Pacific Journal of Mathematics

Volume 309 No. 1 November 2020

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

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

The subscription price for 2020 is US $520/year for the electronic version, and $705/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

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

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.1

LIE 2-ALGEBRAS OF VECTOR FIELDS

DANIEL BERWICK-EVANS AND EUGENE LERMAN

We show that the category of vector fields on a geometric stack has the structure of a Lie 2-algebra. This proves a conjecture of R. Hepworth. The construction uses a Lie groupoid that presents the geometric stack. We show that the category of vector fields on the Lie groupoid is equivalent to the category of vector fields on the stack. The category of vector fields on the Lie groupoid has a Lie 2-algebra structure built from known (ordinary) Lie brackets on multiplicative vector fields of Mackenzie and Xu and the global sections of the Lie algebroid of the Lie groupoid. After giving a precise formulation of Morita invariance of the construction, we verify that the Lie 2-algebra structure defined in this way is well-defined on the underlying stack.

1. Introduction

Vector fields on a Lie groupoid G form a category [Hepworth 2009]. We denote it by $\mathbb{X}(G)$. The *objects* of $\mathbb{X}(G)$ are the multiplicative vector fields of Mackenzie and Xu [1998]. These are functors $v : G \to TG$ satisfying $\pi_G \circ v = \mathrm{id}_G$, where TG denotes the tangent groupoid and $\pi_G : TG \to G$ is the projection functor. A *morphism* $\alpha : v \Rightarrow v'$ in this category is a natural transformation α such that $\pi_G(\alpha(x)) = \mathrm{id}_x$ for every object x of the groupoid G. The first result of this paper is:

Theorem 3.4. *The category of vector fields* $\mathbb{X}(G)$ *on a Lie groupoid* G *is a* (*strict*) *Lie* 2*-algebra. That is,* $\mathbb{X}(G)$ *is a category internal to the category of Lie algebras.*

Remark 1.1. When a manifold M is regarded as a discrete Lie groupoid, $\mathbb{X}(M)$ is the usual Lie algebra of vector fields on M regarded as a discrete Lie 2-algebra.

To every Lie groupoid G there corresponds the stack $\mathbb{B}G$ of principal G-bundles, and Morita equivalent Lie groupoids G and H correspond to isomorphic stacks $\mathbb{B}G$ and $\mathbb{B}H$. It is natural to wonder if the Lie 2-algebra $\mathbb{X}(G)$ lives on the stack $\mathbb{B}G$ in some appropriate sense. To start, we can ask whether Morita equivalent Lie groupoids G and H have "Morita equivalent" Lie 2-algebras $\mathbb{X}(G)$ and $\mathbb{X}(H)$. More precisely we could ask for the existence of a (2-)functor $\mathbb{X}$ from the bicategory Bi of Lie groupoids, bibundles and isomorphisms of bibundles to an appropriate

MSC2020: primary 17B66; secondary 18D05.
Keywords: Lie 2-algebra, stack, vector field, Lie groupoid.

bicategory of Lie 2-algebras that sends Morita equivalences to Morita equivalences. It turns out that such a functor is too much to ask for but there is a functor from a sub-bicategory of Bi.

The reasons behind this fact can already be seen in the case of manifolds. Recall that there is no naturally defined functor from the category of manifolds to the category of Lie algebras that assigns to each manifold its Lie algebra of vector fields. However if we restrict ourselves to the category $\mathsf{Man}_{\mathrm{iso}}$ whose objects are manifolds and whose morphisms are diffeomorphisms then there is a perfectly well defined functor with the desired properties.

Getting back to Lie groupoids, recall that there is a localization of the strict 2-category of Lie groupoids, internal functors, and internal natural transformations at the class of functors that are fully faithful and essentially surjective ; denote this localization by Bi. Lie 2-algebras, internal functors and internal natural transformations form the strict 2-category $\mathsf{Lie2Alg}_{\mathrm{strict}}$, and localizing at the essential equivalences produces a bicategory $\mathsf{Lie2Alg}$ (see Subsections 2B and 2C). Let $\mathsf{Bi}_{\mathrm{iso}}$ be the sub-bicategory of Bi whose objects are Lie groupoids, 1-morphisms are (weakly) *invertible* bibundles (i.e., Morita equivalences) and 2-morphisms are isomorphisms of bibundles. We recall that a bicategory with invertible 2-morphism is, by definition, a $(2,1)$-*bicategory*.

Theorem 4.1. *The map $G \mapsto \mathbb{X}(G)$ that assigns to each Lie groupoid its category of vector fields extends to a functor $\mathbb{X} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{Lie2Alg}$. In particular, if $P : G \to H$ is a Morita equivalence of Lie groupoids then $\mathbb{X}(P) : \mathbb{X}(G) \to \mathbb{X}(H)$ is a (weakly) invertible 1-morphism of Lie 2-algebras in the bicategory* $\mathsf{Lie2Alg}$.

Remark 1.2. In the Lie groupoid literature there are two standard constructions that associate a Lie algebra to a Lie groupoid: global sections of its Lie algebroid and Mackenzie and Xu's multiplicative vector fields. The Lie 2-algebra structure on $\mathbb{X}(G)$ is built out of this pair of Lie algebras. At first pass this might seem surprising: neither multiplicative vector fields nor sections of Lie algebroids are well-behaved under Morita equivalence of Lie groupoids. Theorem 4.1 shows that combining this pair of Lie algebras into a Lie 2-algebra gives us an object that *is* preserved by Morita equivalence.

How does the existence of the functor in Theorem 4.1 imply that the Lie 2-algebra $\mathbb{X}(G)$ "lives" on the stack $\mathbb{B}G$? To answer this, we need to recall the relationship between the bicategory Bi and the 2-category Stack of stacks over the site of smooth manifolds. The assignment $G \mapsto \mathbb{B}G$ extends to a fully faithful functor $\mathbb{B} : \mathsf{Bi} \to \mathsf{Stack}$. The essential image of the functor $\mathbb{B}$ is the 2-category $\mathsf{GeomStack}$ of geometric stacks. Restricting the functor $\mathbb{B}$ to the bicategory $\mathsf{Bi}_{\mathrm{iso}}$ of groupoids and Morita equivalences gives us an equivalence of bicategories $\mathbb{B} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{GeomStack}_{\mathrm{iso}}$, where $\mathsf{GeomStack}_{\mathrm{iso}}$ is the (2,1)-category of geometric stacks, isomorphisms of stacks (that is, weakly invertible 1-morphisms of stacks) and 2-morphisms. By inverting this equivalence $\mathbb{B}$ and composing it with the functor $\mathbb{X}$ we get a functor

$$\mathsf{GeomStack}_{\mathrm{iso}} \xrightarrow{\mathbb{B}^{-1}} \mathsf{Bi}_{\mathrm{iso}} \xrightarrow{\mathbb{X}} \mathsf{Lie2Alg}. \tag{1-1}$$

So in particular we get a functorial assignment of a Lie 2-algebra to every geometric stack, with isomorphic stacks being assigned "isomorphic" Lie 2-algebras.

Hepworth [2009] introduced a category of vector fields $\mathsf{Vect}(\mathcal{A})$ on a stack $\mathcal{A}$, which is a groupoid. We introduce a groupoid $\mathsf{Vect}'(\mathcal{A})$ equivalent to $\mathsf{Vect}(\mathcal{A})$ which is more convenient for our purposes. In particular, the assignment $\mathcal{A} \mapsto \mathsf{Vect}'(\mathcal{A})$ easily extends to a functor $\mathsf{Vect}' : \mathsf{GeomStack}_{\mathrm{iso}} \to \mathsf{Gpd}$, where Gpd is the (2,1)-category of groupoids, functors and natural transformations (which are automatically natural isomorphisms). We show that the functor Vect' is compatible with the functors $\mathbb{B} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{GeomStack}_{\mathrm{iso}}$ and $\mathbb{X} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{Lie2Alg}$ in the following sense.

Theorem 6.1. *The diagram of* (2,1)*-bicategories and functors*

$$\begin{array}{ccc} \mathsf{GeomStack}_{\mathrm{iso}} & \xrightarrow{\mathsf{Vect}'} & \mathsf{Gpd} \\ {\scriptstyle\mathbb{B}}\uparrow & \overset{\Upsilon}{\nwarrow\!\!\!\nwarrow} & \uparrow{\scriptstyle u} \\ \mathsf{Bi}_{\mathrm{iso}} & \xrightarrow[\mathbb{X}]{} & \mathsf{Lie2Alg} \end{array}$$

2*-commutes. Here as above* Gpd *denotes the* (2,1)*-category of groupoids, functors and natural isomorphisms, and* $u : \mathsf{Lie2Alg} \to \mathsf{Gpd}$ *denotes the functor that assigns to each Lie* 2*-algebra its underlying groupoid. The components of the transformation* Υ *are weakly invertible functors* (*i.e., equivalences of categories*). *In particular for a geometric stack* $\mathcal{A}$ *the category underlying the Lie* 2*-algebra* $(\mathbb{X} \circ \mathbb{B}^{-1})(\mathcal{A})$ *is equivalent to Hepworth's category* $\mathsf{Vect}(\mathcal{A})$ *of vector fields on the stack.*

Remark 1.3. Consider a geometric stack $\mathcal{A}$. The groupoids $\mathsf{Vect}(\mathcal{A})$ and $\mathsf{Vect}'(\mathcal{A})$ are equivalent. Let $G = \mathbb{B}^{-1}(\mathcal{A})$. Then the stacks $\mathbb{B}G$ and $\mathcal{A}$ are isomorphic and consequently the categories $\mathsf{Vect}'(\mathcal{A})$ and $\mathsf{Vect}'(\mathbb{B}G)$ are equivalent. Since the diagram above 2-commutes and the components of Υ are weakly invertible functors, $\mathsf{Vect}'(\mathcal{A})$ is equivalent to the groupoid $u(\mathbb{X}(G))$ underlying the Lie 2-algebra $\mathbb{X}(G)$. Consequently Hepworth's groupoid $\mathsf{Vect}(\mathcal{A})$ of vector fields on a stack is equivalent to the groupoid underlying the Lie 2-algebra $\mathbb{X}(\mathbb{B}^{-1}(\mathcal{A}))$.

A different choice $(\mathbb{B}^{-1})'$ of the inverse of $\mathbb{B}$ is isomorphic to $\mathbb{B}^{-1}$. Consequently the Lie 2-algebras $\mathbb{X}(\mathbb{B}^{-1}(\mathcal{A}))$ and $\mathbb{X}((\mathbb{B}^{-1})'(\mathcal{A}))$ are naturally weakly isomorphic and their underlying groupoids are equivalent.

Related work. The recent work of Cristian Ortiz and James Waldron [2019] lies in a similar circle of ideas. Recall that an $\mathcal{LA}$-groupoid is a groupoid object in Lie algebroids. Given an $\mathcal{LA}$-groupoid, Ortiz and Waldron introduced its category of multiplicative sections and showed that it carries a natural strict Lie 2-algebra structure in the language of crossed modules of Lie algebras (which affords an equivalent description of the category of strict Lie 2-algebras). They showed that if two $\mathcal{LA}$-groupoids are Morita equivalent then the corresponding crossed modules of Lie algebras are connected by a zig-zag of equivalences. Furthermore, to every stack Ortiz and Waldron assigned an ordinary Lie algebra and showed that in the case of proper geometric stacks the set underlying this Lie algebra is in bijective correspondence with isomorphism classes of vector fields in Hepworth's definition.

2. Background and notation

We assume that the reader is familiar with ordinary categories, strict 2-categories and bicategories (also known as weak 2-categories). We mostly work with (2,1)-bicategories, that is with bicategories whose 2-morphisms are invertible. Standard references for bicategories are [Bénabou 1967; Borceux 1994]. See also [Street 1996]. We assume familiarity with Lie groupoids, with a standard reference being [Moerdijk and Mrčun 2003]. We also assume that the reader is comfortable with stacks over the site of manifolds and the relationship with Lie groupoids, e.g., see [Behrend and Xu 2011]. This said, Sections 3, 4 and 7 do not use stacks.

Given a category C we denote its collection of objects by C_0 and the collection of arrows/morphisms by C_1.[1] We usually denote the source and target maps of C by s and t, respectively. We write $\mathsf{C} = \{\mathsf{C}_1 \rightrightarrows \mathsf{C}_0\}$ and suppress the other structure maps of the category C. The map $1 : \mathsf{C}_0 \to \mathsf{C}_1$ assigns the identity arrow 1_x to each object x of the category C. The composition/multiplication in the category C is defined on the collection C_2 of pairs of composable arrows. Our convention is that $\mathsf{C}_2 := \{(\gamma_2, \gamma_1) \in \mathsf{C}_1 \times \mathsf{C}_1 \mid s(\gamma_2) = t(\gamma_1)\} =: \mathsf{C}_1 \times_{s,\mathsf{C}_0,t} \mathsf{C}_1$. We denote the composition in the category C by m. In particular, we write the composition from right to left: $\gamma_2\gamma_1$ means γ_1 followed by γ_2. If the category C is a groupoid we denote the inversion map by i.

2A. *Bicategories of Lie groupoids.*

Definition 2.1. A *bibundle* $P : G \to H$ from a Lie groupoid G to a Lie groupoid H is a manifold P with two maps a_P^L and a_P^R:

$$\begin{array}{ccccc} G_1 & & & & H_1 \\ \downarrow\downarrow & & P & & \downarrow\downarrow \\ & \swarrow^{a_P^L} & & \searrow^{a_P^R} & \\ G_0 & & & & H_0 \end{array}$$

along with a left action of G and a right action of H on P:

$$G_1 \times_{G_0} P \to P, \quad (g, p) \mapsto g \cdot p, \qquad P \times_{H_0} H_1 \to P, \quad (p, h) \mapsto p \cdot h.$$

[1] We use the words "arrow," "morphism" and "1-cell" interchangeably.

We refer to a^L_P as the *left anchor* and to a^R_P as the *right anchor*. We further require that

(1) the actions of G and H commute: for all $(g, p) \in G_1 \times_{G_0} P$, $(p, h) \in P \times_{H_0} H_1$
$$g \cdot (p \cdot h) = (g \cdot p) \cdot h;$$
(2) the map $a^P_L : P \to G_0$ is a surjective submersion and is H-invariant: $a^P_L(p \cdot h) = a^P_L(p)$ for all $(p, h) \in P \times_{H_0} H_1$;
(3) the map a^P_R is G-invariant: $a^P_R(g \cdot p) = a^P_R(p)$ for all $(g, p) \in G_1 \times_{G_0} P$;
(4) the map
$$P \times_{a^R_P, H_0, t} H_1 \to P \times_{a^L_P, G_0, a^L_P} P, \quad (p, h) \mapsto (p, p \cdot h)$$
is a diffeomorphism.

Definition 2.2. An *isomorphism* of two bibundles $P, P' : G \to H$ is a diffeomorphism $\alpha : P \to P'$ which is G- and H-equivariant.

Given two bibundles $P : G \to H$ and $Q : H \to K$ their composite $Q \circ P$ is defined to be the quotient of the fiber product $P_{a^P_R, H_0, a^Q_L} Q$ by the action of H: $Q \circ P := (P \times_{a^P_R, H_0, a^Q_L} Q)/H$. The composition of bibundles is not associative on the nose: given 3 composable bibundles P, Q and R the composites $(R \circ Q) \circ P$ and $R \circ (Q \circ P)$ are only isomorphic. Therefore Lie groupoids and bibundles do not form a category. One can show that Lie groupoids, bibundles and isomorphisms of bibundles form a bicategory which we denote by Bi.

Notation 2.3. An isomorphism of bibundles $\alpha : P \to Q$ is a smooth map and a 2-arrow (2-cell) in the bicategory Bi described above. For this reason we may sometimes write $\alpha : P \Rightarrow Q$ for an isomorphism of bibundles.

There is also the strict 2-category LieGpd of Lie groupoids, (smooth) functors and (smooth) natural transformations Note that both Bi and LieGpd are (2,1)-bicategories.

Notation 2.4. In the bicategories LieGpd and Bi, we write the horizontal composition of 2-arrows as $\star$. Given a 1-morphism f and a 2-morphism α we abuse notation by writing $f \star \alpha$ for the horizontal composition (whiskering) $1_f \star \alpha$, where 1_f is the identity 2-arrow on the 1-morphism f. The vertical composition of 2-morphisms is denoted by $\circ$. When convenient, we also use arrow notation to denote morphisms in groupoids with specified source or target, e.g., $x \xleftarrow{g} y$ for a morphism g with target x and source y.

Remark 2.5. There is a functor $U :$ LieGpd $\to$ Bi that is the identity on objects. On 1-morphisms U sends a functor $f : G \to H$ to the bibundle

$$\begin{aligned} \langle f \rangle &:= G_0 \times_{f, H_0, t} H_1 \\ &:= \{(x, \gamma) \mid f(x) = t(\gamma)\} = \{(x, f(x) \xleftarrow{\gamma}) \mid x \in G_0, \gamma \in H_1\} \end{aligned} \tag{2-1}$$

whose left and right anchor maps are given respectively by

$$a^L_{\langle f\rangle}(x,\gamma) = x, \qquad a^R_{\langle f\rangle}(x,\gamma) = s(\gamma).$$

Here, as before, $s : H_1 \to H_0$ is the source map. The left action of the groupoid G on the manifold $\langle f\rangle$ is given by $(g, (x, \gamma)) \mapsto (t(g), f(g)\gamma)$. The right action of the groupoid H on $\langle f\rangle$ is given by $((x, \gamma), \nu) \mapsto (x, \gamma\nu)$. Note that $a^L_P : \langle f\rangle \to G_0$ has a canonical section $x \mapsto (x, 1_{f(x)})$. Given a pair of functors $f, k : G \to H$ and a natural isomorphism $\alpha : f \Rightarrow k$, we get an isomorphism of bibundles $\langle\alpha\rangle : \langle f\rangle \Rightarrow \langle k\rangle$. The isomorphism $\langle\alpha\rangle$ is defined by $\langle\alpha\rangle(x, f(x) \xleftarrow{\gamma}) = (x, k(x) \xleftarrow{\alpha(x)\gamma})$.

The functor $U = \langle\ \rangle$ takes vertical and horizontal composition of natural transformations to the composition of isomorphisms of bibundles and horizontal composition of isomorphisms, respectively.

Remark 2.6. By construction of the functor $U = \langle\ \rangle$ the total space of the bibundle $\langle \mathrm{id}_G\rangle$ corresponding to the identity functor $\mathrm{id}_G : G \to G$ on a Lie groupoid G is the fiber product $G_0 \times_{G_0} G_1$. This fiber product is diffeomorphic to G_1. We therefore define the manifold G_1 together with the actions of G by left and right multiplication to be the identity bibundle for a Lie groupoid G.

The functor U is far from being an equivalence of 2-categories. The issue is that for almost all groupoids G and H the functor

$$U : \mathsf{Hom}_{\mathsf{LieGpd}}(G, H) \to \mathsf{Hom}_{\mathsf{Bi}}(G, H) \tag{2-2}$$

fails to be essentially surjective. The failure of essential surjectivity follows from the well-known fact:

Lemma 2.7. *A bibundle $P : G \to H$ is isomorphic to a bibundle $\langle f\rangle$ for some functor $f : G \to H$ if and only if the left anchor $a^L_P : P \to G_0$ has a section.*

We omit the proof.

Remark 2.8. Recall that a functor $f : G \to H$ between two Lie groupoids is an *essential equivalence* if the right anchor $a^R : \langle f\rangle \to H_0$ is a surjective submersion and the action of G on $\langle f\rangle$ is principal. The functor $U =: \mathsf{LieGpd} \to \mathsf{Bi}$ is a localization of the 2-category LieGpd at the class of all essential equivalences; see Example 2.12.

In contrast to failure of the functor U to be surjective on 1-morphisms, for 2-morphisms the following result holds. The result must be known but we are not aware of a reference.

Theorem 2.9 (folklore). *For any pair of functors $f, k : G \to H$ of Lie groupoids the map*

$$U : \mathsf{Hom}_{\mathsf{LieGpd}}(f, k) \to \mathsf{Hom}_{\mathsf{Bi}}(\langle f\rangle, \langle k\rangle), \quad \alpha \mapsto \langle\alpha\rangle$$

is a bijection.

Sketch of proof. Let $\delta : \langle f\rangle \to \langle k\rangle$ be an isomorphism of bibundles. The left anchor $a^L_{\langle f\rangle} : \langle f\rangle \to G_0$ has a natural section σ_f. It is defined by

$$\sigma_f(x) = (x, f(x) \xleftarrow{1_{f(x)}}).$$

Similarly we have a natural section $\sigma_k : \langle k\rangle \to G_0$ of the left anchor $a^L_{\langle k\rangle} : \langle k\rangle \to G_0$. Since $a^L_{\langle k\rangle} : \langle k\rangle \to G_0$ is a principal H bundle, for any $x \in G_0$ there is a unique arrow $\bar{\delta}(x) \in H_1$ so that

$$\delta(\sigma_f(x)) = \sigma_k(x) \cdot \bar{\delta}(x)$$

for all $x \in G_0$. By equivariance of δ, the map

$$\bar{\delta} : G_0 \to H_1, \quad x \mapsto \bar{\delta}(x)$$

is a natural isomorphism from f to k. □

2B. *Localizations of bicategories.*

Definition 2.10. Given a bicategory C and a class of 1-morphisms W in C we define a *localization* of C at the class W (if it exists) to be a pair $(\mathsf{C}[W^{-1}], U : \mathsf{C} \to \mathsf{C}[W^{-1}])$, where $\mathsf{C}[W^{-1}]$ is a bicategory and $U : \mathsf{C} \to \mathsf{C}[W^{-1}]$ is a functor satisfying the following universal property: For any bicategory D the precomposition with U induces an equivalence of bicategories

$$\mathsf{Hom}(\mathsf{C}[W^{-1}], \mathsf{D}) \xrightarrow{-\circ U} \mathsf{Hom}_W(\mathsf{C}, \mathsf{D}),$$

where $\mathsf{Hom}_W(\mathsf{C}, \mathsf{D})$ denotes the bicategory of functors sending elements of W to weakly invertible 1-morphisms in D.

In particular given any functor $F : \mathsf{C} \to \mathsf{D}$ mapping elements of W to invertible morphisms of D there exists a functor $\widetilde{F} : \mathsf{C}[W^{-1}] \to \mathsf{D}$ and a natural isomorphism

$$F \Rightarrow \widetilde{F} \circ U.$$

Following a common abuse of notation, we often denote the localizations of C at W simply as $\mathsf{C}[W^{-1}]$ (and omit the localization functor U).

The localization $\mathsf{C}[W^{-1}]$ is defined up to equivalence of bicategories, so it will be convenient to refer to any functor $F : \mathsf{C} \to \mathsf{C}'$ between bicategories as a localization of C at the class W if it has the same universal property as the localization functor $U : \mathsf{C} \to \mathsf{C}[W^{-1}]$. Namely we ask that for any bicategory D the precomposition with F induces an equivalence of bicategories $\mathsf{Hom}(\mathsf{C}', \mathsf{D}) \xrightarrow{-\circ F} \mathsf{Hom}_W(\mathsf{C}, \mathsf{D})$. Pronk [1996] gave a criterion for a functor $F : \mathsf{C} \to \mathsf{C}'$ to be "the" localization of C at the class W:

Proposition 2.11 [Pronk 1996, Proposition 24]. *A functor $F : \mathsf{C} \to \mathsf{C}'$ between bicategories is a localization of* C *at the class W if*

(1) *F sends the elements of W to (weakly) invertible 1-morphisms in C';*

(2) *F is essentially surjective on objects;*

(3) *for every 1-morphism f in C' there are 1-morphisms w in W and g in C with a 2-morphism $F(g) \Rightarrow f \circ F(w)$;*

(4) *F is fully faithful on 2-morphisms.*

Example 2.12. The functor $U := \langle\ \rangle : \mathsf{LieGpd} \to \mathsf{Bi}$ is the localization of the 2-categories of Lie groupoids, functors and natural transformations at the class of essential equivalences.

Localizations of bicategories will come up several times in this paper.

2C. *2-vector spaces and Lie 2-algebras.*

Definition 2.13. A (real) 2-*vector space* [Baez and Crans 2004] is a category V internal to the category of (real)vector spaces. Hence $V = \{V_1 \rightrightarrows V_0\}$, where V_0 a vector space of objects, V_1 a vector space of morphisms, and all the structure maps (source, target, unit, and composition) are linear.

There is a 2-category 2Vect whose objects are 2-vector spaces, 1-morphisms are (linear) functors and 2-morphisms are (linear) natural transformations. There is a forgetful functor $\mathsf{2Vect} \to \mathsf{Cat}$ from the 2-category of 2-vector spaces to the 2-category Cat of categories that forgets the linear structure.

Remark 2.14. There is an equivalence of categories of 2-vector spaces and of 2-term chain complexes of vector spaces. See, for example, [Baez and Crans 2004]. A similar result was proved much earlier by Deligne [SGA 4_3 1973]. We remind the reader of how this equivalence is defined on objects. Given a 2-term complex $\partial : U \to W$ there is an action of the abelian group U on W given by

$$u \cdot w := \partial(u) + w \tag{2-3}$$

for all $u \in U, w \in W$. The corresponding action groupoid $\{U \times W \rightrightarrows W\}$ is a 2-vector space.

The converse is true as well: any 2-vector space $V = \{V_1 \rightrightarrows V_0\}$ is isomorphic to an action groupoid defined by the 2-term complex $\partial = t|_{\ker s} : \ker s \to V_0$. Here as before $s, t : V_1 \to V_0$ are the source and target map of the category V; see [Baez and Crans 2004] for a proof. In particular a category underlying a 2-vector space is a groupoid.

Definition 2.15. A strict *Lie* 2-*algebra* is a category internal to the category of Lie algebras (over the reals): the space of objects and morphisms of a Lie 2-algebra are ordinary Lie algebras and all the structure maps are maps of Lie algebras.

Notation 2.16. Categories internal to Lie algebras, internal functors and internal natural transformations form a strict 2-category which we denote by $\mathsf{Lie2Alg}_{\text{strict}}$.

Definition 2.17. (see, for example, [Frégier and Wagemann 2011, Definition 15]) A *crossed module* of Lie algebras consists of a Lie algebra homomorphism $\partial : \mathfrak{m} \to \mathfrak{n}$ together with a Lie algebra homomorphism

$$D : \mathfrak{n} \to \mathsf{Der}(\mathfrak{m})$$

from $\mathfrak{n}$ to the Lie algebra $\mathsf{Der}(\mathfrak{m})$ of derivations of $\mathfrak{m}$ so that, for all $m, m' \in \mathfrak{m}$, $n \in \mathfrak{n}$,

(i) $\partial(D(n)m) = [n, \partial(m)]$ and

(ii) $D(\partial(m))m' = [m, m']$.

A crossed module of Lie algebras determines a Lie 2-algebra: see, for example, the proof of Theorem 3 in [Frégier and Wagemann 2011]. The converse is true as well: any Lie 2-algebra canonically defines a crossed module of Lie 2-algebras. In fact more is true: crossed modules form a strict 2-category, and the 2-categories of Lie 2-algebras and of crossed modules are equivalent (see [Frégier and Wagemann 2011, Theorem 3]). We need the following result:

Lemma 2.18. *Let $V = \{V_1 \rightrightarrows V_0\}$ be a 2-vector space. Suppose the corresponding 2-term complex $\partial = t|_{\ker s} : \ker s \to V_0$ is part of the data of a Lie algebra crossed module. That is, suppose that V_0, $\ker s$ are Lie algebras, ∂ is a Lie algebra map, and that there is an action $D : V_0 \to \mathsf{Der}(\ker s)$ of V_0 on $\ker s$ by derivations making $(\partial : \ker s \to V_0, D : V_0 \to \mathsf{Der}(\ker s))$ into a crossed module of Lie algebras. Then V is a Lie 2-algebra.*

Sketch of proof. Since $1 \circ s = \mathrm{id}_{V_0}$, we have $V_1 = \ker s \oplus V_0$. We define a bracket on $\ker s \oplus V_0$ by

$$[(x_1, y_1), (x_2, y_2)] := \big([x_1, x_2] + D(y_1)x_2 - D(y_2)x_1, [y_1, y_2]\big) \tag{2-4}$$

for all $(x_1, y_1), (x_2, y_2) \in \ker s \oplus V_0$. That is, we define the Lie algebra V_1 to be the semidirect product of V_0 and $\ker s$. Checking that source, target and unit maps of V are Lie algebra maps is easy. To check that the composition $m : V_1 \times_{V_0} V_1 \to V_1$ in the category V is a Lie algebra map we observe that m is given by

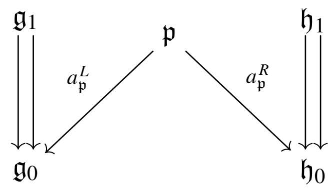
$$m((x_1, y_1), (x_2, y_2)) = (x_1 + x_2, y_2) \tag{2-5}$$

for all $(x_1, y_1), (x_2, y_2) \in \ker s \oplus V_0$ with $y_1 = t(x_2, y_2) = \partial x_2 + y_2$. This fact is not completely obvious. It lies in the heart of the correspondence between 2-vector spaces and 2-term chain complexes. See Remark 2.14 and [Baez and Crans 2004]. A computation now shows that the map m is a Lie algebra map. ☐

There is a problem with the 2-category $\mathsf{Lie2Alg}_{\mathrm{strict}}$ of Lie 2-algebras. Namely, suppose $f : \mathfrak{g} \to \mathfrak{h}$ is a morphism of Lie 2-algebras which is fully faithful and essentially surjective, that is, an essential equivalence. Then f has a weak inverse (as a functor), but there is no reason for that inverse to be a morphism of Lie 2-algebras. In fact it is easy to come up with examples where such morphism of Lie 2-algebras does not exist.

Fortunately the problem has a universal solution: we localize the 2-category $\mathsf{Lie2Alg}_{\mathrm{strict}}$ at the class of essential equivalences and obtain a bicategory $\mathsf{Lie2Alg}$ (see [Pronk 1996] and Section 2B above). This localization has a simple and explicit description: we define a morphism between Lie 2-algebras as a "bibundle internal to the category of Lie algebras." Here are the details.

Definition 2.19. A *bibundle* $\mathfrak{p} : \mathfrak{g} \to \mathfrak{h}$ from a Lie 2-algebra $\mathfrak{g}$ to a Lie 2-algebra $\mathfrak{h}$ is a Lie algebra $\mathfrak{p}$ with *left* and *right anchor maps* $a^L_{\mathfrak{p}}$ and $a^R_{\mathfrak{p}}$ (which are maps of Lie algebras),



along with a left action of the groupoid $\mathfrak{g}$ and right action of the groupoid $\mathfrak{h}$

$$\mathfrak{g}_1 \times_{s,\mathfrak{g}_0,a^L_{\mathfrak{p}}} \mathfrak{p} \to \mathfrak{p}, \quad (g, p) \mapsto g \cdot p, \qquad \mathfrak{p} \times_{a^R_{\mathfrak{p}},\mathfrak{h}_0,t} \mathfrak{h}_1 \to \mathfrak{p}, \quad (p, h) \mapsto p \cdot h.$$

We require that the actions are maps of Lie algebras and commute with each other. We require that $a^L_{\mathfrak{p}}$ is surjective. Finally, we require that the map

$$\mathfrak{p} \times_{a^R_{\mathfrak{p}},\mathfrak{h}_0,t} \mathfrak{h}_1 \to \mathfrak{p} \times_{a^L_{\mathfrak{p}},\mathfrak{g}_0,a^L_{\mathfrak{p}}} \mathfrak{p}, \quad (p, h) \mapsto (p, p \cdot h)$$

is an isomorphism of Lie algebras. Thus in particular we require that $a^L_{\mathfrak{p}} : \mathfrak{p} \to \mathfrak{g}_0$ is a principal $\mathfrak{h}$-bundle.

Remark 2.20. The composition of bibundles between Lie 2-algebras is defined in the same way as in the case of bibundles between Lie groupoids: it is the quotient of the appropriate fiber product. We will omit a proof that Lie 2-algebras, bibundles of Lie algebras and isomorphisms of bibundles form a bicategory. We denote this bicategory by $\mathsf{Lie2Alg}$. We note that biprincipal bibundles are weakly invertible in this bicategory.

As in the case of Lie groupoids there is a functor $\langle\ \rangle : \mathsf{Lie2Alg}_{\mathrm{strict}} \to \mathsf{Lie2Alg}$. It sends a strict map $f : \mathfrak{g} \to \mathfrak{h}$ of Lie 2-algebras to the bibundle

$$\langle f \rangle := \mathfrak{g}_0 \times_{f_0,\mathfrak{h}_0,t} \mathfrak{h}_1 := \{(x, \gamma) \mid f_0(x) = t(\gamma)\}.$$

Lemma 2.21. *Suppose $f : \mathfrak{g} \to \mathfrak{h}$ is a strict map of Lie 2-algebras whose underlying functor is fully faithful and essentially surjective. Then the bibundle of Lie 2-algebras*

$$\langle f \rangle : \mathfrak{g} \to \mathfrak{h}$$

is weakly invertible.

Proof. It is enough to show that $a^R_{\langle f\rangle} : \langle f \rangle \to \mathfrak{h}$ is an $\mathfrak{g}$-principal bundle. That is, it is enough to show that $a^R_{\langle f\rangle}$ is surjective and that the map

$$\varphi : \mathfrak{p} \times_{a^R_{\mathfrak{p}}, \mathfrak{h}_0, t} \mathfrak{h}_1 \to \mathfrak{p} \times_{a^L_{\mathfrak{p}}, \mathfrak{g}_0, a^L_{\mathfrak{p}}} \mathfrak{p}, \quad \varphi(p, h) := (p, p \cdot h)$$

is an isomorphism of Lie algebras. Since $a^R_{\langle f\rangle}(x, \gamma) = s(\gamma)$ the surjectivity of $a^R_{\langle f\rangle}$ is equivalent to the essential surjectivity of the functor f. The fullness of f translates into φ being onto and faithfulness of f translates into φ being 1-1. □

We now apply Proposition 2.11 to conclude that $\langle\ \rangle : \mathsf{Lie2Alg}_{\text{strict}} \to \mathsf{Lie2Alg}$ is the localization of $\mathsf{Lie2Alg}_{\text{strict}}$ at the class of essential equivalences. See also Theorem 4.4 below for a similar argument.

Remark 2.22. A reader familiar with Noohi's butterflies [2013] should not have much trouble showing that the bicategory $\mathsf{Lie2Alg}$ of Lie 2-algebras defined above is equivalent to the bicategory of crossed modules of Lie algebras, butterflies and isomorphisms of butterflies.

2D. *Tangent functors.* Recall that we have a functor T from the category Man of C^∞ manifolds to itself which sends a manifold M to its tangent bundle $TM \xrightarrow{\pi_M} M$. We also have a natural transformation $\pi : T \Rightarrow \mathrm{id}_{\mathsf{Man}}$ whose components are the projections $\pi_M : TM \to M$.

Remark 2.23. Recall that if a point c is a regular value of a smooth map $f : M \to N$ between two manifolds then $(Tf)^{-1}(c, 0) = T(f^{-1}(c))$. Consequently the tangent functor T preserves transverse fiber products.

The following result is well-known. We omit the proof.

Lemma 2.24. *The tangent functor $T : \mathsf{Man} \to \mathsf{Man}$ and the natural transformation $\pi : T \Rightarrow \mathrm{id}_M$ extend to a functor $T^{\mathsf{LieGpd}} : \mathsf{LieGpd} \to \mathsf{LieGpd}$ and to a natural transformation $\pi : T^{\mathsf{LieGpd}} \Rightarrow \mathrm{id}_{\mathsf{LieGpd}}$.*

Lemma 2.25. *The functor $T : \mathsf{Man} \to \mathsf{Man}$ extends to a functor $T^{\mathsf{Bi}} : \mathsf{Bi} \to \mathsf{Bi}$ from the bicategory Bi of Lie groupoids, bibundles and isomorphisms of bibundles to itself.*

Sketch of proof. For a Lie groupoid G we set $T^{\mathsf{Bi}}(G) = TG$. This defines T^{Bi} on 0-cells. Given a bibundle $P : G \to H$ the application of the tangent functor $T : \mathsf{Man} \to \mathsf{Man}$ gives us the bibundle $TP : TG \to TH$. Given two bibundles

$P, Q : G \to H$ and an isomorphism $\alpha : P \to Q$ of bibundles its derivative $T\alpha : TP \to TQ$ is also an isomorphism of bibundles. It is not hard to check that for any two Lie groupoids G and H the map $T : \mathsf{Hom}_{\mathsf{Bi}}(G, H) \to \mathsf{Hom}_{\mathsf{Bi}}(TG, TH)$ defined above is a functor. Given a Lie groupoid G we defined the identity bibundle $\langle \mathrm{id}_G \rangle$ to be the manifold G_1 together with the source and target maps as left and right anchors and left and right multiplications as left and right actions of G on G_1. Then $T\langle \mathrm{id}_G \rangle = TG_1 = \langle \mathrm{id}_{TG} \rangle$. Hence the comparison 2-cells $\mu_G : \mathrm{id}_{T^{\mathsf{Bi}}(G)} \to T^{\mathsf{Bi}}(\mathrm{id}_G)$ are identity 2-cells. We also need the comparison 2-cells $\mu_{Q,P} : T^{\mathsf{Bi}}Q \circ T^{\mathsf{Bi}}P \to T^{\mathsf{Bi}}(Q \circ P)$. They are constructed as follows. Given a pair $G \xrightarrow{P} H \xrightarrow{Q} K$ of composable bibundles $T(P \times_{H_0} Q) = TQ \times_{TH_0} TQ$ (cf. Remark 2.23). Additionally $T(P \times_{H_0} Q)/H \simeq \big(T(P \times_{H_0} Q)\big)/TH$ since for any Lie groupoid H and any H-principal bundle $R \to B$, TR/TH is isomorphic to TB. Consequently

$$T(Q\circ P)=T((P\times_{H_0}Q)/H)\simeq T(P\times_{H_0}Q)/TH\simeq(TP\times_{TH_0}TQ)/TH=TQ\circ TP.$$

This diffeomorphism is the desired invertible 2-cell $\mu_{Q,P}$. □

Lemma 2.26. *The functors $U \circ T^{\mathsf{LieGpd}}$ and $T^{\mathsf{Bi}} \circ U$ are isomorphic. Here $U := \langle\ \rangle : \mathsf{LieGpd} \to \mathsf{Bi}$ is the localization functor.*

Proof. We construct a pseudonatural transformation $\sigma : U \circ T^{\mathsf{LieGpd}} \Rightarrow T^{\mathsf{Bi}} \circ U$ as follows. Since for a groupoid G $U \circ T^{\mathsf{LieGpd}}(G) = TG = T^{\mathsf{Bi}} \circ U(G)$ we set $\sigma_G := \langle \mathrm{id}_{TG} \rangle$. Given a functor $f : G \to H$ we define the 2-cell

$$\sigma_f : \sigma_H \circ T^{\mathsf{Bi}}(U(f)) \Rightarrow U(T^{\mathsf{LieGpd}}(f)) \circ \sigma_G$$

as composite of the isomorphisms of bibundles $\langle \mathrm{id}_{TH} \rangle \circ T\langle f \rangle \to T\langle f \rangle \to \langle Tf \rangle \to \langle Tf \rangle \circ \langle \mathrm{id}_{TG} \rangle$. Here the middle arrow is the diffeomorphism $T(G_0 \times_{f,H_0,t} H_1) \xrightarrow{\simeq} TG_0 \times_{Tf,TH_0,Tt} TH_1$. Given an isomorphism $\alpha : f \Rightarrow f'$ between two functors $f, f' : G \to H$ the diagram

$$\begin{array}{ccc} T(G_0 \times_{f,H_0,t} H_1) & \xrightarrow{\sim} & TG_0 \times_{Tf,TH_0,Tt} TH_1 \\ {\scriptstyle T\langle\alpha\rangle}\downarrow & & \downarrow{\scriptstyle \langle T\alpha\rangle} \\ T(G_0 \times_{f',H_0,t} H_1) & \xrightarrow[\sim]{} & TG_0 \times_{Tf',TH_0,Tt} TH_1 \end{array}$$

commutes. It follows that σ_f's are components of a natural transformation

$$\sigma_{G,H} : (U \circ T^{\mathsf{LieGpd}}) \circ \sigma_G \Rightarrow \sigma_H \circ (T^{\mathsf{Bi}} \circ U).$$ □

Lemma 2.27. *There exists a transformation $\tilde{\pi} : T^{\mathsf{Bi}} \Rightarrow \mathrm{id}_{\mathsf{Bi}}$ whose 1-cells $\tilde{\pi}_G : T^{\mathsf{Bi}}(G) \to G$ are (isomorphic to) the bibundles $\langle \pi_G \rangle$, where as before $\pi_G : TG \to G$ are the projection functors.*

Proof. Since $U = \langle\ \rangle : \mathsf{LieGpd} \to \mathsf{Bi}$ is a localization functor, the pullback by U defines an equivalence of bicategories

$$U^* := - \circ U : \mathsf{Hom}(\mathsf{Bi}, \mathsf{Bi}) \to \mathsf{Hom}_W(\mathsf{LieGpd}, \mathsf{Bi}).$$

Here as before $\mathsf{Hom}_W(\mathsf{LieGpd}, \mathsf{Bi})$ denotes the bicategory of functors that send essential equivalence in LieGpd to invertible bibundles. Consequently for any two functors $F, G : \mathsf{Bi} \to \mathsf{Bi}$ the functor

$$U^* : \mathsf{Hom}(F, G) \to \mathsf{Hom}(F \circ U, G \circ U)$$

is an equivalence of categories. Note that the objects of $\mathsf{Hom}(F, G)$ are pseudonatural transformations and morphisms are modifications. In Lemma 2.24 we constructed a natural transformation $\pi : T^{\mathsf{LieGpd}} \to \mathrm{id}$ and Lemma 2.26 we constructed a natural isomorphism $\sigma : U \circ T^{\mathsf{LieGpd}} \Rightarrow T^{\mathsf{Bi}} \circ U$. Therefore we have a natural transformation $U\pi \circ \sigma^{-1} : T^{\mathsf{Bi}} \circ U \Rightarrow U$. Since $U^* : \mathsf{Hom}(T^{\mathsf{Bi}}, \mathrm{id}) \to \mathsf{Hom}(T^{\mathsf{Bi}} \circ U, U)$ is essentially surjective, there exists a pseudonatural transformation $\tilde{\pi} : T^{\mathsf{Bi}} \to \mathrm{id}$ so that $\tilde{\pi} \circ U$ differs from $U\pi \circ \sigma^{-1}$ by a modification. It will be convenient to fix one such modification throughout the paper. It follows that for each Lie groupoid G we have chosen an isomorphism of bibundles $\tilde{\pi}_G \to \langle \pi_G \rangle$ where $\tilde{\pi}_G$ is the component of the transformation $\tilde{\pi}$ at G. □

3. The Lie 2-algebra $\mathbb{X}(G)$ of vector fields on a Lie groupoid G

In this section we prove Theorem 3.4: the category of multiplicative vector fields on a Lie groupoid underlies a strict Lie 2-algebra.

Definition 3.1 [Hepworth 2009]. Consider a Lie groupoid G with its tangent groupoid $\pi_G : TG \to G$. The *category* $\mathbb{X}(G)$ *of multiplicative vector fields* is defined as follows. The set of *objects* of $\mathbb{X}(G)$ is

$$\mathbb{X}(G)_0 := \{v : G \to TG \mid v \text{ is a functor and } \pi_G \circ v = \mathrm{id}_G\}.$$

This is the set of multiplicative vector fields of Mackenzie and Xu [1998]. A *morphism* in $\mathbb{X}(G)$ from a multiplicative vector field v to a multiplicative vector field w is a natural transformation $\alpha : v \Rightarrow w$ such that for every point $x \in G_0$

$$\pi_G(\alpha_x) = 1_x. \tag{3-1}$$

The composition of morphisms is the vertical composition of natural transformations. Note that the category $\mathbb{X}(G)$ is a groupoid.

Notation 3.2. We denote the source and target maps in the category $\mathbb{X}(G)$ by $\mathfrak{s}$ and $\mathfrak{t}$, respectively. The unit map is denoted by $\mathbf{1}$, the inversion by $(\ )^{-1}$ and the composition/multiplication of morphisms by $\circ$.

Lemma 3.3. *The category of multiplicative vector fields $\mathbb{X}(G)$ on a Lie groupoid G is a 2-vector space.*

Proof. Mackenzie and Xu [1998] proved that the set $\mathbb{X}(G)_0$ of multiplicative vector fields is a real vector space.

We next argue that the set of morphisms $\mathbb{X}(G)_1$ of the category $\mathbb{X}(G)$ is a vector space as well. Suppose $\alpha_1 : v_1 \Rightarrow w_1$ and $\alpha_2 : v_2 \Rightarrow w_2$ are morphisms between multiplicative vector fields. Equation (3-1) says that α_1 and α_2 are both sections of the vector bundle $TG_1|_{G_0} \to G_0$ where we have suppressed the unit map $1_G : G_0 \to G_1$. Clearly the linear combination $\lambda_1\alpha_1 + \lambda_2\alpha_2$ is again a section of the bundle $TG_1|_{G_0} \to G_0$ for any choice of scalars λ_1, λ_2. We need to check that it is actually a natural transformation from $\lambda_1 v_1 + \lambda_2 v_2$ to $\lambda_1 w_1 + \lambda_2 w_2$. That is, we need to check that for any arrow $y \xleftarrow{\gamma} x$ in the groupoid G

$$(\lambda_1\alpha_1 + \lambda_2\alpha_2)_y \bullet ((\lambda_1 v_1 + \lambda_2 v_2)(\gamma)) = ((\lambda_1 w_1 + \lambda_2 w_2)(\gamma)) \bullet (\lambda_1\alpha_1 + \lambda_2\alpha_2)_x.$$

Here and below $\bullet : TG_1 \times_{TG_0} TG_1 \to TG_1$ denotes the multiplication in the Lie groupoid TG.

Since $\bullet$ is the derivative of the multiplication $m : G_1 \times_{G_0} G_1 \to G_1$ in the groupoid G, it is fiberwise linear: for any $(\gamma_2, \gamma_1) \in G_1 \times_{G_0} G_1$ and $(a_1, a_2), (b_1, b_2) \in T_{\gamma_2}G_1 \times_{TG_0} T_{\gamma_1}G_1 = T_{(\gamma_1,\gamma_2)}(G_1 \times_{G_0} G_1)$ we have (in the prefix notation)

$$\bullet\,(\lambda(a_1, a_2) + \mu(b_1, b_2)) = \lambda(\bullet(a_2, a_1)) + \mu(\bullet(b_1, b_2)) \tag{3-2}$$

for all scalars λ, μ. In the infix notation (3-2) reads

$$(\lambda a_1 + \mu b_1) \bullet (\lambda a_2 + \mu b_2) = \lambda(a_1 \bullet a_2) + \mu(b_1 \bullet b_2). \tag{3-3}$$

Hence

$$\begin{aligned}(\lambda_1\alpha_1 + \lambda_2\alpha_2)_y \bullet ((\lambda_1 v_1 + \lambda_2 v_2)(\gamma)) &= \lambda_1((\alpha_1)_y \bullet v_1(\gamma)) + \lambda_2((\alpha_2)_y \bullet v_2(\gamma)) \\ &= \lambda_1(w_1(\gamma) \bullet (\alpha_1)_x) + \lambda_2(w_2(\gamma) \bullet (\alpha_2)_x) \\ &= ((\lambda_1 w_1 + \lambda_2 w_2)(\gamma)) \bullet (\lambda_1\alpha_1 + \lambda_2\alpha_2)_x.\end{aligned}$$

Here the first and third equalities hold by (3-3). In the second equality we used the fact that $\alpha_1 : v_1 \Rightarrow w_1$ and $\alpha_2 : v_2 \Rightarrow w_2$ are natural transformations. Therefore the space of morphisms $\mathbb{X}(G)_1$ is a vector space.

Moreover the computation above shows that for $\lambda_1, \lambda_2 \in \mathbb{R}$, $\alpha_1 : v_1 \Rightarrow w_1$, $\alpha_2 : v_2 \Rightarrow w_2 \in \mathbb{X}(G)_1$ the source of $\lambda_1\alpha_1 + \lambda_2\alpha_2$ is $\lambda_1 v_1 + \lambda_2 v_2$. That is, the source map $\mathfrak{s} : \mathbb{X}(G)_1 \to \mathbb{X}(G)_0$ of the category $\mathbb{X}(G)$ is linear. Similarly the target map $\mathfrak{t}$ is linear. It is also easy to see that the unit map $\mathbb{X}(G)_0 \to \mathbb{X}(G)_1$ is linear as well.

Finally we need to check that multiplication/composition $\circ$ in the category $\mathbb{X}(G)$, which is the vertical composition of natural transformations, is linear as a map from $\mathbb{X}(G)_1 \times_{\mathbb{X}(G)_0} \mathbb{X}(G)_1$ to $\mathbb{X}(G)_1$. That is, we need to check that

$$(\lambda\alpha_2 + \mu\beta_2) \circ (\lambda\alpha_1 + \mu\beta_1) = \lambda(\alpha_2 \circ \alpha_1) + \mu(\beta_2 \circ \beta_1) \tag{3-4}$$

for all $\lambda, \mu \in \mathbb{R}$, $(\alpha_2, \alpha_1), (\beta_2, \beta_1) \in \mathbb{X}(G)_1 \times_{\mathbb{X}(G)_0} \mathbb{X}(G)_1$. Recall that the vertical composition $\circ$ is computed pointwise: for any composable natural transformations δ_1, δ_1 and any point $x \in G_0$ $(\delta_2 \circ \delta_1)_x = (\delta_2)_x \bullet (\delta_1)_x$, where as before $\bullet$ is the multiplication in TG. Since $\bullet$ is fiberwise linear (3-4) follows. □

Theorem 3.4. *The category of vector fields $\mathbb{X}(G)$ on a Lie groupoid G is a* (*strict*) *Lie 2-algebra.*

Proof. Recall the notation: $s : G_1 \to G_0$ is the source map for the groupoid G, its differential $Ts : TG_1 \to TG_0$ is the source map for the tangent groupoid TG. We use $\mathfrak{s}$, $\mathfrak{t}$ to denote the source and target maps of the groupoid $\mathbb{X}(G)$, respectively.

By Lemma 2.18 it is enough to: give the vector spaces $\ker(\mathfrak{s} : \mathbb{X}(G)_1 \to \mathbb{X}(G)_0)$ and $\mathbb{X}(G)_0$ the structure of Lie algebras; check that $\partial := \mathfrak{t}|_{\ker \mathfrak{s}} : \ker \mathfrak{s} \to \mathbb{X}(G)_0$ is a Lie algebra map; define an action $D : \mathbb{X}(G)_0 \to \mathsf{Der}(\ker \mathfrak{s})$ on $\ker \mathfrak{s}$ by derivations; and check the compatibility of ∂ and D:

$$\partial(D(X)\alpha) = [X, \partial(\alpha)], \tag{3-5}$$

$$D(\partial\alpha_1)\alpha_2 = [\alpha_1, \alpha_2] \tag{3-6}$$

for all $\alpha, \alpha_1, \alpha_2 \in \ker \mathfrak{s}$ and all multiplicative vector fields X on the Lie groupoid G (compare with Definition 2.17).

The fact that the vector space $\mathbb{X}(G)_0$ of multiplicative vector fields carries a Lie bracket is due to Mackenzie and Xu [1998]. We argue next that $\ker \mathfrak{s}$ is the space of sections of the Lie algebroid $A_G \to G_0$. By definition of the source map $\mathfrak{s}$, $\ker \mathfrak{s} = \{\alpha : X \Rightarrow Y \mid X = 0\}$. Therefore $\alpha \in \ker \mathfrak{s}$ if and only if there is a multiplicative vector field Y so that the diagram

$$\begin{array}{ccc} 0_x & \xrightarrow{0_\gamma} & 0_y \\ \alpha_x \downarrow & & \downarrow \alpha_y \\ Y(x) & \xrightarrow[Y(\gamma)]{} & Y(y) \end{array} \tag{3-7}$$

commutes for all arrows $y \xleftarrow{\gamma} x$ in G_1. Hence if $\alpha \in \ker \mathfrak{s}$ then $Ts(\alpha_x) = 0_x$ for all $x \in G_0$. That is, α is a section of $A_G \to G_0$. Conversely if $\alpha : G_0 \to A_G$ is a section of the Lie algebroid we can define a multiplicative vector field $Y : G \to TG$ so that (3-7) commutes. Namely on objects we define $Y(x) := Tt(\alpha_x)$ for all $x \in G_0$.

And for $y \xleftarrow{\gamma} x$ in G_1 we set

$$Y(\gamma) = \alpha_y \bullet 0_\gamma \bullet (\alpha_x)^{-1}. \tag{3-8}$$

Here as before $\bullet$ is the multiplication in TG and $()^{-1} = Ti : TG_1 \to TG_1$ is the inverse map, which is the derivative of the inverse map i of the groupoid G. We conclude that $\ker(\mathfrak{s} : \mathbb{X}(G)_1 \to \mathbb{X}(G)_0) = \Gamma(A_G)$ and that $\partial := \mathfrak{t}|_{\ker \mathfrak{s}} : \ker \mathfrak{s} \to \mathbb{X}(G)_0$ is given by

$$(\partial\alpha)(\gamma) = \alpha_{t(\gamma)} \bullet 0_\gamma \bullet (\alpha_{s(\gamma)})^{-1} \tag{3-9}$$

for all $\gamma \in G_1$. Note that (3-9) can be written as

$$\partial\alpha = \vec{\alpha} + \overleftarrow{\alpha}. \tag{3-10}$$

where $\vec{\alpha}(\gamma) = TR_\gamma\, \alpha(t(\gamma))$ and $\overleftarrow{\alpha}(\gamma) = T(L_\gamma \circ i)\, \alpha(t(\gamma))$ for all $\gamma \in G_1$. Here R_γ and L_γ are right and left multiplications by γ, respectively. The bracket on the space of sections $\Gamma(A_G)$ of the Lie algebroid $A_G \to G_0$ is defined by requiring that the injective map $^{\to} : \Gamma(A_G) \to \Gamma(TG_1)$, $\alpha \mapsto \vec{\alpha}$ is a map of Lie algebras. Consequently $^{\leftarrow} : \Gamma(A_G) \to \Gamma(TG_1)$, $\alpha \mapsto \overleftarrow{\alpha}$ is also a map of Lie algebras. Since left- and right-invariant vector fields commute (cf. [Mackenzie 2005]) we conclude that $\partial = \mathfrak{t}|_{\ker \mathfrak{s}} : \ker \mathfrak{s} = \Gamma(A_G) \to \mathbb{X}(G)_0$ is a Lie algebra map.

Following Mackenzie and Xu we define the map D from the space $\mathbb{X}(G)_0$ of multiplicative vector fields to $\mathsf{Hom}(\Gamma(TG_1|_{G_0}), \Gamma(TG_1|_{G_0}))$ by setting $D(X)\alpha := [X, \vec{\alpha}]|_{G_0}$ for all multiplicative vector fields X and all sections $\alpha \in \Gamma(A_G)$. Mackenzie and Xu [1998, Proposition 3.7] proved that $[X, \vec{\alpha}]$ is tangent to the fibers of s and is right invariant. Hence $[X, \vec{\alpha}]|_{G_0}$ is a section of the Lie algebroid $A_G \to G_0$. They furthermore showed that $D(X)$ is a derivation of $\Gamma(A_G)$ and that $D : \mathbb{X}(G)_0 \to \mathsf{Der}(\Gamma(A_G))$ is a map of Lie algebras [Mackenzie and Xu 1998, Proposition 3.8]. Since left- and right-invariant vector fields commute, for any $\alpha_1, \alpha_2 \in \Gamma(A_G)$ we have $[\partial\alpha_1, \vec{\alpha}_2] = [\vec{\alpha}_1 + \overleftarrow{\alpha}_1, \vec{\alpha}_2] = [\vec{\alpha}_1, \vec{\alpha}_2]$ and (3-6) follows.

We end the proof by showing that (3-5) holds. On the right-hand side we have $[X, \partial\alpha] = [X, \vec{\alpha} + \overleftarrow{\alpha}] = [X, \vec{\alpha}] + [X, \overleftarrow{\alpha}]$ while on the left, $\partial(D(X)\,\alpha) = (D(X)\,\alpha)^{\to} + (D(X)\,\alpha)^{\leftarrow}$. By definition of D, $(D(X)\,\alpha)^{\to} = [X, \vec{\alpha}]$, so it remains to prove that $[X, \overleftarrow{\alpha}] = (D(X)\,\alpha)^{\leftarrow}$. Since X is a functor, $Ti \circ X = X \circ i$. The inversion map i relates right- and left-invariant vector fields. That is, $Ti \circ \vec{\alpha} = \overleftarrow{\alpha} \circ i$ for all α. Consequently

$$\begin{aligned}(D(X)\,\alpha)^{\leftarrow}(g) &= T(L_g \circ i)(D(X)\,\alpha)(1_{s(g)}) \\ &= T(L_g)Ti[X, \vec{\alpha}](1_{s(g)}) = TL_g[X, \overleftarrow{\alpha}](i(1_{s(g)})).\end{aligned}$$

Since $[X, \overleftarrow{\alpha}]$ is left-invariant, $TL_g[X, \overleftarrow{\alpha}](i(1_{s(g)})) = [X, \overleftarrow{\alpha}](g)$. It follows that $(D(X)\,\alpha)^{\leftarrow}(g) = [X, \overleftarrow{\alpha}](g)$ for all $g \in G_1$ and we are done. □

4. Morita invariance of the Lie 2-algebra of vector fields

The goal of this section is to prove:

Theorem 4.1. *The assignment $G \mapsto \mathbb{X}(G)$ of the category of vector fields to a Lie groupoid extends to a functor*

$$\mathbb{X} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{Lie2Alg} \tag{4-1}$$

from the bicategory $\mathsf{Bi}_{\mathrm{iso}}$ *of Lie groupoids, invertible bibundles and isomorphisms of bibundles to the bicategory* $\mathsf{Lie2Alg}$ *of Lie* 2*-algebras. Hence, in particular, if* $P : G \to H$ *is a Morita equivalence of Lie groupoids then* $\mathbb{X}(P) : \mathbb{X}(G) \to \mathbb{X}(H)$ *is a (weakly) invertible* 1*-morphism of Lie* 2*-algebras in the bicategory* $\mathsf{Lie2Alg}$.

Our strategy for constructing the functor $\mathbb{X}$ is to first construct it on a simpler category.

Definition 4.2. An *essentially surjective open embedding* of Lie groupoids is a functor $f : U \to G$ such that

(1) the maps on objects $f_0 : U_0 \to G_0$ and on morphisms $f_1 : U_1 \to G_1$ are open embeddings and
(2) the functor f is an essential equivalence, i.e., the corresponding bibundle

$$\langle f \rangle := U_0 \times_{f_0, G_0, t} G_1 : U \to G$$

is weakly invertible. Equivalently $\langle a^R \rangle : \langle f \rangle \to G_0$ is a principal U-bundle.

Remark 4.3. It is clear that Lie groupoids, essentially surjective open embeddings and natural transformations form a strict 2-category which we denote by $\mathscr{E}mb$.

Theorem 4.4. *The localization of the* 2*-category* $\mathscr{E}mb$ *at the class* W *of all* 1*-morphisms is the bicategory* $\mathsf{Bi}_{\mathrm{iso}}$ *of bicategory of Lie groupoids, invertible bibundles and isomorphisms of bibundles.*

Proof. We apply Proposition 2.11. Consider the localization functor $\langle\ \rangle : \mathsf{LieGpd} \to \mathsf{Bi}$ introduced in Remark 2.5. By definition of the 2-category $\mathscr{E}mb$ the restriction of the functor $\langle\ \rangle$ to $\mathscr{E}mb$ sends every 1-morphism $w : U \to G$ of $\mathscr{E}mb$ to an invertible bibundle $\langle w \rangle$ (and a 2-morphism to an isomorphism of bibundles). This gives us a functor

$$\langle\ \rangle : \mathscr{E}mb \to \mathsf{Bi}_{\mathrm{iso}}, \quad (G \xrightarrow{w} H) \mapsto (G \xrightarrow{\langle w \rangle} H). \tag{4-2}$$

The functor is surjective on objects. By Theorem 2.9 the functor is fully faithful on 2-morphisms.

It remains to check that given an invertible bibundle $P : G \to H$ there exist essentially surjective open embeddings w_G, w_H so that $\langle w_H \rangle \circ P$ is isomorphic to $\langle w_G \rangle$. Since the bibundle P is weakly invertible, it gives rise to the *linking* *groupoid* [Weinstein 2009, Proposition 4.3], denoted $G *_P H$ and recalled presently. The manifold of objects $(G *_P H)_0$ is the disjoint union $G_0 \sqcup H_0$ of the objects of the groupoids G and H. The manifold of arrows $(G *_P H)_1$ is the disjoint union $G_1 \sqcup P \sqcup P^{-1} \sqcup H_1$. We think of the manifold P as the space of arrows from the points of H_0 to the points of G_0. We think of the elements of P^{-1} as the inverses of the elements of P. The multiplication in $G *_P H$ comes from the multiplications in the groupoids G and H and the actions of G and H on P and on P^{-1}. The inclusion $w_G : G \to G *_P H$ is given by the open embeddings

$$G_0 \hookrightarrow G_0 \sqcup H_0, \quad G_1 \hookrightarrow G_1 \sqcup P \sqcup P^{-1} \sqcup H_1.$$

It is easy to see that w_G is an essential equivalence, i.e., that the bibundle $\langle w_G \rangle$ is biprincipal, hence weakly invertible. Similarly we have the essentially surjective open embedding $w_H : H \hookrightarrow G *_P H$. A computation shows that the bibundles $\langle w_H \rangle \circ P$ and $\langle w_G \rangle$ are isomorphic. □

Proposition 4.5. *The assignment $G \mapsto \mathbb{X}(G)$ of the Lie 2-algebra of vector fields to a Lie groupoid extends to a contravariant functor*

$$(\mathscr{E}mb)^{op} \to \mathsf{Lie2Alg}_{\mathrm{strict}}, \quad (G \xrightarrow{w} H) \mapsto (\mathbb{X}(H) \xrightarrow{w^*} \mathbb{X}(G)) \tag{4-3}$$

from the 2-category $\mathscr{E}mb$ of Lie groupoids, essentially surjective open embeddings and natural isomorphism to the strict 2-category $\mathsf{Lie2Alg}_{\mathrm{strict}}$ of Lie 2-algebras.

Proof. Consider an essentially surjective open embedding $w : G \to H$. Then $w(G) \subset H$ is an open Lie subgroupoid and $w : G \to w(G)$ is an isomorphism of Lie groupoids. We now assume without any loss of generality that G is an open subgroupoid of H. Then the tangent bundle TG is an open subgroupoid of TH. Moreover, any multiplicative vector field $v : H \to TH$ restricts to a multiplicative vector field $v|_G : G \to TG$. Similarly, a morphism $\alpha : v \Rightarrow u$ of multiplicative vector fields restricts to a morphism $\alpha|_G : v|_G \Rightarrow u|_G$. This gives us a functor

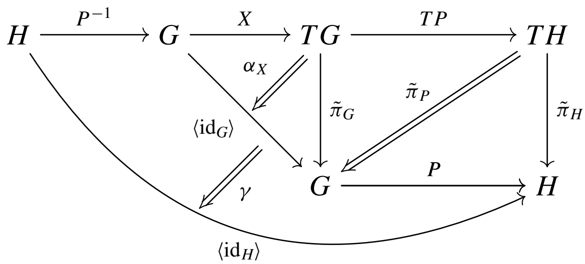
$$w^* : \mathbb{X}(H) \to \mathbb{X}(G), \quad w^*(\alpha : v \Rightarrow u) = (\alpha|_G : v|_G \Rightarrow u|_G). \tag{4-4}$$

The restriction to an open subgroupoid is a map of 2-vector spaces and preserves the brackets. Hence (4-4) is a map of Lie 2-algebras. □

Definition 4.6 (the category $\mathbb{X}_{\mathsf{gen}}(G)$ of *generalized vector fields on a Lie groupoid* G). Recall that there is a natural transformation $\tilde{\pi} : T^{\mathsf{Bi}} \Rightarrow \mathrm{id}_{\mathsf{Bi}}$ (Lemmas 2.25 and 2.27). An object of the category of generalized vector fields $\mathbb{X}_{\mathsf{gen}}(G)$ on a Lie groupoid G is a pair (P, α_P) where $P : G \to TG$ is a bibundle and $\alpha_P : \tilde{\pi}_G \circ P \Rightarrow \langle \mathrm{id}_G \rangle$ an isomorphism of bibundles. A morphism β in $\mathbb{X}_{\mathsf{gen}}(G)$ from (P, α_P) to (Q, α_Q) is a map of bibundles $\beta : P \Rightarrow Q$ so that $\alpha_Q = \alpha_P \circ (\tilde{\pi}_G \star \beta)$. Here as before $\star$ denotes whiskering in Bi, and $\circ$ is the composition of isomorphisms of bibundles.

Lemma 4.7. *A weakly invertible bibundle $P : G \to H$ between two Lie groupoids induces an equivalence of categories $P_* : \mathbb{X}_{\mathsf{gen}}(G) \to \mathbb{X}_{\mathsf{gen}}(H)$ between the corresponding categories of generalized vector fields.*

Proof. Since the 1-morphism P is (weakly) invertible, there is a 2-morphism $\gamma : (P \circ \langle \mathrm{id}_G \rangle) \circ P^{-1} \Rightarrow \langle \mathrm{id}_H \rangle$. Given an object (X, α_X) of $\mathbb{X}_{\mathsf{gen}}(G)$ we define $P_* X := TP \circ (X \circ P^{-1})$ The 2-morphism $\alpha_{P_* X} : \tilde{\pi}_H \circ P_* X \Rightarrow \langle \mathrm{id}_H \rangle$ comes from the 2-commutative diagram



We set

$$\alpha_{P_* X} := \gamma \circ (P \star \alpha_X \star P^{-1}) \circ (\tilde{\pi}_P \star (X \circ P^{-1})).$$

Given a morphism $\beta : (X, \alpha_X) \to (Y, \alpha_Y)$ in the category $\mathbb{X}_{\mathsf{gen}}(G)$ we define

$$P_* \beta := TP \star \beta \star P^{-1}.$$

A diagram chase ensures that $P_* \beta$ is a morphism in $\mathbb{X}_{\mathsf{gen}}(H)$ from $(P_* X, \alpha_{P_* X})$ to $(P_* Y, \alpha_{P_* Y})$. Finally one checks that the functor $(P^{-1})_* : \mathbb{X}(H) \to \mathbb{X}(G)$ is a weak inverse of P_*. Hence P_* is an equivalence of categories as claimed. ☐

Definition 4.8. The "inclusion" functor $\iota_G : \mathbb{X}(G) \hookrightarrow \mathbb{X}_{\mathsf{gen}}(G)$ is defined as follows. Suppose $v : G \to TG$ is a multiplicative vector field. Since $U : \mathsf{LieGpd} \to \mathsf{Bi}$ is a functor the composition $\langle \pi_G \rangle \circ \langle v \rangle$ is isomorphic to $\langle \pi_g \circ v \rangle = \langle \mathrm{id}_G \rangle$. Since we fixed the modification from $\tilde{\pi}$ to $\pi \circ U$ there is a canonical isomorphism of bibundles $\tilde{\pi}_G \to \langle \pi_G \rangle$. Consequently there is a canonical isomorphism $\alpha_{\langle v \rangle} : \tilde{\pi}_G \circ \langle v \rangle \Rightarrow \langle \mathrm{id}_G \rangle$. We define $\iota_G(v) := (\langle v \rangle, \alpha_{\langle v \rangle})$. Given a morphism $v \xrightarrow{\gamma} v'$ in $\mathbb{X}(G)$ we define $\iota_G(\gamma) = \langle \gamma \rangle$; it is a morphism in $\mathbb{X}_{\mathsf{gen}}(G)$ from $(\langle v \rangle, \alpha_{\langle v \rangle})$ to $(\langle v' \rangle, \alpha_{\langle v' \rangle})$.

Theorem 4.9. *For any Lie groupoid G the inclusion functor $\iota_G : \mathbb{X}(G) \hookrightarrow \mathbb{X}_{\mathsf{gen}}(G)$ defined above is an equivalence of categories.*

In the case where the groupoid G is proper Theorem 4.9 can be deduced from [Hepworth 2009, Theorem 15]. The proof of Theorem 4.9 is technical; we carry it out in Section 7 below.

Lemma 4.10. *Let $w : G \to G'$ be an essentially surjective open embedding. Then (the 1-morphism of categories underlying) the pull-back/restriction functor $w^* : \mathbb{X}(G') \to \mathbb{X}(G)$, which is given by $w^*(v \xrightarrow{\gamma} v') = (v|_G \xrightarrow{\gamma|_G} v'|_G)$, is fully faithful and essentially surjective.*

Proof. We first argue that w^* is fully faithful. We want to show that given a morphism $\delta : v|_G \to v'|_G$ in $\mathbb{X}(G)$ there exists a unique arrow $\tilde{\delta} : v \to v'$ in $\mathbb{X}(G')$ so that $\tilde{\delta}|_G = \delta$.

We deal with uniqueness first. Suppose $\gamma, \gamma' : v \to v'$ are two morphisms in $\mathbb{X}(G')$ with $\gamma|_G = \gamma'|G$. Fix an object y of G'. Since $w : G \to G'$ is essentially surjective, for any object y of G' there is an arrow $\mu : y \to x$ with x an object of G. Then $\gamma_x = \gamma'_x$ and therefore $\gamma_y = v'(\mu^{-1}) \circ \gamma_x \circ v(\mu) = v'(\mu^{-1}) \circ \gamma'_x \circ v(\mu) = \gamma'_y$. We conclude that $\gamma = \gamma'$. Now given $\delta : v|_G \to v'|_G$ we define $\tilde{\delta} : v \to v'$ at an object y by choosing as above $\mu : y \to x$ and setting $\tilde{\delta}_y = v'(\mu^{-1}) \circ \delta_x \circ v(\mu)$. If $\nu : y \to x$ is another arrow then $v'(\nu \circ \mu^{-1}) \circ \delta_x = \delta_x \circ v(\nu \circ \mu^{-1})$. Therefore $v'(\mu^{-1}) \circ \delta_x \circ v(\mu) = v'(\nu^{-1}) \circ \delta_x \circ v(\nu)$. Moreover the dependence of $\tilde{\delta}$ on y is smooth: since $w : G \to G'$ is an essential equivalence the right anchor $a^R_{\langle w \rangle} : \langle w \rangle \to G'_0$ is a surjective submersion. Note that $\langle w \rangle = t^{-1}(G_0) \subset G'_1$ and $a^R_{\langle w \rangle}(\mu) = s(\mu)$. Choose a local section $\sigma : U \to \langle w \rangle$ of $a^R_{\langle w \rangle}$ with $y \in U$. Then for all points $z \in U$ $\tilde{\delta}(z) = v'((\sigma(z))^{-1}) \circ \delta_{t(\sigma(y))} \circ v(\sigma(z))$, which is smooth. We conclude that w^* is fully faithful.

To prove essential surjectivity we need to argue that for any multiplicative vector field $u : G \to TG$ there is a multiplicative vector field $\tilde{u} : G' \to TG'$ and an isomorphism $\tilde{u}|_G \xrightarrow{\sim} u$. The functor $\iota_{G'} : \mathbb{X}(G') \to \mathbb{X}_{\mathsf{gen}}(G')$ is an equivalence of categories by Theorem 4.9. The functor $(\langle w \rangle)_* : \mathbb{X}_{\mathsf{gen}}(G) \to \mathbb{X}_{\mathsf{gen}}(G')$ is an equivalence of categories by Lemma 4.7. Therefore for any $u \in \mathbb{X}(G)_0$ there is a vector field $\tilde{u} \in \mathbb{X}(G')$ and an isomorphism $\langle \tilde{u} \rangle \xrightarrow{\sim} \langle w \rangle_* \langle u \rangle$.

The diagram

$$\begin{array}{ccc} TG & \xrightarrow{Tw} & TG' \\ {\scriptstyle \tilde{u}|_G}\uparrow & & \uparrow{\scriptstyle \tilde{u}} \\ G & \xrightarrow[w]{} & G' \end{array}$$

commutes. This together with the fact that $U : \mathsf{LieGpd} \to \mathsf{Bi}$ is a functor and $T\langle w \rangle$ is isomorphic to $\langle Tw \rangle$ shows that there exists an isomorphism $T\langle w \rangle \circ \langle \tilde{u}|_G \rangle \xrightarrow{\sim} \langle \tilde{u} \rangle \circ \langle w \rangle$. Thus $\langle \tilde{u} \rangle$ is isomorphic to $T\langle w \rangle \circ (\langle \tilde{u}|_G \rangle \circ \langle w \rangle^{-1})$. Hence $T\langle w \rangle \circ (\langle \tilde{u}|_G \rangle \circ \langle w \rangle^{-1}) \xrightarrow{\sim} T\langle w \rangle \circ (\langle u \rangle \circ \langle w \rangle^{-1})$. Consequently there is an isomorphism $\beta : \langle \tilde{u}|_G \rangle \to \langle u \rangle$ of bibundles. Note that this *does not* yet imply that the generalized vector fields $(\langle \tilde{u}|_G \rangle, \alpha_{\langle \tilde{u}|_G \rangle})$ and $(\langle u \rangle, \alpha_{\langle u \rangle})$ are isomorphic in the category $\mathbb{X}_{\mathsf{gen}}(G)$. The issue is $(\langle \tilde{u}|_G \rangle, \alpha_{\langle \tilde{u}|_G \rangle})$ need not equal $\alpha_{\langle u \rangle} \circ (\tilde{\pi}_G \star \beta)$. But $\beta : (\langle \tilde{u}|_G \rangle, \tau) \to (\langle u \rangle, \alpha_{\langle u \rangle})$ is a morphism in $\mathbb{X}_{\mathsf{gen}}(G)$ if we set $\tau = \alpha_{\langle u \rangle} \circ (\tilde{\pi}_G \star \beta)$. By Lemma 7.4 $(\langle \tilde{u}|_G \rangle, \tau)$ is isomorphic to $(\langle \tilde{u}|_G\rangle, \alpha_{\rangle \tilde{u}|_G\rangle})$. Consequently $\iota_G(\tilde{u}|_G)$ is isomorphic to $\iota_G(u)$. Since ι_G is an equivalence of categories $\tilde{u}|_G$ is isomorphic to u and we are done. □

Lemma 4.11. *The functor (4-3) of Proposition 4.5 takes every essentially surjective open embedding to a weakly invertible* 1*-morphism of Lie* 2*-algebras.*

Proof. By Lemma 4.10 the functor $w^* : \mathbb{X}(G') \to \mathbb{X}(G)$ associated to an essentially surjective open embedding $w : G \to G'$ is fully faithful and essentially surjective, hence an essential equivalence of Lie 2-algebras. The localization functor $\langle\ \rangle : \mathsf{Lie2Alg}_{strict} \to \mathsf{Lie2Alg}$ takes all essential equivalences to weakly invertible 1-morphisms. □

We are now in position to extend the assignment $G \mapsto \mathbb{X}(G)$ to a (covariant) functor $\mathbb{X} : \mathscr{E}mb \to \mathsf{Lie2Alg}$.

Definition 4.12. We define the functor $\mathbb{X} : \mathscr{E}mb \to \mathsf{Lie2Alg}$ on objects to be the assignment $G \mapsto \mathbb{X}(G)$. Given an essentially surjective open embedding $G \xrightarrow{w} G'$, the bibundle $\langle w^*\rangle$ is weakly invertible in $\mathsf{Lie2Alg}$ by Lemma 4.11. We set $\mathbb{X}(w) := (\langle w^*\rangle)^{-1}$.

Proof of Theorem 4.1. Since $\langle\ \rangle : \mathscr{E}mb \to \mathsf{Bi}_{\mathrm{iso}}$ is a localization of the bicategory $\mathscr{E}mb$ at the class $\mathscr{E}mb_1$ of all 1-morphisms and since the functor $\mathbb{X} : \mathscr{E}mb \to \mathsf{Lie2Alg}$ sends every 1-morphism of $\mathscr{E}mb$ to an invertible morphism there exists by Proposition 2.11 a functor $\widetilde{\mathbb{X}} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{Lie2Alg}$ (which is unique up to isomorphism) and an isomorphism $\widetilde{\mathbb{X}} \circ \langle\ \rangle \overset{\sim}{\Longleftrightarrow} \mathbb{X}$ of functors. It is no loss of generality to assume that $\widetilde{\mathbb{X}}(G) = \mathbb{X}(G)$ for every Lie groupoid G. We now drop the $\widetilde{\ }$ and obtain the desired functor $\mathbb{X} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{Lie2Alg}$. □

We end the section with a result that we will need in Section 5.

Lemma 4.13. *Let* $w : G \to G'$ *be an essentially surjective open embedding,* $w^* : \mathbb{X}(G') \to \mathbb{X}(G)$ *the pull-back/restriction functor of Proposition 4.5 and* $(\langle w\rangle)_* : \mathbb{X}_{\mathsf{gen}}(G) \to \mathbb{X}_{\mathsf{gen}}(G')$ *the push-forward along the bibundle* $\langle w\rangle$ *constructed in Lemma 4.7. Then the diagram*

$$\begin{array}{ccc} \mathbb{X}(G) & \xrightarrow{\quad \iota_G \quad} & \mathbb{X}_{\mathsf{gen}}(G) \\ {\scriptstyle w^*}\uparrow & \Downarrow \sigma & \downarrow {\scriptstyle (\langle w\rangle)_*} \\ \mathbb{X}(G') & \xrightarrow[\quad \iota_{G'} \quad]{} & \mathbb{X}_{\mathsf{gen}}(G') \end{array} \tag{4-5}$$

2*-commutes.*

Proof. We will show that the functors $(\langle w\rangle)_* \circ \iota_G \circ w^*$ and $\iota_{G'}$ are isomorphic by directly constructing a natural isomorphism $\sigma : (\langle w\rangle)_* \circ \iota_G \circ w^* \Rightarrow \iota_{G'}$.

We have already seen in the proof of Lemma 4.10 that for any multiplicative vector field $u : G' \to TG'$ the bibundles $(\langle w\rangle)_*(\langle u|_G\rangle) = T\langle w\rangle \circ (\langle u|_G\rangle \circ \langle w\rangle^{-1})$ and $\langle u\rangle$ are isomorphic. We need to be more precise about choosing these isomorphisms: we have to make sure that they are isomorphisms in $\mathbb{X}_{\mathsf{gen}}(G')$ from $(\langle w\rangle)_*(\langle u|_G\rangle, \alpha_{\langle u|_G\rangle})$ to $(\langle u\rangle, \alpha_{\langle u\rangle})$ and that they assemble into a natural transformation. As before we may assume that G is an open subgroupoid of G' and $w : G \to G'$ is an inclusion.

The category $\mathbb{X}_{\mathsf{gen}}(G')$ implicitly depends on the component $\tilde{\pi}_{G'}$ of the transformation $\tilde{\pi} : T^{\mathsf{Bi}} \Rightarrow \mathrm{id}_{\mathsf{Bi}}$. Recall that the bibundle $\tilde{\pi}_{G'}$ is isomorphic to the bibundle $\langle \pi_{G'}\rangle$. Replacing $\tilde{\pi}_{G'}$ in the definition of $\mathbb{X}_{\mathsf{gen}}(G')$ by $\langle \pi_{G'}\rangle$ and of $\tilde{\pi}_G$ by $\langle \pi_G\rangle$ in the definition of $\mathbb{X}_{\mathsf{gen}}(G)$ results in isomorphic categories. Consequently we may assume that $\tilde{\pi}_{G'} = \langle \pi_{G'}\rangle$ and $\tilde{\pi}_G = \langle \pi_G\rangle$.

By definition $\langle w\rangle$ is the fiber product $G_0 \times_{w, G'_0, t} G'_1$. Here as usual $t : G'_1 \to G'_0$ is the target map. Since the fiber product is defined by its universal property we may assume that $\langle w\rangle = t^{-1}(G_0) = \{\gamma \in G'_1 \mid t(\gamma) \in G_0\}$. Then the left anchor $a^L_{\langle w\rangle}$ is the restriction of the target map t and the right anchor is the restrictions of the source map s. The inverse of $\langle w\rangle$ is then $t^{-1}(G_0)$ with the left and right anchors reversed.

Some notation for elements of the composites of two bibundles is necessary. Given two composable bibundles $A \xrightarrow{P} B \xrightarrow{Q} C$ their composite is $Q \circ P = (P \times_{a^R_P, B_0, a^L_Q} Q)/B$. So an element of $Q \circ P$ is the B-orbit $[p, q]$ of an element (p, q) in the fiber product $P \times_{a^R_P, B_0, a^L_Q} Q \subset P \times Q$. For example given a multiplicative vector field $u : G' \to TG'$, an element of the composite $\langle \pi_{G'}\rangle \circ \langle u\rangle$ is of the form $[(x, u(x) \xleftarrow{\mu} \dot{y}), (\dot{y}, \pi_{G'}(\dot{y}) \xleftarrow{\nu} z)]$ for some $x \in G'_0$, $\mu \in TG_1$, $\dot{y} \in TG_0$, $\nu \in G'_1$ and $z \in G'_0$. The natural isomorphism $\alpha_{\langle u\rangle} : \langle \pi_{G'}\rangle \circ \langle u\rangle \to G'_1 = \langle \mathrm{id}_{G'}\rangle$ is then given by

$$\alpha_{\langle u\rangle}\big([(x, u(x) \xleftarrow{\mu} \dot{y}), (\dot{y}, \pi_{G'}(\dot{y}) \xleftarrow{\nu} z)]\big) = x \xleftarrow{\pi_{G'}(\mu)\circ\nu} z.$$

For any morphism $\beta : u \Rightarrow v$ in the category $\mathbb{X}(G')$ the isomorphism of bibundles $\langle \beta\rangle : \langle u\rangle \to \langle v\rangle$ is given by

$$\langle \beta\rangle(x, u(x) \xleftarrow{\Gamma} \dot{z}) = (x, v(x) \xleftarrow{\beta(x)\circ\Gamma} \dot{z}).$$

Given an element $[[y \xrightarrow{\gamma} x, (x, u(x) \xleftarrow{\dot{\mu}} \dot{z})], \dot{z} \xleftarrow{\dot{\nu}} \dot{d}]$ of

$$\big(\langle w\rangle^{-1} \times_{G_0} \langle u|_G\rangle/G\big) \times_{TG_0} T\langle w\rangle/TG = T\langle w\rangle \circ (\langle u|_G\rangle \circ \langle w\rangle^{-1}),$$

we define

$$\sigma_u\big([[y \xrightarrow{\gamma} x, (x, u(x) \xleftarrow{\dot{\mu}} \dot{z})], \dot{z} \xleftarrow{\dot{\nu}} \dot{d}]\big) := (y, u(y) \xleftarrow{u(\gamma^{-1})\circ\dot{\mu}\circ\dot{\nu}} \dot{d}) \in \langle u\rangle.$$

It is not very difficult to check that σ_u is well-defined isomorphism of the bibundles.

It remains to check that the isomorphisms $\{\sigma_u\}$ are components of a natural isomorphism $\sigma : (\langle w\rangle_* \circ \iota_G \circ w^*) \Rightarrow \iota_{G'}$. Thus for a morphism $\beta : u \Rightarrow v$ in $\mathbb{X}(G')$ we need to check that $\iota_{G'}(\beta) \circ \sigma_u = \sigma_v \circ \big((\langle w\rangle_* \circ \iota_G \circ w^*)(\beta)\big)$. The isomorphism of bibundles

$$\big((\langle w\rangle_* \circ \iota_G \circ w^*)(\beta)\big) = \langle w\rangle_*(\langle \beta|_G\rangle) : \langle w\rangle_*(\langle w\rangle_*(\langle u|_G\rangle)) \to \langle w\rangle_*(\langle w\rangle_*(\langle v|_G\rangle))$$

is given by

$$\big([[y \xrightarrow{\gamma} x, (x, u(x) \xleftarrow{\dot{\mu}} \dot{z})], \dot{z} \xleftarrow{\dot{\nu}} \dot{d}]\big) \mapsto \big([[y \xrightarrow{\gamma} x, (x, v(x) \xleftarrow{\beta(x)\circ\dot{\mu}} \dot{z})], \dot{z} \xleftarrow{\dot{\nu}} \dot{d}]\big).$$

Hence

$$\sigma_v\Big(\langle w\rangle_*(\langle\beta|_G\rangle)\,\big([[y \xrightarrow{\gamma} x, (x, u(x) \xleftarrow{\dot\mu} \dot z)], \dot z \xleftarrow{\dot\nu} \dot d]\big)\Big) = (y, v(y) \xleftarrow{v(\gamma^{-1})\circ\beta(x)\circ\dot\mu\circ\dot\nu} \dot d).$$

On the other hand

$$\langle\beta\rangle\Big(\sigma_u\big([[y \xrightarrow{\gamma} x, (x, u(x) \xleftarrow{\dot\mu} \dot z)], \dot z \xleftarrow{\dot\nu} \dot d]\big)\Big) = (y, v(y) \xleftarrow{\beta(y)\circ u(\gamma^{-1})\circ\dot\mu\circ\dot\nu} \dot d).$$

Since β is a natural transformation, $\beta(y)\circ u(\gamma^{-1}) = v(\gamma^{-1})\circ\beta(x)$ for any arrow $y\xrightarrow{\gamma}x$ of G'. Thus $\sigma_v\circ(\langle w\rangle_*(\langle\beta|_G\rangle)) = \langle\beta\rangle\circ\sigma_u$ for any morphism $\beta : u \Rightarrow v$ in $\mathbb{X}(G')$. We conclude that $\{\sigma_u\}$ are components of the desired natural isomorphism. □

5. Categories of vector fields on stacks and Lie 2-algebras

We recall Hepworth's construction [2009] of the category of vector fields $\mathsf{Vect}(\mathcal{A})$ on a stack $\mathcal{A}$. The first step is to extend the tangent functor $T : \mathsf{Man} \to \mathsf{Man}$ on the category of manifolds to a functor $T^{\mathsf{Stack}} : \mathsf{Stack} \to \mathsf{Stack}$ on the 2-category of stacks over manifolds along the Yoneda embedding $y : \mathsf{Man} \to \mathsf{Stack}$. This results in a 2-commuting diagram

$$\begin{array}{ccc} \mathsf{Stack} & \xrightarrow{T^{\mathsf{Stack}}} & \mathsf{Stack} \\ {\scriptstyle y}\uparrow & \nearrow\!\!\!\!\nearrow & \uparrow{\scriptstyle y} \\ \mathsf{Man} & \xrightarrow[T]{} & \mathsf{Man} \end{array}$$

and there is a natural transformation $\pi : T^{\mathsf{Stack}} \Rightarrow \mathrm{id}_{\mathsf{Stack}}$.

Definition 5.1 (Hepworth). The objects of the *category of vector fields* $\mathsf{Vect}(\mathcal{A})$ on a stack $\mathcal{A}$ are pairs (v, α_v) where $v : \mathcal{A} \to T^{\mathsf{Stack}}\mathcal{A}$ is a 1-morphism of stacks and $\alpha_v : \pi_{\mathcal{A}} \circ v \Rightarrow \mathrm{id}_{\mathcal{A}}$ is a 2-morphism. A morphism in $\mathsf{Vect}(\mathcal{A})$ from (v, α_v) to (u, α_u) is a 2-morphism $\beta : v \Rightarrow u$ so that $\alpha_u \circ (\pi_{\mathcal{A}} \star \beta) = \alpha_v$. Here $\circ$ is the vertical composition and $\star$ is whiskering.

Recall that for any Lie groupoid G there is a stack $\mathbb{B}G$ of principal G-bundles. The assignment $G \mapsto \mathbb{B}G$ can be promoted to a functor $\mathbb{B}$ in different ways depending on which source 2-category one chooses. Hepworth took the source to be the 2-category LieGpd of Lie groupoids, smooth functors and natural isomorphisms and considered the functor

$$\mathbb{B} : \mathsf{LieGpd} \to \mathsf{Stack}. \tag{5-1}$$

The essential image of this functor consists of the 2-category $\mathsf{GeomStack}$ of geometric stacks. The functor $\mathbb{B}$ is faithful but not full. In particular the functor $\mathbb{B}$ maps essential equivalences of Lie groupoids (which need not be invertible in LieGpd, even weakly) to isomorphisms of stacks.[2] The tangent functor $T : \mathsf{Man} \to \mathsf{Man}$ is easily extended to a functor $T^{\mathsf{LieGpd}} : \mathsf{LieGpd} \to \mathsf{LieGpd}$. We have a natural transformation $\pi^{\mathsf{LieGpd}} : T^{\mathsf{LieGpd}} \Rightarrow \mathrm{id}_{\mathsf{LieGpd}}$.

Hepworth [2009, Theorem 3.11] proved that there is a natural isomorphism

$$\mathbb{B} \circ T^{\mathsf{LieGpd}} \Leftrightarrow T^{\mathsf{Stack}} \circ \mathbb{B}. \tag{5-2}$$

Consequently given a vector field $v : G \to TG$ on a Lie groupoid G we get a map of stacks $\mathbb{B}v : \mathbb{B}G \to \mathbb{B}TG$. Composing v with the isomorphism $\mathbb{B}TG \to T^{\mathsf{Stack}}(\mathbb{B}G)$ gives us a functor that we again denoted by $\mathbb{B}v : \mathbb{B}G \to T^{\mathsf{Stack}}(\mathbb{B}G)$. This determines an object $(\mathbb{B}v, a_{Bv})$ in the category $\mathsf{Vect}(\mathbb{B}G)$ of vector fields on the stack $\mathbb{B}G$. Hepworth showed that the assignment $v \mapsto (\mathbb{B}v, a_{Bv})$ can be promoted to a functor

$$\mathbb{X}(G) \to \mathsf{Vect}(\mathbb{B}G). \tag{5-3}$$

Here as before $\mathbb{X}(G)$ denotes the category of vector fields on a Lie groupoid G (see Definition 3.1). Hepworth [2009, Theorem 4.15] proved that if the groupoid G is proper then the functor (5-3) is an equivalence of categories.[3] Another important consequence of the existence of the isomorphism (5-2) is that for any geometric stack $\mathcal{A}$ the tangent stack $T^{\mathsf{Stack}}\mathcal{A}$ is geometric as well.

We can promote the assignment $G \to \mathbb{B}G$ to a functor out of a different bicategory, which at a slight risk of confusion we will again denote by $\mathbb{B}$. Namely we can choose as our source the bicategory Bi of Lie groupoids, bibundles and isomorphisms of bibundles. The advantage is that the functor $\mathbb{B} : \mathsf{Bi} \to \mathsf{Stack}$ is fully faithful: for Lie groupoids G and H, the functor $\mathbb{B} : \mathsf{Hom}_{\mathsf{Bi}}(G, H) \to \mathsf{Hom}_{\mathsf{Stack}}(\mathbb{B}G, \mathbb{B}H)$ is an equivalence of categories. Consequently the functor $\mathbb{B} : \mathsf{Bi} \to \mathsf{GeomStack}$ is an equivalence of bicategories. It is not hard to adapt [Hepworth 2009, Theorem 3.11] to this setting: the diagram

$$\begin{array}{ccc} \mathsf{Bi} & \xrightarrow{T^{\mathsf{Bi}}} & \mathsf{Bi} \\ {\scriptstyle \mathbb{B}}\downarrow & \nearrow\!\!\!\nearrow & \downarrow{\scriptstyle \mathbb{B}} \\ \mathsf{GeomStack} & \xrightarrow[T^{\mathsf{Stack}}]{} & \mathsf{GeomStack} \end{array}$$

2-commutes. For convenience, we will choose a weak inverse $\mathbb{B}^{-1} : \mathsf{GeomStack} \to \mathsf{Bi}$ and consider the functor

$$T^{\mathsf{GeomStack}} : \mathsf{GeomStack} \to \mathsf{GeomStack}, \quad T^{\mathsf{GeomStack}} := \mathbb{B} \circ T^{\mathsf{Bi}} \circ \mathbb{B}^{-1},$$

[2] Recall that by tradition a weakly invertible 1-morphism of stacks is called an *isomorphism.*

[3] The hypothesis that the groupoid G is proper is not explicit in the statement of [Hepworth 2009, Theorem 4.15]. However the proof depends on several lemmas: [Hepworth 2009, 4.11, 4.12, 2.11, 2.12]. In particular the proof uses the existence of partitions of unity and Weinstein–Zung linearization, both of which require properness.

which by construction is isomorphic to Hepworth's functor T^{Stack} restricted to geometric stacks. As in the case of T^{Stack} we have a transformation $\pi : T^{\mathsf{GeomStack}} \Rightarrow \mathrm{id}_{\mathsf{GeomStack}}$: namely $\pi := \mathbb{B} \star \tilde{\pi} \star \mathbb{B}^{-1}$. Given a geometric stack $\mathcal{A}$ we now define a category of vector fields $\mathsf{Vect}'(\mathcal{A})$ on $\mathcal{A}$ as follows (compare with Definition 5.1).

Definition 5.2. The *category of vector fields* $\mathsf{Vect}'(\mathcal{A})$ on a geometric stack $\mathcal{A}$, has as objects pairs (X, α_X) where $X : \mathcal{A} \to T^{\mathsf{GeomStack}}\mathcal{A}$ is a 1-morphism of stacks and $\alpha_X : \pi_{\mathcal{A}} \circ X \Rightarrow \mathrm{id}_{\mathcal{A}}$ is a 2-morphism. A morphism from (X, α_X) to (Y, α_Y) in $\mathsf{Vect}'(\mathcal{A})$ is a 2-morphism $\beta : X \Rightarrow Y$ so that $\alpha_Y \circ (\pi_{\mathcal{X}} \star \beta) = \alpha_X$.

It is easy to see that for a geometric stack $\mathcal{A}$ the categories $\mathsf{Vect}(\mathcal{A})$ and $\mathsf{Vect}'(\mathcal{A})$ are equivalent (and even isomorphic). For us there are several advantages in working with $\mathsf{Vect}'(\mathcal{A})$. First of all, the functor Vect' is more explicit than T^{Stack}: the latter involves 2-limits and stackification. Additionally the following result is easy to prove:

Lemma 5.3. *For a Lie groupoid G the classifying stack functor $\mathbb{B} : \mathsf{Bi} \to \mathsf{GeomStack}$ induces an equivalence of categories $(\mathbb{B}_*)_G : \mathbb{X}_{\mathsf{gen}}(G) \to \mathsf{Vect}'(\mathbb{B}G)$, where $\mathbb{X}_{\mathsf{gen}}(G)$ is the category of generalized vector fields (Definition 4.6).*

Proof. Consider a generalized vector field (P, α_P) on the Lie groupoid G. By definition we have an isomorphism $\alpha_P : \tilde{\pi}_G \circ P \Rightarrow \langle \mathrm{id}_G \rangle$ of bibundles. Apply the classifying stack functor $\mathbb{B}$ to the 2-morphism α_P. We get the 2-morphism of stacks $\mathbb{B}\alpha_P : \mathbb{B}(\tilde{\pi}_G \circ P) \Rightarrow \mathbb{B}\langle \mathrm{id}_G \rangle$. Since $\mathbb{B}$ is functor between bicategories, we have canonical 2-arrows $\mathbb{B}\langle \mathrm{id}_G \rangle \Rightarrow \mathrm{id}_{\mathbb{B}G}$ and $\mathbb{B}\langle \pi_G \rangle \circ \mathbb{B}P \Rightarrow \mathbb{B}(\tilde{\pi}_G \circ P)$. Note that these 2-morphisms are 2-isomorphisms since all 2-arrows in the 2-category of stacks are invertible. Composing the three 2-arrows we get a 2-arrow $\mathbb{B}\tilde{\pi}_G \circ \mathbb{B}P \Rightarrow \mathrm{id}_{\mathbb{B}G}$ which we denote by $\alpha_{\mathbb{B}P}$. By definition the pair $(\mathbb{B}P, \alpha_{\mathbb{B}P})$ is an object of $\mathsf{Vect}'(\mathbb{B}G)$.

Similarly a morphism $\beta : (P, \alpha_P) \to (Q, \alpha_Q)$ in $\mathbb{X}_{\mathsf{gen}}(G)$ gives rise to a morphism $\mathbb{B}\beta : \mathbb{B}P \Rightarrow \mathbb{B}Q$. One checks that $\alpha_{\mathbb{B}Q} \circ (\pi_{\mathbb{B}G} \star \mathbb{B}\beta) = \alpha_{\mathbb{B}Q}$. Consequently $\mathbb{B}\beta$ is a morphism in $\mathsf{Vect}'(\mathbb{B}G)$ from $(\mathbb{B}P, \alpha_{\mathbb{B}P})$ to $(\mathbb{B}Q, \alpha_{\mathbb{B}Q})$. We therefore get a functor $(\mathbb{B}_*)_G : \mathbb{X}_{\mathsf{gen}}(G) \to \mathsf{Vect}'(\mathbb{B}G)$. A weak inverse $\mathbb{B}^{-1} : \mathsf{GeomStack} \to \mathsf{Bi}$ gives rise to the functor $((\mathbb{B}^{-1})_*)_G : \mathsf{Vect}'(\mathbb{B}G) \to \mathbb{X}_{\mathsf{gen}}(G)$ in the other direction. The induced functors $(\mathbb{B}_*)_G$ and $((\mathbb{B}^{-1})_*)_G$ are weak inverses of each other. □

We now address the issue of giving the category of vector fields $\mathsf{Vect}'(\mathcal{A})$ on a geometric stack $\mathcal{A}$ the structure of a Lie 2-algebra. We study the functoriality of the assignment $\mathcal{A} \mapsto \mathbb{X}(G)$ of a Lie 2-algebra of vector fields to a geometric stack by a choice of an atlas $G_0 \to \mathcal{A}$. Consider the 2-category $\mathsf{GeomStack}_{\mathrm{iso}}$ of geometric stacks, *isomorphisms* of stacks and 2-morphisms of stacks. The classifying stack functor $\mathbb{B} : \mathsf{Bi} \to \mathsf{GeomStack}$ restricts to an equivalence of bicategories $\mathbb{B} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{GeomStack}_{\mathrm{iso}}$. A choice of a weak inverse $\mathbb{B}^{-1}$ of $\mathbb{B}$ amounts to choosing an atlas for each geometric stack. Once the inverse $\mathbb{B}^{-1}$ is chosen, we have the composite functor $\mathsf{GeomStack}_{\mathrm{iso}} \xrightarrow{\mathbb{B}^{-1}} \mathsf{Bi}_{\mathrm{iso}} \xrightarrow{\mathbb{X}} \mathsf{Lie2Alg}$. By construction, for a stack $\mathcal{A}$

the Lie 2-algebra $\mathbb{X}(\mathbb{B}^{-1}(\mathcal{A}))$ is the Lie 2-algebra of vector fields on the Lie groupoid $G = \mathbb{B}^{-1}(\mathcal{A})$. By the discussion above the category underlying the Lie 2-algebra $\mathbb{X}(\mathbb{B}^{-1}(\mathcal{A}))$ is equivalent to the category of vector fields $\mathsf{Vect}'(\mathcal{A})$ on the stack $\mathcal{A}$.

A different choice of a weak inverse $(\mathbb{B}^{-1})'$ of $\mathbb{B}$ amounts to choosing a possibly different atlas for each geometric stack. Once $(\mathbb{B}^{-1})'$ is chosen we have a natural isomorphism $\alpha : \mathbb{B}^{-1} \Rightarrow (\mathbb{B}^{-1})'$. For each geometric stack $\mathcal{A}$ the component $\alpha_{\mathcal{A}}$ of the natural transformation α is an invertible bibundle $\alpha_{\mathcal{A}} : \mathbb{B}^{-1}(\mathcal{A}) \to (\mathbb{B}^{-1})'(\mathcal{A})$. Applying the functor $\mathbb{X} : \mathsf{Bi}_{\text{iso}} \to \mathsf{Lie2Alg}$ to $\alpha_{\mathcal{A}}$ we get an invertible bibundle $\mathbb{X}(\alpha_{\mathcal{A}}) : \mathbb{X}(\mathbb{B}^{-1}(\mathcal{A})) \to \mathbb{X}((\mathbb{B}^{-1})'(\mathcal{A}))$ in the bicategory $\mathsf{Lie2Alg}$.

One can be fairly explicit as to what the bibundle $\mathbb{X}(\alpha_{\mathcal{A}})$ actually is. Namely let $G_0 \to \mathcal{A}$ be the atlas giving rise to the Lie groupoid $G = \mathbb{B}^{-1}(\mathcal{A})$ and $H_0 \to \mathcal{A}$ be the atlas giving rise to $H = (\mathbb{B}^{-1})'(\mathcal{A})$. Then the total space of the bibundle $\alpha_{\mathcal{A}} : G \to H = (\mathbb{B}^{-1})'(\mathcal{A})$ represents the fiber product $G_0 \times_{\mathcal{A}} H_0$. The linking groupoid $G *_{\alpha_{\mathcal{A}}} H$ is the groupoid corresponding to the atlas $G_0 \sqcup H_0 \to \mathcal{A}$. The linking groupoid comes with two canonical essentially surjective open embeddings $i_G : G \hookrightarrow G *_{\alpha_{\mathcal{A}}} H$ and $i_H : H \hookrightarrow G *_{\alpha_{\mathcal{A}}} H$. By Lemma 4.10 the pullback/restriction functors $i_G^* : \mathbb{X}(G *_{\alpha_{\mathcal{A}}} H) \to \mathbb{X}(G)$, $i_H^* : \mathbb{X}(G *_{\alpha_{\mathcal{A}}} H) \to \mathbb{X}(G)$ are 1-morphisms of Lie 2-algebras that are fully faithful and essentially surjective. Hence the bibundle $\langle i_G^* \rangle$ is invertible in the bicategory $\mathsf{Lie2Alg}$. On the other hand, as was noted in the proof of Theorem 4.4, the bibundles $\langle i_H \rangle \circ \alpha_{\mathcal{A}}$ and $\langle i_G \rangle$ are isomorphic. Hence $\mathbb{X}(\langle i_H \rangle) \circ \mathbb{X}(\alpha_{\mathcal{A}}) \simeq \mathbb{X}(\langle i_G \rangle)$. By construction of the functor $\mathbb{X} : \mathsf{Bi}_{\text{iso}} \to \mathsf{Lie2Alg}$ we have $\mathbb{X}(\langle i_G \rangle) = \langle i_G^* \rangle^{-1}$ and $\mathbb{X}(\langle i_H \rangle) = \langle i_H^* \rangle^{-1}$. Hence $\mathbb{X}(\alpha_{\mathcal{A}}) \simeq \langle i_H^* \rangle \circ \langle i_G^* \rangle^{-1}$.

6. Lie 2-algebras of vector fields on stacks

In the previous section we constructed a functor $\mathbb{X} \circ \mathbb{B}^{-1} : \mathsf{GeomStack}_{\text{iso}} \to \mathsf{Lie2Alg}$. Recall that there is a forgetful functor $u : \mathsf{Lie2Alg} \to \mathsf{Gpd}$ that assigns to a Lie 2-algebra its underlying groupoid. Therefore for every geometric stack $\mathcal{A}$ we have the groupoid $(u \circ \mathbb{X} \circ \mathbb{B}^{-1})(\mathcal{A})$. We should make sure that this groupoid is equivalent to the groupoid of vector fields $\mathsf{Vect}'(\mathcal{A})$ (and hence to Hepworth's groupoid $\mathsf{Vect}(\mathcal{A})$ of vector fields on the stack $\mathcal{A}$).

We start by promoting the assignment $\mathcal{A} \mapsto \mathsf{Vect}'(\mathcal{A})$ to a functor $\mathsf{Vect}' : \mathsf{GeomStack}_{\text{iso}} \to \mathsf{Gpd}$ whose source is the 2-category of geometric stacks and isomorphisms and whose target is the (2,1)-category Gpd of (small) groupoids. We then prove the following theorem:

Theorem 6.1. *The diagram of* (2,1)*-bicategories and functors*

$$\begin{array}{ccc}
\mathsf{GeomStack}_{\text{iso}} & \xrightarrow{\mathsf{Vect}'} & \mathsf{Gpd} \\
\uparrow{\scriptstyle \mathbb{B}} & \overset{\Upsilon}{\nwarrow\!\!\!\nwarrow} & \uparrow{\scriptstyle u} \\
\mathsf{Bi}_{\text{iso}} & \xrightarrow[\mathbb{X}]{} & \mathsf{Lie2Alg}
\end{array}$$

2-commutes. Here as above Gpd *denotes the* (2,1)*-category of groupoids, functors and natural isomorphisms, and* $u : \mathsf{Lie2Alg} \to \mathsf{Gpd}$ *denotes the functor that assigns to each Lie* 2*-algebra its underlying groupoid. The* 1*-components of the transformation* Υ *are weakly invertible functors* (*i.e., equivalences of categories*)*. In particular for a geometric stack* $\mathcal{A}$ *the category underlying the Lie* 2*-algebra* $(\mathbb{X} \circ \mathbb{B}^{-1})(\mathcal{A})$ *is equivalent to Hepworth's category* $\mathsf{Vect}(\mathcal{A})$ *of vector fields on the stack.*

We now construct the 2-functor $\mathsf{Vect}' : \mathsf{GeomStack} \to \mathsf{Gpd}$. An isomorphism $f : \mathcal{A}_1 \to \mathcal{A}_2$ of stacks induces an equivalence of categories $f_* : \mathsf{Vect}'(\mathcal{A}_1) \to \mathsf{Vect}'(\mathcal{A}_2)$: one adapts the proof of Lemma 4.7 to the setting of geometric stacks. Note that if $f = \mathrm{id}_{\mathcal{A}}$ we may take $f_* = \mathrm{id}_{\mathsf{Vect}'(\mathcal{A})}$.

Given isomorphisms $f : \mathcal{A}_1 \to \mathcal{A}_2$ and $g : \mathcal{A}_2 \to \mathcal{A}_3$ of stacks we get equivalences of categories: $(g \circ f)_*$ and $g_* \circ f_*$. We need to produce a natural transformation $\mu_{gf} : g_* \circ f_* \Rightarrow (g \circ f)_*$. So given an object (v, a_v) of $\mathsf{Vect}'(\mathcal{A}_1)$ we need to produce a 2-cell

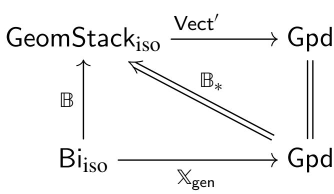
$$(\mu_{gf})_{(v,a_v)} : g_*(f_*(v, a_v)) \Rightarrow (g \circ f)_*(v, a_v)$$

in the category $\mathsf{Vect}'(\mathcal{A}_3)$. By the (adapted) proof of Lemma 4.7 $g_*(f_*(v)) = T^{\mathsf{GeomStack}} g \circ (T^{\mathsf{GeomStack}} f \circ v \circ f^{-1}) \circ g^{-1}$. Since $T^{\mathsf{GeomStack}}$ is a (pseudo-) functor, there is a natural isomorphism $T^{\mathsf{GeomStack}} g \circ T^{\mathsf{GeomStack}} f \Rightarrow T^{\mathsf{GeomStack}}(g \circ f)$. Consequently there is an isomorphism

$$T^{\mathsf{GeomStack}} g \circ (T^{\mathsf{GeomStack}} f \circ v \circ f^{-1}) \circ g^{-1} \Rightarrow T^{\mathsf{GeomStack}}(g \circ f) \circ v \circ (g \circ f)^{-1}.$$

This isomorphism is the desired 2-cell $(\mu_{gf})_{(v,a_v)}$. We are now ready to describe the functor Vect'. To a geometric stack $\mathcal{A}$ it assigns the category $\mathsf{Vect}'(\mathcal{A})$. To an arrow $f : \mathcal{A}_1 \to \mathcal{A}_2$ it assigns the equivalence of categories $\mathsf{Vect}'(f) := f_*$. Additionally for each pair (g, f) we have a natural isomorphism $\mu_{gf} : g_* \circ f_* \Rightarrow (g \circ f)_*$ constructed above. Proceeding similarly (and keeping track of the coherence data) we can promote the assignment $\mathsf{Bi}_{\mathrm{iso}} \ni (G \xrightarrow{P} H) \mapsto (\mathbb{X}_{\mathsf{gen}}(G) \xrightarrow{P_*} \mathbb{X}_{\mathsf{gen}}(H))$ to a functor $\mathbb{X}_{\mathsf{gen}} : \mathsf{Bi}_{\mathrm{iso}} \to \mathsf{Gpd}$. Lemma 5.3 now generalizes as follows:

Lemma 6.2. *The equivalences of categories* $(\mathbb{B}_*)_G : \mathbb{X}_{\mathsf{gen}}(G) \to \mathsf{Vect}'(\mathbb{B}G)$ (*one for each Lie groupoid* G) *assemble into a transformation* $\mathbb{B}_* : \mathbb{X}_{\mathsf{gen}} \Rightarrow \mathbb{B} \circ \mathsf{Vect}'$*. That is, the diagram*



2-commutes. Here as before Gpd *is the* (2,1)*-category of groupoids, functors and natural isomorphisms.*

Lemma 6.3. *The diagram*

$$\begin{array}{ccc} & & \mathsf{Gpd} \\ & \nearrow^{\mathbb{X}_{\mathsf{gen}}} & \\ \mathsf{Bi}_{\text{iso}} & \Uparrow \iota & \uparrow u \\ & \searrow_{\mathbb{X}} & \\ & & \mathsf{Lie2Alg} \end{array}$$

2-commutes and the 1-components of ι *are equivalences of categories.*

Proof. We have the underlying category functor $u_{\text{strict}} : \mathsf{Lie2Alg}_{\text{strict}} \to \mathsf{Gpd}$ which sends Lie 2-algebras to their underlying groupoids and morphisms of Lie 2-algebras to the underlying functors. The functor u_{strict} sends essential equivalences of Lie 2-algebras to weakly invertible functors. By the universal property of the localization $\langle\ \rangle : \mathsf{Lie2Alg}_{\text{strict}} \to \mathsf{Lie2Alg}$ we get the underlying category functor $u : \mathsf{Lie2Alg} \to \mathsf{Gpd}$ with $u(\langle f \rangle)$ isomorphic to $u_{\text{strict}}(f)$ for every essential equivalence of Lie 2-algebras. It follows that for any essential equivalence f in $\mathsf{Lie2Alg}_{\text{strict}}$ the functor $u(\langle f \rangle^{-1})$ is a weak inverse of $u_{\text{strict}}(f)$. We proved that for any essentially surjective open embedding $w : G \to G'$ of Lie groupoids the pullback functor $w^* : \mathbb{X}(G') \to \mathbb{X}(G)$ is an essential equivalence. We defined $\mathbb{X}(w) = \langle w^* \rangle^{-1}$. It follows that $u(\mathbb{X}(w))$ is a weak inverse of $u_{\text{strict}}(w^*)$.

By Lemma 4.13 the diagram (4-5) 2-commutes for any 1-morphism $w : G \to G'$ in $\mathscr{E}mb$. Hence the diagram

$$\begin{array}{ccc} \mathbb{X}(G) & \xrightarrow{\iota_G} & \mathbb{X}_{\mathsf{gen}}(G) \\ {\scriptstyle u(\mathbb{X}(w))}\big\downarrow & & \big\downarrow{\scriptstyle \langle w \rangle_* = \mathbb{X}_{\mathsf{gen}}(w)} \\ \mathbb{X}(G') & \xrightarrow[\iota_{G'}]{} & \mathbb{X}_{\mathsf{gen}}(G') \end{array}$$

2-commutes as well. It follows that the functors $u \circ \mathbb{X}$, $\mathbb{X}_{\mathsf{gen}} \circ \langle\ \rangle \in \mathsf{Hom}_W(\mathscr{E}mb, \mathsf{Gpd})$ are isomorphic (i.e., differ by a transformation whose components are equivalences of categories). Here $\mathsf{Hom}_W(\mathscr{E}mb, \mathsf{Gpd})$ denotes the bicategory of functors that send the collection W of all 1-cells in $\mathscr{E}mb$ to weakly invertible functors.

