Pacific Journal of Mathematics

Volume 325 No. 2 August 2023

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 2023 is US $605/year for the electronic version, and $820/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. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.191

TOTALLY GEODESIC HYPERBOLIC 3-MANIFOLDS IN HYPERBOLIC LINK COMPLEMENTS OF TORI IN S^4

MICHELLE CHU AND ALAN W. REID

We prove that certain hyperbolic link complements of 2-tori in S^4 do not contain closed embedded totally geodesic hyperbolic 3-manifolds.

1. Introduction

A classical problem in 4-dimensional topology is (see [14, Problem 3.20]): *Under what conditions does a closed, orientable 3-manifold M smoothly embed in S^4?* As an example of an obstruction, it is an old result of Hantzsche [8] that if a closed orientable 3-manifold M embeds in S^4, then $\mathrm{Tor}(H_1(M, \mathbb{Z})) \cong A \oplus A$ for some finite abelian group A. The focus of this paper is obstructing the embedding of closed hyperbolic 3-manifolds in S^4 via embeddings in hyperbolic link complements of 2-tori in S^4.

A simple but elegant argument (see [9, Proposition 4.10]) shows that if X is a hyperbolic link complement of 2-tori in S^4 then $\chi(X) = \chi(S^4) = 2$, and so there are only finitely many hyperbolic link complements of 2-tori in S^4. This finiteness statement holds more generally for hyperbolic link complements of 2-tori and Klein bottles in any fixed 4-manifold. By way of comparison, Thurston's hyperbolization theorem shows that many links in S^3 have hyperbolic complements, and although it is known that many hyperbolic link complements in S^3 do not contain a closed embedded totally geodesic surface (e.g., alternating links [17]), examples do exist (see [15; 17]). The main result of this paper (see Theorem 1.1 below) provides more examples of hyperbolic link complements of 2-tori in S^4 that do not contain a closed embedded totally geodesic hyperbolic 3-manifold (our previous paper [5] provided one such example). We note that [5] shows that the hyperbolic link complements of 2-tori in S^4 in Theorem 1.1 do contain infinitely many immersed closed totally geodesic hyperbolic 3-manifolds. To state Theorem 1.1 we need to recall some additional notation.

Chu is supported by NSF grant DMS 2243188. Reid was supported by NSF grant DMS 1812397.
MSC2020: primary 57K32, 57K45; secondary 57K40.
Keywords: cusped hyperbolic manifold, arithmetic manifold, totally geodesic submanifold.

Ratcliffe and Tschantz [19] provided a census of 1171 so-called integral congruence two hyperbolic 4-manifolds that are all obtained from face-pairings of the ideal 24-cell in $\mathbb{H}^4$. These are all commensurable cusped, arithmetic, hyperbolic 4-manifolds of Euler characteristic 1 (i.e., minimal volume). Ivanšić [10] provided an example of a cusped, orientable, hyperbolic 4-manifold of Euler characteristic 2 that is the complement of five 2-tori in S^4 (with the standard smooth structure [11]) and is constructed as the orientable double cover of the nonorientable manifold 1011 in the census of integral congruence two hyperbolic 4-manifolds mentioned above. In [5] we proved that this link complement of 2-tori in S^4 does not contain any closed embedded totally geodesic hyperbolic 3-manifolds (it does contain embedded orientable noncompact finite volume totally geodesic hyperbolic 3-manifolds). In [12], four additional examples of link complements of 2-tori in manifolds homeomorphic to S^4 were found. These arise as the orientable double covers of the nonorientable manifolds in the census of [19] with numbers 23, 71, 1091 and 1092. The main result of this note is to extend the result of [5] to these four other examples.

We fix the following notation. For $n \in \{23, 71, 1091, 1092\}$, we denote by $p_n : W_n \to N_n$ the orientation double coverings of the nonorientable integral congruence two hyperbolic 4-manifolds N_n. By construction, $\chi(N_n) = 1$ and $\chi(W_n) = 2$, with W_n a link complement of 2-tori in S^4.

Theorem 1.1. *For $n \in \{23, 71, 1091, 1092\}$ the manifolds W_n do not contain a closed embedded totally geodesic hyperbolic* 3*-manifold.*

As with the case of the manifold 1011 of [19], the W_n of Theorem 1.1 all contain embedded noncompact finite volume totally geodesic hyperbolic 3-manifolds.

The strategy of the proof of Theorem 1.1 is similar to that of [5] but additional complications arise with these four examples (see Section 2B for a fuller discussion). Moreover, a different argument is needed to handle these cases, and this requires a detailed analysis of possible closed totally geodesic surfaces that can embed in certain arithmetic hyperbolic 3-manifolds that cover the Picard orbifold $\mathbb{H}^3/\mathrm{PSL}(2, \mathbb{Z}[i])$.

We finish the introduction by posing a question prompted by our work:

Question 1.2. *Does there exist a hyperbolic link complement of* 2*-tori in a closed* (*smooth*) *simply connected* 4*-manifold that contains a closed embedded totally geodesic hyperbolic* 3*-manifold*?

Examples of hyperbolic link complements of 2-tori in closed (smooth) simply connected 4-manifolds are given in [12; 20]. Indeed, [20, Theorem 1.2] shows that such link complements exist only in S^4, $\#_r(S^2 \times S^2)$, or $\#_r(\mathbb{CP}^2 \# \overline{\mathbb{CP}^2})$, with $r > 0$. Furthermore, using the examples of [10], examples of link complements of 2-tori in $\#_r(S^2 \times S^2)$ for r even were exhibited in [20] (these cover the link complement of [10]). It is unknown whether there exists a finite volume hyperbolic link complement of 2-tori in $\#_r(\mathbb{CP}^2\#\overline{\mathbb{CP}^2})$, for some $r > 0$.

2. Recap from [5]

The hyperboloid model of $\mathbb{H}^4$ is defined using the quadratic form

$$J = x_1^2 + x_2^2 + x_3^2 + x_4^2 - x_5^2$$

with

$$\mathbb{H}^4 = \{x \in \mathbb{R}^5 : J(x) = -1, x_5 > 0\}$$

equipped with the Riemannian metric induced from the Lorentzian inner product associated to J. The full group of isometries of $\mathbb{H}^4$ is then identified with $\mathrm{O}^+(4, 1)$, the subgroup of

$$\mathrm{O}(4, 1) = \{A \in \mathrm{GL}(5, \mathbb{R}) : A^t J A = J\},$$

preserving the upper sheet of the hyperboloid $J(x) = -1$, and where we abuse notation and use J to denote the symmetric matrix associated to the quadratic form. The full group of orientation-preserving isometries is given by

$$\mathrm{SO}^+(4, 1) = \{A \in \mathrm{O}^+(4, 1) : \det(A) = 1\}.$$

The groups $\mathrm{O}^+(3, 1)$ and $\mathrm{SO}^+(3, 1)$ are defined in a similar manner.

2A. ***Integral congruence two hyperbolic 4-manifolds.*** The manifolds $p_n : W_n \to N_n$ where $n \in \{23, 71, 1091, 1092\}$ of interest to us all arise as face-pairings of the regular ideal 24-cell in $\mathbb{H}^4$ (with all dihedral angles $\pi/2$), and are regular $(\mathbb{Z}/2\mathbb{Z})^4$ covers of the orbifold $\mathbb{H}^4/\Lambda(2)$, where $\Lambda(2)$ is the level two congruence subgroup of the group $\mathrm{O}^+(J, \mathbb{Z}) = \mathrm{O}^+(4, 1) \cap \mathrm{O}(J, \mathbb{Z})$. These manifolds are referred to as *integral congruence two hyperbolic* 4-*manifolds* in [19, Table 1]. It will be useful to describe the $(\mathbb{Z}/2\mathbb{Z})^4$ action, and this is best described in the ball model as follows.

Locate the 24-cell in the ball model of hyperbolic space with vertices

$$(\pm 1, 0, 0, 0),\ (0, \pm 1, 0, 0),\ (0, 0, \pm 1, 0),\ (0, 0, 0, \pm 1) \text{ and } \left(\pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}\right).$$

The four reflections in the coordinate planes of $\mathbb{R}^4$ can be taken as generators of this $(\mathbb{Z}/2\mathbb{Z})^4$ group of isometries. Passing to the hyperboloid model, these reflections are elements of $\Lambda(2)$ and are listed as the first four matrices in [19, page 110]. Following [19] we denote this $(\mathbb{Z}/2\mathbb{Z})^4$ group of isometries by $K < \Lambda(2)$.

As noted in [19] (see also [18]) all of the face-pairings of *any* of the integral congruence two hyperbolic 4-manifolds are invariant under the group K. This implies that each of the coordinate hyperplane cross sections of the 24-cell extends in each of the integral congruence two hyperbolic 4-manifolds to a totally geodesic hypersurface which is the fixed point set of one of the reflections described above.

Following [18] we call these hypersurfaces *cross sections*. As described in [18], these cross sections can be identified with integral congruence two hyperbolic 3-manifolds which are also described in [19]. Moreover, it is possible to use [19] or [18] to identify these in any given example.

Lemma 2.1. (1) *N_{23} has 4 nonorientable cross-sections all isometric to each other.*

(2) *N_{71} has one orientable cross-section isometric to the complement of the link 8^4_2 and three nonorientable cross-sections.*

(3) *N_{1091} has one orientable cross-section isometric to the complement of the link 8^4_2 and three nonorientable cross-sections, two of which are isometric to each other.*

(4) *N_{1092} has two orientable cross-sections isometric to the complement of the link 8^4_2 and two nonorientable cross-sections which are isometric to each other.*

Proof. The proof of this lemma is similar to the proof of [5, Lemma 7.1]. We give a fairly detailed discussion of case (1), and only mention salient points in the remaining cases.

(1) In [19, Table 3] the manifold N_{23} is given by the code `1569A4` which represents the side pairing `111155556666999AAAA4444` for the 24 sides of the ideal 24-cell Q^4. In the notation of [19], the four cross sections have $k_1k_5k_9$ codes `352`, `352`, `156`, `156`, which correspond to the side pairings r_ik_i for the 12 sides of the polytope Q^3 where r_i is the reflection in side i and $k_1 = k_2 = k_3 = k_4$, $k_5 = k_6 = k_7 = k_8$, $k_9 = k_{10} = k_{11} = k_{12}$. Since r_i is a reflection, the side pairing r_ik_i is orientation preserving if and only if the corresponding k_i is orientation reversing. But this happens only if $k_i \in \{1, 2, 4, 7\}$ since then it corresponds to the diagonal matrices with $1 \leftrightarrow \mathrm{diag}(-1, 1, 1, 1)$, $2 \leftrightarrow \mathrm{diag}(1, -1, 1, 1)$, $4 \leftrightarrow \mathrm{diag}(1, 1, -1, 1)$, $7 \leftrightarrow \mathrm{diag}(-1, -1, -1, 1)$. Therefore, all four cross-sections of N_{23} are nonorientable.

From [19, Table 1], we see that the code `156` corresponds to the nonorientable integral congruence two 3-manifold M^3_4 of [19]. As in the proof of [5, Lemma 7.1], it can be checked that code `352` is equivalent to the code `156` via a symmetry of Q^3 (the polyhedron in [5, Figure 2]), and hence determine isometric manifolds.

(2) In [19, Table 3] the manifold N_{71} is given by the code `13EB34`. In the notation of [19], the four cross sections have $k_1k_5k_9$ codes `712`, `152`, `173`, `136`. It can be checked that the 3-manifold with code `712` is orientable and isometric to M^3_{10} of [19], and that the remaining codes determine nonorientable manifolds. Thus, N_{71} has one orientable cross-section, and three nonorientable cross-sections.

(3) In [19, Table 3] the manifold N_{1091} is given by the code `53FF35`. In the notation of [19], the four cross sections have $k_1k_5k_9$ codes `712`, `173`, `173`, `537`. As noted in case (2), the code `712` determines an orientable 3-manifold, and it can be checked that the remaining determine nonorientable ones. Hence, N_{1091} has one orientable cross-section, and three nonorientable cross-sections.

(4) In [19, Table 3] the manifold N_{1092} is given by the code `53FFCA`. In the notation of [19], the four cross sections have $k_1k_5k_9$ codes `765`, `174`, `174`, `537`. Thus, the manifolds with code `174` is orientable and isometric to M^3_{10} of [19], and the remaining two determine nonorientable manifolds. Furthermore, it can be checked that the codes `765` and `537` are equivalent via a symmetry of Q^3 and determine isometric manifolds. Thus, N_{1092} has two isometric orientable cross-sections, and two isometric nonorientable cross-sections. □

In what follows, let A (with orientable double cover A^+) denote the nonorientable manifold given by the code `537`, and similarly let B, C, D, and E (with orientable double covers B^+, C^+, D^+, and E^+) denote the nonorientable manifolds given by the codes `152`, `173`, `136`, and `156` respectively.

2B. ***Volume from tubular neighbourhoods.*** As in [5], to prove Theorem 1.1, we will make use of a result of Basmajian [2] which provides disjoint collars about closed embedded orientable totally geodesic hypersurfaces in hyperbolic manifolds. We state this only for hyperbolic 4-manifolds.

Following [2], let $r(x) = \log\coth(x/2)$, and let $V(r)$ denote the volume of a ball of radius r in $\mathbb{H}^3$. It is noted in [2] that, $V(r) = \omega_3 \int_0^r \sinh^2(r)\,dr$, where ω_3 is the area of the unit sphere in $\mathbb{R}^3$ (i.e., $\omega_3 = 4\pi$).

In [2, pages 213–214], the volume of a tubular neighbourhood of a closed embedded orientable totally geodesic hyperbolic 3-manifold of 3-dimensional hyperbolic volume A in a hyperbolic 4-manifold is given in terms of the 4-dimensional tubular neighbourhood function $c_4(A) = \left(\frac{1}{2}\right)(V \circ r)^{-1}(A)$. Moreover, as noted in [2, Remark 2.1], when the totally geodesic submanifold separates, an improved estimate can be obtained using the tubular neighbourhood function $d_4(A) = \left(\frac{1}{2}\right)(V \circ r)^{-1}(A/2)$ and we record this as follows.

Lemma 2.2. *Let X be an orientable finite volume hyperbolic 4-manifold containing a closed embedded separating orientable totally geodesic hyperbolic 3-manifold of 3-dimensional hyperbolic volume A. Then X contains a tubular neighbourhood of M of volume*

$$\mathcal{V}'(A) = 2A \int_0^{d_4(A)} \cosh^3(t)\,dt.$$

Moreover, [2] also proves that disjoint embedded closed orientable totally geodesic hyperbolic 3-manifolds in an orientable finite volume hyperbolic 4-manifold have disjoint collars, thereby contributing additional volume. For our purposes we summarize what we need in the following.

Corollary 2.3. *Let X be an orientable finite volume hyperbolic 4-manifold of Euler characteristic χ containing K disjoint copies of a closed embedded orientable totally geodesic hyperbolic 3-manifold of 3-dimensional hyperbolic volume A. Assume that all of these disjoint copies separate X. Then*

$$\mathrm{Vol}(X) = \left(\tfrac{4}{3}\pi^2\right)\chi \geq K\mathcal{V}'(A).$$

Given this set up we recall the basic strategy of [5]. To that end, let N (resp. W) denote one of the manifolds N_{23}, N_{71}, N_{1091}, or N_{1092} (resp. W_{23}, W_{71}, W_{1091}, or W_{1092}) and $M \hookrightarrow W$ a closed embedded totally geodesic hyperbolic 3-manifold. Since $W \subset S^4$, M is orientable and the embedding separates S^4.

As in [5, Lemma 7.2] since W is the orientable double cover of N, it is a characteristic cover of N and hence a regular cover of $\mathbb{H}^4/\Lambda(2)$ (using Section 2A). If it can be shown that M is disjoint from the preimages of all of the cross-sections in N, then since W is a regular cover of $\mathbb{H}^4/\Lambda(2)$, using the isometries of W induced from the reflections in the coordinate hyperplanes we get 16 disjoint copies of M, all embedded and separating in W (since it is a submanifold of S^4).

Now the minimal volume of a closed hyperbolic 3-manifold is that of the Weeks manifold and is approximately 0.9427... [7]. Using this estimate for $\mathrm{Vol}(M)$, and applying Corollary 2.3 we see that $\mathrm{Vol}(W) \geq 16\mathcal{V}'(0.94)$, which is approximately 28.9. On the other hand, since $\chi(W) = 2$, $\mathrm{Vol}(W) = \frac{8}{3}\pi^2$ which is approximately 26.3, a contradiction.

As proved in [5, Lemma 3.2], any M (as above) is disjoint from the lift of any orientable cross-section in N. To prove Theorem 1.1 we need to show that M is disjoint from the preimage of a nonorientable cross-section in N. This follows from our next lemma, since if M (as above) was not disjoint this would give rise to a closed embedded totally geodesic (possibly nonorientable) surface in the preimage of the cross-section.

Lemma 2.4. *Let Y be any of the nonorientable cross-sections listed in Lemma 2.1 and Y^+ the orientable double cover. Then Y^+ (and hence Y) does not contain a closed embedded totally geodesic surface.*

The strategy to prove Lemma 2.4 is this: we first identify $\Gamma = \pi_1(Y^+)$ as a congruence subgroup of the Picard group $\mathrm{PSL}(2, \mathbb{Z}[i])$, and identify matrices (up to sign) that correspond to a generating set for Γ. We next use the classification of circles left invariant by nonelementary Fuchsian subgroups of $\mathrm{PSL}(2, \mathbb{Z}[i])$ given in [16] to limit the possibilities for what circles can be associated to a closed embedded totally geodesic surface in Y^+. Finally we use a criterion given by [13, Corollary 3.3] to prove that any candidate totally geodesic surface cannot be embedded.

The proof of Lemma 2.4 occupies the remainder of this paper.

3. The Picard group and the fundamental groups of the cross-sections

The Picard group $\mathrm{PSL}(2, \mathbb{Z}[i])$ has a presentation from [21]:

$$\mathrm{PSL}(2,\mathbb{Z}[i]) = \langle \alpha, l, t, u \mid \alpha^2 = l^2 = (\alpha l)^2 = (tl)^2 = (ul)^2 = (\alpha t)^3 = (u\alpha l)^3 = [t,u] = 1 \rangle,$$

where these generators can be represented by the matrices (up to sign) shown below:

$$\alpha = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad u = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \quad l = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

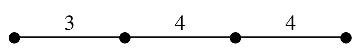
3A. ***Locating subgroups.*** It will be helpful to prove the following result which helps "locate" the fundamental groups of the manifolds Y^+ as subgroups of $\mathrm{PSL}(2, \mathbb{Z}[i])$.

Proposition 3.1. *Let Y be any of the nonorientable cross-sections listed in Lemma 2.1 and Y^+ the orientable double cover. Then $\pi_1(Y^+)$ admits a faithful representation with image group Γ of index* 48 *in* $\mathrm{PSL}(2, \mathbb{Z}[i])$ *and* $[\Gamma(1+i), \Gamma(1+i)] \triangleleft \Gamma \triangleleft \Gamma(1+i)$.

Proof. To establish that $\pi_1(Y^+)$ admits a faithful representation with image group Γ with $\Gamma \triangleleft \Gamma(1+i)$, recall from [19, Section 3] that these integral congruence two hyperbolic 3-manifolds are constructed as follows. As in the case of dimension 4 described above these manifolds arise as regular covers (all with covering group $(\mathbb{Z}/2\mathbb{Z})^3$) of a certain congruence quotient of $\mathbb{H}^3$, namely the subgroup $\Lambda(2) < \mathrm{O}^+(3, 1; \mathbb{Z})$. As shown in [19, page 105]:

$$[\mathrm{O}^+(3, 1; \mathbb{Z}) : \Lambda(2)] = 12 \quad \text{with } \mathrm{O}^+(3, 1; \mathbb{Z})/\Lambda(2) \cong S_3 \times \mathbb{Z}/2\mathbb{Z}.$$

In addition, [19, Section 3] identifies the group $\mathrm{O}^+(3, 1; \mathbb{Z})$ with the Coxeter group T, generated by reflections in the faces of the noncompact tetrahedron with Coxeter diagram:



Using a presentation of this Coxeter group, one can find a presentation for the subgroup $\mathrm{SO}^+(3, 1; \mathbb{Z})$ (consisting of orientation preserving isometries) of index 2 in $\mathrm{O}^+(3, 1; \mathbb{Z})$, and it follows from this that the abelianization of $\mathrm{SO}^+(3, 1; \mathbb{Z})$ is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Indeed, it is known (see, e.g., [4]) that the subgroup $\mathrm{SO}^+(3, 1; \mathbb{Z})$ can be identified with $\mathrm{PGL}(2, \mathbb{Z}[i])$, which in turn contains $\mathrm{PSL}(2, \mathbb{Z}[i])$ of index 2.

Since $\Lambda(2)$ contains elements of determinant -1, the group

$$\Lambda^+(2) = \Lambda(2) \cap \mathrm{SO}^+(3, 1; \mathbb{Z})$$

has index 2 in $\Lambda(2)$, and so $\Lambda^+(2)$ is isomorphic to a normal subgroup of index 12 in $\mathrm{PGL}(2, \mathbb{Z}[i])$ with quotient group $S_3 \times \mathbb{Z}/2\mathbb{Z}$.

As noted above, the abelianization of $\mathrm{SO}^+(3,1;\mathbb{Z}) \cong \mathrm{PGL}(2,\mathbb{Z}[i])$ is $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. We claim that this implies that $\Lambda^+(2)$ is contained in the commutator subgroup of $\mathrm{SO}^+(3,1;\mathbb{Z})$. To see this, first recall that the abelianization of S_3 is $\mathbb{Z}/2\mathbb{Z}$ so we get the epimorphism $a : S_3 \times \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Since the surjective composition

$$\mathrm{SO}^+(3,1;\mathbb{Z}) \xrightarrow{p} S_3 \times \mathbb{Z}/2\mathbb{Z} \xrightarrow{a} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$$

lands in an abelian group, it factors through the abelianization of $\mathrm{SO}^+(3,1;\mathbb{Z})$, with the second map an isomorphism. It then follows that $\Lambda^+(2)$ must be contained in the commutator subgroup of $\mathrm{SO}^+(3,1;\mathbb{Z})$.

Furthermore, via the above identifications, $\Lambda^+(2)$ must be isomorphic to a normal subgroup of index 6 in $\mathrm{PSL}(2,\mathbb{Z}[i])$. By [6, Theorem 2] there is a unique normal subgroup of index 6 in $\mathrm{PSL}(2,\mathbb{Z}[i])$ and it is the principal congruence subgroup $\Gamma(1+i)$ (i.e., those elements in $\mathrm{PSL}(2,\mathbb{Z}[i])$ congruent to the identity modulo the ideal $(1+i)$).

From [19, page 105], the group $\mathrm{O}^+(3,1;\mathbb{Z})$ is generated by the 4 reflections

$$a=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\quad b=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix},\quad c=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix},\quad d=\begin{pmatrix}0&-1&-1&1\\-1&0&-1&1\\-1&-1&0&1\\-1&-1&-1&2\end{pmatrix},$$

and the subgroup $\mathrm{SO}^+(3,1;\mathbb{Z})$ can be identified with the group $\mathrm{PGL}(2,\mathbb{Z}[i])$ via the isomorphism defined by

$$ab \mapsto \begin{pmatrix}1&-1\\1&0\end{pmatrix},\quad ac \mapsto \begin{pmatrix}0&1\\-i&0\end{pmatrix},\quad ad \mapsto \begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

From [19, page 106] the group $\Lambda(2)$ is generated by the reflections $r_1 = abcba$, $r_2 = bcb$, $r_3 = c$, $r_4 = abdcdba$, $r_5 = bdcdb$, $r_6 = dcd$, and the subgroup $\Lambda^+(2)$ can thus be identified with the group $\Gamma(1+i)$ via the induced isomorphism defined by

$$r_1r_2 \mapsto \begin{pmatrix}-1&1+i\\-1+i&1\end{pmatrix},\quad r_1r_3 \mapsto \begin{pmatrix}-i&-1+i\\0&i\end{pmatrix},\quad r_1r_4 \mapsto \begin{pmatrix}-i&2i\\0&i\end{pmatrix},$$
$$r_1r_5 \mapsto \begin{pmatrix}2-i&-1+i\\1-i&i\end{pmatrix},\quad r_1r_6 \mapsto \begin{pmatrix}1&1+i\\0&1\end{pmatrix}.$$

As noted above, $\pi_1(Y) \triangleleft \Lambda(2)$ with quotient group $(\mathbb{Z}/2\mathbb{Z})^3$, so it follows that $\pi_1(Y^+) \triangleleft \Lambda^+(2) \cong \Gamma(1+i)$ with quotient group $(\mathbb{Z}/2\mathbb{Z})^3$.

Now $\Gamma(1+i)/[\Gamma(1+i),\Gamma(1+i)] \cong (\mathbb{Z}/2\mathbb{Z})^5$ (see the Magma [3] routine following this proof). Since the map $\Gamma(1+i) \to \Gamma(1+i)/\pi_1(Y^+) \cong (\mathbb{Z}/2\mathbb{Z})^3$ goes to an abelian group, the commutator subgroup $[\Gamma(1+i),\Gamma(1+i)]$ is sent to 1 and it follows that $[\Gamma(1+i),\Gamma(1+i)] \triangleleft \pi_1(Y^+)$.

$$
\begin{tikzcd}
 & \mathrm{O}^+(3, 1; \mathbb{Z}) \arrow[ddl, dash, "12"'] \arrow[d, dash, "2"] & & \\
 & \mathrm{SO}^+(3, 1; \mathbb{Z}) \arrow[rr, dash, "\cong"] \arrow[dd, dash, "12"] & & \mathrm{PGL}(2, \mathbb{Z}[i]) \arrow[d, dash, "2"] \\
\Lambda(2) \arrow[dr, dash, "2"] & & & \mathrm{PSL}(2, \mathbb{Z}[i]) \arrow[dll, dash, "6"'] \\
 & \Lambda^+(2) \cong \Gamma(1+i) \arrow[d, dash, "2^3"] & & \\
 & \Gamma \arrow[d, dash, "2^2"] & & \\
 & {[\Gamma(1+i), \Gamma(1+i)]} & &
\end{tikzcd}
$$

Figure 1. Lattice of subgroups.

Putting all of this together we obtain the lattice of subgroups shown in Figure 1. We have the image group $\Gamma = \pi_1(Y^+)$ as claimed. That the index in $\mathrm{PSL}(2, \mathbb{Z}[i])$ is 48 is clear from the lattice of subgroups above (or from volume consideration discussed in [19, page 108]). □

Remark 3.2. The subgroup $[\Gamma(1+i), \Gamma(1+i)]$ can also be identified as the principal congruence subgroup $\Gamma(2+2i)$ (i.e., those elements in $\mathrm{PSL}(2, \mathbb{Z}[i])$ congruent to the identity modulo the ideal $(2+2i) = (1+i)^3$). This follows from [1, Propositon 3.1] where the group $\Gamma(2+2i)$ is identified as a link group, and arises as the normal closure in $\mathrm{PSL}(2, \mathbb{Z}[i])$ of the subgroup $\langle t^2u^2, t^4\rangle$. Now $t^2u^2, t^4 \in \Gamma(1+i)$, and since $\Gamma(1+i)^{\mathrm{ab}} \cong (\mathbb{Z}/2\mathbb{Z})^5$, it follows that t^2u^2 and t^4 are both mapped trivially under the abelianization map $\Gamma(1+i) \to \Gamma(1+i)^{\mathrm{ab}}$. Hence $\langle t^2u^2, t^4\rangle \subset [\Gamma(1+i), \Gamma(1+i)]$.

Now the subgroup $[\Gamma(1+i), \Gamma(1+i)]$ is a characteristic subgroup of $\Gamma(1+i)$ and hence is normal in $\mathrm{PSL}(2, \mathbb{Z}[i])$, and so it follows that the normal closure in $\mathrm{PSL}(2, \mathbb{Z}[i])$ of the subgroup $\langle t^2u^2, t^4\rangle$ is contained in $[\Gamma(1+i), \Gamma(1+i)]$; i.e., $\Gamma(2+2i) \subset [\Gamma(1+i), \Gamma(1+i)]$. However, both these groups have index 192 in $\mathrm{PSL}(2, \mathbb{Z}[i])$.

We now provide the short Magma [3] routine referred to in the proof of Proposition 3.1. Referring below, the group g is the group $\mathrm{PSL}(2, \mathbb{Z}[i])$, and the presentation used is that given above. The group `N` is $\Gamma(1+i)$.

```
> g<a, l, t, u> := Group<a, l, t, u | a^2, l^2, (a*l)^2, (t*l)^2, (u*l)^2, (a*t)^3,
(u*a*l)^3, (t, u)>;
> h := sub<g | t*u, t^2, u^2>;
> N := NormalClosure(g, h);
> printIndex(g, N);
6
> printAbelianQuotientInvariants(N);
[2, 2, 2, 2, 2]
```

3B. ***Generators for the fundamental groups of the cross-sections.*** Using the description of each nonorientable cross-section from Lemma 2.1, the notation of Section 2A together with the isomorphism described in the proof of Proposition 3.1, we find generators for the fundamental groups of A, B, C, D and E as side pairings of the 24-cell. We then use Magma to find generators for the orientation double covers A^+, B^+, C^+, D^+ and E^+ and also to eliminate redundant generators. The generators are written below both as elements of $\mathrm{PSL}(2, \mathbb{Z}[i])$ and as words in the reflections r_i's:

$$\pi_1(A^+) = \left\langle \begin{pmatrix} -i & -1+i \\ 1+i & -i \end{pmatrix}, \begin{pmatrix} 3i & 1-i \\ 1+i & -i \end{pmatrix}, \begin{pmatrix} 1-2i & -2-4i \\ -2i & -3-2i \end{pmatrix}, \begin{pmatrix} -1-4i & -4+4i \\ -2-2i & -1+4i \end{pmatrix}, \begin{pmatrix} -2-5i & -1+9i \\ -3-3i & 2+7i \end{pmatrix} \right\rangle$$
$$= \langle r_3r_6r_2r_1, r_1r_3r_6r_2, r_2r_1r_5r_3r_4r_3, r_2r_1r_5r_3r_2r_5r_3r_1, r_2r_1r_5r_1r_6r_2r_3r_5r_1r_2 \rangle,$$

$$\pi_1(B^+) = \left\langle \begin{pmatrix} 3 & -1-i \\ 2-2i & -1 \end{pmatrix}, \begin{pmatrix} -2+i & 1+i \\ 1+i & -i \end{pmatrix}, \begin{pmatrix} 1+2i & 2-2i \\ 4i & 5-2i \end{pmatrix}, \begin{pmatrix} 1+2i & -1-i \\ 4+4i & -3-2i \end{pmatrix}, \begin{pmatrix} -2-i & 1+i \\ -1-i & i \end{pmatrix} \right\rangle$$
$$= \langle r_1r_2r_6r_2, r_3r_4r_3r_2, (r_5r_3r_1)^2, r_5r_3r_2r_6r_2r_1r_3r_5, r_5r_1r_4r_2r_1r_5 \rangle,$$

$$\pi_1(C^+) = \left\langle \begin{pmatrix} 1 & -1-i \\ 2-2i & -3 \end{pmatrix}, \begin{pmatrix} -1 & 1+i \\ -2 & 1+2i \end{pmatrix}, \begin{pmatrix} 3+2i & -2-2i \\ 2+2i & -1-2i \end{pmatrix}, \begin{pmatrix} 1 & 1+i \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -3 & 5-3i \\ -2 & 3-2i \end{pmatrix} \right\rangle$$
$$= \langle r_2r_6r_2r_1, r_2r_5r_3r_1, (r_1r_4r_2)^2, r_1r_4r_6r_1r_4r_1, r_1r_4r_5r_3r_1r_2r_4r_1 \rangle,$$

$$\pi_1(D^+) = \left\langle \begin{pmatrix} 1 & -1-i \\ 2-2i & -3 \end{pmatrix}, \begin{pmatrix} -2+3i & -2i \\ -1+i & -i \end{pmatrix}, \begin{pmatrix} 2+i & -2 \\ 1-i & i \end{pmatrix}, \begin{pmatrix} 2+i & -2i \\ 3+i & -3i \end{pmatrix}, \begin{pmatrix} -3+2i & 4 \\ -5+i & 5+2i \end{pmatrix} \right\rangle$$

$$= \langle r_2r_6r_2r_1,\, r_2r_3r_2r_3r_4r_3r_2r_1r_5r_2,\, r_2r_3r_2r_4r_2r_1r_5r_2,\, r_2r_5r_1r_2r_3r_2r_3r_4r_3r_2,\, r_2r_5r_1r_2r_6r_2r_5r_1\rangle,$$

$$\pi_1(E^+) = \left\langle \begin{pmatrix} 1 & -1-i \\ 2-2i & -3 \end{pmatrix}, \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}, \begin{pmatrix} 1+2i & 2-2i \\ 4i & 5-2i \end{pmatrix}, \begin{pmatrix} -1+2i & 2 \\ -4+2i & 3+2i \end{pmatrix}, \begin{pmatrix} 3+2i & -1-i \\ 4+4i & -1-2i \end{pmatrix} \right\rangle$$

$$= \langle r_2r_6r_2r_1,\, r_4r_3r_2r_1r_3r_5,\, (r_5r_3r_1)^2,\, r_5r_3r_1r_4r_3r_2,\, r_5r_3r_1r_2r_6r_2r_3r_5\rangle.$$

4. Fuchsian subgroups of the Picard group

To understand totally geodesic surfaces in Y^+, we recall from [16] that every nonelementary Fuchsian subgroup of $\mathrm{PSL}(2, \mathbb{Z}[i])$ fixes a circle or straight line $\mathcal{C}$ with an equation of the form: $a|z|^2+\bar{B}z+B\bar{z}+c=0$, with $a, c \in \mathbb{Z}$ and $B \in \mathbb{Z}[i]$, and vice versa. We caution the reader that the normalization of the equation of the circle follows [13] (which we will use in the proof) rather than [16], the normalization of [16] uses $Bz + \bar{B}\bar{z}$.

Two such circles (or straight-lines) $\mathcal{C}$ and $\mathcal{C}'$ are said to be *equivalent* if there exists $\gamma \in \mathrm{PSL}(2, \mathbb{Z}[i])$ such that $\gamma\mathcal{C} = \mathcal{C}'$. Define $D = |B|^2 - ac$ to be the *discriminant* of $\mathcal{C}$. This is preserved by the action of $\mathrm{PSL}(2, \mathbb{Z}[i])$, and hence equivalent circles have the same discriminant (see [16]). If $\Delta < \mathrm{PSL}(2, \mathbb{Z}[i])$ is a torsion-free subgroup of finite index, then a Δ-equivalence class of circles and straight-lines can be associated to a totally geodesic surface in $\mathbb{H}^3/\Delta$ and vice versa. Hence we can also refer to the discriminant of the associated totally geodesic surface. When $a \neq 0$, $\mathcal{C}$ is a circle centered at $-B/a$, with radius $\sqrt{D}/|a|$. This is the case when the totally geodesic surface is closed [13, Lemma 3.1]. Note that if the surface associated to a circle $\mathcal{C}$ is closed and embedded in $\mathbb{H}^3/\Delta$, then for every element $\delta \in \Delta$ we must have $\delta\mathcal{C} = \mathcal{C}$ or $\delta\mathcal{C} \cap \mathcal{C} = \varnothing$.

It is shown in [16] that every circle (or straight-line) as above is $\mathrm{PSL}(2, \mathbb{Z}[i])$-equivalent to one of the following:

- $\mathcal{C}_D : |z|^2 - D = 0$.
- $\mathcal{C}_{D,1} : 2|z|^2 + z + \bar{z} - \frac{1}{2}(D-1) = 0$ (when $D \equiv 1 \pmod 4$).
- $\mathcal{C}_{D,2} : 2|z|^2 + iz - i\bar{z} - \frac{1}{2}(D-1) = 0$ (when $D \equiv 1 \pmod 4$).
- $\mathcal{C}_{D,3} : 2|z|^2 + (1+i)z + (1-i)\bar{z} - \frac{1}{2}(D-2) = 0$ (when $D \equiv 2 \pmod 4$).

The radius of the first circle listed above is $\sqrt{D}$ and for the others it is $\sqrt{D}/2$.

Lemma 4.1. *Let Δ be a normal subgroup of finite index in* $\mathrm{PSL}(2, \mathbb{Z}[i])$, *and assume that $S \hookrightarrow \mathbb{H}^3/\Delta$ is an embedded totally geodesic surface associated to the circle $\mathcal{C}$ of discriminant D. Then there exists an embedded totally geodesic surface S' associated to one of the circles $\mathcal{C}_D$ or $\mathcal{C}_{D,j}$ for one of $j = 1, 2, 3$.*

Proof. Observe that, since $\mathcal{C}$ is associated to an embedded totally geodesic surface in $\mathbb{H}^3/\Delta$, for any $\alpha \in \mathrm{PSL}(2, \mathbb{Z}[i])$, the circle $\alpha\mathcal{C}$ is associated to an embedded totally geodesic surface in $\mathbb{H}^3/\alpha\Delta\alpha^{-1}$. Since Δ is assumed to be normal in $\mathrm{PSL}(2, \mathbb{Z}[i])$, the surface associated to $\alpha\mathcal{C}$ is actually embedded in $\mathbb{H}^3/\Delta$.

From the classification of circles given above there is an $\alpha \in \mathrm{PSL}(2, \mathbb{Z}[i])$ such that $\alpha\mathcal{C}$ is one of $\mathcal{C}_D$ or $\mathcal{C}_{D,j}$ for one of $j = 1, 2, 3$. The result follows. $\square$

Associated to the circle $\mathcal{C}$ with equation $a|z|^2 + \bar{B}z + B\bar{z} + c = 0$ as above, is the Hermitian matrix $A = \left(\begin{smallmatrix} a & B \\ \bar{B} & c \end{smallmatrix}\right)$ with an action of $\mathrm{PSL}(2, \mathbb{Z}[i])$ given by $\gamma^* A \gamma$, where $*$ denotes conjugate-transpose and the given action sends $\mathcal{C}$ to $\gamma^{-1}\mathcal{C}$. Here $\mathcal{C}$ is the set of all points $z \in \mathbb{C}$ such that $A\left(\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} \bar{z} \\ 1 \end{smallmatrix}\right) = 0$. Now [13, Corollary 3.3] provides a criterion for a totally geodesic surface S associated to a circle $\mathcal{C}$ to be embedded in $\mathbb{H}^3/\Delta$ or not; namely if $\gamma \in \Delta$ does not leave $\mathcal{C}$ invariant and satisfies

$$|\mathrm{tr}(\gamma^* A \gamma A^{-1})| < 2,$$

then $\gamma^{-1}\mathcal{C} \cap \mathcal{C} \neq \varnothing$ and S is not embedded. Furthermore, if S is closed, and γ does not leave $\mathcal{C}$ invariant, then $\gamma^{-1}\mathcal{C} \cap \mathcal{C}$ is two points.

We will make use of the following lemma.

Lemma 4.2. *Let Δ be a subgroup of finite index in* $\mathrm{PSL}(2, \mathbb{Z}[i])$ *which contains the group* $[\Gamma(1+i), \Gamma(1+i)]$, *and let $M = \mathbb{H}^3/\Delta$. If $S \hookrightarrow M$ is a closed embedded totally geodesic surface* (*not necessarily orientable*) *associated to a circle $\mathcal{C}$, then there exists $\alpha \in$* $\mathrm{PSL}(2, \mathbb{Z}[i])$ *such that $\alpha\mathcal{C} = \mathcal{C}_{6,3}$.*

Proof. Recall from Remark 3.2 that $[\Gamma(1+i), \Gamma(1+i)] = \Gamma(2+2i)$, and so $\Gamma(2+2i) \subset \Delta$ by hypothesis. Therefore S gives rise to a closed embedded totally geodesic surface associated to the circle $\mathcal{C}$ in the cover $\mathbb{H}^3/\Gamma(2+2i)$. Assuming that $\mathcal{C}$ has discriminant D, there exists $\alpha \in \mathrm{PSL}(2, \mathbb{Z}[i])$ such that $\alpha\mathcal{C}$ is one of $\mathcal{C}_D$ or $\mathcal{C}_{D,j}$ for $j = 1, 2, 3$

Now $\Gamma(2+2i) \lhd \mathrm{PSL}(2, \mathbb{Z}[i])$, and Lemma 4.1 shows that one of $\mathcal{C}_D$ or $\mathcal{C}_{D,j}$ for $j = 1, 2, 3$ also gives rise to a closed embedded totally geodesic surface in $\mathbb{H}^3/\Gamma(2+2i)$.

Now the element $\left(\begin{smallmatrix} 1 & 2+2i \\ 0 & 1 \end{smallmatrix}\right)$ is in $\Gamma(2+2i)$, and using this element, a simple calculation shows that for the surface to be embedded, the radius of the associated circle must be $\leq \sqrt{2}$. From above, the radii of the circles $\mathcal{C}_D$ or $\mathcal{C}_{D,j}$ for $j = 1, 2, 3$ is $\sqrt{D}$ or $\sqrt{D}/2$. Hence, amongst the circles $\mathcal{C}_D$ or $\mathcal{C}_{D,j}$, $j = 1, 2, 3$, the only possibilities are

$$\mathcal{C}_1, \quad \mathcal{C}_2, \quad \mathcal{C}_{1,1}, \quad \mathcal{C}_{1,2}, \quad \mathcal{C}_{2,3}, \quad \mathcal{C}_{5,1}, \quad \mathcal{C}_{5,2}, \quad \mathcal{C}_{6,3}$$

and the only one of these that can give rise to a *closed* surface is $\mathcal{C}_{6,3}$ (see [16, Lemma 8]).

