# Geometry & Topology

Volume 27 (2023)

Issue 4 (pages 1273–1655)

msp

# GEOMETRY & TOPOLOGY
msp.org/gt



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

The subscription price for 2023 is US $740/year for the electronic version, and $1030/year (+$70, 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) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

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

PUBLISHED BY

**mathematical sciences publishers**

**nonprofit scientific publishing**

http://msp.org/

*Geometry & Topology* **27**:4 (2023) 1273–1390
DOI: 10.2140/gt.2023.27.1273
Published: 15 June 2023

# Floer theory and reduced cohomology on open manifolds

YOEL GROMAN

We construct Hamiltonian Floer complexes associated to continuous, and even lower semicontinuous, time-dependent exhaustion functions on geometrically bounded symplectic manifolds. We further construct functorial continuation maps associated to monotone homotopies between them, and operations which give rise to a product and unit. The work rests on novel techniques for energy confinement of Floer solutions as well as on methods of non-Archimedean analysis. The definition for general Hamiltonians utilizes the notion of reduced cohomology familiar from Riemannian geometry, and the continuity properties of Floer cohomology. This gives rise, in particular, to local Floer theory. We discuss various functorial properties as well as some applications to existence of periodic orbits and to displaceability.

53D40; 32P05

| | | |
|---|---|---|
| 1. | Introduction | 1274 |
| 2. | Symplectic cohomology | 1281 |
| 3. | Overview | 1289 |
| 4. | I–bounded almost complex structures | 1296 |
| 5. | Floer solutions and the Gromov metric | 1305 |
| 6. | Loopwise dissipativity | 1318 |
| 7. | Proof of Theorem 1.1 | 1336 |
| 8. | Hamiltonian Floer cohomology by approximation | 1348 |
| 9. | The product structure | 1361 |
| 10. | Computations and applications | 1375 |
| | References | 1386 |

# 1 Introduction

Floer theory is a machine that associates algebraic structures to objects of symplectic geometry. Over the years, it has come to play central role in all aspects of the field, from mirror symmetry to quantitative symplectic topology. In this paper we extend the range of applicability of Floer theory, focusing on Hamiltonian Floer theory, to open symplectic manifolds which are geometrically bounded.

There are a number of reasons why one would be interested in studying Floer theory on open manifolds. First and foremost, many symplectic manifolds which arise naturally are open. Among these we count the cotangent bundle, the magnetic cotangent bundle, affine varieties, coadjoint orbits of noncompact groups, Hitchin moduli spaces, and many more. More specifically related to Floer theory, there are phenomena which only become apparent on open manifolds. For example, on a closed manifold the Hamiltonian Floer cohomology reduces as an abelian group to the singular cohomology, and is thus often too coarse to see much of the symplectic topology. On open manifolds, new invariants, such as symplectic cohomology, make their appearance and encode purely symplectic phenomena; see for instance Cieliebak, Floer and Hofer [15], Oancea [41], Seidel [56] and Viterbo [63].

There is a vast literature studying these invariants and their structural properties. For example, symplectic cohomology of a Liouville domain has been shown to play a key role in homological mirror symmetry by encoding the Hochschild homology of the Fukaya category; see for instance Abouzaid [1], Ganatra [24; 25] and Seidel [55]. In another related direction, there are numerous results relating Floer theory of a symplectic manifold with the Floer theory of embedded local models; see eg Cieliebak and Oancea [18], Ganatra, Pardon and Shende [26], Seidel [55] and Varolgunes [61]. But the existing literature focuses mostly on examples which are convex at infinity, a condition which does not cover, for instance, most of the examples mentioned in the previous paragraph. As another example, in the study of the Fukaya category by the method of localization away from a divisor as in [55], one does not necessarily wish to restrict attention to ample divisors; see eg Auroux [8; 7], Daemi and Fukaya [19] and Groman [29]. Studying Floer theory in more general settings would contribute to our understanding of mirror symmetry, the geometric Langlands program, and many branches of symplectic topology.

The class of geometrically bounded manifolds contains those that are convex at infinity, but is much larger. Geometric boundedness has appeared as a relevant condition already in Gromov's seminal paper [31], and in numerous works since. A couple of early ones are Audin, Lalonde and Polterovich [6] and Sikorav [57]. Geometrically bounded symplectic manifolds are the most general setting in which holomorphic curve theory is known to work without resorting to the methods of symplectic field theory.

The novelty in the present paper is twofold. One is showing how to apply geometric boundedness of the underlying manifold to carry out Floer-theoretic constructions beyond $J$–holomorphic curves. Such constructions are central, for instance, to the notion of wrapped Floer theory. The other is showing that the invariants constructed by choosing a geometrically bounded metric at infinity are independent of the choice. The results here provide a unified and flexible framework which incorporates the various constructions in the literature (see eg Cieliebak, Floer and Hofer [15], Oancea [42], Ritter [47] and Viterbo [63]), works in full generality, and has transparent symplectic invariance properties.

It should be emphasized that while we do not mention the Fukaya category elsewhere in this paper, the difficulties posed by noncompactness are virtually the same for the Hamiltonian version of Floer theory as for its Lagrangian intersection version. Thus, this paper sets the stage for the study of the (wrapped) Fukaya category on open manifolds such as those mentioned above, insofar as one can overcome the usual difficulties already present in the closed case.

We shall assume familiarity with the basic machinery of Hamiltonian Floer theory and symplectic cohomology such as can be acquired from the first three lectures in Salamon [52] together with Oancea [41]. For the discussion of the product structure we shall assume also some familiarity with treatments such as Abouzaid [4] or Ritter [48]. The latter is not necessary for most of the novel ideas in this paper.

## 1.1 The main result

A symplectic manifold $(M,\omega)$ is said to be *geometrically bounded* if there is an $\omega$–compatible almost complex structure $J$, a constant $C>1$, and a complete Riemannian metric $g$ with sectional curvature bounded from above and injectivity radius bounded away from 0, such that

$$\frac{1}{C}g(v,v)\le\omega(v,Jv)\le Cg(v,v)$$

for all tangent vectors $v$. Note that the almost complex structure $J$ is *not* part of the data. Examples include closed symplectic manifolds, cotangent bundles of arbitrary smooth manifolds, manifolds whose end is modeled after the convex half of the symplectization of a compact contact manifold (as in Sikorav [57]), twisted cotangent bundles (as in Cieliebak, Ginzburg and Kerman [17]), and there are many more. The class of geometrically bounded symplectic manifolds is closed under products and coverings.

It should be emphasized at the outset that an open symplectic manifold of finite volume, such as the unit ball in $\mathbb{C}^n$, cannot be endowed with a metric that is at once complete and satisfies the above bounds on sectional curvature and radius of injectivity and thus is *not* geometrically bounded. Floer theory for open finite-volume symplectic manifolds, such as Liouville domains, will be discussed, in the context of local Floer theory, when they are embedded in a geometrically bounded symplectic manifold.

Recall that $(M,\omega)$ is said to be *semimonotone* if there exists a constant $\tau \geq 0$ such that for any $A \in \pi_2(M)$ we have $\omega(A) = \tau c_1(A)$, where $c_1$ is the first Chern class. $(M,\omega)$ is said to be *Calabi–Yau* if $c_1(A) = 0$ for every $A \in \pi_2(M)$. *Henceforth*, $(M,\omega)$ *is a geometrically bounded symplectic manifold which is either semimonotone or Calabi–Yau.* In particular, for any class $A \in \pi_2(M)$ we have

$$c_1(A) < 0 \implies \omega(A) \leq 0.$$

We hasten to emphasize that this assumption is made for definiteness only. The methods introduced herein are orthogonal to the usual questions of transversality and can be adapted to any regularization scheme.

Fix a field $R$ and denote by $\Lambda_R$ the universal Novikov field and by $\Lambda_{R,\omega}$ the Novikov field associated with $\omega$; see Section 7.3. We shall use the notation $\mathbb{K}$ to denote either $\Lambda_R$ or $\Lambda_{R,\omega}$. Note that $\mathbb{K}$ is a graded field. That is, a commutative even graded field in which every nonzero homogeneous element is invertible. *In the entire text*, $\Lambda_R$ *coefficients should be assumed by default whenever the coefficient field is not indicated in the notation.*

Denote by $\mathcal{J}$ the space of smooth $\mathbb{R}/\mathbb{Z}$ parametrized families of almost complex structures which are compatible with $\omega$. Denote by $\mathcal{H}_{\mathrm{sm}}$ the space of smooth functions on $\mathbb{R}/\mathbb{Z} \times M$ which are proper and bounded from below. Consider the category $\mathcal{F}$ of Floer data whose objects are elements $(H,J) \in \mathcal{H}_{\mathrm{sm}} \times \mathcal{J}$ and in which there is a single morphism $(H_1,J_1) \to (H_2,J_2)$ between objects whenever the order relation defined by

$$(H_1,J_1) \leq (H_2,J_2) \quad \iff \quad H_{1,t}(x) \leq H_{2,t}(x) \quad \text{for all } (t,x) \in \mathbb{R}/\mathbb{Z} \times M$$

is satisfied.

The main contribution of this paper is summarized in the following theorem.

**Theorem 1.1** *There exists a full subcategory $\mathcal{F}_{\mathrm{d,reg}} \subset \mathcal{F}$, referred to as the regular dissipative Floer data, for which the Floer cohomology*

$$(H, J) \mapsto HF^*(H, J; \mathbb{K})$$

*is well defined as a functor to $\mathbb{Z}$–graded non-Archimedean (semi)normed $\mathbb{K}$–modules. Namely, there is a functorial norm-preserving continuation map*

$$HF^*(H_1, J_1; \mathbb{K}) \to HF^*(H_2, J_2; \mathbb{K})$$

*whenever $H_1 \le H_2 \in \mathcal{F}_{\mathrm{d,reg}}$. The subcategory $\mathcal{F}_{\mathrm{d,reg}}$ satisfies the following:*

(a) *It is invariant under the action of the symplectomorphism group*

$$\psi \cdot (H, J) \mapsto (H \circ \psi, \psi^* J).$$

(b) *It is final and cofinal in $\mathcal{F}$.*

(c) *It contains all pairs $(H, J)$ for which $J$ is geometrically bounded and $H$ has sufficiently small Lipschitz constant and is (nearly) time independent near infinity.*

(d) *The continuation map $HF^*(H_1, J_1; \mathbb{K}) \to HF^*(H_2, J_2; \mathbb{K})$ is an isomorphism if $H_2 - H_1$ is bounded on $M$.*

Theorem 1.1 relies on the dissipative method introduced herein for controlling compactness of various Floer moduli spaces. This is done by systematically replacing the more conventional reliance on maximum principles by a combination of the monotonicity inequality for $J$–holomorphic curves and a certain quantitative nondegeneracy condition to control the ends. A more detailed discussion of this method is given in Section 3.1. This method should be of independent interest for researchers wishing to apply Floer theory methods in any way to open symplectic manifolds. We emphasize that the Floer data which are typically used in the literature on symplectic cohomology mostly fit into the dissipative framework. For a discussion of the case of Liouville domains, see Example 6.14 below.

The true power of the dissipative method is revealed when considering the functoriality statement in Theorem 1.1, which is one of the main contributions of this paper. To demonstrate this we first remark that a particular consequence of the functoriality is the independence of Floer cohomology of a dissipative $(H, J)$ on the choice of $J$. This is new even for the case $H = 0$, ie for $J$–holomorphic curves. We use this in

Theorem 4.15 below to show that in a number of contexts where invariants on open manifolds are defined using a geometrically bounded almost complex structures $J$, the resulting invariants do not depend on the choice of such a $J$. The difficulty in proving this is that given two geometrically bounded compatible almost complex structures which are not metrically equivalent, it is not likely they can be homotoped to one another through geometrically bounded almost complex structures. Our solution is to introduce the notion of intermittent boundedness, or, i–boundedness, which requires boundedness only on an appropriate infinite sequence of hypersurfaces. We then show that any two such almost complex structures can be homotoped to one another through intermittently bounded almost complex structures. For the rest of the introduction we thus drop $J$ from the notation and consider dissipativity as a property of a Hamiltonian function.

As an illustration of the use of functoriality we indicate an easy proof of the Künneth formula in symplectic cohomology of Liouville domains; cf [42]. This requires comparing the direct limit of Floer cohomologies over a sequence of linear Hamiltonians on the smoothing of a product of Liouville domains to the direct limit over a sequence of linear split Hamiltonians on the product itself, all with slope going to infinity. Since one can squeeze a sequence of linear Hamiltonians between a sequence of split linear ones, the Künneth formula follows from Theorem 1.1 as soon as one establishes that both linear and split linear Hamiltonians are dissipative. The latter is immediately implied by Examples 5.24 and 6.20. The line of argument can be shown to extend to Liouville domains with arbitrary corners.

## 1.2 Reduced Floer cohomology for general Hamiltonians

Our next theorem combines the result of Theorem 1.1 with certain continuity properties of Floer cohomology, to extend the definition of Floer cohomology to more arbitrary Floer data. Namely, we extend a certain version of Floer cohomology as a functor on the category $(\mathcal{H}_{\mathrm{d,reg}}, \leq)$ of regular dissipative Hamiltonians to the category $(\mathcal{H}_{\mathrm{sc}}, \leq)$ of all generalized lower semicontinuous functions $\mathbb{R}/\mathbb{Z} \times M \to \mathbb{R} \cup \infty$ which are proper and bounded from below.

Before proceeding we introduce the concept of *reduced Floer cohomology* $\overline{HF}^*(H)$ of a nondegenerate dissipative Hamiltonian $H$. The ordinary Floer cohomology $HF^*(H)$ is the homology of a chain complex which is complete with respect to a non-Archimedean norm. Thus the group $HF^*(H;\mathbb{K})$ is naturally seminormed. However, in general, the differential needn't have closed image. In such a case $HF^*(H;\mathbb{K})$ contains nontrivial elements of norm 0. The reduced Floer cohomology $\overline{HF}^*(H)$ is the quotient of $HF^*(H)$ by the elements of norm 0. A similar construction is familiar from Riemannian geometry in the context of $L^2$–cohomology; see Cheeger, Goresky and MacPherson [11], Dai [20] and Lück [36]. The precise statement of the following result requires some further preparation. We therefore present first an informal statement. Theorem 3.3 below is a more precise restatement.

**Theorem 1.2** *The reduced Floer cohomology functor $H \mapsto \overline{HF}^*(H;\mathbb{K})$ extends in a natural way from a functor on the category $\mathcal{H}_d$ of dissipative Hamiltonians to a functor on the category $\mathcal{H}_{\mathrm{sc}}$ of all lower semicontinuous exhaustion functions. Moreover, $\overline{HF}^*(H;\mathbb{K})$ arises as the reduced cohomology of a certain non-Archimedean Banach chain complex, which is associated to $H$ up to an appropriate notion of quasi-isomorphism, and which specializes to the Floer chain complex for dissipative Hamiltonians.*

Theorem 1.2 employs the method of Floer theory by approximation. This is explained in more detail in section Section 3.2. Among other things, Theorem 1.2 can be interpreted as saying that, at least if one is concerned with reduced cohomology, one needn't worry about the question of whether a given Floer datum is dissipative or not. In a forthcoming note, joint with U Varolgunes, we show that Theorem 1.2 actually holds for the unreduced version of Floer cohomology. For a more extensive discussion of this, see comment (d) right after the statement of Theorem 3.3.

## 1.3 The product structure

To state the final main theorem we introduce the notion of symplectic cohomology on an arbitrary geometrically bounded symplectic manifold. Let $\mathcal{H} \subset \mathcal{H}_{\mathrm{sc}}$ be a subset consisting of time-independent Hamiltonians such that for any $H_1, H_2 \in \mathcal{H}$ we have that $2\max\{H_1, H_2\} \in \mathcal{H}$. We call $\mathcal{H}$ a *monoidal indexing set*. For each monoidal indexing set $\mathcal{H}$ we define a group

$$SH^*(M;\mathcal{H}) := \varinjlim_{H \in \mathcal{H}} \overline{HF}^*(M).$$

The set of monoidal indexing sets is partially ordered by the relation $\mathcal{H}_1 \preceq \mathcal{H}_2$, which is defined to hold if and only if for any $H \in \mathcal{H}_1$ there is a constant $C$ and an $H_2 \in \mathcal{H}_2$

such that $H_1 \le H_2 + C$. Before proceeding to the statement of the following theorem, we note that there is a decomposition

$$SH^*(M;\mathcal{H}) = \bigoplus_{\alpha\in[S^1,M]} SH^{*,\alpha}(M;\mathcal{H}),$$

where for a free homotopy class $\alpha$, the group $SH^{*,\alpha}(M;\mathcal{H})$ arises from the periodic orbits in the homotopy class $\alpha$. In particular, in the following theorem we refer to the subgroups $SH^{*,0}(M;\mathcal{H})$ which arise from considering only contractible periodic orbits.

**Theorem 1.3** *The groups $SH^{*,0}(M;\mathcal{H})$ have the following properties*:

(a) *For any monoidal indexing set $\mathcal{H}$, there is a product structure*

$$*\colon SH^{*,0}(M;\mathcal{H}) \otimes SH^{*,0}(M;\mathcal{H}) \to SH^{*,0}(M;\mathcal{H}),$$

*which is associative*, *and supercommutative.*

(b) *The small quantum product on $QH^*(M;\mathbb{K}) := H^*(M;\mathbb{K})$ is well defined and for any monoidal indexing set $\mathcal{H}$ there is a natural PSS homomorphism*

$$QH^*(M;\mathbb{K}) \to SH^{*,0}(M;\mathcal{H})$$

*such that the image of $1 \in QH^*(M;\mathbb{K})$ acts as the unit in $SH^{*,0}(M;\mathcal{H})$.*

(c) *Given monoidal indexing sets $\mathcal{H}_1 \preceq \mathcal{H}_2$, the natural continuation map*

$$SH^{*,0}(M;\mathcal{H}_1) \to SH^{*,0}(M;\mathcal{H}_2)$$

*is a unital algebra homomorphism.*

The proof of Theorem 1.3 is carried out in Section 9.4. The restriction to contractible periodic orbits is done for the sake of brevity in the proof. See Remark 9.6 below for an extended discussion. Theorem 1.3 allows the construction of various flavors of symplectic cohomology as a unital algebra. Essentially the same proof can be used to construct operations associated with any family of nodal Riemann surfaces and parametrized Floer data. Moreover, it is possible to carry out a Lagrangian intersection variant of the results of this paper. Thus, the TQFT structure presented for the case of Liouville domains in [48] can be transferred in its entirety to the present setting.

### Organization of the paper

The paper is organized as follows. Section 2 discusses various notions of symplectic cohomology and gives some applications of the main theorems. Section 3.1 provides an overview of the techniques going into the proof of Theorem 1.1 while Section 3.2 is devoted to explaining Theorem 1.2. Sections 4 through 6 are devoted to constructing the dissipative Floer data featuring in Theorem 1.1. The proof of the latter is carried out in Section 7. In Section 8 we prove Theorem 1.2 (restated as Theorem 3.3). In Section 9 we prove Theorem 1.3. In Section 10 we carry out the proofs of the properties and applications mentioned in Section 2.

**Acknowledgements** This work subsumes part of the author's PhD thesis carried out in the Hebrew University of Jerusalem. The author is grateful to the Azrieli foundation for granting him an Azrieli fellowship at that time.

Various stages of the work were carried out at the math departments of Hebrew University of Jerusalem, the ETH of Zurich, and Columbia University. They were supported by the ERC Starting Grant 337560, ISF Grant 1747/13, Swiss National Science Foundation grant 200021_156000, and the Simons Foundation/SFARI (grant #385571, Abouzaid).

The author would like to thank his thesis advisor Jake P Solomon for his guidance throughout his graduate work and for a great number of discussions and suggestions. The author is grateful to Mohammed Abouzaid for suggesting some key ideas and to Paul Biran, Lev Buhovski, Sheel Ganatra, Alex Oancea, Dietmar Salamon, Ran Tessler, Amitai Zernik, Umut Varolgunes and Sara Venkatesh, for useful input and enlightening discussions. In preparing Section 9 of the current version, the author is greatly indebted to M Abouzaid and U Varolgunes for discussions on how to set up Floer-theoretic operations for noncommuting Hamiltonians. The discussion in previous versions was erroneous on this count.

Finally, the author would like to thank the anonymous referee for the thorough and detailed review, and numerous helpful suggestions. These have helped catch numerous gaps and have no doubt greatly improved the exposition.

## 2 Symplectic cohomology

In the following subsections we use Theorems 1.2 and 1.3 to construct symplectic cohomology rings and discuss some of their functorial properties and applications. One of the main lessons is that there are two different notions of symplectic cohomology associated with two different topologies that one can consider on the colimit of a sequence of Banach chain complexes. The first of these involves completing at the chain level so as to obtain a Banach space. The details at the chain level are described in Section 8.3, or, at the cohomology level, in equation (64). The Banach topology gives rise to local invariants and corresponds under mirror symmetry to locally defined analytic polyvector fields. Similar constructions have been carried independently in [62; 61], but the construction has roots back in [15]. We refer to this as *local symplectic cohomology*.

A second topology one may consider is one in which no completion is applied to the direct limit. As a topological space we consider the direct limit with the *final topology* described in Section 9.5. We refer to this as *global symplectic cohomology*. Global symplectic cohomology is a generalization of the construction in [63]; see also [56, Section 3e], which is more explicit in this regard. It gives rise to global invariants and can be thought to correspond under mirror symmetry to the ring of algebraic polyvector fields. While this distinction, referred to in [56] as quantitative vs qualitative symplectic cohomology, has been previously known, its significance appears to have been masked to a large extent in the literature so far due to the emphasis on Liouville domains with trivially valued coefficient fields, where various different invariants coincide. In general, different topologies may give rise to completely different vector spaces. For an example of this phenomenon see [62].

## 2.1 Local symplectic cohomology

Let $K \subset M$ be a compact set. Let

$$H_K(x) := \begin{cases} 0 & \text{if } x \in K, \\ \infty & \text{if } x \in M \setminus K. \end{cases}$$

The *local symplectic cohomology* at $K$ is defined by

$$SH^*(M\,|\,K;\mathbb{K}) := \overline{HF}^*(H_K;\mathbb{K}).$$

The following theorem lists the basic properties of $SH^*(M\,|\,K;\mathbb{K})$, which can be almost readily read off Theorems 1.2 and 1.3. As before, there is a decomposition

$$SH^*(M;\mathcal{H}) = \bigoplus_{\alpha\in[S^1,M]} SH^{*,\alpha}(M;\mathcal{H}),$$

and we denote by 0 the class of contractible loops.

**Theorem 2.1** (a) *The map $K \mapsto SH^*(M|K;\mathbb{K})$ is contravariantly functorial with respect to inclusions.*

(b) *Any symplectomorphism $\psi\colon M \to M$ induces an isometry*

$$\psi_*\colon SH^*(M|K;\mathbb{K}) \to SH^*(M|\psi(K);\mathbb{K}).$$

(c) *The group $SH^{*,0}(M|K;\mathbb{K})$ is a unital $\mathbb{K}$–algebra with respect to the operation $*$ induced from the identification $SH^{*,0}(M|K;\mathbb{K}) = SH^{*,0}(M;\{H_K\})$.*

(d) *We have a commutative triangle of $\mathbb{K}$–algebras*

$$\begin{tikzcd} H^*(M;\mathbb{K}) \arrow[d] \arrow[dr] & \\ SH^{*,0}(M|K_2;\mathbb{K}) \arrow[r] & SH^{*,0}(M|K_1;\mathbb{K}) \end{tikzcd} \tag{1}$$

(e) *For any $H \in \mathcal{H}$ which is bounded on $K$ we have a continuous and functorial map*

$$\overline{HF}^*(H;\mathbb{K}) \to SH^*(M|K;\mathbb{K}),$$

*which increases the valuation*[1] *by at most $c = \sup_K H$.*

The proof of Theorem 2.1 appears at the end of Section 9.

**Remark 2.2** Suppose $M$ is symplectically aspherical and for a pair of compact sets $K = K_1, K_2$, we have that $H_K$ can be approximated by Hamiltonians whose nontrivial periodic orbits have action positive and bounded away from 0. Then the commutative triangle (1) can be refined to a commutative square

$$\begin{tikzcd} H^*(K_2;\mathbb{K}) \arrow[r] \arrow[d] & H^*(K_1;\mathbb{K}) \arrow[d] \\ SH^*(M|K_2;\mathbb{K}) \arrow[r] & SH^*(M|K_1;\mathbb{K}) \end{tikzcd}$$

Combined with (3) below, this reproduces Viterbo's commutative square for Liouville domains [63].

**Remark 2.3** We comment on the name local symplectic cohomology. Assume $M$ is symplectically aspherical and the boundary of $K$ is stable Hamiltonian. Then it can be shown that elements of $SH^*(M|K)$ are represented by linear combinations of

[1] By definition, the valuation is $\mathrm{val} := \log\|\cdot\|$.

constant periodic orbits inside $K$ and the Reeb orbits of $\partial K$. If the boundary is not stable Hamiltonian, these can be represented by, in addition to constant orbits inside, periodic orbits lying arbitrarily close to $\partial K$. Thus, at least when $M$ is aspherical, $SH^*(M|K)$ can be thought of as symplectic cohomology relative to the complement of $K$; that is, as localized at $K$. When $M$ is not aspherical, this type of locality is far from clear. This question is taken up in forthcoming work. Theorem 2.4 below can be seen as a particular manifestation of locality in the general case.

**Theorem 2.4** *Let $H$ be a smooth Hamiltonian such that $H^{-1}(0) = \partial K$.*

(a) *Suppose $\alpha$ is a nontrivial free homotopy class of loops. Suppose also that $SH^{*,\alpha}(M|K;\mathbb{K}) \neq 0$. Then there is a sequence $a_n > 0$ converging to $0$ such that $H^{-1}(a_n)$ has a periodic orbit representing $\alpha$.*

(b) *If $SH^{*,0}(M|K;\mathbb{K}) \neq H^*(K;\mathbb{K})$, then there is a sequence $a_n > 0$ converging to $0$ such that $H^{-1}(a_n)$ has a contractible periodic orbit.*

Theorem 2.4 is proven in Section 10.6. Some applications of local Floer cohomology to embedding and displaceability problems are given in Section 2.3 below.

We conclude with some comments on the relation of these groups with similar work of others.

(a) When $M$ is symplectically aspherical and $K$ is the closure of an open set $U$, the groups $SH^*_{[a,b)}(M|K)$ coincide with the corresponding symplectic cohomology groups of $U$ as defined in [15] using Hamiltonians which are constant at infinity.

(b) In [62] the notion of completed symplectic cohomology is introduced and studied for Liouville cobordisms $\mathcal{W}$ inside monotone symplectic manifolds. The computations in [62] show that the local symplectic cohomology groups depend nontrivially on $K$. The choice of Floer data in [62] is such that the Floer chain complexes have finite boundary depth; see Remark 3.2. In particular, ordinary and reduced Floer cohomology coincide for these Floer data. A consequence of Theorem 3.3 is that the invariant of [62] is the local Floer cohomology as defined here.

(c) In [61] an invariant which is closely related to local symplectic cohomology as defined here is studied and is shown to fit into a local-to-global spectral sequence when the compact sets involved are invariant sets of an involutive system of Hamiltonians.

## 2.2 Global symplectic cohomology

Consider the set $\mathcal{H}_{\mathrm{univ}} \subset \mathcal{H}_{\mathrm{sc}}$ of *smooth* time-independent exhaustion functions on $M$. Then $\mathcal{H}_{\mathrm{univ}}$ is a monoidal indexing set. By Theorem 1.3 we thus obtain for any geometrically bounded symplectic manifold a $k$–algebra which is a symplectic invariant.

**Definition 2.5** The *universal symplectic cohomology* is defined by

$$SH^*_{\mathrm{univ}}(M) := SH^*(M;\mathcal{H}_{\mathrm{univ}}).$$

For any choice of $\mathcal{H}$ the algebra $SH(M;\mathcal{H})$ carries a topology, called the *final topology*, as a direct limit of topological vector spaces. This topology is not guaranteed a priori to be Hausdorff, and its Hausdorff completion is not guaranteed to be metrizable. However, in the few cases where something is known about it, $SH^*_{\mathrm{univ}}$ turns out to be a reasonably well-behaved object. Example 9.20 below should give some sense of what universal symplectic cohomology is like in nice cases.

Note that for a compact set $K$ we have $\mathcal{H}_{\mathrm{univ}} \preceq \mathcal{H}_K$. Thus there is a natural unital map $SH^*_{\mathrm{univ}}(M) \to SH^*(M|K)$ for any compact set. One way to utilize it is if one finds a monoidal indexing set $\mathcal{H} \subset \mathcal{H}_{\mathrm{univ}}$ for which $SH^*(M;\mathcal{H}) = 0$, it then follows that $SH^*(M|K) = 0$ for all compact $K \subset M$. Observe that since we do not complete after taking the direct limit, the algebra $SH^*(M;\mathcal{H})$ is not sensitive to behavior on compact sets. Indeed, if we define an equivalence relation $\mathcal{H}_1 \sim \mathcal{H}_2$ by $\mathcal{H}_1 \preceq \mathcal{H}_2$ and $\mathcal{H}_2 \preceq 1$, then the associated symplectic cohomologies are canonically isomorphic. On the other hand, if $\mathcal{H}$ consists of continuous functions, the $\sim$–equivalence class of $\mathcal{H}$ is unaffected by any alterations on any compact set. Thus, for $\mathcal{H} \subset \mathcal{H}_{\mathrm{univ}}$ the algebra $SH^*(M;\mathcal{H})$ is only sensitive to the growth at infinity. For this reason we refer to this type of symplectic cohomology as global $SH$.

Before applying $SH^*_{\mathrm{univ}}(M)$ we discuss some settings where something can be said about it.

Let $(M,\omega)$ be a compact symplectic manifold and let $\psi\colon M \to M$ be a symplectomorphism. Denote by $\widetilde{M}_\psi$ the associated symplectic mapping torus; see Section 10.2 for the definition. Denote by $HF^*(M,\psi)$ the fixed-point Floer homology of $\psi$ as introduced in [21]. The following theorem allows us to distinguish mapping tori by fixed-point Floer homology. $\widetilde{M}_\psi$ carries a distinguished closed 1–form $dt$ pulled back from the natural map $\widetilde{M}_\psi \to S^1$. The 1–form $dt$ induces a grading on $SH^*_{\mathrm{univ}}$ since continuation maps are homotopies. We denote by $SH^{*,k}_{\mathrm{univ}}(\widetilde{M}_\psi)$ the $k^{\mathrm{th}}$ graded piece.

**Theorem 2.6** (cf [22]) *There is a map*

$$\bigoplus_{k\in\mathbb{Z}} HF^*(M,\psi^k)\to SH^*_{\mathrm{univ}}(\widetilde{M}_\psi),$$

*which is injective and dense. Moreover, for each $k\in\mathbb{Z}$ there is a natural isomorphism*

$$HF^*(M,\psi^k)=SH^{*,k}_{\mathrm{univ}}(M).$$

*In particular, let $\psi_i\colon M\to M$ be a symplectomorphism for $i=0,1$. Suppose there exists a symplectomorphism*

$$\phi\colon \widetilde{M}_{\psi_1}\to\widetilde{M}_{\psi_2}$$

*which preserves the class of $dt$. Then $\phi$ induces an isomorphism $HF^*(M,\psi_1)=HF^*(M,\psi_2)$.*

Theorem 2.6 is proven in Section 10.2.

**Theorem 2.7** (Künneth formula) *Let $M_1$ and $M_2$ be geometrically bounded symplectic manifolds. Then there is a natural map*

$$SH^*_{\mathrm{univ}}(M_1)\otimes SH^*_{\mathrm{univ}}(M_2)\to SH^*_{\mathrm{univ}}(M_1\times M_2), \tag{2}$$

*which is injective with dense image. A similar claim holds if one restricts to $SH^{*,0}_{\mathrm{univ}}$.*

Theorem 2.7 is proven in Section 10.4.

The following theorem refers to the additional grading on $SH^*_{\mathrm{univ}}(M)$ by free homotopy classes of loops as discussed in the paragraph preceding Theorem 1.3.

**Theorem 2.8** (nearby existence)

(a) *Suppose $SH^{*,0}_{\mathrm{univ}}(M)=0$. Then for any Hamiltonian $H\colon M\to\mathbb{R}$ which is proper and bounded from below, the subset of levels containing a contractible periodic orbit is dense in $H(M)\subset\mathbb{R}$. The claim holds also if we merely assume $SH^{*,0}_{\mathrm{univ}}(M)=\overline{\{0\}}$, where $\overline{\{0\}}$ is the closure of $0\in SH^{*,0}_{\mathrm{univ}}$ with respect to the final topology on $SH^{*,0}_{\mathrm{univ}}$.*

(b) *Suppose $\alpha$ is a nontrivial free homotopy class of loops. Suppose also that $SH^{*,\alpha}(M)\neq\{0\}$. Then there is a compact $K\subset M$ such that for any smooth proper and bounded below $H\colon M\to\mathbb{R}$ and any $a\in\mathbb{R}$ for which $H(K)\subset(-\infty,a]$, the set of $x\in[a,\infty)$ for which $H^{-1}(x)$ has a periodic orbit representing $\alpha$ is dense in $[a,\infty)$.*

Theorem 2.8 is proven in Section 10.6.

**Remark 2.9** Examples satisfying the hypotheses of the first part of Theorem 2.8 are complete toric varieties $M$ such that $c_1(M)=0$. This follows from the vanishing criterion of Theorem 10.9. See Example 10.12. There are manifolds in this class of examples which, unlike $\mathbb{C}$, contain nondisplaceable sets. Examples are the canonical bundles over $\mathbb{P}^2$ and over $\mathbb{P}^1\times\mathbb{P}^1$. By the Künneth formula, the product of such a manifold with any geometrically bounded symplectic manifold of vanishing Chern class will again satisfy the hypothesis.

An example of an $M$ and $\alpha$ satisfying the hypotheses of the second part of Theorem 2.8 is given by the cotangent bundle of the torus and any nontrivial homotopy class $\alpha$. This can be deduced from Theorems 2.6 and 2.7. From this we obtain many examples by taking the product with an arbitrary compact manifold or with a geometrically bounded one for which symplectic cohomology does not vanish, and considering homotopy classes pulled from the cotangent factor.

We can also use the methods of this paper to produce periodic orbits with prescribed action. Namely, for a dissipative Hamiltonian $H$, call a class $a\in\overline{HF}^*(H)$ *essential* if it maps to a nonzero class in $SH^*_{\mathrm{univ}}(M)$. Suppose $M$ is symplectically aspherical. If $H_1\le H_2$ are dissipative then for any essential class $a$ in $HF^*(H_1)$ there is a periodic orbit of $H_2$ in the same homotopy class with action bounded by $\mathrm{val}(a)$. Indeed, the map $HF^*(H_1)\to SH^*_{\mathrm{univ}}(M)$ factors through $HF^*(H_2)$ by the continuation map which is action decreasing.

**Example 2.10** On a Liouville domain, for any function $H$ which is convex at infinity, all nonzero classes in $HF^*(H)$ are essential. This follows from Theorem 2.11 below. The same holds for the product of a Liouville domain with a compact aspherical manifold. These claims require working over $R$ instead of over $\Lambda_R$, but this is not problematic in this restricted setting since the action spectrum is bounded below and so the topology is discrete.

## 2.3 Liouville domains and displaceability

Let $M$ be the completion of a Liouville domain $U$. Denote by $SH^*_{\mathrm{Viterbo}}(U;\mathbb{K})$ the symplectic cohomology as defined in [63] by taking a direct limit of the Floer cohomology groups $HF^*(H,J)$ over all $(H,J)$ where $H$ is linear at infinity and $J$ is of contact type. See Section 10.1 for notation and definitions. Denote by $\mathcal{L}\subset\mathcal{H}$ the set of Hamiltonians which are linear at infinity. Then $\mathcal{L}$ is a monoidal indexing set. We have

$$SH^*_{\mathrm{Viterbo}}(U;\mathbb{K})=SH^*(M;\mathcal{L},\mathbb{K}),$$

and therefore a natural map

$$f : SH^*_{\mathrm{Viterbo}}(U;\mathbb{K}) \to SH^*_{\mathrm{univ}}(M;\mathbb{K}).$$

We prove in Theorem 10.2 below:

**Theorem 2.11** *The map $f$ is an isomorphism for $\mathbb{K} = R$ coefficients.*

**Corollary 2.12** *For a Liouville manifold $M$ of finite type, $SH^*_{\mathrm{Viterbo}}(M; R)$ is independent of the choice of primitive of $\omega$.*

**Remark 2.13** Theorem 2.11 can generally *not* be expected to be true over a nontrivially valued field. See Remark 10.3 for an explanation on this point.

It is also not hard to show that for any Liouville subdomain $V \subset M$ we have a natural isomorphism of vector spaces

$$SH^*(M|V; R) = SH^*_{\mathrm{Viterbo}}(V; R). \tag{3}$$

Note however that the left-hand side of (3) is naturally a normed vector space while the right-hand side is not. The equation will thus cease to be true over a nontrivially valued field. The generalization of (3) for the nontrivially valued case is the excision principle

$$SH^*(M|V;\Lambda_R) = SH^*(\widehat{V}|V;\Lambda_R) \tag{4}$$

whenever $M$ is a Liouville manifold and $V$ is a Liouville subdomain with $\widehat{V}$ its completion. This follows by the no-escape lemma near the concave boundary of $M \setminus V$. See [48]. We now formulate a theorem showing that this independence of the ambient manifold holds under more general conditions for skeleta of Liouville domains. In the following, we denote by $SH^{*,0}(M|V)$ the subgroup consisting of periodic orbits that are contractible in $M$. The proofs of the following theorems as well as some pertinent definitions are given in Section 10.1.

**Theorem 2.14** *Let $M$ be symplectically aspherical and let $U$ be a Liouville domain with Liouville field $Z$. Let $i : U \to M$ be an embedding with the property that $i_* : H_1(U;\mathbb{R}) \to H_1(M;\mathbb{R})$ is injective. Then, denoting by $\mathrm{Skel}(U,Z)$ the skeleton of $U$ with respect to $Z$,*

$$SH^{*,0}(M\,|\,\mathrm{Skel}(U,Z);\mathbb{K}) = SH^{*,0}(U\,|\,\mathrm{Skel}(U,Z);\mathbb{K}).$$

**Remark 2.15** The restriction to contractible periodic orbits in Theorem 2.14 can be removed by adding the assumption that $M$ is symplectically atoroidal.

Theorem 2.14 implies:

**Theorem 2.16** *Let $U$, $Z$ and $M$ be as in Theorem 2.14 and suppose that*

$$SH^*_{\mathrm{Viterbo}}(U) \neq 0. \tag{5}$$

*Then $\mathrm{Skel}(U, Z)$ is not displaceable.*

Taking $U$ the cotangent disc bundle, this is a well known theorem by Gromov. Namely, $\mathbb{C}^n$ contains no simply connected Lagrangians. The particular case $M = \widehat{U}$, the completion of $U$, is a theorem of [33]. We remark that Theorem 2.16 follows from Theorem 2.14 by a general vanishing principle for the localized Floer cohomology of a displaceable set. We prove this for $M$ aspherical. In [61] this is proven without the asphericity assumption. Note however that the asphericity assumption in the last two theorems cannot be removed. Indeed, there are examples of displaceable Lagrangian spheres [2; 44]. However, there are quantitative counterparts which should hold assuming essentially only geometric boundedness.

**Theorem 2.17** *Suppose that $M$ is monotone or Calabi–Yau, and let $U \hookrightarrow M$ be a Liouville domain. Then there is a $\delta > 0$ for which $SH^{*,0}_{\mathrm{Viterbo}}(U; R)$ embeds into $SH^{*,0}_{[0,\delta)}(M \,|\, \mathrm{Skel}(U, Z); \mathbb{K})$ with valuation $0$ as an $R$–subspace.*

**Theorem 2.18** *Let $M$ be aspherical and let $U \hookrightarrow M$ be a Liouville domain satisfying $SH^*_{\mathrm{Viterbo}}(U) \neq 0$. Then $\mathrm{Skel}(U, Z)$ has positive displacement energy.*

**Remark 2.19** It should not be hard to remove the asphericity assumption. Once this is done, and taking $U$ to be the cotangent disk bundle, we recover a classical theorem by Chekanov [13] stating that Lagrangian submanifolds have positive displacement energy.

# 3 Overview

## 3.1 Diameter control of Floer trajectories

In the next couple of sections we wish to investigate the conditions under which a Floer datum $F \in \mathcal{F}$ gives rise to Floer homology groups. What it comes down to are conditions under which Gromov compactness holds. To sketch an outline of what is

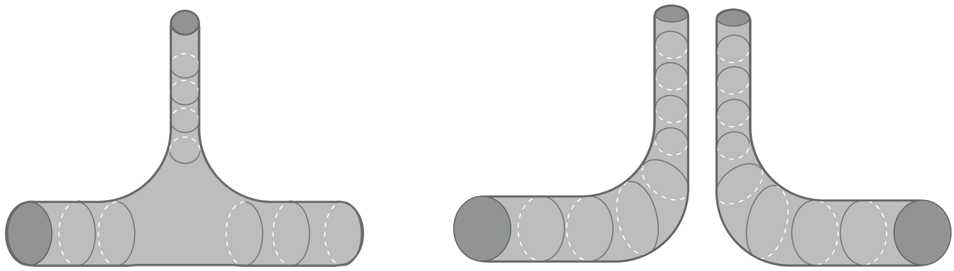

Figure 1: Two types of divergence: type 1, left; type 2, right.

to come, let us first discuss how compactness might fail. Fix the coordinates $(s,t)$ on $\mathbb{R}\times\mathbb{R}/\mathbb{Z}$. Let $u_n\colon \mathbb{R}\times\mathbb{R}/\mathbb{Z}$ be a sequence of solutions to Floer's equation

$$\partial_s u + J(\partial_t u - X_H) = 0 \tag{6}$$

which satisfies, for some positive number $E$ and some compact set $K\subset M$,

$$E(u) := \frac{1}{2}\int \|\partial_s u\|^2 \le E, \quad u(\mathbb{R}\times\mathbb{R}/\mathbb{Z})\cap K \neq \varnothing.$$

In general there are two ways in which such a sequence may diverge. First there might be (after possibly reparametrizing) a fixed value $s$ and a compact set $K'$ such that $u_n(s,\cdot)$ intersects $K'$ but the diameter of $u_n(s,\cdot)$ is not bounded uniformly in $n$. We refer to this as a divergence of type 1; see Figure 1, left. Second, there might be a sequence $s_n\to\infty$ such that $u_n(s_n,\cdot)$ converges to infinity. This is referred to as type 2 divergence; see Figure 1, right.

In the text below we introduce two conditions, one for ruling out each type of divergence. For the first type of divergence we introduce the condition of intermittent boundedness, or i–boundedness. It involves bounds on the geometry of an associated metric on the Hamiltonian mapping torus which are required to hold on a sufficiently large subset of $M$. This condition is introduced first for the case where $H=0$ in Section 4, where we show that it provides diameter control for pseudoholomorphic curves. The condition of i–boundedness is framed so as to allow homotopies between any two elements, as well as higher homotopies, for which the diameter estimate continues to hold. This is the content of Theorem 4.7. Note that it is not reasonable to expect that any two almost complex structures which induce a geometrically bounded metric are connected by a path of the same kind of almost complex structures. Figure 2 illustrates the kind of homotopy that intermittent boundedness allows.

So far, the discussion only pertains to pseudoholomorphic curves. In Section 5 we discuss a trick which allows us to obtain the same diameter control for $H$ nonzero, provided we restrict attention to fixed compact sets of the domain. When $H$ is nonzero, we are considering a geometry which is determined not just by $J$ but also by $H$. Most of Section 5 is devoted to studying the geometry of this metric.

To rule out the divergence of type 2 we introduce a condition called loopwise dissipativity. It is a variant of the Palais–Smale condition, which has played a role in early variational arguments for existence of symplectic capacities [38, Chapter 12]. This condition is not contractible, but this is not a problem since it only needs to be satisfied on the ends. In this it is similar to the nondegeneracy condition which is usually required in Floer theory. Note that unlike the property of i–boundedness, the property of loopwise dissipativity is not readily verifiable on nonexact submanifolds for Hamiltonians that do not have a small Lipschitz constant. In those cases it requires some understanding of the Hamiltonian flow.

Floer data satisfying these conditions are called *dissipative*. Theorem 6.3 states that dissipative Floer data satisfy a priori $C^0$ estimates. A variant which works under a slightly weaker condition on exact symplectic manifolds is given in Theorem 6.12.

We discuss three classes of examples of dissipative $(H, J)$.

(a) $H$ is Lipschitz with respect to $g_J$ with sufficiently small Lipschitz constant outside of a compact set. More generally, mainly to allow a cofinal set, we require the Lipschitz condition only on a sufficiently large subset of $M$. This class of examples is sufficient for all the theoretical constructions of this paper.

(b) $M$ is exact and the action functional satisfies the Palais–Smale condition. For the details see Section 6.4. Strictly speaking, as noted in the beginning of Section 6.4, the Palais–Smale condition is slightly weaker than the dissipativity condition. Nevertheless it fits into the general dissipative framework.

(c) The Hamiltonian flow of $H$ is sufficiently close to being invariant with respect to a radial parameter. See Section 6.5.

## 3.2 Floer theory by approximation

This subsection is devoted to clarifying the statement of Theorem 1.2 and its underlying philosophy. We first discuss the notion of reduced Floer cohomology. The Floer cohomology associated by Theorem 1.1 to a dissipative Floer datum $(H, J)$ is the homology of the Floer complex $CF^*(H,J)$ constructed in Section 7.3. The complexes $CF^*(H,J)$ can be considered as non-Archimedean Banach spaces over $\Lambda_R$, as we explain momentarily. The chain complex $CF^*(H,J)$ is generated by an appropriate Novikov covering of the space of 1–periodic orbits of $H$. Thus Floer cohomology can be considered as the Morse cohomology of a single-valued action functional $\mathcal{A}_H$. Our conventions are set up so that action decreases along gradient lines. See Section 5.1 for precise definitions. Thus our chain complexes carry a decreasing filtration by $\mathcal{A}_H$. Moreover, the continuation maps of Theorem 1.2 are induced by certain chain maps which preserve this filtration.

$CF^*(H,J)$ is thus normed with norm given by (54). The fact that the differential and continuation maps are action decreasing means they are bounded with respect to this norm, and, in particular, continuous. On an open manifold, $CF^*(H,J)$ will typically not be finitely generated over the Novikov ring. For the differential and continuation maps to be well defined we need to consider the completed complex $\widehat{CF}^*(H,J)$. Moreover, the differential can generally not be expected to have a closed image.

**Definition 3.1** Let $(C^*,d)$ be a normed complex. The *reduced cohomology* of $C^*$ is

$$\overline{H}^*(C^*,d) := \widehat{\ker d^*}/\overline{\operatorname{im} d^{*-1}},$$

with the hat denoting completion with respect to the norm, and the bar denoting the closure inside the completion. For a dissipative $H$, we denote the reduced Floer cohomology by $\overline{HF}^*(H)$.[2]

**Remark 3.2** When the Floer complex is finitely generated over a field, the differential has closed image. So, in that case, reduced Floer cohomology coincides with ordinary Floer cohomology. The same is true whenever the Floer complex has finite boundary depth, meaning that the differential has a bounded right inverse [60]. For Liouville domains, the Floer differential for a strictly convex Hamiltonian has a closed image if one is working over $R$, but not necessarily when working over $\Lambda_R$.

Denote by $\mathcal{H}^{\mathbb{N},\geq}_{\mathrm{d,reg}}$ the set of sequences $\{H_i\}$ of regular dissipative Hamiltonians satisfying $H_i(x) \leq H_{i+1}(x)$ for all $i$ and for all $x \in \mathbb{R}/\mathbb{Z} \times M$. The set $\mathcal{H}^{\mathbb{N},\geq}_{\mathrm{d,reg}}$ carries a natural order relation. Namely, $\{H^1_i\} \leq \{H^2_i\}$ is defined to hold if and only if for any $i$ there is

[2]We suppress $J$ in the notation since the homology is independent of $J$ as a consequence of part (c) of Theorem 1.1.

a $j$ such that $H_i^1 \leq H_j^2$. General nonsense about filtered complexes leads to a certain extension of the functor $\overline{HF}^*$ to the category $\mathcal{H}_{\mathrm{d,reg}}^{\mathbb{N},\geq}$ as follows. For a dissipative Floer datum $(H, J)$ and an interval $[a,b) \subset \mathbb{R}$ we can consider the *action-truncated Floer cohomology* $HF^*_{[a,b)}(H)$. See (55) for its definition. Given intervals $[a,b)$ and $[a',b')$ such that $a' \leq a$ and $b' \leq b$, there is a natural map $HF^*_{[a,b)}(H) \to HF^*_{[a',b')}(H)$ which behaves functorially with respect to continuation maps. We then define

$$\overline{HF}^*(\{H_i\}) := \varprojlim_{a} \varinjlim_{b,i} HF^*_{[a,b)}(H_i). \tag{7}$$

The motivation behind this definition will be clarified in Theorem 3.3 and the comments following it.

On the other hand, by Dini's theorem from basic calculus, there exists a functor, that is, an ordered map

$$\sup\colon (\mathcal{H}_{\mathrm{d,reg}}^{\mathbb{N},\geq}, \leq) \to (\mathcal{H}_{\mathrm{sc}}, \leq),$$

which takes $\{H_i\}$ to the function $x \mapsto \sup_i H_i(x)$.

**Theorem 3.3** *The map* $\sup$ *is surjective. Moreover, if* $\sup(\{H_i^1\}) = \sup(\{H_i^2\})$, *there is a natural isomorphism*

$$\overline{HF}^*(\{H_i^1\}) = \overline{HF}^*(\{H_i^2\}). \tag{8}$$

*In particular, there is an induced Floer cohomology functor* $\overline{HF}^*$ *from the category* $\mathcal{H}_{\mathrm{sc}}$ *to the category of* $\mathbb{Z}$*–graded non-Archimedean Banach spaces over* $\mathbb{K}$*. This definition of* $\overline{HF}^*$ *coincides on the subcategory* $\mathcal{H}_{\mathrm{d,reg}} \subset \mathcal{H}_{\mathrm{sc}}$ *with the previous definition which is implied by Theorem 1.1.*

Theorem 3.3 is proved towards the end of Section 8.2 right before Lemma 8.14.

Let us unpack the meaning of Theorem 3.3.

(a) Theorem 3.3 allows one to talk about reduced Floer cohomology of a smooth proper exhaustion Hamiltonian $H$ *without first establishing that $H$ is dissipative*. The further extension to lower semicontinuous functions is of interest since the characteristic function of an open set is lower semicontinuous. This is used in the discussion of local Floer cohomology of compact sets (defined as Floer cohomology of the characteristic function of the complement).

(b) The heart of the proof of the isomorphism (8) is Theorem 8.9, which can be interpreted as saying that the *truncated* Floer cohomology is continuous with respect to *convergence on compact sets*. This continuity is a consequence of the quantitative nature of our main $C^0$ estimate Theorem 6.3. Namely, Floer trajectories connecting regions that are far apart must have high energy. Thus, for fixed action truncation, regions that are sufficiently far apart don't interact Floer-theoretically.

This continuity statement is not true for the *reduced* Floer cohomology $\overline{HF}^*(H)$. Indeed, it is easy to construct examples of a monotone sequence $H_i$ of regular dissipative Hamiltonians converging on compact sets to a regular dissipative Hamiltonian $H$ for which

$$\lim_i \overline{HF}^*(H_i) \neq \overline{HF}^*(H) = \overline{HF}(\{H_i\}).$$

The discrepancy between the leftmost side and rightmost side in the last equation arises because of the interchange of direct and inverse limits. For example, on a Liouville domain, $H$ can be taken to be a quadratic Hamiltonian, while the sequence $H_i$ can be taken to consist of Hamiltonians whose slope near infinity is constant and less than the smallest period of a Reeb orbit. This can be done so that for each compact set the sequence still converges uniformly to $H$. Then it can be shown that for each $i$ we have $\overline{HF}^*(H_i) = H^*(M)$. This is the case since, up to isomorphism, the Floer cohomology $\overline{HF}^*(H_i) = H^*(M)$ depends only on the slope at infinity. Thus the left-hand side is isomorphic to singular cohomology, whereas the right-hand side is not generally isomorphic to singular cohomology. The reason for the discrepancy is that the Hamiltonians $H_i$ will have many periodic orbits whose action is arbitrarily close to $-\infty$. These cancel the contribution to $\overline{HF}^*(H_i)$ coming from the high-action nontrivial periodic orbits. However, when truncating below at any fixed value as in the procedure described by (7), the contribution of the high-action periodic orbits remains uncanceled.

(c) Continuity of truncated Floer cohomology with respect to *uniform* convergence, and hence an extension of the definition of truncated Floer homology to $C^0$ Hamiltonians, has to the author's knowledge first been observed in [63].

(d) In the text, a much stronger statement than the isomorphism of (8) is proven. Namely, it is shown that to each of $j = 1, 2$ one can associate a complete filtered chain complex $\widehat{CF}^*(H_i^j)$, after making some additional choices, such that $\overline{HF}^*(\{H_i^j\})$ is the reduced cohomology of $\widehat{CF}^*(H_i^j)$. It is then shown that these complexes are filtered quasi-isomorphic. See Definition 8.15. Filtered quasi-isomorphism is an equivalence relation which implies isomorphism of the reduced Floer cohomology. In a forthcoming note, joint with U Varolgunes, we show that filtered quasi-isomorphism in fact implies quasi-isomorphism in the usual sense. Thus Theorem 3.3 can be strengthened to a statement concerning unreduced Floer cohomology. This will be a great advantage as it will allow the application of tools from homological algebra. The notion of reduced cohomology is still central however to our construction of the product in symplectic cohomology, as it is purely cohomological. Chain level constructions involving colimits are generally extremely involved, as one needs to keep track of higher homotopical data. The construction at the chain level is carried out in Theorem 8.16 by taking an appropriate kind of chain level limit, which takes the Banach topology of the complexes $CF^*(H_i^j)$ into account. This builds on a construction from [5] and has also been utilized in [62; 61].

(e) Theorem 3.3 relies on the possibility of approximating any element in $\mathcal{H}_{\mathrm{sc}}$ from below by a sequence of functions which have small Lipschitz constant and are thus dissipative by Theorem 1.1. Proper functions which are not bounded below would require considering, in addition, inverse limits. We do not pursue this here.

(f) The comment (a) allows one to adapt Floer-theoretic constructions to the geometry of the specific setting one is interested in without having to worry about complicated compactness questions. For an example of this, see the derivation of the Künneth formula in Hamiltonian Floer cohomology in Section 10.3. Two cautionary remarks are in order, however:

(i) For there to be a relation between the reduced Floer cohomology and periodic orbits of the Hamiltonian we are investigating, we must at least rule out divergence of the second type, described in Section 3.1 below. Namely, we need to establish loopwise dissipativity, or some related property. In the geometrically interesting settings that the author is aware of, this is straightforward, but it would be interesting to have a better understanding of this property.

(ii) It is theoretically possible for there to exist Floer data $(H, J)$ which are not dissipative, but for which, due to some accident, all the Floer moduli spaces are compact and, moreover, give rise to reduced Floer homologies differing from $\overline{HF}^*(H, J)$ as stipulated by Theorem 3.3. This cannot happen for $(H, J)$ which satisfy the following robustness property enjoyed by dissipative Floer data:

> *The set of Floer solutions intersecting a given compact set $K$ and having energy at most $E$ does not change if the Floer datum is changed outside of a sufficiently large ball around $K$.*

Note that the usually employed maximum principles are global in nature and so do not imply this property.

Sections 4 through 6 are devoted to the construction of dissipative Floer data. They are organized as follows. Sections 4 and 5 are concerned with ruling out type 1 divergence. In Section 4 we introduce the notion of i–boundedness, establish its contractibility and derive various versions of diameter estimate it implies. In Section 5 we introduce the Floer equation and the Gromov metric. We introduce the notions of i–bounded and geometrically dissipative almost Floer data. Finally, we study the geometry of the Gromov metric for translation-invariant Floer data. Section 6 is concerned with ruling out type 2 divergence. In it we introduce and study the property of loopwise dissipativity, and establish a diameter estimate as well as some effective criteria.

# 4 I–bounded almost complex structures

For a Riemannian metric $g$ on a manifold $M$ and a point $p \in M$ we denote by $\mathrm{inj}_g(p)$ the radius of injectivity, and by $\mathrm{Sec}_g(p)$ the maximal sectional curvature at $p$. We drop $g$ from the notation when it is clear from the context.

**Definition 4.1** Let $(M, g)$ be a complete Riemannian manifold. For $a > 0$, the metric $g$ is said to be *a–bounded* at a point $p \in M$ if $\mathrm{inj}(x) \geq 1/a$ and $|\mathrm{Sec}(x)| \leq a^2$ for all $x \in B_{1/a}(p)$.

We say that $g$ is *intermittently bounded*, abbreviated *i–bounded*, if there is an exhaustion $K_1 \subset K_2 \subset \cdots$ of $M$ by precompact sets and a sequence $\{a_i\}_{i\geq 1}$ of positive numbers such that:

(a) We have $d(K_i, \partial K_{i+1}) > \dfrac{1}{a_i} + \dfrac{1}{a_{i+1}}$.

(b) The metric $g$ is $a_i$–bounded on $\partial K_i$.

(c) We have

$$\sum_{i=1}^{\infty} \frac{1}{{a_i}^2} = \infty. \tag{9}$$

The data set $\{K_i, a_i\}_{i\geq 1}$ is called taming data for $(M, g)$.

More generally we allow a slight weakening in the definition and say that a Riemannian metric $g$ is i–bounded if there exists a metric $g'$ that is i–bounded as above, with taming data $(K_i, a_i)$ and a sequence of constants $C_i$ such that:

(a) We have

$$\sum_{i=1}^{\infty} \frac{1}{(C_i a_i)^2} = \infty. \tag{10}$$

(b) The metric $g$ is $C_i$–quasi-isometric to $g'$ on $B(\partial K_i, 1/a_i)$. Namely,

$$\frac{1}{C_i}\|X\|_g \le \|X\|_{g'} \le C_i\|X\|_g$$

on $B(\partial K_i, 1/a_i)$.

In this case we will refer to the sequence $(K_i, a_i, C_i)$ as the taming data of $g$.

For a symplectic manifold $(M, \omega)$, an $\omega$–compatible almost complex structure $J$ is called *i–bounded* if $g_J$ is i–bounded. The symplectic form $\omega$ is said to be *i–bounded* if it admits an i–bounded almost complex structure. For an i–bounded $(M, \omega)$, denote by $\mathcal{J}_{\mathrm{ib}}(M, \omega)$ the space of i–bounded almost complex structures.

A $k$–parameter family $(g_t)_{t\in[0,1]^k}$ of i–bounded Riemannian metrics on $M$ is said to be *uniformly i–bounded*, or *u.i.b.*, if there is an $\epsilon > 0$ such that for each $t_0 \in [0,1]^k$ the taming data $\{K_i, a_i, C_i\}$ can be chosen fixed on the $\epsilon$ neighborhood of $t_0$. A family $\{J_t\}$ of almost complex structures is called *u.i.b.* if the corresponding family $\{g_{J_t}\}$ of Riemannian metrics is uniformly i–bounded.

**Example 4.2** If $J$ is geometrically bounded, meaning that $g_J$ is $a$–bounded everywhere for some $a$, it is i–bounded. In this case, we can take the taming data to be $\{K_i = B_{3i/a}(p), a_i = a\}$, for some arbitrary point $p \in M$.

**Example 4.3** Suppose now that $f : M \to \mathbb{R}$ is the distance from some point $p \in M$ and that at each point $x \in M$ the metric $g_J$ is $f(p)$–bounded. Then $g_J$ is still i–bounded. For this case consider the sequence of real numbers $b_i$ obtained from the set $\bigcup_{n=1}^{\infty}\{n + k/n \mid 0 \le k < n\} \subset \mathbb{R}$ with its standard order. Then the sequence $(K_i = f^{-1}(0, b_{3i}), a_i = \lceil b_{3i} \rceil)$ constitutes taming data for $g_J$. Indeed, by assumption, the metric is $a_i$ bounded on $K_i$ and the series $\sum 1/a_i^2$ is readily seen to diverge.

**Remark 4.4** The condition of uniform i–boundedness is framed so that it simultaneously guarantees the conclusions of Theorems 4.7 and 4.11 below. Namely, on the one hand, the condition is contractible in the sense that any two homotopies satisfying the condition are connected by a homotopy satisfying the same condition. On the other hand, it still allows a priori control of the diameters of $J$–holomorphic curves. If we were to require boundedness everywhere, not just near $\partial K_i$, it appears unlikely that we would get a contractible condition as required in invariance proofs.[3]

**Remark 4.5** Theorem 4.7 below will remain true if we impose more stringent requirements on the numbers $a_i$, say, that they be bounded by a given constant. The reason we allow the numbers $a_i$ to diverge (subject to (10)) is that in the context of Floer theory, sometimes there naturally arise almost complex structures with associated metrics that do not have uniformly bounded sectional curvature. Examples are the Sasaki metric on the cotangent bundle and the induced metric on the mapping torus of a quadratic Hamiltonian on the completion of a Liouville domain.

**Remark 4.6** If $J$ is i–bounded and $J'$ is such that $\|J - J'\|_{g_J}$ is bounded, then $J'$ is i–bounded.

**Theorem 4.7** *The space $\mathcal{J}_{\mathrm{ib}}(M,\omega)$ is connected. Moreover, any two elements can be connected by a u.i.b. family. Similarly, any two u.i.b. $k$–parameter families can be connected by a u.i.b. $(k+1)$–parameter family.*

**Remark 4.8** The idea of the proof is very similar to the that of [14, Proposition 11.22].

**Proof** Let $J_0, J_1 \in \mathcal{J}_{\mathrm{ib}}$. Suppose we are given taming data $\{K^i_n, a^i_n, C^i_n\}_{n\geq 1}$ for $J_i$, $i = 0, 1$. For the rest of the proof we assume $C^i_n = 1$, the adjustment to the general case being trivial. Let $(c^i_n, d^i_n)_{n\geq 1}$ be sequences of positive integers constructed inductively such that:

(a) $\overline{K^0_{d^0_n + c^0_n}} \subset K^1_{d^1_n}$ and $\overline{K^1_{d^1_n + c^1_n}} \subset K^0_{d^0_{n+1}}$ for all $n$.

(b) $\displaystyle\sum_{k=d^i_n+1}^{d^i_{n+c_n-1}} \left(\frac{1}{a^i_k}\right)^2 \geq \frac{1}{n}$ for $i = 0, 1$.

Write

$$V^i_n := \overline{K^i_{d^i_n + c^i_n}} \setminus K^i_{d^i_n} \quad \text{for } i = 0, 1.$$

The sets $V^i_n$ are all disjoint by (a). Let $\{J_s\})_{s\in[0,1]}$ be a smooth homotopy connecting $J_0$ and $J_1$ which is fixed and equal to $J_1$ on the subsets $V^0_n$ for all $s \in \left[0, \frac{2}{3}\right]$ and to $J_1$

[3] As evidence for this, consider that one can show, using the result of [40], that the space of complete Riemannian metrics inducing a given volume form and having bounded geometry is disconnected. In fact, it has infinitely many connected components.

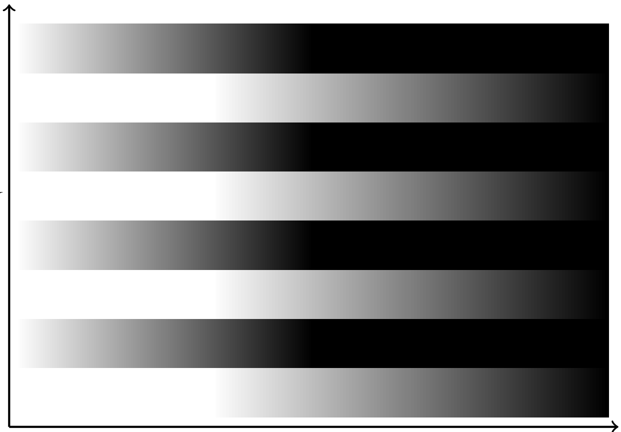


Figure 2: A zigzag homotopy from $J_0$ (light) to $J_1$ (dark).

on the subsets $V_n^1$ for all $s \in \left[\frac{1}{3}, 1\right]$. We refer to such a homotopy as a zigzag homotopy; see Figure 2. Set

$$A^i := \bigcup_n [d_n^i + 1, d_n^i + c_n^i - 1] \cap \mathbb{N}$$

for $i = 0, 1$. By (a) and (b), the data

$$\{K^i_{n^i_k}, a^i_{n^i_k}\}_{n^i_k \in A^i}, \quad i = 0, 1,$$

constitute taming data for $J_s$ on the intervals $\left[0, \frac{2}{3}\right]$ and $\left[\frac{1}{3}, 1\right]$, respectively. Moreover, for each $s \in [0, 1]$ the metric $g_{J_s}$ is complete. Indeed, the distance of $\partial K^i_{n_k}$ from any fixed point goes to $\infty$ for $i = 0$ and $s \in \left[0, \frac{2}{3}\right]$, and for $i = 1$ and $s \in \left[\frac{1}{3}, 1\right]$. We have thus connected $J_0$ and $J_1$ in a uniformly i–bounded way.

We now generalize to the $k$–parameter case. Let $\{J_{i,t}\}_{t \in [0,1]^k}$ for $i = 0, 1$ be two smooth $k$–parameter families. Let $C$ be an open cover of the cube $[0, 1]^k$ by open cubes of side length $\epsilon$ for $\epsilon$ so small that the taming data for both families can be chosen fixed on each such cube. For each $c \in C$ and $i \in \{0, 1\}$, we construct precompact open subsets $\{V_n^{i,c}\}$ in such a way that:

(a) $V_{n_1}^{i_1,c_1}$ is disjoint from $V_{n_2}^{i_2,c_2}$ whenever $(i_1, c_1, n_1) \neq (i_2, c_2, n_2)$.

(b) There is taming data supported in $\bigcup_{n=1}^{\infty} V_n^{i,c}$ for $\{J_{i,t}\}_{t \in c}$, where we say that the taming data $\{K_i, a_i\}_{i \geq 1}$ is supported in an open set $V \subset M$ if $V$ contains all the balls $B_{1/a_i}(\partial K_i)$.

Such sets can be constructed inductively along the same lines as in the 0–parameter case. We can then take any smooth homotopy

$$\{J_{s,t}\}_{(s,t)\in[0,1]\times[0,1]^k}$$

which is fixed on all the subsets $\{V_n^{0,c}\}$ for $s \in \left[0, \frac{2}{3}\right]$, and on all the subsets $\{V_n^{1,c}\}$ for $s \in \left[\frac{1}{3}, 1\right]$. □

For a $J$–holomorphic curve $u\colon S \to M$, denote by $E(u; S)$ the energy

$$\int_S u^*\omega$$

of $u$ on $S$. We drop $S$ from the notation when it is clear from the context.

The following theorem is taken from [57].

**Theorem 4.9** (monotonicity) *Let $g_J$ be $a$–bounded*[4] *at $p \in M$. Let $\Sigma$ be a compact Riemann surface with boundary and let $u\colon \Sigma \to M$ be $J$–holomorphic such that $p$ is in the image of $u$ and such that*

$$u(\partial\Sigma) \cap B_{1/a}(p) = \varnothing.$$

*Then there is a universal constant $c$ such that*

$$E\big(u; u^{-1}(B_{1/a}(p))\big) \geq \frac{c}{a^2}.$$

*If $g_J$ is quasi-isometric to an $a$–bounded metric with quasi-isometry constant $A$, the same inequality holds but with $c$ replaced by $c/A^2$.*

**Proof** This is just a reformulation of the monotonicity inequality in [57]; see Proposition 4.3.1(ii) and the comment right after Definition 4.1.1 there. For completeness, we add a statement and proof of that comment in Lemma 4.10, as we didn't find a proof of it in the literature. □

**Lemma 4.10** *Let $g$ be Riemannian metric which is $a$–bounded at $p \in M$. Then any loop $\gamma\colon S^1 \to B_{1/(2a)}(p)$ bounds a disk of area less than $\frac{1}{2}\ell^2(\gamma)$.*

Our proof is taken from [37], the only addition being the precise dependence on the curvature.

[4]As the proof shows, we only need an estimate from above on the sectional curvature. The stronger requirement is needed later in Section 5.5.

**Proof** Set $\gamma(0)=q$. Let $\widetilde{\gamma}\colon S^1\to T_qM$ be the unique path such that $\exp_q\widetilde{\gamma}(\theta)=\gamma(\theta)$. Consider the disk $u(t,\theta)=\exp_q t\widetilde{\gamma}(\theta)$. Using the triangle inequality one shows that $u$ maps into the ball $B_{1/a}(p)$. Since the geodesics emanating from $q$ minimize distance within $B_{1/a}(q)$, we have

$$\|\partial_t u\| = \|\widetilde{\gamma}(\theta)\| = d(q,\gamma(\theta)) \le \tfrac{1}{2}\ell(\gamma). \tag{11}$$

We need to estimate the Jacobi field $J(t,\theta) := \partial_\theta u(t,\theta)$. More precisely, we need to estimate the component $J^\perp$ which is perpendicular to $\widetilde{\gamma}(\theta)$. For this we use the generalized Rauch estimate [34, 1.8.2], according to which the function

$$f(t) = \frac{\|J^\perp(t,\theta)\|}{\sin t}$$

is nondecreasing on the interval $(0,\pi)$.[5] Observe that $\gamma'(\theta)=J(1,\theta)$. So,

$$\|J^\perp(t,\theta)\| \le \frac{\sin t}{\sin 1}\|\gamma'(\theta)\| \quad \text{for } t\le 1.$$

Applying the last estimate and equation (11) we get

$$\begin{aligned}
\mathrm{Area}(u) &= \int_0^1\int_0^{2\pi} \|\partial_\theta u\|\,\|\partial_t u\|\sin(\angle(\partial_\theta u,\partial_t u))\,d\theta\,dt\\
&= \int_0^1\int_0^{2\pi} \|\partial_\theta u^\perp\|\,\|\partial_t u\|\sin(\angle(\partial_\theta u^\perp,\partial_t u))\,d\theta\,dt\\
&\le \int_0^1\int_0^{2\pi} \|\gamma'(\theta)\|\tfrac{1}{2}\ell(\gamma)\,d\theta\,dt = \tfrac{1}{2}\ell^2(\gamma).
\end{aligned}$$ □

The following theorem is fundamental for all that follows. It gives a priori control over the diameter of a $J$–holomorphic curve $u\colon\Sigma\to M$ with free boundary in terms of its energy.

**Theorem 4.11** *Let $J\in\mathcal{J}_{\mathrm{ib}}$.*

(a) *For any compact set $K\subset M$ and $E\in\mathbb{R}_+$ there exists an $R>0$ such that for any connected compact Riemann surface $\Sigma$ with boundary, and any $J$–holomorphic map*

$$u\colon(\Sigma,\partial\Sigma)\to(M,K)$$

*satisfying $E(u;\Sigma)\le E$, we have $u(\Sigma)\subset B_R(K)$. Moreover, if the geometry is uniformly bounded, $R$ can be taken to be independent of $K$. In fact, it can be taken to be proportional to $E$.*

[5] Observe the coordinate $t$ is related to the coordinate $r$ of [34, 1.8.2] by $r=\|\widetilde{\gamma}\|t$, and that $\|\widetilde{\gamma}\|\le a$, where $a^2$ plays the role of $\Delta$ in [34, 1.8.2].

(b) *Let $\Sigma$ be a connected compact Riemann surface with boundary. For any compact set $K \subset M$, any compact subset $S$ of the interior of $\Sigma$, and any $E \in \mathbb{R}_+$, there exists an $R$ such that for any $J$–holomorphic map*

$$u\colon \Sigma \to M$$

*satisfying $E(u;\Sigma) \leq E$ and $u(S) \cap K \neq \varnothing$, we have $u(S) \subset B_R(K)$.*

*In both cases, besides the dependence on $E$ and on $S$, $R$ depends only on taming data of $J$ inside $B_R(K)$. That is, given $J'$ which has the same taming data as $J$ on $B_R(K)$, the claim will hold with the same $R$ for $J'$–holomorphic curves with energy at most $E$.*

**Remark 4.12** The reader is cautioned that in case (b), where there is no control over the image of the boundary, to control the diameter of $u(S)$ we need control of the energy in the larger surface $\Sigma$.

**Remark 4.13** Concerning the dependence of $R$ on the geometry in case (b), in addition to the dependence on the taming data and on $E$, we have that $R$ depends on an estimate from below of the distance $d(S, \partial\Sigma)$, and from above on the area and curvature of $\Sigma$, all with respect to an arbitrarily chosen conformal metric.

**Proof** Let $\{K_i, a_i, C_i\}$ be taming data for $J$. The argument will be given for the case $C_i = 1$ for all $i \in \mathbb{N}$. Let $N \in \mathbb{Z}$ be such that $K \subset K_N$. Let $i_0 > 0$ and $x_{i_0} \in \Sigma$ be such that $u(x_{i_0}) \in \partial K_{i_0+N}$. If no such $i_0$ and $x_{i_0}$ exist, we take $R = d(K, K_{N+1})$ and we are done. Otherwise, there is a sequence $x_i \in \Sigma$ such that $u(x_i) \in \partial K_{N+i}$ for $0 < i \leq i_0$. In case (a) we argue as follows. For each $1 \leq i \leq i_0$, we have $B_{1/a_{N+i}}(u(x_i)) \cap u(\partial\Sigma) = \varnothing$. Also,

$$d(u(x_i), u(x_j)) > \frac{1}{a_{N+i}} + \frac{1}{a_{N+j}}$$

whenever $i \neq j$. By Theorem 4.9 we obtain

$$E(u;\Sigma) \geq \sum_{i=1}^{i_0} E\bigl(u; u^{-1}\bigl(B_{1/a_{N+i}}(u(x_i))\bigr)\bigr) \geq \sum_{i=1}^{i_0} \frac{c}{a_{i+N}^2}.$$

By (10) this implies an a priori upper bound on $i_0$. Let $i_0$ be the largest possible such. The claim then holds with $R = d(K, K_{N+i_0+1})$.

In case (b) we argue as follows. Pick an area form $\omega_\Sigma$ on $\Sigma$ which together with $j_\Sigma$ determines a metric whose sectional curvature is bounded in absolute value by 1.

Let $A = \int_\Sigma \omega_\Sigma$ and let $\epsilon := d(S, \partial\Sigma)$. Let

$$\tilde{u} := \mathrm{Id} \times u \colon \Sigma \to \Sigma \times M$$

be the graph of $u$, and let $\tilde{J}$ be the product almost complex structure on $\Sigma \times M$. Then $\tilde{u}$ is $J$–holomorphic and $E(\tilde{u}) = E(u) + A$. For any $x \in \Sigma$, any $p \in M$ and any $a \geq 1$ such that $(M, g_J)$ is $a$–bounded at $p$, we have that $(\Sigma \times M, g_{\tilde{J}})$ is $a$–bounded at $(x, p)$. Moreover, defining $x_i$ as before for points $x_i \in S$, the ball of radius $\min\{1/a_{i+N}, \epsilon\}$ around $\tilde{u}(x_i) = (x_i, u(x_i))$ does not meet $\tilde{u}(\partial\Sigma)$. Thus, arguing as before, we have

$$E(u; \Sigma) + A = E(\tilde{u}; \Sigma) \geq \sum_{i=1}^{i_0} c \min\left\{\frac{1}{a_{i+N}^2}, \epsilon^2\right\}.$$

The claim follows as before. □

The final ingredient we shall need is the following elementary observation, whose proof we leave for the reader.

**Theorem 4.14** *The pullback of a u.i.b. family by a uniformly continuous map is u.i.b.*

Theorems 4.7 and 4.11 have consequences for symplectic invariants on open manifolds, which we state as the following theorem.

**Theorem 4.15** *The following invariants, whose definition requires fixing a geometrically bounded almost complex structure $J$, are independent of the choice of such $J$.*

(a) *The Gromov–Witten theory on geometrically bounded manifolds studied in* [35].

(b) *Symplectic homology of relatively compact open sets studied in* [17].

(c) *Rabinowitz Floer homology of tame stable Hamiltonian hypersurfaces in geometrically bounded manifolds* [16].[6]

**Proof** By Theorem 4.7 we can connect any two such almost complex structures $J_0$ and $J_1$ via a uniformly i–bounded path $J_s$ of compatible almost complex structures. We outline how to prove invariance in each case separately.

(a) As a particular consequence of Theorem 4.11, the $J_s$–holomorphic curves representing a given homology class and nontrivially intersecting a compactly supported cohomology class are all contained in a fixed compact set $K$. Thus the moduli space of such spheres with $s$ varying from 0 to 1 generically gives rise to a cobordism between the moduli spaces associated with $J_0$ and $J_1$.

[6] See Remark 3.3 in [17] and likewise the beginning of Section 4.5 in [16], where this question is raised.

(b) The symplectic homology is defined by considering compactly supported Hamiltonians. In that setting, geometric boundedness gives rise to $C^0$ estimates as follows. Suppose $u$ is a Floer trajectory connecting periodic orbits inside some open set $U$ where some Hamiltonian $H$ is supported. Then the intersection of the image of $u$ with $M \setminus U$ is $J$–holomorphic. Moreover, the symplectic energy is bounded a priori in terms of the action difference across $u$. It is clear by Theorem 4.11 that i–boundedness is sufficient to obtain the same type of $C^0$ estimate. We show that the continuation map associated with the 1–parameter family $J_s$ fixing $H$ also satisfies a $C^0$ estimate. Note that we cannot directly appeal to Theorem 4.11, since we are now considering a domain-dependent $J$. To overcome this difficulty we apply the Gromov trick. Namely, we consider the graph

$$\widetilde{u}\colon \mathbb{R}\times\mathbb{R}/\mathbb{Z} \to \mathbb{R}\times\mathbb{R}/\mathbb{Z}\times M$$

for a continuation map $u$. Let $\widetilde{J} := j \times J$, where $j$ is the standard complex structure on the cylinder. Then $\widetilde{u}$ is $\widetilde{J}$–holomorphic outside of $U$. Consider the area form on the cylinder obtained by identifying it with the twice punctured sphere. Then the associated metric $g_{\widetilde{J}}$ is i–bounded. Moreover, the energy of the part of $\widetilde{u}$ mapping outside of $U$ is still bounded a priori in terms of the periodic orbits connected by $u$.[7] Appealing to Theorem 4.11, the path $J_s$ gives rise to a chain homotopy between the Floer homologies of any fixed Hamiltonian with respect to the two choices of $J$. In the same way, given $H \geq K$, and a homotopy $H_s$ of Hamiltonians, the concatenations $(K, J_s)\,\#\,(H_s, J_0)$ and $(H_s, J_1)\#(H, J_s)$ can be interpolated by a homotopy $(H_{s,\tau}, J_{s,\tau})$ such that $H_{s,\tau}$ is compactly supported and $J_{s,\tau}$ is uniformly i–bounded. A $C^0$ estimate for the homotopy is immediate from Theorem 4.11. Thus the continuation maps in the directed system defining symplectic homology also coincide generically for different choices of $J$. It follows that the two invariants coincide.

(c) Rabinowitz Floer homology for a stable Hamiltonian hypersurface $\Sigma$ in a geometrically bounded manifold is considered with Hamiltonian vector fields that are supported on some compact set $K$ containing $\Sigma$. A gradient trajectory in Rabinowitz Floer homology consists of a pair $(v, \eta)$, where $v\colon \mathbb{R}\times\mathbb{R}/\mathbb{Z} \to M$ and $\eta\colon \mathbb{R} \to \mathbb{R}$ satisfy a certain equation. The part of $V$ which maps out of $K$ is $J$–holomorphic with a priori bounded energy for a gradient connecting critical points. Compactness of the space of gradient trajectories consists in first establishing a $C^0$ estimate on $V$ and once

[7]Later, when considering a nonzero Hamiltonian, we will generally have i–boundedness only if we consider the cylindrical metric, which has infinite area. For this reason we will need to complement i–boundedness with the additional requirement of loopwise dissipativity.

$V$ is confined to a compact region, deriving an estimate on $\eta$ and appealing to Gromov compactness. As in the previous part, i–boundedness is sufficient for the $C^0$ estimates on $V$. This holds as well for $s$ dependent $J$. The argument for invariance now follows as before. □

**Remark 4.16** The question of what kind of deformation of the symplectic structure preserves which of these invariants appears to be more subtle and is not studied here. In [30], the question is taken up for a particular type of deformation on Liouville domains.

**Remark 4.17** It is not known to the author whether the class of i–bounded symplectic manifolds is strictly larger than the class of geometrically bounded symplectic manifolds. It appears likely that it might be easier to characterize the class of i–bounded symplectic manifolds in terms of the topology of $\omega$. It is easy to see that a punctured Riemann surface cannot be assigned an i–bounded compatible metric, even though there is a compatible complete metric of bounded curvature. This motivates the following question. Suppose $M$ is such that for any disconnecting compact hypersurfaces $\Sigma$, a component of $M \setminus \Sigma$ which has finite volume is precompact in $M$. Are there any obstructions to finding a compatible i–bounded $J$?

It is also an interesting question whether finiteness of the total volume is an obstruction to weak boundedness, as it is to boundedness. In dimension 2 the answer is positive, as remarked above, but in higher dimension this is not clear to the author. If the answer is negative, it is possible that there are contact manifolds whose symplectizations admit i–bounded almost complex structures, allowing one to define Floer-theoretic invariants on them without recourse to symplectic field theory. This remark is due to A Oancea.

# 5 Floer solutions and the Gromov metric

## 5.1 Floer's equation

Let $(M,\omega)$ be a symplectic manifold. Let $\mathcal{L}M := C^\infty(\mathbb{R}/\mathbb{Z}, M)$ denote the free loop space. For a smooth function $H \in C^\infty(\mathbb{R}/\mathbb{Z} \times M)$ and for any $t \in \mathbb{R}/\mathbb{Z}$, denote by $X_{H_t}$ its Hamiltonian vector field. This is the unique vector field satisfying $dH_t(\cdot) = \omega(X_{H_t}, \cdot)$. For each component $\mathcal{L}M_a$ of $\mathcal{L}M$ pick a base loop $\gamma_a$ and define a (multivalued) functional $\mathcal{A}_H : \mathcal{L}M \to \mathbb{R}$ by

$$\mathcal{A}_H(\gamma) := -\int \omega - \int_0^{2\pi} H(\gamma(t))\, dt,$$

where the integral of $\omega$ is taken over a path in loop space from $\gamma_a$ to $\gamma$. Later, in Section 7.3, we will consider $\mathcal{A}_H$ as a single-valued functional on an appropriate cover of the loop space.

Denote by $\mathcal{P}(H) \subset \mathcal{L}M$ the set of 1–periodic orbits of $X_H$. This is the same as the critical point set of $\mathcal{A}_H$. Given an $\mathbb{R}/\mathbb{Z}$–parametrized family of almost complex structures $J_t$ on $M$, the gradient of $\mathcal{A}_H(\gamma)$ at $\gamma$ is the vector field

$$\nabla\mathcal{A}_H(\gamma)(t) := J_t\big(\dot{\gamma}(t) - X_{H_t}(\gamma(t))\big)$$

along $\gamma$. Note that the gradient field is independent of the choice of base paths and is single-valued. A gradient trajectory is a path in (a covering of) loop space, whose tangent vector at each point is the negative gradient at that point. Explicitly a gradient trajectory is a map

$$u\colon \mathbb{R} \times \mathbb{R}/\mathbb{Z} \to M$$

satisfying Floer's equation

$$\partial_s u + J_t(\partial_t u - X_{H_t} \circ u) = 0. \tag{12}$$

We refer to such solutions as *Floer trajectories*. A Floer trajectory is *nontrivial* if there is a point such that $\partial_t u \neq X_H$.

More generally, let $\Sigma$ be a finite type Riemann surface with cylindrical ends. This means that $\Sigma$ is obtained from a compact Riemann surface $\overline{\Sigma}$ by removing a finite number of punctures. Moreover, near each puncture we fix a conformal coordinate system $(s,t)\colon (a,b)\times\mathbb{R}/\mathbb{Z} \to \Sigma$ such that either $(a,b) = (-\infty, 0)$ or $(a,b) = (0,\infty)$. In the first case we call the puncture negative, and in the second, positive. Let $\mathfrak{H} \in \Omega^1(\Sigma, C^\infty(M))$ be a 1–form with values in smooth Hamiltonians such that near each puncture there is an $H \in C^\infty(\mathbb{R}/\mathbb{Z} \times M)$ for which $\mathfrak{H} = H\,dt$ in the cylindrical coordinates. We denote by $X_{\mathfrak{H}}$ the corresponding 1–form with values in Hamiltonian vector fields. Let $J \in C^\infty(\Sigma, \mathcal{J}(\omega))$ and suppose $J$ is independent of the coordinate $s$ on the cylindrical ends. The datum $(\mathfrak{H}, J)$ is called a *domain-dependent Floer datum*.

Let $u\colon \Sigma \to M$ be smooth. For a 1–form $\gamma$ on $\Sigma$ with values in $u^*TM$, write

$$\gamma^{0,1} := \tfrac{1}{2}(\gamma + J \circ \gamma \circ j_\Sigma).$$

A *Floer solution* on $\Sigma$ is a map $u\colon \Sigma \to M$ satisfying Floer's equation

$$(du - X_{\mathfrak{H}}(u))^{0,1} = 0. \tag{13}$$

Note that equation (12) is equivalent to a special case of equation (13). We refer to $J$ and $\mathfrak{H}$ as the Floer data of $u$. The *geometric energy* of $u$ on a subset $S \subset \Sigma$ is defined as

$$E_{\mathfrak{H},J}(u;S) := \frac{1}{2}\int_S \|du - X_{\mathfrak{H}}\|_J^2 \, d\mathrm{vol}_\Sigma. \tag{14}$$

We omit any one of $\mathfrak{H}$, $J$ or $S$ from the notation when they are clear from the context. We define the topological energy $E_{\mathrm{top}}(u)$ of a Floer solution $u$ as follows. Consider $\mathfrak{H}$ as a 1–form on $\Sigma \times M$ and let $\widetilde{u}\colon \Sigma \to \Sigma \times M$ be the product map $\widetilde{u} = \mathrm{Id} \times u$. Then

$$E_{\mathrm{top}}(u) := \int u^*\omega + d\widetilde{u}^*\mathfrak{H}. \tag{15}$$

Floer's equation reduces to the nonlinear Cauchy–Riemann equation when $\mathfrak{H}(v) = \gamma \otimes \mathrm{const}$ for $\gamma$ a 1–form on $\Sigma$. In this case the two definitions of the energy coincide. Namely, we have the identity

$$\frac{1}{2}\int_S \|du - X_{\mathfrak{H}}\|^2 = \frac{1}{2}\int_S \|du\|^2 = \int_S u^*\omega. \tag{16}$$

## 5.2 The Gromov metric

Let $u\colon \Sigma \to M$ be a Floer solution for the Floer data $F = (\mathfrak{H}, J)$. Define an almost complex structure $J_F$ on $\Sigma \times M$ by

$$J_F(z,x) := \begin{pmatrix} j_\Sigma(z) & 0 \\ X_{\mathfrak{H}}(z,x) \circ j_\Sigma(z) - J(z,x) \circ X_{\mathfrak{H}}(z,x) & J(x) \end{pmatrix}.$$

Let

$$\widetilde{u} = (\mathrm{Id}, u)\colon \Sigma \to \Sigma \times M.$$

Then $\widetilde{u}$ is $J_F$–holomorphic. This construction is known as Gromov's trick; see for instance [37, Chapter 8.1].

> *Henceforth, given a Riemann surface $\Sigma$ with cylindrical ends, we shall assume that it comes equipped with an area form which is compatible with the complex structure and coincides with the standard one, $ds \wedge dt$, on the ends.*

Note that $J_F$ is not generally tamed by the product symplectic structure

$$\omega_{\widetilde{M}} = \pi_1^*\omega_\Sigma + \pi_2^*\omega_M.$$

However, we have the following lemma.

**Lemma 5.1** *Suppose $\{\mathfrak{H},\mathfrak{H}\}=0$, ie for any $z\in\Sigma$ and any pair $v_1,v_2\in T_z\Sigma$, we have*

$$\{\mathfrak{H}(v_1),\mathfrak{H}(v_2)\}=0. \tag{17}$$

*Now consider $\mathfrak{H}$ as a 1–form on $\Sigma\times M$ which is trivial in the directions tangent to $M$. Assume that for each $(z,x)\in\Sigma\times M$ we have*

$$d\mathfrak{H}(z,x)|_{T_z\Sigma}\geq 0. \tag{18}$$

*That is, it is positive with respect to the orientation determined by $j_\Sigma$, the complex structure. Then the 2–form*

$$\omega_{\mathfrak{H}}:=\pi_1^*\omega_\Sigma+\pi_2^*\omega_M+d\mathfrak{H}$$

*is a symplectic form on $\Sigma\times M$, which is compatible with $J_F$.*

**Proof** We only show that $\omega_{\mathfrak{H}}$ is a symplectic form. Closedness is clear, so we only need to show nondegeneracy. In local coordinates on $\Sigma$ write

$$\mathfrak{H}=H\,dt+G\,ds.$$

Then

$$d\mathfrak{H}=dH\wedge dt+dG\wedge ds+(\partial_sH-\partial_tG)\,ds\wedge dt.$$

Suppose there is a vector $v=(v_1,v_2)\in T(\Sigma\times M)$ for which $\iota_v\omega_{\mathfrak{H}}=0$. Then, in particular, the restrictions of $\iota_v\widetilde{\omega}$ to the fibers of $\pi_2$ vanish, giving

$$-\iota_{v_2}\omega_M=dt(v_1)\,dH+ds(v_1)\,dG.$$

So, $v_2=aX_H+bX_G$ for appropriate constants $a,b\in\mathbb{R}$. Since $\{H,G\}=0$ it follows that $\iota_{v_2}(dH\wedge dt+dG\wedge ds)=0$. Thus,

$$\iota_{v_1}(\omega_\Sigma+(\partial_sH-\partial_tG)\,ds\wedge dt)=0.$$

Our assumption is that the coefficient of $ds\wedge dt$ is nonnegative. It follows that $v_1=0$, which in turn implies $v_2=0$. □

**Remark 5.2** More generally, if we replace the estimate (18) by

$$d\mathfrak{H}(z,x)|_{T_z\Sigma}\geq -a\,ds\wedge dt \tag{19}$$

for some constant $a$, we have that the form $\omega_{\mathfrak{H},a}:=\omega_{\mathfrak{H}}+a\,ds\wedge dt$ is symplectic.

The Poisson bracket condition (17) may also be weakened to the requirement that for any point $z\in\sigma$ and vectors $v_1,v_2\in T_z\Sigma$, we have

$$|\{\mathfrak{H}(v_1),\mathfrak{H}(v_2)\}(x)|<a\|v_1\|\|v_2\|\quad\text{for all } x\in M. \tag{20}$$

In that case, the form $\omega_{\mathfrak{H},a}$ will again be a symplectic form.

**Lemma 5.3** *Let $\Sigma$ be a Riemann surface with cylindrical ends and let $(\mathfrak{H}, J)$ be a domain-dependent Floer datum on $\Sigma$. For any $(\mathfrak{H}, J)$–Floer solution $u\colon \Sigma \to M$ satisfying* (17) *and* (18), *and for any Borel subset $A \subset \Sigma$, we have*

$$E(u; A) := \int_A \|du - X_{\mathfrak{H}}\|^2 \, d\mathrm{vol}_\Sigma \le E_{\mathrm{top}}(u; A). \tag{21}$$

**Proof** Write in local coordinates $\mathfrak{H} = H\,dt + G\,ds$. Then using the Floer equation and denoting by $d'$ the exterior derivative in the $M$ direction,

$$\begin{aligned} \|du - X_{\mathfrak{H}}\|^2 \, ds \wedge dt &= \omega(\partial_t u - X_H, X_G - \partial_s u)\, ds \wedge dt \\ &= u^*\omega + (d'H(\partial_s u) + d'G(\partial_t u))\, ds \wedge dt \\ &= u^*\omega + d\mathfrak{H} - (\partial_s H \circ u - \partial_t G \circ u)\, ds \wedge dt \\ &\le u^*\omega + d\mathfrak{H}. \end{aligned}$$

We have used the conformal invariance of energy,

$$\|du - X_{\mathfrak{H}}\|^2 \, d\mathrm{vol}_\Sigma = \|du - X_{\mathfrak{H}}\|^2 \, ds \wedge dt. \qquad \square$$

Henceforth, we shall denote by $g_{J_F}$ the Riemannian metric determined by $\omega_{\mathfrak{H}}$ and $J_F$ and refer to it as the *Gromov metric*. When $\mathfrak{H} = H\,dt$ we will also use the notation $J_H$ and $g_{J_H}$.

**Example 5.4** Let $\mathfrak{H} = H\,dt$, where $H\colon M \to \mathbb{R}$ is smooth. Then one finds by a straightforward computation that

$$g_{J_H} = \pi_1^* g_j + \pi_2^* g_J + g_{\mathrm{mixed}}, \tag{22}$$

where $\pi_i$ are the natural projections and

$$g_{\mathrm{mixed}} = -g_J(X_H, \cdot)\, dt - dt\, g_J(X_H, \cdot) + \|X_H\|^2 dt^2. \tag{23}$$

In order to define the notion of i–boundedness for Floer data we need a relative notion of intermittent boundedness.

**Definition 5.5** Let $\Sigma$ be a Riemann surface with cylindrical ends. A Riemannian metric on $\Sigma \times M$ is said to be *intermittently bounded relative to the projection $\pi : \Sigma \times M \to M$* if there is an exhaustion of $\Sigma \times M$ by a sequence of open sets $K_i$ such that for any precompact open $U \subset \Sigma$, the sets $\pi^{-1}(U \cap K_i)$ are precompact, and such that the rest of Definition 4.1 holds for these $K_i$. Let $\widetilde{\omega}$ be a symplectic form on $\Sigma \times M$.

An $\widetilde{\omega}$–compatible almost complex structure $J$ on $\Sigma \times M$ is said to be intermittently bounded relative to $\pi$ if the associated metric $g_J$ is. We denote the set of all these by $\mathcal{J}_{\mathrm{ib}}(\Sigma \times M, \widetilde{\omega}, \pi)$. For an open set $U \subset \Sigma$ we denote by $\mathcal{J}_{\mathrm{ib}}(U \times M, \widetilde{\omega})$ the set of $\widetilde{\omega}$–compatible almost complex structures on $U \times M$ that are restrictions of a $J \in \mathcal{J}_{\mathrm{ib}}(\Sigma \times M, \widetilde{\omega}, \pi)$.

The following lemma is an obvious variant of Theorem 4.11(b), the only difference being the need to restrict to $J$–holomorphic sections.

**Lemma 5.6** *Let $U \subset \Sigma$ be an open precompact subset. Let $J \in \mathcal{J}_{\mathrm{ib}}(U \times M, \widetilde{\omega}, \pi)$. Suppose $\|d\pi\|$ is uniformly bounded from above with respect to some fixed conformal metric on $\Sigma$. For any compact set $K \subset M$, any compact subset $S \subset U$, and any $E \in \mathbb{R}_+$, there exists an $R$ such that for any $J$–holomorphic section*

$$u : U \to U \times M$$

*satisfying $E(u; U) \leq E$ and $u(S) \cap K \neq \varnothing$, we have $u(S) \subset B_R(S \times K)$.*

**Proof** The assumption on $\|d\pi\|$ guarantees that for any $z \in \Sigma$ we have $B_\epsilon(u(z)) \subset u(B_\epsilon(z))$. The argument is then word for word that of Theorem 4.11(b) . □

**Remark 5.7** The dependence of $R$ on $U$ and $S$ is spelled out in Remark 4.13.

**Lemma 5.8** *For the Gromov metric $g_{J_F}$ associated with $F$ any Floer datum satisfying* (17) *and* (18)*, we have $\|d\pi\| \leq 1$.*

**Proof** For any vector $v$ tangent to $\Sigma \times M$ we have $\|v\| = \pi^*\omega_\Sigma(v, J_F v) + \omega_F(v, J_F v)$. The second term is nonnegative by Lemma 5.1 since Lemma 5.1 holds for *any* choice of $\omega_\Sigma$. The first term equals $\omega_\Sigma(\pi_* v, j_\Sigma \pi_* v)$ by holomorphicity of $\pi$. □

**Definition 5.9** Let $\Sigma$ be a Riemann surface with cylindrical ends. A domain-dependent Floer datum $(\mathfrak{H}, J)$ on $\Sigma$ is called *i–bounded* if:

(a) $\mathfrak{H}$ satisfies (17) and (18) (or, more generally, inequalities (20) and (19)).

(b) There exists a finite open cover $C$ of $\Sigma$ such that for each $U \in C$ we have $J_{\mathfrak{H}}|_{U \times M} \in \mathcal{J}_{\mathrm{ib}}(U \times M, \omega_{\mathfrak{H}})$ (or, more generally, $J_{\mathfrak{H}} \in \mathcal{J}_{\mathrm{ib}}(U \times M, \omega_{\mathfrak{H},a})$).

**Definition 5.10** Let $\mathcal{S}$ be a compact manifold with corners. A smooth family $\Sigma_{\{s \in \mathcal{S}\}}$ of (broken) Riemann surfaces with cylindrical ends together with a smooth choice of domain-dependent i–bounded Floer data $(\mathfrak{H}_s, J_s)$ is called *admissible* if the following holds. Denote by $\pi : \widetilde{\mathcal{S}} \to \mathcal{S}$ the tautological bundle. Then we assume there is a smooth choice $\sigma_s$ of area forms on $\Sigma_s$ and a finite cover of $\widetilde{\mathcal{S}}$ by connected opens consisting of elements of two types: $\mathrm{Thick}_{\mathcal{S}}$ and $\mathrm{Thin}_{\mathcal{S}}$. The elements of $\mathrm{Thick}_{\mathcal{S}}$ are assumed to be subsets of $\widetilde{\mathcal{S}}$ which are trivializable to the form $U = V \times W$, where $W \subset \mathcal{S}$ and $V$ is a bordered Riemann surface whose area is uniformly bounded on $W$. The fibers of $\pi$ restricted to elements of $\mathrm{Thin}_{\mathcal{S}}$ are generically cylinders (of finite, half-infinite or infinite length), which may degenerate to nodes at the corners. Moreover, for the thin elements we require that the Floer data be translation invariant on the fibers of $\pi$ and that the area forms coincide with $ds \wedge dt$. We say that the family $\mathcal{S}$ is *uniformly i–bounded* if for any thick element $U = V \times W$ there exist taming data on $V \times M$ which are constant on $W$, and for any thin element $U$ there exist taming data on $[-1, 1] \times \mathbb{R}/\mathbb{Z} \times M$ which are constant on $\pi(U)$.

For the rest of the section we wish to establish criteria for i–boundedness of $J_F$. This is not strictly necessary for the proof of the main theorems in the introduction. Lemma 5.11, to be stated presently, is all we shall need for that purpose. The proof is left to the reader.

**Lemma 5.11** *Let $(H_1, J)$ be a Floer datum, and $H_2$ a time-dependent Hamiltonian such that $\|X_{H_2}\| \le C$ for some constant $C$. Then $g_{H_1+H_2}$ is $C^2$–quasi-isometric to $g_{H_1}$. In particular, when $J$ is i–bounded and $H$ is such that $\|X_H\|$ is bounded, we have that $J_H$ is i–bounded.*

However, for applications in practice we need effective criteria. For example, we need to show that Floer data that has been hitherto used in the literature fits into the dissipative framework. To do this we need, first of all, a criterion for completeness of the metric $g_{J_F}$. Then we need to discuss how to compute the curvature of $g_{J_F}$ and control its radius of injectivity in terms of the Floer data $J$ and $\mathfrak{H}$. We do this in the case where $\mathfrak{H} = H\,dt$ for a time-independent Hamiltonian $H$ as in Example 5.4. Since intermittent boundedness is preserved under quasi-isometry, this is quite sufficient for applications insofar as Floer trajectories are concerned. The consideration of more general Floer solutions will be reduced to that of Floer trajectories.

## 5.3 Completeness

**Definition 5.12** Let $J$ be an almost complex structure. We say that an exhaustion function $H : M \to \mathbb{R}$ is *$J$–proper* if $H$ factors as $H = f \circ h$ for some proper smooth function $h : M \to \mathbb{R}$ satisfying $\|\nabla h\| \le 1$ with respect to the metric $g_J$ on $M$.

**Lemma 5.13** *Let $H$ be a smooth time-independent $J$–proper Hamiltonian, ie $H = f \circ h$ with $\|\nabla h\| \le 1$. For any function $g\colon [a,b] \to \mathbb{R}$ and any $\gamma\colon [a,b] \to M$, we have*

$$|h(\gamma(b)) - h(\gamma(a))|^2 \le (b-a) \int_a^b \|g(t)X_H - \gamma'(t)\|^2 \, dt. \tag{24}$$

**Remark 5.14** More generally, if $H$ is time-dependent, and factors as $H_t = f \circ h_t$ where $h_t$ is a smooth proper time-dependent function satisfying $|\nabla h_t| \le 1$, we have

$$\begin{aligned}(25)\quad & |h_b(\gamma(b)) - h_a(\gamma(a))|^2 \\ &\le (b-a)\left( \int_a^b \|g(t)X_H - \gamma'(t)\|^2 \, dt + \sup_{t\in[a,b]} \partial_t h_t \circ \gamma(t)\right).\end{aligned}$$

**Proof** We have

$$\begin{aligned} |h(\gamma(b)) - h(\gamma(a))|^2 &= \left| \int_a^b \langle \nabla h, \gamma'(t)\rangle \, dt \right|^2 \\ &= \left| \int_a^b \langle \nabla h, \gamma'(t) - g(t)X_H\rangle \, dt \right|^2 \\ &\le (b-a) \int_a^b \|g(t)X_H - \gamma'(t)\|^2 \, dt. \end{aligned}$$

We used Cauchy–Schwarz, $\|\nabla h\| \le 1$, and the fact that $X_H \perp \nabla h$. □

**Lemma 5.15** *Suppose $H$ is smooth time-independent and $J$–proper. Then the metric $g_{J_H}$ on $\widetilde{M} := \mathbb{R} \times \mathbb{R}/\mathbb{Z} \times M$ is complete.*

**Proof** Let $H = f \circ h$, where $h\colon M \to \mathbb{R}$ is proper and satisfies $\|\nabla h\| \le 1$. We show that the pullback $\widetilde{h}$ of $h$ to $\widetilde{M}$ is Lipschitz with respect to $g_{J_H}$. It suffices to show that for any path $\widetilde{\gamma}\colon [a,b] \to \widetilde{M}$ lifting a path $\gamma\colon [a,b] \to M$, we have

$$|\widetilde{h}(\widetilde{\gamma}(b)) - \widetilde{h}(\widetilde{\gamma}(a))|^2 \le (b-a) \int \|\widetilde{\gamma}'\|^2_{g_{J_H}}.$$

For each $x \in [a,b]$ we can $g_{H_J}$–orthogonally decompose

$$\widetilde{\gamma}'(x) = v(x) + g(x)(X_H + \partial_t) + \partial_s,$$

where $v(x) \in TM$. We have

$$\|\widetilde{\gamma}'(x)\|^2_{g_{H_J}} \ge \|v(x)\|^2 = \|\gamma'(x) - g(x)X_H\|^2_{g_J}.$$

Since $\widetilde{h}$ is independent of $s$, the claim follows by Lemma 5.13.

To see that $g_{J_H}$ is complete note first that by translation invariance it suffices to prove completeness of the restriction of $g_{J_H}$ to the mapping torus $s = \text{const}$. For this, note that the restriction of $\widetilde{h}$ to the set $s = \text{const}$ is still Lipschitz and, moreover, it is proper since $H$ is. Thus for any $x$ we have that the ball of radius $R$ around $x$ in $\mathbb{R}/\mathbb{Z} \times M$ is contained in the compact subset

$$\widetilde{h}^{-1}([\widetilde{h}(x) - R, \widetilde{h}(x) + R]).$$

Completeness now follows by Hopf–Rinow. $\square$

We conclude with a criterion for $J$–properness. Call a function $f : \mathbb{R} \to [1, \infty)$ tame if the primitive of $1/f$ is unbounded from above.

**Lemma 5.16** *Suppose there is a tame function such that*

$$\|\nabla H\|_{g_J} \leq f(H).$$

*Then $H$ is $J$–proper.*

**Proof** Let $g$ be a primitive of $1/f$. We have

$$\|\nabla(g \circ H)\| = g' \circ H \|\nabla H\| = \frac{1}{f \circ H} \|\nabla H\| \leq 1.$$

By assumption, $h := g \circ H$ is proper. Moreover, $g$ is monotone (primitive of a positive function). So $H = g^{-1} \circ h$. $\square$

## 5.4 Curvature

We introduce some notation and recall some basic formulae in Riemannian geometry. We refer to [45] for details. Let $(M, g)$ be a Riemannian manifold and let $r : M \to \mathbb{R}$ be a distance function; that is, a function satisfying $\|\nabla r\| = 1$. Write $\partial_r := \nabla r$ and denote by $S$ the tensor $\nabla \partial_r$. Denote by $U_r$ the level sets of $r$. Denote by $R$ the curvature tensor of $M$, by $R^t$ the tangential component of $R$ restricted to $TU_r$ and by $R^r$ the curvature tensor of $U_r$. Also, write $R_{\partial_r} = R(\cdot, \partial_r)\partial_r$.

The following formulae, together with the symmetries of the curvature tensor, show that the full curvature tensor on $M$ is determined by the curvature of the level sets of $r$, by the tensor $S$ and by its first derivative:

$$-R_{\partial_r} = S^2 + \nabla_{\partial_r} S, \tag{26}$$

$$R^t(X, Y)Z = R^r(X, Y)Z - S(X) \wedge S(Y)Z, \tag{27}$$

$$R(X, Y)\partial_r = (\nabla_X S)(Y) - (\nabla_Y S)(X). \tag{28}$$

The vectors $X$, $Y$ and $Z$ in the above formulae are all tangent to $U_r$. In what follows, given a vector $V \in TM$ we will use the notation $\theta_{g,V}$ for the dual to $V$ with respect to $g$ and will drop $g$ from the notation when there is no ambiguity. We utilize the following formula for the covariant derivative of a vector field $X$:

$$2\theta_{g,\nabla X} = d\theta_{g,X} + \mathcal{L}_X g. \tag{29}$$

This formula presents the decomposition of $\theta_{g,\nabla X}$ into a symmetric and an antisymmetric bilinear form. For a proof see [45, page 26].

Let $\mathfrak{H} = H\,dt$, where $H\colon M \to \mathbb{R}$ is smooth. Since $g_{J_H}$ is translation invariant with respect to $s$, we restrict attention to submanifolds of $\mathbb{R} \times \mathbb{R}/\mathbb{Z} \times M$ with fixed values of $s$, or, in other words, to $\mathbb{R}/\mathbb{Z} \times M$ with the metric $g_{J_H}$ as computed in Example 5.4. The function $t$ (which is locally single-valued) is a distance function on $\mathbb{R}/\mathbb{Z} \times M$ with respect to this metric. To see this, note that by (22) we have $dt = g_{J_H}(X_H + \partial/\partial t, \cdot)$. That is, $\nabla t = X_H + \partial/\partial t$. One verifies that $\|X_H + \partial_t\|^2_{g_{J_H}} = 1$.

**Theorem 5.17** *We have $\nabla\nabla t = \frac{1}{2}(\nabla^{g_J} X_H + \nabla^{g_J} X_H^*) \circ \pi$, where the superscript denotes conjugation with respect to the metric $g_J$ and $\pi\colon T(\mathbb{R}/\mathbb{Z} \times M) \to TM$ is the $g_{J_{\mathfrak{H}}}$ orthogonal projection.*

**Proof** Write $N = \nabla t$. By equation (29) we have

$$2\theta_{\nabla N} = d\theta_N + \mathcal{L}_N g_{J_F}.$$

Since $\theta_N = dt$, we have $d\theta_N = 0$. We claim that $\mathcal{L}_N g_{J_F} = \pi^* \mathcal{L}_{X_H} g_J$. To see this, denote by $\psi_t$ the time $t$ flow of $X_H$ and let

$$\phi\colon (-\epsilon, \epsilon) \times M \subset \mathbb{R}/\mathbb{Z} \times M \to \mathbb{R}/\mathbb{Z} \times M$$

be the map $(t, p) \mapsto (t, \psi_t(p))$. Then

$$\phi_*|_{T(\{t_0\} \times M)} = \psi_{t_0,*} \quad \text{and} \quad \phi_* \partial_t = \partial_t + X_H = N.$$

In particular, $\phi^* g_{J_F}|_{\mathbb{R}/\mathbb{Z} \times M} = \pi^* \psi_t^* g_J + dt^2$. Thus,

$$\begin{aligned} \phi^* \mathcal{L}_N g_{J_F} &= \mathcal{L}_{\partial_t} \phi^* g_{J_F} \\ &= \partial_t (\pi^* \psi_t^* g_J + dt^2) \\ &= \pi^* \psi_t^* \mathcal{L}_{X_H} g_J \\ &= \phi^* \pi^* \mathcal{L}_{X_H} g_J. \end{aligned} \tag{30}$$

By (29) we have

$$\mathcal{L}_{X_H} g_J = [\theta_{\nabla X_H, g_J}],$$

where $[\alpha(\cdot,\cdot)]$ denotes the symmetrization. Thus,

$$S = \tfrac{1}{2}(\nabla^{g_J} X_H + \nabla^{g_J} X_H^*) \circ \pi. \qquad \square$$

We say that a Hamiltonian $H\colon M \to \mathbb{R}$ is Killing (with respect to some compatible almost complex structure $J$) if the flow of $X_H$ preserves $g_J$.

**Corollary 5.18** *Suppose $H$ is Killing. Then $\nabla\nabla t \equiv 0$.*

## 5.5 Injectivity radius

We turn to discussing the control of the radius of injectivity of $g_{J_H}$. In the following lemmas fix a point $x_0 = (s_0, t_0, p_0) \in \mathbb{R} \times \mathbb{R}/\mathbb{Z} \times M$.

**Lemma 5.19** *For any $r < \frac{1}{2}$ we have*

$$\operatorname{Vol}_{g_{J_H}}(B_r(x_0)) > \tfrac{1}{9} r^2 \operatorname{Vol}_{g_J}(B_{r/3}(p_0)).$$

**Proof** Denote by $\psi_t$ the Hamiltonian flow of $H$. Since $X_H + \partial_t$ is perpendicular, with respect to $g_{J_H}$, to hypersurfaces of constant $t$, we have that $B_r(x_0)$ contains the cylinder

$$C = \bigcup_{t \in [t_0 - r/3, t_0 + r/3]} \left[s_0 - \tfrac{1}{3}r, s_0 + \tfrac{1}{3}r\right] \times \{t\} \times \psi_t(B_{r/3}(p_0)).$$

Since $\psi_t$ preserves the $g_J$–volume we have

$$\operatorname{Vol}_{g_{J_H}}(C) = \tfrac{1}{9} r^2 \operatorname{Vol}_{g_J}(B_{r/3}(p_0)). \qquad \square$$

**Lemma 5.20** *Let $(M, g)$ be an $n$–dimensional Riemannian manifold. Let $a > 0$ and let $p \in M$ be such that*

$$\operatorname{Vol}_g\big(B_{1/a}(p)\big) \ge v_0 \Big(\frac{1}{a}\Big)^n,$$

*and such that $|\mathrm{Sec}_g(x)| \le a^2$ on $B_{1/a}(p)$. Then there is a constant $f = f(v_0, n)$, independent of $a$, such that $\operatorname{inj}_g(p) \ge f(v_0, n)$.*

**Proof** For $a = 1$, this is an immediate consequence of [12, Theorem 4.3]. The claim follows by scaling. $\square$

**Lemma 5.21** *Suppose $(M, g)$ satisfies, for some $p \in M$, that*

$$\mathrm{inj}_g(p) \geq a \quad \text{and} \quad |\mathrm{Sec}_g(x)| \leq a^2$$

*on $B_{1/a}(p)$. Then there is a constant $C = C(n) > 0$ such that*

$$\mathrm{Vol}_g(B_{1/a}(p)) \geq C\left(\frac{1}{a}\right)^n.$$

**Proof** By scaling, the claim is equivalent to the claim that there is a constant $C(n) > 0$ such that a geodesic ball of radius 1 with sectional curvature bounded by 1 has volume at least $C(n)$. By the Jacobi equation, sectional curvature controls the derivatives of the metric in geodesic coordinates [10, Chapter 5]. In particular there is an a priori estimate from below on the determinant of the metric in these coordinates for a small enough ball around the origin. The claim follows. □

**Theorem 5.22** *There is a constant $i = i(n)$ such that if $g_J$ is $a$–bounded at $p_0 \in M$, then $\mathrm{inj}_{g_{J_H}}(x) \geq i(n)/a$.*

**Proof** Combining Lemmas 5.19 and 5.21 we have that there is a constant such that

$$\mathrm{Vol}_{g_{J_H}}(B_{1/a}(x)) \geq \frac{1}{3^{n+2}} C(n)\left(\frac{1}{a}\right)^{n+2}.$$

The claim follows by Lemma 5.20. □

## 5.6 Some criteria for boundedness

**Lemma 5.23** *Suppose $g_J$ is $a$–bounded at $p \in M$ and $H$ is a time-independent Hamiltonian such that*

$$\max\{\|\nabla X_H(p)\|^2, \|\nabla^2 X_H\|, \|\nabla_{X_H}(\nabla X_H + \nabla X_H^T)\|\} < a^2. \tag{31}$$

*Then for a constant $c = c(n)$ independent of $a$, we have that $g_{J_H}$ is $ca$–bounded at $p$.*

**Proof** We need to estimate the sectional curvature and radius of injectivity of $g_{J_H}$. Up to multiplication by a constant dependent on $n$, estimating sectional curvature is the same as estimating the coefficients of the curvature tensor in an orthonormal basis. Since $J$ is $a$–bounded, it remains to estimate only coefficients involving the direction $\partial_t + X_H$ at least once. In light of formulae (26)–(28) we need to estimate $\nabla S$ and $S^2$, where $S = \nabla t$. Theorem 5.17 provides us with an estimate on $S^2$ and the tangential restriction of $\nabla S$ in terms of $\nabla X_H$ and $\nabla^2 X_H$. It remains to estimate the right-hand side of (26). For this it is preferable to use the formula

$$-R_N = \mathcal{L}_N S - S^2.$$

See [45]. Each summand vanishes on $N$. So it remains to estimate $\mathcal{L}_N S$ applied to a tangential vector. Let $V$ be a tangential vector field which commutes with $N$. Then

$$\begin{aligned}\mathcal{L}_{X_H+\partial_t}(SV) = \mathcal{L}_{X_H}(SV) &= \nabla^{g_J}_{X_H}(SV) - \nabla^{g_J}_{SV}(X_H)\\ &= (\nabla^{g_J}_{X_H}S)V + S(\nabla^{g_J}_V X_H) - \nabla^{g_J}_{SV}(X_H).\end{aligned}$$

This shows that estimate (31) implies $\mathrm{Sec}_{g_{J_H}}(p) \le c^2a^2$ for an appropriate $c = c(n)$. Theorem 5.22 provides us with the estimate on $\mathrm{inj}_{g_{J_H}}$ in terms of $\mathrm{inj}_{g_J}(p)$. The claim follows. □

**Example 5.24** Let $(\Sigma, \alpha)$ be a contact manifold and let

$$(M = \mathbb{R}_+ \times \Sigma, \omega = e^r(d\alpha + dr \wedge \alpha))$$

be the convex end of its symplectization. Let $R$ be the Reeb vector field on $\Sigma$. Fix an $\omega$–compatible translation-invariant almost complex structure $J$ satisfying $JR = \partial r$. Then

$$g_J = e^r(dr^2 + g_\Sigma) \tag{32}$$

for some metric $g_\Sigma$ on $\Sigma$. Since the metric $g_J$ scales up, the radius of injectivity of $g_J$ is bounded away from 0, and in fact goes to $\infty$ with $r$. Pick local coordinates on $\Sigma$ and use the function $r$ as the coordinate on the $\mathbb{R}_+$ factor. Then the Christoffel symbols of the metric (32) are $O(1)$ in these coordinates. Therefore

$$\langle \nabla_{\partial_i}\partial j, \partial_k\rangle \sim e^r.$$

Since $\|\partial_i\|^2 \sim e^r$, for some constant $C$ we have $\|\nabla\partial_i\| \le C$. Similarly,

$$\|\nabla^2_{ij}\partial_k\|^2 \sim e^r,$$

allowing us to deduce that

$$\|\nabla^2\partial_k\| \sim e^{-r/2}.$$

Suppose $H$ is a function on the symplectization which is given outside of a compact set by $H = h(e^r)$. Then there are some constants $a_i$ such that

$$X_H = h'(e^r)\sum a_i\partial_i.$$

First suppose $h'(e^r)$ is constant. Then by the reasoning above, we conclude that *the induced metric $g_{J_H}$ is uniformly bounded for Hamiltonians which are linear at infinity with a bound that is proportional to the slope $h'(e^r)$.*

**Example 5.25** Continuing with the previous example, assume now that $h'(e^r)$ is at most linear in the distance from $\Sigma$. *Caution*: *this means it is at most linear in $e^{r/2}$.* Then there is a bound on the geometry of $g_{J_H}$, which is also linear in the distance.

To see this note that for a point $p$ which is a distance $d$ from $\Sigma$, the metric $g_{J_H}$ is uniformly equivalent on a neighborhood of size 1 to the metric associated with the slope $h'(e^{r(p)})$. It follows by Example 4.3 that the metric associated to $H$ is i–bounded. Note that while this allows superlinear Hamiltonians, it does not include quadratic Hamiltonians $h' \sim e^r$.

**Example 5.26** Consider the cotangent bundle $T^*M$ of a compact manifold $M$, let $g$ be a Riemannian metric on $M$ and let $J$ be the Sasaki almost complex structure on $T^*M$. It is defined as follows: the Levi-Civita connection on $T^*M$ induces a splitting $TT^*M = V \oplus H$ into horizontal and vertical vectors. Moreover, we take $J : V \simeq H$ to be the natural isomorphism identifying an element of $V$ with an element of $T^*M$, then via $\omega$ with an element of $TM$ and finally with an element of $H$ via horizontal lifting. Identifying $TM = T^*M$, in standard local Darboux coordinates $\{q_i, p_i = dq_i\}$, where $q_i$ are geodesic coordinates centered at a point $q$, $J$ is given in the fiber over $q$ by

$$J\frac{\partial}{\partial p_i} = \frac{\partial}{\partial q_i}.$$

Then it is easy to show that the metric $g_J$ is $\|p\|$–bounded at the point $(p,q)$. In particular, $J$ is i–bounded (but not bounded). Consider a Hamiltonian of the form $H = \sqrt{a|p|^2 + V \circ \pi}$, where $\pi : T^*M \to M$ is the standard projection and $V : M \to \mathbb{R}$ is smooth. Then $J_H$ is i–bounded. Indeed, denoting by $M$ the maximum of $\sqrt{|V|}$ over $M$, we have in local coordinates as above,

$$\|X_H\| = aM\frac{1}{2\|p\|}\left\|\sum_i p_i \frac{\partial}{\partial q_i}\right\| \le aM.$$

So, the claim follows from Lemma 5.11. Note that mechanical Hamiltonians of the form $|p|^2 + V \circ \pi$ are not i–bounded with respect to the Sasaki metric.

# 6 Loopwise dissipativity

## 6.1 Diameter control of Floer trajectories

Suppose $(H, J)$ is i–bounded, let $u$ be a Floer trajectory, and let $\widetilde{u}$ be its graph. Suppose that for some precompact $U \subset \mathbb{R} \times \mathbb{R}/\mathbb{Z}$, we have control over $u(\partial U)$. Theorem 4.11 above then provides control over $u(U)$ in terms of

$$E(\widetilde{u}; U) = E(u; U) + \mathrm{Area}(U).$$

This indicates that the only source of noncompactness in the moduli space of finite-energy Floer trajectories comes from the potential existence of finite-energy solutions with one end converging to infinity. This motivates the following definition.

We refer henceforth to a Floer solution on a possibly finite cylinder $[a,b]\times\mathbb{R}/\mathbb{Z}$ as a *partial Floer trajectory*. For $H\colon \mathbb{R}/\mathbb{Z}\times M\to\mathbb{R}$ proper and an $\omega$–compatible almost complex structure, define a function $\Gamma_{H,J}(r_1,r_2)$ as the infimum over all $E$ for which there is a partial Floer trajectory $u$ of energy $E$ with one end of $u\times t$ contained in $H^{-1}([-r_1,r_1])$ and the other end in $H^{-1}(\mathbb{R}\setminus(-r_2,r_2))$. Note that $\Gamma_{H,J}(r_1,r_2)$ may take the value of infinity.

**Definition 6.1** We say that $(H,J)$ is *loopwise dissipative* (*LD*) if for any fixed $r_1$ we have $\Gamma_{H,J}(r_1,r)\to\infty$ as $r\to\infty$. If this holds for some function $\Gamma\colon\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}_+\cup\{\infty\}$ satisfying $\Gamma_{H,J}\geq\Gamma$, we say that $(H,J)$ is $\Gamma$–LD. We say that $(H,J)$ is *robustly loopwise dissipative* (*RLD*) if there is a function $\Gamma\colon\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ and an open neighborhood[8] of $(H,J)$ in $C^1\times C^0$, all elements of which are $\Gamma$–LD.

**Definition 6.2** Denote by $\mathcal{F}^{(0)}_d(M)$ the set of i–bounded Floer data $(H,J)$ which are RLD. Elements of $\mathcal{F}^{(0)}_d(M)$ are referred to as *dissipative Floer data*.

Our next theorem shows that dissipativity is all we need for diameter control. In the ensuing sections we show both that on a geometrically bounded manifold there is always a sufficient supply of dissipative Floer data, and that this property can be verified directly in various settings.

In the following theorem, recall Definitions 5.9 and 5.10 of an i–bounded Floer datum and family of Floer data.

**Theorem 6.3** *Let $(\mathcal{S},F_{s\in\mathcal{S}}=(\mathfrak{H}_s,J_s))$ be a uniformly i–bounded family of connected (broken) Riemann surfaces decorated with Floer data and equipped with a thick–thin decomposition as in Definition 5.10. Let $(H_i,J_i)\in\mathcal{F}^{(0)}_d$ be Floer data such that on the $i^{th}$ component of* $\mathrm{Thin}_{\mathcal{S}}$, *we have that $F_s$ coincides with $(H_i,J_i)$ for all $s\in\mathcal{S}$. Then for any compact $K\subset M$ and any real number $E$, there is an $R=R(E,K)$ such that for any $s\in\mathcal{S}$, any $F_s$–Floer solution $(\Sigma_s,u)$ with $E(u)\leq E$ and intersecting $\partial K$ is contained in $B_R(K)$. Moreover, if $K$ is a level set of $H$ with no degenerate periodic orbits in a neighborhood of $\partial K$, we can take $R(E,K)\to 0$ as $E\to 0$.*

[8] Here and hereafter the topology can be taken to be the uniform topology with respect to $g_J$. However what is truly necessary is that open sets are sufficiently thick to allow perturbations for achieving regularity.

**Definition 6.4** We refer to families $(F_{s\in\mathcal{S}} = (\mathfrak{H}_s, J_s))$ satisfying the hypotheses of Theorem 6.3 as *dissipative families*. When $\mathcal{S}$ consists of a single element, we refer to it as a *dissipative Floer datum*.

**Proof of Theorem 6.3** By assumption, we can decompose $\Sigma_s$ into a thick part consisting of components $A_i$ with area bounded by a constant $C_{\mathcal{S}}$, independently of $s\in\mathcal{S}$, and a thin part consisting of components $B_i$ on which the Floer data are translation invariant and given by $(H_i, J_i)$. Moreover, there is a number $N_{\mathcal{S}}$ which bounds the number of components independently of $s\in\mathcal{S}$. For each $A_i$, the graph $\widetilde{u} = \mathrm{Id}\times u : A_i \to A_i\times M$ is $J_F$–holomorphic and satisfies $E(\widetilde{u}) \le E + \mathrm{Area}(A_i)$. Furthermore we may assume for some $\epsilon > 0$ that we have $J_F|_{B_\epsilon(A_i)\times M} \in \mathcal{J}_{\mathrm{ib}}(B_\epsilon(A_i)\times M, \omega_{\mathfrak{H}})$.

We construct an $R_0 = R_0(E, K)$ such that if $u(A_i)\cap K \ne \varnothing$ for some component $A_i$, then $u(A_i) \subset B_{R_0}(K)$ and moreover for all components $B_j$ which share a boundary with $A_i$, we have $u(B_j) \subset B_{R_0}(K)$. A similar $R_0$ can be constructed starting with a component $B_i$. Since the total number of components is bounded by $N_{\mathcal{S}}$ this will inductively give rise to an $R$ as in the statement of the present theorem.

To construct $R_0$, suppose $u(A_i)$ intersects some compact set $K$. Then, since $\widetilde{u}$ extends to a neighborhood of $A_i$, by Lemma 5.6 we have that $\widetilde{u}(A_i)$ is contained in a ball $B_{\widetilde{R}}(K\times A_i)$ for some $\widetilde{R} = \widetilde{R}(K, E + 2C_{\mathcal{S}})$ depending additionally on the taming data associated with $F_s$ and thus on $\mathcal{S}$. From this we deduce the same for $u(A_i)$, with perhaps a different radius $R$. It follows that each of the components $B_j$ whose closure intersects $A_i$ has a boundary component contained in $B_R(K)$. Let $a_j$ be such that $B_R(K) \subset H_j^{-1}([-a_j, a_j])$. By loopwise dissipativity there is a $b_j > a_j$ such that $\Gamma_{H_j,J_j}(a_j, b_j) > E$. Writing $B_j = I\times\mathbb{R}/\mathbb{Z}$ for some interval $I$, we have, by definition of $\Gamma_{H_j,J_j}$, that $u(\{s\}\times\mathbb{R}/\mathbb{Z})$ intersects $H^{-1}(-b_j, b_j)$ for each $s\in I$. Restricting $u$ to $(s-1, s+1)\times\mathbb{R}/\mathbb{Z}$ and invoking Lemma 5.6 again, we obtain an $R'$ such that for any $s\in I$, we have $u((s-1,s+1)\times\mathbb{R}/\mathbb{Z}) \subset B_{R'}(H^{-1}(-b_j, b_j))$. It follows that the same holds for $u(B_j)$. Now take $R_0$ such that $B_{R'}(H^{-1}(-b_j, b_j)) \subset B_{R_0}(K)$ for each $j$.

For the last statement of the theorem we rely on the following property of Floer trajectories, which is stated in [52]. There are constants $c$ and $\hbar$ such that

$$\int_{B_r(s,t)} \|\partial_s u\|^2 < \hbar \implies \|\partial_s u\|^2(s,t) < \frac{8}{\pi r^2}\int_{B_r(s,t)} \|\partial_s u\|^2 + cr^2.$$

Once we know that a solution is contained in an a priori compact set, we can take all the constants to be fixed by that compact set. By taking $r = E(u)^{1/4}$ we deduce that for an appropriate constant,

$$\|\partial_t u - X_H\|^2 = \|\partial_s u\|^2 < CE(u)^{1/2}$$

once $E(u)$ is small enough. It follows that making $E(u)$ arbitrarily small, $u$ will be contained in an arbitrarily small neighborhood of some periodic orbit. □

We conclude this subsection with a counterexample showing that geometric boundedness alone does not guarantee loopwise dissipativity.

**Example 6.5** Consider $(M,\omega) = (\mathbb{R} \times \mathbb{R}/\mathbb{Z}, ds \wedge dt)$. Let $H$ be a smoothing of the function

$$(s,t) \mapsto s - \ln(|s|+1),$$

and let $J$ be multiplication by $i$. Then $\|X_H\|$ is bounded, so $(H,J)$ is i–bounded by Lemma 5.11. But it is not LD. Indeed, the map $u\colon \mathbb{R}_+ \times \mathbb{R}/\mathbb{Z} \to M$ defined by $u(s,t) = (\ln(s+1), t)$ is an $(H,J)$–partial Floer trajectory of finite energy and infinite diameter.

## 6.2 Hamiltonians with small Lipschitz constant

**Theorem 6.6** *Let $J$ be a geometrically bounded almost complex structure compatible with $\omega$. There is an $\epsilon > 0$ such that for any Hamiltonian $H\colon M \to \mathbb{R}$ which is proper and satisfies, with respect to $g_J$, that $\|X_H\| < \epsilon$ outside of some compact set, the datum $(H,J)$ is dissipative. The claim remains true when $H$ is $C^0$–close to a time-independent Hamiltonian.*

The proof of Theorem 6.6 is carried out at the end of this section.

**Lemma 6.7** *Let $u\colon [a,b] \times \mathbb{R}/\mathbb{Z} \to M$ be a differentiable map. Then we have*

$$(b-a) \geq \frac{\displaystyle\int_{t\in\mathbb{R}/\mathbb{Z}} d^2(u(a,t),u(b,t))}{\displaystyle\int_{[a,b]\times\mathbb{R}/\mathbb{Z}} \|\partial_s u\|^2}.$$

**Proof** By the Cauchy–Schwarz inequality we have

$$\begin{aligned}(b-a)\int_{[a,b]\times\mathbb{R}/\mathbb{Z}} \|\partial_s u\|^2 &\geq \int_{t\in\mathbb{R}/\mathbb{Z}} \ell^2(u(t\times[a,b]))\,dt \\ &\geq \int_{t\in\mathbb{R}/\mathbb{Z}} d^2(u(a,t),u(b,t)).\end{aligned}$$ □

**Lemma 6.8** *Let $H\colon \mathbb{R}/\mathbb{Z}\times M\to\mathbb{R}$ be a proper smooth function. Suppose $J$ is a compatible almost complex structure. Suppose $H$ factors as $H=f\circ k$, where $k\colon \mathbb{R}/\mathbb{Z}\times M\to\mathbb{R}$ is proper and has uniformly bounded gradient with respect to $g_J$, and $f\colon\mathbb{R}\to\mathbb{R}$ is monotone. Then $(H,J)$ is LD if and only if there exists a sequence $h_i\to\infty$ and a constant $\delta>0$ such that*

$$\Gamma_{H,J}(h_{2i},h_{2i+1})>\delta. \tag{33}$$

*If $h_i$ and $\delta$ can be fixed for an open neighborhood of $(H,J)$, it is RLD.*

**Proof** The forward implication is obvious from the definition. For the other direction we use the following characterization of loopwise dissipativity:

*Let $K_i:=H^{-1}[-h_{2i},h_{2i}]\subset\mathbb{R}/\mathbb{Z}\times M$. For any $E\geq 0$ and for any natural $i$ there is an $i'(i,E)$ such that if $u$ is a partial solution with one end of $t\times u$ contained in $K_i$ and satisfying $E(u)<E$, then the other end intersects $K_{i+i'}$.*

For ease of exposition we assume for the rest of the proof that $H$ is time-independent, the general case being similar. We prove loopwise dissipativity by induction on the smallest integer $n$ bounding $E(u)/\delta$.

When $n=1$, this is just reformulating the assumption. Suppose we have proven the statement for all solutions $u$ satisfying $E(u)\leq n\delta$. Let $u$ be a solution with one end in $K_i$ and $E(u)\leq(n+1)\delta$. Without loss of generality we assume $u_a\subset K_i$. Here and henceforth $u_a:=u(a,\cdot)$. Let

$$s_1=\inf\{s\in[a,b]:u_s\subset M\setminus K_{i+1}\}.$$

If this set is empty there is nothing to prove. Otherwise, let

$$s_2=\inf\left(\{s\in[s_1,b]:\|\partial_s u\|<1\}\cup\{b\}\right).$$

Finally, take

$$s_0=\sup\left(\{s\in[a,s_1]:\|\partial_s u\|<1\}\cup\{a\}\right).$$

We have

$$E(u)\geq\int_{s_0}^{s_2}\|\partial_s u\|^2\,ds>\int_{s_0}^{s_2}ds=s_2-s_0.$$

So, by Lemma 6.7, there is a $t\in\mathbb{R}/\mathbb{Z}$ such that

$$d(u_{s_0}(t),u_{s_2}(t))<E. \tag{34}$$

We find an $i_0(i)$ such that $u_{s_0}\subset K_{i_0}$. Indeed, if $a=s_0$ there is nothing to prove. Otherwise, we have $\|\nabla\mathcal{A}_H(u_{s_0})\|\leq 1$. Since $s_0<s_1$, we have that $u_{s_0}$ intersects $K_{i+1}$.

Factor $H$ as $H = f \circ k$, as in the formulation of the present lemma. Since $f$ is monotone,

$$\min_t k_t(u_{s_0}(t)) < f^{-1}(h_{2(i+1)}) \quad \text{and} \quad \max_t k_t(u_{s_0}(t)) > f^{-1}(-h_{2(i+1)}). \tag{35}$$

From Lemma 5.13 we get an a priori estimate $c$ on the oscillation of $k$ on $u_{s_0}$ for the time-independent case. Here $c$ depends only on the bound on $|\nabla k|$. In the time-dependent case we appeal to (25) for this a priori estimate. Let $i_0$ satisfy

$$h_{i_0} \geq \lceil \max\{f(f^{-1}(h_{2(i+1)}) + c), -f(f^{-1}(-h_{2(i+1)}) - c)\} \rceil.$$

Combined with (35), this gives the a priori estimate

$$u_{s_0} \subset K_{i_0}.$$

By (34) we get from this an $i_1 = i_1(i, E)$ such that $u_{s_2}$ meets $K_{i_1}$. If $s_2 = b$, this concludes the proof. Otherwise, as for $s_1$, we find an $i_2$ such that $u_{s_2} \subset K_{i_2}$. We have $E(u|_{[s_2,b]\times\mathbb{R}/\mathbb{Z}}) \leq n\delta$ since $s_2 > s_1$ and by the hypothesis of the lemma $E(u|_{[a,s_1\times\mathbb{R}/\mathbb{Z}]}) > \delta$. So, by the inductive hypothesis, there is an $i_3$ depending on $i$ and $n$ such that $u_b$ meets $K_{i_3}$. The first part of the claim now follows. The second part is clear since $i'(i, E)$ is constructed using only the data of $\{K_i\}$ and $\delta$. $\square$

**Lemma 6.9** *Let $J$ be a geometrically bounded almost complex structure compatible with $\omega$. There are constants $R$, $\epsilon$ and $\delta$, depending on the bounds on the geometry of $g_J$, such that the following holds. Let $H: M \to \mathbb{R}$ be a proper Hamiltonian satisfying, for some $h \in \mathbb{R}$,*

$$\|X_H\| < \epsilon \quad \text{for all } x \in H^{-1}([h, h+R]). \tag{36}$$

*Then $\Gamma_{H,J}(h, h+R) > \delta$. This remains true if $H$ is merely assumed to be $C^0$–close to a time-independent Hamiltonian. Moreover, the estimate is unaffected if $H$ is arbitrarily time-dependent away from $H^{-1}([h, h+R])$.*

**Proof** We first prove the claim when the left-hand side of (36) is taken to hold for all $x \in M$. We consider the strictly time-independent case, leaving adjustments for the slightly more general case to the reader. For some $R > 0$ let $u: [a,b] \times \mathbb{R}/\mathbb{Z} \to M$ be a solution to Floer's equation with one end in $H^{-1}(-h, h)$ and the other end in $H^{-1}(\mathbb{R} \setminus [-h - R, h + R])$. Write $A = [a,b] \times \mathbb{R}/\mathbb{Z}$. Then by positivity of energy,

$$E(u; A) = \int_A u^*\omega + \int_{\partial A} u^* H\, dt \geq \left| \left| \int_A u^*\omega \right| - \left| \int_{\partial A} u^* H\, dt \right| \right|. \tag{37}$$

We will show that if we take $\epsilon$ small enough, there are constants $\delta_1$ and $\delta_2$ such that

$$E(u; A) < \delta_1 \implies \left|\int u^*\omega\right| < \delta_2. \tag{38}$$

Since

$$\left|\int_{\partial A} u^* H\, dt\right| \geq R,$$

it will then follow from (37) that if $R > 2\delta_2$, then $E(u; A) > \min\{\delta_1, \delta_2\}$. This will prove the claim.

Let $\delta > 0$ be so small that any loop of length $2\delta$ has diameter less than a tenth of the radius of injectivity of $M$ with respect to $g_J$. The isoperimetric inequality of Lemma 4.10 guarantees that any loop of length $< 2\delta$ is fillable by a disk $v\colon D \to M$ such that

$$\mathrm{Area}(v) < 2\delta^2.$$

We take $\epsilon = \delta$. Given $u$ as above, and denoting by $\ell(u(s,\cdot))$ the length of the loop $t \mapsto u(s,t)$, let

$$I = \{s \in [a,b] \mid \ell(u(s,\cdot)) > 2\delta\}.$$

For any interval $(c,d) \subset I$ we have the estimate

$$\left|\int_{(c,d)\times\mathbb{R}/\mathbb{Z}} u^*\omega\right| \leq \mathrm{Area}(u|_{(c,d)\times\mathbb{R}/\mathbb{Z}}) \leq 3E(u; (c,d)\times\mathbb{R}/\mathbb{Z}). \tag{39}$$

The first of these is Wirtinger's inequality, which says that for a compatible metric the symplectic area is dominated by the Riemannian area. Note that the Riemannian area is not sensitive to orientation, while the symplectic area is. For the second, note that

$$\mathrm{Area}(u) \leq \int_{(c,d)\times\mathbb{R}/\mathbb{Z}} \|\partial_s u\|\|\partial_t u\| \leq \frac{1}{2}\int_{(c,d)\times\mathbb{R}/\mathbb{Z}} (\|\partial_s u\|^2 + \|\partial_t u\|^2).$$

But

$$\begin{aligned} 4\delta^2 \leq \int_{\mathbb{R}/\mathbb{Z}} \|\partial_t u\|^2 &\leq \int_{\mathbb{R}/\mathbb{Z}} \|\partial_t u - X_H\|^2 + \int_{\mathbb{R}/\mathbb{Z}} \|X_H\|^2 \\ &\leq \int_{\mathbb{R}/\mathbb{Z}} \|\partial_s u\|^2 + \epsilon^2 = \int_{\mathbb{R}/\mathbb{Z}} \|\partial_s u\|^2 + \delta^2 \\ &\leq \int_{\mathbb{R}/\mathbb{Z}} \|\partial_s u\|^2 + \frac{1}{4}\int_{\mathbb{R}/\mathbb{Z}} \|\partial_t u\|^2, \end{aligned} \tag{40}$$

so

$$\int_{\mathbb{R}/\mathbb{Z}} \|\partial_t u\|^2 \leq 2\int_{\mathbb{R}/\mathbb{Z}} \|\partial_s u\|^2,$$

which implies the desired inequality.

Suppose now that

$$\left|\int_{(a,b)\times\mathbb{R}/\mathbb{Z}} u^*\omega\right| > 20E(u;(a,b)\times\mathbb{R}/\mathbb{Z}),$$

and that for some constant $c_2$ to be determined shortly, $E(u) < c_2\epsilon^2$. Denote by $\mu$ the Lebesgue measure on $\mathbb{R}$. Then by these hypotheses and by equations (39) and (40) we have $\mu(I) < \min\{\frac{1}{4}(a-b), c_2\}$. Let

$$a' = \inf[a,b]\setminus I \quad \text{and} \quad b' = \sup[a,b]\setminus I.$$

We will show that if $c_2$ is assumed small enough, then

$$\left|\int_{(a',b')\times\mathbb{R}/\mathbb{Z}} u^*\omega\right| < 4\delta^2. \tag{41}$$

We then have

$$\begin{aligned}\left|\int_{(a,b)\times\mathbb{R}/\mathbb{Z}} u^*\omega\right| < 4\delta^2 + \left|\int_I u^*\omega\right| &< 4\delta^2 + \tfrac{1}{3}E(u) \\ &< 4\delta^2 + \frac{1}{60}\left|\int_{(a,b)\times\mathbb{R}/\mathbb{Z}} u^*\omega\right|.\end{aligned}$$

By picking $\delta_1 = c_2\epsilon^2$ and $\delta_2 = \min\{5\delta^2, 20\delta_1\}$, we get that with these values (38) holds in any case.