By the definition of the functor $\mathbb{X} : \mathsf{Bi}_{\text{iso}} \to \mathsf{Lie2Alg}$ its precomposition with the localization functor $\langle\ \rangle : \mathscr{E}mb \to \mathsf{Bi}_{\text{iso}}$ is isomorphic to $\mathbb{X} : \mathscr{E}mb \to \mathsf{Lie2Alg}$. It follows that the functors $u \circ \mathbb{X} \circ \langle\ \rangle$ and $\mathbb{X}_{\mathsf{gen}} \circ \langle\ \rangle$ are isomorphic in $\mathsf{Hom}_W(\mathscr{E}mb, \mathsf{Gpd})$. By the universal property of the localization $\langle\ \rangle : \mathscr{E}mb \to \mathsf{Bi}_{\text{iso}}$, the functors $u \circ \mathbb{X}$ and $\mathbb{X}_{\mathsf{gen}}$ are isomorphic in $\mathsf{Hom}(\mathsf{Bi}_{\text{iso}}, \mathsf{Gpd})$. □

Theorem 6.1 now follows directly from Lemmas 6.2 and 6.3.

7. Generalized and multiplicative vector fields on a Lie groupoid

In this section we prove Theorem 4.9. In the case of *proper* Lie groupoids Theorem 4.9 follows from [Hepworth 2009, Theorem 4.15].

The fact that the functor $\iota_G : \mathbb{X}(G) \hookrightarrow \mathbb{X}_{\mathsf{gen}}(G)$ is fully faithful is an easy consequence of Theorem 2.9. We now address essential surjectivity. We first prove:

Lemma 7.1. *Let $V = \{V_1 \rightrightarrows V_0\}$ be a 2-vector space, $v_1, \dots, v_s \in V_0$ a finite collection of objects and $\{v_i \xleftarrow{w_{ij}} v_j\}_{i,j=1}^s$ a collection of morphisms satisfying the cocycle conditions: $w_{ii} = 1_{v_i}$ for all i; $w_{ji} = {w_{ij}}^{-1}$ for all i, j; $w_{ij}w_{jk} = w_{ik}$ for all i, j, k. Then for any $\lambda_1, \dots, \lambda_s \in [0, 1]$ with $\sum \lambda_k = 1$ there are morphisms $v_i \xleftarrow{z_i} \sum \lambda_k v_k$ $(i = 1, \dots, s)$ with $w_{ij} = z_i z_j^{-1}$ for all i, j.*

Proof. By Remark 2.14 the category V is isomorphic to the action groupoid $\{U \times V_0 \rightrightarrows V_0\}$ where $U_0 = \ker(s : V_1 \to V_0)$, $\partial : U \to V_0$ is $t|_U$ and the action of U on V_0 is given by $u \cdot v := v + \partial(u)$. Note that the multiplication/composition in $\{U \times V_0 \rightrightarrows V_0\}$ is given by $(u', v + \partial(u))(u, v) = (u' + u, v)$ for all $v \in V_0$, $u, u' \in U$. Consequently $(u, v)^{-1} = (-u, v + \partial(u))$. The isomorphism $f : V \to \{U \times V_0 \rightrightarrows V_0\}$ is given on morphisms by

$$f(w) = (w - 1_{s(w)}, s(w)) \in U \times V_0 \qquad \text{for all } w \in V_1.$$

The isomorphism f followed by the projection onto U sends the morphisms w_{ij} to vectors $u_{ij} \in U$. It is easy to see that the cocycle conditions translate into: $u_{ii} = 0$ for all i; $u_{ji} = -u_{ij}$ for all i, j and $u_{ik} - u_{jk} = u_{ij}$ for all i, j, k. Moreover $\partial(u_{ij}) = v_i - v_j$ for all i, j. Now consider $y_i = \left(\sum \lambda_k u_{ik}, \sum \lambda_k v_k\right) \in U \times V_0$ and set $z_i := f^{-1}(y_i) \in V_1$. We now verify that the z_i's are the desired morphisms. By definition the source of y_i is $\sum \lambda_k v_k$. The target of y_i is

$$\begin{aligned}\partial\Big(\sum_k \lambda_k u_{ik}\Big) + \sum_k \lambda_k v_k &= \sum_k \lambda_k \partial(u_{ik}) + \sum_k \lambda_k v_k \\ &= \sum_k \lambda_k (v_i - v_k) + \sum_k \lambda_k v_k = \sum_k \lambda_k v_i = v_i.\end{aligned}$$

Hence z_i is an arrow from $\sum \lambda_k v_k$ to v_i. Finally

$$\begin{aligned}y_i y_j^{-1} &= \Big(\sum_k \lambda_k u_{ik}, \sum \lambda_k v_k\Big)\Big(-\sum_k \lambda_k u_{jk}, v_j\Big) \\ &= \Big(\sum_k \lambda_k (u_{ik} - u_{jk}), v_j\Big) = \Big(\sum_k \lambda_k u_{ij}, v_j\Big) = (u_{ij}, v_j),\end{aligned}$$

and so $z_i z_j^{-1} = w_{ij}$ as desired. □

Proposition 7.2. *Let $G = \{G_1 \rightrightarrows G_0\}$ be a Lie groupoid, $U_0 \subset G_0$ an open submanifold and $U = \{U_1 \rightrightarrows U_0\}$ the restriction of G to U_0 (that is, U_1 consists of arrows of G with source and target in U_0). Given a functor $X : U \to TG$ together with a natural isomorphism $\alpha : (i : U \hookrightarrow G) \Rightarrow \pi_G \circ X$ there exists a functor $Y : U \to TU$ so that $\pi_U \circ Y = \mathrm{id}_U$ and a natural isomorphism $\beta : Ti \circ Y \Rightarrow X$.*

Proof. By definition of α the diagram

$$\begin{array}{ccc} U_0 & \xrightarrow{X} & TG_0 \\ {\scriptstyle \alpha}\downarrow & & \downarrow{\scriptstyle \pi_G} \\ G_1 & \xrightarrow[t]{} & G_0 \end{array}$$

commutes. Hence there is a smooth map $(\alpha, X) : U_0 \to G_1 \times_{t,G_0,\pi} TG_0 = t^*TG_0$. Since the target map $t : G_1 \to G_0$ is a submersion, its differential $Tt_\gamma : T_\gamma G_1 \to T_{t(\gamma)}G_0$ is a surjective linear map for each $\gamma \in G_1$. Consequently the map $\Phi : TG_1 \to t^*TG_0$, $\Phi(\gamma, v) = (\gamma, Tt_\gamma v)$ is a surjective map of vector bundles over G_1. Choose a smooth section $\sigma : t^*TG_0 \to TG_1$ of Tt and consider the composite $\beta := \sigma \circ (\alpha, X) : U_0 \to TG_1$. By construction of β,

$$\beta(x) \in T_{\alpha(x)}G_1 \quad \text{and} \quad Tt_{\alpha(x)}\beta(x) = X(x)$$

for any $x \in U_0$. We now define a functor $Y : U \to TU$. On objects we set

$$Y(x) = Ts(\beta(x)).$$

For an arrow $x \xrightarrow{\gamma} y \in U_1$ we set

$$Y(\gamma) = \beta(y)^{-1}X(\gamma)\beta(x).$$

It is easy to check that Y is indeed functor, $\beta : Ti \circ Y \Rightarrow X$ is a natural transformation and $\pi_U \circ Y = \mathrm{id}_U$. □

Proposition 7.3. *Let G be a Lie groupoid and*

$$\begin{array}{ccccc} G_1 & & & & TG_1 \\ \downdownarrows & & P & & \downdownarrows \\ & \swarrow{\scriptstyle a_P^L} & & {\scriptstyle a_P^R}\searrow & \\ G_0 & & & & TG_0 \end{array}$$

be a bibundle from G to the tangent groupoid TG such that the composite $\langle\pi\rangle \circ P$ is isomorphic to $\langle \mathrm{id}_G\rangle$ by way of a bibundle isomorphism $\mathbf{a} : \langle\pi\rangle \circ P \Rightarrow \langle \mathrm{id}_G\rangle$. Then the left anchor $a_P^L : P \to G_0$ has a global section $\tau : G_0 \to P$. Moreover we may choose τ so that the corresponding functor $X_\tau : G \to TG$ is a multiplicative vector field (i.e., $\pi_G \circ X_\tau = \mathrm{id}_G$). Consequently the functor $\iota_G : \mathbb{X}(G) \to \mathbb{X}_{\mathsf{gen}}(G)$ of Definition 4.8 is essentially surjective.

Proof. Since $a_P^L : P \to G_0$ is a surjective submersion, it has local sections. Choose a collection of local sections $\{\sigma_i : U_0^{(i)} \to P\}$ of a_P^L so that $\{U_0^{(i)}\}$ is an open cover of G_0. It is no loss of generality to assume that the cover is locally finite. Denote the restriction of the groupoid G to $U_0^{(i)}$ by $U^{(i)}$. That is, the manifold of objects of $U^{(i)}$ is $U_0^{(i)}$ and the manifold of arrows $U_1^{(i)}$ consists of all arrows of G with source and target in $U_0^{(i)}$, so $U_1^{(i)} := s^{-1}(U_0^{(i)}) \bigcap t^{-1}(U_0^{(i)})$.

For each section σ_i we get a functor $X_i : U^{(i)} \to TG$ whose value on objects is $X_i(x) = a_P^R(\sigma_i(x))$. The value of X_i on an arrow $y \xleftarrow{\gamma} x \in U_1^{(i)}$ is uniquely defined by the equation $\gamma \cdot \sigma_i(x) = \sigma_i(y) \cdot X_i(\gamma)$ (see Lemma 2.7). We next observe that the isomorphism $\mathbf{a} : \langle \pi_G \rangle \circ P \to \langle \mathrm{id}_G \rangle$ gives rise to natural isomorphisms $\alpha_j : \pi_G \circ X_j \Rightarrow (\iota_j : U^{(j)} \hookrightarrow G)$ where $\iota_j : U^{(j)} \hookrightarrow G$ is the inclusion functor. This can be seen as follows.

Recall that the composite $Q \circ P$ of bibundles $P : K \to L$ and $Q : L \to M$ is the quotient of the fiber product $P \times_{a_P^R, L_0, a_Q^L} Q$ by the action of L. We denote by $[p, q]$ the orbit of $(p, q) \in P \times_{a_P^R, L_0, a_Q^L} Q$ in $Q \circ P = (P \times_{a_P^R, L_0, a_Q^L} Q)/L$. The bibundle $\langle \pi_G \rangle$ is the fiber product $TG_0 \times_{\pi_G, G_0, t} G_1$ with the anchor maps $a^R_{\langle \pi_G \rangle}(v, \gamma) = v$, $a^L_{\langle \pi_G \rangle}(v, \gamma) = s(\gamma)$. Consequently in our case

$$\langle \pi_G \rangle \circ P = (P \times_{a_P^R, TG_0, a^L_{\langle \pi_G \rangle}} (TG_0 \times_{\pi_G, G_0, s} G_1))/TG.$$

It is convenient to identify $P \times_{a_P^R, TG_0, a^L_{\langle \pi_G \rangle}} (TG_0 \times_{\pi_G, G_0, s} G_1)$ with $P \times_{\pi_G \circ a_P^R, G_0, s} G_1$ by way of the TG-equivariant isomorphism $(p, (\pi_G \circ a_P^R)(p), \gamma) \mapsto (p, \gamma)$. We then have a $G \times G$ equivariant diffeomorphism

$$\mathbf{a} : (P \times_{\pi_G \circ a_P^R, G_0, s} G_1)/TG \to G_1, \quad [p, \gamma] \mapsto \mathbf{a}([p, \gamma])$$

with

$$s(\mathbf{a}([p, \gamma])) = s(\gamma) \quad \text{and} \quad t(\mathbf{a}([p, \gamma])) = a_L^P(p).$$

A local section $\sigma_i : U_0^{(i)} \to P$ also defines a local section $\bar{\sigma}_i : U_0^{(i)} \to (P \times_{G_0} G_1)/TG$ of $a^L_{\langle \pi_G \rangle \circ P} : (P \times_{G_0} G_1)/TG \to G_0$. It is given by

$$\bar{\sigma}_i(x) = [\sigma_i(x), 1_{(\pi_G \circ a_P^R \circ \sigma_i)(x)}] (= [\sigma_i(x), 1_{\pi_G \circ X_i\,(x)}]).$$

The arrow $\mathbf{a}(\bar{\sigma}_i(x)) \in G_1 = \langle \mathrm{id}_G \rangle$ is an arrow with the target $a_P^L(\sigma_i(x)) = x$ and the source $s(1_{\pi_G \circ X_i\,(x)}) = \pi_G \circ X_i\,(x)$. We define the desired natural isomorphism α_i by setting $\alpha_i(x) = (\mathbf{a}(\bar{\sigma}_i(x)))^{-1}$. By Proposition 7.2 there are smooth maps $\beta_i : U_0^{(i)} \to TG_1$ so that $\pi_G \circ \beta_i = \alpha_i$ and $Tt \circ \beta_i = X_i$. Moreover the functors $Y_i : U^{(i)} \to TG$ given by $Y_i = Ts \circ \beta_i$ define multiplicative vector fields on each groupoid $U^{(i)}$. This is because their images land in $TU^{(i)} \subset TG$. In particular $\pi_G(Y_i(x)) = x$ for all $x \in U_0^{(i)}$.

Define the local sections $\nu_i : U_0^{(i)} \to P$ of a_P^L by $\nu_i(x) := \sigma_i(x) \cdot \beta_i(x)$ for all $x \in U_0^{(i)}$. Then by definition

$$a_P^R(\nu_i(x)) = Y_i(x) \quad \text{and} \quad \gamma \cdot \nu_i(x) = \nu_i(y) \cdot Y_i(x)$$

for all arrows $y \xleftarrow{\gamma} x$. For all i and all $x \in U_0^{(i)}$, we have

$$\begin{aligned}\mathbf{a}\big([\nu_i(x), 1_{\pi_G \circ a_P^R \circ \nu_i(x)}]\big) &= \mathbf{a}\big([\sigma_i(x)\beta_i(x), 1_{\pi_G \circ Y_i(x)}]\big) = \mathbf{a}\big([\sigma_i(x), \pi_G(\beta(x))]\big) \\ &= \mathbf{a}\big([\sigma_i(x), 1_{\pi_G \circ X_i(x)}]\big)\pi_G(\beta(x)) = \mathbf{a}(\bar{\sigma}_i(x))\alpha_i(x) = 1_x.\end{aligned}$$

Hence $\mathbf{a}([\nu_i(x), 1_{\pi_G \circ Y_i(x)}] = 1_x$. Finally, we construct a global section $\tau : G_0 \to P$ of a_P^L and the corresponding global multiplicative vector fields $X_\tau : G \to TG$ using a partition of unity argument. Choose a partition of unity $\{\lambda_i\}$ on G_0 subordinate to the cover $\{U_0^{(i)}\}$. Since the cover is locally finite it is no loss of generality to assume that the cover is in fact finite.

Consider a point $x \in U_0^{(i)} \cap U_0^{(j)}$. Then

$$\pi_G \circ a_P^R \circ \nu_i(x) = x = \pi_G \circ a_P^R \circ \nu_j(x).$$

Moreover

$$\mathbf{a}\big([\nu_i(x), 1_x]\big) = \mathbf{a}\big([\nu_i(x), 1_{\pi_G \circ a_R \circ \nu_i(x)}]\big) = 1_x.$$

Similarly $\mathbf{a}([\nu_j(x), 1_x]) = 1_x$. It follows that $[\nu_j(x), 1_x] = [\nu_i(x), 1_x]$ in the orbit space $(P \times_{G_0} G_1)/TG$ since $\mathbf{a}$ is a diffeomorphism. Therefore there is an arrow $w_{ij} \in TG_1$ so that

$$(\nu_i(x)w_{ij}(x), 1_x) = (\nu_j(x), \pi_G(w_{ij}(x))1_x).$$

Consequently

$$\nu_i(x)w_{ij}(x) = \nu_j(x) \quad \text{and} \quad \pi_G(w_{ij}(x)) = 1_x,$$

that is, $w_{ij}(x) \in T_{1_x}G_1$. Moreover since $a_P^L : P \to G_0$ is a principal TG_1 bundle, the arrow $w_{ij}(x)$ with this property is unique and depends smoothly on x. Note that the source of w_{ij} is $Y_j(x)$ and the target is $Y_i(x)$. The uniqueness of the $w_{ij}(x)$'s implies that the collection $\{w_{ij}(x)\}$ satisfies the cocycle conditions of Lemma 7.1. Therefore there exist arrows $Y_i(x) \xleftarrow{z_i(x)} \sum_k \lambda_k Y_k(x)$ with $z_i(x)z_j(x)^{-1} = w_{ij}(x)$. A quick look at the proof of Lemma 7.1 should convince the reader that $z_i(x)$'s depend smoothly on x.

For $x \in U_0^{(i)}$ we set $\tau(x) = \nu_i(x) \cdot z_i(x)$. Note that for $x \in U_0^{(i)} \cap U_0^{(j)}$, we have

$$\nu_j(x) = \nu_i(x)w_{ij}(x) = \nu_i(x)z_i(x)z_j(x)^{-1}.$$

Therefore $\nu_j(x) \cdot z_j(x) = \nu_i(x) \cdot z_i(x)$. It follows that τ is a globally defined section of $a_P^L : P \to G_0$. It remains to show that the corresponding functor $X_\tau : G \to TG$ is a multiplicative vector field. By construction for each index i we have a natural isomorphism $z_i : Y_i \Rightarrow X_\tau|_{U^{(i)}}$. Since $z_i(x) \in T_{1_x}G_1$ and since Y_i is a multiplicative vector field, the restriction $X_\tau|_{U^{(i)}}$ is also a multiplicative vector field. We conclude that X_τ is a multiplicative vector field globally.

We now argue that $\iota_G : \mathbb{X}(G) \to \mathbb{X}_{\mathsf{gen}}(G)$ is essentially surjective. Given a generalized vector field (P, α_P) the isomorphism $\alpha_P : \tilde{\pi}_G \circ P \to \langle \mathrm{id}_G \rangle$ defines an isomorphism $\mathbf{a} : \langle \pi_G \rangle \circ P \to \langle \mathrm{id}_G \rangle$ (remember that we fixed a 2-cell $\tilde{\pi}_G \to \langle \pi_G \rangle$ for every Lie groupoid G). By the above argument we have a multiplicative vector field $X : G \to TG$ and an isomorphism of bibundles $\gamma : P \to \langle X \rangle$. It remains to show that the vector fields (P, α_P) and $(\langle X \rangle, \alpha_{\langle X \rangle})$ are isomorphic. This is not entirely obvious since $\alpha_{\langle X \rangle} \circ (\tilde{\pi}_G \star \gamma)$ need not equal to α_P. Nonetheless, γ can be modified to a new isomorphism $\beta : P \to \langle X \rangle$ so that $\alpha_{\langle X \rangle} \circ (\tilde{\pi}_G \star \beta) = \alpha_{\langle X \rangle}$. This follows from Lemma 7.4 below. □

Lemma 7.4. *Let G be a Lie groupoid, $P : G \to TG$ a bibundle, $\alpha, \alpha' : \widetilde{P} \circ P \to \langle \mathrm{id}_G \rangle$ two isomorphisms of bibundles. Then there exists an isomorphism $\beta : P \to P$ of bibundles so that $\alpha \circ (\tilde{\pi}_G \star \beta) = \alpha'$. Moreover we may take $\beta = P \star (\alpha' \circ \alpha^{-1})$.*

Proof. For any two isomorphisms $\gamma, \delta : \langle \mathrm{id}_G \rangle \to \langle \mathrm{id}_G \rangle$ of bibundles, we have

$$\gamma \circ \delta = \gamma \star \delta.$$

Consequently for any $\gamma : \langle \mathrm{id}_G \rangle \to \langle \mathrm{id}_G \rangle$ we have

$$\begin{aligned}
(\alpha \circ (\alpha')^{-1}) \circ \gamma &= (\alpha \circ (\alpha')^{-1}) \star \gamma \\
&= (\alpha \circ 1_{\tilde{\pi}_G \circ P} \circ (\alpha')^{-1}) \star (1_{\langle \mathrm{id}_G \rangle} \circ \gamma \circ 1_{\langle \mathrm{id}_G \rangle}) \\
&= (\alpha \star 1_{\langle \mathrm{id}_G \rangle}) \circ ((\tilde{\pi_G} \circ P) \star \gamma) \circ ((\alpha')^{-1} \star 1_{\langle \mathrm{id}_G \rangle}) \\
&= \alpha \circ (\tilde{\pi}_G \star (P \star \gamma)) \circ (\alpha')^{-1}.
\end{aligned}$$

Hence

$$1_{\langle \mathrm{id}_G \rangle} = \alpha \circ (\tilde{\pi}_G \star (P \star (\alpha' \circ \alpha^{-1}))) \circ (\alpha')^{-1}$$

if $\gamma = \alpha' \circ \alpha^{-1}$. Therefore

$$\alpha' = \alpha \circ (\tilde{\pi}_G \star (P \star (\alpha' \circ \alpha^{-1}))).$$ □

Acknowledgments

We thank Henrique Bursztyn for many helpful discussions. In particular this paper has partially originated from conversations of one of us (E.L.) with Henrique at Poisson 2014. We thank James Waldron for making us exercise more care with the equivalences of Lie 2-algebras. We thank the referees for helpful comments.

References

[Baez and Crans 2004] J. C. Baez and A. S. Crans, "Higher-dimensional algebra, VI: Lie 2-algebras", *Theory Appl. Categ.* **12** (2004), 492–538. MR Zbl

[Behrend and Xu 2011] K. Behrend and P. Xu, "Differentiable stacks and gerbes", *J. Symplectic Geom.* **9**:3 (2011), 285–341. MR Zbl

[Bénabou 1967] J. Bénabou, "Introduction to bicategories", pp. 1–77 in *Reports of the Midwest Category Seminar*, edited by A. Dold and B. Eckmann, Lecture Notes in Math. **47**, Springer, 1967. MR Zbl

[Borceux 1994] F. Borceux, *Handbook of categorical algebra, I: Basic category theory*, Encycl. Math. Appl. **50**, Cambridge Univ. Press, 1994. MR Zbl

[Frégier and Wagemann 2011] Y. Frégier and F. Wagemann, "On Hopf 2-algebras", *Int. Math. Res. Not.* **2011**:15 (2011), 3471–3501. MR Zbl

[Hepworth 2009] R. Hepworth, "Vector fields and flows on differentiable stacks", *Theory Appl. Categ.* **22** (2009), 542–587. MR Zbl

[Mackenzie 2005] K. C. H. Mackenzie, *General theory of Lie groupoids and Lie algebroids*, Lond. Math. Soc. Lect. Note Ser. **213**, Cambridge Univ. Press, 2005. MR Zbl

[Mackenzie and Xu 1998] K. C. H. Mackenzie and P. Xu, "Classical lifting processes and multiplicative vector fields", *Q. J. Math. Oxford* (2) **49**:193 (1998), 59–85. MR Zbl

[Moerdijk and Mrčun 2003] I. Moerdijk and J. Mrčun, *Introduction to foliations and Lie groupoids*, Cambridge Stud. Adv. Math. **91**, Cambridge Univ. Press, 2003. MR Zbl

[Noohi 2013] B. Noohi, "Integrating morphisms of Lie 2-algebras", *Compos. Math.* **149**:2 (2013), 264–294. MR Zbl

[Ortiz and Waldron 2019] C. Ortiz and J. Waldron, "On the Lie 2-algebra of sections of an $\mathscr{LA}$-groupoid", *J. Geom. Phys.* **145** (2019), art. id. 103474. MR Zbl

[Pronk 1996] D. A. Pronk, "Etendues and stacks as bicategories of fractions", *Compos. Math.* **102**:3 (1996), 243–303. MR Zbl

[SGA 4_3 1973] P. Deligne, "La formule de dualité globale", pp. 481–587 in *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (Séminaire de Géométrie Algébrique du Bois Marie 1963–1964), edited by M. Artin et al., Lecture Notes in Math. **305**, Springer, 1973. MR Zbl

[Street 1996] R. Street, "Categorical structures", pp. 529–577 in *Handbook of algebra, I*, edited by M. Hazewinkel, Elsevier, Amsterdam, 1996. MR Zbl

[Weinstein 2009] A. Weinstein, "The volume of a differentiable stack", *Lett. Math. Phys.* **90**:1-3 (2009), 353–371. MR Zbl

Received November 29, 2019. Revised June 9, 2020.

DANIEL BERWICK-EVANS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, IL
UNITED STATES
danbe@illinois.edu

EUGENE LERMAN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, IL
UNITED STATES
lerman@illinois.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.35

LOWER REGULARITY SOLUTIONS OF THE BIHARMONIC SCHRÖDINGER EQUATION IN A QUARTER PLANE

ROBERTO DE A. CAPISTRANO-FILHO,
MÁRCIO CAVALCANTE AND FERNANDO A. GALLEGO

We deal with the initial-boundary value problem of the biharmonic cubic nonlinear Schrödinger equation in a quarter plane with inhomogeneous Dirichlet–Neumann boundary data. We prove local well-posedness in the low regularity Sobolev spaces by introducing Duhamel boundary forcing operator associated to the linear equation in order to construct solutions in the whole line. With this in hand, the energy and nonlinear estimates allow us to apply the Fourier restriction method, introduced by J. Bourgain, to obtain our main result. Additionally, we discuss adaptations of this approach for the biharmonic cubic nonlinear Schrödinger equation on star graphs.

1. Introduction

1A. ***Presentation of the model.*** The fourth-order nonlinear Schrödinger (4NLS) equation or biharmonic cubic nonlinear Schrödinger equation

$$i\partial_t u + \partial_x^2 u - \partial_x^4 u = \lambda |u|^2 u, \tag{1-1}$$

was introduced in [Karpman 1996; Karpman and Shagalov 2000] to take into account the role of small fourth-order dispersion terms in the propagation of intense laser beams in a bulk medium with Kerr nonlinearity. Equation (1-1) arises in many scientific fields such as quantum mechanics, nonlinear optics and plasma physics, and has been intensively studied with fruitful references (see [Ben-Artzi et al. 2000; Cui and Guo 2007; Karpman 1996; Pausader 2007; 2009a]).

The past twenty years such 4NLS equations have been deeply studied from different mathematical viewpoints. For example, Fibich et al. [2002] worked on various properties of the equation in the subcritical regime, with part of their analysis relying on very interesting numerical developments. The well-posedness and existence of solutions for different domains have been shown (see, for instance, [Capistrano-Filho et al. 2019; Kwak 2018; Özsarı and Yolcu 2019; Pausader 2007;

MSC2010: 35A07, 35C15, 35G15, 35G30, 35Q55.
Keywords: biharmonic Schrödinger equation, initial-boundary value problem, local well-posedness, quarter plane.

2009a; Tsutsumi 2014; Oh and Tzvetkov 2017; Wen et al. 2014]) by means of the Fourier restriction method, energy method, forcing boundary operators, Laplace transform, harmonic analysis, Fokas method, etc.

It is interesting to point out that there are many works related to (1-1) not only dealing with well-posedness theory. For example, Natali and Pastor [2015] considered the fourth-order dispersive cubic nonlinear Schrödinger equation on the line with mixed dispersion. They proved the orbital stability, in the $H^2(\mathbb{R})$-energy space, by constructing a suitable Lyapunov function. Considering (1-1) on the circle, Oh and Tzvetkov [2017] showed that the mean-zero Gaussian measures on Sobolev spaces $H^s(\mathbb{T})$, for $s > \frac{3}{4}$, are quasi-invariant under the flow. There has been significant progress over recent years; see for instance [Burq et al. 2002; 2013] for the nonlinear Schrödinger equation.

In addition to these works, two of us worked recently with the intent of proving controllability results for the 4NLS equation. More precisely, we proved that the solutions of the associated linear system (1-1) is globally exponentially stable in a periodic domain $\mathbb{T}$, by using certain properties of propagation of compactness and regularity in Bourgain spaces. Theses properties together with the local exact controllability ensure that fourth order nonlinear Schrödinger is globally exactly controllable; for details, see [Capistrano-Filho and Cavalcante 2019].

Özsarı and Yolcu [2019] proposed (1-1) without the term $\partial_x^2 u$. This system has an interesting physical point of view, precisely, the model corresponds to a situation in which wave is generated from a fixed source such that it moves into the medium in one specific direction.

1B. ***Setting of the problem.*** We mainly consider the biharmonic Schrödinger equation on the right half-line

$$\begin{cases} i\partial_t u - \partial_x^4 u + \lambda |u|^2 u = 0, & (t,x) \in (0,T) \times (0,\infty), \\ u(0,x) = u_0(x), & x \in (0,\infty), \\ u(t,0) = f(t),\ u_x(t,0) = g(t), & t \in (0,T). \end{cases} \tag{1-2}$$

With suitable choices of $f(t)$ and $g(t)$ in (1-2), we are interested on the following initial-boundary value problem (IBVP):

Is the IBVP (1-2) local well-posed in the low regularity Sobolev space, more precisely, in $H^s(\mathbb{R}^+)$ for $0 \le s < \frac{1}{2}$?

Before presenting the answer for this question, let us present some brief comments on the techniques to solve IBVPs on the half-line.

1C. ***Comments about the techniques to solve IBVPs on the half-line.*** Different techniques have been developed in the last years in order to solve IBVPs associated to some dispersive models on the half-line. Fokas [2008] introduced an approach to solve IBVPs associated to integrable nonlinear evolution equations, which is known as the unified transform method (UTM) or as Fokas transform method. The UTM provides a generalization of the inverse scattering transform method from initial value problems (IVP) to IBVPs. The classical method based on the Laplace transform was used successfully in [Bona et al. 2006; 2018; Erdoğan and Tzirakis 2017; Compaan and Tzirakis 2017]. A new approach was introduced by Colliander and Kenig [2002] by recasting the IBVP on the half-line by a forced IVP defined in the line $\mathbb{R}$. To see other applications of this technique, we refer the results established in [Cavalcante 2017; Cavalcante and Corcho 2019; Holmer 2005; 2006]. On the other hand, Faminskii [2019] used an approach based on the investigation of special solutions of a "boundary potential" type for solution of linearized Korteweg–de Vries (KdV) equation in order to obtain global results for the IBVP associated to the KdV equation on the half-line with more general boundary conditions. Fokas et al. [2016] introduced a method which combines the UTM with a contraction mapping principle. We caution that this is only a small sample of the extant works on these techniques.

1D. ***Biharmonic NLS equation.*** As mentioned in the beginning of this introduction, the 4NLS equation or biharmonic NLS equation

$$i\partial_t u + \partial_x^2 u - \partial_x^4 u = \lambda |u|^2 u, \tag{1-3}$$

was introduced in [Karpman 1996; Karpman and Shagalov 2000]. Huo and Jia [2005] studied the Cauchy problem of one-dimensional fourth-order nonlinear Schrödinger equation related to the vortex filament. They proved the local well-posedness for initial data in $H^s(\mathbb{R})$ for $s \geq \frac{1}{2}$ by using the Fourier restriction norm method under certain coefficient condition. Concerning local well-posedness of the nonlinear fourth order Schrödinger equations, we cite [Hao et al. 2006; Segata 2004]. With respect of the global well-posedness, in the one-dimensional case with some restriction in the initial data for various nonlinearities, we refer to [Hayashi and Naumkin 2015a; 2015b; 2015c; 2015d] and, finally, for the study n-dimensional case the reader can see [Pausader 2009b; Pausader and Shao 2010].

Lastly, in a recent work of IBVP for biharmonic Schrödinger equation on the half-line

$$i\partial_t u + \partial_x^4 u = \lambda |u|^p u, \tag{1-4}$$

Özsarı and Yolcu [2019], proved local well-posedness on the high regularity function spaces $H^s(\mathbb{R}^+)$, for $\frac{1}{2} < s < \frac{9}{2}$, with $s \neq \frac{3}{2}$. The authors used the Fokas method [1997; 2008] combined with contraction arguments to achieve the result.

1E. ***Main result.*** Now, let us present the main result of this article. Consider the biharmonic Schrödinger equation on the right half-line

$$\begin{cases} i\partial_t u + \gamma\partial_x^4 u + \lambda|u|^2 u = 0, & (t,x)\in(0,T)\times(0,\infty), \\ u(0,x) = u_0(x), & x\in(0,\infty), \\ u(t,0) = f(t),\ u_x(t,0) = g(t), & t\in(0,T), \end{cases} \tag{1-5}$$

for $\gamma,\lambda\in\mathbb{R}$. We say that system (1-5) is focusing if $\gamma\lambda < 0$ and defocusing when $\gamma\lambda > 0$. In this paper we will study the case when $\gamma = -1$, however the approach used here can be applied when $\gamma\in\mathbb{R}\setminus\{0\}$.

The presence of two boundary conditions in (1-5) can be motivated by integral identities on smooth decaying solutions for the linear equation

$$i\partial_t u - \partial_x^4 u = 0. \tag{1-6}$$

Indeed, for a smooth decaying solution u of (1-6) and $T > 0$, we have

$$\int_0^\infty |u(T,x)|^2\,dx = \int_0^\infty |u(0,x)|^2\,dx - \int_0^T \operatorname{Im}(\partial_x^3 u(t,0)\bar{u}(t,0))\,dt + \int_0^T \operatorname{Im}(\partial_x^2 u(t,0)\partial_x\bar{u}(t,0))\,dt. \tag{1-7}$$

Thus, from (1-7) we can conclude that if we assume $u(0,x) = u(t,0) = u_x(t,0) = 0$ the linear solution for (1-6) is the trivial one.

It is well-known by [Kenig et al. 1991] that the *local smoothing effect* for the fourth-order linear group operator $e^{it\partial_x^4}$

$$\|\partial_x^j e^{it\partial_x^4}\phi\|_{L_x^\infty \dot{H}^{\frac{1}{8}(2s+3-2j)}(\mathbb{R}_t)} \le c\|\phi\|_{H^s(\mathbb{R})} \quad \text{for } j = 0, 1 \text{ and } s\in\mathbb{R}, \tag{1-8}$$

which motivates the relation of regularities among initial and boundary data.

Thus, we are able to present the main goal in the paper: to answer the problem cited in the beginning of this introduction, that is, to show the local well-posedness of (1-5) in the low regularity Sobolev space $H^s(\mathbb{R}^+)$, for $0 \le s < \frac{1}{2}$.

We state the main theorem for IBVP (1-5) as follows.

Theorem 1.1. *Let $s\in\left[0,\frac{1}{2}\right)$. For given initial-boundary data*

$$(u_0, f, g)\in H^s(\mathbb{R}^+)\times H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+)\times H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+),$$

there exist a positive time

$$T := T\big(\|u_0\|_{H^s(\mathbb{R}^+)}, \|f\|_{H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+)}, \|g\|_{H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+)}\big),$$

and unique solution $u(t,x) \in C((0,T); H^s(\mathbb{R}^+))$ *of the IBVP* (1-5), *when* $\gamma = -1$, *satisfying*

$$u \in C\big(\mathbb{R}^+; H^{\frac{1}{8}(2s+3)}(0,T)\big) \cap X^{s,b}\big((0,T)\times\mathbb{R}^+\big) \ \ and \ \ \partial_x u \in C\big(\mathbb{R}^+; H^{\frac{1}{8}(2s+1)}(0,T)\big),$$

for some $b(s) < \frac{1}{2}$. *Moreover, the map* $(u_0, f, g) \mapsto u$ *is analytic from* $H^s(\mathbb{R}^+) \times H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+) \times H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+)$ *to* $C\big((0,T);\ H^s(\mathbb{R}^+)\big)$.

Remarks. Finally, the following comments are now in order:

1. The proof of Theorem 1.1 is based on the Fourier restriction method for a suitable extension of solutions. We first convert the IBVP of (1-5) posed in $\mathbb{R}^+ \times \mathbb{R}^+$ to the initial value problem (IVP) of (1-5) (integral equation formula) in the whole space $\mathbb{R} \times \mathbb{R}$ (see Section 3) by using the Duhamel boundary forcing operator. The energy and nonlinear estimates (established in Section 4) allow us to apply the Picard iteration method for IVP of (1-5), and hence we can complete the proof. The new tools used here are the Duhamel boundary forcing operator for the fourth-order linear equation and its analysis.

2. Note that Theorem 1.1 give us the local well-posedness in low regularity for the biharmonic nonlinear Schrödinger equation. However, in [Özsarı and Yolcu 2019], the authors showed the local well-posedness in the Sobolev spaces, by using Fokas approach. We point out that the low regularity in our main result is obtained using the boundary forcing operator, proposed by Holmer, which has been obtained in an independent way and with a different approach to that of [Özsarı and Yolcu 2019].

3. The approach used in our result, together with some extension as it was done in [Cavalcante 2017; Cavalcante and Kwak 2019; Holmer 2005; 2006] also guarantee the local well-posedness result in high regularity.

1F. *Notations.* In all this paper, we will consider $\mathbb{R}^+$ as $(0,\infty)$. Moreover, for positive real numbers $x, y \in \mathbb{R}^+$, we mean $x \lesssim y$ by $x \le Cy$ for some $C > 0$. Also, denote $x \sim y$ by $x \lesssim y$ and $y \lesssim x$. Similarly, $\lesssim_s$ and $\sim_s$ can be defined, where the implicit constants depend on s.

Our work is outlined in the following way: In Section 2, we introduce some function spaces defined on the half-line and construct the solution spaces. Section 3 is devoted to the introduction of the boundary forcing operator for the biharmonic Schrödinger equation. In Section 4, we show the energy estimates and present the trilinear estimates, respectively. The main result of this article, Theorem 1.1, is proved in Section 5. Finally, in Section 6, we present some open problems which seem to be of interest from the mathematical point of view.

2. Preliminaries

Throughout the paper, we fix a cut-off function $\psi(t) := \psi$,

$$\psi \in C_0^\infty(\mathbb{R}) \quad \text{such that} \quad 0 \le \psi \le 1, \quad \psi \equiv 1 \text{ on } [0,1], \quad \psi \equiv 0 \text{ for } |t| \ge 2, \tag{2-1}$$

and for $T > 0$ we denote $\psi_T(t) = \frac{1}{T}\psi\left(\frac{t}{T}\right)$.

2A. ***Sobolev spaces on the half-line.*** For $s \ge 0$, we define the homogeneous L^2-based Sobolev spaces $\dot{H}^s = \dot{H}^s(\mathbb{R})$ by the norm $\|\phi\|_{\dot{H}^s} = \big\| |\xi|^s \hat{\psi}(\xi) \big\|_{L^2_\xi}$ and the L^2-based inhomogeneous Sobolev spaces $H^s = H^s(\mathbb{R})$ by the norm $\|\phi\|_{H^s} = \|(1+|\xi|^2)^{s/2}\hat{\psi}(\xi)\|_{L^2_\xi}$, where $\hat{\phi}$ denotes the Fourier transform of ϕ. Moreover, we say that $f \in H^s(\mathbb{R}^+)$ if there exists $F \in H^s(\mathbb{R})$ such that $f(x) = F(x)$ for $x > 0$, in this case we set

$$\|f\|_{H^s(\mathbb{R}^+)} = \inf_F \|F\|_{H^s(\mathbb{R})}.$$

On the other hand for $s \in \mathbb{R}$, we have $f \in H_0^s(\mathbb{R}^+)$ provided that there exists $F \in H^s(\mathbb{R})$ such that F is the extension of f on $\mathbb{R}$ and $F(x) = 0$ for $x < 0$. In this case, we set $\|f\|_{H_0^s(\mathbb{R}^+)} = \inf_F \|F\|_{H^s(\mathbb{R})}$. For $s < 0$, we define $H^s(\mathbb{R}^+)$ as the dual space of $H_0^{-s}(\mathbb{R}^+)$.

Let us also define the sets $C_0^\infty(\mathbb{R}^+) = \{f \in C^\infty(\mathbb{R});\ \operatorname{supp} f \subset [0,\infty)\}$ and $C_{0,c}^\infty(\mathbb{R}^+)$ as the subset of $C_0^\infty(\mathbb{R}^+)$, whose members have a compact support on $(0,\infty)$. We remark that $C_{0,c}^\infty(\mathbb{R}^+)$ is dense in $H_0^s(\mathbb{R}^+)$ for all $s \in \mathbb{R}$.

We finish this subsection with some elementary properties of the Sobolev spaces.

Lemma 2.1 [Jerison and Kenig 1995, Lemma 3.5]. *For $-\frac{1}{2} < s < \frac{1}{2}$ and $f \in H^s(\mathbb{R})$,*

$$\|\chi_{(0,\infty)} f\|_{H^s(\mathbb{R})} \le c\|f\|_{H^s(\mathbb{R})}. \tag{2-2}$$

Lemma 2.2 [Colliander and Kenig 2002, Lemma 2.8]. *If $0 \le s < \frac{1}{2}$, then, for the cut-off function ψ defined in (2-1), $\|\psi f\|_{H^s(\mathbb{R})} \le c\|f\|_{\dot{H}^s(\mathbb{R})}$ and $\|\psi f\|_{\dot{H}^{-s}(\mathbb{R})} \le c\|f\|_{H^{-s}(\mathbb{R})}$, where the constant c depends only on s and ψ.*

Remark. Lemma 2.2 is equivalent to

$$\|f\|_{H^s(\mathbb{R})} \sim \|f\|_{\dot{H}^s(\mathbb{R})},$$

for $-\frac{1}{2} < s < \frac{1}{2}$, where $f \in H^s(\mathbb{R})$ with $\operatorname{supp} f \subset [0,1]$.

The following two auxiliaries lemmas can be found in [Colliander and Kenig 2002] and their proofs will be omitted.

Lemma 2.3 [Colliander and Kenig 2002, Proposition 2.4]. *If $\frac{1}{2} < s < \frac{3}{2}$, the following statements are valid*:

(a) $H_0^s(\mathbb{R}^+) = \{f \in H^s(\mathbb{R}^+);\ f(0) = 0\}$.

(b) *If $f \in H^s(\mathbb{R}^+)$ with $f(0) = 0$, then* $\|\chi_{(0,\infty)} f\|_{H_0^s(\mathbb{R}^+)} \le c\|f\|_{H^s(\mathbb{R}^+)}$.

Lemma 2.4 [Colliander and Kenig 2002, Proposition 2.5]. *Let* $-\infty < s < \infty$ *and* $f \in H_0^s(\mathbb{R}^+)$. *For the cut-off function* ψ *defined in* (2-1), *we have* $\|\psi f\|_{H_0^s(\mathbb{R}^+)} \le c\|f\|_{H_0^s(\mathbb{R}^+)}$.

2B. *Solution spaces.* For $f \in \mathcal{S}(\mathbb{R}^2)$, we denote by $\tilde{f}$ or $\mathcal{F}(f)$ the Fourier transform of f with respect to both spatial and time variables

$$\tilde{f}(\tau,\xi) = \int_{\mathbb{R}^2} e^{-ix\xi} e^{-it\tau} f(t,x)\,dx\,dt.$$

Moreover, we use $\mathcal{F}_x$ and $\mathcal{F}_t$ to denote the Fourier transform with respect to space and time variable respectively (also we use $\widehat{\ }$ for both cases).

Bourgain [1993a; 1993b] established a way to prove the well-posedness of a classes of dispersive systems. More precisely, on the Sobolev spaces H^s, for smaller values of s, Bourgain found a yet more suitable smoothing property for solutions of the Korteweg–de Vries equation.

In this spirit, for $s, b \in \mathbb{R}$, we introduce the classical Bourgain spaces $X^{s,b}$ associated to (1-2) as the completion of $\mathcal{S}'(\mathbb{R}^2)$ under the norm

$$\|f\|_{X^{s,b}}^2 = \int_{\mathbb{R}^2} \langle\xi\rangle^{2s} \langle\tau+\xi^4\rangle^{2b} |\tilde{f}(\tau,\xi)|^2\,d\xi\,d\tau,$$

where $\langle\cdot\rangle = (1+|\cdot|^2)^{1/2}$.

One basic property of $X^{s,b}$ can be read as follows:

Lemma 2.5 [Tao 2006, Lemma 2.11]. *Let* $\psi(t)$ *be a Schwartz function in time. Then*, *we have*

$$\|\psi(t) f\|_{X^{s,b}} \lesssim_{\psi,b} \|f\|_{X^{s,b}}.$$

Ginibre et al. [1997], while establishing local well-posedness results for the Zakharov system, showed the following important estimate:

Lemma 2.6. *Let* $-\frac{1}{2} < b' < b \le 0$ *or* $0 \le b' < b < \frac{1}{2}$, $w \in X^{s,b}(\phi)$ *and* $s \in \mathbb{R}$. *Then*

$$\|\psi_T w\|_{X^{s,b'}(\phi)} \le cT^{b-b'} \|w\|_{X^{s,b}(\phi)}.$$

As is well-known, the space $X^{s,b}$ with $b > \frac{1}{2}$ is well-adapted to study the IVP of dispersive equations. However, in the study of IBVP, the standard argument cannot be applied directly. This is due to the lack of hidden regularity, more precisely, the control of (derivatives) time trace norms of the Duhamel boundary operator requires to work in $X^{s,b}$-type spaces for $b < \frac{1}{2}$, since the full regularity range cannot be covered (see Lemma 4.2 inequality (4-5)).

Therefore, to treat the solution of our problem, set the solution space denoted by $Z^{s,b}$ with the norm

$$\|f\|_{Z^{s,b}(\mathbb{R}^2)} = \sup_{t\in\mathbb{R}} \|f(t,\cdot)\|_{H^s(\mathbb{R})} + \sum_{j=0}^{1} \sup_{x\in\mathbb{R}} \|\partial_x^j f(\cdot,x)\|_{H^{\frac{1}{8}(2s+3-2j)}(\mathbb{R})} + \|f\|_{X^{s,b}}.$$

The spatial and time restricted space of $Z^{s,b}(\mathbb{R}^2)$ is defined in the standard way:

$$Z^{s,b}((0,T)\times\mathbb{R}^+) = Z^{s,b}\big|_{(0,T)\times\mathbb{R}^+}$$

equipped with the norm

$$\|f\|_{Z^{s,b}((0,T)\times\mathbb{R}^+)} = \inf_{g\in Z^{s,b}} \{\|g\|_{Z^{s,b}} : g(t,x) = f(t,x) \text{ on } (0,T)\times\mathbb{R}^+\}.$$

2C. *Riemann–Liouville fractional integral.* Before we begin our study of the IBVP for (1-2), we give a brief summary of the Riemann–Liouville fractional integral operator; see [Colliander and Kenig 2002; Holmer 2006] for more details.

Let us define the function t_+ as follows:

$$t_+ = \begin{cases} t & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases}$$

The tempered distribution $t_+^{\alpha-1}/\Gamma(\alpha)$ is defined like a locally integrable function for $\operatorname{Re}\alpha > 0$ by

$$\left\langle \frac{t_+^{\alpha-1}}{\Gamma(\alpha)}, f \right\rangle = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} f(t)\,dt.$$

It follows that

$$\frac{t_+^{\alpha-1}}{\Gamma(\alpha)} = \partial_t^k \left(\frac{t_+^{\alpha+k-1}}{\Gamma(\alpha+k)} \right), \tag{2-3}$$

for all $k\in\mathbb{N}$. Expression (2-3) can be used to extend the definition of $t_+^{\alpha-1}/\Gamma(\alpha)$ for all $\alpha\in\mathbb{C}$ in the sense of distributions. In fact, a change of contour shows that the Fourier transform of $t_+^{\alpha-1}/\Gamma(\alpha)$ is

$$\left(\frac{t_+^{\alpha-1}}{\Gamma(\alpha)} \right)^{\widehat{}}(\tau) = e^{-\frac{1}{2}\pi i\alpha} (\tau - i0)^{-\alpha}, \tag{2-4}$$

where

$$(\tau - i0)^{-\alpha} = |\tau|^{-\alpha}\chi_{(0,\infty)} + e^{\alpha\pi i}|\tau|^{-\alpha}\chi_{(-\infty,0)} \tag{2-5}$$

is the distributional limit. For $\alpha\notin\mathbb{Z}$, by using (2-5), we rewrite (2-4) in the following way:

$$\left(\frac{t_+^{\alpha-1}}{\Gamma(\alpha)} \right)^{\widehat{}}(\tau) = e^{-\frac{1}{2}\alpha\pi i}|\tau|^{-\alpha}\chi_{(0,\infty)} + e^{\frac{1}{2}\alpha\pi i}|\tau|^{-\alpha}\chi_{(-\infty,0)}. \tag{2-6}$$

For $f \in C_0^\infty(\mathbb{R}^+)$, define $\mathcal{I}_\alpha f$ as

$$\mathcal{I}_\alpha f = \frac{t_+^{\alpha-1}}{\Gamma(\alpha)} * f.$$

Thus, for $\operatorname{Re}\alpha > 0$, we have

$$\mathcal{I}_\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s)\,ds. \tag{2-7}$$

The following properties easily hold:

$$\mathcal{I}_0 f = f, \quad \mathcal{I}_1 f(t) = \int_0^t f(s)\,ds, \quad \mathcal{I}_{-1} f = f' \quad \text{and} \quad \mathcal{I}_\alpha \mathcal{I}_\beta = \mathcal{I}_{\alpha+\beta}.$$

Moreover, the lemmas below can be found in [Holmer 2006], and we will omit the proofs.

Lemma 2.7 [Holmer 2006, Lemma 2.1]. *If $f \in C_0^\infty(\mathbb{R}^+)$, then $\mathcal{I}_\alpha f \in C_0^\infty(\mathbb{R}^+)$, for all $\alpha \in \mathbb{C}$.*

Lemma 2.8 [Holmer 2006, Lemma 5.3]. *If $0 \le \operatorname{Re}\alpha < \infty$ and $s \in \mathbb{R}$, then $\|\mathcal{I}_{-\alpha} h\|_{H_0^s(\mathbb{R}^+)} \le c\|h\|_{H_0^{s+\alpha}(\mathbb{R}^+)}$, where $c = c(\alpha)$.*

Lemma 2.9 [Holmer 2006, Lemma 5.4]. *If $0 \le \operatorname{Re}\alpha < \infty$, $s \in \mathbb{R}$ and $\mu \in C_0^\infty(\mathbb{R})$, then $\|\mu \mathcal{I}_\alpha h\|_{H_0^s(\mathbb{R}^+)} \le c\|h\|_{H_0^{s-\alpha}(\mathbb{R}^+)}$, where $c = c(\mu, \alpha)$.*

2D. *Oscillatory integral.* In this subsection, we will define the oscillatory integral which is the key to defining, in the next section, the Duhamel boundary forcing operator. Let

$$B(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} e^{-i\xi^4}\,d\xi. \tag{2-8}$$

We first calculate $B(0)$. A change of variable ($\eta = \xi^4$), gives us the following:

$$B(0) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi^4}\,d\xi = \frac{1}{4\pi} \int_0^{+\infty} e^{-i\eta} \eta^{-3/4}\,d\eta.$$

Now, a change of contour yields

$$B(0) = \frac{(-i)^{1-3/4}}{4\pi} \int_0^{+\infty} e^{-t} t^{(1/4)-1}\,dt = \frac{(-i)^{1/4}}{4\pi} \Gamma\Big(\frac{1}{4}\Big) = -\frac{i^{7/4}}{\pi} \Gamma\Big(\frac{5}{4}\Big).$$

Let us obtain the Mellin transform of $B(x)$.

Lemma 2.10. *For $\operatorname{Re}\lambda > 0$ we have*

$$\int_0^\infty x^{\lambda-1} B(x)\,dx = \frac{\Gamma(\lambda)\Gamma\big(\frac{1}{4} - \frac{\lambda}{4}\big)}{8\pi} \big(e^{-i\frac{\pi}{8}(1+3\lambda)} + e^{-i\frac{\pi}{8}(1-5\lambda)}\big). \tag{2-9}$$

Proof. By analytic argument, we can assume that λ is a real number in the set $\left(0, \frac{3}{8}\right)$.

Consider

$$B_1(x) = \frac{1}{2\pi} \int_0^{+\infty} e^{ix\xi} e^{-i\xi^4} \, d\xi$$

and

$$B_2(x) = \frac{1}{2\pi} \int_{-\infty}^{0} e^{ix\xi} e^{-i\xi^4} \, d\xi = \frac{1}{2\pi} \int_0^{\infty} e^{-ix\xi} e^{-i\xi^4} \, d\xi,$$

then we have $B(x) = B_1(x) + B_2(x)$. Define

$$B_{1,\epsilon}(x) = \frac{1}{2\pi} \int_0^{+\infty} e^{ix\xi} e^{-i\xi^4} e^{-\epsilon\xi} \, d\xi.$$

By using the dominated convergence theorem and Fubini's theorem we have

$$\begin{aligned} \text{(2-10)} \quad \int_0^{\infty} x^{\lambda-1} B_1(x) \, dx &= \lim_{\epsilon\to 0} \lim_{\delta\to 0} \int_0^{\infty} e^{-\delta x} x^{\lambda-1} B_{1,\epsilon}(x) \, dx \\ &= \lim_{\epsilon\to 0} \lim_{\delta\to 0} \frac{1}{2\pi} \int_0^{+\infty} e^{-i\xi^4} e^{-\epsilon\xi} \int_0^{\infty} e^{ix\xi} e^{-\delta x} x^{\lambda-1} \, dx \, d\xi. \end{aligned}$$

Using a change of contour, we get that

$$\text{(2-11)} \quad \int_0^{+\infty} e^{ix\xi} e^{-\delta x} x^{\lambda-1} \, dx = \xi^{-\lambda} e^{i\lambda\frac{\pi}{2}} \Gamma\left(\lambda, -\frac{\delta}{\xi}\right),$$

where $\Gamma(\lambda, z) = \int_0^{+\infty} r^{\lambda-1} e^{irz} e^{-r} \, dr$. Again, thanks to the dominated convergence theorem it follows that

$$\text{(2-12)} \quad \lim_{\delta\to 0} \int_0^{+\infty} e^{ix\xi} e^{-\delta x} x^{\lambda-1} \, dx = \xi^{-\lambda} e^{i\lambda\frac{\pi}{2}} \Gamma(\lambda).$$

Once more applying the dominated convergence theorem and changing the contour we conclude that

$$\begin{aligned} \text{(2-13)} \quad \int_0^{+\infty} x^{\lambda-1} B_1(x) \, dx &= \frac{\Gamma(\lambda)}{2\pi} e^{i\lambda\frac{\pi}{2}} \lim_{\epsilon\to 0} \int_0^{+\infty} e^{-i\xi^4} e^{-\epsilon\xi} \xi^{-\lambda} \, d\xi \\ &= \frac{\Gamma(\lambda)}{2\pi} e^{i\lambda\frac{\pi}{2}} \frac{1}{4} \lim_{\epsilon\to 0} \int_0^{+\infty} e^{-i\eta} e^{-\epsilon\eta^{1/4}} (\eta)^{-(\lambda+3)/4} \, d\eta \\ &= \frac{\Gamma(\lambda)}{2\pi} e^{i\lambda\frac{\pi}{2}} \frac{1}{4} e^{-\frac{\pi i}{2}((1-\lambda)/4)} \Gamma\left(\frac{1}{4} - \frac{\lambda}{4}\right) \\ &= \frac{\Gamma(\lambda)\Gamma\left(\frac{1}{4} - \frac{\lambda}{4}\right)}{8\pi} e^{-i\frac{\pi}{8}(1-5\lambda)}. \end{aligned}$$

In a similar way, by using the identity

$$\int_0^{+\infty} e^{-ix\xi} e^{-\delta x} x^{\lambda-1}\, dx = \xi^{-\lambda} e^{-i\lambda\frac{\pi}{2}} \Gamma\Big(\lambda, \frac{\delta}{\xi}\Big), \tag{2-14}$$

we obtain

$$\begin{aligned}\int_0^{+\infty} x^{\lambda-1} B_2(x)\, dx &= \frac{\Gamma(\lambda)}{2\pi} e^{-i\lambda\frac{\pi}{2}} \frac{1}{4} e^{-\frac{\pi i}{2}(\frac{1-\lambda}{4})} \Gamma\Big(\frac{1}{4} - \frac{\lambda}{4}\Big) \\ &= \frac{\Gamma(\lambda)\Gamma\big(\frac{1}{4} - \frac{\lambda}{4}\big)}{8\pi} e^{-i\frac{\pi}{8}(1+3\lambda)}.\end{aligned} \tag{2-15}$$

Finally, as we can split by $B(x) = B_1(x) + B_2(x)$, equation (2-9) holds. □

3. Duhamel boundary forcing operator

In this section, we study the Duhamel boundary forcing operator, which was introduced by Colliander and Kenig [2002], in order to construct the solution to (1-2) forced by boundary conditions. We refer to [Cavalcante 2017; Cavalcante and Corcho 2019; Holmer 2005] for further exposition about this topic.

3A. ***Duhamel boundary forcing operator class.*** Let us introduce the Duhamel boundary forcing operator associated to the linearized biharmonic Schrödinger equation. Consider

$$M = \frac{1}{B(0)\Gamma\big(\frac{3}{4}\big)}. \tag{3-1}$$

For $f \in C_0^\infty(\mathbb{R}^+)$, define the boundary forcing operator $\mathcal{L}^0$ (of order 0) as

$$\mathcal{L}^0 f(t,x) := M \int_0^t e^{i(t-t')\partial_x^4} \delta_0(x) \mathcal{I}_{-3/4} f(t')\, dt', \tag{3-2}$$

where $e^{it\partial_x^4}$ denotes the group associated to (1-6) given by

$$e^{it\partial_x^4}\psi(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} e^{-it\xi^4} \hat{\psi}(\xi)\, d\xi.$$

Note that the property of convolution operator $(\partial_x(f * g) = (\partial_x f) * g = f * (\partial_x g))$ and the integration by parts in t' of (3-2) yield that

$$i\mathcal{L}^0(\partial_t f)(t,x) = iM\delta_0(x)\mathcal{I}_{-3/4} f(t) + \partial_x^4 \mathcal{L}^0 f(t,x). \tag{3-3}$$

By a change of variable and using (2-8), we get that

$$\begin{aligned}\mathcal{L}^0 f(t,x) &= M\int_0^t e^{i(t-t')\partial_x^4}\delta_0(x)\mathcal{I}_{-3/4}f(t')\,dt' \\ &= M\int_0^t B\left(\frac{x}{(t-t')^{1/4}}\right)\frac{\mathcal{I}_{-3/4}f(t')}{(t-t')^{1/4}}\,dt'.\end{aligned} \tag{3-4}$$

We are now in a position to make it precise when the boundary forcing term is continuous or discontinuous. More precisely, the following lemma holds.

Lemma 3.1. *Let* $f\in C_{0,c}^\infty(\mathbb{R}^+)$.

(a) *For fixed* $0\le t\le 1$, *we have that* $\partial_x^k\mathcal{L}^0 f(t,x)$, $k=0,1,2$, *is continuous in* $x\in\mathbb{R}$ *and has the decay property in terms of the spatial variable as follows*:

$$|\partial_x^k\mathcal{L}^0 f(t,x)|\lesssim_N \|f\|_{H^{N+k}}\langle x\rangle^{-N},\qquad N\ge 0. \tag{3-5}$$

(b) *For fixed* $0\le t\le 1$, *we have that* $\partial_x^3\mathcal{L}^0 f(t,x)$ *is continuous in* x *for* $x\neq 0$ *and it is discontinuous at* $x=0$ *satisfying*

$$\lim_{x\to 0^-}\partial_x^3\mathcal{L}^0 f(t,x) = -i\frac{M}{2}\mathcal{I}_{-3/4}f(t),\qquad \lim_{x\to 0^+}\partial_x^3\mathcal{L}^0 f(t,x) = i\frac{M}{2}\mathcal{I}_{-3/4}f(t).$$

Also, $\partial_x^3\mathcal{L}^0 f(t,x)$ *has the decay property in terms of the spatial variable*

$$|\partial_x^3\mathcal{L}^0 f(t,x)|\lesssim_N \|f\|_{H^{N+3}}\langle x\rangle^{-N},\qquad N\ge 0. \tag{3-6}$$

Proof. In fact, the continuity of $\partial_x^k\mathcal{L}^0 f(t,x)$ follows from (3-4), for $k=0,1,2$, and the proof of (3-5) exactly follows the idea introduced by Holmer [2005, Lemma 12]. Moreover, (3-5) and (3-3) yield that $\partial_x^4\mathcal{L}^0 f(t,x)$ is discontinuous only at $x=0$ of size $M\mathcal{I}_{-3/4}f(t)$ (where M is defined as (3-1)), and the decay bound (3-6) holds. □

Remark. Lemma 3.1 ensures that $\mathcal{L}^0 f(t,0)=f(t)$.

We are now in position to generalize the boundary forcing operator (3-2). For $\operatorname{Re}\lambda>-4$ and given $g\in C_0^\infty(\mathbb{R}^+)$, we define

$$\mathcal{L}^\lambda g(t,x) = \left[\frac{x_-^{\lambda-1}}{\Gamma(\lambda)} * \mathcal{L}^0(\mathcal{I}_{-\lambda/4}g)(t,\cdot)\right](x), \tag{3-7}$$

where $*$ denotes the convolution operator and $x_-^{\lambda-1}/\Gamma(\lambda) = (-x)_+^{\lambda-1}/\Gamma(\lambda)$. In particular, for $\operatorname{Re}\lambda>0$, we have

$$\mathcal{L}^\lambda g(t,x) = \frac{1}{\Gamma(\lambda)}\int_x^\infty (y-x)^{\lambda-1}\mathcal{L}^0(\mathcal{I}_{-\lambda/4}g)(t,y)\,dy. \tag{3-8}$$

A property of the convolution operator ($\partial_x^4(f*g)=(\partial_x^4 f)*g=f*(\partial_x^4 g)$) and (3-3) give us

$$\begin{aligned}\mathcal{L}^\lambda g(t,x) &= \left[\frac{x_-^{(\lambda+4)-1}}{\Gamma(\lambda+4)} * \partial_x^4 \mathcal{L}^0(\mathcal{I}_{-\lambda/4} g)(t,\cdot)\right](x) \\ &= iM\frac{x_-^{(\lambda+4)-1}}{\Gamma(\lambda+4)}\mathcal{I}_{-3/4-\lambda/4} g(t) \\ &\qquad + i\int_x^\infty \frac{(y-x)^{(\lambda+4)-1}}{\Gamma(\lambda+4)}\mathcal{L}^0(\partial_t \mathcal{I}_{-\lambda/4} g)(t,y)\,dy,\end{aligned} \tag{3-9}$$

for $\operatorname{Re}\lambda > -4$, where M is defined as in (3-1). From (3-3) and (3-7), we have

$$(i\partial_t - \partial_x^4)\mathcal{L}^\lambda g(t,x) = iM\frac{x_-^{\lambda-1}}{\Gamma(\lambda)}\mathcal{I}_{-3/4-\lambda/4} g(t),$$

in the distributional sense.

To finish this subsection, we will give two lemmas concerning the spatial continuity and decay properties of the $\mathcal{L}^\lambda g(t,x)$ and the explicit values for $\mathcal{L}^\lambda f(t,0)$, respectively.

Lemma 3.2. *Let $g \in C_0^\infty(\mathbb{R}^+)$ and M be as in (3-1). Then, we have*

$$\mathcal{L}^{-k} g = \partial_x^k \mathcal{L}^0 \mathcal{I}_{k/4} g, \qquad k = 0, 1, 2, 3. \tag{3-10}$$

Moreover, $\mathcal{L}^{-3} g(t,x)$ is continuous in $x \in \mathbb{R}\setminus\{0\}$ and has a step discontinuity at $x=0$. For real λ satisfying $\lambda > -3$, $\mathcal{L}^\lambda g(t,x)$ is continuous in $x\in\mathbb{R}$. For $-3 \le \lambda \le 1$ and $0 \le t \le 1$, $\mathcal{L}^\lambda g(t,x)$ satisfies the following decay bounds:

$$\begin{aligned}|\mathcal{L}^\lambda g(t,x)| &\le c_{\lambda,g}\langle x\rangle^{\lambda-1} && \text{for all } x \ge 0,\\ |\mathcal{L}^\lambda g(t,x)| &\le c_{m,\lambda,g}\langle x\rangle^{-m} && \text{for all } x \ge 0 \text{ and } m \ge 0.\end{aligned}$$

Proof. We give a sketch of the proof. The detailed argument can be found in [Holmer 2006]. By using (3-9), we have that (3-10) follows. Moreover, Lemma 3.1 together with (3-10) guarantee the continuity (except for $x=0$ when $\lambda=-3$) and discontinuity at $x=0$ of $\mathcal{L}^\lambda g$ for $\lambda \ge -3$ and $\lambda=-3$, respectively. The proof of decay bounds can be obtained by using (3-9), (3-3) and Lemma 3.1. □

Lemma 3.3. *For $\operatorname{Re}\lambda > -4$ and $f \in C_0^\infty(\mathbb{R}^+)$, we have the following value of $\mathcal{L}^\lambda f(t,0)$*:

$$\mathcal{L}^\lambda f(t,0) = \frac{M}{8} f(t)\left(\frac{e^{-i\frac{\pi}{8}(1+3\lambda)} + e^{-i\frac{\pi}{8}(1-5\lambda)}}{\sin\left(\frac{1}{4}(1-\lambda)\pi\right)}\right). \tag{3-11}$$

Proof. By using (3-9) we get

$$\mathcal{L}^\lambda f(t,0) = i\int_0^\infty \frac{y^{(\lambda+4)-1}}{\Gamma(\lambda+4)}\mathcal{L}^0(\partial_t\mathcal{I}_{-\lambda/4}f)(t,y)\,dy.$$

This show that $\mathcal{L}^\lambda f(t,0)$ is analytic, in λ, for $\operatorname{Re}\lambda > -4$.

By analytic argument, it suffices to consider the case when λ is a positive real number and (3-4), where M is defined as in (3-1). In fact, in order to use (2-9), we take $\lambda \in \left(0, \frac{3}{8}\right)$ in (3-8). Thus, in the calculations, we use the representation (3-8) for $\lambda > 0$. Fubini's theorem, the change of variable, (2-10) and (2-7), yield that

$$\begin{aligned}
\mathcal{L}^\lambda f(t,0) &= \frac{M}{\Gamma(\lambda)}\int_0^\infty y^{\lambda-1}\int_0^t B\left(\frac{y}{(t-t')^{1/4}}\right)\frac{\mathcal{I}_{(-\lambda-3)/4}f(t')}{(t-t')^{1/4}}\,dt'\,dy\\
&= \frac{M}{\Gamma(\lambda)}\int_0^t (t-t')^{\frac{\lambda+3}{4}-1}\mathcal{I}_{(-\lambda-3)/4}f(t')\int_0^\infty y^{\lambda-1}B(y)\,dy\,dt'\\
&= \frac{M}{\Gamma(\lambda)}\Gamma\left(\frac{\lambda}{4}+\frac{3}{4}\right)f(t)\frac{\Gamma(\lambda)\Gamma\left(\frac{1}{4}-\frac{\lambda}{4}\right)}{8\pi}\left(e^{-i\frac{\pi}{8}(1+3\lambda)}+e^{-i\frac{\pi}{8}(1-5\lambda)}\right)\\
&= \frac{M}{8}f(t)\left(\frac{e^{-i\frac{\pi}{8}(1+3\lambda)}+e^{-i\frac{\pi}{8}(1-5\lambda)}}{\sin\left(\frac{1}{4}(1-\lambda)\pi\right)}\right),
\end{aligned}$$

where in the last equality we used the fact that

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin\,\pi z}.$$

Thus, the proof is complete. □

3B. ***Construction of the solution.*** Let us describe how we can construct the solution for the linear fourth order Schrödinger equation

$$i\partial_t u - \partial_x^4 u = 0. \tag{3-12}$$

3B1. *Linear version.* First, we define the unitary group associated to (3-12) as

$$e^{it\partial_x^4}\phi(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}e^{-it\xi^4}\hat{\phi}(\xi)\,d\xi,$$

which allows

$$\begin{cases} (i\partial_t - \partial_x^4)e^{it\partial_x^4}\phi(x) = 0, & (t,x)\in\mathbb{R}\times\mathbb{R},\\ e^{it\partial_x^4}\phi(x)|_{t=0} = \phi(x), & x\in\mathbb{R}.\end{cases} \tag{3-13}$$

Recall $\mathcal{L}^\lambda$ in (3-9) for the right half-line problem. Let

$$\begin{aligned}
u(t,x) &= \mathcal{L}^{\lambda_1}\gamma_1(t,x) + \mathcal{L}^{\lambda_2}\gamma_2(t,x),\\
\partial_x u(t,x) &= \mathcal{L}^{\lambda_1-1}\mathcal{I}_{-1/4}\,\gamma_1(t,x) + \mathcal{L}^{\lambda_2-1}\mathcal{I}_{-1/4}\,\gamma_2(t,x),
\end{aligned}$$

where γ_j $(j=1,2)$ will be chosen later in terms of the given boundary data f and g.

Let a_j and b_j be constants depending on λ_j, $j = 1, 2$, given by

$$
\begin{aligned}
a_j &= \frac{M}{8}\left(\frac{e^{-i\frac{\pi}{8}(1+3\lambda_j)} + e^{-i\frac{\pi}{8}(1-5\lambda_j)}}{\sin\big(\frac{1}{4}(1-\lambda_j)\pi\big)}\right),\\
b_j &= \frac{M}{8}\left(\frac{e^{-i\frac{\pi}{8}(-2+3\lambda)} + e^{-i\frac{\pi}{8}(6-5\lambda)}}{\sin\big(\frac{1}{4}(2-\lambda_j)\pi\big)}\right).
\end{aligned}
\tag{3-14}
$$

By Lemmas 3.2 and 3.3, we get

$$f(t) = u(t, 0) = a_1\gamma_1(t) + a_2\gamma_2(t), \tag{3-15}$$

$$g(t) = \partial_x u(t, 0) = b_1\mathcal{I}_{-1/4}\,\gamma_1(t) + b_2\mathcal{I}_{-1/4}\,\gamma_2(t). \tag{3-16}$$

Thanks to (3-15) and (3-16), we can write a matrix in the form

$$\begin{bmatrix} f(t) \\ \mathcal{I}_{1/4}\, g(t) \end{bmatrix} = A \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix},$$

where

$$A(\lambda_1, \lambda_2) = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}.$$

Choosing appropriate λ_j, $j = 1, 2$, such that A is invertible, we have that u solves

$$
\begin{cases}
(i\partial_t - \partial_x^4)u(t, x) = iM\frac{x_-^{\lambda_1-1}}{\Gamma(\lambda_1)}\mathcal{I}_{-3/4-\lambda/4}\,\gamma_1(t) \\
\qquad\qquad + iM\frac{x_-^{\lambda_2-1}}{\Gamma(\lambda)}\mathcal{I}_{-3/4-\lambda/4}\,\gamma_2(t), & (t, x) \in \mathbb{R}^+ \times \mathbb{R}, \\
u(0, x) = 0, & x \in \mathbb{R}, \\
u(t, 0) = f(t), \quad \partial_x u(t, 0) = g(t), & t \in \mathbb{R}^+.
\end{cases}
\tag{3-17}
$$

By restriction of the function u on the set $\mathbb{R}^+ \times \mathbb{R}^+$, we can construct a solution for the linear fourth order dispersive equation (3-12) posed on the right half-line.

3B2. *Nonlinear version.* Now, we define the classical Duhamel inhomogeneous solution operator $\mathcal{D}$ by

$$\mathcal{D}w(t, x) = -i\int_0^t e^{i(t-t')\partial_x^4} w(t', x)\, dt'.$$

It follows that

$$
\begin{cases}
(i\partial_t - \partial_x^4)\mathcal{D}w(t, x) = w(t, x), & (t, x) \in \mathbb{R} \times \mathbb{R}, \\
\mathcal{D}w(x, 0) = 0, & x \in \mathbb{R}.
\end{cases}
$$

Let

$$u(t, x) = \mathcal{L}^{\lambda_1}\gamma_1(t, x) + \mathcal{L}^{\lambda_2}\gamma_2(t, x) + e^{it\partial_x^4}\phi(x) + \mathcal{D}w.$$

Similar to what was done in Section 3B, taking γ_1 and γ_2 appropriately, depending on f, g, $e^{it\partial_x^4}\phi(x)$ and $\mathcal{D}w$, we see that u solves

$$\begin{cases} (i\partial_t - \partial_x^4)u(t,x) = w(t,x), & (t,x) \in \mathbb{R}^+ \times \mathbb{R}^+, \\ u(0,x) = \phi(x), & x \in \mathbb{R}^+, \\ u(t,0) = f(t), \ \ \partial_x u(t,0) = g(t), & t \in \mathbb{R}^+. \end{cases} \tag{3-18}$$

The discussion about the structure of the system (3-18) can be found in Section 5

4. Energy estimates

The main purpose of this section is to prove the energy estimate of the solutions of the fourth order nonlinear Schrödinger equation in the Bourgain spaces $X^{s,b}$.

Lemma 4.1. *Let $s \in \mathbb{R}$ and $b \in \mathbb{R}$. If $\phi \in H^s(\mathbb{R})$, then the following estimates hold*:

$$\|\psi(t)e^{it\partial_x^4}\phi(x)\|_{C_t(\mathbb{R};\, H_x^s(\mathbb{R}))} \lesssim_\psi \|\phi\|_{H^s(\mathbb{R})}, \tag{4-1}$$

$$\|\psi(t)\partial_x^j e^{it\partial_x^4}\phi(x)\|_{C_x(\mathbb{R}; H_t^{(2s+3-2j)/8}(\mathbb{R}))} \lesssim_{\psi,s,j} \|\phi\|_{H^s(\mathbb{R})}, \qquad j \in \{0,1\}; \tag{4-2}$$

$$\|\psi(t)e^{it\partial_x^4}\phi(x)\|_{X^{s,b}} \lesssim_{\psi,b} \|\phi\|_{H^s(\mathbb{R})}. \tag{4-3}$$

Estimates (4-1), (4-2) and (4-3) are so-called space traces, derivative time traces and Bourgain spaces estimates, respectively.

Proof. The proofs of (4-1) and (4-3) are standard and the proof of (4-2) follows from the smoothness of ψ and the local smoothing estimate (1-8), thus we will omit the details. □

Lemma 4.2. *Let $0 < b < \frac{1}{2}$ and $j = 0, 1$, we have the following inequalities*

$$\|\psi(t)\mathcal{D}w(t,x)\|_{C(\mathbb{R}_t;\, H^s(\mathbb{R}_x))} \lesssim \|w\|_{X^{s,-b}}, \tag{4-4}$$

for $s \in \mathbb{R}$;

$$\|\psi(t)\partial_x^j\mathcal{D}w(t,x)\|_{C(\mathbb{R}_x; H^{(2s+3-2j)/8}(\mathbb{R}_t))} \lesssim \|w\|_{X^{s,-b}}, \tag{4-5}$$

for $-\frac{3}{2} + j < s < \frac{1}{2} + j$;

$$\|\psi(t)\partial_x^j\mathcal{D}w(t,x)\|_{X^{s,b}} \lesssim \|w\|_{X^{s,-b}}, \tag{4-6}$$

for $s \in \mathbb{R}$.

Estimates (4-4), (4-5) and (4-6) are so-called space traces, derivative time traces and Bourgain spaces estimates, respectively.

Proof. The idea to prove this lemma follows a variation of the proof due to [Kenig et al. 1991]. Here, we will give the sketch of the proof for sake of completeness.

Estimate (4-4): By using $2\chi_{(0,t)}(t') = \operatorname{sgn} t' + \operatorname{sgn}(t-t')$, $\widehat{\operatorname{sgn}}(\tau) = \text{p.v.}\frac{2}{i\tau}$ and $e^{i\tau\xi^4}\hat{f}(\tau) = \hat{f}(\tau+\xi^4)$ we have

$$\psi(t)\mathcal{D}w(t,x) = c\int e^{ix\xi}e^{-it\xi^4}\psi(t)\int \widetilde{w}(\tau',\xi)\frac{e^{it(\tau'+\xi^4)}-1}{(\tau'+\xi^4)}\,d\tau'\,d\xi. \tag{4-7}$$

We denote by $w = w_1 + w_2$, where

$$\widetilde{w}_1(\tau,\xi) = \eta_0(\tau+\xi^4)\widetilde{w}(\tau,\xi),$$
$$\widetilde{w}_2(\tau,\xi) = (1-\eta_0(\tau+\xi^4))\widetilde{w}(\tau,\xi).$$

Here, $\eta_0 : \mathbb{R} \to [0, 1]$ is a smooth bump function supported in $[-2, 2]$ and equal to 1 in $[-1, 1]$. For w_1, we use the Taylor expansion of e^x at $x = 0$. Then, we can rewrite (4-7) for w_1 as

$$\begin{aligned}
\psi(t)\mathcal{D}w(t,x) &= c\int e^{ix\xi}e^{-it\xi^4}\psi(t)\int \widetilde{w}_1(\tau',\xi)\frac{e^{it(\tau'+\xi^4)}-1}{(\tau'+\xi^4)}\,d\tau'\,d\xi \\
&= c\sum_{k=1}^{\infty}\frac{i^{k-1}}{k!}\psi^k(t)\int e^{ix\xi}e^{-it\xi^4}\widehat{F}_1^k(\xi)\,d\xi \\
&= c\sum_{k=1}^{\infty}\frac{i^{k-1}}{k!}\psi^k(t)e^{it\partial_x^4}F_1^k(x),
\end{aligned}$$

where $\psi^k(t) = t^k\psi(t)$ and

$$\widehat{F}_1^k(\xi) = \int \widetilde{w}_1(\tau,\xi)(\tau+\xi^4)^{k-1}\,d\tau. \tag{4-8}$$

Since

$$\|F_1^k\|_{H^s} = \left(\int \langle\xi\rangle^{2s}\left|\int \widetilde{w}_1(\tau,\xi)(\tau+\xi^4)^{k-1}\,d\tau\right|^2 d\xi\right)^{\frac{1}{2}} \lesssim \|w\|_{X^{s,-b}}, \tag{4-9}$$

we have from (4-1) that

$$\|\psi(t)\mathcal{D}w(t,x)\|_{C_tH^s} \lesssim \sum_{k=1}^{\infty}\frac{1}{k!}\|F_1^k\|_{H_x^s} \lesssim \|w\|_{X^{s,-b}}.$$

For w_2, a direct calculation gives

$$\mathcal{F}[\psi\mathcal{D}w](\tau,\xi) = c\int \widetilde{w}_2(\tau',\xi)\frac{\widehat{\psi}(\tau-\tau') - \widehat{\psi}(\tau+\xi^4)}{(\tau'+\xi^4)}\,d\tau'. \tag{4-10}$$

Since $\|\psi\mathcal{D}w\|_{C_tH^s} \lesssim \|\langle\xi\rangle^s\mathcal{F}[\psi\mathcal{D}w](\tau,\xi)\|_{L_\xi^2L_\tau^1}$, it suffices to bound the term

$$\left(\int \langle\xi\rangle^{2s}\left|\int |\widetilde{w}_2(\tau',\xi)|\int \frac{|\widehat{\psi}(\tau-\tau') - \widehat{\psi}(\tau+\xi^4)|}{|\tau'+\xi^4|}\,d\tau\,d\tau'\right|^2 d\xi\right)^{\frac{1}{2}}, \tag{4-11}$$

due to (4-10). We use the L^1 integrability of $\widehat{\psi}$, to bound (4-11) by

$$c\left(\int \langle\xi\rangle^{2s}\left|\int_{|\tau'+\xi^4|>1} \frac{|\widetilde{w}_2(\tau',\xi)|}{|\tau'+\xi^4|}\,d\tau'\right|^2 d\xi\right)^{\frac{1}{2}} \lesssim \|w\|_{X^{s,-b}}.$$

Estimate (4-5): We only consider the case $j=0$, since the estimate for $j=1$ is a direct consequence of the case $j=0$. Initially, take $\theta(\tau)\in C^\infty(\mathbb{R})$ such that $\theta(\tau)=1$ for $|\tau|<\frac{1}{2}$ and $\operatorname{supp}\theta\subset\left[-\frac{2}{3},\frac{2}{3}\right]$. A standard calculation gives

$$\begin{aligned}
\mathcal{F}_x\left(\psi(t)\int_0^t e^{(t-t')\partial_x^4}w(x,t')\right)(\xi)&\\
&= c\psi(t)\int_\tau \frac{e^{it\tau}-e^{-it\xi^4}}{\tau+\xi^4}\widetilde{w}(\xi,\tau)\,d\tau\\
&= c\psi(t)e^{it\xi^4}\int_\tau \frac{e^{-it(\tau+\xi^4)}-1}{\tau+\xi^4}\theta(\tau+\xi^4)\widetilde{w}(\xi,\tau)\,d\tau\\
&\qquad + c\psi(t)\int_\tau e^{it\tau}\frac{1-\theta(\tau+\xi^4)}{\tau+\xi^4}\widetilde{w}(\xi,\tau)\,d\tau\\
&\qquad\qquad - c\psi(t)e^{it\xi^4}\int_\tau \frac{1-\theta(\tau+\xi^4)}{\tau+\xi^4}\widetilde{w}(\xi,\tau)\,d\tau\\
&:= \mathcal{F}_x w_1+\mathcal{F}_x w_2-\mathcal{F}_x w_3.
\end{aligned}$$

By the power series expansion for $e^{-it(\tau+\xi^4)}$, we have

$$w_1(x,t)=\sum_{k=1}^{\infty}\frac{\psi_k(t)}{k!}e^{it\partial_x^4}\phi_k(x).$$

Here, $\psi_k(t)=i^k t^k\theta(t)$ and

$$\hat{\phi}_k(\xi)=\int_\tau (\tau+\xi^4)^{k-1}\theta(\tau+\xi^4)\widetilde{w}(\xi,\tau)\,d\tau.$$

By using (4-2), it suffices to show that $\|\phi_k\|_{H^s(\mathbb{R})}\le c\|u\|_{X^{s,-b}}$, for $b<\frac{1}{2}$. Using the definition of ϕ_k and the Cauchy–Schwarz inequality, it follows that

$$\begin{aligned}
\|\phi_k\|^2_{H^s(\mathbb{R}_x)} &= c\int_\xi \langle\xi\rangle^{2s}\left(\int_{\{\tau:|\tau+\xi^4|\le\frac{2}{3}\}}\sum_{k=1}^{\infty}(\tau+\xi^4)^{k-1}\theta(\tau+\xi^4)\tilde{u}(\xi,\tau)\right)^2 d\xi\\
&\le c\int_\xi \langle\xi\rangle^{2s}\int_\tau \langle\tau+\xi^4\rangle^{2c}|\tilde{u}(\xi,\tau)|^2\,d\tau\,d\xi.
\end{aligned}$$

This completes the estimate of w_1. Now we treat w_2. By using the change of variable $\eta = \xi^4$ and the Cauchy–Schwarz inequality we obtain

$$\|w_2\|^2_{C(\mathbb{R}_x;\, H^{(2s+3)/8}(\mathbb{R}_t))} \le c \int_\tau \langle\tau\rangle^{(2s+3)/4} G(\tau) \int_\xi \langle \tau+\xi^4\rangle^{-2b} \langle\xi\rangle^{2s} |\widetilde{w}_2(\xi,\tau)|^2 \, d\xi \, d\tau,$$

where $G(\tau) = c\int_\eta \langle\tau+\eta\rangle^{-2+2b}|\eta|^{-3/4}\langle\eta\rangle^{-s/2}\,d\eta$. To conclude the estimate of w_2, we need to prove the following estimate:

$$G(\tau) \le c\langle\tau\rangle^{-(2s+3)/4}. \tag{4-12}$$

We split it in two cases. In the first case, we consider $|\tau| < 1$. For this, we use $\langle\tau+\eta\rangle \sim \langle\eta\rangle$ to get

$$G(\tau) \le c\int \langle\eta\rangle^{-2+2b-(s/2)} |\eta|^{-3/4}\, d\eta.$$

The above integral is bounded in the case $s > -\frac{7}{2} + 4b$, since $-b > -\frac{1}{2}$. Also, this estimate is valid for $s > -\frac{3}{2}$.

Now, the second case $|\tau| \ge 1$ can be estimated by separating the integral into three regions $|\eta| \le 1$, $2|\eta| \le |\tau|$, $|\tau| \le 2|\eta|$ and using that $-\frac{3}{2} < s \le \frac{1}{2}$, so (4-12) follows.

Finally, to bound w_3, let us rewrite w_3 like $w_3 = \psi(t)e^{it\partial_x^4}\phi(x)$, where

$$\hat{\phi}(\xi) = \int \frac{1-\theta(\tau+\xi^4)}{\tau+\xi^4}\widetilde{w}(\xi,\tau)\,d\tau.$$

Thanks to (4-2) and Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\|w_3\|^2_{C(\mathbb{R}_x;H^{(2s+3)/8}(\mathbb{R}_t))} &= c\|\psi(t)e^{it\partial_x^4}\phi(x)\|^2_{C(\mathbb{R}_x;H^{(2s+3)/8}(\mathbb{R}_t))} \le c\|\phi\|^2_{H^s(\mathbb{R})} \\ &\le c\int_\xi \langle\xi\rangle^{2s}\left(\int_\tau |\widetilde{w}(\xi,\tau)|^2\langle\tau+\xi^4\rangle^{-2b}\,d\tau \int \frac{d\tau}{\langle\tau+\xi^4\rangle^{2-2b}}\right) d\xi.\end{aligned}$$

Since $b < \frac{1}{2}$, we have

$$\int \frac{1}{\langle\tau+\xi^4\rangle^{2-2b}}\,d\tau \le c.$$

By using (4-12), estimate (4-5) for w_3 follows and, consequently, (4-5) holds true for $w = w_1 + w_2 + w_3$.

Estimate (4-6): Finally, again we split $w = w_1 + w_2$, similar to what was done in the proof of (4-4). For w_1, estimates (4-3) and (4-9) yield that

$$\|\psi(t)\mathcal{D}w_1(t,x)\|_{X^{s,b}} \lesssim \sum_{k=1}^{\infty}\frac{1}{k!}\|F_1^k\|_{H_x^s} \lesssim \|w\|_{X^{s,-b}},$$

where F_1^k is defined as in (4-8).

For w_2, note that

$$\begin{aligned}\psi\partial_x^j\mathcal{D}w(t,x) &= c\int e^{ix\xi}e^{-it\xi^4}(i\xi)^j\psi(t)\int\frac{\widetilde{w}(\tau',\xi)}{(\tau'+\xi^4)}\big(e^{it(\tau'+\xi^4)}-1\big)\,d\tau'\,d\xi\\ &= c\int e^{ix\xi}e^{-it\xi^4}(i\xi)^j\psi(t)\int\frac{\widetilde{w}(\tau',\xi)}{(\tau'+\xi^4)}e^{it(\tau'+\xi^4)}\,d\tau'\,d\xi\\ &\qquad -c\int e^{ix\xi}e^{-it\xi^4}(i\xi)^j\psi(t)\int\frac{\widetilde{w}(\tau',\xi)}{(\tau'+\xi^4)}\,d\tau'\,d\xi\\ &= I-II.\end{aligned}$$

Let

$$\widehat{W}(\xi)=\int\frac{\widetilde{w}_2(\tau,\xi)}{(\tau+\xi^4)}\,d\tau.$$

Therefore, we use (4-3) in II to obtain

$$\|\psi e^{it\partial_x^4}W\|_{X^{s,b}}\lesssim\|W\|_{H^s}\lesssim\|w\|_{X^{s,-b}}$$

for $b<\frac{1}{2}$.

Now, it remains to show the following estimate:

$$\left(\int_{|\xi|>1}|\xi|^{2s}\int\langle\tau+\xi^4\rangle^{2b}\left|\int\frac{\widetilde{w}_2(\tau',\xi)}{i(\tau'+\xi^4)}\widehat{\psi}(\tau-\tau')\,d\tau'\right|^2d\tau\,d\xi\right)^{\frac{1}{2}}\lesssim\|w\|_{X^{s,-b}}. \tag{4-13}$$

This follows by using the same argument as we used to prove (4-5). In fact, the proof of (4-13) is easier than proof of (4-5), since the L^2 integral with respect to ξ is negligible and hence it is enough to consider the relation between $\tau+\xi^4$ and $\tau'+\xi^4$. Thus, as a consequence, we have

$$\|\psi\mathcal{D}w\|_{X^{s,b}}\lesssim\|w\|_{X^{s,-b}}.$$

Therefore, Lemma 4.2 is proved. □

Lemma 4.3. *Let $s\in\mathbb{R}$.*

(a) *For $\frac{1}{2}(2s-7)<\lambda<\frac{1}{2}(1+2s)$ and $\lambda<\frac{1}{2}$ the following inequality holds*:

$$\|\psi(t)\mathcal{L}^\lambda f(t,x)\|_{C(\mathbb{R}_t;\,H^s(\mathbb{R}_x))}\le c\|f\|_{H_0^{(2s+3)/8}(\mathbb{R}^+)}. \tag{4-14}$$

(b) *For $-4+j<\lambda<1+j$, $j=0,1$, we have*

$$\|\psi(t)\partial_x^j\mathcal{L}^\lambda f(t,x)\|_{C(\mathbb{R}_x;\,H_0^{(2s+3-2j)/8}(\mathbb{R}_t^+))}\le c\|f\|_{H_0^{(2s+3)/8}(\mathbb{R}^+)}. \tag{4-15}$$

(c) *If $s<4-4b$, $b<\frac{1}{2}$, $-5<\lambda<\frac{1}{2}$ and $s+4b-2<\lambda<s+\frac{1}{2}$ yields that*

$$\|\psi(t)\mathcal{L}^\lambda f(t,x)\|_{X^{s,b}}\le c\|f\|_{H_0^{(2s+3)/8}(\mathbb{R}^+)}. \tag{4-16}$$

Estimates (4-14), (4-15) and (4-16) are so-called space traces, derivative time traces and Bourgain spaces estimates, respectively.

Proof. Let us first prove (4-14). By density, we may assume that $f \in C^{\infty}_{0,c}(\mathbb{R}^+)$. Moreover, from definition of $\mathcal{L}^{\lambda}$, it suffices to consider $\mathcal{L}^{\lambda} f(t,x)$ (removing ψ) for $\operatorname{supp} f \subset [0,1]$, thanks to Lemma 2.4.

From (2-4), (3-2) and (3-7), we see that

$$\mathcal{F}_x(\mathcal{L}^{\lambda} f)(t,\xi) = M e^{-i\pi\lambda/2}(\xi - i0)^{-\lambda} \int_0^t e^{-i(t-t')\xi^4} \mathcal{I}_{-\lambda/4-3/4} f(t')\,dt'.$$

By using the following change of variable $\eta = \xi^4$, (2-5) and the definition of the Fourier transform we have that

$$\begin{aligned}\|\mathcal{L}^{\lambda} f(t,\cdot)\|^2_{H^s(\mathbb{R})} &\le c \int_{\eta} |\eta|^{-\lambda/2-3/4} \langle \eta \rangle^{s/2} \left| \int_0^t e^{-i(t-t')\eta} \mathcal{I}_{-\lambda/4-3/4} f(t')\,dt' \right|^2 d\eta \\ &= c \int_{\eta} |\eta|^{-\lambda/2-3/4} \langle \eta \rangle^{s/2} \left| (\chi_{(-\infty,t)} \mathcal{I}_{-\lambda/4-3/4} f)\widehat{\ }(\eta) \right|^2 d\eta,\end{aligned}$$

for a fixed t. Note that, by Lemma 2.2, we can replace $|\eta|^{-\lambda/2-3/4}$ by $\langle \eta \rangle^{-\lambda/2-3/4}$, since

$$-1 < -\frac{\lambda}{2} - \frac{3}{4} \Leftrightarrow \lambda < \frac{1}{2}.$$

Moreover, Lemma 2.1 (under the condition $-1 < -\frac{\lambda}{2} - \frac{3}{4} + \frac{s}{2} < 1$ for removing $\chi_{(-\infty,t)}$) and Lemma 2.8 (under the condition $-5 < \lambda$) yield that

$$\begin{aligned}&\int_{\eta} |\eta|^{-\lambda/2-3/4} \langle \eta \rangle^{s/2} \left| (\chi_{(-\infty,t)} \mathcal{I}_{-\lambda/4-3/4} f)\widehat{\ }(\eta) \right|^2 d\eta \\ &\qquad \le c \int_{\eta} \langle \eta \rangle^{s/2-\lambda/2-3/4} \left| (\chi_{(-\infty,t)} \mathcal{I}_{-\lambda/4-3/4} f)\widehat{\ }(\eta) \right|^2 d\eta \\ &\qquad \le c \|\mathcal{I}_{-\lambda/4-3/4} f\|^2_{H^{s/4-\lambda/4-3/8}} \le c \|f\|^2_{H_0^{(2s+3)/8}},\end{aligned}$$

which proves (4-14) thanks to the definition of $H_0^s(\mathbb{R}^+)$-norm.

Now we prove (4-15). A direct calculation gives

$$\partial_x^j \mathcal{L}^{\lambda} f = \mathcal{L}^{\lambda - j}(\mathcal{I}_{-j/4} f).$$

With the previous equality in hand and Lemma 2.8, it suffices to show (4-15) for $j = 0$. Lemma 2.4 ensures us to ignore the cut-off function ψ. The change of variable $t \to t - t'$ gives

$$\begin{aligned}&(I - \partial_t^2)^{(2s+3)/16} \left(\frac{x_-^{\lambda-1}}{\Gamma(\lambda)} * \int_{-\infty}^t e^{i(t-t')\partial_x^4} \delta(x) h(t')dt' \right) \\ &\qquad = \left(\frac{x_-^{\lambda-1}}{\Gamma(\lambda)} * \int_{-\infty}^t e^{i(t-t')\partial_x^4} \delta(x) (I - \partial_{t'}^2)^{(2s+3)/16} h(t')dt' \right).\end{aligned}$$

So, we just need to prove that

$$\left\| \int_\xi e^{ix\xi}(\xi - i0)^{-\lambda} \int_{-\infty}^{t} e^{-i(t-t')\xi^4} (\mathcal{I}_{-\frac{\lambda}{4}-\frac{3}{4}} f)(t')\,dt'\,d\xi \right\|_{L_x^\infty L_t^2(\mathbb{R})} \le c\|f\|_{L_t^2(\mathbb{R}^+)}, \tag{4-17}$$

thanks to $\partial_t^\sigma(\mathcal{I}_\alpha f) = \mathcal{I}_\alpha(\partial_t^\sigma f)$. We use $\chi_{(-\infty,t)} = \frac{1}{2}\operatorname{sgn}(t-t') + \frac{1}{2}$ to obtain

$$\begin{aligned}
\int_\xi e^{ix\xi}(\xi - i0)^{-\lambda} \int_{-\infty}^{t} & e^{-i(t-t')\xi^4} (\mathcal{I}_{-\lambda/4-3/4} f)(t')\,dt'\,d\xi \\
&= \frac{1}{2}\int_\xi e^{ix\xi}(\xi - i0)^{-\lambda} \int_{-\infty}^{\infty} \operatorname{sgn}(t-t') e^{-i(t-t')\xi^4} (\mathcal{I}_{-\lambda/4-3/4} f)(t')\,dt'\,d\xi \\
&\qquad + \frac{1}{2}\int_\xi e^{ix\xi}(\xi - i0)^{-\lambda} \int_{-\infty}^{\infty} e^{-i(t-t')\xi^4} (\mathcal{I}_{-\lambda/4-3/4} f)(t')\,dt'\,d\xi \\
&:= I(t,x) + II(t,x).
\end{aligned}$$

We will treat $I(t,x) := I$ and $II(t,x) := II$ separately. To estimate I, we can rewrite it as

$$I(t,x) = \frac{1}{2}\int_\xi e^{ix\xi}(\xi - i0)^{-\lambda} \big((e^{-i\cdot\xi^4}\operatorname{sgn}(\cdot)) * \mathcal{I}_{-\lambda/4-3/4} f\big)(t)\,d\xi.$$

A direct calculation gives

$$\mathcal{F}_t\big((e^{-i\cdot\xi^4}\operatorname{sgn}(\cdot)) * \mathcal{I}_{-\lambda/4-3/4} f\big)(\tau) = \frac{(\tau - i0)^{\frac{1}{4}(3+\lambda)}\hat{f}(\tau)}{i(\tau+\xi^4)}.$$

Fubini's theorem and the dominated converge theorem imply that

$$I(t,x) = \int_\tau e^{it\tau} \lim_{\epsilon\to 0}\int_{|\tau+\xi^4|>\epsilon} \frac{e^{ix\xi}(\tau - i0)^{\frac{1}{4}(\lambda+3)}(\xi - i0)^{-\lambda}}{i(\tau+\xi^4)} \hat{f}(\tau)\,d\xi\,d\tau.$$

Thus, once we show that the function

$$g(\tau) := \lim_{\epsilon\to 0}\int_{|\tau+\xi^4|>\epsilon} \frac{e^{ix\xi}(\tau - i0)^{\frac{1}{4}(\lambda+3)}(\xi - i0)^{-\lambda}}{(\tau+\xi^4)}\,d\xi$$

is bounded independently of τ variable, the Plancherel's theorem enables us to obtain (4-17). The change of variable $\xi \mapsto |\tau|^{\frac{1}{4}}\xi$ and the fact that

$$(|\tau|^{\frac{1}{4}}\xi - i0)^{-\lambda} = |\tau|^{-\frac{\lambda}{4}}(\xi_+^{-\lambda} + e^{i\pi\lambda}\xi_-^{-\lambda})$$

gives

$$\begin{aligned} g(\tau) = \chi_{\{\tau>0\}} \int_\xi e^{ix|\tau|^{\frac{1}{4}}\xi} \, \frac{\xi_+^{-\lambda} + e^{i\pi\lambda}\xi_-^{-\lambda}}{1+\xi^4} \, d\xi \\ - e^{-\frac{1}{4}(i\pi(\lambda+3))} \chi_{\{\tau<0\}} \int_\xi e^{ix|\tau|^{\frac{1}{4}}\xi} \, \frac{\xi_+^{-\lambda} + e^{i\pi\lambda}\xi_-^{-\lambda}}{1-\xi^4} \, d\xi \\ := g_1 - e^{-\frac{1}{4}(i\pi(\lambda+3))} g_2. \end{aligned}$$

We only consider g_2, since g_1 is uniformly bounded in τ for $-3 < \lambda < 1$. Let us define the following cut-off function $\zeta \in C^\infty(\mathbb{R})$ such that

$$\zeta := \begin{cases} 1 & \text{in } \left[\frac{3}{4}, \frac{4}{3}\right], \\ 0 & \text{outside } \left(\frac{1}{2}, \frac{3}{2}\right). \end{cases}$$

Then, we obtain

$$\begin{aligned} g_2 &= \chi_{\{\tau<0\}} \int_\xi e^{ix|\tau|^{\frac{1}{4}}\xi} \zeta(\xi) \frac{\xi_+^{-\lambda}}{1-\xi^4} \, d\xi + \chi_{\{\tau<0\}} \int_\xi e^{ix|\tau|^{\frac{1}{4}}\xi} (1-\zeta(\xi)) \frac{\xi_+^{-\lambda} + e^{i\pi\lambda}\xi_-^{-\lambda}}{1-\xi^4} \, d\xi \\ &= g_{21} + g_{22}. \end{aligned}$$

It is clear that g_{22} is bounded independently of τ when $\lambda > -3$, and hence it remains to deal with g_{21}. Consider the functions

$$\widehat{\Theta}(\xi) = \frac{\zeta(\xi)\xi_+^{-\lambda}}{1+\xi+\xi^2+\xi^3} \quad \text{and} \quad \widehat{\Psi}(\xi) = \frac{1}{i(\xi-1)}.$$

We remark that $\widehat{\Theta}$ is a Schwartz function, and hence $\Theta \in \mathcal{S}(\mathbb{R})$. Moreover, we immediately know that

$$\Psi(x) = \frac{1}{2} e^{ix} \operatorname{sgn}(x),$$

since $\mathcal{F}_x[\operatorname{sgn}(x)](\xi) = \text{v.p.}\frac{2}{i\xi}$. Then, g_{21} can be written as

$$g_{21}(\tau) = -i\chi_{\{\tau<0\}} \int_\xi e^{ix|\tau|^{\frac{1}{4}}\xi} \widehat{\Theta}(\xi)\widehat{\Psi}(\xi) \, d\xi = -2i\pi \chi_{\{\tau<0\}} (\Theta * \Psi)(|\tau|^{\frac{1}{4}} x),$$

which implies

$$|g_{21}(\tau)| \lesssim \left| \int \Theta(y)\Psi(|\tau|^{\frac{1}{4}} x - y) \, dy \right| \lesssim \int |\Theta(y)| \, dy \lesssim_\zeta 1.$$

Now, we bound II. By using the definition of Fourier transform and (2-5) we have, after the change of variable $\eta=\xi^4$ and contour, that

$$II(t,x)=\frac{1}{2}\int_\xi e^{ix\xi}e^{-it\xi^4}(\xi^4-i0)^{\frac{1}{4}(\lambda+3)}\hat{f}(\xi^4)(\xi-i0)^{-\lambda}\,d\xi$$
$$=\frac{1}{2}\int_0^{+\infty} e^{it\eta}e^{-ix\eta^{\frac{1}{4}}}(\eta-i0)^{\frac{1}{4}(\lambda+3)}(\eta^{\frac{1}{4}}-i0)^{-\lambda}\eta^{-\frac{3}{4}}\hat{f}(\eta)\,d\eta=cf(t),$$

for some $c\in\mathbb{C}$, implying $\|II(\cdot,x)\|_{L^2_t}\lesssim\|f\|_{L^2_t}$. This completes the proof of (4-15).

Lastly, let us show (4-16). A direct calculation ensures that

$$\mathcal{F}_x(\psi(t)\mathcal{L}^\lambda f)(t,\xi)$$
$$=Me^{-\frac{1}{2}(i\pi\lambda)}e^{\frac{1}{10}(i\pi(\lambda+4))}(\xi-i0)^{-\lambda}\psi(t)e^{-it\xi^4}\int\frac{e^{it(\tau'+\xi^4)}-1}{i(\tau'+\xi^4)}(\tau'-i0)^{\frac{\lambda}{4}+\frac{3}{4}}\hat{f}(\tau')\,d\tau',$$

which can be divided into the following quantities:

$$\hat{f}_1(t,\xi)=Me^{-\frac{1}{2}i\pi\lambda}e^{\frac{1}{10}i\pi(\lambda+4)}(\xi-i0)^{-\lambda}\psi(t)$$
$$\times\int\frac{e^{it\tau'}-e^{-it\xi^4}}{i(\tau'+\xi^4)}\theta(\tau'+\xi^4)(\tau'-i0)^{\frac{\lambda}{4}+\frac{3}{4}}\hat{f}(\tau')\,d\tau',$$
$$\hat{f}_2(t,\xi)=Me^{-\frac{1}{2}i\pi\lambda}e^{\frac{1}{10}i\pi(\lambda+4)}(\xi-i0)^{-\lambda}\psi(t)$$
$$\times\int\frac{e^{it\tau'}}{i(\tau'+\xi^4)}(1-\theta(\tau'+\xi^4))(\tau'-i0)^{\frac{\lambda}{4}+\frac{3}{4}}\hat{f}(\tau')\,d\tau'$$
$$\hat{f}_3(t,\xi)=Me^{-\frac{1}{2}i\pi\lambda}e^{\frac{1}{10}i\pi(\lambda+4)}(\xi-i0)^{-\lambda}\psi(t)$$
$$\times\int\frac{e^{-it\xi^4}}{i(\tau'+\xi^4)}(1-\theta(\tau'+\xi^4))(\tau'-i0)^{\frac{\lambda}{4}+\frac{3}{4}}\hat{f}(\tau')\,d\tau'.$$

Here $\theta\in\mathcal{S}(\mathbb{R})$ is defined by

$$\theta(\tau):=\begin{cases}1 & \text{for } |\tau|\le 1,\\ 0 & \text{for } |\tau|\ge 2.\end{cases}\tag{4-18}$$

It follows that $\psi(t)\mathcal{L}^\lambda f=f_1+f_2-f_3$.

For f_1, we use the same argument as was done for w_1, in the proof of inequality (4-6). By the Taylor series expansion for $e^{it(\tau'+\xi^4)}$ at $it(\tau'+\xi^4)=0$, we write

$$\psi(t)\mathcal{L}^\lambda f_1(t,x)=c\sum_{k=1}^{\infty}\frac{i^{k-1}}{k!}\psi^k(t)e^{it\partial_x^4}F_1^k(x),$$

for some constant $c\in\mathbb{C}$, where $\psi^k(t)=t^k\psi(t)$ and

$$\widehat{F_1^k}(\xi)=(\xi-i0)^{-\lambda}\int\theta(\tau'+\xi^4)(\tau'+\xi^4)^{k-1}\tau'^{\frac{\lambda}{4}+\frac{3}{4}}\hat{f}(\tau')\,d\tau'.$$

By using (2-5), (4-6) and (4-18), it is enough to show that

$$\text{(4-19)}\quad \int_\xi \langle\xi\rangle^{2s}|\xi|^{-2\lambda}\left|\int_{|\tau'+\xi^4|\le 1}|\tau'+\xi^4|^{k-1}|\tau'|^{\frac14(\lambda+3)}|\hat f(\tau')|\,d\tau'\right|^2 d\xi \lesssim \|f\|^2_{H_0^{(2s+3)/8}}.$$

Let us split $|\xi|$ into two regions: $|\xi|\le 1$ and $|\xi|>1$. For the region $|\xi|\le 1$ and $|\tau'|\lesssim 1$ ($|\xi|\le 1$ and $|\tau'+\xi^4|\le 1$ imply $|\tau'|\lesssim 1$) we have that both $|\xi|^{-2\lambda}$ and $|\tau'|^{\frac12(\lambda+3)}$ are integrable, for $-5<\lambda<\frac12$, respectively. So, we obtain (4-19) by using the Cauchy–Schwarz inequality in τ'.

Assume that $|\xi|>1$, which in addition with $|\tau'+\xi^4|\le 1$ implies $|\tau'|\sim|\xi|^4>1$. Let $\hat f^*(\tau')=\langle\tau'\rangle^{\frac18(2s+3)}\hat f(\tau')$. Then the change of variable $\xi^4\mapsto\eta$ gives that the left-hand side of (4-19) is bounded by

$$\int_{|\xi|>1}|\xi|^3|\mathcal{M}\hat f^*(\xi^4)|^2\,d\xi \lesssim \int_{|\eta|>1}|\mathcal{M}\hat f^*(\eta)|^2\,d\eta \lesssim \|f^*\|^2_{L^2}=\|f\|^2_{H_0^{(2s+3)/8}},$$

where $\mathcal{M}\hat f^*$ is the Hardy–Littlewood maximal function of $\hat f^*$, and f_1 is controlled.

For f_2, from (2-5), the definition of inverse Fourier transform and Lemma 2.5, it follows that

$$\begin{aligned}\|f_2\|^2_{X^{s,b}} &\lesssim \int\int \langle\xi\rangle^{2s}|\xi|^{-2\lambda}\langle\tau+\xi^4\rangle^{2b}\frac{(1-\theta(\tau+\xi^4))^2}{|\tau+\xi^4|^2}|\tau|^{\frac12(\lambda+3)}|\hat f(\tau)|^2\,d\tau d\xi\\ &\lesssim \int |\tau|^{\frac12(\lambda+3)}\left(\int\frac{\langle\xi\rangle^{2s}|\xi|^{-2\lambda}}{\langle\tau+\xi^4\rangle^{2-2b}}\,d\xi\right)|\hat f(\tau)|^2\,d\tau.\end{aligned}$$

Thus, by the change of variable $\eta=\xi^4$ and Lemma 2.2, for $-5<\lambda$ (we may assume $\operatorname{supp} f\subset[0,1]$, thanks to Lemma 2.4), it suffices to show

$$I(\tau)=\int\frac{|\eta|^{-\frac34-\frac{\lambda}{2}}\langle\eta\rangle^{\frac{s}{2}}}{\langle\tau+\eta\rangle^{2-2b}}\,d\eta\lesssim\langle\tau\rangle^{\frac{2s}{4}-\frac{2\lambda}{4}-\frac34}.$$

Here, we split $|\tau|$ into two regions: $|\tau|\le 2$ and $|\tau|>2$. When $|\tau|\le 2$, we have $\langle\tau+\eta\rangle\sim\langle\eta\rangle$. For $s<4-4b$ and $s+4b-\frac72<\lambda<\frac12$, we get

$$I(\tau)\lesssim\int_{|\eta|\le 1}|\eta|^{\frac14(-3-2\lambda)}+\int\frac{d\eta}{\langle\eta\rangle^{2-2b-\frac{s}{2}+\frac34+\frac{\lambda}{2}}}\lesssim 1.$$

Now, working in the region $|\tau|>2$, we divide the integral region in η into $|\eta|<\frac{|\tau|}{2}$ and $|\eta|\ge\frac{|\tau|}{2}$. In the first region, for $b<\frac12$ and $\lambda<\min\bigl(\frac12,s+\frac12\bigr)$, we bound in the following way:

$$\langle\tau\rangle^{2b-2}\left(\int_{|\eta|\le 1}|\eta|^{-\frac34-\frac{2\lambda}{4}}\,d\eta+\int_{1<|\eta|\le\frac{|\tau|}{2}}|\eta|^{\frac14(-3-2\lambda+2s)}d\eta\right)\lesssim\langle\tau\rangle^{\frac14(-3-2\lambda+2s)}.$$

On the other hand, in the second region, we have that $|\tau+\eta| \geq \frac{1}{2}|\tau| > 1$. Then, for $s-2<\lambda$ and $b<\frac{1}{2}$, it holds that

$$I(\tau) \lesssim \langle\tau\rangle^{\frac{1}{4}(-3-2\lambda+2s)} \int \frac{d\eta}{\langle\tau+\eta\rangle^{2-2b}} \lesssim \langle\tau\rangle^{\frac{1}{4}(-3-2\lambda+2s)} \int_{|s|>1} \frac{ds}{|s|^{2-2b}} \lesssim \langle\tau\rangle^{\frac{1}{4}(-3-2\lambda+2s)},$$

so

$$\|f_2\|_{X^{s,b}} \lesssim \|f\|_{H_0^{(2s+3)/8}}.$$

This completes the estimate for f_2.

Finally, let us show that f_3 can be controlled. Similarly as for f_1, it suffices to show

$$\int \langle\xi\rangle^{2s}|\xi|^{-2\lambda}\left|\int (1-\theta(\tau'+\xi^4))|\tau'+\xi^4|^{-1}|\tau'|^{\frac{1}{4}(\lambda+3)}|\hat{f}(\tau')|\,d\tau'\right|^2 d\xi \lesssim \|f\|^2_{H_0^{(2s+3)/8}}. \tag{4-20}$$

Again, we split the region $|\xi|$ as follows: $|\xi| \leq 1$ and $|\xi| > 1$. Considering $|\xi| \leq 1$, since $|\xi|^{-2\lambda}$ is integrable, for $\lambda < \frac{1}{2}$, and we may ignore the integration in ξ. Let us work in the region $|\tau'| \leq 1$. In this region $|\tau'|^{\frac{1}{2}(\lambda+3)}$ is integrable, for $\lambda > -5$, and hence we get (4-20).

On the region $|\tau'| > 1$, since $|\tau'+\xi^4| \sim |\tau'|$ and $|\tau'|^{\frac{1}{5}(-s+\lambda-3)}$ are L^2 integrable, for $\lambda < s+\frac{1}{2}$, we also get (4-20) by using the Cauchy–Schwarz inequality in τ'. Still looking on the region $|\tau'| > 1$, since $|\tau'+\xi^4| \sim |\tau'|$ we have that the left-hand side of (4-20) is bounded by

$$c\left(\int_{|\tau'|>1} |\tau'|^{\frac{\lambda-1}{4}}|\hat{f}|^2 d\tau'\right)^2 \sim \left(\int_{|\tau'|>1} \frac{|\tau'|^{\frac{\lambda-1}{4}}}{\langle\tau'\rangle^{\frac{2s+3}{8}}}\langle\tau'\rangle^{\frac{2s+3}{8}}|\hat{f}|^2 d\tau'\right)^2 \lesssim \int \frac{\langle\tau'\rangle^{\frac{\lambda-1}{2}}}{\langle\tau'\rangle^{\frac{2s+3}{8}}} d\tau' \|f\|^2_{H^{(2s+3)/4}} \lesssim \|f\|^2_{H_0^{(2s+3)/8}}, \tag{4-21}$$

where we have used that $\lambda < s+\frac{1}{2}$, and the result follows on $|\xi| \leq 1$. On the other hand, in the region $|\xi| > 1$ and $|\tau'| \leq 1$, since $|\tau'+\xi^4| \sim |\xi|^4 \sim \langle\xi\rangle^4$ and $\langle\xi\rangle^{2s-2\lambda-8}$ are integrable for $\lambda > -\frac{7}{2}+s$, we also get (4-20).

