrm{reg}}(1,1) \hookrightarrow \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^{k+1})$ is T_{k+1}–equivariant, and over the flag $\boldsymbol{f}_\sigma$ the weight of $f_j^{[i]}$ is $\lambda_j - \lambda_{\sigma(i)}$. In the iterated residue formula of Proposition 5.11 we substitute $\lambda_{\sigma(i)}$ for z_i over $\boldsymbol{f}_\sigma$ and therefore $\lambda_j - z_i$ for this weight and therefore the T_{k+1}–equivariant Euler class transforms into

$$\mathrm{Euler}_{\boldsymbol{z}}^{T_{k+1}}(\overline{E}) = \prod_{i=1}^{k} \prod_{j=n+1}^{k} (\lambda_j - z_i)$$

over the flag $\boldsymbol{f}_\sigma$ corresponding to an iterated pole $\boldsymbol{z} = (z_1, \dots z_k)$. If

$$\alpha = \alpha(\theta_1, \dots, \theta_r, \eta_1, \dots, \eta_k)$$

is a bisymmetric polynomial in the Chern roots θ_i of the pull-back of F over $\widetilde{CX}_p^{[k+1]} \subset \mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ and the Chern roots η_j of the tautological rank-k bundle $\mathcal{E}$, then the trivial extension of α to $\mathrm{Flag}_k(\mathrm{Sym}^{\le k}\mathbb{C}^{k+1})$ is closed and therefore Remark 6.3, Proposition 6.7 and Theorem 6.1 tell us

$$\int_{\overline{CX}_p^{[n]}} \alpha = \operatorname*{Res}_{z=\infty} \frac{Q_k(z)\prod_{m<l}(z_m-z_l)\alpha(\theta,z)\,dz}{\prod_{m+r\le l\le k}(z_m+z_r-z_l)\prod_{l=1}^k\prod_{i=1}^k(\lambda_i-z_l)}\cdot\prod_{l=1}^k\prod_{i=n+1}^k(\lambda_i-z_l)$$
$$= \operatorname*{Res}_{z=\infty} \frac{Q_k(z)\prod_{m<l}(z_m-z_l)\alpha(\theta,z)}{\prod_{m+r\le l\le k}(z_m+z_r-z_l)\prod_{l=1}^k\prod_{i=1}^n(\lambda_i-z_l)}\,dz. \qquad \square$$

7 Proof of Theorem 1.2 and examples

Let $P=P(c_1,\dots,c_{r(k+1)})$ be a Chern polynomial of degree $\dim\overline{CX}^{[k+1]}$, which equals $n+(n-1)k$, where the $c_i=c_i(F^{[k+1]})$ are the Chern classes of the tautological rank-$r(k+1)$ bundle on the curvilinear Hilbert scheme. To evaluate the integral $\int_{\overline{CX}^{[k+1]}} P$ we can first integrate (push forward) along the fibres of $\pi\colon \overline{CX}^{[k+1]}\to X$ followed by integration over X. By fixing local holomorphic coordinates on X near p these fibres are canonically isomorphic to $\overline{CX}_p^{[k+1]}\subset \mathrm{Grass}_k(\mathrm{Sym}^{\le k}(\mathbb{C}^n)$ endowed with a natural $\mathrm{GL}(n)$–action induced by the standard $\mathrm{GL}(n)$–action on $\mathbb{C}^n$. We can use this action to perform torus equivariant localisation on $\overline{CX}_p^{[k+1]}$ to integrate along the fibres. According to Remark 6.3, $\int_{\overline{CX}_p^{[k+1]}} P=\int_{\widetilde{CX}_p^{[k+1]}}\rho^*P$ holds, where $\rho\colon \widetilde{CX}_p^{[k+1]}\to\overline{CX}_p^{[k+1]}$ is the partial resolution constructed in Section 4.3. Applying Theorem 6.1 and its extension Theorem 6.8 and the expression in (14) for the Chern classes of $F^{[k+1]}$ we get

$$\int_{\overline{CX}_p^{[k+1]}} P = \operatorname*{Res}_{z=\infty}\frac{Q_k(z)\prod_{m<l}(z_m-z_l)P(c_l(z+\theta,\theta))}{\prod_{m+r\le l\le k}(z_m+z_r-z_l)\prod_{l=1}^k\prod_{i=1}^n(\lambda_i-z_l)}\,dz, \tag{50}$$

where $\theta_1,\dots,\theta_r$ are the Chern roots of F and $c_l(z+\theta,\theta)$ denotes the l^{th} symmetric polynomial in the formal Chern roots $\{z_i+\theta_j,\theta_j : 1\le i\le k, 1\le j\le r\}$.

According to Corollary 3.14 $\overline{CX}_p^{[k+1]}=J_k^{\mathrm{reg}}X_p/\mathrm{Diff}_k(1)$, which sits in $\mathrm{Grass}_k(\mathcal{D}^k_{X,p})$. By choosing local coordinates on X near p we identify $\mathcal{D}^k_{X,p}$ with $\mathrm{Sym}^{\le k}\mathbb{C}^n$ and the weights $\lambda_1,\dots,\lambda_n$ of the $\mathrm{GL}(n)$–action on $\mathbb{C}^n$ intuitively correspond to the Chern roots of $\mathcal{D}^{\le 1}_{X,p}/\mathcal{D}^{\le 0}_{X,p}=T_pX$. To finish the proof of Theorem 1.2 we simply substitute the λ_i with the Chern roots of T_X. Indeed, this is what we have to do, but this intuitive step needs further explanation.

The crucial observation is that $\mathrm{GL}(n)$ is a (strong) deformation retract of Diff_k via the homotopy

$$\mathrm{Diff}_k\times[0,1]\to\mathrm{Diff}_k$$

which sends (ϕ, t) to the ϕ_t whose linear part is identical to the linear part of ϕ but whose quadratic and higher order terms are those of ϕ multiplied by t. This homotopy contracts the quadratic and higher order terms of ϕ to zero. This induces a retraction of the classifying spaces

$$\tau\colon\ B\mathrm{Diff}_k \to B\mathrm{GL}(n)$$

which is a homotopy equivalence.

Given a Diff_k–module V the embedding $\mathrm{GL}(n) \hookrightarrow \mathrm{Diff}_k$ also defines a $\mathrm{GL}(n)$–module structure on V and the corresponding universal bundles

$$E_{\mathrm{Diff}}\, V = E\mathrm{Diff}_k \times_{\mathrm{Diff}_k} V \quad \text{and} \quad E_{\mathrm{GL}(n)}\, V = E\mathrm{GL}(n) \times_{\mathrm{GL}(n)} V$$

are homotopy equivalent. In particular,

$$E_{\mathrm{Diff}}\, \overline{CX}_p^{[k+1]} = E\mathrm{Diff}_k \times_{\mathrm{Diff}_k} \overline{CX}_p^{[k+1]}$$

and

$$E_{\mathrm{GL}}\, \overline{CX}_p^{[k+1]} = E\mathrm{GL}(n) \times_{\mathrm{GL}(n)} \overline{CX}_p^{[k+1]}$$

are homotopy equivalent and therefore their pull-backs along the classifying map $\xi\colon X \to B\mathrm{Diff}_k$,

$$\overline{CX}^{[k+1]} = \xi^*\, E_{\mathrm{Diff}}\, \overline{CX}_p^{[k+1]} \quad \text{and} \quad \overline{CX}_{\mathrm{GL}}^{[k+1]} = (\tau \circ \xi)^*\, E_{\mathrm{GL}}\, \overline{CX}_p^{[k+1]},$$

are also homotopy equivalent. They sit in the corresponding Grassmannian bundles:

$$\begin{aligned} \overline{CX}^{[k+1]} &\subset \xi^*\, E_{\mathrm{Diff}}\, \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \mathrm{Grass}_k(\mathcal{D}_X^k),\\ \overline{CX}_{\mathrm{GL}}^{[k+1]} &\subset (\tau \circ \xi)^*\, E_{\mathrm{GL}}\, \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n) = \mathrm{Grass}_k(\mathrm{Sym}^{\le k} T_X). \end{aligned}$$

If α is a polynomial in the Chern classes of the tautological rank-k bundle $\mathcal{E}$ over $E_{\mathrm{GL}}\, \mathrm{Grass}_k(\mathrm{Sym}^{\le k}\mathbb{C}^n)$ then

$$\int_{\overline{CX}_{\mathrm{GL}}^{[k+1]}} (\tau \circ \xi)^* \alpha = \int_{\overline{CX}^{[k+1]}} \xi^* \alpha$$

holds, and therefore we can replace integration over $\overline{CX}^{[k+1]}$ with integration over $\overline{CX}_{\mathrm{GL}}^{[k+1]}$. The commutative diagram

$$\begin{array}{ccc} \overline{CX}_{\mathrm{GL}}^{[k+1]} & \longrightarrow & E_{\mathrm{GL}}\, \overline{CX}_p^{[k+1]} \\ \downarrow & & \downarrow \\ X & \xrightarrow{\ \tau\circ\xi\ } & B\mathrm{GL}(n) \end{array}$$

induces a diagram of cohomology maps

$$\begin{tikzcd}
H^*(\overline{CX}^{[k+1]}_{\mathrm{GL}}) \arrow[d, "\int_p"] & H^*(E_{\mathrm{GL}}\,\overline{CX}^{[k+1]}_p) \arrow[l] \arrow[d, "\mathrm{Res}"] \\
H^*(X) & H^*(B\mathrm{GL}(n)) \arrow[l, "\mathrm{Sub}"']
\end{tikzcd}$$

where:

- Res is the residue operator sending a polynomial P in the Chern classes of $\mathcal{E}$ to
$$\operatorname*{Res}_{\mathbf{z}=\infty} \frac{Q_k(\mathbf{z})\prod_{m<l}(z_m-z_l)P(c_l(\mathbf{z}))}{\prod_{m+r\le l\le k}(z_m+z_r-z_l)\prod_{l=1}^k\prod_{i=1}^n(\lambda_i-z_l)}\,d\mathbf{z}.$$
- Sub is the substitution of the Chern roots of X into the weights $\lambda_1,\dots,\lambda_n$.
- $\int_p$ is integration along the fibre.

Commutativity tells us that integration along the fibre $\overline{CX}^{[k+1]}_p$ of a class pulled back from the universal bundle $\mathcal{E}$ over $E_{\mathrm{GL}}\,\overline{CX}^{[k+1]}_p$ is given by applying the residue operation followed by the substitution of the Chern roots of X into the weights λ_i of the torus action.

To get the final version of the iterated residue formula we replace the variables z_i by $-z_i$ for $i=1,\dots,k$. This changes the sign of the iterate residue (50) with $(-1)^k$ as this substitution corresponds to changing the orientation of the contour circles. Then the terms involving the λ_i in (50) can be rewritten as

$$\frac{1}{\prod_{i=1}^n(\lambda_i+z_j)}=\frac{1}{z_j^n c(1/z_j)}=\frac{s_X(1/z_j)}{z_j^n},$$

where $s_X(1/z_i)=1+s_1(X)/z_i+s_2(X)/z_i^2+\dots+s_n(X)/z_i^n$ is the total Segre class of X. Next observe that the denominator and the numerator of the fraction

$$\frac{\prod_{i<j}(z_i-z_j)Q_k(z)}{\prod_{i+j\le l\le k}(z_i+z_j-z_l)}$$

are homogeneous polynomials of the same degree; hence this substitution will leave this rational expression unchanged and only replaces z_i by $-z_i$ in $P(c_l(\mathbf{z}+\theta,\theta))$. So (50) can be rewritten as

$$\int_{\overline{CX}^{[k+1]}} P = \int_X \operatorname*{Res}_{\mathbf{z}=\infty} \frac{(-1)^k\prod_{1\le i<j\le k}(z_i-z_j)Q_k(\mathbf{z})P(c_l(\theta-\mathbf{z},\theta))\,d\mathbf{z}}{\prod_{i+j\le l\le k}(z_i+z_j-z_l)(z_1\cdots z_k)^n}\prod_{i=1}^k s_X\Big(\frac{1}{z_i}\Big),$$

and Theorem 1.2 is proved.

Remark 7.1 (1) Note that if we give the z_i and θ_j degree 1 then the total degree of the rational expression

$$\frac{(-1)^k \prod_{i<j}(z_i-z_j)Q_k(z)P(c_l(\theta-\boldsymbol{z},\theta))}{\prod_{i+j\le l\le k}(z_i+z_j-z_l)(z_1\cdots z_k)^n}$$

in the formula is $n-k$, so taking the iterated residue indeed gives us a bisymmetric homogeneous polynomial of degree n in the θ_j and s_i.

(2) The Chern class $c_l(\theta-\boldsymbol{z},\theta)$ is the coefficient of t^l in

$$c(F^{[k+1]})(t)=\prod_{j=1}^{r}(1+\theta_j t)\prod_{i=1}^{k}\prod_{j=1}^{r}(1-z_i t+\theta_j t),$$

that is, the i^{th} Chern class of the bundle with formal Chern roots θ_j, θ_j-z_i. For example,

$$c_1(\theta-\boldsymbol{z},\theta)=(k+1)\sum_{j=1}^{r}\theta_j-r\sum_{i=1}^{k}z_i,$$

and in general $c_l(\theta-\boldsymbol{z},\theta)$ is a degree-l polynomial of the form

$$c_l(\theta-\boldsymbol{z},\theta)=A_l c_l(\boldsymbol{z})+A_{l-1}c_{l-1}(\boldsymbol{z})+\cdots+A_0,$$

where $c_j(\boldsymbol{z})$ is the j^{th} elementary symmetric polynomial in $z_1,\dots,z_k$ and A_j is a degree-$(n-j)$ symmetric polynomial in $\theta_1,\dots,\theta_r$.

In certain special cases, however, we do not need this expansions of the Chern classes. We finish this paper with showing a particularly nice example of this, the Segre classes of tautological bundles over the curvilinear Hilbert schemes.

Example 7.2 (top Segre classes of tautological bundles) Top Segre classes

$$s_{\text{top}}(F^{[k+1]})=\int_{\overline{CX}^{[k+1]}} s(F^{[k+1]})$$

of tautological bundles have been long studied and they are of special interest because of their role in Donaldson–Thomas theory of counting sheaves on surfaces; see Marian, Oprea and Pandharipande [29] for details. Here $s(F^{[k+1]})=1/c(F^{[k+1]})$ is the total Segre class of $F^{[k+1]}$, that is,

$$s(F^{[k+1]})=s(\theta-\boldsymbol{z},\theta)=\prod_{j=1}^{r}\frac{1}{1+\theta_j}\cdot\prod_{i=1}^{k}\prod_{j=1}^{r}\frac{1}{1+\theta_j-z_i}=s_F\cdot(z_1\cdots z_k)^{-r}\prod_{i=1}^{k}\mathcal{S}\Big(\frac{1}{z_i}\Big),$$

where s_F is the total Segre class of F and

$$\mathcal{S}\Big(\frac{1}{z_i}\Big) = -\prod_{j=1}^{r}\Big(1 + \frac{1+\theta_j}{z_i} + \frac{(1+\theta_j)^2}{z_i^2} + \cdots + \frac{(1+\theta_j)^n}{z_i^n}\Big)$$

is a polynomial in $1/z_i$ with coefficients polynomials in the Chern classes of F, that is, $\mathcal{S}(x) \in \mathbb{C}[c_1(F), \dots, c_r(F)][x]$.

Substituting into Theorem 1.2 we arrive at the following expression:

$$s_{\mathrm{top}}(F^{[k+1]}) = \int_X \operatorname*{Res}_{\mathbf{z}=\infty} \frac{(-1)^k \prod_{1\le i<j\le k}(z_i - z_j) Q_k(\mathbf{z}) s_F \, d\mathbf{z}}{\prod_{i+j\le l\le k}(z_i + z_j - z_l)(z_1 \cdots z_k)^{r+n}} \prod_{i=1}^{k} \mathcal{S}\Big(\frac{1}{z_i}\Big) s_X\Big(\frac{1}{z_i}\Big).$$

References

[1] **M F Atiyah**, **R Bott**, *The moment map and equivariant cohomology*, Topology 23 (1984) 1–28 MR

[2] **G Bérczi**, *Thom polynomials of Morin singularities and the Green–Griffiths–Lang conjecture*, preprint (2010) arXiv

[3] **G Bérczi**, **B Doran**, **T Hawes**, **F Kirwan**, *Geometric invariant theory for graded unipotent groups and applications*, preprint (2016) arXiv

[4] **G Bérczi**, **B Doran**, **T Hawes**, **F Kirwan**, *Projective completions of graded unipotent quotients*, preprint (2016) arXiv

[5] **G Bérczi**, **F Kirwan**, *Graded unipotent groups and Grosshans theory*, preprint (2015) arXiv

[6] **G Bérczi**, **A Szenes**, *Tautological integrals on Hilbert schemes and counting singular hypersurfaces*, in preparation

[7] **G Bérczi**, **A Szenes**, *Thom polynomials of Morin singularities*, Ann. of Math. 175 (2012) 567–629 MR

[8] **N Berline**, **E Getzler**, **M Vergne**, *Heat kernels and Dirac operators*, corrected 1st edition, Grundl. Math. Wissen. 298, Springer (2004) MR

[9] **N Berline**, **M Vergne**, *Zéros d'un champ de vecteurs et classes caractéristiques équivariantes*, Duke Math. J. 50 (1983) 539–549 MR

[10] **R Bott**, **L W Tu**, *Differential forms in algebraic topology*, Graduate Texts in Mathematics 82, Springer (1982) MR

[11] **J-P Demailly**, *Algebraic criteria for Kobayashi hyperbolic projective varieties and jet differentials*, from "Algebraic geometry" (J Kollár, R Lazarsfeld, D R Morrison, editors), Proc. Sympos. Pure Math. 62, Amer. Math. Soc., Providence, RI (1997) 285–360 MR

[12] **M Duflo**, **M Vergne**, *Orbites coadjointes et cohomologie équivariante*, from "The orbit method in representation theory" (M Duflo, N V Pedersen, M Vergne, editors), Progr. Math. 82, Birkhäuser, Boston (1990) 11–60 MR

[13] **D Edidin**, **W Graham**, *Characteristic classes in the Chow ring*, J. Algebraic Geom. 6 (1997) 431–443 MR

[14] **D Eisenbud**, *Commutative algebra*, Graduate Texts in Mathematics 150, Springer (1995) MR

[15] **G Ellingsrud**, **L Göttsche**, **M Lehn**, *On the cobordism class of the Hilbert scheme of a surface*, J. Algebraic Geom. 10 (2001) 81–100 MR

[16] **L M Fehér**, **R Rimányi**, *Thom series of contact singularities*, Ann. of Math. 176 (2012) 1381–1426 MR

[17] **W Fulton**, *Intersection theory*, Ergeb. Math. Grenzgeb. 2, Springer (1984) MR

[18] **W Fulton**, *Equivariant cohomology in algebraic geometry*, Eilenberg lectures, Columbia University (2007) Available at `https://people.math.osu.edu/anderson.2804/eilenberg/`

[19] **T Gaffney**, *The Thom polynomial of* $\overline{\Sigma^{1111}}$, from "Singularities, 1" (P Orlik, editor), Proc. Sympos. Pure Math. 40, Amer. Math. Soc., Providence, RI (1983) 399–408 MR

[20] **L Göttsche**, *A conjectural generating function for numbers of curves on surfaces*, Comm. Math. Phys. 196 (1998) 523–533 MR

[21] **M Green**, **P Griffiths**, *Two applications of algebraic geometry to entire holomorphic mappings*, from "The Chern Symposium 1979" (H H Wu, S Kobayashi, I M Singer, A Weinstein, J Wolf, editors), Springer (1980) 41–74 MR

[22] **M E Kazarian**, *Characteristic classes of singularity theory*, from "The Arnold–Gelfand mathematical seminars" (V I Arnold, I M Gelfand, V S Retakh, M Smirnov, editors), Birkhäuser, Boston (1997) 325–340 MR

[23] **M E Kazarian**, *Multisingularities, cobordisms, and enumerative geometry*, Uspekhi Mat. Nauk 58 (2003) 29–88 MR In Russian; translated in Russian Math. Surveys 58 (2003) 665–724

[24] **S Kleiman**, **R Piene**, *Enumerating singular curves on surfaces*, from "Algebraic geometry: Hirzebruch 70" (P Pragacz, M Szurek, J Wiśniewski, editors), Contemp. Math. 241, Amer. Math. Soc., Providence, RI (1999) 209–238 MR

[25] **M Kool**, **V Shende**, **R P Thomas**, *A short proof of the Göttsche conjecture*, Geom. Topol. 15 (2011) 397–406 MR

[26] **M Lehn**, *Chern classes of tautological sheaves on Hilbert schemes of points on surfaces*, Invent. Math. 136 (1999) 157–207 MR

[27] **J Li**, *Zero dimensional Donaldson–Thomas invariants of threefolds*, Geom. Topol. 10 (2006) 2117–2171 MR

[28] **A-K Liu**, *Family blowup formula, admissible graphs and the enumeration of singular curves, I*, J. Differential Geom. 56 (2000) 381–579 MR

[29] **A Marian**, **D Oprea**, **R Pandharipande**, *Segre classes and Hilbert schemes of points*, preprint (2015) arXiv

[30] **V Mathai**, **D Quillen**, *Superconnections, Thom classes, and equivariant differential forms*, Topology 25 (1986) 85–110 MR

[31] **E Miller**, **B Sturmfels**, *Combinatorial commutative algebra*, Graduate Texts in Mathematics 227, Springer (2005) MR

[32] **H Nakajima**, *Heisenberg algebra and Hilbert schemes of points on projective surfaces*, Ann. of Math. 145 (1997) 379–388 MR

[33] **J Rennemo**, *Universal polynomials for tautological integrals on Hilbert schemes*, Geom. Topol. 21 (2017) 253–314 MR

[34] **W Rossmann**, *Equivariant multiplicities on complex varieties*, from “Orbites unipotentes et représentations, III”, Astérisque 173–174, Société Mathématique de France, Paris (1989) 313–330 MR

[35] **A Szenes**, *Iterated residues and multiple Bernoulli polynomials*, Internat. Math. Res. Notices (1998) 937–956 MR

[36] **Y-J Tzeng**, *A proof of the Göttsche–Yau–Zaslow formula*, J. Differential Geom. 90 (2012) 439–472 MR

[37] **M Vergne**, *Polynômes de Joseph et représentation de Springer*, Ann. Sci. École Norm. Sup. 23 (1990) 543–562 MR

Mathematical Institute, University of Oxford
Andrew Wiles Building, OX2 6GG Oxford, UK

Current address: *Department of Mathematics, ETH Zürich*
Raemistrasse 101, HG J 16.4, CH-8092 Zürich, Switzerland

gergely.berczi@math.ethz.ch

Proposed: Lothar Göttsche
Seconded: Frances Kirwan, Dan Abramovich

Received: 12 November 2015
Revised: 18 August 2016

Geometry & Topology 21 (2017) 2945–2988

Convexity of the extended K-energy and the large time behavior of the weak Calabi flow

ROBERT J BERMAN
TAMÁS DARVAS
CHINH H LU

Let (X, ω) be a compact connected Kähler manifold and denote by $(\mathcal{E}^p, d_p)$ the metric completion of the space of Kähler potentials $\mathcal{H}_\omega$ with respect to the L^p–type path length metric d_p. First, we show that the natural analytic extension of the (twisted) Mabuchi K-energy to $\mathcal{E}^p$ is a d_p–lsc functional that is convex along finite-energy geodesics. Second, following the program of J Streets, we use this to study the asymptotics of the weak (twisted) Calabi flow inside the CAT(0) metric space $(\mathcal{E}^2, d_2)$. This flow exists for all times and coincides with the usual smooth (twisted) Calabi flow whenever the latter exists. We show that the weak (twisted) Calabi flow either diverges with respect to the d_2–metric or it d_1–converges to some minimizer of the K-energy inside $\mathcal{E}^2$. This gives the first concrete result about the long-time convergence of this flow on general Kähler manifolds, partially confirming a conjecture of Donaldson. We investigate the possibility of constructing destabilizing geodesic rays asymptotic to diverging weak (twisted) Calabi trajectories, and give a result in the case when the twisting form is Kähler. Finally, when a cscK metric exists in $\mathcal{H}_\omega$, our results imply that the weak Calabi flow d_1–converges to such a metric.

53C55; 32W20, 32U05

1 Introduction

Given a compact connected Kähler manifold (X, ω), we denote by $\mathcal{H}$ the space of smooth Kähler metrics in the cohomology class $[\omega]$. As follows from the $\partial\bar\partial$–lemma of Hodge theory, up to a constant, this space is in one-to-one correspondence with the space of Kähler potentials

$$\mathcal{H}_\omega = \{u \in C^\infty(X) : \omega_u := \omega + i\partial\bar\partial u > 0\}.$$

As $\mathcal{H}_\omega$ is an open subset of $C^\infty(X)$, it is a Fréchet manifold and it is possible to endow it with different L^p–type Finsler metrics for $p \geq 1$, via

$$\|\xi\|_{p,u} := \left(V^{-1}\int_X |\xi|^p \omega_u^n\right)^{1/p}, \quad \xi \in T_u\mathcal{H}_\omega = C^\infty(X). \tag{1}$$

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2945

For $p = 2$ one recovers the Riemannian structure of Mabuchi which turns $\mathcal{H}$ into a Riemannian symmetric space of constant negative curvature (see Donaldson [38], Mabuchi [55] and Semmes [59]), but as will be explained below, the Finsler case $p = 1$ will also play a key role in the present paper.

One of the central questions of Kähler geometry, going back to Calabi, is to understand under what conditions $\mathcal{H}$ contains a constant scalar curvature Kähler (csc-K) metric. From a variational point of view this amounts to looking for critical points (minimizers) of Mabuchi's K-energy functional $\mathcal{K}\colon \mathcal{H}_\omega \to \mathbb{R}$ [38; 55], whose first variation is defined by the formula

$$\langle D\mathcal{K}(u), \delta u\rangle = V^{-1}\int_X \delta u(\bar{S} - S_{\omega_u})\omega_u^n,$$

where $V = \int_X \omega^n$ is the total volume and $\bar{S} = nV^{-1}\int_X \operatorname{Ric}\omega \wedge \omega^{n-1} = V^{-1}\int_X S_\omega \omega^n$ is the mean scalar curvature. According to a formula of Chen and Tian, the K-energy can be expressed explicitly in terms of the Kähler potential as

$$\mathcal{K}(u) := \operatorname{Ent}(\omega^n, \omega_u^n) + \bar{S}\,\mathrm{AM}(u) - n\,\mathrm{AM}_{\operatorname{Ric}\omega}(u), \tag{2}$$

where $\operatorname{Ent}(\omega^n, \omega_u^n) = V^{-1}\int_X \log(\omega_u^n/\omega^n)\omega_u^n$ is the entropy of the measure ω_u^n with respect to ω^n and $\mathrm{AM}, \mathrm{AM}_\gamma\colon \mathcal{H}_\omega \to \mathbb{R}$ are the Aubin–Mabuchi (also Aubin–Yau) energy and its "γ–contracted" version:

$$\mathrm{AM}(u) = \frac{1}{(n+1)V}\sum_{j=0}^{n}\int_X u\omega_u^j \wedge \omega^{n-j}, \quad \mathrm{AM}_\gamma(u) = \frac{1}{nV}\sum_{j=0}^{n-1}\int_X u\gamma \wedge \omega_u^j \wedge \omega^{n-1-j}.$$

As shown by Mabuchi, the K-energy is convex along geodesics in $\mathcal{H}_\omega$ when the geodesics are defined in terms of the corresponding L^2–Riemannian structure. However, a major technical stumbling block in this infinite-dimensional setting is that the Riemannian structure on $\mathcal{H}_\omega$ is not geodesically complete, and this is one of the reasons that we will be forced to work with various completions of $\mathcal{H}_\omega$, as discussed below.

In the finite-dimensional Riemannian setting, a time-honored approach to finding minimizers of convex functions is to follow their negative (downward) gradient flow. In the present infinite-dimensional Riemannian setting the negative gradient flow of the K-energy is precisely the Calabi flow $t \to c_t$:

$$\frac{d}{dt}c_t = S_{\omega_{c_t}} - \bar{S}.$$

Given arbitrary initial potential $c_0 \in \mathcal{H}_\omega$, short-time existence of the flow, assuming the initial potential is $C^{3,\alpha}$, is due to Chen and He [24], but long-time existence is still an open conjecture due to Calabi and Chen. In the case where $\dim X = 1$, long-time existence and convergence of the flow was first explored by Chruściel [28].

Fine [42] used finite-dimensional flows to approximate the Calabi flow. Under various restrictive conditions, convergence and existence theorems for the Calabi flow have been extensively studied. We refer the reader to Chen and He [25], Feng and Huang [40], He [48], Huang [49], Huang and Zheng [50], Li, Wang and Zheng [54], Székelyhidi [63] and Tosatti and Weinkove [65], to cite a few works from a very fast-growing literature.

The main motivation of our paper is the following conjecture of Donaldson on the long-time asymptotics and convergence of the Calabi flow, which, roughly stated, says:

Conjecture 1.1 [39] *Let $[0,\infty) \ni t \to c_t \in \mathcal{H}$ be a Calabi flow trajectory. Exactly one of the following alternatives holds:*

(i) *The curve $t \to c_t$ converges smoothly to some csc-K potential $c_\infty \in \mathcal{H}_\omega$ as $t \to \infty$.*

(ii) *The curve $t \to c_t$ diverges as $t \to \infty$ and encodes destabilizing information about the Kähler structure.*

We refer to Donaldson [39] for a precise statement and further details about this conjecture. To avoid the difficulties arising in PDE theory related to long-time existence, we recast the Calabi flow in the metric completion of $(\mathcal{H}_\omega, d_2)$ following Streets [61; 62], who applied the work of Mayer [56] and Bačák [3] concerning gradient flows of convex functionals on Hadamard spaces (ie $\mathrm{CAT}(0)$ spaces) to the setting of the "minimizing movement" Calabi flow. Before we can do this, however, we need to understand how the K-energy extends to certain spaces of singular potentials. The key new feature of our approach is that we take advantage of the fact that the corresponding abstract metric space (defined in terms of Cauchy sequences in [61]) can be realized concretely in terms of certain singular Kähler potentials, ie using pluripotential theory, which in particular allows us to improve on the abstract convergence result in [62].

Finite-energy spaces and extensions of the twisted K-energy In order to briefly introduce our setting, we denote by $\mathcal{E}^p$ the space of ω–psh functions on X which have finite energy with respect to the standard p–homogenous weight, as introduced by Guedj and Zeriahi [45]. As shown in Darvas and He [30], the abstract metric completion of the L^p–type Finsler metric (1) on $\mathcal{H}_\omega$ may be identified with the finite-energy space $\mathcal{E}^p$ equipped with a natural distance function that we will denote by d_p, which is comparable to an explicit energy-type expression (8). When $p = 2$, this identification was conjectured by Guedj in [44]. Furthermore, in the case $p = 1$, it yields a Finsler realization $(\mathcal{E}^1, d_1)$ of the strong topology on $\mathcal{E}^1$ introduced in Berman, Boucksom, Eyssidieux, Guedj and Zeriahi [9] (which can be seen as a higher-dimensional "nonlinear" generalization of the classical strong topology defined by the Dirichlet norm on a Riemann surface).

Moreover, as shown in Darvas [29] and Darvas and He [30], for any pair of potentials $u_0, u_1 \in \mathcal{E}^p$ one can construct a d_p–geodesic segment (in the metric sense) explicitly, as a decreasing pointwise limit of $C^{1,\bar{1}}$–weak geodesics (in the sense of Chen [20], ie as $C^{1,\bar{1}}$–solutions to certain complex Monge–Ampère equations). These d_p–geodesic segments will be referred to as *finite-energy geodesics* in the future and we direct the reader to Theorem 2.3 for more details. A recurrent theme in the present work is the interaction between the cases $p = 2$ and $p = 1$, which in particular will allow us to exploit the energy/entropy compactness theorem from [9] to get a convergence result for the Calabi flow with respect to the d_1–topology. This strengthens the general convergence result of [62], concerning the weak d_2–topology, which does not imply any convergence in the sense of pluripotential theory (Remark 5.4).

Our starting point is the observation that the K-energy functional $\mathcal{K}$ originally defined on $\mathcal{H}_\omega$ admits a natural "analytic extension" to the finite-energy space $\mathcal{E}^1$ (and hence by restriction to all spaces $\mathcal{E}^p$). This is simply the extension obtained by interpreting the entropy part (the first term) and the energy part (the second two terms) in formula (2) in the general sense of probability theory and pluripotential theory, respectively, essentially as in the Fano setting previously considered in Berman [7] and in [9]. As we will see, the energy part is d_1–continuous, whereas the entropy part is only d_1–lsc, and in the particular case of $C^{1,\bar{1}}$–potentials, this extension coincides with the one introduced by Chen [19]. We then go on to show that the restriction to $\mathcal{E}^p$ of the analytic extension coincides with the canonical "topological extension" of the K-energy, ie the greatest d_p–lsc extension from $\mathcal{H}_\omega$. In particular, applied to the case $p = 2$, which is the one relevant to the Calabi flow, this yields an analytic formula for Streets' extension of the K-energy.

The analytic extension formula allows us to establish the convexity of the extended K-energy along finite-energy geodesics, using an approximation argument and the $C^{1,\bar{1}}$–case recently settled in Berman and Berndtsson [8, Theorem 1.1] (originally conjectured by Chen).

Before we state our first theorem, recall that in various applications of Kähler geometry it is necessary to deal with the more general concept of twisted csc-K metrics and the corresponding twisted-K energy (see eg Chen [23], Chen, Paun and Zeng [26], Dervan [36], Fine [41] and Stoppa [60]). As it takes little extra effort, throughout this paper we work at this level of generality, with χ denoting a very general twisting form (3) and $\mathcal{K}_\chi$ the corresponding twisted K-energy (5). The relevant terminology will be recalled in Section 2.1.

Theorem 1.2 (Theorem 4.7) *Suppose (X, ω) is a compact connected Kähler manifold. The K-energy can be extended to a functional $\mathcal{K}\colon \mathcal{E}^1 \to (-\infty, \infty]$ using (2). The* *restricted functional* $\mathcal{K}|_{\mathcal{E}^p}$ *is the greatest* d_p*–lsc extension of* $\mathcal{K}|_{\mathcal{H}_\omega}$ *for any* $p \geq 1$*. Additionally,* $\mathcal{K}|_{\mathcal{E}^p}$ *is convex along the finite-energy geodesics of* $\mathcal{E}^p$*. If* $\chi = \beta + i\partial\overline{\partial} f$ *satisfies (3), the corresponding result also holds for the twisted K-energy* $\mathcal{K}_\chi$*.*

An important ingredient in the proof of Theorem 1.2 is understanding approximation of potentials of $\mathcal{E}^p$ while also approximating entropy. In this direction, we note the following theorem. More precise results can be obtained using the flow techniques of Guedj and Zeriahi [46] and Nezza and Lu [57], and will be discussed elsewhere.

Theorem 1.3 (Theorem 3.2) *Suppose* $u \in \mathcal{E}^p$ *and* f *is a usc function on* X *satisfying* $e^{-f} \in L^1(X, \omega^n)$*. Then one can find* $u_k \in \mathcal{H}_\omega$ *with* $d_p(u_k, u) \to 0$ *and* $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_{u_k}) \to \mathrm{Ent}(e^{-f}\omega^n, \omega^n_u)$*.*

Finally, as a consequence of Theorem 1.2 we obtain that the space of finite χ–entropy potentials $\mathrm{Ent}_\chi(X, \omega)$ is geodesically closed, and if $\mathrm{Ric}\,\omega \geq \beta$ then the twisted entropy is convex along finite-energy geodesics, giving the Kähler analog of a central result of Lott, Sturm and Villani in optimal transport theory; see Villani [66]. For details on notation and a detailed discussion on relationship with the literature, we refer to Section 4.4.

Theorem 1.4 (Theorem 4.10) *If* $\chi = \beta + i\partial\overline{\partial} f$ *satisfies (3), then* $(\mathrm{Ent}_\chi(X, \omega), d_1)$ *is a geodesic sub-metric space of* $(\mathcal{E}^1(X, \omega), d_1)$*. Additionally, if* $\mathrm{Ric}\,\omega \geq \beta$ *then the map* $\mathrm{Ent}_\chi(X, \omega) \ni u \to \mathrm{Ent}(e^{-f}\omega^n, \omega^n_u) \in \mathbb{R}$ *is convex along finite-energy geodesics.*

Convergence and large-time behavior of the weak twisted Calabi flow As advertised above, using Theorem 1.2, we can run the weak twisted Calabi flow $[0, \infty) \ni t \to c_t \in \mathcal{E}^2$ for any starting point $c_0 \in \mathcal{E}^2$. Indeed, $(\mathcal{E}^2, d_2)$ is a $\mathrm{CAT}(0)$-space and the extended functional $\mathcal{K}_\chi$ is convex along d_2–geodesics, hence we are in the setting of Mayer [56], as detailed in Section 2.5. This yields a flow of (possibly singular) Kähler potentials which is uniquely determined by the corresponding normalized Monge–Ampère measures, which in turn yields a flow of probability measures which is regularizing in the sense that the entropy immediately becomes finite and in particular the measures have an L^1–density for positive times.

When χ is smooth and X is a Riemann surface, the smooth twisted Calabi flow was recently explored by Pook [58]. To provide consistency, we will show that the weak twisted Calabi flow agrees with the smooth version whenever the latter exists (Proposition 6.1), generalizing a result of Streets [62] in the case where $\chi = 0$. Providing additional consistency, as an application of Theorem 1.2, in Section 6 we show that Streets' (a priori different) minimizing movement Calabi flow coincides with our weak Calabi flow.

Generalizing twisted csc-K metrics, by $\mathcal{M}^p_\chi$ we denote the minimizers of the extended K-energy on $\mathcal{E}^p$:

$$\mathcal{M}^p_\chi = \{u \in \mathcal{E}^p : \mathcal{K}_\chi(u) = \inf_{v \in \mathcal{E}^p} \mathcal{K}_\chi(v)\}.$$

In the case $\chi = 0$ we will simply use $\mathcal{M}^p := \mathcal{M}^p_0$. Concerning the convergence and blow-up behavior of the weak twisted Calabi flow, we prove the following concrete result:

Theorem 1.5 (Theorem 6.3) *Suppose (X, ω) is a compact connected Kähler manifold and $\chi = \beta + i\partial\bar{\partial} f$ satisfies (3). The following statements are equivalent:*

(i) *$\mathcal{M}^2_\chi$ is nonempty.*

(ii) *For any weak twisted Calabi flow trajectory $t \to c_t$, there exists $c_\infty \in \mathcal{M}^2_\chi$ such that $d_1(c_t, c_\infty) \to 0$ and $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_{c_t}) \to \mathrm{Ent}(e^{-f}\omega^n, \omega^n_{c_\infty})$.*

(iii) *Any weak twisted Calabi flow trajectory $t \to c_t$ is d_2–bounded.*

(iv) *There exists a weak twisted Calabi flow trajectory $t \to c_t$ and $t_j \to \infty$ for which the sequence $\{c_{t_j}\}_j$ is d_2–bounded.*

- By the consistency result discussed above, the previous theorem in particular applies to the smooth Calabi flow (when it exists) and it should be stressed that the result and its elaborations discussed below are new also in this smooth case. In particular, it generalizes results of the first author on the smooth Calabi flow on Fano manifolds without nontrivial holomorphic vector fields; see Berman [7]. One new feature of our result is that the latter assumption, which guarantees the uniqueness of csc-K metrics, is not needed. This means that the limit c_∞ is not uniquely determined by X and will, in general, depend on the initial data c_0.

- By Darvas and He [30, Theorem 5] and part (ii) of the above theorem, if a csc-K potential exists in $\mathcal{H}_\omega$ then the weak Calabi flow $t \to c_t$ converges pointwise ae to some potential $c_\infty \in \mathcal{M}^2$, and the measures $\omega^n_{c_t}$ converge weakly and in entropy to $\omega^n_{c_\infty}$. In the Fano case it additionally follows that c_∞ is csc-K. However, due to progress on the regularity Conjecture 1.8 discussed in the companion paper Berman, Darvas and Lu [11], this result also holds on general Kähler manifolds as well, making further progress on Donaldson's conjecture (see Theorem 1.10 and Theorem 1.11).

- Finally, in light of Theorem 1.6, we mention that Proposition 2.11(ii) strengthens the corresponding convergence result of Streets in [62]. Given a CAT(0) metric space (M, d), it is possible to introduce a notion of weak d–convergence, generalizing the concept of weak convergence on Hilbert spaces (Section 2.4). In general, little concrete is known about this type of convergence; see Kirk and Panyanak [51]. Streets, however, observed that one can adapt the result of Bačák [3] to our setting, ie whenever $\mathcal{M}^2$ is nonempty, each weak Calabi flow trajectory converges d_2–weakly to an element of $\mathcal{M}^2$ [62]. Though weak d_2 convergence does not even imply weak L^1 convergence of the potentials (Remark 5.4), we use this idea in the proof of the above theorem together with the following result, which sheds light on the relationship between all the different topologies involved:

Theorem 1.6 (Theorem 5.3) *Suppose $\{u_k\}_k \subset \mathcal{E}^2$ is d_2–bounded and $u \in \mathcal{E}^2$. Then $d_1(u_k, u) \to 0$ if and only if $\|u_j - u\|_{L^1(X)} \to 0$ and u_k converges to u d_2–weakly.*

The conjectural picture of Donaldson Before we proceed, let us note a last corollary of Theorem 1.5, a consequence of the equivalence between (i) and (iv):

Corollary 1.7 *Suppose that (X, ω) is a compact connected Kähler manifold and that $[0, \infty) \ni t \to c_t \in \mathcal{E}^2$ is a weak twisted Calabi flow trajectory. Exactly one of the following holds:*

(i) *The curve $t \to c_t$ d_1–converges to some $c_\infty \in \mathcal{M}^2_\chi$.*

(ii) *$d_2(c_0, c_t) \to \infty$ as $t \to \infty$.*

Though this corollary is in line with Donaldson's conjectural picture, one would like to understand how a diverging Calabi flow trajectory "destabilizes" the Kähler structure, as proposed in Conjecture 1.1. In this direction we recall the following concept from Darvas and He [31]: suppose (M, d) is a geodesic metric space and $[0, \infty) \ni t \to \gamma_t \in M$ is a continuous curve. We say that the unit speed d–geodesic ray $[0, \infty) \ni t \to g_t \in M$ is *d–weakly asymptotic* to the curve $t \to \gamma_t$ if there exists $t_j \to \infty$ and unit speed d–geodesic segments $[0, d(\gamma_0, \gamma_{t_j})] \ni t \to g^j_t \in M$ connecting γ_0 and γ_{t_j} such that $\lim_{j\to\infty} d(g^j_t, g_t) = 0$ for $t \in [0, \infty)$.

Clearly, to have a geodesic ray weakly asymptotic to $t \to \gamma_t$, we need $t \to d(\gamma_0, \gamma_t)$ to be unbounded. By the above corollary, this condition makes diverging weak Calabi flow trajectories $t \to c_t$ perfect candidates for this construction. However, more needs to be known about $t \to c_t$ before we can proceed. In Darvas and Rubinstein [32, Conjecture 2.8] it was pointed out that an important roadblock in resolving Tian's properness conjecture for csc-K metrics is a conjecture about regularity of minimizers of $\mathcal{K}$. The twisted version of this conjecture should also hold:

Conjecture 1.8 (Darvas and Rubinstein [32]) *Suppose (X, ω) is a compact connected Kähler manifold and χ is smooth. Then $\mathcal{M}^1_\chi \subset \mathcal{H}_\omega$, ie $\mathcal{M}^1_\chi$ contains only smooth twisted csc-K potentials.*

We note that this conjecture generalizes an earlier conjecture of Chen [22, Conjecture 6.3] about $C^{1,1}$ minimizers of $\mathcal{K}$. When (X,ω) is Fano, Conjecture 1.8 was proved in Berman [7] and Berman, Boucksom, Eyssidieux, Guedj and Zeriahi [9]. The next result partially confirms Donaldson's conjecture in the Fano case and also in the case when χ is a Kähler form.

Theorem 1.9 (Theorem 6.5) *Suppose (X,ω) is a compact connected Kähler manifold, $\chi \geq 0$ is smooth and Conjecture 1.8 holds. Let $[0,\infty) \ni t \to c_t \in \mathcal{E}^2$ be a weak twisted Calabi flow trajectory. Exactly one of the following holds:*

(i) *The curve $t \to c_t$ d_1–converges to a smooth twisted csc-K potential c_∞.*

(ii) *$d_1(c_0,c_t) \to \infty$ as $t \to \infty$ and the curve $t \to c_t$ is d_1–weakly asymptotic to a finite-energy geodesic $[0,\infty) \ni t \to u_t \in \mathcal{E}^1$ along which $\mathcal{K}_\chi$ decreases.*

If $\chi > 0$, then, independently of Conjecture 1.8, exactly one of the following holds:

(i′) *The curve $t \to c_t$ d_1–converges to a unique minimizer in $\mathcal{E}^1$ of $\mathcal{K}_\chi$.*

(ii′) *$d_1(c_0,c_t) \to \infty$ as $t \to \infty$ and the curve $t \to c_t$ is d_1–weakly asymptotic to a finite-energy geodesic $[0,\infty) \ni t \to u_t \in \mathcal{E}^1$ along which $\mathcal{K}_\chi$ strictly decreases.*

Though stated differently, when (X,ω) is Fano and $\chi = 0$ the analog of this result for the Kähler–Ricci flow has been obtained in [31, Theorem 2]. There we have smooth convergence in (i) and along the geodesic ray of (ii) the potentials are bounded, all thanks to the Perelman estimates available for the Kähler–Ricci flow. It would be interesting to compare the above theorem to the results in Chen and Sun [27], where the authors construct in a specific situation a geodesic ray asymptotic to the Calabi flow and are able to draw geometric conclusions based on this.

Concluding remarks and additional results Based on geometric considerations, and the analogous picture in case of the Kähler–Ricci flow (see Guedj and Zeriahi [46]), it is natural to speculate that for any starting point $c_0 \in \mathcal{E}^2$, the weak Calabi flow $t \to c_t$ is instantly smooth, ie $c_t \in \mathcal{H}_\omega$ for $t > 0$ (see also Chen [23, Conjecture 3.5]). Such a result would instantly give the $\mathcal{E}^2$ version of Conjecture 1.8, that $\mathcal{E}^2$–minimizers of $\mathcal{K}$ are smooth csc-K metrics. Indeed, by the general result of Mayer [56], the weak Calabi flow $t \to c_t$ starting at a minimizer $c_0 \in \mathcal{E}^2$ has to be stationary. If $t \to c_t$ was instantly smooth, then we could conclude that $c_0 \in \mathcal{H}_\omega$.

In the companion paper [11] we make progress on Conjecture 1.8 using different techniques from the ones presented in this paper:

Theorem 1.10 *Suppose (X,ω) is a Kähler manifold and $\mathcal{H}_\omega$ contains a csc-K potential. Then $\mathcal{M}^1$ contains only smooth csc-K potentials.*

The consequences of this theorem related to K-stability and energy properness will be discussed in [11]. As $\mathcal{M}^2 \subset \mathcal{M}^1$, here we just mention the following consequence of this result and Theorem 1.5(ii), making further progress on Conjecture 1.1 (see also Streets [62, Remark 1.10]):

Theorem 1.11 *Suppose (X,ω) is a Kähler manifold and $\mathcal{H}_\omega$ contains a csc-K potential u. Then any weak Calabi flow trajectory $t \to c_t$ d_1–converges to a smooth csc-K potential $c_\infty \in \mathcal{H}_\omega$. In addition, the densities $\omega^n_{c_t}/\omega^n$ converge in L^1 to the density $\omega^n_{c_\infty}/\omega^n$.*

Organization of the paper In the first part of Section 2 we recall recent results on complex Monge–Ampère theory which we will use in this paper. In the second part we briefly recall Mayer's theory of gradient flows on nonpositively curved metric spaces. The approximation of finite-energy ω–plurisubharmonic functions with convergent entropy is presented in Section 3. The twisted Mabuchi energy is studied in Section 4. The weak d_2 topology is explored in Section 5, while the last section is devoted to the weak twisted Calabi flow.

Acknowledgments We would like to thank László Lempert for his careful suggestions regarding the presentation of the paper. Berman and Lu are supported by the European Research Council. Lu is on leave from Chalmers University of Technology. Darvas is partially supported by BSF grant 2012236 and NSF grant DMS 1610202.

2 Preliminaries

2.1 The twisted K-energy

Suppose χ is a closed positive $(1,1)$–current and β is a smooth closed $(1,1)$–form in the same cohomology class as χ. In most applications of Kähler geometry, the twisting current χ can be smooth, but in order to treat the case of smooth and singular canonical metrics (eg conical csc-K metrics) together, it is natural to ask for the following more general restriction on χ:

$$\chi = \beta + i\partial\bar\partial f, \quad \text{where } f \in \mathrm{PSH}(X,\beta) \text{ with } e^{-f} \in L^1(X,\omega^n). \tag{3}$$

We observe that the integrability condition $e^{-f} \in L^1(X,\omega^n)$ implies that $e^{-f} \in L^p(X,\omega^n)$ for some $p > 1$, as follows from the openness conjecture, recently proved by Berndtsson [13] (see also [43]). We note that some of our results, in particular Theorem 1.2 above, hold for more general χ. However, it is unlikely that greater generality will have applications, and we leave it to the reader to find optimal conditions for χ in our theorems.

The twisted K-energy $\mathcal{K}_\chi\colon \mathcal{H}_\omega \to \mathbb{R}$ can now be defined as

$$\mathcal{K}_\chi(u) = \mathrm{Ent}(e^{-f}\omega^n, \omega_u^n) + \bar{S}_\chi \,\mathrm{AM}(u) - n\,\mathrm{AM}_{\mathrm{Ric}\,\omega-\beta}(u) - \int_X f\omega^n, \tag{4}$$

where $\bar{S}_\chi = nV^{-1}\int_X (\mathrm{Ric}_\omega - \chi)\wedge\omega^{n-1}$. Notice that for $\beta = 0$, $f = 0$ we get back the usual K-energy (2). Using the identity $n\,\mathrm{AM}_\chi(u) = n\,\mathrm{AM}_\beta(u) + \int f\omega_u^n - \int f\omega^n$ one can give an alternative formula for $\mathcal{K}_\chi$, perhaps more familiar from the literature:

$$\mathcal{K}_\chi(u) = \mathrm{Ent}(\omega^n, \omega_u^n) + \bar{S}_\chi \,\mathrm{AM}(u) - n\,\mathrm{AM}_{\mathrm{Ric}\,\omega-\chi}(u). \tag{5}$$

The virtue of this formula is that it shows that $\mathcal{K}_\chi$ is independent of the choice of β and f. As will be made clear shortly, when trying to extend $\mathcal{K}_\chi$, our original definition is more advantageous, however. Note that when χ is smooth, the first-order variation of $\mathcal{K}_\chi$ is given by the formula

$$\langle D\mathcal{K}_\chi(u), \delta v\rangle = V^{-1}\int_X \delta v(\bar{S}_\chi - S_{\omega_u} + \mathrm{Tr}^{\omega_u}\chi)\omega_u^n.$$

Hence, the critical points of this functional are the twisted csc-K potentials, as these satisfy $\bar{S}_\chi - S_{\omega_u} + \mathrm{Tr}^{\omega_u}\chi = 0$. The smooth twisted Calabi flow is defined analogously.

2.2 The complete geodesic metric spaces $(\mathcal{E}^p, d_p)$

In this section we summarize results from [29; 30; 17; 9] needed the most in this paper. Formula (1) introduces L^p–type weak Finsler metrics on the Fréchet manifold $\mathcal{H}_\omega$. A curve $[0,1] \ni t \to \alpha_t \in \mathcal{H}_\omega$ is called smooth if $\alpha(t,z) = \alpha_t(z) \in C^\infty([0,1]\times M)$. The L^p–length of a smooth curve $t \to \alpha_t$ is given by

$$l_p(\alpha) := \int_0^1 \|\dot{\alpha}_t\|_{p,\alpha_t}\,dt.$$

Definition 2.1 The path length pseudo-distance of $(\mathcal{H}_\omega, d_p)$ is defined by

$$d_p(u_0,u_1) := \inf\{l_p(\alpha) : [0,1]\ni t \to \alpha_t \in \mathcal{H}_\omega \text{ is a smooth curve with } \alpha_0 = u_0,\ \alpha_1 = u_1\}.$$

It turns out that d_p is an honest metric [30, Theorem 3.5]. To state the result, consider $[0,1]\times\mathbb{R}\times X$ as a complex manifold of dimension $n+1$, and let $\pi_2\colon [0,1]\times\mathbb{R}\times X \to X$ be the natural projection.

Theorem 2.2 *$(\mathcal{H}_\omega, d_p)$ is a metric space. Moreover, for any $t \in [0,1]$,*

$$d_p(u_0,u_1) = \|\dot{u}_t\|_{p,u_t} \geq 0,$$

where $\dot{u}_t = du_t/dt$ is the "tangent" at time t of $t \to u_t$, the $\mathbb{R}$–invariant solution of the Monge–Ampère equation

$$\begin{cases} \varphi \in \mathrm{PSH}(\pi_2^\star \omega, [0,1] \times \mathbb{R} \times X), \\ (\pi_2^\star \omega + \sqrt{-1}\partial\overline{\partial}\varphi)^{n+1} = 0, \\ \varphi|_{\{i\}\times\mathbb{R}} = u_i, \quad i = 0, 1. \end{cases} \tag{6}$$

Some comments are in order. By the main result of [20] (see also [14]), the equation (6) has a unique $\mathbb{R}$–invariant solution for which $u(t,x) = u_t(x)$ has bounded Laplacian in $[0,1] \times \mathbb{R} \times X$. We can look at this solution as a curve

$$[0,1] \ni t \to u_t \in \mathcal{H}^\Delta_\omega = \{u \in \mathrm{PSH}(X,\omega) : \Delta_\omega u \in L^\infty(X)\}.$$

We call this curve the *weak geodesic* connecting $u_0, u_1 \in \mathcal{H}_\omega$. Recall that

$$\mathrm{PSH}(X,\omega) = \{\varphi \in L^1(X,\omega^n) : \varphi \text{ is usc and } \omega_\varphi \geq 0\}.$$

Given $\varphi_k \in \mathrm{PSH}(X,\omega)$, $k = 1, \dots, n$, one can introduce the following *nonpluripolar product* [17], generalizing the Bedford–Taylor product [6] concerning the case with bounded potentials:

$$\omega_{\varphi_1} \wedge \omega_{\varphi_2} \wedge \cdots \wedge \omega_{\varphi_n} := \lim_{j \to -\infty} \mathbf{1}_{\bigcap_k \{\varphi_k > j\}} \omega_{\max(\varphi_1, j)} \wedge \omega_{\max(\varphi_2, j)} \wedge \cdots \wedge \omega_{\max(\varphi_n, j)}. \tag{7}$$

The measures $\omega_{\max(\varphi_1,j)} \wedge \cdots \wedge \omega_{\max(\varphi_n,j)}$ are defined by the work of Bedford and Taylor [6] since $\max\{\varphi, j\}$ is bounded. Restricted to $\bigcap_k \{\varphi_k > j\}$, these measures are increasing, hence the above limit is well defined [45; 17] and $\int_X \omega_{\varphi_1} \wedge \cdots \wedge \omega_{\varphi_n} \leq \int_X \omega^n$.

Following Guedj and Zeriahi [45, Definition 1.1] we introduce the class of potentials with "full volume", $\mathcal{E}(X,\omega) := \{\varphi \in \mathrm{PSH}(X,\omega) : \int_X \omega_\varphi^n = \int_X \omega^n\}$, and the corresponding finite-energy classes

$$\mathcal{E}^p := \left\{\varphi \in \mathcal{E}(X,\omega) : \int_X |\varphi|^p \omega_\varphi^n < \infty\right\}.$$

The next result characterizes the d_p–metric completion of $\mathcal{H}_\omega$:

Theorem 2.3 [30, Theorem 2] *The metric completion of $(\mathcal{H}_\omega, d_p)$ equals $(\mathcal{E}^p, d_p)$, where*

$$d_p(u_0, u_1) := \lim_{k \to \infty} d_p(u_0^k, u_1^k),$$

for any smooth decreasing sequences $\{u_i^k\}_{k \in \mathbb{N}} \subset \mathcal{H}_\omega$ converging pointwise to $u_i \in \mathcal{E}^p$, $i = 0, 1$. Moreover, for each $t \in (0,1)$, define

$$u_t := \lim_{k \to \infty} u_t^k,$$

where u_t^k is the weak geodesic connecting u_0^k and u_1^k. Then $u_t \in \mathcal{E}^p$, the curve $[0,1] \ni t \to u_t \in \mathcal{E}^p$ is well-defined independently of the choices of approximating sequences, and this curve is a d_p–geodesic.

In the rest of the paper we will call the d_p–geodesics constructed in this theorem *finite-energy geodesics*. As mentioned in [30], for arbitrary p and u_0, u_1, the finite-energy geodesic joining these potentials may not be unique as a d_p–geodesic.

By [34; 15] it is always possible to find approximating sequences $\{u_0^k\}_k, \{u_1^k\}_k$ as in the above theorem. We now recall [30, Theorem 3], giving a concrete characterization of the growth of all d_p metrics:

Theorem 2.4 *There exists $C > 1$ such that, for all $u, v \in \mathcal{E}^p$,*

$$C^{-1} d_p(u,v) \le \left(\int_X |u-v|^p \omega_u^n\right)^{1/p} + \left(\int_X |u-v|^p \omega_v^n\right)^{1/p} \le C d_p(u,v). \tag{8}$$

The inequalities in (8) have an important consequence: $|\sup_X u| \le C d_p(u,0)$ for all $u \in \mathcal{E}^p$. Also, when $p=1$, d_1–convergence is equivalent to convergence with respect to the quasidistance $I(u,v) = \int_X (u-v)(\omega_v^n - \omega_u^n)$ introduced in [9], as shown in [30, Theorem 5.5].

Monotonic sequences behave well with respect to all d_p–metrics [30, Proposition 4.9]:

Proposition 2.5 *Suppose $u_k, u \in \mathcal{E}^p$. If $\{u_k\}_k$ is monotone decreasing/increasing and converges to u ae then $d_p(u_k, u) \to 0$.*

Given $u_0, u_1, \dots, u_k \in \mathrm{PSH}(X,\omega)$, by $P(u_0, u_1, \dots, u_k) \in \mathrm{PSH}(X,\omega)$ we define the upper envelope

$$P(u_0, u_1, \dots, u_k) = \sup\{v \in \mathrm{PSH}(X,\omega) \text{ such that } v \le u_0, \dots, v \le u_k\}.$$

According to the next proposition it is possible to sandwich a subsequence of any d_p–convergent sequence between two monotone sequences converging to the same limit.

Proposition 2.6 *Suppose $u_k, u \in \mathcal{E}^p$. If $d_p(u_k, u) \to 0$ then there exists a subsequence $k_j \to \infty$ and $\{w_{k_j}\}_j \subset \mathcal{E}^p$ decreasing, $\{v_{k_j}\}_j \subset \mathcal{E}^p$ increasing with $v_{k_j} \le u_{k_j} \le w_{k_j}$ and $d_p(w_{k_j}, u), d_p(v_{k_j}, u) \to 0$.*

Proof By (8) there exists $C > 0$ such that $|\sup_X u_j| \le C$ for $j \ge 1$. We introduce the sequence

$$w_k = \mathrm{usc}(\sup_{j \ge k} u_j).$$

As $u_k \leq w_k \leq C$, by [45] it follows that $w_k \in \mathcal{E}^p$. As $d_p(u_k, u) \to 0$, we have that $u_k \to u$ pointwise ae, hence w_k decreases to u. Proposition 2.5 then gives $d_p(w_k, u) \to 0$.

Now we construct the increasing sequence v_{k_j}. To do this, first take a subsequence u_{k_j} of u_k satisfying $d_p(u_{k_j}, u) \leq 2^{-j}$. As follows from the proof of [30, Theorem 4.17] and [29, Theorem 9.2], the following limit exists:

$$v_{k_j} = P(u_{k_j}, u_{k_{j+1}}, u_{k_{j+2}}, \ldots) := \lim_{h\to\infty} P(u_{k_j}, u_{k_{j+1}}, \ldots, u_{k_{j+h}}).$$

Additionally, $\{v_{k_j}\}_j \subset \mathcal{E}^p$ and v_{k_j} increases ae to u. The previous proposition now gives $d_p(u, v_{k_j}) \to 0$. □

Though stated differently, the next proposition is essentially contained in [17]:

Proposition 2.7 *Suppose $p \geq 1$, $\{u_j\}_j \subset \mathcal{E}^p$ is a d_p–bounded sequence and $u \in \operatorname{PSH}(X, \omega)$ with $\|u_j - u\|_{L^1(X,\omega^n)} \to 0$. Then $u \in \mathcal{E}^p$.*

Proof Boundedness with respect to d_p implies that $|\sup_X u_j| \leq B$ for some $B \in \mathbb{R}$ (8). For simplicity assume that $B = 0$. The following sequence converges ae to u:

$$w_k = \operatorname{usc}(\sup_{j\geq k} u_j) \leq 0.$$

This sequence is additionally decreasing, and because $u_k \leq w_k \leq 0$, we have that $w_k \in \mathcal{E}^p$. If we could argue that $\{w_k\}_k$ is uniformly d_p–bounded then we would be finished by [30, Lemma 4.16]. But d_p–boundedness follows from (8). Indeed,

$$\int_X |w_k|^p \omega^n \leq \int_X |u_k|^p \omega^n \quad \text{and} \quad \int_X |w_k|^p \omega_{w_k}^n \leq C(p) \int_X |u_k|^p \omega_{u_k}^n$$

by [45, Lemma 3.5], hence by (8) the quantity $d_p(0, w_k)$ is uniformly bounded. □

Given two Borel measures μ, ν on X, if ν is not subordinate to μ, then by definition $\operatorname{Ent}(\mu, \nu) = \infty$. On the other hand, if ν is subordinate to μ then $\operatorname{Ent}(\mu, \nu) = \int_X \log(f)\nu$, where f is the Radon–Nikodym density of ν with respect to μ. The entropy functional $\mu \to \operatorname{Ent}(\mu, \nu)$ is lsc with respect to weak convergence of measures [35]. Related to entropy, we recall the following crucial compactness result [9, Theorem 2.17]:

Theorem 2.8 *Let $p > 1$ and suppose $\mu = f\omega^n$ is a probability measure with $f \in L^p(X, \omega^n)$. Suppose there exists $C > 0$ such that $\{u_k\}_k \subset \mathcal{E}^1$ satisfies*

$$|\sup_X u_k| < C, \quad \operatorname{Ent}(\mu, \omega_{u_k}^n) < C.$$

Then $\{u_k\}_k$ contains a d_1–convergent subsequence.

2.3 The complex Monge–Ampère equation in $\mathcal{E}^p$

We summarize in this section basic results concerning solutions of degenerate complex Monge–Ampère equations that are needed in this paper.

A subset $E \subset X$ is called pluripolar if it is contained in the singular set of a function $\varphi \in \mathrm{PSH}(X,\omega)$, ie $E \subset \{\varphi = -\infty\}$. Let μ be a positive measure on X with total mass $\mu(X) = \int_X \omega^n$. We consider the equation

$$\omega_\varphi^n = \mu. \tag{9}$$

It was proved in [45, Theorem A] that when μ does not charge pluripolar sets the equation (9) has a solution $\varphi \in \mathcal{E}(X,\omega)$. The solution turns out to be unique up to an additive constant [37]. For each $\varepsilon > 0$, the same variational approach as in the proof of Theorem C on page 222 of [10] (see also [47, Corollary 11.9]) applied to the functional

$$F_\varepsilon(u) := \mathrm{AM}(u) - \frac{1}{\varepsilon}\log\int_X e^{\varepsilon u}\,d\mu, \quad u \in \mathcal{E}^1,$$

shows that there exists a solution $\varphi_\varepsilon \in \mathcal{E}^1$ to the equation

$$\omega_{\varphi_\varepsilon}^n = e^{\varepsilon\varphi_\varepsilon}\mu. \tag{10}$$

The solution is uniquely determined as follows from the comparison principle (see [12, Proposition 4.1]). The following version of the comparison principle will be useful later.

Lemma 2.9 *Let $\varepsilon > 0$. Assume that $\varphi \in \mathcal{E}(X,\omega)$ is a solution of (10), while $\psi \in \mathcal{E}(X,\omega)$ is a subsolution, ie $\omega_\psi^n \geq e^{\varepsilon\psi}\mu$. Then $\varphi \geq \psi$ on X.*

This result might be well known to experts in Monge–Ampère theory. As a courtesy to the reader we give a proof below.

Proof By the comparison principle for the class $\mathcal{E}(X,\omega)$ (see [45, Theorem 1.5]) we have

$$\int_{\{\varphi<\psi\}} \omega_\psi^n \leq \int_{\{\varphi<\psi\}} \omega_\varphi^n.$$

As φ is a solution and ψ is a subsolution to (10) we also have

$$\int_{\{\varphi<\psi\}} e^{\varepsilon\psi}\,d\mu \leq \int_{\{\varphi<\psi\}} \omega_\psi^n \leq \int_{\{\varphi<\psi\}} \omega_\varphi^n = \int_{\{\varphi<\psi\}} e^{\varepsilon\varphi}\,d\mu \leq \int_{\{\varphi<\psi\}} e^{\varepsilon\psi}\,d\mu.$$

It follows that all inequalities above are equalities, hence $\varphi \geq \psi$ μ–almost everywhere on X. By Dinew's domination principle [16, Proposition 5.9] we get $\varphi \geq \psi$ everywhere on X. □

One might wonder whether the solution of (9) arises as a limit of solutions of (10) as $\varepsilon \to 0$. The following result answers this affirmatively.

Lemma 2.10 *Let $p \geq 1$. Assume that $\mu = \omega_\varphi^n$ with $\varphi \in \mathcal{E}^p$ and $\int_X \varphi \, d\mu = 0$. For each $\varepsilon > 0$, let $\varphi_\varepsilon \in \mathcal{E}^1$ be the unique solution to (10). Then in fact $\varphi_\varepsilon \in \mathcal{E}^p$ and $d_p(\varphi_\varepsilon, \varphi) \to 0$ as $\varepsilon \to 0$.*

Proof As $\varphi - \sup_X \varphi$ is a subsolution of (10), it follows from Lemma 2.9 that $\varphi_\varepsilon \geq \varphi - \sup_X \varphi$ for all $\varepsilon > 0$, hence $\varphi_\varepsilon \in \mathcal{E}^p$. We claim that φ_ε is uniformly bounded from above for $\varepsilon \in [0, 1]$. Assume on the contrary that we can extract a subsequence denoted by $\varphi_j = \varphi_{\varepsilon_j}$ such that $\sup_X \varphi_j \to \infty$. The sequence $\psi_j := \varphi_j - \sup_X \varphi_j$ stays in a compact set in $L^1(X, \omega^n)$, hence a subsequence (still denoted by φ_j) converges to some $\psi \in \mathrm{PSH}(X, \omega)$. It then follows that $\varphi_j = \psi_j + \sup_X \varphi_j$ converges uniformly to ∞. In the other hand, by Jensen's inequality (for simplicity we may assume that $\mu(X) = 1$) we have

$$\int_X \varphi_j \, d\mu \leq 0.$$

Since φ_j is bounded from below by $\varphi - \sup_X \varphi$, which is integrable with respect to $d\mu$, the above inequality contradicts the fact that φ_j converges uniformly to ∞. Hence the claim follows.

Now the family φ_ε stays in a compact set of $L^1(X, \omega^n)$. As $\varepsilon \to 0$ each cluster point φ_0 satisfies

$$\omega_{\varphi_0}^n \geq \left(\liminf_{\varepsilon \to 0} e^{\varepsilon \varphi_\varepsilon}\right) \mu = \mu,$$

as follows from [17, Corollary 2.21]. As the two measures have the same total mass, one obtains equality. That $\varphi_0 = \varphi$ follows from uniqueness of complex Monge–Ampère measures [37] and the identity

$$0 = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \log \int_X e^{\varepsilon \varphi_\varepsilon} \, d\mu = \int_X \varphi_0 \, d\mu.$$

Finally, the last statement can be addressed using the identity

$$\int_X |\varphi_\varepsilon - \varphi|^p (\omega_{\varphi_\varepsilon}^n + \omega_\varphi^n) = \int_X |\varphi_\varepsilon - \varphi|^p (e^{\varepsilon \varphi_\varepsilon} + 1) \omega_\varphi^n.$$

Using this, (8) and the fact that $\sup_X \varphi_\varepsilon$ is bounded from above, by the dominated convergence theorem we conclude that $d_p(\varphi_\varepsilon, \varphi) \to 0$. □

2.4 Weak convergence in a CAT(0) space

Let us recall that a geodesic metric space (M, d) is a metric space for which any two points can be connected with a geodesic. By a geodesic connecting two points $a, b \in M$ we understand a curve α: $[0, 1] \to M$ such that $\alpha(0) = a$, $\alpha(1) = b$ and

$$d(\alpha(t_1), \alpha(t_2)) = |t_1 - t_2| d(a, b)$$

for any $t_1, t_2 \in [0,1]$. Furthermore, a geodesic metric space (M,d) is nonpositively curved (in the sense of Alexandrov) or CAT(0) if for any distinct points $q, r \in M$ there exists a geodesic $\gamma\colon [0,1] \to M$ joining q and r such that for any $s \in \{\gamma\}$ and $p \in M$ the inequality

$$d(p,s)^2 \le \lambda d(p,r)^2 + (1-\lambda) d(p,q)^2 - \lambda(1-\lambda) d(q,r)^2$$

is satisfied, where $\lambda = d(q,s)/d(q,r)$. A basic property of CAT(0) spaces is that geodesic segments joining different points are unique. For more about these spaces we refer to [18].

Let $\{x_n\}_n$ be a bounded sequence in a CAT(0) metric space (M,d). For $x \in M$, we set

$$r(x, \{x_n\}_n) = \limsup d(x, x_n).$$

The *asymptotic radius* of $\{x_n\}_n$ is given by $r(\{x_n\}_n) = \inf\{r(x, \{x_n\}_n) : x \in M\}$, and the *asymptotic center* $A(\{x_n\}_n)$ of $\{x_n\}_n$ is the set

$$A(\{x_n\}_n) = \{x \in M : r(x, \{x_n\}_n) = r(\{x_n\}_n)\}.$$

It is well known (see eg [62, Lemma 4.3]) that, in a CAT(0) space, $A(\{x_n\}_n)$ consists of exactly one point. A sequence $\{x_n\}_n$ *converges d–weakly* to $x \in M$ if x is the asymptotic center of all subsequences of $\{x_n\}_n$.

For a more detailed account of weak d–convergence we refer to [51], and for results related to the Calabi flow to [62, Section 4]. If (M,d) is a Hilbert space then weak d–convergence is the same as weak convergence in the sense of Hilbert spaces. With this in mind, the contents of the next result may seem less surprising:

Proposition 2.11 *Suppose (M,d) is a* CAT(0) *space. The following hold:*

(i) [51, Proposition 3.5] *If $\{x_n\}_n$ is a d–bounded sequence then it has a weak d–convergent subsequence.*

(ii) [51, Proposition 3.2] *Suppose $C \subset M$ is a geodesically convex closed set and $\{x_n\}_n \subset C$ converges d–weakly to $x \in M$. Then $x \in C$.*

2.5 General weak gradient flows

Let G be a d–lsc function on a complete metric space (M,d). In this generality there are, as explained in [1], various notions of weak gradient flows c_t for G, emanating from an initial point c_0 in M. A natural approximation scheme (the so-called minimizing movement) for obtaining such a candidate $t \to c_t$ was introduced by De Giorgi [33]. It can be seen as a variational formulation of the (backward) Euler scheme: given

$t \in [0,\infty)$ and a positive integer m, one first defines a discrete version c_t^m of c_t as the m^{th} step in the following (m–dependent) iteration with initial data $c_t^{m,0} = c_0$: given $c_t^{m,j} \in M$, the next step $c_t^{m,j+1}$ is obtained by minimizing on M the functional

$$v \to \frac{1}{2}d(v, c_t^{m,j})^2 + \frac{t}{m}G(v). \tag{11}$$

If such a minimizer always exists then the corresponding minimizing movement c_t is defined as the large m limit of $c_t^m = c_t^{m,m}$, if the limit exists in (M,d). As shown by Mayer [56], if (M,d) is a CAT(0) metric space and G is convex this procedure indeed produces a unique limit c_t with a number of useful properties.

Theorem 2.12 [56, Theorem 1.13] *If (M,d) is* CAT(0)*, G is a d–lsc convex function on (M,d), then for any initial point c_0 with $G(c_0) < \infty$ the corresponding minimizing movement $t \to c_t$ exists and defines a contractive continuous semigroup (which is locally Lipschitz continuous on $[0,\infty)$).*

Moreover, as shown in [56], the curve $t \to c_t$ can be thought of as the curve of steepest descent with respect to G in the sense that

$$-\frac{d}{dt}(G(c_t)) = |(\partial G)(c_t)|\left|\frac{dc_t}{dt}\right|, \qquad \left|\frac{dc_t}{dt}\right| = |(\partial G)(c_t)| \tag{12}$$

for almost every t, where $|(\partial G)(y)|$ is the local upper gradient of G at y and $|dc_t/dt|$ is the metric derivative of $t \to c_t$ at t (in the sense of [1]):

$$|(\partial G)(y)| := \limsup_{z\to y} \frac{(G(y)-G(z))^+}{d(y,z)}, \qquad \left|\frac{dc_t}{dt}\right| = \lim_{s\to t}\left|\frac{d(c_s,c_t)}{s-t}\right|.$$

In the case when (M,d) is a finite-dimensional Riemannian manifold and G is smooth, the relations (12) are equivalent to the usual gradient flow formulation for G. In the terminology of [1] the relations (12) imply that the minimizing movement $t \to c_t$ provided by Mayer's theorem is a curve of *maximal slope* with respect to the upper gradient $|(\partial G)|$ (see [1, Definition 1.3.2]). Moreover, by [1, Theorem 4.0.4] the curve $t \to c_t$ is the unique solution of the *evolution variational inequality*

$$\frac{1}{2}\frac{d}{dt}d^2(c_t, v) \le G(v) - G(c_t) \quad \text{for ae } t > 0 \text{ and all } v \text{ such that } G(v) < \infty \tag{13}$$

among all locally absolutely continuous curves in (M,d) such that $\lim_{t\to 0} c_t = c_0$. Among other things, this inequality shows that

$$\lim_{t\to\infty} G(c_t) = \inf_{y\in M} G(y).$$

Remark 2.13 A necessary condition for the solvability of the minimization steps (11) is to have $G(c_0) < \infty$. An approximation argument using contractivity of the minimizing movement yields that it is possible to uniquely define $t \to c_t$ for any c_0 in the d–closure of the set $\{G < \infty\}$. This slightly more general movement satisfies all the above-mentioned properties and additionally $G(c_t) < \infty$ for any $t > 0$ (for more details see [1]).

Lastly, we recall a theorem of Bačák, central in our later developments:

Theorem 2.14 [3, Theorem 1.5] *Given a* CAT(0) *space* (M, d) *and a* d*–lsc convex function* $G\colon M \to (-\infty, \infty]$*, assume that* G *attains its minimum on* M*. Then any minimizing movement trajectory* $t \to c_t$ *weakly* d*–converges to some minimizer of* G *as* $t \to \infty$*.*

3 Approximation in d_p with convergent entropy

The approximation results in this section will be used in the proof of Theorem 1.2. Our main tools will come from Sections 2.1–2.3. We begin with the simplified situation of approximation in $\mathcal{E}^1$:

Lemma 3.1 *Suppose* f *is usc on* X *with* $e^{-f} \in L^1(X, \omega^n)$*. Given* $u \in \mathcal{E}^1$*, there exists* $u_k \in \mathcal{H}_\omega$ *such that* $d_1(u_k, u) \to 0$ *and* $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_{u_k}) \to \mathrm{Ent}(e^{-f}\omega^n, \omega^n_u)$*.*

Proof If $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_u) = \infty$ then any sequence $u_k \in \mathcal{H}_\omega$ with $d_1(u_k, u) \to 0$ satisfies the requirements, as the entropy is d_1–lsc. Indeed, this follows from the classical fact that the entropy is lsc with respect to weak convergence of measures [35], and d_1–convergence implies weak convergence of the complex Monge–Ampère measures [30].

We can suppose that $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_u) < \infty$. Let $g = \omega^n_u/\omega^n \geq 0$ be the density function of ω^n_u. We will show that there exist positive functions $g_k \in C^\infty(X)$ such that $|g - g_k|_{L^1} \to 0$ and

$$\int_M g_k \log \frac{g_k}{e^{-f}} \omega^n \to \int_M g \log \frac{g}{e^{-f}} \omega^n = \mathrm{Ent}(e^{-f}\omega^n, \omega^n_u).$$

First introduce $h_k = \min\{k, g\}$, $k \in \mathbb{N}$. As $\phi(t) = t \log t$, $t > 0$, is bounded from below by $-e^{-1}$ and increasing for $t > 1$, we get

$$-e^{-1}e^{-f} \leq h_k \log \frac{h_k}{e^{-f}} \leq \max\left\{0, g \log \frac{g}{e^{-f}}\right\}.$$

Clearly $|h_k - g|_{L^1} \to 0$, and as $e^{-f} \in L^1(X, \omega^n)$ and $g \log(g/e^{-f}) \in L^1(X, \omega^n)$, the Lebesgue dominated convergence theorem gives that

$$\int_M h_k \log \frac{h_k}{e^{-f}} \omega^n \to \int_M g \log \frac{g}{e^{-f}} \omega^n = \operatorname{Ent}(e^{-f}\omega^n, \omega_u^n).$$

Using the density of $C^\infty(M)$ in $L^1(M)$, by another application of the dominated convergence theorem, we find a positive sequence $g_k \in C^\infty(X)$ such that $|g_k - h_k|_{L^1} \le 1/k$ and

$$\left| \int_M h_k \log \frac{h_k}{e^{-f}} \omega^n - \int_M g_k \log \frac{g_k}{e^{-f}} \omega^n \right| \le \frac{1}{k}.$$

Using the Calabi–Yau theorem we find potentials $v_k \in \mathcal{H}_\omega$ with $\sup_M v_k = 0$ and $\omega_{v_k}^n = g_k \omega^n / \int_M g_k \omega^n$. Theorem 2.8 now guarantees that (after possibly passing to a subsequence) $d_1(v_k, h) \to 0$ for some $h \in \mathcal{E}^1(X)$. But [30, Theorem 5(i)] implies the equality of measures $\omega_h^n = \omega_u^n$. Finally, by the uniqueness theorem [45, Theorem B] we get that h and u can differ by at most a constant. Hence, after possibly adding a constant, we can suppose that $d_1(v_k, u) \to 0$. □

The key point in this proof is that a bound on the entropy implies compactness in $(\mathcal{E}^1, d_1)$. There are examples showing that the d_2 version of this compactness result does not hold in general. Therefore, to approximate functions in $(\mathcal{E}^p, d_p)$, $p > 1$ with convergent entropy, a new approach is necessary:

Theorem 3.2 *Suppose $\varphi \in \mathcal{E}^p$, $p \ge 1$ and f is usc on X with $e^{-f} \in L^1(X, \omega^n)$. Then there exists $\varphi_j \in \mathcal{H}_\omega$ such that $d_p(\varphi_j, \varphi) \to 0$ and*

$$\operatorname{Ent}(e^{-f}\omega^n, \omega_{\varphi_j}^n) \to \operatorname{Ent}(e^{-f}\omega^n, \omega_\varphi^n).$$

Proof We divide the approximation procedure into three steps.

Step 1 Assume that $u \in \mathcal{E}^p$ has finite twisted entropy $\operatorname{Ent}(e^{-f}\omega^n, \omega_u^n) < \infty$ and

$$\omega_u^n = e^g \omega^n$$

for some measurable function g. We also normalize u so that $\int_X u\omega_u^n = 0$. For each $\varepsilon > 0$ let $u_\varepsilon \in \mathcal{E}^p(X, \omega)$ be the unique solution to

$$\omega_{u_\varepsilon}^n = e^{\varepsilon u_\varepsilon + g} \omega^n.$$

Then we claim that $d_p(u_\varepsilon, u) \to 0$ and $\operatorname{Ent}(e^{-f}\omega^n, \omega_{u_\varepsilon}^n) \to \operatorname{Ent}(e^{-f}\omega^n, \omega_u^n)$ as $\varepsilon \to 0$.

Indeed, from Lemma 2.10, u_ε is uniformly bounded from above for $\varepsilon \in [0,1]$, and converges in d_p to u as $\varepsilon \to 0$. Also, by the comparison principle (Lemma 2.9), $u_\varepsilon \ge u - \sup_X u$. As in the proof of Lemma 3.1 we can show using the dominated convergence theorem that $\operatorname{Ent}(e^{-f}\omega^n, \omega_{u_\varepsilon}^n)$ converges to $\operatorname{Ent}(e^{-f}\omega^n, \omega_u^n)$ as $\varepsilon \to 0$.

Step 2 Let g be a measurable function such that $\int_X e^g \omega^n < \infty$. Assume that $u \in \mathcal{E}^p$ has finite twisted entropy $\operatorname{Ent}(e^{-f}\omega^n, \omega_u^n) < \infty$ and

$$\omega_u^n = e^{\varepsilon u + g}\omega^n$$

for some $\varepsilon > 0$. Consider $g_k := \min(g, k)$, $k \in \mathbb{N}$. Let $u_k \in \operatorname{PSH}(X, \omega) \cap \mathcal{C}^0(X)$ be the unique solution to

$$\omega_{u_k}^n = e^{\varepsilon u_k + g_k}\omega^n.$$

The fact that u_k is continuous follows from Kołodziej's $\mathcal{C}^0$ estimate [52]. By the comparison principle u_k is decreasing and converges to u as $k \to \infty$. It follows from Proposition 2.5 that $d_p(u_k, u) \to 0$ as $k \to \infty$. Again, the proof of Lemma 3.1 shows that $\operatorname{Ent}(e^{-f}\omega^n, \omega_{u_k}^n)$ converges to $\operatorname{Ent}(e^{-f}\omega^n, \omega_u^n)$ as $k \to \infty$.

Step 3 Assume that g is bounded from above, $u \in \mathcal{E}^p$ has finite twisted entropy $\operatorname{Ent}(e^{-f}\omega^n, \omega_u^n) < \infty$ and

$$\omega_u^n = e^{\varepsilon u + g}\omega^n \tag{14}$$

for some $\varepsilon > 0$. Let $\{g_k\}_{k \in \mathbb{N}}$ be a sequence of smooth functions, uniformly bounded above, such that e^{g_k} converges to e^g in $L^2(X, \omega^n)$ and $\int_X e^{g_k}\omega^n = \int_X e^g \omega^n$. Let $u_k \in \mathcal{H}_\omega$ be the unique smooth solution to

$$\omega_{u_k}^n = e^{\varepsilon u_k + g_k}\omega^n. \tag{15}$$

The fact that u_k is smooth on X is well known (see [2] or [64; 47, Chapter 14] for other proofs).

We claim that $\sup_X u_k$ is bounded above. Indeed, by an argument similar to that of Lemma 2.10, suppose that for some subsequence (again denoted by u_k) we have that $\sup_X u_k \to \infty$. Then a subsequence of $v_k := u_k - \sup_X u_k$ (again denoted by v_k) L^1–converges to some $v \in \operatorname{PSH}(X, \omega)$. As all L^p topologies are equivalent on $\operatorname{PSH}(X, \omega)$, we actually have $v_k \to_{L^p} v$ for any $p \geq 1$. However, using Jensen's inequality and (15), we obtain that

$$\int_X u_k e^{g_k}\omega^n = \int_X v_k e^{g_k}\omega^n + \sup_X u_k \int_X e^g \omega^n$$

is uniformly bounded above. We have that $v_k \to_{L^2} v$ and $e^{g_k} \to_{L^2} e^g$, hence using Hölder's inequality we obtain $\int_X v_k e^{g_k}\omega^n \to \int_X v e^g \omega^n \neq \infty$. As $\sup_X u_k \to \infty$ we arrive at a contradiction with the upper bound on $\int_X u_k e^{g_k}\omega^n$, finishing the proof of the claim.

As $\sup_X u_k$ is bounded above, using Kołodziej's estimates [53] for (15), we obtain a uniform upper bound on $\|u_k\|_{C^{0,\alpha}}$ for some $\alpha > 0$. After perhaps choosing a subsequence, $\{u_k\}_k$ will converge uniformly to some $v \in C^{0,\alpha} \cap \mathrm{PSH}(X,\omega)$, ultimately giving $d_p(u_k, v) \to 0$. As both the left- and right-hand sides of (15) converge, we get that $\omega_v^n = e^{\varepsilon v+g}\omega^n$, hence by uniqueness of solutions to (14) (Lemma 2.9) we get $v = u$. By a repeated use of the dominated convergence theorem, the corresponding twisted entropies also converge.

Now, we come back to the proof of Theorem 3.2. If $\mathrm{Ent}(e^{-f}\omega^n, \omega_\varphi^n) = \infty$ then any decreasing sequence $\varphi_j \in \mathcal{H}_\omega$ which converges pointwise to φ satisfies our requirement since the entropy is lsc with respect to weak convergence of measures. We can thus assume that $\mathrm{Ent}(e^{-f}\omega^n, \omega_\varphi^n) < \infty$. Then we can write $\omega_\varphi^n = e^g\omega^n$. We can also assume that $\int_X \varphi\omega_\varphi^n = 0$. Fix $\delta > 0$ arbitrarily small. Denoting $\varphi_0 = \varphi$, by the three steps above we can find $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{E}^p$, with $\varphi_3 \in \mathcal{H}_\omega$, such that

$$d_p(\varphi_j, \varphi_{j+1}) \le \delta \quad \text{and} \quad \left|\mathrm{Ent}(e^{-f}\omega^n, \omega_{\varphi_j}^n) - \mathrm{Ent}(e^{-f}\omega^n, \omega_{\varphi_{j+1}}^n)\right| \le \delta, \quad j = 0, 1, 2.$$

From this the result follows. □

4 Extension of the twisted K-energy

The main goal of this section is to prove Theorem 1.2. Before we can attempt a proof, we need to understand the d_1–continuity properties of each functional appearing in the right-hand side of (4). Some of the preliminary results below are well known, but as a courtesy to the reader we give a detailed account.

4.1 The AM functional

The Aubin–Mabuchi functional is given by the formula (see [55, Theorem 2.3])

$$\mathrm{AM}(u) := \frac{V^{-1}}{n+1} \sum_{j=0}^{n} \int_X u\,\omega^j \wedge \omega_u^{n-j}, \quad u \in \mathcal{H}_\omega. \tag{16}$$

A series of integrations by parts gives

$$\mathrm{AM}(v) - \mathrm{AM}(u) = \frac{V^{-1}}{n+1} \int_X (v-u) \sum_{k=0}^{n} \omega_u^{n-k} \wedge \omega_v^k, \quad u, v \in \mathcal{H}_\omega. \tag{17}$$

Among other things, this formula shows that

$$u \le v \implies \mathrm{AM}(u) \le \mathrm{AM}(v),$$

and by computing $\lim_{t\to 0}(\mathrm{AM}(v_t) - \mathrm{AM}(v))/t$ we arrive at the first-order variation of AM:

$$\langle D\,\mathrm{AM}(v), \delta v\rangle = V^{-1} \int_X \delta v \omega_v^n, \quad v \in \mathcal{H}_\omega, \delta v \in C^\infty(X). \tag{18}$$

Suppose $u \in \mathcal{E}^1$ and let $u_j \in \mathcal{H}_\omega$ be pointwise decreasing to u. Using Proposition 2.5 we have $d_1(u, u_j) \to 0$. We hope to extend AM to $\mathcal{E}^1$ via

$$\mathrm{AM}(u) = \lim_j \mathrm{AM}(u_j). \tag{19}$$

As it turns out, this choice of extension is justified by the following precise result:

Proposition 4.1 *The map* AM: $\mathcal{H}_\omega \to \mathbb{R}$ *is* d_1*–Lipschitz continuous. Thus,* (19) *gives* d_1*–Lipschitz extension of* AM *to* $\mathcal{E}^1$*.*

Proof First we argue that $|\mathrm{AM}(u_0) - \mathrm{AM}(u_1)| \le d_1(u_0, u_1)$ for $u_0, u_1 \in \mathcal{H}_\omega$. Let $[0,1] \ni t \to \gamma_t \in \mathcal{H}_\omega$ be a smooth curve connecting u_0 and u_1. By (18) we can write

$$|\mathrm{AM}(u_1) - \mathrm{AM}(u_0)| = \left| V^{-1} \int_0^1 \int_X \dot\gamma_t \omega^n_{\gamma_t}\, dt \right| \le V^{-1} \int_0^1 \int_X |\dot\gamma_t| \omega^n_{\gamma_t}\, dt = l(\gamma).$$

Taking the infimum over all smooth curves connecting u_0 and u_1, we obtain that

$$|\mathrm{AM}(u_1) - \mathrm{AM}(u_0)| \le d_1(u_0, u_1).$$

The density of $\mathcal{H}_\omega$ in $\mathcal{E}^1$ implies that AM extends to $\mathcal{E}^1$ using the formula (19). The extension has to be d_1–Lipschitz continuous. □

Before we proceed, we note that the "abstract" d_1–continuous extension AM: $\mathcal{E}^1 \to \mathbb{R}$ given by the above result is the same as the "concrete" one given by the expression of (16) after replacing the smooth products $\omega^j \wedge \omega_u^{n-j}$ with the nonpluripolar products from (7), as done in [17]. Moving on, we give a kind of "domination principle" for the extended Aubin–Mabuchi energy on $\mathcal{E}^1$:

Proposition 4.2 *Suppose* $\phi, \psi \in \mathcal{E}^1$ *with* $\phi \ge \psi$*. If* $\mathrm{AM}(\phi) = \mathrm{AM}(\psi)$*, then* $\phi = \psi$*.*

Proof Suppose $\phi_k, \psi_k \in \mathcal{H}_\omega$ are sequences pointwise decreasing to ϕ and ψ, respectively, with $\phi_k \ge \psi_k$. Then (17) gives that

$$0 \le \frac{1}{(n+1)V} \int_X (\phi_k - \psi_k) \omega^n_{\psi_k} \le \mathrm{AM}(\phi_k) - \mathrm{AM}(\psi_k).$$

Using the previous proposition and [30, Lemma 5.2] with $\chi(t) = |t|$, $v_k = \phi_k$, $u_k = \psi_k$, $w_k = \psi_k$, we may take the limit in this estimate to obtain

$$0 \le \frac{1}{(n+1)V} \int_X (\phi - \psi) \omega^n_{\psi} \le \mathrm{AM}(\phi) - \mathrm{AM}(\psi) = 0,$$

hence $\psi \ge \phi$ ae with respect to ω^n_ψ. The domination principle of the class $\mathcal{E}$ [16, Proposition 5.9] gives now that $\psi \ge \phi$ globally on X, hence $\psi = \phi$. □

The last result of this subsection points out that the family of finite-energy geodesics inside $\mathcal{E}^p$ is in fact "endpoint-stable". We note that in the case $p = 2$ this follows from the fact that $(\mathcal{E}^2, d_2)$ is CAT(0) [18].

Proposition 4.3 *Suppose $[0,1] \ni t \to u^j_t \in \mathcal{E}^p$ is a sequence of finite-energy geodesic segments such that $d_p(u^j_0, u_0), d_p(u^j_1, u_1) \to 0$. Then $d_p(u^j_t, u_t) \to 0$ for all $t \in [0,1]$, where $[0,1] \ni t \to u_t \in \mathcal{E}^p$ is the finite-energy geodesic segment connecting u_0 and u_1.*

Proof Let $t \in [0,1]$. Notice that we only have to show that any subsequence of $\{u^j_t\}_j$ contains a subsubsequence d_p–converging to u_t.

Let $\{u^{j_k}_t\}_k$ be an arbitrary subsequence of $\{u^j_t\}_j$. Let j_{k_l} be a subsequence of j_k with the following property: for $i = 0, 1$, there exist a monotone increasing sequence $\{v^{j_{k_l}}_i\}_l$ and a monotone decreasing sequence $\{w^{j_{k_l}}_i\}_l$ such that

$$v^{j_{k_l}}_i \le u^{j_{k_l}}_i \le w^{j_{k_l}}_i \quad \text{for all } j_{k_l} \qquad \text{and} \qquad v^{j_{k_l}}_i, w^{j_{k_l}}_i \to_{d_p} u_i.$$

This is possible to arrange according to Proposition 2.6.

By $[0,1] \ni t \to v^{j_{k_l}}_t \in \mathcal{E}^p$ and $[0,1] \ni t \to w^{j_{k_l}}_t \in \mathcal{E}^p$ we denote finite-energy geodesics connecting $v^{j_{k_l}}_0$ to $v^{j_{k_l}}_1$ and $w^{j_{k_l}}_0$ to $w^{j_{k_l}}_1$, respectively. By the maximum principle of finite-energy geodesics we can write

$$v_t := \operatorname{usc}\Big(\lim_l v^{j_{k_l}}_t\Big) \le u_t \le w_t := \lim_l w^{j_{k_l}}_t.$$

As AM is d_p–continuous it follows that

$$\lim_l \mathrm{AM}(v^{j_{k_l}}_i) = \mathrm{AM}(u_i) = \lim_l \mathrm{AM}(w^{j_{k_l}}_i) \quad \text{for } i = 0, 1.$$

As AM is also linear along finite-energy geodesics we get

$$\mathrm{AM}(v_t) = \mathrm{AM}(u_t) = \mathrm{AM}(w_t) \quad \text{for any } t \in [0,1].$$

Proposition 4.2 gives that $v_t = u_t = w_t$, hence

$$d_p(v^{j_{k_l}}_t, u_t) \to 0 \quad \text{and} \quad d_p(w^{j_{k_l}}_t, u_t) \to 0.$$

Using $v^{j_{k_l}}_t \le u^{j_{k_l}}_t \le w^{j_{k_l}}_t$, [30, Lemma 4.2] gives that

$$d_p(v^{j_{k_l}}_t, u^{j_{k_l}}_t) \le d_p(v^{j_{k_l}}_t, w^{j_{k_l}}_t) \to 0,$$

hence $d_p(u^{j_{k_l}}_t, u_t) \to 0$, as desired. □

4.2 The AM_γ functional

For the moment we fix a closed $(1,1)$–current γ on X, not necessarily positive. Recall from the introduction that the functional AM_γ is defined as follows:

$$\mathrm{AM}_\gamma(u) := \frac{1}{nV}\sum_{j=0}^{n-1}\int_X u\,\gamma\wedge\omega^j\wedge\omega_u^{n-1-j}, \quad u\in\mathcal{H}_\omega. \tag{20}$$

Similarly to AM, integrating by parts gives

$$\mathrm{AM}_\gamma(v)-\mathrm{AM}_\gamma(u) = \frac{1}{nV}\int_X (v-u)\sum_{k=0}^{n-1}\gamma\wedge\omega_u^{n-k-1}\wedge\omega_v^k. \tag{21}$$

When $\gamma\geq 0$ this last formula gives

$$u\leq v \implies \mathrm{AM}_\gamma(u)\leq\mathrm{AM}_\gamma(v).$$

By computing $\lim_{t\to 0}(\mathrm{AM}_\gamma(v_t)-\mathrm{AM}_\gamma(v))/t$ we arrive at the first-order variation of AM_γ:

$$\langle D\,\mathrm{AM}_\gamma(v),\delta v\rangle = V^{-1}\int_X \delta v\gamma\wedge\omega_v^{n-1}, \quad v\in\mathcal{H}_\omega, \delta v\in C^\infty(X). \tag{22}$$

Extension of AM_γ to $\mathcal{E}^1$ when γ is smooth For this paragraph suppose γ is smooth. Suppose $u\in\mathcal{E}^1$ and let $u_j\in\mathcal{H}_\omega$ be pointwise decreasing to u. Using Proposition 2.5 we have $d_1(u,u_j)\to 0$. We hope to extend AM_γ to $\mathcal{E}^1$ via

$$\mathrm{AM}_\gamma(u) = \lim_j \mathrm{AM}_\gamma(u_j). \tag{23}$$

As it turns out, this extension is rigorous as we have the following precise result:

Proposition 4.4 *Formula (23) gives a d_1–continuous functional $\mathrm{AM}_\gamma\colon \mathcal{E}^1\to\mathbb{R}$. Additionally, AM_γ thus extended is bounded on d_1–bounded subsets of $\mathcal{E}^1$.*

Proof We argue that for any $R>0$ there exists $f_R\colon\mathbb{R}\to\mathbb{R}$ continuous with $f_R(0)=0$ such that

$$|\mathrm{AM}_\gamma(u_0)-\mathrm{AM}_\gamma(u_1)|\leq f_R(d_1(u_0,u_1)) \tag{24}$$

for any $u_0,u_1\in\mathcal{H}_\omega\cap\{v: d_1(0,v)\leq R\}$. We have $-C\omega\leq\gamma\leq C\omega$ for some $C>1$. Using (21) and the observation $\omega_{(u_0+u_1)/4}=\frac{1}{2}\omega+\frac{1}{4}\omega_{u_0}+\frac{1}{4}\omega_{u_1}$ it follows that

$$|\mathrm{AM}_\gamma(u_0)-\mathrm{AM}_\gamma(u_1)|\leq C\int_X |u_0-u_1|\omega^n_{(u_0+u_1)/4}.$$

By [30, Corollary 5.7] and its proof, for each $R > 0$ there exists a continuous function $f_R\colon \mathbb{R} \to \mathbb{R}$ with $f_R(0) = 0$ such that

$$\int_X |v - w|\omega_h^n \leq f_R(d_1(v, w))$$

for any $v, w, h \in \mathcal{E}^1 \cap \{v : d_1(0, v) \leq R\}$. Using this last fact, to argue that (24) holds, it is enough to show that $d_1\big(0, \frac{1}{4}(u_0 + u_1)\big)$ is bounded in terms of $d_1(0, u_0)$ and $d_1(0, u_1)$. We recall [30, Lemma 5.3], which says that there exists $D > 1$ such that $d_1\big(a, \frac{1}{2}(a+b)\big) \leq D d_1(a, b)$ for any $a, b \in \mathcal{E}^1$. Using this several times along with the triangle inequality, we can write

$$\begin{aligned} d_1\big(0, \tfrac{1}{4}(u_0 + u_1)\big) &\leq C d_1\big(0, \tfrac{1}{2}(u_0 + u_1)\big) \leq C\big(d_1(0, u_0) + d_1\big(u_0, \tfrac{1}{2}(u_0 + u_1)\big)\big) \\ &\leq C^2(d_1(0, u_0) + d_1(u_0, u_1)) \leq 2C^2(d_1(0, u_0) + d_1(0, u_1)), \end{aligned}$$

finishing the proof. □

As in the case of AM, the “abstract” d_1–continuous extension $\mathrm{AM}_\gamma\colon \mathcal{E}^1 \to \mathbb{R}$ given by the above result is identical to the one given by the “concrete” expression of (20) after replacing the smooth products $\gamma \wedge \omega^j \wedge \omega_u^{n-j-1}$ with nonpluripolar products similar to (7).