The upshot of this discussion is that if $S \hookrightarrow M$ is a closed embedded totally geodesic surface with associated circle $\mathcal{C}$, then there exists $\alpha \in \mathrm{PSL}(2, \mathbb{Z}[i])$ such that $\mathcal{C} = \alpha\mathcal{C}_{6,3}$. □

The proof of Lemma 2.4 will be completed in the sections below. To that end we make some additional comments and introduce some notation. From Proposition 3.1 and Remark 3.2, we need to consider certain groups Δ with $\Gamma(2+2i) \lhd \Delta \lhd \Gamma(1+i)$.

A complete system of (left or right) coset representatives for $\Gamma(1+i)$ in $\mathrm{PSL}(2, \mathbb{Z}[i])$ is provided by the following 6 matrices:

$$T_0 = \mathrm{id}, \qquad T_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix},$$

$$T_3 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \qquad T_4 = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \qquad T_5 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

Using this and the normality of Δ in $\Gamma(1+i)$ (see the proof of Lemma 4.1), it follows that if $\mathcal{C} = \alpha\mathcal{C}_{6,3}$ corresponds to a closed embedded totally geodesic surface in $\mathbb{H}^3/\Delta$ then one of $T_i\mathcal{C}_{6,3}$ also corresponds to such a surface. Briefly, since $\mathcal{C}$ corresponds to a closed embedded totally geodesic surface in $\mathbb{H}^3/\Delta$, $\mathcal{C}_{6,3} = \alpha^{-1}\mathcal{C}$ corresponds to a closed embedded totally geodesic surface in $\mathbb{H}^3/\alpha^{-1}\Delta\alpha$. Writing $\alpha = \gamma T_i$ for some $\gamma \in \Gamma(1+i)$, and using $\Delta \lhd \Gamma(1+i)$ we deduce that $\mathcal{C}_{6,3}$ corresponds to a closed embedded totally geodesic surface in $\mathbb{H}^3/T_i^{-1}\Delta T_i$ from which it follows that $T_i\mathcal{C}_{6,3}$ corresponds to a closed embedded totally geodesic surface in $\mathbb{H}^3/\Delta$.

Using the action on the Hermitian forms described above, the action by the matrices T_i is given by $(T_i^{-1})^* A T_i^{-1}$, which, since the entries of the matrices T_i are integers, is simply $(T_i^{-1})^t A T_i^{-1}$. Hence, the circles $T_i\mathcal{C}_{6,3}$ for $i = 0, 1, \dots, 5$ are represented by the matrices

$$A = A_0 = \begin{pmatrix} 2 & 1-i \\ 1+i & -2 \end{pmatrix}, \quad A_1 = \begin{pmatrix} -2 & -1-i \\ -1+i & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 3-i \\ 3+i & 2 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 2 & -1-i \\ -1+i & -2 \end{pmatrix}, \quad A_4 = \begin{pmatrix} -2 & -3-i \\ -3+i & -2 \end{pmatrix}, \quad A_5 = \begin{pmatrix} -2 & 1-i \\ 1+i & 2 \end{pmatrix}.$$

5. Proving no closed embedded totally geodesic surfaces

From Lemma 2.1 to prove Lemma 2.4 (i.e., that the link complements W_{23}, W_{71}, W_{1091} and W_{1092} do not contain a closed embedded totally geodesic hyperbolic 3-manifold), we are reduced to showing that the hyperbolic 3-manifolds A^+, B^+, C^+, D^+ and E^+ do not contain a closed embedded totally geodesic surface (which could be nonorientable).

In what follows in each of the subsections below we list elements of the groups $\pi_1(A^+)$, $\pi_1(B^+)$, $\pi_1(C^+)$, $\pi_1(D^+)$ and $\pi_1(E^+)$ that provide self-intersections of the circles $T_i\mathcal{C}_{6,3}$. This is done using the matrix generators for each of $\pi_1(A^+)$, $\pi_1(B^+)$, $\pi_1(C^+)$, $\pi_1(D^+)$ and $\pi_1(E^+)$ listed in Section 3B and the criteria of [13] stated Section 4:

$$|\mathrm{tr}(\gamma^* A_i \gamma A_i^{-1})| < 2.$$

These calculations were performed in Mathematica [22] and the notebook is available from the authors upon request. For convenience, we shall simply denote the generators for each of the groups $\pi_1(A^+)$, $\pi_1(B^+)$, $\pi_1(C^+)$, $\pi_1(D^+)$ and $\pi_1(E^+)$ in Section 3B by $g_1, g_2, \dots, g_5$ in the order that they are listed. What is listed below are the Hermitian forms $A = A_0, A_1, \dots, A_5$ and those elements γ, written in terms of $g_1, g_2, \dots, g_5$, for which $|\mathrm{tr}(\gamma^* A_i \gamma A_i^{-1})| < 2$.

We will also make use of the element $l = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, which being an element of $\Gamma(1+i)$ normalizes each of the groups $\pi_1(A^+)$, $\pi_1(B^+)$, $\pi_1(C^+)$, $\pi_1(D^+)$ and $\pi_1(E^+)$ by Proposition 3.1. Additional explanation of elements that are not visibly in the groups $\pi_1(A^+)$, $\pi_1(B^+)$, $\pi_1(C^+)$, $\pi_1(D^+)$ and $\pi_1(E^+)$ is provided when needed.

Finally, we remark that we also need to ensure that the elements do not leave the circles in question invariant. This is clear if the elements are parabolic (since the surface is closed) and when the trace is a nonreal complex number that is not purely imaginary. In the cases where the element has trace that is pure imaginary we check to see whether the circle is left invariant.

The manifold A^+:

circle/form	element	|trace value|
A	g_3	2/3
A_1	g_1	2/3
A_2	$g_1 g_4$	2/3
A_3	g_1	2/3
A_4	$l\beta l^{-1}$	2/3
A_5	g_3	2/3

Note that the element g_1 has trace $-2i$. However, a calculation shows that g_1 does not leave invariant any of the circles $T_i\mathcal{C}_{6,3}$ for $i = 0, 1, \dots, 5$. That $\beta \in \pi_1(A^+)$, can be checked by noting that

$$g_2^{-1} g_1 g_3 \beta = \begin{pmatrix} 5+8i & 10-2i \\ 18+6i & 13-16i \end{pmatrix} = \begin{pmatrix} 1+4(1+2i) & (2+2i)(2-3i) \\ (2+2i)(6-3i) & 1+4(3-4i) \end{pmatrix} \in \Gamma(2+2i).$$

By Remark 3.2 and Proposition 3.1, $\Gamma(2+2i) = [\Gamma(1+i), \Gamma(1+i)] < \pi_1(A^+)$.

None of the elements g_3, $g_1 g_4$ and β (and hence also $l\beta l^{-1} \in \pi_1(A^+)$) have purely imaginary trace.

The manifold B^+:

circle/form	element	|trace value|
A	g_1	$2/3$
A_1	g_1	$2/3$
A_2	g_2	$2/3$
A_3	g_2	$2/3$
A_4	lg_1l^{-1}	$2/3$
A_5	g_1	$2/3$

From Section 3B, g_1 and g_2 are both parabolic, and as above $lg_1l^{-1} \in \pi_1(B^+)$.

The manifold C^+: In this case the parabolic element g_4 works for all the forms $A, A_1, \ldots, A_5$ with trace value $2/3$.

The manifold D^+:

circle/form	element	|trace value|
A	g_1	$2/3$
A_1	g_1	$2/3$
A_2	lg_1l^{-1}	$2/3$
A_3	lg_1l^{-1}	$2/3$
A_4	lg_1l^{-1}	$2/3$
A_5	lg_1l^{-1}	$2/3$

From Section 3B, g_1 is parabolic, and as noted above $lg_1l^{-1} \in \pi_1(D^+)$.

The manifold E^+: Since the parabolic element $g_1 \in \pi_1(E^+)$ is exactly the same as for $\pi_1(D^+)$, the same table holds for E^+ as that shown for D^+.

Remark 5.1. From Remark 3.2 and Proposition 3.1 we know that $\mathbb{H}^3/\Gamma(2+2i)$ covers each of the manifolds A^+, B^+, C^+, D^+, and E^+, which we have shown do not contain a closed embedded totally geodesic surface. On the other hand, as pointed out in [13] the link complement $\mathbb{H}^3/\Gamma(2+2i)$ does contain a closed totally geodesic surface of genus 3 associated to $\mathcal{C}_{6,3}$.

Acknowledgements

This work began whilst the authors were both visiting the Institut de Mathématiques, Université de Neuchâtel, and we would like to thank the Insititut for its hospitality.

We also wish to thank A. Kolpakov, S. Riolo and L. Slavich for many helpful conversations on topics related to this work. We are particularly grateful to a referee who spotted some errors in the calculations in Sections 4 and 5 of the first version of this paper, and made several helpful clarifying comments and suggestions.

References

[1] M. D. Baker and A. W. Reid, "Principal congruence link complements", *Ann. Fac. Sci. Toulouse Math.* (6) **23**:5 (2014), 1063–1092. MR Zbl

[2] A. Basmajian, "Tubular neighborhoods of totally geodesic hypersurfaces in hyperbolic manifolds", *Invent. Math.* **117**:2 (1994), 207–225. MR Zbl

[3] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[4] A. M. Brunner, Y. W. Lee, and N. J. Wielenberg, "Polyhedral groups and graph amalgamation products", *Topology Appl.* **20**:3 (1985), 289–304. MR Zbl

[5] M. Chu and A. W. Reid, "Embedding closed hyperbolic 3-manifolds in small volume hyperbolic 4-manifolds", *Algebr. Geom. Topol.* **21**:5 (2021), 2627–2647. MR Zbl

[6] B. Fine and M. Newman, "The normal subgroup structure of the Picard group", *Trans. Amer. Math. Soc.* **302**:2 (1987), 769–786. MR Zbl

[7] D. Gabai, R. Meyerhoff, and P. Milley, "Minimum volume cusped hyperbolic three-manifolds", *J. Amer. Math. Soc.* **22**:4 (2009), 1157–1215. MR Zbl

[8] W. Hantzsche, "Einlagerung von Mannigfaltigkeiten in euklidische Räume", *Math. Z.* **43**:1 (1938), 38–58. MR Zbl

[9] D. Ivanšić, "Embeddability of noncompact hyperbolic manifolds as complements of codimension-1 and -2 submanifolds", *Topology Appl.* **120**:1-2 (2002), 211–236. MR Zbl

[10] D. Ivanšić, "Hyperbolic structure on a complement of tori in the 4-sphere", *Adv. Geom.* **4**:1 (2004), 119–139. MR Zbl

[11] D. Ivanšić, "A topological 4-sphere that is standard", *Adv. Geom.* **12**:3 (2012), 461–482. MR Zbl

[12] D. Ivanšić, J. G. Ratcliffe, and S. T. Tschantz, "Complements of tori and Klein bottles in the 4-sphere that have hyperbolic structure", *Algebr. Geom. Topol.* **5** (2005), 999–1026. MR Zbl

[13] J. Jung and A. W. Reid, "Embedding closed totally geodesic surfaces in Bianchi orbifolds", 2020. To appear in *Math. Research Letters*. arXiv 2003.05427

[14] R. Kirby, "Problems in low-dimensional topology", pp. 35–473 in *Geometric topology* (Athens, GA, 1993), AMS/IP Stud. Adv. Math. **2**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[15] C. J. Leininger, "Small curvature surfaces in hyperbolic 3-manifolds", *J. Knot Theory Ramifications* **15**:3 (2006), 379–411. MR Zbl

[16] C. Maclachlan and A. W. Reid, "Parametrizing Fuchsian subgroups of the Bianchi groups", *Canad. J. Math.* **43**:1 (1991), 158–181. MR Zbl

[17] W. Menasco and A. W. Reid, "Totally geodesic surfaces in hyperbolic link complements", pp. 215–226 in *Topology '90* (Columbus, OH, 1990), edited by B. Apanasov et al., Ohio State Univ. Math. Res. Inst. Publ. **1**, de Gruyter, Berlin, 1992. MR Zbl

[18] J. G. Ratcliffe and S. T. Tschantz, "Gravitational instantons of constant curvature", pp. 2613–2627 in *Topology of the Universe Conference* (Cleveland, OH, 1997), vol. 15, edited by G. D. Starkman, 1998. MR Zbl

[19] J. G. Ratcliffe and S. T. Tschantz, "The volume spectrum of hyperbolic 4-manifolds", *Experiment. Math.* **9**:1 (2000), 101–125. MR Zbl

[20] H. Saratchandran, "Finite volume hyperbolic complements of 2-tori and Klein bottles in closed smooth simply connected 4-manifolds", *New York J. Math.* **24** (2018), 443–450. MR

[21] R. G. Swan, "Generators and relations for certain special linear groups", *Advances in Math.* **6** (1971), 1–77. MR Zbl

[22] Wolfram Research, Inc., "Mathematica", software, Champaign, Illinois, 2017. Version 11.2.

Received March 10, 2022. Revised August 14, 2023.

MICHELLE CHU
SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN
UNITED STATES
mchu@umn.edu

ALAN W. REID
DEPARTMENT OF MATHEMATICS
RICE UNIVERSITY
HOUSTON, TX
UNITED STATES
alan.reid@rice.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.209

INFINITE HOMOTOPY STABLE CLASS FOR 4-MANIFOLDS WITH BOUNDARY

ANTHONY CONWAY, DIARMUID CROWLEY AND MARK POWELL

We show that for every odd prime q, there exists an infinite family $\{M_i\}_{i=1}^{\infty}$ of topological 4-manifolds that are all stably homeomorphic to one another, all the manifolds M_i have isometric rank one equivariant intersection pairings and boundary $L(2q, 1)\#(S^1 \times S^2)$, but they are pairwise not homotopy equivalent via any homotopy equivalence that restricts to a homotopy equivalence of the boundary.

1. Introduction

In what follows a manifold is understood to mean a compact, connected, oriented, topological manifold. Let $W_g := \#_g(S^2 \times S^2)$ be the g-fold connected sum of $S^2 \times S^2$ with itself. Two 4-manifolds M and N with the same Euler characteristic are *stably homeomorphic*, denoted $M \cong_{\mathrm{st}} N$, if there exists a nonnegative integer g and a homeomorphism

$$M\#W_g \cong N\#W_g.$$

Surgery theory suggests two ways to classify 4-manifolds. The classical Browder–Novikov–Sullivan–Wall [Wall 1999] approach is to classify up to homotopy equivalence and then employ the surgery exact sequence. Kreck's modified surgery approach [1999] seeks to classify up to stable homeomorphism, and then attempt to destabilise. A natural question then arising is to compare the homotopy and stable classifications. To do this precisely for 4-manifolds with boundary we fix a 4-manifold M and define the *homotopy stable class*:

$$\mathcal{S}_h^{\mathrm{st}}(M) := \{N \mid N \cong_{\mathrm{st}} M\}/\text{homotopy equivalence of pairs.}$$

Here, we understand a *homotopy equivalence of pairs* $N_1 \simeq N_2$ to be one that restricts to a homotopy equivalence between the boundaries. When the manifolds are closed, this recovers the usual notion of homotopy equivalence.

Using the equivariant intersection form λ_N of N as an invariant, $\mathcal{S}_h^{\mathrm{st}}(M)$ can be arbitrarily large: for example, one can use Freedman's work [1982] to realise

MSC2020: 57K40, 57R65.
Keywords: stable homeomorphism, homotopy equivalence, 4-manifold.

distinct positive definite symmetric bilinear forms with the same signature and rank by simply connected closed 4-manifolds with identical Kirby–Siebenmann invariant. For this reason, we study the homotopy stable class one intersection form at a time and set

$$\mathcal{S}^{\mathrm{st}}_{h,\lambda}(M) := \{N \mid N \cong_{\mathrm{st}} M, \lambda_N \cong \lambda_M\}/\text{homotopy equivalence of pairs.}$$

If M is closed and has $\pi_1(M) = 1$, $\mathbb{Z}$, or $\mathbb{Z}/n$, or $\pi_1(M)$ is a solvable Baumslag–Solitar group, then $|\mathcal{S}^{\mathrm{st}}_{h,\lambda}(M)| = 1$: stably homeomorphic manifolds with isometric equivariant intersection forms are homeomorphic for $\pi_1 = 1$ by [Freedman 1982], for $\pi_1 \cong \mathbb{Z}$ by [Freedman and Quinn 1990], for $\pi_1 \cong \mathbb{Z}/n$ by [Hambleton and Kreck 1993, Theorem C], and for solvable Baumslag–Solitar group by [Hambleton, Kreck and Teichner 2009, Theorem A]. On the other hand, Kreck and Schafer [1984] found pairs of smooth closed 4-manifolds with finite π_1 and isometric equivariant intersection forms that are stably diffeomorphic but not homotopy equivalent. When the boundary is nonempty and $\pi_1 = 1$, one can use work of Boyer [1993] to produce simply connected 4-manifolds M with boundary and arbitrarily large (but necessarily finite) $\mathcal{S}^{\mathrm{st}}_{h,\lambda}(M)$. Until now however, there have been no examples of 4-manifolds with infinite $\mathcal{S}^{\mathrm{st}}_{h,\lambda}(M)$. For every odd prime q, our main result describes a 4-manifold M with fundamental group $\mathbb{Z}$ and infinite $\mathcal{S}^{\mathrm{st}}_{h,\lambda_{2q}}(M)$, where the fixed Hermitian form is

$$\lambda_{2q} : \mathbb{Z}[t^{\pm 1}] \times \mathbb{Z}[t^{\pm 1}] \to \mathbb{Z}[t^{\pm 1}]; \quad (x, y) \mapsto 2qx\bar{y}.$$

Theorem 1.1. *For every odd prime q, there exists an infinite family $\{M_i\}_{i=1}^{\infty}$ of 4-manifolds with fundamental group $\mathbb{Z}$ that are all stably homeomorphic, and all the manifolds M_i have equivariant intersection pairing isometric to λ_{2q} and boundary $L(2q, 1)\#(S^1 \times S^2)$, but they are pairwise not homotopy equivalent via any homotopy equivalence that restricts to a homotopy equivalence on the boundary. In other words,*

$$|\mathcal{S}^{\mathrm{st}}_{h,\lambda_{2q}}(M_1)| = \infty.$$

For a fixed odd prime q, the manifolds in Theorem 1.1 all have fundamental group $\mathbb{Z}$, boundary

$$Y_q := L(2q, 1)\#(S^1 \times S^2),$$

equivariant intersection form isometric to λ_{2q}, and integral intersection form isometric to

$$\lambda^{\mathbb{Z}}_{2q} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}; \quad (x, y) \mapsto 2qxy,$$

but are distinguished by an invariant, first introduced in [Conway, Piccirillo and Powell 2022] and inspired by [Boyer 1993], related to the Blanchfield form of Y_q. While the manifold M_1 is smooth, we cannot tell whether any of the other M_i admit smooth structures. Their construction uses surgery methods, in particular a recent realisation result from [Conway, Piccirillo and Powell 2022], which a priori only works in the topological category.

Before giving more details and describing the main steps in the proof of Theorem 1.1, we briefly compare the study of the homotopy stable class in dimension 4 with the situation in higher dimensions.

Remark 1.2. Kreck and Schafer [1984] found pairs of closed smooth $4k$-manifolds, for $k \geq 1$, that are stably diffeomorphic and have hyperbolic equivariant intersection forms, but are pairwise not homotopy equivalent. In [Conway, Crowley, Powell and Sixt 2023], we gave the first examples of simply connected, closed, smooth $4k$-manifolds, for $k \geq 2$, with hyperbolic intersection form and arbitrarily large homotopy stable class $\mathcal{S}^{\mathrm{st}}_{h,\lambda}$. In [Conway, Crowley, Powell and Sixt 2021] for $k \geq 2$, we produced smooth closed $4k$-manifolds with fundamental group $\mathbb{Z}$, again with hyperbolic intersection form, and such that the homotopy stable class $\mathcal{S}^{\mathrm{st}}_{h,\lambda}$ is infinite. In those papers we were unable to obtain examples in dimension 4. In [Conway, Crowley, Powell and Sixt 2023], in lieu of this we defined a spin^c version of the stable class in dimension 4, and we showed that this spin^c stable class can be arbitrarily large. This article shows that a variation on those methods, with analogous underlying algebra, does produce examples of 4-manifolds with nonempty boundary and fundamental group $\mathbb{Z}$ that have infinite homotopy stable class.

Next we describe the main steps in the proof of Theorem 1.1. Fix an odd prime q. The first observation is that if N_1, N_2 are 4-manifolds with integral intersection forms isometric to $\lambda^{\mathbb{Z}}_{2q}$, then there can be no orientation-reversing homotopy equivalence between N_1 and N_2. For this reason, and for the purpose of proving our main theorem, we restrict to orientation-preserving homotopy equivalences (o.p. homotopy eq. for short) and therefore consider

$$\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) := \{N \mid N \cong_{\mathrm{st}} M,\, \lambda_N \cong \lambda_M\}/\text{o.p. homotopy eq. of pairs.}$$

We now restrict to 4-manifolds M with fundamental group $\mathbb{Z}$ such that the inclusion $\partial M \subseteq M$ induces a surjection $\varphi : \pi_1(\partial M) \twoheadrightarrow \pi_1(M) \xrightarrow{\cong} \mathbb{Z}$ (we say that M has *ribbon boundary*) and for which $H_1(\partial M; \mathbb{Z}[t^{\pm 1}])$ is a $\mathbb{Z}[t^{\pm 1}]$-torsion module. Here and throughout the paper we assume that the fundamental groups of our 4-manifolds are equipped with a preferred isomorphism to $\mathbb{Z}$; to indicate this we write $\pi_1(M) = \mathbb{Z}$.

Given two such manifolds N_1 and N_2, we write $\partial N_1 \cong_B \partial N_2$ if there exists an orientation-preserving homeomorphism $f : \partial N_1 \xrightarrow{\cong} \partial N_2$ that intertwines the inclusion induced epimorphisms $\varphi_i : \pi_1(\partial N_i) \twoheadrightarrow \pi_1(N_i)$ and, in the case that N_1 and N_2 are spin, such that the union $N_1 \cup_f -N_2$ is spin. The terminology $\cong_B$ is motivated by modified surgery theory [Kreck 1999], in which B is the standard notation for the normal 1-type.

Next, if N_1 and N_2 have fundamental group $\mathbb{Z}$, $\partial N_1 \cong_B \partial N_2$, the same Kirby–Siebenmann invariant, and $\lambda_{N_1} \cong \lambda_{N_2}$, then they are stably homeomorphic. Indeed, N_1 and N_2 must have isometric integral intersection forms (in particular with the same type and the same signature) and the same Kirby–Siebenmann invariant, so [Kreck 1999, Theorem 2] ensures they are stably homeomorphic; see Lemma 4.1 for details.

Put differently, if M is a 4-manifold with infinite cyclic fundamental group, then

$$\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) = \frac{\{N \mid \partial N \cong_B \partial M, \pi_1(N) = \mathbb{Z}, \lambda_N \cong \lambda_M, \mathrm{ks}(N) = \mathrm{ks}(M)\}}{\text{o.p. homotopy eq. of pairs}}.$$

This next step is to recast $\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M)$ in terms of the group $\mathrm{Aut}(\mathrm{Bl}_{\partial M})$ of isometries of the Blanchfield form $\mathrm{Bl}_{\partial M}$ (whose definition we recall in Section 3). Firstly, as we recall in Section 5, the group $\mathrm{hAut}^+_\varphi(\partial M)$ of orientation-preserving homotopy equivalences $h : \partial M \simeq \partial M$ that intertwine the inclusion induced map $\varphi : \pi_1(\partial M) \to \pi_1(M) = \mathbb{Z}$ acts on $\mathrm{Aut}(\mathrm{Bl}_{\partial M})$. Secondly, as we also recall in Section 5, the group $\mathrm{Aut}(\lambda_M)$ of isometries of λ_M also acts on $\mathrm{Aut}(\mathrm{Bl}_{\partial M})$, and the two actions commute with one another. Quotienting out by these two actions leads to an orbit set $\mathrm{Aut}(\mathrm{Bl}_{\partial M})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M))$. Note that it need not be group.

In order to account for our 4-manifolds being spin, we will in fact need to work with a smaller set of isometries. Namely, if M is spin, then $\mathrm{Bl}_{\partial M}$ admits a quadratic enhancement

$$\mu_{\mathrm{Bl}_{\partial M}} : H_1(\partial M; \mathbb{Z}[t^{\pm 1}]) \to \frac{\{b \in \mathbb{Q}(t) \mid b = \bar{b}\}}{\{a + \bar{a} \mid a \in \mathbb{Z}[t^{\pm 1}]\}}$$

and we write $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}}) \subseteq \mathrm{Aut}(\mathrm{Bl}_{\partial M})$ for those isometries of $\mathrm{Bl}_{\partial M}$ that also preserve $\mu_{\mathrm{Bl}_{\partial M}}$. Writing $\mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)$ for those homotopy equivalences whose induced map on the Alexander module preserves $\mu_{\mathrm{Bl}_{\partial M}}$ leads to the orbit set

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)).$$

One of the main steps in the proof of Theorem 1.1 is the following partial description of $\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M)$ for a large class of 4-manifolds M with infinite cyclic fundamental group and ribbon boundary. As we will explain in Proposition 5.2, this result follows fairly promptly from the machinery developed in [Conway, Piccirillo and Powell 2022]. In the following proposition, and throughout the paper, *spin* refers to a manifold that admits a spin structure compatible with the orientation.

Proposition 1.3. *If M is a spin 4-manifold with ribbon boundary, $\pi_1(M) = \mathbb{Z}$, and nondegenerate equivariant intersection form λ_M, then there is a surjection*

$$b : \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) \twoheadrightarrow \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)).$$

The surjection is described explicitly in Construction 5.1.

Fix an odd prime q and let $X_{2q}(U)$ denote the $2q$-trace on the unknot U, i.e. the smooth 4-manifold obtained from D^4 by attaching a $2q$-framed 2-handle along the unknot. The final part of the proof of Theorem 1.1, which is carried out in Proposition 6.8, consists of proving that for $M = X_{2q}(U)\natural(S^1 \times D^3)$, the set

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M))$$

is countably infinite. Modulo this statement, we can now conclude the proof of Theorem 1.1, which states that $\mathcal{S}^{\mathrm{st}}_{h,\lambda_{2q}}(M)$ is infinite.

Proof of Theorem 1.1. Fix an odd prime q and consider $M := X_{2q}(U)\natural(S^1 \times D^3)$. This 4-manifold is spin, has ribbon boundary, admits an identification $\pi_1(M) = \mathbb{Z}$, and has nondegenerate equivariant intersection $\lambda_M \cong (2q)$. Since for any two 4-manifolds N_1 and N_2 with integral intersection forms isometric to $\lambda^{\mathbb{Z}}_{2q}$, there is no orientation reversing homotopy equivalence between them, $\mathcal{S}^{\mathrm{st}}_{h,\lambda_{2q}}(M) = \mathcal{S}^{\mathrm{st}}_{h^+,\lambda_{2q}}(M)$. We therefore prove that $\mathcal{S}^{\mathrm{st}}_{h^+,\lambda_{2q}}(M)$ is infinite. To prove this we apply Proposition 1.3, which implies that $\mathcal{S}^{\mathrm{st}}_{h^+,\lambda_{2q}}(M)$ surjects onto the orbit set

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M))$$

and this latter set is countably infinite by Proposition 6.8. □

Remark 1.4. The existence of M with infinite

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M))$$

is what makes it possible for us to obtain an example where the homotopy stable class $\mathcal{S}^{\mathrm{st}}_{h,\lambda}$ is infinite. While an analogue of Proposition 1.3 can be proved in the simply connected case using results of [Boyer 1993], the corresponding algebra always remains finite for trivial fundamental group.

All of the infinite sets we discuss are necessarily countable. Primarily, this has to be the case because there are only countably many compact manifolds [Cheeger and Kister 1970]. On the algebraic side it is also evident that the orbit set onto which the homotopy stable class surjects in Proposition 1.3 is countable, essentially because all the homology groups involved are finitely generated over $\mathbb{Z}[t^{\pm 1}]$.

Next we discuss a variation on Proposition 1.3 that may be of independent interest. The surjection in Proposition 1.3 can be improved to a bijection if we require the homotopy equivalences $N_1 \simeq N_2$ to restrict to homeomorphisms on the boundary; i.e. if we consider

$$\mathcal{S}^{\mathrm{st},\partial}_{h^+,\lambda}(M) := \frac{\{N \mid N \cong_{\mathrm{st}} M,\ \lambda_N \cong \lambda_M\}}{\text{o.p. homotopy eq. that restricts to a homeo. on the boundary}}$$

and change the target accordingly, i.e. consider

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{Homeo}^{+,\mathfrak{q}}_{\varphi}(\partial M))$$

instead of

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathfrak{q}}_{\varphi}(\partial M)).$$

In fact, the same result is obtained with

$$\mathcal{S}^{\mathrm{st}}_{+,\lambda}(M) := \frac{\{N \mid N \cong_{\mathrm{st}} M,\ \lambda_N \cong \lambda_M\}}{\text{o.p. homeomorphism}}.$$

Proposition 1.5. *If M is a spin 4-manifold with $\pi_1(M) = \mathbb{Z}$, ribbon boundary and nondegenerate equivariant intersection form λ_M, then there are bijections*

$$\mathcal{S}^{\mathrm{st},\partial}_{h^+,\lambda}(M) \xrightarrow{\approx} \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{Homeo}^{+,\mathfrak{q}}_{\varphi}(\partial M)),$$
$$\mathcal{S}^{\mathrm{st}}_{+,\lambda}(M) \xrightarrow{\approx} \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{Homeo}^{+,\mathfrak{q}}_{\varphi}(\partial M)).$$

The bijections are induced by the map b that will be introduced in Construction 5.1. For $M := X_{2q}(U)\natural(S^1 \times D^3)$, with q an odd prime, the sets above are countably infinite.

Proof. The surjectivity follows from the same argument that we will use in Proposition 5.2. We prove injectivity. If $b(N_1) = b(N_2)$, then [Conway, Piccirillo and Powell 2022, Theorem 1.1] shows that the manifolds N_1 and N_2 are orientation-preserving homeomorphic. Since the quotient with $\mathrm{Homeo}^{+,\mathfrak{q}}_{\varphi}(\partial M)$ replaced by $\mathrm{hAut}^{+,\mathfrak{q}}_{\varphi}(\partial M)$ is infinite, and since $\mathrm{Homeo}^{+,\mathfrak{q}}_{\varphi}(\partial M) \subseteq \mathrm{hAut}^{+,\mathfrak{q}}_{\varphi}(\partial M)$, it follows that the sets in the statement are infinite. □

We now characterise $M := X_{2q}(U)\natural(S^1 \times D^3)$ within $\mathcal{S}^{\mathrm{st}}_{+,\lambda}(M)$ in terms of the knottedness of the sphere $S^2_l := \{\mathrm{pt}\} \times S^2 \subseteq ((S^1 \times S^2) \setminus \mathrm{Int}(D^3)) \subseteq \partial M$ and the connect sum sphere $S^2_c \subseteq M$:

Theorem 1.6. *For $M = X_{2q}(U)\natural(S^1 \times D^3)$ and $N \in \mathcal{S}^{\mathrm{st}}_{+,\lambda}(M)$, the following are equivalent*:

(1) *N is homeomorphic to M.*

(2) *$S^2_l \subseteq \partial N$ bounds a locally flat $D^3 \subseteq N$.*

(3) *S^2_c bounds a locally flat $D^3 \subseteq N$.*

Proof. The implications (1) ⇒ (2) and (1) ⇒ (3) are immediate.

We prove the implication (2) ⇒ (1). Cut N along the D^3 with boundary S^2_l to obtain a simply connected 4-manifold with boundary $L(2q, 1)$ and $H_2 = \mathbb{Z}$. Theorem 0.1 of [Boyer 1986] implies that such a manifold is homeomorphic to $X_{2q}(U)$. Glue back the $D^3 \times [0, 1]$ that we removed to recover N as M.

Finally, we prove the implication (3) $\Rightarrow$ (1). Cut N open along the separating D^3, resulting in a disjoint union of two 4-manifolds. The first is simply connected with $H_2 = \mathbb{Z}$ and boundary $L(2q, 1)$ and is therefore homeomorphic to $X_{2q}(U)$ [Boyer 1986, Theorem 0.1]. The second has $\pi_1 = \mathbb{Z}$, no H_2 and boundary $S^1 \times S^2$; it is thus homeomorphic to $S^1 \times D^3$ [Freedman and Quinn 1990, §11.6]. Glue back the $D^3 \times [0, 1]$ that we removed to recover N as M. □

Organisation. In Sections 2 and 3, we review some facts about linking forms and in particular the Blanchfield form. In Section 4 we give a criterion that implies stable homeomorphism of 4-manifolds with fundamental group $\mathbb{Z}$ and nonempty boundary. In Section 5, we prove Proposition 1.3. In Section 6 we show that for $M = X_{2q}(U)\natural(S^1 \times D^3)$, with q an odd prime, the set

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathfrak{q}}_{\varphi}(\partial M))$$

is infinite.

Conventions. We work in the topological category unless otherwise stated. All manifolds are assumed to be compact, connected, based, and oriented. If a manifold has a nonempty boundary, then the basepoint is assumed to be in the boundary. For a 4-manifold M with fundamental group $\mathbb{Z}$, we fix an identification an write $\pi_1(M) = \mathbb{Z}$. We say that M is spin if M admits a spin structure compatible with the orientation. We write $p \mapsto \bar{p}$ for the involution on $\mathbb{Z}[t^{\pm 1}]$ induced by $t \mapsto t^{-1}$. Given a $\mathbb{Z}[t^{\pm 1}]$-module H, we write $\overline{H}$ for the $\mathbb{Z}[t^{\pm 1}]$-module whose underlying abelian group is H but with module structure given by $p \cdot h = \bar{p}h$ for $h \in H$ and $p \in \mathbb{Z}[t^{\pm 1}]$. We write $H^* := \overline{\operatorname{Hom}_{\mathbb{Z}[t^{\pm 1}]}(H, \mathbb{Z}[t^{\pm 1}])}$.

2. Linking forms and unions

Since a large part of this paper is concerned with the Blanchfield form and isometries thereof, we start by recalling terminology related to the underlying algebra. In Section 2.1 we recall symmetric and quadratic linking forms. In Section 2.2 we recall how a Hermitian form has a boundary which is a symmetric linking form, and the boundary of an even form has the additional structure of a quadratic refinement. In Section 2.3 we recall how isometries of these linking forms can be used to glue two linking forms together, and we show that the union of two even forms along an isometry of their boundary quadratic linking forms is again an even form.

2.1. ***Symmetric and quadratic linking forms.*** Everything in this subsection is the special case for $\mathbb{Z}[t^{\pm 1}]$ of a general theory for arbitrary rings with involution developed by Ranicki [1981, §3.4].

Definition 2.1. A *symmetric linking form* over $\mathbb{Z}[t^{\pm 1}]$ is a pair (T, ℓ), where T is a torsion $\mathbb{Z}[t^{\pm 1}]$-module, and $\ell : T \times T \to \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]$ is a Hermitian, sesquilinear, nonsingular pairing.

We write $S := \mathbb{Z}[t^{\pm 1}] \setminus \{0\}$, and set

$$\begin{aligned}
Q^1(\mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]) &:= \frac{\{b \in \mathbb{Q}(t) \mid b - \bar{b} \in \mathbb{Z}[t^{\pm 1}]\}}{\mathbb{Z}[t^{\pm 1}]}, \\
Q_1(\mathbb{Z}[t^{\pm 1}], S) &:= \frac{\{b \in \mathbb{Q}(t) \mid b = \bar{b}\}}{\{a + \bar{a} \mid a \in \mathbb{Z}[t^{\pm 1}]\}}, \\
Q^1(\mathbb{Z}[t^{\pm 1}], S) &:= \frac{\{b \in \mathbb{Q}(t) \mid b - \bar{b} = a - \bar{a} \text{ for some } a \in \mathbb{Z}[t^{\pm 1}]\}}{\mathbb{Z}[t^{\pm 1}]} \\
&\subseteq Q^1(\mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]).
\end{aligned}$$

For a symmetric linking form (T, ℓ), we have that $\ell(x, x) \in Q^1(\mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}])$ for all $x \in T$. The symmetric linking form is called *even* if $\ell(x, x) \in Q^1(\mathbb{Z}[t^{\pm 1}], S)$ for all $x \in T$. We define a map

$$q : Q_1(\mathbb{Z}[t^{\pm 1}], S) \to Q^1(\mathbb{Z}[t^{\pm 1}], S), \quad [b] \mapsto [b].$$

Definition 2.2. A *quadratic refinement* of an even symmetric linking form (T, ℓ) is a function $\mu : T \to Q_1(\mathbb{Z}[t^{\pm 1}], S)$ satisfying

(i) $\mu(rx) = r\mu(x)\bar{r} \in Q_1(\mathbb{Z}[t^{\pm 1}], S)$ for all $x \in T$ and for all $r \in \mathbb{Z}[t^{\pm 1}]$;

(ii) $\mu(x + y) = \mu(x) + \mu(y) + \ell(x, y) + \overline{\ell(x, y)} \in Q_1(\mathbb{Z}[t^{\pm 1}], S)$ for all $x, y \in T$;

(iii) $q(\mu(x)) = \ell(x, x) \in Q^1(\mathbb{Z}[t^{\pm 1}], S)$ for all $x \in T$.

A triple (T, ℓ, μ) consisting of a symmetric linking form together with a quadratic refinement is called a *quadratic linking form* over $\mathbb{Z}[t^{\pm 1}]$.

For aficionados of [Ranicki 1981], we emphasise that we are using the nonsplit version of quadratic linking forms.

We will also need to consider isometries and automorphisms of symmetric and quadratic linking forms.

Definition 2.3. Let (T, ℓ) and (T', ℓ') be symmetric linking forms over $\mathbb{Z}[t^{\pm 1}]$ and let $\mu : T \to Q_1(\mathbb{Z}[t^{\pm 1}], S)$ and $\mu' : T' \to Q_1(\mathbb{Z}[t^{\pm 1}], S)$ be respective quadratic refinements.

(1) An isomorphism $f : T \to T'$ is an *isometry of symmetric linking forms* if

$$\ell'(f(x), f(y)) = \ell(x, y)$$

for every $x, y \in T$.

(2) The isometry of symmetric linking forms f is moreover an *isometry of quadratic linking forms*, $f : (T, \ell, \mu) \cong (T', \ell', \mu')$ if $\mu'(f(x)) = \mu(x)$ for every $x \in T$.

(3) If $(T, \ell) = (T', \ell')$, then f as in (1) is an *automorphism of symmetric linking forms*. We write $\mathrm{Aut}(T, \ell)$ for the group of automorphisms.

(4) If $(T, \ell, \mu) = (T', \ell', \mu')$, then f as in (2) is an *automorphism of quadratic linking forms*. We write $\mathrm{Aut}(T, \ell, \mu)$ for the group of automorphisms.

Remark 2.4. Given a quadratic linking form (T, ℓ, μ) over $\mathbb{Z}[t^{\pm 1}]$ with underlying symmetric linking form (T, ℓ), we note that $\mathrm{Aut}(T, \ell, \mu) \subseteq \mathrm{Aut}(T, \ell)$. We give an example showing that this can be a proper inclusion: multiplication by 3 induces an isomorphism $\mathbb{Z}[t^{\pm 1}]/8 \to \mathbb{Z}[t^{\pm 1}]/8$ that preserves the linking form $\ell(x, y) = \frac{1}{8}x\bar{y}$ but does not preserve the quadratic refinement $\mu(x) = \frac{1}{8}x\bar{x}$. Indeed $\mu(3) = \frac{9}{8} \neq \frac{1}{8} = \mu(1) \in Q_1(\mathbb{Z}[t^{\pm 1}], S)$ because 1 cannot be written as $a + \bar{a}$ with $a \in \mathbb{Z}[t^{\pm 1}]$.

2.2. *Boundaries of quadratic forms.* We recall some terminology about Hermitian forms. A *Hermitian form* refers to a pair (H, λ), where H is a free $\mathbb{Z}[t^{\pm 1}]$-module and $\lambda : H \times H \to \mathbb{Z}[t^{\pm 1}]$ is a sesquilinear Hermitian pairing. Given a Hermitian form (H, λ) over $\mathbb{Z}[t^{\pm 1}]$, we use $\hat{\lambda} : H \to H^* =: \overline{\mathrm{Hom}_{\mathbb{Z}[t^{\pm 1}]}(H, \mathbb{Z}[t^{\pm 1}])}$ to denote the linear map defined by $\hat{\lambda}(y)(x) = \lambda(x, y)$. We often refer to $\hat{\lambda}$ as the *adjoint* of λ. We say that λ is *nondegenerate* if $\hat{\lambda}$ is injective and *nonsingular* if $\hat{\lambda}$ is an isomorphism. We also recall that a Hermitian form (H, λ) is called *even* if for all $x \in H$, there exists $a \in \mathbb{Z}[t^{\pm 1}]$ such that $\lambda(x, x) = a + \bar{a}$.

We describe how a nondegenerate even Hermitian form over $\mathbb{Z}[t^{\pm 1}]$ determines a quadratic linking form, following [Ranicki 1981, p. 243].