Consider the region $|\xi| > 1$ and $|\tau'| > 1$. There are two possibilities:

(I) $|\tau'| \leq \frac{1}{2}|\xi|^4$.

(II) $\frac{1}{2}|\xi|^4 < |\tau'|$.

In view of the proof of [Holmer 2006, Lemma 5.8(d)] (see also [Holmer 2005]), one can replace

$$\frac{1-\theta(\tau'+\xi^4)}{\tau'+\xi^4}$$

by $\beta(\tau'+\xi^4)$ for some $\beta\in\mathcal{S}(\mathbb{R})$. Hence, the left-hand side of (4-20) is dominated by

$$c\int_{|\xi|>1}|\xi|^{2s-2\lambda}\left|\int_{|\tau'|>1}|\tau'+\xi^4|^{-N}|\tau'|^{\frac{1}{4}(\lambda+3)}|\hat{f}(\tau')|\,d\tau'\right|^2 d\xi, \tag{4-22}$$

for $N\geq 0$. By the Cauchy–Schwarz inequality and choosing $N=N(s,\lambda)\gg 1$, we have (4-20) for both cases. Indeed, for the case I (in this case $|\tau'+\xi^4|\sim|\xi|^4$), (4-22) can be controlled by

$$c\|f\|^2_{H^{\frac{1}{8}(2s+3)}}\int_{|\xi|>1}|\xi|^{2s-2\lambda-4N}\int_{1<|\tau'|\leq\frac{1}{2}|\xi|^4}|\tau'|^{\frac{1}{4}(-2s-3+2\lambda+6-4N)}d\tau' d\xi\lesssim\|f\|^2_{H_0^{\frac{1}{8}(2s+3)}}.$$

For case II (in this case $|\tau'+\xi^4|\sim|\tau'|$), (4-22) is bounded by

$$c\int_{|\xi|>1}|\xi|^{2s-2\lambda}\left|\int_{\frac{1}{2}|\xi|^4<|\tau'|}|\tau'|^{\frac{1}{4}(2\lambda+6-2s-3-4N)}|\tau'|^{\frac{1}{8}(2s+3)}|\hat{f}(\tau')|\,d\tau'\right|^2 d\xi \lesssim\|f\|^2_{H_0^{\frac{1}{8}(2s+3)}},$$

thus

$$\|f_3\|_{X^{s,b}}\lesssim\|f\|_{H_0^{\frac{1}{8}(2s+3)}},$$

finishing the estimate for f_3.

Remembering that $\psi(t)\mathcal{L}^\lambda f=f_1+f_2-f_3$, and using the estimates of f_i, $i=1,2,3$, equation (4-16) follows and the proof is complete. □

To close this section, let us enunciate the trilinear estimates associated to the fourth order nonlinear Schrödinger equation (1-2). The proof of this estimate can be found in [Oh and Tzvetkov 2017] (see also [Capistrano-Filho and Cavalcante 2019]), thus we will omit it.

Proposition 4.4. *For $s\geq 0$, there exists $b=b(s)<\frac{1}{2}$ such that we have*

$$\|u_1u_2\bar{u}_3\|_{X^{s,-b}}\leq c\|u_1\|_{X^{s,b}}\|u_2\|_{X^{s,b}}\|u_3\|_{X^{s,b}}. \tag{4-23}$$

5. Proof of Theorem 1.1

Initially, we pick an extension $\tilde{u}_0\in H^s(\mathbb{R})$ of u_0 such that

$$\|\tilde{u}_0\|_{H^s(\mathbb{R})}\leq 2\|u_0\|_{H^s(\mathbb{R}^+)}.$$

Let $b=b(s)<\frac{1}{2}$ such that the estimates given in Proposition 4.4 are valid.

By using similar arguments to those in Section 3B, let

$$u(t,x)=\mathcal{L}^{\lambda_1}\gamma_1(t,x)+\mathcal{L}^{\lambda_2}\gamma_2(t,x)+F(t,x), \tag{5-1}$$

where γ_i $(i=1,2)$ will be chosen in terms of initial and boundary data u_0, f, g and $F(t,x)=e^{it\partial_x^4}\tilde{u}_0+\lambda\mathcal{D}(|u|^2u)$.

Remember that a_j and b_j are defined by

$$a_j = \frac{M}{8}\left(\frac{e^{-i\frac{\pi}{8}(1+3\lambda_j)} + e^{-i\frac{\pi}{8}(1-5\lambda_j)}}{\sin\left(\frac{1}{4}(1-\lambda_j)\pi\right)}\right),$$
(5-2)
$$b_j = \frac{M}{8}\left(\frac{e^{-i\frac{\pi}{8}(-2+3\lambda_j)} + e^{-i\frac{\pi}{8}(6-5\lambda_j)}}{\sin\left(\frac{1}{4}(2-\lambda_j)\pi\right)}\right).$$

By Lemmas 3.2 and 3.3, we get

(5-3) $$f(t) = u(t,0) = a_1\gamma_1(t) + a_2\gamma_2(t) + F(t,0)$$

(5-4) $$g(t) = \partial_x u(t,0) = b_1\mathcal{I}_{-1/4}\,\gamma_1(t) + b_2\mathcal{I}_{-1/4}\,\gamma_2(t) + \partial_x F(t,0).$$

Putting together (5-3) and (5-4), we can write a matrix in the form

$$\begin{bmatrix} f(t) - F(t,0) \\ \mathcal{I}_{1/4}\,g(t) - \mathcal{I}_{1/4}\,\partial_x F(t,0)\end{bmatrix} = A\begin{bmatrix}\gamma_1(t)\\ \gamma_2(t)\end{bmatrix},$$

where

$$A(\lambda_1,\lambda_2) = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2\end{bmatrix}.$$

By using a mathematical software, the determinant of matrix $A(\lambda_1,\lambda_2)$ is given by

$$\det A = 2(-1)^{\frac{15}{8}} e^{-\frac{1}{8}i(6+3\lambda_1+\lambda_2)\pi}(1+e^{i\lambda_1\pi})(-1+e^{\frac{1}{2}i\lambda_2\pi})\sec\left(\frac{(1+\lambda_1)\pi}{4}\right) + 4(-1)^{\frac{3}{8}} e^{-\frac{1}{8}i(\lambda_1+\lambda_2)\pi}(-1+e^{\frac{1}{2}i\lambda_1\pi})(1-ie^{\frac{1}{2}i\lambda_2\pi}).$$

Note that the following graphics, with real and imaginary parts, of the determinant function $A(\lambda_1,\lambda_2)$, help us to see when the matrix A is invertible.

Thus, matrix $A(\lambda_1,\lambda_2)$ is invertible if we get

(5-5)
$$\lambda_2 \neq \frac{2}{\pi}\left(2\pi n - i\log\left\{\frac{-2(-1)^{\frac{1}{4}}e^{\frac{i\pi\lambda_1}{4}} + 2(-1)^{\frac{1}{4}}e^{\frac{3i\pi\lambda_1}{4}} + \left(e^{i\pi\lambda_1}+1\right)\sec\left(\frac{(1+\lambda_1)\pi}{4}\right)}{-2(-1)^{\frac{3}{4}}e^{\frac{i\pi\lambda_1}{4}} + 2(-1)^{\frac{3}{4}}e^{\frac{3i\pi\lambda_1}{4}} + \left(e^{i\pi\lambda_1}+1\right)\sec\left(\frac{(1+\lambda_1)\pi}{4}\right)}\right\}\right)$$

and

(5-6) $$\lambda_j \neq 1-4n, \quad \lambda_j \neq 2-4n, \quad j = 1,2,$$

for all $n \in \mathbb{Z}$.

Figure 1 helps us to see that there are an infinite set of parameters which satisfy the relations (5-5) and (5-6). In fact, for example, pick $\lambda_1 \approx 0$ and $\lambda_2 \approx \frac{1}{3}$. Thus, for $0 \le s < \frac{1}{2}$, the choice of parameters λ_1 and λ_2 satisfying the conditions

$$-3 < \lambda_j < \tfrac{1}{2}, \quad s + 4b - 2 < \lambda_j < s + \tfrac{1}{2}, \quad j = 1,2,$$

ensures that Lemma 4.3 holds. Thus, for fixed $s \in \left[0, \frac{1}{2}\right)$, we can choose λ_1 and λ_2 as before and define the forcing functions $\gamma_1(t)$ and $\gamma_2(t)$ for any λ_j, $j = 1, 2$, given by

$$\begin{bmatrix}\gamma_1(t)\\ \gamma_2(t)\end{bmatrix} = A^{-1}\begin{bmatrix} f(t) - F(t,0) \\ \mathcal{I}_{1/4}\, g(t) - \mathcal{I}_{1/4}\, \partial_x F(t,0)\end{bmatrix}, \tag{5-7}$$

which shows that formula (5-1) restricted on the set $(0, +\infty) \times (0, +\infty)$ satisfies

$$(i\partial_t - \partial_x^4)u = \lambda |u|^2 u,$$

in the sense of distributions.

Thus, we define the solution operator by

$$\Lambda u(t,x) = \psi(t)\mathcal{L}^{\lambda_1}\gamma_1(t,x) + \psi(t)\mathcal{L}^{\lambda_2}\gamma_2(t,x) + \psi(t)F(t,x), \tag{5-8}$$

where

$$\begin{bmatrix}\gamma_1(t)\\ \gamma_2(t)\end{bmatrix} = A^{-1}\begin{bmatrix} f(t) - F(t,0) \\ \mathcal{I}_{1/4}\, g(t) - \mathcal{I}_{1/4}\, \partial_x F(t,0)\end{bmatrix},$$

$F(t,x) = e^{it\partial_x^4}\tilde{u}_0 + \lambda \mathcal{D}(\psi_T |u|^2 u)$ and ψ is defined by (2-1).

Recall the solution space $Z^{s,b}$, defined in Section 2B, under the norm

$$\|v\|_{Z^{s,b}} = \sup_{t\in\mathbb{R}} \|v(t,\cdot)\|_{H^s} + \sum_{j=0}^{1} \sup_{x\in\mathbb{R}} \|\partial_x^j v(\cdot,x)\|_{H^{\frac{1}{8}(2s+3-2j)}} + \|v\|_{X^{s,b}}.$$

The estimates obtained in Section 2 together with estimates of Section 4 and (4-23) yield that

$$\|\Lambda u\|_{Z^{s,b}} \le c\big(\|u_0\|_{H^s(\mathbb{R}^+)} + \|f\|_{H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+)} + \|g\|_{H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+)}\big) + C_1 T^{\epsilon} \|u\|^3_{Z^{s,b}},$$

for ϵ adequately small. Similarly,

$$\|\Lambda u_1 - \Lambda u_2\|_{Z^{s,b}} \le C_2 T^{\epsilon} \big(\|u_1\|^2_{Z^{s,b}} + \|u_2\|^2_{Z^{s,b}}\big)\|u_1 - u_2\|_{Z^{s,b}},$$

for $u_1(0,x) = u_2(0,x)$.

Consider in Z the ball defined by $B = \{u \in Z^{s,b};\ \|u\|_{Z^{s,b}} \le M\}$, where

$$M = 2c\Big(\|u_0\|_{H^s(\mathbb{R}^+)} + \|f\|_{H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+)} + \|g\|_{H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+)}\Big).$$

Lastly, choosing $T = T(M)$ sufficiently small, such that

$$\|\Lambda u\|_{Z^{s,b}} \le M \quad \text{and} \quad \|\Lambda u_1 - \Lambda u_2\|_{Z^{s,b}} \le \tfrac{1}{2}\|u_1 - u_2\|_{Z^{s,b}},$$

it follows that Λ is a contraction map on B, finishing the proof of Theorem 1.1. □

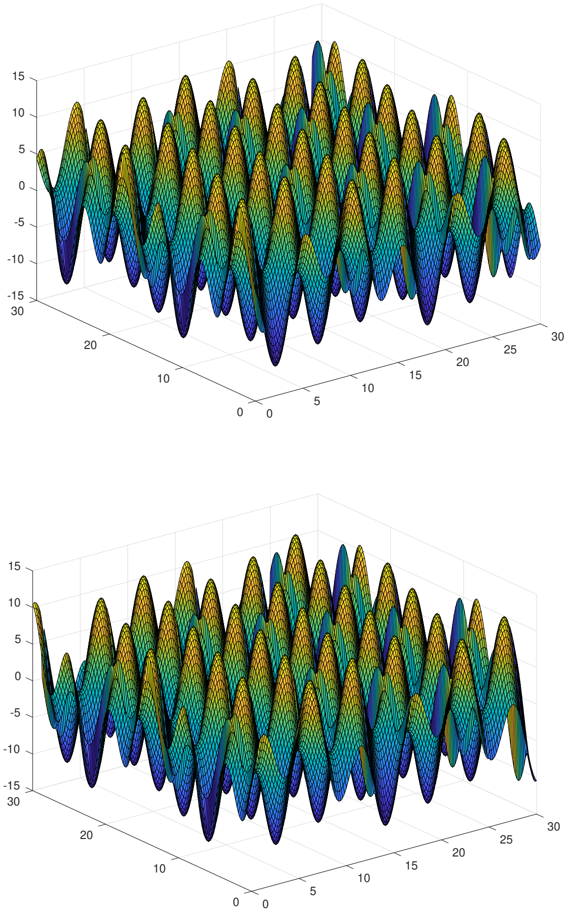


Figure 1. Real (top) and imaginary (bottom) parts of $\det A$.

Remarks. Concerning our main result, Theorem 1.1, the following remarks are now in order:

1. An important point to treat in dispersive systems is the analysis of the scaling. For our case, that is, for the biharmonic Schrödinger equation on the half-line, we have the following: if $u(t,x)$ is solution for IBVP (1-2) on $[0,T)\times(0,\infty)$, then, for $\lambda>0$, the function $u_\lambda(t,x)=\lambda^2 u(\lambda^4 t,\lambda x)$ is solution for (1-2) on $[0,T/\lambda^4)\times(0,\infty)$ with initial-boundary conditions $u_\lambda(0,x)=\lambda^2 u_0(\lambda x):=u_{0\lambda}$, $u_\lambda(t,0)=\lambda^2 f(\lambda^4 t):=f_\lambda$ and $u_{x,\lambda}(t,0)=\lambda^3 g(\lambda^4 t):=g_\lambda$. A straightforward calculation gives

$$\begin{aligned}(5\text{-}9)\quad &\|u_{0\lambda}\|_{H^s(\mathbb{R}^+)}+\|f_\lambda\|_{H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+)}+\|g_\lambda\|_{H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+)}\\ &\lesssim \lambda^{\frac{3}{2}}\langle\lambda\rangle^s\|u_0\|_{H^s(\mathbb{R}^+)}+\langle\lambda\rangle^{\frac{1}{2}(2s+3)}\|f\|_{H^{\frac{1}{8}(2s+3)}(\mathbb{R}^+)}+\lambda\langle\lambda\rangle^{\frac{1}{8}(2s+1)}\|g\|_{H^{\frac{1}{8}(2s+1)}(\mathbb{R}^+)}.\end{aligned}$$

2. In order to make the norms of our initial data u_0, f and g small, we rescale the data u_0 and g by choosing λ adequately small, by using (5-9). However, we can not rescale the function f since a positive power of λ does not appear in (5-9). To overcome this difficulty in our context, we introduce the cut-off function ψ_T, defined by (2-1), in the operator Λ (see (5-8)) to prove that Λ is thus a contraction, proving the main result of the article.

3. It is important to note that the scaling argument was successful in the cases of the quadratic NLS equation [Cavalcante 2017], KdV equation [Holmer 2006] and Kawahara equation [Cavalcante and Kwak 2019] posed on the half-line.

4. Finally, in view of (5-3), (5-4) and (5-7), it is necessary to check $\gamma_i(t)$, $i = 1, 2$ to be well-defined in $H_0^{(2s+3)/8}(\mathbb{R}^+)$. However, it follows from Lemmas 4.1, 4.2 and 4.3, Propositions 4.4 and Lemmas 2.1 and 2.3.

6. Further comments and open problems

In this section, our plan is to present four problems that can be treated with the approach used in this article.

6A. *Biharmonic NLS on star graphs.* The authors [Capistrano-Filho et al. 2019] considered the biharmonic Schrödinger equation on star graphs, given by N edges $(0, \infty)$ connected with a common vertex $(0, 0, \dots, 0)$ (see Figure 2), namely

$$(6\text{-}1)\quad \begin{cases} i\partial_t u_j - \partial_x^4 u_j + \lambda |u_j|^2 u_j = 0, & (t,x) \in (0,T) \times (0,\infty),\ \ j = 1, 2, \dots, N, \\ u_j(0,x) = u_{j0}(x), & x \in (0,\infty), \end{cases}$$

with initial conditions $(u_1(0,x), u_2(0,x), \dots, u_N(0,x)) \in H^s(\mathbb{R}^+)$.

For a better understanding, we are interested in solving (6-1) with the following three classes boundary conditions:

$$(6\text{-}2)\quad \begin{cases} \partial_x^k u_1(t,0) = \partial_x^k u_2(t,0) = \cdots = u_N(t,0), & k = 0, 1,\ \ t \in (0,T), \\ \sum_{j=1}^n \partial_x^k u_j(t,0) = 0, & k = 2, 3,\ \ t \in (0,T); \end{cases}$$

$$(6\text{-}3)\quad \begin{cases} \partial_x^k u_1(t,0) = \partial_x^k u_2(t,0) = \cdots = u_N(t,0), & k = 2, 3,\ \ t \in (0,T), \\ \sum_{j=1}^n \partial_x^k u_j(t,0) = 0, & k = 0, 1,\ \ t \in (0,T); \end{cases}$$

$$(6\text{-}4)\quad \begin{cases} \partial_x^k u_1(t,0) = \partial_x^k u_2(t,0) = \cdots = u_N(t,0), & k = 0, 3,\ \ t \in (0,T), \\ \sum_{j=1}^n \partial_x^k u_j(t,0) = 0, & k = 1, 2,\ \ t \in (0,T). \end{cases}$$

The motivation of these boundary conditions and how we can choose it, follows the ideas contained in [Cavalcante 2018], and are detailed in [Capistrano-Filho et al. 2019].

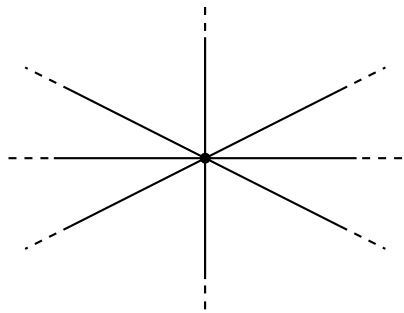


Figure 2. Star graphs connected with a common vertex $(0, 0, \ldots, 0)$.

6B. *Control theory.* We split this section in two parts: control theory for the biharmonic NLS on star graphs and on an unbounded domain, respectively.

6B1. *Control theory of biharmonic NLS on star graphs.* First, let us consider the controllability problem associated to (6-1) with three possibilities of boundary conditions, namely, (6-2), (6-3) and (6-4). Due of the recent development of graph theory for the Korteweg–de Vries equation, in the following paragraph we present a few comments about this study.

In three interesting papers Ammari and Crépeau [2018], Cavalcante [2018] and Mugnolo et al. [2018] dealt with the study of the KdV and Airy equations in graphs. In summary, in the first work, the authors proposed a model using the Korteweg–de Vries equation on a finite star-shaped network and proved the well-posedness of the system. Also, as the main result of the work, by using properties of the energy, they showed that the solutions of the system decays exponentially to zero (as $t \to \infty$) and they studied an exact boundary controllability problem. In the second work, Cavalcante showed local well-posedness for the Cauchy problem associated with Korteweg–de Vries equation on a metric star graph. More precisely, he used the Duhamel boundary forcing operator, in the context of half-line, introduced by Colliander and Kenig [2002] and Holmer [2006] to achieve his result. Finally, Mugnolo et al. obtained a characterization of all boundary conditions under which the Airy-type evolution equation $u_t = \alpha u_{xxx} + \beta u_x$, for $\alpha \in \mathbb{R} \setminus \{0\}$ and $\beta \in \mathbb{R}$ on star graphs, generates contraction semigroups.

In this spirit, looking for the energy identity of the system (6-1), namely the L^2-energy, which satisfies an equality given by

$$\begin{aligned} \text{(6-5)}\quad & E(u_1(T,x), u_2(T,x), \ldots, u_N(T,x)) \\ &= -\sum_{i=1}^{N} \int_0^T \operatorname{Im}(\partial_x^3 u_j(t,0)\bar{u}_j(t,0))\,dt + \sum_{i=1}^{N} \int_0^T \operatorname{Im}(\partial_x^2 u_j(t,0)\partial_x \bar{u}_j(t,0))\,dt \\ &\qquad - E(u_1(0,x), u_2(0,x), \ldots, u_N(0,x)), \end{aligned}$$

where

$$E(u_1(t,x), u_2(t,x), \dots, u_N(t,x)) := \sum_{i=1}^{N} \int_0^{+\infty} |u(t,x)|^2 \, dx,$$

the following natural questions arise.

Problem A: Which are the boundary conditions that we can impose in (6-2), (6-3) and (6-4) such that the energy is a nonincreasing function of the time variable t?

Problem B: If we can impose some boundary conditions such that the energy (6-5) is a nonincreasing function of the time variable t, is the system (6-1), with appropriate boundary conditions, asymptotically stable when the time tends to infinity?

Problem C: Can we find appropriate boundary controls such that the system (6-1) is controllable in some sense?

6B2. *Control theory of biharmonic NLS in unbounded domain.* In the context of control in unbounded domain Faminskii [2019] considered the initial-boundary value problem posed on infinite domain for the Korteweg–de Vries equation. Precisely, he elected a function f_0 on the right-hand side of the equation as an unknown function, regarded as a control. Thus he proved that this function must be chosen such that the corresponding solution should satisfy certain additional integral condition.

These techniques probably work well for the following biharmonic NLS system:

$$\text{(6-6)} \qquad \begin{cases} i\partial_t u + \gamma \partial_x^4 u + \lambda |u|^2 u = f_0(t) v(x,t), & (t,x) \in (0,T) \times (0,\infty), \\ u(0,x) = u_0(x), & x \in (0,\infty), \\ u(t,0) = h(t), \;\; u_x(t,0) = g(t) & t \in (0,T), \end{cases}$$

for $\gamma, \lambda \in \mathbb{R}$, where v is a given function and f_0 is an unknown control function. Therefore, the issue here is:

Problem D: Is (6-6) controllable in the sense of Faminskii's work? Namely, can we find a pair $\{f_0, u\}$, satisfying appropriate additional integral conditions (for details see [Faminskii 2019])?

Acknowledgments

The authors wish to thank the referee for valuable comments which improved this paper. Capistrano–Filho was supported by CNPq (Brazil) grants 306475/2017-0, 408181/2018-4, CAPES-PRINT (Brazil) grant 88881.311964/2018-01 and Qualis A - Propesq (UFPE). Gallego was partially supported under projects SIGP 58907 and 45511. This work was carried out during some visits of the authors to the Federal University of Pernambuco and Universidad Nacional de Colombia - Sede Manizales. The authors would like to thank both Universities for their hospitality.

References

[Ammari and Crépeau 2018] K. Ammari and E. Crépeau, "Feedback stabilization and boundary controllability of the Korteweg–de Vries equation on a star-shaped network", *SIAM J. Control Optim.* **56**:3 (2018), 1620–1639. MR Zbl

[Ben-Artzi et al. 2000] M. Ben-Artzi, H. Koch, and J.-C. Saut, "Dispersion estimates for fourth order Schrödinger equations", *C. R. Acad. Sci. Paris Sér. I Math.* **330**:2 (2000), 87–92. MR Zbl

[Bona et al. 2006] J. L. Bona, S. M. Sun, and B.-Y. Zhang, "Boundary smoothing properties of the Korteweg-de Vries equation in a quarter plane and applications", *Dyn. Partial Differ. Equ.* **3**:1 (2006), 1–69. MR

[Bona et al. 2018] J. L. Bona, S.-M. Sun, and B.-Y. Zhang, "Nonhomogeneous boundary-value problems for one-dimensional nonlinear Schrödinger equations", *J. Math. Pures Appl.* (9) **109** (2018), 1–66. MR Zbl

[Bourgain 1993a] J. Bourgain, "Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, I: Schrödinger equations", *Geom. Funct. Anal.* **3**:2 (1993), 107–156. MR

[Bourgain 1993b] J. Bourgain, "Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, II: The KDV-equation", *Geom. Funct. Anal.* **3**:3 (1993), 209–262. MR Zbl

[Burq et al. 2002] N. Burq, P. Gérard, and N. Tzvetkov, "An instability property of the nonlinear Schrödinger equation on S^d", *Math. Res. Lett.* **9**:2-3 (2002), 323–335. MR Zbl

[Burq et al. 2013] N. Burq, L. Thomann, and N. Tzvetkov, "Long time dynamics for the one dimensional non linear Schrödinger equation", *Ann. Inst. Fourier* (*Grenoble*) **63**:6 (2013), 2137–2198. MR Zbl

[Capistrano-Filho and Cavalcante 2019] R. A. Capistrano-Filho and M. Cavalcante, "Stabilization and control for the biharmonic Schrödinger equation", *Appl. Math. Optim.* (2019).

[Capistrano-Filho et al. 2019] R. A. Capistrano-Filho, M. Cavalcante, and F. A. Gallego, "Forcing operators on star graphs applied for the cubic fourth order Schrödinger equation", 2019. arXiv

[Cavalcante 2017] M. Cavalcante, "The initial boundary value problem for some quadratic nonlinear Schrödinger equations on the half-line", *Differential Integral Equations* **30**:7-8 (2017), 521–554. MR Zbl

[Cavalcante 2018] M. Cavalcante, "The Korteweg–de Vries equation on a metric star graph", *Z. Angew. Math. Phys.* **69**:5 (2018), Paper No. 124, 22. MR Zbl

[Cavalcante and Corcho 2019] M. Cavalcante and A. J. Corcho, "The initial-boundary value problem for the Schrödinger–Korteweg–de Vries system on the half-line", *Commun. Contemp. Math.* **21**:8 (2019), 1850066, 47. MR Zbl

[Cavalcante and Kwak 2019] M. Cavalcante and C. Kwak, "Local well-posedness of the fifth-order KdV-type equations on the half-line", *Commun. Pure Appl. Anal.* **18**:5 (2019), 2607–2661. MR

[Colliander and Kenig 2002] J. E. Colliander and C. E. Kenig, "The generalized Korteweg-de Vries equation on the half line", *Comm. Partial Differential Equations* **27**:11-12 (2002), 2187–2266. MR

[Compaan and Tzirakis 2017] E. Compaan and N. Tzirakis, "Well-posedness and nonlinear smoothing for the "good" Boussinesq equation on the half-line", *J. Differential Equations* **262**:12 (2017), 5824–5859. MR Zbl

[Cui and Guo 2007] S. Cui and C. Guo, "Well-posedness of higher-order nonlinear Schrödinger equations in Sobolev spaces $H^s(\mathbb{R}^n)$ and applications", *Nonlinear Anal.* **67**:3 (2007), 687–707. MR Zbl

[Erdoğan and Tzirakis 2017] M. B. Erdoğan and N. Tzirakis, "Regularity properties of the Zakharov system on the half line", *Comm. Partial Differential Equations* **42**:7 (2017), 1121–1149. MR Zbl

[Faminskii 2019] A. V. Faminskii, "Control problems with an integral condition for Korteweg–de Vries equation on unbounded domains", *J. Optim. Theory Appl.* **180**:1 (2019), 290–302. MR Zbl

[Fibich et al. 2002] G. Fibich, B. Ilan, and G. Papanicolaou, "Self-focusing with fourth-order dispersion", *SIAM J. Appl. Math.* **62**:4 (2002), 1437–1462. MR Zbl

[Fokas 1997] A. S. Fokas, "A unified transform method for solving linear and certain nonlinear PDEs", *Proc. Roy. Soc. London Ser. A* **453**:1962 (1997), 1411–1443. MR Zbl

[Fokas 2008] A. S. Fokas, *A unified approach to boundary value problems*, CBMS-NSF Regional Conference Series in Applied Mathematics **78**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. MR Zbl

[Fokas et al. 2016] A. S. Fokas, A. A. Himonas, and D. Mantzavinos, "The Korteweg–de Vries equation on the half-line", *Nonlinearity* **29**:2 (2016), 489–527. MR Zbl

[Ginibre et al. 1997] J. Ginibre, Y. Tsutsumi, and G. Velo, "On the Cauchy problem for the Zakharov system", *J. Funct. Anal.* **151**:2 (1997), 384–436. MR Zbl

[Hao et al. 2006] C. Hao, L. Hsiao, and B. Wang, "Wellposedness for the fourth order nonlinear Schrödinger equations", *J. Math. Anal. Appl.* **320**:1 (2006), 246–265. MR Zbl

[Hayashi and Naumkin 2015a] N. Hayashi and P. I. Naumkin, "Factorization technique for the fourth-order nonlinear Schrödinger equation", *Z. Angew. Math. Phys.* **66**:5 (2015), 2343–2377. MR Zbl

[Hayashi and Naumkin 2015b] N. Hayashi and P. I. Naumkin, "Global existence and asymptotic behavior of solutions to the fourth-order nonlinear Schrödinger equation in the critical case", *Nonlinear Anal.* **116** (2015), 112–131. MR Zbl

[Hayashi and Naumkin 2015c] N. Hayashi and P. I. Naumkin, "Large time asymptotics for the fourth-order nonlinear Schrödinger equation", *J. Differential Equations* **258**:3 (2015), 880–905. MR Zbl

[Hayashi and Naumkin 2015d] N. Hayashi and P. I. Naumkin, "On the inhomogeneous fourth-order nonlinear Schrödinger equation", *J. Math. Phys.* **56**:9 (2015), 093502, 25. MR Zbl

[Holmer 2005] J. Holmer, "The initial-boundary-value problem for the 1D nonlinear Schrödinger equation on the half-line", *Differential Integral Equations* **18**:6 (2005), 647–668. MR Zbl

[Holmer 2006] J. Holmer, "The initial-boundary value problem for the Korteweg-de Vries equation", *Comm. Partial Differential Equations* **31**:7-9 (2006), 1151–1190. MR Zbl

[Huo and Jia 2005] Z. Huo and Y. Jia, "The Cauchy problem for the fourth-order nonlinear Schrödinger equation related to the vortex filament", *J. Differential Equations* **214**:1 (2005), 1–35. MR Zbl

[Jerison and Kenig 1995] D. Jerison and C. E. Kenig, "The inhomogeneous Dirichlet problem in Lipschitz domains", *J. Funct. Anal.* **130**:1 (1995), 161–219. MR Zbl

[Karpman 1996] V. I. Karpman, "Stabilization of soliton instabilities by higher order dispersion: KdV-type equations", *Phys. Lett. A* **210**:1-2 (1996), 77–84. MR Zbl

[Karpman and Shagalov 2000] V. I. Karpman and A. G. Shagalov, "Stability of solitons described by nonlinear Schrödinger-type equations with higher-order dispersion", *Phys. D* **144**:1-2 (2000), 194–210. MR Zbl

[Kenig et al. 1991] C. E. Kenig, G. Ponce, and L. Vega, "Oscillatory integrals and regularity of dispersive equations", *Indiana Univ. Math. J.* **40**:1 (1991), 33–69. MR Zbl

[Kwak 2018] C. Kwak, "Periodic fourth-order cubic NLS: Local well-posedness and non-squeezing property", *J. Math. Anal. Appl.* **461**:2 (2018), 1327–1364. MR Zbl

[Mugnolo et al. 2018] D. Mugnolo, D. Noja, and C. Seifert, "Airy-type evolution equations on star graphs", *Anal. PDE* **11**:7 (2018), 1625–1652. MR Zbl

[Natali and Pastor 2015] F. Natali and A. Pastor, "The fourth-order dispersive nonlinear Schrödinger equation: orbital stability of a standing wave", *SIAM J. Appl. Dyn. Syst.* **14**:3 (2015), 1326–1347. MR Zbl

[Oh and Tzvetkov 2017] T. Oh and N. Tzvetkov, "Quasi-invariant Gaussian measures for the cubic fourth order nonlinear Schrödinger equation", *Probab. Theory Related Fields* **169**:3-4 (2017), 1121–1168. MR Zbl

[Özsarı and Yolcu 2019] T. Özsarı and N. Yolcu, "The initial-boundary value problem for the biharmonic Schrödinger equation on the half-line", *Commun. Pure Appl. Anal.* **18**:6 (2019), 3285–3316. MR

[Pausader 2007] B. Pausader, "Global well-posedness for energy critical fourth-order Schrödinger equations in the radial case", *Dyn. Partial Differ. Equ.* **4**:3 (2007), 197–225. MR Zbl

[Pausader 2009a] B. Pausader, "The cubic fourth-order Schrödinger equation", *J. Funct. Anal.* **256**:8 (2009), 2473–2517. MR Zbl

[Pausader 2009b] B. Pausader, "The cubic fourth-order Schrödinger equation", *J. Funct. Anal.* **256**:8 (2009), 2473–2517. MR Zbl

[Pausader and Shao 2010] B. Pausader and S. Shao, "The mass-critical fourth-order Schrödinger equation in high dimensions", *J. Hyperbolic Differ. Equ.* **7**:4 (2010), 651–705. MR Zbl

[Segata 2004] J.-i. Segata, "Remark on well-posedness for the fourth order nonlinear Schrödinger type equation", *Proc. Amer. Math. Soc.* **132**:12 (2004), 3559–3568. MR Zbl

[Tao 2006] T. Tao, *Nonlinear dispersive equations*, CBMS Regional Conference Series in Mathematics **106**, Amer. Math. Soc, Providence, RI, 2006. Local and global analysis. MR Zbl

[Tsutsumi 2014] Y. Tsutsumi, "Strichartz estimates for Schrödinger equation of fourth order with periodic boundary condition", pp. 104–113 in *Regularity and Singularity for Partial Differential Equations with Conservation Laws* (Kyoto), vol. 1914, 2014.

[Wen et al. 2014] R. Wen, S. Chai, and B.-Z. Guo, "Well-posedness and exact controllability of fourth order Schrödinger equation with boundary control and collocated observation", *SIAM J. Control Optim.* **52**:1 (2014), 365–396. MR Zbl

Received December 23, 2018. Revised December 29, 2019.

ROBERTO DE A. CAPISTRANO-FILHO
UNIVERSIDADE FEDERAL DE PERNAMBUCO
RECIFE, PE
BRAZIL
roberto.capistranofilho@ufpe.br

MÁRCIO CAVALCANTE
UNIVERSIDADE FEDERAL DE ALAGOAS (UFAL)
MACEIÓ, AL
BRAZIL
marcio.melo@im.ufal.br

FERNANDO A. GALLEGO
UNIVERSIDAD NACIONAL DE COLOMBIA, SEDE MANIZALES
MANIZALES
COLOMBIA
fagallegor@unal.edu.co

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.71

THE ARITHMETIC HODGE INDEX THEOREM AND RIGIDITY OF DYNAMICAL SYSTEMS OVER FUNCTION FIELDS

ALEXANDER CARNEY

In one of the fundamental results of Arakelov's arithmetic intersection theory, Faltings and Hriljac (independently) proved the Hodge index theorem for arithmetic surfaces by relating the intersection pairing to the negative of the Néron–Tate height pairing. More recently, this has been generalized to higher dimensions by Moriwaki and by Yuan and Zhang. We extend these results to projective varieties over transcendence degree one function fields. The new challenge is dealing with nonconstant but numerically trivial line bundles coming from the constant field via Chow's K/k-trace functor. As an application of the Hodge index theorem, we also prove a rigidity theorem for the set of canonical height zero points of polarized algebraic dynamical systems over function fields. For function fields over finite fields, this gives a rigidity theorem for preperiodic points, generalizing previous work of Mimar, of Baker and DeMarco, and of Yuan and Zhang.

1. Introduction

The Hodge index theorem states classically that the divisor intersection pairing on an algebraic surface has signature $+1, -1, \ldots, -1$. The corresponding result for line bundles on arithmetic surfaces, i.e., relative curves over the ring of integers of a number field, was proven independently by Faltings [1984] and Hriljac [1985], and is a fundamental result in Arakelov theory. More recently, Moriwaki [1996] extended this to higher-dimensional arithmetic varieties, and Yuan and Zhang [2017] proved a Hodge index theorem for adelic metrized line bundles over $\overline{\mathbb{Q}}$.

In their work, Yuan and Zhang also conjectured that a similar result should hold over function fields. Here we prove their conjecture. Our theorem statement differs slightly from their conjecture, however, so as to be stated more directly and to avoid reliance on a noncanonical isogeny between Chow's function field K/k-trace and K/k-image.

MSC2020: 11G50, 14G40, 37P30.
Keywords: arithmetic dynamics, intersection theory, Arakelov theory.

Let k be an algebraically closed field of arbitrary characteristic, and let $K = k(B)$ be the function field of B, a smooth, projective curve over k. Let $\pi : X \to \operatorname{Spec}(K)$ be a geometrically normal, geometrically integral, projective variety of dimension $n \geq 1$. We will consider the group $\widehat{\operatorname{Pic}}(X)$ of adelic metrized line bundles on X in the sense of [Zhang 1995]; definitions will be recalled in Section 2.1. Since an adelic metric can be specified, for example, by a line bundle on a model $\mathcal{X} \to B$ of X, this setting also covers relative varieties fibered over B, in the same way that Yuan and Zhang's work over number fields encompasses Arakelov's setting of arithmetic varieties over the spectrum of the ring of integers of a number field.

Chow's K/k-trace functor $\operatorname{Tr}_{K/k}$ identifies the part of the Picard variety of X which is defined over k, and the line bundles in $\operatorname{Tr}_{K/k}(\operatorname{Pic}^0(X))$ can all be given adelic metrics in a well-defined canonical way using isotrivial models over B. This construction is detailed in Section 2.5. Let $\operatorname{Pic}^\tau(X)$ be the numerically trivial subgroup of $\operatorname{Pic}(X)$. We prove the following result, with more detailed versions stated in Section 3:

Theorem 1.1. *Let $M, N \in \operatorname{Pic}^\tau(X)$, and let $L_1, \dots L_{n-1} \in \operatorname{Pic}(X)$ be ample. There exist canonical metrics on M and N so that*

$$\langle M, N\rangle_{L_1,\dots,L_{n-1}} := \overline{M} \cdot \overline{N} \cdot \bar{L}_1 \cdots \bar{L}_{n-1}$$

is a well-defined bilinear pairing, independent of the choice of the metrics on $L_1, \dots, L_{n-1}$. This extends to a symmetric $\mathbb{R}$-bilinear form on $\operatorname{Pic}^\tau(X) \otimes_{\mathbb{Z}} \mathbb{R}$ which is negative semidefinite with kernel

$$\operatorname{Tr}_{K/k}(\operatorname{Pic}^0(X)) \otimes_{\mathbb{Z}} \mathbb{R}.$$

If one removes the function field trace (so that the kernel is trivial), this is the same result that Yuan and Zhang prove for number fields. It is straightforward to see that $\operatorname{Tr}_{K/k}(\operatorname{Pic}^0(X)) \otimes_{\mathbb{Z}} \mathbb{R}$ is in the kernel. Thus, the main new difficulty is showing that numerically trivial adelic metrized line bundles which are nonconstant must all come from isotrivial subgroups of the Picard group of X. In essence, all arguments of the proof must commute with the K/k-trace functor.

1.1. ***Arithmetic dynamics.*** Again let X be a projective variety over a function field K. A polarized dynamical system (f, L, q) is an endomorphism $f : X \to X$ along with an ample line bundle $L \in \operatorname{Pic}(X)$ such that $f^*L \cong L^{\otimes q}$ for some $q > 1$. The set of preperiodic points of f is defined as

$$\operatorname{Prep}(f) := \{x \in X(\overline{K}) \mid x \text{ has a finite forward orbit under } f\}.$$

Call and Silverman [1993] show that such a polarized endomorphism defines a canonical Weil height $\widehat{h}_f$. Here we show that L can be given an *admissible* metric $\bar{L}_f$ so that the height $h_{\bar{L}_f}$ defined by $\bar{L}_f$ via arithmetic intersections agrees with $\widehat{h}_f$ on $X(\overline{K})$. The advantage to our definition is that $h_{\bar{L}_f}$ defines not only heights of points, but heights of subvarieties of X as well. By applying the Hodge index theorem to compare the canonical heights defined by two different polarized dynamical systems, we prove the following rigidity theorem:

Theorem 1.2. *Let X be a projective variety defined over a transcendence degree one function field K over any base k, and let (f, L, q) and (g, M, r) be two polarized dynamical systems on X. If the points with height zero under $h_{\bar{L}_f}$ and the points with height zero under $h_{\overline{M}_g}$ agree on a Zariski dense subset of $X(\overline{K})$, then they are identical.*

This is stated more generally in Section 4. When k is the algebraic closure of a finite field, the Northcott property implies that the points with canonical height zero under f are exactly the preperiodic points $\mathrm{Prep}(f)$, giving an immediate corollary.

Corollary 1.3. *In the same setting as Theorem 1.2 but with the additional hypothesis that $k = \overline{\mathbb{F}}_q$, if* $\mathrm{Prep}(f) \cap \mathrm{Prep}(g)$ *is Zariski dense in $X(\overline{K})$, then* $\mathrm{Prep}(f) = \mathrm{Prep}(g)$.

This was conjectured by Yuan and Zhang, and they prove a similar result over number fields.

Over general function fields the corollary does not hold, as not all canonical height zero points are preperiodic. The proofs differ as well, as while it is clear that the set $\mathrm{Prep}(f)$ does not depend on the choice of polarization L, this must be proven for canonical height zero points, and then the heights compared in a more indirect way. Even so, some limited things can be said.

Corollary 1.4. *Let K be the function field of a smooth projective curve over any field k, and let f and g be two rational functions $\mathbb{P}^1_K \to \mathbb{P}^1_K$ which are not isotrivial. If* $\mathrm{Prep}(f)$ *and* $\mathrm{Prep}(g)$ *intersect on an infinite subset of $\mathbb{P}^1(\overline{K})$, they are equal.*

This is a direct consequence of our theorem and a theorem of Baker [2009]. Chatzidakis and Hrushovski [2008a; 2008b] prove results comparing preperiodic points and height zero points using a model-theoretic nonisotriviality condition in a much more general setting, but it is difficult to combine that result into a useful rigidity statement. This is discussed further in Section 5.

1.2. ***Outline of paper and sketch of methods.*** Definitions and basic properties of adelic metrized line bundles and Chow's K/k-image and trace are recalled in Section 2. Additionally, this section includes technical lemmas, such as the existence of flat metrics, which will be needed throughout the paper.

Our main Hodge index theorem, a classification of numerically trivial line bundles, and an $\mathbb{R}$-linear variant (Theorems 3.1, 3.2, and 3.3) are fully stated and proven in Section 3. We begin with the case of X being a curve. Decomposing adelic metrized line bundles into flat and vertical pieces, and addressing intersections of the vertical parts using the local Hodge index theorem of [Yuan and Zhang 2017, Theorem 2.1], we reduce to the case of flat metrics. Then, following the methods of Faltings [1984] and Hriljac [1985], we relate the intersection pairing to the Néron–Tate height pairing on the Jacobian variety of X, and complete the result for curves using properties of heights on the Jacobian.

Next we prove the inequality part of Theorem 3.1 by induction on the dimension of X, using a Bertini-type theorem of Seidenberg [1950] to find sections which cut out nice subvarieties of X. Along the way we prove a Cauchy–Schwarz inequality for this intersection pairing. Theorem 3.2 and the equality part of Theorem 3.1 are then also proved by induction, where we again decompose into flat and vertical metrics and must show that the K/k-trace and image functors behave nicely when restricted to a subvariety. This is more difficult than the inequality, however. For the inequality, we write each metrized line bundle as a limit of model metrics, and prove the result for model metrics, thus getting the same inequality on their limit. We can write the same limit in the equality case, but we cannot assume that the same equality hypothesis holds for the model metrics, and must argue by other means. Finally, Theorem 3.3 is easily deduced from Theorem 3.1 and its proof.

Section 4 proves the application of our result to polarized algebraic dynamical systems. We first describe and prove the existence of admissible metrics for a given polarized algebraic dynamical system, which generalize flat metrics, and give rise to canonical heights defined by intersections. This transforms the rigidity statement into a statement comparing two different admissible adelic metrized line bundles, which is proved using the Hodge index theorem.

Section 5 gives corollaries of the main results proven here, and discusses what can still be said about preperiodic points over larger fields without the Northcott property.

2. Preliminaries

Here we introduce the definitions, basic properties, and lemmas which will be needed throughout the paper. The core theory used in this paper is built on local intersection theory as developed by Gubler [1998; 2007b], Chambert-Loir [2006], Chambert-Loir and Thuillier [2009], and Zhang [1995]. More generally, one can find an introduction to Arakelov theory in [Moriwaki 2014; Lang 1988; Soulé 1992].

2.1. ***Metrized line bundles over local fields.*** Let K be a complete non-Archimedean field with nontrivial absolute value $|\cdot|$. Denote the valuation ring of K by

$$K^{\circ} := \{a \in K : |a| \leq 1\},$$

and its maximal ideal by

$$K^{\circ\circ} := \{a \in K : |a| < 1\},$$

so that $\widetilde{K} := K^{\circ}/K^{\circ\circ}$ is the residue field.

Let X be a variety over K and denote by X^{an} its Berkovich analytification [1990]. For $x \in X^{\mathrm{an}}$, write $K(x)$ for the residue field of x. A line bundle L on X has an analytification, denoted L^{an}, as a line bundle on X^{an}.

Definition 2.1. A *continuous metric* $\|\cdot\|$ on L consists of a $K(x)$-metric $\|\cdot\|_x$ on $L^{\mathrm{an}}(x)$ for every $x \in X^{\mathrm{an}}$, where this collection of metrics is continuous in the sense that for every rational section s of L, the map $X^{\mathrm{an}} \to \mathbb{R}$ defined by $x \mapsto \|s(x)\|_x$ is continuous away from the poles of s. We call L with a continuous metric a *metrized line bundle* and denote this by $\bar{L} = (L, \|\cdot\|)$. For a fixed line bundle L, limits of metrics are taken with respect to the topology induced by the supremum norm.

An important example of a continuous metric is a *model metric*: Let $\mathcal{X}$ be a model of X over K°, i.e., a projective, flat, finitely presented, integral scheme over $\operatorname{Spec} K^\circ$ whose generic fiber $\mathcal{X}_K$ is isomorphic to X, and let $\mathcal{L}$ be a line bundle on $\mathcal{X}$ whose generic fiber $\mathcal{L}_K$ is isomorphic to L. Then we can define a continuous metric on L by specifying that for any trivialization $\mathcal{L}_{\mathcal{U}} \xrightarrow{\sim} \mathcal{O}_{\mathcal{U}}$ on an open set $\mathcal{U} \subset \mathcal{X}$ given by a rational section ℓ, we have $\|\ell(x)\|_x = 1$ for any x reducing to $\mathcal{U}_{\widetilde{K}}$ in the reduction $\widetilde{X}$ over $\widetilde{K}$.

We now define some important properties and notation.

Definition 2.2. Let $\bar{L} = (L, \|\cdot\|)$ and $\overline{M}$ be metrized line bundles on X.

(1) A model metric is *nef* if it is given by a relatively nef line bundle on the corresponding model.

(2) Call both $\bar{L}$ and $\|\cdot\|$ *nef* if $\|\cdot\|$ is equal to a limit of nef model metrics.

(3) $\bar{L}$ is *arithmetically positive* if it is nef and L is ample.

(4) $\bar{L}$ is *integrable* if it can be written as $\bar{L} = \bar{L}_1 - \bar{L}_2$ with $\bar{L}_1$ and $\bar{L}_2$ nef.

(5) $\overline{M}$ is $\bar{L}$*-bounded* if there exists a positive integer m such that $m\bar{L} + \overline{M}$ and $m\bar{L} - \overline{M}$ are both nef.

(6) $\bar{L}$ is *vertical* if it is integrable and $L \cong \mathcal{O}_X$.

(7) $\bar{L}$ is *constant* if it is isometric to the pull-back of a metrized line bundle on $\operatorname{Spec} K$.

(8) $\widehat{\operatorname{Pic}}(X)$ is defined to be the group of isometry classes of integrable metrized line bundles.

Remark. When we say a line bundle is relatively ample or nef, we always mean with respect to the structure morphism; here $\mathcal{X} \to \operatorname{Spec} K^\circ$. A concise discussion of relative amplitude and nefness can be found in [Lazarsfeld 2004, Chapter 1.7].

We also have a local intersection theory for metrized line bundles on X, due to [Gubler 1998; 2007a], and to [Zhang 1995] when K has a discrete valuation. Let Z

be a d-dimensional cycle on X, let $\bar{L}_0, \dots, \bar{L}_d$ be integrable metrized line bundles on X, and $\ell_0, \dots, \ell_d$ rational sections of each, respectively, such that

$$\left(\bigcap_i |\mathrm{div}(\ell_i)|\right) \cap |Z| = \varnothing,$$

where $|Z|$ means the underlying topological space of the cycle Z. Then Z has a local height $\widehat{\mathrm{div}}(\ell_0) \cdots \widehat{\mathrm{div}}(\ell_d) \cdot [Z]$ with the following properties:

(1) The local height is linear in $\widehat{\mathrm{div}}(\ell_i)$ and Z.

(2) For fixed sections, it is continuous with respect to the metrics.

(3) When $\bar{L}_i$ has a model metric given by $\mathcal{L}_i$ on a common model $\mathcal{X}$, the height is given by classical intersections:

$$\widehat{\mathrm{div}}(\ell_0) \cdots \widehat{\mathrm{div}}(\ell_d) \cdot [Z] = \mathrm{div}_{\mathcal{X}}(\ell_0) \cdots \mathrm{div}_{\mathcal{X}}(\ell_d) \cdot [\mathcal{Z}],$$

where $\mathcal{Z}$ is the Zariski closure of Z in $\mathcal{X}$.

(4) If the support of $\mathrm{div}(\ell_0)$ contains no component of Z, there is a measure $c_1(\bar{L}_1) \cdots c_1(\bar{L}_d)\delta_Z$ on X^{an} due to [Chambert-Loir 2006] which allows us to compute $\widehat{\mathrm{div}}(\ell_0) \cdots \widehat{\mathrm{div}}(\ell_d) \cdot [Z]$ inductively as

$$\widehat{\mathrm{div}}(\ell_1) \cdots \widehat{\mathrm{div}}(\ell_d) \cdot [\mathrm{div}(\ell_0) \cdot Z] - \int_{X^{\mathrm{an}}} \log\|\ell_0(x)\|_x c_1(\bar{L}_1) \cdots c_1(\bar{L}_d)\, \delta_Z.$$

This notation is meant to suggest that $c_1(\bar{L}_i)$ should be thought of as the arithmetic version of the classical Chern form $c_1(L_i)$.

(5) If $\ell_0|_{Z_j}$ is constant and $c_1(L_1) \cdots c_1(L_d) \cdot [Z_j] = 0$ for every irreducible component Z_j of Z, then this pairing does not depend on the choice of sections, so we may simply write

$$\bar{L}_0 \cdots \bar{L}_d \cdot Z = \widehat{\mathrm{div}}(\ell_0) \cdots \widehat{\mathrm{div}}(\ell_d) \cdot [Z].$$

When $Z = X$, we typically omit Z in all of the above notation.

By definition, every integrable metric can be written as a limit of model metrics (with respect to the supremum norm). Properties (3) and (4) above guarantee that intersections of integrable metrized line bundles are equal to the corresponding limits of intersections of models which approximate them.

2.2. ***Adelic metrized line bundles.*** We now move to the global theory, which is built from the theory of metrized line bundles over each localization, discussing first models and then adelic metrized line bundles. We return to the setting of the main theorems of this paper, where k is any algebraically closed field, B is a smooth projective curve over k, $K = k(B)$ is its function field, and $\pi : X \to \mathrm{Spec}(K)$ is a geometrically normal, geometrically integral, projective variety.

Let $\mathcal{X}$ be a model for X, meaning that $\mathcal{X} \to B$ is geometrically integral, projective, and flat, and the generic fiber $\mathcal{X}_K$ is isomorphic to X. Given a geometrically integral subvariety $\mathcal{Y}$ of dimension $d+1$ in $\mathcal{X}$ and line bundles $\mathcal{L}_0, \ldots, \mathcal{L}_d$ on $\mathcal{X}$ each with a respective section $\ell_0, \ldots, \ell_d$ such that their common support has empty intersection with $\mathcal{Y}_K$, the arithmetic intersection pairing on $\operatorname{Pic}(\mathcal{X})$ is defined locally as

$$\mathcal{L}_0 \cdots \mathcal{L}_d \cdot \mathcal{Y} := \widehat{\operatorname{div}}(\ell_0) \cdots \widehat{\operatorname{div}}(\ell_d) \cdot [\mathcal{Y}] := \sum_{\nu} \bigl(\widehat{\operatorname{div}}(\ell_0) \cdots \widehat{\operatorname{div}}(\ell_d) \cdot [\mathcal{Y}]\bigr)_{\nu},$$

where ν ranges over the closed points (places) of B, and

$$\bigl(\widehat{\operatorname{div}}(\ell_0) \cdots \widehat{\operatorname{div}}(\ell_d) \cdot [\mathcal{Y}]\bigr)_{\nu}$$

means the local intersection number after base-change to the complete field K_ν. As the notation suggests, this does not depend on the choice of sections. Again we typically drop $\mathcal{Y}$ in the notation if $\mathcal{Y} = \mathcal{X}$, and when $\mathcal{X}$ is one-dimensional, we call $\widehat{\deg}(\mathcal{L}_0) := \mathcal{L}_0 \cdot \mathcal{X}$ the arithmetic degree of $\mathcal{L}_0$.

Remark. This arithmetic intersection theory for $\mathcal{X} \to B$ is equal to the classical intersection theory given by viewing $\mathcal{X}$ as a variety over the field k, but is written using the fibration so as to align notationally with Arakelov's arithmetic intersection theory [1974; 1975]. In the function field setting there are no Archimedean places to consider, as B is projective.

Given a line bundle L on X we call a line bundle $\mathcal{L}$ on $\mathcal{X}$ a model for L provided that $\mathcal{L}_K \cong L$. For each place ν of B, completing with respect to ν induces a model over K_ν° and a model metric $\|\cdot\|_{\mathcal{L},\nu}$ of $L_{K_\nu}^{\mathrm{an}}$ on $X_\nu^{\mathrm{an}} := X_{K_\nu}^{\mathrm{an}}$.

Definition 2.3. The collection $\|\cdot\|_{\mathcal{L},\mathbb{A}} = \{\|\cdot\|_{\mathcal{L},\nu}\}_\nu$ of continuous metrics for every place ν of B given by $(\mathcal{X}, \mathcal{L})$ is called a *model adelic metric* on L. More generally, an *adelic metric* $\|\cdot\|_{\mathbb{A}}$ on L is a collection of continuous metrics $\|\cdot\|_\nu$ of $L_{K_\nu}^{\mathrm{an}}$ on X_ν^{an} for every place ν, which agrees with some model adelic metric at all but finitely many places. A line bundle on X with an adelic metric is called an *adelic metrized line bundle*, and is denoted $\bar{L} = (L, \|\cdot\|_{\mathbb{A}})$. When the context is clear we will frequently drop adelic and simply write *metrized line bundle*. For a fixed line bundle L, limits of adelic metrics are taken with respect to the topology induced by $\max_\nu \|\cdot\|_{\sup}$, the maximum of the supremum norm on each fiber. Such a limit does not require fixing a single model $\mathcal{X}$.

We extend our local definitions of properties of metrized line bundles to the global case.

Definition 2.4. Let $\bar{L}$ be an adelic metrized line bundle.

(1) $\bar{L}$ is *nef* if it is equal to a limit of model metrics induced by nef line bundles on models of X.

(2) $\bar{L}$ is *integrable* if it can be written as $\bar{L} = \bar{L}_1 - \bar{L}_2$, where each $\bar{L}_i$ is nef.

(3) $\bar{L}$ is *arithmetically positive* if L is ample and $\bar{L} - \pi^* \overline{N}$ is nef for some adelic metrized line bundle $\overline{N}$ on Spec K with $\widehat{\deg}(\overline{N}) > 0$.

(4) $\overline{M}$ is $\bar{L}$*-bounded* if there exists a positive integer m such that $m\bar{L} + \overline{M}$ and $m\bar{L} - \overline{M}$ are both nef.

(5) $\bar{L}$ is *vertical* if it is integrable and $L \cong \mathcal{O}_X$

(6) $\bar{L}$ is *constant* if it is isometric to the pull-back of a metrized line bundle on Spec K.

(7) $\widehat{\mathrm{Pic}}(X)$ is defined to be the group of isometry classes of integrable metrized line bundles.

Remark. In the definition of arithmetically positive, we have thus far only defined the arithmetic degree in the model case, but every adelic metrized line bundle in Spec K has a model metric, so we may use that definition. The definition is also made more general in the following material.

Remark. The definition of arithmetically positive is equivalent to requiring that L is ample and for every $\overline{N} \in \widehat{\mathrm{Pic}}(K)$, there exists some positive integer m such that $m\bar{L} - \pi^*\overline{N}$ is nef. This means, in particular, that all of $\pi^*\widehat{\mathrm{Pic}}(K)$ is $\bar{L}$-bounded for arithmetically positive $\bar{L}$.

Remark. To avoid confusion, note that the preceding definitions are not equivalent to requiring that the local property of the same name holds at every fiber. In fact, since relative amplitude (resp. nefness) holds if and only if the restriction to every fiber is ample (resp. nef), if a property holds in the global setting then the corresponding property holds locally at every place, but the converse is false. For example, if $\bar{L}_v$ is nef on X_v for every place, each $\bar{L}_v$ can be written as a limit of nef models $\mathcal{L}_{v,i}$ on $\mathcal{X}_{v,i}$, but it may not be possible to assemble these into global models $\mathcal{L}_i$ on models $\mathcal{X}_i$ of X.

Global intersections are defined similarly to the model case, except with the local metrics given explicitly by the adelic metric instead of induced by a model. Given a d-dimensional integral subvariety Z of X and integrable adelic metrized line bundles $\bar{L}_0, \ldots, \bar{L}_d$ with respective sections $\ell_0, \ldots, \ell_d$ with empty common intersection with Z, their intersection is

$$\begin{aligned}\bar{L}_0 \cdots \bar{L}_d \cdot Z :&= \widehat{\mathrm{div}}(\ell_0) \cdots \widehat{\mathrm{div}}(\ell_d) \cdot [Z] \\ &= \sum_v \widehat{\mathrm{div}}(\ell_0|_{X_v}) \cdots \widehat{\mathrm{div}}(\ell_d|_{X_v}) \cdot [Z|_{X_v}],\end{aligned}$$

where again this is independent of the choice of sections. Summing the local induction formula at each place produces a global induction formula: letting ℓ_0 be a rational section of $\bar{L}_0$ whose support does not contain Z,

$$\begin{aligned}\bar{L}_0 \cdots \bar{L}_d \cdot Z \\ &= \bar{L}_1 \cdots \bar{L}_d \cdot (Z \cdot \operatorname{div}(\ell_0)) - \sum_{\nu} \int_{X_\nu^{\mathrm{an}}} \log\|\ell_0(x)\|_\nu c_1(\bar{L}_1, \nu) \cdots c_1(\bar{L}_d, \nu)\, \delta_Z|_{X_\nu}.\end{aligned}$$

As before, we drop Z when $Z = X$, and when X is zero-dimensional, we call $\widehat{\deg}(\bar{L}_0) := \bar{L}_0 \cdot X$ the arithmetic degree of $\bar{L}_0$.

As in the local case, we can always compute intersections of adelic metrized line bundles by approximating them with model metrics and computing the limit of the corresponding arithmetic intersections of the models.

Definition 2.5. An adelic metrized line bundle $\overline{M}$ on X of dimension n is called *numerically trivial* if for any $\bar{L}_1, \dots, \bar{L}_n \in \widehat{\mathrm{Pic}}(X)$,

$$\overline{M} \cdot \bar{L}_1 \cdots \bar{L}_n = 0.$$

Call two adelic metrized line bundles *numerically equivalent* if their difference is numerically trivial.

2.3. *Flat metrics.* Adelic metrized line bundles with flat metrics form an especially nice class of adelic metrized line bundles. We will often be able to split a metrized line bundle into a bundle with a flat metric plus a vertical bundle, and then work with each of these separately, as flatness will tell us that these have trivial intersection.

Definition 2.6. Let X be a projective variety over a complete field K, and let $\bar{L}$ be a metrized line bundle on X. Then $\bar{L}$ is *flat* if for any morphism $f : C \to X$ of a projective curve over K into X, we have $c_1(f^*\bar{L}) = 0$ on the Berkovich analytification C^{an}. If now X is a projective variety over a global field and $\bar{L}$ an adelic metrized line bundle on X, call $\bar{L}$ flat provided it is flat at every place.

Note that if $\bar{L}$ is flat, L must be numerically trivial, as

$$\deg(L|_C) = \int_{C^{\mathrm{an}}} c_1(\bar{L}|_C) = 0.$$

Lemma 2.7. *Let L be a numerically trivial line bundle on a projective, normal variety X over a function field K. Then L has a flat metric, which is unique up to constant multiple.*

Remark. When X is a curve, this lemma has a much simpler proof using linear algebra; see for example [Hriljac 1985, Theorem 1.3]. If $\mathcal{X} \to B$ is a model for X and $\mathcal{X}_\nu$ is geometrically normal (for example, every place ν of good reduction), then the flat metric on L at ν is induced by the model metric corresponding to the closure in $\mathcal{X}$ of a divisor on X in the class of L.

To prove the lemma in general, the following related notion will be useful. We will show that it is equivalent to flatness for abelian varieties.

Definition 2.8. Let $\bar{L}$ be a metrized line bundle on an abelian variety A such that L is algebraically trivial. We call $\bar{L}$ *admissible* if $[2]^*\bar{L} \cong 2\bar{L}$.

Proof of Lemma 2.7. First, suppose X is an abelian variety. Then L is algebraically trivial, and we have an isomorphism $\phi : [2]^*L \cong 2L$. Take any metric $\|\cdot\|_1$ on L. Then Tate's limiting argument, as in [Zhang 1995, Theorem 2.2], shows that

$$\|\cdot\|_n := \phi^*[2]^*\|\cdot\|_{n-1}^{\frac{1}{2}}$$

converges to an admissible metric $\|\cdot\|_0$ on L, and that further this is the unique admissible metric on L up to constant multiples.

Let $C \to X$ be a smooth projective curve in X. After a translation and extension of K, we can fix a point $x_0 \in C(K)$ which maps to $0 \in X$. By the universal property of the Jacobian, $C \to X$ factors through the Jacobian map $C \to \mathrm{Jac}(C)$ taking $x_0 \to 0$, and the pullback of $(L, \|\cdot\|_0)$ to $\mathrm{Jac}(C)$ is also admissible. Then by [Gubler 2007b, Remark 3.14], $c_1(L, \|\cdot\|_0) = 0$, and hence L has a flat metric. By taking the tensor product of this metric with the inverse of any other flat metric on L, uniqueness up to constant multiple is reduced to showing that $\|1\|$ is constant for any flat metric on $\mathcal{O}_X$. Any two points on X are connected by a curve; let D be its normalization. Then $\|1\|$ is constant by the local Hodge index theorem [Yuan and Zhang 2017, Theorem 2.1] in dimension one at each place.

Now let X be an arbitrary projective, normal variety, choose a point $x_0 \in X(K)$ (extending K if necessary) and recall the Albanese map $i : X \to \mathrm{Alb}(X)$ taking x_0 to 0. Since L is numerically trivial, we may replace it by a multiple and assume it is algebraically trivial. Then L corresponds to a K point ξ of $\mathrm{Pic}^0_{\mathrm{red},X} = \mathrm{Alb}(X)^\vee$. By definition, L is (isomorphic to) the Poincaré bundle P on $\mathrm{Alb}(X) \times \mathrm{Alb}(X)^\vee$ restricted to $\mathrm{Alb}(X) \times \{\xi\}$, then pulled back through

$$i \times \mathrm{id} : X \times \mathrm{Alb}(X)^\vee \to \mathrm{Alb}(X) \times \mathrm{Alb}(X)^\vee.$$

$P|_{\mathrm{Alb}(X)\times\{\xi\}}$ is algebraically trivial, and hence has a flat metric. But the pullback of a flat metric is also flat, so this defines a flat metric for L. □

The reason we care about flat metrics is shown by Lemma 2.9 and Corollary 2.10:

Lemma 2.9. *Let K be a complete non-Archimedean field, and $X \to \operatorname{Spec} K$ a geometrically connected, geometrically normal, projective variety of dimension n, with a flat metrized line bundle $\overline{M}$. Then given any integrable metrized line bundles $\bar{L}_1, \dots, \bar{L}_{n-1}$ on X,*

$$c_1(\overline{M})c_1(\bar{L}_1)\cdots c_1(\bar{L}_{n-1}) = 0.$$

Proof. We show that

$$\int_{X^{\mathrm{an}}} \log\|\ell_n(x)\|_x c_1(\overline{M}) \cdot c_1(\bar{L}_1)\cdots c_1(\bar{L}_{n-1}) = 0$$

for every section ℓ_n of any metrized line bundle $\bar{L}_n$. Proceed by induction on n.

Since any integral metrized line bundle can be written as a difference of arithmetically positive metrized line bundles and the measure is additive with respect to the metrized line bundles, we may assume that $\bar{L}_{n-1}$ is arithmetically positive without loss of generality. Further, by approximation, it suffices to treat the case where $\bar{L}_{n-1}$ is a model metric, induced by some ample line bundle $\mathcal{L}$ on a model $\mathcal{X}$ for X. By Seidenberg's Bertini theorem [Seidenberg 1950, Theorem 7'], $\mathcal{L}$ has a section s which cuts out a horizontal, geometrically integral, normal subvariety $\mathcal{Y}$. After base changing to a finite extension K' of K, we may assume that this subvariety is geometrically normal. Since this extension merely scales the intersection number by $[K' : K]$ it has no effect on the proof of this lemma. Let Y be the generic fiber of $\mathcal{Y}$, and let Z be $\mathrm{div}(\ell_n)$ restricted to Y.

We compute an intersection product in two different ways. First,

$$\begin{aligned}\overline{M}\cdot\bar{L}_1\cdots\bar{L}_n &= \overline{M}|_Y\cdot\bar{L}_1|_Y\cdots\bar{L}_{n-2}|_Y\cdot\bar{L}_n|_Y\\ &= \overline{M}|_Z\cdot\bar{L}_1|_Z\cdots\bar{L}_{n-2}|_Z-\int_{X^{\mathrm{an}}}\log\|\ell_n(x)\|_x c_1(\overline{M})c_1(\bar{L}_1)\cdots c_1(\bar{L}_{n-2})\,\delta_Y\\ &= \overline{M}|_Z\cdot\bar{L}_1|_Z\cdots\bar{L}_{n-2}|_Z,\end{aligned}$$

where the first equality follows from $\mathcal{Y}$ being horizontal, the second from the induction formula for local intersection numbers, and the third from the induction hypothesis. We now compute this in a different order:

$$\begin{aligned}&\overline{M}\cdot\bar{L}_1\cdots\bar{L}_n\\ &\quad= \overline{M}\cdot\bar{L}_1\cdots\bar{L}_{n-1}\cdot(\mathrm{div}(\ell_n)) - \int_{X^{\mathrm{an}}}\log\|\ell_n(x)\|_x c_1(\overline{M})c_1(\bar{L}_1)\cdots c_1(\bar{L}_{n-1})\\ &\quad= \overline{M}|_Z\cdot\bar{L}_1|_Z\cdots\bar{L}_{n-2}|_Z - \int_{X^{\mathrm{an}}}\log\|\ell_n(x)\|_x c_1(\overline{M})c_1(\bar{L}_1)\cdots c_1(\bar{L}_{n-1}),\end{aligned}$$

where now the first inequality follows from the induction formula, and the second from $\mathcal{Y}$ being horizontal. Comparing the two equalities completes the proof. □

Corollary 2.10. *Let $\overline{M}$ be flat, $\overline{N}$ be vertical, and $\bar{L}_1, \ldots, \bar{L}_{n-1}$ be any integrable adelic metrized line bundles. Then*

$$\overline{M}\cdot\overline{N}\cdot\bar{L}_1\cdots\bar{L}_{n-1}=0.$$

Proof. Since $\overline{N}$ is vertical, $N=\mathcal{O}_X$. Compute this intersection using the induction formula with the section $s=1$ of $\mathcal{O}_X$:

$$\begin{aligned}&\overline{M}\cdot\overline{N}\cdot\bar{L}_1\cdots\bar{L}_{n-1}\\ &\qquad= \overline{M}\cdot\bar{L}_1\cdots\bar{L}_{n-1}\cdot(\mathrm{div}(s)) - \int_{X^{\mathrm{an}}}\log\|1\|_x c_1(\overline{M})\cdot c_1(\bar{L}_1)\cdots c_1(\bar{L}_{n-1}).\end{aligned}$$

The first term is zero since $\mathrm{div}(s)$ is empty, and the integral is zero by Lemma 2.9. □

2.4. ***Heights of points and subvarieties.*** An important application of the intersection theory of adelic metrized line bundles is to define height functions.

Definition 2.11. Let $\overline{N} \in \widehat{\mathrm{Pic}}(X)$. We define the *height* of a point $x \in X(\overline{K})$ by

$$h_{\overline{N}}(x) := \frac{1}{[K(x):K]} \overline{N} \cdot \tilde{x},$$

where $\tilde{x}$ is the image of x in X via $X_{K(x)} \to X_K = X$.

Remark. The heights produced by this definition are Weil heights, which can be defined without intersection theory [Bombieri and Gubler 2006; Call and Silverman 1993], but we use the above definition as it generalizes to define heights of subvarieties.

Definition 2.12. Let $d = \dim Y$. The *height* of Y with respect to $\overline{N}$ is defined to be

$$h_{\overline{N}}(Y) := \frac{(\overline{N}|_Y)^{d+1}}{(d+1)(N|_Y)^d}$$

and the *essential minimum* of Y with respect to $\overline{N}$ is

$$\lambda_1(Y, \overline{N}) := \sup_{\substack{U \subset Y \\ \text{open}}} \Big(\inf_{x \in U(\overline{K})} h_{\overline{N}|_Y}(x) \Big).$$

By the successive minima of Zhang [1995, Theorem 1.1], and proven in the function field setting by Gubler [2007a, Theorem 4.1], we can state the following.

Proposition 2.13. *When $\overline{N}$ is nef,*

$$\lambda_1(Y, \overline{N}) \geq h_{\overline{N}}(Y) \geq 0.$$

2.5. ***Abelian varieties and Chow's K / k-trace and image.*** Proofs of the existence and properties of the trace and image can be found in [Lang 1983] and [Conrad 2006]. Let A be an abelian variety defined over K. The K/k-image $(\mathrm{Im}_{K/k}(A), \lambda)$ consists of an abelian variety $\mathrm{Im}_{K/k}(A)$ over k and a surjective morphism

$$\lambda : A \to \mathrm{Im}_{K/k}(A)_K$$

with the following universal property: If V is an abelian variety defined over k, and $\phi : A \to V_K$ is a morphism, then ϕ factors through λ. Provided the fields K and k are clear, we will often drop the K/k subscript and just write $\mathrm{Im}(A)$.

The K/k-trace is $(\mathrm{Tr}_{K/k}(A), \tau)$ where $\mathrm{Tr}_{K/k}(A)$ is an abelian variety over k, and

$$\tau : \mathrm{Tr}_{K/k}(A)_K \to A$$

is universal among all morphisms from k-abelian varieties to A. Again we will often drop the K/k when the fields are unambiguous. The image can be thought of as the largest quotient of A that can be defined over k and the trace as the largest abelian subvariety that can be defined over k. This heuristic is literally true in characteristic zero, but in positive characteristic the trace map may have an infinitesimal kernel; see [Conrad 2006, Section 6].

These constructions are dual to each other in the sense that

$$\mathrm{Tr}(A^\vee) = \mathrm{Im}(A)^\vee,$$

and the image and trace are isogenous via the composition $\lambda \circ \tau$ (descended to the k-varieties).

Given a morphism of abelian varieties $f : A \to B$, we get morphisms $f_{\mathrm{Tr}} : \mathrm{Tr}(A) \to \mathrm{Tr}(B)$ and $f_{\mathrm{Im}} : \mathrm{Im}(A) \to \mathrm{Im}(B)$ commuting with τ and λ.

Now suppose X is a geometrically normal projective variety over K of dimension n, and assume that K is large enough so that $X(K)$ is nonempty. We write $\mathbf{Pic}_X$ for the Picard scheme of X, representing the Picard functor on X. This scheme exists (i.e., the Picard functor is representable), and its reduced neutral component, denoted $\mathbf{Pic}^0_{\mathrm{red},X}$, is an abelian variety [Kleiman 2005; Grothendieck 1962, Lecture 236]. Note that we do require the reduction, as $\mathbf{Pic}^0_X$ may fail to be reduced in positive characteristic. Write $\mathrm{Pic}(X)$ and $\mathrm{Pic}^0(X)$ for the abelian groups of K points of $\mathbf{Pic}_X$ and $\mathbf{Pic}^0_{\mathrm{red},X}$, respectively.

We can then define Alb_X, called the Albanese variety of X, to be the abelian variety dual to $\mathbf{Pic}^0_{\mathrm{red},X}$. Choosing a point $x_0 \in X(K)$ fixes an Albanese morphism

$$\iota : X \to \mathrm{Alb}_X$$

taking x_0 to 0, and then (Alb_X, ι) uniquely satisfies the Albanese universal property: any morphism from X to an abelian variety taking x_0 to zero must factor through ι [Wittenberg 2008].

We now have the language to differentiate between metrized line bundles defined over the constant field k and those which are not. Define a group homomorphism

$$\widehat{\tau}_{K/k} : \mathrm{Tr}_{K/k}(\mathbf{Pic}^0_{\mathrm{red},X})(k) \to \widehat{\mathrm{Pic}}(X)$$

as follows. First, by the duality of the K/k-trace and image,

$$\mathrm{Tr}_{K/k}(\mathbf{Pic}^0_{\mathrm{red},X})(k) = \mathrm{Pic}^0(\mathrm{Im}_{K/k}(\mathrm{Alb}_X)).$$

Then we can map

$$\mathrm{Pic}^0(\mathrm{Im}_{K/k}(\mathrm{Alb}_X)) \hookrightarrow \mathrm{Pic}(\mathrm{Im}_{K/k}(\mathrm{Alb}_X)) \to \mathrm{Pic}(\mathrm{Im}_{K/k}(\mathrm{Alb}_X) \times_k B),$$

where the map on the right is the pullback of projection onto the first factor. Since $\mathrm{Im}_{K/k}(\mathrm{Alb}_X)$ is defined over k, the fibered product $\mathrm{Im}_{K/k}(\mathrm{Alb}_X) \times_k B$ is a model for $\mathrm{Im}_{K/k}(\mathrm{Alb}_X) \times_k K$, and thus we get a map

$$\mathrm{Pic}(\mathrm{Im}_{K/k}(\mathrm{Alb}_X) \times_k B) \to \widehat{\mathrm{Pic}}(\mathrm{Im}_{K/k}(\mathrm{Alb}_X)_K)$$

given by taking model metrics. Finally, X maps to $\mathrm{Im}_{K/k}(\mathrm{Alb}_X)_K$ via the Albanese map followed by the image map, and pulling this back gives

$$\widehat{\mathrm{Pic}}(\mathrm{Im}_{K/k}(\mathrm{Alb}_X)_K) \to \widehat{\mathrm{Pic}}(X).$$

We can thus define $\widehat{\tau}_{K/k}$ as the composition of the above maps. While it took several steps to formally define $\widehat{\tau}_{K/k}$, it is very natural; if we define

$$\phi : \widehat{\mathrm{Pic}}(X) \to \mathrm{Pic}(X)$$

by forgetting the metric, then

$$\phi \circ \widehat{\tau}_{K/k} = \tau_{K/k}$$

is the K/k-trace morphism (on field-valued points), and the image of this composition lands in $\mathrm{Pic}^0(X)$. To simplify notation, we write $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))$ to mean the image of $\widehat{\tau}_{K/k}$ in $\widehat{\mathrm{Pic}}(X)$. By construction $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))$ is flat and numerically trivial, as on every fiber X_ν this group restricts to $\mathrm{Tr}_{K/k}(\mathbf{Pic}^0_{\mathrm{red},X})_{K_\nu}(k)$, which is algebraically trivial, and so in particular $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))$ has zero intersection with every vertical metrized line bundle.

3. Proof of Hodge index theorem

3.1. ***Statement of results.*** Let k be any algebraically closed field, let B be a smooth projective curve over that field, and let $K = k(B)$ be the corresponding function field. Let X be a geometrically normal projective variety over K of dimension n, and assume that K is large enough so that $X(K)$ is nonempty. Then choosing a point $x \in X(K)$ we may fix an Albanese morphism $\iota : X \to \mathrm{Alb}_X$. We impose these conditions on X as well as this choice of Albanese morphism throughout the rest of the paper.

We can now state our main theorem:

Theorem 3.1 (arithmetic Hodge index theorem for function fields). *Let $\overline{M}$ be an integrable adelic $\mathbb{Q}$-line bundle on X and $\bar{L}_1, \ldots, \bar{L}_{n-1}$ nef adelic $\mathbb{Q}$-line bundles on X. Suppose if $n \geq 2$ that $M \cdot L_1 \ldots L_{n-1} = 0$ and each L_i is big, or that* $\deg M = 0$ *if $n = 1$. Then*

$$\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} \leq 0.$$

Further, if every $\bar{L}_i$ is arithmetically positive, and $\overline{M}$ is $\bar{L}_i$-bounded for every i, then

$$\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = 0$$

if and only if

$$\overline{M} \in \pi^* \widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}.$$

When $n = 1$ so that X is a curve,

$$\overline{M}^2 = -2h_{\mathrm{NT}}(M),$$

where h_{NT} is the Néron–Tate height on the Jacobian of X.

Remark. When k is the algebraic closure of a finite field, $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$ is zero, since all elements are torsion.

Remark. For the "if" direction of the equality, note that all of

$$\pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$$

is $\bar{L}$ bounded for any arithmetically positive $\bar{L}$. This follows from the remark after Definition 2.4 for $\pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}}$, and from the fact that $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$ is numerically trivial by construction.

Call a metrized line bundle $\overline{M}$ on X numerically trivial if

$$\overline{M} \cdot \bar{L}_1 \cdots \bar{L}_n = 0$$

for every choice of metrized line bundles $\bar{L}_1, \ldots, \bar{L}_n$. The classical Hodge index theorem says that the only divisors on a surface with zero self-intersection are the numerically trivial divisors. We show that that is nearly, but not quite the case here:

Theorem 3.2. *The following are equivalent for $\overline{M} \in \widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$:*

(1) *$\overline{M}$ is numerically trivial.*

(2) *The height $h_{\overline{M}}$ is identically zero on $X(\bar{K})$.*

(3) *$\overline{M} \in \pi^*\widehat{\mathrm{Pic}}^0(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$, where $\widehat{\mathrm{Pic}}^0(K)_{\mathbb{Q}}$ is defined to be the elements of $\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}}$ with arithmetic degree zero.*

Define $\mathrm{Pic}^\tau(X)$ to be the group of isomorphism classes of numerically trivial line bundles on X. We define a pairing on $\mathrm{Pic}^\tau(X)$ to give an $\mathbb{R}$-linear version of Theorem 3.1. Let $M, N \in \mathrm{Pic}^\tau(X)_{\mathbb{R}}$, and let $L_1, \ldots, L_{n-1} \in \mathrm{Pic}(X)_{\mathbb{Q}}$ be nef. Then define a pairing by

$$\langle M, N\rangle_{L_1,\ldots,L_{n-1}} := \overline{M} \cdot \overline{N} \cdot \bar{L}_1 \cdots \bar{L}_{n-1},$$

using any choice of flat metrics on M and N, and any choice of metrics on L_i. By Lemma 5.19 of [Yuan and Zhang 2017], (proven as a simple consequence of Lemma 2.9 here) this pairing does not depend on the choice of metric.

Theorem 3.3. *For any $M \in \mathrm{Pic}^\tau(X)_{\mathbb{R}}$ and nef $L_1, \ldots, L_{n-1} \in \mathrm{Pic}(X)_{\mathbb{Q}}$,*

$$\langle M, M\rangle_{L_1,\ldots,L_{n-1}} \leq 0.$$

Further, if every L_i is ample, then equality holds if and only if $M \in \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{R}}$. When X is a curve,

$$\langle\cdot,\cdot\rangle = -2\langle\cdot,\cdot\rangle_{\mathrm{NT}},$$

where $\langle\cdot,\cdot\rangle_{\mathrm{NT}}$ is the Néron–Tate height pairing on the Jacobian of X.

These results are proven over the next three subsections, with the bulk of the work going into proving Theorem 3.1, with Theorems 3.2 and 3.3 following as corollaries.

3.2. ***Curves.*** We begin when X is a curve. Here we can work directly in $\widehat{\mathrm{Pic}}(X)$ as opposed to $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$. Then the theorem discusses the self-intersection $\overline{M}^2$ when $\deg M = 0$. By Lemma 2.7, M has a flat metric $\overline{M}_0 = (M, \|\cdot\|_0)$.

Let $\overline{N}$ be the vertical line bundle defined by

$$\overline{M} = \overline{M}_0 + \overline{N}.$$

Since $\overline{M}_0$ is flat, $\overline{M}_0 \cdot \overline{N} = 0$ so that

$$\overline{M}^2 = \overline{M}_0^2 + \overline{N}^2 = \overline{M}_0^2 + \sum_{\nu} \overline{N}_{\nu}^2,$$

where $\overline{N}_{\nu}$ is the restriction of $\overline{N}$ to $X_{\nu} := X \otimes_K K_{\nu}$ for each place ν of K (i.e., each closed point of B). Now $\overline{N}_{\nu}^2 \leq 0$ with equality if and only if $\overline{N}_{\nu}$ is constant by the local Hodge index theorem [Yuan and Zhang 2017, Theorem 2.1]. Hence,

$$\sum_{\nu} \overline{N}_{\nu}^2 \leq 0,$$

with equality if and only if $\overline{N} \in \pi^* \widehat{\mathrm{Pic}}(K)$.

Next, we consider $\overline{M}_0^2$. Since M has degree zero, it corresponds naturally to a K-point on the Jacobian, Jac_X, of X. Given any two points $P, Q \in \mathrm{Jac}_X(K)$, let L_P and L_Q be the corresponding algebraically trivial line bundles on X. These each have a flat metric, $\bar{L}_P$ and $\bar{L}_Q$, respectively, unique up to constant metric, and thus we get a well-defined symmetric bilinear pairing

$$(P, Q) \mapsto -\bar{L}_P \cdot \bar{L}_Q$$

on $\mathrm{Jac}_X(K)$, as the intersection does not depend on the choice of flat metric. As is noted in [Faltings 1984] and [Hriljac 1985] in the arithmetic setting, this pairing is exactly the Néron–Tate height pairing. Then the Shioda–Tate theorem [Shioda 1999, Theorem 7] states that this pairing descends to a positive definite pairing on $\mathrm{Jac}_X(K)_{\mathbb{Q}} / \mathrm{Tr}_{K/k}(\mathrm{Jac}_X)(k)$, and

$$\overline{M}^2 = -2h_{\mathrm{NT}}(M).$$

Since $\widehat{\tau}_{K/k}$ produces elements of $\widehat{\mathrm{Pic}}(X)$ with flat metrics, our pairing on

$$\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))$$

matches that considered by Shioda, and this completes the proof of Theorem 3.1 in dimension one. Since Shioda's pairing extends $\mathbb{R}$-linearly, this also proves Theorem 3.3 in dimension one.

We now turn to Theorem 3.2 in dimension one. (1) $\Rightarrow$ (2), as heights are defined using intersections. In particular, fix any model $\mathcal{X} \to B$ for X, and for $x \in X(\overline{K})$ let $\bar{L}_x$ be the model metric corresponding to the Zariski closure of $\tilde{x}$ in $\mathcal{X}$. Then for $\overline{M} \in \widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$ the height $h_{\overline{M}}(x)$ is just $\overline{M} \cdot \bar{L}_x / [K(x) : K] = 0$.

Now suppose $h_{\overline{M}}$ is identically zero on $X(\overline{K})$. Then $\deg M = 0$ as otherwise M or $-M$ is ample and $\overline{M}$ defines an unbounded Weil height. Suppose $\bar{L}$ is a model metric induced by a very ample line bundle $\mathcal{L}$ on a model $\mathcal{X} \to B$. Then, extending K if necessary, by Seidenberg's Bertini theorem [Seidenberg 1950], $\mathcal{L}$ has a section which cuts out a normal, irreducible horizontal subvariety of $\mathcal{X}$. Since $\mathcal{X}$ is a surface, this is just the closure of a point $x_0 \in X(\overline{K})$ on the generic fiber. Then

$$\overline{M} \cdot \bar{L} = \overline{M}|_{x_0} = h_{\overline{M}}(x_0) = 0.$$

Since $\mathrm{Pic}(\mathcal{X})_{\mathbb{Q}}$ is generated by linear combinations of very ample line bundles and $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$ consists of limits of such, this proves (2) $\Rightarrow$ (1).

If $\overline{M}$ is numerically trivial, then $\overline{M} \in \widehat{\mathrm{Pic}}(K) + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$, as necessarily $\overline{M}^2 = 0$. If $\overline{M}_1 \in \widehat{\mathrm{Pic}}(K)$, then $h_{\overline{M}_1}$ is constant, with value equal to $\widehat{\deg}(\overline{M}_1)$. Thus (2) $\Rightarrow$ (3).

Finally, if $\overline{M} \in \pi^*\widehat{\mathrm{Pic}}^0(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$ then $\overline{M}$ is numerically trivial, as $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$ is numerically trivial on every fiber by construction, and the intersection of $\overline{N} \in \pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}}$ with $\bar{L} \in \widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$ is the arithmetic degree of $\overline{N}$ times the degree of L. Then (3) $\Rightarrow$ (1), completing the proof.

3.3. ***Inequality and Cauchy–Schwarz.*** We will now prove the inequality part of Theorem 3.1 by induction on $n = \dim X$, and get a version of the Cauchy–Schwarz inequality as a corollary. As in [Yuan and Zhang 2017, Section 3.3, Assumption (2)], we may assume that each $\bar{L}_i$ is arithmetically positive (instead of just big) by a limiting argument. Thus we may assume L_i is ample.

Since $\overline{M}$ and each $\bar{L}_i$ can be approximated by model metrics, it suffices to prove

$$\mathcal{M}^2 \cdot \mathcal{L}_1 \cdots \mathcal{L}_{n-1} \le 0,$$

under the assumption that $\mathcal{M}$ and every $\mathcal{L}_i$ are line bundles on a model $\mathcal{X}$ for X, that $\mathcal{L}_i$ is ample with respect to k, and that the intersection $\mathcal{M}_K \cdot (\mathcal{L}_1)_K \cdots (\mathcal{L}_{n-1})_K$ on the generic fiber is zero.

Replacing $\mathcal{L}_1$ by a positive tensor power if necessary, we may assume it is very ample. Then by a Bertini-type result of Seidenberg [1950, Theorem 7'], a generic section of $\mathcal{L}_1$ cuts out an integral normal subvariety $\mathcal{Y}$ of $\mathcal{X}$, and we may further stipulate that $\mathcal{Y}$ is horizontal. Then

$$\mathcal{M}^2 \cdot \mathcal{L}_1 \cdots \mathcal{L}_{n-1} = \mathcal{M}|_{\mathcal{Y}}^2 \cdot \mathcal{L}_1|_{\mathcal{Y}} \cdots \mathcal{L}_{n-2}|_{\mathcal{Y}}.$$

This reduces the problem to a lower dimension, but we require that $\mathcal{Y}_K$ have a K point to conclude the result by induction. This is certainly true if we replace K with a finite extension K', or equivalently replace B with a finite cover. Since intersection numbers simply scale by $[K' : K]$ and the subgroup $\widehat{\mathrm{Pic}}(K) + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))$ is equal to $\widehat{\mathrm{Pic}}(K') + \mathrm{Tr}_{K'/k}(\mathrm{Pic}^0(X_{K'}))$ intersected with $\widehat{\mathrm{Pic}}(X)$, such a base change is permissible.

Given $M \in \mathrm{Pic}^\tau(X)_\mathbb{R}$, we can write it as an $\mathbb{R}$-linear combination of numerically trivial line bundles on X, and each has a flat metric by Lemma 2.7. Then the inequality of Theorem 3.1 immediately implies the inequality of Theorem 3.3 when every L_i is big. If L_i is merely nef, choose any ample line bundle A and $\epsilon > 0$, and then $L_{i_\epsilon} := L_i + \epsilon A$ is big. Thus the inequality holds with $L_{i,\epsilon}$ replacing L_i, and taking the limit as $\epsilon \to 0$, it holds in general.

As a corollary, we have the following Cauchy–Schwarz inequality:

Corollary 3.4. *Let $\overline{M}$ and $\overline{N}$ be two integral adelic line bundles on X, and let $\bar{L}_1, \dots, \bar{L}_{n-1}$ be nef adelic line bundles on X such that*

$$M \cdot L_1 \cdots L_{n-1} = N \cdot L_1 \cdots L_{n-1} = 0.$$

Then

$$(\overline{M} \cdot \overline{N} \cdot \bar{L}_1 \cdots \bar{L}_{n-1})^2 \le (\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1})(\overline{N}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1}).$$

Proof. This follows from the inequality part of the Hodge index theorem proven above, and from the standard proof of the Cauchy–Schwarz inequality using the (negative semidefinite) inner product

$$\langle M, N \rangle_{\bar{L}_1,\dots,\bar{L}_{n-1}} := \overline{M} \cdot \overline{N} \cdot \bar{L}_1 \cdots \bar{L}_{n-1}. \qquad \square$$

3.4. ***Equality.*** We now proceed to the equality part of Theorem 3.1. To prove the "if" direction, suppose $\overline{M} \in \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_\mathbb{Q}$. Then $\overline{M}$ is numerically trivial, as it is numerically trivial on every fiber by construction. If $\overline{M} \in \pi^*\widehat{\mathrm{Pic}}(K)_\mathbb{Q}$, then $\overline{M}^2$ consists of self-intersections of whole fibers, which are equal to zero.

To prove "only if," suppose that each $\bar{L}_i$ is arithmetically positive, that $\overline{M}$ is $\bar{L}_i$-bounded for all i, and that

$$\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = 0.$$

Note that as a consequence of the Cauchy–Schwarz inequality above, the set of metrized line bundles $\overline{M}$ satisfying these properties forms a group via tensor products.

By [Yuan and Zhang 2017, Lemma 3.7] (this requires that $\bar{L}_i$ is arithmetically positive), M is numerically trivial on X. Thus it has a flat metric; let $\overline{M}_0 = (M, \|\cdot\|)$ be flat. Then, similar to the curve case, $\overline{N} := \overline{M} - \overline{M}_0$ is vertical, and

$$\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = \overline{M}_0^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} + \overline{N}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1}.$$

The inequality part of the Hodge index theorem guarantees that both terms on the right are zero, and then by the local Hodge index theorem at every place occurring in $\overline{N}$, we have $\overline{N} \in \widehat{\mathrm{Pic}}(K)_\mathbb{Q}$. Hence we are reduced to proving the statement in the flat metric case $\overline{M} = \overline{M}_0$.

We again replace $\bar{L}_1$ by a positive multiple to assume that L_1 is very ample, then apply Seidenberg's Bertini theorem to conclude that $(L_1)_{\overline{K}}$ has a section s which cuts out an integral, normal subvariety Y. Such Y is defined over some finite extension of K'/K, and thus after a base change from K to K', we may assume that L_1 has a section which cuts out a geometrically integral and geometrically normal subvariety Y. As in the proof of the inequality, this finite extension merely scales the intersection numbers by a positive factor. We thus continue writing K, assuming it has been made large enough, to avoid excessive additional notation.

Lemma 3.5. *If $\overline{M}$ is flat, and Y is a geometrically normal subvariety of X, then*

$$\overline{M}|_Y^2 \cdot \bar{L}_2|_Y \cdots \bar{L}_{n-1}|_Y = 0.$$

Proof. By the induction formula of Chambert-Loir [2006], recalled in Section 2.1,

$$\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = \overline{M}|_Y^2 \cdot \bar{L}_2|_Y \cdots \bar{L}_{n-1}|_Y - \sum_v \int_{X_v^{\mathrm{an}}} \log\|s\|_v c_1(\overline{M})^2 c_1(\bar{L}_2) \cdots c_1(\bar{L}_{n-1}).$$

Since $\overline{M}$ is flat, all the integrals are zero. □

Thus we may assume

$$\overline{M}|_Y \in \pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(Y))_{\mathbb{Q}}$$

by induction.

Write $\overline{M}|_Y = \overline{M}' + \pi^*\overline{M}_1$, with $\overline{M}' \in \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(Y))_{\mathbb{Q}}$ and $\overline{M}_1 \in \widehat{\mathrm{Pic}}(K)_{\mathbb{Q}}$. Then define $\overline{M}_2 = \overline{M} - \overline{\pi}^*\overline{M}_1$. Since M is numerically trivial, replacing $\overline{M}$ by a positive integer multiple if necessary, we may further assume M is algebraically trivial, and then that π^*M_1, $M_2 \in \mathrm{Pic}^0(X)$.

As noted earlier, if we drop the metric structure the map $\widehat{\tau}$ is simply the K/k-trace map on field-valued points. The following lemma then proves that $M_2|_Y = M'$ lifts via the pullback of $Y \hookrightarrow X$ to an element of $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$.

Lemma 3.6. *Let $f : A \to B$ be a morphism of abelian varieties defined over K. In the commutative diagram*

$$\begin{array}{ccc} \mathrm{Tr}(A)(k)_{\mathbb{Q}} & \xrightarrow{\tau_A} & A(K)_{\mathbb{Q}} \\ \downarrow{\scriptstyle f_{\mathrm{Tr}}} & & \downarrow{\scriptstyle f} \\ \mathrm{Tr}(B)(k)_{\mathbb{Q}} & \xrightarrow{\tau_B} & B(K)_{\mathbb{Q}} \end{array}$$

$(f \circ \tau_A)(\mathrm{Tr}(A)(k)_{\mathbb{Q}})$ *is equal to* $f(A(K)_{\mathbb{Q}}) \cap \tau_B(\mathrm{Tr}(B)(k)_{\mathbb{Q}})$.

Proof. To shorten notation, we will drop writing the map τ_A and consider $\mathrm{Tr}(A)(k)$ directly as a subgroup of $A(K)$ (and similarly for B). First reduce to the case where f is surjective: let B' be the image of f, an abelian subvariety of B. By Poincaré reducibility, B is isogenous to $B' \times B''$, for some abelian variety B''. Then $\mathrm{Tr}(B)$ is isogenous to $\mathrm{Tr}(B') \times \mathrm{Tr}(B'')$, and the intersection of $\mathrm{Tr}(B')(k) \times \mathrm{Tr}(B'')(k)$ with $B'(K)$ is just $\mathrm{Tr}(B')(k)$.

Now assume f is surjective. By [Conrad 2006, Theorem 6.4], $\mathrm{Tr}(A)_K$ is isogenous to an abelian subvariety $A' \subset A$ such that $\mathrm{Tr}(A) \cong \mathrm{Tr}(A')$ and $\mathrm{Tr}(A/A') = 0$. Similarly, B has an abelian subvariety B' with the same properties. Composing with these isogenies, we get a surjection

$$\mathrm{Tr}(A)_K \times (A/A') \twoheadrightarrow \mathrm{Tr}(B)_K \times (B/B')$$

where the map on the first component is f descended to the traces. Now consider the map $A/A' \to \mathrm{Tr}(B)_K$ obtained from the above map composed with projection onto the first component. This map must factor through $\mathrm{Im}(A/A')_K$, which is trivial as $\mathrm{Im}(A/A')$ is isogenous to $\mathrm{Tr}(A/A') = 0$. Thus $\mathrm{Tr}(A)_K \to \mathrm{Tr}(B)_K$ is surjective, and we get a surjection $\mathrm{Tr}(A)(k) \twoheadrightarrow \mathrm{Tr}(B)(k)$, proving the lemma. □

Hence we may lift $\overline{M}_2|_Y$ to an element $\overline{M}'_2 \in \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$, and we must have

$$\overline{M}_2 - \overline{M}'_2 \in \ker\bigl(\widehat{\mathrm{Pic}}(X) \to \widehat{\mathrm{Pic}}(Y)\bigr).$$

Since $\mathrm{Pic}^0(X) \to \mathrm{Pic}^0(Y)$ has finite kernel [Kleiman 2005, Remark 9.5.8], replacing M with a positive integer multiple, we may assume $M_2 - M'_2 = \mathcal{O}_X$ and thus $\overline{M}_2 - \overline{M}'_2$ is vertical. Additionally, by the Cauchy–Schwarz inequality, Corollary 3.4,

$$(\overline{M}_2 - \overline{M}'_2)^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = (\overline{M} - \pi^*\overline{M}_1 - \overline{M}'_2)^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = 0,$$

so that by the local Hodge index theorem the metric must be constant at each place and $\overline{M}_2 - \overline{M}'_2 \in \pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}}$. Note that the local Hodge index theorem requires that $\overline{M}_2 - \overline{M}'_2$ be $\bar{L}_i$-bounded, but this holds, as $\overline{M}$ is $\bar{L}_i$-bounded by hypothesis, and all of $\pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$ is $\bar{L}_i$-bounded as well. This means that

$$\overline{M} = (\pi^*\overline{M}_1 + \overline{M}_2 - \overline{M}'_2) + \overline{M}'_2 \in \pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}.$$

This proves that when $\overline{M}$ is $\bar{L}_i$-bounded and $\bar{L}_i$ is arithmetically positive for all i, then

$$\overline{M}^2 \cdot \bar{L}_1 \cdots \bar{L}_{n-1} = 0$$

if and only if $\overline{M} \in \pi^*\widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$, which completes the proof of Theorem 3.1.

The inequality part of Theorem 3.3 is implied immediately by the inequality of Theorem 3.1 provided each L_i is big. To accomplish this, chose an ample line bundle A on X, and define $L_{i,\epsilon} := L_i + \epsilon A$ for $\epsilon > 0$. Then extending these to nef metrics $\bar{L}_{i,\epsilon}$, we have

$$\langle M, M\rangle_{L_{1,\epsilon},\dots,L_{n-1,\epsilon}} = \overline{M}^2 \cdot \bar{L}_{1,\epsilon} \cdots \bar{L}_{n-1,\epsilon} \le 0,$$

and the result follows letting $\epsilon \to 0$.

To prove the equality of Theorem 3.3, again split $M \in \mathrm{Pic}^{\tau}(X)_{\mathbb{R}}$ into an $\mathbb{R}$-linear combination of numerically trivial line bundles. Using the inequality, the equality can be proven for each of these individually. L_1 is ample, so Lemma 3.5 applies, and then by the induction hypothesis $M|_Y \in \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(Y))_{\mathbb{R}}$. Then by Lemma 3.6 we conclude $M \in \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{R}}$.

Finally, we prove Theorem 3.2. If $\overline{M}$ is numerically trivial it is flat by Theorem 3.1, and then its restriction to any geometrically normal subvariety is also numerically trivial by the proof of Lemma 3.5. Thus (1) $\Rightarrow$ (2) follows from the dimension one case, as we can compute the height of a point on any curve passing through that point.

To show (2) $\Rightarrow$ (3), assume $h_{\overline{M}}$ is trivial on $X(\overline{K})$, and chose a curve $C \subset X$. Since the height is trivial on all of C, we have $\overline{M}|_C \in \pi^*\widehat{\mathrm{Pic}}^0(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(C))_{\mathbb{Q}}$, as was proven earlier for curves. Then by the induction argument above,

$$\overline{M} \in \pi^*\widehat{\mathrm{Pic}}^0(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}.$$

We have established previously that $\pi^*\widehat{\mathrm{Pic}}^0(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$ is numerically trivial, so (3) $\Rightarrow$ (1).

4. Algebraic dynamical systems

As before, K is the function field of a smooth projective curve B over an algebraically closed field k, and let X be a projective variety over K. Suppose (X, f, L) and (X, g, M) are two polarized dynamical systems on X, so that f and g are endomorphisms of X, and L and M are ample line bundles such that $f^*L \cong L^q$ and $g^*M \cong M^r$ for some $q, r > 1$.

Remark. If X is not normal, we may replace X by its normalization $\psi : X' \to X$, replace f by the normalization $f' : X' \to X'$ of $f \circ \psi$, and replace L by $L' = \psi^* L$ to get a new polarized algebraic dynamical system (X', f', L') with $\mathrm{Prep}(f') = \psi^{-1}\,\mathrm{Prep}(f)$, and similarly for (X, g, M). By first replacing K with an extension if necessary, we may further assume that the normalization is geometrically normal. Hence from here on out we assume without loss of generality that X is geometrically normal.

Our main goal in this section is to prove a comparison theorem for the points with dynamical height 0 under f and g, with an important corollary comparing the preperiodic points of f and g when k is the algebraic closure of a finite field. We begin with general properties of polarized algebraic dynamical systems, then define the particular arithmetic dynamical heights involved before stating the theorem.

4.1. ***An f^*-splitting of the Néron–Severi sequence.*** We first show that the projection from $\mathrm{Pic}(X)$ onto the Néron–Severi group has a unique f^* equivariant section.

The pullback f^* preserves the exact sequence

$$0 \to \mathrm{Pic}^0(X) \to \mathrm{Pic}(X) \to \mathrm{NS}(X) \to 0,$$

defining the Néron–Severi group $\mathrm{NS}(X)$, and the Néron–Severi theorem [SGA 6 1971, Exposé XII, Théorème 5.1, p. 650] tells us that $\mathrm{NS}(X)$ is a finitely generated $\mathbb{Z}$-module. For arbitrary k, the $\mathbb{Z}$-module $\mathrm{Pic}^0(X)$ need not be finitely generated, but by the Lang–Néron theorem [1959],

$$\mathrm{Pic}^0(X) \big/ \mathrm{Tr}_{K/k} \mathrm{Pic}^0(X) \cong \mathrm{Pic}^0(X) \big/ \mathrm{Pic}^0(\mathrm{Im}_{K/k}(\mathrm{Alb}(X)))$$

is a finitely generated $\mathbb{Z}$-module. To shorten our notation, define

$$\mathrm{Pic}^0_{\mathrm{tr}}(X) := \mathrm{Pic}^0(X) \big/ \mathrm{Tr}_{K/k} \mathrm{Pic}^0(X),$$

$$\mathrm{Pic}_{\mathrm{tr}}(X) := \mathrm{Pic}(X) \big/ \mathrm{Tr}_{K/k} \mathrm{Pic}^0(X),$$

so that we have an exact sequence of finite-dimensional $\mathbb{C}$-vector spaces

$$0 \to \mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{C}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{C}} \to \mathrm{NS}(X)_{\mathbb{C}} \to 0,$$

which is also an exact sequence of f^*-modules.

Lemma 4.1. *The operator f^* is semisimple on $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{C}}$ with eigenvalues of absolute value $q^{1/2}$, and is semisimple on $\mathrm{NS}(X)$ with eigenvalues of absolute value q.*

Proof. As usual, let $n = \dim X$. By the classical Hodge index theorem [SGA 6 1971, Exposé XIII, Corollaire 7.4], we can decompose $\mathrm{NS}(X)_{\mathbb{R}}$ as

$$\mathrm{NS}(X)_{\mathbb{R}} := \mathbb{R}L \oplus P(X), \qquad P(X) := \{\xi \in \mathrm{NS}(X)_{\mathbb{R}} : \xi \cdot L^{n-1} = 0\},$$

and define a negative definite pairing on $P(X)$ by

$$\langle \xi_1, \xi_2 \rangle := \xi_1 \cdot \xi_2 \cdot L^{n-2}.$$

The projection formula for intersection numbers applied to L^n gives us $\deg f = q^n$, and then applied to this pairing, we have

$$\langle f^*\xi_1, f^*\xi_2 \rangle = q^2 \langle \xi_1, \xi_2 \rangle.$$

Hence $\frac{1}{q} f^*$ is orthogonal with respect to this pairing, and $\frac{1}{q} f^*$ is diagonalizable on $\mathrm{NS}(X)_{\mathbb{C}}$ with eigenvalues all of absolute value 1.

On $\mathrm{Pic}^0(X)_{\mathbb{R}}$ we can define a pairing as follows: for $\xi_1, \xi_2 \in \mathrm{Pic}^0(X)_{\mathbb{R}}$, let $\overline{\xi}_1$ and $\overline{\xi}_2$ be flat metrized extensions, and let $\bar{L}$ be any integrable adelic line bundle extending L. Then define

$$\langle \xi_1, \xi_2 \rangle := \overline{\xi}_1 \cdot \overline{\xi}_2 \cdot \bar{L}^{n-1}.$$

It follows from Corollary 2.10 that this pairing does not depend on the choice of metrics. Since $\mathrm{Tr}_{K/k} \mathrm{Pic}^0(X)$ is numerically trivial, this pairing descends to $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{R}}$, and by Theorem 3.1, it is negative definite on this quotient.

Again applying the projection formula,

$$(f^*\overline{\xi}_1) \cdot (f^*\overline{\xi}_2) \cdot (f^*\bar{L})^{n-1} = q^n(\overline{\xi}_1 \cdot \overline{\xi}_2 \cdot \bar{L}^{n-1}),$$

since each $f^*\overline{\xi}_i$ is still flat. We may also replace $f^*\bar{L}$ by $\bar{L}^q$ because the pairing is independent of the choice of metric on L, and have

$$\langle f^*\xi_1, f^*\xi_2 \rangle = q \langle \xi_1, \xi_2 \rangle.$$

Hence, $q^{-\frac{1}{2}} f^*$ is orthogonal on $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{R}}$ with respect to the negative of this pairing, making it diagonalizable with eigenvalues of absolute value 1 as a transformation on $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{C}}$. □

By the theorem,

$$0 \to \mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{C}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{C}} \to \mathrm{NS}(X)_{\mathbb{C}} \to 0$$

has a unique splitting as f^*-modules by a section

$$\ell_f : \mathrm{NS}(X)_{\mathbb{C}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{C}}.$$

Let $P, Q \in \mathbb{Q}[T]$ be the minimal polynomials of f^* on $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{Q}}$ and $\mathrm{NS}(X)_{\mathbb{Q}}$ respectively. Because the eigenvalues of f^* are different on $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{Q}}$ and $\mathrm{NS}(X)_{\mathbb{Q}}$, we see that P and Q are coprime, and $R := PQ$ is the minimal polynomial of f^* on $\mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$. Define

$$\mathrm{Pic}_{\mathrm{tr},f}(X)_{\mathbb{Q}} := \ker Q(f^*)|_{\mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}}$$

and then this splitting can be given over $\mathbb{Q}$ as

$$\ell_f : \mathrm{NS}(X)_{\mathbb{Q}} \xrightarrow{\sim} \mathrm{Pic}_{\mathrm{tr},f}(X)_{\mathbb{Q}} \hookrightarrow \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}.$$

4.2. *Admissible metrics.* Adding to the notation above, define

$$\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)_{\mathbb{Q}} := \widehat{\mathrm{Pic}}(X)_{\mathbb{Q}} / \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}.$$

Theorem 4.2. *The projection* $\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)_{\mathbb{Q}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$ *has a unique section* $M \mapsto \overline{M}_f$ *as* f^**-modules, satisfying*:

(1) *If* $M \in \mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{Q}}$ *then* $\overline{M}_f$ *is flat.*

(2) *If* $M \in \mathrm{Pic}_{\mathrm{tr},f}(X)_{\mathbb{Q}}$ *is ample then* $\overline{M}_f$ *is nef.*

Adelic metrized line bundles of the form $\overline{M}_f$ *are called* f*-admissible.*

Remark. Since $\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}} \subset \widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$ is flat and numerically trivial, and the underlying line bundles in $\mathrm{Pic}(X)$ are also numerically trivial, the notions of ampleness, nefness, and flatness are all well defined modulo the trace. While $\overline{M}_f$ represents a coset of the trace in $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$ instead of a single metrized line bundle, all coset representatives will produce the same height functions and intersection numbers, by Theorem 3.2.

Proof. Define $\widehat{\mathrm{Pic}}(X)'$ to be the group of adelic line bundles on X with continuous (but not necessarily integrable) metrics, and $\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)' := \widehat{\mathrm{Pic}}(X)'/\,\mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))$. This contains $\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)$. We will show that if the projection $\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$ has a unique section, then properties (1) and (2) of the theorem hold for this section. Since $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{Q}}$ and the ample classes in $\mathrm{Pic}_{\mathrm{tr},f}(X)_{\mathbb{Q}}$ generate $\mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$, the section does in fact produce integrable metrics, proving the theorem.

The kernel of the projection $\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$ is

$$D(X) = \widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} \bigoplus_v C(X_v^{\mathrm{an}}),$$

where $C(X_v^{\mathrm{an}})$ is the ring of continuous $\mathbb{R}$-valued functions on X_v^{an}, via the association $\|\cdot\|_v \to -\log\|1\|_v$. Recall that $R = PQ$ was defined to be the minimal polynomial of f^* on $\mathrm{Pic}(X)_{\mathbb{Q}}$ and now consider the action of $R(f^*)$ on $D(X)$.

Lemma 4.3. *$R(f^*)$ is invertible on $D(X)$.*

Proof. The pullback f^* acts as the identity on $\widehat{\mathrm{Pic}}(K)$, hence $R(f^*)$ acts as $R(1)$, and this is not zero because the roots of R all have absolute value q or $q^{1/2}$. So it suffices to show that $R(f^*)$ is invertible on $C(X)_{\mathbb{C}} := \left(\bigoplus_v C(X_v^{\mathrm{an}})\right) \otimes_{\mathbb{R}} \mathbb{C}$. Factor R over $\mathbb{C}$ as

$$R(T) = a\prod_i \left(1 - \frac{T}{\lambda_i}\right),$$

where $a \neq 0$, and by Lemma 4.1, $|\lambda_i|$ is either $q^{1/2}$ or q. $R(f^*)$ is invertible provided each term $1 - f^*/\lambda_i$ is, and each term has inverse

$$\left(1 - \frac{f^*}{\lambda_i}\right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{f^*}{\lambda_i}\right)^k,$$

provided this series converges absolutely with respect to the operator norm, which is defined with respect to the supremum norm $\|\cdot\|_{\sup}$ on $C(X_v^{\mathrm{an}})_{\mathbb{C}}$ for every place v. The pullback f^* does not change the supremum norm, so the operator norm of f^* is 1, and

$$\left\| \left(\frac{f^*}{\lambda_i}\right)^k \right\| = \frac{1}{|\lambda_i|^k} \le q^{-\frac{k}{2}},$$

so the series converges absolutely. □

Corollary 4.4. *The exact sequence*

$$0 \to D(X) \to \widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}} \to 0$$

has a unique f^*-equivariant splitting.

Proof. Define

$$E(X) := \ker\bigl(R(f^*) : \widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}} \to \widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}}\bigr).$$

Since $R(f^*)$ kills all of $\mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$, this gives an f^*-invariant decomposition

$$\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}} = D(X) \bigoplus E(X)$$

such that the projection onto $\mathrm{Pic}_{\mathrm{tr}}(X)$ gives an isomorphism $E(X) \xrightarrow{\sim} \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$, whose inverse is the desired splitting.