Convexity and extension of AM_χ to $\mathcal{H}_\omega^\Delta$ when χ satisfies (3) Suppose that $\chi = \beta + i\partial\bar\partial f$ is a $(1,1)$–current satisfying (3). Observe that it is not possible to extend AM_χ to $\mathcal{H}_\omega^\Delta$ using the techniques of the previous paragraph directly. Instead, using integration by parts, we notice that, given $u \in \mathcal{H}_\omega$, we have an alternative formula for $\mathrm{AM}_\chi(u)$:

$$\begin{aligned} \mathrm{AM}_\chi(u) &= \frac{1}{nV} \sum_{j=0}^{n-1} \int_X u\, \beta \wedge \omega^j \wedge \omega_u^{n-1-j} + \frac{1}{nV}\int_X f(\omega_u^n - \omega^n) \\ &= \mathrm{AM}_\beta(u) + \frac{1}{nV}\int_X f(\omega_u^n - \omega^n). \end{aligned} \tag{25}$$

As β is smooth, AM_β extends d_1–continuously to $\mathcal{H}_\omega^\Delta$ by the previous paragraph. The map $u \to \int_X f\omega_u^n$ clearly makes sense and is finite for all $u \in \mathcal{H}_\omega^\Delta$, hence using (25) it is possible to extend AM_χ to $\mathcal{H}_\omega^\Delta$. Though not needed, it can be further shown that this extension is independent of the choice of β and f.

Given $u_0, u_1 \in \mathcal{H}_\omega$, for the weak geodesic $[0,1] \ni t \to u_t \in \mathcal{H}_\omega^\Delta$ connecting u_0 and u_1 we would like to show that $t \to \mathrm{AM}_\chi(t)$ is convex. When χ is smooth this follows from the result of [21]. It turns out that for more general χ the same proof gives an analogous result:

Proposition 4.5 *Suppose $\chi = \beta + i\partial\bar\partial f \geq 0$ satisfies (3). Equation (25) gives an extension $\mathrm{AM}_\chi\colon \mathcal{H}_\omega^\Delta \to \mathbb{R}$ for which $t \to \mathrm{AM}_\chi(u_t)$ is convex for any weak geodesic segment $[0,1] \ni t \to u_t \in \mathcal{H}_\omega^\Delta$.*

Proof Suppose $t_1 \geq t_0$. When χ is smooth, it is well known that for $[0,1] \ni t \to v_t \in \mathcal{H}_\omega$ smooth subgeodesic (ie $\pi^*\omega + i\partial\bar\partial v \geq 0$) we actually have

$$\frac{d}{dt}\Big|_{t=t_1} \mathrm{AM}_\chi(v_t) - \frac{d}{dt}\Big|_{t=t_0} \mathrm{AM}_\chi(v_t) = \int_{S_{t_0,t_1}\times X} \pi^*\chi \wedge (\pi^*\omega + i\partial\bar\partial v)^n,$$

where $S_{t_0,t_1} \subset \mathbb{C}$ is the strip $\{t_0 \leq \operatorname{Re} z \leq t_1\}$. Hence, $t \to \mathrm{AM}_\chi(v_t)$ is convex. We claim that the same proof goes through for any positive closed current $\chi = \beta + i\partial\bar\partial f$ as well.

When dealing with a weak geodesic $[0,1] \ni t \to u_t \in \mathcal{H}_\omega^\Delta$, it is possible to approximate it uniformly with a decreasing sequence of smooth subgeodesics $t \to u_t^\varepsilon$ called ε–geodesics (see [20]). All measures $\omega^n_{u_t^\varepsilon} = g_t^\varepsilon \omega^n$ have uniformly bounded density g_t^ε, and converge weakly to $\omega^n_{u_t}$. Hence, by the dominated convergence theorem we can write

$$\lim_{\varepsilon\to 0} \int_X f\omega^n_{u_t^\varepsilon} = \int_X f\omega^n_{u_t} \quad \text{and} \quad \lim_{\varepsilon\to 0} \mathrm{AM}_\beta(u_t^\varepsilon) = \mathrm{AM}_\beta(u_t),$$

where in the last limit we have used the continuity property of the mixed Monge–Ampère operator (see [4; 5] for the original statement and [45] for the corresponding theory on compact Kähler manifolds). Hence, after repeatedly taking limit in (25), it follows that $t \to \lim_{\varepsilon\to 0} \mathrm{AM}_\chi(u_t^\varepsilon) = \mathrm{AM}_\chi(u_t)$ is convex. □

Finally, we note the following useful inequality for AM_γ.

Lemma 4.6 *Let $\psi \in \mathcal{E}^1$ and set $\theta = \omega_\psi$. For any $u, v \in \mathcal{E}^1$ we have*

$$\frac{1}{V}\int_X (u-v)\omega_u^{n-1} \wedge \theta \leq \mathrm{AM}_\theta(u) - \mathrm{AM}_\theta(v) \leq \frac{1}{V}\int_X (u-v)\omega_v^{n-1} \wedge \theta.$$

For AM *we have similar inequalities*

$$\frac{1}{V}\int_X (u-v)\omega_u^n \leq \mathrm{AM}(u) - \mathrm{AM}(v) \leq \frac{1}{V}\int_X (u-v)\omega_v^n.$$

Proof Using (17) and (21) the desired inequalities simply follow from the fact that

$$\int_X (u-v) i\partial\bar\partial(u-v) \wedge T \leq 0$$

for any $T = \omega_{\varphi_1} \wedge \cdots \wedge \omega_{\varphi_{n-1}}$ with $\varphi_j \in \mathcal{E}^1$ for all j. □

4.3 The twisted K-energy

For the remainder of the paper suppose $\chi = \beta + i\partial\bar\partial f$ satisfies (3) unless specified otherwise. Recall that the twisted K-energy $\mathcal{K}_\chi\colon \mathcal{H}_\omega \to \mathbb{R}$ is defined as

$$\mathcal{K}_\chi = \mathrm{Ent}(e^{-f}\omega^n, \omega_u^n) + \bar{S}_\chi \,\mathrm{AM}(u) - n\,\mathrm{AM}_{\mathrm{Ric}\,\omega - \beta}(u) - \int_X f\omega^n.$$

When f is smooth, recall the following formula for the variation of the entropy:

$$\langle D\,\mathrm{Ent}(e^{-f}\omega^n, \omega_v^n), \delta v\rangle = nV^{-1}\int_X \delta v(\mathrm{Ric}\,\omega - \mathrm{Ric}\,\omega_v + i\partial\bar\partial f)\wedge\omega_v^{n-1}.$$

When χ is smooth, putting the above formula, (18) and (22) together we obtain

$$\begin{aligned}\langle D\mathcal{K}_\chi(v), \delta v\rangle &= \frac{n}{V}\int_X \delta v(\bar{S}_\chi\omega_v - \mathrm{Ric}\,\omega_v + \chi)\wedge\omega_v^{n-1}\\ &= V^{-1}\int_X \delta v(\bar{S}_\chi - S_{\omega_v} + \mathrm{Tr}^{\omega_u}\chi)\omega_v^n.\end{aligned}$$

We arrive at the main theorem of this section:

Theorem 4.7 *Suppose (X,ω) is a compact connected Kähler manifold and $\chi = \beta + i\partial\bar\partial f$ satisfies (3). The twisted K-energy can be extended to a functional $\mathcal{K}_\chi\colon \mathcal{E}^1 \to \mathbb{R}\cup\{\infty\}$ using the formula*

$$\mathcal{K}_\chi(u) = \mathrm{Ent}(e^{-f}\omega^n, \omega_u^n) + \bar{S}_\chi\,\mathrm{AM}(u) - n\,\mathrm{AM}_{\mathrm{Ric}\,\omega-\beta}(u) - \int_X f\omega^n. \tag{26}$$

Thus extended, $\mathcal{K}_\chi|_{\mathcal{E}^p}$ is the greatest d_p–lsc extension of $\mathcal{K}_\chi|_{\mathcal{H}_\omega}$ for any $p \geq 1$. Additionally, $\mathcal{K}_\chi|_{\mathcal{E}^p}$ is convex along the finite-energy geodesics of $\mathcal{E}^p$.

Proof First we argue that the expression given by (26) does give a d_1–lsc function on $\mathcal{E}^1$. Indeed, by Propositions 4.1 and 4.4 the functionals AM and $\mathrm{AM}_{\mathrm{Ric}\,\omega-\beta}$ admit a d_1–continuous extension to $\mathcal{E}^1$. Lastly, as d_1–convergence of potentials implies weak convergence of the corresponding complex Monge–Ampère measures, it follows that the correspondence $u \to \mathrm{Ent}(e^{-f}\omega^n, \omega_u^n)$ is d_1–lsc. When restricted to $\mathcal{E}^p$, (26) is additionally d_p–lsc, because d_p–convergence dominates d_1–convergence for any $p > 1$.

We now show that, thus extended, $\mathcal{K}_\chi|_{\mathcal{E}^p}$ is indeed the greatest d_p–lsc extension of $\mathcal{K}_\chi|_{\mathcal{H}_\omega}$. For this we only have to argue that for any $u \in \mathcal{E}^p$ there exists $\{u_j\} \subset \mathcal{H}_\omega$ such that $d_p(u_j, u) \to 0$ and

$$\mathcal{K}_\chi(u) = \lim_j \mathcal{K}_\chi(u_j).$$

As $\mathrm{AM}(\cdot)$, $\mathrm{AM}_{\mathrm{Ric}\,-\beta}(\cdot)$ are d_p–continuous, this is exactly the content of Theorem 3.2.

Since finite-energy geodesics of $\mathcal{E}^p$ are also finite-energy geodesics in $\mathcal{E}^1$, it remains to show that for any finite-energy geodesic $[0,1] \ni t \to u_t \in \mathcal{E}^1$ the curve $t \to \mathcal{K}_\chi(u_t)$ is convex and continuous.

Suppose $t_0, t_1 \in [0,1]$ with $t_0 \le t_1$. As $\mathcal{K}_\chi$ was extended in the greatest d_1–lsc manner, we can find $u^k_{t_0}, u^k_{t_0} \in \mathcal{H}_\omega$ with $d_1(u^k_{t_0}, u_{t_0}) \to 0$, $d_1(u^k_{t_1}, u_{t_1}) \to 0$ and

$$\mathcal{K}_\chi(u_{t_0}) = \lim_k \mathcal{K}_\chi(u^k_{t_0}), \quad \mathcal{K}_\chi(u_{t_1}) = \lim_k \mathcal{K}_\chi(u^k_{t_1}).$$

Let $[t_0, t_1] \ni t \to u^k_t \in \mathcal{H}^\Delta_\omega$ be the weak geodesics connecting $u^k_{t_0}$ and $u^k_{t_1}$. By Proposition 4.3 we get that $d_1(u^k_t, u_t) \to 0$ for any $t \in [t_0, t_1]$. Note that for $u \in \mathcal{H}^\Delta_\omega$ we can write

$$\mathcal{K}_\chi(u) = \mathrm{Ent}(\omega^n, \omega^n_u) + \bar{S}_\chi \,\mathrm{AM}(u) - n\,\mathrm{AM}_{\mathrm{Ric}\,\omega}(u) + \left(n\,\mathrm{AM}_\beta(u) + \frac{1}{V}\int_X f\omega^n_u\right).$$

Using this, Proposition 4.5, [8, Theorem 1.1] and the linearity of AM along finite-energy geodesics, it follows that $t \to \mathcal{K}_\chi(u^k_t)$ is convex on $[0,1]$. As $\mathcal{K}_\chi\colon \mathcal{E}^1 \to \mathbb{R} \cup \{\infty\}$ is d_1–lsc, it follows that

$$\begin{aligned}\mathcal{K}_\chi(u_t) \le \liminf_k \mathcal{K}_\chi(u^k_t) &\le \frac{t-t_0}{t_1-t_0}\lim_k \mathcal{K}_\chi(u^k_{t_0}) + \frac{t_1-t}{t_1-t_0}\lim_k \mathcal{K}_\chi(u^k_{t_1}) \\ &\le \frac{t-t_0}{t_1-t_0}\mathcal{K}_\chi(u_{t_0}) + \frac{t_1-t}{t_1-t_0}\mathcal{K}_\chi(u_{t_1}),\end{aligned}$$

hence $[0,1] \ni t \to \mathcal{K}_\chi(u_t) \in (-\infty, \infty]$ is convex. As $\mathcal{K}_\chi$ is d_1–lsc it follows additionally that $t \to \mathcal{K}_\chi(u_t)$ is continuous up to the boundary of $[0,1]$. □

Finally, we bring Theorem 2.8 into a form that will be most convenient to use in our later developments:

Corollary 4.8 *Suppose $\chi = \beta + i\partial\bar\partial f$ satisfies (3) and $\{u_k\}_k \subset \mathcal{E}^1$ is a sequence for which*

$$d_1(0, u_k) < C, \quad \mathcal{K}_\chi(u_k) < C.$$

Then $\{u_k\}_k$ contains a d_1–convergent subsequence.

Proof By (8) it follows that $|\sup_X u_k| < C$. From (26) and Propositions 4.1 and 4.4 we get that $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_{u_k})$ is also uniformly bounded. Now we can invoke Theorem 2.8 to finish the argument. □

4.4 Convexity in the finite entropy space.

Suppose $\chi = \beta + i\partial\bar\partial f$ satisfies (3). Denote by $\mathrm{Ent}_\chi(X,\omega)$ the space of finite-entropy potentials:

$$\mathrm{Ent}_\chi(X,\omega) = \{u \in \mathcal{E}(X,\omega) : \mathrm{Ent}(e^{-f}\omega^n, \omega^n_u) < \infty\}.$$

Observe that $\mathrm{Ent}_\chi(X,\omega)$ is independent of the choice of β and f. Also, we show that $\mathrm{Ent}_\chi(X,\omega)$ is contained in the finite-energy space $\mathcal{E}^1$:

Lemma 4.9 *Suppose* $\chi=\beta+i\partial\bar\partial f$ *satisfies* (3). *Then* $\mathrm{Ent}_\chi(X,\omega)\subset\mathcal{E}^1$.

Proof Suppose $u\in\mathrm{Ent}_\chi(X,\omega)$ with $\omega_u^n=h\omega^n$. The functions $\phi,\psi\colon[0,\infty)\to[0,\infty)$ given by $\phi(t)=(t+1)\log(t+1)-t$ and $\psi(t)=e^t-t-1$ are convex conjugates of each other, implying that $ab\le\phi(a)+\psi(b)$. Using this, we can write

$$\begin{aligned}\int_X |u|\omega_u^n &= \int_X |u|(he^f)e^{-f}\omega^n\\ &\le \int_X (e^{|u|}-|u|-1)e^{-f}\omega^n+\int_X((he^f+1)\log(he^f+1)-h)e^{-f}\omega^n.\end{aligned}$$

To finish the proof it is enough to argue that both terms in this last expression are bounded. For the first term, suppose $1/p+1/q=1$. Using Young's inequality, we arrive at

$$\int_X(e^{|u|}-|u|-1)e^{-f}\omega^n\le\frac{1}{q}\int_X(e^{|u|}-|u|-1)^q\omega^n+\frac{1}{p}\int_X e^{-pf}\omega^n.$$

As u has zero Lelong numbers [45, Corollary 1.8], the first integral is finite by Skoda's theorem. For an appropriate p the second integral is bounded, as $e^{-f}\in L^p(X,\omega^n)$ for some $p>1$.

For the second term, observe that $\phi(t)\le 2t\log t$ for t big enough, hence we can write

$$\begin{aligned}\int_X((he^f+1)\log(he^f+1)-h)e^{-f}\omega^n &\le 2\int_X h\log(he^f)\omega^n+C\\ &= 2V\,\mathrm{Ent}(e^{-f}\omega^n,\omega_u^n)+C.\end{aligned}$$

□

As a consequence of Theorem 4.7 we obtain that $\mathrm{Ent}_\chi(X,\omega)\subset\mathcal{E}^1$ is to some extent "geodesically convex":

Theorem 4.10 *Suppose* $\chi=\beta+i\partial\bar\partial f$ *satisfies* (3). *Then* $(\mathrm{Ent}_\chi(X,\omega),d_1)$ *is a geodesic sub-metric space of* $(\mathcal{E}^1(X,\omega),d_1)$. *Additionally, if* $\mathrm{Ric}\,\omega\ge\beta$ *then the map* $\mathrm{Ent}_\chi(X,\omega)\ni u\to\mathrm{Ent}(e^{-f}\omega^n,\omega_u^n)\in\mathbb{R}$ *is convex along finite-energy geodesics.*

Proof Suppose $u_0,u_1\in\mathrm{Ent}_\chi(X,\omega)$. Let $[0,1]\ni t\to u_t\in\mathcal{E}^1$ be the finite-energy geodesic connecting u_0 and u_1. By Theorem 4.7 it follows that $t\to\mathcal{K}_\chi(u_t)$ is convex on $[0,1]$, hence $\mathcal{K}_\chi(u_t)$ is finite for all $t\in[0,1]$. Using the finiteness of AM and $\mathrm{AM}_{\mathrm{Ric}\,\omega-\beta}$, this necessarily gives that $\mathrm{Ent}(e^{-f}\omega^n,\omega_{u_t}^n)$ is also finite for all $t\in[0,1]$.

For the last statement, notice that $t\to n\,\mathrm{AM}_{\mathrm{Ric}\,\omega-\beta}(u_t)-\bar{S}_\chi\,\mathrm{AM}(u_t)$ is convex, as follows from Proposition 4.5. As $t\to\mathcal{K}_\chi(u_t)$ is also convex, from (26) it follows that $t\to\mathrm{Ent}(e^{-f}\omega^n,\omega_{u_t}^n)$ is also convex. □

In the case $\beta = 0$, this convexity result can be seen as the complex version of one of the central results of the theory of optimal transport of measure, which says that, if g_0 is a given Riemannian metric on a compact real manifold X with nonnegative Ricci curvature and whose normalized volume form is denoted by μ_0, then the relative entropy function $\mu \to \mathrm{Ent}(\mu_0, \mu)$ is convex along curves $t \to \mu_t$ defined by McCann's displacement interpolation (which may be formulated in terms of optimal transport maps). The latter curves can be seen as weak geodesics for Otto's Riemannian metric on the space of all normalized volume forms on X. More precisely, the curves $t \to \mu_t$ are the geodesics in the metric space $(\mathcal{P}(M), d_{W_2})$ defined by the space $\mathcal{P}(M)$ of all probability measures on X equipped with the Wasserstein 2–metric, which can be viewed as a completion of Otto's Riemannian structure [66]. Hence, the role of Otto's Riemannian metric is in the present complex setting played by Mabuchi's Riemannian metric.

4.5 Uniqueness of twisted K-energy minimizers

In this subsection we suppose χ is a Kähler form. We are going to prove that there is at most one minimizer in $\mathcal{E}^1$ of the twisted K-energy $\mathcal{K}_\chi$. We need the following result, which may be of independent interest.

Lemma 4.11 *Let $\varphi_0, \varphi_1 \in \mathcal{E}^1$ and let $[0,1] \ni t \to \varphi_t$ be the finite-energy geodesic connecting φ_0 and φ_1. Suppose that $\omega^n_{\varphi_t}$ is absolutely continuous with respect to ω^n for every $t \in [0,1]$. Then for almost every $t \in (0,1)$ we have*

$$\mathrm{AM}(\varphi_1) - \mathrm{AM}(\varphi_0) = \frac{1}{V}\int_X \dot\varphi_t^+ \omega^n_{\varphi_t} = \frac{1}{V}\int_X \dot\varphi_t^- \omega^n_{\varphi_t}, \tag{27}$$

where, for fixed $x \in X$, $\dot\varphi_t^+(x)$ and $\dot\varphi_t^-(x)$ are the right and left derivatives of $\varphi(\cdot, x)$, respectively.

Proof For simplicity we assume that $V = 1$. Fix two real numbers a, b such that $0 < a < b < 1$. We first observe that for $t \in (a,b)$ and $h > 0$ small enough, by convexity we have

$$\frac{\varphi_t - \varphi_0}{t} \le \frac{\varphi_{t+h} - \varphi_t}{h} \le \frac{\varphi_1 - \varphi_t}{1-t}.$$

It thus follows that both $\dot\varphi_t^+$ and $\dot\varphi_t^-$ are integrable with respect to $\omega^n_{\varphi_t}$. From Lemma 4.6 we obtain

$$\mathrm{AM}(\varphi_{t+h}) - \mathrm{AM}(\varphi_t) \le \int_X (\varphi_{t+h} - \varphi_t)\omega^n_{\varphi_t}.$$

Since AM is linear along the weak geodesic φ_t, by dividing the above inequality by h and letting $h \to 0$ we obtain

$$\int_X \dot\varphi_t^- \omega^n_{\varphi_t} \le \mathrm{AM}(\varphi_1) - \mathrm{AM}(\varphi_0) \le \int_X \dot\varphi_t^+ \omega^n_{\varphi_t}. \tag{28}$$

For each $x \in X$ the function $t \to \varphi_t(x)$ is convex, hence differentiable almost everywhere in $[0,1]$. It follows that the set

$$\{(x,t) \in X \times [a,b] : \dot\varphi_t^-(x) < \dot\varphi_t^+(x)\}$$

has zero measure (where the measure here is the product of ω^n and dt). Let $f(t,x)$ be the density of the Monge–Ampère measure $(\omega + i\partial\bar\partial\varphi_t)^n$. We then have

$$\int_{X\times[a,b]} \dot\varphi_t^- f(t,x)\omega^n\,dt = \int_{X\times[a,b]} \dot\varphi_t^+ f(t,x)\omega^n\,dt. \tag{29}$$

Now, by Fubini's theorem, (28) and (29) we see that the inequalities in (28) become equalities for almost every t in $[a,b]$, completing the proof. □

Theorem 4.12 *Let α be a Kähler form. Let $\varphi_0, \varphi_1 \in \mathcal{E}^1$ and let φ_t be the finite-energy geodesic connecting φ_0 and φ_1. Suppose that $\omega^n_{\varphi_t}$ is subordinate to ω^n for any $t \in [0,1]$. If AM_α is linear along $t \to \varphi_t$ then $\varphi_1 - \varphi_0$ is constant.*

Proof We can assume that $\mathrm{AM}(\varphi_0) = \mathrm{AM}(\varphi_1)$ and we normalize ω so that $V = 1$. We claim that AM_β is also linear along φ_t, where β is any Kähler form. Indeed, multiplying β by some small positive constant, we can assume that $\gamma := \alpha - \beta > 0$. It follows from Proposition 4.5 that both $t \to \mathrm{AM}_\gamma(\varphi_t)$ and $t \to \mathrm{AM}_\beta(\varphi_t)$ are convex. Because $\mathrm{AM}_\alpha = \mathrm{AM}_\beta + \mathrm{AM}_\gamma$ is linear along φ_t, it follows that in fact $t \to \mathrm{AM}_\beta(\varphi_t)$ is linear as well. By approximation it follows that $\mathrm{AM}_{\omega_\psi}$ is linear along φ_t for any $\psi \in \mathcal{E}^1$.

Fix $s \in (0,1)$ such that (27) holds in Lemma 4.11. For $h > 0$ small enough we have

$$\begin{aligned}\int_X \frac{\varphi_{s+h} - \varphi_s}{h}\omega^n_{\varphi_s} &\geq \frac{\mathrm{AM}_{\omega_{\varphi_s}}(\varphi_{s+h}) - \mathrm{AM}_{\omega_{\varphi_s}}(\varphi_s)}{h}\\ &= -\frac{1}{n}\int_X \frac{\varphi_{s+h} - \varphi_s}{h}\omega^n_{\varphi_{s+h}}\\ &\geq -\frac{1}{n}\int_X \frac{\varphi_{s+h} - \varphi_s}{h}\omega^n_{\varphi_s}.\end{aligned}$$

In the first line we have used Lemma 4.6. In the second line we have used the assumption that AM is constant along $s \to \varphi_s$. In the last line we have used again Lemma 4.6. Now, letting $h \to 0$ and using Lemma 4.11 we see that the right derivative of $l \to \mathrm{AM}_{\omega_{\varphi_s}}(\varphi_l)$ at s is zero. Thus $l \to \mathrm{AM}_{\omega_{\varphi_s}}(\varphi_l)$ is in fact constant. This combined with $l \to \mathrm{AM}(\varphi_l)$ being constant imply that

$$\begin{aligned}0 &= (n+1)\big(\mathrm{AM}(\varphi_1) - \mathrm{AM}(\varphi_s)\big) - n\big(\mathrm{AM}_{\omega_{\varphi_s}}(\varphi_1) - \mathrm{AM}_{\omega_{\varphi_s}}(\varphi_s)\big)\\ &= \int_X (\varphi_1 - \varphi_s)\omega^n_{\varphi_1}.\end{aligned} \tag{30}$$

A computation similar to the one in Lemma 4.6 gives that all terms in the expression of $\mathrm{AM}(\varphi_1) - \mathrm{AM}(\varphi_s)$ from (17) are greater than $\int_X (\varphi_1 - \varphi_s)\omega^n_{\varphi_1}$. Using this, (30) and $\mathrm{AM}(\varphi_1) - \mathrm{AM}(\varphi_s) = 0$ we obtain $\int_X (\varphi_1 - \varphi_s)\omega^n_{\varphi_s} = 0$. Together with (30) this gives

$$I(\varphi_1, \varphi_s) = \int_X (\varphi_1 - \varphi_s)(\omega^n_{\varphi_s} - \omega^n_{\varphi_1}) = 0.$$

Hence, by the results in [9, Section 2.1], the difference $\varphi_s - \varphi_1$ is constant. In fact $\varphi_s = \varphi_1$, as we have assumed that the Aubin–Mabuchi energy is constant along the geodesic $l \to \varphi_l$. Now, φ_1 can be replaced by φ_0 in (30), and the same arguments as above show that $\varphi_0 = \varphi_s$, ultimately giving $\varphi_0 = \varphi_1$. □

We are now ready to prove the uniqueness result.

Theorem 4.13 *Assume that χ is a Kähler form. If φ_0 and φ_1 are minimizers in $\mathcal{E}^1$ of the twisted Mabuchi energy $\mathcal{K}_\chi$ then $\varphi_1 - \varphi_0$ is constant.*

Proof Let $t \to \varphi_t$ be the finite-energy geodesic connecting φ_0 and φ_1. By the convexity of $\mathcal{K}_\chi$ it follows that $\mathcal{K}_\chi$ is linear along $t \to \varphi_t$. Since $t \to \mathrm{AM}(\varphi_t), \mathcal{K}_\chi(\varphi_t)$ are linear and $t \to \mathrm{AM}_\chi(\varphi_t), \mathcal{K}(\varphi_t)$ are convex, the decomposition

$$\mathcal{K}_\chi = \mathcal{K} + (\bar{S}_\chi - \bar{S})\,\mathrm{AM} + n\,\mathrm{AM}_\chi$$

then reveals that AM_χ is also linear along $t \to \varphi_t$ and $\omega^n_{\varphi_t}$ is subordinate to ω^n. The result now follows from Theorem 4.12. □

Remark 4.14 When χ is a Kähler form, using this last theorem, it can be seen that the conditions (A1)–(A4) and (P1)–(P7) are verified in [32, Theorem 3.4] for the data $(\mathcal{E}^1, d_1, \mathcal{K}_\chi, \{\mathrm{Id}\})$ to give that a minimizer of $\mathcal{K}_\chi$ exists in $\mathcal{E}^1$ if and only if there exist $C, D > 0$ such that

$$\mathcal{K}_\chi(u) \geq C d_1(0, u) - D, \quad u \in \mathcal{H}_\omega.$$

This verifies a weak version of [23, Conjecture 1.21] going back to [22, Conjecture 6.1]. For related partial results, see also [36].

5 Relating d_1–convergence to weak d_2–convergence

Before we get into the details of our particular situation, we start with a pedagogical example: suppose (M, μ) is a measure space with finite volume. By $(L^p(M, \mu), \|\cdot\|_p)$ we denote the usual L^p spaces on M. From Hölder's inequality it follows that on $L^2(M, \mu)$ the $\|\cdot\|_2$ norm dominates the $\|\cdot\|_1$ norm. Our focus however is on the weak-L^2–topology. As it turns out, the L^1–topology dominates the weak-L^2–topology. The simple explanation for this is that L^1–balls inside $L^2(M,\mu)$ are closed convex sets, and it is a classical fact that weak L^2–limits do not exit closed convex sets. Though much simplified, as it turns out, this idea generalizes to the setting of the metric spaces $(\mathcal{E}^p, d_p)$. As we show below, the d_1–metric balls have a certain convexity property that will make these sets d_2–convex and closed inside $\mathcal{E}^2$. This will imply that d_1–convergence dominates weak d_2–convergence. In the next section, coupled with Theorem 2.14, this fact will have implications for the convergence of the weak twisted Calabi flow.

As advocated in [29; 30], a proper understanding of the "rooftop" envelopes $P(u_0,u_1)$ gives insight into the geometry of the spaces $(\mathcal{E}^p, d_p)$. Furthering this relationship, we state the following proposition:

Proposition 5.1 *Suppose* $[0,1] \ni t \to u_t, v_t \in \mathcal{E}^1$ *are finite-energy geodesics. Then the map* $t \to \mathrm{AM}(P(u_t,v_t))$ *is concave. Consequently, the map* $t \to d_1(u_t,v_t)$ *is convex.*

The significance of this result comes from the fact that the d_1 metric, unlike the d_2 metric, is not CAT(0). Indeed, by the results of [30], the geodesic segments with fixed endpoints inside $(\mathcal{E}^1, d_1)$ are not even unique. On the other hand, by the above proposition, the d_1 metric structure has some geometric convexity that can be exploited.

Proof Let $a,b \in [0,1]$. As shown in [29, Theorem 3], we have $P(u_a,v_a), P(u_b,v_b) \in \mathcal{E}^1(X,\omega)$. Let $[0,1] \ni t \to w_t \in \mathcal{E}^1(X,\omega)$ be a finite-energy geodesic connecting $w_0 = P(u_a,v_a)$ and $w_1 = P(u_b,v_b)$. By the maximum principle of finite-energy geodesics we have $w_t \le u_{ta+(1-t)b}, v_{ta+(1-t)b}$, hence also $w_t \le P(u_{ta+(1-t)b}, v_{ta+(1-t)b})$. By the monotonicity of the Aubin–Mabuchi energy and since $t \to \mathrm{AM}(w_t)$ is linear we obtain

$$\begin{aligned} t\,\mathrm{AM}(P(u_a,v_a)) + (1-t)\,\mathrm{AM}(P(u_b,v_b)) &= \mathrm{AM}(w_t) \\ &\le \mathrm{AM}(P(u_{ta+(1-t)b}, v_{ta+(1-t)b})). \end{aligned}$$

The last statement of the proposition follows from the linearity of AM along finite-energy geodesics, the concavity we just established and the explicit formula for d_1 given in [30, Corollary 4.14], according to which

$$d_1(u_t,v_t) = \mathrm{AM}(u_t) + \mathrm{AM}(v_t) - 2\,\mathrm{AM}(P(u_t,v_t)). \qquad \square$$

The geodesic convexity and closedness of d_1–balls inside $\mathcal{E}^2$ is an immediate consequence:

Corollary 5.2 *For any $\rho > 0$ and $u \in \mathcal{E}^2(X,\omega)$, the set*

$$B_\rho(u) = \{v \in \mathcal{E}^2(X,\omega) : d_1(v,u) \le \rho\}$$

is d_2–closed and d_2–convex, ie for any $v_0, v_1 \in B_\rho(u)$ the finite-energy geodesic $[0,1] \ni t \to v_t \in \mathcal{E}^2$ connecting v_0 and v_1 is contained in $B_\rho(u)$.

Proof Closedness with respect to d_2 follows from the fact that d_2 dominates d_1. Let $[0,1] \ni t \to v_t \in \mathcal{E}^2$ be a finite-energy geodesic with $v_0, v_1 \in B_\rho(u)$. By definition, since $\mathcal{E}^2 \subset \mathcal{E}^1$, the curve $t \to v_t$ is a finite-energy geodesic inside $\mathcal{E}^1$ as well. By the previous proposition $t \to d_1(u, v_t)$ is convex, hence $d_1(u, v_t) \le \rho$. □

The main result of this subsection is the following:

Theorem 5.3 *Suppose $\{u_k\}_k \subset \mathcal{E}^2$ is d_2–bounded and $u \in \mathcal{E}^2$. Then $d_1(u_k, u) \to 0$ if and only if $\|u_j - u\|_{L^1(X)} \to 0$ and u_k converges to u d_2–weakly.*

Proof Assume first that $d_1(u_k, u) \to 0$. From [30, Theorem 5(ii)] it follows that $\|u_j - u\|_{L^1(X)} \to 0$. As recalled in Proposition 2.11, any subsequence of $\{u_k\}_k$ contains a d_2–weakly convergent subsubsequence u_{k_l}, converging d_2–weakly to some $v \in \mathcal{E}^2$. We show that $v = u$. Indeed, for any $j \in \mathbb{N}$ the set $B_{1/j}(u)$ is d_2–closed and d_2–convex by the previous corollary, and for large enough k_l we have $u_{k_l} \in B_{1/j}(u)$. As recalled in Proposition 2.11, it follows now that $v \in B_{1/j}(u)$ for all j, hence $v = u$.

For the reverse direction, as d_2–boundedness gives that $\mathrm{AM}(u_j)$ is uniformly bounded, by [30, Proposition 5.9] it suffices to show that any convergent subsequence of $\mathrm{AM}(u_j)$ converges to $\mathrm{AM}(u)$. Assume that u_{j_k} is such a subsequence and set $c = \lim_k \mathrm{AM}(u_{j_k})$. By definition, u_{j_k} still converges d_2–weakly to u. For each $\varepsilon > 0$, consider the set

$$E_\varepsilon := \{\phi \in \mathcal{E}^2 : c - \varepsilon \le \mathrm{AM}(\phi) \le c + \varepsilon\}.$$

Since d_2 dominates d_1 and AM is d_1–continuous and linear along finite-energy geodesics, it follows that E_ε is d_2–closed and d_2–convex. By Proposition 2.11, it follows that $u \in E_\varepsilon$. Letting $\varepsilon \to 0$ we get $\mathrm{AM}(u) = c$, finishing the proof. □

Remark 5.4 Using (the proof of) this last result, it is possible to construct a d_2–bounded sequence $u_j \in \mathcal{E}^2$ converging d_2–weakly to some $u \in \mathcal{E}^2$, but for which $\|u_j - u\|_{L^1(X)} \not\to 0$. Indeed, one can construct a d_2–bounded sequence $u_j \in \mathcal{E}^2$ such that $\|u_j - v\|_{L^1(X)} \to 0$ for some $v \in \mathcal{E}^2$ but $\omega^n_{u_j}$ does not converge weakly to ω^n_v, in particular $\mathrm{AM}(u_j)$ cannot converge to $\mathrm{AM}(v)$. By Proposition 2.11 we can extract a subsequence, again denoted by u_j, such that u_j converges d_2–weakly to some $u \in \mathcal{E}^2$. By the last step in the proof of the previous theorem AM is weak d_2–continuous, hence $\mathrm{AM}(u_j) \to \mathrm{AM}(u)$, but we cannot have $u = v$ as $\mathrm{AM}(u) \ne \mathrm{AM}(v)$.

6 The weak twisted Calabi flow

As shown in [29], the metric completion $(\mathcal{E}^2, d_2) = \overline{(\mathcal{H}, d_2)}$ is a CAT(0) space. Suppose χ satisfies (3). By Theorem 4.7, the extended $\mathcal{K}_\chi$ is d_2–lsc and convex on $\mathcal{E}^2$. By Theorem 2.12 and Remark 2.13, the weak gradient flow $t \to c_t$ of $\mathcal{K}_\chi$ emanating from any $c_0 \in \mathcal{E}^2$ is well defined and uniquely determined by the evolution variational inequality (13).

When χ is smooth, the *smooth twisted Calabi flow* is just a simple generalization of the usual smooth Calabi flow:

$$\frac{d}{dt} c_t = S_{\omega_{c_t}} - \bar{S}_\chi - \operatorname{Tr}^{\omega_{c_t}} \chi.$$

Comparison with Streets' setting In [61] another (a priori different) extension $\bar{\mathcal{K}}$ of the Mabuchi functional $\mathcal{M}$ on $\mathcal{H}$ to the completion $\overline{(\mathcal{H}, d_2)} = (\mathcal{E}^2, d_2)$ was considered, defined by

$$\bar{\mathcal{K}}(\bar{u}) := \liminf_{d(u_j, \bar{u}) \to 0} \mathcal{K}(u_j),$$

where the infimum is taken over all sequences u_j in $\mathcal{H}$ converging to $\bar{u}$ in $\overline{(\mathcal{H}, d_2)}$. It is shown in [61] that the functional $\bar{\mathcal{K}}$ thus defined is d_2–lsc on $\overline{(\mathcal{H}, d_2)}$, and then the author proceeds to study the gradient flow of $\bar{\mathcal{K}}$, dubbed the *minimizing movement Calabi flow*. By Theorem 4.7 we actually have $\bar{\mathcal{K}} = \mathcal{K}$, thus our finite-energy Calabi flow coincides with the minimizing movement Calabi flow considered in [61]. One of the advantages of our consideration is that computations in $\mathcal{E}^2$ are explicit and avoid the difficulties of using Cauchy sequences.

We show that the weak version of the twisted Calabi flow agrees with the smooth version as long as the latter exists. The following result was proved by Streets in the case $\chi = 0$ using different methods.

Proposition 6.1 *Suppose $\chi \geq 0$ is a smooth closed $(1,1)$–form. Given any initial point $c_0 \in \mathcal{H}_\omega$, the corresponding weak twisted Calabi flow $t \to c_t$ coincides with the smooth twisted Calabi flow, as long as the latter exists.*

Proof By the uniqueness property in [1, Theorem 4.0.4] for curves $t \to c_t$ satisfying the *evolution variational inequality* (13) (which is shown by differentiating $d(c_t^1, c_t^2)$ for two different solutions $t \to c_t^1$ and $t \to c_t^2$) it is enough to show that a solution $t \to h_t$ to the ordinary twisted Calabi flow with starting point $h_0 = c_0$ satisfies the inequality (13).

Suppose $v \in \mathcal{H}_\omega$ is arbitrary, fix a time $t = t_0$ and let $[0,1] \ni s \to u_s \in \mathcal{H}_\omega^\Delta$ be the weak geodesic connecting $u_0 = h_{t_0}$ and $u_1 = v$. From [8, Lemma 3.5] we get the following "slope inequality":

$$\mathcal{K}_\chi(v) - \mathcal{K}_\chi(h_{t_0}) \geq \int_X (\bar{S} - S_{\omega_{h_{t_0}}} + \operatorname{Tr}^{\omega_{h_{t_0}}} \chi) \frac{du_s}{ds}\Big|_{s=0} \omega^n_{h_{t_0}}.$$

Now, by the definition of the twisted Calabi flow the right-hand side above may be written as minus the scalar product $\int_X (dh_t/dt)_{|t=t_0}(du_s/ds)_{|s=0}\omega^n_{h_{t_0}}$. Since $v \in \mathcal{H}_\omega$, the latter scalar product coincides with the derivative at $t = t_0$ of the function $t \to \frac{1}{2} d_2^2(h_t, v)$ (by [20, Theorem 6], or rather by a formula appearing in the proof of the latter theorem). This concludes the proof in the case when $v \in \mathcal{H}_\omega$.

We handle the general case: suppose $v \in \mathcal{E}^2$ and $\mathcal{K}_\chi(v) < \infty$. Notice that it is enough to show the following "integral" version of (13) (with $G = \mathcal{K}_\chi$):

$$\tfrac{1}{2}(d_2^2(c_{t_1}, v) - d_2^2(c_{t_0}, v)) \leq (t_1 - t_0)\mathcal{K}_\chi(v) - \int_{t_0}^{t_1} \mathcal{K}_\chi(c_t)\, dt \tag{31}$$

for any $t_0, t_1 \in [0, \infty)$, $t_0 \leq t_1$. Indeed, the left-hand side is locally Lipschitz, whereas $t \to \mathcal{K}_\chi(c_t)$ is smooth, hence we may divide both sides by $t_1 - t_0$ and take the limit $t_1 \to t_0$ to obtain (13). By Theorem 3.2 there exists a sequence $v_j \in \mathcal{H}_\omega$ that d_2–converges to v such that $\mathcal{K}_\chi(v_j)$ converges to $\mathcal{K}_\chi(v)$. After integrating, by the first part of the proof estimate (31) holds for v_j in place of v. Letting $j \to \infty$, we obtain (31) for v as well. □

Lemma 6.2 *The functional* AM *is constant along any weak twisted Calabi flow trajectory $t \to c_t$.*

Proof For a *smooth* Calabi flow this follows directly from differentiating along the flow, but here we have to proceed in a different manner. We can assume that $\mathrm{AM}(c_0) = 0$, as $\mathcal{K}_\chi$ is invariant under adding constants. On the other hand, for any $u, v \in \mathcal{E}^2$, $d_2(u - \mathrm{AM}(u), v - \mathrm{AM}(v)) \leq d_2(u, v)$. Thus the variational construction of the weak Calabi flow (see Section 2.5) gives "minimizing movement" c_t^m with $\mathrm{AM}(c_t^m) = 0$ for all m. Since AM is continuous with respect to d_2 it follows that $\mathrm{AM}(c_t) = 0$ for all t. □

Now we arrive at the main result of this section:

Theorem 6.3 *Suppose (X, ω) is a compact connected Kähler manifold and $\chi = \beta + i\partial\bar{\partial} f$ satisfies (3). The following statements are equivalent:*

(i) $\mathcal{M}^2_\chi \neq \varnothing$.

(ii) *For any weak twisted Calabi flow trajectory $t \to c_t$ there exists $c_\infty \in \mathcal{M}^2_\chi$ such that $d_1(c_t, c_\infty) \to 0$ and $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_{c_t}) \to \mathrm{Ent}(e^{-f}\omega^n, \omega^n_{c_\infty})$.*

(iii) *Any weak twisted Calabi flow trajectory $t \to c_t$ is d_2–bounded.*

(iv) *There exists a weak twisted Calabi flow trajectory $t \to c_t$ and $t_j \to \infty$ for which the sequence $\{c_{t_j}\}_j$ is d_2–bounded.*

Proof We start with the implication (i) $\Longrightarrow$ (ii). Let $t \to c_t$ be a weak twisted Calabi flow trajectory. Let $v \in \mathcal{M}^2_\chi$. From (13) it follows that $d_2(v, c_t) \le d_2(v, c_0)$, hence $t \to c_t$ is a d_2–bounded curve.

As observed in [62], Theorem 2.14 guarantees the existence of $c_\infty \in \mathcal{M}^2_\chi$ such that $c_t \to c_\infty$ d_2–weakly. But $\{c_t\}_t$ is bounded in the d_2 metric and also $\mathcal{K}_\chi(c_t)$ is bounded. By Corollary 4.8 it follows that $\{c_t\}_t$ is d_1–relatively compact, ie each subsequence has a d_1–convergent subsubsequence. By Theorem 5.3 we must have $d_1(c_t, c_\infty) \to 0$.

In the definition of $\mathcal{K}_\chi$ all terms are d_1–continuous except for the entropy term. Since c_∞ is a minimizer, lower semicontinuity gives $\lim_{t\to\infty} \mathcal{K}_\chi(c_t) = \mathcal{K}_\chi(c_\infty)$. All this additionally implies $\mathrm{Ent}(e^{-f}\omega^n, \omega^n_{c_t}) \to \mathrm{Ent}(e^{-f}\omega^n, \omega^n_{c_\infty})$.

The implications (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (iv) are trivial. We finish the proof by arguing that (iv) $\Longrightarrow$ (i). Let $t \to c_t$ be a weak twisted Calabi flow trajectory and $\{c_{t_j}\}_j$ be a d_2–bounded sequence with $t_j \to \infty$. From Proposition 2.7 and Corollary 4.8 it follows that there exists $c_\infty \in \mathcal{E}^2$ such that $d_1(c_{t_j}, c_\infty) \to 0$, and by the lower semicontinuity of $\mathcal{K}_\chi$, we get that in fact $c_\infty \in \mathcal{M}^2_\chi$. □

In Theorem 6.3(ii) one would like to have convergence with respect to d_2. The next result confirms this in the case when the flow is bounded from below by some potential:

Proposition 6.4 *Suppose (X, ω) is a compact connected Kähler manifold and that χ satisfies (3). Let $t \to c_t$ be a weak twisted Calabi flow trajectory. If there exists $\psi \in \mathcal{E}^2$ such that $c_t \ge \psi$ for all t, then c_t converges in d_2 to a minimizer of $\mathcal{K}_\chi$.*

Proof Without loss of generality we can assume that $\psi \le 0$. By hypothesis we have in particular that $\psi \in \mathcal{E}^1$ and $\mathrm{AM}(\psi)$ is finite.

We first claim that $d_2(c_t, 0)$ is uniformly bounded in t. Indeed, by [30, Corollary 4.14] the d_1–distance $d_1(c_t, 0)$ can be expressed as

$$d_1(c_t, 0) = \mathrm{AM}(c_t) + \mathrm{AM}(0) - 2\,\mathrm{AM}(P(c_t, 0)).$$

Since $\psi \le P(c_t, 0) \le 0$ it follows from monotonicity of AM that $\mathrm{AM}(P(c_t, 0))$ is uniformly bounded in t. As $\mathrm{AM}(c_t)$ is constant, it follows that $d_1(c_t, 0)$ is uniformly bounded. This together with [30, Corollary 4] implies that $\sup_X c_t$ is bounded. Finally, applying [30, Theorem 3] finishes the proof of the claim.

By Theorem 6.3 we know that $t \to c_t$ converges in d_1 to some $u \in \mathcal{E}^2$, a minimizer of $\mathcal{K}_\chi$. As $c_t \geq \psi$, by the dominated convergence theorem and Theorem 2.4 we only have to prove that $\int_X (c_t - c)^2 \omega_{c_t}^n \to 0$. For a fixed $s > 0$ we have

$$\int_{\{|c_t - c| \leq s\}} (c_t - c)^2 \omega_{c_t}^n \leq s \int_X |c_t - c| \omega_{c_t}^n \to 0 \quad \text{as } t \to \infty,$$

since $d_1(c_t, c) \to 0$. Thus it suffices to show that

$$\sup_{t>0} \int_{\{|c_t - c| > s\}} (c_t - c)^2 \omega_{c_t}^n \to 0 \tag{32}$$

as $s \to \infty$. Since $d^2(c, c_t)$ is bounded, by Theorem 2.4 one can find a positive constant C_1 such that $\sup_X c_t \leq C_1$ for all $t > 0$. By the comparison principle in $\mathcal{E}$ (see [45]) one has

$$\int_{\{c_t - c > s\}} \omega_{c_t}^n \leq \int_{\{c_t - c > s\}} \omega_c^n \leq \int_{\{c < C_1 - s\}} \omega_c^n,$$

which yields

$$\int_s^\infty \omega_{c_t}^n (c_t - c > r) r \, dr \leq \int_s^\infty \omega_c^n (c < C_1 - r) r \, dr.$$

The right-hand side converges to 0 as $s \to \infty$ because $c \in \mathcal{E}^2$. Therefore, to prove (32) it remains to show that

$$\sup_{t>0} \int_s^\infty \omega_{c_t}^n (c_t - c < -r) r \, dr \to 0. \tag{33}$$

Since $\sup_X c_t$ is bounded from above and $c_t \geq \psi$, we can find $C_2 > 0$ such that

$$\{c_t - c < -r\} \subset \left\{\psi \leq C_2 + \tfrac{1}{2}(c_t - r)\right\}.$$

Using $\omega_{c_t}^n \leq 2^n \omega_{c_t/2}^n$ and the comparison principle, we arrive at

$$\begin{aligned} \int_s^\infty \omega_{c_t}^n (c_t - c < -r) r \, dr &\leq \int_s^\infty \omega_{c_t}^n \big(\psi < C_2 + \tfrac{1}{2}(c_t - r)\big) r \, dr \\ &\leq \int_s^\infty \omega_\psi^n \big(\psi < C_3 - \tfrac{1}{2} r\big) r \, dr, \end{aligned}$$

where $C_3 = C_2 + \frac{1}{2} C_1$. The last term converges to 0 as $s \to \infty$ because $\psi \in \mathcal{E}^2$. This proves (33) and completes the proof. □

Finally, we prove a result about geodesic rays weakly asymptotic to diverging weak Calabi flow trajectories.

Theorem 6.5 *Suppose (X,ω) is a compact connected Kähler manifold, $\chi \geq 0$ is smooth and Conjecture 1.8 holds. Let $[0,\infty) \ni t \to c_t \in \mathcal{E}^2$ be a weak twisted Calabi flow trajectory. Exactly one of the following holds:*

(i) *The curve $t \to c_t$ d_1–converges to a smooth twisted csc-K potential c_∞.*

(ii) *$d_1(c_0, c_t) \to \infty$ as $t \to \infty$ and the curve $t \to c_t$ is d_1–weakly asymptotic to a finite-energy geodesic $[0,\infty) \ni t \to u_t \in \mathcal{E}^1$ along which $\mathcal{K}_\chi$ decreases.*

If $\chi > 0$, then, independently of Conjecture 1.8, exactly one of the following holds:

(i′) *The curve $t \to c_t$ d_1–converges to a unique minimizer in $\mathcal{E}^1$ of $\mathcal{K}_\chi$.*

(ii′) *$d_1(c_0, c_t) \to \infty$ as $t \to \infty$ and the curve $t \to c_t$ is d_1–weakly asymptotic to a finite-energy geodesic $[0,\infty) \ni t \to u_t \in \mathcal{E}^1$ along which $\mathcal{K}_\chi$ strictly decreases.*

Proof Suppose (i) holds. Then $t \to c_t$ is d_2–bounded hence also d_1–bounded, hence it is impossible for (ii) to hold.

Now suppose (i) does not hold. By Corollary 4.8 we must have $d_t := d_1(c_0, c_t) \to \infty$, otherwise there would exist $c_\infty \in \mathcal{M}^1_\chi$ smooth twisted csc-K; in particular, $c_\infty \in \mathcal{M}^2_\chi$. By Theorem 1.5 this would imply that (i) holds, a contradiction.

Let $[0, d_t] \ni l \to u^t_l \in \mathcal{E}^1$ be the d_1–unit finite-energy geodesic connecting c_0 and c_t. By convexity of $l \to \mathcal{K}_\chi(u^t_l)$ it follows that

$$\frac{\mathcal{K}_\chi(u^t_l) - \mathcal{K}_\chi(c_0)}{l} = \frac{\mathcal{K}_\chi(u^t_l) - \mathcal{K}_\chi(u^t_0)}{l} \leq \frac{\mathcal{K}_\chi(u^t_{d_t}) - \mathcal{K}_\chi(u^t_0)}{d_1(c_0, c_t)} = \frac{\mathcal{K}_\chi(c_t) - \mathcal{K}_\chi(c_0)}{d_1(c_0, c_t)} \leq 0,$$

hence $\{\mathcal{K}_\chi(u^t_l)\}_{t\in[0,\infty)}$ is uniformly bounded. As $d_1(u^t_l, u^t_0) = l$, we can apply Corollary 4.8 to find a subsequence d_1–converging to some $u_l \in \mathcal{E}^1$. Using a Cantor process, we can arrange for a subsequence t_k such that for all $l \in \mathbb{Q}$ there exists $u_l \in \mathcal{E}^1$ such that $d_1(u^{t_k}_l, u_l) \to 0$ as $k \to \infty$ for each l. As we are dealing with the limit of d_1–unit speed geodesic segments, we will clearly have

$$d_1(u_{l_1}, u_{l_2}) = |l_1 - l_2|, \quad l_1, l_2 \in \mathbb{Q}_+.$$

Using equicontinuity, in the complete metric space $\mathcal{E}^1$ we can extend the curve $\mathbb{Q}_+ \ni l \to u_l \in \mathcal{E}^1$ to a d_1–geodesic ray $[0,\infty) \ni l \to u_l \in \mathcal{E}^1$, satisfying $d_1(u^{t_k}_l, u_l) \to 0$ for all $l \in [0,\infty)$.

Using Proposition 4.3 we additionally obtain that $l \to u_l$ is in fact a finite-energy geodesic. Because all functions $l \to \mathcal{K}_\chi(u^{t_k}_l)$ are uniformly bounded above and $\mathcal{K}_\chi$ is d_1–lsc, it necessarily follows that $l \to \mathcal{K}_\chi(u_l)$ is also bounded above. Convexity and boundedness now give that $l \to \mathcal{K}_\chi(u_l)$ is actually decreasing.

Lastly, we focus on the case when $\chi > 0$ is Kähler. In Theorem 4.13 we have proved that a minimizer of the twisted Mabuchi functional is unique if exists. Also, when (i′) holds then by Remark 4.14 the curve $t \to c_t$ is d_1–bounded, hence it is impossible for (ii′) to hold.

We assume that (i′) does not hold. Let $t \to c_t$ be a weak twisted Calabi flow trajectory. We can assume that c_t is d_1–divergent, otherwise Theorem 2.8 would imply existence of a minimizer in $\mathcal{E}^1$. By the same argument as above, we can construct a weakly asymptotic finite-energy geodesic ray $t \to u_t$ along which $\mathcal{K}_\chi$ is decreasing. We claim that in fact $\mathcal{K}_\chi$ is strictly decreasing along $t \to u_t$. Indeed, if it were not the case, by convexity of $t \to \mathcal{K}_\chi(u_t)$ we would obtain that $t \to \mathcal{K}_\chi(u_t)$ is constant for t greater than some $t_0 > 0$. By Lemma 6.2 AM is constant along $t \to c_t$, hence also along $t \to u_t$. As both $t \to \mathcal{K}(u_t), \mathrm{AM}_\chi(u_t)$ are convex, we obtain that $t \to \mathrm{AM}_\chi(u_t)$ is in fact linear and Theorem 4.12 then reveals that u_t is stationary after t_0, contradicting the d_1–divergence of the ray $t \to u_t$. □

References

[1] **L Ambrosio**, **N Gigli**, **G Savaré**, *Gradient flows in metric spaces and in the space of probability measures*, 2nd edition, Birkhäuser, Basel (2008) MR

[2] **T Aubin**, *Équations du type Monge–Ampère sur les variétés kählériennes compactes*, Bull. Sci. Math. 102 (1978) 63–95 MR

[3] **M Bačák**, *The proximal point algorithm in metric spaces*, Israel J. Math. 194 (2013) 689–701 MR

[4] **E Bedford**, **B A Taylor**, *The Dirichlet problem for a complex Monge–Ampère equation*, Invent. Math. 37 (1976) 1–44 MR

[5] **E Bedford**, **B A Taylor**, *A new capacity for plurisubharmonic functions*, Acta Math. 149 (1982) 1–40 MR

[6] **E Bedford**, **B A Taylor**, *Fine topology, Šilov boundary, and $(dd^c)^n$*, J. Funct. Anal. 72 (1987) 225–251 MR

[7] **R J Berman**, *A thermodynamical formalism for Monge–Ampère equations, Moser–Trudinger inequalities and Kähler–Einstein metrics*, Adv. Math. 248 (2013) 1254–1297 MR

[8] **R J Berman**, **B Berndtsson**, *Convexity of the K-energy on the space of Kähler metrics and uniqueness of extremal metrics*, preprint (2014) arXiv

[9] **R J Berman**, **S Boucksom**, **P Eyssidieux**, **V Guedj**, **A Zeriahi**, *Kähler–Einstein metrics and the Kähler–Ricci flow on log Fano varieties*, preprint (2011) arXiv

[10] **R J Berman**, **S Boucksom**, **V Guedj**, **A Zeriahi**, *A variational approach to complex Monge–Ampère equations*, Publ. Math. Inst. Hautes Études Sci. 117 (2013) 179–245 MR

[11] **R J Berman**, **T Darvas**, **C H Lu**, *Regularity of weak minimizers of the K-energy and applications to properness and K-stability*, preprint (2016) arXiv

[12] **R J Berman**, **H Guenancia**, *Kähler–Einstein metrics on stable varieties and log canonical pairs*, Geom. Funct. Anal. 24 (2014) 1683–1730 MR

[13] **B Berndtsson**, *The openness conjecture and complex Brunn–Minkowski inequalities*, from "Complex geometry and dynamics" (J E Fornæss, M Irgens, E F Wold, editors), Abel Symposia 10, Springer (2015) 29–44

[14] **Z Błocki**, *On geodesics in the space of Kähler metrics*, from "Advances in geometric analysis" (S Janeczko, J Li, D H Phong, editors), Adv. Lect. Math. 21, International Press, Somerville, MA (2012) 3–19 MR

[15] **Z Błocki**, **S Kołodziej**, *On regularization of plurisubharmonic functions on manifolds*, Proc. Amer. Math. Soc. 135 (2007) 2089–2093 MR

[16] **T Bloom**, **N Levenberg**, *Pluripotential energy*, Potential Anal. 36 (2012) 155–176 MR

[17] **S Boucksom**, **P Eyssidieux**, **V Guedj**, **A Zeriahi**, *Monge–Ampère equations in big cohomology classes*, Acta Math. 205 (2010) 199–262 MR

[18] **M R Bridson**, **A Haefliger**, *Metric spaces of non-positive curvature*, Grundl. Math. Wissen. 319, Springer (1999) MR

[19] **X Chen**, *On the lower bound of the Mabuchi energy and its application*, Internat. Math. Res. Notices (2000) 607–623 MR

[20] **X Chen**, *The space of Kähler metrics*, J. Differential Geom. 56 (2000) 189–234 MR

[21] **X X Chen**, *A new parabolic flow in Kähler manifolds*, Comm. Anal. Geom. 12 (2004) 837–852 MR

[22] **X Chen**, *Space of Kähler metrics, IV: On the lower bound of the K-energy*, preprint (2008) arXiv

[23] **X Chen**, *On the existence of constant scalar curvature Kähler metric: a new perspective*, preprint (2015) arXiv

[24] **X X Chen**, **W Y He**, *On the Calabi flow*, Amer. J. Math. 130 (2008) 539–570 MR

[25] **X Chen**, **W He**, *The Calabi flow on toric Fano surfaces*, Math. Res. Lett. 17 (2010) 231–241 MR

[26] **X Chen**, **M Paun**, **Y Zeng**, *On deformation of extremal metrics*, preprint (2015) arXiv

[27] **X Chen**, **S Sun**, *Calabi flow, geodesic rays, and uniqueness of constant scalar curvature Kähler metrics*, Ann. of Math. 180 (2014) 407–454 MR

[28] **P T Chruściel**, *Semi-global existence and convergence of solutions of the Robinson–Trautman (2–dimensional Calabi) equation*, Comm. Math. Phys. 137 (1991) 289–313 MR

[29] **T Darvas**, *The Mabuchi completion of the space of Kähler potentials*, preprint (2015) arXiv

[30] **T Darvas**, *The Mabuchi geometry of finite energy classes*, Adv. Math. 285 (2015) 182–219 MR

[31] **T Darvas**, **W He**, *Geodesic rays and Kähler–Ricci trajectories on Fano manifolds*, Trans. Amer. Math. Soc. (online publication March 2017)

[32] **T Darvas**, **Y A Rubinstein**, *Tian's properness conjectures and Finsler geometry of the space of Kähler metrics*, J. Amer. Math. Soc. 30 (2017) 347–387 MR

[33] **E De Giorgi**, *New problems on minimizing movements*, from "Boundary value problems for partial differential equations and applications" (J-L Lions, C Baiocchi, editors), Res. Notes Appl. Math. 29, Masson, Paris (1993) 81–98 MR

[34] **J-P Demailly**, *Regularization of closed positive currents of type $(1,1)$ by the flow of a Chern connection*, from "Contributions to complex analysis and analytic geometry" (H Skoda, J-M Trépreau, editors), Aspects Math. E26, Vieweg, Braunschweig (1994) 105–126 MR

[35] **A Dembo**, **O Zeitouni**, *Large deviations: techniques and applications*, 1st edition, Jones and Bartlett, Boston (1993) MR

[36] **R Dervan**, *Uniform stability of twisted constant scalar curvature Kähler metrics*, Int. Math. Res. Not. 2016 (2016) 4728–4783 MR

[37] **S Dinew**, *Uniqueness in $\mathcal{E}(X,\omega)$*, J. Funct. Anal. 256 (2009) 2113–2122 MR

[38] **S K Donaldson**, *Symmetric spaces, Kähler geometry and Hamiltonian dynamics*, from "Northern California Symplectic Geometry Seminar" (Y Eliashberg, D Fuchs, T Ratiu, A Weinstein, editors), Amer. Math. Soc. Transl. Ser. 2 196, Amer. Math. Soc., Providence, RI (1999) 13–33 MR

[39] **S K Donaldson**, *Conjectures in Kähler geometry*, from "Strings and geometry" (M Douglas, J Gauntlett, M Gross, editors), Clay Math. Proc. 3, Amer. Math. Soc., Providence, RI (2004) 71–78 MR

[40] **R Feng**, **H Huang**, *The global existence and convergence of the Calabi flow on $\mathbb{C}^n/\mathbb{Z}^n+i\mathbb{Z}^n$*, J. Funct. Anal. 263 (2012) 1129–1146 MR

[41] **J Fine**, *Constant scalar curvature Kähler metrics on fibred complex surfaces*, J. Differential Geom. 68 (2004) 397–432 MR

[42] **J Fine**, *Calabi flow and projective embeddings*, J. Differential Geom. 84 (2010) 489–523 MR

[43] **Q Guan**, **X Zhou**, *A proof of Demailly's strong openness conjecture*, Ann. of Math. 182 (2015) 605–616 MR

[44] **V Guedj**, *The metric completion of the Riemannian space of Kähler metrics*, preprint (2014) arXiv

[45] **V Guedj**, **A Zeriahi**, *The weighted Monge–Ampère energy of quasiplurisubharmonic functions*, J. Funct. Anal. 250 (2007) 442–482 MR

[46] **V Guedj**, **A Zeriahi**, *Regularizing properties of the twisted Kähler–Ricci flow*, preprint (2013) arXiv

[47] **V Guedj**, **A Zeriahi**, *Degenerate complex Monge–Ampère equations*, EMS Tracts in Math. 26, Eur. Math. Soc., Zürich (2017)

[48] **W He**, *On the convergence of the Calabi flow*, Proc. Amer. Math. Soc. 143 (2015) 1273–1281 MR

[49] **H Huang**, *Convergence of the Calabi flow on toric varieties and related Kähler manifolds*, J. Geom. Anal. 25 (2015) 1080–1097 MR

[50] **H Huang**, **K Zheng**, *Stability of the Calabi flow near an extremal metric*, Ann. Sc. Norm. Super. Pisa Cl. Sci. 11 (2012) 167–175 MR

[51] **W A Kirk**, **B Panyanak**, *A concept of convergence in geodesic spaces*, Nonlinear Anal. 68 (2008) 3689–3696 MR

[52] **S Kołodziej**, *The complex Monge–Ampère equation*, Acta Math. 180 (1998) 69–117 MR

[53] **S Kołodziej**, *Hölder continuity of solutions to the complex Monge–Ampère equation with the right-hand side in L^p: the case of compact Kähler manifolds*, Math. Ann. 342 (2008) 379–386 MR

[54] **H Li**, **B Wang**, **K Zheng**, *Regularity scales and convergence of the Calabi flow*, preprint (2015) arXiv

[55] **T Mabuchi**, *Some symplectic geometry on compact Kähler manifolds, I*, Osaka J. Math. 24 (1987) 227–252 MR

[56] **U F Mayer**, *Gradient flows on nonpositively curved metric spaces and harmonic maps*, Comm. Anal. Geom. 6 (1998) 199–253 MR

[57] **E D Nezza**, **C H Lu**, *Uniqueness and short time regularity of the weak Kähler–Ricci flow*, preprint (2014) arXiv

[58] **J Pook**, *Twisted Calabi flow on Riemann surfaces*, Int. Math. Res. Not. 2016 (2016) 83–108 MR

[59] **S Semmes**, *Complex Monge–Ampère and symplectic manifolds*, Amer. J. Math. 114 (1992) 495–550 MR

[60] **J Stoppa**, *Twisted constant scalar curvature Kähler metrics and Kähler slope stability*, J. Differential Geom. 83 (2009) 663–691 MR

[61] **J Streets**, *Long time existence of minimizing movement solutions of Calabi flow*, Adv. Math. 259 (2014) 688–729 MR

[62] **J Streets**, *The consistency and convergence of K–energy minimizing movements*, Trans. Amer. Math. Soc. 368 (2016) 5075–5091 MR

[63] **G Székelyhidi**, *Remark on the Calabi flow with bounded curvature*, Univ. Iagel. Acta Math. 50 (2013) 107–115 MR

[64] **G Székelyhidi**, **V Tosatti**, *Regularity of weak solutions of a complex Monge–Ampère equation*, Anal. PDE 4 (2011) 369–378 MR

[65] **V Tosatti**, **B Weinkove**, *The Calabi flow with small initial energy*, Math. Res. Lett. 14 (2007) 1033–1039 MR

[66] **C Villani**, *Topics in optimal transportation*, Graduate Studies in Mathematics 58, Amer. Math. Soc., Providence, RI (2003) MR

RJB and CHL: *Chalmers University of Technology and University of Gothenburg*
SE-412 96 Gothenburg, Sweden

TD: *Department of Mathematics, University of Maryland*
College Park, MD 20742-4015, United States

Current address for CHL: *Département de Mathématiques, Université Paris-Sud*
Bâtiment 425, Bureau 144, 91405 Paris Orsay Cedex, France

robertb@chalmers.se, tdarvas@math.umd.edu,
hoang-chinh.lu@math.u-psud.fr

Proposed: Gang Tian
Seconded: John Lott, Bruce Kleiner

Received: 19 November 2015
Revised: 10 July 2016

msp

Geometry & Topology 21 (2017) 2989–3008

On 5–manifolds with free fundamental group and simple boundary links in S^5

MATTHIAS KRECK
YANG SU

We classify compact oriented 5–manifolds with free fundamental group and π_2 a torsion-free abelian group in terms of the second homotopy group considered as a π_1–module, the cup product on the second cohomology of the universal covering, and the second Stiefel–Whitney class of the universal covering. We apply this to the classification of simple boundary links of 3–spheres in S^5. Using this we give a complete algebraic picture of closed 5–manifolds with free fundamental group and trivial second homology group.

57R65; 57R40

1 Introduction

There is a close relationship between classical links and closed 3–manifolds since all 3–manifolds are obtained by surgeries on links and Kirby calculus determines when two links give the same 3–manifold. We consider a special case of such a relation in dimension 5. The special condition on the side of links is that we only consider *simple boundary links* L of a disjoint union of 3–spheres in S^5, which means that the fundamental group of the complement is freely generated by the meridians of the link components. As in dimension 3 we can perform surgery on the link L to obtain a closed smooth manifold $M(L)$. It is easy to see that the fundamental group of $M(L)$ is a free group and $H_2(M(L);\mathbb{Z})=0$. In addition, the second homotopy group is that of the complement X of the link and this is torsion-free as an abelian group. One can ask which 5–manifolds are obtained this way and for the classification of the links and the determination of the fibers of the map from links to 5–manifolds given by surgery.

We answer this question by giving a classification of a more general class of closed 5–manifolds, namely we classify all 5–manifolds M with $\pi_1(M)$ a free group and $\pi_2(M)$ torsion-free as an abelian group, in terms of an invariant we call *generalized Milnor pairing*, since it is a generalization of the Milnor pairing for knots. We also consider compact manifolds with boundary the disjoint union of copies of $S^1\times S^3$ and free fundamental group that is freely generated by the circles in the boundary, and, as

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.2989

before, $\pi_2(M)$ is torsion-free as an abelian group. We also define a topological version of the generalized Milnor pairing, called topological generalized Milnor pairing, and prove a corresponding result for topological manifolds.

A second well-known class of examples are fibered 5–manifolds M over the circle with simply connected fiber. These are in the image of the surgery construction above if and only if we have a fibered knot and $H_2(M;\mathbb{Z}) = 0$. But in general fibered 5–manifolds over the circle have nontrivial second homology. Thus our more general class of manifolds also occurs naturally. See Remark 1.4 and the appendix for more on this class of manifolds.

To give a feeling for the generalized Milnor pairing, we define it in a special case, where M is spin. Then it is represented by the triple

$$\big(\pi_1(M),\, \pi_2(M),\, b_M\colon \pi_2(M)^* \times \pi_2(M)^* \to \big(H^1(B\pi_1 M;\mathbb{Q}[\pi_1 M])\big)^*\big),$$

where b_M is given by the cup product. For details we refer to Section 2. Now we formulate our main result.

Theorem 1.1 *Let M_0 and M_1 be two smooth (or topological), compact, oriented 5–manifolds with free fundamental group of rank n and torsion-free π_2, with empty boundary or boundary consisting of n copies of $S^1 \times S^3$ such that the circles in the boundary generate $\pi_1(M_i)$. Then there is an orientation-preserving diffeomorphism (homeomorphism) between M_0 and M_1 if and only if there is an isomorphism between their (topological) generalized Milnor pairings.*

We actually prove a stronger result (Theorem 2.4) about the realization of isomorphisms between the generalized Milnor pairings.

Levine [11] has classified 3–dimensional simple knots in S^5 in terms of S–equivalence classes of Seifert matrices and Liang [12] has extended this to higher-dimensional simple boundary links in terms of l–equivalence classes of Seifert matrices. The general case of 3–dimensional simple boundary links in S^5 seems to be open. Our classification result implies that Liang's result extends to dimension 3. Also, by extending Liang's argument to higher dimension we can characterize the Seifert matrices occurring from links. We call the corresponding conditions *unimodularity conditions*. Thus we obtain a complete algebraic picture of simple boundary links in S^5.

Theorem 1.2 *The l–equivalence classes of Seifert matrices of simple boundary links of 3–spheres in S^5 determine the isotopy type of the link. Moreover, the l–equivalence classes of Seifert matrices give a bijection from the set of isotopy classes of simple boundary links of 3–spheres in S^5 to the set of l–equivalence classes of square integral matrices D satisfying the unimodularity conditions.*

We would also like to give an algebraic picture of our closed 5–manifolds. In general we don't know which values the generalized Milnor pairing takes. But if we require that $H_2(M;\mathbb{Z})=0$, these manifolds are all results of surgeries on links and we can use the realization of the link invariants to give a complete answer.

Let D be an $m\times m$ integral matrix satisfying the unimodularity conditions; then there is associated to D a $\mathbb{Z}[F_n]$–module map $\varphi_D\colon (\mathbb{Z}[F_n])^m\to(\mathbb{Z}[F_n])^m$ and a generalized Milnor pairing

$$\big(F_n,\ \operatorname{coker}\varphi_D,\ b_D\colon (\operatorname{coker}\varphi_D)^*\times(\operatorname{coker}\varphi_D)^*\to\big(H^1(BF_n;\mathbb{Q}[F_n])\big)^*\big)$$

We will give a detailed description of this in Section 4.

Theorem 1.3 *There is a bijection between the diffeomorphism classes of closed oriented 5–manifolds M with $\pi_1(M)$ a free group of rank n and $H_2(M;\mathbb{Z})=0$, and the isomorphism classes of generalized Milnor pairings $(F_n,\operatorname{coker}\varphi_D,b_D)$ for all matrices D (with various sizes m) fulfilling the unimodularity conditions.*

We will give more details of the generalized Milnor paring in Section 2, and prove the main classification theorem in Section 3. The discussion of 3–links and their relation with 5–manifolds will be the contents of Section 4.

Remark 1.4 A special case of Theorem 2.4 is when $\pi_1(M)\cong\mathbb{Z}$ and $\pi_2(M)$ is a finitely generated abelian group. In this case we can show that $\pi_2(M)$ is torsion-free and the bilinear form on $\pi_2(M)$ is unimodular, $w_2(\widetilde{M})$ is determined by the bilinear form on $\pi_2(M)$, and the realization problem of the invariants can be solved. This gives a complete classification of closed 5–manifolds with $\pi_1=\mathbb{Z}$ and π_2 a finitely generated abelian group. As an application, this reproves the fibration theorems in dimension 5 in the topological and smooth category given by Hsu [7], Weinberger [20] and Shaneson [17], respectively. See more details in the appendix.

Remark 1.5 The notions of *Borel manifolds* and *strongly Borel manifolds* were coined by Kreck and Lück [10, Definition 0.2]. A manifold M is called a Borel manifold if for any homotopy equivalence $f\colon N\to M$ there exists a homeomorphism $h\colon N\to M$ such that f and h induce the same map on the fundamental groups up to conjugation. It is called strongly Borel if all homotopy equivalences are homotopic to a homeomorphism. If M^5 is a closed oriented spin topological 5–manifold with free fundamental group and torsion-free π_2, then it is Borel. Since for any homotopy equivalence $f\colon N^5\xrightarrow{\simeq}M^5$, f induces an isomorphism between the topological generalized Milnor pairings (in this case the Kirby–Siebenmann invariant is determined by the bilinear form b_M; see the proof of Theorem 2.4), the statement follows from Theorem 1.1. On the other hand, for a closed oriented topological 5–manifold M^5 with free fundamental group, a computation of the topological structure set of M using the surgery exact sequence gives $\mathscr{S}^{\mathrm{TOP}}(M^5) = H^2(M;\mathbb{Z}/2)$. Therefore, by [10, Theorem 1.1], M is strongly Borel if and only if $H_2(M;\mathbb{Z}/2)=0$.

One often hears the statement that the classification of high-dimensional manifolds is completely understood. What people mean is that with the s–cobordism theorem one has a criterion of when two manifolds are diffeomorphic and with surgery theory one has a reduction of the problem of finding an s–cobordism to problems in homotopy theory (unstable and stable) and algebra (surgery obstruction groups), and the analysis of certain maps relating the homotopy theory and the algebra. But this doesn't mean that even for some very explicit manifolds, like for example complete intersections, the procedure can be carried out successfully. Given the complications of the homotopy groups of spheres, in higher dimensions the problems get harder and harder. But in comparatively low dimensions (say up to 8) one has a chance, which doesn't mean that it is routine. Most results in that dimension range concern simply connected manifolds. In this paper we make a first step towards a classification of 5–manifolds with fundamental group the free group F_n. This class is particular interesting, since such manifolds occur on the one hand as total spaces of bundles over the circle and on the other hand as fundamental groups of links of 3–spheres in S^5. We classify both in the smooth and topological category. It might be interesting to note that the topological classification of 4–manifolds with fundamental group the free group F_n is completely open for $n>1$. The question of whether the group F_n is good in the sense of Freedman and Quinn [6] is the key question for topological 4–manifolds. If this is the case then one can use similar methods as in the present paper to attack the classification of 4–manifolds with fundamental group F_n.

Acknowledgement Kreck would like to thank the Mathematical Institute of the Chinese Academy of Sciences in Beijing and the Max-Planck-Institute for Mathematics in Bonn for their support while this research was carried out. Su would like to thank the Max-Planck Institute for Mathematics in Bonn for a research visit in August and September, 2015. Both authors would like to thank the referee and the editor for their suggestions to improve the paper.

Su was partially supported by NSFC11571343.

2 The generalized Milnor pairing and the statement of the main theorem

Now we describe the generalized Milnor pairing which we use to classify our manifolds. First we give the general algebraic definition. A *generalized Milnor pairing* is a quadruple (π_1, π_2, b, w_2) consisting of the following:

(1) π_1 a free group of rank n; let $\Lambda = \mathbb{Z}[\pi_1]$ be the integral group ring and $\Lambda_{\mathbb{Q}} = \mathbb{Q}[\pi_1]$ be the rational group ring.
(2) π_2 a finitely generated Λ–module, which is torsion-free as an abelian group.
(3) $b\colon \pi_2^* \times \pi_2^* \to (H^1(B\pi_1; \Lambda_{\mathbb{Q}}))^*$ a symmetric Λ–equivariant pairing, where $*$ stands for the $\mathbb{Q}$–dual $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{Q})$, and by Λ–equivariant we mean that b is a Λ–module map under the diagonal action of Λ on $\pi_2^* \times \pi_2^*$ and the natural Λ–module structure on $(H^1(B\pi_1; \Lambda_{\mathbb{Q}}))^*$.
(4) $w_2 \in \mathrm{Hom}(\pi_2, \mathbb{Z}/2)$.

An isomorphism $(\alpha, \beta)\colon (\pi_1, \pi_2, b, w_2) \to (\pi_1', \pi_2', b', w_2')$ between generalized Milnor pairings consists of

(1) an isomorphism $\alpha\colon \pi_1 \to \pi_1'$;
(2) an isomorphism $\beta\colon \pi_2 \to \pi_2'$, which is compatible with the Λ– and Λ'–module structure and the pairings b and b', and maps w_2' to w_2.

Let M^5 be a smooth closed oriented 5–manifold with $\pi_1(M) \cong F_n$ and $\pi_2(M)$ a torsion-free abelian group; we associate a generalized Milnor pairing $\varphi(M) = (\pi_1(M), \pi_2(M), b_M, w_2(\widetilde{M}))$ to M as follows. Let $\widetilde{M}$ be the universal cover of M. By Poincaré duality we have an isomorphism $H_4(\widetilde{M}; \mathbb{Q}) = H_4(M; \Lambda) \otimes \mathbb{Q} \cong H^1(M; \Lambda_{\mathbb{Q}})$ and the latter group is isomorphic to $H^1(B\pi_1(M); \Lambda_{\mathbb{Q}})$, because M has a CW–structure $M \simeq \bigvee_n S^1 \vee \bigvee S^2 \cup e^3 \cup \cdots$ [19, Proposition 3.3]. Next we use the Kronecker isomorphism to identify $H^4(\widetilde{M}; \mathbb{Q})$ with $H_4(\widetilde{M}; \mathbb{Q})^*$, where $*$ stands for the $\mathbb{Q}$–dual, and the isomorphism above to obtain an isomorphism $H^4(\widetilde{M}; \mathbb{Q}) \cong (H^1(B\pi_1(M); \Lambda_{\mathbb{Q}}))^*$. The cup product and this identification together define a symmetric Λ–equivariant form

$$H^2(\widetilde{M}; \mathbb{Q}) \times H^2(\widetilde{M}; \mathbb{Q}) \to (H^1(B\pi_1 M; \Lambda_{\mathbb{Q}}))^*.$$

Using the Kronecker isomorphism and the Hurewicz isomorphism we obtain a symmetric Λ–equivariant form

$$b_M\colon \pi_2(M)^* \times \pi_2(M)^* \to (H^1(B\pi_1 M; \Lambda_{\mathbb{Q}}))^*,$$

where $*$ is again the vector space of homomorphisms to $\mathbb{Q}$, We will discuss more about this bilinear form in the beginning of Section 3.

To this we add the second Stiefel Whitney class

$$w_2(\widetilde{M}) \in \mathrm{Hom}(H_2(\widetilde{M}; \mathbb{Z}), \mathbb{Z}/2) = \mathrm{Hom}(\pi_2(M), \mathbb{Z}/2)$$

to obtain our invariant and get the quadruple

$$\varphi(M) = (\pi_1(M), \pi_2(M), b_M, w_2(\widetilde{M}));$$

we call this the *generalized Milnor pairing of* M. The group of self-isomorphisms of $\varphi(M)$ is denoted by $\mathrm{Aut}(\varphi(M))$.

Remark 2.1 In the case where only spin manifolds are concerned, $w_2(\widetilde{M})$ is always 0, and the generalized Milnor pairing is actually a triple $\varphi(M)=(\pi_1(M),\pi_2(M),b_M)$. This is the case in Theorem 1.3.

Remark 2.2 It's easy to see from the Leray–Serre spectral sequence of the fibration $\widetilde{M}\xrightarrow{p} M\to\bigvee_n S^1$ that $p^*\colon H^2(M;\mathbb{Z}/2)\to H^2(\widetilde{M};\mathbb{Z}/2)$ is injective. Therefore $w_2(M)$ and $w_2(\widetilde{M})$ determine each other.

We also classify a special case of compact oriented manifolds M with boundary which is relevant for classifying links in S^5. The boundary has to be a disjoint union of n copies of $S^1\times S^3$ and we require that the circles in the boundary components generate the fundamental group F_n of M. Here we replace $H_4(\widetilde{M};\mathbb{Q})$ by $H_4(\widetilde{M},\partial\widetilde{M};\mathbb{Q})$ and we note that $H^2(\widetilde{M};\mathbb{Q})\cong H^2(\widetilde{M},\partial\widetilde{M};\mathbb{Q})$, so that the definition of b_M makes sense. With this modification we can consider the quadruple defining $\varphi(M)$ as before. But we have to observe that the identification of the fundamental groups of M and M' is now given by an identification of the boundary components.

Remark 2.3 When X is the complement of a simple 3–knot, we have a bilinear paring $b\colon H^2(\widetilde{X};\mathbb{Q})\times H^2(\widetilde{X};\mathbb{Q})\to\mathbb{Q}$, which is the Milnor paring [15]; see also [13].

We also classify the corresponding topological manifolds. Here we add a fifth term to our invariant, the Kirby–Siebenmann invariant

$$KS(M)\in H^4(M;\mathbb{Z}/2)\cong\pi_1(M)/[\pi_1(M),\pi_1(M)]\otimes\mathbb{Z}/2.$$

We call the quintuple $(\pi_1(M),\pi_2(M),b_M,w_2(\widetilde{M}),KS(M))$ the *topological generalized Milnor pairing* of the topological manifold M. Of course in the definition of an isomorphism (α,β) between two topological generalized Milnor pairings we require that the isomorphism $\alpha\colon\pi_1(M)\to\pi_1(M')$ respects the Kirby–Siebenmann invariant, too.

Now we restate the classification theorem of the manifolds under consideration and add the realization statement for induced maps.

Theorem 2.4 *Let M_0 and M_1 be two smooth (or topological), closed, oriented 5–manifolds with free fundamental group of rank n and torsion-free π_2. Then M_0 and M_1 are oriented-diffeomorphic (-homeomorphic) if and only if their (topological) generalized Milnor pairings are isomorphic. Any isomorphism between the (topological) generalized Milnor parings can be realized by an orientation-preserving diffeomorphism (homeomorphism) from M_0 to M_1.*

If M_0 and M_1 are compact with boundary consisting of n copies of $S^1 \times S^3$ such that the circles in the boundary generate $\pi_1(M_i)$, then M_0 and M_1 are oriented-diffeomorphic (-homeomorphic) if and only if there exists an isomorphism (α, β) between their (topological) generalized Milnor pairings, where α is induced by identifying the boundary components. Any such isomorphism can be realized by an orientation-preserving diffeomorphism (homeomorphism).

The isomorphism α above actually sends free generators x_i of $\pi_1(M_0)$ to conjugates of free generators x_i' of $\pi_1(M_1)$, which are represented by different arcs in the interior to a basepoint.

Remark 2.5 In the definition of the invariant $\varphi(M)$ we use the cup product on the cohomology with rational coefficients. Usually one loses information when passing from integral coefficients to rational coefficients. But in our situation the rational cohomology contains essentially more information than the integral cohomology. This can be illuminated by the following example.

Example Let

$$A = \begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix};$$

then $A + A'$ (where A' is the transpose of A) is unimodular and has signature 0. Therefore by [11, Theorem 2] there is a simple 3–knot $K \subset S^5$ with Seifert matrix S–equivalent to A. The Alexander polynomial of K is $\Delta_K(t) = \det(A - tA') = 2t^2 + 5t + 2$. Let X be the complement of K; then, by [4, Theorem 1.5], $H_2(\widetilde{X}) \cong \mathbb{Z}\left[\frac{1}{2}\right] \oplus \mathbb{Z}\left[\frac{1}{2}\right]$. Let M^5 be the result of surgery on K; then $\pi_1(M) \cong \mathbb{Z}$ and $\pi_2(M) \cong H_2(\widetilde{M}) \cong H_2(\widetilde{X}) \cong \mathbb{Z}[1/2] \oplus \mathbb{Z}\left[\frac{1}{2}\right]$. We see that $H^2(\widetilde{M}; \mathbb{Z}) = 0$ but $H^2(\widetilde{M}; \mathbb{Q}) \cong \mathbb{Q}^2$.

3 Proof of Theorem 2.4

Before giving the proof of the main theorem we first rephrase the bilinear form b_M in a more explicit form. Fix an identification $\pi_1(M) \xrightarrow{\cong} F_n$ and consider the classifying map of the fundamental group $f\colon M \to BF_n = \bigvee_{i=1}^n S_i^1$. From the Leray–Serre spectral sequence (with twisted coefficients, which we denote by an underline) of the fibration $\widetilde{M} \to M \to \bigvee_{i=1}^n S_i^1$, we get an isomorphism $H_5(M) \to H_1\big(\bigvee_n S^1; \underline{H_4(\widetilde{M})}\big)$. Note that

$$H_1\Big(\bigvee_n S^1; \underline{H_4(\widetilde{M})}\Big) = \operatorname{Ker}\Big(\bigoplus_n H_4(\widetilde{M}) \xrightarrow{d} H_4(\widetilde{M})\Big),$$

where $d(x_1, \dots, x_n) = \sum_i (g_i - 1)x_i$ with $g_1, \dots, g_n$ the corresponding generators of F_n. This leads to an injection $H_5(M) \to \bigoplus_n H_4(\widetilde{M})$. Denote the image of the fundamental class $[M]$ by $(\sigma_1,\dots,\sigma_n)$. Now denote by $I_i(M)$ the symmetric bilinear form $H^2(\widetilde{M};\mathbb{Q})\times H^2(\widetilde{M};\mathbb{Q})\to\mathbb{Q}$ given by $I_i(\alpha,\beta)=\langle\alpha\cup\beta,\sigma_i\rangle$. From the relation $\sum_i(g_i-1)\sigma_i=0$ we see that the bilinear forms satisfy the relation $\sum_i I_i(\alpha,\beta)=\sum_i I_i(g_i^*\alpha,g_i^*\beta)$.

Geometrically, we choose regular values $q_i\in S^1_i$ and let $F_i=f^{-1}(q_i)$. Let E be the complement of an open tubular neighborhood of $\bigcup_i F_i$; then E has boundary $\partial E=\bigcup_i F_i^{\pm}$, where $F_i^{\pm}$ are the positive and negative boundary components of the tubular neighborhood of F_i. We obtain $\widetilde{M}$ by gluing infinitely many copies of E under the deck transformation, ie $\widetilde{M}=\bigcup_{g\in F_n}E_g$. Let $\overline{M}_i\to M$ be the $\mathbb{Z}$–covering of M corresponding to $M\to\bigvee_{i=1}^n S^1_i\to S^1_i$; then it's easy to see that the Leray–Serre spectral sequence of this covering gives an isomorphism $H_5(M)\xrightarrow{\cong}H_4(\overline{M}_i)$, with $[M]\mapsto[F_i^-]$. Furthermore the commutative diagram

$$\begin{tikzcd}
\widetilde{M} \arrow[rr] \arrow[dr] & & \overline{M}_i \arrow[dl] \\
& M &
\end{tikzcd}$$

induces

$$\begin{tikzcd}
& \bigoplus_n H_4(\widetilde{M}) \arrow[d, "\text{projection to the } i^{\text{th}} \text{ component}"] \\
H_5(M) \arrow[ur] \arrow[dr] & H_4(\widetilde{M})_i \arrow[d] \\
& H_4(\overline{M}_i)
\end{tikzcd}$$

From this we see that each σ_i is represented by F_i^- in $E\subset\widetilde{M}$.

By [19, Proposition 3.3] we know that M has a CW–structure of the form $M\simeq\bigvee_{i=1}^n S^1_i\vee\bigvee S^2\cup e^3\cup\cdots$. Therefore we have isomorphisms $H_4(\widetilde{M})\cong H^1_c(\widetilde{M})\cong H^1\bigl(\bigvee_n S^1,\Lambda\bigr)$, where Λ denotes the group ring $\mathbb{Z}[F_n]$. Thus we have a surjection $\Lambda^n\to H_4(\widetilde{M})$. Let e_i be the standard basis of Λ^n; then e_i is mapped to σ_i. Therefore $\sigma_1,\dots,\sigma_n$ form a set of generators of the Λ–module $H_4(\widetilde{M})$. For any $\alpha,\ \beta\in H^2(\widetilde{M};\mathbb{Q})$ and $x\in H_4(\widetilde{M})$, we may assume that $x=\sum_i\lambda_i\sigma_i$, with $\lambda_i=\sum_g a_g^{(i)}\cdot g\in\Lambda$. Then $\langle\alpha\cup\beta,x\rangle=\bigl\langle\alpha\cup\beta,\sum_i\lambda_i\sigma_i\bigr\rangle=\sum_{i,g}a_g^{(i)}\langle g^{-1}\alpha\cup g^{-1}\beta,\sigma_i\rangle=\sum_{i,g}a_g^{(i)}I_i(g^{-1}\alpha,g^{-1}\beta)$. Thus we have shown:

Lemma 3.1 *The sequence of bilinear forms $(I_1,\dots,I_n)$ contains the same information as the bilinear pairing b_M together with an identification of $\pi_1(M)$ with the free group F_n.*

Next we relate the signature of forms I_i to the signatures of the fiber F_i^4.

Lemma 3.2 *The bilinear form $I_i\colon H^2(\widetilde{M};\mathbb{Q}) \times H^2(\widetilde{M};\mathbb{Q}) \to \mathbb{Q}$ has the same signature as the intersection form of F_i^4.*

Proof We use homology and cohomology with $\mathbb{Q}$–coefficients.

Let E be the exterior of an open tubular neighborhood of $\bigcup_i F_i$. Then the universal cover $\widetilde{M}$ is $\widetilde{M} = \bigcup_{g \in F_n} E_g$, where each E_g is a copy of E. Since $H_2(F_i)$ is finite-dimensional, there exists a connected compact submanifold $M_0 \subset \widetilde{M}$ which is a union of finitely many of the E_g and $F_i \subset M_0$ such that any $x \in \operatorname{Ker}(H_2(F_i) \to H_2(\widetilde{M}))$ is in $\operatorname{Ker}(H_2(F_i) \to H_2(M_0))$. Therefore

$$\operatorname{Ker}(H_2(F_i) \to H_2(\widetilde{M})) = \operatorname{Ker}(H_2(F_i) \to H_2(M_0)).$$

Dually on cohomology, we have

$$\operatorname{Im}(H^2(\widetilde{M}) \to H^2(F_i)) = \operatorname{Im}(H^2(M_0) \to H^2(F_i)).$$

The boundary ∂M_0 has a component F_0 which is the image of F_i under a deck transformation by $g \in \pi_1(M)$. There is a commutative diagram

$$\begin{tikzcd} H^2(M_0) \arrow[r] \arrow[dr] & H^2(F_i) \arrow[d, "g^*"] \\ & H^2(F_0) \end{tikzcd}$$

where g^* is an isometry. So we have

$$\begin{aligned} H^2(\widetilde{M})/\mathrm{rad}(I_i) &= \operatorname{Im}(H^2(\widetilde{M}) \to H^2(F_i))/\mathrm{rad} \\ &= \operatorname{Im}(H^2(M_0) \to H^2(F_i))/\mathrm{rad} \\ &\cong \operatorname{Im}(H^2(M_0) \to H^2(F_0))/\mathrm{rad}. \end{aligned}$$

Note that $\operatorname{Ker}(H_2(\partial M_0) \to H_2(M_0))$ is a Lagrangian in $H_2(\partial M_0)$. Therefore

$$\operatorname{Ker}(H_2(F_0) \to H_2(M_0)) = \operatorname{Ker}(H_2(\partial M_0) \to H_2(M_0)) \cap H_2(F_0)$$

is isotropic. A standard argument in linear algebra shows that dually on cohomology, $\operatorname{Im}(H^2(M_0) \to H^2(F_0))$ has a complement which is isotropic. Let's denote it by K; it generates a hyperbolic form $H(K)$ in $H^2(F_0)$ and we have

$$\operatorname{Im}(H^2(M_0) \to H^2(F_0))/\mathrm{rad} \oplus H(K) = H^2(F_0).$$

Therefore $\operatorname{sign}(I_i) = \operatorname{sign}(H^2(F_0))$. □

The proof of Theorem 2.4 is based on modified surgery theory. We refer to [9] for the details of this machinery for classifying manifolds. For the convenience of the reader we summarize the basic concepts and the main theorem we apply. The basic idea is to weaken the normal homotopy type, which is the first basic invariant of a manifold M in classical surgery, to the normal k–type. This is roughly given by the k–skeleton of M together with the restriction of the normal bundle. Since the k–skeleton is not well-defined we pass to Postnikov towers instead or, better, Moore–Postnikov decompositions. The normal bundle is equivalent to the normal Gauss map $\nu\colon M \to BO$. The *normal k–type* is the k^{th} stage of the Moore–Postnikov tower of $\overline{\nu}$, which is a fibration $p\colon B_k(M) \to BO$ which is completely characterized by the property that there is a lift $\overline{\nu}\colon M \to B_k(M)$ of ν which induces an isomorphism on homotopy groups up to degree k and is surjective in degree $k+1$. Note that if k is larger than the dimension of M the normal k–type is equivalent to the normal homotopy type, thus modified surgery generalizes classical surgery. Such a lift is called a *normal k–smoothing*.

Given two normal k–smoothings $(M, \overline{\nu}_M)$ and $(M', \overline{\nu}_{M'})$ in the same fibration B_k, the first step is to decide whether these normal k–smoothings are bordant. This means that there is a coboundary W together with a lift of the normal Gauss map $\overline{\nu}_W$ (but this is not highly connected). The main theorem of modified surgery is that if $k \geq \frac{1}{2}\dim M - 1$, then there is a surgery obstruction in a monoid $l_{\dim M+1}(\pi_1(M), w_1(M))$, from which one can decide whether W is B_k–bordant to an s–cobordism.

Now we return to our situation of 5–manifolds. We will work with the normal 2–type of M. Then the obstruction is actually in the classical Wall group $L_5(\pi_1(M), w_1(M))$. We prepare the proof with a construction of the normal 2–type (cf [9, Proposition 2]) of a smooth manifold M (of arbitrary dimension), which might be of separate interest elsewhere. Let $u\colon M \to P$ be the second-stage Postnikov tower of M; there are unique cohomology classes $w_i \in H^i(P;\mathbb{Z}/2)$ for $i = 1, 2$ such that $u^*(w_i) = w_i(M)$. Let $w_1 \times w_2\colon P \to K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,2)$ be the classifying map of these classes, and $w_1(EO) \times w_2(EO)\colon BO \to K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,2)$ be the classifying map of the universal Stiefel–Whitney classes. Consider the following pullback square:

$$\begin{tikzcd}
B(\pi_1(M), \pi_2(M), k_1, w_1(M), w_2(M)) \arrow[r, "h"] \arrow[d, "p"] & P \arrow[d, "w_1 \times w_2"] \\
BO \arrow[r, "w_1(EO) \times w_2(EO)"] & K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,2)
\end{tikzcd}$$

There is a lift $\overline{\nu}\colon M \to B(\pi_1(M), \pi_2(M), k_1, w_1(M), w_2(M))$ of the normal Gauss map $\nu\colon M \to BO$ of M, which a 3–equivalence, and p is 3–coconnected. Thus we have shown:

Lemma 3.3 *The fibration*

$$p\colon B(\pi_1(M), \pi_2(M), k_1, w_1(M), w_2(M)) \to BO$$

is the normal 2–type of M. There is a corresponding construction in the topological category, if one replaces BO by $B\mathrm{Top}$.

Now we are ready to prove Theorem 2.4.

Proof of Theorem 2.4 We begin with the smooth category. In our situation, the second-stage Postnikov tower P of M is a fibration over $\bigvee_{i=1}^n S_i^1$ with fiber $K = K(\pi_2(M), 2)$ and monodromy given by the $\pi_1(M)$–module structure of $\pi_2(M)$. We denote the normal 2–type by $p\colon B \to BO$ and recall that by the lemma above it is determined by $\pi_1(M)$, $\pi_2(M)$ as a $\mathbb{Z}[\pi_1(M)]$–module, and $w_2(M)$.

Now we compute the bordism group $\Omega_5(B, p)$. Note that $\Omega_5(B, p) = \pi_5^S(M(p))$. We consider the fibration $\widetilde{B} \to B \to \bigvee_{i=1}^n S_i^1$; the Wang sequence of the generalized homology theory π_*^S is

$$\cdots \to \Omega_5(\widetilde{B}, \widetilde{p}) \to \Omega_5(B, p) \to \bigoplus_n \Omega_4(\widetilde{B}, \widetilde{p}) \to \cdots,$$

where $\widetilde{B}$ is the pullback

$$\begin{tikzcd} \widetilde{B} \arrow[rr] \arrow[d, "\widetilde{p}"] & & K \arrow[d, "\mathrm{const}\times w_2"] \\ BO \arrow[r] & K(\mathbb{Z}/2,1) \times K(\mathbb{Z}/2,2) & \end{tikzcd}$$

where $w_2 \in H^2(K;\mathbb{Z}/2)$ is the image of $w_2 \in H^2(P;\mathbb{Z}/2)$ under the injection $H^2(P;\mathbb{Z}/2) \to H^2(K;\mathbb{Z}/2)$. From this we have $\Omega_n(\widetilde{B}, \widetilde{p}) = \Omega_n^{\mathrm{spin}}(K;\eta)$, where the latter group is the bordism group of $f\colon M \to K$ together with a spin structure on $f^*\eta \oplus \nu M$, where η is a complex line bundle over K such that $w_2(\eta) = w_2 \in H^2(K;\mathbb{Z}/2)$.