Definition 2.5. The *boundary symmetric linking form* of a nondegenerate Hermitian form (H, λ) over $\mathbb{Z}[t^{\pm 1}]$ is the symmetric linking form $(\mathrm{coker}(\hat{\lambda}), \partial\lambda)$, where $\partial\lambda$ is defined as

$$\partial\lambda : \mathrm{coker}(\hat{\lambda}) \times \mathrm{coker}(\hat{\lambda}) \to \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}], \quad ([x], [y]) \mapsto \frac{1}{s}(y(z)),$$

where, as $\mathrm{coker}(\hat{\lambda})$ is $\mathbb{Z}[t^{\pm 1}]$-torsion, there exists $s \in \mathbb{Z}[t^{\pm 1}]$ and $z \in H$ such that $sx = \hat{\lambda}(z)$.

If (H, λ) is additionally assumed to be even, then its *boundary quadratic linking form* is the quadratic linking form $(\mathrm{coker}(\hat{\lambda}), \partial\lambda, \mu_\partial)$, where the quadratic refinement of $\partial\lambda$ is

$$\mu_\partial : \mathrm{coker}(\hat{\lambda}) \to Q_1(\mathbb{Z}[t^{\pm 1}], S), \quad [y] \mapsto \frac{1}{s}(y(z)),$$

with $s \in \mathbb{Z}[t^{\pm 1}]$ and $z \in H$ such that $sy = \hat{\lambda}(z)$.

We assert that $\partial\lambda$ is independent of the choices involved, and is nonsingular, sesquilinear, and Hermitian. These can all be verified directly. To enable us to give a reference to existing literature, note that the boundary symmetric linking form is the linking form on $H^1(C^*)$ associated with the 1-dimensional $\mathbb{Q}(t)$-acyclic symmetric Poincaré complex (C_*, φ) over $\mathbb{Z}[t^{\pm 1}]$ given by

$$\begin{array}{ccc} C^0 = H & \xrightarrow{\lambda} & C^1 = H^* \\ \Big\downarrow{\scriptstyle \varphi_0 = \mathrm{Id}} & & \Big\downarrow{\scriptstyle \varphi_0 = \mathrm{Id}} \\ C_1 = H & \xrightarrow{\lambda^* = \lambda} & C_0 = H^* \end{array}$$

In [Powell 2016, Propositions 3.3, 3.4, and 3.8] it was shown that such a linking form is well-defined, nonsingular, sesquilinear, and Hermitian.

The next proposition is implicit in [Ranicki 1981, p. 243]. As far as we know such a proof has not appeared in the literature, so we provide the details of the proof.

Proposition 2.6. *The function μ_∂ is a well-defined function* $\operatorname{coker}(\hat{\lambda}) \to Q_1(\mathbb{Z}[t^{\pm 1}], S)$, *and is a quadratic refinement of the boundary symmetric linking form $\partial\lambda$ of (H, λ), i.e. μ_∂ satisfies the requirements of Definition 2.2.*

Proof. First we show it is well-defined. Let y, z, and s be as in Definition 2.5. Express $\hat{\lambda}$ as a Hermitian matrix $A = \bar{A}^T$ over $\mathbb{Z}[t^{\pm 1}]$ with respect to some basis of H. Then A is invertible over $\mathbb{Q}(t)$ because λ is nondegenerate, so $\det(A) \neq 0$. We have $\mu_\partial(y) = y^T A^{-1} \bar{y}$. Therefore

$$\overline{\mu_\partial(y)} = \overline{\mu_\partial(y)}^T = (\overline{y^T A^{-1} \bar{y}})^T = y^T \overline{A^{-1}}^T \bar{y} = y^T (\bar{A}^T)^{-1} \bar{y} = y^T A^{-1} \bar{y} = \mu_\partial(y).$$

Hence $\mu_\partial(y) \in Q_1(\mathbb{Z}[t^{\pm 1}], S)$.

Next we show that the choices of z and s do not change $\mu_\partial(y)$. Let $y \in H$ and let $y \in \operatorname{coker}(\hat{\lambda})$, $s \in \mathbb{Z}[t^{\pm 1}]$, and $z \in H$ be as in Definition 2.5, with $\hat{\lambda}(z) = sy$. Let $s' \in \mathbb{Z}[t^{\pm 1}]$ and $z' \in H$ be another pair of choices, such that $\hat{\lambda}(z') = s'y$. Since $\mu_\partial(y) \in Q_1(\mathbb{Z}[t^{\pm 1}], S)$, we have

$$\frac{1}{s'} y(z') = \overline{\frac{1}{s'} y(z')}.$$

The difference between the two computations of $\mu_\partial(y)$ yields

$$\begin{aligned} \frac{1}{s} y(z) - \frac{1}{s'} y(z') &= \frac{1}{s} y(z) - \overline{\frac{1}{s'} y(z')} = \frac{1}{s} y(z) \frac{\bar{s}'}{\bar{s}'} - \overline{\frac{1}{s'} y(z')} \frac{\bar{s}}{\bar{s}} \\ &= \frac{1}{s} (s'y)(z) \frac{1}{\bar{s}'} - \overline{\frac{1}{s'} (sy)(z')} \frac{1}{\bar{s}} = \frac{1}{s} \hat{\lambda}(z')(z) \frac{1}{\bar{s}'} - \overline{\frac{1}{s'} \hat{\lambda}(z)(z')} \frac{1}{\bar{s}} \\ &= \frac{1}{s} \lambda(z, z') \frac{1}{\bar{s}'} - \overline{\frac{1}{s'} \overline{\lambda(z, z')}} \frac{1}{\bar{s}} = 0 \in Q_1(\mathbb{Z}[t^{\pm 1}], S). \end{aligned}$$

Next we show that μ_∂ does not depend on the representative y for the class in $\operatorname{coker}(\hat{\lambda})$. Replace $y \in \operatorname{coker}(\hat{\lambda})$ by another representative $y + \hat{\lambda}(u)$, for some $u \in H$. Then $\hat{\lambda}(z + su) = s(y + \hat{\lambda}(u))$. Therefore

$$\mu_\partial(y + \hat{\lambda}(u)) = \frac{1}{s}(y + \hat{\lambda}(u))(z + su) = \frac{1}{s} y(z) + \frac{1}{s} \hat{\lambda}(u)(z) + y(u) + \hat{\lambda}(u)(u).$$

We have

$$\hat{\lambda}(u)(z) = \lambda(z, u) = \overline{\lambda(u, z)} = \overline{\hat{\lambda}(z)(u)} = \overline{(sy)(u)} = \overline{y(u)\bar{s}} = s\overline{y(u)}.$$

Substituting, we obtain that

$$\mu_\partial(y + \hat{\lambda}(u)) = \frac{1}{s} y(z) + \overline{y(u)} + y(u) + \lambda(u, u).$$

The last term is symmetric over $\mathbb{Z}[t^{\pm 1}]$, which implies it is of the form $a + \bar{a}$. Hence up to terms of the form $a + \bar{a}$, we have

$$\mu_\partial(y + \hat{\lambda}(u)) = \frac{1}{s} y(z) = \mu_\partial(y),$$

as desired.

Now we know that μ_∂ is well-defined, we prove that it satisfies the conditions in Definition 2.2 for it to be a quadratic refinement of the boundary symmetric linking form $\partial\lambda$. For (i), let $y \in \operatorname{coker}(\hat{\lambda})$, let $r \in \mathbb{Z}[t^{\pm 1}]$, and let $z \in H$ be such that $\hat{\lambda}(z) = sy$. Then $\hat{\lambda}(rz) = rsy = sry$. Thus

$$\mu_\partial(ry) = \frac{1}{s}((ry)(rz)) = \frac{r}{s}(y(z))\bar{r} = r\mu_\partial(y)\bar{r},$$

as desired. Next, aiming for (ii), we compute $\mu_\partial(x + y)$, for $x, y \in \operatorname{coker}(\hat{\lambda})$. Let $r, s \in \mathbb{Z}[t^{\pm 1}]$ and $w, z \in H$ be such that $\hat{\lambda}(w) = rx$ and $\hat{\lambda}(z) = sy$. Then $\hat{\lambda}(sw + ry) = srx + rsy = rs(x + y)$. Hence

$$\begin{aligned}\mu_\partial(x + y) &= \frac{1}{rs}(x + y)(sw + rz) = \frac{1}{r}x(w) + \frac{1}{s}(y)(z) + \frac{1}{r}y(w) + \frac{1}{s}x(z) \\ &= \mu_\partial(x) + \mu_\partial(y) + \partial\lambda(x, y) + \partial\lambda(y, x).\end{aligned}$$

Since $\partial\lambda$ is Hermitian, this proves (ii). Condition (iii) is immediate from the formulae. □

Remark 2.7. We note for later use that an isomorphism $F : H_0 \to H_1$ induces an isomorphism $F^{-*} := (F^*)^{-1} : H_0^* \to H_1^*$ and that if additionally the isomorphism F is an isometry, then F^{-*} descends to an isomorphism

$$\partial F := F^{-*} : \operatorname{coker}(\hat{\lambda}_0) \to \operatorname{coker}(\hat{\lambda}_1)$$

which determines an isometry of quadratic linking forms. Hence $\operatorname{Aut}(\lambda)$ acts both on $\operatorname{Aut}(\operatorname{coker}(\hat{\lambda}), \partial\lambda)$ and on the subset $\operatorname{Aut}(\operatorname{coker}(\hat{\lambda}), \partial\lambda, \partial\mu_\partial)$ by $F \cdot h = h \circ \partial F^{-1}$.

2.3. ***Algebraic unions.*** We recall the definition of the union of two Hermitian forms along an isometry of their boundary linking forms. The definition appears for the ring $\mathbb{Z}$ in [Crowley 2002, Lemma 3.6] and was generalised to the ring $\mathbb{Z}[t^{\pm 1}]$ in [Conway and Powell 2023, Construction 2.7]. The goal of this section is to prove that if the isometry preserves the quadratic refinements, then the union is an even form.

Construction 2.8. Let (H_0, λ_0) and (H_1, λ_1) be nondegenerate Hermitian forms over $\mathbb{Z}[t^{\pm 1}]$, and let $h : (\operatorname{coker}(\hat{\lambda}_0), \partial\lambda_0) \to (\operatorname{coker}(\hat{\lambda}_1), \partial\lambda_1)$ be an isometry of their boundary symmetric linking forms. Consider the pair $(H_0 \cup_h H_1, \lambda_0 \cup_h -\lambda_1)$ with

$$H_0 \cup_h H_1 := \ker\big(h\pi_0 - \pi_1 : H_0^* \oplus H_1^* \to \operatorname{coker}(\hat{\lambda}_1)\big)$$
$$\lambda_0 \cup_h -\lambda_1\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}\right) = \frac{1}{s_0} y_0(z_0) - \frac{1}{s_1} y_1(z_1) \in \mathbb{Q}(t),$$

where, since $\operatorname{coker}(\hat{\lambda}_i)$ is torsion, there exists $s_i \in \mathbb{Z}[t^{\pm 1}]$ and $z_i \in H_i$ such that $s_i x_i = \hat{\lambda}_i(z_i)$. Since the Hermitian forms λ_0 and λ_1 are nondegenerate, it is not difficult to prove that the pairing $\lambda_0 \cup_h -\lambda_1$ does not depend on the choice of s_0, s_1, z_0, z_1. One verifies that $\lambda_0 \cup_h -\lambda_1$ is a sesquilinear, Hermitian form and takes values in $\mathbb{Z}[t^{\pm 1}]$; see [Conway and Powell 2023, Proposition 2.8]. This pairing will be referred to as the *algebraic union* of λ_0 and λ_1.

Lemma 2.9. *Let (H_0, λ_0), (H_1, λ_1) and (H, λ) be nondegenerate Hermitian forms over $\mathbb{Z}[t^{\pm 1}]$, and let $h : (\operatorname{coker}(\hat{\lambda}_0), \partial\lambda_0) \to (\operatorname{coker}(\hat{\lambda}_1), \partial\lambda_1)$ be an isometry of the boundary linking forms. If $F : \lambda_0 \cong \lambda_1$ is an isometry, then there is an isometry*

$$\lambda_0 \cup_h -\lambda_1 \cong \lambda_1 \cup_{h \circ \partial F^{-1}} -\lambda_1.$$

Proof. See [Conway and Powell 2023, Proposition 2.8]. □

Lemma 2.10. *Let (H_0, λ_0) and (H_1, λ_1) be two nondegenerate even Hermitian forms over $\mathbb{Z}[t^{\pm 1}]$. Suppose that $F : (H_0, \lambda_0) \cong (H_1, \lambda_1)$ is an isometry and that*

$$h : (\operatorname{coker}(\hat{\lambda}_0), \partial\lambda_0, (\mu_\partial)_0) \to (\operatorname{coker}(\hat{\lambda}_1), \partial\lambda_1, (\mu_\partial)_1)$$

is an isometry of quadratic linking forms. Then the algebraic union $\lambda_0 \cup_h -\lambda_1$ is even.

Proof. Using Lemma 2.9 we can assume without loss of generality that $H_0 = H_1$ and $\lambda_0 = \lambda_1$ and $(\mu_\partial)_0 = (\mu_\partial)_1$. Write them both as $(H, \lambda, \mu_\partial)$. This means composing h with ∂F, but since ∂F is an isometry of quadratic linking forms too, this does not affect the argument. We abuse notation and without loss of generality use h to denote the new isometry of boundary quadratic linking forms.

It thus suffices to check that for every $x_0, x_1 \in H^*$ such that $h \circ \pi_0(x_0) = \pi_1(x_1)$, the self-intersection $\lambda \cup_h -\lambda\big((x_0, x_1), (x_0, x_1)\big)$ is of the form $a + \bar{a}$ for some $a \in \mathbb{Z}[t^{\pm 1}]$.

Pick $z_0, z_1 \in H$ and $s_0, s_1 \in \mathbb{Z}[t^{\pm 1}]$ such that $s_0 x_0 = \hat{\lambda}(z_0)$ and $s_1 x_1 = \hat{\lambda}(z_1)$. This implies both that

$$\mu_\partial(\pi_i(x_i)) = \frac{1}{s_i}(x_i(z_i)) \in Q_1(\mathbb{Z}[t^{\pm 1}], S)$$

and that

$$\lambda \cup_h -\lambda\big((x_0, x_1), (x_0, x_1)\big) = \frac{1}{s_0} x_0(z_0) - \frac{1}{s_1} x_1(z_1) \in \mathbb{Z}[t^{\pm 1}].$$

Passing to $Q_1(\mathbb{Z}[t^{\pm 1}], S)$, by the definition of μ_∂ this equals

$$\mu_\partial(\pi_0(x_0)) - \mu_\partial(\pi_1(x_1)) = \mu_\partial(\pi_0(x_0)) - \mu_\partial(h \circ \pi_0(x_0)) = 0.$$

The first equality used $h \circ \pi_0(x_0) = \pi_1(x_1)$. The second equality used that h is an isometry of quadratic linking forms.

Now we just have to note that the indeterminacy in $Q_1(\mathbb{Z}[t^{\pm 1}], S)$ consists entirely of even elements $\{a + \bar{a} \mid a \in \mathbb{Z}[t^{\pm 1}]\}$, and therefore $\lambda \cup_h -\lambda\big((x_0, x_1), (x_0, x_1)\big) \subseteq \{a + \bar{a} \mid a \in \mathbb{Z}[t^{\pm 1}]\}$. It follows that $\lambda \cup_h -\lambda$ is even, as desired. □

3. The Blanchfield form

In Section 3.1 we review the definition of the Blanchfield form and how it is related to the equivariant intersection form of a 4-manifold with fundamental group $\mathbb{Z}$. Then in Section 3.2 we review isometries of Blanchfield forms. In Section 3.3 we prove promised the spin gluing result, which demonstrates how isometries of the Blanchfield form together with a quadratic refinement can be used to ensure that the union of two spin 4-manifolds is again spin.

3.1. ***The Blanchfield form.*** We recall the definition of the Blanchfield form Bl_Y of a closed 3-manifold Y equipped with an epimorphism $\varphi : \pi_1(Y) \twoheadrightarrow \mathbb{Z}$ and how, if M is a 4-manifold with ribbon boundary, then $\mathrm{Bl}_{\partial M}$ is related to the equivariant intersection form λ_M of M.

Construction 3.1. Let Y be a closed 3-manifold and let $\varphi : \pi_1(Y) \twoheadrightarrow \mathbb{Z}$ be an epimorphism. Assume that the Alexander module $H_1(Y; \mathbb{Z}[t^{\pm 1}])$ is torsion so that, in particular, the Bockstein homomorphism $\mathrm{BS} : H^1(Y; \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]) \to H^2(Y; \mathbb{Z}[t^{\pm 1}])$ is an isomorphism. Consider the composition of Poincaré duality, the inverse Bockstein homomorphism and the evaluation homomorphism:

$$\begin{aligned}\Phi : H_1(Y; \mathbb{Z}[t^{\pm 1}]) &\xrightarrow{\mathrm{PD}, \cong} H^2(Y; \mathbb{Z}[t^{\pm 1}]) \xrightarrow{\mathrm{BS}^{-1}, \cong} H^1(Y; \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]) \\ &\xrightarrow{\mathrm{ev}} \overline{\mathrm{Hom}(H_1(Y; \mathbb{Z}[t^{\pm 1}]), \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}])}.\end{aligned}$$

Using the universal coefficient spectral sequence, one can check that the evaluation map is an isomorphism, and thus so is Φ. Thus the pairing $(x, y) \mapsto \Phi(y)(x)$ is nonsingular. It is straightforward to see that this pairing is sesquilinear. It is also Hermitian; see e.g. [Powell 2016].

Definition 3.2. Let Y be a closed 3-manifold and let $\varphi : \pi_1(Y) \twoheadrightarrow \mathbb{Z}$ be an epimorphism such that the Alexander module $H_1(Y; \mathbb{Z}[t^{\pm 1}])$ is torsion. The *Blanchfield form*

$$\mathrm{Bl}_Y : H_1(Y; \mathbb{Z}[t^{\pm 1}]) \times H_1(Y; \mathbb{Z}[t^{\pm 1}]) \to \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]$$

is the sesquilinear, nonsingular, Hermitian form defined by $\mathrm{Bl}_Y(x, y) = \Phi(y)(x)$.

Let M be a 4-manifold with $\pi_1(M) = \mathbb{Z}$, nondegenerate equivariant intersection form λ_M and ribbon boundary (meaning that the inclusion induced map $\pi_1(\partial M) \to \pi_1(M)$ is surjective). We now outline why the symmetric boundary linking form $\partial\lambda_M$ (that was described in Section 2.2) is isometric to $-\mathrm{Bl}_{\partial M}$.

As explained in [Conway and Powell 2023, Remark 3.3], the connecting homomorphism δ in the long exact sequence of the pair $(M, \partial M)$, together with Poincaré duality and the evaluation map, determines an isomorphism

$$D_M : \operatorname{coker}(\hat{\lambda}_M) \xrightarrow{\cong} H_1(\partial M; \mathbb{Z}[t^{\pm1}]), \quad [x] \mapsto \delta \circ \mathrm{PD} \circ \mathrm{ev}^{-1}(x)$$

that fits into the following commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] & H_2(M; \mathbb{Z}[t^{\pm1}]) \arrow[r, "\hat{\lambda}_M"] \arrow[d, "="'] & H_2(M; \mathbb{Z}[t^{\pm1}])^* \arrow[r] \arrow[d, "\cong"', "\mathrm{PD}\circ\mathrm{ev}^{-1}"] & \operatorname{coker}(\hat{\lambda}_M) \arrow[r] \arrow[d, "\cong"', "D_M"] & 0 \\
0 \arrow[r] & H_2(M; \mathbb{Z}[t^{\pm1}]) \arrow[r, "j"] & H_2(M, \partial M; \mathbb{Z}[t^{\pm1}]) \arrow[r, "\delta"] & H_1(\partial M; \mathbb{Z}[t^{\pm1}]) \arrow[r] & 0
\end{tikzcd} \tag{1}$$

Proposition 3.3 [Conway and Powell 2023, Proposition 3.5]. *Let M be a 4-manifold with $\pi_1(M) = \mathbb{Z}$, whose boundary is ribbon and with torsion Alexander module. The isomorphism D_M induces an isometry of symmetric linking forms*

$$D_M : \partial(H_2(M; \mathbb{Z}[t^{\pm1}]), \lambda_M) = (\operatorname{coker}(\hat{\lambda}_M), \partial\lambda_M) \xrightarrow{\cong} (H_1(\partial M; \mathbb{Z}[t^{\pm1}]), -\mathrm{Bl}_{\partial M}).$$

Construction 3.4. Suppose M is a spin 4-manifold with $\pi_1(M) = \mathbb{Z}$, whose boundary is ribbon and with torsion Alexander module. Since M is spin, the equivariant intersection form of M is even. Then $\mathrm{Bl}_{\partial M}$ admits a preferred quadratic refinement

$$\mu_{\mathrm{Bl}_{\partial M}} : H_1(\partial M; \mathbb{Z}[t^{\pm1}]) \to Q_1(\mathbb{Z}[t^{\pm1}], S), \quad x \mapsto \mu_\partial(D_M^{-1}(x)).$$

Here recall from Definition 2.5 that μ_∂ refers to the quadratic refinement of the symmetric linking form $\partial\lambda_M$; it exists because λ_M is even.

By construction

$$D_M : (\operatorname{coker}(\hat{\lambda}_M), \partial\lambda_M, \mu_\partial) \xrightarrow{\cong} (H_1(\partial M; \mathbb{Z}[t^{\pm1}]), -\mathrm{Bl}_{\partial M}, -\mu_{\mathrm{Bl}_{\partial M}})$$

is an isometry of quadratic linking forms.

3.2. ***Homotopy equivalences and isometries of the Blanchfield form.*** Given 3-manifolds Y_0, Y_1 equipped with epimorphisms $\varphi_i : \pi_1(Y_i) \twoheadrightarrow \mathbb{Z}$, we recall when homotopy equivalences of 3-manifolds induce isometries of the corresponding Blanchfield forms. We then apply these considerations to boundaries of 4-manifolds with $\pi_1 = \mathbb{Z}$.

Proposition 3.5 [Conway and Powell 2023, Proposition 3.7]. *Let Y_0, Y_1 be 3-manifolds equipped with epimorphisms $\varphi_i : \pi_1(Y_i) \twoheadrightarrow \mathbb{Z}$ and assume that the resulting*

Alexander modules are torsion for $i = 0, 1$. If an orientation-preserving homotopy equivalence $f : Y_0 \to Y_1$ satisfies $\varphi_1 \circ f_* = \varphi_0$ on $\pi_1(Y_0)$, then it induces an isometry between the Blanchfield forms

$$\tilde{f}_* : H_1(Y_0; \mathbb{Z}[t^{\pm 1}]) \to H_1(Y_1; \mathbb{Z}[t^{\pm 1}]).$$

The proof of Proposition 3.5 is fairly straightforward: the condition that

$$\varphi_1 \circ f_* = \varphi_0$$

ensures that f lifts to the infinite cyclic covers, and the required isometry is then obtained by taking the induced map on $H_1(-; \mathbb{Z}[t^{\pm 1}])$.

Remark 3.6. Let M and N be 4-manifolds with fundamental group $\mathbb{Z}$, whose boundaries are ribbon and with torsion Alexander modules.

- A consequence of Proposition 3.5 is that an orientation-preserving homotopy equivalence $f : \partial M \to \partial N$ that intertwines the epimorphisms $\pi_1(\partial M) \twoheadrightarrow \pi_1(M)$ and $\pi_1(\partial N) \twoheadrightarrow \pi_1(N)$ induces an isometry $\tilde{f}_* : \mathrm{Bl}_{\partial M} \cong \mathrm{Bl}_{\partial N}$. However, in general $\tilde{f}_*$ need not preserve the boundary quadratic refinements.
- Consider the group $\mathrm{hAut}^+_\varphi(\partial M)$ of orientation-preserving homotopy equivalences $f : \partial M \to \partial M$ that satisfy $\varphi \circ f_* = f_* : \pi_1(\partial M) \to \mathbb{Z}$. We write

  $$\mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M) \subseteq \mathrm{hAut}^+_\varphi(\partial M)$$

  for the subset consisting of homotopy equivalences such that $\tilde{f}_*$ preserves $\mu_{\mathrm{Bl}_{\partial M}}$. Proposition 3.5 implies that $\mathrm{hAut}^+_\varphi(\partial M)$ acts on $\mathrm{Aut}(\mathrm{Bl}_{\partial M})$ and that $\mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)$ acts on $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}}) \subseteq \mathrm{Aut}(\mathrm{Bl}_{\partial M})$.
- In fact $\mathrm{Aut}(\mathrm{Bl}_{\partial M})$ and $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$ also admit actions of $\mathrm{Aut}(\lambda_M)$. In more detail, we can use D_M to transport the action on $\mathrm{Aut}(\partial\lambda_M)$ (recall Remark 2.7) to an action on $\mathrm{Aut}(\mathrm{Bl}_{\partial M})$: the action of F on $h \in \mathrm{Aut}(\mathrm{Bl}_{\partial M})$ is by $F \cdot h := h \circ (D_M \circ \partial F \circ D_M^{-1})$. Since ∂F also preserves μ_∂ it follows that $\mathrm{Aut}(\partial\lambda_M)$ also acts on $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}}) \subseteq \mathrm{Aut}(\mathrm{Bl}_{\partial M})$.
- Remark 2.4 leads to an example where $\mu_{\mathrm{Bl}_{\partial M}}$, the quadratic refinement of the Blanchfield form, depends on the choice of coboundary M. Consider the 4-manifold $M := (S^2 \widetilde{\times}_8 D^2) \natural (S^1 \times D^3)$, where $S^2 \widetilde{\times}_8 D^2$ denotes the total space of the D^2-bundle over S^2 with Euler number 8. Multiplication by 3 induces an automorphism of the symmetric linking form of $\partial M = L(8, 1) \# (S^1 \times S^2)$ that is not induced by $\mathrm{Aut}(\lambda_M)$ nor by $\mathrm{Homeo}^+_\varphi(\partial M)$. Hence [Conway, Piccirillo and Powell 2022] implies there is a nonhomeomorphic 4-manifold M' with the same boundary and intersection form as M. The quadratic refinements of $\mathrm{Bl}_{\partial M}$ induced by M and M' do not even lie in the same $\mathrm{Homeo}^+_\varphi(\partial M)$ orbits, and so depend strongly on the choice of coboundary.

- It is an interesting question whether it is possible to define the quadratic refinement on $\mathrm{Bl}_{\partial M}$ intrinsically, using only the spin structure M induces on ∂M. An analogue of this result is known for the standard linking form on ∂M; see [Deloup and Massuyeau 2005, §2.4 & 2.5]. We leave this problem for later work.

3.3. ***Unions of spin* 4*-manifolds.*** We describe how the union construction from Section 2.3 and quadratic isometries of the Blanchfield form can be used to ensure that certain unions of spin 4-manifolds with fundamental group $\mathbb{Z}$ remain spin.

In order to cut down on notation, we identify $(H_1(\partial M; \mathbb{Z}[t^{\pm 1}]), -\mathrm{Bl}_{\partial M}, -\mu_{\mathrm{Bl}_{\partial M}})$ with $(\operatorname{coker}(\hat{\lambda}_M), \partial\lambda_M, \mu_\partial)$ i.e. we temporarily omit the isometry D_M mentioned in Proposition 3.3 from the notation. In particular, given an isometry $h : \mathrm{Bl}_{\partial M} \cong \mathrm{Bl}_{\partial N}$, we allow ourselves to write $\lambda_M \cup_h -\lambda_N$.

The next proposition recalls how the algebraic union can be used to understand the equivariant intersection form of a union of two 4-manifolds with fundamental group $\mathbb{Z}$.

Lemma 3.7 [Conway and Powell 2023, Proposition 3.9]. *Let M and N be two* 4*-manifolds with fundamental group $\mathbb{Z}$, nondegenerate equivariant intersection forms, and whose boundaries are ribbon. If there is an orientation-preserving homeomorphism $f : \partial M_0 \cong \partial M_1$ that intertwines the inclusion-induced epimorphisms $\pi_1(\partial M_0) \twoheadrightarrow \pi_1(M_0)$ and $\pi_1(\partial M_1) \twoheadrightarrow \pi_1(M_1)$, then there is an isometry*

$$\lambda_{M_0} \cup_{\tilde{f}_*} -\lambda_{M_1} \cong \lambda_{M_0 \cup_f -M_1}$$

Note that the intersection form being nondegenerate implies that the Alexander modules are $\mathbb{Z}[t^{\pm 1}]$-torsion, via the long exact sequences of the pairs $(M, \partial M)$ and $(N, \partial N)$ with $\mathbb{Z}[t^{\pm 1}]$ coefficients.

Proposition 3.8. *Let M and N be spin* 4*-manifolds with fundamental group $\mathbb{Z}$, nondegenerate equivariant intersection forms, and whose boundaries are ribbon. If $\lambda_M \cong \lambda_N$ and $f : \partial M \cong \partial N$ is a homeomorphism that induces an isometry between the boundary quadratic linking forms of ∂M and ∂N, then $M \cup_f -N$ is spin.*

Proof. Pick an isometry $F : \lambda_M \cong \lambda_N$. Applying successively Lemma 3.7 and Lemma 2.9 we obtain

$$\lambda_{M \cup_f -N} \cong \lambda_M \cup_{\tilde{f}_*} -\lambda_N \cong \lambda_N \cup_{\tilde{f}_* \circ \partial F^{-1}} -\lambda_N.$$

Since $\tilde{f}_*$ and ∂F preserve the quadratic linking forms, so does $\tilde{f}_* \circ \partial F^{-1}$.

As N is spin, it follows that λ_N is even. Lemma 2.10 implies that so is $\lambda_N \cup_{\tilde{f}_* \circ \partial F^{-1}} -\lambda_N$. Therefore $\lambda_{M \cup_f -N}$ is even. Since the fundamental group of $M \cup_f -N$ is $\mathbb{Z}$ (see e.g. [Conway and Powell 2023, Proposition 3.8]) and therefore has no 2-torsion, this implies that $M \cup_f -N$ is spin, as required. □

4. Stable homeomorphism

In this section, we collect some facts about stable homeomorphism of 4-manifolds with fundamental group $\mathbb{Z}$ and nonempty boundary. We then focus on the case of 4-manifolds whose boundary is ribbon and has torsion Alexander module.

Given two 4-manifolds M and N with fundamental group $\pi_1(M) = \mathbb{Z} = \pi_1(N)$ that are either both spin or both nonspin, we call a homeomorphism $f : \partial M \xrightarrow{\cong} \partial N$ a *B-compatible* homeomorphism if the diagram

$$\begin{tikzcd} \pi_1(\partial M) \arrow[rr, "f_*"] \arrow[d] & & \pi_1(\partial N) \arrow[d] \\ \pi_1(M) \arrow[r, "="] & \mathbb{Z} & \pi_1(N) \arrow[l, "="'] \end{tikzcd}$$

commutes, and if, in the case that M and N are spin, the union $M \cup_f -N$ is spin. We write $\partial M \cong_B \partial N$ if such a homeomorphism exists, and say that ∂M and ∂N are *B-homeomorphic*.

The terminology "B-compatible" and "B-homeomorphic" is motivated by our use of modified surgery below, where B is shorthand for the normal 1-type of the manifolds involved.

Lemma 4.1. *Let M and N be two 4-manifolds with fundamental group $\mathbb{Z}$ and nonempty B-homeomorphic boundaries. Suppose that M and N have the same Kirby–Siebenmann invariant and isometric equivariant intersection forms. Then M and N are stably homeomorphic.*

In particular, there is an equality

$$\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) = \frac{\{N \mid \partial N \cong_B \partial M, \pi_1(N) = \mathbb{Z}, \lambda_N \cong \lambda_M, \mathrm{ks}(N) = \mathrm{ks}(M)\}}{\textit{o.p. homotopy eq. of pairs}}.$$

Proof. Theorem 2 of [Kreck 1999] ensures that M and N are stably homeomorphic if and only if there is a B-compatible homeomorphism $f : \partial M \xrightarrow{\cong} \partial N$ such that the union $M \cup_f -N$ vanishes in the bordism group $\Omega_4^{\mathrm{STOP}}(S^1) \cong \mathbb{Z} \oplus \mathbb{Z}_2$ (when M and N are nonspin) or $\Omega_4^{\mathrm{TOPSpin}}(S^1) \cong \mathbb{Z}$ (when M and N are spin). In the nonspin case, this bordism group is detected by the signature and Kirby–Siebenmann invariant whereas in the spin case, it is detected by the signature alone. Since $\lambda_M \cong \lambda_N$, we deduce that M and N have the same signature and Wall additivity implies that the union $M \cup_f -N$ has vanishing signature. The fact that $M \cup_f -N$ has vanishing Kirby–Siebenmann invariant follows from the additivity of this invariant; see e.g. [Friedl, Nagel, Orson and Powell 2019, Theorem 8.2]. □

Given a 4-manifold M with $\pi_1(M) = \mathbb{Z}$, nondegenerate equivariant intersection form and ribbon boundary, as we will describe below, the methods of [Conway, Piccirillo and Powell 2022] produce 4-manifolds N with $\pi_1(N) = \mathbb{Z}$, ribbon boundary, $\lambda_M \cong \lambda_N$, $\mathrm{ks}(M) = \mathrm{ks}(N)$ and a preferred identification $g : \partial M \cong \partial N$. Since

$\lambda_M \cong \lambda_N$ and $\pi_1(M) = \mathbb{Z} = \pi_1(N)$ has no 2-torsion, either M and N are both spin or both nonspin. By Lemma 4.1, in order to prove that M and N are stably homeomorphic, it therefore suffices to know that in the spin case, the union $M \cup_g -N$ is spin, for which we use Proposition 3.8.

Proposition 4.2. *Let M and N be spin 4-manifolds with fundamental group $\mathbb{Z}$, nondegenerate equivariant intersection forms, and whose boundaries are ribbon. If $\lambda_M \cong \lambda_N$, and $g : \partial M \cong \partial N$ is a homeomorphism that induces an isometry between the boundary quadratic linking forms of ∂M and ∂N, then M and N are stably homeomorphic.*

Proof. Proposition 3.8 implies that the homeomorphism $\partial M \cong \partial N$ is B-compatible and the result therefore follows from Lemma 4.1. □

5. From stable homeomorphism to isometries of the Blanchfield form

In this section, M denotes a spin 4-manifold with an identification $\pi_1(M) = \mathbb{Z}$, ribbon boundary (the inclusion induced map $\varphi : \pi_1(\partial M) \to \pi_1(M)$ is surjective), and nondegenerate equivariant intersection form λ_M. As we recalled in Section 3, since λ_M is nondegenerate, the Alexander module $H_1(\partial M; \mathbb{Z}[t^{\pm 1}])$ is torsion and supports the Blanchfield form,

$$\mathrm{Bl}_{\partial M} : H_1(\partial M; \mathbb{Z}[t^{\pm 1}]) \times H_1(\partial M; \mathbb{Z}[t^{\pm 1}]) \to \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}],$$

which is nonsingular, sesquilinear, and Hermitian. The goal of this section is to describe in more detail the set $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M))$ that was mentioned in the introduction and then to prove Proposition 1.3.

We start by recalling the aforementioned actions of $\mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M)$ and $\mathrm{Aut}(\lambda_M)$ on the group $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$ of isometries of the quadratic refinement of the Blanchfield form of ∂M induced by the even form λ_M. Here recall that $\mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M)$ denotes the group of orientation-preserving homotopy equivalences $f : \partial M \to \partial M$ that satisfy $\varphi \circ f_* = f_* : \pi_1(\partial M) \to \mathbb{Z}$ and preserve $\mu_{\mathrm{Bl}_{\partial M}}$, while $\mathrm{Aut}(\lambda_M)$ denotes the set of isometries of λ_M.

- We recall the action of $\mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M)$ on $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$. Since any homotopy equivalence $f \in \mathrm{hAut}^{+}_{\varphi}(\partial M)$ satisfies $\varphi \circ f_* = f_*$, it lifts to a homotopy equivalence $\tilde{f}$ on the $\mathbb{Z}$-covers that induces a $\mathbb{Z}[t^{\pm 1}]$-linear map on homology; we denote this map by $\tilde{f}_*$. Since f is orientation-preserving, so is $\tilde{f}$ and it follows that $\tilde{f}_*$ is an isometry of the Blanchfield form. The action of f on $h \in \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$ is then by $f \cdot h = \tilde{f}_* \circ h$.
- We describe the action of $\mathrm{Aut}(\lambda_M)$ on $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$. Recall from Section 3 that the even Hermitian form λ_M determines an adjoint map

$$\hat{\lambda}_M : H_2(M; \mathbb{Z}[t^{\pm 1}]) \to H_2(M; \mathbb{Z}[t^{\pm 1}])^*,$$

and a $\mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]$-valued quadratic linking form

$$\begin{aligned} \partial\lambda_M : \operatorname{coker}(\hat{\lambda}_M) \times \operatorname{coker}(\hat{\lambda}_M) \to \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}]; \quad & \partial\lambda_M([x],[y]) = y(z)/p, \\ \mu_\partial : \operatorname{coker}(\hat{\lambda}_M) \to Q_+(\mathbb{Z}[t^{\pm 1}], S); \quad & \mu_\partial([x]) = x(z)/p. \end{aligned} \tag{2}$$

Here $p \in \mathbb{Z}[t^{\pm 1}]$ and $z \in H_2(M; \mathbb{Z}[t^{\pm 1}])$ satisfy $px = \hat{\lambda}_M(z)$. In Section 3 we also recalled the definition of the isometry

$$D_M : -\partial\lambda_M \cong \mathrm{Bl}_{\partial M},$$

and that an isometry $F \in \operatorname{Aut}(\lambda_M)$ induces an isometries $\partial F : -\partial\lambda_M \cong -\mathrm{Bl}_{\partial M}$ and $\partial F : \mu_\partial \cong -\mu_{\mathrm{Bl}_{\partial M}}$ by noting that the isomorphism $(F^*)^{-1}$ descends to an isometry on the cokernels. The action of F on $h \in \operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$ is then by $F \cdot h := h \circ (D_M \circ \partial F \circ D_M^{-1})$.

Note that the actions commute, because one acts by precomposition and the other acts by postcomposition. We obtain an action of the product $\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)$ on $\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})$. Now that we have made sense of the orbit set

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)),$$

we describe its relation to the homotopy stable class of M. Recall from Lemma 4.1 that

$$\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) = \frac{\{N \mid \partial N \cong_B \partial M, \pi_1(N) = \mathbb{Z}, \lambda_N \cong \lambda_M, \mathrm{ks}(N) = \mathrm{ks}(M)\}}{\text{o.p. hom. equiv. of pairs}}. \tag{3}$$

In order to relate $\mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M)$ to the orbit set

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)),$$

we recall a construction from [Conway, Piccirillo and Powell 2022, Construction 1] which has its origins in [Boyer 1993].

Construction 5.1. We describe a map

$$b : \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) \to \operatorname{Aut}(\mathrm{Bl}_{\partial M})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M)).$$

Given a 4-manifold $N \in \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M)$, pick a homeomorphism $g : \partial N \cong \partial M$ and an isometry $F : \lambda_M \cong \lambda_N$. Since $N \in \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M)$, it also has ribbon boundary and torsion Alexander module, thus ensuring that the isometry $D_N : -\partial\lambda_N \cong \mathrm{Bl}_{\partial N}$ is defined. Now set

$$b(N) := g_* \circ D_N \circ \partial F \in \operatorname{Aut}(\mathrm{Bl}_{\partial M})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M)).$$

One verifies that $b(N)$ is independent of the choices of F, g and the orientation-preserving homotopy equivalence class of $(N, \partial N)$. See [Conway, Piccirillo and Powell 2022] for details.

Suppose that in addition we assume the homeomorphism g is such that the induces homomorphism $g_* : H_1(\partial N; \mathbb{Z}[t^{\pm 1}]) \to H_1(\partial M; \mathbb{Z}[t^{\pm 1}])$ intertwines the quadratic refinements $\mu_{\mathrm{Bl}_{\partial N}}$ and $\mu_{\mathrm{Bl}_{\partial M}}$. To see that such a homeomorphism exists, restrict a stable homeomorphism $\Phi : M\#W_r \cong N\#W_r$ to the boundaries: this will intertwine the quadratic refinements because stabilising a Hermitian form by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ does not affect the boundary quadratic linking form. Since both D_N and ∂F preserve the quadratic refinements, it follows that b in fact defines a map

$$b : \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) \to \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M)).$$

The next proposition proves Proposition 1.3 from the introduction.

Proposition 5.2. *If M is a spin 4-manifold with $\pi_1(M) = \mathbb{Z}$, ribbon boundary and nondegenerate equivariant intersection form λ_M, then the map b from Construction 5.1 defines a surjection*

$$b : \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M) \twoheadrightarrow \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M)).$$

Proof. According to [Conway, Piccirillo and Powell 2022, Theorem 1.15], after fixing one choice of isometry $\partial\lambda_M \cong -\mathrm{Bl}_{\partial M}$ (we will use D_M) every element $\mathrm{Aut}(\mathrm{Bl}_{\partial M})/\mathrm{Aut}(\lambda_M)$ is realised by a 4-manifold N with $\pi_1(N) = \mathbb{Z}$, ribbon boundary ∂N homeomorphic to ∂M via a homeomorphism $g : \partial N \cong \partial M$, equivariant intersection form λ_N isometric to λ_M via an isometry $F : \lambda_M \cong \lambda_N$, and $\mathrm{ks}(N) = \mathrm{ks}(M)$. In particular we can realise every element

$$b \in \mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M))$$

by such a manifold N.

As indicated in [Conway, Piccirillo and Powell 2022, Proof of Proposition 4.14], we have $b = g_* \circ D_N \circ \partial F$. Since b, D_N, and ∂F are isometries that preserve quadratic refinements, so does g.