We can write this down even more explicitly. For $M \in \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}}$, let $\overline{M}$ be any choice of metric in $\widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'_{\mathbb{Q}}$. Then define

$$\overline{M}_f := \overline{M} - R(f^*)|^{-1}_{D(X)} R(f^*)\overline{M}. \qquad \square$$

It now remains to show that this splitting satisfies (1) and (2). To start, suppose M is in $\mathrm{Pic}^0_{\mathrm{tr}}(X)_{\mathbb{Q}}$. After extending K if necessary, we can find a preperiodic point $x_0 \in X(K)$ (in fact, by [Fakhruddin 2003], $\mathrm{Prep}(f)$ is dense in $X(\overline{K})$), and by replacing f with an iterate we may assume that x_0 is a fixed point. Let $i : X \to \mathrm{Alb}(X)$ be the Albanese map taking $x_0 \mapsto 0$, then f^* and i^* induce the following commutative diagram, where $f' := (f^*)^\vee$:

$$\begin{array}{ccc}
\mathrm{Pic}^0_{\mathrm{tr}}(\mathrm{Alb}(X)) & \underset{\sim}{\overset{i^*}{\rightleftarrows}} & \mathrm{Pic}^0_{\mathrm{tr}}(X) \\
{\scriptstyle (f')^*}\downarrow & & \downarrow{\scriptstyle f^*} \\
\mathrm{Pic}^0_{\mathrm{tr}}(\mathrm{Alb}(X)) & \underset{\sim}{\overset{i^*}{\rightleftarrows}} & \mathrm{Pic}^0_{\mathrm{tr}}(X) \\
{\scriptstyle M \mapsto \overline{M}_{f'}}\downarrow & & \downarrow{\scriptstyle M \mapsto \overline{M}_f} \\
\widehat{\mathrm{Pic}}_{\mathrm{tr}}(\mathrm{Alb}(X))' & \xrightarrow{i^*} & \widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)'
\end{array}$$

Because this commutes, it suffices to show (1) for abelian varieties, as i^* takes $M_{f'}$ to $\overline{M}_f$, and the pullback of a flat metric is also flat. Now $[2]^*M = 2M$, and since $[2]$ commutes with f',

$$[2]^*\overline{M}_{f'} = 2\overline{M}_{f'},$$

so that as in the proof of Lemma 2.7, we have that $\overline{M}_{f'}$, and hence also $\overline{M}_f$ is flat.

Finally, we show that (2) also holds. This is proven when K is a number field in [Yuan and Zhang 2017, Theorem 4.9], however the proof works identically in our geometric setting, as it only relies on the fact that $\mathrm{Pic}_f(X)_{\mathbb{Q}}$ (here $\mathrm{Pic}_{\mathrm{tr},f}(X)_{\mathbb{Q}}$) is a finite-dimensional $\mathbb{Q}$-vector space on which the operator $q^{-1}f^*$ has eigenvalues with absolute value one. □

Thus, we have an f^*-equivariant linear map

$$\widehat{\ell}_f : \mathrm{NS}(X)_{\mathbb{Q}} \to \widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)_{\mathbb{Q}}$$

given by the composition of the section developed in Theorem 4.2 and the map just preceding it. Importantly, we can think of this as a map into $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$ which is well defined up to a numerically trivial factor, and thus sufficient to specify heights and intersections. Given $M \in \mathrm{Pic}(X)_{\mathbb{Q}}$, we will write $\overline{M}_f$ to mean any lift of the image of M under $\mathrm{Pic}(X)_{\mathbb{Q}} \to \mathrm{Pic}_{\mathrm{tr}}(X)_{\mathbb{Q}} \to \widehat{\mathrm{Pic}}_{\mathrm{tr}}(X)_{\mathbb{Q}}$ to $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$.

4.3. *Rigidity of height zero points and preperiodic points.* Heights given by f-admissible metrized line bundles have particularly nice properties and correspond to the dynamical canonical heights defined by Call and Silverman [1993].

Proposition 4.5. *Let $M \in \mathrm{Pic}(X)_{\mathbb{Q}}$. Then*:

(1) *If $f^*M = M^{\lambda}$ for some $\lambda \in \mathbb{Q}$, then $f^*\overline{M}_f = \overline{M}_f^{\lambda}$ in $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$, and*

$$h_{\overline{M}_f}(f(\cdot)) = \lambda h_{\overline{M}_f}(\cdot).$$

(2) *For $x \in \mathrm{Prep}(f)$, $\overline{M}_f|_x$ is trivial on $\widehat{\mathrm{Pic}}(x)_{\mathbb{Q}}$, and in particular $h_{\overline{M}_f}$ is zero on* $\mathrm{Prep}(f)$.

Further, if M is ample and $f^*M = \lambda M$ for some $\lambda > 1$ (in particular, if $M = L$), then

(3) *$h_{\overline{M}_f}(x) \geq 0$ for all $x \in X(K)$, and*

(4) *if k is finite, $h_{\overline{M}_f}(x) = 0$ if and only if $x \in \mathrm{Prep}(f)$.*

Call and Silverman [1993] establish that our height agrees with the dynamical canonical height $\widehat{h}_f$, and then the above properties all follow from well-known properties of dynamical heights proven in [loc. cit.].

We can now state and prove our main theorem of this section.

Theorem 4.6. *Let (f, L) and (g, M) be two polarized algebraic dynamical systems on X. Define $Z_f := \{x \in X(\overline{K}) | h_{\overline{L}_f}(x) = 0\}$ to be the set of height zero points with respect to $\overline{L}_f$, and Z_g the set of height zero points with respect to $\overline{M}_g$, and let Z be the Zariski closure of $Z_f \cap Z_g$ in X. Then*

$$Z_f \cap Z(\overline{K}) = Z_g \cap Z(\overline{K}).$$

When k is finite, $Z_f = \mathrm{Prep}(f)$ and $Z_g = \mathrm{Prep}(g)$, so Corollary 1.3 stated in the introduction follows as an immediate consequence. If k is not finite, it is still true that $Z_f \supseteq \mathrm{Prep}(f)$, but there may be height zero points with infinite forward orbit. See Section 5 for further discussion.

Proof. We begin by proving a simpler lemma, justifying the notation that Z_f does not depend on the polarization L.

Lemma 4.7. *Let $f : X \to X$, and let L and M be two ample line bundles which polarize f. Then*

$$\{x \in X(\bar{K}) \mid h_{\bar{L}_f}(x) = 0\} \text{ is equal to } \{x \in X(\bar{K}) \mid h_{\overline{M}_f}(x) = 0\},$$

and we unambiguously call both sets Z_f.

Proof. Since L is ample, there exists a constant $c > 0$ such that $cL - M$ is also ample. Then by Proposition 4.5, the canonical heights $h_{\overline{M}_f}$ and $h_{\overline{cL}_f} = ch_{\bar{L}_f}$ are related by

$$0 \le h_{\overline{M}_f}(x) \le ch_{\bar{L}_f}(x)$$

for all $x \in X(\bar{K})$. Thus

$$\{x \in X(\bar{K}) | h_{\bar{L}_f}(x) = 0\} \subseteq \{x \in X(\bar{K}) | h_{\overline{M}_f}(x) = 0\}.$$

By symmetry, we also have containment in the other direction. □

We now prove the theorem.

Let Y be the normalization of an irreducible component of Z, assume K is replaced by a finite extension if necessary so that Y is geometrically normal, and say $\dim Y = d$. Let ξ be the image of L in $\mathrm{NS}(X)$. Then ξ has two different lifts $\widehat{\ell_f}(\xi)$ and $\widehat{\ell_g}(\xi)$ to $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}} / \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}$, and we can pick representatives $\bar{L}_f$ and $\bar{L}_g$ in $\widehat{\mathrm{Pic}}(X)_{\mathbb{Q}}$. By Theorem 4.2, $\bar{L}_f$ and $\bar{L}_g$ are both nef, and are f- and g-admissible, respectively. Since L, L_g, and L_f are all in the same numerical equivalence class in $\mathrm{Pic}(X)_{\mathbb{Q}}$, all are ample.

Their sum $\overline{N} := \bar{L}_f + \bar{L}_g$ is also nef, and defines a height function $h_{\overline{N}}$, which does not depend on the choice of representatives of cosets modulo the trace.

By Lemma 4.7 and the premise that $Z_f \cap Z_g \cap Z(\bar{K})$ is dense, Y has a dense set of points which have height zero under $h_{\overline{N}}$. By the successive minima (see Proposition 2.13),

$$\lambda_1(Y, \overline{N}) = h_{\overline{N}}(Y) = 0.$$

Rewriting the height of Y in terms of intersections,

$$0 = (\bar{L}_f|_Y + \bar{L}_g|_Y)^{d+1} = \sum_{i=0}^{d+1} \binom{d+1}{i} (\bar{L}_f|_Y)^i \cdot (\bar{L}_g|_Y)^{d+1-i}.$$

Since both $\bar{L}_f$ and $\bar{L}_g$ are nef, every term in the sum on the right is nonnegative, hence all must be zero. Then

$$(\bar{L}_f|_Y - \bar{L}_g|_Y)^2 \cdot (\bar{L}_f|_Y + \bar{L}_g|_Y)^{d-1} = 0,$$

as well. Because $L_f - L_g$ is zero in the Néron–Severi group, and thus numerically trivial, we also have

$$(L_f|_Y - L_g|_Y)\cdot(L_f|_Y + L_g|_Y)^{d-1} = 0.$$

Additionally, $(\bar{L}_f - \bar{L}_g)$ is clearly $(\bar{L}_f + \bar{L}_g)$-bounded, and we are nearly in the right setting to apply Theorem 3.1, except that $(\bar{L}_f + \bar{L}_g)$ is nef, but not necessarily arithmetically positive.

To fix this, we simply adjust the metric by a small positive factor: let $\bar{C} \in \widehat{\mathrm{Pic}}(K)$ with $\widehat{\deg}(\bar{C}) > 0$. Replace the pair $(\bar{L}_f - \bar{L}_g, \bar{L}_f + \bar{L}_g)$ by $(\bar{L}_f - \bar{L}_g, \bar{L}_f + \bar{L}_g + \bar{\pi}^* C)$. Since $L_f - L_g$ is numerically trivial, the metric on $\bar{L}_f - \bar{L}_g$ is flat, so adding $\bar{\pi}^* C$, which is vertical, does not change the intersection number. All the conditions of the theorem are now satisfied, so that the theorem tells us

$$(\bar{L}_f - \bar{L}_g) \in \widehat{\mathrm{Pic}}(K)_{\mathbb{Q}} + \mathrm{Tr}_{K/k}(\mathrm{Pic}^0(X))_{\mathbb{Q}}.$$

We therefore conclude by Theorem 3.2 that $h_{\bar{L}_f} - h_{\bar{L}_g}$ is a constant height function on Y. Since these two heights both take value zero on a dense set in Z, they must be equal on Y. Thus these heights define the same sets of height zero points, and then by Lemma 4.7, Z_f and Z_g agree on Y, and hence on all of Z. □

5. Related results and further questions

5.1. ***Rigidity of preperiodic points over global function fields.*** We first summarize some basic consequences of Theorem 4.6 when K is a global function field, particularly in the case when $\mathrm{Prep}(f) \cap \mathrm{Prep}(g)$ is dense in X.

Lemma 5.1. *Let K be a global function field, and let f and g be two polarized algebraic dynamical systems on a projective variety X. Then the following are equivalent*:

(1) $\mathrm{Prep}(f) = \mathrm{Prep}(g)$.

(2) $\mathrm{Prep}(f) \cap \mathrm{Prep}(g)$ *is dense in* X.

(3) $\mathrm{Prep}(f) \subset \mathrm{Prep}(g)$.

(4) $g(\mathrm{Prep}(f)) \subset \mathrm{Prep}(f)$.

Proof. The equivalence of (1) and (2) is an immediate consequence of Theorem 4.6 and the fact that over a global function field, all dynamical height zero points are preperiodic. Clearly (1) implies (4). By Fakhruddin [2003], $\mathrm{Prep}(f)$ is always dense in X, hence (3) implies (2). We now show (4) implies (3).

Stratify $\mathrm{Prep}(f)$ by degree, writing

$$\mathrm{Prep}(f) = \bigcup_{d \geq 0} \mathrm{Prep}(f, d),$$

where

$$\mathrm{Prep}(f, d) := \{x \in \mathrm{Prep}(f) \mid [K(x) : K] \leq d\}.$$

Since each $\mathrm{Prep}(f, d)$ has height zero and bounded degree, it is finite. Now (4) says that g fixes $\mathrm{Prep}(f)$, but since g is defined over K, it fixes each $\mathrm{Prep}(f, d)$ as well. Thus every point of $\mathrm{Prep}(f)$ has finite forward orbit under g. □

This lemma suggests two related questions which we do not answer here.

(1) When is $\mathrm{Prep}(f)$ equal to $\mathrm{Prep}(g)$?

(2) If $\mathrm{Prep}(f) = \mathrm{Prep}(g)$, how closely related must f and g be?

In the case of $f : \mathbb{P}^1 \to \mathbb{P}^1$, Mimar [2013] gives a variety of partial answers to these questions, with the general implication being that if f and g have the same preperiodic points, their Julia sets must also be very similar. But this is likely very difficult in dimension greater than one.

5.2. ***Preperiodic points over larger function fields.*** Theorem 3.1 and most of the proof of Theorem 4.6 hold over all transcendence degree one function fields, not just global function fields. But because the Northcott principal fails when k is not a finite field or the algebraic closure of a finite field, we cannot equate height zero points with preperiodic points over arbitrary function fields, and thus Theorem 4.6 is a statement about height zero points and not preperiodic points. In this broader setting, however, some things can still be said.

Baker [2009] proves the following theorem, first proven by Benedetto [2005] in the case of polynomials.

Theorem 5.2. *Let $f : \mathbb{P}^1_K \to \mathbb{P}^1_K$ be a rational function of degree ≥ 2, and suppose that f is not isotrivial, in the sense that there exists no finite extension K' of K and Möbius transformation $M \in \mathrm{PGL}_2(K')$ such that*

$$f' := M^{-1} \circ f \circ M$$

is defined over k. Then

$$\mathrm{Prep}(f) = Z_f.$$

Thus Theorem 4.6 proven here immediately implies Corollary 1.4.

In higher-dimension isotriviality is less straightforward to classify. When A is an abelian variety, its K/k-trace classifies how isotrivial it is, and then the Lang–Néron theorem provides a Northcott-like result for the Néron–Tate canonical height (the dynamical height induced by $[n]$): height zero points fall into only finitely many cosets of $\mathrm{Tr}_{K/k}(A)(k) \hookrightarrow A(K)$.

There is no notion of a trace for general varieties, however, and $\iota^{-1}\,\mathrm{Tr}_{K/k}(\mathrm{Alb}(X))$ is not a sufficient substitute, as $\mathrm{Alb}(X)$ will often be trivial. Chatzidakis and Hrushovski [2008a; 2008b] instead use model theory, and a variant of isotriviality called *constructible descent* to k. Their theorem generalizes both Baker's result and the Lang–Néron theorem.

Theorem 5.3. *Let K be any function field and let k be its field of constants. Let $f : X \to X$ be an algebraic dynamical system defined over K, and assume f does not constructibly descend to k. Then for every point $x \in X(\overline{K})$ with dynamical height zero there exists a proper Zariski closed subset $Y_x \subsetneq X$ such that the orbit of x is contained in Y_x.*

The author is optimistic that the methods of arithmetic heights and rigidity theorem of this paper, combined with model-theoretic treatment of isotriviality will yield stronger dynamics results over general function fields in the future.

Acknowledgements

The author expresses his gratitude to Xinyi Yuan, his doctoral thesis advisor, for an introduction to this subject and for support throughout his PhD and beyond. Thanks go also to the referee for their comments, corrections, and clarifying suggestions. The author was supported by an NSF GRFP and NSF RTG grant during the preparation of this paper.

References

[Arakelov 1974] S. J. Arakelov, "Intersection theory for divisors on an arithmetic surface", *Izv. Akad. Nauk SSSR Ser. Mat.* **38** (1974), 1179–1192. In Russian; translated in *Math. USSR, Izv.* **8**:6 (1974), 1167–1180. MR Zbl

[Arakelov 1975] S. J. Arakelov, "Theory of intersections on the arithmetic surface", pp. 405–408 in *Proceedings of the International Congress of Mathematicians, I* (Vancouver, B.C. 1974), 1975. MR

[Baker 2009] M. Baker, "A finiteness theorem for canonical heights attached to rational maps over function fields", *J. Reine Angew. Math.* **626** (2009), 205–233. MR

[Benedetto 2005] R. L. Benedetto, "Heights and preperiodic points of polynomials over function fields", *Int. Math. Res. Not.* **2005**:62 (2005), 3855–3866. MR

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI, 1990. MR

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs **4**, Cambridge University Press, 2006. MR

[Call and Silverman 1993] G. S. Call and J. H. Silverman, "Canonical heights on varieties with morphisms", *Compositio Math.* **89**:2 (1993), 163–205. MR

[Chambert-Loir 2006] A. Chambert-Loir, "Mesures et équidistribution sur les espaces de Berkovich", *J. Reine Angew. Math.* **595** (2006), 215–235. MR

[Chambert-Loir and Thuillier 2009] A. Chambert-Loir and A. Thuillier, "Mesures de Mahler et équidistribution logarithmique", *Ann. Inst. Fourier* (*Grenoble*) **59**:3 (2009), 977–1014. MR

[Chatzidakis and Hrushovski 2008a] Z. Chatzidakis and E. Hrushovski, "Difference fields and descent in algebraic dynamics, I", *J. Inst. Math. Jussieu* **7**:4 (2008), 653–686. MR

[Chatzidakis and Hrushovski 2008b] Z. Chatzidakis and E. Hrushovski, "Difference fields and descent in algebraic dynamics, II", *J. Inst. Math. Jussieu* **7**:4 (2008), 687–704. MR

[Conrad 2006] B. Conrad, "Chow's K/k-image and K/k-trace, and the Lang–Néron theorem", *Enseign. Math.* (2) **52**:1-2 (2006), 37–108. MR

[Fakhruddin 2003] N. Fakhruddin, "Questions on self maps of algebraic varieties", *J. Ramanujan Math. Soc.* **18**:2 (2003), 109–122. MR

[Faltings 1984] G. Faltings, "Calculus on arithmetic surfaces", *Ann. of Math.* (2) **119**:2 (1984), 387–424. MR

[Grothendieck 1962] A. Grothendieck, *Fondements de la géométrie algébrique*, Secrétariat mathématique, Paris, 1962. MR Zbl

[Gubler 1998] W. Gubler, "Local heights of subvarieties over non-Archimedean fields", *J. Reine Angew. Math.* **498** (1998), 61–113. MR

[Gubler 2007a] W. Gubler, "The Bogomolov conjecture for totally degenerate abelian varieties", *Invent. Math.* **169**:2 (2007), 377–400. MR

[Gubler 2007b] W. Gubler, "Tropical varieties for non-Archimedean analytic spaces", *Invent. Math.* **169**:2 (2007), 321–376. MR

[Hriljac 1985] P. Hriljac, "Heights and Arakelov's intersection theory", *Amer. J. Math.* **107**:1 (1985), 23–38. MR

[Kleiman 2005] S. L. Kleiman, "The Picard scheme", pp. 235–321 in *Fundamental algebraic geometry*, edited by B. Fantechi et al., Math. Surveys Monogr. **123**, Amer. Math. Soc., Providence, RI, 2005. MR

[Lang 1983] S. Lang, *Abelian varieties*, Springer, 1983. Reprint of the 1959 original. MR

[Lang 1988] S. Lang, *Introduction to Arakelov theory*, Springer, 1988. MR

[Lang and Néron 1959] S. Lang and A. Néron, "Rational points of abelian varieties over function fields", *Amer. J. Math.* **81** (1959), 95–118. MR

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I*, Ergebnisse der Mathematik (3) **48**, Springer, 2004. MR Zbl

[Mimar 2013] A. Mimar, "On the preperiodic points of an endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ which lie on a curve", *Trans. Amer. Math. Soc.* **365**:1 (2013), 161–193. MR

[Moriwaki 1996] A. Moriwaki, "Hodge index theorem for arithmetic cycles of codimension one", *Math. Res. Lett.* **3**:2 (1996), 173–183. MR

[Moriwaki 2014] A. Moriwaki, *Arakelov geometry*, Translations of Mathematical Monographs **244**, American Mathematical Society, Providence, RI, 2014. Translated from the 2008 Japanese original. MR

[Seidenberg 1950] A. Seidenberg, "The hyperplane sections of normal varieties", *Trans. Amer. Math. Soc.* **69** (1950), 357–386. MR

[SGA 6 1971] A. Grothendieck, P. Berthelot, and L. Illusie, *Théorie des intersections et théorème de Riemann–Roch* (Séminaire de Géométrie Algébrique du Bois Marie 1966–1967), Lecture Notes in Math. **225**, Springer, 1971. MR Zbl

[Shioda 1999] T. Shioda, "Mordell–Weil lattices for higher genus fibration over a curve", pp. 359–373 in *New trends in algebraic geometry* ((Warwick, 1996)), edited by C. P. Klaus Hulek, Fabrizio Catanese and M. Reid, London Math. Soc. Lecture Note Ser. **264**, Cambridge Univ. Press, 1999. MR

[Soulé 1992] C. Soulé, *Lectures on Arakelov geometry*, Cambridge Studies in Advanced Mathematics **33**, Cambridge University Press, 1992. MR

[Wittenberg 2008] O. Wittenberg, "On Albanese torsors and the elementary obstruction", *Math. Ann.* **340**:4 (2008), 805–838. MR

[Yuan and Zhang 2017] X. Yuan and S.-W. Zhang, "The arithmetic Hodge index theorem for adelic line bundles", *Math. Ann.* **367**:3-4 (2017), 1123–1171. MR

[Zhang 1995] S. Zhang, "Small points and adelic metrics", *J. Algebraic Geom.* **4**:2 (1995), 281–300. MR

Received December 1, 2019. Revised July 24, 2020.

ALEXANDER CARNEY
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ROCHESTER
ROCHESTER, NY
UNITED STATES
alexanderjcarney@rochester.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.103

ON THE VANISHING OF THE THETA INVARIANT AND A CONJECTURE OF HUNEKE AND WIEGAND

OLGUR CELIKBAS

Huneke and Wiegand conjectured that, if M is a finitely generated, nonfree, torsion-free module with rank over a one-dimensional Cohen–Macaulay local ring R, then the tensor product of M with its algebraic dual has torsion. This conjecture, if R is Gorenstein, is a special case of a celebrated conjecture of Auslander and Reiten on the vanishing of self-extensions that stems from the representation theory of finite-dimensional algebras.

If R is a one-dimensional Cohen–Macaulay ring such that $R = S/(f)$ for some local ring $(S, \mathfrak{n})$, and a non-zero-divisor $f \in \mathfrak{n}^2$ on S, we make use of Hochster's theta invariant and prove that such R-modules M which have finite projective dimension over S satisfy the proposed torsion conclusion of the conjecture. Along the way we give several applications of our argument pertaining to torsion properties of tensor products of modules.

1. Introduction 103
2. Preliminaries 106
3. Proof of the main result 111
4. Some corollaries of the main result 117
5. On tensor products of totally reflexive modules 129
Appendix A. On the vanishing of the theta invariant 136
Appendix B. Some examples of torsion-free tensor products 139
Acknowledgements 141
References 141

1. Introduction

This paper concerns commutative Noetherian local rings $(R, \mathfrak{m}, k)$ and finitely generated R-modules.

The aim of this paper is to study the torsion-freeness property of tensor products of modules, a subtle topic which stems from the beautiful work of Auslander [1961].

MSC2010: primary 13D07; secondary 13C13, 13C14, 13H10.

Keywords: complete intersection dimension, complexity, theta invariant, torsion in tensor products of modules, vanishing of Ext and Tor.

Our focus is on the torsion of tensor products of the form $M \otimes_R M^*$ over one-dimensional Cohen–Macaulay local rings R, where M^* denotes $\mathrm{Hom}_R(M, R)$. In particular, we are concerned with the following long-standing conjecture of Huneke and Wiegand.

Conjecture 1.1 [Huneke and Wiegand 1994, page 473]. *Let R be a one-dimensional local ring and let M be a finitely generated, nonfree, torsion-free R-module. If M has rank (e.g., if R is a domain), then $M \otimes_R M^*$ has torsion.*

Recall that a finitely generated R-module M is said to have *rank* if there is a nonnegative integer r such that $M_{\mathfrak{p}} \cong R_{\mathfrak{p}}^{\oplus r}$ for all associated primes $\mathfrak{p}$ of R; see [Bruns and Herzog 1993, 1.4.3].

Conjecture 1.1 stems from the seminal works of Auslander [1961], and Huneke and Wiegand [1994]. The conjecture is true over hypersurface rings [Huneke and Wiegand 1994, 3.7], but it is very much open in general, even for ideals over complete intersection domains of codimension two. It is worth noting that Conjecture 1.1 is a special version of the celebrated conjecture of Auslander and Reiten [1975] on the vanishing of Ext when the ring in question is a one-dimensional Gorenstein domain; see [Celikbas and Wiegand 2015, 8.6] for details.

There is strong evidence that Conjecture 1.1 should be true over complete intersections; see [Celikbas and Wiegand 2015; Huneke et al. 2019]. Moreover, there are various examples supporting the conjecture over rings that are not necessarily complete intersections. For example, it is proved in [Huneke et al. 2019, 3.6] that Conjecture 1.1 is true over Cohen–Macaulay rings with minimal multiplicity, e.g., over local Arf rings [Lipman 1971]. For some further examples, we refer to [Celikbas et al. 2019a] and point out the following:

Example 1.2. Let R be a one-dimensional, reduced, nonregular, local ring.

(i) If R is complete, and has prime characteristic p and perfect residue field, then it follows ${}^{\varphi^n}\!R \otimes_R ({}^{\varphi^n}\!R)^*$ has torsion for all $n \gg 0$. Here $\varphi^n : R \to R$ is the n-th iterate of the Frobenius endomorphism given by $r \mapsto r^{p^n}$, and ${}^{\varphi^n}\!R$ denotes R with the R-action given by $r \cdot s = r^{p^n} s$ for all $r, s \in R$; see [Celikbas et al. 2019a, 2.15; Miller 2003, 2.1.3 and 2.2.12].

(ii) If R is a Gorenstein domain and I is an Ulrich ideal of R which is not principal, then I is a self-dual R-module, i.e., $I \cong I^*$, and so $I \otimes_R I^*$ has torsion. In particular, if $R = \mathbb{C}[\![t^4, t^5, t^6]\!]$ and $I = (t^4, t^6)$, then $I \otimes_R I^*$ has torsion; see Example 4.17 and Proposition 4.18.

The purpose of this paper is to prove Theorem 1.3 and give some observations about Conjecture 1.1; see Theorem 3.2 for a higher dimensional version of the next result. The tool we employ to prove Theorem 1.3 is the Hochster's θ invariant, which was initially defined by Hochster [1981] to study the direct summand conjecture; it was further developed by Dao [2008; 2013], and more recently by Buchweitz and van Straten [2012], and Walker et al. [Moore et al. 2011; Walker 2017]. The invariant $\theta^R(M, N)$ for R-modules M and N is defined as $\operatorname{length}_R(\operatorname{Tor}^R_{2n+2}(M, N)) - \operatorname{length}_R(\operatorname{Tor}^R_{2n+1}(M, N))$ for $n \gg 0$; see §2.13 for the details.

Theorem 1.3. *Let R be a one-dimensional Cohen–Macaulay local ring, where $R = S/(f)$ for some local ring $(S, \mathfrak{n})$ and a non-zero-divisor $f \in \mathfrak{n}$ on S. Let M and N be nonzero finitely generated R-modules, and assume the following conditions hold*:

(i) $\operatorname{pd}_S(M) < \infty$ *or* $\operatorname{pd}_S(N) < \infty$ (*e.g., S is regular*).

(ii) $\operatorname{length}_R(\operatorname{Tor}^R_i(M, N)) < \infty$ *for all $i \gg 0$* (*e.g., R is reduced*).

(iii) $\theta^R(M, N) = 0$.

If $M \otimes_R N$ is torsion-free, then $\operatorname{Tor}^R_i(M, N) = 0$ for all $i \geq 1$, and M and N are torsion-free.

To the best of our knowledge, Theorem 1.3 is new, even if S is a ramified regular ring; see [Celikbas et al. 2015a, 3.6] and Section 3. Next is a corollary of Theorem 1.3 concerning Conjecture 1.1; see Corollaries 4.6 and 4.8.

Corollary 1.4. *Let R be a one-dimensional Cohen–Macaulay ring such that, for some local ring $(S, \mathfrak{n})$ and a non-zero-divisor $f \in \mathfrak{n}^2$ on S, we have $R = S/(f)$. Assume M is a finitely generated R-module that has rank.*

If M is a nonfree torsion-free R-module and $\operatorname{pd}_S(M) < \infty$, then $M \otimes_R M^*$ has torsion. In particular, if $M = \operatorname{coker}(\alpha)$, where (α, β) is a reduced matrix factorization of f over S (i.e., a matrix factorization of f with entries in $\mathfrak{n}$), then $M \otimes_R M^*$ has torsion.

As mentioned previously, if S is regular, Corollary 1.4 follows from a result of Huneke and Wiegand [1994, 3.7]. In this case, as is well-known, maximal Cohen–Macaulay R-modules with no free summands occur as reduced matrix factorizations of f over S; see [Eisenbud 1980]. Similarly, if S is *G-regular* (i.e., when there are no nonfree totally reflexive S-modules), Takahashi [2008] proved that there is a one-to-one correspondence between reduced matrix factorizations of f and totally reflexive R-modules without free summands. Note that, if the ring R is as in Corollary 1.4, reduced matrix factorizations of f exist due to a result of Herzog, Ulrich, and Backelin; see [Herzog et al. 1991, 1.2 and 2.2], and also [Avramov 1998, 5.1.3, Avramov et al. 1997, 3.1, Yoshino 1990, Chapter 8].

In Sections 2 and 3 we collect some preliminary results and give a proof of Theorem 1.3, respectively. Sections 4 and 5 are devoted to several applications of Theorem 1.3 pertaining to torsion properties of tensor products of modules. As Theorem 1.3 relies upon the vanishing of theta invariant, in Appendix A we point out by an example that $\theta^R(M, N)$ can vanish nontrivially: in Example A.3, we record an example of a one-dimensional reduced hypersurface ring R, and finitely generated R-modules M and N such that $\theta^R(M, N) = 0$, but neither M nor N has rank, or equivalently, neither M nor N has zero class in the reduced Grothendieck group $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$. Moreover, in Appendix B, building on an argument of Huneke and Wiegand [1994, 4.7], we recall how to obtain examples of nonfree, torsion-free R-modules M with rank such that $M \otimes_R M$ is torsion-free over certain one-dimensional local rings R; see §B.1.

2. Preliminaries

In this section we recall definitions and collect some basic facts that will be used throughout the paper. We have, by definition, $\operatorname{depth}(0) = \infty$ and $\operatorname{pd}(0) = -\infty$. Moreover, ΩM denotes the syzygy of a given finitely generated R-module M.

2.1 *Torsion submodule*. Let R be a local ring and let M be a finitely generated R-module. The *torsion submodule* $\top_R M$ of M is the kernel of the natural map $M \to \mathrm{Q}(R) \otimes_R M$, where $\mathrm{Q}(R)$ is the total quotient ring of R. Hence there is an exact sequence of R-modules:

$$0 \to \top_R M \to M \to \perp_R M \to 0. \tag{2.1.1}$$

M is said to have *torsion* (or be, *torsion-free*) if $\top_R M \neq 0$ (respectively, $\top_R M = 0$). Note, M is torsion, i.e., $\top_R M = M$, if and only if $M_{\mathfrak{p}} = 0$ for each associated prime $\mathfrak{p}$ of R. Note also that $M = 0$ if and only if M is both torsion and torsion-free.

2.2. Let R be a local ring. If $0 \to X \to Y \to Z \to 0$ is a short exact sequence of finitely generated R-modules, then it follows that the sequence

$$0 \to \top_R X \to \top_R Y \to \top_R Z$$

is exact. In particular, if X and Z are torsion-free, then so is Y.

The next fact will be used several times throughout, for example, for Corollary 4.6.

2.3. Let R be a local ring and let M be a finitely generated R-module. Set $M^* = \operatorname{Hom}(M, R)$, the algebraic dual of M. If $M^* = 0$, then there is an $x \in R$, which is a non-zero-divisor on R, such that $xM = 0$; see [Bruns and Herzog 1993, 1.2.3(b)]. In other words, $M^* = 0$ if and only if M is a torsion R-module. In particular, if $M \neq 0$ and M is torsion-free, then $M \otimes_R M^* \neq 0$.

The following argument is from [Huneke and Wiegand 1994]; we will invoke it in the proofs of Theorem 3.2, Remark 3.5, Proposition 4.4, and Corollary 5.12.

2.4 [Huneke and Wiegand 1994, 1.1]. Let R be a local ring, and let M and N be nonzero finitely generated R-modules. Assume $M \otimes_R N$ is torsion-free. Set $U = \perp_R M$ and $V = \top_R M$. Then:

(i) $M \otimes_R N \cong U \otimes_R N$.

(ii) If $\operatorname{Tor}_1^R(U, N) = 0$, then M is torsion-free, i.e., $M = U$.

(iii) If $\operatorname{Tor}_i^R(U, N) = 0$ for all $i \geq 1$, then $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$.

To establish part (i), we tensor the exact sequence $0 \to V \to M \to U \to 0$ by N, and obtain the following exact sequence of R-modules:

$$\operatorname{Tor}_1^R(U, N) \to V \otimes_R N \xrightarrow{\alpha} M \otimes_R N \xrightarrow{\beta} U \otimes_R N \to 0.$$

As $V \otimes_R N$ is torsion, we see that the image of α is also torsion. Hence, since $M \otimes_R N$ is torsion-free, it follows that $\alpha = 0$. Therefore, β is an isomorphism and part (i) follows.

Now assume $\operatorname{Tor}_1^R(U, N) = 0$. Then α is both zero and injective so that $V \otimes_R N = 0$. This implies, as $N \neq 0$, that $V = 0$, i.e., M is torsion-free, i.e., $U = M$. This proves part (ii). Notice, part (iii) is a consequence of part (ii).

2.5 ***Gorenstein and complete intersection dimensions*** [Auslander and Bridger 1969; Avramov et al. 1997]. Let R be a local ring and let M be a finitely generated R-module.

M is said to be *totally reflexive* provided that the natural map $M \to M^{**}$ is bijective and $\operatorname{Ext}_R^i(M, R) = 0 = \operatorname{Ext}_R^i(M^*, R)$ for all $i \geq 1$. The infimum of n for which there exists an exact sequence $0 \to X_n \to \cdots \to X_0 \to M \to 0$ such that each X_i is totally reflexive is called the *Gorenstein dimension* of M. If M has Gorenstein dimension n, we write $\operatorname{G-dim}_R(M) = n$. Therefore, M is totally reflexive if and only if $\operatorname{G-dim}_R(M) \leq 0$, where it follows by convention that $\operatorname{G-dim}_R(0) = -\infty$.

A diagram of local ring maps $R \to R' \twoheadleftarrow S$ is called a *quasi-deformation* if $R \to R'$ is flat and the kernel of the surjection $R' \twoheadleftarrow S$ is generated by a regular sequence on S. The *complete intersection dimension* of M is defined as follows:

$$\operatorname{CI-dim}_R(M) = \inf\{\operatorname{pd}_S(M \otimes_R R') - \operatorname{pd}_S(R') : R \to R' \twoheadleftarrow S \text{ is a quasi-deformation}\}.$$

The following inequalities hold in general:

$$\operatorname{G-dim}_R(M) \leq \operatorname{CI-dim}_R(M) \leq \operatorname{pd}_R(M). \tag{2.5.1}$$

Moreover, if any of the dimensions in (2.5.1) is finite, then it is equal to those to its left.

2.6 ***Complexity*** [Avramov 1989]. Let R be a local ring and let M be a finitely generated R-module. The *complexity* $\operatorname{cx}_R(M)$ of M is the smallest nonnegative integer r such that there exists a real number A with $\beta_n(M) \leq A \cdot n^{r-1}$ for all $n \gg 0$. Here $\beta_n(M)$ is the n-th Betti number of M. It follows that $\operatorname{cx}_R(M) = 0$ if and only if $\operatorname{pd}_R(M) < \infty$, and $\operatorname{cx}_R(M) \leq 1$ if and only if M has bounded Betti numbers.

Next we collect certain properties of complexity and complete intersection dimension. Prior to that we recall the definition of the transpose:

2.7 ***Auslander transpose*** [Auslander and Bridger 1969]. Let R be a local ring and let M be a finitely generated R-module with a projective presentation

$$P_1 \xrightarrow{f} P_0 \to M \to 0.$$

The *transpose* $\operatorname{Tr} M$ of M is the cokernel of $f^* = \operatorname{Hom}_R(f, R)$, and so is given by the following exact sequence:

$$0 \to M^* \to P_0^* \to P_1^* \to \operatorname{Tr} M \to 0. \tag{2.7.1}$$

In particular, up to projectives, $\operatorname{Tr} M$ is uniquely defined and $\operatorname{Tr}\operatorname{Tr} M \cong M$.

2.8. Let R be a local ring and let M and N be finitely generated R-modules such that $M \neq 0$.

(i) If $\text{CI-dim}_R(M) < \infty$, then it follows that $\operatorname{cx}_R(M) \le \mathsf{embdim}(R) - \operatorname{depth}(R)$; see [Avramov et al. 1997, 5.6].

(ii) If $\text{CI-dim}_R(M) < \infty$, then it follows that $\text{CI-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}) \le \text{CI-dim}_R(M)$ for all $\mathfrak{p} \in \operatorname{Spec}(R)$; see [Avramov et al. 1997, 1.6].

(iii) If $\text{CI-dim}_R(M) < \infty$, then it follows that $\text{CI-dim}_R(M) = \operatorname{depth}(R) - \operatorname{depth}_R(M)$, which also equals $\sup\{i : \operatorname{Ext}^i_R(M, R) \neq 0\}$; see [Avramov et al. 1997, 1.4].

(iv) Assume $\text{CI-dim}_R(M) < \infty$. If f is a non-zero-divisor on R and $fM = 0$, then it follows that $\text{CI-dim}_{R/fR}(M) < \infty$. Also, if f is a non-zero-divisor on both R and M, then it follows that $\text{CI-dim}_{R/fR}(M/fM) = \text{CI-dim}_R(M)$; see [Avramov et al. 1997, 1.12.2–3].

(v) If $R \to R'$ is a flat local map of local rings and $\text{CI-dim}_{R'}(M \otimes_R R') < \infty$, then it follows that $\text{CI-dim}_R(M) = \text{CI-dim}_{R'}(M \otimes_R R')$; see [Avramov et al. 1997, 1.11].

(vi) If $\text{CI-dim}_R(M) = 0$, then it follows that $\text{CI-dim}(M^*) = \text{CI-dim}_R(\operatorname{Tr} M) = 0$, and also $\operatorname{cx}(M) = \operatorname{cx}(M^*) < \infty$. Moreover, $\text{CI-dim}_R(M) = 0$ if and only if $\text{CI-dim}_R(\operatorname{Tr} M) = 0$; see [Bergh and Jorgensen 2011, 3.5; 2014, 3.2; Celikbas et al. 2015b, 3.2(i)].

(vii) If $\text{CI-dim}_R(M) < \infty$ and $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \gg 0$, then it follows $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \ge \text{CI-dim}_R(M) + 1$; see [Avramov and Buchweitz 2000, 4.9]. Hence, if $\text{CI-dim}_R(M) < \infty$, then $\operatorname{Tor}_i^R(M, N)$ is torsion for all $i \gg 0$ if and only if $\operatorname{Tor}_i^R(M, N)$ is torsion for all $i \ge 1$; see §2.1.

(viii) If $\text{CI-dim}_R(M) < \infty$ and $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \ge 1$, then the *depth formula* for M and N holds, i.e., $\operatorname{depth}(M) + \operatorname{depth}(N) = \operatorname{depth}(R) + \operatorname{depth}(M \otimes_R N)$; see [Araya and Yoshino 1998, 2.5].

2.9 [Avramov and Martsinkovsky 2002, 3.1]. Let R be a local ring and let N be a finitely generated R-module such that $\text{G-dim}_R(N) < \infty$. Then there is an exact sequence of finitely generated R-modules $0 \to L \to Z \to N \to 0$, where $\text{G-dim}_R(Z) = 0$ and $\text{pd}_R(L) < \infty$.

The next result is due to Sather-Wagstaff [2004]; here we record the module case, but in fact his result holds for homologically finite complexes.

2.10 [Sather-Wagstaff 2004, 3.6]. Let R be a local ring and let

$$0 \to X_1 \to X_2 \to X_3 \to 0$$

be a short exact sequence of finitely generated R-modules. Assume i, j, s are integers with $\{i, j, s\} = \{1, 2, 3\}$. If $\text{pd}_R(X_i) < \infty$ and $\text{CI-dim}_R(X_j) < \infty$, then $\text{CI-dim}_R(X_s) < \infty$.

We use §2.11 in the proofs of Theorem 3.2, Corollary 4.1, and Proposition 4.4.

2.11. Let R be a one-dimensional Cohen–Macaulay local ring, and let N be a finitely generated R-module such that $\text{CI-dim}_R(N) < \infty$. Then there is an exact sequence of finitely generated R-modules

$$0 \to F \to Z \to N \to 0, \tag{2.11.1}$$

where F is free, $\text{CI-dim}_R(Z) = 0$, and $\text{cx}_R(Z) = \text{cx}_R(N)$.

To see this, first note that we have $\text{G-dim}_R(N) < \infty$ since $\text{CI-dim}_R(N) < \infty$; see (2.5.1). Hence a short exact sequence as in (2.11.1) exists by § 2.9, where $\text{G-dim}_R(Z) = 0$ and $\text{pd}_R(F) < \infty$. Now § 2.10 implies that $\text{CI-dim}_R(Z) < \infty$. Consequently, we deduce that $\text{CI-dim}_R(Z) = \text{G-dim}_R(Z) = 0$; see (2.5.1). Note, by §2.8(iii), we have that $\text{depth}_R(Z) = 1$. Thus, the depth lemma applied to (2.11.1) yields $\text{depth}_R(F) = 1$. So, we conclude that F is free by the Auslander–Buchsbaum formula. Finally, notice that as F is free, by tensoring (2.11.1) with k, we obtain that $\beta_i^R(Z) = \beta_i^R(N)$ for each $i \geq 2$. This yields the equality $\text{cx}_R(Z) = \text{cx}_R(N)$; see §2.6.

We use the following exact sequence in § 2.13, and also in the proofs of Theorem 3.2 and Proposition 5.9.

2.12 [Rotman 1979, 11.65]. Let $(S, \mathfrak{n})$ be a local ring and let $R = S/(f)$ for some non-zero-divisor $f \in \mathfrak{n}$ on S. If M and N are finitely generated R-modules, then we have the change of rings long exact sequence of Tors:

$$\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\text{Tor}_i^R(M,N) & \to & \text{Tor}_{i+1}^S(M,N) & \to & \text{Tor}_{i+1}^R(M,N) & \to & \\
\vdots & & \vdots & & \vdots & & \\
\text{Tor}_0^R(M,N) & \to & \text{Tor}_1^S(M,N) & \to & \text{Tor}_1^R(M,N) & \to & 0
\end{array}$$

In the following, we recall the definition of a version of Hochster's θ pairing [1981], developed by Dao [2007]. This pairing can be defined in a more general setting, but the definition recorded here suffices for our argument; see [Dao 2008; 2013] for more details.

2.13 θ ***pairing*** [Dao 2007; Hochster 1981]. Let M and N be finitely generated R-modules. Assume $R \to R' \twoheadleftarrow S$ is a codimension one quasi-deformation with zero-dimensional closed fibre, i.e., we have a diagram of local ring maps such that $R \to R'$ is flat, $R' \cong S/(f)$ for some non-zero-divisor f on S, and $\dim(R'/\mathfrak{m}R') = 0$. We set $(-)' = - \otimes_R R'$ and assume the following conditions hold:

(a) $\text{CI-dim}_S(N') < \infty$ and $\text{Tor}_i^S(M', N') = 0$ for all $i \gg 0$ (e.g., $\text{pd}_S(N') < \infty$).

(b) $\text{length}_R(\text{Tor}_i^R(M, N)) < \infty$ for all $i \gg 0$ (e.g., R is an isolated singularity).

It follows that

$$\text{CI-dim}_{R'}(N') < \infty \quad \text{and} \quad \text{CI-dim}_R(N) = \text{CI-dim}_{R'}(N');$$

see §2.8(iv, v). Note we have, by (a) and §2.8(vii), that $\text{Tor}_i^S(M', N') = 0$ for all $i > \text{CI-dim}_S(N')$. Therefore §2.12 yields the following isomorphisms:

$$\text{Tor}_i^{R'}(M', N') \cong \text{Tor}_{i+2}^{R'}(M', N') \quad \text{for all } i > \text{CI-dim}_{R'}(N'). \tag{2.13.1}$$

For a nonmaximal prime ideal $\mathfrak{p}$, we have

$$\text{Tor}_i^R(M, N)_{\mathfrak{p}} = 0 \quad \text{for all } \ i > \text{CI-dim}_R(N);$$

see §2.8(ii, vii). Thus $\text{length}(\text{Tor}_i^R(M, N)) < \infty$ for all $i > \text{CI-dim}_R(N)$, and hence

$$\text{length}_{R'}(\text{Tor}_i^{R'}(M', N')) < \infty \quad \text{for all } i > \text{CI-dim}_{R'}(N'). \tag{2.13.2}$$

Let $\ell = \text{length}_{R'}(R'/\mathfrak{m}R')$. Then, by (2.13.1) and (2.13.2), we see that the difference

$$\begin{aligned}\text{length}_R(\text{Tor}_{2n+2}^R(M, N)) &- \text{length}_R(\text{Tor}_{2n+1}^R(M, N)) \\ &= \frac{1}{\ell} \cdot \big(\text{length}_{R'}(\text{Tor}_{2n+2}^{R'}(M', N')) - \text{length}_{R'}(\text{Tor}_{2n+1}^{R'}(M', N'))\big)\end{aligned}$$

is independent of n if $2n > \text{CI-dim}_{R'}(N') - 1$. One defines the theta pairing over R as

$$\theta^R(M, N) = \text{length}_R(\text{Tor}_{2n+2}^R(M, N)) - \text{length}_R(\text{Tor}_{2n+1}^R(M, N)), \tag{2.13.3}$$

where n is an integer such that $2n > \text{CI-dim}_R(N) - 1$.

It follows θ^R is additive on short exact sequence of finitely generated R-modules, whenever it is well-defined on each pair of modules in question; see, for example, [Dao 2007, 4.3(2)].

2.14. Let M and N be finitely generated R-modules. Assume the following conditions hold:

(a) $$\text{CI-dim}_R(N) < \infty \quad \text{and} \quad \text{cx}_R(N) = 1.$$

(b) $$\text{length}_R(\text{Tor}_i^R(M, N)) < \infty \quad \text{for all } i \gg 0.$$

Then we can choose a codimension one quasi-deformation of the form $R \to R' \xleftarrow{\alpha} S$, where $\text{pd}_S(N') < \infty$; see [Avramov and Buchweitz 2000, 4.1.3]. Localizing at some $\mathfrak{p} \in \text{Min}_{R'}(R'/\mathfrak{m}R')$, set $\mathfrak{q} = \alpha^{-1}(\mathfrak{p})$: we see that $R \to R'_{\mathfrak{p}} \xleftarrow{\alpha} S_{\mathfrak{q}}$ is a codimension one quasi-deformation with $\text{pd}_{S_{\mathfrak{q}}}(N \otimes_R R'_{\mathfrak{p}}) < \infty$; see the proof of [Sather-Wagstaff 2004, 2.11]. Therefore, replacing the original quasi-deformation with the aforementioned one, we may assume $\dim(R'/\mathfrak{m}R') = 0$.

So it follows from (2.13.3) that $\theta^R(M, N)$ is well-defined, as long as n is an integer such that $2n > \text{CI-dim}_R(N) - 1$.

Next we record two more preliminary results prior to moving to the next section; both of these results are used in the next section for our proof of Theorem 3.2.

2.15. Let R be a local ring, and let L and M be finitely generated R-modules. If M is maximal Cohen–Macaulay and $\text{pd}_R(L) < \infty$, then $\text{Tor}_i^R(L, M) = 0$ for all $i \geq 1$; see, for example, [Celikbas 2011, 3.8].

2.16. Let R be a local ring and let M be a finitely generated R-module. Assume M has a finite free resolution, i.e., there is an exact sequence

$$0 \to F_n \to \cdots \to F_1 \to F_0 \to M \to 0,$$

where each F_i is a finitely generated free R-module. The *Euler number* of M is defined as $\chi(M) = \sum (-1)^i \operatorname{rank} F_i$.

It is known that $\chi(M)$ is independent of the choice of the finite free resolution of M. Moreover, it follows that $\chi(M) = 0$ if and only if there is a non-zero-divisor f on R such that $fM = 0$; see [Matsumura 1986, 19.8 and page 158].

3. Proof of the main result

In this section we prove the main result of this paper; see Theorem 3.2. Our motivation comes from the following result, which is recorded for the one-dimensional case:

3.1 [Celikbas et al. 2015a, 3.6]. Let R be a one-dimensional local ring with $\widehat{R} = S/(f)$ for some unramified regular local ring S, and a non-zero-divisor $f \in \mathfrak{n}$ on S. Let M and N be finitely generated R-modules. If $\text{length}_R(\text{Tor}_i^R(M, N)) < \infty$ for all $i \gg 0$, $\theta^R(M, N) = 0$, and $M \otimes_R N$ is torsion-free, then $\text{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, and M and N are both torsion-free.

A consequence of our argument gives an extension of §3.1 and establishes the vanishing of $\text{Tor}_i^R(M, N)$ when S is an arbitrary two-dimensional Cohen–Macaulay local ring, and $\mathrm{pd}_S(N) < \infty$ or $\mathrm{pd}_S(M) < \infty$; see Theorem 3.2. As is clear, since we do not work over hypersurface rings, our method of proof is different from that employed to prove §3.1. Among other things, one of the properties that is not available to us under our setup is that, when R is a hypersurface, every torsion-free module can be embedded in a free R-module; see [Huneke and Wiegand 1994, 1.5]. Also, over a ring R as in §3.1, for a pair of finitely generated R-modules (M, N), if $\theta^R(M, N)$ is defined and vanishes, then the pair (M, N) is Tor-rigid, i.e., if $\mathrm{Tor}_n^R(M, N) = 0$ for some $n \geq 0$, then $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq n$ [Dao 2013, 2.8]; this Tor-rigidity result depends on the fact that S is an unramified regular ring. Thus the properties that play an important role in the proof of §3.1 do not apply directly under our setup.

The following is our main result; although we are interested in the one dimensional case (due to Conjecture 1.1), our argument works over Cohen–Macaulay local rings of arbitrary positive dimension if the modules considered have sufficiently large depth; see also Remark 3.5.

Theorem 3.2. *Let R be a Cohen–Macaulay local ring, and let M and N be finitely generated R modules. Assume* $\dim(R) = d \geq 1$ *and the following conditions hold*:

(i) $\text{CI-dim}_R(N) < \infty$ *and* $\mathrm{cx}_R(N) = 1$.

(ii) $\mathrm{length}_R(\mathrm{Tor}_i^R(M, N)) < \infty$ *for all* $i \gg 0$.

(iii) $\mathrm{depth}_R(M) \geq d - 1$ *and* $\mathrm{depth}_R(N) \geq d - 1$.

(iv) *If* $d = 1$, *assume further* $\theta^R(M, N) = 0$.

If $M \otimes_R N$ is (nonzero) maximal Cohen–Macaulay, then $\mathrm{Tor}_i^R(M, N) = 0$ *for all* $i \geq 1$, *and M and N are both maximal Cohen–Macaulay.*

Proof. It suffices to prove the vanishing of $\mathrm{Tor}_i^R(M, N)$ for all $i \geq 1$; see §2.8(viii).

First we assume $d \geq 2$, and choose a non-zero-divisor x on R, M, N, and $M \otimes_R N$. Set $T = R/xR$, $A = M/xM$, and $B = N/xN$. Notice

$$\text{CI-dim}_T(B) = \text{CI-dim}_R(N) < \infty \quad \text{and} \quad \mathrm{cx}_T(B) = \mathrm{cx}_R(N) = 1;$$

see §2.8(iv). We have the following exact sequence:

$$0 \to M \xrightarrow{x} M \to A \to 0. \tag{3.2.1}$$

Tensoring (3.2.1) with N, we obtain the following long exact sequence for all $i \geq 0$:

$$\begin{aligned} \cdots \to \mathrm{Tor}_{i+1}^R(M,N) \xrightarrow{x} \mathrm{Tor}_{i+1}^R(M,N) \to \mathrm{Tor}_{i+1}^R(A,N) \to \mathrm{Tor}_i^R(M,N) \to \\ \cdots \to \mathrm{Tor}_1^R(A,N) \to M \otimes_R N \xrightarrow{x} M \otimes_R N \to A \otimes_R N \to 0. \end{aligned} \tag{3.2.2}$$

As x is a non-zero-divisor on R and N, and $xA = 0$, we have that

$$\mathrm{Tor}_i^T(A, B) \cong \mathrm{Tor}_i^R(A, N) \quad \text{for all } i \geq 0.$$

It follows from (3.2.2) that $\mathrm{length}_T(\mathrm{Tor}_i^T(A, B)) < \infty$ for all $i \gg 0$. Moreover, $\theta^R(A, N)$ is well-defined and we see that $\theta^R(M, N) = \theta^R(M, N) + \theta^R(A, N)$, by additivity applied to (3.2.1). Therefore, $0 = \theta^R(A, N) = \theta^T(A, B)$. It follows from (3.2.2) that $(M \otimes_R N)/x(M \otimes_R N) \cong A \otimes_R N \cong A \otimes_T B$. This implies that $A \otimes_T B$ is maximal Cohen–Macaulay over T. Moreover, $\mathrm{depth}_T(A) \geq \mathrm{depth}(T) - 1$ and $\mathrm{depth}_T(B) \geq \mathrm{depth}(T) - 1$. Consequently, we may use induction to go all the way down to dimension one. More precisely, we may replace the pair (M, N) over the ring R with the pair (A, B) over the ring T, and we may assume $\dim(T) = 1$: this case yields the vanishing of $\mathrm{Tor}_i^T(A, B)$ for all $i \geq 1$, and in view of Nakayama's lemma, we conclude by (3.2.2) that $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, as claimed.

We now proceed by assuming $d = 1$. Set $U = \perp_R M$ and $V = \top_R M$. Note that U is a nonzero maximal Cohen–Macaulay R-module. Choose a quasi-deformation $R \to R' \twoheadleftarrow S$ such that $\dim(R'/\mathfrak{m}R') = 0$, $R' \cong S/(f)$ for some local ring $(S, \mathfrak{n})$ and a non-zero-divisor $f \in \mathfrak{n}$ on S such that $\mathrm{pd}_S(N \otimes_R R') < \infty$. So we may assume $R = R' = S/(f)$, where $\mathrm{pd}_S(N) < \infty$. As $\dim(R) = 1$, we know that $\mathrm{length}_R(V) < \infty$. Thus $\theta^R(V, N)$ is well-defined; see §2.13.

Next we record two claims; we use these claims to prove the vanishing of $\mathrm{Tor}_i^R(M, N)$ for all $i \geq 1$, and defer the proofs of the claims until the end.

Claim 1. $$\theta^R(V, N) = 0.$$

Assuming Claim 1 is true, consider the following short exact sequence of R-modules:

$$0 \to V \to M \to U \to 0. \tag{3.2.3}$$

Recall that we have $\mathrm{length}_R(V) < \infty$ and $\mathrm{length}(\mathrm{Tor}_i^R(M, N)) < \infty$ for all $i \gg 0$. Hence it follows from (3.2.3) that:

$$\mathrm{length}_R(\mathrm{Tor}_i^R(U, N)) < \infty \quad \text{for all } i \gg 0. \tag{3.2.4}$$

In particular, $\theta^R(U, N)$ is well-defined; see §2.13. Moreover, by the additivity of θ pairing applied to the short exact sequence in (3.2.3), we see that:

$$\theta^R(M, N) = \theta^R(U, N) + \theta^R(V, N). \tag{3.2.5}$$

We know, by the hypothesis, that $\theta^R(M, N) = 0$ and, by Claim 1, that $\theta^R(V, N) = 0$. Therefore, it follows from (3.2.5) that:

$$\theta^R(U, N) = 0. \tag{3.2.6}$$

As $\dim(R) = 1$ and $\text{CI-dim}_R(N) < \infty$, it follows from §2.11 that there is a short exact sequence of finitely generated R-modules

$$0 \to F \to Z \to N \to 0, \tag{3.2.7}$$

where F is free, $\text{CI-dim}_R(Z) = 0$, and $\text{cx}_R(Z) = \text{cx}_R(N) = 1$.

Claim 2. $U \otimes_R Z$ is a torsion-free R-module, $\text{length}_R(\text{Tor}_i^R(U, Z)) < \infty$ for each $i \geq 1$, and $\text{Tor}_i^R(U, Z) \cong \text{Tor}_i^R(U, N)$ for all $i \geq 1$.

Assuming Claim 2 is true, note that (3.2.7) implies $\text{pd}_S(Z) < \infty$ since $\text{pd}_S(N)$ and $\text{pd}_S(F)$ are finite. As $\text{length}_R(\text{Tor}_i^R(U, Z)) < \infty$ for all $i \geq 1$, we see from §2.13 that $\theta^R(U, Z)$ is well-defined and the following holds:

(3.2.8) $$\theta^R(U, Z) = \text{length}_R(\text{Tor}_{2n+2}^R(U, Z)) - \text{length}_R(\text{Tor}_{2n+1}^R(U, Z))$$

for each integer n with $2n > \text{CI-dim}_R(Z) - 1$, i.e., for each $n \geq 0$ (because $\text{CI-dim}_R(Z) = 0$).

We know, by Claim 2, that $\text{Tor}_i^R(U, Z) \cong \text{Tor}_i^R(U, N)$ for all $i \geq 1$. Hence it follows that $\theta^R(U, N) = \theta^R(U, Z)$. Thus (3.2.6) and (3.2.8) yield the following equalities of lengths:

(3.2.9) $$\text{length}_R(\text{Tor}_{2i+2}^R(U, Z)) = \text{length}_R(\text{Tor}_{2i+1}^R(U, Z)) \quad \text{for all } i \geq 0.$$

Notice $0 = \text{CI-dim}_R(Z) = \text{depth}(R) - \text{depth}_R(Z)$, i.e., $\text{depth}_R(Z) = 1$; see §2.8(iii). Hence $\text{depth}_S(Z) = 1$. Since $\dim(S) = 2$ and $\text{pd}_S(Z) < \infty$, we conclude, by the Auslander–Buchsbaum formula, that $\text{pd}_S(Z) = 1$. Now we consider the following exact sequence that follows from §2.12 applied for the pair (U, Z):

(3.2.10) $$\text{Tor}_2^S(U, Z) \to \text{Tor}_2^R(U, Z) \to U \otimes_R Z \to \text{Tor}_1^S(U, Z) \to \text{Tor}_1^R(U, Z) \to 0.$$

As $\text{pd}_S(Z) = 1$, we have $\text{Tor}_2^S(U, Z) = 0$. So, by (3.2.10), we see that $\text{Tor}_2^R(U, Z)$ embeds in $U \otimes_R Z$. Moreover, we know by Claim 2 that $\text{length}_R(\text{Tor}_2^R(U, Z)) < \infty$ and $U \otimes_R Z$ is a torsion-free R-module. So we conclude from (3.2.10) that

(3.2.11) $$\text{Tor}_2^R(U, Z) = 0.$$

Consequently, (3.2.9) and (3.2.11) yield

(3.2.12) $$\text{Tor}_1^R(U, Z) = 0 = \text{Tor}_2^R(U, Z).$$

On the other hand, as $\text{pd}_S(Z) = 1$, we can use §2.12 once more, this time for the pair (U, Z), and obtain

(3.2.13) $$\text{Tor}_i^R(U, Z) \cong \text{Tor}_{i+2}^R(U, Z) \quad \text{for all } i \geq 1.$$

Therefore, we see from (3.2.12) and (3.2.13) that $\text{Tor}_i^R(U, Z) = 0$ for all $i \geq 1$. Thus Claim 2 yields the vanishing of $\text{Tor}_i^R(U, N)$ for all $i \geq 1$. Finally we can now invoke §2.4(iii) and deduce that $\text{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, as required.

Now we establish the claims and complete the proof of the theorem.

Proof of Claim 1. To prove the claim, we follow the argument of [Celikbas and Dao 2014, 4.6]. Note that $\Omega_R N$ is a maximal Cohen–Macaulay R-module. So

$\operatorname{depth}_R(\Omega_R N) = 1$, which equals $\operatorname{depth}_S(\Omega_R N)$. Note also that there is a short exact sequence $0 \to \Omega_R N \to T \to N \to 0$, where T is a finitely generated free R-module. This exact sequence, since both $\operatorname{pd}_S(N)$ and $\operatorname{pd}_S(T)$ are finite, implies that $\operatorname{pd}_S(\Omega_R N) < \infty$. Hence, since $\operatorname{depth}(S) = 2$, we conclude from the Auslander-Buchsbaum formula that $\operatorname{pd}_S(\Omega_R N) = 1$. It now follows from §2.12 that there is an exact sequence:

$$(3.2.14)\quad 0 \to \operatorname{Tor}_2^R(\Omega_R N, k) \to \Omega_R N \otimes_R k \to \operatorname{Tor}_1^S(\Omega_R N, k) \to \operatorname{Tor}_1^R(\Omega_R N, k) \to 0.$$

Taking the alternating sum of lengths of modules in (3.2.14), we obtain

$$(3.2.15)\qquad \theta^R(k, \Omega_R N) = \beta_2^R(\Omega_R N) - \beta_1^R(\Omega_R N) = \beta_0^S(\Omega_R N) - \beta_1^S(\Omega_R N),$$

where $\beta_i^R(\Omega_R N)$ denotes the i-th Betti number of $\Omega_R N$.

As $\Omega_R N$ has a finite free resolution over S, we can now apply § 2.16 for the module $\Omega_R N$ over the ring S: the Euler number of $\Omega_R N$ over S, which is $\beta_0^S(\Omega_R N) - \beta_1^S(\Omega_R N)$, vanishes since $f \cdot \Omega_R N = 0$. So, by (3.2.15), we have $\theta^R(k, \Omega_R N) = 0$. As $\theta^R(k, N) = -\theta^R(k, \Omega_R N)$, we see $\theta^R(k, N) = 0$. Moreover, since V has a finite filtration by copies of k, it follows that $\theta^R(V, N)$ vanishes. This justifies Claim 1. □

Proof of Claim 2. Notice, as F is free, tensoring (3.2.7) with U over R, we see that there are isomorphisms

$$(3.2.16)\qquad \operatorname{Tor}_i^R(U, Z) \cong \operatorname{Tor}_i^R(U, N) \quad \text{for each } i \geq 2,$$

and there is an exact sequence of the form

$$(3.2.17)\quad 0 \to \operatorname{Tor}_1^R(U, Z) \to \operatorname{Tor}_1^R(U, N) \xrightarrow{\gamma} U \otimes_R F \to U \otimes_R Z \to U \otimes_R N \to 0.$$

Let $\mathfrak{p}$ be a minimal prime ideal of R. Then it follows from (3.2.4) that

$$\operatorname{Tor}_i^R(U, N)_{\mathfrak{p}} = 0 \quad \text{for all } i \gg 0,$$

and hence (3.2.16) shows that $\operatorname{Tor}_i^R(U, Z)_{\mathfrak{p}} = 0$ for all $i \gg 0$. So it follows from §2.8(vii) that $\operatorname{Tor}_i^R(U, Z)_{\mathfrak{p}} = 0$ for all $i \geq \text{CI-dim}_{R_{\mathfrak{p}}}(Z_{\mathfrak{p}}) + 1$. Also, by §2.8(ii), we know $\text{CI-dim}_{R_{\mathfrak{p}}}(Z_{\mathfrak{p}}) \leq \text{CI-dim}_R(Z) = 0$. So we see that $\operatorname{Tor}_i^R(U, Z)_{\mathfrak{p}} = 0$ for all $i \geq 1$. As $\mathfrak{p}$ is an arbitrary minimal prime ideal of R, this argument shows that $\operatorname{length}_R(\operatorname{Tor}_i^R(U, Z)) < \infty$ for all $i \geq 1$, as claimed.

Note that it follows from (3.2.16) that $\operatorname{length}_R(\operatorname{Tor}_i^R(U, N)) < \infty$ for all $i \geq 2$. In particular, $\operatorname{Tor}_i^R(U, N)$ is torsion for all $i \geq 2$. However, this forces $\operatorname{Tor}_i^R(U, N)$ to be torsion for each $i \geq 1$; see §2.8(vii). Thus the image of the map γ in (3.2.17) is torsion. On the other hand, $U \otimes_R F$, being a finite direct sum of copies of U, is torsion-free. So $\gamma = 0$, and it follows from (3.2.17) that $\operatorname{Tor}_i^R(U, Z) \cong \operatorname{Tor}_i^R(U, N)$, for $i = 1$ as well. Hence, by (3.2.16), we establish that $\operatorname{Tor}_i^R(U, Z) \cong \operatorname{Tor}_i^R(U, N)$ for all $i \geq 1$.

In light of the fact that $\gamma = 0$, the following exact sequence is induced from (3.2.17):

$$0 \to U \otimes_R F \to U \otimes_R Z \to U \otimes_R N \to 0. \tag{3.2.18}$$

As $M \otimes_R N$ and $U \otimes_R F$ are torsion-free and $U \otimes_R N \cong M \otimes_R N$, we conclude from (3.2.18) that $U \otimes_R Z$ is torsion-fee; see §2.2 and §2.4(i). This completes the proof of Claim 2. □

We finish this section by recording some remarks concerning Theorem 3.2.

Remark 3.3. It is worth mentioning that the conclusion of Theorem 3.2 is not necessarily true if the ring in question is Artinian. For example, if $R = \mathbb{C}[\![x]\!]/(x^2)$ and $M = N = R/(x)$, then R is an Artinian hypersurface (so that each R-module has finite complete intersection dimension and $\theta^R(-, -)$ is well-defined), $\operatorname{cx}_R(N) = 1$, $\operatorname{Tor}_i^R(M, N) \cong N \neq 0$ for all $i \geq 0$, and $\theta^R(M, N) = 0$.

In Section 4 we refer to the next fact to prove Proposition 4.4 and Remark 4.7.

Remark 3.4. Let R be a local ring and let N be a nonzero finitely generated R-module. Assume $\text{CI-dim}_R(N) = 0$ and $\operatorname{cx}_R(N) = 1$. Then it follows that $\beta_i^R(N) = \beta_{i+1}^R(N)$ and $\Omega_R^i(N) \cong \Omega_R^{i+2}(N)$, for all $i \geq 0$; see [Avramov et al. 1997, 7.3].

Note that, in Theorem 3.2, we have $\text{CI-dim}_R(\Omega_R N) = 0$ and $\operatorname{cx}_R(\Omega_R N) = 1$. Therefore Remark 3.4 implies $\beta_2^R(\Omega_R N) = \beta_1^R(\Omega_R N)$, and so $\theta^R(k, \Omega_R N) = 0$; see (3.2.15) in the proof of Theorem 3.2. This gives an alternative way of establishing the vanishing of $\theta^R(k, \Omega_R N)$ without appealing to the property of the Euler number recorded in §2.16.

Next we consider Theorem 3.2 for the case where $\operatorname{cx}_R(N) = 0$, i.e., $\operatorname{pd}_R(N) < \infty$. In this case we can obtain the vanishing of $\operatorname{Tor}_i^R(M, N)$ without any depth assumption on M.

Remark 3.5. Let R be a d-dimensional Cohen–Macaulay local ring, and let M and N be finitely generated R-modules. Assume $\operatorname{pd}_R(N) < \infty$ and $\operatorname{depth}_R(N) \geq d - 1$. If $M \otimes_R N$ is (nonzero) maximal Cohen–Macaulay, then N is free and M is maximal Cohen–Macaulay.

To establish this, we may assume $\operatorname{pd}_R(N) \neq 0$, as otherwise N would be free. In particular, we may assume $d \geq 1$. Note, by the Auslander–Buchsbaum formula and the hypothesis, we have that $\operatorname{pd}_R(N) = 1$. Set $U = \perp_R M$. Then, since U is a torsion-free R-module, [Celikbas and Takahashi 2019, 2.7] implies that $\operatorname{Tor}_i^R(U, N) = 0$ for all $i \geq 1$. Hence §2.4(iii) gives the vanishing of $\operatorname{Tor}_i^R(M, N)$ for all $i \geq 1$. As $\operatorname{pd}_R(N) < \infty$, the depth formula for M and N over R holds; see §2.8(viii). This shows, since $M \otimes_R N$ is maximal Cohen–Macaulay, that both M and N

are maximal Cohen–Macaulay R-modules. Consequently, N is free due to the Auslander–Buchsbaum formula.

4. Some corollaries of the main result

In this section we proceed to give various corollaries of Theorem 3.2 concerning the torsion in tensor products of modules, especially those of the form $M \otimes_R M^*$ over one dimensional local rings. In particular, we give a proof of Corollary 1.4; see Corollaries 4.6 and 4.8. Along the way we extend results of Huneke and Wiegand [1994], and Auslander [1961] on the reflexivity of tensor products of modules which justify a higher dimensional version of Conjecture 1.1 over normal domains; see Proposition 4.12 and Conjecture 4.13.

We denote by $\mathsf{G}(R)$ the *Grothendieck group* of finitely generated R-modules, i.e., the quotient of the free abelian group of all isomorphism classes of finitely generated R-modules by the subgroup generated by the relations coming from short exact sequences of finitely generated R-modules. We write $[M]$ for the class of a finitely generated R-module M in $\mathsf{G}(R)$ and denote by $\overline{\mathsf{G}}(R)$ the group $\mathsf{G}(R)/\mathbb{Z}\cdot[R]$, the reduced Grothendieck group of R. We set $\overline{\mathsf{G}}(R)_{\mathbb{Q}} = (\mathsf{G}(R)/\mathbb{Z}\cdot[R]) \otimes_{\mathbb{Z}} \mathbb{Q}$.

The next corollary corroborates [Celikbas 2011, 1.2], which examines the vanishing of Tor for modules of complexity at most one over complete intersection rings.

Corollary 4.1. *Let R be a one-dimensional local ring, and let M and N be nonzero finitely generated R-modules. Assume* $\text{CI-dim}_R(N) < \infty$, $\text{cx}_R(N) \leq 1$, *and $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime ideal $\mathfrak{p}$ of R.*

(i) *If* $\text{pd}_R(N) < \infty$ *and $M \otimes_R N$ is torsion-free, then M is torsion-free and N is free.*

(ii) *Assume* $\text{pd}_R(N) = \infty$, *or equivalently,* $\text{cx}_R(N) = 1$, *and* $[M] = 0$ *in* $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$. *Then:*

   (a) *If $M \otimes_R N$ is torsion-free, then so are M and N, and* $\text{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$.*

   (b) *If* $\text{Tor}_n^R(M, N) = 0$ *for some $n \geq 1$, then* $\text{Tor}_i^R(M, N) = 0$ *for all $i \geq n$, i.e., the pair (M, N) is Tor-rigid.*

Proof. We may assume R is Cohen–Macaulay; as otherwise N would be free and all the claims follow. In particular, part (i) is a special case of Remark 3.5. Hence we assume $\text{cx}_R(N) = 1$ and $[M] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$, and proceed to prove part (ii).

Let X be a finitely generated R-module. As $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime ideal $\mathfrak{p}$ of R, we have $\text{length}_R(\text{Tor}_i^R(N, X)) < \infty$ for all $i \geq 1$. In particular, $\theta^R(N, X)$ is well-defined; see §2.14. This yields, since θ is additive on short exact sequence of finitely generated R-modules, a linear map $\theta^R(N, -) : \mathsf{G}(R) \to \mathbb{Z}$.

Moreover, as $\theta^R(N, R) = 0$, this map induces a map $\theta^R(N, -) : \overline{\mathsf{G}}(R)_{\mathbb{Q}} \to \mathbb{Q}$. Hence, since $[M] = 0$ in $\overline{G}(R)_{\mathbb{Q}}$, we have that $\theta^R(M, N) = 0$. Now, if $M \otimes_R N$ is torsion-free, then it follows from Theorem 3.2 that M and N are torsion-free, and $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. This establishes part (ii)(a).

To prove part (ii)(b), we proceed by assuming $\operatorname{Tor}_n^R(M, N) = 0$ for some $n \geq 1$. Then we have $\operatorname{Tor}_1^R(M, \Omega_R^{n-1}(N)) = 0$ and $\text{CI-dim}_R(\Omega_R^{n-1}(N)) < \infty$; see §2.10. Hence we use the exact sequence that follows from §2.11 for the module $\Omega_R^{n-1}(N)$:

$$0 \to F \to Z \to \Omega_R^{n-1}(N) \to 0. \tag{4.1.1}$$

Here F is free, $\text{CI-dim}_R(Z) = 0$, and $\operatorname{cx}_R(Z) = \operatorname{cx}_R(\Omega_R^{n-1}(N)) = 1$. Note that, by (2.5.1), we have that $\text{G-dim}_R(Z) = 0$, i.e., Z is totally reflexive and hence Z is torsion-free. Moreover, $\operatorname{Tor}_1^R(M, Z)$ vanishes since

$$\operatorname{Tor}_1^R(M, Z) \hookrightarrow \operatorname{Tor}_1^R(M, \Omega_R^{n-1}(N)).$$

Applying $- \otimes_R Z$ to the syzygy exact sequence $0 \to \Omega_R(M) \to R^{\oplus v} \to M \to 0$, where v is a positive integer, we see that $\Omega_R(M) \otimes_R Z$ is contained in $Z^{\oplus v}$. So, $\Omega_R(M) \otimes_R Z$ is torsion-free. As $[M] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$, it follows from the syzygy sequence that $[\Omega_R(M)] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$. Furthermore, by (4.1.1), we know $Z_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime ideal $\mathfrak{p}$ of R. Hence, by using part (ii)(a) for the pair $(\Omega_R(M), Z)$, we conclude that $\operatorname{Tor}_i^R(\Omega_R(M), Z) = 0$ for all $i \geq 1$. This implies the vanishing of $\operatorname{Tor}_i^R(M, Z)$ for all $i \geq 2$. Consequently, we deduce $\operatorname{Tor}_i^R(M, Z) = 0$ for all $i \geq 1$ since we already know that $\operatorname{Tor}_1^R(M, Z)$ vanishes. This implies that $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq n + 1$, as claimed. □

The next two remarks are concerned with Corollary 4.1.

Remark 4.2. If R is a one-dimensional local ring and M is a finitely generated R-module (not necessarily torsion-free) which has rank, then it follows that $[M] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$; see [Celikbas and Dao 2011, 2.5; Huneke and Wiegand 1994, 1.3]. In view of this fact and Corollary 4.1, we have the following result:

If R is a one-dimensional local ring, and M and N are nonzero finitely generated R-modules such that $\text{CI-dim}_R(N) < \infty$, $\operatorname{cx}_R(N) = 1$, $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime $\mathfrak{p}$ of R, M has rank and $M \otimes_R N$ is torsion-free, then $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, and both M and N are torsion-free.

Remark 4.3. Let R be a one-dimensional complete intersection domain, and let M and N be nonzero finitely generated R-modules.

It is not known whether R-modules are Tor-rigid, in general. Moreover, if $M \otimes_R N$ is torsion-free, it is also not known whether M or N must be torsion-free; see [Celikbas and Wiegand 2015, 2.10]. In fact, Tor-rigidity yields an affirmative answer to the latter query: if $M \otimes_R N$ is torsion-free and N is Tor-rigid, then it follows $\operatorname{Ext}^1_R(\operatorname{Tr} M, N) = 0$ and hence $\operatorname{Ext}^1_R(\operatorname{Tr} M, R) = 0$, i.e., M is torsion-free; see, for example, [Auslander and Bridger 1969, 2.8; Celikbas et al. 2019b, 3.4].

Corollary 4.1 gives a partial affirmative answer to the aforementioned open problems. It points out that modules of complexity at most one, i.e., modules of bounded Betti numbers, are Tor-rigid over R. Moreover, it shows that, if $M \otimes_R N$ is torsion-free, and M or N has complexity at most one, then both M and N are torsion-free (note that, for a one-dimensional local domain R, one has $\overline{\mathsf{G}}(R)_{\mathbb{Q}} = 0$; see, for example, [Celikbas and Dao 2011, 2.5]).

It is known that the conclusion of Corollary 4.1 may fail in case $[M] \neq 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$. For example, if $R = k[\![x, y]\!]/(xy)$, $M = R/(x)$, and $N = R/(x^2)$, then $M_{\mathfrak{p}}$ and $N_{\mathfrak{p}}$ are both free over $R_{\mathfrak{p}}$ for each associated prime ideal $\mathfrak{p}$ of R (as R is reduced) and $M \otimes_R N$ is torsion-free, but $\operatorname{Tor}^R_{2i-1}(M, N) \neq 0 = \operatorname{Tor}^R_{2i}(M, N)$ for all $i \geq 1$; see [Huneke and Wiegand 1997, page 164] and also §A.1. This example also illustrates the following:

Proposition 4.4. *Let R be a one-dimensional local ring, and let M and N be finitely generated R-modules. Assume* $\operatorname{CI-dim}_R(N) < \infty$ *and* $\operatorname{cx}_R(N) = 1$. *Assume further $M_{\mathfrak{p}}$ or $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime $\mathfrak{p}$ of R (e.g., R is reduced). If $M \otimes_R N$ is torsion-free, then* $\operatorname{Tor}^R_{2i}(\perp_R M, N) = 0$ *for all $i \geq 1$.*

Proof. Note that we may assume $M \neq 0 \neq N$. We may further assume R is Cohen–Macaulay, as otherwise M or N would be free and the claim would follow. Moreover, if $M_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime ideal $\mathfrak{p}$ of R, then so is $\perp_R M$. Therefore, we may replace M with $\perp_R M$, and assume M is a nonzero torsion-free R-module; see §2.4(i).

There is a short exact sequence of R-modules of the form

$$0 \to F \to Z \to N \to 0, \tag{4.4.1}$$

where F is free, $\operatorname{CI-dim}_R(Z) = 0$, and $\operatorname{cx}_R(Z) = \operatorname{cx}_R(N)$; see §2.11. By tensoring (4.4.1) with M over R, we obtain an exact sequence,

$$\operatorname{Tor}^R_1(M, N) \xrightarrow{\nu} M \otimes_R F \to M \otimes_R Z \to M \otimes_R N \to 0. \tag{4.4.2}$$

As $\operatorname{Tor}^R_1(M, N)$ is torsion and M is torsion-free, we conclude ν is the zero map and that $M \otimes_R Z$ is torsion-free; see §2.2.

Next we consider a pushforward sequence of Z, i.e., a short exact sequence of R-modules as

$$0 \to Z \to G \to Z_1 \to 0, \tag{4.4.3}$$

where G is free, $\operatorname{CI-dim}_R(Z_1) = 0$, and also $\operatorname{cx}_R(Z_1) = \operatorname{cx}_R(Z) = \operatorname{cx}_R(N) = 1$; see §2.8(iii), §2.10, and §B.3. Notice it follows from Remark 3.4 that $Z_1 \cong \Omega^2_R(Z_1)$.

Therefore, as $Z \cong \Omega_R(Z_1)$, we conclude that $\Omega_R(Z) \cong \Omega_R^2(Z_1) \cong Z_1$. This yields the short exact sequence of R-modules

$$0 \to Z \to G \to \Omega_R(Z) \to 0. \tag{4.4.4}$$

Tensoring (4.4.4) with M over R, we obtain an injection

$$\operatorname{Tor}_1^R(\Omega_R(Z), M) \hookrightarrow M \otimes_R Z.$$

As $M \otimes_R Z$ is torsion-free and $\operatorname{Tor}_1^R(\Omega_R(Z), M)$ is torsion, we see $\operatorname{Tor}_1^R(\Omega_R(Z), M)$ vanishes, i.e., $0 = \operatorname{Tor}_1^R(\Omega_R(Z), M) \cong \operatorname{Tor}_2^R(Z, M)$. This forces $\operatorname{Tor}_{2i}^R(Z, M) = 0$ for all $i \geq 1$ since $Z \cong \Omega_R^2(Z)$. This completes the proof of the proposition: due to (4.4.1), we have that $\operatorname{Tor}_{2i}^R(Z, M) \cong \operatorname{Tor}_{2i}^R(N, M)$ for all $i \geq 1$. □

Our next observation may be of independent interest: the first part examines the complete intersection dimension of a torsion-free module with its algebraic dual over one-dimensional local rings without any depth assumption on their tensor products. The second part of Lemma 4.5 is our first step to establish consequences of Theorem 3.2 concerning Conjecture 1.1 — it is used in the proof of Corollary 4.6.

Lemma 4.5. *Let R be a one-dimensional local ring, and let M be a finitely generated R-module.*

(i) *Assume M is torsion-free. Then*

$$\text{CI-dim}_R(M) < \infty \quad \textit{if and only if} \quad \text{CI-dim}_R(M^*) < \infty.$$

(ii) *Assume* $\text{CI-dim}_R(M) < \infty$, *and $M \otimes_R M^*$ is a nonzero torsion-free R-module. If* $\operatorname{Tor}_i^R(M, M^*) = 0$ *for all $i \gg 0$, then M is free.*

Proof. (i) If $\text{CI-dim}_R(M) < \infty$, then, since $\operatorname{depth}_R(M) = 1$, it follows from §2.8(iii) that $\text{CI-dim}_R(M) = 0$; this implies, in view of §2.8(vi), that $\text{CI-dim}_R(M^*) = 0$. Hence it suffices to assume $\text{CI-dim}_R(M^*) < \infty$ and show that $\text{CI-dim}_R(M) < \infty$.

Assume $\text{CI-dim}_R(M^*) < \infty$. Then (2.7.1) and §2.10 show that

$$\text{CI-dim}_R(\operatorname{Tr} M) < \infty.$$

Let $\mathfrak{p}$ be an associated prime ideal of R. Then, by (2.5.1), we have

$$\text{G-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}) \leq \text{CI-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}),$$

and it follows from §2.8(iii) that $\text{CI-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}) \leq \operatorname{depth}(R_\mathfrak{p}) = 0$. Hence $M_\mathfrak{p}$ is totally reflexive over $R_\mathfrak{p}$. Therefore, it follows that $\operatorname{Ext}_R^1(\operatorname{Tr} M, R)_\mathfrak{p} = 0$. This shows, as $\mathfrak{p}$ is an arbitrary associated prime ideal of R, that $\operatorname{Ext}_R^1(\operatorname{Tr} M, R)$ is a torsion R-module. Now, since $\operatorname{Ext}_R^1(\operatorname{Tr} M, R) \hookrightarrow M$ and as M is torsion-free, it follows that $\operatorname{Ext}_R^1(\operatorname{Tr} M, R) = 0$; see [Auslander and Bridger 1969, 2.8]. Consequently §2.8(iii) shows that $\text{CI-dim}_R(\operatorname{Tr} M) = 0$. Note, up to free summands, we have that $\operatorname{Tr}\operatorname{Tr} M \cong M$. Thus §2.8(vi) implies that $\text{CI-dim}_R(\operatorname{Tr}\operatorname{Tr} M) = 0$, i.e., $\text{CI-dim}_R(M) = 0$, as required.

(ii) Notice, since $\operatorname{Tor}_i^R(M, M^*) = 0$ for all $i \gg 0$, we have that $\operatorname{Tor}_i^R(M, \operatorname{Tr} M) = 0$ for all $i \gg 0$; see (2.7.1). Hence it suffices to prove that $\text{CI-dim}_R(M) = 0$: in this case, §2.8(vii) implies $\operatorname{Tor}_1^R(M, \operatorname{Tr} M) = 0$ so that M is free; by, for example, [Yoshino 1990, 3.9]. Consequently, as $\text{CI-dim}_R(M) \leq \operatorname{depth}(R)$, we may assume $\operatorname{depth}(R) \neq 0$, i.e., we may assume R is a one-dimensional Cohen–Macaulay local ring.

Let $\mathfrak{p}$ be a minimal prime ideal of R. Then $\text{CI-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}) = 0$; see §2.8(iii). Moreover, as $\operatorname{Tor}_i^R(M, M^*)_\mathfrak{p} = 0$ for all $i \gg 0$, we conclude from § 2.8(vii) that $\operatorname{Tor}_i^R(M, M^*)_\mathfrak{p} = 0$ for all $i \geq \text{CI-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}) + 1$. This implies that each $\operatorname{Tor}_i^R(M, M^*)$ has finite length for $i \geq 1$. Thus it follows from [Celikbas et al. 2015b, 3.6] that $\operatorname{Tor}_i^R(M, M^*) = 0$ for all $i \geq 1$. Since $M \otimes_R M^*$ is nonzero and torsion-free, the depth formula implies $\operatorname{depth}_R(M) = 1$; see §2.8(viii). Finally, we conclude by §2.8(iii) that $\text{CI-dim}_R(M) = 0$, as claimed. □

If $R = S/(f)$, where $(S, \mathfrak{n})$ is a two-dimensional regular local ring and $0 \neq f \in \mathfrak{n}$, it follows from a result of Huneke and Wiegand [1994, 3.7] that $M \otimes_R M^*$ has torsion for each nonfree, torsion-free finitely generated R-module M with rank. In particular, Conjecture 1.1 is true over hypersurface rings. In Corollary 4.6, we generalize this fact and show that it carries over to Cohen–Macaulay rings (not necessarily hypersurcases) under mild conditions; see also Corollaries 4.8 and 4.10 for related results.

In Corollary 4.6, we assume $R = S/(f)$, where $(S, \mathfrak{n})$ is a two-dimensional Cohen–Macaulay local ring, and $f \in \mathfrak{n}$ is a non-zero-divisor on S. We assume further M is a finitely generated R-module such that $\operatorname{pd}_S(M) < \infty$. Then, by using the quasi-deformation $R \xrightarrow{=} R \twoheadleftarrow S$, we have that $\text{CI-dim}_R(M) < \infty$; see §2.5. Moreover, [Avramov 1989, 3.2(3)] shows that $\operatorname{cx}_R(X) \leq \operatorname{cx}_S(X) + 1$ for each finitely generated R-module X, in particular, $\operatorname{cx}_R(M) \leq 1$. In Corollary 4.6, we also consider the case where $\operatorname{pd}_S(M^*) < \infty$, i.e., $\operatorname{pd}_S(\operatorname{Hom}_R(M, R)) < \infty$. It is worth noting that, in general, a module can have finite projective dimension, even though its algebraic dual has infinite projective dimension; see, for example, [Huneke and Jorgensen 2003, 2.3].

Corollary 4.6. *Let R be a one-dimensional Cohen–Macaulay ring such that, for some local ring $(S, \mathfrak{n})$ and a non-zero-divisor $f \in \mathfrak{n}$ on S, we have $R = S/(f)$. Let M be a finitely generated R-module. Assume the following hold*:

(a) $\operatorname{pd}_S(M) < \infty$ *or* $\operatorname{pd}_S(M^*) < \infty$.

(b) *$M \otimes_R M^*$ is a torsion-free R-module.*

Then M is free provided that at least one of the following conditions hold:

(i) *$M \otimes_R M^*$ is not zero, and M has rank over R.*

(ii) *M is a torsion-free module that has rank over R.*

(iii) *$M \otimes_R M^*$ is not zero,*

$$\operatorname{length}_R(\operatorname{Tor}_i^R(M, M^*)) < \infty \quad \textit{for all } i \gg 0, \quad \textit{and} \quad \theta^R(M, M^*) = 0.$$

(iv) *M is torsion-free,*

$$\operatorname{length}_R(\operatorname{Tor}_i^R(M, M^*)) < \infty \quad \textit{for all } i \gg 0, \quad \textit{and} \quad \theta^R(M, M^*) = 0.$$

(v) *M is torsion-free, and*

$$\operatorname{length}_R(\operatorname{Tor}_n^R(M, M^*)) = \operatorname{length}_R(\operatorname{Tor}_{n+q}^R(M, M^*)) < \infty$$

for an even integer $n \geq 2$ and an odd integer $q \geq 1$.

Proof. We may assume $M \neq 0$. Recall, if $M \neq 0$ and M is torsion-free, then $M \otimes_R M^* \neq 0$; see §2.3. This implies, for each part, we have that $M \otimes_R M^*$ is a nonzero torsion-free R-module. Furthermore, it shows that part (ii) follows from part (i), and part (iv) follows from part (iii).

Note, as $\operatorname{pd}_S(M) < \infty$ or $\operatorname{pd}_S(M^*) < \infty$, we have $\operatorname{CI-dim}_R(M) < \infty$ or $\operatorname{CI-dim}_R(M^*) < \infty$; see §2.5. Hence it suffices to show $\operatorname{Tor}_i^R(M, M^*) = 0$ for all $i \gg 0$, and M is torsion-free: in that case we can use Lemma 4.5: the first part of the lemma implies $\operatorname{CI-dim}_R(M) < \infty$, and hence the second part shows that M is free.

If $\operatorname{pd}_R(M^*) < \infty$, then M^* is free by the Auslander–Buchsbaum formula. This implies, since $M \otimes_R M^*$ is a nonzero torsion-free R-module, that M is torsion-free. So M must be free. Similarly, if $\operatorname{pd}_R(M) < \infty$, then Remark 3.5 shows that M is free. Moreover, we know that $\operatorname{cx}_R(M) \leq 1$, or $\operatorname{cx}_R(M^*) \leq 1$; see [Avramov 1989, 3.2(3)]. Consequently, we may assume $\operatorname{CI-dim}_R(M) < \infty$ and $\operatorname{cx}_R(M) = 1$, or $\operatorname{CI-dim}_R(M^*) < \infty$ and $\operatorname{cx}_R(M^*) = 1$.

(i) Assume M has rank over R. Then M^* also has rank, which equals the rank of M. In particular, both M and M^* are free when localized at each associated prime ideal of R. Thus Remark 4.2 yields the vanishing of $\operatorname{Tor}_i^R(M, M^*)$ for all $i \geq 1$ and the fact that M is torsion-free.

(iii) Assume $\operatorname{length}_R(\operatorname{Tor}_i^R(M, M^*)) < \infty$ for all $i \gg 0$, and $\theta^R(M, M^*) = 0$. In that case Theorem 3.2 yields the vanishing of $\operatorname{Tor}_i^R(M, M^*)$ for all $i \geq 1$, and that M is torsion-free.

(v) Recall that, by Lemma 4.5(i), we know $\operatorname{CI-dim}_R(M) < \infty$. Let $\mathfrak{p} \in \operatorname{Supp}_R(M)$ be a minimal prime ideal of R. Then it follows

$$\operatorname{CI-dim}_{R_\mathfrak{p}}(M_\mathfrak{p}) = 0 \quad \text{and} \quad \operatorname{Tor}_n^R(M, M^*)_\mathfrak{p} = \operatorname{Tor}_{n+q}^R(M, M^*)_\mathfrak{p} = 0;$$

see §2.8(ii). Therefore, by [Bergh 2008, 3.2], we see $\mathrm{Tor}_i^R(M, M^*)_{\mathfrak{p}} = 0$ for all $i \geq 1$. This implies that $\mathrm{length}_R(\mathrm{Tor}_i^R(M, M^*)) < \infty$ for each $i \geq 1$. We also know that $\mathrm{Tor}_i^R(M, M^*) \cong \mathrm{Tor}_{i+2}^R(M, M^*)$ for each $i \geq 1$; see Remark 3.4. So the hypothesis $\mathrm{length}_R(\mathrm{Tor}_n^R(M, M^*)) = \mathrm{length}_R(\mathrm{Tor}_{n+q}^R(M, M^*))$ implies that

$$\mathrm{length}_R(\mathrm{Tor}_i^R(M, M^*)) = \mathrm{length}_R(\mathrm{Tor}_{i+1}^R(M, M^*)) \quad \text{for each } i \geq 1.$$

Hence $\theta^R(M, M^*) = 0$ and the result follows from part (iv). □

If M and N are finitely generated modules over a local ring R such that M is totally reflexive, then it follows from [Avramov and Buchweitz 2000, 4.4.7] that

$$\widehat{\mathrm{Tor}}_i^R(M, N) \cong \widehat{\mathrm{Ext}}_R^{-i-1}(M^*, N) \quad \text{for all } i \in \mathbb{Z}.$$

Here $\widehat{\mathrm{Tor}}$ and $\widehat{\mathrm{Ext}}$ denote the Tate Tor and Ext, respectively; see, for example, [Avramov and Buchweitz 2000] for details.

In the following, for Gorenstein rings, we provide an alternative proof of Corollary 4.6(iv) that does not appeal to Theorem 3.2.

Remark 4.7. Let R be a one-dimensional Cohen–Macaulay ring such that, for some local ring $(S, \mathfrak{n})$ and $f \in \mathfrak{n}$ a non-zero-divisor on S, we have $R = S/(f)$. Let M be a nonzero finitely generated torsion-free R-module such that $\mathrm{pd}_S(M) < \infty$ or $\mathrm{pd}_S(M^*) < \infty$.

If $\mathrm{pd}_R(M) < \infty$, then M is free by the Auslander–Buchsbaum formula. So we may assume $\mathrm{pd}_R(M) = \infty$. Then it follows that $\text{CI-dim}_R(M) = 0$ and $\mathrm{cx}_R(M) = 1$; see Lemma 4.5(i). Then we have

$$\widehat{\mathrm{Tor}}_i^R(M, N) \cong \widehat{\mathrm{Tor}}_{i+2}^R(M, N) \quad \text{and} \quad \widehat{\mathrm{Ext}}_R^i(M, N) \cong \widehat{\mathrm{Ext}}_R^{i+2}(M, N) \quad \text{for all } i \in \mathbb{Z};$$

see Remark 3.4. Therefore, as M is totally reflexive, we have the following isomorphisms for all $i \geq 1$:

$$\mathrm{Tor}_{2i-1}^R(M^*, M) \cong \widehat{\mathrm{Tor}}_{2i-1}^R(M^*, M) \cong \widehat{\mathrm{Ext}}_R^{-2i}(M, M) \cong \widehat{\mathrm{Ext}}_R^{2i}(M, M) \cong \mathrm{Ext}_R^{2i}(M, M),$$

$$\mathrm{Tor}_{2i}^R(M^*, M) \cong \widehat{\mathrm{Tor}}_{2i}^R(M^*, M) \cong \widehat{\mathrm{Ext}}_R^{-2i-1}(M, M) \cong \widehat{\mathrm{Ext}}_R^{2i-1}(M, M) \cong \mathrm{Ext}_R^{2i-1}(M, M).$$

In particular, if $\mathrm{Tor}_{2i-1}^R(M^*, M) = 0$ for some $i \geq 1$, then M is free; see [Avramov and Buchweitz 2000, 4.2].

Now assume R is Gorenstein and $\mathrm{length}_R(\mathrm{Tor}_i^R(M, M^*)) < \infty$ for all $i \gg 0$. Then it follows $\mathrm{length}_R(\mathrm{Tor}_i^R(M, M^*)) < \infty$ for all $i \geq 1$; see §2.8(ii, viii). In particular, we have that $\mathrm{length}_R(\mathrm{Ext}_R^1(M, M)) < \infty$.

Now assume $M \otimes_R M^*$ is torsion-free. Then [Huneke and Jorgensen 2003, 5.9] implies $\mathrm{Ext}_R^1(M, M) = 0$. Therefore, $M \otimes_R M^*$ is torsion-free if and only if $\mathrm{Tor}_{2i}^R(M^*, M) = 0$ for all $i \geq 1$. Consequently, if $\theta^R(M, M^*) = 0$ and $M \otimes_R M^*$ is torsion-free, then $\mathrm{Tor}_j^R(M, M^*) = 0$ for all $j \geq 1$, and hence M is free, by, for example, Lemma 4.5(ii).

If $(S, \mathfrak{n})$ is a Cohen–Macaulay local ring and $f \in \mathfrak{n}^2$ is a non-zero-divisor on S, then f has a reduced matrix factorization (φ, ψ) over S. In this case, $\operatorname{coker}(\varphi)$ is a nonfree, maximal Cohen–Macaulay module over $S/(f)$ which has projective dimension one over S; see [Herzog et al. 1991, 1.2 and 2.2].

A local ring S is called *G-regular* [Takahashi 2008] if each totally reflexive S-module is free. It is known that each regular ring, as well as each Golod ring that is not a hypersurface, is G-regular. In particular, every non-Gorenstein Cohen–Macaulay local ring with minimal multiplicity is G-regular; see [Takahashi 2008, 5.1]. Note that, if $R = S/(f)$, where $(S, \mathfrak{n})$ is a G-regular ring, and $f \in \mathfrak{n}^2$ is a non-zero-divisor on S, then R is not G-regular; see [Takahashi 2008, 4.6].

The following, advertised in Corollary 1.4, follows from Corollary 4.6 and [Takahashi 2008, 2.13].

Corollary 4.8. *Let $R = S/(f)$ be a one-dimensional Cohen–Macaulay ring, where $(S, \mathfrak{n})$ is a local ring and $f \in \mathfrak{n}^2$ is a non-zero-divisor on S. Assume M is a finitely generated R-module that has rank. Then $M \otimes_R M^*$ has torsion if at least one of the following holds*:

(i) $M = \operatorname{coker}(\varphi)$, *where (φ, ψ) is a reduced matrix factorization of f.*

(ii) *S is G-regular and M is a nonfree totally reflexive R-module.*

Proof. (i) We know M, the cokernel of φ, is a nonfree, torsion-free module over R. Since $\operatorname{pd}_S(M) < \infty$ and M has rank over R, it follows from Corollary 4.6(ii) that $M \otimes_R M^*$ has torsion.

(ii) As M is a totally reflexive R-module, it follows that $\operatorname{G-dim}_S(M) < \infty$; see, for example, [Takahashi 2008, 1.5(3)(ii)]. Hence, since S is G-regular, we conclude that $\operatorname{pd}_S(M) < \infty$, and so the claim follows from Corollary 4.6(ii). □

Here is an example for which we can employ Corollary 4.8(i); note the ring in question is a complete intersection, but is not a hypersurface; see also [Celikbas 2011, 4.17; Huneke and Wiegand 1994, 3.7].

Example 4.9. Let $R = S/(f)$, where $S = \mathbb{C}[[x, y, z]]/(xz - y^2)$ and $f = x^3 - z^2$. Then it follows that $R \cong \mathbb{C}[[t^4, t^5, t^6]]$ is a one-dimensional local domain. Moreover,

$$(\varphi, \psi) = \left(\begin{pmatrix} -z & x \\ x^2 & -z \end{pmatrix}, \begin{pmatrix} z & x \\ x^2 & z \end{pmatrix} \right)$$

is a reduced matrix factorization of f over S. Therefore, by Corollary 4.8(i), the tensor product $M \otimes_R M^*$ has torsion, where M is the finitely generated R-module given by the exact sequence $0 \to S^{\oplus 2} \xrightarrow{\varphi} S^{\oplus 2} \to M \to 0$.

Next is a reformulation of Corollary 4.6(i, ii); it shows that Conjecture 1.1 is true for modules that have finite complete intersection dimension and bounded Betti numbers. We separate this result for the convenience of the reader as it is stated slightly different to Corollary 4.6.

Corollary 4.10. *Let R be a one-dimensional local ring, and let M be a non-free finitely generated R-module that has rank (e.g., R is a domain). Assume* $\mathrm{CI\text{-}dim}_R(M) < \infty$ *and* $\mathrm{cx}_R(M) \le 1$ *(e.g., R is a hypersurface ring). If $M \otimes_R M^*$ is not zero, then $M \otimes_R M^*$ has torsion. In particular, if M is torsion-free, then $M \otimes_R M^*$ has torsion.*

Proof. Note that, since M is not free but has rank, R is a Cohen–Macaulay ring. Note also that, if M is torsion-free, then $M \otimes_R M^* \neq 0$; see §2.3. Hence it suffices to assume $M \otimes_R M^*$ is not zero and prove that $M \otimes_R M^*$ has torsion.

Suppose $M \otimes_R M^*$ is a nonzero torsion-free R-module, and seek a contradiction. It follows from Remark 3.5 that $\mathrm{pd}_R(M) = \infty$, i.e., $\mathrm{cx}_R(M) = 1$. Then we may choose a codimension one quasi-deformation $R \to R' \twoheadleftarrow S$ with zero-dimensional closed fibre such that $\mathrm{pd}_S(M \otimes_R R') < \infty$; see §2.14. Thus $R' \cong S/(f)$ for some local ring $(S, \mathfrak{n})$, and a non-zero-divisor $f \in \mathfrak{n}$ on S. Moreover, it follows that R' is a one-dimensional Cohen–Macaulay ring, $M' = M \otimes_R R'$ has rank over R', and $M' \otimes_{R'} (M')^* \neq 0$. Now Corollary 4.6(i), applied to the module M' over R', shows that M' is free over R', which implies M is free over R, i.e., a contradiction. Hence, if $M \otimes_R M^*$ is not zero, then $M \otimes_R M^*$ must have torsion. $\square$

Further remarks related to Conjecture 1.1. Huneke and Wiegand, motivated by a theorem of Auslander [1961, 3.3], proved that, if R is a local domain satisfying Serre's condition (S_2) such that $R_{\mathfrak{p}}$ is a hypersurface for each height-one prime ideal $\mathfrak{p}$ of R, and M is a finitely generated torsion-free R-module such that $M \otimes_R M^*$ is reflexive, then M is free; see [Huneke and Wiegand 1994, 5.2]. In this subsection we slightly strengthen this result, and show that it holds for R-modules M (not necessarily torsion-free) such that $M \otimes_R M^*$ is nonzero and reflexive; see Proposition 4.12. We also discuss a higher dimensional version of Conjecture 1.1; see Conjecture 4.13 and also Proposition 4.14.

We proceed with a lemma:

Lemma 4.11. *Let R be a local ring, and let M be a finitely generated R-module such that $M^* \neq 0$. If $\mathrm{Ext}^1_R(\mathrm{Tr}\, M, M^*) = \mathrm{Ext}^2_R(\mathrm{Tr}\, M, M^*) = 0$, then M is free.*

Proof. There is an exact sequence $0 \to M^* \to F \to G \to \mathrm{Tr}\, M \to 0$, where F and G are finitely generated free R-modules; see (2.7.1). This yields the following exact sequences:

$$(4.11.1) \qquad 0 \to M^* \to F \to L \to 0 \quad \text{and} \quad 0 \to L \to G \to \mathrm{Tr}\, M \to 0.$$

As $0 = \mathrm{Ext}^2_R(\mathrm{Tr}\, M, M^*) \cong \mathrm{Ext}^1_R(L, M^*)$, the exact sequence $0 \to M^* \to F \to L \to 0$ splits; this implies M^* is free. Since $\mathrm{Ext}^1_R(\mathrm{Tr}\, M, M^*) = \mathrm{Ext}^2_R(\mathrm{Tr}\, M, M^*) = 0$, we conclude that $\operatorname{Ext}^1_R(\operatorname{Tr} M, R) = \operatorname{Ext}^2_R(\operatorname{Tr} M, R) = 0$. Therefore, the natural map $M \to M^{**}$ is bijective, i.e., M is reflexive. Note, as M^* is free, so is M^{**}. Consequently, we deduce that M is free. □

Note that, if R is a local ring and M is a finitely generated R-module such that M has positive rank, then $\operatorname{Supp}_R(M) = \operatorname{Spec}(R)$: to see this, notice, given $\mathfrak{q} \in \operatorname{Spec}(R)$, there is a minimal prime ideal $\mathfrak{p}$ of R such that $\mathfrak{p} \subseteq \mathfrak{q}$. As $M_{\mathfrak{p}} \neq 0$, we conclude that $M_{\mathfrak{q}} \neq 0$. In particular, if R is a local ring and M is a finitely generated R-module such that M has rank r and $M^* \neq 0$, then $r \geq 1$ and M^* has rank r, so that $\operatorname{Supp}_R(M) = \operatorname{Supp}_R(M^*) = \operatorname{Spec}(R)$; see §2.3. We make use of this observation in the next results.

Proposition 4.12. *Let R be a local ring satisfying Serre's condition (S_2) and let M be a finitely generated R-module such that $M^* \neq 0$ and $M \otimes_R M^*$ is reflexive. Then M is free if one of the following conditions holds*:

(i) *$M_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each prime ideal $\mathfrak{p}$ of R of height at most one.*

(ii) *M has rank, and $R_{\mathfrak{p}}$ is a hypersurface ring for each height-one prime ideal $\mathfrak{p}$ of R.*

Proof. For part (i), we may assume $\dim(R) \geq 2$. Consider the following exact sequence which follows from [Auslander and Bridger 1969, 2.6(a)]:

$$0 \to \operatorname{Ext}^1_R(\operatorname{Tr} M, M^*) \to M \otimes_R M^* \xrightarrow{\phi} \operatorname{Hom}_R(M^*, M^*) \to \operatorname{Ext}^2_R(\operatorname{Tr} M, M^*) \to 0. \tag{4.12.1}$$

It follows by part (i) that $\operatorname{Ext}^1_R(\operatorname{Tr} M, M^*)_{\mathfrak{p}} = 0 = \operatorname{Ext}^2_R(\operatorname{Tr} M, M^*)_{\mathfrak{p}}$. Hence the map $\phi_{\mathfrak{p}}$ is an isomorphism for each prime ideal $\mathfrak{p}$ of R of height at most one. Notice $\operatorname{Hom}_R(M^*, M^*)$ is a torsion-free R-module since M^* is torsion-free. This implies ϕ is an isomorphism; see, for example, [Celikbas and Wiegand 2015, page 446]. Therefore, $\operatorname{Ext}^1_R(\operatorname{Tr} M, M^*) = \operatorname{Ext}^2_R(\operatorname{Tr} M, M^*) = 0$, and hence M is free by Lemma 4.11.

For part (ii), let $\mathfrak{p}$ be a height-one prime ideal of R. Then $M_{\mathfrak{p}}$ has rank over $R_{\mathfrak{p}}$ and $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} M^*_{\mathfrak{p}}$ is a nonzero torsion-free $R_{\mathfrak{p}}$ module. Hence it follows from [Huneke and Wiegand 1994, 3.1] that $M_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$; see also Corollary 4.10. Now part (i) shows that M is free. □

In passing we consider a higher dimensional version of Conjecture 1.1.

Conjecture 4.13. *Let R be a local ring satisfying (S_2) and let M be a finitely generated R-module. If M has rank and $M \otimes_R M^*$ is a nonzero reflexive R-module, then M is free.*

It is known that Conjecture 1.1 can be stated over local rings of arbitrary dimension under extra assumptions; see, for example, [Celikbas and Wiegand 2015, 8.6]. However, we could not find a suitable reference that proves, if Conjecture 1.1

is true, then so is Conjecture 4.13. Next we use Proposition 4.12(i) and give an argument to point out this fact.

Proposition 4.14. *If Conjecture 1.1 is true, then Conjecture 4.13 is also true.*

Proof. Assume Conjecture 1.1 is true. Let R be a d-dimensional local ring satisfying (S_2), and let M be a finitely generated R-module such that M has rank and $M \otimes_R M^*$ is a nonzero reflexive R-module. To show M must be free, we proceed by induction on d.

If $d = 0$, then M is free since M has rank. Hence assume $d = 1$. Then R is a one-dimensional Cohen–Macaulay ring and $M \otimes_R M^*$ is a nonzero torsion-free R-module. Set $U = \perp_R M$. Then U is a nonzero torsion-free R-module that has rank. Moreover, we have that $M \otimes_R M^* \cong U \otimes_R M^*$; see § 2.3. By dualizing the short exact sequence (2.1.1), we obtain the following exact sequence: $0 \to U^* \to M^* \to (\mathsf{T}_R M)^*$. As $(\mathsf{T}_R M)^* = 0$, we see that $M^* \cong U^*$. Thus we have

$$M \otimes_R M^* \cong U \otimes_R M^* \cong U \otimes_R U^*.$$

In particular, $U \otimes_R U^*$ is torsion-free. As Conjecture 1.1 is true, we conclude that U is a free R-module. This forces M to be free; see [Huneke and Wiegand 1994, 1.1].

Next assume $d \geq 2$ and let $\mathfrak{p}$ be a height-one prime ideal of R. Then $R_{\mathfrak{p}}$ is a local ring satisfying (S_2), and $M_{\mathfrak{p}}$ is a finitely generated R-module such that $M_{\mathfrak{p}}$ has rank over $R_{\mathfrak{p}}$ and $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} (M_{\mathfrak{p}})^*$ is a nonzero reflexive $R_{\mathfrak{p}}$-module. Hence the induction hypothesis shows that $M_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$. Consequently Proposition 4.12(i) shows that M is free. □

The fact that maximal Cohen–Macaulay modules are reflexive over Gorenstein rings shows that, if Conjecture 4.13 is true, then so is Conjecture 1.1 over (one-dimensional) Gorenstein rings. However, in general, over Cohen–Macaulay rings (not necessarily Gorenstein), we do not know whether or not this implication is true.

Our next aim is to establish Example 1.2(ii) from the introduction, and hence obtain a new class of ideals that satisfy the torsion conclusion of Conjecture 1.1.

Let R be a Cohen–Macaulay local ring, and let I be an $\mathfrak{m}$-primary ideal of R containing a parameter ideal $\mathfrak{q}$ of R as a reduction. Then I is said to be an *Ulrich ideal* provided that $I^2 = \mathfrak{q}I$ and I/I^2 is a free R/I-module; see [Goto et al. 2014, 1.1]. We refer the reader to [Goto et al. 2014] for details about Ulrich ideals; here we record a few observations about them related to our argument. For our purpose, we only consider Ulrich ideals that are not parameter ideals.

Remark 4.15. If R is a Gorenstein ring and I is an Ulrich ideal of R, then I has bounded Betti numbers, i.e., $\mathrm{cx}_R(I) \leq 1$; see [Goto et al. 2014, 7.4].

Corollary 4.16. *Let R be a one-dimensional complete intersection domain, and let I be an Ulrich ideal of R. Then R/I is a Tor-rigid R-module. Moreover, if M is a finitely generated R-module that has torsion, then $M \otimes_R I$ has torsion.*

Proof. This follows from Corollary 4.1(ii) and Remark 4.15 (note that $\bar{\mathsf{G}}(R)_{\mathbb{Q}} = 0$; see [Celikbas and Dao 2011, 2.5]). □

Example 4.17. Let $R = \mathbb{C}[[t^4, t^5, t^6]] \cong \mathbb{C}[[x, y, z]]/(xz - y^2, x^3 - z^2)$ and let $I = (t^4, t^6)$. Then R is a one-dimensional complete intersection domain, and I is an Ulrich ideal of R; see [Goto et al. 2014, 6.3]. Hence, R/I is Tor-rigid, and $I \otimes_R M$ has torsion for each finitely generated R-module M that has torsion; see Corollary 4.16. This fact can fail if M does not have torsion. For example, letting J be the ideal (t^4, t^5) of R, we have that $I \otimes_R J$ is torsion-free, i.e., $\operatorname{Tor}_2^R(R/I, R/J) = 0$; see [Huneke and Wiegand 1994, 4.3]. Hence, since R/I is Tor-rigid, we conclude that $\operatorname{Tor}_i^R(R/I, R/J) = 0$ for all $i \geq 2$, i.e., $\operatorname{Tor}_i^R(I, J) = 0$ for all $i \geq 1$.