Now $\pi_2(M)$ is the direct limit of its finitely generated subgroups, and by assumption $\pi_2(M)$ is a torsion-free abelian group, hence it is a direct limit of finitely generated free abelian groups $\varinjlim G_\alpha$. Therefore K is a direct limit of spaces $K = \varinjlim K(G_\alpha, 2)$. In general there is an Atiyah–Hirzebruch spectral sequence computing $\Omega_n^{\mathrm{spin}}(X;\eta)$ with E_2–terms $H_p(X;\Omega_q^{\mathrm{spin}})$, and the differential d_2 is dual to $\mathrm{Sq}^2 + w_2(\eta)\cdot$; see [18]. An easy computation with this spectral sequence shows that $\Omega_5^{\mathrm{spin}}(K(G_\alpha,2);\eta) = 0$ for a finitely generated free abelian group G_α, and henceforth $\Omega_5^{\mathrm{spin}}(K;\eta) = \varinjlim \Omega_5^{\mathrm{spin}}(K(G_\alpha,2);\eta) = 0$.

Therefore we have an injection $\Omega_5(B,p) \to \bigoplus_n \Omega_4^{\mathrm{spin}}(K;\eta)$. There is a commutative diagram

$$\begin{array}{ccc} \Omega_5(B,p) & \longrightarrow & \bigoplus_n \Omega_4^{\mathrm{spin}}(K;\eta) \\ \downarrow & & \downarrow \\ H_5(P) & \longrightarrow & \bigoplus_n H_4(K) \end{array}$$

with the horizontal arrows injective and the vertical arrows the edge homomorphisms. Following the definition of the boundary map in the Mayer–Vietoris sequence of the bordism theory, we see that a bordism class $[f\colon M \to B]$ is mapped to

$$([h\circ f\colon F_1 \to K],\dots,[h\circ f\colon F_n \to K]) \in \bigoplus_n \Omega_4^{\mathrm{spin}}(K;\eta),$$

where $h\colon B \to P$ is the map in the pullback square, $\pi\colon P \to \bigvee_n S^1$ is the projection map, and $F_i = (\pi\circ h\circ f)^{-1}(q_i)$ is the preimage of a regular value $q_i \in S^1_i$. A direct calculation with the Atiyah–Hirzebruch spectral sequence shows that a bordism class

$$[\varphi\colon N^4 \to K] \in \Omega_4^{\mathrm{spin}}(K(G_\alpha,2);\eta)$$

is determined by $\mathrm{sign}(N)$ and $\varphi_*[N] \in H_4(K(G_\alpha,2))$. Passing to the limit we see that a bordism class $[\varphi\colon N^4 \to K] \in \Omega_4^{\mathrm{spin}}(K;\eta)$ is determined by $\mathrm{sign}(N)$ and $\varphi_*[N] \in H_4(K)$. Now $H_4(K) = \varinjlim H_4(K(G_\alpha,2))$ is a direct limit of free abelian groups, hence is torsion-free, therefore $\varphi_*[N]$ is determined by its image in $H_4(K;\mathbb{Q})$, which is further determined by the evaluation with elements in $H^4(K;\mathbb{Q})$. Note that $H^4(K;\mathbb{Q}) = H^4(K(\pi_2(M)\otimes\mathbb{Q},2);\mathbb{Q})$, where $\pi_2(M)\otimes\mathbb{Q}$ is a $\mathbb{Q}$–vector space. From this it's easy to see that the cup product map

$$H^2(K;\mathbb{Q}) \otimes H^2(K;\mathbb{Q}) \xrightarrow{\cup} H^4(K;\mathbb{Q})$$

is surjective, therefore $\varphi_*[N] \in H_4(K;\mathbb{Q})$ is determined by $\langle \varphi^*\alpha \cup \varphi^*\beta, [N]\rangle$ for $\alpha,\beta \in H^2(K;\mathbb{Q})$.

For a normal 2–smoothing $\bar{\nu}\colon M \to B$, let $f\colon M \xrightarrow{\bar{\nu}} B \xrightarrow{h} P$ be the composition; we have a commutative diagram

$$\begin{array}{ccc} \widetilde{M} & \xrightarrow{\tilde{f}} & K \\ \downarrow & & \downarrow \\ M & \xrightarrow{f} & P \end{array}$$

and $f = \tilde{f}\circ i\colon F_i \to K$ for $F_i \subset \widetilde{M}$. Notice that $\tilde{f}^*\colon H^2(K;\mathbb{Q}) \to H^2(\widetilde{M};\mathbb{Q})$ is an isomorphism, therefore the evaluation $\langle f^*\alpha \cup f^*\beta, [F_i]\rangle = \langle \tilde{f}^*\alpha \cup \tilde{f}^*\beta, i_*[F_i]\rangle = \langle \tilde{f}^*\alpha \cup \tilde{f}^*\beta, \sigma_i\rangle$ is exactly the bilinear form $I_i\colon H^2(\widetilde{M};\mathbb{Q}) \otimes H^2(\widetilde{M};\mathbb{Q}) \to \mathbb{Q}$.

By Lemma 3.2, $\operatorname{sign}(F_i)$ equals the signature of the bilinear form I_i. This shows that the bordism class $[M, \overline{\nu}]$ is determined by the bilinear forms I_i for $i = 1, \dots, n$.

Now, given two manifolds M and M' with isomorphic algebraic invariants and — depending on an ordering of the boundary components in the bounded case — equal boundary as in Theorem 2.4, they have the same normal 2–type (B, p). We identify the boundaries (one of the manifolds with opposite orientation) to obtain a closed manifold and use the normal 2–smoothings $\overline{\nu}\colon M \to B$ and $\overline{\nu}'\colon M' \to B$ to obtain an element in $\Omega_5(B, p)$. We note here that we controlled the restriction of the normal 2–smoothings to the boundary by requiring that the identification of the boundary components be compatible with the identification of the fundamental groups. By the consideration above this is the zero element if our invariant φ agrees for M_0 and M_1 with the normal 2–smoothings chosen so that the invariants agree.

Let W be a B–null-bordism of the glued manifold, then there is an obstruction $\theta(W) \in l_6(F_n)$. If this is elementary, then W is B–bordant rel boundary to an s–cobordism [9, Theorem 3]. In our situation with $\pi_1(M) \cong F_n$, the Whitehead group $\mathrm{Wh}(F_n) = \bigoplus \mathrm{Wh}(\mathbb{Z}) = 0$, and so we won't have to consider the preferred bases. Furthermore by the remark on [9, page 730] the obstruction sits in the ordinary L–group $L_6(F_n)$. This group is isomorphic to $\mathbb{Z}/2$ and the obstruction is detected by the Arf-invariant [2, Theorem 16]. Since there is a simply connected closed 6–manifold with Arf-invariant 1 we can change W by disjoint sum with this, if necessary, to show that $\theta(W) = 0 \in L_6(F_n)$. This implies that $\theta(W)$ is elementary and finishes the proof in the smooth case.

The proof of the topological case is similar, since the modified surgery method also applies to topological manifolds (cf [9]). The only difference is that an element $[\varphi\colon F^4 \to K] \in \Omega_4^{\mathrm{TopSpin}}(K; \eta)$ is determined by the image of the fundamental class $\varphi_*[F] \in H_4(K)$, the signature $\operatorname{sign}(F)$ and the Kirby–Siebenmann invariant $KS(F)$. Each F_i has trivial normal bundle in M, therefore under the isomorphism $H^4(M; \mathbb{Z}/2) \xrightarrow{\cong} \bigoplus_{i=1}^n H^4(F_i; \mathbb{Z}/2)$, $KS(M)$ is mapped to

$$(KS(F_1), \dots, KS(F_n)).$$

The rest is the same as in the smooth case. □

4 Proofs of Theorems 1.2 and 1.3

The Seifert matrix of a boundary link is defined as follows (cf [12]): choose Seifert manifolds F_i of the link L; then there are linking forms

$$\theta_{ij}\colon H_q(F_i) \otimes H_q(F_j) \to \mathbb{Z}, \quad (\alpha, \beta) \mapsto L(z_1, z_2),$$

defined by linking numbers between z_1, representing α, and z_2, representing $i_+\beta$. With respect to a basis of the torsion-free part of $H_q(F_i)$ the linking forms θ_{ij} are represented by a matrix A_{ij}; then the Seifert matrix $D=(A_{ij})$ of L is an integral square matrix formed by the blocks A_{ij}, and D is $(-1)^q$–symmetric. Different choices of Seifert manifolds will lead to different Seifert matrices, but they are related by a sequence of "algebraic moves" and are *l–equivalent*. The l–equivalence class of the Seifert matrix D is a well-defined invariant of L [12, Theorem 1].

Given a square integral matrix $D=(A_{ij})$, consisting of matrices blocks A_{ij}, the *unimodularity condition* of D requires that $A_{ii}+A'_{ii}$ for $i=1,\dots,n$ and $D+D'$ are unimodular. It's shown in [12] that there is a boundary simple $(2q-1)$–link L whose Seifert matrix is $D=(A_{ij})$ when $q\geq 3$ [12, Theorem 1].

Given a link $f\colon \bigcup_{i=1}^n S^3 \hookrightarrow S^5$ we note that up to isotopy there is a unique tubular neighborhood U of Image(f). We denote the complement of the interior of this tubular neighborhood by X_f and use the tubular neighborhood to identify ∂X_f with $\bigcup_n(S^1\times S^3)$.

If two links $f\colon \bigcup_{i=1}^n S^3 \hookrightarrow S^5$ and $f'\colon \bigcup_{i=1}^n S^3 \hookrightarrow S^5$ are isotopic, the isotopy extension theorem implies that the identification $\partial X_f\to\partial X_{f'}$ extends to a diffeomorphism $X_f\to X_{f'}$. In turn, if there is an orientation-preserving diffeomorphism $g\colon X_f\to X_{f'}$ extending the identification on the boundary, then we can extend this by the identification on the tubular neighborhoods to an orientation-preserving diffeomorphism $\hat{g}\colon S^5\to S^5$ mapping the first link to the second. Now we use the fact that $\pi_0(\mathrm{Diff}^+(S^5))$ is isomorphic to the group of homotopy 6–spheres (using the h–cobordism theorem and Cerf's theorem [3] that pseudoisotopy implies isotopy) and that the group of 6–dimensional homotopy spheres is trivial [8]. Thus the two links are isotopic.

Now note that the link complement X has free fundamental group of rank n, generated by the circles in the boundary components. Furthermore, from Farber [5, Theorem 5.7] we know that π_2 of the complement of a simple boundary link is torsion-free. Thus Theorem 2.4 applies to complements of simple boundary 3–links in S^5.

The meridians give rise to an identification $\pi_1(X_f)\xrightarrow{\cong}F_n$; under this identification, by the reinterpretation of the invariants in the beginning of Section 3, we have an invariant

$$\psi(X_f)=(\pi_2(X_f),\, b_i\colon \pi_2(X_f)^*\times\pi_2(X_f)^*\to\mathbb{Q},\ i=1,\dots,n).$$

Here we consider $\pi_2(X_f)$ as an F_n–module and $*$ stands for the $\mathbb{Q}$–dual. The link complement X_f is a Spin–manifold, thus Theorem 2.4 implies that this invariant determines the oriented diffeomorphism type mod boundary, meaning that the identification on the boundary extends to an orientation-preserving diffeomorphism between the whole manifolds. Thus we have proved the following:

Lemma 4.1 *Two simple boundary* 3*–links* $f\colon \bigcup_n S^3 \hookrightarrow S^5$ *and* $f'\colon \bigcup_n S^3 \hookrightarrow S^5$ *are isotopic if and only if under certain identifications of* $\pi_1(X_f)$ *and* $\pi_1(X_{f'})$ *with* F_n *coming from enumerating the link components,* $\psi(X_f)$ *and* $\psi(X_{f'})$ *are isomorphic.*

Proof of Theorem 1.2 By Lemma 4.1, to prove that the l–equivalence class of the Seifert matrices determines the isotopy type of the link, we need to show that the l–equivalence class of the Seifert matrices determines $\psi(X_f)$. Let F_i be Seifert manifolds of a link given by an embedding f. Let X_f be the complement of the tubular neighborhood of the link; then the universal cover $\widetilde{X}_f$ is obtained by gluing infinitely many copies of Y via the deck transformation, where Y is obtained from X_f by cutting up along the Seifert manifolds. We identify $\pi_1(X_f)$ with F_n by sending the meridian (with the induced orientation from that of S^5 and S^3) of the i^{th} component of the link to the i^{th} standard generator t_i of F_n. The Mayer–Vietoris sequence computing $H_2(\widetilde{X}_f)$ is

$$\bigoplus_{i=1}^{n} H_2(F_i)\otimes_{\mathbb{Z}}\mathbb{Z}[F_n] \xrightarrow{\varphi} H_2(Y)\otimes_{\mathbb{Z}}\mathbb{Z}[F_n] \to H_2(\widetilde{X}_f) \to 0,$$

where, under the basis of $H_2(F_i)$ and the Alexander dual basis of $H_2(Y)$, φ is given by $(A_{ij} - t_i A'_{ij})$. Therefore the $\mathbb{Z}[F_n]$–module $H_2(\widetilde{X}_f)$ is determined by $D=(A_{ij})$. Also we see that the map $H_2(F_i) \to H_2(\widetilde{X}_f)$ is determined by $D=(A_{ij})$, hence the dual map $H^2(\widetilde{X}_f;\mathbb{Q}) \to H^2(F_i;\mathbb{Q})$. And the intersection form of F_i is given by $A_{ii}+A'_{ii}$. It's easy to see from the definition that the bilinear pairing b_i is given by the composition of $H^2(\widetilde{X}_f;\mathbb{Q}) \to H^2(F_i;\mathbb{Q})$ with the intersection form on $H^2(F_i;\mathbb{Q})$. Therefore the bilinear form b_i is determined by the Seifert matrix $D=(A_{ij})$.

Given two simple boundary 3–links L_0 and L_1, with l–equivalent Seifert matrices $D_0 = (A^{(0)}_{ij})$ and $D_1 = (A^{(1)}_{ij})$, then by [12, Lemma 1] we may choose Seifert manifolds $\{F^0_i\}$ and $\{F^1_i\}$ of L_0 and L_1, respectively, such that the corresponding Seifert matrices are equal. Then, by the above discussion, L_0 and L_1 are equivalent.

Using a stabilization trick introduced by Levine in the case of knots, we can extend the construction of links with given Seifert matrix in [12] to the case $q=2$. The construction goes as follows.

Firstly, by [11, Lemma 16] we may find embeddings $F^4_i \subset B^5_i \subset S^5$ with $\partial F_i = S^3$ a simple 3–knot whose Seifert matrix A_i is S–equivalent to A_{ii}. After stabilization by connected sum with copies of $S^2 \times S^2$, these Seifert manifolds F^4_i are diffeomorphic to connected sums of $S^2\times S^2$ and the Kummer surface with a 4–ball B^4 deleted. These smooth 4–manifolds all have a handle decomposition of the form $F_i = D^4 \cup h_1 \cup \cdots \cup h_k$ where the h_i are 2–handles (see eg [14]). Then by the same argument in the proof of [12, Theorem 1] we can show that the new Seifert matrix D', which is l–equivalent to D, is the Seifert matrix of a boundary simple 3–link L. □

Now we describe the Milnor pairing associated to an $m \times m$ integral matrix $D = (A_{ij})$ satisfying the unimodularity conditions. Let $\varphi_D\colon (\mathbb{Z}[F_n])^m \to (\mathbb{Z}[F_n])^m$ be the $\mathbb{Z}[F_n]$–module map given by the matrix $(A_{ij} - t_i A'_{ij})$. Assume the square matrix A_{ii} has dimension m_i; then $A_{ii} + A'_{ii}$ defines a symmetric bilinear form I_i on $\mathbb{Z}^{m_i}$. Let ι_i be the composition

$$\iota_i\colon \mathbb{Z}^{m_i} \xrightarrow{\bigoplus_j A_{ij}} \bigoplus_j \mathbb{Z}^{m_j} \to \bigoplus_j \mathbb{Z}^{m_j} \otimes_{\mathbb{Z}} \mathbb{Z}[F_n] = (\mathbb{Z}[F_n])^m \to \operatorname{coker} \varphi_D$$

The $\mathbb{Q}$–dual of ι_i is $\iota_i^*\colon (\operatorname{coker}\varphi_D)^* \to \mathbb{Q}^{m_i}$. Let $C_1 = (\mathbb{Z}[F_n])^n \xrightarrow{d_1} C_0 = \mathbb{Z}[F_n]$ be the standard chain complex computing $H_*(BF_n; \mathbb{Z}[F_n])$, $\{e_i \mid i = 1, \dots, n\}$ be the standard basis of $(\mathbb{Z}[F_n])^n$, $\{e_i^* \mid i = 1, \dots, n\}$ be the dual basis, and $[e_i^*] \in H^1(BF_n; \mathbb{Z}[F_n])$ be the corresponding cohomology class. Then the bilinear form

$$b_D\colon (\operatorname{coker}\varphi_D)^* \times (\operatorname{coker}\varphi_D)^* \to (H^1(BF_n; \mathbb{Q}[F_n]))^*$$

is given by $\langle b_D(u,v), [e_i^*]\rangle = I_i(\iota_i^*(u), \iota_i^*(v))$. (See Lemma 3.1.)

Proof of Theorem 1.3 There is a surjective map from the set of isotopy classes of simple boundary n–components links $L \subset S^5$ to the set of diffeomorphism classes of smooth oriented closed 5–manifolds M^5 with free fundamental group of rank n and $H_2(M;\mathbb{Z}) = 0$. This is given by surgery: given a link L, we may do surgery on L and obtain a 5–manifold M with $H_2(M;\mathbb{Z}) = 0$. If L is simple boundary, then it's easy to see that $\pi_1(M)$ is isomorphic to F_n. The meridians of the link components form an embedding $\bigcup_n S^1 \subset M$, and these circles generate $\pi_1(M)$. On the other hand, given such an M^5 we may choose an embedding $\bigcup_n S^1 \subset M^5$ such that the circle generate $\pi_1(M)$. Then we do surgery on this embedding and obtain S^5; the complementary spheres $\bigcup_n S^3 \subset S^5$ form a link L. Clearly this is a simple boundary link.

By comparing the definitions, we see that the generalized Milnor pairing $\varphi(M)$ of M is the same as the generalized Milnor pairing $\psi(X_f)$ of the link complement defined before Lemma 4.1. In the proof of Theorem 1.2 we have shown how the generalized Milnor pairing $\psi(X_f)$ is determined by the Seifert matrix $D = (A_{ij})$. This is exactly $(F_n, \operatorname{coker}\varphi_D, b_D)$, which was described before the proof of Theorem 1.3. By Theorem 1.2, all such matrices satisfying the unimodular conditions are realized by simple boundary links. This finishes the proof. □

Appendix

In this appendix we show some basic properties of the class of manifolds mentioned in Remark 1.4, ie oriented closed 5–manifolds M with $\pi_1(M) \cong \mathbb{Z}$ and $\pi_2(M)$ a finitely generated abelian group.

Lemma A.1 *Let M^5 be a 5–manifold with $\pi_1(M) = \mathbb{Z}$ and $\pi_2(M)$ a finitely generated abelian group; then all higher homotopy groups $\pi_i(M)$ for $i \geq 2$ are finitely generated abelian groups.*

Proof By Serre's mod $\mathscr{C}$ theory [16], we only need to show that $H_i(\widetilde{M})$ for $i \geq 3$ are finitely generated abelian groups. The only problem is $H_3(\widetilde{M})$. We have $H_3(\widetilde{M}) = H_3(M;\Lambda) \cong H^2(M;\Lambda)$, where $\Lambda = \mathbb{Z}[\mathbb{Z}] = \mathbb{Z}[t,t^{-1}]$ is the group ring. By [19, Proposition 3.3], the CW–structure of M has the form

$$M = S^1 \vee \Big(\bigvee S^2\Big) \cup \cdots .$$

Therefore the cellular chain complex $C_*(M;\Lambda)$ has the form

$$\cdots \to C_3 \xrightarrow{d} C_2 \xrightarrow{0} C_1 \to C_0.$$

From the exact sequence $C_3 \xrightarrow{d} C_2 \to \operatorname{coker} d \to 0$ we have the dual exact sequence $0 \to (\operatorname{coker} d)^* \to C_2^* \xrightarrow{d^*} C_3^*$, hence $H^2(M;\Lambda) = \ker d^* = (\operatorname{coker} d)^*$. Now $\operatorname{coker} d = H_2(M;\Lambda) = \pi_2(M)$ is a finitely generated abelian group; the proof is done given the following lemma. □

Lemma A.2 *If a Λ–module G is a finitely generated abelian group, then*

$$\operatorname{Hom}_\Lambda(G,\Lambda) = 0.$$

Proof The torsion subgroup T is a sub-Λ–module and the exact sequence

$$0 \to T \to G \to G/T \to 0$$

induces an exact sequence

$$0 \to \operatorname{Hom}_\Lambda(G/T,\Lambda) \to \operatorname{Hom}_\Lambda(G,\Lambda) \to \operatorname{Hom}_\Lambda(T,\Lambda),$$

therefore $\operatorname{Hom}_\Lambda(G/T,\Lambda) \cong \operatorname{Hom}_\Lambda(G,\Lambda)$ since $\operatorname{Hom}_\Lambda(T,\Lambda) = 0$. Therefore we may assume that G is a finitely generated free abelian group.

Let $x_1,\dots,x_n$ be a basis of G; a Λ–module structure on G is given by $A \in \mathrm{GL}_n(\mathbb{Z})$, which specifies the action of the generator t on the basis. A Λ–homomorphism $G \to \Lambda$ is given by the images $v_1,\dots,v_n \in \Lambda$ of $x_1,\dots,x_n$. The n–tuple $v = (v_1,\dots,v_n)$ should satisfy the equation $(tI - A)v = 0$. Clearly $\det(tI - A) \neq 0$, thus the equation has no nonzero solution in the quotient field (Λ is an integral domain), hence also has no nonzero solution in Λ. Therefore $\mathrm{Hom}_\Lambda(G,\Lambda)=0$. □

Now let M^5 be a closed orientable 5–manifold with $\pi_1(M)=\mathbb{Z}$ and $\pi_2(M)$ a finitely generated abelian group. Fix an orientation of M and a generator t of $\pi_1(M)$; these choices determine a generator (a fundamental class) $\sigma_M\in H_4(\widetilde{M})=\mathbb{Z}$. Then, on the finitely generated free abelian group $H^2(\widetilde{M})$, a symmetric bilinear form $H^2(\widetilde{M})\times H^2(\widetilde{M})\to\mathbb{Z}$ is defined by $(\alpha,\beta)\mapsto\langle\alpha\cup\beta,\sigma_M\rangle$. By the following proposition, we see that this bilinear form is unimodular and $\pi_2(M)$ is a free abelian group. Thus this bilinear form induces a symmetric bilinear form on $\pi_2(M)=\pi_2(\widetilde{M})=H_2(\widetilde{M})=H^2(\widetilde{M})^*$, denoted by $I(M)$.

Proposition A.3 *Let M^5 be an orientable 5–manifold with $\pi_1(M)=\mathbb{Z}$ and $\pi_2(M)$ a finitely generated abelian group. Then we have the following:*

(1) *$\pi_2(M)$ is torsion-free.*

(2) *The symmetric bilinear form $I(M)$ is unimodular; $I(M)$ is even if and only if $w_2(M)=0$.*

(3) *$\langle p_1(M),\sigma_M\rangle=3\cdot\mathrm{sign}(I(M))$, where $\sigma_M\in H_4(\widetilde{M})$ is the generator determined by the orientation of M and the generator t of $\pi_1(M)$.*

Proof Consider $M\times\mathbb{C}\mathrm{P}^2$. By Lemma A.1 and Browder and Levine's fibration theorem [1], we know that this manifold is a fiber bundle over S^1 with simply connected fiber F^8. Therefore $\widetilde{M}\times\mathbb{C}\mathrm{P}^2$ is homotopy equivalent to F.

(1) By the Künneth formula and Poincaré duality, we have

$$H^3(\widetilde{M})\cong H^7(\widetilde{M}\times\mathbb{C}\mathrm{P}^2)\cong H^7(F)\cong H_1(F)=0.$$

This proves that $\mathrm{tors}\,\pi_2(M)=\mathrm{tors}\,H_2(\widetilde{M})=\mathrm{tors}\,H^3(\widetilde{M})=0$.

(2) On $H^4(\widetilde{M}\times\mathbb{C}\mathrm{P}^2)$ there is defined a symmetric bilinear form $I(M\times\mathbb{C}\mathrm{P}^2)$, which is isometric to the tensor product of $I(M)$ and the intersection form of $\mathbb{C}\mathrm{P}^2$ plus a hyperbolic form. On the other hand, the bilinear form $I(M\times\mathbb{C}\mathrm{P}^2)$ is isometric to the intersection form of F, which is unimodular by Poincaré duality. Therefore the bilinear form $I(M)$ is unimodular.

From the discussion above we see that $I(M)$ is even if and only if the Wu class $v_4(F)$ is zero. The Wu classes and Stiefel–Whitney classes of M and F are related as follows. Let $i\colon F\to M\times\mathbb{C}\mathrm{P}^2$ be the inclusion of the fiber; then $TF\oplus\underline{\mathbb{R}}=i^*T(M\times\mathbb{C}\mathrm{P}^2)$. We have

$$w_2(M)=v_2(M),\quad w_3(M)=\mathrm{Sq}^1w_2(M),\quad w_4(M)=w_2(M)^2,$$
$$v_2(F)=w_2(F)=i^*(w_2(M)+w_2(\mathbb{C}\mathrm{P}^2)),\quad w_3(F)=\mathrm{Sq}^1w_2(F)+v_3(F);$$

on the other hand, $w_3(F) = i^* w_3(M)$, from which we have

$$v_3(F) = i^*(\mathrm{Sq}^1 w_2(M) + w_3(M)).$$

By the Wu formula

$$w_4(F) = v_2(F)^2 + \mathrm{Sq}^1 v_3(F) + v_4(F);$$

on the other hand,

$$w_4(F) = i^*(w_4(M) + w_2(M) w_2(\mathbb{C}\mathrm{P}^2) + w_4(\mathbb{C}\mathrm{P}^2)).$$

Comparing these two equations we have

$$v_4(F) = i^*(w_2(M) w_2(\mathbb{C}\mathrm{P}^2)).$$

But $H^3(F;\mathbb{Z}_2) \cong H^3(\widetilde{M} \times \mathbb{C}\mathrm{P}^2;\mathbb{Z}_2) \cong H^3(\widetilde{M};\mathbb{Z}_2) = 0$ (the last identity is a consequence of the fact that $H_2(\widetilde{M})$ is free and $H_3(\widetilde{M}) = 0$; see Lemma A.1). From the Wang sequence we see that $i^*\colon H^4(M \times \mathbb{C}\mathrm{P}^2;\mathbb{Z}_2) \to H^4(F;\mathbb{Z}_2)$ is injective. Thus $v_4(F) = 0$ if and only if $w_2(M) = 0$.

(3) Since $I(M)$ and $I(M \times \mathbb{C}\mathrm{P}^2)$ differ by a hyperbolic form, we have

$$\mathrm{sign}(I(M)) = \mathrm{sign}(I(M \times \mathbb{C}\mathrm{P}^2)) = \mathrm{sign}(F) = \tfrac{1}{45}\langle 7p_2(F) - p_1(F)^2, [F]\rangle,$$

where the last identity is the Hirzebruch index formula. Since F has trivial normal bundle in $M \times \mathbb{C}\mathrm{P}^2$, we have

$$p_1(F) = i^* p_1(M \times \mathbb{C}\mathrm{P}^2) = i^*(p_1(M) + p_1(\mathbb{C}\mathrm{P}^2)),$$
$$p_2(F) = i^* p_2(M \times \mathbb{C}\mathrm{P}^2) = i^*(p_1(M) p_1(\mathbb{C}\mathrm{P}^2)).$$

A straightforward calculation shows that $3\,\mathrm{sign}(I(M)) = \langle p_1(M), \sigma_M\rangle$. □

References

[1] **W Browder**, **J Levine**, *Fibering manifolds over a circle*, Comment. Math. Helv. 40 (1966) 153–160 MR

[2] **S E Cappell**, *Mayer–Vietoris sequences in hermitian K–theory*, from "Algebraic K–theory, III: Hermitian K–theory and geometric applications" (H Bass, editor), Lecture Notes in Math. 343, Springer (1973) 478–512 MR

[3] **J Cerf**, *La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie*, Inst. Hautes Études Sci. Publ. Math. 39 (1970) 5–173 MR

[4] **R H Crowell**, *The group G'/G'' of a knot group G*, Duke Math. J. 30 (1963) 349–354 MR

[5] **M Farber**, *Stable-homotopy and homology invariants of boundary links*, Trans. Amer. Math. Soc. 334 (1992) 455–477 MR

[6] **M H Freedman**, **F Quinn**, *Topology of 4–manifolds*, Princeton Math. Series 39, Princeton Univ. Press (1990) MR

[7] **F Hsu**, PhD thesis, Michigan State University (1985)

[8] **M A Kervaire**, **J W Milnor**, *Groups of homotopy spheres, I*, Ann. of Math. 77 (1963) 504–537 MR

[9] **M Kreck**, *Surgery and duality*, Ann. of Math. 149 (1999) 707–754 MR

[10] **M Kreck**, **W Lück**, *Topological rigidity for non-aspherical manifolds*, Pure Appl. Math. Q. 5 (2009) 873–914 MR

[11] **J Levine**, *An algebraic classification of some knots of codimension two*, Comment. Math. Helv. 45 (1970) 185–198 MR

[12] **C C Liang**, *An algebraic classification of some links of codimension two*, Proc. Amer. Math. Soc. 67 (1977) 147–151 MR

[13] **R A Litherland**, *Cobordism of satellite knots*, from "Four-manifold theory" (C Gordon, R Kirby, editors), Contemp. Math. 35, Amer. Math. Soc., Providence, RI (1984) 327–362 MR

[14] **R Mandelbaum**, *Special handlebody decompositions of simply connected algebraic surfaces*, Proc. Amer. Math. Soc. 80 (1980) 359–362 MR

[15] **J W Milnor**, *Infinite cyclic coverings*, from "Conference on the topology of manifolds" (J G Hocking, editor), Prindle, Weber & Schmidt, Boston (1968) 115–133 MR

[16] **J-P Serre**, *Groupes d'homotopie et classes de groupes abéliens*, Ann. of Math. 58 (1953) 258–294 MR

[17] **J L Shaneson**, *Non-simply-connected surgery and some results in low dimensional topology*, Comment. Math. Helv. 45 (1970) 333–352 MR

[18] **P Teichner**, *On the signature of four-manifolds with universal covering spin*, Math. Ann. 295 (1993) 745–759 MR

[19] **C T C Wall**, *Finiteness conditions for CW–complexes*, Ann. of Math. 81 (1965) 56–69 MR

[20] **S Weinberger**, *On fibering four- and five-manifolds*, Israel J. Math. 59 (1987) 1–7 MR

Hausdorff Research Institute for Mathematics, Universität Bonn
D-53115 Bonn, Germany
Institute of Mathematics, Chinese Academy of Sciences
Bejing, 100190, China

kreck@mathi.uni-bonn.de, suyang@math.ac.cn

Proposed: András I. Stipsicz — Received: 9 February 2016
Seconded: Peter Teichner, Rob Kirby — Revised: 22 November 2016

Geometry & Topology 21 (2017) 3009–3045

On the Fano variety of linear spaces contained in two odd-dimensional quadrics

CAROLINA ARAUJO
CINZIA CASAGRANDE

We describe the geometry of the $2m$–dimensional Fano manifold G parametrizing $(m-1)$–planes in a smooth complete intersection Z of two quadric hypersurfaces in the complex projective space $\mathbb{P}^{2m+2}$ for $m \geq 1$. We show that there are exactly 2^{2m+2} distinct isomorphisms in codimension one between G and the blow-up of $\mathbb{P}^{2m}$ at $2m+3$ general points, parametrized by the 2^{2m+2} distinct m–planes contained in Z, and describe these rational maps explicitly. We also describe the cones of nef, movable and effective divisors of G, as well as their dual cones of curves. Finally, we determine the automorphism group of G.

These results generalize to arbitrary even dimension the classical description of quartic del Pezzo surfaces ($m = 1$).

14E30, 14J45; 14M15, 14N20, 14E05

1 Introduction

In this paper we describe the geometry of the $2m$–dimensional Fano manifold $G^{(2m)}$ parametrizing $(m-1)$–planes in a smooth complete intersection of two quadric hypersurfaces in the complex projective space $\mathbb{P}^{2m+2}$ for $m \geq 1$. The case $m = 1$ is classical:

1.1 The surface $S = G^{(2)}$ is itself a smooth complete intersection of two quadric hypersurfaces in $\mathbb{P}^4$, and hence a quartic del Pezzo surface. It is well-known that $\rho(S) = 6$, and that the cone of effective curves of S is generated by the classes of its 16 lines. These 16 lines have a very special incidence relation: each line intersects properly exactly 5 lines. The del Pezzo surface S can also be described as the blow-up of $\mathbb{P}^2$ at 5 points in general linear position. In fact, there are 16 different ways to realize S as this blow-up: For every line $\ell \subset S$, there is a birational morphism $\pi_\ell\colon S \to \mathbb{P}^2$, unique up to projective transformation of $\mathbb{P}^2$, contracting the 5 lines incident to ℓ to points $p_1^\ell, \dots, p_5^\ell \in \mathbb{P}^2$ in general linear position. The image of ℓ under π_ℓ is the unique conic through the p_i, and the image of the other 10 lines are the 10 lines through 2 of the p_i. Moreover, for any two lines $\ell, \ell' \subset S$, the sets of points $\{p_1^\ell, \dots, p_5^\ell\}$ and $\{p_1^{\ell'}, \dots, p_5^{\ell'}\}$ are related by a projective transformation of $\mathbb{P}^2$.

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.3009

The automorphism group $\mathrm{Aut}(S)$ of S is also well understood (see for instance Dolgachev [12, Section 8.6.4]). In order to describe it, we view $\mathrm{Pic}(S)$ with the intersection product as a unimodular lattice. Its primitive sublattice $K_S^{\perp}$ is a D_5–lattice. We denote by $W(D_5)$ the Weyl group of automorphisms of this lattice. For any $\zeta \in \mathrm{Aut}(S)$, the induced isomorphism $\zeta^*\colon \mathrm{Pic}(S) \to \mathrm{Pic}(S)$ preserves the intersection product and fixes K_S. This yields an inclusion of groups $\mathrm{Aut}(S) \hookrightarrow W(D_5) \cong (\mathbb{Z}/2\mathbb{Z})^4 \rtimes S_5$, whose image contains the normal subgroup $(\mathbb{Z}/2\mathbb{Z})^4$. Moreover, if S is general, then $\mathrm{Aut}(S) \cong (\mathbb{Z}/2\mathbb{Z})^4$.

We will show that the picture described in Section 1.1 above generalizes to arbitrary even dimension. We start by fixing some notation. Let m be a positive integer, set $n = 2m$ and fix $n+3$ distinct points in $\mathbb{P}^1$, up to order and projective equivalence,

$$(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1) \in \mathbb{P}^1.$$

With this fixed data, we introduce the two main characters of this paper, $G^{(n)}$ and $X^{(n)}$:

1.2 ($G^{(n)}$) Let $Z^{(n)}$ be a smooth complete intersection of the two quadric hypersurfaces in $\mathbb{P}^{n+2}$

$$Q_1 : \sum_{i=1}^{n+3} x_i^2 = 0 \quad \text{and} \quad Q_2 : \sum_{i=1}^{n+3} \lambda_i x_i^2 = 0.$$

(Up to projective transformation of $\mathbb{P}^{n+2}$, any smooth complete intersection of two quadric hypersurfaces can be written in this way; see Section 2.) Then consider the subvariety $G^{(n)}$ of the Grassmannian $\mathrm{Gr}(m-1, \mathbb{P}^{n+2})$ parametrizing $(m-1)$–planes contained in $Z^{(n)}$. It is well known that $G^{(n)}$ is a smooth n–dimensional Fano variety with Picard number $\rho(G^{(n)}) = n+4$ (see Section 3 and references therein).

1.3 ($X^{(n)}$) Fix a Veronese embedding $\nu_n\colon \mathbb{P}^1 \hookrightarrow \mathbb{P}^n$, and set $p_i = \nu_n((\lambda_i : 1)) \in \mathbb{P}^n$. The points $p_1, \dots, p_{n+3}$ are in general linear position. (In fact, this gives a natural correspondence between sets of $n+3$ distinct points in $\mathbb{P}^1$, up to projective equivalence, and $n+3$ points in general linear position in $\mathbb{P}^n$, up to projective equivalence.) Let $X^{(n)}$ be the blow-up of $\mathbb{P}^n$ at the points $p_1, \dots, p_{n+3}$.

Our starting point is the following:

1.4 Theorem (Bauer [3], Casagrande [8]) *The varieties $G^{(n)}$ and $X^{(n)}$ are isomorphic in codimension* 1.

The proof of Theorem 1.4 makes use of moduli spaces of parabolic vector bundles. By [8], $G^{(n)}$ is isomorphic to the moduli space $\mathcal{M}^{(n)}$ of stable rank 2 parabolic vector bundles on $(\mathbb{P}^1, (\lambda_1 : 1), \dots, (\lambda_{n+3} : 1))$ of degree zero and weights $\left(\frac{1}{2}, \dots, \frac{1}{2}\right)$. On the other hand, by [3] (see also Mukai [22, Theorem 12.56]), $X^{(n)}$ is isomorphic to the moduli space of stable rank 2 parabolic vector bundles on $(\mathbb{P}^1, (\lambda_1 : 1), \dots, (\lambda_{n+3} : 1))$ of degree zero and weights $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$, which is isomorphic to $\mathcal{M}^{(n)}$ in codimension 1.

This proof, however, does not give much information about the possible isomorphisms in codimension 1 between $G^{(n)}$ and $X^{(n)}$. We call an isomorphism in codimension 1 a *pseudoisomorphism*. In this paper we describe explicitly the birational maps $G^{(n)} \dashrightarrow \mathbb{P}^n$ inducing a pseudoisomorphism $G^{(n)} \dashrightarrow X^{(n)}$. As we shall see, up to automorphism of $\mathbb{P}^n$, there are exactly 2^{n+2} distinct such birational maps, parametrized by the 2^{n+2} linear copies of $\mathbb{P}^m$ contained in $Z^{(n)}$. In order to state this precisely, we need to recall some facts about $Z^{(n)}$ (see Section 2 and references therein).

The set $\mathcal{F}_m(Z^{(n)})$ of m–planes in $Z^{(n)}$ has cardinality 2^{n+2}. For each $i = 1, \dots, n+3$, consider the involution $\sigma_i \colon Z^{(n)} \to Z^{(n)}$ switching the sign of the coordinate x_i. The group generated by these involutions is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{n+2}$, and acts on $\mathcal{F}_m(Z^{(n)})$ freely and transitively. For every subset $I \subseteq \{1, \dots, n+3\}$, we set $\sigma_I := \prod_{i \in I} \sigma_i = \prod_{j \in I^c} \sigma_j$. For every $M \in \mathcal{F}_m(Z^{(n)})$ and $I \subset \{1, \dots, n+3\}$ with $|I| \le m+1$, we have $\dim(M \cap \sigma_I(M)) = m - |I|$. Consider the incidence variety $\mathcal{I} := \{([L], p) \in G^{(n)} \times Z^{(n)} \mid p \in L\}$ and the associated diagram:

$$\begin{tikzcd} & \mathcal{I} \arrow[dl, "\pi"'] \arrow[dr, "e"] & \\ G^{(n)} & & Z^{(n)} \end{tikzcd}$$

We show that for every m–plane $M \in \mathcal{F}_m(Z^{(n)})$, $E_M := \pi_*(e^*(M))$ is the class of a unique prime divisor on $G^{(n)}$, which we denote by the same symbol (see Proposition 5.5).

Now we can state our main result. See Theorem 5.7 for more details, including explicit descriptions of the linear systems on $G^{(n)}$ defining the birational maps $G^{(n)} \dashrightarrow \mathbb{P}^n$.

1.5 Theorem (Theorem 5.7 and Corollary 5.8) *Let $M \in \mathcal{F}_m(Z^{(n)})$, in the notation above. Up to a unique permutation of the p_i, there is a unique birational map $\rho_M \colon G^{(n)} \dashrightarrow \mathbb{P}^n$, inducing a pseudoisomorphism $G^{(n)} \dashrightarrow X^{(n)}$, with the following properties:*

- *The image of E_M under ρ_M is $\mathrm{Sec}_{m-1}(C)$, the $(m-1)^{st}$ secant variety of the unique rational normal curve C through $p_1, \dots, p_{n+3}$ in $\mathbb{P}^n$.*
- *The map ρ_M contracts $E_{\sigma_i(M)}$ to the point $p_i \in \mathbb{P}^n$.*
- *For each $I \subseteq \{1, \dots, n+3\}$ of even cardinality $|I| \le n$, the image of $E_{\sigma_I(M)}$ under ρ_M is the join of $\langle p_i \rangle_{i \in I}$ and $\mathrm{Sec}_{s-1}(C)$, where $s = \frac{1}{2}(n - |I|)$.*

Moreover, any pseudoisomorphism between $G^{(n)}$ and any blow-up $\widetilde{X}$ of $\mathbb{P}^n$ at $n+3$ points is of this form. In particular, $\widetilde{X} \cong X^{(n)}$.

As immediate corollaries of Theorem 1.5, we obtain the following:

1.6 Corollary *Let $\mathcal{P}_1, \mathcal{P}_2 \subset \mathbb{P}^n$ be subsets of $n+3$ distinct points and let $X_{\mathcal{P}_i}$ be the blow-up of $\mathbb{P}^n$ along $\mathcal{P}_i$ for $i = 1,2$. Assume that the points in $\mathcal{P}_1$ are in general linear position. Then the following are equivalent:*

(i) $X_{\mathcal{P}_1} \cong X_{\mathcal{P}_2}$.

(ii) *$X_{\mathcal{P}_1}$ and $X_{\mathcal{P}_2}$ are pseudoisomorphic.*

(iii) *$\mathcal{P}_1$ and $\mathcal{P}_2$ are projectively equivalent (as unordered sets).*

1.7 Corollary *Let $\mathcal{S}_i = \{(\lambda_1^i : 1), \dots, (\lambda_{n+3}^i : 1)\} \subset \mathbb{P}^1$ for $i = 1,2$ be subsets of $n+3$ distinct points. For each $i \in \{1,2\}$, let $Z_{\mathcal{S}_i} \subset \mathbb{P}^{n+2}$ be the smooth complete intersection of the two quadrics*

$$Q_1 : \sum_{j=1}^{n+3} x_j^2 = 0 \quad \text{and} \quad Q_2^i : \sum_{j=1}^{n+3} \lambda_j^i x_j^2 = 0,$$

and let $G_{\mathcal{S}_i}$ be the variety of $(m-1)$–planes contained in $Z_{\mathcal{S}_i}$. Then the following are equivalent:

(i) $G_{\mathcal{S}_1} \cong G_{\mathcal{S}_2}$.

(ii) *$G_{\mathcal{S}_1}$ and $G_{\mathcal{S}_2}$ are pseudoisomorphic.*

(iii) *$\mathcal{S}_1$ and $\mathcal{S}_2$ are projectively equivalent (as unordered sets).*

Notice that Corollary 1.6 is a classical result, originally due to Coble (see Dolgachev and Ortland [14]). See also Biswas, Holla and Kumar [4] for a result related to Corollary 1.7, in terms of moduli spaces of rank 2 parabolic vector bundles on $\mathbb{P}^1$.

To prove Theorem 1.5, we determine the nef cone of $G^{(n)}$ explicitly, and then compare it with the Mori chamber decomposition of the effective cone of $X^{(n)}$ described by Mukai [23]. This decomposition encodes the nef cones of all varieties pseudoisomorphic to $X^{(n)}$. In order to determine the cone of effective curves and the nef cone of $G^{(n)}$, we generalize to arbitrary dimension a construction of Borcea [6] in dimension $n = 4$. We define isomorphisms

$$H^{2n-2}(G^{(n)}, \mathbb{Z}) \xrightarrow{\alpha} H^n(Z^{(n)}, \mathbb{Z}) \xrightarrow{\beta} H^2(G^{(n)}, \mathbb{Z})$$

such that $\beta(M) = E_M$ and $\alpha^{-1}(M)$ is the class of a line on the dual m–plane $M^* \subset G^{(n)}$ for every $M \in \mathcal{F}_m(Z^{(n)})$. These isomorphisms are dual with respect to the intersection products, ie $x \cdot \beta(y) = \alpha(x) \cdot y$ for every $x \in H^{2n-2}(G^{(n)}, \mathbb{Z})$ and $y \in H^n(Z^{(n)}, \mathbb{Z})$. They allow us to describe explicitly special cones of curves and divisors on $G^{(n)}$:

1.8 Theorem (Theorem 5.1 and Proposition 5.5) *Let $\mathcal{E} \subset H^n(Z,\mathbb{R})$ be the polyhedral cone generated by the classes $\{M\}_{M \in \mathcal{F}_m(Z)}$, and denote by $\mathcal{E}^\vee \subset H^n(Z,\mathbb{R})$ its dual cone. Then $\mathcal{E}^\vee \subset \mathcal{E}$, and the cones of nef and effective divisors of $G^{(n)}$ and their dual cones of effective and moving curves satisfy*

$$\mathrm{Nef}(G^{(n)}) = \beta(\mathcal{E}^\vee) \subset \beta(\mathcal{E}) = \mathrm{Eff}(G^{(n)}),$$
$$\mathrm{Mov}_1(G^{(n)}) = \alpha^{-1}(\mathcal{E}^\vee) \subset \alpha^{-1}(\mathcal{E}) = \mathrm{NE}(G^{(n)}).$$

We give a geometric description of the extremal rays and facets of these cones, and the associated contractions in Section 6. In Proposition 6.6 and its following subsection, we also describe the cone $\mathrm{Mov}^1(G^{(n)})$ of movable divisors of $G^{(n)}$, and give a geometric description of the curves corresponding to its facets.

We end this paper by determining the automorphism group of the Fano variety $G^{(n)}$, generalizing the description of the automorphism group of a quartic del Pezzo surface in Section 1.1. In what follows, we write $W(D_{n+3})$ for the Weyl group of automorphisms of a D_{n+3}–lattice, and we denote by the same symbol the involution of $G^{(n)}$ induced by the involution σ_i of $Z^{(n)}$.

1.9 Proposition (Proposition 7.1) *There is an inclusion of groups*

$$\mathrm{Aut}(G^{(n)}) \hookrightarrow W(D_{n+3}) \cong (\mathbb{Z}/2\mathbb{Z})^{n+2} \rtimes S_{n+3},$$

whose image contains the normal subgroup $(\mathbb{Z}/2\mathbb{Z})^{n+2}$ generated by the involutions σ_i of $G^{(n)}$.

Moreover, if the points $(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1) \in \mathbb{P}^1$ are general, then $\mathrm{Aut}(G^{(n)}) \cong (\mathbb{Z}/2\mathbb{Z})^{n+2}$.

This paper is organized as follows. Section 2 is dedicated to smooth complete intersections $Z \subset \mathbb{P}^{n+2}$ for $n = 2m$ of two quadric hypersurfaces in even-dimensional projective spaces. In particular, we investigate the set $\mathcal{F}_m(Z)$ of m–planes in Z, and the cone it spans in $H^n(Z,\mathbb{R})$. In Section 3, we address the Fano variety G of $(m-1)$–planes in Z. We construct the isomorphisms $H^{2n-2}(G,\mathbb{Z}) \xrightarrow{\alpha} H^n(Z,\mathbb{Z}) \xrightarrow{\beta} H^2(G,\mathbb{Z})$, and determine some extremal rays of the cone of effective curves of G. In Section 4, we consider the blow-up X of $\mathbb{P}^n$ at $n+3$ points in general linear position. We describe the Mori chamber decomposition of $\mathrm{Eff}(X)$, following Mukai [23] and Bauer [3]. From this we can write the nef cone of G in terms of a natural basis for $\mathcal{N}^1(X)$. In Section 5, we put together the results from the previous sections to prove Theorem 1.5. In Section 6, we study cones of curves and divisors in G, giving a geometric description of their facets and extremal rays. In Section 7, we describe the automorphism group of the Fano variety G.

Notation and conventions We always work over the field $\mathbb{C}$ of complex numbers.

Given a subvariety $Z \subset \mathbb{P}^n$ and a nonnegative integer $d < n$, we denote by $\mathcal{F}_d(Z)$ the closed subset of the Grassmannian $\mathrm{Gr}(d, \mathbb{P}^n)$ parametrizing d–planes contained in Z.

Acknowledgements We thank Ana-Maria Castravet, Alex Massarenti, Elisa Postinghel and the referee for useful comments and discussions.

Araujo was partially supported by CNPq and Faperj Research Fellowships and an ICTP Simons Associateship. This work started during Araujo's visit to Università di Torino; the authors are grateful to INdAM (Istituto Nazionale di Alta Matematica) for the support for this visit.

2 Smooth complete intersections of two quadrics

In this section we describe the geometry of smooth complete intersections of two quadric hypersurfaces in even dimensional complex projective spaces. Many of the results are well known and can be found in Reid [25, Chapter 3] or Borcea [6, Section 1], to which we refer for details and proofs. See also the recent paper by Dolgachev and Duncan [13] for a study of these complete intersections over a field of characteristic 2.

Let $n = 2m \geq 2$ be an even integer, and let $Z = Q_1 \cap Q_2 \subset \mathbb{P}^{n+2}$ be a smooth complete intersection of two quadric hypersurfaces. Up to a projective transformation of $\mathbb{P}^{n+2}$, we can assume that the quadrics have equations

$$(2.1) \qquad Q_1 : \sum_{i=1}^{n+3} x_i^2 = 0, \quad Q_2 : \sum_{i=1}^{n+3} \lambda_i x_i^2 = 0,$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$. Thus Z is determined by $n+3$ distinct points

$$(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1) \in \mathbb{P}^1.$$

Acting on these points by permutations and projective automorphisms of $\mathbb{P}^1$ yields projectively isomorphic varieties $Z \subset \mathbb{P}^{n+2}$.

2.2 (involutions and double covers) For each $i = 1, \dots, n+3$, let $\sigma_i \colon Z \to Z$ be the involution switching the sign of the coordinate x_i. Then $\sigma_1, \dots, \sigma_{n+3}$ commute and have the unique relation $\sigma_1 \cdots \sigma_{n+3} = \mathrm{Id}_Z$, so they generate a subgroup W' of $\mathrm{Aut}(Z)$ isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{n+2}$. For every subset $I \subseteq \{1, \dots, n+3\}$, we set $\sigma_I := \prod_{i \in I} \sigma_i$. Notice that $\sigma_I = \sigma_{I^c}$.

For each $i = 1, \dots, n+3$, the projection from the i^{th} coordinate point in $\mathbb{P}^{n+2}$ yields a double cover $\pi_i \colon Z \to Q^n$, where $Q^n \subset \mathbb{P}^{n+1}$ is the smooth quadric having equation $\sum_{j \neq i} (\lambda_j - \lambda_i) x_j^2 = 0$ for projective coordinates $(x_1 : \dots : \hat{x}_i : \dots : x_{n+3})$ in $\mathbb{P}^{n+1}$. The involution associated to this double cover is σ_i.

2.3 (the set of m–planes in Z) Consider the set $\mathcal{F}_m(Z)$ of m–planes in Z. It is a finite set with cardinality 2^{n+2}. The group W' generated by the involutions σ_i acts on $\mathcal{F}_m(Z)$ freely and transitively.

For every $M \in \mathcal{F}_m(Z)$ and $I \subset \{1, \dots, n+3\}$ with $|I| \le m+1$, we have

$$\dim(M \cap \sigma_I(M)) = m - |I|. \tag{2.4}$$

2.5 For each $i = 1, \dots, n+3$, the double cover $\pi_i \colon Z \to Q^n$ induces a map

$$\mathcal{F}_m(Z) \to \mathcal{F}_m(Q^n).$$

Recall that $\mathcal{F}_m(Q^n)$ has two connected components T^φ and T^ψ, and that two m–planes $\Lambda, \Lambda' \subset Q^n$ belong to the same connected component if and only if $\dim(\Lambda \cap \Lambda') \equiv m \bmod 2$ (see for instance Reid [25, Theorem 1.2(b)] or Harris [17, Theorem 22.14]).

Let $M \in \mathcal{F}_m(Z)$. We have $\pi_i(\sigma_i(M)) = \pi_i(M)$. On the other hand, if j is in $\{1, \dots, n+3\} \smallsetminus \{i\}$, then M and $\sigma_j(M)$ intersect in codimension one, by (2.4), and the same holds for $\pi_i(M)$ and $\pi_i(\sigma_j(M))$. Therefore $\pi_i(M)$ and $\pi_i(\sigma_j(M))$ belong to different connected components of $\mathcal{F}_m(Q^n)$. In general, if $I \subseteq \{1, \dots, n+3\}$ does not contain i, then $\pi_i(M)$ and $\pi_i(\sigma_I(M))$ belong to the same connected component of $\mathcal{F}_m(Q^n)$ if and only if $|I|$ is even. This shows that the image of $\mathcal{F}_m(Z)$ in $\mathcal{F}_m(Q^n)$ consists of 2^{n+1} points, half in each connected component.

2.6 (the cohomology group $H^n(Z, \mathbb{Z})$) The cohomology group $H^n(Z, \mathbb{Z})$ is isomorphic to $\mathbb{Z}^{n+4}$, and is generated over $\mathbb{Z}$ by the classes of the m–planes in Z. Moreover $H^n(Z, \mathbb{Z})$ is a unimodular lattice with respect to the intersection form.

For every $M \in \mathcal{F}_m(Z)$ we denote by the same symbol M the corresponding fundamental class in $H^n(Z, \mathbb{Z})$. We denote by $\eta \in H^n(Z, \mathbb{Z})$ the class of a codimension-m linear section of $Z \subset \mathbb{P}^{n+2}$, so that

$$\eta^2 = 4 \qquad \text{and} \qquad \eta \cdot M = 1 \quad \text{for every } M \in \mathcal{F}_m(Z).$$

The sublattice $\eta^\perp$ (namely the primitive part $H^n(Z, \mathbb{Z})_0$) is a D_{n+3}–lattice. We denote by $W(D_{n+3})$ its Weyl group of automorphisms, which is generated by the reflections in the roots of $\eta^\perp$. It is the full group of automorphisms of the triple $(H^n(Z, \mathbb{Z}), \cdot, \eta)$, and it is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{n+2} \rtimes S_{n+3}$.

The group $W' \cong (\mathbb{Z}/2\mathbb{Z})^{n+2}$ generated by the involutions σ_i acts naturally and faithfully on $H^n(Z,\mathbb{Z})$. We still denote by σ_I the involution of $H^n(Z,\mathbb{Z})$ induced by $\sigma_I\colon Z \to Z$. So we view W' as a subgroup of $W(D_{n+3})$. It is a normal subgroup with quotient $W(D_{n+3})/W'$ isomorphic to the symmetric group S_{n+3}.

For every $M \in \mathcal{F}_m(Z)$ and $i, j \in \{1,\dots,n+3\}$ with $i \neq j$, we have

$$\eta = M + \sigma_i(M) + \sigma_j(M) + \sigma_{ij}(M). \tag{2.7}$$

2.8 Notation Fix $M_0 \in \mathcal{F}_m(Z)$. For every $i = 1,\dots,n+3$, we set $M_i := \sigma_i(M_0)$. More generally, for every subset $I \subseteq \{1,\dots,n+3\}$, we set $M_I := \sigma_I(M_0)$. Notice again that $M_I = M_{I^c}$. We also set

$$\varepsilon_i := M_0 + M_i - \tfrac{1}{2}\eta \in H^n(Z,\mathbb{R}) \quad \text{for every } i = 1,\dots,n+3. \tag{2.9}$$

Then $\{\eta, \varepsilon_1, \dots, \varepsilon_{n+3}\}$ is an orthogonal basis for $H^n(Z,\mathbb{R})$, which is useful for computations. We have

$$\eta^2 = 4 \qquad \text{and} \qquad \varepsilon_i^2 = (-1)^m \quad \text{for every } i = 1,\dots,n+3. \tag{2.10}$$

In particular, the intersection form on $H^n(Z,\mathbb{R})$ is positive definite when $n \equiv 0 \bmod 4$, and has signature $(1, n+3)$ when $n \equiv 2 \bmod 4$. Notice that this basis depends on the choice of M_0.

Let $G_0 \subset W(D_{n+3})$ be the stabilizer of M_0. Then $G_0 \cong S_{n+3}$ and G_0 acts by (the same) permutations both on $\{M_1,\dots,M_{n+3}\}$ and on $\{\varepsilon_1,\dots,\varepsilon_{n+3}\}$. We have $W(D_{n+3}) = W' \rtimes G_0$. Moreover, for every $I \subseteq \{1,\dots,n+3\}$ of even cardinality, we have

$$\sigma_I(\varepsilon_i) = \begin{cases} \varepsilon_i & \text{if } i \notin I, \\ -\varepsilon_i & \text{if } i \in I. \end{cases} \tag{2.11}$$

Thus we see the usual action of $W(D_{n+3})$ on the linear span of $\varepsilon_1,\dots,\varepsilon_{n+3}$ by permutation and even sign changes of $\varepsilon_1,\dots,\varepsilon_{n+3}$ (see for instance Humphreys [19, Section 12.1]).

We collect some identities in $H^n(Z,\mathbb{R})$ that we will use in later computations.

$$M_I = \frac{1}{4}\eta + \frac{(-1)^{|I|}}{2}\left(\sum_{j \notin I} \varepsilon_j - \sum_{i \in I} \varepsilon_i\right) \quad \text{for every } I \subseteq \{1,\dots,n+3\}, \tag{2.12}$$

$$M_I = \frac{1}{n+1}\left((n+2-|I|)\left(\frac{1}{2}\eta - \sum_{i \in I} M_i\right) + (|I|-1)\sum_{j \in I^c} M_j\right) \tag{2.13}$$

for every $I \subseteq \{1,\dots,n+3\}$ with even cardinality,

$$\varepsilon_i = \frac{1}{2(n+1)}\eta - \frac{1}{n+1}\sum_{j=1}^{n+3} M_j + M_i \quad \text{for every } i = 1, \dots, n+3. \tag{2.14}$$

Our next goal is to describe the polyhedral cone $\mathcal{E}$ in $H^n(Z,\mathbb{R})$ generated by the classes of m–planes in Z. As we shall see below, this is a cone over a $(n+3)$–dimensional *demihypercube*. Before we start discussing the cone $\mathcal{E}$, we gather some results about demihypercubes.

2.15 (the demihypercube) Let $N \geq 4$ be an integer. Write $(\alpha_1, \dots, \alpha_N)$ for coordinates in $\mathbb{R}^N$. The vertices of the hypercube $\left[-\frac{1}{2}, \frac{1}{2}\right]^N \subset \mathbb{R}^N$ are the points of the form $v_I = ((v_I)_1, \dots, (v_I)_N)$, where $I \subseteq \{1, \dots, N\}$, $(v_I)_i = \frac{1}{2}$ if $i \in I$, and $(v_I)_i = -\frac{1}{2}$ otherwise. The parity of the vertex v_I is the parity of $|I|$. For each subset $I \subseteq \{1, \dots, N\}$, define the degree 1 polynomial in the α_i

$$H_I := \sum_{j \notin I}\left(\tfrac{1}{2} + \alpha_j\right) + \sum_{i \in I}\left(\tfrac{1}{2} - \alpha_i\right). \tag{2.16}$$

Notice that, for any two subsets $I, J \subset \{1, \dots, N\}$,

$$H_I(v_J) = \#(I \smallsetminus J) + \#(J \smallsetminus I) \tag{2.17}$$

is the graph distance of v_I and v_J in the skeleton of the hypercube $\left[-\frac{1}{2}, \frac{1}{2}\right]^N$.

The *demihypercube* is the polytope $\Delta \subset \left[-\frac{1}{2}, \frac{1}{2}\right]^N$ generated by the odd vertices of the hypercube. The polytope Δ has $2^{N-1} + 2N$ facets (see for instance Green [15, Lemma 2.3]). More precisely, the polytope Δ is defined in a minimal way by the set of inequalities

$$\Delta = \begin{cases} -\frac{1}{2} \leq \alpha_i \leq \frac{1}{2}, & i \in \{1, \dots, N\}, \\ H_I \geq 1, & |I| \text{ even.} \end{cases} \tag{2.18}$$

Notice that the facets of Δ supported on the hyperplanes $\left(\alpha_i = \pm\frac{1}{2}\right)$ are isomorphic to the $(N-1)$–dimensional demihypercube. In particular, they are not simplicial. On the other hand, the facet supported on the hyperplane $(H_I = 1)$ for $|I|$ even is the $(N-1)$–dimensional simplex generated by the N vertices of $\left[-\frac{1}{2}, \frac{1}{2}\right]^N$ at graph distance 1 to v_I.

The demihypercube can also be described as a weight polytope of the root system of type D_N; see Green [16, Example 8.5.13].

Now we go back to $H^n(Z,\mathbb{R})$ and consider the convex rational polyhedral cone

$$\mathcal{E} := \operatorname{Cone}(M)_{M \in \mathcal{F}_m(Z)} \subset H^n(Z,\mathbb{R}).$$

It is the cone over the $(n+3)$–dimensional polytope

$$\mathcal{E}_0 = \mathrm{Conv}(M)_{M \in \mathcal{F}_m(Z)}$$

obtained by intersecting $\mathcal{E}$ with the affine hyperplane $\mathcal{H} := \{\gamma \mid \gamma \cdot \eta = 1\}$. Note that the Weyl group $W(D_{n+3})$ preserves $\mathcal{E}$, $\mathcal{H}$ and $\mathcal{E}_0$.

We fix $M_0 \in \mathcal{F}_m(Z)$ and consider the orthogonal basis $\{\eta, \varepsilon_1, \dots, \varepsilon_{n+3}\}$ for $H^n(Z, \mathbb{R})$ introduced in (2.9). Then $\frac{1}{4}\eta \in \mathcal{H}$ and $\{\varepsilon_1, \dots, \varepsilon_{n+3}\}$ is a basis for $\eta^{\perp}$, so that $\left(\frac{1}{4}\eta, \{\varepsilon_1, \dots, \varepsilon_{n+3}\}\right)$ induces affine coordinates

$$(\alpha_1, \dots, \alpha_{n+3}) \tag{2.19}$$

on the hyperplane $\mathcal{H} \cong \mathbb{R}^{n+3}$. With these coordinates, $\frac{1}{4}\eta$ is identified with the origin and, by (2.12), for every $I \subset \{1, \dots, n+3\}$ with $|I|$ even, M_I is identified with v_{I^c}. Thus the polytope $\mathcal{E}_0$ is identified with the demihypercube Δ described in Section 2.15, and $\mathcal{E}$ with the cone over Δ.

2.20 Example (the surface case) When $n = 2$, $Z \subset \mathbb{P}^4$ is a smooth quartic del Pezzo surface (see Section 1.1). The cone $\mathcal{E} \subset H^2(Z, \mathbb{R})$, generated by the classes of the 16 lines in Z, is the cone of effective curves of Z. In this case the polytope $\mathcal{E}_0$ is a 5–dimensional demihypercube, and coincides with the 5–dimensional Gosset polytope (see Dolgachev [12, Sections 8.2.5 and 8.2.6]). In higher dimensions, demihypercubes and Gosset polytopes are different polytopes.

Let us explicitly describe the facets of $\mathcal{E}$, or equivalently the generators of the dual cone $\mathcal{E}^{\vee} \subset H^n(Z, \mathbb{R})$. Let $(y, x_1, \dots, x_{n+3})$ be the coordinates on $H^n(Z, \mathbb{R}) \cong \mathbb{R}^{n+4}$ induced by the basis $\{\eta, \varepsilon_1, \dots, \varepsilon_{n+3}\}$. It follows from (2.18) that the cone $\mathcal{E}$ is defined in a minimal way by the set of inequalities

$$\mathcal{E} = \begin{cases} 2y + x_i \geq 0, & i \in \{1, \dots, n+3\}, \\ 2y - x_i \geq 0, & i \in \{1, \dots, n+3\}, \\ 2(n+1)y + \sum\limits_{j \notin I} x_j - \sum\limits_{i \in I} x_i \geq 0, & I \subset \{1, \dots, n+3\} \text{ even}. \end{cases} \tag{2.21}$$

This is equivalent to saying that the dual cone $\mathcal{E}^{\vee} \subset H^n(Z, \mathbb{R})$ is the convex polyhedral cone generated by the classes

$$\begin{cases} \frac{1}{2}\eta + \varepsilon_i \text{ and } \frac{1}{2}\eta - \varepsilon_i, & i \in \{1, \dots, n+3\}, \\ \frac{1}{2}(n+1)\eta + (-1)^m \sum\limits_{j \notin I} \varepsilon_j - (-1)^m \sum\limits_{i \in I} \varepsilon_i, & I \subset \{1, \dots, n+3\}, \ |I| \text{ even}. \end{cases} \tag{2.22}$$

2.23 Remark Using (2.7), (2.9) and (2.12), we can write the generators (2.22) of $\mathcal{E}^{\vee}$ in terms of η and the M_I:

$$\begin{cases} \frac{1}{2}\eta + \varepsilon_i = M_0 + M_i, \\ \frac{1}{2}\eta - \varepsilon_i = M_j + M_{ij} \quad \text{for any } j \neq i, \\ \frac{1}{2}(n+1)\eta + (-1)^m \sum_{j \notin I} \varepsilon_j - (-1)^m \sum_{i \in I} \varepsilon_i = 2\big(\lfloor \frac{1}{2}(m+1) \rfloor \eta + (-1)^m M_I\big). \end{cases}$$

Note in particular that $\mathcal{E}^\vee \subset \mathcal{E}$.

For $I \subseteq \{1, \dots, n+3\} \smallsetminus \{i\}$, it follows from (2.10) and (2.12) that

$$\begin{aligned} \big(\tfrac{1}{2}\eta + \varepsilon_i\big) \cdot M_I &= \begin{cases} 1 & \text{if } |I| \equiv m \bmod 2, \\ 0 & \text{otherwise}, \end{cases} \\ \big(\tfrac{1}{2}\eta - \varepsilon_i\big) \cdot M_I &= \begin{cases} 0 & \text{if } |I| \equiv m \bmod 2, \\ 1 & \text{otherwise}. \end{cases} \end{aligned} \tag{2.24}$$

This describes the generators of the (nonsimplicial) facets of $\mathcal{E}$, corresponding to the extremal rays of $\mathcal{E}^\vee$ generated by $\frac{1}{2}\eta \pm \varepsilon_i$.

For each $M \in \mathcal{F}_m(Z)$, set

$$\delta_M := \lfloor \tfrac{1}{2}(m+1) \rfloor \eta + (-1)^m M.$$

The facet of the cone $\mathcal{E}$ corresponding to the extremal ray of $\mathcal{E}^\vee$ generated by δ_M is simplicial, and given by

$$\operatorname{Cone}(\sigma_i(M))_{i \in \{1,\dots,n+3\}}.$$

Indeed, for $I \subseteq \{1, \dots, n+3\}$ with $|I|$ odd, one computes, using (2.12),

$$\delta_M \cdot \sigma_I(M) = \tfrac{1}{2}(|I| - 1).$$

Let $(z, t_1, \dots, t_{n+3})$ be the coordinates induced by the basis $\{\eta, M_1, \dots, M_{n+3}\}$ on $H^n(Z, \mathbb{R})$. In the sequel we need equations for $\mathcal{E}^\vee$ in these coordinates. Let $I \subseteq \{1, \dots, n+3\}$ be such that $|I| \equiv m \bmod 2$. Using (2.12), one computes

$$\left(z\eta + \sum_{i=1}^{n+3} t_i M_i \right) \cdot M_I = 2z + (|I| - m) \sum_{i=1}^{n+3} t_i - 2 \sum_{i \in I} t_i.$$

So we get the following:

2.25 Lemma *An element $z\eta + \sum_{i=1}^{n+3} t_i M_i$ is in $\mathcal{E}^\vee$ if and only if*

$$2z + (|I| - m) \sum_{i=1}^{n+3} t_i - 2 \sum_{i \in I} t_i \geq 0 \tag{2.26}$$

for every $I \subseteq \{1, \dots, n+3\}$ such that $|I| \equiv m \bmod 2$.

We conclude this section with the following elementary description of the symmetry group of the cone $\mathcal{E}$:

2.27 Lemma *Let $f\colon H^n(Z,\mathbb{R}) \to H^n(Z,\mathbb{R})$ be a linear map. The following are equivalent:*

(i) *$f(\mathcal{E}) = \mathcal{E}$ and $f(x)\cdot\eta = x\cdot\eta$ for every $x \in H^n(Z,\mathbb{R})$.*

(ii) *$f(\mathcal{E}^\vee) = \mathcal{E}^\vee$ and $f(\eta) = \eta$.*

(iii) *$f \in W(D_{n+3})$.*

Proof The implications (iii) $\Longrightarrow$ (i) and (iii) $\Longrightarrow$ (ii) are clear.

We prove (i) $\Longrightarrow$ (iii). Let f be an endomorphism of $H^n(Z,\mathbb{R})$ satisfying (i). Then f permutes the vertices of $\mathcal{E}_0$, and hence $f(\mathcal{F}_m(Z)) = \mathcal{F}_m(Z)$.

Recall Notation 2.8; let $M_0 \in \mathcal{F}_m(Z)$. By Remark 2.23, $\delta_{M_0} = \lfloor \frac{1}{2}(m+1) \rfloor \eta + (-1)^m M_0$ generates an extremal ray of $\mathcal{E}^\vee$, and the corresponding facet of $\mathcal{E}$ is simplicial, given by

$$\mathrm{Cone}(M_1,\dots,M_{n+3}).$$

Then $f(\mathrm{Cone}(M_1,\dots,M_{n+3}))$ must be another simplicial facet of $\mathcal{E}$, of the form

$$\mathrm{Cone}(\sigma_1(M_I),\dots,\sigma_{n+3}(M_I)) = \sigma_I(\mathrm{Cone}(M_1,\dots,M_{n+3}))$$

for some $I \subseteq \{1,\dots,n+3\}$. By composing f with the involution $\sigma_I \in W(D_{n+3})$, we may assume that f fixes the facet $\mathrm{Cone}(M_1,\dots,M_{n+3})$ of $\mathcal{E}$. In particular, f induces a permutation on the set $\{M_1,\dots,M_{n+3}\}$. Let $\omega \in W(D_{n+3})$ be the element in the stabilizer of M_0 inducing the same permutation as f on the set $\{M_1,\dots,M_{n+3}\}$. Then, by composing f with ω^{-1}, we may assume that f fixes each of $M_1,\dots,M_{n+3}$.

We also have $f(\mathcal{F}_m(Z) \smallsetminus \{M_1,\dots,M_{n+3}\}) = \mathcal{F}_m(Z) \smallsetminus \{M_1,\dots,M_{n+3}\}$, therefore f must fix the point

$$v := \sum_{M \in \mathcal{F}_m(Z) \smallsetminus \{M_1,\dots,M_{n+3}\}} M.$$

Since $\delta_{M_0}\cdot v > 0$, v is not contained in the linear span of $M_1,\dots,M_{n+3}$ (see Remark 2.23). This implies that $f = \mathrm{Id}_{H^n(Z,\mathbb{R})} \in W(D_{n+3})$.

Finally we prove (ii) $\Longrightarrow$ (iii). Let f be an endomorphism of $H^n(Z,\mathbb{R})$ satisfying (ii). Then the dual map $g := f^t\colon H^n(Z,\mathbb{R}) \to H^n(Z,\mathbb{R})$ satisfies (i), hence, by what precedes, $g \in W(D_{n+3})$. In particular g is orthogonal, and $f = g^t = g^{-1} \in W(D_{n+3})$. $\square$

3 The Fano variety G of $(m-1)$–planes in $Q_1 \cap Q_2 \subset \mathbb{P}^{2m+2}$

Let $n = 2m \geq 2$ be an even integer, and let $Z = Q_1 \cap Q_2 \subset \mathbb{P}^{n+2}$ be a smooth complete intersection of two quadric hypersurfaces as in (2.1). In this section we consider the variety G of $(m-1)$–planes in Z:

$$G := \mathcal{F}_{m-1}(Z) = \{[L] \in \mathrm{Gr}(m-1, \mathbb{P}^{n+2}) \mid L \subset Z\}.$$

This is a smooth n–dimensional Fano variety that has been much studied. In particular, it is known that $\mathrm{Pic}(G) \cong H^2(G,\mathbb{Z}) \cong \mathbb{Z}^{n+4}$, $\mathcal{N}^1(G) \cong H^2(G,\mathbb{R})$ and $-K_G$ is the restriction of $\mathcal{O}(1)$ on $\mathrm{Gr}(m-1, \mathbb{P}^{n+2})$ (see Reid [25, Theorem 2.6], Borcea [5, Theorem 4.1 and Remark 4.3] and Jiang [20, Proposition 3.2]). Moreover G is rational, hence $H^{2n-2}(G,\mathbb{Z})$ is torsion-free (Artin and Mumford [2, Proposition 1]) and generated by fundamental classes of one-cycles (Soulé and Voisin [26, Lemma 1]). Thus we also have $H^{2n-2}(G,\mathbb{Z}) \cong \mathbb{Z}^{n+4}$ and $\mathcal{N}_1(G) \cong H^{2n-2}(G,\mathbb{R})$.

For each $M \in \mathcal{F}_m(Z)$ we set

$$M^* := \{[L] \in G \mid L \subset M\}. \tag{3.1}$$

It is an m–plane in G (under the Plücker embedding). Let $\ell_M \in H^{2n-2}(G,\mathbb{Z})$ be the class of a line in M^*. By (2.4), for every M, $M' \in \mathcal{F}_m(Z)$ we have

$$M^* \cap (M')^* \neq \varnothing \qquad \Longleftrightarrow \qquad M' = \sigma_i(M) \quad \text{for some } i = 1, \dots, n+3,$$

and $M^* \cap \sigma_i(M)^*$ is the point $[M \cap \sigma_i(M)] \in G$.

3.2 (the fibrations φ_i and ψ_i on G) We define $2(n+3)$ fibrations on G, generalizing a construction by Borcea in the case $n = 4$ [6, Section 3]. For each $i = 1, \dots, n+3$, the double cover $\pi_i \colon Z \to Q^n$ introduced in Section 2.2 induces a map

$$\Pi_i \colon G \to \mathcal{F}_{m-1}(Q^n).$$

Each $(m-1)$–plane in Q^n is contained in exactly one m–plane of each of the two families T^φ and T^ψ of m–planes in Q^n (see for instance Harris [17, Theorem 22.14]). This yields two morphisms

$$\mathcal{F}_{m-1}(Q^n) \to T^\varphi \subset \mathrm{Gr}(m, \mathbb{P}^{n+1}) \quad \text{and} \quad \mathcal{F}_{m-1}(Q^n) \to T^\psi \subset \mathrm{Gr}(m, \mathbb{P}^{n+1}).$$

By composing them with $\Pi_i \colon G \to \mathcal{F}_{m-1}(Q^n)$, we get two distinct morphisms

$$\overline{\varphi}_i,\ \overline{\psi}_i \colon G \to \mathrm{Gr}(m, \mathbb{P}^{n+1})$$

such that $\overline{\varphi}_i(G) \subseteq T^\varphi$ and $\overline{\psi}_i(G) \subseteq T^\psi$. Let

$$G \xrightarrow{\varphi_i} Y_{\varphi_i} \to \overline{\varphi}_i(G) \qquad \text{and} \qquad G \xrightarrow{\psi_i} Y_{\psi_i} \to \overline{\psi}_i(G)$$

be the Stein factorizations of $\overline{\varphi}_i$ and $\overline{\psi}_i$, respectively.

3.3 Lemma *The morphism $\varphi_i\colon G \to Y_{\varphi_i}$ has general fiber $\mathbb{P}^1$, and has exactly 2^n singular fibers, each isomorphic to a union of two copies of $\mathbb{P}^m$ meeting transversally at one point. More precisely, the singular fibers of φ_i are of the form $M^* \cup \sigma_i(M)^*$, with $M \in \mathcal{F}_m(Z)$ such that $[\pi_i(M)] \in T^\varphi$. An analogous statement holds for ψ_i.*

As a consequence, the cone $\mathrm{NE}(\varphi_i)$ is the convex cone generated by the classes ℓ_M for $M \in \mathcal{F}_m(Z)$ such that $[\pi_i(M)] \in T^\varphi$, and similarly for $\mathrm{NE}(\psi_i)$.

Proof For simplicity we assume in the proof that $m \geq 2$ and $n \geq 4$, the case $n = 2$ being classical.

Let $[\Lambda] \in T^\varphi \subset \mathrm{Gr}(m, \mathbb{P}^{n+1})$, and let $\Lambda' \subset \mathbb{P}^{n+2}$ be the $(m+1)$–plane through the i^{th} coordinate point that projects onto $\Lambda \subset \mathbb{P}^{n+1}$. Then Λ' is contained in a singular quadric of the pencil of quadrics through Z, so that $\Lambda' \cap Z = \Lambda' \cap Q_1$ is an m–dimensional quadric in Λ'. Hence $[\Lambda] \in \overline{\varphi}_i(G)$ if and only if $\Lambda' \cap Z$ contains an $(m-1)$–plane. This happens if and only if the quadric $\Lambda' \cap Z$ has rank at most 4.

If the m–dimensional quadric $\Lambda' \cap Z$ has rank 4, then it is the join of an $(m-3)$–plane with a smooth quadric surface $\cong \mathbb{P}^1 \times \mathbb{P}^1$. So it contains two distinct 1–dimensional families of $(m-1)$–planes, each parametrized by $\mathbb{P}^1$. Therefore $\overline{\varphi}_i^{-1}([\Lambda])$ is the disjoint union of two copies of $\mathbb{P}^1$, and this yields two smooth fibers of φ_i, each isomorphic to $\mathbb{P}^1$.

If $\Lambda' \cap Z$ has rank 3, then it is the join of an $(m-2)$–plane with a plane conic. So it contains a one-dimensional family of $(m-1)$–planes, parametrized by the conic. Thus in this case $\overline{\varphi}_i^{-1}([\Lambda])_{\mathrm{red}} \cong \mathbb{P}^1$, and this yields a fiber of φ_i with reduced structure isomorphic to $\mathbb{P}^1$.

If $\Lambda' \cap Z$ has rank 2, then it is the union of two m–planes intersecting in codimension one, both projecting onto Λ. Thus there exists $M \in \mathcal{F}_m(Z)$ such that $\Lambda = \pi_i(M)$, $\Lambda' \cap Z = M \cup \sigma_i(M)$ and $\overline{\varphi}_i^{-1}([\Lambda]) = M^* \cup \sigma_i(M)^*$. It follows from (2.4) that M^* and $\sigma_i(M)^*$ intersect in one point.

Finally, if $\Lambda' \cap Z$ has rank 1, then set-theoretically we should have $\Lambda' \cap Z = M$ for some $M \in \mathcal{F}_m(Z)$, and hence $\overline{\varphi}_i^{-1}([\pi_i(M)]) = M^*$, which is impossible because we have already seen that $\overline{\varphi}_i^{-1}([\pi_i(M)]) = M^* \cup \sigma_i(M)^*$.

Now set

$$U := Y_{\varphi_i} \smallsetminus \{\varphi_i(M^* \cup \sigma_i(M)^*) \mid M \in \mathcal{F}_m(Z) \text{ and } [\pi_i(M)] \in T^\varphi\}.$$

We have shown that φ_i has one-dimensional fibers over U, and since G is Fano, φ_i is a conic bundle over U. A general singular fiber should be reduced with two irreducible components. Since there are no such fibers, φ_i is smooth over U. □

In Section 6.5 we will characterize the varieties Y_{φ_i} and Y_{ψ_i}.

Fix $M_0 \in \mathcal{F}_m(Z)$ such that $[\pi_i(M_0)] \in T^{\psi}$, and follow Notation 2.8. It follows from Section 2.5 that, for every $I \subseteq \{1, \dots, n+3\}$ such that $i \notin I$,

$$[\pi_i(M_I)] \in \begin{cases} T^{\varphi} & \text{if } |I| \text{ is odd,} \\ T^{\psi} & \text{if } |I| \text{ is even.} \end{cases}$$

So we get the following corollary of Lemma 3.3:

3.4 Corollary *We have*

$$\mathrm{NE}(\varphi_i) = \mathrm{Cone}(\ell_{M_I})_{|I| \text{odd}, i \notin I} \quad \textit{and} \quad \mathrm{NE}(\psi_i) = \mathrm{Cone}(\ell_{M_I})_{|I| \text{even}, i \notin I}.$$

The general fiber of φ_i has class $\ell_{M_j} + \ell_{M_{ij}}$ for $j \neq i$, and the general fiber of ψ_i has class $\ell_{M_0} + \ell_{M_i}$.

3.5 (the isomorphisms between $H^{2n-2}(G, \mathbb{Z})$, $H^n(Z, \mathbb{Z})$ and $H^2(G, \mathbb{Z})$) Recall that, by Poincaré duality, the intersection product gives a perfect pairing

$$H^2(G, \mathbb{Z}) \times H^{2n-2}(G, \mathbb{Z}) \to \mathbb{Z}.$$

We will define natural isomorphisms $H^{2n-2}(G, \mathbb{Z}) \cong H^n(Z, \mathbb{Z})$ and $H^2(G, \mathbb{Z}) \cong H^n(Z, \mathbb{Z})$, which behave well with respect to the intersection products. This construction is due to Borcea in the case $n = 4$ [6, Section 2]. Throughout this section, we use the same notation as in Section 2.

Consider the incidence variety

$$\mathcal{I} := \{([L], p) \in G \times Z \mid p \in L\}$$

and the associated diagram:

$$\begin{tikzcd} & \mathcal{I} \arrow[dl, "\pi"'] \arrow[dr, "e"] & \\ G & & Z \end{tikzcd}$$

The morphism π is a $\mathbb{P}^{m-1}$–bundle, hence $\mathcal{I}$ is smooth, irreducible, of dimension $3m - 1 = \frac{3}{2}n - 1$. Consider the morphisms given by pull-backs and Gysin homomorphisms

$$\alpha := e_* \circ \pi^* \colon H^{2n-2}(G, \mathbb{Z}) \xrightarrow{\pi^*} H^{2n-2}(\mathcal{I}, \mathbb{Z}) \xrightarrow{e_*} H^n(Z, \mathbb{Z}),$$

$$\beta := \pi_* \circ e^* \colon H^n(Z, \mathbb{Z}) \xrightarrow{e^*} H^n(\mathcal{I}, \mathbb{Z}) \xrightarrow{\pi_*} H^2(G, \mathbb{Z}),$$

so that we have

$$H^{2n-2}(G, \mathbb{Z}) \xrightarrow{\alpha} H^n(Z, \mathbb{Z}) \xrightarrow{\beta} H^2(G, \mathbb{Z}). \tag{3.6}$$

Note that $\alpha(\ell_M) = M$ for every $M \in \mathcal{F}_m(Z)$. We set $E_M := \beta(M) \in H^2(G,\mathbb{Z})$ for every $M \in \mathcal{F}_m(Z)$.

3.7 Proposition [6, Proposition 2.2] *Both α and β are isomorphisms, and they are dual to each other with respect to the intersection products. Namely,*

$$x \cdot \beta(y) = \alpha(x) \cdot y \quad \textit{for every } x \in H^{2n-2}(G,\mathbb{Z}) \textit{ and } y \in H^n(Z,\mathbb{Z}).$$

Proof Since $\alpha(\ell_M) = M$ and the classes $\{M\}_{M \in \mathcal{F}_m(Z)}$ generate $H^n(Z,\mathbb{Z})$, the homomorphism α is surjective. Then α must be an isomorphism, because $H^{2n-2}(G,\mathbb{Z})$ and $H^n(Z,\mathbb{Z})$ are free of the same rank.

It follows from properties of Poincaré duality that $\alpha^t = (e_* \circ \pi^*)^t = (\pi^*)^t \circ (e_*)^t = \pi_* \circ e^* = \beta$, so α is the transpose homomorphism of β. It follows that β must be an isomorphism too. □

3.8 Corollary *We have $\beta(\eta) = -K_G$.*

Proof Using Proposition 3.7, for every $M \in \mathcal{F}_m(Z)$ we have

$$1 = \eta \cdot M = \eta \cdot \alpha(\ell_M) = \beta(\eta) \cdot \ell_M = -K_G \cdot \ell_M.$$

Since α is an isomorphism, and the classes $\{M\}_{M \in \mathcal{F}_m(Z)}$ generate $H^n(Z,\mathbb{Z})$, the classes $\{\ell_M\}_{M \in \mathcal{F}_m(Z)}$ generate $H^{2n-2}(G,\mathbb{Z})$. This yields the statement. □

Consider the involution $\sigma_I\colon Z \to Z$ for $I \subseteq \{1,\dots,n+3\}$ defined in Section 2.2. It induces an involution of G, which we denote by the same symbol,

$$\sigma_I\colon G \to G, \quad [L] \mapsto [\sigma_I(L)].$$

Therefore the group $W' \cong (\mathbb{Z}/2\mathbb{Z})^{n+2}$ generated by the involutions σ_i acts on G, $H^2(G,\mathbb{Z})$ and $H^{2n-2}(G,\mathbb{Z})$. It also acts on the incidence variety $\mathcal{I}$ in such a way that both morphisms π and e are W'–equivariant. It follows that the isomorphisms α and β are W'–equivariant.

3.9 Proposition *For every $M \in \mathcal{F}_m(Z)$, ℓ_M generates an extremal ray of* $\mathrm{NE}(G)$.

Proof Fix $M_0 \in \mathcal{F}_m(Z)$ and $i \in \{1,\dots,n+3\}$ such that $[\pi_i(M_0)] \in T^\psi$, and follow Notation 2.8. By Corollary 3.4, we have

$$\alpha(\mathrm{NE}(\varphi_i)) = \mathrm{Cone}(M_I)_{|I|\mathrm{odd},\, i\notin I} \quad \text{and} \quad \alpha(\mathrm{NE}(\psi_i)) = \mathrm{Cone}(M_I)_{|I|\mathrm{even},\, i\notin I}.$$

By (2.24), these are facets of the cone $\mathcal{E} \subset H^n(Z,\mathbb{R})$, whose extremal rays are generated by the classes $M = \alpha(\ell_M)$ contained in these facets. Thus, for every $M \in \mathcal{F}_m(Z)$ the class ℓ_M generates an extremal ray of either $\mathrm{NE}(\varphi_i)$ or $\mathrm{NE}(\psi_i)$, and hence of $\mathrm{NE}(G)$. □

4 The blow-up X of $\mathbb{P}^n$ at $n+3$ points

Let $n \geq 3$ be an integer. Unless otherwise stated, in this section we do not assume that n is even. Let $\mathcal{P} = \{p_1, \dots, p_{n+3}\} \subset \mathbb{P}^n$ be a set of distinct points in general linear position, and denote by C the unique rational normal curve in $\mathbb{P}^n$ through these points. Let $X = X_{\mathcal{P}}$ be the blow-up of $\mathbb{P}^n$ at $p_1, \dots, p_{n+3}$. Notice that acting on $\mathcal{P} = \{p_1, \dots, p_{n+3}\}$ by permutations and projective automorphisms of $\mathbb{P}^n$ yields isomorphic varieties $X_{\mathcal{P}}$. The variety X and its birational geometry have been widely studied. We refer the reader to Dolgachev [10], Bauer [3], Mukai [21; 23], Castravet and Tevelev [9], Araujo and Massarenti [1] and Brambilla, Dumitrescu and Postinghel [7].

We have $\mathrm{Pic}(X) \cong H^2(X, \mathbb{Z})$ and $\mathcal{N}^1(X) \cong H^2(X, \mathbb{R})$. We denote by H the pullback to X of the hyperplane class in $\mathbb{P}^n$, and by E_i the exceptional divisor over the point p_i (as well as its class in $H^2(X, \mathbb{Z})$).

4.1 (special subvarieties of X) Given a subset $I \subset \{1, \dots, n+3\}$, with $|I| = d \leq n$, and an integer $0 \leq s \leq \frac{1}{2}(n-d)$, we consider the join

$$\mathrm{Join}(\langle p_i \rangle_{i \in I}, \mathrm{Sec}_{s-1}(C)) \subset \mathbb{P}^n.$$

(Here we write $\mathrm{Sec}_k(C)$ for the subvariety of $\mathbb{P}^n$ obtained as the closure of the union of all k–planes spanned by $k+1$ general points of C for $k \geq 0$; in particular $\mathrm{Sec}_0(C) = C$. We also set $\mathrm{Sec}_{-1}(C) = \varnothing$.)

This join has dimension equal to $d+2s-1$. We denote by $J_{I,s} \subset X$ the strict transform of $\mathrm{Join}(\langle p_i \rangle_{i \in I}, \mathrm{Sec}_{s-1}(C))$. When $d + 2s = n$ (so that $|I^c| = n+3-3 = 2s+3$ is odd) we denote the divisor $J_{I,s}$ and its class in $H^2(X, \mathbb{Z})$ by E_I; in particular, for $n = 2m$ even, $E_\varnothing = J_{\varnothing,m}$ is the strict transform of $\mathrm{Sec}_{m-1}(C)$. For $I = \{i\}^c$, we set $E_I = E_i$. For every $I \subset \{1, \dots, n+3\}$ with $|I^c| = 2s+3$ odd and $s \geq 0$, we have the following identity in $H^2(X, \mathbb{Z})$:

$$E_I = (s+1)H - (s+1)\sum_{i \in I} E_i - s \sum_{j \notin I} E_j. \tag{4.2}$$

By Castravet and Tevelev [9, Theorem 1.2], each E_I generates an extremal ray of $\mathrm{Eff}(X)$, and all extremal rays are of this form. Moreover, by [9, Theorem 1.3] and Mukai [23], X is a Mori dream space (MDS for short). We refer to Hu and Keel [18] for the definition and basic properties of MDSs. Here we only recall an important feature of a MDS, the Mori chamber decomposition of its effective cone.

4.3 (the Mori chamber decomposition) Let Y be a projective, normal and $\mathbb{Q}$–factorial MDS. The effective cone $\mathrm{Eff}(Y)$ admits a fan structure, called *Mori chamber decomposition* and denoted by $\mathrm{MCD}(Y)$, which can be described as follows (see Hu and

Keel [18, Proposition 1.11(2)] and Okawa [24, Section 2.2]). There are finitely many birational contractions (ie birational maps whose inverses do not contract any divisor) from Y to projective, normal and $\mathbb{Q}$–factorial MDSs, denoted by $g_i\colon Y \dashrightarrow Y_i$. The set $\operatorname{Exc}(g_i)$ of classes of exceptional prime divisors of g_i has cardinality $\rho(Y)-\rho(Y_i)$. The maximal cones $\mathcal{C}_i$ of the fan $\operatorname{MCD}(Y)$ are of the form:

$$\mathcal{C}_i = \operatorname{Cone}\big(g_i^*(\operatorname{Nef}(Y_i)), \operatorname{Exc}(g_i)\big).$$

By abuse of notation, we often write $\operatorname{Nef}(Y_i) \subset \operatorname{Eff}(Y)$ for $g_i^*(\operatorname{Nef}(Y_i)) \subset \operatorname{Eff}(Y)$. If $\operatorname{Exc}(g_i) = \varnothing$, then we say that $g_i\colon Y \dashrightarrow Y_i$ is a small $\mathbb{Q}$–factorial modification of Y. The movable cone $\operatorname{Mov}(Y)$ of Y is the union

$$\operatorname{Mov}(Y) = \bigcup_{\operatorname{Exc}(g_i)=\varnothing} \mathcal{C}_i.$$

An arbitrary cone $\sigma \in \operatorname{MCD}(Y)$ is of the form

$$\sigma = \operatorname{Cone}\big(f^*(\operatorname{Nef}(W)), \mathcal{E}\big),$$

where $f\colon Y \dashrightarrow W$ is a dominant rational map to a normal projective variety, which factors as $Y \overset{g_i}{\dashrightarrow} Y_i \xrightarrow{f_i} W$ for some i, where $f_i\colon Y_i \to W$ is the contraction of an extremal face of $\operatorname{Nef}(Y_i)$, and $\mathcal{E} \subset \operatorname{Exc}(g_i)$.

Given an effective divisor D on Y, its class in $\mathcal{N}^1(Y)$ lies in the relative interior of some cone in $\operatorname{MCD}(Y)$, say $\operatorname{Cone}\big(f^*(\operatorname{Nef}(W)), \mathcal{E}\big)$. The map $f\colon Y \dashrightarrow W$ coincides with the map $\varphi_{|mD|}$ for $m \gg 1$ divisible enough. In this case, we write Y_D for the variety W.

Now we go back to X. Our next goal is to describe the Mori chamber decomposition of $\operatorname{Eff}(X)$, following Mukai [23] and Bauer [3] (see also Araujo and Massarenti [1, Section 3]).

Let us consider the coordinates $(y, x_1, \dots, x_{n+3})$ in $H^2(X,\mathbb{R})$ induced by the basis $(H, E_1, \dots, E_{n+3})$, and consider the affine hyperplane

$$\mathcal{H} = \Big((n+1)y + \sum x_i = 1\Big) \subset H^2(X,\mathbb{R}).$$

It contains all the generators E_I of $\operatorname{Eff}(X)$ described above, as well as $\frac{1}{4}(-K_X)$.

We now observe that the convex hull of the E_I in $\mathcal{H}$ is a demihypercube. To see this, we need suitable coordinates in $\mathcal{H}$. For $i = 1, \dots, n+3$, set

$$\widetilde{\varepsilon}_i := \frac{1}{2}\Big(H - \sum_{j\neq i} E_j + E_i\Big). \tag{4.4}$$

Then $\{\widetilde{\varepsilon}_1, \dots, \widetilde{\varepsilon}_{n+3}\}$ is a basis for the linear subspace $((n+1)y + \sum x_i = 0)$, so that $\left(\frac{1}{4}(-K_X), \{\widetilde{\varepsilon}_1, \dots, \widetilde{\varepsilon}_{n+3}\}\right)$ induces affine coordinates $(\alpha_1, \dots, \alpha_{n+3})$ in $\mathcal{H} \cong \mathbb{R}^{n+3}$. The radial projection

$$H^2(X, \mathbb{R}) \smallsetminus \Big((n+1)y + \sum x_i = 0\Big) \to \mathcal{H}$$

is given in coordinates by

$$\alpha_i = \frac{y + x_i}{(n+1)y + \sum x_i} - \frac{1}{2} \quad \text{for } i = 1, \dots, n+3. \tag{4.5}$$

In the coordinates α_i, $\frac{1}{4}(-K_X)$ is identified with the origin, and E_I with v_{I^c}, with the notation introduced in Section 2.15. Thus $\operatorname{Eff}(X) \cap \mathcal{H}$ is identified with the demihypercube $\Delta \subset \mathbb{R}^{n+3}$ described in Section 2.15,

$$\Delta = \begin{cases} -\frac{1}{2} \le \alpha_i \le \frac{1}{2}, & i \in \{1, \dots, N\}, \\ H_I \ge 1, & |I| \text{ even.} \end{cases}$$

Recall the degree 1 polynomials H_I introduced in (2.16), and consider the hyperplane arrangement

$$(H_I = k)_{I \subset \{1, \dots, n+3\},\, k \in \mathbb{N},\, 2 \le k \le (n+3)/2,\, |I| \not\equiv k \bmod 2}\,. \tag{4.6}$$

It defines a subdivision of Δ in polytopes, and a fan structure on $\operatorname{Eff}(X)$, given by the cones over these polytopes. By Mukai [23] and Bauer [3], this fan coincides with $\operatorname{MCD}(X)$. Moreover, one has the following description of the wall crossings (see [23, Propositions 2 and 3] and also [3, Section 2]):

(1) The intersection of $\operatorname{Mov}(X)$ with the hyperplane $\mathcal{H}$ is given by

$$\Delta_{\mathrm{Mov}} = \operatorname{Mov}(X) \cap \mathcal{H} = \begin{cases} -\frac{1}{2} \le \alpha_i \le \frac{1}{2}, & i \in \{1, \dots, n+3\}, \\ H_I \ge 2, & |I| \text{ odd.} \end{cases}$$

(2) All small $\mathbb{Q}$–factorial modifications of X are smooth.

(3) Let $\mathcal{C}$ be a maximal cone of $\operatorname{MCD}(X)$, contained in $\operatorname{Mov}(X)$, corresponding to a small $\mathbb{Q}$–factorial modification $\widetilde{X}$ of X. Let $\sigma \subset \partial\mathcal{C}$ be a wall such that $\sigma \subset \partial \operatorname{Mov}(X)$, and let $f\colon \widetilde{X} \to Y$ be the corresponding elementary contraction. Then $\sigma \cap \mathcal{H} \subset \Delta_{\mathrm{Mov}}$ is supported on a hyperplane of one of the following forms:

- (a) $\left(\alpha_i = -\frac{1}{2}\right)$ or $\left(\alpha_i = \frac{1}{2}\right)$.
- (b) $(H_I = 2)$, with $|I|$ odd.

In case (a), $f\colon \widetilde{X} \to Y$ is a $\mathbb{P}^1$–bundle. In case (b), $f\colon \widetilde{X} \to Y$ is the blow-up of a smooth point, and the exceptional divisor of f is the strict transform in $\widetilde{X}$ of the divisor $E_{I^c} \subset X$.