Proposition 4.2 now ensures that M and N are stably homeomorphic. We have therefore produced $N \in \mathcal{S}^{\mathrm{st}}_{h^+,\lambda}(M)$ with $b(N) = b$, as required. □

6. Infinite automorphism sets

In this section, we conclude the proof of Theorem 1.1 by showing that

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M))$$

is countably infinite when

$$M = X_{2q}(U)\natural(S^1 \times D^3),$$

with q an odd prime. Here $X_{2q}(U)$ denotes the $2q$-trace on the unknot U, i.e. the smooth 4-manifold obtained from D^4 by attaching a $2q$-framed 2-handle along the unknot. The plan is to first study $\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/\operatorname{Aut}(\lambda_M)$ and then to consider the action by the self-homotopy equivalences of ∂M. In fact we will show in Lemma 6.3 that $\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/\operatorname{Aut}(\lambda_M) = \operatorname{Aut}(\mathrm{Bl}_{\partial M})/\operatorname{Aut}(\lambda_M)$. So first we will consider the latter set, ignoring quadratic refinements for a moment.

To study $\operatorname{Aut}(\mathrm{Bl}_{\partial M})/\operatorname{Aut}(\lambda_M)$, recall from Section 5 that $\mathrm{Bl}_{\partial M}$ is isometric to the linking form $\partial\lambda_M$ defined in (2), at the start of Section 5. In particular, the isometry $D_M : -\partial\lambda_M \cong \mathrm{Bl}_{\partial M}$ induces a bijection

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M})/\operatorname{Aut}(\lambda_M) \cong \operatorname{Aut}(\partial\lambda_M)/\operatorname{Aut}(\lambda_M).$$

For rank one forms (such as $(H_2(M; \mathbb{Z}[t^{\pm 1}]), \lambda_M) = (\mathbb{Z}[t^{\pm 1}], \lambda_{2q})$), this set admits a particularly convenient description.

Given a ring R with involution $x \mapsto \bar{x}$, the group of *unitary units* $U(R)$ refers to those $u \in R$ such that $u\bar{u} = 1$, with the group operation given by restricting the multiplication on R. For example, when $R = \mathbb{Z}[t^{\pm 1}]$, all units are unitary and are of the form $\pm t^k$ with $k \in \mathbb{Z}$. In what follows, we make no distinction between rank one Hermitian forms and symmetric Laurent polynomials. The next lemma follows by unwinding the definition of $\operatorname{Aut}(\partial\lambda)$; see also [Conway and Powell 2023, Remark 1.16; Conway, Piccirillo and Powell 2022, Lemma 7.1].

Lemma 6.1. *If $\lambda \in \mathbb{Z}[t^{\pm 1}]$ is a symmetric Laurent polynomial, then*

$$\operatorname{Aut}(\partial\lambda)/\operatorname{Aut}(\lambda) = U(\mathbb{Z}[t^{\pm 1}]/\lambda)/U(\mathbb{Z}[t^{\pm 1}]).$$

Proposition 6.2. *Given an odd prime q, the map*

$$\Theta : \mathbb{Z} \xrightarrow{\cong} U(\mathbb{Z}[t^{\pm 1}]/2q)/U(\mathbb{Z}[t^{\pm 1}]), \quad n \mapsto (q-1)t^n + q$$

is a group isomorphism.

Proof. One verifies that $(q-1)t^n + q$ is a unitary unit by using that $q(q-1) \equiv 0 \bmod 2q$ (recall that q is odd). We then check that Θ is a homomorphism:

$$\begin{aligned} n+m \mapsto ((q-1)t^n + q)((q-1)t^m + q) &= (q-1)^2 t^{n+m} + q(q-1)(t^m + t^n) + q^2 \\ &\sim -(q-1)t^{n+m} - q \\ &\sim (q-1)t^{n+m} + q. \end{aligned}$$

Here the penultimate equivalence uses that $(q-1)^2 = q(q-1) - (q-1) \equiv -(q-1) \bmod 2q$ and $q^2 \equiv q \equiv -q \bmod 2q$. The last equivalence uses that $-1 \in U(\mathbb{Z}[t^{\pm 1}])$.

Next we show that Θ is injective. If $(q-1)t^n + q$ were trivial, we would have $(q-1)t^n + q = \pm t^k \in \mathbb{Z}[t^{\pm 1}]/2q$ for some k, but this is true only if $n = 0$.

Now we show that Θ is surjective. An explicit verification shows that the following map is an isomorphism:

$$U(\mathbb{Z}[t^{\pm 1}]/2) \times U(\mathbb{Z}[t^{\pm 1}]/q) \to U(\mathbb{Z}[t^{\pm 1}]/2q), \quad (a, b) \mapsto qa - (q-1)b.$$

To see this one should check that $(qa-(q-1)b)(q\bar{a}-(q-1)\bar{b}) \equiv 1$ when $a\bar{a} = 1 = b\bar{b}$, which implies that the map lands in the claimed target. The inverse is given by $x \mapsto ([x]_2, [x]_q)$, i.e. considering the coefficients modulo 2 and q respectively. Checking that this is the inverse homomorphism implies that the map is an isomorphism as asserted.

The units of $\mathbb{Z}[t^{\pm 1}]/2$ are of the form t^m for $m \in \mathbb{Z}$. On the other hand, since q is an odd prime, the unitary units of $\mathbb{Z}[t^{\pm 1}]/q$ are of the form $\pm t^n$ for $n \in \mathbb{Z}$. It follows that

$$U(\mathbb{Z}[t^{\pm 1}]/2q) \cong \{qt^m + (q-1)\varepsilon t^n \mid n, m \in \mathbb{Z}, \varepsilon \in \{\pm 1\}\}.$$

Passing to the quotient by $U(\mathbb{Z}[t^{\pm 1}])$ yields the required isomorphism, because once we can multiply by $\pm t^k$ for any $k \in \mathbb{Z}$, we have $qt^m + (q-1)\varepsilon t^n \sim (q-1)\varepsilon t^{n-m} + q$. Also

$$-(q-1)t^{n-m} + q \sim -(q-1)t^{n-m} - q \sim (q-1)t^{n-m} + q,$$

so we can ignore the ε, and every element of $U(\mathbb{Z}[t^{\pm 1}]/2q)$ is of the form $(q-1)t^k + q$ for some $k \in \mathbb{Z}$. So Θ is indeed surjective, which completes the proof that Θ is an isomorphism. □

Lemma 6.3. *Given an odd prime $q \in \mathbb{Z}$, for the Hermitian form $\lambda = 2q \in \mathbb{Z}[t^{\pm 1}]$, one has*

$$\operatorname{Aut}(\partial\lambda, \mu_\partial)/\operatorname{Aut}(\lambda) = \operatorname{Aut}(\partial\lambda)/\operatorname{Aut}(\lambda).$$

Proof. The inclusion $\operatorname{Aut}(\partial\lambda, \mu_\partial)/\operatorname{Aut}(\lambda) \subseteq \operatorname{Aut}(\partial\lambda)/\operatorname{Aut}(\lambda)$ always holds. So, thanks to Proposition 6.2, the lemma reduces to proving that for every $n \in \mathbb{Z}$ multiplication by $(q-1)t^n + q$ as a map $\mathbb{Z}[t^{\pm 1}]/2q \to \mathbb{Z}[t^{\pm 1}]/2q$ preserves the quadratic refinement

$$\mu_\partial(x) = \frac{1}{2q}x.$$

Writing $q = 2k+1$, a direct calculation in $Q_1(\mathbb{Z}[t^{\pm 1}], S)$ now shows that

$$\begin{aligned}\frac{1}{2q}((q-1)t^n + q)((q-1)t^{-n} + q) &= \frac{1}{2q}((q-1)^2 + q^2) + \frac{1}{2q}(q(q-1)(t^n + t^{-n})) \\ &= \frac{1}{2q} + 2k + k(t^n + t^{-n}) \equiv \frac{1}{2q}.\end{aligned}$$

This concludes the proof of the lemma. □

For q an odd prime, the combination of Lemma 6.1, Proposition 6.2, and Lemma 6.3 implies that for $M = X_{2q}(U)\natural(S^1 \times D^3)$ we have

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/\operatorname{Aut}(\lambda_M) = \operatorname{Aut}(\mathrm{Bl}_{\partial M})/\operatorname{Aut}(\lambda_M) \cong \mathbb{Z}.$$

We now study the effect of factoring out by $\mathrm{hAut}^{+,\mathrm{q}}_{\varphi}(\partial M)$.

Remark 6.4. Our strategy will be to determine $\mathrm{hAut}^+_\varphi(\partial M)$ and show that the effect of its action on $\mathrm{Aut}(\mathrm{Bl}_{\partial M})/\mathrm{Aut}(\lambda_M)$ is given by multiplication by elements of the form $\pm t^n$, $n \in \mathbb{Z}$. This will automatically imply that $\mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)$ also acts on $\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/\mathrm{Aut}(\lambda_M)$ by multiplication by elements of the form $\pm t^n$.

In fact we will make this argument in a slightly more general setting. Consider the 3-manifold $Y := N\#(S^1 \times S^2)$, where N is a 3-manifold with finite fundamental group. We fix an identification $H_1(S^1 \times S^2) = \mathbb{Z}$, an identification $\pi_1(Y) = \pi_1(N) * \mathbb{Z}$, and consider the finite abelian group $A := TH_1(Y) \cong H_1(N)$. Let $\varphi : \pi_1(Y) \to H_1(Y)/TH_1(Y) = \mathbb{Z}$ be the canonical projection onto the free part of $H_1(Y)$. In what follows, to distinguish $H_1(S^1 \times S^2) = \mathbb{Z}$ from the free $\mathbb{Z}$-factor of $\pi_1(Y) \cong \pi_1(N) * \mathbb{Z}$, we will exclusively write $H_1(S^1 \times S^2)$ as $\langle t \rangle$.

Summarising the notation, we have

$$A := TH_1(Y) \cong H_1(N), \quad \varphi : \pi_1(Y) \twoheadrightarrow \langle t \rangle,$$
$$\theta : \pi_1(N) \xrightarrow{\mathrm{ab}} H_1(\pi_1(N)) = H_1(N) = A.$$

The example we have in mind is $Y_q = \partial M = L(2q, 1)\#(S^1 \times S^2)$, where $M = X_{2q}(U)\natural S^1 \times D^3$, so that $A \cong \mathbb{Z}/2q$ and $\varphi : \pi_1(Y_q) \twoheadrightarrow \langle t \rangle$ coincides with the inclusion induced map $\pi_1(\partial M) \twoheadrightarrow \pi_1(M) = \mathbb{Z}$.

Returning to the more general setting where $Y = N\#(S^1 \times S^2)$ with N a 3-manifold with $\pi_1(N)$ finite, the epimorphism $\varphi : \pi_1(Y) \twoheadrightarrow \langle t \rangle$ induces an infinite cyclic cover Y^∞ with

$$H_1(Y^\infty) \cong H_1(Y; \mathbb{Z}[t^{\pm 1}]) \cong H_1(\ker(\varphi)).$$

Our goal is now to describe the isomorphism type of this $\mathbb{Z}[t^{\pm 1}]$-module (this is the content of Construction 6.5 and Lemma 6.6 below) and to then deduce the effect of the action of $\mathrm{hAut}^+_\varphi(Y)$ on $H_1(Y; \mathbb{Z}[t^{\pm 1}])$ in Proposition 6.7.

In what follows, we write $A[t^{\pm 1}]$ for the abelian group of Laurent polynomials with coefficients in the finite abelian group A.

Construction 6.5. We construct a group homomorphism $\Psi : A[t^{\pm 1}] \to H_1(\ker(\varphi))$. Elements of $A[t^{\pm 1}]$ are of the form $\sum_i a_i t^i$ with $a_i \in A$. As the map $\varphi : \pi_1(N)*\mathbb{Z} \to \mathbb{Z}$ is surjective, we can write each t^i as $\varphi(g_i) = t^i$ for some $g_i \in \pi_1(N) * \mathbb{Z}$. The abelianisation $\theta : \pi_1(N) \to A = H_1(\pi_1(N))$ is also surjective, so we can write each $a \in A$ as $a = \theta(p)$ for some $p \in \pi_1(N)$. We can therefore write an element of $A[t^{\pm 1}]$ as $\sum_i \theta(p_i)\varphi(g_i)$. Since $p_i \in \pi_1(N)$, $g_i \in \pi_1(N) * \mathbb{Z}$ and $A \subseteq \ker(\varphi)$, we can consider the element $g_i p_i g_i^{-1}$ as an element of $\ker(\varphi) \subseteq \pi_1(N) * \mathbb{Z}$ and use $[g_i p_i g_i^{-1}] \in H_1(\ker(\varphi))$ to denote its image in the abelianisation. Define the map Ψ as

$$\Psi : A[t^{\pm 1}] \to H_1(\ker(\varphi)), \quad \sum_i a_i t^i = \sum_i \theta(p_i)\varphi(g_i) \mapsto \sum_i [g_i p_i g_i^{-1}].$$

We show that Ψ does not depend on the choice of the p_i and the g_i. First we argue that the definition of Ψ does not depend on the choice of the p_i. It suffices to show that if $\theta(p)=\theta(p')$, then $\Psi(\theta(p)\varphi(g))=\Psi(\theta(p')\varphi(g))$ for every $g \in \pi_1(N)*\mathbb{Z}$. Since $\theta(p(p')^{-1})=0$, we know that pp'^{-1} lies in the commutator subgroup $\pi_1(N)^{(1)}=[\pi_1(N),\pi_1(N)]$. Therefore, since $\pi_1(N)^{(1)}$ is normal, $gpp'^{-1}g^{-1}=(gpg^{-1})(gp'^{-1}g^{-1}) \in \pi_1(N)^{(1)}$ for all $g \in \pi_1(N)*\mathbb{Z}$. Since $\pi_1(N) \subseteq \ker(\varphi)$, it follows that $\pi_1(N)^{(1)} \subseteq (\ker(\varphi))^{(1)}$, and therefore $(gpg^{-1})(gp'^{-1}g^{-1}) \in (\ker(\varphi))^{(1)}$, from which it follows $(gpg^{-1})(gp'^{-1}g^{-1})=0 \in H_1(\ker(\varphi))$. We deduce that $[gpg^{-1}]=[gp'g^{-1}] \in H_1(\ker(\varphi))$ and thus

$$\Psi(\theta(p)\varphi(g))=[gpg^{-1}]=[gp'g^{-1}]=\Psi(\theta(p')\varphi(g)) \in H_1(\ker(\varphi)).$$

This proves that Ψ does not depend on the choice of the p_i.

Next, we argue that the definition of Ψ does not depend on the choice of the g_i. This time, it suffices to prove that if $\varphi(g)=\varphi(g')$ and $p \in \pi_1(N)$, then $\Psi(\theta(p)\varphi(g))=\Psi(\theta(p)\varphi(g'))$. This latter equality holds if and only if $[gpg^{-1}g'p^{-1}g'^{-1}]=0 \in H_1(\ker(\varphi))$, which in turn, by conjugating with g^{-1}, holds if and only if $[pg^{-1}g'p^{-1}g'^{-1}g]=0 \in H_1(\ker(\varphi))$. But since $pg^{-1}g'p^{-1}g'^{-1}g$ is a commutator of p and $g^{-1}g'$, which both lie in $\ker(\varphi)$, we indeed obtain $[pg^{-1}g'p^{-1}g'^{-1}g]=0 \in H_1(\ker(\varphi))$.

This concludes the verification that Ψ does not depend on any of the choices we made. One also verifies readily that Ψ is a group homomorphism. This completes Construction 6.5.

As in Construction 6.5, for each $h \in \langle t\rangle$, we fix a $g \in \pi_1(N)*\mathbb{Z}$ such that $\varphi(g)=h$. This choice will be used again in the next lemma which establishes that the map Ψ is an isomorphism.

Lemma 6.6. *The map $\Psi : A[t^{\pm 1}] \to H_1(\ker(\varphi))$ from Construction 6.5 is an isomorphism.*

Proof. We construct an inverse $\Theta : H_1(\ker(\varphi)) \to A[t^{\pm 1}]$ to Ψ. A word $w \in \ker(\varphi) \subseteq \pi_1(N)*\mathbb{Z}$ representing an element of $H_1(\ker(\varphi))$ is a product of elements of $\pi_1(N)$ and $\mathbb{Z}$.

By introducing cancelling pairs of the type $g_i^{-1}g_i$ in between each occurrence of a $p'_k \in \pi_1(N)$ in w, we can arrange that for some elements $\tilde{g}_k \in \pi_1(N)*\mathbb{Z}$ and $p'_k \in \pi_1(N)$, the word w is of the form

$$w=\prod_k \tilde{g}_k p'_k \tilde{g}_k^{-1}.$$

Here it is crucial to use that $w \in \ker(\varphi)$. For example if $w=p'_1n_1p'_2n_2p'_3n_3$, for $p_i \in \pi_1(N)$ and $n_j \in \mathbb{Z}$, then since $w \in \ker(\varphi)$ we know that $n_3=(n_1n_2)^{-1}=(n_2n_1)^{-1}$. Therefore we can express w as $w=p'_1n_1p'_2n_1^{-1}(n_1n_2)p'_3(n_1n_2)^{-1}$.

As was mentioned before the lemma, we fixed a preferred $g_j \in \pi_1(N) * \mathbb{Z}$ with $\varphi(g_j) = \varphi(\tilde{g}_k)$. Arguing as in Construction 6.5 (when we showed that the choice of the g_i is immaterial), up to commutators in $[\ker(\varphi), \ker(\varphi)]$, we can replace $\tilde{g}_k p'_k \tilde{g}_k^{-1}$ with $g_j p'_k g_j^{-1}$. Next, working in $H_1(\ker(\varphi)) = \ker(\varphi)_{ab}$ and collecting terms with the same conjugating element g_j, we obtain an element of the form $\sum_j [g_j p_j g_j^{-1}]$, where $p_j = \prod_{\{k \mid \varphi(\tilde{g}_k) = \varphi(g_j)\}} p'_k$. We can therefore define a map

$$\Theta : H_1(\ker(\varphi)) \to A[t^{\pm 1}], \quad [w] \mapsto \sum_j \theta(p_j)\varphi(g_j).$$

One checks that the map Θ is a homomorphism and is the inverse to Ψ. Thus Ψ is an isomorphism. □

We are now able to describe the action of $\mathrm{hAut}^+_\varphi(Y)$ on $H_1(Y; \mathbb{Z}[t^{\pm 1}])$.

Proposition 6.7. *Let $at^\ell \in H_1(Y; \mathbb{Z}[t^{\pm 1}]) \cong A[t^{\pm 1}]$. The action of $f \in \mathrm{hAut}^+_\varphi(Y)$ sends $at^\ell \mapsto a' t^{k+\ell}$, for some $k \in \mathbb{Z}$ and for some element $a' \in A$ having the same order as a.*

Proof. As in Construction 6.5, we can represent any element of $A[t^{\pm 1}]$ as a sum of $\theta(p)\varphi(g)$, where $p \in \pi_1(N)$ and $g \in \pi_1(N) * \mathbb{Z}$. We will describe $f_*(\theta(p)\varphi(g))$.

In fact, since we have the commutative diagram of isomorphisms

$$\begin{tikzcd}
A[t^{\pm 1}] \arrow[r, "\cong", "\Psi"'] \arrow[d, "f_*"', "\cong"] & H_1(\ker(\varphi)) \arrow[d, "f_*"', "\cong"] \\
A[t^{\pm 1}] \arrow[r, "\cong", "\Psi"'] & H_1(\ker(\varphi))
\end{tikzcd}$$

it is equivalent to describe $\Psi^{-1} \circ f_* \circ \Psi(\theta(p)\varphi(g))$. First, the definition of Ψ implies that $\Psi(\theta(p)\varphi(g)) = [gpg^{-1}] \in H_1(\ker(\varphi))$. Applying f_*, we then obtain $[f_*(g) f_*(p) f_*(g)^{-1}] \in H_1(\ker(\varphi))$.

But now, under an isomorphism $\pi_1(N) * \mathbb{Z} \xrightarrow{\cong} \pi_1(N) * \mathbb{Z}$, every element of $\pi_1(N)$ is sent to an element of finite order, since $\pi_1(N)$ is finite. This implies that for every $p \in \pi_1(N)$, we have that $f_*(p) = hp'h^{-1} \in \pi_1(N) * \mathbb{Z}$ for some $p' \in \pi_1(N)$ and some $h \in \pi_1(N) * \mathbb{Z}$. This follows by considering the cyclic subgroup generated by $f_*(p)$ and applying the Kurosh subgroup theorem, which implies that a finite subgroup of a free product of nontrivial groups is a conjugate of a finite subgroup of one of the factors.

Next, since f_* is an isomorphism, $[f_*(p)] = [hp'h^{-1}]$ has the same order as $[p]$ in $H_1(\ker(\varphi))$. Since they are conjugate, in $\pi_1(N) * \mathbb{Z}$, we know that $hp'h^{-1}$ and p' have the same order. We claim that $[hp'h^{-1}]$ has the same order as $[p']$ in $H_1(\ker(\varphi))$.

To prove the claim, suppose that $\mathrm{ord}([p']) = k$. Then $[(p')^k] = 0 \in H_1(\ker(\varphi))$, i.e. $(p')^k \in \ker(\varphi)^{(1)}$. Since $\ker(\varphi)$ is normal, for every $x \in \ker(\varphi)$ we have that

$hxh^{-1} \in \ker(\varphi)$, and therefore since $h[x, y]h^{-1} = [hxh^{-1}, hyh^{-1}]$, for every $z \in \ker(\varphi)^{(1)}$ we have that $hzh^{-1} \in \ker(\varphi)^{(1)}$. Thus $h(p')^k h^{-1} = (hp'h^{-1})^k \in \ker(\varphi)^{(1)}$, and therefore $\operatorname{ord}([hp'h^{-1}]) \leq k = \operatorname{ord}([p'])$. Since p' is also a conjugate of $hp'h^{-1}$, by symmetry we also have $\operatorname{ord}([p']) \leq \operatorname{ord}([hp'h^{-1}])$, and so we have equality. This completes the proof of the claim.

The claim implies that in $H_1(\ker(\varphi))$ we have

$$\operatorname{ord}([p']) = \operatorname{ord}([hp'h^{-1}]) = \operatorname{ord}([f_*(p)]) = \operatorname{ord}([p]).$$

Returning to the main arc of the proof, so far we have

$$f_* \circ \Psi(\theta(p)\varphi(g)) = [f_*(g) f_*(p) f_*(g)^{-1}] = [f_*(g) hp'h^{-1} f_*(g)^{-1}]$$

and it remains to apply Ψ^{-1}. The effect of Ψ^{-1} is $\theta(p')\varphi(h)\varphi(f_*(g)) \in A[t^{\pm 1}]$. Since $f \in \mathrm{hAut}^+_\varphi(Y)$, we have $\varphi \circ f_* = \varphi$ and therefore

$$f_*(\theta(p)\varphi(g)) = \Psi^{-1} \circ f_* \circ \Psi(\theta(p)\varphi(g)) = \theta(p')\varphi(h)\varphi(g) \in A[t^{\pm 1}].$$

We can now calculate $f_*(at^\ell)$. Pick $g \in \pi_1(N) * \mathbb{Z}$ and $p \in \pi_1(N)$ such that we have $\varphi(g) = t^\ell$ and $\theta(p) = a$. Now $f_*(at^\ell) = f_*(\theta(p)\varphi(g)) = \theta(p')\varphi(h)t^\ell$, so the lemma follows by writing $\varphi(h) = t^k$ and $a' := \theta(p')$. Then since $[p']$ has the same order as $[p]$, it follows that a' has the same order as a. □

We can now prove the main result of this section.

Proposition 6.8. *Fix an odd prime q. For $M = X_{2q}(U) \natural (S^1 \times D^3)$, the sets*

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M)),$$
$$\operatorname{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M))$$

are countably infinite.

Proof. Fix identifications $\pi_1(M) = \mathbb{Z}$ and $(H_2(M; \mathbb{Z}[t^{\pm 1}]), \lambda_M) = (\mathbb{Z}[t^{\pm 1}], \lambda_{2q})$. Lemma 6.1 implies that $\operatorname{Aut}(\mathrm{Bl}_{\partial M}) / \operatorname{Aut}(\lambda_M) = U(\mathbb{Z}[t^{\pm 1}]/2q)/U(\mathbb{Z}[t^{\pm 1}])$. We know from Proposition 6.2 that $U(\mathbb{Z}[t^{\pm 1}]/2q)/U(\mathbb{Z}[t^{\pm 1}]) \cong \mathbb{Z}$, every element of which is of the form $(q-1)t^n + q$ with $n \in \mathbb{Z}$. We will now show that there is a bijection of sets

$$\operatorname{Aut}(\mathrm{Bl}_{\partial M})/(\operatorname{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M)) \cong \mathbb{Z}.$$

In the notation of Proposition 6.7, we have $N = L(2q, 1)$ with $\pi_1(L(2q, 1)) \cong \mathbb{Z}/2q$ as well as $A = H_1(\pi_1(N)) = \pi_1(N) = \mathbb{Z}/2q$.

Using Proposition 6.7, we will argue that any automorphism of the group $H_1(\partial M; \mathbb{Z}[t^{\pm 1}]) \cong (\mathbb{Z}/2q)[t^{\pm 1}]$ induced by a homotopy equivalence $f \in \operatorname{Aut}^+_\varphi(\partial M)$ is of the form $p(t) \mapsto \pm t^k p(t)$, for some $k \in \mathbb{Z}$. To see this, given $p(t) \in (\mathbb{Z}/2q)[t^{\pm 1}]$, by $\mathbb{Z}[t^{\pm 1}]$-linearity of f_* we have $f_*(p(t)) = p(t) f_*(1)$. By Proposition 6.7,

$f_*(1) = a \cdot t^k$, for some $k \in \mathbb{Z}$ and some $a \in \mathbb{Z}/2q$. We need to show that $a = \pm 1$. Since f_* is an isometry of $\mathrm{Bl}_{\partial M}$, we also know that $a^2 = 1 \in \mathbb{Z}[t^{\pm 1}]/2q$; this holds because

$$\begin{aligned}\frac{-1}{2q} &= \mathrm{Bl}_{\partial M}(1, 1) = \mathrm{Bl}_{\partial M}(f_*(1), f_*(1)) \\ &= \mathrm{Bl}_{\partial M}(a \cdot t^k, a \cdot t^k) = \frac{-a^2}{2q} \in \mathbb{Q}(t)/\mathbb{Z}[t^{\pm 1}],\end{aligned}$$

which implies that $a^2 = 1 \in \mathbb{Z}[t^{\pm 1}]/2q$. Then since $a \in \mathbb{Z}/2q$ we have that $a^2 = 1 \in \mathbb{Z}/2q$. Here we used that $\mathrm{Bl}_{\partial M} \cong -\partial\lambda_{2q}$ to compute the Blanchfield form [Conway and Powell 2023, Proposition 3.5].

However the only elements of $A = \mathbb{Z}/2q$ with $a^2 = 1$ are $\pm 1 \in \mathbb{Z}/2q$. Indeed such an a belongs to $U(\mathbb{Z}/2q) \cong U(\mathbb{Z}/q) \times U(\mathbb{Z}/2)$. However $U(\mathbb{Z}/2)$ is trivial, so in fact $U(\mathbb{Z}/2q) \cong U(\mathbb{Z}/q)$. We will show that $U(\mathbb{Z}/q) = \{\pm 1\}$. To see this, recall that for q an odd prime the units $(\mathbb{Z}/q)^\times$ is a cyclic group of order $q - 1$, and in such a group there is precisely one element of order 2. Taken together with the trivial element there are therefore precisely two solutions to $x^2 = 1 \in (\mathbb{Z}/q)^\times$, namely ± 1. So we see that $U(\mathbb{Z}/2q) \cong U(\mathbb{Z}/q) = \{\pm 1\}$. It follows that $a = \pm 1$ and

$$f_*(p(t)) = p(t) f_*(1) = \pm t^k p(t),$$

as asserted above. In particular, observe that the action of a homotopy equivalence $f \in \mathrm{hAut}^+_\varphi(\partial M)$ is the same as the action by an element of $\mathrm{Aut}(\lambda_M) \cong U(\mathbb{Z}[t^{\pm 1}])$. We deduce that

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M)) \cong \mathrm{Aut}(\mathrm{Bl}_{\partial M})/\mathrm{Aut}(\lambda_M).$$

But in Proposition 6.2 we computed the latter set to be

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M})/\mathrm{Aut}(\lambda_M) \cong U(\mathbb{Z}[t^{\pm 1}]/2q)/U(\mathbb{Z}[t^{\pm 1}]) \cong \mathbb{Z}.$$

The inverse of these isomorphisms sends $n \in \mathbb{Z}$ to the automorphism given by multiplying by $(q-1)t^n + q$. Since the action of an element of $\mathrm{hAut}^+_\varphi(\partial M)$ is the same as the action by an element $\mathrm{Aut}(\lambda_M) \cong U(\mathbb{Z}[t^{\pm 1}])$, the same can be said for elements of $\mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M) \subseteq \mathrm{hAut}^+_\varphi(\partial M)$. As we mentioned in Remark 6.4, Lemma 6.3 now implies that

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M}, \mu_{\mathrm{Bl}_{\partial M}})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^{+,\mathrm{q}}_\varphi(\partial M)) \cong \mathrm{Aut}(\mathrm{Bl}_{\partial M})/(\mathrm{Aut}(\lambda_M) \times \mathrm{hAut}^+_\varphi(\partial M)).$$

The second assertion in Proposition 6.8 therefore follows from the first. $\square$

Acknowledgements

We thank the referee for helpful comments which helped improve the paper. Conway was partially supported by the NSF Standard Grant DMS-2303674. Powell was partially supported by the EPSRC New Investigator grant EP/T028335/2 and EPSRC New Horizons grant EP/V04821X/2.

References

[Boyer 1986] S. Boyer, "Simply-connected 4-manifolds with a given boundary", *Trans. Amer. Math. Soc.* **298**:1 (1986), 331–357. MR Zbl

[Boyer 1993] S. Boyer, "Realization of simply-connected 4-manifolds with a given boundary", *Comment. Math. Helv.* **68**:1 (1993), 20–47. MR Zbl

[Cheeger and Kister 1970] J. Cheeger and J. M. Kister, "Counting topological manifolds", *Topology* **9** (1970), 149–151. MR Zbl

[Conway and Powell 2023] A. Conway and M. Powell, "Embedded surfaces with infinite cyclic knot group", *Geom. Topol.* **27**:2 (2023), 739–821. MR Zbl

[Conway, Crowley, Powell and Sixt 2021] A. Conway, D. Crowley, M. Powell, and J. Sixt, "Stably diffeomorphic manifolds and modified surgery obstructions", 2021. arXiv 2109.05632

[Conway, Crowley, Powell and Sixt 2023] A. Conway, D. Crowley, M. Powell, and J. Sixt, "Simply connected manifolds with large homotopy stable classes", *J. Aust. Math. Soc.* **115**:2 (2023), 172–203. MR Zbl

[Conway, Piccirillo and Powell 2022] A. Conway, L. Piccirillo, and M. Powell, "4-manifolds with boundary and fundamental group $\mathbb{Z}$", 2022. arXiv 2205.12774

[Crowley 2002] D. Crowley, *The classification of highly connected manifolds in dimensions 7 and 15*, Ph.D. thesis, Indiana University, 2002, available at https://www.proquest.com/docview/305508775.

[Deloup and Massuyeau 2005] F. Deloup and G. Massuyeau, "Quadratic functions and complex spin structures on three-manifolds", *Topology* **44**:3 (2005), 509–555. MR Zbl

[Freedman 1982] M. H. Freedman, "The topology of four-dimensional manifolds", *J. Differential Geometry* **17**:3 (1982), 357–453. MR Zbl

[Freedman and Quinn 1990] M. H. Freedman and F. Quinn, *Topology of 4-manifolds*, Princeton Mathematical Series **39**, Princeton University Press, 1990. MR Zbl

[Friedl, Nagel, Orson and Powell 2019] S. Friedl, M. Nagel, P. Orson, and M. Powell, "A survey of the foundations of four-manifold theory in the topological category", 2019. arXiv 1910.07372

[Hambleton and Kreck 1993] I. Hambleton and M. Kreck, "Cancellation, elliptic surfaces and the topology of certain four-manifolds", *J. Reine Angew. Math.* **444** (1993), 79–100. MR Zbl

[Hambleton, Kreck and Teichner 2009] I. Hambleton, M. Kreck, and P. Teichner, "Topological 4-manifolds with geometrically two-dimensional fundamental groups", *J. Topol. Anal.* **1**:2 (2009), 123–151. MR Zbl

[Kreck 1999] M. Kreck, "Surgery and duality", *Ann. of Math.* (2) **149**:3 (1999), 707–754. MR Zbl

[Kreck and Schafer 1984] M. Kreck and J. A. Schafer, "Classification and stable classification of manifolds: some examples", *Comment. Math. Helv.* **59**:1 (1984), 12–38. MR Zbl

[Powell 2016] M. Powell, "Twisted Blanchfield pairings and symmetric chain complexes", *Q. J. Math.* **67**:4 (2016), 715–742. MR Zbl

[Ranicki 1981] A. Ranicki, *Exact sequences in the algebraic theory of surgery*, Mathematical Notes **26**, Princeton University Press, 1981. MR Zbl

[Wall 1999] C. T. C. Wall, *Surgery on compact manifolds*, 2nd ed., Mathematical Surveys and Monographs **69**, American Mathematical Society, Providence, RI, 1999. MR Zbl

Received October 7, 2022. Revised September 18, 2023.

ANTHONY CONWAY
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TX
UNITED STATES
anthony.conway@austin.utexas.edu

DIARMUID CROWLEY
SCHOOL OF MATHEMATICS & STATISTICS
UNIVERSITY OF MELBOURNE
PARKVILLE, VIC
AUSTRALIA
dcrowley@unimelb.edu.au

MARK POWELL
SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF GLASGOW
GLASGOW
UNITED KINGDOM
mark.powell@glasgow.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.239

COUNTEREXAMPLES TO THE NONSIMPLY CONNECTED DOUBLE SOUL CONJECTURE

JASON DEVITO

A double disk bundle is any smooth closed manifold obtained as the union of the total spaces of two disk bundles, glued together along their common boundary. The double soul conjecture asserts that a closed simply connected manifold admitting a metric of nonnegative sectional curvature is necessarily a double disk bundle. We study a generalization of this conjecture by dropping the requirement that the manifold be simply connected. Previously, a unique counterexample was known to this generalization, the Poincaré dodecahedral space S^3/I^*. We find infinitely many 3-dimensional counterexamples, as well as another infinite family of flat counterexamples whose dimensions grow without bound.

1. Introduction

Suppose B_- and B_+ are closed smooth manifolds and that $DB_\pm \to B_\pm$ are disk bundles over them, possibly of different ranks. Suppose in addition that the boundaries $\partial DB_\pm$ of $DB_\pm$ are diffeomorphic, say via a diffeomorphism $f : \partial DB_- \to \partial DB_+$. Then we may form a smooth closed manifold $M = DB_- \cup_f DB_+$. A manifold diffeomorphic to one obtained from this construction is called a *double disk bundle*. For example, $\mathbb{R}P^2$ is a double disk bundle, for it is a union of a disk and a closed Möbius band. That is, $\mathbb{R}P^2$ is a union of a trivial 2-disk bundle over a point together with nontrivial 1-disk bundle over S^1.

Double disk bundles arise naturally in many diverse fields of geometry and topology. We refer the reader to the introduction of [DeVito et al. 2023] for numerous examples of this. Our main interest stems from Grove's double soul conjecture [2002].

Conjecture 1.1 (double soul conjecture). *Suppose M is a closed simply connected manifold which admits a Riemannian metric of nonnegative sectional curvature. Then M is a double disk bundle.*

MSC2020: 53C21, 53C30, 55R25.
Keywords: homogeneous spaces, double soul conjecture, disk bundles.

Evidence for this conjecture includes the fact that cohomogeneity one manifolds (and free isometric quotients by a subaction of the cohomogeneity one action [Wilking 2007]), which are one of two main building blocks for nonnegatively curved manifolds, admit such a structure [Mostert 1957; Galaz-García and Zarei 2018]. In addition, Cheeger [1973] showed that the connect sum of two compact rank one symmetric spaces (CROSS) admits a metric of nonnegative sectional curvature. As a CROSS with a small open ball removed has the structure of a disk bundle, these manifolds also verify the double soul conjecture. In addition, Cheeger and Gromoll's soul theorem [1972] gives an analogous theorem for noncompact complete Riemannian manifolds of nonnegative sectional curvature. The conjecture has also been verified for many other examples, including all known simply connected positively curved manifolds [DeVito et al. 2023, Theorem 3.3], simply connected biquotients in dimension at most 7 [González-Álvaro and Guijarro 2023], and simply connected homogeneous spaces of dimension at most 10 [González-Álvaro and Guijarro 2023]. We recall that a biquotient is the quotient of a Riemannian homogeneous space by a free isometric action, and comprise the other main building block of nonnegatively curved manifolds.

The conjecture also implies some classification results. For example, if true, then it would follow that our known list of nonnegatively curved simply connected 4 and 5-dimensional manifolds is complete [Ge and Radeschi 2015, Theorem 1.1; DeVito et al. 2023, Theorem B].

Grove [2002] noted that the natural generalization of Conjecture 1.1 to nonsimply connected manifolds is false: the Poincaré dodecahedral space S^3/I^* admits a positively curved Riemannian metric, but does not support a double disk bundle structure. However, this was previously the only known counterexample to the generalized conjecture. As such, it is natural to search for more, with various topological and geometric properties. Our main result supplies infinitely many counterexamples to the generalized conjecture, on opposite ends of the nonnegative curvature landscape.

Theorem 1.2. *There are infinitely many closed Riemannian* 3*-manifolds of positive sectional curvature which are not double disk bundles. In addition, there are infinitely many closed flat manifolds which are not double disk bundles.*

The 3-manifold family consists of infinitely many nontrivial isometric quotients of a round S^3. The homogeneous spaces S^3/I^*, S^3/O^*, S^3/T^*, where I^*, O^*, and T^* are the binary icosahedral, octahedral, and tetrahedral groups are among these examples.

In fact, S^3/I^*, S^3/O^*, and S^3/T^* are the only homogeneous spaces among our examples. It is thus natural to wonder if there are more. This leads to the obvious question:

Question 1.3. *Are there infinitely many homogeneous spaces which are not double disk bundles*?

Given that the three homogeneous examples of Theorem 1.2 are quotients of S^3, one is tempted to answer Question 1.3 by looking at homogeneous quotients of spheres of higher dimension. However, we prove that S^3/I^*, S^3/O^*, and S^3/T^* are the only homogeneous quotients of a sphere, in any dimension, which are not double disk bundles; see Proposition 4.9.

The other infinite family, the closed flat manifolds, are precisely those with trivial first homology group. The construction of such flat manifolds is rather abstract, so we have been unable to determine which dimensions these examples appear. However, we can show they exist in arbitrarily large dimensions.

We stress that all of our examples have nontrivial fundamental groups, so the double soul conjecture remains open. In fact, all of our examples have nonnilpotent fundamental groups, so the generalized double soul conjecture is still open for nilpotent manifolds.

We now give an outline of the proof of Theorem 1.2, beginning with the 3-dimensional examples. We first prove that if M^3 has a metric of positive sectional curvature and is a double disk bundle, then it must have a double disk bundle structure where the common boundary $\partial DB_- \cong \partial DB_+$ is diffeomorphic to a sphere S^2 or to a torus T^2. We then classify all disk bundles whose total space has boundary diffeomorphic to S^2 or T^2, and then consider all possible ways of gluing these together. The double disk bundle decomposition lends itself to the use of the Seifert–van Kampen theorem, so we are able to compute presentations for all the resulting fundamental groups. The end conclusion is that a positively curved M^3 admits a double disk bundle decomposition if and only if it is a lens space or a particular $\mathbb{Z}/2\mathbb{Z}$ quotient of a lens space, a so-called prism manifold. From the known classification of fundamental groups of spherical 3-manifolds [Wolf 2011, Section 7.5], we obtain infinitely many examples which are not double disk bundles. It is worth noting that the examples we find are the only 3-dimensional counterexamples to the double soul conjecture, even under the weaker assumption that M has a Riemannian metric of nonnegative sectional curvature; see Remark 4.3.

For the flat examples, enumerating all the possibilities for the common boundary $\partial DB_- \cong \partial DB_+$ is not feasible, so we proceed differently. We first show in Proposition 3.6 that for any manifold covered by a contractible manifold, any double disk bundle decomposition must have both disk bundles of rank 1. On the other hand, we also establish (Proposition 3.2) that if a manifold admits a double disk bundle structure with at least one double disk bundle has rank 1, then the manifold must have a nontrivial double cover, which in turn implies that the first homology group surjects onto $\mathbb{Z}/2\mathbb{Z}$. Thus, any flat manifold with trivial first homology group cannot be a double disk bundle. Such flat manifolds have been constructed by Igor Belegradek [2022], providing the examples.

An outline of the paper follows. In Section 2, we cover the required background and set up notation. Section 3 contains general results on the topology of double disk bundles especially in the case where at least one disk bundle has rank 1. In Section 4, we classify the nonnegatively curved 3-manifolds which are double disk bundles, finding that some positively curved examples are not double disk bundles. Finally, Section 5 contains the results concerning flat manifolds.