Notice, Remark 4.15, together with Corollary 4.10, establishes Example 1.2(ii) over complete intersection rings. In fact, this result is true over Gorenstein rings that are not necessarily complete intersections. This fact can be shown as follows:

Proposition 4.18. *Let $(R, \mathfrak{m})$ be a one-dimensional Gorenstein local ring. If I is a nonprincipal Ulrich ideal of R, then it follows that $I \cong I^*$, and $I \otimes_R I^*$ has torsion.*

Proof. Note that, since $\mathfrak{q} \subsetneq I$, we conclude from [Goto et al. 2014, 2.6(b)] that I is generated by two elements. Moreover, since $I^2 = \mathfrak{q}I$, there is an exact sequence $0 \to \mathfrak{q}/I^2 \to I/I^2 \to I/\mathfrak{q} \to 0$, where $I/I^2 \cong (R/I)^{\oplus 2}$, $\mathfrak{q}/I^2 \cong R/I$, and $I/\mathfrak{q} \cong R/I$. Thus the multiplicity of I is equal to $2 \cdot \operatorname{length}_R(R/I) = \operatorname{length}_R(I/I^2)$. Hence [Ooishi 1996, 2.3] implies that I is a self-dual R-module, i.e., $I \cong I^*$. This yields the isomorphism $I \otimes_R I \cong I^* \otimes_R I$.

We now follow the idea discussed in the paragraph preceding [Huneke and Wiegand 1994, 4.4] and observe that $I \otimes_R I$ has torsion; this implies that $I \otimes_R I^*$ has torsion, as claimed. We see, by applying $- \otimes_R I$ to the exact sequence $0 \to I \to R \to R/I \to 0$, that there is an exact sequence

$$0 \to \operatorname{Tor}_1^R(R/I, I) \to I \otimes_R I \xrightarrow{\alpha} I \xrightarrow{\beta} I \otimes_R R/I \to 0. \tag{4.18.1}$$

Note that I contains a non-zero-divisor on R (as it contains a parameter ideal). Hence $\operatorname{Tor}_1^R(R/I, I)$ is a torsion R-module. Now suppose $I \otimes_R I$ is torsion-free. Then it follows from (4.18.1) that $\operatorname{Tor}_1^R(R/I, I) = 0$, α is injective, and $I \otimes_R I \cong \ker(\beta) \cong I^2$. In particular, we have that $\mu_R(I \otimes_R I) = \mu_R(I^2)$, where $\mu_R(-)$ denotes the number of elements in a minimal generating set. It follows from Nakayama's lemma that $\mu_R(I \otimes_R I) = \mu_R(I)^2$. Therefore we obtain $\mu_R(I)^2 = \mu_R(I^2)$, which forces I to be principal since $\mu_R(I^2) \leq \mu_R(I)(\mu_R(I) + 1)/2$. Hence, since I is not principal, we conclude that $I \otimes_R I$ has torsion. □

5. On tensor products of totally reflexive modules

Huneke and Wiegand proved that tensor products of two nonfree modules over a local domain — that is a quotient of a regular ring by a nonzero element — cannot be maximal Cohen–Macaulay. In fact, this result is true over such rings that are not domains as long as one of the modules in question has rank; see [Huneke and Wiegand 1994, 3.1].

The main purpose of this section is to prove a consequence of Theorem 3.2 that is somewhat of a different nature. Namely, we would like to show that tensor products of two nonfree totally reflexive modules over a Cohen–Macaulay local domain – that is a quotient of a G-regular ring by a non-zero-divisor – cannot be totally reflexive; see Proposition 5.7. Recall that R is called G-regular [Takahashi 2008] if there are no nonfree totally reflexive R-modules. Since each regular ring is G-regular, and each totally reflexive module is maximal Cohen–Macaulay over Cohen–Macaulay rings, our conclusion may be considered as a G-hypersurface version of the result of Huneke and Wiegand [1994, 3.1] mentioned above.

We start by giving a few examples which illustrate the fact that, in general, tensor products of nonfree totally reflexive modules may or may not be totally reflexive.

Example 5.1. Let $R = \mathbb{C}[\![x, y]\!]/(xy)$ and $M = R/(x)$. Then R is a Gorenstein ring and $M \otimes_R M \cong M$ is totally reflexive.

Recall, over a local ring $(R, \mathfrak{m})$, an element $0 \neq x \in \mathfrak{m}$ is said to be an *exact zero-divisor* [Henriques and Şega 2011] if there exists $y \in R$ such that $(0 :_R x) = (y)$ and $(0 :_R y) = (x)$.

Example 5.2. Let $R = \mathbb{C}[\![x, y, z, w]\!]/(x^2, xy, y^2, z^2, w^2)$. Note that R is not Gorenstein, and z and w are exact zero-divisors on R. Set $M = R/(z)$ and $N = R/(w)$. Then M and N are both totally reflexive R-modules. Moreover, $M \otimes_R N$ is a totally reflexive R-module since

$$\begin{aligned}\text{G-dim}_R(M \otimes_R N) = \text{G-dim}_R(R/(z, w)) &= \text{G-dim}_{R/zR}(R/(z, w)) \\ &= \text{G-dim}_{R/zR}((R/zR)/w(R/zR)) = 0.\end{aligned}$$

Here the second equality follows from [Soto 2000, Corollary on page 53], while the last one is due to the fact that w is an exact zero-divisor on $R/(z)$.

In the next example, we observe that over local rings with $\mathfrak{m}^3 = 0$, the tensor product of two totally reflexive modules, given by a pair of exact zero-divisors, is not totally reflexive.

Example 5.3. Let $(R, \mathfrak{m})$ be a local ring such that $\mathfrak{m}^3 = 0$ and R is not Gorenstein; e.g., $R = \mathbb{C}[\![x, y, z]\!]/(x^2, y^2, z^2, yz)$. Assume $\{x, y\}$ is a pair of distinct exact zero-divisors. Let $M = R/(x)$ and $N = R/(y)$, and consider the following short exact sequence of R-modules:

$$0 \to (x, y)/(x) \to R/(x) \to R/(x, y) \to 0. \tag{5.3.1}$$

Notice $(x, y)/(x) \cong R/(x :_R y)$, and $y \cdot \mathfrak{m}^2 = 0$ so that $y\mathfrak{m} \subseteq (0 :_R y) = (x)$, i.e., $(x :_R y) = \mathfrak{m}$. So, if $\text{G-dim}_R(M \otimes_R N) < \infty$, then (5.3.1) shows

$$\text{G-dim}_R(R/(x :_R y)) = \text{G-dim}_R(R/\mathfrak{m}) < \infty,$$

i.e., R is Gorenstein; see [Christensen 2000, 1.2.9 and 1.4.9]. Hence it follows $\text{G-dim}_R(M \otimes_R N) = \infty$.

It also seems worth noting, even if $M \otimes_R N$ has finite Gorenstein dimension, M or N may not have finite Gorenstein dimension. We record such an example next.

Example 5.4. Let $R = \mathbb{C}[\![x, y, z]\!]/(x^2, xy, y^2)$, $M = R/(xz)$, and let $N = R/(z)$. Then R is not Gorenstein. Moreover, it follows that $M \otimes_R N \cong N$ so that

$$\text{pd}_R(M \otimes_R N) = \text{pd}_R(N) = 1 < \infty$$

since z is a non-zero-divisor on R. We proceed to show that $\text{G-dim}_R(M) = \infty$.

Note that we have the following isomorphisms:

$$(xz) \cong (x) \cong R/(x, y). \tag{5.4.1}$$

The first isomorphism in (5.4.1) between the ideals of R holds since z is a non-zero-divisor on R, while the second one is due to the fact that $(0 :_R x) = (x, y)$. Therefore it follows from (5.4.1) that there is a short exact sequence of R-modules

$$0 \to R/(x, y) \to R \to M \to 0. \tag{5.4.2}$$

Set $T = R/(x, y)$. Then it follows from (5.4.2) that $\text{G-dim}_R(M) < \infty$ if and only if $\text{G-dim}_R(T) < \infty$; see [Christensen 2000, 1.2.9]. Hence it suffices to observe that $\text{G-dim}_R(T) = \infty$.

As z is a non-zero-divisor on R and T, we have that

$$\text{G-dim}_R(T) = \text{G-dim}_{R/(z)}(T/zT);$$

see [Christensen 2000, 1.4.5]. Note also $R/(z) \cong \mathbb{C}[\![x, y]\!]/(x^2, xy, y^2)$ is a non-Gorenstein local ring. Therefore, since T/zT is isomorphic to the residue field of the ring $R/(z)$, we conclude that $\text{G-dim}_{R/(z)}(T/zT) = \infty$; see [Christensen 2000, 1.4.9]. Consequently we deduce $\text{G-dim}_R(M) = \infty$, as claimed.

The next observation is known; see, for example, the proof of [Miller 1998, 1.1]. Recall that, if $R = S/(f)$,where $(S, \mathfrak{n})$ is a local ring and $f \in \mathfrak{n}$ is a non-zero-divisor on S, then it follows that $\text{cx}_R(M) \leq \text{cx}_S(M) + 1$ for each finitely generated R-module M; see [Avramov 1989, 3.2.3].

Remark 5.5. Let $R = S/(f)$, where $(S, \mathfrak{n})$ is a local ring and $f \in \mathfrak{n}$ is a non-zero-divisor on S. Let M and N be finitely generated R-modules such that

$$\operatorname{pd}_S(M \otimes_R N) < \infty.$$

If $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, then it follows that $\operatorname{pd}_R(M) < \infty$ or $\operatorname{pd}_R(N) < \infty$.

To see this, let P and Q be minimal free resolutions of M and N, respectively, over R. Then, as $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, we see that $P \otimes_R Q$ is a minimal free resolution of $M \otimes_R N$. Moreover, for each $n \geq 0$, it follows that

$$\beta_n^R(M \otimes_R N) = \operatorname{rank}(P \otimes_R Q)_n = \sum_{i+j=n} \operatorname{rank}(P_i \otimes_R Q_j) = \sum_{i=0}^{n} \beta_i^R(M)\beta_{n-i}^R(N) \tag{5.5.1}$$

Now, if P and Q are infinite resolutions, then (5.5.1) shows that $\beta_n^R(M \otimes_R N) \geq n+1$ for each $n \geq 0$. However, since $\operatorname{pd}_S(M \otimes_R N) < \infty$, we have that $\operatorname{cx}_R(M \otimes_R N) \leq 1$, i.e., there is a real number A such that $\beta_n^R(M \otimes_R N) \leq A$ for each $n \geq 0$; see §2.6. So, P or Q must a finite complex, i.e., $\operatorname{pd}_R(M) < \infty$ or $\operatorname{pd}_R(N) < \infty$. Furthermore, if $\operatorname{pd}_R(N) < \infty$, it follows from (5.5.1) that $\beta_i^R(M)$ is bounded by a real number for each $i \geq 0$, i.e., $\operatorname{cx}_R(M) \leq 1$.

Next is a corollary of Theorem 3.2 and Remark 5.5.

Corollary 5.6. *Let $R = S/(f)$ be a Cohen–Macaulay ring, where $(S, \mathfrak{n})$ is a local ring and $f \in \mathfrak{n}$ is a non-zero-divisor on S. Let M and N be finitely generated R-modules such that*:

(a) $\operatorname{pd}_S(N) < \infty$ *and* $\operatorname{pd}_S(M \otimes_R N) < \infty$.

(b) *M, N, and $M \otimes_R N$ are maximal Cohen–Macaulay R-modules.*

Then M or N is free provided that at least one of the following holds:

(i) *M has rank and $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime ideal $\mathfrak{p}$ of R.*

(ii) $\dim(R) \geq 2$ *and* $\operatorname{length}_R(\operatorname{Tor}_i^R(M, N)) < \infty$ *for all $i \gg 0$.*

Proof. Note, by Remark 5.5, it suffices to prove $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. Note also that $\operatorname{CI-dim}_R(N) < \infty$ and $\operatorname{cx}_R(N) \leq 1$.

For part (i), since M has rank, we may assume $\dim(R) \geq 1$. First, consider the case where $\dim(R) = 1$. Then, if $\operatorname{pd}_R(N) < \infty$, Remark 3.5 shows that N is free. Hence we assume $\operatorname{cx}_R(N) = 1$. In that case Remark 4.2 yields $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$.

Next assume $\dim(R) \geq 2$. Localizing at a nonmaximal prime ideal $\mathfrak{p}$ of R, we see that the hypotheses are preserved. Hence, by the induction hypothesis, we have that $M_{\mathfrak{p}}$ or $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$. In particular, $\operatorname{length}_R(\operatorname{Tor}_i^R(M, N)) < \infty$ for all

$i \geq 1$. Thus, to prove part (i), it suffices to prove part (ii). The fact that part (ii) is an immediate consequence of Theorem 3.2 completes the proof. □

Here is the main result of this section: recall, if $R = S/(f)$, where $(S, \mathfrak{n})$ is a G-regular local ring and $f \in \mathfrak{n}^2$ is a non-zero-divisor on S, then R is not G-regular; see [Takahashi 2008, 4.6].

Proposition 5.7. *Let $R = S/(f)$, where $(S, \mathfrak{n})$ is a Cohen–Macaulay G-regular local ring and $f \in \mathfrak{n}^2$ is a non-zero-divisor on S. Let M and N be finitely generated, nonfree, and totally reflexive R-modules.*

If M has rank and $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for each associated prime $\mathfrak{p}$ of R, then $M \otimes_R N$ is not totally reflexive. In particular, if R is a domain, then $M \otimes_R N$ is not totally reflexive.

Proof. If $M \otimes_R N$ is totally reflexive, then M, N, and $M \otimes_R N$ have finite projective dimension over S; see, for example, [Takahashi 2008, 1.5(4)]; in that case Corollary 5.6 implies that M or N is free. Therefore, $M \otimes_R N$ is not totally reflexive. □

Example 5.8. Let $R = S/(x^2+y^2+z^2+w^2)$, where $S = \mathbb{C}[\![x, y, z, w]\!]/(xy, yz, zw)$. Note that R is reduced, $\dim(S) = 2$, and $\{x + y + z, y + z + w\}$ is an S-regular sequence. Moreover, it follows that

$$S/(x + y + z, y + z + w) \cong \mathbb{C}[\![z, w]\!]/(-zw - w^2, -z^2 - zw, zw)$$

is an Artinian ring with radical square zero. Hence, S is G-regular but R is not G-regular; see [Takahashi 2008, 4.2 and 4.6]. Therefore, if M and N are nonfree totally reflexive R-modules, either of which has rank, then Proposition 5.7 shows that $M \otimes_R N$ is not totally reflexive.

We finish this section by proving a result similar to Theorem 3.2: our aim is to show that, in case the module M in Theorem 3.2 is maximal Cohen–Macaulay, then one can prove the vanishing of Tor under weaker assumptions, for example, regardless of the depth of N. Subsequently, we give an application of our result concerning tensor products of totally reflexive modules over hypersurfaces; see Corollary 5.12.

Note, by §2.13, $\theta^R(M, N)$ is well-defined under the hypotheses of Proposition 5.9.

Proposition 5.9. *Let R be a Cohen–Macaulay local ring of dimension $d \geq 1$ such that $R = S/(f)$ for some local ring $(S, \mathfrak{n})$ and a non-zero-divisor $f \in \mathfrak{n}$ on S. Let M be a (finitely generated) maximal Cohen–Macaulay R-module, and let N be a finitely generated R-module. Assume the following conditions hold:*

(i) $\text{CI-dim}_S(N) < \infty$ *and* $\operatorname{Tor}_i^S(M, N) = 0$ *for all* $i \gg 0$ (*e.g.*, $\operatorname{pd}_S(N) < \infty$).

(ii) $\operatorname{length}_R(\operatorname{Tor}_i^R(M, N)) < \infty$ *for all* $i \gg 0$.

(iii) $\theta^R(M, N) = 0$.

If $M \otimes_R N$ *is torsion-free, then* $\operatorname{Tor}_i^R(M, N) = 0$ *for all* $i \geq 1$.

Proof. We start by noting $\text{CI-dim}_R(N) < \infty$; see §2.8(iv). Hence $\text{G-dim}_R(N) < \infty$ so that there is an exact sequence of finitely generated R-modules

$$0 \to L \to Z \to N \to 0, \tag{5.9.1}$$

where Z is a totally reflexive R-module and $\operatorname{pd}_R(L) < \infty$; see (2.5.1) and §2.9.

As $\operatorname{pd}_R(L) < \infty$, it follows that $\operatorname{pd}_S(L) < \infty$ [Rotman 1979, 9.32]. Hence, by (5.9.1), we have $\text{CI-dim}_S(Z) < \infty$ as $\text{CI-dim}_S(N) < \infty$; see § 2.10. Thus $\text{CI-dim}_S(Z) = 1$ and $\text{CI-dim}_R(Z) = 0$; see §2.8(iii). Moreover, since $\operatorname{pd}_S(L) < \infty$ and $\operatorname{Tor}_i^S(M, N) = 0$ for all $i \gg 0$, it follows from (5.9.1) that $\operatorname{Tor}_i^S(M, Z) = 0$ for all $i \gg 0$. Consequently $\operatorname{Tor}_i^S(M, Z) = 0$ for all $i \geq 2$; see §2.8(vii). Note also, as $\operatorname{length}_R(\operatorname{Tor}_i^R(M, N)) < \infty$ for all $i \gg 0$, we have that $\operatorname{Tor}_i^R(M, N)$ is torsion for all $i \geq 1$; see §2.8(vii).

Claim. $M \otimes_R Z$ is torsion-free, and $\operatorname{Tor}_i^R(M, Z) \cong \operatorname{Tor}_i^R(M, N)$ for all $i \geq 1$.

We first show that the claim is sufficient to complete the proof. Note, by the claim, we have that $\operatorname{length}_R(\operatorname{Tor}_i^R(M, Z)) < \infty$ for all $i \gg 0$. Hence the fact $\text{CI-dim}_R(Z) = 0$ forces $\operatorname{length}_R(\operatorname{Tor}_i^R(M, Z)) < \infty$ for all $i \geq 1$; see §2.8(ii, vii). In particular, as $\operatorname{Tor}_i^S(M, Z) = 0$ for all $i \geq 2$ and $\text{CI-dim}_S(Z) < \infty$, we conclude that $\theta^R(M, Z)$ is well-defined; see § 2.13. Hence it follows by the claim $\theta^R(M, Z) = \theta^R(M, N)$ so that $\theta^R(M, Z) = 0$.

As $M \otimes_R Z$ is torsion-free, $\operatorname{Tor}_2^S(M, Z) = 0$ and $\operatorname{Tor}_2^R(M, N)$ is torsion, we use §2.12 for the pair (M, Z), and deduce that $\operatorname{Tor}_2^R(M, Z)$ vanishes. So, in view of (2.13.3), we have $\theta^R(M, Z) = \operatorname{length}_R(\operatorname{Tor}_2^R(M, Z)) - \operatorname{length}_R(\operatorname{Tor}_1^R(M, Z))$ and hence $\operatorname{Tor}_1^R(M, Z) = 0$. Now, as $\operatorname{Tor}_i^R(M, Z) \cong \operatorname{Tor}_{i+2}^R(M, Z)$ for all $i \geq 1$, we see that $\operatorname{Tor}_i^R(M, Z) = 0$ for all $i \geq 1$. Therefore, it remains to justify the claim.

To prove the claim, we will first show $M \otimes_R L$ is torsion-free, or equivalently, $M \otimes_R L$ satisfies (S_1), i.e., $\operatorname{depth}_{R_\mathfrak{p}}(M_\mathfrak{p} \otimes_{R_\mathfrak{p}} L_\mathfrak{p}) \geq \min\{1, \dim(R_\mathfrak{p})\}$ for each $\mathfrak{p} \in \operatorname{Spec}(R)$. Let $\mathfrak{p} \in \operatorname{Supp}_R(M \otimes_R L)$ and assume $\dim(R_\mathfrak{p}) \geq 1$ (recall $\operatorname{depth}(0) = \infty$). Since $\operatorname{Tor}_i^R(M, L) = 0$ for all $i \geq 1$, the equality

$$\operatorname{depth}_{R_\mathfrak{p}}(L_\mathfrak{p}) + \operatorname{depth}_{R_\mathfrak{p}}(M_\mathfrak{p}) = \operatorname{depth}(R_\mathfrak{p}) + \operatorname{depth}_{R_\mathfrak{p}}(L_\mathfrak{p} \otimes_{R_\mathfrak{p}} M_\mathfrak{p})$$

holds; see § 2.8(viii) and § 2.15. So $\operatorname{depth}_{R_\mathfrak{p}}(L_\mathfrak{p}) = \operatorname{depth}_{R_\mathfrak{p}}(L_\mathfrak{p} \otimes_{R_\mathfrak{p}} M_\mathfrak{p})$. Notice (5.9.1) localized at $\mathfrak{p}$ shows that $L_\mathfrak{p} \hookrightarrow Z_\mathfrak{p} \neq 0$. Since $Z_\mathfrak{p}$ is a torsion-free module over $R_\mathfrak{p}$, we have that $\operatorname{depth}_{R_\mathfrak{p}}(L_\mathfrak{p}) \geq 1$. Consequently this shows $\operatorname{depth}_{R_\mathfrak{p}}(M_\mathfrak{p} \otimes_{R_\mathfrak{p}} L_\mathfrak{p}) \geq 1$, and hence $\operatorname{depth}_{R_\mathfrak{p}}(M_\mathfrak{p} \otimes_{R_\mathfrak{p}} L_\mathfrak{p}) \geq \min\{1, \dim(R_\mathfrak{p})\}$ for all $\mathfrak{p} \in \operatorname{Spec}(R)$. In particular, $M \otimes_R L$ is torsion-free.

Now, as $M \otimes_R L$ is torsion-free and $\mathrm{Tor}_1^R(M, N)$ is torsion, by tensoring (5.9.1) with M, we obtain the exact sequence

$$0 \to M \otimes_R L \to M \otimes_R Z \to M \otimes_R N \to 0.$$

This implies $M \otimes_R Z$ is torsion-free; see §2.2. Moreover, as $\mathrm{Tor}_i^R(M, L) = 0$ for all $i \geq 1$, in view of (5.9.1), we conclude that $\mathrm{Tor}_i^R(M, Z) \cong \mathrm{Tor}_i^R(M, N)$ for all $i \geq 1$. This proves the claim and completes the proof. □

We use the next observation to prove Corollary 5.11:

Remark 5.10. Let R be a local ring, M a finitely generated reflexive R-module, and let x be a non-zero-divisor on R. Then M/xM is a torsionless R/xR-module. In particular, M/xM is a torsion-free module over R/xR. One can show this as follows:

Note that, since M is torsionless, there is a short exact sequence

$$0 \to M \to F \to C \to 0,$$

where F is free and $\mathrm{Ext}_R^1(C, R) = 0$; see, for example, §B.3. Dualizing this short exact sequence, we have the following commutative diagram, where λ denotes the natural map:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \longrightarrow & F & \longrightarrow & C & \longrightarrow & 0 \\ & & \cong\downarrow \lambda_M & & \cong\downarrow \lambda_F & & \downarrow \lambda_C & & \\ 0 & \longrightarrow & M^{**} & \longrightarrow & F^{**} & \longrightarrow & C^{**} & \longrightarrow & \mathrm{Ext}_R^1(M^*, R) \end{array}$$

This shows λ_C is injective, i.e., C is torsionless. So this implies that

$$\mathrm{Tor}_1^R(C, R/xR) = 0.$$

Hence the top row yields an injection $M/xM \hookrightarrow F/xF$, as claimed.

Corollary 5.11. *Let R be a Cohen–Macaulay ring of dimension $d \geq 2$ such that $R = S/(f)$ for some local ring $(S, \mathfrak{n})$ and a non-zero-divisor $f \in \mathfrak{n}$ on S. Let M and N be finitely generated R-modules, and assume the following hold*:

(i) $\mathrm{pd}_S(N) < \infty$ *and* $\mathrm{depth}_R(N) \geq 1$.

(ii) *M is a maximal Cohen–Macaulay R-module.*

(iii) $\mathrm{length}_R(\mathrm{Tor}_i^R(M, N)) < \infty$ *for all* $i \gg 0$.

If $M \otimes_R N$ is reflexive, then $\mathrm{Tor}_i^R(M, N) = 0$ *for all* $i \geq 1$.

Proof. Let $x \in \mathfrak{m}$ be a non-zero-divisor on R, M, and N. For any R-module X we let $\overline{X}$ denote X/xX. Then it follows that $\overline{M}$ is a maximal Cohen–Macaulay

$\bar{R}$-module and $\overline{M \otimes_R N} \cong \bar{M} \otimes_{\bar{R}} \bar{N}$, where $\bar{M} \otimes_{\bar{R}} \bar{N}$ is a torsion-free $\bar{R}$-module by Remark 5.10. Now consider the following short exact sequence of R-modules:

$$0 \to N \xrightarrow{x} N \to \bar{N} \to 0. \tag{5.11.1}$$

We see from (iii) and (5.11.1) that $\operatorname{length}_R(\operatorname{Tor}_i^R(M, \bar{N})) < \infty$ for all $i \gg 0$. Hence $\theta^R(M, \bar{N})$ is well-defined, and the additivity of the θ-pairing applied to the short exact sequence (5.11.1) yields $\theta^R(M, N) = \theta^R(M, N) + \theta^R(M, \bar{N})$, i.e., $\theta^R(M, \bar{N}) = 0$; see §2.13.

Write $x = y + (f)$ for some $y \in \mathfrak{n}$. Then $\{y, f\}$ is an S-regular sequence. Hence we can write $\bar{R} = T/(f)$, where $T = S/(y)$ and f is a non-zero-divisor on T contained in the maximal ideal of T.

Notice $\operatorname{Tor}_i^{\bar{R}}(\bar{M}, \bar{N}) \cong \operatorname{Tor}_i^R(M, \bar{N})$ for all $i \geq 0$; this implies $\theta^{\bar{R}}(\bar{M}, \bar{N})$ is well-defined, and hence $\theta^{\bar{R}}(\bar{M}, \bar{N}) = \theta^R(M, \bar{N}) = 0$. Moreover, since y is a non-zero-divisor on S and N, it follows that

$$\operatorname{pd}_T(\bar{N}) = \operatorname{pd}_T(N/xN) = \operatorname{pd}_{S/(y)}(N/yN) = \operatorname{pd}_S(N) < \infty.$$

So we have $\operatorname{pd}_T(\bar{N}) < \infty$, $\operatorname{length}_{\bar{R}}(\operatorname{Tor}_i^{\bar{R}}(\bar{M}, \bar{N})) < \infty$ for all $i \gg 0$, and also $\theta^{\bar{R}}(\bar{M}, \bar{N}) = 0$. Hence we use Proposition 5.9 with the pair $(\bar{M}, \bar{N})$ over the ring $\bar{R} = T/(f)$, and conclude that $\operatorname{Tor}_i^{\bar{R}}(\bar{M}, \bar{N}) = 0$ for all $i \geq 1$. Consequently, by using (5.11.1) and Nakayama's lemma, we see that $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. □

The following result corroborates with the Tor vanishing conclusion of [Huneke and Wiegand 1994, 2.7].

Corollary 5.12. *Let R be a local hypersurface ring, i.e., $R = S/(f)$ for some regular local ring $(S, \mathfrak{n})$ and $0 \neq f \in \mathfrak{n}$. Let M be a nonfree maximal Cohen–Macaulay R-module, and let N be a finitely generated R-module such that $\operatorname{pd}_R(N) = \infty$. Assume $\dim(R) \geq 2$, and $R_\mathfrak{p}$ is a domain for each $\mathfrak{p} \in \operatorname{Spec}(R) - \{\mathfrak{m}\}$. Then $M \otimes_R N$ is not reflexive.*

Proof. We assume $M \otimes_R N$ is reflexive and seek a contradiction. Note that it follows $M \otimes_R \perp_R N$ is reflexive; see §2.4(i).

Pick a prime ideal $\mathfrak{p} \in \operatorname{Supp}_R(M \otimes_R \perp_R N) - \{\mathfrak{m}\}$. Then $M_\mathfrak{p} \otimes_{R_\mathfrak{p}} (\perp_R N)_\mathfrak{p}$ is a reflexive $R_\mathfrak{p}$-module over the hypersurface domain $R_\mathfrak{p}$. So [Huneke and Wiegand 1994, 2.7] implies that $\operatorname{Tor}_i^R(M, \perp_R N)_\mathfrak{p} = 0$ for all $i \geq 1$. This shows that $\operatorname{length}_R(\operatorname{Tor}_i^R(M, \perp_R N)) < \infty$ for all $i \geq 1$.

Now, as $\operatorname{pd}_S(\perp_R N) < \infty$ and $\operatorname{depth}_R(\perp_R N) \geq 1$, from Corollary 5.11 we conclude that $\operatorname{Tor}_i^R(M, \perp_R N) = 0$ for all $i \geq 1$. Consequently, by §2.4(iii), we obtain the vanishing of $\operatorname{Tor}_i^R(M, N)$ for all $i \geq 1$. This forces M or N to have finite projective dimension; see [Huneke and Wiegand 1997, 1.9]. Consequently $M \otimes_R N$ cannot be reflexive. □

The conclusion of Corollary 5.12 can fail over arbitrary hypersurfaces that are not locally domains on the punctured spectrum.

Example 5.13. Let $R = \mathbb{C}[\![x, y, z]\!]/(xy)$, $M = R/(x)$ and let $N = R/(x^2)$. Then R is a two-dimensional hypersurface, M is a nonfree maximal Cohen–Macaulay R-module and $\mathrm{pd}_R(N) = \infty$. Note that, $M \otimes_R N$, being isomorphic to M, is reflexive, and $R_{\mathfrak{p}}$ is not a domain for $\mathfrak{p} = (x, y) \in \mathrm{Spec}(R)$.

It is worth noting that totally reflexive modules over a ring as in Corollary 5.12 cannot be defined by exact zero-divisors.

Remark 5.14. Let $(R, \mathfrak{m})$ be a local ring such that $\mathrm{depth}(R) \geq 2$ and $R_{\mathfrak{p}}$ is a domain for each $\mathfrak{p} \in \mathrm{Spec}(R) - \{\mathfrak{m}\}$. Then R does not admit exact zero divisors.

To see this, suppose $x \in \mathfrak{m}$ is an exact zero divisor on R, and seek a contradiction. It follows from the definition that there is $0 \neq y \in \mathfrak{m}$ such that $(0 :_R x) = (y)$ and $(0 :_R y) = (x)$; i.e., the following is the minimal free resolution of R/xR over R:

$$\cdots \to R \xrightarrow{x} R \xrightarrow{y} R \xrightarrow{x} R \xrightarrow{y} R \xrightarrow{x} R \to R/xR \to 0.$$

First assume that there is a prime ideal $\mathfrak{p} \in \mathrm{Spec}(R) - \{\mathfrak{m}\}$ such that $x, y \in \mathfrak{p}$. Then, by localizing the minimal free resolution of R/xR at $\mathfrak{p}$, we obtain the minimal free resolution of $(R/xR)_{\mathfrak{p}}$ over $R_{\mathfrak{p}}$:

$$\cdots \to R_{\mathfrak{p}} \xrightarrow{x} R_{\mathfrak{p}} \xrightarrow{y} R_{\mathfrak{p}} \xrightarrow{x} R_{\mathfrak{p}} \xrightarrow{y} R_{\mathfrak{p}} \xrightarrow{x} R_{\mathfrak{p}} \to R_{\mathfrak{p}}/xR_{\mathfrak{p}} \to 0.$$

In particular, we see that (x, y) is also pair of exact zero divisors on $R_{\mathfrak{p}}$. However, this is not possible since $R_{\mathfrak{p}}$ is a domain and local domains cannot admit exact zero divisors.

Now let I be the ideal of R generated by x and y. Suppose $\mathfrak{m}$ is minimal prime over I. Then $2 \leq \mathrm{depth}(R) \leq \dim(R) = \mathrm{height}_R(\mathfrak{m}) \leq 2$, i.e., R is Cohen–Macaulay of dimension two, and that $\mathrm{height}_R(I) = 2$. This implies that $\{x, y\}$ is a regular sequence on R, which is not possible. Therefore $\mathfrak{m}$ is not minimal over I. Consequently, there is a prime ideal $\mathfrak{p} \in \mathrm{Spec}(R) - \{\mathfrak{m}\}$ such that $x, y \in \mathfrak{p}$; this gives a contradiction by the previous argument.

Appendix A: On the vanishing of the theta invariant

Recall that, if R is a one-dimensional reduced hypersurface ring, then $\theta^R(M, N)$ is defined and vanishes for all finitely generated R-modules M and N, either of which has rank; see Remark 4.2. Since Theorem 3.2 relies upon the vanishing of θ pairing, we would like to find out whether θ can vanish nontrivially. More precisely, we would like to find out whether there is a one-dimensional reduced hypersurface ring R, and modules M and N over R — neither of which has rank — such that $\theta^R(M, N) = 0$. We were unable to find an example (or a result) from the literature that addresses our query. The aim of this section is to record such an example suggested to us by Hailong Dao; see Example A.3. First, in §A.1, we will record a related fact that was shown to us by Mark Walker: over one-dimensional reduced local rings R, a finitely generated R-module M has rank if and only if its class is zero in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$. A similar result that makes use of θ pairing is established in [Dao 2013, 3.3] for hypersurface rings.

A.1. Let R be a one-dimensional Cohen–Macaulay local ring, and let M be a finitely generated R-module.

(i) There exists a rational number r such that $\text{length}_{R_\mathfrak{p}}(M_\mathfrak{p}) = r \cdot \text{length}(R_\mathfrak{p})$ for each associated prime ideal $\mathfrak{p}$ of R if and only if $[M] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$.

(ii) Assume R is reduced. Then M has rank if and only if $[M] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$.

Proof. (i) Let $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_n\}$ be the set of all minimal (associated) prime ideals of R. Note that $\mathsf{G}(k) = \mathbb{Z} \cdot [k]$ and $\mathsf{G}(R_{\mathfrak{p}_j}) = \mathbb{Z} \cdot [k(\mathfrak{p}_j)]$, where $k(\mathfrak{p}_j)$ is the residue field of $R_{\mathfrak{p}_j}$, for all $j = 1, \ldots, n$.

There is a right exact sequence of the form

$$\mathsf{G}(k) \xrightarrow{\alpha} \mathsf{G}(R) \xrightarrow{\beta} \bigoplus_{j=1}^{n} \mathsf{G}(R_{\mathfrak{p}_j}) \to 0. \tag{A.1.1}$$

Here α is the natural inclusion with

$$\alpha([k]) = [k] \quad \text{and} \quad \beta([M]) = \big(\text{length}_{R_{\mathfrak{p}_j}}(M_{\mathfrak{p}_j})[k(\mathfrak{p}_j)]\big)_j.$$

In (A.1.1), by identifying $\mathsf{G}(k)$ with $\mathbb{Z}$, and $\bigoplus_{j=1}^{n} \mathsf{G}(R_{\mathfrak{p}_j})$ with $\mathbb{Z}^{\oplus n}$, we obtain another right exact sequence of the form

$$\mathbb{Z} \xrightarrow{\alpha} G(R) \xrightarrow{\beta} \mathbb{Z}^{\oplus n} \to 0, \tag{A.1.2}$$

where $\alpha(1) = [k]$ and $\beta([M]) = \big(\text{length}_{R_{\mathfrak{p}_j}}(M_{\mathfrak{p}_j})\big)_j$. Applying $- \otimes_{\mathbb{Z}} \mathbb{Q}$ to (A.1.2), we see there is a right exact sequence of the form

$$\mathbb{Q} \xrightarrow{\alpha \otimes 1} \mathsf{G}(R)_{\mathbb{Q}} \xrightarrow{\beta \otimes 1} \mathbb{Q}^{\oplus n} \to 0. \tag{A.1.3}$$

Here $\alpha \otimes 1(1) = [k]$, which is zero in $\mathsf{G}(R)_{\mathbb{Q}}$. Hence $\alpha \otimes 1$ is the zero map so that $\beta \otimes 1$ is an isomorphism.

Consequently $[M] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$ if and only if $[M] = r \cdot [R]$ for some rational number r if and only if $\beta \otimes 1([M]) = r \cdot \beta \otimes 1([R])$ if and only if

$$\text{length}_{R_{\mathfrak{p}_j}}(M_{\mathfrak{p}_j}) = r \cdot \text{length}(R_{\mathfrak{p}_j}) \quad \text{for all } j = 1, \ldots, n.$$

(ii) If M has rank, then $[M]=0$ in $\overline{G}(R)_{\mathbb{Q}}$; see Remark 4.2. To see the converse, let $\mathfrak{p}$ be an associated prime ideal of R. Then, by (i), we have

$$\text{length}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = r \cdot \text{length}(R_{\mathfrak{p}})$$

for some rational number r. Since $M_{\mathfrak{p}} \cong R_{\mathfrak{p}}^{\oplus n}$ for some positive integer n, we see $n=r$ and hence M has rank r. □

The next example shows that the conclusion of § A.1(ii) can fail if R is not reduced. It also shows that Conjecture 1.1 would fail if the module M in question does not have rank, even if $[M]=0$ in $\overline{G}(R)_{\mathbb{Q}}$.

Example A.2. Let $R=\mathbb{C}[\![x,y]\!]/(x^2)$ and let $M=R/(x)$. Then $M\cong M^*$, and so M is torsion-free. The exact sequence $0\to M\to R\to M\to 0$ implies that $[M]=0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$. Moreover, M does not have rank. Note also that $\text{Tor}_i^R(M,M^*)\cong M$ for all $i\geq 0$, and hence $\text{length}(\text{Tor}_i^R(M,M^*))=\infty$ for all $i\geq 0$.

Here is an example we seek on the vanishing of θ invariant.

Example A.3. Let

$$R=\mathbb{C}[\![x,y]\!]/(xy(x-y)), \quad M=R/(x) \quad \text{and,} \quad N=M\oplus R/(y)\oplus R/(y).$$

Then R is a one-dimensional reduced hypersurface ring, and M and N are nonfree, finitely generated, torsion-free R-modules.

The minimal free resolution of M is given by

$$F= \ \cdots \xrightarrow{(x-y)y} R \xrightarrow{x} R \xrightarrow{(x-y)y} R \xrightarrow{x} R \longrightarrow 0 \ .$$

Thus $\text{Tor}_1^R(M,M)\cong k[\![y]\!]/(y^2)$ and $\text{Tor}_2^R(M,M)=0$ so that $\theta^R(M,M)=-2$. Similarly one can check $\text{Tor}_1^R(R/(y),R/(y))\cong k[\![x]\!]/(x^2)$ and $\text{Tor}_2^R(R/(y),R/(y))=0$. So it follows $\theta^R(R/(y),R/(y))=-2$. Tensoring F with $R/(y)$, we see

$$\text{Tor}_2^R(M,R/(y))\cong k \quad \text{and} \quad \text{Tor}_1^R(M,R/(y))=0.$$

This yields $\theta^R(M,R/(y))=1$.

Therefore we have

$$\theta^R(N,N)=-6 \quad \text{and} \quad \theta^R(M,N)=\theta^R(M,M)+2\theta^R(M,R/(y))=0.$$

Note that, since $\theta^R(M,M)\neq 0$ and $\theta^R(N,N)\neq 0$, neither M nor N has rank.

Remark A.4. It seems interesting that, contrary to Example A.3, over certain reduced hypersurface rings, $\theta(M,N)$ can vanish only when M and N have rank. For example, if $R=\mathbb{C}[\![x,y]\!]/(xy)$, and M and N are finitely generated R-modules, then one can check that $\theta^R(M,N)$ vanishes if and only if M and N both have rank. Note, by § A.1, one concludes for this particular hypersurface ring R, and

R-modules M and N that, $\theta^R(M, N) = 0$ if and only if $[M] = [N] = 0$ in $\overline{\mathsf{G}}(R)_{\mathbb{Q}}$ if and only if M and N both have rank.

Appendix B: Some examples of torsion-free tensor products

In this section we recall that Conjecture 1.1 may fail if one considers the tensor product $M \otimes_R M$ instead of $M \otimes_R M^*$. Huneke and Wiegand showed that, if R is a one-dimensional local domain that is not Gorenstein, then there exists a finitely generated torsion-free module R-module M such that M is not free and $M \otimes_R M$ is torsion-free; see the proof of [Huneke and Wiegand 1994, 4.7]. However their argument seems to not yield an explicit example of such a module M. Building on the proof of Huneke and Wiegand, we will point out how to construct nonfree torsion-free R-modules M with rank such that $M \otimes_R M$ is torsion-free over certain one-dimensional local rings R.

B.1. Let R be a one-dimensional Cohen–Macaulay local ring with canonical module ω. Set $M = \operatorname{Tr} \Omega \operatorname{Tr} \Omega\omega$. If R is generically Gorenstein but not Gorenstein, then M is a finitely generated, nonfree, torsion-free R-module with rank such that $M \otimes_R M$ is torsion-free.

Proof. It follows from [Auslander and Bridger 1969, 2.21] that there is an exact sequence of the form

$$0 \to F \to M \oplus G \to \omega \to 0, \tag{B.1.1}$$

where F and G are finitely generated free R-modules. In particular, M and M^* are torsion-free modules such that M has rank and M^* is nonzero. As syzygy modules are torsionless, we have $\operatorname{Ext}^1_R(M, R) = 0$. It follows that $M \otimes_R \omega$ is torsion-free, and the sequence (B.1.1) does not split; see [Avramov et al. 2005, B.4; Araya et al. 2018, 2.5]. Now tensoring (B.1.1) with M, we see that $M \otimes_R M$ is torsion-free; see §2.2. □

The observation in §B.1 raises the following question; an affirmative answer to this question yields a counterexample to Conjecture 1.1.

Question B.2. Is there a one-dimensional, generically Gorenstein, Cohen–Macaulay local ring R with a canonical module $\omega \ncong R$ such that $(\operatorname{Tr} \Omega \operatorname{Tr} \Omega\omega)^* \cong \operatorname{Tr} \Omega \operatorname{Tr} \Omega\omega$?

Modules yielding torsion-free tensor products as in §B.1 can also be obtained without appealing to the short exact sequence involving the transpose. Such a module can be realized as the pushforward of the first syzygy of the canonical module of the ring R. We observe this below by including a few additional details to the argument of [Huneke and Wiegand 1994, 4.7].

B.3 [Huneke and Wiegand 1994, 4.7]. Let M be a finitely generated R-module, and let $\pi : F \to M^*$ be a minimal free presentation of M^*. Denote $\mu : M \to F^*$ by the composition of the natural map $\delta_M : M \to M^{**}$ and $\pi^* : M^{**} \to F^*$. Then μ^* is surjective, and the cokernel of μ, denoted by $\mathrm{PF}(M)$, is called the *pushforward* of M (pushforward is unique up to free summands; see, for example, [Celikbas 2011, page 174]).

Now assume M is torsion-free and $\mathrm{Ext}^1_R(M, R) \neq 0$. Take a minimal generating set $\alpha_1, \ldots, \alpha_t$ of $\mathrm{Ext}^1_R(M, R)$. Then each α_i represents a short exact sequence of the form $0 \to R \to N_i \to M \to 0$. Let $\alpha : 0 \to R^{\oplus t} \to N \to M \to 0$ be a pullback of the short exact sequence $\bigoplus_{i=1}^t \alpha_i : 0 \to R^{\oplus t} \to \bigoplus_{i=1}^t N_i \to M^{\oplus t} \to 0$ by the diagonal map $\Delta : M \to M^{\oplus t}$. Then $\alpha = (\alpha_1, \ldots, \alpha_t) \in \mathrm{Ext}^1_R(M, R^{\oplus t}) \cong \mathrm{Ext}^1_R(M, R)^{\oplus t}$. Next consider the induced exact sequence

$$0 \to M^* \to N^* \to (R^{\oplus t})^* \xrightarrow{\alpha} \mathrm{Ext}^1_R(M, R) \to \mathrm{Ext}^1_R(N, R) \to 0.$$

Since the map $(R^{\oplus t})^* \xrightarrow{\alpha} \mathrm{Ext}^1_R(M, R)$ is surjective, we see that $\mathrm{Ext}^1_R(N, R) = 0$. Thus, in the following pullback diagram, W, being a direct sum of $R^{\oplus s}$ and $R^{\oplus t}$, is free. So the vanishing of $\mathrm{Ext}^1_R(N, R)$ shows that $N = \mathrm{PF}(\Omega M)$.

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & R^{\oplus t} & \longrightarrow & N & \longrightarrow & M & \longrightarrow & 0\\
 & & \| & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & R^{\oplus t} & \longrightarrow & W & \longrightarrow & R^{\oplus s} & \longrightarrow & 0\\
 & & & & \uparrow & & \uparrow & & \\
 & & & & \Omega M & = & \Omega M & & \\
 & & & & \uparrow & & \uparrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

Now, if R is as in §B.1 and $M = \omega$, via the argument there, $N \otimes_R N$ is torsion-free.

In the next example we record a nonfree, torsion-free module N over a one-dimensional local domain R, where $N \otimes_R N$ is torsion-free, but $N \otimes_R N^*$ has torsion.

Example B.4. Let $R = \mathbb{C}[\![t^3, t^4, t^5]\!] \cong \mathbb{C}[\![x, y, z]\!]/(y^2 - xz, x^3 - yz, x^2y - z^2)$. Then R is a one-dimensional local domain which is not Gorenstein. Let N be the R-module given by the following exact sequence:

$$R^{\oplus 3} \xrightarrow{\begin{bmatrix} -y & x & z \\ x^2 & -z & -xy \\ -z & y & x^2 \end{bmatrix}} R^{\oplus 3} \longrightarrow N \longrightarrow 0.$$

One can check, for example, by using [Macaulay2 1993], that both N and $N \otimes_R N$ are torsion-free R-modules. Moreover, it follows that $N \otimes_R N^*$ has torsion; see [Huneke et al. 2019, 3.6].

Acknowledgements

The author is grateful to Hailong Dao and Mark Walker for their comments and suggestions about the results in Appendix A, especially about §A.1 and Example A.3; to Naoki Endo for discussions about Example 4.17 and Proposition 4.18; to Mohsen Gheibi and Ryo Takahashi for discussions concerning the examples in Section 5; to W. Frank Moore for discussions about Example B.4; to Arash Sadeghi for his suggestions on Lemma 4.11 and Proposition 4.12, and to Yongwei Yao for comments on Example 1.2(i).

The author is also grateful to Tokuji Araya, Shiro Goto, Hiroki Matsui, Li Liang, and Greg Piepmeyer for their valuable comments and suggestions at different stages of this project.

The author sincerely thanks to the anonymous referee; the comments and the suggestions of the referee significantly improved the presentation of the manuscript, as well as clarified various points throughout the first version of the manuscript.

References

[Araya and Yoshino 1998] T. Araya and Y. Yoshino, "Remarks on a depth formula, a grade inequality and a conjecture of Auslander", *Comm. Algebra* **26**:11 (1998), 3793–3806. MR Zbl

[Araya et al. 2018] T. Araya, O. Celikbas, A. Sadeghi, and R. Takahashi, "On the vanishing of self extensions over Cohen–Macaulay local rings", *Proc. Amer. Math. Soc.* **146**:11 (2018), 4563–4570. MR Zbl

[Auslander 1961] M. Auslander, "Modules over unramified regular local rings", *Illinois J. Math.* **5** (1961), 631–647. MR

[Auslander and Bridger 1969] M. Auslander and M. Bridger, *Stable module theory*, Mem. Amer. Math. Soc. **94**, Amer. Math. Soc., Providence, RI, 1969. MR Zbl

[Auslander and Reiten 1975] M. Auslander and I. Reiten, "On a generalized version of the Nakayama conjecture", *Proc. Amer. Math. Soc.* **52** (1975), 69–74. MR Zbl

[Avramov 1989] L. L. Avramov, "Modules of finite virtual projective dimension", *Invent. Math.* **96**:1 (1989), 71–101. MR Zbl

[Avramov 1998] L. L. Avramov, "Infinite free resolutions", pp. 1–118 in *Six lectures on commutative algebra* (Bellaterra, Spain, 1996), edited by J. M. Giral et al., Progr. Math. **166**, Birkhäuser, Basel, 1998. MR Zbl

[Avramov and Buchweitz 2000] L. L. Avramov and R.-O. Buchweitz, "Support varieties and cohomology over complete intersections", *Invent. Math.* **142**:2 (2000), 285–318. MR Zbl

[Avramov and Martsinkovsky 2002] L. L. Avramov and A. Martsinkovsky, "Absolute, relative, and Tate cohomology of modules of finite Gorenstein dimension", *Proc. Lond. Math. Soc.* (3) **85**:2 (2002), 393–440. MR Zbl

[Avramov et al. 1997] L. L. Avramov, V. N. Gasharov, and I. V. Peeva, "Complete intersection dimension", *Inst. Hautes Études Sci. Publ. Math.* **86** (1997), 67–114. MR Zbl

[Avramov et al. 2005] L. L. Avramov, R.-O. Buchweitz, and L. M. Şega, "Extensions of a dualizing complex by its ring: commutative versions of a conjecture of Tachikawa", *J. Pure Appl. Algebra* **201**:1-3 (2005), 218–239. MR Zbl

[Bergh 2008] P. A. Bergh, "On the vanishing of (co)homology over local rings", *J. Pure Appl. Algebra* **212**:1 (2008), 262–270. MR Zbl

[Bergh and Jorgensen 2011] P. A. Bergh and D. A. Jorgensen, "On the vanishing of homology for modules of finite complete intersection dimension", *J. Pure Appl. Algebra* **215**:3 (2011), 242–252. MR Zbl

[Bergh and Jorgensen 2014] P. A. Bergh and D. A. Jorgensen, "On growth in minimal totally acyclic complexes", *J. Commut. Algebra* **6**:1 (2014), 17–31. MR Zbl

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Stud. Adv. Math. **39**, Cambridge Univ. Press, 1993. MR Zbl

[Buchweitz and van Straten 2012] R.-O. Buchweitz and D. van Straten, "An index theorem for modules on a hypersurface singularity", *Mosc. Math. J.* **12**:2 (2012), 237–259. MR Zbl

[Celikbas 2011] O. Celikbas, "Vanishing of Tor over complete intersections", *J. Commut. Algebra* **3**:2 (2011), 169–206. MR Zbl

[Celikbas and Dao 2011] O. Celikbas and H. Dao, "Asymptotic behavior of Ext functors for modules of finite complete intersection dimension", *Math. Z.* **269**:3-4 (2011), 1005–1020. MR Zbl

[Celikbas and Dao 2014] O. Celikbas and H. Dao, "Necessary conditions for the depth formula over Cohen–Macaulay local rings", *J. Pure Appl. Algebra* **218**:3 (2014), 522–530. MR Zbl

[Celikbas and Takahashi 2019] O. Celikbas and R. Takahashi, "On the second rigidity theorem of Huneke and Wiegand", *Proc. Amer. Math. Soc.* **147**:7 (2019), 2733–2739. MR Zbl

[Celikbas and Wiegand 2015] O. Celikbas and R. Wiegand, "Vanishing of Tor, and why we care about it", *J. Pure Appl. Algebra* **219**:3 (2015), 429–448. MR Zbl

[Celikbas et al. 2015a] O. Celikbas, S. B. Iyengar, G. Piepmeyer, and R. Wiegand, "Criteria for vanishing of Tor over complete intersections", *Pacific J. Math.* **276**:1 (2015), 93–115. MR Zbl

[Celikbas et al. 2015b] O. Celikbas, A. Sadeghi, and R. Takahashi, "Bounds on depth of tensor products of modules", *J. Pure Appl. Algebra* **219**:5 (2015), 1670–1684. MR Zbl

[Celikbas et al. 2019a] O. Celikbas, S. Goto, R. Takahashi, and N. Taniguchi, "On the ideal case of a conjecture of Huneke and Wiegand", *Proc. Edinb. Math. Soc.* (2) **62**:3 (2019), 847–859. MR Zbl

[Celikbas et al. 2019b] O. Celikbas, H. Matsui, and A. Sadeghi, "On an example concerning the second rigidity theorem", preprint, 2019. arXiv

[Christensen 2000] L. W. Christensen, *Gorenstein dimensions*, Lecture Notes in Math. **1747**, Springer, 2000. MR Zbl

[Dao 2007] H. Dao, "Asymptotic behavior of Tor over complete intersections and applications", preprint, 2007. arXiv

[Dao 2008] H. Dao, "Some observations on local and projective hypersurfaces", *Math. Res. Lett.* **15**:2 (2008), 207–219. MR Zbl

[Dao 2013] H. Dao, "Decent intersection and Tor-rigidity for modules over local hypersurfaces", *Trans. Amer. Math. Soc.* **365**:6 (2013), 2803–2821. MR Zbl

[Eisenbud 1980] D. Eisenbud, "Homological algebra on a complete intersection, with an application to group representations", *Trans. Amer. Math. Soc.* **260**:1 (1980), 35–64. MR Zbl

[Goto et al. 2014] S. Goto, K. Ozeki, R. Takahashi, K.-I. Watanabe, and K.-I. Yoshida, "Ulrich ideals and modules", *Math. Proc. Cambridge Philos. Soc.* **156**:1 (2014), 137–166. MR Zbl

[Henriques and Şega 2011] I. B. D. A. Henriques and L. M. Şega, "Free resolutions over short Gorenstein local rings", *Math. Z.* **267**:3-4 (2011), 645–663. MR Zbl

[Herzog et al. 1991] J. Herzog, B. Ulrich, and J. Backelin, "Linear maximal Cohen–Macaulay modules over strict complete intersections", *J. Pure Appl. Algebra* **71**:2-3 (1991), 187–202. MR Zbl

[Hochster 1981] M. Hochster, "The dimension of an intersection in an ambient hypersurface", pp. 93–106 in *Algebraic geometry* (Chicago, 1980), edited by A. Libgober and P. Wagreich, Lecture Notes in Math. **862**, Springer, 1981. MR Zbl

[Huneke and Jorgensen 2003] C. Huneke and D. A. Jorgensen, "Symmetry in the vanishing of Ext over Gorenstein rings", *Math. Scand.* **93**:2 (2003), 161–184. MR Zbl

[Huneke and Wiegand 1994] C. Huneke and R. Wiegand, "Tensor products of modules and the rigidity of Tor", *Math. Ann.* **299**:3 (1994), 449–476. Correction in **338**:2 (2007), 291–293. MR Zbl

[Huneke and Wiegand 1997] C. Huneke and R. Wiegand, "Tensor products of modules, rigidity and local cohomology", *Math. Scand.* **81**:2 (1997), 161–183. MR Zbl

[Huneke et al. 2019] C. Huneke, S. B. Iyengar, and R. Wiegand, "Rigid ideals in Gorenstein rings of dimension one", *Acta Math. Vietnam.* **44**:1 (2019), 31–49. MR Zbl

[Lipman 1971] J. Lipman, "Stable ideals and Arf rings", *Amer. J. Math.* **93**:3 (1971), 649–685. MR Zbl

[Macaulay2 1993] D. R. Grayson and M. E. Stillman, "Macaulay2, a software system for research in algebraic geometry", 1993, available at http://www.math.uiuc.edu/Macaulay2.

[Matsumura 1986] H. Matsumura, *Commutative ring theory*, Cambridge Stud. Adv. Math. **8**, Cambridge Univ. Press, 1986. MR Zbl

[Miller 1998] C. Miller, "Complexity of tensor products of modules and a theorem of Huneke–Wiegand", *Proc. Amer. Math. Soc.* **126**:1 (1998), 53–60. MR Zbl

[Miller 2003] C. Miller, "The Frobenius endomorphism and homological dimensions", pp. 207–234 in *Commutative algebra* (Grenoble/Lyon, 2001), edited by L. L. Avramov et al., Contemp. Math. **331**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Moore et al. 2011] W. F. Moore, G. Piepmeyer, S. Spiroff, and M. E. Walker, "Hochster's theta invariant and the Hodge–Riemann bilinear relations", *Adv. Math.* **226**:2 (2011), 1692–1714. MR Zbl

[Ooishi 1996] A. Ooishi, "On the self-dual maximal Cohen–Macaulay modules", *J. Pure Appl. Algebra* **106**:1 (1996), 93–102. MR Zbl

[Rotman 1979] J. J. Rotman, *An introduction to homological algebra*, Pure Appl. Math. **85**, Academic Press, New York, 1979. MR Zbl

[Sather-Wagstaff 2004] S. Sather-Wagstaff, "Complete intersection dimensions for complexes", *J. Pure Appl. Algebra* **190**:1-3 (2004), 267–290. MR Zbl

[Soto 2000] J. J. M. Soto, "Gorenstein quotients by principal ideals of free Koszul homology", *Glasg. Math. J.* **42**:1 (2000), 51–54. MR Zbl

[Takahashi 2008] R. Takahashi, "On G-regular local rings", *Comm. Algebra* **36**:12 (2008), 4472–4491. MR Zbl

[Walker 2017] M. Walker, "On the vanishing of Hochster's θ invariant", *Ann. K-Theory* **2**:2 (2017), 131–174. MR Zbl

[Yoshino 1990] Y. Yoshino, *Cohen–Macaulay modules over Cohen–Macaulay rings*, Lond. Math. Soc. Lect. Note Ser. **146**, Cambridge Univ. Press, 1990. MR Zbl

Received May 5, 2019. Revised February 4, 2020.

OLGUR CELIKBAS
DEPARTMENT OF MATHEMATICS
WEST VIRGINIA UNIVERSITY
MORGANTOWN, WV
UNITED STATES

olgur.celikbas@math.wvu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.145

ALGEBRAIC AND GEOMETRIC PROPERTIES OF FLAG BOTT–SAMELSON VARIETIES AND APPLICATIONS TO REPRESENTATIONS

NAOKI FUJITA, EUNJEONG LEE AND DONG YOUP SUH

We define and study flag Bott–Samelson varieties which generalize both Bott–Samelson varieties and flag varieties. Using a birational morphism from an appropriate Bott–Samelson variety to a flag Bott–Samelson variety, we compute the Newton–Okounkov bodies of flag Bott–Samelson varieties as generalized string polytopes, which are applied to give polyhedral expressions for irreducible decompositions of tensor products of G-modules. Furthermore, we show that flag Bott–Samelson varieties degenerate into flag Bott manifolds with higher rank torus actions, and we describe the Duistermaat–Heckman measures of the moment map images of flag Bott–Samelson varieties with torus actions and invariant closed 2-forms.

1. Introduction 145
2. Newton–Okounkov bodies of flag Bott–Samelson varieties 148
3. Applications to representation theory 164
4. Flag Bott–Samelson varieties and flag Bott towers 176
5. Torus actions and Duistermaat–Heckman measure 186
Acknowledgments 192
References 192

1. Introduction

Bott–Samelson varieties provide fruitful connections between representation theory and algebraic geometry. They are nonsingular towers of $\mathbb{C}P^1$-fibrations, and studied in [Bott and Samelson 1958; Demazure 1974; Hansen 1973] to find resolutions of singularities of Schubert varieties. Moreover, the set of global sections of a holomorphic line bundle over a Bott–Samelson variety is the dual of a *generalized Demazure*

Lee is the corresponding author. Fujita was partially supported by Grant-in-Aid for JSPS Fellows (No. 16J00420). Lee was partially supported by IBS-R003-D1. Suh was partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (No. 2016R1A2B4010823).
MSC2020: primary 05E10; secondary 14M15, 57S25.
Keywords: flag Bott–Samelson varieties, Bott–Samelson varieties, Newton–Okounkov bodies, flag Bott manifolds.

module. This leads to worthwhile connections between representation theory and algebraic geometry exemplified by the character formulas of Demazure modules [Andersen 1985; Kumar 1987], the standard monomial theory [Lakshmibai et al. 1979; Lakshmibai and Seshadri 1991; Lakshmibai and Raghavan 2008; Seshadri 2007], and the theory of Newton–Okounkov bodies [Fujita and Naito 2017; Kaveh 2015].

On the other hand, for a given Bott–Samelson variety, it is presented by Grossberg and Karshon [1994] that there is a complex one-parameter family of smooth varieties, which are all diffeomorphic, such that a generic fiber is the Bott–Samelson variety and the special fiber is a nonsingular toric variety, called a *Bott manifold*. It should be noted that the Bott manifold is toric, while the Bott–Samelson variety is not toric in general. Using this connection, [Grossberg and Karshon 1994] also introduces a combinatorial object, called a *Grossberg–Karshon twisted cube*, which is used to compute the multiplicities of generalized Demazure modules (see [Grossberg and Karshon 1994, Theorem 3]).

One of the primary goals of this paper is to generalize the notion of Bott–Samelson varieties and to extend its rich connections with representation theory. Moreover, the generalization also supports the Grossberg–Karshon type degeneration into flag Bott manifolds. Indeed, we consider a *flag Bott–Samelson variety* (see Definition 2.1) which is a nonsingular projective tower of products of full flag manifolds. Moreover, under a certain condition, the flag Bott–Samelson variety is a desingularization of a Schubert variety. Because of the definition, both the full flag varieties and Bott–Samelson varieties are flag Bott–Samelson varieties. Hence we may regard flag Bott–Samelson varieties as the generalization of both flag varieties and Bott–Samelson varieties.

This notion of flag Bott–Samelson varieties is not new. Actually, in [Jantzen 2003], flag Bott–Samelson varieties are treated in a more general setting without naming them. Indeed, flag Bott–Samelson varieties are iterated flag manifold fibrations but Jantzen [2003] considers iterated Schubert varieties fibrations. Perrin [2007] uses these varieties to obtain small resolutions of Schubert varieties. In fact, they are called *generalized Bott–Samelson varieties* (see [Brion and Kannan 2019a; 2019b]). Moreover, flag Bott–Samelson varieties are generalized Bott–Samelson varieties (see Remark 2.2).

Let G be a simply connected semisimple algebraic group of rank n over $\mathbb{C}$. Bott–Samelson varieties $Z_{\boldsymbol{i}}$ are parametrized by words $\boldsymbol{i} = (i_1, \dots, i_r)$, where $i_1, \dots, i_r$ are elements in the set $[n] := \{1, \dots, n\}$. On the other hand, flag Bott–Samelson varieties $Z_{\mathcal{I}}$ are parametrized by sequences $\mathcal{I} = (I_1, \dots, I_r)$ of subsets of $[n]$. Even though the class of flag Bott–Samelson varieties is much larger than that of Bott–Samelson varieties, for each flag Bott–Samelson variety, there exists a Bott–Samelson variety such that there is a birational morphism from the Bott–Samelson variety to the flag Bott–Samelson variety (see Proposition 2.7).

Using the above mentioned birational morphism, we provide Theorem 2.20, which describes the set of holomorphic sections of a holomorphic line bundle over the flag Bott–Samelson variety in terms of generalized Demazure modules. The theory of Newton–Okounkov bodies of projective varieties has been used to present a connection between representation theory and algebraic geometry (see Section 2C for the definition of Newton–Okounkov bodies). The description of holomorphic sections of flag Bott–Samelson varieties is used to compute their Newton–Okounkov bodies. Indeed, using the result of Newton–Okounkov bodies of Bott–Samelson varieties by Fujita [2018], we obtain Theorem 2.22, which shows that the Newton–Okounkov bodies of flag Bott–Samelson varieties with an appropriate valuation agree with generalized string polytopes up to sign.

One of the fundamental questions in group representation theory is to find the multiplicities of irreducible representations in the tensor product of two representations. Berenstein and Zelevinsky [2001] describe the multiplicities in terms of the numbers of lattice points in some explicit rational convex polytope. In Theorem 3.19 we give a different description of the multiplicities using the integral lattice points of the Newton–Okounkov bodies, hence generalized string polytopes, of flag Bott–Samelson varieties. We notice that our results give concrete constructions of convex bodies, appearing in [Kaveh and Khovanskii 2012a], which encode multiplicities of irreducible representations.

As is mentioned before, we degenerate the complex structures of flag Bott–Samelson varieties. The notion of Bott manifolds is generalized to that of *flag Bott manifolds* in terms of iterated flag manifold fibrations described in [Kaji et al. 2020; Kuroki et al. 2020]. More precisely, a flag Bott manifold is the total space of an iterated flag manifold fibrations which are taken by the full flag fibration of a sum of line bundles (see Definition 4.1). In general, a flag Bott manifold is not toric but admits an action of a certain torus. For a given flag Bott–Samelson variety, we provide a complex one-parameter family of smooth varieties, which are all diffeomorphic, such that a generic fiber is the flag Bott–Samelson variety and the special fiber is a flag Bott manifold (see Corollary 4.7). Moreover, when the Levi subgroup L_{I_k} of the parabolic subgroup P_{I_k} is of type A, we explicitly describe such flag Bott manifolds in Theorem 4.10 in terms of the Chern classes of the line bundles used in the construction of the flag Bott manifold.

For a given flag Bott–Samelson variety, there exists a Bott–Samelson variety which is birationally equivalent to the flag Bott–Samelson variety. Moreover, using the result of Grossberg and Karshon [1994], and our one-parameter family, we obtain two manifolds: a flag Bott manifold and a Bott manifold, and a map between them. We study a relation between these manifolds. Actually, considering torus actions on these manifolds, we describe the Duistermaat–Heckman measure of the flag Bott manifold with a certain closed 2-form using a Grossberg–Karshon twisted cube in Theorem 5.5.

This paper is organized as follows. In Section 2, we provide the definition of flag Bott–Samelson varieties and their properties. In particular, we investigate a relation between flag Bott–Samelson varieties and Bott–Samelson varieties. Moreover, we describe the set of holomorphic sections of a line bundle over a flag Bott–Samelson variety using generalized Demazure modules in Theorem 2.20. Using this association, we describe the Newton–Okounkov bodies of flag Bott–Samelson varieties in Theorem 2.22. In Section 3, we give an application of Newton–Okounkov bodies of flag Bott–Samelson varieties to representation theory. Indeed, we provide a way to compute the multiplicities of representations in the tensor product of representations counting certain lattice points in the Newton–Okounkov bodies of flag Bott–Samelson varieties in Theorem 3.19. In Section 4, we present a Grossberg–Karshon type degeneration of complex structures on flag Bott–Samelson varieties, and explicitly describe the corresponding flag Bott manifold when all Levi subgroups of parabolic subgroups P_{I_k} are of type A in Theorem 4.10. In Section 5, we study torus actions on flag Bott manifolds which are obtained by the degeneration of flag Bott–Samelson manifolds. Moreover, we describe the Duistermaat–Heckman measure of flag Bott manifolds using Grossberg–Karshon twisted cubes in Theorem 5.5.

2. Newton–Okounkov bodies of flag Bott–Samelson varieties

2A. ***Definition of flag Bott–Samelson varieties.*** In this subsection we introduce flag Bott–Samelson varieties which generalize both Bott–Samelson varieties and flag varieties, and study their properties. We notice that the notion of flag Bott–Samelson varieties is already considered in Jantzen's book [2003, II.13] without naming it.

Let G be a simply connected semisimple algebraic group of rank n over $\mathbb{C}$. Choose a Cartan subgroup H, and let $\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha} \mathfrak{g}_{\alpha}$ be the decomposition of the Lie algebra $\mathfrak{g}$ of G into root spaces, where $\mathfrak{h}$ is the Lie algebra of H. Let $\Delta \subset \mathfrak{h}^*$ denote the roots of G. Choose a set of positive roots $\Delta^+ \subset \Delta$, and let $\Sigma = \{\alpha_1, \dots, \alpha_n\} \subset \Delta^+$ denote the simple roots. Let $\Delta^- := -\Delta^+$ be the set of negative roots. Let B be the Borel subgroup whose Lie algebra is $\mathfrak{b} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta^+} \mathfrak{g}_{\alpha}$. Let $\{\alpha_1^{\vee}, \dots, \alpha_n^{\vee}\}$ denote the coroots, and $\{\varpi_1, \dots, \varpi_n\}$ the fundamental weights which are characterized by the relation $\langle \varpi_i, \alpha_j^{\vee} \rangle = \delta_{ij}$. Here, δ_{ij} denotes the Kronecker symbol. Let $s_i \in W$ denote the simple reflection in the Weyl group W of G corresponding to the simple root α_i.

For a subset I of $[n] := \{1, \dots, n\}$, define the subtorus $H_I \subset H$ as

$$H_I := \{h \in H \mid \alpha_i(h) = 1 \text{ for all } i \in I\}^0. \tag{2-1}$$

Here, for a group G, G^0 is the connected component which contains the identity element of G. Then the centralizer $C_G(H_I) = \{g \in G \mid gh = hg \text{ for all } h \in H_I\}$ of

H_I is a connected reductive subgroup of G whose Weyl group is isomorphic to $W_I := \langle s_i \mid i \in I\rangle$. We set $L_I := C_G(H_I)$ for simplicity. Then the Borel subgroup B_I of L_I is $B \cap L_I$ (see [Springer 1998, §8.4.1]). Let Δ_I be the subset $\Delta \cap \mathrm{span}_{\mathbb{Z}}\{\alpha_i \mid i \in I\}$ of Δ. The set of roots $\Delta^+ \setminus \Delta_I$ defines the unipotent subgroup U_I of G satisfying the condition

$$\mathrm{Lie}(U_I) = \bigoplus_{\alpha \in \Delta^+ \setminus \Delta_I} \mathfrak{g}_\alpha.$$

The *parabolic subgroup* P_I of G corresponding to I is defined to be $P_I := L_I U_I$. The subgroup L_I is called a *Levi subgroup* of P_I.

Note that the Lie algebra of the parabolic subgroup P_I is

$$\mathrm{Lie}(P_I) = \mathfrak{b} \oplus \bigoplus_{\alpha \in \Delta^- \cap \Delta_I} \mathfrak{g}_\alpha.$$

Moreover the parabolic subgroup P_I can be described that

$$P_I = \bigcup_{w \in W_I} BwB = \overline{Bw_I B} \subset G,$$

where w_I be the longest element in W_I (see [Springer 1998, Theorem 8.4.3]).

We now define the flag Bott–Samelson variety using a sequence of parabolic subgroups. Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$, and let $\boldsymbol{P}_{\mathcal{I}} = P_{I_1} \times \cdots \times P_{I_r}$. Define a right action Θ of $B^r = \underbrace{B \times \cdots \times B}_{r}$ on $\boldsymbol{P}_{\mathcal{I}}$ as

$$\Theta\big((p_1, \ldots, p_r), (b_1, \ldots, b_r)\big) = \big(p_1 b_1, b_1^{-1} p_2 b_2, \ldots, b_{r-1}^{-1} p_r b_r\big) \tag{2-2}$$

for $(p_1, \ldots, p_r) \in \boldsymbol{P}_{\mathcal{I}}$ and $(b_1, \ldots, b_r) \in B^r$.

Definition 2.1. Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$. The *flag Bott–Samelson variety* $Z_{\mathcal{I}}$ is defined to be the orbit space

$$Z_{\mathcal{I}} := \boldsymbol{P}_{\mathcal{I}}/\Theta.$$

For instance, suppose that $\mathcal{I} = ([n])$. Then we have $\boldsymbol{P}_{\mathcal{I}} = G$ and the action Θ is the right multiplication of B. Therefore the flag Bott–Samelson variety $Z_{\mathcal{I}}$ is the flag variety G/B. Moreover, for the case when $|I_k| = 1$ for all k, the flag Bott–Samelson variety is a *Bott–Samelson variety*, see [Bott and Samelson 1958] for the definition of a Bott–Samelson variety. In this case we use a sequence $(i_1, \ldots, i_r)$ of elements of $[n]$ rather than $(\{i_1\}, \ldots, \{i_r\})$, and we write $Z_{(i_1,\ldots,i_r)}$ for the corresponding Bott–Samelson variety.

For the subsequence $\mathcal{I}' = (I_1, \ldots, I_{r-1})$ of $\mathcal{I}$, there is a fibration structure on the flag Bott–Samelson variety $Z_{\mathcal{I}}$:

$$P_{I_r}/B \hookrightarrow Z_{\mathcal{I}} \xrightarrow{\pi} Z_{\mathcal{I}'}, \tag{2-3}$$

where the projection map $\pi : Z_{\mathcal{I}} \to Z_{\mathcal{I}'}$ is defined as

$$\pi([p_1, \dots, p_{r-1}, p_r]) = [p_1, \dots, p_{r-1}].$$

On the other hand, we have that

$$Z_{\mathcal{I}} = P_{I_1} \times^B Z_{(I_2,\dots,I_r)}.$$

Remark 2.2. For a finite sequence $\hat{w} = (w_1, \dots, w_r)$ of elements of W, Perrin [2007] considers a tower $\widehat{X}(\hat{w})$ of Schubert varieties $X(w_1), \dots, X(w_r)$ fibrations and call it a *generalized Bott–Samelson variety*. Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$. When we take w_k to be the longest element in W_{I_k} for $1 \le k \le r$, the generalized Bott–Samelson variety $\widehat{X}(w_1, \dots, w_r)$ is the flag Bott–Samelson variety $Z_{\mathcal{I}}$. Indeed, using the notation in [Perrin 2007, §5.2], we obtain $P^{w_k} = P_{[n]\setminus I_k}$, $P_{w_k} = P_{I_k}$, $G_{w_k} = L_{I_k}$. Therefore, $P^{w_k} \cap G_{w_k}$ is a Borel subgroup B_{I_k} of L_{I_k} and $P_{w_k} \cap L_{I_k} = L_{I_k}$. Thus, we obtain

$$\begin{aligned}
\widehat{X}(w_1, \dots, w_r) &= \overline{(P_{w_1} \cap G_{w_1}) w_1 (P^{w_1} \cap G_{w_1})} \times^{(P^{w_1} \cap G_{w_1})} \widehat{X}(w_2, \dots, w_r) \\
&= \overline{B_{I_1} w_1 L_{I_1}} \times^{B_{I_1}} \widehat{X}(w_2, \dots, w_r) \\
&= L_{I_1} \times^{B_{I_1}} \widehat{X}(w_2, \dots, w_r) \\
&\simeq P_{I_1} \times^B \widehat{X}(w_2, \dots, w_r).
\end{aligned}$$

Continuing this procedure, we get $\widehat{X}(w_1, \dots, w_r) \simeq Z_{\mathcal{I}}$. This shows that flag Bott–Samelson varieties are generalized Bott–Samelson varieties. Because we are considering sequences $\hat{w} = (w_1, \dots, w_r)$ consisting of longest elements, not all generalized Bott–Samelson varieties are flag Bott–Samelson varieties.

Let $w_k \in W_{I_k}$ be the longest element in W_{I_k} for $1 \le k \le r$. Consider the following subset of $\boldsymbol{P}_{\mathcal{I}}$:

$$\boldsymbol{P}'_{\mathcal{I}} := B w_1 B \times \cdots \times B w_r B \subset \boldsymbol{P}_{\mathcal{I}}.$$

One can check that $\boldsymbol{P}'_{\mathcal{I}}$ is closed under the action Θ of B^r in (2-2), so we consider the orbit space

$$Z'_{\mathcal{I}} := \boldsymbol{P}'_{\mathcal{I}} / \Theta.$$

It is known that flag Bott–Samelson varieties $Z_{\mathcal{I}}$ have following properties (see [Jantzen 2003, II.13] for details).

Proposition 2.3. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$. Then the flag Bott–Samelson variety $Z_{\mathcal{I}}$ has following properties*:

(1) *$Z_{\mathcal{I}}$ is a smooth projective variety.*

(2) *$Z'_{\mathcal{I}}$ is a dense open subset in $Z_{\mathcal{I}}$.*

(3) *$Z'_{\mathcal{I}} \simeq \mathbb{C}^{\sum_{k=1}^{r} \ell(w_k)}$, where $\ell(w_k)$ is the length of the element w_k.*

Consider the multiplication map

$$\eta : Z_{\mathcal{I}} \to G/B, \quad [p_1, \dots, p_r] \mapsto p_1 \cdots p_r \tag{2-4}$$

which is a well-defined morphism. The following proposition says that certain flag Bott–Samelson varieties are birationally equivalent to Schubert varieties via the map η.

Proposition 2.4 [Jantzen 2003, II.13.5]. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$, and let $w_k \in W_{I_k}$ be the longest element in $W_{I_k} = \langle s_i \mid i \in I_k \rangle$. Set $w = w_1 \cdots w_r$. If $\ell(w) = \ell(w_1) + \cdots + \ell(w_r)$, then the morphism η induces an isomorphism between $Z'_{\mathcal{I}}$ and $BwB/B \subset G/B$. Indeed, the morphism η maps $Z_{\mathcal{I}}$ birationally onto its image $X(w) := \overline{BwB/B} \subset G/B$.*

Example 2.5. Let $G = \mathrm{SL}(4)$.

(1) Suppose that $\mathcal{I}_1 = (\{1\}, \{2\}, \{1\}, \{3\})$. Then we have $w_1 = s_1,\ w_2 = s_2,\ w_3 = s_1$, $w_4 = s_3$, and $w = s_1 s_2 s_1 s_3$, which is a reduced decomposition. Hence the morphism η gives a birational morphism between $Z_{\mathcal{I}_1}$ and $X(s_1 s_2 s_1 s_3)$.

(2) Let $\mathcal{I}_2 = (\{1, 2\}, \{3\})$. Then we have that $w_1 = s_1 s_2 s_1,\ w_2 = s_3$, and $w = w_1 w_2 = s_1 s_2 s_1 s_3$. Again, this is a reduced decomposition, so the morphism η gives a birational morphism between $Z_{\mathcal{I}_2}$ and $X(s_1 s_2 s_1 s_3)$.

Remark 2.6. Example 2.5 gives two different choices of flag Bott–Samelson varieties each of which has a birational morphism onto the same Schubert variety $X(s_1 s_2 s_1 s_3)$. For a given Schubert variety $X(w)$, there are different choices of flag Bott–Samelson varieties which define birational morphisms onto $X(w)$, and there are several studies about such different choices. See, for example, [Elnitsky 1997; Escobar et al. 2018; Tenner 2006].

We now define a multiplication map between two flag Bott–Samelson varieties. Let

$$\mathcal{J} = \left(J_{1,1}, \dots, J_{1,N_1}, \dots, J_{r,1}, \dots, J_{r,N_r}\right) \tag{2-5}$$

be a sequence of subsets of $[n]$ such that $J_{k,l} \subset I_k$ for $1 \le l \le N_k$ and $1 \le k \le r$. Since each $J_{k,l}$ is contained in I_k, we have $P_{J_{k,l}} \subset P_{I_k}$ by the definition of parabolic subgroups. Hence we have a multiplication map

$$\eta_{\mathcal{J},\mathcal{I}} : Z_{\mathcal{J}} \to Z_{\mathcal{I}} \tag{2-6}$$

defined as

$$[(p_{k,l})_{1 \le k \le r, 1 \le l \le N_k}] \mapsto \left[\prod_{l=1}^{N_1} p_{1,l}, \dots, \prod_{l=1}^{N_r} p_{r,l}\right].$$

The following proposition describes a birational morphism between two flag Bott–Samelson varieties.

Proposition 2.7. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$, and let $\mathcal{J} = \big(J_{1,1}, \dots, J_{1,N_1}, \dots, J_{r,1}, \dots, J_{r,N_r}\big)$ be a sequence of subsets of $[n]$ such that $J_{k,1}, \dots, J_{k,N_k} \subset I_k$ for $1 \le k \le r$. Let $w_{k,l}$, respectively v_k, be the longest element in $W_{J_{k,l}}$, respectively in W_{I_k}. Suppose that*

$$w_{k,1} \cdots w_{k,N_k} = v_k \quad \textit{and} \quad \ell(w_{k,1}) + \cdots + \ell(w_{k,N_k}) = \ell(v_k) \quad \textit{for } 1 \le k \le r.$$

Then the multiplication map $\eta_{\mathcal{J},\mathcal{I}} : Z_{\mathcal{J}} \to Z_{\mathcal{I}}$ in (2-6) induces an isomorphism between dense open subsets $Z'_{\mathcal{J}} \xrightarrow{\sim} Z'_{\mathcal{I}}$.

There always exists a sequence $(i_{k,1}, \dots, i_{k,N_k}) \in [n]^{N_k}$ which is a reduced word for the longest element in W_{I_k} for $1 \le k \le r$. Concatenating such sequences we get a sequence $\boldsymbol{i} = (i_{k,l})_{1 \le k \le r, 1 \le l \le N_k} \in [n]^{N_1 + \cdots + N_r}$. Hence for a given flag Bott–Samelson variety $Z_{\mathcal{I}}$ one can always find a Bott–Samelson variety $Z_{\boldsymbol{i}}$ which is birationally isomorphic to $Z_{\mathcal{I}}$.

Proof of Proposition 2.7. First we recall from [Bourbaki 2002, VI. §1, Corollary 2 of Proposition 17; Jantzen 2003, II.13.1] that for a reduced decomposition $w = s_{i_1} \cdots s_{i_N} \in W$, the subgroup $U(w) \subset G$ is defined to be

$$U(w) := U_{\alpha_{i_1}} \cdot U_{s_{i_1}(\alpha_{i_2})} \cdot U_{s_{i_1} s_{i_2}(\alpha_{i_3})} \cdots U_{s_{i_1} \cdots s_{i_{N-1}}(\alpha_{i_N})}.$$

Moreover, we have an isomorphism

$$\psi(w) : U_{\alpha_{i_1}} \times U_{\alpha_{i_2}} \times \cdots \times U_{\alpha_{i_N}} \xrightarrow{\sim} U(w) \tag{2-7}$$

which is defined to be $(u_1, \dots, u_N) \mapsto u_1 s_{i_1} u_2 s_{i_2} \cdots u_N s_{i_N} w^{-1}$. Also we have another isomorphism $\psi_{\mathcal{I}}$ between varieties:

$$\psi_{\mathcal{I}} : U(v_1) \times \cdots \times U(v_r) \xrightarrow{\sim} Z'_{\mathcal{I}} \tag{2-8}$$

which sends $(g_1, \dots, g_r)$ to $[g_1 v_1, \dots, g_r v_r]$ (see [Jantzen 2003, II.13.5]).

Because of the assumption, the concatenation $w_{k,1} \cdots w_{k,N_k}$ is a reduced decomposition of the element v_k. Hence we have an isomorphism induced by (2-7):

$$\psi_k : U(w_{k,1}) \times \cdots \times U(w_{k,N_k}) \xrightarrow{\sim} U(v_k)$$

which maps $(u_1, \dots, u_{N_k})$ to $u_1 w_{k,1} u_2 w_{k,2} \cdots u_{N_k} w_{k,N_k} v_k^{-1}$ for $1 \le k \le r$. Combining isomorphisms ψ_k and (2-8) we have the following commutative diagram:

$$\begin{array}{ccc} Z'_{\mathcal{J}} & \xleftarrow[\sim]{\psi_{\mathcal{J}}} & U(w_{1,1}) \times \cdots \times U(w_{1,N_1}) \times \cdots \times U(w_{r,1}) \times \cdots \times U(w_{r,N_r}) \\ {\scriptstyle \eta_{\mathcal{J},\mathcal{I}}}\downarrow & & \wr\downarrow{\scriptstyle \psi_1 \times \cdots \times \psi_r} \\ Z'_{\mathcal{I}} & \xleftarrow[\sim]{\psi_{\mathcal{I}}} & U(v_1) \times \cdots \times U(v_r) \end{array}$$

Hence the result follows. □

Example 2.8. Let $G = \mathrm{SL}(4)$, and let $\mathcal{I} = (\{1,2\},\{3\})$. Then $w_1 = s_1 s_2 s_1$, respectively $w_2 = s_3$, is a reduced decomposition of the longest element of $W_{\{1,2\}}$, respectively $W_{\{3\}}$. Then we have the birational morphism $\eta_{(1,2,1,3),\mathcal{I}} : Z_{(1,2,1,3)} \to Z_{\mathcal{I}}$. Together with the birational morphism η described in Example 2.5(2), we can see that three varieties $Z_{(1,2,1,3)}$, $Z_{\mathcal{I}}$, and $X(s_1 s_2 s_1 s_3)$ are birationally equivalent:

$$Z_{(1,2,1,3)} \to Z_{\mathcal{I}} \to X(s_1 s_2 s_1 s_3).$$

On the other hand, we have another reduced decomposition $w_1' = s_2 s_1 s_2$ of the longest element of $W_{\{1,2\}}$. This also gives the birational morphism

$$\eta_{(2,1,2,3),\mathcal{I}} : Z_{(2,1,2,3)} \to Z_{\mathcal{I}}.$$

Hence we have the following diagram whose maps are all birational morphisms:

$$\begin{array}{ccccc} Z_{(1,2,1,3)} & & & & \\ & \searrow & & & \\ & & Z_{\mathcal{I}} & \longrightarrow & X(s_1 s_2 s_1 s_3) \\ & \nearrow & & & \\ Z_{(2,1,2,3)} & & & & \end{array}$$

2B. ***Line bundles over flag Bott–Samelson varieties.*** Let $\mathcal{I}$ be a sequence of subsets of $[n]$. In this subsection we study line bundles over a flag Bott–Samelson variety $Z_{\mathcal{I}}$ and their pullbacks in Proposition 2.10. For an integral weight $\lambda \in \mathbb{Z}\varpi_1 + \cdots + \mathbb{Z}\varpi_n$, we have the homomorphism $e^{\lambda} : H \to \mathbb{C}^*$. We extend it to the homomorphism $e^{\lambda} : B \to \mathbb{C}^*$ by composing with the homomorphism

$$\Upsilon : B \to H \tag{2-9}$$

induced by the canonical projection of Lie algebras $\mathfrak{b} \to \mathfrak{h}$ as in [Jantzen 2003, II.1.8]. Suppose that $\lambda_1, \dots, \lambda_r$ are integral weights. Define a representation $\mathbb{C}_{\lambda_1,\dots,\lambda_r}$ of $B^r = B \times \cdots \times B$ (r factors) on $\mathbb{C}$ as

$$(b_1, \dots, b_r) \cdot v = e^{\lambda_1}(b_1) \cdots e^{\lambda_r}(b_r) v.$$

From this we can build a line bundle over $Z_{\mathcal{I}}$ by setting

$$\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} = \boldsymbol{P}_{\mathcal{I}} \times_{B^r} \mathbb{C}_{-\lambda_1,\dots,-\lambda_r}, \tag{2-10}$$

where an action of B^r is defined as

$$\begin{aligned} (p_1, \dots, p_r, w) \cdot (b_1, \dots, b_r) \\ = \big(\Theta((p_1, \dots, p_r), (b_1, \dots, b_r)), e^{\lambda_1}(b_1) \cdots e^{\lambda_r}(b_r) w\big). \end{aligned} \tag{2-11}$$

For simplicity, we use the following notation:

$$\mathcal{L}_{\mathcal{I},\lambda} := \mathcal{L}_{\mathcal{I},0,\dots,0,\lambda}. \tag{2-12}$$

Remark 2.9. Recall from [Fulton 1998, Example 19.1.11(d)] that for a flag bundle X over Y, the cycle map $\mathrm{cl}_X : A_k(X) \to H_{2k}(X)$ is an isomorphism if and only if cl_Y is an isomorphism. Moreover, the cycle map is isomorphic for an arbitrary flag manifold. Since a flag Bott–Samelson variety is an iterated bundle of flags P_{I_k}/B over a point, the cycle map $\mathrm{cl}_{Z_{\mathcal{I}}} : A_k(Z_{\mathcal{I}}) \to H_{2k}(Z_{\mathcal{I}})$ is an isomorphism. On the other hand, since flag Bott–Samelson varieties are smooth (see Proposition 2.3(1)), we obtain the following isomorphisms

$$\mathrm{Pic}(Z_{\mathcal{I}}) \xrightarrow{\cong} A_{(\dim_{\mathbb{C}} Z_{\mathcal{I}})-1}(Z_{\mathcal{I}}) \xrightarrow[\mathrm{cl}_{Z_{\mathcal{I}}}]{\cong} H_{2(\dim_{\mathbb{C}} Z_{\mathcal{I}})-2}(Z_{\mathcal{I}}) \xrightarrow{\cong} H^2(Z_{\mathcal{I}}). \tag{2-13}$$

Here, the first isomorphism comes from [Fulton 1998, Example 2.1.1] and the last isomorphism is obtained by the Poincaré duality. Indeed, $c_1 : \mathrm{Pic}(Z_{\mathcal{I}}) \to H^2(Z_{\mathcal{I}})$ is the isomorphism (2-13). When the Levi subgroup L_{I_k} of P_{I_k} has Lie type A for all k, we present the association (2-13) using a certain generator of $H^2(Z_{\mathcal{I}})$ in (4-3), and we present the first Chern class of the line bundle $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ in Proposition 4.9.

Specifically when a flag Bott–Samelson variety is a usual Bott–Samelson variety, we will choose the weights to be of special form. We recall a description of the Picard group of $Z_{\boldsymbol{i}}$ from [Lauritzen and Thomsen 2004]. For given an integer vector $\boldsymbol{a} = (a_1, \dots, a_r) \in \mathbb{Z}^r$, we define a sequence of weights $\lambda_1, \dots, \lambda_r$ associated to the word $\boldsymbol{i} = (i_1, \dots, i_r)$ and the vector $\boldsymbol{a}$ by setting

$$\lambda_1 := a_1\varpi_{i_1}, \dots, \lambda_r := a_r\varpi_{i_r}.$$

For such λ_j we use the notation

$$\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}} := \mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}. \tag{2-14}$$

Since a Bott–Samelson variety is an iterated sequence of projective bundles, the Picard group of Bott–Samelson variety $Z_{\boldsymbol{i}}$ is a free abelian group of rank r by [Hartshorne 1977, Exercise II.7.9]. Indeed, the association between $\boldsymbol{a} \in \mathbb{Z}^r$ and $\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$ gives an isomorphism between $\mathbb{Z}^r$ and $\mathrm{Pic}(Z_{\boldsymbol{i}})$ (see [Lauritzen and Thomsen 2004, §3.1]).

Let $\boldsymbol{i} = (i_{k,l})_{1 \le k \le r, 1 \le l \le N_k} \in [n]^{N_1+\cdots+N_r}$ be a sequence such that $(i_{k,1}, \dots, i_{k,N_k})$ is a reduced word for the longest element in W_{I_k} for $1 \le k \le r$. Recall from Proposition 2.7 that we have the birational morphism $\eta_{\boldsymbol{i},\mathcal{I}} : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$. The following proposition describes the pullback bundle $\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ under the morphism $\eta_{\boldsymbol{i},\mathcal{I}}$ in terms of an integer vector.

Proposition 2.10. *Let $\mathcal{I}$, $\boldsymbol{i}$, and $\lambda_1, \dots, \lambda_r$ be as above. The pullback bundle $\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ over the Bott–Samelson variety $Z_{\boldsymbol{i}}$ is isomorphic to the line bundle $\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$ for the integer vector*

$$\boldsymbol{a} = \big(\boldsymbol{a}_1(1), \dots, \boldsymbol{a}_1(N_1), \dots, \boldsymbol{a}_r(1), \dots, \boldsymbol{a}_r(N_r)\big) \in \mathbb{Z}^{N_1} \oplus \cdots \oplus \mathbb{Z}^{N_r}$$

given by

$$\boldsymbol{a}_k(l) = \begin{cases} \langle \lambda_k, \alpha_s^\vee \rangle + \displaystyle\sum_{\substack{k<j\le r \\ s\notin\{i_{t,u} \mid k<t\le j,\, 1\le u\le N_t\}}} \langle \lambda_j, \alpha_s^\vee \rangle & \text{if } l = \max\{q \mid i_{k,q} = s\}, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.11. Let $G = \mathrm{SL}(4)$, $\mathcal{I} = (\{1, 2\}, \{3\})$ and $\boldsymbol{i} = (1, 2, 1, 3)$. Consider the line bundle $\mathcal{L}_{\mathcal{I},\lambda_1,\lambda_2}$. Then the pullback line bundle $\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\lambda_2}$ corresponds to the integer vector

$$\begin{aligned} \boldsymbol{a} &= (\boldsymbol{a}_1(1), \boldsymbol{a}_1(2), \boldsymbol{a}_1(3), \boldsymbol{a}_2(1)) \\ &= \big(0, \langle \lambda_1, \alpha_2^\vee \rangle + \langle \lambda_2, \alpha_2^\vee \rangle, \langle \lambda_1, \alpha_1^\vee \rangle + \langle \lambda_2, \alpha_1^\vee \rangle, \langle \lambda_2, \alpha_3^\vee \rangle\big). \end{aligned}$$

Remark 2.12. It is known from [Lauritzen and Thomsen 2004, Theorem 3.1, Corollary 3.3] that the line bundle $\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$ is very ample, respectively generated by global sections, if and only if $\boldsymbol{a} \in \mathbb{Z}^{|\boldsymbol{i}|}_{>0}$, respectively $\boldsymbol{a} \in \mathbb{Z}^{|\boldsymbol{i}|}_{\ge 0}$. Suppose that $\boldsymbol{i}$ is a sequence satisfying the condition in Proposition 2.10. As we saw in Example 2.11, we cannot ensure that the pullback line bundle $\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is very ample even if the weights $\lambda_1, \dots, \lambda_r$ are regular dominant weights.

Proof of Proposition 2.10. By the definition of pullback line bundles, we have

$$\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} = \big\{(p, q) \in Z_{\boldsymbol{i}} \times \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \mid \eta_{\boldsymbol{i},\mathcal{I}}(p) = \pi_{\mathcal{I},\lambda_1,\dots,\lambda_r}(q)\big\},$$

where $\pi_{\mathcal{I},\lambda_1,\dots,\lambda_r} : \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \to Z_{\mathcal{I}}$. In other words,

$$\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} = \left\{ \big([(p_{k,l})_{1\le k\le r, 1\le l\le N_k}], [p_1, \dots, p_r, w]\big) \;\middle|\; \left[\prod_{l=1}^{N_1} p_{1,l}, \dots, \prod_{l=1}^{N_r} p_{r,l}\right] = [p_1, \dots, p_r] \text{ in } Z_{\mathcal{I}} \right\}. \tag{2-15}$$

Define the line bundle $\mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$ on $Z_{\boldsymbol{i}}$ by

$$\begin{aligned} \mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} &:= \mathcal{L}_{\boldsymbol{i},\underbrace{0,\dots,0,\lambda_1}_{N_1},\underbrace{0,\dots,0,\lambda_2}_{N_2},\dots,\underbrace{0,\dots,0,\lambda_r}_{N_r}} \\ &= \big(\boldsymbol{P_i} \times \mathbb{C}_{0,\dots,0,\lambda_1,0,\dots,0,\lambda_2,\dots,0,\dots,0,\lambda_r}\big)/B^{N_1+\cdots+N_r}. \end{aligned}$$

Claim 1. $$\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \cong \mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}.$$

Consider a well-defined morphism $f : \eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \to \mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$ given by

$$f\big([(p_{k,l})_{k,l}], [p_1, \dots, p_r, w]\big) := [(p_{k,l})_{k,l}, Cw]. \tag{2-16}$$

Here, the value C is defined as follows. Because of the description in (2-15), for each element $([(p_{k,l})_{k,l}], [p_1, \dots, p_r, w])$ in the pullback bundle $\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, there exist $b_1, \dots, b_r \in B$ such that

$$p_1 b_1 = \prod_{l=1}^{N_1} p_{1,l}, \quad b_1^{-1} p_2 b_2 = \prod_{l=1}^{N_2} p_{2,l}, \quad \dots, \quad b_{r-1}^{-1} p_r b_r = \prod_{l=1}^{N_r} p_{r,l}. \tag{2-17}$$

Using these elements $b_1, \dots, b_r$, the value C is defined by

$$C := e^{\lambda_1}(b_1) e^{\lambda_2}(b_2) \cdots e^{\lambda_r}(b_r). \tag{2-18}$$

On the other hand, we have a well-defined morphism

$$g : \mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} \to \eta^*_{\boldsymbol{i},\mathcal{I}} \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$$

defined by

$$g([(p_{k,l})_{k,l}, w]) := \left([(p_{k,l})_{k,l}], \left[\prod_{l=1}^{N_1} p_{1,l}, \prod_{l=1}^{N_2} p_{2,l}, \dots, \prod_{l=1}^{N_r} p_{r,l}, w\right]\right). \tag{2-19}$$

We claim that both compositions $f \circ g$ and $g \circ f$ are identities. First we consider the composition $f \circ g$:

$$\begin{aligned} f \circ g([(p_{k,l})_{k,l}, w]) & \\ &= f\left([(p_{k,l})_{k,l}], \left[\prod_{l=1}^{N_1} p_{1,l}, \prod_{l=1}^{N_2} p_{2,l}, \dots, \prod_{l=1}^{N_r} p_{r,l}, w\right]\right) \quad \text{(by (2-19))} \\ &= ([(p_{k,l})_{k,l}, w]). \end{aligned}$$

Here, the last equality holds because all the elements $b_1, \dots, b_r$ satisfying the equations (2-17) are the identity element, and so $C = 1$. For the composition $g \circ f$, we obtain

$$\begin{aligned} g \circ f\big([(p_{k,l})_{k,l}], [p_1, \dots, p_r, w]\big) & \\ &= g([(p_{k,l})_{k,l}, Cw]) \\ &= \left([(p_{k,l})_{k,l}], \left[\prod_{l=1}^{N_1} p_{1,l}, \prod_{l=1}^{N_2} p_{2,l}, \dots, \prod_{l=1}^{N_r} p_{r,l}, Cw\right]\right) \\ &= \big([(p_{k,l})_{k,l}], [p_1 b_1, b_1^{-1} p_2 b_2, \dots, b_{r-1}^{-1} p_r b_r, Cw]\big) \quad \text{(by (2-17))} \\ &= \big([(p_{k,l})_{k,l}], [p_1, p_2, \dots, p_r, e^{\lambda_1}(b_1^{-1}) \cdots e^{\lambda_r}(b_r^{-1}) Cw]\big) \\ &= \big([(p_{k,l})_{k,l}], [p_1, \dots, p_r, w]\big) \quad \text{(by (2-18))}. \end{aligned}$$

Accordingly, f is an isomorphism, and Claim 1 holds.

Claim 2. $\mathcal{L}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} \cong \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$, *where* $\boldsymbol{a}$ *is the integer vector given in the statement of the proposition.*

To present a concrete isomorphism, we set

$$\begin{aligned} k(j,s) &:= \max\bigl(\{k \mid 1 \le k \le j,\ i_{k,l} = s \text{ for some } 1 \le l \le N_k\} \cup \{0\}\bigr), \\ m(j,s) &:= \max\{q \mid i_{k(j,s),q} = s\} \end{aligned} \tag{2-20}$$

for $1 \le j \le r$ and $s \in [n]$. We define certain products $\zeta(j,s)$ of $p_{k,l}$ using $k(j,s)$ as follows:

$$\zeta(j,s) := \Biggl(\prod_{m=m(j,s)+1}^{N_{k(j,s)}} p_{k(j,s),m}\Biggr)\Biggl(\prod_{k=k(j,s)+1}^{j}\prod_{l=1}^{N_k} p_{k,l}\Biggr).$$

Here, $p_{0,l}$ is the identity element.

We denote the integral weight λ_k by $d_{k,1}\varpi_1 + \cdots + d_{k,n}\varpi_n$ using integers $d_{k,j}$ for $1 \le k \le r$. We consider the following morphism:

$$f_2 : \mathcal{L}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \to \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}, \quad [(p_{k,l})_{k,l}, w] \mapsto [(p_{k,l})_{k,l}, C'w], \tag{2-21}$$

where the value C' is defined to be

$$C' := \prod_{s=1}^{n}\prod_{j=1}^{r} e^{d_{j,s}\varpi_s}(\zeta(j,s))^{-1}. \tag{2-22}$$

We note that if $I \subset [n]$ and $s \notin I$, then the map $e^{\varpi_s} : B \to \mathbb{C}^*$ is naturally extended to $e^{\varpi_s} : P_I \to \mathbb{C}^*$ by setting $e^{\varpi_s}(\exp(\mathfrak{g}_\alpha)) = \{1\}$ for all $\alpha \in \Delta^- \cap \Delta_I$. Hence $e^{d_{j,s}\varpi_s}(\zeta(j,s))$ is defined.

If the map f_2 is well-defined, then we obtain Claim 2 because the inverse of f_2 is attained by multiplying $(C')^{-1}$. Therefore, to prove Claim 2, it is enough to show that f_2 is well-defined. Suppose that

$$\bigl[(b_{k,l-1}^{-1} p_{k,l} b_{k,l})_{k,l}, e^{\lambda_1}(b_{1,N_1}) \cdots e^{\lambda_r}(b_{r,N_r})w\bigr]$$

is another representative of the element $[(p_{k,l})_{k,l}, w]$ in $\mathcal{L}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r}$. Here, we use the convention that $b_{k,0} = b_{k-1,N_{k-1}}$ and $b_{0,l}$ is the identity element. To show the well-definedness of f_2, we have to prove that the following equality holds:

$$\begin{aligned} &\Biggl(\prod_{k=1}^{r}\prod_{l=1}^{N_k} e^{\boldsymbol{a}_k(l)\varpi_{i_{k,l}}}(b_{k,l})\Biggr)\Biggl(\prod_{s=1}^{n}\prod_{j=1}^{r} e^{d_{j,s}\varpi_s}(\zeta(j,s))^{-1}\Biggr) \\ &\qquad = \Biggl(\prod_{j=1}^{r}\prod_{s=1}^{n} e^{d_{j,s}\varpi_s}(\zeta(j,s)')^{-1}\Biggr)\Biggl(\prod_{k=1}^{r} e^{\lambda_k}(b_{k,N_k})\Biggr). \end{aligned} \tag{2-23}$$

Here, $\zeta(j,s)'$ is defined by

$$\begin{aligned}\zeta(j,s)' &= \Biggl(\prod_{m=m(j,s)+1}^{N_{k(j,s)}} b_{k(j,s),m-1}^{-1} p_{k(j,s),m} b_{k(j,s),m}\Biggr)\Biggl(\prod_{k=k(j,s)+1}^{j}\prod_{l=1}^{N_k} b_{k,l-1}^{-1} p_{k,l} b_{k,l}\Biggr)\\ &= b_{k(j,s),m(j,s)}^{-1}\zeta(j,s) b_{j,N_j}.\end{aligned}$$

Furthermore, since the weight λ_j is the sum of $d_{j,s}\varpi_s$, we have that

$$\prod_{j=1}^{r}\prod_{s=1}^{n} e^{d_{j,s}\varpi_s}(b_{j,N_j}) = \prod_{j=1}^{r} e^{\lambda_j}(b_{j,N_j}).$$

Therefore, the right hand side of (2-23) becomes

$$\Biggl(\prod_{j=1}^{r}\prod_{s=1}^{n} e^{d_{j,s}\varpi_s}(b_{k(j,s),m(j,s)})\Biggr)\Biggl(\prod_{s=1}^{n}\prod_{j=1}^{r} e^{d_{j,s}\varpi_s}(\zeta(j,s))^{-1}\Biggr).$$

This implies that to show the equality (2-23), it is enough to show that

$$\prod_{k=1}^{r}\prod_{l=1}^{N_k} e^{\boldsymbol{a}_k(l)\varpi_{i_{k,l}}}(b_{k,l}) = \prod_{s=1}^{n}\prod_{j=1}^{r} e^{d_{j,s}\varpi_s}(b_{k(j,s),m(j,s)}). \tag{2-24}$$

The left hand side of (2-24) is written as

$$\prod_{k=1}^{r}\prod_{l=1}^{N_k} e^{\boldsymbol{a}_k(l)\varpi_{i_{k,l}}}(b_{k,l}) = \prod_{s=1}^{n}\prod_{\substack{1\le k\le r,\ 1\le l\le N_k\\ i_{k,l}=s}} e^{\boldsymbol{a}_k(l)\varpi_s}(b_{k,l}).$$

Using this observation, we verify the equality (2-24) by showing

$$\prod_{\substack{1\le k\le r,\ 1\le l\le N_k\\ i_{k,l}=s}} e^{\boldsymbol{a}_k(l)\varpi_s}(b_{k,l}) = \prod_{j=1}^{r} e^{d_{j,s}\varpi_s}(b_{k(j,s),m(j,s)}) \tag{2-25}$$

for all $s\in[n]$. Let s be an arbitrary index in $[n]$. If s does not appear in $(i_{k,l})_{k,l}$, then $k(j,s)=0$ for all j, and so the equality (2-25) holds. Otherwise, let $1\le j_1<\cdots<j_x\le r$ be the indices such that $s\in\{i_{j,1},\dots,i_{j,N_j}\}$ if and only if $j\in\{j_1,\dots,j_x\}$. By the definition of the number $k(j,s)$, we have that

$$\begin{aligned}&0=k(1,s)=\cdots=k(j_1-1,s),\\ &j_u=k(j_u,s)=\cdots=k(j_{u+1}-1,s)\quad\text{for } 1\le u\le x.\end{aligned}$$

Here, we set $j_{x+1}=r+1$. Therefore, we have that

$$\prod_{j=1}^{r} e^{d_{j,s}\varpi_s}(b_{k(j,s),m(j,s)}) = \prod_{u=1}^{x} e^{(d_{j_u,s}+\cdots+d_{j_{u+1}-1,s})\varpi_s}(b_{j_u,m(j_u,s)}).$$

On the other hand, by the definition of the integer vector $\boldsymbol{a}$, if $i_{k,l} = s$, then $k \in \{j_1, \dots, j_x\}$ and we have that

$$\boldsymbol{a}_k(l) = \begin{cases} d_{j_u,s} + d_{j_u+1,s} + \cdots + d_{j_{u+1}-1,s} & \text{if } k = j_u \text{ and } l = m(j_u, s), \\ 0 & \text{otherwise.} \end{cases}$$

Accordingly, we obtain the equality (2-25) for all $s \in [n]$, and we get the equality (2-24). Therefore, the morphism f_2 is a well-defined isomorphism. This proves Claim 2. By combining Claims 1 and 2, the result follows. □

For the reader's convenience, we provide an example for explaining notation $C, k(j, s), C'$ in the proof of Proposition 2.10.

Example 2.13. Let $G = \mathrm{SL}(4)$. Suppose that $\mathcal{I}$ and $\boldsymbol{i}$ are given as in Example 2.11. Then for an element $([p_{1,1}, p_{1,2}, p_{1,3}, p_{2,1}], [p_1, p_2, w])$ in $\eta^*_{\boldsymbol{i},\mathcal{I}} \mathcal{L}_{\mathcal{I},\lambda_1,\lambda_2}$ the value C in (2-18) is given by

$$C = e^{\lambda_1}\big(p_1^{-1} p_{1,1} p_{1,2} p_{1,3}\big) e^{\lambda_2}\big(p_2^{-1} p_1^{-1} p_{1,1} p_{1,2} p_{1,3} p_{2,1}\big).$$

Moreover the indices $k(j, s)$ in (2-20) are computed by

$$k(1,1) = 1, \quad k(1,2) = 1, \quad k(1,3) = 0, \quad k(2,1) = 1, \quad k(2,2) = 1, \quad k(2,3) = 2.$$

Hence the value C' in (2-22) is

$$C' = e^{d_{1,2}\varpi_2}(p_{1,3})^{-1} e^{d_{1,3}\varpi_3}(p_{1,1} p_{1,2} p_{1,3})^{-1} e^{d_{2,1}\varpi_1}(p_{2,1})^{-1} e^{d_{2,2}\varpi_2}(p_{1,3} p_{2,1})^{-1},$$

where $\lambda_k = d_{k,1}\varpi_1 + d_{k,2}\varpi_2 + d_{k,3}\varpi_3$ for $k = 1, 2$.

2C. ***Newton–Okounkov bodies of flag Bott–Samelson varieties.*** Here we study the Newton–Okounkov bodies of flag Bott–Samelson varieties in Theorem 2.22. First we recall the definition and background of Newton–Okounkov bodies. We refer the reader to [Fujita and Naito 2017; Harada and Kaveh 2015; Kaveh 2015; Kaveh and Khovanskii 2012b] for more details. Let R be a $\mathbb{C}$-algebra without nonzero zero-divisors, and fix a total order $<$ on $\mathbb{Z}^r$, $r \geq 1$, respecting the addition.

Definition 2.14. A map $v : R \setminus \{0\} \to \mathbb{Z}^r$ is called a *valuation* on R if the following conditions hold. For every $f, g \in R \setminus \{0\}$ and $c \in \mathbb{C} \setminus \{0\}$,

(1) $v(f \cdot g) = v(f) + v(g)$,

(2) $v(cf) = v(f)$, and

(3) $v(f + g) \geq \min\{v(f), v(g)\}$ unless $f + g = 0$.

Moreover we say the valuation v has *one-dimensional leaves* if it satisfies that if $v(f) = v(g)$ then there exists a nonzero constant $\lambda \in \mathbb{C}$ such that $v(g - \lambda f) > v(g)$ or $g - \lambda f = 0$.

Let X be a projective variety of dimension r over $\mathbb{C}$ equipped with a line bundle $\mathcal{L}$ which is generated by global sections. Fix a valuation v which has one-dimensional leaves on the function field $\mathbb{C}(X)$ of X. Using the valuation v one can associate a semigroup $S \subset \mathbb{N} \times \mathbb{Z}^r$ as follows. Fix a nonzero element $\tau \in H^0(X, \mathcal{L})$. We use τ to identify $H^0(X, \mathcal{L})$ with a finite-dimensional subspace of $\mathbb{C}(X)$ by mapping

$$H^0(X, \mathcal{L}) \to \mathbb{C}(X), \quad \sigma \mapsto \sigma/\tau.$$

Similarly we have the map

$$H^0(X, \mathcal{L}^{\otimes k}) \to \mathbb{C}(X), \quad \sigma \mapsto \sigma/\tau^k.$$

Using these identifications we define the semigroup:

$$S = S(v, \tau) = \bigcup_{k>0} \bigl\{(k, v(\sigma/\tau^k)) \mid \sigma \in H^0(X, \mathcal{L}^{\otimes k}) \setminus \{0\}\bigr\} \subset \mathbb{N} \times \mathbb{Z}^r,$$

and denote by $C = C(v, \tau) \subset \mathbb{R}_{\geq 0} \times \mathbb{R}^r$ the smallest real closed cone containing $S(v, \tau)$. Now we have the definition of Newton–Okounkov body:

Definition 2.15. The *Newton–Okounkov body associated to* $(X, \mathcal{L}, v, \tau)$ is defined to be

$$\{\boldsymbol{x} \in \mathbb{R}^r \mid (1, \boldsymbol{x}) \in C(v, \tau)\}.$$

We denote the Newton–Okounkov body by $\Delta(X, \mathcal{L}, v, \tau)$.

If we take another section $\tau' \in H^0(X, \mathcal{L}) \setminus \{0\}$ then

$$\Delta(X, \mathcal{L}, v, \tau') = \Delta(X, \mathcal{L}, v, \tau) + v(\tau/\tau').$$

Hence the Newton–Okounkov body $\Delta(X, \mathcal{L}, v, \tau)$ does not fundamentally depend on the choice of the nonzero section $\tau \in H^0(X, \mathcal{L}) \setminus \{0\}$. Accordingly, we sometimes denote the Newton–Okounkov body by $\Delta(X, \mathcal{L}, v)$.

Remark 2.16. If we choose a very ample line bundle $\mathcal{L}$ in the construction, then it is known in [Harada and Kaveh 2015, Theorem 3.9] that the Newton–Okounkov body has maximal dimension, i.e., it has real dimension r. Since we do not necessarily assume that the line bundle $\mathcal{L}$ is very ample in this paper, the real dimension of the Newton–Okounkov body may be less than r.

There are many possible valuations with one-dimensional leaves. We recall one of them introduced in [Kaveh 2015]. One can construct a valuation on the function field $\mathbb{C}(X)$ using a regular system of parameters $u_1, \dots, u_r$ in a neighborhood of a smooth point p on X. Fix a total ordering on $\mathbb{Z}^r$ respecting the addition. Let f be a polynomial in $u_1, \dots, u_r$. Suppose that $c_k u_1^{k_1} \cdots u_r^{k_r}$ is the term in f with the largest exponent $k = (k_1, \dots, k_r)$. Then

$$v(f) := (-k_1, \dots, -k_r)$$

defines a valuation on $\mathbb{C}(X)$, called the *highest term valuation* with respect to the parameters $u_1, \dots, u_r$.

Example 2.17. Let $X = Z_{\boldsymbol{i}}$ be the Bott–Samelson variety determined by a word $\boldsymbol{i} = (i_1, \dots, i_r)$. Let f_i be a nonzero element in $\mathfrak{g}_{-\alpha_i}$. Then the following map $\Phi_{\boldsymbol{i}} : \mathbb{C}^r \to Z_{\boldsymbol{i}}$ defines a coordinate system as in [Fujita 2018, §2.3; Kaveh 2015, §2.2]:

$$\Phi_{\boldsymbol{i}} : (t_1, \dots, t_r) \mapsto (\exp(t_1 f_{i_1}), \dots, \exp(t_r f_{i_r})) \mod B^r.$$

We denote the highest term valuation with respect to the lexicographic order on $\mathbb{Z}^r$ by $v_{\boldsymbol{i}}^{\text{high}}$.