(4) Let $\mathcal{C}$ and $\mathcal{C}'$ be two maximal cones of $\mathrm{MCD}(X)$, contained in $\mathrm{Mov}(X)$ and having a common facet. Let $f\colon X \dashrightarrow \widetilde{X}$ and $f'\colon X \dashrightarrow \widetilde{X}'$ be the corresponding small $\mathbb{Q}$–factorial modifications of X. The intersections of these cones with $\mathcal{H}$ are separated in Δ by a hyperplane of the form $(H_I = k)$, with $3 \le k \le \frac{1}{2}(n+3)$ and $|I| \not\equiv k \mod 2$. Suppose that $\mathcal{C} \cap \mathcal{H} \subset (H_I \le k)$ and $\mathcal{C}' \cap \mathcal{H} \subset (H_I \ge k)$. Then the birational map $f' \circ f^{-1}\colon \widetilde{X} \dashrightarrow \widetilde{X}'$ flips a $\mathbb{P}^{k-2}$ into a $\mathbb{P}^{n+1-k}$.

4.7 Remark It is possible to give a more precise description of the flipping locus $\mathbb{P}^{k-2} \subset \widetilde{X}$ (or $\mathbb{P}^{n+1-k} \subset \widetilde{X}'$) in the situation described under (4) above (see [3, Proposition 2.6(iv) and Theorem 2.9]): Consider the nef cone of X and its section with $\mathcal{H}$,

$$\Delta_{\mathrm{Nef}} = \mathrm{Nef}(X) \cap \mathcal{H} = \begin{cases} H_{\{i\}} \ge 2, & i \in \{1, \dots, n+3\}, \\ H_{\{i,j\}} \le 3, & i, j \in \{1, \dots, n+3\},\ i \ne j. \end{cases}$$

Suppose that $\Delta_{\mathrm{Nef}} \subset (H_I \le k)$. Then the $\mathbb{P}^{k-2} \subset \widetilde{X}$ flipped by $f' \circ f^{-1}$ is the strict transform in $\widetilde{X}$ of the special variety $J_{I,s} \subset X$, where $s = \frac{1}{2}(k - |I| - 1) \ge 0$.

Suppose that $\Delta_{\mathrm{Nef}} \subset (H_I \ge k)$. Then the $\mathbb{P}^{n+1-k} \subset \widetilde{X}'$ flipped by $f \circ (f')^{-1}$ is the strict transform in $\widetilde{X}'$ of the special variety $J_{I^c,s'} \subset X$, where $s' = \frac{1}{2}(|I| - k - 1) \ge 0$.

4.8 Remark Recall from Section 2.15 the description of the facets of Δ. Each of the $2(n+3)$ facets of Δ supported on the hyperplanes $\left(\alpha_i = \pm\frac{1}{2}\right)$ intersects Δ_{Mov} along a facet, while the other facets of Δ, supported on the hyperplanes $(H_I = 1)$ for $|I|$ even, are disjoint from Δ_{Mov}. Let us describe the rational maps associated to the facets of Δ_{Mov} supported on the hyperplanes $\left(\alpha_i = \pm\frac{1}{2}\right)$.

Fix $i \in \{1, \dots, n+3\}$ and let $\mathcal{P}_i \subset \mathbb{P}^{n-1}$ be the image of the set $\mathcal{P} \smallsetminus \{p_i\}$ under the projection $\pi_{p_i}\colon \mathbb{P}^n \dashrightarrow \mathbb{P}^{n-1}$ from p_i. Let $Y = (X_{\mathcal{P}_i})^{n-1}$ be the blow-up of $\mathbb{P}^{n-1}$ at the $n+2$ points in $\mathcal{P}_i$.

There is a small $\mathbb{Q}$–factorial modification $X \dashrightarrow X_i$ and a $\mathbb{P}^1$–bundle $X_i \to Y$ extending π_{p_i} (see [23, Example 1]). Let $\pi_i\colon X \dashrightarrow Y$ be the composite map. The general fiber of π_i is the strict transform in X of a general line in $\mathbb{P}^n$ through p_i. The hyperplane $(\pi_i)^* H^2(Y, \mathbb{R})$ has equation $y + x_i = 0$. Using (4.5), we see that $(\pi_i)^* H^2(Y, \mathbb{R}) \cap \mathcal{H}$ is the hyperplane $\left(\alpha_i = -\frac{1}{2}\right)$. Thus the cone $(\pi_i)^* \mathrm{Eff}(Y)$ is the cone over the polytope $\Delta \cap \left(\alpha_i = -\frac{1}{2}\right)$, which is an $(n+2)$–dimensional demihypercube.

Similarly, there is a map $\pi_i'\colon X \dashrightarrow Y$ whose general fiber is the strict transform in X of a general rational normal curve through the points p_λ for $\lambda \ne i$. Indeed, fix $j \ne i$ and let $\varphi\colon \mathbb{P}^n \dashrightarrow \mathbb{P}^n$ be the standard Cremona transformation centered at the points p_λ for $\lambda \ne i, j$. This map sends rational normal curves through the points p_λ for $\lambda \ne i$ to lines through $\varphi(p_j)$. There is an automorphism of $\mathbb{P}^n$ fixing p_λ for $\lambda \ne i, j$, sending p_j to $\varphi(p_i)$ and sending p_i to $\varphi(p_j)$ (see Remark 7.2). By composing φ with the projection from $\varphi(p_j)$, we obtain a rational map $\pi'_{p_i}\colon \mathbb{P}^n \dashrightarrow Y$ whose general fiber is a general rational normal curve through the points p_λ for $\lambda \neq i$. This yields a $\mathbb{P}^1$–bundle $X'_i \to Y$ on a small $\mathbb{Q}$–factorial modification of X, and the desired map $\pi'_i\colon X \dashrightarrow Y$. As before, one checks that $(\pi_i)^* \operatorname{Eff}(Y)$ is the cone over the demihypercube $\Delta \cap \left(\alpha_i = \frac{1}{2}\right)$.

The center of the polytopes Δ_{Mov} and Δ is the origin $\overline{0} \in \mathbb{R}^{n+3}$, which corresponds to $\frac{1}{4}(-K_X)$. In particular, the divisor $-K_X$ is movable. We want to describe the Fano model $X^n_{\text{Fano}} := X_{-K_X}$.

If n is odd, then $\overline{0}$ is a vertex in the subdivision of Δ and is contained in the intersection of the hyperplanes

$$\left(H_I = \tfrac{1}{2}(n+3)\right)_{|I| \not\equiv (n+3)/2 \mod 2}.$$

Thus $-K_X$ lies in a one-dimensional cone of the fan $\operatorname{MCD}(X)$, contained in the interior of $\operatorname{Mov}(X)$. Therefore X^n_{Fano} is non-$\mathbb{Q}$–factorial and has Picard number 1.

For the remainder of this section, we assume that $n = 2m \geq 2$ is even. Then $\overline{0}$ lies in the interior of a maximal polytope in the subdivision of Δ_{Mov}, namely the polytope defined by

$$\Delta_{\text{Fano}} = (H_I \geq m+1)_{|I| \equiv m \mod 2}. \tag{4.9}$$

Then X^n_{Fano} is a small $\mathbb{Q}$–factorial modification of X, it is a smooth Fano manifold, and $\operatorname{Nef}(X^n_{\text{Fano}}) \subset \operatorname{Eff}(X)$ is the cone over the polytope Δ_{Fano}.

4.10 Remark By Theorem 1.4, when $\mathcal{P}$ is the image of $\{(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1)\} \subset \mathbb{P}^1$ under a Veronese embedding $\mathbb{P}^1 \hookrightarrow \mathbb{P}^n$, X is pseudoisomorphic to the Fano variety G addressed in Section 3. This implies that X^n_{Fano} is isomorphic to G.

4.11 Using the properties of MDSs, and the description of $\operatorname{MCD}(X)$ above, we can deduce many properties of X^n_{Fano}:

- The Mori cone $\operatorname{NE}(X^n_{\text{Fano}})$ admits exactly 2^{n+2} extremal rays, whose corresponding contractions all contract a $\mathbb{P}^m$ to a point.
- The variety X^n_{Fano} admits $2(n+3)$ distinct (nontrivial) contractions of fiber type. Indeed, the points in $\partial\Delta_{\text{Mov}} \cap \Delta_{\text{Fano}}$ are those of the form $\alpha = (\alpha_1, \dots, \alpha_{n+3})$, where $\alpha_i = -\frac{1}{2}$ or $\frac{1}{2}$ for some fixed i, and $\alpha_j = 0$ for $j \neq i$. These points all lie in $\partial\Delta$. We denote the corresponding contractions by ϕ_i and ϕ'_i, respectively.

4.12 Lemma *The morphisms ϕ_i and ϕ'_i are generic $\mathbb{P}^1$–bundles over $(X_{\mathcal{P}_i})^{n-1}_{\text{Fano}}$, where $\mathcal{P}_i \subset \mathbb{P}^{n-1}$ is as in Remark 4.8. The general fiber of ϕ_i is the strict transform in X^n_{Fano} of a general line in $\mathbb{P}^n$ through p_i. The general fiber of ϕ'_i is the strict transform in X^n_{Fano} of a general rational normal curve in $\mathbb{P}^n$ through $\mathcal{P} \smallsetminus \{p_i\}$.*

Proof Let $\alpha=(\alpha_1,\dots,\alpha_{n+3})$, where $\alpha_i=-\frac{1}{2}$ and $\alpha_j=0$ for $j\neq i$, and consider the corresponding fibration $\phi_i\colon X^n_{\mathrm{Fano}}\to X_D$, where D is an effective divisor such that $\mathbb{R}_{\geq 0}[D]\cap\mathcal{H}=\alpha$.

Consider the map $\pi_i\colon X \dashrightarrow Y:=(X_{\mathcal{P}_i})^{n-1}$ introduced in Remark 4.8, and recall that $(\pi_i)^*\operatorname{Eff}(Y)$ is the cone over the $(n+2)$–dimensional demihypercube $\Delta\cap\big(\alpha_i=-\frac{1}{2}\big)$. The center of this demihypercube is α, hence D is a positive multiple of $(\pi_i)^*(-K_Y)$. So the image X_D of ϕ_i is precisely the Fano model $(X_{\mathcal{P}_i})^{n-1}_{\mathrm{Fano}}$ of Y.

A similar argument shows the statement for ϕ_i'. □

4.13 Let $(z,t_1,\dots,t_{n+3})$ be new coordinates in $H^2(X,\mathbb{R})$, induced by the basis $\{-K_X,E_1,\dots,E_{n+3}\}$. These are related to $(y,x_1,\dots,x_{n+3})$ by $y=z(n+1)$ and $x_i=t_i-(n-1)z$. Using the defining inequalities for Δ_{Fano} in (4.9), and the expression for the radial projection onto $\mathcal{H}$ in (4.5), we conclude that $\operatorname{Nef}(X^n_{\mathrm{Fano}})\subset H^2(X,\mathbb{R})$ is defined by the inequalities

$$2z+(|I|-m)\sum_{i=1}^{n+3}t_i-2\sum_{i\in I}t_i\geq 0 \tag{4.14}$$

for every $I\subseteq\{1,\dots,n+3\}$ such that $|I|\equiv m \bmod 2$.

4.15 We end this section by describing the birational map $X\dashrightarrow X^n_{\mathrm{Fano}}$. First notice that to go from the interior of the polytope $\Delta_{\mathrm{Nef}}=\operatorname{Nef}(X)\cap\mathcal{H}$ to the interior of the polytope $\Delta_{\mathrm{Fano}}=\operatorname{Nef}(X^n_{\mathrm{Fano}})\cap\mathcal{H}$, we must cross the wall $(H_I=k)$ for every $I\subset\{1,\dots,n+3\}$ and $3\leq k\leq m+1$ such that $|I|\not\equiv k \bmod 2$ and $|I|\leq k-1$. By Remark 4.7 and [3, Theorem 2.9], we conclude that the rational map $X\dashrightarrow X^n_{\mathrm{Fano}}$ factors as

$$X=X_0\overset{\varphi_1}{\dashrightarrow}X_1\overset{\varphi_2}{\dashrightarrow}X_2\dashrightarrow\cdots\overset{\varphi_{m-1}}{\dashrightarrow}X_{m-1}=X^n_{\mathrm{Fano}},$$

where each $\varphi_i\colon X_{i-1}\dashrightarrow X_i$ flips the strict transforms in X_{i-1} of all special subvarieties $J_{I,s}\subset X$ of dimension i. These strict transforms are disjoint in X_{i-1} and each isomorphic to $\mathbb{P}^i$. The flipped locus on X_i is a disjoint union of copies of $\mathbb{P}^{n-1-i}$, one for each $J_{I,s}$ of dimension i. Notice that in general the map φ_i is not the flip of a small contraction: it is a pseudoisomorphism that can be factored as a sequence of flips.

In particular, we can describe the 2^{n+2} copies of $\mathbb{P}^m$ in X^n_{Fano} corresponding to the 2^{n+2} extremal rays of $\operatorname{NE}(X^n_{\mathrm{Fano}})$. These are the strict transforms of the special subvarieties $J_{I,s}\subset X$ of dimension m, and the flipped locus of the flips of the strict transforms of the special subvarieties $J_{I,s}\subset X$ of dimension $m-1$. These are, respectively,

$$\sum_{\substack{d=0 \\ d\not\equiv m \bmod 2}}^{m+1} \binom{n+3}{d} \quad \text{for } m\text{–dimensional } J_{I,s},$$

$$\sum_{\substack{d=0 \\ d\equiv m \bmod 2}}^{m} \binom{n+3}{d} \quad \text{for } (m-1)\text{–dimensional } J_{I,s}.$$

We can also describe the strict transforms in X^n_{Fano} of the divisors $\mathbb{P}^{n-1} \cong E_i \subset X$ under the rational map $X \dashrightarrow X^n_{\mathrm{Fano}}$. There are $n+3$ special points $q_1, \dots, q_{n+3} \subset E_i$: q_j is the intersection of E_i with the strict transform of the line through p_i and p_j when $j \neq i$, and q_i is the intersection of E_i with the strict transform of C. The points q_i all lie in a rational normal curve C' of degree $n-1$ in $E_i \cong \mathbb{P}^{n-1}$. Given a subset $I \subset \{1, \dots, n+3\}$, with $|I| \le n-1$, and an integer $0 \le s \le \frac{1}{2}(n-1-|I|)$, we denote by $J^i_{I,s}$ the join $\mathrm{Join}(\langle q_j \rangle_{j\in I}, \mathrm{Sec}_{s-1}(C')) \subset E_i$. One can check that

$$E_i \cap J_{I,s} = \begin{cases} J^i_{I\smallsetminus\{i\},s} & \text{if } i \in I, \\ \varnothing & \text{if } i \notin I \text{ and } s = 0, \\ J^i_{I\cup\{i\},s-1} & \text{if } i \notin I \text{ and } s \ge 1. \end{cases}$$

Therefore, the strict transform of E_i under φ_1 is the blow-up of $\mathbb{P}^{n-1}$ at the points $q_1, \dots, q_{n+3}$. For $2 \le j \le m-1$, the restriction of φ_j to the strict transform of E_i in X_{j-1} flips the strict transforms of every $J^i_{I,s}$ of dimension $j-1$.

4.16 When $n = 4$, the birational map $\varphi_1\colon X = X_0 \dashrightarrow X_1 = X^4_{\mathrm{Fano}}$ flips $J_{\{ij\},0}$ (strict transform of the line $\overline{p_i p_j} \subset \mathbb{P}^4$) for $1 \le i, j \le 7$, and $J_{\varnothing,1}$ (strict transform of $C \subset \mathbb{P}^4$); this yields 22 among the 64 special copies of $\mathbb{P}^2$ in X^4_{Fano}, corresponding to the 64 extremal rays of $\mathrm{NE}(X^4_{\mathrm{Fano}})$. The remaining ones are the strict transforms of the 7 surfaces $\mathrm{Join}(\langle p_i \rangle, C)$ and of the 35 planes $\langle p_i, p_j, p_h \rangle$ in $\mathbb{P}^4$.

Notice in particular that $E_i \subset X$ does not contain any special subvariety $J_{I,s}$, while the strict transform of E_i in X^4_{Fano} contains 7 special copies of $\mathbb{P}^2$, namely the flipped loci of the flips of $J_{\{ij\},0}$ for $j \neq i$ and of $J_{\varnothing,1}$.

5 Pseudoisomorphisms between G and X

Let m be a positive integer, and set $n = 2m$. Fix $n+3$ distinct points

$$(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1) \in \mathbb{P}^1,$$

and let $p_1, \dots, p_{n+3} \in \mathbb{P}^n$ be their images under a Veronese embedding $\mathbb{P}^1 \hookrightarrow \mathbb{P}^n$. Let Z, G and X be the varieties introduced in Sections 2, 3 and 4. We follow the notation introduced in those sections. In this section we determine the nef cone of G, and then we prove Theorem 1.5, which follows from Theorem 5.7 and Corollary 5.8. Our aim is to identify the line bundles on G whose linear systems define rational maps $G \dashrightarrow \mathbb{P}^n$ inducing a pseudoisomorphism $G \dashrightarrow X$. This is achieved by combining the description of $\mathrm{Nef}(G) \subset H^2(G,\mathbb{R})$ given by Theorem 5.1 and the description of $\mathrm{Nef}(X^n_{\mathrm{Fano}}) \subset H^2(X,\mathbb{R})$ in terms of the basis $\{-K_X, E_1, \dots, E_{n+3}\}$ for $H^2(X,\mathbb{R})$, which was obtained from the Mori chamber decomposition of $\mathrm{Eff}(X)$ in Section 4.

We first describe the cones $\mathrm{Nef}(G)$ and $\mathrm{NE}(G)$. For $n=4$, this was proved by Borcea [6, Theorem 4.3].

5.1 Theorem *Let the notation be as above. Then*

$$\mathrm{NE}(G) = \mathrm{Cone}(\ell_M)_{M\in\mathcal{F}_m(G)} = \alpha^{-1}(\mathcal{E}) \quad \textit{and} \quad \mathrm{Nef}(G) = \beta(\mathcal{E}^\vee).$$

Proof By Proposition 3.9, the class ℓ_M generates an extremal ray of $\mathrm{NE}(G)$ for every $M \in \mathcal{F}_m(G)$. This yields 2^{n+2} distinct extremal rays of $\mathrm{NE}(G)$. On the other hand, $G \cong X_{\mathrm{Fano}}$ by Remark 4.10, and $\mathrm{NE}(X_{\mathrm{Fano}})$ has precisely 2^{n+2} extremal rays, as explained in Section 4.11. So we have

$$\mathrm{NE}(G) = \mathrm{Cone}(\ell_M)_{M\in\mathcal{F}_m(G)} = \alpha^{-1}(\mathcal{E}).$$

The equality $\mathrm{Nef}(G) = \beta(\mathcal{E}^\vee)$ follows from the duality between $\mathrm{Nef}(G)$ and $\mathrm{NE}(G)$ and from Proposition 3.7. □

Similarly, we will show in Proposition 5.5 that $\mathrm{Eff}(G) = \beta(\mathcal{E})$ and $\mathrm{Mov}_1(G) = \alpha^{-1}(\mathcal{E}^\vee)$. So the cones $\mathrm{NE}(G)$ and $\mathrm{Eff}(G)$ are isomorphic under $\beta\circ\alpha$, and the same holds for $\mathrm{Mov}_1(G)$ and $\mathrm{Nef}(G)$.

Recall from Section 3 that $E_M = \beta(M) \in H^2(G,\mathbb{Z})$ for every $M \in \mathcal{F}_m(Z)$. For each $M \in \mathcal{F}_m(Z)$, consider the linear map

$$h_M\colon H^2(X,\mathbb{R}) \to H^2(G,\mathbb{R})$$

defined by

$$h_M(-K_X) = -K_G \qquad \text{and} \qquad h_M(E_i) = E_{\sigma_i(M)} \quad \text{for every } i = 1,\dots,n+3.$$

One can check that h_M respects the integral points, namely that it is induced by an isomorphism $H^2(X,\mathbb{Z}) \to H^2(G,\mathbb{Z})$, and that $h_{\sigma_I(M)} = \sigma_I \circ h_M$ for every $I \subseteq \{1,\dots,n+3\}$.

We also set

$$\tilde{h}_M := \beta^{-1}\circ h_M\colon H^2(X,\mathbb{R}) \to H^n(Z,\mathbb{R}), \tag{5.2}$$

so that $\tilde{h}_M(-K_X) = \eta$ and $\tilde{h}_M(E_i) = \sigma_i(M)$ for every $i = 1,\dots,n+3$.

5.3 Lemma *For every $M \in \mathcal{F}_m(Z)$ and $I \subseteq \{1,\dots,n+3\}$ of even cardinality, we have*

$$h_M(E_I) = E_{\sigma_I(M)}, \quad h_M(\mathrm{Eff}(X)) = \beta(\mathcal{E}), \quad h_M(\mathrm{Nef}(X_{\mathrm{Fano}})) = \mathrm{Nef}(G).$$

Proof Let $I \subseteq \{1,\dots,n+3\}$ be such that $|I| = n-2s$ is even with $s \geq 0$. We can rewrite (4.2) as

$$E_I = \frac{1}{n+1}\left((s+1)(-K_X) - 2(s+1)\sum_{i\in I} E_i + (n-1-2s)\sum_{j\in I^c} E_j\right).$$

It follows from (2.13) that $\tilde{h}_M(E_I) = \sigma_I(M)$, and hence $h_M(E_I) = E_{\sigma_I(M)}$. This implies that $h_M(\mathrm{Eff}(X)) = \beta(\mathcal{E})$.

By comparing (4.14) and (2.26), we see that $\tilde{h}_M(\mathrm{Nef}(X^n_{\mathrm{Fano}})) = \mathcal{E}^\vee$. Therefore $h_M(\mathrm{Nef}(X^n_{\mathrm{Fano}})) = \beta(\mathcal{E}^\vee) = \mathrm{Nef}(G)$ by Theorem 5.1. □

5.4 Proposition *Let $\xi\colon G \dashrightarrow X$ be a pseudoisomorphism, and consider the induced linear map*

$$\xi^*\colon H^2(X,\mathbb{R}) \to H^2(G,\mathbb{R}).$$

Then, up to a unique permutation of $E_1,\dots,E_{n+3} \subset X$, there is a unique $M \in \mathcal{F}_m(Z)$ such that $\xi^* = h_M$.

Proof We have $\xi^*(-K_X) = -K_G$, and hence $\xi^*(\mathrm{Nef}(X^n_{\mathrm{Fano}})) = \mathrm{Nef}(G)$.

Recall Notation 2.8; fix $M_0 \in \mathcal{F}_m(Z)$. Consider $\xi^* \circ (h_{M_0})^{-1}\colon H^2(G,\mathbb{R}) \to H^2(G,\mathbb{R})$. By Lemma 5.3, this map fixes $-K_G$ and sends $\mathrm{Nef}(G)$ to itself. Using the isomorphism $\beta\colon H^n(Z,\mathbb{R}) \to H^2(G,\mathbb{R})$ and Theorem 5.1, we obtain a linear map $f\colon H^n(Z,\mathbb{R}) \to H^n(Z,\mathbb{R})$ such that $f(\eta) = \eta$ and $f(\mathcal{E}^\vee) = \mathcal{E}^\vee$:

$$\begin{tikzcd}
 & H^2(X,\mathbb{R}) \arrow[dl, "h_{M_0}"'] \arrow[dr, "\xi^*"] & \\
H^2(G,\mathbb{R}) \arrow[rr, "\xi^*\circ(h_{M_0})^{-1}"] & & H^2(G,\mathbb{R}) \\
H^n(Z,\mathbb{R}) \arrow[u, "\beta"] \arrow[rr, "f"] & & H^n(Z,\mathbb{R}) \arrow[u, "\beta"']
\end{tikzcd}$$

By Lemma 2.27, we have $f \in W(D_{n+3})$.

Consider the stabilizer $G_0 \subset W(D_{n+3})$ of M_0, and recall that $W(D_{n+3}) = W' \rtimes G_0$ and $G_0 \cong S_{n+3}$. Thus there are uniquely defined $\omega \in G_0$, $\sigma_I \in W'$ and $\kappa \in S_{n+3}$

such that $f = \sigma_I \circ \omega$ and $\omega(M_i) = M_{\kappa(i)}$ for every $i = 1, \dots, n+3$. Since β is W'–equivariant, this means that

$$\xi^*(E_i) = \beta(f(M_i)) = \beta(\sigma_{\kappa(i)}(M_I)) = \sigma_{\kappa(i)}(\beta(M_I)) = \sigma_{\kappa(i)}(E_{M_I})$$

for every $i = 1, \dots, n+3$. Apply the permutation κ^{-1} to $E_1, \dots, E_{n+3} \subset X$. After this reordering, we get $f = \sigma_I \in W'$ and $\xi^* = \sigma_I \circ h_{M_0} = h_{M_I}$. □

From now on we order the divisors $E_1, \dots, E_{n+3} \subset X$, and correspondingly the points $p_1, \dots, p_{n+3} \in \mathbb{P}^n$, as in Proposition 5.4. At this point we can determine the cone of effective divisors and the cone of moving curves of G.

5.5 Proposition *For every $M \in \mathcal{F}_m(Z)$, there is a unique effective divisor in G with class $E_M \in H^2(G, \mathbb{Z})$. This is a fixed prime divisor, which we still denote by $E_M \subset G$. We have*

$$\mathrm{Eff}(G) = \beta(\mathcal{E}) = \mathrm{Cone}(E_M)_{M \in \mathcal{F}_m(Z)} \quad \textit{and} \quad \mathrm{Mov}_1(G) = \alpha^{-1}(\mathcal{E}^\vee).$$

Proof By Theorem 1.4, there exists a pseudoisomorphism $\xi\colon G \dashrightarrow X$. By Proposition 5.4 there exists $M \in \mathcal{F}_m(Z)$ such that $\xi^* = h_M$. In particular, for every $I \subset \{1, \dots, n+3\}$ with $|I|$ even, we have $\xi^*(E_I) = E_{\sigma_I(M)}$ by Lemma 5.3. Thus the strict transform in G of $E_I \subset X$ is a fixed prime divisor, and it is the unique effective divisor with class $E_{\sigma_I(M)}$. It also follows from Lemma 5.3 that

$$\mathrm{Eff}(G) = \xi^* \, \mathrm{Eff}(X) = \beta(\mathcal{E}) = \mathrm{Cone}(E_M)_{M \in \mathcal{F}_m(Z)}.$$

The equality $\mathrm{Mov}_1(G) = \alpha^{-1}(\mathcal{E}^\vee)$ follows from the duality $\mathrm{Mov}_1(G) = \mathrm{Eff}(G)^\vee$ and from Proposition 3.7. □

For each $M \in \mathcal{F}_m(Z)$, we set

$$H_M := h_M(H) = \frac{1}{n+1}\left(-K_G + (n-1)\sum_{i=1}^{n+3} E_{\sigma_i(M)}\right) \tag{5.6}$$
$$= m(-K_G) - (n-1)E_M \in H^2(G, \mathbb{Z}),$$

where the last equality follows from (2.13) (taking $M = M_0$ and $I = \varnothing$), using the isomorphism $\beta\colon H^n(Z, \mathbb{R}) \to H^2(G, \mathbb{R})$.

5.7 Theorem *For every $M \in \mathcal{F}_m(Z)$, the divisor class H_M is movable, and its complete linear system defines a birational map*

$$\rho_M\colon G \dashrightarrow \mathbb{P}^n,$$

with exceptional divisors $E_{\sigma_1(M)}, \dots, E_{\sigma_{n+3}(M)}$*, inducing a pseudoisomorphism*

$$\xi_M \colon G \dashrightarrow X$$

whose induced map $\xi_M^* \colon H^2(X,\mathbb{R}) \to H^2(G,\mathbb{R})$ *coincides with* h_M.

For every $I \subseteq \{1,\dots,n+3\}$, $\rho_{\sigma_I(M)} = \rho_M \circ \sigma_I$ *and* $\xi_{\sigma_I(M)} = \xi_M \circ \sigma_I$.

Proof By Theorem 1.4, there exists a pseudoisomorphism $\xi \colon G \dashrightarrow X$. Let the map $\rho \colon G \dashrightarrow \mathbb{P}^n$ be the composition of ξ with the blow-up morphism $X \to \mathbb{P}^n$.

By Proposition 5.4, there exists $M_0 \in \mathcal{F}_m(Z)$ such that $\xi^* = h_{M_0}$. This implies that $\rho^*(\mathcal{O}_{\mathbb{P}^n}(1)) = H_{M_0}$. Hence the class H_{M_0} is movable, and $H^0(G, H_{M_0}) \cong H^0(\mathbb{P}^n, \mathcal{O}_{\mathbb{P}^n}(1))$. This proves the first statement for $M = M_0$, with $\rho_{M_0} = \rho$ and $\xi_{M_0} = \xi$.

Let $I \subseteq \{1,\dots,n+3\}$. We use Notation 2.8. The automorphism $\sigma_I \colon G \to G$ fixes $-K_G$ and maps E_{M_0} to E_{M_I}, hence it maps H_{M_0} to H_{M_I}. This yields the first statement for $M = M_I$, with $\rho_{M_I} = \rho \circ \sigma_I$ and $\xi_{M_I} = \xi \circ \sigma_I$.

The last statement is clear. □

5.8 Corollary *Let* $\widetilde{X}$ *be any blow-up of* $\mathbb{P}^n$ *at* $n+3$ *points. If* $\widetilde{X}$ *is pseudoisomorphic to* G*, then* $\widetilde{X}$ *is isomorphic to* X.

Proof Let $\tilde{\xi} \colon G \dashrightarrow \widetilde{X}$ be a pseudoisomorphism, and let $\tilde{\rho} \colon G \dashrightarrow \mathbb{P}^n$ be the composition of $\tilde{\xi}$ with the blow-up morphism $\widetilde{X} \to \mathbb{P}^n$. Then $\tilde{\rho}$ has $n+3$ exceptional prime divisors, whose classes must generate a simplicial facet of $\mathrm{Eff}(G)$. By Proposition 5.5 and the description of the facets of $\mathcal{E}$ in Remark 2.23, every simplicial facet of $\mathrm{Eff}(G)$ is generated by $E_{\sigma_1(M)}, \dots, E_{\sigma_{n+3}(M)}$ for some $M \in \mathcal{F}_m(Z)$. Since each $E_{\sigma_i(M)}$ is unique in its linear system, $\tilde{\rho} \colon G \dashrightarrow \mathbb{P}^n$ and $\rho_M \colon G \dashrightarrow \mathbb{P}^n$ have the same exceptional divisors. This means that $\tilde{\rho}$ and ρ_M coincide up to a projective transformation of $\mathbb{P}^n$, and therefore $\widetilde{X} \cong X$. □

5.9 Remark (comparing the intersection product in $H^n(Z,\mathbb{Z})$ with Dolgachev's pairing on $H^2(X,\mathbb{Z})$) In [10], Dolgachev defined a nondegenerate symmetric bilinear form $(\ ,\ )$ on $H^2(X,\mathbb{Z})$, by imposing that the basis $H, E_1, \dots, E_{n+3}$ is orthogonal,

$$(H,H) = n-1 \qquad \text{and} \qquad (E_i,E_i) = -1 \quad \text{for all } i = 1,\dots,n+3.$$

This pairing has signature $(1, n+3)$, and $(-K_X, -K_X) = 4(n-1)$. Consider $\tilde{\varepsilon}_i \in H^2(X,\mathbb{R})$, defined in (4.4),

$$\tilde{\varepsilon}_i := \frac{1}{2}\Big(H - \sum_{j \neq i} E_j + E_i\Big) \quad \text{for } i = 1,\dots,n+3.$$

Then we have

$$(-K_X, \widetilde{\varepsilon}_i) = 0 \qquad \text{and} \qquad (\widetilde{\varepsilon}_i, \widetilde{\varepsilon}_j) = -\delta_{ij} \quad \text{for every } i, j = 1, \dots, n+3,$$

thus $-K_X, \widetilde{\varepsilon}_1, \dots, \widetilde{\varepsilon}_{n+3}$ is another orthogonal basis for $H^2(X, \mathbb{R})$.

Fix $M_0 \in \mathcal{F}_m(Z)$, and consider the orthogonal basis $\eta, \varepsilon_1, \dots, \varepsilon_{n+3}$ for $H^n(Z, \mathbb{R})$ introduced in (2.9). Recall that $\eta^2 = 4$ and $\varepsilon_i^2 = (-1)^m$ for every $i = 1, \dots, n+3$. Consider the isomorphism introduced in (5.2),

$$\tilde{h}_{M_0}\colon H^2(X, \mathbb{R}) \to H^n(Z, \mathbb{R}).$$

From (5.6) and (2.14) we have $\tilde{h}_{M_0}(\widetilde{\varepsilon}_i) = \varepsilon_i$ for every $i = 1, \dots, n+3$. Therefore $\tilde{h}_{M_0}$ maps an orthogonal basis for Dolgachev's pairing in $H^2(X, \mathbb{R})$ to an orthogonal basis for the intersection product in $H^n(Z, \mathbb{R})$. In particular, $\tilde{h}_{M_0}$ sends the D_{n+3}–lattice $(-K_X)^{\perp} \subset H^2(X, \mathbb{Z})$ to the D_{n+3}–lattice $\eta^{\perp} \subset H^n(Z, \mathbb{Z})$, and the restriction of $\tilde{h}_{M_0}$ to these lattices is an isometry up to the sign $(-1)^{m-1}$. (Notice that $\tilde{h}_{M_0}$ is globally an isometry if and only if $n = 2$.) This also shows that $\tilde{h}_{M_0}$ is $W(D_{n+3})$–equivariant.

6 Cones of curves and divisors in G

Let the setup be as in Section 5. Recall that in Section 4 we considered the cones

$$\mathrm{Nef}(X^n_{\mathrm{Fano}}) \subset \mathrm{Mov}^1(X) \subset \mathrm{Eff}(X) \subset H^2(X, \mathbb{R}),$$

the affine hyperplane $\mathcal{H} \subset H^2(X, \mathbb{R})$ containing all the E_I, and the polytopes given by the intersections of these cones with $\mathcal{H}$,

$$\Delta_{\mathrm{Fano}} \subset \Delta_{\mathrm{Mov}} \subset \Delta \subset \mathcal{H} \cong \mathbb{R}^{n+3}.$$

From the linear inequalities defining these polytopes in $\mathbb{R}^{n+3}$ and the expression (4.5) of the radial projection onto $\mathcal{H}$, one can write explicitly the linear inequalities defining the cones $\mathrm{Nef}(X^n_{\mathrm{Fano}}) \cong \mathrm{Nef}(G)$, $\mathrm{Mov}^1(X) \cong \mathrm{Mov}^1(G)$ and $\mathrm{Eff}(X) \cong \mathrm{Eff}(G)$ with respect to the basis $H, E_1, \dots, E_{n+3}$ of $H^2(X, \mathbb{R})$. Inequalities defining $\mathrm{Mov}^1(X)$ and $\mathrm{Eff}(X)$ were obtained in a different way by Brambilla, Dumitrescu and Postinghel [7]. In this section, we reinterpret the facets and extremal rays of these cones in terms of special divisors and curves in G.

Recall from Section 2 that $\mathcal{E} \subset H^n(Z, \mathbb{R})$ is the cone over the demihypercube Δ with vertices $\{M\}_{M \in \mathcal{F}_m(Z)}$. Its dual cone $\mathcal{E}^{\vee} \subset \mathcal{E}$ has $2(n+3) + 2^{n+2}$ extremal rays, generated by the classes

$$\big\{M + \sigma_i(M) \mid M \in \mathcal{F}_m(Z),\ i \in \{1, \dots, n+3\}\big\} \\ \cup \big\{\delta_M = \big\lfloor \tfrac{1}{2}(m+1) \big\rfloor \eta + (-1)^m M\big\}_{M \in \mathcal{F}_m(Z)}.$$

For a fixed $i \in \{1, \dots, n+3\}$, there are two distinct classes $M + \sigma_i(M)$ as M varies in $\mathcal{F}_m(Z)$, and they form an orbit for the action of W' on $H^n(Z, \mathbb{Z})$. The stabilizer of this orbit is the subgroup $G_i := \{\sigma_I \mid i \notin I \text{ and } |I| \text{ is even}\}$. The group W' acts transitively and freely on the set $\{\delta_M\}_{M \in \mathcal{F}_m(Z)}$. The facet of $\mathcal{E}$ corresponding to each extremal ray of $\mathcal{E}^\vee$ was described in Remark 2.23:

- $(M + \sigma_i(M))^\perp \cap \mathcal{E}$ is the cone over the $(n+2)$–dimensional demihypercube with vertices $\{\sigma_I(M) \mid I \subset \{1, \dots, n+3\} \smallsetminus \{i\},\ |I| \not\equiv m \bmod 2\}$.
- $(\delta_M)^\perp \cap \mathcal{E}$ is a simplicial cone generated by the classes $\sigma_i(M)$, $i \in \{1, \dots, n+3\}$.

Now we turn to cones of curves and divisors in G. We showed in Theorem 5.1 and Proposition 5.5 that

$$\operatorname{Nef}(G) = \beta(\mathcal{E}^\vee) \subset \beta(\mathcal{E}) = \operatorname{Eff}(G),$$
$$\operatorname{Mov}_1(G) = \alpha^{-1}(\mathcal{E}^\vee) \subset \alpha^{-1}(\mathcal{E}) = \operatorname{NE}(G).$$

We give a geometric description of the facets and extremal rays of these cones in terms of special divisors and curves in G.

6.1 (Eff(G)) The cone Eff(G) has 2^{n+2} extremal rays, generated by the classes $\{E_M\}_{M \in \mathcal{F}_m(Z)}$. Each E_M is a fixed prime divisor. The group $W' \subset \operatorname{Aut}(G)$ acts transitively and freely on the set $\{E_M\}_{M \in \mathcal{F}_m(Z)}$. In particular, all these divisors are isomorphic, and they can be described as a small modification of the blow-up of $\mathbb{P}^{n-1}$ at $n+3$ points contained in a rational normal curve (see Section 4.15 for a precise description).

6.2 (the divisor E_M when $n = 4$) Set $n = 4$; in this case E_M is isomorphic to the blow-up of $\mathbb{P}^3$ at 7 points contained in a rational normal curve. To describe geometrically E_M inside G, consider the closed subset

$$\{[L] \in G \mid L \cap M \neq \varnothing\}.$$

Then this locus is not equidimensional, and E_M is its unique divisorial component.

Indeed, let us consider again the incidence diagram

$$\begin{tikzcd} & \mathcal{I} \arrow[dl, "\pi"'] \arrow[dr, "e"] & \\ G & & Z \end{tikzcd}$$

as in 3.5, so that $\dim \mathcal{I} = 5$, π is a $\mathbb{P}^1$–bundle and $\{[L] \in G \mid L \cap M \neq \varnothing\} = \pi(e^{-1}(M))$. For the purposes of this subsection only, it is better to denote by $[M] \in H^4(Z, \mathbb{Z})$ the fundamental class of the plane $M \subset Z$.

It is not difficult to see that e is flat, so that $e^{-1}(M)$ is equidimensional of dimension 3, and $e^*([M]) = [e^{-1}(M)] \in H^4(\mathcal{I}, \mathbb{Z})$. Then $\beta([M]) = \pi_* e^*([M]) = [\pi_*(e^{-1}(M))]$. By Proposition 5.5, we have $E_M = \pi_*(e^{-1}(M))$, so that E_M is the unique divisorial component of $\pi(e^{-1}(M))$.

Now let us consider the planes $M^*, \sigma_1(M)^*, \dots, \sigma_7(M)^* \subset G$ (see (3.1)); they are all contained in $\pi(e^{-1}(M))$.

Let $i \in \{1, \dots, 7\}$. Recall that $\ell_{\sigma_i(M)} \subset \sigma_i(M)^*$ is a line and that $\ell_{\sigma_i(M)} = \alpha(\sigma_i(M))$. By Proposition 3.7, using for instance (2.12), we have

$$E_M \cdot \ell_{\sigma_i(M)} = M \cdot \sigma_i(M) = -1,$$

so that $\sigma_i(M)^* \subset E_M$. On the other hand E_M contains only 7 planes $(M')^*$ (see Section 4.16), therefore M^* cannot be contained in E_M. This shows that M^* is a 2–dimensional irreducible component of $\pi(e^{-1}(M))$.

6.3 $(\mathrm{NE}(G))$ The cone $\mathrm{NE}(G)$ has 2^{n+2} extremal rays, generated by the classes $\{\ell_M\}_{M \in \mathcal{F}_m(Z)}$, on which $W' \subset \mathrm{Aut}(G)$ acts transitively. The contraction of the extremal ray generated by ℓ_M contracts $M^* \cong \mathbb{P}^m$ to a point.

Fix $M \in \mathcal{F}_m(Z)$ and consider the pseudoisomorphism $\xi_M \colon G \dashrightarrow X$ from Theorem 5.7. This fixes an identification of G with X^n_{Fano}, which identifies each divisor $E_{\sigma_I(M)} \subset G$ with the strict transform of the divisor $E_I \subset X$. Let $I \subset \{1, \dots, n+3\}$ be such that $|I| \le m+1$. It follows from the discussion in Section 4.15 that

- If $|I| \not\equiv m \bmod 2$, then $(\sigma_I(M))^* \subset G$ is the strict transform of $J_{I,s} \subset X$, where $s = \frac{1}{2}(m+1-|I|)$.
- If $|I| \equiv m \bmod 2$, then $(\sigma_I(M))^* \subset G$ is the flipped locus of the flip of the strict transform of $J_{I,s} \subset X$, where $s = \frac{1}{2}(m-|I|)$.

In particular, we see that $(M')^* \subset E_M$ if and only if $M' = \sigma_I(M)$ for some $I \subset \{1, \dots, n+3\}$ with $|I| \le m-1$ and $|I| \not\equiv m \bmod 2$.

6.4 $(\mathrm{Nef}(G))$ The cone $\mathrm{Nef}(G)$ has $2^{n+2} + 2(n+3)$ extremal rays, generated by the classes

$$\{D_M = \beta(\delta_M)\}_{M \in \mathcal{F}_m(Z)} \cup \big\{E_M + E_{\sigma_i(M)} \mid M \in \mathcal{F}_m(Z),\ i = 1, \dots, n+3\big\}.$$

For fixed i, the morphisms associated to the extremal rays generated by $E_M + E_{\sigma_i(M)}$ and $E_{\sigma_j(M)} + E_{\sigma_{ij}(M)}$ for $j \neq i$ are the generic $\mathbb{P}^1$–bundles $\varphi_i \colon G \to Y_{\varphi_i}$ and $\psi_i \colon G \to Y_{\psi_i}$ described in Lemma 3.3. The morphism associated to the extremal ray generated by D_M is the composition of the (disjoint) small contractions of $\sigma_i(M)^* \subset G$ to a point for $i = 1, \dots, n+3$.

6.5 $(\mathrm{Mov}_1(G))$ The cone $\mathrm{Mov}_1(G)$ has $2(n+3)+2^{n+2}$ extremal rays, generated by the curve classes

$$\{\ell_M+\ell_{\sigma_i(M)} \mid M\in\mathcal{F}_m(Z),\ i=1,\dots,n+3\}\cup\{d_M \mid M\in\mathcal{F}_m(Z)\},$$

where

$$d_M:=\alpha^{-1}(\delta_M)=\big\lfloor\tfrac{1}{2}(m+1)\big\rfloor\alpha^{-1}(\eta)+(-1)^m\ell_M\in\mathcal{N}_1(G).$$

For a fixed $i\in\{1,\dots,n+3\}$, there are two distinct classes $\ell_M+\ell_{\sigma_i(M)}$ as M varies in $\mathcal{F}_m(Z)$, and they form an orbit for the action of W' on $\mathcal{N}_1(G)$. By Corollary 3.4, these are the classes of the fibers of the generic $\mathbb{P}^1$–bundles $\varphi_i\colon G\to Y_{\varphi_i}$ and $\psi_i\colon G\to Y_{\psi_i}$. Under the identification $G\cong X^n_{\mathrm{Fano}}$ induced by a pseudoisomorphism $G\dashrightarrow X$, these correspond to the generic $\mathbb{P}^1$–bundles $\phi_i,\phi_i'\colon X^n_{\mathrm{Fano}}\to(X_{\mathcal{P}_i})^{n-1}_{\mathrm{Fano}}$ described in Lemma 4.12. In particular, we see that $Y_{\varphi_i}\cong Y_{\psi_i}\cong(X_{\mathcal{P}_i})^{n-1}_{\mathrm{Fano}}$.

As for the class d_M, using Proposition 3.7 and Remark 2.23, one computes

$$-K_G\cdot d_M=\eta\cdot\delta_M=n+1,$$
$$E_{\sigma_i(M)}\cdot d_M=\sigma_i(M)\cdot\delta_M=0\quad\text{for every } i=1,\dots,n+3.$$

Therefore d_M is the class of the strict transform in G of a general line in $\mathbb{P}^n$ under the map $\rho_M\colon G\dashrightarrow\mathbb{P}^n$.

In order to complete the picture, next we describe equations for the movable cone $\mathrm{Mov}^1(G)\subset H^2(G,\mathbb{R})$ and give a geometric description of the extremal rays of the dual cone $\mathrm{Mov}^1(G)^\vee\subset\mathcal{N}_1(G)$. We do this for $n\geq 4$, since when $n=2$ we have $\mathrm{Mov}^1(G)=\mathrm{Nef}(G)$ and $\mathrm{Mov}^1(G)^\vee=\mathrm{NE}(G)$.

6.6 Proposition *Suppose $n\geq 4$. The cone $\mathrm{Mov}^1(G)^\vee\subset\mathcal{N}_1(G)$ has $2^{n+2}+2(n+3)$ extremal rays, generated by the classes*

$$\{e_M \mid M\in\mathcal{F}_m(Z)\}\cup\{\ell_M+\ell_{\sigma_i(M)} \mid M\in\mathcal{F}_m(Z), i=1,\dots,n+3\},$$

where $e_M:=\big\lfloor\tfrac{1}{2}(m)\big\rfloor\alpha^{-1}(\eta)+(-1)^{m-1}\ell_M$.

Proof Recall from Section 4 that the intersection of $\mathrm{Mov}^1(X)$ with the affine hyperplane $\mathcal{H}\subset H^2(X,\mathbb{R})$ is given by

$$\Delta_{\mathrm{Mov}}=\begin{cases}-\frac{1}{2}\leq\alpha_i\leq\frac{1}{2}, & i\in\{1,\dots,n+3\},\\ H_I\geq 2, & |I| \text{ odd}.\end{cases}$$

So $\mathrm{Mov}^1(G)=\beta(\mathcal{M})$, where $\mathcal{M}$ is the cone over Δ_{Mov}, now viewed as a polytope in the hyperplane $\{\gamma\mid\gamma\cdot\eta=1\}\subset H^n(Z,\mathbb{R})$.

Notice that the facet $(H_I = 2) \cap \Delta_{\mathrm{Mov}}$ of Δ_{Mov} is the convex hull of the vertices v_J such that $\#(I \setminus J) + \#(J \setminus I) = 2$. This follows from (2.17). In the same way as done in Section 2 for $\mathcal{E}$, one can use the linear inequalities defining Δ_{Mov} to compute the linear inequalities defining $\mathcal{M}$, or equivalently the generators of the dual cone $\mathcal{M}^\vee$. These are

$$\{M + \sigma_i(M) \mid M \in \mathcal{F}_m(Z),\ i \in \{1, \dots, n+3\}\} \cup \{\eta_M\}_{M \in \mathcal{F}_m(Z)},$$

where $\eta_M = \lfloor \frac{m}{2} \rfloor \eta + (-1)^{m-1} M$ (notice that $e_M = \alpha(\eta_M)$). Indeed, one can check using (2.12) that

$$\eta_M \cdot \sigma_{ij}(M) = 0 \quad \text{for all } i \neq j. \tag{6.7}$$

By the duality properties of α and β, we have $\mathrm{Mov}^1(G)^\vee = \alpha^{-1}(\mathcal{M}^\vee)$, and the result follows. □

6.8 The classes $\ell_M + \ell_{\sigma_i(M)}$ were described in Section 6.5 above. Now we want to describe the classes e_M.

Given $M \in \mathcal{F}_m(Z)$ and $i \in \{1, \dots, n+3\}$, set $M_0 = \sigma_i(M)$, and follow Notation 2.8, so that $M = M_i$. Consider the pseudoisomorphism $\xi_{M_0}\colon G \dashrightarrow X$ from Theorem 5.7, and note that the divisor $E_M \subset G$ is the strict transform of the divisor $E_i \subset X$ under ξ_{M_0}. By (6.7) above, we have that

$$E_{M_j} \cdot e_M = 0 \quad \text{for all } j \neq i.$$

Similarly one computes that $E_M \cdot e_M = -1$. We conclude that e_M is the class of the strict transform under $\xi_{M_0}^{-1}$ of a general line in $E_i \cong \mathbb{P}^{n-1}$.

6.9 Remark Set $c := \alpha^{-1}(\eta) \in \mathcal{N}_1(G)$. We have

$$-K_G \cdot c = 4 \qquad \text{and} \qquad E_M \cdot c = 1 \quad \text{for every } M \in \mathcal{F}_m(Z).$$

The class c is fixed by the action of $W(D_{n+3})$ and sits in the interior of the cone $\mathrm{Mov}_1(G) \subset \mathrm{NE}(G)$. Let $M \in \mathcal{F}_m(Z)$ and consider the rational map $\rho_M\colon G \dashrightarrow \mathbb{P}^n$ from Theorem 5.7. Then c is the class of the strict transform via ρ_M^{-1} of an elliptic curve of degree $n+1$ in $\mathbb{P}^n$ through $p_1, \dots, p_{n+3}$. There is a 4–dimensional family of such curves (see Dolgachev [11]).

6.10 Remark Brambilla, Dumitrescu and Postinghel [7] describe the effective cone $\mathrm{Eff}^1(X) \subset H^2(X, \mathbb{R})$ by 3 sets of linear inequalities (A_n), (B_n) and $(C_{n,t})$. Similarly, the movable cone $\mathrm{Mov}^1(X) \subset H^2(X, \mathbb{R})$ is described by 3 sets of linear inequalities (A_n), (B_n) and $(D_{n,t})$ (see [7, Theorems 5.1 and 5.3]). These are related to the extremal rays of $\mathrm{Mov}_1(G)$ and $\mathrm{Mov}^1(G)^\vee$ described in Section 6.5 and 6.8 as follows. A divisor class $D \in H^2(G,\mathbb{R})$ satisfies the inequalities (A_n) and (B_n) if and only if

$$D \cdot (\ell_M + \ell_{\sigma_i(M)}) \geq 0 \quad \text{for every } M \in \mathcal{F}_m(Z) \text{ and } i = 1, \dots, n+3.$$

It satisfies the inequalities $(C_{n,t})$ if and only if

$$D \cdot d_M \geq 0 \quad \text{for every } M \in \mathcal{F}_m(Z).$$

Finally, it satisfies the inequalities $(D_{n,t})$ if and only if

$$D \cdot e_M \geq 0 \quad \text{for every } M \in \mathcal{F}_m(Z).$$

6.11 (MCD(G)) Consider the subdivision in polytopes of the demihypercube $\Delta \subset \mathcal{H} \subset H^n(Z,\mathbb{R})$ given by the hyperplane arrangement (4.6). By taking the cones over these polytopes and using the isomorphism $\beta\colon H^n(Z,\mathbb{R}) \to H^2(G,\mathbb{R})$, this subdivision yields the fan MCD(G).

Fix $M_0 \in \mathcal{F}_m(Z)$ and consider the orthogonal basis $\varepsilon_1, \dots, \varepsilon_{n+3}$ of $\eta^\perp \subset H^n(Z,\mathbb{R})$ introduced in (2.9) and the affine coordinates $\alpha_1, \dots, \alpha_{n+3}$ in the hyperplane $\mathcal{H} := \{\gamma \mid \gamma \cdot \eta = 1\}$ described in (2.19). The group W' fixes $\mathcal{H}$ and η, thus it acts linearly in the coordinates α_i. More precisely it follows from (2.11) that, if $I \subset \{1, \dots, n+3\}$ has even cardinality, then $\sigma_I(\alpha_1, \dots, \alpha_{n+3}) = (\alpha'_1, \dots, \alpha'_{n+3})$ with

$$\alpha'_i = \begin{cases} \alpha_i & \text{if } i \notin I, \\ -\alpha_i & \text{if } i \in I. \end{cases}$$

The group W' fixes both Δ and Δ_{Mov}, while the 2^{n+2} polytopes $\sigma_I(\Delta_{\mathrm{Nef}})$ are all distinct. The corresponding cones in MCD(G) are $\xi^*_{M_I}(\mathrm{Nef}(X)) = \sigma^*_I(\xi^*_{M_0}(\mathrm{Nef}(X))$.

7 The automorphism group of G

Let the setup be as in Section 5. In this section we describe the automorphism group of the Fano variety G, generalizing the description of the automorphism group of a quartic del Pezzo surface in Section 1.1.

7.1 Proposition *There are inclusion of groups*

$$(\mathbb{Z}/2\mathbb{Z})^{n+2} \cong W' \subseteq \mathrm{Aut}(G) \subseteq W(D_{n+3}) \cong (\mathbb{Z}/2\mathbb{Z})^{n+2} \rtimes S_{n+3}.$$

Moreover, if the points $(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1) \in \mathbb{P}^1$ are general, then $\mathrm{Aut}(G) = W' \cong (\mathbb{Z}/2\mathbb{Z})^{n+2}$.

Notice that in the general case we also have $\mathrm{Aut}(Z) = W'$ (see Reid [25, Lemma 3.1]), so that Z and G have the same automorphism group.

Proof Clearly we have $W' \subseteq \mathrm{Aut}(G)$.

For any $\zeta \in \mathrm{Aut}(G)$, the induced isomorphism $\zeta^*\colon H^2(G,\mathbb{R}) \to H^2(G,\mathbb{R})$ preserves $-K_G$ and $\mathrm{Eff}(G)$. As in the proof of Proposition 5.4, one shows that $\zeta^* \in W(D_{n+3})$. This yields a group homomorphism

$$\mathrm{Aut}(G) \to W(D_{n+3}).$$

Fix $M_0 \in \mathcal{F}_m(Z)$. Consider the stabilizer G_0 of M_0 in $W(D_{n+3})$, and recall that $W(D_{n+3}) = W' \rtimes G_0 \cong (\mathbb{Z}/2\mathbb{Z})^{n+2} \rtimes S_{n+3}$. So, given $\zeta \in \mathrm{Aut}(G)$, there are unique elements $\omega \in G_0$ and $\sigma_I \in W'$ such that $\zeta^* = \omega \circ \sigma_I$. Set $\tilde{\zeta} := \sigma_I \circ \zeta \in \mathrm{Aut}(G)$. Then $\tilde{\zeta}^* = \zeta^* \circ \sigma_I = \omega$, so $\tilde{\zeta}^*$ fixes E_{M_0}, and hence it also fixes H_{M_0}.

Consider the rational map $\rho_{M_0}\colon G \dashrightarrow \mathbb{P}^n$ induced by H_{M_0}, which contracts the divisors $E_{M_1}, \dots, E_{M_{n+3}}$ to the points $p_1, \dots, p_{n+3}$ (see Theorem 5.7). Then $\tilde{\zeta}^*(\rho_{M_0}^*(\mathcal{O}_{\mathbb{P}^n}(1))) = \rho_{M_0}^*(\mathcal{O}_{\mathbb{P}^n}(1)) = H_{M_0}$, so ρ_{M_0} and $\rho_{M_0} \circ \tilde{\zeta}$ differ by a projective transformation $f \in \mathrm{Aut}(\mathbb{P}^n)$ preserving the set of points $\{p_1, \dots, p_{n+3}\}$:

$$\begin{tikzcd} G \arrow[r, "\tilde{\zeta}"] \arrow[d, dashed, "\rho_{M_0}"'] & G \arrow[d, dashed, "\rho_{M_0}"] \\ \mathbb{P}^n \arrow[r, "f"] & \mathbb{P}^n \end{tikzcd}$$

In particular, if the points $p_1, \dots, p_{n+3}$ are general, then $f = \mathrm{Id}_{\mathbb{P}^n}$, and so $\zeta = \sigma_I$.

Suppose that $\zeta^* = \mathrm{Id}_{H^2(G,\mathbb{R})}$. Then $\tilde{\zeta} = \zeta$ and f must fix each p_i. Since $p_1, \dots, p_{n+3}$ are in general linear position, this implies that $f = \mathrm{Id}_{\mathbb{P}^n}$, and hence $\zeta = \tilde{\zeta} = \mathrm{Id}_G$. This shows that the homomorphism $\mathrm{Aut}(G) \to W(D_{n+3})$ is injective, yielding the statement. $\square$

Every automorphism of X is induced by a projective transformation of $\mathbb{P}^n$ preserving the set $\{p_1, \dots, p_{n+3}\}$. This in turns corresponds to a projective transformation of $\mathbb{P}^1$ preserving the set of points $\{(\lambda_1 : 1), \dots, (\lambda_{n+3} : 1)\} \subset \mathbb{P}^1$. In particular, if $\lambda_1, \dots, \lambda_{n+3}$ are general, then $\mathrm{Aut}(X) = \{\mathrm{Id}_X\}$.

For any projective variety Y, we denote by $\mathrm{Bir}^0(Y)$ the group of *pseudoautomorphisms* of Y. These are birational maps $Y \dashrightarrow Y$ which are isomorphisms in codimension one.

Since X and G are pseudoisomorphic, we have $\mathrm{Bir}^0(X) \cong \mathrm{Bir}^0(G)$. On the other hand, since G is a Fano manifold, we have $\mathrm{Bir}^0(G) = \mathrm{Aut}(G)$. Indeed if $\zeta \in \mathrm{Bir}^0(G)$, then $\zeta^*(-K_G) = -K_G$. Since ζ is an isomorphism in codimension one and $-K_G$ is ample, ζ must be regular, and similarly for ζ^{-1}.

7.2 Remark (explicit description of pseudoautomorphisms of X) The action of W' on X by pseudoautomorphisms is described by Dolgachev in [11, Sections 4.4–4.6]. Up to a projective transformation, we may assume that $p_1, \dots, p_{n+1}$ are the coordinate points, $p_{n+2} = (1 : \dots : 1)$ and $p_{n+3} = (a_0 : \dots : a_{n+3})$. Since no $n+1$ of the points lie on a hyperplane, all the a_j are nonzero.

Consider the standard Cremona map centered at $p_1, \dots, p_{n+1}$,

$$s\colon (z_0 : \dots : z_n) \mapsto \left(\frac{1}{z_0} : \dots : \frac{1}{z_n}\right).$$

It is regular at p_{n+2} and p_{n+3}, which map to itself and $(1/a_0 : \dots : 1/a_n)$, respectively. The projective transformation

$$r\colon (z_0 : \dots : z_n) \mapsto (a_0 z_0 : \dots : a_n z_n)$$

fixes $p_1, \dots, p_{n+1}$, maps p_{n+2} to p_{n+3}, and maps $(1/a_0 : \dots : 1/a_n)$ to p_{n+2}. So the composition

$$f_{n+2,n+3} = r \circ s\colon \mathbb{P}^n \dashrightarrow \mathbb{P}^n$$

induces a pseudoautomorphism $\omega_{n+2,n+3}\colon X \dashrightarrow X$.

Similarly, for every $i, j \in \{1, \dots, n+3\}$ with $i < j$, we can define a birational involution $f_{ij}\colon \mathbb{P}^n \dashrightarrow \mathbb{P}^n$, which is not regular only at $\{p_1, \dots, p_{n+3}\} \smallsetminus \{p_i, p_j\}$ and exchanges p_i and p_j. This induces a pseudoautomorphism $\omega_{ij}\colon X \dashrightarrow X$.

One can check that ω_{ij}^* acts on $H^2(X, \mathbb{Z})$ as follows:

$$\omega_{ij}^*(-K_X) = -K_X, \quad \omega_{ij}^*(E_i) = E_j, \quad \omega_{ij}^*(E_j) = E_i$$

$$\omega_{ij}^*(H) = nH - (n-1)\left(\sum_{h=1}^{n+1} E_h - E_i - E_j\right)$$

$$\begin{aligned}\omega_{ij}^*(E_r) &= H - \sum_{h=1}^{n+3} E_h + E_i + E_j + E_r \\ &= \frac{1}{n+1}(-K_X) - \frac{2}{n+1}\sum_{h=1}^{n+3} E_h + E_i + E_j + E_r \quad \text{for } r \neq i, j.\end{aligned}$$

Consider the isomorphism $\tilde{h}_{M_0}\colon H^2(X, \mathbb{R}) \to H^n(Z, \mathbb{R})$ defined in (5.2), and the corresponding action of ω_{ij}^* on $H^n(Z, \mathbb{R})$. We have

$$\omega_{ij}^*(\eta) = \eta \quad \text{and} \quad \omega_{ij}^*(\varepsilon_r) = \begin{cases} -\varepsilon_r & \text{if } r = i, j, \\ \varepsilon_r & \text{if } r \neq i, j. \end{cases}$$

(The latter can be checked using (2.14).) Hence $\omega_{ij}^* = \sigma_{ij}$ and ω_{ij} is the pseudoautomorphism of X induced by $\sigma_{ij} \in W'$. In particular, the pseudoautomorphism of X induced by $\sigma_1 \in W'$ is $\omega_{23}\omega_{45}\cdots\omega_{n+2,n+3}$, and so on.

References

[1] **C Araujo**, **A Massarenti**, *Explicit log Fano structures on blow-ups of projective spaces*, Proc. Lond. Math. Soc. 113 (2016) 445–473 MR

[2] **M Artin**, **D Mumford**, *Some elementary examples of unirational varieties which are not rational*, Proc. London Math. Soc. 25 (1972) 75–95 MR

[3] **S Bauer**, *Parabolic bundles, elliptic surfaces and* $\mathrm{SU}(2)$*–representation spaces of genus zero Fuchsian groups*, Math. Ann. 290 (1991) 509–526 MR

[4] **I Biswas**, **Y I Holla**, **C Kumar**, *On moduli spaces of parabolic vector bundles of rank* 2 *over* $\mathbb{CP}^1$, Michigan Math. J. 59 (2010) 467–479 MR

[5] **C Borcea**, *Deforming varieties of* k*–planes of projective complete intersections*, Pacific J. Math. 143 (1990) 25–36 MR

[6] **C Borcea**, *Homogeneous vector bundles and families of Calabi–Yau threefolds, II*, from "Several complex variables and complex geometry, II" (E Bedford, J P D'Angelo, R E Greene, S G Krantz, editors), Proc. Sympos. Pure Math. 52, Amer. Math. Soc., Providence, RI (1991) 83–91 MR

[7] **M C Brambilla**, **O Dumitrescu**, **E Postinghel**, *On the effective cone of* $\mathbb{P}^n$ *blown-up at* $n+3$ *points*, Exp. Math. 25 (2016) 452–465 MR

[8] **C Casagrande**, *Rank* 2 *quasiparabolic vector bundles on* $\mathbb{P}^1$ *and the variety of linear subspaces contained in two odd-dimensional quadrics*, Math. Z. 280 (2015) 981–988 MR

[9] **A-M Castravet**, **J Tevelev**, *Hilbert's* 14^{th} *problem and Cox rings*, Compos. Math. 142 (2006) 1479–1498 MR

[10] **I V Dolgachev**, *Weyl groups and Cremona transformations*, from "Singularities, I" (P Orlik, editor), Proc. Sympos. Pure Math. 40, Amer. Math. Soc., Providence, RI (1983) 283–294 MR

[11] **I V Dolgachev**, *On certain families of elliptic curves in projective space*, Ann. Mat. Pura Appl. 183 (2004) 317–331 MR

[12] **I V Dolgachev**, *Classical algebraic geometry: a modern view*, Cambridge University Press (2012) MR

[13] **I Dolgachev**, **A Duncan**, *Regular pairs of quadratic forms on odd-dimensional spaces in characteristic* 2, preprint (2015) arXiv

[14] **I Dolgachev**, **D Ortland**, *Point sets in projective spaces and theta functions*, Astérisque 165, Soc. Math. France, Paris (1988) MR

[15] **R M Green**, *Homology representations arising from the half cube*, Adv. Math. 222 (2009) 216–239 MR

[16] **R M Green**, *Combinatorics of minuscule representations*, Cambridge Tracts in Mathematics 199, Cambridge University Press (2013) MR

[17] **J Harris**, *Algebraic geometry: a first course*, Graduate Texts in Mathematics 133, Springer (1992) MR

[18] **Y Hu**, **S Keel**, *Mori dream spaces and GIT*, Michigan Math. J. 48 (2000) 331–348 MR

[19] **J E Humphreys**, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics 9, Springer (1972) MR

[20] **Z Jiang**, *A Noether–Lefschetz theorem for varieties of r–planes in complete intersections*, Nagoya Math. J. 206 (2012) 39–66 MR

[21] **S Mukai**, *Counterexample to Hilbert's fourteenth problem for the 3–dimensional additive group*, preprint, RIMS, Kyoto University (2001) Available at `http://www.kurims.kyoto-u.ac.jp/preprint/file/RIMS1343.pdf`

[22] **S Mukai**, *An introduction to invariants and moduli*, Cambridge Studies in Advanced Mathematics 81, Cambridge University Press (2003) MR

[23] **S Mukai**, *Finite generation of the Nagata invariant rings in A–D–E cases*, preprint, RIMS, Kyoto University (2005) Available at `http://www.kurims.kyoto-u.ac.jp/preprint/file/RIMS1502.pdf`

[24] **S Okawa**, *On images of Mori dream spaces*, Math. Ann. 364 (2016) 1315–1342 MR

[25] **M Reid**, *The complete intersection of two or more quadrics*, PhD thesis, University of Cambridge (1972) Available at `http://homepages.warwick.ac.uk/~masda/3folds/qu.pdf`

[26] **C Soulé**, **C Voisin**, *Torsion cohomology classes and algebraic cycles on complex projective manifolds*, Adv. Math. 198 (2005) 107–127 MR

Instituto de Matemática Pura e Aplicada
22460-320 Rio de Janeiro, Brazil

Dipartimento di Matematica, Università di Torino
I-10123 Torino, Italy

`caraujo@impa.br`, `cinzia.casagrande@unito.it`

Proposed: Lothar Göttsche
Seconded: Dan Abramovich, Gang Tian

Received: 14 February 2016
Revised: 29 July 2016

Geometry & Topology 21 (2017) 3047–3092

Stable homology of surface diffeomorphism groups made discrete

SAM NARIMAN

We answer affirmatively a question posed by Morita on homological stability of surface diffeomorphisms made discrete. In particular, we prove that C^∞–diffeomorphisms of surfaces as family of discrete groups exhibit *homological stability*. We show that the stable homology of C^∞–diffeomorphisms of surfaces as discrete groups is the same as homology of certain infinite loop space related to Haefliger's classifying space of foliations of codimension 2. We use this infinite loop space to obtain new results about (non)triviality of characteristic classes of flat surface bundles and codimension-2 foliations.

58D05, 57R32, 55P35, 55R40, 57R19, 57R32, 57R50; 57R20

0 Statements of the main results

This paper is a continuation of the project initiated in Nariman [33] on the homological stability and the stable homology of discrete surface diffeomorphisms.

0.1 Homological stability for surface diffeomorphisms made discrete

To fix some notations, let $\Sigma_{g,n}$ denote a surface of genus g with n boundary components and let $\mathrm{Diff}^\delta(\Sigma_{g,n},\partial)$ denote the discrete group of orientation-preserving diffeomorphisms of $\Sigma_{g,n}$ that are supported away from the boundary.

The starting point of this paper is a question posed by Morita [32, Problem 12.2] about an analogue of Harer stability for surface diffeomorphisms made discrete. In light of the fact that all known cohomology classes of $\mathrm{BDiff}^\delta(\Sigma_g)$ are stable with respect to g, Morita [32] asked:

Question *Do the homology groups of* $\mathrm{BDiff}^\delta(\Sigma_g)$ *stabilize with respect to* g*?*

In order to prove homological stability for a family of groups, it is more convenient to have a map between them. To define a map inducing homological stability, let $j\colon \Sigma_{g,1}\hookrightarrow \Sigma_{g+1,1}\backslash\partial\Sigma_{g+1,1}$ be an embedding such that the complement of $j(\Sigma_{g,1})$ in $\Sigma_{g+1,1}\backslash\partial\Sigma_{g+1,1}$ is diffeomorphic to the interior of $\Sigma_{1,2}$. By extending diffeomorphisms via the identity, this embedding induces a group homomorphism between

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.3047

diffeomorphism groups, $s\colon \mathrm{Diff}^{\delta}(\Sigma_{g,1},\partial)\to \mathrm{Diff}^{\delta}(\Sigma_{g+1,1},\partial)$. Although the stabilization map s depends on the embedding j, it is not hard to see that different choices of embeddings induce the same map on the group homology (see [33, Theorem 2.5] for more details); hence, by abuse of notation, we denote the induced map between classifying spaces $s\colon \mathrm{BDiff}^{\delta}(\Sigma_{g,1},\partial)\to \mathrm{BDiff}^{\delta}(\Sigma_{g+1,1},\partial)$ by the same letter. Our first theorem affirmatively answers Morita's question.

Theorem 0.1 *The stabilization map*

$$s_*\colon H_k(\mathrm{BDiff}^{\delta}(\Sigma_{g,1},\partial);\mathbb{Z})\to H_k(\mathrm{BDiff}^{\delta}(\Sigma_{g+1,1},\partial);\mathbb{Z})$$

induces an isomorphism as long as $k\le \frac{1}{3}(2g-2)$.

Remark 0.2 Bowden [3] proved stability for $k\le 3$ if $g\ge 8$. Here, we give a proof with the same stability range as that of the mapping class groups.

These homological stability results hold for surface diffeomorphisms with any order of regularity, ie the stabilization map induces a homology isomorphism for C^r–diffeomorphisms of surfaces as $r>0$. However, the remarkable theorem of Tsuboi [45] implies that the classifying space of C^1–diffeomorphisms with discrete topology, $\mathrm{BDiff}^{\delta,1}(\Sigma_{g,n},\partial)$, is homology equivalent to the classifying space of the mapping class group of $\Sigma_{g,n}$. Hence, for regularity $r=1$, the homological stability is already implied by Harer stability [12] for mapping class groups.

It should be further noted that the proof of [33, Theorem 1.1], which is a similar theorem for high-dimensional analogue of surfaces, does not carry over verbatim to prove homological stability of surface diffeomorphisms. In proving homological stability for a family of groups, one key step is to build a highly connected simplicial complex on which the family acts. To prove the highly connectedness of the simplicial complex used in [33, Theorem 1.1], it is essential to work in dimension higher than 5 so that certain surgery arguments work.

However, Randal-Williams [37] proved a homological stability theorem for moduli spaces of surfaces equipped with a "tangential structure". We use Thurston's generalization [43] of Mather's theorem in foliation theory and Randal-Williams's theorem [37, Theorem 7.1] to establish homological stability of $\mathrm{Diff}^{\delta}(\Sigma_{g,1},\partial)$. The advantage of this high-powered approach is that it describes the limiting homology in terms of an infinite loop space related to codimension-2 foliations.

0.2 Stable homology of $\mathrm{Diff}^{\delta}(\Sigma_{g,n},\partial)$

Analogously to [33, Theorem 1.2], we describe the stable homology of $\mathrm{BDiff}^{\delta}(\Sigma_{g,n},\partial)$ in terms of an infinite loop space related to the Haefliger category. Let us recall the definition of the Haefliger classifying space of foliations.

Definition 0.3 The Haefliger category Γ_n^r is a topological category whose objects are points in $\mathbb{R}^n$ with its usual topology and morphisms between two points, say x and y, are germs of C^r diffeomorphisms that send x to y. The space of morphisms is equipped with the sheaf topology (see Section 1.2.1 for more details). If we do not decorate the Haefliger category with r, we usually mean the space of morphisms has the C^∞–regularity. By $\mathrm{S}\Gamma_n^r$, we mean the subcategory of Γ_n^r with the same objects, but the morphisms are germs of orientation-preserving diffeomorphisms (see Haefliger [10] for more details).

The classifying space of the Haefliger category classifies Haefliger structures up to concordance. The normal bundle to the Haefliger structure induces a map

$$\nu\colon \mathrm{BS}\Gamma_n \to \mathrm{BGL}_n^+(\mathbb{R}),$$

where $\mathrm{GL}_n^+(\mathbb{R})$ is the group of real matrices with positive determinants.

Let γ_2 denote the tautological bundle over $\mathrm{GL}_2^+(\mathbb{R})$. Recall that the *Madsen–Tillmann spectrum* $\mathrm{MTSO}(2)$ is the Thom spectrum of the virtual bundle $-\gamma_2$ over $\mathrm{BGL}_2^+(\mathbb{R})$. Let $\mathrm{MT}\nu$ denote the Thom spectrum of the virtual bundle $\nu^*(-\gamma_2)$ over $\mathrm{BS}\Gamma_2$ (see Definition 2.2 for a more detailed description). We denote the basepoint component of the infinite loop space associated to this spectrum by $\Omega_0^\infty \mathrm{MT}\nu$. As we shall explain in Section 2.2, there exists a *parametrized Pontryagin–Thom construction*, which induces a continuous map

$$\alpha\colon \mathrm{BDiff}^\delta(\Sigma_{g,n}, \partial) \to \Omega_0^\infty \mathrm{MT}\nu.$$

Our second theorem is an analogue of the Madsen–Weiss theorem [20] for discrete surface diffeomorphisms.

Theorem 0.4 *The map α induces a homology isomorphism in the stable range of Theorem 0.1.*

For any topological group G, let G^δ denote the same group with the discrete topology. The identity map defines a continuous homomorphism $G^\delta \to G$. Thus, for the topological group $\mathrm{Diff}(\Sigma_{g,n}, \partial)$ with C^∞ topology, the identity induces a map

$$\iota\colon \mathrm{BDiff}^\delta(\Sigma_{g,n}, \partial) \to \mathrm{BDiff}(\Sigma_{g,n}, \partial).$$

To study the effect of this map on cohomology in the stable range, we study the following natural map between infinite loop spaces that is induced by ν:

$$\Omega^\infty \boldsymbol{\nu}\colon \Omega_0^\infty \mathrm{MT}\nu \to \Omega_0^\infty \mathrm{MTSO}(2).$$

Remark 0.5 In foliation theory, Mather and Thurston studied the homotopy fiber of ι (see Theorem 1.3), and they already translated understanding the homotopy fiber of ι to a homotopy-theoretic question about $\mathrm{B}\mathrm{S}\Gamma_2$. However, we pursue a different approach by replacing $\mathrm{BDiff}^{\delta}(\Sigma_{g,n},\partial)$ and $\mathrm{BDiff}(\Sigma_{g,n},\partial)$ with appropriate infinite loop spaces to study ι in the stable range. As we shall see in Sections 0.3 and 0.4, this approach is much more amenable to actual calculations in the stable range instead of understanding the homotopy fiber of ι.

Theorem 0.6 *The map $\Omega^\infty \boldsymbol{\nu}$ induces an injection on $\mathbb{F}_p$–cohomology, ie*

$$H^k(\Omega^\infty_0\mathrm{MTSO}(2);\mathbb{F}_p)\hookrightarrow H^k(\Omega^\infty_0\mathrm{MT}\nu;\mathbb{F}_p)$$

for any k.

Corollary 0.7 *The map ι induces an injection*

$$H^k(\mathrm{BDiff}(\Sigma_{g,n},\partial);\mathbb{F}_p)\hookrightarrow H^k(\mathrm{BDiff}^{\delta}(\Sigma_{g,n},\partial);\mathbb{F}_p)$$

as long as $k\leq \frac{1}{3}(2g-2)$.

Remark 0.8 Theorems 0.1, 0.4 and 0.6 in fact hold for C^r–diffeomorphisms for any $r>0$. For applications in Section 0.3 and 0.4, we formulated the theorems for C^∞–diffeomorphisms but in fact all of them hold for C^r–diffeomorphisms while $r>1$ and $r\neq 3$. As we mentioned earlier, the case of $\mathrm{Diff}^{\delta,1}(\Sigma_{g,n},\partial)$ is an exception that thanks to Tsuboi's theorem [45] this group has the same homology of the mapping class group of $\Sigma_{g,n}$. The reason we also exclude $r=3$ is that for most of the applications, we need the perfectness of the identity component $\mathrm{Diff}_0^{\delta,r}(\Sigma_{g,n},\partial)$, which in the smooth case is a consequence of Thurston's work [43] and for $r\neq 3$ is a consequence of Mather's work [21].

0.3 Applications to characteristic classes of flat surface bundles

The theory of characteristic classes of fiber bundles and foliated fiber bundles (ie fiber bundle with a foliation transverse to the fibers) whose fibers are diffeomorphic to a C^∞–manifold M is equivalent to understanding the cohomology groups $H^*(\mathrm{BDiff}(M))$ and $H^*(\mathrm{BDiff}^{\delta}(M))$, respectively. The theory of characteristic classes of manifold bundles and surface bundles in particular have been studied extensively (see Galatius and Randal-Williams [8] and Morita [30]). Therefore, we have some understanding of $H^*(\mathrm{BDiff}(M))$ for certain classes of manifolds. For foliated (flat) manifold bundles, however, there seems to be very little known about the existence of nontrivial characteristic classes in $H^*(\mathrm{BDiff}^{\delta}(M))$.

The abstract results in Section 0.2 shed new light on the (non)triviality of characteristic classes of flat surface bundles. They also provide a unified approach to previous results of Kotschick and Morita [17] and Bowden [3]. We pursue the study of $H^*(\mathrm{BSymp}^\delta(\Sigma_g))$ in light of these theorems elsewhere.

Morita [30] showed that finite index subgroups of mapping class group of surfaces cannot be realized as subgroups of diffeomorphisms by showing that the Mumford–Miller–Morita classes $\kappa_i \in H^{2i}(\mathrm{BDiff}(\Sigma_g);\mathbb{Q})$ for $i > 2$ get sent to zero via the induced map

$$H^*(\mathrm{BDiff}(\Sigma_g);\mathbb{Q}) \to H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q}).$$

Unlike the cohomology with rational coefficients, Corollary 0.7 implies that all monomials of the κ_i are nontorsion classes in $H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$.

Theorem 0.9 *There is an injection*

$$\mathbb{Z}[\kappa_1,\kappa_2,\dots] \hookrightarrow H^k(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$$

as long as $k \le \frac{1}{3}(2g-2)$.

Remark 0.10 Not all nontorsion classes in $H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ can be realized by an element in $\mathrm{Hom}(H_*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z});\mathbb{Z})$. As $\mathrm{BDiff}^\delta(\Sigma_g)$ is not a finite type space, the universal coefficient theorem implies that the Ext term in $H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ might have nontorsion classes too. In particular, nontriviality of κ_i in $H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ does not imply that there exists a flat surface bundle whose κ_i is nonzero. But in fact one can use the method of Akita, Kawazumi and Uemura [1] to prove such flat surface bundles exist.

Corollary 0.11 *The group* $H_{2k-1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ *is not finitely generated as long as* $k > 2$ *and* $k \le \frac{2}{3}g$.

Kotschick and Morita [16] constructed a flat surface bundle over a surface whose signature is nonzero. Hence, they conclude that κ_1, which is 3 times the signature of the total space, is nonzero in $H^2(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$. We can use Theorem 0.4 to give a homotopy-theoretic proof of their result in the stable range.

Theorem 0.12 (Morita and Kotschick) *The image of* κ_1^n *in* $H^{2n}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ *is nonzero for all positive integer* n *provided that* $g \ge 3n$.

To summarize the (non)vanishing results of MMM classes for flat surface bundles, recall that the Bott vanishing theorem implies that κ_i vanishes in $H^{2i}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ for $i > 2$ and Theorem 0.12 implies that the first MMM class κ_1 does not vanish in $H^2(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$. Moreover, Theorem 0.9 implies that all κ_i are nontorsion classes in $H^{2i}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$. By Theorem 0.9 we know that κ_2 is nonzero in $H^4(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$; however, we still do not know the answer to the following problem, posed by Kotschick and Morita [16]:

Problem *Determine whether the second MMM class κ_2 is nontrivial in the space $H^4(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$.*

We prove that this problem is equivalent to an open problem in foliation theory related to the cube of the Euler class of the normal bundle of codimension-2 foliations; see Hurder [14, Problem 15.4].

Theorem 0.13 *The MMM class κ_2 in $H^4(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ is nonzero for $g>6$ if and only if a C^2–foliation $\mathcal{F}$ of codimension 2 on a 6–manifold exists such that $e(\nu(\mathcal{F}))^3\neq 0$, where $\nu(\mathcal{F})$ is the normal bundle of the foliation $\mathcal{F}$ and $e(\nu(\mathcal{F}))$ is its Euler class.*

Remark 0.14 Using the universal coefficient theorem, Thom's result on representing cycles by manifolds and Thurston's h–principle for foliations of codimension greater than 2, one can show that proving the existence of a codimension-2 foliation $\mathcal{F}$ with $e(\nu(\mathcal{F}))^3\neq 0$ is in fact equivalent to proving $\nu^*(e^3)\in H^6(\mathrm{B}S\Gamma_2;\mathbb{Q})$ does not vanish, where $e\in H^2(\mathrm{BGL}_2^+(\mathbb{R});\mathbb{Q})$ is the universal Euler class. We then show that nonvanishing of κ_2 and $\nu^*(e^3)$ are equivalent.

Theorem 0.4 and well-known results about the continuous variation of foliations of codimension 2 can be used to construct more nontrivial classes on flat surface bundles. For example, Rasmussen [38] showed that the two Godbillon–Vey classes $h_1.c_2$ and $h_1.c_1^2$ in $H^5(\mathrm{B}S\Gamma_2;\mathbb{R})$ (see Bott [2, Section 10] for the definition of Godbillon–Vey classes) continuously vary for families of foliations of codimension 2, ie the map

$$(h_1c_1^2,h_1c_2)\colon H_5(\mathrm{B}S\Gamma_2;\mathbb{Q})\twoheadrightarrow\mathbb{R}^2$$

induced by the evaluation of $h_1.c_2$ and $h_1.c_1^2$ is surjective. We use this theorem of Rasmussen to simplify the proof of Bowden's theorem [3], which says, for all g, the fiber integration of the two Godbillon–Vey classes $h_1.c_2$ and $h_1.c_1^2$ induce a surjective homomorphism

$$H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})\twoheadrightarrow\mathbb{R}^2. \tag{0-15}$$

In the stable range though, we use Theorem 0.4 to prove a stronger result, that essentially all secondary classes in $H^3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R})$ come from secondary classes of flat disk bundles; more precisely:

Theorem 0.16 *Let $\mathbb{R}^2 \hookrightarrow \Sigma_g$ be an embedding of an open disk in the surface Σ_g. For $g \geq 6$, the induced map*

$$H_3(\mathrm{BDiff}_c^\delta(\mathbb{R}^2); \mathbb{Q}) \longrightarrow H_3(\mathrm{BDiff}^\delta(\Sigma_g); \mathbb{Q})$$

is surjective.

Using the surjectivity of (0-15) and the Hopf algebra structure on $H_*(\Omega_0^\infty \mathrm{MT}\nu; \mathbb{Q})$, we construct *discontinuous* classes (see Morita [29] for applications of discontinuous invariants) in $H_{3k}(\mathrm{BDiff}^\delta(\Sigma_g); \mathbb{Q})$.

Theorem 0.17 *There exists a surjective map*

$$H_{3k}(\mathrm{BDiff}^\delta(\Sigma_g); \mathbb{Q}) \twoheadrightarrow \textstyle\bigwedge_{\mathbb{Q}}^k \mathbb{R}^2$$

provided $k \leq \frac{1}{9}(2g-2)$, where $\bigwedge_{\mathbb{Q}}^k \mathbb{R}^2$ is the k^{th} exterior power of $\mathbb{R}^2$ as a vector space over $\mathbb{Q}$.

0.4 Applications to the foliated cobordism of codimension 2

Let $\mathcal{F}\Omega_{n,k}$ be the cobordism group of n–manifolds with a foliation of codimension k and $\mathrm{MSO}_n(X)$ be the oriented cobordism group of n–manifolds equipped with a map to X. Using Theorem 0.4, we compare codimension-2 foliations with foliated surface bundles. By the Atiyah–Hirzebruch spectral sequence, the homology equivalence in Theorem 0.4 implies the equivalence in bordism theory

$$\mathrm{MSO}_n(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathrm{MSO}_n(\Omega_0^\infty \mathrm{MT}\nu),$$

which is an isomorphism in the stable range. Let e_n be the map that associates to every flat surface bundle the foliated cobordism class of the codimension-2 foliation on the total space of the surface bundle

$$e_n\colon \mathrm{MSO}_n(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathcal{F}\Omega_{n+2,2}.$$

Using Theorem 0.4, we will determine the image of e_2 and e_3 up to torsions. More precisely:

Theorem 0.18 *For $g \geq 4$, the map*

$$e_3\colon \mathrm{MSO}_3(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathcal{F}\Omega_{5,2},$$

is rationally surjective and, for $g \geq 6$, it is rationally an isomorphism.

Remark 0.19 To geometrically interpret the theorem, let $\mathcal{F}$ be any codimension-2 foliation on a manifold of dimension 5 and let $k\mathcal{F}$ denote a disjoint union of k copies of $\mathcal{F}$. Then $k\mathcal{F}$, for some integer k, is foliated cobordant to a flat surface bundle of genus at most 4. And the injection for $g \geq 6$ means that if $\mathcal{F}$ is a flat surface bundle over a 3–manifold which bounds a codimension-2 foliation on a 6–manifold, then $k\mathcal{F}$ for some integer k bounds a flat surface bundle over a 4–manifold where the genus of the fibers is at least 6.

Remark 0.20 This theorem might be compared to the result of Mizutani, Morita and Tsuboi [28], in which they proved that any codimension-one foliation almost without holonomy is homologous to a disjoint union of flat circle bundles over tori.

Theorem 0.21 *Let ϕ be the map*

$$\phi\colon \mathcal{F}\Omega_{4,2} \otimes \mathbb{Q} \to \mathbb{Q}$$

that sends a foliation $\mathcal{F}$ on a 4–manifold M to the difference of the Pontryagin numbers $\int_M p_1(M) - p_1(\nu(\mathcal{F}))$. For $g \geq 3$, we have the short exact sequence

$$0 \to \mathrm{MSO}_2(\mathrm{BDiff}^\delta(\Sigma_g)) \otimes \mathbb{Q} \xrightarrow{e_2} \mathcal{F}\Omega_{4,2} \otimes \mathbb{Q} \xrightarrow{\phi} \mathbb{Q} \to 0.$$

Remark 0.22 Roughly speaking, up to torsion the only obstruction for a codimension-2 foliation $\mathcal{F}$ on a 4–manifold M to be foliated cobordant to a flat surface bundle is $\int_M p_1(M) - p_1(\nu(\mathcal{F}))$.

We also prove that in low dimensions, we can change surface bundles up to cobordism to obtain a flat surface bundle; more precisely we prove:

Theorem 0.23 *For $g > 5$, every surface bundle of genus g over a 3–manifold is cobordant to a flat surface bundle.*

Remark 0.24 Using the perfectness of the identity component of C^∞–diffeomorphisms, Kotschick and Morita [16] proved that every surface bundle over a surface is foliated cobordant to a flat surface bundle.

0.5 Outline

This paper is organized as follows: In Section 1, we obtain a short proof of the homological stability of discrete surface diffeomorphisms using a deep theorem of Mather and Thurston and a version of twisted stability of mapping class group due to Randal-Williams. In Section 2, we derive a Madsen–Weiss-type theorem for discrete surface diffeomorphisms. In Section 3, we explore the consequences of having a Madsen–Weiss-type theorem for discrete surface diffeomorphisms in flat surface bundles and their characteristic classes.

Acknowledgments It is my pleasure to thank my thesis advisor, Søren Galatius, for his patience, encouragement and technical support during this project. Without his help and support, this paper would not have existed. I also would like to thank Oscar Randal-Williams for many sage remarks. I would like to thank Alexander Kupers for reading the first draft of this paper and Jonathan Bowden for pointing out to me an obstruction for a codimension-2 foliation on a 4–manifold to be foliated cobordant to a flat surface bundle. I am also very grateful to the referee, whose many helpful comments improve the exposition of the paper. This work was partially supported by NSF grants DMS-1105058 and DMS-1405001.

1 Homological stability from foliation theory

Our goal in this section is to show that the homological stability of discrete surface diffeomorphisms is implied by a "twisted" homological stability of mapping class groups developed in [37]. We use foliation theory to show that $\mathrm{BDiff}^{\delta}(\Sigma_{g,k},\partial)$ is homology equivalent to a moduli space of a certain tangential structure in the sense of [37, Definition 1.1].

1.1 Stabilization maps

In the introduction, we formulated the homological stability for discrete diffeomorphisms of $\Sigma_{g,1}$. Let us describe the stabilization maps for surfaces with any positive number of boundary components.

For a surface Σ with boundary, we shall write $\mathrm{Diff}^{\delta}(\Sigma,\partial)$ to denote the discrete group of compactly supported orientation-preserving diffeomorphisms of $\Sigma\backslash\partial\Sigma$ and if Σ is a closed compact surface, $\mathrm{Diff}^{\delta}(\Sigma,\partial)$ will just mean all orientation-preserving diffeomorphisms of Σ equipped with discrete topology. Let $\Sigma\hookrightarrow\Sigma'$ be a subsurface in a collared surface Σ' (ie with a choice of a collar neighborhood of the boundary) such that each of the boundary components of the subsurface either coincides with one of the boundary components of the bigger surface or entirely lies in its interior. If we extend diffeomorphisms of Σ via the identity over the cobordism $K=\Sigma'\backslash\overline{\Sigma}$, we obtain a map

$$t\colon \mathrm{Diff}^{\delta}(\Sigma,\partial)\to\mathrm{Diff}^{\delta}(\Sigma',\partial).$$

Let $\Sigma_{g,k}$ be a fixed model for an orientable surface of genus g and k boundary components with a chosen collar neighborhood of the boundary. If we choose a diffeomorphism f from Σ to $\Sigma_{g,k}$ and a diffeomorphism h from Σ' to $\Sigma_{g',k'}$, we obtain a stabilization map

$$s_{f,h}(t)\colon \mathrm{Diff}^{\delta}(\Sigma_{g,k},\partial)\to\mathrm{Diff}^{\delta}(\Sigma_{g',k'},\partial).$$

Different choices of f and h induce different stabilization maps, but one can show, similar to [33, Theorem 2.5], that if for some choices of f and h the map $s_{f,h}(t)$ induces a homology isomorphism in some homological degrees, then the stabilization map induces a homology isomorphism in the same homological degrees for all choices of f and h. Therefore, we shall not write the dependence of the stabilization maps on choices of f and h.

1.2 Homological stability of the moduli space of Γ_2–structures

Because we are interested in homological stability of $\mathrm{Diff}^\delta(\Sigma_{g,k},\partial)$, we may replace $\mathrm{BDiff}^\delta(\Sigma_{g,k},\partial)$ by a homology equivalent space which is more convenient from the point of view of homotopy theory. To do so, we recall what we need from foliation theory.

1.2.1 Mather–Thurston theory Recall from Definition 0.3 that $\mathrm{S}\Gamma_n$ is the groupoid of germs of orientation-preserving diffeomorphisms of $\mathbb{R}^n$. Let $\mathrm{Mor}_{\mathrm{S}\Gamma_n}$ denote the space of morphisms in the topological groupoid $\mathrm{S}\Gamma_n$. To recall the topology on the space of morphisms, let $g\in\mathrm{Mor}_{\mathrm{S}\Gamma_n}$ be a germ sending x to y. One can represent g as a local diffeomorphism $\tilde{g}\colon U\to V$, where U and V are open sets containing x and y, respectively. The set of germs of $\tilde{g}$ at all points in U gives an open neighborhood of the germ g.

Definition 1.1 Let X be a topological space. A 1–cocycle on X with values in $\mathrm{S}\Gamma_n$ consists of an open cover $\{U_\alpha\}_I$ of X, and for any two indices α and β in I, a continuous map $\gamma_{\alpha\beta}\colon U_\alpha\cap U_\beta\to\mathrm{Mor}_{\mathrm{S}\Gamma_n}$ satisfying the cocycle condition, for any α, β and δ,

$$\gamma_{\alpha\beta}\gamma_{\beta\delta}=\gamma_{\alpha\delta}\quad\text{on } U_\alpha\cap U_\beta\cap U_\delta.$$

In particular, the left-hand side is defined, ie the source of the map $\gamma_{\alpha\beta}$ is the same as the target of $\gamma_{\beta\delta}$.

Two cocycles $c=\{U_\alpha,\gamma_{\alpha\beta}\}_I$ and $c'=\{U_{\alpha'},\gamma_{\alpha'\beta'}\}_J$ are said to be equivalent if there exists a cocycle $c''=\{U_{\alpha''},\gamma_{\alpha''\beta''}\}_K$ such that $K=I\cup J$ and c'' restricts to c on $\{U_\alpha\}_I$ and to c' on $\{U_{\alpha'}\}_J$.

Definition 1.2 An $\mathrm{S}\Gamma_n$–structure on X is an equivalence class of 1–cocycles with values in $\mathrm{Mor}_{\mathrm{S}\Gamma_n}$ on X.

Note that a cooriented foliation of codimension n can be specified by a covering of X by open sets U_α, together with a submersion f_α from each open set U_α to $\mathbb{R}^n$, such that for each α and β there is a map $g_{\alpha\beta}$ from $U_\alpha\cap U_\beta$ to local diffeomorphisms satisfying

$$f_\alpha(v) = g_{\alpha\beta}(u)(f_\beta(v))$$

whenever v is close enough to $u \in U_\alpha \cap U_\beta$. Then the covering U_α and the germs of $g_{\alpha\beta}$ defines a $\mathrm{S}\Gamma_n$–structure on X. An advantage of Haefliger structures over foliations is that they are closed under pullbacks.

Two $\mathrm{S}\Gamma_n$–structures c_0 and c_1 on X are concordant if there exists a $\mathrm{S}\Gamma_n$–structure on $X \times [0,1]$ such that the restriction of c to $X \times \{i\}$ is c_i for $i = 0, 1$. Homotopy classes of maps to the classifying space of the groupoid $\mathrm{S}\Gamma_n$ classify $\mathrm{S}\Gamma_n$–structures on X up to concordance (for further details consult [10]).

One can associate a foliated space to every $\mathrm{S}\Gamma_n$–structure $c = \{U_\alpha, \gamma_{\alpha\beta}\}_I$ as follows. Note that for all α, the cocycle condition implies that $\gamma_{\alpha\alpha}$ is the germ of the identity at some point in $\mathbb{R}^n$, hence $\gamma_{\alpha\alpha}$ induces a map from U_α to $\mathbb{R}^n$. Consider the space

$$\coprod_\alpha U_\alpha \times \mathbb{R}^n/\sim,$$

where the identification is given by $(x \in U_\alpha, y_\alpha) \sim (x \in U_\beta, y_\beta)$ if $y_\alpha = \gamma_{\alpha\beta} y_\beta$. We now consider $(x, \gamma_{\alpha\alpha}(x)) \in U_\alpha \times \mathbb{R}^n$, the graph of $\gamma_{\alpha\alpha}$. Let E be the space obtained by the union of the neighborhoods of these graphs by the identification. The space E is germinally well-defined and the horizontal foliation on $U_\alpha \times \mathbb{R}^n$ induces a foliation on E. Hence, we obtain the data of a microbundle $X \xrightarrow{s} E \xrightarrow{p} X$, where s is the section given by the graphs and p is the projection to the first factor. The foliation on E is transverse to the fibers and its pullback to X via the section s is the Haefliger structure c. We call this microbundle the foliated microbundle associated to c (see [46, Section 4] for more details on foliated microbundles).

To state the Mather–Thurston theorem, we let $\overline{\mathrm{BS}\Gamma_n}$ be the homotopy fiber of the natural map

$$\nu\colon \mathrm{BS}\Gamma_n \to \mathrm{BGL}_n^+(\mathbb{R}),$$

which is induced by the map between groupoids that sends every germ to its derivative. By replacing spaces with homotopy equivalent spaces, we may assume that ν is a Serre fibration. For any orientable manifold M, we let $\overline{\mathrm{BDiff}(M,\partial)}$ be the homotopy fiber of the map

$$\mathrm{BDiff}^\delta(M,\partial) \to \mathrm{BDiff}(M,\partial).$$

Let $\tau_M\colon M \to \mathrm{BGL}_n^+(\mathbb{R})$ be the map that classifies the tangent bundle of M. Thus, we obtain a Serre fibration $\tau_M^*(\nu)$ with the base M. The manifold structure on M or the foliation by points on M induces a homotopy class of maps $M \to \mathrm{BS}\Gamma_n$. Let s_0 be one such map that is induced by the point foliation. Let also $\mathrm{Sect}_c(\tau_M^*(\nu))$ denote the compactly supported sections which differ only on a compact set from s_0.

Theorem 1.3 (Mather and Thurston [22]) *There exists a map*

$$f_M\colon \overline{\mathrm{BDiff}(M,\partial)} \to \mathrm{Sect}_c(\tau_M^*(\nu))$$

which induces an isomorphism in homology with integer coefficients.

Remark 1.4 Roughly, the map f_M is given by thinking of a diffeomorphism as a collection of germs at each point of M. Since the elements of $\overline{\mathrm{BDiff}(M,\partial)}$ can be thought of as integrable sections of the fiber bundle $\tau_M^*(\nu)$, this theorem is very similar in spirit to Gromov h–principle-type theorems.