It remains to prove (41). Let $[s_0,s_1] \subset [a,b]$ be any interval such that $s_0, s_1 \notin \{a,b\}$ and $s_1 - s_0 \le 2\mu(I)$. Call such an interval *admissible*. Denote by $u_{[s_0,s_1]}$ the restriction $u|_{[s_0,s_1]\times\mathbb{R}/\mathbb{Z}}$. Each component of the boundary of $u_{[s_0,s_1]}$ is contained in a geodesic ball $B_\delta(x_i) \subset M$. We claim that if $c_2$ is taken small enough, then

$$u_{[s_0,s_1]} \subset B_{2\delta}(x_0) \cup B_{2\delta}(x_1). \tag{42}$$

Indeed, otherwise there is a point $(s,t) \in [s_0,s_1]\times\mathbb{R}/\mathbb{Z}$ such that writing $x_2 = u(s,t)$ we have $d(x_2,\{x_0,x_1\}) > 2\delta$, that is, the ball $B_\delta(x_2) \subset M$ does not meet the boundary of $u_{[s_0,s_1]}$. As in Lemma 5.11, the metric $g_{J_H}$ is quasi-isometric to the product metric of $g_J$ on $M$ with the flat metric on the cylinder, where we can take the quasi-isometry constant to equal 2 if $\epsilon < 2$. Thus we can apply the monotonicity inequality of Theorem 4.9 to obtain, for an appropriate constant $c'$ which is independent of $\epsilon$,

$$E(\widetilde{u};[s_0,s_1]\times\mathbb{R}/\mathbb{Z}) = s_1 - s_0 + E(u;[s_0,s_1]\times\mathbb{R}/\mathbb{Z}) \ge c'\delta^2,$$

where $\widetilde{u}$ is the graph of $u$. This implies

$$E(u) \ge c'\delta_2 - c_2.$$

Thus, if we take $c_2 \le \frac{1}{4}c'\delta_2$, equation (42) follows.

Denote by $u^*$ a filling of $u_{[s_0,s_1]}$ by discs contained in $B_\delta(x_i)$. Then $u^* \subset B_{5\delta}(x_0)$ and so is contractible. In particular, the integral of $\omega$ over $u_{[s_0,s_1]}$ can be replaced by the integral of $\omega$ over these filling discs. Since $[a',b']$ can be subdivided into admissible intervals, and the integrals over the filling discs cancel in pairs for all but two, (41) follows.

This proves the theorem for the case when the left-hand side of (36) is taken to hold for all $x \in M$.

For the more general case we argue as follows. Write $K_x := H^{-1}([-x,x])$. For some $R > 0$ let $u\colon [a,b]\times\mathbb{R}/\mathbb{Z} \to M$ be a solution to Floer's equation with one end in $K_h$ and the other in $M \setminus K_{h+R}$. Let $[a',b'] \subset [a,b]$ be such that $u([a',b']\times\mathbb{R}/\mathbb{Z})$ has one end in $K_{h+R/4}$ and the other in $M \setminus K_{h+3R/4}$. In each case assume the relevant boundary of $u([a',b']\times\mathbb{R}/\mathbb{Z})$ meets the boundary of the region $K_x$. We separate into two cases. If $u([a',b']\times\mathbb{R}/\mathbb{Z}) \subset K_R$, we have the estimate $\|X_H\| < \epsilon$ for $u|_{[a',b']\times\mathbb{R}/\mathbb{Z}}$, and the entire argument goes through with no change. By taking $R$ big enough, the claim follows since $\Gamma(h,h+R) \geq \Gamma(h+R/4,h+3R/4)$. Otherwise, for some $c \in \{a',b'\}$ we have that the oscillation of $H$ along $u_c$ is at least $R/4$. Moreover, by the bound on $\nabla H$ inside $K_R$, a similar estimate applies to the diameter of $u_c$ with respect to the metric $g_{J_H}$. By the argument of Theorem 4.11 this implies a lower bound on the energy $E(u;[a,b]\times\mathbb{R}/\mathbb{Z})$. We spell out the details, since the present case doesn't fit precisely into the stipulations of Theorem 4.11.

As above, denote by

$$\tilde{u}\colon [a,b]\times\mathbb{R}/\mathbb{Z} \to \mathbb{R}\times\mathbb{R}/\mathbb{Z}\times M$$

the graph of $u$. Since $H$ has Lipschitz constant $\epsilon$ on $K$, Lemma 5.16 implies the metric $g_{J_H}$ is equivalent on $K$ to the product metric with quasi-isometry constant depending only on $\epsilon$. Thus, by Theorem 4.9 there are constants $\delta_0$ and $r_0$, depending only on $J$ and $\epsilon$, such that for any point $x$ in the domain of $u$ for which

$$(*) \qquad A_x := \tilde{u}^{-1}(B_{r_0}(x,u(x)) \subset (a,b)\times\mathbb{R}/\mathbb{Z},$$

we have

$$E(u;A_x) + \mathrm{Area}(A_x) \geq \delta_0.$$

Here we consider the ball $B_{r_0}(x,u(x) \subset \widetilde{M}$ with respect to the metric $g_{J_H}$. Call a point for which the hypothesis $(*)$ holds a good point. Since the ends of the $u$ map entirely outside of $K_{h+R}$ it follows that for any $x$ for which $u(x) \in K_{h+R-(1+\epsilon)r_0}$, we have that $A_x$ is a good point and moreover, $A_x \subset B_{r_0}(x)$. Indeed, for any $x, x'$ in the domain we have

$$d_{g_{J_H}}\big((x, u(x)), (x', u(x'))\big) \geq (1-\epsilon) d_{g_J}(u(x), u(x')).$$

For any $N$, by assuming $R$ is large enough, we can find $N$ good points $x_i \in \{c\} \times \mathbb{R}/\mathbb{Z}$ such that $d_{g_H}(\widetilde{u}(x_i), \widetilde{u}(x_j) > 2r_0$; that is, $A_{x_i} \cap A_{x_j} = \varnothing$ whenever $i \neq j$. We then have

$$E(u) + 2r_0 \geq E(u; \cup A_{x_i}) + \mathrm{Area}(\cup A_{x_i}) \geq N\delta_0.$$

By taking $N$ large enough so that $N\delta_0 - 2r_0 > \delta$ for some chosen $\delta > 0$, the claim follows. □

**Proof of Theorem 6.6** Lemmas 6.9 and 6.8 imply that $(H, J)$ is RLD if $\|X_H\| < \epsilon$ for $\epsilon$ small enough. To establish dissipativity, we need to prove, in addition, i–boundedness. This follows immediately from Lemma 5.11. □

## 6.3 Bidirectedness

**Theorem 6.10** *For **any** smooth exhaustion function $H: M \times /\mathbb{R}/\mathbb{Z} \to \mathbb{R}$ and any geometrically bounded $\omega$–compatible almost complex structure $J$, there are exhaustion functions $H_+, H_-$ such that $(H_\pm, J)$ are dissipative Floer data and $H_- \leq H \leq H_+$ pointwise. In other words, the set of Hamiltonians which taken together with $J$ are dissipative Floer data is both final and cofinal in the set of all exhaustion functions.*

**Proof** According to [27] there exists an exhaustion function $f: M \to \mathbb{R}$ such that $\|\nabla f\| = \|X_f\| < \epsilon_0$ with respect to the metric $g_J$. Moreover, we may find a constant $R_0$ such that $d(f^{-1}(x), f^{-1}(x+R_0))$ is bounded away from 0 for $x \in \mathbb{R}$. Indeed, $f$ can be taken to be $C^0$–close to a multiple of the distance function to some point. So, $(f, J)$ is dissipative by Theorem 6.6. Let $h: \mathbb{R} \to \mathbb{R}$ be any monotone function such that $h'(x) = 1$ on any of the intervals $[2nR, (2n+1)R)$ and is arbitrary otherwise. Here $R$ is a constant as in Lemma 6.9, and without loss of generality $R > R_0$. Then the set of functions of the form $h \circ f$ is cofinal in the set of all exhaustion functions. On the other hand, $(h \circ f, J)$ is dissipative. Indeed, $h \circ f$ is clearly $J$–proper, since $f$ is. The metric $g_{X_{h \circ f}}$ is uniformly bounded on each of the regions $f^{-1}(h, h+R)$ by Lemma 5.11. So this metric is i–bounded. Lemmas 6.9 and 6.8 imply that $h \circ f$ is RLD. This completes the proof of cofinality.

By Theorem 6.6, to prove finality it suffices to exhibit an exhaustion function $H_- \leq H$ which has sufficiently small gradient. Fix a point $p \in M$ and let $R_i$ be a monotone increasing sequence such that $B_{R_i}$ contains $H^{-1}((-\infty, i])$. Denote by $h\colon M \to \mathbb{R}$ the distance function $h(x) = d(x, p)$. Define $a_i$ inductively by $a_0 = 0$ and

$$a_i = \min\{i-1, a_{i-1} + R_i - R_{i-1}\} \quad \text{for } i \geq 1.$$

Let $f\colon \mathbb{R}_+ \to \mathbb{R}$ be the piecewise linear function which is smooth at noninteger points and satisfies $f_i = a_i$ for $i \geq 1$. Note that $f$ is monotone increasing, proper and has slope at most 1 wherever the slope is defined. So the function $g = f \circ h$ is Lipschitz with Lipschitz constant 1. Moreover, $g \leq H$ everywhere. The function $g$ can be $C^0$–approximated by a smooth function $k$ with $\|\nabla k\| \leq 2$; see [27]. Then $k$ is an exhaustion function, so taking $H_- := k/C$ for $C$ sufficiently large gives a function as required. □

## 6.4 Dissipativity on exact symplectic manifolds

Let $(M, \omega = d\alpha)$ be an exact symplectic manifold. In this subsection we prove Theorem 6.12, which is variant of Theorem 6.3 that works on exact symplectic manifolds under slightly weaker hypotheses. Fix an $\omega$–compatible almost complex structure and let $H\colon \mathbb{R}/\mathbb{Z} \times M \to \mathbb{R}$. The pair $(H, J)$ is said to be *Palais–Smale* if any sequence of loops $\gamma_n$ with $\mathcal{A}_H(\gamma_n) < c < \infty$ and $\|\nabla \mathcal{A}_H(\gamma_n)\|_{L^2} \to 0$ has a subsequence converging to a periodic orbit of $H$. If $J_0, J_1$ are almost complex structures which are quasi-isometric, and $H_0, H_1$ are Hamiltonians such that $\|\nabla(H_0 - H_1)\|$ converges to 0 with respect to either, then $(H_0, J_0)$ is Palais–Smale if and only if $(H_1, J_1)$ is.

**Lemma 6.11** *Suppose $(H, J)$ is i–bounded and Palais–Smale. Then for any $c$ and $d$ there is a real number $\ell$ and a compact set $K$ with the following significance. For any segment $[a, b]$ of length at least $\ell$, and any solution $u\colon [a-1, b+1] \times \mathbb{R}/\mathbb{Z} \to M$ to Floer's equation such that $\mathcal{A}_H(u(s, \cdot)) \in [c, d]$ for $s \in [a-1, b+1]$, we have $u([a, b] \times \mathbb{R}/\mathbb{Z}) \subset K$.*

**Proof** First, by the Palais–Smale condition, there are an $\epsilon > 0$ and a compact $K' \subset M$ such that any loop $\alpha$ with $\|\nabla \mathcal{A}_H(\alpha)\|_{L^2} < \epsilon$ and $\mathcal{A}_H(\alpha) < d$ is contained in $K'$. Indeed the negation of this statement would allow us to produce a sequence of loops $\gamma_n$ satisfying the hypotheses of the Palais–Smale condition which nevertheless has no convergent subsequence. Suppose $b - a > (d - c)/\epsilon$. Then for all but a subset $I \subset [a, b]$

of total measure $(d-c)/\epsilon$ we have that $u(s,\cdot)$ is contained in $K'$. This follows by the energy estimate

$$d-c \geq \mathcal{A}_H(u(b+1)) - \mathcal{A}_H(u(a-1)) = \int_{a-1}^{b+1} \|\nabla \mathcal{A}_H(\alpha)\|_{L^2}.$$

Indeed, taking $I$ to be the set of $s$ for which $u(s,\cdot)$ is not contained in $K'$, the right-hand side of the last equation dominates:

$$\int_I \|\nabla \mathcal{A}_H(\alpha)\|_{L^2} \geq \epsilon \int_I dt.$$

It remains to control $u(s,\cdot)$ for $s \in I$. Each connected component $I'$ of $I$ has at least one boundary point $s$ for which $u(s,\cdot) \subset K$. Moreover, $I' \subset I$ has a priori bounded measure. Thus applying part (b) of Theorem 4.11 to the graph $\widetilde{u}|_{I'\times\mathbb{R}/\mathbb{Z}}$, we deduce the image of $I' \cap [a,b] \times \mathbb{R}/\mathbb{Z}$ is contained in some larger compact set $K$ depending only on $K'$ and $d-c$. □

Note that the Palais–Smale condition produces, by Lemma 6.11, an estimate which is slightly weaker than loopwise dissipativity because it depends not only on energy but also on action. Nevertheless, this is sufficient for proving the following variant of Theorem 6.3.

**Theorem 6.12** *Suppose that $(M, \omega = d\alpha)$ is an exact symplectic manifold and that $(\mathcal{S}, F_{s\in\mathcal{S}} = (\mathfrak{H}_s, J_s))$ is a uniformly i–bounded family of connected (broken) Riemann surfaces with a thick–thin decomposition as in Definition 5.10. Let $(H_i, J_i)$ be Palais–Smale Floer data such that on the $i^{th}$ component of* $\mathrm{Thin}_{\mathcal{S}}$*, we have that $F_s$ coincides with $(H_i, J_i)$ for all $s \in \mathcal{S}$. Then for any interval $[c,d]$, there is a compact set $K \subset M$ such that for any $s \in \mathcal{S}$ and any solution $(\Sigma, u)$ associated with $F_s$ for which the actions of the periodic orbits on the ends all occur in the interval $[c,d]$, the image of $u$ is contained in $K$.*

**Proof** First observe that without loss of generality we may assume the all the components of $\mathrm{Thin}_{\mathcal{S}}$ are of the form $I \times \mathbb{R}/\mathbb{Z}$ for $I$ an interval of length at least $\ell$, where $\ell$ is as in Lemma 6.11. Namely, with this assumption, the areas of the elements of $\mathrm{Thick}_{\mathcal{S}}$ are bounded a priori in terms of $c$ and $d$. Under such identification it is a consequence of Lemma 5.3 that for any $s \in I$ we have $\mathcal{A}_{H_i}(u_s) \in [c,d]$. It thus follows from Lemma 6.11 that there is a compact set $K$, depending only on $c$ and $d$, such that the images of the components of $\mathrm{Thin}_{\mathcal{S}}$ are all contained in $K$. As a consequence, the image of each component $A$ of $\mathrm{Thick}_S$ meets $K$. Since the energy of $A$ is at most $d-c$, we can as in the proof of Theorem 6.3 apply Theorem 4.11(b) to the graph of $u|_A$ to obtain an $R = R(d-c)$ such that $u(A) \subset B_R(K)$. □

**Example 6.13** Let $\alpha$ be a primitive of $\omega$ and let $Z$ be the $\omega$–dual of $\alpha$. For any time-independent Hamiltonian $H$, define the function $f : M \to \mathbb{R}$ by

$$f(x) = \omega(Z(x), X_H(x)) - H(x). \tag{43}$$

Suppose $f$ is proper and bounded below and $J$ is such that for some constant $C$,

$$\|Z(x)\|^2 < C f(x)$$

outside a compact set. Then $H$ is Palais–Smale.

**Proof** We have

$$\begin{aligned} \mathcal{A}_H(\gamma) &= \int_{\mathbb{R}/\mathbb{Z}} f(\gamma(t))\,dt + \int_{\mathbb{R}/\mathbb{Z}} \omega\big(Z(\gamma(t)), \gamma'(t) - X_H(t)\big)\,dt \\ &\geq \int_{\mathbb{R}/\mathbb{Z}} f(\gamma(t))\,dt - \|\nabla \mathcal{A}_H(\gamma)\| \sqrt{C \int_{\mathbb{R}/\mathbb{Z}} f(\gamma(t))\,dt}. \end{aligned} \tag{44}$$

Suppose $\mathcal{A}_H(\gamma) \leq c$ and $\|\nabla \mathcal{A}_H(\gamma)\| \leq 1$. Since $f$ is proper, estimate (44) implies that there is a compact set $K$ depending only on $c$ such that $\gamma$ intersects $K$. Given a sequence $\gamma_n$ of loops intersecting $K$ such that

$$\|\nabla \mathcal{A}_H(\gamma)\|_{L^2} \geq \int_{\mathbb{R}/\mathbb{Z}} \|X_H(t) - \gamma_n'(t)\| \to 0,$$

it is a standard fact that the sequence converges to an integral loop of $X_H$. □

In particular, consider the convex end of a symplectization $\mathbb{R}_+ \times \Sigma$ as in Example 5.24. Denote by $r$ the coordinate on $\mathbb{R}_+$ and by $\sigma$ the coordinate on $\Sigma$. If $H$ satisfies

$$\lim_{r\to\infty} e^r(\partial_r H)(e^r, \sigma) - H(e^r, \sigma) \to \infty,$$

and $J$ is any almost complex structure satisfying

$$e^r \alpha(J\partial r) \leq C(e^r \partial_r H(e^r, \sigma) - H(e^r, \sigma))$$

for some $C$, then $H$ is Palais–Smale. This holds in particular for contact-type $J$, ie satisfying $J\partial_r = R$, where $R$ is the Reeb flow of $\alpha$ on $\Sigma$. After a $C^2$–small perturbation, $(H, J)$ will satisfy the same estimates, so it will remain Palais–Smale. In addition, it will be nondegenerate.

**Example 6.14** Continuing with the convex end of a symplectization, any function which is of the form $h(e^r)$ such that $e^r h'(e^r) - h(e^r) \geq Ce^r$ for some constant $C$ is Palais–Smale. This holds, for instance, for $h(x) = x^\alpha$ with $\alpha > 1$. If we have $h'(e^r) \leq e^{r/2}$ then by Example 5.25 $H$ is dissipative. The cutoff appears to be $\alpha = \frac{3}{2}$, which unfortunately excludes quadratic Hamiltonians, which are central in classical mechanics. See also the discussion in Example 5.26. Nevertheless, as we shall see below, Floer cohomology can be defined by approximation by slow Hamiltonians. Moreover, similarly to the proof of Remark 10.3, it can be shown that for an arbitrary convex Hamiltonian the resulting Floer cohomology coincides with the Floer cohomology defined using contact-type $J$ and relying on maximum principles.

When $e^r h'(e^r) - h(e^r) \to c < \infty$ for some $c$ which is not in the period spectrum, $H$ is still Palais–Smale even though this is not covered by the previous example, and in particular, it is dissipative. A proof of this fact is given below in Example 6.20.

## 6.5 Some not necessarily exact examples

Let $(M, g)$ be a Riemannian manifold and $V$ a time-dependent vector field on $M$. For $p$ in $M$ define

$$f(p, V, g) := \inf_{\{\gamma : [0,1] \to M \mid \gamma(0) = \gamma(1) = p\}} \left\{ \int_0^1 \|\gamma'(t) - V_t \circ \gamma(t)\|^2 \, dt \right\}.$$

Clearly, $f$ is continuous with respect to all variables in the $C^0$ norm. We drop $g$ from the notation when there is no ambiguity.

**Lemma 6.15** *Let $(H, J)$ be such that $g_{J_H}$ has uniformly bounded geometry. Suppose that there is a compact $K \subset M$ and a $\delta > 0$ such that for all $p \in M \setminus K$ we have $f(p, X_H, g_J) \geq \delta$. Then $(H, J)$ is RLD.*

**Proof** Let $u : [a, b] \times \mathbb{R}/\mathbb{Z}$ be a partial solution with one boundary in a compact set $K_0 \supset K$ and with energy $E(u) \leq E$ for some $E$. Without loss of generality $u_a \subset K_0$. Suppose $u_b \subset M \setminus B_{R_0}(K_0)$ for some $R_0$. Then considering the graph of $u$ as a $J_H$–holomorphic map, it has energy $E + (b - a)$. Theorem 4.11(a) applied to the compact set $\partial(B_{R_0}(K_0) \setminus K_0)$ then implies that for some constant $C$ depending on the bound on the geometry, we have $R_0 \leq C(E + (b - a))$. The assumption on $f(p, X_H, g_J)$ implies $b - a \leq E/\delta$. Taken together we obtain the estimate $R_0 \leq CE(1 + 1/\delta)$. □

The quantity $f(p,V,g)$ can sometimes be estimated from below by the following procedure. We say that the pair $(V,g)$ is of *Lyapunov type* if there exists a constant $\lambda\geq 0$ such that for any $x,y\in M$ and any $t\geq 0$ we have

$$d_g(\phi_t(x),\phi_t(y))\leq e^{\lambda t}d_g(x,y), \tag{45}$$

where $\phi_t$ denotes the time $t$ flow of $V$. We refer to $\lambda$ as a Lyapunov constant for $V$.

**Lemma 6.16** *If $V$ is time-independent and $\|\nabla V\|\leq\lambda$, then $\lambda$ is a Lyapunov constant for $V$.*

**Proof** For $x\neq y$ close enough and for sufficiently short times, the function $h(t)=d(\phi_t(x),\phi_t(y))$ is differentiable. Moreover, for each $t$ there is a unique geodesic $s\mapsto\alpha_t(s)$ realizing the distance between $d(\phi_t(x),\phi_t(y))$. Denote by $\overline{V}_{\phi_t(x)}$ the parallel transport of $V_{\phi_t(x)}$ along $\alpha_t$. Considering that the gradient of the distance function $d(x,y)$ for, say, $x$ fixed is the tangent vector to the unit-speed geodesic from $x$ to $y$, it follows that

$$\frac{dh}{dt}=\langle\alpha_t'(1),V_{\phi_t(y)}\rangle-\langle\alpha_t'(0),V_{\phi_t(x)}\rangle. \tag{46}$$

From this we obtain the differential inequality

$$\frac{dh}{dt}\leq|V_{\phi_t(y)}-\overline{V}_{\phi_t(x)}|\leq\lambda h. \tag{47}$$

The claim for $x,y$ sufficiently close and for sufficiently short times now follows by Grönwall's inequality. The claim for arbitrary $x,y$ and sufficiently short times follows by the triangle inequality. The claim for arbitrary long time follows since the flow $\phi_t$ is autonomous. □

**Lemma 6.17** *Suppose that $V$ is a time-independent vector field of Lyapunov type with Lyapunov constant $\lambda\geq 0$. Then*[9]

$$d_g(p,\phi_1(p))^2\leq\frac{e^{2\lambda}-1}{2\lambda}f(p,V,g). \tag{48}$$

**Remark 6.18** For $V=X_H$ with $H$ time-independent and uniformly Lipschitz we can replace the global requirement that (45) hold everywhere with the requirement that it hold for points $x,y\in U:=H^{-1}([H(p)-\epsilon,H(p)+\epsilon]$ for some $\epsilon>0$. We then get an estimate from below on $f(p,V,g)$ by combining the present lemma, to estimate

[9]When $\lambda=0$ the coefficient on the right-hand side tends to 1.

the energy of loops which map into $U$, with Lemma 5.13, to estimate the energy of the loops at $p$ which do not remain within $U$. Moreover, this estimate depends only on the Lipschitz constant of the restriction $H|_U$ and remains valid if $H$ is arbitrarily time-dependent outside of $U$.

**Proof of Lemma 6.17** Fix some $\epsilon > 0$, which will be later taken to be arbitrarily small. Let $\gamma : [0,1] \to M$ be a loop based at $p$. Let $r \in \mathbb{R}$ be small enough so that for each point $q \in \gamma([0,1])$ there is a chart $(U_q \subset M, \psi_q : B_{2r}(0) \to U_q)$ with coordinate map $\psi_q$ which is bi-Lipschitz with Lipschitz constant $1+\epsilon$. By compactness of $\gamma([0,1])$, there is a constant $K$ such that for any $q$ the vector field $d\psi_q^{-1} V$, considered as a map $B_{2r}(0) \to \mathbb{R}^{2n}$, is Lipschitz with constant $K$.

Write

$$g(t) := \|\gamma'(t) - V_{\gamma(t)}\| \quad \text{and} \quad f(t) = \int_0^t g(s)\, ds.$$

Let

$$\Delta t \ll r \min\left\{ \frac{1}{K}, \frac{1}{\sup\|\gamma(t)\|}, \frac{1}{\sup\|V_{\gamma(t)}\|}, \frac{1}{\max g(t)} \right\}.$$

Without loss of generality suppose $N := 1/\Delta t$ is an integer. Suppose $\Delta t$ is made smaller still so that $f(t)$ has an approximation by a piecewise linear function $h(t)$ such that

$$(1-\epsilon) h'(t) < g(t) \le h'(t) \tag{49}$$

and such that $h$ is linear of slope $\epsilon_i$ on the intervals $[i/N, (i+1)/N]$. Let $t_i = i/N$. Let $\gamma_i(t) := \phi_{t-t_i}(\gamma(t_i))$ and let $x_i = \gamma_i(1)$. Writing $\Delta x_i := d(x_i, x_{i-1})$ for $i = 1, \dots N$ we have, by the Lyapunov condition,

$$\Delta x_i \le e^{\lambda(1-t_i)} d_g(\gamma_i(t_i), \gamma_{i-1}(t_i)).$$

On the other hand we have an estimate

$$d_g(\gamma_i(t_i), \gamma_{i-1}(t_i)) \le (1+\epsilon) \frac{\epsilon_i}{K} (e^{K\Delta t} - 1). \tag{50}$$

To see this note that both the path $\gamma$ and the path $\gamma_{i-1}$ map the interval $[t_{i-1}, t_i]$ into the coordinate chart $U_{\gamma(t_{i-1})}$. Let

$$k(t) = d_0(\gamma(t), \gamma_{i-1}(t)), \quad t \in [t_{i-1}, t_i], \tag{51}$$

be the Euclidean distance. Then $k(t)$ satisfies the differential inequality

$$\frac{dk}{dt} \le |\gamma'(t) - V_{\gamma(t)}| + |V_{\gamma(t)} - V_{\gamma_{i-1}(t)}| \le g(t) + Kk(t),$$

with initial condition $k(t_{i-1})=0$. By Grönwall's inequality we get, for $t\in[t_{i-1},t_i]$,

$$\begin{aligned} d_g(\gamma(t),\gamma_{i-1}(t)) \le (1+\epsilon)k(t) &\le (1+\epsilon)e^{K(t-t_{i-1})}\int_{t_{i-1}}^t e^{-Ks}g(s)\,ds \\ &\le (1+\epsilon)\frac{\epsilon_i}{K}(e^{K(t-t_{i-1})}-1), \end{aligned}$$

implying (50). The right-hand side of (50) is $\le (1+\epsilon)^2\epsilon_i\Delta t$ since $\Delta t \ll 1/K$.

We have $\phi_1(\gamma(0))=x_0$ and $\gamma(1)=\gamma(0)=x_N$. Thus,

$$d(x_0,x_N)\le \sum \Delta x_i \le \sum_{i=1}^N (1+\epsilon)^2\epsilon_i e^{\lambda(1-t_i)}\Delta t.$$

The last expression approximates the integral

$$(1+\epsilon)^2\int_0^1 h'(t)e^{\lambda(1-t)}\,dt \le \sqrt{\int_0^1 (h'(t))^2\,dt}\sqrt{\frac{e^{2\lambda}-1}{2\lambda}}.$$

Combining the last two inequalities gives the estimate

$$d_g(p,\phi_1(p))^2 = d_g(x_0,x_N)^2 \le \frac{(1+\epsilon)^2}{(1-\epsilon)^2}\frac{e^{2\lambda}-1}{2\lambda}\|\gamma'-V_\gamma\|_{L^2}^2.$$

Since $\epsilon$ is arbitrary this proves the claim. □

We say that a Floer datum $(H,J)$ is of Lyapunov type if $(X_H,g_J)$ is of Lyapunov type.

**Corollary 6.19** *Suppose $J$ is geometrically bounded, $(H,J)$ is of Lyapunov type, and that outside of a compact set the quantity $d(p,\psi_1(p))$ is bounded away from 0. Then $(H,J)$ is RLD.*

**Proof** This is an immediate consequence of Lemmas 6.17 and 6.15. □

**Example 6.20** Using the notation of Example 5.24, let $M$ have an end modeled on $\Sigma\times\mathbb{R}_+$ and let $H_0$ be a function which is linear at infinity, with slope $a$ not in the period spectrum. Then $H$ is of Lyapunov type. Indeed the flow on any level set of $H$ is of Lyapunov type by Lemma 6.16 and compactness. Since the flows on different level sets are conjugate, the existence of a Lyapunov estimate follows also for $x$ and $y$ not on the same level set. Let $X=X_{H_0}|_{\Sigma\times\{1\}}$. Let $J$ be a translation-invariant almost complex structure. Then from (48) it follows that $f(p, X_{H_0})$ is bounded away from 0, and so $H_0$ is LD.

Let $\delta$ be the distance of $c$ to the period spectrum of $\Sigma_\alpha$ and let $H_1$ be any Hamiltonian such that

$$\frac{\|X_{H_1} - X_{H_0}\|}{\|X_{H_0}\|} \ll \tfrac{1}{2}\delta.$$

For example, this inequality will hold for our choice of $J$ whenever $\|X_{H_1 - H_0}\|$ is bounded. Then $f(p, X_{H_1})$ is bounded away from 0 at infinity. So, $H_1$ is also LD.

**Example 6.21** Let $M_1$ be as in the previous example and let $M_2$ be a compact symplectic manifold. Let $a$ be a real number not in the period spectrum of $M_1$ and let $f\colon \mathbb{R}/\mathbb{Z} \times M_1 \times M_2$ be any function which tends to 1 at infinity with derivatives dominated by $o(e^{-r/2})$. Then, reasoning as in the previous example, the function $H := afe^r$ is LD.

**Lemma 6.22** *Let the end of $M$ be diffeomorphic to $\Sigma \times \mathbb{R}_+$, with $\Sigma$ a compact hypersurface. Suppose the projection $\pi\colon \Sigma \times \mathbb{R}_+ \to \Sigma$ satisfies $\|\pi_* v\| \le \|v\|$ for any tangent vector $v \in T(\Sigma \times \mathbb{R}_+)$. Let $X$ be vector field on $\Sigma$ with no 1–periodic orbits and let $H$ be such that $\pi_* X_H$ converges uniformly to $X$. Then for some $\delta > 0$ we have $f(p, X_H) > \delta > 0$ and, in particular, $X_H$ is LD.*

**Proof** Let $\epsilon$ be such that $f(p, X) > \epsilon$. For $r$ large enough, the convergence assumption implies

$$f(p, \pi_* X_H) > \tfrac{1}{2}\epsilon.$$

The nonincreasing assumption implies $f(p, X_H) > f(p, \pi_* X_H)$. □

**Example 6.23** Let $M$, $\Sigma$, $\alpha$, $H_0$ and $H_1$ be as in Example 6.20. Let $\sigma$ be a closed two-form on $\Sigma$. Suppose $\sigma$ extends to a closed form on $M$ which is invariant under the Liouville flow near $\Sigma$. Then $\sigma$ can be extended in a translation-invariant way to a closed two-form on the completion of $M$, still denoted by $\sigma$. For $t$ small enough, the form $\omega_{t\sigma} = -d\alpha + t\sigma$ defines a symplectic form on the completion of $M$. By rescaling $\sigma$, assume this holds for $t = 1$. Then $H_0$ and $H_1$ are LD for the symplectic form $\omega_\sigma$. Indeed, write $X'_{H_0}$ for the Hamiltonian vector field with respect to $\omega_\sigma$. Let $X$ be as in Example 6.20. Then all the requirements of Lemma 6.22 are satisfied for the pair $X'_{H_0}, X$. The claim for $H_1$ now follows by comparison to $H_0$.

**Example 6.24** In Example 6.20 assume the pair $\Sigma, \alpha$ is not necessarily contact, but stable Hamiltonian for the restriction $\omega_1 := \omega|_{\Sigma\times 1}$, with stabilizing form $\alpha$. Namely, $\alpha$ satisfies $\ker \omega \subset \ker d\alpha$ and $\alpha \wedge \omega^{n-1} > 0$. Assume $\omega$ is of the form $\omega_\alpha := \omega + d(e^r\alpha)$ on $\Sigma \times \mathbb{R}_{\geq 0}$ and is symplectic for all $r \geq 1$. Assume further that there exists a translation-invariant $\omega_\alpha$–compatible almost complex structure $J$ on $\Sigma \times \mathbb{R}_{\geq 0}$. Then the forms $\omega(\cdot\,, J)$ and $d\alpha(\cdot\,, J)$ are separately nonnegative. So the projection $\pi_*$ is norm nonincreasing. So if $H_0$ is linear at infinity with slope not in the period spectrum, then $f(p, X_{H_0})$ is bounded away from 0 and $H_0$ is dissipative. The same will hold under a sufficiently small deformation of $\omega$ or a sufficiently small Hamiltonian perturbation of $X_{H_0}$.

In all the examples of this section we have considered Hamiltonians which are roughly linear at infinity. It is easy to use these examples to construct superlinear Hamiltonians which are LD. It is an interesting question as to what Hamiltonians can be perturbed to become LD. The property of being LD is clearly related to the behavior of the function $f(p, X_H, g_J)$. Namely, if one can find an exhaustion for which this function is appropriately bounded away from 0 near the boundaries, the Floer datum will be LD.

# 7 Proof of Theorem 1.1

## 7.1 Floer systems

For a symplectic manifold $(M, \omega)$, denote by $\mathcal{J}(M, \omega)$ the set of $\omega$–compatible almost complex structures on $M$. Let

$$\mathcal{F} \subset C^\infty(\mathbb{R}/\mathbb{Z} \times M) \times C^\infty(\mathbb{R}/\mathbb{Z}, \mathcal{J}(M, \omega))$$

denote the set of Floer data $(H, J)$ such that $H$ is proper and bounded from below, the Hamiltonian flow of $H$ is defined for all time, and the metric $g_{J_t} := \omega(\cdot\,, J_t\,\cdot)$ is complete for any $t \in \mathbb{R}/\mathbb{Z}$. Denoting by $\Delta^i$ the standard simplex, let

$$\mathcal{F}^i \subset C^\infty(\Delta^i, \mathcal{F})$$

be the subset consisting of elements which are constant in a neighborhood of the vertices. Furthermore, we require that for any $F \in \mathcal{F}^{(1)}$, $\partial_s F \geq 0$. Denote by $\Delta^{i,\circ}$ the interior of the simplex. Fix once and for all diffeomorphisms

$$\sigma : \mathbb{R} \to \Delta^{1,\circ}, \quad \psi : \mathbb{R} \times (0,1) \to \Delta^{2,\circ},$$

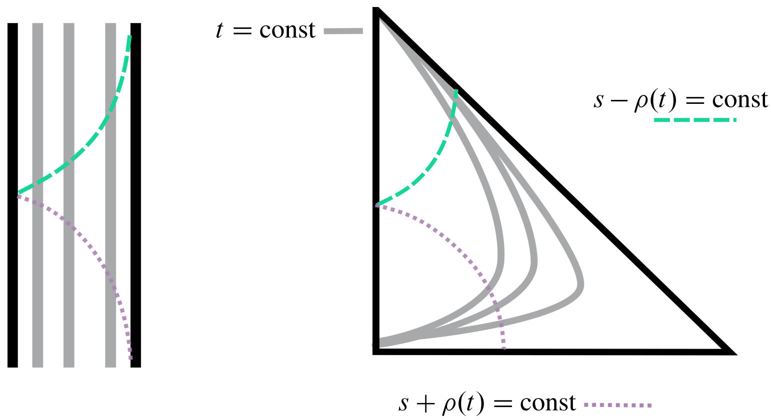


Figure 3: The source and target of the map $\psi$.

and an increasing diffeomorphism

$$\rho\colon (0,1) \to (0,\infty)$$

for which

$$\lim_{t\to 1} \psi(s+\rho(t),t) = f^0 \circ \sigma(s),$$
$$\lim_{t\to 1} \psi(s-\rho(t),t) = f^2 \circ \sigma(s),$$
$$\lim_{t\to 0} \psi(s\pm\rho(t),t) = f^1 \circ \sigma(s),$$

uniformly on compact subsets of $\mathbb{R}$. Here $f^i\colon \Delta^1 \to \partial\Delta^2$ is the standard embedding of the face missing the $i^{\text{th}}$ vertex. We extend the maps $\psi^{\pm} := \psi(\cdot \pm \rho(\cdot),\cdot)$ to the closure $\mathbb{R}\times[-1,1]$ in the obvious way. See Figure 3.

**Definition 7.1** A Floer datum $(H,J) \in \mathcal{F}^{(0)}$ is called *well-behaved* if for any $E > 0$ and any compact $K \subset M$ there is an $R = R(E,K) > 0$ such that any solution $u$ to Floer's equation

$$\partial_s u + J(\partial_t u - X_H) = 0$$

satisfying

$$E(u) := \frac{1}{2}\int \|\partial_s u\|^2 \le E, \quad u(\mathbb{R}\times S^1) \cap K \ne \varnothing$$

is contained in the ball $B_R(K)$.

A homotopy $F = (H_s, J_s) \in \mathcal{F}^{(1)}$ is called well-behaved if the corresponding condition holds for the solutions to

$$\partial_s u + J_{\sigma(s)}(t)\big(\partial_t u - X_{H_{\sigma(s)}}(t,u(s,t))\big) = 0.$$

Finally, an element $\{F_p\}_{p\in\Delta^2} \in \mathcal{F}^{(2)}$ is called well-behaved if the corresponding condition holds for the set of solutions to

$$\partial_s u + J_{\psi^\pm(s,\tau)}(t)\big(\partial_t u - X_{H_{\psi^\pm(s,\tau)}}(t, u(s,t))\big) = 0, \quad \tau \in [0,1],$$

with $R(E,K)$ independent of $\tau$. Denote by $\mathcal{F}^{(i)}_{\mathrm{wb}} \subset \mathcal{F}$ the subset consisting of well-behaved elements.

**Definition 7.2** A *Floer system* $\mathcal{D}$ on $M$ consists of the data of subsets $\mathcal{D}^{(i)} \subset \mathcal{F}^{(i)}_{\mathrm{wb}}$ for $i = 0,1,2$ such that the following hold:

(a) For any element $F \in \mathcal{D}^{(i)}$ there is an open neighborhood $F \in V \subset C^1 \times C^0$ such that $V \subset \mathcal{D}^{(i)}$.

(b) A face of an element of $\mathcal{D}^{(i)}$ is an element of $\mathcal{D}^{(i-1)}$.

(c) For any pair $F_i = (H_i, J_i) \in \mathcal{D}^{(0)}$, $i = 0,1$, such that $H_1 \geq H_0$, there is a homotopy $\{F_s\}_{s\in[0,1]} \in \mathcal{D}^{(1)}$ with endpoints $F_0$ and $F_1$.

(d) Given a pair $F', F'' \in \mathcal{D}^{(1)}$ such that $F'_1 = F''_0$, there is a $G \in \mathcal{D}^{(2)}$ whose restriction to the $\{0,1\}$ and $\{1,2\}$ faces coincides with $F'$ and $F''$, respectively.

(e) Given homotopies $F_{01}, F_{12}, F_{02} \in \mathcal{D}^{(1)}$ such that the endpoints of $F_{ij}$ are $F_i$ and $F_j$, there is a $G \in \mathcal{D}^2$ whose face $ij$ coincides with $F_{ij}$.

A Floer system $\mathcal{D}$ is said to be *invariant* if it is invariant under the action of the symplectomorphism group given by

$$\phi \cdot (H,J) = (H \circ \phi, \phi^* J).$$

Elements of $\mathcal{D}^0$ will be referred to as $\mathcal{D}$–admissible. A function $H \in C^\infty(M)$ is said to be $\mathcal{D}$–admissible if there is an almost complex structure $J$ such that $(H,J) \in \mathcal{D}$. A *bi-directed* Floer system is one in which for any admissible $H_1$ and $H_2$ there are admissible $H_3$ and $H_0$ such that

$$H_3 \geq \max\{H_1, H_2\} \quad \text{and} \quad H_0 \leq \min\{H_1, H_2\}.$$

In Theorem 7.5 below we show that on any geometrically bounded manifold there is a canonically defined invariant bidirected Floer system.

**Definition 7.3** Define the *dissipative Floer system* $\mathcal{F}_d$ to consist of the following data. Let $\mathcal{F}^{(0)}_d(M)$ be the set of i–bounded Floer data $(H,J)$ which are RLD. Let $\mathcal{F}^{(1)}_d$ be the set of monotone paths $(H_s, J_s)_{s\in[0,1]}$ in $\mathcal{F}^{(0)}$ with endpoints in $\mathcal{F}^{(0)}_d$ such that the domain-dependent Floer datum $(s,t) \mapsto (H(t,\cdot)_{\sigma(s)}\,dt, J(t,\cdot)_{\sigma(s)})$ is i–bounded as in Definition 5.9. Finally, $\mathcal{F}^{(2)}_d$ is defined as follows. Let $F_{p\in\Delta^2} = (H_p, J_p)_{p\in\Delta^2} \in \mathcal{F}^{(2)}$

with edges in $\mathcal{D}^{(1)}$. Associate to $\Delta^2$ and the map $\psi\colon \mathbb{R}\times(0,1)\to\Delta^2$ a family $C_{\tau\in[0,1]}$ of cylinders over the unit interval degenerating to a broken cylinder in the obvious way. Let the domain-dependent Floer datum on $C_\tau$ be defined by $F_\tau(s,t) = (H_{\psi(s,\tau)}(t,\cdot)\,dt, J_{\psi(s,\tau)}(t,\cdot))$. Then $F\in\mathcal{F}_d^{(2)}$ if and only if the family $C$ with this choice of domain-dependent Floer data is uniformly i–bounded as in Definition 5.10.

**Remark 7.4** The set of all well-behaved Floer data is not necessarily connected. Thus it is possible that there exist other Floer systems perhaps giving rise to inequivalent theories. However, any Floer system for which the well-behavedness property of Definition 7.1 holds in a sufficiently domain-local manner is equivalent to the dissipative system by an argument similar to the proof of Theorem 4.7.

**Theorem 7.5** *Let $(M,\omega)$ be a monotone or Calabi–Yau geometrically bounded symplectic manifold. Then $\mathcal{F}_d(M)$ is an invariant bi-directed Floer system on $M$.*

Before proving Theorem 7.5 we need the following lemma.

**Lemma 7.6** *Let $(H_i,J_i)\in\mathcal{F}_d^{(0)}(M)$ be such that $H_0\le H_1$. There exists an i–bounded monotone Floer datum on $\mathbb{R}\times S^1$ which coincides with $(H_0,J_0)$ on $\{s\ll 0\}$ and with $(H_1,J_1)$ on $\{s\gg 0\}$. Moreover, the set of such Floer data is contractible in the same sense as in Theorem 4.7.*

**Proof** To conform with Definition 5.9, it suffices to produce an almost complex structure on $[0,1]\times\mathbb{R}/\mathbb{Z}\times M$ of the form $J_{H_s}$ for some $(H_s,J_s)$, such that the following are satisfied:

- $\partial_s J_{H_s}$ vanishes identically near the boundary of $[0,1]\times\mathbb{R}/\mathbb{Z}\times M$, and thus extends to an almost complex structure on $\mathbb{R}\times\mathbb{R}/\mathbb{Z}\times M$ interpolating between $J_{H_0}$ and $J_{H_1}$. We continue to denote this extended almost complex structure by $J_{H_s}$.
- Denoting by $\pi\colon\mathbb{R}\times\mathbb{R}/\mathbb{Z}\times M$ the projection to $\mathbb{R}\times\mathbb{R}/\mathbb{Z}$, we have that the restriction of $J_{H_s}$ to each of $\pi^{-1}\big((\tfrac13,\infty)\times\mathbb{R}/\mathbb{Z}\big)$ and to $\pi^{-1}\big((-\infty,\tfrac23)\times\mathbb{R}/\mathbb{Z}\big)$ is intermittently bounded relative to $\pi$.
- $\partial_s H_s\ge 0$.

Other than the last condition, the construction would be the same as in the proof of Theorem 4.7. We show that the monotonicity requirement does not affect the proof of Theorem 4.7. For simplicity, assume $H_i$ is time-independent. As in the proof of

Theorem 4.7 fix two disjoint open sets $V_1, V_2 \subset M$ such that there is taming data for $J_{H_i}$ supported in $[0,1] \times \mathbb{R}/\mathbb{Z} \times V_i$ for $i = 0, 1$. We may assume that each of the $V_i$ is a disjoint union of precompact sets. Let $\chi\colon M \to [0,1]$ be a function which equals 0 on $V_0$ and 1 on $V_1$. Let $f\colon [0,1] \to [0,1]$ be a monotone function which is identically 0 near 0 and identically 1 on $\left[\frac{1}{3}, 1\right]$. Let $g\colon M \times [0,1] \to [0,1]$ be defined by

$$g(x,s) = f(1-s)f(s)\chi(x) + 1 - f(1-s).$$

Then $g$ is monotone increasing in $s$, identically 0 for all $x$ when $s$ is near 0, and identically 1 for all $x$ when $s$ is near 1. Take $H_s = g(x,s)H_1 + (1 - g(x,s))H_0$. Then $H_s$ is also monotone increasing in $s$. Moreover, $H$ is fixed and equal to $H_0$ on $\left[0, \frac{2}{3}\right] \times V_0$ and to $H_1$ on $\left[\frac{1}{3}, 1\right] \times V_1$. Let $J_s$ be any homotopy which is fixed and equal to $J_0$ on $\left[0, \frac{2}{3}\right]$ and to $J_1$ on $\left[\frac{1}{3}, 1\right]$. Then $J_{H_s}$ is i–bounded since it coincides with $J_{H_0}$ on $\left[0, \frac{2}{3}\right] \times V_0$ and with $J_{H_1}$ on $\left[\frac{1}{3}, 1\right] \times V_1$. Contractibility of the set of all such homotopies is similar. $\square$

**Proof of Theorem 7.5** Let $\mathcal{F}_d(M)$ be as in Definition 7.3. We verify that $\mathcal{F}_d(M)$ has all the required properties. Namely, that it is a Floer system and that it satisfies the properties guaranteed in Theorem 7.5.

**Well-behavedness** This follows from the definition and Theorem 6.3.

**Condition (a)** We need to show that if $(H, J)$ is dissipative, so is a nearby $(H', J')$. The most involved case is when $i = 2$ which we treat. Near each vertex, we have fixed Floer data so by definition we can pick an open neighborhood which maintains RLD-ness for all three of these. The property of being u.i.b. depends on the metric only up to quasi-isometry which is preserved for any uniform open neighborhood.

**Condition (b)** This follows by definition.

**Condition (c)** This is just Lemma 7.6.

**Condition (d)** Pick an $R_0 > 0$ for which $\sigma^{-1}(\operatorname{supp} \partial_s F') \subset [-R_0, R_0]$, and similarly for $F''$. For any $R > 0$ define the homotopy $I_R = F' \#_R F''$ by

$$I_{R,s} := \begin{cases} F'_{\sigma(0)} & \text{if } s \le -R - R_0, \\ F'_{\sigma(s+R)} & \text{if } s \in [-R - R_0, -R], \\ F'_{\sigma(1)} = F''_{\sigma(0)} & \text{if } s \in [-R, R], \\ F''_{\sigma(s-R)} & \text{if } s \in [R, R + R_0], \\ F''_{\sigma(1)} & \text{if } s \ge R + R_0. \end{cases}$$

Define $G$ by $G_{\psi(s,\tau)} := I_{\rho(\tau),s}$. It is immediate that $G \in \mathcal{F}_d^{(2)}$.

**Condition (e)** First merge $F_{01}$ with $F_{12}$ as in the previous part. Then homotope to $F_{02}$ relying on contractibility in Lemma 7.6.

**Invariance** Evident from the definition.

**Bi-directedness** This follows by Theorem 6.10. □

## 7.2 Transversality and control of bubbling

**Definition 7.7** Denote by $\mathcal{J}_{\mathrm{reg}}$ the set of almost complex structures for which all moduli spaces

$$\mathcal{M}^*(A; J)$$

of non-multiply-covered $J$–holomorphic spheres representing any class $A \in H^2(M;\mathbb{Z})$ are smooth manifolds of expected dimension. For $J \in \mathcal{J}_{\mathrm{reg}}$, let $\mathcal{H}_{\mathrm{reg}}(J)$ denote that set of all nondegenerate Hamiltonians satisfying the following conditions:

(a) The linearization $D_u$ of Floer's equation at a Floer trajectory $u$ is surjective for all $(H, J)$–Floer trajectories.

(b) No Floer trajectory with index difference $\leq 2$ intersects a $J$–holomorphic sphere of Chern number 0.

(c) No periodic orbit of $H$ intersects a $J$–holomorphic sphere of Chern number $\leq 1$.

Write

$$\mathcal{F}^{(0)}_{\mathrm{reg}} := \bigcup_{J \in \mathcal{J}_{\mathrm{reg}}} \mathcal{H}_{\mathrm{reg}}(J) \times \{J\}.$$

Recall that $M$ is said to be *semipositive* if for any class $A \in \pi_2(M)$ we have

$$3 - n \leq c_1(A) < 0 \implies \omega(A) \leq 0.$$

Observe that if $M$ is monotone or Calabi–Yau, then it is semipositive.

**Theorem 7.8** *Suppose $M$ is semipositive. Let $(H, J) \in \mathcal{D}^{(0)}$ and let $V \subset \mathcal{F}^{(0)}_{\mathrm{wb}}$ be an open neighborhood of $(H, J)$ in $C^\infty \cap \mathcal{F}_{\mathrm{wb}}$. Shrinking $V$, write $V = V_1 \times V_2 \subset \mathcal{H} \times \mathcal{J}$. The set $\mathcal{J}_{\mathrm{reg}}$ is of second category in $V_2$ and for each $J \in \mathcal{J}_{\mathrm{reg}} \cap V_2$, the set $\mathcal{H}_{\mathrm{reg}}(J)$ is of second category in $V_1$.*

**Remark 7.9** Theorem 7.8 is formulated for time-independent almost complex structures following [32]. The same claim holds for time-dependent Hamiltonians after appropriately modifying the regularity requirement. However, if we wish to construct homotopies of such, we need to restrict to the case where $M$ is monotone or Calabi–Yau.

See Remark 7.12 below. One reason why one would wish to work with time-dependent $J$ is that once $H$ has nondegenerate periodic orbits, for generic Floer data of the form $(H, J)$ with $J$ time-dependent the moduli space of smooth Floer trajectories is a smooth manifold of the expected dimension. See Theorem 5.1 in [23]. It then follows easily that generic such $(H, J)$ are regular.

**Proof** Since all the moduli spaces for all the Floer data in $V$ intersecting a compact set $K$ and possessing energy $E$ are contained, for some $R < \infty$, in $B_R(K)$, this follows from the compact case. For the compact case see eg [32]. $\square$

Suppose that for $i = 0, 1$, we have well-behaved elements $F_i \in \mathcal{F}^{(0)}_{\mathrm{reg}}$, and let

$$F_{01} := \{F_s = (H_s, J_s)\}_{s\in\Delta^1}$$

be a well-behaved homotopy between them.

**Definition 7.10** Call such a homotopy regular if the following hold:

(a) For any $A \in H^2(M;\mathbb{Z})$ write
$$\mathcal{M}^*(A;\{J_s\}) := \{(s,u) \mid u \in \mathcal{M}^*(A; J_{s,t})\}.$$
Then $\mathcal{M}^*(A;\{J_s\})$ is smooth and of the expected dimension.

(b) For any $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ the moduli spaces
$$\mathcal{M}(\widetilde{\gamma}_1, \widetilde{\gamma}_2, F = \{H_s, J_s\})$$
of nontrivial continuation trajectories are smooth and of the expected dimension.

(c) There is no continuation trajectory $u$ of index 0 or 1 for which there is a point $(s,t)$ such that $u(s,t)$ is in the image of a $J_s$–holomorphic sphere of Chern number 0.

Similarly, let $F \in \mathcal{F}^{(2)}$, with edges corresponding to regular homotopies. For such an $F$, $\Delta^{(2)}$ parametrizes a family of time-dependent Floer data $(H, J)$. Write $(H_{s,\lambda}, J_{s,\lambda}) := F_{\psi(s,\lambda)}$.

We say that $F$ is regular if:

(a) The moduli space $\mathcal{M}^*(A;\{J_{s,\lambda}\})$ is smooth of the expected dimension.

(b) The corresponding family $\{u_\lambda\}$ of Floer solutions is smooth of the expected dimension.

(c) There is no $\lambda$ for which there is a point $(s,t)$ and a continuation trajectory $u_\lambda$ of index $-1$ or $0$ such that $u_\lambda(s,t)$ intersects a $J_{s,\lambda}$–holomorphic sphere of Chern number 0.

Denote by $\mathcal{F}^{(i)}_{\mathrm{reg}}$ for $i=1,2$, respectively, the regular 1– and 2–simplices of Floer data.

For $i=1,2$ let $F^i \in \mathcal{F}^{(i)}_{\mathrm{wb}}$ and let $V(F^i) \subset \mathcal{F}^{(i)}_{\mathrm{wb}}$ be an open neighborhood. Let $V_2(F^i) \subset V(F^i)$ be the set of elements whose $H$ component coincides with that of $F^i$.

To achieve transversality in the definition of continuation maps, we wish to avoid perturbing $H_{01}$ since it is required to satisfy a monotonicity condition. Thus we will perturb $J_{01}$ in an $s$–dependent manner.

**Theorem 7.11** *Suppose $M$ is monotone or Calabi–Yau. Then $\mathcal{F}^{(i)}_{\mathrm{reg}} \cap V_2(F^i)$ is of second category in $V_2(F^i)$ for $i=1,2$.*

**Remark 7.12** The strengthening of the assumption relative to Theorem 7.8 is required in the case $i=2$. Indeed, in this case, the assumption of semipositivity does not rule out the possibility that for an isolated $(s,\lambda)$ there is a $J_{s,\lambda}$–holomorphic sphere with negative Chern number. Once such a sphere is present, its multiple covers interact with Floer trajectories in a nontransverse way. Invariance of Floer cohomology under homotopies of $J$ can still be established for the semipositive case by constructing chain homotopies for truncated Floer homologies, since regularity for that case is easily seen to be an open condition. We do not pursue this here.

**Proof** We need to verify that we can achieve regularity even though we avoid perturbing $H$. For the generic smoothness of the moduli spaces see Section 16 in [48]. The nonintersection property is a variation of the corresponding claim in [32]. Namely, for $i=1$, to show this is to show that the universal moduli space

$$\mathcal{N} := \{(s,z,F=(H_s,J_s),u_1,u_2 \mid F \in V_2, u_1(z)=u_2(s,t)\}$$

is a smooth separable Banach space of the expected codimension. Here $u_1 \in \mathcal{M}^*(J_{s,t})$ and $u_2$ is an $F$ continuation trajectory. For this it suffices that the evaluation map

$$\mathbb{R} \times S^1 \times S^2 \times \{u \in \mathcal{M}^*(J_s) \mid J_s \in V_2\} \to \mathbb{R} \times S^1 \times M$$

defined as

$$(s,t,z,u) \mapsto (s,t,u(z))$$

is a submersion. For this, apply Lemma 3.4.3 from [37]. For $i=2$ the argument is similar. □

## 7.3 Floer's theorem

*In this section we assume throughout that $M$ is monotone or Calabi–Yau. Moreover, we assume $M$ is connected.*

Denote by $\mathcal{L}M$ the free loop space $C^\infty(\mathbb{R}/\mathbb{Z}, M)$. Let $I_\omega, I_c\colon \pi_1(\mathcal{L}M) \to \mathbb{R}$ be given by integrating $\omega$ and the Chern class, respectively. Denote by $\widetilde{\mathcal{L}M}$ the *Floer–Novikov covering* of $\mathcal{L}M$; that is, the abelian covering space of $\mathcal{L}M$ for which $i_*\pi_1(\widetilde{\mathcal{L}M}) = \ker I_\omega \cap \ker I_c$, where $i_*\colon \pi_1(\widetilde{\mathcal{L}M}) \hookrightarrow \pi_1(\mathcal{L}M)$ is the natural inclusion. Explicitly, the space $\widetilde{\mathcal{L}M}$ is constructed as follows. For each component $\mathcal{L}M_a$ of $\mathcal{L}M$ choose a base loop $\gamma_a$. Then $\widetilde{\mathcal{L}M_a}$ consists of equivalence classes of pairs $(\gamma, A)$ such that $\gamma \in \mathcal{L}M_a$, $A$ is a homotopy class of paths in $\mathcal{L}M_a$ starting at $\gamma_a$ and ending at $\gamma$, and the equivalence relation is $(\gamma, A_1) \sim (\gamma, A_2)$ if and only if $\omega(A_1) = \omega(A_2)$ and $c_1(A_1) = c_1(A_2)$.

For a smooth function $H \in C^\infty(\mathbb{R}/\mathbb{Z} \times M)$ and for any $t \in \mathbb{R}/\mathbb{Z}$, denote by $X_{H_t}$ its Hamiltonian vector field. This is the unique vector field satisfying $dH_t(\cdot) = \omega(X_{H_t}, \cdot)$. Define a functional $\mathcal{A}_H\colon \widetilde{\mathcal{L}M} \to \mathbb{R}$ by

$$\mathcal{A}_H([\gamma, A]) := -\omega(A) - \int_0^{2\pi} H(\gamma(t))\, dt.$$

Note that this functional depends on the choice of base loop $\gamma_a$ for the connected component $a \in \pi_0(\mathcal{L}M)$.

Denote by $\mathcal{P}(H) \subset \mathcal{L}M$ the set of 1–periodic orbits of $X_H$. Denoting by

$$\pi\colon \widetilde{\mathcal{L}M} \to \mathcal{L}M$$

the covering map, set

$$\widetilde{\mathcal{P}(H)} = \pi^{-1}(\mathcal{P}(H)).$$

This is the same as the critical point set of $\mathcal{A}_H$.

We define an index

$$i_{\mathrm{RS}}\colon \widetilde{\mathcal{P}(H)} \to \mathbb{Z}$$

as follows. For each homotopy class $a \in \pi_0(\mathcal{L}M)$ fix a trivialization of $\gamma_a^* TM$. Then if $\widetilde{\gamma} = (\gamma, A)$, trivialize $\gamma^*TM$ along $A$ by extending the existing trivialization from $\gamma_a$. With respect to this trivialization, the linearization $t \mapsto D\psi_{t,\gamma(t)}$ of the flow along $\gamma$ is a path of symplectic matrices, to which is associated its Robbin–Salamon index [51]. We take $i_{\mathrm{RS}}(\widetilde{\gamma})$ to be the Robbin–Salamon index in this trivialization. Note that $i_{\mathrm{RS}}$ is independent of choices up to an integer shift $n_a$ for each $a \in \pi_0(\mathcal{L}M)$.

For each homotopy class $a \in \pi_1(M)$ let $\Gamma_a \subset \mathbb{R} \times 2\mathbb{Z}$ be the image of $\pi_1(\mathcal{L}M_a)$ under $I_\omega \times I_c$. We identify elements of $\Gamma_a$ with equivalence classes in $\pi_1(\mathcal{L}M_a)$ modulo $\ker I_\omega \cap \ker I_c$. For any ring $R$, define the Novikov ring $\Lambda_{R,\Gamma_a}$ by the set of formal sums

$$\sum_{A \in \Gamma_a} \lambda_A T^{I_\omega(A)} e^{2I_c(A)}, \quad \text{with } \lambda_A \in R,$$

which satisfy for each constant $c$ that

$$\#\{A \in \Gamma_a \mid \lambda_A \neq 0,\ \omega(A) < c\} < \infty.$$

We have an action of $\Gamma_a$ on $\widetilde{\mathcal{L}M}_a$ by

$$A \cdot [x, B] := [x, A \# B].$$

This is a covering action, so it restricts to an action on $\widetilde{\mathcal{P}H}$.

Fix a Floer system $\mathcal{D}$. Write $\mathcal{D}_{\mathrm{reg}} := \mathcal{F}_{\mathrm{reg}} \cap \mathcal{D}$ and let $F = (H, J) \in \mathcal{D}^{(0)}_{\mathrm{reg}}$. We define the Floer chain complex $CF^*(H, J; R)$ as the set of formal sums

$$\sum_{\tilde{x} \in \widetilde{\mathcal{P}(H)}} \lambda_{\tilde{x}} \langle \tilde{x} \rangle, \quad \text{with } \lambda_{\tilde{x}} \in R,$$

satisfying for each constant $c$ that

$$\#\{\tilde{x} \in \widetilde{\mathcal{P}(H)} \mid \lambda_{\tilde{x}} \neq 0,\ \mathcal{A}_H(\tilde{x}) > c\} < \infty. \tag{52}$$

$CF^*(H, J; R)$ is a graded vector space over $R$ with grading given by

$$i(\tilde{x}) := i_{\mathrm{RS}}(\tilde{x}) + n. \tag{53}$$

Here $n = \frac{1}{2}\dim M$. $CF^*(H, J; R)$ can be considered as a non-Archimedean Banach space over $R$ with its trivial valuation. The norm on $CF^*(H, J; R)$ for a linear combination of generators is given by

$$\left\| \sum_i a_i \tilde{\gamma}_i \right\| := \max_{\{i \mid a_i \neq 0\}} e^{\mathcal{A}_H(\tilde{\gamma}_i)}. \tag{54}$$

For each homotopy class $a$, the vector space $CF^{*,a}(H, J; R)$ generated by $\widetilde{\mathcal{P}_a(H)}$ is a graded Banach module over the Novikov ring $\Lambda_{R,\Gamma_a}$ via the action of $\Gamma_a$ on $\widetilde{\mathcal{P}_a(H)}$. The set $\mathcal{P}_a(H)$ noncanonically defines a basis of $CF^{*,a}(H, J)$ over $\Lambda_{R,\Gamma_a}$ by choosing a lift to $\widetilde{\mathcal{P}_a(H)}$.

Let $\Gamma \subset \mathbb{R} \times \mathbb{Z}$ be a subgroup. Denote by $\Lambda_{R,\Gamma}$ the ring

$$\Lambda_{R,\Gamma} := \left\{ \sum_i a_i T^{\lambda_i} e^{2n_i} \;\middle|\; (\lambda_i, n_i) \in \Gamma,\, a_i \in R,\, \lim_{i\to\infty} \lambda_i = \infty \right\}.$$

Let $\Gamma_\omega \subset \mathbb{R} \times \mathbb{Z}$ be the subgroup generated by $\bigcup_{a\in\pi_1(M)} \Gamma_a$. Write $\Lambda_{R,\omega} := \Lambda_{R,\Gamma_\omega}$ and $\Lambda_R := \Lambda_{R,\mathbb{R}\times\mathbb{Z}}$. Assume $R$ is a field. $\Lambda_R$ is referred to as the *universal Novikov field over $R$*. Strictly speaking $\Lambda_{R,\mathbb{R}\times\mathbb{Z}}$ is only a graded field; that is, only homogeneous elements with respect to the grading induced by projection $\mathbb{R}\times\mathbb{Z} \to \mathbb{Z}$ are invertible. Henceforth let $\mathbb{K}$ be either $\Lambda_R$ or $\Lambda_{R,\omega}$. Note that $\mathbb{K}$ carries a non-Archimedean norm induced from $\|T^\lambda\| := e^{-\lambda}$. That is,[10] $\mathrm{val}(T^\lambda) = -\lambda$.

Let

$$CF^*(H, J; \mathbb{K}) := \bigoplus_{a\in\pi_1(M)} CF^{*,a}(H, J; \Lambda_{R,\Gamma_a}) \mathbin{\widehat{\otimes}}_{\Lambda_{R,\Gamma_a}} \mathbb{K},$$

where the hat denotes completion with respect to the induced valuation.

**Remark 7.13** The approach we follow here to Floer theory over the Novikov ring is the one originally introduced by [32]. In the literature (compare [47; 50; 61]) there is a slightly different construction of the Floer chain complexes over the Novikov ring, where one tensors the space generated by $\mathcal{P}(H)$ with $\Lambda_R$, instead of passing to a covering space. In that version, the chain complexes do not have an action filtration nor a grading, but they do have a Novikov filtration over $\Lambda_R$. We do not pursue the latter approach here.

We define a linear operator $d$ on $CF^*(H, J; R)$ by counting Floer trajectories in the usual way. Namely, for any two elements

$$\tilde{x}_1, \tilde{x}_2 \in \widetilde{\mathcal{P}(H)}$$

of index difference 1, denote by $\mathcal{M}(\tilde{x}_1, \tilde{x}_2; J)$ the moduli space of Floer trajectories which at $-\infty$ are asymptotic to $\tilde{x}_1$ and at $+\infty$ to $\tilde{x}_2$, divided by the action of $\mathbb{R}$. By the inclusion $\mathcal{D}^0 \subset \mathcal{F}^{(0)}_{\mathrm{wb}}$ and Gromov–Floer compactness, $\mathcal{M}(\tilde{x}_1, \tilde{x}_2; J)$ is compact. Since $F \in \mathcal{F}^{(0)}_{\mathrm{reg}}$ we get that when the virtual dimension is 0,

$$\#\mathcal{M}(\tilde{x}_1, \tilde{x}_2; J) < \infty.$$

We thus define

$$d\tilde{x}_1 = \sum_{\tilde{x}_2 \,|\, i_{\mathrm{RS}}(\tilde{x}_2) = i_{\mathrm{RS}}(\tilde{x}_1)+1} \#\mathcal{M}(\tilde{x}_1, \tilde{x}_2; J)\langle \tilde{x}_2 \rangle.$$

[10] Caution: in many texts the convention is $\|\cdot\| = e^{-\mathrm{val}(\cdot)}$.

**Theorem 7.14** *The Floer boundary map $d$ is well defined and satisfies $d^2 = 0$.*

**Proof** We need to show that for any $\tilde{x}_1 \in \widetilde{\mathcal{P}(H)}$ we have that $d\tilde{x}_2$ satisfies the finiteness condition (52). For any $c$, let

$$A_c := \{\tilde{x}_2 \in \widetilde{\mathcal{P}(H)} \mid \mathcal{M}(\tilde{x}_1, \tilde{x}_2; J) \neq \varnothing,\ \mathcal{A}_H(\tilde{x}_2) > c\}.$$

For any $\tilde{x} \in A_c$ there is a Floer trajectory of energy at most $\mathcal{A}_H(\tilde{x}_1) - c$ connecting $\tilde{x}_1$ and $\tilde{x}_2$. Well-behavedness of $F$ thus implies that there is a compact set $K \subset M$ such that any $\tilde{x} \in A_c$ is contained in $K$. The claim now follows by Gromov–Floer compactness. Thus $d$ is well defined. That $d^2 = 0$ follows from the compact case [52; 37] since all Floer trajectories under a given energy level are contained in an a priori compact set. □

By its definition, $d$ commutes with the action of $\Lambda_{R,\Gamma_a}$ and thus induces a well-defined operator on $CF^*(H, J; \mathbb{K})$.

**Theorem 7.15** (Floer's theorem) *Let $M$ be a monotone or Calabi–Yau symplectic manifold. Let $\mathcal{D}$ be a Floer system. Then there exists a dense subsystem $\mathcal{D}_{\mathrm{reg}}$ such that:*

(a) *For any $F = (H, J) \in \mathcal{D}^{(0)}_{\mathrm{reg}}$, the graded filtered complex*

$$(CF^*(H, J; \mathbb{K}), d)$$

*is well defined.*

(b) *For any pair of elements $F_1 \leq F_2 \in \mathcal{D}^{(0)}_{\mathrm{reg}}$ we have that*

$$\mathcal{D}^{(1)}_{\mathrm{reg}}(F_1, F_2) \neq \varnothing.$$

*Associated with any homotopy $F^{12} \in \mathcal{D}^{(1)}_{\mathrm{reg}}(F_1, F_2)$ is a chain map*

$$f_{F^{12}} \colon CF^*(F^1) \to CF^*(F^2),$$

*defined by counting the corresponding rigid Floer solutions. If $F_1 - F_2 = c$ and $F^{12}$ is of the form $F^{12}_s = (H + f(s), J)$, then $f_{F^{12}}$ is the identity.*

(c) *For any triple $F_0 \leq F_1 \leq F_2 \in \mathcal{D}^{(0)}_{\mathrm{reg}}$ and elements $F_{ij} \in \mathcal{D}^{(1)}_{\mathrm{reg}}(F_i, F_j)$, the set $\mathcal{D}^{(2)}_{\mathrm{reg}}(F_{01}, F_{12}, F_{02})$ is nonempty. Any element*

$$F \in \mathcal{D}^{(2)}_{\mathrm{reg}}(F_{01}, F_{12}, F_{02})$$

*defines a chain homotopy between the map $f_{F_{02}}$ and the composition $f_{F_{12}} \circ f_{F_{01}}$ by counting the corresponding rigid Floer solutions.*

**Proof** We take as above $\mathcal{D}_{\mathrm{reg}} := \mathcal{D} \cap \mathcal{F}_{\mathrm{reg}}$. By the theorems in the previous subsection, $\mathcal{D}_{\mathrm{reg}}$ is dense in $\mathcal{D}$.

(a) This is just Theorem 7.14.

(b) Given $\tilde{x}_i \in \widetilde{\mathcal{P}(H_i)}$ for $i = 1, 2$, let $\mathcal{M}(\tilde{x}_1, \tilde{x}_2; \{H_s, J_s\})$ be the moduli space of Floer solutions for the Floer data $\{H_s, J_s\}$ with $\tilde{x}_1$ and $\tilde{x}_2$ as asymptotes. For any element

$$u \in \mathcal{M}(\tilde{x}_1, \tilde{x}_2; \{H_s, J_s\}),$$

we have the a priori estimate

$$E(u) \leq \mathcal{A}_H(\tilde{x}_1) - \mathcal{A}_H(\tilde{x}_2)$$

by Lemma 5.3 with $F = H_s\, dt$. By the assumption that $\mathcal{D}$ consists of Floer data that are well-behaved as in Definition 7.1, it follows that

$$\mathcal{M}(\tilde{x}_1, \tilde{x}_2; \{H_s, J_s\})$$

is compact. We define the continuation map by counting the 0–dimensional moduli space. Since action decreases along continuation maps, the finiteness condition is the same as the case of the differential. The fact that $f_{I^{12}}$ is a chain map is the same as in the compact case [52; 37].

(c) It follows again from well-behavedness that all Floer solutions under a given energy level for all the elements of the family are contained in an a priori compact set. The claim thus follows again from the compact case. □

**Proof of Theorem 1.1** This follows by Definition 7.3 from Theorems 7.5, 6.6 and 7.15, by passing to homology. □

# 8 Hamiltonian Floer cohomology by approximation

## 8.1 Reduced cohomology

**Definition 8.1** We refer to a chain complex which is a topological vector space with continuous differential as a *topological complex*. For a topological complex $(C, d)$, the differential is in general not closed. To stay within the realm of complete Hausdorff topological vector spaces we define the *reduced cohomology* of complete Hausdorff topological complexes $(C, d)$ by

$$\overline{H}(C, d) := \frac{\ker d}{\overline{\operatorname{im} d}},$$

where the bar denotes the closure. For general $C^*$ we first take the Hausdorff completion and then take its reduced cohomology. Note that if $C^*$ is a Banach space then $\overline{H}(C,d)$ is itself a Banach space with the induced quotient norm, which is defined by

$$\|[a]\| := \inf_{b\in[a]} \|b\|.$$

For the ensuing discussion, we consider the vector space $CF^*(H,J;\mathbb{K})$ as a vector space over $R$, forgetting the action of the Novikov parameter. For $a>0$ define $CF^*_{(-\infty,a)}(H,J;\mathbb{K})$ to be the $R$–subcomplex of $CF^*(H,J;\mathbb{K})$ generated by periodic orbits of action less than $a$. Define by $CF^*_{[a,b)}(H;\mathbb{K})$ the quotient complex

$$CF^*_{[a,b)}(H,J;\mathbb{K}) := CF^*_{(-\infty,b)}(H,J;\mathbb{K})/CF^*_{(-\infty,a)}(H,J;\mathbb{K}). \tag{55}$$

Denote by $HF^*_{[a,b)}(H,J;\mathbb{K})$ the corresponding cohomology groups. These are vector spaces over $R$.

**Theorem 8.2** *We have that*:

(a) *A continuous chain map between topological complexes induces a well-defined map on the reduced cohomologies.*

(b) *A nullhomotopic map induces the zero map on reduced cohomology.*

**Proof** For the first assertion, continuity implies that $\overline{\operatorname{im} d}$ is mapped into $\overline{\operatorname{im} d}$. For the second assertion, note that $f$ maps all cycles into $\operatorname{im} d \subset \overline{\operatorname{im} d}$. □

**Remark 8.3** A short exact sequence of topological complexes with continuous maps induces a long sequence of reduced cohomologies. However, exactness of this sequence only holds under special assumptions. A reference in the case of Hilbert complexes is [36].

Let $C^*$ be a topological complex whose topology is induced by a filtration by subcomplexes $\{C^*_t\}_{t\in\mathbb{R}}$ such that $C^*_t \subset C^*_{t'}$ whenever $t<t'$. For any element $a\in C^*$ let $\operatorname{val}(a) := \inf\{t \mid a \in C^*_t\}$. Then val naturally induces a filtration on $\overline{H}^*(C^*)$ defined by $\operatorname{val}([a]) = \inf_{c\in[a]} \operatorname{val} c$. Define a filtration on the vector space $\varprojlim_t H^*(C^*/C^*_t)$ by

$$\operatorname{val}(x) = \inf\bigl\{t_0 \mid x \in \ker\bigl(\varprojlim_t H^*(C^*/C^*_t) \to H^*(C^*/C^*_{t_0})\bigr)\bigr\}$$

for $x \in \varprojlim_t H^*(C^*/C^*_t)$. Observe that the spaces $\varprojlim_t H^*(C^*/C^*_t)$ and $H^*(C^*/C^*_{t_0})$ are both complete with respect to the norm $\|a\| := e^{\operatorname{val}(a)}$, and are thus Banach spaces.

**Theorem 8.4** *As Banach spaces,*

$$\overline{H}^*(C^*) = \varprojlim_t H^*(C^*/C^*_t).$$

**Proof** Any two cycles in the Hausdorff completion of $C^*$ representing the same element in $\overline{H}^*(C^*)$ represent the same element of $H^*(C^*/C^*_t)$ for each $t$. We thus get a well-defined morphism

$$f\colon \overline{H}^*(C^*) \to \varprojlim_t H^*(C^*/C^*_t).$$

If $c \in C^*$ is a cycle and $f(c) = 0$, then $c$ is a boundary mod $C^*_t$ for each $t$. In particular it is in the closure of the space of boundaries of $C^*$, so $[c] = 0$ in the reduced homology. Thus $f$ is injective. We now show that $f$ is surjective. Let $a \in \varprojlim_t H^*(C^*/C^*_t)$. We can compute the inverse limit by taking a subset $\{t_i\}$ of the index set $\mathbb{R}$ which is discrete, bounded above and unbounded below. In this presentation, $a$ is a sequence

$$(\dots, [a_i], [a_{i+1}], \dots, [a_0]), \quad \text{with } [a_i] \in H^*(C^*/C^*_{t_i}),$$

where $[a_i]$ maps to $[a_{i+1}]$ under the natural map induced on homology. We consider the representatives $a_i$ as living in $C^*$ and claim that they can be chosen so that, already at the chain level, $a_i$ maps to $a_{i+1} \mod C^*_{t_{i+1}}$. Inductively, suppose this holds for all $i_0 < i \le 0$. We have that there is a $b_{i_0} \in C^*$ such that $a_{i_0+1} - a_{i_0} = db_{i_0} \mod C^*_{t_{i_0+1}}$. Replace $a_{i_0}$ by $a_{i_0} + db_{i_0}$ to get the claim for $i_0$. The sequence $\{a_i\}$ converges as $i \to -\infty$ to an element $\hat{a}$ in the completion of $\widehat{C}^*$. By construction, $d\hat{a} = 0$ and $f([\hat{a}]) = a$. Unwinding definitions one verifies that $f$ preserves the valuations and is thus a Banach space isometry. □

**Example 8.5** In this example we illustrate how reduced and unreduced cohomology may differ from one another. Fix a field $R$ and consider the vector space

$$C^* = \bigoplus_{i=1}^{\infty} R\langle x^i, y^i\rangle[q]/q^2,$$

where the $x^i$ and $y^i$ are formal symbols of degree 0 and 1, respectively, for all $i$, and $q$ is a formal symbol of degree $-1$. Define a non-Archimedean valuation on $C^*$ by taking $\operatorname{val}(x^i) = 0 = \operatorname{val}(q)$ and $\operatorname{val}(y^i) = -i$. Define a differential by

$$dx^i := y^i, \quad dy^i := 0, \quad d(qx^i) = qy^i + x^{i+1} - x^i, \quad d(qy^i) = y^i - y^{i+1}.$$

Suppose $\gamma \in C^*$ is a *finite* sum of generators satisfying $d\gamma = 0$. Then $\gamma$ is a linear combination of the $y^i$ and elements of the form $qy^i - x^i + x^{i+1}$; that is, a linear combination of $dx^i$ and $d(qx^i)$. Thus the homology of $C^*$ vanishes. Consider now the complex $\widehat{C}^*$ obtained by completing $C^*$ with respect to the valuation, and let

$$\gamma := x^1 + \sum_{i=1}^{\infty} (-1)^i q y^i.$$

Then $\gamma$ is well defined in $\widehat{C}^*$, and $d\gamma = 0$. We show that $\gamma$ is not a coboundary and is not even approximated by a sequence of coboundaries. For this, consider its image in $C^*/C^*_t$. It is straightforward to verify that the class of $\gamma$ is equivalent mod $C^*_t$ to the class of $x^i$ for any $i > -t$, and that $\operatorname{val}([x^i]) = 0$. In this case one verifies that the image of the differential is closed. It follows that the reduced homology of the completion $\widehat{C}^*$ (which in this case equals the ordinary homology) is nonzero.

Consider now the subcomplex $\widehat{R\langle x^i, y^i\rangle}$ of $\widehat{C}^*$. Let

$$\gamma = \sum y^i.$$

Then $\gamma$ is again a convergent sum. Moreover, for any $t$ we have that $\gamma$ is a boundary mod $C^*_t$. Indeed, for any $N > t$ we have

$$\gamma = d \sum_{i=1}^{N} x^i \mod C^*_t.$$

But the sum on the right-hand side does not converge as $i \to \infty$. One verifies in this case that the reduced cohomology $\overline{H}^*(\widehat{R\langle x^i, y^i\rangle})$ vanishes while the unreduced one does not.

## 8.2 Floer cohomology of lower semicontinuous exhaustion functions

**Theorem 8.6** *Let $(H_0, J_0)$ and $(H_1, J_1)$ be dissipative. Suppose*

$$H_1 - c \le H_0 \le H_1. \tag{56}$$

*Then the canonical continuation map $HF^*(H_0, J_0) \to HF^*(H_1, J_1)$ is an isomorphism which decreases norms by a factor of at most $e^{-c}$. In particular, when $H_0 = H_1$, the continuation map is an isometry.*

**Proof** Recall that for a monotone homotopy, the induced continuation map is valuation decreasing. Consider the composition of continuation maps

$$HF^*(H_0, J_0) \to HF^*(H_1, J_1) \to HF^*(H_0 + c, J_0).$$

It coincides with the continuation map

$$HF^*(H_0, J_0) \to HF^*(H_0 + c, J_0).$$

The latter stems from a naive identification of the underlying complexes with the norm scaled by $e^{-c}$. This shows that the map $HF^*(H_0, J_0) \to HF^*(H_1, J_1)$ is right-invertible and decreases norm by at most $e^{-c}$. Left-invertibility is shown similarly. □

Henceforth we drop $J$ from the notation and talk about $HF^*(H)$. Abusing notation we will also drop $J$ from the chain level notation. Accordingly, we refer to a Hamiltonian $H$ as dissipative if there exists a compatible almost complex structure $J$ such that $(H, J)$ is dissipative.

As a consequence of Theorem 8.6 we may extend the definition of Floer cohomology to some Hamiltonians which are degenerate or even nonsmooth.

**Lemma 8.7** *Suppose that $H_i$ and $F_i$ are pointwise monotone increasing sequences of nondegenerate dissipative Hamiltonians, both converging uniformly in $C^0$ to the same continuous function $H$. Then there is an isomorphism*

$$\varinjlim_i HF^*(H_i) \to \varinjlim_i HF^*(F_i),$$

*which is natural in that it commutes with all continuation maps involving dissipative and nondegenerate Hamiltonians. We may thus define*

$$HF^*(H) := \varinjlim_n HF^*(H_n). \tag{57}$$

*Define a seminorm on $HF^*(H)$ by*

$$\|a\| := \inf_i \|a_i\|, \tag{58}$$

*where $a_i \in HF^*(H_i)$ maps to $a$ under the natural map. Then $\|\cdot\|$ is a non-Archimedean seminorm on $HF^*(H)$ which is independent of the choice of $H_i$. Moreover, when $H$ is smooth, dissipative and nondegenerate, the two definitions of $HF^*(H)$ as a seminormed space coincide.*

**Proof** Call a sequence $H_n$ as in the hypothesis *admissible* if for each $n$ there is a constant $c_n > 0$ for which $H - H_n \geq c_n$. Given any two admissible sequences we can squeeze a subsequence of one into a subsequence of the other. The first part of the statement then follows by the universal property of the direct limit. Given a not necessarily admissible monotone sequence $H_n$ converging to $H$, the sequence $H_{nk} := H_n - 1/k$ is admissible, monotone and converges uniformly to $H_n$. We thus have a natural isomorphism

$$\varinjlim_n HF^*(H_n) = \varinjlim_n \varinjlim_k HF^*(H_{nk}).$$

But the sequence $H_{nn}$ is admissible and cofinal in the doubly indexed sequence $H_{nk}$, and so we have a natural isomorphism

$$\varinjlim_n \varinjlim_k HF^*(H_{nk}) = \varinjlim_n HF^*(H_{nn}).$$

We have a similar relation for $F_{nk} := F_n - 1/k$ and $F_{nn}$. The sequences $H_{nn}$ and $F_{nn}$ are admissible, monotone and converge uniformly to $H$. Combined with the isomorphisms we just deduced, we obtain the isomorphism

$$\varinjlim_n HF^*(H_n) = \varinjlim_n HF^*(H_{nn}) = \varinjlim_n HF^*(F_{nn}) = \varinjlim HF^*(F_n),$$

where all the isomorphisms are natural.

To see that $\|\cdot\|$ defines a non-Archimedean seminorm, note that by Theorem 8.6, the sequence $\|a_i\|$ is bounded below. Since it is monotone decreasing, it is convergent. So

$$\|a+b\| = \lim_i \|a_i + b_i\| \le \lim_i \max\{\|a_i\|, \|b_i\|\} = \max\{\|a\|, \|b\|\}.$$

The homogeneity of $\|\cdot\|$ is obvious. In light of Theorem 8.6, the argument for the independence of val on the choice of sequence is similar to the claim concerning the natural isomorphism. Finally, for the last part of the claim, take $F_n$ to be the constant sequence $F_n = H$. □

The definition of action-truncated Floer homology groups also extends.

**Lemma 8.8** *Let $H$ and $H_n$ be smooth nondegenerate and dissipative, and suppose the sequence $H_n$ is monotone and converges uniformly to $H$. Then the natural map*

$$\varinjlim_n HF^*_{[a,b)}(H_n) \to HF^*_{[a,b)}(H) \tag{59}$$

*is an isomorphism. If we drop the assumption that $H$ is dissipative and define $HF^*_{[a,b)}(H)$ by* (59), *the right-hand side is independent of the choice of $H_n$.*

*Caution*: the claim does not necessarily hold if we consider other segments such as $(a,b]$.

**Proof** As in the proof of Lemma 8.7, the first part is proven by squeezing a sequence of the form $H_n - c_{nk}$ into a sequence $H - c_n$. The second part is proven by a similar squeezing. The argument is spelled out in the proof of Lemma 8.7. $\square$

The next theorem is key for what follows. It shows that truncated Floer homology is continuous with respect to convergence on compact sets.

**Theorem 8.9** *Let $\{H_n\}$ be a monotone increasing sequence of dissipative Hamiltonians converging pointwise to a dissipative Hamiltonian $H$. Then for any real $a < b$, we have that the natural map*

$$f : \varinjlim_i HF^*_{[a,b)}(H_n) \to HF^*_{[a,b)}(H)$$

*is an isomorphism.*

**Proof** Fix an almost complex structure $J$ for which $(H, J)$ and $(H_i, J)$ are dissipative. Without loss of generality we may assume that all the involved Floer data are regular and nondegenerate. As in Lemma 8.8 we reduce to the case where $H - H_n \geq c_n > 0$ for some $c_n > 0$. By Dini's theorem, the $H_i$ converge to $H$ uniformly on compact sets. By squeezing in an appropriate sequence we may assume that there is an exhaustion of $M$ by compact sets $K_n$ such that $H_n = H - c_n$ on $K_n$. For such a sequence we have that for a fixed real number $E > 0$ and compact set $K$, the numbers $R(E, K)$ of Theorem 6.3, defined for each of the $H_n$, stabilize as $n \to \infty$. So, given an $i$ and a cocycle $\gamma \in CF^*_{[a,b)}(H_i)$, there is a compact set $K$ and an $i_0 > i$ such that any continuation trajectory $f_{i,i'}(\gamma)$ or $f_i(\gamma)$ is contained mod $CF^*_{(-\infty,a)}$ in $K$. Here $f_i : CF^*_{[a,b)}(H_i, J) \to CF^*_{[a,b)}(H, J)$ and $f_{i,i'} : CF^*_{[a,b)}(H_i, J) \to CF^*_{[a,b)}(H_{i'}, J)$ are the natural continuation maps. Indeed, since we are considering only trajectories of energy less than $b - a$, Theorem 6.3 provides an estimate on the diameter as required. The same claim holds for composite trajectories of the form $d \circ f_{i,i'}$ and $d \circ f_i$, etc.

Since $H_i|_{K_i} = H|_{K_i} - c_i$, we may identify those periodic orbits of $H_i$ which are inside $K_i$ with the periodic orbits of $H$ in the same region. For each periodic orbit $\gamma$ of $H$, the corresponding periodic orbit of $H_i$ is mapped mod $a$ by the continuation map to $\gamma$. Indeed, for $i$ large enough, any continuation trajectory emanating from $\gamma$ and having energy at most $b - a$ is contained in $K_i$ and so satisfies a translation-invariant equation. To be rigid it must be trivial. Moreover, taking $i$ still larger, the same claim is true for the periodic orbits appearing in the expansion of $d\gamma$ mod $i$. This shows that $f$ is surjective. Injectivity follows in the same way. Namely, suppose there is an $i$ and a $\delta \in CF^*_{[a,b]}(H)$ such that $f_i(\gamma) = d\delta$ mod $a$ in $CF^*(H)$. For $i$ large enough, the same relation will hold in $CF^*(H_i)$. $\square$

We use the notation $\overline{HF}^*(H)$ for the reduced cohomology of $CF^*(H,J)$.

**Corollary 8.10** *For any dissipative Hamiltonian $H$ and any sequence $H_i$ of dissipative Hamiltonians converging to $H$ uniformly on compact sets, the natural map*

$$\varprojlim_{a} \varinjlim_{b,i} HF^*_{[a,b)}(H_i) \to \overline{HF}^*(H) \tag{60}$$

*is an isomorphism. Moreover, we have for any $\alpha \in \overline{HF}^*(H)$ that*

$$\operatorname{val}(\alpha) = \inf_a \bigl\{\alpha \in \ker\bigl(\overline{HF}^*(H) \to \varinjlim_{b,i} HF^*_{[a,b)}(H_i)\bigr)\bigr\}, \tag{61}$$

*upgrading the isomorphism* (60) *to an isometry of Banach spaces.*

**Proof** We have by Theorem 8.9 and by exactness of the direct limit,

$$\varinjlim_{b,i} HF^*_{[a,b)}(H_i) = \varinjlim_{b} HF^*_{[a,b)}(H) = HF^*_{[a,\infty)}(H). \tag{62}$$

So by Theorem 8.4 we obtain the isomorphism of Banach spaces. □

The last corollary will allow us to extend the definition of reduced Floer homology to arbitrary lower semicontinuous exhaustion functions; that is, a lower semicontinuous function $H\colon \mathbb{R}/\mathbb{Z} \times M \to \mathbb{R}$ which is proper and bounded from below. But first we need to formulate an approximation lemma for such functions.

**Lemma 8.11** *Let $H\colon \mathbb{R}/\mathbb{Z}\times M \to \mathbb{R}\cup\{\infty\}$ be a lower semicontinuous function which is proper and bounded below. Let $F\colon \mathbb{R}/\mathbb{Z} \times M \to \mathbb{R}$ be a smooth proper exhaustion function such that $F < H - \epsilon$ pointwise for some $\epsilon > 0$. Then there is a pointwise monotone sequence of smooth exhaustion functions $H_n\colon \mathbb{R}/\mathbb{Z} \times M \to \mathbb{R}$ such that $H_n$ converges pointwise to $H$ everywhere and such that for each $n$ there is a compact set $K_n$ such that $H_n|_{M\setminus K_n} = F|_{M\setminus K_n}$.*

**Proof** It is a standard fact that $H$ is the supremum of a monotone increasing sequence $H'_n$ of smooth functions. For each $n$ take $H_n$ to coincide with $H'_n$ on $H^{-1}((-\infty,n))$, to equal $n$ on the set

$$\{(t,x) : H_t(x) \geq n \geq F_t(x)\},$$

and to equal $F$ on $F^{-1}(n,\infty)$. Then $H_n$ is well defined and continuous. After a slight perturbation it is smooth and satisfies all the requirements. □

By Theorem 6.6 we can take $F$ in the previous lemma to be a function with sufficiently small Lipschitz constant outside of a compact set so as to be dissipative. Then all the functions in the sequence $H_n$ are also dissipative, as they coincide with $F$ outside of a compact set. Thus, *any lower semicontinuous exhaustion function is the pointwise limit of a monotone sequence of dissipative Hamiltonians*.

**Lemma 8.12** *Let $H$ be a lower semicontinuous exhaustion function. Let $H_n \le F_n \le H$ be a pair of monotone sequences of dissipative Hamiltonians each converging pointwise to $H$. Then for any $a < b \in \mathbb{R}$, the natural continuation map*

$$\varinjlim_i HF^*_{[a,b)}(H_i) \to \varinjlim_i HF^*_{[a,b)}(F_i) \tag{63}$$

*is an isomorphism.*

**Proof** For each $n$ choose monotone sequences $H_{kn}$ and $F_{kn}$ such that the following properties hold. First, for fixed $n$, they converge on compact sets to $H_n$ and $F_n$, respectively, as $k \to \infty$. Secondly, there is an exhaustion of $M$ by precompact sets $U_k$ such that $H_{kn}$ and $F_{kn}$ coincide with $H_1 - 1/k$ on the complement of $U_k$. Such sequences exist by Lemma 8.11. By Theorem 8.9 we have natural isomorphisms

$$\lim_{k\to\infty} HF^*_{[a,b)}(H_{kn}) = HF^*_{[a,b)}(H_n),$$

and a similar isomorphism relating the Floer cohomologies of $F_{kn}$ and $F_n$. The map appearing in equation (63) corresponds under this isomorphism to the natural map

$$\lim_{k,n\to\infty} HF^*_{[a,b)}(H_{kn}) \to \lim_{k,n\to\infty} HF^*_{[a,b)}(F_{kn}).$$

For each $k$ we have that $H - H_{kn}$ and $H - F_{kn}$ are bounded away from 0. Moreover, for each compact set $K$ we can make $H_{kn}$ and $F_{kn}$ arbitrarily close to $H$ on $K$ by adjusting $k$ and $n$. It follows that for each $k$ and $n$, we can find numbers $k'$, $n'$, $k''$ and $n''$ such that $F_{k''n''} > H_{k'n'} > F_{kn}$. Thus we can squeeze a cofinal subsequence of $H_{kn}$ into a cofinal subsequence of $F_{kn}$. The claim follows by the same argument as in Lemma 8.7. □

**Lemma 8.13** *Let $H$ be a lower semicontinuous exhaustion function. Let $F_n$ and $G_n$ be a pair of monotone sequences of dissipative Hamiltonians each converging pointwise to $H$. Then there exists a monotone sequence of dissipative Hamiltonian $H_n$ such that $H_n \le \min\{F_n, G_n\}$ and $H_n$ converges pointwise to $H$.*

**Proof** The function $H'_n = \min\{F_n, G_n\}$ is continuous. As in Lemma 8.11, let $H''_n$ be a continuous function which coincides with $H'_n$ on $H_n'^{-1}(-\infty, n)$, and with some fixed smooth dissipative function $F \le H_1$ everywhere else. Then $H''_n$ is continuous, the sequence $H''_n$ is monotone, and it still converges to $H$. Finally, replace $H''_n$ by a smooth $H_n$ satisfying $H''_n - 1/n \le H_n \le H''_n$ and which equals $F$ outside some compact set. Then all the requirements are satisfied. $\square$

**Proof of Theorem 3.3** The surjectivity statement follows from Lemma 8.11 as stated in the paragraph right after the proof of Lemma 8.11.

For a pointwise monotone sequence $\{H_i\}$ of regular dissipative Hamiltonians, define

$$\overline{HF}^*(\{H_i\}) := \varprojlim_a \varinjlim_{b,i} HF^*_{[a,b)}(H_i), \tag{64}$$

with the norm given by the right-hand side of (61).

Let $\sup(\{H^1_i\}) = \sup(\{H^2_i\}) = H$ for some $H \in \mathcal{H}_{\mathrm{sc}}$. By Lemma 8.13 we can find a third sequence $\{H_i\}$ of regular dissipative Hamiltonians such that $H_i \le \min\{H^1_i, H^2_i\}$. By Lemma 8.12 and equation (64) it follows that we have natural isomorphisms

$$\overline{HF}^*(H^1_i) = \overline{HF}^*(H_i) = \overline{HF}^*(H^2_i).$$

For an arbitrary $H \in \mathcal{H}_{\mathrm{sc}}$ we define $\overline{HF}^*(H)$ as the pushout over all approximating sequences $\{H_i\}$ of $\overline{HF}^*(\{H_i\})$ under the natural isomorphisms just described. By naturality we get an induced functorial continuation map for $H_1 \le H_2$. This defines the functor $\overline{HF}^*$ on $(\mathcal{H}_{\mathrm{sc}}, \le)$. To see that the restriction to $(\mathcal{H}_{\mathrm{d,reg}}, \le)$ agrees with the previous definition, note that any element $H \in \mathcal{H}_{\mathrm{d,reg}}$ can be considered as a constant sequence $\{H_i\}$ with $H_i = H$. $\square$

In fact we have proven the following stronger lemma, which is used below.

**Lemma 8.14** *Let $H$ be a lower semicontinuous exhaustion function. Let $H_n$ and $F_n$ be a pair of monotone sequences of dissipative Hamiltonians each converging pointwise to $H$. Such sequences are guaranteed to exist by Lemma 8.11. Then for any segment $[a,b)$ there exists an isomorphism*

$$\varinjlim_i HF^*_{[a,b)}(H_i) \to \varinjlim_i HF^*_{[a,b)}(F_i).$$

*This isomorphism is natural in the sense that it commutes with all induced continuation maps. We thus define*

$$HF^*_{[a,b)}(H) := \varinjlim_i HF^*_{[a,b)}(H_i). \tag{65}$$

*In other words, we have*

$$\overline{HF}^*(H) = \varprojlim_a \varinjlim_b HF^*_{[a,b)}(H). \tag{66}$$

In the next section we will see that it is actually possible to define $\overline{HF}^*$ as the reduced cohomology of an appropriate chain complex which is well defined up to filtered quasi-isomorphism.

## 8.3 The chain level construction

We apply the telescope construction appearing in [5] to define $HF^*(H)$, for general lower semicontinuous $H$, as the cohomology of a certain complex. Namely, let $(H_i, J_i)$ be a sequence of dissipative Floer data. Let $q$ be a formal variable of degree $-1$ satisfying $q^2 = 0$. Write[11]

$$SC^*(\{H_i\}) := \bigoplus_{i=1}^{\infty} CF^*(H_i)[q],$$

and equip it with the differential

$$\delta(a + qb) := (-1)^{\deg a} da + (-1)^{\deg b}(qdb + \kappa(b) - b),$$

where $\kappa$ denotes the continuation map $CF^*(H_i) \to CF^*(H_{i+1})$ for each $i$. Let $\widehat{SC}^*(\{H_i\})$ denote the completion with respect to the action filtration. It is shown in [5] that, ignoring topology, there is a natural isomorphism

$$\varinjlim_i HF^*(H_i) = H^*(SC^*(\{H_i\}), \delta). \tag{67}$$

This isomorphism arises as follows. Consider the underlying complexes $CF^*(H_i)$ with differential $\delta(a) := (-1)^{\deg a} d(a)$. This change does not affect anything at the cohomology level, and continuation maps remain chain maps. The obvious embeddings $(CF^*(H_i), \delta) \hookrightarrow SC^*(\{H_i\})$ commute up to homotopy with the continuation maps, thus giving rise to the map in (67). For more details see [5].

**Definition 8.15** Let $(C_i^*, d)$ for $i = 1, 2$ be complexes filtered by a valuation. We say that a valuation-decreasing chain map $f : C_1^* \to C_2^*$ is a *filtered quasi-isomorphism* if it induces an isomorphism on filtered homologies $H^*_{[a,b)}$ for $a > -\infty$. We say that $(C_1^*, d)$ is filtered quasi-isomorphic to $(C_2^*, d)$ if there is a zigzag of filtered quasi-isomorphisms starting at one and ending at the other.

**Theorem 8.16** *Let $H$ be a lower semicontinuous exhaustion function and let $(H_i^1, J_i^1)$ and $(H_i^2, J_i^2)$ be monotone increasing sequences of dissipative Floer data such that $H_i^j$ converges to $H$ pointwise for $j = 1, 2$. Then $\widehat{SC}^*(\{H_i^1\})$ is filtered quasi-isomorphic*

[11] As usual we abuse notation, omitting mention of the almost complex structures.

*to $\widehat{SC}^*(\{H_i^2\})$. If $H$ is itself dissipative, they are both filtered quasi-isomorphic to $\widehat{CF}^*(H)$.*

The proof of Theorem 8.16 is carried out after establishing the following few lemmas, which are of interest in their own right.

**Lemma 8.17** *We have for any interval $-\infty < a < b \leq \infty$,*

$$H^*_{[a,b)}(\widehat{SC}^*(\{H_i\}), \delta) = \varinjlim_i HF^*_{[a,b)}(H_i).$$

*Further, we have an isometry of Banach spaces*

$$\overline{H}^*(\widehat{SC}^*(\{H_i\}), \delta) = \varprojlim_a \varinjlim_i HF^*_{[a,\infty)}(H_i),$$

*where on the right-hand side we define the norm by equation* (61), *and on the left-hand side we take the norm induced from the $CF^*(H_i)$.*

**Proof** The first half of the claim is what is shown in [5], since no topology is involved. For the second half, the isomorphism of topological vector spaces follows from the first half and Theorem 8.4. The fact that this is an isometry also follows from the first half by unwinding definitions. □

**Lemma 8.18** *Let $F^0 = \{(H_i^0, J_i^0)\}$ and $F^1 = \{(H_i^1, J_i^1)\}$ be two monotone sequences of dissipative Floer data such that $H_i^0 \leq H_i^1$. Let $\mathfrak{H}_{i,s}$ be a monotone dissipative interpolating family. Then there is a filtration-decreasing continuation map*

$$\phi_{\mathfrak{H}}\colon \widehat{SC}^*(\{H_i^0\}) \to \widehat{SC}^*(\{H_i^1\}),$$

*inducing, for each interval $[a,b)$, the canonical continuation map*

$$\varinjlim_i HF^*_{[a,b)}(H_i^0) \to \varinjlim_i HF^*_{[a,b)}(H_i^1).$$

*If $\mathfrak{H}^1$ and $\mathfrak{H}^2$ are two homotopies interpolating between $F^0$ and $F^1$, there exists a filtration-decreasing chain homotopy operator*

$$\mathfrak{K}\colon \widehat{SC}^*(\{H_i^0\}) \to \widehat{SC}^{*+1}(\{H_i^1\})$$

*such that*

$$\phi_{\mathfrak{H}^1} - \phi_{\mathfrak{H}^2} = \delta \circ \mathfrak{K} + \mathfrak{K} \circ \delta.$$

**Proof** Let $j_i\colon CF^*(H_i^0) \to CF^{*+1}(H_{i+1}^1)$ be the chain homotopy operator satisfying

$$f_{\mathfrak{H}_{i+1}} \circ \kappa - \kappa \circ f_{\mathfrak{H}_i} = d \circ j_i + j_i \circ d.$$

Define

$$\phi\colon \widehat{SC}^*(F^0) \to \widehat{SC}^*(F^1)$$

by $\phi(a+qb) := f(a) + qf(b) + j(b)$. One verifies that this is indeed a chain map. Inspecting isomorphism (67) one finds that the homology-level square

$$\begin{tikzcd}
\varinjlim HF^*_{[a,b)}(H_i^0) \arrow[r] \arrow[d] & \varinjlim HF^*_{[a,b)}(H_i^1) \arrow[d] \\
H^*_{[a,b)}(SC^*(\{H_i^0\})) \arrow[r] & H^*_{[a,b)}(SC^*(\{H_i^1\}))
\end{tikzcd}$$

commutes. This proves the first half of the claim. To define the chain homotopy operator $\mathfrak{K}$, let $l_i\colon CF^*(H_i^1) \to CF^{*+1}(H_i^2)$ be the chain homotopy operator associated to a family of homotopies interpolating between $\mathfrak{H}^1$ and $\mathfrak{H}^2$. Let $j_i^n$ be the homotopies between $f_{i+1}^n \circ \kappa$ and $\kappa \circ f_i^n$ for $n = 1, 2$. Let

$$m_i\colon CF^*(H_i^1) \to CF^{*+2}(H_{i+1}^2)$$

be a degree 2 operator satisfying

$$d \circ m + m \circ d = j^1 + \kappa \circ l - (l \circ \kappa + j^2).$$

We show that such an $m$ exists before proceeding. To see this note that each term on the right-hand side is a chain homotopy operator from $\kappa \circ f_1$ to $f_2 \circ \kappa$ coming from appropriate one-dimensional families of interpolating homotopies. By standard Floer theoretic machinery, a generic two-dimensional family interpolating these one-dimensional homotopies gives rise to an operator $m$ as required. By energy considerations, $m$ is action decreasing.

Having established the existence of $m$, we define the chain homotopy

$$\mathfrak{K}(a+qb) := (-)^{\deg(a+1)} l(a) + (-1)^{\deg(b+1)}(ql(b) + m(b)).$$

A straightforward but somewhat tedious calculation shows that $\mathfrak{K}$ is indeed a chain homotopy operator, as required. □

**Proof of Theorem 8.16** Use Lemma 8.13 to find a monotone sequence of dissipative Hamiltonians $\{H_i^0\}$ dominated by $\{H_i^j\}$ for $j = 1, 2$ and still converging pointwise to $H$. By Lemmas 8.12 and 8.18, the continuation map induces a quasi-isomorphism for each finite truncation. □

# 9 The product structure

## 9.1 Floer data for the pair-of-pants product

For time-dependent Hamiltonians $H_1, H_2$, let $H_1 * H_2 \colon \mathbb{R}/\mathbb{Z} \times M \to \mathbb{R}$ be the time-dependent function

$$(H_1 * H_2)_t = \begin{cases} 2H_1(t) & \text{if } t \in \left[0, \frac{1}{2}\right), \\ 2H_2(2t-1) & \text{if } t \in \left[\frac{1}{2}, 1\right). \end{cases}$$

Note that $H_1 * H_2$ depends discontinuously on $t$ with jump discontinuities at $t = \frac{1}{2}, 0 \sim 1$. The operation is introduced for notational convenience. A triple $(H_0, H_1, H_2)$ is called a (strict) product triple if $H_2 \geq H_1 * H_0$ ($H_2 > H_1 * H_0$).

Denote by $\Sigma$ the pair of pants $S^2 \setminus \{0, 1, \infty\}$. For our convenience we pick cylindrical ends which extend globally as follows. Consider the holomorphic map $\psi \colon \Sigma \to \mathbb{R} \times \mathbb{R}/\mathbb{Z}$ given by

$$z \mapsto \frac{1}{2\pi} \operatorname{Log} z(z-1).$$

This defines cylindrical coordinates

$$s = \frac{1}{2\pi} \log |z(z-1)|, \quad t = \frac{1}{2\pi} \arg(z(z-1))$$

in punctured neighborhoods of 0 and 1, coordinatized as inputs. For the cylindrical end at $\infty$ we take

$$s = \frac{1}{2\pi} \log |z(z-1)|, \quad t = \frac{1}{4\pi} \arg(z(z-1)).$$

Thus $\infty$ is coordinatized as an output. Henceforth we write

$$\alpha_\Sigma := \frac{1}{2\pi} d \arg z(z-1)$$

and take $h_\Sigma \colon \Sigma \to \mathbb{R}$ to be the function

$$z \mapsto s = \frac{1}{2\pi} \log |z(z-1)|.$$

Then $dh_\Sigma \wedge \alpha_\Sigma \geq 0$, and $h_\Sigma$ has a single critical point at $z = \frac{1}{2}$. Note also that at the output we have $\alpha_\Sigma = 2\, dt$, while at each input we have $\alpha_\Sigma = dt$. We consider the coordinate $t$ to be well- defined on the complement of $s = h_\Sigma = \frac{1}{2}$.

**Definition 9.1** A Floer datum $(H, J)$ is called superdissipative if for any $f \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ we have that $(fH, J)$ is dissipative.

**Lemma 9.2** *If $H$ is a Lipschitz exhaustion function and $h\colon \mathbb{R} \to \mathbb{R}$ is a monotone function satisfying* $\lim_{t\to\infty} h'(t) \to 0$, *then $h \circ H$ is superdissipative.*

**Proof** The Lipschitz constant of $f \cdot h \circ H$ is arbitrarily small outside of a sufficiently large compact set. The claim follows by Theorem 6.6. $\square$

For a superdissipative $(H, J)$, let $\mathcal{F}(H, J)$ be the set of all pairs in $C^\infty(\mathbb{R}/\mathbb{Z} \times M) \times \mathcal{J}$ which coincide outside of some compact set with $(fH, J)$ for some $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}_+$. Drop $J$ from the notation when there is no ambiguity. A 1–form $\mathfrak{H} \in \Omega^1(\Sigma, C^\infty(M))$ is called *$H$–admissible* if there is a function $G\colon \Sigma \times M \to \mathbb{R}$ such that $\mathfrak{H} = G \otimes \alpha_\Sigma$ and such that for each $x \in M$ we have $d\mathfrak{H}(x) \geq 0$, and for each $z \in \Sigma$ we have $G(z, \cdot) \in \mathcal{F}(H)$. A Floer datum $(\mathfrak{H}, J')$ is called $H$–admissible if, in addition, $J'$ is quasi-isometric to $J$. For an $H$–admissible product triple we denote by $\mathcal{P}(H_0, H_1, H_2)$ the set of $H$–admissible data on the pair of pants which for $i = 0, 1, 2$ equals $H_i\, dt$ at the $i^{\text{th}}$ end. We refer to the set $\mathcal{P}(H_0, H_1, H_2)$ as *product data* for the triple $(H_0, H_1, H_2)$. Included in the set $\mathcal{P}(H_0, H_1, H_2)$ are broken Floer data, which are concatenations of monotone continuation data in $\mathcal{F}(H)$ with $H$–admissible pairs of pants.

**Lemma 9.3** (a) *$H$–admissible one-forms satisfy the hypotheses of Lemma 5.3.*

(b) *$H$–admissible product data are dissipative.*

(c) *If $(H^0, H^1, H^2) \in \mathcal{F}(H)^3$ is a strict product triple, then*

$$\mathcal{P}(H^0, H^1, H^2) \neq \varnothing.$$

(d) *$\mathcal{P}(H^0, H^1, H^2)$ is connected, and the path connecting any two elements is dissipative.*

**Proof** (a) Equations (17) and (18) hold by construction.

(b) Loopwise dissipativity follows by the assumption of superdissipativity of $H$. As for i–boundedness, observe that the metric $g_{J_\mathfrak{H}}$ is uniformly equivalent to the metric $g_{J_H}$. In fact, if $H$ is Lipschitz, $g_{J_\mathfrak{H}}$ is uniformly equivalent to the product metric on $\Sigma \times M$.

(c) Let $F \in \mathcal{F}(H)$ be a time-independent Hamiltonian satisfying for all $x \in M$

$$2\max\{H_0^0(x), H_0^1(x)\} < F(x) < \min\{H_0^2(x), H_{1/2}^2(x)\}.$$

Such a Hamiltonian exists by assumption. Let

$$G'_t(x) := \begin{cases} \max\{2H_{2t}^0(x), F(x)\} & \text{if } t \in \left[0, \tfrac{1}{2}\right], \\ \max\{2H_{2t}^1(x), F(x)\} & \text{if } t \in \left[\tfrac{1}{2}, 1\right]. \end{cases}$$

Then $G'$ satisfies

(i) $(H^0, H^1, G')$ is a product triple,

(ii) for each $x \in M$ there is a neighborhood $U \subset M$ and $I \subset \mathbb{R}/\mathbb{Z}$ of $\{0, \frac{1}{2}\}$ such that $G'_t(y) = F(x)$ for $(t, y) \in I \times U$, and

(iii) $G' < H_2$.

The function $G'$ has nonsmooth points away from $t \in \{0, \frac{1}{2}\}$ but it can be smoothed to a function $G$ so that the properties (i), (ii) and (iii) still hold. Let $G^s_t$ be a family such that for $s \gg \frac{1}{2}$ we have $G^s = H_2$ and for $s \le \frac{1}{2}$ we have $G^s = G$, and such that $\partial_s G^s \ge 0$. This can again be pieced together by an appropriate Urysohn function.

We define $H'_z$ to equal $G_{h_\Sigma(z),t}$ for $z$ such that $h_\Sigma(z) \ge \frac{1}{2}$. By property (ii), $H'_z$ extends smoothly beyond the branchpoint at $z = \frac{1}{2}$. On each input end we can smoothly interpolate between the input $H_i$ and $\frac{1}{2}G_s$ in a monotone way by employing an Urysohn function. The result $\mathfrak{H} := H' \otimes \alpha_\Sigma$ is an $H$–admissible product datum.

(d) If $H^1\alpha$ and $H^2\alpha$ are two (unbroken) $H$–admissible product data with the same inputs and output, so is any convex combination. If $H^1\alpha$ is a broken product datum, it can connected by a path to an unbroken one by gluing. Thus $\mathcal{P}(H_0, H_1, H_2)$ is connected. Dissipativity is immediate as in part (b). □

## 9.2 Construction of the pair-of-pants product

**Lemma 9.4** *Fix a superdissipative $H$. Suppose that $(H_0, H_1, H_2) \in \mathcal{F}(H)^3$ consists of nondegenerate Hamiltonians. Then for generic choice of element in $\mathcal{P}(H_0, H_1, H_2)$ and for a generic path in $\mathcal{P}(H_0, H_1, H_2)$, the associated 0– and 1–dimensional moduli spaces of Floer solutions are compact, smooth and of the expected dimension.*

**Proof** By construction, an admissible product datum satisfies the hypotheses of Lemma 5.3. Therefore, given generators $\tilde{\gamma}_i \in CF^*(H_i)$, the pairs of pants of a fixed dissipative Floer datum with $i^{\text{th}}$ end asymptotic to $\tilde{\gamma}_i$ have energy estimated, according to Lemma 5.3, by $E^{\text{top}}(U)$, equal to the action difference $\mathcal{A}_{H_3}(\tilde{\gamma}_3) - \mathcal{A}_{H_1}(\tilde{\gamma}_1) - \mathcal{A}_{H_2}(\tilde{\gamma}_2)$. The first part of Lemma 9.3 and Theorem 6.3 thus imply that they are all contained in an a priori compact set $K$ depending on the differences $b_i - a_i$. The story is now the same as the closed case, which is dealt with in the aspherical case in [54]. Sphere bubbling is treated in the exact same way as for the differential, continuation maps and chain homotopy operators, which was done in detail in Section 7.2. The upshot is that in the monotone or Calabi–Yau case, whether one is working with a single Floer datum or with a family of parametrized ones, spheres occur in codimension 4. All the arguments involve at most 1–dimensional families of Floer solutions. Generically, there is no sphere bubbling for such families. □

**Theorem 9.5** *Fix a superdissipative $H$. For any Floer datum $(H, J)$, denote by $CF^{*,0}(H, J)$ the subcomplex generated by contractible periodic orbits. A generic choice of admissible product datum for a generic product triple $(H_0, H_1, H_2) \in \mathcal{F}(H)^3$ determines a bilinear map*

$$*\colon CF^{*,0}(H_0, J_0) \otimes CF^{*,0}(H_1, J_1) \to CF^{*,0}(H_2, J_2)$$

*satisfying*

$$\mathrm{val}(x_1 * x_2) \le \mathrm{val}(x_1) + \mathrm{val}(x_2). \tag{68}$$

*Moreover, we have $d(x_1 * x_2) = dx_1 * x_2 + (-1)^{\deg(x_1)} x_1 * dx_2$. The induced map on homology satisfies the following properties*:

(a) *It is independent of the choice of admissible product datum.*

(b) *If, in addition, $H_2 \ge H_1 * H_0$, it is supercommutative.*

(c) *It commutes with all continuation maps in $\mathcal{F}(H)$.*

> *For the remainder of this section, unless indicated otherwise, all the Floer cohomology groups are those arising from contractible orbits. We do not indicate this further in the notation.*

**Remark 9.6** The reason we restrict our discussion of the pair-of-pants product to contractible periodic orbits is in the formulation of the Floer complex we chose in Section 7.3. In that formulation the Floer complex is generated over $R$ by appropriate equivalence classes $[\gamma, A]$, where $\gamma$ is a periodic orbit and $A$ a path from a base loop in the component of $\gamma$. Concatenating two such data $(\gamma_0, A_0), (\gamma_1, A_1)$ with a pair of pants having output on some periodic orbit $\gamma_2$ does not give rise to an appropriate path $A_2$ unless all the involved periodic orbits $\gamma_i$ are contractible. To obtain a well-defined product involving noncontractible orbits, additional choices need to be made. This should not be hard, but we do not pursue the details. An alternative approach which avoids this issue altogether is indicated in Remark 7.13. Note also that if the symplectic form is exact or even merely aspherical and atoroidal, there is no issue. Moreover, if one of the inputs is contractible, the pair-of-pants product is well defined without additional choices. Thus, as a result of the present subsection, we do get the module structure of the full symplectic cohomology over the contractible part without any additional work.

**Proof of Theorem 9.5** Fix a regular pair-of-pants datum $P \in \mathcal{P}(H_0, H_1, H_2)$. Let $\tilde{\gamma}_i \in CF^*(H_i)$ for $i = 0, 1, 2$ be such that

$$i_{\mathrm{RS}}(\tilde{\gamma}_2) = i_{\mathrm{RS}}(\tilde{\gamma}_0) + i_{\mathrm{RS}}(\tilde{\gamma}_1).$$

The product $*$ is defined by counting the Floer solutions associated with $p$. The Leibnitz rule is obtained by analyzing the boundary of the 1–dimensional moduli spaces. For details see [4, Section 2.3.5]. Note that while [4] concerns the cotangent bundle, once we fix a regular product datum, the analysis of the moduli spaces is exactly the same.

Behavior of the valuation under $*$ follows by Lemma 5.3. Namely, by monotonicity of the product data, any solution must have nonnegative energy.

Given two choices of admissible product data, Lemma 9.3 allows us to construct a dissipative homotopy. We can perturb while maintaining dissipativity to get a sufficiently generic homotopy inducing a chain homotopy between the appropriate complexes. Commutation with continuation maps follows in the same way from Lemma 9.3 and a standard gluing argument.

The claim about supercommutativity follows by pulling back the product datum $P$ by a biholomorphism of $S^2$ which fixes $\infty$ and commutes 0 and 1. For details see [4, Lemma 2.3.24]. $\square$

**Lemma 9.7** *The pair-of-pants product induces a map*

$$*\colon HF^*_{[a_1,b_1)}(H_1) \otimes HF^*_{[a_2,b_2)}(H_2) \to HF^*_{[\max\{a_1+b_2,a_2+b_1\},b_1+b_2)}(H_3)$$

*for all $\mathcal{F}(H)$ admissible triples. Moreover, the product $*$ fits into a commutative diagram*

$$\begin{tikzcd} HF^*_{[a_1,b_1)}(H_1) \otimes HF^*_{[a_2,b_2)}(H_2) \arrow[r] \arrow[d] & HF^*_{[\max\{a_1+b_2,a_2+b_1\},b_1+b_2)}(H_3) \arrow[d] \\ HF^*_{[a_1',b_1')}(H_1) \mathbin{\hat{\otimes}} HF^*_{[a_2',b_2')}(H_2) \arrow[r] & HF^*_{[\max\{a_1'+b_2',a_2'+b_1'\},b_1'+b_2')}(H_3) \end{tikzcd} \tag{69}$$

*whenever $a_i' > a_i$ and $b_i' > b_i$.*

**Proof** Recall the identity for any $R$–modules $A, B, C, D$,

$$A/B \otimes C/D = A \otimes C/(B \otimes C + A \otimes D).$$

The first part of the claim follows by definition of $CF^*_{[a,b)}$ and the estimate (68). The second part is clear if one works with representatives. $\square$

**Lemma 9.8** *Suppose $H_1, H_2, H_3$ are a triple of lower semicontinuous exhaustion functions satisfying*

$$H_3 \geq H_1 * H_2 \tag{70}$$

*and $H_i \geq H$. Then there is a monotone sequence of admissible product triples $H_{ik} \in \mathcal{F}(H), i = 1, 2, 3$ such that $H_{ik}$ converges pointwise as $k \to \infty$ to $H_i$.*

**Proof** Pick constants $a_1, a_2, a_3 < 1$ such that $a_1 + a_2 < a_3$. According to Lemma 8.11, we can find sequences $H_{ik}, i = 1, 2, 3$, which are monotone in $k$ converging pointwise to $H_i$ and coinciding with $a_i H$ outside of a compact set. From (70) it follows that for each $k$ there exists an index $i_k$ such that $H_{3,i_k} > H_{1,k} * H_{2,k}$. The sequence $H_{1,k}, H_{2,k}, H_{3,i_k}$ is as required. $\square$

**Lemma 9.9** *The pair-of-pants product for Hamiltonians in $\mathcal{F}(H)$ induces a map*

$$*_H : HF^*_{[a_1,b_1)}(H_1) \otimes HF^*_{[a_2,b_2)}(H_2) \to HF^*_{[\max\{a_1+b_2, a_2+b_1\}, b_1+b_2)}(H_3)$$

*for all triples $(H_1, H_2, H_3) \in \mathcal{H}^3_{\mathrm{sc}}$ that satisfy $H_i \geq H$. Moreover, the product $*_H$ fits into a commutative diagram as in* (69).

**Proof** Pick a monotone sequence $(H_{0k}, H_{1k}, H_{2k})$ of $H$–admissible triples converging to $(H_0, H_1, H_2)$. As in (65),

$$HF^*_{[a,b)}(H_i) := \varinjlim_k HF^*_{[a,b)}(H_{ik}).$$

Since tensor product commutes with direct limits, we get an induced product as in the statement of the lemma. Moreover, since the pair-of-pants product commutes with all continuation maps in $\mathcal{F}(H)$, the product $*_H$ is independent of the choice of approximating sequence. $\square$

**Lemma 9.10** *Fix a superdissipative $H$. The pair-of-pants product on $\mathcal{F}(H)$ induces a canonical product*

$$*_H : \overline{HF}^*(H_0) \mathbin{\widehat{\otimes}} \overline{HF}^*(H_1) \to \overline{HF}^*(H_2)$$

*for all product triples $(H_0, H_1, H_2) \in \mathcal{H}^3_{\mathrm{sc}}$. The operation $*_H$ commutes with all continuation maps and is supercommutative.*

**Proof** For $i=0,1$, let $\gamma_i\in\overline{HF}^*(H_i)$. By Lemma 8.17, $\overline{HF}^*(H_i)$ is the reduced cohomology of an appropriate chain complex $\widehat{CF}^*(H_i)$ well defined up to filtered quasi-isomorphism. Pick such chain complexes for $H_0$ and $H_1$. Let $\widetilde{\gamma}_0,\widetilde{\gamma}_1$ be representatives of $\gamma_0,\gamma_1$ respectively. We construct an element $\gamma_2=[\widetilde{\gamma}_1]*_H[\widetilde{\gamma}_2]\in\overline{HF}^*(H_2)$ as follows. By (66), to give an element in $\overline{HF}^*(H_2)$ it suffices to fix some $c$ and give for each $a<c<b$ an element $\gamma_2^{ab}\in HF^*_{[a,b)}(H_2)$ so that the $\gamma_2^{ab}$ agree under the natural maps

$$HF^*_{[a,b)}(H_2)\to HF^*_{[a',b')}(H_2), \tag{71}$$

defined whenever $a<a'$ and $b<b'$. For some $\epsilon>0$ write $b_i=\mathrm{val}(\widetilde{\gamma}_i)+\epsilon$ for $i=0,1$. Fix some $b>b_0+b_1$ and for any $a<b$ let $a_0=a-b_1$ and $a_1=a-b_0$. Then by applying the operation $*_H$ of Lemma 9.9 to the classes of $\widetilde{\gamma}_i$ in $HF^*_{[a_i,b_i)}(H_i)$ we obtain an element $\gamma_2^{ab}\in HF^*_{[a,b)}$. Moreover, $\gamma_2^{ab}$ agrees with $\gamma_2^{a'b'}$ under the natural maps (71). We thus obtain an element $\gamma_2\in\overline{HF}^*(H_2)$ which is well defined after fixing representatives $\widetilde{\gamma}_i$. We need to verify that $\gamma_2$ is independent of the choice of representatives. For this it suffices to show that if $\widetilde{\gamma}_i$ is in the closure of the image of the boundary for either $i=0$ or $i=1$, then $\gamma_2=0$. This amounts to showing that for each $a$ there exists a $b$ such that $\gamma_2^{ab}=0$. For definiteness assume $[\widetilde{\gamma}_0]=0\in\overline{HF}^*(H_0)$. By (66) we need to show that for each $a$ there is a $b$ such that

$$\gamma_2^{ab}=0\in HF_{[a,b)}(H_2).$$

For this it suffices to show that there is a $b$ such that we can find numbers $a_i,b_i$ for $i=0,1$ such that

$$\max\{a_0+b_1,a_1+b_0\}\le a<b_0+b_1\le b,\quad b_i>\mathrm{val}(\widetilde{\gamma}_i),$$

and such that $[\widetilde{\gamma}_0]=0\in HF^*_{[a_0,b_0)}(H_0)$. We choose $b_1=\mathrm{val}(\widetilde{\gamma}_1)+\epsilon$ and $a_0=a-b_1$. Since $\widetilde{\gamma}_0$ vanishes in reduced cohomology, there exists a $b_0=b_0(a_0)>\mathrm{val}(\widetilde{\gamma}_1)$ such that it vanishes in $HF^*_{[a_0,b_0)}(H_0)$. Pick $a_1=a-b_0$. Then all the requirements are satisfied. It is clear that changing the underlying complexes for $\overline{HF}^*(H_i)$ up to filtered quasi-isomorphism does not affect the definition of $*$. □

### 9.2.1 Independence of the choice of $H$ at infinity

**Lemma 9.11** *Let $F_0\le F_1$ be superdissipative. For $i=0,1$ let $(H_0^i,H_1^i,H_2^i)\in\mathcal{F}(F_i)$, and suppose $H_j^0\le H_j^1$ for $j=0,1,2$. Then the operation $*$ commutes with the continuation maps $H_j^0\to H_j^1$. In particular, the definition of the pair of products is independent at the homology level of the choice of $H$ in the approximating scheme.*

**Proof** Fix dissipative homotopies $H_j^0 \to H_j^1$. Gluing either the first two homotopies to the input of the pair of pants in $\mathcal{F}(F_0)$ or the last one to the pair of pants in $\mathcal{F}(F_1)$ gives rise to two pairs of pants from $H_0^0, H_1^0$ to $H_2^1$ with 1–forms of the form $\mathfrak{H}_i = G_i\alpha_\Sigma$ such that $d\mathfrak{H}_i \wedge dh_\Sigma \geq 0$. We need a family $(G_s, J_s)$ such that $(\mathfrak{H}_s := G_s\alpha_\Sigma, J_s)$ form a dissipative interpolating family still satisfying

$$d\mathfrak{H}_s \wedge dh_\Sigma \geq 0.$$

The proof of existence of such a $(G_s, H_s)$ is exactly as in Lemma 7.6. $\square$

#### 9.2.2 Associativity

**Lemma 9.12** *Let $H^i$ for $i = 1, 2, 3$ and $H^{1,2}, H^{2,3}$ be elements of $\mathcal{H}_{\mathrm{sc}}$. Suppose $(H^1, H^2, H^{1,2})$ and $(H^2, H^3, H^{2,3})$ are product triples. Let $H^4$ be such that*

$$H^4 \geq 2 \max_{t \in \mathbb{R}/\mathbb{Z}} \{H_t^{1,2}, H_t^{2,3}\}.$$

*Then the maps*

$$\overline{HF}^*(H^1) \otimes \overline{HF}^*(H^2) \otimes \overline{HF}^*(H^3) \to \overline{HF}^*(H^4)$$

*coming from the two compositions in Figure 4 coincide.*

**Proof** It suffices to prove the claim under the assumption that $H^i, H^{i,j} \in \mathcal{F}(H)$ for some superdissipative $H$, since we can replace all the involved Hamiltonians by approximating sequences in $\mathcal{F}(H)$. Moreover, we may assume all inequalities are strict. The assumption implies there is a time-independent $H$ such that $H^4 > H > 2\max\{H^{1,2}, H^{2,3}\}$. Thus if we prove associativity for the case where all the functions are positive multiples of a single function, the general claim will follow by naturality of the pair-of-pants product with respect to continuation maps. For this case the proof is standard in the literature (see eg [5; 4]) but we spell out the details.

Consider $H^1 = H^2 = H^3 = H$, $H^{1,2} = H^{2,3} = 2H$ and $H^4 = 4H$. Now let $f^{1,2}, f^{2,1} : \Sigma \to \mathbb{R}$ be functions such that $d(f^{i,j}\alpha_\Sigma) \geq 0$, $df^{i,j}$ is compactly supported, and $f^{i,j}$ is equal to $i$ at the input $z_0$, to $j$ at the input $z_1$, and to 4 at the output. We consider the 1–forms $\mathfrak{H}^{1,1,2} := H\alpha_\Sigma$ and $\mathfrak{H}^{i,j,4} := f^{i,j}H\alpha$. The two possible compositions correspond to the gluing of $\mathfrak{H}^{1,1,2}$ to either $\mathfrak{H}^{2,1,4}$ at the first input, or to $\mathfrak{H}^{1,2,4}$ at the second input. We must show that there exists a homotopy between the two compositions. We write these glued 1–forms as $H\alpha_0$ and $H\alpha_1$; here $\alpha_i$ are 1–forms on $S^2 \setminus \{0, 1, z', \infty\}$, where $z'$ is a point $z_0$ near 0 for the first composition and $z_1$ near

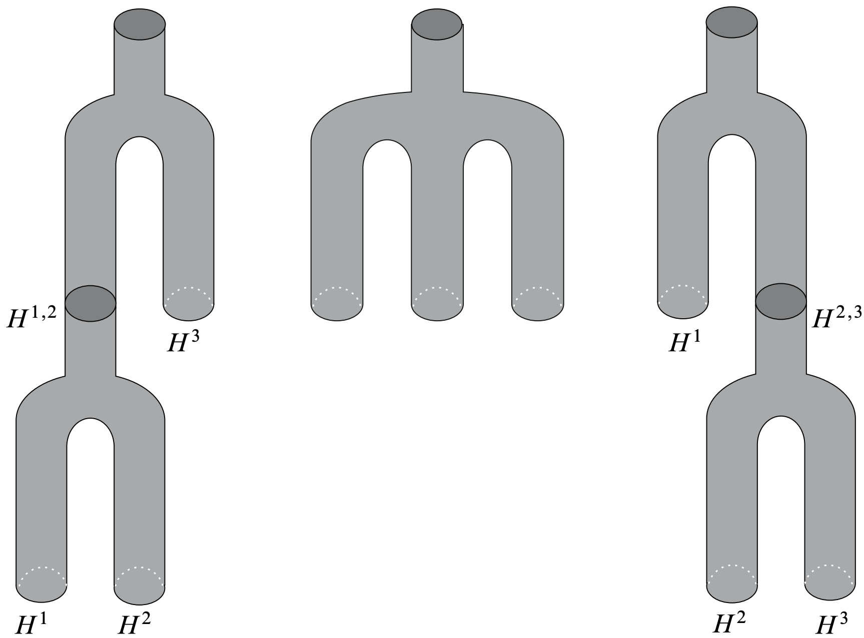


Figure 4: Homotopy for associativity.

1 for the second composition. Note that $d\alpha_i \geq 0$. We pick a smooth path $z_{\sigma\in[0,1]}$ in the Riemann sphere. We show that we can lift this to a path $\alpha_\sigma$ of 1–forms satisfying $d\alpha_\sigma \geq 0$ and connecting $\alpha_0$ to $\alpha_1$. Smooth four-punctured spheres are diffeomorphic by a diffeomorphism which preserves the cylindrical ends. Thus the claim reduces to finding such a homotopy on a fixed surface. But the condition $d\alpha \geq 0$ is convex. Thus we can find a path $\alpha_\sigma$ as required. The family $\mathfrak{H}_\sigma := H\alpha_\sigma$ gives the required homotopy. □

## 9.3 The PSS map

**Theorem 9.13** *Let $M$ be geometrically bounded. Then the small quantum product on $H^*(M;\mathbb{K})$ is well defined.*

**Proof** We take as our model of $H^*(M)$ the homology of the Morse complex $CM^*$ arising from considering the positive gradient flow of some proper exhaustion function $f: M \to \mathbb{R}$ with nondegenerate critical points, together with a geometrically bounded $J$ such that the pair $(f, g_J)$ is Morse–Smale. For this to compute cohomology (and, indeed, for the Morse differential to be well defined) we take $CM^* := \mathbb{K}^{\mathrm{crit}(f)}$. Indeed,

$CM^*$ thus defined is the dual of $CM_*$, which consists of finite formal sums of critical points with the differential defined by counting negative gradient lines. Since $g$ is proper and bounded below, the subcomplexes $CM^*_a \subset CM^*$ generated by critical points in the sublevel set $M_a := f^{-1}(-\infty, a)$ compute the singular homology of $M_a$ by standard Morse theory [39]. By taking colimits it follows that $CM_*$ computes singular homology of $M$, and therefore, $CM^*$ computes singular cohomology. The small quantum product is defined by counting the $J$–holomorphic spheres with three marked points intersecting the unstable manifolds of some input critical points $p, q$ and the stable manifold of an output critical point $r$. Since $f$ is proper and bounded below, the stable manifold of any critical point is precompact. Thus by Theorem 4.11 all the spheres are contained in a priori compact sets. The fact that the operation thus defined is indeed an associative product is now standard; see eg [37, Section 12.2] and [46]. □

**Remark 9.14** An alternative way to think of the construction of the small quantum product is to observe that cohomology of a noncompact manifold is Poincaré dual to Borel–Moore homology. That is, a homology where one allows locally finite sums of singular chains. Given a triple $\gamma_1, \gamma_2, \gamma_3$ of Borel–Moore homology classes, the coefficient of $\gamma_3$ in the quantum product $\gamma_1 *_{QH} \gamma_2$ is the three-pointed Gromov–Witten associated with the triple $\gamma_1, \gamma_2, \gamma_3^*$, where $\gamma_3^* \in H_*(M)$ is the Poincaré dual of $\gamma_3$. Now $\gamma_3^*$ is a cycle in ordinary homology and thus a finite chain. Therefore, by geometric boundedness, the number of $J$–holomorphic spheres intersecting $\gamma_3^*$ and representing a given homology class is finite.

**Theorem 9.15** *For any $H \in \mathcal{H}_{\mathrm{sc}}$ there is a natural map*

$$f_H^{\mathrm{PSS}} \colon H^*(M;\mathbb{K}) \to \overline{HF}^*(H).$$

*Denote by $*$ the product in Floer cohomology and by $*_{QH}$ the small quantum product. Then $f^{\mathrm{PSS}}$ satisfies for any product triple $H_0, H_1, H_2$ and for any pair of classes $a, b \in H^*(M;\mathbb{K})$,*

$$f_{H_0}^{\mathrm{PSS}}(a) * f_{H_1}^{\mathrm{PSS}}(b) = f_{H_2}^{\mathrm{PSS}}(a *_{QH} b).$$

*In addition, for any $x \in \overline{HF}(H_1)$,*

$$f_{H_0}^{\mathrm{PSS}}(1) * x = f_{H_1,H_2}(x),$$

*where*

$$f_{H_1,H_2} \colon \overline{HF}^*(H_1) \to \overline{HF}^*(H_2)$$

*is the natural continuation map.*

**Proof** In the compact case for smooth nondegenerate Hamiltonians this is [46]. In the noncompact case, we achieve $C^0$ estimates for smooth dissipative nondegenerate Hamiltonians by considering appropriate dissipative data on the plane. Namely, we pick a geometrically bounded $J$ and monotone homotopy $H_z$ going from 0 for $z$ in a neighborhood of the origin and to $H$ for $z$ near $\infty$, and such that the associated Gromov metric $g_{J_H}$ is i–bounded. This is done as in Lemma 7.6. Alternatively, we can restrict the direct definition to the superdissipative case, and just take $H_z = f(|z|)H$ for some monotone increasing function. The definition for arbitrary $H$ is by an approximation scheme as in the definition of the pair-of-pants product. Write $dt := d\arg z$. Then the Floer datum $(J, H_z\,dt)$ on the complex plane is dissipative. Moreover, it is monotone, so by Lemma 5.3 and Theorem 6.3 the solutions emanating from any critical point of $f$ to any critical point of $\mathcal{A}_H$ are confined to an a priori compact set. This reduces the claims to the compact case. □

## 9.4 Proof of Theorems 1.3 and 2.1

Let $\mathcal{H} \subset \mathcal{H}_{\mathrm{sc}}$ be a subset consisting of time-independent Hamiltonians such that for any $H_1, H_2 \in \mathcal{H}$ we have that $2\max\{H_1, H_2\} \in \mathcal{H}$. We call $\mathcal{H}$ a *monoidal indexing set*. For each monoidal indexing set $\mathcal{H}$, we define a group

$$SH^*(M;\mathcal{H}) := \varinjlim_{H\in\mathcal{H}} \overline{HF}^*(M).$$

We denote by $\mathcal{H}_{\mathrm{sm}}$ the monoidal indexing set consisting of all smooth functions which are proper and bounded from below, and define

$$SH^*_{\mathrm{univ}}(M) := SH^*(M;\mathcal{H}_{\mathrm{sm}}).$$

We now prove Theorem 1.3 from the introduction, which states that $SH^*(M;\mathcal{H})$ is a unital algebra over $QH^*(M;\mathbb{K})$.

**Proof of Theorem 1.3** (a) Given $\gamma_0, \gamma_1 \in SH^*(M;\mathcal{H})$, we can find $H_0, H_1 \in \mathcal{H}$ such that $\gamma_i$ lifts to an element still denoted by $\gamma_i \in \overline{HF}^*(H_i)$ for $i = 0, 1$. Since $\mathcal{H}$ is a monoidal indexing set we can find an $H_2 \in \mathcal{H}$ such that $(H_0, H_1, H_2)$ form a product triple. Pick a superdissipative Hamiltonian $H \le H_i$, $i = 0, 1, 2$, and let $\gamma_2 := \gamma_0 *_H \gamma_1 \in \overline{HF}^*(H_2)$, using the induced product $*_H$ from Lemma 9.10. By Lemma 9.11, $\gamma_2$ is independent of the choice of $H$. We define $\gamma_0 * \gamma_1 \in SH^*(M;\mathcal{H})$ to be the image of $\gamma_2$ under the natural map $\overline{HF}^*(H_2) \to SH^*(M;\mathcal{H})$. Since $*_H$ commutes with all continuation maps, $\gamma_1 * \gamma_2$ is independent of the choice of product triple $(H_0, H_1, H_2) \in \mathcal{H}^3$. Associativity and supercommutativity hold up to continuation maps in $\mathcal{H}$ by Lemmas 9.12 and 9.10. Indeed, if $H_1, H_2, H_3 \in \mathcal{H}$, then so is $4\max\{H_1, H_2, H_3\}$.

(b) This is an immediate consequence of Theorem 9.15.

(c) This is an immediate consequence of the naturality of the pair-of-pants product with respect to the continuation maps. □

We next turn to the proof of Theorem 2.1 concerning local symplectic cohomology, but first we recall some definitions. Let $K \subset M$ be a compact set. Let

$$H_K(x) := \begin{cases} 0 & \text{if } x \in K, \\ \infty & \text{if } x \in M \setminus K. \end{cases}$$

The *local symplectic cohomology* at $K$ is defined by

$$SH^*(M|K;\mathbb{K}) := \overline{\overline{HF}}^*(H_K;\mathbb{K}).$$

**Proof of Theorem 2.1** (a) We have

$$K_1 \subset K_2 \iff H_{K_2} \le H_{K_1}.$$

Thus there is a continuation map

$$SH^*(M|K_2) = \overline{\overline{HF}}^*(H_{K_2}) \to SH^*(M|K_1) = \overline{\overline{HF}}^*(H_{K_1}).$$

(b) This is the symplectic invariance of the construction of $\overline{\overline{HF}}^*$.

(c) We have that $\mathcal{H}_K := \{H_K\}$ forms a monoidal indexing set. So $SH^*(\mathcal{H}_K) = \overline{\overline{HF}}^*(H_K)$, and the claim follows from Theorem 1.3.

(d) This is an immediate consequence of Theorem 1.3 and the functoriality of the continuation maps.

(e) We have $H - \sup_K H \le H_K$. On the other hand the map corresponding to $H \to H + c$ is a conformal isomorphism decreasing valuation by $c$. □

## 9.5 Symplectic cohomology as a topological vector space

There is more than one natural way to put a topology on $SH^*(\mathcal{H})$ depending on the purpose one has in mind. In the rest of the paper we shall consider the *final topology* on $SH^*(M;\mathcal{H})$. That is, the strongest topology for which all the continuation maps $\overline{\overline{HF}}^*(H) \to SH^*(M;\mathcal{H})$ for $H \in \mathcal{H}$ are continuous. Note that this topology is not necessarily Hausdorff. We also do not address the continuity of the pair-of-pants product. For applications later in this paper we will only need the following lemma.

**Lemma 9.16** *Let $\mathcal{H}$ be a monoidal indexing set consisting of continuous Hamiltonians. Let $K \subset M$ be a compact set. Then the natural map*

$$SH^*(\mathcal{H}) \to SH^*(M|K)$$

*is continuous.*

**Proof** By Theorem 2.1(e), for any continuous Hamiltonian $H$, the continuation map $HF^*(H) \to SH^*(M|K) = HF^*(H_K)$ is continuous. Continuity of the induced map follows by definition of the final topology. □

**Remark 9.17** Since the spaces $SH^*(M|K)$ are all Banach spaces, Lemma 9.16 will still hold if one considers topologies on $SH^*(\mathcal{H})$ which are weaker than the final topology. We do not pursue this further, however.

We illustrate the various notions of symplectic cohomology by considering the case $M = \mathbb{R} \times \mathbb{R}/\mathbb{Z}$. We will compare symplectic cohomologies for three different monoidal indexing systems. *Since $M$ is a Liouville domain we no longer restrict the discussion of the pair-of-pants product to contractible orbits.*

**Example 9.18** Consider the monoidal indexing set $\mathcal{L}$ consisting of Hamiltonians which, outside of a compact set, are of the form

$$H(s,t) = a|s| + b.$$

According to a theorem by Viterbo [64] equating symplectic cohomology of a cotangent with loopspace homology of the underlying manifold, for any coefficient field $R$ we have that

$$SH^*(M;\mathcal{L}) = R[x, x^{-1}, \partial_x]/\partial_x^2. \tag{72}$$

Thus, *$SH^*(M;\mathcal{L})$ is the exterior algebra of polyvector fields on $R \setminus \{0\}$.* By the Künneth formula [42], or by Viterbo's theorem again, the same holds for $M = T^*\mathbb{T}^n$.