2. Background and notation

Suppose B_- and B_+ are closed manifolds and that $D^{\ell_\pm+1} \to DB_\pm \to B_\pm$ are disk bundles. We assume their boundaries are diffeomorphic, say by a diffeomorphism $f : \partial DB_- \to \partial DB_+$. Then we can form the closed manifold $M = DB_- \cup_f DB_+$ by gluing DB_- and DB_+ along their boundary. A manifold obtained via this construction is called a *double disk bundle*.

Restricting the projection maps to their respective boundaries, we obtain sphere bundles $S^{\ell_\pm} \to \partial DB_\pm \to B_\pm$. The numbers $\ell_\pm \geq 0$ will always refer to the dimension of these fiber spheres. We will use L to denote the diffeomorphism type of the common boundary. We will borrow language from the field of singular Riemannian foliations, and refer to L as the regular leaf and the $B_\pm$ as the singular leaves.

As was shown in [DeVito et al. 2023, Proposition 4.1], if a connected closed manifold M admits a double disk bundle decomposition, then it necessarily admits one where both $B_\pm$ are connected. Thus we can and will always assume that in any double disk bundle decomposition, both singular leaves $B_\pm$ are connected. Using the sphere bundles $S^{\ell_\pm} \to L \to B_\pm$, the condition that both $B_\pm$ are connected implies that L has at most 2 components, and that L is connected unless B_- and B_+ are diffeomorphic, $\ell_- = \ell_+ = 0$, and $L \cong S^0 \times B_- \cong S^0 \times B_+$.

The decomposition of M into two disk bundles is ideal for applying the Mayer–Vietoris sequence in cohomology, as well as the Seifert–van Kampen theorem for fundamental groups, at least when L is connected. In this context, we note that contracting the fiber disks in either $DB_\pm$ provides a deformation retract of $DB_\pm$ to $B_\pm$, and the inclusion map $L \cong \partial DB_\pm \subseteq DB_\pm$ becomes homotopic to the sphere bundle projection $L \to B_\pm$ under this deformation retract.

3. Some general structure results for double disk bundles

In this section, we will collect several needed facts regarding the relationship between the fiber sphere dimensions $\ell_\pm$ and coverings. We begin with some general structure results where at least one $\ell_\pm = 0$.

Lemma 3.1. *Suppose $S^0 \to L \to B$ is a sphere bundle with $\ell = 0$ and B a connected smooth manifold. There is a smooth free involution $\sigma : L \to L$ with L/σ diffeomorphic to B.*

Proof. Because S^0 consists of two points, the sphere bundle is nothing but a double cover. If L is disconnected, it follows that $L \cong S^0 \times B$ and the required involution σ simply interchanges the two copies of B.

On the other hand, if L is connected, the covering $L \to B$ is characterized by an index 2-subgroup of $\pi_1(B)$, which is necessarily normal. Hence, the covering is regular, so the deck group is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. Then one can take σ to be the nontrivial element of the deck group. □

Proposition 3.2. *Suppose M is a connected manifold and $M = DB_- \cup_f DB_+$ is a double disk bundle with $\ell_- = 0$. Then M admits a nontrivial double cover of the form $\overline{M} = DB_+ \cup_g DB_+$ for some diffeomorphism $g : L \to L$. That is, $\overline{M}$ has a double disk bundle decomposition where each half is a copy of DB_+.*

Proof. Because $\ell_- = 0$, Lemma 3.1 gives a free involution $\sigma : L \to L$ with quotient B_-. We now form $\overline{M}$ as the union

$$\overline{M} = (DB_+ \times \{-1\}) \cup_{\sigma\circ f} L \times [-1, 1] \cup_f (DB_+ \times \{1\}),$$

where $DB_+ \times \{-1\}$ is glued to $L \times \{-1\}$ and $DB_+ \times \{1\}$ is glued to $L \times \{1\}$. From [Kosinski 1993, Chapter VI, Section 5], the union $(DB_+ \times \{-1\}) \cup_{\sigma\circ f} L \times [-1, 1]$ is diffeomorphic to DB_+, so $\overline{M}$ is diffeomorphic to a double disk bundle with both halves a copy of DB_+.

Thus, we need only show that $\overline{M}$ is a double cover of M. To that end, we define a free involution ρ on $\overline{M}$ whose quotient is M. Given $(x, \pm 1) \in DB_+ \times \{\pm 1\}$, we define $\rho(x, \pm 1) = (x, \mp 1)$. In other words, ρ interchanges the two copies of DB_+ on the "ends" of $\overline{M}$. In addition, we define the action of ρ on $L \times [-1, 1]$ by mapping a point (y, t) to $(\sigma(y), -t)$. It is easy to verify that this is the required involution. □

If both $\ell_\pm = 0$, then applying Proposition 3.2 gives a double cover which again has both $\ell_\pm = 0$. Hence, we can iterate this procedure. This shows that a manifold can only admit a double disk bundle decomposition with both $\ell_\pm = 0$ if $\pi_1(M)$ is infinite. In fact, while it will not be needed in the remainder of the paper, it turns out that a double cover of M fibers over S^1.

Proposition 3.3. *Suppose M is a connected manifold which admits a double disk bundle structure with both $\ell_- = \ell_+ = 0$ and regular leaf L. Then $\pi_1(M)$ is infinite, and M has a double cover $\overline{M}$ which fibers over S^1 with fiber L.*

Proof. We have already proven the first statement, so we focus on the second. By assumption, we may write $M = DB_+ \cup_f DB_-$ for some diffeomorphism $f : L \to L$.

As both $\ell_\pm = 0$, Lemma 3.1 gives a pair of free involutions $\sigma_\pm : L \to L$ with $L/\sigma_\pm$ diffeomorphic to $B_\pm$. Both $\sigma_\pm$ extend to involutions on $L \times [-1, 1]$ defined by $(y, t) \mapsto (\sigma_\pm(y), -t)$. The quotient $(L \times [-1, 1])/\sigma_\pm$ is clearly diffeomorphic to $DB_\pm$.

Now, take two copies of $L \times [-1, 1]$, which we will refer to as the left copy and right copy. We glue $(y, 1)$ in the left copy to $(f(y), 1)$ in the right copy, and we glue $(y, -1)$ in the left copy to $(\sigma_+(f(\sigma_-(y))), -1)$ to form the manifold $\overline{M}$.

From [Kosinski 1993, Chapter VI, Section 5], if we only do the gluing of $(y, 1)$ to $(f(y), 1)$, the resulting manifold is diffeomorphic to $L \times [-1, 1]$. Thus, $\overline{M}$ has the structure of a mapping torus for some self diffeomorphism of L, so is a bundle over S^1 with fiber L.

It remains to see that $\overline{M}$ is a double cover of M. To that end, we define a free involution ρ on $\overline{M}$ with quotient M as follows. On the left copy of $L \times [-1, 1]$, ρ acts by $(y, t) \mapsto (\sigma_-(y), -t)$. On the right copy, ρ acts by $(y, t) \mapsto (\sigma_+(y), -t)$. Once again, it is easy to verify this has the desired properties. $\square$

Remark 3.4. In Proposition 3.3, if L is disconnected, then M itself fibers over S^1. On the other hand, if L is connected, passing to a double cover is sometimes necessary to obtain the bundle structure. For example, if $M = \mathbb{R}P^n \# \mathbb{R}P^n$ with $n \geq 3$, then M has a double disk bundle structure with both $\ell_\pm = 0$. Indeed, $\mathbb{R}P^n$ with a ball removed is a diffeomorphic to the total space of the disk bundle in the tautological bundle over $\mathbb{R}P^{n-1}$. But M does not fiber over S^1 because its fundamental group $\pi_1(M) \cong (\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/2\mathbb{Z})$ has abelianization $(\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2\mathbb{Z})$, so does not surject onto $\mathbb{Z}$.

The next proposition describes how double disk bundles act with respect to covering maps.

Proposition 3.5. *Suppose M is a connected manifold which admits a double disk bundle structure with both $\ell_\pm \geq 1$. If $\rho : M' \to M$ is any nontrivial covering (in the sense that M' is connected), then M' is a double disk bundle with regular leaf $L' := \rho^{-1}(L)$, singular leaves $B'_\pm := \rho^{-1}(B_\pm)$, and with $\ell'_\pm = \ell_\pm$. In addition, each of L', B'_+, and B'_- are connected.*

Proof. Since a covering map is a submersion, everything except the connectedness of L', $B'_\pm$ is a direct consequence of [DeVito et al. 2023, Proposition 3.1d]. Thus, we need only show the connectedness of L' and $B'_\pm$. As both $B'_\pm$ are the continuous image of the sphere bundle projections $L' \to B'_\pm$, it is sufficient to show that L' is connected.

So, we now show that L' is connected. Because ρ is a covering, so is $\rho|_{L'} : L' \to L$. In addition, since at least one $\ell_\pm \geq 1$, L must be connected. Thus, to show L' is connected, it is sufficient to select $x \in L$, and show that any pair of points in $\rho^{-1}(x)$ can be connected by a path in L'. Let $x_1, x_2 \in \rho^{-1}(x)$.

Because M' is connected, we may connect x_1 and x_2 by a path $\gamma' : [0, 1] \to M'$ in M'. Then $\gamma := \rho \circ \gamma'$ is a closed curve in M.

We claim that γ is homotopic rel endpoints to a closed curve α lying entirely in L. To see this, note that γ represents an element of $\pi_1(M, x)$, so we need to show the map $\pi_1(L, x) \to \pi_1(M, x)$ induced by the inclusion $L \to M$ is surjective.

The Seifert–van Kampen theorem applied to the double disk bundle decomposition of M shows that any curve in M is, up to homotopy rel endpoints, a finite concatenation of curves in DB_+ and DB_-. Because both $\ell_\pm \geq 1$, the long exact sequence in homotopy groups implies the maps $\pi_1(L) \to \pi_1(DB_\pm) \cong \pi_1(B_\pm)$ are surjective, so each curve in DB_+ or DB_- is homotopic rel end points to one lying entirely in L. In particular, γ is homotopic rel end points to a curve α in L.

Now, since $\rho : L' \to L$ is a covering, it is, in particular, a fibration. As γ has a lift to M', α must lift to a curve $\alpha' : [0, 1] \to M'$. Since the homotopy from γ to α fixed the end points and the fiber of ρ is discrete, α' must have the same endpoints as γ'. That is, α' is a curve connecting x_1 and x_2 with image in L'. This completes the proof that L' is connected, and thus, of the proposition. □

In the special case that M is aspherical, i.e., the universal cover of M is contractible, we can completely characterize the possibilities for the fiber sphere dimensions $\ell_\pm$ for any double disk bundle structure on it.

Proposition 3.6. *Suppose M is an aspherical manifold which admits a double disk bundle structure. Then both $\ell_- = \ell_+ = 0$. That is, both fiber spheres are 0-dimensional.*

Proof. We assume for a contradiction that M has a double disk bundle decomposition with say, $\ell_- > 0$. This implies that the regular leaf L is connected. If $\ell_+ = 0$, then Proposition 3.2 implies that M has a double cover admitting a double disk bundle structure with both $\ell_\pm > 0$. Noting that the double cover of an aspherical manifold is aspherical, we may therefore assume that both $\ell_\pm > 0$.

In this case, we consider the universal cover $\rho : M' \to M$. From Proposition 3.5, we obtain a double disk bundle structure on M' with regular leaf L' and singular leaves $B'_\pm$ connected. We will conclude the proof by showing that M' has no such double disk bundle structure. Specifically, we will show that $H^{t(\ell_+ + \ell_-)}(L'; \mathbb{Z}/2\mathbb{Z})$ is nontrivial for all $t \geq 0$, contradicting the fact that L' is a finite-dimensional manifold. Set $R = \mathbb{Z}/2\mathbb{Z}$ for legibility.

Because M' is contractible, the Mayer–Vietoris sequence for the double disk bundle decomposition of M' yields isomorphisms

$$\psi_k : H^k(B'_-; R) \oplus H^k(B'_+; R) \to H^k(L'; R)$$

for each $k \geq 1$ (and that ψ_0 is surjective). Recalling that ψ_k is nothing but the difference in the maps induced by the sphere bundle projections $L' \to B'_\pm$, it follows that each map $H^k(B'_\pm; R) \to H^k(L'; R)$ must injective. Since both $B'_\pm$ are connected, we have Gysin sequences associated to $L' \to B'_\pm$; injectivity of $H^*(B'_\pm; R) \to H^*(L'; R)$ then implies via the Gysin sequence that the R-Euler classes $e_\pm$ of both bundles $L' \to B'_\pm$ are trivial. In more detail, in the sequence

$$H^0(B'_\pm; R) \xrightarrow{\cup e_\pm} H^{\ell_\pm+1}(B'_\pm; R) \to H^{\ell_\pm+1}(L'; R),$$

the first map has image generated by $e_\pm$ while the while the second map is injective. Exactness at the middle term then forces $e_\pm = 0$. We thus have group isomorphisms

$$H^*(L'; R) \cong H^*(B'_+; R) \otimes H^*(S^{\ell_+}; R) \cong H^*(B'_-; R) \otimes H^*(S^{\ell_-}; R),$$

where the inclusions $H^*(B'_\pm; R) \to H^*(B'_\pm; R) \otimes H^*(S^{\ell_\pm}; R)$ are the obvious ones.

We will now prove that $H^{t(\ell_-+\ell_+)}(L'; R) \neq 0$ for all $t \geq 0$ by induction. The base case is clear, as it is simply the assertion that $H^0(L'; R) \neq 0$.

Now, assume that $H^{t(\ell_-+\ell_+)}(L'; R)$ is nonzero for some $t \geq 0$. Since ψ_k for $k := t(\ell_+ + \ell_-)$ is surjective, there must therefore be a nonzero element x in at least one of $H^k(B'_\pm; R)$. We assume without loss of generality that $x \in H^k(B'_+; R)$. If $y_\pm \in H^{\ell_\pm}(S^{\ell_\pm}; R) \cong R$ is the nonzero element, then the element $x \otimes y_+ \in H^{k+\ell_+}(L'; R)$ is nonzero, and not in the image of $H^{k+\ell_+}(B'_+; R)$. Since $\psi_{k+\ell_+}$ is surjective, it now follows that $H^{k+\ell_+}(B'_-; R) \neq 0$. Suppose $z \in H^{k+\ell_+}(B'_-; R)$ is such a nonzero element. Then the element $z \otimes y_- \in H^{(t+1)(\ell_-+\ell_+)}(L'; R)$ is nonzero, completing the induction. □

We will also need a proposition regarding orientability.

Proposition 3.7. *Suppose M is a double disk bundle and that M is orientable. Then so is the regular leaf L.*

Proof. Because L is the boundary of both disk bundles, L must have trivial normal bundle. Then $TM|_L = TL \oplus 1$ with 1 denoting a trivial rank 1 bundle. Computing the first Stiefel–Whitney class using the Whitney sum formula, we find

$$0 = w_1(TM|_L) = w_1(TL) + w_1(1) = w_1(TL).$$

Thus $w_1(TL) = 0$, so L is orientable. □

4. 3-dimensional examples

The goal of this section is to prove the following theorem.

Theorem 4.1. *Suppose M^3 is a closed manifold admitting a metric of positive sectional curvature. Then M is a double disk bundle if and only if M is S^3, a lens space $L(p, q)$, or a prism manifold.*

By definition, a lens space $L(p, q)$ (where $\gcd(p, q)$ is necessarily 1) is the quotient of S^3 by a free isometric action by the cyclic group $\mathbb{Z}/p\mathbb{Z} \subseteq S^1 \subseteq \mathbb{C}$ acting on $S^3 \subseteq \mathbb{C}^2$ via $\mu * (z_1, z_2) = (\mu z_1, \mu^q z_2)$. Also, by definition, a prism manifold is an isometric quotient of a round S^3 with fundamental group isomorphic to $\langle a, b \mid aba^{-1}b = 1, a^{2\beta} = b^{\alpha}\rangle$, where $\gcd(\alpha, \beta) = 1$. Prism manifolds include the homogeneous spaces S^3/D^*_{4n}, where D^*_{4n} is the order $4n$ group generated by $e^{2\pi i/n}$ and j in the group $Sp(1)$ of unit length quaternions.

From, e.g., [McCullough 2002, Table 1], the homogeneous 3-manifolds which are covered by S^3 consists of precisely the lens space $L(p, 1)$, the prism manifolds S^3/D^*_{4n}, and the spaces S^3/T^*, S^3/O^*, or S^3/I^*, where T^*, O^*, and I^* are the binary tetrahedral, octahedral, and icosahedral groups respectively. In addition, from e.g., [Wolf 2011, Section 7.5], the product of any of these fundamental groups with a cyclic group of relatively prime order is again the fundamental group of a positively curved 3-manifold. Thus, Theorem 4.1has the following corollary.

Corollary 4.2. *There are infinitely many positively curved* 3*-manifolds which do not admit a double disk bundle structure. These examples include precisely three homogeneous examples*: S^3/T^*, S^3/O^*, *and* S^3/I^*, *where* T^*, O^*, *and* I^* *are the binary tetrahedral*, *octahedral*, *and icosahedral groups respectively.*

Remark 4.3. By using work of others, it is easy to extend Theorem 4.1 to nonnegatively curved 3-manifolds. Hamilton [1982, main theorem; 1986, Theorem 1.2] showed a closed 3-manifold M admitting a metric of nonnegative sectional curvature is covered by S^3, $S^2 \times S^1$, or T^3. If M is covered by $S^2 \times S^1$, then M is diffeomorphic to $S^2 \times S^1$, $\mathbb{R}P^2 \times S^1$, $\mathbb{R}P^3 \# \mathbb{R}P^3$, or to the unique nontrivial S^2 bundle over S^1 [Tollefson 1974]. Clearly for each of these possibilities, M is a double disk bundle. If M is covered by T^3, then from [Scott 1983, p. 448], M is a double disk bundle.

We now work towards proving Theorem 4.1. For the remainder of this section, M denotes a 3-manifold of positive sectional curvature. From [Hamilton 1982, main theorem], M is finitely covered by S^3, so has finite fundamental group. A simple application of the Lefschetz fixed point theorem implies that M must be orientable. From Proposition 3.3, at least one of $\ell_\pm > 0$, which, in particular, implies that L is connected.

Proposition 4.4. *Suppose M is a closed orientable* 3*-manifold which admits a double disk bundle decomposition with at least one fiber sphere of positive dimension. The regular leaf L must be diffeomorphic to either S^2 or T^2.*

Proof. Assume without loss of generality that $\ell_+ > 0$. This implies that L is connected. Since L is 2-dimensional and an S^{ℓ_+}-bundle over B_+, we must have $\ell_+ \in \{1, 2\}$. If $\ell_+ = 2$, the fiber inclusion map $S^2 \to L$ is an embedding between closed manifolds of the same dimension, hence a diffeomorphism. If $\ell_+ = 1$, then the Euler characteristic $\chi(L) = \chi(S^1)\chi(B_+) = 0$, so L must be T^2 or a Klein bottle. But L must be orientable from Proposition 3.7. $\square$

We will proceed by breaking into cases depending on whether $L = S^2$ or $L = T^2$. We will classify all disk bundles whose boundary is diffeomorphic to L, and then classify ways of gluing the corresponding disk bundles. Using a collar neighborhood, it easy to see that if two gluing maps are isotopic, then the corresponding double disk bundles are diffeomorphic. The following lemma provides another circumstance where the double disk bundles are diffeomorphic.

Lemma 4.5. *Suppose X and Y are manifolds with boundary and $f : \partial X \to \partial Y$ is a diffeomorphism. Assume in addition that $G : X \to X$ is a diffeomorphism with $g := G|_{\partial X} : \partial X \to \partial X$. Then the manifolds $X \cup_f Y$ and $X \cup_{f\circ g} Y$ are diffeomorphic.*

Proof. We define a diffeomorphism $\phi : X \cup_{f\circ g} Y \to X \cup_f Y$ by mapping $x \in X$ to $\phi(x) = G(x)$ and mapping $y \in Y$ to $\phi(y) = y$. It is obvious that ϕ is a diffeomorphism, if it is well defined.

We now check that it is well-defined. If we first identify $x \in \partial X$ with $f(g(x))$ and then apply ϕ, we obtain the point $f(g(x))$. On the other hand, if we first apply ϕ and then identify with ∂Y, we get $\phi(x) = G(x) = g(x) \sim f(g(x))$. $\square$

Proposition 4.6. *Suppose M is a double disk bundle with regular leaf $L = S^2$. Then, M is diffeomorphic to S^3, $\mathbb{R}P^3$, or $\mathbb{R}P^3\#\mathbb{R}P^3$.*

Proof. To begin with, note there are precisely two isomorphism types of sphere bundles with total space S^2: they are $S^2 \to S^2 \to \{p\}$, and $S^0 \to S^2 \to \mathbb{R}P^2$. Since a diffeomorphism of either S^0 or S^1 extends to a diffeomorphism of the corresponding disk, both of these extend uniquely to disk bundles. Moreover, $\mathrm{Diff}(S^2)$ deformation retracts to $O(2)$ [Smale 1959], so we may assume our gluing map is either the identity or the antipodal map. Both options extend to a diffeomorphism of the 3-ball B^3, so by Lemma 4.5 the choice of gluing map is irrelevant if either $B_\pm = \{p\}$.

If we have $B_+ = B_- = \{p\}$, then M is obtained by gluing two 3-balls along their boundary S^2, so M is diffeomorphic to S^3 in this case. If we have $B_+ = \{p\}$ and $B_- = \mathbb{R}P^2$, then gluing gives $\mathbb{R}P^3$. Finally, if we have $B_+ = B_- = \mathbb{R}P^2$, we obtain $\mathbb{R}P^3\#\pm\mathbb{R}P^3$. But $\mathbb{R}P^3$ admits an orientation reversing diffeomorphism, so $\mathbb{R}P^3\#-\mathbb{R}P^3$ is diffeomorphic to $\mathbb{R}P^3\sharp\mathbb{R}P^3$. $\square$

We now classify all double disk bundles with regular leaf $L = T^2$ and with at least one $\ell_\pm > 0$, which completes the proof of Theorem 4.1.

Proposition 4.7. *Suppose M admits a double disk bundle structure with regular leaf $L = T^2$ and with $\ell_+ > 0$. Then either $\pi_1(M)$ is abelian, or M is a prism manifold.*

Remark 4.8. The classification of 3-manifolds with $\pi_1(M)$ abelian is well known [Aschenbrenner et al. 2015, Section 1.7, Table 2]. The only such examples which are covered by S^3 are the lens spaces $L(p, q)$. Each of these is well-known to be a double disk bundle, e.g., they are all quotients of S^3 via a subaction of the well-known cohomogeneity one action of T^2 on S^3. The examples which are not covered by S^3 are covered by $S^2 \times S^1$, so are all double disk bundles by Remark 4.3.

Proof. The assumption that $\ell_+ > 0$ implies that $\ell_+ = 1$, so $B_+ = S^1$. An S^1-bundle over S^1 is determined by an element of $\pi_0(\mathrm{Diff}(S^1))$. Since $\mathrm{Diff}(S^1)$ deformation retracts to $O(2)$, there are precisely two S^1-bundles over S^1. Of course, one has total space K, the Klein bottle. Thus, there is a unique S^1 bundle over S^1 with total space T^2, the trivial bundle.

If $\ell_- = 2$, the fiber inclusion $S^2 \to T^2$ must be an embedding, giving an obvious contradiction. Hence, $\ell_- \in \{0, 1\}$. Of course, if $\ell_- = 1$, then the bundle $L \to B_-$ must be the trivial bundle as in the previous paragraph. On the other hand, if $\ell_- = 0$, then $L \to B_-$ is a 2-fold covering, so B_- is diffeomorphic to either T^2 or K.

Each of these S^1-bundles extends to a disk bundle in a unique way. In addition, $\mathrm{Diff}(T^2)$ deformation retracts to $Gl_2(\mathbb{Z})$ [Farb and Margalit 2012, Theorem 2.5], so we can always assume our gluing map lies in $Gl_2(\mathbb{Z})$. Moreover, the diffeomorphism $\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ of $T^2 = \partial(D^2 \times S^1)$ extends to a diffeomorphism of $DB_+ \cong D^2 \times S^1$, so Lemma 4.5 implies that we may assume our gluing map lies in $Gl_2^+(\mathbb{Z})$.

Applying the Seifert–van Kampen theorem to the double disk decomposition of M, we note that since $\ell_+ = 1$, the map $\pi_1(L) \to \pi_1(B_+)$ is surjective. This implies that $\pi_1(M)$ is isomorphic to a quotient of $\pi_1(DB_-) = \pi_1(B_-)$. Thus, if $B_- \neq K$, then $\pi_1(M)$ is necessarily abelian.

So, we assume $B_- = K$, and that the gluing map is determined by a matrix

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in Gl_2^+(\mathbb{Z}).$$

We have presentations

$$\pi_1(S^1) = \langle a \rangle, \pi_1(T^2) \cong \langle b, c \mid [b, c] = 1 \rangle, \text{ and } \pi_1(K) = \langle d, e \mid ded^{-1}e = 1 \rangle.$$

The unique abelian index 2 subgroup of $\pi_1(K)$, which is isomorphic to $\pi_1(T^2)$, is generated by $\{d^2, e\}$. Therefore, by picking a new generating set for $\pi_1(T^2)$ if necessary, we may assume the map $\pi_1(T^2) \to \pi_1(K)$ maps b to d^2 and c to e.

We claim that, in addition, we may choose the projection map $T^2 \to S^1$ to map b to a and c to the identity. We begin with the standard projection onto the second factor $p_2 : T^2 = S^1 \times S^1 \to S^1$. This maps $(1, 0) \in \pi_1(T^2) \cong \mathbb{Z}^2$ to the identity and $(0, 1)$ to a. Since $\{b, c\}$ generates $\pi_1(T^2)$, there is an element of $f \in Gl_2(\mathbb{Z}) \simeq \mathrm{Diff}(T^2)$ which maps c to $(1, 0)$ and b to $(0, 1)$. Then the composition $p_2 \circ f : T^2 \to S^1$ is a (trivial) fiber bundle with fiber S^1 which maps c to the identity and b to a.

Note that under the gluing map $\left[\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right]$, the map

$$\pi_1(T^2) \xrightarrow{\left[\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right]} \pi_1(T^2) \to \pi_1(S^1)$$

is therefore given by $b \mapsto b^\alpha c^\gamma \mapsto a^\alpha$, and $c \mapsto b^\beta c^\delta \mapsto a^\beta$, where we have used multiplicative notation rather than additive for both $\pi_1(T^2) \cong \mathbb{Z}^2$ and $\pi_1(S^1) \cong \mathbb{Z}$. Thus, the Seifert–van Kampen theorem gives

$$\pi_1(M) \cong \langle a, d, e \mid ded^{-1}e = 1, a^\alpha = d^2, a^\beta = e\rangle.$$

We claim that this is isomorphic to

$$\langle d, e \mid ded^{-1}e = 1, d^{2\beta} = e^\alpha\rangle,$$

so that M has the fundamental group of a prism manifold.

To that end, we first note that the generator a in the first presentation is unnecessary. Indeed, we have $\alpha\delta - \beta\gamma = 1$, so

$$a^1 = a^{\alpha\delta-\beta\gamma} = (a^\alpha)^\delta (a^\beta)^{-\gamma} = d^{2\delta} e^{-\gamma}.$$

Thus, we need only demonstrate that the relations in the first presentation are consequences of the relations in the second, and vice versa.

So, assume initially that both $a^\alpha = d^2$ and $a^\beta = e$. Raising the first relation to the power of β, and the second to the power of α, we obtain

$$d^{2\beta} = a^{\alpha\beta} = e^\alpha,$$

so the relations in the first presentation imply those in the second. Conversely, assuming $d^{2\beta} = e^\alpha$, noting that d^2 commutes with everything, and setting $a = d^{2\delta}e^{-\gamma}$, we find

$$\begin{aligned} a^\alpha &= d^{2\alpha\delta} e^{-\gamma\alpha} \\ &= d^{2(1+\beta\gamma)} e^{-\gamma\alpha} \\ &= d^2 (d^{2\beta})^\gamma (e^\alpha)^{-\gamma} \\ &= d^2 (e^\alpha)^\gamma (e^\alpha)^{-\gamma} \\ &= d^2 \end{aligned}$$

and likewise, we find that $a^\beta = e$.

Thus, $\pi_1(M)$ is isomorphic to the fundamental group of a prism manifold, as defined above. Since such manifolds are classified up to diffeomorphism by their fundamental group [Aschenbrenner et al. 2015, Theorem 2.2], M must be a prism manifold in these cases. □

We conclude this section by proving that the three homogeneous examples S^3/T^*, S^3/O^*, and S^3/I^* of Corollary 4.2 are the only homogeneous examples in any dimension which are covered by a sphere but are not double disk bundles.

Proposition 4.9. *Suppose M is a closed homogeneous space which is covered by a sphere. Then M admits a double disk bundle decomposition, except when M is diffeomorphic to one of S^3/T^*, S^3/O^*, or S^3/I^*.*

Proof. From [Wilking and Ziller 2018, Table 2], we see that the homogeneous spaces nontrivially covered by a sphere are

(a) real projective spaces,

(b) homogeneous lens spaces, or

(c) quotients of $S^{4n-1} \subseteq \mathbb{H}^n$ by a nonabelian finite subgroup of $Sp(1)$ acting diagonally.

Here, a homogeneous lens space is a quotient $S^{2n+1}/(\mathbb{Z}/m\mathbb{Z})$, where $\mathbb{Z}/m\mathbb{Z} = \{(z, z, \dots, z) \in \mathbb{C}^{n+1} : z^m = 1\}$, and $\mathbb{H}$ denotes the skew-field of quaternions.

We have a uniform description of these actions: let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and set $k = \dim_{\mathbb{R}}(\mathbb{K})$. Let G denote any finite subgroup of $O(1)$, $U(1)$ or $Sp(1)$ respectively. Then G acts freely on $S^{kn-1} \subseteq \mathbb{K}^n$ via the diagonal action in each coordinate and the cases (a), (b), and (c) above correspond to the choice of $\mathbb{K}$.

We first claim that if $n \geq 2$ then all such quotients S^{nk-1}/G admit a double disk bundle decomposition. Indeed, one can simply observe that the block action by $O(n-1) \times O(1)$, $U(n-1) \times U(1)$, or $Sp(n-1) \times Sp(1)$ on $S^{nk-1} \subseteq \mathbb{K}^n = \mathbb{K}^{n-1} \oplus \mathbb{K}$ is cohomogeneity one, and G acts via a subaction of the block action.

This leaves the case $n = 1$, which gives the manifolds S^0/G, S^1/G, or S^3/G. Of course, the first is 0-dimensional, and any quotient S^1/G is diffeomorphic to S^1, and thus admits a double disk bundle decomposition. The final case S^3/G is given by Corollary 4.2. □

5. Flat examples

The goal of this section is to prove the following theorem.

Theorem 5.1. *There are infinitely many closed flat manifolds, in arbitrarily large dimension, which are not double disk bundles.*

We begin with a proposition which allows us to recognize when a flat manifold does not admit a double disk bundle decomposition.

Proposition 5.2. *Suppose M is a closed flat manifold with $H_1(M)$ finite of odd order. Then M cannot admit a double disk bundle decomposition.*

Proof. Assume for a contradiction that M admits a double disk bundle decomposition. Since M is flat, the Cartan–Hadamard theorem implies that M is aspherical. Thus, Proposition 3.6 applies: any double disk bundle decomposition on M must have both $\ell_\pm = 0$. Then, from Proposition 3.2, M admits a nontrivial double cover. In particular, $\pi_1(M)$ must have an index 2 subgroup, so admits a surjection to $\mathbb{Z}/2\mathbb{Z}$. Since $H_1(M)$ is the abelianization of $\pi_1(M)$, this surjection must factor through $H_1(M)$. But no finite group of odd order admits a surjection to $\mathbb{Z}/2\mathbb{Z}$, giving a contradiction. □

In order to prove Theorem 5.1, we need only establish the existence of infinitely many flat manifolds M in arbitrarily large dimensions with first homology group $H_1(M)$ finite of odd order. In fact, we will find examples with $H_1(M)$ trivial. As $H_1(M)$ is the abelianization of $\pi_1(M)$, we are thus tasked with finding an infinite family of flat manifolds for which $\pi_1(M) = [\pi_1(M), \pi_1(M)]$ is perfect. These examples are furnished by the following theorem.

Theorem 5.3. *Suppose ϕ is any finite perfect group. Then there is a closed flat manifold M_ϕ for which $H_1(M_\phi) = 0$ and for which M_ϕ has holonomy ϕ.*

Recall that the alternating group on n letters, A_n is perfect if $n \geq 4$. We claim that for $n \geq 7$, that $\dim M_{A_n} \geq n-1$, so Theorem 5.1 immediately follows from Proposition 5.2 and Theorem 5.3. Indeed, the holonomy group of an n-manifold is a subgroup of the orthogonal group $O(n)$, and for $n \geq 7$, the smallest nontrivial representation of A_n occurs in dimension $n-1$ [Fulton and Harris 1991, Problem 5.5].

Thus, to prove Theorem 5.1, we need only to prove Theorem 5.3. We do this using an argument due to Igor Belegradek [2022].

We will use the following characterization of the fundamental group of a closed flat manifold.

Theorem 5.4 [Bieberbach 1911; Auslander and Kuranishi 1957]. *An abstract group π is the fundamental group of a closed flat n-manifold if and only if both of the following conditions are satisfied*:

(1) *π is torsion free.*

(2) *π fits into a short exact sequence of the form $0 \to \mathbb{Z}^n \to \pi \to \phi \to 0$, where ϕ is a finite group.*

The finite group ϕ is called the holonomy of π as it is isomorphic to the holonomy group of the flat manifold n-manifold with fundamental group π.

We need a lemma, which is [Holt and Plesken 1989, Proposition 2.3.13].

Lemma 5.5. *Suppose a group π fits into a short exact sequence of the form*

$$0 \to \mathbb{Z}^n \to \pi \to \phi \to 0,$$

where ϕ is a finite group. Then the commutator subgroup $\pi' = [\pi, \pi]$ also fits into a short exact sequence of the form

$$0 \to \mathbb{Z}^m \to \pi' \to \phi' = [\phi, \phi] \to 0.$$

In addition, if ϕ is perfect, then so is π'.

We may now prove Theorem 5.3.

Proof of Theorem 5.3. Let ϕ denote any finite perfect group. From [Auslander and Kuranishi 1957, Theorem 3] there is an abstract group π satisfying both conditions of Theorem 5.4. The commutator $\pi' = [\pi, \pi]$ is a subgroup of the torsion free group π, so is torsion free. From Lemma 5.5, π' is also perfect, and satisfies the second condition of Theorem 5.4 with finite quotient $\phi' = [\phi, \phi] = \phi$. Hence, by Theorem 5.4, there is a flat manifold M_ϕ with fundamental group π'. Since π' is perfect, $H_1(M_\phi) = 0$. □

Acknowledgment

The research is partially supported by NSF DMS-2105556. We are grateful for the support. We would also like to thank Martin Kerin and Karsten Grove for numerous comments on an earlier version of this article.

References

[Aschenbrenner et al. 2015] M. Aschenbrenner, S. Friedl, and H. Wilton, *3-manifold groups*, European Mathematical Society, Zürich, 2015. MR Zbl

[Auslander and Kuranishi 1957] L. Auslander and M. Kuranishi, "On the holonomy group of locally Euclidean spaces", *Ann. of Math.* (2) **65** (1957), 411–415. MR

[Belegradek 2022] I. Belegradek, "Is there a flat manifold with trivial first homology?", MathOverflow thread, 2022, available at https://mathoverflow.net/q/435038.

[Bieberbach 1911] L. Bieberbach, "Über die Bewegungsgruppen der Euklidischen Räume", *Math. Ann.* **70**:3 (1911), 297–336. MR Zbl

[Cheeger 1973] J. Cheeger, "Some examples of manifolds of nonnegative curvature", *J. Differential Geometry* **8** (1973), 623–628. MR Zbl

[Cheeger and Gromoll 1972] J. Cheeger and D. Gromoll, "On the structure of complete manifolds of nonnegative curvature", *Ann. of Math.* (2) **96** (1972), 413–443. MR Zbl

[DeVito et al. 2023] J. DeVito, F. Galaz-García, and M. Kerin, "Manifolds that admit a double disk-bundle decomposition", *Indiana Univ. Math. J.* **72**:4 (2023), 1503–1551. MR

[Farb and Margalit 2012] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series **49**, Princeton University Press, 2012. MR Zbl

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics **129**, Springer, 1991. MR Zbl

[Galaz-García and Zarei 2018] F. Galaz-García and M. Zarei, “Cohomogeneity one topological manifolds revisited”, *Math. Z.* **288**:3-4 (2018), 829–853. MR Zbl

[Ge and Radeschi 2015] J. Ge and M. Radeschi, “Differentiable classification of 4-manifolds with singular Riemannian foliations”, *Math. Ann.* **363**:1-2 (2015), 525–548. MR Zbl

[González-Álvaro and Guijarro 2023] D. González-Álvaro and L. Guijarro, “On the double soul conjecture”, Chapter 8, pp. 227–244 in *New trends in geometric analysis*, edited by A. Alarcón et al., Springer, 2023.

[Grove 2002] K. Grove, “Geometry of, and via, symmetries”, pp. 31–53 in *Conformal*, *Riemannian and Lagrangian geometry* (Knoxville, TN, 2000), edited by A. Freire, Univ. Lecture Ser. **27**, Amer. Math. Soc., Providence, RI, 2002. MR

[Hamilton 1982] R. S. Hamilton, “Three-manifolds with positive Ricci curvature”, *J. Differential Geometry* **17**:2 (1982), 255–306. MR Zbl

[Hamilton 1986] R. S. Hamilton, “Four-manifolds with positive curvature operator”, *J. Differential Geom.* **24**:2 (1986), 153–179. MR Zbl

[Holt and Plesken 1989] D. F. Holt and W. Plesken, *Perfect groups*, Oxford University Press, 1989. MR Zbl

[Kosinski 1993] A. A. Kosinski, *Differential manifolds*, Pure and Applied Mathematics **138**, Academic Press, Boston, 1993. MR Zbl

[McCullough 2002] D. McCullough, “Isometries of elliptic 3-manifolds”, *J. London Math. Soc.* (2) **65**:1 (2002), 167–182. MR Zbl

[Mostert 1957] P. S. Mostert, “On a compact Lie group acting on a manifold”, *Ann. of Math.* (2) **65** (1957), 447–455. MR Zbl

[Scott 1983] P. Scott, “The geometries of 3-manifolds”, *Bull. London Math. Soc.* **15**:5 (1983), 401–487. MR Zbl

[Smale 1959] S. Smale, “Diffeomorphisms of the 2-sphere”, *Proc. Amer. Math. Soc.* **10** (1959), 621–626. MR Zbl

[Tollefson 1974] J. L. Tollefson, “The compact 3-manifolds covered by $S^2 \times R^1$”, *Proc. Amer. Math. Soc.* **45** (1974), 461–462. MR

[Wilking 2007] B. Wilking, “Nonnegatively and positively curved manifolds”, pp. 25–62 in *Surveys in differential geometry*, edited by J. Cheeger and K. Grove, Surv. Differ. Geom. **11**, International Press, Somerville, MA, 2007. MR Zbl

[Wilking and Ziller 2018] B. Wilking and W. Ziller, “Revisiting homogeneous spaces with positive curvature”, *J. Reine Angew. Math.* **738** (2018), 313–328. MR Zbl

[Wolf 2011] J. A. Wolf, *Spaces of constant curvature*, 6th ed., AMS Chelsea, Providence, RI, 2011. MR Zbl

Received February 2, 2023. Revised August 28, 2023.

JASON DEVITO
UNIVERSITY OF TENNESSEE AT MARTIN
MARTIN, TN
UNITED STATES

jdevito1@ut.utm.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.255

THE FUNDAMENTAL GROUP OF AN EXTENSION IN A TANNAKIAN CATEGORY AND THE UNIPOTENT RADICAL OF THE MUMFORD–TATE GROUP OF AN OPEN CURVE

PAYMAN ESKANDARI AND V. KUMAR MURTY

In the first part, we give a self-contained account of Tannakian fundamental groups of extensions, generalizing a result of Hardouin (2008; 2011). In the second part, we use Hardouin's characterization of Tannakian groups of extensions to give a characterization of the unipotent radical of the Mumford–Tate group of an open complex curve. Consequently, we prove a formula that relates the dimension of the unipotent radical of the Mumford–Tate group of an open complex curve $X \setminus S$ with X smooth and projective and S a finite set of points to the rank of the subgroup of the Jacobian of X supported on S.

1. Introduction

Let X be a smooth complex projective curve and $S \subset X(\mathbb{C})$ a finite nonempty set of points. There is an exact sequence

$$0 \to H^1(X) \to H^1(X \setminus S) \xrightarrow{\text{residue}} \mathbb{Q}(-1)^{|S|-1} \to 0 \tag{1}$$

of (rational) mixed Hodge structures, where the first arrow is induced by the inclusion $X \setminus S \subset X$. In connection to a new proof of the Manin–Drinfeld theorem for modular curves, Deligne proved in the 1970s that this sequence splits (or equivalently, $H^1(X \setminus S)$ is semisimple) if and only if the rank of the subgroup of the Jacobian of X supported on S is zero (see [7, Section 10.3] and [8, Remarque 7.5], and also [11] for another argument).