Remark 1.5 Haefliger [10] proved that $\overline{\mathrm{B}\mathrm{S}\Gamma_n}$ is n–connected and he conjectured that it is $2n$–connected. Thurston could improve the connectivity of $\overline{\mathrm{B}\mathrm{S}\Gamma_n}$ by one. To explain his idea, note that Theorem 1.3 for an n–disk D^n implies that the space $\overline{\mathrm{BDiff}(D^n,\partial)}$ is homology isomorphic to the n–fold loop space $\Omega^n\overline{\mathrm{B}\mathrm{S}\Gamma_n}$. Given that Thurston [43] also proved that the identity component of diffeomorphisms of manifolds is a perfect group (in fact he showed it is even simple), one can deduce that $H_1(\overline{\mathrm{BDiff}(D^n,\partial)};\mathbb{Z})=0$. Therefore, using Theorem 1.3 and the Hurewicz theorem, we obtain

$$H_1(\Omega^n\overline{\mathrm{B}\mathrm{S}\Gamma}_n;\mathbb{Z}) = H_{n+1}(\overline{\mathrm{B}\mathrm{S}\Gamma}_n;\mathbb{Z}) = \pi_{n+1}(\overline{\mathrm{B}\mathrm{S}\Gamma}_n) = 0.$$

As we shall see, the topological group $\mathrm{Diff}(M,\partial)$ acts on suitable models for

$$\overline{\mathrm{BDiff}(M,\partial)} \quad \text{and} \quad \mathrm{Sect}_c(\tau_M^*(\nu)).$$

Our goal is to show that the homotopy quotients of these actions are also homology equivalent. In order to achieve this goal, it is convenient to work with simplicial sets instead of topological spaces, and we will explain how to define a map of simplicial sets modeling f_M, which is equivariant for an action of a simplicial group modeling Diff(M) (see [33, Section 5.1] for a different model of the map f_M which is equivariant). Henceforth, we substitute spaces with their singular simplicial complex.

1.2.2 Construction of the map f_M

Since $\overline{\mathrm{BDiff}(M,\partial)}$ and $\mathrm{Sect}_c(\tau_M^*(\nu))$ classify certain geometric structures, it is more convenient to describe their singular simplicial complex geometrically. To do so, we need to recall a few notions from [40; 22].

Definition 1.6 We say a $\mathrm{S}\Gamma_n$–structure c on the total space of the fiber bundle $E\to B$ is *transverse* to the fibers if its restriction to the fibers is a foliation. If the fiber bundle $E\to B$ is a smooth bundle, this is equivalent to the condition that c is a smooth foliation and its leaves are transverse to the fibers.

Definition 1.7 Let M be a smooth n–manifold, X a topological space and c a $\mathrm{S}\Gamma_n$ on $X \times M$. We say c is horizontal if c is the inverse image of the differentiable structure of M via the projection $X \times M \to M$. If $t \in X$ and $x \in M$, we will say c is locally horizontal at (t, x) if there is an open neighborhood $U \times N$ of (t, x) in $X \times M$ such that $c|_{U \times N}$ is horizontal. The support of c, denoted by $\mathrm{supp}(c)$, will mean the closure in M of the set of $x \in M$ for which there is at least one $t \in X$ such that c is not locally horizontal at (t, x).

Since $\overline{\mathrm{BDiff}(M, \partial)}$ is the homotopy fiber of the map

$$\mathrm{BDiff}^{\delta}(M, \partial) \to \mathrm{BDiff}(M, \partial),$$

the p–simplices of the singular simplicial complex $S_{\bullet}(\overline{\mathrm{BDiff}(M, \partial)})$ are uniquely given by $\mathrm{S}\Gamma_n$–structures on $\Delta^p \times M$ transverse to the fiber of the projection $\Delta^p \times M \to \Delta^p$ and have support in the interior of M.

The p–simplices of the simplicial group $S_{\bullet}(\mathrm{Diff}(M, \partial))$, namely the singular complex of $\mathrm{Diff}(M, \partial)$, can be described as the commutative diagrams

$$\begin{tikzcd} \Delta^p \times M \arrow[rr, "\phi"] \arrow[dr, "\mathrm{pr}_1"'] & & \Delta^p \times M \arrow[dl, "\mathrm{pr}_1"] \\ & \Delta^p & \end{tikzcd}$$

where ϕ is a diffeomorphism which is the identity on $\Delta^p \times U$, where U is a neighborhood of the boundary ∂M. We can pull back $\mathrm{S}\Gamma_n$–structures on $\Delta^p \times M$ via ϕ. Hence, we have an action of $S_{\bullet}(\mathrm{Diff}(M, \partial))$ on $S_{\bullet}(\overline{\mathrm{BDiff}(M, \partial)})$. Using the theorem of Milnor [25], we know that $|S_{\bullet}(\overline{\mathrm{BDiff}(M, \partial)})|$ is a model for $\overline{\mathrm{BDiff}(M, \partial)}$, hence we obtain an action of the group $|S_{\bullet}(\mathrm{Diff}(M, \partial))|$ which is weakly equivalent to $\mathrm{Diff}(M, \partial)$ on $\overline{\mathrm{BDiff}(M, \partial)}$. Therefore, the homotopy quotient[1]

$$|S_{\bullet}(\overline{\mathrm{BDiff}(M, \partial)})| /\!/ |S_{\bullet}(\mathrm{Diff}(M, \partial))| \tag{1-8}$$

is weakly equivalent to $\mathrm{BDiff}^{\delta}(M, \partial)$.

To describe the simplicial set $S_{\bullet}(\mathrm{Sect}_c(\tau_M^*(\nu)))$ geometrically, we consider the tangent bundle as the tangent microbundle (see [26] for the definition of microbundles). Recall that the tangent microbundle of the manifold M is the data

$$M \xrightarrow{\Delta} M \times M \xrightarrow{\mathrm{pr}_1} M,$$

[1] For a topological group G acting on a topological space X, the homotopy quotient is denoted by $X /\!/ G$ and is given by $X \times_G \mathrm{E}G$, where $\mathrm{E}G$ is a contractible space on which G acts freely.

which we denote by t_M. Milnor [26, Theorem 2.2] showed that the underlying microbundle of the tangent bundle TM is isomorphic to t_M. For an element $f \in \mathrm{Diff}(M, \partial)$ we have an action of f on TM so that it acts by f on M and by the differential df on the fiber of TM; the corresponding action on t_M acts by f on the base M and by $f \times f$ on $M \times M$.

Recall that every section in $\mathrm{Sect}_c(\tau_M^*(\nu))$ is a lift of the tangent bundle in

$$\begin{tikzcd} & \mathrm{BS}\Gamma_n \arrow[d, "\nu"] \\ M \arrow[ur, dotted, "s"] \arrow[r, "\tau"] & \mathrm{BGL}_n^+(\mathbb{R}) \end{tikzcd}$$

to $\mathrm{BS}\Gamma_n$, which gives a map $s\colon M \to \mathrm{BS}\Gamma_n$ and an isomorphism between t_M and the underlying microbundle of $s^* \circ \nu^*(\gamma_n)$, where γ_n is the tautological bundle on $\mathrm{BGL}_n^+(\mathbb{R})$. This means that the graph of the $\mathrm{S}\Gamma_n$–structure induced by s is a foliated microbundle in the neighborhood of the diagonal $\Delta(M) \subset M \times M$ that is transverse to the fibers of $\mathrm{pr}_1\colon M \times M \to M$.

Definition 1.9 A germ of $\mathrm{S}\Gamma_n$–structure c on $\Delta^p \times M \times M$ at $\Delta^p \times \mathrm{diag}\, M$ which is transverse to the fiber of the projection $\mathrm{id} \times \mathrm{pr}_1\colon \Delta^p \times M \times M \to \Delta^p \times M$ is said to be horizontal at $x \in M$ if there exists a neighborhood U around x such that the restriction of the $\mathrm{S}\Gamma_n$–structure to $\Delta^p \times U \times U$ is induced by the projection $\Delta^p \times U \times U \to U$ on the last factor. By the support of c, we mean the set of $x \in M$ where c is not horizontal. Note that $\mathrm{supp}(c)$ is a closed subset.

Hence, p–simplices in the simplicial set $S_\bullet(\mathrm{Sect}_c(\tau_M^*(\nu)))$ can be described as the germ of $\mathrm{S}\Gamma_n$–structures on $\Delta^p \times M \times M$ at $\Delta^p \times \mathrm{diag}\, M$ which are transverse to the fiber of the projection $\mathrm{id} \times \mathrm{pr}_1\colon \Delta^p \times M \times M \to \Delta^p \times M$ and have compact support. This gives a model for the compactly supported sections $\mathrm{Sect}_c(\tau_M^*(\nu))$ [22, Section 16]. Similar to the previous case, there is an obvious action of $S_\bullet(\mathrm{Diff}(M, \partial))$ on $S_\bullet(\mathrm{Sect}_c(\tau_M^*(\nu)))$.

Construction 1.10 Let $f_{M,\bullet}\colon S_\bullet(\overline{\mathrm{BDiff}(M, \partial)}) \to S_\bullet(\mathrm{Sect}_c(\tau_M^*(\nu)))$ be the simplicial map that sends a p–simplex c in $S_\bullet(\overline{\mathrm{BDiff}(M, \partial)})$ to the germ of the $\mathrm{S}\Gamma_n$–structure induced by $(\mathrm{id} \times \mathrm{pr}_2)^*(c)$ at $\Delta^p \times \mathrm{diag}\, M$, where pr_2 is the projection to the second factor. This map is obviously $S_\bullet(\mathrm{Diff}(M, \partial))$–equivariant. Hence, using the Mather–Thurston theorem, the map $f_{M,\bullet}$ also induces a homology isomorphism between homotopy quotients,

$$|S_\bullet(\overline{\mathrm{BDiff}(M, \partial)})/\!\!/|S_\bullet(\mathrm{Diff}(M, \partial))| \to |S_\bullet(\mathrm{Sect}_c(\tau_M^*(\nu)))|/\!\!/|S_\bullet(\mathrm{Diff}(M, \partial))|.$$

Therefore, Mather and Thurston's theorem imply that $\mathrm{BDiff}^{\delta}(M,\partial)$ is homology equivalent to $|S_\bullet(\mathrm{Sect}_c(\tau_M^*(\nu)))/\!\!/|S_\bullet(\mathrm{Diff}(M,\partial))|$.

1.2.3 Homological stability for tangential structures Recall from [37] that a *tangential structure* is a map $\theta\colon B\to \mathrm{BGL}_2(\mathbb{R})$ from a path-connected space B to $\mathrm{BGL}_2(\mathbb{R})$. A θ–structure on a surface Σ is a bundle map $\mathrm{T}\Sigma\to\theta^*\gamma_2$, where γ_2 is the universal bundle over $\mathrm{BGL}_2(\mathbb{R})$. We denote the space of θ–structures on a surface Σ by $\mathrm{Bun}(\mathrm{T}\Sigma,\theta^*\gamma_2)$ and equip it with the compact–open topology. For a collared surface Σ with a choice of collar $c\colon \partial\Sigma\times[0,1)\to\Sigma$, we fix a boundary condition $\ell_{\partial\Sigma}\colon \epsilon^1\oplus \mathrm{T}(\partial\Sigma)\to\theta^*\gamma_2$. And we define $\mathrm{Bun}_\partial(\mathrm{T}\Sigma,\theta^*\gamma_2;\ell_{\partial\Sigma})$ to be the space of bundle maps $\ell\colon \mathrm{T}\Sigma\to\theta^*\gamma_2$ such that $\ell_{\partial\Sigma}=Dc|_{\{0\}\times\partial\Sigma}\circ\ell|_{\partial\Sigma}$. Note that the group $\mathrm{Diff}(\Sigma,\partial)$ naturally acts on $\mathrm{Bun}_\partial(\mathrm{T}\Sigma,\theta^*\gamma_2;\ell_{\partial\Sigma})$. The moduli space of θ–structures on surfaces of topological type Σ with boundary condition $\ell_{\partial\Sigma}$ is the homotopy quotient of the action $\mathrm{Diff}(\Sigma,\partial)$ on $\mathrm{Bun}_\partial(\mathrm{T}\Sigma,\theta^*\gamma_2;\ell_{\partial\Sigma})$ and we denote it by

$$\mathcal{M}^{\theta}(\Sigma;\ell_{\partial\Sigma}) := \mathrm{Bun}_\partial(\mathrm{T}\Sigma,\theta^*\gamma_2;\ell_{\partial\Sigma})/\!\!/\mathrm{Diff}(\Sigma,\partial). \tag{1-11}$$

If we do not mention the boundary condition $\ell_{\partial\Sigma}$, we mean the standard boundary condition on $\partial\Sigma$ in the sense of [37, Definition 4.1]. Henceforth, we consider ν–structures, where $\nu\colon \mathrm{B}S\Gamma_2\to\mathrm{BGL}_2^+(\mathbb{R})$.

Recall that a foliated microbundle in the neighborhood of the diagonal of $\Sigma\times\Sigma$ which is transverse to the fibers of the projection pr_1 is a section in $\mathrm{Sect}_c(\tau_\Sigma^*(\nu))$. This then gives a bundle map from $T\Sigma$ to $\nu^*(\gamma_2)$. Therefore, there is a canonical map

$$\epsilon\colon |S_\bullet(\mathrm{Sect}_c(\tau_\Sigma^*(\nu)))| \xrightarrow{\simeq} \mathrm{Bun}_\partial(\mathrm{T}\Sigma,\nu^*\gamma_2;l_{\partial\Sigma}).$$

Using the fact that $\mathrm{Bun}_\partial(\mathrm{T}\Sigma,\gamma_2)$ is contractible [9, Lemma 5.1], one can show that there exists a homotopy inverse to the map ϵ, hence it is a weak equivalence. There is an action of $|S_\bullet(\mathrm{Diff}(\Sigma,\partial))|$ on the left-hand side and there is an action of $\mathrm{Diff}(\Sigma,\partial)$ on the right-hand side, and also there is a canonical map $\epsilon'\colon |S_\bullet(\mathrm{Diff}(\Sigma,\partial))|\to\mathrm{Diff}(\Sigma,\partial)$ which is weakly equivalent. The augmentation map ϵ is readily seen to be equivariant with respect to the map ϵ'. Hence, using the Serre spectral sequence we deduce that the induced map

$$|S_\bullet(\mathrm{Sect}_c(\tau_\Sigma^*(\nu)))|/\!\!/|S_\bullet(\mathrm{Diff}(\Sigma,\partial))|\to\mathcal{M}^{\nu}(\Sigma) \tag{1-12}$$

is a homology isomorphism. As we saw in Construction 1.10, we have a map from $\mathrm{BDiff}^{\delta}(\Sigma,\partial)$ to $|S_\bullet(\mathrm{Sect}_c(\tau_\Sigma^*(\nu)))|/\!\!/|S_\bullet(\mathrm{Diff}(\Sigma,\partial))|$ which is homology equivalent. For the future reference, we record this fact as a lemma.

Lemma 1.13 *There is a map from* $\mathrm{BDiff}^{\delta}(\Sigma,\partial)$ *to* $\mathcal{M}^{\nu}(\Sigma)$ *that induces an isomorphism on homology.*

Hence Lemma 1.13 reduces the proof of Theorem 0.1 to the homological stability of the moduli space of ν–structures, which then follows from a general theorem due to Randal-Williams [37, Theorem 7.1] about homological stability of moduli spaces of θ–structures satisfying certain properties. Qualitatively, he proved that if the connected components of $\mathcal{M}^{\nu}(\Sigma)$ stabilizes with respect to the genus of the surface Σ, the moduli space $\mathcal{M}^{\nu}(\Sigma)$ exhibits homological stability in a certain range.

To show the stability of connected components, consider the exact sequence of homotopy groups

$$\pi_1(\mathrm{BDiff}(\Sigma,\partial)) \to \pi_0(\mathrm{Sect}_c(\tau_{\Sigma}^*(\nu))) \to \pi_0(\mathcal{M}^{\nu}(\Sigma)) \to \pi_0(\mathrm{BDiff}(\Sigma,\partial)).$$

The classifying space $\mathrm{BDiff}(\Sigma,\partial)$ is path-connected and since by Remark 1.5 the space $\overline{\mathrm{B}\Gamma}_2$ is at least 3–connected, the section space $\mathrm{Sect}_c(\tau_{\Sigma}^*(\nu))$ is also path-connected. Hence, $\pi_0(\mathcal{M}^{\nu}(\Sigma))$ is trivial.

To find a stability range, Randal-Williams [37, Definition 6.2] defined a notion of k–triviality and proved that if a θ–structure stabilizes at genus h, then it would be $(2h+1)$–trivial. Since ν–structure stabilizes at genus 0, by [37, Theorem 7.1] the stability range for stabilization maps is the same as the stability range for the orientation structure $\mathrm{BSO}(2) \to \mathrm{BO}(2)$. Thus, we have:

Theorem 1.14 *Let Σ be a collared surface and let $\iota\colon \Sigma \hookrightarrow \Sigma'$ be an embedding of Σ into a surface Σ' which may not have boundary. As we explained in Section 1.1, this embedding induces a map*

$$H_*(\mathrm{BDiff}^{\delta}(\Sigma,\partial);\mathbb{Z}) \to H_*(\mathrm{BDiff}^{\delta}(\Sigma',\partial);\mathbb{Z})$$

which is an isomorphism as long as $* \le \frac{1}{3}(2g(\Sigma)-2)$ and an epimorphism provided that $* \le \frac{2}{3}g(\Sigma)$.

Remark 1.15 Using the same idea and the theorem of McDuff [24] about volume-preserving diffeomorphisms, we could show that the discrete group of symplectomorphisms $\mathrm{Symp}^{\delta}(\Sigma,\partial)$ that are supported away from the boundary also exhibit homological stability. We pursue the study of the group of symplectomorphisms in a different paper [34].

2 Stable homology of surface diffeomorphisms made discrete

Given that we established the relation between $\mathrm{BDiff}^{\delta}(\Sigma,\partial)$ for a collared surface Σ and the moduli space of $\mathrm{S}\Gamma_2$–structures on surfaces of the topological type Σ in Lemma 1.13, we can use the machinery developed in [9; 7] to study the stable homology of the moduli space of tangential structures.

2.1 Cobordism category with $S\Gamma_2$–structure

Recall the definition of the cobordism category equipped with θ–structures from [9, Definition 5.2];

Definition 2.1 Let $\mathcal{C}_\theta$ be the topological category whose space of objects is given by the pairs of real numbers a and 1–dimensional closed submanifolds M such that $(a, M) \subset \mathbb{R} \times \mathbb{R}^\infty$ and whose space of morphisms from (a, M) to (a', M') for $a < a'$ is given by a cobordism Σ so that $\Sigma \subset [a, a'] \times \mathbb{R}^\infty$ is a surface equipped with a θ–structure and is *collared* near the boundary, which means that it coincides with $[a, a'] \times M$ near $\{a\} \times \mathbb{R}^\infty$ and with $[a, a'] \times M'$ near $\{a'\} \times \mathbb{R}^\infty$. For a careful treatment of how this category is enriched over topological spaces consult [7, Section 2]. We shall write $\mathcal{C}_+$ for the cobordism category with the orientation structure.

Definition 2.2 For the map $\nu\colon \mathrm{BS\Gamma}_2 \to \mathrm{BGL}_2^+(\mathbb{R}) = \widetilde{\mathrm{Gr}}_2(\mathbb{R}^\infty)$, where $\widetilde{\mathrm{Gr}}_2(\mathbb{R}^\infty)$ is the oriented Grassmannian of two planes in $\mathbb{R}^\infty$, we can associate a Thom spectrum $\mathrm{MT}\nu$ as follows: First let $\mathrm{BS\Gamma}_2(\mathbb{R}^n) := \nu^{-1}(\widetilde{\mathrm{Gr}}_2(\mathbb{R}^n))$, where it sits in the pullback diagram

$$\begin{array}{ccc} \mathrm{BS\Gamma}_2(\mathbb{R}^n) & \longrightarrow & \mathrm{BS\Gamma}_2 \\ {\scriptstyle \nu_n}\downarrow & & \downarrow{\scriptstyle \nu} \\ \widetilde{\mathrm{Gr}}_2(\mathbb{R}^n) & \longrightarrow & \mathrm{BGL}_2^+(\mathbb{R}) \end{array}$$

Let U_n be the orthogonal complement of the tautological 2–plane bundle over $\widetilde{\mathrm{Gr}}_2(\mathbb{R}^n)$. Then the n^{th} space of the spectrum $\mathrm{MT}\nu$ is the Thom space of the pullback bundle $\nu_n^*(\mathrm{U}_n)$.

The main theorem of [9] implies that there exists a weak equivalence

$$\mathrm{B}\mathcal{C}_\nu \xrightarrow{\simeq} \Omega^{\infty-1}\mathrm{MT}\nu,$$

which is induced by a functor from the category $\mathcal{C}_\nu$ to the category $\mathrm{Path}(\Omega^{\infty-1}\mathrm{MT}\nu)$, whose objects are points in $\Omega^{\infty-1}\mathrm{MT}\nu$ and whose morphisms are continuous paths. We shall briefly recall below how this functor is constructed and refer the reader to [19, Section 2] for further details.

2.2 The map α in Theorem 0.4

A morphism Σ in the category $\mathcal{C}_\nu$ is a surface with a collared boundary that is embedded in $[a, a'] \times \mathbb{R}^{n-1}$ for some n. We say that this morphism is *fatly embedded* if the canonical map from the normal bundle $N\Sigma$ to $\mathbb{R}^n$ restricts to an embedding of the unit disk bundle into $[a,a']\times\mathbb{R}^{n-1}$. In the definition of the cobordism category, one can consider only fatly embedded morphisms without changing the homotopy type of the realization of the category. Thus the Pontryagin–Thom construction gives a map from $[a,a']_+\wedge S^{n-1}$ to the Thom space of $N\Sigma$. Since the surface Σ is equipped with a ν–structure, the Gauss map $\Sigma\to\widetilde{\mathrm{Gr}}_2(\mathbb{R}^n)$ that classifies the tangent bundle can be lifted to $\mathrm{BS}\Gamma_2(\mathbb{R}^n)$. Therefore, we have the pullback diagram

$$\begin{tikzcd} N\Sigma \arrow[r] \arrow[d] & \nu_n^*(U_n) \arrow[d] \\ \Sigma \arrow[r] & \mathrm{BS}\Gamma_2(\mathbb{R}^n) \end{tikzcd}$$

Hence, one obtains the map

$$[a,a']_+\wedge S^{n-1}\to \mathrm{Th}(N\Sigma)\to \mathrm{Th}(\nu_n^*(U_n)).$$

By the adjointness, we obtain a path

$$[a,a']\to\Omega^{n-1}\mathrm{Th}(\nu_n^*(U_n))\subset\Omega^{\infty-1}\mathrm{MT}\nu.$$

This construction gives rise to a functor from the modification of $\mathcal{C}_\nu$ to the path category $\mathrm{Path}(\Omega^{\infty-1}\mathrm{MT}\nu)$. Since the modification of $\mathcal{C}_\nu$ does not change its homotopy type and the geometric realization of $\mathrm{Path}(\Omega^{\infty-1}\mathrm{MT}\nu)$ has the same homotopy type as $\Omega^{\infty-1}\mathrm{MT}\nu$, the functor induces a well-defined map up to homotopy between geometric realizations,

$$\mathrm{B}\mathcal{C}_\nu\to\Omega^{\infty-1}\mathrm{MT}\nu.$$

One can choose a certain model for the homotopy quotient in (1-11) (see [9, Section 5]) so that the space $\mathcal{M}^\nu(\Sigma)$ becomes a subspace of the morphism space in $\mathcal{C}_\nu$. Therefore, we obtain a natural map

$$\mathcal{M}^\nu(\Sigma)\to\Omega\mathrm{B}\mathcal{C}_\nu\to\Omega^\infty\mathrm{MT}\nu. \tag{2-3}$$

Note that the map $\nu\colon \mathrm{BS}\Gamma_2\to\mathrm{BGL}_2^+(\mathbb{R})$ induces a functor $\mathcal{C}_\nu\to\mathcal{C}_+$, hence by the naturality of the above constructions, we have the homotopy commutative diagram

$$\begin{tikzcd} \mathcal{M}^\nu(\Sigma) \arrow[r] \arrow[d] & \Omega\mathrm{B}\mathcal{C}_\nu \arrow[r] \arrow[d] & \Omega^\infty\mathrm{MT}\nu \arrow[d] \\ \mathcal{M}^+(\Sigma) \arrow[r] & \Omega\mathrm{B}\mathcal{C}_+ \arrow[r] & \Omega^\infty\mathrm{MTSO}(2) \end{tikzcd}$$

Recall that the space $\mathcal{M}^+(\Sigma)$ is a model for $\mathrm{BDiff}(\Sigma,\partial)$ and the space

$$|S_\bullet(\overline{\mathrm{BDiff}(M,\partial)})|/\!\!/|S_\bullet(\mathrm{Diff}(M,\partial))|$$

is a model for $\mathrm{BDiff}^\delta(\Sigma,\partial)$. Hence, we have the homotopy commutative diagram

(2-4)
$$\begin{tikzcd} |S_\bullet(\overline{\mathrm{BDiff}(M,\partial)})/\!\!/|S_\bullet(\mathrm{Diff}(M,\partial))| \arrow[r, "g"] \arrow[d] & \mathcal{M}^\nu(\Sigma) \arrow[d] \\ \mathrm{B}|S_\bullet(\mathrm{Diff}(M,\partial))| \arrow[r, "\simeq"] & \mathcal{M}^+(\Sigma) \end{tikzcd}$$

where the bottom horizontal map is a weak equivalence by Milnor's theorem [25] and the top horizontal map g is given by the composition of the map in Construction 1.10 and (1-12). The map α now is given by the composition of the map $g\colon \mathrm{BDiff}^\delta(\Sigma)\to\mathcal{M}^\nu(\Sigma)$ in (2-4) and the maps in (2-3). Hence, we obtain a homotopy commutative diagram

(2-5)
$$\begin{tikzcd} \mathrm{BDiff}^\delta(\Sigma,\partial) \arrow[r, "\alpha"] \arrow[d] & \Omega^\infty\mathrm{MT}\nu \arrow[d] \\ \mathrm{BDiff}(\Sigma,\partial) \arrow[r] & \Omega^\infty\mathrm{MTSO}(2) \end{tikzcd}$$

Theorem 2.6 *In diagram (2-5) the horizontal maps, in the stable range of Theorem 1.14, induce homology isomorphisms onto the connected components that they hit.*

Remark 2.7 The volume-preserving case reproduces [17, Theorem 4] and more, which we will pursue elsewhere [34].

Sketch of the proof of Theorem 2.6 The fact that the bottom horizontal map in the stable range induces an isomorphism on homology is the celebrated Madsen–Weiss theorem [20; 9, Theorem 7.2]. Hence, we only sketch the proof for the similar statement for the map α. We replace $\mathrm{BDiff}^\delta(\Sigma_{g,k},\partial)$ by the homology equivalent space $\mathcal{M}^\nu(\Sigma_{g,k})$. Recall that the main theorem of [9] implies that the geometric realization of $\mathcal{C}_\nu$ is weakly homotopy equivalent to $\Omega^{\infty-1}\mathrm{MT}\nu$. Therefore, from the above discussion, we only need to prove that the map

(2-8)
$$\mathcal{M}^\nu(\Sigma_{g,k})\to\Omega\mathrm{B}\mathcal{C}_\nu$$

in the stable range induces an isomorphism on homology. As we shall briefly explain, this follows from applying the argument in [9, Section 7] to the category $\mathcal{C}_\nu$. Following Tillmann [44], we need to consider a smaller category, which is called *the positive boundary* subcategory $\mathcal{C}_{\nu,\partial}\subset\mathcal{C}_\nu$, whose space of objects is the same as $\mathcal{C}_\nu$ and whose space of morphisms from M_0 to M_1 consists of those pairs $(\Sigma,t)\in\mathcal{C}_\nu$ where $\pi_0(M_1)\to\pi_0(\Sigma)$ is surjective. By [9, Theorem 6.1] the inclusion of $\mathcal{C}_{\nu,\partial}$ in $\mathcal{C}_\nu$ induces a map between geometric realizations

(2-9)
$$\mathrm{B}\mathcal{C}_{\nu,\partial}\xrightarrow{\simeq}\mathrm{B}\mathcal{C}_\nu,$$

which is a weak equivalence. There is nothing special about the tangential structure ν in (2-9). Now given the homological stability in Theorem 1.14, the standard group completion argument (see [9, Proposition 7.1]) implies that the first map in

$$\mathbb{Z} \times \mathcal{M}^{\nu}(\Sigma_{\infty,k}) \xrightarrow{H_*-\text{iso}} \Omega \mathrm{B}\mathcal{C}_{\nu,\partial} \xrightarrow{\simeq} \Omega^{\infty}\mathrm{MT}\nu,$$

induces an isomorphism on homology, and hence the map in (2-8) induces a homology isomorphism in the stable range. □

Remark 2.10 There is a more direct description of the diagram (2-5) without invoking the cobordism category. To briefly explain this alternative description, let $\Sigma \to E \xrightarrow{\pi} M$ be a surface bundle and let $T\pi$ denote the vertical tangent bundle. For paracompact base M, one can find a fiberwise embedding $f\colon E \hookrightarrow M \times \mathbb{R}^N$ with a tubular neighborhood. Collapsing the complement of the tubular neighborhood to the basepoint and identifying the tubular neighborhood with the open disk bundle of the fiberwise normal bundle Nf, gives the map $\Sigma^N(M_+) \to \mathrm{Th}(Nf)$. Stably it gives *a pretransfer map*, well-defined up to homotopy,

$$\mathrm{pretrf}_{\pi}\colon \Sigma^{\infty}(M_+) \to \mathbf{Th}(-T\pi),$$

where $\mathbf{Th}(-T\pi)$ is the Thom spectrum of the virtual bundle $-T\pi$. Let

$$\theta\colon B \to \mathrm{BGL}_2^+(\mathbb{R})$$

be a tangential structure. Recall that γ_2 is the tautological bundle over $\mathrm{BGL}_2^+(\mathbb{R})$. If the vertical tangent bundle is equipped with a θ–structure, ie the map $E \to \mathrm{BGL}_2^+(\mathbb{R})$ classifying $T\pi$ has a choice of lift to B, then we obtain a well-defined map up to homotopy

$$\Sigma^{\infty}(M_+) \xrightarrow{\mathrm{pretrf}_{\pi}} \mathbf{Th}(-T\pi) \to \mathbf{Th}(-\theta^*(\gamma_2)). \tag{2-11}$$

Now consider the pullback diagram

$$\begin{array}{ccc} \Sigma/\!/\mathrm{Diff}^{\delta}(\Sigma,\partial) & \longrightarrow & \Sigma/\!/\mathrm{Diff}(\Sigma,\partial) \\ \pi' \big\downarrow & & \pi \big\downarrow \\ \mathrm{BDiff}^{\delta}(\Sigma,\partial) & \longrightarrow & \mathrm{BDiff}(\Sigma,\partial) \end{array}$$

Note that $T\pi'$ is the pullback of $T\pi$ and it has a ν–structure, therefore, by the naturality of the pretransfer map, we have a commutative diagram of spectra

$$\begin{tikzcd}
\Sigma^\infty(\mathrm{BDiff}^\delta(\Sigma,\partial)_+) \arrow[r] \arrow[d] & \mathbf{Th}(-\nu^*(\gamma_2)) \arrow[d] \\
\Sigma^\infty(\mathrm{BDiff}(\Sigma,\partial)_+) \arrow[r] & \mathbf{Th}(-\gamma_2)
\end{tikzcd}$$

which gives a homotopy commutative diagram of spaces:

$$\begin{tikzcd}
\mathrm{BDiff}^\delta(\Sigma,\partial) \arrow[r] \arrow[d] & \Omega^\infty \mathrm{MT}\nu \arrow[d] \\
\mathrm{BDiff}(\Sigma,\partial) \arrow[r] & \Omega^\infty \mathrm{MTSO}(2)
\end{tikzcd}$$

The fact this diagram is the same as diagram (2-5) follows from the standard fact that the construction using pretransfer and the construction via the cobordism category are homotopic (see [19, Section 2] for more details).

2.3 Comparison of $\mathbf{BDiff}^\delta(\Sigma_{g,k},\partial)$ and $\mathbf{BDiff}(\Sigma_{g,k},\partial)$ in the stable range

Let $\mathrm{Diff}^\delta(\Sigma_{\infty,k},\partial)$ denote the colimit of the groups $\mathrm{Diff}^\delta(\Sigma_{g,k},\partial)$ as g varies using the stabilization map between them. Thus by taking the colimit of the diagram (2-5), we have a homotopy commutative diagram

$$\begin{tikzcd}
\mathrm{BDiff}^\delta(\Sigma_{\infty,k},\partial) \arrow[r, "\iota"] \arrow[d] & \mathrm{BDiff}(\Sigma_{\infty,k},\partial) \arrow[d] \\
\Omega^\infty_0 \mathrm{MT}\nu \arrow[r, "\Omega^\infty \boldsymbol{\nu}"] & \Omega^\infty_0 \mathrm{MTSO}(2)
\end{tikzcd}$$

where the right vertical map is a homology isomorphism by the Madsen–Weiss theorem and the left vertical map is also a homology isomorphism by Theorem 2.6. For a prime p, let

$$\iota^p_*\colon H_*(\mathrm{BDiff}^\delta(\Sigma_{\infty,k},\partial);\mathbb{F}_p) \to H_*(\mathrm{BDiff}(\Sigma_{\infty,k},\partial);\mathbb{F}_p)$$

be the map induced by ι. To study the map ι^p_*, instead we study the map

$$\Omega^\infty \boldsymbol{\nu}\colon \Omega^\infty_0 \mathrm{MT}\nu \to \Omega^\infty_0 \mathrm{MTSO}(2)$$

between infinite loop spaces after p–completion (see [23] for a definition of p–completion).

Our third main theorem is the following splitting theorem after p–completion.

Theorem 2.12 *For all primes p, after p–completion the map $\Omega^\infty \boldsymbol{\nu}$ admits a section, ie it is split surjection after p–adic completion.*

Corollary 2.13 *The map ι_*^p is a split surjection for all primes p.*

Proof of Theorem 2.12 To show that $\Omega^\infty \boldsymbol{\nu}$ has a section after p–completion, it is sufficient to prove this on the spectrum level, ie, as we shall see, it is enough to prove that the map

$$\boldsymbol{\nu}\colon \mathrm{MT}\nu \to \mathrm{MTSO}(2) \tag{2-14}$$

is a split surjection after p–completion of spectra. The reason is, in general, if A is a spectrum and $A_p^\wedge$ is its p–completion, then $\Omega_0^\infty(A_p^\wedge)$ is a p–completed space and it receives a map from $(\Omega_0^\infty A)_p^\wedge$ which is weakly equivalent. The map is induced by

$$\Omega_0^\infty A \to \Omega_0^\infty(A_p^\wedge),$$

which factors through $(\Omega_0^\infty A)_p^\wedge$ by the universal property of p–completion. Suppose the map $\boldsymbol{\nu}_p^\wedge$ which is induced by $\boldsymbol{\nu}$ between p–completions of the Thom spectra has a section denoted by $\boldsymbol{s}$. Consider the diagram

$$\begin{tikzcd}
\Omega_0^\infty(\mathrm{MTSO}(2))_p^\wedge \arrow[r, "\Omega^\infty \boldsymbol{s}"] & \Omega_0^\infty(\mathrm{MT}\nu)_p^\wedge \\
(\Omega_0^\infty \mathrm{MTSO}(2))_p^\wedge \arrow[u, "\simeq"] \arrow[r, dashed] & (\Omega_0^\infty \mathrm{MT}\nu)_p^\wedge \arrow[u, "\simeq"']
\end{tikzcd}$$

where the vertical maps are induced by the universal property of the p–completion. Hence, $\Omega^\infty \boldsymbol{s}$ induces a section for $\Omega^\infty \boldsymbol{\nu}$ after p–completion. We are left to prove the map of spectra (2-14) is a split surjection after p–completion.

Recall that the map

$$\nu\colon \mathrm{BS}\Gamma_2 \to \mathrm{BGL}_2^+(\mathbb{R}) \simeq \mathrm{BSO}(2)$$

that classifies the normal bundle of the codimension-2 Haefliger structures is induced by a map between groupoids $\overline{\nu}\colon \mathrm{S}\Gamma_2 \to \mathrm{GL}_2^+(\mathbb{R})$, where $\overline{\nu}$ sends a germ to its derivative.

The key point is that there is an obvious map between groupoids

$$\overline{\alpha}\colon \mathrm{SO}(2)^\delta \to \mathrm{S}\Gamma_2$$

that sends every rotation to its germ as a diffeomorphism at the origin and makes the diagram

$$\begin{tikzcd}
\mathrm{SO}(2)^\delta \arrow[r, "\overline{\alpha}"] \arrow[dr, "\overline{\beta}"'] & \mathrm{S}\Gamma_2 \arrow[d, "\overline{\nu}"] \\
& \mathrm{GL}_2(\mathbb{R})
\end{tikzcd} \tag{2-15}$$

commute, where $\bar{\beta}$ is the composition $\mathrm{SO}(2)^\delta \to \mathrm{SO}(2) \hookrightarrow \mathrm{GL}_2(\mathbb{R})$. The map $\bar{\alpha}$ induces a map between classifying spaces of the groupoids, which we denote by α. The commutativity of (2-15) implies that the composition $\nu \circ \alpha$ is homotopic to the identity.

The map $\nu \circ \alpha$ gives a tangential structure and let $\mathrm{MT}(\nu \circ \alpha)$ denote the Thom spectrum of the virtual bundle $(\nu \circ \alpha)^*(-\gamma_2)$ over $\mathrm{BSO}(2)^\delta$. Consider the maps between Thom spectra

$$\mathrm{MT}(\nu \circ \alpha) \xrightarrow{\boldsymbol{\alpha}} \mathrm{MT}\nu \xrightarrow{\boldsymbol{\nu}} \mathrm{MTSO}(2).$$

If we show that the map of spectra $\boldsymbol{\nu} \circ \boldsymbol{\alpha}$ is a split surjection after p–completion, we are done. We prove that $(\boldsymbol{\nu} \circ \boldsymbol{\alpha})^*$ is an isomorphism on mod-p cohomology, hence $\boldsymbol{\nu} \circ \boldsymbol{\alpha}$ actually induces a weak equivalence after p–completion. Let $\overline{\mathrm{BS}}^1$ denote the homotopy fiber of the map

$$\mathrm{B}(\mathrm{S}^1)^\delta \to \mathrm{BS}^1.$$

It is a special case of [27, Lemma 3] that the space $\overline{\mathrm{BS}}^1$ has a mod-p homology of a point. Thus, we deduce that $H^*(\mathrm{B}(\mathrm{S}^1)^\delta; \mathbb{F}_p) = H^*(\mathrm{BS}^1; \mathbb{F}_p)$. Hence, using Thom isomorphism, it follows that

$$(\boldsymbol{\nu} \circ \boldsymbol{\alpha})^*\colon H^*(\mathrm{MTSO}(2); \mathbb{F}_p) \to H^*(\mathrm{MT}(\nu \circ \alpha); \mathbb{F}_p)$$

is an isomorphism. □

3 Applications to flat surface bundles

In this section, we explore the consequences of Theorem 2.6 and Theorem 2.12 for flat surface bundles.

3.1 The spectrum MT$\boldsymbol{\nu}$ and a fiber sequence of infinite loop spaces

By studying the homotopy groups of $\Omega_0^\infty \mathrm{MT}\nu$, we prove Theorem 0.18 and also we find a new description for $H_2(\mathrm{BDiff}^\delta(\Sigma_g); \mathbb{Z})$.

Let $\mathrm{MTSO}(n)$ denote the Madsen–Tillmann spectrum for the orientation structure $\mathrm{SO}(n) \to \mathrm{O}(n)$. There exists a cofiber sequence of spectra and a fiber sequence of infinite loop space (see [9, Proposition 3.1])

$$\mathrm{MTSO}(n) \to \Sigma^\infty(\mathrm{BSO}(n)_+) \to \mathrm{MTSO}(n-1), \tag{3-1}$$

$$\Omega^\infty \mathrm{MTSO}(n) \to \Omega^\infty \Sigma^\infty(\mathrm{BSO}(n)_+) \to \Omega^\infty \mathrm{MTSO}(n-1), \tag{3-2}$$

where here $+$ means a disjoint basepoint. For $n = 2$, the fiber sequence of the infinite loop space plays an important role in computing mod p homology of the mapping class group [5]. In dimension 2, we prove that there exists a similar fiber sequence for $\Omega^\infty\mathrm{MT}\nu$.

Theorem 3.3 *There is a homotopy fibration sequence*

$$\Omega^\infty\mathrm{MT}\nu \to \Omega^\infty\Sigma^\infty((\mathrm{B}\mathrm{S}\Gamma_2)_+) \to \Omega^\infty\Sigma^{\infty-1}(\overline{\mathrm{B}\mathrm{S}\Gamma_2}_+). \tag{3-4}$$

Proof Let η and ξ be two vector bundles over a topological space X; we have the general cofiber sequence of Thom spaces (see [42, Lemma 4.3.1])

$$\mathrm{Th}(p^*(\xi)) \to \mathrm{Th}(\xi) \to \mathrm{Th}(\xi\oplus\eta), \tag{3-5}$$

where $p\colon \mathrm{S}(\eta)\to X$ is the sphere bundle of η.

Let X be $\mathrm{B}\mathrm{S}\Gamma_2$ and ξ and η be the virtual bundles $\nu^*(-\gamma_2)$ and $\nu^*(\gamma_2)$, respectively. Note that

$$S^1 \to \overline{\mathrm{B}\mathrm{S}\Gamma_2} \xrightarrow{p} \mathrm{B}\mathrm{S}\Gamma_2$$

is the sphere bundle of $\nu^*(\gamma_2)$, hence $p^*(\nu^*(-\gamma_2))$ is a trivial bundle over $\overline{\mathrm{B}\mathrm{S}\Gamma_2}$. Using (3-5), we obtain the following cofiber sequence of the spectra:

$$\Sigma^{\infty-2}(\overline{\mathrm{B}\mathrm{S}\Gamma_2}_+) \to \mathrm{MT}\nu \to \Sigma^\infty((\mathrm{B}\mathrm{S}\Gamma_2)_+) \to \Sigma^{\infty-1}(\overline{\mathrm{B}\mathrm{S}\Gamma_2}_+).$$

Applying Ω^∞, we obtain the associated homotopy fibration sequence of infinite loop spaces. □

Lemma 3.6 *We have*

$$\pi_1(\Omega^\infty_0\mathrm{MT}\nu)=0,$$
$$0\to\pi_4(\overline{\mathrm{B}\mathrm{S}\Gamma_2})\to\pi_2(\Omega^\infty_0\mathrm{MT}\nu)\to\pi_2(\Omega^\infty_0\mathrm{MTSO}(2))\to 0,$$
$$\pi_3(\Omega^\infty_0\mathrm{MT}\nu)\twoheadrightarrow\pi_3(\Omega^\infty_0\mathrm{MTSO}(2)).$$

Proof Recall that for $g\geq 3$, the group $\mathrm{Diff}^\delta(\Sigma_g)$ is a perfect group because the identity component ${\mathrm{Diff}^\delta}_0(\Sigma_g)$ is a simple group [43] and the mapping class group $\pi_0(\mathrm{Diff}(\Sigma_g))$ is perfect for $g\geq 3$ [36]. Hence $H_1(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})=0$ for $g\geq 3$. On the other hand, Theorem 2.6 implies that $H_1(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})=H_1(\Omega^\infty_0\mathrm{MT}\nu;\mathbb{Z})=\pi_1(\Omega^\infty_0\mathrm{MT}\nu)$, where the second equality holds because $\Omega^\infty_0\mathrm{MT}\nu$ is an H–space. Therefore we have

$$\pi_1(\Omega^\infty_0\mathrm{MT}\nu)=0.$$

Also recall that, for stable homotopy groups, we have $\pi_i^s(X_+)=\pi_i^s(X)\oplus\pi_i^s$, $\pi_4^s=\pi_5^s=0$, $\pi_3^s=\mathbb{Z}/24$ and $\pi_2^s=\mathbb{Z}/2$. Given these facts, the long exact sequence of the homotopy groups of the fibration (3-4) is as follows:

$$\begin{tikzcd}
\cdots \arrow[r] & \pi_5^s(\overline{\mathrm{BS\Gamma}_2}) \arrow[r] & \pi_3(\Omega_0^\infty \mathrm{MT}\nu) \arrow[r] & \pi_3^s(\mathrm{BS\Gamma}_2) \oplus \mathbb{Z}/24 \arrow[dlll, rounded corners] \\
\pi_4^s(\overline{\mathrm{BS\Gamma}_2}) \arrow[r] & \pi_2(\Omega_0^\infty \mathrm{MT}\nu) \arrow[r] & \pi_2^s(\mathrm{BS\Gamma}_2) \oplus \mathbb{Z}/2 \arrow[r, two heads] & \pi_3^s(\overline{\mathrm{BS\Gamma}_2}_+)
\end{tikzcd}$$

By Remark 1.5, we know that the map $\nu\colon \mathrm{BS\Gamma}_2 \to \mathrm{BSO}(2)$ is at least 4–connected. Hence, the long exact sequence of the homotopy groups of the fibrations (3-4) becomes

(3-7)
$$\begin{tikzcd}
\cdots \arrow[r] & \pi_5^s(\overline{\mathrm{BS\Gamma}_2}) \arrow[r] & \pi_3(\Omega_0^\infty \mathrm{MT}\nu) \arrow[r] & \mathbb{Z}/24 \arrow[dlll, rounded corners] \\
\pi_4^s(\overline{\mathrm{BS\Gamma}_2}) \arrow[r] & \pi_2(\Omega_0^\infty \mathrm{MT}\nu) \arrow[r] & \pi_2^s(\mathbb{C}\mathrm{P}^\infty) \oplus \mathbb{Z}/2 \arrow[r, two heads, "\overset{d}{e}"] & \mathbb{Z}/24
\end{tikzcd}$$

and the long exact sequence of homotopy groups of the fibration in (3-2) is as follows:

(3-8)
$$\begin{tikzcd}
\cdots \arrow[r] & 0 \arrow[r] & \pi_3(\Omega_0^\infty \mathrm{MTSO}(2)) \arrow[r] & \mathbb{Z}/24 \arrow[dlll, rounded corners] \\
0 \arrow[r] & \pi_2(\Omega_0^\infty \mathrm{MTSO}(2)) \arrow[r] & \pi_2^s(\mathbb{C}\mathrm{P}^\infty) \oplus \mathbb{Z}/2 \arrow[r, two heads, "\overset{d'}{e'}"] & \mathbb{Z}/24
\end{tikzcd}$$

There are natural maps from corresponding terms in (3-7) to that of (3-8). In Lemma 3.11 below, we will prove that the map d has to be zero. Given Lemma 3.11, we obtain

(3-9)
$$\pi_3(\Omega_0^\infty \mathrm{MT}\nu) \twoheadrightarrow \pi_3(\Omega_0^\infty \mathrm{MTSO}(2)) = \mathbb{Z}/24.$$

The homology of $\Omega_0^\infty \mathrm{MTSO}(2)$ is the same as the stable homology of the mapping class group. Since the second stable homology of the mapping class group is $\mathbb{Z}$ [11], by the Hurewicz theorem we obtain that $\pi_2(\Omega_0^\infty \mathrm{MTSO}(2)) = \mathbb{Z}$. Hence, by comparing the maps e and e', we deduce that $\ker(e) = \mathbb{Z}$. Therefore, we have the short exact sequence

(3-10)
$$0 \to \pi_4^s(\overline{\mathrm{BS\Gamma}_2}) \to \pi_2(\Omega_0^\infty \mathrm{MT}\nu) \to \mathbb{Z} \to 0,$$

where the last map is induced by the map $\Omega_0^\infty \mathrm{MT}\nu \to \Omega_0^\infty \mathrm{MTSO}(2)$ on the second homotopy groups. Since $\overline{\mathrm{BS\Gamma}_2}$ is 3–connected, by the Hurewicz theorem, the fourth stable homotopy group of $\overline{\mathrm{BS\Gamma}_2}$ is the same as its fourth homotopy group, hence we obtain the second equality of the lemma. □

Lemma 3.11 *The map d in the long exact sequence (3-7) is zero.*

Proof Let Q denote the functor $\Omega^\infty \Sigma^\infty$. One can associate to the circle bundle

$$\overline{\mathrm{BS\Gamma}_2} \to \mathrm{BS\Gamma}_2$$

the circle transfer (ie the pretransfer for a circle bundle; see Remark 2.10), which is a map
$$\tau\colon Q((\mathrm{B}S\Gamma_2)_+) \to QS^{-1}((\overline{\mathrm{B}S\Gamma_2})_+),$$
where QS^{-1} denotes the functor $\Omega^\infty\Sigma^{\infty-1}$. Recall the fiber sequence (3-4)

$$\Omega^\infty \mathrm{MT}\nu \to Q((\mathrm{B}S\Gamma_2)_+) \xrightarrow{\tau} QS^{-1}(\overline{\mathrm{B}S\Gamma_2}_+) \tag{3-12}$$

where the second map is the circle transfer τ by the construction of the fiber sequence. Hence the map d is the map induced by the circle transfer τ on the third homotopy groups. In order to show that d is zero, we consider the pullback diagram

$$\begin{tikzcd} S^1 \arrow[r] \arrow[d] & \overline{\mathrm{B}S\Gamma_2} \arrow[d] \\ * \arrow[r] & \mathrm{B}S\Gamma_2 \end{tikzcd}$$

where the bottom horizontal map is a basepoint of $\mathrm{B}S\Gamma_2$. By naturality of the circle transfer, we have the commutative diagram

$$\begin{tikzcd} QS^{-1}(S^1_+) \arrow[r, "g"] & QS^{-1}(\overline{\mathrm{B}S\Gamma_2}_+) \\ QS^0 \arrow[r] \arrow[u, "f"] & Q((\mathrm{B}S\Gamma_2)_+) \arrow[u, "\tau"'] \end{tikzcd}$$

where f is the circle transfer for the trivial circle bundle over a point and g is the induced by naturality of the transfer map. Given that the map $\nu\colon \mathrm{B}S\Gamma_2 \to \mathrm{BGL}_2^+(\mathbb{R}) \simeq \mathbb{C}\mathrm{P}^\infty$ is 4–connected (Remark 1.5), we obtain that
$$\pi_3^s(\mathrm{B}S\Gamma_2) = \pi_3^s(\mathbb{C}\mathrm{P}^\infty) = 0.$$
Therefore, the bottom map in the above diagram induces an isomorphism on π_3. Hence, to show that d is zero, it is enough to show that g induces zero on π_3. Note that the disjoint basepoint splits off naturally, ie $QS^{-1}(S^1_+) \simeq QS^{-1}(S^1) \times QS^{-1}$ and similarly we have $QS^{-1}(\overline{\mathrm{B}S\Gamma_2}_+) \simeq QS^{-1}(\overline{\mathrm{B}S\Gamma_2}) \times QS^{-1}$. Recall that $\pi_3(QS^{-1}) = \pi_4^s = 0$, so we need to show that the induced map
$$\pi_3(QS^{-1}(S^1)) \to \pi_3(QS^{-1}(\overline{\mathrm{B}S\Gamma_2}))$$
is zero. Given that $\overline{\mathrm{B}S\Gamma_2}$ is 3–connected (Remark 1.5) and in particular simply connected, the map $S^1 \to \overline{\mathrm{B}S\Gamma_2}$ is null-homotopic. Therefore, the map
$$QS^{-1}(S^1) \to QS^{-1}(\overline{\mathrm{B}S\Gamma_2})$$
is also nullhomotopic, which implies that the map g induces zero on π_3. □

Theorem 3.13 *Fix an embedding of* $\mathbb{R}^2 \hookrightarrow \Sigma_{g,k}$. *This embedding induces a map*

$$\mathrm{BDiff}_c^\delta(\mathbb{R}^2) \to \mathrm{BDiff}^\delta(\Sigma_{g,k}, \partial),$$

which for $g \geq 4$ *gives the short exact sequence*

$$0 \to H_2(\mathrm{BDiff}_c^\delta(\mathbb{R}^2); \mathbb{Z}) \to H_2(\mathrm{BDiff}^\delta(\Sigma_{g,k}, \partial); \mathbb{Z}) \to H_2(\mathrm{BDiff}(\Sigma_{g,k}, \partial); \mathbb{Z}) \to 0.$$

Remark 3.14 Using Theorem 1.3 for $M \cong \mathbb{R}^2$ and the connectivity of $\overline{\mathrm{B}S\Gamma_2}$, one can in fact show that there is a natural map

$$H_2(\mathrm{BDiff}^\delta(\Sigma_{g,k}, \partial); \mathbb{Z}_{(p)}) \to H_4(\mathrm{B}S\Gamma_2; \mathbb{Z}_{(p)}),$$

which induces an isomorphism for $g \geq 4$ and primes $p > 3$ (see [34]).

Proof By Theorem 2.6, we know that, for $g \geq 4$,

$$H_2(\mathrm{BDiff}^\delta(\Sigma_{g,k}, \partial); \mathbb{Z}) \xrightarrow{\cong} H_2(\Omega_0^\infty \mathrm{MT}\nu; \mathbb{Z}).$$

Recall by Thurston's theorem on the perfectness of the identity component of diffeomorphism groups and Powell's theorem on the perfectness of the mapping class group, the group $\mathrm{Diff}(\Sigma_{g,k}; \partial)$ is a perfect group for $g > 2$, which implies that $\Omega_0^\infty \mathrm{MT}\nu$ is simply connected. Hence, by the Hurewicz theorem and (3-10), we have the short exact sequence

$$0 \to \pi_4(\overline{\mathrm{B}S\Gamma_2}) \to H_2(\Omega_0^\infty \mathrm{MT}\nu) \to H_2(\Omega_0^\infty \mathrm{MTSO}(2)) \to 0. \tag{3-15}$$

By Theorem 1.3 or Remark 1.5, we know that $\overline{\mathrm{BDiff}_c(\mathbb{R}^2)}$ is homology equivalent to $\Omega^2\overline{\mathrm{B}S\Gamma_2}$. Since $\mathrm{Diff}_c(\mathbb{R}^2)$ is contractible [41], we have $\overline{\mathrm{BDiff}_c(\mathbb{R}^2)} \simeq \mathrm{BDiff}_c^\delta(\mathbb{R}^2)$. Given that $\overline{\mathrm{B}S\Gamma_2}$ is 3–connected, by the Hurewicz theorem we have $\pi_4(\overline{\mathrm{B}S\Gamma_2}) = H_2(\Omega^2\overline{\mathrm{B}S\Gamma_2}; \mathbb{Z})$. Therefore, the exact sequence (3-15) is the same as the exact sequence in the theorem. □

Remark 3.16 It is easy to see that the Serre spectral sequence for the fibration

$$\overline{\mathrm{BDiff}(\Sigma_g)} \to \mathrm{BDiff}^\delta(\Sigma_g) \to \mathrm{BDiff}(\Sigma_g)$$

and the perfectness of the identity component, ${\mathrm{Diff}^\delta}_0(\Sigma_g)$, implies that the map

$$H_2(\mathrm{BDiff}^\delta(\Sigma_g)) \twoheadrightarrow H_2(\mathrm{BDiff}(\Sigma_g))$$

is surjective for all g, which means every surface bundle over a surface is cobordant to a flat surface bundle (see [16] for a more explicit construction of such cobordisms).

We derive a geometric consequence of (3-10) and (3-9) for flat surface bundles.

Theorem 3.17 *Let* $\mathrm{MSO}_k(X)$ *denote the oriented cobordism group of* k *dimensional manifolds equipped with a map to* X. *Then the map*

$$\mathrm{MSO}_3(\mathrm{BDiff}^\delta(\Sigma_g)) \twoheadrightarrow \mathrm{MSO}_3(\mathrm{BDiff}(\Sigma_g))$$

is surjective for $g \geq 6$. *In other words, every surface bundle of genus at least* 6 *over a* 3*–manifold is cobordant to a flat surface bundle.*

Proof Note that $\mathrm{MSO}_i(X) = H_i(X;\mathbb{Z})$ for $i = 3$. Thus, we only need to prove it for homology. In order to prove that the map

$$H_3(\mathrm{BDiff}^\delta(\Sigma_{g,k},\partial);\mathbb{Z}) \to H_3(\mathrm{BDiff}(\Sigma_{g,k},\partial);\mathbb{Z})$$

is surjective in the stable range, we need a little lemma:

Lemma 3.18 *Let* X *be a simply connected space; then we have the exact sequence*

$$\pi_3(X) \to H_3(X) \to H_3(\mathrm{K}(\pi_2(X),2);\mathbb{Z}),$$

where $\mathrm{K}(\pi_2(X),2)$ *is the Eilenberg–Mac Lane space whose second homotopy group is* $\pi_2(X)$.

The proof of the lemma is an easy Serre spectral sequence argument for the map, turned into a fibration,

$$X \to \mathrm{K}(\pi_2(X),2).$$

Recall that since $\pi_2(\Omega_0^\infty \mathrm{MTSO}(2)) = \mathbb{Z}$ (see [11]), we can deduce that

$$H_3(\mathrm{K}(\pi_2(\Omega_0^\infty \mathrm{MTSO}(2)),2);\mathbb{Z}) = 0.$$

If we apply Lemma 3.18 for $X = \Omega_0^\infty \mathrm{MTSO}(2)$, we obtain that every degree 3 homology class of $\Omega_0^\infty \mathrm{MTSO}(2)$ is spherical, ie

$$\pi_3(\Omega_0^\infty \mathrm{MTSO}(2)) \twoheadrightarrow H_3(\Omega_0^\infty \mathrm{MTSO}(2);\mathbb{Z}).$$

We have the following commutative diagram by naturality of Hurewicz maps:

$$\begin{tikzcd}
\pi_3(\Omega_0^\infty \mathrm{MT}\nu) \arrow[r] \arrow[d, two heads] & H_3(\Omega_0^\infty \mathrm{MT}\nu;\mathbb{Z}) \arrow[d] \\
\pi_3(\Omega_0^\infty \mathrm{MTSO}(2)) \arrow[r, two heads] & H_3(\Omega_0^\infty \mathrm{MTSO}(2);\mathbb{Z})
\end{tikzcd}$$

The left vertical map is surjective by comparing exact sequences of (3-8) and (3-7), so the right vertical map has to be surjective. □

3.2 Applications to characteristic classes of flat surface bundles

In this section, we study two types of characteristic classes for flat surface bundles. The first type is constructed by forgetting the flat structure on the total space of the bundle and considering it just as a surface bundle. The second type of class is secondary characteristic classes, or the so-called Godbillon–Vey classes of the codimension-2 foliation induced by a flat structure on the total space.

3.2.1 The MMM classes of the flat surface bundles Consider the following natural map again:

$$\iota\colon \mathrm{BDiff}^{\delta}(\Sigma_g) \to \mathrm{BDiff}(\Sigma_g).$$

The first type of characteristic classes of the flat surface bundles is the pullback of MMM classes via the map ι. Recall the definition of MMM classes is as follows. Let $\pi\colon E \to B$ be a surface bundle whose fibers are diffeomorphic to Σ_g. Let $T\pi$ denote the vertical tangent bundle and let $e(T\pi)$ denote its Euler class. Then the i^{th} MMM class is defined to be

$$\kappa_i(E \to B) := \pi_!(e(T\pi)^{i+1}) \in H^{2i}(B;\mathbb{Z}),$$

where $\pi_!$ is the push-forward map or the integration along the fiber, which is defined since the fibers are compact closed manifolds. We denote the i^{th} MMM class of the universal surface bundle over $\mathrm{BDiff}(\Sigma_g)$ by κ_i. Let κ_i^{δ} denote the pullback of κ_i via the map ι. One of the consequences of the Madsen–Weiss theorem and Harer stability is that the natural map

$$\mathbb{Q}[\kappa_1, \kappa_2, \dots] \to H^*(\mathrm{BDiff}(\Sigma_g);\mathbb{Q})$$

is injective in the stable range. However, Morita [30] observed that κ_i^{δ} vanishes in rational cohomology for $i > 2$, ie the map

$$\mathbb{Q}[\kappa_1^{\delta}, \kappa_2^{\delta}, \dots] \to H^*(\mathrm{BDiff}^{\delta}(\Sigma_g);\mathbb{Q})$$

sends κ_i^{δ} to zero if $i > 2$. For the above observation, it is essential to work with diffeomorphisms that are at least twice differentiable. It follows from Tsuboi's theorem [45], mentioned in the introduction, and the Madsen–Weiss theorem that the similar map for C^1–diffeomorphisms is in fact an isomorphism.

Morita and Kotschick [16] proved there exists a flat surface bundle over a surface with nonzero signature, hence they conclude that κ_1^{δ} does not vanish in $H^2(\mathrm{BDiff}^{\delta}(\Sigma_g);\mathbb{Q})$ for $g \geq 3$. First, let us give a homotopy-theoretic proof of their theorem using Theorem 2.6,

Proposition 3.19 *The class κ_1^{δ} does not vanish in $H^2(\mathrm{BDiff}^{\delta}(\Sigma_g);\mathbb{Q})$ for $g > 3$.*

Proof By Theorems 2.6 and 1.14, we know

$$H^2(\Omega^\infty_0 \mathrm{MT}\nu;\mathbb{Q}) \to H^2(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$$

is surjective for $g \geq 3$ and an isomorphism for $g > 3$. Thus, we need to show that the corresponding class in $H^2(\Omega^\infty_0 \mathrm{MT}\nu;\mathbb{Q})$ is nonzero. Consider the sequence

$$H^{2i+2}(\mathrm{B}\mathrm{S}\Gamma_2;\mathbb{Q}) \xrightarrow{\cong} H^{2i}(\mathrm{MT}\nu;\mathbb{Q}) \xrightarrow{\sigma^*} H^{2i}(\Omega^\infty_0 \mathrm{MT}\nu;\mathbb{Q}),$$

where the first map is the Thom isomorphism and the second map is induced by the suspension map. Let e denote the generator of $H^2(\mathrm{BSO}(2);\mathbb{Q})$. It is easy to see that the image of $\nu^*(e^{i+1}) \in H^{2i+2}(\mathrm{B}\mathrm{S}\Gamma_2;\mathbb{Q})$ in $H^{2i}(\Omega^\infty_0 \mathrm{MT}\nu;\mathbb{Q})$ is κ_i^δ (see [6, Theorem 3.1]). Since the map ν is at least 4–connected, $\nu^*(e^2)$ is not zero. Thus, to prove that κ_1^δ is nonzero, it is enough to show that the map

$$H_2(\Omega^\infty_0 \mathrm{MT}\nu;\mathbb{Q}) \xrightarrow{\sigma_*} H_2(\mathrm{MT}\nu;\mathbb{Q}) \xrightarrow{\kappa_1^\delta} \mathbb{Q}$$

is nontrivial. To prove that the suspension map is surjective on rational homology, let us consider the commutative diagram

$$\begin{tikzcd}
\pi_*(\Omega^\infty \mathrm{MT}\nu)\otimes\mathbb{Q} \arrow[r] \arrow[d] & \pi_*(\mathrm{MT}\nu)\otimes\mathbb{Q} \arrow[d] \\
H_*(\Omega^\infty \mathrm{MT}\nu;\mathbb{Q}) \arrow[r, two heads] & H_*(\mathrm{MT}\nu;\mathbb{Q})
\end{tikzcd}$$

The horizontal maps are induced by the suspension map and the vertical maps are induced by the Hurewicz map. The top horizontal map is an isomorphism by the definition of the homotopy groups of spectra and the right vertical map is also an isomorphism because of the rational Hurewicz theorem (see [39, Theorem 7.11]). Therefore, σ_*, the bottom horizontal map, is surjective, which implies $\kappa_1^\delta \circ \sigma_*$ is nontrivial. □

Remark 3.20 Morita and Kotschick [16], by a formal argument, showed that nontriviality of κ_1^δ implies that all its powers are nontrivial in the stable range. This can also be deduced from the fact that $H_*(\Omega^\infty \mathrm{MT}\nu;\mathbb{Q})$ is a Hopf algebra over $\mathbb{Q}$.

Regarding κ_2^δ, Morita and Kotschick asked the following problem:

Problem *Does κ_2^δ vanish in $H^4(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$?*

Toward answering this problem, we prove that it is equivalent to an open problem in foliation theory.

Theorem 3.21 *The MMM class κ_2 in $H^4(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ is nonzero for $g \geq 6$ if and only if a C^2–foliation $\mathcal{F}$ of codimension 2 on a 6–manifold exists such that $e(\nu(\mathcal{F}))^3 \neq 0$, where $\nu(\mathcal{F})$ is the normal bundle of the foliation $\mathcal{F}$.*

Proof Let $\mathcal{F}\Omega_{n,k}$ denote the group of foliated cobordism group of codimension-k foliations on n–dimensional manifolds. Every oriented codimension-k foliation on a manifold M gives a well-defined homotopy class of maps from M to $\mathrm{BS}\Gamma_k$ (see [10, Theorem 7]). Hence, we have the well-defined map

$$\mathcal{F}\Omega_{n,k} \to \mathrm{MSO}_n(\mathrm{BS}\Gamma_k).$$

By a result of [15, Theorem 1′], the map $\mathcal{F}\Omega_{6,2} \to \mathrm{MSO}_6(\mathrm{BS}\Gamma_2)$ is rationally bijective. By the Atiyah–Hirzebruch spectral sequence, $\mathrm{MSO}_4(\mathrm{BDiff}^\delta(\Sigma_g)) \to H_4(\mathrm{BDiff}^\delta(\Sigma_g))$ is surjective. Hence, in order to prove (non)triviality of κ_2^δ, we need to study (non)triviality of the map

$$\kappa_2^\delta\colon \mathrm{MSO}_4(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathbb{Q}.$$

By Theorem 2.6, we know

$$\mathrm{MSO}_4(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathrm{MSO}_4(\Omega_0^\infty \mathrm{MT}\nu)$$

is an isomorphism for $g > 6$ and surjective for $g \geq 6$. We have the commutative diagram

$$\begin{tikzcd}
\mathrm{MSO}_4(\Omega_0^\infty \mathrm{MT}\nu) \otimes \mathbb{Q} \arrow[r, "\kappa_2^\delta"] \arrow[d, "\sigma_*"] & \mathbb{Q} \\
\mathrm{MSO}_4(\mathrm{MT}\nu) \otimes \mathbb{Q} \arrow[r, "\cong"] & \mathrm{MSO}_6(\mathrm{BS}\Gamma_2) \otimes \mathbb{Q} \arrow[u, "e^3"']
\end{tikzcd}$$

where σ is the suspension map, the bottom map is the Thom isomorphism and the right vertical map is the map given by the cube of the Euler class of the codimension-2 Haefliger structures. Hence, if we show that the map σ_* is rationally surjective, the (non)triviality of κ_2^δ and e^3 become equivalent. We know that

$$H_4(\Omega_0^\infty \mathrm{MT}\nu;\mathbb{Q}) \to H_4(\mathrm{MT}\nu;\mathbb{Q})$$

is surjective by the same argument as in Proposition 3.19 and, since the Atiyah–Hirzebruch spectral sequence implies that $\mathrm{MSO}_4(X) = H_4(X) \oplus \mathbb{Z}$, we have

$$\sigma_*\colon \mathrm{MSO}_4(\Omega_0^\infty \mathrm{MT}\nu) \otimes \mathbb{Q} \twoheadrightarrow \mathrm{MSO}_4(\mathrm{MT}\nu) \otimes \mathbb{Q}. \qquad \square$$

Remark 3.22 One possible way to show that e^3 is nonzero in $H^6(\mathrm{BS}\Gamma_2;\mathbb{Q})$, which implies that a flat surface bundle with a nontrivial κ_2^δ exists, is to look at the Gysin sequence for the circle bundle

$$S^1 \to \overline{\mathrm{BS}\Gamma_2} \to \mathrm{BS}\Gamma_2.$$

Part of the Gysin sequence that is relevant to us is

$$H_6(\mathrm{BS}\Gamma_2;\mathbb{Q}) \xrightarrow{\cap e} H_4(\mathrm{BS}\Gamma_2;\mathbb{Q}) \xrightarrow{\tau} H_5(\overline{\mathrm{BS}\Gamma_2};\mathbb{Q}),$$

where τ is the transgression map for the circle bundle. If we show that a class $c \in H_4(\mathrm{BS}\Gamma_2;\mathbb{Q})$ that satisfies $e^2(c) \neq 0$ maps to zero via the transgression τ, then the Gysin sequence implies that e^3 has to be nonzero in $H^6(\mathrm{BS}\Gamma_2;\mathbb{Q})$. To construct a class c for which $e^2(c) \neq 0$, we consider a map $f\colon \mathbb{C}\mathrm{P}^2 \to \mathrm{BGL}_2^+(\mathbb{R})$ that is nontrivial on the fourth homology with rational coefficients. Since $\mathrm{BS}\Gamma_2 \to \mathrm{BGL}_2^+(\mathbb{R})$ is 4–connected (Remark 1.5), we can lift f to a map $g\colon \mathbb{C}\mathrm{P}^2 \to \mathrm{BS}\Gamma_2$. If we take the pullback of the circle bundle $\overline{\mathrm{BS}\Gamma_2} \to \mathrm{BS}\Gamma_2$ via the map g, we obtain the commutative diagram

$$\begin{tikzcd} S^5 \arrow[r, "\widehat{g}"] \arrow[d] & \overline{\mathrm{BS}\Gamma_2} \arrow[d] \\ \mathbb{C}\mathrm{P}^2 \arrow[r, "g"] & \mathrm{BS}\Gamma_2 \end{tikzcd}$$

The transgression maps the class $g([\mathbb{C}\mathrm{P}^2]) \in H_4(\mathrm{BS}\Gamma_2;\mathbb{Q})$ to the class $\widehat{g}([S^5]) \in H_5(\overline{\mathrm{BS}\Gamma_2};\mathbb{Q}) \cong \pi_5(\overline{\mathrm{BS}\Gamma_2}) \otimes \mathbb{Q}$. Since $\pi_5(\overline{\mathrm{BS}\Gamma_2}) = \pi_5(\mathrm{BS}\Gamma_2)$, we conclude that e^3 is nontrivial in $H^6(\mathrm{BS}\Gamma_2;\mathbb{Q})$ if

$$g_*\colon \pi_5(\mathbb{C}\mathrm{P}^2) \otimes \mathbb{Q} \to \pi_5(\mathrm{BS}\Gamma_2) \otimes \mathbb{Q}$$

is trivial. Note that since the image of g maps to zero in $H_5(\mathrm{BS}\Gamma_2;\mathbb{Q})$, the Godbillon–Vey classes vanish on the image of g, so it cannot be detected by the GV classes.

3.3 MMM classes as integral cohomology classes

The situation, however, is surprisingly different with integer coefficients. To study integral MMM classes, we first reduce them to classes with finite field coefficients. Consider the commutative diagram

$$\begin{tikzcd} H^*(\Omega_0^\infty \mathrm{MTSO}(2);\mathbb{F}_p) \arrow[r, "\Omega^\infty \boldsymbol{\nu}^*"] \arrow[d] & H^*(\Omega_0^\infty \mathrm{MT}\nu;\mathbb{F}_p) \arrow[d] \\ H^*(\mathrm{BDiff}(\Sigma_g);\mathbb{F}_p) \arrow[r, "\iota^*"] & H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{F}_p) \end{tikzcd}$$

where the vertical maps are isomorphisms in the stable range. Thus, to study ι^*, we need to study $\Omega^\infty \boldsymbol{\nu}^*$. In Theorem 2.12, we proved that $\Omega^\infty \boldsymbol{\nu}^*$ is injective on cohomology with $\mathbb{F}_p$ coefficients. Given the injectivity on $\mathbb{F}_p$–cohomology for all primes p, we shall see the map

$$H^*(\Omega_0^\infty \mathrm{MTSO}(2);\mathbb{Z}) \to H^*(\Omega_0^\infty \mathrm{MT}\nu;\mathbb{Z})$$

is injective. Hence, we summarize the situation with finite coefficients and with integer coefficients as follows:

Theorem 3.23 *For every prime p, the map*

$$H^*(\mathrm{BDiff}(\Sigma_g);\mathbb{F}_p) \xrightarrow{\iota^*} H^*(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{F}_p)$$

is injective in the stable range.

Theorem 3.24 *The induced map*

$$\mathbb{Z}[\kappa_1^\delta,\kappa_2^\delta,\dots] \hookrightarrow H^*(\mathrm{BDiff}^\delta(\Sigma_{\infty,1},\partial);\mathbb{Z})$$

is injective.

Proof We even prove a stronger result, that the induced map

$$\iota^*\colon H^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{Z}) \to H^*(\mathrm{BDiff}^\delta(\Sigma_{\infty,1},\partial);\mathbb{Z})$$

is injective. It follows from Theorem 2.12 that the map

$$\iota_p^*\colon H^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{F}_p) \to H^*(\mathrm{BDiff}^\delta(\Sigma_{\infty,1},\partial);\mathbb{F}_p),$$

is injective for all primes p. To prove that ι^* is injective, consider the induced map between the Bockstein exact sequences

$$\begin{tikzcd}
H^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{F}_p) \arrow[d, "\iota_p^*"] & H^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{Z}) \arrow[l] \arrow[d, "\iota^*"] & H^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{Z}) \arrow[l, "\times p"'] \arrow[d, "\iota^*"] \\
H^*(\mathrm{BDiff}^\delta(\Sigma_{\infty,1},\partial);\mathbb{F}_p) & H^*(\mathrm{BDiff}^\delta(\Sigma_{\infty,1},\partial);\mathbb{Z}) \arrow[l] & H^*(\mathrm{BDiff}^\delta(\Sigma_{\infty,1},\partial);\mathbb{Z}) \arrow[l, "\times p"']
\end{tikzcd}$$

Let $a \in \mathrm{Ker}(\iota^*)$; since ι_p^* is injective for all p, it follows from the above diagram that $a \in pH^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{Z})$ for all p. Since $H^*(\mathrm{BDiff}(\Sigma_{\infty,1},\partial);\mathbb{Z})$ is finitely generated by the Madsen–Weiss theorem [20], we deduce $a=0$. □

Remark 3.25 Akita, Kawazumi and Uemura [1] proved the algebraic independence of MMM classes by using finite cyclic subgroups of the mapping class groups. Their method can be also applied to prove the above theorem.

Remark 3.26 Recall that, as Morita [30] observed, the classes κ_i^δ for $i > 2$ in cohomology with $\mathbb{R}$ coefficients or even with $\mathbb{Q}$ coefficients vanish. This observation implies that there is a class in $H^{2i-1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R}/\mathbb{Z})$ for $i > 2$ which maps to κ_i^δ in the Bockstein exact sequence

$$H^{2i-1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R}/\mathbb{Z}) \to H^{2i}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z}) \to H^{2i}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R}).$$

Cheeger–Simons character theory helps us to find a canonical lift of κ_{2i+1}^δ for $i > 0$ in $H^{4i+1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R}/\mathbb{Z})$. In [4, Proposition 7.3], Cheeger and Simons showed that there are canonical classes, known as Pontryagin characters, $\widehat{p}_i \in H^{4i-1}(\mathrm{B}\mathrm{S}\Gamma_n;\mathbb{R}/\mathbb{Z})$ for $2i > n$ such that the image of $\widehat{p}_i$ under the Bockstein map

$$H^{4i-1}(\mathrm{B}\mathrm{S}\Gamma_n;\mathbb{R}/\mathbb{Z}) \xrightarrow{\beta} H^{4i}(\mathrm{B}\mathrm{S}\Gamma_n;\mathbb{Z})$$

is $-p_i$ of the normal bundle of the universal Haefliger structure. Let $\widehat{\kappa}_{2i+1}$ denote the image of the class $\widehat{p}_1^{i+1} \in H^{4i+3}(\mathrm{B}\mathrm{S}\Gamma_2;\mathbb{R}/\mathbb{Z})$ under the maps

$$H^{4i+3}(\mathrm{B}\mathrm{S}\Gamma_2;\mathbb{R}/\mathbb{Z}) \to H^{4i+1}(\mathrm{MT}\nu;\mathbb{R}/\mathbb{Z}) \to H^{4i+1}(\Omega_0^\infty\mathrm{MT}\nu;\mathbb{R}/\mathbb{Z}).$$

Hence, the class $\widehat{\kappa}_{2i+1}$ gives a canonical lift of κ_{2i+1}^δ if its degree lies in the stable range and $\beta(\widehat{\kappa}_{2i+1}) = -\kappa_{2i+1}^\delta$, where β is the Bockstein map in the sequence

$$H^{4i+1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R}/\mathbb{Z}) \xrightarrow{\beta} H^{4i+2}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z}).$$

We showed that κ_{2i+1}^δ is a nontorsion class, so β is a nontrivial map. Therefore, for $i > 0$ we obtain a regulator-type map for discrete surface diffeomorphisms,

$$\text{regulator map: } H_{4i+1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z}) \xrightarrow{\widehat{\kappa}_{2i+1}} \mathbb{R}/\mathbb{Z}.$$

Question *What is the cocycle formula for $\widehat{\kappa}_{2i+1}$?*

Nontriviality of this regulator map implies that $H_{4i+1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ in the stable range is not a trivial group, but in fact it follows easily from Theorem 3.24 that $H_{2i+1}(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ in the stable range with $i > 1$ is in fact uncountable (see [33, Theorem 6.4]).

3.3.1 Secondary classes of flat surface bundles In this section, we prove Theorem 0.18 from the introduction. Let us first recall a preliminary result about continuous variation of secondary characteristic classes of foliations of codimension 2. The cohomology of $\mathrm{B}\mathrm{S}\Gamma_n$ is not yet very well understood but it has been extensively studied via secondary characteristic classes of foliations, known as the Godbillon–Vey classes. For codimension-2 foliations there are two GV classes, h_1c_2, $h_1c_1^2 \in H^5(\mathrm{B}\mathrm{S}\Gamma_2;\mathbb{R})$

(for the definition of these classes look at [35; 18]). Rasmussen [38] proved that these two classes vary continuously and independently, ie

$$(h_1c_2, h_1c_1^2)\colon H_5(\mathrm{BS}\Gamma_2;\mathbb{Z}) \twoheadrightarrow \mathbb{R}^2.$$

If we take the universal flat surface bundle over $\mathrm{BDiff}^\delta(\Sigma_g)$, we can integrate h_1c_2 and $h_1c_1^2$ along the compact fibers, and we write their integration along the fiber as

$$\int_{\text{fiber}} h_1c_2,\ \int_{\text{fiber}} h_1c_1^2 \in H^3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{R}).$$

Morita [32, Problem 44] posed the question of whether the map induced by the above two cohomology classes from $H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Z})$ to $\mathbb{R}^2$ is surjective. Bowden [3] used a curious spectral sequence that only converges in low homological degrees to answer Morita's question affirmatively. Here, we simplify his proof using the Mather–Thurston theorem.

Theorem 3.27 (Bowden) *For any g, the induced map*

$$k = \left(\int_{\textit{fiber}} h_1c_2,\ \int_{\textit{fiber}} h_1c_1^2\right)\colon H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q}) \twoheadrightarrow \mathbb{R}^2$$

is surjective.

Proof Embed $\mathbb{R}^2$ as the interior of a small disk in Σ_g. The restriction of the map k to the embedded disk gives the commutative diagram

$$\begin{tikzcd} H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q}) \arrow[rr, "k"] & & \mathbb{R}^2 \\ H_3(\mathrm{BDiff}^\delta_c(\mathbb{R}^2);\mathbb{Q}) \arrow[u] \arrow[urr, "k'"'] & & \end{tikzcd}$$

where the map k' is also induced by the fiber integration of the GV classes along the embedded disk. If we show k' is surjective then it implies that k is also surjective. Let $\overline{\mathrm{BDiff}_c(\mathbb{R}^2)}$ be the homotopy fiber of the map

$$\mathrm{BDiff}^\delta_c(\mathbb{R}^2) \to \mathrm{BDiff}_c(\mathbb{R}^2).$$

But the topological group $\mathrm{Diff}_c(\mathbb{R}^2)$ is contractible [41], so $\overline{\mathrm{BDiff}_c(\mathbb{R}^2)} \simeq \mathrm{BDiff}^\delta_c(\mathbb{R}^2)$. By Thurston's theorem [43], we know that there is a map

$$\overline{\mathrm{BDiff}_c(\mathbb{R}^2)} \to \Omega^2\overline{\mathrm{BS}\Gamma}_2$$

that induces a homology isomorphism. By Remark 1.5, we know $\overline{\mathrm{B}\mathrm{S}\Gamma_2}$ is at least 3–connected. Therefore, we have

$$H_3(\mathrm{BDiff}_c^{\delta}(\mathbb{R}^2);\mathbb{Q}) \xrightarrow{\cong} H_3(\Omega^2\overline{\mathrm{B}\mathrm{S}\Gamma_2};\mathbb{Q}) \twoheadrightarrow H_5(\overline{\mathrm{B}\mathrm{S}\Gamma_2};\mathbb{Q}) \twoheadrightarrow \mathbb{R}^2. \tag{3-28}$$

The first map is an isomorphism by Thurston's theorem. The second map is the suspension map and, because $\overline{\mathrm{B}\mathrm{S}\Gamma_2}$ is at least 3–connected, the rational Hurewicz theorem implies that in the diagram

$$\begin{tikzcd}
\pi_3(\Omega^2\overline{\mathrm{B}\mathrm{S}\Gamma_2})\otimes\mathbb{Q} \arrow[r] \arrow[d, "\cong"] & H_3(\Omega^2\overline{\mathrm{B}\mathrm{S}\Gamma_2};\mathbb{Q}) \arrow[d] \\
\pi_5(\overline{\mathrm{B}\mathrm{S}\Gamma_2})\otimes\mathbb{Q} \arrow[r, two heads] & H_5(\overline{\mathrm{B}\mathrm{S}\Gamma_2};\mathbb{Q})
\end{tikzcd}$$

the bottom map is surjective, hence so is the right vertical map. The third map in (3-28) is given by the Godbillon–Vey classes

$$\left(\int h_1c_2, \int h_1c_1^2\right)\colon H_5(\overline{\mathrm{B}\mathrm{S}\Gamma_2};\mathbb{Q}) \to \mathbb{R}^2,$$

which is surjective as a corollary of the theorem of Rasmussen [38]. □

Corollary 3.29 *There exists a surjective map*

$$H_{3k}(\mathrm{BDiff}^{\delta}(\Sigma_g);\mathbb{Q}) \twoheadrightarrow \textstyle\bigwedge_{\mathbb{Q}}^k \mathbb{R}^2$$

provided $k \le \frac{1}{9}(2g-2)$, *where* $\bigwedge_{\mathbb{Q}}^k \mathbb{R}^2$ *is the* k^{th} *exterior power of* $\mathbb{R}^2$ *as a vector space over* $\mathbb{Q}$.

Proof For $k \le \frac{1}{9}(2g-2)$, using Theorem 2.6 we know that

$$H_{3k}(\mathrm{BDiff}^{\delta}(\Sigma_g);\mathbb{Q}) \simeq H_{3k}(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q}).$$

By Theorem 3.27, we obtain a surjective map

$$H_3(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q}) \twoheadrightarrow \mathbb{R}^2.$$

Note that $H_*(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q})$ is a simply connected Hopf algebra over $\mathbb{Q}$. Hence, elements in $H_3(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q})$ are primitive in the Hopf algebra. If we choose a basis for the vector space $H_3(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q})$, their exterior powers are nontrivial and provide us with a map

$$H_{3k}(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q}) \twoheadrightarrow \textstyle\bigwedge_{\mathbb{Q}}^k \mathbb{R}^2,$$

which is surjective. □

Recall from [13] that $H_3(\mathrm{BDiff}(\Sigma_g);\mathbb{Q}) = 0$, but, as we showed in the proof of Theorem 3.27, secondary characteristic classes in $H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ vary continuously and independently on diffeomorphisms of Σ_g that are only supported in a disk. Bowden asked the author if there is a nontrivial class in $H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ that cannot be detected by an embedding of a disk. To give an answer to his question, we prove that, at least in the stable range, all classes in $H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$ are essentially supported in a disk; more precisely:

Theorem 3.30 *Let $\mathbb{R}^2 \hookrightarrow \Sigma_g$ be an embedding of an open disk into the surface Σ_g. For $g \geq 6$, the induced map*

$$H_3(\mathrm{BDiff}^\delta_c(\mathbb{R}^2);\mathbb{Q}) \longrightarrow H_3(\mathrm{BDiff}^\delta(\Sigma_g);\mathbb{Q})$$

is surjective.

Proof Recall that since the topological group $\mathrm{Diff}_c(\mathbb{R}^2)$ is contractible (see [41]), we have $\overline{\mathrm{BDiff}_c(\mathbb{R}^2)} \simeq \mathrm{BDiff}^\delta_c(\mathbb{R}^2)$. Hence, all the $\mathbb{R}^2$–bundles trivialized at the infinity over $\mathrm{BDiff}^\delta_c(\mathbb{R}^2)$ are topologically trivial bundle, therefore Pontryagin–Thom theory for the trivial bundle $\mathrm{BDiff}^\delta_c(\mathbb{R}^2)\times\mathbb{R}^2 \to \mathrm{BDiff}^\delta_c(\mathbb{R}^2)$, as we shall explain in Lemma 3.35, provides a map

$$\beta\colon \mathrm{BDiff}^\delta_c(\mathbb{R}^2) \to \Omega^2\overline{\mathrm{B}S\Gamma}_2,$$

which is a homology isomorphism by the theorem of Thurston [43]. We showed that there exists a Madsen–Weiss-type map

$$\alpha_{\Sigma_g}\colon \mathrm{BDiff}^\delta(\Sigma_g) \to \Omega^\infty_0\mathrm{MT}\nu$$

for the universal flat Σ_g–bundle over $\mathrm{BDiff}^\delta(\Sigma_g)$, which induces a homology isomorphism in the stable range. In Lemma 3.35 below, we shall prove that there exists a homotopy commutative diagram

$$\begin{tikzcd}
 & \mathrm{BDiff}^\delta_c(\mathbb{R}^2) \arrow[r] \arrow[dl, "\beta"'] \arrow[d, "\alpha_{\mathbb{R}^2}"] & \mathrm{BDiff}^\delta(\Sigma_g) \arrow[d, "\alpha_{\Sigma_g}"] \\
\Omega^2\overline{\mathrm{B}S\Gamma}_2 \arrow[r, "\lambda"] & Q_0S^{-2}(\overline{\mathrm{B}S\Gamma}_{2+}) \arrow[r] & \Omega^\infty_0\mathrm{MT}\nu
\end{tikzcd} \tag{3-31}$$

where $Q = \Omega^\infty\Sigma^\infty$ and the subscript 0 means the basepoint component. The maps $\alpha_{\mathbb{R}^2}$ and α_{Σ_g} are defined as in Remark 2.10 and the map λ is the natural composition

$$\Omega^2\overline{\mathrm{B}S\Gamma}_2 \to Q(\Omega^2\overline{\mathrm{B}S\Gamma}_2) \to QS^{-2}(\overline{\mathrm{B}S\Gamma}_{2+}).$$

Thus, to prove the theorem, we need to show that the induced map

$$H_3(\Omega^2\overline{\mathrm{B}S\Gamma}_2;\mathbb{Q}) \to H_3(\Omega^\infty_0\mathrm{MT}\nu;\mathbb{Q})$$

is surjective. Consider the commutative diagram

$$\begin{tikzcd} \pi_3(\Omega^2\overline{\mathrm{BS\Gamma}_2})\otimes\mathbb{Q} \arrow[r, "g"] \arrow[d] & \pi_3(\Omega^\infty_0\mathrm{MT}\nu)\otimes\mathbb{Q} \arrow[d, two heads] \\ H_3(\Omega^2\overline{\mathrm{BS\Gamma}_2};\mathbb{Q}) \arrow[r, "f"] & H_3(\Omega^\infty_0\mathrm{MT}\nu;\mathbb{Q}) \end{tikzcd} \tag{3-32}$$

Because we have $H_1(\Omega^\infty_0\mathrm{MT}\nu;\mathbb{Q}) = H_3(\mathrm{BS\Gamma}_2;\mathbb{Q}) = 0$, the rational Hurewicz theorem implies that the right vertical map is surjective. Therefore, to prove that the map f in (3-32) is surjective, it is sufficient to prove that the map g is surjective.

Recall from Theorem 3.3 that we have the fibration sequence

$$QS^{-2}(\overline{\mathrm{BS\Gamma}_2}_+) \xrightarrow{h} \Omega^\infty\mathrm{MT}\nu \to Q(\mathrm{BS\Gamma}_{2+}),$$

and its long exact sequence of homotopy groups in (3-7) implies that the map h induces a surjection between third rational homotopy groups

$$\pi_5^s(\overline{\mathrm{BS\Gamma}_2})\otimes\mathbb{Q} \xrightarrow{h_*} \pi_3(\Omega^\infty_0\mathrm{MT}\nu)\otimes\mathbb{Q}. \tag{3-33}$$

As for the surjectivity of the map g, we observed in the diagram (3-31) that the map g is induced by the composition

$$\begin{aligned} \pi_3(\Omega^2\overline{\mathrm{BS\Gamma}_2})\otimes\mathbb{Q} = \pi_5(\overline{\mathrm{BS\Gamma}_2})\otimes\mathbb{Q} &\xrightarrow{\lambda_*} \pi_3(QS^{-2}(\overline{\mathrm{BS\Gamma}_2}_+))\otimes\mathbb{Q} \\ &\xrightarrow{h_*} \pi_3(\Omega^\infty_0\mathrm{MT}\nu)\otimes\mathbb{Q}. \end{aligned} \tag{3-34}$$

Note that $\pi_3(QS^{-2}(\overline{\mathrm{BS\Gamma}_2}_+)) = \pi_5^s(\overline{\mathrm{BS\Gamma}_2})\oplus\pi_5^s$ and, since $\pi_5^s = 0$, the map g is surjective if in (3-34) the map λ_* is surjective. Recall $\overline{\mathrm{BS\Gamma}_2}$ is 3–connected (Remark 1.5); therefore the rational Hurewicz theorem implies that the Hurewicz map

$$\pi_5(\overline{\mathrm{BS\Gamma}_2})\otimes\mathbb{Q} \xrightarrow{\cong} \pi_5^s(\overline{\mathrm{BS\Gamma}_2})\otimes\mathbb{Q},$$

is an isomorphism. Hence $g = h_*\circ\lambda_*$ is surjective. □

Lemma 3.35 *The diagram (3-31) is homotopy commutative.*

Proof Since the group $\mathrm{Diff}^\delta_c(\mathbb{R}^2)$ acts on the surface Σ_g via the embedding $\mathbb{R}^2\hookrightarrow\Sigma_g$, we obtain a map between Borel constructions

$$\mathbb{R}^2/\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2) \to \Sigma_g/\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2).$$

On the other hand, $\Sigma_g/\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2)$ is a flat surface bundle induced by the pullback of the universal flat surface bundle over $\mathrm{BDiff}^\delta(\Sigma_g)$ via the map $\mathrm{BDiff}^\delta_c(\mathbb{R}^2)\to\mathrm{BDiff}^\delta(\Sigma_g)$.

Therefore, we have the homotopy commutative diagram

$$\begin{tikzcd}
\mathbb{R}^2/\!\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2) \arrow[r] \arrow[dr, "\pi'"'] & \Sigma_g/\!\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2) \arrow[r] \arrow[d] & \Sigma_g/\!\!/\mathrm{Diff}^\delta(\Sigma_g) \arrow[d, "\pi"'] \\
& \mathrm{BDiff}^\delta_c(\mathbb{R}^2) \arrow[r] & \mathrm{BDiff}^\delta(\Sigma_g)
\end{tikzcd}$$

By the naturality of the pretransfer construction in Remark 2.10, we have a commutative diagram of spectra

$$\begin{tikzcd}
\Sigma^\infty(\mathrm{BDiff}^\delta_c(\mathbb{R}^2)_+) \arrow[r] \arrow[d] & \mathbf{Th}(-T\pi') \arrow[d] \\
\Sigma^\infty(\mathrm{BDiff}^\delta(\Sigma_g)_+) \arrow[r] & \mathbf{Th}(-T\pi)
\end{tikzcd} \tag{3-36}$$

Note that, by the flatness of the bundles, the classifying maps for $T\pi$ and $T\pi'$ lift to $\mathrm{BS}\Gamma_2$. But, since $\mathrm{BDiff}_c(\mathbb{R}^2)$ is contractible, topologically the bundle $\mathbb{R}^2/\!\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2)$ is trivial, which implies that the bundle $T\pi'$ is trivial, hence the classifying map for $T\pi'$ further lifts to $\overline{\mathrm{BS}\Gamma_2}$. Thus, we have a commutative diagram

$$\begin{tikzcd}
\mathbf{Th}(-T\pi') \arrow[r] \arrow[d] & \mathbf{Th}(-\underline{\mathbb{R}}^2 \to \overline{\mathrm{BS}\Gamma_2}) \arrow[d] \\
\mathbf{Th}(-T\pi) \arrow[r] & \mathbf{Th}(-\nu^*(\gamma))
\end{tikzcd} \tag{3-37}$$

where $\underline{\mathbb{R}}^2$ denotes the trivial 2–dimensional vector bundle over $\overline{\mathrm{BS}\Gamma_2}$. From the diagrams (3-36) and (3-37), we obtain a homotopy commutative diagram

$$\begin{tikzcd}
\mathrm{BDiff}^\delta_c(\mathbb{R}^2) \arrow[r, "\alpha_{\mathbb{R}^2}"] \arrow[d] & Q_0 S^{-2}(\overline{\mathrm{BS}\Gamma_2}_+) \arrow[d] \\
\mathrm{BDiff}^\delta(\Sigma_g) \arrow[r, "\alpha_{\Sigma_g}"] & \Omega^\infty_0 \mathrm{MT}\nu
\end{tikzcd} \tag{3-38}$$

We are left to show that the map $\alpha_{\mathbb{R}^2}$ equals $\lambda \circ \beta$ up to homotopy. Recall that $\mathbb{R}^2/\!\!/\mathrm{Diff}^\delta_c(\mathbb{R}^2) \simeq \mathbb{R}^2 \times \mathrm{BDiff}^\delta_c(\mathbb{R}^2)$ is topologically trivial bundle and the flatness (transverse foliation on the total space) of the bundle gives rise to a map

$$f\colon \mathbb{R}^2 \times \mathrm{BDiff}^\delta_c(\mathbb{R}^2) \to \overline{\mathrm{BS}\Gamma_2},$$

which classifies the vertical tangent bundle $T\pi'$. Since the foliation on $\mathbb{R}^2 \times \mathrm{BDiff}^\delta_c(\mathbb{R}^2)$ is trivial outside of a compact set of the fiber, the map f factors through the map

$\mathbb{R}^2 \times \mathrm{BDiff}_c^\delta(\mathbb{R}^2) \to \Sigma^2(\mathrm{BDiff}_c^\delta(\mathbb{R}^2)_+)$. Therefore, the spectrum map

$$\Sigma^\infty(\mathrm{BDiff}_c^\delta(\mathbb{R}^2)_+) \to \mathbf{Th}(-\underline{\mathbb{R}}^2 \to \mathbb{R}^2 \times \mathrm{BDiff}_c^\delta(\mathbb{R}^2)) \to \mathbf{Th}(-\underline{\mathbb{R}}^2 \to \overline{\mathrm{BS}\Gamma_2}),$$

whose adjoint is $\alpha_{\mathbb{R}^2}$, factors as

$$\Sigma^\infty(\mathrm{BDiff}_c^\delta(\mathbb{R}^2)_+) \xrightarrow{\mathrm{id}} \Sigma^{\infty-2}\Sigma^2(\mathrm{BDiff}_c^\delta(\mathbb{R}^2)_+) \to \mathbf{Th}(-\underline{\mathbb{R}}^2 \to \overline{\mathrm{BS}\Gamma_2}).$$

Using Ω^∞–Σ^∞–adjointness, the map

$$\alpha_{\mathbb{R}^2}\colon \mathrm{BDiff}_c^\delta(\mathbb{R}^2) \to Q(\mathrm{BDiff}_c^\delta(\mathbb{R}^2)_+) \to QS^{-2}(\overline{\mathrm{BS}\Gamma}_{2+}),$$

can be factored as

$$\begin{tikzcd} \mathrm{BDiff}_c^\delta(\mathbb{R}^2) \arrow[r] \arrow[d, "\beta"'] & Q(\mathrm{BDiff}_c^\delta(\mathbb{R}^2)_+) \arrow[d] \\ \Omega^2\overline{\mathrm{BS}\Gamma_2} \arrow[r, "\lambda"] & QS^{-2}(\overline{\mathrm{BS}\Gamma}_{2+}) \end{tikzcd} \tag{3-39}$$

Hence, $\alpha_{\mathbb{R}^2} \simeq \lambda \circ \beta$. □

Using the same idea, we will show that, up to torsion, every codimension-2 foliation on a manifold of dimension 5 is foliated cobordant to a flat surface bundle of genus higher than 5.

Any flat Σ_g–bundle $[\Sigma_g \to E^{n+2} \xrightarrow{\pi} B^n]$ over an n–manifold B^n gives a codimension-2 foliation on E^{n+2}. We let e_n be the map that assigns to a flat surface the foliated cobordism class of the codimension-2 foliation on the total space. Hence, e_n induces a well-defined map

$$e_n\colon \mathrm{MSO}_n(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathcal{F}\Omega_{n+2,2}.$$

Let $E^5 \xrightarrow{\pi} B^3$ be a flat Σ_g–bundle over a 3–dimensional manifold B^3; then characteristic classes $h_1.c_2, h_1.c_1^2$ of the codimension 2 foliation on E^5 live in $H^5(E^5;\mathbb{R})$, hence

$$x = \left\langle \int_{\text{fiber}} h_1.c_1^2, [B] \right\rangle, \quad y = \left\langle \int_{\text{fiber}} h_1.c_2, [B] \right\rangle$$

are real characteristic numbers associated to $E^5 \xrightarrow{\pi} B^3$. Consider the diagram

$$\begin{tikzcd} & \mathcal{F}\Omega_{5,2} \arrow[dr, "(\int h_1c_2, \int h_1c_1^2)"] & \\ \mathrm{MSO}_3\big(\coprod_g \mathrm{BDiff}^\delta(\Sigma_{g,1},\partial)\big) \arrow[ur, "e_3"] \arrow[rr, "{(x,y)}"] & & \mathbb{R}^2 \end{tikzcd}$$

Since the classes x and y are invariants of the foliated cobordism class, it is easy to see that the map induced by (x, y) factors through $\mathcal{F}\Omega_{5,2}$. The surjectivity of the map induced by the integration of h_1c_2 and $h_1c_1^2$ is the statement of Rasmussen's theorem. In Theorem 3.27, we showed that the map induced by (x, y) is rationally surjective. Now we prove that in fact e_3 is also rationally surjective.