**Example 9.19** Now let $K \subset M = [-a,a] \times \mathbb{R}/\mathbb{Z}$. To keep track of actions choose the primitive $s\,dt$ of the standard symplectic form. With this choice, one shows that

$SH^*(M|K)$ is obtained from $SH^*(M;\mathcal{L})$ by completing with respect to the valuation $\operatorname{val}(x^i) := |i|a$ and $\operatorname{val}(\partial_x) := 0$. When the underlying ring $R$ is trivially valued, this completion is of no effect. However, when working over the universal Novikov ring we obtain for example that $SH^0(M|K)$ consists of all infinite Laurent series

$$\sum_{i=-\infty}^{\infty} b_i x^i \quad \text{with } b_i \in \Lambda_{0,\mathrm{nov}}$$

satisfying

$$\lim \operatorname{val}(b_i) + |i|a \to -\infty.$$

This is the same as the set of analytic functions on the rigid analytic torus $\Lambda^*$ which converge on the subtorus $\{z \in \Lambda^* \mid \operatorname{val}(z) \in [-a, a]\}$. A reference for rigid analytic geometry is [9]. Closer to home, [58] and [3] provide a closely related point of view from the vantage point of Lagrangian Floer homology of the $\mathbb{R}/\mathbb{Z}$ fibers.

**Example 9.20** Finally we study $SH^*_{\mathrm{univ}}(\mathbb{R}\times S^1)$. We claim that *over the universal Novikov ring, $SH^0_{\mathrm{univ}}(\mathbb{R}\times S^1)$ consists of formal Laurent series $\sum b_i x^i$ that are rapidly decreasing in the sense that there exists a superlinear convex function $g$ such that $\operatorname{val}(b_n) = -g(n)$.*

To see this, observe that $SH^*_{\mathrm{univ}}(M)$ is computable by direct limit of $HF^*(H)$ over functions $H$ which outside of a compact set are of the form $H(s,t) = h(|s|)$ for $h\colon \mathbb{R}_+ \to \mathbb{R}_+$ a convex function such that $h'(t)$ is unbounded as $t \to \infty$.[12] For any such $H$ there is a natural map

$$SH^*(M;\mathcal{L}) \to HF^*(H),$$

by the universal property of the direct limit and the fact that for any $H' \in \mathcal{L}$ we have $H' \le H$ outside of a compact set. Each monomial $x^i$ maps to a class associated with a unique periodic orbit $\gamma^i$ of $H$. It is not hard to show that $HF^*(H)$ is in fact the completion of the algebra $SH^*(M;\mathcal{L}) = R[x, x^{-1}, \partial_x]$ with respect to the norm $\operatorname{val}(x_i) = \mathcal{A}_H(\gamma_i)$. This is computed as follows. Let $s_i \in \mathbb{R}$ be the unique real number such that $\partial_s H(s,t) = i$. Then

$$\operatorname{val}(x^i) = s_i h'(s_i) - h(|s_i|) = i s_i - h(|s_i|).$$

Writing $f = h'$ we have $s_i = f^{-1}(i)$ and the right-hand side of the last equation is exactly $g(i) := \int_0^i f^{-1}(t)\,dt$. Since $f^{-1}$ is monotone and unbounded this means $g_i$ is convex and superlinear. From this it is not hard to deduce the claim.

[12] Such a function is not necessarily dissipative. However, we may still talk about its Floer cohomology by approximating by linear Hamiltonians.

# 10 Computations and applications

## 10.1 Liouville domains

Let $(\Sigma,\alpha)$ be a contact manifold with contact form $\alpha$. Let $(U,\omega=d\lambda)$ be a compact exact symplectic manifold with $\Sigma$ as boundary such that $\alpha=\lambda|_{\partial U=\Sigma}$ and such that the Liouville field $Z$, which is defined by $\iota_Z\omega=\lambda$, points outward at the boundary. Let $\widehat{U}$ be the completion of $U$ by attaching the cone $\Sigma\times\mathbb{R}_{\geq 0}$ with the symplectic form $\omega_\alpha=e^r(d\alpha+dr\wedge\alpha)$. The vector field $Z$ extends to $\widehat{U}$ and is given by $Z=\partial/\partial r$.

Denoting by $\phi_t$ the time $t$ flow of $Z$, the skeleton of $U$ relative to $\lambda$ is defined by

$$\mathrm{Skel}(U,Z)=\bigcap_{t>0}\phi_t(U).$$

The map $\Sigma\times\mathbb{R}\to\widehat{U}$ defined by $(x,r)\mapsto\phi_r(x)$ is a symplectic embedding of the symplectization of $\Sigma$, whose image is $\widehat{U}\setminus\mathrm{Skel}(U,\lambda)$. A reference for these basic definitions and claims is [14]. In particular, the function $\phi_r(x)\mapsto e^r$ is defined and smooth on $\widehat{U}\setminus\mathrm{Skel}(U,Z)$. Moreover, it extends to a continuous function of $\widehat{U}$, still denoted by $e^r$, by defining $e^r(p)=0$ for $p\in\mathrm{Skel}(U,Z)$. Denote by $\mathcal{L}$ the set of Hamiltonians that outside of a compact set containing the skeleton are of the form $ae^r+b$ for $a>0$ and $b\in\mathbb{R}$. We refer to these Hamiltonians as being *linear at infinity*. Similarly, Hamiltonians which outside of a compact set are of the form $h(e^r)$ for $h$ convex are referred to as *convex at infinity*. Let $J$ be of contact type; that is, $J$ is an $\omega$–compatible translation-invariant almost complex structure $J$ satisfying $JR=\partial r$ for $R$ the Reeb flow. As in [63], define $SH^*_{\mathrm{Viterbo}}(U)=\varinjlim_{H\in\mathcal{L}}HF^*(H,J)$.

**Lemma 10.1** $$SH^*_{\mathrm{Viterbo}}(U)=SH^*(\widehat{U};\mathcal{L}).$$

**Proof** By Example 5.25, when paired with a contact type $J$, the elements of $\mathcal{L}$ are i–bounded. Any $H\in\mathcal{L}$ with slope at infinity not in the period spectrum of $\Sigma$ is dissipative by Example 6.14. It follows that $\overline{HF}^*(H)=HF^*(H)$. So the directed systems computing each side coincide. □

In particular, we have a natural map

$$f\colon SH^*_{\mathrm{Viterbo}}(U;R)=SH^*(\widehat{U};\mathcal{L},R)\to SH^*_{\mathrm{univ}}(\widehat{U};R).$$

**Theorem 10.2** *The map $f$ is an isomorphism.*

**Remark 10.3** The proof below of Theorem 10.2 relies crucially on the fact that for Hamiltonians which are convex at infinity, the action spectrum is bounded from below, rendering the topology of $HF^*(H)$ discrete. This fails when working over a nontrivially valued field. To see what sets trivially valued fields apart, consider the following. Given two Hamiltonians $H_1 \leq H_2$ such that $H_i = h_i(e^r)$, the continuation map

$$f_{12}\colon HF^*(H_1; R) \to HF^*(H_2; R)$$

can be shown two be an isomorphism of vector spaces, and thus, since the topology is discrete, of topological vector spaces. However the inverse of $f_{12}$ will generally not be bounded. Thus when working over $\Lambda_R$, the map $f_{12}$ will no longer be a homeomorphism.

**Proof of Theorem 10.2** We consider the set $\mathcal{C}$ of smooth Hamiltonians $H$ for which there is a compact $K = \{e^r < \epsilon\}$ for some $\epsilon > 0$ such that $H$ is $C^2$–small and negative on $K$, and of the form $H = h(e^r)$ outside of $K$. The action of any 1–periodic orbit of such a Hamiltonian is positive. The set of Hamiltonians $\mathcal{C}$ is cofinal in the set of all smooth Hamiltonians with respect to the order relation $\preceq$, defined by

$$H_1 \preceq H_2 \iff H_2(x) - H_1(x) \geq C > -\infty \quad \text{for all } x \in M. \tag{73}$$

Pick a sequence $F_i \in \mathcal{C}$ given outside of a compact set by $F_i = h_i(e^r)$ so that the sequence $F_i$ converges to $H$ on compact subsets of $M$ and such that near infinity $h_i$ is linear of slope not in the period spectrum of $\alpha$. The action of a periodic orbit is given by the right-hand side of (43), which in this case specializes, for a nontrivial periodic orbit $\gamma$ of $F_i$ occurring at some level set $e^r = t$, to

$$\mathcal{A}_{F_i}(\gamma) = t h_i'(t) - h_i(t),$$

which is positive for $h_i$ convex. Positivity also holds for the trivial periodic orbits, since they occur inside $U$ where $F_i < 0$. We thus have that $HF^*_{[k,\infty)}(F_i) = HF^*(F_i)$ for all $k \leq 0$. A similar statement holds for $H$. From this we deduce, first, that $\overline{HF}^*(H) = HF^*(H)$, and, second, that $HF^*(H) = \varinjlim_i HF^*(F_i)$.

The set $F_i$ is cofinal in $\mathcal{L}$ with respect the order relation $\preceq$. Therefore, we obtain an isomorphism of $R$–modules

$$HF^*(H) = \varinjlim_i HF^*(F_i) \to SH^*(\mathcal{L}_{\mathrm{reg}}; R).$$

Moreover, given two convex functions $H_1 \leq H_2$, the continuation maps from $H_1$ to $H_2$ will commute with the above isomorphisms since they are all defined via continuation maps between functions which are linear near infinity. The claim follows. $\square$

We similarly prove:

**Theorem 10.4**
$$\begin{aligned} SH^*(\widehat{U}\,|\,\mathrm{Skel}(U,Z);R) &= SH^*(\widehat{U}|U;R) \\ &= SH^*_{\mathrm{Viterbo}}(U;R) \\ &= SH^*_{\mathrm{univ}}(\widehat{U};R). \end{aligned}$$

**Proof** Consider a monotone sequence $H_i$ belonging to the set of convex Hamiltonians $\mathcal{C}$ defined in the proof of Theorem 10.2, so that

$$\lim_{x\to\infty} H_i(x) = \begin{cases} 0 & \text{if } x \in U, \\ \infty & \text{if } x \in \widehat{U} \setminus U. \end{cases}$$

Then, by positivity of the action spectrum,

$$SH^*(\widehat{U}|U;R) = \varprojlim_k \varinjlim_i HF^*_{[k,\infty)}(H_i) = \varinjlim_i HF^*(H_i).$$

The right-hand side equals $SH^*_{\mathrm{Viterbo}}(U;R)$ by the same argument as Theorem 10.2. In a similar way, $SH^*(\widehat{U}\,|\,\mathrm{Skel}(U,Z);R) = SH^*_{\mathrm{Viterbo}}(U;R)$ by considering a sequence $H_i \in \mathcal{C}$ such that

$$\lim_{x\to\infty} H_i(x) = \begin{cases} 0 & \text{if } x \in \mathrm{Skel}(U,Z), \\ \infty & \text{if } x \in \widehat{U} \setminus \mathrm{Skel}(U,Z). \end{cases}$$

Finally, the equality $SH^*_{\mathrm{Viterbo}}(U;R) = SH^*_{\mathrm{univ}}(\widehat{U};R)$ is Theorem 10.2. $\square$

**Proof of Theorem 2.14** We consider the radial coordinate $t = e^r$ on $U$, which we may assume surjects onto $(0,1)$, with $\mathrm{Skel}(U)$ corresponding to $t=0$. We use the notation $U(t_0) := \{p \in U \mid t(p) \le t_0\}$. We will consider a family of dissipative $S$–shaped Hamiltonians $H_{c,\epsilon}$, which are defined as follows. $H$ is equal to 0 on $U(\epsilon)$, to $ct - c\epsilon$ on $U\left(\frac{1}{2}\right) \setminus U(\epsilon)$, and has small gradient and Hessian outside $U\left(\frac{1}{2}\right)$. Here it is understood that we perturb slightly to get a smooth Hamiltonian which is transversely nondegenerate on $U\left(\frac{1}{2}\right)$. By Theorem 1.1(c), the Hamiltonians $H_{c,\epsilon}$ are dissipative. We construct a monotone increasing sequence $c_i$ going to $\infty$ and a monotone decreasing sequence $\epsilon_i$ going to 0, so that the distance of $c_i$ to the period spectrum of $\partial U$ is more than $2\epsilon_i$. We take $\epsilon_i$ even smaller so that the energy required according to Theorem 6.3 for a Floer trajectory to meet both sides of $U(1) \setminus U\left(\frac{1}{2}\right)$ is more than $\epsilon_i c_i$. Observe now that by our assumption, the action functional on $M$ restricted to loops in $U$ which are contractible in $M$ coincides up to a constant with the action functional defined using the Liouville form. Moreover, the periodic orbits outside of $U\left(\frac{1}{2}\right)$ are constants with large value of $H$. Thus, the set of periodic orbits of $H_{c_i,\epsilon_i}$ having nonnegative action are the constants inside $U(\epsilon_i)$ as well as the periodic orbits appearing as the slope goes from 0 to $c_i$. Their actions are all at most $c_i\delta_i$. Thus, the Floer trajectories connecting orbits of nonnegative action all remain inside $U(1)$. So $SH^{*,0}_{[0,\infty)}(M \mid \mathrm{Skel}(U); R) = SH^{*,0}_{\mathrm{Viterbo}}(U; R)$.

It remains to show that the negative-action periodic orbits form an acyclic complex. Consider an increasing 1–parameter family of Hamiltonians $H_t = H_{c(t),\epsilon(t)}$ with $c(t) \to \infty$ as $t \to \infty$, and fix an action window $[a, 0)$. We cannot show that for an arbitrary $a$ there is a fixed $t$ such that $HF^*_{[a,0)}(H_t) = 0$ since as we increase the slope, new negative periodic orbits keep appearing with action not far from 0. However, we claim that for each $t_0$ there is a $t_1 > t_0$ such that, denoting by $f_{t_0,t_1}$ the continuation map from $t_0$ to $t_1$, we have

$$f_{t_0,t_1}(HF^*_{[a,0)}(H_t)) = \{0\}.$$

Indeed, let $\gamma$ be a periodic orbit of $\partial U$ with period $T \le c(0)$. It will appear as a periodic orbit of $H_{c(t),\epsilon(t)}$ of action $\frac{1}{2}(T - c(t)) - \epsilon(t)$. Consider the cohomology class $\alpha(t) = f_{0,t}(\alpha)$. By functoriality of the continuation maps, its action is a monotone decreasing function of $t$. Moreover, since for any $t$ the complex $CF^*_{[a,0)}(H_t)$ is finitely generated, there is a discrete set of points $\{t_i \in [0, \infty)\}$ such that on the interval $(t_i, t_{i+1})$, the cocycle $\alpha$ is represented by the action minimizing cycle $\sum \gamma^i_j$. Let $T^i_j$ be the period of $\gamma^i_j$ as a Reeb orbit of $\partial U$ and let $T^i$ be the maximal of these. Along the interval $(t_i, t_{i+1})$, the action of $\alpha(t)$ will be given by $\frac{1}{2}(T^i - c(t)) - \epsilon(t)$. Since the action of $\alpha(t)$ is nonincreasing, we must have that $T^i$ is nonincreasing. Since $c(t) \to \infty$ it follows that the negative periodic orbits in $U(1)$ eventually fall out of any action window under the continuation maps. The periodic orbits outside of $U(1)$ are constants with action going to negative infinity. This means all the periodic orbits which lie outside $U(1)$ are in the closure of the boundary operator in $SC^*(\{H_i\})$. Upon tensoring $SC^*(\{H_i\})$ with $\Lambda_R$, the same remains true. □

**Proof of Theorem 2.16** Let $K$ be a compactly supported displacing Hamiltonian for $\mathrm{Skel}(U)$. Then $K$ displaces an open neighborhood of $\mathrm{Skel}(U)$, which we may take to be $U$ itself. Let $F$ be a Hamiltonian which vanishes on a neighborhood of $U \cup \operatorname{supp} K$ and has small enough gradient and Hessian to be dissipative and have only critical points as periodic orbits. Let $H_i$ be a sequence of Hamiltonians which vanishes on $\mathrm{Skel}(U)$, increases on $U(\epsilon_i)$ for some $\epsilon_i \to 0$, and becomes a constant $C_i$ outside the $u(\epsilon_i)$, with $C_i \to \infty$. The sequence $H_i + F$ is monotone and converges to $H_{\mathrm{Skel}(U)}$, and so computes $SH^*(M \mid \mathrm{Skel}(U); \mathbb{K})$.

Recall the notation $H_1 \# H_2 := H_1 + H_2 \circ \psi_{H_1}$, where $\psi_H$ is the time 1 Hamiltonian flow of $H$. We have $\psi_{H_1 \# H_2} = \psi_{H_2} \circ \psi_{H_1}$. Observe that the sequence $(H_i + F) \# K$ computes $SH^*(M \,|\, \mathrm{Skel}(U); \mathbb{K})$ as a topological vector space. To see this, note that there is a constant $C$ such that $|H_i \# K - H_i| < C$, so we can factor

$$SC^*(\{H_i + F - C\}) \to SC^*(\{(H_i + F) \# K\} \to SC^*(\{H_i + F + C\}),$$

and vice versa.

$F + K = F \# K$ is also a displacing Hamiltonian which, after slightly perturbing, we can take to be nondegenerate. Moreover, all the fixed points of $(H_i + F) \# K$ coincide with those of $F \# K$. By a standard argument [59], adding $H_i$ has the effect of shifting the action spectrum by $-C_i$. The action spectrum of $F \# K$ is bounded from above since all the positive action orbits lie in the compact set defined by $F = 0$. Thus the whole action spectrum of $(H_i + F) \# K$ moves to negative infinity. So $SH^*(M \,|\, \mathrm{Skel}(U)) = 0$. By Theorem 2.14, this implies $SH^*_{\mathrm{Viterbo}}(U) = 0$. □

**Proof of Theorem 2.17** We consider the family of Hamiltonians $H_{c,\epsilon}$ as in the proof of Theorem 2.14. The periodic orbits of $U$ that are contractible in $U$ embed in an action-preserving manner in $\mathcal{L}M$. We take $\delta > 0$ such that any Floer trajectory of energy at most $\delta$ which meets $U\left(\frac{1}{2}\right)$ is contained in $U(1)$. The classes $[x, A]$ with $x$ a contractible periodic orbit in $U\left(\frac{1}{2}\right)$ and $A$ a path in $\mathcal{L}U(1) \subset \mathcal{L}(M)$ thus form a direct summand of $SC^*_{[0,\delta)}(\{H_{c_i,\epsilon_o}\})$. Moreover, the proof of Theorem 2.14 shows that the differential applied to contractible periodic orbits in $U\left(\frac{1}{2}\right)$ coincides mod $\delta$ with the differential computing $SH^*(\widehat{U} | U)$. This proves the claim. □

**Proof of Theorem 2.18** Consider Hamiltonians as in the proof of Theorem 2.16, and denote by

$$f : SC^*(\{H_i + F\}) \to SC^*(\{(H_i + F) \# K\})$$

a continuation map induced by an appropriate homotopy, by $g$ the continuation map in the other direction, and let $\mathfrak{H}$ be the chain homotopy operator between the identity and $g \circ f$. Fix an action value $a$. By taking $i_0$ large enough we have, as in the proof of Theorem 2.16, that $f_i$ vanishes mod $a$ for all $i \geq i_0$. Therefore, starting our sequences at $i_0$, we have $\mathrm{Id} = \mathfrak{H} \circ d + d \circ \mathfrak{H} \bmod c$ for all $c \geq a$. But $\mathfrak{H}$ can increase the valuation by at most the Hofer norm of $K$. It follows that if $\operatorname{val} \alpha = c > a$ and $d\alpha = 0 \bmod c$, the largest possible window $[c, d)$ for which $\alpha \neq 0 \in HF^*_{[c,d)}(H_U)$ has $d - c < d(U) \leq \|H\|_{\mathrm{Hofer}}$. So taking $a < 0$ and $\delta$ as in Theorem 2.17, we get $d(e) > \delta$. □

## 10.2 Mapping tori

Let $(M,\omega)$ be a compact symplectic manifold and let $\psi\colon M\to M$ be a symplectomorphism. Denote by $M_\psi$ the mapping torus

$$M_\psi := [0,1]\times M/(0,p)\simeq(1,\psi(p)).$$

Let $\widetilde{\omega}$ be the 2–form on $M_{\psi_i}$ obtained by pulling back $\omega$ via projection to $M$, and let

$$\widetilde{M}_\psi := \mathbb{R}\times M_\psi,$$

with the symplectic structure $\widetilde{\omega}+ds\wedge dt$. Denote by $HF^*(M,\psi)$ the fixed-point Floer homology of $\psi$ as introduced in [21]. The closed 1–form $dt$ induces a grading of the Floer homologies by integrating over periodic orbits.

Denote by $S\colon \widetilde{M}_\psi\to\mathbb{R}$ the natural coordinate $(s,t,p)\mapsto s$. Let $f\colon\mathbb{R}\to\mathbb{R}$ be a proper convex function which is linear at infinity of slope at greater than $k$ for some integer $k$. Let $J$ be an almost complex structure for which the map $\pi\colon\widetilde{M}_\psi\to\mathbb{R}\times S^1$ defined by $(s,t,p)\mapsto(s,t)$ is $J$–holomorphic. Let $H=f\circ S$. The following theorem is due to M Abouzaid.

**Theorem 10.5** *Denoting conformal isomorphism by $\simeq$, we have*

$$\overline{HF}^{*,k}(H;\Lambda_{\boldsymbol{R}})=HF^{*,k}(H;\Lambda_{\boldsymbol{R}})\simeq HF^*(M,\psi^k;\Lambda_{\boldsymbol{R}}). \tag{74}$$

**Proof** The Hamiltonian vector field of $S$ is $\partial_t$. So the periodic orbits of $H=f\circ S$ are contained in fibers of $H$ for which $f'$ is an integer. The periodic orbits corresponding to an integer $k$ are the fixed points of $\psi^k$. The periodic orbits corresponding to different values of $k$ have different homotopy classes. Thus the Floer differential only connects orbits within a fiber. The $(H,J)$–Floer trajectories in $\widetilde{M}_\psi$ project under $\pi$ to maps satisfying the inhomogeneous Cauchy–Riemann equation

$$\partial_s S=\partial_t T-1,\quad \partial_s T=\partial_t S.$$

Thus the function $s+it\mapsto S+i(T-t)$ is holomorphic. By the maximum principle, $S$ must be constant. In particular, Floer trajectories connecting orbits within a fiber of $H$ must stay within that fiber. Also, we have $T=t$.

Recall the definition of the differential in fixed-point Floer homology. Namely, for fixed points $x$ and $y$ of $\psi^k$, it counts $J$–holomorphic strips asymptotic to $x$ and $y$ satisfying the boundary condition $u(s,1)=\psi(u(s,0))$. Given such a $u$ we obtain a

Floer cylinder $\widetilde{u}$ in the mapping torus by $\widetilde{u}(s,t) := (t, u(s,t))$. This sets up a bijection between the Floer trajectories connecting periodic orbits in a fiber and fixed-point holomorphic strips. The rightmost equality in (74) follows. For the other equality note that since $\psi^k$ has a finite number of fixed points, $CF^{*,k}(H;\Lambda_R)$ is finite-dimensional and thus the differential has closed image. □

**Proof of Theorem 2.6** Let $f$ be any proper convex function and let $H = f \circ S$. Consider a monotone sequence of convex functions $f_n$ which are linear of slope larger than $k$ near infinity, and which converge to $f$. Write $H_n := f_n \circ S$.

Since Floer trajectories remain in fibers of $S$, we have by the isomorphism (74) that

$$HF^{*,k}(H) = HF^{*,k}(H_n) = HF^*(M, \psi^k).$$

By the same reasoning, given convex functions $f_1 \le f_2$ and writing $H_i = f_i \circ S$ for $i = 1, 2$, we get that the natural map $HF^{*,k}(H_1) \to HF^{*,k}(H_2)$ is just the identity under the above identification.

It follows that

$$SH^{*,k}_{\mathrm{univ}}(\widetilde{M}_\psi) = HF^{*,k}(M, \psi^k).$$

Observe that $SH^*_{\mathrm{univ}}$ amounts to completing the direct sum

$$\bigoplus_{k\in\mathbb{Z}} SH^{*,k}_{\mathrm{univ}}(\widetilde{M}_\psi)$$

by allowing certain infinite sums. The claim follows. □

## 10.3 The Künneth formula for split Hamiltonians

For $i = 1, 2$, let $M_i$ be symplectic manifolds and let $(H_i, J_i)$ be dissipative Floer data on $M_i$. Unless $(H_i, J_i)$ are strictly bounded, the data $(H_1 \circ \pi_1 + H_2 \circ \pi_2, \pi_1^* J_1 + \pi_2^* J_2)$ will not be i–bounded. In that case we replace $CF^*(H)$ via the telescope construction by a sequence of Hamiltonians which are strictly bounded, and continue to denote this by $CF^*(H)$. We have

$$\widehat{CF}^*(H_1 \circ \pi_1 + H_2 \circ \pi_2, \pi_1^* J_1 + \pi_2^* J_2) = CF^*(H_1, J_1) \mathbin{\widehat{\otimes}} CF^*(H_2, J_2), \tag{75}$$

where the hat denotes here and later the complete tensor product. This is defined by taking the Banach norm $\|\cdot\|$ on the tensor product $X \otimes Y$ to be defined by

$$\|z\| := \inf\Bigl\{\max_i\{\|x_i\|\|y_i\|\} : z = \sum x_i \otimes y_i\Bigr\} \quad \text{for } z \in X \otimes Y.$$

It is straightforward to verify using (54) that this is indeed the norm induced by (75).

**Theorem 10.6** *We have a natural isometry of Banach spaces over $\Lambda_R$,*

$$\overline{HF}^*(H_1+H_2;\Lambda_R)=\overline{HF}^*(H_1;\Lambda_R)\mathbin{\hat{\otimes}}\overline{HF}^*(H_2;\Lambda_R).$$

**Proof** We follow the proof of the finite-dimensional case from [28]. Isomorphism (75) induces a norm-preserving map

$$\overline{HF}^*(H_1)\mathbin{\hat{\otimes}}\overline{HF}^*(H_2)\to\overline{H}^*(CF^*(H_1)\mathbin{\hat{\otimes}}CF^*(H_2))=\overline{HF}^*(H_1+H_2).$$

We show that this map is surjective. All spaces considered here are countably generated. In particular, every closed subspace has a closed complement [53, Proposition 10.5]. We thus decompose the chain complexes $CF^*(M_i;H_i)$ into a direct sum $C_i\oplus Z_i$ of chains and cycles, and then further decompose $Z_i=K_i\oplus B_i$, where $B_i=\overline{\partial C_i}$.

Any cycle $\gamma\in CF^*(H_1)\mathbin{\hat{\otimes}}CF^*(H_2)$ is, up to the closure of the boundary, an element of

$$B_1\mathbin{\hat{\otimes}}C_2\oplus K_1\mathbin{\hat{\otimes}}C_2\oplus C_1\mathbin{\hat{\otimes}}K_2\oplus K_1\mathbin{\hat{\otimes}}K_2\oplus C_1\mathbin{\hat{\otimes}}C_2.$$

Now note that the images of the spaces under $\partial$ are contained, respectively, in

$$B_1\mathbin{\hat{\otimes}}B_2,\quad K_1\mathbin{\hat{\otimes}}B_2,\quad B_1\mathbin{\hat{\otimes}}K_2,\quad 0,\quad B_1\mathbin{\hat{\otimes}}C_2\oplus B_2\mathbin{\hat{\otimes}}C_1,$$

which are pairwise disjoint. So each component of the boundary must vanish separately. Thus if $\gamma$ is a cycle it must actually be an element of $K_1\mathbin{\hat{\otimes}}K_2$, up to the closure of the boundary. In particular, the map is indeed surjective. □

## 10.4 The Künneth formula for universal symplectic cohomology

We shall need the following lemma. The author is grateful to Lev Buhovski for its proof.

**Lemma 10.7** *Let $M$ and $N$ be smooth manifolds, and let $P=M\times N$. The set of functions of the form $f\circ\pi_1+g\circ\pi_2$ is cofinal in $C^\infty(P)$.*

**Proof** Take an exhaustion $K_1\subset K_2\subset\cdots\subset M$ of $M$ and an exhaustion $L_1\subset L_2\subset\cdots\subset N$ of $N$ by compact sets, and define positive locally bounded functions $g_1\colon M\to\mathbb{R}$ and $g_2\colon N\to\mathbb{R}$ by $g_1(x)=\max_{K_i\times L_i}f$ and $g_2(y)=\max_{K_r\times L_r}f$, where $i$ is the minimal positive integer such that $x\in K_i$, and $r$ is the minimal positive integer such that $y\in L_r$. Then we have $f(x,y)\leqslant g_1(x)+g_2(y)$ for any $(x,y)\in M\times N$. Now, since $g_1$ and $g_2$ are locally bounded, one can find smooth functions $f_1\colon M\to\mathbb{R}$ and $f_2\colon N\to\mathbb{R}$ such that $g_1(x)\leqslant f_1(x)$ for any $x\in M$, and $g_2(y)\leqslant f_2(y)$ for any $y\in N$, so then we have $f(x,y)\leqslant f_1(x)f_2(y)$ for any $(x,y)\in M\times N$. □

**Proof of Theorem 2.7** By the discussion preceding Theorem 10.6, we have a natural map

$$\varinjlim_{(H_1,H_2)\in\mathcal{H}(M_1)\times\mathcal{H}(M_2)} \overline{HF}^*(H_1;\Lambda_R)\otimes\overline{HF}^*(H_2;\Lambda_R)\to SH^*_{\mathrm{univ}}(M_1\times M_2;\Lambda_R).$$

By Lemma 10.7 we can consider the right-hand side as a direct limit over the same indexing set of split Hamiltonians. So an element of the right-hand side lifts, for some pair $(H_1,H_2)$, to an element of $\gamma\in HF^*(H_1\circ\pi_1+H_2\circ\pi_2)$. By Theorem 10.6 the image of the natural map

$$HF^*(H_1)\otimes HF^*(H_2)\to HF^*(H_1\circ\pi_1+H_2\circ\pi_2)$$

is sequentially dense. The density part of the claim follows. Now suppose some element $x$ of the left-hand side maps to 0 in the right-hand side. Then there is an $(H_1,H_2)$ such that the lift $\tilde{x}$ of $x$ to $\overline{HF}^*(H_1)\otimes\overline{HF}^*(H_2)$ maps to 0 in $\overline{HF}^*(H_1+H_2)$. It follows from Theorem 10.6 that $\tilde{x}=0$. □

**Corollary 10.8** *Suppose that $SH^*_{\mathrm{univ}}(M_1)=\{0\}$. Then $SH^*_{\mathrm{univ}}(M_1\times M_2)=\overline{\{0\}}$.*

## 10.5 Vanishing results

**Theorem 10.9** *Let $M$ be a geometrically bounded manifold such that $c_1(M)=0$. Suppose there exists a proper dissipative nondegenerate Hamiltonian on $M$ carrying no periodic orbits of index 0. Then $SH^*_{\mathrm{univ}}(M;\mathbb{K})=0$.*

**Proof** By definition, the natural map $H^*(M;\mathbb{K})\to SH^*(M;\mathbb{K})$ factors through $HF^*(M;\mathbb{K})$. Since $HF^*(H;\mathbb{K})=0$, we get from Theorem 1.3 that $SH^*_{\mathrm{univ}}(M;\mathbb{K})$ is a unital algebra in which $1=0$. □

**Lemma 10.10** *Let $M$ be a geometrically bounded manifold such that $c_1(M)=0$. The hypotheses of Theorem 10.9 are satisfied if $M$ carries a circle action $\psi_{\theta\in S^1}$ with the following properties:*

(a) *It is generated by a Hamiltonian $H$ which is proper and bounded from below.*

(b) *There is an equivariant compatible geometrically bounded almost complex structure $J$ such that the distance $d(p,\psi_{1/2}(p))$ under $g_J$ is bounded away from 0 outside of a compact set, and $\|\nabla H\|_{g_J}\leq f(H)$ for some function $f:\mathbb{R}\to[1,\infty)$ for which the primitive of $1/f$ is unbounded from above.*

**Proof** Assume that the flow of $H$ has minimal period 1. Our assumptions imply that for any integer $k$, the function $\left(k+\frac{1}{2}\right)H$ is dissipative. Indeed, invariance of $J$ under the flow implies the flow of $H$ is Killing. Thus, by Corollary 5.18 and Lemma 5.16 the metric $g_{J_H}$ is geometrically bounded. The estimate on $d(p,\psi_{1/2}(p))$ implies loopwise dissipativity by Lemma 6.17 and Corollary 6.19. Indeed, in this case, the Lyapunov constant vanishes. Let $P\subset M$ be a connected component of the set of critical points of $H$. Then $P$ is compact and Morse–Bott. Since the flow of $H$ is 1–periodic, the Robbin–Salamon index $i_{\mathrm{RS}}(p)$ for any $p\in P$ is related to the Morse index of $H$ by $2i_{\mathrm{Morse}}(p)+\dim P=i_{\mathrm{RS}}(p)+2n$. Suppose first that $i_{\mathrm{RS}}(p)\neq 0$. The Robbin–Salamon index is additive with respect to concatenation, and invariant under reparametrization. Thus, the absolute value of the Robbin–Salamon index of the critical points $p\in P$ can be made arbitrarily large by multiplying $H$ by a large enough constant. Suppose now that $i_{\mathrm{RS}}(p)=0$. Then $0\le i_{\mathrm{Morse}}(p)=n-\frac{1}{2}\dim P$. We have $\dim P<2n$ since the action is nontrivial. So we can perturb $P$ and obtain fixed points with Robbin–Salamon indices lying in $\left[-\frac{1}{2}\dim P,\frac{1}{2}\dim P\right]\subset(-n,n)$. Since the grading defined in equation (53) (for which the unit has degree 0) is by $i_{\mathrm{RS}}(p)+n$, we get that in either case, for $k$ large enough, there are no periodic orbits of index 0. □

**Remark 10.11** Lemma 10.10 has the curious implication that on a closed symplectic manifold $M$ with $c_1(M)=0$ there are no Hamiltonian circle actions. This is in fact proven in [43].

**Example 10.12** Let $M$ be a toric Calabi–Yau manifold obtained as the symplectic reduction of $\mathbb{C}^N$ by a torus preserving the holomorphic volume. Then $M$ has an induced almost complex structure which preserves the action of the residual torus. With the induced Kahler metric, $M$ can be shown to have bounded geometry, and the circle action given by the diagonal action

$$\theta\cdot[z_1,\dots,z_N]=[e^{i\theta}z_1,\dots,e^{i\theta}z_N]$$

can be shown to satisfy the conditions of Lemma 10.10. Thus $SH^*_{\mathrm{univ}}(M)=0$. This generalizes the vanishing of the symplectic cohomology of $\mathbb{C}^n$ as well as the more general result of [49] concerning the case where $M$ is total space of a negative line bundle over projective space and $c_1(M)=0$.

## 10.6 Existence of periodic orbits

**Proof of Theorem 2.4** Let $H\colon M\to\mathbb{R}$ be a proper smooth function such that $H^{-1}(-\infty,0)=K$. Suppose there is a $\delta>0$ such that the flow of $H$ on $H^{-1}(0,\delta)$ has no periodic orbits representing $\alpha$ in the first part or contractible in the second. We may assume without loss of generality that $H$ has sufficiently small Hessian everywhere so that the only periodic orbits are critical points. Let $h_n\colon \mathbb{R} \to \mathbb{R}$ be a monotone function constructed inductively so that $-1/n < h_n(x) < 0$ for $x \in (-\infty, a]$, $h_n(x) = x + n$ on $(a + \delta/n, \infty)$ and $h_n(x) \geq h_{n-1}(x)$ everywhere. Let $H_n = h_n \circ H$. Note that by our assumption the only periodic orbits of $H_n$ are critical points, or, in the first part, periodic orbits not representing $\alpha$. We have that $H_n$ converges in a monotone way to $H_K$. So, by Lemma 8.17,

$$SH^*(M|K;\mathbb{K}) = \widehat{SC}^*(\{H_n\}) = \varprojlim_a \varinjlim_n HF^*_{[a,\infty)}(H_n;\mathbb{K}).$$

The first part of the theorem now follows, since the complex $\widehat{SC}^{*,\alpha}(\{H_n\})$ computing $SH^{*,\alpha}(M|K)$ is the zero complex. We prove the second part. We claim that for any $n$, any $-\infty < a < b$, and any $x \in HF^*_{[a,b)}(H_n;\mathbb{K})$ supported on critical points lying outside of $K$, there is an $n'$ such that $x \mapsto 0$ in $HF^*_{[a,b)}(H_{n'};\mathbb{K})$. Indeed, if we choose sufficiently generic time-independent almost complex structures we may assume that for any triple of integers $m, n_1, n_2$, any simple Floer trajectory in the differential of $CF^*((1/m)H_{n_i})$ or in the continuation map $CF^*((1/m)H_{n_1}) \to CF^*((1/m)H_{n_2})$ is of the expected dimension. By a standard argument in Floer theory [32], all the solutions are time-independent. Namely, since the Floer data and the asymptotic data are all time-independent, a solution $\tilde{u}$ is either time-independent as well, or part of a nontrivial $S^1$ family of solutions. In the latter case, $\tilde{u}$ is an $m$–fold cover of a simple time-dependent solution $u$ associated to Hamiltonians $\frac{1}{m}H_{n_i}$, which also appears in an $S^1$ family and is thus not of the expected dimension, contradicting the assumption. Any time-independent trajectory is gradient-like for $H$. So if it emanates from a critical point outside of $K$, it remains outside of $K$. Moreover, the action difference for a continuation trajectory going from a critical point $x_1$ of $H_{n_1}$ to a critical point $x_2$ of $H_{n_2}$, both lying outside of $K$, is just

$$-H_{n_2}(x_2) + H_{n_1}(x_1) < -(n_2 - n_1).$$

Thus, if $n_2 - n_1 > \mathrm{val}(x) + a$, then $x$ will map to 0 in $HF^*_{[a,b)}(H_{n_2};\mathbb{K})$. By similar reasoning, if $x$ is supported in $K$, it will map to itself under the obvious identification of critical points of $H_{n_i}$ with those of $H$. The claim follows. □

**Proof of Theorem 2.8(a)** For any compact set $K$, the map $SH^*_{\mathrm{univ}} \to SH^*(M|K)$ is unital. Moreover, the map $SH^*_{\mathrm{univ}} \to SH^*(M|K)$ is continuous by Lemma 9.16. Therefore, $\overline{\{0\}}$ maps to 0 under this map. So the hypothesis implies $SH^*(M|K) = 0$ for all $K \subset M$. The claim follows from Theorem 2.4. □

The proof of part (b) of Theorem 2.8 relies on the following lemma.

**Lemma 10.13** *Under the assumptions of Theorem 2.8(b), we have that for any smooth $J$–proper Hamiltonian $H\colon M \to \mathbb{R}_+$, there is an $a \in \mathbb{R}_+$ such that the set of $x \in [a,\infty)$ for which $H^{-1}(x)$ has a periodic orbit representing $\alpha$ is dense in $[a,\infty)$.*

**Proof** Suppose otherwise. Then there is a monotone increasing unbounded sequence $\{a_i\}$ such that for $x \in (a_{2i}, a_{2i+1})$ the flow of $H$ has no periodic orbits representing $\alpha$. Fix a geometrically bounded almost complex structure $J$. For any $R > 0$ we may assume without loss of generality that $B_R(H^{-1}(a_{2i-1})) \subset H^{-1}(a_{2i})$. Fix a constant $\epsilon$ and consider the set $\mathcal{E}$ of functions $h\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $\|\nabla h \circ H\| < \epsilon$ outside of the segments $(a_{2i}, a_{2i+1})$. Then the set $\{h \circ H \mid h \in \mathcal{E}\}$ is $\preceq$–cofinal in $\mathcal{H}$, where $\preceq$ is as defined in (73). Taking $R$ large enough and epsilon small enough, the Floer data $(h \circ H, J)$ will be dissipative by the proof of Theorem 6.10. Moreover, these compositions have no periodic orbits representing $\alpha$. Thus $SH^{*,\alpha}(M) = 0$, contradicting the assumption. □

**Proof of Theorem 2.8(b)** Suppose otherwise. Then for any $K$ there is a proper Hamiltonian $H$ and real numbers $0 < a < b$ such that $K \subset H^{-1}([0,a])$ and there are no periodic orbits in the interval $(a,b)$. Inductively choose an exhaustion by compact sets $K_i$, and exhaustion Hamiltonians $H_i$ with gaps $(a_i, b_i)$ so that for all $i$ we have

$$K_i \subset H^{-1}([0,a_i]) \subset H^{-1}([0,b_i]) \subset K_{i+1} \quad \text{and} \quad a_i < b_i < a_{i+1}.$$

Let $H$ be any proper Hamiltonian coinciding with $H_i$ on $H_i^{-1}([a_i,b_i])$ and satisfying

$$b_i < H(x) < a_{i+1}$$

on the region

$$\{H_i(x) > b_i\} \cap \{H_{i+1} < a_{i+1}\}.$$

By taking a subsequence, we can assume further that $H$ is $J$–proper. There is no $a > 0$ for which $H$ satisfies nearby existence on $[a,\infty)$, in contradiction to Lemma 10.13. □

## References

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

[2] **M Abouzaid**, *Framed bordism and Lagrangian embeddings of exotic spheres*, Ann. of Math. 175 (2012) 71–185 MR Zbl

[3] **M Abouzaid**, *Family Floer cohomology and mirror symmetry*, from "Proceedings of the International Congress of Mathematicians, II" (S Y Jang, Y R Kim, D-W Lee, I Yie, editors), Kyung Moon Sa, Seoul (2014) 813–836 MR Zbl

[4] **M Abouzaid**, *Symplectic cohomology and Viterbo's theorem*, from "Free loop spaces in geometry and topology" (J Latschev, A Oancea, editors), IRMA Lect. Math. Theor. Phys. 24, Eur. Math. Soc., Zürich (2015) 271–485 MR Zbl

[5] **M Abouzaid**, **P Seidel**, *An open string analogue of Viterbo functoriality*, Geom. Topol. 14 (2010) 627–718 MR Zbl

[6] **M Audin**, **F Lalonde**, **L Polterovich**, *Symplectic rigidity: Lagrangian submanifolds*, from "Holomorphic curves in symplectic geometry" (M Audin, J Lafontaine, editors), Progr. Math. 117, Birkhäuser, Basel (1994) 271–321 MR Zbl

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

[8] **D Auroux**, *Special Lagrangian fibrations, wall-crossing, and mirror symmetry*, from "Surveys in differential geometry, XIII: Geometry, analysis, and algebraic geometry" (H-D Cao, S-T Yau, editors), Surv. Differ. Geom. 13, International, Somerville, MA (2009) 1–47 MR Zbl

[9] **S Bosch**, *Lectures on formal and rigid geometry*, Lecture Notes in Math. 2105, Springer (2014) MR Zbl

[10] **M P do Carmo**, *Riemannian geometry*, Birkhäuser, Boston, MA (1992) MR Zbl

[11] **J Cheeger**, **M Goresky**, **R MacPherson**, *$L^2$–cohomology and intersection homology of singular algebraic varieties*, from "Seminar on differential geometry" (S T Yau, editor), Ann. of Math. Stud. 102, Princeton Univ. Press (1982) 303–340 MR Zbl

[12] **J Cheeger**, **M Gromov**, **M Taylor**, *Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds*, J. Differential Geometry 17 (1982) 15–53 MR Zbl

[13] **Y V Chekanov**, *Lagrangian intersections, symplectic energy, and areas of holomorphic curves*, Duke Math. J. 95 (1998) 213–226 MR Zbl

[14] **K Cieliebak**, **Y Eliashberg**, *From Stein to Weinstein and back: symplectic geometry of affine complex manifolds*, American Mathematical Society Colloquium Publications 59, Amer. Math. Soc., Providence, RI (2012) MR Zbl

[15] **K Cieliebak**, **A Floer**, **H Hofer**, *Symplectic homology, II: A general construction*, Math. Z. 218 (1995) 103–122 MR Zbl

[16] **K Cieliebak**, **U Frauenfelder**, **G P Paternain**, *Symplectic topology of Mañé's critical values*, Geom. Topol. 14 (2010) 1765–1870 MR Zbl

[17] **K Cieliebak**, **V L Ginzburg**, **E Kerman**, *Symplectic homology and periodic orbits near symplectic submanifolds*, Comment. Math. Helv. 79 (2004) 554–581 MR Zbl

[18] **K Cieliebak**, **A Oancea**, *Symplectic homology and the Eilenberg–Steenrod axioms*, Algebr. Geom. Topol. 18 (2018) 1953–2130 MR Zbl

[19] **A Daemi**, **K Fukaya**, *Monotone Lagrangian Floer theory in smooth divisor complements*, *II*, preprint (2018) arXiv 1809.03409

[20] **X Dai**, *An introduction to $L^2$ cohomology*, from "Topology of stratified spaces" (G Friedman, E Hunsicker, A Libgober, L Maxim, editors), Math. Sci. Res. Inst. Publ. 58, Cambridge Univ. Press (2011) 1–12 MR Zbl

[21] **S Dostoglou**, **D A Salamon**, *Self-dual instantons and holomorphic curves*, Ann. of Math. 139 (1994) 581–640 MR Zbl

[22] **O Fabert**, *Contact homology of Hamiltonian mapping tori*, Comment. Math. Helv. 85 (2010) 203–241 MR Zbl

[23] **A Floer**, **H Hofer**, **D Salamon**, *Transversality in elliptic Morse theory for the symplectic action*, Duke Math. J. 80 (1995) 251–292 MR Zbl

[24] **S Ganatra**, *Symplectic cohomology and duality for the wrapped Fukaya category*, preprint (2013) arXiv 1304.7312

[25] **S Ganatra**, *Cyclic homology*, *$S^1$–equivariant Floer cohomology*, *and Calabi–Yau structures*, preprint (2019) arXiv 1912.13510

[26] **S Ganatra**, **J Pardon**, **V Shende**, *Covariantly functorial wrapped Floer theory on Liouville sectors*, Publ. Math. Inst. Hautes Études Sci. 131 (2020) 73–200 MR Zbl

[27] **R E Greene**, **H Wu**, *$C^\infty$ approximations of convex, subharmonic, and plurisubharmonic functions*, Ann. Sci. École Norm. Sup. 12 (1979) 47–84 MR Zbl

[28] **P Griffiths**, **J Harris**, *Principles of algebraic geometry*, Wiley, New York (1978) MR Zbl

[29] **Y Groman**, *The wrapped Fukaya category for semi-toric SYZ fibrations*, preprint (2018) arXiv 1805.03635

[30] **Y Groman**, **W J Merry**, *The symplectic cohomology of magnetic cotangent bundles*, preprint (2018) arXiv 1809.01085

[31] **M Gromov**, *Pseudo holomorphic curves in symplectic manifolds*, Invent. Math. 82 (1985) 307–347 MR Zbl

[32] **H Hofer**, **D A Salamon**, *Floer homology and Novikov rings*, from "The Floer memorial volume" (H Hofer, C H Taubes, A Weinstein, E Zehnder, editors), Progr. Math. 133, Birkhäuser, Basel (1995) 483–524 MR Zbl

[33] **J Kang**, *Symplectic homology of displaceable Liouville domains and leafwise intersection points*, Geom. Dedicata 170 (2014) 135–142 MR Zbl

[34] **H Karcher**, *Riemannian comparison constructions*, from "Global differential geometry" (S S Chern, editor), MAA Stud. Math. 27, Math. Assoc. America, Washington, DC (1989) 170–222 MR Zbl

[35] **G Lu**, *Virtual moduli cycles and Gromov–Witten invariants of noncompact symplectic manifolds*, Comm. Math. Phys. 261 (2006) 43–131 MR Zbl

[36] **W Lück**, *$L^2$–invariants: theory and applications to geometry and $K$–theory*, Ergebnesse der Math. (3) 44, Springer (2002) MR Zbl

[37] **D McDuff**, **D Salamon**, *$J$–holomorphic curves and symplectic topology*, American Mathematical Society Colloquium Publications 52, Amer. Math. Soc., Providence, RI (2004) MR Zbl

[38] **D McDuff**, **D Salamon**, *Introduction to symplectic topology*, 3rd edition, Oxford Univ. Press (2017) MR Zbl

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

[40] **A Nabutovsky**, *Disconnectedness of sublevel sets of some Riemannian functionals*, Geom. Funct. Anal. 6 (1996) 703–725 MR Zbl

[41] **A Oancea**, *A survey of Floer homology for manifolds with contact type boundary or symplectic homology*, from "Symplectic geometry and Floer homology: a survey of the Floer homology for manifolds with contact type boundary or symplectic homology", Ensaios Mat. 7, Soc. Brasil. Mat., Rio de Janeiro (2004) 51–91 MR Zbl

[42] **A Oancea**, *The Künneth formula in Floer homology for manifolds with restricted contact type boundary*, Math. Ann. 334 (2006) 65–89 MR Zbl

[43] **K Ono**, *Some remarks on group actions in symplectic geometry*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 35 (1988) 431–437 MR Zbl

[44] **M Pabiniak**, *Displacing (Lagrangian) submanifolds in the manifolds of full flags*, Adv. Geom. 15 (2015) 101–108 MR Zbl

[45] **P Petersen**, *Riemannian geometry*, 2nd edition, Graduate Texts in Math. 171, Springer (2006) MR Zbl

[46] **S Piunikhin**, **D Salamon**, **M Schwarz**, *Symplectic Floer–Donaldson theory and quantum cohomology*, from "Contact and symplectic geometry" (C B Thomas, editor), Publ. Newton Inst. 8, Cambridge Univ. Press (1996) 171–200 MR Zbl

[47] **A F Ritter**, *Deformations of symplectic cohomology and exact Lagrangians in ALE spaces*, Geom. Funct. Anal. 20 (2010) 779–816 MR Zbl

[48] **A F Ritter**, *Topological quantum field theory structure on symplectic cohomology*, J. Topol. 6 (2013) 391–489 MR Zbl

[49] **A F Ritter**, *Floer theory for negative line bundles via Gromov–Witten invariants*, Adv. Math. 262 (2014) 1035–1106 MR Zbl

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

[51] **J Robbin**, **D Salamon**, *The Maslov index for paths*, Topology 32 (1993) 827–844 MR Zbl

[52] **D Salamon**, *Lectures on Floer homology*, from "Symplectic geometry and topology" (Y Eliashberg, L Traynor, editors), IAS/Park City Math. Ser. 7, Amer. Math. Soc., Providence, RI (1999) 143–229 MR Zbl

[53] **P Schneider**, *Nonarchimedean functional analysis*, Springer (2002) MR Zbl

[54] **M Schwarz**, *Cohomology operations from $S^1$–cobordisms in Floer homology*, PhD thesis, Swiss Federal Institute of Technology Zurich (1995) Available at `http://www.math.uni-leipzig.de/~schwarz/diss.pdf`

[55] **P Seidel**, *Fukaya categories and deformations*, from "Proceedings of the International Congress of Mathematicians" (T Li, editor), volume 2, Higher Ed., Beijing (2002) 351–360 MR Zbl

[56] **P Seidel**, *A biased view of symplectic cohomology*, from "Current developments in mathematics, 2006" (B Mazur, T Mrowka, W Schmid, R Stanley, S-T Yau, editors), International, Somerville, MA (2008) 211–253 MR Zbl

[57] **J-C Sikorav**, *Some properties of holomorphic curves in almost complex manifolds*, from "Holomorphic curves in symplectic geometry" (M Audin, J Lafontaine, editors), Progr. Math. 117, Birkhäuser, Basel (1994) 165–189 MR Zbl

[58] **J Tu**, *On the reconstruction problem in mirror symmetry*, Adv. Math. 256 (2014) 449–478 MR Zbl

[59] **M Usher**, *The sharp energy-capacity inequality*, Commun. Contemp. Math. 12 (2010) 457–473 MR Zbl

[60] **M Usher**, *Boundary depth in Floer theory and its applications to Hamiltonian dynamics and coisotropic submanifolds*, Israel J. Math. 184 (2011) 1–57 MR Zbl

[61] **U Varolgunes**, *Mayer–Vietoris property for relative symplectic cohomology*, Geom. Topol. 25 (2021) 547–642 MR Zbl

[62] **S Venkatesh**, *Rabinowitz Floer homology and mirror symmetry*, J. Topol. 11 (2018) 144–179 MR Zbl

[63] **C Viterbo**, *Functors and computations in Floer homology with applications, I*, Geom. Funct. Anal. 9 (1999) 985–1033 MR Zbl

[64] **C Viterbo**, *Functors and computations in Floer homology with applications*, *II*, preprint (2018) arXiv 1805.01316

*Einstein Institute of Mathematics, The Hebrew University of Jerusalem - Givat Ram*
*Jerusalem, Israel*

ygroman@gmail.com

Proposed: Yakov Eliashberg Received: 20 September 2017
Seconded: Leonid Polterovich, Gang Tian Revised: 2 November 2021

*Geometry & Topology* **27**:4 (2023) 1391–1478
DOI: 10.2140/gt.2023.27.1391
Published: 15 June 2023

# Geodesic coordinates for the pressure metric at the Fuchsian locus

XIAN DAI

We prove that the Hitchin parametrization provides geodesic coordinates at the Fuchsian locus for the pressure metric in the Hitchin component $\mathcal{H}_3(S)$ of surface group representations into $\mathrm{PSL}(3,\mathbb{R})$.

The proof consists of the following elements: We compute first derivatives of the pressure metric using the thermodynamic formalism. We invoke a gauge-theoretic formula to compute the first and second variations of the reparametrization functions by studying flat connections from Hitchin's equations and their parallel transports. We then extend these expressions of integrals over closed geodesics to integrals over the two-dimensional surface. Symmetries of the Liouville measure then provide cancellations, which show that the first derivatives of the pressure metric tensors vanish at the Fuchsian locus.

53B20; 37D35

| | | |
|---|---|---|
| 1. | Introduction | 1391 |
| 2. | Background and notation | 1396 |
| 3. | More thermodynamic formalism | 1411 |
| 4. | Proof of the main theorem: initial steps | 1422 |
| 5. | Computation of the variations for a model case | 1429 |
| 6. | Evaluation on the Poincaré disk for the model case | 1445 |
| 7. | The remaining cases | 1454 |
| References | | 1476 |

## 1 Introduction

The Weil–Petersson metric on Teichmüller space is a central object in classical Teichmüller theory. Quite a bit is known about it: it is a negatively curved real analytic

Kähler metric with isometry group induced from the extended mapping class group (see Ahlfors [1], Tromba [36] and Masur and Wolf [25]). Although it is not complete (see Wolpert [38] and Chu [11]), it resembles a complete negative curved metric and shares many similar nice properties (see Wolpert [38; 39]).

In recent years, considerable attention has focused on higher-rank Teichmüller spaces; see Goldman [13], Hitchin [15] and Labourie [19]. It is natural to seek metric structures on these spaces with the hope that such structure will reflect important properties of the spaces. To that end, Bridgeman, Canary, Labourie and Sambarino [8] have extended the Weil–Petersson metric from Teichmüller space to an analytic Riemannian metric by techniques from thermodynamic formalism, called the pressure metric on Hitchin components. The Hitchin component $\mathcal{H}_n(S)$, defined by Hitchin in [15], is a special component of the representation space of the fundamental group of a closed surface $S$ of genus $g \geq 2$ into $\mathrm{PSL}(n, \mathbb{R})$. In particular, the Teichmüller space $\mathcal{T}(S)$, identified as representations into $\mathrm{PSL}(2, \mathbb{R})$, embeds in this component and is called the Fuchsian locus. To define the pressure metric, we associate a geodesic flow to each Hitchin representation and describe these reparametrized geodesic flows by some Hölder functions, called reparametrization functions. Our pressure metric is defined on the tangent space of a Hitchin component by taking the variance of the first variations of the reparametrization functions that record the infinitesimal change of the representations.

Bridgeman, Canary, Labourie and Sambarino have proved that the pressure metric in fact restricts to a multiple of the Weil–Petersson metric on the Fuchsian locus and is invariant under the action of the mapping class group. Despite this nice coincidence, very little is presently known about the pressure metric. Some $C^0$ properties of the pressure metric have recently been identified by Labourie and Wentworth [20]. In particular, they show that, when restricted to the Fuchsian locus, the pressure metric is proportional to a Petersson-type pairing for variation given by holomorphic differentials. Building upon their work, our goal in this paper is to investigate some variational $C^1$ properties of the pressure metric using tools from thermodynamic formalism.

One may be curious to what extent the pressure metric in Hitchin components resembles Weil–Petersson geometry. Inspired by Ahlfors' work [1] showing the Bers coordinates are geodesic for Weil–Petersson metric, we will show that, for one particular case of the Hitchin component, similar coordinates are geodesic for the pressure metric near the Fuchsian locus. The Hitchin component we consider is $\mathcal{H}_3(S)$, which coincides with the space of convex real projective structures; see Choi and Goldman [10]. It is a prototypical example of higher-rank Teichmüller spaces. We expect similar results will hold for general cases of Hitchin components $\mathcal{H}_n(S)$.

Inspired by the methods of Labourie and Wentworth [20] for the $C^0$ properties of the pressure metric, we will find and evaluate expressions for the derivatives of the pressure metric at the Fuchsian locus for the case of $\mathrm{PSL}(3,\mathbb{R})$ and its Hitchin component $\mathcal{H}_3(S)$.

The coordinates we choose are very natural in the setting of Hitchin components from a Higgs bundle perspective. Picking $(q_1,\dots,q_{6g-6})$ to be a basis for $H^0(X,K^2)$ over $\mathbb{R}$ and $(q_{6g-5},\dots,q_{16g-16})$ to be a basis for $H^0(X,K^3)$ over $\mathbb{R}$, every element of $\mathcal{H}_3(S)$ corresponds to some

$$m(\xi)=\xi_1q_1+\cdots+\xi_lq_l$$

with $\xi=(\xi_1,\dots,\xi_l)\in\mathbb{R}^l$ and $l=16g-16$.

The $\xi_i$ are coordinate functions and the coordinate system is realized by the Hitchin parametrization $\mathcal{H}_3(S)\cong H^0(X,K^2)\oplus H^0(X,K^3)$. The Hitchin parametrization is given by the Hitchin section of the Hitchin fibration, which was defined by Hitchin in [15] and will be explained in the next section.

We will show:

**Theorem 1.1** *Let $S$ be a closed oriented surface with genus $g\geq 2$. For any point $\sigma\in\mathcal{T}(S)\subset\mathcal{H}_3(S)$, let $X$ be the Riemann surface corresponding to $\sigma$. Then the Hitchin parametrization $H^0(X,K^2)\oplus H^0(X,K^3)$ provides geodesic coordinates for the pressure metric at $\sigma$.*

More explicitly, if we denote components of the pressure metric at $\sigma$ by $g_{ij}(\sigma)$ with respect to the coordinates given by Hitchin parametrization, then $\partial_k g_{ij}(\sigma)=0$ for all possible $i$, $j$ and $k$ ranging from 1 to $16g-16$.

The proof will be a combination of techniques from the theory of thermodynamic formalism and the theory of Higgs bundles. On the one hand, we will use thermodynamic formalism to study the pressure metric and investigate its $C^1$ properties. On the other hand, reparametrization functions and their variations need to be understood via their Higgs bundle invariants. We now outline some important ingredients of our computations and proofs.

Since there are two types of tangential directions in $\mathcal{H}_3(S)$ — directions given by quadratic differentials and directions given by cubic differentials (corresponding to directions along the Fuchsian locus and transverse to it, respectively) — the derivatives of the metric tensor will be divided into different cases according to this distinction:

- The vanishing of a few types of first derivative of the metric tensor follows easily from the geometric facts that the Fuchsian locus is a totally geodesic embedding into the Hitchin component and that the Bers coordinates on Teichmüller space are geodesic.

- On the other hand, to compute the bulk of the components, we need to invoke thermodynamic formalism to obtain an explicit formula for first derivatives of the pressure metric. We find a formula for the first variations of the pressure metric by computing third derivatives of pressure functions using the theory of the Ruelle operator. This expression involves the first and second variations of the reparametrization functions.

- We start from studying the first and second variations of the reparametrization functions on closed geodesics. Because vectors tangent to periodic geodesics are dense in tangent bundles of hyperbolic surfaces, the computation of the first and second variations of the reparametrization functions on closed geodesics can be extended to the unit tangent bundle after an argument that the natural extensions are Hölder functions.

- To study the first variations of the reparametrization functions on closed geodesics, we recall a gauge-theoretic formula from [20]. We then interpret the resulting formula as defining a system of homogeneous ordinary differential equations, which we proceed to solve.

- Finding the second variations of the reparametrization functions is equivalent to understanding the first variations of our gauge-theoretic formula from the previous paragraph. The difficulty here is in describing how projections onto the eigenvectors for the holonomy map vary when we have a family of representations in the Hitchin component. Indeed, it turns out that we need to understand the variations of all of the eigenvectors of our holonomy map. We interpret this problem in terms of solving a system of nonhomogeneous ordinary differential equations with suitable boundary conditions, which we then proceed to solve.

- For some types of metric tensors that involve both the tangential directions and transverse directions to the Fuchsian locus, analyzing flat connections associated to these directions require understanding the corresponding harmonic metrics that are solutions of Hitchin's equations. The harmonic metrics are no longer diagonalizable when leaving the Fuchsian locus along these mixed directions. We break up the infinitesimal version of Hitchin's equation system and obtain nine scalar equations. We analyze them by maximum principles and Bochner techniques to compute the second variations of the reparametrization functions.

- The evaluation of first derivatives of the pressure metric can be lifted to the Poincaré disk following an idea from [20]. Here is where it becomes important that we are taking first derivatives of the pressure metric rather than zero derivatives of the pressure metric. In particular, we find formulas involving iterated integrals of these holomorphic differentials. Specifying a point on the unit tangent bundle, we can identify the Poincaré disk as our coordinate chart and write down the analytic expansions of our holomorphic differentials on this chart. Using geodesic flow invariance and rotational invariance of the Liouville measure, we find that no nonzero coefficients of our analytic expansions remain after integration.

There are more cases of tangential directions along Fuchsian locus in $\mathcal{H}_n(S)$ for $n \geq 4$, where the harmonic metrics are not known to be diagonalizable. Despite the fact that this makes the analysis difficult, the $n = 3$ case suggests the following conjecture:

**Conjecture 1.2** *Let $S$ be a closed oriented surface with genus $g \geq 2$ and $n \geq 4$. For any point $\sigma \in \mathcal{T}(S) \subset \mathcal{H}_n(S)$, let $X$ be the Riemann surface corresponding to $\sigma$; the Hitchin parametrization $\bigoplus_{i=2}^{i=n} H^0(X, K^i)$ provides geodesic coordinates for the pressure metric at $\sigma$.*

Recently, a Riemannian metric in $\mathcal{H}_n(S)$ associated to periods given by the first simple root length, $L_{\alpha_1}(\rho(\gamma)) = \log\big(\lambda_1(\rho(\gamma))/\lambda_2(\rho(\gamma))\big)$, has been defined by Bridgeman, Canary, Labourie and Sambarino [9], where $\lambda_1(\rho(\gamma))$ and $\lambda_2(\rho(\gamma))$ are the largest and second largest moduli of eigenvalues of $\rho(\gamma)$. This Riemannian metric is called the Liouville pressure quadratic form in [9]. Our methods of computing first derivatives of metric tensors can be applied to the Liouville pressure quadratic form. We expect similar geodesic coordinate results to hold in that setting as well.

**Structure of the article** In Section 2, we recall some fundamental results from the theory of thermodynamic formalism and reparametrizations of geodesic flows. We define the pressure metric. We also introduce Higgs bundles and Hitchin deformation for defining our coordinates in Hitchin components. Section 3 is devoted to preliminary proofs by thermodynamic formalism machinery. We compute the formula for third derivatives of the pressure function. In Section 4, we start the proof of the main theorem and divide the components of first derivatives of metric tensors into several types. We also include a gauge-theoretic formula given by Labourie and Wentworth [20] here. Then, in Section 5, we derive the second variations of the reparametrization functions by studying infinitesimal variation of parallel transport equations. In Section 6, we evaluate the first derivatives of the pressure metric and show they are zero following the steps explained above. We finally generalize the arguments to all types of metric tensors in Section 7.

**Acknowledgements** The author would like to thank her advisor, Michael Wolf, for his help and kind support. The weekly meetings were an important source of encouragement and guidance. The author also wants to thank Martin Bridgeman for his warm introduction to the pressure metric. The author also wants to express her appreciation to Siqi He, Qiongling Li, Andrea Tamburelli and Siran Li for helpful conversations with them. Finally the author acknowledges support from US National Science Foundation (NSF) grant DMS-1564374 as well as *Geometric structures and representation varieties* (the GEAR network). The paper would not have been possible without this support. Finally we would like to thank the reviewers for their careful reading and comments.

## 2 Background and notation

In this section, we develop the notation and background material that we will need. We begin in Section 2.1 with a discussion of reparametrization of geodesic flows. Then, in Section 2.2, we recall the elements of thermodynamic formalism that we will need, and finally, in Section 2.3, we conclude with some notation from the theory of Higgs bundles which arises in our arguments.

Let $S$ be a closed oriented surface with genus $g \geq 2$. We will define all the concepts for introducing the pressure metric in the context of Hitchin components $\mathcal{H}_n(S)$. The reader can find a more general version in [8]. The Hitchin components $\mathcal{H}_n(S)$ will be briefly introduced in Section 2.3.

Equip $S$ with a complex structure $J$ such that $X = (S, J)$ is a Riemann surface and thus a point in Teichmüller space. Let $\sigma$ be the hyperbolic metric in the conformal class of $X$. We denote the unit tangent bundle of $X$ with respect to $\sigma$ by $UX$ and the geodesic flow on $(X, \sigma)$ by $\Phi$.

### 2.1 Reparametrization function

We now introduce how we reparametrize the geodesic flow $\Phi$ by reparametrization functions. In particular, we introduce Livšic's theorem and geodesic flows for Hitchin representations.

Suppose $f: UX \to \mathbb{R}$ is a positive Hölder function and $a$ a closed orbit. We will reparametrize the flow $\Phi$ by the function $f$ so that, for the new flow $\Phi^f$, the flow's direction remains the same everywhere but the speed of the flow changes. In particular, for a $\Phi$–periodic orbit $a$, denoting its period with respect to $\Phi$ by $l(a)$, we want the period of $a$ for the new flow $\Phi^f$ to be

$$l_f(a) = \int_0^{l(a)} f(\Phi_s(x))\,ds,$$

where $x$ is any point on $a$.

This leads to the following definition of reparametrization:

**Definition 2.1** Let $f: UX \to \mathbb{R}$ be a positive Hölder continuous function. We define the reparametrization of $\Phi$ by $f$ to be the flow $\Phi^f$ on $UX$ such that, for any $(x,t) \in UX \times \mathbb{R}$,

$$\Phi_t^f(x) = \Phi_{\alpha_f(x,t)}(x),$$

where $\kappa_f(x,t) = \int_0^t f(\Phi_s(x))\,ds$ and $\alpha_f: UX \times \mathbb{R} \to \mathbb{R}$ satisfies

$$\alpha_f(x, \kappa_f(x,t)) = t.$$

**Remark 2.2** Suppose $O$ is the set of periodic orbits of $\Phi$. If $a \in O$, then its period as a $\Phi_t^f$–periodic orbit is $l_f(a)$ because

$$\Phi_{l_f(a)}^f(x) = \Phi_{\alpha_f(x, l_f(a))}(x) = \Phi_{l(a)}(x) = x.$$

We introduce Livšic cohomology classes [22]. Livšic-cohomologous Hölder functions turn out to reparametrize a flow in "equivalent" ways.

Let $C^h(UX)$ denote the set of real-valued Hölder functions on $UX$.

**Definition 2.3** For $f, g \in C^h(UX)$, we say they are Livšic cohomologous if there exists a Hölder continuous function $V: UX \to \mathbb{R}$ that is differentiable in the flow's direction such that

$$f(x) - g(x) = \left.\frac{\partial V(\Phi_T(x))}{\partial t}\right|_{t=0}.$$

If $f$ is Livšic cohomologous to $g$, then we will denote it by $f \sim g$.

We have the following important properties of Livšic-cohomologous functions:

(1) (Livšic's theorem [23]) Two Hölder continuous function $f$ and $g$ are Livšic cohomologous if and only if $l_f(a) = l_g(a)$ for every $a \in O$.

(2) If $f$ and $g$ are Livšic cohomologous, then they have the same integral over any $\Phi$–invariant measure. This is because $\int_{UX} V(\Phi_t(x))\,dm = \text{const}$ for any $\Phi$–invariant measure $m$ and any $t \in \mathbb{R}$.

(3) [17, Proposition.19.2.8] If $f$ and $g$ are positive and Livšic cohomologous, then the reparametrized flows $\Phi^f$ and $\Phi^g$ are Hölder conjugate, ie there exists a Hölder homeomorphism $h\colon UX \to UX$ such that, for all $x \in UX$ and $t \in \mathbb{R}$,
$$h(\Phi_t^f(x)) = \Phi_t^g(h(x)).$$

The procedure of reparametrizing geodesic flows can be applied to Hitchin components $\mathcal{H}_n(S)$ and provides reparametrization functions as codings for representations. This idea was first introduced by Sambarino to study counting problems associated to Anosov representations [33]. It has also been elaborated later in [34; 31] and other work of Sambarino. In the setting we are working in, similar ideas lead to a construction of a geodesic flow $\Phi^\rho$ associated to each (conjugacy class of a) Hitchin representation $\rho \in \mathcal{H}_n(S)$. We refer the reader to [8] for the explicit construction. In particular, this flow relates $\mathcal{H}_n(S)$ to thermodynamic formalism. We will describe here some of the important properties of $\Phi^\rho$:

- $\Phi^\rho$ is an Anosov flow.
- There exists a Hölder function $f_\rho\colon UX \to \mathbb{R}^+$, called the reparametrization function of $\rho$, such that the reparametrized flow $\Phi^{f_\rho}$ of $\Phi$ is Hölder conjugate to $\Phi^\rho$ [33].
- The period of the orbit associated to $[\gamma] \in \pi_1(S)$ is $\log \Lambda_\gamma(\rho)$, where $\Lambda_\gamma(\rho)$ is the spectral radius of $\rho(\gamma)$, ie the largest modulus of the eigenvalues of $\rho(\gamma)$.

**Remark 2.4** One can also reparametrize the geodesic flow by a Hölder function with periods given by simple root lengths $L_{\alpha_1}(\rho(\gamma)) = \log\bigl(\lambda_1(\rho(\gamma))/\lambda_2(\rho(\gamma))\bigr)$, where $\lambda_1(\rho(\gamma))$ and $\lambda_2(\rho(\gamma))$ are the largest and second largest moduli of eigenvalues of $\rho(\gamma)$. This will lead to the Liouville pressure quadratic form, which also gives rise to a Riemannian metric in $\mathcal{H}_n(S)$ (see [9, Theorem 1.6]). However we will mainly focus on the spectrum radius length $\Lambda_\gamma(\rho)$ and its associated pressure metric in this paper.

## 2.2 Thermodynamic formalism

Next we will introduce some concepts arising from the thermodynamic formalism needed for our proofs. The introduction of most of the material here can also be found in [8]. After the introduction, we will define the pressure metric on Hitchin components.

As usual, we let $\Phi$ denote the geodesic flow on a hyperbolic surface $(X, \sigma)$. We denote by $\mathcal{M}^{\Phi}$ the set of $\Phi$–invariant probability measures on $UX$. Recall $l(a)$ denotes the period of the periodic point $a$ with respect to $\Phi$. Let

$$R_T = \{a \text{ closed orbit of } \Phi \mid l(a) \le T\}.$$

**Definition 2.5** The topological entropy of $\Phi$ is defined as

$$h(\Phi) = \limsup_{T\to\infty} \frac{\log \# R_T}{T}.$$

Recall, for a Hölder function $f : UX \to \mathbb{R}$, we write

$$l_f(a) = \int_0^{l(a)} f(\Phi_s(x))\, ds.$$

**Definition 2.6** The topological pressure (or simply pressure) of a continuous function $f : UX \to \mathbb{R}$ with respect to $\Phi$ is defined by

$$\boldsymbol{P}(\Phi, f) = \limsup_{T\to\infty} \frac{1}{T} \log\biggl(\sum_{a\in R_T} e^{l_f(a)}\biggr).$$

**Remark 2.7** From this definition, we see the pressure of a function f only depends on the periods of $f$, ie the collection of numbers $\{l_f(a)\}$ for any $a \in O$. From Livšic's theorem, we conclude the pressure of a function only depends on its Livšic cohomology class.

In statistical mechanics, suppose we are given a physical system with different possible states $i = 1, \dots, n$ and the energies of these states are $E_1, E_2, \dots, E_n$ with probability $p_i$ that state $i$ occurs. When energy is fixed, the principle "nature maximizes entropy $h$" says that the entropy $h(p_1, \dots, p_n) = \sum_{i=1}^{n} -p_i \log p_i$ of the distribution will be maximized with right choices of $p_i$. However, when the physical system is put in contact with a much larger "heat source" which is at a fixed temperature $T$ and energy is allowed to pass between the original system and the heat source, "nature minimizes the free energy" will instead apply by reaching the "Gibbs distribution". The free energy is $E - kTh$, where $k$ is a physical constant and $E = \sum_{i=1}^{n} p_i E_i$ is the average energy. In the thermodynamic formalism, energy potentials $E_i$ of different states are encoded by continuous functions and "Gibbs distributions" for discrete probability spaces are generalized to equilibrium states. The principle "nature minimizes free energy" motivates the following:

**Proposition 2.8** (variational principle) *Denoting the measure-theoretic entropy of $\Phi$ with respect to a measure $m \in \mathcal{M}^{\Phi}$ as $h(\Phi, m)$, the (topological) pressure of a continuous function $f : UX \to \mathbb{R}$ satisfies*

$$\boldsymbol{P}(\Phi, f) = \sup_{m \in \mathcal{M}^{\Phi}} \left( h(\Phi, m) + \int_{UX} f \, dm \right).$$

*In particular, the topological entropy is the supremum of all measure-theoretic entropies,*

$$\boldsymbol{P}(\Phi, 0) = \sup_{m \in \mathcal{M}^{\Phi}} (h(\Phi, m)) = h(\Phi).$$

**Remark 2.9** One can also take Proposition 2.8 as definitions of pressure and topological entropies.

We shall omit the background geodesic flow $\Phi$ in the notation of pressure and simply write

$$\boldsymbol{P}(\cdot) = \boldsymbol{P}(\Phi, \cdot).$$

**Definition 2.10** A measure $m \in \mathcal{M}^{\Phi}$ on $UX$ such that

$$\boldsymbol{P}(f) = h(\Phi, m) + \int_{UX} f \, dm$$

is called an equilibrium state of $f$.

**Proposition 2.11** (Bowen and Ruelle [6]) *For any Hölder function $f : UX \to \mathbb{R}$, with respect to the geodesic flow $\Phi$, there exists a unique equilibrium state for $f$, denoted by $m_f$. Moreover, $m_f$ is ergodic.*

**Remark 2.12** By the definition of equilibrium states, if $f - g$ is Livšic cohomologous to a constant, then $f$ and $g$ have the same equilibrium states.

**Definition 2.13** The equilibrium state $m_0$ for $f = 0$ is called a probability measure of maximal entropy. It is also called the Bowen–Margulis measure of $\Phi$. We also denote it by $m_{\Phi}$. It satisfies

$$\boldsymbol{P}(0) = \boldsymbol{P}(\Phi, 0) = h(\Phi, m_{\Phi}) = h(\Phi).$$

**Remark 2.14** The Liouville measure $m_L$, the normalized Riemannian measure on $UX$, is a probability measure of maximal entropy for geodesic flows of closed hyperbolic manifolds (see [16, Section 2]). Thus, when considering the geodesic flow $\Phi$ of a hyperbolic surface $(X, \sigma)$, we have $m_L = m_{\Phi}$.

Given $f$ a positive Hölder continuous function on $UX$, denoting $h(f) = h(\Phi^f)$ to be the topological entropy of the reparametrized flow $\Phi^f$, we have the following lemma, which allows us to "normalize" a Hölder function to have pressure zero:

**Lemma 2.15** (Sambarino [33]; Bowen and Ruelle [6]) *The pressure satisfies*

$$\boldsymbol{P}(-hf) = 0$$

*if and only if* $h = h(f) = h(\Phi^f)$.

Potrie and Sambarino show, in the Hitchin component $\mathcal{H}_n(S)$, the topological entropy is maximized only along the Fuchsian locus. In particular, it is a constant on the Fuchsian locus.

**Theorem 2.16** (Potrie and Sambarino [31]) *If* $\rho \in \mathcal{H}_n(S)$, *then* $h(\rho) \leq 2/(n-1)$. *Moreover, if* $h(\rho) = 2/(n-1)$, *then* $\rho$ *lies in the Fuchsian locus.*

We start to define variance and covariance which will be important. The convergence of them for mean zero functions is classical.

**Definition 2.17** For $g$ a Hölder continuous function on $UX$ with mean zero with respect to $m_f$ (ie $\int_{UX} g\, dm_f = 0$), the variance of $g$ with respect to $f$ is defined as

$$\operatorname{Var}(g, m_f) = \lim_{T\to\infty} \frac{1}{T} \int_{UX} \left( \int_0^T g(\Phi_s(x))\, ds \right)^2 dm_f(x). \tag{2-1}$$

**Definition 2.18** For $g_1$ and $g_2$ Hölder continuous functions on $UX$ with mean zero with respect to $m_f$ (ie $\int_{UX} g_1\, dm_f = \int_{UX} g_2\, dm_f = 0$), the covariance of $g_1, g_2$ with respect to $f$ is defined as

$$\begin{aligned} \operatorname{Cov}(g_1, g_2, m_f) \\ = \lim_{T\to\infty} \frac{1}{T} \int_{UX} \left( \int_0^T g_1(\Phi_s(x))\, ds \right) \left( \int_0^T g_2(\Phi_s(x))\, ds \right) dm_f(x). \end{aligned} \tag{2-2}$$

Note these expressions are finite:

**Proposition 2.19** *For* $g_1$ *and* $g_2$ *Hölder continuous function on* $UX$ *with mean zero with respect to* $m_f$, *the covariance of* $g_1$ *and* $g_2$ *is finite*:

$$\operatorname{Cov}(g_1, g_2, m_f) < \infty.$$

The convergence is guaranteed by decay of correlations (see [26]).

**Definition 2.20** We define an operator $P_m\colon C^h(UX) \to C^h(UX)$ associated to a probability measure $m$ on $UX$ to be

$$P_m(g)(x) = g(x) - m(g),$$

where we use the notation $m(g) = \int_{UX} g\,dm$ for a probability measure $m$.

The following corollary will be useful:

**Corollary 2.21** *It suffices to have $m_f(g_1) = 0$ and $m_f(g_2) < \infty$ to guarantee the convergence of covariance and*

$$\operatorname{Cov}(g_1, g_2, m_f) = \operatorname{Cov}(g_1, P_{m_f}(g_2), m_f) < \infty. \tag{2-3}$$

The same applies to the case $m_f(g_2) = 0$ and $m_f(g_1) < \infty$.

**Proof** We have

$$\begin{aligned}
&\frac{1}{T}\int_{UX}\biggl(\int_0^T g_1(\Phi_s(x))\,ds\biggr)\biggl(\int_0^T g_2(\Phi_s(x)) - P_{m_f}\bigl(g_2(\Phi_s(x))\bigr)\,ds\biggr)\,dm_f(x)\\
&= \frac{1}{T}\int_{UX}\biggl(\int_0^T g_1(\Phi_s(x))\,ds\biggr)\biggl(\int_0^T m_f(g_2)\,ds\biggr)\,dm_f(x)\\
&= m_f(g_2)\int_{UX}\int_0^T g_1(\Phi_s(x))\,ds\,dm_f(x) && \text{(as } m_f(g_2) \text{ is a constant)}\\
&= m_f(g_2)\int_0^T\int_{UX} g_1(\Phi_s(x))\,dm_f(x)\,ds && \text{(by Fubini's theorem)}\\
&= m_f(g_2)\int_0^T\int_{UX} g_1(x)\,dm_f(x)\,ds && \text{(as } m_f \text{ is } \Phi\text{–invariant)}\\
&= 0.
\end{aligned}$$

Letting $T \to \infty$, we obtain the desired result. □

We will also need the following characterization of covariance for later use:

**Proposition 2.22** (Pollicott [29]) *For $g_1$ and $g_2$ Hölder continuous functions with mean zero with respect to $m_f$* (ie $\int_{UX} g_1\,dm_f = \int_{UX} g_2\,dm_f = 0$), *the covariance of $g_1$ and $g_2$ may also be written as*

$$\operatorname{Cov}(g_1, g_2, m_f) = \lim_{T\to\infty}\int_{UX} g_2(x)\biggl(\int_{-T/2}^{T/2} g_1(\Phi_s(x))\,ds\biggr)\,dm_f(x).$$

**Proof** We have

$$
\begin{aligned}
\operatorname{Cov}(g_1, g_2, m_f) &= \lim_{T\to\infty} \frac{1}{T} \int_{UX} \left(\int_0^T g_1(\Phi_s(x))\,ds\right)\left(\int_0^T g_2(\Phi_s(x))\,ds\right) dm_f(x) \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{UX} \left(\int_{-T/2}^{T/2} g_1(\Phi_s(x))\,ds\right)\left(\int_{-T/2}^{T/2} g_2(\Phi_s(x))\,ds\right) dm_f(x) \\
&\hspace{20em} (\text{as } m_f \text{ is } \Phi\text{–invariant}) \\
&= \lim_{T\to\infty} \int_{-T/2}^{T/2}\int_{UX} g_1(\Phi_t(x)) \frac{1}{T}\left(\int_{-T/2}^{T/2} g_2(\Phi_s(x))\,ds\right) dm_f(x)\,dt.
\end{aligned}
$$

Because $m \in \mathcal{M}^\Phi$, the following does not vary with $s$:

$$
\begin{aligned}
\text{const} &= \lim_{T\to\infty} \int_{-T/2}^{T/2}\int_{UX} g_1(\Phi_t(x)) g_2(\Phi_s(x))\,dm_f(x)\,dt \qquad (\text{for all } s \in \mathbb{R}) \\
&= \lim_{T\to\infty} \int_{-T/2}^{T/2}\int_{UX} g_1(\Phi_t(x)) \frac{1}{S}\left(\int_{-S/2}^{S/2} g_2(\Phi_s(x))\,ds\right) dm_f(x)\,dt \\
&\hspace{16em} \left(\text{average over } s \in \left[-\tfrac{1}{2}S, \tfrac{1}{2}S\right]\right) \\
&= \lim_{S\to\infty}\lim_{T\to\infty} \int_{-T/2}^{T/2}\int_{UX} g_1(\Phi_t(x)) \frac{1}{S}\left(\int_{-S/2}^{S/2} g_2(\Phi_s(x))\,ds\right) dm_f(x)\,dt \\
&= \lim_{T\to\infty} \int_{-T/2}^{T/2}\int_{UX} g_1(\Phi_t(x)) \frac{1}{T}\left(\int_{-T/2}^{T/2} g_2(\Phi_s(x))\,ds\right) dm_f(x)\,dt \\
&= \operatorname{Cov}(g_1, g_2, m_f).
\end{aligned}
$$

In particular, setting $s = 0$ gives

$$\operatorname{Cov}(g_1, g_2, m_f) = \lim_{T\to\infty} \int_{-T/2}^{T/2}\int_{UX} g_1(\Phi_t(x)) g_2((x))\,dm_f(x)\,dt.$$

Rearranging the integrals gives the desired result. □

Higher correlation and higher covariance are introduced for Anosov diffeomorphism in [18]. For geodesic flows, we define:

**Definition 2.23** For $g_1$, $g_2$ and $g_3$ Hölder continuous functions with mean zero with respect to $m_f$, we define the higher covariance by

$$
\begin{aligned}
&\operatorname{Cov}(g_1, g_2, g_3, m_f) \\
&\quad = \lim_{T\to\infty} \frac{1}{T} \int_{UX} \int_0^T g_1(\Phi_t(x))\,dt \int_0^T g_2(\Phi_t(x))\,dt \int_0^T g_3(\Phi_t(x))\,dt\,dm_f(x).
\end{aligned}
$$

Equivalently,

$$\begin{aligned}\mathrm{Cov}(g_1, g_2, g_3, m_f) \\ = \lim_{T\to\infty} \int_{UX} g_1(x)\left(\int_{-T/2}^{T/2} g_2(\Phi_s(x))\,ds\right)\left(\int_{-T/2}^{T/2} g_3(\Phi_s(x))\,ds\right) dm_f(x).\end{aligned}$$

This equivalence is clear from the proof of Proposition 2.22. The convergence of $\mathrm{Cov}(h_1, h_2, h_3, m)$ is guaranteed by "exponential multiple mixing" for geodesic flow on negatively curved compact surfaces (see Pollicott's note [30]). These definitions will be used later when we introduce first derivatives of the pressure metric.

We use the general notation in the sequel

$$\partial_s f(0) = \frac{df(s)}{ds}\bigg|_{s=0}, \quad \partial_s^2 f(0) = \frac{d^2 f(s)}{ds^2}\bigg|_{s=0}. \tag{2-4}$$

If there is more than one parameter, for example $f(s_1, s_2, \dots, s_k)$ and $k \geq 2$, then we specify the indexes that we are taking derivatives of, such as

$$\partial_{s_{i_1}\dots s_{i_j}} f(0) = \frac{\partial^j f(s_1, s_2, \dots, s_k)}{\partial s_{i_1} \cdots \partial s_{i_j}}\bigg|_{s_1=s_2=\dots=0}. \tag{2-5}$$

**Theorem 2.24** (Parry and Pollicott [27]; McMullen [26]) *Let $f_s$ be a smooth family of functions in $C^h(UX)$. Then*:

(1) *The first derivative of $\boldsymbol{P}(f_s)$ at $s = 0$ is given by*

$$\frac{d\boldsymbol{P}(f_s)}{ds}\bigg|_{s=0} = \int_{UX} \partial_s f_0 \, dm_{f_0}. \tag{2-6}$$

(2) *If the first derivative is zero, then*

$$\frac{d^2\boldsymbol{P}(f_s)}{ds^2}\bigg|_{s=0} = \mathrm{Var}(\partial_s f_0, m_{f_0}) + \int_{UX} \partial_s^2 f_0 \, dm_{f_0}. \tag{2-7}$$

(3) *If the first derivative is zero, then $\mathrm{Var}(\partial_s f_0, m_{f_0}) = 0$ if and only if $\partial_s f_0$ is Livšic cohomologous to zero.*

**Remark 2.25** If $f(s,t)$ is a smooth two-parameter family in $C^h(UX)$, then

$$\begin{aligned}\frac{\partial \boldsymbol{P}(f(s,t))}{\partial t\,\partial s}\bigg|_{s=t=0} \\ = \mathrm{Cov}\big(P_{m_{f(0)}}(\partial_s f(0)), P_{m_{f(0)}}(\partial_t f(0)), m_{f(0)}\big) + \int_{UX} \partial_{st} f(0)\, dm_{f(0)}.\end{aligned} \tag{2-8}$$

Define $\mathcal{P}(UX)$ to be the set of pressure zero Hölder functions on $UX$, ie

$$\mathcal{P}(UX) = \{f \in C^h(UX) : \boldsymbol{P}(f) = 0\}.$$

The tangent space of $\mathcal{P}(UX)$ at $f$ is the set

$$T_f\mathcal{P}(UX) = \ker d_f \boldsymbol{P} = \left\{h \in C^h(UX) \;\middle|\; \int_{UX} h\, dm_f = 0\right\}.$$

We define a pressure seminorm on the tangent space of $\mathcal{P}(UX)$ at $f$, by letting:

**Definition 2.26** The pressure seminorm of $g \in T_f\mathcal{P}(UX)$ is defined as

$$\langle g, g\rangle_P = -\frac{\operatorname{Var}(g, m_f)}{\int_{UX} f\, dm_f}.$$

One notices, for $g \in T_f\mathcal{P}(UX)$, the variance $\operatorname{Var}(g, m_f) = 0$ if and only if $g$ is Livšic cohomologous to 0, ie $g \sim 0$.

## 2.3 Higgs bundles and Hitchin deformation

We next introduce all the notation from the theory of Higgs bundles that will arise in our arguments. We also introduce a coordinate system on the Hitchin component at the end of the section.

Recall $S$ is a closed oriented surface with genus $g \geq 2$ and $X = (S, J)$ is a Riemann surface.

**Definition 2.27** A rank $n$ Higgs bundle over $X$ is a pair $(E, \Phi)$, where $E$ is a holomorphic vector bundle of rank $n$ and $\Phi \in H^0(X, \operatorname{End}(E) \otimes K)$ is called a Higgs field. An $\mathrm{SL}(n, \mathbb{C})$–Higgs bundle is a Higgs bundle $(E, \Phi)$ satisfying $\det E = \mathcal{O}$ and $\operatorname{Tr} \Phi = 0$.

**Definition 2.28** (1) A Higgs bundle $(E, \Phi)$ is semistable if every proper $\Phi$–invariant holomorphic subbundle $F$ of $E$ satisfies

$$\frac{\deg(F)}{\operatorname{rank}(F)} \leq \frac{\deg(E)}{\operatorname{rank}(E)}$$

and stable if this inequality is strict.

(2) A semistable Higgs bundle $(E, \Phi)$ is polystable if it decomposes as a direct sum of stable Higgs bundles.

**Theorem 2.29** *It is classical that, for a holomorphic vector bundle $E$ with holomorphic structure $\overline{\partial}_E$ and a Hermitian metric H, there exists a unique connection $\nabla_{\overline{\partial}_E,H}$, called the Chern connection, such that:*

(1) $\nabla^{0,1}_{\overline{\partial}_E,H} = \overline{\partial}_E$.
(2) $\nabla_{\overline{\partial}_E,H}$ *is unitary.*

We will from now on restrict our interest to degree zero Higgs bundles.

**Theorem 2.30** (Hitchin [14]; Simpson [35]) *Let $(E,\Phi)$ be a rank n, degree zero Higgs bundle on X. Then $E$ admits a Hermitian metric $H$ satisfying Hitchin's equation if and only if $(E,\Phi)$ is polystable. Here Hitchin's equation is*

$$F_{\overline{\partial},H} + [\Phi, \Phi^{*H}] = 0, \tag{2-9}$$

*where $F_{\overline{\partial},H}$ is the curvature of the Chern connection $\nabla_{\overline{\partial}_E,H}$ and $\Phi^{*H}$ is the Hermitian adjoint of $\Phi$.*

**Remark 2.31** Define a connection $D_H$ on $(E,\Phi,H)$ as

$$D_H = \nabla_{\overline{\partial}_E,H} + \Phi + \Phi^{*H}. \tag{2-10}$$

$D_H$ is flat if and only if Hitchin's equation is satisfied.

We define the Higgs bundles moduli space and de Rham moduli space as:

**Definition 2.32**
- The space of gauge equivalence classes of polystable $\mathrm{SL}(n,\mathbb{C})$–Higgs bundles is called the moduli space of $\mathrm{SL}(n,\mathbb{C})$–Higgs bundles and is denoted by $\mathcal{M}_{\mathrm{Higgs}}(\mathrm{SL}(n,\mathbb{C}))$.
- The space of gauge equivalence classes of reductive flat $\mathrm{SL}(n,\mathbb{C})$ connections is called the de Rham moduli space and is denoted by $\mathcal{M}_{\mathrm{de\,Rham}}(\mathrm{SL}(n,\mathbb{C}))$.

**Remark 2.33** The Hitchin–Simpson theorem gives a one-to-one correspondence between $\mathcal{M}_{\mathrm{Higgs}}(\mathrm{SL}(n,\mathbb{C}))$ and $\mathcal{M}_{\mathrm{de\,Rham}}(\mathrm{SL}(n,\mathbb{C}))$ from the above remark. It is also called the Hitchin–Kobayashi correspondence.

We will introduce the Hitchin fibration and Hitchin section following Baraglia's work [2]. We refer the reader to [2, Section 2] for a more comprehensive exposition.

Given a principal 3–dimensional subalgebra $\mathfrak{s} = \mathrm{span}\{x, e, \tilde{e}\}$ of $\mathfrak{sl}(n,\mathbb{C})$ consisting of a semisimple element $x$ and regular nilpotent elements $e$ and $\tilde{e}$ with commutation relations

$$[x,e] = e, \quad [x,\tilde{e}] = -\tilde{e}, \quad [e,\tilde{e}] = x,$$

the Lie algebra $\mathfrak{sl}(n,\mathbb{C})$ decomposes into a direct sum of irreducible subspaces under the adjoint representation of $\mathfrak{s}$,

$$\mathfrak{sl}(n,\mathbb{C}) = \bigoplus_{i=1}^{n-1} V_i.$$

We take $e_1,\dots,e_{n-1}$ as highest-weight elements of $V_1,\dots,V_{n-1}$, where $e_1 = e$. With these defined, there exists a basis of $\mathrm{SL}(n,\mathbb{C})$–invariant homogeneous polynomials $p_i$ of degree $i$ on $\mathfrak{sl}(n,\mathbb{C})$, where $2 \le i \le n$, such that, for all elements $f \in \mathfrak{sl}(n,\mathbb{C})$ of the form

$$f = \tilde{e} + \alpha_2 e_1 + \dots + \alpha_n e_{n-1},$$

we have $p_i(f) = \alpha_i$.

**Definition 2.34** The Hitchin fibration is a map from the moduli space of $\mathrm{SL}(n,\mathbb{C})$–Higgs bundles over $X$ to the direct sum of holomorphic differentials given by

$$p\colon \mathcal{M}_{\mathrm{Higgs}}(\mathrm{SL}(n,\mathbb{C})) \to \bigoplus_{i=2}^{i=n} H^0(X,K^i), \quad (E,\Phi)\mapsto (p_2(\Phi),\dots,p_n(\Phi)),$$

where $p_i$ are the homogeneous invariant polynomials defined above.

**Definition 2.35** A Hitchin section $s$ of the Hitchin fibration is a map back from $\bigoplus_{i=2}^{i=n} H^0(X,K^i)$ to $\mathcal{M}_{\mathrm{Higgs}}(\mathrm{SL}(n,\mathbb{C}))$. For $q=(q_2,q_3,\dots,q_n)\in\bigoplus_{i=2}^{i=n} H^0(X,K^i)$, we define $s(q)$ to be a Higgs bundle $E = K^{(n-1)/2}\oplus K^{(n-3)/2}\oplus\dots\oplus K^{(1-n)/2}$ with its Higgs field given by

$$\Phi(q) = \tilde{e} + q_2 e_1 + q_3 e_2 + \dots + q_n e_{n-1}.$$

More explicitly, we have

$$\Phi(q) = \begin{bmatrix} 0 & r_1q_2 & r_1r_2q_3 & r_1r_2r_3q_4 & \cdots & \prod_{i=1}^{n-2} r_iq_{n-1} & \prod_{i=1}^{n-1} r_iq_n \\ 1 & 0 & r_2q_2 & r_2r_3q_3 & \cdots & \cdots & \prod_{i=2}^{n-1} r_iq_{n-1} \\ 0 & 1 & 0 & r_3q_2 & r_3r_4q_3 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & r_{n-1}q_2 \\ 0 & 0 & \cdots & \cdots & 0 & 1 & 0 \end{bmatrix} : E \to E\otimes K,$$

where $r_i = \frac{1}{2}i(n-i)$ and $K^{1/2}$ is a holomorphic line bundle with its square to be the canonical line bundle $K$. The notation for $e_i$ we use here can be found in [2; 21].

**Remark 2.36** There exists an involutive automorphism $\sigma$ on $\mathfrak{sl}(n,\mathbb{C})$ such that

$$\sigma(e_i) = -e_i, \quad \sigma(\tilde{e}) = -\tilde{e}.$$

Composing with the compact real form $\rho$ on $\mathfrak{sl}(n,\mathbb{C})$ given by $\rho(X) = -X^*$, we can obtain the split real involution given by $\lambda = \rho \circ \sigma$. The fixed-point set of $\lambda$ is the real split form $\mathfrak{sl}(n,\mathbb{R})$. A detailed exposition for this can be found in [2].

From the fact that $\lambda(\Phi(q)) = \Phi(q)^*$, one can see the flat connection (2-10) has holonomy in the split real form of $\mathfrak{sl}(n,\mathbb{C})$. Hitchin therefore shows that the Higgs bundles in the image of the Hitchin section have holonomy in $\mathrm{SL}(n,\mathbb{R})$ (see [15]). The representation space of these Higgs bundles up to conjugacy equivalence forms a connected component of the representation variety $\mathrm{Rep}(\pi_1(S),\mathrm{SL}(n,\mathbb{R}))$, called the Hitchin component $\mathcal{H}_n(S)$. Here we recall that the representation variety $\mathrm{Rep}(\pi_1(S),\mathrm{SL}(n,\mathbb{R}))$ is the space of conjugacy classes of reductive representations from $\pi_1(S)$ to $\mathrm{SL}(n,\mathbb{R})$.

**Remark 2.37** The isomorphism between $\mathcal{H}_n(S)$ and $\bigoplus_{i=2}^{i=n} H^0(X,K^i)$ yields a parametrization of the Hitchin component $\mathcal{H}_n(S)$. We call $\bigoplus_{i=2}^{i=n} H^0(X,K^i)$ the Hitchin base. In particular, the tangent space at the Fuchsian point $X$ is identified with the Hitchin base.

Fixing $E = K^{(n-1)/2} \oplus K^{(n-3)/2} \cdots \oplus K^{(1-n)/2}$, we consider the following map as an infinitesimal change of a family of Higgs fields $\Phi_\epsilon$ associated to $q$:

$$\chi\colon \bigoplus_{i=2}^{i=n} H^0(X,K^i) \to \Omega^{1,0}(X,\mathfrak{sl}(\mathfrak{n},\mathbb{R})), \quad \chi(q) = \sum_{i=2}^{n} q_i \otimes e_{i-1}.$$

In particular, the infinitesimal change of a family of flat connections (2-10) in the space $\mathcal{M}_{\mathrm{de\,Rham}}(\mathrm{SL}(n,\mathbb{C}))$ associated to $q$ defines an isomorphism of $\bigoplus_{i=2}^{i=n} H^0(X,K^i)$ with the tangent space of the Hitchin component $T_X\mathcal{H}_n(S)$. Associated to $\chi(q)$, the deformation of flat connections which is the infinitesimal version of (2-10) is:

**Definition 2.38** At the Fuchsian point $X$, we define our Hitchin deformation associated to $q$ to be

$$\varphi(q) := \chi(q) + \lambda(\chi(q)),$$

where $\lambda$ is the antilinear involution for the split real form of $\mathfrak{sl}(\mathfrak{n},\mathbb{C})$ defined above.

This type of deformation will be the tangential objects we consider for the pressure metric.

**Remark 2.39** The Hitchin parametrization in Remark 2.37 gives a coordinate system for $\mathcal{H}_n(S)$ based at $X$. More explicitly, given a basis $\{q_i\}_{i=1}^{i=l}$ of $\bigoplus_{i=2}^{i=n} H^0(X, K^i)$ with $l = 2(n^2-1)(g-1)$, the coordinate system is given by

$$m(\xi) = \xi_1 q_1 + \cdots + \xi_l q_l,$$

where $\xi = (\xi_1, \dots, \xi_l) \in \mathbb{R}^l$. Because of the isomorphism between $\mathcal{H}_n(S)$ and $\bigoplus_{i=2}^{i=n} H^0(X, K^i)$, the vector $\xi = (\xi_1, \dots, \xi_l)$ provides local parameters on $\mathcal{H}_n(S)$ and $\xi_i \colon \mathcal{H}_n(S) \to \mathbb{R}$ is a coordinate function for $1 \le i \le l$.

## 2.4 The pressure metric on Hitchin components

We define the pressure metric for Hitchin components $\mathcal{H}_n(S)$ in this subsection and state some known results about it.

Recall $\mathcal{H}(UX)$ is the space of pressure zero Hölder functions modulo Livšic coboundaries. We relate $\mathcal{H}(UX)$ to the Hitchin component $\mathcal{H}_n(S)$ by the following thermodynamic mapping:

**Definition 2.40** The thermodynamic mapping $\Psi \colon \mathcal{H}_n(S) \to \mathcal{H}(UX)$ from a Hitchin component $\mathcal{H}_n(S)$ to the space $\mathcal{H}(UX)$ of Livšic cohomology classes of pressure zero Hölder functions on $UX$ is defined as

$$\Psi(\rho) = [-h(\rho) f_\rho],$$

where $h(\rho) = h(f_\rho) = h(\Phi^{f_\rho})$ is the topological entropy of the reparametrized flow $\Phi^{f_\rho}$.

The mapping $\Psi$ admits local analytic lifts to the space $\mathcal{P}(UX)$ of pressure zero Hölder functions. In particular, the map $\widetilde{\Psi} \colon \mathcal{H}_n(S) \to \mathcal{P}(UX)$ given by $\widetilde{\Psi}(\rho) = -h(\rho) f_\rho$ is an analytic local lift of $\Psi$. This enables us to pull back the pressure form on $\mathcal{P}(UX)$ to obtain a pressure form on $\mathcal{H}_n(S)$.

We will from now on write $f_\rho^N = -h(\rho) f_\rho$ for the normalized reparametrization function.

Given an analytic family $\{\rho_s\}_{s \in (-1,1)}$ of (conjugacy classes of) representations in the Hitchin component $\mathcal{H}_n(S)$, we define $\dot\rho_0 = \partial_s \rho_0 = \partial_s \rho_s(0)$. Let $\{f_{\rho_s}\}_{s \in (-1,1)}$ be associated reparametrization functions, we pull back the pressure form on $\mathcal{P}(UX)$ to obtain

$$\begin{aligned}\langle \dot{\rho}_0, \dot{\rho}_0 \rangle_P &= \langle d\widetilde{\Psi}(\dot{\rho}_0), d\widetilde{\Psi}(\dot{\rho}_0) \rangle_P \\ &= \left\langle \frac{\partial(-h(\rho_s) f_{\rho_s})}{\partial s}\bigg|_{s=0}, \frac{\partial(-h(\rho_s) f_{\rho_s})}{\partial s}\bigg|_{s=0} \right\rangle_P \\ &= \langle \partial_s (f^N_{\rho_s})(0), \partial_s (f^N_{\rho_s})(0) \rangle_P \\ &= -\frac{\mathrm{Var}(\partial_s (f^N_{\rho_s})(0), m_{f^N_{\rho_0}})}{\int_{UX} f^N_{\rho_0} \, dm_{f^N_{\rho_0}}}.\end{aligned}$$

It is proved in [8] that the pullback pressure form is nondegenerate and thus defines a Riemannian metric on $\mathcal{H}_n(S)$:

**Definition 2.41** If $\{\rho_s\}_{s \in (-1,1)}$ and $\{\eta_t\}_{t \in (-1,1)}$ are two analytic families of (conjugacy classes of) representations in the Hitchin component $\mathcal{H}_n(S)$ such that $\rho_0 = \eta_0$, the pressure metric for $\dot{\rho_0}, \dot{\eta_0} \in T_{\rho_0}\mathcal{H}_n(S)$ is defined as

$$\langle \dot{\rho_0}, \dot{\eta_0} \rangle_P = -\frac{\mathrm{Cov}\big(\partial_s (f^N_{\rho_s})(0), \partial_s (f^N_{\eta_s})(0), m_{f^N_{\rho_0}}\big)}{\int_{UX} f^N_{\rho_0} \, dm_{f^N_{\rho_0}}}.$$

For simplicity, later we will also write $\partial_s (f^N_{\rho_s})(0) = \partial_s f^N_{\rho_0}$ and $\partial_s (f^N_{\eta_s})(0) = \partial_s f^N_{\eta_0}$. The principle is that we always first normalize a family of reparametrization functions to be pressure zero and then take derivatives.

Because of the identification of $\bigoplus_{i=2}^{i=n} H^0(X, K^i)$ with the tangent space of the Hitchin component $T_X\mathcal{H}_n(S)$, our Hitchin deformation $\varphi(q)$ introduced in Definition 2.38 can be thought of as tangent vectors in $T_X\mathcal{H}_n(S)$. With this understood, we introduce the following important results of Labourie and Wentworth [20]:

Let $q_i$ be a holomorphic differential of degree $k$ on $X$ and let $\varphi(q_i)$ be the associated Hitchin deformation. Labourie and Wentworth [20] show the pressure metric satisfies

$$\langle \varphi(q_i), \varphi(q_i) \rangle_P = C(n,k) \langle q_i, q_i \rangle_X,$$

where $C(n,k) > 0$ is a constant that does not depend on $\sigma$ and $\langle q_i, q_i \rangle_X$ is the Petersson pairing

$$\langle q_i, q_i \rangle_X = \int_X q_i \bar{q}_i \sigma^{-k}(z) \, dA_\sigma$$

with $dA_\sigma = \sigma(z) \, dx \wedge dy$ denoting the area form for the hyperbolic metric $\sigma$.

If $q_i$ and $q_j$ are holomorphic differentials of the same degree, then

$$\begin{aligned}\langle\varphi(q_i),\varphi(q_j)\rangle_P &= \tfrac{1}{4}[\langle\varphi(q_i+q_j),\varphi(q_i+q_j)\rangle_P - \langle\varphi(q_i-q_j),\varphi(q_i-q_j)\rangle_P] \\ &= \tfrac{1}{4}C(n,k)\langle q_i+q_j,q_i+q_j\rangle_X - \tfrac{1}{4}C(n,k)\langle q_i-q_j,q_i-q_j\rangle_X \\ &= C(n,k)\langle q_i,q_j\rangle_X.\end{aligned}$$

If $q_i$ and $q_j$ are holomorphic differentials of different degrees on $X$, Labourie and Wentworth [20] show that

$$\langle\varphi(q_i),\varphi(q_j)\rangle_P = 0. \tag{2-11}$$

We denote the pressure metric components with respect to the coordinates introduced in Remark 2.39 by $g_{ij}$. Equivalently, the metric tensor $g_{ij}(\xi)$ means that the pressure metric $\langle\cdot,\cdot\rangle_P$ is evaluated at $\xi$ with tangential vectors parallel to the $q_i$–axis and $q_j$–axis. In particular, at the point $X$, we have $g_{ij}(0) = g_{ij}(\sigma) = \langle\varphi(q_i),\varphi(q_j)\rangle_P$. It is always possible to choose an orthonormal basis $\{q_i\}$ with respect to our pressure metric from the vector space $\bigoplus_{i=2}^{i=n} H^0(X,K^i)$ so that $g_{ij}(\delta) = \delta_{ij}$.

# 3 More thermodynamic formalism

Bowen and Ruelle's work [3; 4; 6] guarantees that many of the results in the thermodynamic formalism proved for subshifts of finite type by the Ruelle operator still hold for Axiom A diffeomorphisms and Axiom A flows. We adopt this idea of simplifying the rather complicated object "flow" by discretizing it and studying a relative simple object "shift" given by symbolic coding. We will compute the formula for the third derivatives of pressure functions using subshifts of finite type. The reader can find an introduction for modeling hyperbolic diffeomorphisms by subshifts of finite type and modeling hyperbolic flows by suspension flows through Markov partition and symbolic dynamics in [5, Sections 3 and 4; 27, Appendix III].

Section 3.1 is devoted to the Ruelle operator and Ruelle–Perron–Frobenius theorem. These are important tools for studying subshifts of finite types. Then, in Section 3.2, we will compute the third derivatives of pressure functions in Lemma 3.8. These will be important for the proof of the main theorem in the next section.

## 3.1 Ruelle operator and others

We start with a cursory introduction to the elements of thermodynamic formalism for subshifts of finite types. A complete description is in [26; 27].

**Definition 3.1** Let $A$ be a $k \times k$ matrix of zeros and ones; we define the associated two-sided shift of finite type $(\Sigma, \sigma_A)$, where $\Sigma$ is the set of sequences

$$\Sigma = \{x = (x_n)_{n=-\infty}^{\infty} : x_n \in \{1, \dots, k\}, n \in \mathbb{Z}, A(x_n, x_{n+1}) = 1\}$$

and $\sigma_A \colon \Sigma \to \Sigma$ is defined by $\sigma_A(x) = y$, where $y_n = x_{n+1}$.

If instead we consider $x = (x_n)_{n=0}^{\infty}$ with the same restriction given by the matrix $A$ and $\sigma(x) = y$, ie $y_n = x_{n+1}$ for $n \geq 0$, then we obtain a one-sided shift of finite type.

The set $\{1, \dots, k\}$ is equipped with the discrete topology and the two-sided (or one-sided) shift space $\Sigma_A$ is equipped with the associated product topology.

Given $\alpha \in (0, 1)$, we can metrize the topology on the two-sided shift space $\Sigma$ by defining a metric $d_\alpha(x, y) = \alpha^N$, where $N$ is the largest nonnegative integer such that $x_i = y_i$ for $|i| < N$. Similarly, we have a metric $d_\alpha$ defined for one-sided shift space.

We let $C(\Sigma)$ be the space of real-valued continuous functions on $\Sigma$ and $C^\alpha(\Sigma)$ be the space of real-valued Hölder functions on $\Sigma$ with Hölder exponent $\alpha$ with respect to $d_\alpha$.

The two-sided (one-sided) shift of finite type $(\Sigma, \sigma_A)$ is called a subshift of finite type if $\sigma_A$ is topologically transitive.

We define the pullback operator on $C^\alpha(\Sigma)$ by $(\sigma_A^* f)(y) = f(\sigma_A(y))$. Similarly to Definition 2.3, we define:

**Definition 3.2** $f_1$ and $f_2$ in $C^\alpha(\Sigma)$ are (Livšic) cohomologous if

$$f_1 - f_2 = f_3 - \sigma_A^* f_3$$

for some $f_3 \in C^\alpha(\Sigma)$.

From now on, we assume our subshift of finite type $(\Sigma, \sigma_A)$ to be one-sided unless otherwise specified.

**Definition 3.3** Given $w \in C^\alpha(\Sigma)$, the Ruelle operator (or transfer operator) on $f \in C^\alpha(\Sigma)$ is defined by

$$\mathcal{L}_w(f)(x) = \sum_{\sigma_A(y)=x} e^{w(y)} f(y).$$

**Theorem 3.4** (Ruelle, Perron and Frobenius) *Suppose $(\Sigma, \sigma_A)$ is topologically mixing (ie $A_{i,j}^M > 0$ for all $i$ and $j$ for some $M > 0$, also called irreducible and aperiodic) and $w \in C^\alpha(\Sigma)$. Then:*

(1) *There is a simple maximal positive eigenvalue* $\rho(\mathcal{L}_w)$ *of* $\mathcal{L}_w \colon C^\alpha(\Sigma) \to C^\alpha(\Sigma)$ *with a corresponding strictly positive eigenfunction* $e^\psi$*:*

$$\mathcal{L}_w(e^\psi) = \rho(\mathcal{L}_w)e^\psi.$$

(2) *The remainder of the spectrum of* $\mathcal{L}_w$ *(excluding* $\rho(\mathcal{L}_w)$*) is contained in a disk of radius strictly smaller than* $\rho(w)$*.*

(3) *There is a unique probability measure* $\mu_w$ *on* $\Sigma$ *such that*

$$\mathcal{L}_w{}^*\mu_w = e^\psi \mu_w.$$

The pressure $\boldsymbol{P}(w)$ of $w$, which can be defined in an analogous way as the pressure of functions on $UX$ by the variational principle Proposition 2.8, turns out to be related to the spectral radius of the Ruelle operator: $\boldsymbol{P}(w) = \log \rho(\mathcal{L}_w)$ (see [5, Theorem 1.22]).

Associated to $\mu_w$ is another measure $m_w = e^\psi \mu_w$. It is called the equilibrium measure of $w$. It is a $\sigma_A$–invariant and ergodic probability measure and satisfies $\mathcal{L}_w{}^* m_w = m_w$.

We will from now on assume $\boldsymbol{P}(w) = 0$. As pressure functions and equilibrium measures depend only on cohomology class, we can modify $w$ by a coboundary so that $\mathcal{L}_w(1) = 1$ and $\mu_w = m_w$. One notices this implies $\mathcal{L}_w(\sigma_A^* f) = f$.

Fixing $m_w$, we define an inner product $\langle f_1, f_2\rangle := \int_\Sigma f_1 f_2 \, dm_w$ on the Banach space $C^\alpha(\Sigma)$.

For convenience, we also write $S_n(f, x) = \sum_{i=0}^{n-1} f(\sigma_A^i x)$.

The following two lemmas are applications of Ruelle operators and will be useful in the next subsection:

**Lemma 3.5** (McMullen [26, Theorems 3.2 and 3.3]) *For any* $g \in C(\Sigma)$ *and* $f \in C^\alpha(\Sigma)$ *with* $\int_\Sigma f \, dm_w = 0$*,*

$$\lim_{n\to\infty} \left\langle g, \frac{S_n(f)^2}{n} \right\rangle = \mathrm{Var}(f, m_w) \int_\Sigma g \, dm_w = 0,$$

*where* $\mathrm{Var}(f, m_w) = \lim_{n\to\infty} (1/n)\langle S_n(f), S_n(f)\rangle$*.*

**Lemma 3.6** *For any* $f \in C^\alpha(\Sigma)$ *with* $\int_\Sigma f \, dm_w = 0$*,*

$$\lim_{n\to\infty} \frac{1}{n} \int_\Sigma (S_n(f))^3 \, dm_w < \infty.$$

**Proof** This proof is similar to Theorem.3.3 of [26]. We have

$$\frac{1}{n} \int_\Sigma (S_n(f))^3 \, dm = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \langle f \circ \sigma_A^i \cdot f \circ \sigma_A^j, f \circ \sigma_A^k \rangle.$$

When $k > j > i$,

$$\begin{aligned}
\langle f\circ\sigma_A^i\cdot f\circ\sigma_A^j, f\circ\sigma_A^k\rangle &= \langle \sigma_A^{*i}(f\cdot f\circ\sigma_A^{j-i}), \sigma_A^{*i}(f\circ\sigma_A^{k-i})\rangle \\
&= \langle f\cdot f\circ\sigma_A^{j-i}, f\circ\sigma_A^{k-i}\rangle \qquad \text{(by $\sigma_A$–invariance of $m_w$)} \\
&= \langle f, f\circ\sigma_A^{j-i}\cdot f\circ\sigma_A^{k-i}\rangle \\
&= \langle f, \sigma_A^{*(j-i)}(f\cdot f\circ\sigma_A^{k-j})\rangle \\
&= \langle \mathcal{L}_w^{j-i}(f), f\cdot f\circ{\sigma_A}^{k-j}\rangle \\
&\qquad\qquad \text{(as $\mathcal{L}_w(\sigma_A^* f) = f$ and $\mathcal{L}_w{}^* m_w = m_w$)} \\
&= \langle f\cdot \mathcal{L}_w^{j-i}(f), f\circ{\sigma_A}^{k-j}\rangle.
\end{aligned}$$

We define a projection operator on $C^\alpha(\Sigma)$ by $P_{m_w}(h)(x) = h(x) - \int_\Sigma h\,dm_w$. Because $P_{m_w}(h)$ has mean zero with respect to $m_w$, the spectrum of the operator $T_w = \mathcal{L}_w\circ P_{m_w}$ lies in a disk of radius $r < 1$ by the Ruelle–Perron–Frobenius theorem.

One has

$$\langle h_1, h_2\circ\sigma\rangle = \langle T_w(h_1), h_2\rangle \tag{3-1}$$

whenever $h_1$ or $h_2$ has mean zero.

Because $f$ is mean zero with respect to $m_w$, $T_w(f) = \mathcal{L}_w(f)$. Moreover,

$$\begin{aligned}
\langle f\cdot \mathcal{L}_w^{j-i}(f), f\circ{\sigma_A}^{k-j}\rangle &= \langle f\cdot T_w^{j-i}(f), f\circ{\sigma_A}^{k-j}\rangle \\
&= \langle T_w^{k-j}(f\cdot T_w^{j-i}(f)), f\rangle \qquad \text{(by (3-1))} \\
&\le \|T^{k-j}\|\|T^{j-i}\|\|f\|^3 \\
&\le C r^{k-i} \qquad\qquad \text{(for some $C > 0$),}
\end{aligned}$$

where the norm for $T$ is the operator norm.

Thus,

$$\begin{aligned}
&\frac{1}{n}\sum_{0\le i<j<k\le n-1} \langle f\circ\sigma_A^i\cdot f\circ\sigma_A^j, f\circ\sigma_A^k\rangle \\
&\qquad\qquad \le \frac{C}{n}\sum_{k=0}^{n-1}\sum_{i=0}^{k}(k-i)r^{k-i} \qquad \text{(by the estimate above)} \\
&\qquad\qquad = \frac{C}{n}\sum_{k=0}^{n-1}\sum_{s=0}^{k} s r^s
\end{aligned}$$

$$= \frac{Cr}{1-r}\left(1 - \frac{1}{n}\sum_{k=1}^{n} r^k\right)$$
$$< \infty \qquad\qquad (\text{when } n \to \infty).$$

This shows $\lim_{n\to\infty}(1/n)\int_\Sigma (S_n(f))^3\, dm_w < \infty$. □

## 3.2 Third derivatives of pressure functions

Our goal in this subsection is to compute the third derivatives of pressure functions in Lemma 3.8. For this, we first need to compute the third derivatives of pressure functions for subshifts of finite type by the method of the Ruelle operator and generalize it to our setting of suspension flows.

We start from introducing suspension flows. We will also recall Bowen's celebrated results, applied to our setting, that suspension flows efficiently model the geodesic flow on $UX$.

**Definition 3.7** Suppose $(\Sigma, \sigma_A)$ is a two-sided shift of finite type. Given a roof function $r\colon \Sigma \to \mathbb{R}^+$, the suspension flow of $(\Sigma, \sigma_A)$ under $r$ is the quotient space

$$\Sigma_r = \{(x,t) \in \Sigma \times \mathbb{R} : 0 \le t \le r(x),\ x \in \Sigma\}/(x, r(x)) \sim (\sigma_A(x), 0)$$

equipped with the natural flow $\sigma^r_{A,s}(x,t) = (x, t+s)$

Any $\sigma_A$–invariant probability measure $m$ on $\Sigma$ induces a natural $\sigma^r_{A,s}$–invariant probability measure on $\Sigma_r$

$$dm_r = \frac{dm\, dt}{\int_\Sigma r\, dm}. \tag{3-2}$$

This correspondence gives a bijection between $\sigma_A$–invariant probability measures and $\sigma^r_{A,s}$–invariant probability measures.

Bowen [3] shows the construction of Markov partitions for Axiom A diffeomorphisms. He then shows how to model Axiom A flows via the Markov partition and symbolic dynamics in [4]. We illustrate the version of this celebrated result in our context (see also [32]): the geodesic flow $\Phi$ admits a Markov coding $(\Sigma_A, \pi, r)$, where $(\Sigma_A, \sigma_A)$ is a topologically mixing two-sided shift of finite type, the roof function $r\colon \Sigma_A \to \mathbb{R}^+$ is Hölder continuous, and the map $\pi\colon \Sigma_A \to UX$ is also Hölder continuous. The suspension flow $\sigma^r_{A,t}$ models $\Phi_t$ effectively in the following sense:

- $\pi$ is surjective.
- $\pi$ is one-to-one on a set of full measure (for any ergodic measure of full support) and on a residual set.
- $\pi$ is finite-to-one.
- $\pi\sigma^r_{A,t} = \Phi_t\pi$ for all $t \in \mathbb{R}$.

Now we are able to state and prove the major result in this subsection:

**Lemma 3.8** *Let $F_s$ be a smooth family in $C^h(UX)$ such that $\boldsymbol{P}(F_0) = 0$ and $\partial_s \boldsymbol{P}(F_s)(0)$. Then*

$$\left.\frac{d^3\boldsymbol{P}(F_s)}{ds^3}\right|_{s=0} = \int_{UX} \partial_s^3 F_0(x)\, dm_{F_0}(x) + \lim_{r\to\infty}\frac{1}{r}\biggl(3\int_{UX}\int_0^r \partial_s F_0(\Phi_t(x))\,dt \int_0^r \partial_s^2 F_0(\Phi_t(x))\,dt\,dm_{F_0}(x) + \int_{UX}\biggl(\int_0^r \partial_s F_0(\Phi_t(x))\,dt\biggr)^3 dm_{F_0}(x)\biggr). \tag{3-3}$$

*In particular, if $F(u,v,w)$ is a smooth three-parameter family of Hölder functions on $UX$ such that $\boldsymbol{P}(F(0,0,0)) = 0$ and all of the first variations of $\boldsymbol{P}(F(u,v,w))$ are zero, then*

$$\begin{aligned}\left.\frac{\partial^3\boldsymbol{P}(F(u,v,w))}{\partial u\,\partial v\,\partial w}\right|_{u=v=w=0} ={}& \int_{UX} \partial_u\partial_v\partial_w F(0)(x)\,dm_{F(0)}(x) \\ &+ \lim_{r\to\infty}\frac{1}{r}\biggl(\int_{UX}\biggl(\int_0^r \partial_u F(0)(\Phi_t(x))\,dt\biggr)\biggl(\int_0^r \partial_v F(0)(\Phi_t(x))\,dt\biggr) \\ &\qquad\cdot\biggl(\int_0^r \partial_w F(0)(\Phi_t(x))\,dt\biggr)dm_{F(0)}(x) \\ &+ \int_{UX}\biggl(\int_0^r \partial_u F(0)(\Phi_t(x))\,dt\biggr)\biggl(\int_0^r \partial_{vw} F(0)(\Phi_t(x))\,dt\biggr)dm_{F(0)}(x) \\ &+ \int_{UX}\biggl(\int_0^r \partial_v F(0)(\Phi_t(x))\,dt\biggr)\biggl(\int_0^r \partial_{uw} F(0)(\Phi_t(x))\,dt\biggr)dm_{F(0)}(x) \\ &+ \int_{UX}\biggl(\int_0^r \partial_w F(0)(\Phi_t(x))\,dt\biggr)\biggl(\int_0^r \partial_{uv} F(0)(\Phi_t(x))\,dt\biggr)dm_{F(0)}(x)\biggr).\end{aligned} \tag{3-4}$$

**Proof** The proof proceeds in two steps. In the first step, we find a formula for the third derivatives of pressure functions for topologically mixing shifts of finite type. In the second step, we show how the computation can be carried to geodesic flows through symbolic coding and suspension flows.

**Step 1** The computation of the first and second derivatives of pressure functions for aperiodic shifts of finite type is shown by Parry and Pollicott [27] using the Ruelle operator. We will give a computation of the third derivative by the same method and then generalize it to our flow case.

Let $(\Sigma_A, \sigma_A)$ be a (one-sided or two-sided) shift of finite type that is topologically mixing. We assume $f_s$ is a smooth family of functions on $C^\alpha(\Sigma_A)$ such that $\boldsymbol{P}(f_0) = 0$ and $\partial_s \boldsymbol{P}(f_s)(0)$. We will prove

$$\begin{aligned}\partial_s^3 \boldsymbol{P}(f_s)(0) = \lim_{n\to\infty} \frac{1}{n} \int_X (S_n(\partial_s f_0))^3 \, dm_{f_0} \\ + \lim_{n\to\infty} \frac{3}{n} \int_X S_n(\partial_s f_0) S_n(\partial_s^2 f_0) \, dm_{f_0} + \int_X \partial_s^3 f_0 \, dm_{f_0}.\end{aligned} \tag{3-5}$$

Any Hölder function on a two-sided shift space is cohomologous to a Hölder function depending only on the corresponding one-sided shift space (see [27, Proposition 1.2]). It suffices to prove (3-5) for one-sided shifts of finite type. We assume $(\Sigma_A, \sigma_A)$ is one-sided and $f_s$ is a smooth family of Hölder functions (with possibly a different Hölder exponent from $\alpha$) on $\Sigma_A$.

We change $f_0$ in its cohomology class so that $\mathcal{L}_{f_0}(1) = 1$.

Following the method in [27], let $Q(s)$ be a projection-valued function which is analytic in $s$ and satisfies

$$\mathcal{L}_{f_s} Q(s) = Q(s) \mathcal{L}_{f_s}.$$

Let $w(s)\colon \Sigma_A \to \mathbb{R}$ be $w(s)(x) := Q(s) \cdot 1$. So

$$\mathcal{L}_{f_s} w(s) = e^{\boldsymbol{P}(f_s)} w(s) \tag{3-6}$$

and $w(0)(x) = Q(0) \cdot 1 = 1$.

Iterate (3-6) $n$ times and take third $s$–derivatives of both sides at $s = 0$:

$$\partial_s^3 \biggl( \sum_{\sigma_A y = x} e^{S_n(f_s)(y)} w(s)(y) \biggr)(0) = \partial_s^3 (e^{n\boldsymbol{P}(f_s)} w(s))(0). \tag{3-7}$$

Notice $\boldsymbol{P}(f_0)=0$, $\partial_s \boldsymbol{P}(f_s)(0)=0$ and $\int_{UX} \partial_s f_0 \, dm_{f_0}=0$. Integrating both sides of (3-7) with respect to $m_{f_0}$ yields

$$\begin{aligned}
&3n\partial_s^2 \boldsymbol{P}(f_s)(0) \int_X \partial_s w(0)\, dm_{f_0} + n\partial_s^3 \boldsymbol{P}(f_s)(0) \\
&= \int_X S_n(\partial_s^3 f_0)\, dm_{f_0} + 3\int_X (S_n(\partial_s f_0)^2 + S_n(\partial_s^2 f_0))\partial_s w(0)\, dm_{f_0} \\
&\qquad + 3\int_X S_n(\partial_s f_0)\partial_s^2 w(0)\, dm_{f_0} + 3\int_X S_n(\partial_s f_0) S_n(\partial_s^2 f_0)\, dm_{f_0} \\
&\qquad\qquad + \int_X S_n(\partial_s f_0)^3\, dm_{f_0}.
\end{aligned}$$

Divide by $n$ and take $n \to \infty$. From ergodicity of $m_{f_0}$, we may evaluate two of the resulting terms:

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \int_X S_n(\partial_s f_0)\partial_s^2 w(0)\, dm_{f_0} &= \int_X \partial_s f_0\, dm_{f_0} \int_X \partial_s^2 w(0)\, dm_{f_0} = 0, \\
\lim_{n\to\infty} \frac{1}{n} \int_X S_n(\partial_s^2 f_0)\partial_s w(0)\, dm_{f_0} &= \int_X \partial_s^2 f_0\, dm_{f_0} \int_X \partial_s w(0)\, dm_{f_0}.
\end{aligned}$$

We also notice that, by applying Lemma 3.5 and the formula for second derivatives of pressure functions,

$$\begin{aligned}
&\partial_s^2 \boldsymbol{P}(f_s)(0) \int_X \partial_s w(0)\, dm_{f_0} \\
&\qquad = \lim_{n\to\infty} \frac{1}{n} \int_X S_n(\partial_s f_0)^2 \partial_s w(0)\, dm_{f_0} + \lim_{n\to\infty} \frac{1}{n} \int_X S_n(\partial_s^2 f_0)\partial_s w(0)\, dm_{f_0}.
\end{aligned}$$

Therefore, we obtain a formal expression

$$\begin{aligned}
\partial_s^3 \boldsymbol{P}(f_s)(0) = \lim_{n\to\infty} \frac{1}{n} \int_X (S_n(\partial_s f_0))^3\, dm_{f_0} + \lim_{n\to\infty} \frac{3}{n} \int_X S_n(\partial_s f_0) S_n(\partial_s^2 f_0)\, dm_{f_0} \\
+ \int_X \partial_s^3 f_0\, dm_{f_0}.
\end{aligned}$$

We observe each term of the right-hand side converges: finiteness of the first limit has been shown in Lemma 3.6 and that of the second is guaranteed by Corollary 2.21.

**Step 2** We now explain how we obtain the flow version of the above formula.

Suppose $F_s$ is a smooth family of functions in $C^h(UX)$ such that $\boldsymbol{P}(F_s)=0$. We have a topologically mixing Markov coding $(\Sigma_A, \pi, r)$ for $UX$. Because of the conjugacy $\pi\sigma^r_{A,t} = \Phi_t \pi$ between geodesic flow and the suspension flow of $(\Sigma_A, \pi, r)$, it suffices to prove (3-3) for $F_s \circ \pi\colon \Sigma_{A,r} \to \mathbb{R}$ on suspension space with pullback measure $\pi^* m_{F_0}$. For simplicity, we still write $F_s \circ \pi$ as $F_s$ and $\pi^* m_{F_0}$ as $m_{F_0}$.

We then want to reduce the problem of proving (3-3) for suspension flows to proving it for subshifts of finite type. We construct a function $\widehat{F}_s\colon \Sigma_A \to \mathbb{R}$ from the function $F_s$ on the suspension space as

$$\widehat{F}_s(x) = \int_0^{r(x)} F_s(x,t)\,dt. \tag{3-8}$$

As $F_s$ and $r$ are Hölder on $\Sigma_{A,r}$ and $\Sigma_A$, respectively, the function $\widehat{F}_s$ is clearly Hölder. Denoting the set of $\sigma_A^r$–invariant probability measures by $\mathcal{M}^{\sigma_A^r}$ and the set of $\sigma_A$–invariant probability measures by $\mathcal{M}^{\sigma_A}$, we have

$$\begin{aligned} \boldsymbol{P}(\sigma_{A,t}^r, F_s) &= \sup_{m_r \in \mathcal{M}^{\sigma_A^r}} \left( h(\sigma_{A,1}^r, m_r) + \int_{\Sigma_{A,r}} F_s\,dm_r \right) \\ &= \sup_{m \in \mathcal{M}^{\sigma_A}} \frac{h(\sigma_A, m) + \int_{\Sigma_A} F_s\,dm}{\int_{\Sigma_A} r\,dm}. \end{aligned}$$

Let $c_s = \boldsymbol{P}(\sigma_{A,t}^r, F_s)$, we have the relation between the pressure function of $F_s$ and the pressure function of $\widehat{F}_s$ (also see [6])

$$\boldsymbol{P}(\sigma_A, \widehat{F}_s - c_s r) = 0. \tag{3-9}$$

Let $\partial_s c_0 = \partial_s(c_s)(0)$ and $\partial_{ss} c_0 = \partial_s^2(c_s)(0)$.

We have the assumption $\partial_s c_0 = 0$. Without loss of generality, we can also assume $\partial_s^2 c_0 = 0$. Otherwise, we consider the family of functions $\widetilde{F}_s := F_s - \frac{1}{2}s^2 \partial_s^2 c_0$. Clearly $\partial_s \boldsymbol{P}(\widetilde{F}_s)(0) = \partial_s^2 \boldsymbol{P}(\widetilde{F}_s)(0) = 0$ and $\partial_s^3 \boldsymbol{P}(\widetilde{F}_s)(0) = \partial_s^3 \boldsymbol{P}(F_s)(0)$.

Now let's take the third $s$–derivative of (3-9) with the assumptions $\partial_s c_0 = \partial_s^2 c_0 = 0$. By (3-5),

$$\begin{aligned} 0 &= \partial_s^3 \boldsymbol{P}(\widehat{F}_s - c_s r)(0) \\ &= \lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_A} (S_n(\partial_s \widehat{F}_0))^3\,dm_{\widehat{F}_0} + \lim_{n\to\infty} \frac{3}{n} \int_{\Sigma_A} S_n(\partial_s \widehat{F}_0) S_n(\partial_s^2 \widehat{F}_0)\,dm_{\widehat{F}_0} \\ &\qquad + \int_{\Sigma_A} (\partial_s^3 \widehat{F}_0 - \partial_s^3 c_0 r)\,dm_{\widehat{F}_0}. \end{aligned}$$

This yields

$$\partial_s^3 c_0 = \partial_s^3 \boldsymbol{P}(\sigma_{A,t}^r, F_s)(0) = \left( \int_{\Sigma_A} r\,dm_{\widehat{F}_0} \right)^{-1} \partial_s^3 \boldsymbol{P}(\sigma_A, \widehat{F}_s)(0).$$

Therefore, proving (3-3) for $F_s$ is equivalent to proving

$$\lim_{r\to\infty}\frac{1}{r}\int_{\Sigma_{A,r}}\left(\int_0^r \partial_s F_0\,dt\right)^3 dm_{F_0}+\lim_{r\to\infty}\frac{3}{r}\int_{\Sigma_{A,r}}\int_0^r \partial_s F_0\,dt\int_0^r \partial_{ss}F_0\,dt\,dm_{F_0}+\int_{\Sigma_{A,r}}\partial_s^3 F_0\,dm_{F_0}$$
$$=\left(\int_{\Sigma_A} r\,dm_{\widehat{F}_0}\right)^{-1}\cdot\left(\lim_{n\to\infty}\frac{1}{n}\int_{\Sigma_A}(S_n(\partial_s\widehat{F}_0))^3\,dm_{\widehat{F}_0}+\lim_{n\to\infty}\frac{3}{n}\int_{\Sigma_A}S_n(\partial_s\widehat{F}_0)S_n(\partial_s^2\widehat{F}_0)\,dm_{\widehat{F}_0}\right)+\left(\int_{\Sigma_A} r\,dm_{\widehat{F}_0}\right)^{-1}\int_{\Sigma_A}\partial_s^3\widehat{F}_0\,dm_{\widehat{F}_0}.$$

Each term on the left is actually equal to the corresponding term on the right. We show here how to obtain

$$\lim_{r\to\infty}\frac{1}{r}\int_{\Sigma_{A,r}}\left(\int_0^r \partial_s F_0(\sigma_t^r(y))\,dt\right)^3 dm_{F_0}(y)=\left(\int_{\Sigma_A} r\,dm_{\widehat{F}_0}\right)^{-1}\lim_{n\to\infty}\frac{1}{n}\int_{\Sigma_A}(S_n(\partial_s\widehat{F}_0)(x))^3\,dm_{\widehat{F}_0}(x). \tag{3-10}$$

The other two terms follow a similar analysis.

To see (3-10), we begin by noting the identity [28], where $y=(x,u)$,

$$\partial_s F_0(\sigma_{A,t}^r(x,u))=\sum_{n\in\mathbb{Z}}\left(\int_0^{r(\sigma_A^n x)}\partial_s F_0(\sigma_A^n x,v)\delta(u+t-v-r^n(x))\,dv\right),$$

where $r^n(x)=r(x)+r(\sigma_A x)+\cdots+r(\sigma_A^{n-1}x)$ for $n>0$ and $r^0(x)=0$ and $r^{-n}(x)=-(r(\sigma_A{}^{-1}x)+\cdots+r(\sigma_A^{-n}x))$ for $n\ge 1$.

One has from Proposition 2.22, the measure correspondence (3-2) and (3-8) that

$$\lim_{r\to\infty}\frac{1}{r}\int_{\Sigma_{A,r}}\left(\int_0^r \partial_s F_0(\sigma_{A,t}^r(y))\,dt\right)^3 dm_{F_0}(y)$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\Sigma_{A,r}}\partial_s F_0(y)\partial_s F_0(\sigma_{A,t}^r(y))\partial_s F_0(\sigma_{A,v}^r(y))\,dm_{F_0}(y)\,dt\,dv$$
$$=\left(\int_{\Sigma_A} r\,dm_{\widehat{F}_0}\right)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\Sigma_A}\int_0^{r(x)}\partial_s F_0(x,u)\partial_s F_0(\sigma_{A,t}^r(x,u))\cdot\partial_s F_0(\sigma_{A,v}^r(x,u))\,du\,dm_{\widehat{F}_0}(x)\,dt\,dv$$

$$\begin{aligned}
&= \left(\int_{\Sigma_A} r\, dm_{\widehat{F}_0}\right)^{-1} \sum_{m,n\in\mathbb{Z}} \int_{\Sigma_A} dm_{\widehat{F}_0}(x) \int_0^{r(x)} \partial_s F_0(x,u)\, du \\
&\qquad\qquad \cdot \int_0^{r(\sigma_A^n x)} \partial_s F_0(\sigma_A^n x, v)\, dv \int_0^{r(\sigma_A^m x)} \partial_s F_0(\sigma_A^m x, v)\, dv \\
&= \left(\int_{\Sigma_A} r\, dm_{\widehat{F}_0}\right)^{-1} \sum_{m,n\in\mathbb{Z}} \int_{\Sigma_A} \partial_s \widehat{F}_0(x) \partial_s \widehat{F}_0(\sigma_A^n x) \partial_s \widehat{F}_0(\sigma_A^m x)\, dm_{\widehat{F}_0}(x) \\
&= \left(\int_{\Sigma_A} r\, dm_{\widehat{F}_0}\right)^{-1} \lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_A} (S_n(\partial_s \widehat{F}_0)(x))^3\, dm_{\widehat{F}_0}(x).
\end{aligned}$$

We therefore obtain a suspension flow version of (3-5) for $F_s$.

The arguments for three-parameter families are the same as the one-parameter case. In fact, since the operator $\partial_u \partial_v \partial_w$ is a symmetric multilinear map in $u$, $v$ and $w$ that is completely characterized by its values on the diagonal, one can deduce (3-4) for multivariable cases directly from (3-3) for one-parameter families. □

Next we introduce a formula for taking derivatives of integrals over varying measures by tools of thermodynamic formalism. This formula will be very useful in later proofs.

**Lemma 3.9** *Suppose $\{f_s\}_{s\in(-1,1)}$ is a smooth family of pressure zero Hölder functions over $UX$ and suppose $\{m_{f_s}\}_{s\in(-1,1)}$ is the associated family of equilibrium states. Suppose furthermore that $\{w_s\}_{s\in(-1,1)}$ is another smooth family of Hölder functions over $UX$. Then*

$$\partial_s\left(\int_{UX} w_s\, dm_{f_s}\right)(0) = \mathrm{Cov}(w_0, \partial_s f_0, m_{f_0}) + \int_{UX} \partial_s w_0\, dm_{f_0}. \tag{3-11}$$

**Proof** We have

$$\begin{aligned}
&\partial_s\left(\int_{UX} w_s\, dm_{f_s}\right)(0) \\
&\quad = \partial_s\left(\frac{\partial \boldsymbol{P}(f_s + t w_s)}{\partial t}\bigg|_{t=0}\right)(0) && \text{(by (2-6))} \\
&\quad = \frac{\partial^2 \boldsymbol{P}(f_s + t w_s)}{\partial s\, \partial t}\bigg|_{s=t=0} \\
&\quad = \mathrm{Cov}(P_{m_{f_0}}(w_0), P_{m_{f_0}}(\partial_s f_0), m_{f_0}) + \int_{UX} \partial_s w_0\, dm_{f_0} && \text{(by (2-8))} \\
&\quad = \mathrm{Cov}(P_{m_{f_0}}(w_0), \partial_s f_0, m_{f_0}) + \int_{UX} \partial_s w_0\, dm_{f_0} \\
&\quad = \mathrm{Cov}(w_0, \partial_s f_0, m_{f_0}) + \int_{UX} \partial_s w_0\, dm_{f_0} && \text{(by Corollary 2.21).}
\end{aligned}$$ □

# 4 Proof of the main theorem: initial steps

We first restate our main theorem:

**Theorem 1.1** *Let $S$ be a closed oriented surface with genus $g \geq 2$. For any point $\sigma \in \mathcal{T}(S) \subset \mathcal{H}_3(S)$, let $X$ be the Riemann surface corresponding to $\sigma$. Then the Hitchin parametrization $H^0(X, K^2) \oplus H^0(X, K^3)$ provides geodesic coordinates for the pressure metric at $\sigma$.*

We want to show $\partial_k g_{ij}(\sigma) = 0$ for the pressure metric components $g_{ij}$ with respect to the coordinates introduced in Remark 2.39 for all possible $i$, $j$ and $k$.

## 4.1 Some geometrical observation

In this subsection, we conclude some derivatives of metric tensors vanish by a geometric observation. Starting from the next section, we will develop a general method to compute first derivatives of the pressure metric via the thermodynamic formalism.

From now on, we restrict ourselves to the Hitchin component $\mathcal{H}_3(S)$. Suppose $\{q_i\}$ is a basis of holomorphic differentials in $H^0(X, K^2) \oplus H^0(X, K^3)$ and suppose $\{\varphi(q_i)\}$ is the associated Hitchin deformation given in Definition 2.38. Recall we use the notation $g_{ij}(\sigma) = \langle \varphi(q_i), \varphi(q_j) \rangle_P$ to emphasize the metric tensor is evaluated at $\sigma \in \mathcal{T}(S)$. We also assume $g_{ij}(\delta) = \delta_{ij}$.

Furthermore, instead of using the Latin letters $i$, $j$ and $k$ to denote arbitrary holomorphic differentials of degree 2 and 3, we let the Latin letters $i$, $j$ and $k$ only refer to quadratic differentials $q_i, q_j, q_k \in H^0(X, K^2)$ from now on. Therefore, the corresponding Hitchin deformations $\varphi(q_i)$, $\varphi(q_j)$ and $\varphi(q_k)$ are tangential directions to the Fuchsian locus in $T_X\mathcal{H}_3(S)$. We use the Greek letters $\alpha$, $\beta$ and $\gamma$ to refer to cubic differentials $q_\alpha, q_\beta, q_\gamma \in H^0(X, K^3)$. Then the corresponding Hitchin deformations $\varphi(q_\alpha)$, $\varphi(q_\beta)$ and $\varphi(q_\gamma)$ are normal directions to the Fuchsian locus in $T_X\mathcal{H}_3(S)$ with respect to the pressure metric.