To any mixed Hodge structure H, one associates an algebraic group called the Mumford–Tate group of H, which we denote by $\mathcal{MT}(H)$. This group can be defined in at least two equivalent ways: In the original definition, due to Mumford (and then refined by Serre) in the pure case, $\mathcal{MT}(H)$ is the subgroup of $\mathrm{GL}(H_{\mathbb{Q}})$ (where as usual, $H_{\mathbb{Q}}$ denotes the underlying rational vector space of H) which fixes all Hodge classes of weight zero in finite direct sums of objects of the form

$$H^{\otimes m} \otimes (H^{\vee})^{\otimes n} \quad (m, n \in \mathbb{Z}_{\geq 0}).$$

MSC2020: primary 11G30, 14C30; secondary 18D30.
Keywords: Tannakian categories, Mumford–Tate groups, mixed Hodge structures.

The second definition, which is somewhat more natural and more conceptual, is in terms of Tannakian formalism: $\mathcal{MT}(H)$ is the fundamental group of the Tannakian subcategory $\langle H\rangle$ of the category of mixed Hodge structures generated by H (see Section 2 for a brief reminder on Tannakian fundamental groups; see [1] for the equivalence of the two definitions). This means that one has a canonical equivalence of categories between $\langle H\rangle$ and the category of finite-dimensional representations of $\mathcal{MT}(H)$.

The unipotent radical of $\mathcal{MT}(H)$ measures how far H is from being semisimple. In particular, H is semisimple if and only if the unipotent radical of $\mathcal{MT}(H)$ is trivial. Thus Deligne's result about $H^1(X\setminus S)$ can be paraphrased as follows: the unipotent radical of $\mathcal{MT}(H^1(X\setminus S))$ is trivial if and only if the rank of the subgroup of the Jacobian of X supported on S is zero.

The unipotent radical of the Mumford–Tate group of a 1-motive (of which the Mumford–Tate group of $H^1(X\setminus S)$ is an example) has been studied in great generality by Bertolin [3; 4] and Jossen [18]. On his path to prove the main theorem of [18], Jossen gives a characterization of this unipotent radical in Theorem 6.2 of the same article.

In the case of $H^1(X\setminus S)$, Jossen's characterization is the following: Suppose $S=\{p_0,\dots,p_n\}$. Let P be the identity connected component of the Zariski closure of the subgroup generated by

$$(p_1-p_0,\dots,p_n-p_0)$$

in $\mathrm{Jac}(X)^n$, where $\mathrm{Jac}(X)$ is the Jacobian of X. Then P itself is an abelian subvariety of $\mathrm{Jac}(X)^n$. Jossen's theorem asserts that the Lie algebra of the unipotent radical of $\mathcal{MT}(H^1(X\setminus S))$ is canonically isomorphic to $H_1(P)$. In particular, the dimension of the unipotent radical of $\mathcal{MT}(H^1(X\setminus S))$ is twice the dimension of P.

To get a more concrete description (one that does not involve the Zariski closure) of the dimension of the unipotent radical of $\mathcal{MT}(H^1(X\setminus S))$, one can note that linear relations between the points $p_1-p_0,\dots,p_n-p_0$ with coefficients in the endomorphism algebra of $\mathrm{Jac}(X)$ cut down the dimension of P.

One of the main results of this paper gives a more explicit description of the unipotent radical of $\mathcal{MT}(H^1(X\setminus S))$ that avoids the Zariski closure (see Theorem 4.9.1). As a consequence, in the case where $\mathrm{Jac}(X)$ is simple, we get the following clean formula for the dimension of the unipotent radical (see Theorem 4.9.2(b)):

Theorem A. *Let X, S, and $\mathrm{Jac}(X)$ be as above. Let g, E, and $\mathcal{U}(H^1(X\setminus S))$ be respectively the genus of X, the endomorphism algebra $\mathrm{End}(\mathrm{Jac}(X))\otimes\mathbb{Q}$ of $\mathrm{Jac}(X)$, and the unipotent radical of the Mumford–Tate group of $H^1(X\setminus S)$. Suppose that $\mathrm{Jac}(X)$ is simple. Then the dimension of $\mathcal{U}(H^1(X\setminus S))$ is equal to $2g$ times the E-rank of the E-submodule of $\mathrm{Jac}(X)(\mathbb{C})\otimes\mathbb{Q}$ generated by the subgroup supported on S.*

In the general case where $\mathrm{Jac}(X)$ is not necessarily simple, for every division algebra in $\mathrm{End}(\mathrm{Jac}(X))\otimes\mathbb{Q}$ we get an upper bound for the dimension of $\mathcal{U}(H^1(X\setminus S))$ (see Theorem 4.9.2(a)).

Let us put this discussion on hold for the moment and go to the abstract setting of (neutral) Tannakian categories. Let $\mathbf{T}$ be a Tannakian category over a field K of characteristic zero, and ω a fiber functor over K (the example relevant to the earlier discussion being the category of mixed Hodge structures and the forgetful functor $H \mapsto H_{\mathbb{Q}}$). Suppose we have an extension

$$0 \to L \to M \to N \to 0 \tag{2}$$

in $\mathbf{T}$. Denoting the Tannakian fundamental groups of objects with respect to ω by $\mathcal{G}(-)$, we have a natural surjection

$$\mathcal{G}(M) \to \mathcal{G}(L \oplus N).$$

Let $\mathcal{U}(M)$ be the kernel of this map (if N and L are semisimple, then $\mathcal{U}(M)$ is the unipotent radical of $\mathcal{G}(M)$). By Tannakian formalism, there is an object

$$\underline{\mathrm{Lie}}(\mathcal{U}(M)) \subset \underline{\mathrm{Hom}}(N, L),$$

whose image under ω is the Lie algebra of $\mathcal{U}(M)$. The question of characterization of $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ has been studied and answered earlier by Hardouin and Bertrand in the case where $N = \mathbb{1}$ and L is semisimple: A theorem of Hardouin [15, Theorem 2] (see also [14]) asserts that in this case, $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is the smallest subobject of

$$\underline{\mathrm{Hom}}(\mathbb{1}, L) \cong L$$

such that the pushforward of (2) along the quotient map

$$L \to L/\underline{\mathrm{Lie}}(\mathcal{U}(M))$$

splits. The result was earlier proved by Bertrand [5, Theorem 1.1] in the setting of D-modules.

The case of arbitrary semisimple N (with L continued to be semisimple as well) can be deduced from Hardouin's result. In this case, the characterization becomes as follows: If ν is the extension of $\mathbb{1}$ by $\underline{\mathrm{Hom}}(N, L)$ corresponding to (2) under the canonical isomorphism

$$\mathrm{Ext}(N, L) \cong \mathrm{Ext}(\mathbb{1}, \underline{\mathrm{Hom}}(N, L))$$

(where Ext means the Yoneda Ext^1 group in $\mathbf{T}$), then $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is the smallest subobject of $\underline{\mathrm{Hom}}(N, L)$ such that the pushforward of ν under the quotient map

$$\underline{\mathrm{Hom}}(N, L) \to \underline{\mathrm{Hom}}(N, L)/\underline{\mathrm{Lie}}(\mathcal{U}(M))$$

splits.

The goal of this paper is twofold. Our first goal, to which the first part of the paper is devoted, is to give a self-contained and general treatment of Tannakian groups of extensions in characteristic zero. More precisely, in the general setting of the extension (2) in a Tannakian category, in Theorem 3.3.1 we give a characterization of $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ as a subobject of $\underline{\mathrm{Hom}}(N, L)$, without assuming that N or L is semisimple. In the semisimple case, the result simplifies to Hardouin's characterization (see Corollary 3.4.1). We also discuss a dual variant of the characterization of $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ (Theorem 3.5.1 and in the semisimple case, Corollary 3.5.2), which is more convenient in some settings.

We should point out that the generalization to the nonsemisimple situation is indeed useful in practical applications: extensions as in (2) with nonsemisimple L and N arise naturally, for example, in a nonsemisimple Tannakian category with a weight filtration, e.g., the category of mixed motives. In fact, in [12] we build on Theorem 3.3.1 to refine a result of Deligne from [18, Appendix] on unipotent radicals of Tannakian fundamental groups in a Tannakian category with a weight filtration, and then give applications to mixed motives which have "large" unipotent radicals of motivic Galois groups (see the aforementioned paper for more details).

The second goal of the paper, to which the second part of the paper is devoted, is to apply the method of the first part to study the unipotent radical of the Mumford–Tate group of an open curve. Here we take $\mathbf{T}$ to be the category of mixed Hodge structures and apply results about Tannakian groups of extensions to the extension (1). This approach leads to a characterization of the unipotent radical of the Mumford–Tate group of an open curve (see Theorem 4.9.1). The dimension formula and upper bounds mentioned above follow from this characterization of the unipotent radical (see Theorem 4.9.2).

The proof of Theorem 4.9.1 has two ingredients: The first ingredient is the semisimple case of Theorem 3.3.1 due to Hardouin (or more precisely, its dual variant given in Corollary 3.5.2). This gives a characterization of $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X\setminus S))\big)$ as follows: if μ is the element of

$$\mathrm{Ext}(H^1(X)^{|S|-1}, \mathbb{1})$$

corresponding to (1) under the canonical isomorphisms

$$\begin{aligned}\mathrm{Ext}(\mathbb{Q}(-1)^{|S|-1}, H^1(X))& \\ &\cong \mathrm{Ext}(H_1(X)\otimes\mathbb{Q}(-1)^{|S|-1}, \mathbb{1}) \overset{\text{Poincaré duality}}{\cong} \mathrm{Ext}(H^1(X)^{|S|-1}, \mathbb{1}),\end{aligned}$$

then the orthogonal complement (see Section 3.5) of $\underline{\mathrm{Lie}}(\mathcal{U}(H^1(X\setminus S))$ is the largest subobject of $H^1(X)^{|S|-1}$ on which μ restricts to a split extension. The second ingredient of the argument is now the calculation of the restrictions of the extension μ along different maps $H^1(X)\to H^1(X)^{|S|-1}$.

Theorem 4.9.1 can be deduced alternatively from Jossen's general characterization of the unipotent radical of the Mumford–Tate group of an arbitrary 1-motive given in [18, Theorem 6.2]. Although Theorem 4.9.1 is weaker than Jossen's [18, Theorem 6.2], we hope that the reader might find some value in the simplicity of our approach and exposition, which solely rely on the general material on Tannakian groups and the calculation of the relevant extensions in the category of mixed Hodge structures. This approach can be applied to any situation where the relevant extensions can be calculated and described nicely. It is also hopefully more accessible to some audiences.

The paper is organized as follows. In the next section, we recall some basic generalities about Tannakian categories. In Section 3 we prove the characterizations of $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ in a general Tannakian category and for general L and N (with notation as above). A reader not familiar with the language of Tannakian categories but familiar with properties of the category of mixed Hodge structures may assume in Sections 2 and 3 that $\mathbf{T}$ is the latter category and ω is the forgetful functor. In Section 4, we come back to the problem of studying the unipotent radical of the Mumford–Tate group of an open curve, and prove Theorems 4.9.1 and 4.9.2.

2. Preliminaries

In this section we briefly recall a few facts and constructions about Tannakian categories. For any commutative ring R, let $\mathbf{Mod}_R$ denote the category of R-modules. Throughout, K is a field of characteristic zero. The categories of groups and commutative K-algebras are respective denoted by **Groups** and $\mathbf{Alg}_K$. For an affine group scheme $\mathcal{G}$ over K, let $\mathbf{Rep}(\mathcal{G})$ be the category of finite-dimensional representations of $\mathcal{G}$ over K. We use the language of [10] for the theory of Tannakian categories. Our Tannakian categories are all neutral.

2.1. Let $\mathbf{T}$ be a Tannakian category over K with unit object $\mathbb{1}$; thus $\mathbf{T}$ is a K-linear rigid abelian tensor category with the identity $\mathbb{1}$ of the tensor structure satisfying $\mathrm{End}(\mathbb{1}) = K$, for which there exists a fiber functor, i.e., a K-linear exact faithful tensor functor

$$\mathbf{T} \to \mathbf{Mod}_K.$$

Let ω be such a functor. Let

$$\underline{\mathrm{Aut}}^{\otimes}(\omega) : \mathbf{Alg}_K \to \mathbf{Groups}$$

be the functor that sends a commutative K-algebra R to

$$\mathrm{Aut}^{\otimes}(\omega \otimes 1_R) := \begin{array}{l}\text{the group of automorphisms of the functor}\\ \omega \otimes 1_R : \mathbf{T} \to \mathbf{Mod}_R\\ \text{respecting the tensor structures.}\end{array}$$

The fundamental theorem of the theory of Tannakian categories [10, Theorem 2.11] asserts that $\underline{\mathrm{Aut}}^{\otimes}(\omega)$ is representable by an affine group scheme $\mathcal{G}(\mathbf{T}, \omega)$ over K (so that $\underline{\mathrm{Aut}}^{\otimes}(\omega)$ is the functor of points of $\mathcal{G}(\mathbf{T}, \omega)$), and that the functor

$$\mathbf{T} \to \mathbf{Rep}(\mathcal{G}(\mathbf{T}, \omega))$$

sending

$$M \mapsto \omega M$$

(with the natural action of $\mathcal{G}(\mathbf{T}, \omega)$ on ωM) is an equivalence of tensor categories. We call $\mathcal{G}(\mathbf{T}, \omega)$ the fundamental (or the Tannakian) group of $\mathbf{T}$ with respect to ω.

If $\mathbf{T}'$ is also a Tannakian category over K, a tensor functor $\phi : \mathbf{T}' \to \mathbf{T}$ gives rise to a morphism

$$\phi^{\#} : \mathcal{G}(\mathbf{T}, \omega) \to \mathcal{G}(\mathbf{T}', \omega \circ \phi)$$

of group schemes over K, sending an automorphism of $\omega \otimes 1_R$ for any K-algebra R to the obvious automorphism induced on $(\omega \otimes 1_R) \circ \phi = (\omega \circ \phi) \otimes 1_R$. The morphism $\phi^{\#}$ is surjective (or faithfully flat) if and only if ϕ is fully faithful and moreover, satisfies the following property: for every $M \in \mathbf{T}'$, every subobject of $\phi(M)$ is isomorphic to $\phi(L)$ for some subobject L of M (see [10, Proposition 2.21], for instance). In particular, if $\mathbf{T}'$ is a full Tannakian subcategory of $\mathbf{T}$ which is closed under taking subobjects, then the inclusion $\mathbf{T}' \subset \mathbf{T}$ gives rise to a surjective morphism $\mathcal{G}(\mathbf{T}, \omega) \to \mathcal{G}(\mathbf{T}', \omega|_{\mathbf{T}'})$, where $\omega|_{\mathbf{T}'}$ is the restriction of ω to $\mathbf{T}'$.

2.2. Let M be an object of $\mathbf{T}$. Let $\langle M \rangle$ denote the full Tannakian subcategory of $\mathbf{T}$ generated by M, that is, the smallest full Tannakian subcategory of $\mathbf{T}$ that contains M, and is closed under taking subobjects (or subquotients). Set

$$\mathcal{G}(M, \omega) := \mathcal{G}(\langle M \rangle, \omega|_{\langle M \rangle}) = \underline{\mathrm{Aut}}^{\otimes}(\omega|_{\langle M \rangle});$$

we refer to this group as the fundamental (or the Tannakian) group of M with respect to ω. Starting with M and $\mathbb{1}$, we can obtain every object of $\langle M \rangle$ by finitely many iterations of taking direct sums, duals, tensor products, and subquotients. It follows that the natural map

$$\mathcal{G}(M, \omega) \to \mathrm{GL}_{\omega M}, \quad \sigma \mapsto \sigma_M$$

(restricting to the action on ωM) is injective, so that, indeed, $\mathcal{G}(M, \omega)$ is an algebraic group over K. (Here, complying with the standard notation for natural transformations, $\sigma_M : \omega M \to \omega M$ is how σ acts on ωM.) Often we will identify $\mathcal{G}(M, \omega)$ as a subgroup of $\mathrm{GL}_{\omega M}$ via the injection above.

Since $\langle M \rangle$ is closed under taking subobjects, the natural map $\mathcal{G}(\mathbf{T}, \omega) \to \mathcal{G}(M, \omega)$ (induced by the inclusion $\langle M \rangle \subset \mathbf{T}$) is surjective. The kernel of this map consists of all $\sigma \in \mathcal{G}(\mathbf{T}, \omega)$ such that σ_M is identity (then by functoriality, σ_N is also identity for every $N \in \langle M \rangle$).

2.3. For any algebraic group G, let $\mathrm{Lie}(G)$ be the Lie algebra of G. Let $\mathcal{N}$ be a normal subgroup of $\mathcal{G}(M, \omega)$. Consider the adjoint representation

$$\mathrm{Ad} : \mathcal{G}(M, \omega) \to \mathrm{GL}_{\mathrm{Lie}(\mathcal{N})}\,. \tag{3}$$

In view of the equivalence of categories

$$\langle M\rangle \to \mathbf{Rep}(\mathcal{G}(M, \omega)), \quad A \mapsto \omega A, \tag{4}$$

there is a canonical object $\underline{\mathrm{Lie}}(\mathcal{N})$ in $\langle M\rangle$ with

$$\omega\underline{\mathrm{Lie}}(\mathcal{N}) = \mathrm{Lie}(\mathcal{N}),$$

such that the natural action of $\mathcal{G}(M, \omega)$ on $\omega\underline{\mathrm{Lie}}(\mathcal{N})$ (through the definition of $\mathcal{G}(M, \omega)$ as the group of tensor automorphisms of the functor ω) coincides with the adjoint representation (3).

3. The fundamental group of an extension

The goal of this section is to study the fundamental group of an extension in a Tannakian category. As before, let $\mathbf{T}$ be a Tannakian category over a field K of characteristic zero. Fix a fiber functor $\omega : \mathbf{T} \to \mathbf{Mod}_K$. We shall drop ω from the notation for fundamental groups, and simply write $\mathcal{G}(M)$ (for M an object of $\mathbf{T}$). We use the notation I_A for the identity map on an object A of a given category. We use an unadorned Hom to denote a Hom group in a category of modules, with the coefficient ring understood from the context. In $\mathbf{T}$ or any category of modules, the dual of an object A is denoted by $A^\vee$.

3.1. Let L, M and N be objects of $\mathbf{T}$ given in an exact sequence

$$0 \to L \xrightarrow{i} M \xrightarrow{q} N \to 0, \tag{5}$$

where (as indicated in the diagram) the morphisms $L \to M$ and $M \to N$ are respectively denoted by i and q.

The inclusion $\iota : \langle L \oplus N\rangle \subset \langle M\rangle$ induces a surjective morphism

$$\iota^\# : \mathcal{G}(M) \to \mathcal{G}(L \oplus N).$$

Let $\mathcal{U}(M)$ be the kernel of this map; it consists of those $\sigma \in \mathcal{G}(M)$ which act trivially on $\omega L \oplus \omega N$, or equivalently, on both ωL and ωN (i.e., $\sigma_L = I_{\omega L}$ and $\sigma_N = I_{\omega N}$). Note that while for simplicity we did not incorporate L and N in the notation for $\mathcal{U}(M)$, in general, $\mathcal{U}(M)$ will also depend on L and N. Our goal in this section is to study the group $\mathcal{U}(M)$.

First, let us describe the map $\iota^\#$ more concretely. Use the map i (see (5)) to identify ωL as a subspace of ωM. Moreover, once and for all, choose a section of the surjection $\omega q : \omega M \to \omega N$ to identify

$$\omega M = \omega L \oplus \omega N$$

(as vector spaces). Then the functor ω applied to the sequence (5) gives

$$0 \to \omega L \to \omega L \oplus \omega N \to \omega N \to 0, \tag{6}$$

where the second and third arrows are the inclusion and projection maps.

Let σ be an element of $\mathcal{G}(M)$. Since σ is an automorphism of the functor ω, we have a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \omega L \arrow[r] \arrow[d, "\sigma_L"] & \omega L \oplus \omega N \arrow[r] \arrow[d, "\sigma_M"] & \omega N \arrow[r] \arrow[d, "\sigma_N"] & 0 \\
0 \arrow[r] & \omega L \arrow[r] & \omega L \oplus \omega N \arrow[r] & \omega N \arrow[r] & 0
\end{tikzcd}$$

It follows that

$$\sigma_M = \begin{pmatrix} \sigma_L & f \\ 0 & \sigma_N \end{pmatrix} \in \mathrm{GL}_{\omega L \oplus \omega N}$$

for some $f \in \mathrm{Hom}(\omega N, \omega L)$. Let

$$G(M) \subset \mathrm{GL}_{\omega L \oplus \omega N}$$

be the subgroup consisting of the elements which stabilize ωL. Regarding $\mathcal{G}(M)$ as a subgroup of $\mathrm{GL}_{\omega M} = \mathrm{GL}_{\omega L \oplus \omega N}$ (via $\sigma \mapsto \sigma_M$), we have

$$\mathcal{G}(M) \subset G(M).$$

Similarly, for any σ in $\mathcal{G}(L \oplus N)$,

$$\sigma_{L \oplus N} = \begin{pmatrix} \sigma_L & 0 \\ 0 & \sigma_N \end{pmatrix} \in \mathrm{GL}_{\omega L \oplus \omega N}\,.$$

Thinking of $\mathcal{G}(L \oplus N)$ (resp. $\mathrm{GL}_{\omega L} \times \mathrm{GL}_{\omega N}$) as a subgroup of $\mathrm{GL}_{\omega L \oplus \omega N}$ via $\sigma \mapsto \sigma_{L \oplus N}$ (resp. the diagonal embedding), we have

$$\mathcal{G}(L \oplus N) \subset \mathrm{GL}_{\omega L} \times \mathrm{GL}_{\omega N}\,.$$

The map $\iota^{\#}$ is then the restriction of

$$\begin{aligned}
\varphi : G(M) &\to \mathrm{GL}_{\omega L} \times \mathrm{GL}_{\omega N} \\
\begin{pmatrix} g & * \\ 0 & g' \end{pmatrix} &\mapsto \begin{pmatrix} g & 0 \\ 0 & g' \end{pmatrix} \qquad (g \in \mathrm{GL}_{\omega L},\ g' \in \mathrm{GL}_{\omega N}).
\end{aligned}$$

Let

$$U(M) := \ker(\varphi).$$

Thus $U(M)$ is the subgroup of $\mathrm{GL}_{\omega L \oplus \omega N}$ consisting of the elements of the form

$$\begin{pmatrix} I_{\omega L} & * \\ 0 & I_{\omega N} \end{pmatrix},$$

and in particular, is an abelian unipotent group. We have a commutative diagram

$$\begin{tikzcd} 1 \arrow[r] & \mathcal{U}(M) \arrow[r] & \mathcal{G}(M) \arrow[r, "\iota^{\#}"] \arrow[d, hook] & \mathcal{G}(L \oplus N) \arrow[r] \arrow[d, hook] & 1 \\ 1 \arrow[r] & U(M) \arrow[r] & G(M) \arrow[r, "\varphi"] & \mathrm{GL}_{\omega L} \times \mathrm{GL}_{\omega N} \arrow[r] & 1 \end{tikzcd} \tag{7}$$

where the injective arrows are inclusion maps and the rows are exact. Thus

$$\mathcal{U}(M) \subset U(M).$$

Being a subgroup of an abelian unipotent group, $\mathcal{U}(M)$ is abelian and unipotent.

As discussed in Section 2.3, the adjoint representation of $\mathcal{G}(M)$ gives a canonical object $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ of $\langle M \rangle$ whose image under ω is $\mathrm{Lie}(\mathcal{U}(M))$. Since $\mathcal{U}(M)$ is abelian, the action of $\mathcal{G}(M)$ on $\mathrm{Lie}(\mathcal{U}(M))$ factors through an action of $\mathcal{G}(L \oplus N)$, so that indeed, the object $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ belongs to the subcategory $\langle L \oplus N \rangle$.

The Lie algebra of $U(M)$ can be identified with

$$\mathrm{Hom}(\omega N, \omega L)$$

(with trivial Lie bracket). The exponential map

$$\exp : \mathrm{Lie}(U(M)) = \mathrm{Hom}(\omega N, \omega L) \to U(M)(K)$$

(with its inverse denoted by log) is given by

$$\exp(f) = \begin{pmatrix} I_{\omega L} & f \\ 0 & I_{\omega N} \end{pmatrix}. \tag{8}$$

Let $\underline{\mathrm{Hom}}(N, L)$ denote the internal hom object in the category **T**. We identify

$$\omega(\underline{\mathrm{Hom}}(N, L)) = \mathrm{Hom}(\omega N, \omega L)$$

via the canonical isomorphism between the two.

The following observation is standard.

Proposition 3.1.1. *The inclusion map*

$$\mathrm{Lie}(\mathcal{U}(M)) \to \mathrm{Hom}(\omega N, \omega L)$$

is ω of a morphism

$$\underline{\mathrm{Lie}}(\mathcal{U}(M)) \to \underline{\mathrm{Hom}}(N, L).$$

(*In other words,* $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ *can be identified as a subobject of* $\underline{\mathrm{Hom}}(N, L)$.)

Proof. In view of the equivalence of categories (4), it is enough to show that the natural actions of $\mathcal{G}(M)$ on $\mathrm{Lie}(\mathcal{U}(M))$ and $\mathrm{Hom}(\omega N, \omega L)$ are compatible. In other words, we need to show that for any commutative K-algebra R and $\sigma \in \mathcal{G}(M)(R)$, we have

$$\sigma_{\underline{\mathrm{Lie}}(\mathcal{U}(M))} = \sigma_{\underline{\mathrm{Hom}}(N,L)}|_{\mathrm{Lie}(\mathcal{U}(M))_R},$$

where for any vector space V over K, we denote $V_R := V \otimes R$. We may identify

$$(\omega\underline{\mathrm{Hom}}(N, L))_R = \mathrm{Hom}((\omega N)_R, (\omega L)_R))$$

(Hom in R-modules). Considering the evaluation map $N \otimes N^\vee \to \mathbb{1}$ and the canonical isomorphism $\underline{\mathrm{Hom}}(N, L) \cong N^\vee \otimes L$ (which after applying ω, are the corresponding maps in linear algebra), one easily sees that the map $\sigma_{\underline{\mathrm{Hom}}(N,L)}$ is given by

$$f \mapsto \sigma_L \circ f \circ \sigma_N^{-1} \quad (f \in \mathrm{Hom}((\omega N)_R, (\omega L)_R)).$$

We now calculate the map $\sigma_{\underline{\mathrm{Lie}}(\mathcal{U}(M))}$. By definition, the action of $\mathcal{G}(M)$ on $\mathrm{Lie}(\mathcal{U}(M))$ is the restriction of the adjoint representation of $\mathcal{G}(M)$ to $\mathrm{Lie}(\mathcal{U}(M))$. Let

$$f \in \mathrm{Lie}(\mathcal{U}(M))_R \subset \mathrm{Lie}(U(M))_R = \mathrm{Hom}((\omega N)_R, (\omega L)_R).$$

Then $\sigma_{\underline{\mathrm{Lie}}(\mathcal{U}(M))}(f)$ is characterized by

$$\exp(\sigma_{\underline{\mathrm{Lie}}(\mathcal{U}(M))}(f)) = \sigma_M \exp(f) \sigma_M^{-1},$$

where exp is the isomorphism between $\mathrm{Lie}(\mathcal{U}(M))$ and $\mathcal{U}(M)$ as varieties over K, and via the inclusion $\mathcal{U}(M) \subset U(M)$, is given by (8) (with coefficients extended to R). Writing

$$\sigma_M = \begin{pmatrix} \sigma_L & h \\ 0 & \sigma_N \end{pmatrix},$$

where $h \in \mathrm{Hom}((\omega N)_R, (\omega L)_R)$, we have

$$\begin{aligned} \sigma_M \exp(f) \sigma_M^{-1} &= \begin{pmatrix} \sigma_L & h \\ 0 & \sigma_N \end{pmatrix} \begin{pmatrix} I_{(\omega L)_R} & f \\ 0 & I_{(\omega N)_R} \end{pmatrix} \begin{pmatrix} \sigma_L^{-1} & -\sigma_L^{-1} \circ h \circ \sigma_N^{-1} \\ 0 & \sigma_N^{-1} \end{pmatrix} \\ &= \begin{pmatrix} I_{(\omega L)_R} & \sigma_L \circ f \circ \sigma_N^{-1} \\ 0 & I_{(\omega N)_R} \end{pmatrix} = \exp(\sigma_L \circ f \circ \sigma_N^{-1}). \end{aligned}$$

Thus

$$\sigma_{\underline{\mathrm{Lie}}(\mathcal{U}(M))}(f) = \sigma_L \circ f \circ \sigma_N^{-1},$$

as desired. □

Remark 3.1.2. (1) The embedding

$$\mathrm{Lie}(\mathcal{U}(M)) \subset \mathrm{Hom}(\omega N, \omega L)$$

is independent of the section of ωq used to identify $\omega M = \omega L \oplus \omega N$. Indeed, if we had chosen a different section of ωq and hence a different identification of ωM

as $\omega L \oplus \omega N$, then the resulting embedding $\mathcal{G}(M) \hookrightarrow \mathrm{GL}_{\omega L \oplus \omega N}$ would differ from the previous one by conjugation by an element of $U(M)$. Since $U(M)$ is abelian, the two embeddings agree on $\mathcal{U}(M)$. Thus our identification of $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ as a subobject of $\underline{\mathrm{Hom}}(N, L)$ is independent of the choice of the section of ωq.

(2) If L and N are semisimple, then $\mathcal{U}(M)$ is the unipotent radical of $\mathcal{G}(M)$, and in particular will only depend on M (and not on the choices of L or N). (Recall that L and N are semisimple if and only if the category $\langle L \oplus N \rangle$ is semisimple if and only if $\mathcal{G}(L \oplus N)$ is reductive.)

3.2. Before we proceed any further, let us recall a categorical construction. The extension (5) gives an element of

$$\mathrm{Ext}(N, L),$$

where Ext denotes the Yoneda Ext^1 group in $\mathbf{T}$. Recall that one has a canonical isomorphism

$$\mathrm{Ext}(N, L) \cong \mathrm{Ext}(\mathbb{1}, \underline{\mathrm{Hom}}(N, L)). \tag{9}$$

Let

$$\nu \in \mathrm{Ext}(\mathbb{1}, \underline{\mathrm{Hom}}(N, L))$$

be the extension class corresponding to (5) under the canonical isomorphism (9). Then ν is the class of the extension obtained by first applying $\underline{\mathrm{Hom}}(N, -)$ to the sequence (5):

$$0 \to \underline{\mathrm{Hom}}(N, L) \to \underline{\mathrm{Hom}}(N, M) \to \underline{\mathrm{Hom}}(N, N) \to 0,$$

and then pulling back along the canonical morphism

$$e : \mathbb{1} \to \underline{\mathrm{Hom}}(N, N)$$

characterized by the fact that

$$\omega e(1) \in \omega\underline{\mathrm{Hom}}(N, N) = \mathrm{Hom}(\omega N, \omega N)$$

is the identity map. Going through this procedure, assuming $N \neq 0$, we see that ν is the class of the extension

$$0 \to \underline{\mathrm{Hom}}(N, L) \to \underline{\mathrm{Hom}}(N, M)^{\dagger} \to \mathbb{1} \to 0, \tag{10}$$

where

- $\underline{\mathrm{Hom}}(N, M)^{\dagger}$ is the subobject of $\underline{\mathrm{Hom}}(N, M)$ characterized by

$$\begin{aligned} &\omega\underline{\mathrm{Hom}}(N, M)^{\dagger} \\ &\qquad = \mathrm{Hom}(\omega N, \omega M)^{\dagger} \\ &\qquad := \left\{ f \in \mathrm{Hom}(\omega N, \omega M) : (\omega q) \circ f = \lambda(f)\,\mathrm{Id}_{\omega N} \text{ for some } \lambda(f) \in K \right\}, \end{aligned}$$

- after applying ω, the injective arrow is $f \mapsto (\omega i) \circ f$, and

• after applying ω, the surjective arrow is the map $f \mapsto \lambda(f)$, where $\lambda(f) \in K$ is as in the definition of $\underline{\mathrm{Hom}}(N, M)^\dagger$ above.

If N (and hence $\underline{\mathrm{Hom}}(N, L)$) is zero, then ν is the trivial extension

$$0 \to 0 \to \mathbb{1} \to \mathbb{1} \to 0.$$

For convenience, we set $\underline{\mathrm{Hom}}(N, M)^\dagger := \mathbb{1}$ in this case.

3.3. We are ready to give the characterization of the subobject $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ of $\underline{\mathrm{Hom}}(N, L)$. To simplify the notation, we identify $\underline{\mathrm{Hom}}(N, L)$ with its image under the injection $\underline{\mathrm{Hom}}(N, L) \to \underline{\mathrm{Hom}}(N, M)^\dagger$.

Theorem 3.3.1. *Let A be a subobject of* $\underline{\mathrm{Hom}}(N, L)$. *Then A contains* $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ *if and only if the quotient*

$$\underline{\mathrm{Hom}}(N, M)^\dagger/A$$

belongs to the subcategory $\langle L \oplus N\rangle$. (Thus $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ *is the smallest subobject of* $\underline{\mathrm{Hom}}(N, L)$ *with this property.)*

Proof. The theorem is trivial if $N = 0$, so we may assume $N \neq 0$. An object X of $\langle M\rangle$ belongs to the subcategory $\langle L \oplus N\rangle$ if and only if the subgroup $\mathcal{U}(M)$ of $\mathcal{G}(M)$ acts trivially on ωX. Thus the assertion in the theorem can be paraphrased as that A contains $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ if and only if the action of $\mathcal{U}(M)$ on $\omega(\underline{\mathrm{Hom}}(N, M)^\dagger/A)$ is trivial.

Let $\sigma \in \mathcal{G}(M)(K)$. Let $A \subset \underline{\mathrm{Hom}}(N, L)$. The morphism

$$\underline{\mathrm{Hom}}(N, M)^\dagger \to \underline{\mathrm{Hom}}(N, M)^\dagger/A$$

gives rise to a commutative diagram

$$\begin{tikzcd}
\mathrm{Hom}(\omega N, \omega M)^\dagger \arrow[r] \arrow[d, "\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger}"] & \mathrm{Hom}(\omega N, \omega M)^\dagger/\omega A \arrow[d, "\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger/A}"] \\
\mathrm{Hom}(\omega N, \omega M)^\dagger \arrow[r] & \mathrm{Hom}(\omega N, \omega M)^\dagger/\omega A
\end{tikzcd}$$

Thus

$$\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger/A}(f + \omega A) = \sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger}(f) + \omega A$$

for every $f \in \mathrm{Hom}(\omega N, \omega M)^\dagger$.

As before, we use our fixed section of $\omega q : \omega M \to \omega N$ to identify $\omega M = \omega L \oplus \omega N$. Then we have

$$\begin{tikzcd}
\mathrm{Hom}(\omega N, \omega M) \arrow[r, equal] & \mathrm{Hom}(\omega N, \omega L) \oplus \mathrm{Hom}(\omega N, \omega N) \\
\mathrm{Hom}(\omega N, \omega M)^\dagger \arrow[r, equal] \arrow[u, hook] & \mathrm{Hom}(\omega N, \omega L) \oplus K \cdot I_{\omega N} \arrow[u, hook]
\end{tikzcd} \tag{11}$$

Suppose $\sigma \in \mathcal{U}(M)(K)$. Then σ_L and σ_N are both identity, and the action of σ on the $\mathcal{G}(M)$-invariant subspace $\mathrm{Hom}(\omega N, \omega L)$ of $\mathrm{Hom}(\omega N, \omega M)^\dagger$ is trivial. Thus

$$\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger/A} = I_{\mathrm{Hom}(\omega N,\omega M)^\dagger/\omega A}$$

if and only if

$$\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger/A}(I_{\omega N} + \omega A) = I_{\omega N} + \omega A,$$

where here, as well as in the rest of this argument except in (12) below, $I_{\omega N}$ is considered as an element of $\mathrm{Hom}(\omega N, \omega M)^\dagger$ via the decomposition (11). This is equivalent to

$$\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger}(I_{\omega N}) - I_{\omega N} \in \omega A.$$

Note that $\sigma_{\underline{\mathrm{Hom}}(N,M)}$ (and hence $\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger}$) is given by

$$f \mapsto \sigma_M \circ f \circ \sigma_N^{-1} = \sigma_M \circ f \quad (f \in \mathrm{Hom}(\omega N, \omega M)).$$

We have

$$\sigma_M = \begin{pmatrix} I_{\omega L} & \log(\sigma_M) \\ 0 & I_{\omega N} \end{pmatrix} \in \mathrm{GL}_{\omega L \oplus \omega N}(K), \tag{12}$$

where $\log(\sigma_M) \in \mathrm{Hom}(\omega N, \omega L)$. Then

$$\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger}(I_{\omega N}) = \sigma_M \circ I_{\omega N} = \log(\sigma_M) + I_{\omega N},$$

so that

$$\sigma_{\underline{\mathrm{Hom}}(N,M)^\dagger}(I_{\omega N}) - I_{\omega N} = \log(\sigma_M).$$

We have shown that any element $\sigma \in \mathcal{U}(M)(K)$ acts trivially on $\omega(\underline{\mathrm{Hom}}(N, M)^\dagger/A)$ if and only if $\log(\sigma_M)$ is in ωA. The group $\mathcal{U}(M)$ is unipotent and hence $\mathcal{U}(M)(K)$ is dense in $\mathcal{U}(M)$. It follows that $\mathcal{U}(M)$ acts trivially on $\omega(\underline{\mathrm{Hom}}(N, M)^\dagger/A)$ if and only if for every $\sigma \in \mathcal{U}(M)(K)$, we have $\log(\sigma_M) \in \omega A$, i.e., if and only if $\mathrm{Lie}(\mathcal{U}(M) \subset \omega A$. This completes the proof. □

3.4. For every subobject A of $\underline{\mathrm{Hom}}(N, L)$, pushing extensions forward along the natural map $\underline{\mathrm{Hom}}(N, L) \to \underline{\mathrm{Hom}}(N, L)/A$ we have a map

$$\mathrm{Ext}(\mathbb{1}, \underline{\mathrm{Hom}}(N, L)) \to \mathrm{Ext}(\mathbb{1}, \underline{\mathrm{Hom}}(N, L)/A).$$

We denote the image of ν under this map by ν/A. Theorem 3.3.1 has the following corollary:

Corollary 3.4.1. (a) *If A is a subobject of $\underline{\mathrm{Hom}}(N, L)$ such that ν/A is trivial, then $\underline{\mathrm{Lie}}(\mathcal{U}(M)) \subset A$.*

(b) *Suppose L and N are semisimple. Then $\nu/\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is trivial (and hence $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is the smallest subobject of $\underline{\mathrm{Hom}}(N, L)$ with this property).*

Proof. We may assume $N \neq 0$. Suppose ν/A is trivial. Then

$$\underline{\mathrm{Hom}}(N, M)^{\dagger}/A \simeq \underline{\mathrm{Hom}}(N, L)/A \oplus \mathbb{1},$$

and hence $\underline{\mathrm{Hom}}(N, M)^{\dagger}/A$ belongs to the subcategory generated by L and N. Thus (a) follows from Theorem 3.3.1.

As for (b), the theorem implies that $\underline{\mathrm{Hom}}(N, M)^{\dagger}/\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is in $\langle L \oplus N\rangle$, which is a semisimple category by the hypothesis of semisimplicity of L and N. Thus $\nu/\underline{\mathrm{Lie}}(\mathcal{U}(M))$ splits. □

Remark 3.4.2. The semisimple case of Corollary 3.4.1 is originally due to Hardouin (see Theorem 2 of [15] as well as Théorème 2.1 of [14]). Hardouin shows that when L and N are semisimple, $\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is the smallest subobject of $\underline{\mathrm{Hom}}(N, L)$ such that $\nu/\underline{\mathrm{Lie}}(\mathcal{U}(M))$ is trivial. The same statement was earlier proved by Bertrand [5, Theorem 1.1] in the special case where $\mathbf{T}$ is the category of D-modules over a differential field of characteristic zero. (Both Hardouin and Bertrand take $N = \mathbb{1}$, but one can deduce the case of arbitrary (semisimple) N from that.)

3.5. Consider the canonical nondegenerate pairing

$$(L^{\vee} \otimes N) \otimes \underline{\mathrm{Hom}}(N, L) \to \mathbb{1} \tag{13}$$

given (after applying ω) by

$$(\gamma \otimes x) \otimes f \mapsto \gamma(f(x)).$$

For any subobject A of $\underline{\mathrm{Hom}}(N, L)$ (resp. $L^{\vee} \otimes N$), we denote by $A^{\perp}$ the subobject of $L^{\vee} \otimes N$ (resp. $\underline{\mathrm{Hom}}(N, L)$) orthogonal to A with respect to the above pairing. It is clear that A can be recovered from $A^{\perp}$ by $A^{\perp\perp} = A$.

In particular, we have a subobject

$$\underline{\mathrm{Lie}}(\mathcal{U}(M))^{\perp} \subset L^{\vee} \otimes N.$$

In this subsection we shall give a dual variant of Theorem 3.3.1 which characterizes this object. In some situations (such as the application in Section 4), this variant might be more convenient to use than the original version.