There are some results on computing the Newton–Okounkov bodies using the valuation $v_{\boldsymbol{i}}^{\text{high}}$. We recall a result of Kaveh [2015]:

Example 2.18. Let $X = G/B$ be the full flag variety, and let $\mathcal{L}$ be the line bundle over X given by a dominant weight λ. Suppose that $\boldsymbol{i} = (i_1, \dots, i_m)$ is a reduced word for the longest element in the Weyl group W of G. Then the Bott–Samelson variety $Z_{\boldsymbol{i}}$ and the full flag variety G/B are birational by Proposition 2.4. Hence their function fields are isomorphic, i.e., $\mathbb{C}(Z_{\boldsymbol{i}}) \cong \mathbb{C}(G/B)$. Using the valuation $v_{\boldsymbol{i}}^{\text{high}}$ in Example 2.17, Kaveh [2015, Corollary 4.2] proves that the Newton–Okounkov body $\Delta(G/B, \mathcal{L}, v_{\boldsymbol{i}}^{\text{high}})$ can be identified with the string polytope.

The following lemma directly comes from the definition of Newton–Okounkov bodies.

Lemma 2.19. *Let $f : X \to Y$ be a birational morphism between varieties of dimension r, and let $\mathcal{L}$ be a line bundle on Y generated by global sections. Suppose that the canonical morphism $H^0(Y, \mathcal{L}^{\otimes k}) \to H^0(X, f^*\mathcal{L}^{\otimes k})$ is an isomorphism for every $k > 0$. Then their Newton–Okounkov bodies coincide, i.e.,*

$$\Delta(X, f^*\mathcal{L}, v, f^*\tau) = \Delta(Y, \mathcal{L}, v, \tau)$$

for any valuation $v : \mathbb{C}(X) \setminus \{0\} \to \mathbb{Z}^r$ and $\tau \in H^0(Y, \mathcal{L}) \setminus \{0\}$. Here v is regarded also as a valuation on $\mathbb{C}(Y)$ under the isomorphism $\mathbb{C}(Y) \cong \mathbb{C}(X)$.

Now we define left actions of P_{I_1} on $Z_{\mathcal{I}}$ and $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ by

$$\begin{aligned} p \cdot [p_1, \dots, p_r] &:= [pp_1, p_2, \dots, p_r], \\ p \cdot [p_1, \dots, p_r, v] &:= [pp_1, p_2, \dots, p_r, v] \end{aligned} \tag{2-26}$$

for $p, p_1 \in P_{I_1}, \ p_2 \in P_{I_2}, \dots, p_r \in P_{I_r}$, and $v \in \mathbb{C}$. Since the projection $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \twoheadrightarrow Z_{\mathcal{I}}$ is compatible with these actions, it follows that the space $H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})$ of global sections has the natural P_{I_1}-module structure.

Theorem 2.20. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$, and let $\boldsymbol{i} = (i_{k,l})_{1 \le k \le r, 1 \le l \le N_k} \in [n]^{N_1 + \cdots + N_r}$ be a sequence such that $(i_{k,1}, \dots, i_{k,N_k})$ is a reduced word for the longest element in W_{I_k} for $1 \le k \le r$. Let $\eta_{\boldsymbol{i},\mathcal{I}} : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$ be the birational morphism in Proposition 2.7. Then for integral weights $\lambda_k :=$*

$d_{k,1}\varpi_1+\cdots+d_{k,n}\varpi_n$ for $1\le k\le r$, and the corresponding integer vector $\boldsymbol{a}$ given in Proposition 2.10,

(1) *the canonical morphism $H^0(Z_{\mathcal{I}},\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})\to H^0(Z_{\boldsymbol{i}},\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}})$ is an isomorphism.*

(2) *The isomorphism in* (1) *induces the B-module isomorphism*

$$H^0(Z_{\mathcal{I}},\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})\cong H^0(Z_{\boldsymbol{i}},\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}})\otimes\mathbb{C}_{-\mu},$$

where μ is the weight defined by

$$\mu=\sum_{j=1}^{r}\ \sum_{s\in[n]\setminus\{i_{k,l}\mid 1\le k\le j,1\le l\le N_k\}} d_{j,s}\varpi_s.$$

To prove the theorem, we recall the following lemma.

Lemma 2.21 [Jantzen 2003, II.14.5.(a)]. *Let $\varphi:Y\to X$ be a dominant and projective morphism of noetherian and integral schemes such that φ induces an isomorphism $\mathbb{C}(X)\xrightarrow{\sim}\mathbb{C}(Y)$ of function fields. If X is normal, then $\varphi_*\mathcal{O}_Y=\mathcal{O}_X$.*

Proof of Theorem 2.20. (1) Because of Propositions 2.3 and 2.7, the morphism $\eta=\eta_{\boldsymbol{i},\mathcal{I}}:Z_{\boldsymbol{i}}\to Z_{\mathcal{I}}$ satisfies all the conditions in Lemma 2.21. Hence we have that

$$\eta_*\mathcal{O}_{Z_{\boldsymbol{i}}}=\mathcal{O}_{Z_{\mathcal{I}}}. \tag{2-27}$$

Then we have the following:

$$\begin{aligned}\eta_*(\eta^*\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})&=\eta_*(\mathcal{O}_{Z_{\boldsymbol{i}}}\otimes_{\mathcal{O}_{Z_{\boldsymbol{i}}}}\eta^*\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})\\&\cong\eta_*\mathcal{O}_{Z_{\boldsymbol{i}}}\otimes_{\mathcal{O}_{Z_{\mathcal{I}}}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}\quad\text{(by [Hartshorne 1977, Exercise II.5.1(d)])}\\&=\mathcal{O}_{Z_{\mathcal{I}}}\otimes_{\mathcal{O}_{Z_{\mathcal{I}}}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}\quad\text{(by (2-27))}\\&=\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}.\end{aligned}$$

Taking global sections we have an isomorphism between $H^0(Z_{\mathcal{I}},\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})$ and $H^0(Z_{\boldsymbol{i}},\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})$ as $\mathbb{C}$-vector spaces. And the later one is isomorphic to $H^0(Z_{\boldsymbol{i}},\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}})$ as $\mathbb{C}$-vector spaces by Proposition 2.10.

(2) Note that there is a bijective correspondence between the set $H^0(Z_{\mathcal{I}},\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})$ of holomorphic sections and the set of morphisms $f:\boldsymbol{P}_{\mathcal{I}}\to\mathbb{C}$ satisfying

$$f((p_1,\dots,p_r)\cdot(b_1,\dots,b_r))=e^{\lambda_1}(b_1)\cdots e^{\lambda_r}(b_r)f(p_1,\dots,p_r) \tag{2-28}$$

for $(p_1,\dots,p_r)\in\boldsymbol{P}_{\mathcal{I}}$ and $(b_1,\dots,b_r)\in B^r$. Indeed, a morphism f defines a section $[p_1,\dots,p_r]\mapsto[p_1,\dots,p_r,f(p_1,\dots,p_r)]$. Using C and C' defined in the proof of Proposition 2.10 (see (2-18) and (2-22)), for a morphism f satisfying

(2-28), we associate a morphism $\tilde{f} : \boldsymbol{P_i} \to \mathbb{C}$

$$\tilde{f}((p_{k,l})_{k,l}) = C'Cf\left(\prod_{l=1}^{N_1} p_{1,l}, \ldots, \prod_{l=1}^{N_r} p_{r,l}\right)$$

which also gives a section in $H^0(Z_{\boldsymbol{i}}, \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}})$. Actually, this association is the isomorphism in (1).

On the other hand, the left action of P_{I_1} on $Z_{\mathcal{I}}$ and that of $P_{i_{1,1}}$ on $Z_{\boldsymbol{i}}$ given in (2-26) define actions of B on the sets of holomorphic sections. For $b \in B$, $f : \boldsymbol{P}_{\mathcal{I}} \to \mathbb{C}$, and $\tilde{f} : \boldsymbol{P_i} \to \mathbb{C}$, we set

$$(b \cdot f)(p_1, \ldots, p_r) := f(b^{-1}p_1, p_2, \ldots, p_r),$$
$$(b \cdot \tilde{f})((p_{k,l})_{k,l}) := \tilde{f}(b^{-1}p_{1,1}, p_{1,2}, \ldots, p_{1,N_1}, \ldots, p_{r,N_r}).$$

Recall from (2-22) that C' is the product of $e^{d_{j,s}\varpi_s}(\zeta(j,s))^{-1}$. For each $s \in [n]$ and $j \in [r]$, by (2-20), the following three conditions are equivalent:

- $p_{1,1}$ is involved in $\zeta(j,s)$;
- $k(j,s) = 0$;
- $s \in [n] \setminus \{i_{1,1}, \ldots, i_{1,N_1}, \ldots, i_{j,N_j}\}$.

Using this observation, we obtain that

$$\begin{aligned}
&(b \cdot \tilde{f})((p_{k,l})_{k,l}) \\
&\quad = \tilde{f}(b^{-1}p_{1,1}, p_{1,2}, \ldots, p_{1,N_1}, \ldots, p_{r,N_r}) \\
&\quad = \left(\prod_{j=1}^{r} \prod_{s \in [n] \setminus \{i_{k,l} \mid 1 \le k \le j, 1 \le l \le N_k\}} e^{d_{j,s}\varpi_s}(b)\right) C'Cf\left(b^{-1}\prod_{l=1}^{N_1} p_{1,l}, \ldots, \prod_{l=1}^{N_r} p_{r,l}\right) \\
&\quad = e^{\mu}(b)\big(\widetilde{b \cdot f}((p_{k,l})_{k,l})\big),
\end{aligned}$$

where C and C' are values determined by $(p_{k,l})_{k,l}$, and μ is the weight given in the statement. This proves the desired equality $\widetilde{b \cdot f} = e^{-\mu}(b)(b \cdot \tilde{f})$. □

As a direct consequence of Theorem 2.20(1) and Lemma 2.19 we have the following theorem.

Theorem 2.22. *Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$, and let $\boldsymbol{i} = (i_{k,l})_{1 \le k \le r, 1 \le l \le N_k} \in [n]^{N_1 + \cdots + N_r}$ be a sequence such that $(i_{k,1}, \ldots, i_{k,N_k})$ is a reduced word for the longest element in W_{I_k} for $1 \le k \le r$. Let $\eta_{\boldsymbol{i},\mathcal{I}} : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$ be the birational morphism defined in Proposition 2.7. Then for integral dominant weights λ_k, $1 \le k \le r$, a valuation v on $\mathbb{C}(Z_{\mathcal{I}})$, and a nonzero section $\tau \in H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r})$, we have the equality*

$$\Delta(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}, v, \tau) = \Delta(Z_{\boldsymbol{i}}, \eta_{\boldsymbol{i},\mathcal{I}}^* \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}, v, \eta_{\boldsymbol{i},\mathcal{I}}^* \tau).$$

Remark 2.23. Even if the line bundle $\mathcal{L} = \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$ is very ample, the pullback bundle $\eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}$ is not necessarily very ample when $Z_{\mathcal{I}}$ is not a Bott–Samelson variety (see Remark 2.12). Therefore the real dimension of Newton–Okounkov body $\Delta(Z_{\boldsymbol{i}}, \eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}, v)$ can possibly be smaller than the complex dimension of $Z_{\boldsymbol{i}}$ as is mentioned in Remark 2.16. However, by Theorem 2.22, when $\mathcal{L}$ is very ample, we can see that

$$\dim_{\mathbb{R}} \Delta(Z_{\boldsymbol{i}}, \eta^*_{\boldsymbol{i},\mathcal{I}}\mathcal{L}, v) = \dim_{\mathbb{R}} \Delta(Z_{\mathcal{I}}, \mathcal{L}, v) = \dim_{\mathbb{C}} Z_{\mathcal{I}} = \dim_{\mathbb{C}} Z_{\boldsymbol{i}}$$

for any valuation v which has one-dimensional leaves.

By Theorem 2.22 and [Fujita 2018, Corollary 5.4], we have the following corollary.

Corollary 2.24. *Suppose that the line bundle $\mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$ constructed by weights $\lambda_1, \ldots, \lambda_r$ is very ample. Then, the Newton–Okounkov body $\Delta(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}, v^{\mathrm{high}}_{\boldsymbol{i}})$ is a rational convex polytope of real dimension equal to the complex dimension of $Z_{\mathcal{I}}$.*

3. Applications to representation theory

In this section, we give applications of Newton–Okounkov bodies of flag Bott–Samelson varieties to representation theory, using the theory of generalized string polytopes introduced in [Fujita 2018]. We restrict ourselves to a specific class of flag Bott–Samelson varieties $Z_{\mathcal{I}}$, that is, to the case of a sequence $\mathcal{I} = (I_1, \ldots, I_r)$ of subsets of $[n]$ such that $I_1 = [n]$. In this case, we have $P_{I_1} = P_{[n]} = G$. Hence the space $H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r})$ of global sections has a natural G-module structure. Let

$$\chi(H) := \mathbb{Z}\varpi_1 + \cdots + \mathbb{Z}\varpi_n$$

be the character lattice, and let

$$\chi_+(H) := \mathbb{Z}_{\geq 0}\varpi_1 + \cdots + \mathbb{Z}_{\geq 0}\varpi_n$$

be the set of integral dominant weights. Fix nonzero elements $e_i \in \mathfrak{g}_{\alpha_i},\ f_i \in \mathfrak{g}_{-\alpha_i}$ for $i \in [n]$. For $\lambda \in \chi_+(H)$, let $V(\lambda)$ denote the irreducible highest weight G-module over $\mathbb{C}$ with the highest weight λ, and let $v_\lambda \in V(\lambda)$ be a highest weight vector. Recall that every finite-dimensional irreducible G-module is isomorphic to $V(\lambda)$ for some $\lambda \in \chi_+(H)$, see [Humphreys 1975, §31.3], and that every finite-dimensional G-module is completely reducible, that is, isomorphic to a direct sum of irreducible G-modules (see [Humphreys 1975, §14.3]). For $\lambda_1, \ldots, \lambda_r \in \chi_+(H)$, we denote by $\tau_{\mathcal{I},\lambda_1,\ldots,\lambda_r} \in H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r})$ the section corresponding to $\tau_{\boldsymbol{i},\boldsymbol{a}} \in H^0(Z_{\boldsymbol{i}}, \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}})$ under the isomorphism in Theorem 2.20 (1), where $\tau_{\boldsymbol{i},\boldsymbol{a}}$ is the section defined in

[Fujita 2018, §2.3]. Let $\pi_{\geq 2} : \mathbb{R}^{N_1+\cdots+N_r} \twoheadrightarrow \mathbb{R}^{N_2+\cdots+N_r}$ be the canonical projection given by $\pi_{\geq 2}((x_{k,l})_{1\leq k\leq r, 1\leq l\leq N_k}) := (x_{k,l})_{2\leq k\leq r, 1\leq l\leq N_k}$, and set

$$\hat{\Delta}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} := \pi_{\geq 2}\bigl(-\Delta(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}, v_{\boldsymbol{i}}^{\mathrm{high}}, \tau_{\mathcal{I},\lambda_1,\ldots,\lambda_r})\bigr).$$

Since $\Delta(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}, v_{\boldsymbol{i}}^{\mathrm{high}}, \tau_{\mathcal{I},\lambda_1,\ldots,\lambda_r})$ is a rational convex polytope, the image $\hat{\Delta}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r}$ is also a rational convex polytope. The following is the main result in this section.

Theorem 3.1. *Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$ such that $I_1 = [n]$, and fix $\boldsymbol{i} = (i_{1,1}, \ldots, i_{1,N_1}, \ldots, i_{r,1}, \ldots, i_{r,N_r}) \in [n]^{N_1+\cdots+N_r}$ such that $(i_{k,1}, \ldots, i_{k,N_k})$ is a reduced word for the longest element in W_{I_k} for $1 \leq k \leq r$. For $\lambda_1, \ldots, \lambda_r \in \chi_+(H)$, write*

$$H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r})^* \simeq \bigoplus_{\nu\in\chi_+(H)} V(\nu)^{\oplus c^{\nu}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}}$$

as a G-module. Then, the multiplicity $c^{\nu}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$ equals the cardinality of

$$\Bigl\{\boldsymbol{x} = (x_{k,l})_{2\leq k\leq r, 1\leq l\leq N_k} \in \hat{\Delta}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \cap \mathbb{Z}^{N_2+\cdots+N_r} \;\Big|\; \lambda_1 + \cdots + \lambda_r - \sum_{2\leq k\leq r, 1\leq l\leq N_k} x_{k,l}\alpha_{i_{k,l}} = \nu\Bigr\}.$$

Remark 3.2. Since $\Delta(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}, v_{\boldsymbol{i}}^{\mathrm{high}}, \tau_{\mathcal{I},\lambda_1,\ldots,\lambda_r}) = \Delta(Z_{\boldsymbol{i}}, \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}, v_{\boldsymbol{i}}^{\mathrm{high}}, \tau_{\boldsymbol{i},\boldsymbol{a}})$ by Theorem 2.22, it is natural to ask why we consider not only $Z_{\boldsymbol{i}}$ but also $Z_{\mathcal{I}}$. The reason is that the space $H^0(Z_{\boldsymbol{i}}, \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}})$ of global sections does not have a natural G-module structure because $Z_{\boldsymbol{i}}$ is not a G-variety. The theory of flag Bott–Samelson varieties gives a natural framework to relate the usual Bott–Samelson variety $Z_{\boldsymbol{i}}$ with G-modules.

In order to prove Theorem 3.1, we use the theory of crystal bases, see [Kashiwara 1995] for a survey on this topic. Lusztig [1990; 1991; 1993] and Kashiwara [1991] constructed a specific $\mathbb{C}$-basis of $V(\lambda)$ via the quantized enveloping algebra associated with $\mathfrak{g}$. This is called (the specialization at $q = 1$ of) the *lower global basis* (= the *canonical basis*), and denoted by $\{G_\lambda^{\mathrm{low}}(b) \mid b \in \mathcal{B}(\lambda)\} \subset V(\lambda)$. See, for example, [Kashiwara 1995, §12] for the definition of $G_\lambda^{\mathrm{low}}(b)$. In this manuscript, we put "low" to emphasize that we are considering the *lower* global basis while Kashiwara [1995] denoted it by $G_\lambda(b)$. The index set $\mathcal{B}(\lambda)$ is endowed with specific maps

$$\mathrm{wt} : \mathcal{B}(\lambda) \to \chi(H), \quad \varepsilon_i, \varphi_i : \mathcal{B}(\lambda) \to \mathbb{Z}_{\geq 0},$$
$$\tilde{e}_i, \tilde{f}_i : \mathcal{B}(\lambda) \to \mathcal{B}(\lambda) \cup \{0\} \quad \text{for } i \in [n],$$

which have the following properties:

$$\begin{aligned}
\mathrm{wt}(b_\lambda) &= \lambda, \\
\mathrm{wt}(\tilde{e}_i b) &= \mathrm{wt}(b) + \alpha_i && \text{if } \tilde{e}_i b \neq 0, \\
\mathrm{wt}(\tilde{f}_i b) &= \mathrm{wt}(b) - \alpha_i && \text{if } \tilde{f}_i b \neq 0, \\
\varepsilon_i(b) &= \max\{k \in \mathbb{Z}_{\geq 0} \mid \tilde{e}_i^k b \neq 0\}, \\
\varphi_i(b) &= \max\{k \in \mathbb{Z}_{\geq 0} \mid \tilde{f}_i^k b \neq 0\}, \\
e_i \cdot G_\lambda^{\mathrm{low}}(b) &\in \mathbb{C}^* G_\lambda^{\mathrm{low}}(\tilde{e}_i b) + \sum_{\substack{b' \in \mathcal{B}(\lambda);\ \mathrm{wt}(b') = \mathrm{wt}(b) + \alpha_i \\ \varphi_i(b') > \varphi_i(b) + 1}} \mathbb{C} G_\lambda^{\mathrm{low}}(b'), \\
f_i \cdot G_\lambda^{\mathrm{low}}(b) &\in \mathbb{C}^* G_\lambda^{\mathrm{low}}(\tilde{f}_i b) + \sum_{\substack{b' \in \mathcal{B}(\lambda);\ \mathrm{wt}(b') = \mathrm{wt}(b) - \alpha_i \\ \varepsilon_i(b') > \varepsilon_i(b) + 1}} \mathbb{C} G_\lambda^{\mathrm{low}}(b')
\end{aligned}$$

for $i \in [n]$ and $b \in \mathcal{B}(\lambda)$, where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, and $b_\lambda \in \mathcal{B}(\lambda)$ is defined as $G_\lambda^{\mathrm{low}}(b_\lambda) \in \mathbb{C}^* v_\lambda$, called the *highest element*. We call $\mathcal{B}(\lambda)$ the *crystal basis* for $V(\lambda)$, which satisfies the axiom of *crystals*, see [Kashiwara 1993, Definition 1.2.1] for the definition of crystals. The operations $\tilde{e}_i$ and $\tilde{f}_i$ are called the *Kashiwara operators*.

Definition 3.3 (see [Kashiwara 1995, §4.2]). The *crystal graph* of a crystal $\mathcal{B}$ is the $[n]$-colored, directed graph with vertex set $\mathcal{B}$ whose directed edges are given by: $b \xrightarrow{i} b'$ if and only if $b' = \tilde{f}_i b$.

In this paper, we identify a crystal $\mathcal{B}$ with its crystal graph. By [Kashiwara 1991, Theorem 3], for a G-module $V = V(\nu_1) \oplus \cdots \oplus V(\nu_M)$, the crystal graph of the corresponding crystal basis $\mathcal{B}(V)$ is the disjoint union of the crystal graphs $\mathcal{B}(\nu_1), \ldots, \mathcal{B}(\nu_M)$.

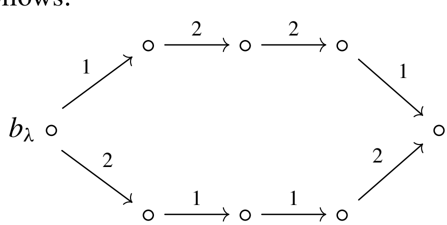
Proposition 3.4 (see [Kashiwara 1993, Proposition 3.2.3]). *Let* $\boldsymbol{i} = (i_1, \ldots, i_r) \in [n]^r$ *be a reduced word for* $w \in W$, *and* $\lambda \in \chi_+(H)$. *Then, the subset*

$$\mathcal{B}_w(\lambda) := \left\{ \tilde{f}_{i_1}^{x_1} \cdots \tilde{f}_{i_r}^{x_r} b_\lambda \mid x_1, \ldots, x_r \in \mathbb{Z}_{\geq 0} \right\} \setminus \{0\} \subset \mathcal{B}(\lambda)$$

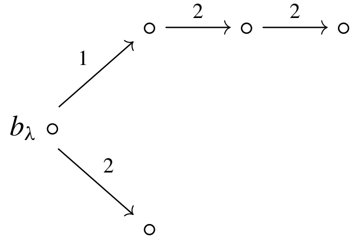
is independent of the choice of a reduced word $\boldsymbol{i}$.

The subset $\mathcal{B}_w(\lambda)$ is called a *Demazure crystal*.

Example 3.5. Let $G = \mathrm{SL}(3)$, and $\lambda = \alpha_1 + \alpha_2 = \varpi_1 + \varpi_2$. Then, the crystal graph of $\mathcal{B}(\lambda)$ is given as follows:



In addition, for $w = s_2 s_1 \in W$, the following directed graph gives the Demazure crystal $\mathcal{B}_w(\lambda)$:



The following is an immediate consequence of [Kashiwara 1993, Proposition 3.2.3].

Lemma 3.6. *Let $\boldsymbol{i} = (i_1, \ldots, i_N) \in [n]^N$ be a reduced word for the longest element $w_0 \in W$. Then, the following equalities hold for all $\lambda \in \chi_+(H)$:*

$$\mathcal{B}(\lambda) = \mathcal{B}_{w_0}(\lambda) = \left\{ \tilde{f}_{i_1}^{x_1} \cdots \tilde{f}_{i_N}^{x_N} b_\lambda \mid x_1, \ldots, x_N \in \mathbb{Z}_{\geq 0} \right\} \setminus \{0\}.$$

In particular, the following equality holds for all $w \in W$:

$$\left\{ \tilde{f}_{i_1}^{x_1} \cdots \tilde{f}_{i_N}^{x_N} b \mid x_1, \ldots, x_N \in \mathbb{Z}_{\geq 0},\ b \in \mathcal{B}_w(\lambda) \right\} \setminus \{0\} = \mathcal{B}(\lambda).$$

For two crystals $\mathcal{B}_1, \mathcal{B}_2$, we can define another crystal $\mathcal{B}_1 \otimes \mathcal{B}_2$, called the *tensor product* of $\mathcal{B}_1$ and $\mathcal{B}_2$, see [Kashiwara 1993, §1.3] for the definition. For $\lambda_1, \ldots, \lambda_r \in \chi_+(H)$, the tensor product $\mathcal{B}(\lambda_1) \otimes \cdots \otimes \mathcal{B}(\lambda_r)$ is identical to the crystal basis for the tensor product module $V(\lambda_1) \otimes \cdots \otimes V(\lambda_r)$ by [Kashiwara 1991, Theorem 1]. Let us recall the definitions of generalized Demazure crystals and generalized string polytopes.

Definition 3.7 (see [Lakshmibai et al. 2002, §1.2]). Let $\boldsymbol{i} = (i_1, \ldots, i_r) \in [n]^r$ be an arbitrary word, and $\boldsymbol{a} = (a_1, \ldots, a_r) \in \mathbb{Z}_{\geq 0}^r$. We define

$$\mathcal{B}_{\boldsymbol{i}, \boldsymbol{a}} \subset \mathcal{B}(a_1 \varpi_{i_1}) \otimes \cdots \otimes \mathcal{B}(a_r \varpi_{i_r})$$

to be the subset

$$\Big\{ \tilde{f}_{i_1}^{x_1}\big(b_{a_1\varpi_{i_1}} \otimes \tilde{f}_{i_2}^{x_2}\big(b_{a_2\varpi_{i_2}} \otimes \cdots \otimes \tilde{f}_{i_{r-1}}^{x_{r-1}}(b_{a_{r-1}\varpi_{i_{r-1}}} \otimes \tilde{f}_{i_r}^{x_r}(b_{a_r\varpi_{i_r}})) \cdots \big)\big) \mid x_1, \ldots, x_r \in \mathbb{Z}_{\geq 0} \Big\} \setminus \{0\};$$

this is called a *generalized Demazure crystal*.

Definition 3.8 [Fujita 2018, Definition 4.4]. Let $\boldsymbol{i} = (i_1, \ldots, i_r) \in [n]^r$ be an arbitrary word, and $\boldsymbol{a} = (a_1, \ldots, a_r) \in \mathbb{Z}_{\geq 0}^r$. For $b \in \mathcal{B}_{\boldsymbol{i}, \boldsymbol{a}}$, we set $b(1) := b$,

$$x_1 := \max\{x \in \mathbb{Z}_{\geq 0} \mid \tilde{e}_{i_1}^{x} b(1) \neq 0\}, \quad \tilde{e}_{i_1}^{x_1} b(1) = b_{a_1\varpi_{i_1}} \otimes b(2),$$
$$x_2 := \max\{x \in \mathbb{Z}_{\geq 0} \mid \tilde{e}_{i_2}^{x} b(2) \neq 0\}, \quad \tilde{e}_{i_2}^{x_2} b(2) = b_{a_2\varpi_{i_2}} \otimes b(3),$$
$$\vdots$$

$$x_r := \max\{x \in \mathbb{Z}_{\geq 0} \mid \tilde{e}_{i_r}^{x} b(r) \neq 0\},$$

and define the *generalized string parametrization* $\Omega_{\boldsymbol{i}}(b)$ of b with respect to $\boldsymbol{i}$ by $\Omega_{\boldsymbol{i}}(b) := (x_1, \ldots, x_r)$.

Definition 3.9 [Fujita 2018, Definition 4.7]. For an arbitrary word $\boldsymbol{i} \in [n]^r$ and $\boldsymbol{a} \in \mathbb{Z}_{\geq 0}^r$, define a subset $\mathcal{S}_{\boldsymbol{i},\boldsymbol{a}} \subset \mathbb{Z}_{>0} \times \mathbb{Z}^r$ by

$$\mathcal{S}_{\boldsymbol{i},\boldsymbol{a}} := \bigcup_{k>0} \{(k, \Omega_{\boldsymbol{i}}(b)) \mid b \in \mathcal{B}_{\boldsymbol{i},k\boldsymbol{a}}\},$$

and denote by $\mathcal{C}_{\boldsymbol{i},\boldsymbol{a}} \subset \mathbb{R}_{\geq 0} \times \mathbb{R}^r$ the smallest real closed cone containing $\mathcal{S}_{\boldsymbol{i},\boldsymbol{a}}$. Let us define a subset $\Delta_{\boldsymbol{i},\boldsymbol{a}} \subset \mathbb{R}^r$ by

$$\Delta_{\boldsymbol{i},\boldsymbol{a}} := \{\boldsymbol{x} \in \mathbb{R}^r \mid (1, \boldsymbol{x}) \in \mathcal{C}_{\boldsymbol{i},\boldsymbol{a}}\};$$

this is called the *generalized string polytope* associated to $\boldsymbol{i}$ and $\boldsymbol{a}$.

The following is a fundamental property of generalized string polytopes.

Proposition 3.10 (see [Fujita 2018, Corollaries 4.16, 5.4(3)]). *The generalized string polytope $\Delta_{\boldsymbol{i},\boldsymbol{a}}$ is a rational convex polytope, and the equality $\Omega_{\boldsymbol{i}}(\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}) = \Delta_{\boldsymbol{i},\boldsymbol{a}} \cap \mathbb{Z}^r$ holds.*

Fujita proved the following relation between the generalized string polytope and a Newton–Okounkov body of the Bott–Samelson variety $Z_{\boldsymbol{i}}$.

Theorem 3.11 (see [Fujita 2018, Corollary 5.3]). *Let $Z_{\boldsymbol{i}}$ be the Bott–Samelson variety determined by a word $\boldsymbol{i} \in [n]^r$, and let $\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$ be the line bundle on $Z_{\boldsymbol{i}}$ determined by an integer vector $\boldsymbol{a} \in \mathbb{Z}_{\geq 0}^r$ as in (2-14). Then we have that*

$$\Delta(Z_{\boldsymbol{i}}, \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}, v_{\boldsymbol{i}}^{\text{high}}, \tau_{\boldsymbol{i},\boldsymbol{a}}) = -\Delta_{\boldsymbol{i},\boldsymbol{a}}.$$

Remark 3.12. The combinatorial structure of generalized string polytopes is quite complicated that even their real dimensions are not easy to be determined. By Remark 2.23, Theorem 3.11 determines the dimensions of generalized string polytopes of the type $\Delta(Z_{\boldsymbol{i}}, \eta_{\boldsymbol{i},\mathcal{I}}^{*}\mathcal{L}, v_{\boldsymbol{i}}^{\text{high}}, \tau_{\boldsymbol{i},\boldsymbol{a}})$, where $\mathcal{I}$ is a sequence of subsets of $[n]$ and $\mathcal{L}$ is a very ample line bundle over $Z_{\mathcal{I}}$.

Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$, and fix a sequence $\boldsymbol{i} = (i_{k,l})_{1 \leq k \leq r, 1 \leq l \leq N_k} \in [n]^{N_1 + \cdots + N_r}$ such that $(i_{k,1}, \ldots, i_{k,N_k})$ is a reduced word for the longest element in W_{I_k} for $1 \leq k \leq r$. Given $\lambda_1, \ldots, \lambda_r \in \chi_{+}(H)$, we denote the dual P_{I_1}-module $H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\ldots,\lambda_r})^*$ by $V_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$, and define $\mathcal{B}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \subset \mathcal{B}(\lambda_1) \otimes \cdots \otimes \mathcal{B}(\lambda_r)$ to be the set of elements of the form

$$\tilde{f}_{i_{1,1}}^{x_{1,1}} \cdots \tilde{f}_{i_{1,N_1}}^{x_{1,N_1}} \big(b_{\lambda_1} \otimes \cdots \otimes \tilde{f}_{i_{r-1,1}}^{x_{r-1,1}} \cdots \tilde{f}_{i_{r-1,N_{r-1}}}^{x_{r-1,N_{r-1}}} \big(b_{\lambda_{r-1}} \otimes \tilde{f}_{i_{r,1}}^{x_{r,1}} \cdots \tilde{f}_{i_{r,N_r}}^{x_{r,N_r}} (b_{\lambda_r})\big) \cdots \big) \tag{3-1}$$

for some $x_{1,1}, \ldots, x_{1,N_1}, \ldots, x_{r,1}, \ldots, x_{r,N_r} \in \mathbb{Z}_{\geq 0}$.

Proposition 3.13. *For $\lambda_1, \dots, \lambda_r \in \chi_+(H)$, let $\boldsymbol{a} \in \mathbb{Z}^{N_1+\cdots+N_r}$ be the integer vector such that $\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}} \simeq \eta_{\boldsymbol{i},\mathcal{I}}^* \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ as given in Proposition 2.10, and let $\mu \in \chi_+(H)$ be the weight defined in Theorem 2.20(2).*

(1) *The B-module $V_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is naturally isomorphic to $\mathbb{C}_\mu \otimes V_{\boldsymbol{i},\boldsymbol{a}}$, where $V_{\boldsymbol{i},\boldsymbol{a}}$ is the generalized Demazure module defined in [Lakshmibai et al. 2002, §1.1].*

(2) *There is a natural bijective map*

$$\mathcal{B}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} \xrightarrow{\sim} b_\mu \otimes \mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$$

compatible with the crystal structures.

(3) *The crystal graph of $\mathcal{B}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$ is identical to that of $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$.*

Proof. (1) The assertion is an immediate consequence of Theorem 2.20 and [Lakshmibai et al. 2002, Theorem 6].

(2) For $\lambda, \mu \in \chi_+(H)$, the crystal basis $\mathcal{B}(\lambda+\mu)$ can be regarded as a connected component of $\mathcal{B}(\lambda) \otimes \mathcal{B}(\mu)$ by identifying $b_{\lambda+\mu}$ with $b_\lambda \otimes b_\mu$ (see [Kashiwara 1995, §4.5]). If we identify b_λ with $b_{\lambda-\langle\lambda,\alpha_i^\vee\rangle\varpi_i} \otimes b_{\langle\lambda,\alpha_i^\vee\rangle\varpi_i}$ for $i \in [n]$ and $\lambda \in \chi_+(H)$, then the definition of tensor product crystals implies that

$$\tilde{f}_i^a b_\lambda = b_{\lambda-\langle\lambda,\alpha_i^\vee\rangle\varpi_i} \otimes \tilde{f}_i^a b_{\langle\lambda,\alpha_i^\vee\rangle\varpi_i} \quad \text{for all } a \in \mathbb{Z}_{\geq 0}$$

(see [Fujita 2018, Appendix A]). Hence it follows that

$$\begin{aligned}\tilde{f}_{i_{1,1}}^{x_{1,1}} \cdots \tilde{f}_{i_{1,N_1}}^{x_{1,N_1}} (b_{\lambda_1} \otimes b) \\ = b_{\lambda_1 - \sum_{1\leq l\leq N_1} \mu_l} \otimes \tilde{f}_{i_{1,1}}^{x_{1,1}} \big(b_{\mu_1} \otimes \tilde{f}_{i_{1,2}}^{x_{1,2}} \big(b_{\mu_2} \otimes \cdots \otimes \tilde{f}_{i_{1,N_1}}^{x_{1,N_1}} (b_{\mu_{N_1}} \otimes b) \cdots \big)\big)\end{aligned}$$

for $b \in \mathcal{B}_{\boldsymbol{i}_{\geq 2},\lambda_2,\dots,\lambda_r}$ and $x_{1,1}, \dots, x_{1,N_1} \in \mathbb{Z}_{\geq 0}$, where

$$\mu_l := \begin{cases} \langle\lambda_1, \alpha_{i_{1,l}}^\vee\rangle\varpi_{i_{1,l}} & \text{if } l = \max\{1 \leq q \leq N_1 \mid i_{1,q} = i_{1,l}\}, \\ 0 & \text{otherwise} \end{cases}$$

for $1 \leq l \leq N_1$, and $\boldsymbol{i}_{\geq 2} := (i_{k,l})_{2\leq k\leq r, 1\leq l\leq N_k}$. By repeating this deformation, all the elements of the form (3-1) can be naturally written as elements in $b_\mu \otimes \mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$. This proves part (2).

(3) Let us prove that $\tilde{e}_i(b_\mu \otimes b) = b_\mu \otimes \tilde{e}_i b$ for all $i \in [n]$ and $b \in \mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$. By the definition of $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$, we have

$$\mathrm{wt}(b) - \mathrm{wt}(b') \in \sum_{j\in\{i_{k,l} \mid 1\leq k\leq r,\ 1\leq l\leq N_k\}} \mathbb{Z}\alpha_j$$

for all $b, b' \in \mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$. Hence $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$ does not have edges labeled by

$$j \notin \{i_{k,l} \mid 1 \leq k \leq r,\ 1 \leq l \leq N_k\}.$$

From this, we may assume that $i \in \{i_{k,l} \mid 1 \le k \le r,\ 1 \le l \le N_k\}$. Then, we have $\langle \mu, \alpha_i^\vee \rangle = 0$ by the definition of μ, which implies by the definition of tensor product crystals that $\tilde{e}_i(b_\mu \otimes b) = b_\mu \otimes \tilde{e}_i b$. Thus, we have proved that the crystal graph of $b_\mu \otimes \mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$ is identical to that of $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$. Then, part (3) follows immediately from part (2). □

Proposition 3.13 implies that all the results in [Lakshmibai et al. 2002] for $V_{\boldsymbol{i},\boldsymbol{a}}$ and $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$ are applicable also for $V_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ and $\mathcal{B}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$.

Proposition 3.14. *The set $\mathcal{B}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$ depends only on $\mathcal{I}, \lambda_1, \dots, \lambda_r$, that is, does not depend on the choice of $\boldsymbol{i}$.*

Proof. We proceed by induction on r. If $r = 1$, then the assertion is an immediate consequence of Proposition 3.4. Assume that $r \ge 2$, and that $\mathcal{B}_{\boldsymbol{i}_{\ge 2},\lambda_2,\dots,\lambda_r}$ is independent of the choice of $\boldsymbol{i}_{\ge 2}$. By [Lakshmibai et al. 2002, Theorem 2] and Proposition 3.13, it follows that $b_{\lambda_1} \otimes \mathcal{B}_{\boldsymbol{i}_{\ge 2},\lambda_2,\dots,\lambda_r}$ is a disjoint union of Demazure crystals. Hence it suffices to prove that for each connected component $\mathcal{B}_v(\lambda)$ of $b_{\lambda_1} \otimes \mathcal{B}_{\boldsymbol{i}_{\ge 2},\lambda_2,\dots,\lambda_r}$ the set

$$\mathcal{B}_{v,i_{1,1},\dots,i_{1,N_1}}(\lambda) := \left\{ \tilde{f}_{i_{1,1}}^{x_1} \cdots \tilde{f}_{i_{1,N_1}}^{x_{N_1}} b \mid x_1, \dots, x_{N_1} \in \mathbb{Z}_{\ge 0},\ b \in \mathcal{B}_v(\lambda) \right\} \setminus \{0\}$$

does not depend on the choice of $(i_{1,1}, \dots, i_{1,N_1})$. We define $v_1, \dots, v_{N_1} \in W$ inductively by

$$v_1 := \begin{cases} s_{i_{1,N_1}} v & \text{if } \ell(s_{i_{1,N_1}} v) > \ell(v), \\ v & \text{if } \ell(s_{i_{1,N_1}} v) < \ell(v), \end{cases}$$
$$v_l := \begin{cases} s_{i_{1,N_1-l+1}} v_{l-1} & \text{if } \ell(s_{i_{1,N_1-l+1}} v_{l-1}) > \ell(v_{l-1}), \\ v_{l-1} & \text{if } \ell(s_{i_{1,N_1-l+1}} v_{l-1}) < \ell(v_{l-1}). \end{cases}$$

Then, we deduce by [Kashiwara 1993, Proposition 3.2.3 (iii)] that $\mathcal{B}_{v,i_{1,1},\dots,i_{1,N_1}}(\lambda) = \mathcal{B}_{v_{N_1}}(\lambda)$. In addition, it follows by [Kashiwara 1993, Lemma 3.2.1 and Proposition 3.2.3 (i)] that

$$\sum_{x_1,\dots,x_{N_1} \in \mathbb{Z}_{\ge 0}} f_{i_{1,1}}^{x_1} \cdots f_{i_{1,N_1}}^{x_{N_1}} \Bigg(\sum_{b \in \mathcal{B}_v(\lambda)} \mathbb{C} G_\lambda^{\mathrm{low}}(b) \Bigg) = \sum_{b \in \mathcal{B}_{v_{N_1}}(\lambda)} \mathbb{C} G_\lambda^{\mathrm{low}}(b).$$

From these, we have

$$\sum_{b \in \mathcal{B}_{v,i_{1,1},\dots,i_{1,N_1}}(\lambda)} \mathbb{C} G_\lambda^{\mathrm{low}}(b) = \sum_{x_1,\dots,x_{N_1} \in \mathbb{Z}_{\ge 0}} f_{i_{1,1}}^{x_1} \cdots f_{i_{1,N_1}}^{x_{N_1}} \Bigg(\sum_{b \in \mathcal{B}_v(\lambda)} \mathbb{C} G_\lambda^{\mathrm{low}}(b) \Bigg);$$

the right hand side does not depend on the choice of $(i_{1,1}, \dots, i_{1,N_1})$ by [Kashiwara 1993, Proposition 3.2.5(v)], which implies that the set $\mathcal{B}_{v,i_{1,1},\dots,i_{1,N_1}}(\lambda)$ is also independent. This proves the proposition. □

We denote $\mathcal{B}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$ by $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, which is also called a *generalized Demazure crystal*. By definition, we have

$$\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r} = \left\{ \tilde{f}_{i_{1,1}}^{x_1} \cdots \tilde{f}_{i_{1,N_1}}^{x_{N_1}} (b_{\lambda_1} \otimes b) \mid x_1, \dots, x_{N_1} \in \mathbb{Z}_{\geq 0},\ b \in \mathcal{B}_{(I_2,\dots,I_r),\lambda_2,\dots,\lambda_r} \right\} \setminus \{0\}.$$

Assume that $I_1 = [n]$, and hence that $(i_{1,1}, \dots, i_{1,N_1})$ is a reduced word for $w_0 \in W$. By [Lakshmibai et al. 2002, Theorem 2] and Proposition 3.13, the set $b_{\lambda_1} \otimes \mathcal{B}_{(I_2,\dots,I_r),\lambda_2,\dots,\lambda_r}$ is a disjoint union of Demazure crystals. Hence the second assertion of Lemma 3.6 implies that each connected component of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is of the form $\mathcal{B}(\nu)$ for some $\nu \in \chi_+(H)$. Note that the character of $V_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ equals the formal character of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ by [Lakshmibai et al. 2002, Theorem 5 and Corollary 10] and Proposition 3.13. Since finite-dimensional G-modules are characterized by their characters, we obtain the following.

Proposition 3.15. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$ such that $I_1 = [n]$. Then, the generalized Demazure crystal $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is isomorphic to the crystal basis for the G-module $V_{\mathcal{I},\lambda_1,\dots,\lambda_r}$. In particular, if $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is the disjoint union of $\mathcal{B}(\nu_1), \dots, \mathcal{B}(\nu_M)$, then $V_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is isomorphic to $V(\nu_1) \oplus \cdots \oplus V(\nu_M)$.*

Since, by Proposition 3.13(3), the crystal graph of $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$ is identical to that of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, the generalized string parametrization $\Omega_{\boldsymbol{i}}$ of $\mathcal{B}_{\boldsymbol{i},\boldsymbol{a}}$ can be regarded as a parametrization of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$. We denote $\Delta_{\boldsymbol{i},\boldsymbol{a}}$ by $\Delta_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$. Then, we have $\hat{\Delta}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} = \pi_{\geq 2}(\Delta_{\boldsymbol{i},\lambda_1,\dots,\lambda_r})$ by Theorems 2.22, 3.11.

Proof of Theorem 3.1. By Proposition 3.15, the multiplicity $c^{\nu}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ equals the number of connected components of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ isomorphic to $\mathcal{B}(\nu)$. For $b \in \mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, we write $\Omega_{\boldsymbol{i}}(b) = (x_{1,1}, \dots, x_{1,N_1}, \dots, x_{r,1}, \dots, x_{r,N_r})$. By the definition of $\Omega_{\boldsymbol{i}}$, we have

$$x_{1,l} = \max\left\{ x \in \mathbb{Z}_{\geq 0} \mid \tilde{e}_{i_{1,l}}^{x} \tilde{e}_{i_{1,l-1}}^{x_{1,l-1}} \cdots \tilde{e}_{i_{1,1}}^{x_{1,1}} b \neq 0 \right\} \tag{3-2}$$

for $1 \leq l \leq N_1$. Let $\mathcal{C}_b$ denote the connected component of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ containing b. Since $I_1 = [n]$, it follows that $(i_{1,1}, \dots, i_{1,N_1})$ is a reduced word for $w_0 \in W$. So we deduce by [Kashiwara 1993, Proposition 3.2.3] that $\tilde{e}_{i_{1,N_1}}^{x_{1,N_1}} \cdots \tilde{e}_{i_{1,1}}^{x_{1,1}} b$ is the highest element in $\mathcal{C}_b$. Hence

$$\left(0, \dots, 0, x_{2,1}, \dots, x_{2,N_2}, \dots, x_{r,1}, \dots, x_{r,N_r}\right) \tag{3-3}$$

is the generalized string parametrization of the highest element. In particular, the surjective map $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \twoheadrightarrow \hat{\Delta}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} \cap \mathbb{Z}^{N_2+\cdots+N_r}$ given by $b \mapsto \pi_{\geq 2}(\Omega_{\boldsymbol{i}}(b))$ induces a bijective map

$$\Psi : \{\text{connected components of } \mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}\} \xrightarrow{\sim} \hat{\Delta}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} \cap \mathbb{Z}^{N_2+\cdots+N_r}.$$

In addition, for a connected component $\mathcal{C}$ of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, the weight of the highest element in $\mathcal{C}$ is determined by $\Psi(\mathcal{C})$ due to the definition of generalized string parametrizations (see Definition 3.8). Indeed, if

$$\Psi(\mathcal{C}) = \big(x_{2,1}, \ldots, x_{2,N_2}, \ldots, x_{r,1}, \ldots, x_{r,N_r}\big),$$

then the weight of the highest element in $\mathcal{C}$ is given by

$$\lambda_1 + \cdots + \lambda_r - \sum_{2 \le k \le r,\ 1 \le l \le N_k} x_{k,l} \alpha_{i_{k,l}}$$

since the generalized string parametrization of this element is given by (3-3). By these reasons, we deduce the assertion of the theorem. □

The following is an immediate consequence of the proof of Theorem 3.1.

Corollary 3.16. *Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$ such that $I_1 = [n]$. Then, the number of connected components of $\mathcal{B}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$ equals the cardinality of*

$$\hat{\Delta}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \cap \mathbb{Z}^{N_2+\cdots+N_r}.$$

Let $\pi_1 : \mathbb{R}^{N_1+\cdots+N_r} \twoheadrightarrow \mathbb{R}^{N_1}$ denote the canonical projection given by

$$\big(x_{1,1}, \ldots, x_{1,N_1}, \ldots, x_{r,1}, \ldots, x_{r,N_r}\big) \mapsto (x_{1,1}, \ldots, x_{1,N_1}).$$

Proposition 3.17. *For $\boldsymbol{x} \in \hat{\Delta}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \cap \mathbb{Z}^{N_2+\cdots+N_r}$, the set $\pi_1(\pi_{\ge 2}^{-1}(\boldsymbol{x}) \cap \Delta_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r})$ is identical to the string polytope for the connected component $\Psi^{-1}(\boldsymbol{x})$ of $\mathcal{B}_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$ with respect to the reduced word $(i_{1,1}, \ldots, i_{1,N_1})$ for $w_0 \in W$; see* [Kaveh 2015, Definition 3.5; Littelmann 1998, §1] *for the definition of string polytopes.*

Proof. Recall that, by Proposition 3.10, $\Omega_{\boldsymbol{i}} : \mathcal{B}_{\mathcal{I},\lambda_1,\ldots,\lambda_r} \to \Delta_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \cap \mathbb{Z}^{N_1+\cdots+N_r}$ is bijective. Hence by the definition of Ψ, we obtain the following bijective map:

$$\begin{aligned} \Psi^{-1}(\boldsymbol{x}) &\to \pi_{\ge 2}^{-1}(\boldsymbol{x}) \cap \Delta_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r} \cap \mathbb{Z}^{N_1+\cdots+N_r}, \\ b &\mapsto \Omega_{\boldsymbol{i}}(b). \end{aligned}$$

In addition, we see by (3-2) that $\pi_1(\Omega_{\boldsymbol{i}}(b))$ is the string parametrization of $b \in \Psi^{-1}(\boldsymbol{x})$ with respect to the reduced word $(i_{1,1}, \ldots, i_{1,N_1})$; see [Littelmann 1998, §1; Kaveh 2015, Definition 3.2] for the definition of string parametrizations. From these, we obtain the assertion of the proposition. □

Remark 3.18. Kaveh and Khovanskii [2012a] gave a general framework to describe multiplicities of irreducible representations by using the Newton–Okounkov bodies. Our results give concrete constructions of convex bodies appearing in [Kaveh and Khovanskii 2012a]. Indeed, by the proof of Theorem 3.1 and [Fujita 2018, Theorem 5.2], it is not hard to prove that the rational convex polytope $\hat{\Delta}_{\boldsymbol{i},\lambda_1,\ldots,\lambda_r}$ is identical to the multiplicity convex body $\hat{\Delta}_G(A)$ in [Kaveh and Khovanskii 2012a,

§4.1] for the valuation $v_{\boldsymbol{i}}^{\text{high}}$, where

$$A := \bigoplus_{k\geq 0} H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}^{\otimes k}).$$

From this and Proposition 3.17, we deduce that the generalized string polytope $\Delta_{\boldsymbol{i},\lambda_1,\dots,\lambda_r}$ equals the string convex body $\tilde{\Delta}(A)$ in [Kaveh and Khovanskii 2012a, §5.2].

In representation theory, it is a fundamental problem to determine the G-module structure of the tensor product module $V(\lambda)\otimes V(\mu)$, which is equivalent to determining the multiplicity $c_{\lambda,\mu}^{\nu}$ of $V(\nu)$ in $V(\lambda)\otimes V(\mu)$. Berenstein and Zelevinsky [2001, Theorems 2.3, 2.4] describes the multiplicity $c_{\lambda,\mu}^{\nu}$ as the number of lattice points in some explicit rational convex polytope. In the following, we see that Theorem 3.1 gives a different approach to such polyhedral expressions for $c_{\lambda,\mu}^{\nu}$. Let us consider the case $\mathcal{I} = ([n],[n])$. In this case, the flag Bott–Samelson variety $Z_{\mathcal{I}}$ is identical to $G\times_B G/B$, and the following map is an isomorphism of varieties:

$$Z_{\mathcal{I}} \xrightarrow{\sim} G/B\times G/B, \quad [g_1, g_2]\mapsto (g_1B/B, g_1g_2B/B);$$

the inverse map is given by $(g_1B/B, g_2B/B)\mapsto [g_1, g_1^{-1}g_2]$. It is easily seen that under the isomorphism $Z_{\mathcal{I}}\simeq G/B\times G/B$, the G-action on $Z_{\mathcal{I}}$ coincides with the diagonal action on $G/B\times G/B$, and the line bundle $\mathcal{L}_{\mathcal{I},\lambda,\mu}$ corresponds to the direct product of $\mathcal{L}_{\lambda}$ and $\mathcal{L}_{\mu}$, where $\mathcal{L}_{\nu}$ denotes the line bundle $\mathcal{L}_{([n]),\nu}$ over G/B for $\nu\in\chi_+(H)$. Hence we obtain the following isomorphisms of G-modules:

$$\begin{aligned} H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda,\mu})^* &\simeq H^0(G/B\times G/B, \mathcal{L}_{\lambda}\times\mathcal{L}_{\mu})^* \\ &\simeq H^0(G/B,\mathcal{L}_{\lambda})^*\otimes H^0(G/B,\mathcal{L}_{\mu})^* \\ &\simeq V(\lambda)\otimes V(\mu), \end{aligned}$$

by the Borel–Weil theorem (see [Jantzen 2003, Corollary II.5.6]). If we write

$$V(\lambda)\otimes V(\mu)\simeq \bigoplus_{\nu\in\chi_+(H)} V(\nu)^{\oplus c_{\lambda,\mu}^{\nu}}$$

as a G-module, then we obtain the following by Theorem 3.1:

Theorem 3.19. *Let $\mathcal{I} = ([n],[n])$, and let $(i_1,\dots,i_N), (j_1,\dots,j_N)\in[n]^N$ be reduced words for $w_0\in W$. Then, the tensor product multiplicity $c_{\lambda,\mu}^{\nu}$ equals the cardinality of*

$$\Bigl\{(y_1,\dots,y_N)\in\hat{\Delta}_{\boldsymbol{i},\lambda,\mu}\cap\mathbb{Z}^N \Bigm| \lambda+\mu-\sum_{1\leq l\leq N} y_l\alpha_{j_l}=\nu\Bigr\},$$

where $\boldsymbol{i} := (i_1,\dots,i_N, j_1,\dots,j_N)$.

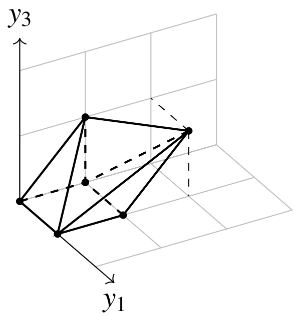


Figure 1. The polytope $\hat{\Delta}_{\boldsymbol{i},\lambda,\mu}$ in Example 3.20.

Example 3.20. Let $G = \mathrm{SL}(3)$, $\mathcal{I} = ([2], [2])$, and $\boldsymbol{i} = (1, 2, 1, 1, 2, 1)$. By [Fujita 2018, Corollary 4.15], the generalized string polytope $\Delta_{\boldsymbol{i},\lambda,\mu}$ is identical to the set of $(x_1, x_2, x_3, y_1, y_2, y_3) \in \mathbb{R}^6_{\geq 0}$ satisfying the following inequalities:

$$\begin{aligned} 0 \leq y_3 &\leq \min\{\lambda_2, \mu_1\}, \\ y_3 \leq y_2 &\leq y_3 + \mu_2, \\ y_2 - \lambda_2 \leq y_1 &\leq \min\{\lambda_1, y_2 - 2y_3 + \mu_1\}, \\ \max\{y_3 - \lambda_2, -y_1 + y_2 - \lambda_2\} \leq x_3 &\leq -2y_1 + y_2 - 2y_3 + \lambda_1 + \mu_1, \\ x_3 \leq x_2 &\leq x_3 + y_1 - 2y_2 + y_3 + \lambda_2 + \mu_2, \\ 0 \leq x_1 &\leq x_2 - 2x_3 - 2y_1 + y_2 - 2y_3 + \lambda_1 + \mu_1, \end{aligned}$$

where $\lambda_i := \langle \lambda, \alpha_i^\vee \rangle$ and $\mu_i := \langle \mu, \alpha_i^\vee \rangle$ for $i = 1, 2$. Hence the polytope $\hat{\Delta}_{\boldsymbol{i},\lambda,\mu}$ is identical to the set of $(y_1, y_2, y_3) \in \mathbb{R}^3_{\geq 0}$ satisfying the following inequalities:

$$\begin{aligned} 0 \leq y_3 &\leq \min\{\lambda_2, \mu_1\}, \\ y_3 \leq y_2 &\leq y_3 + \mu_2, \\ y_2 - \lambda_2 \leq y_1 &\leq \min\{\lambda_1, y_2 - 2y_3 + \mu_1\}. \end{aligned}$$

We deduce by Theorem 3.19 that the tensor product multiplicity $c_{\lambda,\mu}^{\nu}$ equals the cardinality of $(y_1, y_2, y_3) \in \hat{\Delta}_{\boldsymbol{i},\lambda,\mu} \cap \mathbb{Z}^3$ such that $\lambda + \mu - (y_1 + y_3)\alpha_1 - y_2\alpha_2 = \nu$.

If $\lambda = \mu = \varpi_1 + \varpi_2$, then the polytope $\hat{\Delta}_{\boldsymbol{i},\lambda,\mu}$ is identical to the set of $(y_1, y_2, y_3) \in \mathbb{R}^3_{\geq 0}$ satisfying the following inequalities:

$$0 \leq y_3 \leq 1, \quad y_3 \leq y_2 \leq y_3 + 1, \quad y_2 - 1 \leq y_1 \leq \min\{1, y_2 - 2y_3 + 1\};$$

see Figure 1. Hence we deduce that

$$V(\varpi_1 + \varpi_2)^{\otimes 2} \simeq V(2\varpi_1 + 2\varpi_2) \oplus V(3\varpi_1) \oplus V(3\varpi_2) \oplus V(\varpi_1 + \varpi_2)^{\oplus 2} \oplus V(0).$$

Theorem 3.1 can be applied to a more general class of representations than [Berenstein and Zelevinsky 2001]. We next consider the case $\mathcal{I} = ([n], [n], \dots, [n])$ (an r-tuple). In this case, we have

$$Z_{\mathcal{I}} = \underbrace{G \times_B G \times_B \cdots \times_B G}_{r}/B,$$

and this is isomorphic to $(G/B)^r := G/B \times G/B \times \cdots \times G/B$ (r factors) as follows:

$$Z_{\mathcal{I}} \xrightarrow{\sim} (G/B)^r, \quad [g_1, g_2, \dots, g_r] \mapsto \big(g_1 B/B, g_1 g_2 B/B, \dots, g_1 g_2 \cdots g_r B/B\big);$$

the inverse map is given by

$$\big(g_1 B/B, g_2 B/B, \dots, g_r B/B\big) \mapsto \big[g_1, g_1^{-1} g_2, g_2^{-1} g_3, \dots, g_{r-1}^{-1} g_r\big].$$

As in the case $r = 2$, under the isomorphism $Z_{\mathcal{I}} \simeq (G/B)^r$, the G-action on $Z_{\mathcal{I}}$ coincides with the diagonal action on $(G/B)^r$, and the line bundle $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ corresponds to the direct product of $\mathcal{L}_{\lambda_1}, \dots, \mathcal{L}_{\lambda_r}$. From this, we have the following isomorphisms of G-modules:

$$\begin{aligned} H^0(Z_{\mathcal{I}}, \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r})^* &\simeq H^0((G/B)^r, \mathcal{L}_{\lambda_1} \times \cdots \times \mathcal{L}_{\lambda_r})^* \\ &\simeq H^0(G/B, \mathcal{L}_{\lambda_1})^* \otimes \cdots \otimes H^0(G/B, \mathcal{L}_{\lambda_r})^* \\ &\simeq V(\lambda_1) \otimes \cdots \otimes V(\lambda_r). \end{aligned}$$

If we write

$$V(\lambda_1) \otimes \cdots \otimes V(\lambda_r) \simeq \bigoplus_{\nu \in \chi_+(H)} V(\nu)^{\oplus c^{\nu}_{\lambda_1,\dots,\lambda_r}}$$

as a G-module, then Theorem 3.1 implies the following.

Corollary 3.21. *Let $\mathcal{I} = ([n], [n], \dots, [n])$, an r-tuple, and take reduced words $(i_{k,1}, \dots, i_{k,N}) \in [n]^N$, $1 \le k \le r$, for $w_0 \in W$. Then, the multiplicity $c^{\nu}_{\lambda_1,\dots,\lambda_r}$ equals the cardinality of*

$$\Big\{ \boldsymbol{x} = (x_{k,l})_{2 \le k \le r, 1 \le l \le N} \in \hat{\Delta}_{\boldsymbol{i},\lambda_1,\dots,\lambda_r} \cap \mathbb{Z}^{(r-1)N} \;\Big|\; \lambda_1 + \cdots + \lambda_r - \sum_{2 \le k \le r, 1 \le l \le N} x_{k,l} \alpha_{i_{k,l}} = \nu \Big\},$$

where $\boldsymbol{i} := (i_{1,1}, \dots, i_{1,N}, \dots, i_{r,1}, \dots, i_{r,N})$.

The following gives an application to $Z_{\mathcal{I}}$ for general $\mathcal{I}$ which does not necessarily start with $[n]$:

Corollary 3.22. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$, and set $I_0 := [n]$. Fix $\boldsymbol{i}_0 = (i_{k,l})_{0 \le k \le r, 1 \le l \le N_k} \in [n]^{N_0 + \cdots + N_r}$ such that $(i_{k,1}, \dots, i_{k,N_k})$ is a* *reduced word for the longest element in* W_{I_k} *for* $0 \le k \le r$. *Then, the number of connected components of* $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ *equals the cardinality of*

$$\hat{\Delta}_{\boldsymbol{i}_0,0,\lambda_1,\dots,\lambda_r} \cap \mathbb{Z}^{N_1+\cdots+N_r}.$$

Proof. We set $\mathcal{I}_0 := (I_0, I_1, \dots, I_r)$. By the definition of tensor product crystals, the bijective map $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \xrightarrow{\sim} b_0 \otimes \mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, $b \mapsto b_0 \otimes b$, is compatible with their crystal structures, where we mean by $b_0 \in \mathcal{B}(0)$ the element b_λ for $\lambda = 0$. Hence we may identify $b_0 \otimes \mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ with $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$. This implies by the definition that the crystal basis $\mathcal{B}_{\mathcal{I}_0,0,\lambda_1,\dots,\lambda_r}$ is obtained from $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ by actions of $\tilde{f}_i$, $i \in [n]$. By [Lakshmibai et al. 2002, proof of Theorem 2] and Proposition 3.13, all connected components of $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ are Demazure crystals in connected components of $\mathcal{B}(\lambda_1) \otimes \cdots \otimes \mathcal{B}(\lambda_r)$. Hence they are not joined by $\tilde{f}_i$, $i \in [n]$, since they have different highest elements. From these, the crystal basis $\mathcal{B}_{\mathcal{I}_0,0,\lambda_1,\dots,\lambda_r}$ has the same number of connected components as $\mathcal{B}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$, which implies the assertion of the corollary by Corollary 3.16. □

4. Flag Bott–Samelson varieties and flag Bott towers

In this section, we study complex structures on the flag Bott–Samelson variety $Z_{\mathcal{I}}$, and its relation with a flag Bott tower in Theorem 4.10. We first recall flag Bott manifolds introduced in [Kuroki et al. 2020]. Let M be a complex manifold and E a holomorphic vector bundle over M. The *associated flag bundle* $\mathcal{F}\ell(E) \to M$ is a fiber bundle obtained from E by replacing each fiber E_p over a point $p \in M$ by the full flag manifold $\mathcal{F}\ell(E_p)$.

Definition 4.1 [Kuroki et al. 2020, Definition 2.1]. A *flag Bott tower* $\{F_k\}_{0 \le k \le r}$ of height r (or an *r-stage flag Bott tower*) is a sequence,

$$F_r \xrightarrow{p_r} F_{r-1} \xrightarrow{p_{r-1}} \cdots \xrightarrow{p_2} F_1 \xrightarrow{p_1} F_0 = \{\text{a point}\}$$

of manifolds $F_k = \mathcal{F}\ell\big(\bigoplus_{l=1}^{m_k+1} \xi_k^{(l)}\big)$, where $\xi_k^{(l)}$ is a holomorphic line bundle over F_{k-1} for each $1 \le l \le m_k + 1$ and $1 \le k \le r$. We call F_k the *k-stage flag Bott manifold* of the flag Bott tower.

For example, the flag manifold $\mathcal{F}\ell(\mathbb{C}^{m+1}) = \mathcal{F}\ell(m+1)$ is a 1-stage flag Bott manifold, and the product of flag manifolds $\mathcal{F}\ell(m_1+1) \times \cdots \times \mathcal{F}\ell(m_r+1)$ is an r-stage flag Bott manifold. Also an r-stage Bott manifold is an r-stage flag Bott manifold (see [Grossberg and Karshon 1994] for the definition of Bott manifolds). We call two flag Bott towers $\{F_k\}_{0 \le k \le r}$ and $\{F'_k\}_{0 \le k \le r}$ *isomorphic* if there is a collection of diffeomorphisms $\varphi : F_k \to F'_k$ which commutes with the maps $p_k : F_k \to F_{k-1}$ and $p'_k : F'_k \to F'_{k-1}$.

Remark 4.2. In [Kaji et al. 2020], an iterated flag bundle whose fibers are not only full flag manifolds of type A but also other flag manifolds of general Lie type is considered. We recall their construction briefly. For $1 \leq k \leq r$, let K_k be a compact connected Lie group, $T_k \subset K_k$ a maximal torus, and $Z_k \subset K_k$ the centralizer of a circle subgroup of T_k. Recall from [Kaji et al. 2020, Definition 3.1] that an *r-stage flag Bott tower* $\{F_k\}_{0\leq k\leq r}$ *of general Lie type associated to* $\{(K_k, Z_k)\}_{0\leq k\leq r}$ is defined recursively:

(1) F_0 is a point.

(2) F_k is the flag bundle over F_{k-1} with fiber K_k/Z_k associated to a map

$$f_k : F_{k-1} \to BK_k,$$

where f_k factors through BT_k.

Here, the map f_k induces the flag bundle $F_k \to F_{k-1}$ from the universal flag bundle $K_k/Z_k \hookrightarrow BZ_k \to BK_k$.

$$\begin{array}{ccc} K_k/Z_k & = & K_k/Z_k \\ \downarrow & & \downarrow \\ F_k & \longrightarrow & BZ_k \\ \downarrow & & \downarrow \\ F_{k-1} & \xrightarrow{f_k} & BK_k \\ & \searrow \qquad \nearrow & \\ & BT_k & \end{array} \tag{4-1}$$

Because the map f_k factors through BT_k, the bundle F_k is the associated K_k/Z_k-flag bundle of the sum of complex line bundles over F_{k-1}. A flag Bott tower defined in Definition 4.1 is a flag Bott tower of general Lie type associated to $\{(U(m_k+1), T^{m_k+1})\}_{0\leq k\leq r}$.

Lemma 4.3. *Let M be a complex manifold and E a holomorphic vector bundle over M. Let $\mathcal{L}$ be a holomorphic line bundle over M. Then we have that $\mathcal{F}\ell(E) \cong \mathcal{F}\ell(E \otimes \mathcal{L})$ as differentiable manifolds.*

Proof. It is well-known that for a holomorphic vector bundle $E \to M$ over a smooth manifold M and a holomorphic line bundle $\mathcal{L} \to M$, there is a diffeomorphism $\mathbb{P}(E \otimes \mathcal{L}) \cong \mathbb{P}(E)$ (see, for example, [Choi et al. 2010, Lemma 2.1]). Since the induced flag bundle is a sequence of projective bundles as shown in [Bott and Tu 1982, Proposition 21.15], we have a diffeomorphism $\mathcal{F}\ell(E) \cong \mathcal{F}\ell(E \otimes \mathcal{L})$. □

The flag manifold $\mathcal{F}\ell(m+1)$ and an orbit space $\mathrm{GL}(m+1)/B_{\mathrm{GL}(m+1)}$ can be identified. Similarly, an r-stage flag Bott manifold F_r can also be considered as an orbit space. We briefly review the orbit space construction of [Kuroki et al. 2020, §2.2]. Recall from [Kuroki et al. 2020, Lemma 2.12] that for a given Bott tower $\{F_k\}_{0\le k\le r}$ such that $\mathcal{F}\ell(m_k+1)\hookrightarrow F_k \to F_{k-1}$, there is a surjective group homomorphism:

$$\psi : \mathbb{Z}^{m_1+1}\times\cdots\times\mathbb{Z}^{m_k+1} \twoheadrightarrow \mathrm{Pic}(F_k) \quad \text{for } 1\le k\le r. \tag{4-2}$$

We briefly explain the geometric meaning of the homomorphism (4-2). For the flag bundle $\mathcal{F}\ell(E)\xrightarrow{p} M$ obtained by a vector bundle E of rank n over a complex manifold M, consider the universal flag of bundles $0\subset E_1\subset E_2\subset\cdots\subset E_n=p^*E$ on $\mathcal{F}\ell(E)$. Then every element of $\mathrm{Pic}(\mathcal{F}\ell(E))$ can be written as a polynomial in $x_i=c_1(E_i/E_{i-1})$ for $1\le i\le n$ with coefficients in $\mathrm{Pic}(M)$ (see, for example, [Fulton 1998, Example 3.3.5]). Because F_k is an iterated flag bundle, applying this procedure recurrently, we obtain the homomorphism (4-2). Moreover, for $\xi\in\mathrm{Pic}(F_k)$, if we have $\xi=\psi(\boldsymbol{a}_1,\dots,\boldsymbol{a}_k)$, where $\boldsymbol{a}_j$ is an integer vector $(\boldsymbol{a}_j(1),\dots,\boldsymbol{a}_j(m_j+1))\in\mathbb{Z}^{m_j+1}$ for $1\le j\le k$, then

$$c_1(\xi)=\sum_{j=1}^{k}\sum_{l=1}^{m_j+1}\boldsymbol{a}_j(l)x_{j,l}. \tag{4-3}$$

Here, $x_{j,l}$ is the first Chern class of the quotient bundle $E_{j,l}/E_{j,l-1}$ obtained by the universal flag of bundles $0\subset E_{j,1}\subset E_{j,2}\subset\cdots\subset E_{j,m_j+1}$ on F_j.[1]

Suppose that $c_1(\xi_l^{(k)})$ is determined by a set of integer vectors

$$\{\boldsymbol{a}_{k,j}^{(l)}\in\mathbb{Z}^{m_j+1}\}_{1\le l\le m_k+1,\,1\le j<k\le r}.$$

Then

$$\psi(\boldsymbol{a}_{k,1}^{(l)},\boldsymbol{a}_{k,2}^{(l)},\dots,\boldsymbol{a}_{k,k-1}^{(l)})=\xi_k^{(l)}\to F_{k-1}$$

for each $1\le l\le m_k+1$ and $2\le k\le r$. Using this set of integer vectors, we define a right action Φ_k of $B_{\mathrm{GL}(m_1+1)}\times\cdots\times B_{\mathrm{GL}(m_k+1)}$ on $\mathrm{GL}(m_1+1)\times\cdots\times\mathrm{GL}(m_k+1)$ as

$$\begin{aligned}\Phi_k((g_1,\dots,g_k),(b_1,\dots,b_k)) &:= \big(g_1b_1,\ \Lambda_{2,1}(b_1)^{-1}g_2b_2,\ \Lambda_{3,1}(b_1)^{-1}\Lambda_{3,2}(b_2)^{-1}g_3b_3,\dots,\\ &\qquad \Lambda_{k,1}(b_1)^{-1}\Lambda_{k,2}(b_2)^{-1}\cdots\Lambda_{k,k-1}(b_{k-1})^{-1}g_kb_k\big)\end{aligned}$$

[1] The classes $x_{j,l}$ generate the cohomology $H^2(F_j;\mathbb{Z})$ with the relations $x_{j,1}+\cdots+x_{j,m_j+1}=c_1(\xi_1^{(j)})+\cdots+c_1(\xi_{m_j+1}^{(j)})$ for $1\le j\le k$ (see [Fulton 1998, Example 3.3.5] or [Kaji et al. 2020, Corollary 2.4]).

for $1 \le k \le r$. Here $\Lambda_{k,j}$ is a homomorphism $B_{\mathrm{GL}(m_j+1)} \to H_{\mathrm{GL}(m_k+1)}$ which sends $b \in B_{\mathrm{GL}(m_j+1)}$ to

$$\operatorname{diag}\bigl(\Upsilon(b)^{\boldsymbol{a}_{k,j}^{(1)}}, \Upsilon(b)^{\boldsymbol{a}_{k,j}^{(2)}}, \dots, \Upsilon(b)^{\boldsymbol{a}_{k,j}^{(m_k+1)}}\bigr) \in H_{\mathrm{GL}(m_k+1)},$$

where $\Upsilon : B_{\mathrm{GL}(m_j+1)} \to H_{\mathrm{GL}(m_j+1)}$ is the canonical projection in (2-9), and

$$h^{\boldsymbol{a}} := h_1^{\boldsymbol{a}(1)} h_2^{\boldsymbol{a}(2)} \cdots h_{m+1}^{\boldsymbol{a}(m+1)}$$

for $h = \operatorname{diag}(h_1, \dots, h_{m+1}) \in H_{\mathrm{GL}(m+1)}$ and $\boldsymbol{a} = (\boldsymbol{a}(1), \dots, \boldsymbol{a}(m+1)) \in \mathbb{Z}^{m+1}$. Now we can describe the flag Bott manifold F_r as an orbit space as follows:

Proposition 4.4 [Kuroki et al. 2020, Propositions 2.8 and 2.11]. *Let $\{F_k\}_{0\le k\le r}$ be a flag Bott tower. Suppose that $c_1(\xi_l^{(k)})$ is determined by a set of integer vectors $\{\boldsymbol{a}_{k,j}^{(l)} \in \mathbb{Z}^{m_j+1}\}_{1\le l\le m_k+1, 1\le j<k\le r}$ and let Φ_k be the action determined by these integer vectors. Then the flag Bott tower $\{F_k\}_{0\le k\le r}$ is isomorphic to*

$$\bigl\{\bigl(\mathrm{GL}(m_1+1) \times \cdots \times \mathrm{GL}(m_k+1)\bigr)/\Phi_k\bigr\}_{0\le k\le r}$$

as flag Bott towers.

A Bott–Samelson variety has a family of complex structures which gives a toric degeneration (see [Grossberg and Karshon 1994, §3.4; Pasquier 2010]). Now we study a family of complex structures on a given flag Bott–Samelson variety. Since the simple roots are linearly independent elements in $\mathfrak{h}^*$, there exist $q \in \mathbb{Z}_{>0}$ and an injective homomorphism $\lambda : \mathbb{C}^* \to H$ such that

$$e^{\alpha}(\lambda(t)) = t^q \tag{4-4}$$

for all simple roots α and $t \in \mathbb{C}^*$. Here $e^{\alpha} : H \to \mathbb{C}^*$ is a character induced from $\alpha : \mathfrak{h} \to \mathbb{C}$. For example, when $G = \mathrm{SL}(2k+1)$ and $q = 1$, consider the homomorphism $\lambda : \mathbb{C}^* \to H$ defined by

$$\lambda : t \mapsto \operatorname{diag}\bigl(t^k, t^{k-1}, \dots, t, 1, t^{-1}, \dots, t^{-k+1}, t^{-k}\bigr). \tag{4-5}$$

Then this homomorphism satisfies the condition on (4-4). We define $\Upsilon_t : B \to B$ by

$$\Upsilon_t : b \mapsto \lambda(t) b (\lambda(t))^{-1}$$

for $t \in \mathbb{C}^*$. It is proved in [Grossberg and Karshon 1994, Proposition 3.5] that $\Upsilon = \lim_{t\to 0} \Upsilon_t$, where $\Upsilon : B \to H$ is the homomorphism in (2-9). We put $\Upsilon_0 := \Upsilon$.

Example 4.5. Suppose that $G = \mathrm{SL}(3)$ and $q = 1$. Considering the homomorphism $\lambda : \mathbb{C}^* \to H$ defined in (4-5), the homomorphism $\Upsilon_t : B \to B$ is given by

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{bmatrix} \mapsto \begin{bmatrix} b_{11} & t b_{12} & t^2 b_{13} \\ 0 & b_{22} & t b_{23} \\ 0 & 0 & b_{33} \end{bmatrix}.$$

Hence we have that $\lim_{t\to 0} \Upsilon_t = \Upsilon$.

We use the homomorphism $\Upsilon_t : B \to B$ to construct a family of complex structures on the flag Bott–Samelson manifold $Z_{\mathcal{I}} = \boldsymbol{P}_{\mathcal{I}}/B^r$. For $t \in \mathbb{C}$, we define a right action Θ_t of B^r on $\boldsymbol{P}_{\mathcal{I}}$ as

$$\Theta_t((p_1,\dots,p_r),(b_1,\dots,b_r)) \\ = \big(p_1b_1, \Upsilon_t(b_1)^{-1}p_2b_2,\dots,\Upsilon_t(b_{r-1})^{-1}p_rb_r\big) \tag{4-6}$$

for $(p_1,\dots,p_r) \in \boldsymbol{P}_{\mathcal{I}}$ and $(b_1,\dots,b_r) \in B^r$. Then Θ_1 coincides with the right action in (2-2) because $\lambda(1) = e \in H$ and hence $\Upsilon_1 = \mathrm{Id}_B$. Again we consider the family of orbit spaces

$$Z^t_{\mathcal{I}} := \boldsymbol{P}_{\mathcal{I}}/\Theta_t$$

for $t \in \mathbb{C}$. The holomorphic line bundle $\mathcal{L}^t_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ over $Z^t_{\mathcal{I}}$ can be defined in a way similar to $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ in (2-10) for integral weights $\lambda_1,\dots,\lambda_r$. Set $\mathcal{L}^t_{\mathcal{I},\lambda} := \mathcal{L}^t_{\mathcal{I},0,\dots,0,\lambda}$ for simplicity.

Proposition 4.6. *For a given sequence $\mathcal{I} = (I_1,\dots,I_r)$, the manifolds $Z^t_{\mathcal{I}}$ are all diffeomorphic for $t \in \mathbb{C}$.*

Proof. We use the similar argument to the proof of Proposition 3.7 in [Grossberg and Karshon 1994]. Let K_{I_j} be the maximal compact subgroup of P_{I_j}. Let T be the maximal compact torus in G, i.e., $T = (S^1)^n$. Recall that $K_{I_j} \cap B = T$. Define a right action of $T^{(r)} := T \times T \times \cdots \times T$ (r factors) on $K_{\mathcal{I}} := K_{I_1} \times \cdots \times K_{I_r}$ as

$$(g_1,\dots,g_r)\cdot(a_1,\dots,a_r) = (g_1a_1, a_1^{-1}g_2a_2,\dots,a_{r-1}^{-1}g_ra_r). \tag{4-7}$$

Let $X_{\mathcal{I}}$ be the orbit space

$$X_{\mathcal{I}} := (K_{I_1} \times \cdots \times K_{I_r})/(T \times \cdots \times T). \tag{4-8}$$

The inclusion map

$$K_{\mathcal{I}} = K_{I_1} \times \cdots \times K_{I_r} \hookrightarrow \boldsymbol{P}_{\mathcal{I}} = P_{I_1} \times \cdots \times P_{I_r}$$

is $T^{(r)}$-equivariant with respect to the $T^{(r)}$-action of (4-7) on $K_{\mathcal{I}}$ and the restricted $T^{(r)}$-action of (4-6) on $\boldsymbol{P}_{\mathcal{I}}$ via the inclusion $T^{(r)} \hookrightarrow B^r$ because $\Upsilon_t(a) = a$ for all $a \in T$. Therefore we get a map

$$f^t_{\mathcal{I}} : X_{\mathcal{I}} \to Z^t_{\mathcal{I}}. \tag{4-9}$$

Since, for all k, the inclusion $K_{I_k} \hookrightarrow P_{I_k}$ induces a diffeomorphism $K_{I_k}/T \cong P_{I_k}/B$, the map $f^t_{\mathcal{I}}$ is a diffeomorphism. □

The manifold $Z^t_{\mathcal{I}}$ has a fibration structure, similar to a flag Bott–Samelson manifold in (2-3):

$$P_{I_r}/B \hookrightarrow Z^t_{\mathcal{I}} \xrightarrow{\pi} Z^t_{\mathcal{I}'}, \tag{4-10}$$

where $\mathcal{I}' = (I_1, \dots, I_{r-1})$ is the subsequence of $\mathcal{I}$ and π is the first $r-1$ coordinates projection for all $t \in \mathbb{C}$.

Let $\mathcal{I} = (I_1, \dots, I_{r-1}, I_r)$ and $\mathcal{I}' = (I_1, \dots, I_{r-1})$. We note that the orbit space $X_{\mathcal{I}}$ has a bundle structure.

$$\begin{array}{ccc} X_{\mathcal{I}} = \boldsymbol{P}_{\mathcal{I}'} \times_T (K_{I_r}/T) & \hookleftarrow & K_{I_r}/T \\ \big\downarrow & & \\ X_{\mathcal{I}'} & & \end{array}$$

Because the structure group T of this bundle is an abelian group, the map f_k inducing the flag bundle $X_{\mathcal{I}} \to X_{\mathcal{I}'}$ from the universal flag bundle factors through BT.

$$\begin{array}{ccccc} K_{I_r}/T & & = & & K_{I_r}/T \\ \big\downarrow & & & & \big\downarrow \\ X_{\mathcal{I}} & & \longrightarrow & & BT \\ \big\downarrow & & & & \big\downarrow \\ X_{\mathcal{I}'} & & \xrightarrow{f_k} & & BK_{I_r} \\ & \searrow & & \nearrow & \\ & & BT & & \end{array}$$

Continuing this procedure, we obtain the following corollary.

Corollary 4.7. *The manifold $X_{\mathcal{I}}$ is an r-stage flag Bott tower of general Lie type associated to $\{(K_{I_j}, T)\}_{0 \le j \le r}$, and so are $Z_{\mathcal{I}}^t$ for all $t \in \mathbb{C}$* (see Remark 4.2 for the definition of flag Bott towers of general Lie type).

For the remaining part of this section, we consider the case when the Levi subgroup L_{I_k} of the parabolic subgroup P_{I_k} has Lie type A, that is, the flag Bott tower $X_{\mathcal{I}}$ is a flag Bott manifold whose fibers are all full flag manifolds of Lie type A. Moreover, we describe the line bundles appearing in the construction explicitly (see Theorem 4.10). We can always take an enumeration $I_k = \{u_{k,1}, \dots, u_{k,m_k}\}$ so that

$$\langle \alpha_{u_{k,s}}, \alpha^\vee_{u_{k,t}} \rangle = \begin{cases} 2 & \text{if } s = t, \\ -1 & \text{if } s - t = \pm 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4-11}$$

Proposition 4.8. *Let $Z_{\mathcal{I}}$ be a flag Bott–Samelson manifold. Let $\mathcal{I}' = (I_1, \dots, I_{r-1})$ be the subsequence of $\mathcal{I}$. Assume that the Levi subgroup L_{I_k} of the parabolic subgroup P_{I_k} has Lie type A_{m_k} for all $1 \le k \le r$. Then the manifold $Z_{\mathcal{I}}^0$ is diffeomorphic to the induced flag bundle over $Z_{\mathcal{I}'}^0$:*

$$Z_{\mathcal{I}}^0 \cong \mathcal{F}\ell\big(\mathcal{L}^0_{\mathcal{I}', \chi_1} \oplus \dots \oplus \mathcal{L}^0_{\mathcal{I}', \chi_{m_r}} \oplus \underline{\mathbb{C}}\big),$$

where $\chi_j = \alpha_{u_{r,j}} + \cdots + \alpha_{u_{r,m_r}} \in \mathfrak{h}^*$ *for* $1 \le j \le m_r$, $\mathcal{L}^0_{\mathcal{I}',\chi} = \mathcal{L}^0_{\mathcal{I}',0,\ldots,0,\chi}$, *and* $\underline{\mathbb{C}}$ *is the trivial line bundle.*

Before proving the proposition, we observe the following. Suppose that the Levi subgroup L_I of the parabolic subgroup P_I for a subset $I \subset [n]$ has Lie type A_m. Then we can label the elements of I as $u_1, \ldots, u_m$ which satisfy the relation (4-11). Also we have the group homomorphism $F : \mathrm{SL}(m+1) \to L_I \hookrightarrow P_I$. Then the map F induces the homomorphism $F_* : \mathfrak{h}_{\mathrm{SL}(m+1)} \to \mathfrak{h}$. We label the coroots of $\mathrm{SL}(m+1)$ as $\beta_1^\vee, \beta_2^\vee, \ldots, \beta_m^\vee$ so that F_* sends $\beta_l^\vee$ to $\alpha_{u_l}^\vee$ for $1 \le l \le m$. Then we have that

$$\langle F^*\lambda, \beta_l^\vee \rangle = \langle \lambda, F_*\beta_l^\vee \rangle = \langle \lambda, \alpha_{u_l}^\vee \rangle$$

for a weight $\lambda \in \mathfrak{h}^*$ and $1 \le l \le m$. Here, we note that $F^*\lambda = \lambda \circ F$ for $\lambda \in \mathfrak{h}^*$. Let $\varpi_1, \varpi_2, \ldots, \varpi_m \in \mathfrak{h}^*_{\mathrm{SL}(m+1)}$ be the fundamental weights. Then the pullback $F^*\lambda$ is given by

$$F^*\lambda = \sum_{l=1}^{m} \langle \lambda, \alpha_{u_l}^\vee \rangle \varpi_l \in \mathfrak{h}^*_{\mathrm{SL}(m+1)}. \tag{4-12}$$

Proof of Proposition 4.8. We write $I = I_r$, $m = m_r$, and $u_j = u_{r,j}$ for $1 \le j \le m$. Note that we have $P_I = L_I U_I$ (see Section 2A). Since we have an isomorphism of varieties

$$P_I/B = (L_I U_I)/B = L_I/(B \cap L_I) = L_I/B_I,$$

we get a diffeomorphism

$$F_1 : \mathrm{SL}(m+1)/B_{\mathrm{SL}(m+1)} \to P_I/B.$$

Moreover, the map which sends an element g in $\mathrm{SL}(m+1)$ to a full flag $(V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_m)$, where $V_l = \langle c_1, \ldots, c_l \rangle$ and c_l is the l-th column vector of g, descends to a diffeomorphism

$$F_2 : \mathrm{SL}(m+1)/B_{\mathrm{SL}(m+1)} \to \mathcal{F}\ell(m+1).$$

The map F_2 is equivariant with respect to the following actions of the torus $H_{\mathrm{SL}(m+1)}$: each element

$$h = \mathrm{diag}(h_1, h_2, \ldots, h_{m+1}) \in H_{\mathrm{SL}(m+1)}$$

acts on $\mathrm{SL}(m+1)/B_{\mathrm{SL}(m+1)}$ by the left multiplication, and on $\mathcal{F}\ell(m+1)$ as the induced action from the representation space $\mathbb{C}^{m+1}$ with weights

$$\big(\varpi_1, -\varpi_1 + \varpi_2, \ldots, -\varpi_{m-1} + \varpi_m, -\varpi_m\big), \tag{4-13}$$

namely $h \cdot v = (h_1 v_1, h_2 v_2, \ldots, h_{m+1} v_{m+1})$ for $v = (v_1, \ldots, v_{m+1}) \in \mathbb{C}^{m+1}$. On the other hand, the map F_1 is equivariant with respect to the left multiplication actions of $H_{\mathrm{SL}(m+1)}$ and of H via the homomorphism $H_{\mathrm{SL}(m+1)} \to H$ given by the map F.

By the relation (4-12) between weights in $\mathfrak{h}^*$ and $\mathfrak{h}^*_{\mathrm{SL}(m+1)}$, we have the following:

$$\begin{aligned} F^*(\chi_j) &= F^*(\alpha_{u_j} + \cdots + \alpha_{u_m}) \\ &= \sum_{l=1}^{m} \langle \alpha_{u_j} + \cdots + \alpha_{u_m}, \alpha_{u_l}^\vee \rangle \varpi_l \\ &= -\varpi_{j-1} + \varpi_j + \varpi_m, \end{aligned}$$

where $\varpi_0 = 0$ for $1 \le j \le m$. Here the third equality follows by considering the Cartan matrix of $\mathrm{SL}(m+1)$. The $H_{\mathrm{SL}(m+1)}$-representation on $\mathbb{C}^{m+1}$ with weights (4-13) becomes an H-representation on $\mathbb{C}^{m+1}$ with weights

$$\big(\chi_1 - \chi', \chi_2 - \chi', \ldots, \chi_m - \chi', -\chi'\big),$$

where χ' is a weight which maps to ϖ_m under the map F^* such that $F_2 \circ F_1^{-1}$ is equivariant with respect to the actions of elements in $H \setminus F(H_{\mathrm{SL}(m+1)})$. This proves that $F_2 \circ F_1^{-1}$ is a left H-equivariant diffeomorphism

$$F_2 \circ F_1^{-1} : P_I/B \to \mathcal{F}\ell\big(\mathbb{C}_{\chi_1-\chi'} \oplus \cdots \oplus \mathbb{C}_{\chi_m-\chi'} \oplus \mathbb{C}_{-\chi'}\big).$$

We notice that the construction of twisted product is functorial, i.e., for a topological group G and a right G-space X, if $f : Y \to Y'$ is an equivariant map of left G-spaces then we have the induced map $X \times_G Y \to X \times_G Y'$, see, for example, [Bredon 1972, §II.2]. Since the unipotent part of B acts trivially on P_I/B and $\mathcal{F}\ell(\mathbb{C}_{\chi_1-\chi'} \oplus \cdots \oplus \mathbb{C}_{\chi_m-\chi'} \oplus \mathbb{C}_{-\chi'})$, the left H-equivariant diffeomorphism $F_2 \circ F_1^{-1}$ induces a diffeomorphism

$$\boldsymbol{P}_{\mathcal{I}}/\Theta_0 \cong \mathcal{F}\ell\big(\mathcal{L}^0_{\mathcal{I}',\chi_1-\chi'} \oplus \cdots \oplus \mathcal{L}^0_{\mathcal{I}',\chi_m-\chi'} \oplus \mathcal{L}^0_{\mathcal{I}',-\chi'}\big).$$

Moreover we have that

$$\begin{aligned} &\mathcal{F}\ell\big(\mathcal{L}^0_{\mathcal{I}',\chi_1-\chi'} \oplus \cdots \oplus \mathcal{L}^0_{\mathcal{I}',\chi_m-\chi'} \oplus \mathcal{L}^0_{\mathcal{I}',-\chi'}\big) \\ &\qquad\qquad = \mathcal{F}\ell\big((\mathcal{L}^0_{\mathcal{I}',\chi_1} \oplus \cdots \oplus \mathcal{L}^0_{\mathcal{I}',\chi_m} \oplus \underline{\mathbb{C}}) \otimes \mathcal{L}^0_{\mathcal{I}',-\chi'}\big). \end{aligned}$$

Then by Lemma 4.3, we are done. □

By Proposition 4.8, we can conclude that $Z^0_{\mathcal{I}}$ is an r-stage flag Bott manifold. For given integral weights $\lambda_1, \ldots, \lambda_r$, consider the line bundle $\mathcal{L}^0_{\mathcal{I},\lambda_1,\ldots,\lambda_r}$ over a flag Bott manifold $Z^0_{\mathcal{I}}$. By (4-2) there is a set of integer vectors $\{\boldsymbol{a}_k \in \mathbb{Z}^{m_k+1}\}_{1 \le k \le r}$ determined by $c_1(\mathcal{L}^0_{\mathcal{I},\lambda_1,\ldots,\lambda_r})$. Indeed, we have

$$\psi(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r) \cong \mathcal{L}^0_{\mathcal{I},\lambda_1,\ldots,\lambda_r}.$$

The following proposition computes these integer vectors in terms of integral weights $\lambda_1, \dots, \lambda_r$ and a sequence $\mathcal{I}$ of subsets of $[n]$.

Proposition 4.9. *Let $\mathcal{I} = (I_1, \dots, I_r)$ be a sequence of subsets of $[n]$. Assume that the Levi subgroup L_{I_k} of the parabolic subgroup P_{I_k} has Lie type A_{m_k} for all $1 \le k \le r$. For given integral weights $\lambda_1, \dots, \lambda_r \in \mathbb{Z}\varpi_1 + \cdots + \mathbb{Z}\varpi_n$, the first Chern class of the line bundle $\mathcal{L} = \mathcal{L}^0_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ is given by integer vectors $\boldsymbol{a}_k = (\boldsymbol{a}_k(1), \dots, \boldsymbol{a}_k(m_k+1)) \in \mathbb{Z}^{m_k+1}$ for $1 \le k \le r$, where*

$$\boldsymbol{a}_k(l) = \big\langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,l}} + \cdots + \alpha^\vee_{u_{k,m_k}} \big\rangle \quad \text{for } 1 \le l \le m_k,$$
$$\boldsymbol{a}_k(m_k+1) = 0.$$

Here, we take an enumeration $I_k = \{u_{k,1}, \dots, u_{k,m_k}\}$ which satisfies (4-11). Indeed, $\mathcal{L}$ is isomorphic to the line bundle $\psi(\boldsymbol{a}_1, \dots, \boldsymbol{a}_r)$.

Proof. Since the Levi subgroup L_{I_k} of P_{I_k} is Lie type A_{m_k}, we have a Lie group homomorphism $F_k : \mathrm{SL}(m_k+1) \to P_{I_k}$. For each $1 \le k \le r$, consider the homomorphism $\psi_k : \mathrm{SL}(m_k+1) \to P_{I_1} \times \cdots \times P_{I_r}$ defined as

$$p \mapsto (e, \dots, e, \underbrace{F_k(p)}_{k\text{-th}}, e, \dots, e)$$

and consider

$$\varphi_k : B_{\mathrm{SL}(m_k+1)} \to \underbrace{B \times \cdots \times B}_{r} = B^r \tag{4-14}$$

which sends b to

$$\big(e, \dots, e, \underbrace{F_k(b)}_{k\text{-th}}, \underbrace{F_k(h)}_{(k+1)\text{-th}}, \dots, \underbrace{F_k(h)}_{r\text{-th}}\big),$$

where $h = \Upsilon(b)$. Then the map ψ_k is φ_k-equivariant, namely, for $b \in B_{\mathrm{SL}(m_k+1)}$ and $g \in \mathrm{SL}(m_k+1)$ we have that

$$\begin{aligned}
&\Theta_0(\psi_k(g), \varphi_k(b)) \\
&\qquad = \Theta_0\big((e, \dots, e, F_k(g), e, \dots, e), (e, \dots, e, F_k(b), F_k(h), \dots, F_k(h))\big) \\
&\qquad = \big(e, \dots, e, F_k(g)F_k(b), \Upsilon(F_k(b))^{-1}F_k(h), e, \dots, e\big) \\
&\qquad = (e, \dots, e, F_k(gb), e, e, \dots, e) \\
&\qquad = \psi_k(gb).
\end{aligned}$$

Here the third equality comes from the fact that F_k is a homomorphism and $\Upsilon(F_k(b)) = F_k(\Upsilon(b))$.

Under the map (4-14) the weight $(\lambda_1, \dots, \lambda_r)$ of H^r pulls back to the weight

$$\sum_{l=1}^{m_k} \langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,l}} \rangle \varpi_l \in \mathfrak{h}^*_{\mathrm{SL}(m_k+1)} \tag{4-15}$$

by (4-12). The integer vector $\boldsymbol{a}_k \in \mathbb{Z}^{m_k+1}$ is completely determined by the weight in (4-15) because of the construction of a flag Bott manifold (see [Kuroki et al. 2020, §2.2]). Indeed, the integer vector $\boldsymbol{a}_k \in \mathbb{Z}^{m_k+1}$ should satisfy the equality

$$\sum_{l=1}^{m_k} \langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,l}} \rangle \varpi_l = \sum_{l=1}^{m_k+1} \boldsymbol{a}_k(l) \varepsilon_l, \tag{4-16}$$

where $\varepsilon_i \in \mathfrak{h}^*_{\mathrm{SL}(m_k+1)}$ sends $\mathrm{diag}(h_1, \dots, h_{m_k+1})$ in $\mathfrak{h}_{\mathrm{SL}(m_k+1)}$ to h_i. Using the identification $\varpi_l = \varepsilon_1 + \cdots + \varepsilon_l$, we have that

$$\begin{aligned}
\sum_{l=1}^{m_k} &\langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,l}} \rangle \varpi_l \\
&= \langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,1}} \rangle \varepsilon_1 + \langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,2}} \rangle (\varepsilon_1 + \varepsilon_2) \\
&\qquad + \cdots + \langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,m_k}} \rangle (\varepsilon_1 + \cdots + \varepsilon_{m_k}) \\
&= \big\langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,1}} + \alpha^\vee_{u_{k,2}} + \cdots + \alpha^\vee_{u_{k,m_k}} \big\rangle \varepsilon_1 \\
&\qquad + \big\langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,2}} + \cdots + \alpha^\vee_{u_{k,m_k}} \big\rangle \varepsilon_2 \\
&\qquad + \cdots + \langle \lambda_k + \cdots + \lambda_r, \alpha^\vee_{u_{k,m_k}} \rangle \varepsilon_{m_k}.
\end{aligned} \tag{4-17}$$

Comparing (4-16) and (4-17), we obtain the assertion of the proposition. □

By combining Propositions 4.8 and 4.9, we can prove the following theorem:

Theorem 4.10. *Suppose that $\mathcal{I} = (I_1, \dots, I_r)$ is a sequence of subsets of $[n]$ such that the Levi subgroup L_{I_k} of the parabolic subgroup P_{I_k} has Lie type A_{m_k} for all $1 \le k \le r$. Take an enumeration $I_k = \{u_{k,1}, \dots, u_{k,m_k}\}$ which satisfies (4-11). Then the manifold $Z^0_{\mathcal{I}}$ is an r-stage flag Bott manifold which is determined by*

$$\big\{\boldsymbol{a}^{(l)}_{k,j} = \big(\boldsymbol{a}^{(l)}_{k,j}(1), \boldsymbol{a}^{(l)}_{k,j}(2), \dots, \boldsymbol{a}^{(l)}_{k,j}(m_j+1)\big)\big\}_{1 \le l \le m_k+1,\, 1 \le j < k \le r}$$

in the sense of Proposition 4.4, where $\boldsymbol{a}^{(l)}_{k,j}(p)$ is

$$\big\langle \alpha_{u_{k,l}} + \cdots + \alpha_{u_{k,m_k}}, \alpha^\vee_{u_{j,p}} + \cdots + \alpha^\vee_{u_{j,m_j}} \big\rangle$$

if $1 \le l \le m_k$ and $1 \le p \le m_j$, and 0 otherwise.

Proof. Consider the subsequence $\mathcal{I}_k := (I_1, \dots, I_k)$ of the sequence $\mathcal{I}$ for all $1 \le k \le r$. Recall from Proposition 4.8 that the flag Bott manifold $Z^0_{\mathcal{I}_k}$ is the induced flag bundle over $Z^0_{\mathcal{I}_{k-1}}$:

$$Z^0_{\mathcal{I}_k} = \mathcal{F}\ell\big(\mathcal{L}^0_{\mathcal{I}_{k-1},\chi_1} \oplus \cdots \oplus \mathcal{L}^0_{\mathcal{I}_{k-1},\chi_{m_k}} \oplus \underline{\mathbb{C}}\big),$$

where $\chi_l = \alpha_{u_{k,l}} + \cdots + \alpha_{u_{k,m_k}}$ for $1 \le l \le m_k$. By Proposition 4.9, the integer vectors $\{\boldsymbol{a}_{k,j}^{(l)} \in \mathbb{Z}^{m_j+1}\}_{1\le j\le k-1}$ which define the line bundle $\mathcal{L}^0_{\mathcal{I}_{k-1},\chi_l}$ are given by

$$\begin{aligned}\boldsymbol{a}_{k,j}^{(l)}(p) &= \langle \chi_l, \alpha^\vee_{u_{j,p}} + \cdots + \alpha^\vee_{u_{j,m_j}}\rangle && \text{(by Proposition 4.9)}\\ &= \big\langle \alpha_{u_{k,l}} + \cdots + \alpha_{u_{k,m_k}}, \alpha^\vee_{u_{j,p}} + \cdots + \alpha^\vee_{u_{j,m_j}}\big\rangle && \text{(by the definition of } \chi_l)\end{aligned}$$

for $1 \le l \le m_k$ and $1 \le p \le m_j$. Moreover we have $\boldsymbol{a}_{k,j}^{(l)}(p) = 0$ if $l = m_k + 1$ or $p = m_j + 1$ by Proposition 4.9. Hence the result follows. □

Example 4.11. Let $G = \mathrm{SL}(4)$. Consider the sequence $\mathcal{I} = (\{1, 2\}, \{1, 2\})$. Hence $u_{1,1} = 1,\ u_{1,2} = 2,\ u_{2,1} = 1,\ u_{2,2} = 2$. The manifold $Z^0_{\mathcal{I}}$ is a 2-stage flag Bott manifold with $F_2 = \mathcal{F}\ell(\xi_2^{(1)} \oplus \xi_2^{(2)} \oplus \mathbb{C})$, where line bundles $\xi_2^{(1)}$ and $\xi_2^{(2)}$ are determined by the following integer vectors:

$$\begin{aligned}\boldsymbol{a}_{2,1}^{(1)} &= \big(\langle \alpha_1 + \alpha_2, \alpha_1^\vee + \alpha_2^\vee\rangle, \langle \alpha_1 + \alpha_2, \alpha_2^\vee\rangle, 0\big) = (2, 1, 0),\\ \boldsymbol{a}_{2,1}^{(2)} &= \big(\langle \alpha_2, \alpha_1^\vee + \alpha_2^\vee\rangle, \langle \alpha_2, \alpha_2^\vee\rangle, 0\big) = (1, 2, 0).\end{aligned}$$

Remark 4.12. Suppose that the flag Bott–Samelson variety $Z_{\mathcal{I}}$ is a Bott–Samelson variety, i.e., $m_1 = \cdots = m_r = 1$. Then integer vectors $\{\boldsymbol{a}_{k,j}^{(l)} \in \mathbb{Z}^2\}_{l\in[2],1\le j<k\le r}$ determining the flag Bott tower $Z^0_{\mathcal{I}}$ is

$$\boldsymbol{a}_{k,j}^{(l)} = \begin{cases} (\langle \alpha_{u_{k,1}}, \alpha^\vee_{u_{j,1}}\rangle, 0) & \text{if } l = 1,\\ (0, 0) & \text{if } l = 2\end{cases}$$

by Theorem 4.10. This computation of $\boldsymbol{a}_{k,j}^{(1)}(1)$ for $1 \le j < k \le r$ coincides with the known result in [Grossberg and Karshon 1994, §3.7].

5. Torus actions and Duistermaat–Heckman measure

Let $\mathcal{I} = (I_1, \ldots, I_r)$ be a sequence of subsets of $[n]$ such that $|I_k| = m_k$. In this section we study torus actions on the manifold $Z^0_{\mathcal{I}}$. We define a torus invariant closed 2-form induced from a given complex line bundle, and we consider the Duistermaat–Heckman measure of the flag Bott–Samelson manifold using a Bott–Samelson variety $Z_{\boldsymbol{i}}$ admitting the birational morphism $\eta_{\boldsymbol{i},\mathcal{I}} : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$ (see Theorem 5.5).

We first study torus actions on $Z^0_{\mathcal{I}}$. Let T be the maximal compact torus of G contained in H. Define an action of $T^{(r)}$ on $Z^0_{\mathcal{I}}$ as

$$\begin{aligned}(s_1, \ldots, s_r) \cdot [p_1, \ldots, p_r] &= [s_1 p_1, s_1^{-1} s_2 p_2, \ldots, s_{r-1}^{-1} s_r p_r]\\ &= [s_1 p_1 s_1^{-1}, \ldots, s_r p_r s_r^{-1}].\end{aligned} \tag{5-1}$$

This action is smooth but not effective. We now find the subtorus which acts trivially on $Z^0_{\mathcal{I}}$. Define a subtorus $T_I \subset T$ for a subset $I \subset [n]$ as

$$T_I := \big\{ s \in T \mid \alpha_i(s) = 1 \text{ for all } i \in I \big\}^0$$

which is similar to (2-1). Here, we consider a simple root $\alpha \in \chi(H)$ as a homomorphism $T \to S^1$. For a given sequence $\mathcal{I} = (I_1, \dots, I_r)$ of subsets of $[n]$, we define the subtorus $T_{\mathcal{I}}$ of $T^{(r)}$ as

$$T_{\mathcal{I}} := T_{I_1} \times \cdots \times T_{I_r}.$$

Similarly, we set $T_{\boldsymbol{i}} := T_{\{i_1\}} \times \cdots \times T_{\{i_r\}}$ for a sequence $\boldsymbol{i} = (i_1, \dots, i_r) \in [n]^r$. Then the following proposition comes from (5-1).

Proposition 5.1. *The torus $T_{\mathcal{I}}$ acts trivially on $Z^0_{\mathcal{I}}$.*

By Proposition 5.1, we have the torus action on $Z^0_{\mathcal{I}}$:

$$T^{(r)}/T_{\mathcal{I}} \curvearrowright Z^0_{\mathcal{I}}. \tag{5-2}$$

Note that $T^{(r)}/T_{\mathcal{I}} \cong (S^1)^{m_1+\cdots+m_r}$.

Suppose that

$$\boldsymbol{i} = (i_{k,l})_{1\le k\le r,\, 1\le l\le N_k} \in [n]^{N_1+\cdots+N_r}$$

is a sequence such that $(i_{k,1}, \dots, i_{k,N_k})$ is a reduced word for the longest element in W_{I_k} for $1 \le k \le r$. From now on, we ignore the complex structure on the flag Bott–Samelson manifold $Z_{\mathcal{I}}$ and regard it as a smooth manifold. Therefore we can identify $Z_{\mathcal{I}}$ with $Z^0_{\mathcal{I}}$ and $Z_{\boldsymbol{i}}$ with $Z^0_{\boldsymbol{i}}$ by Proposition 4.6. Using the observation (5-2), we have the torus action on the Bott–Samelson manifold $Z_{\boldsymbol{i}}$:

$$(S^1)^N \cong T^{(N)}/T_{\boldsymbol{i}} \curvearrowright Z_{\boldsymbol{i}},$$

where $N := N_1 + N_2 + \cdots + N_r$. We denote $\widetilde{\boldsymbol{T}} := (T^{(N)})/T_{\boldsymbol{i}}$ and $\boldsymbol{T} := (T^{(r)})/T_{\mathcal{I}}$ for simplicity.

Lemma 5.2. *There is a homomorphism $A : \boldsymbol{T} \to \widetilde{\boldsymbol{T}}$ such that the map $\eta_{\boldsymbol{i},\mathcal{I}} : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$ is equivariant with respect to the action of $\boldsymbol{T}$, i.e.,*

$$\eta_{\boldsymbol{i},\mathcal{I}}(A(t)\cdot x) = t \cdot \eta_{\boldsymbol{i},\mathcal{I}}(x)$$

for any $t \in \boldsymbol{T}$ and $x \in Z_{\boldsymbol{i}}$.

Proof. Define an inclusion map $\iota : T^{(r)} \hookrightarrow T^{(N)}$ as

$$(a_1, \dots, a_r) \overset{\iota}{\mapsto} (\underbrace{a_1, \dots, a_1}_{N_1}, \dots, \underbrace{a_k, \dots, a_k}_{N_k}, \dots, \underbrace{a_r, \dots, a_r}_{N_r}).$$

Then we have the action $T^{(r)} \curvearrowright Z_{\boldsymbol{i}}$ via the inclusion ι and the map $\eta_{\boldsymbol{i},\mathcal{I}} : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$ is equivariant with respect to the action of $T^{(r)}$ by the definition of torus action in (5-1).

We claim that $\iota(T_{\mathcal{I}}) \subset T_{\boldsymbol{i}}$. For an element $(a_1, \dots, a_r) \in T^{(r)}$, we have that

$$(a_1, \dots, a_r) \in T_{\mathcal{I}} \Leftrightarrow a_k \in T_{I_k} \quad \text{for all } 1 \le k \le r.$$

Hence we have $a_k \in T_{i_{k,1}}, \dots, a_k \in T_{i_{k,N_k}}$ since $\{i_{k,1}, \dots, i_{k,N_k}\} = I_k$ for all $1 \le k \le r$. This gives that $\iota(T_{\mathcal{I}}) \subset T_{\boldsymbol{i}}$ as claimed. We thus have the homomorphism

$$A : T^{(r)}/T_{\mathcal{I}} \to T^{(N)}/T_{\boldsymbol{i}} \tag{5-3}$$

induced from the inclusion ι. Moreover the projection map $Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$ is equivariant with respect to the action of $\boldsymbol{T}$ because of the $T^{(r)}$-equivariance of the projection. □

We set $A_k : T/T_{I_k} \to T^{(N_k)}/T_{(i_{k,1},\dots,i_{k,N_k})}$ for $1 \le k \le r$. By the definition of $T_{(i_{k,1},\dots,i_{k,N_k})}$, the torus $T^{(N_k)}/T_{(i_{k,1},\dots,i_{k,N_k})}$ has dimension N_k. Suppose that $\{f_{k,1}, \dots, f_{k,N_k}\}$ is the standard basis of $\mathrm{Lie}((S^1)^{N_k})^* \cong \mathbb{R}^{N_k}$. Then it is known from [Grossberg and Karshon 1994, §3.7] that the pullback of $f_{k,l}$ is $\alpha_{i_{k,l}}$ for $1 \le l \le N_k$. Since the homomorphism A can be identified with $A_1 \times \cdots \times A_r$, the Lie algebra homomorphism $(dA)^* : \mathbb{R}^N \to \mathbb{R}^{m_1+\cdots+m_r}$ maps $f_{k,l}$ to $\alpha_{i_{k,l}}$ for $1 \le k \le r$ and $1 \le l \le N_k$.

Example 5.3. Recall from Example 2.8 that we have a morphism $\eta_{(1,2,1,3),\mathcal{I}}$ from $Z_{(1,2,1,3)}$ to $Z_{\mathcal{I}}$, where $\mathcal{I} = (\{1,2\},\{3\})$. Suppose that $A : T^{(2)}/T_{\mathcal{I}} \to T^{(4)}/T_{(1,2,1,3)}$ is the homomorphism in Lemma 5.2. Then the Lie algebra homomorphism $(dA)^*$: $\mathbb{R}^4 \to \mathbb{R}^3$ is defined using the integer matrix:

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We now consider Duistermaat–Heckman measures corresponding to flag Bott–Samelson manifolds. We recall definitions from [Audin 2004]. Suppose that M is an oriented, compact manifold of real dimension $2d$ with an action of a compact torus T. Let ω be a presymplectic form, i.e., a T-invariant closed not necessarily nondegenerate 2-form. Then we call the manifold (M, ω, T) *presymplectic T-manifold*. A *moment map* on (M, ω, T) is defined to be a map $\Phi : M \to \mathrm{Lie}(T)^*$ such that

$$\langle d\Phi, \xi \rangle = -\iota(\xi_M)\omega \quad \text{for all } \xi \in \mathrm{Lie}(T),$$

where ξ_M is the vector field on M which generates the action of the one-parameter subgroup $\{\exp(t\xi) \mid t \in \mathbb{R}\}$ of T. Note that the Liouville measure on M is defined to be $\int_A \omega^d/d!$ for an open subset $A \subset M$, and its push-forward $\Phi_*\omega^d/d!$ is called the *Duistermaat–Heckman measure* in $\mathrm{Lie}(T)^*$.

Consider the line bundle $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r}$ over $Z_{\mathcal{I}}$ determined by integral weights $\lambda_1, \dots, \lambda_r$. Then we have an integer vector $\boldsymbol{a} = (\boldsymbol{a}^{(1)}, \dots, \boldsymbol{a}^{(r)}) \in \mathbb{Z}^{N_1} \oplus \cdots \oplus \mathbb{Z}^{N_r}$ such that $\eta^*\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} = \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$ by Proposition 2.10. Let $\omega'_{\boldsymbol{i}}$, respectively $\omega'_{\mathcal{I}}$, be a closed 2-form corresponding to the first Chern class of the line bundle $\mathcal{L}_{\boldsymbol{i},\boldsymbol{a}} \to Z_{\boldsymbol{i}}$, respectively $\mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} \to Z_{\mathcal{I}}$. By taking averages of $\omega'_{\boldsymbol{i}}$ and $\omega'_{\mathcal{I}}$ by corresponding torus actions we have the following two 2-forms:

$$\omega_{\boldsymbol{i}} := \int_{a\in\widetilde{\boldsymbol{T}}} (a^*\omega'_{\boldsymbol{i}})\,da \quad \text{and} \quad \omega_{\mathcal{I}} := \int_{t\in\boldsymbol{T}} (t^*\omega'_{\mathcal{I}})\,dt. \tag{5-4}$$

Then the form $\omega_{\boldsymbol{i}}$, respectively $\omega_{\mathcal{I}}$, is a $\widetilde{\boldsymbol{T}}$-invariant, respectively $\boldsymbol{T}$-invariant, closed 2-form on $(Z_{\boldsymbol{i}}, \widetilde{\boldsymbol{T}})$, respectively $(Z_{\mathcal{I}}, \boldsymbol{T})$. Since compact tori $\widetilde{\boldsymbol{T}}$ and $\boldsymbol{T}$ are connected, we have that

$$[\omega_{\boldsymbol{i}}] = [\omega'_{\boldsymbol{i}}] \quad \text{in } H^2(Z_{\boldsymbol{i}};\mathbb{R}), \qquad [\omega_{\mathcal{I}}] = [\omega'_{\mathcal{I}}] \quad \text{in } H^2(Z_{\mathcal{I}};\mathbb{R}) \tag{5-5}$$

(see [Guillemin et al. 2002, Corollary B.13]).