Theorem 3.40 *The map*

$$e_3\colon \mathrm{MSO}_3(\mathrm{BDiff}^{\delta}(\Sigma_g)) \to \mathcal{F}\Omega_{5,2}$$

is rationally surjective if $g \geq 5$, and is a rational isomorphism if $g \geq 6$.

Proof Recall that, by a result of [15, Theorem 1′], the map

$$\mathcal{F}\Omega_{k,2} \to \mathrm{MSO}_k(\mathrm{BS}\Gamma_2)$$

is rationally bijective for $k > 2$. Furthermore, by Theorems 2.6 and 1.14, we know that, for $n \leq \frac{1}{3}(2g-2)$, the map

$$\mathrm{MSO}_n(\mathrm{BDiff}^{\delta}(\Sigma_g)) \to \mathrm{MSO}_n(\Omega_0^{\infty}\mathrm{MT}\nu)$$

is bijective and, for $n \leq \frac{2}{3}g$, it is surjective. Consider the commutative diagram

$$\begin{array}{ccc} \mathrm{MSO}_3(\mathrm{BDiff}^{\delta}(\Sigma_g)) & \xrightarrow{e_3} & \mathcal{F}\Omega_{5,2} \\ \downarrow & & \downarrow \\ \mathrm{MSO}_3(\Omega_0^{\infty}\mathrm{MT}\nu) & \longrightarrow & \mathrm{MSO}_4(\mathrm{BS}\Gamma_2) \end{array} \tag{3-41}$$

Therefore, the statement of the theorem is equivalent to proving that the bottom map in (3-41) is rationally an isomorphism. Note that the bottom map is given by composing the suspension map and the Thom isomorphism

$$\mathrm{MSO}_3(\Omega_0^{\infty}\mathrm{MT}\nu) \xrightarrow{\sigma} \mathrm{MSO}_3(\mathrm{MT}\nu) \xrightarrow{\text{Thom iso}} \mathrm{MSO}_5(\mathrm{BS}\Gamma_2).$$

Recall that $\mathrm{MSO}_*(X) \otimes \mathbb{Q} \cong H_*(X;\mathbb{Q}) \otimes \mathrm{MSO}_*(\mathrm{pt})$ for any topological space X. Since both $\mathrm{BS}\Gamma_2$ and $\Omega_0^{\infty}\mathrm{MT}\nu$ are simply connected, one can easily see that

$$\mathrm{MSO}_3(\Omega_0^{\infty}\mathrm{MT}\nu) \otimes \mathbb{Q} \xrightarrow{\cong} H_3(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q}),$$
$$\mathrm{MSO}_5(\mathrm{BS}\Gamma_2) \otimes \mathbb{Q} \xrightarrow{\cong} H_5(\mathrm{BS}\Gamma_2;\mathbb{Q}).$$

Therefore, we need to show that the natural map

$$H_3(\Omega_0^{\infty}\mathrm{MT}\nu;\mathbb{Q}) \to H_3(\mathrm{MT}\nu;\mathbb{Q}),$$

is an isomorphism. We know from the proof of Proposition 3.19 that the above is surjective. To prove that it is also injective consider the commutative diagram

$$\begin{tikzcd}
\pi_3(\Omega_0^\infty \mathrm{MT}\nu)\otimes\mathbb{Q} \arrow[r] \arrow[d] & \pi_3(\mathrm{MT}\nu)\otimes\mathbb{Q} \arrow[d] \\
H_3(\Omega_0^\infty \mathrm{MT}\nu;\mathbb{Q}) \arrow[r, two heads] & H_3(\mathrm{MT}\nu;\mathbb{Q})
\end{tikzcd}$$

As we observed in the proof of Proposition 3.19, the composition of the top horizontal map and the right vertical map is an isomorphism. Given that the left vertical map is surjective by the rational Hurewicz theorem, the bottom horizontal map must be an isomorphism. □

Unlike e_3, the map

$$e_2\colon \mathrm{MSO}_2(\mathrm{BDiff}^\delta(\Sigma_g)) \to \mathcal{F}\Omega_{4,2}$$

is not rationally surjective. To every codimension-2 foliation $\mathcal{F}$ on a 4–manifold M, we can assign the difference of the Pontryagin classes $p_1(M)-p_1(\nu(\mathcal{F}))$, where $\nu(\mathcal{F})$ is the normal bundle of $\mathcal{F}$. It is easy to see that the number $\int_M p_1(M)-p_1(\nu(\mathcal{F}))$ is an invariant of the foliated cobordism class of $\mathcal{F}$, hence it induces a map

$$\phi\colon \mathcal{F}\Omega_{4,2}\otimes\mathbb{Q}\to\mathbb{Q}.$$

Suppose we have a flat surface bundle $\Sigma_g \to M \xrightarrow{\pi} \Sigma_h$; then the normal bundle of the codimension-2 foliation on M is the vertical tangent bundle $T\pi$. It is easy to see that $p_1(M)=p_1(T\pi)=p_1(\nu(\mathcal{F}))$ (see [31, Proposition 4.11]). Hence ϕ vanishes on flat surface bundles. We prove below that the vanishing of ϕ is essentially the only obstruction for a codimension-2 foliation $\mathcal{F}$ on a 4–manifold M to be foliated cobordant to a flat surface bundle.

Theorem 3.42 *For $g \geq 4$, there is a short exact sequence*

$$0\to \mathrm{MSO}_2(\mathrm{BDiff}^\delta(\Sigma_g))\otimes\mathbb{Q} \xrightarrow{e_2} \mathcal{F}\Omega_{4,2}\otimes\mathbb{Q} \xrightarrow{\phi} \mathbb{Q}\to 0. \tag{3-43}$$

Proof By [15, Theorem 1′], we know that

$$\mathcal{F}\Omega_{4,2}\otimes\mathbb{Q}\xrightarrow{\cong}\mathrm{MSO}_4(\mathrm{B}\mathrm{S}\Gamma_2)\otimes\mathbb{Q}.$$

To show that ϕ is surjective, we just need to find a 4–manifold M and a map $f\colon M\to \mathrm{BS}\Gamma_2$ such that ϕ does not vanish on $[M,f]\in\mathrm{MSO}_4(\mathrm{BS}\Gamma_2)$. Let M be $\mathbb{C}\mathrm{P}^2$, which is a 4–manifold whose signature is not zero, and let f be a nullhomotopic map.

Since f is trivial the normal bundle of the Haefliger structure $\mathcal{H}$ induced by f is trivial. Thus, $p_1(\nu(\mathcal{H}))=0$. Hence, we have

$$\phi([\mathbb{C}\mathrm{P}^2, f]) = p_1(\mathbb{C}\mathrm{P}^2) - p_1(\nu(\mathcal{H})) = 3 \neq 0.$$

To prove the injectivity of e_2, note that we have

$$\mathrm{MSO}_2(\Omega^\infty_0 \mathrm{MT}\nu)\otimes\mathbb{Q} \xrightarrow{\cong} H_2(\Omega^\infty_0\mathrm{MT}\nu;\mathbb{Q}) \xrightarrow{\cong} H_2(\mathrm{MT}\nu;\mathbb{Q}) \xrightarrow{\cong} H_4(\mathrm{BS}\Gamma_2;\mathbb{Q}),$$

where the second isomorphism is given by the Hurewicz theorem, as $\pi_1(\Omega^\infty_0\mathrm{MT}\nu)=0$. Hence, using Theorems 2.6 and 1.14, for $g\geq 4$, the map

$$\mathrm{MSO}_2(\mathrm{BDiff}^\delta(\Sigma_g))\otimes\mathbb{Q}\to H_4(\mathrm{BS}\Gamma_2;\mathbb{Q})$$

is an isomorphism. On the other hand, by the Atiyah–Hirzebruch spectral sequence, we have a short exact sequence

$$0\to\mathbb{Q}\to\mathrm{MSO}_4(\mathrm{BS}\Gamma_2)\otimes\mathbb{Q}\to H_4(\mathrm{BS}\Gamma_2;\mathbb{Q})\to 0.$$

Therefore, we have a commutative diagram

$$\begin{tikzcd}
\mathrm{MSO}_2(\mathrm{BDiff}^\delta(\Sigma_g))\otimes\mathbb{Q} \arrow[r, "e_2"] \arrow[d, "\cong"] & \mathcal{F}\Omega_{4,2}\otimes\mathbb{Q} \arrow[d, "\cong"] \\
H_4(\mathrm{BS}\Gamma_2;\mathbb{Q}) & \mathrm{MSO}_4(\mathrm{BS}\Gamma_2)\otimes\mathbb{Q} \arrow[l]
\end{tikzcd}$$

Hence, the map e_2 is injective with cokernel $\mathbb{Q}$. Since ϕ vanishes on the image of e_2, the exactness in the middle term of (3-43) is also readily implied. □

Remark 3.44 Jonathan Bowden pointed out to the author that in fact there is an example of a codimension-2 foliation (not just $\mathrm{S}\Gamma_2$–structure) which is not in the image of e_2. To include his example, let $\mathcal{F}$ be the foliation by fibers of a surface bundle over a surface whose signature is nonzero. It is easy to see that $\phi(\mathcal{F})\neq 0$, hence $\mathcal{F}$ is not in the image of e_2.

References

[1] **T Akita**, **N Kawazumi**, **T Uemura**, *Periodic surface automorphisms and algebraic independence of Morita–Mumford classes*, J. Pure Appl. Algebra 160 (2001) 1–11 MR

[2] **R Bott**, *Lectures on characteristic classes and foliations*, from "Lectures on algebraic and differential topology" (S Gitler, editor), Lecture Notes in Math. 279, Springer (1972) 1–94 MR

[3] **J Bowden**, *The homology of surface diffeomorphism groups and a question of Morita*, Proc. Amer. Math. Soc. 140 (2012) 2543–2549 MR

[4] **J Cheeger**, **J Simons**, *Differential characters and geometric invariants*, from "Geometry and topology" (J Alexander, J Harer, editors), Lecture Notes in Math. 1167, Springer (1985) 50–80 MR

[5] **S Galatius**, *Mod p homology of the stable mapping class group*, Topology 43 (2004) 1105–1132 MR

[6] **S Galatius**, **I Madsen**, **U Tillmann**, *Divisibility of the stable Miller–Morita–Mumford classes*, J. Amer. Math. Soc. 19 (2006) 759–779 MR

[7] **S Galatius**, **O Randal-Williams**, *Monoids of moduli spaces of manifolds*, Geom. Topol. 14 (2010) 1243–1302 MR

[8] **S Galatius**, **O Randal-Williams**, *Stable moduli spaces of high-dimensional manifolds*, Acta Math. 212 (2014) 257–377 MR

[9] **S Galatius**, **U Tillmann**, **I Madsen**, **M Weiss**, *The homotopy type of the cobordism category*, Acta Math. 202 (2009) 195–239 MR

[10] **A Haefliger**, *Homotopy and integrability*, from "Manifolds" (N H Kuiper, editor), Lecture Notes in Math. 197, Springer (1971) 133–163 MR

[11] **J Harer**, *The second homology group of the mapping class group of an orientable surface*, Invent. Math. 72 (1983) 221–239 MR

[12] **J L Harer**, *Stability of the homology of the mapping class groups of orientable surfaces*, Ann. of Math. 121 (1985) 215–249 MR

[13] **J Harer**, *The third homology group of the moduli space of curves*, Duke Math. J. 63 (1991) 25–55 MR

[14] **S Hurder**, *Foliation geometry/topology problem set*, preprint (2003) `http://homepages.math.uic.edu/~hurder/papers/58manuscript.pdf`

[15] **U Koschorke**, *Singularities and bordism of q–plane fields and of foliations*, Bull. Amer. Math. Soc. 80 (1974) 760–765 MR

[16] **D Kotschick**, **S Morita**, *Signatures of foliated surface bundles and the symplectomorphism groups of surfaces*, Topology 44 (2005) 131–149 MR

[17] **D Kotschick**, **S Morita**, *Characteristic classes of foliated surface bundles with area-preserving holonomy*, J. Differential Geom. 75 (2007) 273–302 MR

[18] **H B Lawson, Jr**, *The quantitative theory of foliations*, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics 27, Amer. Math. Soc., Providence, RI (1977) MR

[19] **I Madsen**, **U Tillmann**, *The stable mapping class group and $Q(\mathbb{C}\mathrm{P}^\infty_+)$*, Invent. Math. 145 (2001) 509–544 MR

[20] **I Madsen**, **M Weiss**, *The stable moduli space of Riemann surfaces: Mumford's conjecture*, Ann. of Math. 165 (2007) 843–941 MR

[21] **J N Mather**, *Commutators of diffeomorphisms*, Comment. Math. Helv. 49 (1974) 512–528 MR

[22] **J N Mather**, *On the homology of Haefliger's classifying space*, from "Differential topology" (V Villani, editor), CIME Summer Schools 73, Springer (2011) 71–116

[23] **J P May**, **K Ponto**, *More concise algebraic topology: localization, completion, and model categories*, University of Chicago Press (2012) MR

[24] **D McDuff**, *Local homology of groups of volume-preserving diffeomorphisms, III*, Ann. Sci. École Norm. Sup. 16 (1983) 529–540 MR

[25] **J Milnor**, *The geometric realization of a semi-simplicial complex*, Ann. of Math. 65 (1957) 357–362 MR

[26] **J Milnor**, *Microbundles, I*, Topology 3 (1964) 53–80 MR

[27] **J Milnor**, *On the homology of Lie groups made discrete*, Comment. Math. Helv. 58 (1983) 72–85 MR

[28] **T Mizutani**, **S Morita**, **T Tsuboi**, *On the cobordism classes of codimension one foliations which are almost without holonomy*, Topology 22 (1983) 325–343 MR

[29] **S Morita**, *Discontinuous invariants of foliations*, from "Foliations" (I Tamura, editor), Adv. Stud. Pure Math. 5, North-Holland, Amsterdam (1985) 169–193 MR

[30] **S Morita**, *Characteristic classes of surface bundles*, Invent. Math. 90 (1987) 551–577 MR

[31] **S Morita**, *Geometry of characteristic classes*, Translations of Mathematical Monographs 199, Amer. Math. Soc., Providence, RI (2001) MR

[32] **S Morita**, *Cohomological structure of the mapping class group and beyond*, from "Problems on mapping class groups and related topics" (B Farb, editor), Proc. Sympos. Pure Math. 74, Amer. Math. Soc. (2006) 329–354 MR

[33] **S Nariman**, *Homological stability and stable moduli of flat manifold bundles*, preprint (2014) arXiv

[34] **S Nariman**, *On the moduli space of flat symplectic surface bundles*, preprint (2016) arXiv

[35] **H V Pittie**, *Characteristic classes of foliations*, Research Notes in Mathematics 10, Pitman, London (1976) MR

[36] **J Powell**, *Two theorems on the mapping class group of a surface*, Proc. Amer. Math. Soc. 68 (1978) 347–350 MR

[37] **O Randal-Williams**, *Resolutions of moduli spaces and homological stability*, J. Eur. Math. Soc. 18 (2016) 1–81 MR

[38] **O H Rasmussen**, *Continuous variation of foliations in codimension two*, Topology 19 (1980) 335–349 MR

[39] **Y B Rudyak**, *On Thom spectra, orientability, and cobordism*, Springer (1998) MR

[40] **F Sergeraert**, *$B\Gamma$ [d'après John N Mather et William Thurston]*, from "Séminaire Bourbaki (1977/78)", Lecture Notes in Math. 710, Springer (1979) exposé 524, 300–315 MR

[41] **S Smale**, *Diffeomorphisms of the 2–sphere*, Proc. Amer. Math. Soc. 10 (1959) 621–626 MR

[42] **P Teichner**, *Topological four-manifolds with finite fundamental group*, Doktorarbeit, Johannes Gutenberg-Universität Mainz (1992) http://people.mpim-bonn.mpg.de/teichner/Papers/phd.pdf

[43] **W Thurston**, *Foliations and groups of diffeomorphisms*, Bull. Amer. Math. Soc. 80 (1974) 304–307 MR

[44] **U Tillmann**, *On the homotopy of the stable mapping class group*, Invent. Math. 130 (1997) 257–275 MR

[45] **T Tsuboi**, *On the foliated products of class C^1*, Ann. of Math. 130 (1989) 227–271 MR

[46] **T Tsuboi**, *Classifying spaces for groupoid structures*, from "Foliations, geometry, and topology" (N C Saldanha, L Conlon, R Langevin, T Tsuboi, P Walczak, editors), Contemp. Math. 498, Amer. Math. Soc., Providence, RI (2009) 67–81 MR

Department of Mathematics, Northwestern University
Evanston, IL, United States

sam@math.northwestern.edu

Proposed: Shigeyuki Morita
Seconded: Stefan Schwede, Dmitri Burago

Received: 3 March 2016
Revised: 17 October 2016

msp

Geometry & Topology 21 (2017) 3093–3158

Kato–Nakayama spaces, infinite root stacks and the profinite homotopy type of log schemes

DAVID CARCHEDI
SARAH SCHEROTZKE
NICOLÒ SIBILLA
MATTIA TALPO

For a log scheme locally of finite type over $\mathbb{C}$, a natural candidate for its profinite homotopy type is the profinite completion of its Kato–Nakayama space. Alternatively, one may consider the profinite homotopy type of the underlying topological stack of its infinite root stack. Finally, for a log scheme not necessarily over $\mathbb{C}$, another natural candidate is the profinite étale homotopy type of its infinite root stack. We prove that, for a fine saturated log scheme locally of finite type over $\mathbb{C}$, these three notions agree. In particular, we construct a comparison map from the Kato–Nakayama space to the underlying topological stack of the infinite root stack, and prove that it induces an equivalence on profinite completions. In light of these results, we define the profinite homotopy type of a general fine saturated log scheme as the profinite étale homotopy type of its infinite root stack.

14F35, 55U35; 55P60

1 Introduction

Log schemes are an enlargement of the category of schemes due to Fontaine, Illusie and Kato; see Kato [27]. The resulting variant of algebraic geometry, "logarithmic geometry", has applications in a variety of contexts ranging from moduli theory to arithmetic and enumerative geometry (see Abramovich, Chen, Gillam, Huang, Olsson, Satriano and Sun [1] for a recent survey).

In the past years there have been several attempts to capture the "log" aspect of these objects and translate it into a more familiar terrain. In the complex analytic case, Kato and Nakayama [28] introduced a topological space $X_{\log}$ (where X is a log analytic space), which may be interpreted as the "underlying topological space" of X, and over which, in some cases, one can write a comparison between logarithmic de Rham cohomology and ordinary singular cohomology. In a different direction, for a log scheme X, Kato introduced two sites, the Kummer-flat site X_{Kfl} and the Kummer-étale

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.3093

site X_{Ket}, that are analogous to the small fppf and étale site of a scheme, and were used later by Hagihara [19] and Nizioł [35] to study the K–theory of log schemes.

Recently, the fourth author together with Vistoli [48] introduced and studied a third incarnation of the "log" aspect of a log structure, namely the *infinite root stack* $\sqrt[\infty]{X}$, and used it to reinterpret Kato's Kummer sites and link them to parabolic sheaves on X. This stack is defined as the limit of an inverse system of algebraic stacks, $\sqrt[\infty]{X} = \varprojlim_n \sqrt[n]{X}$, parametrizing n^{th} roots of the log structure of X.

The infinite root stack can be thought of as an "algebraic incarnation" of the Kato–Nakayama space: if X is a log scheme locally of finite type over $\mathbb{C}$, both $X_{\log}$ and $\sqrt[\infty]{X}$ have a map to X. The fiber of $X_{\log} \to X_{\mathrm{an}}$ over a point $x \in X_{\mathrm{an}}$ is homeomorphic to $(S^1)^r$, where r is the rank of the log structure at x. For all n, the reduced fiber of $\sqrt[n]{X} \to X$ over the corresponding closed point of X is equivalent to the classifying stack $B(\mathbb{Z}/n\mathbb{Z})^r$ (for the same r). Regarding the infinite root stack not as the limit $\varprojlim_n \sqrt[n]{X}$, but instead as the diagram of stacks

$$n \mapsto \sqrt[n]{X},$$

ie as a pro-object or "formal limit", yields then that the reduced fiber of $\sqrt[\infty]{X} \to X$ is the diagram of stacks

$$n \mapsto B(\mathbb{Z}/n\mathbb{Z})^r,$$

which regarded as a pro-object is simply $B\widehat{\mathbb{Z}}^r \simeq \widehat{B\mathbb{Z}^r}$, the profinite completion of $(S^1)^r$.

In this paper we formalize this analogy and prove a comparison result between the profinite completions of $X_{\log}$ and $\sqrt[\infty]{X}$ for a fine saturated log scheme X locally of finite type over $\mathbb{C}$. Furthermore, we put this result in a wider circle of ideas, centered around the concept of the *profinite homotopy type* of a log scheme.

Our approach relies in a crucial way on a careful reworking of the foundations of the theory of topological stacks and profinite completions within the framework of ∞–categories; see Lurie [31]. This allows us to have greater technical control than earlier and more limited treatments, and plays an important role in the proof of our main result. In the second half of the paper we construct a comparison map between $X_{\log}$ and $\sqrt[\infty]{X}$ and show that it is induces an equivalence between their profinite completions. The proof involves an analysis of the local geometry of log schemes, and a local-to-global argument which reduces the statement to a local computation. Next, we review the main ideas in the paper in greater detail.

1.1 Topological stacks and profinite completions of homotopy types

The first ingredient that we need in order to compare $X_{\log}$ and $\sqrt[\infty]{X}$ is the notion of a *topological stack* (see Noohi [36]) associated with an algebraic stack. This is an extension of the analytification functor defined on schemes and algebraic spaces, which equips algebraic stacks with a topological counterpart, and allows one, for example, to talk about their homotopy type. Given an algebraic stack $\mathcal{X}$ locally of finite type over $\mathbb{C}$, let us denote by $\mathcal{X}_{\mathrm{top}}$ its "underlying topological stack". This formalism allows us to carry (infinite root stack) over $\sqrt[\infty]{X}$ to the topological world, where $X_{\log}$ lives.

The second ingredient we need is a functorial way of associating to a topological stack its homotopy type. Although this is in principle accomplished by Noohi [37] and Coyne and Noohi [14], the construction is a bit complicated and it is difficult to notice the nice formal properties this functor has from the construction. We instead construct a functor Π_∞ associating to a topological stack $\mathcal{X}$ its *fundamental* ∞*–groupoid*. The source of this functor is a suitable ∞–category of higher stacks on topological spaces, and the target is the ∞–category $\mathcal{S}$ of spaces. Using the language and machinery of ∞–categories makes the construction and functoriality of Π_∞ entirely transparent; it is the unique colimit-preserving functor which sends each space T to its weak homotopy type.

The third ingredient we need is a way of associating to a space its profinite completion. Combining this with the functor Π_∞ gives a way of associating to a topological stack a profinite homotopy type. The notion of profinite completion of homotopy types is originally due to Artin and Mazur [5]. Profinite homotopy types have since played many important roles in mathematics, perhaps most famously in relation to the Adams conjecture from algebraic topology; see Friedlander [16], Quillen [42] and Sullivan [47]. A more modern exposition using model categories is given by Isaksen [25] and Quick [40; 41]; however, the notion of profinite completion is a bit complicated in this framework. Finally, Lurie [32] briefly introduces an ∞–categorical model for profinite homotopy types, which has recently been shown to be equivalent to Quick's model by Barnea, Harpaz and Horel [6] (and also to a special case of Isaksen's). The advantage of Lurie's framework is that the definition of profinite spaces and the notion of profinite completion become very simple. A π*–finite space* is a space X with finitely many connected components, and finitely many homotopy groups, all of whom are finite, and a *profinite space* is simply a pro-object in the ∞–category of π–finite spaces. The profinite completion functor

$$\widehat{(\cdot)}\colon \mathcal{S} \to \mathrm{Prof}(\mathcal{S})$$

from the ∞–category of spaces to the ∞–category of profinite spaces preserves colimits, and composing this functor with Π_∞ gives a colimit-preserving functor $\widehat{\Pi}_\infty$ which assigns to a topological stack its profinite homotopy type. This property is used in an essential way in the proof of our main theorem. Using this machinery, we are able to derive some nontrivial properties of profinite spaces that are used in a crucial way to prove our main result; in particular we show that profinite spaces can be glued along hypercovers (Lemma 6.1).

1.2 The comparison map and the equivalence of profinite completions

Our main result states:

Theorem (see Theorem 6.4) *Let X be a fine saturated log scheme locally of finite type over $\mathbb{C}$. Then there is a canonical map of pro-topological stacks*

$$\Phi_X\colon X_{\log} \to \sqrt[\infty]{X}_{\mathrm{top}}$$

that induces an equivalence upon profinite completion,

$$\widehat{\Pi}_\infty(X_{\log}) \xrightarrow{\sim} \widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}}).$$

This theorem makes precise the idea that the infinite root stack is an algebraic incarnation of the Kato–Nakayama space, and that it completely captures the "profinite homotopy type" (à la Artin–Mazur) of the corresponding log scheme.

The construction of the comparison map Φ_X is first performed étale locally on X, where there is a global chart for the log structure, and then globalized by descent. The local construction uses the quotient stack description of the root stacks, that reduces the problem of finding a map to constructing a (topological) torsor on $X_{\log}$ with an equivariant map to a certain space.

This permits the construction of Φ_X as a canonical morphism of pro-topological stacks over X_{an}:

$$\begin{tikzcd} X_{\log} \arrow[rr, "\Phi_X"] \arrow[dr, "\pi_{\log}"'] & & \sqrt[\infty]{X}_{\mathrm{top}} \arrow[dl, "\pi_\infty"] \\ & X_{\mathrm{an}} & \end{tikzcd}$$

The jump patterns of the fibers of $\pi_{\log}$ and π_∞ reflect the way in which the rank of the log structure varies over X_{an}. More formally, the log structure defines a canonical stratification on X_{an} called the *rank stratification*, which makes $X_{\log}$ and $\sqrt[\infty]{X}_{\mathrm{top}}$ into stratified fibrations. After profinite completion, the fibers of $\pi_{\log}$ and π_∞ on each stratum become equivalent; indeed they are equivalent respectively to real tori of dimension n and to the (pro-)classifying stacks $B\widehat{\mathbb{Z}}^n$. The fact that the fibers of $\pi_{\log}$ and π_∞ are profinite homotopy equivalent was in fact our initial intuition as to why the main result should be true. Extracting from this fiberwise statement a proof that Φ_X induces an equivalence of profinite homotopy types requires a local-to-global argument that makes full use of the ∞–categorical framework developed in the first half of the paper.

The Kato–Nakayama space models the topology of log schemes, but its applicability is limited to schemes over the complex numbers. Our results suggest that the infinite root stack encodes all the topological information of log schemes (or at least its profinite completion) in a way that is exempt from this limitation. More precisely, if X is a log scheme locally of finite type over $\mathbb{C}$, there are three natural candidates for its "profinite homotopy type": the profinite completion of the Kato–Nakayama space $X_{\log}$, the profinite étale homotopy type of $\sqrt[\infty]{X}$ and the profinite completion of the (pro-)topological stack $\sqrt[\infty]{X}_{\mathrm{top}}$. Theorems 6.4 and 7.2 (the latter proven by Carchedi [11]) imply that these three constructions give the same result. This justifies the definition of the profinite homotopy type for a log scheme X, even outside of the complex case, as the profinite étale homotopy type of its infinite root stack $\sqrt[\infty]{X}$.

Another possible approach to this would be to define the homotopy type of a log scheme via Kato's Kummer-étale topos. As proved by Talpo and Vistoli [48, Section 6.2], this topos is equivalent to an appropriately defined small étale topos of the infinite root stack. It is not immediate, however, to link the resulting profinite homotopy type and the one that we define in the present paper. We plan to address this point in future work.

We believe that our results hold in the framework of *log analytic spaces* as well. Even though root stacks of those have not been considered anywhere yet, the construction and results about them that we use in the present paper should carry through without difficulty, using some notion of "analytic stacks" instead of algebraic ones.

In recent unpublished work, Howell and Vologodsky give a definition of the motive of a log scheme inside Voevodsky's triangulated category of motives. Based on our results we expect that infinite root stacks should provide an alternative encoding of the motive of log schemes, or a profinite approximation of it. It is an interesting question to explore possible connections between these two viewpoints.

Description of content

The paper is structured as follows.

In the first two sections we develop the framework necessary to associate profinite homotopy types to (pro-)algebraic and topological stacks. Along the way, in Section 3.4 we prove an interesting result (Theorem 3.25) which expresses the homotopy type of the Kato–Nakayama space of a log scheme as the classifying space of a natural category.

As a first step towards the main theorem, we construct in Section 4 (Proposition 4.1) a canonical map of pro-topological stacks

$$\Phi_X \colon X_{\log} \to \sqrt[\infty]{X}_{\mathrm{top}} \tag{1}$$

by exploiting the local quotient stack presentations of the root stacks $\sqrt[n]{X}$, and gluing the resulting maps.

Section 5 contains results about the topology of the Kato–Nakayama space and the topological infinite root stack that we use in an essential way in the proof of our main result.

In Section 6, we give the proof of Theorem 6.4: we show that the canonical map (1) induces an equivalence after profinite completion. The proof is based on a local-to-global analysis: we use a suitable hypercover $U^\bullet$ of X_{an} constructed in Section 5 to reduce the question to the restriction of the map Φ_X to each element of this hypercover. We then use the results about the topology of the Kato–Nakayama space and the topological infinite root stack proven in the same section to reduce to showing that the map induces a profinite homotopy equivalence along fibers. This concludes the proof.

Finally, in Section 7 we make some remarks about the definition of the profinite homotopy type of a general log scheme.

In the appendix, we gather definitions and facts that we use throughout the paper about log schemes, the analytification functor, the Kato–Nakayama space, root stacks, and topological stacks. In particular, in Appendix A.6, we carefully construct the "rank stratification" of X (and X_{an}), over which the characteristic monoid $\overline{M}$ of the log structure is locally constant.

Notations and conventions We will always work over a field k, which will almost always be the complex numbers $\mathbb{C}$. In particular all our log schemes will be fine and saturated, and locally of finite type over $\mathbb{C}$, unless otherwise stated.

If P is a monoid we denote by P^{gp} the associated group. Our monoids will typically be integral, finitely generated, saturated and sharp (hence torsion-free). A monoid P with these properties has a distinguished "generating set", consisting of all its indecomposable elements. This gives a presentation of any such monoid P through generators and relations.

If F is a sheaf of sets on the small étale site of a scheme, its "stalks" will always be stalks on geometric points.

By an ∞–category, we mean a quasicategory or inner-Kan complex. These are a model for $(\infty, 1)$–categories. We will follow very closely the notational conventions and terminology from Lurie [31], and refer the reader to the index and notational index in [31]. One slight deviation from the notational conventions just mentioned that will be made is that, for C and D objects of an ∞–category $\mathscr{C}$, we will denote by $\mathrm{Hom}_{\mathscr{C}}(C, D)$ the space of morphisms from C to D in $\mathscr{C}$, rather than using the notation $\mathrm{Map}_{\mathscr{C}}(C, D)$, in order to highlight the analogy with classical category theory. A very brief heuristic introduction to ∞–categories can be found in Appendix A of Carchedi [12]. See also Groth [18].

Acknowledgements All of the authors would like to thank their respective home institutions for their support.

We are also happy to thank Kai Behrend, Thomas Goodwillie, Marc Hoyois, Jacob Lurie, Thomas Nikolaus, Behrang Noohi, Gereon Quick, Angelo Vistoli and Kirsten Wickelgren for useful conversations.

We are grateful to the anonymous referee for a careful reading and useful comments, in particular for pointing out the short proof of Proposition 4.4.

2 Profinite homotopy types

In this section we will introduce the ∞–categorical model for profinite spaces that we will use in this article. This ∞–category is introduced in [32, Section 3.6]; a profinite space will succinctly be a pro-object in the ∞–category of π–finite spaces. This notion is equivalent to the notion of profinite space introduced by Quick [40; 41] (see [6]), but the machinery and language of ∞–categories is much more convenient to work with. Most importantly, the notion of profinite completion becomes completely transparent in this set up, and it is left adjoint to the canonical inclusion of profinite spaces into pro-spaces, and hence in particular preserves all colimits. We use this fact in an essential way in the proof of our main result, and we do not know how to prove the analogous fact about profinite completion in any other formalism.

We start first by reviewing the notion of ind-objects and pro-objects.

We will interchangeably use the notation $\mathcal{S}$ and Gpd_∞ for the ∞–category of spaces, and the ∞–category of ∞–groupoids. These two ∞–categories are one and the same, and we will use the different notations solely to emphasize in what way we are viewing the objects.

Recall that for $\mathscr{D}$ a small category, the category of ind-objects is essentially the category obtained from $\mathscr{D}$ by freely adjoining formal filtered colimits. This construction carries over for ∞–categories. Moreover, if $\mathscr{D}$ is an essentially small ∞–category, the ∞–category of *ind-objects* in $\mathscr{D}$, $\mathrm{Ind}(\mathscr{D})$, admits a canonical functor

$$j\colon \mathscr{D} \to \mathrm{Ind}(\mathscr{D})$$

satisfying the following universal property:

For every ∞–category $\mathscr{E}$ which admits small filtered colimits, composition with j induces an equivalence of ∞–categories

$$\mathrm{Fun}_{\mathrm{filt}}(\mathrm{Ind}(\mathscr{D}), \mathscr{E}) \to \mathrm{Fun}(\mathscr{D}, \mathscr{E}),$$

where $\mathrm{Fun}_{\mathrm{filt}}(\mathrm{Ind}(\mathscr{D}), \mathscr{E})$ denotes the ∞–category of all functors $\mathrm{Ind}(\mathscr{D}) \to \mathscr{E}$ which preserve filtered colimits.

A more concrete description of the ∞–category $\mathrm{Ind}(\mathscr{D})$ is as follows. First, recall the following proposition:

Proposition 2.1 [31, Corollary 5.3.5.4] *Denote by $\mathrm{Psh}_\infty(\mathscr{D})$ the ∞–category of ∞–presheaves on $\mathscr{D}$, that is, the functor category*

$$\mathrm{Fun}(\mathscr{D}^{\mathrm{op}}, \mathrm{Gpd}_\infty).$$

Let $\mathscr{D}$ be an essentially small ∞–category and let $F\colon \mathscr{D}^{\mathrm{op}} \to \mathrm{Gpd}_\infty$ be an ∞–presheaf. Then the following conditions are equivalent:

(i) *In the right fibration*

$$\int_{\mathscr{D}} F \to \mathscr{D}$$

classified by F, $\int_{\mathscr{D}} F$ is a filtered ∞–category.

(ii) *There exists a small filtered ∞–category $\mathfrak{J}$ and a functor*

$$f\colon \mathfrak{J} \to \mathscr{D}$$

such that F is the colimit of the composite

$$\mathfrak{J} \xrightarrow{f} \mathscr{D} \xhookrightarrow{y} \mathrm{Psh}_\infty(\mathscr{D})$$

(where y denotes the Yoneda embedding),

and, if $\mathscr{D}$ has finite colimits, (i) *and* (ii) *are equivalent to*

(iii) *F is left exact (ie preserves finite limits).*

The ∞–category $\mathrm{Ind}(\mathscr{D})$ may be described as the full subcategory of $\mathrm{Psh}_\infty(\mathscr{D})$ satisfying the equivalent conditions (i) and (ii) (or (iii) if $\mathscr{D}$ has finite colimits). In particular, this implies that j is full and faithful, since it is a restriction of the Yoneda embedding. In a nutshell $\mathrm{Psh}_\infty(\mathscr{D})$ is the ∞–category obtained from $\mathscr{D}$ by freely adjoining formal colimits, and (ii) above states that $\mathrm{Ind}(\mathscr{D})$ is the full subcategory thereof on those formal colimits of objects in $\mathscr{D}$ which are filtered colimits.

The notion of a pro-object is dual to that of an ind-object; it is a formal cofiltered limit. By definition, the ∞–category of *pro-objects* of an essentially small ∞–category $\mathscr{D}$ is

$$\mathrm{Pro}(\mathscr{D}) := \mathrm{Ind}(\mathscr{D}^{\mathrm{op}})^{\mathrm{op}}.$$

If $\mathscr{D}$ has small limits, we see that $\mathrm{Pro}(\mathscr{D})$ can be described as the full subcategory of $\mathrm{Fun}(\mathscr{D}, \mathrm{Gpd}_\infty)^{\mathrm{op}}$ on those functors

$$F\colon \mathscr{D} \to \mathrm{Gpd}_\infty$$

such that F preserves finite limits. Since this definition makes sense even when $\mathscr{D}$ is not essentially small, we make the following definition, due to Lurie:

Definition 2.2 If $\mathscr{E}$ is any accessible ∞–category with finite limits, then we define the ∞–category of *pro-objects* of $\mathscr{E}$, $\mathrm{Pro}(\mathscr{E})$, to be the full subcategory of $\mathrm{Fun}(\mathscr{E}, \mathrm{Gpd}_\infty)^{\mathrm{op}}$ on those functors $F\colon \mathscr{E} \to \mathrm{Gpd}_\infty$ which are accessible and preserve finite limits.

Remark 2.3 If $\mathscr{E}$ is any accessible ∞–category and E is an object of $\mathscr{E}$, then the functor

$$\mathrm{Hom}(E, \cdot)\colon \mathscr{E} \to \mathrm{Gpd}_\infty$$

corepresented by E is accessible and preserves all limits. This induces a fully faithful functor

$$\mathscr{E} \xhookrightarrow{j} \mathrm{Pro}(\mathscr{E}).$$

The functor j satisfies the following universal property:

If $\mathcal{D}$ is any ∞–category admitting small cofiltered limits, then composition with j induces an equivalence of ∞–categories

$$\mathrm{Fun}_{\mathrm{cofilt}}(\mathrm{Pro}(\mathscr{E}), \mathscr{D}) \to \mathrm{Fun}(\mathscr{E}, \mathscr{D}), \tag{2}$$

where $\mathrm{Fun}_{\mathrm{cofilt}}(\mathrm{Pro}(\mathscr{E}), \mathscr{D})$ is the full subcategory of $\mathrm{Fun}(\mathrm{Pro}(\mathscr{E}), \mathscr{D})$ spanned by those functors which preserve small cofiltered limits; see [32, Proposition 3.1.6].

Remark 2.4 If $\mathscr{C}$ is any (not necessarily accessible) ∞–category, there always exists an ∞–category $\mathrm{Pro}(\mathscr{C})$ satisfying the universal property (2). This is a special case of [31, Proposition 5.3.6.2].

Remark 2.5 Let $\mathscr{E}$ be any accessible ∞–category which is not necessarily essentially small. Let $\mathcal{U}$ be the Grothendieck universe of small sets and let $\mathcal{V}$ be a Grothendieck universe such that $\mathcal{U} \in \mathcal{V}$, so that we may regard $\mathcal{V}$ as the Grothendieck universe of large sets. Let $\widehat{\mathrm{Gpd}}_\infty$ denote the ∞–category of ∞–groupoids in the universe $\mathcal{V}$. By the proof of [32, Proposition 3.1.6], it follows that the essential image of the composition

$$\mathrm{Pro}(\mathscr{E}) \hookrightarrow \mathrm{Fun}(\mathscr{E}, \mathrm{Gpd}_\infty)^{\mathrm{op}} \hookrightarrow \mathrm{Fun}(\mathscr{E}, \widehat{\mathrm{Gpd}}_\infty)^{\mathrm{op}}$$

consists of those functors $F\colon \mathscr{E} \to \widehat{\mathrm{Gpd}}_\infty$ for which there exists a small filtered ∞–category $\mathcal{J}$ and a functor

$$f\colon \mathcal{J} \to \mathscr{E}^{\mathrm{op}}$$

such that F is the colimit of the composite

$$\mathcal{J} \xrightarrow{f} \mathscr{E}^{\mathrm{op}} \hookrightarrow \mathrm{Fun}(\mathscr{E}, \widehat{\mathrm{Gpd}}_\infty).$$

Remark 2.6 In light of Remark 2.5, any object X of $\mathrm{Pro}(\mathscr{E})$, for $\mathscr{E}$ an accessible ∞–category, can be written as a cofiltered limit of a diagram of the form

$$F\colon \mathcal{I} \to \mathscr{E} \xhookrightarrow{j} \mathrm{Pro}(\mathscr{E}),$$

or, in more informal notation,

$$X = \varprojlim_{i \in \mathcal{I}} X_i.$$

Unwinding the definitions, we see that if $Y = \varprojlim_{j \in \mathcal{J}} Y_j$ is another such object of $\mathrm{Pro}(\mathscr{E})$, then the usual formula for the morphism space holds:

$$\mathrm{Hom}_{\mathrm{Pro}(\mathscr{E})}(X, Y) \simeq \varprojlim_{j \in \mathcal{J}} \varinjlim_{i \in \mathcal{I}} \mathrm{Hom}_{\mathscr{E}}(X_i, Y_j).$$

Now suppose that $\mathscr{E}$ has a terminal object 1. Then

$$\mathrm{Hom}_{\mathrm{Pro}(\mathscr{E})}(X, j(1)) \simeq \varinjlim_{i \in \mathcal{I}} \mathrm{Hom}_{\mathscr{E}}(X_i, 1).$$

Notice that each space $\mathrm{Hom}_{\mathscr{E}}(X_i, 1)$ is contractible since 1 is terminal, and, since (-2)–truncated objects (ie terminal objects) are closed under filtered colimits in $\mathcal{S}$ by [31, Corollary 5.5.7.4], it follows that $\mathrm{Hom}_{\mathrm{Pro}(\mathscr{E})}(X, j(1))$ itself is a contractible space, and hence we conclude that $j(1)$ is a terminal object.

Example 2.7 Let $\mathscr{E} = \mathcal{S}$ be the ∞–category of spaces. Then the ∞–category of *pro-spaces*, $\mathrm{Pro}(\mathcal{S})$, can be identified with the opposite category of functors $F\colon \mathcal{S} \to \mathcal{S}$ such that F is accessible and left exact. Notice that any space X gives rise to a pro-space

$$\mathrm{Hom}(X, \cdot)\colon \mathcal{S} \to \mathcal{S}$$

which moreover preserves all limits. Moreover if $F\colon \mathcal{S} \to \mathcal{S}$ is *any* functor which preserves all limits, then by the Adjoint Functor theorem for ∞–categories [31, Corollary 5.5.2.9], F must have a right adjoint G, and is moreover accessible by [31, Proposition 5.4.7.7]. This then implies that

$$\mathrm{Hom}(G(*), X) \simeq \mathrm{Hom}(*, F(X)) \simeq F(X).$$

Hence $F \simeq j(G(*))$. We conclude that the essential image of

$$j\colon \mathcal{S} \hookrightarrow \mathrm{Pro}(\mathcal{S})$$

is precisely those ∞–functors $\mathcal{S} \to \mathcal{S}$ which preserve all small limits.

Proposition 2.8 *The functor*

$$T\colon \mathrm{Pro}(\mathcal{S}) \xrightarrow{\mathrm{Hom}(j(*),\cdot)} \mathcal{S}$$

is right adjoint to the canonical inclusion $j\colon \mathcal{S} \to \mathrm{Pro}(\mathcal{S})$.

Proof By Remark 2.5, we may identify $\mathrm{Pro}(\mathcal{S})^{\mathrm{op}}$ with a subcategory of the ∞–category $\mathrm{Fun}(\mathcal{S}, \widehat{\mathrm{Gpd}}_\infty)$ of large ∞–copresheaves, and since limits commute with limits, this subcategory is stable under small limits. Note that this implies that $\mathrm{Pro}(\mathcal{S})$ is cocomplete. Since the Yoneda embedding into large ∞–presheaves

$$\mathcal{S}^{\mathrm{op}} \overset{y}{\hookrightarrow} \widehat{\mathrm{Psh}}_\infty(\mathcal{S})$$

preserves small limits, it follows that

$$j\colon \mathcal{S} \hookrightarrow \mathrm{Pro}(\mathcal{S})$$

preserves small colimits. Since $\mathcal{S} \simeq \mathrm{Psh}_\infty(1)$, where 1 is the terminal ∞–category, and since $\mathrm{Pro}(\mathcal{S})$ is cocomplete, one has by [31, Theorem 5.1.5.6] that $j \simeq \mathrm{Lan}_{y_1}(t)$, where y_1 is the Yoneda embedding $1 \to \mathcal{S}$ and $t\colon 1 \to \mathrm{Pro}(\mathcal{S})$ is the functor picking out the object $j(*)$. It follows immediately from the Yoneda lemma that $\mathrm{Hom}(j(*),\cdot)$ is right adjoint to $\mathrm{Lan}_{y_1}(t)$. □

Remark 2.9 Let $P\colon \mathcal{S} \to \mathcal{S}$ be a pro-space. By [31, Proposition 5.4.6.6], since P is accessible it follows that the associated left fibration

$$\int_{\mathcal{S}} P$$

is accessible, and hence has a small cofinal subcategory

$$r\colon \mathcal{C}_P \hookrightarrow \int_{\mathcal{S}} P,$$

and P may be identified with the limit of the composite

$$\mathcal{C}_P \overset{r}{\hookrightarrow} \int_{\mathcal{S}} P \xrightarrow{\pi_P} \mathcal{S} \overset{j}{\hookrightarrow} \mathrm{Pro}(\mathcal{S}).$$

We claim that

$$T(P) \simeq \varprojlim \pi_P \circ r.$$

Indeed,

$$\begin{aligned} T(P) &= \mathrm{Hom}(j(*), P) \\ &\simeq \mathrm{Hom}(j(*), \varprojlim j \circ \pi_P \circ r) \\ &\simeq \varprojlim \mathrm{Hom}(j(*), j \circ \pi_P \circ r) \\ &\simeq \varprojlim \mathrm{Hom}(*, \pi_P \circ r) \\ &\simeq \varprojlim \pi_P \circ r. \end{aligned}$$

By the same proof, if one has P presented as a cofiltered limit $P = \varprojlim j(X_\alpha)$ of spaces, then $T(P) \simeq \varprojlim X_\alpha$. In fact, this holds more generally, by the next proposition.

Proposition 2.10 *Let $\mathscr{C}$ be an accessible ∞–category which admits small filtered limits. Then the canonical inclusion*

$$j\colon \mathscr{C} \hookrightarrow \mathrm{Pro}(\mathscr{C})$$

has a right adjoint T and if $F\colon \mathcal{I} \to \mathscr{C}$ is a cofiltered diagram corresponding to an object in $\mathrm{Pro}(\mathscr{C})$*, then $T(F) = \varprojlim F$.*

Proof By Remark 2.5, composition with

$$j\colon \mathscr{C} \hookrightarrow \mathrm{Pro}(\mathscr{C})$$

induces an equivalence of ∞–categories

$$\mathrm{Fun}_{\mathrm{cofilt}}(\mathrm{Pro}(\mathscr{C}), \mathscr{C}) \to \mathrm{Fun}(\mathscr{C}, \mathscr{C}),$$

so we can find a functor $T\colon \mathrm{Pro}(\mathscr{C}) \to \mathscr{C}$ and an equivalence

$$\eta\colon \mathrm{id}_{\mathscr{C}} \xrightarrow{\sim} T \circ j.$$

Let Z be an arbitrary object of $\mathrm{Pro}(\mathscr{C})$; then we can write $Z = \varprojlim_{i \in \mathcal{I}} j(X_i)$. First note that since η is an equivalence and T preserves cofiltered limits (by definition), we have that for such a Z,

$$T(Z) \simeq \varprojlim_{i \in \mathcal{I}} X_i.$$

This shows that T has the desired properties on pro-objects. Let us now show that T is a right adjoint to j. Let C be an object of $\mathscr{C}$; then we have

$$\mathrm{Hom}_{\mathscr{C}}(D, T(Z)) \simeq \mathrm{Hom}_{\mathscr{C}}(D, \varprojlim_{i \in \mathcal{I}} X_i) \simeq \varprojlim_{i \in \mathcal{I}} \mathrm{Hom}_{\mathscr{C}}(D, X_i),$$

and since j is fully faithful, we have for each i

$$\mathrm{Hom}_{\mathscr{C}}(D, X_i) \simeq \mathrm{Hom}_{\mathrm{Prof}(\mathscr{C})}(j(D), j(X_i)).$$

It follows then that

$$\begin{aligned}\mathrm{Hom}_{\mathscr{C}}(D, T(Z)) &\simeq \varprojlim_{i \in \mathcal{I}} \mathrm{Hom}_{\mathrm{Prof}(\mathscr{C})}(j(D), j(X_i)) \\ &\simeq \mathrm{Hom}_{\mathrm{Prof}(\mathscr{C})}(j(D), \varprojlim_{i \in \mathcal{I}} j(X_i)) \\ &= \mathrm{Hom}_{\mathrm{Prof}(\mathscr{C})}(j(D), Z).\end{aligned}$$

□

Definition 2.11 A space X in $\mathcal{S}$ is *π–finite* if all of its homotopy groups are finite, it has only finitely many nontrivial homotopy groups, and finitely many connected components.

Definition 2.12 Let $\mathcal{S}^{\mathrm{fc}}$ denote the full subcategory of the ∞–category $\mathcal{S}$ on the π–finite spaces. $\mathcal{S}^{\mathrm{fc}}$ is essentially small and idempotent complete (and hence accessible). The ∞–category of *profinite spaces* is defined to be the ∞–category

$$\mathrm{Prof}(\mathcal{S}) := \mathrm{Pro}(\mathcal{S}^{\mathrm{fc}}).$$

Proposition 2.13 *Let V be a π–finite space. Note that V is n–truncated for some n, since it has only finitely many homotopy groups. The associated profinite space $j(V)$ is also n–truncated.*

Proof Let $X = \varprojlim_{i \in \mathcal{I}} X_i$ be a profinite space. Then, by Remark 2.6, we have that

$$\mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}(X, j(V)) \simeq \underset{i \in \mathcal{I}}{\underrightarrow{\mathrm{colim}}}\, \mathrm{Hom}_{\mathcal{S}^{\mathrm{fc}}}(X_i, V).$$

Each space $\mathrm{Hom}_{\mathcal{S}^{\mathrm{fc}}}(X_i, V)$ is n–truncated since V is, and n–truncated spaces are stable under filtered colimits by [31, Corollary 5.5.7.4], so it follows that $\mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}(X, j(V))$ is also n–truncated. □

Remark 2.14 The assignment $\mathscr{C} \mapsto \mathrm{Pro}(\mathscr{C})$ is functorial among accessible ∞–categories with finite limits. Given a functor $f \colon \mathscr{C} \to \mathscr{D}$, the composite

$$\mathscr{C} \xrightarrow{f} \mathscr{D} \xhookrightarrow{j} \mathrm{Pro}(\mathscr{D})$$

corresponds to an object of the ∞–category $\mathrm{Fun}(\mathscr{C}, \mathrm{Pro}(\mathscr{D}))$, which by Remark 2.3 is equivalent to the ∞–category $\mathrm{Fun}_{\mathrm{cofilt}}(\mathrm{Pro}(\mathscr{C}), \mathrm{Pro}(\mathscr{D}))$. Hence, one gets an induced functor

$$\mathrm{Pro}(f)\colon \mathrm{Pro}(\mathscr{C}) \to \mathrm{Pro}(\mathscr{D})$$

which preserves cofiltered limits. Moreover, $\mathrm{Pro}(f)$ is fully faithful if f is. If f happens to be accessible and left exact, then there is an induced functor in the opposite direction, given by

$$f^*\colon \mathrm{Pro}(\mathscr{D}) \to \mathrm{Pro}(\mathscr{C}), \quad (\mathscr{D} \xrightarrow{F} \mathrm{Gpd}_\infty) \mapsto (\mathscr{C} \xrightarrow{f} \mathscr{D} \xrightarrow{F} \mathrm{Gpd}_\infty),$$

and f^* is left adjoint to $\mathrm{Pro}(f)$. See [32, Remark 3.1.7] (but note there is a typo, since f^* is in fact a left adjoint, not a right adjoint).

Example 2.15 The canonical inclusion $i\colon \mathcal{S}^{\mathrm{fc}} \hookrightarrow \mathcal{S}$ induces a fully faithful embedding

$$\mathrm{Pro}(i)\colon \mathrm{Prof}(\mathcal{S}) \hookrightarrow \mathrm{Pro}(\mathcal{S})$$

of profinite spaces into pro-spaces. Moreover, i is accessible and preserves finite limits, hence the above functor has a left adjoint

$$i^*\colon \mathrm{Pro}(\mathcal{S}) \to \mathrm{Prof}(\mathcal{S}).$$

This functor sends a pro-space P to its *profinite completion*.

Definition 2.16 We denote by $\widehat{(\cdot)}$ the composite

$$\mathcal{S} \xhookrightarrow{j} \mathrm{Pro}(\mathcal{S}) \xrightarrow{i^*} \mathrm{Prof}(\mathcal{S})$$

and call it the *profinite completion functor*. Concretely, if X is a space in $\mathcal{S}$, then $\widehat{X}$ corresponds to the composite

$$\mathcal{S}^{\mathrm{fc}} \xhookrightarrow{i} \mathcal{S} \xrightarrow{\mathrm{Hom}(X,\cdot)} \mathcal{S}.$$

This functor has a right adjoint given by the composite

$$\mathrm{Prof}(\mathcal{S}) \xhookrightarrow{\mathrm{Pro}(i)} \mathrm{Pro}(\mathcal{S}) \xrightarrow{T} \mathcal{S}.$$

We will denote this right adjoint simply by U.

Remark 2.17 We will sometimes abuse notation and denote the profinite completion of a pro-space Y also by $\widehat{Y}$ rather than i^*Y, when no confusion will arise.

2.1 The relationship with profinite groups

In this subsection, we will touch briefly upon the relationship between profinite groups and profinite spaces. Recall the notion of profinite completion of a group. A profinite group is a pro-object in the category of finite groups. Equivalently, a profinite group is a group object in profinite sets; see [32, Proposition 3.2.12].

Denote by $i\colon \mathit{FinGp} \hookrightarrow \mathit{Gp}$ the fully faithful inclusion of the category of finite groups into the category of groups. The composite

$$\mathit{Gp} \hookrightarrow \mathrm{Pro}(\mathit{Gp}) \xrightarrow{i^*} \mathrm{Pro}(\mathit{FinGp}) \simeq \mathit{Gp}(\mathrm{Pro}(\mathit{FinSet}))$$

is the functor assigning to a group its profinite completion. We also denote this functor by $\widehat{(\cdot)}$ when no confusion will arise. Recall that the profinite completion of a group has a classical concrete description as follows: Let G be a group, then its profinite completion is the limit $\varprojlim_N j(G/N)$, where N ranges over all the finite index normal subgroups of G. Similarly, denote by $i_{\mathrm{ab}}\colon \mathit{FinAbGp} \hookrightarrow \mathit{AbGp}$ the fully faithful inclusion of the category of finite abelian groups into the category of abelian groups. By the analogous construction to the above, there is an induced profinite completion functor

$$\widehat{(\cdot)}_{\mathrm{ab}}\colon \mathit{AbGp} \to \mathrm{Pro}(\mathit{FinAbGp}).$$

It can be described classically by the same formula as in the nonabelian case. If

$$\phi\colon \mathit{AbGp} \hookrightarrow \mathit{Gp}$$

is the canonical inclusion of abelian groups into groups, it follows that the following diagram commutes up to canonical natural equivalence:

$$\begin{tikzcd}
\mathit{AbGp} \arrow[r, "\widehat{(\cdot)}_{\mathrm{ab}}"] \arrow[d, "\phi"'] & \mathrm{Pro}(\mathit{FinAbGp}) \arrow[d, "\mathrm{Pro}(\phi)"] \\
\mathit{Gp} \arrow[r, "\widehat{(\cdot)}"] & \mathrm{Pro}(\mathit{FinGp})
\end{tikzcd}$$

By [32, Proposition 3.2.14], there is a canonical equivalence of categories

$$\mathrm{Pro}(\mathit{FinAbGp}) \simeq \mathit{AbGp}(\mathrm{Pro}(\mathit{FinSet}))$$

between the category of pro-objects in finite abelian groups and the category of abelian group objects in profinite sets. Thus, in particular, finite coproducts (direct sums) in $\mathrm{Pro}(\mathit{FinAbGp})$ coincide with finite products. Since $\widehat{(\cdot)}_{\mathrm{ab}}$ is a left adjoint, it preserves direct sums, and by Remark 2.14, $\mathrm{Pro}(\phi)$ is a right adjoint (since ϕ preserves finite limits), so $\mathrm{Pro}(\phi)$ preserves products. It follows that the composite

$$\widehat{(\cdot)} \circ \phi\colon \mathit{AbGp} \to \mathrm{Pro}(\mathit{FinGp})$$

preserves finite products.

Corollary 2.18 *Let k be a nonnegative integer. Then there is a canonical isomorphism of profinite groups*

$$\widehat{\mathbb{Z}^k} \cong \hat{\mathbb{Z}}^k.$$

We now note a recent result which compares the ∞–categorical model for profinite spaces just presented with the model categorical approach developed by Quick [40; 41]:

Theorem 2.19 [6, Corollary 7.4.6] *The ∞–category associated to the model category presented in* [40; 41] *is equivalent to* $\mathrm{Prof}(\mathcal{S})$.

The details of Quick's model category need not concern us here, but we cite the above theorem in order to freely use results of [40; 41] about profinite spaces.

Proposition 2.20 *Let k be a nonnegative integer. There is a canonical equivalence of profinite spaces*

$$\widehat{B(\mathbb{Z}^k)} \simeq B(\hat{\mathbb{Z}}^k).$$

Proof Since $\mathbb{Z}^k$ is a finitely generated free abelian group, it is *good* in the sense of Serre [45]. It follows from [41, Proposition 3.6] and Theorem 2.19 that the canonical map

$$\widehat{B(\mathbb{Z}^k)} \to B(\widehat{\mathbb{Z}^k})$$

is an equivalence of profinite spaces. The result now follows from Corollary 2.18. □

The following lemma will be used in an essential way several times in this paper:

Lemma 2.21 *Let $f\colon \Delta \to \mathscr{C}$ be a cosimplicial diagram and suppose that $\mathscr{C}$ is an $(n+1,1)$–category, ie an ∞–category whose mapping spaces are all n–truncated. Then, provided both limits exist, the canonical map*

$$\varprojlim f \to \varprojlim (f|_{\Delta_{\leq n}})$$

is an equivalence.

Proof Let $\mathscr{C}$ be an arbitrary $(\infty, n+1)$–category. Notice that for any diagram $f\colon \Delta \to \mathscr{C}$ and any object C of $\mathscr{C}$, we have

$$\mathrm{Hom}\big(C, \varprojlim_{k \in \Delta} f(k)\big) \simeq \varprojlim_{k \in \Delta} \mathrm{Hom}(C, f(k)),$$

and since $\mathscr{C}$ is an $(\infty, n+1)$–category, each $\mathrm{Hom}(C, f(k))$ is an n–truncated space. Therefore the general case follows from the case when $\mathscr{C}$ is the full subcategory $\mathcal{S}^{\leq n}$ of $\mathcal{S}$ on the n–truncated spaces. By [31, Theorem 4.2.4.1], to prove the lemma for the special case $\mathscr{C} = \mathcal{S}^{\leq n}$, it suffices to prove the corresponding statement about homotopy limits in the Quillen model structure on the category of compactly generated spaces $\mathbf{CG}$, since the associated ∞–category is $\mathcal{S}$.

Suppose that

$$X^\bullet\colon \Delta \to \mathbf{CG}$$

is a cosimplicial space which is fibrant with respect to the projective model structure on $\mathrm{Fun}(\Delta, \mathbf{CG})$ (with respect to the Quillen model structure on $\mathbf{CG}$), ie the diagram $X^\bullet$ consists entirely of Serre fibrations. Then the homotopy limit of $X^\bullet$ may be computed as $\mathrm{Tot}(X)$, and moreover, $\mathrm{Tot}(X)$ can be written as the (homotopy) limit of a tower of fibrations

$$\cdots \to \mathrm{Tot}(X)_k \to \mathrm{Tot}(X)_{k-1} \to \cdots \to \mathrm{Tot}(X)_1 \to \mathrm{Tot}(X)_0 = X,$$

where each $\mathrm{Tot}(X)_k$ is a model for the homotopy limit of $X|_{\Delta_{\leq k}}$. Moreover, the (homotopy) fiber of each map

$$\mathrm{Tot}(X)_k \to \mathrm{Tot}(X)_{k-1}$$

is homotopy equivalent to the k–fold loop space $\Omega^k(M^k X^\bullet)$, where

$$M^k X^\bullet = \varprojlim_{\substack{[k+1] \twoheadrightarrow [j] \\ j \leq k}} X^j$$

is the k^{th} matching object of $X^\bullet$ (see eg the introduction of [33]).

Now let us assume that each X_k is n–truncated. Then, as $X^\bullet$ is fibrant, the diagram involved in the limit above consists entirely of fibrations, so the limit is a homotopy limit, hence each matching object is also n–truncated (since n–truncated objects are stable under limits in $\mathcal{S}$ by [31, Proposition 5.5.6.5]). It follows then that each homotopy fiber

$$\mathrm{Tot}(X)_k \to \mathrm{Tot}(X)_{k-1}$$

is weakly contractible for $k > n$, and hence the natural map

$$\underleftarrow{\mathrm{holim}}\, X^\bullet = \mathrm{Tot}(X) \to \mathrm{Tot}(X)_n = \underleftarrow{\mathrm{holim}}\, X^\bullet|_{\Delta_{\leq n}}$$

is a weak homotopy equivalence. $\square$

Proposition 2.22 *Let* $\varprojlim_{i \in \mathcal{I}} G_i$ *be a pro-object in the category of finite groups, or, equivalently, a group object in* $\operatorname{Pro}(\mathit{FinSet})$. *Consider the profinite space*

$$B\big(\varprojlim_i G_i\big) := \varinjlim\big[\ldots \varprojlim_i G_i \times \varprojlim_i G_i \substack{\rightarrow\\\rightarrow\\\rightarrow} \varprojlim_i G_i \rightrightarrows *\big],$$

where the colimit is computed in $\operatorname{Prof}(\mathcal{S})$ *and* $*$ *denotes the terminal profinite space. In more detail, the diagram whose colimit is being taken is the simplicial diagram which is the Čech nerve of the unique map* $\varprojlim_i G_i \to *$ *in* $\operatorname{Prof}(\mathcal{S})$. *Consider for each* i *the object in* $\mathcal{S}$

$$B(G_i) := \varinjlim\big[\ldots G_i \times G_i \substack{\rightarrow\\\rightarrow\\\rightarrow} G_i \rightrightarrows *\big],$$

ie the colimit in $\mathcal{S}$ *of the Čech nerve of* G_i. *Then these spaces assemble into a profinite space* $\varprojlim_i B(G_i)$, *and we have a canonical equivalence*

$$B\big(\varprojlim_i G_i\big) \simeq \varprojlim_i B(G_i)$$

in $\operatorname{Prof}(\mathcal{S})$.

Proof It suffices to prove that for each π–finite space V we have an equivalence

$$\operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(B\big(\varprojlim_i G_i\big), j(V)\big) \simeq \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(\varprojlim_i B(G_i), j(V)\big).$$

Recall that, by Proposition 2.13, $j(V)$ is n–truncated for some n. As such, we have

$$\begin{aligned}\operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(B\big(\varprojlim_i G_i\big), j(V)\big) &\simeq \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(\varinjlim_{\Delta^{\mathrm{op}}} N\big(\varprojlim_i G_i\big), j(V)\big)\\ &\simeq \varprojlim_{\Delta} \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(N\big(\varprojlim_i G_i\big), j(V)\big)\\ &\simeq \varprojlim_{\Delta_{\le n}} \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(N\big(\varprojlim_i G_i\big), j(V)\big),\end{aligned}$$

that last equivalence following from Lemma 2.21. Expanding this out we get

$$\begin{aligned}&\varprojlim_{\Delta_{\le n}}\big[\operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}(1, j(V)) \rightrightarrows \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(\varprojlim_i G_i, j(V)\big)\\ &\qquad \substack{\rightarrow\\\rightarrow\\\rightarrow} \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(\big(\varprojlim_i G_i\big)^2, j(V)\big) \ldots \operatorname{Hom}_{\operatorname{Prof}(\mathcal{S})}\big(\big(\varprojlim_i G_i\big)^n, j(V)\big)\big]\end{aligned}$$

which is equivalent to

$$\begin{aligned}&\varprojlim_{\Delta_{\le n}}\big[\operatorname{Hom}_{\mathcal{S}}(*, V) \rightrightarrows \varinjlim_i \operatorname{Hom}_{\mathcal{S}}(G_i, j(V))\\ &\qquad \substack{\rightarrow\\\rightarrow\\\rightarrow} \varinjlim_i \operatorname{Hom}_{\mathcal{S}}(G_i^2, j(V)) \ldots \varinjlim_i \operatorname{Hom}_{\mathcal{S}}(G_i^n, j(V))\big]\end{aligned}$$

and since by [31, Proposition 5.3.3.3] finite limits commute with filtered colimits in $\mathcal{S}$, we get

$$\varinjlim_i \varprojlim_{\Delta_{\le n}} [\operatorname{Hom}_{\mathcal{S}}(*, V) \rightrightarrows \operatorname{Hom}_{\mathcal{S}}(G_i, V) \substack{\rightarrow\\\rightarrow\\\rightarrow} \operatorname{Hom}_{\mathcal{S}}(G_i^2, j(V)) \ldots \operatorname{Hom}_{\mathcal{S}}(G_i^n, V)].$$

Now, since $j(V)$ is n–truncated by Proposition 2.13, it follows from Lemma 2.21 that we can rewrite this as

$$\underset{\overrightarrow{i}}{\operatorname{colim}}\, \underset{\overleftarrow{\Delta}}{\lim}\big[\operatorname{Hom}_{\mathcal{S}}(*, V) \rightrightarrows \operatorname{Hom}_{\mathcal{S}}(G_i, V) \mathrel{\substack{\textstyle\rightarrow\\[-0.6ex]\textstyle\rightarrow\\[-0.6ex]\textstyle\rightarrow}} \operatorname{Hom}_{\mathcal{S}}(G_i^2, j(V)) \ldots \operatorname{Hom}_{\mathcal{S}}(G_i^n, V) \ldots\big],$$

which is equivalent to

$$\begin{aligned}\underset{\overrightarrow{i}}{\operatorname{colim}}\, \operatorname{Hom}_{\mathcal{S}}\big(\underset{\overrightarrow{\Delta^{\mathrm{op}}}}{\operatorname{colim}}\, N(G_i), V\big) &\simeq \underset{\overrightarrow{i}}{\operatorname{colim}}\, \operatorname{Hom}_{\mathcal{S}}(B(G_i), V)\\ &\simeq \operatorname{Hom}_{\mathrm{Prof}(\mathcal{S})}\big(\underset{\overleftarrow{i}}{\lim}\, B(G_i), j(V)\big).\end{aligned}$$ □

3 The homotopy type of topological stacks

In this section we use the formalism of ∞–categories to produce two important constructions necessary for our paper. Firstly, we extend the construction of analytification, which sends a complex variety to its set of closed points, equipped with the analytic topology, to a colimit-preserving functor

$$(\cdot)_{\mathrm{top}}\colon \mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$$

from the ∞–category of ∞–sheaves over the étale site of affine schemes of finite type over $\mathbb{C}$ to the ∞–category of hypersheaves over an appropriate topological site. This functor, in particular, sends an Artin stack locally of finite type over $\mathbb{C}$ to its underlying topological stack in the sense of Noohi [36]. Using this functor, one associates to the infinite root stack $\sqrt[\infty]{X}$ of a log scheme a pro-topological stack $\sqrt[\infty]{X}_{\mathrm{top}}$. In Section 4, we produce a map

(3) $$X_{\log} \to \sqrt[\infty]{X}_{\mathrm{top}}$$

from the Kato–Nakayama space to the underlying (pro-)topological stack of the infinite root stack. The main result of the paper is that this map is a profinite homotopy equivalence, but to make sense of such a statement, one first needs to associate to each of these objects a (pro-)homotopy type, in a functorial way. To achieve this, the second construction we produce is a colimit-preserving functor

$$\Pi_\infty\colon \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}) \to \mathcal{S}$$

which sends every topological space X to its underlying homotopy type, and sends every topological stack to its homotopy type in the sense of Noohi [37]. Using this construction and the map (3), one has an induced map in $\mathrm{Pro}(\mathcal{S})$,

$$\Pi_\infty(X_{\log}) \to \Pi_\infty(\sqrt[\infty]{X}_{\mathrm{top}}),$$

which we prove in Section 6 becomes an equivalence after applying the profinite completion functor, ie the map (3) is a profinite homotopy equivalence.

3.1 The underlying topological stack of an algebraic stack

Let $\mathbf{Top}$ be the category of topological spaces and let $\mathbf{Top}^s_{\mathbb{C}}$ denote the full subcategory of $\mathbf{Top}$ of all contractible and locally contractible spaces which are homeomorphic to a subspace of $\mathbb{R}^n$ for some n. Note that $\mathbf{Top}^s_{\mathbb{C}}$ is essentially small. Denote by $\mathbf{Top}_{\mathbb{C}}$ the following subcategory of topological spaces:

- A topological space T is in $\mathbf{Top}_{\mathbb{C}}$ if T has an open cover $(U_\alpha \hookrightarrow T)_\alpha$ such that each U_α is an object of $\mathbf{Top}^s_{\mathbb{C}}$.

We use the subscript $\mathbb{C}$ to highlight the fact that $\mathbf{Top}_{\mathbb{C}}$ will serve as the target of the analytification functor from the category of algebraic spaces over $\mathbb{C}$. Note that the objects of $\mathbf{Top}_{\mathbb{C}}$ are closed under taking open subspaces. As such, it makes sense to equip $\mathbf{Top}_{\mathbb{C}}$ with the Grothendieck topology generated by open covers. Denote by $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$ the ∞–topos of hypersheaves on $\mathbf{Top}_{\mathbb{C}}$, ie the hypercompletion of the ∞–topos of ∞–sheaves. There is also a natural structure of a Grothendieck site on $\mathbf{Top}^s_{\mathbb{C}}$ as follows:

- Let T be a space in $\mathbf{Top}^s_{\mathbb{C}}$. A covering family of T consists of an open cover $(U_\alpha \hookrightarrow T)$ such that each U_α is in $\mathbf{Top}^s_{\mathbb{C}}$.

Note that every open cover of T can be refined by such a cover. We denote the associated ∞–topos of hypersheaves by $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}})$. By the comparison lemma of [3, Exposé III], we have that restriction along the canonical inclusion

$$\mathbf{Top}^s_{\mathbb{C}} \hookrightarrow \mathbf{Top}_{\mathbb{C}}$$

induces an equivalence between their respective categories of sheaves of sets. It then follows from [26, Theorem 5; 31, Proposition 6.5.2.14] that this lifts to an equivalence

$$\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}) \xrightarrow{\sim} \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}}),$$

and in particular, $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$ is an ∞–topos (which is not a priori clear for sites which are not essentially small).

Denote by $\mathbf{Aff}^{\mathrm{LFT}}_{\mathbb{C}}$ the category of affine schemes of finite type over $\mathbb{C}$. Note that it is a small category with finite limits. Denote by

$$(\,\cdot\,)_{\mathrm{an}}\colon \mathbf{Aff}^{\mathrm{LFT}}_{\mathbb{C}} \to \mathbf{Top}$$

the functor associating to such an affine scheme its space of $\mathbb{C}$–points, equipped with the analytic topology. The above functor preserves finite limits, and is the restriction of a functor defined for all algebraic spaces locally of finite type over $\mathbb{C}$; see [49, page 12]. Note also that if V is a scheme which is separated and locally of finite type, then V_{an} is locally (over any affine) a triangulated space by [30], so in particular V_{an} is locally contractible. Also observe that V_{an} is locally cut-out of $\mathbb{C}^n$ by polynomials, so it follows that V_{an} is in $\mathbf{Top}_{\mathbb{C}}$. Consequently $(\,\cdot\,)_{\mathrm{an}}$ restricts to a functor

$$(\,\cdot\,)_{\mathrm{an}}\colon \mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}} \to \mathbf{Top}_{\mathbb{C}},$$

which preserves finite limits.

Note that the category $\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}$ can be equipped with the Grothendieck topology generated by étale covering families. Denote the associated ∞–topos of ∞–sheaves on this site by $\mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$.

The following theorem is an extension of [36, Proposition 20.2]:

Theorem 3.1 *The functor*

$$(\,\cdot\,)_{\mathrm{an}}\colon \mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}} \to \mathbf{Top}_{\mathbb{C}}$$

lifts to a left exact colimit-preserving functor

$$(\,\cdot\,)_{\mathrm{top}}\colon \mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}).$$

Proof Note that the image under $(\,\cdot\,)_{\mathrm{an}}$ of an étale map is a local homeomorphism. Also note that if

$$S \to T$$

is a local homeomorphism and T is in $\mathbf{Top}_{\mathbb{C}}$, so is S. Furthermore, since the inclusion of any open subspace of a topological space is a local homeomorphism, and since any cover by local homeomorphisms can be refined by a cover by open subspaces, it follows that open covers and local homeomorphisms generate the same Grothendieck topology on $\mathbf{Top}_{\mathbb{C}}$. It follows that any ∞–sheaf on $\mathbf{Top}_{\mathbb{C}}$, so in particular any hypersheaf, satisfies descent with respect to covers by local homeomorphisms. The result now follows from [31, Proposition 6.2.3.20]. □

Remark 3.2 Denote by Y the Yoneda embedding

$$Y\colon \mathbf{Top}_{\mathbb{C}} \hookrightarrow \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$$

and denote by y the Yoneda embedding

$$y\colon \mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}} \hookrightarrow \mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}).$$

Explicitly, $(\,\cdot\,)_{\mathrm{top}}$ is the left Kan extension of $Y \circ (\,\cdot\,)_{\mathrm{an}}$ along y,

$$\mathrm{Lan}_y[Y \circ (\,\cdot\,)_{\mathrm{an}}]\colon \mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}),$$

or more concretely, it is the unique colimit-preserving functor such that for a representable $y(X)$, ie an affine scheme,

$$y(X)_{\mathrm{top}} \cong Y(X_{\mathrm{an}}).$$

By the proof of Theorem 3.1, we see that given any hypersheaf F on $\mathbf{Top}_{\mathbb{C}}$, the functor

$$F \circ (\,\cdot\,)_{\mathrm{an}}$$

is an ∞–sheaf on $(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$, ie we have a well-defined functor

$$(\,\cdot\,)_{\mathrm{an}}^{*}\colon \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}}) \to \mathrm{Sh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}).$$

Proposition 3.3 *The functor $(\,\cdot\,)_{\mathrm{top}}$ is left adjoint to $(\,\cdot\,)_{\mathrm{an}}^{*}$.*

Proof Since $(\,\cdot\,)_{\mathrm{top}}$ is colimit-preserving, it follows from [31, Corollary 5.5.2.9] that it has a right adjoint. Let us denote the right adjoint by R. By the Yoneda lemma, we have that if F is a hypersheaf F on $\mathbf{Top}_{\mathbb{C}}$, then $R(F)$ is the ∞–sheaf on $(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$ such that, if X is an affine scheme,

$$R(F)(X) \simeq \mathrm{Hom}(y(X), R(F)) \simeq \mathrm{Hom}((y(X))_{\mathrm{top}}, F) \simeq \mathrm{Hom}(Y(X_{\mathrm{an}}), F) \simeq F(X_{\mathrm{an}}).$$

□

Remark 3.4 The adjoint pair $(\,\cdot\,)_{\mathrm{top}} \dashv (\,\cdot\,)_{\mathrm{an}}^{*}$ assembles into a geometric morphism of ∞–topoi

$$f\colon \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}}) \to \mathrm{Sh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}),$$

with direct image functor

$$f_{*} = (\,\cdot\,)_{\mathrm{an}}^{*}$$

and inverse image functor

$$f^{*} = (\,\cdot\,)_{\mathrm{top}}.$$

Lemma 3.5 *Let $\mathbf{AlgSp}_{\mathbb{C}}^{\mathrm{LFT}}$ denote the category of algebraic spaces locally of finite type over $\mathbb{C}$. Equip $\mathbf{AlgSp}_{\mathbb{C}}^{\mathrm{LFT}}$ with the étale topology. Then restriction along the canonical inclusion*

$$\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}} \hookrightarrow \mathbf{AlgSp}_{\mathbb{C}}^{\mathrm{LFT}}$$

induces an equivalence of ∞–categories

$$\mathrm{Sh}_{\infty}(\mathbf{AlgSp}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \xrightarrow{\sim} \mathrm{Sh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}).$$

Proof The inclusion satisfies the conditions of the comparison lemma of [3, Exposé III], so we have an induced equivalence

$$\mathrm{Sh}(\mathbf{AlgSp}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \xrightarrow{\sim} \mathrm{Sh}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$$

between sheaves of sets. Since both sites have finite limits, the result now follows from [31, Proposition 6.4.5.4]. □

Proposition 3.6 *Let X be any algebraic space locally of finite type over $\mathbb{C}$. Then $X_{\mathrm{top}} \cong X_{\mathrm{an}}$.*

Proof Let $\mathcal{U}$ be the Grothendieck universe of small sets and let $\mathcal{V}$ be the Grothendieck universe of large sets with $\mathcal{U} \in \mathcal{V}$. Denote by $\widehat{\mathrm{Gpd}}_\infty$ the ∞–category of large ∞–groupoids, and denote by $\widehat{\mathbb{H}\mathrm{ypSh}}_\infty(\mathbf{Top}_\mathbb{C})$ the ∞–category of hypersheaves on $\mathbf{Top}_\mathbb{C}$ with values in $\widehat{\mathrm{Gpd}}_\infty$, and similarly let $\widehat{\mathrm{Sh}}_\infty(\mathbf{AlgSp}_\mathbb{C}^{\mathrm{LFT}}, \text{ét})$ denote the ∞–category of sheaves on the étale site of algebraic spaces with values in $\widehat{\mathrm{Gpd}}_\infty$. Then by the same proof as Theorem 3.1, by left Kan extension there is a $\mathcal{V}$–small colimit-preserving functor

$$L\colon \widehat{\mathrm{Sh}}_\infty(\mathbf{AlgSp}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \to \widehat{\mathbb{H}\mathrm{ypSh}}_\infty(\mathbf{Top}_\mathbb{C})$$

such that, for all representable sheaves $y(P)$ on $(\mathbf{AlgSp}_\mathbb{C}^{\mathrm{LFT}}, \text{ét})$,

$$L(y(P)) \cong Y(P_{\mathrm{an}}).$$

By [31, Remark 6.3.5.17], both inclusions

$$\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C}) \hookrightarrow \widehat{\mathbb{H}\mathrm{ypSh}}_\infty(\mathbf{Top}_\mathbb{C})$$

and

$$\mathrm{Sh}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \hookrightarrow \widehat{\mathrm{Sh}}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét})$$

preserve $\mathcal{U}$–small colimits. Hence both composites

$$\mathrm{Sh}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \hookrightarrow \widehat{\mathrm{Sh}}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \simeq \widehat{\mathrm{Sh}}_\infty(\mathbf{AlgSp}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \xrightarrow{L} \widehat{\mathbb{H}\mathrm{ypSh}}_\infty(\mathbf{Top}_\mathbb{C})$$

and

$$\mathrm{Sh}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \xrightarrow{(\,\cdot\,)_{\mathrm{top}}} \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C}) \hookrightarrow \widehat{\mathbb{H}\mathrm{ypSh}}_\infty(\mathbf{Top}_\mathbb{C})$$

are $\mathcal{U}$–small colimit-preserving, and agree up to equivalence on every representable $y(X)$, for X an affine scheme. It follows from [31, Theorem 5.1.5.6] that both compositions must in fact be equivalent. However, the inclusion

$$\mathrm{Sh}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \hookrightarrow \widehat{\mathrm{Sh}}_\infty(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét}) \simeq \widehat{\mathrm{Sh}}_\infty(\mathbf{AlgSp}_\mathbb{C}^{\mathrm{LFT}}, \text{ét})$$

carries an algebraic space P to its representable sheaf $y(P)$. The result follows. □

The next lemma follows immediately from the fact that $(\,\cdot\,)_{\mathrm{top}}$ preserves finite limits.

Lemma 3.7 *Let $\mathcal{G}$ be a groupoid object in sheaves of sets on the étale site $(\mathbf{Aff}_\mathbb{C}^{\mathrm{LFT}}, \text{ét})$. Then applying $(\,\cdot\,)_{\mathrm{top}}$ levelwise produces a groupoid object in sheaves of sets on $\mathbf{Top}_\mathbb{C}$, denoted by $\mathcal{G}_{\mathrm{top}}$. Moreover, if the original groupoid $\mathcal{G}$ is a groupoid object in algebraic spaces, then $\mathcal{G}_{\mathrm{top}}$ is degreewise representable, ie a topological groupoid.*

Proposition 3.8 *Let $\mathcal{G}$ be a groupoid object in sheaves of sets on the étale site $(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$. Denote by $[\mathcal{G}]$ its associated stack of torsors, and denote by $[\mathcal{G}_{\mathrm{top}}]$ the stack of groupoids on $\mathbf{Top}_{\mathbb{C}}$ associated to $\mathcal{G}_{\mathrm{top}}$, ie the stack on $\mathbf{Top}_{\mathbb{C}}$ of principal $\mathcal{G}_{\mathrm{top}}$–bundles. Then $[\mathcal{G}]_{\mathrm{top}} \simeq [\mathcal{G}_{\mathrm{top}}]$.*

Proof The stack $[\mathcal{G}]$ is the stackification of the presheaf of groupoids $\tilde{y}(\mathcal{G})$ which sends an affine scheme X to the groupoid $\mathcal{G}(X)$. Denote by $N(\mathcal{G})$ the simplicial presheaf which is the nerve of this presheaf of groupoids. Consider the diagram

$$\Delta^{\mathrm{op}} \xrightarrow{N(\mathcal{G})} \mathrm{Psh}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \mathit{Set}) \hookrightarrow \mathrm{Psh}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \mathrm{Gpd}_{\infty}).$$

We claim that the colimit of the above functor is $\tilde{y}(\mathcal{G})$. Since colimits are computed objectwise, it suffices to show that if $\mathcal{H}$ is any discrete groupoid, then $N(\mathcal{H})$ is the homotopy colimit of the diagram

$$\Delta^{\mathrm{op}} \xrightarrow{N(\mathcal{H})} \mathit{Set} \hookrightarrow \mathit{Set}^{\Delta^{\mathrm{op}}},$$

which follows easily from the well-known fact that the homotopy colimit of a simplicial diagram of simplicial sets can be computed by taking the diagonal. It follows then that $[\mathcal{G}]$ is the colimit of the diagram

$$\Delta^{\mathrm{op}} \xrightarrow{N(\mathcal{G})} \mathrm{Sh}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \hookrightarrow \mathrm{Sh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}),$$

since ∞–sheafification preserves colimits, as it is a left adjoint. By the same argument, we have that $[\mathcal{G}_{\mathrm{top}}]$ is the colimit of the diagram

$$\Delta^{\mathrm{op}} \xrightarrow{N(\mathcal{G}_{\mathrm{top}})} \mathrm{Sh}(\mathbf{Top}_{\mathbb{C}}) \hookrightarrow \mathrm{Sh}_{\infty}(\mathbf{Top}_{\mathbb{C}}).$$

Notice that for all n we have

$$N(\mathcal{G}_{\mathrm{top}})_n = (N(\mathcal{G})_n)_{\mathrm{top}}.$$

The result now follows from the fact that $(\cdot)_{\mathrm{top}}$ preserves colimits. □

Definition 3.9 A *topological stack* is a stack on $\mathbf{Top}_{\mathbb{C}}$ of the form $[\mathcal{G}]$ for $\mathcal{G}$ a groupoid object in $\mathbf{Top}_{\mathbb{C}}$. Denote the associated $(2,1)$–category of topological stacks by $\mathfrak{TopSt}$.

Remark 3.10 In the literature, typically there is no restriction on a topological stack to come from a topological groupoid which is locally contractible, and such a stack is represented by its functor of points on the Grothendieck site of all topological spaces. However, the $(2,1)$–category of topological stacks in the sense we defined above embeds fully faithfully into the larger $(2,1)$–category of all topological stacks in the classical sense.

Corollary 3.11 *The functor*

$$(\,\cdot\,)_{\mathrm{top}}\colon \mathrm{Sh}_\infty(\mathbf{Aff}^{\mathrm{LFT}}_{\mathbb{C}}, \text{ét}) \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$$

restricts to a left exact functor

$$(\,\cdot\,)_{\mathrm{top}} : \mathfrak{AlgSt}^{\mathrm{LFT}}_{\mathbb{C}} \to \mathfrak{TopSt}$$

from Artin stacks locally of finite type over $\mathbb{C}$ *to topological stacks.*

Up to the identification mentioned in Remark 3.10, the construction in the above corollary agrees with that of Noohi in [36, Section 20].

3.2 The fundamental infinity-groupoid of a stack

The following proposition will allow us to talk about homotopy types of topological stacks:

Proposition 3.12 *There is a colimit-preserving functor*

$$\overline{L}\colon \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}}) \to \mathcal{S}$$

sending every representable sheaf $y(T)$ *for* T *in* $\mathbf{Top}^s_{\mathbb{C}}$ *to its weak homotopy type.*

Proof The proof is essentially the same as [12, Proposition 3.1]. By Lemma 3.1 in [12], there is a functor

$$\mathbf{Top}^s_{\mathbb{C}} \hookrightarrow \mathbf{Top} \xrightarrow{h} \mathcal{S}$$

assigning to each space T its associated weak homotopy type. Denote this functor by π. Since $\mathbf{Top}^s_{\mathbb{C}}$ is essentially small, by left Kan extension there is a colimit-preserving functor

$$\mathrm{Lan}_Y\,\pi\colon \mathrm{Psh}_\infty(\mathbf{Top}^s_{\mathbb{C}}) \to \mathcal{S}$$

sending every representable presheaf $Y(T)$ to the underlying weak homotopy type of T. It follows from the Yoneda lemma that this functor has a right adjoint R_π which sends an ∞–groupoid Z to the ∞–presheaf

$$R_\pi(Z)\colon T \mapsto \mathrm{Hom}(\pi(T), Z).$$

We claim that $R_\pi(Z)$ is a hypersheaf. To see this, it suffices to observe that if

$$U^\bullet\colon \Delta^{\mathrm{op}} \to \mathbf{Top}^s_{\mathbb{C}}/T$$

is a hypercover of T with respect to the coverage of contractible open coverings, then the colimit of the composite

$$\Delta^{\mathrm{op}} \xrightarrow{U^\bullet} \mathbf{Top}^s_{\mathbb{C}}/T \to \mathbf{Top}^s_{\mathbb{C}} \xrightarrow{\pi} \mathcal{S}$$

is $\pi(T)$, which follows from [15, Theorem 1.3]. We thus have that $\mathrm{Lan}_y\,\pi$ and R_π restrict to an adjunction

$$\bar{L} \dashv \Delta$$

between $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}})$ and $\mathcal{S}$, so in particular, $\bar{L}$ preserves colimits. □

Corollary 3.13 *Let $\mathcal{G}$ be an ∞–groupoid. Denote by $\Delta(\mathcal{G})$ the constant presheaf on $\mathbf{Top}^s_{\mathbb{C}}$. Then $\Delta(\mathcal{G})$ is a hypersheaf.*

Proof Following the proof of the above theorem, we have that $R_\pi(\mathcal{G})$ is a hypersheaf. Moreover, for each space T in $\mathbf{Top}^s_{\mathbb{C}}$, we have that

$$R_\pi(\mathcal{G})(T) \simeq \mathrm{Hom}(Y(T), R_\pi(\mathcal{G})) \simeq \mathrm{Hom}(\bar{L}(Y(T)), \mathcal{G}) \simeq \mathrm{Hom}(*, \mathcal{G}) \simeq \mathcal{G},$$

since each such T is in fact contractible. □

Remark 3.14 The ∞–category $\mathcal{S}$ of spaces is the terminal ∞–topos. In particular, if $\mathscr{C}$ is any ∞–category equipped with a Grothendieck topology, then the unique geometric morphism

$$\mathrm{Sh}_\infty(\mathscr{C}) \to \mathcal{S}$$

has as direct image functor the global sections functor

$$\Gamma\colon \mathrm{Sh}_\infty(\mathscr{C}) \to \mathcal{S}$$

defined by $\Gamma(F) = \mathrm{Hom}(1, F)$, which is the same as $F(1)$ if $\mathscr{C}$ has a terminal object. The inverse image functor is given by

$$\Delta\colon \mathcal{S} \to \mathrm{Sh}_\infty(\mathscr{C})$$

and it sends an ∞–groupoid $\mathcal{G}$ to the sheafification of the constant presheaf with value $\mathcal{G}$. Similarly, the unique geometric morphism

$$\mathbb{H}\mathrm{ypSh}_\infty(\mathscr{C}) \to \mathcal{S}$$

has its direct image functor Γ given by the same construction as for ∞–sheaves, and the inverse image functor Δ assigns an ∞–groupoid $\mathcal{G}$ the hypersheafification of the constant presheaf with values $\mathcal{G}$. In either case we have $\Delta \dashv \Gamma$. In particular, Corollary 3.13 implies that for the ∞–topos $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}})$ we have a triple of adjunctions

$$\bar{L} \dashv \Delta \dashv \Gamma.$$

Although we will not prove it here, there is in fact a further right adjoint to Γ, $\mathit{coDisc} \vdash \Gamma$, and moreover the quadruple

$$\bar{L} \dashv \Delta \dashv \Gamma \dashv \mathit{coDisc}$$

exhibits $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}})$ as a *cohesive* ∞*–topos* in the sense of [44].

Proposition 3.15 *The composite*

$$\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}) \xrightarrow{\sim} \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}}) \xrightarrow{\bar{L}} \mathcal{S}$$

is colimit-preserving and sends a representable sheaf $Y(X)$ for X in $\mathbf{Top}_{\mathbb{C}}$ to its underlying weak homotopy type.

Proof By [12, Lemma 3.1], there is a functor

$$\mathbf{Top}_{\mathbb{C}} \hookrightarrow \mathbf{Top} \xrightarrow{h} \mathcal{S}$$

assigning to each space X its associated weak homotopy type. Denote this functor by Π. By exactly the same proof as Proposition 3.12, by using that $\mathbf{Top}_{\mathbb{C}}$ is $\mathcal{V}$–small, with $\mathcal{V}$ the Grothendieck universe of large sets, we construct a colimit-preserving functor

$$\mathbb{L}\colon \widehat{\mathbb{H}\mathrm{ypSh}}_\infty(\mathbf{Top}_{\mathbb{C}}) \to \widehat{\mathcal{S}},$$

where $\widehat{\mathcal{S}}$ is the ∞–category of large spaces (or large ∞–groupoids), which sends every representable sheaf $Y(X)$ to its underlying weak homotopy type. The rest of the proof is analogous to that of Proposition 3.6. $\square$

Definition 3.16 We denote the composite from Proposition 3.15 by

$$\Pi_\infty\colon \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}) \to \mathcal{S}.$$

For F a hypersheaf on $\mathbf{Top}_{\mathbb{C}}$, we call $\Pi_\infty(F)$ its *fundamental* ∞*–groupoid*.

Remark 3.17 In light of Remark 3.14, we have that $\Pi_\infty \dashv \Delta \dashv \Gamma$, where Γ is global sections, and Δ assigns an ∞–groupoid $\mathcal{G}$ the hypersheafification of the constant presheaf with value $\mathcal{G}$. In particular, we have a formula for $\Delta(\mathcal{G})$, namely, if X is a space in $\mathbf{Top}_{\mathbb{C}}$,

$$\Delta(\mathcal{G})(X) \simeq \mathrm{Hom}(\Pi_\infty(X), \mathcal{G}),$$

that is, the space of maps from the homotopy type of X to $\mathcal{G}$.

The following proposition may be seen as an extension of the results of [37]:

Proposition 3.18 *Let $\mathcal{G}$ be a groupoid object in $\mathbf{Top}_{\mathbb{C}}$ and denote by $[\mathcal{G}]$ denote the associated stack of groupoids on $\mathbf{Top}_{\mathbb{C}}$, ie the stack of principal $\mathcal{G}$–bundles. Then $\Pi_\infty([\mathcal{G}])$ has the same weak homotopy type as*

$$B\mathcal{G} = \|N(\mathcal{G})\|,$$

the fat geometric realization of the simplicial space arising as the topologically enriched nerve of $\mathcal{G}$.

Proof We know that $[\mathcal{G}]$ is the colimit in $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$ of the diagram

$$\Delta^{\mathrm{op}} \xrightarrow{N(\mathcal{G})} \mathbf{Top}_{\mathbb{C}} \overset{Y}{\hookrightarrow} \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$$

(as in the proof of Proposition 3.8). The result now follows from Proposition 3.15 and [12, Lemma 3.3] □

Lemma 3.19 *Let F be a hypersheaf on $\mathbf{Top}^s_{\mathbb{C}}$. Then $\overline{L}(F)$ is the colimit of F, ie the colimit of the diagram*

$$F\colon (\mathbf{Top}^s_{\mathbb{C}})^{\mathrm{op}} \to \mathcal{S}.$$

Proof By the proof of Proposition 3.12, $\overline{L}$ factors as the composition

$$\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}^s_{\mathbb{C}}) \hookrightarrow \mathrm{Psh}_\infty(\mathbf{Top}^s_{\mathbb{C}}) \xrightarrow{\mathrm{Lan}_Y \pi} \mathcal{S} = \mathrm{Gpd}_\infty .$$

Note however that every space in $\mathbf{Top}^s_{\mathbb{C}}$ is contractible, so the canonical morphism $\pi \to t$ to the terminal functor

$$t\colon \mathbf{Top}^s_{\mathbb{C}} \to \mathrm{Gpd}_\infty$$

(ie the functor with constant value the one-point set) is an equivalence, and hence $\mathrm{Lan}_Y \pi$ is left adjoint to the constant functor t^* which sends an ∞–groupoid $\mathcal{G}$ to the constant presheaf with value $\mathcal{G}$. Since $\mathrm{Psh}_\infty(\mathbf{Top}^s_{\mathbb{C}}) = \mathrm{Fun}((\mathbf{Top}^s_{\mathbb{C}})^{\mathrm{op}}, \mathrm{Gpd}_\infty)$, the result now follows from the universal property of $\underrightarrow{\mathrm{colim}}(\,\cdot\,)$. □

Corollary 3.20 *Let F be a hypersheaf on $\mathbf{Top}_{\mathbb{C}}$. Then $\Pi_\infty(F)$ is the colimit of $F|_{\mathbf{Top}^s_{\mathbb{C}}}$.*

3.3 The profinite homotopy type of a (pro-)stack

Let us define the profinite version of the homotopy type of a stack.

Definition 3.21 We denote the composite

$$\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}) \xrightarrow{\Pi_\infty} \mathcal{S} \xrightarrow{\widehat{(\cdot)}} \mathrm{Prof}(\mathcal{S})$$

by $\widehat{\Pi}_\infty$. For F a hypersheaf on $\mathbf{Top}_{\mathbb{C}}$, we call $\widehat{\Pi}_\infty(F)$ its *profinite fundamental ∞–groupoid* or simply its *profinite homotopy type*.

Let us extend the constructions of this section to pro-objects. Note that the functor

$$(\cdot)_{\mathrm{top}}\colon \mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$$

extends to a functor on pro-objects, which by abuse of notation we will denote by the same symbol,

$$(\cdot)_{\mathrm{top}}\colon \mathrm{Pro}(\mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})) \to \mathrm{Pro}(\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})).$$

We will now describe how to define the profinite homotopy type of a pro-object in $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$. First, we may extend the profinite fundamental ∞–groupoid functor on $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})$ to pro-objects. This can be achieved easily by the universal property of $\mathrm{Pro}(\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}))$. Indeed, consider the functor

$$\widehat{\Pi}_\infty\colon \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}}) \to \mathrm{Prof}(\mathcal{S})$$

and denote its unique cofiltered limit-preserving extension, by abuse of notation, again by

$$\widehat{\Pi}_\infty\colon \mathrm{Pro}(\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_{\mathbb{C}})) \to \mathrm{Prof}(\mathcal{S}).$$

Unwinding the definitions, we see that if $\varprojlim_{i\in\mathcal{I}} \mathcal{Y}^i$ is a pro-object in hypersheaves on $\mathbf{Top}_{\mathbb{C}}$, then its profinite homotopy type is

$$\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} \mathcal{Y}^i\big) = \varprojlim_{i\in\mathcal{I}} \widehat{\Pi}_\infty(\mathcal{Y}^i).$$

3.4 The homotopy type of Kato–Nakayama spaces

In this subsection, we will give a formula expressing the homotopy type of the Kato–Nakayama space of a log scheme in terms of algebro-geometric data. We first start by reviewing a functorial approach to Kato–Nakayama spaces which is due to Kato, Illusie and Nakayama. Let (X, M, α) be a log scheme, and let X_{an} be the analytification of X, which is an object of $\mathbf{Top}_{\mathbb{C}}$.

Consider the slice category $\mathbf{Top}_{\mathbb{C}}/X_{\mathrm{an}}$. If $(T \xrightarrow{p} X_{\mathrm{an}})$ is an object in $\mathbf{Top}_{\mathbb{C}}/X_{\mathrm{an}}$, one can pullback M to T and take the sectionwise group completion. In this way we obtain a sheaf of abelian groups on T, which we denote by p^*M^{gp}. Note that p^*M^{gp} contains $p^*\mathcal{O}_X^\times$ as a subsheaf of abelian groups.

Let G be any abelian topological group. If T is a topological space, we denote $\underline{G}_T$ the sheaf on T of continuous maps to G equipped with the group structure coming from addition in G. Note that we have $f^*(\underline{G}_S) = \underline{G}_T$.