With the above notation understood, we have in total six types of first derivative of metric tensors that need to be considered: $\partial_k g_{ij}$, $\partial_j g_{i\alpha}$, $\partial_\alpha g_{ij}$, $\partial_i g_{\alpha\beta}$, $\partial_\beta g_{i\alpha}$ and $\partial_\gamma g_{\alpha\beta}$. Our goal is to prove they all vanish.

We first notice the following facts:

(1) $\partial_k g_{ij}(\sigma) = 0$.

To see this, note that the pressure metric is a constant multiple of the Weil–Petersson metric on Teichmüller space $\mathcal{T}(S)$. Because the coordinates system in terms of quadratic differentials from the Hitchin reparametrization agrees with Bers coordinates through second order in the case of $\mathcal{T}(S)$ [37, Corollaries 5.2 and 5.4]. That the Bers coordinates are geodesic [1] for the Weil–Petersson metric implies that, for the pressure metric, $\partial_k g_{ij}(\sigma) = 0$.

(2) $\partial_j g_{\alpha i}(\sigma) = 0$ implies $\partial_\alpha g_{ij}(\sigma) = 0$.

The contragredient involution $\kappa\colon \mathrm{PSL}(3,\mathbb{R}) \to \mathrm{PSL}(3,\mathbb{R})$ given by $\kappa(g) = (g^{-1})^t$ induces an involution $\hat{\kappa}$ on $\mathcal{H}_3(S)$ by $\hat{\kappa}(\rho)(\gamma) = \kappa(\rho(\gamma))$. Because $\hat{\kappa}$ is an isometry of $\mathcal{H}_3(S)$ with respect to the pressure metric and the fixed-point set of $\hat{\kappa}$ is $\mathcal{T}(S)$, the Fuchsian locus is in fact totally geodesic in $\mathcal{H}_3(S)$ (see [7]). So, for $\tilde{\nabla}$ the Levi-Civita connection of the pressure metric and any $X, Y \in T_\sigma\mathcal{T}(S)$, we have

$$\Pi(X,Y) = (\tilde{\nabla}_X Y)^\perp = 0. \tag{4-1}$$

Thus, the Christoffel symbols for the connection $\tilde{\nabla}$ satisfy $\Gamma^\alpha_{ij}(\sigma) = 0$ and, because

$$\begin{aligned}\Gamma^\alpha_{ij} &= \tfrac{1}{2} g^{\beta\alpha}(\partial_j g_{i\beta} + \partial_i g_{j\beta} - \partial_\beta g_{ji}) \quad (\text{since } g_{k\alpha}(\sigma) = 0 \text{ and } g^{k\alpha}(\sigma) = 0)\\ &= \tfrac{1}{2} g^{\alpha\alpha}(\partial_j g_{i\alpha} + \partial_i g_{j\alpha} - \partial_\alpha g_{ji}) \quad (\text{since } g_{\alpha\beta} = \sigma_{\alpha\beta}),\end{aligned}$$

it suffices to know $\partial_j g_{i\alpha}(\sigma) = 0$ and $\partial_i g_{j\alpha}(\sigma) = 0$ to conclude $\partial_\alpha g_{ij}(\sigma) = 0$.

(3) $\partial_\beta g_{\alpha\alpha}(\sigma) = 0$ implies $\partial_\gamma g_{\alpha\beta}(\sigma) = 0$, and $\partial_i g_{\alpha\alpha}(\sigma) = 0$ implies $\partial_i g_{\alpha\beta}(\sigma) = 0$.

This is because

$$\begin{aligned}\partial_\gamma g_{\alpha\beta} &= \tfrac{1}{2}(\partial_\gamma g_{\alpha+\beta,\alpha+\beta} - \partial_\gamma g_{\alpha\alpha} - \partial_\gamma g_{\beta\beta}),\\ \partial_i g_{\alpha\beta} &= \tfrac{1}{2}(\partial_i g_{\alpha+\beta,\alpha+\beta} - \partial_i g_{\alpha\alpha} - \partial_i g_{\beta\beta}).\end{aligned}$$

The remaining four cases left to prove are as follows:

(i) $\partial_\beta g_{\alpha\alpha}(\sigma) = 0$.

(ii) $\partial_i g_{\alpha\alpha}(\sigma) = 0$.

(iii) $\partial_j g_{\alpha i}(\sigma) = 0$.

(iv) $\partial_\beta g_{\alpha i}(\sigma) = 0$.

We will have a general method to prove them. We first give a general formula for first derivatives of the pressure metric in the next subsection. The computation for the model case $\partial_\beta g_{\alpha\alpha}(\sigma)$ will be shown in Sections 5 and 6. The other three cases will be discussed in Section 7.

## 4.2 First derivatives of the pressure metric

This subsection is devoted to a formula for first derivatives of the pressure metric. We also prove we have some freedom to choose representatives for the variations of the reparametrization functions from the Livšic cohomology classes.

Suppose $\{\rho(u,v,w)\}_{(u,v,w)\in\{(-1,1)\}^3}$ is an analytic three-parameter family of representations in the Hitchin component $\mathcal{H}_n(S)$ with basepoint $\rho(0,0,0)\in\mathcal{T}(S)$ corresponding to $X$. Suppose $\{f_{\rho(u,v,w)}\}_{(u,v,w)\in\{(-1,1)\}^3}$ are associated reparametrization functions. For simplicity of notation, we denote the renormalized reparametrization functions by

$$F(u,v,w)=f^N_{\rho(u,v,w)}=-h(\rho(u,v,w))\,f_{\rho(u,v,w)}.$$

We also write $F(0)=F(0,0,0)$ and $\rho(0)=\rho(0,0,0)$.

In the case of the Fuchsian representation, the topological entropy and the reparametrization function are simple. We have $h(\rho(0))=1$ (see Theorem 2.16). Since $\Phi_{\rho(0)}=\Phi$, the reparametrization function $f_{\rho(0)}$ can be chosen to be 1 in the Livšic cohomology class. Therefore, one can choose $F(0)=-1$.

The following characterization of the equilibrium measure for $F(0)$ is important:

**Lemma 4.1** *The equilibrium state $m_{F(0)}$ for $F(0)$ is the Liouville measure $m_L$.*

**Proof** Since the Liouville measure $m_L$ coincides with the Bowen–Margulis measure (Remark 2.14), this follows easily from the variational principle (Proposition 2.8). □

The Liouville measure $m_L$ is both $\Phi_t$–invariant and rotationally invariant on $UX$, ie $(e^{i\theta})^*m_L=m_L$, where $e^{i\theta}$ acts on $UX$ by usual multiplication. We will repeatedly use these important properties of the Liouville measure for our proofs later.

**Proposition 4.2** *The first derivatives of the pressure metric at $\rho(0)$ satisfy*

$$\begin{aligned}\partial_w\big(\langle\partial_u\rho(0,0,w),\partial_v\rho(0,0,w)\rangle_P\big)(0)&\\ =\lim_{r\to\infty}\frac{1}{r}\bigg(\int_{UX}\int_0^r\partial_uF(0)\,dt\int_0^r\partial_vF(0)\,dt\int_0^r\partial_wF(0)\,dt\,dm_0&\\ +\int_{UX}\int_0^r\partial_uF(0)\,dt\int_0^r\partial_{wv}F(0)\,dt\,dm_0&\\ +\int_{UX}\int_0^r\partial_vF(0)\,dt\int_0^r\partial_{wu}F(0)\,dt\,dm_0\bigg)&,\end{aligned}$$

*where the flow $\Phi_t(x)$ is omitted for simplicity.*

**Proof** Starting from the Fuchsian point $\rho(0)$, along the ray with parametrization $\{(0,0,w)\}_{w\in(-1,1)}$, the pressure metric $\langle\,\cdot\,,\cdot\,\rangle_P : T_{(0,0,w)}\mathcal{H}_n(S)\times T_{(0,0,w)}\mathcal{H}_n(S)\to\mathbb{R}$ satisfies

$$\begin{aligned}\langle\partial_u\rho(0,0,w),\partial_v\rho(0,0,w)\rangle_P &= -\frac{\operatorname{Cov}(\partial_u F(0,0,w),\partial_v F(0,0,w),m_{F(0,0,w)})}{\int_{UX} F(0,0,w)\,dm_{F(0,0,w)}}\\ &= -\frac{\partial_v\partial_u \boldsymbol{P}(F(0,0,w)) - \int_{UX}\partial_{uv}F(0,0,w)\,dm_{F(0,0,w)}}{\int_{UX} F(0,0,w)\,dm_{F(0,0,w)}} \quad \text{(by (2-8)).}\end{aligned}$$

We first notice $\int_{UX} F(0)\,dm_0 = -1$ and, from (3-11),

$$\begin{aligned}\partial_w\Big(\int_{UX} F(0,0,w)\,dm_{F(0,0,w)}\Big)(0) &= \operatorname{Cov}(F(0),\partial_w F(0),m_0) + \int_{UX}\partial_w F(0)\,dm_0\\ &= 0.\end{aligned}$$

Therefore,

$$\begin{aligned}&\partial_w\big(\langle\partial_u\rho(0,0,w),\partial_v\rho(0,0,w)\rangle_P\big)(0)\\ &= \partial_w\partial_v\partial_u\boldsymbol{P}(F(0)) - \partial_w\Big(\int_{UX}\partial_{uv}F(0)\,dm_{F(0,0,w)}\Big)(0)\\ &= \partial_w\partial_v\partial_u\boldsymbol{P}(F(0)) - \operatorname{Cov}(\partial_{uv}F(0),\partial_w F(0),m_0) - \int_{UX}\partial_{uvw}F(0)\,dm_0 \quad \text{(by (3-11))}\\ &= \lim_{r\to\infty}\frac{1}{r}\Big(\int_{UX}\int_0^r\partial_u F(0)\,dt\int_0^r\partial_v F(0)\,dt\int_0^r\partial_w F(0)\,dt\,dm_0\\ &\qquad + \int_{UX}\int_0^r\partial_u F(0)\,dt\int_0^r\partial_{vw}F(0)\,dt\,dm_0\\ &\qquad + \int_{UX}\int_0^r\partial_v F(0)\,dt\int_0^r\partial_{wu}F(0)\,dt\,dm_0\Big) \quad \text{(by (3-4)).} \qquad\square\end{aligned}$$

**Proposition 4.3** *The formula in Proposition 4.2 for the first derivatives of the pressure metric only depends on the Livšic class of each component function $\partial_u F(0)$, $\partial_v F(0)$, $\partial_w F(0)$, $\partial_{wv}F(0)$ and $\partial_{wu}F(0)$.*

**Proof** We know from the proof of Proposition 4.2 that

$$\begin{aligned}&\partial_w\big(\langle\partial_u\rho(0,0,w),\partial_v\rho(0,0,w)\rangle_P\big)(0)\\ &\qquad = \partial_w\partial_v\partial_u\boldsymbol{P}(F(0)) - \int_{UX}\partial_{uvw}F(0)\,dm_0 - \operatorname{Cov}(\partial_{uv}F(0),\partial_w F(0),m_0).\end{aligned}$$

By (2-8), in general, if we take two mean-zero Hölder functions $h_1$ and $h_2$ with respect to $m_0$, then

$$\mathrm{Cov}(h_1, h_2, m_0) = \partial_u \partial_v \boldsymbol{P}(F(0) + uh_1 + vh_2)(0).$$

As the value of the pressure function $\boldsymbol{P}$ only depends on the Livšic class, we see changing $h_1$ and $h_2$ in its cohomology class does not change $\mathrm{Cov}(h_1, h_2, m_0)$. In particular, this holds for $\mathrm{Cov}(\partial_{uv} F(0), \partial_w F(0), m_0)$.

Similarly, from (3-3), it is clear that

$$\begin{aligned} \partial_w \partial_v \partial_u \boldsymbol{P}(F(0)) - \int_{UX} \partial_{uvw} F(0)\, dm_0 \\ = \partial_w \partial_v \partial_u \boldsymbol{P}\big(F(0) + u\partial_u F(0) + v\partial_v F(0) + w\partial_w F(0) \\ + uv\partial_{uv} F(0) + uw\partial_{uw} F(0) + vw\partial_{vw} F(0)\big)(0). \end{aligned}$$

Again the above pressure function $\boldsymbol{P}$ does not change value if we change each component function. So, altogether, we know the first derivatives of the pressure metric only depend on the Livšic class of each component function $\partial_u F(0)$, $\partial_v F(0)$, $\partial_w F(0)$, $\partial_{wv} F(0)$ and $\partial_{wu} F(0)$. □

## 4.3 A gauge-theoretical formula

In [20], Labourie and Wentworth show the variations of the reparametrization functions can be expressed by a gauge-theoretical formula. This formula will be crucial for our computation in the next section. We include the formula and its proof here for completeness. We add some assumptions which are natural for our case of Hitchin components $\mathcal{H}_n(S)$.

We consider $(E, H)$ a rank $n$ Hermitian bundle over the surface $S$ equipped with a Riemannian metric $g$. We let $\gamma$ be a closed curve on $S$ with arc-length parametrization $\gamma(t)$. Suppose $D_{A^0}$ is a flat connection on $E$ whose holonomy has distinct eigenvalues along $\gamma$. Suppose $\lambda_\gamma$ is one eigenvalue with a corresponding eigenline $\mathcal{L}_\gamma$ and $\mathcal{H}_\gamma$ is the complementary hyperplane stabilized by the holonomy. We denote by $\mathcal{L}_\gamma(t)$ the line generated by the parallel transports of $\mathcal{L}_\gamma$ along $\gamma$ at time $t$, by $\mathcal{H}_\gamma(t)$ the hyperplane generated by complementary eigenvectors, and by $\pi(t)$ the projection on $\mathcal{L}_\gamma(t)$ along $\mathcal{H}_\gamma(t)$. Then we have:

**Proposition 4.4** (Labourie and Wentworth [20]) *For $D_{A^s}$ a smooth one-parameter family of flat connections, we have a unique smooth function $\lambda_\gamma(s)$ such that, for*

*$s$ small enough, $\lambda_\gamma(s)$ is the eigenvalue of the holonomy of $D_{A^s}$ with $\lambda_\gamma(0)=\lambda_\gamma$. Moreover,*

$$\frac{d\log\lambda_\gamma(s)}{ds}\bigg|_{s=0} = -\int_0^{l_\gamma} \mathrm{Tr}(\partial_s D_{A^0}(t)\cdot\pi(t))\,dt. \tag{4-2}$$

*Here the notation is $\partial_s D_{A^0}(t) := \partial_s D_{A^s}(\dot\gamma(t)\,\partial/\partial t)(0)$, where $\partial_s D_{A^0}$ is an $\mathrm{End}(E)$–valued $1$–form and $\dot\gamma(t)\,\partial/\partial t$ is the tangent vector field along $\gamma(t)$.*

**Proof** We prove (4-2) here.

Let $\{g_s\}$ be a family of gauge transformations acting on $\{D_{A^s}\}$ with $g_0=\mathrm{id}$. Define the new connection 1–forms $\widetilde{A}^s := g_s^*A^s$. We first prove

$$\int_0^{l_\gamma} \mathrm{Tr}(\partial_s D_{A^0}(t)\cdot\pi(t))\,dt = \int_0^{l_\gamma} \mathrm{Tr}(\partial_s D_{\widetilde{A}^0}(t)\cdot\pi(t))\,dt.$$

Note here $\partial_s D_{A^0}(t)$ is a 0–form since we have contracted the 1–form $\partial_s D_{A^s}(0)$ with the tangential vector field. Therefore, $\mathrm{Tr}(\partial_s D_{A^0}(t)\cdot\pi(t))$ is a function in $t$ or in $\dot\gamma(t)$.

Taking the derivative of $\widetilde{A}^s := g_s^*A^s$ at $s=0$ yields

$$\partial_s D_{\widetilde{A}^0} = \partial_s D_{A^0} + D_{A^0}\dot g,$$

where $\dot g$, denoting $\partial g_s/\partial_s|_{s=0}$, is a section of $\mathrm{End}(E)$ and the connection $D_{A^0}$ acts on $\dot g$ as $D_{A^0}\dot g = d\dot g + [A^0,\dot g]$.

We want to show

$$\int_0^{l_\gamma} \mathrm{Tr}((D_{A^0}\dot g)\pi)\,dt = 0.$$

To simplify the notation, we will always omit the variable $t$ when writing our formulas. For example, here $(D_{A^0}\dot g)\pi := (D_{A^0}(t)\dot g(t))\pi(t)$.

We start by proving that $\pi$ is a $D_{A^0}$–parallel section in $\mathrm{End}(E)$. Given any section $v\in\Gamma(E)$, we can write it as a linear combination of eigenvectors of holonomy. Set $v(t)=\sum_{i=1}^n a_i(t)e_i(t)$, where $e_i(t)$ satisfies the parallel transport equation $D_{A_0}e_i=0$ with boundary conditions $e_i(l_\gamma)=\lambda_\gamma^i e_i(0)$ and $\|e_i(0)\|=1$. In particular, we assume $\lambda_\gamma^1=\lambda_\gamma$ and $\mathcal{L}_\gamma(t)$ is generated by $e_1(t)$. Then

$$\begin{aligned}(D_{A^0}\pi)(v) &= [D_{A^0},\pi]v = D_{A^0}(\pi v) - \pi(D_{A^0}v)\\ &= D_{A^0}(a_1(t)e_1(t)) - \pi\biggl(\sum_{i=1}^n (da_i(t)e_i(t) + a_i(t)D_{A^0}e_i(t))\biggr)\\ &= da_1(t)e_1(t) - da_1(t)e_1(t)\\ &= 0.\end{aligned}$$

Thus,

$$\begin{aligned}
\int_0^{l_\gamma} \frac{d}{dt}(\mathrm{Tr}(\dot g\cdot\pi))\,dt
&= \int_0^{l_\gamma} \mathrm{Tr}\Big(\frac{\partial}{\partial t}(\dot g\pi)\Big)\,dt \\
&= \int_0^{l_\gamma} \mathrm{Tr}(D_{A^0}(\dot g\pi))\,dt && (\text{since } \mathrm{Tr}([A^0,\dot g\pi])=0) \\
&= \int_0^{l_\gamma} \mathrm{Tr}([D_{A^0},\dot g\pi])\,dt && (\text{notice } \dot g\pi\in\Gamma(\mathrm{End}(E))) \\
&= \int_0^{l_\gamma} \mathrm{Tr}([D_{A^0},\dot g]\pi+\dot g[D_{A^0},\pi])\,dt \\
&= \int_0^{l_\gamma} \mathrm{Tr}((D_{A^0}\dot g)\pi+\dot g(D_{A^0}\pi))\,dt && (\text{action of a connection on } \Gamma(\mathrm{End}(E))) \\
&= \int_0^{l_\gamma} \mathrm{Tr}((D_{A^0}\dot g)\pi)\,dt && (\text{since } D_{A^0}\pi=0).
\end{aligned}$$

So

$$\begin{aligned}
\int_0^{l_\gamma} \mathrm{Tr}((D_{A^0}\dot g)\pi)\,dt &= \int_0^{l_\gamma} \frac{d}{dt}(\mathrm{Tr}(\dot g\cdot\pi))\,dt \\
&= \mathrm{Tr}(\dot g(l_\gamma)\pi(l_\gamma)) - \mathrm{Tr}(\dot g(0)\pi(0)) \\
&= 0.
\end{aligned}$$

As $s$ varies, the eigenline $\mathcal{L}^s_\gamma(t)$ corresponding to $\lambda_\gamma(s)$ varies according to s and so does the complementary hyperplane $\mathcal{H}^s_\gamma(t)$. By picking suitable gauges $\{g_s\}$, we can assume, for $\widetilde{A}^s := g_s^* A^s$, the eigenlines $\widetilde{\mathcal{L}}^s_\gamma(t)$ and complementary hyperplanes $\widetilde{\mathcal{H}}^s_\gamma(t)$ satisfy $\widetilde{\mathcal{L}}^s_\gamma(t)=\mathcal{L}_\gamma(t)$ and $\widetilde{\mathcal{H}}^s_\gamma(t)=\mathcal{H}_\gamma(t)$.

Without loss of generality, we assume $D_{A^s}$ is itself the connection for a suitable gauge and $\{e_i^s\}$ are eigenvectors for $A^s$ with $e_1^s$ corresponding to $\mathcal{L}^s_\gamma$. Thus,

$$D_{A^s} e_i^s(t)=0, \quad e_i^s(l_\gamma)=\lambda^i_\gamma(s)e_i^s(0).$$

In particular, we can assume

$$D_{A^s} e_1^s(t)=0, \quad e_1^s(t)=c_s(t)e_1^0(t), \quad e_1^s(l_\gamma)=\lambda^1_\gamma(s)e_1^s(0), \quad e_1^s(0)=e_1^0(0).$$

So

$$e_1^s(l_\gamma)=c_s(l_\gamma)e_1^0(l_\gamma)=c_s(l_\gamma)\lambda^1_\gamma(0)e_1^0(0)=\lambda^1_\gamma(s)e_1^s(0)=\lambda^1_\gamma(s)e_1^0(0)$$

and thus $c_s(l_\gamma) = \lambda^1_\gamma(s)/\lambda^1_\gamma(0)$ and $c_0(l_\gamma) = 1$. Notice

$$\frac{H(e_1^0(t), D_{A^s} e_1^0(t))}{H(e_1^0(t), e_1^0(t))} = \frac{H\big(e_1^0(t), D_{A^s}(e_1^s(t)/c_s(t))\big)}{H(e_1^0(t), e_1^s(t)/c_s(t))} = \frac{\partial_t(1/c_s(t))}{1/c_s(t)} = -\frac{\partial(\log c_s(t))}{\partial t}.$$

So

$$\begin{aligned}\int_0^{l_\gamma} \mathrm{Tr}(\partial_s D_{A^0}\pi)\,dt &= \int_0^{l_\gamma} \frac{H(e_1^0(t), \partial_s D_{A^0} e_1^0(t))}{H(e_1^0(t), e_1^0(t))}\,dt \\ &= -\int_0^{l_\gamma} \frac{\partial}{\partial s}\left(\frac{\partial(\log c_s(t))}{\partial t}\right)\bigg|_{s=0} dt \\ &= -\frac{d \log \lambda^1_\gamma(s)}{ds}\bigg|_{s=0}. \end{aligned}$$

□

# 5 Computation of the variations of the reparametrization functions for a model case

In this section and the next, we consider the model case $\partial_\beta g_{\alpha\alpha}(\sigma)$. Note the treatment of this case will involve all the steps needed for the other cases. This justifies the expositional decision that we consider it here first and in isolation.

In this case, we are given parameters $(u, v) \in \{(-1, 1)\}^2$ with (conjugacies classes of) representations $\{\rho(u, v)\}$ in $\mathcal{H}_3(S)$ corresponding to

$$\{(0, uq_\alpha + vq_\beta)\} \subset H^0(X, K^2) \oplus H^0(X, K^3)$$

by Hitchin parametrization (see Remark 2.37). In particular, at the Fuchsian point $\rho(0) = X$, we identify $\partial_u\rho(0,0)$ with $\varphi(q_\alpha)$ and $\partial_v\rho(0,0)$ with $\varphi(q_\beta)$, where $\varphi$ is the Hitchin deformation given in Definition 2.38. We suppose $\{f_{\rho(u,v)}\}$ is an associated two-parameter family of reparametrization functions. By Proposition 4.2, the formula for $\partial_\beta g_{\alpha\alpha}(\sigma)$ is

$$\begin{aligned}\partial_\beta g_{\alpha\alpha}(\sigma) &= \partial_v\big(\langle \partial_u\rho(0,v), \partial_u\rho(0,v)\rangle_P\big)(0) \\ &= \lim_{r\to\infty} \frac{1}{r}\bigg[\int_{UX}\left(\int_0^r \partial_u f^N_{\rho(0)}\,dt\right)^2 \int_0^r \partial_v f^N_{\rho(0)}\,dt\,dm_0 \\ &\qquad\qquad + 2\int_{UX}\int_0^r \partial_u f^N_{\rho(0)}\,dt \int_0^r \partial_{uv} f^N_{\rho(0)}\,dt\,dm_0\bigg].\end{aligned}$$

Because $\partial_u h(\rho(u,0)) = \partial_v h(\rho(0,v)) = 0$ on Fuchsian locus $\mathcal{T}(S)$. By Theorem 2.16, the variations of the reparametrization functions that need to be computed are the following:

(i) $\partial_u f^N_{\rho(0)} = -\partial_u f_{\rho(0)}$.

(ii) $\partial_v f^N_{\rho(0)} = -\partial_v f_{\rho(0)}$.

(iii) $\partial_{uv} f^N_{\rho(0)} = -\partial_{uv} h(\rho(0)) - \partial_{uv} f_{\rho(0)}$.

Before proceeding to compute (i), (ii) and (iii), we explain our general strategy to compute the variations of the reparametrization functions. Our computation will be based on Proposition 4.4 and tools from Higgs bundles theory. Let us first set up our Higgs bundles.

In the component $\mathcal{H}_3(S)$ we are considering, the rank-3 holomorphic vector bundle is fixed as $E = K \oplus \mathcal{O} \oplus K^{-1}$. Associated to a representation $\rho$ in $\mathcal{H}_3(S)$ is a Hermitian metric $H$ on $E$ that solves Hitchin's equation (2-9) and a flat connection $D_H = \nabla_{\bar\partial_E, H} + \Phi + \Phi^{*H}$, where $\nabla_{\bar\partial_E, H}$ is the Chern connection (see Theorem 2.29).

Given a parameter $s \in (-1, 1)$, suppose we are considering a family of conjugacy classes of representations $\{\rho_s\}$ in $\mathcal{H}_3(S)$. On the one hand, there is a family of flat connections $\{D_{H(s)}\}$ given by (2-10) associated to $\{\rho_s\}$. On the other hand, there is a family of reparametrization functions $\{f_{\rho_s}\}_{s\in(-1,1)}$ associated to $\{\rho_s\}$ from the thermodynamical point of view. Recall our notation (2-4)–(2-5). For a family of flat connections $\{D_{H(s)}\}$, we write

$$\partial_s D_{H(0)} = \frac{\partial D_{H(s)}}{\partial s}\bigg|_{s=0}$$

and, for a family of reparametrization functions $\{f_{\rho_s}\}$,

$$\partial_s f_{\rho_0} = \frac{\partial f_{\rho_s}}{\partial s}\bigg|_{s=0}.$$

By Proposition 4.4 and Livšic's theorem, the Hölder function $-\mathrm{Tr}(\partial_s D_{H(0)}\pi)(x)$ and $\partial_s f_{\rho_0}(x)$ are in the same Livšic cohomology class. Recalling our notation in Definition 2.3,

$$\partial_s f_{\rho_0}(x) \sim -\mathrm{Tr}(\partial_s D_{H(0)}\pi)(x). \tag{5-1}$$

Here we define $\mathrm{Tr}(\partial_s D_{H(0)}\pi)(\Phi_t(x)) := \mathrm{Tr}(\partial_s D_{H(0)}(t)\pi(t))$, following Proposition 4.4. The curve $\gamma(t)$ in Proposition 4.4 from now on will be a unit-speed geodesic starting from $x$. Therefore, $x = \dot\gamma(0)\,\partial/\partial t$ and $\Phi_t(x) = \dot\gamma(t)\,\partial/\partial t$.

Proposition 4.3 allows us to consider the first and second variations of the reparametrization functions in terms of Livšic cohomology classes instead of individual functions.

From now on, for the first and second variations of the reparametrization functions, we will no longer distinguish cohomologous elements.

Because $X$ is a hyperbolic surface and the geodesic flow is Anosov, the vectors tangent to periodic geodesics are dense in $TX$. To recover the information of $\partial_s f_{\rho_0}$, it suffices to compute $\mathrm{Tr}(\partial_s D_{H(0)}\pi)$ on each closed geodesic. Similarly, to compute the second variations of the reparametrization functions, it suffices to compute them on each closed geodesic.

Now we start to give a complete computation of the first and second variations of the reparametrization functions for the case $\partial_\beta g_{\alpha\alpha}(\sigma)$. The steps of our argument are divided into different subsections:

(1) We set up coordinates adapted to the closed geodesics we study and conclude special properties of affine metrics with respect to chosen coordinates on these geodesics.

(2) We first construct a homogeneous ODE arising from the parallel transport equation for the base flat connection at $\rho(0) = \sigma \in \mathcal{T}(S)$. This leads to formulas for the first variations of the reparametrization functions proved in [20].

(3) We consider a family of parallel transport equations associated to a family of flat connections by solving Hitchin's equations based at $\rho(0) = \sigma \in \mathcal{T}(S)$. The variation of this family of parallel transport equations at $\sigma$ gives rise to some nonhomogeneous ODEs and yields solutions for the second variations of the reparametrization functions on the closed geodesics we consider.

(4) We extend our computation from the closed geodesics to the surface.

## 5.1 Setting up coordinates on surfaces

In this subsection, we set up coordinates adapted to the closed geodesics we study. We will obtain some important properties for the affine metric after setting up the coordinates. They can be used in the computation of the first and second variations of the reparametrization functions in the following sections. The first variations have been computed in [20] by advanced Lie-theoretic methods.

The convention we use for a Hermitian metric $H$ on $E$ is it is $\mathbb{C}$–linear in the second variable and conjugate-linear in the first variable. Suppose on a coordinate chart $(U, z)$, the bundle $E = K \oplus \mathcal{O} \oplus K^{-1}$ is trivialized as $E|_U \cong U \times \mathbb{C}^3$. Locally we have a holomorphic frame $(s_1, s_2, s_3)$ on $U$. With respect to the local holomorphic frame and our convention of the Hermitian metric, the $(1,0)$–part of the Chern connection $\nabla_{\bar\partial_E,H}$ is $H^{-1}\,\partial H$. The Hermitian conjugate is $\Phi^{*H} = H^{-1}\bar{\Phi}^t H$. The connection 1–form $A$ of the flat connection $D_H$ is thus

$$A = H^{-1}\,\partial H + \Phi + \Phi^{*H}.$$

Associated to representations $\{\rho(u,v)\}$ are a two-parameter family of flat connections $\{D_{H(u,v)}\}$. We will study their connection 1–forms in holomorphic frames with respect to some carefully chosen coordinates on the surface $X$.

When the Higgs field is

$$\Phi(u,v) = \begin{bmatrix} 0 & 0 & uq_\alpha + vq_\beta \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

Baraglia proves the Hermitian metric $H(u,v)$ that solves Hitchin's equation (2-9) is diagonal (see [2]). Following Baraglia's notation [2], we denote the Hermitian metric by $H(u,v) = e^{2\Omega(u,v)}$. We have

$$H(u,v) = \begin{bmatrix} h(u,v)^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & h(u,v) \end{bmatrix},$$

where $h = h(u,v)$ is a section of $\bar{K} \otimes K$ and

$$\Omega(u,v) = \begin{bmatrix} -\omega(u,v) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \omega(u,v) \end{bmatrix}$$

with $\omega(u,v) = \frac{1}{2}\log h(u,v)$.

We denote the corresponding flat connection by

$$D_{H(u,v)} = \nabla_{\bar\partial_E, H(u,v)} + \Phi(u,v) + \Phi(u,v)^{*H(u,v)}.$$

The connection 1–form $A(u,v) \in \Gamma(T^*X \otimes \operatorname{End} E)$ is thus

$$A(u,v) = \begin{bmatrix} -2\,\partial\omega(u,v) & h(u,v) & uq_\alpha + vq_\beta \\ 1 & 0 & h(u,v) \\ h^{-2}(u\bar{q}_\alpha + v\bar{q}_\beta) & 1 & 2\,\partial\omega(u,v) \end{bmatrix}. \tag{5-2}$$

In fact, $2h(u,v)$ is an affine metric for some hyperbolic affine sphere in the conformal class of $\sigma$ (see [2]).

We let $\phi = \log(2h/\sigma)$. Note $\phi = \phi(u,v,z)$ is actually a globally well-defined function on $X$ that does not depend on coordinate systems. Hitchin's equation (2-9), using the integrability condition for affine sphere (see [24]), can be written as an equation of $\phi$ as

$$\Delta_\sigma \phi + 16\|uq_\alpha + vq_\beta\|_\sigma^2 e^{-2\phi} - 2e^\phi + 2 = 0, \tag{5-3}$$

where $\|\cdot\|_\sigma$ is the induced norm on cubic differentials. It satisfies $\|q\|_\sigma^2 = |q|^2/\sigma^3$. The notation we adopt for Laplacian is $\Delta_\sigma = 4\partial_{\bar z}\partial_z/\sigma$.

For simplicity of notation, we sometimes omit variables and write $\phi$ as $\phi(u,v)$ or $\phi(z)$ depending on our needs.

We have the following observation from (5-3):

- When $(u,v) = (0,0)$, the only solution of (5-3) is $\phi = \phi(0,0) = 0$. The affine metric $2h = \sigma$ is indeed the hyperbolic metric of constant curvature $-1$.
- Taking the $u$–derivative or $v$–derivative of (5-3) at $(u,v) = (0,0)$ yields

$$\Delta_\sigma \phi_u - 2e^\phi \phi_u = 0, \tag{5-4}$$

$$\Delta_\sigma \phi_v - 2e^\phi \phi_v = 0. \tag{5-5}$$

  Therefore, the fact that $\phi = \phi(0,0) = 0$ implies $\phi_u = \phi_u(0,0) = 0$ and $\phi_v = \phi_v(0,0) = 0$.

We now choose a special coordinate system that facilitates the study of holonomy problems on a closed geodesic. Let $z$ be a local holomorphic coordinate on $X$. Suppose the affine metric in this coordinate is $e^{\psi(u,v,z)}|dz|^2$ and the hyperbolic metric in this coordinate is $\sigma = e^{\delta(z)}|dz|^2$. Suppose $\gamma(t)$ is any closed geodesic with respect to the hyperbolic metric $\sigma$ on the Riemann surface $X$. Then, written in the $z$–coordinate, it is

$$\gamma(t) = z(t) = \operatorname{Re}\gamma(t) + i \operatorname{Im}\gamma(t)$$

and

$$\dot\gamma(t)\frac{\partial}{\partial t} = (\operatorname{Re}\dot\gamma(t) + i\operatorname{Im}\dot\gamma(t))\frac{\partial}{\partial z} + (\operatorname{Re}\dot\gamma(t) - i\operatorname{Im}\dot\gamma(t))\frac{\partial}{\partial \bar z}.$$

In particular, we can model $\gamma(t)$ on a strip $S = \{x + iy : |y| < \frac{\pi}{2}\}$ with the hyperbolic metric $ds = |dz|/\cos y$ and $\gamma(t) = (t,0)$. This coordinate around $\gamma$ is called a Fermi coordinate and satisfies $\operatorname{Re}\dot\gamma(t) = 1$ and $\operatorname{Im}\dot\gamma(t) = 0$. Thus, it's easy to check that, on $\gamma$, one has $\gamma^* ds = |dz|$ and $\delta(z) = 0$.

The variable $t$ is then the arc-length parameter for our choice of coordinates. Therefore, if one writes $\dot\gamma(0)\,\partial/\partial t = x \in UX$, then $\dot\gamma(t)\,\partial/\partial t = \Phi_t(x)$. We will always assume $\dot\gamma(0)\,\partial/\partial t = x$ in our discussion.

With the Fermi coordinate understood, from the fact that the only solution of (5-3) is $\phi = 0$, we conclude

$$\psi(z) = \phi(z) + \delta(z) = \delta(z) = 0.$$

From (5-4) together with (5-5) and their solutions $\phi_u = \phi_v = 0$, we obtain

$$\psi_u(z) = \phi_u(z) = 0, \quad \psi_v(z) = \phi_v(z) = 0.$$

Also $\psi(z) = 0$ implies

$$\psi_z(z) = \delta_z(z) = 0.$$

All this information about the affine metric $\psi$ with respect to the Fermi coordinate will be important in computation in later sections.

## 5.2 Homogeneous ODEs for holonomy and first variations of the reparametrization functions

In this subsection, we show a formula for the first variations of the reparametrization functions from [20]. We also construct homogeneous ODEs arising from the parallel transport equations for the base flat connection at $\sigma \in \mathcal{T}(S)$. These serve as the first step for the computation of the second variations in later subsections.

We first explain our notation. For $q_i = q_i(z)\,dz^2$ any quadratic differential and $q_\alpha = q_\alpha(z)\,dz^3$ any cubic differential, we also use $q_i$ and $q_\alpha$ to denote Hölder functions on the unit tangent bundle $UX$ as follows. We let $q_i : UX \to \mathbb{C}$ and $q_\alpha : UX \to \mathbb{C}$ be

$$q_i(x) := q_i(z)\,dz^2(x,x) = q_i(z)(dz(x))^2, \tag{5-6}$$

$$q_\alpha(x) := q_\alpha(z)\,dz^3(x,x,x) = q_\alpha(z)(dz(x))^3. \tag{5-7}$$

The first variations of the reparametrization functions for our cases have been computed in [20] as follows:

**Proposition 5.1** [20, Theorem 4.0.2] *The first variations of the reparametrization functions $\partial_u f_{\rho(0)} : UX \to \mathbb{R}$ and $\partial_v f_{\rho(0)} : UX \to \mathbb{R}$ for our model case $\partial_\beta g_{\alpha\alpha}(\sigma)$ satisfy*

$$-\partial_u f_{\rho(0)}(x) \sim \operatorname{Re} q_\alpha(x), \quad -\partial_v f_{\rho(0)}(x) \sim \operatorname{Re} q_\beta(x),$$

*where the notation $\sim$ is Livšic equivalence* (Definition 2.3).

Proposition 5.1 is proved in [20] as a consequence of (5-1).

We then study parallel transport equations for the connection $D_{H(0)}$ arising from holonomy problems based at $\rho(0) \in \mathcal{T}(S)$. With the coordinates introduced in the last section, they become homogenous ODE systems that are easy to solve. We list some important computations involved here. These will be important for the second variations of the reparametrization functions.

The parallel transport equation for the connection $D_{H(0)}$ on the closed geodesic $\gamma$ is

$$D_{H(0),\dot{\gamma}} V = 0, \tag{5-8}$$

where $V \in \Gamma(E)$ is a parallel section with boundary conditions

$$V(l_\gamma) = \lambda_i(\gamma, \rho(0)) V(0).$$

Here $\lambda_i(\gamma, \rho(0))$ is one of the eigenvalues for holonomy of $D_{H(0)}$ on $\gamma$ for $i = 1, 2, 3$. We want to write (5-8) on a specific holomorphic frame, which can be constructed as follows.

We cover $\gamma$ by $m$ charts $\{(U_i, z_i)\}_{i=1}^m$ such that $z_i : U_i \to z_i(U_i) \subset \mathbb{C}$ is a diffeomorphism for $1 \le i \le m$. We assume our holomorphic bundle $E$ is trivialized on each $U_i$. Furthermore, we assume the transition map on every overlap is either the identity or a hyperbolic translation viewed on the universal cover $\mathbb{D}$. Since $dz_i$ is a local holomorphic section of $K$ on $U_i$ and $\partial/\partial z_i$ is a local holomorphic section of $K^{-1}$ on $U_i$, we can define a local holomorphic frame $s^i = (s_1^i, s_2^i, s_3^i)$ for $E = K \oplus \mathcal{O} \oplus K^{-1}$ on $U_i$, where $s_1^i = dz_i$ and $s_2^i = 1$ and $s_3^i = \partial/\partial z_i$. Setting $(U_{m+1}, z_{m+1}) = (U_1, z_1)$ and $s_j^{m+1} = s_j^1$, this yields a well-defined holomorphic frame for $\gamma$ because, on each overlap and for $j = 1, 2, 3$, we have $s_j^i = s_j^{i+1}$ on $\gamma|_{U_i} \cap \gamma|_{U_i+1}$ with $1 \le i \le m$.

We will simply write the holomorphic frame on $\gamma$ as $s_j$ for $j = 1, 2, 3$. With respect to this frame, the parallel transport equation for $V(t) = \sum_{i=1}^3 V_i(t) s_i(t)$ becomes

$$\partial_t \begin{bmatrix} V^1(t) \\ V^2(t) \\ V^3(t) \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} V^1(t) \\ V^2(t) \\ V^3(t) \end{bmatrix} = 0.$$

There are three eigenvalues for this ODE system: $\lambda_1(\gamma, \rho(0)) = e^{l_\gamma}$, $\lambda_2(\gamma, \rho(0)) = 1$ and $\lambda_3(\gamma, \rho(0)) = e^{-l_\gamma}$. The solutions for $V$ (assuming norm 1 at the starting point with respect to the Hermitian metric $H(0)$), denoted by $e_i$ corresponding to $\lambda_i(\gamma)$ for $i = 1, 2, 3$, are

$$e_1 = \frac{\sqrt{2}}{2} e^t \begin{bmatrix} \frac{1}{2} \\ -1 \\ 1 \end{bmatrix}, \quad e_2 = \frac{1}{2} \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \quad e_3 = \frac{\sqrt{2}}{2} e^{-t} \begin{bmatrix} \frac{1}{2} \\ 1 \\ 1 \end{bmatrix}.$$

We note at the Fuchsian point $\rho(0) \in \mathcal{T}(S)$, the eigenvectors $e_1$, $e_2$ and $e_3$ are orthogonal. In our holomorphic frame, the projection $\pi(0) = \pi(\rho(0))$ can be computed as

$$\pi(0) = \frac{1}{2}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{4} \\ -1 & 1 & -\frac{1}{2} \\ 1 & -1 & \frac{1}{2} \end{bmatrix}.$$

The eigenvectors $e_i$ and projection $\pi$ will play important roles in later sections.

## 5.3 Inhomogeneous ODEs and the second variations of the reparametrization functions

We will compute the second variation of the reparametrization functions $\partial_{uv} f_{\rho(0)}$ in this and the next subsection. With our formula (5-1), we have

$$\begin{aligned} \partial_{uv} f_{\rho(0)} &\sim -\partial_v\big(\mathrm{Tr}(\partial_u D_{H(0,v)}\pi(0,v))\big)(0) \\ &= -\mathrm{Tr}\bigg(\frac{\partial^2 D_{H(0)}}{\partial u\,\partial v}\pi(0)\bigg) - \mathrm{Tr}(\partial_u D_{H(0)}\partial_v \pi(0)) \\ &=: -\mathrm{I} - \mathrm{II}. \end{aligned} \tag{5-9}$$

In this subsection, we compute $\partial_{uv} f_{\rho(0)}$ along a closed geodesic by computing I and II. We study variation of holonomy problems along a closed geodesic and construct associated inhomogeneous ODEs. In the next subsection, we extend the computation of $\partial_{uv} f_{\rho(0)}$ to the whole surface.

**Compute I** With the holomorphic frames and Fermi coordinates setup as before, one obtains, on $\gamma$,

$$\partial_{uv} D_{H(0)}(x) = \begin{bmatrix} -(\psi_z)_{uv}(z) & \frac{1}{2}\psi_{uv}(z) & 0 \\ 0 & 0 & \frac{1}{2}\psi_{uv}(z) \\ 0 & 0 & (\psi_z)_{uv}(z) \end{bmatrix}.$$

Thus,

$$\mathrm{Tr}\bigg(\frac{\partial^2 D_{H(0)}}{\partial u\,\partial v}(x)\pi(0)\bigg) = -\tfrac{1}{2}\psi_{uv}(z).$$

More explicitly, $\mathrm{Tr}((\partial^2 D_{H(0)}/\partial u\,\partial v)\pi(0))\colon UX \to \mathbb{R}$ satisfies

$$\mathrm{Tr}\bigg(\frac{\partial^2 D_{H(0)}}{\partial u\,\partial v}\pi(0)\bigg)(x) = -\tfrac{1}{2}\psi_{uv}\big(z(p(x))\big) = -\tfrac{1}{2}\phi_{uv}(p(x)),$$

where $p\colon UX \to X$ is the projection from the unit tangent bundle to our surface and $z$ is the Fermi coordinate we choose evaluating at the point $p(x) \in X$. Note that the affine metric $\psi$ is always real and $\phi = \psi - \sigma$ does not depend on the coordinates we choose.

**Compute II** To study $\partial_v \pi(0)$ takes some effort. We set $u = 0$ and take a family of flat connections $\{D_{H(v)}\}$ with connection 1–forms $A(0, v)$ (recall (5-2)). Associated to each of them is a parallel transport equation along the closed geodesic $\gamma$ on $(S, \sigma)$,

$$D_{H(v),\dot{\gamma}} V(v, t) = 0, \tag{5-10}$$

with the assumption $\|V(v, 0)\|_{H(0)} = 1$.

In [19], Labourie proves the images of every Hitchin representation are purely loxodromic. For $\rho(0, v)$ in $\mathcal{H}_3(S)$, we know $\rho(0, v)(\gamma)$ has distinct eigenvalues $\lambda_1(\gamma, v) > \lambda_2(\gamma, v) > \lambda_3(\gamma, v)$. The holonomy problem for $\rho(0, v)$ has three distinct eigenvectors which are parallel sections $\{e_i(v, t)\}_{i=1}^3$ along $\gamma(t)$. Each section $V(v, t) = e_i(v, t)$ satisfies (5-10). In addition to the norm 1 condition at the starting point, $\|V(v, 0)\|_{H(0)} = 1$, we also impose another boundary condition in order to guarantee these are eigenvectors. The boundary conditions are, for $i = 1, 2, 3$,

(i) $\|e_i(v, 0)\|_{H(0)} = 1$;

(ii) $e_i(v, l_\gamma) = \lambda_i(\gamma, v) e_i(v, 0)$.

The reader may notice that, up to now, there are two frames for $E$ along $\gamma$ mentioned, the holomorphic frame $(s_1, s_2, s_3)$ and the frame spanned by eigenvectors $(e_1, e_2, e_3)$. On the one hand, we can write our holomorphic frames as linear combinations of eigenvectors $s_i(t) = \sum_{j=1}^3 a_{ij}(v, t) e_j(v, t)$ for $i = 1, 2, 3$. On the other hand, we can write the eigenvectors as linear combinations of our holomorphic frames $e_j(v, t) = \sum_{k=1}^3 e_{jk}(v, t) s_k(t)$ for $j = 1, 2, 3$. We have the following observation:

With respect to the holomorphic frame $(s_1, s_2, s_3)$, the projection onto $e_1$ along the hyperplane spanned by $(e_2, e_3)$ in matrix form is

$$\begin{aligned}
\pi(v, t) &= \begin{bmatrix} \pi(v, t) s_1(t) & \pi(v, t) s_2(t) & \pi(v, t) s_3(t) \end{bmatrix} \\
&= \begin{bmatrix} a_{11}(v, t) e_1(v, t) & a_{21}(v, t) e_1(v, t) & a_{31}(v, t) e_1(v, t) \end{bmatrix} \\
&= \begin{bmatrix} a_{11}(v, t) e_{11}(v, t) & a_{21}(v, t) e_{11}(v, t) & a_{31}(v, t) e_{11}(v, t) \\ a_{11}(v, t) e_{12}(v, t) & a_{21}(v, t) e_{12}(v, t) & a_{31}(v, t) e_{12}(v, t) \\ a_{11}(v, t) e_{13}(v, t) & a_{21}(v, t) e_{13}(v, t) & a_{31}(v, t) e_{13}(v, t) \end{bmatrix}.
\end{aligned}$$

To understand $\partial_v \pi(0)$, we need to know $\partial_v e_1(0)$ and $\partial_v \alpha_{i1}(0)$ for $i = 1, 2, 3$. One can check, in the holomorphic frame,

$$\begin{aligned}
\operatorname{Tr}(\partial_u D_{A(0)} \partial_v \pi(0)) = {} & q_\alpha (\partial_v a_{11}(0) e_{13}(0) + a_{11}(0) \partial_v e_{13}(0)) \\
& + 4\bar{q}_\alpha (\partial_v a_{31}(0) e_{11}(0) + a_{31}(0) \partial_v e_{11}(0)),
\end{aligned} \tag{5-11}$$

where $e_{11}(0)$ and $e_{13}(0)$ are known. Thus, we need to compute $\partial_v e_1(0)$ and $\partial_v a_{11}(0)$ and $\partial_v a_{31}(0)$.

We first show how to obtain $\partial_v e_1(0,t)$ as the solution of an inhomogeneous ODE system arising from taking the $v$–derivative for a family of parallel transport equations (5-10) at $v=0$,

$$\partial_t \begin{bmatrix} \partial_v e_{11}(0,t) \\ \partial_v e_{12}(0,t) \\ \partial_v e_{13}(0,t) \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \partial_v e_{11}(0,t) \\ \partial_v e_{12}(0,t) \\ \partial_v e_{13}(0,t) \end{bmatrix} = -\frac{\sqrt{2}}{2} e^t \begin{bmatrix} q_\beta(\Phi_t(x)) \\ 0 \\ 2\bar{q}_\beta(\Phi_t(x)) \end{bmatrix},$$

with boundary conditions

$$\begin{aligned} &H(\partial_v e_1(0,0), e_1(0,0)) = 0, \\ &\qquad \partial_v e_1(0,l_\gamma) = -e^{l_\gamma}\left(\int_0^{l_\gamma} \operatorname{Re} q_\beta(\Phi_s(x))\,ds\right) e_1(0,0) + e^{l_\gamma}\partial_v e_1(0,0). \end{aligned}$$

The boundary conditions arise from taking the $v$–derivative for boundary conditions (i) and (ii) of the parallel transport equation (5-10) that the maximum eigenvector $e_1$ satisfies.

With these boundary conditions, we solve

$$\begin{aligned} \begin{bmatrix} \partial_v e_{11}(t) \\ \partial_v e_{12}(t) \\ \partial_v e_{13}(t) \end{bmatrix} &= \begin{bmatrix} -\frac{\sqrt{2}}{2}\int_0^t e^s(\cosh(t-s)\operatorname{Re} q_\beta + i \operatorname{Im} q_\beta)\,ds \\ \sqrt{2}\int_0^t e^s \sinh(t-s)\operatorname{Re} q_\beta\,ds \\ -\sqrt{2}\int_0^t e^s(\cosh(t-s)\operatorname{Re} q_\beta - i \operatorname{Im} q_\beta)\,ds \end{bmatrix} \\ &\quad + \begin{bmatrix} -\frac{\sqrt{2}}{4}(e^{2l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{2s-t}\operatorname{Re} q_\beta\,ds - \frac{\sqrt{2}}{2} i (e^{l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^s \operatorname{Im} q_\beta\,ds \\ -\frac{\sqrt{2}}{2}(e^{2l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{2s-t}\operatorname{Re} q_\beta\,ds \\ -\frac{\sqrt{2}}{2}(e^{2l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{2s-t}\operatorname{Re} q_\beta\,ds + \sqrt{2} i (e^{l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^s \operatorname{Im} q_\beta\,ds \end{bmatrix}. \end{aligned}$$

Here $q_\beta$ refers to $q_\beta(\Phi_s(x))$ defined in (5-7).

We continue to compute $\partial_v \alpha_{11}(0)$ and $\partial_v \alpha_{31}(0)$. Combining

$$e_j(v,t) = \sum_{k=1}^{3} e_{jk}(v,t) s_k(t) \quad \text{and} \quad s_i(t) = \sum_{j=1}^{3} a_{ij}(v,t) e_j(v,t)$$

gives

$$\sum_{j=1}^{j=3} a_{ij}(v,t) e_{jk}(v,t) = \sigma_{ik}. \tag{5-12}$$

Recall the $e_{jk}(0,t)$ are known:

$$e_1(0,t)=\begin{bmatrix} e_{11}(0,t)\\ e_{12}(0,t)\\ e_{13}(0,t)\end{bmatrix}=\tfrac{\sqrt{2}}{2}e^t\begin{bmatrix}\frac{1}{2}\\ -1\\ 1\end{bmatrix},$$

$$e_2(0,t)=\begin{bmatrix} e_{21}(0,t)\\ e_{22}(0,t)\\ e_{23}(0,t)\end{bmatrix}=\tfrac{1}{2}\begin{bmatrix}-1\\ 0\\ 2\end{bmatrix},$$

$$e_3(0,t)=\begin{bmatrix} e_{31}(0,t)\\ e_{32}(0,t)\\ e_{33}(0,t)\end{bmatrix}=\tfrac{\sqrt{2}}{2}e^{-t}\begin{bmatrix}\frac{1}{2}\\ 1\\ 1\end{bmatrix}.$$

Then one obtains

$$a(0,t)=\begin{bmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}=\begin{bmatrix} \frac{\sqrt{2}}{2}e^{-t} & -1 & \frac{\sqrt{2}}{2}e^{t}\\ -\frac{\sqrt{2}}{2}e^{-t} & 0 & \frac{\sqrt{2}}{2}e^{t}\\ \frac{\sqrt{2}}{4}e^{-t} & \frac{1}{2} & \frac{\sqrt{2}}{4}e^{t}\end{bmatrix}.$$

Taking the $v$–derivative of (5-12) at $v=0$,

$$\sum_{j=1}^{j=3}\partial_v a_{ij}(0,t)e_{jk}(0,t)+\sum_{j=1}^{j=3}a_{ij}(0,t)\partial_v e_{jk}(0,t)=0.$$

Solutions of $\partial_v a_{ij}(0,t)$ can be expressed in terms of $\partial_v e_1(0,t)$, $\partial_v e_2(0,t)$ and $\partial_v e_3(0,t)$. We have just solved $\partial_v e_1$. Similarly,$\partial_v e_2(0)$ and $\partial_v e_3(0)$ are solutions of another two systems of nonhomogeneous ODEs deduced from (5-10). We now proceed to solve $\partial_v e_2(0,t)$ and $\partial_v e_3(0,t)$.

(1) For $\partial_v e_2(0,t)$, we have

$$\partial_t\begin{bmatrix}\partial_v e_{21}(0,t)\\ \partial_v e_{22}(0,t)\\ \partial_v e_{23}(0,t)\end{bmatrix}+\begin{bmatrix}0 & \frac{1}{2} & 0\\ 1 & 0 & \frac{1}{2}\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}\partial_v e_{21}(0,t)\\ \partial_v e_{32}(0,t)\\ \partial_v e_{23}(0,t)\end{bmatrix}=\begin{bmatrix}-q_\beta(\Phi_t(x))\\ 0\\ 2\bar{q}_\beta(\Phi_t(x))\end{bmatrix}$$

with boundary conditions

$$H(\partial_v e_2(0,0),e_2(0,0))=0,$$

$$\partial_v e_2(0,l_\gamma)=2\biggl(\int_0^{l_\gamma}\operatorname{Re}q_\beta(\Phi_s(x))\,ds\biggr)e_2(0,0)+\partial_v e_2(0,0).$$

(2) For $\partial_v e_3(0,t)$, we get

$$\partial_t\begin{bmatrix}\partial_v e_{31}(0,t)\\ \partial_v e_{32}(0,t)\\ \partial_v e_{33}(0,t)\end{bmatrix}+\begin{bmatrix}0 & \frac{1}{2} & 0\\ 1 & 0 & \frac{1}{2}\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}\partial_v e_{31}(0,t)\\ \partial_v e_{32}(0,t)\\ \partial_v e_{33}(0,t)\end{bmatrix}=-\tfrac{\sqrt{2}}{2}e^{-t}\begin{bmatrix}q_\beta(\Phi_t(x))\\ 0\\ 2\bar{q}_\beta(\Phi_t(x))\end{bmatrix}$$

with boundary conditions

$$
\begin{aligned}
&H(\partial_v e_3(0,0), e_3(0,0)) = 0,\\
&\qquad\partial_v e_3(0,l_\gamma) = -e^{-l_\gamma}\left(\int_0^{l_\gamma} \operatorname{Re} q_\beta(\Phi_s(x))\,ds\right) e_3(0,0) + e^{-l_\gamma}\partial_v e_3(0,0).
\end{aligned}
$$

We obtain respective solutions from

$$
\begin{bmatrix}\partial_v e_{21}(t)\\ \partial_v e_{22}(t)\\ \partial_v e_{23}(t)\end{bmatrix} = \begin{bmatrix} -\int_0^t \operatorname{Re} q_\beta + i\cosh(t-s)\operatorname{Im} q_\beta\,ds\\ 2\int_0^t i\sinh(t-s)\operatorname{Im} q_\beta\,ds\\ 2\int_0^t \operatorname{Re} q_\beta - i\cosh(t-s)\operatorname{Im} q_\beta\,ds\end{bmatrix}
+ \begin{bmatrix} -\frac{i}{2}\int_0^{l_\gamma}\operatorname{Im} q_\beta((1-e^{l_\gamma})^{-1}e^{l_\gamma+t-s} + (1-e^{-l_\gamma})^{-1}e^{-l_\gamma-t+s})\,ds\\ i\int_0^{l_\gamma}\operatorname{Im} q_\beta((1-e^{l_\gamma})^{-1}e^{l_\gamma+t-s} - (1-e^{-l_\gamma})^{-1}e^{-l_\gamma-t+s})\,ds\\ -i\int_0^{l_\gamma}\operatorname{Im} q_\beta((1-e^{l_\gamma})^{-1}e^{l_\gamma+t-s} + (1-e^{-l_\gamma})^{-1}e^{-l_\gamma-t+s})\,ds\end{bmatrix}
$$

and

$$
\begin{bmatrix}\partial_v e_{31}(t)\\ \partial_v e_{32}(t)\\ \partial_v e_{33}(t)\end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{2}}{2}\int_0^t e^{-s}(\cosh(t-s)\operatorname{Re} q_\beta + i\operatorname{Im} q_\beta)\,ds\\ \sqrt{2}\int_0^t e^{-s}\sinh(t-s)\operatorname{Re} q_\beta\,ds\\ -\sqrt{2}\int_0^t e^{-s}(\cosh(t-s)\operatorname{Re} q_\beta - i\operatorname{Im} q_\beta)\,ds\end{bmatrix}
+ \begin{bmatrix} -\frac{\sqrt{2}}{4}(e^{-2l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{t-2s}\operatorname{Re} q_\beta\,ds - \frac{\sqrt{2}}{2} i(e^{-l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{-s}\operatorname{Im} q_\beta\,ds\\ \frac{\sqrt{2}}{2}(e^{-2l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{t-2s}\operatorname{Re} q_\beta\,ds\\ -\frac{\sqrt{2}}{2}(e^{-2l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{t-2s}\operatorname{Re} q_\beta\,ds + \sqrt{2} i(e^{-l_\gamma}-1)^{-1}\int_0^{l_\gamma} e^{-s}\operatorname{Im} q_\beta\,ds\end{bmatrix},
$$

where $q_\beta$ in the solutions again refers to $q_\beta(\Phi_s(x))$ defined in (5-7).

We are therefore able to solve $\partial_v a_{ij}(0,t)$ from $\partial_v e_1(0,t)$, $\partial_v e_2(0,t)$ and $\partial_v e_3(0,t)$. For a closed geodesic $\gamma$ of length $l_\gamma$ starting from $\dot\gamma(0) = x$, we compute, from (5-11),

$$
\begin{aligned}
\text{(5-13)}\quad &\operatorname{Tr}(\partial_u D_{A(0)}\partial_v\pi(0))(\Phi_t(x))\\
&= \operatorname{Re} q_\alpha(\Phi_t(x))\int_0^t (e^{2(t-s)} - e^{2(s-t)})\operatorname{Re} q_\beta(\Phi_s(x))\,ds\\
&\quad + 2\operatorname{Im} q_\alpha(\Phi_t(x))\int_0^t (e^{t-s} - e^{s-t})\operatorname{Im} q_\beta(\Phi_s(x))\,ds\\
&\quad + \operatorname{Re} q_\alpha(\Phi_t(x))\int_0^{l_\gamma}\left(\frac{e^{2(t-s)}}{e^{-2l_\gamma}-1} - \frac{e^{2(s-t)}}{e^{2l_\gamma}-1}\right)\operatorname{Re} q_\beta(\Phi_s(x))\,ds\\
&\quad + 2\operatorname{Im} q_\alpha(\Phi_t(x))\int_0^{l_\gamma}\left(\frac{e^{t-s}}{e^{-l_\gamma}-1} - \frac{e^{s-t}}{e^{l_\gamma}-1}\right)\operatorname{Im} q_\beta(\Phi_s(x))\,ds.
\end{aligned}
$$

In particular, at $t = 0$,

$$\text{(5-14)}\quad \begin{aligned}\operatorname{Tr}(\partial_u D_{A(0)} \partial_v \pi(0))(x) &= \operatorname{Re} q_\alpha(x) \int_0^{l_\gamma} \left( \frac{e^{-2s}}{e^{-2l_\gamma} - 1} - \frac{e^{2s}}{e^{2l_\gamma} - 1} \right) \operatorname{Re} q_\beta(\Phi_s(x))\, ds \\ &\quad + 2 \operatorname{Im} q_\alpha(x) \int_0^{l_\gamma} \left( \frac{e^{-s}}{e^{-l_\gamma} - 1} - \frac{e^{s}}{e^{l_\gamma} - 1} \right) \operatorname{Im} q_\beta(\Phi_s(x))\, ds.\end{aligned}$$

**Remark 5.2** Every point on the closed geodesic $\gamma$ plays an equivalent role. We can always let $y = \Phi_t(x)$ be the initial point of our $\gamma$ and set up boundary conditions for our ODEs based at $y$ instead of $x$. The solution of this new ODE system is (5-14), treating $y = \Phi_t(x)$ as the initial point. It is in fact the same as starting from $x$ and obtaining $\operatorname{Tr}(\partial_u D_{A(0)} \partial_v \pi(0))(\Phi_t(x))$ from (5-13).

## 5.4 Hölder extension to the surface

The holonomy problems only yield solutions on closed geodesics as they can be simplified as linear ODEs with boundary conditions. However, it is still possible to extend the computation for the second variations of the reparametrization functions from closed geodesics to the Riemann surface $X$. This will be our goal in this subsection. In particular, We will prove in the end of this subsection the main proposition about second variations of the reparametrization functions.

**Proposition 5.3** *The second variation of the reparametrization functions*

$$\partial_{uv} f_{\rho(0)} \colon UX \to \mathbb{R}$$

*for our model case $\partial_\beta g_{\alpha\alpha}(\sigma)$ satisfies*

$$\partial_{uv} f_{\rho(0)}(x) \sim \tfrac{1}{2}\phi_{uv}(p(x)) - \eta(x),$$

*where we recall that $\phi$ is defined in* (5-3) *and $p\colon UX \to X$ is the projection from the unit tangent bundle $UX$ to our Riemann surface $X$, and $\eta\colon UX \to \mathbb{R}$ is given by*

$$\begin{aligned}\eta(x) = -\operatorname{Re} q_\alpha(x) \int_0^\infty e^{-2s} \operatorname{Re} q_\beta(\Phi_s(x))\, ds - \operatorname{Re} q_\alpha(x) \int_{-\infty}^0 e^{2s} \operatorname{Re} q_\beta(\Phi_s(x))\, ds \\ -2 \operatorname{Im} q_\alpha(x) \int_0^\infty e^{-s} \operatorname{Im} q_\beta(\Phi_s(x))\, ds - 2 \operatorname{Im} q_\alpha(x) \int_{-\infty}^0 e^{s} \operatorname{Im} q_\beta(\Phi_s(x))\, ds.\end{aligned}$$

We will prove that $\eta(x)$ coincides with $\mathrm{Tr}(\partial_u D_{A(0)}\partial_v \pi(0))(x)$ on periodic orbits and that $\eta(x)$ is a Hölder function. Denoting the subset of $UX$ that consists of all unit tangent vectors to closed geodesics by $W$, we first show:

**Proposition 5.4** *For any $x \in W$, $\eta(x) = \mathrm{Tr}(\partial_u D_{A(0)}\partial_v \pi(0))(x)$.*

To prove Proposition 5.4, from the computation of $\mathrm{Tr}(\partial_u D_{A(0)}\partial_v \pi(0))(x)$ in (5-14), we introduce an intermediate function $\psi\colon W \times \mathbb{R}^+ \to \mathbb{R}$, given by

$$\begin{aligned}\psi(x,r) = \mathrm{Re}\, q_\alpha(x) \int_0^r \left(\frac{e^{-2s}}{e^{-2r}-1} - \frac{e^{2s}}{e^{2r}-1}\right) \mathrm{Re}\, q_\beta(\Phi_s(x))\, ds \\ + 2\,\mathrm{Im}\, q_\alpha(x) \int_0^r \left(\frac{e^{-s}}{e^{-r}-1} - \frac{e^{s}}{e^{r}-1}\right) \mathrm{Im}\, q_\beta(\Phi_s(x))\, ds.\end{aligned}$$

Given $x \in W$, if we denote the closed geodesic that $x$ is tangential to by $\gamma_x$, with length $l_{\gamma_x}$, then clearly $\mathrm{Tr}(\partial_u D_{A(0)}\partial_v \pi(0))(x) = \psi(x, l_{\gamma_x})$. To prove Proposition 5.4 for the set $W$, we need the following lemma, which states that $\psi(x,r)$ attains the same value when $r$ is any positive integer multiple of $l_{\gamma_x}$:

**Lemma 5.5** *$\psi(x, k l_{\gamma_x}) = \psi(x, l_{\gamma_x})$ for all $x \in W$ and $k \in \mathbb{Z}^+$.*

**Proof** For any $k \in \mathbb{Z}^+$, we have

$$\begin{aligned}&\int_0^{k l_{\gamma x}} \left(\frac{e^{-2s}}{e^{-2k l_{\gamma x}}-1} - \frac{e^{2s}}{e^{2k l_{\gamma x}}-1}\right) \mathrm{Re}\, q_\beta(\Phi_s(x))\, ds \\ &= \sum_{i=1}^{k} \int_{(i-1) l_{\gamma x}}^{i l_{\gamma x}} \left(\frac{e^{-2s}}{e^{-2k l_{\gamma x}}-1} - \frac{e^{2s}}{e^{2k l_{\gamma x}}-1}\right) \mathrm{Re}\, q_\beta(\Phi_s(x))\, ds \\ &= \frac{1}{e^{-2k l_{\gamma x}}-1} \sum_{i=1}^{k} \int_{(i-1) l_{\gamma x}}^{i l_{\gamma x}} e^{-2s}\, \mathrm{Re}\, q_\beta(\Phi_s(x))\, ds \\ &\qquad - \frac{1}{e^{2k l_{\gamma x}}-1} \sum_{i=1}^{k} \int_{(i-1) l_{\gamma x}}^{i l_{\gamma x}} e^{2s}\, \mathrm{Re}\, q_\beta(\Phi_s(x))\, ds \\ &= \int_0^{l_{\gamma x}} \left(\frac{e^{-2s}}{e^{-2 l_{\gamma x}}-1} - \frac{e^{2s}}{e^{2 l_{\gamma x}}-1}\right) \mathrm{Re}\, q_\beta(\Phi_s(x))\, ds.\end{aligned}$$

Similar arguments hold for $\int_0^{l_{\gamma x}} (e^{-s}/(e^{-l_{\gamma x}}-1) - e^{s}/(e^{l_{\gamma x}}-1))\, \mathrm{Im}\, q_\beta(\Phi_s(x))\, ds$. Thus, we obtain $\psi(x, k l_{\gamma_x}) = \psi(x, l_{\gamma_x})$. □

**Remark 5.6** This equality is clear if one understands that $\psi(x, k l_\gamma)$ is the solution of the holonomy problem that goes around our closed geodesic $\gamma$ $k$ times with the same boundary conditions.

**Proof of Proposition 5.4** Instead of flowing from $x$ to $\Phi_{l_{\gamma_x}}(x)$, we view $x$ as our midpoint and consider our flow from $\Phi_{-l_{\gamma_x}/2}(x)$ to $x$ and then from $x$ to $\Phi_{l_{\gamma_x}/2}(x)$. From this point of view, we can write $\psi(x, l_{\gamma_x})$ as

$$\begin{aligned}\psi(x,l_{\gamma_x}) = \operatorname{Re} q_\alpha(x) \int_0^{l_{\gamma_x}/2} &\left(\frac{e^{-2s}}{e^{-2l_{\gamma_x}}-1} - \frac{e^{2s}}{e^{2l_{\gamma_x}}-1}\right) \operatorname{Re} q_\beta(\Phi_s(x))\, ds \\ &+ \operatorname{Re} q_\alpha(x) \int_{-l_{\gamma_x}/2}^{0} \left(\frac{e^{2s}}{e^{-2l_{\gamma_x}}-1} - \frac{e^{-2s}}{e^{2l_{\gamma_x}}-1}\right) \operatorname{Re} q_\beta(\Phi_s(x))\, ds \\ &+ 2\operatorname{Im} q_\alpha(x) \int_0^{l_{\gamma_x}/2} \left(\frac{e^{-s}}{e^{-l_{\gamma_x}}-1} - \frac{e^{s}}{e^{l_{\gamma_x}}-1}\right) \operatorname{Im} q_\beta(\Phi_s(x))\, ds \\ &+ 2\operatorname{Im} q_\alpha(x) \int_{-l_{\gamma_x}/2}^{0} \left(\frac{e^{s}}{e^{-l_{\gamma_x}}-1} - \frac{e^{-s}}{e^{l_{\gamma_x}}-1}\right) \operatorname{Im} q_\beta(\Phi_s(x))\, ds.\end{aligned}$$

The above also holds if we replace $l_{\gamma_x}$ by $kl_{\gamma_x}$. We will now conclude by taking $k \to \infty$ in the above formula.

Suppose $\max_{x \in UX}\{|\operatorname{Re} q_\alpha(x)|, |\operatorname{Im} q_\alpha(x)|, |\operatorname{Re} q_\beta(x)|, |\operatorname{Im} q_\beta(x)|\} = M$. Then notice

$$\begin{aligned}|\psi(x,kl_{\gamma_x}) - \eta(x)| \le 2M^2 \int_{kl_{\gamma_x}/2}^{\infty} e^{-2s}\, ds + 4M^2 \int_{kl_{\gamma_x}/2}^{\infty} e^{-s}\, ds \\ + 2M^2 \int_0^{kl_{\gamma_x}/2} \left|\frac{e^{-2s}}{e^{-2kl_{\gamma_x}}-1} - \frac{e^{2s}}{e^{2kl_{\gamma_x}}-1} + e^{-2s}\right| ds \\ + 4M^2 \int_0^{kl_{\gamma_x}/2} \left|\frac{e^{-s}}{e^{-kl_{\gamma_x}}-1} - \frac{e^{s}}{e^{kl_{\gamma_x}}-1} + e^{-s}\right| s \\ \to 0 \quad \text{when } k \to \infty.\end{aligned}$$

Thus, by Lemma 5.5, we obtain, for any $x \in W$,

$$\operatorname{Tr}(\partial_u D_{A(0)} \partial_v \pi(0))(x) = \psi(x, l_{\gamma_x}) = \lim_{k \to \infty} \psi(x, kl_{\gamma_x}) = \eta(x). \qquad \square$$

We also need the following proposition about regularity of the function $\eta$:

**Proposition 5.7** $\eta(x)\colon UX \to \mathbb{R}$ *is a Hölder function.*

**Proof** We start by showing $\int_0^\infty e^{-s} \operatorname{Im} q_\beta(\Phi_s(x))\, ds$ is Hölder. Let $x$ and $y$ be close, with $d(x,y) = \epsilon \ll 1$. It is classical for a hyperbolic surface $(S,\sigma)$ that we have standard ODE estimates on the geodesic flow

$$d(\Phi_s(x), \Phi_s(y)) \le Ne^s d(x,y) = \epsilon N e^s,$$

where $N > 0$ is some constant and the distance function $d$ on $UX$ is induced from the canonical (Sasaki) metric $\langle \cdot, \cdot \rangle$ on $UX$.

Consider $T = -\log(\epsilon)$. Then, dividing the integral into two parts, from 0 to $T$ and from $T$ to $\infty$, yields

$$\begin{aligned}
&\left| \int_0^\infty e^{-s} \operatorname{Im} q_\beta(\Phi_s(x))\, ds - \int_0^\infty e^{-s} \operatorname{Im} q_\beta(\Phi_s(y))\, ds \right| \\
&\quad = \left| \int_0^T e^{-s} \big(\operatorname{Im} q_\beta(\Phi_s(x)) - \operatorname{Im} q_\beta(\Phi_s(y))\big)\, ds \right| \\
&\qquad\qquad + \left| \int_T^\infty e^{-s} \big(\operatorname{Im} q_\beta(\Phi_s(x)) - \operatorname{Im} q_\beta(\Phi_s(y))\big)\, ds \right| \\
&\quad \le \int_0^T e^{-s} N_1 N \epsilon e^s\, ds + 2N_2 e^{-T} \\
&\quad \le -N_1 N \epsilon \log(\epsilon) + 2N_2 \epsilon \\
&\quad \le (N_1 N + 2N_2) d(x,y)^{1/2}.
\end{aligned}$$

Here we use the fact that $\operatorname{Im} q_\beta$ is smooth, so we can assume its Lipschitz constant to be $N_1$. We also use that $UX$ is compact and we assume $\sup_{x \in UX} \operatorname{Im} q_\beta(x) = N_2$.

It then follows easily that $\operatorname{Im} q_\alpha(x) \int_0^\infty e^{-2s} \operatorname{Im} q_\beta(\Phi_s(x))\, ds$ is also a Hölder function. The arguments to show that the other three terms in $\eta(x)$ are Hölder are the same. We therefore conclude that $\eta(x)$ is a Hölder function. □

Finally, with Propositions 5.4 and 5.7, we are able to prove Proposition 5.3 about the second variations of the reparametrization functions on the Riemann surface $X$.

**Proof of Proposition 5.3** We have most of the necessary elements for this proof in previous estimates. We assemble everything here. Because $\operatorname{Tr}(\partial_u D_{A(0)} \partial_v \pi(0))(\Phi_t(x))$ is a Hölder function and it equals the Hölder function $\eta(x)$ on a dense subset of $UX$, we conclude it coincides with $\eta(x)$ everywhere on $UX$. We obtain

$$\begin{aligned}
\partial_{uv} f_{\rho(0)} &\sim -\partial_u \big(\operatorname{Tr}(\partial_v D_{H(0)} \pi(0))\big) \\
&= -\operatorname{Tr}\left( \frac{\partial^2 D_{H(0)}}{\partial u\, \partial v} \pi(0) \right) - \operatorname{Tr}(\partial_v D_{H(0)} \partial_u \pi(0)) \\
&= \tfrac{1}{2} \phi_{uv}(p(x)) - \eta(x),
\end{aligned}$$

where we recall here $\phi = \log(2h/\sigma)$ is a globally well-defined function defined in (5-3) evaluating at the point $p(x) \in X$ and $p \colon UX \to X$ is the projection from the unit tangent bundle to our surface. □

## 6 Evaluation on the Poincaré disk for the model case

After the computation of the first and second variations of the reparametrization functions on $UX$ in the last two sections, we are able to evaluate $\partial_\beta g_{\alpha\alpha}(\sigma)$. Our goal in this section is to show the following:

**Proposition 6.1** *For $\sigma \in \mathcal{T}(S)$, $\partial_\beta g_{\alpha\alpha}(\sigma) = 0$.*

Let's first write down the expression for $\partial_\beta g_{\alpha\alpha}(\sigma)$,

$$\begin{aligned}
&\partial_\beta g_{\alpha\alpha}(\sigma)\\
&= \partial_v\big(\langle \partial_u \rho(0,v), \partial_u \rho(0,v)\rangle_P\big)(0)\\
&= \lim_{r\to\infty} \frac{1}{r}\biggl[\int_{UX}\biggl(\int_0^r \partial_u f^N_{\rho(0)}\,dt\biggr)^2 \int_0^r \partial_v f^N_{\rho(0)}\,dt\,dm_0\\
&\qquad\qquad\qquad\qquad + 2\int_{UX}\int_0^r \partial_u f^N_{\rho(0)}\,dt \int_0^r \partial_{uv} f^N_{\rho(0)}\,dt\,dm_0\biggr]\\
&= \lim_{r\to\infty}\frac{1}{r}\int_{UX}\biggl(\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\biggr)^2 \int_0^r \operatorname{Re} q_\beta(\Phi_t(x))\,dt\,dm_0\\
&\quad + \lim_{r\to\infty}\frac{1}{r}\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt \int_0^r -\partial_{uv} h(\rho(0)) - \partial_{uv} f_{\rho(0)}(\Phi_t(x))\,dt\,dm_0\\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}$$

The formula for $\partial_{uv} f_{\rho(0)}$ is given in Proposition 5.3.

We aim to prove both I and II are zero. The following lemma will be crucial:

**Lemma 6.2** *For any $t, s \in \mathbb{R}$, we have*

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\,dm_0(x) = 0, \tag{6-1}$$

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Im} q_\alpha(\Phi_t(x)) \operatorname{Im} q_\beta(\Phi_s(x))\,dm_0(x) = 0. \tag{6-2}$$

We use the methods in [20] to show the integrals are zero. Similarly to the proof of Theorem 6.3.1 in [20], the key is to use the symmetry properties of the Liouville measure $m_0 = m_L$ and homogeneity of holomorphic differentials viewed as functions on $UX$. We transfer the problem of evaluating the integrals in (6-1) and (6-2) to analyzing the Fourier coefficients of holomorphic differentials.

Before we start our proof, we first explain the coordinates we will use to do the computation following [20]. We take the Poincaré disk as our charts. Pick a point $x \in UX$. We identify the universal cover of $(X, \sigma)$ with $\mathbb{D}$ by the unique isometry that takes $\pi(x) \in X$ to $0 \in \mathbb{D}$ and identify the vector $x \in UX$ with the vector $(1,0) \in T_0\mathbb{D}$.

We express our holomorphic differentials in these coordinates. The holomorphic cubic differential $q_\alpha$ has the analytic expansion in the coordinate based on $x$,

$$q_{\alpha,x}(z) = \sum_{n=1}^{\infty} a_n(x) z^n \, dz^3.$$

Recall the hyperbolic distance $d_H$ in the Poincaré disk model satisfies

$$d_H(0, Re^{i\theta}) = r(R) = \tfrac{1}{2}\log\Big(\frac{1+R}{1-R}\Big).$$

Thus, $\partial/\partial r = (1-R^2)\,\partial/\partial R$ and

$$dz\Big(\frac{\partial}{\partial r}\Big)\Big|_{Re^{i\theta}} = (1-R^2)e^{i\theta}.$$

Denoting $\tilde{q}_{\alpha,x}(z) := \operatorname{Re}(q_{\alpha,x}(z)(\partial/\partial r, \partial/\partial r, \partial/\partial r))$, one has

$$\operatorname{Re} q_\alpha(\Phi_r(e^{i\theta}x)) = \tilde{q}_{\alpha,x}(Re^{i\theta}) = \operatorname{Re}\Big(\sum_{n=0}^{\infty} a_n(x) R^n (1-R^2)^3 e^{i(n+3)\theta}\Big). \tag{6-3}$$

In particular, when $r = 0$,

$$\lim_{R\to 0} dz\Big(\frac{\partial}{\partial r}\Big)\Big|_{Re^{i\theta}} = e^{i\theta}.$$

Therefore,

$$\begin{aligned}\operatorname{Re} q_\alpha(e^{i\theta}x) &= \tilde{q}_{\alpha,x}(0\cdot e^{i\theta}) \\ &= \lim_{R\to 0} \operatorname{Re}\Big(q_{\alpha,x}(Re^{i\theta})\Big(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}, \frac{\partial}{\partial r}\Big)\Big) \\ &= \operatorname{Re}(a_0(x)e^{i3\theta}).\end{aligned} \tag{6-4}$$

Suppose the coefficients of the analytic expansion for $q_\beta$ are $b_n$; then

$$\operatorname{Re} q_\beta(\Phi_r(e^{i\theta}x)) = \tilde{q}_{\beta,x}(Re^{i\theta}) = \operatorname{Re}\Big(\sum_{n=0}^{\infty} b_n(x) R^n (1-R^2)^3 e^{i(n+3)\theta}\Big). \tag{6-5}$$

For the convenience of computation later for other cases, we also write down here two analytic expansions for holomorphic quadratic differentials $q_i$ and $q_j$, with coefficients

$c_n$ and $d_n$, respectively,

$$\operatorname{Re} q_i(\Phi_r(e^{i\theta}x)) = \tilde{q}_{i,x}(Re^{i\theta}) = \operatorname{Re}\biggl(\sum_{n=0}^{\infty} c_n(x) R^n (1-R^2)^2 e^{i(n+2)\theta}\biggr). \tag{6-6}$$

$$\operatorname{Re} q_j(\Phi_r(e^{i\theta}x)) = \tilde{q}_{j,x}(Re^{i\theta}) = \operatorname{Re}\biggl(\sum_{n=0}^{\infty} d_n(x) R^n (1-R^2)^2 e^{i(n+2)\theta}\biggr). \tag{6-7}$$

**Proof of Lemma 6.2** We begin with showing (6-1).

The proof of it will be divided into two cases:

(1) $t \geq 0$ and $s \geq 0$.

(2) $t < 0$ or $s < 0$.

In the first case, we work with the analytic expansions (6-3) and (6-5). We choose two special situations: $s = t$ and $s = \frac{1}{2}t$. We observe some symmetries in these two situations and argue from these symmetries that (6-1) holds for the first case. We then apply the results for the first case to the second case by flow-invariance properties of $m_L$. Equation (6-2) then follows easily from (6-1) once we find the relation between them.

Since $m_0 = m_L$ is rotationally invariant, ie $(e^{i\theta})^* m_L = m_L$, we have

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\, dm_0(x) \\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_{UX} \operatorname{Re} q_\alpha(e^{i\theta}x) \operatorname{Re} q_\alpha(\Phi_t(e^{i\theta}x)) \operatorname{Re} q_\beta(\Phi_s(e^{i\theta}x))\, dm_0(x)\, d\theta.\end{aligned}$$

(1) We restrict ourselves to the case $t, s \geq 0$ of (6-1) so that we can work with the analytic expansions (6-3) and (6-5).

We let $t(T) = \frac{1}{2}\log((1+T)/(1-T))$ and $s(S) = \frac{1}{2}\log((1+S)/(1-S))$. We first consider $t > 0$ and $s > 0$. Then, if we first integrate over the $\theta$–variable, in terms of the analytic expansion, we get

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\, dm_0(x) \\ &= \frac{1}{4}\sum_{n=0}^{\infty}\biggl(\int_{UX} \operatorname{Re}(a_0 a_n \bar{b}_{n+3})\, dm_0\, T^n(1-T^2)^3 S^{n+3}(1-S^2)^3 \\ &\qquad + \int_{UX} \operatorname{Re}(a_0 \bar{a}_{n+3} b_n)\, dm_0\, T^{n+3}(1-T^2)^3 S^n(1-S^2)^3\biggr).\end{aligned} \tag{6-8}$$

We let $A_n = \int_{UX} \operatorname{Re}(a_0 a_n \bar{b}_{n+3})\, dm_0$ and $B_n = \int_{UX} \operatorname{Re}(a_0 \bar{a}_{n+3} b_n)\, dm$. To show (6-1) holds for $t, s \geq 0$, it suffices to prove, for $n \geq 0$,

$$A_n = B_n = 0. \tag{6-9}$$

If $t = 0$ or $s = 0$, equation (6-1) is equivalent to

$$A_0 = B_0 = 0,$$

which are included in (6-9). To prove (6-9), we consider two special cases of (6-1): flow times $s = t$ and $s = \frac{1}{2}t$.

By the $\Phi_t$–invariance of $m_0$, flow time $s = t$ satisfies

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_t(x))\, dm_0(x) = \int_{UX} \operatorname{Re} q_\alpha(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(x)\, dm_0(x).$$

A convenient observation is that flowing from $x$ backwards for time $t$ is the opposite of flowing forwards for time $t$ from $-x$, ie $\Phi_{-t}(x) = -\Phi_t(-x)$. Let $y = -x$ and notice $(e^{i\pi})^* m_0 = m_0$, so we have

$$\int_{UX} \operatorname{Re} q_\alpha(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(x)\, dm_0(x) = -\int_{UX} \operatorname{Re} q_\alpha(\Phi_t(y)) \operatorname{Re} q_\alpha(y) \operatorname{Re} q_\beta(y)\, dm_0(y).$$

Therefore,

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_t(x))\, dm_0(x) = -\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(x)\, dm_0(x).$$

This implies

$$\sum_{n=0}^{\infty} (A_n + B_n) T^{2n+3} (1 - T^2)^6 = -B_0 T^3 (1 - T^2)^3.$$

The coefficient of $T^0$ yields

$$A_0 + 2B_0 = 0. \tag{6-10}$$

Similarly, for flow time $s = \frac{1}{2}t$, we let $y = -x$ and again use the fact $(e^{i\pi})^* m_0 = m_0$:

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_{t/2}(x))\, dm_0(x) \\ &= \int_{UX} \operatorname{Re} q_\alpha(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_{-t/2}(x))\, dm_0(x) \\ &= -\int_{UX} \operatorname{Re} q_\alpha(\Phi_t(-x)) \operatorname{Re} q_\alpha(-x) \operatorname{Re} q_\beta(\Phi_{t/2}(-x))\, dm_0(x) \\ &= -\int_{UX} \operatorname{Re} q_\alpha(\Phi_t(y)) \operatorname{Re} q_\alpha(y) \operatorname{Re} q_\beta(\Phi_{t/2}(y))\, dm_0(y).\end{aligned}$$

Thus, $\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_{t/2}(x))\, dm_0(x) = 0$.

Recall $t(T) = \frac{1}{2}\log((1+T)/(1-T))$ and $s = \frac{1}{2}\log((1+S)/(1-S))$. In the case $s = \frac{1}{2}t$, we have $T = 2S/(S^2+1)$. The analytic expansion for

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_{t/2}(x))\, dm_0(x) = 0$$

with condition $T = 2S/(S^2+1)$ simplifies to

$$\sum_{n=0}^{\infty} (A_n(S^2+1)^3 + 8B_n)\left(\frac{2S^2}{S^2+1}\right)^n = 0.$$

Let $W = S^2/(S^2+1)$ with $0 < W < \frac{1}{2}$. Then the above is equivalent to

$$\sum_{n=0}^{\infty}\left(A_n \sum_{k=0}^{\infty} \tfrac{1}{2}(k+1)(k+2)W^k + 8B_n\right)2^n W^n = 0.$$

This give relations

$$2^{n+3}B_n + \sum_{k=0}^{n}(n-k+1)(n-k+2)2^{k-1}A_k = 0, \quad n \geq 0.$$

When $n = 0$, combining with (6-10), we obtain $A_0 = B_0 = 0$. Then (6-10) yields $A_n + B_n = 0$ for all $n \in \mathbb{N}$. This fact, combined with the above formula, gives $A_n = B_n = 0$ and (6-1) holds for $t, s \geq 0$.

(2) For $t < 0$ or $s < 0$, there are three cases we need to discuss.

- If $t \leq s$ and $t < 0$, then, as $m_0$ is $\Phi_t$–invariant,

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\, dm_0(x) \\ &= \int_{UX} \operatorname{Re} q_\alpha(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_{s-t}(x))\, dm_0(x).\end{aligned}$$

This is the same as the $s, t \geq 0$ case.

- If $s < t \le 0$, then

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\, dm_0(x)\\ &= \int_{UX} \operatorname{Re} q_\alpha(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_{s-t}(x))\, dm_0(x)\\ &= -\int_{UX} \operatorname{Re} q_\alpha(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_{t-s}(-x))\, dm_0(x)\\ &= 0.\end{aligned}$$

This is from the observation that the analytic expansion of $\operatorname{Re} q_\beta(\Phi_r(-e^{i\theta}x))$ based at $x$ for $r > 0$ is

$$\begin{aligned}\operatorname{Re} q_\beta(\Phi_r(-e^{i\theta}x)) = \operatorname{Re} q_\beta(\Phi_r(e^{i(\theta+\pi)}x)) &= \tilde{q}_{\beta,x}(Re^{i(\theta+\pi)})\\ &= \operatorname{Re}\biggl(\sum_{n=0}^{\infty} b_n(x) R^n (1-R^2)^3 e^{i(n+3)(\theta+\pi)}\biggr)\end{aligned}$$

and that, for $n \ge 0$,

$$e^{-i(n+6)\pi} \int_{UX} \operatorname{Re}(a_0 a_n \bar{b}_{n+3})\, dm_0 = 0, \quad e^{i(n+3)\pi} \int_{UX} \operatorname{Re}(a_0 \bar{a}_{n+3} b_n)\, dm_0 = 0.$$

- If $s < 0 \le t$, then we consider

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\, dm_0(x)\\ &= \int_{UX} \operatorname{Re} q_\alpha(\Phi_t(-x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_{t-s}(-x))\, dm_0(x)\\ &= 0.\end{aligned}$$

The argument is essentially the same as the other cases. This finishes the proof of (6-1).

Equation (6-2) follows easily from (6-1) since, for all $t, s \in \mathbb{R}$,

$$\begin{aligned}\operatorname{Re}&\biggl(\int_{UX} \operatorname{Re} q_\alpha(x) q_\alpha(\Phi_t(x)) q_\beta(\Phi_s(x))\, dm_0(x)\biggr)\\ &= \int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\beta(\Phi_s(x))\, dm_0(x)\\ &\qquad - \int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Im} q_\alpha(\Phi_t(x)) \operatorname{Im} q_\beta(\Phi_s(x))\, dm_0(x)\end{aligned}$$

and

$$\int_{UX} \operatorname{Re} q_\alpha(x) q_\alpha(\Phi_t(x)) q_\beta(\Phi_s(x))\, dm_0(x) = 0. \tag{6-11}$$

This is easy to see from the fact that $\int_0^{2\pi} \operatorname{Re}(a_0(x)e^{i3\theta})e^{i(n+3)\theta}e^{i(m+3)\theta}\,d\theta = 0$ for all $n, m \geq 0$ and thus, for $t, s > 0$,

$$\begin{aligned}
&\int_{UX} \operatorname{Re} q_\alpha(x) q_\alpha(\Phi_t(x)) q_\beta(\Phi_s(x))\,dm_0(x) \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_{UX} \operatorname{Re} q_\alpha(e^{i\theta}x) q_\alpha(\Phi_t(e^{i\theta}x)) q_\beta(\Phi_s(e^{i\theta}x))\,dm_0(x)\,d\theta \\
&= \frac{1}{2\pi}\sum_{m,n\geq 0}\int_{UX}\int_0^{2\pi} \operatorname{Re}(a_0(x)e^{i3\theta})(a_n(x)T^n(1-T^2)^3e^{i(n+3)\theta}) \\
&\qquad\qquad\qquad\qquad \cdot (b_m(x)S^{m+3}(1-S^2)^3e^{i(m+3)\theta})\,d\theta\,dm_0(x) \\
&= 0.
\end{aligned}$$

The argument for $t \leq 0$ or $s \leq 0$ can be transferred back to the $t > 0$ and $s > 0$ cases. One needs the observation that $-\Phi_{-t}(-x) = \Phi_t(x)$ and $-e^{i\theta}x = e^{i(\theta+\pi)}x$. We conclude (6-11) holds for all $t, s \in \mathbb{R}$ and thus (6-2) holds. □

**Proof of Proposition 6.1** We start to show $\mathrm{I} = \mathrm{II} = 0$.

$\mathrm{I} = 0$ reduces to (6-1) of Lemma 6.2 if we take $r \to \infty$ in

$$\begin{aligned}
&\frac{1}{r}\int_{UX}\left(\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\right)^2 \int_0^r \operatorname{Re} q_\beta(\Phi_t(x))\,dt\,dm_0 \\
&= \frac{1}{r}\int_0^r\int_0^r\int_0^r\int_{UX} \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x)) \operatorname{Re} q_\beta(\Phi_\mu(x))\,dm_0\,d\mu\,dt\,ds \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Fubini's theorem)} \\
&= \frac{1}{r}\int_0^r\int_0^r\int_0^r\int_{UX} \operatorname{Re} q_\alpha(\Phi_{t-s}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_{\mu-s}(x))\,dm_0\,d\mu\,dt\,ds \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(since } m_0 \text{ is } \Phi_t\text{–invariant)} \\
&= 0.
\end{aligned}$$

We next look into II:

$$\begin{aligned}
\mathrm{II} &= \lim_{r\to\infty}\frac{1}{r}\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt \int_0^r -\partial_{uv}h(\rho(0)) - \partial_{uv} f_{\rho(0)}(\Phi_t(x))\,dt\,dm_0 \\
&= -\lim_{r\to\infty}\frac{1}{r}\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt \int_0^r \partial_{uv}h(\rho(0))\,dt\,dm_0 \\
&\quad - \lim_{r\to\infty}\frac{1}{r}\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt \int_0^r \operatorname{Tr}\left(\frac{\partial^2 D_{A(0)}}{\partial u\,\partial v}\pi(0)\right)(\Phi_t(x))\,dt\,dm_0 \\
&\quad + \lim_{r\to\infty}\frac{1}{r}\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Tr}(\partial_v D_{A(0)}\partial_u\pi(0))(\Phi_t(x))\,dt\,dm_0.
\end{aligned}$$

There are three terms here. Since $\partial_{uv}h(\rho(0))$ is a constant, the first term is

$$\lim_{r\to\infty}\frac{1}{r}\int_{UX}2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\int_0^r \partial_{uv}h(\rho(0))\,dt\,dm_0 = \lim_{r\to\infty}2\partial_{uv}h(\rho(0))\int_{UX}\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\,dm_0.$$

Recall our expressions given by (6-3) and (6-6). Then

$$\begin{aligned}
\int_{UX}\int_0^r &\operatorname{Re} q_\alpha(\Phi_t(x))\,dt\,dm_0\\
&=\int_0^r\int_{UX}\operatorname{Re} q_\alpha(\Phi_t(x))\,dm_0\,dt\\
&=\int_0^r\int_{UX}\operatorname{Re} q_\alpha(x)\,dm_0\,dt && \text{(since } m_0 \text{ is } \Phi_t\text{–invariant)}\\
&=\frac{1}{2\pi}\int_0^r\int_{UX}\int_0^{2\pi}\operatorname{Re} q_\alpha(e^{i\theta}x)\,d\theta\,dm_0\,dt && \text{(since } m_0 \text{ is rotationally invariant)}\\
&=\frac{r}{2\pi}\int_{UX}\int_0^{2\pi}\operatorname{Re}(a_0(x)e^{i3\theta})\,d\theta\,dm_0\\
&=0.
\end{aligned}$$

The second term in II is

$$-\lim_{r\to\infty}\frac{1}{r}\int_{UX}2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\int_0^r \operatorname{Tr}\biggl(\frac{\partial^2 D_{A(0)}}{\partial u\,\partial v}\pi(0)\biggr)(\Phi_t(x))\,dt\,dm_0 = \lim_{r\to\infty}\frac{1}{r}\int_{UX}2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\int_0^r \tfrac{1}{2}\phi_{uv}(\Phi_t(x))\,dt\,dm_0,$$

recalling that $\phi$ is a globally well-defined function on $X$ (see formula (5-3)), and

$$\tfrac{1}{2}\phi_{uv}\bigl(p(\Phi_t(x))\bigr)=\tfrac{1}{2}\phi_{uv}\bigl(p(\Phi_t(e^{i\theta}x))\bigr).$$

So

$$\begin{aligned}
\frac{1}{r}\int_{UX}2\int_0^r &\operatorname{Re} q_\alpha(\Phi_t(x))\,dt\int_0^r \tfrac{1}{2}\phi_{uv}(\Phi_t(x))\,dt\,dm_0\\
&=\frac{1}{r}\int_{UX}\int_0^{2\pi}2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(e^{i\theta}x))\,dt\int_0^r \tfrac{1}{2}\phi_{uv}\bigl(p(\Phi_t(e^{i\theta}x))\bigr)\,dt\,d\theta\,dm_0\\
&=\frac{1}{r}\int_{UX}\int_0^{2\pi}2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(e^{i\theta}x))\,dt\int_0^r \tfrac{1}{2}\phi_{uv}\bigl(p(\Phi_t(x))\bigr)\,dt\,d\theta\,dm_0\\
&=\frac{1}{r}\int_0^r\int_0^r\int_{UX}\phi_{uv}\bigl(p(\Phi_{t-s}(x))\bigr)\int_0^{2\pi}\operatorname{Re} q_\alpha(e^{i\theta}x)\,d\theta\,dm_0\,ds\,dt.
\end{aligned}$$

Again by the fact $\int_0^{2\pi} \operatorname{Re} q_\alpha(e^{i\theta}x)\, d\theta = \int_0^{2\pi} \operatorname{Re}(a_0(x)e^{i3\theta})\, d\theta = 0$, we conclude

$$\lim_{r\to\infty} \frac{1}{r} \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha\, dt \int_0^r \operatorname{Tr}\biggl(\frac{\partial^2 D_{A(0)}}{\partial u\, \partial v}\pi(0)\biggr)\, dt\, dm_0 = 0.$$

It remains to show

$$\lim_{r\to\infty} \frac{1}{r} \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha\, dt \int_0^r \operatorname{Tr}(\partial_v D_{A(0)} \partial_u \pi(0))\, dt\, dm_0 = 0.$$

This is

$$\begin{aligned}
&\lim_{r\to\infty} \frac{1}{r} \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt \int_0^r \eta(\Phi_t(x))\, dt\, dm_0\\
&= -\lim_{r\to\infty} \frac{1}{r}\biggl(\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt\\
&\qquad\qquad \cdot \int_0^r \operatorname{Re} q_\alpha(\Phi_\mu(x)) \int_0^\infty e^{-2s} \operatorname{Re} q_\beta(\Phi_{\mu+s}(x))\, ds\, d\mu\, dm_0\\
&\qquad + \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt\\
&\qquad\qquad \cdot \int_0^r \operatorname{Re} q_\alpha(\Phi_\mu(x)) \int_{-\infty}^0 e^{2s} \operatorname{Re} q_\beta(\Phi_{\mu+s}(x))\, ds\, d\mu\, dm_0\\
&\qquad + \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt\\
&\qquad\qquad \cdot \int_0^r 2\operatorname{Im} q_\alpha(\Phi_\mu(x)) \int_0^\infty e^{-s} \operatorname{Im} q_\beta(\Phi_{\mu+s}(x))\, ds\, d\mu\, dm_0\\
&\qquad + \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt\\
&\qquad\qquad \cdot \int_0^r 2\operatorname{Im} q_\alpha(\Phi_\mu(x)) \int_{-\infty}^0 e^{s} \operatorname{Im} q_\beta(\Phi_{\mu+s}(x))\, ds\, d\mu\, dm_0\biggr).
\end{aligned}$$

We have estimates for these tail terms

$$\begin{aligned}
&\frac{1}{r} \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt \int_0^r \operatorname{Re} q_\alpha(\Phi_\mu(x)) \int_r^\infty e^{-2s} \operatorname{Re} q_\beta(\Phi_{\mu+s}(x))\, ds\, d\mu\, dm_0\\
&\quad + \frac{1}{r} \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt \int_0^r \operatorname{Re} q_\alpha(\Phi_\mu(x)) \int_{-\infty}^{-r} e^{2s} \operatorname{Re} q_\beta(\Phi_{\mu+s}(x))\, ds\, d\mu\, dm_0\\
&\qquad \le \frac{4M^3}{r} r^2 \int_r^\infty e^{-2s}\, ds\\
&\qquad = 2M^3 r e^{-2r} \xrightarrow{r\to\infty} 0.
\end{aligned}$$

The other two tail terms with integrals involving $\operatorname{Im} q_\alpha$ and $\operatorname{Im} q_\beta$ also go to zero for the same reason. So, in fact,

$$\begin{aligned}
&\lim_{r\to\infty}\frac{1}{r}\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha\,dt \int_0^r \operatorname{Tr}(\partial_v D_{A(0)}\partial_u \pi(0))\,dt\,dm_0\\
&= -\lim_{r\to\infty}\frac{1}{r}\biggl(\int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\\
&\qquad\qquad\cdot\int_0^r \operatorname{Re} q_\alpha(\Phi_\mu(x))\int_0^r e^{-2s}\operatorname{Re} q_\beta(\Phi_{\mu+s}(x))\,ds\,d\mu\,dm_0\\
&\qquad + \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\\
&\qquad\qquad\cdot\int_0^r \operatorname{Re} q_\alpha(\Phi_\mu(x))\int_{-r}^0 e^{2s}\operatorname{Re} q_\beta(\Phi_{\mu+s}(x))\,ds\,d\mu\,dm_0\\
&\qquad + \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\\
&\qquad\qquad\cdot\int_0^r 2\operatorname{Im} q_\alpha(\Phi_\mu(x))\int_0^r e^{-s}\operatorname{Im} q_\beta(\Phi_{\mu+s}(x))\,ds\,d\mu\,dm_0\\
&\qquad + \int_{UX} 2\int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\,dt\\
&\qquad\qquad\cdot\int_0^r 2\operatorname{Im} q_\alpha(\Phi_\mu(x))\int_{-r}^0 e^{s}\operatorname{Im} q_\beta(\Phi_{\mu+s}(x))\,ds\,d\mu\,dm_0\biggr).
\end{aligned}$$

Similar to I, the above equaling 0 reduces to (6-2). This finishes our proof of Proposition 6.1 and so concludes the discussion of the model case $\partial_\beta g_{\alpha\alpha}(\sigma)$. □

# 7 The remaining cases

We will show in this section the proofs of the remaining three cases, ie $\partial_i g_{\alpha\alpha}(\sigma)=0$, $\partial_j g_{\alpha i}(\sigma)=0$ and $\partial_\beta g_{\alpha i}(\sigma)=0$. They provide a complete proof of Theorem 1.1.

## 7.1 The case of $\partial_i g_{\alpha\alpha}(\sigma)$

In this case, given parameters $(u,v)\in\{(-1,1)\}^2$, we obtain a family of (conjugacy classes of) representations $\{\rho(u,v)\}$ in $\mathcal{H}_3(S)$ corresponding to $\{(vq_i,uq_\alpha)\}\subset H^0(X,K^2)\oplus H^0(X,K^3)$ by the Hitchin parametrization. In particular, $\partial_u\rho(0,0)$ is identified with $\varphi(q_\alpha)$ and $\partial_v\rho(0,0)$ is identified with $\varphi(q_i)$. The formula for $\partial_i g_{\alpha\alpha}(\sigma)$

is

$$\begin{aligned}\partial_i g_{\alpha\alpha}(\sigma) &= \partial_v\big(\langle \partial_u \rho(0,v), \partial_u \rho(0,v)\rangle_P\big)(0)\\ &= \lim_{r\to\infty}\frac{1}{r}\bigg[\int_{UX}\bigg(\int_0^r \partial_u f^N_{\rho(0)}\,dt\bigg)^2\int_0^r \partial_v f^N_{\rho(0)}\,dt\,dm_0\\ &\qquad\qquad + 2\int_{UX}\int_0^r \partial_u f^N_{\rho(0)}\,dt\int_0^r \partial_{uv} f^N_{\rho(0)}\,dt\,dm_0\bigg],\end{aligned}$$

where the first and second variations are

(i) $\partial_u f^N_{\rho(0)} = -\partial_u f_{\rho(0)}$;

(ii) $\partial_v f^N_{\rho(0)} = -\partial_v f_{\rho(0)}$;

(iii) $\partial_{uv} f^N_{\rho(0)} = -\partial_{uv} h(\rho(0)) - \partial_{vu} f_{\rho(0)}$.

**7.1.1 First and second variations of the reparametrization functions** We compute the first and second variations for the case of $\partial_i g_{\alpha\alpha}(\sigma)$ in this subsection.

We have Higgs field

$$\Phi(u,v) = \begin{bmatrix} 0 & vq_i & uq_\alpha\\ 1 & 0 & vq_i\\ 0 & 1 & 0\end{bmatrix}.$$

Following the steps and methods for our model case $\partial_\beta g_{\alpha\alpha}(\sigma)$ in Section 5, we show in this subsection:

**Proposition 7.1** *The first variations of the reparametrization functions $\partial_u f_{\rho(0)}\colon UX\to\mathbb{R}$ and $\partial_v f_{\rho(0)}\colon UX\to\mathbb{R}$ for the case $\partial_i g_{\alpha\alpha}(\sigma)$ satisfy*

$$\partial_u f_{\rho(0)}(x) \sim -\mathrm{Re}\, q_\alpha(x),\quad \partial_v f_{\rho(0)}(x)\sim 2\,\mathrm{Re}\, q_i(x)$$

*and the second variation of the reparametrization functions $\partial_{vu} f_{\rho(0)}\colon UX\to\mathbb{R}$ for the case $\partial_i g_{\alpha\alpha}(\sigma)$ satisfies*

$$\begin{aligned}\partial_{uv} f_{\rho(0)}(x) &\sim \tfrac12\,\mathrm{Re}\, y_{21}(x)\\ &\qquad - 2\,\mathrm{Im}\, q_\alpha(x)\bigg(\int_0^\infty \mathrm{Im}\, q_i(\Phi_s(x))e^{-s}\,ds + \int_{-\infty}^0 \mathrm{Im}\, q_i(\Phi_s(x))e^{s}\,ds\bigg),\end{aligned}$$

*where $p\colon UX\to X$ is the projection from the unit tangent bundle $UX$ to our Riemann surface $X$. Understanding a section of $\mathrm{End}(E)$ as a linear map on each fiber of $E = K\oplus\mathcal{O}\oplus K^{-1}$ over a point of $X$, the element $y_{21}$ is the component of the section $Y = H^{-1}\partial_{uv}H$ that takes $K$ to $\mathcal{O}$. As a function on $UX$, $y_{21}$ transforms as $y_{21}(e^{i\theta}x) = e^{-i\theta}y_{21}(x)$.*

**Proof** The first variations are found in [20]. The computation of the second variation of the reparametrization functions $\partial_{uv} f_{\rho(0)}$ of (5-9) is again divided into computations of I and II.

**Compute I** The major difference between the case $\partial_i g_{\alpha\alpha}(\sigma)$ and $\partial_\beta g_{\alpha\alpha}(\sigma)$ is the computation of this term. As before, our flat connection is

$$D_{H(u,v)} = \nabla_{\bar{\partial}_E, H(u,v)} + \Phi(u,v) + \Phi(u,v)^{*H(u,v)}.$$

For the computation of $\partial_u f_{\rho(0)}$ and $\partial_v f_{\rho(0)}$, when $u = 0$ or $v = 0$, the harmonic metric $H(u,v)$ is diagonal and one obtains

$$\partial_u D_{H(0)} = \begin{bmatrix} 0 & 0 & q_\alpha \\ 0 & 0 & 0 \\ 4\bar{q}_\alpha & 0 & 0 \end{bmatrix}, \quad \partial_v D_{H(0)} = \begin{bmatrix} 0 & q_i & 0 \\ 2\bar{q}_i & 0 & q_i \\ 0 & 2\bar{q}_i & 0 \end{bmatrix}.$$

However, when $u \neq 0$ and $v \neq 0$ both hold, the harmonic metric $H(u,v)$ corresponding to our Higgs field $\Phi(u,v)$ is not diagonal. The computation of $\partial^2 D_{H(0)}/\partial u\, \partial v$ requires an analysis of Hitchin's equations.

We start from the family of Hitchin's equations

$$F_{D_{H(u,v)}} + [\Phi(u,v), \Phi(u,v)^{*H(u,v)}] = 0. \tag{7-1}$$

We take $u$– and $v$–derivatives of Hitchin's equations (7-1) at $u, v = 0$:

$$\partial_u \partial_v \big(F_{D_{H(u,v)}} + [\Phi(u,v), \Phi(u,v)^{*H(u,v)}]\big)(0,0)(0) = 0. \tag{7-2}$$

We consider taking $H^{-1}\partial_{vu} H$ as a variable. We define

$$Y = H^{-1}\partial_{vu} H = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix}.$$

$Y = H^{-1}\partial_{uv} H$ is a section of $\mathrm{End}(E)$.

We now work with local coordinates and local trivialization. When varying the real parameters $u$ and $v$, the holomorphic structure of our bundle $E$ does not change. Thus, fixing a local holomorphic frame for all $u$ and $v$, the Chern connection 1–form under this frame compatible with the Hermitian metric $H(u,v)$ is $A(u,v) = H(u,v)^{-1}\,\partial H(u,v)$. The curvature term in our holomorphic frame is

$$F_{D_{H(u,v)}} = dA(u,v) + A(u,v) \wedge A(u,v) = \bar{\partial}(H(u,v)^{-1}\,\partial H(u,v)).$$

The section $Y \in \Gamma(\mathrm{End}(E))$ in a local holomorphic frame has the following properties:

(i) $\mathrm{Tr}(Y) = 0$.

(ii) $H(u,v)^* = H(u,v)$. Also, because $u$ and $v$ are real parameters, we have $\partial_{uv} H = \partial_{uv}(H^*) = (\partial_{uv} H)^*$ and $Y^* = HYH^{-1}$.

We can express $\partial^2 D_{H(0)}/\partial u\,\partial v$ in terms of $Y$ on $\gamma$. With respect to the local holomorphic frame introduced in the model case adapted to the Fermi coordinate, we have $\partial H = 0$ on $\gamma$. So

$$\begin{aligned}
\text{(7-3)}\quad \partial_{uv}(D_{H(0)})|_\gamma
&= \partial_{uv}(H(u,v)^{-1}\,\partial H(u,v) + \Phi(u,v) + \Phi(u,v)^{*H(u,v)})(0,0)(0)\\
&= -YH^{-1}\,\partial H + H^{-1}\,\partial H Y + \partial Y + \Phi^{*H} Y - Y\Phi^{*H}\\
&= \partial Y + \Phi^{*H} Y - Y\Phi^{*H}.
\end{aligned}$$

We want to simplify (7-2) as an equation about $Y$ and then solve $Y$ from (7-2).

Before we continue, we first fix some notation. We will write

$$\begin{aligned}
H &= H(0,0), & \partial_u H &= \frac{\partial H(u,v)}{\partial u}\bigg|_{u,v=0}, & & \\
\Phi &= \Phi(0,0), & \partial_v H &= \frac{\partial H(u,v)}{\partial v}\bigg|_{u,v=0}, & \partial_{uv} H &= \frac{\partial H(u,v)}{\partial u\,\partial v}\bigg|_{u,v=0}.
\end{aligned}$$

As a generalization of the classic result of Ahlfors, the first variations of the harmonic metric vanish at the Fuchsian point (see [20, Theorem 3.5.1]). In particular,

$$\partial_u H = \partial_v H = 0.$$

Taking $H^{-1}\partial_{uv} H$ as a variable, one can verify from (7-2) that

$$\begin{aligned}
\text{(7-4)}\quad 0 = {}&\bar\partial\partial(H^{-1}\partial_{uv}H) - H^{-1}\partial H \wedge \bar\partial(H^{-1}\partial_{uv}H) - \bar\partial(H^{-1}\partial_{uv}H)\wedge H^{-1}\,\partial H\\
&+ \bar\partial(H^{-1}\,\partial H)H^{-1}\partial_{uv}H - H^{-1}\partial_{uv}H\,\bar\partial(H^{-1}\,\partial H)\\
&+ [\partial_u\Phi, (\partial_v\Phi)^{*H}] + [\partial_v\Phi, (\partial_u\Phi)^{*H}] + [\Phi, [-H^{-1}\partial_{uv}H, \Phi^{*H}]].
\end{aligned}$$

Equation (7-4) can be simplified by the observation

$$\begin{aligned}
&\bar\partial(H^{-1}\,\partial H)H^{-1}\partial_{uv}H - H^{-1}\partial_{uv}H\,\bar\partial(H^{-1}\,\partial H) + [\Phi, [-H^{-1}\partial_{uv}H, \Phi^{*H}]]\\
&\qquad = [H^{-1}\partial_{uv}H, [\Phi, \Phi^{*H}]] - [\Phi, [H^{-1}\partial_{uv}H, \Phi^{*H}]] \quad \text{(Hitchin's equation)}\\
&\qquad = [[H^{-1}\partial_{uv}H, \Phi], \Phi^{*H}] \quad \text{(Jacobi identity).}
\end{aligned}$$

As $Y = H^{-1}\partial_{uv}H$, this yields

$$\begin{aligned}
\text{(7-5)}\quad &\bar\partial\partial Y + [\Phi^{*H}, [Y,\Phi]] - H^{-1}\,\partial H \wedge \bar\partial Y - \bar\partial Y \wedge H^{-1}\,\partial H\\
&\qquad = -[\partial_u\Phi, (\partial_v\Phi)^{*H}] - [\partial_v\Phi, (\partial_u\Phi)^{*H}].
\end{aligned}$$

The PDE system (7-5) in local holomorphic frames is equivalent to the following nine scalar equations about $y_{ij}$:

(1) $\bar\partial\partial y_{11}+h(y_{22}-y_{11})=0$.

(2) $\bar\partial\partial y_{22}+h(y_{33}-2y_{22}+y_{11})=0$.

(3) $\bar\partial\partial y_{33}+h(y_{22}-y_{33})=0$.

(4) $\bar\partial\partial y_{21}+h(y_{32}-y_{21})+h^{-1}\,\partial h\,\bar\partial y_{21}=h^{-2}q_i\bar q_\alpha$.

(5) $\bar\partial\partial y_{32}+h(y_{21}-y_{32})+h^{-1}\,\partial h\,\bar\partial y_{32}=-h^{-2}q_i\bar q_\alpha$.

(6) $\bar\partial\partial y_{12}+h(y_{23}-2y_{12})-h^{-1}\,\partial h\,\bar\partial y_{12}=h^{-1}q_\alpha\bar q_i$.

(7) $\bar\partial\partial y_{23}+h(y_{12}-2y_{23})-h^{-1}\,\partial h\,\bar\partial y_{23}=-h^{-1}q_\alpha\bar q_i$.

(8) $\bar\partial\partial y_{31}+2h^{-1}\,\partial h\,\bar\partial y_{31}=0$.

(9) $\bar\partial\partial y_{13}+2hy_{13}-2h^{-1}\,\partial h\,\bar\partial y_{13}=0$.

From property (ii) of $Y$, one can thus verify (4) is equivalent to (6), (5) is equivalent to (7), and (8) is equivalent to (9). Thus, it suffices to consider the following six equations:

- $\bar\partial\partial y_{11}+h(y_{22}-y_{11})=0$.
- $\bar\partial\partial y_{22}+h(y_{33}-2y_{22}+y_{11})=0$.
- $\bar\partial\partial y_{33}+h(y_{22}-y_{33})=0$.
- $\bar\partial\partial y_{21}+h(y_{32}-y_{21})+h^{-1}\,\partial h\,\bar\partial y_{21}=h^{-2}q_i\bar q_\alpha$.
- $\bar\partial\partial y_{32}+h(y_{21}-y_{32})+h^{-1}\,\partial h\,\bar\partial y_{32}=-h^{-2}q_i\bar q_\alpha$.
- $\bar\partial\partial y_{31}+2h^{-1}\,\partial h\,\bar\partial y_{31}=0$.

We first take a look at the first three equations. We deduce from them

$$\begin{aligned}\bar\partial\partial(y_{11}+y_{22}+y_{33})&=0,\\ \bar\partial\partial(y_{11}-y_{33})-h(y_{11}-y_{33})&=0,\\ \bar\partial\partial(y_{11}+y_{33})+h(2y_{22}-(y_{11}+y_{33}))&=0.\end{aligned}$$

As $Y=H^{-1}\partial_{uv}H$ is a section of $\mathrm{End}(E)$, the components $y_{ii}\in\Gamma(\mathcal{O})$ are actually just functions on the surface $X$ for $i=1,2,3$. Recall our notation $\Delta_\sigma=4\partial_z\partial_{\bar z}/\sigma$ and the fact $h=h(0,0)=\frac{1}{2}\sigma$, so the above equations can be written independent of coordinate charts on our surface as

$$\begin{aligned}\Delta_\sigma(y_{11}+y_{22}+y_{33})&=0,\\ \Delta_\sigma(y_{11}-y_{33})-2(y_{11}-y_{33})&=0,\\ \Delta_\sigma(y_{11}+y_{33})+2(2y_{22}-(y_{11}+y_{33}))&=0.\end{aligned}$$

We have the following observations:

- From the first equation, we obtain $y_{11}+y_{22}+y_{33}=C$, where $C$ is a constant.
- Since all eigenvalues of $\Delta_\sigma$ should be nonpositive, the second equation can hold only when $y_{11}-y_{33}=0$.
- The third equation is $\Delta_\sigma(y_{11}+y_{33})-6(y_{11}+y_{33})=-4C$. By a maximum principle argument, one gets $y_{11}+y_{33}=\frac{2}{3}C$.

Thus, property (i) of $Y$ gives $y_{11}=y_{22}=y_{33}=0$.

We then continue on the other three equations. From them, we deduce

$$\begin{aligned}\bar\partial\partial(y_{21}+y_{32})+h^{-1}\,\partial h\,\bar\partial(y_{21}+y_{32})&=0,\\ \bar\partial\partial(y_{21}-y_{32})-2h(y_{21}-y_{32})+h^{-1}\,\partial h\,\bar\partial(y_{21}-y_{32})&=2h^{-2}q_i\bar q_\alpha,\\ \bar\partial\partial y_{31}+2h^{-1}\,\partial h\,\bar\partial y_{31}&=0.\end{aligned}$$

Let $w=y_{21}+y_{32}$. We want to compute $\Delta_h\|w\|_h^2$, where the $h$–norm $\|\cdot\|_h$ is defined as

$$\|s\|_h^2=h^{-i}s\bar s$$

for a section $s\in\Gamma(K^i)$ and $i\in\mathbb{Z}$.

Because $h=h(0,0)=\frac{1}{2}\sigma$ and $\sigma=e^{\delta(z)}|dz|^2$ is a hyperbolic metric with curvature $K(\sigma)=-\Delta_\sigma(\log\sigma)=-1$, we have that $h$ satisfies

$$\bar\partial\partial h=\frac{\partial h\,\bar\partial h}{h}+\tfrac{1}{2}h^2. \tag{7-6}$$

Note $w\in\Gamma(K^{-1})$. The metric $h$ induces a Chern connection $\nabla^h$ on $K^{-1}$ and, in our local holomorphic frames, one has

$$\nabla^{h,(1,0)}w=\partial w+h^{-1}\,\partial h\,w.$$

One recognizes $\nabla^{h,(1,0)}w$ is a section of $\Omega^{(1,0)}(K^{-1})=\Gamma(\mathcal{O})$. Therefore,

$$\|\nabla^{h,(1,0)}w\|_h^2=(\partial w+h^{-1}\,\partial h\,w)(\overline{\partial w+h^{-1}\,\partial h\,w}). \tag{7-7}$$

Combining (7-6) and (7-7) gives

$$\Delta_h\|w\|_h^2=\frac{4\bar\partial\partial(hw\bar w)}{h}=2\|w\|_h^2+4\|\partial\bar w\|_h+4\|\nabla^{h,(1,0)}w\|_h^2\geq 0.$$

This is an inequality independent of coordinates valid on the Riemann surface. By a maximum principle argument, $\|w\|_h^2$ must be a constant $M$. If $M\neq 0$, then $0=\Delta_h(M)\geq 2M>0$, leading to a contradiction. Thus, $M=0$ and $y_{21}+y_{32}=0$.

We have similar arguments for $\bar\partial\partial y_{31} + 2h^{-1}\,\partial h\,\bar\partial y_{31} = 0$. We begin with computing $\Delta_h \|y_{31}\|_h^2$.

Since $y_{31}$ is a section of $\Gamma(K^{-2})$, in local holomorphic frames, the Chern connection $\nabla^h$ induced from $h$ in this case acts as $\nabla^{h,(1,0)} y_{31} = \partial y_{31} + h^{-2}\,\partial(h^2) y_{31}$.

We obtain

$$\Delta_h \|y_{31}\|_h^2 = \frac{4\bar\partial\partial(h^2 y_{31}\bar y_{31})}{h} = \|y_{31}\|_h^2 + 4\|\bar\partial y_{31}\|_h + 4\|\nabla^{h,(1,0)} y_{31}\|_h^2 \geq 0.$$

Similar to the argument for $w$, this leads to $y_{31} = 0$.

We conclude up to this point that $Y = H^{-1}\partial_{vu} H \in \Gamma(\mathrm{End}(E))$ in our local frame is of the form

$$Y = H^{-1}\partial_{uv} H = \begin{bmatrix} 0 & h\bar y_{21} & 0 \\ y_{21} & 0 & -h\bar y_{21} \\ 0 & -y_{21} & 0 \end{bmatrix}$$

with $\bar\partial\partial y_{21} - 2h y_{21} + h^{-1}\,\partial h\,\bar\partial y_{21} = h^{-2} q_i \bar q_\alpha$.

With respect to the Fermi coordinate, we have $h(z) = \frac{1}{2}$ and $\partial_z h = 0$ on $\gamma$. Also, we know $Y^* = HYH^{-1}$, so we finally obtain on $\gamma$, from (7-3),

$$\mathrm{Tr}\biggl(\frac{\partial^2 D_{H(0)}}{\partial u\,\partial v}\pi(0)\biggr)(x) = \mathrm{Tr}\biggl(\frac{\partial^2 D_{H(0)}}{\partial u\,\partial v}(x)\pi(0)\biggr) = -\tfrac{1}{2}\,\mathrm{Re}\, y_{21}(x).$$

**Remark 7.2** We have $y_{21}(x) = y_{21}(z)$, where $x = \dot\gamma(0)$ is the starting point of $\gamma$. Recall $y_{21}$ is the component of $Y \in \Gamma(\mathrm{End}(E))$ taking $\mathcal{K}$ to $\mathcal{O}$ and $y_{21}(z)$ is $y_{21}$ evaluating at $p(x)$ in the trivialization given by the holomorphic frame adapted to the Fermi coordinate $z$ for $\gamma$.

In particular, if we consider another closed geodesic $\gamma_2$ starting from $\gamma_2'(0) = e^{i\theta}x$ with its Fermi coordinate around $\gamma_2$ to be $w$, then $y_{21}(e^{i\theta}x) = y_{21}(w)$. We have $y_{21}(w) = y_{21}(z)\,dw/dz = y_{21}(z)e^{i\theta}$.

Because the vectors tangent to periodic orbits are dense in $TX$, we can extend $y_{21}$ to be everywhere defined on $UX$. We conclude that, as a function on $UX$, $y_{21}$ transfers as

$$y_{21}(e^{i\theta}x) = e^{-i\theta} y_{21}(x).$$

This finishes the computation of I on $UX$. We now move to II; together, these provide an expression for the second variations of the reparametrization functions.

**Compute II** We have

$$\begin{aligned}\operatorname{Tr}(\partial_u D_{H(0)}\partial_v\pi(0))\\ &= q_\alpha(\partial_v a_{11}(0)e_{13}(0)+a_{11}(0)\partial_v e_{13}(0))+4\bar{q}_\alpha(\partial_v a_{31}(0)e_{11}(0)+a_{31}(0)\partial_v e_{11}(0)).\end{aligned}$$

Similar to the model case $\partial_\beta g_{\alpha\alpha}$, here $\partial_v e_1(0)=y$ is the solution of a nonhomogeneous ODE system which arises from taking a $v$–derivative on the system of the parallel transport equation (5-10) at $v=0$,

$$\partial_t\begin{bmatrix} y_1(t)\\ y_2(t)\\ y_3(t)\end{bmatrix}+\begin{bmatrix}0 & \frac{1}{2} & 0\\ 1 & 0 & \frac{1}{2}\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix} y_1(t)\\ y_2(t)\\ y_3(t)\end{bmatrix}=\frac{\sqrt{2}}{2}e^t\begin{bmatrix} q_i(\Phi_t(x))\\ -2\operatorname{Re}q_i(\Phi_t(x))\\ 2\bar{q}_i(\Phi_t(x))\end{bmatrix},$$

with boundary conditions

$$H(y(0),e_1(0,0))=0,\quad y(l_\gamma)=e^{l_\gamma}\biggl(\int_0^{l_\gamma}2\operatorname{Re}q_i(\Phi_s(x))\,ds\biggr)e_1(0,0)+e^{l_\gamma}y(0).$$

The boundary conditions are set up based on the same consideration as the case of $\partial_\beta g_{\alpha\alpha}(\sigma)$. The solution is

$$\begin{aligned}\begin{bmatrix}\partial_v e_{11}(t)\\ \partial_v e_{12}(t)\\ \partial_v e_{13}(t)\end{bmatrix}\\ &=\begin{bmatrix}\frac{\sqrt{2}}{2}\int_0^t(e^t\operatorname{Re}q_i+ie^s\operatorname{Im}q_i)\,ds\\ -\sqrt{2}\int_0^t e^t\operatorname{Re}q_i\,ds\\ \sqrt{2}\int_0^t(e^t\operatorname{Re}q_i-ie^s\operatorname{Im}q_i)\,ds\end{bmatrix}+\begin{bmatrix}\frac{\sqrt{2}}{2}(e^{l_\gamma}-1)^{-1}\int_0^{l_\gamma}ie^s\operatorname{Im}q_i\,ds\\ 0\\ -\sqrt{2}(e^{l_\gamma}-1)^{-1}\int_0^{l_\gamma}ie^s\operatorname{Im}q_i\,ds\end{bmatrix}.\end{aligned}$$

Similarly, one can compute $\partial_v e_2(0)$ and $\partial_v e_3(0)$ by this method. It turns out that

$$\begin{aligned}\operatorname{Tr}(\partial_u D_{H(0)}\partial_v\pi(0))(\Phi_t(x))\\ &=2\operatorname{Im}q_\alpha(\Phi_t(x))\int_0^t(e^{s-t}-e^{t-s})\operatorname{Im}q_i(\Phi_s(x))\,ds\\ &\qquad+2\operatorname{Im}q_\alpha(\Phi_t(x))\int_0^{l_\gamma}\biggl(\frac{e^{s-t}}{e^{l_\gamma}-1}-\frac{e^{t-s}}{e^{-l_\gamma}-1}\biggr)\operatorname{Im}q_i(\Phi_s(x))\,ds.\end{aligned}$$

We therefore obtain, for a closed geodesic $\gamma$ of length $l_\gamma$ starting from $\dot{\gamma(0)}=x$,

$$\operatorname{Tr}(\partial_u D_H(0)\partial_v\pi(0))(x)=2\operatorname{Im}q_\alpha(x)\int_0^{l_\gamma}\biggl(\frac{e^s}{e^{l_\gamma}-1}-\frac{e^{-s}}{e^{-l_\gamma}-1}\biggr)\operatorname{Im}q_i(\Phi_s(x))\,ds.$$

Similar to our model case of $g_{\alpha\alpha,\beta}(\sigma)$, one can define a function $\eta\colon W\to\mathbb{R}$,

$$\eta(x)=2\operatorname{Im}q_\alpha(x)\biggl(\int_0^\infty e^{-s}\operatorname{Im}q_i(\Phi_s(x))\,ds+\int_{-\infty}^0 e^s\operatorname{Im}q_i(\Phi_s(x))\,ds\biggr),$$

and we verify that $\eta(x)$ is Hölder, so that $\operatorname{Tr}(\partial_u D_{H(0)}\partial_v\pi(0))(x)\equiv\eta(x)$ on $UX$.

We conclude

$$\begin{aligned}
&\partial_{uv} f_{\rho(0)}(x)\\
&\sim -\partial_v\bigl(\mathrm{Tr}(\partial_u D_{H(0)}\pi(0))\bigr)(x)\\
&= \tfrac{1}{2}\operatorname{Re} y_{21}(x) - 2\operatorname{Im} q_\alpha(x)\biggl(\int_0^\infty e^{-s}\operatorname{Im} q_i(\Phi_s(x))\,ds + \int_{-\infty}^0 e^{s}\operatorname{Im} q_i(\Phi_s(x))\,ds\biggr).
\end{aligned}$$

This finishes the proof of Proposition 7.1. □

**Remark 7.3** Instead of starting from the first variation of the reparametrization functions $\partial_u f_{\rho(0)}(x) \sim -\mathrm{Tr}(\partial_u D_{H(0)}\pi(0))(x)$, we can take the first variation of the reparametrization functions to be $\partial_v f_{\rho(0)}(x) \sim -\mathrm{Tr}(\partial_v D_{H(0)}\pi(0))(x)$ by (5-1) and consider

$$\begin{aligned}
\partial_{vu} f_{\rho(0)}(x) &\sim -\partial_u\bigl(\mathrm{Tr}(\partial_v D_{H(0)}\pi(0))\bigr)(x)\\
&= -\mathrm{Tr}\biggl(\frac{\partial^2 D_{H(0)}}{\partial v\,\partial u}\pi(0)\biggr)(x) - \mathrm{Tr}(\partial_v D_{H(0)}\partial_u\pi(0))(x).
\end{aligned}$$

By the same method, we get

$$\begin{aligned}
&\mathrm{Tr}(\partial_v D_{H(0)}\partial_u\pi(0))(\Phi_t(x))\\
&\quad= 2\operatorname{Im} q_i(\Phi_t(x))\int_0^t (e^{s-t} - e^{t-s})\operatorname{Im} q_\alpha(\Phi_s(x))\,ds\\
&\qquad\qquad + 2\operatorname{Im} q_i(\Phi_t(x))\int_0^{l_\gamma}\biggl(\frac{e^{s-t}}{e^{l_\gamma}-1} - \frac{e^{t-s}}{e^{-l_\gamma}-1}\biggr)\operatorname{Im} q_\alpha(\Phi_s(x))\,ds.
\end{aligned}$$

One can verify, by Fubini's theorem,

$$\int_0^{l_\gamma}\mathrm{Tr}(\partial_u D_{H(0)}\partial_v\pi(0))(\Phi_t(x))\,dt = \int_0^{l_\gamma}\mathrm{Tr}(\partial_v D_{H(0)}\partial_u\pi(0))(\Phi_t(x))\,dt.$$

This agrees with the fact that $\partial_v\bigl(\mathrm{Tr}(\partial_u D_{H(0)}\pi(0))\bigr)(x)$ and $\partial_u\bigl(\mathrm{Tr}(\partial_v D_{H(0)}\pi(0))\bigr)(x)$ should be in the same Livšic class by Livšic's theorem.

#### 7.1.2 Evaluation on the Poincaré disk

With the computation in the last section, we have

$$\begin{aligned}
\partial_i g_{\alpha\alpha}(\sigma) &= \partial_v\bigl(\langle\partial_u\rho(0,v), \partial_u\rho(0,v)\rangle_P\bigr)(0)\\
&= \lim_{r\to\infty}\frac{1}{r}\biggl[\int_{UX}\biggl(\int_0^r \operatorname{Re} q_\alpha\,dt\biggr)^2\int_0^r -2\operatorname{Re} q_i\,dt\,dm_0\\
&\qquad\qquad - 2\int_{UX}\int_0^r \operatorname{Re} q_\alpha\,dt\int_0^r \partial_{uv} f^N_{\rho(0)}\,dt\,dm_0\biggr],
\end{aligned}$$

where $\partial_{uv} f^N_{\rho(0)} = -\partial_{uv} h(\rho(0)) - \partial_{vu} f_{\rho(0)}$ and

$$\partial_{uv} f_{\rho(0)}(x)$$
$$\sim \tfrac{1}{2} \operatorname{Re} y_{21}(x) - 2 \operatorname{Im} q_\alpha(x) \bigg( \int_0^\infty \operatorname{Im} q_i(\Phi_s(x)) e^{-s}\, ds + \int_{-\infty}^0 \operatorname{Im} q_i(\Phi_s(x)) e^{s}\, ds \bigg).$$

We show in this subsection:

**Proposition 7.4** *For $\sigma \in \mathcal{T}(S)$, $\partial_i g_{\alpha\alpha}(\sigma) = 0$.*

The argument for this proposition boils down to the following lemma:

**Lemma 7.5** *For any $t, s \in \mathbb{R}$,*

$$\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0(x) = 0, \tag{7-8}$$

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0(x) = 0, \tag{7-9}$$

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Im} q_\alpha(\Phi_t(x)) \operatorname{Im} q_i(\Phi_s(x))\, dm_0(x) = 0. \tag{7-10}$$

**Proof** The proof of this lemma is basically the same as the proof of Lemma 6.2 except that flow time $s = \frac{1}{2}t$ tells us nothing in this case. We instead choose the flow times to be the three special cases $s = t$, $s = 2t$ and $s = 3t$. We recall our analytic expansions for $q_i$ and $q_\alpha$ are

$$q_i(\Phi_r(e^{i\theta} x)) = \bigg( \sum_{n=0}^\infty c_n(x) R^n (1 - R^2)^2 e^{i(n+2)\theta} \bigg),$$
$$q_\alpha(\Phi_r(e^{i\theta} x)) = \bigg( \sum_{n=0}^\infty a_n(x) R^n (1 - R^2)^3 e^{i(n+3)\theta} \bigg).$$

We have, when $t, s > 0$,

$$\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0(x)$$
$$= \frac{1}{2\pi} \int_0^{2\pi} \int_{UX} \operatorname{Re} q_i(e^{i\theta} x) \operatorname{Re} q_\alpha(\Phi_t(e^{i\theta} x)) \operatorname{Re} q_\alpha(\Phi_s(e^{i\theta} x))\, dm_0(x)\, d\theta$$
$$= \frac{1}{4} \sum_{n=0}^\infty \int_{UX} \operatorname{Re}(c_0 a_n \bar{a}_{n+2})\, dm_0 T^n S^n (1 - T^2)^3 (1 - S^2)^3 (S^2 + T^2).$$

Consider $t = s > 0$. Then

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_i(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_t(x))\, dm_0(x) \\ &= \int_{UX} \operatorname{Re} q_i(\Phi_{-t}(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(x)\, dm_0(x) \\ &= \int_{UX} \operatorname{Re} q_i(-\Phi_t(-x)) \operatorname{Re} q_\alpha(-x) \operatorname{Re} q_\alpha(-x)\, dm_0(x) \\ &= \int_{UX} \operatorname{Re} q_i(\Phi_t(y)) \operatorname{Re} q_\alpha(y) \operatorname{Re} q_\alpha(y)\, dm_0(y) \qquad \text{(with } y = -x).\end{aligned}$$

The analytic expansions of the left- and right-hand sides of the above equation give

$$\tag{7-11} \frac{1}{2}\sum_{n=0}^{\infty} \int_{UX} \operatorname{Re}(c_0 a_n \bar{a}_{n+2})\, dm_0 T^{2n}(1-T^2)^6 T^2 = \frac{1}{4}\int_{UX} \operatorname{Re}(a_0 a_0 \bar{c}_4)\, dm_0 (1-T^2)^2 T^4.$$

We let $C_n = \int_{UX} \operatorname{Re}(c_0 a_n \bar{a}_{n+2})\, dm_0$ and $D_n = \int_{UX} \operatorname{Re}(a_0 a_n \bar{c}_{n+4})\, dm_0$ for $n \geq 0$. We proceed to prove $C_n = 0$ for $n \geq 0$.

The coefficients of $T^0$ and $T^2$ yield

$$C_0 = 0, \quad 2C_1 - 8C_0 = D_0.$$

On the other hand, if we consider $s = 2t$ and $s = 3t$, they lead to

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_i(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_{2t}(x))\, dm_0(x) \\ &= -\int_{UX} \operatorname{Re} q_i(\Phi_t(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(-x))\, dm_0(y)\end{aligned}$$

and

$$\begin{aligned}\int_{UX} &\operatorname{Re} q_i(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_{3t}(x))\, dm_0(x) \\ &= -\int_{UX} \operatorname{Re} q_i(\Phi_t(x)) \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_{2t}(-x))\, dm_0(y).\end{aligned}$$

When $s = 2t$, we have $S = 2T/(T^2+1)$ and $S = 2T + O(T^3)$. When $s = 3t$, we have $S = (3T + T^3)/(3T^2+1)$ and $S = 3T + O(T^3)$.

Compare coefficients of $T^4$ of the analytic expansions of the above two equations and use the relations $S = 2T + O(T^3)$ and $S = 3T + O(T^3)$ to obtain $D_0 = 0$. Therefore, from (7-11), we conclude $C_n = 0$ for $n \geq 0$ and (7-8) holds for $t, s > 0$. For $s \leq 0$ or

$t \leq 0$, the argument for (7-8) to hold is an analogy of the $\partial_\beta g_{\alpha\alpha}(\sigma)$ case. We omit it here.

Equation (7-9) then follows from (7-8) by a $\Phi_t$–invariance argument for $m_0$. To prove (7-10), we just need

$$\int_{UX} \operatorname{Re} q_\alpha(x) q_i(\Phi_t(x)) q_i(\Phi_s(x))\, dm_0(x) = 0 \quad \text{for all } t, s \in \mathbb{R}.$$

The argument is the same as the argument for Lemma 6.2. This finishes the proof of Lemma 7.5. □

**Proof of Proposition 7.4** We begin by showing

$$\lim_{r\to\infty} \frac{1}{r} \int_{UX} 2 \int_0^r \partial_u f^N_{\rho(0)}\, dt \int_0^r \operatorname{Tr}\left(\frac{\partial^2 D_{H(0)}}{\partial u\, \partial v} \pi(0)\right) dt\, dm_0 = 0$$

by evaluating the integral on the Poincaré disk.

Recall from the last subsection that $y_{21}$ is the solution of $\bar\partial \partial y_{21} - 2h y_{21} + h^{-1}\, \partial h\, \bar\partial y_{21} = h^{-2} q_i \bar q_\alpha$. Because $q_i$ and $\bar q_\alpha$ are real analytic and because $h = h(0,0) = \frac{1}{2}\sigma$ is also real analytic, we know $y_{21}$ is real analytic by analytic elliptic regularity theory [12].

As discussed before, the function $y_{21}$ on $UX$ transfers as $y_{21}(e^{i\theta}x) = e^{-i\theta} y_{21}(x)$. Similarly to the model case of $g_{\alpha\alpha,\beta}$, we write the real analytic expansion for $y_{21}$ in the coordinates given by the Poincaré disk model based on $x$,

$$y_{21,x}(z) = \sum_{n,m\geq 0} b_{n,m}(x) z^n \bar z^m \frac{\partial}{\partial z}.$$

Define $\tilde y_{21,x}(z) := \operatorname{Re}(y_{21,x}(z)(dr))$. Recall $r(R) = \frac{1}{2}\log((1-R)/(1+R))$. One has

$$y_{21}(\Phi_r(e^{i\theta}x)) = \tilde y_{21,x}(Re^{i\theta}) = \operatorname{Re}\left(\sum_{n,m\geq 0} b_{n,m}(x) R^{n+m} (1-R^2)^{-1} e^{i(n-m-1)\theta}\right).$$

Thus,

$$\begin{aligned}
&\lim_{r\to\infty} \frac{1}{r} \int_{UX} 2 \int_0^r \partial_u f^N_{\rho(0)}\, dt \int_0^r \operatorname{Tr}\left(\frac{\partial^2 D_{H(0)}}{\partial u\, \partial v} \pi(0)\right) dt\, dm_0 \\
&\qquad = \lim_{r\to\infty} \frac{1}{r} \int_{UX} \int_0^r \operatorname{Re} q_\alpha(\Phi_t(x))\, dt \int_0^r \operatorname{Re} y_{21}(\Phi_t(x))\, dt\, dm_0 \\
&\qquad = \lim_{r\to\infty} \frac{1}{r} \int_0^r \int_0^r \int_{UX} \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} y_{21}(\Phi_s(x))\, dm_0\, dt\, ds \\
&\qquad = \lim_{r\to\infty} \frac{1}{r} \int_0^r \int_0^r \int_{UX} \operatorname{Re} q_\alpha(\Phi_{t-s}(x)) \operatorname{Re} y_{21}(x)\, dm_0\, dt\, ds.
\end{aligned}$$

When $\mu = t - s \geq 0$,

$$\int_{UX} \operatorname{Re} q_\alpha(\Phi_\mu(x)) \operatorname{Re} y_{21}(x)\, dm_0 = \frac{1}{2\pi} \int_{UX} \int_0^{2\pi} \operatorname{Re} q_\alpha(\Phi_\mu(e^{i\theta} x)) \operatorname{Re} y_{21}(e^{i\theta} x)\, d\theta\, dm_0.$$

However, $\int_0^{2\pi} \operatorname{Re}(e^{-i\theta} b_{0,0}) \operatorname{Re}(a_n e^{i(n+3)\theta})\, d\theta = 0$ for all $n \geq 0$, which implies the above is zero.

It also holds for $\mu \leq 0$ by simply observing that $\operatorname{Re} q_\alpha(\Phi_{-\mu}(-x)) = -\operatorname{Re} q_\alpha(\Phi_\mu(x))$. Therefore, we conclude

$$\lim_{r\to\infty} \frac{1}{r} \int_{UX} 2 \int_0^r \partial_u f^N_{\rho(0)}\, dt \int_0^r \operatorname{Tr}\biggl(\frac{\partial^2 D_{H(0)}}{\partial u\, \partial v} \pi(0)\biggr)\, dt\, dm_0 = 0.$$

Arguments for the other terms in $\partial_i g_{\alpha\alpha}(\sigma)$ to be equal to zero are analogous to the model case of $\partial_\beta g_{\alpha\alpha}(\sigma)$. They all reduce to Lemma 7.5. We thus finish the proof of Proposition 7.4. □

## 7.2 The case of $\partial_j g_{\alpha i}(\sigma)$

The proofs for the case of $\partial_j g_{\alpha i}(\sigma)$ in this subsection and the case of $\partial_\beta g_{\alpha i}(\sigma)$ in the next subsection are basically the same as the cases for $\partial_\beta g_{\alpha\alpha}(\sigma)$ and $\partial_i g_{\alpha\alpha}(\sigma)$. Although there are no new ingredients in the proofs, we include them here for completeness.