Let

$$\mu \in \mathrm{Ext}(L^{\vee} \otimes N, \mathbb{1})$$

be the extension class corresponding to the defining extension of M (i.e., (5)) under the canonical isomorphism

$$\mathrm{Ext}(N, L) \cong \mathrm{Ext}(L^{\vee} \otimes N, \mathbb{1}). \tag{14}$$

The extension class μ is obtained as follows. Let

$$ev : L^{\vee} \otimes L \to \mathbb{1}$$

be the evaluation pairing between L and its dual. Then μ is the class of the pushforward of the extension

$$0 \to L^\vee \otimes L \xrightarrow{I_{L^\vee}\otimes i} L^\vee \otimes M \xrightarrow{I_{L^\vee}\otimes q} L^\vee \otimes N \to 0$$

(obtained by tensoring (5) by $L^\vee$) through the morphism ev. More explicitly, when L is not zero, μ is given by the extension

$$0 \to \mathbb{1} \to (L^\vee \otimes M)^\dagger \to L^\vee \otimes N \to 0, \tag{15}$$

where

- $(L^\vee \otimes M)^\dagger$ is the quotient of $L^\vee \otimes M$ by $(I_{L^\vee} \otimes i)(\ker(ev))$,
- the injective arrow is the composition

$$\mathbb{1} \xrightarrow{\simeq,\ \text{induced by } ev} (L^\vee \otimes L)\big/\ker(ev) \xrightarrow{\text{induced by } I_{L^\vee}\otimes i} (L^\vee \otimes M)^\dagger, \text{ and}$$

- the surjective arrow is induced by $I_{L^\vee} \otimes q$.

If $L = 0$, then μ is given by the extension

$$0 \to \mathbb{1} \to \mathbb{1} \to 0 \to 0.$$

For convenience, in this case we set $(L^\vee \otimes M)^\dagger := \mathbb{1}$.

We shall use the following notation for restrictions of extensions. For every subobject B of $L^\vee \otimes N$, let $\mu|_B$ be the restriction of μ to B (i.e., the pullback of μ along the inclusion map $B \to L^\vee \otimes N$).

We can now state the dual variants of Theorem 3.3.1 and Corollary 3.4.1.

Theorem 3.5.1. *Let B be a subobject of $L^\vee \otimes N$. Then*

$$B \subset \underline{\mathrm{Lie}}(\mathcal{U}(M))^\perp$$

if and only if the preimage of B under the surjective arrow in (15) *belongs to the subcategory $\langle L \oplus N\rangle$.*

Proof. One can prove this directly, similar to the proof of Theorem 3.3.1, by calculating the action of $\mathcal{U}(M)$ on $(L^\vee \otimes M)^\dagger$ (and its subobjects) explicitly. We shall instead use a few categorical considerations to show that the statement is equivalent to Theorem 3.3.1. Let T be an object of $\mathbf{T}$. For any subobject A of T, denote by $A^\perp$ the orthogonal complement of A with respect to the evaluation pairing

$$T^\vee \otimes T \to \mathbb{1}.$$

Dualizing the exact sequence

$$0 \to A \to T \to T/A \to 0,$$

we get

$$0 \to (T/A)^\vee \to T^\vee \to A^\vee \to 0.$$

Use this to identify

$$A^{\perp} \overset{\text{(by definition)}}{=} \ker(T^{\vee} \to A^{\vee}) \cong (T/A)^{\vee}.$$

There is a commutative diagram

$$\begin{tikzcd}
\mathrm{Ext}(\mathbb{1}, T) \arrow[r, "\text{dualizing}, \simeq"] \arrow[d, "\text{pushforward}"'] & \mathrm{Ext}(T^{\vee}, \mathbb{1}) \arrow[d, "\text{pullback}"] \\
\mathrm{Ext}(\mathbb{1}, T/A) \arrow[r, "\text{dualizing}, \simeq"] & \mathrm{Ext}(A^{\perp}, \mathbb{1})
\end{tikzcd}$$

where the horizontal maps dualize extensions. Apply this with $T = \underline{\mathrm{Hom}}(N, L)$, and use the pairing (13) to identify $L^{\vee} \otimes N$ as $T^{\vee}$ (so that (13) becomes simply the evaluation pairing between T and $T^{\vee}$). It is easy to see that ν and μ are duals of one another, with the isomorphism between $(L^{\vee} \otimes M)^{\dagger}$ and the dual of $\underline{\mathrm{Hom}}(N, M)^{\dagger}$ defined by the pairing

$$(L^{\vee} \otimes M)^{\dagger} \otimes \underline{\mathrm{Hom}}(N, M)^{\dagger} \to \mathbb{1},$$

which after applying ω is given by

$$\overline{g \otimes x} \otimes f \mapsto g\big(\lambda(f)x - f((\omega q)(x))\big).$$

(Here $\overline{g \otimes x}$ is the image of $g \otimes x \in \omega(L^{\vee}) \otimes \omega M$ in $\omega(L^{\vee} \otimes M)^{\dagger}$, and f is in $\mathrm{Hom}(\omega N, \omega M)^{\dagger}$.) Thus by the above diagram, for any subobject A of $\underline{\mathrm{Hom}}(N, L)$, we have an isomorphism between

$$(\underline{\mathrm{Hom}}(N, M)^{\dagger}/A)^{\vee}$$

and the preimage of $A^{\perp}$ under the surjective arrow in (15). The equivalence of Theorems 3.3.1 and 3.5.1 is clear from this. $\square$

The argument also gives the following dual variant of Corollary 3.4.1:

Corollary 3.5.2. (a) *If B is a subobject of $L^{\vee} \otimes N$ such that $\mu|_B$ is trivial, then $B \subset \underline{\mathrm{Lie}}(\mathcal{U}(M))^{\perp}$.*

(b) *Suppose L and N are semisimple. Then the restriction of μ to $\underline{\mathrm{Lie}}(\mathcal{U}(M))^{\perp}$ is trivial.* (*Hence $\underline{\mathrm{Lie}}(\mathcal{U}(M))^{\perp}$ is the largest subobject of $L^{\vee} \otimes N$ with this property.*)

4. The unipotent radical of the Mumford–Tate group of H^1 of an algebraic curve

Let **MHS** be the category of rational mixed Hodge structures. The category **MHS** is a neutral Tannakian category over $\mathbb{Q}$. The forgetful functor $\omega_B : \mathbf{MHS} \to \mathbf{Mod}_{\mathbb{Q}}$ sending an object to its underlying rational vector space is a fiber functor. For any rational mixed Hodge structure M, the group $\mathcal{G}(M)$ with $(\mathbf{T}, \omega) = (\mathbf{MHS}, \omega_B)$ is called the Mumford–Tate group of M. In this section, we will use the results of the previous section to study the unipotent radical of the Mumford–Tate group of the degree one cohomology of a smooth complex projective curve minus a finite set of points.

4.1. ***Notation.*** By a mixed Hodge structure we always mean a rational one. As usual, $\mathbb{Q}(-n)$ denotes the Hodge structure of weight $2n$ with underlying rational vector space $(2\pi i)^{-n}\mathbb{Q}$, with its complexification identified with $\mathbb{C}$ via

$$(2\pi i)^{-n} \otimes 1 \mapsto (2\pi i)^{-n}.$$

The unit object $\mathbb{1}$ is $\mathbb{Q}(0)$. For any object M of **MHS**, we denote by $M_{\mathbb{Q}}$ the underlying rational vector space of M. If R is a commutative $\mathbb{Q}$-algebra, M_R denotes $M_{\mathbb{Q}} \otimes R$.

Given a pure Hodge structure H of weight -1, we denote by JH the intermediate Jacobian

$$JH := \frac{H_{\mathbb{C}}}{F^0 H_{\mathbb{C}} + H_{\mathbb{Q}}},$$

where $F^{\cdot}$ is the Hodge filtration.

Given any smooth complex variety X, by $H^i(X)$ we mean the mixed Hodge structure on the degree i Betti cohomology of X (with underlying rational vector space $H^i(X, \mathbb{Q})$). We shall identify $H^i(X)_{\mathbb{C}} = H^i(X, \mathbb{C})$ with $H^i_{dR}(X)$ (= smooth complex de Rham cohomology) via the isomorphism of de Rham. By $H_i(X)$ we mean the dual of $H^i(X)$; it is a mixed Hodge structure with underlying rational vector space $H_i(X, \mathbb{Q})$.

All the Ext (= Yoneda Ext^1) groups in this section are in **MHS**.

4.2. Carlson [6] gives an explicit description of Ext groups in **MHS**. We briefly recall this description here in a special case that is of interest to us.

Let A be a pure Hodge structure of weight 1. Carlson gives a canonical isomorphism

$$\mathrm{Ext}(A, \mathbb{1}) \to J(A^{\vee}),$$

where $A^{\vee}$ is the dual Hodge structure to A. The isomorphism is functorial in A. See [6] for details. (Carlson [6] proves the analogous result for integral mixed Hodge structures. The proof of the rational case is identical.)

4.3. From this point on, let X be a smooth complex projective curve. We denote the Jacobian variety of X by $\mathrm{Jac}(X)$. Let $CH_0^{\mathrm{hom}}(X)$ be the group of divisors of degree 0 on X modulo the subgroup of principal divisors. (In other words, $CH_0^{\mathrm{hom}}(X)$ is the homologically trivial subgroup of the Chow group $CH_0(X)$.) The group $CH_0^{\mathrm{hom}}(X)$ is the group of complex points of $\mathrm{Jac}(X)$.

Set $CH_0^{\hom}(X)_{\mathbb{Q}} := CH_0^{\hom}(X) \otimes \mathbb{Q}$. The Abel–Jacobi map on X gives an isomorphism

$$AJ_X : CH_0^{\hom}(X)_{\mathbb{Q}} \to JH_1(X),$$

sending the class of $p-q$, with $p, q \in X$, to the class of the functional $\int_q^p$ on the space of harmonic 1-forms on X. (See, for instance, [2, Chapter 1]. Note that here, said integral means the integral over any path from q to p. The choice of the path will not matter in $JH_1(X)$.)

Composing AJ_X with Carlson's isomorphism we get an isomorphism

$$\operatorname{Ext}(H^1(X), \mathbb{1}) \cong CH_0^{\hom}(X)_{\mathbb{Q}}. \tag{16}$$

We shall identify these two groups to simplify the notation.

4.4. Let S be a finite nonempty set of (complex) points of X. We identify $H^1(X)$ as a subobject of $H^1(X \setminus S)$ via the map induced by the inclusion $(X \setminus S) \subset X$. The reader can refer to Deligne's [7, Section 10.3] for a thorough study of the mixed Hodge structure $H^1(X \setminus S)$.

Since $X \setminus S$ is affine, every element of $H^1(X \setminus S)_{\mathbb{C}}$ can be represented by a meromorphic differential form on X with possible singularities only along S, and has a well-defined residue at every $p \in X$. Indeed, if $c = [\omega]$ with ω a meromorphic form, set $\operatorname{res}_p(c) := \operatorname{res}_p(\omega)$ (= the residue of ω at p, which is $1/(2\pi i)$ times the integral of ω along a small positively oriented loop around p). The subspace $H^1(X)_{\mathbb{C}}$ of $H^1(X \setminus S)_{\mathbb{C}}$ consists of the cohomology classes with zero residue everywhere (in other words, classes of differentials of the second kind).

For any vector space or mixed Hodge structure V, we denote by $(V^S)'$ the kernel of the map

$$V^S \to V, \quad (v_p)_{p\in S} \mapsto \sum_{p\in S} v_p$$

(where the v_p are in V).

One has a short exact sequence of mixed Hodge structures

$$0 \to H^1(X) \to H^1(X \setminus S) \xrightarrow{\operatorname{res}_S} (\mathbb{Q}(-1)^S)' \to 0, \tag{17}$$

where the injective arrow is inclusion and $\operatorname{res}_S : H^1(X \setminus S)_{\mathbb{C}} \to (\mathbb{C}^S)'$ is the map $c \mapsto (\operatorname{res}_p(c))_{p\in S}$.

4.5. We shall apply the results of Section 3.3 to the exact sequence (17). The Hodge structure $H^1(X)$ is polarizable and hence semisimple (see, for instance, [19, Section 7.1.2] and [9, Proposition 3.6]). Thus the group

$$\mathcal{U}(H^1(X \setminus S)) := \ker\big(\mathcal{G}(H^1(X \setminus S)) \to \mathcal{G}(H^1(X) \oplus \mathbb{Q}(-1))\big)$$

is the unipotent radical of the Mumford–Tate group of $H^1(X \setminus S)$. In view of Section 3.5, the determination of the group $\mathcal{U}(H^1(X \setminus S))$ amounts to finding

$$\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp} \subset H^1(X)^{\vee} \otimes (\mathbb{Q}(-1)^S)'.$$

We use the Poincaré duality isomorphism

$$PD : H^1(X)(1) \to H^1(X)^{\vee}, \quad [\eta] \mapsto \frac{1}{2\pi i}\int_X \eta \wedge -,$$

where η is a closed smooth 1-form on X and the isomorphism

$$H^1(X)(1) \otimes (\mathbb{Q}(-1)^S)' \to (H^1(X)^S)', \quad c \otimes (a_p)_{p\in S} \mapsto (a_p c)_{p\in S}$$

to identify

$$H^1(X)^{\vee} \otimes (\mathbb{Q}(-1)^S)' \cong (H^1(X)^S)'.$$

Following the notation of Section 3.5, we let

$$\mu \in \mathrm{Ext}\big((H^1(X)^S)', \mathbb{1}\big)$$

be the element corresponding to the sequence (17) under the canonical isomorphism

$$\begin{aligned} \mathrm{Ext}\big((\mathbb{Q}(-1)^S)', H^1(X)\big) &\cong \mathrm{Ext}\big(H^1(X)^{\vee} \otimes (\mathbb{Q}(-1)^S)', \mathbb{1}\big) \\ &= \mathrm{Ext}\big((H^1(X)^S)', \mathbb{1}\big). \end{aligned} \tag{18}$$

By Corollary 3.5.2 (and on recalling that $H^1(X)$ is semisimple), we have that $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp}$ is the largest subobject of $(H^1(X)^S)'$ with the property that the restriction of μ to it is trivial.

4.6. Let us consider the restrictions of μ to some obvious subobjects of $(H^1(X)^S)'$. For each $p \in S$, let $\iota_p : H^1(X) \to H^1(X)^S$ be the embedding into the p-coordinate. Given $p, q \in S$, we have a morphism

$$\iota_p - \iota_q : H^1(X) \to (H^1(X)^S)'$$

(which is an embedding if $p \neq q$).

Proposition 4.6.1. *Let $p, q \in S$. Via the identification* (16), *we have*

$$(\iota_p - \iota_q)^*(\mu) = p - q$$

(*where $(\iota_p - \iota_q)^*(\mu)$ is the pullback of μ along $\iota_p - \iota_q$, and with abuse of notation the class of $p - q$ in $CH_0^{hom}(X)_{\mathbb{Q}}$ is also denoted by $p - q$*).

Proof. This is a reformulation of a well-known result about Hodge theory of open curves, which in turn is a special case of general results about equivalence of various definitions of the Abel–Jacobi map (see the remark below). With abuse of notation, let ι_p also denote the embedding of $\mathbb{Q}(-1)$ as the p-coordinate of $\mathbb{Q}(-1)^S$. Then we have a commutative diagram

$$\begin{tikzcd} \mathrm{Ext}\big((\mathbb{Q}(-1)^S)', H^1(X)\big) \arrow[r, "(\iota_p-\iota_q)^*"] \arrow[d, "\cong"'] & \mathrm{Ext}(\mathbb{Q}(-1), H^1(X)) \arrow[d, "\cong"] \\ \mathrm{Ext}\big((H^1(X)^S)', \mathbb{1}\big) \arrow[r, "(\iota_p-\iota_q)^*"] & \mathrm{Ext}(H^1(X), \mathbb{1}) \end{tikzcd}$$

where the vertical isomorphisms are given by (14) and Poincaré duality. Under the isomorphism on the left (i.e., (18)), μ and (17) correspond to each other. The pullback of the extension (17) along $\iota_p - \iota_q : \mathbb{Q}(-1) \to (\mathbb{Q}(-1)^S)'$ is the extension

$$\begin{array}{c} 0 \to H^1(X) \to (\mathrm{res}_S)^{-1}\big((\iota_p - \iota_q)(\mathbb{Q}(-1))\big) \xrightarrow{\mathrm{res}_p} \mathbb{Q}(-1) \to 0 \\ \Vert \\ H^1(X \setminus \{p, q\}) \end{array} \tag{19}$$

This extension corresponds to $p - q$ under

$$\mathrm{Ext}(\mathbb{Q}(-1), H^1(X)) \cong \mathrm{Ext}(H^1(X), \mathbb{1}) \cong JH_1(X) \cong CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}}. \tag{20}$$

See, for example, Sections 9.0–9.2 of Jannsen [17]. □

Remark 4.6.2. The fact that the extension (19) corresponds to $p - q$ under (20) is already stated in Section 4.3 of Deligne's [8]. The same paragraph outlines a motivically inspired definition of the Abel–Jacobi map, which naturally takes values in Ext groups in any suitable cohomology theory. Via this approach and in the case of Hodge theory (or more precisely, cohomology with values in **MHS**), the Abel–Jacobi image of $p-q$ is by definition the extension (19) (in other cohomology theories, by definition the Abel–Jacobi image is the analogous extension). A detailed description of this motivic approach towards the Abel–Jacobi map for any smooth complex variety can be found in Sections 9.0 and 9.1 of Jannsen's book [17]. The fact that for Hodge theory the Abel–Jacobi map defined in terms of extensions coincides with the classical (Griffiths) Abel–Jacobi map with values in intermediate Jacobians is asserted in Lemma 9.2 of [17] and follows from the works [16] and [13] of Jannsen and Esnault–Viehweg. (See Section 9.2 of [17] for more details.)

4.7. We now calculate the slightly more complicated restrictions of μ. Let

$$E = \mathrm{End}^0(\mathrm{Jac}(X)) := \mathrm{End}(\mathrm{Jac}(X)) \otimes \mathbb{Q}$$

be the endomorphism algebra of the Jacobian of X. We have an (anti-) isomorphism

$$E \to \mathrm{End}(H^1(X)), \quad f \mapsto f^\star,$$

where for any element f of the endomorphism algebra of $\mathrm{Jac}(X)$, by $f^\star$ we mean the pullback map on cohomology.[1] This induces an isomorphism

$$(E^S)' \to \mathrm{Hom}\big(H^1(X), (H^1(X)^S)'\big), \quad (f_p)_{p\in S} \mapsto \sum_{p\in S} \iota_p f_p^\star. \tag{21}$$

Consider the composition

$$\begin{aligned} &\mathrm{Hom}\big(H^1(X), (H^1(X)^S)'\big) \\ &\qquad \xrightarrow{\phi\mapsto\phi^*\mu} \mathrm{Ext}(H^1(X), \mathbb{1}) \overset{(16)}{\cong} CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}} = \mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q}. \end{aligned} \tag{22}$$

Since every simple subobject of $(H^1(X)^S)'$ is the image of a morphism

$$H^1(X) \to (H^1(X)^S)'$$

(because $\langle H^1(X)\rangle$ is semisimple), the following corollary of Proposition 4.6.1 can be used to describe all restrictions of μ.

Corollary 4.7.1. *Let $(f_p)_{p\in S} \in (E^S)'$. Then*

$$\Big(\sum_{p\in S} \iota_p f_p^\star\Big)^* \mu = \sum_{p\in S} f_p(p-e),$$

where e is any point in X.

Proof. Let $(f_p)_{p\in S} \in (E^S)'$ and $e \in S$. Since $\sum_{p\in S} f_p = 0$, we have

$$\sum_{p\in S} \iota_p f_p^\star = \sum_{p\in S} (\iota_p - \iota_e) f_p^\star.$$

Thus

$$\Big(\sum_{p\in S} \iota_p f_p^\star\Big)^* \mu = \sum_{p\in S} ((\iota_p - \iota_e) f_p^\star)^* \mu = \sum_{p\in S} (f_p^\star)^* (\iota_p - \iota_e)^* \mu = \sum_{p\in S} f_p(p-e),$$

where in the last line we used Proposition 4.6.1 together with the commutativity of the diagram

$$\begin{array}{ccc} \mathrm{Ext}(H^1(X), \mathbb{1}) \cong JH_1(X) & \xleftarrow{\simeq, AJ_X} & CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}} = \mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q} \\ {\scriptstyle (f^\star)^*}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ \mathrm{Ext}(H^1(X), \mathbb{1}) \cong JH_1(X) & \xleftarrow[\simeq, AJ_X]{} & CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}} = \mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q} \end{array} \qquad \square$$

For any $(f_p)_{p\in S} \in (E^S)'$, the value of $\sum_{p\in S} f_p(p-e)$ does not depend on the choice of $e \in X$. To simplify the notation, let us denote this common value by

$$\sum_{p\in S} f_p(p).$$

[1] We use the symbol $*$ for pullback of extensions and the symbol $\star$ for pullback of cohomology induced by morphisms of varieties.

Note that if

$$(f_p)_{p\in S} \in (\mathbb{Q}^S)' \subset (E^S)',$$

then $\sum_{p\in S} f_p(p)$ defined above agrees with the other possible interpretation of the notation (i.e., the image of the divisor $\sum_{p\in S} f_p p$ of degree zero with coefficients in $\mathbb{Q}$ in $CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}}$).

4.8. Having computed the restrictions of μ, we return to the problem of determination of $\mathcal{U}(H^1(X\setminus S))$.

Proposition 4.8.1. *Let $(f_p)_{p\in S} \in (E^S)'$. The following statements are equivalent*:

(i) *The restriction of μ to the image of $\sum_{p\in S} \iota_p f_p^\star$ splits.*

(ii) *$\sum_{p\in S} f_p(p)$ is zero in $CH_0^{hom}(X)_{\mathbb{Q}}$.*

(iii) *The image of $\sum_{p\in S} \iota_p f_p^\star$ is contained in $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X\setminus S))\big)^\perp$.*

Proof. Recall that by Corollary 3.5.2, $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X\setminus S))\big)^\perp$ is the largest subobject of $(H^1(X)^S)'$ with the property that the restriction of μ to it splits (see Section 4.5). This gives the equivalence of statements (i) and (iii) of the proposition. The equivalence of (i) and (ii) follows from Corollary 4.7.1, on noting (by weight considerations) that for any quotient B of $H^1(X)$, the canonical map

$$\mathrm{Ext}(B, \mathbb{1}) \to \mathrm{Ext}(H^1(X), \mathbb{1})$$

is injective. □

In particular, the proposition recovers the following well-known result, originally due to Deligne (see the remark below), which gives an arithmetic criterion for when $\mathcal{U}(H^1(X\setminus S))$ is trivial (or equivalently, for when the sequence (17) splits):

Corollary 4.8.2. *The group $\mathcal{U}(H^1(X\setminus S))$ is trivial if and only if the subgroup of the Jacobian of X supported on S has zero rank.*

Proof. Note that $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X\setminus S))\big)^\perp = (H^1(X)^S)'$ if and only if $\mathrm{Im}(\iota_p - \iota_q)$ is contained in $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X\setminus S))\big)^\perp$ for every $p, q \in S$, which in turn is equivalent to $p-q$ being zero in $CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}}$ for every $p, q \in S$. □

Remark 4.8.3. Corollary 4.8.2 is originally due to Deligne, implicit in [7] and announced explicitly in [8, Remarque 7.5], in relation to a new proof of the Manin–Drinfeld theorem on modular curves. See [11] for a more detailed discussion of this.

4.9. We are ready to give the main result of this part of the paper. The results gives a characterization of $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X\setminus S))\big)^\perp$ (and hence $\mathcal{U}(H^1(X\setminus S))$).

Theorem 4.9.1. *Let A be the subobject of $(H^1(X)^S)'$ which is the sum of the images of all the maps of the form*

$$\sum_{p\in S} \iota_p f_p^\star \in \mathrm{Hom}\big(H^1(X), (H^1(X)^S)'\big),$$

with $(f_p)_{p\in S} \in (E^S)'$ *and* $\sum_{p\in S} f_p(p) = 0$ *(see Section 4.7). Then*

$$A = \underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp}.$$

Proof. The inclusion

$$A \subset \underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp}$$

is immediate from Proposition 4.8.1. To see the reverse inclusion, first note that since $\langle H^1(X)\rangle$ is semisimple, $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp}$ is a direct sum of simple subobjects. Let B be a simple subobject of $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp}$. Then B is the image of a map $H^1(X) \to (H^1(X)^S)'$. Any such map is of the form $\sum_{p\in S} \iota_p f_p^\star$ for some $(f_p)_{p\in S} \in (E^S)'$. By Proposition 4.8.1, for the image of such a map to be in $\underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp}$ we must have

$$\sum_{p\in S} f_p(p) = 0$$

in $CH_0^{\mathrm{hom}}(X)_{\mathbb{Q}}$. Thus $B \subset A$. □

We end the paper by deducing the following result about the dimension of $\mathcal{U}(H^1(X \setminus S))$ (note that part (b) is Theorem A of the introduction).

Theorem 4.9.2. *Let g be the genus of X. Recall that E is the endomorphism algebra of the Jacobian* $\mathrm{Jac}(X)$.

(a) *Suppose D is any division algebra contained in E. Then the dimension of $\mathcal{U}(H^1(X \setminus S))$ is at most $2g$ times the D-rank of the D-submodule of* $\mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q}$ *generated by the subgroup supported on S.*

(b) *Suppose $H^1(X)$ is simple. Then the dimension of $\mathcal{U}(H^1(X \setminus S))$ is equal to $2g$ times the E-rank of the E-submodule of* $\mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q}$ *generated by the subgroup supported on S.*

Proof. Let A be as in Theorem 4.9.1.

(a) For any subalgebra R of E, let Λ_R be the composition

$$(R^S)' \hookrightarrow (E^S)' \xrightarrow{(21)} \mathrm{Hom}\big(H^1(X), (H^1(X)^S)'\big) \xrightarrow{(22)} \mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q}.$$

This is R-linear by Corollary 4.7.1. The image of Λ_R is the R-submodule of $\mathrm{Jac}(X)(\mathbb{C}) \otimes \mathbb{Q}$ generated by the subgroup supported on S. Let A_R be the subobject of $(H^1(X)^S)'$ which is the sum of the images of the maps $\sum_{p\in S} \iota_p f_p^\star$ with $(f_p)_{p\in S}$ in $\ker(\Lambda_R)$, so that $A_R \subset A$ and $A_E = A$. If $\beta = \{(f_p^{(r)})_{p\in S}\}_{1\le r\le d}$ is an R-spanning set for $\ker(\Lambda_R)$, then A_R is the sum of the images of $\sum_{p\in S} \iota_p (f_p^{(r)})^\star$ for $1 \le r \le d$. Moreover, if $R = D$ is a division algebra and β is D-linearly independent, then A_D is the direct sum of the images of the previous d maps. Since each of these images is then a copy of $H^1(X)$ (because D is a division algebra), we have

$$\begin{aligned}\dim \underline{\mathrm{Lie}}\big(\mathcal{U}(H^1(X \setminus S))\big)^{\perp} = \dim(A) \ge \dim(A_D) &= 2g \cdot \dim_D(\ker(\Lambda_D)) \\ &= 2g(|S| - 1 - \dim_D \mathrm{Im}(\Lambda_D)).\end{aligned}$$

Taking orthogonal complements we get the desired bound.

(b) Since $H^1(X)$ is simple, E is a division algebra. Taking $D = E$, by the proof of part (b) we have

$$\dim(A) = \dim(A_E) = 2g(|S| - 1 - \dim_E \operatorname{Im}(\Lambda_E)).$$

The claimed formula follows. □

Acknowledgements

We thank Madhav Nori for his comments on a draft of the paper. When we wrote the first draft of this paper, we were not aware of the related works of Bertrand [5], Hardouin [14; 15] and Jossen [18]. We thank the referees of the earlier versions of the article for bringing these works to our attention, as well as for providing comments that helped us improve the results and exposition of the article. We also thank a referee for the *Pacific Journal of Mathematics* for providing us with many comments that helped us significantly improve the exposition of the second part of the paper.

Eskandari was at the University of Toronto and the Fields Institute for Research in Mathematical Sciences when this work was done. He wishes to thank both institutions for providing a pleasant working environment.

This work was a prequel to the work [12].

References

[1] Y. André, "Mumford–Tate groups of mixed Hodge structures and the theorem of the fixed part", *Compositio Math.* **82**:1 (1992), 1–24. MR Zbl

[2] E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves, I*, Grundl. Math. Wissen. **267**, Springer, New York, 1985. MR Zbl

[3] C. Bertolin, "The Mumford–Tate group of 1-motives", *Ann. Inst. Fourier* (*Grenoble*) **52**:4 (2002), 1041–1059. MR Zbl

[4] C. Bertolin, "Le radical unipotent du groupe de Galois motivique d'un 1-motif", *Math. Ann.* **327**:3 (2003), 585–607. MR Zbl

[5] D. Bertrand, "Unipotent radicals of differential Galois group and integrals of solutions of inhomogeneous equations", *Math. Ann.* **321**:3 (2001), 645–666. MR

[6] J. A. Carlson, "Extensions of mixed Hodge structures", pp. 107–127 in *Journées de Géometrie Algébrique d'Angers* (Juillet, 1979), Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. MR Zbl

[7] P. Deligne, "Théorie de Hodge, III", *Inst. Hautes Études Sci. Publ. Math.* **44** (1974), 5–77. MR Zbl

[8] P. Deligne, "Valeurs de fonctions L et périodes d'intégrales", pp. 313–346 in *Proc. Sympos. Pure Math.*, *part* 2, vol. 33, Amer. Math. Soc., Providence, R.I., 1979. MR Zbl

[9] P. Deligne, "Hodge cycles on abelian varieties", pp. 9–100 in *Hodge cycles*, *motives*, *and Shimura varieties*, Lecture Notes in Mathematics **900**, Springer, Berlin, 1982. Zbl

[10] P. Deligne and J. S. Milne, "Tannakian categories", pp. 101–228 in *Hodge cycles, motives, and Shimura varieties*, Lecture Notes in Mathematics **900**, Springer, Berlin, 1982. Zbl

[11] R. Elkik, "Le théorème de Manin–Drinfeld", pp. 59–67 in *Séminaire sur les Pinceaux de Courbes Elliptiques*, Astérisque **183**, Société Mathématique de France, Paris, 1990. MR Zbl

[12] P. Eskandari and V. K. Murty, "On unipotent radicals of motivic Galois groups", *Algebra Number Theory* **17**:1 (2023), 165–215. MR

[13] H. Esnault and E. Viehweg, "Deligne–Beilinson cohomology", pp. 43–91 in *Beilinson's conjectures on special values of L-functions*, Perspect. Math. **4**, Academic Press, Boston, MA, 1988. MR Zbl

[14] C. Hardouin, "Hypertranscendance et Groupes de Galois aux différences", 2006. arXiv 0609646

[15] C. Hardouin, "Unipotent radicals of Tannakian Galois groups in positive characteristic", pp. 223–239 in *Arithmetic and Galois theories of differential equations*, Sémin. Congr. **23**, Soc. Math. France, Paris, 2011. MR Zbl

[16] U. Jannsen, "Deligne homology, Hodge-$\mathscr{D}$-conjecture, and motives", pp. 305–372 in *Beilinson's conjectures on special values of L-functions*, Perspect. Math. **4**, Academic Press, Boston, MA, 1988. MR Zbl

[17] U. Jannsen, *Mixed motives and algebraic K-theory*, Lecture Notes in Mathematics **1400**, Springer, Berlin, 1990. MR Zbl

[18] P. Jossen, "On the Mumford–Tate conjecture for 1-motives", *Invent. Math.* **195**:2 (2014), 393–439. MR Zbl

[19] C. Voisin, *Hodge theory and complex algebraic geometry, I*, Cambridge Studies in Advanced Mathematics **76**, Cambridge University Press, Cambridge, 2002. MR Zbl

Received August 20, 2021. Revised March 23, 2023.

PAYMAN ESKANDARI
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF WINNIPEG
WINNIPEG, MB
CANADA
p.eskandari@uwinnipeg.ca

V. KUMAR MURTY
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO, ON
CANADA
murty@math.toronto.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.281

THE SIZE OF SEMIGROUP ORBITS MODULO PRIMES

WADE HINDES AND JOSEPH H. SILVERMAN

Let V be a projective variety defined over a number field K, let S be a polarized set of endomorphisms of V all defined over K, and let $P \in V(K)$. For each prime $\mathfrak{p}$ of K, let $m_{\mathfrak{p}}(S, P)$ denote the number of points in the orbit of P mod $\mathfrak{p}$ for the semigroup of maps generated by S. Under suitable hypotheses on S and P, we prove an analytic estimate for $m_{\mathfrak{p}}(S, P)$ and use it to show that the set of primes for which $m_{\mathfrak{p}}(S, P)$ grows subexponentially as a function of $\mathsf{N}_{K/\mathbb{Q}}\,\mathfrak{p}$ is a set of density zero. For $V = \mathbb{P}^1$ we show that this holds for a generic set of maps S provided that at least two of the maps in S have degree at least four.

1. Introduction

A general expectation in arithmetic dynamics over number fields is that the dynamical systems generated by "unrelated" self-maps $f_1, f_2 : V \to V$ should not be too similar. For example, they should not have identical canonical heights [16], they should not have infinitely many common preperiodic points [2; 8; 11], their orbits should not have infinite intersection [10], and arithmetically their orbits should not have unexpectedly large common divisors [15]. It is not always clear what "unrelated" should mean, but in any case it includes the assumption that f_1 and f_2 do not share a common iterate.

Similarly, we expect that the points in semigroup orbits generated by all finite compositions of "unrelated" maps f_1 and f_2 should be asymptotically large [4; 13] when ordered by height, where now unrelated means that the semigroup is not unexpectedly small. For example, the semigroup is small if it contains no free

Silverman's research was supported by Simons Collaboration Grant #712332, and by National Science Foundation Grant #1440140 while in residence at MSRI in spring 2023.
MSC2020: primary 37P25; secondary 37P05, 37P55.
Keywords: arithmetic dynamics, finite field dynamics, semigroup dynamics.

subsemigroups requiring at least 2 generators; cf. [3].

In this note, we study the size of semigroup orbits over finite fields. In particular, we show that, under suitable hypotheses, a free semigroup of maps defined over a number field generates many large orbits when reduced modulo primes. See [1; 6; 7; 21] for additional results in this vein.

Definition 1. We set notation that will remain in effect throughout this note.

$K/\mathbb{Q}$	a number field
V/K	a smooth projective variety defined over K
$r \geq 1$	an integer
$S = \{f_1, \ldots, f_r\}$	a set of morphisms $f_i : V \to V$ defined over K
$d_1, \ldots, d_r$	real numbers satisfying $d_i > 1$ for $i = 1, \ldots, r$
$\mathcal{L} \in \operatorname{Pic}_K(V) \otimes \mathbb{R}$	line bundles satisfying $f_i^* \mathcal{L} \cong \mathcal{L}^{\otimes d_i}$
M_S	the semigroup generated by S under composition
$\operatorname{Orb}_S(P)$	the orbit $\{f(P) : f \in M_S\}$ of a point $P \in V$

The following property will play a crucial role in some of our results.

Definition 2. A point $P \in V$ is called *strongly S-wandering* if the evaluation map

$$M_S \to V, \quad f \mapsto f(P), \tag{1}$$

is injective.

Remark 3. If $V = \mathbb{P}^1$ and S is any sufficiently generic set of maps as described in Section 3, then the set of points that fail to be strongly S-wandering is a set of bounded height. In particular, it follows in this case that all infinite orbits contain strongly wandering points, and this weaker condition is sufficient for our orbit bounds.

Our goal is to study the number of points in the reduction of $\operatorname{Orb}_S(P)$ modulo primes. We set some additional notation, briefly recall a standard definition, and then define our principal object of study.

R_K	the ring of integers of K
$\operatorname{Spec}(R_K)$	the set of prime ideals of R_K
$\mathsf{N}\mathfrak{p}$	the norm of $\mathfrak{p} \in \operatorname{Spec}(R_K)$, that is, $\mathsf{N}\mathfrak{p} := \#R_K/\mathfrak{p}$.

Definition 4. Let $\mathfrak{p} \in \operatorname{Spec}(R_K)$, and let $R_\mathfrak{p}$ denote the localization of R_K at $\mathfrak{p}$, and let $k_\mathfrak{p} = R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p}$ denote the residue field. A finite K-morphism $f : V \to V$ has *good reduction at* $\mathfrak{p}$ if there is a scheme $\mathcal{V}_\mathfrak{p}/R_\mathfrak{p}$ that is proper and smooth over $R_\mathfrak{p}$, and there is an $R_\mathfrak{p}$-morphism $F_\mathfrak{p} : \mathcal{V}_\mathfrak{p} \to \mathcal{V}_\mathfrak{p}$ whose generic fiber is $f : V \to V$.[1]

[1]Intuitively, this means that we can find equations for V and for f that have coefficients in R_K, and so that when we reduce the equations modulo $\mathfrak{p}$, the reduced variety $\tilde{V}$ mod $\mathfrak{p}$ is non-singular and the reduced map $\tilde{f} : \tilde{V} \to \tilde{V}$ is a morphism having the same degree as f. Of course, when we say “find equations”, this needs to be done locally on an appropriately fine cover by affine neighborhoods.

We write $\tilde{\mathcal{V}}_\mathfrak{p} = \mathcal{V}_\mathfrak{p} \times_{R_\mathfrak{p}} k_\mathfrak{p}$ for the special fiber of $\mathcal{V}_\mathfrak{p}$. Properness implies that each point $Q \in V(K)$ extends to a section $\mathcal{Q}_\mathfrak{p} \in \mathcal{V}_\mathfrak{p}(R_\mathfrak{p})$, and the *reduction* $\tilde{Q}_\mathfrak{p} \in \tilde{\mathcal{V}}_\mathfrak{p}(k_\mathfrak{p})$ *of* Q *modulo* $\mathfrak{p}$ is the intersection of the image of $\mathcal{Q}_\mathfrak{p}$ with the fiber $\tilde{\mathcal{V}}_\mathfrak{p}$, i.e.,

$$\{\tilde{Q}_\mathfrak{p}\} = \mathcal{Q}_\mathfrak{p}(\operatorname{Spec} R_\mathfrak{p}) \cap \mathcal{V}_\mathfrak{p}.$$

Similarly, the *reduction* $\tilde{f}_\mathfrak{p}$ *of* f *modulo* $\mathfrak{p}$ is the restriction of $F_\mathfrak{p}$ to the special fiber $\tilde{\mathcal{V}}_\mathfrak{p}$.

Remark 5. Continuing with notation from Definition 4, we note that if f has good reduction at $\mathfrak{p}$, then reduction modulo $\mathfrak{p}$ commutes with evaluation,

$$\widetilde{f(Q)}_\mathfrak{p} = \tilde{f}_\mathfrak{p}(\tilde{Q}_\mathfrak{p}).$$

Further, composition commutes with reduction for good reduction maps. In other words, if f and g have good reduction at $\mathfrak{p}$, then

$$\big(\widetilde{f \circ g}\big)_\mathfrak{p} = \tilde{f}_\mathfrak{p} \circ \tilde{g}_\mathfrak{p}.$$

Definition 6. Let $\mathfrak{p} \in \operatorname{Spec}(R_K)$. Continuing with notation from Definition 4, let $f_1, \dots, f_r : V \to V$ be maps that have good reduction modulo $\mathfrak{p}$, and let $P \in V(K)$. Then the reduction of the S-orbit of P modulo $\mathfrak{p}$ is the set

$$\operatorname{Orb}_{\tilde{S}}(\tilde{P} \bmod \mathfrak{p}) := \big\{\tilde{f}_\mathfrak{p}(\tilde{P}_\mathfrak{p}) : f \in M_S\big\}.$$

We define

$$m_\mathfrak{p} := m_\mathfrak{p}(S, P) = \#\operatorname{Orb}_{\tilde{S}}(\tilde{P} \bmod \mathfrak{p})$$

to be the size of the mod $\mathfrak{p}$ reduction of $\operatorname{Orb}_S(P)$. (If any of the maps $f_1, \dots, f_r$ has bad reduction at $\mathfrak{p}$, then we formally set $m_\mathfrak{p} = \infty$.)

Our main result is an analytic formula that implies that $m_\mathfrak{p}$ is not too small on average.

Theorem 7. *Assume that M_S is a free semigroup, that $P \in V(K)$ is a strongly S-wandering point, and that $r = \#S \geq 2$. Then there exists a constant $C_1 = C_1(K, V, S, P)$ such that, for all $\epsilon > 0$,*

$$\sum_{\mathfrak{p} \in \operatorname{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p} \cdot m_\mathfrak{p}(S, P)^\epsilon} \leq C_1 \epsilon^{-1}. \tag{2}$$

Remark 8. The principal result of the paper [21] is an estimate exponentially weaker than (2) in the case that $r = \#S = 1$, while a principal result of the paper [17] is an estimate that exactly mirrors (2) with $m_\mathfrak{p}$ equal to the number of points on the mod $\mathfrak{p}$ reduction of the multiples of a point on an abelian variety. Thus the present paper, as well as the papers [4; 13], suggest that the analogy

$$\begin{pmatrix}\text{arithmetic of points}\\ \text{of an abelian variety}\end{pmatrix} \iff \begin{pmatrix}\text{arithmetic of points in orbits}\\ \text{of a dynamical system}\end{pmatrix}$$

described in [5] and [22, §6.5] may be more accurate when the dynamical system on the right-hand side is generated by at least two non-commuting maps, rather than using orbits coming from iteration of a single map.

Estimate (2) can be used to show that there are few primes $\mathfrak{p}$ for which $m_{\mathfrak{p}}(S,P)$ is subexponential compared to $\mathsf{N}\mathfrak{p}$. We quantify this assertion in the following corollary.

Corollary 9. *Let S, M_S and P be as in Theorem 7, and let $\overline{\delta}$ and δ denote the (upper) logarithmic analytic densities on sets of primes as described in Definition 12.*

(a) *There is a constant $C_2 = C_2(K, V, S, P)$ such that*

$$\overline{\delta}\big(\big\{\mathfrak{p} \in \operatorname{Spec}(R_K) : m_{\mathfrak{p}}(S, P) \le \mathsf{N}\mathfrak{p}^{\gamma}\big\}\big) \le C_2\gamma$$

holds for all $0 < \gamma < 1$.