Definition 3.22 We denote by $F_{\log}$ the presheaf of sets on $\mathbf{Top}_{\mathbb{C}}/X_{\mathrm{an}}$ that is defined on objects by the following assignment:

$$(T \xrightarrow{p} X_{\mathrm{an}}) \mapsto \left\{ \text{morphisms of sheaves } s\colon p^*M^{\mathrm{gp}} \to \underline{S}^1_T \text{ such that } s(f) = \frac{f}{|f|} \text{ for } f \in \mathcal{O}_X^\times \right\}.$$

Theorem 3.23 [23, Section 1.2] *The presheaf $F_{\log}$ is represented by $X_{\log}$.*

Since $X_{\log}$ is an object of $\mathbf{Top}_{\mathbb{C}}$, the functor $F_{\log}$ completely determines $X_{\log}$. Moreover, we can use this functorial description to give an expression for the homotopy type of $X_{\log}$, as we will now show.

Definition 3.24 Denote by $\mathscr{C}_{\mathrm{KN}}(X)$ the following category: the objects consist of triples (T, p, s) where

- T is a topological space in $\mathbf{Top}^s_{\mathbb{C}}$,
- $p\colon T \to X_{\mathrm{an}}$ is a continuous map,
- and s is a morphism of sheaves of abelian groups
$$s\colon p^*M^{\mathrm{gp}} \to \underline{S}^1_T$$
such that $s(f) = f/|f|$ for $f \in \mathcal{O}_X^\times$.

The morphisms $(T, p, s) \to (S, q, r)$ are continuous maps $f\colon T \to S$ such that $f^*(r) = s$.

Theorem 3.25 *Let X be a log scheme. The weak homotopy type of the Kato–Nakayama space is that of $B\mathscr{C}_{\mathrm{KN}}(X)$.*

Proof The reader may notice that $\mathscr{C}_{\mathrm{KN}}(X)$ is simply the Grothendieck construction

$$\int_{\mathbf{Top}^s_{\mathbb{C}}} (F_{\log}|_{\mathbf{Top}^s_{\mathbb{C}}/X_{\mathrm{an}}}).$$

Notice also that

$$\mathbf{Top}^s_{\mathbb{C}}/X_{\mathrm{an}} \to \mathbf{Top}^s_{\mathbb{C}}$$

is the Grothendieck construction of $Y(X_{\mathrm{an}})|_{\mathbf{Top}^s_{\mathbb{C}}}$ (where Y denotes the Yoneda embedding) ie the corresponding fibered category. Now, there is a canonical equivalence of categories

$$\mathrm{Sh}(\mathbf{Top}^s_{\mathbb{C}}/X_{\mathrm{an}}) \simeq \mathrm{Sh}(\mathbf{Top}^s_{\mathbb{C}})/Y(X_{\mathrm{an}})|_{\mathbf{Top}^s_{\mathbb{C}}},$$

and it follows that $\int_{\mathbf{Top}^s_{\mathbb{C}}} (F_{\log}|_{\mathbf{Top}^s_{\mathbb{C}}/X_{\mathrm{an}}})$ is equivalent to the Grothendieck construction of $Y(X_{\log})$. We have from Proposition 3.15 that $\Pi_\infty(Y(X_{\log}))$ is the weak homotopy type of $X_{\log}$. The result now follows from Corollary 3.20 and [12, Corollary 3.2]. □

Corollary 3.26 *Let X be a log scheme. The profinite homotopy type of its Kato–Nakayama space $X_{\log}$ is that of the profinite completion of $B\mathscr{C}_{\mathrm{KN}}(X)$.*

4 Construction of the map

In all that follows X will be a fine and saturated log scheme over $\mathbb{C}$ that is locally of finite type. See the appendix for a condensed introduction to the main concepts and notations that we will use in this section. Our goal is to prove the following proposition:

Proposition 4.1 *There is a canonical morphism of pro-topological stacks*

$$\Phi_X\colon X_{\log} \to (\sqrt[\infty]{X})_{\mathrm{top}}.$$

Later (Section 6) we will show that this map induces a weak equivalence of profinite homotopy types. The proof of Proposition 4.1 will take up the rest of this section.

Our strategy will be to construct the morphism Φ_X étale locally on X, where the log structure has a Kato chart, and then to show that the locally defined morphisms glue together to give a global one.

Step 1 (local case) First let us assume that $X \to \operatorname{Spec}\mathbb{C}[P]$ is a Kato chart for X, where P is a fine saturated sharp monoid. In this case everything is very explicit: as explained in Appendix A.4, there is an isomorphism $\sqrt[n]{X} \simeq [X_n/\mu_n(P)]$, where $X_n = X \times_{\operatorname{Spec}\mathbb{C}[P]} \operatorname{Spec}\mathbb{C}\big[\frac{1}{n}P\big]$, the group $\mu_n(P)$ is the Cartier dual of the cokernel of $P^{\mathrm{gp}} \to \frac{1}{n}P^{\mathrm{gp}}$, and the action on X_n is induced by the natural one on $\operatorname{Spec}\mathbb{C}\big[\frac{1}{n}P\big]$.

By following Noohi's construction (see Proposition A.19) we see that $\sqrt[n]{X}_{\mathrm{top}}$ is canonically isomorphic to the quotient $[(X_n)_{\mathrm{an}}/\mu_n(P)_{\mathrm{an}}]$, where $\mu_n(P)_{\mathrm{an}} \cong (\mathbb{Z}/n)^r$. Note that the finite morphism $\operatorname{Spec}\mathbb{C}\big[\frac{1}{n}P\big] \to \operatorname{Spec}\mathbb{C}[P]$ is étale on the open torus $\operatorname{Spec}\mathbb{C}[P^{\mathrm{gp}}] \subseteq \operatorname{Spec}\mathbb{C}[P]$, and ramified exactly on the complement.

Now let us construct a morphism of topological stacks $X_{\log} \to \sqrt[n]{X}_{\mathrm{top}}$. By the quotient stack description of the target, this is equivalent to giving a $\mu_n(P)_{\mathrm{an}}$–torsor (ie principal bundle) on $X_{\log}$, together with a $\mu_n(P)_{\mathrm{an}}$–equivariant map to $(X_n)_{\mathrm{an}}$.

Let us look at a couple of examples first.

Example 4.2 Let X be the standard log point $\operatorname{Spec}\mathbb{C}$ with log structure given by $\mathbb{N} \oplus \mathbb{C}^{\times} \to \mathbb{C}$ sending (n,a) to $0^n \cdot a$. Then $X_{\log} \cong S^1$, and $\sqrt[n]{X}_{\mathrm{top}} \simeq B(\mathbb{Z}/n)$. In this case the morphism $S^1 \to B(\mathbb{Z}/n)$ corresponds to the $(\mathbb{Z}/n)$–torsor $S^1 \to S^1$ defined by $z \mapsto z^n$.

Example 4.3 Let X be $\mathbb{A}^1$ with the divisorial log structure at the origin. Then $X_{\log} \cong \mathbb{R}_{\geq 0} \times S^1$ and $\sqrt[n]{X}_{\mathrm{top}} \simeq [\mathbb{C}/(\mathbb{Z}/n)]$, where the morphism $[\mathbb{C}/(\mathbb{Z}/n)] \to (\mathbb{A}^1)_{\mathrm{an}} = \mathbb{C}$ is induced by $z \mapsto z^n$, and $\mathbb{Z}/n$ acts by roots of unity.

In this case the map $\mathbb{R}_{\geq 0} \times S^1 \to [\mathbb{C}/(\mathbb{Z}/n)]$ corresponds to the $(\mathbb{Z}/n)$–torsor $\mathbb{R}_{\geq 0} \times S^1 \to \mathbb{R}_{\geq 0} \times S^1$ defined by $(a,b) \mapsto (a^n, b^n)$ and the equivariant map $\mathbb{R}_{\geq 0} \times S^1 \to \mathbb{C}$ given by $(a,b) \mapsto a \cdot b$.

Note that the map $\mathbb{R}_{\geq 0} \times S^1 \to \mathbb{R}_{\geq 0} \times S^1$ coincides with $z \mapsto z^n$ outside of the "origin" $\{0\} \times S^1$, and this is étale even on the algebraic side. Over the "origin", it is precisely the presence of the S^1 introduced by the Kato–Nakayama construction that allows the map to be a $(\mathbb{Z}/n)$–torsor. This is what happens in general (see also [28, Lemma 2.2]).

Proposition 4.4 *Consider the map $\phi_{\log}\colon (X_n)_{\log} \to X_{\log}$ induced by the morphism of log schemes $\phi\colon X_n \to X$. The map $\phi_{\log}$ is a $\mu_n(P)_{\mathrm{an}}$–torsor, and the projection $(X_n)_{\log} \to (X_n)_{\mathrm{an}}$ is a $\mu_n(P)_{\mathrm{an}}$–equivariant map.*

Note (see Definition A.9) that if P is a monoid, $\mathbb{C}(P)$ will denote the log analytic space $(\operatorname{Spec} \mathbb{C}[P])_{\mathrm{an}}$ with the induced natural log structure.

Proof The action of $\mu_n(P)$ on $\operatorname{Spec} \mathbb{C}\big[\frac{1}{n}P\big]$ induces an action on X_n, and the map $X_n \to X$ is invariant. Consequently we have an induced action of $\mu_n(P)_{\mathrm{an}}$ on $(X_n)_{\log}$, and the map $\phi_{\log}\colon (X_n)_{\log} \to X_{\log}$ is invariant.

Moreover, since taking $(\,\cdot\,)_{\log}$ commutes with strict base change (see Proposition A.12), we have a cartesian diagram

$$\begin{tikzcd}
(X_n)_{\log} \arrow[r] \arrow[d, "\phi_{\log}"'] & \mathbb{C}\big(\frac{1}{n}P\big)_{\log} \arrow[d, "\phi_{P,\log}"] \\
X_{\log} \arrow[r] & \mathbb{C}(P)_{\log}
\end{tikzcd}$$

and because the action of $\mu_n(P)_{\mathrm{an}}$ on $(X_n)_{\log}$ comes from the one on $\mathbb{C}\big(\frac{1}{n}P\big)_{\log}$, it suffices to prove the statement for the right-hand column.

Similarly, in order to verify that $(X_n)_{\log} \to (X_n)_{\mathrm{an}}$ is $\mu_n(P)_{\mathrm{an}}$–equivariant we are reduced to checking that $\mathbb{C}\big(\frac{1}{n}P\big)_{\log} \to \mathbb{C}\big(\frac{1}{n}P\big)$ is $\mu_n(P)_{\mathrm{an}}$–equivariant.

Now note that $\mu_n(P)_{\mathrm{an}}$ is precisely the kernel of $\operatorname{Hom}\big(\frac{1}{n}P, S^1\big) \to \operatorname{Hom}(P, S^1)$, so the action of $\mu_n(P)_{\mathrm{an}}$ on $\operatorname{Hom}\big(\frac{1}{n}P, S^1\big)$ is free and transitive. It is also not hard to check that there are local sections (note that $\operatorname{Hom}(P, S^1) = \operatorname{Hom}(P^{\mathrm{gp}}, S^1) \cong (S^1)^k$ non-canonically), so the map is a $\mu_n(P)_{\mathrm{an}}$–torsor.

Furthermore, $\phi_{P,\log}\colon \mathbb{C}\big(\frac{1}{n}P\big)_{\log} \to \mathbb{C}(P)_{\log}$ is the restriction map

$$\operatorname{Hom}\Big(\frac{1}{n}P, \mathbb{R}_{\geq 0} \times S^1\Big) \to \operatorname{Hom}(P, \mathbb{R}_{\geq 0} \times S^1),$$

and this is the product of the two maps $\mathrm{Hom}\big(\frac{1}{n}P, \mathbb{R}_{\geq 0}\big) \to \mathrm{Hom}(P, \mathbb{R}_{\geq 0})$ (which is a homeomorphism) and $\mathrm{Hom}\big(\frac{1}{n}P, S^1\big) \to \mathrm{Hom}(P, S^1)$. The action of $\mu_n(P)_{\mathrm{an}}$ is trivial on the factor $\mathrm{Hom}\big(\frac{1}{n}P, \mathbb{R}_{\geq 0}\big)$ and the one given by the aforementioned inclusion as a subgroup on the factor $\mathrm{Hom}\big(\frac{1}{n}P, S^1\big)$. Consequently, $\phi_{P,\log}$ is a $\mu_n(P)_{\mathrm{an}}$–torsor for the natural action, as required.

The map $\mathbb{C}\big(\frac{1}{n}P\big)_{\log} \to \mathbb{C}\big(\frac{1}{n}P\big)$ coincides with the map

$$\mathrm{Hom}\Big(\frac{1}{n}P, \mathbb{R}_{\geq 0} \times S^1\Big) \to \mathrm{Hom}\Big(\frac{1}{n}P, \mathbb{C}\Big)$$

induced by the natural map $\mathbb{R}_{\geq 0} \times S^1 \to \mathbb{C}$, and thus it is manifestly $\mathrm{Hom}\big(\frac{1}{n}P, S^1\big)$–equivariant, and in particular $\mu_n(P)_{\mathrm{an}}$–equivariant. $\square$

This proposition gives a morphism of pro-topological stacks $\Phi_{n,P}\colon X_{\log} \to \sqrt[n]{X}_{\mathrm{top}}$. It is clear from the construction that if $n \mid m$, then the diagram

$$\begin{tikzcd} X_{\log} \arrow[r, "\Phi_{m,P}"] \arrow[dr, "\Phi_{n,P}"'] & \sqrt[m]{X}_{\mathrm{top}} \arrow[d] \\ & \sqrt[n]{X}_{\mathrm{top}} \end{tikzcd}$$

is 2–commutative, so we obtain a morphism $(\Phi_X)_P\colon X_{\log} \to (\sqrt[\infty]{X})_{\mathrm{top}}$ of pro-topological stacks.

Step 2 (compatibility of the local constructions) Let us extend this local construction to a global one. The idea is of course to use descent and glue the local constructions, and intuitively, one would expect that these local maps patch together to define a global one without incident. However, writing down all the necessary 2–categorical coherences gets pretty technical quickly, and it is much cleaner to use the machinery of ∞–categories.

We will need some preliminary lemmas and constructions.

Lemma 4.5 *Let X be a fine saturated log scheme over a field k with two Kato charts $X \to \operatorname{Spec} k[P]$ and $X \to \operatorname{Spec} k[Q]$ for the log structure. Then for every geometric point x of X, after passing to an étale neighborhood of x, there is a third chart $X \to \operatorname{Spec} k[R]$ with maps of monoids $P \to R$ and $Q \to R$ inducing a commutative diagram*

$$\begin{tikzcd} & & \operatorname{Spec} k[P] \\ X \arrow[r] \arrow[urr, bend left] \arrow[drr, bend right] & \operatorname{Spec} k[R] \arrow[ur] \arrow[dr] & \\ & & \operatorname{Spec} k[Q] \end{tikzcd}$$

Proof We can take $R = \overline{M}_x$. There is a chart with monoid R in an étale neighborhood of x by Proposition A.6, and we have maps $P \to R$ and $Q \to R$ that induce a commutative diagram as in the statement, possibly after further localization. □

Now let us define a category $\mathfrak{I}$ of étale open subsets of X with a global chart: objects are triples $(\phi\colon U \to X, P, f)$ where $\phi\colon U \to X$ is étale, P is a fine saturated sharp monoid and $f\colon U \to \operatorname{Spec}\mathbb{C}[P]$ is a chart for the log structure on U (pulled back via ϕ).

A morphism $(\phi\colon U \to X, P, f) \to (\psi\colon V \to X, Q, g)$ is given by a (necessarily étale) map $U \to V$ over X and a morphism $Q \to P$ such that the diagram

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d] & \operatorname{Spec}\mathbb{C}[P] \arrow[d] \\ V \arrow[r, "g"] & \operatorname{Spec}\mathbb{C}[Q] \end{tikzcd}$$

is commutative.

We have two (lax) functors $(\cdot)_{\log}$ and $(\sqrt[n]{\cdot})_{\mathrm{top}}\colon \mathfrak{I} \to \mathfrak{TopSt}/X_{\mathrm{an}}$, as follows: for each $A = (\phi\colon U \to X, P, f) \in \mathfrak{I}$ we get, via strict pullback through the chart morphism, a local model for the Kato–Nakayama space $X^A_{\log}$ (over U) and one for the n^{th} root stack $\sqrt[n]{X}^A_{\mathrm{top}}$. We set $A_{\log} = X^A_{\log}$ and $\sqrt[n]{A}_{\mathrm{top}} = \sqrt[n]{X}^A_{\mathrm{top}}$. The maps to X_{an} are given by the composites of the projections to U_{an} and the local homeomorphism $U_{\mathrm{an}} \to X_{\mathrm{an}}$. The action of these two functors on morphisms is clear.

The construction in the local case (ie Step 1 above) gives an assignment, for each $A \in \mathfrak{I}$, of a morphism of topological stacks $\alpha^n_A\colon A_{\log} \to \sqrt[n]{A}_{\mathrm{top}}$.

Lemma 4.6 *The family (α^n_A) gives a lax natural transformation*

$$\alpha^n\colon (\cdot)_{\log} \Rightarrow (\sqrt[n]{\cdot})_{\mathrm{top}},$$

in the sense of [20, Appendix A].

Proof By translating the definition, in the present case this means the following: if $a\colon A = (\phi\colon U \to X, P, f) \to (\psi\colon V \to X, Q, g) = B$ is a morphism in $\mathfrak{I}$, then the diagram

$$\begin{tikzcd} A_{\log} \arrow[r, "\alpha^n_A"] \arrow[d] \arrow[dr, phantom, "\swarrow \alpha^n(a)"] & \sqrt[n]{A}_{\mathrm{top}} \arrow[d] \\ B_{\log} \arrow[r, "\alpha^n_B"'] & \sqrt[n]{B}_{\mathrm{top}} \end{tikzcd}$$

2–commutes, and the 2–cells $\alpha^n(a)$ satisfy a compatibility condition.

This follows from the fact that the morphism $a = (U \to V,\, Q \to P)$ gives a commutative diagram

$$
\begin{tikzcd}
& (U_n)_{\log} \arrow[rr] \arrow[dl] \arrow[dd] & & (U_n)_{\mathrm{an}} \arrow[dl] \\
(V_n)_{\log} \arrow[rr] \arrow[dd] & & (V_n)_{\mathrm{an}} & \\
& X^A_{\log} \arrow[dl] & & \\
X^B_{\log} & & &
\end{tikzcd}
$$

between the two objects corresponding to the functors α^n_A and α^n_B. This gives a canonical natural transformation that makes the diagram

$$
\begin{tikzcd}
X^A_{\log} \arrow[r, "\alpha^n_A"] \arrow[d] \arrow[dr, phantom, "\Swarrow \alpha^n(a)"] & {[(U_n)_{\mathrm{an}}/\mu_n(P)_{\mathrm{an}}] = \sqrt[n]{X}^A_{\mathrm{top}}} \arrow[d] \\
X^B_{\log} \arrow[r, "\alpha^n_B"] & {[(V_n)_{\mathrm{an}}/\mu_n(Q)_{\mathrm{an}}] = \sqrt[n]{X}^B_{\mathrm{top}}}
\end{tikzcd}
$$

2–commutative, and this is the required diagram.

Now if $C = (\eta\colon W \to X, R, h)$ is a third object of $\mathfrak{I}$ with a morphism $b\colon B \to C$ in $\mathfrak{I}$, then the fact that the diagram

$$
\begin{tikzcd}
& & (U_n)_{\log} \arrow[rr] \arrow[dl] \arrow[ddd] & & (U_n)_{\mathrm{an}} \arrow[dl] \\
& (V_n)_{\log} \arrow[rr] \arrow[dl] \arrow[ddd] & & (V_n)_{\mathrm{an}} \arrow[dl] & \\
(W_n)_{\log} \arrow[rr] \arrow[ddd] & & (W_n)_{\mathrm{an}} & & \\
& & X^A_{\log} \arrow[dl] & & \\
& X^B_{\log} \arrow[dl] & & & \\
X^C_{\log} & & & &
\end{tikzcd}
$$

commutes implies that the composite of the two 2–cells $\alpha^n(b)$ and $\alpha^n(a)$ is equal to $\alpha^n(b \circ a)$. $\square$

By composition with the natural functor

$$\mathfrak{TopSt}/X_{\mathrm{an}} \hookrightarrow \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}})/X_{\mathrm{an}}$$

to hypersheaves on $\mathbf{Top}_{\mathbb{C}}$ (see Section 3) and by abuse of notation we get a natural transformation of functors of ∞–categories:

$$\begin{tikzcd} \mathfrak{I} \arrow[r, bend left, "(\cdot)_{\log}"] \arrow[r, bend right, "\sqrt[n]{\cdot}_{\mathrm{top}}"'] \arrow[r, phantom, "\Downarrow \alpha^n"] & \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}})/X_{\mathrm{an}} \end{tikzcd}$$

Step 3 (the global case) We will now use the natural transformation α^n above to construct a global map

$$\Phi_X\colon X_{\log} \to \sqrt[n]{X}.$$

We will first need a crucial lemma:

Lemma 4.7 *Let $\iota\colon \mathfrak{I} \to \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}})$ be the functor $(\phi\colon U \to X, P, f) \mapsto U_{\mathrm{an}}$. Then the canonical map $\underrightarrow{\mathrm{colim}}\, \iota \to X_{\mathrm{an}}$ is an equivalence.*

Before proving the above lemma, we will show how we may use this lemma to produce the global morphism we seek. The key idea is the following basic fact about ∞–topoi:

Proposition 4.8 (colimits are universal) *Let $\underrightarrow{\mathrm{colim}}_{i \in I} A_i \to B$ be a morphism in an ∞–topos $\mathcal{E}$, and let $C \to B$ be another morphism. Then the canonical map*

$$\underset{i \in I}{\underrightarrow{\mathrm{colim}}}(C \times_B A_i) \to C \times_B \underset{i \in I}{\underrightarrow{\mathrm{colim}}}\, A_i$$

is an equivalence.

The above fact is standard and is an immediate consequence of the fact that any ∞–topos is locally cartesian closed.

Let us now see how we may complete the construction. Suppose we know that the canonical map $\underrightarrow{\mathrm{colim}}\, \iota \to X_{\mathrm{an}}$ is an equivalence. We can write this informally as

$$\underset{(\phi\colon U \to X, P, f)}{\underrightarrow{\mathrm{colim}}} U_{\mathrm{an}} \xrightarrow{\sim} X_{\mathrm{an}}.$$

Consider the morphism $X_{\log} \to X_{\mathrm{an}}$. Then since colimits are universal we have that the following is a pullback diagram:

$$\begin{tikzcd}
\varinjlim_{(\phi\colon U\to X,P,f)} U_{\mathrm{an}} \times_{X_{\mathrm{an}}} X_{\log} \arrow[r] \arrow[d] & X_{\log} \arrow[d] \\
\varinjlim_{(\phi\colon U\to X,P,f)} U_{\mathrm{an}} \arrow[r, "\sim"] & X_{\mathrm{an}}
\end{tikzcd}$$

It follows that the top map

$$\varinjlim_{(\phi\colon U\to X,P,f)} U_{\mathrm{an}} \times_{X_{\mathrm{an}}} X_{\log} \to X_{\log}$$

is also an equivalence. However, notice that we have a canonical identification

$$U_{\mathrm{an}} \times_{X_{\mathrm{an}}} X_{\log} \cong U_{\log},$$

hence

$$X_{\log} \simeq \varinjlim_{(\phi\colon U\to X,P,f)} U_{\log} = \varinjlim (\,\cdot\,)_{\log}.$$

By a completely analogous argument, one sees that

$$\sqrt[n]{X}_{\mathrm{top}} \simeq \varinjlim_{(\phi\colon U\to X,P,f)} \sqrt[n]{U}_{\mathrm{top}} = \varinjlim \sqrt[n]{\cdot}_{\mathrm{top}}.$$

For each n, the global map is then defined to be

$$\varinjlim \alpha^n\colon \varinjlim (\,\cdot\,)_{\log} \to \varinjlim \sqrt[n]{\cdot}_{\mathrm{top}}.$$

Just as in the local case, one easily sees that the maps

$$\varinjlim \alpha^n\colon X_{\log} \to \sqrt[n]{X}_{\mathrm{top}}$$

assemble into a morphism of pro-objects

$$\Phi_X\colon X_{\log} \to \sqrt[\infty]{X}_{\mathrm{top}}.$$

It is immediate from the construction that this map agrees locally with the map constructed in Step 1. In the next sections we will prove that Φ_X induces an equivalence of profinite spaces.

To finish the proof of the existence of the above map, it suffices to prove Lemma 4.7. Without further ado, we present the proof below.

Proof of Lemma 4.7 Equip $\mathfrak{I}$ with the following Grothendieck topology: a collection of morphisms

$$((\phi_i\colon U_i \to X, P_i, f_i) \to (\phi\colon U \to X, P, f))_i$$

will be a covering family if the induced family

$$(U_i \to U)_i$$

is an étale covering family. Note that there is a canonical morphism of sites

$$F\colon \mathfrak{I} \to X_{\text{ét}}$$

from $\mathfrak{I}$ to the small étale site of X. Moreover, by Lemma 4.5, one easily checks that F satisfies the conditions of the comparison lemma of [29, page 151], so the induced geometric morphism of topoi

$$\mathrm{Sh}(\mathfrak{I}) \to \mathrm{Sh}(X_{\text{ét}})$$

is an equivalence. It then follows from [26, Theorem 5; 31, Proposition 6.5.2.14] that the induced geometric morphism between the respective ∞–topoi of hypersheaves

$$\mathbb{H}\mathrm{ypSh}_{\infty}(\mathfrak{I}) \to \mathbb{H}\mathrm{ypSh}_{\infty}(X_{\text{ét}})$$

is an equivalence. By Remark 3.4, the analytification functor is the inverse image part of a geometric morphism

$$f\colon \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}}) \to \mathrm{Sh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}).$$

By [31, Proposition 6.5.2.13], there is an induced geometric morphism

$$\tilde{f}\colon \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}}) \to \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}).$$

By left Kan extension of the canonical functor

$$X_{\text{ét}} \to \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})/X$$

which sends each étale open $U \to X$ to itself, one produces a colimit-preserving functor

$$\omega\colon \mathbb{H}\mathrm{ypSh}_{\infty}(X_{\text{ét}}) \to \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})/X.$$

Consider the composite

$$\mathbb{H}\mathrm{ypSh}_{\infty}(\mathfrak{I}) \simeq \mathbb{H}\mathrm{ypSh}_{\infty}(X_{\text{ét}}) \xrightarrow{\omega} \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})/X \to \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét}) \xrightarrow{\tilde{f}^*} \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}}),$$

where $\mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})/X \to \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$ is the canonical projection. Denote the composite by Θ. The functor Θ is colimit-preserving as it is the composite of colimit-preserving functors, and, unwinding definitions, one sees that the composite

$$\mathfrak{I} \xrightarrow{y} \mathbb{H}\mathrm{ypSh}_{\infty}(\mathfrak{I}) \xrightarrow{\Theta} \mathbb{H}\mathrm{ypSh}_{\infty}(\mathbf{Top}_{\mathbb{C}})$$

is canonically equivalent to ι. It follows that there is a canonical equivalence

$$\varinjlim \iota \simeq \Theta(\varinjlim y).$$

But y is strongly generating, so by the proof of [10, Lemma 5.3.5] the colimit of y is the terminal object. Unwinding the definitions, one sees that the terminal object gets sent to X_{an} by Θ. This completes the proof. □

5 The topology of log schemes

This section contains preliminaries about some topological properties of fine saturated log schemes locally of finite type over $\mathbb{C}$, the Kato–Nakayama space and the root stacks.

5.1 Stratified fibrations

The following proposition is a consequence of the material in Appendix A.6.

Recall that if X is a fine saturated log scheme locally of finite type over $\mathbb{C}$, there is a stratification $\mathcal{R} = \{R_n\}_{n\in\mathbb{N}}$ of X, the *rank stratification* (Definition A.25), given by $R_n = \{x \in X \mid \operatorname{rank}_{\mathbb{Z}} \overline{M}^{\mathrm{gp}}_{\overline{x}} \geq n\}$.

Proposition 5.1 *The Kato–Nakayama space $X_{\log}$, the topological root stacks $\sqrt[m]{X}_{\mathrm{top}}$ and the topological infinite root stack $\sqrt[\infty]{X}_{\mathrm{top}}$ are stratified fibrations over X_{an} with respect to the stratification $\mathcal{R}$, ie they are fibrations over the strata $(S_n)_{\mathrm{an}} = (R_n \setminus R_{n+1})_{\mathrm{an}}$ of the stratification $\mathcal{R}_{\mathrm{an}}$.*

Proof All constructions are compatible with arbitrary base change along strict morphisms, so

$$X_{\log}|_{(S_n)_{\mathrm{an}}} \cong (S_n)_{\log},$$
$$\sqrt[m]{X}|_{S_n} \simeq \sqrt[m]{S_n},$$

where m can be ∞ and S_n has the log structure pulled back from X. It suffices then to show that the two maps $(S_n)_{\log} \to (S_n)_{\mathrm{an}}$ and $(\sqrt[m]{S_n})_{\mathrm{top}} \to (S_n)_{\mathrm{an}}$ are fibrations over S_n.

Let us cover $(S_n)_{\mathrm{an}}$ with open subsets over which the sheaf $\overline{M}$ is constant, and recall that by definition of S_n it will have rank n. We can choose such opens so that we have a cartesian diagram

$$\begin{array}{ccc} (S_n)_{\log} & \longrightarrow & (\operatorname{Spec} k[P])_{\log} \\ \downarrow & & \downarrow \\ (S_n)_{\mathrm{an}} & \longrightarrow & (\operatorname{Spec} k[P])_{\mathrm{an}} \end{array}$$

over each of them, where the bottom horizontal arrow sends everything to the vertex v_P (as in the proof of Proposition A.27). It follows that $(S_n)_{\log} \cong (S^1)^n \times (S_n)_{\mathrm{an}}$, and that the map $(S_n)_{\log} \to (S_n)_{\mathrm{an}}$ is identified with the projection. The factor $(S^1)^n$ is the fiber of the map $(\operatorname{Spec} k[P])_{\log} \to (\operatorname{Spec} k[P])_{\mathrm{an}}$ over the point v_P.

The analogous diagram

$$\begin{tikzcd} (\sqrt[m]{S_n})_{\mathrm{top}} \arrow[r] \arrow[d] & \sqrt[m]{\operatorname{Spec} k[P]}_{\mathrm{top}} \arrow[d] \\ (S_n)_{\mathrm{an}} \arrow[r] & (\operatorname{Spec} k[P])_{\mathrm{an}} \end{tikzcd}$$

shows the same conclusion for root stacks. In this case we get an isomorphism

$$(\sqrt[m]{S_n})_{\mathrm{top}} \simeq \mathcal{X} \times (S_n)_{\mathrm{an}},$$

where $\mathcal{X}$ is the fiber of the map $\sqrt[m]{\operatorname{Spec} k[P]}_{\mathrm{top}} \to (\operatorname{Spec} k[P])_{\mathrm{an}}$ over the vertex v_P. $\square$

We need a similar (local) statement for groupoid presentations of the root stacks.

Take $x \in X$, and an open étale neighborhood $U \to$X of x where there is a global chart $U \to \operatorname{Spec} \mathbb{C}[P]$ for the log structure, where P is fine, saturated and sharp. Then we have a quotient stack presentation for the topological n^{th} root stack $\sqrt[n]{U}_{\mathrm{top}} \simeq (\sqrt[n]{X}|_U)_{\mathrm{top}}$ for every n (see the discussion preceding Proposition A.18). Let us denote by $\mathbb{G}(n)$ the simplicial topological space associated with this quotient presentation. There are compatible maps $\mathbb{G}(m) \to \mathbb{G}(n)$ whenever $n \mid m$, and the whole system gives a (simplicial) presentation for the topological infinite root stack $\sqrt[\infty]{U}_{\mathrm{top}}$.

Explicitly, the simplicial space $\mathbb{G}(n)$ is obtained from the action of $\mu_n(P)$ on the scheme $U_n = U \times_{\operatorname{Spec} \mathbb{C}[P]} \operatorname{Spec} \mathbb{C}\big[\frac{1}{n} P\big]$ (see the local description of the root stacks in Appendix A.4), so that

$$\mathbb{G}(n)_k \cong (U_n \times \mu_n(P) \times \cdots \times \mu_n(P))_{\mathrm{an}},$$

where there are k copies of $\mu_n(P)$ and the map $\mathbb{G}(n)_k \to U_{\mathrm{an}}$ is the composite of the projection to $(U_n)_{\mathrm{an}}$ followed by the map $(U_n)_{\mathrm{an}} \to U_{\mathrm{an}}$.

Proposition 5.2 *Every $x \in U_{\mathrm{an}}$ has arbitrarily small neighborhoods over which, for every n and k, the map $\mathbb{G}(n)_k \to U_{\mathrm{an}}$ is a product over $U_{\mathrm{an}} \cap (S_r)_{\mathrm{an}}$, where $x \in (S_r)_{\mathrm{an}}$.*

In particular, for every n and k the topological space $\mathbb{G}(n)_k$ is a stratified fibration over U_{an}.

Proof Note first of all that since the map $U \to \operatorname{Spec} \mathbb{C}[P]$ is strict, the rank stratification of $\operatorname{Spec} \mathbb{C}[P]$ with its natural log structure is pulled back to the rank stratification of U, in the obvious sense.

Moreover, from the cartesian diagram

$$\begin{tikzcd} (U_n)_{\mathrm{an}} \arrow[r] \arrow[d] & \mathbb{C}\left(\frac{1}{n}P\right) \arrow[d] \\ U_{\mathrm{an}} \arrow[r] & \mathbb{C}(P) \end{tikzcd}$$

and from the fact that $\mathbb{G}(n)_k \to U_{\mathrm{an}}$ is the projection

$$\mathbb{G}(n)_k \cong (U_n \times \mu_n(P) \times \cdots \times \mu_n(P))_{\mathrm{an}} \to (U_n)_{\mathrm{an}}$$

followed by $(U_n)_{\mathrm{an}} \to U_{\mathrm{an}}$, we see that it suffices to prove that the map

$$\pi\colon \mathbb{C}\left(\frac{1}{n}P\right) = \left(\operatorname{Spec}\mathbb{C}\left[\frac{1}{n}P\right]\right)_{\mathrm{an}} \to \mathbb{C}(P) = (\operatorname{Spec}\mathbb{C}[P])_{\mathrm{an}}$$

is a stratified fibration. The proof will show that for a stratum S we can find an open subset $V \subseteq \mathbb{C}(P)$ such that the map π is a product over $V \cap S$ for *every* n.

Let us pick $\phi \in \mathbb{C}(P) = \operatorname{Hom}(P, \mathbb{C})$, and call $p_1, \dots, p_l$ the (finitely many) indecomposable elements of P (see [38, Proposition 2.1.2]). Assume (by reordering) that the first h of those get sent to 0 by ϕ, and the last ones are sent to nonzero complex numbers. Call r the rank of the group associated to the quotient $P/\langle p_i \mid i = h+1, \dots, l\rangle$ (ie the rank of the log structure of $\mathbb{C}(P)$ at ϕ).

The stratum of the rank stratification of $\mathbb{C}(P)$ to which ϕ belongs will then be S_r, the set of points of $\mathbb{C}(P)$ where the log structure has rank exactly equal to r. It is clear that ϕ actually belongs to the open subset S_ϕ of S_r consisting of the morphisms $\psi \in \operatorname{Hom}(P, \mathbb{C})$ such that $\psi(p_i) = 0$ for $1 \le i \le h$ and $\psi(p_i) \neq 0$ for $h < i \le l$.

Note also that the same condition on images of indecomposables of $\frac{1}{n}P$ will determine a subset $S'_\phi \subseteq \mathbb{C}\left(\frac{1}{n}P\right) = \operatorname{Hom}\left(\frac{1}{n}P, \mathbb{C}\right)$ (of those morphisms such that the image of p_i/n is zero exactly when $1 \le i \le h$) that a moment's reflection will show to be exactly the preimage $\pi^{-1}(S_\phi)$. Let us check that we can choose a neighborhood of ϕ in $\mathbb{C}(P)$ over which the restriction of $\pi\colon \pi^{-1}(S_\phi) \to S_\phi$ is a product.

For each $i = h+1, \dots, l$ let us choose a small open disk D_i around $\phi(p_i)$ in $\mathbb{C}$ that does not contain the origin, and for $i = 1, \dots, h$ let D_i be a small open disk around the origin. These define an open neighborhood W of ϕ in $\mathbb{C}(P)$, made up of those functions ψ such that $\psi(p_i) \in D_i$ for every i.

Let us also choose an n^{th} root $\sqrt[n]{\phi(p_i)}$ of the nonzero complex number $\phi(p_i)$ for $i = h+1, \dots, l$. There are a finite number of such choices, and there is a subset of those choices for which the homomorphism $\frac{1}{n}P \to \mathbb{C}$ given by $p_i/n \mapsto \sqrt[n]{\phi(p_i)}$ is well-defined (note that this assignment might not give a well-defined homomorphism due to the relations among the indecomposable elements of the monoid P). Let us call A this set of "good" choices.

Any element of A determines for each $i = h+1, \dots, l$ an n^{th} root function $\sqrt[n]{-}_i$ defined on the small disk D_i. Let us define a map $W \cap S_\phi \to \pi^{-1}(W \cap S_\phi) \subseteq \operatorname{Hom}\bigl(\frac{1}{n}P, \mathbb{C}\bigr)$ by sending ψ to the morphism defined by $p_i/n \mapsto \sqrt[n]{\psi(p_i)}_i$. This is a section of the projection $\pi^{-1}(W \cap S_\phi) \to W \cap S_\phi$, and one can check that this induces a homeomorphism $W \cap S_\phi \times A \cong \pi^{-1}(W \cap S_\phi)$, where A is viewed as a discrete set. We leave the details to the reader.

These arguments are uniform in $n \in \mathbb{N}$, so the open subset W that we identified will work for any n. □

5.2 A system of open neighborhoods for X_{an}

In this subsection we will prove the following crucial lemma:

Lemma 5.3 *For all $x \in X_{\text{an}}$ there exists a fundamental system of contractible analytic open neighborhoods $\mathcal{U}_x$ of x with global charts $f\colon U \to (\operatorname{Spec} \mathbb{C}[P])_{\text{an}}$ for $U \in \mathcal{U}_x$ such that*

(1) *the map f sends x into the* vertex *of $(\operatorname{Spec} \mathbb{C}[P])_{\text{an}}$ (ie the maximal ideal generated by all nonzero elements of P), and*

(2) *the maps*
$$(X_{\log})_x \to X_{\log}|_U$$
and
$$(\mathbb{G}(n)_i)_x \to (\mathbb{G}(n)_i)|_U$$
are weak homotopy equivalences, where $\{\mathbb{G}(n)\}_{n \in \mathbb{N}}$ is the family of topological groupoid presentations for the topological n^{th} root stack coming from the chart f, as in Proposition 5.2.

First of all we review some standard facts on triangulations and open covers. Let M be a topological space equipped with a triangulation $\mathcal{T}$. Denote by $\mathcal{V}_{\mathcal{T}}$ the set of vertices of $\mathcal{T}$. If f is a simplex of $\mathcal{T}$, we denote by $s(f)$ the union of the relative interiors of the simplices of $\mathcal{T}$ that contain f. We call $s(f)$ the *star* of f. Note that $s(f)$ is a contractible open subset of M. If v is a vertex of $\mathcal{T}$, we set $U_v := s(v)$. The star of a simplex f is naturally stratified by the simplices containing f: the strata are the relative interiors of the simplices containing f.

We say that a subspace of $\mathbb{R}^n$ is a cone if it is invariant under the action of $\mathbb{R}_{>0}$ by rescaling. We say that a cone is linear if it can be expressed as an intersection of finitely many linear spaces and linear half-spaces.

Lemma 5.4 *Let v be in $\mathcal{V}_{\mathcal{T}}$. Then there exists an $N \in \mathbb{N}$ such that U_v can be embedded as a linear cone in $\mathbb{R}^N$. Further we can choose this embedding in such a way that $s(f) \subset U_v$ is mapped to a linear subcone for all simplices f containing v.*

Proof Let $v_1, \dots, v_N$ be the one-dimensional simplices that contain v and let $e_1, \dots, e_N$ be the standard basis of $\mathbb{R}^N$. If I is a subset of $\{1, \dots, N\}$ we write

$$O_I := \Big\{ \sum_{i \in I} \alpha_i e_i \mid \alpha_i \geq 0 \Big\} \subset \mathbb{R}^N.$$

Every simplex Δ containing v determines a subset I_Δ of $\{1, \dots, N\}$ in the following way: i belongs to I_Δ if and only if Δ contains v_i. We obtain an embedding of U_v into $\mathbb{R}^N$ by considering a piecewise linear homeomorphism

$$U_v \simeq \bigcup_{v \in \Delta} O_{I_\Delta}.$$

This embedding has all the properties claimed by the lemma. □

Lemma 5.5 *Let x be in M, and let f be the lowest dimensional simplex such that x belongs to f. Then there exists a system of open neighborhoods $\mathcal{U}_x$ of x such that all U in $\mathcal{U}_x$ have the following properties:*

(1) *U is contractible.*

(2) *U does not intersect simplices of $\mathcal{T}$ that do not contain f.*

Proof Let v be a vertex incident to f. By Lemma 5.4 the open neighborhood U_v can be embedded as a linear cone in $\mathbb{R}^n$ in such a way that $s(f) \subset U_v$ is a linear subcone. Equip $\mathbb{R}^n$ with a Euclidean metric. Then $\mathcal{U}_x$ can be obtained by intersecting $s(f)$ with a system of open neighborhoods given by open balls in $\mathbb{R}^n$ centered at x. □

Next we turn to the log scheme X. Let x be in X_{an}. Since we are interested in constructing a system of open neighborhoods for x we can assume, by étale localizing around x, that

- X is affine, and
- we have a global chart $f\colon X \to \operatorname{Spec} \mathbb{C}[P]$, where $P = \overline{M}_x$ (see Proposition A.6), which sends x to the vertex of $\operatorname{Spec} \mathbb{C}[P]$.

The fact that X is affine is key in order to produce triangulations, which we do in Lemma 5.6.

By Lemma A.23 the log structure determines a stratification $\mathcal{R}_X$ of X.

Lemma 5.6 *There exists a triangulation $\mathcal{T}_X$ of X that refines $\mathcal{R}_X$.*

Proof The existence of triangulations refining stratifications of affine schemes goes back to Lojasiewicz [30]. See also Shiota's work [46] for a more recent reference. □

By Lemma 5.5, the triangulation $\mathcal{T}_X$ gives us a system of open neighborhoods $\mathcal{U}_x$ of x in X_{an} satisfying the two properties stated there. We claim that $\mathcal{U}_x$ has all the properties required by Lemma 5.3. Note that, since we assumed without loss of generality that X is affine and has a global chart to $\operatorname{Spec}\mathbb{C}[P]$ sending x to the vertex of $\operatorname{Spec}\mathbb{C}[P]$, we only need to prove that property (2) holds. We do this next.

The following lemma was proved in [43]:

Lemma 5.7 [43, Lemma 3.25] *Let W_1 and W_2 be locally compact and locally contractible Hausdorff spaces. Let $p\colon W_1 \to W_2$ be a continuous map, and let $K_2 \subset W_2$ be a closed deformation retract. Suppose that the restriction $p^{-1}(W_2 \setminus K_2) \to W_2 \setminus K_2$ is homeomorphic to the projection from a product $F \times (W_2 \setminus K_2) \to W_2 \setminus K_2$. Then $K_1 := p^{-1}(K_2)$ is a deformation retract of W_1.*

We will actually need a slight variant of Lemma 5.7. Assume that $W_2 \setminus K_2$ decomposes as a finite disjoint union of m components, which we denote by $(W_2 \setminus K_2)_i$,

$$W_2 \setminus K_2 = \bigcup_{i=1}^{m} (W_2 \setminus K_2)_i.$$

Then the claim still holds if the restriction $p^{-1}(W_2 \setminus K_2) \to W_2 \setminus K_2$ is homeomorphic to the projection from a disjoint union of products

$$\bigcup_{i=1}^{m} F_i \times (W_2 \setminus K_2)_i \to \bigcup_{i=1}^{m} (W_2 \setminus K_2)_i.$$

This stronger statement is proved exactly as Lemma 5.7, and, in fact, follows from it through an induction on the number of connected components of $W_2 \setminus K_2$.

We conclude the proof of Lemma 5.3 by showing that the following proposition holds:

Proposition 5.8 *For all U in $\mathcal{U}_x$, each of the maps*

$$(X_{\log})_x \to X_{\log}|_U, \quad (\mathbb{G}(n)_i)_x \to (\mathbb{G}(n)_i)|_U$$

is a weak homotopy equivalence, where $\{\mathbb{G}(n)\}_{n \in \mathbb{N}}$ is the family of topological groupoid presentations for the topological n^{th} root stack coming from the chart f.

Proof The proof is the same for both $X_{\log}$ and $\mathbb{G}(n)_i$. The argument relies exclusively on the fact that $X_{\log}$ and $\mathbb{G}(n)_i$ give stratified fibrations on X_{an} with respect to the stratification $\mathcal{R}_X$. To avoid repetition, we prove the statement only for $X_{\log}$ but the argument remains valid if we substitute $\mathbb{G}(n)_i$ in all occurrences of $X_{\log}$.

Let f be the lowest-dimensional simplex of $\mathcal{T}_X$ such that x lies on f. Recall from the proof of Lemma 5.5 that, in order to define $\mathcal{U}_x$, we pick a vertex v of the triangulation $\mathcal{T}_X$ that is incident to f. By construction, U is an open subset of U_v. Thus U carries a stratification which is obtained by restricting to it the stratification on U_v by the simplices containing v.

For all $k \in \mathbb{N}$, denote by $U_k \subset U$ the k–skeleton of U: that is, U_k is the union of the strata of dimension less than or equal to k. Note that U_k is empty if $k < \dim(f)$ and is contractible if $\dim(f) \leq k$. Further, U_k is a strong deformation retract of $U_{k'}$ if $\dim(f) \leq k \leq k'$. Indeed both U_k and $U_{k'}$ are CW complexes (up to compactifying), and any contractible subcomplex of a contractible CW complex is a strong deformation retract, see eg [34, Lemma 1.6].

We prove next that if $\dim(f) \leq k-1$, the map

$$X_{\log}|_{U_{k-1}} \to X_{\log}|_{U_k}$$

is a deformation retract. Note that $U_k \setminus U_{k-1}$ is equal to the disjoint union of k–dimensional strata. That is, $U_k \setminus U_{k-1}$ can be written as a disjoint union of m components,

$$U_k \setminus U_{k-1} = \bigcup_{i=1}^{m} (U_k \setminus U_{k-1})_i .$$

The restriction of the map $X_{\log}|_U \to X_{\mathrm{an}}|_U$ to each stratum of U is a principal bundle. Indeed, the stratification on U is finer that the restriction to U of $\mathcal{R}_X$. Further, it is a trivializable principal bundle, since the strata are paracompact Hausdorff and contractible.

Thus the restriction

$$X_{\log}|_{U_k \setminus U_{k-1}} \to U_k \setminus U_{k-1}$$

is homeomorphic to a projection from a disjoint union of products

$$X_{\log}|_{U_k \setminus U_{k-1}} \simeq \bigcup_{i=1}^{m} F_i \times (U_k \setminus U_{k-1})_i \to \bigcup_{i=1}^{m} (U_k \setminus U_{k-1})_i .$$

We have showed that the map $U_{k-1} \to U_k$ is a deformation retract. We apply Lemma 5.7, or rather the variant that was discussed immediately after the statement of Lemma 5.7, (note that $X_{\log}$ is locally compact Hausdorff and locally contractible by Proposition A.13), and deduce that the map

$$X_{\log}|_{U_{k-1}} \to X_{\log}|_{U_k}$$

is also a deformation retract, as we claimed.

There exists an $N \in \mathbb{N}$ such that $U_N = U$. By applying recursively the retractions that we have constructed in the previous paragraph, we obtain a deformation retract

$$X_{\log}|_{U_{\dim(f)}} \to X_{\log}|_U.$$

By property (2) of Lemma 5.5, $U_{\dim(f)}$ is connected. Further, it is contractible and paracompact, and thus $X_{\log}|_{U_{\dim(f)}}$ is homeomorphic to a product $F \times U_{\dim(f)}$. This implies that there are homotopy equivalences

$$(X_{\log})|_x \simeq F \times \{x\} \xhookrightarrow{\sim} F \times U_{\dim(f)} \simeq X_{\log}|_{U_{\dim(f)}},$$

and this concludes the proof. □

6 The equivalence

At last, in this section we will prove the main result of this paper, namely that there is an equivalence

$$\widehat{\Pi}_\infty(\Phi_X)\colon \widehat{\Pi}_\infty(X_{\log}) \to \widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}})$$

of profinite spaces, where $\widehat{\Pi}_\infty$ is the "profinite homotopy type" functor defined in Section 3.3 and Φ_X is the morphism of pro-topological stacks constructed in Section 4.

The main idea is to use the basis of open subsets constructed in Lemma 5.3 to produce a suitable hypercover of X_{an} and to use this to reduce to checking that one has a profinite homotopy equivalence along fibers. First, we will need a few more technical lemmas.

The following lemma makes precise in what way one can glue profinite spaces together using hypercovers:

Lemma 6.1 *Let $\mathcal{X}$ be a hypersheaf in $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C})$. Let $\mathcal{I}$ be a cofiltered ∞–category and let*

$$f_\bullet\colon \mathcal{I} \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C})/\mathcal{X}$$

be an $\mathcal{I}$–indexed pro-system with associated pro-object $\varprojlim_{i\in\mathcal{I}}(f_i\colon \mathcal{Y}^i \to \mathcal{X})$. Let

$$U^\bullet\colon \Delta^{\mathrm{op}} \to \mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C})/\mathcal{X}$$

be a hypercover of $\mathcal{X}$. For each i, denote by $f_i^*U^\bullet$ the pullback of the hypercover $U^\bullet$ to a hypercover of $\mathcal{Y}^i$. Consider the underlying pro-object $\varprojlim_{i\in\mathcal{I}} \mathcal{Y}^i$ in $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C})$. Then there is a canonical equivalence of profinite spaces

$$\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} \mathcal{Y}^i\big) \simeq \varinjlim_{n\in\Delta^{\mathrm{op}}}\big[\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} f_i^*U^n\big)\big],$$

where

$$\widehat{\Pi}_\infty\colon \mathrm{Pro}(\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C})) \to \mathrm{Prof}(\mathcal{S})$$

is the functor constructed in Section 3.3.

Proof It suffices to show that for every π–finite space V there is a canonical equivalence

$$\mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}\big(\varinjlim_{n\in\Delta^{\mathrm{op}}}\big[\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} f_i^*U^n\big)\big], j(V)\big) \simeq \mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}\big(\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} \mathcal{Y}^i\big), j(V)\big)$$

which is natural in V. We have that

$$\mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}\big(\varinjlim_{n\in\Delta^{\mathrm{op}}}\big[\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} f_i^*U^n\big)\big], j(V)\big) \simeq \varprojlim_{n\in\Delta}\big[\varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty f_i^*U^n, V)\big].$$

Notice that V is k–truncated for some k, and hence so is $j(V)$ by Proposition 2.13. Since filtered colimits of k–truncated spaces are k–truncated, it follows that, for all n,

$$\varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty f_i^*U^n, V)$$

is k–truncated. By Lemma 2.21, it then follows that

$$\mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}\big(\varinjlim_{n\in\Delta^{\mathrm{op}}}\big[\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}} f_i^*U^n\big)\big], j(V)\big) \simeq \varprojlim_{n\in\Delta_{\le k}}\big[\varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty f_i^*U^n, V)\big].$$

By using that filtered colimits commute with finite limits, we then have that this is in turn equivalent to

$$\varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\big[\varprojlim_{n\in\Delta_{\le k}}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty f_i^*U^n, V)\big].$$

Again by Lemma 2.21 this is equivalent to

$$\varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\big[\varprojlim_{n\in\Delta}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty f_i^*U^n, V)\big].$$

Finally, we have the following string of natural equivalences:

$$\begin{aligned}\varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\big[\varprojlim_{n\in\Delta}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty f_i^*U^n, V)\big] &\simeq \varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\mathrm{Hom}_{\mathcal{S}}\big(\varinjlim_{n\in\Delta^{\mathrm{op}}}\Pi_\infty f_i^*U^n, V\big)\\ &\simeq \varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\mathrm{Hom}_{\mathcal{S}}\big(\Pi_\infty \varinjlim_{n\in\Delta^{\mathrm{op}}} f_i^*U^n, V\big)\\ &\simeq \varinjlim_{i\in\mathcal{I}^{\mathrm{op}}}\mathrm{Hom}_{\mathcal{S}}(\Pi_\infty \mathcal{Y}^i, V)\\ &\simeq \mathrm{Hom}_{\mathrm{Prof}(\mathcal{S})}\big(\widehat{\Pi}_\infty\big(\varprojlim_{i\in\mathcal{I}}\mathcal{Y}^i\big), j(V)\big).\end{aligned}$$ □

Let X be a log scheme. Denote by $\mathcal{U}$ the basis of contractible open subsets of X_{an} given by Lemma 5.3.

Lemma 6.2 *There is a hypercover*

$$U^\bullet\colon \Delta^{\mathrm{op}} \to \mathbf{Top}_{\mathbb{C}}/X_{\mathrm{an}}$$

such that, for all n, the map $U^n \to X_{\mathrm{an}}$ is isomorphic to the coproduct of inclusions of open neighborhoods in the basis $\mathfrak{U}$ and all the structure maps are local homeomorphisms.

Proof Using standard techniques, since $\mathfrak{U}$ is a basis for the topology of X_{an} we can construct a split hypercover satisfying the above by induction (see [13]). □

Remark 6.3 The image under the Yoneda embedding of the hypercover of topological spaces $U^\bullet$ just constructed is a hypercover of $Y(X_{\mathrm{an}})$ in the ∞–topos $\mathbb{H}\mathrm{ypSh}_\infty(\mathbf{Top}_\mathbb{C})$. We will abuse notation by identifying the two.

We now prove our main result:

Theorem 6.4 *Let X be a fine saturated log scheme locally of finite type over $\mathbb{C}$. The induced map*

$$\widehat{\Pi}_\infty(\Phi_X)\colon \widehat{\Pi}_\infty(X_{\log}) \xrightarrow{\sim} \widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}})$$

is an equivalence of profinite spaces.

Proof Consider now the hypercover $U^\bullet$ of X_{an} just constructed. Then each $U^n = \coprod_\alpha V_\alpha$, where each V_α is in $\mathfrak{U}$. Let us restrict to one such $V = V_\alpha$. Since V is in $\mathfrak{U}$, there exists an $x \in V$ such that $(X_{\log})_x \to X_{\log}|_V$ is a weak homotopy equivalence, and such that there is a Kato chart $U \to \operatorname{Spec}\mathbb{C}[P]$, with $U \to X$ étale, such that $U_{\mathrm{an}} \to X_{\mathrm{an}}$ admits a section σ over V and with the property that the composite

$$V \xrightarrow{\sigma} U_{\mathrm{an}} \to (\operatorname{Spec}\mathbb{C}[P])_{\mathrm{an}}$$

carries x to the vertex point of the toric variety $\operatorname{Spec}\mathbb{C}[P]$. Let us fix this x, and call it the *center* of V. Suppose that the monoid P has rank k; then the log structure at x also has rank k. Moreover, the fiber of the map

$$V_n := V \times_{(\operatorname{Spec}\mathbb{C}[P])_{\mathrm{an}}} \left(\operatorname{Spec}\mathbb{C}\left[\frac{1}{n}P\right]\right)_{\mathrm{an}} \to V$$

over x consists of a single point (see [24, Lemma 1.2]).

Let us fix an n; then we have that

$$\sqrt[n]{X}_{\mathrm{top}}|_V \simeq [(\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n] = [V_n/(\mathbb{Z}/n\mathbb{Z})^k].$$

Hence our groupoid presentation $\mathbb{G}(n)$ for $\sqrt[n]{X}_{\mathrm{top}}|_V$ guaranteed by Proposition 5.2 is the topological action groupoid $(\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n$. This groupoid admits a continuous functor to V (viewing V as a topological groupoid with only identity arrows) which on objects is simply the canonical map $V_n \to V$. Similarly, regard the one-point space $*$ also as a topological groupoid, and consider the canonical map

$$* \to V$$

picking out x. Since V and $*$ have no nonidentity arrows, the lax fibered product of topological groupoids

$$* \times_V^{(2,1)} ((\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n)$$

is equivalent to the strict fibered product

$$* \times_V ((\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n),$$

which is canonically equivalent to the action groupoid

$$(\mathbb{Z}/n\mathbb{Z})^k \ltimes (V_n)_x,$$

where $(V_n)_x$ is the fiber over $V_n \to V$. Since this fiber consists of a single point, we conclude that the lax fibered product may be identified with $(\mathbb{Z}/n\mathbb{Z})^k$, where we are identifying the group $(\mathbb{Z}/n\mathbb{Z})^k$ with its associated 1–object groupoid.

Consider the continuous functor of topological groupoids

$$(\mathbb{Z}/n\mathbb{Z})^k \simeq * \times_V^{(2,1)} ((\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n) \to (\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n.$$

This induces a map of simplicial topological spaces between their simplicially enriched nerves

$$N((\mathbb{Z}/n\mathbb{Z})^k) \to N((\mathbb{Z}/n\mathbb{Z})^k \ltimes V_n).$$

By Lemma 5.3, this map is degreewise a weak homotopy equivalence. It follows from Proposition 3.18 and [12, Lemma 3.2] that the induced map

$$B((\mathbb{Z}/n\mathbb{Z})^k) \simeq \Pi_\infty((\mathbb{Z}/n\mathbb{Z})^k \ltimes *) \to \Pi_\infty\big([(\mathbb{Z}/n\mathbb{Z})^k \ltimes V]\big) \simeq \Pi_\infty(\sqrt[n]{X}_{\mathrm{top}}|_V)$$

is an equivalence in $\mathcal{S}$. Since the topological groupoid presentations for $\sqrt[n]{X}_{\mathrm{top}}$ constructed in Section 5.1 are compatible with the natural maps $\sqrt[m]{X}_{\mathrm{top}} \to \sqrt[n]{X}_{\mathrm{top}}$ when $n\,|\,m$, it follows that we have a natural identification

$$\widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}}|_V) \simeq \varprojlim_n B((\mathbb{Z}/n\mathbb{Z})^k)$$

in $\mathrm{Prof}(\mathcal{S})$. Consider the pro-system of finite groups

$$n \mapsto (\mathbb{Z}/n\mathbb{Z})^k.$$

This is the k^{th} cartesian power of the pro-system

$$n \mapsto (\mathbb{Z}/n\mathbb{Z}),$$

which is simply $\widehat{\mathbb{Z}}$. By Proposition 2.22, it follows that

$$\widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}}|_V) \simeq B(\widehat{\mathbb{Z}}^k),$$

and hence, by Proposition 2.20, we have that
$$\widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}}|_V) \simeq B(\widehat{\mathbb{Z}^k}).$$
We also have that
$$(X_{\log})_x \cong (S^1)^k.$$
It follows that
$$\Pi_\infty(X_{\log}|_V) \simeq \Pi_\infty(S^1)^k \simeq B(\mathbb{Z}^k),$$
and so
$$\widehat{\Pi}_\infty(X_{\log}|_V) \simeq \widehat{B(\mathbb{Z}^k)}.$$

Since $\mathbb{Z}^k$ is a finitely generated free abelian group, it is *good* in the sense of Serre [45]. It follows from [41, Proposition 3.6] and Theorem 2.19 that the canonical map
$$\widehat{B(\mathbb{Z}^k)} \to B(\widehat{\mathbb{Z}^k})$$
is an equivalence of profinite spaces, hence
$$\widehat{\Pi}_\infty(X_{\log}|_V) \simeq B(\widehat{\mathbb{Z}^k}).$$
It now follows that
$$\widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}}|_V) \simeq \widehat{\Pi}_\infty(X_{\log}|_V),$$
which is a local version of our statement.

Now let us globalize using the hypercover $U^\bullet$. For each n, denote by q_n the natural map
$$q_n\colon \sqrt[n]{X}_{\mathrm{top}} \to X_{\mathrm{an}}.$$

Since $\widehat{\Pi}_\infty$ preserves colimits, it follows that the induced map
$$\varinjlim_{l\in\Delta^{\mathrm{op}}} \widehat{\Pi}_\infty \circ \tau^* U^l \to \varinjlim_{l\in\Delta^{\mathrm{op}}} \bigl(\widehat{\Pi}_\infty \circ \varprojlim_n q_n^* U^l\bigr)$$
is an equivalence of profinite spaces, where τ is the canonical map $\tau\colon X_{\log} \to X_{\mathrm{an}}$. However,
$$\varinjlim_{l\in\Delta^{\mathrm{op}}} \widehat{\Pi}_\infty \circ \tau^* U^l \simeq \widehat{\Pi}_\infty\bigl(\varinjlim_{l\in\Delta^{\mathrm{op}}} \tau^* U^l\bigr) \simeq \widehat{\Pi}_\infty(X_{\log}),$$
since $\tau^* U^\bullet$ is a hypercover of $X_{\log}$. Finally, by Lemma 6.1,
$$\varinjlim_{l\in\Delta^{\mathrm{op}}} \bigl(\widehat{\Pi}_\infty \circ \varprojlim_n q_n^* U^l\bigr) \simeq \widehat{\Pi}_\infty\bigl(\varprojlim_n \sqrt[n]{X}_{\mathrm{top}}\bigr) = \widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}}). \qquad \square$$

7 The profinite homotopy type of a log scheme

We conclude this paper by defining the profinite homotopy type of an arbitrary log scheme over a ground ring k, by using the notion of étale homotopy type.

Étale homotopy theory, as originally introduced by Artin and Mazur [5], is a way of associating to a suitably nice scheme a pro-homotopy type. In this seminal work they proved a generalized Riemann existence theorem:

Theorem 7.1 [5, Theorem 12.9] *Let X be scheme of finite type over $\mathbb{C}$; then the profinite completion of the étale homotopy type of X agrees with the profinite completion of X_{an}.*

In light of the above theorem, the étale homotopy type of a complex scheme of finite type gives a way of accessing homotopical information about its analytic topology by using only algebro-geometric information, and for a setting where the analytic topology is not available, such as a scheme over an arbitrary base, the profinite completion of its étale homotopy type serves as a suitable replacement.

In the original work of Artin and Mazur, for X a locally Noetherian scheme, one associates a pro-object in the homotopy category of spaces $\mathrm{Ho}(\mathcal{S})$. This definition was later refined by Friedlander [17] to produce a pro-object in the category $\mathit{Set}^{\Delta^{\mathrm{op}}}$ of simplicial sets, and a generalized Riemann existence theorem is also proven in this context. In recent work of Lurie [32], the étale homotopy type of an arbitrary higher Deligne–Mumford stack is defined by using *shape theory* to produce an object in the ∞–category $\mathrm{Pro}(\mathcal{S})$ (in fact the definition in [32] is for *spectral* Deligne–Mumford stacks — analogues of Deligne–Mumford stacks for algebraic geometry over $\mathbb{E}_\infty$–rings), and Hoyois has recently proven that, up to profinite completion, this definition agrees with that of Friedlander for a classical locally Noetherian scheme in [22]. See also recent work of Barnea, Harpaz and Horel [6].

In recent work of the first author [11], the étale homotopy type of an arbitrary higher stack on the étale site of affine k–schemes is defined, and is shown to agree with the definition of Lurie when restricted to higher Deligne–Mumford stacks. In particular, there is shown to be a functor

$$\widehat{\Pi}^{\text{ét}}_\infty\colon \mathrm{Sh}_\infty(\mathbf{Aff}_k^{\mathrm{LFT}}, \text{ét}) \to \mathrm{Prof}(\mathcal{S})$$

associating to a higher stack $\mathcal{X}$ on the étale site of affine k–schemes of finite type a profinite space $\widehat{\Pi}^{\text{ét}}_\infty(\mathcal{X})$ called its *profinite homotopy type*, and an even more generalized Riemann existence theorem is proven:

Theorem 7.2 [11, Theorem 4.13] *Let $\mathcal{X}$ be higher stack on affine schemes of finite type over $\mathbb{C}$; then there is a canonical equivalence of profinite spaces*

$$\widehat{\Pi}^{\text{ét}}_\infty(\mathcal{X}) \simeq \widehat{\Pi}_\infty(\mathcal{X}_{\mathrm{top}})$$

between its profinite étale homotopy type and the profinite homotopy type of its underlying topological stack $\mathcal{X}_{\mathrm{top}}$ in the sense of Theorem 3.1.

Now let X be a log scheme locally of finite type over $\mathbb{C}$. Its infinite root stack $\sqrt[\infty]{X}$ is a pro-object in $\mathrm{Sh}_\infty(\mathbf{Aff}_{\mathbb{C}}^{\mathrm{LFT}}, \text{ét})$. Notice that the functor $\widehat{\Pi}_\infty^{\text{ét}}$ canonically extends to a functor

$$\widehat{\Pi}_\infty^{\text{ét}}\colon \mathrm{Pro}(\mathrm{Sh}_\infty(\mathbf{Aff}_k^{\mathrm{LFT}}, \text{ét})) \to \mathrm{Prof}(\mathcal{S}).$$

In light of the above theorem, we conclude that there is a canonical equivalence of profinite spaces

$$\widehat{\Pi}_\infty^{\text{ét}}(\sqrt[\infty]{X}) \simeq \widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}})$$

between the profinite étale homotopy type of the infinite root stack $\sqrt[\infty]{X}$ and the profinite homotopy type of the underlying topological stack of the infinite root stack $\sqrt[\infty]{X}_{\mathrm{top}}$. Combining this with Theorem 6.4 yields the following theorem:

Theorem 7.3 *Let X be a log scheme locally of finite type over $\mathbb{C}$. Then the following three profinite spaces are canonically equivalent:*

- (i) *The profinite completion $\widehat{X_{\log}}$ of its Kato–Nakayama space.*
- (ii) *The profinite homotopy type $\widehat{\Pi}_\infty(\sqrt[\infty]{X}_{\mathrm{top}})$ of the underlying topological stack of its infinite root stack $\sqrt[\infty]{X}$.*
- (iii) *The profinite étale homotopy type $\widehat{\Pi}_\infty^{\text{ét}}(\sqrt[\infty]{X})$ of its infinite root stack $\sqrt[\infty]{X}$.*

In light of the above theorem, we make the following definition:

Definition 7.4 Let X be a log scheme over a ground ring k. Then the *profinite homotopy type* of X is the profinite étale homotopy type of its infinite root stack $\sqrt[\infty]{X}$.

Appendix

In this appendix we gather some definitions and results about log schemes, analytification, the Kato–Nakayama space, root stacks and topological stacks.

A.1 Log schemes

Log (short for "logarithmic") schemes were first defined and studied systematically in [27]. A modern introduction (with a view towards moduli theory) can be found in [1].

Remark A.1 We will give definitions and facts in the algebraic category, but we will apply them to the complex-analytic context as well. The only difference is that instead of the étale topology we will be using the analytic topology.

Definition A.2 A *log scheme* is a scheme X with a sheaf of monoids M on the small étale site $X_{\text{ét}}$ and a homomorphism $\alpha\colon M \to \mathcal{O}_X$ of sheaves of monoids, where $\mathcal{O}_X$

is seen as a monoid with respect to multiplication of regular functions, such that α induces an isomorphism

$$\alpha|_{\alpha^{-1}(\mathcal{O}_X^\times)}\colon \alpha^{-1}(\mathcal{O}_X^\times) \to \mathcal{O}_X^\times.$$

Note that the last condition gives us a canonical embedding $\mathcal{O}_X^\times \hookrightarrow M$ as a subsheaf of groups.

We denote a log scheme by (X, M, α) or sometimes simply by X.

Example A.3
- Any scheme X is a log scheme with $M = \mathcal{O}_X^\times$ and α the inclusion. This is the trivial log structure on X.
- Any effective Cartier divisor $D \subseteq X$ induces a log structure, by taking M to be the subsheaf of $\mathcal{O}_X$ given by functions that are invertible outside of D.
- If P is a monoid, the spectrum of the monoid algebra $X_P := \operatorname{Spec} k[P]$ has a natural log structure. The sheaf M is obtained by considering the natural map $P \to k[P] = \Gamma(\mathcal{O}_{X_P})$ and taking the "associated log structure" (see below for a few more details).

Log structures can be pulled back and pushed forward along morphisms of schemes. In particular:

- Any open subscheme of a log scheme can be equipped with the restriction of the log structure.
- If we have a morphism of schemes $f\colon X \to \operatorname{Spec} k[P]$ we get an induced log structure on X. This happens in the following way: f gives a morphism of monoids $P \to \mathcal{O}_X(X)$, which induces $\overline{\alpha}\colon \underline{P} \to \mathcal{O}_X$, where $\underline{P}$ is the constant sheaf. It is typically not true that $\overline{\alpha}$ induces an isomorphism between $\overline{\alpha}^{-1}\mathcal{O}_X^\times$ and $\mathcal{O}_X^\times$, but there is a procedure to fix the behavior of the units, and this produces a log structure $\alpha\colon M \to \mathcal{O}_X$. See [27, Example 1.5] for details.

Remark A.4 In the situation of the last bullet, the quotient $M/\mathcal{O}_X^\times$ is obtained from $\underline{P}$ by locally "killing the sections of $\underline{P}$ that become invertible in $\mathcal{O}_X$", so in particular all the stalks of $M/\mathcal{O}_X^\times$ are quotients of the monoid P.

We consider only coherent log structures, which are those that, étale locally, come by pullback from the spectrum of the monoid algebra of a monoid.

Definition A.5 A log scheme X is *quasicoherent* if there is an étale cover U_i of X, monoids P_i and morphisms of log schemes $f_i\colon U_i \to \operatorname{Spec} k[P_i]$ that are strict, ie the log structure on U_i is pulled back from $\operatorname{Spec} k[P_i]$ via f_i. The monoid P_i and the map f_i are a *chart* for the log structure over U_i.

A log scheme X is *coherent* (resp. *fine*, resp. *fine and saturated*) if the monoids P_i above can be taken to be finitely generated (resp. finitely generated and integral, resp. finitely generated, integral and saturated).

A morphism such as f_i in the definition above that identifies the pullback of the log structure on the target with the one of the source will be called *strict*.

We are interested only in fine and saturated log schemes.

Proposition A.6 [39, Proposition 2.1] *Let X be a fine saturated log scheme and x a geometric point. Then there exists an étale neighborhood U of x over which there is a chart for the log structure with monoid $P = (M/\mathcal{O}_X^\times)_x$.*

This says in particular that, if X is fine and saturated, we can locally find charts with P fine, saturated and sharp.

The quotient sheaf $\overline{M} = M/\mathcal{O}_X^\times$ is called the characteristic sheaf of the log structure. Taking the quotient (in an appropriate sense) by $\mathcal{O}_X^\times$ of the map α, we get an alternative definition of a (quasi-integral) log scheme, introduced in [9].

Let us denote by Div_X the fibered category over $X_{\text{ét}}$ whose objects over $U \to X$ are pairs (L, s) where L is an invertible sheaf of $\mathcal{O}_U$–modules on U and s is a global section. This is a symmetric monoidal fibered category, where the monoidal operation is given by tensor product.

Definition A.7 A *log scheme* is a scheme X together with a sheaf of monoids A and a symmetric monoidal functor $L\colon A \to \mathrm{Div}_X$ with trivial kernel.

The phrasing "trivial kernel" in the definition means that if a section a is such that $L(a)$ is isomorphic to $(\mathcal{O}_X, 1)$ in Div_X, then $a = 0$.

Given a (quasi-integral) log scheme (X, M, α), by taking the "stacky quotient" of $\alpha\colon M \to \mathcal{O}_X$ by $\mathcal{O}_X^\times$ we get the functor $L\colon A = \overline{M} = [M/\mathcal{O}_X^\times] \to [\mathcal{O}_X/\mathcal{O}_X^\times] = \mathrm{Div}_X$. Quasi-integrality ensures that the quotient $[M/\mathcal{O}_X^\times]$ is actually a sheaf. Of course integral log structures are quasi-integral. See [9, Theorem 3.6] for details.

One can give a notion of charts in this context as well. For many purposes these two notions of chart can be used indifferently. We mostly use charts as in the first definition above. These are called "Kato charts" in [9].

Remark A.8 A first approximation of how one should "visualize" a log scheme is by thinking about the stalks of the sheaf $\overline{M}$. This sheaf is locally constant on a stratification of X (see Proposition A.27) and the stalks are fine saturated sharp monoids. Of course this disregards the particular extension M of $\overline{M}$ by $\mathcal{O}_X^\times$ and the map α (or equivalently the functor L), so it is indeed just a crude approximation.

A.2 Analytification

We are mainly concerned with log schemes locally of finite type over $\mathbb{C}$, and with their analytifications.

Recall that if X is a scheme locally of finite typer over $\mathbb{C}$, the associated analytic space X_{an} is defined as a set as the $\mathbb{C}$ points $X(\mathbb{C}) = X(\operatorname{Spec}\mathbb{C})$ of X. This has an "analytic" topology coming from the local embeddings into $\mathbb{C}^n$. Moreover this construction extends also to algebraic spaces locally of finite type over $\mathbb{C}$ (see [2; 49]).

If X is a log scheme locally of finite type over $\mathbb{C}$, the analytification X_{an} inherits a log structure, because of the relationship between the étale topos of X and the analytic topos of X_{an}. An étale morphism $X \to Y$ induces a local homeomorphism $X_{\mathrm{an}} \to Y_{\mathrm{an}}$, which consequently has local sections in the analytic topology. This gives a functor from the étale site of X to the analytic site of X_{an}, and induces a morphism of topoi. The log structure on X_{an} is obtained via this functor. Thus, in what follows, every time something holds étale locally for the log scheme X, it will hold analytically locally for the log analytic space X_{an}.

We will use this without further mention, and will use the same letter to denote the sheaf of monoids M on X and the induced one on X_{an}. This should cause no real confusion.

Definition A.9 For a monoid P we denote by $\mathbb{C}(P)$ the analytification of the spectrum of the monoid algebra $\operatorname{Spec}\mathbb{C}[P]$.

As sets we have $\mathbb{C}(P) = \operatorname{Hom}(P, \mathbb{C})$, the set of homomorphisms of monoids, where $\mathbb{C}$ is given the multiplicative structure.

A basis of opens of $\mathbb{C}(P)$ (where P is fine, saturated and sharp) can be described as follows: call $p_1, \dots, p_k$ the indecomposable elements of P (see [38, Proposition 2.1.2]), and choose open disks D_i in the complex plane $\mathbb{C}$. Then the set of homomorphisms $\phi \in \operatorname{Hom}(P, \mathbb{C})$ such that $\phi(p_i) \in D_i$ is open in $\mathbb{C}(P)$. Letting the disks D_i vary we get a basis for the open subsets of $\mathbb{C}(P)$.

Lemma A.10 [49, page 12] *Analytification commutes with finite limits.*

We will need the following result on the topological properties of analytifications of schemes locally of finite type over $\mathbb{C}$. As a reference, we point out [30].

Proposition A.11 *Let X be an affine scheme of finite type over $\mathbb{C}$ and $Y \subseteq X$ be a closed subscheme. Then there exist compatible triangulations of X_{an} and Y_{an}, realizing Y_{an} as a subcomplex.*

We can apply this iteratively to a stratification, to get compatible triangulations of the ambient affine scheme and of all the (closed) strata.

A.3 Kato–Nakayama space

From now on all log schemes will be fine and saturated unless we specify otherwise. Just for this subsection, X will denote an analytic space rather than a scheme.

The Kato–Nakayama space $X_{\log}$ of a log analytic space X (for example of the form Y_{an} for some log scheme Y locally of finite type over $\mathbb{C}$) is a topological space introduced in [28]. The idea is to define a topological space that "embodies" the log structure of X in a topological way (ie without using the sheaf of monoids, but only "points").

What comes out is a topological space $X_{\log}$ (that also comes with a natural sheaf of rings, but we do not use this in the present work) with a continuous map $\tau\colon X_{\log} \to X$ that is proper and surjective. Moreover if $U \subseteq X$ is the trivial locus of the log structure (the largest open subset over which $\mathcal{O}_X^\times \hookrightarrow M$ is an isomorphism), the open embedding $i\colon U \to X$ factors through τ, so that $X_{\log}$ can be considered as a " relative compactification" of the open immersion i.

Let us denote by $p^\dagger$ the log analytic space given by the point $\mathrm{pt} = (\operatorname{Spec}\mathbb{C})_{\mathrm{an}}$ with monoid $M = \mathbb{R}_{\geq 0} \times S^1$, and map $\alpha\colon M \to \mathbb{C}$ described by $(r, a) \mapsto r \cdot a$. Note that this log structure is not integral.

As a set we have $X_{\log} = \operatorname{Hom}(p^\dagger, X)$, the set of morphisms of log analytic spaces from the log point $p^\dagger$ to X. By unraveling this one can also write

$$X_{\log} = \bigl\{(x, c) \mid c\colon M_x^{\mathrm{gp}} \to S^1 \text{ is a group homomorphism such that } c(f) = f/|f| \text{ for all } f \in \mathcal{O}_{X,x}^\times\bigr\}.$$

In particular one can see that $\mathbb{C}(P)_{\log} = \operatorname{Hom}(p^\dagger, \mathbb{C}(P)) = \operatorname{Hom}(P, \mathbb{R}_{\geq 0} \times S^1)$, and the projection $\tau\colon \mathbb{C}(P)_{\log} \to \mathbb{C}(P)$ is given by postcomposition with $\mathbb{R}_{\geq 0} \times S^1 \to \mathbb{C}$.

Note that from the above description $\mathbb{C}(P)_{\log}$ has a natural topology, which by means of local charts for the log structure gives a topology on $X_{\log}$ in general [28, Section 1.2].

From the description one sees easily that, for $x \in X_{\mathrm{an}}$, the fiber $\tau^{-1}(x)$ is homeomorphic to $(S^1)^r$, where r is the rank of the stalk $\overline{M}_x$, defined to be the rank of the free abelian group $\overline{M}_x^{\mathrm{gp}}$.

The construction of the Kato–Nakayama space is clearly functorial, and is also compatible with strict base change.

Proposition A.12 [28, Lemma 1.3] *Let $f\colon X \to Y$ be a strict morphism of fine saturated log analytic spaces. Then the diagram of topological spaces*

$$\begin{tikzcd} X_{\log} \arrow[r] \arrow[d] & Y_{\log} \arrow[d] \\ X \arrow[r] & Y \end{tikzcd}$$

is cartesian.

The description of $X_{\log}$ as a set can actually be enhanced to a description of its functor of points (see Section 3.4).

We now prove the following proposition:

Proposition A.13 *For any log scheme X, the Kato–Nakayama space $X_{\log}$ is locally Hausdorff, locally contractible and locally compact.*

We will start by assuming that X is affine and has a global chart $X \to \operatorname{Spec} \mathbb{C}[P]$ for a fine saturated sharp monoid P, and will prove that $X_{\log}$ is locally compact, Hausdorff and locally contractible. This implies the conclusion for arbitrary X.

Note that since $f\colon X \to \operatorname{Spec} \mathbb{C}[P]$ is strict, there is a cartesian diagram of topological spaces

$$\begin{tikzcd} X_{\log} \arrow[r] \arrow[d] & \mathbb{C}(P)_{\log} \arrow[d, "\tau"] \\ X_{\mathrm{an}} \arrow[r, "f_{\mathrm{an}}"] & \mathbb{C}(P) \end{tikzcd}$$

Our proof will be as follows: We note that X_{an} and $\mathbb{C}(P)$ are semialgebraic, and the map $X_{\mathrm{an}} \to \mathbb{C}(P)$ is a semialgebraic function (this part of the diagram is even algebraic). We will check that $\mathbb{C}(P)_{\log}$ is semialgebraic, and that the projection to $\mathbb{C}(P)$ is a semialgebraic function.

After we do that, it will follow that $X_{\log}$ is semialgebraic as well (being the inverse image of the diagonal $\mathbb{C}(P) \subseteq \mathbb{C}(P) \times \mathbb{C}(P)$, a semialgebraic set, through the semialgebraic map $(f_{\mathrm{an}}, \tau)\colon X_{\mathrm{an}} \times \mathbb{C}(P)_{\log} \to \mathbb{C}(P) \times \mathbb{C}(P)$, see [8, Proposition 2.2.7]), hence triangulable (by the results of [30]), and any triangulable locally semialgebraic set is locally compact, Hausdorff and locally contractible [21].

Lemma A.14 *The topological space $\mathbb{C}(P)_{\log}$ is semialgebraic, and the projection $\mathbb{C}(P)_{\log} \to \mathbb{C}(P)$ is a semialgebraic map.*

Proof We will check this by writing out these spaces explicitly. Let p_i be a finite set of generators for P (for example the indecomposable elements), and assume we have a finite number of relations that present the monoid P, of the form $\sum_j r_{ij} p_j = \sum_j s_{ij} p_j$. Say there are k generators and h relations.

Then we have a map $\mathbb{C}(P) = \mathrm{Hom}(P, \mathbb{C}) \to \mathbb{C}^k$ given by $\phi \mapsto (\phi(p_i))$. This is an embedding, and the closed image is the Zariski closed subset with equations $\prod_j (z_j)^{r_{ij}} = \prod_j (z_j)^{s_{ij}}$ obtained from the h relations of the chosen presentation of P, where (z_j) are the coordinates of $\mathbb{C}^k$.

In the exact same way we have a map $\mathbb{C}(P)_{\log} = \mathrm{Hom}(P, \mathbb{R}_{\geq 0} \times S^1) \to (\mathbb{R}_{\geq 0} \times S^1)^k$ given by $\psi \mapsto (\psi(p_i))$. To describe the image, let us note that we have $\mathbb{R}_{\geq 0} \times S^1 \subseteq \mathbb{R}^3$ in a natural way, as a semialgebraic subset. If we denote by (ζ_j) the "coordinates" of $(\mathbb{R}_{\geq 0} \times S^1)^k$, then the (isomorphic) image of $\mathbb{C}(P)_{\log}$ is again described by the equations $\prod_j (\zeta_j)^{r_{ij}} = \prod_j (\zeta_j)^{s_{ij}}$, so it is semialgebraic (the equations translate into algebraic equations on $(\mathbb{R}^3)^k$).

Of course the diagram

$$\begin{tikzcd}
\mathbb{C}(P)_{\log} \arrow[r] \arrow[d] & (\mathbb{R}_{\geq 0} \times S^1)^k \arrow[d] \\
\mathbb{C}(P) \arrow[r] & \mathbb{C}^k
\end{tikzcd}$$

commutes.

From this, it suffices to check that the map $(\mathbb{R}_{\geq 0} \times S^1)^k \to \mathbb{C}^k$ is semialgebraic, and this is easy: in coordinates (where we see $(\mathbb{R}_{\geq 0} \times S^1)^k \subseteq (\mathbb{R}^3)^k$ and $\mathbb{C}^k \cong (\mathbb{R}^2)^k$) it is given by $(a_i, b_i, c_i) \mapsto (a_i \cdot b_i, a_i \cdot c_i)$. □

A.4 Root stacks

Root stacks of log schemes were introduced in [9]. The infinite root stack, an inverse limit of the ones with finitely generated weight system, is the subject of [48]. We briefly recall the functorial definition and the groupoid presentations coming from local charts.

Let us fix a natural number n and a log scheme X with log structure $L\colon A \to \mathrm{Div}_X$. We can consider a sheaf $\frac{1}{n}A$ of "fractions" of sections of A: the sections of $\frac{1}{n}A$ are formal fractions $\frac{a}{n}$ where a is a section of A. There is a natural inclusion $i_n\colon A \to \frac{1}{n}A$.

Note that $\frac{1}{n}A$ is isomorphic to A via $a \mapsto \frac{a}{n}$. Through this isomorphism, the inclusion i_n corresponds to multiplication by $n\colon A \to A$. The fact that this map is injective follows from torsion-freeness of stalks of A, which are fine saturated sharp monoids.

Definition A.15 The n^{th} *root stack* $\sqrt[n]{X}$ of the log scheme X is the stack over **Sch**, the category of schemes (with the étale topology), whose functor of points sends a scheme T to the groupoid whose objects are pairs (ϕ, N, a) where $\phi\colon T \to X$ is a morphism of schemes, $N\colon \frac{1}{n}\phi^* A \to \mathrm{Div}_X$ is a symmetric monoidal functor with trivial kernel and a is a natural isomorphism between $\phi^* L$ and the composite $N \circ i_n$:

$$\begin{tikzcd} \phi^* A \arrow[r] \arrow[d] \arrow[dr, phantom, "\swArrow a"] & \mathrm{Div}_X \\ \frac{1}{n}\phi^* A \arrow[ur] & \end{tikzcd}$$

Morphisms are defined in the obvious way.

In other words the n^{th} root stack parametrizes extensions of the symmetric monoidal functor $L\colon A \to \mathrm{Div}_X$ to the sheaf $\frac{1}{n}A$. The pair (N, a) in the definition above could be called an "n^{th} root" of the log structure $L\colon A \to \mathrm{Div}_X$.

Every time $n \mid m$ there is a morphism $\sqrt[m]{X} \to \sqrt[n]{X}$, and by letting n and m vary, these maps give an inverse system of stacks over **Sch**.

Definition A.16 The *infinite root stack* $\sqrt[\infty]{X}$ of the log scheme X is the pro-algebraic stack $(\sqrt[n]{X})_{n\in\mathbb{N}}$.

Remark A.17 In [48] the infinite root stack is defined as the actual limit of the inverse system in the 2–category of fibered categories, but in the present paper it will always be the pro-object. The two contain the same information, since by the results of [48, Section 5] the limit of the system of n^{th} root stacks recovers the log scheme completely, and hence recovers the pro-object as well.

The n^{th} root stack $\sqrt[n]{X}$ is a tame Artin stack with coarse moduli space X. Moreover there are presentations of $\sqrt[n]{X}$ for each n that assemble into a pro-object in groupoids in schemes, and can be regarded as a presentation of the pro-object $\sqrt[\infty]{X}$. This follows from the following local descriptions as quotient stacks [48, Corollary 3.12].

Let us fix a monoid P, and let us denote by C_n the cokernel of the injective map $P^{\mathrm{gp}} \to \frac{1}{n}P^{\mathrm{gp}}$. Furthermore, denote by $\mu_n(P)$ the Cartier dual of C_n. This acts on the monoid algebra $\operatorname{Spec} k\big[\frac{1}{n}P\big]$ (k here is some base field, but this works the same way over $\mathbb{Z}$).

If X is a log scheme with a global chart $X \to \operatorname{Spec} k[P]$, then there is a cartesian diagram

$$\begin{tikzcd} \sqrt[n]{X} \arrow[r] \arrow[d] & \big[\operatorname{Spec} k\big[\frac{1}{n}P\big]/\mu_n(P)\big] \arrow[d] \\ X \arrow[r] & \operatorname{Spec} k[P] \end{tikzcd}$$

presenting $\sqrt[n]{X}$ as a quotient stack $[X_n/\mu_n(P)]$, where $X_n = X \times_{\operatorname{Spec} k[P]} \operatorname{Spec} k\big[\frac{1}{n}P\big]$.

As we mentioned, these quotient stack presentations are all compatible, in the sense that they give a pro-object in groupoids in schemes $(X_n \times \mu_n(P) \rightrightarrows X_n)_{n\in\mathbb{N}}$, which can be seen as a groupoid presentation of $\sqrt[\infty]{X}$.

If X does not have a global chart we cover it with étale opens U_i where there is a chart with monoid P_i and assemble together the corresponding groupoid presentations.

Proposition A.18 [9, Proposition 4.19] *The n^{th} root stack $\sqrt[n]{X}$ is a tame Artin stack, and is Deligne–Mumford when we are over a field of characteristic* 0.

A.5 Topological stacks

The main reference for this section is [36].

The two preceding subsections were about the objects that we would like to compare, namely the Kato–Nakayama space and the infinite root stack of a log scheme locally of finite type over $\mathbb{C}$. Note that the former is of topological nature, and the latter is algebraic. In order to find a map between them, we carry over the root stacks to the topological side.

One can talk about stacks over any Grothendieck site. Algebraic stacks (also known as Artin stacks) are stacks on the category of schemes over a base with the étale topology[1] that admit a representable smooth epimorphism from a scheme and whose diagonal is representable by algebraic spaces (and often one imposes some conditions on this diagonal morphism, like being quasicompact or locally of finite type). Equivalently, one can describe algebraic stacks as stacks of (étale) torsors for certain groupoid objects in algebraic spaces, whose structure maps are smooth.

If instead of schemes over a base with the étale topology we start from topological spaces with the étale topology (where covers are local homeomorphisms), and we require a representable epimorphism from a topological space, we obtain the theory of topological stacks.[2] Such a stack will always have diagonal representable by a topological space. As on the algebraic side, a topological stack can be defined through a groupoid presentation: a topological stack is a stack of principal $\mathcal{G}$–bundles for $\mathcal{G}$ a topological groupoid, and much of the basic yoga that one learns when working with algebraic stacks carries over in close analogy in this context.

In particular if G is a topological group acting on a space X, the functor of points of the quotient stack $[X/G]$ is described as principal G–bundles (the topological analogue

[1] Sometimes, rather than working with the étale topology, one defines algebraic stacks with the *fppf* topology. However, the resulting 2–category of stacks is the same; see [7, Tag 076U].

[2] Noohi [36] demands further conditions for such a stack to be called a topological stack; however, in subsequent papers (eg [37]), he relaxes these conditions to the ones just described.

of G–torsors) with an equivariant map to X. In the same fashion, if $R \rightrightarrows U$ is a topological groupoid, one can characterize the associated stack $[U/R]$ as the stack of principal bundles for this groupoid.

There is a procedure to produce a topological stack starting from an algebraic one, that extends the analytification functor. We apply this in particular to the n^{th} root stacks of a log scheme.

Denote by $\mathbf{\mathfrak{AlgSt}}_{\mathbb{C}}^{\mathrm{LFT}}$ the 2–category of algebraic stack locally of finite type over $\mathbb{C}$ and by $\mathfrak{TopSt}$ the 2–category of topological stacks.

Proposition A.19 [36, Section 20] *There is a functor of* 2*–categories*

$$(\,\cdot\,)_{\mathrm{top}}\colon \mathbf{\mathfrak{AlgSt}}_{\mathbb{C}}^{\mathrm{LFT}} \to \mathfrak{TopSt}$$

that associates a topological stack to an algebraic stack locally of finite type over $\mathbb{C}$*.*

In Section 3, we extend Noohi's results to produce a left exact colimit-preserving functor from ∞–sheaves (also known as stacks of ∞–groupoids) on the algebraic étale site to hypersheaves on a suitable topological site. See Theorem 3.1 and Corollary 3.11.

This functor has several nice properties. We point out the ones that we use:

1. If X is a scheme (or algebraic space) locally of finite type over $\mathbb{C}$, then $X_{\mathrm{top}} \simeq X_{\mathrm{an}}$ is the analytification
2. The functor $(\,\cdot\,)_{\mathrm{top}}$ preserves all finite limits (ie is left exact).
3. The preceding properties give us a procedure for calculating $\mathcal{X}_{\mathrm{top}}$ for an algebraic stack $\mathcal{X}$. If $R \rightrightarrows U$ is a groupoid presentation of $\mathcal{X}$, where R and U are locally of finite type and the maps are smooth, then by the first property we can apply the analytification functor to the diagram, and, by the second one, this will result in another groupoid, namely the groupoid in topological spaces $R_{\mathrm{an}} \rightrightarrows U_{\mathrm{an}}$. The topological stack $\mathcal{X}_{\mathrm{top}}$ is then the associated stack $[U_{\mathrm{an}}/R_{\mathrm{an}}]$.

In particular, if $\mathcal{X} = [U/G]$ for an action of an algebraic group locally of finite type G on a scheme locally of finite type X, we have $\mathcal{X}_{\mathrm{top}} = [U_{\mathrm{an}}/G_{\mathrm{an}}]$.

Definition A.20 Let X be a log scheme locally of finite type over $\mathbb{C}$. The *topological* n^{th} *root stack* of X is the topological stack $\sqrt[n]{X}_{\mathrm{top}}$. As for the algebraic ones, the topological root stacks form an inverse system. The pro-topological stack $\sqrt[\infty]{X}_{\mathrm{top}} := (\sqrt[n]{X}_{\mathrm{top}})_{n\in\mathbb{N}}$ is the *topological infinite root stack* of X.

A.6 The rank stratification

In this section we will prove that the characteristic sheaf $\overline{M}$ is locally constant on a stratification over the log scheme X. This is used in the main body of this article to prove that the Kato–Nakayama space and the infinite root stack are "stratified fibrations" over X, and that the map that we construct between them induces an equivalence of profinite completions.

The results of this part are probably known to experts, and we are including them because of the lack of a suitable reference.

Definition A.21 By a *stratification* of a topological space T we mean a collection of closed subsets $\mathcal{S} = \{S_i \subseteq T\}_{i \in I}$ where I is partially ordered and

- if $i \leq j$ then $S_i \subseteq S_j$, and
- the stratification is locally finite: every point $t \in T$ has an open neighborhood U such that only finitely many of the intersections $U \cap S_i$ are nonempty.

The locally closed subsets $S_j \setminus S_i$ will be called the *strata* of the stratification.

If in the above definition T is the underlying topological space of a scheme X and each S_i is Zariski closed, we will say that $\mathcal{S}$ is an *algebraic* stratification of the scheme X. Note that an algebraic stratification on X will induce a stratification on the analytification X_{an}.

Definition A.22 Let T be a topological space equipped with a stratification $\mathcal{S}$, and let $f\colon T' \to T$ be a morphism, where T' is a topological space or stack. We will say that f is a *stratified fibration* with respect to $\mathcal{S}$ if the restrictions of f to the strata of $\mathcal{S}$ are fibrations (in our case, this will always mean "locally the projection from a product").

Now let X be a log scheme locally of finite type over a field k. We will describe an algebraic stratification of X over which the sheaf $\overline{M}$ is locally constant.

The basic idea is that we are stratifying by the rank of the stalks $\overline{M}^{\mathrm{gp}}_x$ of the sheaf of abelian groups $\overline{M}^{\mathrm{gp}}$.

Lemma A.23 [39, Lemma 3.5] *The sheaf $\overline{M}^{\mathrm{gp}}$ is a constructible sheaf of $\mathbb{Z}$–modules* [4, Exposé IX, Definition 2.3]. *This means that (Zariski locally) there is a decomposition of X into locally closed subsets over which $\overline{M}^{\mathrm{gp}}$ is a locally constant sheaf.*

Lemma A.24 [38, Theorem 2.3.2] *If ξ is a generalization of η in X, meaning that $\eta \in \overline{\{\xi\}}$, then there is a natural morphism of the stalks $\overline{M}_{\overline{\eta}} \to \overline{M}_{\overline{\xi}}$, and this is surjective (more specifically, it is a quotient by a* face*).*

This last lemma follows from Proposition A.6 and from the explicit description of the stalks of the monoid $\overline{M}$ of the log structure obtained from a chart; see Remark A.4.

In particular the rank "only jumps up in closed subsets", ie for every $n \in \mathbb{N}$ the subset R_n of points of X where the rank of the group $\overline{M}^{\mathrm{gp}}_{\overline{x}}$ is at least n is closed: it is constructible by Lemma A.23, and stable under specialization by Lemma A.24, so it is closed. Note also that $R_{n+1} \subseteq R_n$.

Definition A.25 The *rank stratification* of a log scheme X is the algebraic stratification $\mathcal{R} = \{R_n\}_{n\in\mathbb{N}}$, where

$$R_n = \{x \in X \mid \operatorname{rank}_{\mathbb{Z}} \overline{M}^{\mathrm{gp}}_{\overline{x}} \geq n\}.$$

We will denote the strata by $S_n := R_n \setminus R_{n+1}$.

For example, $R_0 = X$ and the complement $X \setminus R_1$ is the open subset of X where the log structure is trivial (which might be empty). In general S_n is the locally closed subset of X over which the rank of $\overline{M}^{\mathrm{gp}}_{\overline{x}}$ is equal to n.

We claim that both sheaves $\overline{M}$ and $\overline{M}^{\mathrm{gp}}$ are locally constant on the strata S_n.

To check this, let us describe the canonical log structure $\overline{M}_P \to \mathrm{Div}_{X_P}$ on $X_P = \operatorname{Spec} k[P]$ in more detail: the log structure is induced by the morphism of monoids $P \to k[P]$, which gives a morphism of sheaves of monoids $\underline{P} \to \mathcal{O}_{X_P}$ (here $\underline{P}$ denotes the constant sheaf), from which we get the sheaf $\overline{M}_P$ by killing the preimage of the units in $\mathcal{O}_{X_P}$. More precisely, denote by $\{p_i\}_{i\in I}$ the finitely many indecomposable elements of the fine saturated monoid P; these are generators of P. For a geometric point $x \to X_P$ call $S \subseteq I$ the subset of indices such that the image of $t^{p_i} \in k[P]$ is invertible in the residue field $k(x)$. Then the stalk $(\overline{M}_P)_x$ is the quotient $P/\langle p_i \mid i \in S\rangle$.

In particular we note the following:

Lemma A.26 *The only point x of X_P where the stalk $(\overline{M}_P)_{\overline{x}}$ has rank $n = \operatorname{rank}_{\mathbb{Z}} P^{\mathrm{gp}}$ is the "vertex" v_P, the point given by the maximal ideal $\langle t^{p_i} \mid i \in I\rangle$ generated by the variables corresponding to the indecomposable elements of P.*

The point v_P is also sometimes referred to as the "torus-fixed point".