For $\partial_j g_{\alpha i}(\sigma)$, we have three parameters $\{(u,v,w)\} \in \{(-1,1)\}^3$. The representations $\{\rho(u,v,w)\}$ in $\mathcal{H}_3(S)$ correspond to $\{(vq_i + wq_j, uq_\alpha)\} \subset H^0(X,K^2) \oplus H^0(X,K^3)$ by Hitchin parametrization. In particular, we have $\partial_u \rho(0,0,0)$ is identified with $\varphi(q_\alpha)$ and $\partial_v \rho(0,0,0)$ is identified with $\varphi(q_i)$. Also $\partial_w \rho(0,0,0)$ is identified with $\varphi(q_j)$. The formula for $\partial_j g_{\alpha i}(\sigma)$ is

$$\begin{aligned}
\partial_j g_{\alpha i}(\sigma) &= \partial_w\bigl(\langle \partial_u \rho(0,0,w), \partial_v \rho(0,0,w)\rangle_P\bigr)(0) \\
&= \lim_{r\to\infty} \frac{1}{r} \biggl( \int_{UX} \int_0^r \partial_u f^N_{\rho(0)}\, dt \int_0^r \partial_v f^N_{\rho(0)}\, dt \int_0^r \partial_w f^N_{\rho(0)}\, dt\, dm_0 \\
&\qquad + \int_{UX} \int_0^r \partial_u f^N_{\rho(0)}\, dt \int_0^r \partial_{vw} f^N_{\rho(0)}\, dt\, dm_0 \\
&\qquad + \int_{UX} \int_0^r \partial_v f^N_{\rho(0)}\, dt \int_0^r \partial_{uw} f^N_{\rho(0)}\, dt\, dm_0 \biggr),
\end{aligned}$$

where the first and second variations are:

(i) $\partial_u f^N_{\rho(0)} = -\partial_u f_{\rho(0)}$.

(ii) $\partial_v f^N_{\rho(0)} = -\partial_v f_{\rho(0)}$.

(iii) $\partial_{uw} f^N_{\rho(0)} = -\partial_{uw} h(\rho(0)) - \partial_{uw} f_{\rho(0)}$.

(iv) $\partial_{vw} f^N_{\rho(0)} = -\partial_{vw} h(\rho(0)) - \partial_{vw} f_{\rho(0)}$.

**7.2.1 First and second variations of the reparametrization functions** Our Higgs field in this case is

$$\Phi(u,v,w) = \begin{bmatrix} 0 & vq_i + wq_j & uq_\alpha \\ 1 & 0 & vq_i + wq_j \\ 0 & 1 & 0 \end{bmatrix}.$$

Following the steps and methods from the cases $\partial_\beta g_{\alpha\alpha}(\sigma)$ and $\partial_\beta g_{\alpha\alpha}(\sigma)$, we have:

**Proposition 7.6** *The first variations of the reparametrization functions $\partial_u f_{\rho(0)}\colon UX \to \mathbb{R}$, $\partial_v f_{\rho(0)}\colon UX \to \mathbb{R}$ and $\partial_w f_{\rho(0)}\colon UX \to \mathbb{R}$ for the case $\partial_j g_{\alpha i}(\sigma)$ satisfy*

$$\partial_u f_{\rho(0)}(x) \sim -\operatorname{Re} q_\alpha(x), \quad \partial_v f_{\rho(0)}(x) \sim 2\operatorname{Re} q_i(x), \quad \partial_w f_{\rho(0)}(x) \sim 2\operatorname{Re} q_j(x),$$

*and the second variations of the reparametrization functions $\partial_{uw} f_{\rho(0)}\colon UX \to \mathbb{R}$ and $\partial_{vw} f_{\rho(0)}\colon UX \to \mathbb{R}$ satisfy*

$$\begin{aligned}\partial_{uw} f_{\rho(0)} &\sim \tfrac{1}{2}\operatorname{Re} y_{21}(x)\\ &\qquad -2\operatorname{Im} q_\alpha(x)\biggl(\int_0^\infty \operatorname{Im} q_i(\Phi_s(x))e^{-s}\,ds + \int_{-\infty}^0 \operatorname{Im} q_i(\Phi_s(x))e^{s}\,ds\biggr),\\ \partial_{vw} f_{\rho(0)}(x) &\sim \tfrac{1}{2}\phi_{vw}(p(x))\\ &\qquad +2\operatorname{Im} q_i(x)\biggl(\int_0^\infty \operatorname{Im} q_j(\Phi_s(x))e^{-s}\,ds + \int_{-\infty}^0 \operatorname{Im} q_j(\Phi_s(x))e^{s}\,ds\biggr),\end{aligned}$$

*where $p\colon UX \to X$ and $y_{21}$ are defined as before.*

**Proof** For the second variations of the reparametrization functions, we have computed $\partial_{uw} f_{\rho(0)}$ in the $\partial_i g_{\alpha\alpha}(\sigma)$ case:

$$\begin{aligned}\partial_{uw} f_{\rho(0)} &\sim \tfrac{1}{2}\operatorname{Re} y_{21}(x)\\ &\qquad -2\operatorname{Im} q_\alpha(x)\biggl(\int_0^\infty \operatorname{Im} q_i(\Phi_s(x))e^{-s}\,ds + \int_{-\infty}^0 \operatorname{Im} q_i(\Phi_s(x))e^{s}\,ds\biggr).\end{aligned}$$

The computation of

$$\partial_{vw} f_{\rho(0)} \sim -\operatorname{Tr}\biggl(\frac{\partial^2 D_{A(0)}}{\partial w\,\partial v}\pi(0)\biggr) - \operatorname{Tr}(\partial_v D_{H(0)}\partial_w \pi(0)) =: -\mathrm{I} - \mathrm{II}$$

is divided into computations of I and II.

**Compute I** We set $u=0$; the Higgs field is

$$\Phi(v,w)=\begin{bmatrix}0 & vq_i+wq_j & 0\\ 1 & 0 & vq_i+wq_j\\ 0 & 1 & 0\end{bmatrix}.$$

The harmonic metric $H(v,w)$ is diagonalizable and the computation of $\partial_{vw}D_{H(0)}$ is the same as in the model case of $\partial_\beta g_{\alpha\alpha}(\sigma)$.

With respect to the notation defined in the model case of $\partial_\beta g_{\alpha\alpha}(\sigma)$, one obtains

$$\operatorname{Tr}\Bigl(\frac{\partial^2 D_{H(0)}}{\partial v\,\partial w}\pi(0)\Bigr)(x)=-\tfrac{1}{2}\psi_{vw}\bigl(z(p(x))\bigr)=-\tfrac{1}{2}\phi_{vw}(p(x)),$$

where $p\colon UX\to X$ is the projection from the unit tangent bundle to our surface and $z$ is the Fermi coordinate we choose evaluating at the point $p(x)\in X$.

**Compute II** Both $\partial_v D_{H(0)}$ and $\partial_w\pi(0)$ have been computed in the $\partial_i g_{\alpha\alpha}(\sigma)$ case. One can check

$$\begin{aligned}
&\operatorname{Tr}(\partial_v D_{H(0)}\partial_w\pi(0))(\Phi_t(x))\\
&=q_i(\partial_w a_{11}(0)e_{12}(0)+a_{11}(0)\partial_w e_{12}(0)+\partial_w a_{21}(0)e_{13}(0)+a_{21}(0)\partial_w e_{13}(0))\\
&\qquad+2\bar q_i(\partial_w a_{21}(0)e_{11}(0)+a_{21}(0)\partial_w e_{11}(0)+\partial_w a_{31}(0)e_{12}(0)+a_{31}(0)\partial_w e_{12}(0))\\
&=2\operatorname{Im}q_i(\Phi_t(x))\int_0^t\operatorname{Im}q_j(\Phi_s(x))(e^{t-s}-e^{s-t})\,ds\\
&\qquad\qquad+2\operatorname{Im}q_i(\Phi_t(x))\int_0^{l_\gamma}\operatorname{Im}q_j(\Phi_s(x))\Bigl(\frac{e^{t-s}}{e^{-l_\gamma}-1}-\frac{e^{s-t}}{e^{l_\gamma}-1}\Bigr)\,ds.
\end{aligned}$$

In particular,

$$\operatorname{Tr}(\partial_v D_{H(0)}\partial_w\pi(0))(x)=2\operatorname{Im}q_i(x)\int_0^{l_\gamma}\operatorname{Im}q_j(\Phi_s(x))\Bigl(\frac{e^{-s}}{e^{-l_\gamma}-1}-\frac{e^{s}}{e^{l_\gamma}-1}\Bigr)\,ds.$$

Similarly to the cases of $\partial_\beta g_{\alpha\alpha}(\sigma)$ and $\partial_i g_{\alpha\alpha}(\sigma)$, one can then define a function $\eta\colon UX\to\mathbb{R}$,

$$\eta(x)=-2\operatorname{Im}q_i(x)\Bigl(\int_0^\infty\operatorname{Im}q_j(\Phi_s(x))e^{-s}\,ds+\int_{-\infty}^0\operatorname{Im}q_j(\Phi_s(x))e^{s}\,ds\Bigr),$$

and verify that $\eta(x)$ is Hölder and such that $\operatorname{Tr}(\partial_v D_{H(0)}\partial_w\pi(0))(x)\equiv\eta(x)$ on $UX$.

We finally obtain

$$\begin{aligned}
&\partial_{vw}f_{\rho(0)}(x)\\
&\quad\sim-\operatorname{Tr}\Bigl(\frac{\partial^2 D_{H(0)}}{\partial v\,\partial w}\pi(0)\Bigr)(x)-\operatorname{Tr}(\partial_v D_{H(0)}\partial_w\pi(0))(x)
\end{aligned}$$

$$= \tfrac{1}{2}\phi_{vw}(p(x)) + 2\operatorname{Im} q_i(x)\left(\int_0^\infty \operatorname{Im} q_j(\Phi_s(x))e^{-s}\,ds + \int_{-\infty}^0 \operatorname{Im} q_j(\Phi_s(x))e^{s}\,ds\right). \qquad \square$$

**7.2.2 Evaluation on the Poincaré disk** We show in this subsection:

**Proposition 7.7** *For $\sigma \in \mathcal{T}(S)$, $\partial_j g_{\alpha i}(\sigma) = 0$.*

For the same reasoning as before, the proof of the above proposition reduces to the following lemma:

**Lemma 7.8** *For any $t, s \in \mathbb{R}$,*

$$\int_{UX} \operatorname{Re} q_i(x)\operatorname{Re} q_j(\Phi_t(x))\operatorname{Re} q_\alpha(\Phi_s(x))\,dm_0(x) = 0, \tag{7-12}$$

$$\int_{UX} \operatorname{Re} q_i(x)\operatorname{Im} q_j(\Phi_t(x))\operatorname{Im} q_\alpha(\Phi_s(x))\,dm_0(x) = 0, \tag{7-13}$$

$$\int_{UX} \operatorname{Re} q_\alpha(x)\operatorname{Re} q_i(\Phi_t(x))\operatorname{Re} q_j(\Phi_s(x))\,dm_0(x) = 0, \tag{7-14}$$

$$\int_{UX} \operatorname{Re} q_\alpha(x)\operatorname{Im} q_i(\Phi_t(x))\operatorname{Im} q_j(\Phi_s(x))\,dm_0(x) = 0. \tag{7-15}$$

**Proof** We just need to show (7-12). Equations (7-13), (7-14) and (7-15) follow easily using the methods we developed in the former cases.

We start from a special case of (7-12), with $q_i = q_j$:

$$\int_{UX} \operatorname{Re} q_i(x)\operatorname{Re} q_i(\Phi_t(x))\operatorname{Re} q_\alpha(\Phi_s(x))\,dm_0(x) = 0 \quad \text{for all } t, s \in \mathbb{R}. \tag{7-16}$$

The proof of this case is an analogy of the case $\partial_\beta g_{\alpha\alpha}(\sigma)$ since, for flow times $s = t$ and $s = \frac{1}{2}t$,

$$\int_{UX} \operatorname{Re} q_i(x)\operatorname{Re} q_i(\Phi_t(x))\operatorname{Re} q_\alpha(\Phi_t(x))\,dm_0(x) = -\int_{UX} \operatorname{Re} q_i(x)\operatorname{Re} q_i(\Phi_t(x))\operatorname{Re} q_\alpha(x)\,dm_0(x)$$

and

$$\int_{UX} \operatorname{Re} q_i(x)\operatorname{Re} q_i(\Phi_t(x))\operatorname{Re} q_\alpha(\Phi_{t/2}(x))\,dm_0(x) = 0.$$

For $t, s > 0$, recall our analytic expansions given in (6-3) and (6-6) lead to

$$\begin{aligned}
&\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_i(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0(x) \\
&= \frac{1}{2\pi} \int_0^{2\pi} \int_{UX} \operatorname{Re} q_i(e^{i\theta}x) \operatorname{Re} q_i(\Phi_t(e^{i\theta}x)) \operatorname{Re} q_\alpha(\Phi_s(e^{i\theta}x))\, dm_0(x)\, d\theta \\
&= \frac{1}{4} \sum_{n=0}^{\infty} \biggl( \int_{UX} \operatorname{Re}(c_0 c_n \bar{a}_{n+1})\, dm_0 T^n (1-T^2)^2 S^{n+1} (1-S^2)^3 \\
&\qquad\qquad + \int_{UX} \operatorname{Re}(c_0 \bar{c}_{n+3} a_n)\, dm_0 T^{n+3} (1-T^2)^2 S^n (1-S^2)^3 \biggr).
\end{aligned}$$

Let $E_n = \int_{UX} \operatorname{Re}(c_0 c_n \bar{a}_{n+1})\, dm_0$ and $F_n = \int_{UX} \operatorname{Re}(c_0 \bar{c}_{n+3} a_n)\, dm_0$. We argue, for $n \geq 0$,

$$E_n = F_n = 0. \tag{7-17}$$

The case $t = 0$ or $s = 0$ of (7-16) is included in the $n = 0$ case of (7-17).

For flow time $s = t$, we have

$$\sum_{n=0}^{\infty} (E_n T^{2n+1} (1-T^2)^5 + F_n T^{2n+3} (1-T^2)^5) = -F_0 T^3 (1-T^2)^2. \tag{7-18}$$

This implies

$$E_0 = 0, \quad E_1 = -2F_0.$$

For flow time $s = \frac{1}{2}t$, we obtain

$$\sum_{n=0}^{\infty} (E_n T^n (1-T^2)^2 S^{n+1} (1-S^2)^3 + F_n T^{n+3} (1-T^2)^2 S^n (1-S^2)^3) = 0,$$

where $T = 2S/(1+S^2)$.

It simplifies to

$$\sum_{n=0}^{\infty} \left( E_n + 8F_n \frac{S^2}{(S^2+1)^3} \right) \left( \frac{2S^2}{S^2+1} \right)^n = 0.$$

Let $W = S^2/(1+S^2)$; we have

$$\sum_{n=0}^{\infty} \left( E_n \sum_{k=0}^{\infty} (k+1) W^k + 8F_n W \right) (2W)^n = 0.$$

This gives relations

$$E_0 = 0, \qquad \sum_{k=0}^{n} 2^k (n-k+1) E_k + 2^{n+2} F_{n-1} = 0, \quad n \geq 1. \tag{7-19}$$

Combining with (7-18), we get $E_1 = F_0 = 0$. Therefore, the right-hand side of (7-18) is zero and we obtain from it $E_{n+1} + F_n = 0$ for $n \geq 0$. Combining this with (7-19) and by an induction argument, one concludes $E_n = F_n = 0$. This proves (7-17) for $s, t \geq 0$. The case $s, t < 0$ is similar to before.

Now we proceed to prove (7-12). The above case implies, for $q_i \neq q_j$,

$$\begin{aligned}
\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_i(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0 &= 0,\\
\int_{UX} \operatorname{Re} q_j(x) \operatorname{Re} q_j(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0 &= 0,\\
\int_{UX} \operatorname{Re}(q_i + q_j)(x) \operatorname{Re}(q_i + q_j)(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0 &= 0.
\end{aligned}$$

Therefore, for all $t, s \in \mathbb{R}$,

$$\begin{aligned}
\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_j(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0 \\
+ \int_{UX} \operatorname{Re} q_j(x) \operatorname{Re} q_i(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0 = 0.
\end{aligned} \tag{7-20}$$

Recall the analytic expansion for $q_j$ is given in (6-7). Consider $t, s > 0$:

$$\begin{aligned}
&\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_j(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_s(x))\, dm_0(x)\\
&= \frac{1}{2\pi} \int_0^{2\pi} \int_{UX} \operatorname{Re} q_i(e^{i\theta} x) \operatorname{Re} q_j(\Phi_t(e^{i\theta} x)) \operatorname{Re} q_\alpha(\Phi_s(e^{i\theta} x))\, dm_0(x)\, d\theta\\
&= \frac{1}{4} \sum_{n=0}^{\infty} \biggl( \int_{UX} \operatorname{Re}(c_0 d_n \bar{a}_{n+1})\, dm_0 T^n (1-T^2)^2 S^{n+1} (1-S^2)^3\\
&\qquad + \int_{UX} \operatorname{Re}(c_0 \bar{d}_{n+3} a_n)\, dm_0 T^{n+3} (1-T^2)^2 S^n (1-S^2)^3 \biggr).
\end{aligned}$$

Let $G_n = \int_{UX} \operatorname{Re}(c_0 d_n \bar{a}_{n+1})\, dm_0$ and $H_n = \int_{UX} \operatorname{Re}(c_0 \bar{d}_{n+3} a_n)\, dm_0$. We want to show $G_n = H_n = 0$ for $n \geq 0$.

Let $m$ be an integer and $m \geq 2$. Consider the flow time $s = mt$. Observe

$$\begin{aligned}
&\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_j(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_{mt}(x))\, dm_0(x)\\
&= \int_{UX} \operatorname{Re} q_i(\Phi_{-t}(x)) \operatorname{Re} q_j(x) \operatorname{Re} q_\alpha(\Phi_{(m-1)t}(x))\, dm_0(x)
\end{aligned}$$

$$= -\int_{UX} \operatorname{Re} q_j(y) \operatorname{Re} q_i(\Phi_t(y)) \operatorname{Re} q_\alpha(\Phi_{-(m-1)t}(y))\, dm_0(y)$$

(let $y = -x$ and $\Phi_t(-y) = -\Phi_{-t}(y)$)

$$= \int_{UX} \operatorname{Re} q_i(y) \operatorname{Re} q_j(\Phi_t(y)) \operatorname{Re} q_\alpha(\Phi_{-(m-1)t}(y))\, dm_0(y) \qquad \text{(by (7-20))}$$

$$= -\int_{UX} \operatorname{Re} q_j(x) \operatorname{Re} q_i(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_{mt}(x))\, dm_0(x)$$

(exchange the roles of $q_i$ and $q_j$)

$$= -\int_{UX} \operatorname{Re} q_i(x) \operatorname{Re} q_j(\Phi_t(x)) \operatorname{Re} q_\alpha(\Phi_{(m-1)t}(-x))\, dm_0(x)$$

When $s = mt$, we have $S = S(m) = ((1+T)^m - (1-T)^m)/((1+T)^m + (1-T)^m) = mT + O(T^3)$. From the analytic expansion

$$\sum_{n=0}^{\infty} \bigl(G_n T^n S(m)^{n+1}(1-S(m)^2)^3 + G_n e^{in\pi} T^n S(m-1)^{n+1}(1-S(m-1)^2)^3\bigr)$$
$$= -\sum_{n=0}^{\infty} \bigl(-H_n e^{in\pi} T^{n+3} S(m-1)^n (1-S(m-1)^2)^3 + H_n T^{n+3} S(m)^n (1-S(m)^2)^3\bigr),$$

the coefficients of $T^1$ and $T^3$ and $T^5$ yield, respectively,

$$\begin{aligned} G_0 &= 0, \\ (m^2 - (m-1)^2) G_1 &= 0, \\ (m^3 + (m-1)^3) G_2 &= -(2m-1) H_1 + (6m-3) H_0. \end{aligned}$$

The cases $m = 2$, $m = 3$ and $m = 4$ together give $H_0 = H_1 = G_2 = 0$. By induction, assuming $G_k = H_{k-1} = 0$ for $1 \le k < n$, the coefficient of $T^{2n+1}$ gives

$$(m^{n+1} + e^{in\pi}(m-1)^{n+1}) G_n = (e^{i(n-1)\pi}(m-1)^{n-1} - m^{n-1}) H_{n-1}.$$

We conclude $G_n = H_n = 0$ for $n \ge 0$ by choosing two different $m$. This finishes the proof of (7-12) for $t, s > 0$. Equation (7-12) for $t \le 0$ and $s \le 0$ can be proved similarly to the former cases. □

### 7.3 The case of $\partial_\beta g_{\alpha i}(\sigma)$

This is the last case. In this case, the representations $\{\rho(u, v, w)\}$ in $\mathcal{H}_3(S)$ correspond to $\{(v q_i, u q_\alpha + w q_\beta)\} \subset H^0(X, K^2) \oplus H^0(X, K^3)$ by Hitchin parametrization. Our metric tensor is

$$\begin{aligned}
&\partial_\beta g_{\alpha i}(\sigma)\\
&= \partial_w\big(\langle \partial_u\rho(0,0,w), \partial_v\rho(0,0,w)\rangle_P\big)(0)\\
&= \lim_{r\to\infty}\frac{1}{r}\bigg(\int_{UX}\int_0^r \partial_u f^N_{\rho(0)}\,dt\int_0^r \partial_v f^N_{\rho(0)}\,dt\int_0^r \partial_w f^N_{\rho(0)}\,dt\,dm_0\\
&\qquad\qquad + \int_{UX}\int_0^r \partial_u f^N_{\rho(0)}\,dt\int_0^r \partial_{vw} f^N_{\rho(0)}\,dt\,dm_0\\
&\qquad\qquad\qquad + \int_{UX}\int_0^r \partial_v f^N_{\rho(0)}\,dt\int_0^r \partial_{uw} f^N_{\rho(0)}\,dt\,dm_0\bigg),
\end{aligned}$$

where the first and second variations are

(i) $\partial_u f^N_{\rho(0)} = -\partial_u f_{\rho(0)}$;

(ii) $\partial_v f^N_{\rho(0)} = -\partial_v f_{\rho(0)}$;

(iii) $\partial_{uw} f^N_{\rho(0)} = -\partial_{uw} h(\rho(0)) - \partial_{uw} f_{\rho(0)}$;

(iv) $\partial_{vw} f^N_{\rho(0)} = -\partial_{vw} h(\rho(0)) - \partial_{vw} f_{\rho(0)}$.

**7.3.1 First and second variations of the reparametrization functions** Our Higgs field in this case is

$$\Phi(u,v,w) = \begin{bmatrix} 0 & vq_i & uq_\alpha + wq_\beta \\ 1 & 0 & vq_i \\ 0 & 1 & 0 \end{bmatrix}.$$

**Proposition 7.9** *The first variations of the reparametrization functions $\partial_u f_{\rho(0)}\colon UX \to \mathbb{R}$ and $\partial_v f_{\rho(0)}\colon UX \to \mathbb{R}$ for the case $\partial_\beta g_{\alpha i}(\sigma)$ satisfy*

$$\partial_u f_{\rho(0)}(x) \sim -\operatorname{Re} q_\alpha(x), \quad \partial_v f_{\rho(0)}(x) \sim 2\operatorname{Re} q_i(x), \quad \partial_w f_{\rho(0)}(x) \sim -\operatorname{Re} q_\beta(x),$$

*and the second variations of the reparametrization functions $\partial_{uw} f_{\rho(0)}\colon UX \to \mathbb{R}$ and $\partial_{vw} f_{\rho(0)}\colon UX \to \mathbb{R}$ satisfy*

$$\begin{aligned}
&\partial_{uw} f_{\rho(0)}(x)\\
&\sim \tfrac{1}{2}\phi_{uw}(p(x)) + \operatorname{Re} q_\alpha(x)\int_0^\infty e^{-2s}\operatorname{Re} q_\beta(\Phi_s(x))\,ds\\
&\qquad + \operatorname{Re} q_\alpha(x)\int_{-\infty}^0 e^{2s}\operatorname{Re} q_\beta(\Phi_s(x))\,ds + 2\operatorname{Im} q_\alpha(x)\int_0^\infty e^{-s}\operatorname{Im} q_\beta(\Phi_s(x))\,ds\\
&\qquad\qquad + 2\operatorname{Im} q_\alpha(x)\int_{-\infty}^0 e^{s}\operatorname{Im} q_\beta(\Phi_s(x))\,ds
\end{aligned}$$

*and*

$$\begin{aligned}&\partial_{vw} f_{\rho(0)}(x)\\&= \partial_{wv} f_{\rho(0)}(x)\\&\sim \tfrac{1}{2}\operatorname{Re} y_{21}(x) - 2\operatorname{Im} q_\beta(x)\left(\int_0^\infty \operatorname{Im} q_i(\Phi_s(x))e^{-s}\,ds + \int_{-\infty}^0 \operatorname{Im} q_i(\Phi_s(x))e^{s}\,ds\right),\end{aligned}$$

*where* $p\colon UX \to X$ *and* $y_{21}$ *are defined as before.*

**Proof** All of the computations have been done in the former cases. □

**7.3.2 Evaluation on the Poincaré disk** We show in this subsection:

**Proposition 7.10** *For* $\sigma \in \mathcal{T}(S)$, $\partial_\beta g_{\alpha i}(\sigma) = 0$.

For the same reasoning as before, the proof of the above proposition reduces to the following lemma:

**Lemma 7.11** *For any* $t, s \in \mathbb{R}$,

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0(x) = 0, \tag{7-21}$$

$$\int_{UX} \operatorname{Im} q_\alpha(x) \operatorname{Im} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0(x) = 0, \tag{7-22}$$

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Im} q_\beta(\Phi_t(x)) \operatorname{Im} q_i(\Phi_s(x))\, dm_0(x) = 0. \tag{7-23}$$

**Proof** We just need to show (7-21). Equations (7-22) and (7-23) follow easily, similar to the former cases.

From the computation of $\partial_i g_{\alpha\alpha}(\sigma)$, we know

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0 = 0,$$

$$\int_{UX} \operatorname{Re} q_\beta(x) \operatorname{Re} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0 = 0,$$

$$\int_{UX} \operatorname{Re}(q_\alpha + q_\beta)(x) \operatorname{Re}(q_\alpha + q_\beta)(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0 = 0.$$

We deduce

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0 + \int_{UX} \operatorname{Re} q_\beta(x) \operatorname{Re} q_\alpha(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0 = 0.$$

Similar to $\partial_j g_{\alpha i}(\sigma)$, we consider $s = mt$ for $m \in \mathbb{N}$ and $m \geq 2$. We observe

$$\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_{mt}(x))\, dm_0 = \int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_{(m-1)t}(-x))\, dm_0.$$

We recall the Poincaré disk model and our analytic expansion for $q_\alpha$, $q_\beta$ and $q_i$ in (6-3), (6-5) and (6-7). For $t, s \geq 0$, the analytic expansion

$$\begin{aligned}
&\int_{UX} \operatorname{Re} q_\alpha(x) \operatorname{Re} q_\beta(\Phi_t(x)) \operatorname{Re} q_i(\Phi_s(x))\, dm_0(x) \\
&\quad= \frac{1}{2\pi} \int_0^{2\pi} \int_{UX} \operatorname{Re} q_\alpha(e^{i\theta} x) \operatorname{Re} q_\beta(\Phi_t(e^{i\theta} x)) \operatorname{Re} q_i(\Phi_s(e^{i\theta} x))\, dm_0(x)\, d\theta \\
&\quad= \frac{1}{4} \sum_{n=0}^{\infty} \Big( \int_{UX} \operatorname{Re}(a_0 b_n \bar{c}_{n+4})\, dm_0 T^n (1-T^2)^3 S^{n+4} (1-S^2)^2 \\
&\qquad\qquad + \int_{UX} \operatorname{Re}(a_0 \bar{b}_{n+2} c_n)\, dm_0 S^n (1-S^2)^2 T^{n+2} (1-T^2)^3 \Big).
\end{aligned}$$

Denoting $I_n = \int_{UX} \operatorname{Re}(a_0 b_n \bar{c}_{n+4})\, dm_0$ and $J_n = \int_{UX} \operatorname{Re}(a_0 \bar{b}_{n+2} c_n)\, dm_0$ for $n \geq 0$, we argue

$$I_n = J_n = 0.$$

When $s = mt$, we have $S = S(m) = ((1+T)^m - (1-T)^m)/((1+T)^m + (1-T)^m) = mT + O(T^3)$. The analytic expansions give

$$\begin{aligned}
&\sum_{n=0}^{\infty} \big( I_n T^n S(m)^{n+4} (1-S(m)^2)^2 - I_n e^{in\pi} T^n S(m-1)^{n+4} (1-S(m-1)^2)^2 \big) \\
&\quad= \sum_{n=0}^{\infty} \big( -J_n T^{n+2} S(m)^n (1-S(m)^2)^2 + J_n e^{in\pi} T^{n+2} S(m-1)^n (1-S(m-1)^2)^2 \big).
\end{aligned}$$

The coefficients of $T^4$ yield

$$(m^4 - (m-1)^4) I_0 = -(2m-1) J_1 + (4m-2) J_0.$$

The cases $m = 2$, $m = 3$ and $m = 4$ give $I_0 = J_0 = J_1 = 0$. By induction, assuming $I_k = J_{k+1} = 0$ for $1 \le k < n$, the coefficient of $T^{2n+4}$ gives

$$(m^{n+4} - e^{in\pi}(m-1)^{n+4})I_n = (e^{i(n+1)\pi}(m-1)^{n+1} - m^{n+1})J_{n+1}.$$

We conclude $I_n = J_n = 0$ for $n \ge 0$ by choosing two different $m$. This finishes the proof of (7-21) for $t, s > 0$. Equation (7-21) for $t \le 0$ and $s \le 0$ can be proved similarly to the former cases. Lemma 7.11 and also Proposition 7.6 therefore hold. □

We have shown

(i) $\partial_\beta g_{\alpha\alpha}(\sigma) = 0$,

(ii) $\partial_i g_{\alpha\alpha}(\sigma) = 0$,

(iii) $\partial_j g_{\alpha i}(\sigma) = 0$, and

(iv) $\partial_\beta g_{\alpha i}(\sigma) = 0$.

This finishes the proof of our Theorem 1.1.

# References


[1] **L V Ahlfors**, *Some remarks on Teichmüller's space of Riemann surfaces*, Ann. of Math. 74 (1961) 171–191 MR Zbl

[2] **D Baraglia**, *$G2$ geometry and integrable systems*, preprint (2010) arXiv 1002.1767

[3] **R Bowen**, *Markov partitions for Axiom* A *diffeomorphisms*, Amer. J. Math. 92 (1970) 725–747 MR Zbl

[4] **R Bowen**, *Symbolic dynamics for hyperbolic flows*, Amer. J. Math. 95 (1973) 429–460 MR Zbl

[5] **R Bowen**, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, revised edition, Lecture Notes in Math. 470, Springer (2008) MR Zbl

[6] **R Bowen**, **D Ruelle**, *The ergodic theory of Axiom A flows*, Invent. Math. 29 (1975) 181–202 MR Zbl

[7] **M Bridgeman**, **R Canary**, **F Labourie**, *Simple length rigidity for Hitchin representations*, Adv. Math. 360 (2020) art. id. 106901 MR Zbl

[8] **M Bridgeman**, **R Canary**, **F Labourie**, **A Sambarino**, *The pressure metric for Anosov representations*, Geom. Funct. Anal. 25 (2015) 1089–1179 MR Zbl

[9] **M Bridgeman**, **R Canary**, **F Labourie**, **A Sambarino**, *Simple root flows for Hitchin representations*, Geom. Dedicata 192 (2018) 57–86 MR Zbl

[10] **S Choi**, **W M Goldman**, *Convex real projective structures on closed surfaces are closed*, Proc. Amer. Math. Soc. 118 (1993) 657–661 MR Zbl

[11] **T Chu**, *The Weil–Petersson metric in the moduli space*, Chinese J. Math. 4 (1976) 29–51 MR Zbl

[12] **D Gilbarg**, **N S Trudinger**, *Elliptic partial differential equations of second order*, 2nd edition, Grundl. Math. Wissen. 224, Springer (1983) MR Zbl

[13] **W M Goldman**, *Convex real projective structures on compact surfaces*, J. Differential Geom. 31 (1990) 791–845 MR Zbl

[14] **N J Hitchin**, *The self-duality equations on a Riemann surface*, Proc. London Math. Soc. 55 (1987) 59–126 MR Zbl

[15] **N J Hitchin**, *Lie groups and Teichmüller space*, Topology 31 (1992) 449–473 MR Zbl

[16] **A Katok**, *Entropy and closed geodesics*, Ergodic Theory Dynam. Systems 2 (1982) 339–365 MR Zbl

[17] **A Katok**, **B Hasselblatt**, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications 54, Cambridge Univ. Press (1995) MR Zbl

[18] **M Kotani**, **T Sunada**, *The pressure and higher correlations for an Anosov diffeomorphism*, Ergodic Theory Dynam. Systems 21 (2001) 807–821 MR Zbl

[19] **F Labourie**, *Anosov flows, surface groups and curves in projective space*, Invent. Math. 165 (2006) 51–114 MR Zbl

[20] **F Labourie**, **R Wentworth**, *Variations along the Fuchsian locus*, Ann. Sci. Éc. Norm. Supér. 51 (2018) 487–547 MR Zbl

[21] **Q Li**, *Harmonic maps for Hitchin representations*, Geom. Funct. Anal. 29 (2019) 539–560 MR Zbl

[22] **A N Livšic**, *Certain properties of the homology of Y–systems*, Mat. Zametki 10 (1971) 555–564 MR Zbl In Russian

[23] **A N Livšic**, *Cohomology of dynamical systems*, Izv. Akad. Nauk SSSR Ser. Mat. 36 (1972) 1296–1320 MR Zbl In Russian; translated in Math. USSR-Izv. 6 (1972) 1278–1301

[24] **J Loftin**, *Survey on affine spheres*, from "Handbook of geometric analysis, II" (L Ji, P Li, R Schoen, L Simon, editors), Adv. Lect. Math. 13, International, Somerville, MA (2010) 161–191 MR Zbl

[25] **H Masur**, **M Wolf**, *The Weil–Petersson isometry group*, Geom. Dedicata 93 (2002) 177–190 MR Zbl

[26] **C T McMullen**, *Thermodynamics, dimension and the Weil–Petersson metric*, Invent. Math. 173 (2008) 365–425 MR Zbl

[27] **W Parry**, **M Pollicott**, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque 187–188, Soc. Math. France, Paris (1990) MR Zbl

[28] **M Pollicott**, *On the rate of mixing of Axiom A flows*, Invent. Math. 81 (1985) 413–426 MR Zbl

[29] **M Pollicott**, *Derivatives of topological entropy for Anosov and geodesic flows*, J. Differential Geom. 39 (1994) 457–489 MR Zbl

[30] **M Pollicott**, *Multiple mixing for hyperbolic flows*, preprint (2012) Available at `http://homepages.warwick.ac.uk/~masdbl/multiplemixing-14Dec.pdf`

[31] **R Potrie**, **A Sambarino**, *Eigenvalues and entropy of a Hitchin representation*, Invent. Math. 209 (2017) 885–925 MR Zbl

[32] **M Ratner**, *Markov partitions for Anosov flows on n–dimensional manifolds*, Israel J. Math. 15 (1973) 92–114 MR Zbl

[33] **A Sambarino**, *Quantitative properties of convex representations*, Comment. Math. Helv. 89 (2014) 443–488 MR Zbl

[34] **A Sambarino**, *The orbital counting problem for hyperconvex representations*, Ann. Inst. Fourier (Grenoble) 65 (2015) 1755–1797 MR Zbl

[35] **C T Simpson**, *Constructing variations of Hodge structure using Yang–Mills theory and applications to uniformization*, J. Amer. Math. Soc. 1 (1988) 867–918 MR Zbl

[36] **A J Tromba**, *Teichmüller theory in Riemannian geometry*, Birkhäuser, Basel (1992) MR Zbl

[37] **M Wolf**, *The Teichmüller theory of harmonic maps*, J. Differential Geom. 29 (1989) 449–479 MR Zbl

[38] **S Wolpert**, *Noncompleteness of the Weil–Petersson metric for Teichmüller space*, Pacific J. Math. 61 (1975) 573–577 MR Zbl

[39] **S A Wolpert**, *Geodesic length functions and the Nielsen problem*, J. Differential Geom. 25 (1987) 275–296 MR Zbl

*Heidelberg University*
*Heidelberg, Germany*

`xian.dai@ruhr-uni-bochum.de`

Proposed: Anna Wienhard
Seconded: Paul Seidel, Leonid Polterovich

Received: 24 January 2020
Revised: 15 July 2021

*Geometry & Topology* **27**:4 (2023) 1479–1538
DOI: 10.2140/gt.2023.27.1479
Published: 15 June 2023

# Birational geometry of the intermediate Jacobian fibration of a cubic fourfold

GIULIA SACCÀ
APPENDIX BY CLAIRE VOISIN

We show that the intermediate Jacobian fibration associated to *any* smooth cubic fourfold $X$ admits a hyper-Kähler compactification $J(X)$ with a regular Lagrangian fibration $\pi\colon J \to \mathbb{P}^5$. This builds upon work of Laza, Saccà and Voisin (2017), where the result is proved for general $X$, as well as on the degeneration techniques introduced in the work of Kollár, Laza, Saccà and Voisin, and the minimal model program. We then study some aspects of the birational geometry of $J(X)$: for very general $X$ we compute the movable and nef cones of $J(X)$, showing that $J(X)$ is not birational to the twisted version of the intermediate Jacobian fibration, nor to an OG10–type moduli space of objects in the Kuznetsov component of $X$; for any smooth $X$ we show, using normal functions, that the Mordell–Weil group $\mathrm{MW}(\pi)$ of the fibration is isomorphic to the integral degree-4 primitive algebraic cohomology of $X$, ie $\mathrm{MW}(\pi) \cong H^{2,2}(X,\mathbb{Z})_0$.

14D06, 14J42

| | |
|---|---|
| Introduction | 1480 |
| 1. A hyper-Kähler compactification | 1484 |
| 2. Moduli spaces of OG10–type | 1497 |
| 3. The relative theta divisor on the intermediate Jacobian fibration | 1502 |
| 4. Birational geometry of $J(X)$ for general $X$ | 1510 |
| 5. The Mordell–Weil group of $J(X)$ | 1519 |
| Appendix. On the Beauville conjecture for LSV varieties | 1530 |
| References | 1533 |

## Introduction

The geometry of smooth cubic fourfolds has ties to that of K3 surfaces and, more generally, to that of higher-dimensional hyper-Kähler manifolds. For example, with certain special cubic fourfolds one can associate a K3 surface via Hodge-theoretic (Hassett [34]) or derived categorical (Kuznetsov [43]) methods. From a more geometric perspective, given a smooth cubic fourfold $X$, hyper-Kähler manifolds of $\mathrm{K3}^{[n]}$–type are constructed geometrically, via parameter spaces of rational curves of certain degrees on $X$ (Beauville and Donagi [12] and Lehn, Lehn, Sorger and van Straten [49]), or as moduli spaces of objects in the Kuznetsov component of $X$ (Bayer, Lahoz, Macrì, Nuer, Perry and Stellari [6] and Lahoz, Lehn, Macrì and Stellari [44]). These constructions give rise to 20–dimensional families of polarized hyper-Kähler manifolds, the maximal possible dimension of families of polarized hyper-Kähler manifolds of $\mathrm{K3}^{[n]}$–type. As the cubic fourfold becomes special, for example when it acquires more algebraic classes, the geometry of these hyper-Kähler manifolds also becomes more interesting. For example, when $X$ has an associated K3 surface in the sense of Addington and Thomas [2], Hassett [34], Huybrechts [37] and Kuznetsov [43], these hyper-Kähler manifolds become isomorphic, or birational, to moduli spaces of objects in the derived category of the corresponding K3 surface; see Addington [1] and Bayer, Lahoz, Macrì, Nuer, Perry and Stellari [6].

Laza, Saccà and Voisin [47] constructed a Lagrangian fibered hyper-Kähler manifold starting from a general cubic fourfold. This hyper-Kähler manifold is a deformation of O'Grady's 10–dimensional exceptional example. More precisely, let $X \subset \mathbb{P}^5$ be a smooth cubic fourfold and let $\pi_U \colon J_U \to U \subset (\mathbb{P}^5)^\vee$ be the family of intermediate Jacobians of the smooth hyperplane sections of $X$. This fibration was considered by Donagi and Markman in [25], where they showed that the total space has a holomorphic symplectic form. The main result of [47] was to construct, for general $X$, a smooth projective hyper-Kähler compactification $J$ of $J_U$, with a flat morphism $J \to (\mathbb{P}^5)^\vee$ extending $\pi_U$, and to show that this hyper-Kähler 10–fold is deformation equivalent to O'Grady's 10–dimensional example. In [80], Voisin constructed a hyper-Kähler compactification $J^T$ of a natural $J_U$–torsor $J_U^T$, which is nontrivial for very general $X$. The two hyper-Kähler manifolds $J$ and $J^T$ are birational over countably many hypersurfaces in the moduli space of cubic fourfolds. These two constructions give rise to two 20–dimensional families of hyper-Kähler manifolds of OG10–type, each of which forms an open subset of a codimension-two locus inside the moduli space of hyper-Kähler manifolds in this deformation class.

If one wishes to study the geometry of these hyper-Kähler manifolds as the cubic fourfold becomes special, a first step is to check if a hyper-Kähler compactification of the fibration $J_U \to U$ can be constructed for an arbitrary smooth cubic fourfold. The starting result of this paper is that this can indeed be done.

**Theorem 1** (Theorem 1.6) *Let $X \subset \mathbb{P}^5$ be a smooth cubic fourfold, and let*

$$\pi_U : J_U \to U \subset (\mathbb{P}^5)^\vee$$

*be the Donagi–Markman fibration. There exists a smooth projective hyper-Kähler compactification $J$ of $J_U$ with a morphism $\pi : J \to (\mathbb{P}^5)^\vee$ extending $\pi_U$.*

The same techniques also give the existence of a Lagrangian fibered hyper-Kähler compactification for the nontrivial $J_U$–torsor $J_U^T \to U$ of [80] for any smooth $X$; see Remark 1.14. Moreover, with little extra work, the theorem is proved also for mildly singular cubic fourfolds such as, for example, cubic fourfolds with a simple node; see Proposition 1.17. For a general cubic fourfolds with one node, the existence of such a Lagrangian fibered hyper-Kähler manifold provides a positive answer to a question of Beauville [11]; see Remark 1.18.

We should point out that as a consequence of the "finite monodromy implies smooth filling" results of Kollár, Laza, Saccà and Voisin [41], we prove in Proposition 1.5 that $J_U$ admits projective birational model that is hyper-Kähler. Theorem 1 shows that there exists a hyper-Kähler model with a Lagrangian fibration extending $\pi_U$.

There are several ingredients in the construction of the hyper-Kähler compactification of [47]: a cycle-theoretic construction of the holomorphic symplectic form, the problem of the existence of so-called very good lines for any hyperplane section of $X$, a smoothness criterion for relative compactified Prym varieties, the independence of the compactification from the choice of a very good line. Here we have pursued a different direction, and instead rely on the existence of a hyper-Kähler compactification for general $X$, use the degeneration techniques introduced in [41], and implement some results from birational geometry and the minimal model program, following Kollár [40] and Lai [45]. One advantage of our method is that it opens the door to using birational geometry to compactify Lagrangian fibrations.

The second result of this paper is concerned with the hyper-Kähler birational geometry of $J$. We show that the relative theta divisor $\Theta$ of the fibration is a prime exceptional divisor and that for general $X$ it can be contracted after a Mukai flop.

**Theorem 2** (Theorem 4.1) *Let $q$ be the Beauville–Bogomolov form on $H^2(J,\mathbb{Z})$. The relative theta divisor $\Theta \subset J$ is a prime exceptional divisor with $q(\Theta) = -2$. For very general $X$, there is a unique other hyper-Kähler birational model of $J$, denoted by $N$, which is the Mukai flop $p\colon J \dashrightarrow N$ of $J$ along the image of the zero section. $N$ admits a divisorial contraction $h\colon N \to \overline{N}$, which contracts the proper transform of $\Theta$ onto an 8–dimensional variety which is birational to the LLSv 8–fold $Z(X)$.*

Thus, for very general $X$, $J$ is the unique hyper-Kähler birational model with a Lagrangian fibration, it is not birational to $J^T$ (Corollary 3.10), and its movable cone is the union of its nef cone and the nef cone of $N$. This answers a question by Voisin [80]. As a consequence of this theorem we show that for very general $X$, $J$ is not birational to a moduli spaces of objects in the Kuznetsov component $\mathcal{K}u(X)$ of $X$; see Corollary 4.2. In the opposite direction, it was recently proved by Li, Pertusi and Zhao [51] that the twisted hyper-Kähler manifold $J^T$ is birational to a moduli space of objects of OG10–type in $\mathcal{K}u(X)$. By objects of OG10–type, we mean objects whose Mukai vector is of the form $2w$, with $w^2 = 2$. As a consequence, the family of intermediate Jacobian fibrations is the only known family of hyper-Kähler manifolds associated with cubic fourfolds whose very general point cannot be described as a moduli space of objects in the Kuznetsov component of $X$.

Given $J = J(X)$, a hyper-Kähler compactification of the intermediate Jacobian fibration for any smooth cubic fourfold $X$, a natural question to ask is how the geometry of $J$ changes as $X$ becomes less general. One way to answer this question is the following theorem, describing the Mordell–Weil group of $\pi$ in terms of the primitive algebraic cohomology of $X$. In Section 5 we prove:

**Theorem 3** (Theorem 5.1) *Let $\mathrm{MW}(\pi)$ be the Mordell–Weil group of $\pi\colon J \to \mathbb{P}^5$, ie the group of rational sections of $\pi$, and let $H^{2,2}(X,\mathbb{Z})_0$ be the primitive degree-4 integral cohomology of $X$. The natural group homomorphism*

$$\phi_X\colon H^{2,2}(X,\mathbb{Z})_0 \to \mathrm{MW}(\pi)$$

*induced by the Abel–Jacobi map is an isomorphism.*

The proof of this result uses the theory of normal functions, as developed by Griffiths and Zucker, as well as the techniques used by Voisin to prove the integral Hodge conjecture for cubic fourfolds. A consequence of this is a geometric description of the Lagrangian fibered hyper-Kähler manifolds with maximal Mordell–Weil rank, whose existence was proved by Oguiso in [64]: indeed, Oguiso's examples are (birationally) given by $J = J(X) \to \mathbb{P}^5$, where $X$ is a smooth cubic fourfold with $H^{2,2}(X, \mathbb{Z})$ of maximal rank.

## Plan of the paper

In Section 1 we prove the existence of a hyper-Kähler compactification for $J_U$ and for $J_U^T$, in the case of any smooth, or mildly singular, $X$. This uses some results from the minimal model program, which are briefly recalled. In Section 2 we review some basic results about moduli spaces of OG10–type and we compute, using the Bayer–Macrì techniques adapted to these singular moduli spaces by Meachan and Zhang [58], the nef and movable cones of certain moduli spaces of OG10–type that appear as limits of the intermediate Jacobian fibration. The main result of Section 3 is the computation that $q(\Theta) = -2$. Section 4 is devoted to the proof of Theorem 2 and its preparation: Given a family of cubic fourfolds degenerating to the chordal cubic, we construct a certain degeneration of the intermediate Jacobian fibration and identify the limit of the corresponding degeneration of the relative Theta divisor. By the results of Section 2, the limiting theta divisor can be contracted after a Mukai flop of the zero section and we deduce the analogous result for $\Theta$. The computation of the Mordell–Weil group occupies Section 5.

Finally, in the appendix by C Voisin, some applications to the Beauville conjecture on the polynomial relations in the Chow group of a projective hyper-Kähler manifold are given for $J = J(X)$, in the case of very general $J$ of Picard number 2 or 3. This is obtained as an application of the computation of $q(\Theta) = -2$ from Theorem 2.

## Acknowledgements

I would like to thank J Kollár for pointing my attention to the techniques of [40; 45], which are used in this paper. It is my pleasure to thank E Arbarello, C Camere, G Di Cerbo, K Hulek, R Laza, E Macrì, C Onorati, G Pearlstein, A Rapagnetta, L Tasin and C Voisin for useful and interesting discussions related to the topic of this paper. I also thank the anonymous referee for having read this paper very carefully and for many useful comments. Finally, I would like to warmly thank Collège de France and École Normal Supérieure for the hospitality and the great working conditions while the final version of this manuscript was being prepared. This work is partially supported by NSF grant DMS-1801818.

# 1 A hyper-Kähler compactification of the intermediate Jacobian fibration for any smooth cubic fourfold

We denote by $X \subset \mathbb{P}^5$ a smooth cubic fourfold, by $(\mathbb{P}^5)^\vee$ the dual projective space parametrizing hyperplane sections $Y = X \cap H \subset X$, and by $U \subset (\mathbb{P}^5)^\vee$ the open subset parametrizing smooth hyperplane sections. The dual hypersurface of $X$, parametrizing singular hyperplane sections, is denoted by $X^\vee \subset (\mathbb{P}^5)^\vee$. Its smooth locus

$$U_1 := (\mathbb{P}^5)^\vee \setminus \mathrm{Sing}(X^\vee) \subset (\mathbb{P}^5)^\vee$$

parametrizes hyperplane sections of $X$ that are smooth or have one simple node and no other singularities. In what follows, we freely drop the $^\vee$ from $(\mathbb{P}^5)^\vee$ and write simply $\mathbb{P}^5$. From the context it will be clear if we are referring to the projective space parametrizing hyperplane sections of $X$ or the projective space containing $X$. For a smooth cubic threefold $Y$, the Griffiths intermediate Jacobian of $Y$ will be denoted by

$$\mathrm{Jac}(Y) \cong H^1(Y, \Omega^2_Y)^\vee / H_3(Y, \mathbb{Z}).$$

It is a principally polarized abelian fivefold which parametrizes rational equivalence classes of homologically trivial 1–cycles on $Y$ [79, Theorem 6.24].

Over $U$ consider the Donagi–Markman fibration

$$\pi_U \colon J_U = J_U(X) \to U, \tag{1-1}$$

whose fiber over a smooth hyperplane section $Y = X \cap H$ is the intermediate Jacobian $\mathrm{Jac}(Y)$. By [25], $J_U$ is quasiprojective and admits a holomorphic symplectic form $\sigma_{J_U}$, with respect to which $\pi_U$ is Lagrangian. The main result of [47] is the following theorem.

**Theorem 1.1** [47] *Let $X$ be a general cubic fourfold. Then there exists a smooth projective compactification $J = J(X)$ of $J_U$, with a flat morphism $\pi \colon J \to (\mathbb{P}^5)^\vee$ extending $\pi_U$, which has irreducible fibers and which admits a rational zero section $s \colon (\mathbb{P}^5)^\vee \dashrightarrow J$. Moreover, $J$ is an irreducible holomorphic symplectic manifold, deformation equivalent to O'Grady's 10–dimensional exceptional example.*

We will say that $X$ is general *in the sense of LSV* if the construction of [47] works for $J_U(X)$, and we refer to $J = J(X)$ as in Theorem 1.1 as the *LSV fibration.* A necessary condition for this to happen is that the hyperplane sections of $X$ are palindromic; see [17]. For example, a cubic fourfold containing a plane is not general in the sense of LSV.

To extend the theorem above for any $X$, we use the existence of a hyper-Kähler compactification for general $X$, the cycle-theoretic description of the holomorphic symplectic form that was given in [47], the degeneration results from [41], and techniques from the minimal model program, following [40; 45]. We start by recalling the construction of a natural partial compactification of $J_U$, which already appeared in [25; 47].

**Lemma 1.2** [25; 47] *For any smooth $X$, there is a canonical partial compactification $J_{U_1} = J_{U_1}(X)$ of $J_U$, with a projective morphism $\pi_{U_1}\colon J_{U_1} \to U_1$ with irreducible fibers extending $\pi_U$. This $J_{U_1}$ is smooth and has a holomorphic symplectic form $\sigma_{J_{U_1}}$ extending $\sigma_{J_U}$.*

**Proof** This is already proved in [25, Section 8.5.2 and Theorem 8.18]. Alternatively, one can use [23, Corollary 2.38], and [47, Definitions 2.2 and 2.9, Proposition 1.4 and Lemma 5.2]. □

Before giving an application of the cycle-theoretic construction of the holomorphic symplectic form [47, Section 1], we recall the definition of symplectic variety.

**Definition 1.3** A normal projective variety $M$ is called symplectic if its smooth locus carries a holomorphic symplectic form which extends to a regular (ie holomorphic) form on any resolution of singularities of $M$.

**Lemma 1.4** *Let $\overline{J}$ be a normal projective compactification of $J_U$. Then*:

(1) *The smooth locus of $\overline{J}$ admits a holomorphic two-form extending $\sigma_{J_U}$. In particular, the canonical class $K_{\overline{J}}$ of $\overline{J}$ is effective and is trivial if and only if $\overline{J}$ is a symplectic variety.*

(2) *$\overline{J}$ is not uniruled.*

**Proof** (1) The first statement is [47, Theorem 1.2(iii)], while the second follows from the fact that the canonical class of $\overline{J}$ is the (closure of the) codimension-one locus where the generically nondegenerate holomorphic two-form is degenerate.

(2) Let $\widetilde{J} \to \overline{J}$ be a resolution of singularities. By (1), $\widetilde{J}$ has effective canonical class and thus by [59] it is not uniruled. □

The following is an application of the degeneration techniques of [41].

**Proposition 1.5** *Let $X$ be a smooth cubic fourfold and let $J_U = J_U(X)$ be as above. Then there exists a smooth projective hyper-Kähler manifold $M$ birational to $J_U$ and of $OG10$–type.*

**Proof** Let $\mathcal{X} \to \Delta$ be a family of smooth cubic fourfolds with $\mathcal{X}_0 = X$. Here $\Delta$ is an open affine subset of a smooth projective curve, or a small disk. We will use the notation $t = 0$ to denote a chosen special point in $\Delta$, and $t \neq 0$ to denote any other point. Up to restricting $\Delta$ if necessary, assume that for $t \neq 0$, $\mathcal{X}_t$ is general in the sense of LSV. By [47, Proposition 2.10], we can assume that for any $t \neq 0$ all the hyperplane sections of $\mathcal{X}_t$ admit a very good line; see [47, Definition 2.9]. Consider the open set $\mathcal{V} = (\mathbb{P}^5)^\vee \times \Delta \setminus \mathrm{Sing}(\mathcal{X}_0^\vee) \times \{0\}$, so that $\mathcal{V}_t = (\mathbb{P}^5)^\vee$ for $t \neq 0$ and $\mathcal{V}_0 = U_1 \times \{0\}$ parametrizes the hyperplane sections of $\mathcal{X}_0 = X$ that have at most one nodal point and no other singularities. The construction of [47, Section 5] can be carried out in families, yielding a projective morphism

$$\mathcal{J}_\mathcal{V} \to \mathcal{V},$$

which is fibered in compactified Prym varieties and is such that, denoting by $\mathcal{J}_t$ the fiber of the induced smooth quasiprojective morphism $\mathcal{J}_\mathcal{V} \to \Delta$ for $t \neq 0$, $\mathcal{J}_t$ is the LSV fibration $J(\mathcal{X}_t)$, and $\mathcal{J}_0 = J_{U_1}(X)$. Let $\widetilde{\mathcal{J}} \to \Delta$ be a projective morphism extending $\mathcal{J}_\mathcal{V} \to \Delta$. The central fiber $\mathcal{J}_0$ has a multiplicity-one component which contains $J_{U_1}$ as dense open subset. By Lemma 1.4, this component is not uniruled. By [41, Corollary 5.2] there is a birational model $M$ of $J_{U_1}(X)$ that is a hyper-Kähler manifold, deformation equivalent to the smooth fibers $\mathcal{J}_t = J(\mathcal{X}_t)$, for $t \neq 0$. □

By [57], given a hyper-Kähler manifold $M$ with a Lagrangian fibration $\pi \colon M \to \mathbb{P}^n$, the locus inside $\mathrm{Def}(M)$ where the Lagrangian fibration deforms is an open subset of the hypersurface where the class $\pi^*\mathcal{O}(1)$ stays of type $(1,1)$. However, this fact alone is not enough to imply the existence of a hyper-Kähler compactification of $J_{U_1}$ for *any* smooth $X$.

This is what we prove in the following theorem, whose proof uses the mmp following Kollár [40, Section 8] and Lai [45]. In Section 1.1 we will recall some basic facts about the mmp that are needed in the proof of Theorems 1.6 and 1.19. We refer to [42] and to [32] for the basic definitions and fundamental results.

**Theorem 1.6** *For any smooth cubic fourfold $X$, there exists a smooth projective hyper-Kähler compactification $J = J(X)$ of $J_U(X)$, with a projective flat morphism $\pi \colon J \to \mathbb{P}^5$ extending $\pi_U$.*

**Proof** Let $\overline{J} \to \mathbb{P}^5$ be any normal projective compactification of $J_{U_1}$ with a regular morphism $\overline{\pi}\colon \overline{J} \to \mathbb{P}^5$. By Lemma 1.4, there is a holomorphic two-form $\overline{\sigma}$ on the smooth locus of $\overline{J}$ extending $\sigma_{J_{U_1}}$, the canonical class $K_{\overline{J}} \geq 0$ is effective, and $K_{\overline{J}} = 0$ if and only if $\overline{J}$ is a symplectic variety. Since $K_{\overline{J}}$ is supported on the complement of $J_{U_1}$, $\operatorname{codim} \overline{\pi}(\operatorname{Supp}(K_{\overline{J}})) \geq 2$. By definition [40, Definition 7], this means that $K_{\overline{J}}$ is $\overline{\pi}$–exceptional, if it is nontrivial. If this is the case, then by [61, III 5.1] (see also [45, Lemma 2.10]), $K_{\overline{J}}$ is not $\overline{\pi}$–nef. More precisely, there is a component of $K_{\overline{J}}$ that is covered by curves that are contracted by $\overline{\pi}$ and that intersect $K_{\overline{J}}$ negatively.

Let $\widetilde{J} \to \mathbb{P}^5$ be a smooth projective compactification of $J_{U_1}$ admitting a regular morphism $\widetilde{\pi}\colon \widetilde{J} \to \mathbb{P}^5$, and let $K_{\widetilde{J}}$ be its canonical class. If the effective divisor $K_{\widetilde{J}}$ is not trivial, we use the mmp to contract $\operatorname{Supp}(K_{\widetilde{J}})$ relatively to $\mathbb{P}^5$. Let $H$ be a $\widetilde{\pi}$–ample $\mathbb{Q}$–divisor such that the pair $(\widetilde{J}, H)$ is klt and $K_{\widetilde{J}} + H$ is relatively big and nef. The mmp with scaling over $\mathbb{P}^5$ (see Section 1.1 below) produces a sequence of birational maps

$$\widetilde{J} = J_0 \overset{\psi_0}{\dashrightarrow} J_1 \overset{\psi_1}{\dashrightarrow} \cdots \dashrightarrow J_i \overset{\psi_i}{\dashrightarrow} \cdots \tag{1-2}$$

over $\mathbb{P}^5$ — ie there are projective morphisms $\pi_i\colon J \to \mathbb{P}^5$ such that $\pi_0 = \widetilde{\pi}$ and $\pi_i := \pi_{i-1} \circ \psi_i^{-1}$ — and a nonincreasing sequence of nonnegative rational numbers $t_0 = 1 \geq t_1 \geq \dots t_i \geq \cdots \geq 0$, with the following properties:

(1) For every $i \geq 0$, $K_{J_i} + t_i H_i$ is $\pi_i$–big and $\pi_i$–nef.

(2) For every $i \geq 0$, $J_i$ is a $\mathbb{Q}$–factorial terminal compactification of $J_{U_1}$. The fact that the birational morphisms $\psi_i$ are isomorphisms away from $J_{U_1}$ follows from the fact that the $K_{J_i}$–negative rays of the mmp correspond to rational curves that are contained in the support of $K_{J_i}$. Thus, by Lemma 1.4, the smooth locus of $J_i$ carries a holomorphic two-form $\sigma_i$ extending $\sigma_{J_{U_1}}$.

(3) $K_{J_i}$ is effective and, if not trivial, it has a component covered by $K_{J_i}$–negative curves which are contracted by $\pi_i$.

(4) The process stops if and only if there exists an $i$ such that $K_{J_i}$ is $\pi_i$–nef. This holds if and only if $K_{J_i} = 0$.

The number of irreducible components of the support of $K_{J_i}$ is nonincreasing, since the birational maps of the mmp extract no divisors. In fact, we claim that this number is eventually strictly decreasing. By (4) above, this happens if and only if the process eventually stops. Suppose that this is not the case. Then by Lemma 1.13, $\lim t_i = 0$. Recall, as already observed, that if $K_{J_k} \neq 0$, then there exists a component that is covered by $K_{J_k}$–negative curves that are contracted by $\pi_k$. Since we are assuming that $\lim t_i = 0$, this implies that for $i \gg 0$, $t_i$ is small enough that this component is contained in the relative stable base locus $\mathbb{B}((K_{J_k} + t_i H_k)/\mathbb{P}^5)$. Since by Lemma 1.12, the divisorial components of $\mathbb{B}((K_{J_k} + t_i H_k)/\mathbb{P}^5)$ are contracted by $J_k \dashrightarrow J_i$, it follows that for $i \gg 0$, the number of irreducible components of the effective divisor $K_{J_i}$ is strictly less than the number of components of $K_{J_k}$. Thus, the claim is proved and for some $i \gg 0$, the process gives a model with $K_{J_i} = 0$. By Lemma 1.4, $\bar{J} := J_i$ is a $\mathbb{Q}$–factorial terminal symplectic compactification of $J_{U_1}$. Finally, by Proposition 1.7 below, $\bar{J}$ is smooth and the theorem is proved. □

**Proposition 1.7** (Greb–Lehn–Rollenske) *Let $\bar{M}$ be a $\mathbb{Q}$–factorial terminal symplectic variety. Suppose that $\bar{M}$ is birational to a smooth hyper-Kähler manifold $M$. Then $\bar{M}$ is smooth.*

**Proof** This is [29, Proposition 6.5]. □

**Remark 1.8** The techniques used to prove the theorem above can be applied to similar contexts to give $\mathbb{Q}$–factorial terminal symplectic compactifications of other quasiprojective Lagrangian fibrations. We plan to come back to this in upcoming work.

As a consequence of Theorem 1.19 below, we will give a slightly stronger version of the theorem just proved (see Remark 1.20) showing that, given a family of smooth cubic fourfolds whose general fiber is general in the sense of [47], then up to a base change and birational transformations, the corresponding family of LSV intermediate Jacobian fibrations can be filled with a Lagrangian fibered smooth projective hyper-Kähler compactification of the Donagi–Markman fibration of the limiting cubic fourfold.

Another approach to Theorem 1 would be to show that the rational map $M \dashrightarrow \mathbb{P}^5$ induced by the birational map $\phi\colon M \dashrightarrow J_{U_1}$ of Proposition 1.5 is almost holomorphic [56, Definition 1]. By [56] this would imply the existence of a birational hyper-Kähler model of $M$ with a regular morphism to $\mathbb{P}^5$. It seems, however, that controlling the mmp of Proposition 1.5 to ensure that $M \dashrightarrow \mathbb{P}^5$ is almost holomorphic is not too far from running the relative mmp as in the proof of Theorem 1.6.

Given a smooth cubic fourfold $X$, we will refer to both the Donagi–Markman fibration $J_U$ and to any hyper-Kähler compactification $J$ of $J_U$ as in Theorem 1.6, as the intermediate Jacobian fibration. Hopefully, it will be clear from the context which one we are referring to.

**Remark 1.9** Unlike the compactification of [47], the proof of Theorem 1.6 is not constructive and, for a given $X$, the hyper-Kähler compactification that we show to exist may not be unique. We will return to this question in Section 4.

## 1.1 The mmp with scaling

In this subsection we recall some basic tools and known results from the minimal model program (mmp) that are used to prove Theorems 1.6 and 1.19. For the basic notions and the fundamental results we refer to [42] and [32]. In this section, by divisor we will mean a $\mathbb{Q}$–divisor.

Let $M$ be a normal $\mathbb{Q}$–factorial variety with a projective morphism $\pi\colon M \to B$ to a normal quasiprojective variety $B$. Let $\Delta$ be an effective divisor on $M$ and let $H$ be a general divisor on $M$ that is ample (or big) over $B$. We assume that the pair $(M, \Delta + H)$ is klt and that $K_M + \Delta + H$ is nef over $B$.

The mmp with scaling of $H$ [32, Section 5.E] produces a sequence of birational maps $\psi_i\colon M_i \dashrightarrow M_{i+1}$ over $B$, such that $M_0 = M$, $\Delta_{i+1} = (\psi_i)_*\Delta_i$, $H_{i+1} = (\psi_i)_*Hi$ and $\psi_i$ is the flip or the divisorial contraction for a $(K_{M_i} + \Delta_i)$–negative relative extremal ray $R_i$ over $B$. We let $\pi_i$ be the induced regular morphism $M_i \to B$. The sequence is defined inductively in the following way. Let

$$t_i = \inf\{t \geq 0 \mid K_{M_i} + \Delta_i + tH_i \text{ is nef over } B\}.$$

If $t_i = 0$, then $K_{M_i} + \Delta_i$ is nef over $B$ and the process stops. Otherwise, there is a $0 < t' \leq t_i$ such that $K_{M_i} + \Delta_i + t'H_i$ is not nef over $B$. By the cone theorem (see [42, Chapter 3] or [32, Theorem 5.4]) $K_{M_i} + \Delta_i + t_i H_i$ is nef over $B$ and there exists a $(K_{M_i} + \Delta_i)$–negative extremal ray $R_i$ over $B$ such that $(K_{M_i} + \Delta_i + t_i H_i) \cdot R_i = 0$.

Let $c_i\colon M_i \to Z_i$ be the extremal contraction over $B$ associated to $R_i$, which exists by the "contraction" part of the cone theorem [32, (5.4.3)–(5.4.4)]. If $\dim Z_i < \dim M_i$, then $c_i$ is a Mori fiber space and we stop. If $c_i$ is not a Mori fiber space then it is either a divisorial or flipping contraction. In the first case, we let $M_{i+1} = Z_i$ and $\psi_i = c_i$. In second case, we let $\psi_i\colon M_i \dashrightarrow M_{i+1}$ be the $(K_{M_i} + \Delta_i + t'H_i)$–flip (which exists by [32, Corollary 5.73]). By construction, $\psi_i$ extracts no divisors, meaning that $\psi_i^{-1}$ contracts no divisors.

By the contraction part of the cone theorem, the divisor $K_{M_{i+1}} + \Delta_{i+1} + t_i H_{i+1}$ is nef over $B$. The pair $(M_i, \Delta_{i+1} + t_i H_{i+1})$ is klt (see [42, Corollaries 3.42–3.44]) and $M_i$ is $\mathbb{Q}$–factorial (see [42, Corollary 3.18]). If $\Delta = 0$ and $M$ is terminal, then so is $M_i$. As long as $K_{M_i} + \Delta_i$ is not $\pi_i$–nef, $t_{i+1}$ is nonzero and $\Delta_{i+1} + t_i H_{i+1}$ is big over $B$. Thus we can keep going, producing a nonincreasing sequence $t_i \geq t_{i+1} \geq \cdots$ of nonnegative rational numbers and a sequence of birational maps $\psi_i \colon M_i \dashrightarrow M_{i+1}$ over $B$. The process stops if there exists an $N$ such that $c_N \colon M_N \to Z_N$ is a Mori fiber space over $B$ or such that $K_{M_N} + \Delta_N$ is nef over $B$. Otherwise, the sequence is infinite.

The pair $(M_i, \Delta_i + t_i H_i)$ is a log terminal model (ltm) for $(M, \Delta + t_i H)$ over $B$; see Definition 5.29 and Lemma 5.31 of [32]. We will need the following lemmas:

**Lemma 1.10** *For any $i > j$, let $\psi_{ij} \colon M_j \dashrightarrow M_i$ be the induced birational morphism over $B$. Then $\psi_{ij}$ is not an isomorphism.*

**Proof** This is [32, Lemma 5.62]. □

**Lemma 1.11** [32, Exercise 5.10] *Let $(M, \Delta)$ be a klt pair as above and suppose that $\Delta$ is big over $B$ and that $K_M + \Delta$ is nef over $B$. Then $K_M + \Delta$ is semiample over $B$, ie there exists a projective morphism $f \colon M \to Z$ over $B$ and an ample divisor $L$ on $B$ such that $K_M + \Delta \sim_{\mathbb{Q},B} f^* L$.*

**Proof** Since $\Delta$ is big over $B$, we can write $\Delta \sim_{\mathbb{Q},B} A + C$, where $A$ is ample over $B$ and $C \geq 0$. Choose an $0 < \epsilon \ll 1$ such that $(M, \Delta')$ is klt, where $\Delta' = (1-\epsilon)\Delta + \epsilon C$. Then

$$(K_M + \Delta) - (K_M + \Delta') = \epsilon A$$

is ample over $B$. By the basepoint-free theorem (see eg [32, Theorem 5.1]), $K_M + \Delta$ is semiample over $B$. □

**Lemma 1.12** *Let the notation be as above and for any $i > 0$, let $\phi_i \colon M \dashrightarrow M_i$ be the induced birational map over $B$. Then the divisors contracted by $\phi_i$ are the divisorial components of $\mathbb{B}((K_{M_i} + \Delta_i + t_i H_i)/B)$, the stable base locus over $B$; cf* [32, Section 2.E]. *Similarly, $\psi_{ij} \colon M_j \dashrightarrow M_i$ contracts the divisorial components of $\mathbb{B}((K_{M_j} + \Delta_j + t_i H_j)/B)$.*

**Proof** Since $(M_i, \Delta_i + t_i H_i)$ is klt, $\Delta_i + t_i H_i$ is big over $B$, and $K_{M_i} + \Delta_i + t_i H_i$ is nef over $B$, by the lemma above, $K_{M_i} + \Delta_i + t_i H_i$ is semiample over $B$.

Let $W$ be a smooth birational model resolving $\phi_i$, and let $p$ and $q$ be the induced birational morphisms to $M$ and $M_i$. By [32, Lemma 5.31] the pair $(M_i, \Delta_i + t_i H_i)$ is a log terminal model for $(M, \Delta + t_i H)$ over $B$; see [32, Definition 5.29]. Thus,

$$p^*(K_M + \Delta + t_i H) = q^*(K_{M_i} + \Delta_i + t_i H_i) + E, \tag{1-3}$$

where

$$E = \sum_F (a(F; M, \Delta + t_i H) - a(F; M_i, \Delta_i + t_i H))F$$

is an effective $q$–exceptional divisor whose support contains the divisors contracted by $\phi_i$. Since

$$\begin{aligned} p^{-1}\mathbb{B}((K_M + \Delta + t_i H)/B) &= \mathbb{B}(p^*(K_M + \Delta + t_i H)/B) \\ &= \mathbb{B}(q^*(K_{M_i} + \Delta_i + t_i H_i) + E/B) \\ &= \operatorname{Supp}(E), \end{aligned}$$

the first statement follows. The second statement is proved in the same way, since by [32, Lemma 5.31], the pair $(M_i, \Delta_i + t_i H_i)$ is a log terminal model for $(M_j, \Delta_j + t_i H_j)$ over $B$ and hence the equivalent of (1-3) holds. □

**Lemma 1.13** *Let the notation be as above. If the mmp with scaling does not terminate, then*

$$\lim_{i \to \infty} t_i = 0.$$

**Proof** This is [26, Proposition 3.2]. The only difference is the relative setting, but the proof is the same: Suppose the mmp does not terminate and that $\lim t_i = t_\infty > 0$. By [13, Theorem E] there are finitely many log terminal models of $(M, \Delta + (t_\infty + t)H)$, with $t \in [0, 1 - t_\infty]$. We have already observed that $(M_i, \Delta_i + t_i H_i)$ is an ltm for $(M, \Delta + t_i H) = (M, \Delta + t_\infty H + (t_i - t_\infty)H))$ over $B$. Thus, if the sequence is infinite there are integers $i > j$ such that the birational map $M_j \dashrightarrow M_i$ is an isomorphism. This gives a contradiction with Lemma 1.10 above. □

## 1.2 Variants

In this section we give some variants of the results of the previous section. First we notice that the compactification result of Theorem 1.6 holds also for the twisted intermediate Jacobian fibration; see Remark 1.14. Then we consider the case of the intermediate Jacobian fibration associated to a mildly singular cubic fourfold; see Proposition 1.15 and Remark 1.16. We then give a slightly stronger version of Theorem 1.6, in that we show that the Lagrangian fibered hyper-Kähler compactification works in families; see Proposition 1.17 and Theorem 1.19. As an application, we give a positive answer to a question of Beauville; see Remark 1.18.

**Remark 1.14** (the twisted case) In [80], Voisin constructed a nontrivial $J_U$–torsor $J_U^T \to U$ defined from a class in $H^1(U, \mathcal{J}_U[3])$, where $\mathcal{J}_U$ is the sheaf of holomorphic sections of $J_U \to U$ and where $\mathcal{J}_U[3] \subset \mathcal{J}_U$ is the sheaf of 3–torsion points. The nontriviality (for very general $X$) of this class corresponds to the nonexistence, for the universal family of hyperplanes sections of $X$, of a relative one-cycle of degree one. The main result of the paper is to produce, for general $X$, a hyper-Kähler compactification $J^T = J^T(X)$ with Lagrangian fibration to $\mathbb{P}^5$ extending $J_U^T \to U$. This builds on the compactification of [47]. We will refer to this hyper-Kähler manifold as the twisted intermediate Jacobian fibration. This hyper-Kähler manifold is deformation equivalent to the nontwisted version $J(X)$, as they agree as soon as $X$ has a two-cycle which restricts to a one-cycle of degree one or two on its hyperplane sections. Lemma 1.4, Proposition 1.5 and Theorem 1.6 work the same for the nontrivial torsor $J_U^T \to U$, giving a Lagrangian fibered hyper-Kähler $J^T = J^T(X)$ for every smooth $X$. In Section 4.1 we will return to the twisted intermediate Jacobian fibration and in Corollary 3.10 we prove that for very general $X$ these two fibrations are not birational and that on $J$ there is a unique isotropic class in the movable cone of $J$. This fact will be used in the appendix.

Finally, we show that the Lagrangian fibered hyper-Kähler compactification exists generically also over $\mathcal{C}_6$, the divisor in the moduli space of cubic fourfolds whose general point parametrizes cubics with one $A_1$ singularity. The following proposition is an adaptation of [47, Section 2] to the case of a cubic fourfold with mild singularities.

**Proposition 1.15** *Let $X_0 \subset \mathbb{P}^5$ be a cubic fourfold with one simple node $o \in X_0$ and no other singularities. Let $U \subset \mathbb{P}^5$ be the open locus parametrizing smooth hyperplane sections, and let $\pi_U \colon J_U = J_U(X_0) \to U$ be the Donagi–Markman fibration. Then there exists a holomorphic symplectic form $\sigma_U$ on $J_U$, which extends to a holomorphic two-form on any smooth projective compactification. As a consequence, Lemma 1.4 holds for $J_U$, namely any projective compactification of $J_U$ has smooth locus admitting a generically nondegenerate holomorphic two-form extending $\sigma_U$, and is not uniruled. Similarly, for the twisted intermediate Jacobian, $J_U^T = J_U^T(X_0)$.*

**Proof** Let $\widetilde{X}_0$ (resp. $\widetilde{\mathbb{P}}^5$) be the blowup of $X_0$ (resp. $\mathbb{P}^5$) at the point $o$. Let $E \subset \widetilde{\mathbb{P}}^5$ be the exceptional divisor. Projection from $o$ determines an isomorphism $\widetilde{X}_0 \cong BL_S\mathbb{P}^4$, where $S$ is the $(2,3)$ complete intersection in $\mathbb{P}^3$ parametrizing lines in $X_0$ by $o$. The surface $S$ is a smooth K3 surface and thus $H^1(\widetilde{X}_0, \Omega^3_{\widetilde{X}_0})$ is one-dimensional; let $\eta$ be a generator. The same argument as in [47, Theorem 1.2 ] shows that $\eta$ induces a holomorphic two-form $\sigma$ on $J_U$, with respect to which the fibers of $J_U \to U$ are isotropic. To show that $\sigma$ is nondegenerate, it suffices to show that for any smooth hyperplane section $Y$ (which in particular does not pass by the point $o$), the map

$$T_{[Y]}U = H^0(Y, \mathcal{O}_Y(1)) \to H^1(Y, \Omega^2_Y) = H^0(J_U, \Omega^1_{J_U}) \tag{1-4}$$

induced by $\sigma$, via the fact that the fibers of $J_U \to U$ are isotropic, is an isomorphism. By [47, Theorem 1.2(ii)], this map is given by the cup product with a class $\eta_Y \in H^1(Y, \Omega^2_Y(-1))$ defined in the following way: let $\eta_{|Y} \in H^1(Y, (\Omega^3_{\widetilde{X}_0})_{|Y})$ be the restriction of $\eta$ to $Y$. Since $H^1(Y, \Omega^3_Y) = 0$, the exact sequence

$$0 \to \Omega^2_Y(-1) \to (\Omega^3_{\widetilde{X}_0})_{|Y} \to \Omega^3_Y \to 0$$

implies that $\eta_{|Y}$ lifts to a class $\eta_Y \in H^1(\Omega^2_Y(-1))$. By Griffiths residue theory [47, Lemma 1.7], $H^1(\Omega^2_Y(-1))$ is one-dimensional and cup product with any nonzero element induces an isomorphism $H^0(Y, \mathcal{O}_Y(1)) \to H^1(Y, \Omega^2_Y)$; more precisely, using the canonical isomorphism $\Omega^2_Y(-1) = T_Y(3)$, this space is spanned by the class of the nontrivial extension $0 \to T_Y \to (T_{\mathbb{P}^4})_{|Y} \to \mathcal{O}_Y(-3) \to 0$. It follows that to show that (1-4) is an isomorphism, we only need to show that $\eta_Y \neq 0$, which amounts to showing that $\eta_{|Y} \neq 0$. Under the isomorphism $\Omega^3_{\widetilde{X}_0} = T_{\widetilde{X}_0}(-3)(2E)$, the class of a generator of $H^1(\widetilde{X}_0, \Omega^3_{\widetilde{X}_0})$ corresponds to the class of the extension $0 \to T_{\widetilde{X}} \to (T_{\widetilde{\mathbb{P}}^5})_{|\widetilde{X}} \to \mathcal{O}_{\widetilde{X}}(3)(-2E) \to 0$. Restricting to $Y$ and considering the tangent bundle sequence for $Y$ in $\mathbb{P}^4$, we get the diagram of short exact sequences

$$\begin{tikzcd}
0 \arrow[r] & (T_{\widetilde{X}})_{|Y} \arrow[r] & (T_{\widetilde{\mathbb{P}}^5})_{|Y} \arrow[r] & \mathcal{O}_Y(3) \arrow[r] & 0 \\
0 \arrow[r] & T_Y \arrow[r] \arrow[u] & (T_{\mathbb{P}^4})_{|Y} \arrow[r] \arrow[u, "\alpha"] & \mathcal{O}_Y(3) \arrow[r] \arrow[u, equal] & 0
\end{tikzcd}$$

where the first two vertical arrows are injective. The extension class of the first row is $\eta_{|Y}$ and the second row is nonsplit, as we already observed. Since $\operatorname{coker}(\alpha) = \mathcal{O}_Y(1)$, we have $\operatorname{Hom}(\mathcal{O}_Y(3), \operatorname{coker}(\alpha)) = 0$. Thus any splitting of the first row would induce a splitting of the second row, giving a contraction. $\square$

**Remark 1.16** Proposition 1.15 holds, more generally, for any cubic fourfold with isolated singularities, as long as a general one-parameter smoothing of it has finite monodromy. This corresponds to the K3 surface $S$ of lines through one of the singular points having canonical singularities. The case of the degeneration to the chordal cubic [34], which has finite monodromy but central fiber with 2–dimensional singular locus, will be discussed at length in Section 4.2.

**Proposition 1.17** *Let $X_0 \subset \mathbb{P}^5$ be as in Proposition 1.15 (or as in Remark 1.16) and let $\pi_U : J_U \to U$ be the corresponding intermediate Jacobian fibration. Then there exists a hyper-Kähler compactification $J = J(X_0)$ of $J_U$, with a regular flat morphism to $(\mathbb{P}^5)^\vee$ extending $\pi_U$. Moreover, if $\mathcal{X} \to \Delta$ is a general family of smooth cubic fourfolds degenerating to $X_0$, then up to a base change, there exists a family of Lagrangian fibered hyper-Kähler manifolds*

$$\mathcal{J} \to \mathbb{P}^5_\Delta \to \Delta$$

*such that for $t \neq 0$, $\mathcal{J}_t = J(\mathcal{X}_t)$ is the LSV compactification and, for $t = 0$, $\mathcal{J}_0$ is a hyper-Kähler compactification of $J_U = J_U(X_0)$. Similarly, the analogous statement holds for the twisted intermediate Jacobian.*

**Proof** By Proposition 1.15 above, $J_U$ has a holomorphic symplectic form that extends to a regular form on any smooth projective compactification. As in Lemma 1.4, it follows that $J_U$ is not uniruled. Let $\mathcal{X} \to \Delta$ be a family of smooth cubic fourfolds degenerating to $\mathcal{X}_0 = X_0$ with the property that for $t \neq 0$, $\mathcal{X}_t$ is general in the sense of LSV. As in the beginning of Theorem 1.6, let $\mathcal{J}_\mathcal{V} \to \mathcal{V}$ be such that the fiber over $t \neq 0$ of $\mathcal{J}_\mathcal{V} \to \Delta$ is the LSV compactification $J(\mathcal{X}_t)$ and, over $t = 0$, is $J_U \to U$. We are thus in the position of applying Theorem 1.19 below, which proves the proposition. □

A consequence of this proposition is a positive answer to a question of Beauville [11], as explained in the following remark.

**Remark 1.18** Given a smooth cubic threefold $Y$, let $\ell \subset Y$ be a line. In [9; 26] it is shown that the moduli space of Ulrich bundles on $Y$ with rank 2, $c_1 = 0$ and $c_2 = 2\ell$ is birational to the intermediate Jacobian of $Y$; more precisely, it can be identified with the blowup of the intermediate Jacobian fibration along the Fano surface. Now let $X_0$ be cubic fourfold with one simple node and let $S \subset \mathbb{P}^4$ be the $(2,3)$ complete intersection K3 surface parametrizing lines through the singular point of $X_0$. Consider the Mukai vector $v = 2v_0 = 2(1, 0, -1) \in H^*(S, \mathbb{Z})$ and let $\widetilde{M}_{2v_0}(S)$ be the symplectic resolution of the singular moduli space of OG10–type; cf Section 2.

By considering the relative moduli spaces of Ulrich bundles supported on the five-dimensional family of cubic threefolds containing $S$ and by restricting the bundles to $S$, Beauville [11, Section 5, Example $d = 3$] shows that there is a birational map $J_U \dashrightarrow M_v(S)$. This induces a rational map $M_{2v_0}(S) \dashrightarrow \mathbb{P}^5$, and Beauville asks whether there exists a hyper-Kähler manifold birational to $M_v(S)$ which admits a *regular* morphism to $\mathbb{P}^5$. Proposition 1.17 thus gives a positive answer to this question.

The proof of the proposition above relies on the following theorem, which is the Lagrangian fibration analogue of results from [41, Theorem 2.1 and Corollary 5.2]. Theorem 1.19 will be used also in Section 4 for the proof of Proposition 4.5 (and thus also of Theorem 4.1). As usual, $\Delta$ is an open affine subset of a smooth curve, or a small analytic disk. In both cases, we keep the notation $t = 0$ to denote a chosen special point in $\Delta$, and $t \neq 0$ to denote any other point.

**Theorem 1.19** *Let $\tilde{f}\colon \tilde{\mathcal{J}} \to \Delta$ be a projective degeneration of hyper-Kähler manifolds of dimension $2n$. Suppose that there is a commutative diagram*

$$\begin{tikzcd} \tilde{\mathcal{J}} \arrow[r, "\tilde{\pi}"] \arrow[dr, "\tilde{f}"'] & \mathbb{P}^n_\Delta \arrow[d, "p"] \\ & \Delta \end{tikzcd}$$

*where $\tilde{\mathcal{J}} \to \mathbb{P}^n_\Delta$ is a projective fibration such that for $t \neq 0$, $\mathcal{J}_t \to \mathbb{P}^n_t$ is a Lagrangian fibration. Assume that the central fiber $\tilde{\mathcal{J}}_0 = Y_0 + \sum_{i\in I} m_i Y_i$ has a reduced component $Y_0$ which is not uniruled. Suppose, furthermore, that there is an open subset of $Y_0 \setminus \bigcup_{i\geq 1}(Y_i \cap Y_0)$ such that the morphism to $\mathbb{P}^n_0$ is a fibration $J_{U_0} \to U_0 \subset \mathbb{P}^n_0$ in abelian varieties. Then*:

1. *There exists a projective degeneration $\bar{f}\colon \bar{\mathcal{J}} \to \Delta$ of hyper-Kähler manifolds such that*
   - (a) *$\bar{\mathcal{J}}$ is $\mathbb{Q}$–factorial, terminal and isomorphic to $\tilde{\mathcal{J}}$ over $\Delta^*$,*
   - (b) *the central fiber $\bar{\mathcal{J}}_0$ is a reduced, irreducible, and a normal symplectic variety with canonical singularities and admitting a symplectic resolution, and*
   - (c) *there is a relative Lagrangian fibration $\bar{\pi}\colon \bar{\mathcal{J}} \to \mathbb{P}^n_\Delta$ compatible, via the birational map $\bar{\mathcal{J}} \dashrightarrow \tilde{\mathcal{J}}$, with $\tilde{\pi}$ and such that, up to restricting the open set $U_0 \subset \mathbb{P}^n_0$, the morphism $\bar{\mathcal{J}}_0 \to \mathbb{P}^n_0$ extends the abelian fibration $J_{U_0} \to U_0$.*
2. *Up to a base change $\Delta' \to \Delta$, there exists a (not necessarily projective) family $\mathcal{J} \to \Delta'$ of hyper-Kähler manifolds, with a birational morphism $\mathcal{J} \to \bar{\mathcal{J}}' := \bar{\mathcal{J}} \times_{\Delta'} \Delta$ over $\Delta'$, which is an isomorphism away from the central fiber and in the central fiber is a symplectic resolution of $\bar{\mathcal{J}}_0$. Moreover, $\mathcal{J}$ has a family of Lagrangian fibrations $\pi'\colon \mathcal{J} \to \mathbb{P}^n_{\Delta'}$ compatible with the base change of $\bar{\pi}$.*

**Proof** The proof follows ideas from [75; 40; 41]. Up to passing to a log resolution of the pair $(\tilde{\mathcal{J}}, \tilde{\mathcal{J}}_0)$, we can assume that $\tilde{\mathcal{J}}_0 = Y_0 + \sum_{i=1}^k m_i Y_i$ is a normal crossing divisor. By [41, Theorem 2.1 and Corollary 5.2], running the mmp over $\Delta$ contracts the components $Y_i$ for $i \geq 1$, and yields a birational model of $\widetilde{\mathcal{J}}$ with an irreducible central fiber which is a symplectic variety. In particular, $Y_0$ is the unique component of $\widetilde{\mathcal{J}}_0$ that is not uniruled; cf [41, Remark 2.2]. To prove the theorem we only need to show that the birational maps required to contract the other components can be preformed relatively to $\mathbb{P}^n_\Delta$ and, furthermore, that they induce isomorphism away from $\bigcup_{i\geq 1} Y_i$. This is to ensure that the central fiber has a Lagrangian fibration extending $J_{U_0} \to U_0$ (maybe up to restricting the open subset $U_0 \subset \mathbb{P}^n_0$).

The canonical class $K_{\widetilde{\mathcal{J}}}$ is trivial over $\Delta^*$, so it is $\widetilde{f}$–equivalent to a divisor of the form $\sum_{i=0}^k a'_i Y_i$. Following [75, Section 2.3, point (1)] we set $r = \min a'_i / m_i$, so[1]

$$K_{\widetilde{\mathcal{J}}} =_{\mathbb{Q},\Delta} \sum_{i=0}^{k} a_i Y_i,$$

where $a_i = a'_i - r m_i \geq 0$ are nonnegative rational numbers and $a_i = 0$ for at least one $i$. Let $J \subsetneq \{0, 1, \dots, k\}$ be the set of indices such that $a_i > 0$ and let $J^c$ be its complement. By [75, Proposition 5.1]:

(1) For every $j \in J$, the irreducible component $Y_j$ is uniruled.

(2) If $|J^c| \geq 2$, then for every $j \in J^c$, the irreducible component $Y_j$ is uniruled.

Since $Y_0$ is not uniruled, it follows that $J = \{1, \dots, k\}$ and thus

$$K_{\widetilde{\mathcal{J}}} =_{\mathbb{Q},\widetilde{\pi}} \sum_{i=1}^{k} a_i Y_i, \quad \text{with } a_i > 0.$$

By assumption, for every $i \geq 1$, the closed subset $Y_0 \cap Y_i$ is in the complement of $J_{U_0}$ and, since the fibers of $\widetilde{\mathcal{J}}_0 \to \mathbb{P}^n_\Delta$ are connected, it follows that the induced map $Y_i \to \mathbb{P}^n_0$ is not dominant. Thus, the codimension of $\widetilde{\pi}(Y_i)$ in $\mathbb{P}^n_\Delta$ is greater or equal to two. In other words, $Y_i$ is $\widetilde{\pi}$–exceptional.

We are in the same setting of Theorem 1.6, namely a projective morphism from a smooth quasiprojective variety with a canonical class that is relatively $\mathbb{Q}$–linearly equivalent to an effective divisor all of whose components are relatively exceptional. We can thus argue as in the proof of Theorem 1.6, running the mmp over $\mathbb{P}^n_\Delta$ with scaling of an ample divisor in order to contract each of the $Y_i$, for $i \geq 1$. This yields a birational map $\widetilde{\mathcal{J}} \dashrightarrow \overline{\mathcal{J}}$ over $\mathbb{P}^n_\Delta$, where $\overline{\mathcal{J}} \to \Delta$ has irreducible fibers and the fibration

[1] For a projective morphism $f\colon A \to B$ and two $\mathbb{Q}$–Cartier divisors $D$ and $D'$ on $A$, we write $D =_{\mathbb{Q},B} D'$ or $D \sim_{\mathbb{Q},f} D'$ if and only if $D$ and $D'$ are $\mathbb{Q}$–linearly equivalent up to the pullback of a $\mathbb{Q}$–Cartier divisor from $B$.

$\bar{\mathcal{J}} \to \mathbb{P}^n_\Delta$ has $\mathbb{Q}$–factorial terminal total space and is such that $K_{\bar{\mathcal{J}}} = \tilde{\pi}^* B$ for some $\mathbb{Q}$–divisor $B$ on $\mathbb{P}^n_\Delta$. Since at each step the $K$–negative rays of the mmp are contained in uniruled components of the central fiber, it follows that the birational map $\tilde{\mathcal{J}} \dashrightarrow \bar{\mathcal{J}}$ is an isomorphism away from $\bigcup_{i \geq 1} Y_i$. In particular, the central fiber $\bar{\mathcal{J}}_0$, which is irreducible, has an open subset which is isomorphic to $J_{U_0}$. Since $(K_{\bar{\mathcal{J}}})_{|\mathcal{J}_t} = 0$ for $t \neq 0$, we get that $B_{|\mathbb{P}^n_t} = 0$ for $t \neq 0$. In particular, $B$ is $p$–trivial, where $p\colon \mathbb{P}^n_\Delta \to \Delta$ is the projection, and thus $K_{\bar{\mathcal{J}}}$ is $\tilde{f}$–trivial. We can now argue as in the last part of the proof of [47, Theorem 1.1] to show that $\bar{\mathcal{J}}_0$ is normal with canonical singularities. As in [47, Corollary 4.2] it follows that $\bar{\mathcal{J}}_0$ is a symplectic variety and that, up to a base change $\Delta' \to \Delta$, there exists a smooth family $\mathcal{J} \to \Delta'$ with a birational morphism $\mathcal{J} \to \bar{\mathcal{J}}' := \bar{\mathcal{J}} \times_{\Delta'} \Delta$ with the desired properties. □

**Remark 1.20** Theorem 1.19 gives another proof of Theorem 1.6, as well as the stronger statement of the existence of a relative intermediate Jacobian fibration $\mathcal{J} \to \mathbb{P}^5_\Delta$ associated to any family $\mathcal{X} \to \Delta$ of smooth cubic fourfolds for which the general fiber is general in the sense of LSV.

## 2 Moduli spaces of OG10–type

By [47, Corollary 6.3] (see also [41, Section 6.3]) any hyper-Kähler compactification $J$ of $J_U$ is deformation equivalent to O'Grady's 10–dimensional example. We start this section by recalling the basic definitions and first properties of those singular moduli spaces of sheaves on a K3 surface whose symplectic resolutions are hyper-Kähler manifolds in this deformation class. Then we use the methods of Bayer and Macrì, as adapted by Meachan and Zhang to this class of singular moduli spaces, to study the movable cone of certain moduli spaces that appear naturally as limits of the intermediate Jacobian fibration, when the underlying cubic fourfold degenerates to the chordal cubic; see Section 4.2.

We start by recalling the following fundamental theorem.

**Theorem 2.1** [60; 82; 63; 50; 38; 66] *Let $(S, H)$ be a general polarized K3 surface and let $v_0 \in H^*_{\mathrm{alg}}(S, \mathbb{Z})$ be a primitive Mukai vector which we suppose to be positive in the sense of* [8, Definition 5.1], *see also* [18, Remark 3.1.1]. *Let $m \geq 2$ be an integer. The moduli space $M_{mv_0,H}(S)$ of $H$–semistable sheaves on $S$ with Mukai vector $mv_0$ is an irreducible normal projective symplectic variety of dimension $m^2 v_0^2 + 2$, which* *admits a symplectic resolution if and only if* $m = 2$ *and* $v_0^2 = 2$*. When this is the case, the symplectic resolution* $\widetilde{M}_{2v_0,H}(S) \to M_{2v_0,H}(S)$ *is the blow up of the singular locus* $\mathrm{Sym}^2 M_{v_0,H}(S) \subset M_{2v_0,H}(S)$*, with its reduced induced structure. Moreover,* $\widetilde{M}_{2v_0,H}(S)$ *is an irreducible holomorphic symplectic manifold and its deformation class is independent of* $(S,H)$ *and of* $v_0$*; in particular,* $\widetilde{M}_{2v_0,H}(S)$ *is deformation equivalent to O'Grady's original* 10*–dimensional exceptional example.*

We will refer to a Mukai vector of the form $2v_0$ with $v_0^2 = 2$ as a Mukai vector of OG10–type and to a hyper-Kähler manifold in this deformation class as a hyper-Kähler of OG10–type.

## 2.1 Contracting the relative theta divisor on the relative Jacobian of curves

It is known [4; 7; 8; 5] that the birational geometry of moduli spaces of pure dimension one sheaves on a K3 surface is related to Brill–Noether loci. For example, on the degree $g-1$ Beauville–Mukai system of a genus $g$ linear system on a K3 surface, the relative theta divisor can be contracted, possibly after performing a finite sequence of birational transformations. This is the content of the following example.

**Example 2.2** [4; 5] Let $(S,C)$ be a general polarized K3 surface of genus $g$, with $\mathrm{NS}(S) = \mathbb{Z}C$. Set $v = (0,C,0) \in H^*(S,\mathbb{Z})$ and let $M_v$ be the moduli space of $C$–stable sheaves on $S$ with Mukai vector[2] $v$. Since we are assuming $(S,C)$ to be general in moduli, we are suppressing the polarization from the notation — thus $M_v$ will denote the moduli space of $C$–semistable sheaves on $S$ with Mukai vector $v$; when we consider instead a Bridgeland stability condition $\sigma$, the corresponding moduli space will be denoted by $M_{v,\sigma}$. This moduli space is smooth and $M_v \to \mathbb{P}^g = |C|$ is the degree $g-1$ relative compactified Jacobian of the genus $g$ linear system $|C|$ on $S$. There is a naturally defined effective, irreducible, relatively ample theta divisor $\theta \subset M_v$ which parametrizes sheaves with a nontrivial global section and which can be realized as the zero locus of a canonical section of the determinant line bundle; see [48, Section 2.3] or [3, Theorem 5.3]. Recall that there is a Hodge isometry $\mathrm{NS}(M_v) \cong v^\perp = \langle (0,0,1), (1,0,0) \rangle$; see for example [7, Theorem 3.6].

[2]This Mukai vector is not positive in the sense defined above, since both the first and last entry are zero. However, since for general $(S,C)$, tensoring by $C$ induces an isomorphism with $M_{v'}$, where $v' = (0,C,g-1)$, the results of [7] still hold. See also [66] for other considerations about the last entry of the Mukai vector.

The class $\ell := (0,0,1)$ is the class of the isotropic line bundle inducing the Lagrangian fibration $M_v \to \mathbb{P}^g$ while the theta divisor $\theta$ corresponds to the class $-(1,0,1) = -v(\mathcal{O}_S)$; see [48, page 643] or also [7, Proposition 7.1 and Theorem 12.3].[3] Since $\theta^2 = -2$, the irreducible effective divisor $\theta$ is prime exceptional. By [27], it can be contracted on a hyper-Kähler birational model of $M_v$. Since the rays corresponding to divisorial contractions and to Lagrangian fibrations must be in the boundary of the movable cone [36], it follows that

$$\overline{\mathrm{Mov}}(M_v) = \mathbb{R}_{\geq 0}\ell + \mathbb{R}_{\geq 0}h,$$

where $h = (-1,0,1) \in \theta^{\perp} \cap v^{\perp}$ is a big line which is nef on some birational model of $M_v$; this also follows from [7, Theorem 12.3]. Using [7], the walls of the nef cones of the various birational models can be computed. Since we don't need this, we omit the computation.

## 2.2 Movable cones of certain moduli spaces of OG10–type

If we consider a nonprimitive genus $g$ linear system $|mC|$, with $m \geq 2$, then the relative compactified Jacobian of degree $g-1$ is singular. For singular moduli spaces of OG10–type, ie when $v = 2w$ with $w^2 = 2$, Meachan and Zhang [58] adapted the techniques of Bayer and Macrì [8; 7] to compute the nef and movable cones of these moduli spaces. We refer to [16; 4; 8; 7] for the relevant definitions and main results on Bridgeland stability conditions on K3 surfaces, and to [58] for the results on moduli spaces of OG10 type.

By [58, Theorem 7.6(3)], all birational models of $M_{2w} = M_{2w,C}$ which are isomorphic to $M_{2w}$ in codimension one are isomorphic to a Bridgeland moduli space $M_{2w,\sigma}$ for some Bridgeland stability condition $\sigma$ on $S$. Moreover, by [58, Corollary 2.8],

$$\mathrm{NS}(M_{2w,\sigma}) \cong w^{\perp}. \tag{2-1}$$

We now apply the results of [58] to describe the nef and movable cones of certain singular models of OG10 appearing as limits of the intermediate Jacobian fibration. By [67], the factoriality properties of a singular moduli space $M_{2w}$ of OG10–type depend on the divisibility of the primitive Mukai vector $w \in H^*_{\mathrm{alg}}(S,\mathbb{Z})$. More precisely, by [67, Theorem 1.1], $M_{2w}$ is factorial if and only if $w \cdot u \in 2\mathbb{Z}$ for every $u \in H^*_{\mathrm{alg}}(S,\mathbb{Z})$. Otherwise, $M_{2w}$ is 2–factorial. Since there can be different birational models with

[3] Compared to [7], there is a difference in a choice of sign in the isomorphism $\mathrm{NS}(M_v) \cong v^{\perp}$.

different factoriality properties (cf Remark 2.3), it is important to choose the correct model to work with.

Now let $(S,C)$ be a general K3 surface of degree 2 and set

$$v_k := (0, C, k-2). \tag{2-2}$$

The Le Potier morphism $\pi\colon M_{2v_k} \to \mathbb{P}^5$ realizes the singular moduli space $M_{2v_k}$ as a compactification of the degree $2k$ relative Jacobian of the genus-five hyperelliptic linear system $|2C|$. Composing $\pi$ with the symplectic resolution $m\colon \widetilde{M}_{2v_k} \to M_{2v_k}$, we get a natural Lagrangian fibration

$$\tilde{\pi}\colon \widetilde{M}_{2v_k} \to \mathbb{P}^5. \tag{2-3}$$

By the result of Perego and Rapagnetta mentioned above, $M_{2v_k}$ is factorial if and only if $k$ is even. It turns out that the birational class of these moduli spaces is independent of $k$, but the isomorphism class depends on the parity of $k$ [18, Proposition 3.2.7]: indeed, tensoring a pure dimension one sheaf by $\mathcal{O}_S(C)$ determines an isomorphism

$$M_{2v_k} \xrightarrow{\sim} M_{2v_{k+2}}. \tag{2-4}$$

**Remark 2.3** Tensoring a line bundle supported on a smooth hyperelliptic curve of genus 5 by the unique $g^1_2$ on the curve defines a birational morphism $M_{2v_k} \dashrightarrow M_{2v_{k+1}}$. (I thank A Rapagnetta for pointing out this to me.) As a side remark, notice that the map thus defined is *not* an isomorphism in codimension one. Indeed, it can be checked that when passing to the birational morphism $\widetilde{M}_{2v_k} \dashrightarrow \widetilde{M}_{2v_{k+1}}$ between the two resolutions, which is an isomorphism in codimension two, the exceptional divisor of one model is exchanged with the proper transform of the locus parametrizing sheaves on reducible curves on the other model.

In view of Lemma 4.4 below and the isomorphism (2-4), we will focus on the case $k=0$.

**Remark 2.4** For general $(S,C)$ it is not hard to check that the structure sheaf of every curve in $|2C|$ satisfies the numerical criterion for $C$–stability and hence that the fibration $M_{2v_0} \to \mathbb{P}^5$ admits a regular zero section. Notice also that the image of this section is not contained in the singular locus of $M_{2v_0}$.

By [58, Corollary 2.8], $\mathrm{NS}(M_{2v_0}) \cong v_0^{\perp} = U = \langle (0,0,1), (-1,C,0)\rangle$, where, as above,

$$\ell = (0,0,1) \tag{2-5}$$

is the line bundle inducing the Lagrangian fibration $\pi\colon M_{2v_0} \to \mathbb{P}^5 = |2C|$. Under the isomorphism $M_{2v_2} \cong M_{2v_0}$ induced by tensoring with $\mathcal{O}_S(-C)$, the relative theta divisor is mapped isomorphically to the prime exceptional divisor

$$\theta := -(1, -C, 2) \tag{2-6}$$

parametrizing sheaves which receive a nontrivial morphism from the spherical object $\mathcal{O}_S(-C)$; see also Lemma 2.6. Indeed, the relative theta divisor in $M_{2v_2}$ parametrizes sheaves with a nontrivial morphism from $\mathcal{O}_S$ and thus its image in $M_{2v_0}$ is exactly the divisor $\theta$. Notice that

$$\theta^2 = -2. \tag{2-7}$$

For later use we highlight the following remark.

**Remark 2.5** The effective divisor $\theta \subset M_{2v_0}$ with cohomology class (2-6) does not contain the singular locus of $M_{2v_0}$: using the description of $\theta$ as the zero locus of a section of the determinant line bundle [3, Theorem 5.3], which is compatible with $S$–equivalence classes, it is enough to show that the section defining $\theta$ is not identically zero on the singular locus of $M_{2v_0}$. It is therefore sufficient to show that there are $S$–equivalence classes of polystable sheaves all of whose members have a zero space of global sections. This is clear, since the generic semistable sheaf with Mukai vector $2v_0$ is an extension of two degree-one line bundles each supported on two distinct curves of genus two.

The following lemma is an application of [58, Theorems 5.1–5.3] to $M_{2v_0}$. (Note that Example 8.6 of loc. cit. is for odd $k$, so in view of Remark 2.3 it is concerned with a birational model of $M_{2v_0}$ which is not isomorphic in codimension one, and hence we cannot immediately apply it here.)

**Lemma 2.6** *Let the notation be as above. Then*

$$\mathrm{Nef}(M_{2v_0}) = \mathbb{R}_{\geq 0}\ell + \mathbb{R}_{\geq 0}h_0, \quad \mathrm{Mov}(M_{2v_0,C}) = \mathbb{R}_{\geq 0}\ell + \mathbb{R}_{\geq 0}h,$$

*where*

$$\ell = (0,0,1), \quad h_0 = (-1, C, 1), \quad h = (-1, C, 0).$$

*Moreover, the wall spanned by $h_0 = (-1, C, 1)$ contracts the zero section of $M_{2v_0} \to \mathbb{P}^5$ and the class corresponding to $h = (-1, C, 0)$ is big and nef on the Mukai flop of $M_{2v_0}$ along the zero section and contracts the proper transform of $\theta$.*

**Proof** Since $\ell = \pi^*\mathcal{O}_{\mathbb{P}^5}(1)$ is nef and isotropic, it is one of the two rays of both the Nef and the movable cones of $M_{2v_0}$. By [58, Theorem 5.3] there is a divisorial contraction of BNU–type (notation as in loc. cit.), determined by the spherical class $s = (1, -C, 2)$, which is orthogonal to $v_0$. The second ray of the movable cone is thus determined by $s^\perp \cap v_0^\perp$. We pick $h = (-1, C, 0)$ as a generator of this ray, since $h \cdot \ell > 0$. By the same theorem in [58], the flopping walls are determined by $w^\perp \cap v_0^\perp$ for $w$ spherical and such that $w \cdot v_0 = 2$. There is a unique ray in $\mathrm{Mov}(M_{2v_0})$ that is of this form. It is determined by $w = (1,0,1) = v(\mathcal{O}_S)$ or, equivalently, by $w' = (-1, 2C, -5) = 2v_0 - w$. We can choose $h_0 = (-1, C, 1)$ as generator of this ray. As in [58, Remark 8.5], we can see that this wall corresponds to the flop of the $\mathbb{P}^5$ corresponding to the sheaves with a morphism from $\mathcal{O}_S$, ie of the image of the zero section. □

**Remark 2.7** It can be shown that the birational model on the other side of the wall can be identified with the Gieseker moduli space $M_{2w_0}$, where $w_0 = (2, C, 0)$. Since we don't need this in the rest of the paper, we omit the proof.

**Remark 2.8** The theta divisor $\theta$ is Cartier, since by [66] $M_{2v_0}$ is factorial; see also Section 2.2. Moreover, it is relatively ample over $\mathbb{P}^5$, since by the description of the Nef cone of Lemma 2.6 we can write $\theta$ as a sum of an ample line bundle and a multiple of $\ell = \pi^*\mathcal{O}_{\mathbb{P}^5}(1)$.

# 3 The relative theta divisor on the intermediate Jacobian fibration

For any smooth cubic threefold $Y$, there is a canonically defined theta divisor in $\mathrm{Jac}(Y)$ which is $(-1)$–invariant and whose unique singular point lies at the origin. For the hyper-Kähler compactification $J = J(X) \to (\mathbb{P}^5)^\vee$ of the intermediate Jacobian fibration associated to a smooth cubic 4–fold $X$, there is an effective relative theta divisor $\Theta \subset J$, which is defined as the closure of the union of the canonical theta divisor in the smooth fibers. More precisely, by [20; 21], see also [47, Lemma 5.4], $\Theta$ can be defined as the closure of the image of the Abel–Jacobi difference mapping

$$\mathcal{F} \times_{(\mathbb{P}^5)^\vee} \mathcal{F} \dashrightarrow J, \quad (\ell, \ell', Y) \mapsto \phi_Y(\ell - \ell'). \tag{3-1}$$

The relative theta divisor $\Theta$ played an important role in [47], where it was shown that for general $X$ the divisor $\Theta$ is $\pi$–ample and $J$ is identified with the relative Proj of the sheaf of $\mathcal{O}_{\mathbb{P}^5}$–algebras associated with this divisor. Another useful way of realizing the theta divisor is using twisted cubics [21]. Let $Z = Z(X)$ be the Lehn–Lehn–Sorger–van Straten 8–fold [49]. Then $Z$ is the blowdown $g\colon Z' \to Z$ of a smooth 10–fold $Z'$ whose points parametrize nets of (generalized) twisted cubics. The exceptional locus of $g$ parametrizes non-ACM cubics and its image in $Z$ is isomorphic to the cubic itself. Let

$$r\colon \mathbb{P}_{Z'} \to Z'$$

be the $\mathbb{P}^1$–bundle over $Z'$ whose fiber over a twisted cubic $[C] \in Z'$ is the pencil $\mathbb{P}^1_C$ of hyperplane sections of $X$ containing $\Sigma_C := X \cap \langle C \rangle$. Here $\langle C \rangle = \mathbb{P}^3$ is the linear span of the curve. By [47, Sublemma 5.5], see also [21, Section 4] or [41, Proposition 6.10], the Abel–Jacobi map

$$\varphi\colon \mathbb{P}_{Z'} \dashrightarrow J, \quad (C, Y) \mapsto \phi_Y(C - h^2), \tag{3-2}$$

is birational onto its image, which is precisely $\Theta$. Here $h^2$ is the class of the intersection of two hyperplanes in $Y$.

**Remark 3.1** For later use, we note the following two facts. First of all, the restriction of $\mathbb{P}_{Z'}$ to the locus of nonCM cubics is mapped to the zero section of $J \to \mathbb{P}^5$ (which lies in $\Theta$). Second, using the Gauss map, see [20, Section 12] or also [33, Section 3], one can see that if $C$ is a twisted cubic in a smooth cubic threefold $Y$ with the property that $\phi_Y(C - h^2) = 0$ in $\mathrm{Jac}(Y)$, then the cubic surface $\Sigma_C = Y \cap \langle C \rangle$ is singular.

For every $X$, the Néron–Severi group of $J = J(X)$ has at least rank two, since

$$\mathrm{NS}(J(X)) \supset \langle L, \Theta \rangle.$$

Here $L = \pi^* \mathcal{O}_{\mathbb{P}^5}(1)$ and $\Theta$ is, as above, the relative theta divisor obtained as the closure of the image of (3-1).

**Lemma 3.2** *For any smooth $X$, there is an isomorphism of rational Hodge structures $H^2(J, \mathbb{Q})_{\mathrm{tr}} \cong H^4(X, \mathbb{Q})_{\mathrm{tr}}$. In particular, $\rho(J) = \mathrm{rk}\, H^{2,2}(X, \mathbb{Q}) + 1$.*

**Proof** The first statement was already noted in [47], while the second follows from the first and the fact that $b_2(J) = 24$ and $b_4(X) = 23$. □

**Remark 3.3** The locus, inside $\mathrm{Def}(J)$, parametrizing intermediate Jacobian fibrations is of codimension two and corresponds to the locus where the classes $L$ and $\Theta$ stay of type $(1,1)$. By [74, Theorem 6], a Lagrangian fibration with a section deforms, as Lagrangian fibration with a section, over a smooth codimension-two locus of the deformation space of the underlying hyper-Kähler manifold. Since by Theorem 1.1 for general $X$ the LSV compactification $J(X)$ has a section, it follows that the codimension-two locus where $L$ and $\Theta$ stay algebraic is exactly the locus where the section deforms.

We highlight the following corollary for future reference.

**Corollary 3.4** *For very general $X$, we have that $\rho(J) = 2$. Thus*:

(1) *$J$ is the only projective hyper-Kähler birational model of $J_U$ where $L$ is nef. In particular, any hyper-Kähler compactification of $J_U$ with a Lagrangian fibration extending $J_U \to U$ is isomorphic to the compactification of* [47].

(2) *There is at most one prime exceptional divisor on $J$.*

**Proof** (1) Since $\rho(J) = 2$, the boundary of the movable cone of $J$ has two rays, of which $L$ is one.

(2) If there is a prime exceptional divisor, its class has to be orthogonal to the second extremal ray of the movable cone [52, Theorem 1.5]. Since two prime exceptional divisors with proportional classes have to be isomorphic [53, Corollary 3.6(3)], there is at most one prime exceptional divisor. □

The following lemma was communicated to me by K Hulek and R Laza. I thank them for sharing this observation with me and for raising the question of computing $q(\Theta)$.

**Lemma 3.5** *We have that $q(L, \Theta) = 1$. In particular, $\langle L, \Theta\rangle$ is a primitive sublattice of $\mathrm{NS}(J)$, isomorphic to the standard hyperbolic lattice $U$ of rank two. For very general $X$, $\mathrm{NS}(J) = U$.*

**Proof** The computation of $q(L, \Theta)$ goes as in [73, Lemma 1]: one expands in $t$ the Fujiki equality $q(L + t\Theta)^5 = c(L + t\Theta)^{10}$, where $c = 945$ is the Fujiki constant [71], and uses the fact that $\Theta^5 L^5 = (\Theta_{|J_{[H]}})^5 = 5!$. The final statement follows from Lemma 3.2. □

I thank C Onorati for many discussions around $\Theta$ and for his interest in the following computation.

**Proposition 3.6** *The irreducible divisor $\Theta \subset J$ is prime exceptional. In particular, it can be contracted on some projective birational hyper-Kähler model of $J$. Moreover, $q(\Theta) = -2$.*

**Proof** Let $\varphi\colon \mathbb{P}_{Z'} \dashrightarrow \Theta$ be the Abel–Jacobi map as in (3-2) and let $V \subset Z'$ be a nonempty open subset such that the restriction of $\varphi$ to $r^{-1}(V) =: \mathbb{P}_V$ is regular. By restricting $V$ if necessary, we can assume that all twisted cubics parametrized by $V$ are such that $\Sigma_C$ is smooth and, in particular, that $C$ is ACM.

Recall that for any twisted cubic $[C] \in V$, we have set $\mathbb{P}^1_C = r^{-1}(C)$. The rational curve $\varphi(\mathbb{P}^1_C) \subset J$ is smooth because it maps to $\pi(\varphi(\mathbb{P}^1_C)) \subset \mathbb{P}^5$, which is the pencil of hyperplane sections of $X$ that contain the curve $C$. Moreover, $(\pi \circ \varphi)(\mathbb{P}_V)$ intersects the dual variety $X^\vee$ in a dense open subset and, similarly, $\varphi(\mathbb{P}_V)$ intersects $\Theta_{X^\vee}$, the restriction of $\Theta$ to $X^\vee$, in a dense open subset. (This statement follows from [21] and the fact, proven there, that for a cubic threefold with one $A_1$ singularity the Abel–Jacobi mapping is birational onto its image.)

We start by showing that for a general $[C] \in V$, the smooth rational curve $\varphi(\mathbb{P}^1_C)$ is contained in the smooth locus of $\Theta$. The singular locus $\Theta_{\mathrm{sing}}$ of $\Theta$ has an irreducible component that is equal to the closure of the zero section of $J_U \to U$, while any other irreducible component of the singular locus is properly contained in $\Theta_{X^\vee}$, the restriction of $\Theta$ to $X^\vee$. Since $V$ parametrizes ACM curves, by Remark 3.1 it follows that the intersection of $\varphi(\mathbb{P}_V)$ with the image of the zero section of $J \to \mathbb{P}^5$ is contained in $J_{X^\vee}$, the restriction of $J$ to $X^\vee \subset \mathbb{P}^5$. Let $B := \varphi^{-1}(\Theta_{\mathrm{sing}} \cap \varphi(\mathbb{P}_V))$ be the locus in $\mathbb{P}_V$ parametrizing points mapped to the singular locus of $\Theta$ and let $W_V := (\pi \circ \varphi)^{-1}(X^\vee \cap (\pi \circ \varphi)(\mathbb{P}_V))$ be the locus in $\mathbb{P}_V$ of pairs $(C, Y)$ such that $Y$ is singular. By what we have observed, it follows that $B \subseteq W_V$. Notice that $W_V$ is irreducible of dimension 8, because it maps to an open subset of $X^\vee$, with fibers parametrizing equivalence classes of twisted cubics contained in a cubic threefold with one $A_1$ singularity; these form an irreducible subset of $Z(X)$, as follows from [21, Section 3]. We have already observed that the general point of $\Theta_{X^\vee}$ is contained in the image $\varphi(\mathbb{P}_V)$. Thus, if $Y_0$ corresponds to a general point in $X^\vee$, there is a twisted cubic $C \subset Y_0$, with $[C] \in V$ and such that $\varphi(C, Y_0) = \phi_{Y_0}(C)$ lies in the smooth locus of $\Theta$. It follows that $B$ is strictly contained in $W_V$. Since $W_V$ is irreducible, $\dim B = 7$. Thus $B$ does not dominate $V$ and hence the image of the open subset $\mathbb{P}_{V'} := r^{-1}(V')$, where $V' := V \setminus r(B)$ is contained in the smooth locus of $\Theta$.

For the general point in $(C, Y) \in \mathbb{P}_V$, let $R := \varphi(\mathbb{P}^1_C) \subset \Theta$ be the corresponding element of the ruling. By generic smoothness, the differential of $\varphi$ is of maximal rank at a general point $x \in R$, so by [39, Chapter II, Proposition 3.4], the vector bundle $(T_\Theta)_{|R}$ is globally generated at $x \in R$. It follows that $(T_\Theta)_{|R} = \bigoplus \mathcal{O}_R(a_i)$, with $a_i \geq 0$. By Lemma 3.8 below, the restriction of the tangent bundle of $J$ to the smooth rational curve $R$ is of the form $\mathcal{O}_R^{\oplus 8} \oplus \mathcal{O}_R(2) \oplus \mathcal{O}_R(-2)$. Using this and the fact that $R$ is contained in the smooth locus of $\Theta$, we find that $(T_\Theta)_{|R} = \mathcal{O}_R(2) \oplus \mathcal{O}_R^{\oplus 8}$ and hence that $N_{R|\Theta} = \oplus \mathcal{O}_R^{\oplus 8}$. In particular, $\Theta \cdot R = -2$.

Consider the lattice embedding $H^2(J,\mathbb{Z}) \subset H^2(J,\mathbb{Z})^\vee = H_2(J,\mathbb{Z})$ induced by the Beauville–Bogomolov form. We claim that under this embedding, the classes of $R$ and of $\Theta$ are equal, ie that $R \cdot x = q(\Theta, x)$ for every $x \in H^2(J,\mathbb{Z})$. This immediately proves the proposition, as it implies that $q(\Theta) = \Theta \cdot R = -2$. By [53, Corollary 3.6(1)] and [27, Proposition 4.5], the class of the ruling of a prime exceptional divisor is proportional, via a positive constant, to the class of the exceptional divisor. Thus, to prove the claim it suffices to show that $\Theta$ is prime exceptional, since the constant would have to be equal to 1, as both $R \cdot L$ and $q(\Theta, L)$ are equal to 1.

To prove that $\Theta$ is prime exceptional we use standard techniques on deformations of maps from rational curves to hyper-Kähler manifolds, following [53, Section 5.1] or also [19, Section 3]. We include a proof because the setting of Markman is different and because the proof in [19, Section 3] is for projective families of hyper-Kähler manifolds. Choose $R \subset \Theta$ a general element in the ruling and let $\mathrm{Def}(J)_R \subset \mathrm{Def}(J)$ be the smooth hypersurface in the deformation space of $J$ where the class of $R$ stays of Hodge type. Let $\mathcal{H}ilb \to \mathrm{Def}(J)_R$ be the component of the relative Douady space containing the point $[R]$. Since $N_{R|J} = \mathcal{O}_R(-2) \oplus \mathcal{O}_R^{\oplus 8}$, then by [70, Theorem 1] it follows that the morphism $\rho\colon \mathcal{H}ilb \to \mathrm{Def}(J)_R$ is smooth at $R$ and of relative dimension 8. Let $T \subset \mathrm{Def}(J)_R$ be a general curve containing 0 (in particular we can assume that for very general $t \in T$, the Néron–Severi of the corresponding deformation $\mathcal{J}_t$ of $J$ is one-dimensional and spanned by a line bundle whose class is proportional to $R_t$, the parallel transport of the class of $R$ to $\mathcal{J}_t$) and let $\rho_T\colon \mathcal{H}ilb_T \to T$ be the component of the base change to $T$ of $\mathcal{H}ilb \to \mathrm{Def}(J)_R$ that contains $[R]$. Since $\rho$ is smooth at $[R]$, $\rho_T$ is dominant of relative dimension 8. Up to a base change and to restricting $T$, we can assume that $\mathcal{H}ilb_T \to T$ has irreducible fibers for $t \neq 0$. Let $\mathcal{J}_T \to T$ be the base change of the universal family to $T \to \mathrm{Def}(J)$ and let $\mathcal{D} \subset \mathcal{J}_T$ be the image of the universal family over $\mathcal{H}ilb_T$ under the evaluation map. Then $\mathcal{D}$ is irreducible of relative codimension one. Moreover, $\mathcal{D}_t$ is irreducible for $t \neq 0$, and $D := \mathcal{D}_0$ is a union of effective uniruled divisors containing $\Theta$ as an irreducible component (with a given multiplicity $m \geq 1$). By the choice of $T$, for very general $t$, $\rho(\mathcal{J}_t) = 1$. It follows that the class of $\mathcal{D}_t$ is proportional to the class of $R_t$ and hence that the class of $D = \mathcal{D}_0$ is proportional to that of $R$. Moreover, the proportionality constant is positive, as both $D$ and $R$ intersect positively with a Kähler class. Hence, since $\Theta \cdot R$ is negative, so is $q(\Theta, D)$. Moreover, since the product of two distinct irreducible uniruled divisors is nonnegative, it follows that $q(\Theta, D) \geq m\, q(\Theta, \Theta)$. Thus $q(\Theta, \Theta) < 0$, ie $\Theta$ is prime exceptional. Thus, as already observed, the classes of $\Theta$ and of $R$ have to be the same and hence $q(\Theta, \Theta) = -2$. $\square$

**Remark 3.7** A posteriori, once we know that $\Theta$ is prime exceptional, we can use [53, Lemma 5.1] to show that $\mathcal{D}_0 = \Theta$.

**Lemma 3.8** *Let $M$ be a hyper-Kähler manifold of dimension $2n$ and $R \subset M$ be a smooth rational curve. Suppose $R$ is a general ruling of a uniruled divisor. Then*

$$(T_M)_{|R} = \mathcal{O}_R^{\oplus 2n-2} \oplus \mathcal{O}_R(2) \oplus \mathcal{O}_R(-2),$$

*and thus*

$$N_{R|M} = \mathcal{O}_R(-2) \oplus \mathcal{O}_R^{\oplus 2n-2}.$$

**Proof** Since $T_M$ is self dual, $(T_M)_{|R} = \bigoplus_i \mathcal{O}_R(a_i) \oplus \bigoplus_i \mathcal{O}_R(-a_i)$, where $a_i \geq 0$. Since $R$ is general and its deformations sweep out a divisor, by [39, Chapter II, Proposition 3.4], the rank of the evaluation map $\mathrm{rk}[H^0(R, (T_M)_{|R}) \otimes \mathcal{O}_R \to (T_M)_{|R}]$ at a general point of $R$ is equal to $2n-1$. Hence $a_2 = \cdots = a_n = 0$ and $a_1 \geq 2$; cf [27, Proposition 4.5]. Since the normal sheaf of $R$ in $M$ is torsion-free and contains the quotient $\mathcal{O}_R(a_1)/T_R = \mathcal{O}_R(a_1)/\mathcal{O}_R(2)$, it follows that $a_1 = 2$. $\square$

Notice that the same argument as the last part of the proof of Proposition 3.6 shows the following.

**Proposition 3.9** *Let $M$ be a hyper-Kähler manifold of dimension $2n$ and let $E \subset M$ be an irreducible uniruled divisor. Suppose that a general curve $R$ in the ruling is smooth and that $E \cdot R < 0$, eg if $R$ is contained in the smooth locus of $E$. Then $E$ is prime exceptional and hence, under the lattice embedding $H^2(M, \mathbb{Z}) \subset H^2(M, \mathbb{Z})^\vee = H_2(M, \mathbb{Z})$ induced by the Beauville–Bogomolov form, the classes of $E$ and $R$ are proportional by a positive constant.*

**Corollary 3.10** *For very general $X$, the movable cone of $J(X)$ is spanned by $L$ and $H$, where $H$ is a generator of $\Theta^\perp \subset \mathrm{NS}(J)$ with $q(H, L) > 0$ and $q(H) > 0$; ie*

$$\mathrm{Mov}(J) = \mathbb{R}_{\geq 0} L + \mathbb{R}_{\geq 0} H.$$

*In particular, there is a unique hyper-Kähler model of $J$ with a Lagrangian fibration, and $J$ is not birational to the twisted intermediate Jacobian fibration $J^T$.*

**Proof** We already know that one of the rays of the movable cone of $J$ is spanned by $L$. By [53, Theorem 1.5] the closure of the movable cone is spanned by classes that intersect nonnegatively with all prime exceptional divisors. Since by Proposition 3.6, $\Theta$ is prime exceptional, the second ray of the movable cone is determined by $\Theta^{\perp}$, which is spanned by a class $H$ which is big and nef on some birational hyper-Kähler model of $J$. Thus, $q(H) > 0$ and $q(H, L) > 0$. In particular, the movable cone is strictly contained in the positive cone, implying that the only isotropic class that is movable is $L$. $\square$

In terms of the other projective hyper-Kähler birational models of $J$, we can actually prove something more precise. The main result of Section 4 describes, for general $X$, which birational model of $J$ the proper transform of $\Theta$ can be contracted on.

## 3.1 Induced automorphisms

For hyper-Kähler manifolds of $\mathrm{K3}^{[n]}$–type, a considerable amount of literature has been devoted to the study and classification of automorphism groups. This includes studying the automorphisms induced from a K3 surface to the moduli spaces of sheaves on it. In view of Theorem 1.6, a natural question is to study the induced action on $J$ of the automorphism group of $X$ in relation to the Lagrangian fibration structures. I thank G Pearlstein for asking questions that led me to the following observations.

Let $X$ be a smooth cubic fourfold and let $\tau$ be an automorphism of $X$. Then $\tau$ acts on the universal family of hyperplane sections of $X$ and thus also on the Donagi–Markman fibration $J_U \to U$, which is identified with the relative $\mathrm{Pic}^0$ of the family of Fano surfaces of the hyperplane sections of $X$. By abuse of notation we denote by

$$\tau\colon J \dashrightarrow J$$

the induced birational morphism. Notice that $\tau$ preserves $\Theta$ and $L$, so the induced action of $\tau^*$ is the identity on $U = \langle L, \Theta\rangle \subset \mathrm{NS}(J)$.

**Proposition 3.11** *Let $X$ be a smooth cubic fourfold and suppose that the fibers of $\pi\colon J \to \mathbb{P}^5$ are irreducible.* (By [47] this happens for general $X$.) *Then*:

1. *$\Theta$ is $\pi$–ample and so is any $B \in \mathrm{NS}(J)$ with $q(L, B) > 0$.*
2. *Any birational automorphism $\tau\colon J \dashrightarrow J$ which fixes $L = \pi^*\mathcal{O}(1)$ extends to a regular automorphism.*

(3) *$L$ is nef on a unique hyper-Kähler birational model of $J$. In other words, if $J'$ is a birational hyper-Kähler model of $J$, with birational map $f: J' \dashrightarrow J$, and the induced map $\pi': J' \dashrightarrow J \to \mathbb{P}^5$ is regular, then $f$ is an isomorphism.*

**Proof** (1) Let $H$ be an ample line bundle on $J$ and let $J_t$ be a smooth fiber of $J \to \mathbb{P}^5$. Then $[H_{|J_t}] = m[\Theta_{|J_t}]$ for a positive integer $m$, so the restrictions of $H$ and $m\Theta$ are topologically equivalent for any smooth fiber. Since the fibers of $\pi$ are irreducible, it follows that the restrictions of $H$ and $m\Theta$ to any fiber are numerically equivalent; see [80, Lemma 4.4]. By Nakai–Moishezon, $m\Theta$ is $\pi$–ample. Similarly, if $q(B, L) > 0$, then there exists positive integers $a$ and $b$ such that $aB$ and $b\Theta$ are numerically equivalent on every fiber.

(2) By assumption, $\tau^*L = L$ so $q(\tau^*\Theta, L) = q(\Theta, L) = 1$. As a consequence, $\tau^*\Theta$ and $\Theta$ are topologically equivalent on the smooth fibers and hence, as above, numerically equivalent on every fiber. Thus, $\tau^*\Theta$ is $\pi$–ample. It follows that $\tau$ is a regular morphism.

(3) Let $H'$ be any ample line bundle on $J'$ and let $L' = f^*L = \pi'^*\mathcal{O}(1)$. Then $0 < q(L', H') = q(L, f^*H')$, so by (1) $f^*H'$ is ample and $f$ is an isomorphism. □

In addition to birational automorphisms induced by the automorphisms of $X$, some examples of birational automorphisms which preserve $L$ are:

(1) The map $\iota: J \to J$ induced by the action of $(-1)$ on the smooth fibers of $J \to \mathbb{P}^5$.

(2) The map $t_\alpha: J \to J$ induced by the translation of a rational section of $\alpha: \mathbb{P}^5 \dashrightarrow J$; cf Section 5.

(3) More generally, any birational automorphism induced by an element of the automorphism group of $J_K$, the generic fiber of $J \to \mathbb{P}^5$.

**Remark 3.12** As already mentioned just below Theorem 1.1, a necessary condition for the irreducibility of the fibers of $J \to \mathbb{P}^5$ is given in [17]. This condition is satisfied if and only if the hyperplane sections $Y$ of $X$ satisfy

$$d(Y) := b_2(Y) - b_4(Y) = 0,$$

where $b_i(Y)$ denotes the $i^{\text{th}}$ Betti number of $Y$ and where $d(Y)$ is called the defect of $Y$. It is easy to see that if $Y$ contains a plane then $d(Y) > 0$.

# 4 Birational geometry of $J(X)$ for general $X$

To describe the birational geometry of the intermediate Jacobian fibration we degenerate the underlying cubic to the chordal cubic, following an idea already contained in [41]. There, it was observed that the central fiber of the corresponding family of intermediate Jacobian fibrations can be chosen to be birational to a moduli space of sheaves of OG10–type on a K3 surface of genus two. As in Section 2, by moduli space of OG10–type we mean a moduli space of sheaves on a K3 surface with Mukai vector $2w$, with $w^2 = 2$. We first refine the construction of this degeneration in order to have a central fiber that is actually *isomorphic* to a certain singular moduli space of sheaves on the associated K3 surface. In this way, we can keep track of the limits of the relative theta divisor and of the line bundle inducing the Lagrangian fibration. This is done in Section 4.2. The results of Meachan and Zhang [58], which were recalled in Lemma 2.6, imply that the central fiber of the relative theta divisor can be contracted after a Mukai flop of the zero section. For $X$ general, we then deduce the same result for $J(X)$ and, for very general $X$, we compute the nef and movable cone of $J(X)$. This is the content of Theorem 4.1.

**Theorem 4.1** *Let $X$ be a smooth cubic fourfold and let $J = J(X) \to \mathbb{P}^5$ be a hyper-Kähler compactification of the intermediate Jacobian fibration as in Section 1.*

*For very general $X$:*

(1) *There is a unique other hyper-Kähler birational model of $J$, denoted by $N$, which is the Mukai flop $p\colon J \dashrightarrow N$ of $J$ along the image of the zero section.*

(2) *There is a divisorial contraction $h\colon N \to \bar{N}$ which contracts the proper transform of $\Theta$ onto an 8–dimensional variety which is birational to the LLSvS 8–fold $Z(X)$.*

*In other words, we have $\mathrm{Mov}(J) = \langle L, H\rangle = \mathrm{Nef}(J) \cup p^* \mathrm{Nef}(N)$, $\mathrm{Nef}(J) = \langle L, H_0\rangle$ and $p^* \mathrm{Nef}(N) = \langle H_0, H\rangle$, where $H_0$ is a big and nef line bundle on $J$ which contracts the zero section of $J \to \mathbb{P}^5$ and $H$ is as in Corollary 3.10.*

*For general $X$, the relative theta divisor $\Theta$ can be contracted after the Mukai flop of the zero section of $J \to \mathbb{P}^5$.*

Before the proof of the theorem, which will be given in Section 4.2, we mention, as a consequence of the theorem above, the relation between the intermediate Jacobian fibration and moduli spaces of objects in the Kuznetsov component of $X$.

### 4.1 Comparison with moduli spaces of objects in the Kuznetsov component of $X$

The recent paper [6] establishes the existence and the fundamental properties of moduli spaces of objects in the Kuznetsov component $\mathcal{K}u(X)$ of a smooth cubic fourfold $X$. We refer the reader to Section 29 of loc. cit. for the relevant definitions and the precise statements of the results.

Given a smooth cubic fourfold $X$, the extended Mukai lattice $\widetilde{H}(\mathcal{K}u(X), \mathbb{Z})$ is a lattice whose underlying group is the topological $K$–theory of $\mathcal{K}u(X)$ and whose Mukai pairing and weight-two Hodge structure are induced from those on $X$. The only classes in $\widetilde{H}(\mathcal{K}u(X), \mathbb{Z})$ that are of type $(1,1)$ for very general $X$ are contained in a rank-two lattice $A_2$, which is spanned by two classes $\lambda_1$ and $\lambda_2$ that satisfy $\lambda_1^2 = \lambda_2^2 = 2$ and $\lambda_1 \cdot \lambda_2 = -1$; see [6, equation (29.1)]. A description of a full connected component of the space of Bridgeland stability conditions on $\mathcal{K}u(X)$ is also produced; see Theorem 29.1 of loc. cit. It is shown that, for a primitive Mukai vector with $v^2 \geq -2$ and for a $v$–generic stability condition $\sigma$ in this component, the moduli space $M_\sigma(\mathcal{K}u(X), v)$ of Bridgeland stable objects in $\mathcal{K}u(X)$ with Mukai vector $v$ is a nonempty smooth projective hyper-Kähler manifold of dimension $v^2+2$, deformation equivalent to a Hilbert scheme of points on a K3 surface; moreover, the formation of these moduli spaces works in families; see Theorem 29.4 of loc. cit. for the precise statement.

For a Mukai vector of OG10–type in the $A_2$ lattice, ie of the form $v = 2\lambda$ with $\lambda^2 = 2$, in [51] it is shown that, for a $\lambda$–generic stability condition $\sigma$, the moduli space $M_\sigma(\mathcal{K}u(X), v)$ is an irreducible normal projective symplectic variety of dimension 10 admitting a symplectic resolution which is deformation equivalent to a manifold of OG10–type. The genericity condition here means that the polystable objects with Mukai vector $v$ are the direct sum of two stable objects with Mukai vector $\lambda$. More precisely, the singular locus of $M_\sigma(\mathcal{K}u(X), v)$ is isomorphic $\operatorname{Sym}^2 M_\sigma(\mathcal{K}u(X), \lambda)$.

Moreover, in [51] it is shown that for general $X$ the twisted intermediate Jacobian fibration $J^T(X)$ is birational to $M_\sigma(\mathcal{K}u(X), 2\lambda)$, for $\lambda^2 = 2$. For the nontwisted case we have the following corollary of Theorem 4.1 that goes in the opposite direction.

**Corollary 4.2** *For very general $X$, $J(X)$ is not birational to a moduli space of the form $M_\sigma(\mathcal{K}u(X), v)$.*

**Proof** First of all, by [6, Remark 29.3], if nonempty, the dimension of a moduli space $M_\sigma(\mathcal{K}u(X), v)$ is $v^2+2$. This dimension is equal to 10 if and only if either

$v$ is primitive (hence $M_\sigma(\mathcal{K}u(X), v)$ is of $\mathrm{K3}^{[5]}$–type and thus cannot be birational to $J(X)$) or else $v = 2\lambda$ with $\lambda^2 = 2$. By the results of [51] cited in the remark above, for $v = 2\lambda$ with $\lambda \in A_2$ and $\lambda^2 = 2$ and $\sigma$ a $\lambda$–generic stability condition, the singular locus of $M_\sigma(\mathcal{K}u(X), v)$ is isomorphic to the second symmetric product of a hyper-Kähler manifold of $\mathrm{K3}^{[2]}$–type. By dimension reasons, the symplectic resolution $\widetilde{M}_\sigma(\mathcal{K}u(X), v) \to M_\sigma(\mathcal{K}u(X), v)$ is not a small contraction. Suppose by contradiction that $J(X)$ is birational to $\widetilde{M}_\sigma(\mathcal{K}u(X), v)$. Then by Theorem 4.1, the symplectic resolution has to coincide with $N \to \overline{N}$, and $M_\sigma(\mathcal{K}u(X), v) \cong \overline{N}$. This implies that the singular locus of $M_\sigma(\mathcal{K}u(X), v)$ has to be birational to the Lehn–Lehn–Sorger–van Straten 8–fold $Z(X)$, which gives a contradiction. Indeed, $Z(X)$ cannot be birational to $\mathrm{Sym}^2 M_\sigma(\mathcal{K}u(X), \lambda)$ since, by Proposition 1.7, this would imply that the latter has a symplectic resolution. This, however, is not true because $\mathrm{Sym}^2 M_\sigma(\mathcal{K}u(X), \lambda)$ is a $\mathbb{Q}$–factorial symplectic variety with singular locus of codimension strictly greater than two and hence does not admit a symplectic resolution (since it does not admit a semismall resolution). $\square$

**Remark 4.3** We expect the more general statement to hold: for very general $X$, $J(X)$ is not birational to a Bridgeland moduli space of objects on a 2–CY category that is deformation equivalent to the derived category of a K3 surface. We present a rough sketch of the argument. Assume there is a family of Bridgeland stability conditions on the family of derived categories realizing the deformation. Then, as in [6, Theorem 21.24], a relative moduli space exists as an algebraic space; by a generalization of a theorem of Mukai [68, Theorem 1.4], the stable locus of each fiber is smooth and has a holomorphic symplectic form; the singular locus parametrizing strictly semistable objects of codimension $\geq 2$. One then expects such moduli spaces to be normal and irreducible. As in the proof of the projectivity in [6, Theorem 29.4] it follows that these moduli spaces are projective. Finally, a similar argument to the one above shows that the contraction $N \to \overline{N}$ cannot be the symplectic resolution of one of these moduli spaces.

In the next subsection we construct the degeneration of the intermediate Jacobian fibration that will allow us to prove Theorem 4.1. The proof of the theorem will be given at the end of the section.

## 4.2 Degeneration to the chordal cubic

The secant variety to the Veronese embedding of $\mathbb{P}^2$ in $\mathbb{P}^5$ is a cubic hypersurface isomorphic to $\mathrm{Sym}^2 \mathbb{P}^2$, called the chordal cubic. Such a singular cubic fourfold is unique up to the action of the projective linear group. Given a one-parameter family of cubic fourfolds degenerating to the chordal cubic, it was proved in [41] that, up to a base change, one can fill the corresponding degeneration of intermediate Jacobian fibrations with a smooth central fiber that is birational to $\widetilde{M}_{2v_0} = \widetilde{M}_{2v_0}(S)$, where $(S, C)$ is the degree-two $K3$ surface associated to the degeneration of cubic fourfolds as in [22; 34; 46], and where $v_0 = (0, C, -2)$ is as in (2-2). We will use this degeneration to study the birational properties of the intermediate Jacobian fibration, at least for general $X$. For this purpose, we need to control what happens to the line bundles $L$ and $\Theta$ under the corresponding degeneration of intermediate Jacobian fibrations. We achieve this by constructing a particular degeneration whose central fiber is precisely the singular moduli space $M_{2v_0}$ and is such that the Lagrangian fibrations of the members of this degeneration fit in a relative Lagrangian fibration. This is done in Proposition 4.5. With this degeneration, we are not only able to identify precisely the limits of $L$ and $\Theta$ (see Lemma 4.7), but we are also able to deform the results about the birational geometry of $M_{2v_0}$ away from the central fiber (see Proposition 4.9), eventually proving Theorem 4.1.

Let $\mathcal{X} \to \Delta$ be a one-parameter family of cubic fourfolds degenerating to the chordal cubic. By this we will mean that $\Delta$ is a small disk or an open affine subset in the base of a pencil of cubic fourfolds with the property that the general fiber is smooth and the central fiber is isomorphic to the chordal cubic. The following facts were proved in [34], see also [46] and [41]:

(a) The monodromy of this family has order two.

(b) To such a degeneration one can associate a degree-two polarized K3 surface $(S, C)$.

(c) For a general pencil, the polarized K3 surface $(S, C)$ is general in moduli.