Grossberg and Karshon [1994] proved that the Duistermaat–Heckman measure of the presymplectic manifold $(Z_{\boldsymbol{i}}, \omega_{\boldsymbol{i}}, \widetilde{\boldsymbol{T}})$ can be computed by considering a combinatorial object, called a *Grossberg–Karshon twisted cube*. We use it to compute the Duistermaat–Heckman measure of the presymplectic manifold $(Z_{\mathcal{I}}, \omega_{\mathcal{I}}, \boldsymbol{T})$.

We recall from [Grossberg and Karshon 1994, §2.5] the definition of Grossberg–Karshon twisted cubes. Let $\boldsymbol{i} = (i_1, \dots, i_N)$ be a sequence of elements in $[n]$ and $\boldsymbol{a} = (a_1, \dots, a_N) \in \mathbb{Z}^N$. A Grossberg–Karshon twisted cube is a pair $(C(\boldsymbol{i}, \boldsymbol{a}), \rho)$, where $C(\boldsymbol{i}, \boldsymbol{a})$ is a subset of $\mathbb{R}^N$ and $\rho : \mathbb{R}^N \to \mathbb{R}$ is a density function with support equal to $C(\boldsymbol{i}, \boldsymbol{a})$. We define the following functions on $\mathbb{R}^N$:

$$\begin{aligned} A_N(x) &= A_N(x_1, \dots, x_N) = -\langle a_N \varpi_{i_N}, \alpha_{i_N}^\vee \rangle, \\ A_\ell(x) &= A_\ell(x_1, \dots, x_N) \\ &= -\langle a_\ell \varpi_{i_\ell} + \cdots + a_N \varpi_{i_N}, \alpha_{i_\ell}^\vee \rangle - \sum_{j>\ell} \langle \alpha_{i_j}, \alpha_{i_\ell}^\vee \rangle x_j \quad \text{for } 1 \le \ell \le N-1. \end{aligned}$$

We also define a function $\operatorname{sign} : \mathbb{R} \to \{\pm 1\}$ as $\operatorname{sign}(x) = -1$ for $x \le 0$ and $\operatorname{sign}(x) = 1$ for $x > 0$.

Definition 5.4. Let $C(\boldsymbol{i}, \boldsymbol{a})$ be the following subset of $\mathbb{R}^N$:

$$C(\boldsymbol{i}, \boldsymbol{a}) := \bigl\{ x = (x_1, \dots, x_N) \in \mathbb{R}^N \mid A_j(x) \le x_j \le 0 \text{ or } 0 < x_j < A_j(x) \text{ for } 1 \le j \le N \bigr\}.$$

We define a density function $\rho : \mathbb{R}^N \to \mathbb{R}$ whose support is $C(\boldsymbol{i}, \boldsymbol{a})$ and $\rho(x) = (-1)^N \operatorname{sign}(x_1) \cdots \operatorname{sign}(x_N)$ on the set $C(\boldsymbol{i}, \boldsymbol{a})$. We call the pair $(C(\boldsymbol{i}, \boldsymbol{a}), \rho)$ the *Grossberg–Karshon twisted cube associated to* $\boldsymbol{i}$ *and* $\boldsymbol{a}$. Also we define a measure

$$m_{C(\boldsymbol{i}, \boldsymbol{a})} = \rho(\alpha)|d\alpha|,$$

where $|d\alpha|$ is the Lebesgue measure in $\mathbb{R}^N$.

Now we have the following theorem.

Theorem 5.5. *Let $(Z_{\mathcal{I}}, \omega_{\mathcal{I}}, \boldsymbol{T})$ be as above, and let $\Phi : Z_{\mathcal{I}} \to \mathbb{R}^{m_1+\cdots+m_r}$ be a moment map of $(Z_{\mathcal{I}}, \omega_{\mathcal{I}}, \boldsymbol{T})$. Then there is a Grossberg–Karshon twisted cube $(C(\boldsymbol{i}, \boldsymbol{a}), \rho)$ and an affine projection $L : \mathbb{R}^N \to \mathbb{R}^{m_1+\cdots+m_r}$ such that the Duistermaat–Heckman measure in* $\mathrm{Lie}(\boldsymbol{T})^* \cong \mathbb{R}^{m_1+\cdots+m_r}$ *is $L_* m_{C(\boldsymbol{i},\boldsymbol{a})}$.*

To give a proof, we need the following theorem.

Theorem 5.6 [Grossberg and Karshon 1994, Theorem 2]. *Let $\widetilde{\Phi} : Z_{\boldsymbol{i}} \to \mathbb{R}^N$ be a moment map of $(Z_{\boldsymbol{i}}, \omega_{\boldsymbol{i}}, \widetilde{\boldsymbol{T}})$. Then the Duistermaat–Heckman measure in* $\mathrm{Lie}(\widetilde{\boldsymbol{T}})^* \cong \mathbb{R}^N$ *coincides with the measure $m_{C(\boldsymbol{i},\boldsymbol{a})}$ for the Grossberg–Karshon twisted cube $C(\boldsymbol{i}, \boldsymbol{a})$.*

Proof of Theorem 5.5. Suppose that $\boldsymbol{i} \in [n]^N$ defines a Bott–Samelson manifold $Z_{\boldsymbol{i}}$ which has a birational morphism $\eta : Z_{\boldsymbol{i}} \to Z_{\mathcal{I}}$. For given weights $\lambda_1, \dots, \lambda_r$, let $\boldsymbol{a} \in \mathbb{Z}^N$ be an integer vector such that $\eta^* \mathcal{L}_{\mathcal{I},\lambda_1,\dots,\lambda_r} = \mathcal{L}_{\boldsymbol{i},\boldsymbol{a}}$. Consider the pullback of $\omega_{\mathcal{I}}$ under the map η. Then we have $[\omega_{\boldsymbol{i}}] = [\eta^*(\omega_{\mathcal{I}})]$ in $H^2(Z_{\boldsymbol{i}}; \mathbb{R})$ by (5-5).

Now we have the following diagram which does not necessarily commute because two forms $\eta^* \omega_{\mathcal{I}}$ and $\omega_{\boldsymbol{i}}$ do not necessarily coincide because of taking averages:

$$\begin{array}{ccc} Z_{\boldsymbol{i}} & \xrightarrow{\widetilde{\Phi}} & \mathbb{R}^N \cong \mathrm{Lie}(\widetilde{\boldsymbol{T}})^* \\ \big\downarrow \eta & & \big\downarrow L \\ Z_{\mathcal{I}} & \xrightarrow{\Phi} & \mathbb{R}^{m_1+\cdots+m_r} \cong \mathrm{Lie}(\boldsymbol{T})^* \end{array}$$

Here, the map $L : \mathbb{R}^N \to \mathbb{R}^{m_1+\cdots+m_r}$ is defined as dA^*, where $A : \boldsymbol{T} \to \widetilde{\boldsymbol{T}}$ in (5-3).

But one can see that $L \circ \widetilde{\Phi}$, respectively $\Phi \circ \eta$, is a moment map for $(Z_{\boldsymbol{i}}, \omega_{\boldsymbol{i}}, \boldsymbol{T})$, respectively $(Z_{\boldsymbol{i}}, \eta^* \omega_{\mathcal{I}}, \boldsymbol{T})$. Recall from [Grossberg and Karshon 1994, Theorem 1] that the push-forward of Liouville measure only depends on the cohomology class, so we have that

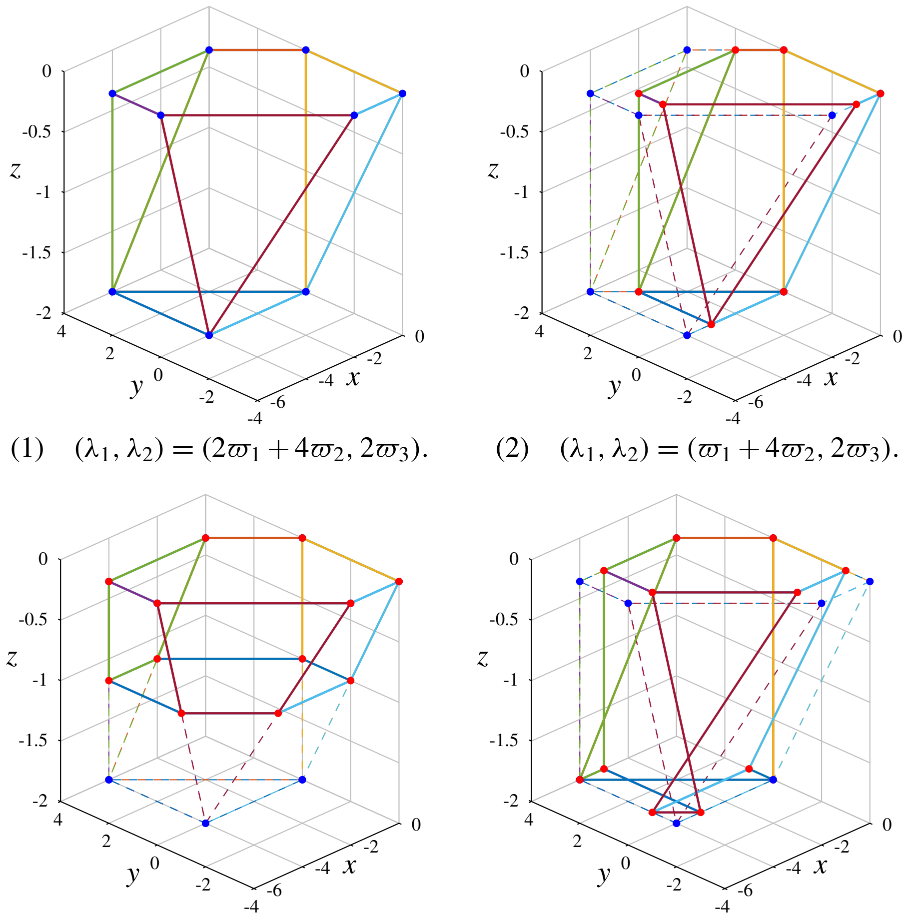
$$(L \circ \widetilde{\Phi})_* \omega_{\boldsymbol{i}}^N = (\Phi \circ \eta)_* (\eta^* \omega_{\mathcal{I}})^N = \Phi_* \omega_{\mathcal{I}}^N.$$

Here the last equality holds since η induces a diffeomorphism between Zariski open dense subsets, and a Zariski closed subset is measure zero. By Theorem 5.6, we have that $\Phi_* \omega_{\mathcal{I}}^N / N! = L_* m_{C(\boldsymbol{i},\boldsymbol{a})}$, so the result follows. □

Example 5.7. Let $G = \mathrm{SL}(4)$, $\mathcal{I} = (\{1, 2\}, \{3\})$ and $\boldsymbol{i} = (1, 2, 1, 3)$. The projection map $L = (dA)^* : \mathbb{R}^4 \to \mathbb{R}^3$ is given by the integer matrix

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

as in Example 5.3. In Figure 2 we draw figures for four different pairs of weights $(\lambda_1, \lambda_2) = (2\varpi_1 + 4\varpi_2, 2\varpi_3), (\varpi_1 + 4\varpi_2, 2\varpi_3), (2\varpi_1 + 4\varpi_2, \varpi_3), (2\varpi_1 + 3\varpi_2, 2\varpi_3)$



(1) $(\lambda_1, \lambda_2) = (2\varpi_1 + 4\varpi_2, 2\varpi_3)$. (2) $(\lambda_1, \lambda_2) = (\varpi_1 + 4\varpi_2, 2\varpi_3)$.

(3) $(\lambda_1, \lambda_2) = (2\varpi_1 + 4\varpi_2, \varpi_3)$. (4) $(\lambda_1, \lambda_2) = (2\varpi_1 + 3\varpi_2, 2\varpi_3)$.

Figure 2. The projection images of Grossberg–Karshon twisted cubes.

which determine line bundles $\mathcal{L}_{\mathcal{I},\lambda_1,\lambda_2}$. The polytope in Figure 2(1) has eight facets. When we change an integer vector (λ_1, λ_2) a little bit, some facets move as one can see in the figure. In Figure 2(2)–(4) the red dots represent vertices of the projection for the corresponding integer vector, and the blue dots represent vertices of the projection for $(\lambda_1, \lambda_2) = (2\varpi_1 + 4\varpi_2, 2\varpi_3)$. For pairs $(2\varpi_1 + 4\varpi_2, 2\varpi_3)$, $(\varpi_1 + 4\varpi_2, 2\varpi_3)$, and $(2\varpi_1 + 4\varpi_2, \varpi_3)$, the projections are honest polytopes while the projection for $(2\varpi_1 + 3\varpi_2, 2\varpi_3)$ is not.

Remark 5.8. Note that a Grossberg–Karshon twisted cube is neither closed not convex. When the Grossberg–Karshon twisted cube is a closed convex polytope, then we say it is *untwisted*. In [Lee 2020], an interpretation of untwistedness of Grossberg–Karshon twisted cubes $C(\boldsymbol{i}, \boldsymbol{a})$ using combinatorics of $\boldsymbol{i}$ and $\boldsymbol{a}$ is provided. (Also, see [Harada and Yang 2015; Harada and Lee 2015].) Using the result [Lee 2020, Theorem 1], Grossberg–Karshon twisted cubes appearing in Example 5.7 are all *twisted*. However, their projections can be honest polytopes as we saw in Figure 2. Determining whether the projection of a Grossberg–Karshon twisted cube is an honest polytope is a still open problem.

Acknowledgments

We are grateful to the anonymous referee for a careful reading and helpful comments to improve our initial manuscript.

References

[Andersen 1985] H. H. Andersen, "Schubert varieties and Demazure's character formula", *Invent. Math.* **79**:3 (1985), 611–618. MR Zbl

[Audin 2004] M. Audin, *Torus actions on symplectic manifolds*, 2nd revised ed., Progr. Math. **93**, Birkhäuser, Basel, 2004. MR Zbl

[Berenstein and Zelevinsky 2001] A. Berenstein and A. Zelevinsky, "Tensor product multiplicities, canonical bases and totally positive varieties", *Invent. Math.* **143**:1 (2001), 77–128. MR Zbl

[Bott and Samelson 1958] R. Bott and H. Samelson, "Applications of the theory of Morse to symmetric spaces", *Amer. J. Math.* **80** (1958), 964–1029. MR Zbl

[Bott and Tu 1982] R. Bott and L. W. Tu, *Differential forms in algebraic topology*, Grad. Texts Math. **82**, Springer, 1982. MR Zbl

[Bourbaki 2002] N. Bourbaki, *Lie groups and Lie algebras: Chapters 4–6*, Springer, 2002. MR Zbl

[Bredon 1972] G. E. Bredon, *Introduction to compact transformation groups*, Pure Appl. Math. **46**, Academic, New York, 1972. MR Zbl

[Brion and Kannan 2019a] M. Brion and S. S. Kannan, "Minimal rational curves on generalized Bott–Samelson varieties", 2019. To appear in *Compos. Math.* arXiv

[Brion and Kannan 2019b] M. Brion and S. S. Kannan, "Some combinatorial aspects of generalised Bott–Samelson varieties", preprint, 2019. arXiv

[Choi et al. 2010] S. Choi, M. Masuda, and D. Y. Suh, "Topological classification of generalized Bott towers", *Trans. Amer. Math. Soc.* **362**:2 (2010), 1097–1112. MR Zbl

[Demazure 1974] M. Demazure, "Désingularisation des variétés de Schubert généralisées", *Ann. Sci. École Norm. Sup.* (4) **7** (1974), 53–88. MR Zbl

[Elnitsky 1997] S. Elnitsky, "Rhombic tilings of polygons and classes of reduced words in Coxeter groups", *J. Combin. Theory Ser. A* **77**:2 (1997), 193–221. MR Zbl

[Escobar et al. 2018] L. Escobar, O. Pechenik, B. E. Tenner, and A. Yong, "Rhombic tilings and Bott–Samelson varieties", *Proc. Amer. Math. Soc.* **146**:5 (2018), 1921–1935. MR Zbl

[Fujita 2018] N. Fujita, "Newton–Okounkov bodies for Bott–Samelson varieties and string polytopes for generalized Demazure modules", *J. Algebra* **515** (2018), 408–447. MR Zbl

[Fujita and Naito 2017] N. Fujita and S. Naito, "Newton–Okounkov convex bodies of Schubert varieties and polyhedral realizations of crystal bases", *Math. Z.* **285**:1-2 (2017), 325–352. MR Zbl

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Mathematik (3) **2**, Springer, 1998. MR Zbl

[Grossberg and Karshon 1994] M. Grossberg and Y. Karshon, "Bott towers, complete integrability, and the extended character of representations", *Duke Math. J.* **76**:1 (1994), 23–58. MR Zbl

[Guillemin et al. 2002] V. Guillemin, V. Ginzburg, and Y. Karshon, *Moment maps, cobordisms, and Hamiltonian group actions*, Math. Surv. Monogr. **98**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Hansen 1973] H. C. Hansen, "On cycles in flag manifolds", *Math. Scand.* **33** (1973), 269–274. MR Zbl

[Harada and Kaveh 2015] M. Harada and K. Kaveh, "Integrable systems, toric degenerations and Okounkov bodies", *Invent. Math.* **202**:3 (2015), 927–985. MR Zbl

[Harada and Lee 2015] M. Harada and E. Lee, "Grossberg–Karshon twisted cubes and hesitant walk avoidance", *Pacific J. Math.* **278**:1 (2015), 119–136. MR Zbl

[Harada and Yang 2015] M. Harada and J. J. Yang, "Grossberg–Karshon twisted cubes and basepoint-free divisors", *J. Korean Math. Soc.* **52**:4 (2015), 853–868. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Grad. Texts Math. **52**, Springer, 1977. MR Zbl

[Humphreys 1975] J. E. Humphreys, *Linear algebraic groups*, Grad. Texts Math. **21**, Springer, 1975. MR Zbl

[Jantzen 2003] J. C. Jantzen, *Representations of algebraic groups*, 2nd ed., Mathematical Surveys and Monographs **107**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Kaji et al. 2020] S. Kaji, S. Kuroki, E. Lee, and D. Y. Suh, "Flag Bott manifolds of general Lie type and their equivariant cohomology rings", *Homology Homotopy Appl.* **22**:1 (2020), 375–390. MR Zbl

[Kashiwara 1991] M. Kashiwara, "On crystal bases of the q-analogue of universal enveloping algebras", *Duke Math. J.* **63**:2 (1991), 465–516. MR Zbl

[Kashiwara 1993] M. Kashiwara, "The crystal base and Littelmann's refined Demazure character formula", *Duke Math. J.* **71**:3 (1993), 839–858. MR Zbl

[Kashiwara 1995] M. Kashiwara, "On crystal bases", pp. 155–197 in *Representations of groups* (Banff, AB, 1994), edited by B. N. Allison and G. H. Cliff, CMS Conf. Proc. **16**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Kaveh 2015] K. Kaveh, "Crystal bases and Newton–Okounkov bodies", *Duke Math. J.* **164**:13 (2015), 2461–2506. MR Zbl

[Kaveh and Khovanskii 2012a] K. Kaveh and A. G. Khovanskii, "Convex bodies associated to actions of reductive groups", *Mosc. Math. J.* **12**:2 (2012), 369–396. MR Zbl

[Kaveh and Khovanskii 2012b] K. Kaveh and A. G. Khovanskii, "Newton–Okounkov bodies, semi-groups of integral points, graded algebras and intersection theory", *Ann. of Math.* (2) **176**:2 (2012), 925–978. MR Zbl

[Kumar 1987] S. Kumar, "Demazure character formula in arbitrary Kac–Moody setting", *Invent. Math.* **89**:2 (1987), 395–423. MR Zbl

[Kuroki et al. 2020] S. Kuroki, E. Lee, J. Song, and D. Y. Suh, "Flag Bott manifolds and the toric closure of a generic orbit associated to a generalized Bott manifold", *Pacific J. Math.* **308**:2 (2020), 347–392.

[Lakshmibai and Raghavan 2008] V. Lakshmibai and K. N. Raghavan, *Standard monomial theory: invariant theoretic approach*, Encycl. Math. Sci. **137**, Springer, 2008. MR Zbl

[Lakshmibai and Seshadri 1991] V. Lakshmibai and C. S. Seshadri, "Standard monomial theory", pp. 279–322 in *Proceedings of the Hyderabad Conference on Algebraic Groups* (Hyderabad, India, 1989), edited by S. Ramanan et al., Prakashan, Madras, 1991. MR Zbl

[Lakshmibai et al. 1979] V. Lakshmibai, C. Musili, and C. S. Seshadri, "Geometry of G/P", *Bull. Amer. Math. Soc.* (*N.S.*) **1**:2 (1979), 432–435. MR Zbl

[Lakshmibai et al. 2002] V. Lakshmibai, P. Littelmann, and P. Magyar, "Standard monomial theory for Bott–Samelson varieties", *Compos. Math.* **130**:3 (2002), 293–318. MR Zbl

[Lauritzen and Thomsen 2004] N. Lauritzen and J. F. Thomsen, "Line bundles on Bott–Samelson varieties", *J. Algebraic Geom.* **13**:3 (2004), 461–473. MR Zbl

[Lee 2020] E. Lee, "Grossberg–Karshon twisted cubes and hesitant jumping walk avoidance", *Electron. J. Combin.* **27**:3 (2020), art. id. P3.34. Zbl

[Littelmann 1998] P. Littelmann, "Cones, crystals, and patterns", *Transform. Groups* **3**:2 (1998), 145–179. MR Zbl

[Lusztig 1990] G. Lusztig, "Canonical bases arising from quantized enveloping algebras", *J. Amer. Math. Soc.* **3**:2 (1990), 447–498. MR Zbl

[Lusztig 1991] G. Lusztig, "Quivers, perverse sheaves, and quantized enveloping algebras", *J. Amer. Math. Soc.* **4**:2 (1991), 365–421. MR Zbl

[Lusztig 1993] G. Lusztig, *Introduction to quantum groups*, Progr. Math. **110**, Birkhäuser, Boston, 1993. MR Zbl

[Pasquier 2010] B. Pasquier, "Vanishing theorem for the cohomology of line bundles on Bott–Samelson varieties", *J. Algebra* **323**:10 (2010), 2834–2847. MR Zbl

[Perrin 2007] N. Perrin, "Small resolutions of minuscule Schubert varieties", *Compos. Math.* **143**:5 (2007), 1255–1312. MR Zbl

[Seshadri 2007] C. S. Seshadri, *Introduction to the theory of standard monomials*, Texts Readings Math. **46**, Hindustan, New Delhi, 2007. MR Zbl

[Springer 1998] T. A. Springer, *Linear algebraic groups*, 2nd ed., Progr. Math. **9**, Birkhäuser, Boston, 1998. MR Zbl

[Tenner 2006] B. E. Tenner, "Reduced decompositions and permutation patterns", *J. Algebraic Combin.* **24**:3 (2006), 263–284. MR Zbl

Received July 14, 2019. Revised April 13, 2020.

NAOKI FUJITA
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
THE UNIVERSITY OF TOKYO
MEGURO-KU
TOKYO
JAPAN
nfujita@ms.u-tokyo.ac.jp

EUNJEONG LEE
CENTER FOR GEOMETRY AND PHYSICS
INSTITUTE FOR BASIC SCIENCE
POHANG
REPUBLIC OF KOREA
eunjeong.lee@ibs.re.kr

DONG YOUP SUH
DEPARTMENT OF MATHEMATICAL SCIENCES
KAIST
YUSEONG-GU
DAEJEON
REPUBLIC OF KOREA
dysuh@kaist.ac.kr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.195

ON A MODULAR FORM OF ZAREMBA'S CONJECTURE

NIKOLAY G. MOSHCHEVITIN AND ILYA D. SHKREDOV

We prove that for any prime p there is a divisible by p number $q = O(p^{30})$ such that for a certain positive integer a coprime with q the ratio a/q has bounded partial quotients. In the other direction we show that there is an absolute constant $C > 0$ such that for any prime p exist divisible by p number $q = O(p^C)$ and a number a, a coprime with q such that all partial quotients of the ratio a/q are bounded by two.

1. Introduction

Let a and q be two positive coprime integers, $0 < a < q$. By the Euclidean algorithm, a rational a/q can be uniquely represented as a regular continued fraction

$$\frac{a}{q} = [0; b_1, \dots, b_s] = \cfrac{1}{b_1 + \cfrac{1}{b_2 + \cfrac{1}{b_3 + \dots + \cfrac{1}{b_s}}}} \qquad b_s \geqslant 2. \tag{1}$$

Assuming q is known, we use $b_j(a)$, $j = 1, \dots, s = s(a)$ to denote the partial quotients of a/q; that is,

$$\frac{a}{q} := [0; b_1(a), \dots, b_s(a)].$$

Zaremba's famous conjecture [1972] posits that there is an absolute constant $\mathfrak{k}$ with the following property: for any positive integer q there exists a coprime to q such that in the continued fraction expansion (1) all partial quotients are bounded:

$$b_j(a) \leqslant \mathfrak{k}, \quad 1 \leqslant j \leqslant s = s(a).$$

In fact, Zaremba conjectured that $\mathfrak{k} = 5$. For large prime q, even $\mathfrak{k} = 2$ should be enough, as conjectured by Hensley [1994; 1996]. This theme is rather popular, especially recently; see, e.g., [Bourgain and Kontorovich 2011; 2014; Frolenkov

This work is supported by the Russian Science Foundation under grant 19-11-00001.
MSC2020: 11B13, 11B75, 11E57, 11J70.
Keywords: continued fractions, Zaremba's conjecture, growth in groups.

and Kan 2014; Hensley 1989; 1996; Kan 2016; Kontorovich 2013; Korobov 1963; Moshchevitin 2007; Niederreiter 1986] and many others. The history of the question can be found, e.g., in [Moshchevitin et al. 2020]. Here we obtain the following "modular" version of Zaremba's conjecture. The first theorem in this direction was proved by Hensley [1994] and after that in [Magee et al. 2014; 2019].

Theorem 1. *There is an absolute constant* $\mathfrak{k}$ *such that for any prime number* p *there exist some positive integers* $q = O(p^{30})$, $q \equiv 0 \pmod p$ *and* a, *a coprime with* q *having the property that the ratio* a/q *has partial quotients bounded by* $\mathfrak{k}$.

Also, we can say something nontrivial about finite continued fractions with $\mathfrak{k} = 2$. It differs our paper from [Bourgain and Kontorovich 2011; 2014; Kan 2016; Magee et al. 2014; 2019].

Theorem 2. *There is an absolute constant* $C > 0$ *such that for any prime number* p *there exist some positive integers* $q = O(p^C)$, $q \equiv 0 \pmod p$ *and* a, *a coprime with* q *having the property that the ratio* a/q *has partial quotients bounded by* 2.

Our proof uses growth results in $\mathrm{SL}_2(\mathbb{F}_p)$ and some well-known facts about the representation theory of $\mathrm{SL}_2(\mathbb{F}_q)$. We study a combinatorial question about intersection of powers of a certain set of matrices $A \subseteq \mathrm{SL}_2(\mathbb{F}_q)$ with an arbitrary Borel subgroup and this seems like a new innovation.

In principle, results from [Hensley 1994] can be written in a form similar to Theorem 1 in an effective way but the dependence of q on p in [Hensley 1994] is rather poor. Thus Theorem 1 can be considered as an explicit version (with very concrete constants) of Hensley's results as well as rather effective Theorem 2 from [Magee et al. 2019]. Also, the methods of [Hensley 1994; Magee et al. 2014; 2019] are very different from ours.

2. Definitions

Let $\mathbf{G}$ be a group with the identity 1. Given two sets $A, B \subset \mathbf{G}$, define the *product set* of A and B as

$$AB := \{ab : a \in A,\ b \in B\}.$$

In a similar way we define the higher product sets, e.g., A^3 is AAA. Let $A^{-1} := \{a^{-1} : a \in A\}$. The Ruzsa triangle inequality [1996] says that

$$|C||AB| \leqslant |AC||C^{-1}B|$$

for any sets $A, B, C \subseteq \mathbf{G}$. As usual, having two subsets A, B of a group $\mathbf{G}$ denote by

$$\mathsf{E}(A, B) = \left|\left\{(a, a_1, b, b_1) \in A^2 \times B^2 : a^{-1}b = a_1^{-1}b_1\right\}\right| \tag{2}$$

the *common energy* of A and B. Clearly, $\mathsf{E}(A, B) = \mathsf{E}(B, A)$ and by the Cauchy–Schwarz inequality

$$\mathsf{E}(A, B)|A^{-1}B| \geqslant |A|^2|B|^2. \tag{3}$$

We use representation function notations like $r_{AB}(x)$ or $r_{AB^{-1}}(x)$, which counts the number of ways $x \in \boldsymbol{G}$ can be expressed as a product ab or ab^{-1} with $a \in A$, $b \in B$, respectively. For example, $|A| = r_{AA^{-1}}(1)$ and $\mathsf{E}(A, B) = r_{AA^{-1}BB^{-1}}(1) = \sum_x r^2_{A^{-1}B}(x)$. In this paper we use the same letter to denote a set $A \subseteq \boldsymbol{G}$ and its characteristic function $A : \boldsymbol{G} \to \{0, 1\}$. We write $\mathbb{F}_q^*$ for $\mathbb{F}_q \setminus \{0\}$. The signs $\ll$ and $\gg$ are the usual Vinogradov symbols. All logarithms are to base 2.

3. On the representation theory of $\mathrm{SL}_2(\mathbb{F}_p)$ and basis properties of its subsets

First of all, we recall some notions and simple facts from the representation theory; see, e.g., [Naimark 2010] or [Serre 1967]. For a finite group $\boldsymbol{G}$ let $\widehat{\boldsymbol{G}}$ be the set of all equivalence classes of irreducible unitary representations of $\boldsymbol{G}$. It is well-known that size of $\widehat{\boldsymbol{G}}$ coincides with the number of all conjugacy classes of $\boldsymbol{G}$. For $\rho \in \widehat{\boldsymbol{G}}$ denote by d_ρ the dimension of this representation. We write $\langle \cdot, \cdot \rangle$ for the corresponding Hilbert–Schmidt scalar product $\langle A, B \rangle = \langle A, B \rangle_{HS} := \mathrm{tr}(AB^*)$, where A, B are any two matrices of the same sizes. Put $\|A\| = \sqrt{\langle A, A \rangle}$. Clearly, $\langle \rho(g)A, \rho(g)B \rangle = \langle A, B \rangle$ and $\langle AX, Y \rangle = \langle X, A^*Y \rangle$. Also, we have $\sum_{\rho \in \widehat{\boldsymbol{G}}} d_\rho^2 = |\boldsymbol{G}|$.

For any $f : \boldsymbol{G} \to \mathbb{C}$ and $\rho \in \widehat{\boldsymbol{G}}$ define the matrix $\hat{f}(\rho)$, which is called the Fourier transform of f at ρ by the formula

$$\hat{f}(\rho) = \sum_{g \in \boldsymbol{G}} f(g)\rho(g). \tag{4}$$

Then the inverse formula takes place

$$f(g) = \frac{1}{|\boldsymbol{G}|} \sum_{\rho \in \widehat{\boldsymbol{G}}} d_\rho \langle \hat{f}(\rho), \rho(g^{-1}) \rangle, \tag{5}$$

and the Parseval identity is

$$\sum_{g \in \boldsymbol{G}} |f(g)|^2 = \frac{1}{|\boldsymbol{G}|} \sum_{\rho \in \widehat{\boldsymbol{G}}} d_\rho \|\hat{f}(\rho)\|^2. \tag{6}$$

The main property of the Fourier transform is the convolution formula

$$\widehat{f * g}(\rho) = \hat{f}(\rho)\hat{g}(\rho), \tag{7}$$

where the convolution of two functions $f, g : \boldsymbol{G} \to \mathbb{C}$ is defined as

$$(f * g)(x) = \sum_{y \in \boldsymbol{G}} f(y) g(y^{-1}x).$$

In terms of representations we can express the common energy of two sets $A, B \subseteq \boldsymbol{G}$, as defined in (2). Indeed, using (6) and (7), we derive

$$\mathsf{E}(A, B) = \sum_x (A^{-1} * B)(x) = \frac{1}{|\,\mathrm{SL}_2(\mathbb{F}_q)|} \sum_\rho d_\rho \|\widehat{A}^*(\rho)\widehat{B}(\rho)\|^2. \tag{8}$$

Finally, it is easy to check that for any matrices A, B one has $\|AB\| \leqslant \|A\|_o \|B\|$ and $\|A\|_o \leqslant \|A\|$, where the operator l^2-norm $\|A\|_o$ is just the absolute value of the maximal singular value of A. In particular, this shows that $\|\cdot\|$ is indeed a matrix norm.

Now consider the group $\mathrm{SL}_2(\mathbb{F}_q)$ of matrices

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ab|cd), \quad a, b, c, d \in \mathbb{F}_q, \quad ad - bc = 1.$$

Clearly, $|\,\mathrm{SL}_2(\mathbb{F}_q)| = q^3 - q$. Denote by B the standard Borel subgroup of all upper-triangular matrices from $\mathrm{SL}_2(\mathbb{F}_q)$, denote by $\mathrm{U} \subset \mathrm{B}$ the standard unipotent subgroup of $\mathrm{SL}_2(\mathbb{F}_q)$ of matrices $(1u|01)$, $u \in \mathbb{F}_q$ and denote by $\Delta \subset \mathrm{B}$ the subgroup of diagonal matrices. B and all its conjugates form all maximal proper subgroups of $\mathrm{SL}_2(\mathbb{F}_p)$. Detailed description of the representation theory of $\mathrm{SL}_2(\mathbb{F}_q)$ can be found in [Naimark 2010, Chapter II, Section 5]. We formulate the main result from [Naimark 2010] concerning this theme.

Theorem 3. *Let $p > 2$ be a prime number and $q = p^n$. There are $q + 3$ nontrivial representations of* $\mathrm{SL}_2(\mathbb{F}_q)$, *namely*:

- $\frac{1}{2}(q-3)$ *representations T_χ of dimension $q+1$ indexed via $\frac{1}{2}(q-3)$ nontrivial multiplicative characters χ on $\mathbb{F}_q^*$, $\chi^2 \neq 1$.*
- *A representation $\widetilde{T}_1$ of dimension q.*
- *Two representations $T^+_{\chi_1}$, $T^-_{\chi_1}$ of dimension $\frac{1}{2}(q+1)$, $\chi_1^2 = 1$.*
- *Two representations $S^+_{\pi_1}$, $S^-_{\pi_1}$ of dimension $\frac{1}{2}(q-1)$.*
- $\frac{1}{2}(q-1)$ *representations S_π of dimension $q-1$ indexed via $\frac{1}{2}(q-1)$ nontrivial multiplicative characters π on an arbitrary quadratic extension of $\mathbb{F}_q$, $\pi^2 \neq 1$.*

By $d_{\min}$ and $d_{\max}$ denote the minimum and maximum over dimensions of all nontrivial representations of a group $\boldsymbol{G}$. Thus the result above tells us that in the case $\boldsymbol{G} = \mathrm{SL}_2(\mathbb{F}_q)$ these quantities differ roughly by a factor of two. Below we assume that $q \geqslant 3$.

Theorem 3 has two consequences, although, a slightly weaker result than Lemma 4 can be obtained via the classical theorem of Frobenius [1896]; see, e.g., [Shkredov 2018]. Originally, similar arguments were suggested in [Gowers 2008; Nikolov and Pyber 2011; Sarnak and Xue 1991].

Lemma 4. *Let $n \geqslant 3$ be an integer, $A \subseteq \mathrm{SL}_2(\mathbb{F}_q)$ be a set and $|A| \geqslant 2(q+1)^2 q^{2/n}$. Then $A^n = \mathrm{SL}_2(\mathbb{F}_q)$. Generally, if for some sets $X_1, \dots, X_n \subseteq \mathrm{SL}_2(\mathbb{F}_q)$ one has*

$$\prod_{j=1}^{n} |X_j| \geqslant (2q(q+1))^n (q-1)^2,$$

then $X_1 \dots X_n = \mathrm{SL}_2(\mathbb{F}_q)$.

Proof. Using (6) with $f = A$ (i.e., according our notation f is the characteristic function of the set A), we have for an arbitrary nontrivial representation ρ that

$$\|\widehat{A}\|_o < \left(\frac{|A||\,\mathrm{SL}_2(\mathbb{F}_q)|}{d_{\min}}\right)^{1/2} = \left(\frac{|A|(q^3-q)}{d_{\min}}\right)^{1/2}. \tag{9}$$

Hence for any $x \in \mathrm{SL}_2(\mathbb{F}_q)$ we obtain via formulae (5), (6) and estimate (9) that

$$r_{A^n}(x) > \frac{|A|^n}{|\,\mathrm{SL}_2(\mathbb{F}_q)|} - \left(\frac{|A|(q^3-q)}{d_{\min}}\right)^{(n-2)/2} |A| \geqslant 0,$$

provided $|A|^n \geqslant 2^{n-2}(q+1)^n q^n (q-1)^2$. The second part of the lemma can be obtained similarly. This completes the proof. □

Remark 5. It is easy to see (or consult Lemma 6 below) that bound (9) is sharp, e.g., take $A = \mathrm{B}$.

For any function $f : \boldsymbol{G} \to \mathbb{C}$ consider the Wiener norm of f defined as

$$\|f\|_W := \frac{1}{|\boldsymbol{G}|} \sum_{\rho \in \widehat{\boldsymbol{G}}} d_\rho \|\hat{f}(\rho)\|. \tag{10}$$

Lemma 6. *Let $\boldsymbol{G}$ be a group and Γ be its subgroup. Then $\|\Gamma\|_W \leqslant 1$. Moreover, $\|\mathrm{B}\|_W = 1$, further $\|\widehat{\mathrm{B}}(\widetilde{T}_1)\| = \|\widehat{\mathrm{B}}(\widetilde{T}_1)\|_o = |\mathrm{B}|$ and the Fourier transform of B vanishes on all other nontrivial representations.*

Proof. Since Γ is a subgroup, we see using (6) twice that

$$|\Gamma|^2 = |\{(\gamma_1, \gamma_2, \gamma_3) \in \Gamma^3 : \gamma_1\gamma_2 = \gamma_3\}| = \frac{1}{|\boldsymbol{G}|} \sum_{\rho \in \widehat{\boldsymbol{G}}} d_\rho \langle \widehat{\Gamma}^2(\rho), \widehat{\Gamma}(\rho)\rangle$$
$$\leqslant \frac{1}{|\boldsymbol{G}|} \sum_{\rho} d_\rho \langle \widehat{\Gamma}(\rho), \widehat{\Gamma}(\rho)\rangle \|\widehat{\Gamma}(\rho)\|_o \leqslant \frac{|\Gamma|}{|\boldsymbol{G}|} \sum_{\rho} d_\rho \langle \widehat{\Gamma}(\rho), \widehat{\Gamma}(\rho)\rangle = |\Gamma|^2,$$

because, clearly, $\|\widehat{\Gamma}(\rho)\|_o \leqslant |\Gamma|$. This means that for any representation ρ either $\|\widehat{\Gamma}(\rho)\| = 0$ (and hence $\|\widehat{\Gamma}(\rho)\|_o = 0$) or $\|\widehat{\Gamma}(\rho)\| \geqslant \|\widehat{\Gamma}(\rho)\|_o = |\Gamma|$ (alternatively, one can use the usual calculations, namely, $\sum_{\gamma \in \Gamma} \rho(\gamma\gamma_*) = \sum_{\gamma \in \Gamma} \rho(\gamma) \cdot \rho(\gamma_*)$

for any $\gamma_* \in \Gamma$ but then one needs to be careful with divisors of zero). Another application of (6) gives us

$$(11)\qquad |\Gamma| = \frac{1}{|\boldsymbol{G}|}\sum_{\rho} d_\rho \|\widehat{\Gamma}(\rho)\|^2 \geqslant |\Gamma|\cdot\frac{1}{|\boldsymbol{G}|}\sum_{\rho} d_\rho \|\widehat{\Gamma}(\rho)\| = |\Gamma|\|\Gamma\|_W.$$

Hence $\|\Gamma\|_W \leqslant 1$ as required.

Now let us prove the second part of the lemma. We write I_n for the identity matrix and let Z_n be the zero matrix of size $n\times n$. Also, we write $\operatorname{diag}(d_1,\dots,d_n)$ for the diagonal matrix with diagonal entries $d_1,\dots,d_n$. Finally, let $e(\cdot)$ be an additive character of $\mathbb{F}_q$. For $u_b \in \mathrm{U}$, $u_b = (1b|01)$, we have [Naimark 2010, pp. 121–123] that in a certain orthogonal basis $\widetilde{T}_1(u_b) = \operatorname{diag}(1, e(b), \dots, e(q-1)b)$ and for $g_\lambda = (\lambda 0|0\lambda^{-1}) \in \Delta$ the matrix $\widetilde{T}_1(g_\lambda)$ is the direct sum of I_1 and a permutation matrix of size $(q-1)\times(q-1)$. Clearly, $\mathrm{B} = \Delta\mathrm{U} = \mathrm{U}\Delta$ and hence $\widehat{\mathrm{B}}(\rho) = \widehat{\Delta}(\rho)\widehat{\mathrm{U}}(\rho)$ for any representation ρ. But from above $\widehat{\mathrm{U}}(\widetilde{T}_1)$ is the direct sum $qI_1 \oplus Z_{q-1}$. Further one can show that $\widehat{\Delta}(\widetilde{T}_1) = (q-1)I_1 \oplus 2\cdot J$, where $J = (J_{ij})_{i,j=1}^{q-1}$ is a certain $(q-1)\times(q-1)$ matrix with all components equal one for i/j belonging to the set of quadratic residues. Such precise description of J is not really important for us, it is enough to see that $\widehat{\Delta}(\widetilde{T}_1)$ is a direct sum of $(q-1)I_1$ and a $(q-1)\times(q-1)$ matrix. Hence

$$\widehat{\mathrm{B}}(\widetilde{T}_1) = \widehat{\Delta}(\widetilde{T}_1)\widehat{\mathrm{U}}(\widetilde{T}_1) = q(q-1)I_1 \oplus Z_{q-1}.$$

Thus $\|\widehat{\mathrm{B}}(\widetilde{T}_1)\| = \|\widehat{\mathrm{B}}(\widetilde{T}_1)\|_o = |\mathrm{B}|$. Applying (11), we obtain

$$(12)\qquad |\mathrm{B}| \geqslant \frac{|\mathrm{B}|^2}{|\operatorname{SL}_2(\mathbb{F}_q)|} + \frac{q}{|\operatorname{SL}_2(\mathbb{F}_q)|}\|\widehat{\mathrm{B}}(\widetilde{T}_1)\|^2 = \frac{|\mathrm{B}|^2}{|\operatorname{SL}_2(\mathbb{F}_q)|}(1+q) = |\mathrm{B}|.$$

It follows that for any other representations Fourier coefficients of B vanish. This completes the proof. □

Lemma 6 gives us an alternative way to show that $A^3 \cap \mathrm{B} \neq \varnothing$. Indeed, just use estimate (9) and write

$$\begin{aligned} r_{A^3\mathrm{B}}(1) &\geqslant \frac{|A|^3|\mathrm{B}|}{|\operatorname{SL}_2(\mathbb{F}_q)|} - \|\mathrm{B}\|_W\left(\frac{|A|(q^3-q)}{d_{\min}}\right)^{3/2} \\ &= \frac{|A|^3|\mathrm{B}|}{|\operatorname{SL}_2(\mathbb{F}_q)|} - \left(\frac{|A|(q^3-q)}{d_{\min}}\right)^{3/2} > 0, \end{aligned}$$

provided

$$(13)\qquad |A| \gg q^{8/3}.$$

We improve this bound in the next section.

4. On intersections of the product set with the Borel subgroup

It was shown in the previous section (see Lemma 4) that for any $A \subseteq \mathrm{SL}_2(\mathbb{F}_q)$ one has $A^3 = \mathrm{SL}_2(\mathbb{F}_q)$, provided $|A|^3 \gg q^8$ and in the same way the last result holds for three different sets, namely, given $X, Y, Z \subseteq \mathrm{SL}_2(\mathbb{F}_q)$ with $|X||Y||Z| \gg q^8$, we have $XYZ = \mathrm{SL}_2(\mathbb{F}_q)$. It is easy to see that in this generality the last result is sharp. Indeed, let $X = S\mathrm{B}$, $Y = \mathrm{B}T$, where S, T are two sets of sizes $\sqrt{q}/2$ which are chosen as $|X| \sim |S||\mathrm{B}|$ and $|Y| \sim |T||\mathrm{B}|$ (e.g., take S, T from left/right cosets of B thanks to the Bruhat decomposition). Then $XY = S\mathrm{B}T$, and hence $|XY| \leqslant |S||T||\mathrm{B}| \leqslant |\mathrm{SL}_2(\mathbb{F}_q)|/2$. Thus we take Z^{-1} to equal the complement to XY in $\mathrm{SL}_2(\mathbb{F}_q)$ and we see that the product set XYZ does not contain 1 but $|X||Y||Z| \gg q^8$.

Nevertheless, in the "symmetric" case of the same set A this 8/3 bound (13) can be improved; see Theorem 9 below. We need a simple lemma and the proof of this result, as well as the proof of Theorem 9 extensively play on noncommutative properties of $\mathrm{SL}_2(\mathbb{F}_q)$.

Lemma 7. *Let $g \notin \mathrm{B}$ be a fixed element from $\mathrm{SL}_2(\mathbb{F}_q)$. Then for any x one has*

$$r_{\mathrm{B}g\mathrm{B}}(x) \leqslant q - 1.$$

Proof. Let $g = (ab|cd)$ and $x = (\alpha\beta|\gamma\delta)$. By our assumption $c \neq 0$. We have

$$\begin{pmatrix} \lambda & u \\ 0 & \lambda^{-1} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \mu & v \\ 0 & \mu^{-1} \end{pmatrix} = \begin{pmatrix} (\lambda a + uc)\mu & * \\ \mu c/\lambda & vc/\lambda + d/(\lambda\mu) \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{14}$$

In other words, $\mu = \lambda\gamma c^{-1} \neq 0$ (hence $\gamma \neq 0$ automatically) and from

$$\alpha = (\lambda a + uc)\mu = \lambda\gamma c^{-1}(\lambda a + uc)$$

we see that having λ we determine u uniquely (then, (14) gives us μ, v automatically). This completes the proof. □

Lemma 7 quickly implies a result on the Bruhat decomposition of $\mathrm{SL}_2(\mathbb{F}_q)$.

Corollary 8. *Let $g \in \mathrm{SL}_2(\mathbb{F}_q) \setminus \mathrm{B}$. Then $\mathrm{B}g\mathrm{B} = \mathrm{SL}_2(\mathbb{F}_q) \setminus \mathrm{B}$.*

Proof. Clearly, $\mathrm{B} \cap \mathrm{B}g\mathrm{B} = \varnothing$ because $g \in \mathrm{SL}_2(\mathbb{F}_q) \setminus \mathrm{B}$. On the other hand, by Lemma 7, we have

$$\mathsf{E}(\mathrm{B}, g\mathrm{B}) = \sum_x r^2_{\mathrm{B}g\mathrm{B}}(x) \leqslant (q-1) \sum_x r_{\mathrm{B}g\mathrm{B}}(x) = (q-1)|\mathrm{B}|^2.$$

Using the last bound and estimate (3), we obtain

$$|\mathrm{B}g\mathrm{B}| \geqslant \frac{|\mathrm{B}|^4}{\mathsf{E}(\mathrm{B}, g\mathrm{B})} \geqslant \frac{|\mathrm{B}|^4}{(q-1)|\mathrm{B}|^2} = q^3 - q^2 = |\mathrm{SL}_2(\mathbb{F}_q) \setminus \mathrm{B}|.$$

This completes the proof. □

Using growth of products of B as in the last corollary, one can combinatorially improve the constant 8/3 (to do this combine Lemma 4 and bound (22) below). We suggest another method which uses the representation theory of $\mathrm{SL}_2(\mathbb{F}_q)$ more extensively and which allows to improve this constant further.

Theorem 9. *Let $A \subseteq \mathrm{SL}_2(\mathbb{F}_q)$ be a set, $|A| \geqslant 4q^{18/7}$. Then $A^3 \cap \mathrm{B} \neq \varnothing$. Generally, $A^n \cap \mathrm{B} \neq \varnothing$ provided $|A| \geqslant 4q^{2+4/(3n-2)}$.*

Proof. Let $g \notin \mathrm{B}$ and put $A^\varepsilon_g = A^\varepsilon \cap g\mathrm{B}$, where $\varepsilon \in \{1, -1\}$. Also, let $\Delta = \max_{\varepsilon,\, g\notin \mathrm{B}} |A^\varepsilon_g|$. Since we can assume $A \cap \mathrm{B} = \varnothing$, it follows that

$$\mathsf{E}(A, \mathrm{B}) = \sum_x r^2_{A^{-1}\mathrm{B}}(x) = \sum_{x\notin \mathrm{B}} r^2_{A^{-1}\mathrm{B}}(x) \leqslant \Delta|\mathrm{B}||A| \tag{15}$$

and similarly for $\mathsf{E}(A^{-1}, \mathrm{B})$. On the other hand, from (8) and by the second part of Lemma 6, we see that

$$\begin{aligned} \Delta|\mathrm{B}||A| \geqslant \mathsf{E}(A, \mathrm{B}) &= \frac{1}{|\mathrm{SL}_2(\mathbb{F}_q)|} \sum_\rho d_\rho \|\widehat{A}^*(\rho)\widehat{\mathrm{B}}(\rho)\|^2 \\ &= \frac{q}{|\mathrm{SL}_2(\mathbb{F}_q)|} \|\widehat{A}^*(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1)\|^2, \end{aligned} \tag{16}$$

and, again, similarly for $\|\widehat{A}(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1)\|^2$. Now consider the equation $b_1 a' a'' a b_2 = 1$ or, equivalently the equation $a'' a b_2 = (a')^{-1} b_1^{-1}$, where $a, a', a'' \in A$ and $b_1, b_2 \in \mathrm{B}$. Clearly, if $A^3 \cap \mathrm{B} = \varnothing$, then this equation has no solutions. Combining Lemma 6 with bound (16) and calculations as in the proof of Lemma 4, we see that this equation can be solved provided

$$\begin{aligned} &\frac{q}{|\mathrm{SL}_2(\mathbb{F}_q)|} |\langle \widehat{A}^2(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1),\, \widehat{A}^*(\widetilde{T}_1)\widehat{\mathrm{B}}^*(\widetilde{T}_1)\rangle| \\ &\qquad\leqslant \frac{q}{|\mathrm{SL}_2(\mathbb{F}_q)|} \|\widehat{A}^2(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1)\| \cdot \|\widehat{A}^*(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1)\| \\ &\qquad\leqslant \frac{q}{|\mathrm{SL}_2(\mathbb{F}_q)|} \|\widehat{A}(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1)\| \, \|\widehat{A}^*(\widetilde{T}_1)\widehat{\mathrm{B}}(\widetilde{T}_1)\| \, \|\widehat{A}\|_o \\ &\qquad\leqslant \Delta|\mathrm{B}||A|\|\widehat{A}\|_o < \frac{|A|^3|\mathrm{B}|^2}{|\mathrm{SL}_2(\mathbb{F}_q)|}. \end{aligned}$$

In other words, in view of (9) it is enough to have

$$|A|^4 \geqslant 2(q+1)^2\Delta^2 \cdot |A| q(q+1) \tag{17}$$

or, equivalently,

$$2q(q+1)^3\Delta^2 \leqslant |A|^3. \tag{18}$$

Now let us obtain another bound which works well when Δ is large. Choose $g \notin \mathrm{B}$ and $\varepsilon \in \{1, -1\}$ such that $\Delta = |A_g^\varepsilon|$. Using Lemma 7, we derive

$$\mathsf{E}(\mathrm{B}, A_g^\varepsilon) = \sum_x r^2_{\mathrm{B}A_g^\varepsilon}(x) \leqslant \sum_x r_{\mathrm{B}A_g^\varepsilon}(x) r_{\mathrm{B}g\mathrm{B}}(x) \leqslant (q-1)|\mathrm{B}||A_g^\varepsilon|, \tag{19}$$

and hence by the Cauchy–Schwarz inequality, we get

$$|\mathrm{B}A_g^\varepsilon| \geqslant \frac{|\mathrm{B}|^2|A_g^\varepsilon|^2}{\mathsf{E}(\mathrm{B}, A_g^\varepsilon)} \geqslant \frac{|\mathrm{B}||A_g^\varepsilon|}{q-1} = q\Delta. \tag{20}$$

Consider the equation $a_g(a'a'')^\varepsilon = b$, where $b \in \mathrm{B}$, $a_g \in A_g^\varepsilon$ and $a', a'' \in A$. Clearly, if $A^3 \cap \mathrm{B} = \varnothing$, then this equation has no solutions. To solve $a_g(a'a'')^\varepsilon = b$ it is enough to solve the equation $z(a'a'')^\varepsilon = 1$, where now $z \in \mathrm{B}A_g^\varepsilon$. Applying the second part of Lemma 4 combining with (20), we obtain that it is enough to have

$$8q^3(q+1)^3(q-1)^2 \leqslant q\Delta|A|^2 \leqslant |\mathrm{B}A_g^\varepsilon||A|^2$$

or, in other words,

$$8q^2(q+1)^3(q-1)^2 \leqslant \Delta|A|^2. \tag{21}$$

Considering the second power of (21) and multiplying it with (18), we obtain

$$|A|^7 \geqslant 2^{14}q^{18} \geqslant 2^7 q^5 (q+1)^9 (q-1)^4$$

as required.

In the general case inequality (18) can be rewritten as

$$|A|^n \geqslant 2^{n-2}\Delta^2(q+1)^n q^{n-2}$$

and using the second part of Lemma 4, we obtain an analogue of (21),

$$|A|^{n-1}\Delta \geqslant 2^n q^{n-1}(q+1)^n(q-1)^2.$$

Combining the last two bounds, we derive the required result. This completes the proof. □

Remark 10. It is easy to see that Theorem 9, as well as Lemma 7 (and also Lemma 6) take place for any Borel subgroup not just for the standard one.

Remark 11. It is easy to see that the arguments of the proof of Theorem 9 give the following combinatorial statement about left/right multiplication of an arbitrary set A by B (just combine bounds (15) and (20)), namely,

$$\max\{|A\mathrm{B}|, |\mathrm{B}A|\} \gg \min\{q^{3/2}|A|^{1/2}, |A|^2 q^{-2}\}. \tag{22}$$

As we have seen by Theorem 9 we know that $A^n \cap \mathrm{B} \neq \varnothing$ for large n but under the condition $|A| \gg q^{2+\varepsilon}$ for a certain $\varepsilon > 0$. For the purpose of the next section we need to break the described q^2-barrier and we do this for prime q, using growth in $\mathrm{SL}_2(\mathbb{F}_p)$. Let us recall quickly what is known about growth of generating sets in $\mathrm{SL}_2(\mathbb{F}_p)$. Helfgott [2008] obtained his famous result in this direction and we proved in [Rudnev and Shkredov 2018] the following form of Helfgott's result.

Theorem 12. *Let $A \subseteq \mathrm{SL}_2(\mathbb{F}_p)$ be a set, $A = A^{-1}$ which generates the whole group. Then $|AAA| \gg |A|^{1+1/20}$.*

Thus in the case of an arbitrary symmetric generating set and a prime number p Theorem 12 combined with Theorem 9, allow to obtain some bounds which guarantee that $A^n = \mathrm{SL}_2(\mathbb{F}_p)$. For example, if A generates $\mathrm{SL}_2(\mathbb{F}_p)$, $A = A^{-1}$, and $|A| \gg p^{2-\epsilon}$, $\epsilon < \frac{2}{21}$, then $A^n \cap \mathrm{B} \neq \varnothing$ for $n \geqslant (84-42\epsilon)/(2-21\epsilon)$. On the other hand, the methods from [Helfgott 2008; Rudnev and Shkredov 2018] allow us to obtain the following result about generation of $\mathrm{SL}_2(\mathbb{F}_p)$ via large and not necessary symmetric sets (the condition of nonsymmetricity of A is rather crucial for us, see the next section).

Theorem 13. *Let $A \subseteq \mathrm{SL}_2(\mathbb{F}_p)$ be a generating set, $p \geqslant 5$ and $|A| \gg p^{2-\epsilon}$, $\epsilon < \frac{2}{25}$. Then $A^n \cap \mathrm{B} \neq \varnothing$ for $n \geqslant (100-50\epsilon)/(2-25\epsilon)$. Also, $A^n = \mathrm{SL}_2(\mathbb{F}_p)$, provided $n \geqslant 144/(2-25\epsilon)$.*

Proof. Put $K = |AAA|/|A|$. We can assume that, say, $|A| \leqslant p^{2+2/35}$ because otherwise one can apply Theorem 9. We call an element $g \in \mathrm{SL}_2(\mathbb{F}_p)$ regular if $\mathrm{tr}(g) \neq 0, \pm 2$ and let $\mathcal{C}_g$ be the correspondent conjugate class, namely,

$$\mathcal{C}_g = \{s \in \mathrm{SL}_2(\mathbb{F}_p) : \mathrm{tr}(s) = \mathrm{tr}(g)\}.$$

Let T be a maximal torus such that there is $g \in T \cap A^{-1}A$ and $g \neq 1$. By [Rudnev and Shkredov 2018, Lemma 5] such torus T_*, containing a regular element g, exists, otherwise $K \gg |A|^{2/3}$. Firstly, suppose that for a certain $h \in A$ the torus $T' = hTh^{-1}$ has no such property, i.e., there are no nontrivial elements from $A^{-1}A \cap T'$. Then for the element $g' = hgh^{-1} \in T'$ (in the case $T = T_*$ the element g' is regular) the projection $a \to ag'a^{-1}$, $a \in A$ is one-to-one. Hence $|A^2A^{-1}AA^{-2} \cap \mathcal{C}_g| \geqslant |A|$. By [Rudnev and Shkredov 2018, Lemma 11], we have $|S \cap \mathcal{C}_g| \ll |S^{-1}S|^{2/3} + p$ for any set S and regular g. Using the Ruzsa triangle inequality, we obtain

$$\begin{aligned} (23)\quad |(A^2A^{-1}AA^{-2})^{-1}(A^2A^{-1}AA^{-2})| &\leqslant |A|^{-1}|A^2A^{-1}AA^{-3}||A^3A^{-1}AA^{-2}| \\ &= |A|^{-1}|A^3A^{-1}AA^{-2}|^2|A|^{-1}(|A|^{-1}|A^3A^{-2}||A^2A^{-2}|)^2 \\ &\leqslant |A|^{-1}(|A|^{-3}|A^4||A^3|^3)^2 \leqslant K^{12}|A| \end{aligned}$$

and hence

$$|A| \ll |(A^2A^{-1}AA^{-2})^{-1}(A^2A^{-1}AA^{-2})|^{2/3} + p \ll K^8|A|^{2/3}.$$

This gives us $K \gg |A|^{1/24}$.

In the complementary second case (see [Rudnev and Shkredov 2018]) thanks to the fact that A is a generating set, we suppose that for *any* $h \in \mathrm{SL}_2(\mathbb{F}_p)$ there is a nontrivial element from $A^{-1}A$ belonging to the torus hTh^{-1}. Then $A^{-1}A$ is partitioned between these tori and hence again by [Rudnev and Shkredov 2018, Lemma 11], as well as the Ruzsa triangle inequality, we obtain

$$\begin{aligned}|(AA^{-1}AA^{-1})^{-1}(AA^{-1}AA^{-1})| &\leqslant |A|^{-1}|A^2A^{-1}AA^{-1}|^2\\ &\leqslant |A|^{-1}(|A|^{-1}|A^2A^{-2}||A^2A^{-1}|)^2\\ &\leqslant |A|^{-1}(|A|^{-3}|A^3|^4)^2 \leqslant K^8|A|\end{aligned}$$

and whence

$$\begin{aligned}K^2|A| \geqslant |A^{-1}A| &\geqslant \sum_{h\in\mathrm{SL}_2(\mathbb{F}_p)/N(T_*)} |A^{-1}A \cap hT_*h^{-1}|\\ &\gg p^2 \cdot \frac{|A|}{|(AA^{-1}AA^{-1})^{-1}(AA^{-1}AA^{-1})|^{2/3}}\\ &\geqslant p^2|A|^{1/3}K^{-16/3},\end{aligned}$$

where $N(T)$ is the normalizer of any torus T, $|N(T)| \asymp |T| \asymp p$. Hence thanks to our assumption $|A| \leqslant p^{2+2/35}$, we have $K \gg p^{3/11}|A|^{-1/11} \gg |A|^{1/24}$. In other words, we always obtain $|AAA| \gg p^{2+(2-25\epsilon)/24}$. After that apply Theorem 9 to find that $A^n \cap \mathrm{B} \neq \varnothing$ for $n \geqslant (100-50\epsilon)/(2-25\epsilon)$. If we use Lemma 4 instead of Theorem 9, then we obtain $A^n = \mathrm{SL}_2(\mathbb{F}_p)$, provided $n \geqslant 144/(2-25\epsilon)$. This completes the proof. □

Thus for sufficiently small $\epsilon > 0$ one can take $n = 51$ to get $A^n \cap \mathrm{B} \neq \varnothing$ (and $n = 73$ to obtain $A^n = \mathrm{SL}_2(\mathbb{F}_p)$). In the next section we improve this bound for a special set A but nevertheless the arguments of the proof of Theorem 13 will be used in the proof of Theorem 2 from the Introduction.

We finish this section showing that generating sets A of sizes close to p^2 (actually, the condition $|A| = \Omega(p^{3/2+\varepsilon})$ is enough) with small tripling constant $K = |A^3|/|A|$ avoid all Borel subgroups.

Lemma 14. *Let $A \subseteq \mathrm{SL}_2(\mathbb{F}_p)$ be a generating set, $p \geqslant 5$ and $K = |A^3|/|A|$. Then for any Borel subgroup B_* one has $|A \cap \mathrm{B}_*| \leqslant 2pK^{5/3}|A|^{1/3}$.*

Proof. We obtain the result for the standard Borel subgroup B and after that apply the conjugation to prove our lemma in full generality. Let $\gamma \in \mathbb{F}_p^*$ be any number and l_γ be the line

$$l_\gamma = \{(\gamma u|0\gamma^{-1}) : u \in \mathbb{F}_p\} \subset \mathrm{SL}_2(\mathbb{F}_p).$$

By [Rudnev and Shkredov 2018, Lemma 7], we have $|A \cap l_\gamma| \leqslant 2|A^3 A^{-1} A|^{1/3}$. Using the last bound, as well as the Ruzsa triangle inequality, we obtain

$$|A \cap \mathrm{B}| \leqslant \sum_{\gamma \in \mathbb{F}_p^*} |A \cap l_\gamma| \leqslant 2p|A^3 A^{-1} A|^{1/3} \leqslant 2p(|A^4||A^{-2}A|/|A|)^{1/3} \leqslant 2pK^{5/3}|A|^{1/3}.$$

This completes the proof. □

Remark 15. Examining the proof of Lemma 7 from [Rudnev and Shkredov 2018] one can equally write $|A \cap l_\gamma| \leqslant 2|A^3 A^{-2}|^{1/3}$ and hence by the calculations above $|A \cap \mathrm{B}_*| \leqslant 2pK^{4/3}|A|^{1/3}$. Nevertheless, this better estimate has no influence to the final bound in Theorem 1.

Remark 16. Bounds for intersections of $A \subseteq \mathrm{SL}_2(\mathbb{F}_q)$, $K = |A^3|/|A|$ with $g\mathrm{B}_*$, where $g \notin \mathrm{B}_*$ are much simpler and follow from Lemma 7 (also, see Remark 10). Indeed, by this result putting $A_* = A \cap g\mathrm{B}_*$, we have

$$K|A| \geqslant |AA| \geqslant |A_* A_*| \geqslant \frac{|A_*|^4}{\mathsf{E}(A_*^{-1}, A_*)} \geqslant \frac{|A_*|^4}{\mathsf{E}(A_*^{-1}, g\mathrm{B}_*)} \geqslant \frac{|A_*|^2}{q-1}$$

without any assumptions on generating properties of A.

5. On Zaremba's conjecture

In this section we apply methods of the proofs of Theorems 9, 13 to Zaremba's conjecture but also we use the specific of this problem, i.e., the special form of the correspondent set of matrices from $\mathrm{SL}_2(\mathbb{F}_p)$.

Denote by $F_M(Q)$ the set of all *rational* numbers $\frac{u}{v}$, $(u, v) = 1$ from $[0, 1]$ with all partial quotients in (1) not exceeding M and with $v \leqslant Q$:

$$F_M(Q) = \left\{ \frac{u}{v} = [0; b_1, \dots, b_s] : (u, v) = 1, 0 \le u \le v \le Q, b_1, \dots, b_s \le M \right\}.$$

By F_M denote the set of all *irrational* numbers from $[0, 1]$ with partial quotients less than or equal to M. From [Hensley 1992] we know that the Hausdorff dimension w_M of the set F_M satisfies

$$w_M = 1 - \frac{6}{\pi^2}\frac{1}{M} - \frac{72}{\pi^4}\frac{\log M}{M^2} + O\left(\frac{1}{M^2}\right), \quad M \to \infty, \tag{24}$$

however here we need a simpler result from [Hensley 1989], which states that

$$1 - w_M \asymp \frac{1}{M} \tag{25}$$

with absolute constants in the sign $\asymp$. Explicit estimates for dimensions of F_M for certain values of M can be found in [Jenkinson 2004; Jenkinson and Pollicott 2001] and in other papers. For example, see [Jenkinson and Pollicott 2001],

$$w_2 = 0.5312805062772051416244686\ldots$$

Hensley [1989; 1990] gave the bound

$$|F_M(Q)| \asymp_M Q^{2w_M}. \tag{26}$$

Now we are ready to prove Theorem 1 from the Introduction. One has

$$\begin{pmatrix} 0 & 1 \\ 1 & b_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & b_s \end{pmatrix} = \begin{pmatrix} p_{s-1} & p_s \\ q_{s-1} & q_s \end{pmatrix}, \tag{27}$$

where $p_s/q_s = [0; b_1, \ldots, b_s]$ and $p_{s-1}/q_{s-1} = [0; b_1, \ldots, b_{s-1}]$. It is clear that $p_{s-1}q_s - p_s q_{s-1} = (-1)^s$. Let $Q = p - 1$ and consider the set $F_M(Q)$. Any $u/v \in F_M(Q)$ corresponds to a matrix from (27) such that $b_j \leqslant M$. The set $F_M(Q)$ splits into ratios with even s and with odd s, in other words $F_M(Q) = F_M^{\text{even}}(Q) \bigsqcup F_M^{\text{odd}}(Q)$. Let $A \subseteq \mathrm{SL}_2(\mathbb{F}_p)$ be the set of matrices of the form above with even s. It is easy to see from (26), multiplying if it is needed the set $F_M^{\text{odd}}(Q)$ by $(01|1b)^{-1}$, $1 \leqslant b \leqslant M$ that $|F_M^{\text{even}}(Q)| \gg_M |F_M(Q)| \gg_M Q^{2w_M}$. It is easy to check that if for a certain n one has $A^n \cap \mathrm{B} \neq \varnothing$, then q_{s-1} from (27) equals zero modulo p and hence there is $u/v \in F_M((2p)^n)$ such that $v \equiv 0 \pmod p$. In a similar way, we can easily assume that for any $g = (ab|cd) \in A$ all entries a, b, c, d are nonzero (and hence by the construction they are nonzero modulo p); see, e.g., [Hensley 1994, p. 46] (the same paper [Hensley 1994] contains the fact that A is a generating subset of $\mathrm{SL}_2(\mathbb{F}_p)$). Analogously, we can suppose that all $g \in A$ are regular, that is, $\mathrm{tr}(g) \neq 0, \pm 2$. Let $K = |AAA|/|A|$ and $\widetilde{K} = |AA|/|A| = K^\alpha$, $0 \leqslant \alpha \leqslant 1$.

We need to estimate from below cardinality of the set of all possible traces of A, that is, cardinality of the set of sums $q_s + p_{s-1}$ (this expression is called "cyclical continuant"). Fix p_{s-1} and q_s. Then $p_{s-1}q_s - 1 = p_s q_{s-1}$ and thus p_s is a divisor of $p_{s-1}q_s - 1$. In particular, the number of such p_s is at most p^ε for any $\varepsilon > 0$. But now knowing the pair (p_s, q_s), we determine the correspondent matrix (27) from A uniquely. Hence the number of different pairs (p_{s-1}, q_s) is at least $\Omega_M(p^{-\varepsilon}|F_M(Q)|)$ and thus the number of different traces of all matrices from A is $\Omega_M(p^{-1-\varepsilon}|A|)$. This holds both in $\mathbb{Z}$ and in $\mathbb{F}_p$ because for any fixed $\lambda \in \mathbb{F}_p$ the equation $p_{s-1} + q_s \equiv \lambda \pmod p$ has at most p solutions. Actually, one can refine the term p^ε in $\Omega_M(p^{-1-\varepsilon}|A|)$ but is has no effect on the final bound and so below we just ignore it.

Now recall [Rudnev and Shkredov 2018, Lemma 12], which is a variant of the Helfgott map [2008] from [Murphy 2017] (we have already used similar arguments in the proof of Theorem 13). For the sake of the completeness we give the proof of a "statistical" version of this result.

Lemma 17. *Let $\mathbf{G}$ be any group and $A \subseteq \mathbf{G}$ be a finite set. Then for an arbitrary $g \in \mathbf{G}$, there is $A_0 \subseteq A$, $|A_0| \geqslant |A|/2$ such that for any $a_0 \in A_0$ the following holds*:

$$|A|/2 \leqslant |\mathrm{Conj}(g) \cap AgA^{-1}| \cdot |\mathrm{Centr}(g) \cap a_0^{-1}A|. \tag{28}$$

Here $\mathrm{Conj}(g)$ *is the conjugacy class and* $\mathrm{Centr}(g)$ *is the centralizer of g in $\mathbf{G}$.*

Proof. Let $\varphi : A \to \mathrm{Conj}(g) \cap AgA^{-1}$ be the Helfgott map $\varphi(a) := aga^{-1}$. One sees that $\varphi(a) = \varphi(b)$ if and only if

$$b^{-1}ag = gb^{-1}a.$$

In other words, $b^{-1}a \in \mathrm{Centr}(g) \cap A^{-1}A$. Clearly, then

$$\begin{aligned} |A| &= \sum_{c \in \mathrm{Conj}(g) \cap AgA^{-1}} |\{a \in A : \varphi(a) = c\}| \\ &\leqslant 2 \sum_{c \in \mathrm{Conj}(g) \cap AgA^{-1} : |\{a \in A : \varphi(a) = c\}| \geqslant |A|/(2|\mathrm{Conj}(g) \cap AgA^{-1}|)} |\{a \in A : \varphi(a) = c\}|. \end{aligned} \tag{29}$$

For $c \in \varphi(A) \subseteq \mathrm{Conj}(g) \cap AgA^{-1}$ put $A(c) = \varphi^{-1}(c) \subseteq A$ and let

$$A_0 = \bigsqcup_{c : |A(c)| \geqslant |A|/(2|\mathrm{Conj}(g) \cap AgA^{-1}|)} A(c).$$

In other words, estimate (29) gives us

$$|A_0| = \sum_c |A(c)| \geqslant |A|/2.$$

But for any $b \in A_0$ one has $|\mathrm{Centr}(g) \cap b^{-1}A| \geqslant |A|/(2|\mathrm{Conj}(g) \cap AgA^{-1}|)$ as required. This completes the proof of the lemma. □

Now summing inequality (28) over all $g \in A$ with different traces, we obtain in view of the Ruzsa triangle inequality that

$$\begin{aligned} |A|^2 p^{-1} &\ll_M |AAA^{-1}| \cdot \max_{g \in A} |\mathrm{Centr}(g) \cap a_0^{-1}(g)A| \\ &\leqslant K\widetilde{K}|A| \cdot \max_{g \in A} |\mathrm{Centr}(g) \cap a_0^{-1}(g)A|. \end{aligned} \tag{30}$$

Here for every $g \in A$ we have taken a concrete $a_0(g) \in A_0(g)$ but in view of Lemma 17 it is known that there are a lot of them and we will use this fact a little bit later. Now by [Helfgott 2008, Lemma 4.7], we see that

$$|(a_0^{-1}(g)A)g_*(a_0^{-1}(g)A)g_*^{-1}(a_0^{-1}(g)A)^{-1}| \gg |\mathrm{Centr}(g) \cap a_0^{-1}(g)A|^3,$$

where $g_* = (ab|cd)$ is any element from A such that $abcd \neq 0$ in the basis where g has the diagonal form. Thanks to Lemma 14 and Remark 16 we can choose $g_* = a_0(g)$, otherwise $|A| \ll p^{3/2}K^{5/2}$. In the last case if, say, $|A| \gg p^{2-1/35}$, then

$K \gg p^{33/175}$ and hence $|A^3| \gg p^{2+4/25}$. Using Theorem 9, we see that one can take $n = 27$ and this is better than we want to prove. Then with this choice of g_*, we have by the Ruzsa triangle inequality

$$|A^2 g_*^{-1} A^{-1}| \leqslant |A^2 A^{-2}| \leqslant K^2 |A|,$$

and hence $|\mathrm{Centr}(g) \cap a_0^{-1}(g) A| \ll K^{2/3} |A|^{1/3}$. Substituting the last bound into (30), we get

$$|A|^2 p^{-1} \ll_M K \widetilde{K} |A| \cdot K^{2/3} |A|^{1/3} \tag{31}$$

and hence

$$K \gg_M (|A|^2 p^{-3})^{1/(5+3\alpha)} \gg p^{4w_M/(5+3\alpha) - 3/(5+3\alpha)}. \tag{32}$$

In other words, $|AAA| \gg_M p^{2+(w_M(14+6\alpha)-13-6\alpha)/(5+3\alpha)}$. Take M sufficiently large such that $w_M(14+6\alpha) - 13 - 6\alpha > 0$. Using Theorem 9, we see that for any

$$n \geqslant \frac{w_M(28+12\alpha) - 6}{w_M(14+6\alpha) - 13 - 6\alpha} \tag{33}$$

one has $A^n \cap \mathrm{B} \neq \varnothing$. On the other hand, from (32), we get

$$|AA| = |A| K^\alpha \gg p^{2+(w_M(10+10\alpha)-10-9\alpha)/(5+3\alpha)}.$$

Suppose that $w_M(10+10\alpha) - 10 - 9\alpha > 0$. It can be done if $\alpha > 0$ and if we take sufficiently large M. Applying Theorem 9 one more time, we derive that for any

$$n \geqslant \frac{2}{3} \cdot \frac{w_M(20+20\alpha) - 6\alpha}{w_M(10+10\alpha) - 10 - 9\alpha} \tag{34}$$

one has $A^n \cap \mathrm{B} \neq \varnothing$. Comparing (33) and (34), we choose α optimally when

$$\alpha^2(120w_M^2 - 12w_M - 72) + \alpha(400w_M^2 - 368w_M + 6) + 280w_M^2 + 180 - 500w_M = 0$$

and it gives

$$18\alpha^2 + 19\alpha - 20 = 0$$

and whence $\alpha = \frac{1}{36}(-19 + \sqrt{1801}) + o_M(1)$ as $M \to +\infty$. Hence from (33), say, we obtain $n \geqslant \frac{1}{3}(47 + \sqrt{1801}) + o_M(1) > 29.81 + o_M(1)$. Taking sufficiently large M, we can choose $n = 30$. If $\alpha = 0$, then for sufficiently large M estimate (33) allows us to take $n = 23$. This completes the proof. □

Combining the arguments above with Theorems 9, 13, we obtain Theorem 2 from the Introduction. Actually, if we apply the second part of Theorem 13, then we generate the whole $\mathrm{SL}_2(\mathbb{F}_p)$ (and this differs our method from [Magee et al. 2019], say). Because in the case $\mathfrak{k} = 2$ we use results about growth in $\mathrm{SL}_2(\mathbb{F}_p)$ for relatively small asymmetric set A ($|A| \gg p^{2w_2} \gg p^{1.062}$) our absolute constant C is large. It is easy to see that the arguments of this section on trace of the set A

begin to work for $w_M > \frac{3}{4}$ (see Lemma 14, as well as estimates (30), (31)) and in this case the constant C can be decreased, although it remains rather large.

Acknowledgement

We thank I.D. Kan for useful discussions and remarks.

References

[Bourgain and Kontorovich 2011] J. Bourgain and A. Kontorovich, "On Zaremba's conjecture", *C. R. Math. Acad. Sci. Paris* **349**:9-10 (2011), 493–495. MR Zbl

[Bourgain and Kontorovich 2014] J. Bourgain and A. Kontorovich, "On Zaremba's conjecture", *Ann. of Math.* (2) **180**:1 (2014), 137–196. MR Zbl

[Frobenius 1896] G. Frobenius, "Über Gruppencharaktere", *Sitzungsber. Preuss. Akad. Wiss. Berlin* **1896** (1896), 985–1021. Zbl

[Frolenkov and Kan 2014] D. A. Frolenkov and I. D. Kan, "A strengthening of a theorem of Bourgain–Kontorovich, II", *Mosc. J. Comb. Number Theory* **4**:1 (2014), 78–117. MR Zbl

[Gowers 2008] W. T. Gowers, "Quasirandom groups", *Combin. Probab. Comput.* **17**:3 (2008), 363–387. MR Zbl

[Helfgott 2008] H. A. Helfgott, "Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$", *Ann. of Math.* (2) **167**:2 (2008), 601–623. MR Zbl

[Hensley 1989] D. Hensley, "The distribution of badly approximable numbers and continuants with bounded digits", pp. 371–385 in *Théorie des nombres* (Quebec, 1987), edited by J.-M. De Koninck and C. Levesque, de Gruyter, Berlin, 1989. MR Zbl

[Hensley 1990] D. Hensley, "The distribution of badly approximable rationals and continuants with bounded digits, II", *J. Number Theory* **34**:3 (1990), 293–334. MR Zbl

[Hensley 1992] D. Hensley, "Continued fraction Cantor sets, Hausdorff dimension, and functional analysis", *J. Number Theory* **40**:3 (1992), 336–358. MR Zbl

[Hensley 1994] D. Hensley, "The distribution mod n of fractions with bounded partial quotients", *Pacific J. Math.* **166**:1 (1994), 43–54. MR Zbl

[Hensley 1996] D. Hensley, "A polynomial time algorithm for the Hausdorff dimension of continued fraction Cantor sets", *J. Number Theory* **58**:1 (1996), 9–45. MR Zbl

[Jenkinson 2004] O. Jenkinson, "On the density of Hausdorff dimensions of bounded type continued fraction sets: the Texan conjecture", *Stoch. Dyn.* **4**:1 (2004), 63–76. MR Zbl

[Jenkinson and Pollicott 2001] O. Jenkinson and M. Pollicott, "Computing the dimension of dynamically defined sets: E_2 and bounded continued fractions", *Ergodic Theory Dynam. Systems* **21**:5 (2001), 1429–1445. MR Zbl

[Kan 2016] I. D. Kan, "A strengthening of a theorem of Bourgain and Kontorovich, IV", *Izv. Ross. Akad. Nauk Ser. Mat.* **80**:6 (2016), 103–126. In Russian; translated in *Izv. Math.* **80**:6 (2016), 1094–1117. MR Zbl

[Kontorovich 2013] A. Kontorovich, "From Apollonius to Zaremba: local-global phenomena in thin orbits", *Bull. Amer. Math. Soc.* (*N.S.*) **50**:2 (2013), 187–228. MR Zbl

[Korobov 1963] N. M. Korobov, Теоретико-числовые методы в приближенном анализе, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1963. MR

[Magee et al. 2014] M. Magee, H. Oh, and D. Winter, "Expanding maps and continued fractions", preprint, 2014. arXiv

[Magee et al. 2019] M. Magee, H. Oh, and D. Winter, "Uniform congruence counting for Schottky semigroups in $SL_2(\mathbb{Z})$", *J. Reine Angew. Math.* **753** (2019), 89–135. MR Zbl

[Moshchevitin 2007] N. G. Moshchevitin, "Sets of the form $\mathscr{A}+\mathscr{B}$ and finite continued fractions", *Mat. Sb.* **198**:4 (2007), 95–116. In Russian; translated in *Sb. Math.* **198**:3-4 (2007), 537–557. MR Zbl

[Moshchevitin et al. 2020] N. G. Moshchevitin, B. Murphy, and I. Shkredov, "Popular products and continued fractions", *Israel J. Math.* **238**:2 (2020), 807–835. MR Zbl

[Murphy 2017] B. Murphy, "Upper and lower bounds for rich lines in grids", 2017. To appear in *Amer. J. Math.* arXiv

[Naimark 2010] M. A. Naimark, Теориja представлениj групп, Fizmatlit, Moscow, 2010.

[Niederreiter 1986] H. Niederreiter, "Dyadic fractions with small partial quotients", *Monatsh. Math.* **101**:4 (1986), 309–315. MR Zbl

[Nikolov and Pyber 2011] N. Nikolov and L. Pyber, "Product decompositions of quasirandom groups and a Jordan type theorem", *J. Eur. Math. Soc.* **13**:4 (2011), 1063–1077. MR Zbl

[Rudnev and Shkredov 2018] M. Rudnev and I. D. Shkredov, "On growth rate in $SL_2(\mathbb{F}_p)$, the affine group and sum-product type implications", preprint, 2018. arXiv

[Ruzsa 1996] I. Z. Ruzsa, "Sums of finite sets", pp. 281–293 in *Number theory* (New York, 1991-1995), edited by D. V. Chudnovsky et al., Springer, 1996. MR Zbl

[Sarnak and Xue 1991] P. Sarnak and X. X. Xue, "Bounds for multiplicities of automorphic representations", *Duke Math. J.* **64**:1 (1991), 207–227. MR Zbl

[Serre 1967] J.-P. Serre, *Représentations linéaires des groupes finis*, Hermann, Paris, 1967. MR Zbl

[Shkredov 2018] I. D. Shkredov, "On asymptotic formulae in some sum-product questions", *Tr. Mosk. Mat. Obs.* **79**:2 (2018), 271–334. In Russian; translated in *Trans. Moscow Math. Soc.* **79** (2018), 231–281.

[Zaremba 1972] S. K. Zaremba, "La méthode des 'bons treillis' pour le calcul des intégrales multiples", pp. 39–119 in *Applications of number theory to numerical analysis* (Montreal, 1971), edited by S. K. Zaremba, Academic, New York, 1972. MR Zbl

Received November 18, 2019. Revised August 12, 2020.

NIKOLAY G. MOSHCHEVITIN
STEKLOV MATHEMATICAL INSTITUTE
MOSCOW
RUSSIA
moshchevitin@gmail.com

ILYA D. SHKREDOV
STEKLOV MATHEMATICAL INSTITUTE
MOSCOW
RUSSIA
ilya.shkredov@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.213

THE FIRST NONZERO EIGENVALUE OF THE p-LAPLACIAN ON DIFFERENTIAL FORMS

SHOO SETO

We introduce a generalization of the p-Laplace operator to act on differential forms and generalize an estimate of Gallot and Meyer (1973) for the first nonzero eigenvalue on closed Riemannian manifolds.

1. Introduction

Let (M, g) be an n-dimensional closed Riemannian manifold. Motivated from the variational characterization of the Laplacian eigenvalue problem, we define the L^p-Dirichlet integral on k-forms (introduced in [Scott 1995]) by

(1) $$\mathcal{F}[\alpha] := \int_M \|d\alpha\|^p + \|d^*\alpha\|^p, \quad \alpha \in \Omega^k(M),$$

where d^* is the L^2-adjoint of the exterior derivative d. Note that $\mathcal{F}[\alpha] = 0$ if and only if $\alpha \in \mathcal{H}^k(M)$, that is, the minimum is zero and is attained for harmonic k-forms, i.e., $\alpha \in \ker(d) \cap \ker(d^*)$. For a nonzero infimum we consider the space

(2) $$A_k := \left\{ \alpha \in \mathcal{W}^{1,p}(\Omega^k(M)) \;\middle|\; \fint_M \|\alpha\|^p = 1,\ \int_M \|\alpha\|^{p-2}\langle \alpha, \omega\rangle = 0,\ \omega \in \mathcal{H}^k(M) \right\},$$

where the space $\mathcal{W}^{1,p}(\Omega^k(M))$ is the $(1, p)$-Sobolev space of differential k-forms defined in [Scott 1995]. See Section 3 for the precise definition. Computing the Euler–Lagrange equation leads us to the defining the following operator:

Definition 1.1 (p-Hodge Laplacian).

(3) $$\Delta_p \alpha := d^*(\|d\alpha\|^{p-2} d\alpha) + d(\|d^*\alpha\|^{p-2} d^*\alpha), \quad \alpha \in \Omega^k(M).$$

When $p = 2$, this becomes the usual Hodge Laplacian. For $p \neq 2$ and $\alpha \in C^\infty(M)$ Δ_p becomes the usual p-Laplacian. The corresponding eigenvalue equation is given by

(4) $$\Delta_p \alpha = \lambda \|\alpha\|^{p-2} \alpha, \quad \alpha \in \Omega^k(M)$$

MSC2010: 47J10, 53C65.
Keywords: differential forms, Hodge Laplacian, p-Laplacian, Weitzenböck curvature.

and the variational principle tells us that

$$\lambda_1 = \inf\{\mathcal{F}[\alpha] \mid \alpha \in A_k\}.$$

See Section 3 for details. When $p = 2$, there is much work on the spectrum of the Hodge Laplacian acting on differential forms. Among many others, we point out the work of Gallot and Meyer [1973; 1975] who show an estimate of the first eigenvalue using bounds from the Weitzenböck curvature on compact Riemannian manifolds. For manifolds with boundary, among many others, see the works [Kwong 2016; Savo 2009; Raulot and Savo 2011].

For $p \neq 2$, the p-Laplace eigenvalue problem on 0-forms (functions) has attracted much attention. See notes by Lindqvist [2006] for a general reference on the p-Laplace equation. For estimates on the first eigenvalue relating to the curvature, among many other works, see [Matei 2000; Naber and Valtorta 2014; Seto and Wei 2017] for eigenvalue estimates with $\mathrm{Ric} \geq K,\ K \in \mathbb{R}$.

Remark 1.2. There is also a related notion of p-harmonic k-forms which looks at the minimizer in a cohomology class of k-forms with finite L^p-norm, i.e.,

$$\inf_{\alpha \in H_d^k(M)} \int_M \|\alpha\|^p.$$

The critical point of the variation leads to the following definition of p-harmonic, for closed k-forms α, if

$$d_p^* := d^*(\|\alpha\|^{p-2}\alpha) = 0$$

then α is p-harmonic. See [Dung 2017].

In this paper we prove the following lower bound estimate for the first eigenvalue:

Theorem 1.3. *Let M^n be a closed Riemannian manifold with the eigenvalues of the curvature operator bounded below by $H \in \mathbb{R}$ and $p \geq 2$. Then*

$$\lambda_1 \geq \left(\frac{k(n-k)}{2^{(2/p)-1}(C + (p-2)/2)} H \right)^{p/2},$$

where

$$C = \max\left\{ \frac{k}{k+1}, \frac{n-k}{n-k+1} \right\}.$$

Remark 1.4. When $p = 2$, the above recovers the estimate due to Gallot and Meyer [1973] (see also [Gallot and Meyer 1975]), for $1 \leq k \leq \frac{n}{2}$,

$$\lambda_1 \geq k(n-k+1)H.$$

The organization of this paper is as follows. In Section 2 we review some known estimates for differential k-forms. In Section 3 we show that the infimum can be characterized as an eigenvalue problem. In Section 4 we give the main estimate. In

Section 5 we give a brief discussion on boundary conditions for differential forms and possible future directions.

2. Some estimates on $\Omega^k(M)$

We first recall the Weitzenböck curvature

Definition 2.1. Let $p \in M$ and let $\{E_i\}_{i=1}^n$ be an orthonormal frame at p. Then for $\alpha \in \Omega^k(M)$, define the Weitzenböck curvature W_k by

$$W_k(\alpha)(X_1, \dots, X_k) := \sum (R(E_j, X_i)\alpha)(X_1, \dots, E_j, \dots, X_k).$$

Note that on 1-forms, this is simply the Ricci tensor.

If the eigenvalues of the curvature operator are bounded by $H \in \mathbb{R}$, we can show that

$$(W_k(\alpha), \alpha) \geq k(n-k)H\|\alpha\|^2. \tag{5}$$

The Weitzenböck curvature appears in the main tool we use in obtaining our estimate, which is the Bochner–Weitzenböck formula for k-forms:

$$\tfrac{1}{2}\Delta\|\alpha\|^2 = (\Delta\alpha, \alpha) - \|\nabla\alpha\|^2 - (W_k(\alpha), \alpha), \tag{6}$$

where $\Delta := \Delta_2 = dd^* + d^*d$. Note that for exact 1-form $\alpha = df$, since $\nabla df = \operatorname{Hess} f$, the usual Cauchy–Schwarz inequality will give us an estimate on the middle term. For k-forms, we will need the following:

Lemma 2.2 [Gallot and Meyer 1973]. *Let* $\alpha \in \Omega^k(M)$, $1 \leq k \leq n-1$. *Then*

$$\|\nabla\alpha\|^2 \geq \frac{1}{k+1}\|d\alpha\|^2 + \frac{1}{n-k+1}\|d^*\alpha\|^2. \tag{7}$$

We give a proof for completeness. The proof we give is in the context of conformal Killing forms and can be found in various sources, for instance, [Moroianu and Semmelmann 2003].

Proof. Consider the two linear maps

$$\iota : TM \otimes \Omega^k(M) \to \Omega^{k-1}(M)$$
$$\iota(v, \alpha) = \iota_v\alpha$$

and

$$\wedge : \Omega^1(M) \otimes \Omega^k(M) \to \Omega^{k+1}(M)$$
$$\wedge(\beta, \alpha) = \beta \wedge \alpha.$$

Let ι^* and $\wedge^*$ be their metric adjoint. Then

$$\wedge \circ \iota^*(\alpha) = 0 \quad \text{and} \quad \iota \circ \wedge^*(\alpha) = 0,$$

so that we get the decomposition

$$TM \otimes \Omega^k(M) \simeq \mathrm{im}(\iota^*) \oplus \mathrm{im}(\wedge^*) \oplus Y,$$

where Y is the orthogonal complement. By direct computation, for $\alpha \in \Omega^k(M)$, we have

$$\iota \circ \iota^*(\alpha) = (n-k+1)\alpha \quad \text{and} \quad \wedge \circ \wedge^*(\alpha) = (k+1)\alpha.$$

Viewing $\nabla\alpha \in \Gamma(TM \otimes \Omega^k(M))$, From the decomposition,

$$\nabla\alpha = \iota^*\beta + \wedge^*\gamma + \delta,$$

applying ι, we have

$$\iota\nabla\alpha = (n-k+1)\beta.$$

So the projection operator onto $\mathrm{im}(\iota^*)$ is given by

$$\pi_{\iota^*}\nabla\alpha = \frac{1}{n-k+1}\iota^*\iota\nabla\alpha$$

and similarly

$$\pi_{\wedge^*}\nabla\alpha = \frac{1}{k+1}\wedge^*\wedge\nabla\alpha.$$

Let $T\alpha := \pi_T(\nabla\alpha)$ the projection onto the orthogonal complement space. Since

$$d\alpha = \wedge(\nabla\alpha) \quad \text{and} \quad d^*\alpha = -\iota(\nabla\alpha),$$

we have the decomposition

$$T\alpha(X) = \nabla_X\alpha - \frac{1}{k+1}\iota_X d\alpha + \frac{1}{n-k+1}X^* \wedge d^*\alpha$$

and taking the norm gives us

$$\|\nabla\alpha\|^2 = \|T\alpha\|^2 + \frac{1}{k+1}\|d\alpha\|^2 + \frac{1}{n-k+1}\|d^*\alpha\|^2,$$

which implies (7). □

Remark 2.3. The projection operator T defined above is called the twistor operator and a form $\alpha \in \Omega^k(M)$ is called a conformal Killing form if $T\alpha = 0$.

The following lemma was pointed out by N.T. Dung and gives us a way to control the interior product by using an orthogonal decomposition of forms as the image under an interior product.

Lemma 2.4 [Dung and Sung 2019, Lemma 3.5]. *Let $V \in TM$, $\alpha \in \Omega^{k+1}$, $\beta \in \Omega^k$. Then*

$$|\langle \iota_V\alpha, \beta\rangle| \le \|V\|\|\alpha\|\|\beta\|.$$

3. Variational characterization of the eigenvalue

In this section we will compute the Euler–Lagrange equation of (1) and show that the extremal problem can be reformulated as an eigenvalue problem. Analogous to the 0-form (function) case, we will look at weak solutions lying the $(1, p)$-Sobolev space of differential k-forms first defined by Scott [1995] as

$$\mathcal{W}^{1,p}(\Omega^k(M)) := \left\{\alpha \in W(\Omega^k(M)) \mid \alpha, d\alpha, d^*\alpha \in L^p(\Omega^*(M))\right\},$$

where $W(\Omega^k(M))$ is the classical Sobolev space of k-forms, i.e., α is locally integrable and admits a generalized gradient.

Definition 3.1. We say that λ is an eigenvalue, if there exists a k-form $\alpha \in \mathcal{W}^{1,p}(\Omega^k(M))$ such that

$$\int_M \|d\alpha\|^{p-2}\langle d\alpha, d\beta\rangle + \int_M \|d^*\alpha\|^{p-2}\langle d^*\alpha, d^*\beta\rangle = \lambda \int_M \|\alpha\|^{p-2}\langle \alpha, \beta\rangle,$$

for any $\beta \in C^\infty(\Omega^k(M))$.

We will show the first nonzero eigenvalue λ_1 can be characterized as the infimum of the L^p-Dirichlet energy over the space A_k given in (2).

Proposition 3.2. *For closed manifolds M and $p \geq 2$,*

$$\lambda_1 = \inf\left\{\int_M \|d\alpha\|^p + \|d^*\alpha\|^p \;\middle|\; \alpha \in A_k\right\}.$$

Proof. Let ω be a fixed harmonic form and let $\beta(t) \in A$ for small $t > 0$ such that $\beta(0) = \alpha$. Computing the first variation of (1), we have

$$\begin{aligned}\frac{d}{dt}\mathcal{F}[\beta(t)]\bigg|_{t=0} &= p\int_M \|d\alpha\|^{p-2}\langle d\alpha, d\beta'(0)\rangle + \|d^*\alpha\|^{p-2}\langle d^*\alpha, d^*\beta'(0)\rangle \\ &= p\int_M \langle \Delta_p\alpha, \beta'(0)\rangle.\end{aligned}$$

Next we compute the variation of the constraints so that

$$\frac{d}{dt}\int_M \|\beta\|^p\bigg|_{t=0} = p\int_M |\alpha|^{p-2}\langle \alpha, \beta'(0)\rangle$$

and

$$\frac{d}{dt}\int_M \|\beta\|^{p-2}\langle \beta, \omega\rangle\bigg|_{t=0} = (p-2)\int_M \|\alpha\|^{p-4}\langle \alpha, \beta'(0)\rangle\langle \alpha, \omega\rangle + \|\alpha\|^{p-2}\langle \beta'(0), \omega\rangle.$$

By the Lagrange multiplier method, there must be some λ and μ such that for $\beta \in \Omega^k(M)$,

$$\int_M \langle \Delta_p\alpha, \beta\rangle = \lambda\int_M \|\alpha\|^{p-2}\langle \alpha, \beta\rangle + \mu\int_M \|\alpha\|^{p-4}\langle \alpha, \beta\rangle\langle \alpha, \omega\rangle + \|\alpha\|^{p-2}\langle \beta, \omega\rangle.$$

Setting $\beta = \omega$, we have

$$0 = \mu \int_M \|\alpha\|^{p-4}\langle\alpha, \omega\rangle^2 + \|\alpha\|^{p-2}\|\omega\|^2$$

so that $\mu = 0$. Therefore,

$$\Delta_p \alpha = \lambda \|\alpha\|^{p-2}\alpha. \qquad \square$$

4. Proof of Theorem 1.3

We will consider the following integral

$$\int_M \langle \Delta_p \alpha, \Delta\alpha\rangle = \int_M \langle \Delta_p \alpha, dd^*\alpha\rangle + \int_M \langle \Delta_p, d^*d\alpha\rangle.$$

Let $\alpha \in \Omega^k(M)$ be an eigenform satisfying (4). Then

$$\begin{aligned}(8)\qquad \int_M \langle \Delta_p \alpha, d^*d\alpha\rangle &= \lambda \int_M \|\alpha\|^{p-2}\langle\alpha, d^*d\alpha\rangle \\ &= \lambda \int_M \big\langle d(\|\alpha\|^{p-2}\alpha), d\alpha\big\rangle \\ &= \lambda \int_M \big\langle d(\|\alpha\|^{p-2}) \wedge \alpha, d\alpha\big\rangle + \lambda \int_M \|\alpha\|^{p-2}\|d\alpha\|^2\end{aligned}$$

and

$$\begin{aligned}(9)\qquad \int_M \langle \Delta_p \alpha, dd^*\alpha\rangle &= \lambda \int_M \|\alpha\|^{p-2}\langle\alpha, dd^*\alpha\rangle \\ &= \lambda \int_M \big\langle d^*(\|\alpha\|^{p-2}\alpha), d^*\alpha\big\rangle \\ &= \lambda \int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2 - \lambda \int_M \big\langle \iota_{\nabla\|\alpha\|^{p-2}}\alpha, d^*\alpha\big\rangle.\end{aligned}$$

On the other hand, by using the Bochner–Weitzenböck formula (6) we have

$$\begin{aligned}(10)\qquad &\int_M \langle \Delta_p \alpha, \Delta\alpha\rangle \\ &= \lambda \int_M \|\alpha\|^{p-2}\langle\alpha, \Delta\alpha\rangle \\ &= \lambda \int_M \big((p-2)\|\alpha\|^{p-2}|\nabla\|\alpha\||^2 + \|\alpha\|^{p-2}\|\nabla\alpha\|^2 + \|\alpha\|^{p-2}(W_k(\alpha), \alpha)\big).\end{aligned}$$

Combining (8), (9), and (10), we obtain

$$\begin{aligned}(11)\qquad \int_M \big\langle d(\|\alpha\|^{p-2}) \wedge \alpha, d\alpha\big\rangle - \int_M \big\langle \iota_{\nabla\|\alpha\|^{p-2}}\alpha, d^*\alpha\big\rangle + \int_M \|\alpha\|^{p-2}\|d\alpha\|^2 \\ + \int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2\end{aligned}$$

$$\text{(11 cont.)} \quad = \int_M \big((p-2)\|\alpha\|^{p-2}|\nabla\|\alpha\||^2 + \|\alpha\|^{p-2}\|\nabla\alpha\|^2 + \|\alpha\|^{p-2}(W_k(\alpha), \alpha)\big).$$

Using Lemma 2.4, the first term of (11) can be estimated as

$$\begin{aligned}
\int_M \langle d(\|\alpha\|^{p-2}) \wedge \alpha, d\alpha\rangle
&= \int_M \langle \alpha, \iota_{\nabla\|\alpha\|^{p-2}}(d\alpha)\rangle \\
&\le \int_M \big\|\nabla\|\alpha\|^{p-2}\big\| \|d\alpha\| \|\alpha\| \\
&= (p-2)\int_M \|\alpha\|^{(p-2)/2}\big\|\nabla\|\alpha\|\big\| \|\alpha\|^{(p-2)/2}\|d\alpha\| \\
&\le \frac{(p-2)}{2}\int_M \|\alpha\|^{p-2}\big\|\nabla\|\alpha\|\big\|^2 + \frac{(p-2)}{2}\int_M \|\alpha\|^{p-2}\|d\alpha\|^2
\end{aligned}$$

and similarly for the second term,

$$\begin{aligned}
-\int_M \langle \iota_{\nabla\|\alpha\|^{p-2}}\alpha, d^*\alpha\rangle &\le \int_M \|\nabla\|\alpha\|^{p-2}\|\alpha\|\|d^*\alpha\| \\
&= (p-2)\int_M \|\alpha\|^{(p-2)/2}\big\|\nabla\|\alpha\|\big\| \|\alpha\|^{(p-2)/2}\|d^*\alpha\| \\
&\le \frac{(p-2)}{2}\int_M \|\alpha\|^{p-2}\big\|\nabla\|\alpha\|\big\|^2 + \frac{(p-2)}{2}\int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2.
\end{aligned}$$

Applying these estimates to (11), we get

$$\begin{aligned}
&\frac{(p-2)+2}{2}\int_M \|\alpha\|^{p-2}\|d\alpha\|^2 + \frac{(p-2)+2}{2}\int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2 \\
&\ge \int_M \|\alpha\|^{p-2}\|\nabla\alpha\|^2 + \int_M \|\alpha\|^{p-2}(W_k(\alpha), \alpha) \\
&\ge \frac{1}{k+1}\int_M \|\alpha\|^{p-2}\|d\alpha\|^2 + \frac{1}{n-k+1}\int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2 + \int_M \|\alpha\|^{p-2}(W_k(\alpha), \alpha).
\end{aligned}$$

Let

$$C := \max\Big\{\frac{k}{k+1}, \frac{n-k}{n-k+1}\Big\}.$$

Using

$$\int_M \|\alpha\|^{p-2}\|d\alpha\|^2 \le \Big(\int_M \|\alpha\|^p\Big)^{1-2/p}\Big(\int_M \|d\alpha\|^p\Big)^{2/p}$$

and

$$\int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2 \le \Big(\int_M \|\alpha\|^p\Big)^{1-2/p}\Big(\int_M \|d^*\alpha\|^p\Big)^{2/p},$$

we have

$$\left(C+\frac{(p-2)}{2}\right)\left(\int_M \|\alpha\|^p\right)^{1-2/p}\left[\left(\int_M \|d\alpha\|^p\right)^{2/p}+\left(\int_M \|d^*\alpha\|^p\right)^{2/p}\right] \geq \int_M \|\alpha\|^{p-2}(W_k(\alpha),\alpha).$$

For $p \geq 2$, and using the lower bound of the Weitzenböck curvature (5), we have

$$2^{(2/p)-1}\left(C+\frac{(p-2)}{2}\right)\left(\int_M \|\alpha\|^p\right)^{1-2/p}\left(\int_M \|d\alpha\|^p+\|d^*\alpha\|^p\right)^{2/p} \geq k(n-k)H\int_M \|\alpha\|^p.$$

Using the fact that $\int_M \|d\alpha\|^p + \|d^*\alpha\|^p = \lambda \int_M \|\alpha\|^p$ for eigenform α, we get

$$\lambda^{2/p} \geq \frac{k(n-k)}{2^{(2/p)-1}(C+(p-2)/2)}.$$

5. Boundary conditions

In this section we briefly discuss the situation of a compact manifold M with nonempty smooth boundary ∂M. Let n denote the unit outer normal vector and let $J : \partial M \to M$ be the inclusion. Then $J^*\alpha$ is the restriction of a form to the boundary. Then d and its adjoint d^* are related with an additional boundary term given by

$$\int_M \langle d\alpha,\beta\rangle = \int_M \langle \alpha, d^*\beta\rangle + \int_{\partial M}\langle J^*(\alpha), \iota_n\beta\rangle, \quad \alpha\in\Omega^k(M),\ \beta\in\Omega^{k+1}(M).$$

and the corresponding Green's formula for the p-Laplacian is

$$(\Delta_p\alpha,\beta) = \int_M \|d\alpha\|^{p-2}\langle d\alpha, d\beta\rangle + \int_M \|d^*\alpha\|^{p-2}\langle d^*\alpha, d^*\beta\rangle - \int_{\partial M}\big\langle \iota_n(\|d\alpha\|^{p-2}d\alpha), J^*(\beta)\big\rangle + \int_{\partial M}\big\langle \|d^*\alpha\|^{p-2}J^*(d^*\alpha), \iota_n\beta\big\rangle.$$

The two most common boundary conditions for the classical Laplacian eigenvalue problem are the Dirichlet and Neumann boundary condition. For the Hodge Laplacian, the analogous boundary conditions are the absolute boundary condition

$$\begin{cases}\iota_n\alpha = 0, \\ \iota_n\, d\alpha = 0, \quad \text{on } \partial M\end{cases}$$

and the relative boundary condition

$$\begin{cases}J^*(\alpha) = 0, \\ J^*(d^*\alpha) = 0, \quad \text{on } \partial M.\end{cases}$$

The essential feature of the boundary condition is that if α satisfies either of the boundary conditions, then $\Delta_p \alpha = 0$ implies $d\alpha = 0$ and $d^*\alpha = 0$. The boundary terms that will be introduced to (11) are

$$\begin{aligned}\int_M \langle d(\|\alpha\|^{p-2}) \wedge \alpha, d\alpha\rangle - \int_M \langle \iota_{\nabla\|\alpha\|^{p-2}}\alpha, d^*\alpha\rangle + \int_M \|\alpha\|^{p-2}\|d\alpha\|^2 \\ + \int_M \|\alpha\|^{p-2}\|d^*\alpha\|^2 - \int_{\partial M} \|\alpha\|^{p-2}\langle J^*(\alpha), \iota_n(d\alpha)\rangle + \int_{\partial M} \|\alpha\|^{p-2}\langle J^*(d^*\alpha), \iota_n(\alpha)\rangle \\ = \int_M \big((p-2)\|\alpha\|^{p-2}|\nabla\|\alpha\||^2 + \|\alpha\|^{p-2}\|\nabla\alpha\|^2 + \|\alpha\|^{p-2}(W_k(\alpha), \alpha)\big).\end{aligned}$$

Since the boundary terms will vanish under either of the boundary conditions, we get the same estimate for the boundary value problem as well. It would be interesting to see what the Reilly formula, for instance a generalization of Theorem 3 in [Raulot and Savo 2011] would be in this context, however due to the asymmetry of the weight function in the p-Laplacian, it is not immediate what the appropriate Bochner–Weitzenböck type formula would be for Δ_p.

Acknowledgements

The author would like to thank Prof. Zhiqin Lu for very helpful discussions, suggestions and constant support, and Prof. Guofang Wei and Prof. Qi Zhang for discussions and encouragement on the problem. We also thank Prof. Nguyen Thac Dung for letting us know of an improvement on the constant, see Lemma 2.4.

References

[Dung 2017] N. T. Dung, "p-harmonic ℓ-forms on Riemannian manifolds with a weighted Poincaré inequality", *Nonlinear Anal.* **150** (2017), 138–150. MR Zbl

[Dung and Sung 2019] N. T. Dung and C.-J. A. Sung, "Analysis of weighted p-harmonic forms and applications", *Int. J. Math.* **30**:11 (2019), art. id. 1950058. MR Zbl

[Gallot and Meyer 1973] S. Gallot and D. Meyer, "Sur la première valeur propre du p-spectre pour les variétés à opérateur de courbure positif", *C. R. Acad. Sci. Paris Sér. A-B* **276** (1973), A1619–A1621. MR Zbl

[Gallot and Meyer 1975] S. Gallot and D. Meyer, "Opérateur de courbure et laplacien des formes différentielles d'une variété riemannienne", *J. Math. Pures Appl.* (9) **54**:3 (1975), 259–284. MR Zbl

[Kwong 2016] K.-K. Kwong, "Some sharp Hodge Laplacian and Steklov eigenvalue estimates for differential forms", *Calc. Var. Partial Differential Equations* **55**:2 (2016), art. id. 38. MR Zbl

[Lindqvist 2006] P. Lindqvist, *Notes on the p-Laplace equation*, Report. Univ. Jyväskylä Dept. Math. Stat. **102**, University of Jyväskylä, 2006. MR Zbl

[Matei 2000] A.-M. Matei, "First eigenvalue for the p-Laplace operator", *Nonlinear Anal.* **39**:8 (2000), 1051–1068. MR Zbl

[Moroianu and Semmelmann 2003] A. Moroianu and U. Semmelmann, "Twistor forms on Kähler manifolds", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **2**:4 (2003), 823–845. MR Zbl

[Naber and Valtorta 2014] A. Naber and D. Valtorta, "Sharp estimates on the first eigenvalue of the p-Laplacian with negative Ricci lower bound", *Math. Z.* **277**:3-4 (2014), 867–891. MR Zbl

[Raulot and Savo 2011] S. Raulot and A. Savo, "A Reilly formula and eigenvalue estimates for differential forms", *J. Geom. Anal.* **21**:3 (2011), 620–640. MR Zbl

[Savo 2009] A. Savo, "On the lowest eigenvalue of the Hodge Laplacian on compact, negatively curved domains", *Ann. Global Anal. Geom.* **35**:1 (2009), 39–62. MR Zbl

[Scott 1995] C. Scott, "L^p theory of differential forms on manifolds", *Trans. Amer. Math. Soc.* **347**:6 (1995), 2075–2096. MR Zbl

[Seto and Wei 2017] S. Seto and G. Wei, "First eigenvalue of the p-Laplacian under integral curvature condition", *Nonlinear Anal.* **163** (2017), 60–70. MR Zbl

Received April 2, 2019. Revised December 19, 2019.

SHOO SETO
DEPARTMENT OF MATHEMATICS
CALIFORNIA STATE UNIVERSITY, FULLERTON
FULLERTON, CA
UNITED STATES
shoseto@fullerton.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 309, No. 1, 2020
https://doi.org/10.2140/pjm.2020.309.223

GLOBAL REGULARITY OF THE NAVIER–STOKES EQUATIONS ON 3D PERIODIC THIN DOMAIN WITH LARGE DATA

NA ZHAO

We consider the Navier–Stokes equations on a 3D periodic thin domain $T_\epsilon = (0, l_1) \times (0, l_2) \times (0, \epsilon)$. We show that there exists an absolute (large) constant C such that for any $C^* > 0$ which can be arbitrarily large, there exists an $\epsilon_0 > 0$ such that the Navier–Stokes equations are globally well-posed for a class of large initial data satisfying

$$\|\partial_h u_0\|_{L^2(T_\epsilon)} \leqslant \frac{C^*}{\epsilon^{\frac{1}{2}} |\ln \epsilon|^{\frac{3}{2}}}, \quad \|\partial_3 u_0\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{1}{2}}},$$

where $\partial_h = (\partial_1, \partial_2)$ and $0 < \epsilon \leqslant \epsilon_0$. This improves the result of Kukavica and Ziane (*Journal of Differential Equations* 234:(2) (2007), 485–506), where the initial data u_0 is required to satisfy

$$\|\nabla u_0\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{1}{2}} |\ln \epsilon|^{\frac{3}{2}}}.$$

1. Introduction

The Navier–Stokes equations describe the time evolution of solutions of mathematical models of viscous incompressible fluids. The research of solutions has attracted many experts. To our knowledge, in the whole space case, Leray [1934] proved that if the divergence-free initial data u_0 belongs to L^2, there exists a weak solution $u(t)$ which is defined for all $t \geqslant 0$ and satisfies a global energy inequality. Hopf [1951] extended the result to the bounded domain case. Furthermore, if the initial data possesses certain regularity, say $u_0 \in H^1(\Omega)$, where Ω is a smooth bounded or periodic domain, then the Leray solution is smooth and unique at least for some short time interval; see [Temam 1984].

The author is supported by China Postdoctoral Science Foundation (Grants No. 2019TQ0042 and 2020M680457).
MSC2010: 35Q30, 76D05, 76N10.
Keywords: Navier–Stokes equations, thin domain.