Proof Since $P^{\mathrm{gp}} \cong \mathbb{Z}^n$ for some n, as soon as at least one of the indecomposable elements p_i is killed, the rank will drop at least by 1. The only point in which no indecomposable is killed is exactly the maximal ideal generated by all the t^{p_i}. □

Proposition A.27 *For every n and every point x of $S_n = R_n \setminus R_{n-1}$, there is an étale neighborhood $U \to S_n$ of x such that the sheaves $\overline{M}|_{S_n}$ and $\overline{M}^{\mathrm{gp}}|_{S_n}$ are constant sheaves.*

Proof If we equip R_n with the reduced subscheme structure, it is a (fine saturated) log scheme with the log structure pulled back from X, and the same is true for the open subset $S_n \subseteq R_n$. Consequently there is an étale neighborhood $U \to S_n$ of x and a chart $U \to \operatorname{Spec} k[P]$ for the induced log structure on U, where $P = \overline{M}_{\overline{x}}$ (Proposition A.6).

If $\overline{M}_P$ is the sheaf of monoids for the canonical log structure on $\operatorname{Spec} k[P]$, there is exactly one point where the stalk has rank $n = \operatorname{rank} P$ $(= \operatorname{rank}_{\mathbb{Z}} P^{\mathrm{gp}})$, corresponding to the vertex v_P (Lemma A.26).

This implies (since over U the rank of the stalks of $\overline{M}$ is always n) that the morphism $U \to \operatorname{Spec} k[P]$ sends everything to v_P, and in turn that the sheaf $\overline{M}|_U$, being a pullback from $\operatorname{Spec} k[P]$, is constant. This implies that $\overline{M}^{\mathrm{gp}}|_U$ is constant as well, and concludes the proof. □

Note that if $k = \mathbb{C}$, the algebraic stratification of X we just constructed induces a stratification of the analytification X_{an}, and the sheaves $\overline{M}$ and $\overline{M}^{\mathrm{gp}}$ of the log analytic space are locally constant over the strata.

References

[1] **D Abramovich**, **Q Chen**, **D Gillam**, **Y Huang**, **M Olsson**, **M Satriano**, **S Sun**, *Logarithmic geometry and moduli*, from "Handbook of moduli, I" (G Farkas, I Morrison, editors), Adv. Lect. Math. 24, International Press, Somerville, MA (2013) 1–61 MR

[2] **M Artin**, *Algebraization of formal moduli, II: Existence of modifications*, Ann. of Math. 91 (1970) 88–135 MR

[3] **M Artin**, **A Grothendieck**, **J L Verdier**, *Théorie des topos et cohomologie étale des schémas, Tome 1: Théorie des topos, Exposés I–IV* (SGA 4_1), Lecture Notes in Math. 269, Springer (1972) MR

[4] **M Artin**, **A Grothendieck**, **J L Verdier**, *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (SGA 4_3), Lecture Notes in Math. 305, Springer (1973) MR

[5] **M Artin**, **B Mazur**, *Etale homotopy*, Lecture Notes in Math. 100, Springer (1969) MR

[6] **I Barnea**, **Y Harpaz**, **G Horel**, *Pro-categories in homotopy theory*, Algebr. Geom. Topol. 17 (2017) 567–643 MR

[7] **P Belmans**, **A J de Jong**, et al., *The Stacks project*, electronic reference, Columbia University (2005–) `http://stacks.math.columbia.edu`

[8] **J Bochnak**, **M Coste**, **M-F Roy**, *Real algebraic geometry*, Ergeb. Math. Grenzgeb. 36, Springer (1998) MR

[9] **N Borne**, **A Vistoli**, *Parabolic sheaves on logarithmic schemes*, Adv. Math. 231 (2012) 1327–1363 MR

[10] **D Carchedi**, *Higher orbifolds and Deligne–Mumford stacks as structured infinity topoi*, preprint (2013) arXiv To appear in Mem. Amer. Math. Soc.

[11] **D Carchedi**, *On the étale homotopy type of higher stacks*, preprint (2015) arXiv

[12] **D Carchedi**, *On the homotopy type of higher orbifolds and Haefliger classifying spaces*, Adv. Math. 294 (2016) 756–818 MR

[13] **B Conrad**, *Cohomological descent*, preprint (2004) `http://math.stanford.edu/~conrad/papers/hypercover.pdf`

[14] **T Coyne**, **B Noohi**, *Singular chains on topological stacks, I*, Adv. Math. 303 (2016) 1190–1235 MR

[15] **D Dugger**, **D C Isaksen**, *Topological hypercovers and $\mathbb{A}^1$–realizations*, Math. Z. 246 (2004) 667–689 MR

[16] **E M Friedlander**, *Fibrations in etale homotopy theory*, Inst. Hautes Études Sci. Publ. Math. 42 (1973) 5–46 MR

[17] **E M Friedlander**, *Étale homotopy of simplicial schemes*, Annals of Mathematics Studies 104, Princeton Univ. Press (1982) MR

[18] **M Groth**, *A short course on ∞–categories*, preprint (2010) arXiv

[19] **K Hagihara**, *Structure theorem of Kummer étale K–group*, K–Theory 29 (2003) 75–99 MR

[20] **A Henriques**, **D Gepner**, *Homotopy theory of orbispaces*, preprint (2007) arXiv

[21] **K R Hofmann**, *Triangulation of locally semi-algebraic spaces*, PhD thesis, University of Michigan (2009) MR `http://search.proquest.com/docview/304931094`

[22] **M Hoyois**, *Higher Galois theory*, preprint (2015) arXiv

[23] **L Illusie**, **K Kato**, **C Nakayama**, *Quasi-unipotent logarithmic Riemann–Hilbert correspondences*, J. Math. Sci. Univ. Tokyo 12 (2005) 1–66 MR

[24] **L Illusie**, **C Nakayama**, **T Tsuji**, *On log flat descent*, Proc. Japan Acad. Ser. A Math. Sci. 89 (2013) 1–5 MR

[25] **D C Isaksen**, *A model structure on the category of pro-simplicial sets*, Trans. Amer. Math. Soc. 353 (2001) 2805–2841 MR

[26] **J F Jardine**, *Simplicial presheaves*, J. Pure Appl. Algebra 47 (1987) 35–87 MR

[27] **K Kato**, *Logarithmic structures of Fontaine–Illusie*, from "Algebraic analysis, geometry, and number theory" (J-I Igusa, editor), Johns Hopkins Univ. Press, Baltimore, MD (1989) 191–224 MR

[28] **K Kato**, **C Nakayama**, *Log Betti cohomology, log étale cohomology, and log de Rham cohomology of log schemes over* $\mathbf{C}$, Kodai Math. J. 22 (1999) 161–186 MR

[29] **A Kock**, **I Moerdijk**, *Presentations of étendues*, Cahiers Topologie Géom. Différentielle Catég. 32 (1991) 145–164 MR

[30] **S Lojasiewicz**, *Triangulation of semi-analytic sets*, Ann. Scuola Norm. Sup. Pisa 18 (1964) 449–474 MR

[31] **J Lurie**, *Higher topos theory*, Annals of Mathematics Studies 170, Princeton Univ. Press (2009) MR

[32] **J Lurie**, *Derived algebraic geometry, XIII: Rational and p–adic homotopy theory*, preprint (2011) `http://tinyurl.com/lurie-dagXIII`

[33] **A Mathew**, **V Stojanoska**, *Fibers of partial totalizations of a pointed cosimplicial space*, Proc. Amer. Math. Soc. 144 (2016) 445–458 MR

[34] **J P McCammond**, *A general small cancellation theory*, Internat. J. Algebra Comput. 10 (2000) 1–172 MR

[35] **W Nizioł**, *K–theory of log-schemes, I*, Doc. Math. 13 (2008) 505–551 MR
[36] **B Noohi**, *Foundations of topological stacks, I*, preprint (2005) arXiv
[37] **B Noohi**, *Homotopy types of topological stacks*, Adv. Math. 230 (2012) 2014–2047 MR
[38] **A Ogus**, *Lectures on logarithmic algebraic geometry*, lecture notes (2006) http://math.berkeley.edu/~ogus/preprints/log_book/logbook.pdf
[39] **M C Olsson**, *Logarithmic geometry and algebraic stacks*, Ann. Sci. École Norm. Sup. 36 (2003) 747–791 MR
[40] **G Quick**, *Profinite homotopy theory*, Doc. Math. 13 (2008) 585–612 MR
[41] **G Quick**, *Some remarks on profinite completion of spaces*, from "Galois–Teichmüller theory and arithmetic geometry" (H Nakamura, F Pop, L Schneps, A Tamagawa, editors), Adv. Stud. Pure Math. 63, Math. Soc. Japan, Tokyo (2012) 413–448 MR
[42] **D G Quillen**, *Some remarks on etale homotopy theory and a conjecture of Adams*, Topology 7 (1968) 111–116 MR
[43] **H Ruddat**, **N Sibilla**, **D Treumann**, **E Zaslow**, *Skeleta of affine hypersurfaces*, Geom. Topol. 18 (2014) 1343–1395 MR
[44] **U Schreiber**, *Differential cohomology in a cohesive infinity-topos*, preprint (2013) arXiv
[45] **J-P Serre**, *Cohomologie galoisienne*, Lecture Notes in Math. 5, Springer (1965) MR
[46] **M Shiota**, *Whitney triangulations of semialgebraic sets*, Ann. Polon. Math. 87 (2005) 237–246 MR
[47] **D Sullivan**, *Genetics of homotopy theory and the Adams conjecture*, Ann. of Math. 100 (1974) 1–79 MR
[48] **M Talpo**, **A Vistoli**, *Infinite root stacks and quasi-coherent sheaves on logarithmic schemes*, preprint (2014) arXiv
[49] **B Toën**, **M Vaquié**, *Algébrisation des variétés analytiques complexes et catégories dérivées*, Math. Ann. 342 (2008) 789–831 MR

Department of Mathematical Sciences, George Mason University
Fairfax, VA 22030, United States
Mathematisches Institut, University of Bonn
Endenicher Allee 60, D-53115 Bonn, Germany
School of Mathematics, Statistics and Actuarial Sciences, University of Kent
Canterbury, CT2 7FS, United Kingdom
Department of Mathematics, Simon Fraser University
Burnaby, BC V5A 1S6, Canada

davidcarchedi@gmail.com, sarah@math.uni-bonn.de, n.sibilla@kent.ac.uk, mtalpo@sfu.ca

Proposed: Dan Abramovich Received: 25 April 2016
Seconded: Paul Goerss, Frances Kirwan Revised: 28 September 2016

Geometry & Topology 21 (2017) 3159–3190

Positive simplicial volume implies virtually positive Seifert volume for 3–manifolds

PIERRE DERBEZ
YI LIU
HONGBIN SUN
SHICHENG WANG

We show that for any closed orientable 3–manifold with positive simplicial volume, the growth of the Seifert volume of its finite covers is faster than the linear rate. In particular, each closed orientable 3–manifold with positive simplicial volume has virtually positive Seifert volume. The result reveals certain fundamental differences between the representation volumes of hyperbolic type and Seifert type. The proof is based on developments and interactions of recent results on virtual domination and on virtual representation volumes of 3–manifolds.

57M50; 51H20

1 Introduction

The representation volume of 3–manifolds is a beautiful theory, exhibiting rich connections with many branches of mathematics. The behavior of those volume functions appears to be quite mysterious; for example, their values are hard to predict except in a very few nice cases. On the other hand, for most motivating applications, it suffices to estimate the growth of such volumes for finite covers of the considered 3–manifold. In this paper, we intend to investigate the possibility of the latter, which is interesting as a topic on its own right.

To be more specific, let us introduce some basic notations and mention some known properties of the representation volume. Let G be either

$$\mathrm{Iso}_+\mathbb{H}^3 \cong \mathrm{PSL}(2;\mathbb{C}),$$

the orientation-preserving isometry group of the 3–dimensional hyperbolic geometry, or

$$\mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R}) \cong \mathbb{R}\times_{\mathbb{Z}}\widetilde{\mathrm{SL}}_2(\mathbb{R}),$$

Published: 15 August 2017 DOI: 10.2140/gt.2017.21.3159

the identity component of the isometry group of the Seifert geometry (see Brooks and Goldman [4]). For any closed orientable 3–manifold N and any representation $\rho\colon \pi_1 N \to G$, denote by $\mathrm{vol}_G(N,\rho)$ the (unsigned) volume of ρ. We denote the set of G–representation volumes of N by

$$\mathrm{vol}(N,G) = \{\mathrm{vol}_G(N,\rho) : \rho \text{ any representation } \pi_1 N \to G\},$$

which is a subset of the interval $[0,+\infty)$.

The following theorem contains a collection of fundamental facts in the theory of representation volumes; see Brooks and Goldman [3] and Reznikov [23].

Theorem 1.1 *Let N be a closed orientable 3–manifold.*

(1) *The sets of values $\mathrm{vol}(N, \mathrm{Iso}_+\mathbb{H}^3)$ and $\mathrm{vol}(N, \mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R}))$ are both finite. Hence the values*

$$\mathrm{HV}(N) = \max \mathrm{vol}(N, \mathrm{Iso}_+\mathbb{H}^3) \quad \textit{and} \quad \mathrm{SV}(N) = \max \mathrm{vol}(N, \mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R}))$$

exist in $[0,+\infty)$, depending only on N.

(2) *If N admits a hyperbolic geometric structure, then $\mathrm{HV}(N)$ equals the usual hyperbolic volume of N, reached by any discrete and faithful representation. A similar statement holds for $\mathrm{SV}(N)$ when N admits a Seifert geometric structure.*

(3) *If $P_1,\dots,P_s$ are the prime factors of N in the Kneser–Milnor decomposition, then*

$$\mathrm{HV}(N) = \mathrm{HV}(P_1) + \cdots + \mathrm{HV}(P_s).$$

A similar formula holds for $\mathrm{SV}(N)$.

(4) *For any map $f\colon M \to N$ between closed orientable 3–manifolds,*

$$\mathrm{HV}(M) \ge |\deg f| \cdot \mathrm{HV}(N).$$

The same comparison holds for $\mathrm{SV}(M)$ and $\mathrm{SV}(N)$.

The values $\mathrm{HV}(N)$ and $\mathrm{SV}(N)$ in the conclusion of Theorem 1.1(1) are called the *hyperbolic volume* and the *Seifert volume* of N, respectively. In light of Theorem 1.1(3), we assume from now on that all the closed orientable 3–manifolds considered are prime, unless specified otherwise. This is especially convenient when we speak of the geometric decomposition of the 3–manifold.

Remark Representation volumes were introduced and studied by R Brooks and W Goldman [3; 4] as a generalization of the simplicial volume originally due to M Gromov [10]. Among the eight 3–dimensional geometries of W P Thurston, $\mathbb{H}^3$ and

$\widetilde{\mathrm{SL}}_2(\mathbb{R})$ are the only two that yield nontrivial invariants, the hyperbolic volume and the Seifert volume, respectively. Recall that the *simplicial volume* of a closed orientable 3–manifold N roughly counts the minimal real number of singular tetrahedra to realize the fundamental class of N, and it is denoted by $\|N\|$. It is known that the sum of the classical hyperbolic volume of the hyperbolic pieces is equal to $v_3\|N\|$ (see Soma [26]), where v_3 is the volume of the ideal regular hyperbolic tetrahedron.

Like the simplicial volume, the volumes of Brooks–Goldman satisfy the *domination property*, as stated by Theorem 1.1(4). It follows that if either of the volumes $\mathrm{HV}(N)$ or $\mathrm{SV}(N)$ is positive, then the set of mapping degrees $D(M,N)$ of N by any given 3–manifold M must be finite. Unlike the simplicial volume, neither the hyperbolic volume nor the Seifert volume satisfies the *covering property*; see Derbez, Liu and Wang [5, Corollary 1.8], and Section 6 for some further discussion.

It can be inferred from Theorem 1.1 and the following remark that nonvanishing $\mathrm{HV}(N)$ or $\mathrm{SV}(N)$ contains interesting information about the topology of the 3–manifold N. However, such information seems difficult to characterize. For example, the vanishing or nonvanishing of $\mathrm{SV}(N)$ implies nothing about the behavior of $\mathrm{HV}(N)$ (see Brooks and Goldman [3, Sections 4 and 5]), and except for the geometric case (Theorem 1.1(2)), the geometry of pieces fails to detect the vanishing or nonvanishing of $\mathrm{HV}(N)$ or $\mathrm{SV}(N)$ either; see Derbez, Liu and Wang [5, Theorem 1.7]. On the other hand, the existence of some finite cover of N with nonvanishing representation volume turns out to be a more accessible question. An affirmative answer would be practically useful: it implies the finiteness of the set of mapping degrees as before. Motivated by that application, it has been discovered that any nongeometric graph manifold admits a finite cover of positive Seifert volume (see Derbez and Wang [7; 8]); a much more general construction that invokes Chern–Simons-theoretic calculations, and virtual properties of 3–manifolds shows that a right geometric piece implies virtually positive volume of the right geometry [5, Theorems 1.6]:

Theorem 1.2 *Suppose that N is a closed orientable nongeometric prime 3–manifold.*

(1) *If N contains at least one hyperbolic geometric piece, then the hyperbolic volume of some finite cover of N is positive.*

(2) *If N contains at least one Seifert geometric piece, then the Seifert volume of some finite cover of N is positive.*

Despite the seeming parallelism so far, the hyperbolic volume and the Seifert volume behave drastically differently with respect to finite covers. In this paper, we support this point by investigating two problems proposed in [5, Section 8]:

Problem 1.3 Estimate the growth of virtual hyperbolic volume and virtual Seifert volume.

Problem 1.4 Is the Seifert volume of a closed prime 3–manifold virtually positive if it has positive simplicial volume?

The main results of this paper address Problem 1.4 affirmatively (Theorem 1.5) and Problem 1.3 partially for 3–manifolds of positive simplicial volume (Theorem 1.7 and the following remark), showing that the growth of virtual Seifert volume is superlinear while the growth of virtual hyperbolic volume is linear. On Problem 1.4, the case of closed hyperbolic 3–manifolds is already known as a direct consequence of the much stronger virtual domination theorem of Sun [27] (quoted as Theorem 1.8 below); so essentially it remains to treat the case of nongeometric 3–manifolds (with only hyperbolic pieces). On Problem 1.3, it is easy to observe that the growth of virtual Seifert volume for a closed Seifert geometric 3–manifold is linear, indeed in a constant rate equal to its Seifert volume. Comparing with our result, we are left with the impression that the growth of virtual hyperbolic volume might be largely governed by the product of the simplicial volume with v_3, and the growth of virtual Seifert volume appears to be more sensitive to the geometric decomposition.

The main results of this paper are stated as Theorems 1.5 and 1.7:

Theorem 1.5 *If M is a closed orientable 3–manifold with positive simplicial volume, then there is a finite cover $\widetilde{M}$ of M with positive Seifert volume.*

Combining with results of Derbez, Liu, Sun and Wang [8; 5; 6], we infer immediately the following characterization:

Corollary 1.6 *Suppose that N is a closed orientable 3–manifold. Then the following three statements are equivalent:*

(1) *The set of mapping degrees $D(M,N)$ is finite for every closed orientable 3–manifold M.*
(2) *The Seifert volume of some finite cover of N is positive.*
(3) *At least one prime factor of N is Seifert geometric, or hyperbolic, or nongeometric.*

Theorem 1.7 *For any closed oriented 3–manifold M with nonvanishing simplicial volume, the set of values*

$$\left\{ \frac{\mathrm{SV}(M')}{[M':M]} \;\middle|\; M' \text{ any finite cover of } M \right\}$$

has no upper bound in $[0,+\infty)$.

Remark By contrast, it is evident by Reznikov [23, Theorem B] and Theorem 1.1 that the set of values

$$\left\{ \frac{\mathrm{HV}(M')}{[M' : M]} \;\middle|\; M' \text{ any finite cover of } M \right\}$$

is contained in the interval $[0, v_3 \|M\|]$.

Theorem 1.7 is significantly stronger than Theorem 1.5. Let us take a closer look at the geometric case to illustrate their difference in the proof. As mentioned, when M is assumed to be geometric, hence hyperbolic, Theorem 1.5 is implied by the following result of Sun [27], by taking N to be a target with positive Seifert volume:

Theorem 1.8 *For any closed oriented hyperbolic* 3*–manifold* M *and any closed oriented* 3*–manifold* N*, there is a finite cover* $\widetilde{M}$ *of* M *with a* π_1*–surjective degree-*2 *map* $f\colon \widetilde{M} \to N$.

Even though Theorem 1.8 is a powerful construction, employing deep theories including Kahn and Markovic [14], Liu and Markovic [17], Agol [1] and Wise [31] on building and separating certain quasiconvex subgroups in closed hyperbolic 3–manifold groups, the construction provides no control on the degree $[\widetilde{M} : M]$. So Theorem 1.7 stays beyond the reach of Theorem 1.8. Armed with a more recent result of A Gaifullin [9], we prove the following Theorem 1.9 based on Theorem 1.8. The improved construction is supplied with a desired efficient control of the mapping degree:

Theorem 1.9 *For any closed oriented hyperbolic* 3*–manifold* M*, there exists a positive constant* $c(M)$ *such that the following statement holds. For any closed oriented* 3*–manifold* N *and any* $\epsilon > 0$*, there exists a finite cover* M' *of* M *which admits a nonzero degree map* $f\colon M' \to N$ *such that*

$$\|M'\| \le c(M) \cdot |\deg f| \cdot (\|N\| + \epsilon).$$

To prove Theorems 1.5 and 1.7 in the nongeometric case, it is tempting to extend Theorems 1.8 and 1.9 to mixed 3–manifolds, but we do not have available tools for that project. Instead, we follow the framework of Derbez, Liu and Wang [5] and integrate the virtual domination theorems. The interaction between Theorem 1.8 and the fundamental construction for Theorem 1.2 is fairly direct and illustrating, so we present it and prove Theorem 1.5 as a warm-up. The proof of Theorem 1.7 is relatively more sophisticated, not only because of Theorem 1.9, but it requires some details of [5]. In particular, we introduce an auxiliary notion called *CI completion* to formalize a useful idea underlying the construction of [5] (see Section 5.2).

All the arguments are based on explicitly stated results, and the exposition is kept otherwise self-contained. The organization of this paper is as follows: The proofs of Theorems 1.5, 1.9 and 1.7 occupy Sections 3, 4 and 5, respectively. Section 2 includes preliminaries on 3–manifold topology and representation volume. Section 6 contains some further questions and observations.

Acknowledgement We are grateful to Ian Agol for valuable conversations. We also thank the anonymous referee for helpful comments. Liu is supported by the Recruitment Program of Global Youth Experts of China. Sun is partially supported by Grant No DMS-1510383 of the National Science Foundation of the United States. Wang is partially supported by Grant No 11371034 of the National Natural Science Foundation of China.

2 Preliminaries

In this section, we review the geometric decomposition of 3–manifolds and the theory of representation volumes.

2.1 Geometry and topology of 3–manifolds after Thurston

Let N be a connected compact prime orientable 3–manifold with toral or empty boundary. As a consequence of the geometrization of 3–manifolds [28; 29] achieved by G Perelman and Thurston, exactly one of the following cases holds:

- N is geometric, supporting one of the following eight geometries: $\mathbb{H}^3$, $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, $\mathbb{H}^2\times\mathbb{R}$, Sol, Nil, $\mathbb{R}^3$, $\boldsymbol{S}^3$ and $\boldsymbol{S}^2\times\mathbb{R}$ (where $\mathbb{H}^n$, $\mathbb{R}^n$ and $\boldsymbol{S}^n$ are the n–dimensional hyperbolic space, Euclidean space and spherical space, respectively).
- N has a canonical nontrivial geometric decomposition. In other words, there is a nonempty minimal union $\mathcal{T}_N\subset N$ of disjoint essential tori and Klein bottles in N, unique up to isotopy, such that each component of $N\setminus\mathcal{T}_N$ is either Seifert fibered or atoroidal. In the Seifert fibered case, the piece supports both the $\mathbb{H}^2\times\mathbb{R}$ geometry and the $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ geometry. In the atoroidal case, the piece supports the $\mathbb{H}^3$ geometry.

When N has nontrivial geometric decomposition, we call the components of $N\setminus\mathcal{T}_N$ the *geometric pieces* of N or, more specifically, *Seifert pieces* or *hyperbolic pieces* according to their geometry.

Traditionally, there is another decomposition introduced by Jaco and Shalen [12] and Johannson [13], known as the *JSJ decomposition*. When N contains no essential Klein bottles and has a nontrivial geometric decomposition, the JSJ decomposition of N

coincides with its geometric decomposition, so the cutting tori and the geometric pieces are also the *JSJ tori* and the *JSJ pieces*, respectively. Possibly after passing to a double cover of N, we may assume that N contains no essential Klein bottle.

A hyperbolic piece J can be realized as a complete hyperbolic 3–manifold of finite volume, unique up to isometry (by Mostow rigidity). Let J be a compact, orientable 3–manifold whose boundary consists of tori $T_1, \dots, T_p$ and whose interior admits a complete hyperbolic metric. Identify J with the complement of p disjoint cusps in the corresponding hyperbolic manifold; then ∂J has a Euclidean metric induced from the hyperbolic structure, and each closed Euclidean geodesic in ∂J has the induced length. The hyperbolic Dehn filling theorem of Thurston [28, Theorem 5.8.2] can be stated in the following form:

Theorem 2.1 *There is a constant $C > 0$ such that the closed 3–manifold $J(\zeta_1, \dots, \zeta_n)$ obtained by Dehn filling each T_i along a slope $\zeta_i \subset T_i$ admits a complete hyperbolic structure if each ζ_i has length greater than C. Moreover, with suitably chosen basepoints, $J(\zeta_1, \dots, \zeta_n)$ converges to the corresponding cusped hyperbolic 3–manifold in the Gromov–Hausdorff sense as the minimal length of ζ_i tends to infinity.*

A Seifert piece J of a nongeometric prime closed 3–manifold N supports both the $\mathbb{H}^2 \times \mathbb{R}$ geometry and the $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ geometry. In this paper, we are more interested in the latter case, so we describe the structure of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ geometric manifolds in the following. All the material can be found in [25].

We consider the group $\mathrm{PSL}(2;\mathbb{R})$ as the orientation-preserving isometries of the hyperbolic 2–space $\mathbb{H}^2 = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ with i as a basepoint. In this way $\mathrm{PSL}(2;\mathbb{R})$ is identified with the unit tangent bundle of $\mathbb{H}^2$, which has a natural Riemannian metric induced from $T\mathbb{H}^2$. Note that $\mathrm{PSL}(2;\mathbb{R})$ is a (topologically trivial) circle bundle over $\mathbb{H}^2$, but not isometric to $\mathbb{H}^2 \times \boldsymbol{S}^1$. Let $p\colon \widetilde{\mathrm{SL}}_2(\mathbb{R}) \to \mathrm{PSL}(2;\mathbb{R})$ be the universal covering of $\mathrm{PSL}(2;\mathbb{R})$ with the induced metric, then $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ is a line bundle over $\mathbb{H}^2$. For any $\alpha \in \mathbb{R}$, denote by $\mathrm{sh}(\alpha)$ the element of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ whose projection into $\mathrm{PSL}(2;\mathbb{R})$ is given by

$$\begin{pmatrix} \cos(2\pi\alpha) & -\sin(2\pi\alpha) \\ \sin(2\pi\alpha) & \cos(2\pi\alpha) \end{pmatrix}.$$

Then the set $\{\mathrm{sh}(n) \mid n \in \mathbb{Z}\}$ is the kernel of p, as well as the center of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, acting by integral translation along the fibers of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$. By extending this $\mathbb{Z}$–action on the fibers by the $\mathbb{R}$–action, we get the whole identity component of the isometry group of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$. To summarize, we have the following diagram of central extensions:

$$\begin{tikzcd}
\{0\} \arrow[r] \arrow[d] & \mathbb{Z} \arrow[r] \arrow[d] & \widetilde{\mathrm{SL}}_2(\mathbb{R}) \arrow[r] \arrow[d] & \mathrm{PSL}(2;\mathbb{R}) \arrow[r] \arrow[d] & \{1\} \arrow[d] \\
\{0\} \arrow[r] & \mathbb{R} \arrow[r] & \mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R}) \arrow[r] & \mathrm{PSL}(2;\mathbb{R}) \arrow[r] & \{1\}
\end{tikzcd}$$

In particular, the group $\mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R})$ is generated by $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ and the image of $\mathbb{R}$, which intersect with each other in the image of $\mathbb{Z}$. More precisely, we state the following useful lemma, which is easy to check.

Lemma 2.2 *We have the identification* $\mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R}) = \mathbb{R}\times_{\mathbb{Z}}\widetilde{\mathrm{SL}}_2(\mathbb{R})$, *where* $(x,h)\sim(x',h')$ *if and only if there exists an integer* $n\in\mathbb{Z}$ *such that* $x'-x=n$ *and* $h'=\mathrm{sh}(-n)\circ h$.

From [4] we know that a closed orientable 3–manifold J supports the $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ geometry — ie there is a discrete and faithful representation $\psi\colon \pi_1 J\to \mathrm{Iso}\,\widetilde{\mathrm{SL}}_2(\mathbb{R})$ — if and only if J is a Seifert fibered space with nonzero Euler number $e(J)$ and the base orbifold $\chi_{O(J)}$ has negative Euler characteristic.

2.2 Representation volumes of closed manifolds

In this subsection, we recall the definition of volume of representations. There are a few equivalent definitions, and we will only state one of them.

Given a semisimple, connected Lie group G and a closed oriented manifold M^n of the same dimension as the contractible space $X^n=G/K$, where K is a maximal compact subgroup of G. We can associate to each representation $\rho\colon \pi_1 M\to G$ a volume $\mathrm{vol}_G(M,\rho)$ in the following way.

First fix a G–invariant Riemannian metric g_X on X, and denote by ω_X the corresponding G–invariant volume form. Let $\widetilde{M}$ denote the universal covering of M. We think of the elements $\tilde{x}$ of $\widetilde{M}$ as the homotopy classes of paths $\gamma\colon[0,1]\to M$ with $\gamma(0)=x_0$, which are acted on by $\pi_1(M,x_0)$ by setting $[\sigma].\tilde{x}=[\sigma.\gamma]$, where the dot denotes the composition of paths.

A developing map $D_\rho\colon \widetilde{M}\to X$ associated to ρ is a $\pi_1 M$–equivariant map such that for any $x\in\widetilde{M}$ and $\alpha\in\pi_1 M$, we have

$$D_\rho(\alpha.x)=\rho(\alpha)D_\rho(x),$$

where $\rho(\alpha)$ acts on X as an isometry. Such a map does exist and can be constructed explicitly as in [2]: Fix a triangulation Δ_M of M; then it lifts to a triangulation $\Delta_{\widetilde{M}}$ of $\widetilde{M}$, which is $\pi_1 M$–invariant. Then fix a fundamental domain Ω of M in $\widetilde{M}$ such that the zero skeleton $\Delta^0_{\widetilde{M}}$ misses the frontier of Ω. Let $\{x_1, \dots, x_l\}$ be the vertices of $\Delta^0_{\widetilde{M}}$ in Ω, and let $\{y_1, \dots, y_l\}$ be any l points in X. We first set

$$D_\rho(x_i) = y_i, \quad i = 1, \dots, l.$$

Then extend D_ρ in a $\pi_1 M$–equivariant way to $\Delta^0_{\widetilde{M}}$: for any vertex x in $\Delta^0_{\widetilde{M}}$, there is a unique vertex x_i in Ω and $\alpha_x \in \pi_1 M$ such that $\alpha_x . x_i = x$, and we set $D_\rho(x) = \rho(\alpha_x)^{-1} D_\rho(x_i)$. Finally we extend D_ρ to edges, faces, etc, and n–simplices of $\Delta_{\widetilde{M}}$ by straightening their images to totally geodesics objects using the homogeneous metric on the contractible space X. This map is unique up to equivariant homotopy. Then $D^*_\rho(\omega_X)$ is a $\pi_1 M$–invariant closed n–form on $\widetilde{M}$, which therefore can be thought of as a closed n–form on M. Then we define

$$\mathrm{vol}_G(M, \rho) = \int_M D^*_\rho(\omega_X) = \sum_{i=1}^{s} \epsilon_i \, \mathrm{vol}_X(D_\rho(\widetilde{\Delta}_i))$$

Here $\{\Delta_1, \dots, \Delta_s\}$ are the n–simplices of Δ_M, $\widetilde{\Delta}_i$ is a lift of Δ_i and $\epsilon_i = \pm 1$ depends on whether $D_\rho|_{\widetilde{\Delta}_i}$ preserves the orientation or not.

3 Positive simplicial volume implies virtually positive Seifert volume

In this section, we adapt Theorem 1.8 to the framework of [5] to prove Theorem 1.5.

3.1 Virtual representation through geometric decomposition

We recall some results from [5]. The following additivity principle allows us to compute the representation volume by information on the JSJ pieces. It is proved by using the relation between the representation volume and the Chern–Simons theory.

Theorem 3.1 (additivity principle [5, Theorem 3.5]; see also [8]) *Let M be an oriented closed 3–manifold with JSJ tori $T_1, \dots, T_r$ and JSJ pieces $J_1, \dots, J_k$, and let $\zeta_1, \dots, \zeta_r$ be slopes on $T_1, \dots, T_r$, respectively.*

Suppose that G is either $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ or $\mathrm{PSL}(2; \mathbb{C})$, that

$$\rho\colon \pi_1(M) \to G$$

is a representation vanishing on the slopes ζ_i, and that $\hat{\rho}_i\colon \pi_1(\hat{J}_i) \to G$ are the induced representations, where $\hat{J}_i$ is the Dehn filling of J_i along slopes adjacent to its boundary, with the induced orientations. Then

$$\mathrm{vol}_G(M, \rho) = \mathrm{vol}_G(\hat{J}_1, \hat{\rho}_1) + \mathrm{vol}_G(\hat{J}_2, \hat{\rho}_2) + \cdots + \mathrm{vol}_G(\hat{J}_k, \hat{\rho}_k).$$

The following simple lemma suggests that we should focus on those JSJ pieces whose groups have nonelementary images under ρ.

Lemma 3.2 [5, Lemma 3.6] *Suppose that G is either $\mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R})$ or $\mathrm{PSL}(2;\mathbb{C})$ and that M is a closed oriented 3–manifold. If $\rho\colon \pi_1 M \to G$ has image either infinite cyclic or finite, then $\mathrm{vol}_G(M,\rho) = 0$.*

The existence of a class inversion for the target group played an important role in [5] for constructing virtual representation of mixed 3–manifold groups. Here we quote the following definition. An intimately related notion called CI completion is introduced and studied in this paper when we prove Theorem 1.7 (see Section 5.2).

Definition 3.3 [5, Definition 5.1] Let $\mathscr{G}$ be an arbitrary group and $\{[A_i]\}_{i\in I}$ be a collection of conjugacy classes of abelian subgroups. By a *class inversion* with respect to $\{[A_i]\}_{i\in I}$, we mean an outer automorphism $[\nu] \in \mathrm{Out}(\mathscr{G})$ such that for any representative abelian subgroup A_i of each $[A_i]$, there is a representative automorphism $\nu_{A_i}\colon \mathscr{G} \to \mathscr{G}$ of $[\nu]$ that preserves A_i, taking every $a \in A_i$ to its inverse. We say $\mathscr{G}$ is *class invertible* with respect to $\{[A_i]\}_{i\in I}$ if there exists a class inversion. We often ambiguously call any collection of representative abelian subgroups $\{A_i\}_{i\in I}$ a class invertible collection, and call any representative automorphism ν a class inversion.

Now we state the following fundamental construction about virtual representation extensions. It uses works of Przytycki and Wise [20; 21; 22] (and [31; 11]) and Rubinstein and Wang [24] (see also [16]) to understand virtual properties of 3–manifolds with nontrivial geometric decomposition.

Theorem 3.4 [5, Theorem 5.2] *Let $\mathscr{G}$ be a group and M be an irreducible orientable closed 3–manifold with nontrivial JSJ decomposition. For a fixed JSJ piece $J_0 \subset M$, suppose a representation*

$$\rho_0\colon \pi_1(J_0) \to \mathscr{G}$$

satisfies the following:

- *ρ_0 has nontrivial kernel restricted to $\pi_1(T)$ for every boundary torus $T \subset \partial J_0$; and*
- *$\rho_0(\pi_1(T))$ forms a class invertible collection of abelian subgroups of $\mathscr{G}$ for every boundary torus $T \subset \partial J_0$.*

Then there exist a finite regular cover

$$\kappa\colon \tilde{M} \to M$$

and a representation

$$\tilde{\rho}\colon \pi_1(\tilde{M}) \to \mathscr{G}$$

satisfying the following:

- *for one or more elevations $\widetilde{J}_0$ of J_0, the restriction of $\widetilde{\rho}$ to $\pi_1(\widetilde{J}_0)$ is, up to a class inversion, conjugate to the pullback $\kappa^*(\rho_0)$; and*
- *for any elevation $\widetilde{J}$ other than the above, of any geometric piece J, the restriction of $\widetilde{\rho}$ to $\pi_1(\widetilde{J})$ is cyclic, possibly trivial.*

3.2 Proof of Theorem 1.5

Now we are ready to prove Theorem 1.5, and here is a sketch of the strategy. Since we can suppose that the manifold has a hyperbolic JSJ piece, Theorem 1.8 gives a virtual representation of the hyperbolic piece with positive Seifert volume. Then, with Lemma 3.6, Theorem 3.4 extends the virtual representation to the whole manifold, and the volume of the virtual representation can be calculated by Theorem 3.1 and Lemma 3.2.

By Theorems 1.2 and 1.8, we may assume that M has nontrivial JSJ decomposition and contains a hyperbolic JSJ piece J in M. Suppose ∂J is a union of tori $T_1, \dots, T_k$. Let α_i be a slope on T_i; then call $\alpha = \{\alpha_1, \dots, \alpha_k\}$ a slope on ∂M. Denote by $J(\alpha)$ the closed orientable 3–manifold obtained by Dehn filling of k solid tori $S_1, \dots, S_k$ to J along α. We can choose α so that $J(\alpha)$ is a hyperbolic 3–manifold (Theorem 2.1).

Take a closed orientable manifold N of nonvanishing Seifert volume. For example, a circle bundle N with Euler class $e \neq 0$ over a closed surface of Euler characteristic $\chi < 0$ works: in fact, for such N,

$$\mathrm{SV}(N) = \frac{4\pi^2|\chi|^2}{|e|} > 0.$$

By Theorem 1.8 there is a finite cover $q\colon Q \to J(\alpha)$ such that Q dominates N, therefore $\mathrm{SV}(Q) > 0$. Let $S = \bigcup S_i$; then $S' = q^{-1}(S) \subset Q$ is a union of solid tori and $J' = Q \setminus S'$ is a connected 3–manifold which covers J. Moreover, Q is obtained by Dehn filling S' to J' along α', where α' is a slope of $\partial J'$ which covers α (ie each component of α' is an elevation of a component of α and $Q = J'(\alpha')$).

Fix J' and α' for the moment. Let $\widetilde{J}$ be a finite covering of J' and $\widetilde{\alpha}$ be the slope of $\partial\widetilde{J}$ which covers α'; then $\mathrm{SV}(\widetilde{J}(\widetilde{\alpha})) > 0$. This is because the covering $\widetilde{J} \to J'$ extends to a branched covering (which is a nonzero degree map) $\widetilde{J}(\widetilde{\alpha}) \to J'(\alpha')$ and $\mathrm{SV}(J'(\alpha')) = SV(Q) > 0$.

According to [5, Proposition 4.2], there is a finite cover $p\colon \widetilde{M} \to M$ such that each JSJ piece $\widetilde{J}$ of $\widetilde{M}$ that covers J factors through J'. In particular, in the notations we have just used, $\mathrm{SV}(\widetilde{J}(\widetilde{\alpha})) > 0$. To simplify the notations, we rewrite $\widetilde{M}$, $\widetilde{J}$ and $\widetilde{\alpha}$ as

M, J and α. Since Theorem 1.5 concludes with a virtual property, we need only to prove the following statement:

Theorem 3.5 *Suppose M is a closed orientable 3–manifold with nontrivial JSJ decomposition and there is a JSJ piece J and a slope α of ∂J such that $\mathrm{SV}(J(\alpha)) > 0$. Then there is a finite cover $\widetilde{M}$ of M such that $\mathrm{SV}(\widetilde{M}) > 0$.*

We are going to apply Theorem 3.4 to prove Theorem 3.5. So we first need to check that the 3–manifold M and the local representation $\rho\colon \pi_1(J) \to G$ (which gives positive representation volume for $J(\alpha)$) in Theorem 3.5 meet the two conditions of Theorem 3.4.

We first write a presentation of $\pi_1(J(\alpha))$ from $\pi_1(J)$ by attaching k relations from Dehn fillings. Let $G = \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ be the identity component of $\mathrm{Iso}\, \widetilde{\mathrm{SL}}_2(\mathbb{R})$, the isometry group of the Seifert space $\widetilde{\mathrm{SL}}_2(\mathbb{R})$. Then the condition $\mathrm{SV}(J(\alpha)) > 0$ implies that there is a representation $\rho\colon \pi_1(J) \to G$ such that, for each component T_i of ∂J, $\rho(\pi_1(T_i))$ is a (possibly trivial) cyclic group. Moreover, ρ extends to $\hat{\rho}\colon \pi_1(J(\alpha)) \to G$ such that $V_G(J(\alpha), \hat{\rho}) > 0$. So the first condition of Theorem 5.2 of [5] is satisfied. The following lemma, which strengthens [5, Lemma 6.1(2)], implies that the second condition of Theorem 3.4 is also satisfied.

Lemma 3.6 $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ *is class invertible with respect to all its cyclic subgroups, and a class inversion can be realized by the conjugation of any* $\nu \in \mathrm{Iso}\, \widetilde{\mathrm{SL}}_2(\mathbb{R}) \setminus \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$. *The corresponding action on* $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ *preserves the orientation.*

Proof There are short exact sequences of groups

$$0 \to \mathbb{R} \to \mathrm{Iso}\, \widetilde{\mathrm{SL}}_2(\mathbb{R}) \xrightarrow{p} \mathrm{Iso}\, \mathbb{H}^2 \to 1$$

and

$$0 \to \mathbb{R} \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}) \xrightarrow{p} \mathrm{Iso}_+ \mathbb{H}^2 \to 1.$$

Recall that there are no orientation-reversing isometries in the $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ geometry.

For each element ν in the component of $\mathrm{Iso}\, \widetilde{\mathrm{SL}}_2(\mathbb{R})$ not containing the identity, ν reverses the orientation of $\mathbb{R}$ (the center of $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$). So $\nu r \nu^{-1} = r^{-1}$ for any $r \in \mathbb{R}$, and $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ is class invertible with respect to its center $\mathbb{R}$. A class inversion can be realized by the conjugation of any $\nu \in \mathrm{Iso}\, \widetilde{\mathrm{SL}}_2(\mathbb{R}) \setminus \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$, and the corresponding action on $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ preserves the orientation. Actually, this part of the proof is the same as the proof of [5, Lemma 6.1(ii)].

In the following, we suppose that $\langle \alpha \rangle$ is a cyclic subgroup of $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ generated by a noncentral element α.

For each nontrivial element a in $\mathrm{Iso}_{+}\mathbb{H}^2$, it is straightforward to see that there exists a reflection about a geodesic l_a in $\mathbb{H}^2$ that conjugates a to its inverse. The l_a can be chosen as (i) passing through the rotation center when a is elliptic; (ii) perpendicular with the axis of a when a is hyperbolic; (iii) passing through the fixed point when a is parabolic.

By the discussion in the last paragraph and the exact sequences, there exists an element $\nu \in \mathrm{Iso}\,\widetilde{\mathrm{SL}}_2(\mathbb{R}) \setminus \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ such that $p(\nu)$ is a reflection of $\mathbb{H}^2$ conjugating $p(\alpha)$ to its inverse, namely $p(\nu^{-1}\alpha\nu) = p(\alpha^{-1})$. We claim that

$$\nu^{-1}\alpha\nu = \alpha^{-1}.$$

In fact, by the short exact sequences above, we have that $\nu^{-1}\alpha\nu = \alpha^{-1}r$ for some r in the center $\mathbb{R}$. Since $p(\nu)$ is a reflection, ν^2 is central, so

$$\alpha = \nu^{-2}\alpha\nu^2 = \nu^{-1}(\alpha^{-1}r)\nu = (\nu^{-1}r\nu)(\nu^{-1}\alpha\nu)^{-1} = r^{-1}(\alpha^{-1}r)^{-1} = \alpha r^{-2}.$$

Here we used the fact that ν is a class inversion for $\langle r\rangle$. So r^{-2} is trivial, and r is trivial as the center is torsion-free. This verifies the claim. We conclude that ν realizes a class inversion of the cyclic subgroup $\langle\alpha\rangle$ of $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$.

For two elements $\alpha_1, \alpha_2 \in \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$, there exist $\nu_1, \nu_2 \in \mathrm{Iso}\,\widetilde{\mathrm{SL}}_2(\mathbb{R}) \setminus \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ such that $\nu_i^{-1}\alpha_i\nu_i = \alpha_i^{-1}$, and there also exists $\beta \in \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ such that $\nu_1 = \beta\nu_2$. Then the conjugation of ν_1 on $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ equals the composition of the conjugation of ν_2 with the conjugation of β. Since $\beta \in \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$, the conjugations of ν_1 and ν_2 represent the same element in $\mathrm{Out}(\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}))$

So $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ is class invertible with respect to all its cyclic subgroups, and a class inversion can be realized by the conjugation of any element in $\mathrm{Iso}\,\widetilde{\mathrm{SL}}_2(\mathbb{R}) \setminus \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$, and the corresponding action on $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ preserves the orientation. □

Proof of Theorem 3.5 Fix J, α and $\rho\colon \pi_1(J) \to G$ as in our previous discussion, and denote them by J_0, α_0, and ρ_0 to match the notations of Theorem 3.4. Since $\rho_0\colon \pi_1(J_0) \to G$ meets the two conditions of Theorem 3.4, we can virtually extend ρ_0 to some $\widetilde{\rho}\colon \pi_1(\widetilde{M}) \to G$ which satisfies the conclusion of Theorem 3.4.

By the additivity principle (Theorem 3.1), we need only to compute the representation volume for each JSJ piece of $\widetilde{M}$, then add the volumes together to compute $V_G(\widetilde{M}, \widetilde{\rho})$. By Theorem 3.4 and Lemma 3.2, only those elevations $\widetilde{J}_0$ of J_0 such that the restriction of $\widetilde{\rho}$ to $\pi_1(\widetilde{J}_0)$ is conjugate to the pullback $\kappa^*(\rho_0)$, up to a class inversion, could contribute to the Seifert representation volume of $\widetilde{M}$.

By Lemma 3.6, the class inversions can be realized by conjugations of orientation-preserving isomorphisms of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, therefore the volumes of all these elevations are positive multiples of $V_G(J_0(\alpha_0), \widehat{\rho}_0) > 0$. So the Seifert representation volume of $\widetilde{M}$ with respect to $\widetilde{\rho}$ is positive, which implies $\mathrm{SV}(\widetilde{M}) > 0$. □

The completion of the proof of Theorem 3.5 also completes the proof of Theorem 1.5. We can reformulate what we have done in this section with the following proposition:

Proposition 3.7 *Let M be an orientable closed mixed 3–manifold and J_0 be a distinguished hyperbolic JSJ piece of M. Suppose that $\widehat{J}_0$ is a closed hyperbolic Dehn filling of J_0 by sufficiently long boundary slopes.*

(1) *For any finite cover $\widehat{J}_0'$ of $\widehat{J}_0$ and any representation*

$$\eta\colon \pi_1(\widehat{J}_0') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}),$$

there exist a finite cover

$$\widetilde{M}' \to M$$

and a representation

$$\rho\colon \pi_1(\widetilde{M}') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

with the following properties:

- *For one or more elevations $\widetilde{J}'$ of J_0 contained in $\widetilde{M}'$, the covering $\widetilde{J}' \to J_0$ factors through a covering $\widetilde{J}' \to J_0'$, where $J_0' \subset \widehat{J}_0'$ denotes the unique elevation of $J_0 \subset \widehat{J}_0$. The restriction of ρ to $\pi_1(\widetilde{J}')$ is conjugate to either the pullback $\beta^*(\eta)$ or the pullback $\beta^*(\nu\eta)$, where ν is a class inversion and β is the composition of the maps*
$$\widetilde{J}' \xrightarrow{\mathrm{cov}} J_0' \xrightarrow{\mathrm{fill}} \widehat{J}_0'.$$
- *For any elevation $\widetilde{J}'$ other than the above, of any JSJ piece J of M, the restriction of ρ to $\pi_1(\widetilde{J})$ has cyclic image, possibly trivial.*

(2) $\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{M}', \rho)$ *is a positive multiple of* $\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0', \eta)$.

Remark The first part of Proposition 3.7 is a specialized refined statement of Theorem 3.4; the second part supplies a slot to connect with Theorem 1.8. Therefore, Theorem 1.5 is a consequence of Proposition 3.7 and Theorem 1.8. The stronger result, Theorem 1.7, will follow from an efficient version of this proposition (Theorem 5.1) plus the efficient virtual domination (Theorem 1.9).

4 Efficient virtual domination by hyperbolic 3–manifolds

In this section, we employ the work of Gaifullin [9] to derive Theorem 1.9 from Theorem 1.8. We quote the statement below for convenience.

Theorem 1.9 *For any closed oriented hyperbolic* 3*–manifold* M*, there exists a positive constant* $c(M)$ *such that the following statement holds. For any closed oriented* 3*–manifold* N *and any* $\epsilon > 0$*, there exists a finite cover* M' *of* M *which admits a nonzero degree map* $f\colon M' \to N$ *such that*

$$\|M'\| \le c(M) \cdot |\deg f| \cdot (\|N\| + \epsilon).$$

Remark In fact, the same statement holds for any closed orientable manifold which virtually dominates all closed orientable manifolds of the same dimension. For dimension 3, all hyperbolic manifolds have such property [27]. For any arbitrary dimension, manifolds with this property have been discovered by Gaifullin [9]. The 3–dimensional example M_{Π^3} of Gaifullin is not a hyperbolic manifold, but we point out that a constant $c_0 = 24v_8/v_3 \approx 86.64$ is sufficient for this case, where v_8 is the volume of the ideal regular hyperbolic octahedron and v_3 is the volume of the ideal regular hyperbolic tetrahedron.

4.1 URC manifolds

As introduced by Gaifullin [9], a closed orientable (topological) n–manifold M is said to have the property of *universal realization of cycles* (*URC*) if every homology class of $H_n(X;\mathbb{Z})$ of an arbitrary topological space X has a positive integral multiple which can be realized by the fundamental class of a finite cover M' of M, via a map $f\colon M' \to X$.

For any arbitrary dimension n, Gaifullin shows that examples of URC n–manifolds can be obtained by taking some 2^n–sheeted cover

$$M_{\Pi^n}$$

of some n–dimensional orbifold Π^n. More precisely, the underlying topology space of Π^n is the *permutahedron*, namely, the polyhedron combinatorially isomorphic to the convex hull of the points $(\sigma(1), \dots, \sigma(n+1))$ of $\mathbb{R}^{n+1}$, where σ runs over all permutations of $\{1, \dots, n+1\}$. The orbifold structure of Π^n is given so that each codimension-1 face is a reflection wall, so each codimension-k face is the local fixed point set of a $\mathbb{Z}_2^k$–action. The abelian characteristic cover of Π^n on which $H_1(\Pi^n;\mathbb{Z}_2) \cong \mathbb{Z}_2^n$ acts is the orientable closed n–manifold M_{Π^n}. In particular, M_{Π^n} can be obtained by facet pairing of 2^n permutahedra.

The following quantitative version of Gaifullin's proof [9, Section 5] is important for our application. Recall that a (compact) *pseudo-n–manifold* is a finite simplicial complex in which each simplex is contained in an n–simplex and each $(n-1)$–simplex is contained in exactly two n–simplices. Topologically, a pseudo-n–manifold is just a manifold away from its codimension-2 skeleton. A *strongly connected orientable* pseudo-n–manifold means that, away from the codimension-2 skeleton, the manifold is connected and orientable, or equivalently that the n–dimensional integral homology is isomorphic to $\mathbb{Z}$. In particular, the concept of (unsigned) mapping degree can be extended similarly to maps between strongly connected orientable pseudo-n–manifolds.

Theorem 4.1 (see [9, Proposition 5.3]) *For any strongly connected orientable pseudo-n–manifold Z, there exists a finite cover M'_{Π^n} of M_{Π^n} and a nonzero degree map $f_1\colon M'_{\Pi^n} \to Z$ such that*

$$\#\{n\text{–permutahedra of } M'_{\Pi^n}\} = (n+1)! \cdot |\deg f_1| \cdot \#\{n\text{–simplices of } Z\}.$$

Remark The map f_1 is as asserted by [9, Proposition 5.3]. The cover $\widehat{M}_{\Pi^n} = \mathcal{U}_{\Pi^n} / \Gamma_H$ there is rewritten as M'_{Π^n} in our notation. To compare with the statement of [9, Proposition 5.3], the index $|W : \Gamma_H|$ there equals the number of permutahedra in M'_{Π^n} here; the notation $|A|$ there stands for the number of n–simplices in the barycentric subdivision of Z, which equals $(n+1)!$ times the number of n–simplices of Z here. For dimension 3, all orientable closed hyperbolic manifold are known to be URC [27].

4.2 Virtual domination through URC 3–manifolds

We combine the results of [9; 27] to prove Theorem 1.9. The following lemma allows us to create an efficient virtual realization of the fundamental class of N.

Lemma 4.2 *For any closed oriented n–manifold N and any $\epsilon > 0$, there exists a connected oriented pseudo-n–manifold Z and a nonzero degree map $f\colon Z \to N$ such that*

$$\#\{n\text{–simplices of } Z\} \le |\deg f| \cdot (\|N\| + \epsilon).$$

Proof By the definition of the simplicial volume, for any $\epsilon > 0$ there exists a singular cycle

$$\alpha = \sum_{i=1}^{k} s_i \sigma_i \in Z_n(N, \mathbb{R})$$

such that $[\alpha] = [N] \in H_n(N, \mathbb{R})$ and

$$\sum_{i=1}^{k} |s_i| < \|N\| + \epsilon.$$

Here the s_i are real numbers and the σ_i are maps from the standard oriented n–simplex to N.

Since

$$\sum_{i=1}^{k} x_i \sigma_i \in Z_n(N, \mathbb{R}) \quad \text{and} \quad \left[\sum_{i=1}^{k} x_i \sigma_i\right] = [N] \in H_n(N, \mathbb{R})$$

can be expressed as linear equations with integer coefficients, they have a rational solution $(r_1, \dots, r_k)$ close to $(s_1, \dots, s_k)$ such that $r_i \in \mathbb{Q}$ and

$$\sum_{i=1}^{k} |r_i| < \|N\| + \epsilon.$$

In particular, $\left[\sum_{i=1}^{k} r_i \sigma_i\right] = [N] \in H_n(N, \mathbb{R})$ holds. Here we can suppose that each r_i is nonnegative, by reversing the orientation of σ_i if necessary.

Let the least common multiple of the denominators of r_i be denoted by m; then

$$\beta = m\left(\sum_{i=1}^{k} r_i \sigma_i\right) = \sum_{i=1}^{k} (m r_i) \sigma_i \in Z_n(N; \mathbb{Z})$$

is an integer linear combination of σ_i and $[\beta] = m[N] \in H_n(N, \mathbb{R})$.

Here we take $m r_i$ copies of the standard oriented n–simplex that is mapped as σ_i for $i = 1, 2, \dots, k$. The condition that $\sum_{i=1}^{k} (m r_i) \sigma_i$ be an n–cycle implies that we can find a pairing of all the $(n-1)$–dimensional faces of the collection of copies of the σ_i such that each such pair is mapped to the same singular $(n-1)$–simplex in N, with opposite orientation.

This pairing allows us to build an oriented pseudomanifold Z' (possibly disconnected). It is given by taking $\sum_{i=1}^{k} m r_i$ copies of the standard oriented n–simplex and pasting them together by the pairing given above. Then the singular n–simplices $\{\sigma_i\}_{i=1}^{k}$ induces a map $f_0\colon Z' \to N$.

Let $[Z']$ be the homology class in $H_n(Z')$ which is represented by the n–cycle which takes each oriented n–simplex in Z' exactly once. It is easy to see that $f_0([Z']) =$

$[\beta] = m[N]$, so f_0 has mapping degree $\deg f_0 = m$. Moreover, the number of n–simplices in Z' is just

$$\sum_{i=1}^{k}(mr_i) = m\left(\sum_{i=1}^{k} r_i\right) < m(\|N\| + \epsilon) = \deg f_0 \cdot (\|N\| + \epsilon).$$

If Z' is connected, we are done with the proof. If Z' is disconnected, take the component Z of Z' such that

$$\frac{\deg(f_0|_Z)}{\#\{n\text{–simplices of } Z\}}$$

is not smaller than the corresponding number for all the other components of Z'. Then $f = f_0|_Z$ satisfies the desired condition in this lemma. □

4.2.1 Construction of (M', f) Let M be a closed orientable hyperbolic 3–manifold and N be any closed orientable 3–manifold. Given any constant $\epsilon > 0$, denote by

$$p\colon Z \to N$$

a virtual realization of the fundamental class of N by a strongly connected orientable pseudo-3–manifold, as guaranteed by Lemma 4.2. Take a finite cover M'_{Π^3} of Gaifullin's URC 3–manifold M_{Π^3} and an efficient domination map

$$f_1\colon M'_{\Pi^n} \to Z,$$

which come from Theorem 4.1. Take a finite cover $\widetilde{M}$ of M and a π_1–surjectively 2–domination map

$$f_2\colon \widetilde{M} \to M_{\Pi^3},$$

which comes from Theorem 1.8. Then there exists a unique finite cover M' of M, up to isomorphism of covering spaces, and a unique π_1–surjective 2–domination map $f'_2\colon M' \to \widetilde{M}_{\Pi^3}$ that fits into the following commutative diagram of maps:

$$\begin{tikzcd} M' \arrow[r, "f'_2"] \arrow[d] & M'_{\Pi^3} \arrow[d] \\ \widetilde{M} \arrow[r, "f_2"] & M_{\Pi^3} \end{tikzcd}$$

Indeed, M' is the cover of $\widetilde{M}$ that corresponds to the subgroup $(f_{2\sharp})^{-1}(\pi_1(M'_{\Pi^3}))$ of $\pi_1(\widetilde{M})$ (after choosing some auxiliary basepoints). The finite cover M' of M and the composed map

$$f\colon M' \xrightarrow{f'_2} \widetilde{M}_{\Pi^3} \xrightarrow{f_1} Z \xrightarrow{p} N$$

are the claimed objects of Theorem 1.9.

4.2.2 Verification With the notations above, the commutative diagram above implies

$$\frac{\|M'\|}{\|\widetilde{M}\|} = [M' : \widetilde{M}] = [M'_{\Pi^3} : M_{\Pi^3}] = \frac{\#\{\text{permutahedra of } M'_{\Pi^3}\}}{\#\{\text{permutahedra of } M_{\Pi^3}\}}.$$

Observe that there are $2^3 = 8$ permutahedra in Gaifullin's URC 3–manifold M_{Π^3}. On the other hand, by Theorem 4.1 and Lemma 4.2, the construction of M'_{Π^3} and Z yields

$$\begin{aligned}\#\{\text{permutahedra of } M'_{\Pi^3}\} &= 4! \cdot |\deg f_1| \cdot \#\{\text{tetrahetra of } Z\} \\ &< 24 \cdot |\deg f_1| \cdot |\deg p| \cdot (\|N\| + \epsilon) \\ &= \tfrac{24}{2} \cdot |\deg f_2'| \cdot |\deg f_1| \cdot |\deg p| \cdot (\|N\| + \epsilon) \\ &= 12 \cdot |\deg f| \cdot (\|N\| + \epsilon).\end{aligned}$$

Therefore,

$$\|M'\| < \tfrac{1}{8}(12 \cdot |\deg f| \cdot (\|N\| + \epsilon) \cdot \|\widetilde{M}\|) = c_0 \cdot |\deg f| \cdot (\|N\| + \epsilon),$$

where the constant c_0 is taken to be

$$c_0 = \tfrac{3}{2}\|\widetilde{M}\|.$$

Note that the constant $c_0 > 0$ depends only on the hyperbolic 3–manifold M, because $\widetilde{M}$ is constructed by Theorem 1.8 without referring to N or ϵ. In this proof, we only applied Theorem 1.8 for the domain M_{Π^3}, not for a general 3–manifold.

This completes the proof of Theorem 1.9.

4.3 Virtual Seifert volume of closed hyperbolic 3–manifolds

We have mentioned in the introduction that Theorem 1.5 for closed hyperbolic 3–manifolds follows directly from Theorem 1.8. Similarly, Theorem 1.7 for hyperbolic closed 3–manifolds is a corollary of Theorem 1.9.

Corollary 4.3 *For any closed oriented hyperbolic 3–manifold M, the set of values*

$$\left\{ \frac{\mathrm{SV}(M')}{[M' : M]} \;\middle|\; M' \text{ any finite cover of } M \right\}$$

is not bounded.

Proof Take a closed orientable manifold N of nonvanishing Seifert volume and vanishing simplicial volume. For example, a circle bundle N with Euler class $e \neq 0$ over a closed surface of Euler characteristic $\chi < 0$ works.

For every positive integer n, apply Theorem 1.9 with $\epsilon = 1/n$. There exists a finite cover $M_n \to M$ and a nonzero degree map $f_n\colon M_n \to N$ such that

$$\|M\| \cdot [M_n : M] = \|M_n\| \leq c(M) \cdot |\deg f_n| \cdot \left(\|N\| + \frac{1}{n}\right) = c(M) \cdot |\deg f_n| \cdot \frac{1}{n}.$$

So we have

$$[M_n : M] \leq \frac{c(M) \cdot |\deg f_n|/n}{\|M\|}.$$

Since $\mathrm{SV}(M_n) \geq |\deg f_n| \cdot \mathrm{SV}(N)$, we have

$$\frac{\mathrm{SV}(M_n)}{[M_n : M]} \geq \frac{|\deg f_n| \cdot \mathrm{SV}(N)}{(c(M) \cdot |\deg f_n|/n)/\|M\|} = n \cdot \frac{\|M\| \cdot \mathrm{SV}(N)}{c(M)}.$$

Since $K = \|M\| \cdot \mathrm{SV}(N)/c(M)$ is a positive constant, $\{\mathrm{SV}(M_n)/[M_n : M]\}$ is not a bounded sequence, so we are done. □

5 Positive simplicial volume implies unbounded virtual Seifert volume

In this section, we prove Theorem 1.7 following the strategy of the proof of Theorem 1.5 summarized in the remark following Proposition 3.7. The main body of the proof is the following theorem which produces virtual Seifert representations with controlled volume, (compare Proposition 3.7).

Theorem 5.1 *Let M be an orientable closed mixed 3–manifold and J_0 be a distinguished hyperbolic JSJ piece of M. Suppose that $\widehat{J}_0$ is a closed hyperbolic Dehn filling of J_0 by sufficiently long boundary slopes.*

(1) *For any finite cover $\widehat{J}_0'$ of $\widehat{J}_0$ and any representation*

$$\eta\colon \pi_1(\widehat{J}_0') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}),$$

there exist a finite cover

$$\widetilde{M}' \to M$$

and a representation

$$\rho\colon \pi_1(\widetilde{M}') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

with the following properties:

- *For one or more elevations $\widetilde{J}'$ of J_0 contained in $\widetilde{M}'$, the covering $\widetilde{J}' \to J_0$ factors through a covering $\widetilde{J}' \to J_0'$, where $J_0' \subset \widehat{J}_0'$ denotes the unique elevation of $J_0 \subset \widehat{J}_0$. The restriction of ρ to $\pi_1(\widetilde{J}')$ is conjugate to either the pullback*

$\beta^*(\eta)$ or the pullback $\beta^*(\nu\eta)$, where ν is a class inversion in Lemma 3.6 and β is the composition of the maps

$$\widetilde{J}' \xrightarrow{\mathrm{cov}} J_0' \xrightarrow{\mathrm{fill}} \widehat{J}_0'.$$

- *For any elevation $\widetilde{J}'$ other than the above, of any JSJ piece J of M, the restriction of ρ to $\pi_1(\widetilde{J})$ has cyclic image, possibly trivial.*

(2) *Furthermore, there exists a positive constant α_0 depending only on M and the Dehn filling $J_0 \to \widehat{J}_0$ such that for any $\widehat{J}_0'$ and η as above, the asserted $\widetilde{M}'$ and ρ can be constructed so that the sum of the covering degrees $[\widetilde{J}' : J_0]$ over all the elevations $\widetilde{J}'$ of the β–pullback type equals $\alpha_0 \cdot [\widetilde{M}' : M]$. Therefore,*

$$\frac{\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{M}', \rho)}{[\widetilde{M}' : M]} = \alpha_0 \cdot \frac{\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0', \eta)}{[\widehat{J}_0' : \widehat{J}_0]}.$$

The rest of this section is devoted to the proof of Theorem 5.1, before which we derive Theorem 1.7 from Theorem 5.1 and Corollary 4.3.

5.1 Proof of Theorem 1.7

Since we have proved Theorem 1.7 for hyperbolic 3–manifolds (Corollary 4.3), we may assume that M is nongeometric with at least one hyperbolic piece, or in other words, *mixed*. The mixed case is derived from the hyperbolic case and Theorem 5.1.

Take a hyperbolic piece J of M and let $\widehat{J}$ be a closed hyperbolic Dehn filling of J. By Corollary 4.3, there are finite covers $\{\widehat{J}_n'\}$ of $\widehat{J}$ such that

$$\frac{\mathrm{SV}(\widehat{J}_n')}{[\widehat{J}_n' : \widehat{J}]} \geq nK$$

for some constant $K > 0$. Let

$$\eta_n \colon \pi_1(\widehat{J}_n') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

be a representation realizing $\mathrm{SV}(\widehat{J}_n')$.

Granted Theorem 5.1, there exist finite covers $\widetilde{M}_n'$ of M and representations

$$\rho_n \colon \pi_1(\widetilde{M}_n') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

such that

$$\frac{|\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{M}_n'; \rho_n)|}{[\widetilde{M}_n' : M]} = \alpha_0 \cdot \frac{|\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_n'; \eta_n)|}{[\widehat{J}_n' : \widehat{J}]} = \alpha_0 \cdot \frac{\mathrm{SV}(\widehat{J}_n')}{[\widehat{J}_n' : \widehat{J}]},$$

where the positive constant α_0 is determined by M and $J_0 \to \widehat{J}_0$. Therefore,

$$\frac{\mathrm{SV}(\widetilde{M}_n')}{[\widetilde{M}_n' : M]} \geq \frac{|\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{M}_n'; \rho_n)|}{[\widetilde{M}_n' : M]} = \alpha_0 \cdot \frac{\mathrm{SV}(\widehat{J}_n')}{[\widehat{J}_n' : \widehat{J}]} \geq n\alpha_0 K,$$

so the sequence $\{\mathrm{SV}(\widetilde{M}_n')/[\widetilde{M}_n' : M]\}$ is unbounded. This completes the proof of Theorem 1.7.

5.2 CI completions of hyperbolic 3–manifolds

The statement of Theorem 5.1(2) suggests a relation between the asserted representation $\rho\colon \pi_1(\widetilde{M}') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ and the given representation $\eta\colon \pi(\widehat{J}_0) \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$. It would certainly hold if ρ factored through the restriction of η to some finite covers of $\widehat{J}_0$. However, the latter is a much stronger requirement that exceeds our ability. To overcome this difficulty, we examine the machinery of Theorem 3.4 and observe that ρ does factor through a finite cover of certain CW complex associated with $\widehat{J}_0$, which looks like $\widehat{J}_0$ attached with a number of Klein bottles. In the following, we formalize the idea and introduce *CI completions*, where CI is an abbreviation for class inversion.

In general, given an arbitrary group with a collection of conjugacy classes of abelian subgroups, it is possible to embed the group into a larger group which possesses a class inversion with respect to the induced collection. For concreteness, we only consider the special case of CI completions for orientable closed hyperbolic 3–manifolds, with respect to a collection of mutually distinct embedded closed geodesics.

5.2.1 Construction of the CI completion Let V be an orientable closed hyperbolic 3–manifold, and let $\gamma_1, \dots, \gamma_s$ be a collection of mutually distinct embedded closed geodesics of V.

The *CI completion* of V with respect to $\gamma_1, \dots, \gamma_s$ is a pair

$$(W, \sigma_W),$$

where W is a specific CW space equipped with a distinguished embedding $V \to W$ and $\sigma_W\colon W \to W$ is a free involution. The construction is as follows.

Take the product space $V \times \mathbb{Z}$, where $\mathbb{Z}$ is endowed with the discrete topology, and for each γ_i, take a cylinder L_i parametrized as $S^1 \times \mathbb{R}$, where S^1 is identified with the unit circle of the complex plane $\mathbb{C}$. We regard each closed geodesic γ_i as a map $S^1 \to V$. Identify the circles $S^1 \times \mathbb{Z}$ of L_i with closed geodesics of $V \times \mathbb{Z}$ by taking any point $(z, n) \in S^1 \times \mathbb{Z}$ to either $(\gamma_i(z), n)$ or $(\gamma_i(\overline{z}), n)$, depending on the parity of n. We agree to use $\gamma_i(z)$ for even n and $\gamma_i(\overline{z})$ for odd n. The resulting space

$\widetilde{W}_{\mathbb{Z}}$ is equipped with a covering transformation $\sigma\colon \widetilde{W}_{\mathbb{Z}} \to \widetilde{W}_{\mathbb{Z}}$, which takes any point $(x,n) \in V \times \mathbb{Z}$ to $(x,n+1)$ and any point $(z,t) \in L_i$ to $(\overline{z},t+1)$. The quotient of $\widetilde{W}_{\mathbb{Z}}$ by the action of $\langle\sigma^2\rangle$ is a space W with a covering transformation σ_W induced by σ.

One may visualize the further quotient space $W/\langle\sigma_W\rangle$ as a 3–manifold V with Klein bottles hanging on the closed geodesics γ_i, one on each. Then W is a double cover of that space into which V lifts, and on which the deck transformation σ_W acts. As a CW space with a free involution, the isomorphism type of (W,σ_W) is independent of the auxiliary parametrizations in the construction, and the isomorphism may further be required to fix the distinguished inclusion of V.

5.2.2 Properties of CI completions We study the relation of CI completions with class inversions and their behavior under finite covers.

Proposition 5.2 *Let V be an orientable closed hyperbolic 3–manifold, and let $\gamma_1,\dots,\gamma_s$ be a collection of mutually distinct embedded closed geodesics of V. Denote by (W,σ_W) the CI completion of V with respect to $\gamma_1,\dots,\gamma_s$.*

(1) *The outer automorphism of $\pi_1(W)$ induced by σ_W is a class inversion of $\pi_1(W)$ with respect to the collection of conjugacy classes of the maximal cyclic subgroups $\pi_1(\gamma_1),\dots,\pi_1(\gamma_s)$ of $\pi_1(W)$ corresponding to the canonically included free loops.*

(2) *Suppose that $\mathscr{G}$ is a group which possesses a class inversion $[\nu] \in \mathrm{Out}(\mathscr{G})$ with respect to the conjugacy classes of all the cyclic subgroups. Then for any homomorphism $\eta\colon \pi_1(V) \to \mathscr{G}$ then there exists an extension of η to $\pi_1(W)$,*

$$\eta\colon \pi_1(W) \to \mathscr{G}.$$

Moreover, for any representative automorphisms $\sigma_{W\sharp}$ and ν of the outer automorphisms $[\sigma_W]$ and $[\nu]$, respectively, the image $\eta\sigma_{W\sharp}(\pi_1(V))$ is conjugate to $\nu\eta(\pi_1(V))$ in $\mathscr{G}$.

(3) *Suppose that $\kappa\colon V' \to V$ is a covering map of finite degree. Denote by $(W',\sigma_{W'})$ the CI completion of V' with respect to all the elevations in V' of $\gamma_1,\dots,\gamma_s$. Then there exists an extension of κ,*

$$\kappa\colon W' \to W,$$

which is a covering map equivariant under the action of $\sigma_{W'}$ and σ_W. In particular, the covering degree is preserved under the extension.

Proof Recall that W is topologically the union of V, $\sigma_W(V)$. and annuli A_i and $\sigma_W(A_i)$. Each annulus A_i has its boundary attached to $V \sqcup \sigma_W(V)$ in such a way that $\gamma_i \subset V$ can be freely homotoped to the orientation-reversal of $\sigma_W(\gamma_i) \subset \sigma_W(V)$ through A_i, and the annuli $\sigma_W(A_i)$ make the homotopy as well.

Statement (1) is now obvious from the above description.

Statement (2) can also be seen topologically. To this end, let X be a CW model for the Eilenberg–Mac Lane CW space $K(\mathscr{G}, 1)$. Uniquely, up to free homotopy, the outer automorphism $[\nu]$ can be realized by a map $R\colon X \to X$, and the homomorphism η can be realized as a map $f\colon V \to X$. With respect to the inclusion $V \to W$, we define a map $F\colon W \to X$, which extends f, as follows. First define the restriction of F to V and $\sigma_W(V)$ to be f and Rf, respectively. Since ν is a class inversion, each $f\gamma_i$ is freely homotopic to the orientation-reversal of $Rf\gamma_i$, as a map $S^1 \to X$, so the homotopy defines maps $F|\colon A_i \to X$ and $F|\colon \sigma_W(A_i) \to X$. The resulting map $F\colon W \to X$ extends $f\colon V \to X$, so on the level of fundamental groups it gives rise to the claimed extension of $\eta\colon \pi_1(V) \to \mathscr{G}$ over $\pi_1(W)$.

Statement (3) follows from a construction on further quotient spaces. Observe that the quotient space $W/\langle\sigma_W\rangle$, rewritten as $\overline{W}$, is topologically the union of V and Klein bottles B_i, where the B_i are projected from A_i. Then any finite covering map $V' \to V$ gives rise to a covering map of the same degree $\overline{W}' \to \overline{W}$. The covering of Klein bottles are induced by the coverings of $\gamma_i \subset \overline{W}$ by their elevations. In fact, the covering $\overline{W}' \to \overline{W}$ is unique up to homotopy. The covering $\overline{W}' \to \overline{W}$ induces two equivariant covering maps $W' \to W$, differing by deck transformation. The one that respects the distinguished inclusions is as claimed. □

5.3 Virtual representations through CI completions

With our gadgets of CI completions, we invoke Theorem 3.4 to derive the asserted virtual representations of Theorem 5.1.

5.3.1 Construction for the basic level Let M be an orientable closed mixed 3–manifold and J_0 be a distinguished hyperbolic JSJ piece of M. Suppose that $\widehat{J}_0$ is a closed hyperbolic Dehn filling of J_0 by sufficiently long boundary slopes, which are denoted by $\gamma_1, \dots, \gamma_s$. Let

$$(W, \sigma_W)$$

be the CI completion of $\widehat{J}_0$ with respect to $\gamma_1, \dots, \gamma_s$, (see Section 5.2.1). Since $\pi_1(W)$ is class invertible with respect to the conjugacy classes of subgroups $\pi_1(\gamma_i)$ (Proposition 5.2(1)), Theorem 3.4 can be applied with the target group $\pi_1(W)$ and the initial homomorphism

$$\pi_1(J_0) \to \pi_1(W)$$

induced by the composition of the Dehn filling inclusion $J_0 \subset \widehat{J}_0$ and the canonical inclusion $\widehat{J}_0 \subset W$. The output is a finite cover $\widetilde{M}$ of M together with a homomorphism

$$\phi\colon \pi_1(\widetilde{M}) \to \pi_1(W),$$

with described restrictions to its JSJ pieces. Since the CI completion W is an Eilenberg–Mac Lane space $K(\pi_1(W), 1)$, it is convenient to realize ϕ as a map

$$f\colon \widetilde{M} \to W,$$

which is unique up to homotopy.

Suppose for the moment that we are provided with a representation

$$\eta_0\colon \pi_1(\widehat{J}_0) \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}),$$

rather than a virtual representation. By Proposition 5.2(2) and Lemma 3.6, there is an extension over $\pi_1(W)$ (which is still denoted by η_0, regarding the original one as restriction), so that the composition

$$\widetilde{\rho}\colon \pi_1(\widetilde{M}) \xrightarrow{\phi} \pi_1(W) \xrightarrow{\eta_0} \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

gives rise to a virtual extension of the representation

$$\rho_0\colon \pi_1(J_0) \to \pi_1(\widehat{J}_0) \xrightarrow{\eta_0|} \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}).$$

At this basic level, the virtual extension is nothing but a finer version of Theorem 3.4 for the special case of Seifert representations of mixed 3–manifolds. It exhibits a factorization of $\widetilde{\rho}$ through the CI completion $\pi_1(W)$. However, Proposition 5.2(3) allows us to promote the above construction to deal with virtual representations of $\pi_1(\widetilde{J}_0)$.

5.3.2 Construction of $(\widetilde{M}', \rho)$ Now suppose as in Theorem 5.1 that $\widehat{J}_0'$ is a finite cover of $\widehat{J}_0$, and

$$\eta\colon \pi_1(\widehat{J}_0') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

is a Seifert representation of $\pi_1(\widehat{J}_0')$. Denote by

$$(W', \sigma_{W'})$$

the CI completion of $\widehat{J}_0'$ with respect to all the elevations of $\gamma_1, \dots, \gamma_s$. By Proposition 5.2(3) there exists a finite covering map

$$\kappa\colon W' \to W$$

which respects the free involutions and the distinguished inclusions. In particular, κ extends the covering $\widehat{J}_0' \to \widehat{J}_0$ preserving the degree.

Remember that we have obtained a finite cover $\widetilde{M}$ and a map $f\colon \widetilde{M} \to W$ for the basic level. Take any elevation of f with respect to κ, denoted by

$$f'\colon \widetilde{M}' \to W'.$$

This means that the following diagram is commutative up to homotopy:

$$\begin{tikzcd} \widetilde{M}' \arrow[r, "f'"] \arrow[d] & W' \arrow[d, "\kappa"] \\ \widetilde{M} \arrow[r, "f"] & W \end{tikzcd}$$

and $\widetilde{M}' \to \widetilde{M}$ is the covering of $\widetilde{M}$ which is minimal in the sense that it admits no intermediate covering with this property. (More concretely, one may replace W with the mapping cylinder $Y_f \simeq W$, and turn the map f into an inclusion $\widetilde{M} \to Y_f$, then any elevation $\widetilde{M}' \to Y_f'$ of $\widetilde{M}$ in the corresponding finite cover $Y_f' \simeq W'$ gives rise to some $f'\colon \widetilde{M}' \to Y_f' \to W'$ up to homotopy.) Since W' is a finite cover of W, there are only finitely many such elevations $(\widetilde{M}', f')$ up to isomorphism between covering spaces and homotopy. Moreover, the covering degree $[\widetilde{M}' : \widetilde{M}]$ is bounded by $[W' : W]$. Denote by

$$\phi'\colon \pi_1(\widetilde{M}') \to \pi_1(W')$$

the homomorphism (up to conjugation) induced by f'.

Provided with η and ϕ' above, we extend η to be

$$\eta\colon \pi_1(W') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

by Proposition 5.2(1) and (3) and Lemma 3.6. The finite cover

$$\widetilde{M}' \to M$$

and the representation

$$\rho\colon \pi_1(\widetilde{M}') \xrightarrow{\phi'} \pi_1(W') \xrightarrow{\eta} \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$$

are the claimed objects in the conclusion of Theorem 5.1.

Homomorphisms which have been presented can be summarized in the following commutative diagram:

$$\begin{tikzcd} & \pi_1(\hat{J}_0') \arrow[r, "\eta|"] \arrow[d, "\mathrm{incl}_\sharp"] & \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}) \arrow[d, "\mathrm{Id}"] \\ \pi_1(\widetilde{M}') \arrow[r, "\phi'"] \arrow[d, "\mathrm{cov}_\sharp"] & \pi_1(W') \arrow[r, "\eta"] \arrow[d, "\kappa_\sharp"] & \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}) \\ \pi_1(\widetilde{M}) \arrow[r, "\phi"] & \pi_1(W) & \end{tikzcd}$$

The homomorphisms ϕ and ϕ' are realized by maps f and f', respectively. The representation ρ that we have constructed is the composition along the middle row.

We are going to verify Theorem 5.1(2) in the next three subsections.

5.3.3 Restriction to JSJ pieces For any elevation $\widetilde{J}' \subset \widetilde{M}'$ of a JSJ piece $J \subset M$, $\widetilde{J}'$ covers a JSJ piece $\widetilde{J}$ of $\widetilde{M}$. Since we have constructed ϕ using Theorem 3.4, either the restriction of ϕ to $\pi_1(\widetilde{J})$ has cyclic image, or J is the distinguished hyperbolic piece J_0 and the restriction of ϕ to $\pi_1(\widetilde{J})$ is one of the following compositions up to conjugation of $\pi_1(W)$:

$$\pi_1(\widetilde{J}) \to \pi_1(\widehat{J}_0) \to \pi_1(W)$$

or

$$\pi_1(\widetilde{J}) \to \pi_1(\widehat{J}_0) \to \pi_1(W) \xrightarrow{\sigma_W} \pi_1(W).$$

In the cyclic case, the restriction of ϕ' to $\pi_1(\widetilde{J}')$ must also have cyclic image as $\kappa_\sharp$ is injective. Then the restriction of ρ to $\pi_1(\widetilde{J}')$ has cyclic image as well. In the other case, the first homomorphism of either composition factors through $\pi_1(J_0)$ via the Dehn filling, so possibly after homotopy of f, we may assume that $\widetilde{J}$ covers either J_0 or $\sigma_W(J_0)$ under the map f. As f' is an elevation of f with respect to κ, the elevation $\widetilde{J}'$ of $\widetilde{J}$ covers either the unique elevation J_0' of J_0 or the unique elevation $\sigma_{W'}(J_0')$ of $\sigma_W(J_0)$ in W'. Note that η is equivariant up to conjugacy with respect to the class inversions $\sigma_{W'}$ and ν (Proposition 5.2 and Lemma 3.6). It follows that by taking

$$\beta\colon \widetilde{J}' \to J_0' \to \widehat{J}_0',$$

the composition of the covering and the inclusion, the restriction of ρ to $\pi_1(\widetilde{J}')$ is either $\beta^*(\eta)$ or $\beta^*(\nu\eta)$. This verifies Theorem 5.1(1).

5.3.4 Count of degree By the consideration about the restriction of ρ to JSJ pieces of $\widetilde{M}'$ above, we have seen that a JSJ piece $\widetilde{J}'$ gives rise to the β–pullback-type restriction of ρ if and only if $\widetilde{J}'$ covers a JSJ piece $\widetilde{J}$ of $\widetilde{M}$ such that $\phi(\pi_1(\widetilde{J}))$ is noncyclic. The union of all such $\widetilde{J}$ in $\widetilde{M}$ form a (disconnected) finite cover $\widetilde{\mathcal{J}}$ of the distinguished piece $J_0 \subset M$, and the union of all β–pullback-types $\widetilde{J}'$ in $\widetilde{M}'$ is nothing but the preimage $\widetilde{\mathcal{J}}'$ of $\widetilde{\mathcal{J}}$ in $\widetilde{M}'$. Therefore, suppose α_0 is the ratio between the total degree of β–pullback-type JSJ pieces of $\widetilde{M}'$ over J_0 and the degree of $\widetilde{M}'$,

$$[\widetilde{\mathcal{J}}' : J_0] = \alpha_0 \cdot [\widetilde{M}' : M],$$

then we observe

$$\alpha_0 = \frac{[\widetilde{\mathcal{J}}' : J_0]}{[\widetilde{M}' : M]} = \frac{[\widetilde{\mathcal{J}}' : \widetilde{\mathcal{J}}] \cdot [\widetilde{\mathcal{J}} : J_0]}{[\widetilde{M}' : \widetilde{M}] \cdot [\widetilde{M} : M]} = \frac{[\widetilde{\mathcal{J}} : J_0]}{[\widetilde{M} : M]}.$$

Note that α_0 depends only on M and $J_0 \to \widehat{J}_0$, since $\widetilde{M}$ and ϕ are constructed according to them, and α_0 is positive because $\widetilde{\mathcal{J}}$ is nonempty by Theorem 3.4.

5.3.5 Count of volume In a very similar situation as in the proof of Theorem 1.5, to compute the volume of the representation

$$\rho\colon \pi_1(\widetilde{M}') \to \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}),$$

it suffices to understand the contribution to the representation volume of ρ from the β–pullback-type JSJ pieces $\widetilde{J}'$ of $\widetilde{M}'$. Note that the map

$$\beta\colon \widetilde{J}' \xrightarrow{\text{cov}} J_0' \xrightarrow{\text{fill}} \widehat{J}_0'$$

factors through a unique hyperbolic Dehn filling $\widetilde{K}'$ of $\widetilde{J}'$, which covers $\widehat{J}_0'$ branching over elevations of the core curves γ_i via a map $\widehat{\beta}$:

$$\beta\colon \widetilde{J}' \xrightarrow{\text{fill}} \widetilde{K}' \xrightarrow{\widehat{\beta}} \widehat{J}_0'$$

The restriction of ρ to $\pi_1(\widetilde{J}')$ thus factors as

$$\pi_1(\widetilde{J}') \xrightarrow{\text{fill}_\sharp} \pi_1(\widetilde{K}') \xrightarrow{\widehat{\rho}} \mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R}),$$

where $\widehat{\rho}$ equals the $\widehat{\beta}$–pullback of η or $\nu\eta$. Note that the class inversion ν of $\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})$ is realized by the conjugation of an orientation-preserving isomorphism of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, so

$$\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \eta) = \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \nu\eta).$$

It follows from the additivity principle (Theorem 3.1) that the contribution to the representation volume of ρ from the piece $\widetilde{J}'$ equals $\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{K}'; \widehat{\rho})$ and

$$\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{K}'; \widehat{\rho}) = |\deg \widehat{\beta}| \cdot \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \eta) = \frac{[\widetilde{J}' : J_0]}{[\widehat{J}_0' : \widehat{J}_0]} \cdot \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \eta).$$

On the other hand, the contribution from any cyclic-type JSJ piece $\widetilde{J}'$ of $\widetilde{M}'$ is always zero by Lemma 3.2. Take the summation of the contribution from all JSJ pieces, using the formula of α_0 in the degree count:

$$\begin{aligned}
\mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{M}; \rho) &= \sum_{\widetilde{J}' \in \widetilde{\mathcal{J}}'} \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{K}'; \widehat{\rho}) \\
&= \sum_{\widetilde{J}' \in \widetilde{\mathcal{J}}'} \frac{[\widetilde{J}' : J_0]}{[\widehat{J}_0' : \widehat{J}_0]} \cdot \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \eta) \\
&= \frac{[\widetilde{\mathcal{J}}' : J_0]}{[\widehat{J}_0' : \widehat{J}_0]} \cdot \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \eta) \\
&= \alpha_0 \cdot \frac{[\widetilde{M}' : M]}{[\widehat{J}_0' : \widehat{J}_0]} \cdot \mathrm{Vol}_{\mathrm{Iso}_e \widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}_0'; \eta),
\end{aligned}$$

or equivalently,

$$\frac{\mathrm{Vol}_{\mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widetilde{M}';\rho)}{[\widetilde{M}':M]}=\alpha_0\cdot\frac{\mathrm{Vol}_{\mathrm{Iso}_e\widetilde{\mathrm{SL}}_2(\mathbb{R})}(\widehat{J}'_0;\eta)}{[\widehat{J}'_0:\widehat{J}_0]}.$$

This completes the proof of Theorem 5.1(2), and therefore the proof of Theorem 5.1.

6 On covering invariants

Although the covering property does not hold for the representation volumes [5, Corollary 1.8], we can stabilize them to obtain covering invariants in the following way.

Definition 6.1 For any closed orientable 3–manifold N, define the *covering Seifert volume* of N to be

$$\mathrm{CSV}(N)=\varprojlim_{\widetilde{N}}\frac{\mathrm{SV}(\widetilde{N})}{[\widetilde{N}:N]},$$

valued in $[0,+\infty]$, where $\widetilde{N}$ runs over all the finite covers of N. Note that the limit exists because $\mathrm{SV}(\widetilde{N})/[\widetilde{N}:N]$ is nondecreasing under passage to finite covers. Similarly one can define the *covering hyperbolic volume* $\mathrm{CHV}(M)$.

Proposition 6.2 *If* CSV*, or* CHV*, is valued on* $[0,+\infty)$ *for a class* $\mathcal{C}$ *of closed orientable* 3*–manifolds, then it satisfies both domination property and covering property for* $\mathcal{C}$*.*

Proof We verify the statement for CSV; the argument for CHV is completely similar.

To verify the domination property, let $f\colon M\to N$ be any map of nonzero degree between $M,N\in\mathcal{C}$. By definition, for any $\epsilon>0$, there is a finite cover $\widetilde{N}$ of N such that

$$\frac{\mathrm{SV}(\widetilde{N})}{[\widetilde{N}:N]}>\mathrm{CSV}(N)-\epsilon.$$

We have the commutative diagram

$$\begin{array}{ccc}\widetilde{M} & \xrightarrow{\tilde{f}} & \widetilde{N}\\ \downarrow & & \downarrow\\ M & \xrightarrow{f} & N\end{array}$$

for the pullback cover $\widetilde{M}$ of M via f, which has degree at most $[\widetilde{N}:N]$. Then we have $[\widetilde{M}:M]\cdot|\deg f|=[\widetilde{N}:N]\cdot|\deg\tilde{f}|$, and $|\deg f|\geq|\deg\tilde{f}|$, and $\mathrm{SV}(\widetilde{M})\geq$

$|\deg \tilde{f}| \cdot \mathrm{SV}(\tilde{N})$. It follows that

$$\frac{\mathrm{SV}(\tilde{M})}{[\tilde{M}:M]} = \frac{\mathrm{SV}(\tilde{M}) \cdot |\deg f|}{[\tilde{N}:N] \cdot |\deg \tilde{f}|} \geq \frac{|\deg f| \cdot \mathrm{SV}(\tilde{N})}{[\tilde{N}:N]} \geq |\deg f| \cdot (\mathrm{CSV}(N) - \epsilon).$$

Taking the limit over all $\tilde{M}$ and $\epsilon \to 0+$, we have

$$\mathrm{CSV}(M) \geq |\deg f| \cdot \mathrm{CSV}(N).$$

To verify the covering property, suppose that $f\colon M \to N$ is a covering map, so $\deg f$ equals $[M:N]$. Then any finite cover $\tilde{M}$ of M is also a finite cover of N. By definition we have

$$\frac{\mathrm{SV}(\tilde{M})}{[\tilde{M}:M]} = [M:N] \cdot \frac{\mathrm{SV}(\tilde{M})}{[\tilde{M}:N]} \leq [M:N] \cdot \mathrm{CSV}(N) = |\deg f| \cdot \mathrm{CSV}(N).$$

Taking the limit over all $\tilde{M}$, we have $\mathrm{CSV}(M) \leq |\deg f| \cdot \mathrm{CSV}(N)$. So indeed we have

$$\mathrm{CSV}(M) = |\deg f| \cdot \mathrm{CSV}(N),$$

where the other direction follows from the domination property. □

We post some further problems, updating those of [5, Section 8].

Problem 6.3 Does $\mathrm{CSV}(M)$ exist in $(0, +\infty)$ for every closed orientable nongeometric graph manifold M?

A positive answer would provide a nowhere-vanishing invariant with the covering property in the class of closed orientable nongeometric graph manifolds. Finding such an invariant was suggested by Thurston [15, Problem 3.16]. See [18; 19; 30] for some attempts motivated by showing the uniqueness of covering degree between graph manifolds. The uniqueness is confirmed by [32] using combinatorial methods and matrix theory.

Problem 6.4 Determine the possible growth types and asymptotics of the virtual Seifert volume for closed orientable 3–manifolds with positive simplicial volume.

We speak of the growth with respect to towers of finite covers, as the covering degree increases. Theorem 1.7 shows that there are towers with superlinear growth. The estimates of [3] imply that the growth must be at most exponential.

Problem 6.5 Is $\mathrm{CHV}(M)$ equal to $v_3\|M\|$ for every closed orientable 3–manifold M?

This quantity is at most $v_3\|M\|$ (see the remark following Theorem 1.7) and we suspect that the equality might be achieved.

References

[1] **I Agol**, *The virtual Haken conjecture*, Doc. Math. 18 (2013) 1045–1087 MR

[2] **G Besson**, **G Courtois**, **S Gallot**, *Inégalités de Milnor–Wood géométriques*, Comment. Math. Helv. 82 (2007) 753–803 MR

[3] **R Brooks**, **W Goldman**, *The Godbillon–Vey invariant of a transversely homogeneous foliation*, Trans. Amer. Math. Soc. 286 (1984) 651–664 MR

[4] **R Brooks**, **W Goldman**, *Volumes in Seifert space*, Duke Math. J. 51 (1984) 529–545 MR

[5] **P Derbez**, **Y Liu**, **S Wang**, *Chern–Simons theory, surface separability, and volumes of 3–manifolds*, J. Topol. 8 (2015) 933–974 MR

[6] **P Derbez**, **H B Sun**, **S C Wang**, *Finiteness of mapping degree sets for 3–manifolds*, Acta Math. Sin. (Engl. Ser.) 27 (2011) 807–812 MR

[7] **P Derbez**, **S Wang**, *Finiteness of mapping degrees and $\mathrm{PSL}(2,\mathbb{R})$–volume on graph manifolds*, Algebr. Geom. Topol. 9 (2009) 1727–1749 MR

[8] **P Derbez**, **S Wang**, *Graph manifolds have virtually positive Seifert volume*, J. Lond. Math. Soc. 86 (2012) 17–35 MR

[9] **A Gaifullin**, *Universal realisators for homology classes*, Geom. Topol. 17 (2013) 1745–1772 MR

[10] **M Gromov**, *Volume and bounded cohomology*, Inst. Hautes Études Sci. Publ. Math. 56 (1982) 5–99 MR

[11] **F Haglund**, **D T Wise**, *Special cube complexes*, Geom. Funct. Anal. 17 (2008) 1551–1620 MR

[12] **W H Jaco**, **P B Shalen**, *Seifert fibered spaces in 3–manifolds*, Mem. Amer. Math. Soc. 220, Amer. Math. Soc., Providence, RI (1979) MR

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

[14] **J Kahn**, **V Markovic**, *Immersing almost geodesic surfaces in a closed hyperbolic three manifold*, Ann. of Math. 175 (2012) 1127–1190 MR

[15] **R Kirby**, *Problems in low dimensional manifold theory*, from "Algebraic and geometric topology, II" (R J Milgram, editor), Proc. Sympos. Pure Math. XXXII, Amer. Math. Soc., Providence, RI (1978) 273–312 MR

[16] **Y Liu**, *A characterization of virtually embedded subsurfaces in 3–manifolds*, Trans. Amer. Math. Soc. 369 (2017) 1237–1264 MR

[17] **Y Liu**, **V Markovic**, *Homology of curves and surfaces in closed hyperbolic 3–manifolds*, Duke Math. J. 164 (2015) 2723–2808 MR

[18] **J Luecke**, **Y-Q Wu**, *Relative Euler number and finite covers of graph manifolds*, from "Geometric topology" (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc., Providence, RI (1997) 80–103 MR

[19] **W D Neumann**, *Commensurability and virtual fibration for graph manifolds*, Topology 36 (1997) 355–378 MR

[20] **P Przytycki**, **D T Wise**, *Mixed 3–manifolds are virtually special*, preprint (2012) arXiv

[21] **P Przytycki**, **D T Wise**, *Graph manifolds with boundary are virtually special*, J. Topol. 7 (2014) 419–435 MR

[22] **P Przytycki**, **D T Wise**, *Separability of embedded surfaces in 3–manifolds*, Compos. Math. 150 (2014) 1623–1630 MR

[23] **A Reznikov**, *Rationality of secondary classes*, J. Differential Geom. 43 (1996) 674–692 MR

[24] **J H Rubinstein**, **S Wang**, *π_1–injective surfaces in graph manifolds*, Comment. Math. Helv. 73 (1998) 499–515 MR

[25] **P Scott**, *The geometries of 3–manifolds*, Bull. London Math. Soc. 15 (1983) 401–487 MR

[26] **T Soma**, *The Gromov invariant of links*, Invent. Math. 64 (1981) 445–454 MR

[27] **H Sun**, *Virtual domination of 3–manifolds*, Geom. Topol. 19 (2015) 2277–2328 MR

[28] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at `http://msri.org/publications/books/gt3m`

[29] **W P Thurston**, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. 6 (1982) 357–381 MR

[30] **S C Wang**, **Y Q Wu**, *Covering invariants and co-Hopficity of 3–manifold groups*, Proc. London Math. Soc. 68 (1994) 203–224 MR

[31] **D T Wise**, *From riches to raags: 3–manifolds, right-angled Artin groups, and cubical geometry*, CBMS Regional Conference Series in Mathematics 117, Amer. Math. Soc., Providence, RI (2012) MR

[32] **F Yu**, **S Wang**, *Covering degrees are determined by graph manifolds involved*, Comment. Math. Helv. 74 (1999) 238–247 MR

PD: *Aix Marseille Université, CNRS UMR 7373, Centrale Marseille, I2M*
Marseille, France

YL: *Beijing International Center for Mathematical Research, Peking University*
Beijing, China

HS: *Department of Mathematics, University of California*
Berkeley, CA, United States

Current address for HS: *Department of Mathematics, Rutgers University*
Piscataway, NJ, United States

SW: *School of Mathematical Sciences, Peking University*
Beijing, China

`pderbez@gmail.com`, `liuyi@math.pku.edu.cn`, `hongbins@math.berkeley.edu`, `wangsc@math.pku.edu.cn`

Proposed: Cameron Gordon — Received: 6 July 2016
Seconded: Gang Tian, Walter Neumann — Accepted: 23 December 2016

Guidelines for Authors

Submitting a paper to Geometry & Topology

Papers must be submitted using the upload page at the GT 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 Geometry & 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 GT is available on the GT 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.

GEOMETRY & TOPOLOGY

Volume 21 Issue 5 (pages 2557–3190) 2017