Suppose that for $t \neq 0$ the cubic fourfold $\mathcal{X}_t$ is general in the sense of LSV — ie in the sense that the construction of the hyper-Kähler compactification of [47] works for $J_U(\mathcal{X}_t)$ — and let $\mathcal{J}^* \to \Delta^*$ be the family of intermediate Jacobians associated to the smooth locus $\mathcal{X}^* \to \Delta^*$ of the pencil, with corresponding family of Lagrangian fibrations $\pi_{\Delta^*}\colon \mathcal{J}^* \to \mathbb{P}^5_{\Delta^*}$.

**Lemma 4.4** [22; 41] *Up to a degree-two base change, we can extend* $\pi_{\Delta^*}\colon \mathcal{J}^* \to \mathbb{P}^5_{\Delta^*}$ *to a projective morphism* $\pi_{\mathcal{V}}\colon \mathcal{J}_{\mathcal{V}} \to \mathcal{V}$, *where* $\mathcal{V} \subset \mathbb{P}^5 \times \Delta$ *is an open subset such that* $\mathcal{V}_t = \mathbb{P}^5$ *for* $t \neq 0$ *and* $\mathcal{V}_0 \subset \mathbb{P}^5$ *is nonempty for* $t = 0$, *and where* $\mathcal{J}_0 \to \mathcal{V}_0 \subset \mathbb{P}^5$ *is* *identified with the restriction of* $M_{2v_0}(S) \to |2C| = \mathbb{P}^5$ (cf (2-3)) *to an open subset* $V \subset |2C|$. *Moreover,* $\mathcal{J}_{\mathcal{V}} \to \mathcal{V}$ *has a zero section and is polarized by a relative principal polarization.*

**Proof** Let $H \subset \mathbb{P}^5$ be a general hyperplane. For the degeneration $\mathcal{Y} := \mathcal{X} \cap (H \times \mathbb{P}^5)$ of a single smooth cubic threefold the statement is due to Collino [22]. In Proposition 1.16 of loc. cit., it is also shown that the class of the limit polarization is the theta divisor of the Jacobian of the genus-five hyperelliptic curve, which is the limiting abelian variety. For the statement about the limit of the intermediate Jacobian fibration, this is [41, Section 6.3]. □

We now compactify the projective family $\mathcal{J}_{\mathcal{V}}$ of the lemma above to construct a family of Lagrangian fibered holomorphic symplectic varieties in such a way that the central fiber is exactly $M_{2v_0} = M_{2v_0}(S)$ (or $\widetilde{M}_{2v_0} = \widetilde{M}_{2v_0}(S)$); cf (2-3).

**Proposition 4.5** *Let* $\mathcal{X} \to \Delta$ *be as above a general family of smooth cubic fourfolds degenerating to the chordal cubic. Suppose that for very general* $t \in \Delta$, $\mathcal{X}_t$ *is very general. Let* $(S, C)$ *be the corresponding* $K3$ *surface of degree two as above. Then, possibly up to a base change, there are two degenerations of the corresponding intermediate Jacobian fibration, fitting in the commutative diagram*

$$\begin{tikzcd} \widetilde{\mathcal{M}} \arrow[r, "m"] \arrow[dr, "\widetilde{f}"'] & \mathcal{M} \arrow[d, "f"] \\ & \Delta \end{tikzcd} \tag{4-1}$$

*where:*

(1) $\widetilde{f} : \widetilde{\mathcal{M}} \to \Delta$ *is a family of smooth hyper-Kähler manifolds, with* $\widetilde{\mathcal{M}}_t = J(\mathcal{X}_t)$ *for* $t \neq 0$ *and* $\widetilde{\mathcal{M}}_0 = \widetilde{M}_{v_0}(S)$. *The family is equipped with a relative Lagrangian fibration* $\widetilde{\mathcal{M}} \to \mathbb{P}^5_\Delta$, *where for each* $t$ *the corresponding Lagrangian fibration is the obvious one.*

(2) $f : \mathcal{M} \to \Delta$ *is a degeneration of hyper-Kähler manifolds, with* $\mathcal{M}_t = J(\mathcal{X}_t)$ *for* $t \neq 0$ *and* $\mathcal{M}_0 = M_{v_0}(S)$. *The morphism* $m : \widetilde{\mathcal{M}} \to \mathcal{M}$ *is proper and birational, for* $t \neq 0$ *it is an isomorphism and for* $t = 0$ *it is the natural symplectic resolution* $m_0 : \widetilde{M}_{2v_0}(S) \to M_{2v_0}(S)$ *of Theorem 2.1. Moreover, there is a relative Lagrangian fibration* $\mathcal{M} \to \mathbb{P}^5_\Delta$ *where for each* $t$ *the corresponding Lagrangian fibration is the obvious one.*

**Proof** Start from the projective morphism $\pi_{\mathcal{V}}\colon \mathcal{J}_{\mathcal{V}} \to \mathcal{V}$ of Lemma 4.4. There is an isomorphism $\mathcal{J}_{\mathcal{V}_0} \cong (\widetilde{M}_{2v_0})_V$, where $(\widetilde{M}_{2v_0})_V$ is the restriction of the Lagrangian fibration $\tilde{\pi}\colon \widetilde{M}_{2v_0} \to \mathbb{P}^5$ (cf (2-3)) to an open subset $V \subset \mathbb{P}^5$. Let $\overline{\mathcal{J}_{\mathcal{V}}} \to \mathbb{P}^5_\Delta$ be any projective morphism extending $\pi_{\mathcal{V}}$. Applying Theorem 1.19(2) to $\overline{\mathcal{J}_{\mathcal{V}}} \to \mathbb{P}^5_\Delta$ yields, possibly up to the base change, a family $\tilde{g}\colon \widetilde{\mathcal{J}} \to \Delta$ of smooth hyper-Kähler manifolds (projective over $(\Delta)^*$), with a relative Lagrangian fibration $\widetilde{\mathcal{J}} \to \mathbb{P}^5_\Delta$. Let $\mathcal{L}$ be the line bundle on $\widetilde{\mathcal{J}}$ inducing it on every fiber. Let

$$\phi_0\colon \widetilde{\mathcal{J}}_0 \dashrightarrow \widetilde{M}_{2v_0} \tag{4-2}$$

be the birational morphism induced by the isomorphism of open subsets $\mathcal{J}_{\mathcal{V}_0} \cong (\widetilde{M}_{2v_0})_V$. Then $(\phi_0)_*\mathcal{L}_0 = \tilde{\ell} := \tilde{\pi}^*\mathcal{O}_{\mathbb{P}^5}(1)$.

We now use an argument very similar to that in the proof of [41, Theorems 1.3 and 1.7], to construct a family which is isomorphic to $\mathcal{J}$ over $\Delta$ and whose central fiber is actually *isomorphic* to $\widetilde{M}_{2v_0}(S)$. Let $\Lambda$ be the OG10 lattice. Fixing a marking of the central fiber and trivializing the local system $R^2\tilde{g}_*\mathbb{Z}$ induces a marking $\eta_t\colon H^2(\widetilde{\mathcal{J}}_t, \mathbb{Z}) \to \Lambda$ of every fiber. Let $\mathcal{D} \subset \mathbb{P}(\Lambda \otimes_{\mathbb{Z}} \mathbb{C})$ be the period domain and let $\mathcal{P}\colon \Delta \to \mathcal{D}$ be the period mapping induced by these markings. Let $\rho_0 = \eta_0(\phi_0)^*\colon H^2(\widetilde{M}_{2v_0}, \mathbb{Z}) \to \Lambda$ be the induced marking on $\widetilde{M}_{2v_0}$. Let $\rho_t\colon H^2(\widetilde{\mathcal{M}}_t, \mathbb{Z}) \to \Lambda$ be markings induced by $\rho_0 = \eta_0(\phi_0)^*$ on fibers of the universal family over $\mathrm{Def}(\widetilde{M})$ and let $\mathcal{P}_{\widetilde{M}}\colon \mathrm{Def}(\widetilde{M}) \to \mathcal{D}$ be the induced period mapping. Since $\mathcal{P}_{\widetilde{M}}$ is a local isomorphism, we can lift $\mathcal{P}$ to a map $\xi\colon \Delta \to \mathrm{Def}(\widetilde{M}_{2v_0})$. Pulling back the universal family gives a family $\tilde{f}\colon \widetilde{\mathcal{M}} \to \Delta$ with central fiber $\widetilde{\mathcal{M}}_0 = \widetilde{M}_{2v_0}$. As in [41] the two families $\tilde{g}\colon \widetilde{\mathcal{J}} \to \Delta$ and $\tilde{f}\colon \widetilde{\mathcal{M}} \to \Delta$ are relatively birational over $\Delta$, since for every $t \in \Delta$, the marked pairs $(\widetilde{\mathcal{J}}_t, \eta_t)$ and $(\widetilde{\mathcal{M}}_t, \rho_t)$ are nonseparated points. To show that the two families $\widetilde{\mathcal{J}}$ and $\widetilde{\mathcal{M}}$ are isomorphic away from the central fiber, first recall that by [35, Theorem 4.3] (cf also [52, Theorem 3.2]), for every $t$ there exists an effective cycle

$$\Gamma_t = Z_t + \sum W_{i,t}$$

of pure dimension 10 in $\widetilde{\mathcal{M}}_t \times \widetilde{\mathcal{J}}_t$ such that $Z_t$ is the graph of a birational map, the codimension of the images of the $W_{i,t}$ in $\widetilde{\mathcal{M}}_t$ and in $\widetilde{\mathcal{J}}_t$ are equal and positive, and $[\Gamma_t]_*$ is a Hodge isometry and is equal to $\rho_t^{-1} \circ \eta_t\colon H^2(\widetilde{\mathcal{J}}_t, \mathbb{Z}) \to H^2(\widetilde{\mathcal{M}}_t, \mathbb{Z})$. Let $\widetilde{\mathcal{L}}$ be the line bundle on $\mathcal{M}$ such that $\widetilde{\mathcal{L}}_t = \rho_t^{-1}\eta_t(\mathcal{L}) = [\Gamma_t]_*(\mathcal{L}_t)$. Since $\widetilde{\mathcal{L}}_0 = \tilde{\pi}^*\mathcal{O}_{\mathbb{P}^5}(1)$ induces a Lagrangian fibration on $\widetilde{\mathcal{M}}_0 = \widetilde{M}_{2v_0}$, by [57] $\widetilde{\mathcal{L}}$ induces a Lagrangian fibration on $\mathcal{M}_t$ for every $t$ (maybe up to restricting $\Delta$). For very general $t$, $\widetilde{\mathcal{L}}_t = [Z_t]_*(\mathcal{L}_t)$, since the isotropic class $[Z_t]_*(\mathcal{L}_t)$ lies in the movable cone of $\widetilde{\mathcal{M}}_t$ and hence by

Corollary 3.10 it has to be equal to $\widetilde{\mathcal{L}}_t$. Corollary 3.4 implies that for very general $t$, $Z_t$ is the graph of an *isomorphism* between $\widetilde{\mathcal{J}}_t$ and $\widetilde{\mathcal{M}}_t$. The same countability argument as in the proof of [41, Theorems 1.3 and 1.7] shows that there exists a component of the Hilbert scheme parametrizing graphs of such cycles $Z_t \subset \mathcal{J}_t \times \mathcal{M}_t$ that dominates $\Delta$. It follows that there is a cycle $\mathcal{Z}$ in the fiber product $\widetilde{\mathcal{J}} \times_\Delta \widetilde{\mathcal{M}}$ which, maybe up to restricting $\Delta$, induces an isomorphism for $t \neq 0$. The conclusion is that the family $\widetilde{\mathcal{M}} \to \Delta$ is such that central fiber is $\widetilde{\mathcal{M}}_0 \cong \widetilde{M}_{2v_0}$ while for $t \neq 0$, we have $\widetilde{\mathcal{M}}_t \cong J(\mathcal{X}_t)$.

Now we construct the second family. By [62, Theorem 2.2] there is a finite morphism

$$\Xi\colon \mathrm{Def}(\widetilde{M}_{2v_0}) \to \mathrm{Def}(M_{2v_0})$$

induced by the symplectic resolution $m_0\colon \widetilde{M}_{2v_0} \to M_{2v_0}$ and compatible with the universal families on the two deformation spaces; for more details see Section 2 of loc. cit. Set $\nu = \Xi \circ \xi\colon \Delta \to \mathrm{Def}(M_{2v_0})$ and let

$$\mathcal{M} \to \Delta$$

be the pullback via $\nu$ of the universal family on $\mathrm{Def}(M_{2v_0})$. Then the birational map $m\colon \widetilde{\mathcal{M}} \to \mathcal{M}$ over $\Delta$ induced by [62, Theorem 2.2] has the desired properties.

Finally, the statement about the Lagrangian fibrations follows from the fact that, since the Lagrangian fibration $\widetilde{M}_{2v_0} \to \mathbb{P}^5$ in the central fiber factors via $\widetilde{M}_{2v_0} \to M_{2v_0}$, the morphism $\widetilde{\mathcal{M}} \to \mathbb{P}^5_\Delta$ factors via a morphism $\mathcal{M} \to \mathbb{P}^5_\Delta$. □

As a consequence of the last part of the proof, notice that there is a line bundle $\mathcal{L}_{\mathcal{M}}$ on $\mathcal{M}$ with

$$m^*\mathcal{L}_{\mathcal{M}} = \widetilde{\mathcal{L}}$$

and whose restriction to the central fiber satisfies $\mathcal{L}_{\mathcal{M}_0} = \ell$, where $\ell$ is as in (2-5).

For any $t \neq 0$, let $\Theta_t$ be the relative theta divisor in $\mathcal{M}_t = J(\mathcal{X}_t)$.

**Lemma 4.6** *For $\star = \widetilde{\mathcal{M}}$ or $\mathcal{M}$, let $\Theta_\star$ be the divisor defined as the closure of $\bigcup_{t\neq 0} \Theta_t$ in $\star$. Then, $\Theta_{\mathcal{M}}$ is a Cartier divisor and hence the following compatibility conditions hold* (notation as in diagram (4-1)):

$$\Theta_{\widetilde{\mathcal{M}}_0} \cong (m^*\Theta_{\mathcal{M}})_{|\widetilde{\mathcal{M}}_0} = m_0^*\Theta_{\mathcal{M}_0}, \tag{4-3}$$

*where $\Theta_{\star_0} := (\Theta_\star)_{|0}$ is the fiber of $\Theta_\star$ over $t = 0$.*

**Proof** Let $\mathcal{I}_{\Theta_{\mathcal{M}}} \subset \mathcal{O}_{\mathcal{M}}$ be the ideal sheaf of $\Theta_{\mathcal{M}}$ in $\mathcal{M}$. Since the morphism $\Theta_{\mathcal{M}} \to \Delta$ is flat, it follows that the restriction $(\mathcal{I}_{\Theta_{\mathcal{M}}})_{|\mathcal{M}_0}$ is the ideal sheaf of $\Theta_{\mathcal{M}_0}$ in $\mathcal{M}_0$. By [66] (cf Section 2.2), $\mathcal{M}_0 = M_{2v_0}$ is factorial so $(\mathcal{I}_{\Theta_{\mathcal{M}}})_{|\mathcal{M}_0}$ is locally free. Hence, so is $\mathcal{I}_{\Theta_{\mathcal{M}}}$. It follows that the divisors $\Theta_{\widetilde{\mathcal{M}}}$ and $m^*\Theta_{\mathcal{M}}$ agree and so do their central fibers. □

The next lemma identifies the limit of $\Theta_t$ in $M_{2v_0} = \mathcal{M}_0$ and shows that all line bundles on $M_{2v_0}$ deform over $\mathcal{M} \to \Delta$. Recall first that by (2-1), $\mathrm{NS}(M_{2v_0}) = U = \langle \ell, \theta \rangle$, and for every $t$,

$$\mathrm{NS}(\mathcal{M}_t) \supseteq U_t = \langle \mathcal{L}_t, \Theta_t \rangle,$$

with equality holding for very general $t$. Here $\Theta_0 = \Theta_{\widetilde{\mathcal{M}}_0}$. In particular, inside $\mathrm{NS}(\widetilde{M}_{2v_0})$ we have the following rank-two sublattices both of which are isomorphic to the hyperbolic lattice $U$: the limit lattice $U_0$ spanned by the limits $\mathcal{L}_0 = \widetilde{\ell}$ and $\Theta_0$, and the pullback lattice $m_0^* \mathrm{NS}(M_{2v_0}) = \langle m_0^*\ell, m_0^*\theta \rangle$.

**Lemma 4.7** *Let the notation be as above. Then*:

1. *The two sublattices $U_0 = \langle \widetilde{\ell}, \Theta_{\widetilde{\mathcal{M}}_0} \rangle$ and $\langle m_0^*\ell, m_0^*\theta \rangle$ of $\mathrm{NS}(\widetilde{M}_{2v_0})$ are the same.*
2. *The limit of the relative theta divisor in $\mathcal{M}_0$ is precisely $\theta$, the relative theta divisor on $M_{2v_0}(S)$ of (2-6).*

**Proof** By Lemmas 4.4 and 4.6, the limit theta divisor $(\Theta_{\mathcal{M}})_0$ is an effective line bundle on $\mathcal{M}_0 = M_{2v_0}$, which restricts to a theta divisor on the smooth fibers of $M_{2v_0} \to \mathbb{P}^5$. Thus $\Theta_{\mathcal{M}_0}$ is linearly equivalent to an effective line bundle of the form $\theta + a\ell$ for some integer $a$. We show that $a = 0$. By (4-3), $\Theta_{\widetilde{\mathcal{M}}_0} = m_0^*\Theta_{\mathcal{M}_0} = m_0^*(\theta + a\ell)$ and $\widetilde{\mathcal{L}}_0 = m_0^*\ell$. This is enough to conclude that the two sublattices

$$U = \langle \Theta_{\widetilde{\mathcal{M}}_0}, \widetilde{\mathcal{L}}_0 \rangle \quad \text{and} \quad U = m_0^*\langle \theta, \ell \rangle = m_0^* \mathrm{NS}(M_{2v_0})$$

of $\mathrm{NS}(\widetilde{M}_{2v_0})$ are the same. This proves the first part of the lemma. By Remark 2.5 above, $\theta$ does not contain the singular locus of $M$, thus $m_0^*\theta$ coincides with its proper transform and is irreducible. Since it has negative Beauville–Bogomolov square (cf (2-7)), it is a prime exceptional divisor. By [53, Section 5.1], a prime exceptional divisor deforms where its first Chern class remains algebraic. Thus $m_0^*\theta$ deforms to a relative effective prime exceptional divisor $\theta_{\widetilde{\mathcal{M}}}$ on $\widetilde{\mathcal{M}}$. By Corollary 3.4 and Proposition 3.6, for very general $t \neq 0$, the fiber over $t$ of the two irreducible effective divisors $\theta_{\widetilde{\mathcal{M}}}$ and $\Theta_{\widetilde{\mathcal{M}}}$ have to agree since there is only one prime exceptional divisor on $\mathcal{M}_t$. Thus $\theta_{\widetilde{\mathcal{M}}}$ and $\Theta_{\widetilde{\mathcal{M}}}$ have to be equal for every $t$. In particular, so are their restrictions to the central fiber. □

**Corollary 4.8** *Let $X$ be a general cubic fourfold and let $\pi\colon J(X) \to \mathbb{P}^5$ be the intermediate Jacobian fibration of* [47]. *The natural rational zero section of $\pi$ is regular.*

**Proof** Consider a degeneration of cubic fourfolds to the chordal cubic as in Proposition 4.5 and let $\mathcal{M} \to \Delta$ be the corresponding family. By Lemma 4.6 the divisor $\Theta_{\mathcal{M}}$ is Cartier and by Remark 2.8 it is relatively ample (up to restricting $\Delta$). Since $\Theta_{\mathcal{M}}$ is $-1$–invariant, it follows that the birational involution $-1$ is biregular. One component of the fixed locus of this involution has the property that its restriction to every fiber is precisely the closure of the corresponding rational zero section. Since by Remark 2.4, in the central fiber the section is regular, it follows that for general $t \in \Delta$ the corresponding rational section is also regular. □

Consider the family $\mathcal{M} \to \Delta$ of Proposition 4.5, with its relative theta divisor $\Theta_{\mathcal{M}}$. By Druel [26] we know that for every $t$, the prime exceptional divisor $\Theta_t$ can be contracted on a hyper-Kähler projective birational model of $\mathcal{M}_t$. In the central fiber $\mathcal{M}_t = M_{2v_0}$ we have, by Lemma 4.7, that $\Theta_{\mathcal{M}_0} = \theta$. By Lemma 2.6 this divisor can be contracted after a Mukai flop. We now show that the same is true for any $t \neq 0$, namely, that after a Mukai flop the relative theta divisor can be contracted, possibly up to restricting $\Delta$.

**Proposition 4.9** *For general $X$, the relative theta divisor $\Theta$ on $J = J(X)$ can be contracted after the Mukai flop of the zero section.*

**Proof** Let $\mathcal{M} \to \mathbb{P}^5_\Delta$ be as in Proposition 4.5. By Corollary 4.8, there is a relative zero section $s\colon \mathbb{P}^5_\Delta \to \mathcal{M}$. Let $T$ be its image. Then $T$ is contained in the smooth locus of the fibers of $\widetilde{\mathcal{M}} \to \Delta$. Let

$$P\colon \mathcal{M} \dashrightarrow \mathcal{N}$$

be the relative Mukai flop of $T$ in $\mathcal{M}$. By Lemma 2.6, the Mukai flop of the zero section in the central fiber $M_{2v_0}$ can be performed in the projective category. Thus, the central fiber of $\mathcal{N}$ is projective and so are all the fibers of $g\colon \mathcal{N} \to \Delta$ (since by Lemma 4.7 there is an ample class on the central fiber that deforms over $\Delta$). For $t \neq 0$, $\mathcal{N}_t$ is smooth while the central fiber $\mathcal{N}_0$ has the same singularities as $\mathcal{M}_0 = M_{2v_0}$, since they are isomorphic away from the flopped locus, which does not meet the singular locus. Via the birational morphism $P$, which is a relative isomorphism in codimension one, we can identify the second integral cohomology group of the fibers of the two families. In particular, for every $t \in \Delta$ we have $P_* U_t \subset \mathrm{NS}(\mathcal{N}_t)$, with equality holding for very general $t$ and for $t = 0$. In what follows we freely restrict $\Delta$, if necessary, without any mention. As in Lemma 2.6, let $H$ be the big and nef line bundle on $\mathcal{N}_0$ that contracts $\theta$, ie $H$ is a generator of the ray $\theta^\perp$. Since $H \in P_* U_0$, by Lemma 4.7, $H$ deforms to a line bundle $\mathcal{H}$ on $\mathcal{N}$. For very general $t$, its restriction $\mathcal{H}_t$ is a generator of the one-dimensional space $(P_t)_* \Theta_t^\perp \subset \mathrm{NS}(\mathcal{N}_t)$. By [27], $(P_t)_* \Theta_t$ can be contracted on a birational model of $\mathcal{N}_t$. We now show that it can be contracted on $\mathcal{N}_t$ itself. For very general $t$, the line bundle inducing the divisorial contraction has to be $\mathcal{H}_t$, or rather its proper transform on an appropriate birational model of $\mathcal{N}_t$. It follows that for very general $t$ (and thus for all $t$) $\mathcal{H}_t$ is big. Moreover, since $\mathcal{H}_0$ is big and nef and $\mathcal{N}_0$ has rational singularities, $H^i(\mathcal{N}_0, \mathcal{H}_0^k) = 0$ for $i > 0$ and any $k \geq 0$. It follows that the locally free sheaf $g_* \mathcal{H}^k$ satisfies base change. Since $\mathcal{H}_0$ is semiample, so is $\mathcal{H}_t$ for all $t$ in $\Delta$. For $k \gg 0$, the regular morphism $\Psi \colon \mathcal{N} \to \mathbb{P}(g^* g_* \mathcal{H}^k)$, relative over $\Delta$, is birational onto its image and contracts $\Theta_t$ for very general $t$ and for $t = 0$. Up to further restricting $t$, we can assume that the locus contracted on $\mathcal{N}_t$ is irreducible, and hence that $\Psi_t$ contracts precisely $(P_t)_* \Theta_t$ for every $t$. $\square$

The proof of Theorem 4.1 is now complete:

**Proof of Theorem 4.1** Let $X$ be general. By Proposition 4.9 the Mukai flop $p \colon J \dashrightarrow N$ of $J$ along the zero section is projective and on $J$ there exists a big and nef line bundle that contracts the zero section. For very general $X$, $H_0$ is unique, up to a positive rational multiple, and $\mathrm{Nef}(J) = \langle L, H_0 \rangle$. Moreover, we have shown that for general $X$ there is a divisorial contraction $N \to \bar{N}$, contracting $p_* \Theta$. Since the divisorial contraction $N \to \bar{N}$ contracts the ruling of $\Theta$ (cf Proposition 3.6), by (3-2) it follows that the image of $\Theta$ in $\bar{N}$ is birational to the LLSvS 8–fold $Z(X)$. For very general $X$, $\mathrm{Nef}(N) = \langle p_* H_0, p_* H \rangle$, where $p^* H$ is the unique (up to a positive multiple) big and nef line bundle inducing the contraction. By [36, Proposition 4.2], $H$ is the second ray of the movable cone of $J$, ie $\mathrm{Mov}(J) = \langle L, H \rangle$. $\square$

# 5 The Mordell–Weil group of $J(X)$

Let $a \colon \mathcal{A} \to B$ be a projective family of abelian varieties over an irreducible basis $B$ and suppose that $a$ admits a zero section. The Mordell–Weil group $\mathrm{MW}(a)$ of $a \colon \mathcal{A} \to B$ is the group of rational sections of $a \colon \mathcal{A} \to B$. Equivalently, if $K$ denotes the function field of $B$, $\mathrm{MW}(a)$ is the group of $K$–rational points of the generic fiber $\mathcal{A}_K$. For Lagrangian hyper-Kähler manifolds, the study of the Mordell–Weil group of abelian fibered hyper-Kähler manifolds was started by Oguiso in [65; 64]. The aim of this section is to prove the following theorem.

**Theorem 5.1** *Let $X$ be a smooth cubic fourfold and let $\pi\colon J = J(X) \to \mathbb{P}^5$ be as in Theorem 1.6, a smooth projective hyper-Kähler compactification of $J_U$. Let $\mathrm{MW}(\pi)$ be the Mordell–Weil group of $\pi$, ie the group of rational sections of $\pi$, and let $H^{2,2}(X,\mathbb{Z})_0$ be the primitive degree-four integral cohomology of $X$. The natural group homomorphism*

$$\phi_X : H^{2,2}(X,\mathbb{Z})_0 \to \mathrm{MW}(\pi)$$

*induced by the Abel–Jacobi map* (5-1) *is an isomorphism.*

**Corollary 5.2** *The group* $\mathrm{MW}(\pi)$ *is torsion-free.*

**Remark 5.3** In [64] Oguiso proved the existence of Lagrangian fibered hyper-Kähler manifolds whose Mordell–Weil group has rank 20. This is the maximal possible rank among all the known examples of hyper-Kähler manifolds, as follows from the Shioda–Tate formula of [65]; see also Proposition 5.4 below. Oguiso considers deformations of the abelian fibration $\widetilde{M}_{2v_0} \to \mathbb{P}^5$ (cf (2-3)) preserving both the Lagrangian fibration structure and the zero section; among these deformations, Oguiso shows the existence of Lagrangian fibration with rank 20 Mordell–Weil group [65, Theorem 1.4(2)]. The general deformation of $\widetilde{M}_{2v_0} \to \mathbb{P}^5$ for which both the Lagrangian fibration structure and the zero section are preserved (this is a codimension-two condition) is, up to birational isomorphism, $J(X)$; see Remark 3.3. By the theorem above, Lagrangian fibrations of the form $J(X)$, for $X$ with $\operatorname{rk} H^{2,2}(X,\mathbb{Z}) = 22$, satisfy $\operatorname{rk}\mathrm{MW}(\pi) = 20$. Thus, they provide an explicit description of Oguiso's examples.

The following proposition is essentially a reformulation of results from [65; 64].

**Proposition 5.4** *Let $\pi : M \to \mathbb{P}^n$ be a projective hyper-Kähler manifold with a fixed (rational) section. Let $K = \mathbb{C}(\mathbb{P}^n)$ be the function field of the base and let $M_K$ be the base change of $M$ to the generic point of $\mathbb{P}^n$. There is a commutative diagram*

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \mathbb{Z}L \oplus \bigoplus_i \mathbb{Z}D_i \arrow[r] \arrow[d, equal] & L^{\perp} \arrow[r] \arrow[d] & \mathrm{Pic}^0(M_K) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbb{Z}L \oplus \bigoplus_i \mathbb{Z}D_i \arrow[r] & \mathrm{NS}(M) \arrow[r, "r_K"] \arrow[d, two heads, "r_b"] & \mathrm{Pic}(M_K) \arrow[r] \arrow[d, two heads] & 0 \\
 & & \mathbb{Z} \arrow[r, equal] & \mathrm{NS}(M_K) &
\end{tikzcd}$$

*where $L = \pi^* \mathcal{O}_{\mathbb{P}^n}(1)$ and where the $D_1, \dots, D_k$ are the irreducible components of the complement of the regular locus of $\pi$ that do not meet the section. In particular,* $\mathrm{rk}(\mathrm{MW}(\pi)) = \mathrm{rk}(\mathrm{NS}(M)) - \mathrm{rk}\, \mathbb{Z}L \oplus \mathbb{Z}D_i - 1 = \mathrm{rk}(\mathrm{NS}(M)) - k$.

**Proof** The column on the left is exact by definition. By [76], for $b$ in the locus $U \subset \mathbb{P}^n$ parametrizing smooth fibers of $\pi$, $\mathrm{Im}[r_b : \mathrm{NS}(M) \to H^2(M_b)] = \mathbb{Z}$. The same argument as in Lemma 3.5 shows that a line bundle $D$ on $M$ lies in $L^\perp$ if and only if $D^n \cdot L^n = (D_{|M_b})^n = 0$. Since $\mathrm{rk}\, r_b = 1$, this holds if and only if $D \cdot L^n = D_{|M_b} = 0$, which is equivalent to $D \in \ker r_b$. This shows that the central column is exact. The same argument of [64, Theorem 1.1], which was used to show that $\mathrm{rk}\, \mathrm{NS}(M_K) = 1$, shows that any element in $\ker(r_b) = L^\perp$ goes to zero in $\mathrm{NS}(M_K)$. Thus there are induced horizontal morphisms $L^\perp \to \mathrm{Pic}^0(M_K)$ and $\mathbb{Z} \to \mathrm{NS}(M_K)$. Since $\mathrm{NS}(M) \to \mathrm{Pic}(M_K)$ is surjective, the bottom horizontal morphism is an isomorphism. The natural morphism $\mathbb{Z}L \oplus \mathbb{Z}D_i \to \mathrm{NS}(M)$ is injective, since by [65, Lemma 2.4] it has maximal rank over $\mathbb{Q}$ and $\mathrm{NS}(M)$ is torsion-free. Clearly, $\mathbb{Z}L \oplus_i \mathbb{Z}D_i \subset \ker(r_K)$. To show the reverse inclusion, let $D$ be any line bundle on $M$ that goes to zero in $\mathrm{Pic}(M_K)$. Then, by what we have already proved, for any smooth fiber we have $r_b(D) = [D_{|M_b}] = 0$. It follows that $D$ is a linear combination of $L = \pi^* \mathcal{O}_{\mathbb{P}^n}(1)$ and boundary divisors, ie $D \in \mathbb{Z}L \oplus_i \mathbb{Z}D_i$. As $\mathrm{rk}(\mathrm{MW}(\pi)) = \mathrm{rk}\, \mathrm{Pic}^0(M_K)$, the last statement also follows. □

**Remark 5.5** The study of the Mordell–Weil group for the Beauville–Mukai system is being carried out in joint work in progress with Chiara Camere.

**Corollary 5.6** *Let $J = J(X) \to \mathbb{P}^5$ be a hyper-Kähler compactification of the intermediate Jacobian fibration. Then*

$$\mathrm{rk}\, \mathrm{MW}(\pi) = \mathrm{rk}\, \mathrm{NS}(J) - 2 = \mathrm{rk}\, H^{2,2}(X, \mathbb{Z})_0.$$

**Proof** The discriminant locus of $\pi$ is irreducible and the fibers of $\pi$ over the general point of the discriminant are also irreducible; cf Lemma 1.2. Thus, in the notation of the proposition above, $\ker r_K = \mathbb{Z}L$ and the equality $\mathrm{rk}\, \mathrm{MW}(\pi) = \mathrm{rk}\, \mathrm{NS}(J) - 2$ follows. The remaining equality follows from Lemma 3.2. □

**Remark 5.7** The corollary just proven, which relies on Oguiso's Shioda–Tate formula above, is the only part of this section where we use that $J_U$ admits a hyper-Kähler compactification with a regular Lagrangian fibration extending $J_U \to U$. Indeed, to define the Abel–Jacobi map $\phi_X$ and to prove that it is injective (Section 5.3), we don't need to assume the existence of a hyper-Kähler compactification. However, we will use this corollary in the proof of the surjectivity (Section 5.4).

**Remark 5.8** An interesting problem is to study the action on $J$ of the birational automorphisms induced by translation by a nontrivial element of $\mathrm{MW}(\pi)$, as well as to study the automorphism group of the generic fiber $J_K$. A consequence of the observations of Section 3.1 is that if $J \to \mathbb{P}^5$ has irreducible fibers, then the birational automorphisms induced by translation are regular morphisms.

## 5.1 The Abel–Jacobi mapping

This sections uses some ingredients from the theory of normal functions (certain holomorphic sections of intermediate Jacobian fibrations), as developed and used by Griffiths [30; 31], Zucker [83; 84] and Voisin [78]. We refer to these papers, as well as to [77, Sections 7.2.1 and 8.2.2], for the relevant theory.

The first task is to define the morphism $\phi_X \colon H^{2,2}(X,\mathbb{Z})_0 \to \mathrm{MW}(\pi)$. One way to do this is to use relative Deligne cohomology, which allows us to define an algebraic section of the fibration $J_U \to U$. See, for example, [78; 28].

A more geometric way to define the morphism $\phi_X$ is in terms of algebraic cycles and Abel–Jacobi maps, which is what we use here. This is possible because the integral Hodge conjecture holds for degree-four Hodge classes on $X$ [78; 84]. It allows us to avoid, in the current presentation, defining the normal function associated with a cohomology class. The reader should keep in mind, however, that constructing an algebraic section of the intermediate Jacobian fibrations with a Hodge class on $X$ is a key ingredient in the proof of the Hodge conjecture of [78; 84], so the shortcut is only at the level of our presentation.

As already mentioned, the integral Hodge conjecture holds for degree-four Hodge classes on $X$. In particular, for every class $\alpha \in H^{2,2}(X,\mathbb{Z})$, there is an algebraic cycle $Z$ such that $[Z]=\alpha$. Let $V \subset \mathbb{P}^5$ be the open subset parametrizing smooth hyperplane sections of $X$ that do not contain any of the components of $Z$. If $\alpha$ is a primitive cohomology class, then for $b=[Y_b]\in V$, the one-cycle $Z_b$ satisfies

$$[Z_{Y_b}]=0 \quad \text{in } H^4(Y_b,\mathbb{Z})=\mathbb{Z},$$

and hence determines a point $\phi_{Y_b}(Z_b)\in \mathrm{Jac}(Y_b)$ in the intermediate Jacobian of $Y_b$. By Griffiths [31] (see also [77, Section 7.2.1]) the assignment

$$\sigma_Z \colon V \to J_V, \quad b \mapsto \phi_{Y_b}(Z_b),$$

defines a holomorphic section of the restriction of $J$ to $V \subset \mathbb{P}^5$. By [83], this section is, in fact, algebraic: indeed, consider a Lefschetz pencil $\mathcal{Y}' \to \mathbb{P}^1$ of hyperplanes of $X$

with $\mathbb{P}^1 \subset V$ and with the property that none of the singular points of the members of the pencil are contained in $Z$. By [83, Proposition (4.58)] the restriction of $\sigma_Z$ to the nonempty open subset $V \cap \mathbb{P}^1$ of the pencil extends to a holomorphic function on all of $\mathbb{P}^1$ and is thus algebraic; see also [28]. Since a holomorphic function that is algebraic in each variable is algebraic (see for example [14, Chapter IX, Theorem 5]), it follows that $\sigma_Z$ actually defines a rational function on $\mathbb{P}^5$, ie

$$\sigma_Z \in \mathrm{MW}(\pi).$$

The holomorphic section $\sigma_Z$ does not depend on the algebraic cycle representing $\alpha$. Indeed, since $\mathrm{CH}_0(X) = \mathbb{Z}$, by [79, Theorem 6.24] it follows that the cycle map $\mathrm{CH}^2(X) \to H^{2,2}(X, \mathbb{Z})$ is injective. It follows that if $Z$ and $Z'$ are homologous, then they are rationally equivalent in $X$ and hence so are their restrictions to a general smooth hyperplane section. The conclusion of this discussion is that the Abel–Jacobi map induces a well-defined group homomorphism

$$\phi_X \colon H^{2,2}(X, \mathbb{Z})_0 \to \mathrm{MW}(\pi), \quad \alpha = [Z] \mapsto \sigma_\alpha := \sigma_Z. \tag{5-1}$$

We prove injectivity of $\phi_X$ in Section 5.3 and surjectivity in Section 5.4. Since we will restrict to general pencils in $\mathbb{P}^5$, we start by recalling a few standard facts about Lefschetz pencils of cubic threefolds.

## 5.2 Preliminaries on Lefschetz pencils

We start by setting up the notation. Let $\mathbb{P}^1 \subset (\mathbb{P}^5)^\vee$ be a Lefschetz pencil with base locus a smooth cubic surface $\Sigma \subset X$. We have the diagram

$$\begin{tikzcd} & \Sigma \times \mathbb{P}^1 \arrow[r, hook, "i"] \arrow[dl, "p_1"'] & \mathcal{Y}' \arrow[dl, "p"'] \arrow[dr, "q"] & \\ \Sigma \arrow[r, hook] & X & & \mathbb{P}^1 \end{tikzcd}$$

where $\mathcal{Y}' = \mathrm{Bl}_\Sigma X$, $q \colon \mathcal{Y}' \to \mathbb{P}^1$ is the fibration of threefolds, and $i \colon \Sigma \times \mathbb{P}^1 \to \mathcal{Y}'$ is the inclusion of the exceptional divisor in $\mathcal{Y}'$. Let $j \colon U' \subset \mathbb{P}^1$ be the open subset parametrizing smooth fibers.

The following lemma is standard. We include a proof for lack of reference.

**Lemma 5.9** *The homology and cohomology groups of a cubic threefold which is smooth or has one $A_1$ singularity have no torsion. Moreover, using notation as above,*

$$R^1 q_* \mathbb{Z} = 0, \quad R^2 q_* \mathbb{Z} = \mathbb{Z}, \quad R^3 q_* \mathbb{Z} = j_* j^* R^3 q_* \mathbb{Z}, \quad R^4 q_* \mathbb{Z} = \mathbb{Z}.$$

**Proof** The statement about the homology groups of a cubic threefold with at most an $A_1$ singularity follow from [24, Example 5.3 and Theorem 2.1]; using the universal coefficient theorem, the statements on the cohomology groups then follow. From loc. cit. it also follows that $H^4(Y,\mathbb{Z}) = H_4(Y,\mathbb{Z})^\vee = \mathbb{Z}$, and hence $R^4 q_*\mathbb{Z} = \mathbb{Z}$ follows by proper base change. The first two statements on the higher direct images follow from the Lefschetz hyperplane section theorem. The third equality, which is also known as the "local invariant cycle" property, is well known to hold with $\mathbb{Q}$–coefficients, and we now show it with $\mathbb{Z}$–coefficients, as follows. By adjunction, there is a natural morphism

$$\varepsilon\colon R^3 q_*\mathbb{Z} \to j_* j^* R^3 q_*\mathbb{Z},$$

which is an isomorphism over $U$. To show $\varepsilon$ is an isomorphism over any point of $B := \mathbb{P}^1 \setminus U'$ we restrict, for every $b_0 \in B$, to a small disk $\Delta$ centered at $b_0$. Then $\varepsilon$ is an isomorphism around $b_0$ if and only if the specialization morphism

$$H^3(Y_{b_0},\mathbb{Z}) \cong H^3(\mathcal{Y}',\mathbb{Z}) \to H^3(Y_b,\mathbb{Z})^{\mathrm{inv}} = (j_* j^* R^3 q_*\mathbb{Z})_{b_0}$$

is an isomorphism (cf [69, pages 439–440]), where $b \in \Delta \cap U'$ and $H^3(Y_b,\mathbb{Z})^{\mathrm{inv}} \subset H^3(Y_b,\mathbb{Z})$ are the local monodromy invariants. Let $\delta \in H_3(Y_b,\mathbb{Z})$ be the vanishing cycle of $\mathcal{Y}'_\Delta$. By the Picard–Lefschetz formula, $H^3(Y_b,\mathbb{Z})^{\mathrm{inv}} = \mathbb{Z}\delta^\perp$, where $\perp$ is taken with respect to the intersection product, which is nondegenerate since $H^3(Y_b,\mathbb{Z})$ is torsion-free. By [77, Corollary 2.17], there is a short exact sequence

$$0 \to \mathbb{Z}\delta \to H_3(Y_b,\mathbb{Z}) \to H_3(\mathcal{Y}'_\Delta,\mathbb{Z}) \cong H_3(Y_{b_0},\mathbb{Z}) \to 0,$$

where $0 \neq \delta \in H_3(Y_b,\mathbb{Z})$ is the class of the vanishing cycle. Dualizing, we get a short exact sequence

$$0 \to H^3(Y_{b_0},\mathbb{Z}) \to H^3(Y_b,\mathbb{Z}) \to (\mathbb{Z}\delta)^\vee \to 0.$$

(Recall the absence of torsion in the homology groups of $Y_b$ and $Y_{b_0}$.) Using the isomorphism $H_3(Y_b,\mathbb{Z}) \cong H^3(Y_b,\mathbb{Z})$ induced by Poincaré duality, we make the identification $\mathbb{Z}\delta^\perp = \ker[H^3(Y_b,\mathbb{Z}) \to \mathbb{Z}\delta^\vee] = \mathrm{Im}[H^3(Y_{b_0},\mathbb{Z}) \to H^3(Y_b,\mathbb{Z})]$. □

It is well known that for a Lefschetz pencil the Leray spectral sequence with coefficients in $\mathbb{Q}$ degenerates at $E_2$. For a Lefschetz pencil of cubic threefolds, this is true also for $\mathbb{Z}$ coefficients. Again, we include a proof for lack of reference. For the whole family of smooth hyperplane sections of $X$ the Leray spectral sequence with integers coefficients does *not* degenerate at $E_2$; this is the starting point of the construction of the nontrivial $J_U$–torsor of [80], cf Remark 1.14.

**Lemma 5.10** *Let $q\colon \mathcal{Y}' \to \mathbb{P}^1$ be as above. The Leray spectral sequence with $\mathbb{Z}$ coefficients degenerates at $E_2$. In particular, the Leray filtration on $H^4(\mathcal{Y}', \mathbb{Z})$ is given by*

$$\begin{gathered}\mathbb{Z} = H^2(\mathbb{P}^1, R^2 f_*\mathbb{Z}) \subset L_1 \subset H^4(\mathcal{Y}', \mathbb{Z}) \twoheadrightarrow H^0(\mathbb{P}^1, R^4 f_*\mathbb{Z}) = \mathbb{Z}, \\ 0 \to H^2(\mathbb{P}^1, R^2 f_*\mathbb{Z}) \to L_1 \xrightarrow{\gamma} H^1(\mathbb{P}^1, R^3 f_*\mathbb{Z}) \to 0.\end{gathered} \tag{5-2}$$

**Proof** Because of the many vanishings in the $E_2$–page of the spectral sequence, the only map we need to show is trivial is $H^0(\mathbb{P}^1, R^4q_*\mathbb{Z}) \to H^2(\mathbb{P}^1, R^3q_*\mathbb{Z})$. For this, it is enough to show that $H^4(Y_b, \mathbb{Z}) \to H^0(\mathbb{P}^1, R^4q_*\mathbb{Z})$ is surjective, which is clearly true since both groups are generated by the class of a line. □

Consider the decomposition

$$H^4(\mathcal{Y}', \mathbb{Z}) = H^4(X, \mathbb{Z}) \oplus H^2(\Sigma, \mathbb{Z}) \oplus H^0(\Sigma, \mathbb{Z}) \tag{5-3}$$

given by the blowup formula. The inclusion of the first summand is given by the pullback $p^*$; we freely omit the symbol $p^*$ when viewing $H^4(X, \mathbb{Z})$ as a subspace of $H^4(\mathcal{Y}', \mathbb{Z})$. The inclusion of the second factor is via the map $H^2(\Sigma, \mathbb{Z}) \to H^4(\mathcal{Y}', \mathbb{Z})$ given by $C \mapsto i_*(C \times \mathbb{P}^1)$. Finally, the inclusion of the last summand is through the map $H^0(\Sigma, \mathbb{Z}) = H^0(\Sigma, \mathbb{Z}) \otimes H^2(\mathbb{P}^1, \mathbb{Z}) \to H^4(\mathcal{Y}', \mathbb{Z})$ that sends $[\Sigma] = [\Sigma \times p] \mapsto i_*([\Sigma \times p])$, where $p \in \mathbb{P}^1$ is a point. We highlight the following results for later use.

**Lemma 5.11** *There is a natural isomorphism $H^0(\Sigma, \mathbb{Z}) \cong H^2(\mathbb{P}^1, R^2q_*\mathbb{Z})$ which allows the identification of the inclusion*

$$H^0(\Sigma, \mathbb{Z}) \cong H^0(\Sigma, \mathbb{Z}) \otimes H^2(\mathbb{P}^1, \mathbb{Z}) \xrightarrow{i_*} H^4(\mathcal{Y}', \mathbb{Z})$$

*of* (5-3) *with the inclusion $H^2(\mathbb{P}^1, R^2q_*\mathbb{Z}) \to H^4(\mathcal{Y}', \mathbb{Z})$ induced by the Leray filtration of Lemma 5.10.*

**Proof** The closed embedding $i\colon \Sigma \times \mathbb{P}^1 \hookrightarrow \mathcal{Y}'$ determines an isomorphism $p_{2*}\mathbb{Z} \cong R^2q_*\mathbb{Z}$ of constant local systems. Here, $p_2\colon \Sigma \times \mathbb{P}^1 \to \mathbb{P}^1$ is the projection onto the section factor. Since $H^2(\mathbb{P}^1, p_{2*}\mathbb{Z}) = H^0(\Sigma, \mathbb{Z}) \otimes H^2(\mathbb{P}^1, \mathbb{Z})$ the lemma follows. □

Via $p^*$, we identify $H^4(X, \mathbb{Z})_0 \cong L_1 \cap H^4(X, \mathbb{Z})$ and set $L_1^{2,2} = L_1 \cap H^{2,2}(\mathcal{Y}', \mathbb{Z})$. Here $L_1 \subset H^4(\mathcal{Y}', \mathbb{Z})$ denotes the second piece of the Leray filtration; see (5-2).

**Corollary 5.12** *The surjective morphism $\gamma\colon L_1 \to H^1(\mathbb{P}^1, R^3q_*\mathbb{Z})$ of* (5-2) *restricts to an injection*

$$\overline{\gamma}\colon L_1 \cap (H^{2,2}(X, \mathbb{Z}) \oplus H^2(\Sigma, \mathbb{Z})) \cong L_1^{2,2}/\ker(\gamma) \to H^1(\mathbb{P}^1, R^3q_*\mathbb{Z}).$$

**Proof** From Lemmas 5.10 and 5.11 above, it follows that $\ker(\gamma) = H^2(\mathbb{P}^1, R^2 q_*\mathbb{Z}) = H^0(\Sigma, \mathbb{Z})$. Thus, by (5-3), it follows that $H^4(X,\mathbb{Z}) \oplus H^2(\Sigma,\mathbb{Z})) \cap \ker(\gamma) = \{0\}$. Since $H^0(\Sigma,\mathbb{Z}) \subset L_1$ and

$$L_1 \cap \bigl(H^{2,2}(X,\mathbb{Z}) \oplus H^2(\Sigma,\mathbb{Z}) \oplus H^0(\Sigma,\mathbb{Z})\bigr) = L_1 \cap \bigl(H^{2,2}(X,\mathbb{Z}) \oplus H^2(\Sigma,\mathbb{Z})\bigr) \oplus H^0(\Sigma,\mathbb{Z}), \tag{5-4}$$

the corollary follows. □

**Lemma 5.13** *The restriction morphism $H^1(\mathbb{P}^1, R^3 q_*\mathbb{Z}) \to H^1(U', R^3 q_{U'*}\mathbb{Z})$ is injective.*

**Proof** The Leray spectral sequence for the open immersion $j: U' \to \mathbb{P}^1$, applied to the sheaf $j^* R^3 q_*\mathbb{Z} = R^3 q_{U'*}\mathbb{Z}$, gives a five-term exact sequence starting with

$$0 \to H^1(\mathbb{P}^1, j_* j^* R^3 q_*\mathbb{Z}) \to H^1(U', R^3 q_{U'*}\mathbb{Z}) \to \cdots.$$

This concludes the proof, since by Lemma 5.9, $R^3 q_*\mathbb{Z} = j_* j^* R^3 q_*\mathbb{Z}$. □

## 5.3 Injectivity of $\phi_X$

The proof of injectivity uses the Hodge class of a normal function; see [83] and [77, Section 8.2.2].

For a pencil $\mathcal{Y}' \to \mathbb{P}^1$ as above, set

$$H^{2,2}(\mathcal{Y}',\mathbb{Z})_0 := L_1^{2,2} = \ker[H^{2,2}(\mathcal{Y}',\mathbb{Z}) \to H^0(\mathbb{P}^1, R^4 q_*\mathbb{Z})] = L_1 \cap H^{2,2}(\mathcal{Y}',\mathbb{Z}),$$

and let

$$\pi' = J' \to \mathbb{P}^1 \quad \text{and} \quad J_{U'} \to U'$$

be the restriction of the intermediate Jacobian fibration to $\mathbb{P}^1$ and to $U'$. Choosing a set of generators for $H^{2,2}(X,\mathbb{Z})_0$, let $\mathcal{Y}' \to \mathbb{P}^1$ be a general enough pencil that the restriction morphism

$$\phi'_X : H^{2,2}(X,\mathbb{Z})_0 \to \mathrm{MW}(\pi') \tag{5-5}$$

is well-defined. Here, $\mathrm{MW}(\pi')$ is the group of rational sections of $\pi'$. Similarly, we get a group homomorphism $\phi'_{\mathcal{Y}'} : H^{2,2}(\mathcal{Y}',\mathbb{Z})_0 \to \mathrm{MW}(\pi')$. Moreover, if $\alpha \in H^{2,2}(X,\mathbb{Z})_0$, then

$$\phi'_{\mathcal{Y}'}(p^*\alpha) = \phi'_X(\alpha) \in \mathrm{MW}(\pi').$$

Recall the Hodge class of a normal function; see for instance [77, Section 8.2.2] and [83, Proposition (3.9)]. Let $\mathcal{H}^3 = R^3 q_{U'*}\mathbb{Z} \otimes_{\mathbb{Z}} \mathcal{O}_{U'}$ be the Hodge bundle associated to the weight-three variation of Hodge structure of the pencil and let $F^*\mathcal{H}^3$ be the Hodge filtration. The sheaf $\mathcal{J}_{U'}$ of holomorphic sections of the intermediate Jacobian fibration fits into the exact sequence

$$0 \to R^3 f'_{U'*}\mathbb{Z} \to \mathcal{H}^3/F^2\mathcal{H}^3 \to \mathcal{J}_{U'} \to 0,$$

and the coboundary morphism

$$H^0(U', \mathcal{J}_{U'}) \xrightarrow{\mathrm{cl}} H^1(U', R^3 f_*\mathbb{Z}), \quad \nu \mapsto \mathrm{cl}(\nu),$$

associates to every holomorphic section $\nu$ of $J_{U'} \to U'$ a class $\mathrm{cl}(\nu)$ in $H^1(U', R^3q_*\mathbb{Z})$, called the Hodge class of $\nu$ — in the present context, this class is of Hodge type with respect to the Hodge structure on $H^1(U', R^3q_*\mathbb{Z})$ induced from that on $H^4(\mathcal{Y}', \mathbb{Z})$ via the degeneracy of the Leray spectral sequence; see [83, Section 3].

**Lemma 5.14** *Let $\mathcal{Y}' \to \mathbb{P}^1$ be a general pencil. The homomorphism*

$$\phi'_X \colon H^{2,2}(X, \mathbb{Z})_0 \xrightarrow{\beta} \mathrm{MW}(\pi')$$

*of* (5-5) *is injective.*

**Proof** By [83, Proposition (3.9)] — see also [77, Lemma 8.20] — the diagram

$$\begin{tikzcd} H^{2,2}(X,\mathbb{Z})_0 \arrow[r, "p^*"] \arrow[dr, "\phi'_X"'] & H^{2,2}(\mathcal{Y}',\mathbb{Z})_0 \arrow[r, "\gamma"] \arrow[d, "\phi'_{\mathcal{Y}'}"'] & H^1(\mathbb{P}^1, R^3p_*\mathbb{Z}) \arrow[d, "\varepsilon"] \\ & H^0(U', \mathcal{J}_{U'}) \arrow[r, "\mathrm{cl}"'] & H^1(U', R^3 f'_*\mathbb{Z}) \end{tikzcd} \tag{5-6}$$

is commutative. The map $\varepsilon$ is injective by Lemma 5.13, and $p^* \circ \gamma$ is injective by Corollary 5.12. Hence, $\mathrm{cl} \circ \phi'_X$ is injective and thus so is $\phi'_X$. $\square$

## 5.4 Surjectivity of $\phi_X$

There are three ingredients in the proof of surjectivity:

- the fact that $\mathrm{rk}\,\mathrm{MW}(\pi) = \mathrm{rk}\, H^{2,2}(X, \mathbb{Z})_0$, as proved in Corollary 5.6;
- the restriction, once again, to Lefschetz pencils;
- the techniques used in [78; 84] for the proof of the integral Hodge conjecture for cubic fourfolds.

We remark that we use their argument in a slightly different way. To prove the Hodge conjecture one starts with a cohomology class, uses it to define a normal function, and then uses the normal function to construct an algebraic cycle representing the cohomology class (possibly up to a multiple of a complete intersection surface). See [78] for more details. Here we start with a rational section of the intermediate Jacobian fibration, we restrict to a general pencil, and use the same method of Voisin to construct an algebraic cycle inducing the section via the Abel–Jacobi map. Then we have to check that the cohomology class representing this cycle is primitive, that it is independent of the pencil, and that it induces, via $\phi_X$, the section we started from.

Since by Corollary 5.6 the cokernel of the injection $\phi_X\colon H^{2,2}(X,\mathbb{Z})_0 \to \mathrm{MW}(\pi)$ is finite, for any $\sigma \in \mathrm{MW}(\pi)$ there is an integer $N$ and a cohomology class $\alpha \in H^{2,2}(X,\mathbb{Z})_0$ such that

$$\sigma_\alpha := \phi_X(\alpha) = N\sigma. \tag{5-7}$$

We will show, again using Lefschetz pencils, that given $\sigma$ and $\alpha$ as above, there exists a $\bar{\beta}' \in H^{2,2}(X,\mathbb{Z})_0$ such that $\alpha = N\bar{\beta}'$. This will give the desired surjectivity. Before we do so, let us introduce some results that we will need.

For a general pencil $\mathcal{Y}' \to \mathbb{P}^1$, let

$$(J^T)' \to \mathbb{P}^1$$

be the restriction of the intermediate Jacobian fibration $J^T \to \mathbb{P}^5$ of [80] (compare with Remark 1.14) to the pencil. For a conic $C \subset \Sigma$, consider the relative one-cycle of degree two in $\mathcal{Y}' \to \mathbb{P}^1$ — any other degree-two relative one-cycle that comes from $\Sigma$ will do. This defines a section of $(J^T)' \to \mathbb{P}^1$, which trivializes the torsor $(J^T)_{U'}$ inducing an isomorphism $J'_{U'} \cong (J^T)_{U'}$. It is easily seen that this extends to an isomorphism $t_C\colon J' \cong (J^T)'$ over $\mathbb{P}^1$. For any $\sigma' \in H^0(U', \mathcal{J}_{U'})$, we may consider the induced section

$$(\sigma^T)' := t_C \circ \sigma' \in H^0(U', \mathcal{J}^T_{U'}).$$

The following result is proved in Voisin [78]; see also [84, Theorem (3.2)], where the result is proved over $\mathbb{Q}$.

**Proposition 5.15** [78, Section 2.3] *For any section $\sigma' \in \mathrm{MW}(\pi')$, there is a relative one-cycle $Z$ on $\mathcal{Y}'$ of degree two such that the cohomology class*

$$\beta' = [Z] - [C \times \mathbb{P}^1] \in H^{2,2}(\mathcal{Y}', \mathbb{Z})_0$$

*satisfies $\phi'_{\mathcal{Y}'}(\beta') = \sigma'$ in $\mathrm{MW}(\pi')$.*

**Proof** For the reader's convenience, we give a brief sketch of the argument. By a result of Markushevich and Tikhomirov [54] and Druel [26] there is a relative birational morphism $c_2\colon \mathcal{M}'_{U'} \to J^T_{U'}$, where $\mathcal{M}'_{U'} \to U'$ is the relative moduli space of sheaves on $\mathcal{Y}'_{U'} \to U'$ with $c_1 = 0$ and $c_2 = 2\ell$. The morphism associates to every sheaf corresponding to a point in $\mathcal{M}'_{U'}$ the Abel–Jacobi invariant of its second Chern class. Given a section $(\sigma^T)' \in H^0(U', \mathcal{J}^T_{U'})$ as above, Voisin uses $\mathcal{M}'_{U'} \to U'$ to construct a family $\mathcal{C}_{U'}$ of degree-two curves in the fibers of $\mathcal{Y}'_{U'} \to U'$ with the property that for every $b \in U'$, the curve $\mathcal{C}_b$ represents the $c_2$ of a sheaf over $(\sigma^T)'(b)$. By construction, letting $Z$ be the closure of $\mathcal{C}_{U'}$ in $\mathcal{Y}'$ and setting $\beta' := [Z] - [C \times \mathbb{P}^1] \in H^{2,2}(\mathcal{Y}', \mathbb{Z})_0$, we have $\phi'_{\mathcal{Y}'}(\beta') = \sigma'$ in $H^0(U', \mathcal{J}_{U'})$. □

Let $\sigma \in \mathrm{MW}(\pi)$. For a general pencil $\mathbb{P}^1 \subset \mathbb{P}^5$, let $\sigma' = \sigma_{|\mathbb{P}^1}$ be the restriction of $\sigma$ to $\mathbb{P}^1$, and let $\beta'$ be as in the proposition above so that $\phi'_{\mathcal{Y}'}(\beta') = \sigma'$. It is tempting to say that, via $\phi_X$, the class $\beta'$ induces $\sigma$ globally and not just on that pencil. This is indeed the case, though we first need to check that $\beta'$ lies in the primitive cohomology of $X$ and that $\beta'$ is independent of the pencil as well as of the chosen isomorphism $t_C\colon J' \cong (J^T)'$. More precisely, we need to check that $\beta'$ induces $\sigma$ over an open subset of $\mathbb{P}^5$ and not just on the chosen pencil. Before checking this, we have the following proposition.

Recall that we have set $H^{2,2}(\mathcal{Y}', \mathbb{Z})_0 = L_1 \cap H^{2,2}(\mathcal{Y}', \mathbb{Z})$.

**Proposition 5.16** [83, Theorem (4.17)] *The Abel–Jacobi morphism*

$$\phi_{\mathcal{Y}'}\colon H^{2,2}(\mathcal{Y}', \mathbb{Z})_0 \to \mathrm{MW}(\pi') \subset H^0(U', \mathcal{J}_{U'})$$

*is surjective and defines an isomorphism*

$$\overline{\phi}_{\mathcal{Y}'}\colon L_1 \cap (H^{2,2}(X, \mathbb{Z}) \oplus H^2(\Sigma, \mathbb{Z})) \to \mathrm{MW}(\pi').$$

**Proof** By diagram (5-6) and the fact that $\varepsilon$ is injective, $\ker(\phi_{\mathcal{Y}'}) = \ker\gamma$, which by Lemma 5.10 is equal to $H^0(\Sigma, \mathbb{Z})$. Since $\phi_{\mathcal{Y}'}$ is surjective by the proposition above, the induced morphism $\overline{\phi}_{\mathcal{Y}'}\colon H^{2,2}(\mathcal{Y}', \mathbb{Z})_0/H^0(\Sigma, \mathbb{Z}) \to \mathrm{MW}(\pi')$ is an isomorphism. Finally, by (5-4), $H^{2,2}(\mathcal{Y}', \mathbb{Z})_0/H^0(\Sigma, \mathbb{Z}) \cong L_1 \cap (H^{2,2}(X, \mathbb{Z}) \oplus H^2(\Sigma, \mathbb{Z}))$. □

We can now end the proof of surjectivity. For $\sigma \in \mathrm{MW}(\pi)$, let $\alpha \in H^{2,2}(X, \mathbb{Z})_0$ be as in (5-7). Restricting to a pencil $\mathcal{Y}' \to \mathbb{P}^1$, set $\sigma' = \sigma_{|\mathbb{P}^1}$ and let $\beta'$ be as in Proposition 5.15 such that $\phi_{\mathcal{Y}'}(\beta') = \sigma'$. Finally, let $\overline{\beta}'$ be the projection of $\beta'$ onto

$L_1 \cap (H^{2,2}(X,\mathbb{Z}) \oplus H^2(\Sigma,\mathbb{Z}))$. In an abuse of notation, we are omitting $p^*$ from the inclusion of $H^4(X,\mathbb{Z})$ in $H^4(\mathcal{Y}',\mathbb{Z})$ and we will write $\alpha$ instead of $p^*\alpha$. We have

$$\phi_{\mathcal{Y}'}(\alpha) = (\phi_X(\alpha))_{|\mathbb{P}^1} = N\sigma' = N\phi_{\mathcal{Y}'}(\beta') = \phi_{\mathcal{Y}'}(N\beta').$$

By Proposition 5.16, $\alpha = N\overline{\beta}' \in L_1 \cap (H^{2,2}(X,\mathbb{Z}) \oplus H^2(\Sigma,\mathbb{Z}))$. Since

$$\alpha \in H^{2,2}(X,\mathbb{Z})_0 \subset L_1 \cap \bigl(H^{2,2}(X,\mathbb{Z}) \oplus H^2(\Sigma,\mathbb{Z})\bigr),$$

it follows that $\overline{\beta}'$, too, has to lie in $H^{2,2}(X,\mathbb{Z})_0 \subset H^{2,2}(\mathcal{Y}',\mathbb{Z})$. Moreover, the class $\overline{\beta}'$, which a priori depends on the chosen Lefschetz pencil, is independent of the pencil. Set $\sigma_{\overline{\beta}'} = \phi_X(\overline{\beta}')$. Then, for *any* sufficiently general Lefschetz pencil $\mathbb{P}^1 \subset \mathbb{P}^5$, we have an equality of sections

$$(\sigma_{\overline{\beta}'})_{|\mathbb{P}^1} = \sigma_{|\mathbb{P}^1},$$

and hence the two rational sections $\sigma_{\overline{\beta}'}$ and $\sigma$ coincide. This proves surjectivity.

## Appendix On the Beauville conjecture for LSV varieties by Claire Voisin

In this appendix we explain a consequence of Corollary 3.10 on the following conjecture made by Beauville in [10].

**Conjecture A.1** Let $M$ be a projective hyper-Kähler manifold. Any polynomial cohomological relation $P(d_1,\dots,d_r)=0$ in $H^*(M,\mathbb{Q})$, where $d_i$ are divisor classes on $M$, already holds in $\mathrm{CH}(M)$.

Here $\mathrm{CH}(M)$ denotes the Chow groups of $M$ with rational coefficients. Let now $M \to B$ be a projective hyper-Kähler manifold of dimension $2n$ equipped with a Lagrangian fibration, and let $L \in \operatorname{Pic} M = \mathrm{NS}(M)$ be the Lagrangian class pulled back from $B$; see [55]. We have $q(L)=0$ by the Beauville–Fujiki relations, since $L^{2n}=0$. Let also $h \in \operatorname{Pic} M = \mathrm{NS}(M)$ be the class of an ample divisor on $M$, so that the intersection pairing $q$ restricted to $\langle L,h\rangle$ is nondegenerate by the Hodge index theorem. The same argument as in [15] shows that the polynomial cohomological relations between $L$ and $h$ are generated by the relations

(A-1) $$\alpha^{n+1} = 0 \quad \text{in } H^{2n+2}(M,\mathbb{Q}) \text{ when } q(\alpha)=0 \text{ for } \alpha \in \langle L,h\rangle.$$

Here we can restrict to rational cohomology classes because we know that there is an isotropic class in $\langle L,h\rangle$. We consider now, more specifically, an LSV variety $J$

(which is of dimension 10, so $n = 5$) constructed in [47] as a Lagrangian fibration over $\mathbb{P}^5$. The Picard group of a very general such variety is a rank-two lattice which contains as above the Lagrangian class $L$ and an ample class, but we take as a basis the classes $L$ and $\Theta$, where $\Theta$ was introduced in [47] and is studied in the present paper. Riess proved in [72] that a hyper-Kähler manifold $M$ which has a Lagrangian fibration and satisfies the "RLF conjecture" characterizing classes associated to Lagrangian fibrations, satisfies Beauville's conjecture. However we do not know that the LSV varieties satisfy the RLF conjecture. We prove here the following result.

**Theorem A.2** *The relations* (A-1) *hold in* $\mathrm{CH}(J)$ *for the lattice* $\langle L, \Theta\rangle$ *of an LSV variety J. Conjecture A.1 is thus satisfied by an LSV variety with Picard number two.*

**Proof** There are, up to multiples, exactly two classes $L$ and $L'$ in $\langle L, \Theta\rangle$ satisfying $q(L) = 0, q(L') = 0$. Obviously $L^6 = 0$ in $\mathrm{CH}(J)$ since $L$ comes from the base which is of dimension 5, so we only have to prove that ${L'}^6 = 0$ in $\mathrm{CH}(J)$. We use Riess' argument in [72], however, in a different way. As a consequence of fundamental results of Huybrechts in [36], Riess proved the following:

**Theorem A.3** [72, Theorem 3.3] *Let $K$ be an isotropic class on a projective hyper-Kähler manifold $M$ of dimension $2n$. Then there exists a cycle $\Gamma \in \mathrm{CH}^{2n}(M \times M)$ such that $\Gamma^*$ acts as an automorphism of $\mathrm{CH}(M)$ preserving the intersection product, the action of $\Gamma^*$ on $H^2(M)$ preserves the Beauville–Bogomolov form $q_M$, and $\Gamma^* K$ belongs to the boundary of the birational Kähler cone of $M$.*

Here the birational Kähler cone of $M$ is defined as the union of the Kähler cones of hyper-Kähler manifolds $M'$ bimeromorphic to $M$ (the bimeromorphic map $M' \dashrightarrow M$ inducing an isomorphism on $H^2$). We apply this theorem to our class $L'$ on $J$ and thus get a correspondence $\Gamma$ as above. The class $\Gamma^* L'$ is an isotropic class, hence it must be proportional to either $L'$ or $L$. Furthermore, it belongs to the boundary of the birational Kähler cone. We now have:

**Lemma A.4** *The class $L'$ does not belong to the boundary of the birational Kähler cone.*

**Proof** This is proved in Corollary 3.10 of the present paper. □

By Lemma A.4, we conclude that $\Gamma^* L'$ is proportional to $L$. As $L^6 = 0$ in $\mathrm{CH}^6(J)$ and $\Gamma^*$ is an automorphism of $\mathrm{CH}(J)$ preserving the intersection product, we conclude that ${L'}^6 = 0$ in $\mathrm{CH}^6(J)$. □

If we consider the case of Picard rank three, where the Picard lattice $N$ of $J$ is generated by three classes $L$, $\Theta$ and $D$ with $q(L,D)=0$ and $q(\Theta,D)=0$, there are now, according to [15], 13 degree-six cohomological relations between $L$, $\Theta$ and $D$, generated by the classes $\alpha^6 \in S^6N \subset S^6H^2(J,\mathbb{Q})$, where $\alpha$ belongs to the conic $q(\alpha)=0$. Among these relations, two of them, namely those involving only $L$ and $\Theta$, are established in $\mathrm{CH}(J)$ by Theorem A.2. We also have the relations

$$L^5D=0 \quad \text{and} \quad L'^5D=0 \quad \text{in } H^{12}(J,\mathbb{Q}), \tag{A-2}$$

which are obtained by differentiating the relation (A-1) at $\alpha=L$ or $\alpha=L'$ in the direction given by $D$, which is tangent to the conic at these points since $q(D,L)=0$ and $q(D,L')=0$. We prove the following:

**Theorem A.5** *The relations* (A-2) *are satisfied in* $\mathrm{CH}^6(J)$.

**Proof** The first relation is proved by applying the following result from [81], which works in a more general context and needs a mild assumption on the infinitesimal variation of Hodge structure of a family of abelian varieties at the generic point of the base. More generally, let $M\to B$ be a fibration into abelian varieties and let $A\in \operatorname{Pic} M$ be a line bundle whose restriction to the general fiber $M_b$ is topologically trivial.

**Proposition A.6** *Assume that at the generic point $t\in B$, there exists a class $\alpha\in H^{1,0}(M_b)$ such that $\overline{\nabla}(\alpha)\colon T_{B,b}\to H^{0,1}(M_b)$ is surjective. Then there exists a point $b\in B$ such that $M_b$ is smooth and $A_{|M_b}$ is a torsion line bundle.*

*If all fibers $M_b$ have the same class $F$ in* $\mathrm{CH}(M)$, *it thus follows that $F.A=0$ in* $\mathrm{CH}(M)$.

Coming back to our situation, we have to check that the assumption on the infinitesimal variation of Hodge structures is satisfied in our situation. Let $J$ be the LSV variety of a cubic fourfold $X$. The infinitesimal variation of Hodge structure for the fibers of the Lagrangian fibration $J\to(\mathbb{P}^5)^\vee$ is thus canonically isomorphic to the variation of Hodge structure on the $H^3$ of the hyperplane section $X_H\subset X$. If $Y$ is a smooth cubic threefold in $\mathbb{P}^4$ defined by an equation $f=0$, Griffiths' theory of IVHS of hypersurfaces says that there are isomorphisms

$$H^{2,1}(Y)\cong R_f^1 \quad \text{and} \quad H^{1,2}(Y)\cong R_f^4$$

such that the infinitesimal variation of Hodge structure on $H^3(Y,\mathbb{C})$ is given (using the identification $R_f^3\cong H^1(Y,T_Y)$) by the multiplication map $R_f^3\to \operatorname{Hom}(R_f^1,R_f^4)$. Now consider the case where $Y$ is a hyperplane section $X_H$, defined by a linear equation $H$, of the cubic fourfold $X$. It is immediate to see that the inclusions $X_H \subset X \subset \mathbb{P}^5$ determine a quadratic polynomial $Q_{X,H} \in R_f^2$ such that the natural map $\rho\colon H^0(X_H, \mathcal{O}_{X_H}(1)) \to R_f^3$, defined as the first-order classifying map for the deformations of $X_H$ in $X$, is given by multiplication by $Q_{X,H}$. Combining these facts, we conclude that the desired infinitesimal criterion for the fibration $J \to (\mathbb{P}^5)^\vee$ holds if there exist a smooth hyperplane section $X_H \subset X$ and a linear form $x \in H^0(X_H, \mathcal{O}_{X_H}(1)) = R_f^1$ such that, with the above notation, the product map

$$xQ_{X,H}\colon R_f^1 \to R_f^4$$

by $xQ_{X,H}$ is an isomorphism. It is quite easy to show that the existence of such a hyperplane section is satisfied by $X$ in codimension one in the moduli space of cubic fourfolds, hence at the generic point of any Hodge locus in this moduli space, or equivalently any Noether–Lefschetz locus for the corresponding LSV variety $J$. The relation $L^5D = 0$ in $\mathrm{CH}^6(J)$ is thus satisfied at the generic point of the deformation locus of $J$ preserving the Hodge class $D$, hence everywhere by specialization.

To conclude the proof of Theorem A.5, we have to prove the relation $L'^5D = 0$ in $\mathrm{CH}^6(J)$. This follows however from the relation $L^5D = 0$ in $\mathrm{CH}^6(J)$ by the same argument as in the proof of Theorem A.2, using the specialization of the cycle $\Gamma$ and observing that $\Gamma^*$ acts by $\pm 1$ on $H^2(J,\mathbb{Q})^{\perp\langle L,\Theta\rangle}$, hence on $D$. □

## References


[1] **N Addington**, *On two rationality conjectures for cubic fourfolds*, Math. Res. Lett. 23 (2016) 1–13 MR Zbl

[2] **N Addington**, **R Thomas**, *Hodge theory and derived categories of cubic fourfolds*, Duke Math. J. 163 (2014) 1885–1927 MR Zbl

[3] **V Alexeev**, *Compactified Jacobians and Torelli map*, Publ. Res. Inst. Math. Sci. 40 (2004) 1241–1265 MR Zbl

[4] **D Arcara**, **A Bertram**, *Bridgeland-stable moduli spaces for K–trivial surfaces*, J. Eur. Math. Soc. 15 (2013) 1–38 MR Zbl

[5] **A Bayer**, *Wall-crossing implies Brill–Noether: applications of stability conditions on surfaces*, from "Algebraic geometry: Salt Lake City 2015" (T de Fernex, B Hassett, M Mustaţă, M Olsson, M Popa, R Thomas, editors), Proc. Sympos. Pure Math. 97, Amer. Math. Soc., Providence, RI (2018) 3–27 MR Zbl

[6] **A Bayer**, **M Lahoz**, **E Macrì**, **H Nuer**, **A Perry**, **P Stellari**, *Stability conditions in families*, Publ. Math. Inst. Hautes Études Sci. 133 (2021) 157–325 MR Zbl

[7] **A Bayer**, **E Macrì**, *MMP for moduli of sheaves on K3s via wall-crossing: nef and movable cones, Lagrangian fibrations*, Invent. Math. 198 (2014) 505–590 MR Zbl

[8] **A Bayer**, **E Macrì**, *Projectivity and birational geometry of Bridgeland moduli spaces*, J. Amer. Math. Soc. 27 (2014) 707–752 MR Zbl

[9] **A Beauville**, *Vector bundles on the cubic threefold*, from "Symposium in honor of C H Clemens" (A Bertram, J A Carlson, H Kley, editors), Contemp. Math. 312, Amer. Math. Soc., Providence, RI (2002) 71–86 MR Zbl

[10] **A Beauville**, *On the splitting of the Bloch–Beilinson filtration*, from "Algebraic cycles and motives, II" (J Nagel, C Peters, editors), London Math. Soc. Lecture Note Ser. 344, Cambridge Univ. Press (2007) 38–53 MR Zbl

[11] **A Beauville**, *Vector bundles on Fano threefolds and K3 surfaces*, Boll. Unione Mat. Ital. 15 (2022) 43–55 MR Zbl

[12] **A Beauville**, **R Donagi**, *La variété des droites d'une hypersurface cubique de dimension* 4, C. R. Acad. Sci. Paris Sér. I Math. 301 (1985) 703–706 MR Zbl

[13] **C Birkar**, **P Cascini**, **C D Hacon**, **J McKernan**, *Existence of minimal models for varieties of log general type*, J. Amer. Math. Soc. 23 (2010) 405–468 MR Zbl

[14] **S Bochner**, **W T Martin**, *Several complex variables*, Princeton Mathematical Series 10, Princeton Univ. Press (1948) MR Zbl

[15] **F A Bogomolov**, *On the cohomology ring of a simple hyper-Kähler manifold (on the results of Verbitsky)*, Geom. Funct. Anal. 6 (1996) 612–618 MR Zbl

[16] **T Bridgeland**, *Stability conditions on K3 surfaces*, Duke Math. J. 141 (2008) 241–291 MR Zbl

[17] **P Brosnan**, *Perverse obstructions to flat regular compactifications*, Math. Z. 290 (2018) 103–110 MR Zbl

[18] **M A A de Cataldo**, **A Rapagnetta**, **G Saccà**, *The Hodge numbers of O'Grady 10 via Ngô strings*, J. Math. Pures Appl. 156 (2021) 125–178 MR Zbl

[19] **F Charles**, **G Mongardi**, **G Pacienza**, *Families of rational curves on holomorphic symplectic varieties and applications to zero-cycles*, preprint (2019) arXiv 1907.10970

[20] **C H Clemens**, **P A Griffiths**, *The intermediate Jacobian of the cubic threefold*, Ann. of Math. 95 (1972) 281–356 MR Zbl

[21] **H Clemens**, *The infinitesimal Abel–Jacobi mapping for hypersurfaces*, from "The Lefschetz centennial conference, I" (D Sundararaman, editor), Contemp. Math. 58, Amer. Math. Soc., Providence, RI (1986) 81–89 MR Zbl

[22] **A Collino**, *The fundamental group of the Fano surface, I, II*, from "Algebraic threefolds" (A Conte, editor), Lecture Notes in Math. 947, Springer (1982) 209–220 MR Zbl

[23] **A Collino**, **J P Murre**, *The intermediate Jacobian of a cubic threefold with one ordinary double point; an algebraic-geometric approach, I*, Nederl. Akad. Wetensch. Proc. Ser. A 81 (= Indag. Math. 40) (1978) 43–55 MR Zbl

[24] **A Dimca**, *On the homology and cohomology of complete intersections with isolated singularities*, Compositio Math. 58 (1986) 321–339 MR Zbl

[25] **R Donagi**, **E Markman**, *Spectral covers, algebraically completely integrable, Hamiltonian systems, and moduli of bundles*, from "Integrable systems and quantum groups" (M Francaviglia, S Greco, editors), Lecture Notes in Math. 1620, Springer (1996) 1–119 MR Zbl

[26] **S Druel**, *Espace des modules des faisceaux de rang 2 semi-stables de classes de Chern* $c_1 = 0,\ c_2 = 2$ *et* $c_3 = 0$ *sur la cubique de* $\mathbb{P}^4$, Int. Math. Res. Not. (2000) 985–1004 MR Zbl

[27] **S Druel**, *Quelques remarques sur la décomposition de Zariski divisorielle sur les variétés dont la première classe de Chern est nulle*, Math. Z. 267 (2011) 413–423 MR Zbl

[28] **F El Zein**, **S Zucker**, *Extendability of normal functions associated to algebraic cycles*, from "Topics in transcendental algebraic geometry" (P Griffiths, editor), Ann. of Math. Stud. 106, Princeton Univ. Press (1984) 269–288 MR Zbl

[29] **D Greb**, **C Lehn**, **S Rollenske**, *Lagrangian fibrations on hyperkähler manifolds—on a question of Beauville*, Ann. Sci. Éc. Norm. Supér. 46 (2013) 375–403 MR Zbl

[30] **P A Griffiths**, *On the periods of certain rational integrals, I, II*, Ann. of Math. 90 (1969) 460–541 MR Zbl

[31] **P A Griffiths**, *Periods of integrals on algebraic manifolds, III: Some global differential-geometric properties of the period mapping*, Inst. Hautes Études Sci. Publ. Math. 38 (1970) 125–180 MR Zbl

[32] **C D Hacon**, **S J Kovács**, *Classification of higher dimensional algebraic varieties*, Oberwolfach Seminars 41, Birkhäuser, Basel (2010) MR Zbl

[33] **J Harris**, **M Roth**, **J Starr**, *Curves of small degree on cubic threefolds*, Rocky Mountain J. Math. 35 (2005) 761–817 MR Zbl

[34] **B Hassett**, *Special cubic fourfolds*, Compositio Math. 120 (2000) 1–23 MR Zbl

[35] **D Huybrechts**, *Compact hyper-Kähler manifolds: basic results*, Invent. Math. 135 (1999) 63–113 MR Zbl

[36] **D Huybrechts**, *The Kähler cone of a compact hyperkähler manifold*, Math. Ann. 326 (2003) 499–513 MR Zbl

[37] **D Huybrechts**, *The K3 category of a cubic fourfold*, Compos. Math. 153 (2017) 586–620 MR Zbl

[38] **D Kaledin**, **M Lehn**, **C Sorger**, *Singular symplectic moduli spaces*, Invent. Math. 164 (2006) 591–614 MR Zbl

[39] **J Kollár**, *Rational curves on algebraic varieties*, Ergebnesse der Math. (3) 32, Springer (1996) MR Zbl

[40] **J Kollár**, *Deformations of elliptic Calabi–Yau manifolds*, from "Recent advances in algebraic geometry" (C D Hacon, M Mustaţă, M Popa, editors), London Math. Soc. Lecture Note Ser. 417, Cambridge Univ. Press (2015) 254–290 MR Zbl

[41] **J Kollár**, **R Laza**, **G Saccà**, **C Voisin**, *Remarks on degenerations of hyper-Kähler manifolds*, Ann. Inst. Fourier (Grenoble) 68 (2018) 2837–2882 MR Zbl

[42] **J Kollár**, **S Mori**, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics 134, Cambridge Univ. Press (1998) MR Zbl

[43] **A Kuznetsov**, *Derived categories of cubic fourfolds*, from "Cohomological and geometric approaches to rationality problems" (F Bogomolov, Y Tschinkel, editors), Progr. Math. 282, Birkhäuser, Boston, MA (2010) 219–243 MR Zbl

[44] **M Lahoz**, **M Lehn**, **E Macrì**, **P Stellari**, *Generalized twisted cubics on a cubic fourfold as a moduli space of stable objects*, J. Math. Pures Appl. 114 (2018) 85–117 MR Zbl

[45] **C-J Lai**, *Varieties fibered by good minimal models*, Math. Ann. 350 (2011) 533–547 MR Zbl

[46] **R Laza**, *The moduli space of cubic fourfolds via the period map*, Ann. of Math. 172 (2010) 673–711 MR Zbl

[47] **R Laza**, **G Saccà**, **C Voisin**, *A hyper-Kähler compactification of the intermediate Jacobian fibration associated with a cubic* 4*–fold*, Acta Math. 218 (2017) 55–135 MR Zbl

[48] **J Le Potier**, *Faisceaux semi-stables de dimension* 1 *sur le plan projectif*, Rev. Roumaine Math. Pures Appl. 38 (1993) 635–678 MR Zbl

[49] **C Lehn**, **M Lehn**, **C Sorger**, **D van Straten**, *Twisted cubics on cubic fourfolds*, J. Reine Angew. Math. 731 (2017) 87–128 MR Zbl

[50] **M Lehn**, **C Sorger**, *La singularité de O'Grady*, J. Algebraic Geom. 15 (2006) 753–770 MR Zbl

[51] **C Li**, **L Pertusi**, **X Zhao**, *Elliptic quintics on cubic fourfolds, O'Grady 10, and Lagrangian fibrations*, Adv. Math. 408 (2022) art. id. 108584 MR Zbl

[52] **E Markman**, *A survey of Torelli and monodromy results for holomorphic-symplectic varieties*, from "Complex and differential geometry" (W Ebeling, K Hulek, K Smoczyk, editors), Springer Proc. Math. 8, Springer (2011) 257–322 MR Zbl

[53] **E Markman**, *Prime exceptional divisors on holomorphic symplectic varieties and monodromy reflections*, Kyoto J. Math. 53 (2013) 345–403 MR Zbl

[54] **D Markushevich**, **A S Tikhomirov**, *The Abel–Jacobi map of a moduli component of vector bundles on the cubic threefold*, J. Algebraic Geom. 10 (2001) 37–62 MR Zbl

[55] **D Matsushita**, *On fibre space structures of a projective irreducible symplectic manifold*, Topology 38 (1999) 79–83 MR Zbl

[56] **D Matsushita**, *On almost holomorphic Lagrangian fibrations*, Math. Ann. 358 (2014) 565–572 MR Zbl

[57] **D Matsushita**, *On deformations of Lagrangian fibrations*, from "K3 surfaces and their moduli" (C Faber, G Farkas, G van der Geer, editors), Progr. Math. 315, Birkhäuser, Cham (2016) 237–243 MR Zbl

[58] **C Meachan**, **Z Zhang**, *Birational geometry of singular moduli spaces of O'Grady type*, Adv. Math. 296 (2016) 210–267 MR Zbl

[59] **Y Miyaoka**, **S Mori**, *A numerical criterion for uniruledness*, Ann. of Math. 124 (1986) 65–69 MR Zbl

[60] **S Mukai**, *Symplectic structure of the moduli space of sheaves on an abelian or K3 surface*, Invent. Math. 77 (1984) 101–116 MR Zbl

[61] **N Nakayama**, *Zariski-decomposition and abundance*, MSJ Memoirs 14, Math. Soc. Japan, Tokyo (2004) MR Zbl

[62] **Y Namikawa**, *Deformation theory of singular symplectic n–folds*, Math. Ann. 319 (2001) 597–623 MR Zbl

[63] **K G O'Grady**, *Desingularized moduli spaces of sheaves on a K3*, J. Reine Angew. Math. 512 (1999) 49–117 MR Zbl

[64] **K Oguiso**, *Picard number of the generic fiber of an abelian fibered hyperkähler manifold*, Math. Ann. 344 (2009) 929–937 MR Zbl

[65] **K Oguiso**, *Shioda–Tate formula for an abelian fibered variety and applications*, J. Korean Math. Soc. 46 (2009) 237–248 MR Zbl

[66] **A Perego**, **A Rapagnetta**, *Deformation of the O'Grady moduli spaces*, J. Reine Angew. Math. 678 (2013) 1–34 MR Zbl

[67] **A Perego**, **A Rapagnetta**, *Factoriality properties of moduli spaces of sheaves on abelian and K3 surfaces*, Int. Math. Res. Not. 2014 (2014) 643–680 MR Zbl

[68] **A Perry**, *The integral Hodge conjecture for two-dimensional Calabi–Yau categories*, Compos. Math. 158 (2022) 287–333 MR Zbl

[69] **C A M Peters**, **J H M Steenbrink**, *Mixed Hodge structures*, Ergebnesse der Math. (3) 52, Springer (2008) MR Zbl

[70] **Z Ran**, *Hodge theory and the Hilbert scheme*, J. Differential Geom. 37 (1993) 191–198 MR Zbl

[71] **A Rapagnetta**, *On the Beauville form of the known irreducible symplectic varieties*, Math. Ann. 340 (2008) 77–95 MR Zbl

[72] **U Rieß**, *On Beauville's conjectural weak splitting property*, Int. Math. Res. Not. 2016 (2016) 6133–6150 MR Zbl

[73] **J Sawon**, *On the discriminant locus of a Lagrangian fibration*, Math. Ann. 341 (2008) 201–221 MR Zbl

[74] **J Sawon**, *Deformations of holomorphic Lagrangian fibrations*, Proc. Amer. Math. Soc. 137 (2009) 279–285 MR Zbl

[75] **S Takayama**, *On uniruled degenerations of algebraic varieties with trivial canonical divisor*, Math. Z. 259 (2008) 487–501 MR Zbl

[76] **C Voisin**, *Sur la stabilité des sous-variétés lagrangiennes des variétés symplectiques holomorphes*, from "Complex projective geometry" (G Ellingsrud, C Peskine, G Sacchiero, S A Strømme, editors), London Math. Soc. Lecture Note Ser. 179, Cambridge Univ. Press (1992) 294–303 MR Zbl

[77] **C Voisin**, *Hodge theory and complex algebraic geometry, II*, Cambridge Studies in Advanced Mathematics 77, Cambridge Univ. Press (2003) MR Zbl

[78] **C Voisin**, *Some aspects of the Hodge conjecture*, Jpn. J. Math. 2 (2007) 261–296 MR Zbl

[79] **C Voisin**, *Chow rings, decomposition of the diagonal, and the topology of families*, Annals of Mathematics Studies 187, Princeton Univ. Press (2014) MR Zbl

[80] **C Voisin**, *Hyper-Kähler compactification of the intermediate Jacobian fibration of a cubic fourfold: the twisted case*, from "Local and global methods in algebraic geometry" (N Budur, T de Fernex, R Docampo, K Tucker, editors), Contemp. Math. 712, Amer. Math. Soc., Providence, RI (2018) 341–355 MR Zbl

[81] **C Voisin**, *Torsion points of sections of Lagrangian torus fibrations and the Chow ring of hyper-Kähler manifolds*, from "Geometry of moduli" (J A Christophersen, K Ranestad, editors), Abel Symp. 14, Springer (2018) 295–326 MR Zbl

[82] **K Yoshioka**, *Moduli spaces of stable sheaves on abelian surfaces*, Math. Ann. 321 (2001) 817–884 MR Zbl

[83] **S Zucker**, *Generalized intermediate Jacobians and the theorem on normal functions*, Invent. Math. 33 (1976) 185–222 MR Zbl

[84] **S Zucker**, *The Hodge conjecture for cubic fourfolds*, Compositio Math. 34 (1977) 199–209 MR Zbl

*Department of Mathematics, Columbia University*
*New York, NY, United States*

*Institut de mathématiques de Jussieu*
*Paris, France*

gs3032@columbia.edu, claire.voisin@imj-prg.fr

Proposed: Richard P Thomas Received: 5 January 2021
Seconded: Mark Gross, Dan Abramovich Revised: 8 November 2021

Geometry & Topology 27:4 (2023) 1539–1586
DOI: 10.2140/gt.2023.27.1539
Published: 15 June 2023

# Cohomological $\chi$–independence for moduli of one-dimensional sheaves and moduli of Higgs bundles

DAVESH MAULIK
JUNLIANG SHEN

We prove that the intersection cohomology (together with the perverse and the Hodge filtrations) for the moduli space of one-dimensional semistable sheaves supported in an ample curve class on a toric del Pezzo surface is independent of the Euler characteristic of the sheaves. We also prove an analogous result for the moduli space of semistable Higgs bundles with respect to an effective divisor $D$ of degree $\deg(D) > 2g-2$. Our results confirm the cohomological $\chi$–independence conjecture by Bousseau for $\mathbb{P}^2$, and verify Toda's conjecture for Gopakumar–Vafa invariants for certain local curves and local surfaces.

For the proof, we combine a generalized version of Ngô's support theorem, a dimension estimate for the stacky Hilbert–Chow morphism, and a splitting theorem for the morphism from the moduli stack to the good GIT quotient.

14H60, 14J60, 14N35; 32S60

| | | |
|---|---|---|
| 0. | Introduction | 1539 |
| 1. | A support theorem for self-dual complexes | 1546 |
| 2. | Moduli of 1–dimensional sheaves and Higgs bundles | 1560 |
| 3. | Intersection cohomology complexes | 1569 |
| 4. | Proof of the main theorem | 1577 |
| | References | 1583 |

## 0 Introduction

We work over the complex numbers $\mathbb{C}$.

## 0.1 Cohomological $\chi$–independence

Let $C$ be a nonsingular irreducible projective curve of genus $g \geq 2$. The moduli space $N_{n,\chi}$ of (slope-)semistable vector bundles $\mathcal{E}$ with

$$\operatorname{rank}(\mathcal{E}) = n \quad \text{and} \quad \chi(\mathcal{E}) = \chi$$

is an irreducible projective variety, whose topology has been studied intensively for decades. When we fix the rank $n$, tensor product and duality induce natural isomorphisms between the moduli spaces indexed by different Euler characteristics (or degrees):

$$N_{n,\chi} \simeq N_{n,\chi+n}, \quad N_{n,\chi} \simeq N_{n,(2-2g)n-\chi}. \tag{1}$$

Under the assumption $\gcd(n,\chi) = 1$, so that the moduli spaces $N_{n,\chi}$ are nonsingular, Harder and Narasimhan proved in [20, Theorem 3.3.2] that the Poincaré polynomials of $N_{n,\chi}$ are distinct unless the moduli spaces are related via (1).

In this paper, we are interested in moduli spaces where the cohomological information does *not* depend on the Euler characteristic $\chi$. More precisely, we consider the following two types of moduli spaces $M^L_{\beta,\chi}$ and $\widetilde{M}_{n,\chi}$:

(A) $M^L_{\beta,\chi}$ is the moduli space of 1–dimensional semistable sheaves $\mathcal{F}$ with

$$[\operatorname{supp}(\mathcal{F})] = \beta \quad \text{and} \quad \chi(\mathcal{F}) = \chi$$

on a nonsingular *toric* del Pezzo surface $S$. Here the semistability is with respect to a polarization $L$ on $S$, $\operatorname{supp}(-)$ denotes the Fitting support, and $\beta$ is an ample curve class.

(B) $\widetilde{M}_{n,\chi}$ is the moduli space of semistable Higgs bundles $(\mathcal{E},\theta)$ with respect to an effective divisor $D$ of degree $\deg(D) > 2g-2$ on $C$ with

$$\operatorname{rank}(\mathcal{E}) = n \quad \text{and} \quad \chi(\mathcal{E}) = \chi.$$

We refer to Section 2 for more details on these moduli spaces. When $\chi$ is chosen so that there are no strictly semistable objects, the moduli spaces $M^L_{\beta,\chi}$ and $\widetilde{M}_{n,\chi}$ are nonsingular, and we consider their singular cohomology. However, for arbitrary values of $\chi$, these moduli spaces can be singular, due to the presence of strictly semistable objects. In this case, it is more natural for us to study their *intersection cohomology*. Our main result states that, unlike the case of curves, the intersection cohomology of these spaces is independent of the choice of $\chi$:

**Theorem 0.1** *For any $\chi, \chi' \in \mathbb{Z}$, there are isomorphisms of graded vector spaces*

$$\mathrm{IH}^*(M^L_{\beta,\chi}) \simeq \mathrm{IH}^*(M^L_{\beta,\chi'}), \quad \mathrm{IH}^*(\widetilde{M}_{n,\chi}) \simeq \mathrm{IH}^*(\widetilde{M}_{n,\chi'}),$$

*where $\mathrm{IH}^*(-)$ denotes the intersection cohomology. Moreover, these isomorphisms respect perverse and Hodge filtrations carried by these vector spaces.*

This phenomenon is surprising, since there is no direct geometric relationship other than those parallel to (1) between these moduli spaces with different Euler characteristics, and the result applies to both smooth and singular moduli spaces. For example in the case (B), the moduli space is nonsingular if and only if $\gcd(n, \chi) = 1$. Nevertheless, the result on intersection cohomology holds uniformly. Regarding the second part of the theorem and compatibility with filtrations, see Theorem 0.4 for further refinements.

Theorem 0.1 proves the cohomological $\chi$–independence conjecture (see Bousseau [3, Conjecture 0.4.3]) of the moduli space of 1–dimensional semistable sheaves on $\mathbb{P}^2$, which further proves [2, Conjecture 0.4.2] on the BPS numbers of the log K3 surface $(\mathbb{P}^2, E)$; see Bousseau [2, Theorem 0.4.5]. Its refinement (Theorem 0.4) proves Toda's conjecture [48, Conjecture 1.2] on the Gopakumar–Vafa invariants in the cases of certain local curves and local toric del Pezzo surfaces with ample curve classes; see Theorem 0.6. In case (A), when $S = \mathbb{P}^2$, it was proven by Bousseau [2, Theorem 0.5.2] that the dependence of the (intersection) Betti numbers on $\chi$ only relies on $\gcd(\deg(\beta), \chi)$, using connections with Gromov–Witten theory for the log K3 surface $(\mathbb{P}^2, E)$ and scattering diagrams. In case (B), when $\gcd(n, \chi) = 1$, the equality of Poincaré polynomials was proved by a direct calculation in work of Mozgovoy and Schiffmann [39] and Mellit [36], as well as in Groechenig, Wyss and Ziegler [19] by $p$–adic integration. We discuss connections between our theorems and enumerative geometry in Section 0.3 in more detail.

**Remark 0.2** By Demazure [15], a nonsingular del Pezzo surface belongs to one of the following types:

(a) $\mathbb{P}^2$.

(b) $\mathbb{P}^1 \times \mathbb{P}^1$.

(c) The blow-up of $\mathbb{P}^2$ at $n$ very general points with $1 \leq n \leq 8$.

Hence Theorem 0.1 recovers the case when a del Pezzo surface belongs to (a), (b), or (c) with $n \leq 3$. We note that the Fano condition is essential (see Section 0.4), but the toric condition is due to a technical result (Proposition 2.6), which we expect to hold for all del Pezzo surfaces. In other words, if the inequality (40) of Proposition 2.6 is proven for any del Pezzo surface $S$, then Theorem 0.1 (as well as Theorem 0.4 below) also holds for any del Pezzo surface.

**Remark 0.3** Although we will not require it further, our proof of Theorem 0.1 actually provides a natural isomorphism between these spaces, well-defined up to a scalar, which is compatible with the perverse and Hodge filtrations.

## 0.2 A support theorem

Comparing to $N_{n,\chi}$, a key feature of a moduli space $M$ of type (A) or (B) is that it admits a morphism $h\colon M \to B$ that behaves like a completely integrable system. Here $h$ is the Hilbert–Chow morphism

$$h\colon M_{\beta,\chi}^{L} \to B := \mathbb{P} H^0(S, \mathcal{O}_S(\beta)), \quad \mathcal{F} \mapsto \operatorname{supp}(\mathcal{F}), \tag{2}$$

in the case (A), and the Hitchin fibration

$$h\colon \widetilde{M}_{n,\chi} \to B := \bigoplus_{i=1}^{n} H^0(C, \mathcal{O}(iD)), \quad (\mathcal{E}, \theta) \mapsto \operatorname{char}(\theta), \tag{3}$$

in the case (B). In either case, there is a maximal Zariski open subset $U \subset B$ parametrizing nonsingular curves in the linear system $|\beta|$ or nonsingular spectral curves over $C$. We denote by $\pi : \mathcal{C} \to U$ the smooth map given by the universal curve over $U$.

**Theorem 0.4** *Let $M$ be a moduli space of* (A) *or* (B), *and let $h\colon M \to B$ be the morphism given by* (2) *or* (3), *respectively. Let $\pi : \mathcal{C} \to U \subset B$ be the universal curve of genus $d$. Then there is an isomorphism*

$$Rh_* \operatorname{IC}_M \simeq \bigoplus_{i=0}^{2d} \operatorname{IC}\big(\textstyle\bigwedge^i R^1\pi_*\mathbb{Q}_{\mathcal{C}}\big)[-i+d] \tag{4}$$

*in the bounded derived category $D^b \operatorname{MHM}(B)$ of mixed Hodge modules on $B$.*

Since the righthand side of (4) clearly does not depend on $L$ or $\chi$, Theorem 0.4 implies Theorem 0.1 immediately by taking global cohomology. The sheaf-theoretic nature of (4) further yields refinements of Theorem 0.1 involving perverse and Hodge filtrations.

Although Theorem 0.4 concerns mixed Hodge modules, it suffices to work with perverse sheaves for the proof. In fact, it is not difficult to check (4) over $U$:

$$Rh_*\mathbb{Q}_{h^{-1}(U)} \simeq \bigoplus_{i=0}^{2d} \bigwedge^i R^1\pi_*\mathbb{Q}[-i], \tag{5}$$

an isomorphism which only concerns the variation of Hodge structures of abelian varieties; see Proposition 2.2. In view of the decomposition theorem of Saito [42] for Hodge modules, to prove (4) from (5), we only need to verify that every semisimple component of $Rh_*\,\mathrm{IC}_M$ has *full support* $B$. This can be checked completely via the decomposition theorem due to Beĭlinson, Bernstein and Deligne [1] of $Rh_*\,\mathrm{IC}_M$ in terms of (shifts of) semisimple perverse sheaves. In particular, Theorem 0.4 can be viewed as a *support theorem* for the moduli spaces (A) and (B).

Ngô [40] introduced *a support theorem*, which determines the supports of the direct image complex $Rf_*\mathbb{Q}$ for certain morphisms $f: M \to B$ called weak abelian fibrations. It played a crucial role in his proof of the fundamental lemma of the Langlands program. After that, support theorems become powerful tools in various branches of mathematics; see for example Maulik and Shen [31], Maulik and Yun [33], Migliorini and Shende [37], Yun [49], Yun and Zhang [50], de Cataldo, Hausel and Migliorini [5] and de Cataldo, Rapagnetta and Saccà [7].

In our proof of Theorem 0.4, we systematically develop techniques for applying Ngô's support theorem to a more general setup. More precisely, we do not assume that the total space $M$ is nonsingular, and we work with more general objects $\mathcal{K} \in D_c^b(M)$ than the trivial local system $\mathbb{Q}$ on $M$. Theorem 1.1 reduces a support inequality of Ngô type to a *relative dimension bound* (see the condition (c)) for the complex $Rf_*\mathcal{K}$. Then we introduce techniques to check this bound when $M$ is a moduli space of type (A) or (B), and $\mathcal{K}$ is the intersection cohomology complex $\mathrm{IC}_M$.

## 0.3 Enumerative geometry

The cohomological $\chi$–independence phenomenon is expected to be part of a much more general phenomenon in the context of enumerative geometry of curves on Calabi–Yau 3–folds, specifically the proposal for Gopakumar-Vafa invariants developed in Maulik and Toda [32] and Toda [48].

Let $X$ be a Calabi–Yau 3–fold with $\beta \in H_2(X,\mathbb{Z})$ a curve class, and let $\sigma \in \mathrm{Pic}(X)_{\mathbb{C}}$ be an element in the complexified ample cone of $X$. Following Davison and Meinhardt [14], and conditional on the conjectural existence of a certain orientation, Toda introduced in [48] the *BPS sheaf*

$$\phi_{\mathrm{BPS}} \in \mathrm{Perv}(M^{\sigma}_{\beta,\chi}),$$

which is a perverse sheaf on the moduli space $M^{\sigma}_{\beta,\chi}$ of $\sigma$–semistable sheaves on $X$. Consider the Hilbert–Chow map

$$h\colon M^{\sigma}_{\beta,\chi} \to \mathrm{Chow}_{\beta}(X), \quad \mathcal{F} \mapsto \mathrm{supp}(\mathcal{F}).$$

For any $\gamma \in \mathrm{Chow}_{\beta}(X)$, the *Gopakumar–Vafa* (*GV*) invariant (see [48, Definition 1.1]) is defined by the identity

$$\Phi_{\sigma}(\gamma,\chi) := \sum_{i\in\mathbb{Z}} \chi\big({}^{\mathfrak{p}}\mathcal{H}^{i}(Rh_{*}\phi_{\mathrm{BPS}})|_{\gamma}\big)y^{i} \in \mathbb{Z}[y,y^{-1}]. \tag{6}$$

If $\chi$ is chosen so there are no strictly semistables, this definition specializes to the definition of Gopakumar–Vafa invariants in Maulik and Toda [32]. For any choice of $\chi$ and $\sigma$, these invariants are conjectured to encode the same information as the Gromov–Witten invariants of $X$ in the curve class $\beta$ and *arbitrary genus*. Since the latter invariants are independent of $\chi$ and $\sigma$, in order for this conjecture to be well-posed, the Gopakumar–Vafa invariants should be independent of this extra data as well. More precisely, in [48, Conjecture 1.2] Toda made the following conjecture concerning the structure of GV invariants, extending [32, Conjecture 3.3].

**Conjecture 0.5** (Toda) *The invariant* (6) *is independent of $\sigma$ and $\chi$.*

The invariants (6) specialize to a certain case of the Joyce–Song generalized Donaldson–Thomas (DT) invariants [24], and Conjecture 0.5 is expected to refine the Joyce–Song conjecture [24, Conjecture 6.20] on the generalized DT invariants, which in turn implies the strong rationality conjecture for Pandharipande–Thomas invariants, in Pandharipande and Thomas [41] and Toda [47].

Although currently the existence of the BPS sheaf is conjectural for most cases, it is known to exist for local curve and surface geometries; Meinhardt [34, Theorem 1.1] proved that when $X$ is a local curve $\mathrm{Tot}_C(\mathcal{O}_C(D)\oplus K_C(-D))$ with $\deg(D) > 2g-2$ or a local del Pezzo surface $\mathrm{Tot}(K_S)$, the BPS sheaf coincides with the intersection cohomology complex of the moduli space.

**Theorem 0.6** *Conjecture 0.5 holds when $X$ is a local curve $\mathrm{Tot}_C(\mathcal{O}_C(D)\oplus K_C(-D))$ with $D$ effective of $\deg(D) > 2g-2$ and $\beta = n[C]$, or a local toric del Pezzo surface $\mathrm{Tot}(K_S)$ and $\beta$ is an ample curve class on $S$.*

In fact, Toda showed in [48, Theorem 7.3] that (6) is independent of the stability parameter $\sigma$ under certain conditions which hold for local curves and local surfaces. Hence we may assume that $\sigma$ is given by a rational polarization, and the $\chi$–independence of (6) follows from the $\chi$–independence of the complex $Rf_* \mathrm{IC}_M$ for a moduli space of type (A) or (B) with the Hilbert–Chow map (2) or (3), respectively. The latter is given by Theorem 0.4.

Cohomological $\chi$–independence for Higgs bundles has also been studied systematically with connections to Kac polynomials and quivers. See Schiffmann [44] for more details. For contractible curves on Calabi–Yau threefolds, cohomological $\chi$–independence has been studied by Davison [12], who has proposed a representation-theoretic approach in that case via the cohomological Hall algebra. For other places where GV invariants arise geometrically, see Shen and Yin [45] and Chuang, Diaconescu and Pan [9] for connections with hyper-Kähler geometries (de Cataldo, Hausel and Migliorini [5]) and the $P = W$ conjecture (de Cataldo, Maulik and Shen [6]), respectively.

## 0.4 K3 surfaces and O'Grady 10

As illustrated in the following example of K3 surfaces, the "Fano" condition for the surface $S$ in (A) and the condition $\deg(D) > 2g - 2$ for Higgs bundles in (B) are essential for the $\chi$–independence to hold for intersection cohomology groups.

Let

$$(S, L) \quad \text{with } L = \mathcal{O}_S(\beta),\ \beta^2 = 2,$$

be a general polarized K3 surface of degree 2. The linear system $|\beta|$ is 2–dimensional whose general member is a genus 2 nonsingular curve. The linear system $|2\beta|$ is 5–dimensional. We consider the moduli space of semistable sheaves on $S$ supported in the curve class $2\beta$.

If $\chi = 1$, the moduli space $M^L_{2\beta,1}$ is nonsingular and deformation equivalent to the Hilbert scheme of 5 points on a K3 surface. When $\chi = 0$, the moduli space $M^L_{2\beta,0}$ is singular which admits a symplectic resolution. The resolved variety provides O'Grady's 10–dimensional "sporadic" example of compact hyper-Kähler manifolds. As a key step in their analysis of the topology of the O'Grady 10 variety, de Cataldo, Rapagnetta and Saccà [7] study the fibrations (2):

$$\begin{tikzcd} M^L_{2\beta,0} \arrow[dr, "h_0"'] & & M^L_{2\beta,1} \arrow[dl, "h_1"] \\ & {|2\beta|} & \end{tikzcd}$$

where $h_i(\mathcal{F}) = \mathrm{supp}(\mathcal{F})$. Combining Corollary 3.6.5 and Proposition 4.7.2 of [7], we observe that

$$Rh_{1*}\,\mathrm{IC} = Rh_{0*}\,\mathrm{IC} \oplus \mathscr{S}[3], \tag{7}$$

where $\mathscr{S}$ is a semisimple object supported on the divisor $\mathrm{Sym}^2(|\beta|) \subset |2\beta|$. Furthermore, by [7, Proposition 4.6.1], the object $\mathscr{S}$ (which is denoted by $\mathscr{S}^-_\Sigma$ in [7]) has nontrivial global cohomology. Hence we see from (7) that the $\chi$–independence fails for the K3 surface $S$ both sheaf theoretically (Theorem 0.4) and cohomologically (Theorem 0.1).

A similar phenomenon as above is expected to hold for the Higgs bundles with $D = K_C$. Failure of the $\chi$–independence for the "Calabi–Yau" case is due to the fact that the BPS sheaf is different from the intersection cohomology complex on the moduli space.

### Plan of the paper

In Section 1, we formulate and prove a generalized version of Ngô's support theorem, which applies to singular varieties and more general complexes. In order to apply this support theorem to intersection cohomology complexes, we need to prove a bound for IC-complexes (which holds automatically in the smooth case). This is accomplished in Sections 2 and 3, where we combine techniques from algebraic stacks, nilpotent Higgs bundles, moduli of framed objects, and unbounded complexes. Then in Section 4, we follow a strategy of Chaudouard and Laumon to show that the support inequalities are sufficient to deduce our theorems for moduli of 1–dimensional sheaves and Higgs bundles.

**Acknowledgements** We are grateful to Bhargav Bhatt, Pierrick Bousseau and Johan de Jong for helpful discussions, and Pinka and Peter Pinkerton for further assistance. We would like to thank Pierrick Bousseau and Tudor Padurariu for their careful reading of an early draft of the paper and pointing out several typos. We also thank the anonymous referee for careful reading and numerous useful suggestions. Shen was supported by the NSF grant DMS-2134315.

## 1 A support theorem for self-dual complexes

### 1.1 Overview

The purpose of this section is to formulate and prove a generalized version of Ngô's support theorem for self-dual complexes. Throughout Section 1, until Section 1.7, we assume that the base field $\boldsymbol{k}$ is a finite field with $\bar{\boldsymbol{k}}$ its algebraic closure. We assume that $l$ is a prime number coprime to the characteristic of $\boldsymbol{k}$ when we work with $l$–adic sheaves. For notational convenience, we omit Tate twists when it does not cause confusion.

Let $B$ be a scheme over $\boldsymbol{k}$. Let $g\colon P \to B$ be a smooth $B$–group scheme with geometrically connected fibers, and let $f\colon M \to B$ be a proper morphism with $M$ quasi-projective. Assume that the group scheme $P$ acts on $M$ via

$$a\colon P \times_B M \to M. \tag{8}$$

We say that the triple $(M, P, B)$ is a *weak abelian fibration* of relative dimension $d$ if

(i) every fiber of the map $g$ is pure of dimension $d$, and $M$ has pure dimension

$$\dim M = d + \dim B, \tag{9}$$

(ii) the action (8) of $P$ on $M$ has *affine* stabilizers, and

(iii) the Tate module $T_{\overline{\mathbb{Q}}_l}(P)$ associated with the group scheme $P$ is polarizable.

The notion of weak abelian fibration was introduced by Ngô [40] modeled on Hitchin's integrable systems [21; 22]. We refer to Section 1.3 for a brief review of Tate modules and their polarizations.

For a closed point $s \in B$, we denote by $\delta(s)$ the dimension of the affine part of the algebraic group $P_s$. This defines an upper-semicontinuous function

$$\delta\colon B \to \mathbb{N}, \quad s \mapsto \delta(s).$$

For a closed subvariety $Z \subset B$, we define $\delta(Z)$ to be the minimal value of the function $\delta$ on $Z$.

The following is our main theorem.

**Theorem 1.1** *Let $(M, P, B)$ be a weak abelian fibration of relative dimension $d$. Let $\mathcal{K} \in D_c^b(M, \overline{\mathbb{Q}}_l)$ be a $P$–equivariant bounded complex satisfying the following properties*:

(a) ***Decomposition theorem*** *The direct image complex admits a (noncanonical) decomposition*

$$Rf_*\mathcal{K} \simeq \bigoplus_i {}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K})[-i]. \tag{10}$$

*Moreover, after a base change to* $B_{\bar{k}} = B \times_k \bar{k}$, *the perverse sheaves* ${}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K})$ *are semisimple of the form*

$$ {}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K}) = \bigoplus_{\alpha} \mathrm{IC}_{Z_{\alpha,i}}(L_{\alpha,i}), $$

*where* $Z_{\alpha,i}$ *is a closed irreducible subvariety of* $B_{\bar{k}}$ *and each* $L_{\alpha,i}$ *is a pure simple local system of weight* $i$ *on an open dense subset of* $Z_{\alpha,i}$. *We call these* $Z_{\alpha,i}$ *the supports of the decomposition* (10).

(b) ***Duality*** *We have an isomorphism*

$$ \mathbb{D}(\mathcal{K}) \simeq \mathcal{K}[2\dim M] $$

*with* $\mathbb{D}(-)$ *the dualizing functor on* $M$.

(c) ***Relative dimension bound*** *For the standard truncation functor* $\tau_{>*}(-)$, *we have*

$$ \tau_{>2d}(Rf_*\mathcal{K}) = 0. $$

*Then for any support* $Z$ *of the decomposition* (10), *we have the inequality*

$$ \operatorname{codim} Z \le \delta_Z. \tag{11} $$

In [40], Ngô worked with the trivial local system $\overline{\mathbb{Q}}_l$ on $M$, where he assumed that the conditions (a) and (b) hold. Furthermore, he assumed that every fiber of $f$ is pure of dimension $d$, where the condition (c) follows automatically by the base change. Therefore, Theorem 1.1 is a generalization of Ngô's support theorem [40, Theorem 7.2.1 and Proposition 7.2.2]. We note that (c) is a crucial condition for the support theorem to hold for general $\mathcal{K}$ as in Theorem 1.1. We first illustrate this in the following special case of Theorem 1.1 — the Goresky–MacPherson inequality.

## 1.2 The Goresky–MacPherson inequality

If the group scheme $P$ is affine and its action on $M$ is trivial, Theorem 1.1 then specializes to the following theorem, which is known as the *Goresky–MacPherson inequality* when $M$ is nonsingular and $\mathcal{K} = \overline{\mathbb{Q}}_l$.

**Theorem 1.2** *Let* $f: M \to B$ *be a proper map with* $\dim M = \dim B + d$. *Assume* $\mathcal{K} \in D^b(M, \overline{\mathbb{Q}}_l)$ *satisfies* (a), (b) *and* (c) *of Theorem 1.1. Then any support* $Z$ *of* (10) *satisfies the inequality*

$$ \operatorname{codim} Z \le d. $$

We first provide a proof of Theorem 1.2 since it contains the main ingredients in the proof of Theorem 1.1, and in particular demonstrates the role played by the conditions (a), (b) and (c).

**Proof** Let $Z$ be a support. We write

$$\begin{aligned} \mathrm{occ}(Z) &:= \{i \in \mathbb{Z} : {}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K}) \text{ contains a simple factor with support } Z\}, \\ \mathrm{amp}(Z) &:= \max(\mathrm{occ}(Z)) - \min(\mathrm{occ}(Z)). \end{aligned}$$

By (b), the set $\mathrm{occ}(Z)$ is symmetric with respect to the integer $\dim M$. This allows us to pick $m \in \mathrm{occ}(Z)$ with $m \geq \dim M$. In particular, we have ${}^{\mathfrak{p}}\mathcal{H}^m(Rf_*\mathcal{K}) \neq 0$. Hence by (a) there exists an open subset $U \subset Z$ and a local system $\mathcal{L}$ on $U$ such that the shifted perverse sheaf

$$(\mathcal{L}[\dim Z])[-m] = \mathcal{L}[\dim Z - m]$$

is a direct-sum component of the complex $(Rf_*\mathcal{K})|_U$. We obtain that

$$\mathcal{H}^{m-\dim Z}(Rf_*\mathcal{K}) \neq 0 \in D_c^b(B, \overline{\mathbb{Q}}_l). \tag{12}$$

By (12) and the condition (c), we conclude that

$$\dim M - \dim Z \leq m - \dim Z \leq 2d,$$

where the first inequality follows from the choice of $m$. This completes the proof of Theorem 1.2 thanks to (9). □

As observed by Ngô [40, Proposition 7.3.2], for a weak abelian fibration $(M, P, B)$ and an object $\mathcal{K}$ as in Theorem 1.1, if we have

$$\mathrm{amp}(Z) \geq 2(d - \delta_Z), \tag{13}$$

then the integer $m$ in the proof of Theorem 1.2 can be chosen so that

$$m \geq \dim M + (d - \delta_Z).$$

An identical argument as above implies (11).

In conclusion, the following proposition implies Theorem 1.1.

**Proposition 1.3** *Under the assumption of Theorem 1.1, the inequality* (13) *holds for any support $Z$ of $Rf_*\mathcal{K}$.*

The rest of Section 1 is devoted to proving Proposition 1.3. We will further reduce Proposition 1.3 to a fiberwise "freeness" statement as stated in Proposition 1.5. Essentially, the arguments of [40, Section 7] can be modified to prove this more generalized version of "freeness" under our assumptions. We point out the necessary modifications (see Propositions 1.4 and 1.6) and sketch all major steps in the proof, which follows [40, Section 7], for the reader's convenience.

## 1.3 Actions of the group scheme

As part of the data of a weak abelian fibration $(M, P, B)$, the group scheme $g\colon P \to B$ is smooth over $B$ with $d$–dimensional geometrically connected fibers, which defines a complex

$$\Lambda_P := Rg_!\overline{\mathbb{Q}}_l[2d]$$

on the base $B$. The stalk of each cohomology sheaf $\mathcal{H}^{-i}(\Lambda_P)$ over a closed point $s \in B$ computes the $i^{\text{th}}$ homology of the group $P_s$,

$$\mathcal{H}^{-i}(\Lambda_P)_s = H_c^{2d-i}(P_s, \overline{\mathbb{Q}}_l) = H_i(P_s, \overline{\mathbb{Q}}_l).$$

The Tate module associated with $g\colon P \to B$ is defined to be

$$T_{\overline{\mathbb{Q}}_l}(P) := \mathcal{H}^{-1}(\Lambda_P). \tag{14}$$

Note that the complex $\Lambda_{(-)}$ and the sheaf $T_{\overline{\mathbb{Q}}_l}(-)$ are defined for any smooth group scheme over $B$. In our setting, as shown in [40, Section 7.4.3], the group structure $\mu\colon P \times_B P \to P$ induces a convolution product

$$\tau\colon \Lambda_P \otimes \Lambda_P \to \Lambda_P.$$

Furthermore, it is also shown there that equation (14) extends to a natural isomorphism

$$\Lambda_P = \bigoplus \bigwedge^i T_{\overline{\mathbb{Q}}_l}(P)[i],$$

compatible with the multiplication action on each side.

We consider the Chevalley decomposition of the nonsingular commutative group $P_s$

$$1 \to R_s \to P_s \to A_s \to 1 \tag{15}$$

over any geometric point $s \in B$ where $R_s$ is affine and $A_s$ is an abelian variety. This induces the short exact sequence of Tate modules

$$0 \to T_{\overline{\mathbb{Q}}_l}(R_s) \to T_{\overline{\mathbb{Q}}_l}(P_s) \to T_{\overline{\mathbb{Q}}_l}(A_s) \to 0. \tag{16}$$

Following [40, Section 7.1.4], we say that the Tate module (14) is polarizable if étale locally there exists a bilinear form

$$T_{\overline{\mathbb{Q}}_l}(P) \times T_{\overline{\mathbb{Q}}_l}(P) \to \overline{\mathbb{Q}}_l$$

which induces a nondegenerate pairing on $T_{\overline{\mathbb{Q}}_l}(A_s)$ for any $s \in B$ via the quotient map of (16).

The following proposition generalizes the cap product action constructed in Section 7.4.2 of [40].

**Proposition 1.4** *Let $(M, P, B)$ be a weak abelian fibration of relative dimension $d$, and let $\mathcal{K} \in D_c^b(M, \overline{\mathbb{Q}}_l)$ be a $P$–equivariant object. Then the $P$–action* (8) *on $M$ induces an action of $\Lambda_P$ on $Rf_*\mathcal{K}$,*

$$c\colon \Lambda_P \otimes Rf_*\mathcal{K} \to Rf_*\mathcal{K}. \tag{17}$$

*Furthermore, the compositions $c \circ (\tau \otimes \mathrm{id})$ and $c \circ (\mathrm{id} \otimes c)$ define the same morphism*

$$\Lambda_P \otimes \Lambda_P \otimes Rf_*\mathcal{K} \to Rf_*\mathcal{K}.$$

**Proof** The trace map

$$Ra_!\overline{\mathbb{Q}}_l[2d] \to \overline{\mathbb{Q}}_l$$

on $M$ associated with (8) induces a morphism

$$Ra_!\overline{\mathbb{Q}}_l[2d] \otimes \mathcal{K} \to \mathcal{K}. \tag{18}$$

We consider the Cartesian diagram

$$\begin{array}{ccc} P \times_B M & \xrightarrow{p_M} & M \\ {\scriptstyle p_P}\downarrow & & \downarrow{\scriptstyle f} \\ P & \xrightarrow{g} & B \end{array} \tag{19}$$

with $p_M$ and $p_P$ the projections. The lefthand side of (18) is equal to

$$Ra_!(\overline{\mathbb{Q}}_l[2d] \otimes a^*\mathcal{K}) = Ra_!(\overline{\mathbb{Q}}_l[2d] \otimes p_M^*\mathcal{K}).$$

Here we used the projection formula and the isomorphism $\phi\colon a^*\mathcal{K} \simeq p_M^*\mathcal{K}$ given by the $P$–equivariance of $\mathcal{K}$. Hence we obtain the morphism

$$Ra_!(\overline{\mathbb{Q}}_l[2d] \otimes p_M^*\mathcal{K}) \to \mathcal{K}.$$

Applying the functor $Rf_*$ to the morphism above and noticing that $Rf_* = Rf_!$, we have

$$Rf_! Ra_!(\overline{\mathbb{Q}}_l[2d] \otimes p_M^* \mathcal{K}) \to Rf_* \mathcal{K},$$

where the lefthand side can be rewritten as

$$\begin{aligned} Rf_! Ra_!(\overline{\mathbb{Q}}_l[2d] \otimes p_M^* \mathcal{K}) &= Rf_! Rp_{M!}(\overline{\mathbb{Q}}_l[2d] \otimes p_M^* \mathcal{K}) \\ &= Rf_!(Rp_{M!}\overline{\mathbb{Q}}_l[2d] \otimes \mathcal{K}) \\ &= Rf_!\big(Rp_{M!}(p_P^*\overline{\mathbb{Q}}_l[2d]) \otimes \mathcal{K}\big) \\ &= Rf_!(f^* Rg_!\overline{\mathbb{Q}}_l[2d] \otimes \mathcal{K}) \\ &= Rg_!\overline{\mathbb{Q}}_l[2d] \otimes Rf_!\mathcal{K}, \end{aligned}$$

where the first equality follows from

$$f \circ a = g \times_B f \colon P \times_B M \to B,$$

the second equality is given by the projection formula, the third equality follows from $p_P^*\overline{\mathbb{Q}}_l = \overline{\mathbb{Q}}_l$, the fourth equality is the base change

$$Rp_{M!}p_P^* = f^* Rg_!$$

with respect to the diagram (19), and the last equality is again given by the projection formula.

To show the second claim of the proposition, we apply the same construction from above to the commutative diagram

(20)
$$\begin{tikzcd} P \times_B P \times_B M \arrow[r, "\mathrm{id}_P \times a"] \arrow[d, "\mu \times \mathrm{id}_M"'] & P \times_B M \arrow[d, "a"] \\ P \times_B M \arrow[r, "a"] & M \end{tikzcd}$$

Again using the trace map, each path defines a morphism

$$\Lambda_P \otimes \Lambda_P \otimes Rf_*\mathcal{K} \to Rf_*\mathcal{K}.$$

The path via the lower-left corner gives the morphism $c \circ (\tau \otimes \mathrm{id})$ and the path via the upper-right corner gives the morphism $c \circ (\mathrm{id} \otimes c)$. The equivariant structure on $\mathcal{K}$ implies a cocycle condition on the isomorphism $\phi$ after pullback to $P \times_B P \times_B M$; this cocycle condition implies that these two morphisms agree.

This completes the proof of Proposition 1.4. $\square$

## 1.4 Actions on each support and freeness

We denote by $I$ the set of the supports $Z \subset B_{\bar{k}}$ of $Rf_*\mathcal{K}$. In view of the condition (a) of Theorem 1.1, we have a canonical decomposition of the perverse sheaf ${}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K})$ in terms of the supports

$$ {}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K}) = \bigoplus_{\alpha \in I} \mathcal{K}^i_\alpha, $$

where $\mathcal{K}^i_\alpha$ has support $Z_\alpha$ indexed by $\alpha \in I$. We collect all the direct summands of $Rf_*\mathcal{K}$ with support $Z_\alpha$,

$$ \mathcal{K}_\alpha := \bigoplus_i \mathcal{K}^i_\alpha. \tag{21} $$

In the following four steps, we prove that Proposition 1.3 can be reduced to a freeness property concerning stalks of (21).

**Step 1** For any support $Z_\alpha$ of $Rf_*\mathcal{K}$, we may find an open dense subset $V_\alpha \subset Z_\alpha$ such that

(i) the restriction of $\mathcal{K}^i_\alpha$ to $V_\alpha$ is of the form $\mathcal{L}^i_\alpha[\dim V_\alpha]$ with $\mathcal{L}^i_\alpha$ a pure local system of weight $i$,

(ii) the restriction $P_\alpha$ of the group scheme $P$ to the support $Z_\alpha$ admits a Chevalley decomposition

$$ 1 \to R_\alpha \to P_\alpha \to A_\alpha \to 1, \tag{22} $$

whose induced short exact sequence of Tate modules

$$ 0 \to T_{\overline{\mathbb{Q}}_l}(R_\alpha) \to T_{\overline{\mathbb{Q}}_l}(P_\alpha) \to T_{\overline{\mathbb{Q}}_l}(A_\alpha) \to 0 $$

satisfies that $T_{\overline{\mathbb{Q}}_l}(R_\alpha)$ is a pure local system of weight $-2$, and

(iii) for any other support $Z_{\alpha'}$, we have $Z_{\alpha'} \cap V_\alpha = \varnothing$ unless $Z_\alpha \subset Z_{\alpha'}$.

Since (i) and (iii) are standard and (ii) only concerns the group scheme $P$, this follows identically from [40, Section 7.4.8].

**Step 2** In [40, Section 7.4.6], we replace $Rf_!\overline{\mathbb{Q}}_l$ by $Rf_!\mathcal{K}$, and we replace the cap product action

$$ \Lambda_P \times Rf_!\overline{\mathbb{Q}}_l \to Rf_!\overline{\mathbb{Q}}_l $$

of [40, Section 7.4.2] by the action (17) constructed in Proposition 1.4:

$$ \Lambda_P \otimes Rf_!\mathcal{K} \to Rf_!\mathcal{K}. $$

As a consequence, for each $\alpha \in I$ and $i$ we obtain an morphism

$$T_{\overline{\mathbb{Q}}_l}(P_\alpha) \otimes \mathcal{K}_\alpha^i \to \mathcal{K}_\alpha^{i-1}. \tag{23}$$

The last statement follows from an identical argument as in Sections 7.4.6 and 7.4.7 of [40]. More precisely, perverse truncation functors yield

$$T_{\overline{\mathbb{Q}}_l}(P_\alpha) \otimes {}^{\mathfrak{p}}\mathcal{H}^i(Rf_!\mathcal{K}) \to {}^{\mathfrak{p}}\mathcal{H}^{i-1}(Rf_!\mathcal{K}),$$

which can be further written as

$$\bigoplus_{\alpha \in I} T_{\overline{\mathbb{Q}}_l}(P_\alpha) \otimes \mathcal{K}_\alpha^i \to \bigoplus_{\alpha \in I} \mathcal{K}_\alpha^{i-1}$$

in terms of the supports $Z_\alpha$. This gives the canonical morphism (23).

**Step 3** Now we combine Steps 1 and 2. Consider the restriction of $\mathcal{K}_\alpha$ to $V_\alpha$ of Step 1,

$$\mathcal{L}_\alpha := \bigoplus_i \mathcal{L}_\alpha^i[-i].$$

Using the last part of Proposition 1.4, the morphisms (23) extend to an action of the local system of graded algebras $\Lambda_{P_\alpha} = \bigoplus \bigwedge^i T_{\overline{\mathbb{Q}}_l}(P_\alpha)[i]$ on $\mathcal{L}_\alpha$.

As explained in [40, Section 7.4.9], the first paragraph of page 121, an argument using weights shows that (23) passes through an action of the abelian variety part $T_{\overline{\mathbb{Q}}_l}(A_\alpha)$ of the Tate module $T_{\overline{\mathbb{Q}}_l}(P_\alpha)$. As a result, we have a graded module structure on $\mathcal{L}_\alpha$ of the graded algebra $\Lambda_{A_\alpha}$ associated with the abelian scheme $A_\alpha$ in (22),

$$\Lambda_{A_\alpha} \otimes \mathcal{L}_\alpha \to \mathcal{L}_\alpha. \tag{24}$$

Note that we use the assumption that $\boldsymbol{k}$ is a finite field here.

**Step 4** As commented in the paragraph after [40, Proposition 7.4.10], Proposition 1.3 can be deduced from the following proposition.

**Proposition 1.5** *We follow the same notation as in Steps* 1–3 *above. Let $u_\alpha \in V_\alpha$ be any geometric point. Then the stalk $\mathcal{L}_{\alpha,u_\alpha}$ of $\mathcal{L}_\alpha$ is a free graded module of the graded algebra $\Lambda_{A_\alpha,u_\alpha}$ under the action* (24).

We complete the proof of Proposition 1.5 in the next two sections.

## 1.5 Freeness

In this section we prove the following proposition, generalizing [40, Proposition 7.5.1]. Then in Section 1.6 we eventually reduce Proposition 1.5 to Proposition 1.6.

**Proposition 1.6** *Assume $X$ is projective over $\overline{k}$ and admits an action of an abelian variety $A$ over $\overline{k}$ with finite stabilizers. Let $\mathcal{E} \in D_c^b(X, \overline{\mathbb{Q}}_l)$ be an $A$–equivariant object. Then the graded cohomology group*

$$\bigoplus_i H^i(X, \mathcal{E})[-i] \tag{25}$$

*is naturally a free graded module of the graded algebra $\Lambda_A = \oplus_i H^i(A, \overline{\mathbb{Q}}_l)[-i]$.*

**Proof** Since the $A$–action preserves the connected components of $X$, we may assume that $X$ is connected. We consider the quotient map

$$q\colon X \to Y := X/A$$

with $X/A$ an Artin stack with finite inertia. Thanks to the projectivity of $A$, the morphism $q$ is smooth and proper. For the $A$–equivariant object $\mathcal{E}$, there exists an object $\mathcal{E}'$ on $Y$ such that

$$q^*\mathcal{E}' = \mathcal{E},$$

and the projection formula yields

$$Rq_*\mathcal{E} = Rq_*\overline{\mathbb{Q}}_l \otimes \mathcal{E}'. \tag{26}$$

In particular, the complex $Rq_*\mathcal{E}$ admits a natural $\Lambda_A$–action through the first factor of the righthand side of (26). This shows that (25) is a natural graded $\Lambda_A$–module.

Now since $q$ is smooth and proper, we have a decomposition[1]

$$Rq_*\overline{\mathbb{Q}}_l \simeq \bigoplus_i R^i q_*\overline{\mathbb{Q}}_l[-i]. \tag{27}$$

Moreover, we consider the Cartesian diagram, with all the arrows smooth maps,

$$\begin{tikzcd} A \times X \arrow[r, "q'"] \arrow[d, "q'"'] & X \arrow[d, "q"] \\ X \arrow[r, "q"] & Y \end{tikzcd}$$

[1]As explained in the proof of [40, Proposition 7.5.1], the decomposition here is induced by the cup-product with an relative ample class. We also refer to [46] as a general reference for the decomposition theorem for Artin stacks with affine stabilizers.

By the base change, we obtain the canonical $A$–equivariant isomorphism of local systems

$$q^* R^i q_* \overline{\mathbb{Q}}_l = R^i q'_* (q'^* \overline{\mathbb{Q}}_l). \tag{28}$$

We have that $q'^* \overline{\mathbb{Q}}_l$ is a trivial local system of rank 1 and an $A$–equivariant structure on it is trivial by the connectedness of $A$; cf [51, Lemma A.1.2]. In particular, $q^* R^i q_* \overline{\mathbb{Q}}_l$ is a trivial local system equipped with the trivial $A$–equivariant structure by (28). Consequently, each $R^i q_* \overline{\mathbb{Q}}_l$ is canonically isomorphic to the trivial local system taking values in $H^i(A, \overline{\mathbb{Q}}_l)$. The filtration $(\tau_{\leq *} Rq_* \overline{\mathbb{Q}}_l) \otimes \mathcal{E}'$ of $Rq_* \mathcal{E}$ induces the spectral sequence

$$H^j(Y, R^i q_* \overline{\mathbb{Q}}_l \otimes \mathcal{E}') = H^j(Y, \mathcal{E}') \otimes H^i(A, \overline{\mathbb{Q}}_l) \Rightarrow H^{i+j}(X, \mathcal{E}),$$

which degenerates thanks to (26) and the decomposition (27). Hence we obtain a filtration stable under the $\Lambda_A$–action, whose graded pieces are the free graded $\Lambda_A$ modules

$$H^j(Y, \mathcal{E}') \otimes \left( \bigoplus_i H^i(A, \overline{\mathbb{Q}}_l) \right).$$

This proves the freeness of the entire module $H^*(X, \mathcal{E}) = \bigoplus_i H^i(X, \mathcal{E})[-i]$. □

## 1.6 Proof of Proposition 1.5

We deduce Proposition 1.5 from Proposition 1.6 by a descending induction on the dimension of the support $Z_\alpha$. This is parallel to [40, Section 7.7].

We complete the induction in the following three steps.

**Step A** The induction base follows from Proposition 1.6, which we explain as follows. We assume $Z_{\alpha_0} = B_{\bar{k}}$ and $V_{\alpha_0}$ is an open dense subset of $Z_{\alpha_0}$ as in Step 1 of Section 1.4. All the other $Z_\alpha$ with $\alpha \neq \alpha_0$ do not intersect with $V_{\alpha_0}$, and

$${}^{\mathfrak{p}}\mathcal{H}^i(Rf_* \mathcal{K})|_{V_{\alpha_0}} = \mathcal{L}^i_{\alpha_0}[\dim B].$$

Therefore, for any geometric point $u_{\alpha_0}$ of $V_{\alpha_0}$ with

$$\iota_{u_{\alpha_0}} : M_{u_{\alpha_0}} \hookrightarrow M$$

the corresponding fiber, we have the identification

$$\bigoplus_i \mathcal{L}^i_{\alpha_0, u_{\alpha_0}}[-i + \dim B] = \bigoplus_i H^i(M_{u_{\alpha_0}}, \iota^*_{u_{\alpha_0}} \mathcal{K})[-i] \tag{29}$$

by the base change. Parallel to Step 3 of Section 1.4, (29) admits a natural $\Lambda_{A_{\alpha_0,u_{\alpha_0}}}$–action induced by the action of $\Lambda_{P_{\alpha_0,u_{\alpha_0}}}$.

As explained in the last paragraph of [40, Section 7.7.1], we may assume that the geometric point $u_{\alpha_0}$ is defined over a finite field. So there exists a quasi-lifting $A_{\alpha_0,u_{\alpha_0}} \to P_{\alpha_0,u_{\alpha_0}}$ (see [40, Proposition 7.5.3]) such that the $\Lambda_{A_{\alpha_0,u_{\alpha_0}}}$–action on (29) is induced by the $A_{\alpha_0,u_{\alpha_0}}$–action on $M_{u_{\alpha_0}}$. By the axiom (ii) of weak abelian fibrations, the $A_{\alpha_0,u_{\alpha_0}}$–action on $M_{u_{\alpha_0}}$ passing through $P_{\alpha_0,u_{\alpha_0}}$ has finite stabilizers. Hence Proposition 1.6 implies that (29) is free over $\Lambda_{A_{\alpha_0,u_{\alpha_0}}}$. This completes the proof of the induction base.

**Step B** Since Proposition 1.5 is a local statement, we may work with a strictly Henselian base. Assume that $B_\alpha$ is the strict Henselization of a geometric point $u_\alpha$ defined over a finite field lying in $V_\alpha \subset Z_\alpha$. By the choice of $V_\alpha$ in Step 1 of Section 1.4, the stalk $\mathcal{K}^i_{\alpha',u_\alpha}$ is nonzero only if $Z_\alpha$ is strictly contained in $Z_{\alpha'}$. In this case, the induction assumption implies that, for any $m \in \mathbb{Z}$, the graded $\overline{\mathbb{Q}}_l$–vector space

$$\bigoplus_i H^m(\mathcal{K}^i_{\alpha',u_\alpha})[-i]$$

is equipped with a natural free $\Lambda_{A_\alpha,u_\alpha}$–action induced by (17). This is explained in [40, Proposition 7.7.4], which essentially relies on the polarizability of $P$, ie the axiom (iii) of weak abelian fibrations. (Since this part only concerns the group scheme $P$, the proof of [40, Proposition 7.7.4] applies identically here.)

**Step C** We complete the induction argument.

The condition (a) of Theorem 1.1 — ie the decomposition theorem for $Rf_*\mathcal{K}$ — implies the degeneracy of the spectral sequence

$$H^j({}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K})_{u_\alpha}) \Rightarrow H^{i+j}(M_{u_\alpha}, \iota^*_{u_\alpha}\mathcal{K}), \tag{30}$$

where $\iota_{u_\alpha}: M_{u_\alpha} \hookrightarrow M$ is the geometric fiber over $u_\alpha$. This induces a $\Lambda_{A_{\alpha,u_\alpha}}$–stable filtration $F^\bullet\mathbb{H}$ on the total cohomology

$$\mathbb{H} := \bigoplus_i H^i(M_{u_\alpha}, \iota^*_{u_\alpha}\mathcal{K})[-i],$$

whose $m^{\text{th}}$ graded piece is

$$F^m\mathbb{H}/F^{m+1}\mathbb{H} = \bigoplus_i H^m({}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K})_{u_\alpha})[-i-m]. \tag{31}$$

In addition, we have the following:

(a) By picking a quasi-lifting $A_{\alpha,u_\alpha} \to P_{\alpha,u_\alpha}$ as in the last paragraph of Step A, it follows from Proposition 1.6 and the $P$–equivariance of $\mathcal{K}$ that $\mathbb{H}$ is a free graded $\Lambda_{A_{\alpha,u_\alpha}}$–module.

(b) Since
$$\bigoplus_i {}^{\mathfrak{p}}\mathcal{H}^i(Rf_*\mathcal{K})|_{B_\alpha}[-i] = \bigoplus_{\alpha'}\bigoplus_i \mathcal{K}^i_{\alpha',u_\alpha}[-i],$$
the graded piece (31), as a graded $\Lambda_{A_{\alpha,u_\alpha}}$–module, is a direct sum of the graded $\Lambda_{A_{\alpha,u_\alpha}}$–modules
$$\bigoplus_i H^m(\mathcal{K}^i_{\alpha',u_\alpha})[-i-m]$$
over all $\alpha'$ with $Z_\alpha \subset Z_{\alpha'}$.

(c) By Step B, the induction assumption implies that each
$$\bigoplus_i H^m(\mathcal{K}^i_{\alpha',u_\alpha})[-i]$$
is a free graded $\Lambda_{A_{\alpha,u_\alpha}}$–module when $Z_\alpha \subset Z_{\alpha'}$.

(d) The graded $\Lambda_{A_\alpha,u_\alpha}$–module vanishes:
$$\bigoplus_i H^m(\mathcal{K}^i_{\alpha,u_\alpha})[-i] = \bigoplus_i H^{m+\dim V_\alpha}(\mathcal{L}^i_{\alpha,u_\alpha})[-i] = 0$$
if $m \neq -\dim V_\alpha$ for degree reasons, since $\mathcal{L}^i_{\alpha,u_\alpha}$ is a skyscraper sheaf supported at $u_\alpha$.

Recall the filtration $F^\bullet\mathbb{H}$ associated with the spectral sequence (30), whose graded pieces are given by (31). We arrive at exactly the situation of the last paragraph of [40, page 131]: the spectral sequence (30) induces a 3–layer filtration of $\Lambda_{A_\alpha,u_\alpha}$–modules
$$0 \subseteq F^{n+1}\mathbb{H} \subseteq F^n\mathbb{H} \subseteq \mathbb{H}, \quad n = -\dim V_\alpha = -\dim Z_\alpha,$$
where

- $\mathbb{H}$ is free by (a), and
- $F^{n+1}\mathbb{H}$ and $\mathbb{H}/F^n\mathbb{H}$ are free by (c) and (d). In fact, (d) ensures that
$$\bigoplus_i H^m(\mathcal{K}^i_{\alpha,u_\alpha})[-i]$$
vanishes when $m \neq n$, and therefore $F^m\mathbb{H}/F^{m+1}\mathbb{H}$ is free by (c).