In this paper, we consider the Navier–Stokes equations of the incompressible fluid flow on a periodic domain T_ϵ,

$$\begin{cases} \partial_t u - \Delta u + u \cdot \nabla u + \nabla p = 0, \\ \nabla \cdot u = 0, \end{cases} \tag{1-1}$$

where u and p denote the velocity and the pressure, respectively, and T_ϵ is a 3D periodic thin domain, $T_\epsilon = T^2 \times T^1_\epsilon$, $T^2 = (0, l_1) \times (0, l_2)$, $0 < l_1, l_2 < \infty$, $T^1_\epsilon = (0, \epsilon)$, $0 < \epsilon < 1$. We assume that the initial data satisfies $u_0 \in H^1_{\mathrm{per}}(T_\epsilon)$ with $\int_{T_\epsilon} u_0 = 0$. As we have mentioned above, there exists a local smooth solution. However, we don't know whether the solution can be global. In fact, in the 3D case, there is a global solution provided the initial data is sufficiently small; see [Fujita and Kato 1964]. It is unknown for the global existence in the large initial data case.

Our goal in this paper is to find how large the initial data can be to ensure the global existence of strong solutions on thin periodic domain. Hale and Raugel [1992a; 1992b] studied reaction diffusion equations and damped wave equations on thin domain. Raugel and Sell [1993; 1994] further studied the existence of strong solutions of the Navier–Stokes equations on thin domain. In particular, in [Raugel and Sell 1993], they proved that, in the periodic boundary condition case, the global existence holds with initial data in a *large set* of $H^1(T_\epsilon)$. Subsequent works concerning various boundary conditions complemented and extended their result; see [Temam and Ziane 1996; Montgomery-Smith 1999; Iftimie 1999; Iftimie and Raugel 2001; Kukavica and Ziane 2006; 2007; Hou et al. 2008; Kukavica et al. 2013; 2014]. It is worth mentioning that Temam and Ziane [1996] proved that in the case with Dirichlet boundary condition, global existence holds if the initial data satisfies

$$\|\nabla u_0\|_{L^2(T_\epsilon)} \leqslant \frac{\nu}{C\epsilon^{\frac{1}{2}}}, \tag{1-2}$$

where ν denotes the viscosity. It would be very interesting to understand how far we can go in the periodic case.

However, the periodic case is quite different with the Dirichlet boundary condition case. In the case of the periodic boundary condition, there is no Poincaré inequality in the vertical direction. For this reason, in the periodic case, the global regularity is still unclear under (1-2). Montgomery and Smith [1999] proved the global existence of solutions if

$$\|\nabla u_0\|_{L^2(T_\epsilon)} \leq \frac{\nu}{C(l_1, l_2)},$$

which was later on improved by Kukavica and Ziane [2006] to

$$\|\nabla u_0\|_{L^2(T_\epsilon)} \leq \frac{\nu}{C(l_1, l_2)\epsilon^{\frac{1}{6}}}.$$

Then after a year, Kukavica and Ziane [2007] improved their result to

$$\|\nabla u_0\|_{L^2(T_\epsilon)} \le \frac{\nu}{C(l_1,l_2)\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}},$$

where C is a sufficiently large constant.

In this paper, we prove that the global existence holds if the initial data satisfies

$$\|\partial_h u_0\|_{L^2(T_\epsilon)} \leqslant \frac{C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}, \qquad \|\partial_3 u_0\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{1}{2}}},$$

where $\partial_h = (\partial_1, \partial_2)$, C^* is an arbitrarily large constant and C is a sufficiently large constant. Here without loss of generality, we have taken the viscosity to be 1. We emphasize that the vertical derivative of the velocity $\partial_3 u$ has reached the desired result with the power of $-\frac{1}{2}$ of the exponent of ϵ. This is due to the observation that the Poincaré inequality for $\partial_3 u$ in the vertical direction holds since the average of $\partial_3 u$ in the vertical direction is automatically 0 for the periodic boundary condition. More precisely, it holds that

$$\frac{1}{\epsilon}\int_0^\epsilon \partial_3 u \,\mathrm{d}x_3 = \frac{1}{\epsilon}\big[u(x_1,x_2,\epsilon)-u(x_1,x_2,0)\big] = 0,$$

since u is periodic in vertical direction. To deal with the horizontal derivative $\partial_h u$, we use the same method as [Kukavica and Ziane 2007]. However, our result allows C^* to be arbitrarily large which is required to be sufficiently small in [Kukavica and Ziane 2007]. The key improvement lies in that in the estimate of $\|u_3\|_{L^\alpha(T_\epsilon)}$, we take $\alpha = 3 + 2|\ln\epsilon|/\ln|\ln\epsilon|$ instead of $\alpha = 3 + |\ln\epsilon|$ to gain more room for C^*.

Before we state our main result, we recall our hypothesis and introduce our notations. We assume that u satisfies the periodic boundary conditions

$$\begin{cases} u(x + l_i e_i, t) = u(x,t), \quad i = 1,2, \\ u(x+\epsilon e_3, t) = u(x,t), \end{cases}$$

where $\{e_1, e_2, e_3\}$ is the natural basis in $\mathbb{R}^3$. In addition, we require that the initial data $u(x,0) = u_0(x)$ satisfies

$$\int_{T_\epsilon} u_0(x)\,\mathrm{d}x = 0. \tag{1-3}$$

It then follows that any solution of (1-1) with the initial data $u_0(x)$ will also satisfy $\int_{T_\epsilon} u(x,t)\,\mathrm{d}x = 0$ for all $t > 0$. Let $L^p(T_\epsilon) \equiv L^p(T_\epsilon, \mathbb{R}^3)$ be the space of L^p vector functions u with the usual norm

$$\|u\|_{L^p(T_\epsilon)} = \left(\int_{T_\epsilon} |u|^p \mathrm{d}x\right)^{\frac{1}{p}}.$$

Let $H^m_{\text{per}}(T_\epsilon) \equiv H^m_{\text{per}}(T_\epsilon, \mathbb{R}^3)$ denote the closure in $H^m(T_\epsilon, \mathbb{R}^3)$ of those smooth functions that are periodic in space, i.e., $u(x+l_ie_i)=u(x)$, $i=1,2,3$, where $l_3=\epsilon$. Throughout this paper, the symbol C denotes a sufficiently large constant, which depends only on l_1 and l_2. Its value may change from one inequality to another. On the other hand, the constant $C_0, C_1, \ldots$, which depend on l_1 and l_2, are fixed.

We are ready to state the main result in this paper.

Theorem 1.1. *Consider the Navier–Stokes equations (1-1) with the initial data $u_0 \in H^1_{\text{per}}(T_\epsilon)$ which satisfies (1-3). For any given arbitrarily large constant C^*, there exists an $\epsilon_0 = \epsilon_0(C^*) \in (0,1)$ such that for every $\epsilon \in (0, \epsilon_0]$, assuming that u_0 satisfies*

$$\begin{cases} \|\partial_h u_0\|_{L^2(T_\epsilon)} \leqslant \dfrac{C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}, \\ \|\partial_3 u_0\|_{L^2(T_\epsilon)} \leqslant \dfrac{1}{C_0\epsilon^{\frac{1}{2}}}, \end{cases}$$

where $\partial_h = (\partial_1, \partial_2)$ and $C_0 > 0$ is a sufficiently large constant which depends only on l_1 and l_2, then (1-1) has a unique global solution u that belongs to $C([0,\infty), H^1_{\text{per}}(T_\epsilon))$.

The following result is a key step in the proof of Theorem 1.1. We emphasize that this theorem is given by Kukavica and Ziane [2007]. However, their proof seems incomplete for us and needs some modifications. For completeness, we will present the details in Section 3.

Theorem 1.2. *Let $3 \leqslant \alpha \leqslant \widetilde{C}|\ln\epsilon|$ be arbitrary, where $\widetilde{C}$ is a large constant. Assume that the initial data $u_0 = (u_{01}, u_{02}, u_{03}) \in H^1_{\text{per}}(T_\epsilon)$ satisfies*

$$\|\nabla u_{0k}\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}}, \qquad k=1,2$$

and

$$\|u_{03}\|_{L^\alpha(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{\alpha-3}{\alpha}}}.$$

Then (1-1) has a unique global solution u. Moreover,

$$\|\nabla u_k(\cdot,t)\|_{L^2(T_\epsilon)} \leqslant \frac{C}{\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}}, \qquad k=1,2$$

and

$$\|u_3(\cdot,t)\|_{L^\alpha(T_\epsilon)} \leqslant \frac{C}{\epsilon^{\frac{\alpha-3}{\alpha}}}$$

for all $t > 0$, where $C > 0$ is a constant which depends only on l_1 and l_2.

The remaining part of this paper is organized as follows. Section 2 focuses on the Sobolev imbedding theorems for thin domain. Section 3 is devoted to proving Theorem 1.2. In Section 4, we finish the proof of Theorem 1.1 by dividing the whole time into three time intervals and using Theorem 1.2 in the third time interval to get the global regularity.

2. Preliminaries

In this section, we will introduce the average operator M and give the Sobolev imbedding theorems for thin domain. In addition, we will give an inequality about the L^α norm of u_3, which will play an important role in proving the main result.

For any $u \in L^1(T_\epsilon)$, as in [Kukavica and Ziane 2006; 2007; Raugel and Sell 1993; Temam and Ziane 1996], the average operator M is defined by

$$(Mu)(x_1, x_2) = \frac{1}{\epsilon} \int_0^\epsilon u(x_1, x_2, x_3)\, \mathrm{d}x_3.$$

We also define the operator N by (see [Kukavica and Ziane 2006; 2007; Raugel and Sell 1993; Temam and Ziane 1996])

$$Nu(x_1, x_2, x_3) = u(x_1, x_2, x_3) - (Mu)(x_1, x_2).$$

It is clear that Mu is independent of x_3 and $MNu = 0$. In addition, we also have

$$\|u\|^2_{L^2(T_\epsilon)} = \|Mu\|^2_{L^2(T_\epsilon)} + \|Nu\|^2_{L^2(T_\epsilon)}.$$

In the following lemma, we will recall the Sobolev imbedding theorems for thin domain which will be frequently used in the proof of the main result. The following estimates can be found in [Kukavica and Ziane 2006; 2007; Temam and Ziane 1996].

Lemma 2.1. *Assume $u \in H^1_{\mathrm{per}}(T_\epsilon)$.*

(i) *We have*

$$\|Nu\|_{L^2(T_\epsilon)} \leqslant C\epsilon \|\partial_3 u\|_{L^2(T_\epsilon)} \quad \textit{and} \quad \|Nu\|_{L^6(T_\epsilon)} \leqslant C\|\nabla u\|_{L^2(T_\epsilon)}.$$

For all $a \in [2, 6]$, we have

$$\|Nu\|_{L^a(T_\epsilon)} \leqslant C\|u\|^{\frac{6-a}{2a}}_{L^2(T_\epsilon)} \|\nabla u\|^{\frac{3a-6}{2a}}_{L^2(T_\epsilon)}.$$

Moreover,

$$\|Nu\|_{L^a(T_\epsilon)} \leqslant C\epsilon^{\frac{6-a}{2a}} \|\nabla u\|_{L^2(T_\epsilon)}.$$

Here C depends only on l_1 and l_2.

(ii) *For all $a \in [2, \infty)$, we have*

$$\|Mu\|_{L^a(T_\epsilon)} \leqslant \frac{Ca^{\frac{1}{2}}}{\epsilon^{\frac{a-2}{2a}}} \|u\|^{\frac{2}{a}}_{L^2(T_\epsilon)} \|\nabla u\|^{\frac{a-2}{a}}_{L^2(T_\epsilon)} + \frac{C}{\epsilon^{\frac{a-2}{2a}}} \|u\|_{L^2(T_\epsilon)}.$$

Moreover, if $\int_{T_\epsilon} u \, \mathrm{d}x = 0$, *then*

$$\|Mu\|_{L^a(T_\epsilon)} \leqslant \frac{Ca^{\frac{1}{2}}}{\epsilon^{\frac{a-2}{2a}}} \|u\|_{L^2(T_\epsilon)}^{\frac{2}{a}} \|\nabla u\|_{L^2(T_\epsilon)}^{\frac{a-2}{a}} \leqslant \frac{Ca^{\frac{1}{2}}}{\epsilon^{\frac{a-2}{2a}}} \|\nabla u\|_{L^2(T_\epsilon)}.$$

Here C *depends only on* l_1 *and* l_2.

(iii) *Assume* $\int_{T_\epsilon} u \, \mathrm{d}x = 0$. *Then for* $a \in [2, 6]$, *we have*

$$\|u\|_{L^a(T_\epsilon)} \leqslant \frac{Ca^{\frac{1}{2}}}{\epsilon^{\frac{a-2}{2a}}} \|u\|_{L^2(T_\epsilon)}^{\frac{6-a}{2a}} \|\nabla u\|_{L^2(T_\epsilon)}^{\frac{3a-6}{2a}} \leqslant \frac{Ca^{\frac{1}{2}}}{\epsilon^{\frac{a-2}{2a}}} \|\nabla u\|_{L^2(T_\epsilon)},$$

where C *depends only on* l_1 *and* l_2.

One can find the proof of the above lemma in [Kukavica and Ziane 2006; 2007; Temam and Ziane 1996]. It should be pointed out that to get the last two inequalities in Lemma 2.1, we used the following Poincaré inequality on the periodic domain $T_\epsilon = (0, l_1) \times (0, l_2) \times (0, \epsilon)$:

$$\|u\|_{L^2(T_\epsilon)} \leq C \|\nabla u\|_{L^2(T_\epsilon)}, \tag{2-1}$$

where C depends on l_1 and l_2. Inequality (2-1) is valid under the assumption $\int_{T_\epsilon} u \, \mathrm{d}x = 0$. To prove this, we first see that $u = Mu + Nu$. This means that the integral average of Nu on the vertical direction and Mu on the horizontal direction are 0, respectively, i.e.,

$$\frac{1}{\epsilon} \int_0^\epsilon Nu \, \mathrm{d}x_3 = 0, \qquad \frac{1}{|T^2|} \int_{T^2} Mu \, \mathrm{d}x_1 \, \mathrm{d}x_2 = 0.$$

Using the Poincaré inequality for Nu on the vertical direction and Mu on T^2, respectively, we get

$$\|Nu\|_{L^2(T_\epsilon)} \leq C\epsilon \|\partial_3 u\|_{L^2(T_\epsilon)}, \qquad \text{(see Lemma 2.1(i))}$$
$$\|Mu\|_{L^2(T^2)} \leq C \|\partial_h Mu\|_{L^2(T^2)} \Rightarrow \|Mu\|_{L^2(T_\epsilon)} \leq C \|\partial_h u\|_{L^2(T_\epsilon)}.$$

Hence,

$$\begin{aligned} \|u\|_{L^2(T_\epsilon)} &\leq \|Nu\|_{L^2(T_\epsilon)} + \|Mu\|_{L^2(T_\epsilon)} \\ &\leq C\epsilon \|\partial_3 u\|_{L^2(T_\epsilon)} + C \|\partial_h u\|_{L^2(T_\epsilon)} \leq C \|\nabla u\|_{L^2(T_\epsilon)}. \end{aligned}$$

Next, we will give an estimate concerning the L^α norm of u_3 which has appeared in [Kukavica and Ziane 2006, Lemma 3] for $\alpha = 6$ and in [Kukavica et al. 2013, Lemma 4.2] for general α in the two-dimensional case. Below is the three-dimensional case. We remark that this has been proven for $\alpha = 6$ in [Kukavica and Ziane 2006, Lemma 4]. For completeness, we will present a proof below which seems even simpler.

Lemma 2.2. *Consider u_3, the third component of the velocity, which is defined on T_ϵ. Let $\alpha \in [2,\infty)$ be arbitrary. Assume that $u_3 \in H^1_{\mathrm{per}}(T_\epsilon) \cap L^\alpha(T_\epsilon)$ satisfies $\nabla(|u_3|^{\frac{\alpha}{2}}) \in L^2_{\mathrm{per}}(T_\epsilon)$ and $\int_{T_\epsilon} u_3 \,\mathrm{d}x = 0$. Then*

$$\big\| |u_3|^{\frac{\alpha}{2}} \big\|^2_{L^2(T_\epsilon)} \leqslant C \big\| \nabla(|u_3|^{\frac{\alpha}{2}}) \big\|^2_{L^2(T_\epsilon)}, \tag{2-2}$$

where C depends only on l_1, l_2 and α.

Remark 2.3. Lemma 2.2 will be used to prove the main result Theorem 1.1. We don't need to add the assumption $\int_{T_\epsilon} u_3 \,\mathrm{d}x = 0$ which appears in Lemma 2.2 to Theorem 1.1. Actually, in Theorem 1.1, we have made an assumption to the initial data $u_0 = (u_{01}, u_{02}, u_{03})$, that is,

$$\int_{T_\epsilon} u_0(x)\,\mathrm{d}x = 0;$$

see (1-3). Under this assumption, we can see that any solution $u = (u_1, u_2, u_3)$ of Navier–Stokes equations with this initial data will satisfy

$$\int_{T_\epsilon} u(x,t)\,\mathrm{d}x = 0$$

for all $t > 0$. Hence when we use Lemma 2.2 to prove Theorem 1.1, we don't need to make extra assumptions.

Proof. Since the size of T_ϵ is not order one, we make a transform to map T_ϵ onto $\widetilde{\Omega} = (0,l_1)\times(0,l_2)\times(0,1)$. The transform is defined by

$$u_3(x_1,x_2,x_3) = u_3(y_1,y_2,\epsilon y_3) = v(y_1,y_2,y_3), \tag{2-3}$$

where $x = (x_1,x_2,x_3) \in T_\epsilon$, $y = (y_1,y_2,y_3) \in \widetilde{\Omega}$ and $x_i = y_i$, $i = 1,2$; $x_3 = \epsilon y_3$. Then we know that v is defined on $\widetilde{\Omega}$ whose size is order one. Let $\tilde{u}(x) = |u_3|^{\frac{\alpha}{2}}(x)$ and $\tilde{v}(y) = |v|^{\frac{\alpha}{2}}(y)$. Since $\int_{T_\epsilon} u_3 \,\mathrm{d}x = 0$, it is obvious $\int_{\widetilde{\Omega}} v(y)\,\mathrm{d}y = 0$. By a similar argument as that of Lemma 3 in [Kukavica and Ziane 2006], we have

$$\|\tilde{v}\|^2_{L^2(\widetilde{\Omega})} \leqslant C \|\nabla \tilde{v}\|^2_{L^2(\widetilde{\Omega})}, \tag{2-4}$$

where C depends only on l_1, l_2 and α. Moreover, we can conclude from (2-3) that

$$\|\tilde{v}\|^2_{L^2(\widetilde{\Omega})} = \frac{1}{\epsilon} \|\tilde{u}\|^2_{L^2(T_\epsilon)}, \tag{2-5}$$

and

$$\|\nabla \tilde{v}\|^2_{L^2(\widetilde{\Omega})} \leqslant \frac{1}{\epsilon} \|\nabla \tilde{u}\|^2_{L^2(T_\epsilon)}. \tag{2-6}$$

It then follows from (2-4)–(2-6) that

$$\|\tilde{u}\|^2_{L^2(T_\epsilon)} \leqslant C\|\nabla\tilde{u}\|^2_{L^2(T_\epsilon)},$$

where C depends only on l_1, l_2 and α. Thus we complete the proof of (2-2). □

3. Proof of Theorem 1.2

In this section, we will prove Theorem 1.2. We follow the idea of [Kukavica and Ziane 2007]. However, from our point of view, compared with the proof in [Kukavica and Ziane 2007], two places need to be modified when we estimate K_3 coming from the estimate of $\|u_3\|_{L^\alpha}$. We will show the details in the following proof.

Proof. Since the initial data $u_0 \in H^1_{\text{per}}(T_\epsilon)$, we know that the solution of (1-1) is smooth and unique on an initial time interval $(0, T_{\max})$, where $T_{\max} > 0$ depends on u_0. Take t_1, $0 < t_1 < T_{\max}$ and suppose $t \in [0, t_1]$. By (1-1), the componentwise Navier–Stokes equations become

$$\partial_t u_k - \Delta u_k + \sum_{j=1}^{3} u_j \partial_j u_k + \partial_k p = 0, \tag{3-1}$$

where $k = 1, 2, 3$.

Consider the Navier–Stokes equations (3-1) for $k = 1, 2$. We multiply the equations with $-\Delta u_k$ respectively and integrate over $T_\epsilon \times [0, t]$, and sum. Let $u_h = (u_1, u_2)$. It then follows that

$$\begin{aligned}\|\nabla u_h(t)\|^2_{L^2_x} - \|\nabla u_{0h}\|^2_{L^2_x} + \|\Delta u_h\|^2_{L^2_t L^2_x} &= \sum_{j=1}^{2}\iint u_j\partial_j u_h \Delta_h u_h + \sum_{j=1}^{2}\iint u_j \partial_j u_h \partial_{33} u_h \\ &\quad + \iint u_3\partial_3 u_h \Delta u_h + \iint \partial_h p \Delta u_h \\ &= J_1 + J_2 + J_3 + J_4,\end{aligned} \tag{3-2}$$

where $\Delta_h = \partial_{11} + \partial_{22}$, $\partial_h = (\partial_1, \partial_2)$ and we abbreviate

$$\|\cdot\|_{L^s_t L^r_x} = \|\cdot\|_{L^s((0,t), L^r(T_\epsilon))}.$$

We remark that above and in the sequel, all unmarked double integrals are understood to be over $T_\epsilon \times [0, t]$ and all unmarked single integrals are understood to be over T_ϵ.

For the term J_1, using integration by parts together with the fact $\nabla\cdot u = 0$, we get

$$\begin{aligned} J_1 &= -\sum_{i,j=1}^{2}\iint \partial_i u_j \partial_j u_h \partial_i u_h - \sum_{i,j=1}^{2}\iint u_j \partial_j \partial_i u_h \partial_i u_h \\ &= -\sum_{i,j=1}^{2}\iint \partial_i u_j \partial_j u_h \partial_i u_h + \frac{1}{2}\sum_{i,j=1}^{2}\iint \partial_j u_j \partial_i u_h \partial_i u_h \\ &= -\sum_{i,j=1}^{2}\iint \partial_i u_j \partial_j u_h \partial_i u_h - \frac{1}{2}\sum_{i=1}^{2}\iint \partial_i u_h \partial_i u_h \partial_3 u_3 \\ &= \frac{1}{2}\sum_{i=1}^{2}\iint \partial_i u_h \partial_i u_h \partial_3 u_3 + \iint \partial_1 u_2 \partial_2 u_1 \partial_3 u_3 - \iint \partial_1 u_1 \partial_2 u_2 \partial_3 u_3 \\ &= J_{11} + J_{12} + J_{13}. \end{aligned}$$

We next estimate J_{11}, J_{12}, J_{13}. Define

$$J(t) = \|\nabla u_h\|_{L_t^\infty L_x^2} + \|\nabla^2 u_h\|_{L_t^2 L_x^2},$$

where $\nabla^2 = (\partial_{ij})$, $i, j = 1, 2, 3$. Then we have the following useful estimate:

$$\|\partial_3 u_k\|_{L_t^2 L_x^a} \leqslant C\epsilon^{\frac{6-a}{2a}} J(t_1), \qquad a \in [2,6],\ k = 1,2,3. \tag{3-3}$$

Since $\int_0^\epsilon \partial_3 u_k \,\mathrm{d}x_3 = 0$ for $k = 1, 2, 3$, by using Lemma 2.1(i), we have

$$\|\partial_3 u_k\|_{L_t^2 L_x^a} \leqslant C\epsilon^{\frac{6-a}{2a}} \|\nabla \partial_3 u_k\|_{L_t^2 L_x^2} \leqslant C\epsilon^{\frac{6-a}{2a}} J(t_1), \quad k = 1, 2.$$

By using the divergence-free condition, we get

$$\|\partial_3 u_3\|_{L_t^2 L_x^a} \leqslant C\epsilon^{\frac{6-a}{2a}} \|\nabla(\partial_1 u_1 + \partial_2 u_2)\|_{L_t^2 L_x^2} \leqslant C\epsilon^{\frac{6-a}{2a}} J(t_1).$$

Thus we finish the proof of the inequality (3-3). For the term J_{11}, we decompose it into three parts:

$$\begin{aligned} J_{11} = \frac{1}{2}\sum_{i=1}^{2}\iint M(\partial_i u_h) M(\partial_i u_h) \partial_3 u_3 + \sum_{i=1}^{2}\iint M(\partial_i u_h) N(\partial_i u_h) \partial_3 u_3 \\ + \frac{1}{2}\sum_{i=1}^{2}\iint N(\partial_i u_h) N(\partial_i u_h) \partial_3 u_3 \\ = J_{111} + J_{112} + J_{113}. \end{aligned} \tag{3-4}$$

Regarding J_{111}, due to the fact that $M(\partial_i u_h)$ is independent of x_3, we have $\partial_3 M(\partial_i u_h) = 0$, thus

$$J_{111} = -\frac{1}{2}\sum_{i=1}^{2}\iint \partial_3 M(\partial_i u_h) M(\partial_i u_h) u_3 - \frac{1}{2}\sum_{i=1}^{2}\iint M(\partial_i u_h)\partial_3 M(\partial_i u_h) u_3 = 0. \tag{3-5}$$

Regarding J_{113}, we have

$$J_{113} \leqslant C\sum_{i=1}^{2} \|N(\partial_i u_h)\|_{L_t^4 L_x^3} \|N(\partial_i u_h)\|_{L_t^4 L_x^3} \|\partial_3 u_3\|_{L_t^2 L_x^3}. \tag{3-6}$$

Since $\int_0^\epsilon N(\partial_i u_h)\,\mathrm{d}x_3 = 0$, we have

$$\begin{aligned}\|N(\partial_i u_h)\|_{L_t^4 L_x^3} &\leqslant C\big\| \|\partial_i u_h\|_{L_x^2}^{\frac{1}{2}} \|\nabla\partial_i u_h\|_{L_x^2}^{\frac{1}{2}} \big\|_{L_t^4} \\ &\leqslant C\|\partial_i u_h\|_{L_t^\infty L_x^2}^{\frac{1}{2}} \|\nabla\partial_i u_h\|_{L_t^2 L_x^2}^{\frac{1}{2}} \\ &\leqslant C(\|\partial_i u_h\|_{L_t^\infty L_x^2} + \|\nabla\partial_i u_h\|_{L_t^2 L_x^2}) \leqslant CJ(t_1).\end{aligned} \tag{3-7}$$

By using the inequality (3-3) with $a = 3$, we have

$$\|\partial_3 u_3\|_{L_t^2 L_x^3} \leqslant C\epsilon^{\frac{1}{2}} J(t_1). \tag{3-8}$$

It then follows from (3-6)–(3-8) that

$$J_{113} \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3. \tag{3-9}$$

Regarding J_{112}, we have

$$J_{112} \leqslant C\sum_{i=1}^{2} \|M(\partial_i u_h)\|_{L_t^4 L_x^4} \|N(\partial_i u_h)\|_{L_t^4 L_x^3} \|\partial_3 u_3\|_{L_t^2 L_x^{\frac{12}{5}}}. \tag{3-10}$$

Since $\int_{T_\epsilon} \partial_i u_h\,\mathrm{d}x = 0$, we can see from Lemma 2.1(ii) with $a = 4$ that

$$\|M(\partial_i u_h)\|_{L_x^4} \leqslant C\epsilon^{-\frac{1}{4}} \|\partial_i u_h\|_{L_x^2}^{\frac{1}{2}} \|\nabla\partial_i u_h\|_{L_x^2}^{\frac{1}{2}}.$$

Therefore,

$$\begin{aligned}\|M(\partial_i u_h)\|_{L_t^4 L_x^4} &\leqslant C\epsilon^{-\frac{1}{4}} \|\partial_i u_h\|_{L_t^\infty L_x^2}^{\frac{1}{2}} \|\nabla\partial_i u_h\|_{L_t^2 L_x^2}^{\frac{1}{2}} \\ &\leqslant C\epsilon^{-\frac{1}{4}}(\|\partial_i u_h\|_{L_t^\infty L_x^2} + \|\nabla\partial_i u_h\|_{L_t^2 L_x^2}) \leqslant C\epsilon^{-\frac{1}{4}} J(t_1).\end{aligned} \tag{3-11}$$

By using the inequality (3-3) with $a = \frac{12}{5}$, we have

$$\|\partial_3 u_3\|_{L_t^2 L_x^{\frac{12}{5}}} \leqslant C\epsilon^{\frac{3}{4}} J(t_1). \tag{3-12}$$

It then follows from (3-7), (3-10)–(3-12) that

$$J_{112} \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3. \tag{3-13}$$

Based on (3-4), (3-5), (3-9) and (3-13), we have

$$J_{11} \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3.$$

The terms J_{12} and J_{13} are estimated in the same way as J_{11}. Therefore, we obtain

$$J_1 \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3. \tag{3-14}$$

For the term J_2, using integration by parts together with $\nabla \cdot u = 0$, we have

$$\begin{aligned} J_2 &= \sum_{j=1}^{2} \iint u_j \partial_j u_h \partial_{33} u_h = -\sum_{j=1}^{2} \iint \partial_3 u_j \partial_j u_h \partial_3 u_h - \sum_{j=1}^{2} \iint u_j \partial_j \partial_3 u_h \partial_3 u_h \\ &= -\sum_{j=1}^{2} \iint \partial_3 u_j \partial_j u_h \partial_3 u_h + \frac{1}{2} \sum_{j=1}^{2} \iint \partial_j u_j \partial_3 u_h \partial_3 u_h = J_{21} + J_{22}. \end{aligned}$$

Regarding J_{21}, we have

$$J_{21} \leqslant \sum_{j=1}^{2} \|\partial_3 u_j\|_{L_t^2 L_x^4} \|\partial_j u_h\|_{L_t^\infty L_x^2} \|\partial_3 u_h\|_{L_t^2 L_x^4} \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3,$$

where we have used (3-3) with $a = 4$. The same estimate holds for J_{22}. Therefore,

$$J_2 \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3. \tag{3-15}$$

For the term J_3, define

$$K(t) = \left(\left\| |u_3|^{\frac{\alpha}{2}} \right\|_{L_t^\infty L_x^2} + \left\| \nabla(|u_3|^{\frac{\alpha}{2}}) \right\|_{L_t^2 L_x^2}\right)^{\frac{2}{\alpha}}, \quad t \in [0, T_{\max}).$$

Then we get

$$J_3 \leqslant \|u_3\|_{L_t^\infty L_x^\alpha} \|\partial_3 u_h\|_{L_t^2 L_x^{\frac{2\alpha}{\alpha-2}}} \|\Delta u_h\|_{L_t^2 L_x^2}.$$

Since $3 \leqslant \alpha \leqslant \widetilde{C}|\ln\epsilon|$ implies that $2 < \frac{2\alpha}{\alpha-2} \leqslant 6$, it follows from (3-3) that

$$\|\partial_3 u_h\|_{L_t^2 L_x^{\frac{2\alpha}{\alpha-2}}} \leqslant C\epsilon^{\frac{\alpha-3}{\alpha}} J(t_1).$$

Thus,

$$J_3 \leqslant C\epsilon^{\frac{\alpha-3}{\alpha}} K(t_1) J(t_1)^2. \tag{3-16}$$

For the term J_4 which includes Δp, we need to take the divergence of (1-1) and obtain that

$$-\Delta p = \nabla\cdot(u\cdot\nabla u) = \sum_{i,j=1}^{3} \partial_i u_j \partial_j u_i.$$

Then we have

$$\begin{aligned} J_4 &= -\iint \Delta p \partial_h u_h = \sum_{i,j=1}^{3} \iint \partial_i u_j \partial_j u_i \partial_h u_h \\ &= \sum_{i,j=1}^{2} \iint \partial_i u_j \partial_j u_i \partial_h u_h + 2\sum_{j=1}^{2} \iint \partial_3 u_j \partial_j u_3 \partial_h u_h \\ &= -\sum_{i,j=1}^{2} \iint \partial_i u_j \partial_j u_i \partial_3 u_3 + 2\sum_{j=1}^{2} \iint \partial_3 u_j \partial_j u_3 \partial_h u_h \\ &= J_{41} + J_{42}. \end{aligned}$$

The term J_{41} can be estimated in a similar way to J_{11}, giving

$$J_{41} \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3. \tag{3-17}$$

Regarding J_{42}, using integration by parts, we have

$$\begin{aligned} J_{42} &= -2\sum_{j=1}^{2} \iint \partial_j \partial_3 u_j u_3 \partial_h u_h - 2\sum_{j=1}^{2} \iint \partial_3 u_j u_3 \partial_j \partial_h u_h \\ &= 2\sum_{j=1}^{2} \iint \partial_j \partial_3 u_j u_3 \partial_3 u_3 - 2\sum_{j=1}^{2} \iint \partial_3 u_j u_3 \partial_j \partial_h u_h \\ &= J_{421} + J_{422}. \end{aligned}$$

Estimate J_{421} and J_{422} to obtain

$$J_{421} \leqslant C\|\partial_j \partial_3 u_j\|_{L_t^2 L_x^2} \|u_3\|_{L_t^\infty L_x^\alpha} \|\partial_3 u_3\|_{L_t^2 L_x^{\frac{2\alpha}{\alpha-2}}} \leqslant C\epsilon^{\frac{\alpha-3}{\alpha}} K(t_1) J(t_1)^2$$

$$J_{422} \leqslant C\sum_{j=1}^{2} \|\partial_3 u_j\|_{L_t^2 L_x^{\frac{2\alpha}{\alpha-2}}} \|u_3\|_{L_t^\infty L_x^\alpha} \|\partial_j \partial_h u_h\|_{L_t^2 L_x^2} \leqslant C\epsilon^{\frac{\alpha-3}{\alpha}} K(t_1) J(t_1)^2.$$

Thus we conclude

$$J_{42} \leqslant C\epsilon^{\frac{\alpha-3}{\alpha}} K(t_1) J(t_1)^2. \tag{3-18}$$

It then follows from (3-17) and (3-18) that

$$J_4 \leqslant C\epsilon^{\frac{1}{2}} J(t_1)^3 + C\epsilon^{\frac{\alpha-3}{\alpha}} K(t_1) J(t_1)^2. \tag{3-19}$$

Therefore, by (3-2), (3-14)–(3-16) and (3-19), we obtain the final estimate about $J(t)$:

$$J(t)^2 \leqslant C\epsilon^{\frac{1}{2}} J(t)^3 + C\epsilon^{\frac{\alpha-3}{\alpha}} K(t) J(t)^2 + J(0)^2, \tag{3-20}$$

where we have used the second derivative estimate

$$\|\nabla^2 u_h\|_{L_t^2 L_x^2} \leqslant C\|\Delta u_h\|_{L_t^2 L_x^2}$$

together with the fact

$$J(0) = \|\nabla u_{0h}\|_{L_x^2}.$$

The next objective is to estimate $K(t)$. Consider the Navier–Stokes equations (3-1) for $k = 3$. We multiply it with $|u_3|^{\alpha-1}\operatorname{sgn} u_3$ and integrate over $T_\epsilon \times [0,t]$. There holds

$$\iint \partial_t u_3 |u_3|^{\alpha-1}\operatorname{sgn} u_3 + \iint u\cdot\nabla u_3 |u_3|^{\alpha-1}\operatorname{sgn} u_3 - \iint \Delta u_3 |u_3|^{\alpha-1}\operatorname{sgn} u_3 \\ = -\iint \partial_3 p |u_3|^{\alpha-1}\operatorname{sgn} u_3.$$

After a short calculation, we have

$$\frac{1}{\alpha}\|u_3(t)\|_{L_x^\alpha}^\alpha + \frac{4(\alpha-1)}{\alpha^2}\|\nabla(|u_3|^{\frac{\alpha}{2}})\|_{L_t^2 L_x^2}^2 \\ = -\iint \partial_3 p|u_3|^{\alpha-1}\operatorname{sgn} u_3 + \frac{1}{\alpha}\|u_{03}\|_{L_x^\alpha}^\alpha. \tag{3-21}$$

It remains to estimate $-\iint \partial_3 p|u_3|^{\alpha-1}\operatorname{sgn} u_3$. From (1-1), we know that

$$p = (-\Delta)^{-1}\nabla\cdot\nabla\cdot(u\otimes u) = \sum_{i,j=1}^{3} R_i R_j(u_i u_j) = \sum_{i,j=1}^{3} R_{i,j}(u_i u_j),$$

where R_1, R_2, R_3 are the Riesz transforms. Since ∂_3 can commute with the Riesz transforms, we have

$$\begin{aligned} \partial_3 p &= \partial_3 \sum_{i,j=1}^{3} R_{i,j}(u_i u_j) = 2\sum_{i,j=1}^{3} R_{i,j}(\partial_3 u_i u_j) \\ &= 2\sum_{i=1}^{3}\sum_{j=1}^{2} R_{i,j}(\partial_3 u_i N u_j) + 2\sum_{i=1}^{3}\sum_{j=1}^{2} R_{i,j}(\partial_3 u_i M u_j) + 2\sum_{i=1}^{3} R_{i,3}(\partial_3 u_i u_3) \\ &= q_1 + q_2 + q_3. \end{aligned}$$

Therefore, we obtain

$$
\begin{aligned}
-\iint \partial_3 p|u_3|^{\alpha-1}\operatorname{sgn} u_3
&= -\iint (q_1+q_2+q_3)u_3|u_3|^{\alpha-2} \\
&= -\iint q_1|u_3|^{\frac{\alpha-2}{2}} N(|u_3|^{\frac{\alpha}{2}})\operatorname{sgn} u_3 - \iint q_1|u_3|^{\frac{\alpha-2}{2}} M(|u_3|^{\frac{\alpha}{2}})\operatorname{sgn} u_3 \\
&\qquad - \iint q_2 u_3|u_3|^{\alpha-2} - \iint q_3|u_3|^{\frac{\alpha-2}{2}} N(|u_3|^{\frac{\alpha}{2}})\operatorname{sgn} u_3 \\
&\qquad - \iint q_3|u_3|^{\frac{\alpha-2}{2}} M(|u_3|^{\frac{\alpha}{2}})\operatorname{sgn} u_3 \\
&= K_1+K_2+K_3+K_4+K_5.
\end{aligned}
$$

For the term K_1,

$$
\begin{aligned}
K_1 &\leqslant \|q_1\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \big\| |u_3|^{\frac{\alpha-2}{2}} \big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}} \big\| N(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2 L_x^6} \\
&\leqslant C\|q_1\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \big\| |u_3|^{\frac{\alpha}{2}} \big\|^{\frac{\alpha-2}{\alpha}}_{L_t^\infty L_x^2} \big\| \nabla(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2 L_x^2} \leqslant C\|q_1\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} K(t_1)^{\alpha-1},
\end{aligned}
$$

where we have used Lemma 2.1(i) for $\big\| N(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2 L_x^6}$. Regarding $\|q_1\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}}$, we have

$$
\begin{aligned}
\text{(3-22)} \qquad \|q_1\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} &\leqslant C \sum_{i=1}^{3}\sum_{j=1}^{2} \|R_{i,j}\|_{L_x^{\frac{3\alpha}{\alpha+3}}} \|\partial_3 u_i N u_j\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \\
&\leqslant C \sum_{i=1}^{3}\sum_{j=1}^{2} \|R_{i,j}\|_{L_x^{\frac{3\alpha}{\alpha+3}}} \|\partial_3 u_i\|_{L_t^2 L_x^{\frac{6\alpha}{\alpha+6}}} \|N u_j\|_{L_t^\infty L_x^6}.
\end{aligned}
$$

As we know, the Riesz transforms $R_{i,j}$ $(i, j = 1, 2, 3)$ are bounded on $L^p(T_\epsilon)$ for $1 < p < \infty$. Furthermore, the bound is given by (see [Grafakos 2004, p. 362])

$$
\text{(3-23)} \qquad \|R_{i,j}\|_{L^p} \leqslant C \max\Big(p, \frac{1}{p-1}\Big),
$$

where C is independent of p. Here, $\frac{3}{2} \leqslant \frac{3\alpha}{\alpha+3} < 3$ when $3 \leqslant \alpha \leqslant \widetilde{C}|\ln\epsilon|$. Thus, $\|R_{i,j}\|_{L_x^{\frac{3\alpha}{\alpha+3}}} \leqslant C$ for $i, j = 1, 2, 3$. Since $2 \leqslant \frac{6\alpha}{\alpha+6} < 6$ when $3 \leqslant \alpha \leqslant \widetilde{C}|\ln\epsilon|$, we can see from (3-3) that

$$
\text{(3-24)} \qquad \|\partial_3 u_i\|_{L_t^2 L_x^{\frac{6\alpha}{\alpha+6}}} \leqslant C\epsilon^{\frac{3}{\alpha}} J(t_1).
$$

Also, by using Lemma 2.1(i) we have

$$\sum_{j=1}^{2} \|Nu_j\|_{L_t^\infty L_x^6} \leqslant \sum_{j=1}^{2} \|\nabla u_j\|_{L_t^\infty L_x^2} \leqslant CJ(t_1). \tag{3-25}$$

Thus, by (3-22)–(3-25), we get

$$\|q_1\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \leqslant C\epsilon^{\frac{3}{\alpha}} J(t_1)^2.$$

Therefore, we obtain the estimate of K_1,

$$K_1 \leqslant C\epsilon^{\frac{3}{\alpha}} J(t_1)^2 K(t_1)^{\alpha-1}. \tag{3-26}$$

For the term K_2, we have

$$K_2 \leqslant \|q_1\|_{L_t^2 L_x^{r_1}} \big\| |u_3|^{\frac{\alpha-2}{2}} \big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}} \big\| M(|u_3|^{\frac{\alpha}{2}}) \big\|_{L_t^2 L_x^b}, \tag{3-27}$$

where r_1 and b satisfy $\frac{1}{r_1} + \frac{\alpha-2}{2\alpha} + \frac{1}{b} = 1$. Let $b \geqslant 2\alpha$, then we have $r_1 \in \big(\frac{2\alpha}{\alpha+2}, \frac{2\alpha}{\alpha+1}\big]$. Now we estimate the three terms on the right-hand side of (3-27). Regarding $\|q_1\|_{L_t^2 L_x^{r_1}}$, we have

$$\|q_1\|_{L_t^2 L_x^{r_1}} \leqslant C \sum_{i=1}^{3} \sum_{j=1}^{2} \|\partial_3 u_i\|_{L_t^2 L_x^{2r_1}} \|Nu_j\|_{L_t^\infty L_x^{2r_1}},$$

Because of the fact that $2r_1 \in \big(\frac{4\alpha}{\alpha+2}, \frac{4\alpha}{\alpha+1}\big] \subset \big[\frac{12}{5}, 4\big)$ when $3 \leqslant \alpha \leqslant C|\ln\epsilon|$, we conclude from (3-3) that

$$\|\partial_3 u_i\|_{L_t^2 L_x^{2r_1}} \leqslant C\epsilon^{\frac{3-r_1}{2r_1}} J(t_1).$$

Also, by using Lemma 2.1(i) we have

$$\sum_{j=1}^{2} \|Nu_j\|_{L_t^\infty L_x^{2r_1}} \leqslant C\epsilon^{\frac{3-r_1}{2r_1}} \sum_{j=1}^{2} \|\nabla u_j\|_{L_t^\infty L_x^2} \leqslant C\epsilon^{\frac{3-r_1}{2r_1}} J(t_1).$$

Thus, we have

$$\|q_1\|_{L_t^2 L_x^{r_1}} \leqslant C\epsilon^{\frac{3-r_1}{r_1}} J(t_1)^2. \tag{3-28}$$

Regarding $\big\| |u_3|^{\frac{\alpha-2}{2}} \big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}}$, we have

$$\big\| |u_3|^{\frac{\alpha-2}{2}} \big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}} = \big\| |u_3|^{\frac{\alpha}{2}} \big\|_{L_t^\infty L_x^2}^{\frac{\alpha-2}{\alpha}} \leqslant CK(t_1)^{\frac{\alpha-2}{2}}. \tag{3-29}$$

Regarding $\big\|M(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2L_x^b}$, by using Lemma 2.1(ii) we have

$$\begin{aligned}
&\big\|M(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2L_x^b}\\
&\quad\leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}\Big\|\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_x^2}^{\frac{2}{b}}\big\|\nabla(|u_3|^{\frac{\alpha}{2}})\big\|_{L_x^2}^{\frac{b-2}{b}}\Big\|_{L_t^2}+C\epsilon^{-\frac{b-2}{2b}}\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_t^2L_x^2}\\
&\quad\leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}\Big\|\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_x^2}^{\frac{2}{b}}\Big\|_{L_t^b}\Big\|\big\|\nabla(|u_3|^{\frac{\alpha}{2}})\big\|_{L_x^2}^{\frac{b-2}{b}}\Big\|_{L_t^{\frac{2b}{b-2}}}+C\epsilon^{-\frac{b-2}{2b}}\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_t^2L_x^2}\\
&\quad\leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_t^2L_x^2}^{\frac{2}{b}}\big\|\nabla(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2L_x^2}^{\frac{b-2}{b}}+C\epsilon^{-\frac{b-2}{2b}}\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_t^2L_x^2}.
\end{aligned}$$

Meanwhile, by using Lemma 2.2, we know that

$$\big\|M(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2L_x^b}\leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}\big\|\nabla(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2L_x^2}\leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}K(t_1)^{\frac{\alpha}{2}}. \tag{3-30}$$

Therefore, it follows from (3-27)–(3-30) that

$$K_2\leqslant Cb^{\frac12}\epsilon^{\frac{3}{\alpha}-\frac{2}{b}}J(t_1)^2K(t_1)^{\alpha-1}. \tag{3-31}$$

For the term K_3, we first rewrite $q_2=2\sum_{i=1}^{3}\sum_{j=1}^{2}R_{i,j}(\partial_3u_iMu_j)$. Since Mu_j is independent of x_3, we have $\partial_3u_iMu_j=\partial_3(Mu_i+Nu_i)Mu_j=\partial_3(Nu_iMu_j)$. Let

$$\widetilde{q_2}=2\sum_{i=1}^{3}\sum_{j=1}^{2}R_{i,j}(Nu_iMu_j).$$

Then we have $q_2=\partial_3\widetilde{q_2}$ as the derivative can commute with the Riesz transforms. Thus, we obtain the following result

$$\begin{aligned}
K_3&=-\iint\partial_3\widetilde{q_2}u_3|u_3|^{\alpha-2}=\iint\widetilde{q_2}\partial_3u_3|u_3|^{\alpha-2}+\iint\widetilde{q_2}u_3\partial_3(|u_3|^{\alpha-2})\\
&=(\alpha-1)\iint\widetilde{q_2}\partial_3u_3|u_3|^{\alpha-2}=\frac{2(\alpha-1)}{\alpha}\iint\widetilde{q_2}\partial_3(|u_3|^{\frac{\alpha}{2}})|u_3|^{\frac{\alpha-2}{2}}\operatorname{sgn}u_3.
\end{aligned}$$

According to the above result, we have

$$\begin{aligned}
K_3&\leqslant\frac{2(\alpha-1)}{\alpha}\|\widetilde{q_2}\|_{L_t^2L_x^\alpha}\big\|\partial_3(|u_3|^{\frac{\alpha}{2}})\big\|_{L_t^2L_x^2}\big\||u_3|^{\frac{\alpha-2}{2}}\big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}}\\
&\leqslant C\|\widetilde{q_2}\|_{L_t^2L_x^\alpha}\big\||u_3|^{\frac{\alpha}{2}}\big\|_{L_t^\infty L_x^2}^{\frac{\alpha-2}{\alpha}}K(t_1)^{\frac{\alpha}{2}}\\
&\leqslant C\|\widetilde{q_2}\|_{L_t^2L_x^\alpha}K(t_1)^{\alpha-1}.
\end{aligned}$$

It remains to estimate $\|\widetilde{q_2}\|_{L_t^2 L_x^\alpha}$.

$$\begin{aligned}\|\widetilde{q_2}\|_{L_x^\alpha} &\leqslant C\sum_{i=1}^3\sum_{j=1}^2 \|R_{i,j}Nu_iMu_j\|_{L_x^\alpha} \leqslant C\sum_{i=1}^3\sum_{j=1}^2 \|R_{i,j}\|_{L_x^\alpha}\|Nu_iMu_j\|_{L_x^\alpha} \\ &\leqslant C\alpha\sum_{i=1}^3\sum_{j=1}^2 \|Nu_iMu_j\|_{L_x^\alpha}.\end{aligned}$$

Here, compared with [Kukavica and Ziane 2007], we modified the estimate of $\|\widetilde{q_2}\|_{L_x^\alpha}$ by adding the L^α norm of Riesz transforms given by (3-23). The reason is that we will take α to be very large, roughly like $|\ln\epsilon|$, when proving the main result. Hence

$$\|\widetilde{q_2}\|_{L_t^2 L_x^\alpha} \leqslant C\alpha\sum_{i=1}^3\sum_{j=1}^2 \|Nu_iMu_j\|_{L_t^2 L_x^\alpha} \leqslant C\alpha\sum_{i=1}^3\sum_{j=1}^2 \|Nu_i\|_{L_t^2 L_x^{r_2}}\|Mu_j\|_{L_t^\infty L_x^b},$$

where $b \geqslant 2\alpha$ and $r_2 = \frac{b\alpha}{b-\alpha} \in (\alpha, 2\alpha]$. By using Lemma 2.1(ii) we have

$$\|Mu_j\|_{L_t^\infty L_x^b} \leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}\|\nabla u_j\|_{L_t^\infty L_x^2} \leqslant Cb^{\frac12}\epsilon^{-\frac{b-2}{2b}}J(t_1).$$

One expects to bound $\|Nu_i\|_{L_t^2 L_x^{r_2}}$ by $\|\nabla^2 u_h\|_{L_t^2 L_x^2}$ and thus by $J(t)$. In [Kukavica and Ziane 2007], the authors considered two cases: $2 \le r_2 \le 6$ and $6 \le r_2 < \infty$. When $2 \le r_2 \le 6$, by Lemma 2.1(i),

$$\|Nu_i\|_{L_t^2 L_x^{r_2}} \le C\epsilon^{\frac{6-r_2}{2r_2}}\sum_{k=1}^3 \|\partial_k(Nu_i)\|_{L_t^2 L_x^2}.$$

For the case $i = 3, k = 1, 2$, they used the Poincaré inequality to get

$$\|\partial_k(Nu_3)\|_{L_t^2 L_x^{r_2}} \le C\epsilon\|\partial_3\partial_k u_3\|_{L_t^2 L_x^2} \le C\epsilon\|\partial_k\partial_h u_h\|_{L_t^2 L_x^2}.$$

When $6 \le r_2 < \infty$, they first used the Gagliardo–Nirenberg inequality to get

$$\|Nu_i\|_{L_t^2 L_x^{r_2}} \le C\sum_{k=1}^3 \|\partial_k(Nu_i)\|_{L_t^2 L_x^{\widetilde{r_2}}}, \quad \widetilde{r_2} = \frac{3r_2}{r_2+3}.$$

Then by Lemma 2.1(i),

$$\|\partial_k(Nu_i)\|_{L_t^2 L_x^{\widetilde{r_2}}} \le C\epsilon^{\frac{6-\widetilde{r_2}}{2\widetilde{r_2}}}\|\nabla\partial_k u_i\|_{L_t^2 L_x^2} = C\epsilon^{\frac{6+r_2}{2r_2}}\|\nabla\partial_k u_i\|_{L_t^2 L_x^2}, \tag{3-32}$$

However, it seems that $\|\nabla\partial_k u_i\|_{L_t^2 L_x^2}$ can't be controlled by $J(t)$ when $i = 3$, $k = 1, 2$ because $J(t)$ doesn't contain the L^2 norm of $\nabla^2 u_3$.

To modify this, we will use the idea of anisotropic interpolations. Obviously

$$\|Nu_i\|_{L^{r_2}_x} = \big\| \|Nu_i\|_{L^{r_2}_{x_3}} \big\|_{L^{r_2}_{x_h}},$$

where we abbreviate $\|\cdot\|_{L^p_{x_3}} = \|\cdot\|_{L^p((0,\epsilon))}$. In the sequel, we will also abbreviate $\|\cdot\|_{L^q_{x_h}} = \|\cdot\|_{L^q(T^2)}$. Interpolating through the vertical direction, we have

$$\|Nu_i\|_{L^{r_2}_{x_3}} \leqslant C\|Nu_i\|^{\frac{1}{2}+\frac{1}{r_2}}_{L^2_{x_3}} \|\partial_3 Nu_i\|^{\frac{1}{2}-\frac{1}{r_2}}_{L^2_{x_3}} \leqslant C\epsilon^{\frac{1}{2}+\frac{1}{r_2}} \|\partial_3 Nu_i\|_{L^2_{x_3}}.$$

This implies that

$$\|Nu_i\|_{L^{r_2}_x} \leqslant C\epsilon^{\frac{1}{2}+\frac{1}{r_2}} \big\| \|\partial_3 Nu_i\|_{L^2_{x_3}} \big\|_{L^{r_2}_{x_h}} \leqslant C\epsilon^{\frac{1}{2}+\frac{1}{r_2}} \big\| \|\partial_3 Nu_i\|_{L^{r_2}_{x_h}} \big\|_{L^2_{x_3}}.$$

Interpolating through the horizontal direction, we obtain that

$$\|\partial_3 Nu_i\|_{L^{r_2}_{x_h}} \leqslant Cr_2^{\frac{1}{2}} \|\partial_3 Nu_i\|^{\frac{2}{r_2}}_{L^2_{x_h}} \|\partial_h\partial_3 Nu_i\|^{1-\frac{2}{r_2}}_{L^2_{x_h}} + C\|\partial_3 Nu_i\|_{L^2_{x_h}}.$$

As a result, we have

$$\begin{aligned}\big\| \|\partial_3 Nu_i\|_{L^{r_2}_{x_h}} \big\|_{L^2_{x_3}} &\leqslant Cr_2^{\frac{1}{2}} \Big\| \|\partial_3 Nu_i\|^{\frac{2}{r_2}}_{L^2_{x_h}} \|\partial_h\partial_3 Nu_i\|^{1-\frac{2}{r_2}}_{L^2_{x_h}} \Big\|_{L^2_{x_3}} + C\|\partial_3 Nu_i\|_{L^2_x} \\ &\leqslant Cr_2^{\frac{1}{2}} \|\partial_3 Nu_i\|^{\frac{2}{r_2}}_{L^2_x} \|\partial_h\partial_3 Nu_i\|^{1-\frac{2}{r_2}}_{L^2_x} + C\|\partial_3 Nu_i\|_{L^2_x} \\ &\leqslant Cr_2^{\frac{1}{2}} \epsilon^{\frac{2}{r_2}} \|\partial_{33} Nu_i\|^{\frac{2}{r_2}}_{L^2_x} \|\partial_h\partial_3 Nu_i\|^{1-\frac{2}{r_2}}_{L^2_x} + C\epsilon\|\partial_{33} Nu_i\|_{L^2_x} \\ &\leqslant Cr_2^{\frac{1}{2}} \epsilon^{\frac{2}{r_2}} \|\nabla\partial_3 u_i\|_{L^2_x}.\end{aligned}$$

Therefore, we get the estimate

$$\|Nu_i\|_{L^{r_2}_x} \leqslant Cr_2^{\frac{1}{2}} \epsilon^{\frac{1}{2}+\frac{3}{r_2}} \|\nabla\partial_3 u_i\|_{L^2_x},$$

which yields

$$\|Nu_i\|_{L^2_t L^{r_2}_x} \leqslant Cr_2^{\frac{1}{2}} \epsilon^{\frac{1}{2}+\frac{3}{r_2}} \|\nabla\partial_3 u_i\|_{L^2_t L^2_x} \leqslant Cr_2^{\frac{1}{2}} \epsilon^{\frac{1}{2}+\frac{3}{r_2}} J(t_1) \leqslant Cb^{\frac{1}{2}} \epsilon^{\frac{1}{2}+\frac{3}{\alpha}-\frac{3}{b}} J(t_1).$$

Thus we have

$$\|\widetilde{q_2}\|_{L^2_t L^\alpha_x} \leqslant C\alpha b\epsilon^{\frac{3}{\alpha}-\frac{2}{b}} J(t_1)^2.$$

It follows that

$$K_3 \leqslant C\alpha b\epsilon^{\frac{3}{\alpha}-\frac{2}{b}} J(t_1)^2 K(t_1)^{\alpha-1}. \tag{3-33}$$

For the term K_4, we have

$$\begin{aligned} K_4 &\leqslant \|q_3\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \big\| |u_3|^{\frac{\alpha-2}{2}} \big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}} \big\| N(|u_3|^{\frac{\alpha}{2}}) \big\|_{L_t^2 L_x^6} \\ &\leqslant C \|q_3\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \big\| |u_3|^{\frac{\alpha-2}{2}} \big\|^{\frac{\alpha-2}{\alpha}}_{L_t^\infty L_x^2} \big\| \nabla(|u_3|^{\frac{\alpha}{2}}) \big\|_{L_t^2 L_x^2} \\ &\leqslant C \|q_3\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} K(t_1)^{\alpha-1}. \end{aligned}$$

Next we estimate $\|q_3\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}}$,

$$\|q_3\|_{L_t^2 L_x^{\frac{3\alpha}{\alpha+3}}} \leqslant C \sum_{i=1}^{3} \|\partial_3 u_i\|_{L_t^2 L_x^3} \|u_3\|_{L_t^\infty L_x^\alpha} \leqslant C\epsilon^{\frac{1}{2}} J(t_1) K(t_1),$$

where we have used (3-3) with $a = 3$. According to this estimate, we have

$$K_4 \leqslant C\epsilon^{\frac{1}{2}} J(t_1) K(t_1)^\alpha. \tag{3-34}$$

For the term K_5, using a similar method as for K_2, we have

$$K_5 \leqslant \|q_3\|_{L_t^2 L_x^{r_1}} \big\| |u_3|^{\frac{\alpha-2}{2}} \big\|_{L_t^\infty L_x^{\frac{2\alpha}{\alpha-2}}} \big\| M(|u_3|^{\frac{\alpha}{2}}) \big\|_{L_t^2 L_x^b},$$

where r_1 and b satisfy $\frac{1}{r_1} + \frac{\alpha-2}{2\alpha} + \frac{1}{b} = 1$. According to (3-29) and (3-30), we also have

$$K_5 \leqslant C b^{\frac{1}{2}} \epsilon^{-\frac{b-2}{2b}} \|q_3\|_{L_t^2 L_x^{r_1}} K(t_1)^{\alpha-1}. \tag{3-35}$$

It remains to estimate $\|q_3\|_{L_t^2 L_x^{r_1}}$:

$$\|q_3\|_{L_t^2 L_x^{r_1}} \leqslant C \sum_{i=1}^{3} \|\partial_3 u_i\|_{L_t^2 L_x^{r_3}} \|u_3\|_{L_t^\infty L_x^\alpha} \leqslant C \sum_{i=1}^{3} \|\partial_3 u_i\|_{L_t^2 L_x^{r_3}} K(t_1),$$

where $\frac{1}{r_3} + \frac{1}{\alpha} = \frac{1}{r_1}$. Since r_3 satisfies $\frac{1}{r_3} + \frac{1}{b} = \frac{1}{2}$, we get that $2 < r_3 \leqslant 3$ when $3 \leqslant \alpha \leqslant \widetilde{C}|\ln\epsilon|$ and $b \geqslant 2\alpha$. Thus we can see from (3-3) that

$$\|\partial_3 u_i\|_{L_t^2 L_x^{r_3}} \leqslant C\epsilon^{\frac{6-r_3}{2r_3}} J(t_1).$$

Hence $\|q_3\|_{L_t^2 L_x^{r_1}} \leqslant C\epsilon^{\frac{6-r_3}{2r_3}} J(t_1) K(t_1)$. Then (3-35) yields that

$$K_5 \leqslant C b^{\frac{1}{2}} \epsilon^{\frac{1}{2}-\frac{2}{b}} J(t_1) K(t_1)^\alpha. \tag{3-36}$$

Finally, by summarizing (3-21), (3-26), (3-31), (3-33), (3-34) and (3-36), we have

$$K(t)^\alpha \leqslant C\alpha^2 b\epsilon^{\frac{3}{\alpha}-\frac{2}{b}} J(t)^2 K(t)^{\alpha-1} + C\alpha b^{\frac{1}{2}}\epsilon^{\frac{1}{2}-\frac{2}{b}} J(t)K(t)^\alpha + K(0)^\alpha$$

for all $t \in (0, T_{\max})$. Letting $b = 2\alpha + |\ln\epsilon|$ and C_1 be a sufficiently large constant, we get the following estimate:

$$K(t)^\alpha \leqslant C_1\alpha^2(\alpha+|\ln\epsilon|)\epsilon^{\frac{3}{\alpha}} J(t)^2 K(t)^{\alpha-1} + C_1\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}\epsilon^{\frac{1}{2}} J(t)K(t)^\alpha + K(0)^\alpha.$$

Meanwhile, by (3-20), we have

$$J(t)^2 \leqslant C_1\epsilon^{\frac{1}{2}} J(t)^3 + C_1\epsilon^{\frac{\alpha-3}{\alpha}} K(t)J(t)^2 + J(0)^2. \tag{3-37}$$

Assume that the initial data u_0 satisfies

$$J(0) = \|\nabla u_{0h}\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}},$$

$$K(0) = \|u_{03}\|_{L^\alpha(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{\alpha-3}{\alpha}}}.$$

We claim that

$$J(t) \leqslant \frac{2}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}}, \tag{3-38}$$

$$K(t) \leqslant \frac{2}{C\epsilon^{\frac{\alpha-3}{\alpha}}}, \tag{3-39}$$

for all $t \in (0, T_{\max})$ provided C is sufficiently large. This fact implies that $T_{\max} = \infty$. Our claim can be established by contradiction. Suppose that the claim is not true, then there exists a time $t^* \in (0, T_{\max})$ such that (3-38) and (3-39) hold for all $t \in [0, t^*]$ and

$$J(t^*) = \frac{2}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}}, \tag{3-40}$$

or

$$K(t^*) = \frac{2}{C\epsilon^{\frac{\alpha-3}{\alpha}}}. \tag{3-41}$$

Using (3-37) with $t = t^*$, we get

$$J(t^*)^2 \leqslant J(t^*)^2\left(\frac{2C_1\epsilon^{\frac{1}{2}}}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}} + C_1\epsilon^{\frac{\alpha-3}{\alpha}}\frac{2}{C\epsilon^{\frac{\alpha-3}{\alpha}}}\right) + \frac{1}{C^2\epsilon\alpha^2(\alpha+|\ln\epsilon|)}.$$

Choose C be large enough such that

$$\frac{2C_1\epsilon^{\frac{1}{2}}}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}} + C_1\epsilon^{\frac{\alpha-3}{\alpha}}\frac{2}{C\epsilon^{\frac{\alpha-3}{\alpha}}} < \frac{3}{4}.$$

Then we get

$$J(t^*)^2 < \frac{4}{C^2\epsilon\alpha^2(\alpha+|\ln\epsilon|)},$$

which contradicts (3-40). Similarly we can also prove

$$K(t^*) < \frac{2}{C\epsilon^{\frac{\alpha-3}{\alpha}}},$$

which contradicts (3-41) provided C is sufficiently large. Therefore we establish our claim and finish the proof of Theorem 1.2. □

4. Proof of Theorem 1.1

In this section, we will prove Theorem 1.1. Our proof will be divided into three steps.

First, we consider the solution on a very small time interval $[0, t_0]$. We will prove that $\|\partial_3 u\|_{L^2(T_\epsilon)}$ decay very fast and $\|\partial_h u\|_{L^2(T_\epsilon)}$ should not increase quickly after a very short time. Furthermore, at the time t_0, we have

$$\begin{cases} \|\partial_h u(t_0)\|_{L^2(T_\epsilon)} \leqslant \dfrac{2C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}, \\ \|\partial_3 u(t_0)\|_{L^2(T_\epsilon)} \leqslant \dfrac{2C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}. \end{cases}$$

This implies that

$$\|\nabla u(t_0)\|_{L^2(T_\epsilon)} \leqslant \frac{4C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}.$$

Second, we regard t_0 as the initial time and $\|\nabla u(t_0)\|_{L^2(T_\epsilon)}$ as the initial data. Consider the solution on a small time interval $[t_0, t_1]$. We will prove that at the time t_1, the data will satisfy the condition of Theorem 1.2.

Finally, we regard t_1 as the initial time and apply Theorem 1.2 directly to get a solution on the time interval $[t_1, \infty)$.

After the above three steps, we will obtain a solution on $[0, \infty)$. Now let us expatiate the details of the proof.

Proof.

Step 1: Solution on $[0, t_0]$.

Our first goal is to estimate $\|\partial_3 u\|_{L^2(T_\epsilon)}$. Applying ∂_3 to (1-1), we obtain a new equation

$$\partial_t\partial_3 u - \Delta\partial_3 u + \partial_3(u\cdot\nabla u) + \nabla\partial_3 p = 0. \tag{4-1}$$

Take the L^2 inner product with $\partial_3 u$ in (4-1) to get

$$\begin{aligned}(4\text{-}2)\quad \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_3 u\|^2_{L^2(T_\epsilon)} + \|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)} &= -\int_{T_\epsilon} \partial_3 u\cdot\nabla u\partial_3 u\,\mathrm{d}x\\ &= -\int_{T_\epsilon} \partial_3 u\cdot\nabla Mu\partial_3 u\,\mathrm{d}x - \int_{T_\epsilon} \partial_3 u\cdot\nabla Nu\partial_3 u\,\mathrm{d}x = I_1 + I_2.\end{aligned}$$

For the term I_1, we note that Mu is independent of x_3, thus we rewrite it as

$$(4\text{-}3)\quad I_1 = -\int_{T_\epsilon} \partial_3 u_h\cdot\partial_h Mu\partial_3 u\,\mathrm{d}x = -\int_{T^2}\int_0^\epsilon \partial_3 u_h\cdot\partial_h Mu\partial_3 u\,\mathrm{d}x_3\,\mathrm{d}x_h,$$

where $\partial_3 u_h = (\partial_3 u_1, \partial_3 u_2)$, $\partial_h = (\partial_1, \partial_2)$ and $\mathrm{d}x_h = \mathrm{d}x_1\,\mathrm{d}x_2$. Using Hölder's inequality to the vertical direction, we get that

$$(4\text{-}4)\quad \int_0^\epsilon \partial_3 u_h\cdot\partial_h Mu\partial_3 u\,\mathrm{d}x_3 \leqslant \|\partial_3 u\|_{L^2_{x_3}}\|\partial_h Mu\|_{L^\infty_{x_3}}\|\partial_3 u\|_{L^2_{x_3}},$$

where $\|\partial_h Mu\|_{L^\infty_{x_3}} = |\partial_h Mu|$ since $\partial_h Mu$ is independent of x_3. Then by applying Hölder's inequality to the horizontal direction, we see from (4-3) and (4-4) that

$$\begin{aligned}I_1 &\leqslant \int_{T^2} \|\partial_3 u\|_{L^2_{x_3}}\|\partial_h Mu\|_{L^\infty_{x_3}}\|\partial_3 u\|_{L^2_{x_3}}\,\mathrm{d}x_h \leqslant \big\|\|\partial_3 u\|_{L^2_{x_3}}\big\|^2_{L^4_{x_h}}\big\||\partial_h Mu|\big\|_{L^2_{x_h}}\\ &\leqslant \epsilon^{-\frac{1}{2}}\big\|\|\partial_3 u\|_{L^4_{x_h}}\big\|^2_{L^2_{x_3}}\|\partial_h Mu\|_{L^2(T_\epsilon)}.\end{aligned}$$

Interpolating through the horizontal direction, we have

$$(4\text{-}5)\quad \|\partial_3 u\|_{L^4_{x_h}} \leqslant C\|\partial_3 u\|^{\frac{1}{2}}_{L^2_{x_h}}\|\partial_h\partial_3 u\|^{\frac{1}{2}}_{L^2_{x_h}} + C\|\partial_3 u\|_{L^2_{x_h}}.$$

It then follows that

$$I_1 \leqslant C\epsilon^{-\frac{1}{2}}\big(\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_h\partial_3 u\|_{L^2(T_\epsilon)} + \|\partial_3 u\|^2_{L^2(T_\epsilon)}\big)\|\partial_h Mu\|_{L^2(T_\epsilon)}.$$

Since $\int_0^\epsilon \partial_3 u\,\mathrm{d}x_3 = 0$, by Lemma 2.1(i) we have that

$$(4\text{-}6)\quad \|\partial_3 u\|_{L^2(T_\epsilon)} \leqslant C\epsilon\|\partial_{33} u\|_{L^2(T_\epsilon)}.$$

Hence

$$\begin{aligned}(4\text{-}7)\quad I_1 &\leqslant C\epsilon^{\frac{1}{2}}(\|\partial_{33} u\|_{L^2(T_\epsilon)}\|\partial_h\partial_3 u\|_{L^2(T_\epsilon)} + \|\partial_{33} u\|^2_{L^2(T_\epsilon)})\|\partial_h u\|_{L^2(T_\epsilon)}\\ &\leqslant C\epsilon^{\frac{1}{2}}\|\partial_h u\|_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)}.\end{aligned}$$

For the term I_2, by using Hölder's inequality to the vertical direction, we have

$$\begin{aligned}(4\text{-}8)\quad I_2 &= -\int_{T^2}\int_0^\epsilon \partial_3 u\cdot\nabla Nu\partial_3 u\,\mathrm{d}x_3\,\mathrm{d}x_h\\ &\leqslant \int_{T^2} \|\partial_3 u\|_{L^2_{x_3}}\|\nabla Nu\|_{L^\infty_{x_3}}\|\partial_3 u\|_{L^2_{x_3}}\,\mathrm{d}x_h.\end{aligned}$$

Regarding $\|\nabla Nu\|_{L^\infty_{x_3}}$, by interpolating through the vertical direction, we have

$$\|\nabla Nu\|_{L^\infty_{x_3}} \leqslant C\|\nabla Nu\|^{\frac{1}{2}}_{L^2_{x_3}}\|\nabla\partial_3 Nu\|^{\frac{1}{2}}_{L^2_{x_3}}. \tag{4-9}$$

Then applying Hölder's inequality to the horizontal direction, we see from (4-8) and (4-9) that

$$\begin{aligned}I_2 &\leqslant C\big\|\|\partial_3 u\|_{L^2_{x_3}}\big\|^2_{L^4_{x_h}}\big\|\|\nabla Nu\|^{\frac{1}{2}}_{L^2_{x_3}}\big\|_{L^4_{x_h}}\big\|\|\nabla\partial_3 Nu\|^{\frac{1}{2}}_{L^2_{x_3}}\big\|_{L^4_{x_h}}\\ &\leqslant C\big\|\|\partial_3 u\|_{L^4_{x_h}}\big\|^2_{L^2_{x_3}}\|\nabla Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}.\end{aligned}$$

To deal with $\|\partial_3 u\|_{L^4_{x_h}}$, we use the same method as (4-5). Therefore

$$\begin{aligned}I_2 &\leqslant C\big(\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_h\partial_3 u\|_{L^2(T_\epsilon)}+\|\partial_3 u\|^2_{L^2(T_\epsilon)}\big)\|\nabla Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\\ &\leqslant C\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_h\partial_3 u\|_{L^2(T_\epsilon)}\|\nabla Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\\ &\qquad\qquad+C\|\partial_3 u\|^2_{L^2(T_\epsilon)}\|\nabla Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\\ &= I_{21}+I_{22}.\end{aligned}$$

Regarding I_{21}, we have

$$\begin{aligned}I_{21} &\leqslant C\|\partial_3 u\|_{L^2(T_\epsilon)}\|\nabla Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^{\frac{3}{2}}_{L^2(T_\epsilon)}\\ &\leqslant C\epsilon^{\frac{1}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\nabla\partial_3 Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^{\frac{3}{2}}_{L^2(T_\epsilon)}\\ &\leqslant C\epsilon^{\frac{1}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)},\end{aligned}$$

where we have used

$$\|\nabla Nu\|_{L^2(T_\epsilon)} \leqslant C\epsilon\|\nabla\partial_3 u\|_{L^2(T_\epsilon)} \tag{4-10}$$

because of $\int_0^\epsilon \nabla Nu\,\mathrm{d}x_3 = 0$. Regarding I_{22}, by (4-6) and (4-10), we have

$$\begin{aligned}I_{22} &\leqslant C\epsilon^{\frac{3}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_{33} u\|_{L^2(T_\epsilon)}\|\nabla\partial_3 Nu\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\\ &\leqslant C\epsilon^{\frac{3}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)}.\end{aligned}$$

Therefore, we get the following estimate of I_2:

$$I_2 \leqslant C\epsilon^{\frac{1}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)}. \tag{4-11}$$

Summarizing (4-2), (4-7) and (4-11), we obtain the estimate of $\|\partial_3 u\|_{L^2(T_\epsilon)}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_3 u\|^2_{L^2(T_\epsilon)} + \big(2-C\epsilon^{\frac{1}{2}}\|\partial_h u\|_{L^2(T_\epsilon)} - C\epsilon^{\frac{1}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\big)\|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)} \leqslant 0.$$

If

$$C\epsilon^{\frac{1}{2}}\|\partial_h u\|_{L^2(T_\epsilon)} + C\epsilon^{\frac{1}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)} < 1 \tag{4-12}$$

for all $t \in [0, t_0]$, where t_0 is given by (4-18), then we can get

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_3 u\|^2_{L^2(T_\epsilon)} + \|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)} \leqslant 0. \tag{4-13}$$

Integrating from 0 to t, we have

$$\|\partial_3 u\|_{L^\infty_t L^2_x} \leqslant \|\partial_3 u_0\|_{L^2(T_\epsilon)}, \tag{4-14}$$

and

$$\|\nabla\partial_3 u\|_{L^2_t L^2_x} \leqslant \|\partial_3 u_0\|_{L^2(T_\epsilon)}. \tag{4-15}$$

In addition, from (4-6), we get

$$\|\nabla\partial_3 u\|^2_{L^2(T_\epsilon)} \geqslant C^{-1}\epsilon^{-2}\|\partial_3 u\|^2_{L^2(T_\epsilon)}.$$

Hence (4-13) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_3 u\|^2_{L^2(T_\epsilon)} + C^{-1}\epsilon^{-2}\|\partial_3 u\|^2_{L^2(T_\epsilon)} \leqslant 0.$$

This implies that

$$\|\partial_3 u\|^2_{L^2(T_\epsilon)} \leqslant \mathrm{e}^{-C^{-1}\epsilon^{-2}t}\|\partial_3 u_0\|^2_{L^2(T_\epsilon)}. \tag{4-16}$$

Assume that the initial data satisfies

$$\|\partial_3 u_0\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C_0\epsilon^{\frac{1}{2}}}. \tag{4-17}$$

Let

$$t_0 = 3C_2\epsilon^2 \ln|\ln\epsilon|, \tag{4-18}$$

where C_2 is the constant C on the right-hand side of (4-16). Then when $t \geqslant t_0$, we have

$$\|\partial_3 u\|_{L^2(T_\epsilon)} \leqslant \frac{2C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}. \tag{4-19}$$

Next, we want to estimate $\|\partial_h u\|_{L^2(T_\epsilon)}$. Similarly, applying ∂_h to (1-1), we get

$$\partial_t\partial_h u - \Delta\partial_h u + \partial_h(u\cdot\nabla u) + \nabla\partial_h p = 0. \tag{4-20}$$

Taking the L^2 inner product with $\partial_h u$ in (4-20), we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_h u\|^2_{L^2(T_\epsilon)} + \|\nabla\partial_h u\|^2_{L^2(T_\epsilon)} = -\int_{T_\epsilon} \partial_h u\cdot\nabla u\,\partial_h u\,\mathrm{d}x = -\int_{T_\epsilon} \partial_h u_3\,\partial_3 u\,\partial_h u\,\mathrm{d}x - \int_{T_\epsilon} \partial_h u_{\bar h}\cdot\partial_{\bar h} u\,\partial_h u\,\mathrm{d}x = I_3 + I_4, \tag{4-21}$$

where $u_{\bar h} = (u_1, u_2)$ and $\partial_{\bar h} = (\partial_1, \partial_2)$. For the term I_3, we rewrite it as

$$I_3 = -\int_{T_\epsilon} \partial_h M u_3\,\partial_3 u\,\partial_h M u\,\mathrm{d}x - \int_{T_\epsilon} \partial_h M u_3\,\partial_3 u\,\partial_h N u\,\mathrm{d}x - \int_{T_\epsilon} \partial_h N u_3\,\partial_3 u\,\partial_h M u\,\mathrm{d}x - \int_{T_\epsilon} \partial_h N u_3\,\partial_3 u\,\partial_h N u\,\mathrm{d}x = I_{31} + I_{32} + I_{33} + I_{34}.$$

Regarding I_{31}, we have

$$I_{31} \leqslant \|\partial_h M u_3\|_{L^4(T_\epsilon)}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_h M u\|_{L^4(T_\epsilon)} \leqslant \|\partial_h M u\|^2_{L^4(T_\epsilon)}\|\partial_3 u\|_{L^2(T_\epsilon)}.$$

Since u satisfies the periodic boundary condition, we know that

$$\int_{T_\epsilon} \partial_h u\,\mathrm{d}x = 0.$$

Hence by Lemma 2.1(ii) with $a = 4$, we have

$$\|\partial_h M u\|_{L^4(T_\epsilon)} \leqslant C\epsilon^{-\frac{1}{4}}\|\partial_h u\|^{\frac{1}{2}}_{L^2(T_\epsilon)}\|\nabla\partial_h u\|^{\frac{1}{2}}_{L^2(T_\epsilon)}.$$

Therefore

$$\begin{aligned} I_{31} &\leqslant C\epsilon^{-\frac{1}{2}}\|\partial_h u\|_{L^2(T_\epsilon)}\|\nabla\partial_h u\|_{L^2(T_\epsilon)}\|\partial_3 u\|_{L^2(T_\epsilon)} \\ &\leqslant C\epsilon^{-1}\|\partial_h u\|^2_{L^2(T_\epsilon)}\|\partial_3 u\|^2_{L^2(T_\epsilon)} + \frac{1}{8}\|\nabla\partial_h u\|^2_{L^2(T_\epsilon)} \\ &\leqslant C\epsilon\|\partial_h u\|^2_{L^2(T_\epsilon)}\|\partial_{33} u\|^2_{L^2(T_\epsilon)} + \frac{1}{8}\|\nabla\partial_h u\|^2_{L^2(T_\epsilon)}, \end{aligned} \tag{4-22}$$

where we have used (4-6). Regarding I_{32}, we have

$$I_{32} \leqslant \|\partial_h M u\|_{L^3(T_\epsilon)}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_h N u\|_{L^6(T_\epsilon)}.$$

Hence by Lemma 2.1(ii) with $a = 3$ and Lemma 2.1(i) with $a = 6$, we have

$$\|\partial_h M u\|_{L^3(T_\epsilon)} \leqslant C\epsilon^{-\frac{1}{6}}\|\partial_h u\|^{\frac{2}{3}}_{L^2(T_\epsilon)}\|\nabla\partial_h u\|^{\frac{1}{3}}_{L^2(T_\epsilon)},$$

and

$$\|\partial_h N u\|_{L^6(T_\epsilon)} \leqslant C\|\nabla\partial_h u\|_{L^2(T_\epsilon)}.$$

Thus

$$I_{32} \leqslant C\epsilon^{-\frac{1}{6}} \|\partial_h u\|_{L^2(T_\epsilon)}^{\frac{2}{3}} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\nabla \partial_h u\|_{L^2(T_\epsilon)}^{\frac{4}{3}}$$
$$\leqslant C\epsilon^{-\frac{1}{2}} \|\partial_h u\|_{L^2(T_\epsilon)}^2 \|\partial_3 u\|_{L^2(T_\epsilon)}^3 + \frac{1}{8}\|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2$$
$$\leqslant C\epsilon^{\frac{3}{2}} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\partial_{33} u\|_{L^2(T_\epsilon)}^2 \|\partial_h u\|_{L^2(T_\epsilon)}^2 + \frac{1}{8}\|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2. \tag{4-23}$$

The estimate of I_{33} is as same as I_{32}, i.e.,

$$I_{33} \leqslant C\epsilon^{\frac{3}{2}} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\partial_{33} u\|_{L^2(T_\epsilon)}^2 \|\partial_h u\|_{L^2(T_\epsilon)}^2 + \frac{1}{8}\|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2. \tag{4-24}$$

Regarding I_{34}, we have

$$I_{34} \leqslant \|\partial_h N u\|_{L^4(T_\epsilon)} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\partial_h N u\|_{L^4(T_\epsilon)}.$$

By using Lemma 2.1(i) with $a = 4$, we have

$$\|\partial_h N u\|_{L^4(T_\epsilon)} \leqslant C\epsilon^{\frac{1}{4}} \|\nabla \partial_h u\|_{L^2(T_\epsilon)}.$$

Thus, by (4-14), we obtain that

$$I_{34} \leqslant C\epsilon^{\frac{1}{2}} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2 \leqslant C\epsilon^{\frac{1}{2}} \|\partial_3 u_0\|_{L^2(T_\epsilon)} \|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2. \tag{4-25}$$

Consequently, summarizing (4-22)–(4-25), we get the estimate of I_3,

$$I_3 \leqslant C(\epsilon \|\partial_{33} u\|_{L^2(T_\epsilon)}^2 + \epsilon^{\frac{3}{2}} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\partial_{33} u\|_{L^2(T_\epsilon)}^2) \|\partial_h u\|_{L^2(T_\epsilon)}^2$$
$$+ C\epsilon^{\frac{1}{2}} \|\partial_3 u_0\|_{L^2(T_\epsilon)} \|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2 + \frac{3}{8}\|\nabla \partial_h u\|_{L^2(T_\epsilon)}^2. \tag{4-26}$$

For the term I_4, we rewrite it as

$$I_4 = -\int_{T_\epsilon} \partial_h u_{\bar{h}} \cdot \partial_{\bar{h}} u \partial_h u \, \mathrm{d}x$$
$$= -\int_{T_\epsilon} \partial_h u_{\bar{h}} \cdot \partial_{\bar{h}} M u \partial_h u \, \mathrm{d}x - \int_{T_\epsilon} \partial_h u_{\bar{h}} \cdot \partial_{\bar{h}} N u \partial_h u \, \mathrm{d}x = I_{41} + I_{42}.$$

Regarding I_{41}, by using Hölder's inequality to the vertical direction and the horizontal direction respectively, we get that

$$I_{41} = -\int_{T^2} \int_0^\epsilon \partial_h u_{\bar{h}} \cdot \partial_{\bar{h}} M u \partial_h u \, \mathrm{d}x_3 \, \mathrm{d}x_h$$
$$\leqslant \int_{T^2} \|\partial_h u\|_{L^2_{x_3}} \|\partial_h M u\|_{L^\infty_{x_3}} \|\partial_h u\|_{L^2_{x_3}} \, \mathrm{d}x_h$$
$$\leqslant \big\| \|\partial_h u\|_{L^2_{x_3}} \big\|_{L^4_{x_h}}^2 \|\partial_h M u\|_{L^2_{x_h}} \leqslant \epsilon^{-\frac{1}{2}} \big\| \|\partial_h u\|_{L^4_{x_h}} \big\|_{L^2_{x_3}}^2 \|\partial_h M u\|_{L^2(T_\epsilon)}.$$

Interpolating through the horizontal direction together with $\int_{T^2} \partial_h u \, \mathrm{d}x_h = 0$, we have

$$\|\partial_h u\|_{L^4_{x_h}} \leqslant C \|\partial_h u\|^{\frac{1}{2}}_{L^2_{x_h}} \|\partial_h \partial_h u\|^{\frac{1}{2}}_{L^2_{x_h}}. \tag{4-27}$$

Thus

$$\begin{aligned} I_{41} &\leqslant C\epsilon^{-\frac{1}{2}} (\|\partial_h u\|_{L^2(T_\epsilon)} \|\partial_h \partial_h u\|_{L^2(T_\epsilon)}) \|\partial_h u\|_{L^2(T_\epsilon)} \\ &\leqslant C\epsilon^{-1} \|\partial_h u\|^4_{L^2(T_\epsilon)} + \frac{1}{8} \|\nabla \partial_h u\|^2_{L^2(T_\epsilon)}. \end{aligned} \tag{4-28}$$

Regarding I_{42}, by using Hölder's inequality to the vertical direction, we have

$$\begin{aligned} I_{42} &= -\int_{T^2} \int_0^\epsilon \partial_h u_{\bar{h}} \cdot \partial_{\bar{h}} N u \partial_h u \, \mathrm{d}x_3 \, \mathrm{d}x_h \\ &\leqslant \int_{T^2} \|\partial_h u\|_{L^2_{x_3}} \|\partial_h N u\|_{L^\infty_{x_3}} \|\partial_h u\|_{L^2_{x_3}} \, \mathrm{d}x_h. \end{aligned}$$

Interpolating through the vertical direction, we have

$$\|\partial_h N u\|_{L^\infty_{x_3}} \leqslant C \|\partial_h N u\|^{\frac{1}{2}}_{L^2_{x_3}} \|\partial_3 \partial_h N u\|^{\frac{1}{2}}_{L^2_{x_3}}.$$

Then by using Hölder's inequality to the horizontal direction, we get

$$\begin{aligned} I_{42} &\leqslant C \big\| \|\partial_h u\|_{L^2_{x_3}} \big\|^2_{L^4_{x_h}} \big\| \|\partial_h N u\|^{\frac{1}{2}}_{L^2_{x_3}} \big\|_{L^4_{x_h}} \big\| \|\partial_h \partial_3 N u\|^{\frac{1}{2}}_{L^2_{x_3}} \big\|_{L^4_{x_h}} \\ &\leqslant C \big\| \|\partial_h u\|_{L^4_{x_h}} \big\|^2_{L^2_{x_3}} \|\partial_h N u\|^{\frac{1}{2}}_{L^2(T_\epsilon)} \|\partial_h \partial_3 N u\|^{\frac{1}{2}}_{L^2(T_\epsilon)}. \end{aligned}$$

By (4-27) and Lemma 2.1(i) with $a = 2$, we have

$$\begin{aligned} I_{42} &\leqslant C \|\partial_h u\|_{L^2(T_\epsilon)} \|\partial_h \partial_h u\|_{L^2(T_\epsilon)} \|\partial_h N u\|^{\frac{1}{2}}_{L^2(T_\epsilon)} \|\partial_h \partial_3 N u\|^{\frac{1}{2}}_{L^2(T_\epsilon)} \\ &\leqslant C\epsilon^{\frac{1}{2}} \|\partial_h u\|_{L^2(T_\epsilon)} \|\partial_h \partial_h u\|_{L^2(T_\epsilon)} \|\partial_h \partial_3 N u\|_{L^2(T_\epsilon)} \\ &\leqslant C\epsilon^{\frac{1}{2}} \|\partial_h u\|_{L^2(T_\epsilon)} \|\nabla \partial_h u\|^2_{L^2(T_\epsilon)}. \end{aligned} \tag{4-29}$$

Consequently, summarizing (4-28) and (4-29), we get the estimate of I_4,

$$\begin{aligned} I_4 \leqslant C\epsilon^{-1} \|\partial_h u\|^2_{L^2(T_\epsilon)} \|\partial_h u\|^2_{L^2(T_\epsilon)} + C\epsilon^{\frac{1}{2}} \|\partial_h u\|_{L^2(T_\epsilon)} \|\nabla \partial_h u\|^2_{L^2(T_\epsilon)} \\ + \frac{1}{8} \|\nabla \partial_h u\|^2_{L^2(T_\epsilon)}. \end{aligned} \tag{4-30}$$

Finally, combining (4-21), (4-26) and (4-30), we get that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_h u\|_{L^2(T_\epsilon)}^2+\Big(\frac{3}{2}-C\epsilon^{\frac{1}{2}}\|\partial_3 u_0\|_{L^2(T_\epsilon)}-C\epsilon^{\frac{1}{2}}\|\partial_h u\|_{L^2(T_\epsilon)}\Big)\|\nabla\partial_h u\|_{L^2(T_\epsilon)}^2 \leqslant C(\epsilon\|\partial_{33}u\|_{L^2(T_\epsilon)}^2+\epsilon^{\frac{3}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_{33}u\|_{L^2(T_\epsilon)}^2 +\epsilon^{-1}\|\partial_h u\|_{L^2(T_\epsilon)}^2)\|\partial_h u\|_{L^2(T_\epsilon)}^2.$$

Assuming that the initial data satisfies

$$\|\partial_3 u_0\|_{L^2(T_\epsilon)}\leqslant\frac{1}{C_0\epsilon^{\frac{1}{2}}}, \tag{4-31}$$

we have

$$C\epsilon^{\frac{1}{2}}\|\partial_3 u_0\|_{L^2(T_\epsilon)}\leqslant\frac{C}{C_0}<\frac{1}{4}$$

provided C_0 is sufficiently large. If

$$C\epsilon^{\frac{1}{2}}\|\partial_h u\|_{L^2(T_\epsilon)}<\frac{1}{4} \tag{4-32}$$

for all $t\in[0,t_0]$, where t_0 is given by (4-18), then we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_h u\|_{L^2(T_\epsilon)}^2+\|\nabla\partial_h u\|_{L^2(T_\epsilon)}^2 \leqslant C(\epsilon\|\partial_{33}u\|_{L^2(T_\epsilon)}^2+\epsilon^{\frac{3}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_{33}u\|_{L^2(T_\epsilon)}^2 +\epsilon^{-1}\|\partial_h u\|_{L^2(T_\epsilon)}^2)\|\partial_h u\|_{L^2(T_\epsilon)}^2.$$

Using Gronwall's inequality, we get that

$$\|\partial_h u\|_{L^2(T_\epsilon)}^2+\int_0^t\|\nabla\partial_h u\|_{L^2(T_\epsilon)}^2\,\mathrm{d}s\leqslant \mathrm{e}^{G(t)}\|\partial_h u_0\|_{L^2(T_\epsilon)}^2,$$

where

$$G(t)=\int_0^t g(s)\,\mathrm{d}s$$

and

$$g(t)=C(\epsilon\|\partial_{33}u\|_{L^2(T_\epsilon)}^2+\epsilon^{\frac{3}{2}}\|\partial_3 u\|_{L^2(T_\epsilon)}\|\partial_{33}u\|_{L^2(T_\epsilon)}^2+\epsilon^{-1}\|\partial_h u\|_{L^2(T_\epsilon)}^2).$$

Our next goal is to show $G(t)$ can be very small when $t\in[0,t_0]$, where t_0 is given by (4-18). Then we will obtain

$$\|\partial_h u(t)\|_{L^2(T_\epsilon)}\leqslant 2\|\partial_h u_0\|_{L^2(T_\epsilon)}. \tag{4-33}$$

We write $G(t)$ as $G_1(t) + G_2(t) + G_3(t)$, where

$$G_1(t) = \int_0^t C\epsilon \|\partial_{33}u\|^2_{L^2(T_\epsilon)} \,\mathrm{d}s,$$
$$G_2(t) = \int_0^t C\epsilon^{\frac{3}{2}} \|\partial_3 u\|_{L^2(T_\epsilon)} \|\partial_{33}u\|^2_{L^2(T_\epsilon)} \,\mathrm{d}s,$$
$$G_3(t) = \int_0^t C\epsilon^{-1} \|\partial_h u\|^2_{L^2(T_\epsilon)} \,\mathrm{d}s.$$

For the term $G_1(t)$, we conclude from (4-15) and (4-31) that

$$G_1(t) \leqslant C\epsilon \|\partial_3 u_0\|^2_{L^2(T_\epsilon)} \leqslant \frac{C}{C_0^2}.$$

For the term $G_2(t)$, by (4-14), (4-15) and (4-31), we have

$$G_2(t) \leqslant C\epsilon^{\frac{3}{2}} \|\partial_3 u\|_{L^\infty_t L^2_x} \|\partial_{33}u\|^2_{L^2_t L^2_x} \leqslant C\epsilon^{\frac{3}{2}} \|\partial_3 u_0\|^3_{L^2(T_\epsilon)} \leqslant \frac{C}{C_0^3}.$$

For the term $G_3(t)$, by (4-33), we have

$$G_3(t) \leqslant C\epsilon^{-1} \|\partial_h u\|^2_{L^\infty_t L^2_x} t_0 \leqslant C\epsilon^{-1} \|\partial_h u_0\|^2_{L^2(T_\epsilon)} t_0.$$

Assume that the initial data satisfies

$$\|\partial_h u_0\|_{L^2(T_\epsilon)} \leqslant \frac{C^*}{\epsilon^{\frac{1}{2}} |\ln\epsilon|^{\frac{3}{2}}}. \tag{4-34}$$

Then

$$G_3(t) \leqslant \frac{3C_1 C (C^*)^2 \ln|\ln\epsilon|}{|\ln\epsilon|^3} \to 0 \quad \text{as } \epsilon \to 0.$$

Based on the above analysis, we can take $C_0 > 1$ and $0 < \epsilon_1 < 1$ such that for every $\epsilon \in (0, \epsilon_1)$, there holds $G(t) < 1$ for all $t \in [0, t_0]$. Then (4-33) holds for all $t \in [0, t_0]$. By (4-14), (4-17), (4-33) and (4-34), we can take $\epsilon_2 \in (0, 1)$ such that for every $\epsilon \in (0, \epsilon_2)$, conditions (4-12) and (4-32) hold for all $t \in [0, t_0]$. Therefore, we completed the a prior estimate. Additionally, by (4-19) and (4-33), we get that at t_0, there hold

$$\|\partial_3 u(t_0)\|_{L^2(T_\epsilon)} \leqslant \frac{2C^*}{\epsilon^{\frac{1}{2}} |\ln\epsilon|^{\frac{3}{2}}}$$

and

$$\|\partial_h u(t_0)\|_{L^2(T_\epsilon)} \leqslant \frac{2C^*}{\epsilon^{\frac{1}{2}} |\ln\epsilon|^{\frac{3}{2}}}.$$

Step 2: Solution on $[t_0, t_1]$.

We consider the solution from t_0. At this time, $\|\nabla u(t_0)\|_{L^2(T_\epsilon)}$ satisfies

$$\|\nabla u(t_0)\|_{L^2(T_\epsilon)} \leqslant \frac{4C^*}{\epsilon^{\frac{1}{2}}|\ln \epsilon|^{\frac{3}{2}}}.$$

In what follows we will estimate $\|\nabla u(t)\|_{L^2(T_\epsilon)}$ for $t \in [t_0, t_0 + T]$, where T will be given by (4-36). We emphasize that in [Kukavica and Ziane 2007], the authors proved the case when C^* is sufficient small. In our case, C^* can be arbitrarily large. Below, we will show that it can be proved by using the same method as [Kukavica and Ziane 2007].

Take the L^2 inner product with $-\Delta u$ in (1-1) to obtain

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla u\|^2_{L^2(T_\epsilon)} + \|\Delta u\|^2_{L^2(T_\epsilon)} \\
&\qquad = \int_{T_\epsilon} u \cdot \nabla u \Delta u \,\mathrm{d}x = -\int_{T_\epsilon} \nabla u \cdot \nabla u \nabla u \,\mathrm{d}x \\
&\qquad = -\int_{T_\epsilon} \nabla Mu \cdot \nabla Mu \nabla u \,\mathrm{d}x - \int_{T_\epsilon} \nabla Mu \cdot \nabla Nu \nabla u \,\mathrm{d}x \\
&\qquad\qquad\qquad - \int_{T_\epsilon} \nabla Nu \cdot \nabla Mu \nabla u \,\mathrm{d}x - \int_{T_\epsilon} \nabla Nu \cdot \nabla Nu \nabla u \,\mathrm{d}x \\
&\qquad = L_1 + L_2 + L_3 + L_4.
\end{aligned}$$

For L_1, by using Hölder's inequality and Lemma 2.1(ii), we have

$$\begin{aligned}
L_1 &\leqslant \|\nabla Mu\|^2_{L^4(T_\epsilon)}\|\nabla u\|_{L^2(T_\epsilon)} \leqslant C\epsilon^{-\frac{1}{2}}\|\nabla u\|^2_{L^2(T_\epsilon)}\|\Delta u\|_{L^2(T_\epsilon)} \\
&\leqslant C\epsilon^{-1}\|\nabla u\|^4_{L^2(T_\epsilon)} + \frac{1}{4}\|\Delta u\|^2_{L^2(T_\epsilon)}.
\end{aligned}$$

For L_2, we have

$$\begin{aligned}
L_2 &\leqslant \|\nabla Mu\|_{L^3(T_\epsilon)}\|\nabla Nu\|_{L^6(T_\epsilon)}\|\nabla u\|_{L^2(T_\epsilon)} \\
&\leqslant C\epsilon^{-\frac{1}{6}}\|\nabla u\|^{\frac{2}{3}}_{L^2(T_\epsilon)}\|\Delta u\|^{\frac{1}{3}}_{L^2(T_\epsilon)}\|\Delta u\|_{L^2(T_\epsilon)}\|\nabla u\|_{L^2(T_\epsilon)} \\
&\leqslant C\epsilon^{-\frac{1}{6}}\|\nabla u\|^{\frac{5}{3}}_{L^2(T_\epsilon)}\|\Delta u\|^{\frac{4}{3}}_{L^2(T_\epsilon)} \leqslant C\epsilon^{-\frac{1}{2}}\|\nabla u\|^5_{L^2(T_\epsilon)} + \frac{1}{4}\|\Delta u\|^2_{L^2(T_\epsilon)}.
\end{aligned}$$

In the same way, we see that

$$L_3 \leqslant C\epsilon^{-\frac{1}{2}}\|\nabla u\|^5_{L^2(T_\epsilon)} + \frac{1}{4}\|\Delta u\|^2_{L^2(T_\epsilon)}.$$

For L_4, by using Hölder's inequality and Lemma 2.1(i), we obtain

$$\begin{aligned} L_4 &\leqslant \|\nabla Nu\|_{L^3(T_\epsilon)}\|\nabla Nu\|_{L^6(T_\epsilon)}\|\nabla u\|_{L^2(T_\epsilon)} \\ &\leqslant C\|\nabla u\|_{L^2(T_\epsilon)}^{\frac{1}{2}}\|\Delta u\|_{L^2(T_\epsilon)}^{\frac{1}{2}}\|\Delta u\|_{L^2(T_\epsilon)}\|\nabla u\|_{L^2(T_\epsilon)} \\ &\leqslant C\|\nabla u\|_{L^2(T_\epsilon)}^{\frac{3}{2}}\|\Delta u\|_{L^2(T_\epsilon)}^{\frac{3}{2}} \leqslant C\|\nabla u\|_{L^2(T_\epsilon)}^6 + \frac{1}{4}\|\Delta u\|_{L^2(T_\epsilon)}^2. \end{aligned}$$

As a result, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla u\|_{L^2(T_\epsilon)}^2 + \|\Delta u\|_{L^2(T_\epsilon)}^2 \leqslant C\epsilon^{-1}\|\nabla u\|_{L^2(T_\epsilon)}^4 + C\|\nabla u\|_{L^2(T_\epsilon)}^6, \tag{4-35}$$

where we have used

$$\begin{aligned} C\epsilon^{-\frac{1}{2}}\|\nabla u\|_{L^2(T_\epsilon)}^5 &= C\epsilon^{-\frac{1}{2}}\|\nabla u\|_{L^2(T_\epsilon)}^2\|\nabla u\|_{L^2(T_\epsilon)}^3 \\ &\leqslant C\epsilon^{-1}\|\nabla u\|_{L^2(T_\epsilon)}^4 + C\|\nabla u\|_{L^2(T_\epsilon)}^6. \end{aligned}$$

Applying Gronwall's inequality to (4-35), we get

$$\|\nabla u(t)\|_{L^2(T_\epsilon)}^2 + \int_{t_0}^t \|\Delta u\|_{L^2(T_\epsilon)}^2\,\mathrm{d}s \leqslant \mathrm{e}^{H(t)}\|\nabla u(t_0)\|_{L^2(T_\epsilon)}^2, \quad t \in (t_0, t_0+T].$$

where

$$H(t) = \int_{t_0}^t C_3\epsilon^{-1}\|\nabla u\|_{L^2(T_\epsilon)}^2 + C_3\|\nabla u\|_{L^2(T_\epsilon)}^4\,\mathrm{d}s,$$

C_3 is the constant C on the right-hand side of (4-35) and

$$T = \min\left\{\frac{\epsilon^2|\ln\epsilon|^3}{128C_3(C^*)^2}, \frac{\epsilon^2|\ln\epsilon|^6}{2\times 64^2 C_3(C^*)^4}\right\}.$$

Take $\epsilon_3 \in (0,1)$ such that for every $\epsilon \in (0,\epsilon_3)$, we have

$$64(C^*)^2 \leqslant |\ln\epsilon|^3.$$

Then we get

$$T = \frac{\epsilon^2|\ln\epsilon|^3}{128C_3(C^*)^2}, \tag{4-36}$$

and $H(t) \leqslant 1$ for $t \in (t_0, t_0+T]$. Consequently, there hold

$$\|\nabla u(t)\|_{L^2(T_\epsilon)} \leqslant 2\|\nabla u(t_0)\|_{L^2(T_\epsilon)} \leqslant \frac{8C^*}{\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}, \quad t \in (t_0, t_0+T],$$

and

$$\int_{t_0}^{t_0+T} \|\Delta u\|_{L^2(T_\epsilon)}^2\,\mathrm{d}s \leqslant 4\|\nabla u(t_0)\|_{L^2(T_\epsilon)}^2 \leqslant \frac{64(C^*)^2}{\epsilon|\ln\epsilon|^3}.$$

Hence, there exists $t_1 \in (t_0, t_0 + T)$ such that

$$\|\Delta u(t_1)\|^2_{L^2(T_\epsilon)} \leqslant \frac{C(C^*)^4}{\epsilon^3 |\ln \epsilon|^6}$$

and

$$\|\nabla u(t_1)\|_{L^2(T_\epsilon)} \leqslant \frac{8C^*}{\epsilon^{\frac{1}{2}} |\ln \epsilon|^{\frac{3}{2}}}. \tag{4-37}$$

Now, let us turn to $\|u(t_1)\|_{L^\alpha(T_\epsilon)} \leqslant \|Nu(t_1)\|_{L^\alpha(T_\epsilon)} + \|Mu(t_1)\|_{L^\alpha(T_\epsilon)}$. For $\|Nu(t_1)\|_{L^\alpha(T_\epsilon)}$, we have

$$\begin{aligned}
\|Nu(t_1)\|_{L^\alpha(T_\epsilon)} &\leqslant C\|Nu(t_1)\|^{\frac{1}{4}+\frac{3}{2\alpha}}_{L^2(T_\epsilon)} \|\Delta Nu(t_1)\|^{\frac{3}{4}-\frac{3}{2\alpha}}_{L^2(T_\epsilon)} \\
&\leqslant C\epsilon^{\frac{1}{4}+\frac{3}{2\alpha}} \|\nabla u(t_1)\|^{\frac{1}{4}+\frac{3}{2\alpha}}_{L^2(T_\epsilon)} \|\Delta u(t_1)\|^{\frac{3}{4}-\frac{3}{2\alpha}}_{L^2(T_\epsilon)} \\
&\leqslant C(C^*)^{\frac{7}{4}-\frac{3}{2\alpha}} \epsilon^{\frac{3-\alpha}{\alpha}} |\ln \epsilon|^{-\frac{21}{8}+\frac{9}{4\alpha}} \\
&\leqslant C_4 (C^*)^{\frac{7}{4}} \epsilon^{\frac{3-\alpha}{\alpha}} |\ln \epsilon|^{-\frac{15}{8}},
\end{aligned}$$

since $3 \leqslant \alpha \leqslant \widetilde{C}|\ln \epsilon|$. Take $\epsilon_4 \in (0, 1)$ such that for every $\epsilon \in (0, \epsilon_4)$, we have

$$|\ln \epsilon|^{\frac{15}{8}} \geqslant C C_4 (C^*)^{\frac{7}{4}}.$$

Then we get

$$\|Nu(t_1)\|_{L^\alpha(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{\alpha-3}{\alpha}}}. \tag{4-38}$$

For $\|Mu(t_1)\|_{L^\alpha(T_\epsilon)}$, we have

$$\|Mu(t_1)\|_{L^\alpha(T_\epsilon)} \leqslant \frac{C\alpha^{\frac{1}{2}}}{\epsilon^{\frac{\alpha-2}{2\alpha}}} \|\nabla u(t_1)\|_{L^2(T_\epsilon)} \leqslant \frac{C_5 C^* \alpha^{\frac{1}{2}}}{\epsilon^{\frac{\alpha-1}{\alpha}} |\ln \epsilon|^{\frac{3}{2}}} = \frac{1}{\epsilon^{\frac{\alpha-3}{\alpha}}} \frac{C_5 C^* \alpha^{\frac{1}{2}}}{\epsilon^{\frac{2}{\alpha}} |\ln \epsilon|^{\frac{3}{2}}}.$$

Fix

$$\alpha = 3 + \frac{2|\ln \epsilon|}{\ln|\ln \epsilon|}, \tag{4-39}$$

then we have

$$\frac{C_5 C^* \alpha^{\frac{1}{2}}}{\epsilon^{\frac{2}{\alpha}} |\ln \epsilon|^{\frac{3}{2}}} \leqslant \frac{2C_5 C^*}{(\ln|\ln \epsilon|)^{\frac{1}{2}}}.$$

Take $\epsilon_5 \in (0, 1)$ such that for every $\epsilon \in (0, \epsilon_5)$, we have

$$(\ln|\ln \epsilon|)^{\frac{1}{2}} \geqslant 2C C_5 C^*.$$

Therefore we have

$$\|Mu(t_1)\|_{L^\alpha(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{\alpha-3}{\alpha}}}. \tag{4-40}$$

Moreover, for the fixed α (4-39), we know that

$$\frac{1}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}} \geqslant \frac{\ln|\ln\epsilon|}{C\epsilon^{\frac{1}{2}}|\ln\epsilon|^{\frac{3}{2}}}.$$

Take $\epsilon_6 \in (0,1)$ such that for every $\epsilon \in (0,\epsilon_6)$, we have

$$\ln|\ln\epsilon| \geqslant 8CC^*.$$

Then by (4-37), we see that

$$\|\nabla u(t_1)\|_{L^2(T_\epsilon)} \leqslant \frac{1}{C\epsilon^{\frac{1}{2}}\alpha(\alpha+|\ln\epsilon|)^{\frac{1}{2}}}. \tag{4-41}$$

Step 3: Solution on $[t_1,\infty)$.

We regard t_1 as the initial time. It follows from (4-38), (4-40) and (4-41) that the data at t_1 satisfies the condition of Theorem 1.2. Then the solution on $[t_1,\infty)$ can be proved by a direct use of Theorem 1.2. Thus, taking

$$\epsilon_0 = \min\{\epsilon_1,\epsilon_2,\epsilon_3,\epsilon_4,\epsilon_5,\epsilon_6\},$$

we finish the proof of Theorem 1.1. □

Acknowledgement

The author would like to thank Professor Zhen Lei for his helpful discussions on this work.

References

[Fujita and Kato 1964] H. Fujita and T. Kato, "On the Navier–Stokes initial value problem, I", *Arch. Rational Mech. Anal.* **16** (1964), 269–315. MR Zbl

[Grafakos 2004] L. Grafakos, *Classical and modern Fourier analysis*, Pearson, Upper Saddle River, NJ, 2004. MR Zbl

[Hale and Raugel 1992a] J. K. Hale and G. Raugel, "A damped hyperbolic equation on thin domains", *Trans. Amer. Math. Soc.* **329**:1 (1992), 185–219. MR

[Hale and Raugel 1992b] J. K. Hale and G. Raugel, "Reaction-diffusion equation on thin domains", *J. Math. Pures Appl.* (9) **71**:1 (1992), 33–95. MR

[Hopf 1951] E. Hopf, "Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen", *Math. Nachr.* **4** (1951), 213–231. MR Zbl

[Hou et al. 2008] T. Y. Hou, Z. Lei, and C. Li, "Global regularity of the 3D axi-symmetric Navier–Stokes equations with anisotropic data", *Comm. Partial Differential Equations* **33**:7-9 (2008), 1622–1637. MR Zbl

[Iftimie 1999] D. Iftimie, "The 3D Navier–Stokes equations seen as a perturbation of the 2D Navier–Stokes equations", *Bull. Soc. Math. France* **127**:4 (1999), 473–517. MR Zbl

[Iftimie and Raugel 2001] D. Iftimie and G. Raugel, "Some results on the Navier–Stokes equations in thin 3D domains", *J. Differential Equations* **169**:2 (2001), 281–331. MR

[Kukavica and Ziane 2006] I. Kukavica and M. Ziane, "Regularity of the Navier–Stokes equation in a thin periodic domain with large data", *Discrete Contin. Dyn. Syst.* **16**:1 (2006), 67–86. MR Zbl

[Kukavica and Ziane 2007] I. Kukavica and M. Ziane, "On the regularity of the Navier–Stokes equation in a thin periodic domain", *J. Differential Equations* **234**:2 (2007), 485–506. MR Zbl

[Kukavica et al. 2013] I. Kukavica, W. Rusin, and M. Ziane, "A class of solutions of the Navier–Stokes equations with large data", *J. Differential Equations* **255**:7 (2013), 1492–1514. MR Zbl

[Kukavica et al. 2014] I. Kukavica, W. Rusin, and M. Ziane, "A class of large BMO^{-1} non-oscillatory data for the Navier–Stokes equations", *J. Math. Fluid Mech.* **16**:2 (2014), 293–305. MR Zbl

[Leray 1934] J. Leray, "Sur le mouvement d'un liquide visqueux emplissant l'espace", *Acta Math.* **63**:1 (1934), 193–248. MR Zbl

[Montgomery-Smith 1999] S. Montgomery-Smith, "Global regularity of the Navier–Stokes equation on thin three-dimensional domains with periodic boundary conditions", *Electron. J. Differential Equations* (1999), No. 11, 19. MR Zbl

[Raugel and Sell 1993] G. Raugel and G. R. Sell, "Navier–Stokes equations on thin 3D domains, I: Global attractors and global regularity of solutions", *J. Amer. Math. Soc.* **6**:3 (1993), 503–568. MR Zbl

[Raugel and Sell 1994] G. Raugel and G. R. Sell, "Navier–Stokes equations on thin 3D domains, II: Global regularity of spatially periodic solutions", pp. 205–247 in *Nonlinear partial differential equations and their applications, XI* (Paris, 1989–1991), edited by J.-L. Lions, Res. Notes in Math. **299**, Longman Sci. Tech., Harlow, 1994. MR Zbl

[Temam 1984] R. Temam, *Navier–Stokes equations*, 3rd ed., Studies in Mathematics and its Applications **2**, North-Holland, Amsterdam, 1984. MR Zbl

[Temam and Ziane 1996] R. Temam and M. Ziane, "Navier–Stokes equations in three-dimensional thin domains with various boundary conditions", *Adv. Differential Equations* **1**:4 (1996), 499–546. MR Zbl

Received June 9, 2018. Revised February 23, 2020.

NA ZHAO
INSTITUTE OF APPLIED PHYSICS AND COMPUTATIONAL MATHEMATICS
BEIJING
CHINA
nzhao13@fudan.edu.cn

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 309 No. 1 November 2020

Lie 2-algebras of vector fields 1
DANIEL BERWICK-EVANS and EUGENE LERMAN

Lower regularity solutions of the biharmonic Schrödinger equation in a quarter plane 35
ROBERTO DE A. CAPISTRANO-FILHO, MÁRCIO CAVALCANTE and FERNANDO A. GALLEGO

The arithmetic Hodge index theorem and rigidity of dynamical systems over function fields 71
ALEXANDER CARNEY

On the vanishing of the theta invariant and a conjecture of Huneke and Wiegand 103
OLGUR CELIKBAS

Algebraic and geometric properties of flag Bott–Samelson varieties and applications to representations 145
NAOKI FUJITA, EUNJEONG LEE and DONG YOUP SUH

On a modular form of Zaremba's conjecture 195
NIKOLAY G. MOSHCHEVITIN and ILYA D. SHKREDOV

The first nonzero eigenvalue of the p-Laplacian on differential forms 213
SHOO SETO

Global regularity of the Navier–Stokes equations on 3D periodic thin domain with large data 223
NA ZHAO