This implies the freeness for

$$F^n\mathbb{H}/F^{n+1}\mathbb{H}=\Big(\bigoplus_i \mathcal{L}^i_{\alpha,u_\alpha}[-i-n]\Big)\oplus\Big(\bigoplus_{\alpha'\neq\alpha}\bigoplus_i H^n(\mathcal{K}^i_{\alpha',u_\alpha})[-i-n]\Big),$$

which completes the induction; see [40, pages 131–132]. $\square$

**Remark 1.7** We fixed some minor typos in [40, Section 7.7.2]: the correct formula for the $m^{\text{th}}$ graded piece [40, page 131, line 9] of the Leray spectral sequence is

$$H^m\Big(\bigoplus_n {}^{\mathfrak{p}}\mathcal{H}^n(Rf_*\overline{\mathbb{Q}}_l)_{s_0}\Big)[-n-m],$$

which is not equal to

$$H^m\Big(\bigoplus_n {}^{\mathfrak{p}}\mathcal{H}^n(Rf_*\overline{\mathbb{Q}}_l[-n])_{s_0}\Big)[-m]$$

as stated in [40]. As consequences, the following statements in the last paragraph of [40, page 131] are incorrect:

- For $\alpha'\neq\alpha$, we have that $H^m\big(\bigoplus_{n\in\mathbb{Z}}\mathcal{K}^n_{\alpha',u_\alpha}[-n]\big)$ is a free $\Lambda_{A_{u_\alpha}}$–module.
- For $\alpha=\alpha'$, we have that $H^m(\mathcal{K}_{\alpha,u_\alpha})=0$ unless $m=-\dim Z_\alpha$.

Their corrected versions are given in (b), (c) and (d) of Step C above.

## 1.7 Spread out for $\mathbb{C}$

As a corollary of Theorem 1.1, the following theorem concerns the intersection cohomology complex of a weak abelian fibration $(M,P,B)$ over the complex numbers $\mathbb{C}$.

**Theorem 1.8** *Suppose that $(M,P,B)$ is a weak abelian fibration over $\mathbb{C}$ of relative dimension $d$, ie the triple satisfies* (i)–(iii) *of Section 1.1. Assume that*

$$\tau_{>2d}(Rf_*\,\mathrm{IC}_M[-\dim M])=0. \tag{32}$$

*Then any support $Z$ of the decomposition for $Rf_*\,\mathrm{IC}_M$ satisfies the inequality*

$$\operatorname{codim} Z\leq\delta_Z.$$

**Proof** By a standard spreading out argument (see for example [1, Section 6]), we may reduce Theorem 1.8 to the same statement over finite fields. More precisely, we spread out the weak abelian fibration $(M,P,B)$ over $\operatorname{Spec} R$ where $R$ is a DVR of characteristic 0, such that the geometric fiber over a general prime $\mathfrak{p}\in\operatorname{Spec} R$ is a weak abelian fibration in characteristic $p$ as in the beginning of Section 1.1.

Moreover, the condition (32) holds over a general prime in Spec $R$. Therefore, if a support $Z$ of the decomposition theorem associated with $(M, P, B)$ violates the inequality $\operatorname{codim} Z \le \delta_Z$, then by spreading out, this inequality is violated by a support over a general prime $\mathfrak{p} \in \operatorname{Spec} R$ as well, which contradicts our assumption that Theorem 1.8 holds over finite fields..

In our setting, note that the complex

$$\mathcal{K} = \mathrm{IC}_M[-\dim M]$$

is $P$–equivariant, which satisfies (a)–(c) of Theorem 1.1 by the decomposition theorem, Verdier duality, and the condition (32), respectively. Hence we conclude Theorem 1.8 from Theorem 1.1. □

In order to apply the support theorem to the intersection cohomology complex for a weak abelian fibration with *singular* ambient space $M$, the crucial point is to verify the "relative dimension bound" (32). We discuss systematically in the next two sections how to obtain such a bound for the moduli of 1–dimensional sheaves and the moduli of semistable Higgs bundles.

# 2 Moduli of 1–dimensional sheaves and Higgs bundles

## 2.1 Overview

Throughout the rest of the paper, we work over the complex numbers $\mathbb{C}$. We show in this section that the morphisms (2) and (3) admit the structures of weak abelian fibrations.

A crucial technical result is Proposition 2.6, concerning a dimension bound for certain moduli of pure 1–dimensional sheaves. As a consequence, we verify in Theorem 2.3 the irreducibility of the moduli spaces $M_{\beta,\chi}^L$ of (A), which may be of independent interest.

The dimension bound given by Proposition 2.6 will be used again in Section 3, which plays an important role in the proof of our main theorems.

## 2.2 Curves in del Pezzo surfaces

Let $S$ be a del Pezzo surface, ie a nonsingular projective surface with $-K_S$ ample.

**Lemma 2.1** *Let $E$ be an effective divisor on $S$. Then*

$$\dim H^1(S, \mathcal{O}_S(E)) = \dim H^2(S, \mathcal{O}_S(E)) = 0.$$

*In particular*, *we have* $\dim H^0(S, \mathcal{O}_S(E)) = \frac{1}{2} E \cdot (E - K_S) + 1.$

**Proof** By Serre duality, we obtain

$$H^2(S, \mathcal{O}_S(E)) = H^0(S, \mathcal{O}_S(K_S - E))^\vee = 0.$$

Now we prove the vanishing of

$$H^1(S, \mathcal{O}_S(E))^\vee = H^1(S, \mathcal{O}_S(K_S - E)). \tag{33}$$

We consider the short exact sequence

$$0 \to \mathcal{O}_S(K_S - E) \to \mathcal{O}_S(K_S) \to \mathcal{O}_E(K_S) \to 0,$$

which induces the long exact sequence

$$\cdots \to H^0(S, \mathcal{O}_E(K_S)) \to H^1(S, \mathcal{O}_S(K_S - E)) \to H^1(S, \mathcal{O}_S(K_S)) \to \cdots. \tag{34}$$

The vanishing $H^0(S, \mathcal{O}_E(K_S)) = 0$ follows from $\deg_E(K_S) < 0$, and Serre duality yields the vanishing $H^1(S, \mathcal{O}_S(K_S)) = 0$. Hence (34) implies the vanishing of (33).

The last statement follows from the Riemann–Roch formula. □

Let $\beta$ be an ample and effective class on $S$. Then Lemma 2.1 implies that the base $B = \mathbb{P} H^0(S, \mathcal{O}_S(\beta))$ of (2) is of dimension

$$\dim B = \tfrac{1}{2}\beta \cdot (\beta - K_S).$$

We define $\pi : \mathcal{C}_B \to B$ to be the universal curve for the linear system $|\beta|$. Since $\beta$ is ample, it is basepoint free on the del Pezzo surface $S$. Hence the Bertini theorem implies that a general member of $|\beta|$ is a nonsingular and integral curve of genus

$$g_\beta = \tfrac{1}{2}\beta \cdot (\beta + K_S) + 1.$$

In particular, there exists a Zariski open dense subset $U \subset S$ such that the restriction of $\mathcal{C}_B$ to $U$ is smooth,

$$\pi : \mathcal{C} \to U \subset B.$$

We consider the relative degree-0 Picard variety

$$P := \mathrm{Pic}^0(\mathcal{C}_B / B)$$

parametrizing line bundles on the fibers of $\pi : \mathcal{C}_B \to B$ whose restrictions to each irreducible component are of degree 0. The projection morphism

$$\pi_P : P \to B$$

has fibers of pure dimension $g_\beta$. The restriction of $P$ to $U$ gives a smooth abelian scheme

$$\pi_P : P_U \big(:= \mathrm{Pic}^0(\mathcal{C}_U / U)\big) \to U.$$

We also consider the relative degree $e$ Picard variety over $U$

$$\pi_{P^e}\colon P_U^e\big(:=\mathrm{Pic}^e(\mathcal{C}_U/U)\big)\to U$$

for any integer $e$. We recall the following well-known fact [5, Lemma 1.3.5] concerning the variation of Hodge structures for Picard varieties of smooth curves.

**Proposition 2.2** *For any $e\in\mathbb{Z}$, we have an isomorphism of variations of Hodge structures on $U$:*

$$R^i\pi_{P^e*}\mathbb{Q}_{P_U^e}\simeq \bigwedge^i R^1\pi_*\mathbb{Q}_{\mathcal{C}}.$$

### 2.3 Moduli spaces of 1–dimensional sheaves

Now assume that $S$ is a *toric* del Pezzo surface with a polarization $L$. The moduli space $M_{\beta,\chi}^L$ parametrizes $S$–equivalence classes of pure 1–dimensional (Gieseker-)semistable sheaves $\mathcal{F}$ on $S$ with

$$\mathrm{supp}(\mathcal{F})=\beta,\quad \chi(\mathcal{F})=\chi.$$

Here the semistability is with respect to the slope function

$$\mu(\mathcal{E})=\frac{\chi(\mathcal{E})}{c_1(\mathcal{E})\cdot L}.$$

We recall the Hilbert–Chow morphism

$$h\colon M_{\beta,\chi}^L\to B,\quad \mathcal{F}\mapsto \mathrm{supp}(\mathcal{F}),$$

defined by taking the Fitting support [29]. The open subvariety $h^{-1}(U)\subset M_{\beta,\chi}^L$ parametrizes line bundles supported on the nonsingular curves in $|\beta|$. Hence every fiber of $h$ over a closed point $b\in U$ is an abelian variety of dimension $g_\beta$, and we have

$$h^{-1}(U)=\mathrm{Pic}^e(\mathcal{C}_U/U),\quad e=\chi-1+g_\beta. \tag{35}$$

The moduli space $M_{\beta,\chi}^L$ can be viewed as a compactification of the relative Picard variety (35).

The following theorem is of independent interest, and we postpone its proof to Section 2.6.

**Theorem 2.3** *The moduli space $M_{\beta,\chi}^L$ is irreducible of dimension*

$$\dim M_{\beta,\chi}^L=\beta^2+1=\dim B+g_\beta.$$

The group scheme $\pi_P\colon P \to B$ acts naturally and fiberwise on the moduli space $M^L_{\beta,\chi}$ via tensor product,

$$\mathcal{L}\cdot\mathcal{F} = \mathcal{L}\otimes\mathcal{F} \quad \text{for } \mathcal{L}\in P_b = \pi_P^{-1}(b),\ \mathcal{F}\in h^{-1}(b),$$

with $b \in B$ a closed point.

**Proposition 2.4** *The triple $(M^L_{\beta,\chi}, P, B)$ with $h$ and $\pi_P$ above form a weak abelian fibration of relative dimension $g_\beta$.*

**Proof** We need to check (i)–(iii) of Section 1.1. Recall from Section 2.2 that the group scheme $\pi_P\colon P \to B$ is smooth and its fibers are of pure dimension $g_\beta$. Hence the condition (i) follows from Theorem 2.3. The affineness of the stabilizers (condition (ii)) is proven in [7, Lemma 3.5.4], and the polarizability of the Tate module (condition (iii)) associated with the group scheme $P$ is given by [4, Theorem 3.3.1] as explained in [7, Lemma 3.5.5]. □

## 2.4 Moduli stacks

For the polarized surface $(S, L)$, the moduli of semistable sheaves can be constructed as a GIT-quotient of the corresponding Quot-scheme (denoted by Quot),

$$M^L_{\beta,\chi} = \mathrm{Quot}^{\mathrm{ss}} /\!\!/ \mathrm{GL}_m,$$

where the semistable part of the Quot scheme $\mathrm{Quot}^{\mathrm{ss}}$ and $m$ rely on the Hilbert polynomial $\dim H^0(S, \mathcal{F}\otimes L^{\otimes n})$ of a semistable sheaf $\mathcal{F}$ with $\mathrm{supp}(\mathcal{F}) = \beta$ and $\chi(\mathcal{F}) = \chi$. We also consider the moduli stack of semistable sheaves

$$\mathcal{M}^L_{\beta,\chi} = [\mathrm{Quot}^{\mathrm{ss}}/\mathrm{GL}_m]$$

such that the natural projection

$$q\colon \mathcal{M}^L_{\beta,\chi} \to M^L_{\beta,\chi}$$

induces a *good moduli space* of the Artin stack $\mathcal{M}^L_{\beta,\chi}$.

**Lemma 2.5** *The stack $\mathcal{M}^L_{\beta,\chi}$ is nonsingular of dimension*

$$\dim \mathcal{M}^L_{\beta,\chi} = \beta^2.$$

**Proof** The obstruction space for a semistable sheaf $\mathcal{F}\in\mathcal{M}^L_{\beta,\chi}$ is

$$\mathrm{Ext}^2_S(\mathcal{F},\mathcal{F}) = \mathrm{Hom}_S(\mathcal{F},\mathcal{F}\otimes\omega_S)^\vee, \quad \omega_S = \mathcal{O}_S(K_S). \tag{36}$$

We prove in the following that (36) vanishes.

By the semicontinuity of $\mathrm{Hom}_S(\mathcal{F}, \mathcal{F} \otimes \omega_S)$, it suffices to show the vanishing

$$\mathrm{Hom}_S(\mathcal{F}, \mathcal{F} \otimes \omega_S) = 0$$

when $\mathcal{F}$ is a polystable sheaf on $S$. Hence we only need to prove the vanishing

$$\mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_2 \otimes \omega_S) = 0 \tag{37}$$

for two stable sheaves $\mathcal{F}_1$ and $\mathcal{F}_2$ on $S$ with the same slope

$$\mu(\mathcal{F}_1) = \mu(\mathcal{F}_2).$$

Since $-K_S$ is effective for the del Pezzo surface $S$, we have a short exact sequence

$$0 \to \mathcal{F}_2 \otimes \omega_S \to \mathcal{F}_2 \to \mathcal{F}_2|_E \to 0,$$

where $E$ is a curve in the linear system $|-K_S|$. The induced long exact sequence gives

$$0 \to \mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_2 \otimes \omega_S) \to \mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_2) \to \mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_2|_E). \tag{38}$$

When $\mathcal{F}_1 \neq \mathcal{F}_2$, by the stability we have $\mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_2) = 0$. When $\mathcal{F}_1 = \mathcal{F}_2$, the second map of (38) is injective:

$$\mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_1) = \mathbb{C} \cdot \mathrm{id} \hookrightarrow \mathrm{Hom}_S(\mathcal{F}_1, \mathcal{F}_1|_E).$$

In particular, (37) vanishes in either case. This implies the vanishing of the obstruction (36) and proves that $\mathcal{M}^L_{\beta,\chi}$ is nonsingular. Consequently, we have

$$\dim \mathcal{M}^L_{\beta,\chi} = \dim \mathrm{Ext}^1_S(\mathcal{F}, \mathcal{F}) - \dim \mathrm{Hom}_S(\mathcal{F}, \mathcal{F}) = -\chi(\mathcal{F}, \mathcal{F}) = \beta^2. \qquad \square$$

Combining with the Hilbert–Chow morphism $h\colon M^L_{\beta,\chi} \to B$, we obtain a morphism

$$h_{\mathcal{M}}\colon \mathcal{M}^L_{\beta,\chi} \to B. \tag{39}$$

**Proposition 2.6** *Let $S$ be a toric del Pezzo surface. For any closed point $b \in B$, we have the following dimension bound for the fiber of* (39):

$$\dim h_{\mathcal{M}}^{-1}(b) \le \tfrac{1}{2}\beta \cdot (\beta + K_S). \tag{40}$$

When $b$ represents an integral nonsingular curve, then $h_{\mathcal{M}}^{-1}(b)$ is exactly a connected component of its Picard stack whose dimension $g_\beta - 1$ matches the righthand side of (40).

We prove Proposition 2.6 in Section 2.5. Then in Section 2.6 we use Proposition 2.6 to complete the proof of Theorem 2.3.

## 2.5 Proof of Proposition 2.6

We reduce Proposition 2.6 to a dimension bound for the nilpotent cone of Higgs bundles.

Let $C$ be a nonsingular curve of genus $g$, and let $D$ be a degree $d$ effective divisor on $C$ with

$$d > 2g-2. \tag{41}$$

We denote by $\mathcal{M}^{\mathrm{nil}}_{n,\chi}$ the moduli stack of nilpotent Higgs bundles

$$(\mathcal{E},\theta) \quad \text{with } \theta\colon \mathcal{E} \to \mathcal{E}\otimes \mathcal{O}_C(D),\ \mathrm{rank}(\mathcal{E}) = n,\ \chi(\mathcal{E}) = \chi,$$

where we do not impose any (semi)stability conditions.

The stack $\mathcal{M}^{\mathrm{nil}}_{n,\chi}$ is essentially the central fiber of the (stacky) Hitchin fibration. Alternatively, by the spectral correspondence, $\mathcal{M}^{\mathrm{nil}}_{n,\chi}$ parametrizes pure 1–dimensional sheaves $\mathcal{F}$ with

$$\mathrm{supp}(\mathcal{F}) = nC, \quad \chi(\mathcal{F}) = \chi,$$

on the total space $\mathrm{Tot}(\mathcal{O}_C(D))$ of the line bundle $\mathcal{O}_C(D)$. Here the spectral correspondence is induced by the pushforward along the standard projection $\mathrm{Tot}(\mathcal{O}_C(D)) \to C$.

**Proposition 2.7** (cf [8]) *We have*

$$\dim \mathcal{M}^{\mathrm{nil}}_{n,\chi} \leq n(g-1) + \tfrac{1}{2}n(n-1)d. \tag{42}$$

**Proof** The dimension formula for the stack of the nilpotent cone and the comparison to the righthand side of (42) are given in lines 2 and 6 of [8, page 725, Section 10]. Although it is assumed in the beginning of [8] that the curve $C$ has genus $g \geq 2$, the dimension calculation of [8, Section 10] does not require this constraint as long as (41) holds.[2] □

Now we prove Proposition 2.6.

We consider the maximal open torus $\mathbb{T} \subset S$ whose action on $S$ induces $\mathbb{T}$–actions on both the moduli stack $\mathcal{M}^L_{\beta,\chi}$ and the base $B$. By a semicontinuity argument (cf [18, Proof of Corollary 1]), it suffices to show (40) for all $T$–fixed points $b \in B$. Since we are only concerned with dimension counts, we prove the following stronger statement for toric divisors without imposing (semi)stability conditions.

[2] An alternative proof of this dimension bound can be obtained using the method of [43, Proposition 3.1]. We note that the last equation of [8, Section 10] shows that (42) still holds if $d = 2g-2$.

**Claim** *For an effective divisor*

$$E = \sum_i n_i E_i, \quad n_i > 0,$$

*with each $E_i$ a nonsingular irreducible toric divisor, we have*

$$\dim \mathcal{M}_{E,\chi} \leq \tfrac{1}{2} E \cdot (E + K_S). \tag{43}$$

*Here $\mathcal{M}_{E,\chi}$ stands for the moduli stack of pure 1–dimensional sheaves supported on $E \subset S$.*

We prove (43) by induction on the number of the irreducible components $\{E_i\}$.

For the induction base, we consider $E = nE'$ with $E'$ irreducible. Then $E' \simeq \mathbb{P}^1$ and the normal bundle $\mathcal{O}_{E'}(d)$ of $E'$ in $S$ satisfies

$$d = E'^2 = -2 + E' \cdot (-K_S) > -2. \tag{44}$$

Since the formal neighborhood of an irreducible toric divisor only depends on the degree of the normal bundle, the thickened curve $E = nE' \subset S$ is isomorphic to the $n^{\text{th}}$ thickening

$$nE' \subset \mathrm{Tot}(\mathcal{O}_{E'}(d))$$

of the 0–section in the total space of $\mathcal{O}_{E'}(d)$. Hence by Proposition 2.7, where the condition (41) is guaranteed by (44), we have

$$\dim \mathcal{M}_{nE',\chi} \leq -n + \tfrac{1}{2} n(n-1)d = nE' \cdot (nE' + K_S).$$

Here we used $E'^2 = d$ and $E' \cdot K_S = -d - 2$ in the last identity. This proves the induction base.

To complete the induction, we assume that $E = E' + E''$. Here

$$E' = \sum_i n_i E'_i \quad \text{and} \quad E'' = \sum_i m_i E''_i,$$

with $E'_i, E''_j$ irreducible toric divisors satisfying $E'_i \neq E''_j$ for any $i, j$.

**Lemma 2.8** *Let $\mathcal{F}$ be a pure 1–dimensional sheaf supported on $E$. Then there exists a canonical short exact sequence*

$$0 \to \mathcal{F}' \to \mathcal{F} \to \mathcal{F}'' \to 0,$$

*where $\mathcal{F}'$ and $\mathcal{F}''$ are pure 1–dimensional sheaves supported on $E'$ and $E''$, respectively.*

**Proof** We take

$$\mathcal{F}'' := (\mathcal{F}|_{E''})/\text{maximal 0–dimensional subsheaf of } \mathcal{F}|_{E''} \in \mathrm{Coh}(E''),$$
$$\mathcal{F}' := \mathrm{Ker}(\mathcal{F} \twoheadrightarrow \mathcal{F}|_{E''} \twoheadrightarrow \mathcal{F}'') \in \mathrm{Coh}(E').$$

Since $\mathcal{F}'$ is a subsheaf of $\mathcal{F}$, it is pure, supported on $E'$. □

Two sheaves $\mathcal{F}'$ and $\mathcal{F}''$ (as in Lemma 2.8) with different supports satisfy

$$\mathrm{Hom}_S(\mathcal{F}'', \mathcal{F}') = \mathrm{Hom}_{E'}(i^*\mathcal{F}'', \mathcal{F}') = 0,$$

where $i : E' \hookrightarrow S$ is the embedding and $i^*\mathcal{F}''$ is 0–dimensional on $E'$. Serre duality further implies

$$\mathrm{Ext}^2_S(\mathcal{F}'', \mathcal{F}') = \mathrm{Hom}_S(\mathcal{F}', \mathcal{F}'' \otimes \omega_S) = 0.$$

Hence by Lemma 2.8, after decomposing the stack $\mathcal{M}_{E,\chi}$ into strata, we obtain a morphism to

$$\bigsqcup_{\chi'+\chi''=\chi} \mathcal{M}_{E',\chi'} \times \mathcal{M}_{E'',\chi''},$$

whose closed fiber over $(\mathcal{F}', \mathcal{F}'')$ has dimension upper bound

$$\dim \mathrm{Ext}^1(\mathcal{F}'', \mathcal{F}') = \chi(\mathcal{F}'', \mathcal{F}') = E' \cdot E''.$$

Combining with the induction assumption on the dimensions of $\mathcal{M}_{E',\chi'}$ and $\mathcal{M}_{E'',\chi''}$, we conclude that

$$\dim \mathcal{M}_{E,\chi} \leq \tfrac{1}{2} E' \cdot (E' + K_S) + \tfrac{1}{2} E'' \cdot (E'' + K_S) + E' \cdot E'' = \tfrac{1}{2} E \cdot (E + K_S).$$

This completes the induction. □

## 2.6 Proof of Theorem 2.3

We first prove the irreducibility of $M^L_{\beta,\chi}$. Equivalently, we prove the irreducibility of the stack $\mathcal{M}^L_{\beta,\chi}$.

Recall the open subset $U \subset B$ formed by nonsingular curves in the linear system $|\beta|$. The open substack $h^{-1}_{\mathcal{M}}(U)$ parametrizes line bundles on these curves with Euler characteristic $\chi$. In particular, $h^{-1}_{\mathcal{M}}(U)$ is Zariski open and dense in an irreducible component of the relative Picard stack associated with the universal curve $\pi : \mathcal{C} \to U$. Assume $\mathcal{M}^L_{\beta,\chi}$ has another irreducible component $\mathcal{M}'$ which does not contain $h^{-1}_{\mathcal{M}}(U)$. By Lemma 2.5 it has dimension

$$\dim \mathcal{M}' = \beta^2,$$

and it maps to the complement $B \setminus U$ under the morphism (39). This implies that a general fiber of

$$h_{\mathcal{M}}|_{\mathcal{M}'} \colon \mathcal{M}' \to B \setminus U$$

has dimension at least

$$\dim \mathcal{M}' - (\dim B - 1) \geq \beta^2 - \tfrac{1}{2}\beta \cdot (\beta - K_S) + 1 > \tfrac{1}{2}\beta \cdot (\beta + K_S),$$

which contradicts Proposition 2.6. This completes the proof of the irreducibility of $M_{\beta,\chi}^L$. $\square$

## 2.7 Higgs bundles

Most of the statements for $M_{\beta,\chi}^L$ discussed above hold identically for the moduli spaces $\widetilde{M}_{n,\chi}$ of Higgs bundles in the case of (B). This is due to the fact that $\widetilde{M}_{n,\chi}$ can be viewed as the moduli space of 1–dimensional semistable sheaves $\mathcal{F}$ on $\mathrm{Tot}(\mathcal{O}_C(D))$ with

$$[\mathrm{supp}(\mathcal{F})] = n[C], \quad \chi(\mathcal{F}) = \chi,$$

via the spectral correspondence. We summarize these results in the following, for the reader's convenience.

Recall the universal spectral curve

$$\pi \colon \mathcal{C}_B \to B$$

with $\pi_P \colon P = \mathrm{Pic}^0(\mathcal{C}_B/B) \to B$ the relative degree 0 Picard variety. Similar to the case of $M_{\beta,n}^L$, the group scheme $P$ acts on $\widetilde{M}_{n,\chi}$ via tensor product

$$\mathcal{L} \cdot \mathcal{F} = \mathcal{L} \otimes \mathcal{F}, \quad \text{with } \mathcal{L} \in \pi_P^{-1}(b) \text{ and } \mathcal{F} \in h^{-1}(b) \text{ for all } b \in B.$$

Here we view a Higgs bundle as a pure 1–dimensional coherent sheaf supported on the spectral curve

$$\pi^{-1}(b) \subset \mathrm{Tot}(\mathcal{O}_C(D)).$$

The moduli stack of semistable Higgs bundles admits a morphism

$$q \colon \widetilde{\mathcal{M}}_{n,\chi} \to \widetilde{M}_{n,\chi},$$

which induces

$$h_{\mathcal{M}} = h \circ q \colon \widetilde{\mathcal{M}}_{n,\chi} \to B.$$

**Proposition 2.9** *Assume* $\deg(D) = d > 2g-2$.*The following statements hold*:

(a) *The fiber of* $h_{\mathcal{M}}$ *over a closed point* $b \in B$ *satisfies*

$$\dim h_{\mathcal{M}}^{-1}(b) \leq n(g-1) + \tfrac{1}{2}n(n-1)d.$$

(b) *The stack $\widetilde{\mathcal{M}}_{n,\chi}$ is irreducible and nonsingular of dimension*

$$\dim \widetilde{\mathcal{M}}_{n,\chi} = n^2 d.$$

(c) *The moduli space $\widetilde{M}_{n,\chi}$ is irreducible of dimension*

$$\dim \widetilde{M}_{n,\chi} = n^2 d + 1.$$

(d) *The triple $(\widetilde{M}_{n,\chi}, P, B)$ form a weak abelian fibration of relative dimension*

$$g_n = n(g-1) + \tfrac{1}{2}n(n-1)d + 1.$$

**Proof** These statements are parallel to Proposition 2.7, Lemma 2.5, Theorem 2.3, and Proposition 2.4. Statement (a) follows from a semicontinuity argument and Proposition 2.7. Statement (b) follows from Serre duality for semistable Higgs bundles [39, Corollary 2.6]. Statement (c) follows from (a) and (b), as explained in Section 2.6. Statement (d) is deduced by an identical proof as for Proposition 2.4. □

### 2.8 Assumptions on the curve class $\beta$

In Section 2, the ampleness assumption of the curve class $\beta$ is used for the following properties:

(I) The linear system $|\beta|$ is basepoint free.

(II) A general curve in $|\beta|$ is integral and nonsingular.

We may replace the ampleness assumption for $\beta$ by the conditions (I) and (II) above.

**Proposition 2.10** *Theorem 2.3 and Proposition 2.4 hold for any curve class $\beta$ which contains an integral curve in the linear system $|\beta|$.*

**Proof** Assume $C_0 \in |\beta|$ is integral. By the adjunction formula, either $C_0 \simeq \mathbb{P}^1$ is an exceptional divisor or $C_0^2 > 0$. In the first case, the moduli space is a reduced point. In the second case, we obtain that the divisor $C$ is integral and nef. Therefore (I) and (II) follow from [16, Corollary 4.7] and the Bertini theorem. □

## 3 Intersection cohomology complexes

We prove in this section a support inequality for the moduli spaces $M^L_{\beta,\chi}$ and $\widetilde{M}_{n,\chi}$.

**Theorem 3.1** *Let $h\colon M \to B$ be the morphism* (2) *or* (3). *We define the $\delta$–function on $B$ from the associated group scheme $P$ as in Section 1. Then any support $Z$ of $Rh_* \mathrm{IC}_M$ satisfies*

$$\operatorname{codim} Z \le \delta_Z.$$

By Theorem 1.8 and Propositions 2.4 and 2.9(d), it suffices to prove the following proposition concerning the intersection cohomology complex.

**Proposition 3.2** *We have*

$$\tau_{>2R}(Rh_* \mathrm{IC}_M[-\dim M]) = 0, \quad R := \dim M - \dim B. \tag{45}$$

## 3.1 Sketch of the proof of Proposition 3.2

Although Proposition 3.2 only concerns bounded complexes on schemes, our proof relies on *unbounded complexes* on *Artin stacks*. From now on, we work with the derived category $D_c(-, \mathbb{Q}_l)$ of constructible sheaves with $\mathbb{Q}_l$–coefficients for Artin stacks as in [26; 27]. We denote by $D_c^b(-, \mathbb{Q}_l)$, $D^-(-, \mathbb{Q}_l)$ and $D^+(-, \mathbb{Q}_l)$ the subcategories of complexes which are bounded, bounded from above, and bounded from below, respectively. In this section, we assume that all Artin stacks are of finite type. We use the six operations for Artin stacks following [26; 27; 28]. Furthermore, by [28], we also have the perverse $t$–structure in the unbounded setting.

Recall the morphism from the moduli stack to the moduli space of 1–dimensional semistable sheaves/Higgs bundles

$$q\colon \mathcal{M} \to M.$$

The composition of $q$ and $h\colon M \to B$ induces a morphism

$$h_{\mathcal{M}} = h \circ q\colon \mathcal{M} \to B.$$

We consider the (unbounded) complexes

$$Rh_{\mathcal{M}!}\mathbb{Q}_l \in D^-(B, \mathbb{Q}_l), \quad Rq_*\mathbb{Q}_l \in D^+(M, \mathbb{Q}_l).$$

We first prove Proposition 3.2 assuming the following two propositions which concern the stack $\mathcal{M}$.

**Proposition 3.3** *We have*

$$\tau_{>2R-2}(Rh_{\mathcal{M}!}\mathbb{Q}_l) = 0. \tag{46}$$

**Proposition 3.4** *There exists a splitting*

$$Rq_*\mathbb{Q}_l \simeq \mathrm{IC}_M[-\dim M] \oplus \mathcal{E} \in D^+(M, \mathbb{Q}_l). \tag{47}$$

**Proof of Proposition 3.2** Applying the dualizing functor to the isomorphism (47), we obtain

$$\mathbb{D}(Rq_*\mathbb{Q}_l) \simeq \mathrm{IC}_M[\dim M] \oplus \mathcal{E}', \quad \mathcal{E}' \in D^-(M, \mathbb{Q}_l).$$

Since $\mathcal{M}$ is nonsingular, the lefthand side is isomorphic to

$$Rq_!\mathbb{D}(\mathbb{Q}_l) = Rq_!\mathbb{Q}_l[2\dim\mathcal{M}] = Rq_!\mathbb{Q}_l[2\dim M - 2].$$

Combining the two equations above, we conclude that

$$Rq_!\mathbb{Q}_l \simeq \mathrm{IC}_M[-\dim M + 2] \oplus \cdots \quad \in D^-(M, \mathbb{Q}_l).$$

Hence, thanks to properness of $h\colon M \to B$, we have $Rh_! = Rh_*$, and the lefthand side of (45) (shifted by degree 2) is a direct sum component of the lefthand side of (46). In particular, Proposition 3.2 follows from Proposition 3.3. □

In the rest of Section 3, we prove Propositions 3.3 and 3.4.

## 3.2 Proof of Proposition 3.3

Proposition 3.3 is a consequence of the following well-known vanishing and the dimension bounds (Proposition 2.6 for (A) and Proposition 2.9(a) for (B)) obtained in Section 2.

**Lemma 3.5** *Let $\mathcal{Y}$ be an irreducible Artin stack of dimension $r$. Then for $n > 2r$ we have the following vanishing for compactly supported cohomology*:

$$H_c^n(\mathcal{Y}, \mathbb{Q}_l) = 0. \tag{48}$$

**Proof** In the special case when $\mathcal{Y}$ is nonsingular, (48) follows from the Verdier duality

$$H_c^n(\mathcal{Y}, \mathbb{Q}_l)^\vee = H^{2r-n}(\mathcal{Y}, \mathbb{Q}_l) = 0, \quad 2r - n < 0.$$

In general, since we are only concerned with the constructible sheaf $\mathbb{Q}_l$, we may assume that $\mathcal{Y}$ is reduced. Then by stratifying $\mathcal{Y}$ into locally closed nonsingular substacks and the excision sequences ([28, Example 2.1(iv)]), we reduce (48) for general $\mathcal{Y}$ to the nonsingular ones. □

Let $b \in B$ be a closed point. We denote by $\mathcal{M}_b$ the substack

$$\mathcal{M}_b := h_{\mathcal{M}}^{-1}(b) \subset \mathcal{M}.$$

Propositions 2.6 and 2.9(a) yield

$$\dim \mathcal{M}_b \leq R - 1, \quad \text{where } R = \dim M - \dim B.$$

Combining with Lemma 3.5, we conclude that the complex

$$(Rh_{\mathcal{M}!}\mathbb{Q}_l)_b = H_c^*(\mathcal{M}_b, \mathbb{Q}_l)$$

is concentrated in degrees $\leq 2(R-1)$ for any closed point $b \in B$. In particular,

$$(\tau_{>2R-2}(Rh_{\mathcal{M}!}\mathbb{Q}_l))_b = \tau_{>2R-2}((Rh_{\mathcal{M}!}\mathbb{Q}_l)_b) = 0 \quad \text{for all } b \in B.$$

This completes the proof of Proposition 3.3. □

## 3.3 Moduli of framed objects

The main difficulty for proving Proposition 3.4 is the nonproperness of the morphism $q : \mathcal{M} \to M$. In order to apply the decomposition theorem [1] to $q$, we use the moduli of framed objects [34] to "approximate" the stack $\mathcal{M}$. See [10; 11; 13; 14; 35; 38; 30] for applications of such techniques in the study of quivers representations and Donaldson–Thomas theory.

Let $M$ and $\mathcal{M}$ be the moduli space and the moduli stack of (A) or (B) in Section 0.1. Since in either case $M$ can be realized as a moduli space of semistable sheaves on an algebraic surface, we obtain (see Section 2.4) that $M$ can be realized as a GIT-quotient of a Quot-scheme

$$M = \mathrm{Quot}^{\mathrm{ss}} /\!\!/ \mathrm{GL}_m,$$

where the semistable locus $\mathrm{Quot}^{\mathrm{ss}} \subset \mathrm{Quot}$ is with respect to a $\mathrm{GL}_m$–linearized polarization $\mathcal{L}_m$ on Quot. The morphism $q$ is induced by the morphism from the stack to the corresponding good GIT-quotient:

$$\mathcal{M} = [\mathrm{Quot}^{\mathrm{ss}}/\mathrm{GL}_m] \xrightarrow{q} \mathrm{Quot}^{\mathrm{ss}} /\!\!/ \mathrm{GL}_m = M. \tag{49}$$

**Proposition 3.6** *For any $N > 0$, there exist a nonsingular scheme $M_f$ and a nonsingular Artin stack $\mathcal{X}_f$ with a commutative diagram*

$$\begin{tikzcd} M_f \arrow[rr, hook, "j"] \arrow[dr, "p_M"'] & & \mathcal{X}_f \arrow[dl, "p_{\mathcal{X}}"] \\ & \mathcal{M} & \end{tikzcd} \tag{50}$$

*satisfying the following properties*:

(a) *$p_{\mathcal{X}}$ is an affine space bundle*,

(b) *$j : M_f \hookrightarrow \mathcal{X}_f$ is an open immersion*,

(c) *the composition* $M_f \xrightarrow{p_M} \mathcal{M} \xrightarrow{q} M$ *is projective, and*

(d) *for the complement* $\mathcal{Z}_f := \mathcal{X}_f \setminus M_f$*, we have*

$$\operatorname{codim}_{\mathcal{X}_f}(\mathcal{Z}_f) > N.$$

**Proof** We complete the proof of Proposition 3.2 in the following three steps.

**Step 1** (quiver moduli) For a fixed integer $f > m$, let $Q_f$ be a quiver of two vertices $P_1$ and $P_2$ such that the dimension vector is $(1, m)$ and there are $f$ arrows from $P_1$ to $P_2$. Following King [25], the representation space

$$\mathbb{A} := \operatorname{Hom}(\mathbb{C}, \mathbb{C}^m)^f \simeq \mathbb{C}^{mf}$$

of the quiver $Q_f$ admits a natural action of the group

$$G_m := \mathrm{GL}_1 \times \mathrm{GL}_m.$$

Moreover, for any $\theta > 0$, the character

$$\chi_\theta \colon G_m \to \mathbb{C}^*, \qquad (g_1, g_m) \mapsto \det(g_1)^{-m\theta} \cdot \det(g_m)^\theta, \quad \text{where } g_i \in \mathrm{GL}_i,$$

yields a stability condition on $\mathbb{A}$. Here the stability is given by GIT associated with the trivial line bundle $\mathcal{O}_{\mathbb{A}}^\theta$, equipped with the $G_m$–linearization induced by $\chi_\theta$. We denote by $\mathbb{A}_\theta^{\mathrm{ss}} \subset \mathbb{A}$ the semistable locus with respect to $\theta$.

**Claim** *We have*

$$\operatorname{codim}_{\mathbb{A}}(\mathbb{A} \setminus \mathbb{A}_\theta^{\mathrm{ss}}) \to \infty \quad \textit{when } f \to \infty.$$

**Proof** If we view $\mathbb{A}$ as the parameter space of $m \times f$ matrices, the GIT-unstable loci are contained in the determinantal variety $D_{m-1} \subset \mathbb{A}$ formed by matrices of rank $< m$. Hence we have

$$\dim(\mathbb{A} \setminus \mathbb{A}_\theta^{\mathrm{ss}}) \le \dim D_{m-1} = (m-1)(f+1),$$

which implies that

$$\operatorname{codim}_{\mathbb{A}}(\mathbb{A} \setminus \mathbb{A}_\theta^{\mathrm{ss}}) \ge mf - (m-1)(f+1) = f + 1 - m \to \infty$$

when $f \to \infty$. □

**Step 2** (moduli of framed objects) The moduli space of framed objects [34] combines the quotients (49) and the quiver $Q_f$, which provides the scheme $M_f$ and the stack $\mathcal{X}_f$ for Proposition 3.6. We recall the construction as follows.

Consider the natural $G_m$–action on the product

$$\mathrm{Quot} \times \mathbb{A},$$

where the action on the first factor passes through the obvious $\mathrm{GL}_m$–action and the action on the second factor is given in Step 1. Since the diagonal torus

$$\mathrm{GL}_1 \hookrightarrow \mathrm{GL}_1 \times \mathrm{GL}_m = G_m$$

acts trivially on both Quot and $\mathbb{A}$, the $G_m$–action on $\mathrm{Quot} \times \mathbb{A}$ further passes through a $\mathrm{PG}_m (:= G_m/\mathrm{GL}_1)$–action.

We choose $a > 0$ such that $a > m\theta$, and we consider the $G_d$–linearization

$$\mathcal{L}^{a,\theta} := \mathcal{L}_m^{\otimes a} \boxtimes \mathcal{O}_{\mathbb{A}}^{\theta} \tag{51}$$

on $\mathrm{Quot} \times \mathbb{A}$. Let

$$(\mathrm{Quot} \times \mathbb{A})^{\mathrm{ss}} \subset \mathrm{Quot} \times \mathbb{A}$$

be the GIT-semistable locus associated with (51). By a calculation using the Hilbert–Mumford criterion, it was proven in [34] (under a more general setup) that we have the open immersions

$$\mathrm{Quot}^{\mathrm{ss}} \times \mathbb{A}_{\theta}^{\mathrm{ss}} \subset (\mathrm{Quot} \times \mathbb{A})^{\mathrm{ss}} \subset \mathrm{Quot}^{\mathrm{ss}} \times \mathbb{A}. \tag{52}$$

Here the first inclusion is given in [34, Remark 3.40] and the second inclusion is given in [34, Proposition 3.39].

We define

$$M_f := (\mathrm{Quot} \times \mathbb{A})^{\mathrm{ss}} /\!\!/ G_m = (\mathrm{Quot} \times \mathbb{A})^{\mathrm{ss}}/\mathrm{PG}_m, \quad \mathcal{X}_f := (\mathrm{Quot}^{\mathrm{ss}} \times \mathbb{A})/\mathrm{PG}_m. \tag{53}$$

Note that by [34, Proposition 3.39], the scheme $M_f$ can be interpreted as the moduli of framed objects, which is nonsingular by [34, Corollary 3.41].

**Step 3** (properties) We show that $M_f$ and $\mathcal{X}_f$ defined in (53) fit into the commutative diagram (50) and satisfy the properties (a)–(c). Moreover, they also satisfy (d) when $f \to \infty$.

The open immersion

$$j : M_f \hookrightarrow \mathcal{X}_f$$

is induced by the second inclusion of (52). The quotient stack $\mathcal{X}_f$ admits a natural map to $\mathcal{M}$ via the natural projection

$$p_{\mathcal{X}} : \mathcal{X}_f = (\mathrm{Quot}^{\mathrm{ss}} \times \mathbb{A})/\mathrm{PG}_m \to \mathrm{Quot}^{\mathrm{ss}}/\mathrm{PG}_m = \mathcal{M},$$

which is an $\mathbb{A}$–bundle. By setting $p_M = p_{\mathcal{X}} \circ j$, we obtain the commutative diagram (50) and (a) and (b) immediately. The property (c) follows from [34, Theorem 3.42].

We note that the morphism from $M_f$ to $M$ has a natural geometric interpretation. In fact, we have

$$M = (\mathrm{Quot}^{\mathrm{ss}} \times \mathbb{A}) /\!\!/ G_m$$

by [34, last paragraph before Section 3.7]. Therefore the contraction

$$q \circ p_M \colon M_f = (\mathrm{Quot} \times \mathbb{A})^{\mathrm{ss}} /\!\!/ G_m \to (\mathrm{Quot}^{\mathrm{ss}} \times \mathbb{A}) /\!\!/ G_m = M$$

can be viewed as a variation of GIT from the $G_m$–linearization (49) to the $G_m$–linearization

$$\mathcal{L}^{a,0} := \mathcal{L}_m^{\otimes a} \boxtimes \mathcal{O}_{\mathbb{A}}$$

with the trivial $G_m$–action on the second factor.

It remains to prove (d). By (52) and the claim in Step 1, we have

$$\mathrm{codim}_{\mathcal{X}_f}(\mathcal{Z}_f) \geq \mathrm{codim}_{\mathbb{A}}(\mathbb{A} \setminus \mathbb{A}_\theta^{\mathrm{ss}}) \to \infty \quad \text{when } f \to \infty. \qquad \square$$

## 3.4 Proof of Proposition 3.4

To construct the desired splitting, we follow the approach of [34]. Namely we use the construction in the previous section to approximate $\mathcal{M}$ with $M_f$ and apply the decomposition theorem to the proper morphism $q \circ p_M \colon M_f \to M$.

Fix $N > 0$ and choose $f$ as in Proposition 3.6. Let $i \colon \mathcal{Z}_f \hookrightarrow M_f$ denote the closed immersion which has codimension larger than $N$. Consider the excision triangle on $\mathcal{X}_f$,

$$i_! i^! \mathbb{Q}_l \to \mathbb{Q}_l \to R j_* j^* \mathbb{Q}_l \to i_! i^! \mathbb{Q}_l[1]. \tag{54}$$

Since $\mathcal{X}_f$ is nonsingular, we have

$$i^! \mathbb{Q}_l = \omega_{\mathcal{Z}_f}[-2 \dim \mathcal{X}_f].$$

Also, from Section V.2 of [23], we have that the complex $\omega_{\mathcal{Z}_f}$ is concentrated in degrees $[-2 \dim \mathcal{Z}_f, \infty]$. By combining these with the codimension bound for $\mathcal{Z}_f$, we have that the complex $i^! \mathbb{Q}_l$ is supported in degrees $[2N, \infty]$.

If we push forward the excision triangle (54) to $M$ along

$$q \circ p_{\mathcal{X}} \colon \mathcal{X}_f \to M,$$

we obtain the triangle

$$(55)\quad R(q \circ p_{\mathcal{X}})_* i_! i^! \mathbb{Q}_l \to R(q \circ p_{\mathcal{X}})_* \mathbb{Q}_l \to R(q \circ p_{\mathcal{X}})_* R j_* j^* \mathbb{Q}_l \to R(q \circ p_{\mathcal{X}})_* i_! i^! \mathbb{Q}_l[1].$$

Since the derived pushforward functor preserves the subcategory $D^{\geq 0}$ and its shifts, the leftmost term of (55) is supported in degrees $[2N, \infty]$ as well. Furthermore, by [28, Lemma 3.3], after changing $N$ by a bounded amount, we have the same support result for perverse cohomology sheaves. In other words, we have the vanishing

$$^{\mathfrak{p}}\tau_{\leq 2N} R(q \circ p_{\mathcal{X}})_* i_! i^! \mathbb{Q}_l = 0$$

and, by (55), the quasi-isomorphism

$$(56)\qquad ^{\mathfrak{p}}\tau_{<2N} R(q \circ p_{\mathcal{X}})_* \mathbb{Q}_l \xrightarrow{\simeq} {}^{\mathfrak{p}}\tau_{<2N} R(q \circ p_{\mathcal{X}})_* R j_* j^* \mathbb{Q}_l.$$

Finally, since $p_{\mathcal{M}} \colon \mathcal{X}_f \to \mathcal{M}$ is an affine space bundle, so that $R p_{\mathcal{X}*} \mathbb{Q}_l = \mathbb{Q}_l$, we can rewrite (56) as

$$(57)\qquad ^{\mathfrak{p}}\tau_{<2N} R q_* \mathbb{Q}_l \xrightarrow{\simeq} {}^{\mathfrak{p}}\tau_{<2N} R(q \circ p_{\mathcal{M}})_* \mathbb{Q}_l.$$

By the decomposition theorem for the proper, surjective morphism $q \circ p_{\mathcal{M}} \colon M_f \to M$, the righthand side of (57) can be noncanonically written as a direct sum of its (shifted) perverse cohomology sheaves:

$$(58)\qquad ^{\mathfrak{p}}\tau_{<2N} R(q \circ p_{\mathcal{M}})_* \mathbb{Q}_l \simeq \bigoplus_{k=\dim M}^{2N-1} \mathcal{P}_k[-k].$$

The lowest perverse cohomology sheaf $\mathcal{P}_{\dim M}$ occurs in degree $\dim M$, because of surjectivity of the morphism. After restricting to an open subset $V \subset M$ over which $q \circ p_{\mathcal{M}}$ is smooth, it is given by the shifted local system whose fiber over $x \in V$ is $H^0(M_{f,x}, \mathbb{Q}_l)$, with $M_{f,x}$ the closed fiber of $M_f$ over $x$. In particular, $\mathcal{P}_{\dim M}$ contains $\mathrm{IC}_M$ as a direct summand.

If we combine the splitting (58) with (57), we see that the composition

$$\mathrm{IC}_M[-\dim M] \to \mathcal{P}_{\dim M}[-\dim M] \to {}^{\mathfrak{p}}\tau_{<2N} R q_* \mathbb{Q}_l$$

admits a splitting

$$u_N \colon {}^{\mathfrak{p}}\tau_{<2N} R q_* \mathbb{Q}_l \to \mathrm{IC}_M[-\dim M].$$

As Ext-groups between perverse sheaves vanish outside of bounded degree, for $N$ sufficiently large, we have the stabilization

$$\mathrm{Hom}({}^{\mathfrak{p}}\tau_{<2N} Rq_*\mathbb{Q}_l, \mathrm{IC}_M[-\dim M]) = \mathrm{Hom}({}^{\mathfrak{p}}\tau_{<2N+1} Rq_*\mathbb{Q}_l, \mathrm{IC}_M[-\dim M]).$$

In other words, for the canonical morphism

$$v_N : {}^{\mathfrak{p}}\tau_{<2N} Rq_*\mathbb{Q}_l \to {}^{\mathfrak{p}}\tau_{<2(N+1)} Rq_*\mathbb{Q}_l,$$

the splittings $u_N$ are compatible in the sense that $u_N = u_{N+1} \circ v_N$.

By [26, Lemma 4.3.2], we have that the unbounded complex $Rq_*\mathbb{Q}_l$ is the homotopy colimit of its truncations, ie

$$Rq_*\mathbb{Q}_l = \operatorname*{hocolim}_{N\to\infty} {}^{\mathfrak{p}}\tau_{<2N} Rq_*\mathbb{Q}_l.$$

So as a result, the splittings $u_N$ yield a splitting

$$u\colon Rq_*\mathbb{Q}_l \to \mathrm{IC}_M[-\dim M],$$

and consequently, a direct summand decomposition (47) as desired. □

# 4 Proof of the main theorem

## 4.1 Overview

We complete the proof of Theorem 0.4. For the approach, we combine the support inequality of Theorem 3.1 and techniques of [8; 4].

## 4.2 $\delta$–inequalities for integral curves

As in Section 2.2, we consider the relative degree 0 Picard variety

$$\mathrm{Pic}^0(\mathcal{C}_B/B) \to B$$

associated with a family of curves $\pi_B : \mathcal{C}_B \to B$. We obtain the $\delta$–invariants computing the dimensions of the affine parts of the group schemes

$$\delta(b) := \dim(\mathrm{Pic}^0(\mathcal{C}_b)^{\mathrm{aff}}), \quad b \in B,$$

where $\mathcal{C}_b$ is the fiber of $\pi_B$ over $b$. For a closed subvariety $Z \subset B$, we define $\delta_Z$ to be $\delta(b)$ for a general point in $Z$.

In Section 4.2, we first focus on the case of a flat family of *integral* curves

$$\pi_B : \mathcal{C}_B \to B$$

satisfying that the compactified Jacobian

$$\mathrm{Pic}^0(\mathcal{C}_B/B) \subset \overline{\mathcal{J}\mathcal{C}}_B$$

(parametrizing degree 0 torsion-free sheaves on the curves $\mathcal{C}_b$) is nonsingular.

The following lemma is the "Severi inequality", which is parallel to [4, (41)] for Higgs bundles. The proof of [4, (41)] works identically here since the only structure used for the spectral curves $\mathcal{C}_B \to B$ in [4] is the smoothness of $\overline{\mathcal{J}\mathcal{C}}_B$. We give a proof for the reader's convenience.

**Lemma 4.1** *For any subvariety $Z \subset B$, we have*

$$\operatorname{codim} Z \geq \delta_Z.$$

**Proof** Let $\mathcal{C}_{\mathrm{univ}} \to B_{\mathrm{univ}}$ be a semiuniversal family of curves such that the family $\mathcal{C}_B \to B$ is induced by a map

$$\phi\colon B \to B_{\mathrm{univ}}.$$

Let $B^{\delta}_{\mathrm{univ}} \subset B_{\mathrm{univ}}$ be the locus given by $\{b \in B : \delta(b) = \delta\}$. Since $\overline{\mathcal{J}\mathcal{C}}_B$ is nonsingular, by the paragraph following [17, Theorem 2], the image $\phi(B) \subset B_{\mathrm{univ}}$ meets $B^{\delta}_{\mathrm{univ}}$ transversally. Hence for any irreducible subvariety $Z \subset B$ whose general points lie in $B^{\delta}_{\mathrm{univ}}$, we have

$$\dim Z \leq \dim(\phi(B) \cap B^{\delta}_{\mathrm{univ}}) = \dim \phi(B) + \dim B^{\delta}_{\mathrm{univ}} - \dim B_{\mathrm{univ}} \leq \dim B - \delta,$$

where the equality follows from the transversality. We conclude that

$$\operatorname{codim} Z = \dim B - \dim Z \geq \delta = \delta_Z. \qquad \square$$

Our major application of Lemma 4.1 is for curves in a linear system on a del Pezzo surface.

**Corollary 4.2** *Let $\beta$ be a curve class on a del Pezzo surface, let*

$$\pi_B : \mathcal{C}_B \to B = \mathbb{P}H^0(S, \mathcal{O}_S(\beta)) \tag{59}$$

*be the universal curve in the linear system, and let $\pi^{\circ}_B : \mathcal{C}^{\circ} \to B^{\circ}$ be the restriction of (59) to the subset $B^{\circ} \subset B$ of integral curves. Then for any irreducible subvariety $Z \subset B$ whose generic point lies in $B^{\circ}$, we have*

$$\operatorname{codim} Z \geq \delta_Z.$$

**Proof** Since the compactified Jacobian $\overline{\mathcal{JC}}_{\mathcal{C}^\circ}$ associated with $\pi_B^\circ\colon \mathcal{C}^\circ \to B^\circ$ is an open subvariety of the moduli of stable pure 1–dimensional sheaves supported in the class $\beta$, we deduce the smoothness of $\overline{\mathcal{JC}}_{\mathcal{C}^\circ}$ from the smoothness of the moduli stack (cf Lemma 2.5). Hence Corollary 4.2 follows by applying Lemma 4.1 to the family $\pi_B^\circ\colon \mathcal{C}^\circ \to B^\circ$. □

When the integral locus $B^\circ \subset B$ is nonempty, by Proposition 2.10 the moduli space $M_{\beta,\chi}^L$ is irreducible for any polarization $L$ and $\chi \in \mathbb{Z}$ satisfying that

$$\dim M_{\beta,\chi}^L = \dim \overline{\mathcal{JC}}_{B^\circ} = \beta^2 + 1.$$

We define the following invariant associated with the curve class $\beta$:

$$\Phi_\beta := \dim \operatorname{Pic}^0(\mathcal{C}^\circ/B^\circ) - 2\dim B = \dim M_{\beta,\chi}^L - 2\dim B = 1 + \beta\cdot K_S, \tag{60}$$

where the last equality follows from Lemma 2.1.

## 4.3 $\delta$–Inequalities for linear systems

In Section 4.3, we assume that $\beta$ is an ample curve class on a del Pezzo surface $S$. We introduce a stratification of

$$B = \mathbb{P}H^0(S, \mathcal{O}_S(\beta))$$

analogous to the stratification introduced in [8, Section 9] and [4, Section 5.2] for Higgs bundles.

We consider the $s$–tuples

$$\underline{\beta} = ((m_1,\beta_1),(m_2,\beta_2),\dots,(m_s,\beta_s)), \tag{61}$$

where $s \ge 1$, $m_i \ge 1$, and $\beta_i$ are (not necessarily distinct) curve classes on $S$ such that

(i) $\sum_{i=1}^s m_i\beta_i = \beta$, and

(ii) there exists an integral curve in $|\beta_i|$ for each $1 \le i \le s$.

The objects (61) are called *types* of the curves in the linear system $B = |\beta|$. Two such objects

$$\underline{\beta} = ((m_1,\beta_1),(m_2,\beta_2),\dots,(m_s,\beta_s)),$$
$$\underline{\beta}' = ((m_1',\beta_1'),(m_2',\beta_2'),\dots,(m_s',\beta_{s'}')),$$

are said to give the same type if $s = s'$ and there exists a bijection

$$\sigma\colon \{1,2,\dots,s\} \to \{1,2,\dots,s\}$$

such that $\beta_i = \beta'_{\sigma(i)}$ and $m_i = m'_{\sigma(i)}$. We have a stratification according to the types of the curves in $|\beta|$,

$$B = \bigsqcup_{\underline{\beta}} B_{\underline{\beta}},$$

where each $B_{\underline{\beta}}$ is a locally closed subset of $B$ formed by curves in $|\beta|$ of type $\underline{\beta}$:

$$B_{\underline{\beta}} = \left\{ E = \sum_i m_i E_i \in |\beta| \;\middle|\; E_i \in |\beta_i|,\ E_i \text{ are distinct integral curves} \right\}.$$

**Proposition 4.3** *Let $Z \subset B$ be an irreducible subvariety whose general points have type*

$$\underline{\beta} = ((m_1, \beta_1), (m_2, \beta_2), \dots, (m_s, \beta_s)).$$

*Then we have*

$$\Phi_\beta + \operatorname{codim} Z \geq \sum_{i=1}^{s} \Phi_{\beta_i} + \delta_Z.$$

**Proof** We apply a similar argument as in [4, Corollary 5.4.4] for Higgs bundles.

For a curve class $\beta_i$, we denote by $|\beta_i|^\circ$ the open subvariety of $|\beta_i| = \mathbb{P} H^0(S, \mathcal{O}_S(\beta))$ consisting of integral curves. We define

$$\mathcal{C}^\circ_{\beta_i} \to |\beta_i|^\circ, \quad \operatorname{Pic}_{\beta_i} \to |\beta_i|^\circ, \tag{62}$$

to be the universal curve and the corresponding relative degree 0 Picard variety over $|\beta_i|^\circ$. For a type $\underline{\beta}$ as in (61), we have a finite morphism

$$\lambda_{\underline{\beta}} : B'_{\underline{\beta_i}} := \prod_{i=1}^{s} |\beta_i| \to B, \quad (E_i)_{i=1}^s \mapsto \sum_{i=1}^{s} m_i E_i,$$

whose image is

$$\operatorname{Im}(\lambda_{\underline{\beta}}) = \overline{B}_{\underline{\beta_i}} \subset B.$$

The morphism $\lambda_{\underline{\beta}}$ sends the open subvariety $\prod_{i=1}^s |\beta_i|^\circ \subset B'_{\underline{\beta_i}}$ to $B_{\underline{\beta_i}} \subset B$.

Now we assume that $\eta \in B$ is the generic point of $Z$. By the first two paragraphs of [4, Proof of Corollary 5.4.4], there exists a point

$$\eta' = (\eta_1, \eta_2, \dots, \eta_s) \in \lambda_{\underline{\beta}}^{-1}(\eta) \subset B'_{\underline{\beta_i}}, \quad \eta_i \in |\beta_i|^\circ,$$

satisfying that

(i) $\dim \overline{\{\eta\}} = \dim \overline{\{\eta'\}} = \dim Z$,

(ii) $\dim(\operatorname{Pic}_{\beta,\eta})^{\mathrm{ab}} = \sum_{i=1}^s \dim(\operatorname{Pic}_{\beta_i,\eta_i})^{\mathrm{ab}}$.

Here for any connected commutative group scheme $P$, we use the notation $P^{\mathrm{ab}}$ to denote its abelian variety part in the Chevalley decomposition (15). Since by definition (see (60)) we have

$$\dim(\mathrm{Pic}_{\beta,\eta})^{\mathrm{ab}} = \Phi_\beta + \dim B - \delta_Z, \quad \dim(\mathrm{Pic}_{\beta_i,\eta_i})^{\mathrm{ab}} = \Phi_{\beta_i} + \dim|\beta_i| - \delta_{\overline{\{\eta_i\}}},$$

we obtain from (ii) that

$$\delta_Z + \sum_{i=1}^{s} \Phi_{\beta_i} - \Phi_\beta = \dim B - \sum_{i=1}^{s} \dim|\beta_i| + \sum_{i=1}^{s} \delta_{\overline{\{\eta_i\}}}. \tag{63}$$

Applying Corollary 4.2 to (62), we have

$$\delta_{\overline{\{\eta_i\}}} \le \dim|\beta_i| - \dim \overline{\{\eta_i\}},$$

which implies that the righthand side of (63) is

$$\le \dim B - \sum_{i=1}^{s} \dim \overline{\{\eta_i\}} = \dim B - \dim Z = \operatorname{codim} Z. \qquad \square$$

## 4.4 Higgs bundles

The analog of Proposition 4.3 for the moduli of Higgs bundles is exactly the inequality (75) of [4, Corollary 5.4.4]. Although the paper [4] works with Higgs bundles with $\gcd(n,\chi) = 1$, Corollary 5.4.4 only concerns the group scheme — the degree 0 Picard variety associated with the universal spectral curve, which is not constrained by the coprime assumption.

We now rewrite the inequality (75) of [4, Corollary 5.4.4] in Proposition 4.4 parallel to the form of Proposition 4.3.

Consider the Hitchin fibration

$$h\colon \widetilde{M}_{n,\chi} \to B$$

associated with $C$, $n$, $\chi$ and an effective divisor $D$ with degree $\deg(D) > 2g-2$. Recall from [4, Section 5.2] that the Hitchin base

$$B = \bigoplus_{i=1}^{n} H^0(C, \mathcal{O}(iD))$$

admits a stratification

$$B = \bigsqcup_{(n_\bullet, m_\bullet)} B_{n_\bullet, m_\bullet}. \tag{64}$$

Here a type of spectral curve is given by $(n_\bullet, m_\bullet)$ with

$$s \geq 1, \quad n_\bullet = (n_1, n_2, \ldots, n_s), \quad m_\bullet = (m_1, m_2 \ \ldots, m_s), \quad \sum_{i=1}^{s} m_i n_i = n,$$

and $B_{n_\bullet, m_\bullet}$ are formed by spectral curves $E \subset \mathrm{Tot}(\mathcal{O}_C(D))$ of the form $E = \sum_i m_i E_i$, with $E_i$ distinct integral spectral curves that are degree $n_i$ covers of the zero section $C$. This actually coincides with the notion of (61) since the class of any spectral curve in the surface $\mathrm{Tot}(\mathcal{O}_C(D))$ is of the form $\beta_i = n_i[C]$ with $[C]$ the curve class of the zero section $C \subset \mathrm{Tot}(\mathcal{O}_C(D))$. We refer to [4, Section 5] for more details about the stratification (64).

We define the invariant, similar to (60),

$$\Phi_n = \dim \widetilde{M}_{n,\chi} - 2 \dim B = 1 + (2g - 2 - \deg(D))n,$$

where we use the dimension formulas of [4, (77)] in the last identity.

**Proposition 4.4** *Let $Z \subset B$ be an irreducible subvariety whose general points have type $(n_\bullet, m_\bullet)$. Then we have*

$$\Phi_n + \operatorname{codim} Z \geq \sum_{i=1}^{s} \Phi_{n_i} + \delta_Z.$$

*Here $\delta_Z$ is defined via the relative degree 0 Picard variety associated with the spectral curves.*

## 4.5 Proof of Theorem 0.4

We complete the proof of Theorem 0.4 in this section. Let

$$h \colon M^L_{\beta,\chi} \to B$$

be the morphism (2). By Lemma 2.5, the open subvariety of stable sheaves

$$M^{L,s}_{\beta,\chi} \subset M^L_{\beta,\chi}$$

is nonsingular. So we have

$$(\mathrm{IC}_{M^L_{\beta,\chi}})|_{M^{L,s}_{\beta,\chi}} = \mathbb{Q}[\dim M^L_{\beta,\chi}].$$

In particular, the restriction of the direct image complex $Rh_* \mathrm{IC}_{M^L_{\beta,\chi}}$ to the open subset $U \subset B$ of nonsingular curves in $|\beta|$ satisfies

$$Rh_* \, \mathrm{IC}_{M^L_{\beta,\chi}}|_U \simeq \bigoplus_{i=0}^{2d} \bigwedge\nolimits^i R^1 \pi_* \mathbb{Q}[\dim M^L_{\beta,\chi} - i]. \tag{65}$$

Here $\pi : \mathcal{C} \to U \subset B$, and (65) is an isomorphism of variations of Hodge structures by Proposition 2.2. Hence, in order to prove Theorem 0.4, it suffices to show that the lefthand side of (4), as a bounded complex of perverse sheaves, has full support $B$.

Assume that the irreducible subvariety $Z \subset B$ is a support whose general point has type $\underline{\beta}$. We combine the inequalities of Proposition 4.3 and Theorem 3.1 to obtain

$$\Phi_\beta + \operatorname{codim} Z \geq \sum_{i=1}^{s} \Phi_{\beta_i} + \delta_Z \geq \sum_{i=1}^{s} \Phi_{\beta_i} + \operatorname{codim} Z,$$

which implies $\Phi_\beta \geq \sum_{i=1}^{s} \Phi_{\beta_i}$. Therefore,

$$1 - s \geq \Big(\beta - \sum_i \beta_i\Big) \cdot (-K_S). \tag{66}$$

Since $-K_S$ is ample and $\beta - \sum_i \beta_i \geq 0$, the only possibility for (66) to hold is $s = 1$ and $m_i = 1$. Equivalently, we have $Z = B$. This completes the proof of Theorem 0.4 for $M^L_{\beta,\chi}$.

The proof for $\widetilde{M}_{n,\chi}$ is identical, where we apply Theorem 3.1 and Proposition 4.4. □

## References


[1] **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 Zbl

[2] **P Bousseau**, *A proof of N Takahashi's conjecture for* $(\mathbb{P}^2, E)$ *and a refined sheaves/Gromov–Witten correspondence*, preprint (2020) arXiv 1909.02992v2

[3] **P Bousseau**, *Scattering diagrams, stability conditions, and coherent sheaves on* $\mathbb{P}^2$, J. Algebraic Geom. 31 (2022) 593–686 MR Zbl

[4] **M A de Cataldo**, *A support theorem for the Hitchin fibration: the case of* $\mathrm{SL}_n$, Compos. Math. 153 (2017) 1316–1347 MR Zbl

[5] **M A A de Cataldo**, **T Hausel**, **L Migliorini**, *Topology of Hitchin systems and Hodge theory of character varieties: the case* $A_1$, Ann. of Math. 175 (2012) 1329–1407 MR Zbl

[6] **M A de Cataldo**, **D Maulik**, **J Shen**, *Hitchin fibrations, abelian surfaces, and the* $P = W$ *conjecture*, J. Amer. Math. Soc. 35 (2022) 911–953 MR Zbl

[7] **M A A de Cataldo**, **A Rapagnetta**, **G Saccà**, *The Hodge numbers of O'Grady* 10 *via Ngô strings*, J. Math. Pures Appl. 156 (2021) 125–178 MR Zbl

[8] **P-H Chaudouard**, **G Laumon**, *Un théorème du support pour la fibration de Hitchin*, Ann. Inst. Fourier (Grenoble) 66 (2016) 711–727 MR Zbl

[9] **W-Y Chuang**, **D-E Diaconescu**, **G Pan**, *BPS states and the $P = W$ conjecture*, from "Moduli spaces" (L Brambila-Paz, O García-Prada, P Newstead, R P Thomas, editors), London Math. Soc. Lecture Note Ser. 411, Cambridge Univ. Press (2014) 132–150 MR Zbl

[10] **B Davison**, *Cohomological Hall algebras and character varieties*, Internat. J. Math. 27 (2016) art. id. 1640003 MR Zbl

[11] **B Davison**, *The integrality conjecture and the cohomology of preprojective stacks*, preprint (2017) arXiv 1602.02110v3

[12] **B Davison**, *Refined invariants of finite-dimensional Jacobi algebras*, preprint (2019) arXiv 1903.00659

[13] **B Davison**, *BPS Lie algebras and the less perverse filtration on the preprojective CoHA*, preprint (2020) arXiv 2007.03289v2

[14] **B Davison**, **S Meinhardt**, *Cohomological Donaldson–Thomas theory of a quiver with potential and quantum enveloping algebras*, Invent. Math. 221 (2020) 777–871 MR Zbl

[15] **M Demazure**, *Surfaces de del Pezzo, I–V*, from "Séminaire sur les singularités des surfaces" (M Demazure, H C Pinkham, B Teissier, editors), Lecture Notes in Math. 777, Springer (1980) MR Zbl

[16] **S Di Rocco**, *$k$–very ample line bundles on del Pezzo surfaces*, Math. Nachr. 179 (1996) 47–56 MR Zbl

[17] **B Fantechi**, **L Göttsche**, **D van Straten**, *Euler number of the compactified Jacobian and multiplicity of rational curves*, J. Algebraic Geom. 8 (1999) 115–133 MR Zbl

[18] **V Ginzburg**, *The global nilpotent variety is Lagrangian*, Duke Math. J. 109 (2001) 511–519 MR Zbl

[19] **M Groechenig**, **D Wyss**, **P Ziegler**, *Mirror symmetry for moduli spaces of Higgs bundles via $p$–adic integration*, Invent. Math. 221 (2020) 505–596 MR Zbl

[20] **G Harder**, **M S Narasimhan**, *On the cohomology groups of moduli spaces of vector bundles on curves*, Math. Ann. 212 (1975) 215–248 MR Zbl

[21] **N J Hitchin**, *The self-duality equations on a Riemann surface*, Proc. London Math. Soc. 55 (1987) 59–126 MR Zbl

[22] **N Hitchin**, *Stable bundles and integrable systems*, Duke Math. J. 54 (1987) 91–114 MR Zbl

[23] **B Iversen**, *Cohomology of sheaves*, Springer (1986) MR Zbl

[24] **D Joyce**, **Y Song**, *A theory of generalized Donaldson–Thomas invariants*, Mem. Amer. Math. Soc. 1020, Amer. Math. Soc., Providence, RI (2012) MR Zbl

[25] **A D King**, *Moduli of representations of finite-dimensional algebras*, Quart. J. Math. Oxford Ser. 45 (1994) 515–530 MR Zbl

[26] **Y Laszlo**, **M Olsson**, *The six operations for sheaves on Artin stacks, I: Finite coefficients*, Publ. Math. Inst. Hautes Études Sci. 107 (2008) 109–168 MR Zbl

[27] **Y Laszlo**, **M Olsson**, *The six operations for sheaves on Artin stacks, II: Adic coefficients*, Publ. Math. Inst. Hautes Études Sci. 107 (2008) 169–210 MR Zbl

[28] **Y Laszlo**, **M Olsson**, *Perverse t–structure on Artin stacks*, Math. Z. 261 (2009) 737–748 MR Zbl

[29] **J Le Potier**, *Faisceaux semi-stables de dimension* 1 *sur le plan projectif*, Rev. Roumaine Math. Pures Appl. 38 (1993) 635–678 MR Zbl

[30] **J Manschot**, **S Mozgovoy**, *Intersection cohomology of moduli spaces of sheaves on surfaces*, Selecta Math. 24 (2018) 3889–3926 MR Zbl

[31] **D Maulik**, **J Shen**, *Endoscopic decompositions and the Hausel–Thaddeus conjecture*, Forum Math. Pi 9 (2021) art. id. e8, MR Zbl

[32] **D Maulik**, **Y Toda**, *Gopakumar–Vafa invariants via vanishing cycles*, Invent. Math. 213 (2018) 1017–1097 MR Zbl

[33] **D Maulik**, **Z Yun**, *Macdonald formula for curves with planar singularities*, J. Reine Angew. Math. 694 (2014) 27–48 MR Zbl

[34] **S Meinhardt**, *Donaldson–Thomas invariants vs intersection cohomology for categories of homological dimension one*, preprint (2015) arXiv 1512.03343

[35] **S Meinhardt**, **M Reineke**, *Donaldson–Thomas invariants versus intersection cohomology of quiver moduli*, J. Reine Angew. Math. 754 (2019) 143–178 MR Zbl

[36] **A Mellit**, *Poincaré polynomials of moduli spaces of Higgs bundles and character varieties (no punctures)*, Invent. Math. 221 (2020) 301–327 MR Zbl

[37] **L Migliorini**, **V Shende**, *A support theorem for Hilbert schemes of planar curves*, J. Eur. Math. Soc. 15 (2013) 2353–2367 MR Zbl

[38] **S Mozgovoy**, **M Reineke**, *Intersection cohomology of moduli spaces of vector bundles over curves*, preprint (2015) arXiv 1512.04076

[39] **S Mozgovoy**, **O Schiffmann**, *Counting Higgs bundles and type A quiver bundles*, Compos. Math. 156 (2020) 744–769 MR Zbl

[40] **B C Ngô**, *Le lemme fondamental pour les algèbres de Lie*, Publ. Math. Inst. Hautes Études Sci. 111 (2010) 1–169 MR Zbl

[41] **R Pandharipande**, **R P Thomas**, *Curve counting via stable pairs in the derived category*, Invent. Math. 178 (2009) 407–447 MR Zbl

[42] **M Saito**, *Mixed Hodge modules*, Publ. Res. Inst. Math. Sci. 26 (1990) 221–333 MR Zbl

[43] **O Schiffmann**, *Indecomposable vector bundles and stable Higgs bundles over smooth projective curves*, Ann. of Math. 183 (2016) 297–362 MR Zbl

[44] **O G Schiffmann**, *Kac polynomials and Lie algebras associated to quivers and curves*, from "Proceedings of the International Congress of Mathematicians, II: Invited lectures" (B Sirakov, P N de Souza, M Viana, editors), World Sci., Hackensack, NJ (2018) 1393–1424 MR Zbl

[45] **J Shen**, **Q Yin**, *Topology of Lagrangian fibrations and Hodge theory of hyper-Kähler manifolds*, Duke Math. J. 171 (2022) 209–241 MR Zbl

[46] **S Sun**, *Generic base change, Artin's comparison theorem, and the decomposition theorem for complex Artin stacks*, J. Algebraic Geom. 26 (2017) 513–555 MR Zbl

[47] **Y Toda**, *Stability conditions and curve counting invariants on Calabi–Yau* 3*–folds*, Kyoto J. Math. 52 (2012) 1–50 MR Zbl

[48] **Y Toda**, *Gopakumar–Vafa invariants and wall-crossing*, J. Differential Geom. 123 (2023) 141–193 MR

[49] **Z Yun**, *Langlands duality and global Springer theory*, Compos. Math. 148 (2012) 835–867 MR Zbl

[50] **Z Yun**, **W Zhang**, *Shtukas and the Taylor expansion of L–functions*, Ann. of Math. 186 (2017) 767–911 MR Zbl

[51] **X Zhu**, *An introduction to affine Grassmannians and the geometric Satake equivalence*, from "Geometry of moduli spaces and representation theory" (R Bezrukavnikov, A Braverman, Z Yun, editors), IAS/Park City Math. Ser. 24, Amer. Math. Soc., Providence, RI (2017) 59–154 MR Zbl

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

*Department of Mathematics, Yale University*
*New Haven, CT, United States*

maulik@mit.edu, junliang.shen@yale.edu

Proposed: Dan Abramovich Received: 10 January 2021
Seconded: Gang Tian, Richard P Thomas Revised: 1 October 2021

Geometry & Topology 27:4 (2023) 1587–1633
DOI: 10.2140/gt.2023.27.1587
Published: 15 June 2023

# Coarse injectivity, hierarchical hyperbolicity and semihyperbolicity

THOMAS HAETTEL
NIMA HODA
HARRY PETYT

We relate three classes of nonpositively curved metric spaces: hierarchically hyperbolic spaces, coarsely injective spaces and strongly shortcut spaces. We show that every hierarchically hyperbolic space admits a new metric that is coarsely injective. The new metric is quasi-isometric to the original metric and is preserved under automorphisms of the hierarchically hyperbolic space. We show that every coarsely injective metric space of uniformly bounded geometry is strongly shortcut. Consequently, hierarchically hyperbolic groups — including mapping class groups of surfaces — are coarsely injective and coarsely injective groups are strongly shortcut.

Using these results, we deduce several important properties of hierarchically hyperbolic groups, including that they are semihyperbolic, they have solvable conjugacy problem and finitely many conjugacy classes of finite subgroups, and their finitely generated abelian subgroups are undistorted. Along the way we show that hierarchically quasiconvex subgroups of hierarchically hyperbolic groups have bounded packing.

20F65, 20F67, 51F30

## 1 Introduction

A principal theme of geometric group theory is the study of groups as metric spaces. This includes studying groups via the types of metric spaces they act on. In this vein, the study of groups acting on spaces satisfying various forms of nonpositive curvature conditions has been especially fruitful. In this article, we are concerned with three classes of spaces exhibiting nonpositive curvature: hierarchically hyperbolic spaces, coarsely injective spaces and strongly shortcut spaces.

## 1.1 The setting

The first of our three classes is that of *hierarchically hyperbolic spaces*, which exhibit hyperbolic-like behaviour. Behrstock, Hagen and Sisto introduced these in [11] and, with Martin [9], showed they include many quotients of mapping class groups, while Hagen and Susse [48] added all known cubical groups to a growing body of interesting examples. The theory has had a number of successes: Behrstock, Hagen and Sisto [12] proved Farb's quasiflats conjecture for mapping class groups, and Abbott, Ng, Spriano, Gupta and Petyt [3] established uniform exponential growth for many cubical groups. We postpone describing the hierarchy structure until Section 3.1.

The next class we consider is that of coarsely injective spaces. A metric space is said to be *coarsely injective* if there is a constant $\delta$ such that, for any family $\{B(x_i, r_i) : i \in I\}$ of balls with $d(x_i, x_j) \leqslant r_i + r_j$ for all $i, j \in I$, the $\delta$–neighbourhoods of those balls have nonempty total intersection. This property was first considered by Chepoi and Estellon in [30].

As the term suggests, this notion is closely related to that of *injective* metric spaces. A metric space is injective (also called *hyperconvex*) if, for any family $\{B(x_i, r_i) : i \in I\}$ of balls with $d(x_i, x_j) \leqslant r_i + r_j$ for all $i, j \in I$, the balls have nonempty total intersection. In other words, it amounts to taking $\delta = 0$ in the definition of coarse injectivity. (There are multiple equivalent ways to define injectivity of a metric space, by a theorem of Aronszajn and Panitchpakdi [6].) A construction of Isbell [57], which was later rediscovered by Dress [36] and by Chrobak and Larmore [32], shows that every metric space has an essentially unique *injective hull*. More precisely, the injective hull of a metric space $X$ is an injective metric space $E(X)$, together with an isometric embedding $e\colon X \to E(X)$, such that no injective proper subspace of $E(X)$ contains $e(X)$. For convenience, we will often identify $X$ with its image $e(X)$. A nice description of the construction of the injective hull is given by Lang in [60, Section 3].

The classes of coarsely injective spaces and injective spaces are tied together by the following useful fact, the proof of which is identical to that of Chalopin, Chepoi, Genevois, Hirai and Osajda [26, Proposition 3.12]. A subset $Y$ of a metric space $X$ is *coarsely dense* if there exists $r$ such that every $x \in X$ is $r$–close to some $y \in Y$.

**Proposition 1.1** *A metric space is coarsely injective if and only if it is coarsely dense in its injective hull.*

Moreover, if a group acts properly and coboundedly on a coarsely injective space, then it acts properly and coboundedly on the injective hull of that space (see Lemma 3.10). Here and throughout the paper, a group $G$ is said to act *properly* on a metric space $X$ if $\{g \in G : gB \cap B \neq \varnothing\}$ is finite for every metric ball $B$ of $X$. This is sometimes referred to as a *metrically proper* action.

Injective metric spaces satisfy a number of properties reminiscent of nonpositive curvature and, in particular, of CAT(0) spaces. For instance, they admit a conical geodesic bicombing [60], and proper injective spaces of finite combinatorial dimension have a canonical convex such bicombing; see Descombes and Lang [34]. Also, every bounded group action on an injective metric space has a fixed point, and the fixed-point set is itself injective [60]. These properties are what allow us to draw our conclusions for hierarchically hyperbolic groups. Although it will not be needed here, it is interesting to note that injective spaces are also complete [6] and contractible [57].

The second author introduced the strong shortcut property for graphs [52] and then generalised it to roughly geodesic metric spaces [51]. A *Riemannian circle* $S$ is the circle $S^1$ endowed with a geodesic metric of some length $|S|$. A roughly geodesic metric space $(X,\sigma)$ is *strongly shortcut* if there exists $K > 1$ such that, for any $C > 0$, there is a bound on the lengths $|S|$ of $(K,C)$–quasi-isometric embeddings $S \to X$ of Riemannian circles $S$ in $(X,\sigma)$. A group is *strongly shortcut* if it acts properly and coboundedly on a strongly shortcut metric space. Many spaces and graphs of interest in geometric group theory and metric graph theory are strongly shortcut, including Gromov-hyperbolic spaces, 1–skeletons of finite-dimensional CAT(0) cube complexes, Cayley graphs of Coxeter groups, and asymptotically CAT(0) spaces. Despite being such a unifying notion, it remains possible to draw conclusions about strongly shortcut groups, including that they are finitely presented and have polynomial isoperimetric function, and so have decidable word problem.

## 1.2 Comparison of the classes

Our main result is the definition of a new metric on hierarchically hyperbolic spaces and, more generally, on coarse median spaces satisfying a nice approximation property of median intervals by CAT(0) cube complexes.

Our construction is directly inspired by work of Bowditch [20], in which he constructs an injective metric on any finite-rank metric median space. Indeed, if one endows a finite-dimensional CAT(0) cube complex with the piecewise $\ell^\infty$ metric, it becomes an injective metric space. The new metric we construct is weakly roughly geodesic and has the property that balls are coarsely median convex; see Theorem 2.13.

We then prove a hierarchical generalisation of a very nice result of Chepoi, Dragan and Vaxès [29] about pairwise close subsets of hyperbolic spaces. Combining this with work of Russell, Spriano and Tran [71] enables us to deduce a coarse Helly property for balls.

**Theorem A** (Proposition 2.16 and Corollary 3.6) *Let $(X,\mathfrak{S})$ be a hierarchically hyperbolic space with metric $d$. There exists a metric $\sigma$ on $X$ such that $(X,\sigma)$ is coarsely injective and quasi-isometric to $(X,d)$. Moreover, $\sigma$ is invariant with respect to the automorphism group of $(X,\mathfrak{S})$.*

Our second result relates the class of coarsely injective spaces to that of strongly shortcut spaces. A metric space has *uniformly bounded geometry* if, for any $r>0$, there exists a uniform $N(r)\in\mathbb{N}$ such that every ball of radius $r$ contains at most $N(r)$ points.

**Theorem B** (Theorem 4.2) *Every coarsely injective metric space of uniformly bounded geometry is strongly shortcut.*

Huang and Osajda [55] proved that weak Garside groups of finite type and Artin groups of FC type are Helly, so we have the following corollary of Theorem B:

**Corollary C** *Weak Garside groups of finite type and Artin groups of* FC *type are strongly shortcut.*

Combining Theorem A with Theorem B, we deduce the following:

**Corollary D** *Every hierarchically hyperbolic space admits a roughly geodesic metric in its quasi-isometry class that satisfies the strong shortcut property.*

In fact, in the case of hierarchically hyperbolic *groups*, the metric we construct is equivariant, by the "moreover" statement of Theorem A (also see Remark 3.9). Therefore, every hierarchically hyperbolic group acts properly cocompactly on a coarsely injective space, and any group admitting such an action is a strongly shortcut group. Moreover, these three classes can be distinguished. Indeed, the second author showed [50] that the $(3,3,3)$ Coxeter triangle group is strongly shortcut but not coarsely injective, and type-preserving uniform lattices in thick buildings of type $C_n$ are coarsely injective [26], but, by work of the first author [46], they cannot be hierarchically hyperbolic groups, because they do not admit any nonelementary actions on hyperbolic spaces.

One can also ask how *Helly groups*, as defined in [26], fit into this framework. A *Helly graph* is a locally finite graph in which any set of pairwise intersecting balls in the vertex set have nonempty total intersection, and a group is Helly if it acts properly cocompactly on a Helly graph. Helly groups have some strong properties, including biautomaticity [26, Theorem 1.5]. Hughes and Valiunas [56] have constructed a hierarchically hyperbolic group that is not biautomatic and hence not Helly, though it is CAT(0) and acts properly cocompactly on an injective metric space.

It is clear that every Helly group is coarsely injective. Recall that, according to Bridson (see [22]), mapping class groups are not CAT(0). Note that any Helly group acts properly cocompactly on a space with a convex geodesic bicombing (see [34]). So we suspect that mapping class groups are not Helly groups.

We can summarise the relations between these classes with the following diagram, in which $A \Rightarrow B$ denotes the statement "any group that is $A$ is necessarily $B$", and $A \not\Rightarrow B$ denotes "there is an example of a group that is $A$ but not $B$":

$$\begin{tikzcd}
\text{HHG} \arrow[d, Rightarrow, shift left, "/" description] \arrow[dr, Rightarrow, shift left] & & \\
& \text{coarsely injective} \arrow[ul, Rightarrow, shift left, "/" description] \arrow[dl, Rightarrow, shift left, "/" description] \arrow[r, Rightarrow, shift left] & \text{strongly shortcut} \arrow[l, Rightarrow, shift left, "/" description] \\
\text{Helly} \arrow[u, Rightarrow, shift left, "/" description] \arrow[ur, Rightarrow, shift left] & &
\end{tikzcd}$$

## 1.3 Metric consequences

We now describe some of the consequences of Theorem A for hierarchically hyperbolic spaces. Recall that a quasigeodesic bicombing on a metric space $(X,\sigma)$ is a map $\gamma: X\times X\times[0,1]\to X$ such that, for each distinct pair $a,b\in X$, the map $[0,\sigma(a,b)]\to X$ given by $t\mapsto \gamma_{a,b}(t/\sigma(a,b))$ is a quasigeodesic from $a$ to $b$ with uniform constants.

There are various fellow-travelling conditions that a bicombing may enjoy. We say that a bicombing is *roughly conical* if there is a $C\geqslant 0$ such that, for all $a,b,a',b'\in X$ and $t\in[0,1]$,

$$\sigma(\gamma_{a,b}(t),\gamma_{a',b'}(t))\leqslant(1-t)\sigma(a,a')+t\sigma(b,b')+C;$$

and it is *roughly reversible* if it satisfies the following coarse version of symmetry: there is a $C \geq 0$ such that, for all $a, b \in X$ and $t \in [0, 1]$,

$$\sigma(\gamma_{a,b}(t), \gamma_{b,a}(1-t)) \leq C.$$

From the existence of conical, reversible, isometry-invariant geodesic bicombings on injective metric spaces [60], we deduce the following:

**Corollary E** (Corollary 3.7) *Let $(X, \mathfrak{S})$ be a hierarchically hyperbolic space. Then $(X, \sigma)$ admits a roughly conical, roughly reversible, quasigeodesic bicombing that is coarsely equivariant under the automorphism group of $(X, \mathfrak{S})$. More strongly, the combing lines are rough geodesics for the metric $\sigma$.*

In particular, this applies to Teichmüller space with either of the standard metrics, with equivariance under the action of the mapping class group. This particular application was unknown to us until comparing results with Durham, Minsky and Sisto [38].

Corollary E gives a positive answer to Question 8.1 of Engel and Wulff [40], as any roughly conical bicombing is *coherent and expanding*, in their terminology. Engel and Wulff proved that the existence of such a bicombing has a large number of $K$–theoretic consequences. This positive answer also allows one to apply work of Fukaya and Oguni (see [42]) to deduce the coarse Baum–Connes conjecture for hierarchically hyperbolic groups. The coarse Baum–Connes conjecture is also a consequence of finite asymptotic dimension, which is a known property of uniformly proper hierarchically hyperbolic spaces [10].

## 1.4 Consequences for groups

We now turn to the case of hierarchically hyperbolic groups, which, as we have seen, act properly cocompactly on coarsely injective spaces. Here we describe some of the consequences of such an action.

Following Alonso and Bridson [4], we say that a bicombing is *bounded* if it satisfies the following weak two-sided fellow-traveller property: there is a $C \geqslant 0$ such that, for all $a, b, a', b' \in X$ and $t \in [0, 1]$,

$$\sigma(\gamma_{a,b}(t), \gamma_{a',b'}(t)) \leqslant C \max(\sigma(a, a'), \sigma(b, b')) + C.$$

Note that, if a bicombing is roughly conical, then it is bounded. A finitely generated group is said to be *semihyperbolic* if it has a Cayley graph that admits an equivariant bounded quasigeodesic bicombing.

Among other results, Alonso and Bridson proved that semihyperbolicity implies the existence of a quadratic isoperimetric function, that the group has soluble word and conjugacy problems, and that an algebraic flat torus theorem holds [4]. For more discussion of the consequences of semihyperbolicity, see Bridson and Haefliger [24]. Semihyperbolicity was introduced as a response to Gromov's call for a weaker form of hyperbolicity in his original essay on hyperbolic groups, and it fits into the framework of algorithmic properties developed by Epstein, Cannon, Holt, Levy, Paterson and Thurston [41]. For example, semihyperbolicity is implied by biautomaticity, but not by automaticity. A survey can be found in [23].

For hierarchically hyperbolic groups $G$, the freeness of the regular action of $G$ on $(G,\sigma)$ allows the bicombing of Corollary E to be pulled back to the Cayley graph of $G$ [4].

**Corollary F** (Corollary 3.11) *Every hierarchically hyperbolic group is semihyperbolic. In particular, the mapping class group of a surface of finite type is semihyperbolic.*

The mapping class group case also follows from unpublished work of Hamenstädt [49], and is related to Mosher's automaticity theorem [62].

We should emphasise that the same result for mapping class groups has been obtained by rather different methods, simultaneously and independently, by Durham, Minsky and Sisto (see [38]). This will be discussed more in Section 1.6.

It is well known that mapping class groups have finitely many conjugacy classes of finite subgroups (see Bridson [21]), a property that they share with hyperbolic groups. However, to the authors' knowledge, all existing proofs of this fact rely on deep results that do not generalise to other settings, such as Kerckhoff's celebrated solution of the Nielsen realisation problem [59]. It is interesting to ask whether there is a proof that avoids such powerful machinery, and indeed a more general question about hierarchically hyperbolic groups was asked by Hagen and Petyt [47]. The question of whether all hierarchically hyperbolic groups have finitely many conjugacy classes of finite subgroups has resisted a number of attempted resolutions.

The fact that hierarchically hyperbolic groups act properly cocompactly on coarsely injective spaces makes the following a simple consequence of Lang's result about bounded actions on injective spaces [60, Proposition 1.2]:

**Theorem G** (Corollary 3.12) *Hierarchically hyperbolic groups have finitely many conjugacy classes of finite subgroups.*

It is interesting to note that this applies in particular to many quotients of mapping class groups [10; 9]. It is also a simple consequence that residually finite hierarchically hyperbolic groups are virtually torsionfree.

We now summarise the consequences for hierarchically hyperbolic groups of the results described above (also see Remark 3.9 for a comment on their generality).

**Corollary H** *Every hierarchically hyperbolic group $G$ has the following properties*:

- *$G$ acts properly cocompactly on a proper coarsely injective space.*
- *$G$ has finitely many conjugacy classes of finite subgroups.*
- *$G$ is semihyperbolic. In particular*,
  - *the conjugacy problem in $G$ is soluble*, *and it can be solved in doubly exponential time*;
  - *any polycyclic subgroup of $G$ is virtually abelian*;
  - *any finitely generated abelian subgroup of $G$ is quasi-isometrically embedded*;
  - *the centraliser of any finite subset of $G$ is finitely generated*, *quasi-isometrically embedded and semihyperbolic.*
- [40, Theorem C] *For any ring $R$*, *if the cohomological dimension $\mathrm{cd}_R(G)$ is finite*, *then $\mathrm{cd}_R(G) \leqslant \mathrm{asdim}(G)+1$.*
- *$G$ is a strongly shortcut group.*

The result about polycyclic subgroups can also be deduced from the Tits alternative for hierarchically hyperbolic groups established by Durham, Hagen and Sisto [37]. The result about finitely generated abelian subgroups was proved by Plummer [67]. The other consequences are new, however. The result about the conjugacy problem extends work of Abbott and Behrstock [1], showing that it can be solved in exponential time for *Morse elements* of hierarchically hyperbolic groups, and generalises the fact that, in mapping class groups, it can always be solved in exponential time; see Masur and Minsky [61] and Tao [74; 8]. In the case of cubical groups, a beautiful result of Niblo and Reeves [63] states that every cubical group is biautomatic, and semihyperbolicity is a direct consequence of this. We emphasise, though, that the class of hierarchically hyperbolic groups is considerably larger than just cubical groups and mapping class groups.

## 1.5 Bounded packing

The bounded packing property for subgroups of finitely generated groups was introduced as a metric abstraction of tools used to prove intersection properties of subgroups of hyperbolic groups by Gitik, Mitra, Rips and Sageev [45] and Rubinstein and Sageev [70], and in turn as a stepping stone towards ensuring cocompactness of the cube complex associated with a finite collection of quasiconvex codimension-1 subgroups; see Sageev [72], Niblo and Reeves [64] and Hruska and Wise [54]. We recall the definition in Section 3; see Hruska and Wise [53; 54] for more motivation and background. The prototypical example is that of a quasiconvex subgroup of a hyperbolic group. That such subgroups have bounded packing was first established by Gitik, Mitra, Rips and Sageev, using compactness of the boundary [45], and another proof was given by Hruska and Wise [53], using induction on the *height* of the subgroups.

More general examples have been provided by Antolín, Mj, Sisto and Taylor [5], who use induction on height to show that finite collections of *stable subgroups* in any finitely generated group have bounded packing. Stable subgroups were introduced by Durham and Taylor [39], and they are always hyperbolic. More generally, *Morse subgroups* were introduced independently by Tran [75] and Genevois [43], and the notion is implicit in earlier work of Sisto [73]. Notably, Tran proved that any finite collection of Morse subgroups has bounded packing [75, Theorem 1.2], again by using induction on height.

**Theorem I** (Corollary 3.13) *Every finite collection of hierarchically quasiconvex subgroups of a group that is a hierarchically hyperbolic space* (*in particular*, *of any hierarchically hyperbolic group*) *has bounded packing.*

For many groups that are HHSs (including all HHGs), every stable subgroup is hierarchically quasiconvex; see Abbott, Behrstock and Durham [2] and Russell, Spriano and Tran [71]. Theorem B also applies to subsurface stabilisers in the mapping class group, which are neither Morse nor stable. See Section 3.1 for the definition of hierarchical quasiconvexity.

Our proof of this result is purely geometric. It relies on a very strong result for quasiconvex subsets of hyperbolic spaces that was proved by Chepoi, Dragan and Vaxès [29]; we state it as Theorem 3.4. Their theorem does not seem to have garnered the notice it deserves in geometric group theory. For instance, it yields what appears to be the simplest and most natural proof of bounded packing for quasiconvex subgroups of hyperbolic groups. One case of our hierarchical generalisation of their result can be stated as follows:

**Theorem J** (Theorem 3.5) *Let $X$ be a hierarchically hyperbolic space, and let $\mathcal{Q}$ be a finite collection of hierarchically quasiconvex subsets of $X$. If every pair of elements of $\mathcal{Q}$ is $r$–close, then there is a point of $X$ that is $R$–close to every element of $\mathcal{Q}$, where $R$ does not depend on the cardinality of $\mathcal{Q}$.*

## 1.6 Comparison to the work of Durham, Minsky and Sisto [38]

Let us now say a few words about the difference between the present article and the work of Durham, Minsky and Sisto [38]. As noted, both articles independently prove that mapping class groups are semihyperbolic, but the approaches differ greatly. In both cases, this fact is deduced from a stronger statement in a more general setting, but those two statements are very different in flavour. Their results hold for hierarchically hyperbolic spaces with the extra assumption of *colorability*, and they deduce interesting corollaries about bicombings on the Teichmüller space with the Teichmüller metric, and the existence of barycentres. These results are also consequences of Theorem A and Corollary E.

Our construction is built on the fact that intervals in hierarchically hyperbolic spaces can be approximated by finite CAT(0) cube complexes (proved in [12]). The main result of Durham, Minsky and Sisto is that these approximations are furthermore *stable*, meaning that a small change in the endpoints of the interval induces a small change in the approximating CAT(0) cube complex. This stability result may prove extremely useful for other purposes.

If we want to compare the bicombing we obtain to the one from [38] in the simplest case of a CAT(0) cube complex, our bicombing looks like the geodesic CAT(0) bicombing, whereas their bicombing is more similar to (but not the same as) Niblo–Reeves normal cube paths [63]. One notable difference is that our bicombing is roughly conical and their bicombing is merely bounded, which is not enough to deduce the consequences of Section 1.3. On the other hand, their bicombing paths are known to be hierarchy paths, whilst ours are not.

## Structure of the article

In Section 2, we recall basic definitions of coarse median spaces, and we explain the extra property we need, a stronger approximation of median intervals by CAT(0) cube complexes. We then define a new distance, and we prove that it is quasi-isometric to the original one and is weakly roughly geodesic, and that its balls are coarsely median convex.

In Section 3, we treat hierarchical hyperbolicity, and prove that hierarchically quasiconvex subsets satisfy a coarse version of the Helly property. We use this to show that the new distance makes hierarchically hyperbolic spaces coarsely injective, and deduce semihyperbolicity of hierarchically hyperbolic groups. We also show that hierarchically quasiconvex subgroups have bounded packing.

In Section 4, we recall the definition of a strongly shortcut group, and prove that coarse injectivity implies the strong shortcut property.

**Acknowledgements** We thank Mark Hagen and Anthony Genevois for many helpful comments and suggestions, and Victor Chepoi, Alexander Engel and Damian Osajda for interesting discussions. We would like to thank Matthew Durham, Yair Minsky and Alessandro Sisto for friendly discussions about our their work and ours. We thank the referees for their careful and thorough reading of the article.

Haettel was supported by the French grant ANR-16-CE40-0022-01 AGIRA.

Hoda was supported by the ERC grant GroIsRan.

# 2 Coarse median spaces with quasicubical intervals

## 2.1 Background on coarse median spaces

*Coarse median spaces*, defined by Bowditch in [17], are a generalisation of CAT(0) cube complexes and Gromov-hyperbolic spaces, and the class is rich enough to encompass mapping class groups of finite-type surfaces. The general idea is to associate to every triple of points in the space a point that satisfies the axioms of a usual median up to controlled error. This point will be called the coarse median.

Let us recall here that a *median* $\mu\colon X^3 \to X$ on a set $X$ is a map satisfying (where we write equivalently $\mu(x,y,z)$ or $\mu_{x,y,z}$ to increase readability)

- $\mu(x,y,z)$ is symmetric in $x$, $y$ and $z$;
- $\mu(x,x,y) = x$ for all $x, y \in X$; and
- $\mu(a,b,\mu_{x,y,z}) = \mu(\mu_{a,b,x},\mu_{a,b,y},z)$ for all $a,b,x,y,z \in X$.

The pair $(X, \mu)$ is called a *median algebra*. The *rank* of $(X, \mu)$ is the supremum of all $\nu \in \mathbb{N}$ such that there exists an injective median homomorphism from the $\nu$–cube $\{0, 1\}^\nu$ into $X$.

Every finite median algebra can be seen as the 0–skeleton of a CAT(0) cube complex (see [28; 69]).

Let $(X, d)$ be a metric space. For any $x, y \in X$, let

$$I_d(x, y) = \{z \in X \mid d(x, z) + d(z, y) = d(x, y)\}$$

denote the interval between $x$ and $y$. The metric space $(X, d)$ is called *metric median* if $I_d(x, y) \cap I_d(y, z) \cap I_d(x, z)$ is a singleton — say $\{\mu(x, y, z)\}$ — for all $x, y, z \in X$. In this case, $\mu$ defines a median on $X$. Examples of median metric spaces include trees, 1–skeletons of CAT(0) cube complexes with the combinatorial distance, and $L^1$ spaces.

In a Gromov-hyperbolic space $X$, the three intervals joining three points may not intersect precisely in a singleton, but by definition they do coarsely intersect with uniformly bounded diameter. This suggests defining a map $X^3 \to X$ that satisfies the axioms of a median up to bounded error. This is made precise by the following definition, due to Bowditch [17], generalising the *centroid* defined for mapping class groups in [14]:

**Definition 2.1** (coarse median space) Let $(X, d)$ be a metric space. A map $\mu\colon X^3 \to X$ is called a *coarse median* if there exists $h\colon \mathbb{N} \to (0, +\infty)$ such that:

- For all $a, b, c, a', b', c' \in X$, we have
$$d(\mu(a, b, c), \mu(a', b', c')) \leqslant h(0)(d(a, a') + d(b, b') + d(c, c')) + h(0).$$
- For each finite nonempty set $A \subset X$ with $|A| \leqslant n$, there exists a finite median algebra $(Q, \mu_Q)$ and maps $\pi\colon A \to Q$ and $\lambda\colon Q \to X$ such that, for every $\alpha, \beta, \gamma \in Q$, we have $d(\lambda\mu_Q(\alpha, \beta, \gamma), \mu(\lambda\alpha, \lambda\beta, \lambda\gamma)) \leqslant h(n)$, and, for every $a \in A$, we have $d\big(a, \lambda(\pi(a))\big) \leqslant h(n)$.

We say that the triple $(X, \mu, d)$ is a *coarse median space*. If $Q$ can always be chosen to have rank at most $\nu$, we say that $\mu$ has rank at most $\nu$. As with median algebras, we shall write $\mu_{a,b,c} = \mu(a, b, c)$ interchangeably. Note that we are also free to assume that $\mu(a, b, c)$ is symmetric in $a$, $b$ and $c$, and that $\mu(a, a, b) = a$ [17, page 73].

We now recall the definitions of intervals and coarse convexity in coarse median spaces.

**Definition 2.2** (median interval) For a pair of points $a, b \in X$, the median interval between $a$ and $b$ is defined as

$$[a, b] = \{\mu(a, b, x) \mid x \in X\}.$$

**Definition 2.3** (coarse median convexity) For a constant $M \geq 0$, a subset $Y$ of $X$ is said to be $M$–coarsely median convex if

$$d(Y, \mu(x, y, y')) \leq M \quad \text{for all } \; y, y' \in Y, \; x \in X.$$

We finish by introducing some terminology.

**Definition 2.4** (weakly roughly geodesic) Recall that a metric space $(X, d)$ (or, more briefly, the metric $d$) is called *roughly geodesic* if there exists a constant $C_d \geqslant 0$ such that, for any $a, b \in X$, there exists a $(1, C_d)$–quasi-isometric embedding of the interval $f : [0, d(a, b)] \to X$ such that $f(0) = a$ and $f(d(a, b)) = b$.

We say that a metric space $(X, d)$ is called *weakly roughly geodesic* if there exists a constant $C'_d \geqslant 0$ such that, for any $a, b \in X$ and any nonnegative $r \leq d(a, b)$, there is a point $c \in X$ with $|d(a, c) - r| \leq C'_d$ and $d(a, c) + d(c, b) \leq d(a, b) + C'_d$.

**Remark 2.5** Every roughly geodesic metric space is weakly roughly geodesic. Moreover, any metric space $(X, d)$ that is weakly roughly geodesic with constant $C'_d$ is necessarily $(4C'_d, 4C'_d)$–quasigeodesic. Indeed, given $x, y \in X$, one can repeatedly take $r = 3C'_d$ in the definition of weak rough geodesicity to get a sequence $x = w_0$, $w_1, \ldots, w_n = y$ such that $d(w_i, w_{i+1}) \in [2C'_d, 4C'_d]$ and $d(w_i, y) \leq d(w_{i-1}, y) - C'_d$, and the points of this sequence form a quasigeodesic from $x$ to $y$.

## 2.2 Construction of a new metric

Let $(X, \mu, d)$ be a coarse median space. Following Bowditch's construction of an injective metric on a median metric space [20], we shall define a new metric $\sigma$ on $X$.

**Definition 2.6** (contraction) For a constant $K \geqslant 0$, a map $\Phi \colon X \to \mathbb{R}$ is called a $K$–contraction if:

- $\Phi$ is $(1, K)$–coarsely Lipschitz, ie $|\Phi(x) - \Phi(y)| \leqslant d(x, y) + K$ for all $a, b \in X$.
- $\Phi$ is a $K$–quasimedian homomorphism, ie

  $$|\Phi(\mu(a, b, c)) - \mu_{\mathbb{R}}(\Phi(a), \Phi(b), \Phi(c))| \leqslant K \quad \text{for all } \; a, b, c \in X,$$

  where $\mu_{\mathbb{R}}$ denotes the standard median on $\mathbb{R}$.

**Definition 2.7** (new metric) For $K > 0$, we define a new metric $\sigma$ on $X$ as follows. Given $a, b \in X$, let $\sigma(a,b)$ denote the supremum of all $r \geqslant 0$ such that there exists a $K$–contraction $\Phi\colon X \to \mathbb{R}$ such that $\Phi(a) = 0$ and $\Phi(b) = r$.

The assumption that $K$ is nonzero is needed to ensure that $\sigma$ separates points in the setting of coarse median spaces. In the special case where $X$ is a CAT(0) cube complex, we may take $K = 0$. More precisely, if $X$ is a CAT(0) cube complex endowed with the piecewise $\ell^p$ length metric for $p \in \{1, 2, \infty\}$, for instance, then the new metric $\sigma$ for $K = 0$ is the piecewise $\ell^\infty$ length metric on $X$.

**Lemma 2.8** *The function $\sigma$ is a metric on $X$.*

**Proof** Let $a, b \in X$ be distinct. Consider the map $\Phi\colon X \to \{0, K\}$ that sends $b$ to $K$ and everything else to 0. It is a $K$–contraction, and so $\sigma(a,b) \geqslant K > 0$.

The proof of the triangle inequality is identical to [20, Lemma 3.1]. For the reader's convenience, we repeat it here. Let $a, b, c \in X$. For each $r < \sigma(a,b)$, there exists a $K$–contraction $\Phi_r\colon X \to \mathbb{R}$ such that $|\Phi_r(a) - \Phi_r(b)| \geq r$. We certainly have $\sigma(a,c) \geq |\Phi_r(c) - \Phi_r(a)|$ and $\sigma(b,c) \geq |\Phi_r(b) - \Phi_r(c)|$, so

$$\begin{aligned}\sigma(a,c) + \sigma(c,b) &\geqslant \sup\{|\Phi_r(c) - \Phi_r(a)| + |\Phi_r(b) - \Phi_r(c)| : r < \sigma(a,b)\}\\ &\geq \sup\{|\Phi_r(b) - \Phi_r(a)| : r < \sigma(a,b)\} = \sigma(a,b).\end{aligned}$$ □

**Remark 2.9** Although the construction of $\sigma$ depends on the choice of a positive constant $K$, the actual choice of $K$ will not matter to us here. If $K_1 < K_2$, then any $K_1$–contraction is automatically a $K_2$–contraction, so $\sigma_{K_1} \leq \sigma_{K_2}$. On the other hand, if $\Phi$ is a $K_2$–contraction, then $(K_1/K_2)\Phi$ is a $K_1$–contraction, so $\sigma_{K_1} \geq (K_1/K_2)\sigma_{K_2}$. Thus, any two choices of $K$ give bi-Lipschitz metrics.

We record the following simple consequence of the definition of $\sigma$:

**Lemma 2.10** *If a group $G$ is acting isometrically on a coarse median space $(X, \mu, d)$ by median isometries, in the sense that $g\mu(x,y,z) = \mu(gx, gy, gz)$ for all $g \in G$ and $x, y, z \in X$, then the induced action of $G$ on $(X, \mu, \sigma)$ is isometric.*

**Proof** For any $g \in G$ and $x, y \in X$, if $\Phi$ is a $K$–contraction with $\Phi(x) = 0$ and $\Phi(y) = r$, then $\Phi' = \Phi g^{-1}$ is a $K$–contraction with $\Phi'(gx) = 0$ and $\Phi'(gy) = r$. □

In order to help understand the metric $\sigma$, we shall work with coarse median spaces that have the following property, which is a strengthening of the second axiom of coarse median spaces for sets $A = \{a, b\}$ with cardinality 2. We require an approximation of the entire median interval $[a, b]$ with uniform constants, and also that the comparison map be a quasi-isometry and not just coarsely invertible.

**Definition 2.11** (quasicubical intervals) Let $(X, \mu, d)$ be a coarse median space. We say that it has *quasicubical intervals* if it has finite rank $\nu$ and there exists $\kappa \geqslant 1$ such that the following holds: for every $a, b \in X$, there exists a finite CAT(0) cube complex $Q$ of dimension at most $\nu$, endowed with the $\ell^1$ metric $d_Q$ and the median $\mu_Q$, such that there exists a map $\lambda\colon Q \to [a, b]$ satisfying:

- $\lambda$ is a $(\kappa, \kappa)$–quasi-isometry, ie $\lambda$ is $\kappa$–coarsely onto and
$$\frac{1}{\kappa} d_Q(\alpha, \beta) - \kappa \leqslant d(\lambda(\alpha), \lambda(\beta)) \leqslant \kappa d_Q(\alpha, \beta) + \kappa \quad \text{for all } \alpha, \beta \in Q.$$
- $\lambda$ is a $\kappa$–quasimedian homomorphism, ie
$$d\big(\lambda(\mu_Q(\alpha, \beta, \gamma)), \mu(\lambda(\alpha), \lambda(\beta), \lambda(\gamma))\big) \leqslant \kappa \quad \text{for all } \alpha, \beta, \gamma \in Q.$$

Obviously this is satisfied by finite-dimensional CAT(0) cube complexes, or indeed by any space with a global quasimedian quasi-isometry to a CAT(0) cube complex.

**Proposition 2.12** *Hierarchically hyperbolic spaces have quasicubical intervals, as do coarse median spaces satisfying the axioms* (B1)–(B10) *in* [18].

**Proof** In hierarchically hyperbolic spaces, the notion of median intervals used here coincides coarsely with the hierarchically quasiconvex hull of a pair of points defined in [13], by [71, Corollary 5.12; 19, Lemma 8.1]. The first statement is thus a special case of [12, Theorem 2.1]. The second statement is exactly [18, Theorem 1.3]. □

As noted by Bowditch, every hierarchically hyperbolic space satisfies the axioms (B1)–(B10) in [18]. It is not known whether all cocompact cube complexes can be given a structure that satisfies these axioms.

We can now state the main result of this section. It sums up Lemma 2.10, Propositions 2.16 and 2.21, and Lemma 2.23, and the proof is split over the next three subsections.

**Theorem 2.13** *Assume that the coarse median space $(X, \mu, d)$ has quasicubical intervals and is roughly geodesic. The metrics $\sigma$ and $d$ are quasi-isometric, $\sigma$ is weakly roughly geodesic, and balls for $\sigma$ are uniformly coarsely median convex. Moreover, $\sigma$ is invariant under the group of median isometries of $(X, \mu, d)$.*

## 2.3 The metrics $d$ and $\sigma$ are quasi-isometric

Here we shall prove that the new distance $\sigma$ is quasi-isometric to the original distance $d$. We need the following technical result for coarse median spaces, which is a special case of Lemmas 2.18 and 2.19 of [65]:

**Lemma 2.14** *In any coarse median space $(X, d, \mu)$, there exists a constant $H_5 \geqslant 0$ such that, for any $a, b, x, y, z \in X$,*

$$d(\mu(a, b, \mu_{x,y,z}), \mu(\mu_{a,b,x}, \mu_{a,b,y}, z)) \leqslant H_5,$$
$$d(\mu(a, b, \mu_{x,y,z}), \mu(\mu_{a,b,x}, \mu_{a,b,y}, \mu_{a,b,z})) \leqslant H_5.$$

We will now prove that, up to multiplicative and additive constants, one can restrict to contractions defined on the interval between two points for the definition of $\sigma$.

**Lemma 2.15** *For each $a, b \in X$, let $\sigma'(a, b)$ denote the supremum of all $r \geqslant 0$ such that there exists a $K$–contraction $\Phi' \colon [a, b] \to \mathbb{R}$ for which $\Phi'(a) = 0$ and $\Phi'(b) = r$. There exists $L \geqslant 1$ such that, for each $a, b \in X$, we have $\sigma(a, b) \leqslant \sigma'(a, b) \leqslant L\sigma(a, b)$.*

**Proof** It is immediate that $\sigma(a, b) \leqslant \sigma'(a, b)$. Consider $r \geqslant 0$ and a $K$–contraction $\Phi' \colon [a, b] \to \mathbb{R}$ such that $\Phi'(a) = 0$ and $\Phi'(b) = r$. Define $\Phi \colon X \to \mathbb{R}$ by $c \mapsto \Phi'(\mu(a, b, c))$. Since the map $c \mapsto \mu(a, b, c)$ is $(h(0), h(0))$–coarsely Lipschitz and $\Phi'$ is $(1, K)$–coarsely Lipschitz, we deduce that $\Phi$ is $(h(0), h(0)+K)$–coarsely Lipschitz.

Now let $x, y, z \in X$. According to Lemma 2.14, we have

$$d(\mu(a, b, \mu_{x,y,z}), \mu(\mu_{a,b,x}, \mu_{a,b,y}, \mu_{a,b,z})) \leqslant H_5.$$

Hence, since $\Phi'$ is $(1, K)$–coarsely Lipschitz,

$$|\Phi'(\mu(a, b, \mu_{x,y,z})) - \Phi'(\mu(\mu_{a,b,x}, \mu_{a,b,y}, \mu_{a,b,z}))| \leqslant H_5 + K.$$

But $\Phi'$ is also a $K$–quasimedian homomorphism, and so

$$|\Phi'(\mu(\mu_{a,b,x}, \mu_{a,b,y}, \mu_{a,b,z})) - \mu_{\mathbb{R}}(\Phi'(\mu_{a,b,x}), \Phi'(\mu_{a,b,y}), \Phi'(\mu_{a,b,z}))| \leqslant K.$$

Combining these and recalling the definition of $\Phi$ enables us to conclude that

$$|\Phi(\mu_{x,y,z}) - \mu_{\mathbb{R}}(\Phi(x), \Phi(y), \Phi(z))| \leqslant H_5 + 2K.$$

Thus, if we set $L = \max\{h(0), 1+h(0)/K, 2+H_5/K\}$, then $(1/L)\Phi$ is a $K$–contraction, and so $\sigma'(a, b) \leqslant L\sigma(a, b)$. □

We can now deduce that $\sigma$ is quasi-isometric to $d$ in the setting of Theorem 2.13.

**Proposition 2.16** *If $(X, \mu, d)$ has quasicubical intervals, then $d$ and $\sigma$ are quasi-isometric.*

**Proof** Fix $a, b \in X$. First of all, since any $K$–contraction is $(1, K)$–coarsely Lipschitz, $\sigma(a,b) \leqslant d(a,b) + K$.

According to the quasicubicality of intervals, there exists a finite CAT(0) cube complex $Q$ of dimension at most $\nu$ and a map $\lambda\colon (Q, d_Q) \to [a,b]$ that is a $(\kappa,\kappa)$–quasi-isometry and a $\kappa$–quasimedian homomorphism. Then $\lambda$ has a quasi-inverse $\pi\colon [a,b] \to (Q, d_Q)$ that is a $(\kappa',\kappa')$–quasi-isometry and a $\kappa'$–quasimedian homomorphism, where $\kappa'$ is a constant depending only on $\kappa$ and $h(0)$.

Note that we shall in fact use $Q$ to denote the vertex set, $d_Q$ to denote the combinatorial (piecewise $\ell^1$) distance on $Q$, and $\mu_Q$ to denote the median on $Q$. Let us denote by $\sigma_Q$ the piecewise $\ell^\infty$ distance on $Q$; we have $\sigma_Q \leqslant d_Q \leqslant \nu\sigma_Q$.

Since $Q$ is a CAT(0) cube complex, there exists a 0–contraction $\Phi_Q\colon (Q, d_Q) \to \mathbb{Z}$ such that $\Phi_Q(\pi(a)) = 0$ and $\Phi_Q(\pi(b)) = \sigma_Q(\pi(a), \pi(b))$ (see [20, Section 7; 7, Corollary 2.5]). Let us consider $\Phi' = (\min\{1, K\}/\kappa')\Phi_Q\pi\colon [a,b] \to \mathbb{R}$. Since $\pi$ is a $(\kappa',\kappa')$–quasi-isometry and $\Phi_Q$ is 1–Lipschitz, we deduce that $\Phi'$ is $(1, K)$–coarsely Lipschitz. Furthermore, for every $x, y, z \in [a,b]$,

$$\begin{aligned}
&|\Phi_Q\pi(\mu_{x,y,z}) - \mu_{\mathbb{R}}(\Phi_Q\pi(x), \Phi_Q\pi(y), \Phi_Q\pi(z))| \\
&\quad \leqslant \bigl|\Phi_Q\pi(\mu_{x,y,z}) - \Phi_Q\bigl(\mu_Q(\pi(x), \pi(y), \pi(z))\bigr)\bigr| \\
&\qquad\qquad\qquad + \bigl|\Phi_Q\bigl(\mu_Q(\pi(x), \pi(y), \pi(z))\bigr) - \mu_{\mathbb{R}}(\Phi_Q\pi(x), \Phi_Q\pi(y), \Phi_Q\pi(z))\bigr| \\
&\quad \leqslant d_Q\bigl(\pi(\mu_{x,y,z}), \mu_Q(\pi(x), \pi(y), \pi(z))\bigr) \\
&\quad \leq \kappa',
\end{aligned}$$

so $\Phi'$ is $K$–quasimedian.

The map $\Phi'$ is therefore a $K$–contraction on $[a,b]$, and $\Phi'(a) = 0$ and $\Phi'(b) = (\min\{1, K\}/\kappa')\sigma_Q(\pi(a), \pi(b)) \geqslant (\min\{1, K\}/\nu\kappa')d_Q(\pi(a), \pi(b))$. Using Lemma 2.15, we deduce that $d_Q(\pi(a), \pi(b)) \leq (\nu\kappa' L/\min\{1, K\})\sigma(a,b)$. But $\pi$ is a $(\kappa',\kappa')$–quasi-isometry, so we also have $d_Q(\pi(a), \pi(b)) \geq (1/\kappa')d(a,b) - \kappa'$.

In conclusion,

$$\frac{\min\{1, K\}}{\nu\kappa'^2 L} d(a,b) - \frac{\min\{1, K\}}{\nu L} \leq \sigma(a,b) \leq d(a,b) + K$$

for all $a, b \in X$. □

## 2.4 The metric $\sigma$ is weakly roughly geodesic

Recall that $(X, \mu, d)$ is a coarse median space with corresponding function $h$, that $X$ has quasicubical intervals (though this will only be used for Proposition 2.21 in this section) and that the metric $d$ is $C_d$–roughly geodesic. We shall prove that the new metric $\sigma$ is weakly roughly geodesic (see Definition 2.4). This will be the most difficult part of the proof of Theorem 2.13.

Let $a, b \in X$, let $E$ be a small positive constant and consider $K$–contractions $\Phi_1\colon X \to [0, r]$ and $\Phi_2\colon X \to [r, r+s]$ (for some $r, s \geqslant E$) such that $\Phi_1(a) \leq E$ and $\Phi_2(b) \geq r + s - E$. We want to find a criterion to ensure that we can combine $\Phi_1$ and $\Phi_2$ into a contraction $\Phi$ such that $\Phi(a) = 0$ and $(r+s) - \Phi(b)$ is bounded above by some constant.

**Lemma 2.17** *Assume that $a$, $b$, $\Phi_1$, $\Phi_2$, $r$, $s$ and $E$ are as above. Let $D = h(0)(3K + 4C_d) + 4K + h(0)$. If $t \in [0, \min\{r, s\} - D + K - E]$ is such that the sets*

$$Z_1 = \{z \in X \mid \Phi_1(z) \leqslant r - t - K\} \quad \textit{and} \quad Z_2 = \{z \in X \mid \Phi_2(z) \geqslant r + t + K\}$$

*are disjoint, then $\sigma(a, b) \geqslant r + s - 2t - 2D - 2E$.*

**Proof** For $m \in \{0, 1, 2\}$, let us write $Y_1^m = \{x \in X \mid \Phi_1(x) \leqslant r - t - D + mK\}$ and $Y_2^m = \{x \in X \mid \Phi_2(x) \geqslant r + t + D - mK\}$. Note that, if $m_1 < m_2$, then $Y_i^{m_1} \subset Y_i^{m_2}$.

**Claim 1** $$d(Y_1^2, Y_2^2) \geqslant D - 4K.$$

**Proof** Let $x_1 \in Y_1^2$ and $x_2 \in Y_2^2$. Since $Y_2^2 \subset Z_2$, we have $x_2 \notin Z_1$, so $\Phi_1(x_2) > r - t - K$. We also have $\Phi_1(x_1) \leqslant r - t - D + 2K$, so $|\Phi_1(x_1) - \Phi_1(x_2)| \geqslant D - 3K$. As $\Phi_1$ is $(1, K)$–coarsely Lipschitz, we have $|\Phi_1(x_1) - \Phi_1(x_2)| \leqslant d(x_1, x_2) + K$, and hence $d(x_1, x_2) \geqslant D - 4K$. ◁

**Claim 2** $$d(Y_1^1, Y_2^1) \geqslant 3K + 4C_d.$$

**Proof** Let $x_1 \in Y_1^1$ and $x_2 \in Y_2^1$, and set $y_1 = \mu(a, b, x_1) \in [a, b]$ and $y_2 = \mu(a, b, x_2) \in [a, b]$. We know that $\Phi_1(y_1) \leqslant \mu_{\mathbb{R}}(\Phi_1(a), \Phi_1(b), \Phi_1(x_1)) + K$. We also have $\Phi_1(a) \leq E$ by assumption, and $\Phi_1(x_1) \leqslant r - t - D + K$. As this latter quantity is at least $E$, $\mu_{\mathbb{R}}(\Phi_1(a), \Phi_1(b), \Phi_1(x_1)) \leqslant r - t - D + K$. Hence, $\Phi_1(y_1) \leqslant r - t - D + 2K$, so $y_1 \in Y_1^2$. A similar argument shows that $y_2 \in Y_2^2$.

According to Claim 1, $d(y_1, y_2) \geqslant D - 4K$. Since $\mu$ is $(h(0), h(0))$–coarsely Lipschitz with respect to each variable, $d(y_1, y_2) \leqslant h(0)d(x_1, x_2) + h(0)$, so

$$d(x_1, x_2) \geqslant \frac{d(y_1, y_2) - h(0)}{h(0)} \geqslant \frac{D - 4K - h(0)}{h(0)} = 3K + 4C_d,$$

as desired. $\triangleleft$

**Claim 3** *The set $\{\mu_{x,y,z} \mid x, y \in Y_1^0, z \in X\}$ is disjoint from $Y_2^0$, and the set $\{\mu_{x,y,z} \mid x, y \in Y_2^0, z \in X\}$ is disjoint from $Y_1^0$.*

**Proof** Fix $x, y \in Y_1^0$ and $z \in X$. Since $\Phi_1(x), \Phi_1(y) \leqslant r - t - D$, we deduce that $\mu_{\mathbb{R}}(\Phi_1(x), \Phi_1(y), \Phi_1(z)) \leqslant r - t - D$, and it follows that $\Phi_1(\mu_{x,y,z}) \leqslant r - t - D + K$, so $\mu_{x,y,z} \in Y_1^1$. Because we showed in Claim 2 that $d(Y_1^1, Y_2^1) \geqslant 3K + 4C_d > 0$, we know that $\mu_{x,y,z} \notin Y_2^1$, and, in particular, $\mu_{x,y,z} \notin Y_2^0$. The other case is similar. $\triangleleft$

Write $Y = X \setminus (Y_1^0 \cup Y_2^0)$, and consider $\Phi\colon X \to [0, r + s - 2t - 2D]$ defined by

$$\Phi(x) = \begin{cases} \Phi_1(x) & \text{if } x \in Y_1^0, \\ \Phi_2(x) - 2t - 2D & \text{if } x \in Y_2^0, \\ r - t - D & \text{if } x \in Y. \end{cases}$$

We have $\Phi(a) \leq E$ and $\Phi(b) \geq r + s - 2t - 2D - E$, so, if we prove that $\Phi$ is a $K$–contraction, then we may deduce that $\sigma(a, b) \geqslant r + s - 2t - 2D - 2E$, the desired conclusion.

**Claim 4** *$\Phi$ is $(1, K)$–coarsely Lipschitz.*

**Proof** Notice that $\Phi$ coincides on $Y_1^0 \cup Y$ with the composition of $\Phi_1\colon X \to [0, r]$ with the 1–Lipschitz map $m_t = \min(\cdot, r - t - D)\colon [0, r] \to [0, r - t - D]$. Hence, if $x, y \in Y_1^0 \cup Y$, then $|\Phi(x) - \Phi(y)| \leqslant |\Phi_1(x) - \Phi_1(y)| \leqslant d(x, y) + K$. A similar argument involving a maximum function applies if $x, y \in Y_2^0 \cup Y$.

Now suppose that $x \in Y_1^0$ and $y \in Y_2^0$. Since $d$ is $C_d$–roughly geodesic, there is a $(1, C_d)$–quasi-isometric embedding $f\colon [0, d(x, y)] \to X$ with $f(0) = x$ and $f(d(x, y)) = y$. For any $\varepsilon > 0$, there exists $\tau$ such that $f(\tau) \in Y_1^0$ but $f(\tau + \delta) \notin Y_1^0$ for any $\delta > \varepsilon$. (Were $f$ continuous, we could take $\varepsilon = 0$ and use the maximal $\tau$ with $f(\tau) \in Y_1^0$.) Write $z_1 = f(\tau)$. We have $d(x, z_1) + d(z_1, y) \leq d(x, y) + C_d$ and $\Phi_1(z_1) \leq r - t - D$. Moreover, for any $\delta > \varepsilon$,

$$\Phi_1(z_1) \geq \Phi_1(f(\tau + \delta)) - \big(d(f(\tau), f(\tau + \delta)) + K\big) > r - t - D - (\delta + C_d + K),$$

and so $\Phi_1(z_1) \geq r - t - D - C_d - K - \varepsilon$. We can now similarly construct $z_2 \in Y_2^0$ such that $d(z_1, z_2) + d(z_2, y) \leq d(z_1, y) + C_d$ and $r + t + D \leq \Phi_2(z_2) \leq r + t + D + C_d + K + \varepsilon$.

With these, we can compute

$$\begin{aligned}
&|\Phi(x)-\Phi(y)| \\
&\quad \le |\Phi(x)-\Phi(z_1)|+|\Phi(z_1)-\Phi(z_2)|+|\Phi(z_2)-\Phi(y)| \\
&\quad = |\Phi_1(x)-\Phi_1(z_1)|+|\Phi_1(z_1)-(\Phi_2(z_2)-2t-2D)|+|\Phi_2(z_2)-\Phi_2(y)| \\
&\quad \leqslant (d(x,z_1)+K)+(2C_d+2K+2\varepsilon)+(d(z_2,y)+K) \\
&\quad \leqslant (d(x,y)+C_d-d(z_1,y)+K)+(2C_d+2K+2\varepsilon)+(d(z_1,y)+C_d-d(z_1,z_2)+K) \\
&\quad = d(x,y)-d(z_1,z_2)+4K+4C_d+2\varepsilon \\
&\quad \leqslant d(x,y)+K+2\varepsilon,
\end{aligned}$$

where the last line comes from Claim 2: $d(z_1,z_2) \geqslant d(Y_1^0,Y_2^0) \geqslant d(Y_1^1,Y_2^1) \geq 3K+4C_d$. This is sufficient, because $\varepsilon$ can be taken to be arbitrarily close to 0. $\lhd$

**Claim 5** *$\Phi$ is $K$–quasimedian.*

**Proof** As noted in the proof of Claim 4, on $Y_1^0 \cup Y$ we have $\Phi = m_t\Phi_1$. As $m_t$ is a median homomorphism with respect to $\mu_{\mathbb{R}}$, if $x,y,z \in Y_1^0 \cup Y$, then

$$\begin{aligned}
|\Phi(\mu_{x,y,z})-\mu_{\mathbb{R}}(\Phi(x),\Phi(y),\Phi(z))| &= |m_t\Phi_1(\mu_{x,y,z})-\mu_{\mathbb{R}}(m_t\Phi_1(x),m_t\Phi_1(y),m_t\Phi_1(z))| \\
&\leqslant |\Phi_1(\mu_{x,y,z})-\mu_{\mathbb{R}}(\Phi_1(x),\Phi_1(y),\Phi_1(z))| \\
&\le K,
\end{aligned}$$

and similarly if $x,y,z \in Y_2^0 \cup Y$.

Assume now that $x,y \in Y_1^0$ and $z \in Y_2^0$. We have that both $\Phi(x)$ and $\Phi(y)$ are at most $r-t-D$. Moreover, $\Phi(z) = \Phi_2(z)-2t-2D \geqslant r-t-D$, and the fact that $z \notin Y_1^0$ implies that $\Phi_1(z) > r-t-D$. Thus, $\mu_{\mathbb{R}}(\Phi(x),\Phi(y),\Phi(z)) = \mu_{\mathbb{R}}(\Phi_1(x),\Phi_1(y),\Phi_1(z)) \le r-t-D$. As $\Phi_1$ is $K$–quasimedian, we deduce that $|\mu_{\mathbb{R}}(\Phi(x),\Phi(y),\Phi(z))-\Phi_1(\mu_{x,y,z})| \leqslant K$. By Claim 3, we know that $\mu_{x,y,z} \notin Y_2^0$, and so $\Phi(\mu_{x,y,z}) = m_t\Phi_1(\mu_{x,y,z})$. But $\mu_{\mathbb{R}}(\Phi(x),\Phi(y),\Phi(z)) \leqslant r-t-D$, so we conclude that $|\mu_{\mathbb{R}}(\Phi(x),\Phi(y),\Phi(z))-\Phi(\mu_{x,y,z})| \leqslant K$. A similar argument applies when $x,y \in Y_2^0$ and $z \in Y_1^0$.

Assume finally that $x \in Y_1^0$, $y \in Y$, and $z \in Y_2^0$. Since $\Phi(x) = \Phi_1(x) \leqslant r-t-D$, $\Phi(y) = r-t-D$ and $\Phi(z) = \Phi_2(z)-2t-2D \geqslant r-t-D$, we have

$$\mu_{\mathbb{R}}(\Phi(x),\Phi(y),\Phi(z)) = r-t-D.$$

If $\mu_{x,y,z} \in Y$, then $\Phi(\mu_{x,y,z}) = r - t - D = \mu_{\mathbb{R}}(\Phi(x), \Phi(y), \Phi(z))$. If $\mu_{x,y,z} \in Y_1^{\mathbf{0}}$, then $\Phi(\mu_{x,y,z}) = \Phi_1(\mu_{x,y,z}) \geqslant \mu_{\mathbb{R}}(\Phi_1(x), \Phi_1(y), \Phi_1(z)) - K \geqslant r - t - D - K$, from which it follows that $|\mu_{\mathbb{R}}(\Phi(x), \Phi(y), \Phi(z)) - \Phi(\mu_{x,y,z})| \leqslant K$. A similar argument applies if $\mu_{x,y,z} \in Y_2^{\mathbf{0}}$.

We have shown that $|\mu_{\mathbb{R}}(\Phi(x), \Phi(y), \Phi(z)) - \Phi(\mu_{x,y,z})| \leqslant K$ in all cases. ◁

We have proved that $\Phi$ is a $K$–contraction. As stated above, this shows that $\sigma(a,b) \geqslant |\Phi(a) - \Phi(b)| \geq r + s - 2t - 2D - 2E$. □

Recall that the *convex hull* $\mathrm{Hull}(A)$ of a subset $A$ of a CAT(0) cube complex $Q$ is the smallest convex subcomplex of $Q$ containing $A$. Equivalently, it is the smallest subcomplex that contains $A$ and which is *median convex*, in the sense that $\mu(q,a,b) \in \mathrm{Hull}(A)$ whenever $a, b \in \mathrm{Hull}(A)$. We regard a subset of $Q^{(0)}$ as convex if the full subcomplex spanned by it is convex. We need the following iterative description convex hulls in CAT(0) cube complexes:

**Lemma 2.18** *Let $Q$ be a CAT$(0)$ cube complex of dimension at most $\nu$, and let $\mu_Q\colon {Q^{(0)}}^3 \to Q^{(0)}$ denote the median. Given $A \subset Q^{(0)}$, set $A_0 = A$, and, for each $i \in \mathbb{N}$, let*

$$A_{i+1} = \mu_Q(Q^{(0)}, A_i, A_i) = \{\mu_Q(x,a,b) \mid a \in A_i,\, b \in A_i,\, x \in Q^{(0)}\}.$$

*Then $A_{\nu'} = \mathrm{Hull}(A)$, where $\nu' = \max\{1, \nu - 1\}$.*

Note that the constant $\nu'$ is probably far from optimal. However, $\nu'$ does depend on $A$ and $Q$. For example, if $A$ is the star of a vertex in a $\nu$–cube, then it can be seen that the optimal value of $\nu'$ is $\lceil \log_2 \nu \rceil$ in this case.

**Proof** The result is trivial if $A$ is convex. Otherwise, fix $x \in \mathrm{Hull}(A) \smallsetminus A$, and let $\mathcal{H}$ be the collection of hyperplanes of $\mathrm{Hull}(A)$ that are adjacent to $x$. For each $H \in \mathcal{H}$, let $Q^{(0)} = H^+ \sqcup H^-$ denote the partition defined by $H$, where $x \in H^+$. Let $\{H_1, \dots, H_n\}$ be a maximal pairwise crossing family in $\mathcal{H}$. We have $n \leq \nu$. For each $i$, let $\mathcal{H}_i$ denote the set of elements of $\mathcal{H}$ that are disjoint from $H_i$, together with $H_i$. An important observation is that $H_i^- \subset H^+$ whenever $H \in \mathcal{H}_i \smallsetminus \{H_i\}$.

If $n = 1$, then $x$ is a cut-point or leaf of $\mathrm{Hull}(A)$, so taking any $a \in A \cap H^+$ and $b \in A \cap H^-$ gives $x = \mu(x,a,b)$, and we are done.

So suppose that $n \geq 2$. If for every $a \in A \cap H_1^-$ we have $a \in H_2^-$, then for every $b \in A \cap H_2^+$ we have $b \in H_1^+$, so, if we take $z_1 \in A \cap H_1^-$ and $z_2 \in A \cap H_2^+$, then $\mu(x, z_1, z_2) \in H^+ \cap H'^+$ for every $H \in \mathcal{H}_1$ and $H' \in \mathcal{H}_2$. We can reason similarly if every element of $A \cap H_2^-$ lies in $H_1^-$. Otherwise there exist $z_1 \in A \cap H_1^- \cap H_2^+$ and $z_2 \in A \cap H_1^+ \cap H_2^-$, and we again have $\mu(x, z_1, z_2) \in H^+ \cap H'^+$ for every $H \in \mathcal{H}_1$ and $H' \in \mathcal{H}_2$. Let $y_1 = \mu(x, z_1, z_2) \in A_1$.

We proceed inductively. Suppose that we have $y_i \in A_i$ such that $y_i \in H^+$ for all $H \in \bigcup_{j \leq i+1} \mathcal{H}_j$. Let $z_{i+2}$ be any point of $A$ that is separated from $y_i$ by $H_{i+2}$, and set $y_{i+1} = \mu(x, y_i, z_{i+2})$. Since $x, y_i \in H^+$ for every $H \in \bigcup_{j \leq i+1} \mathcal{H}_j$, the same is true of $y_{i+1}$, and, since $y_i$ and $z_{i+2}$ lie on opposite sides of $H_{i+2}$, we also have that $y_{i+1} \in H^+$ for all $H \in \mathcal{H}_{i+2}$.

By this procedure, we obtain $y_{n-1} \in A_{n-1} \cap \mathrm{Hull}(A)$ that is not separated from $x$ by any hyperplane of $\mathrm{Hull}(A)$, so we must have $y_{n-1} = x$. □

In order to apply Lemma 2.17, we focus on contractions on CAT(0) cube complexes. Recall that a *chain* of hyperplanes is an ordered sequence $(H_1, \dots, H_n)$ of pairwise disjoint hyperplanes such that $H_j$ separates $H_i$ from $H_k$ whenever $i < j < k$.

**Lemma 2.19** *Let $Q$ be a CAT(0) cube complex of dimension at most $\nu$, and let $\Phi\colon Q^{(0)} \to \mathbb{R}$ be a $K'$–quasimedian, $(K', K')$–coarsely Lipschitz map (for the $\ell^1$ metric) with bounded image. There exists an interval $[u, v]$ of $\mathbb{Z}$ and a chain $(H_n)_{u \leqslant n \leqslant v}$ of hyperplanes of $Q$ satisfying the following:*

- *For each vertex $x$ in $Q$, there exists a unique $n = \Psi(x) \in [u-1, v]$ such that*
  - *either $u \leqslant n \leqslant v-1$ and $x$ is between $H_n$ and $H_{n+1}$,*
  - *or $n = u-1$ and $H_u$ separates $x$ from $H_{u+1}$,*
  - *or $n = v$ and $H_v$ separates $x$ from $H_{v-1}$.*
- *For each vertex $x$ in $Q$, we have $|\Phi(x) - 4K'\nu\Psi(x)| \leqslant 4K'\nu$.*

**Proof** Fix $n \in \mathbb{Z}$, and consider $K_n = \Phi^{-1}\big((2An - A, 2An]\big)$, where $A = 2K'\nu$. Since $Q^{(0)}$ is 1–connected, $\Phi(Q^{(0)})$ is $2K'$–connected. In particular, the set of integers $n \in \mathbb{Z}$ such that $K_n \neq \varnothing$ is an interval $[u-1, v]$. Furthermore, for each $u \leqslant n \leqslant v-1$, we know that $K_n$ disconnects $Q$.

In the notation of Lemma 2.18, for all $i \geq 0$, if $x \in (K_n)_i$, then $|\Phi(x) - \Phi(K_n)| \leqslant K'i$. Indeed, this is clear for $i = 0$ and, if $x \in (K_n)_{i+1}$, so that there exist $a, b \in (K_n)_i$ with

$x = \mu_Q(x, a, b)$, then the fact that $\Phi$ is $K'$–quasimedian implies that

$$|\Phi(x) - \mu_{\mathbb{R}}(\Phi(x), \Phi(a), \Phi(b))| \leq K',$$

yielding the claimed inequality by induction. In particular, Lemma 2.18 tells us that every $x \in \mathrm{Hull}(K_n)$ satisfies $|\Phi(x) - \Phi(K_n)| \leq K'\nu$.

As a consequence, if $n \neq m$, then the convex subcomplexes $\mathrm{Hull}(K_n)$ and $\mathrm{Hull}(K_m)$ are disjoint. Thus, for each $u \leqslant n \leqslant v$, there exists a hyperplane $H_n$ of $Q$ that separates $\mathrm{Hull}(K_{n-1})$ from $\mathrm{Hull}(K_n)$ [27, Corollary 1].

For each vertex $x \in Q^{(0)}$, let $u - 1 \leqslant n \leqslant v$ be such that $\Phi(x) \in (2An - 2A, 2An]$. Then $\Psi(x)$ is equal either to $n-1$ or to $n$. So $|\Phi(x) - 2A\Psi(x)| \leqslant 2A = 4K'\nu$. □

Before stating the next lemma, we remark that, given any chain $\mathcal{H}$ of hyperplanes in a finite CAT(0) cube complex $Q$, there is an associated map $Q^{(0)} \to \mathbb{Z}$: the cube complex dual to $\mathcal{H}$ is a finite interval of $\mathbb{Z}$, and each vertex of $Q$ determines a consistent orientation of the hyperplanes in $\mathcal{H}$. This is a special case of the *restriction quotient* described in [25], and it is clearly a median map. Conversely, any 0–contraction on $Q$ can be realised as restriction quotient in this manner. Moreover, after a translation of $\mathbb{Z}$, we may assume that the codomain is contained in $\mathbb{N}$ if it is bounded.

**Lemma 2.20** *Let $Q$ be a finite CAT(0) cube complex of dimension at most $\nu$. Let $\mathcal{C}$ be a (necessarily finite up to translations of $\mathbb{Z}$) family of 0–contractions on $Q$, ie each $\Psi \in \mathcal{C}$ is a map $Q^{(0)} \to \mathbb{N}$ given by a chain $(H_{\Psi,1}, \dots, H_{\Psi,n_\Psi})$ of hyperplanes of $Q$. Let $\sigma_{\mathcal{C}}$ denote the pseudometric on $Q^{(0)}$ defined by*

$$\sigma_{\mathcal{C}}(\alpha, \beta) = \max_{\Psi \in \mathcal{C}} |\Psi(\alpha) - \Psi(\beta)| \quad \textit{for all } \alpha, \beta \in Q^{(0)}.$$

*Then, for each $\alpha, \beta \in Q^{(0)}$ and for each integer $0 \leqslant r \leqslant \sigma_{\mathcal{C}}(\alpha, \beta)$, there is a vertex $\gamma \in [\alpha, \beta]$ and contractions $\Psi_1, \Psi_2 \in \mathcal{C}$ such that the following hold:*

1. *$\sigma_{\mathcal{C}}(\alpha, \gamma) = r$.*
2. *$\sigma_{\mathcal{C}}(\alpha, \gamma) = |\Psi_1(\alpha) - \Psi_1(\gamma)|$ and $\sigma_{\mathcal{C}}(\gamma, \beta) = |\Psi_2(\gamma) - \Psi_2(\beta)|$.*
3. *If $(H_{1,1}, \dots, H_{1,n_1})$ is the maximal subchain of hyperplanes defining $\Psi_1$ that separate $\alpha$ from $\gamma$ and $(H_{2,1}, \dots, H_{2,n_2})$ is the maximal subchain of hyperplanes defining $\Psi_2$ that separate $\gamma$ from $\beta$, then $(H_{1,1}, \dots, H_{1,n_1}, H_{2,1}, \dots, H_{2,n_2})$ is a chain.*

**Proof** Fix $\alpha, \beta \in Q^{(0)}$ and an integer $0 < r < \sigma_{\mathcal{C}}(\alpha, \beta)$. Since $\sigma_{\mathcal{C}}$ is 1–Lipschitz with respect to the combinatorial distance on $Q^{(0)}$, we know that there exists $\gamma \in [\alpha, \beta]$ such that $\sigma_{\mathcal{C}}(\alpha, \gamma) = r$. Among all possible choices, choose such $\gamma$ as far away from $\alpha$ as possible, in the sense that, for $\gamma' \in [\alpha, \beta]$,

$$\sigma_{\mathcal{C}}(\alpha, \gamma') = r \quad \text{and} \quad \gamma \in [\alpha, \gamma'] \quad \Longrightarrow \quad \gamma' = \gamma.$$

Let $\Psi_2 \in \mathcal{C}$ be such that $\sigma_{\mathcal{C}}(\gamma, \beta) = |\Psi_2(\gamma) - \Psi_2(\beta)|$. Let $(H_{2,1}, \dots, H_{2,n_2})$ be the maximal subchain of hyperplanes defining $\Psi_2$ that separate $\gamma$ from $\beta$, numbered from $\gamma$ to $\beta$.

Let $H$ be a hyperplane of $Q$ adjacent to $\gamma$ and either equal to $H_{2,1}$ or separating $\gamma$ from $H_{2,1}$, and let $\gamma' \in [\alpha, \beta]$ be the vertex adjacent to $\gamma$ such that $H$ crosses the edge $[\gamma, \gamma']$. First note that, since $H_{2,1}$ separates $\gamma$ and $\beta$, we deduce that $H$ separates $\gamma$ and $\beta$. Thus, $H$ does not separate $\alpha$ and $\gamma$, because $\gamma \in [\alpha, \beta]$. In particular, $\gamma \in [\alpha, \gamma']$. Since $\gamma$ is chosen as far from $\alpha$ as possible among points at $\sigma_{\mathcal{C}}$–distance equal to $r$, and every hyperplane separating $\alpha$ and $\gamma$ separates $\alpha$ and $\gamma'$, we deduce that $\sigma_{\mathcal{C}}(\alpha, \gamma') > \sigma_{\mathcal{C}}(\alpha, \gamma) = r$, so $\sigma_{\mathcal{C}}(\alpha, \gamma') = \sigma_{\mathcal{C}}(\alpha, \gamma) + 1$.

Let $\Psi_1 \in \mathcal{C}$ be such that $\sigma_{\mathcal{C}}(\alpha, \gamma') = |\Psi_1(\alpha) - \Psi_1(\gamma')|$. Let $(H_{1,1}, \dots, H_{1,n_1+1})$ be the maximal subchain of hyperplanes defining $\Psi_1$ that separate $\alpha$ from $\gamma'$, numbered from $\alpha$ to $\gamma'$. Since $\sigma_{\mathcal{C}}(\alpha, \gamma') = \sigma_{\mathcal{C}}(\alpha, \gamma) + 1$, we know that $H = H_{1,n_1+1}$ and that $\sigma_{\mathcal{C}}(\alpha, \gamma) = |\Psi_1(\alpha) - \Psi_1(\gamma)|$. In particular, $H$ is disjoint from $H_{1,1}, \dots, H_{1,n_1}$. We deduce that $H$ separates $H_{1,1}, \dots, H_{1,n_1}$ from $H_{2,1}, \dots, H_{2,n_2}$, and the conclusion follows. □

We can now use these lemmas to prove that, in the setting of Theorem 2.13, the metric $\sigma$ is weakly roughly geodesic (Definition 2.4).

**Proposition 2.21** *If $(X, \mu, d)$ has quasicubical intervals and is roughly geodesic, then $\sigma$ is weakly roughly geodesic.*

**Proof** Let $a, b \in X$. Since $X$ has quasicubical intervals, there exists a finite CAT(0) cube complex $Q$ (with the $\ell^1$ metric) of dimension at most $\nu$ and a map $\lambda\colon Q \to [a, b]$ that is a $(\kappa, \kappa)$–quasi-isometry and a $\kappa$–quasimedian homomorphism. We can therefore fix $\alpha, \beta \in Q$ such that $d(\lambda(\alpha), a) \leqslant \kappa$ and $d(\lambda(\beta), b) \leqslant \kappa$. According to Proposition 2.16, there is a constant $q \geq 1$ such that $d$ and $\sigma$ are $(q, q)$–quasi-isometric. It follows that $\sigma(\lambda(\alpha), a) \leq q(\kappa + 1)$ and $\sigma(\lambda(\beta), b) \leq q(\kappa + 1)$.

For each $K$–contraction $\Phi\colon X \to \mathbb{R}$, the composition $\Phi\lambda\colon Q \to \mathbb{R}$ is a $K'$–quasimedian, $(K', K')$–coarsely Lipschitz map, where $K' = K + \kappa$. According to Lemma 2.19, there exists a 0–contraction $\Psi\colon Q \to \mathbb{Z}$ such that $|\Phi\lambda(\xi) - 4K'\nu\Psi(\xi)| \leqslant 4K'\nu$ for all $\xi \in Q^{(0)}$. Let $\mathcal{C}$ denote the set of all 0–contractions $\Psi\colon Q \to \mathbb{Z}$ such that there is some $K$–contraction $\Phi\colon X \to \mathbb{Z}$ with $|\Phi\lambda(\xi) - 4K'\nu\Psi(\xi)| \leqslant 4K'\nu$ for all $\xi \in Q^{(0)}$.

We shall prove that $\sigma$ is weakly roughly geodesic with constant

$$C'_\sigma = 64K'\nu + 4q(\kappa+1) + 4\kappa + 4K + 2D,$$

where $D$ is the constant from Lemma 2.17.

Let $r \in [0, \sigma(a,b)]$. If $r < C'_\sigma$, then clearly we can take $c = a$ for the desired point. Similarly, if $r > \sigma(a,b) - C'_\sigma$, then we can take $c = b$. Otherwise, Lemma 2.20, applied to $\alpha$, $\beta$, the family $\mathcal{C}$ and $r' = \lfloor r/4K'\nu \rfloor$, provides a vertex $\gamma \in [\alpha, \beta]$ and 0–contractions $\Psi_1, \Psi_2 \in \mathcal{C}$. Let $c = \lambda(\gamma) \in [a,b]$.

Let us start by computing $\sigma(a,c)$. By definition of the set $\mathcal{C}$, for any $K$–contraction $\Phi\colon X \to \mathbb{R}$ there is some $\Psi \in \mathcal{C}$ (and, conversely, for any $\Psi \in \mathcal{C}$ there exists a $K$–contraction $\Phi$) such that

$$\begin{aligned}\big||\Phi\lambda(\xi) - \Phi\lambda(\zeta)| - 4K'\nu|\Psi(\xi) - \Psi(\zeta)|\big| &\le \big||\Phi\lambda(\xi) - 4K'\nu\Psi(\xi)| + |\Phi\lambda(\zeta) - 4K'\nu\Psi(\zeta)|\big| \\ &\le 8K'\nu\end{aligned}$$

holds for all $\xi, \zeta \in Q^{(0)}$. It follows that

$$|\sigma(\lambda(\xi), \lambda(\zeta)) - 4K'\nu\sigma_{\mathcal{C}}(\xi,\zeta)| \leqslant 8K'\nu. \tag{1}$$

By the choice of $\gamma$, we have $\sigma_{\mathcal{C}}(\alpha,\gamma) = r'$. Thus, from (1) we obtain

$$\begin{aligned}|\sigma(a,c) - r| &\le |\sigma(\lambda(\alpha), \lambda(\gamma)) - 4K'\nu r'| + q(\kappa+1) + 4K'\nu \\ &\le 12K'\nu + q(\kappa+1) \le C'_\sigma.\end{aligned}$$

The aim for the rest of the proof is to confirm the second restriction on $c$, namely that $\sigma(a,c) + \sigma(c,b) \le \sigma(a,b) + C'_\sigma$. The strategy is to apply Lemma 2.17.

Recall that $\Psi_1, \Psi_2 \in \mathcal{C}$ are the 0–contractions provided by Lemma 2.20: they satisfy $\sigma_{\mathcal{C}}(\alpha,\gamma) = |\Psi_1(\alpha) - \Psi_1(\gamma)| = r'$ and $\sigma_{\mathcal{C}}(\gamma,\beta) = |\Psi_2(\gamma) - \Psi_2(\beta)| = s'$. After translations of $\mathbb{Z}$, we may also assume that $\Psi_1(\alpha) = 0$, $\Psi_1(\gamma) = \Psi_2(\gamma) = r'$ and $\Psi_2(\beta) = r' + s'$. By definition of $\mathcal{C}$, there exist $K$–contractions $\Phi_1$ and $\Phi_2$ on $X$ such that $|\Phi_1\lambda(\xi) - 4K'\nu\Psi_1(\xi)| \leqslant 4K'\nu$ and $|\Phi_2\lambda(\xi) - 4K'\nu\Psi_2(\xi)| \leqslant 4K'\nu$ for all

$\xi \in Q^{(0)}$. In particular, $\Phi_1(a) \le \Phi_1\lambda(\alpha) + \kappa + K \le 4K'\nu + \kappa + K$. Moreover, by using (1) we see that

$$\begin{aligned}\Phi_2(b) &\ge 4K'\nu(r'+s') - 4K'\nu - \kappa - K \\ &= 4K'\nu(\sigma_{\mathcal{C}}(\alpha,\gamma) + \sigma_{\mathcal{C}}(\gamma,\beta)) - 4K'\nu - \kappa - K \\ &\ge \sigma(\lambda(\alpha),\lambda(\gamma)) + \sigma(\lambda(\gamma),\lambda(\beta)) - 16K'\nu - 4K'\nu - \kappa - K \\ &\ge \sigma(a,c) + \sigma(c,b) - 20K'\nu - 2q(\kappa+1) - \kappa - K.\end{aligned}$$

We are now in the setting of Lemma 2.17, with $E = 20K'\nu + 2q(\kappa+1) + \kappa + K$ and with the image of $\Phi_2$ bounded above by $\sigma(a,c)+\sigma(c,b)$. Let us show that the assumptions of the lemma are met if we take $t = 12K'\nu + \kappa + K$.

We must first note that $r - D + K - E \ge C'_\sigma - D + K - E \ge t$, and secondly that $\sigma(a,c)+\sigma(c,b)-r-D+K-E \ge \sigma(a,b)-r-D+K-E \ge C'_\sigma - D + K - E \ge t$.

It remains to prove that the subspaces $Z_1 = \{z \in X \mid \Phi_1(z) \leqslant r - t - K\}$ and $Z_2 = \{z \in X \mid \Phi_2(z) \geqslant r + t + K\}$ are disjoint. Fix $z \in X$, let $x = \mu(z,a,b)$ and pick any $\xi \in Q^{(0)}$ such that $d(\lambda(\xi), x) \leqslant \kappa$.

If $z \in Z_1$, so that $\Phi_1(x) \leqslant \Phi_1(z) + K \le r - t$, then $\Phi_1(\lambda(\xi)) \leqslant r - t + \kappa + K$, and hence

$$\Psi_1(\xi) \leqslant \frac{r - t + \kappa + K + 4K'\nu}{4K'\nu} \leqslant \frac{r - 8K'\nu}{4K'\nu} \leqslant r' - 1.$$

Similarly, if $z \in Z_2$, then $\Psi_2(\xi) \geqslant r' + 1$.

According to property (3) of Lemma 2.20, the halfspace of $H_{\Psi_1,r'}$ containing $\alpha$ is disjoint from the halfspace of $H_{\Psi_2,1}$ containing $\beta$. Thus, if $\xi \in Q^{(0)}$, then we cannot simultaneously have both $\Psi_1(\xi) \leqslant r'-1$ and $\Psi_2(\xi) \geqslant r'+1$. As a consequence, we cannot have both $z \in Z_1$ and $x \in Z_2$. This implies that $Z_1 \cap Z_2 = \varnothing$.

The conditions of Lemma 2.17 are therefore met, and, by applying it, we deduce that $\sigma(a,b) \ge \sigma(a,c) + \sigma(c,b) - 2t - 2D - 2E = \sigma(a,c) + \sigma(c,b) - C'_\sigma$. This completes the proof that $\sigma$ is weakly roughly geodesic with constant $C'_\sigma$. □

## 2.5 Coarse convexity of balls

To complete the proof of Theorem 2.13, it remains to show that balls in $(X,\sigma)$ are uniformly coarsely median convex.

**Lemma 2.22** *There is a constant $\epsilon \geqslant 0$ such that, for any $x, y, z \in X$ with $x \in [y, z]$, we have $d(x, \mu(x, y, z)) \leq \epsilon$.*

**Proof** According to [19, Lemma 8.1], there are constants $r_0$ and $r_0'$ such that $x$ lies at distance at most $r_0'$ from a point $x'$ with $d(x', \mu(x', y, z)) \leq r_0$. Since the coarse median $\mu$ is coarsely Lipschitz, $d(\mu(x, y, z), \mu(x', y, z)) \leqslant h(0)d(x, x') + h(0) \leq h(0)(r_0' + 1)$. We deduce that $d(x, \mu(x, y, z)) \leq \epsilon$, where $\epsilon = r_0' + r_0 + h(0)(r_0' + 1)$. □

**Lemma 2.23** *Suppose that $(X, \mu, d)$ has quasicubical intervals and is roughly geodesic. There is a constant $M$ such that each ball in $(X, \sigma)$ is $M$–coarsely median convex.*

**Proof** Fix $w \in X$ and $R \geqslant 0$. Let $y, z \in B_\sigma(w, R)$. Given any $a \in X$, we want to bound the distance from $x = \mu(a, y, z)$ to $B_\sigma(w, R)$.

Let $r < \sigma(w, x)$ and let $\Phi\colon X \to [0, r]$ be a $K$–contraction such that $\Phi(w) = 0$ and $\Phi(x) \geq r$. Lemma 2.22 tells us that $d(\mu_{x,y,z}, x) \leqslant \epsilon$, so $|\Phi(\mu_{x,y,z}) - \Phi(x)| \leq \epsilon + K$. Since $\Phi$ is a $K$–quasimedian homomorphism,

$$\mu_{\mathbb{R}}(\Phi(x), \Phi(y), \Phi(z)) \geq \Phi(\mu_{x,y,z}) - K \geq \Phi(x) - \epsilon - 2K \geq r - \epsilon - 2K.$$

This means that one of $\Phi(y)$ and $\Phi(z)$ must be at least $r - \epsilon - 2K$, and so $\sigma(w, x) \leqslant \max\{\sigma(w, y), \sigma(w, z)\} + \epsilon + 2K$.

This proves that $x \in B_\sigma(w, R + \epsilon + 2K)$. According to Proposition 2.21, $\sigma$ is weakly roughly geodesic with constant $C_\sigma'$. Applying Definition 2.4 with $a = w$, $b = x$ and $r = \min\{R - C_\sigma', \sigma(w, x)\}$ yields a point $x' \in B_\sigma(w, R)$ with $d(x', x) \leq \epsilon + 2K + 3C_\sigma' = M$, which shows that balls in $(X, \sigma)$ are $M$–coarsely median convex. □

# 3 Quasiconvexity and a coarse Helly property in HHSs

The goal of this section is to prove that *hierarchically quasiconvex* subsets of hierarchically hyperbolic spaces satisfy a coarse version of the Helly property. Since coarsely median convex subsets of a hierarchically hyperbolic space are hierarchically quasiconvex [71, Proposition 5.11], this applies in particular to balls for the metric $\sigma$ constructed in Section 2, by Theorem 2.13, allowing us to deduce Theorem A. We also deduce the bounded packing property for hierarchically quasiconvex subgroups of groups that are HHSs.

## 3.1 Background on hierarchical hyperbolicity

Here we give a description of hierarchically hyperbolic spaces (HHSs) and hierarchically hyperbolic groups (HHGs). For full definitions, see [13, Definitions 1.1 and 1.21]. Briefly, an HHS consists of a quasigeodesic space $(X, d)$, a constant $E$ and a set $\mathfrak{S}$, elements of which are called *domains*. Each domain $U$ has an associated $E$–hyperbolic space $\mathcal{C}U$, and the various axioms give structure for extracting information about $X$ from these hyperbolic spaces. This includes:

- Each domain $U$ has an associated $E$–coarsely onto, $(E, E)$–coarsely Lipschitz *projection* map $\pi_U : X \to \mathcal{C}U$.
- $\mathfrak{S}$ has a partial order $\sqsubset$, called nesting, and a symmetric relation $\perp$, called orthogonality. If $U \sqsubset V$ and $V \perp W$, then $U \perp W$. The relations $\sqsubsetneq$, $\perp$ and $=$ are mutually exclusive, and their complement, denoted by $\pitchfork$, is called transversality.
- There is a bound on the size of $\sqsubsetneq$–chains and pairwise orthogonal sets.
- If $U \sqsubsetneq V$ or $U \pitchfork V$, then there is a set $\rho^U_V \subset \mathcal{C}V$ of diameter at most $E$.
- If $U \sqsubsetneq V$, then there is also a map $\rho^V_U : \mathcal{C}V \to \mathcal{C}U$. If $\gamma \subset \mathcal{C}V$ is a geodesic and $d_{\mathcal{C}V}(\gamma, \rho^U_V) > E$, then $\operatorname{diam} \rho^V_U(\gamma) \leqslant E$.

This last point is referred to as *bounded geodesic image*. For $x, y \in X$, it is standard to write $d_U(x, y)$ in place of $d_{\mathcal{C}U}(\pi_U(x), \pi_U(y))$, and similarly for subsets of $X$. Moreover, we can always assume that $X$ and the associated hyperbolic spaces are graphs (for example by [33, Lemma 3.B.6]). In particular, we can and shall assume that $X$ and the $\mathcal{C}U$ are geodesic.

We say that $X$ admits an HHS structure if there is an HHS whose underlying metric space is $X$, and we write $(X, \mathfrak{S})$ as shorthand for the entirety of a choice of HHS structure. An HHG is a finitely generated group $G$ whose Cayley graph admits an HHS structure $(G, \mathfrak{S})$ such that $G$ acts cofinitely on $\mathfrak{S}$ and elements of $G$ induce isometries $\mathcal{C}U \to \mathcal{C}gU$ for all $U \in \mathfrak{S}$. (There are a couple of other natural regulatory assumptions that we shall not concern ourselves with here.)

The idea behind two domains being orthogonal is that one can see a direct product of associated sub-HHSs inside $X$. This is made precise by the *partial realisation* axiom.

**Axiom** (partial realisation) If $\{U_i\}$ is a set of pairwise orthogonal domains, then, for any choice of points $p_i \in \mathcal{C}U_i$, there is some $x \in X$ with $d_{U_i}(x, p_i) \leq E$ for all $i$, and with $d_V(x, \rho^{U_i}_V) \leq E$ whenever $U_i \sqsubsetneq V$ or $U_i \pitchfork V$.

In fact, one of the main tools for dealing with HHSs is the *realisation theorem* [13, Theorem 3.1], which extends the partial realisation axiom. Roughly, it says that any *consistent tuple* is well approximated by the projections of some point in $X$. In other words, performing constructions in $X$ can be reduced to performing constructions in the associated hyperbolic spaces and checking that the points produced by this process are consistent.

**Definition 3.1** (consistent tuple) For a constant $\kappa \geqslant E$, a tuple $(b_U) \in \prod_{U \in \mathfrak{S}} \mathcal{C}U$ is said to be $\kappa$–consistent if

$$\min\{d_U(b_U, \rho_U^V), d_V(b_V, \rho_V^U)\} \leqslant \kappa \quad \text{whenever } U \pitchfork V$$

and

$$\min\{d_V(b_V, \rho_V^U), \operatorname{diam}(b_U \cup \rho_U^V(b_V))\} \leqslant \kappa \quad \text{whenever } U \subsetneq V.$$

**Axiom** (consistency) For any $x \in X$, the tuple $(\pi_U(x))_{U \in \mathfrak{S}}$ is $E$–consistent.

It will be useful to be able to talk about consistency for subsets of $\mathfrak{S}$. Given $u \in \mathcal{C}U$ and $v \in \mathcal{C}V$, we say that $u$ and $v$ *satisfy the consistency inequalities* for $U$ and $V$ if

- $U \pitchfork V$ and $\min\{d_U(u, \rho_U^V), d_V(v, \rho_V^U)\} \leq E$, or
- (after relabelling) $U \subsetneq V$ and $\min\bigl\{d_V(v, \rho_V^U), \operatorname{diam}(\{u\} \cup \rho_U^V(v))\bigr\} \leq E$.

Let us now state the realisation theorem, which will be the mechanism for our proof of Theorem 3.5. We shall only need the existence part.

**Theorem 3.2** (realisation [13, Theorem 3.1]) *For each $\kappa \geq E$, there are numbers $\theta_e(\kappa)$ and $\theta_u(\kappa)$ such that, if $(b_U)_{U \in \mathfrak{S}}$ is a $\kappa$–consistent tuple, then there is some $x \in X$ with $d_U(x, b_U) \leq \theta_e(\kappa)$ for all domains $U$. Moreover, the set of such $x$ has diameter at most $\theta_u(\kappa)$.*

A key application of the realisation theorem is for the construction of a coarse median operation for HHSs. Given three points $x$, $y$ and $z$ in an HHS $(X, \mathfrak{S})$, let $(m_U)_{U \in \mathfrak{S}}$ be the tuple whose $U$–entry is the median of the triple $(\pi_U(x), \pi_U(y), \pi_U(z))$ in the hyperbolic space $\mathcal{C}U$. This tuple is consistent [13, Theorem 7.3], so we define $\mu(x, y, z)$ to be a point obtained by applying the realisation theorem to the tuple $(m_U)$. (One also needs Proposition 10.1 of Bowditch [17] to conclude that $(X, \mu, d)$ is a coarse median space.) When $X$ is an HHG, one can arrange for $\mu$ to be equivariant.

The action on the index set is what distinguishes HHGs from groups that are HHSs, and this turns out to be an important distinction. For example, the property of being an HHS is invariant under quasi-isometries, but there are groups that are virtually HHGs but not HHGs themselves. Indeed, the $(3,3,3)$ triangle group is virtually abelian, but, as mentioned in the introduction, it is not coarsely injective [50], and it therefore cannot be an HHG by Corollary H. A more direct proof, not relying on the results of this paper, is given in [66]. On the other hand, any group that is an HHS can be equipped with a coarse median [13], but this may fail to be equivariant if the structure is only an HHS structure.

A related notion that is closed under taking subgroups is that of a group that acts on an HHS $(X,\mathfrak{S})$ by *HHS automorphisms*. In other words, it acts on $X$ isometrically and on $\mathfrak{S}$ with the regulatory assumptions alluded to above, but the action on $\mathfrak{S}$ need not be cofinite. The median is still equivariant for such actions.

In the theory of hyperbolic spaces, an important class of subsets are the quasiconvex subsets, because they inherit the structure of the ambient space. The natural analogue in the setting of hierarchical hyperbolicity is that of a *hierarchically quasiconvex* subset.

**Definition 3.3** (hierarchical quasiconvexity) A subset $Y$ of an HHS $(X,\mathfrak{S})$ is said to be hierarchically quasiconvex if there is a function $k$ such that every $\pi_U(Y)$ is $k(0)$–quasiconvex and, if $x \in X$ has $d_U(x,Y) \leqslant r$ for all $U \in \mathfrak{S}$, then $d_X(x,Y) \leqslant k(r)$.

We finish this section with some examples.

All hyperbolic groups are hierarchically hyperbolic, as are the (extended) mapping class groups of finite-type surfaces [11]; Teichmüller space with either of the standard metrics [11]; many graphs defined from curves on surfaces, including the pants graph [76]; quotients of mapping class groups by powers of pseudo-Anosovs [10] and Dehn-twist subgroups [9]; extensions of Veech groups [35]; the genus-two handlebody group [31]; fundamental groups of closed 3–manifolds without Nil or Sol components [13]; right-angled Artin groups [11]; and, in fact, all known cubical groups [48]. Aside from the extensions of Veech groups and some 3–manifold groups, the groups listed here are all known to be HHGs, not merely HHSs.

There are also various ways to combine HHSs and HHGs to produce new ones. For example, both classes are closed under relative hyperbolicity [13], any graph product of HHGs is an HHG [16], and many graphs of groups are HHGs [13; 15; 68].

## 3.2 Coarse injectivity

Here we prove our result on hierarchically quasiconvex subsets of an HHS and deduce that HHSs are coarsely injective when equipped with the metric $\sigma$ from Section 2. We then deduce that every HHG acts properly cocompactly by isometries on a coarsely injective space.

We shall make use of the following powerful result for hyperbolic spaces. The version stated here is a combination of [29, Lemma 5.1] and the proof of [29, Theorem 5.1]. It states in particular that quasiconvex subsets of a hyperbolic graph satisfy a coarse version of the Helly property. Throughout this section, we say that subsets $Z_1$ and $Z_2$ of a metric space $(X, d)$ are $r$–close if there exist $z_1 \in Z_1$ and $z_2 \in Z_2$ with $d(z_1, z_2) \leq r$.

**Theorem 3.4** [29] *The following holds for any nonnegative constants $E$, $r$ and $k_0$: Let $Y$ be an $E$–hyperbolic graph and let $y$ be a vertex of $Y$. Suppose that $\mathcal{Q}$ is a collection of pairwise $2Er$–close $k_0$–quasiconvex subsets of $Y^{(0)}$ with the property that $\{d(y, Q) : Q \in \mathcal{Q}\}$ is bounded. By discreteness, we can fix $Q \in \mathcal{Q}$ with $d(y, Q)$ maximal. Let $z \in Q$ have $d(y,z) = d(y, Q)$, and let $c$ be the point on a geodesic $[y, z]$ with $d(c, z) = \min\{Er, d(y, z)\}$. Then $d(c, Q') \leqslant r'$ for all $Q' \in \mathcal{Q}$, where $r' = \max\{2k_0 + 5E, Er + k_0 + 3E\}$.*

The strength of this theorem is twofold. Firstly, the constant $r'$ is independent of the size of the set $\mathcal{Q}$ — a statement with this independence does not seem to appear elsewhere in the geometric group theory literature. The second strength is that the construction of the point $c$ is both completely explicit and allows for a lot of flexibility in the choice of $y$. Observe that the condition that $\{d(y, Q) : Q \in \mathcal{Q}\}$ is bounded is satisfied automatically if any $Q \in \mathcal{Q}$ is bounded.

We will now prove that hierarchically quasiconvex subsets of a hierarchically hyperbolic space satisfy a coarse version of the Helly property.

**Theorem 3.5** (coarse Helly property) *Let $(X, \mathfrak{S})$ be an HHS with constant $E$, and let $\mathcal{Q}$ be a collection of $k$–hierarchically quasiconvex subsets of $X$ such that either $\mathcal{Q}$ is finite or $\mathcal{Q}$ contains an element with bounded diameter. Suppose that there is a constant $r$ such that any two elements of $\mathcal{Q}$ are $r$–close. There is a constant $R = R(E, k, r)$ such that there is a point $x \in X$ with $d(x, Q) \leqslant R$ for all $Q \in \mathcal{Q}$.*

**Proof** Let us say that a domain $U$ *begets* a domain $V$ if either $U \pitchfork V$ or $U \subsetneq V$. If $U$ begets $V$, then there is a well-defined bounded set $\rho_V^U$.

Let $\mathcal{U} = \{U_1, \dots, U_n\}$ be a maximal collection of pairwise orthogonal, nest-minimal domains. Note that we may choose $\mathcal{U}$ arbitrarily. For any domain $V \in \mathfrak{S} \smallsetminus \mathcal{U}$, there is some $i$ such that $U_i$ begets $V$. By [37, Lemma 1.5], for any domain $V \in \mathfrak{S}$, we have $d_V(\rho_V^{U_i}, \rho_V^{U_j}) \leqslant 2E$ whenever $U_i$ and $U_j$ both beget $V$. Moreover, recall that $\operatorname{diam} \rho_V^{U_i} \leqslant E$. At the cost of increasing the hierarchical hyperbolicity constant to at most $10E$, we can therefore perturb the HHS structure to assume that every $\rho_V^{U_i}$ is a singleton, and that $\rho_V^{U_i} = \rho_V^{U_j}$ whenever both $U_i$ and $U_j$ beget $V$. We write $\rho_V^{\mathcal{U}}$ for the singleton

$$\rho_V^{\mathcal{U}} = \bigcup_{\{i : U_i \text{ begets } V\}} \rho_V^{U_i}.$$

As mentioned, the construction of $\mathcal{U}$ ensures that the point $\rho_V^{\mathcal{U}}$ exists for all $V \in \mathfrak{S} \smallsetminus \mathcal{U}$.

We are free to assume that if $r > 0$ then $r > 1$. Thus, by definition of hierarchical quasiconvexity and the fact that projection maps are $(E, E)$–coarsely Lipschitz, for any domain $V$, the sets $\pi_V(Q)$ for $Q \in \mathcal{Q}$ are pairwise $2Er$–close and $k_0$–quasiconvex, where $k_0 = k(0)$. We assumed that $\mathcal{Q}$ either is finite or it contains an element with bounded diameter, so, for any point $y \in X$ and any domain $V$, the set $\{d_V(y, Q) : Q \in \mathcal{Q}\}$ is bounded. Let $r'$ be as in the statement of Theorem 3.4. That theorem now allows us to choose, for each $U \in \mathcal{U}$, a point $b_U$ in $\mathcal{C}U$ with $d_U(b_U, Q) \leqslant r'$ for all $Q \in \mathcal{Q}$. For any other domain $V$, let $b_V$ be the point of $\mathcal{C}V$ obtained by applying Theorem 3.4 in the hyperbolic graph $\mathcal{C}V$, with quasiconvex subsets $\{\pi_V(Q) : Q \in \mathcal{Q}\}$ and starting vertex $\rho_V^{\mathcal{U}}$.

**Claim** *The tuple $(b_V)_{V \in \mathfrak{S}}$ is $(r' + 7E + Er)$–consistent.*

**Proof** Suppose that $W$ begets $V$ and $d_V(\rho_V^W, \rho_V^{\mathcal{U}}) \leqslant 2E$. Assume that $d_V(b_V, \rho_V^W) > r' + 7E + Er$. By the construction of $b_V$, there exists some $Q \in \mathcal{Q}$ such that $d_V(b_V, Q) \leqslant Er$. As a consequence,

$$d_V(Q, \rho_V^W) \geq d_V(b_V, \rho_V^W) - d_V(b_V, Q) - \operatorname{diam} \rho_V^W > r' + 6E.$$

If $W \pitchfork V$, then $\pi_W(Q)$ is contained in the $E$–neighbourhood of $\rho_W^V$ by consistency for elements of $Q$. In particular, $d_W(\rho_W^V, b_W) \leqslant r' + E$ as $b_W$ is $r'$–close to $\pi_W(Q)$. If $W \subsetneq V$, then, since $\pi_V(Q)$ is $k_0$–quasiconvex and $r' + 6E > k_0 + 6E$, bounded geodesic image and consistency show that the set $\rho_W^V(\pi_V(Q))$ has diameter at most $E$, and its $E$–neighbourhood contains $\pi_W(Q)$. Moreover, its $E$–neighbourhood contains $\rho^V_W(b_V)$ by bounded geodesic image, as witnessed by the geodesic used to construct $b_V$. Thus,

$$\begin{aligned}\operatorname{diam}(b_W \cup \rho^V_W(b_V)) &\leqslant d_W(b_W, Q) + \operatorname{diam}\pi_W(Q) + d_W(Q, \rho^V_W(b_V)) + \operatorname{diam}\rho^V_W(b_V)\\ &\leqslant r' + 3E + 3E + E\\ &= r' + 7E.\end{aligned}$$

The above paragraph will be referred to as $(*)$ for the rest of the proof of the claim. We split the checking of the consistency inequalities for pairs $(V, W)$ of domains into three cases.

**Case 1** ($W \in \mathcal{U}$ begets $V$) In this case, $\rho^W_V = \rho^{\mathcal{U}}_V$, so we are done by $(*)$.

**Case 2** (there is some $U \in \mathcal{U}$ that begets both $V$ and $W$) Proposition 1.8 of [13] states that, if $W$ begets $V$, then $\rho^U_V$ and $\rho^U_W$ satisfy the consistency inequalities for $V$ and $W$. Consequently, by $(*)$, the only case we need to check here is when $U \pitchfork W$, $W \subsetneq V$ and $\operatorname{diam}(\rho^U_W \cup \rho^V_W(\rho^U_V)) \leqslant 2E$. Assuming that $d_V(\rho^W_V, b_V) > r' + 7E + Er$, there are two possibilities, depending on the location of $\rho^U_V$.

If there is a geodesic $[\rho^U_V, b_V]$ that is disjoint from the $E$–neighbourhood of $\rho^W_V$, then $\operatorname{diam}(\rho^V_W(\rho^U_V) \cup \rho^V_W(b_V)) \leqslant E$, so $d_W(\rho^U_W, \rho^V_W(b_V)) \leqslant 3E$. Moreover, for each $Q \in \mathcal{Q}$ there is some $q \in Q$ such that any geodesic $[b_V, \pi_V(q)]$ is disjoint from the $E$–neighbourhood of $\rho^W_V$. In particular, $\rho^V_W(b_V)$ is $2E$–close to each $\pi_W(q)$, and hence $\rho^U_W$ is $5E$–close to each $\pi_W(Q)$. Since $b_W$ lies on a shortest geodesic between $\rho^U_W$ and some $\pi_W(Q)$, we get that $d_W(b_W, \rho^U_W) \leqslant 5E$, and so $b_W$ is $8E$–close to $\rho^V_W(b_V)$.

Otherwise, every geodesic $[\rho^U_V, b_V]$ meets the $E$–neighbourhood of $\rho^W_V$. By construction of $b_V$, there exists $Q \in \mathcal{Q}$ such that $d_V(\rho^W_V, Q) > (r' + 7E + Er) + Er - 2E = r' + 5E + 2Er$. By the same argument as in $(*)$, we now get that $\rho^V_W(b_V)$ is $3E$–close to $\pi_W(Q)$, which has diameter at most $3E$. Hence, $\operatorname{diam}(b_W \cup \rho^V_W(b_V)) \leqslant r' + 7E$.

**Case 3** (no $U_i$ begets both $V$ and $W$, and neither $V$ nor $W$ is in $\mathcal{U}$) After relabelling, we can assume that $U_1$ begets $V$ and $U_2$ begets $W$. Since $U_1$ does not beget $W$, we have $U_1 \perp W$, and, similarly, $U_2 \perp V$. In particular, the only case that needs checking is when $V \pitchfork W$. The partial realisation axiom applied to any points $p_1 \in \mathcal{C}U_1$ and $p_2 \in \mathcal{C}U_2$ provides a point $z \in X$ such that $d_V(z, \rho^{U_1}_V) \leqslant E$ and $d_W(z, \rho^{U_2}_W) \leqslant E$. By consistency for $z$, either $d_V(\rho^W_V, \rho^{U_1}_V) \leqslant 2E$ or $d_W(\rho^V_W, \rho^{U_2}_W) \leq 2E$. We are done by $(*)$. $\lhd$

In light of the claim, Theorem 3.2 provides a point $x \in X$ such that $d_V(x, b_V) \leqslant \theta_e(r' + 7E + Er)$ for all $V \in \mathfrak{S}$. By construction of the points $b_V$, we have that

$d_V(x,Q) \leqslant r' + \theta_e(r'+7E+Er)$ for all $Q \in \mathcal{Q}$. Hierarchical quasiconvexity of $Q$ now tells us that $x$ is $k(r'+\theta_e(r'+7E+Er))$–close to $Q$ for all $Q \in \mathcal{Q}$. $\square$

It is worth noting that the proof of Theorem 3.5 gives flexibility of a similar kind to that in Theorem 3.4. Indeed, we are free in our choice of $\mathcal{U}$ and, once this is chosen, we apply the Chepoi–Dragan–Vaxès construction in each of the hyperbolic spaces associated with $\mathcal{U}$, without restriction on the choice of starting point therein. We shall not need to make use of this in the present paper.

**Corollary 3.6** *If $X$ is an HHS, then $(X,\sigma)$ is coarsely injective, and hence roughly geodesic.*

**Proof** By Proposition 2.12, the geodesic coarse median space $(X,\mu,d)$ has quasicubical intervals, so Theorem 2.13 tells us that the metric $\sigma$ is weakly roughly geodesic on $X$, that it is quasi-isometric to $d$ and that $\sigma$–balls are uniformly coarsely median convex. Let $\{B_\sigma(x_i,r_i) : i \in I\}$ be a family of balls in $(X,\sigma)$ with the property that $\sigma(x_i,x_j) \leq r_i + r_j$ for all $i,j \in I$. Since $\sigma$ is weakly roughly geodesic, there is a constant $\delta$, independent of the family of balls, such that the balls $B_\sigma(x_i,r_i+\delta)$ intersect pairwise.

Let $B_i$ be the image of the ball $B_\sigma(x_i,r_i+\delta)$ under the identity quasi-isometry $(X,\sigma) \to (X,d)$. The $B_i$ are uniformly coarsely median convex, and so they are uniformly hierarchically quasiconvex by [71, Proposition 5.11]. They also intersect pairwise, and each is bounded, so Theorem 3.5 produces a point at uniformly bounded $d$–distance from each $B_i$. As $d$ and $\sigma$ are quasi-isometric, this point is at uniformly bounded $\sigma$–distance from each $B_\sigma(x_i,r_i+\delta)$. Thus, $(X,\sigma)$ is coarsely injective.

Since any injective space is geodesic, we deduce that the coarsely injective metric space $(X,\sigma)$ is not merely weakly roughly geodesic, but actually roughly geodesic, as it is coarsely dense in its injective hull. $\square$

Usually it really is necessary to change the metric: Example 5.13 of [26] shows that $\mathbb{Z}^3$ with the standard $\ell^1$ metric is not coarsely injective, though it is an HHG.

We now explain how to deduce the existence of a bicombing from work of Lang. See Section 1.3 for the definitions of roughly conical and roughly reversible bicombings.

**Corollary 3.7** *If $(X,\mathfrak{S})$ is an HHS, then $(X,\sigma)$ admits a roughly conical, roughly reversible bicombing by rough geodesics that is coarsely equivariant under the automorphism group of $(X,\mathfrak{S})$.*

**Proof** According to Corollary 3.6, the metric space $(X,\sigma)$ is coarsely injective, so it is $D$–coarsely dense in its injective hull for some $D$. A construction of Lang shows that every injective metric space $E$ admits a conical, reversible, geodesic, Isom $E$–invariant bicombing $\gamma'$ [60]. Take $E = E((X,\sigma))$. For each $a, b \in X$ and $t \in [0,1]$, define $\gamma_{a,b}(t)$ as any point of $X$ at distance at most $D$ from $\gamma'_{a,b}(t)$. Since $\gamma(t)$ is at uniform distance $D$ from $\gamma'(t)$, we deduce that $\gamma$ is a bicombing on $(X,\sigma)$ with the listed properties. □

Note that, if the action of the automorphism group of $(X,\mathfrak{S})$ on $X$ is free, then the bicombing may be chosen to be actually equivariant.

Let us now discuss the consequences of our construction for HHGs.

**Corollary 3.8** *If $G$ is an HHG, then $G$ admits a proper, cocompact, isometric action on the coarsely injective space $(G,\sigma)$.*

**Proof** $(G,\sigma)$ is coarsely injective by Corollary 3.6. Since the median is equivariant in an HHG, Lemma 2.10 tells us that the action is isometric. Properness and cocompactness follow from Proposition 2.16. □

**Remark 3.9** In fact, we do not quite need to assume that we have a hierarchically hyperbolic group in Corollary 3.8: we only need a proper cocompact action by median isometries on an HHS. In fact, cocompactness can be relaxed to coboundedness for the sake of the applications in this paper. For example, it would be sufficient to assume that $G$ is a group acting properly coboundedly by HHS automorphisms on an HHS. The consequences for HHGs listed here and in the introduction therefore apply in this generality.

The next lemma is a modified version of [26, Proposition 6.7], in which the assumption that the hull is proper has been dropped.

**Lemma 3.10** *If a group $G$ acts properly coboundedly on a coarsely injective space $X$, then $G$ acts properly coboundedly on the injective hull $E(X)$. In particular, every HHG admits a proper, cobounded action on an injective space.*

**Proof** There is an induced action of $G$ on $E(X)$ and the isometric embedding $e\colon X \to E(X)$ is equivariant with respect to this induced action [60, Proposition 3.7]. To simplify notation, we identify the points of $X$ with their images under $e$ and thus identify $X$ with $E(X)$. The Hausdorff distance between $X$ and $E(X)$ is bounded by some constant $D$, so the action of $G$ on $E(X)$ is cobounded. For properness, let $Y \subset E(X)$ be bounded and let $Y' = \{x \in X : d(Y,x) \leqslant D\} \neq \varnothing$. Since $e$ is an isometric embedding, $Y'$ is bounded. If $g \in G$ has $gY \cap Y \neq \varnothing$, then pick $y \in Y$ with $gy \in Y$ and let $x \in X$ have $d(y,x) \leqslant D$. Then $d(gy,gx) \leqslant D$, so $gx$ is $D$–close to $Y$. That is, $gY' \cap Y' \neq \varnothing$, so, since $Y'$ is bounded and the action of $G$ on $X$ is proper, there are only finitely many such $g$. The final sentence follows from Corollary 3.8. □

Next we strengthen Corollary 3.7 in the case of HHGs. In particular, this applies to (extended) mapping class groups of finite-type surfaces.

**Corollary 3.11** *If $G$ is an HHG, then $G$ is semihyperbolic.*

**Proof** By Lemma 3.10, $G$ acts properly coboundedly on an injective space $E$. Every orbit map $G \to E$ is a $G$–equivariant quasi-isometry. By [60, Proposition 3.8], $E$ has a $G$–invariant, bounded, geodesic bicombing in the sense of [4]. As the action of $G$ on itself is free, it is semihyperbolic by [4, Theorem 4.1]. □

**Corollary 3.12** *If $G$ is an HHG, then $G$ has finitely many conjugacy classes of finite subgroups.*

**Proof** By Lemma 3.10, $G$ acts properly coboundedly on an injective space $E$. Let $x \in E$ and let $r$ be a constant such that $G \cdot x$ is $r$–coarsely dense in $E$. Let $F$ be a finite subgroup of $G$. By [60, Proposition 1.2], there is a point $z \in E$ that is fixed by $F$, and hence $F$ fixes the ball $B(z,r)$ in $E$, which contains a point of $G \cdot x$. It follows that a conjugate of $F$ fixes a point in $B(x,r)$, and we are done by properness of the action. □

## 3.3 Packing subgroups

Here we describe the application to bounded packing mentioned in the introduction. Following Hruska and Wise [53], we say that a finite collection $\mathcal{H}$ of subgroups of a discrete group $G$ has *bounded packing in* $G$ if for each $N$ there is a constant $r$ such that, for any collection of $N$ distinct cosets of elements of $\mathcal{H}$, at least two are separated by a distance of at least $r$ (with respect to some left-invariant, proper distance). If $\mathcal{H}$ consists of a single subgroup $H$, then we say that $H$ has bounded packing in $G$.

**Corollary 3.13** *If $\mathcal{H}$ is a finite collection of hierarchically quasiconvex subgroups of a group $G$ that is an HHS, then $\mathcal{H}$ has bounded packing in $G$.*

**Proof** By Theorem 3.5, any finite collection of cosets of elements of $\mathcal{H}$ that are pairwise $r$–close must all come $R$–close to a single point $x \in G$. In other words, they all intersect the $R$–ball about $x$. Since distinct cosets of a given subgroup are disjoint and balls in $G$ are finite, this bounds the size of the collection of cosets. □

In the case of quasiconvex subgroups of hyperbolic groups, one can use Theorem 3.4 in place of Theorem 3.5 in this argument to provide a new, simpler proof of bounded packing. This type of argument is also implicit in [47, Remark 4.4 and Corollary 4.5], though the coarse Helly property for quasiconvex subgroups of hyperbolic groups is established in a much less efficient way there.

Previous proofs of this result work by induction on the height of subgroups. However, this line of reasoning does not generalise outside the setting of strict negative curvature; indeed, no subgroup of a flat can ever have finite height. Moreover, Theorems 3.4 and 3.5 are purely geometric: there is no group action involved. It therefore seems that the most natural way to establish bounded packing for quasiconvex subgroups of hyperbolic groups is via the Chepoi–Dragan–Vaxès theorem as described above.

If a group $G$ has a codimension-1 subgroup $H$, then Sageev's construction yields an action of $G$ on a CAT(0) cube complex, and, if the conjugates of $H$ satisfy the coarse Helly property, then it follows that the action of $G$ on the CAT(0) cube complex is cocompact [72]. This raises the following question:

**Question** *Does the mapping class group have property* $\mathrm{FW}_\infty$*, ie does any action of the mapping class group on a finite-dimensional CAT*$(0)$ *cube complex have a fixed point*?

Note that property $\mathrm{FW}_\infty$ is intermediate between having no virtual surjection onto $\mathbb{Z}$ and Kazhdan's property (T). There are known restrictions on what an action of the mapping class group on a CAT(0) cube complex could look like. Indeed, the mapping class group of a surface of genus at least three does not admit a properly discontinuous action by semisimple isometries on a complete CAT(0) space [58; 24; 22], nor, more specifically, does it act properly on a CAT(0) cube complex (even an infinite-dimensional one) [44].

More generally, in relationship with property (T) and the Haagerup property, the existence of nontrivial actions of the mapping class group on various generalisations of CAT(0) cube complexes remains mysterious, for example median spaces, Hilbert spaces, CAT(0) spaces, and $L^p$ spaces. The coarse version of the Helly property established here may prove useful in the study of such actions.

# 4 Strong shortcut property

In this section we will prove that coarsely injective spaces of uniformly bounded geometry are strongly shortcut. Recall that a metric space has *uniformly bounded* *geometry* if, for any $r > 0$, there exists a uniform $N(r) \in \mathbb{N}$ such that every ball of radius $r$ contains at most $N(r)$ points.

A *Riemannian circle* $S$ is $S^1$ endowed with a geodesic metric of some length $|S|$. A roughly geodesic metric space $(X, \sigma)$ is *strongly shortcut* if there exists $K > 1$ such that for any $C > 0$ there is a bound on the lengths $|S|$ of $(K, C)$–quasi-isometric embeddings $S \to X$ of Riemannian circles $S$ in $(X, \sigma)$ [51]. A group $G$ is *strongly shortcut* if it acts properly and coboundedly on a strongly shortcut metric space [52; 51].

We will now give a brief description of the injective hull construction of Isbell [57], which was later rediscovered by Dress [36] and Chrobak and Larmore [32]. For a nice discussion on this construction, see Lang [60]. Let $(X, \sigma)$ be a metric space. A *radius function* on $X$ is a function $f : X \to \mathbb{R}_{\geq 0}$ for which

$$\sigma(x, y) \leq f(x) + f(y)$$

for every $x, y \in X$. A radius function $f : A \to \mathbb{R}_{\geq 0}$ on any subspace of $A \subseteq X$ is called a *partial radius function* on $X$. If $f, g\colon X \to \mathbb{R}_{\geq 0}$ are two radius functions, then $f$ *dominates* $g$ if $f(x) \geq g(x)$ for all $x \in X$. A radius function $f : X \to \mathbb{R}_{\geq 0}$ is minimal if the only radius function it dominates is itself.

If $f : A \to \mathbb{R}_{\geq 0}$ is a partial radius function on $X$, then there exists a minimal radius function $g\colon X \to \mathbb{R}_{\geq 0}$ such that $g|_A$ is dominated by $f$. For any $x \in X$, the function $\sigma(\,\cdot\,, x)$ is a minimal radius function. If $f, g\colon X \to \mathbb{R}_{\geq 0}$ are two minimal radius functions, then

$$|f - g|_\infty = \sup_{x \in X} |f(x) - g(x)|$$

is finite. The set of minimal radius functions on $X$, with metric given by $d_{E(X)}(f, g) = |f - g|_\infty$, is the *injective hull* $E(X)$ of $X$. The isometric embedding $e\colon X \hookrightarrow E(X)$ sends $x \in X$ to the minimal radius function $e(x)\colon y \mapsto \sigma(x, y)$ and, for any $x \in X$ and $f \in E(X)$, we have $d_{E(X)}(e(x), f) = f(x)$.

**Lemma 4.1** *Let $(X, \sigma)$ be a metric space. Let $g\colon X \to \mathbb{R}_{\geq 0}$ be a minimal radius function, let $\bar{f}\colon X \to \mathbb{R}_{\geq 0}$ be a radius function and let $f : X \to \mathbb{R}_{\geq 0}$ be any minimal radius function dominated by $\bar{f}$. Then $|g - f|_\infty \leq |g - \bar{f}|_\infty$.*

**Proof** Let $y \in X$. Then $f(y) \leq \bar{f}(y) \leq g(y) + |g - \bar{f}|_\infty$ and so $f(y) - g(y) \leq |g - \bar{f}|_\infty$. It remains to prove that $g(y) - f(y) \leq |g - \bar{f}|_\infty$. By minimality of $g$, for any $\epsilon > 0$, there exists $z \in X$ for which $g(y) + g(z) < \sigma(y,z) + \epsilon$. Then, since $f$ is a radius function dominated by $\bar{f}$,

$$f(y) \ge \sigma(y,z) - f(z) \ge \sigma(y,z) - \bar{f}(z) \ge \sigma(y,z) - g(z) - |g - \bar{f}|_\infty > g(y) - \epsilon - |g - \bar{f}|_\infty$$

and so $g(y) - f(y) < |g - \bar{f}|_\infty + \epsilon$, which completes the proof since we chose $\epsilon > 0$ arbitrarily. □

**Theorem 4.2** *Let $(X,\sigma)$ be a coarsely injective metric space. If $(X,\sigma)$ has uniformly bounded geometry, then $(X,\sigma)$ is strongly shortcut.*

**Proof** In order to prove this theorem, we will show that, for some uniform radius $r$, a $(K,C)$–quasi-isometric embedding of a Riemannian circle $S \to X$ implies the existence of a "centre" point $x$ such that the cardinality of the ball $B(x,r)$ is bounded below by an expression that tends to infinity as $K$ approaches 1 and $|S|$ approaches infinity. If $X$ is not strongly shortcut, then, for any $K > 1$, it will admit $(K, C_K)$–quasi-isometric embeddings of arbitrarily long Riemannian circles, so that we would then contradict the uniformly bounded geometry assumption.

Let $X \to E(X)$ be the embedding of $(X,\sigma)$ into its injective hull and view this embedding as an inclusion of a subspace. By Proposition 1.1, the subspace $X$ is $\delta$–coarsely dense in $E(X)$ for some $\delta > 0$. So there is a retraction $r \colon E(X) \to X$ such that $r$ is a $(1,2\delta)$–quasi-isometry.

Let $\phi \colon S \to X$ be a $(K,C)$–quasi-isometric embedding of a Riemannian circle. Let $f'' \colon \phi(S) \to \mathbb{R}_{\ge 0}$ be the constant function taking the value $K \cdot \frac{1}{4}|S| + C$. Then $f''$ is a radius function on $\phi(S) \subset X$. Let $f' \colon \phi(S) \to \mathbb{R}_{\ge 0}$ be a minimal radius function on $\phi(S)$ dominated by $f''$. Then, for each $x \in \phi(S)$ and each $\epsilon > 0$, there exists a $y \in \phi(S)$ for which $f'(x) + f'(y) < \sigma(x,y) + \epsilon$. Since $f'$ is a partial radius function on $X$, we can let $f \colon X \to \mathbb{R}_{\ge 0}$ be a minimal radius function on $X$ dominated by $f'$. Then $f$ is a point of $E(X)$, by definition of $E(X)$. Moreover, if $x \in \phi(S)$, then $d_{E(X)}(f,x) = f(x) \le f'(x) \le f''(x) = K \cdot \frac{1}{4}|S| + C$, so, if $\bar{s}$ is the antipode in $S$ of any element of $s \in \phi^{-1}(x)$, then

$$d_{E(X)}(x,\phi(\bar{s})) = d_{E(X)}(\phi(s),\phi(\bar{s})) \ge \frac{1}{K} d_S(s,\bar{s}) - C = \frac{|S|}{2K} - C$$

and

$$\begin{aligned} d_{E(X)}(x,\phi(\bar{s})) &\le d_{E(X)}(x,f) + d_{E(X)}(f,\phi(\bar{s})) = f(x) + f(\phi(\bar{s})) \\ &\le f(x) + K \cdot \tfrac{1}{4}|S| + C, \end{aligned}$$

so that $f(x) \geq |S|/2K - K \cdot \frac{1}{4}|S| - 2C = ((2-K^2)/4K)|S| - 2C$. Thus, we have shown that

$$\frac{2-K^2}{4K}|S| - 2C \leq f(x) \leq K \cdot \tfrac{1}{4}|S| + C$$

for any $x \in \phi(S)$.

For $x, y \in X$, let $\ell_{x,y} = f(x) + f(y) - \sigma(x, y)$. Since $f$ is dominated by $f'$, and $f'$ is a minimal radius function on $\phi(S)$, for each $x \in \phi(S)$ and each $\epsilon > 0$ there exists $y \in \phi(S)$ such that $\ell_{x,y} < \epsilon$. Moreover, for $a, b \in S$,

$$\begin{aligned}\frac{2-K^2}{2K}|S| - 4C &\leq f(\phi(a)) + f(\phi(b)) \\ &= \sigma(\phi(a), \phi(b)) + \ell_{\phi(a),\phi(b)} \\ &\leq K d_S(a,b) + C + \ell_{\phi(a),\phi(b)}\end{aligned}$$

and so

$$d_S(a,b) \geq \frac{2-K^2}{2K^2}|S| - \frac{\ell_{\phi(a),\phi(b)} + 5C}{K}.$$

**Claim** *Let $x \in \phi(S)$. There exists a sequence of minimal radius functions $(f_x^k : X \to \mathbb{R}_{\geq 0})_k$, where $k$ ranges in $\{0, 1, \ldots, M_x\}$, such that $M_x = \lfloor f(x)/\delta \rfloor$ and the following properties hold for all $k$, $k'$ and $y$:*

(1) $f_x^0 = f$.

(2) $d_{E(X)}(f_x^k, f_x^{k'}) = \delta|k - k'|$.

(3) $f(y) + k\delta - \ell_{x,y} \leq f_x^k(y) \leq f(y) + \max\{0, k\delta - \ell_{x,y}\}$.

**Proof** We construct the $(f_x^k)_k$ by induction on $k$. By property (1), we must start with $f_x^0 = f$. Assuming we have $f_x^{k-1}$, we will begin by defining a radius function $\bar{f}_x^k$. Set $\bar{f}_x^k(x) = f_x^{k-1}(x) - \delta$. By minimality of $f_x^{k-1}$, there exists $y \in X$ for which the inequality

$$f_x^{k-1}(y) + f_x^{k-1}(x) - \delta < \sigma(x, y) \tag{2}$$

holds. Indeed, if no such $y$ existed then

$$y \mapsto \begin{cases} f_x^{k-1}(y) & \text{if } y \neq x, \\ f_x^{k-1}(y) - \frac{1}{2}\delta & \text{if } y = x, \end{cases}$$

would be a radius function that is dominated by but not equal to $f_x^{k-1}$ and this would contradict minimality of $f_x^{k-1}$. Set $\bar{f}_x^k(x) = f_x^{k-1}(x) - \delta$. For all $y \in X \setminus \{x\}$ satisfying (2), set $\bar{f}_x^k(y) = \sigma(x, y) - f_x^{k-1}(x) + \delta$. For all other $y \in X \setminus \{x\}$, set

$\bar{f}_x^k(y) = f_x^{k-1}(y)$. Then, except for at $y = x$, we have $\bar{f}_x^k(y) \geq f_x^{k-1}(y)$. Thus, to check that $\bar{f}_x^k$ is a radius function, we need only verify that $\bar{f}_x^k(x) + \bar{f}_x^k(y) \geq \sigma(x, y)$ for any $y \in X$. When $y = x$, the inequality $\bar{f}_x^k(x) + \bar{f}_x^k(y) \geq \sigma(x, y)$ is equivalent to $f_x^{k-1}(x) \geq \delta$, which holds by the inductive application of property (2) and the triangle inequality. When $y$ satisfies (2), the inequality $\bar{f}_x^k(x) + \bar{f}_x^k(y) \geq \sigma(x, y)$ is equivalent to $f_x^{k-1}(x) - \delta + \sigma(x, y) - f_x^{k-1}(x) + \delta \geq \sigma(x, y)$, which holds with equality. Finally, when $y$ does not satisfy (2), $\bar{f}_x^k(x) + \bar{f}_x^k(y) = f_x^{k-1}(x) - \delta + f_x^{k-1}(y) \geq f_x^{k-1}(x) - \delta + \sigma(x, y) - f_x^{k-1}(x) + \delta = \sigma(x, y)$. Thus, $\bar{f}_x^k$ is a radius function. Define $f_x^k$ as any minimal radius function that is dominated by $\bar{f}_x^k$.

Since $\bar{f}_x^k(y) = \sigma(x, y) - f_x^{k-1}(x) + \delta = \sigma(x, y) - \bar{f}_x^k(x)$ for some $y \in X$, we must have $f_x^k(x) = \bar{f}_x^k(x) = f_x^{k-1}(x) - \delta$. Thus, $|f_x^{k-1} - f_x^k|_\infty \geq \delta$ and

$$d_{E(X)}(f_x^{M_x}, x) = f_x^{M_x}(x) = f(x) - M_x\delta = f(x) - \left\lfloor \frac{f(x)}{\delta} \right\rfloor \delta < \delta,$$

so $d_{E(X)}(f_x^{M_x}, x) < \delta$. On the other hand, by Lemma 4.1, $|f_x^{k-1} - f_x^k|_\infty \leq |f_x^{k-1} - \bar{f}_x^k|_\infty \leq \delta$ and so $d_{E(X)}(f_x^{k-1}, f_x^k) = |f_x^{k-1} - f_x^k|_\infty = \delta$. Therefore,

$$\begin{aligned} d_{E(X)}(f, x) &= f(x) \\ &= M_x\delta + f(x) - M_x\delta \\ &= M_x\delta + d_{E(X)}(f_x^{M_x}, x) \\ &= \sum_{k=1}^{M_x} d_{E(X)}(f_x^{k-1}, f_x^k) + d_{E(X)}(f_x^{M_x}, x), \end{aligned}$$

where $f_x^0 = f$. Then, by the triangle inequality, property (2) is satisfied.

To verify property (3), let $y \in X$. We have

$$f(y) + k\delta - \ell_{x,y} = \sigma(x, y) + k\delta - f(x) = \sigma(x, y) - f_x^k(x) \leq f_x^k(y),$$

so the lower bound holds. The upper bound on $f_x^k(y)$ given by property (3) is $R_k = f(y) + \max\{0, k\delta - \ell_{x,y}\}$. Suppose property (3) doesn't hold and let $k$ be the least integer for which $f_x^k(y) > R_k$. Then $k > 0$ and $k$ must satisfy $f_x^k(y) - f_x^{k-1}(y) > R_k - R_{k-1} \geq 0$. By the construction of $f_x^k$, the fact that $f_x^k(y) > f_x^{k-1}(y)$ implies that $f_x^{k-1}(y) + f_x^{k-1}(x) - \delta < \sigma(x, y)$ and that $\bar{f}_x^k(y) = \sigma(x, y) - f_x^{k-1}(x) + \delta = \sigma(x, y) - \bar{f}_x^k(x)$. Then we must have

$$f_x^k(y) = \bar{f}_x^k(y) = \sigma(x, y) - \bar{f}_x^k(x) = \sigma(x, y) - f_x^k(x) = f(y) + k\delta - \ell_{x,y} \leq R_k,$$

which contradicts $f_x^k(y) > R_k$. Thus, we have verified property (3). $\triangleleft$

We will now use the sequence $(f_x^k)_k$ of minimal radius functions to prove the theorem. Assume that $a, a' \in S$ satisfy $d_S(a,a') \geq (2(K^2-1)/K^2)|S| + (4\delta+10C)/K$. Such $a$ and $a'$ exist when $K$ is close enough to 1. Take $b \in S$ for which $\ell_{\phi(a),\phi(b)} < \delta$. Then $d_S(a,a') + d_S(a',b) + d_S(b,a) \leq |S|$, so

$$\begin{aligned} d_S(a',b) &\leq |S| - d_S(a,b) - d_S(a,a') \\ &\leq |S| - \frac{2-K^2}{2K^2}|S| + \frac{\ell_{\phi(a),\phi(b)} + 5C}{K} - d_S(a,a') \\ &< |S| - \frac{2-K^2}{2K^2}|S| + \frac{5C+\delta}{K} - d_S(a,a') \\ &\leq |S| - \frac{2-K^2}{2K^2}|S| + \frac{5C+\delta}{K} - \frac{2(K^2-1)}{K^2}|S| - \frac{4\delta+10C}{K} \\ &= \frac{2-K^2}{2K^2}|S| - \frac{3\delta+5C}{K} \end{aligned}$$

and so

$$\frac{2-K^2}{2K^2}|S| - \frac{\ell_{\phi(a'),\phi(b)} + 5C}{K} \leq d_S(a',b) < \frac{2-K^2}{2K^2}|S| - \frac{3\delta+5C}{K},$$

which implies $\ell_{\phi(a'),\phi(b)} > 3\delta$. So

$$\begin{aligned} f^3_{\phi(a')}(\phi(b)) &\leq f(\phi(b)) + \max\{0, 3\delta - \ell_{\phi(a'),\phi(b)}\} \\ &= f(\phi(b)) \\ &\leq f^3_{\phi(a)}(\phi(b)) - 3\delta + \ell_{\phi(a),\phi(b)} \\ &< f^3_{\phi(a)}(\phi(b)) - 2\delta, \end{aligned}$$

where the inequalities are applications of property (3). Thus,

$$d_{E(X)}(f^3_{\phi(a')}, f^3_{\phi(a)}) > 2\delta$$

and so $r(f^3_{\phi(a')})$ and $r(f^3_{\phi(a)})$ are distinct elements of the metric ball $B(r(f), 5\delta)$ of radius $5\delta$ centred at $r(f)$ in $X$. So, if $\{a_i\}_{i=1}^N \subset S$ subdivide $S$ into segments of length at least $(2(K^2-1)/K^2)|S| + (4\delta+10C)/K$, then $B(r(f), 5\delta)$ contains at least $N$ points. Subdividing $S$ evenly, we can achieve $N = \lfloor (2(K^2-1)/K^2 + (4\delta+10C)/K|S|)^{-1} \rfloor$. So we have shown that, if $X$ admits a $(K,C)$–quasi-isometric embedding of a Riemannian circle $S$ and $K$ is close enough to 1, then, for some $x \in X$, we have $|B(x, 5\delta)| \geq \lfloor (2(K^2-1)/K^2 + (4\delta+10C)/K|S|)^{-1} \rfloor$.

To complete the proof, suppose $X$ is not strongly shortcut. Then, for each $K > 1$, there exists $C_K > 0$ and a sequence $(\phi_n \colon S_n \to X)_n$ of $(K, C_K)$–quasi-isometric embeddings of Riemannian circles where $|S_n| \geq n$. The argument above shows that, for each small enough $K > 1$ and each $n \in \mathbb{N}$, there exists $x_{K,n} \in X$ satisfying $|B(x_{K,n}, 5\delta)| \geq \lfloor (2(K^2-1)/K^2 + (4\delta + 10C_k)/K|S_n|)^{-1} \rfloor$. The expression $(2(K^2-1)/K^2+(4\delta+10C_K)/K|S_n|)^{-1}$ tends to $K^2/2(K^2-1)$ as $n$ tends to infinity, so, if $n_K \in \mathbb{N}$ is large enough, then $|B(x_{K,n_K}, 5\delta)| \geq K^2/2(K^2-1) - 1$. But this contradicts the uniform bounded geometry assumption on $X$ since $K^2/2(K^2-1)$ tends to infinity as $K$ tends to 1. □

# References

[1] **C Abbott**, **J Behrstock**, *Conjugator lengths in hierarchically hyperbolic groups*, preprint (2018) arXiv 1808.09604

[2] **C Abbott**, **J Behrstock**, **M G Durham**, *Largest acylindrical actions and stability in hierarchically hyperbolic groups*, Trans. Amer. Math. Soc. Ser. B 8 (2021) 66–104 MR Zbl

[3] **C Abbott**, **T Ng**, **D Spriano**, **R Gupta**, **H Petyt**, *Hierarchically hyperbolic groups and uniform exponential growth*, preprint (2019) arXiv 1909.00439

[4] **J M Alonso**, **M R Bridson**, *Semihyperbolic groups*, Proc. London Math. Soc. 70 (1995) 56–114 MR Zbl

[5] **Y Antolín**, **M Mj**, **A Sisto**, **S J Taylor**, *Intersection properties of stable subgroups and bounded cohomology*, Indiana Univ. Math. J. 68 (2019) 179–199 MR Zbl

[6] **N Aronszajn**, **P Panitchpakdi**, *Extension of uniformly continuous transformations and hyperconvex metric spaces*, Pacific J. Math. 6 (1956) 405–439 MR Zbl

[7] **H-J Bandelt**, **M van de Vel**, *Superextensions and the depth of median graphs*, J. Combin. Theory Ser. A 57 (1991) 187–202 MR Zbl

[8] **J Behrstock**, **C Druţu**, *Divergence, thick groups, and short conjugators*, Illinois J. Math. 58 (2014) 939–980 MR Zbl

[9] **J Behrstock**, **M Hagen**, **A Martin**, **A Sisto**, *A combinatorial take on hierarchical hyperbolicity and applications to quotients of mapping class groups*, preprint (2020) arXiv 2005.00567

[10] **J Behrstock**, **M F Hagen**, **A Sisto**, *Asymptotic dimension and small-cancellation for hierarchically hyperbolic spaces and groups*, Proc. Lond. Math. Soc. 114 (2017) 890–926 MR

[11] **J Behrstock**, **M F Hagen**, **A Sisto**, *Hierarchically hyperbolic spaces, I: Curve complexes for cubical groups*, Geom. Topol. 21 (2017) 1731–1804 MR

[12] **J Behrstock**, **M F Hagen**, **A Sisto**, *Quasiflats in hierarchically hyperbolic spaces*, preprint (2017) arXiv 1704.04271

[13] **J Behrstock**, **M Hagen**, **A Sisto**, *Hierarchically hyperbolic spaces, II: Combination theorems and the distance formula*, Pacific J. Math. 299 (2019) 257–338 MR Zbl

[14] **J A Behrstock**, **Y N Minsky**, *Centroids and the rapid decay property in mapping class groups*, J. Lond. Math. Soc. 84 (2011) 765–784 MR Zbl

[15] **F Berlai**, **B Robbio**, *A refined combination theorem for hierarchically hyperbolic groups*, Groups Geom. Dyn. 14 (2020) 1127–1203 MR Zbl

[16] **D Berlyne**, **J Russell**, *Hierarchical hyperbolicity of graph products*, Groups Geom. Dyn. 16 (2022) 523–580 MR Zbl

[17] **B H Bowditch**, *Coarse median spaces and groups*, Pacific J. Math. 261 (2013) 53–93 MR Zbl

[18] **B H Bowditch**, *Convex hulls in coarse median spaces*, preprint (2019) Available at `http://www.warwick.ac.uk/~masgak/papers/hulls-cms.pdf`

[19] **B H Bowditch**, *Quasiflats in coarse median spaces*, preprint (2019) Available at `http://www.warwick.ac.uk/~masgak/papers/quasiflats.pdf`

[20] **B H Bowditch**, *Median and injective metric spaces*, Math. Proc. Cambridge Philos. Soc. 168 (2020) 43–55 MR Zbl

[21] **M R Bridson**, *Finiteness properties for subgroups of* $\mathrm{GL}(n,\mathbb{Z})$, Math. Ann. 317 (2000) 629–633 MR Zbl

[22] **M R Bridson**, *Semisimple actions of mapping class groups on* $\mathrm{CAT}(0)$ *spaces*, from "Geometry of Riemann surfaces" (F P Gardiner, G González-Diez, C Kourouniotis, editors), London Math. Soc. Lecture Note Ser. 368, Cambridge Univ. Press (2010) 1–14 MR Zbl

[23] **M R Bridson**, *Semihyperbolicity*, from "Beyond hyperbolicity" (M Hagen, R Webb, H Wilton, editors), London Math. Soc. Lecture Note Ser. 454, Cambridge Univ. Press (2019) 25–64 MR Zbl

[24] **M R Bridson**, **A Haefliger**, *Metric spaces of non-positive curvature*, Grundl. Math. Wissen. 319, Springer (1999) MR Zbl

[25] **P-E Caprace**, **M Sageev**, *Rank rigidity for* $\mathrm{CAT}(0)$ *cube complexes*, Geom. Funct. Anal. 21 (2011) 851–891 MR Zbl

[26] **J Chalopin**, **V Chepoi**, **A Genevois**, **H Hirai**, **D Osajda**, *Helly groups*, preprint (2020) arXiv 2002.06895

[27] **V Chepoi**, *Separation of two convex sets in convexity structures*, J. Geom. 50 (1994) 30–51 MR Zbl

[28] **V Chepoi**, *Graphs of some* $\mathrm{CAT}(0)$ *complexes*, Adv. in Appl. Math. 24 (2000) 125–179 MR Zbl

[29] **V Chepoi**, **F F Dragan**, **Y Vaxès**, *Core congestion is inherent in hyperbolic networks*, from "Proceedings of the twenty-eighth annual ACM–SIAM symposium on discrete algorithms" (P N Klein, editor), SIAM, Philadelphia, PA (2017) 2264–2279 MR Zbl

[30] **V Chepoi**, **B Estellon**, *Packing and covering δ–hyperbolic spaces by balls*, from "Approximation, randomization, and combinatorial optimization: algorithms and techniques" (M Charikar, K Jansen, O Reingold, J D P Rolim, editors), Lecture Notes in Computer Science 4627, Springer (2007) 59–73 Zbl

[31] **M Chesser**, *Stable subgroups of the genus* 2 *handlebody group*, Algebr. Geom. Topol. 22 (2022) 919–971 MR

[32] **M Chrobak**, **L L Larmore**, *Generosity helps or an* 11*–competitive algorithm for three servers*, J. Algorithms 16 (1994) 234–263 MR Zbl

[33] **Y Cornulier**, **P de la Harpe**, *Metric geometry of locally compact groups*, EMS Tracts in Mathematics 25, Eur. Math. Soc., Zürich (2016) MR Zbl

[34] **D Descombes**, **U Lang**, *Convex geodesic bicombings and hyperbolicity*, Geom. Dedicata 177 (2015) 367–384 MR Zbl

[35] **S Dowdall**, **M G Durham**, **C J Leininger**, **A Sisto**, *Extensions of Veech groups*, *I*: *A hyperbolic action*, preprint (2020) arXiv 2006.16425

[36] **A W M Dress**, *Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces*, Adv. in Math. 53 (1984) 321–402 MR Zbl

[37] **M G Durham**, **M F Hagen**, **A Sisto**, *Boundaries and automorphisms of hierarchically hyperbolic spaces*, Geom. Topol. 21 (2017) 3659–3758 MR Zbl

[38] **M G Durham**, **Y N Minsky**, **A Sisto**, *Stable cubulations*, *bicombings*, *and barycenters*, preprint (2020) arXiv 2009.13647

[39] **M G Durham**, **S J Taylor**, *Convex cocompactness and stability in mapping class groups*, Algebr. Geom. Topol. 15 (2015) 2839–2859 MR Zbl

[40] **A Engel**, **C Wulff**, *Coronas for properly combable spaces*, preprint (2017) arXiv 1711.06836

[41] **D B A Epstein**, **J W Cannon**, **D F Holt**, **S V F Levy**, **M S Paterson**, **W P Thurston**, *Word processing in groups*, Jones and Bartlett, Boston, MA (1992) MR Zbl

[42] **T Fukaya**, **S-i Oguni**, *A coarse Cartan–Hadamard theorem with application to the coarse Baum–Connes conjecture*, J. Topol. Anal. 12 (2020) 857–895 MR Zbl

[43] **A Genevois**, *Hyperbolicities in* CAT(0) *cube complexes*, Enseign. Math. 65 (2019) 33–100 MR Zbl

[44] **A Genevois**, *Median sets of isometries in CAT*(0) *cube complexes and some applications*, preprint (2019) arXiv 1902.04883

[45] **R Gitik**, **M Mitra**, **E Rips**, **M Sageev**, *Widths of subgroups*, Trans. Amer. Math. Soc. 350 (1998) 321–329 MR Zbl

[46] **T Haettel**, *Hyperbolic rigidity of higher rank lattices*, Ann. Sci. Éc. Norm. Supér. 53 (2020) 439–468 MR Zbl

[47] **M F Hagen**, **H Petyt**, *Projection complexes and quasimedian maps*, Algebr. Geom. Topol. 22 (2022) 3277–3304 MR Zbl

[48] **M F Hagen**, **T Susse**, *On hierarchical hyperbolicity of cubical groups*, Israel J. Math. 236 (2020) 45–89 MR Zbl

[49] **U Hamenstaedt**, *Geometry of the mapping class group*, *II*: *A biautomatic structure*, preprint (2009) arXiv 0912.0137

[50] **N Hoda**, *Crystallographic Helly groups*, preprint (2020) arXiv 2010.07407

[51] **N Hoda**, *Strongly shortcut spaces*, preprint (2020) arXiv 2010.07400

[52] **N Hoda**, *Shortcut graphs and groups*, Trans. Amer. Math. Soc. 375 (2022) 2417–2458 MR Zbl

[53] **G C Hruska**, **D T Wise**, *Packing subgroups in relatively hyperbolic groups*, Geom. Topol. 13 (2009) 1945–1988 MR Zbl

[54] **G C Hruska**, **D T Wise**, *Finiteness properties of cubulated groups*, Compos. Math. 150 (2014) 453–506 MR Zbl

[55] **J Huang**, **D Osajda**, *Helly meets Garside and Artin*, Invent. Math. 225 (2021) 395–426 MR Zbl

[56] **S Hughes**, **M Valiunas**, *Commensurating HNN-extensions*: *hierarchical hyperbolicity and biautomaticity*, preprint (2022) arXiv 2203.11996

[57] **J R Isbell**, *Six theorems about injective metric spaces*, Comment. Math. Helv. 39 (1964) 65–76 MR Zbl

[58] **M Kapovich**, **B Leeb**, *Actions of discrete groups on nonpositively curved spaces*, Math. Ann. 306 (1996) 341–352 MR Zbl

[59] **S P Kerckhoff**, *The Nielsen realization problem*, Ann. of Math. 117 (1983) 235–265 MR Zbl

[60] **U Lang**, *Injective hulls of certain discrete metric spaces and groups*, J. Topol. Anal. 5 (2013) 297–331 MR Zbl

[61] **H A Masur**, **Y N Minsky**, *Geometry of the complex of curves, II: Hierarchical structure*, Geom. Funct. Anal. 10 (2000) 902–974 MR Zbl

[62] **L Mosher**, *Mapping class groups are automatic*, Ann. of Math. 142 (1995) 303–384 MR Zbl

[63] **G A Niblo**, **L D Reeves**, *The geometry of cube complexes and the complexity of their fundamental groups*, Topology 37 (1998) 621–633 MR Zbl

[64] **G A Niblo**, **L D Reeves**, *Coxeter groups act on* CAT(0) *cube complexes*, J. Group Theory 6 (2003) 399–413 MR Zbl

[65] **G A Niblo**, **N Wright**, **J Zhang**, *A four point characterisation for coarse median spaces*, Groups Geom. Dyn. 13 (2019) 939–980 MR Zbl

[66] **H Petyt**, **D Spriano**, *Unbounded domains in hierarchically hyperbolic groups*, Groups Geom. Dyn. 17 (2023) 479–500 MR Zbl

[67] **P Plummer**, *Some graph manifold groups are not hierarchically hyperbolic groups*, preprint (2020)

[68] **B Robbio**, **D Spriano**, *Hierarchical hyperbolicity of hyperbolic-2–decomposable groups*, preprint (2020) arXiv 2007.13383

[69] **M Roller**, *Poc sets*, *median algebras and group actions*, Habilitationsschrift, Universität Regensburg (1998) arXiv 1607.07747

[70] **H Rubinstein**, **M Sageev**, *Intersection patterns of essential surfaces in* 3*–manifolds*, Topology 38 (1999) 1281–1291 MR Zbl

[71] **J Russell**, **D Spriano**, **H C Tran**, *Convexity in hierarchically hyperbolic spaces*, preprint (2018) arXiv 1809.09303

[72] **M Sageev**, *Codimension-*1 *subgroups and splittings of groups*, J. Algebra 189 (1997) 377–389 MR Zbl

[73] **A Sisto**, *Quasi-convexity of hyperbolically embedded subgroups*, Math. Z. 283 (2016) 649–658 MR Zbl

[74] **J Tao**, *Linearly bounded conjugator property for mapping class groups*, Geom. Funct. Anal. 23 (2013) 415–466 MR Zbl

[75] **H C Tran**, *On strongly quasiconvex subgroups*, Geom. Topol. 23 (2019) 1173–1235 MR Zbl

[76] **K M Vokes**, *Hierarchical hyperbolicity of graphs of multicurves*, Algebr. Geom. Topol. 22 (2022) 113–151 MR Zbl

*Institut Montpelliérain Alexander Grothendieck, CNRS, Université de Montpellier*
*Montpellier, France*

*DMA, École normale supérieure, Université PSL, CNRS*
*Paris, France*

Current address: *Department of Mathematics, Cornell University*
*Ithaca, NY, United States*

*School of Mathematics, University of Bristol*
*Bristol, United Kingdom*

Current address: *Mathematical Institute, University of Oxford*
*Oxford, United Kingdom*

thomas.haettel@umontpellier.fr, nima.hoda@mail.mcgill.ca,
petyt@maths.ox.ac.uk

Proposed: Urs Lang Received: 13 January 2021
Seconded: Mladen Bestvina, Anna Wienhard Revised: 24 June 2021

*Geometry & Topology* **27**:4 (2023) 1635–1655
DOI: 10.2140/gt.2023.27.1635
Published: 15 June 2023

# Classifying sufficiently connected PSC manifolds in 4 and 5 dimensions

OTIS CHODOSH
CHAO LI
YEVGENY LIOKUMOVICH

We show that if $N$ is a closed manifold of dimension $n = 4$ (resp. $n = 5$) with $\pi_2(N) = 0$ (resp. $\pi_2(N) = \pi_3(N) = 0$) that admits a metric of positive scalar curvature, then a finite cover $\widehat{N}$ of $N$ is homotopy equivalent to $S^n$ or connected sums of $S^{n-1} \times S^1$. Our approach combines recent advances in the study of positive scalar curvature with a novel argument of Alpert, Balitskiy and Guth.

Additionally, we prove a more general mapping version of this result. In particular, this implies that if $N$ is a closed manifold of dimensions 4 or 5, and $N$ admits a map of nonzero degree to a closed aspherical manifold, then $N$ does not admit any Riemannian metric with positive scalar curvature.

53C21

## Introduction

We are concerned here with the problem of classification of manifolds admitting positive scalar curvature (PSC). For closed (compact, no boundary) 2– and 3–manifolds, this problem is completely resolved; namely, the sphere and projective plane are the only closed surfaces admitting positive scalar curvature and a 3–manifold admits positive scalar curvature if and only if it has no aspherical factors in its prime decomposition. In particular, a 3–manifold admitting positive scalar curvature has a finite cover diffeomorphic to $S^3$ or to a connected sum of finitely many $S^2 \times S^1$.

The main result of this paper is the following partial generalization of this statement to dimensions $n = 4, 5$:

**Theorem 1** *Suppose that $N$ is a closed smooth $n$–manifold admitting a metric of positive scalar curvature and*

- $n = 4$ *and* $\pi_2(N) = 0$, *or*
- $n = 5$ *and* $\pi_2(N) = \pi_3(N) = 0$.

*Then a finite cover $\widehat{N}$ of $N$ is homotopy equivalent to $S^n$ or connected sums of $S^{n-1} \times S^1$.*

Chodosh and Li [7] and Gromov [14] showed that, if a closed $N^n$ is aspherical (ie $\pi_k(N) = 0$ for all $k \geq 2$) and $n = 4, 5$, then there is no Riemannian metric of positive scalar curvature on $N$. Theorem 1 can thus be seen as a refinement of this into a positive result.

**Remark** By Theorem 1.3 of Gadgil and Seshadri [12] (see also Freedman [11], Milnor [26] and Kreck and Lück [22]), we have that if $n = 4$ and $\widehat{N}$ is homotopy equivalent to $S^4$ or $S^3 \times S^1$, or if $n = 5$ (with no further restriction on the homotopy type), then homotopy equivalence in the conclusion to Theorem 1 can be upgraded to homeomorphism.

We also prove a more general "mapping" version of Theorem 1.

**Theorem 2** *Suppose that $N$ is a closed smooth $n$–manifold with a metric of positive scalar curvature and there exists a nonzero degree map $f\colon N \to X$, to a manifold $X$ satisfying*

- $n = 4$ *and* $\pi_2(X) = 0$, *or*
- $n = 5$ *and* $\pi_2(X) = \pi_3(X) = 0$.

*Then a finite cover $\widehat{X}$ of $X$ is homotopy equivalent to $S^n$ or connected sums of $S^{n-1} \times S^1$.*

We note that the following result immediately follows from Theorem 2:

**Corollary 3** *Let $n \in \{4, 5\}$, $X$ and $N$ be closed oriented manifolds of dimension $n$, and $X$ be aspherical. Suppose there exists a map $f\colon N \to X$ with $\deg f \neq 0$. Then $N$ does not admit any Riemannian metric of positive scalar curvature.*

Recall that it was previously shown in [7; 14] that closed aspherical (ie $\pi_k(N) = 0$ for all $k \geq 2$) $n$–manifolds do not admit PSC for $n = 4, 5$. In [14] a related statement was proven for manifolds admitting proper, distance-decreasing maps to uniformly contractible manifolds. In fact, Corollary 3 seems to have been asserted by Gromov [15, page 144–145], but the (relatively simple) lifting argument does not appear there.

## 0.1 Urysohn width bounds

Recall that a metric space $(X, d)$ has *Urysohn q–width* $\leq \Lambda$ if there is a $q$–dimensional simplicial complex $K$ and a continuous map $X \to K$ such that diam $f^{-1}(s) \leq \Lambda$ for all $s \in K$. As such, having finite Urysohn $q$–width implies that a manifold looks $\leq q$–dimensional in some macroscopic sense.

A well-known conjecture (see [15, page 63]) of Gromov posits that an $n$–manifold with scalar curvature $\geq 1$ has finite Urysohn $(n-2)$–width. Various forms of this conjecture are proven for $n = 3$ — see Gromov and Lawson [16], Katz [21], Marques and Neves [24] and Liokumovich and Maximo [23] — while the conjecture is largely open for $n \geq 4$ (some progress has been achieved by Bolotov and Dranishnikov [2; 3]).

A key component in the proof of Theorem 1 is the following result:

**Theorem 4** *For $(N^n, g)$ satisfying the hypothesis of Theorem 1, the universal cover $(\widetilde{N}, \tilde{g})$ has finite Urysohn 1–width.*

This follows by combining Corollary 7 and Proposition 8 below. A simple example where Theorem 1 applies is the product metric on $S^1 \times S^3$, whose universal cover is $\mathbb{R} \times S^3$, clearly of finite Urysohn 1–width. On the other hand, we note that the higher connectivity hypothesis in Theorem 4 is necessary: compare with $T^2 \times S^2$.

**Remark** Consider a metric $g_R$ on $S^3$ formed by capping off a cylinder $[-R, R] \times S^2(1)$ with hemispheres and smoothing out the resulting metric, so that the scalar curvature is $\geq 1$. The product metric $(S^1(1), g_S) \times (S^3, g_R)$ has scalar curvature $\geq 1$ but the universal cover has Urysohn 1–width $\sim R$. As such, the estimate in Theorem 4 cannot be made quantitative (essentially, the issue is that the universal cover converges to $\mathbb{R}^2 \times S^2(1)$, which has nontrivial $\pi_2$).

As we were finishing this paper, we discovered that, recently, Gromov has indicated a proof of the classification of PSC 3–manifolds [15, page 135] by using finiteness of the 1–Urysohn width of the universal cover. Our proof of Theorem 1 follows a similar strategy once Theorem 4 is proven.

## 0.2 Remarks on positive isotropic curvature

Theorem 1 has an interesting relationship to well-known conjectures of Gromov [13, Section 3(b)] and Schoen [29] concerning the topology of closed $n$–manifolds admitting a metric with positive isotropic curvature (PIC). Namely, they (respectively) conjecture that if a closed manifold has a PIC metric then the fundamental group is virtually free and a finite cover is diffeomorphic to either a sphere or connected sums of finitely many $S^1 \times S^{n-1}$.

There have been distinct approaches to such a question, relying on either minimal surface theory or Ricci flow. Using minimal surface theory, Micallef and Moore have shown that, if $M^n$ is a closed PIC manifold then $\pi_k(M) = 0$ for $k = 2, \dots, \left[\frac{1}{2}n\right]$ [25]. In particular, if $M$ is simply connected, then it is homeomorphic to a sphere. In related work, Fraser has proven that an $n$–manifold ($n \geq 5$) with PIC does not contain a subgroup isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$ [10].

On the other hand, using Ricci flow, Hamilton has classified 4–manifolds admitting PIC that do not contain nontrivial incompressible $(n-1)$–dimensional space forms [18]. This was extended to prove the Gromov–Schoen conjectures for $n = 4$ by Chen, Tang and Zhu [6]. In higher dimensions, Brendle and Schoen [5] and Nguyen [27] proved the PIC condition is preserved under the Ricci flow; this is an important ingredient in Brendle and Schoen's proof of the differentiable sphere theorem. Recently, Brendle has achieved a breakthrough in the study of the Ricci flow of PIC manifolds and has extended Hamilton's result to dimensions $n \geq 12$ [4]; as above, this result has been used to prove the Gromov–Schoen conjectures for $n \geq 12$ by Huang [20].

We note that, since PIC implies PSC, combining [25] with Theorem 1 yields an alternative proof of Gromov's conjecture (the fundamental group is virtually free) for $n = 4$ and proves a weak version of Schoen's conjecture for $n = 4$ (ie with homotopy equivalence replacing diffeomorphism). Furthermore, Theorem 1 implies that a PIC 5–manifold with $\pi_3(M) = 0$ satisfies Gromov's conjecture and the same weak version of Schoen's conjecture. It is an interesting question if a 5–manifold with PIC has $\pi_3(M) = 0$ (note that $\pi_2(M) = 0$ by [25]).

## Organization of the paper

In Section 1 we revisit the filling radius estimates from [7; 14]. In Section 2 we show that such estimates imply Theorem 4. Then, we complete the proof of Theorem 1 in Section 3. Finally, in Section 4 we prove Theorem 2.

## Acknowledgements

Chodosh was supported by a Sloan Fellowship, a Terman Fellowship and NSF grant DMS-2016403. Li was supported by NSF grant DMS-2005287. Liokumovich was supported by a NSERC Discovery grant and an NSERC Accelerator award. We are grateful to Hannah Alpert, Larry Guth, Ciprian Manolescu, Kasra Rafi and Boyu Zhang for various discussions related to this article. We would like to thank Georg Frenck for suggesting the statement of Theorem 2 to us as well as the referee for several helpful suggestions.

# 1 Filling estimates

In [7; 14], it was shown that a closed aspherical $n$–manifold does not admit positive scalar curvature for $n = 4, 5$ by combining a linking argument with a filling radius inequality in the presence of positive scalar curvature. In this section we observe that this filling radius inequality carries over to the setting considered here.

We begin by summarizing the results contained in [7] that will be needed in this paper.

**Theorem 5** *Consider $(N^n, g)$ a closed Riemannian $n$–manifold with scalar curvature $R \geq 1$. Fix a Riemannian cover $(\widehat{N}, \hat{g})$.*

(1) *Suppose that $n = 4$. There is a universal constant $L_0 > 0$ with the following property. Consider a closed embedded 2–dimensional submanifold $\widehat{\Sigma}_2 \subset \widehat{N}$ with $[\widehat{\Sigma}_2] = 0 \in H_2(\widehat{N}; \mathbb{Z})$. Then there is a 3–chain $\widehat{\Sigma}'_3 \subset B_{L_0}(\widehat{\Sigma}_2)$ and a closed embedded 2–dimensional submanifold $\widehat{\Sigma}'_2$ with*

$$\partial \widehat{\Sigma}'_3 = \widehat{\Sigma}_2 - \widehat{\Sigma}'_2$$

*as chains such that, for every connected component $S$ of $\widehat{\Sigma}'_2$, the extrinsic diameter of $S$ satisfies* $\operatorname{diam}(S) \leq L_0$.

(2) *Suppose that $n = 5$. There is a universal constant $L_0 > 0$ with the following property. Consider a closed embedded 3–dimensional submanifold $\widehat{\Sigma}_3 \subset \widehat{N}$ with $[\widehat{\Sigma}_3] = 0 \in H_3(\widehat{N}; \mathbb{Z})$. Then there is a 4–chain $\widehat{\Sigma}'_4 \subset B_{L_0}(\widehat{\Sigma}_3)$ and a closed embedded 3–dimensional submanifold $\widehat{\Sigma}'_3$ with*

$$\partial \Sigma'_4 = \Sigma_3 - \Sigma'_3$$

*as chains as well as 3–chains $\widehat{U}_1, \dots, \widehat{U}_m$ with* $\operatorname{diam}(\widehat{U}_j) \leq L_0$ *and 2–cycles*

$$\{\widehat{\Gamma}_j^l : j = 1, \dots, m, l = 1, \dots, k(j)\}$$

*with* $\operatorname{diam}(\widehat{\Gamma}_j^l) \le L_0$ *such that*

$$\widehat{\Sigma}_3 = \sum_{j=1}^{m} \widehat{U}_j \qquad \text{and} \qquad \partial \widehat{U}_j = \sum_{l=1}^{k(j)} \widehat{\Gamma}_j^l \quad \text{for each } j = 1, \dots, m,$$

*where both equalities hold as chains (not just in homology). Finally, there is an integer* $q$ *and a function*

$$u \colon \{(j,l) : j = 1, \dots, m,\ l = 1, \dots, k(j)\} \to \{1, \dots, q\}$$

*such that, for* $r \in \{1, \dots, q\}$*, we have*

$$\operatorname{diam}\biggl( \bigcup_{(j,l) \in u^{-1}(r)} \widehat{\Gamma}_j^l \biggr) \le L_0$$

*and, moreover,*

$$\sum_{(j,l) \in u^{-1}(r)} \widehat{\Gamma}_j^l = 0$$

*as 2–chains for* $r \in \{1, \dots, q\}$*.*

**Proof** When $n = 4$, one can solve Plateau's problem to find $\widehat{\Sigma}_3$ minimizing area with $\partial \widehat{\Sigma}_3 = \widehat{\Sigma}_2$. Applying the "$\mu$–bubble technique" (see [7, Section 3]), we can find $\widehat{\Sigma}_2' \subset \widehat{\Sigma}_3$ with $d_{\widehat{\Sigma}_3}(\widehat{\Sigma}_2', \widehat{\Sigma}_2) \le L_0$ and such that $\widehat{\Sigma}_2' \subset \widehat{\Sigma}_3$ is a "stable $\mu$–bubble" in the sense of [7, Lemma 14]. By [7, Lemma 16], the intrinsic diameter of each component is $\le L_0$ (taking $L_0$ larger if necessary). This proves the assertion (since extrinsic distances are bounded by the intrinsic distances).

Similarly, when $n = 5$, we can solve Plateau's problem to find $\widehat{\Sigma}_4$ minimizing area with $\partial \widehat{\Sigma}_4 = \widehat{\Sigma}_3$. As before, we can find a "stable $\mu$–bubble" $\widehat{\Sigma}_3'$ with $d_{\widehat{\Sigma}_4}(\widehat{\Sigma}_3', \widehat{\Sigma}_3) \le L_0$. Finally, the construction of the $\widehat{U}_j$ and $\widehat{\Gamma}_j^k$ follows from the "slice-and-dice" procedure from [7, Sections 6.3–6.4].

Note that the last conclusion (ie that $\sum_{(j,l) \in u^{-1}(r)} \widehat{\Gamma}_j^l = 0$) was stated slightly differently in [7]. To be precise, it was proven that the cycles $\sum_{(j,l) \in u^{-1}(r)} \widehat{\Gamma}_j^l$ are disjoint for distinct $r$ (see [7, Section 6.4]). Now, by using $\sum_{r=1}^{q} \sum_{(j,l) \in u^{-1}(r)} \widehat{\Gamma}_j^l = \partial\bigl(\sum_{j=1}^{m} U_j\bigr) = 0$, we find that each term in the sum must vanish. □

**Example 1** We illustrate the "slice-and-dice" procedure and its relevance to the statements of Theorem 5 with Figure 1, where $\widehat{\Sigma}_3'$ is diffeomorphic to $S^2 \times S^1$. We first cut (slice) $\widehat{\Sigma}_3'$ by an embedded $S^2$ and view the result as a 3–manifold with boundary, which we further cut (dice) into seven 3–chains $\widehat{U}_1, \dots, \widehat{U}_7$ such that each

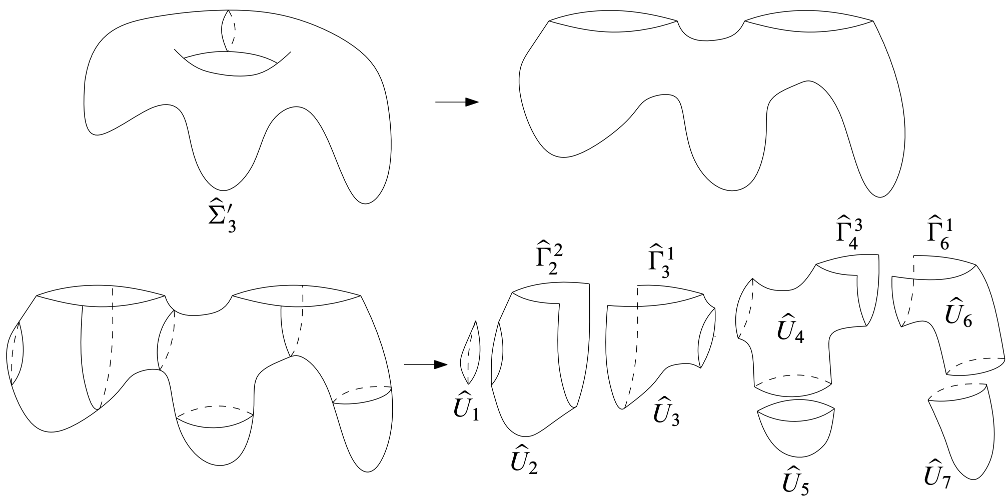


Figure 1: Cutting the 3–cycle $\widehat{\Sigma}'_3$ into small pieces.

$\widehat{U}_j$ satisfies $\operatorname{diam} \widehat{U}_j \le L_0$ (of course, the number of chains may vary in different examples). We label the boundary components of $\widehat{U}_j$, from left to right in the figure, by $\widehat{\Gamma}_j^l$ for $l = 1, \dots, k(j)$. Note that, in this case, there are four such boundary components that are nonsmooth, namely $\widehat{\Gamma}_2^2$, $\widehat{\Gamma}_3^1$, $\widehat{\Gamma}_4^3$ and $\widehat{\Gamma}_6^1$. The function $u$ groups different $\widehat{\Gamma}_j^l$ that glue together into a 2–cycle. For example, we have

$$u(2,2) = u(3,1) = u(4,3) = u(6,1),$$

and

$$u(1,1) = u(2,1), \quad u(3,2) = u(4,1), \quad u(4,2) = u(5,1), \quad u(6,2) = u(7,1).$$

Moreover, the values of $u$ on different groups of $\widehat{\Gamma}_j^l$ are different (eg $u(2,2) \neq u(1,1)$). Note here that $\sum_{(j,l)\in u^{-1}(r)} \widehat{\Gamma}_j^l = 0$ for each $r$.

The following proposition will be used to replace [7, Proposition 10] in the more general setting considered here:

**Proposition 6** *Consider $\pi : (\widehat{N}, \hat{g}) \to (N, g)$ a regular*[1] *Riemannian covering map of $n$–dimensional manifolds, with $(N, g)$ compact. Assume that $H_l(\widehat{N}, \mathbb{Z}) = 0$. Then, for $r > 0$, there is $R = R(r) < \infty$ such that $H_l(B_r(x), \mathbb{Z}) \to H_l(B_R(x), \mathbb{Z})$ is the zero map for any $x \in \widehat{N}$.*

[1] Recall that a cover is regular if the group of deck transformations acts transitively on the fibers. In particular, the universal cover is a regular cover.

**Proof** We first fix $x = x_0$. For any $r > 0$, there is $r_1 \in [r, 2r]$ with $\overline{B_{r_1}(x_0)}$ a compact manifold (with boundary). By Corollaries A.8 and A.9 in [19], the homology groups of $\overline{B_{r_1}(x_0)}$ are finitely generated. Assume that $\alpha_1, \dots, \alpha_J$ generates $H_l(\overline{B_{r_1}(x_0)}, \mathbb{Z})$. By assumption, for each $i$, $\alpha_i = \partial\beta_i$ for some $(l+1)$–chains $\beta_1, \dots, \beta_J$. Choose $R_1 = R_1(r)$ so that $\beta_i \in B_{R_1}(x_0)$ for $i = 1, \dots, J$. Then we see that $H_l(\overline{B_{r_1}(x_0)}, \mathbb{Z}) \to H_l(B_{R_1}(x_0), \mathbb{Z})$ is the zero map, so, in particular,

$$H_l(B_r(x_0), \mathbb{Z}) \to H_l(B_{R_1}(x_0), \mathbb{Z})$$

is the zero map.

Now, for any $x \in \widehat{N}$, we can assume (using a deck transformation) that $d(x, x_0) \leq \operatorname{diam} N$. Thus,

$$B_r(x) \subset B_{r+\operatorname{diam} N}(x_0) \quad \text{and} \quad B_{R_1(r+\operatorname{diam} N)}(x_0) \subset B_{R_1(r+\operatorname{diam} N)+\operatorname{diam} N}(x).$$

Thus, we find that the assertion holds for $R(r) = R_1(r + \operatorname{diam} N) + \operatorname{diam} N$. □

Putting these facts together, we thus obtain the following generalization of the filling estimate obtained in [7; 14]:

**Corollary 7** *Suppose that, for $n \in \{4, 5\}$, $(N^n, g)$ is a closed Riemannian $n$–manifold with positive scalar curvature and $\pi_2(N) = \cdots = \pi_{n-2}(N) = 0$. Then there is $L = L(N, g) > 0$ with the following property. Consider $\Sigma_{n-2}$ an closed embedded $(n-2)$–submanifold in $\widetilde{N}$ the universal cover. Then $\Sigma_{n-2}$ is nullhomologous in $B_L(\Sigma_{n-2})$.*

**Proof** Observe that the universal cover $\widetilde{N}$ has $\pi_1(\widetilde{N}) = \cdots = \pi_{n-2}(\widetilde{N}) = 0$. By the Hurewicz theorem, $H_{n-3}(\widetilde{N}, \mathbb{Z}) = H_{n-2}(\widetilde{N}, \mathbb{Z}) = 0$.

When $n = 4$, the assertion immediately follows from a combination of Theorem 5 with Proposition 6. Indeed, Theorem 5 implies that $\Sigma_2$ is homologous to $\Sigma_2'$ in $B_{L_0}(\Sigma_2)$, where $\operatorname{diam}(\Sigma_2') \leq L_0$. Proposition 6 implies that $\Sigma_2'$ can be filled in an $R(L_0)$–neighborhood. Thus, $\Sigma_2$ can be filled in an $(L_0+R(L_0))$–neighborhood.

When $n = 5$, the proof is more complicated due to the nature of the "slice-and-dice" decomposition in Theorem 5. Fix $\widehat{\Sigma}_3' \subset B_{L_0}(\Sigma_3)$ homologous to $\Sigma_3$ and $\{\widehat{U}_j\}$ and $\{\widehat{\Gamma}_j^l\}$ with the properties described in Theorem 5. We can now fill $\widehat{\Sigma}_3'$ in a bounded neighborhood following [7, Section 6.4], which we explain here. Since $\operatorname{diam}(\widehat{\Gamma}_j^l) \leq L_0$, Proposition 6 implies that $\widehat{\Gamma}_j^l = \partial\widetilde{\Gamma}_j^l$ for a 3–chain with $\operatorname{diam}(\widetilde{\Gamma}_j^l) \leq R(L_0)$. Then, because $\operatorname{diam}(\widehat{U}_j) \leq L_0$,

$$\widehat{U}_j - \sum_{l=1}^{k(j)} \widetilde{\Gamma}_j^l$$

is a 3–cycle of diameter $\leq L_0 + 2R(L_0)$. Thus, by Proposition 6, there is a 4–chain $\widetilde{U}_j$ with $\operatorname{diam}(\widetilde{U}_j) \leq R(L_0 + 2R(L_0))$ and

$$\partial \widetilde{U}_j = \widehat{U}_j - \sum_{l=1}^{k(j)} \widetilde{\Gamma}_j^l.$$

On the other hand, as was proven in Theorem 5, there is

$$u\colon \{(j,l) : j = 1, \dots, m, l = 1, \dots, k(j)\} \to \{1, \dots, q\}$$

such that

$$\operatorname{diam}\biggl(\bigcup_{(j,l)\in u^{-1}(r)} \widehat{\Gamma}_j^l\biggr) \leq L_0$$

and

$$\sum_{(j,l)\in u^{-1}(r)} \widehat{\Gamma}_j^l = 0$$

as 2–chains.

As such, for $r \in \{1, \dots, q\}$, $\sum_{(j,l)\in u^{-1}(r)} \widetilde{\Gamma}_j^l$ is a 3–cycle of diameter bounded by $2R(L_0) + L_0$ and thus there is a 4–chain $\Xi_r$ with $\operatorname{diam}(\Xi_r) \leq R(L_0 + 2R(L_0))$ and

$$\partial \Xi_r = \sum_{(j,l)\in u^{-1}(r)} \widetilde{\Gamma}_j^l.$$

This yields

$$\widehat{\Sigma}_3' = \partial \biggl[\sum_{r=1}^{q} \Xi_q + \sum_{j=1}^{m} \widetilde{U}_j\biggr]$$

with

$$\sum_{r=1}^{q} \Xi_q + \sum_{j=1}^{m} \widetilde{U}_j \subset B_{R(L_0+2R(L_0))}(\widehat{\Sigma}_3').$$

Thus, $\Sigma_3$ is nullhomologous in an $\bigl(R(L_0+2R(L_0))+R(L_0)\bigr)$–neighborhood. $\square$

**Example 2** Continuing Example 1, we illustrate in Figure 2 how Corollary 7 works for $\widehat{\Sigma}_3'$ in Figure 1. Consider all 2–cycles $\widehat{\Gamma}_j^i$ with $u(j,i) = r$. Fill in $\widehat{\Gamma}_j^l$ with a 3–chain $\widetilde{\Gamma}_j^l$. By construction, the sum of these $\widetilde{\Gamma}_j^l$ forms a 3–cycle, which can then be filled in by a 4–chain $\Xi_r$. By Proposition 6 and Corollary 7, the diameter of all these fill-ins are bounded by $R(L_0 + 2R(L_0))$.

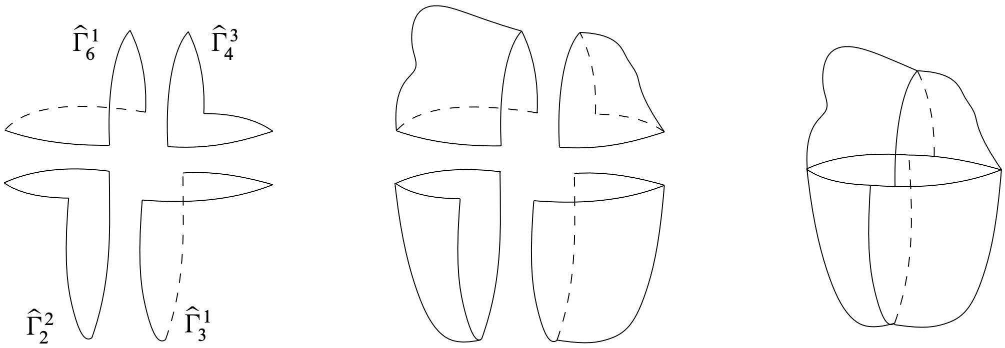


Figure 2: To fill in $\{\widehat{\Gamma}_j^l : u(j,l) = r\}$, we first fill in each $\widehat{\Gamma}_j^l$ with $\widetilde{\Gamma}_j^l$ (center), then fill in by $\Xi_r$ to obtain a 3–cycle (right).

## 2 Filling versus Urysohn width

The next result is inspired by work of Hannah Alpert, Alexey Balitskiy and Larry Guth [1], which we learnt about from a talk by Alpert. The strategy should be compared with [16, Corollary 10.11].

**Proposition 8** *Assume that $(N^n, g)$ has the property that any closed embedded $(n-2)$–submanifold in the universal cover $\Sigma_{n-2} \subset \widetilde{N}$ can be filled in $B_L(\Sigma_{n-2})$. Then the universal cover $(\widetilde{N}, \tilde{g})$ satisfies:*

(∗) *For any point $p \in \widetilde{N}$, each connected component of a level set of $d(p,\cdot)$ has diameter $\leq 20L$.*

Note that Corollary 7 implies a manifold $(N, g)$ in Theorem 1 satisfies the assumptions of Proposition 8. By the argument in [17, Corollary 10.11], this shows that the universal cover $(\widetilde{N}, \tilde{g})$ has Urysohn 1–width $\leq 20L$. In particular, the macroscopic dimension of $\widetilde{N}$ is 1.

**Proof** Let $p \in \widetilde{N}$ be a point and consider level sets of the distance function $f(x) = d(p,x)$.

For the sake of contradiction, suppose that there is a curve $\gamma \subset f^{-1}(t)$ connecting points $x$ and $y$ with $d(x,y) \geq 20L$. Fix a minimizing geodesic $\eta_x$ from $p$ to $x$ (and similarly for $\eta_y$) and consider the triangle $T = \eta_x * \gamma * -\eta_y$. Fix $0 < l < L$ such that $\partial B_{4L+l}(x)$ and $\partial B_{L+l}(\eta_x)$ are smooth hypersurfaces intersecting transversely. Set $\Sigma_{n-2} := \partial B_{4L+l}(x) \cap \partial B_{L+l}(\eta_x)$. Note that we have not ruled out $\Sigma_{n-2} = \varnothing$; in this case we will take $d(\Sigma_{n-2}, \cdot) = \infty$ below.

By construction,

$$d(\Sigma_{n-2}, \eta_x) > L.$$

Set $\Sigma'_{n-1} := \partial B_{4L+l}(x) \cap \overline{B_{L+l}(\eta_x)}$ and note that $\partial \Sigma'_{n-1} = \Sigma_{n-2}$. Observe that, since $\eta_x$ is a minimizing geodesic between $p$ and $x$ and $\Sigma'_{n-1}$ is a subset of $\partial B_{4L+l}(x)$, it must hold that $\eta_x$ intersects $\Sigma'_{n-1}$ exactly once and does so orthogonally (and thus transversally). We will return to this observation below.

**Lemma 9** $$d(\Sigma_{n-2}, \gamma) \geq d(\Sigma'_{n-1}, \gamma) > L.$$

**Proof** We first prove that $d(\Sigma'_{n-1}, \gamma) > L$. Choose $s \in \Sigma'_{n-1}$ with $d(s, \gamma) = d(\Sigma'_{n-1}, \gamma)$. There is $e \in \eta_x$ such that $d(s, e) \leq L + l$. We have

$$d(x, e) \geq d(x, s) - d(s, e) \geq 4L + l - (L + l) = 3L.$$

Since $\eta_x$ is minimizing (and has length $t$), we have $d(p, e) \leq t - 3L$. Thus,

$$d(p, s) \leq d(p, e) + d(e, s) \leq t - 3L + L + l = t - 2L + l.$$

Thus,

$$d(s, \gamma) \geq d(p, \gamma) - d(p, s) \geq t - (t - 2L + l) = 2L - l.$$

This completes the proof of $d(\Sigma'_{n-1}, \gamma) > L$. Since $\Sigma_{n-2} \subset \Sigma'_{n-1}$, it clearly holds that $d(\Sigma_{n-2}, \gamma) \geq d(\Sigma'_{n-1}, \gamma)$. □

**Lemma 10** $$\Sigma'_{n-1} \cap \eta_y = \varnothing.$$

**Proof** Suppose the contrary. Consider $s \in \Sigma'_{n-1} \cap \eta_y$. Note that $d(s, x) = 4L + l$ and there is $e \in \eta_x$ with $d(s, e) \leq L + l$. We have

$$d(x, e) \leq d(x, s) + d(e, s) \leq 5L + 2l.$$

As such,

$$d(p, e) \geq t - 5L - 2l,$$

so

$$d(p, s) \geq d(p, e) - d(e, s) \geq t - 5L - 2l - L - l = t - 6L - 3l.$$

Thus,

$$d(s, y) \leq 6L + 3l.$$

However, this contradicts

$$20L \leq d(x, y) \leq d(x, s) + d(s, y) \leq 4L + l + 6L + 3l = 10L + 4l. \qquad \square$$

**Lemma 11** $$d(\Sigma_{n-2}, \eta_y) > L.$$

**Proof** The proof is similar to the previous argument. Suppose we have $s \in \Sigma_{n-2}$ and $e_y \in \eta_y$ with $d(s, e_y) \leq L$. There is $e_x \in \eta_x$ with $d(s, e_x) = L + l$. Note that

$d(s,x) = 4L + l$. Thus,

$$d(p, e_x) = t - d(x, e_x) \geq t - d(x, s) - d(s, e_x) \geq t - 5L - 2l.$$

Thus,

$$d(p, e_y) \geq d(p, e_x) - d(e_x, e_y) \geq d(p, e_x) - d(e_x, s) - d(s, e_y) = t - 7L - 3l.$$

This implies that

$$d(y, e_y) \leq 7L + 3l.$$

However, this contradicts

$$20L \leq d(x, y) \leq d(x, s) + d(s, e_y) + d(e_y, y) \leq 12L + 4l. \qquad \square$$

We can now complete the proof of Proposition 8. Perturb the triangle $T$ to be a smooth embedded curve $T'$ still intersecting $\Sigma'_{n-1}$ transversely. As long as the perturbation is small, $T' \cap \Sigma'_{n-1}$ will consist of a single point (thanks to Lemmas 9 and 10, along with the observation that $\eta_x$ intersects $\Sigma'_{n-1}$ transversely in exactly one point). Assume first that $\Sigma_{n-2} \neq \varnothing$. By assumption, there is $\Sigma_{n-1} \subset B_L(\Sigma_{n-2})$ with $\partial \Sigma_{n-1} = \Sigma_{n-2}$. Using Lemmas 9 and 11 as well as $d(\Sigma_{n-2}, \eta_x) = L + l$, we find that $\Sigma_{n-1} \cap T' = \varnothing$. As such, $T'$ has nontrivial algebraic intersection with the cycle $\Sigma'_{n-1} - \Sigma_{n-1}$. This is a contradiction since $\widetilde{N}$ is simply connected.

If $\Sigma_{n-2} = \varnothing$, then the argument is similar but simpler. In this case, we note that $\Sigma'_{n-1}$ is a cycle and, combining Lemmas 9 and 10 with the fact that $\Sigma'_{n-1}$ intersects $\eta_x$ transversely exactly once, we see that $\Sigma'_{n-1}$ is a cycle with nontrivial algebraic intersection with $T'$, a contradiction as before. $\square$

## 3 Fundamental group and homotopy type

In this section, we prove Theorem 1. We first prove (see Corollary 14 below) that a manifold $(N^n, g)$ whose universal cover satisfies the conclusion of Proposition 8 has virtually free fundamental group. (Recall that a group is virtually free if it processes a free subgroup of finite index.) This fact seems to be well known among certain experts (in particular, see [15, page 135]). We give a proof here, roughly following the strategy used in [28]. The argument is based on notion of the number of *ends* of a group.

**Definition 12** Given a group $G$, its number of ends, $e(G)$, is defined as the number of topological ends of $\widetilde{K}$, where $\widetilde{K} \to K$ is a regular covering of finite simplicial complexes $K$ and $\widetilde{K}$, and $G$ is the group of deck transformations.

It follows from [9] that a finitely generated group can have 0, 1, 2 or infinitely many ends. Our main result here is as follows:

**Proposition 13** *Suppose $(N, g)$ is a closed Riemannian manifold satisfying the conclusions of Proposition 8. Then any finitely generated subgroup $G$ of $\pi_1(N)$ cannot have one end.*

We will prove this below, but first we note that it yields the desired statement:

**Corollary 14** *Suppose $(N, g)$ is a closed Riemannian manifold satisfying the conclusions of Proposition 8. Then $\pi_1(N)$ is virtually free.*

**Proof** We follow the proof of [28, Theorem 2.5]. Indeed, by combining the main result of [8] (see [30, Section 7]) with Proposition 13, $\pi_1(N)$ is the fundamental group of a finite graph of groups with finite edge and vertex groups. The assertion now follows from Proposition 11 in Chapter II, Section 2.6 of [31] (or eg [30, Theorem 7.3]). $\square$

Moreover, we observe that given these results, we can finish the proof of Theorem 1.

**Proof of Theorem 1** By Corollary 7, Proposition 8, and Corollary 14, $\pi_1(N)$ is virtually free. Let $G \subset \pi_1(N)$ be a finite-index subgroup which is a free group. Consider the finite covering $\widehat{N} \xrightarrow{\hat{p}} N$ such that the image of $\hat{p}_\#$ is $G$. Then $\pi_1(\widehat{N})$ is a finitely generated free group. Since $\pi_2(\widehat{N}) = \cdots = \pi_{n-2}(\widehat{N}) = 0$, Sections 2 and 3 of [12] imply that $\widehat{N}$ is homotopy equivalent to $S^n$ or connected sums of $S^{n-1} \times S^1$. $\square$

We now give the proof of Proposition 13:

**Proof of Proposition 13** Suppose there is a finitely generated subgroup $G$ of $\pi_1(N)$ with one end. We will show that this leads to a contradiction.

We divide the proof into several steps. Take a cover $N_0 \xrightarrow{p} N$ such that $p_\#(\pi_1(N_0)) = G$. Because $p_\#\colon \pi_1(N_0) \to \pi_1(N)$ is injective, this ensures that $\pi_1(N_0) = G$. If $G$ is finite then $e(G) = 0$, so we can assume $G$ is infinite.

Since $G$ is finitely generated, we can find $K \subset N_0$ a compact submanifold with boundary containing representatives of all of the generators of $G$. Write $i\colon K \to N_0$ for the inclusion map and note that $i_\#\colon \pi_1(K) \to \pi_1(N_0) = G$ is surjective. Let $H = \ker i_\#$, so $G = \pi_1(K)/H$. Choose $j\colon \widetilde{K} \to K$ the cover (with $\widetilde{K}$ path connected) of $K$ so that

$j_\#(\pi_1(\widetilde{K})) = H$. Since $H$ is a normal subgroup of $\pi_1(K)$, the covering $j : \widetilde{K} \to K$ is regular, and the group of deck transformations of $j$ is isomorphic to $\pi_1(K)/H = G$. Thus, $\widetilde{K}$ is noncompact. Note that $j_\# \circ i_\# \colon \pi_1(\widetilde{K}) \to \pi_1(N_0)$ is the zero map, so we can lift $i$ to $\tilde{\imath} \colon \widetilde{K} \to \widetilde{N}$, where $\widetilde{N}$ is the universal cover of $N$. (We emphasize that $\widetilde{K}$ is not necessarily the universal cover of $K$.)

As such, we have the diagram of spaces

$$\begin{tikzcd}
\widetilde{K} \arrow[r, "\tilde{\imath}"] \arrow[d, "j"] & \widetilde{N} \arrow[d, "\tilde{p}"] \\
K \arrow[r, "i"] & N_0 \arrow[d, "p"] \\
 & N
\end{tikzcd}$$

The maps $i$ and $\tilde{\imath}$ are inclusions of codimension zero submanifolds with boundary; indeed:

**Lemma 15** *$\tilde{\imath}$ is a proper embedding.*

**Proof** We first show that $\tilde{\imath}$ is injective. Suppose that $\tilde{\imath}(\tilde{a}) = \tilde{\imath}(\tilde{b})$. Connect $a$ and $b$ by a curve $\tilde{\eta}$ in $\widetilde{K}$. By assumption, $\tilde{\imath}(\tilde{\eta})$ is a loop in $\widetilde{N}$, so $\eta = j(\tilde{\eta})$ has $[\eta] = e \in \pi_1(N_0) = G$. Thus, $[\eta] \in H \subset \pi_1(K)$. This implies that $\tilde{\eta}$ is a loop, ie $a = b$. It is straightforward to check that $\tilde{\imath}$ is a closed map, using the fact that $K$ is compact and $\tilde{p}$ is a covering map. Therefore, $\tilde{\imath}$ is proper. □

Note that $N$ is equipped with a Riemannian metric $g$, so that, for any $p \in \widetilde{N}$, each connected component of a level set of $f_p(x) = d_{\tilde{g}}(x, p)$ has diameter $\leq C$, where $C = 20L$ as given in Proposition 8. The embeddings $i, \tilde{\imath}$ induce metric structures on $K$ and $\widetilde{K}$, respectively.

**Lemma 16** *For each $r > 0$, there exists $R(r) > 0$ such that, for any $a, b \in \widetilde{K}$ with $d_{\widetilde{N}}(a, b) \leq r$, we have $d_{\widetilde{K}}(a, b) \leq R(r)$.*

**Proof** Fix $x \in \widetilde{K}$. By applying a deck transformation, we can assume that $d_{\widetilde{N}}(a, x) \leq c_0$ (here $c_0$ only depends on $K$), so $d_{\widetilde{N}}(b, x) \leq r + c_0$. Since $\widetilde{K}$ is connected (and $\tilde{\imath}$ is proper), there exists $R = R(r + c_0)$ so that $d_{\widetilde{K}}(a, x), d_{\widetilde{K}}(b, x) \leq R$. The assertion follows from the triangle inequality. □

In the following lemma, we will call a curve $\tilde{\gamma}\colon \mathbb{R} \to \widetilde{K}$ a *line* if it minimizes length on compact subintervals relative to competitors in $\widetilde{K}$. Note that such a curve is a geodesic in the sense of metric geometry, but not necessarily in the sense of Riemannian geometry, since it could stick to $\partial\widetilde{K}$ in places. Similarly, we will call $\sigma'\colon [0,\infty) \to \widetilde{K}$ a *minimizing ray* if it minimizes length in the same sense.

**Lemma 17** *There exists a line $\tilde{\gamma}$ in $\widetilde{K}$.*

**Proof** Fix $p \in \widetilde{K}$, and choose $p_j \in \widetilde{K}$ diverging. Let $\sigma_j$ denote a curve that minimizes length in $\widetilde{K}$ between $p$ and $p_j$. We assume that $\sigma_j$ is parametrized by unit speed. In particular, $\sigma_j$ is a 1–Lipschitz map from an interval to $\widetilde{K}$. Consider an exhaustion of $\widetilde{K}$ by nested compact sets containing $p$. Applying Arzelà–Ascoli in each compact set and taking a diagonal sequence, we obtain that, after passing to a subsequence, $\sigma_i$ converges to a minimizing ray $\sigma'\colon [0,\infty) \to \widetilde{K}$. Since $G$ is the group of deck transformations of $\widetilde{K} \to K$ acting transitively on $\widetilde{K}$ and $K$ is compact, we can choose $t_i \to \infty$ and deck transformations $\Phi_i$ so that $d_{\tilde{g}}\big(p, \Phi_i(\sigma'(t_i))\big)$ is uniformly bounded. Then $\sigma_i'(t) = \Phi_i(\sigma'(t+t_i))$ subsequentially converges to a geodesic line $\sigma$ (using Arzelà–Ascoli again). □

Parametrize the curve $\tilde{\gamma}$ so that $d_{\widetilde{K}}(\tilde{\gamma}(a), \tilde{\gamma}(b)) = |a-b|$ (note that $d_{\widetilde{N}}(\tilde{\gamma}(a), \tilde{\gamma}(b))$ might be smaller than $|a-b|$). Let $\gamma = \tilde{\imath} \circ \tilde{\gamma}$. Note that $\tilde{\gamma}$ is automatically proper in $\widetilde{K}$, and thus Lemma 15 implies that $\gamma$ is proper in $\widetilde{N}$.

For each $R > 0$, consider the open geodesic ball $B_R(\gamma(0)) \subset \widetilde{N}$. Define parameters

$$t_-(R) = \max\{t : \gamma(-\infty, t) \cap B_R(\gamma(0)) = \varnothing\},$$
$$t_+(R) = \min\{t : \gamma(t, \infty) \cap B_R(\gamma_0) = \varnothing\}.$$

Note that $t_\pm(R) \to \pm\infty$ as $R \to \infty$.

Since $e(\widetilde{K}) = 1$, $\gamma(t_\pm(R))$ can be connected in $\widetilde{K} \setminus B_R(\gamma(0))$. Because $\widetilde{N}$ is simply connected, this implies that $\gamma(t_\pm(R))$ lie in the *same* component of $\partial B_R(\gamma(0)) \subset \widetilde{N}$ and thus

$$d_{\widetilde{N}}\big(\gamma(t_-(R)), \gamma(t_+(R))\big) \le C.$$

On the other hand, we have

$$d_{\widetilde{K}}\big(\gamma(t_-(R)), \gamma(t_+(R))\big) = |t_-(R) - t_+(R)| \to \infty.$$

This contradicts Lemma 16. This completes the proof of Proposition 13. □

## 4 Generalization to the mapping problem

In this section we prove Theorem 2. The proof here is partly motivated by [14, Section 5], where nonexistence of PSC metrics on certain noncompact manifold admitting a proper, distance-decreasing map to a uniformly contractible manifold is established. We first observe that we may assume, without loss of generality, that $\pi_1(X)$ is infinite. Indeed, if $\pi_1(X)$ is finite, then the universal cover $\widetilde{X}$ is compact and satisfies that $\pi_1(\widetilde{X}) = \cdots = \pi_{n-2}(\widetilde{X}) = 0$. By the Hurewicz theorem, we have that $H_1(\widetilde{X}) = \cdots = H_{n-2}(\widetilde{X}) = 0$. Poincaré duality further implies that $H_1(\widetilde{X}) = \cdots = H_{n-1}(\widetilde{X}) = 0$, and hence $\widetilde{X}$ is homeomorphic to $S^n$.

We begin with the following general lemma. Note that it is tempting to try to lift a map of nonzero degree to the universal covers, but this map may not be proper (and hence the degree will not be well defined). We note that the construction of the appropriate cover is somewhat analogous to the construction of $\widetilde{K}$ in Section 3.

**Lemma 18** *Suppose that $X$ and $N$ are closed oriented manifolds and $f: N \to X$ has nonzero degree. Letting $\widetilde{X}$ denote the universal cover of $X$, there exists a connected cover $\widehat{N} \to N$ and a lift $\hat{f}: \widehat{N} \to \widetilde{X}$ such that $\hat{f}$ is proper and $\deg \hat{f} = \deg f$.*

**Proof** Choose a regular value $x \in X$ and set $f^{-1}(x) = \{z_1, \dots, z_k\}$. Consider $H := \ker f_{\#}: \pi_1(N, z_1) \to \pi_1(X, x)$. Choose a covering space $p: \widehat{N} \to N$ so that image $p_{\#}: \pi_1(\widehat{N}, \hat{z}_1) \to \pi_1(N, z_1)$ is $H$. Below we will show that the map $f$ lifts to $\hat{f}: \widehat{N} \to \widetilde{X}$ and that $\hat{f}$ satisfies the assertions made above.

**Noncompactness of $\widehat{N}$** We claim that $\widehat{N}$ is noncompact. We first show that the image of $f_{\#}$ is a subgroup of $\pi_1(X, x)$ with finite index. Let $G = f_{\#}(\pi_1(N, z_1))$ and $\bar{\pi}: (\overline{X}, \bar{x}) \to (X, x)$ be a covering map such that image$((\bar{\pi})_{\#}: \pi_1(\overline{X}, \bar{x}) \to \pi_1(X, x))$ is $G$. The map $f$ lifts to a map $\bar{f}: (N, z_1) \to (\overline{X}, \bar{x})$ such that $f = \bar{\pi} \circ \bar{f}$. Since $N$ is compact and $f$ is surjective, we see that $\overline{X}$ is compact. Hence, we have $\deg f = \deg \bar{\pi} \cdot \deg \bar{f}$. It follows that $\deg \bar{\pi}$ is an integer factor of $\deg f$, and thus $G$ is a subgroup of $\pi_1(X, x)$ of finite index.

The number of sheets of the covering map $p$ is the index of $H = p_{\#}(\pi_1(\widehat{N}, \hat{z}_1))$ in $\pi_1(N, z_1)$. Since $H$ is a normal subgroup, this is equal to the number of elements of the group $\pi_1(N, z_1)/H$, which is isomorphic to $G$ and thus of infinite order. This implies that $\widehat{N}$ is noncompact, as claimed.

**Lifting the map $f$** Consider $f \circ p \colon \widehat{N} \to X$. Note that $(f \circ p)_{\#} \colon \pi_1(\widehat{N}) \to \pi_1(X)$ is the zero map. Thus, we can lift $f \circ p$ to the universal cover of $X$:

$$\begin{tikzcd}
(\widehat{N}, \hat{z}_1) \arrow[r, "\hat{f}"] \arrow[d, "p"'] & (\widetilde{X}, \tilde{x}) \arrow[d, "\pi"] \\
(N, z_1) \arrow[r, "f"'] & (X, x)
\end{tikzcd}$$

Clearly, a loop in $N$ lifts to a loop in $\widehat{N}$ if and only if it is in $H$ (recall that $H$ is normal).

**Counting lifts of preimages** We now claim that $\#(\hat{f}^{-1}(\tilde{x}) \cap p^{-1}(z_j)) = 1$. To this end, suppose that $a, b \in \hat{f}^{-1}(\tilde{x}) \cap p^{-1}(z_j)$. Choose a path $\hat{\gamma}$ in $\widehat{N}$ connecting the two points. Then $\gamma = p \circ \hat{\gamma}$ is a loop in $N$ based at $z_j$. On the other hand, $\tilde{\gamma} := \hat{f} \circ \hat{\gamma}$ is a loop in $\widetilde{X}$ based at $\tilde{x}$. Since $e = \pi_{\#}[\tilde{\gamma}] = f_{\#}[\gamma]$, we thus see that $[\gamma] \in H$. This is a contradiction since this would imply that $\gamma$ lifts to a loop (as remarked above).

**Properness** We now show that $\hat{f}$ is proper. Assume that $\hat{r}_i \to \infty$ in $\widehat{N}$ but $\hat{f}(\hat{r}_i) \to q$ in $\widetilde{X}$. Since $N$ is compact, we can pass to a subsequence such that $p(\hat{r}_i) \to r \in N$. Then $\pi(q) = f(r)$.

Choose a contractible neighborhood $U \subset N$ with $r \in U$. By shrinking $U$, we can assume that $f(U)$ is contained in a contractible open set $W \subset X$. Then $\pi^{-1}(W)$ consists of disjoint copies of $W$. We can assume that $\hat{f}(\hat{r}_i)$ are all contained in the copy containing $q$.

Assume that $p(\hat{r}_i) \in U$ for all $i$. Fix paths $\eta_i$ from $p(\hat{r}_i)$ to $r$ in $U$ and paths $\hat{\gamma}_i$ from $\hat{r}_1$ to $\hat{r}_i$ in $\widehat{N}$. Then

$$\alpha_i := (\eta_i) * (p \circ \hat{\gamma}_i) * (-\eta_1)$$

is a loop from $r$ to $r$. Lift $\alpha_i$ to $\hat{\alpha}_i$ a path in $\widehat{N}$ that agrees with $\hat{\gamma}_i$ on that portion of $\hat{\alpha}_i$. Note that $\hat{\alpha}_i$ cannot be a loop for $i$ large, since the $\hat{r}_i$ are diverging.

We now consider $\tilde{\alpha}_i := \hat{f} \circ \hat{\alpha}_i$ a path in $\widetilde{X}$. By construction, $\tilde{\alpha}_i$ is a loop in $\widetilde{X}$. This is a contradiction as before.

**Degree** Finally, we check that $\deg \hat{f} = \deg f$. The lift $\tilde{x}$ is a regular point for $\hat{f}$ and we have seen that each element of $f^{-1}(x)$ lifts to a unique element of $\hat{f}^{-1}(\tilde{x})$. But the local degree of $\hat{f}$ at each preimage $\hat{z}_i$ is the same as the degree of $f$ at the corresponding point $p(\hat{z}_i)$ (since $p$ is a covering map). □

**Remark** With some trivial modifications in the proof, a similar result holds for possibly nonorientable $X$ and $N$ with a map $f: N \to X$ of nonzero mod 2 degree.

Using the lifted map $\hat{f}$ we can now follow [14, Section 5] to show that the conclusion of Corollary 7 holds in the setting of Theorem 2.

**Lemma 19** *Let $X$ and $N$ be oriented Riemannian manifolds and $f:(N,g)\to(X,g_X)$ with $f$ distance-decreasing and $\deg f \neq 0$. Assume that $N$ admits a metric of positive scalar curvature and that either $n=4$ and $\pi_2(X)=0$, or $n=5$ and $\pi_2(X)=\pi_3(X)=0$.*

*Then there exists $L>0$ with the following property: if $\Sigma_{n-2}$ is an $(n-2)$–dimensional nullhomologous cycle in the universal cover $\widetilde{X}$ of $X$, then the cycle $\deg(f)\Sigma_{n-2}$ can be filled inside $B_L(\Sigma_{n-2})$.*

**Proof** We consider $n=5$ since the $n=4$ case is similar (but simpler). By scaling, we can assume that $(N,g)$ has scalar curvature $R \geq 1$. As in Corollary 7, $H_2(\widetilde{X},\mathbb{Z}) = H_3(\widetilde{X},\mathbb{Z}) = 0$.

By assumption, $\Sigma_3 = \partial\Sigma_4$ in $\widetilde{X}$ for some chain $\Sigma_4$. Up to a small perturbation, we can assume that $\hat{f}$ is transversal to $\Sigma_3$ and $\Sigma_4$. Set $\widehat{\Sigma}_4 := \hat{f}^{-1}(\Sigma_4)$ and similarly $\widehat{\Sigma}_3 = \partial\widehat{\Sigma}_4$. Note that $\widehat{\Sigma}_3$ is nullhomologous in $\widehat{N}$ (by construction). Using Theorem 5, we can find $\widehat{\Sigma}_3' \subset B_{L_0}(\widehat{\Sigma}_3)$ homologous to $\widehat{\Sigma}_3$ as well as 3–chains $\widehat{U}_1,\dots,\widehat{U}_m$ with $\operatorname{diam}(U_j)\leq L_0$ and 2–cycles $\{\widehat{\Gamma}_j^l : j=1,\dots,m,\ l=1,\dots,k(j)\}$ with $\operatorname{diam}(\widehat{\Gamma}_j^l)\leq L_0$ and such that

$$\widehat{\Sigma}_3' = \sum_{j=1}^{m} \widehat{U}_j \qquad \text{and} \qquad \partial\widehat{U}_j = \sum_{l=1}^{k(j)} \widehat{\Gamma}_j^l \quad \text{for each } j=1,\dots,m,$$

where both equalities hold as chains (not just in homology). Finally, there is an integer $q$ and a function

$$u: \{(j,l) : j=1,\dots,m,\ l=1,\dots,k(j)\} \to \{1,\dots,q\}$$

such that, for $r\in\{1,\dots,q\}$, we have

$$\operatorname{diam}\Bigl(\bigcup_{(j,l)\in u^{-1}(r)} \widehat{\Gamma}_j^l\Bigr) \leq L_0$$

and

$$\sum_{(j,l)\in u^{-1}(r)} \widehat{\Gamma}_j^l = 0$$

as 2–cycles for $r\in\{1,\dots,q\}$.

Denote by $\Sigma'_3$ the 3–cycle in $\widetilde{X}$ obtained by pushing $\widehat{\Sigma}'_3$ forward by the map $\hat{f}$ and similarly for $U_j$ and $\Gamma^l_j$.

Since $\hat{f}$ is transversal to $M_3$, it is easy to check that $\deg \hat{f}|_{M_3} = \deg \hat{f}$. Hence, $\hat{f}_\#([\widehat{\Sigma}'_3]) = (\deg \hat{f})[\Sigma_3]$. Moreover, since $f$ (and thus $\hat{f}$) was assumed to be distance-decreasing, we see that $d_{(\widetilde{X}, g_{\widetilde{X}})}(\Sigma_3, \Sigma'_3) \le L_0$. As such, it suffices to bound $\Sigma'_3$ in a controlled neighborhood.

To do so, we follow the argument used in Corollary 7. Because $\operatorname{diam}(\widehat{\Gamma}^l_j) \le L_0$, we can use Proposition 6 to find a 3–chain $\widetilde{\Gamma}^j_l$ with $\operatorname{diam}(\widetilde{\Gamma}^l_j) \le R(L_0)$ and $\partial \widetilde{\Gamma}^j_l = \Gamma^j_l$ and then a 4–chain $\widetilde{U}_j$ with

$$\partial \widetilde{U}_j = U_j - \sum_{l=1}^{k(j)} \widetilde{\Gamma}^l_j$$

and $\operatorname{diam}(\widetilde{U}_j) \le R(L_0 + 2R(L_0))$. Thus,

$$\Sigma'_3 = \sum_{j=1}^{m} \partial \widetilde{U}_j + \sum_{r=1}^{q} \sum_{(j,l)\in u^{-1}(r)} \widetilde{\Gamma}^l_j.$$

and

$$\operatorname{diam}\Bigg( \sum_{(j,l)\in u^{-1}(r)} \widetilde{\Gamma}^l_j \Bigg) \le 2R(L_0) + L_0.$$

We can thus complete the proof as in Corollary 7. □

Granted Lemma 19, Theorem 2 follows. Indeed, in order to prove the Urysohn width estimate of Proposition 8, it is enough to assume that the filling radius estimate holds for a multiple $\deg(f)\Sigma_{n-2}$ of every cycle $\Sigma_{n-2}$. The rest of the proof of Theorem 2 proceeds exactly as the proof of Theorem 1.

# References


[1] **H Alpert**, **A Balitskiy**, **L Guth**, *Macroscopic scalar curvature and codimension 2 width*, preprint (2021) arXiv 2112.04594

[2] **D Bolotov**, *About the macroscopic dimension of certain PSC-manifolds*, Algebr. Geom. Topol. 9 (2009) 21–27 MR Zbl

[3] **D Bolotov**, **A Dranishnikov**, *On Gromov's scalar curvature conjecture*, Proc. Amer. Math. Soc. 138 (2010) 1517–1524 MR Zbl

[4] **S Brendle**, *Ricci flow with surgery on manifolds with positive isotropic curvature*, Ann. of Math. 190 (2019) 465–559 MR Zbl

[5] **S Brendle**, **R Schoen**, *Manifolds with* $1/4$*–pinched curvature are space forms*, J. Amer. Math. Soc. 22 (2009) 287–307 MR Zbl

[6] **B-L Chen**, **S-H Tang**, **X-P Zhu**, *Complete classification of compact four-manifolds with positive isotropic curvature*, J. Differential Geom. 91 (2012) 41–80 MR Zbl

[7] **O Chodosh**, **C Li**, *Generalized soap bubbles and the topology of manifolds with positive scalar curvature*, preprint (2020) arXiv 2008.11888

[8] **M J Dunwoody**, *The accessibility of finitely presented groups*, Invent. Math. 81 (1985) 449–457 MR Zbl

[9] **D B A Epstein**, *Ends*, from "Topology of 3–manifolds and related topics" (M K Fort, Jr, editor), Prentice-Hall, Englewood Cliffs, NJ (1962) 110–117 MR Zbl

[10] **A M Fraser**, *Fundamental groups of manifolds with positive isotropic curvature*, Ann. of Math. 158 (2003) 345–354 MR Zbl

[11] **M H Freedman**, *The topology of four-dimensional manifolds*, J. Differential Geometry 17 (1982) 357–453 MR Zbl

[12] **S Gadgil**, **H Seshadri**, *On the topology of manifolds with positive isotropic curvature*, Proc. Amer. Math. Soc. 137 (2009) 1807–1811 MR Zbl

[13] **M Gromov**, *Positive curvature, macroscopic dimension, spectral gaps and higher signatures*, from "Functional analysis on the eve of the 21st century, II" (S Gindikin, J Lepowsky, R L Wilson, editors), Progr. Math. 132, Birkhäuser, Boston, MA (1996) 1–213 MR Zbl

[14] **M Gromov**, *No metrics with positive scalar curvatures on aspherical* 5*–manifolds*, preprint (2020) arXiv 2009.05332

[15] **M Gromov**, *Four lectures on scalar curvature*, preprint (2021) arXiv 1908.10612v5

[16] **M Gromov**, **H B Lawson, Jr**, *Spin and scalar curvature in the presence of a fundamental group, I*, Ann. of Math. 111 (1980) 209–230 MR Zbl

[17] **M Gromov**, **H B Lawson, Jr**, *Positive scalar curvature and the Dirac operator on complete Riemannian manifolds*, Inst. Hautes Études Sci. Publ. Math. 58 (1983) 83–196 MR Zbl

[18] **R S Hamilton**, *Four-manifolds with positive isotropic curvature*, Comm. Anal. Geom. 5 (1997) 1–92 MR Zbl

[19] **A Hatcher**, *Algebraic topology*, Cambridge Univ. Press (2002) MR Zbl

[20] **H Huang**, *Compact manifolds of dimension* $n \geq 12$ *with positive isotropic curvature*, preprint (2019) arXiv 1909.12265

[21] **M Katz**, *The first diameter of* 3*–manifolds of positive scalar curvature*, Proc. Amer. Math. Soc. 104 (1988) 591–595 MR Zbl

[22] **M Kreck**, **W Lück**, *Topological rigidity for non-aspherical manifolds*, Pure Appl. Math. Q. 5 (2009) 873–914 MR Zbl

[23] **Y Liokumovich**, **D Maximo**, *Waist inequality for 3–manifolds with positive scalar curvature*, preprint (2020) arXiv 2012.12478

[24] **F C Marques**, **A Neves**, *Rigidity of min-max minimal spheres in three-manifolds*, Duke Math. J. 161 (2012) 2725–2752 MR Zbl

[25] **M J Micallef**, **J D Moore**, *Minimal two-spheres and the topology of manifolds with positive curvature on totally isotropic two-planes*, Ann. of Math. 127 (1988) 199–227 MR Zbl

[26] **J Milnor**, *Lectures on the h–cobordism theorem*, Princeton Univ. Press (1965) MR Zbl

[27] **H T Nguyen**, *Isotropic curvature and the Ricci flow*, Int. Math. Res. Not. 2010 (2010) 536–558 MR Zbl

[28] **M Ramachandran**, **J Wolfson**, *Fill radius and the fundamental group*, J. Topol. Anal. 2 (2010) 99–107 MR Zbl

[29] **R Schoen**, *Manifolds with poinwise* $1/4$*-pinched curvature*, preprint (2007)

[30] **P Scott**, **T Wall**, *Topological methods in group theory*, from "Homological group theory" (C T C Wall, editor), London Math. Soc. Lecture Note Ser. 36, Cambridge Univ. Press (1979) 137–203 MR Zbl

[31] **J-P Serre**, *Trees*, Springer (2003) MR Zbl

*Department of Mathematics, Stanford University*
*Stanford, CA, United States*

*Courant Institute of Mathematical Sciences, New York University*
*New York, NY, United States*

*Department of Mathematics, University of Toronto*
*Toronto ON, Canada*

ochodosh@stanford.edu, chaoli@nyu.edu, ylio@math.toronto.edu

Proposed: John Lott Received: 21 June 2021
Seconded: Urs Lang, Gang Tian Revised: 13 September 2021

## 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 27 Issue 4 (pages 1273–1655) 2023

Floer theory and reduced cohomology on open manifolds 1273

YOEL GROMAN

Geodesic coordinates for the pressure metric at the Fuchsian locus 1391

XIAN DAI

Birational geometry of the intermediate Jacobian fibration of a cubic fourfold 1479

GIULIA SACCÀ

Cohomological $\chi$–independence for moduli of one-dimensional sheaves and moduli of Higgs bundles 1539

DAVESH MAULIK and JUNLIANG SHEN

Coarse injectivity, hierarchical hyperbolicity and semihyperbolicity 1587

THOMAS HAETTEL, NIMA HODA and HARRY PETYT

Classifying sufficiently connected PSC manifolds in 4 and 5 dimensions 1635

OTIS CHODOSH, CHAO LI and YEVGENY LIOKUMOVICH