(b) *Let $L(t)$ be a subexponential function, i.e., a function with the property that*

$$\lim_{t\to\infty} \frac{L(t)}{t^{\mu}} = 0 \quad \textit{for all } \mu > 0.$$

Then

$$\delta\big(\big\{\mathfrak{p} \in \operatorname{Spec}(R_K) : m_{\mathfrak{p}}(S, P) \le L(\mathsf{N}\mathfrak{p})\big\}\big) = 0.$$

In the special case that $V = \mathbb{P}^1$, we show that the conclusions of Theorem 7 and Corollary 9 are true for generic sets of maps. In the statement of the next result, we write Rat_d for the space of rational maps of $\mathbb{P}^1$ of degree $d \ge 2$, so in particular Rat_d is an affine variety of dimension $2d+1$; see [20, §4.3] for details.

Theorem 10. *Let $r \ge 2$, and let $d_1, \dots, d_r$ be integers satisfying*

$$d_1, d_2 \ge 4 \quad \textit{and} \quad d_3, \dots, d_r \ge 2.$$

Then there is a Zariski dense subset

$$\mathcal{U} = \mathcal{U}(d_1, \dots, d_r) \subseteq \operatorname{Rat}_{d_1} \times \cdots \times \operatorname{Rat}_{d_r}$$

such that the inequality (2) in Theorem 7 and the density estimates in Corollary 9 are true for all number fields $K/\mathbb{Q}$, all $S \in \mathcal{U}(K)$, and all $P \in \mathbb{P}^1(K)$ for which $\operatorname{Orb}_S(P)$ is infinite.

The contents of this paper are as follows. In Section 2 we build upon prior work [17; 21] of the second author to prove Theorem 7 and Corollary 9. Then in Section 3 we use results from [10; 13; 23] to construct many sets of maps on $\mathbb{P}^1$

for which the bounds in Section 2 apply. The key step is to construct a point in every infinite orbit that is strongly wandering. The construction is explicit, and in particular, Theorem 15 describes an explicit set $\mathcal{U}$ for which Theorem 10 is true.

2. The size of orbits modulo $\mathfrak{p}$

We start with a key estimate.

Proposition 11. *Let S, M_S and P be as in Theorem 7. For each $m \geq 2$, we define an integral ideal*

$$\mathfrak{D}(m) = \mathfrak{D}(m; K, V, S, P) := \prod_{\substack{\mathfrak{p}\in\operatorname{Spec}(R_K)\\ m_\mathfrak{p}(S,P)\leq m}} \mathfrak{p}. \tag{3}$$

There are constants $C_i = C_i(K, V, S, P)$ for $i = 3, 4$ such that the following hold:

(a) *If $r = \#S = 1$, then*

$$\log\log \mathsf{N}\mathfrak{D}(m) \leq C_3 m \quad \textit{for all } m \geq 2.$$

(b) *Assume that S generates a free semigroup, that $P \in V(K)$ is strongly S-wandering, and that $r = \#S \geq 2$. Then*

$$\log\log \mathsf{N}\mathfrak{D}(m) \leq C_4 \log m \quad \textit{for all } m \geq 2.$$

Proof. (a) This is [21, Proposition 10].

(b) Next, since V and S are polarized with respect to some line bundle $\mathcal{L}$, we may choose $N \geq 1$ and an embedding $V \subseteq \mathbb{P}^N$ such that the f_i extend to self-morphisms of $\mathbb{P}^N$. Next, for notational convenience, we write $m_\mathfrak{p}$ for $m_\mathfrak{p}(S, P)$ and use the standard combinatorics notation $[r] = \{1, 2, \ldots, r\}$. Also to ease notation, we write

$$f_{\boldsymbol{i}} := f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_k} \quad \text{for} \quad \boldsymbol{i} = (i_1, \ldots, i_k) \in [r]^k.$$

Let

$$m \geq 1 \quad \text{and} \quad k = k(m) := \left\lceil \frac{\log(m+1)}{\log r} \right\rceil.$$

For each good reduction prime $\mathfrak{p}$, we consider the map that sends a function $f_{\boldsymbol{i}}$ to the image of P under reduction modulo $\mathfrak{p}$,

$$[r]^k \to \mathcal{O}_{\tilde{S}}(\tilde{P} \bmod \mathfrak{p}), \quad \boldsymbol{i} \mapsto f_{\boldsymbol{i}}(P) \pmod{\mathfrak{p}}. \tag{4}$$

If

$$m_\mathfrak{p} \leq m, \quad \text{then} \quad r^k > m_\mathfrak{p} \quad \text{by our choice of } k,$$

so the map (4) cannot be injective (pigeonhole principle) and there exist

$$\boldsymbol{i} \neq \boldsymbol{j} \text{ in } [r]^k \quad \text{satisfying} \quad f_{\boldsymbol{i}}(P) \equiv f_{\boldsymbol{j}}(P) \pmod{\mathfrak{p}}.$$

Sine we have assumed that P is strongly wandering, i.e., that the map

$$M_S \to V(K), \quad f \mapsto f(P),$$

is injective, it follows that the global points are distinct,

$$f_{\boldsymbol{i}}(P) \neq f_{\boldsymbol{j}}(P),$$

so the ideals generated by their "differences" are non-zero.

More formally, [21, Lemma 9] says that there is an integral ideal $\mathfrak{C} = \mathfrak{C}_K \subseteq R_K$ with the property that every point $Q \in \mathbb{P}^N(K)$ can be written with homogeneous coordinates

$$Q = [\alpha_0, \dots, \alpha_N]$$

satisfying[2]

$$\alpha_0, \dots, \alpha_N \in R_K \quad \text{and} \quad (\alpha_0 R_K + \cdots \alpha_N R_K) \mid \mathfrak{C}_K.$$

Applying [21, Lemma 9] to our situation, for the given $P \in V(K)$, we can write

$$f_{\boldsymbol{i}}(P) = \bigl[A_0(\boldsymbol{i}), \dots, A_N(\boldsymbol{i})\bigr]$$

with $A_0(\boldsymbol{i}), \dots, A_N(\boldsymbol{i}) \in R_K$ and such that the ideal

$$\mathfrak{A}(\boldsymbol{i}) := A_0(\boldsymbol{i}) R_K + \cdots + A_N(\boldsymbol{i}) R_K \quad \text{divides the ideal } \mathfrak{C}. \tag{5}$$

Then for $\mathfrak{p} \nmid \mathfrak{C}$ we have

$$\begin{aligned} & f_{\boldsymbol{i}}(P) \equiv f_{\boldsymbol{j}}(P) \pmod{\mathfrak{p}} \\ & \quad \iff \quad A_u(\boldsymbol{i}) A_v(\boldsymbol{j}) \equiv A_v(\boldsymbol{i}) A_u(\boldsymbol{j}) \pmod{\mathfrak{p}} \quad \text{for all } 0 \le u, v \le N. \end{aligned}$$

We define a difference ideal

$$\mathfrak{B}(\boldsymbol{i}, \boldsymbol{j}) := \sum_{0 \le u, v \le N} \bigl(A_u(\boldsymbol{i}) A_v(\boldsymbol{j}) - A_v(\boldsymbol{i}) A_u(\boldsymbol{j})\bigr) R_K,$$

and the product of the difference ideals

$$\mathfrak{D}'(m) := \prod_{\boldsymbol{i}, \boldsymbol{j} \in [r]^k,\ \boldsymbol{i} \neq \boldsymbol{j}} \mathfrak{B}(\boldsymbol{i}, \boldsymbol{j}).$$

Then

$$\mathfrak{p} \nmid \mathfrak{C} \text{ and } m_{\mathfrak{p}} \le m \quad \Longrightarrow \quad \mathfrak{p} \mid \mathfrak{D}'(m),$$

and hence

$$\mathfrak{D}(m) \mid \mathfrak{C} \cdot \mathfrak{D}'(m).$$

[2] If R_K is a PID, then we can take $\mathfrak{C}_K = R_K$, so $\gcd(\alpha_0, \dots, \alpha_N) = 1$. Thus [21, Lemma 9] provides a weaker version of this gcd result that holds for all R_K.

Since $\mathfrak{C}$ depends only on K, it remains to estimate the norm of $\mathfrak{D}'(m)$.

Let $h(\cdot)$ denote the logarithmic Weil height on $\mathbb{P}^N$, and let $\mathfrak{A}(\boldsymbol{i})$ for $\boldsymbol{i} \in [r]^k$ be the ideals defined by (5). Then, using [21, Proposition 7], we find for all $\boldsymbol{i}$ and $\boldsymbol{j}$ that

$$\frac{1}{[K:\mathbb{Q}]} \log \frac{\mathsf{N}\mathfrak{B}(\boldsymbol{i}, \boldsymbol{j})}{\mathsf{N}\mathfrak{A}(\boldsymbol{i}) \cdot \mathsf{N}\mathfrak{A}(\boldsymbol{j})} \le h(f_{\boldsymbol{i}}(P)) + h(f_{\boldsymbol{j}}(P)) + C_5,$$

where C_5 is an absolute constant. Since $\mathsf{N}\mathfrak{A}(\boldsymbol{i})$ and $\mathsf{N}\mathfrak{A}(\boldsymbol{j})$ are smaller than $\mathsf{N}\mathfrak{C}$, this implies that

$$\frac{1}{[K:\mathbb{Q}]} \log \mathsf{N}\mathfrak{B}(\boldsymbol{i}, \boldsymbol{j}) \le h(f_{\boldsymbol{i}}(P)) + h(f_{\boldsymbol{j}}(P)) + C_6.$$

Next we apply the height estimate

$$h(f_{\boldsymbol{i}}(P)) \le C_7 \cdot \prod_{u=1}^{k} d_{i_u},$$

which is a weak form of [12, Lemma 2.1]. This yields

$$\frac{1}{[K:\mathbb{Q}]} \log \mathsf{N}\mathfrak{B}(\boldsymbol{i}, \boldsymbol{j}) \le C_7 \cdot \prod_{u=1}^{k} d_{i_u} + C_7 \cdot \prod_{u=1}^{k} d_{j_u} + C_6.$$

This gives

$$\begin{aligned}
\log \mathfrak{D}'(m) &= \sum_{\boldsymbol{i}, \boldsymbol{j} \in [r]^k,\, \boldsymbol{i} \ne \boldsymbol{j}} \log \mathfrak{B}(\boldsymbol{i}, \boldsymbol{j}) \\
&\le \sum_{\boldsymbol{i}, \boldsymbol{j} \in [r]^k,\, \boldsymbol{i} \ne \boldsymbol{j}} \left(C_7 \cdot \prod_{u=1}^{k} d_{i_u} + C_7 \cdot \prod_{u=1}^{k} d_{j_u} + C_6 \right) \\
&\le C_8 \cdot r^k \cdot \sum_{\boldsymbol{i} \in [r]^k} \prod_{u=1}^{k} d_{i_u} \\
&= C_8 \cdot \left(r \cdot \sum_{i \in [r]} d_i \right)^k \\
&\le C_9 \cdot \left(r \cdot \sum_{i \in [r]} d_i \right)^{1 + \frac{\log(m+1)}{\log r}}.
\end{aligned}$$

Hence

$$\log\log \mathfrak{D}'(m) \le C_{10} \cdot \log(m+1) + C_{11}.$$

Since $m \ge 2$, we can absorb C_{11} into C_{10}, although we remark that if we leave in $C_{11}(K, V, S, P)$, then we can take C_{10} to depend on only the degrees of the maps in S,

$$C_{10} = C_{10}(d_1, \dots, d_r) = 1 + \frac{\log(d_1 + \cdots + d_r)}{\log r}.$$

This completes the proof of Proposition 11. $\square$

Proof of Theorem 7. To ease notation, we let

$$g(t) = \frac{\log t}{t} \quad \text{and} \quad G(t) = \frac{1}{t^\epsilon}.$$

We start with two elementary estimates. First, the mean value theorem gives

$$G(m) - G(m+1) \le \sup_{m \le t \le m+1} -G'(t) = \sup_{m \le t \le m+1} \frac{\epsilon}{t^{1+\epsilon}} = \frac{\epsilon}{m^{1+\epsilon}}. \tag{6}$$

Second, an easy integral calculation gives

$$\sum_{m \ge 1} g(m) G(m) \le \int_1^\infty \frac{\log x}{x^{1+\epsilon}}\, \mathrm{dx} = \frac{1}{\epsilon^2}. \tag{7}$$

We use these and our other calculations to estimate

$$\begin{aligned}
&\sum_{\mathfrak{p} \in \operatorname{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p} \cdot m_{\mathfrak{p}}^{\epsilon}} \\
&\qquad = \sum_{\mathfrak{p} \in \operatorname{Spec}(R_K)} g(\mathsf{N}\mathfrak{p}) \cdot G(m_{\mathfrak{p}}) && \text{(by definition of } g \text{ and } G\text{)} \\
&\qquad = \sum_{m \ge 1} G(m) \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R_K) \\ m_{\mathfrak{p}} = m}} g(\mathsf{N}\mathfrak{p}) \\
&\qquad = \sum_{m \ge 1} \bigl(G(m) - G(m+1)\bigr) \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R_K) \\ m_{\mathfrak{p}} \le m}} g(\mathsf{N}\mathfrak{p}) && \text{(Abel summation)} \\
&\qquad \le \sum_{m \ge 1} \frac{\epsilon}{m^{1+\epsilon}} \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R_K) \\ m_{\mathfrak{p}} \le m}} g(\mathsf{N}\mathfrak{p}) && \text{(from (6))} \\
&\qquad = \sum_{m \ge 1} \frac{\epsilon}{m^{1+\epsilon}} \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R_K) \\ \mathfrak{p} \mid \mathfrak{D}(m)}} g(\mathsf{N}\mathfrak{p}) && \text{(by definition (3) of } \mathfrak{D}(m)\text{)} \\
&\qquad \le \sum_{m \ge 1} \frac{\epsilon}{m^{1+\epsilon}} \cdot \bigl(C_{12} \log\log \mathfrak{D}(m) + C_{13}\bigr) && \text{(from [17, Corollary 2.3])} \\
&\qquad \le C_{14} \sum_{m \ge 1} \frac{\epsilon}{m^{1+\epsilon}} \cdot \log m && \text{(from Proposition 11(b))} \\
&\qquad = C_{14} \cdot \epsilon \cdot \sum_{m \ge 1} g(m) \cdot G(m) && \text{(by definition of } g \text{ and } G\text{)} \\
&\qquad \le C_{15} \epsilon^{-1} && \text{(from (7)).}
\end{aligned}$$

$\square$

Definition 12. Let $\mathcal{P} \subset \operatorname{Spec}(R_K)$ be a set of primes. The *upper logarithmic analytic density of* $\mathcal{P}$ is

$$\overline{\delta}(\mathcal{P}) := \limsup_{s\to 1^+} (s-1) \sum_{\mathfrak{p}\in\mathcal{P}} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}^s}.$$

Similarly, the *logarithmic analytic density of* $\mathcal{P}$, denoted $\delta(P)$, is given by the same formula with a limit, instead of a lim sup.

Proof of Corollary 9. (a) For any $0 < \gamma < 1$, we let

$$\mathcal{P}_\gamma := \left\{\mathfrak{p} \in \operatorname{Spec}(R_K) : m_{\mathfrak{p}} \le \mathsf{N}\mathfrak{p}^\gamma\right\}.$$

Then

$$\begin{aligned}
\frac{C_1}{\epsilon} &\ge \sum_{\mathfrak{p}\in\operatorname{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}\cdot m_{\mathfrak{p}}^{\epsilon}} && \text{(from Theorem 7)}\\
&\ge \sum_{\mathfrak{p}\in\mathcal{P}_\gamma} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}\cdot m_{\mathfrak{p}}^{\epsilon}} && \text{(summing over a smaller set)}\\
&\ge \sum_{\mathfrak{p}\in\mathcal{P}_\gamma} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}^{1+\gamma\epsilon}} && \text{(by definition of } \mathcal{P}_\gamma). \qquad (8)
\end{aligned}$$

This allows us to estimate the upper logarithmic density of $\mathcal{P}_\gamma$ by

$$\begin{aligned}
\overline{\delta}(\mathcal{P}_\gamma) &= \limsup_{s\to 1^+} (s-1) \sum_{\mathfrak{p}\in\mathcal{P}_\gamma} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}^s}\\
&= \limsup_{\epsilon\to 0^+} \gamma\epsilon \sum_{\mathfrak{p}\in\mathcal{P}_\gamma} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}^{1+\gamma\epsilon}} && \text{(setting } s = 1+\gamma\epsilon)\\
&\le \limsup_{\epsilon\to 0^+} \gamma\epsilon \cdot \frac{C_1}{\epsilon} && \text{(from (8))}\\
&= C_1\gamma.
\end{aligned}$$

This completes the proof of Corollary 9(a).

(b) We let

$$\mathcal{P}_L := \left\{\mathfrak{p} \in \operatorname{Spec}(R_K) : m_{\mathfrak{p}} \le L(\mathsf{N}\mathfrak{p})\right\}.$$

The assumption that L is subexponential means that for all $\mu > 0$ there exists a constant $C_{16}(L, \mu)$ depending only on L and μ such that

$$L(t) \le t^\mu \quad \text{for all } t > C_{16}(L, \mu).$$

We also note that

$$
\begin{aligned}
\mathfrak{p} \in \mathcal{P}_L &\iff m_{\mathfrak{p}} \le L(\mathsf{N}\mathfrak{p}) \\
&\implies m_{\mathfrak{p}} \le (\mathsf{N}\mathfrak{p})^{\mu} \quad \text{for all } \mathsf{N}\mathfrak{p} > C_{16}(L,\mu).
\end{aligned} \tag{9}
$$

We now fix a $\mu > 0$ and estimate

$$
\begin{aligned}
\overline{\delta}(\mathcal{P}_L) &= \limsup_{\lambda \to 0^+} \lambda \sum_{\mathfrak{p} \in \mathcal{P}_L} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}^{1+\lambda}} \\
&= \limsup_{\lambda \to 0^+} \lambda \sum_{\substack{\mathfrak{p} \in \mathcal{P}_L \\ \mathsf{N}\mathfrak{p} \ge C_{16}(L,\mu)}} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}^{1+\lambda}} && \text{(since μ is fixed, so we can discard finitely many terms)} \\
&\le \limsup_{\lambda \to 0^+} \lambda \sum_{\substack{\mathfrak{p} \in \mathcal{P}_L \\ \mathsf{N}\mathfrak{p} \ge C_{16}(L,\mu)}} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}} \cdot \frac{1}{m_{\mathfrak{p}}^{\lambda/\mu}} && \text{(from (9))} \\
&\le \limsup_{\lambda \to 0^+} \lambda \sum_{\mathfrak{p} \in \operatorname{Spec} R_K} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}} \cdot \frac{1}{m_{\mathfrak{p}}^{\lambda/\mu}} \\
&\le \limsup_{\lambda \to 0^+} \lambda \cdot C_1 \cdot \left(\frac{\lambda}{\mu}\right)^{-1} && \text{(from Theorem 7)} \\
&= C_1 \mu.
\end{aligned}
$$

This estimate holds for all $\mu > 0$, so we find that

$$\overline{\delta}(\mathcal{P}_L) \le \inf_{\mu > 0} C_1 \cdot \mu = 0,$$

which completes the proof that $\delta(\mathcal{P}_L) = 0$. ☐

3. Orbits of generic families of maps of $\mathbb{P}^1$

In this section, we show that there are many sets of endomorphisms of $\mathbb{P}^1$ for which Theorem 7 holds. To make this statement precise, we need some definitions.

Definition 13. Let f be a non-constant rational map of $\mathbb{P}^1$ defined over $\overline{\mathbb{Q}}$. A point $w \in \mathbb{P}^1(\overline{\mathbb{Q}})$ is a *critical value* of f if $f^{-1}(w)$ contains fewer than $\deg(f)$ elements. It is a *simple critical value* if

$$\# f^{-1}(w) = \deg(f) - 1.$$

The map f is *critically simple* if all of its critical values are simple.

Definition 14. Let f and g be non-constant rational maps of $\mathbb{P}^1$ with respective critical value sets $\operatorname{CritVal}_f$ and $\operatorname{CritVal}_g$. We say that f and g are *critically separated* if

$$\operatorname{CritVal}_f \cap \operatorname{CritVal}_g = \varnothing.$$

Our first result says that the conclusions of Theorem 7 and Corollary 9 hold for certain sets S that contain a pair of critically simple and critically separated maps and initial points P with infinite orbit.

Theorem 15. *Let $K/\mathbb{Q}$ be a number field, let S be a set of endomorphisms of $\mathbb{P}^1$ defined over K containing a pair of critically simple and critically separated maps of degree at least* 4, *and let $P \in \mathbb{P}^1(K)$ be a point with infinite S-orbit. Then there is a constant $C_{17} = C_{17}(K, S, P)$ such that for all $\epsilon > 0$,*

$$\sum_{\mathfrak{p} \in \operatorname{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p} \cdot m_{\mathfrak{p}}(S, P)^{\epsilon}} \le C_{17} \cdot \epsilon^{-1}.$$

Remark 16. In particular, there is a constant $C_{18}(S)$ such that Theorem 15 holds for all $P \in \mathbb{P}^1(K)$ satisfying $h(P) > C_{18}(S)$; see Lemma 18.

We start with a definition and some basic height estimates. In what follows, we fix an embedding $V \subseteq \mathbb{P}^N$ and extend the maps f_i to self-morphisms of $\mathbb{P}^N$; here we use our assumption that S is polarizable with respect to some line bundle $\mathcal{L}$. Moreover, $h(\,\cdot\,)$ denotes the logarithmic Weil height on $\mathbb{P}^N$.

Definition 17. A point $P \in V$ is *moderately S-preperiodic* if

$$g \circ f(P) = f(P) \quad \text{for some } f, g \in M_S \text{ with } g \neq 1.$$

Lemma 18. *Let $V/\overline{\mathbb{Q}}$ be a variety, and let $S = \{f_1, \dots, f_r\}$ be a set of polarized endomorphisms as described in Definition 1. Then there exists a constant $C_{19} = C_{19}(S, V, \mathcal{L})$ such that the following statements hold for all $Q \in V(\overline{\mathbb{Q}})$*:

(a) *If Q is moderately S-preperiodic as described in Definition 17, then $h(Q) \le C_{19}$. In particular, this is true if $\operatorname{Orb}_S(Q)$ is finite.*

(b) *If $h(Q) > C_{19}$, then*

$$h(f(Q)) \ge h(Q) \quad \textit{for all } f \in M_S.$$

Proof. These estimates are proven in [4, Lemma 2.11]. □

We combine Lemma 18 with the techniques in [13; 19] to obtain the following result for pairs of maps that are critically simple and critically separated.

Proposition 19. *Let f_1 and f_2 be endomorphisms of $\mathbb{P}^1$ of degree at least* 4, *let $S = \{f_1, f_2\}$, and suppose that f_1 and f_2 are critically simple and critically separated. Then the following statements hold*:

(a) *The semigroup M_S is free.*

(b) *Let $P \in \mathbb{P}^1(\overline{\mathbb{Q}})$ be a point whose S-orbit $\operatorname{Orb}_S(P)$ is infinite. Then there exists a point $Q \in \operatorname{Orb}_S(P)$ such that Q is strongly S-wandering as described in Definition 2.*

Proof. (a) See [13, Proposition 4.1].

(b) We fix a number field K over which P, f_1, and f_2 are defined. Letting $\Delta \subset \mathbb{P}^1 \times \mathbb{P}^1$ be the diagonal, we define three curves

$$\Gamma_i := (f_i \times f_i)^{-1}(\Delta) \quad \text{for } i = 1, 2,$$
$$\Gamma_{1,2} := (f_1 \times f_2)^{-1}(\Delta).$$

Then the main results in [19] (see also [13, Proposition 4.6]) imply that the curves Γ_1 and Γ_2 are each the union of Δ and an irreducible curve of geometric genus ≥ 2, while $\Gamma_{1,2}$ is itself an irreducible curve of geometric genus ≥ 2.

More specifically, the assumption that f_1 and f_2 are critically simple implies from [19, Corollary 3.6] that $\mathcal{C}_1 \smallsetminus \Delta$ and $\mathcal{C}_2 \smallsetminus \Delta$ are irreducible, while the assumption that f_1 and f_2 are critically separated implies from [19, Proposition 3.1] that $\mathcal{C}_{1,2}$ is irreducible. It then follows from [19, pages 208 and 210] that the geometric genera of these curves are given by the formulas

$$\operatorname{genus}(\Gamma_i \smallsetminus \Delta) = \bigl(\deg(f_i) - 2\bigr)^2 \quad \text{for } i = 1, 2,$$
$$\operatorname{genus}(\Gamma_{1,2} \smallsetminus \Delta) = (\deg(f_1) - 1)(\deg(f_2) - 1).$$

In particular, the assumption that f_1 and f_2 have degree at least 4 ensures that these genera are at least 2.

We now invoke Faltings's theorem [9], [14, Theorem E.0.1] to deduce that the set

$$\Sigma := \Gamma_{1,2}(K) \cup (\Gamma_1 \smallsetminus \Delta)(K) \cup (\Gamma_2 \smallsetminus \Delta)(K)$$

is finite. We note that the definition of Σ says that for all $P, Q \in \mathbb{P}^1(K)$, we have

$$\begin{bmatrix} P \neq Q & \text{and} & f_1(P) = f_1(Q) & \Longrightarrow & (P, Q) \in \Sigma \\ P \neq Q & \text{and} & f_2(P) = f_2(Q) & \Longrightarrow & (P, Q) \in \Sigma \\ & & f_1(P) = f_2(Q) & \Longrightarrow & (P, Q) \in \Sigma \end{bmatrix}. \tag{10}$$

Let $\pi_1, \pi_2 : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^1$ be the two projection maps, and let

$$C_{20} := \max\bigl(\{h(P) : P \in \pi_1(\Sigma)\} \cup \{h(P) : P \in \pi_2(\Sigma)\}\bigr)$$

be the maximum of the heights of the coordinates of the finitely many points in Σ. We then set

$$C_{21} := \max\{C_{19}, C_{20}\},$$

where C_{19} is the constant that appears in Lemma 18.

The fact that $\operatorname{Orb}_S(P) \subseteq \mathbb{P}^1(K)$ is infinite, combined with Northcott's theorem [18] saying that $\mathbb{P}^1(K)$ has only finitely many points of bounded height, implies that there exists a point $Q \in \mathrm{Orb}_S(P)$ satisfying

$$h(Q) > C_{21}. \tag{11}$$

We claim that Q is strongly wandering for S. To see this, suppose that

$$f_{i_1} \circ \cdots \circ f_{i_n}(Q) = f_{j_1} \circ \cdots \circ f_{j_m}(Q), \tag{12}$$

where without loss of generality we may assume that $n \geq m$. Our goal is to prove that $m = n$ and $i_k = j_k$ for all $1 \leq i \leq n$.

To ease notation, we let

$$F = f_{i_2} \circ \cdots \circ f_{i_n} \quad \text{and} \quad G = f_{j_2} \circ \cdots \circ f_{j_m} \tag{13}$$

be the compositions with the initial map omitted. Thus (12) and (13) say that

$$f_{i_1}\big(F(Q)\big) = f_{j_1}\big(G(Q)\big). \tag{14}$$

It follows from (14) and (10) that one of the following is true:

(i) $i_1 = j_1$ and $F(Q) = G(Q)$.

(ii) $i_1 = j_1$ and $F(Q) \neq G(Q)$ and $\big(F(Q), G(Q)\big) \in \Sigma$.

(iii) $i_1 \neq j_1$ and $\big(F(Q), G(Q)\big) \in \Sigma$.

On the other hand, we know that

$$\big(F(Q), G(Q)\big) \in \Sigma \quad \Longrightarrow \quad h\big(F(Q)\big) \leq C_{21} < h(Q)$$

by (11). But this contradicts Lemma 18. Hence (ii) and (iii) are false, so (i) is true.

We recall that $m \leq n$, so repeating this argument, we conclude that

$$i_k = j_k \quad \text{for all } 1 \leq k \leq m.$$

If $m < n$ is a strict inequality, then we see that

$$f_{i_{m+1}} \circ \cdots \circ f_{i_n}(Q) = Q.$$

But then Lemma 18 implies that $h(Q) \leq C_{19} \leq C_{21}$, and we obtain a contradiction of (11). Thus $m = n$ and $i_k = j_k$ for all $1 \leq k \leq n$, which completes the proof that Q is a strongly S-wandering point. □

We now have the tools in place to prove Theorem 15.

Proof of Theorem 15. Let S be the given set of endomorphisms of $\mathbb{P}^1$, and let f_1 and f_2 be the given maps in S that have degree at least 4 and that are critically simple and critically separated. We let

$$S' = \{f_1, f_2\}.$$

We are given that the point $P \in \mathbb{P}^1(K)$ has infinite S-orbit, and hence by Northcott's theorem [18], there are points of arbitrarily large height in $\mathrm{Orb}_S(P)$. We choose a point

$$Q' \in \mathrm{Orb}_S(P) \quad \text{satisfying} \quad h(Q') > C_{19}(S'),$$

where $C_{19}(S')$ is the constant associated to the set S' appearing in Lemma 18. In particular, it follows from Lemma 18(b) and Northcott's theorem that $\mathrm{Orb}_{S'}(Q')$ must be infinite. Then Proposition 19 implies that $M_{S'}$ is free and that there is a point

$$Q \in \mathrm{Orb}_{S'}(Q') \subseteq \mathrm{Orb}_S(P)$$

that is strongly S'-wandering. Applying Theorem 7 to the set S' and the point Q, we deduce that

$$\sum_{\mathfrak{p}\in\mathrm{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}\cdot m_{\mathfrak{p}}(S, P)^{\epsilon}} \le \sum_{\mathfrak{p}\in\mathrm{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p}\cdot m_{\mathfrak{p}}(S', Q)^{\epsilon}} \le C_{17}\epsilon^{-1}$$

for some constant C_{17} depending on S, Q (and so P) and K. For this last conclusion, we have also used the fact that

$$m_{\mathfrak{p}}(S', Q) \le m_{\mathfrak{p}}(S, P),$$

which is immediate from the inclusion $\mathrm{Orb}_{S'}(Q) \subseteq \mathrm{Orb}_S(P)$. $\square$

Proof of Theorem 10. We recall that Rat_d denotes the space of rational maps of degree d. Then it follows from Theorems 1.1–1.4 in [19] that if $d_1, d_2 \ge 4$, then the set

$$\mathcal{V}_{d_1,d_2} := \left\{ (f_1, f_2) \in \mathrm{Rat}_{d_1} \times \mathrm{Rat}_{d_2} : \begin{array}{l} f_1 \text{ and } f_2 \text{ are critically simple} \\ \text{and critically separated} \end{array} \right\}$$

is Zariski dense in $\mathrm{Rat}_{d_1} \times \mathrm{Rat}_{d_2}$. Then for any $d_3, \dots, d_r \ge 2$, the set

$$\mathcal{U}(d_1, \dots, d_r) := \mathcal{V}_{d_1,d_2} \times \mathrm{Rat}_{d_3} \times \cdots \times \mathrm{Rat}_{d_r}$$

is Zariski dense in $\mathrm{Rat}_{d_1} \times \cdots \times \mathrm{Rat}_{d_r}$, and Theorem 15 gives us that the desired inequality (2) for every S generated by a set of maps

$$(f_1, \dots, f_r) \in \mathcal{U}(d_1, \dots, d_r). \qquad \square$$

We conclude with a variant of Theorem 15 in which the maps are polynomials. We start with a definition.

Definition 20. A polynomial $f(x) \in \overline{\mathbb{Q}}[x]$ is *power-like* if there exist polynomials $R(x), C(x), L(x) \in \overline{\mathbb{Q}}[x]$ such that

$$f = R \circ C \circ L, \deg(L) = 1, \deg(C) \ge 2,$$

$$C(x) = \text{a power map or a Chebyshev polynomial}.$$

Theorem 21. *Let $K/\mathbb{Q}$ be a number field, let S be a set of endomorphisms of $\mathbb{P}^1$ defined over K, and let $P \in \mathbb{P}^1(K)$ be a point such that $\operatorname{Orb}_S(P)$ is infinite. Suppose further that S contains polynomials $f_1(x), f_2(x) \in K[x]$ having the following properties*:

(1) *Neither f_1 nor f_2 is power-like*; *see Definition 20.*

(2) *For all $g \in \overline{\mathbb{Q}}[x]$ satisfying $\deg(g) \geq 2$, we have*

$$f_1 \neq f_2 \circ g \quad \text{and} \quad f_2 \neq f_1 \circ g.$$

Then there is a constant $C_{22} = C_{22}(K, S, P)$ such that for all $\epsilon > 0$,

$$\sum_{\mathfrak{p} \in \operatorname{Spec}(R_K)} \frac{\log \mathsf{N}\mathfrak{p}}{\mathsf{N}\mathfrak{p} \cdot m_{\mathfrak{p}}(S, P)^{\epsilon}} \leq C_{22} \cdot \epsilon^{-1}.$$

The proof of Theorem 21 is similar to the proof of Theorem 15, except that we use [10; 23] instead of [13; 19]. As a first step, we need the following result, which is a polynomial analogue of Proposition 19.

Proposition 22. *Let f_1 and f_2 be polynomials satisfying the hypotheses of Theorem 21, and let $S = \{f_1, f_2\}$. Then the following statements hold*:

(a) *The semigroup M_S is free.*

(b) *Let $P \in \mathbb{P}^1(\overline{\mathbb{Q}})$ be a point whose S-orbit $\operatorname{Orb}_S(P)$ is infinite. Then there exists a strongly S-wandering point $Q \in \operatorname{Orb}_S(P)$.*

Proof. (a) See [13, Proposition 4.5].

(b) The proof is very similar to the proof of Proposition 19, so we just give a brief sketch, highlighting the differences. We note that we have picked a coordinate function x on $\mathbb{P}^1$. We let $\infty \in \mathbb{P}^1$ be the pole of x and let $\mathbb{A}^1 = \mathbb{P}^1 \smallsetminus \{\infty\}$. Replacing P with another point in $\operatorname{Orb}_S(P)$ if necessary, we may assume that $P \neq \infty$ is not the point at infinity. We choose a set $\mathfrak{S}$ of primes of K so that the ring of $\mathfrak{S}$-integers $R_{K,\mathfrak{S}}$ satisfies

$$P \in \mathbb{A}^1(R_{K,\mathfrak{S}}) \quad \text{and} \quad f_1(x), f_2(x) \in R_{K,\mathfrak{S}}[x].$$

We use the map $f_1 \times f_2 : \mathbb{A}^2 \to \mathbb{A}^2$ to define three affine curves,

$$\Gamma_1 := (f_1 \times f_1)^{-1}(\Delta), \quad \Gamma_2 := (f_2 \times f_2)^{-1}(\Delta), \quad \Gamma_{1,2} := (f_1 \times f_2)^{-1}(\Delta).$$

Then [13, Proposition 4.5], itself a consequence of the main results of [10; 23], tells us that these are geometrically irreducible curves of geometric genus at least 1. (This is where we use the assumptions (1) and (2) of Theorem 21 on f_1 and f_2.)

The Siegel–Mahler theorem for integral points on affine curves [14, Theorem D.9.1] then implies that

$$\Gamma_1(R_{K,\mathfrak{S}}),\quad \Gamma_2(R_{K,\mathfrak{S}}),\quad \text{and}\quad \Gamma_{1,2}(R_{K,\mathfrak{S}})\quad \text{are finite sets,}$$

and hence that

$$\Sigma := \Gamma_{1,2}(R_{K,\mathfrak{S}}) \cup (\Gamma_1 \smallsetminus \Delta)(R_{K,\mathfrak{S}}) \cup (\Gamma_2 \smallsetminus \Delta)(R_{K,\mathfrak{S}})$$

is finite.

The remainder of the proof of Proposition 22 follows the proof of Proposition 19, starting with the three possibilities described in (10). □

Proof of Theorem 21. The proof of Theorem 21 is identical to that of Theorem 15. We first use Lemma 18, Proposition 22, and the fact that $\mathrm{Orb}_S(P)$ is infinite to find a point $Q \in \mathrm{Orb}_S(P)$ that is strongly wandering for $S' = \{f_1, f_2\}$. We then apply Theorem 7 to the point Q and the set $S' = \{f_1, f_2\}$ to deduce the desired result for P and S. □

Acknowledgements

The authors thank the referee for their detailed and helpful report.

References

[1] A. Akbary and D. Ghioca, “Periods of orbits modulo primes”, *J. Number Theory* **129**:11 (2009), 2831–2842. MR Zbl

[2] M. Baker and L. DeMarco, “Preperiodic points and unlikely intersections”, *Duke Math. J.* **159**:1 (2011), 1–29. MR Zbl

[3] J. P. Bell, K. Huang, W. Peng, and T. J. Tucker, “A Tits alternative for rational functions”, preprint, 2022. arXiv 2103.09994

[4] J. Bell, W. Hindes, and Z. Xiao, “Counting points by height in semigroup orbits”, preprint, 2023. arXiv 2305.04960

[5] R. Benedetto, P. Ingram, R. Jones, M. Manes, J. H. Silverman, and T. J. Tucker, “Current trends and open problems in arithmetic dynamics”, *Bull. Amer. Math. Soc.* (*N.S.*) **56**:4 (2019), 611–685. MR Zbl

[6] M.-C. Chang, “On periods modulo p in arithmetic dynamics”, *C. R. Math. Acad. Sci. Paris* **353**:4 (2015), 283–285. MR Zbl

[7] M.-C. Chang, C. D’Andrea, A. Ostafe, I. E. Shparlinski, and M. Sombra, “Orbits of polynomial dynamical systems modulo primes”, *Proc. Amer. Math. Soc.* **146**:5 (2018), 2015–2025. MR

[8] L. DeMarco, H. Krieger, and H. Ye, “Uniform Manin–Mumford for a family of genus 2 curves”, *Ann. of Math.* (2) **191**:3 (2020), 949–1001. MR Zbl

[9] G. Faltings, “Endlichkeitssätze für abelsche Varietäten über Zahlkörpern”, *Invent. Math.* **73**:3 (1983), 349–366. MR Zbl

[10] D. Ghioca, T. J. Tucker, and M. E. Zieve, “Linear relations between polynomial orbits”, *Duke Math. J.* **161**:7 (2012), 1379–1410. MR Zbl

[11] D. Ghioca, L.-C. Hsia, and T. J. Tucker, "Preperiodic points for families of rational maps", *Proc. Lond. Math. Soc.* (3) **110**:2 (2015), 395–427. MR Zbl

[12] V. O. Healey and W. Hindes, "Stochastic canonical heights", *J. Number Theory* **201** (2019), 228–256. MR Zbl

[13] W. Hindes, "Counting points of bounded height in monoid orbits", *Math. Z.* **301**:4 (2022), 3395–3416. MR Zbl

[14] M. Hindry and J. H. Silverman, *Diophantine geometry: an introduction*, Graduate Texts in Mathematics **201**, Springer, 2000. MR Zbl

[15] K. Huang, "Generalized greatest common divisors for orbits under rational functions", *Monatsh. Math.* **191**:1 (2020), 103–123. MR Zbl

[16] S. Kawaguchi and J. H. Silverman, "Dynamics of projective morphisms having identical canonical heights", *Proc. Lond. Math. Soc.* (3) **95**:2 (2007), 519–544. MR Zbl

[17] M. R. Murty, M. Rosen, and J. H. Silverman, "Variations on a theme of Romanoff", *Internat. J. Math.* **7**:3 (1996), 373–391. MR Zbl

[18] D. G. Northcott, "Periodic points on an algebraic variety", *Ann. of Math.* (2) **51** (1950), 167–177. MR Zbl

[19] F. Pakovich, "Algebraic curves $P(x) - Q(y) = 0$ and functional equations", *Complex Var. Elliptic Equ.* **56**:1-4 (2011), 199–213. MR Zbl

[20] J. H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics **241**, Springer, 2007. MR Zbl

[21] J. H. Silverman, "Variation of periods modulo p in arithmetic dynamics", *New York J. Math.* **14** (2008), 601–616. MR Zbl

[22] J. H. Silverman, *Moduli spaces and arithmetic dynamics*, CRM Monograph Series **30**, American Mathematical Society, Providence, RI, 2012. MR Zbl

[23] U. Zannier, "Integral points on curves $\frac{f(X)-f(Y)}{X-Y}$", *Math. Z.* **301**:4 (2022), 3609–3616. MR Zbl

Received March 31, 2023. Revised September 1, 2023.

WADE HINDES
DEPARTMENT OF MATHEMATICS
TEXAS STATE UNIVERSITY
SAN MARCOS, TX
UNITED STATES

wmh33@txstate.edu

JOSEPH H. SILVERMAN
DEPARTMENT OF MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RI
UNITED STATES

joseph_silverman@brown.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.299

TROPICAL LAGRANGIAN MULTISECTIONS AND TORIC VECTOR BUNDLES

YAT-HIN SUEN

We introduce the notion of tropical Lagrangian multisections over a fan and study its relation with toric vector bundles. We also introduce a "SYZ-type" construction for toric vector bundles which gives a reinterpretation of Kaneyama's linear algebra data. In dimension 2, this "mirror-symmetric" approach provides us a pure combinatorial condition for checking which rank 2 tropical Lagrangian multisections arise from toric vector bundles.

1. Introduction

Toric geometry is an interaction between algebraic geometry and combinatorics. Difficult problems in algebraic geometry can usually be simplified in the toric world. Toric geometry also plays a key role in the current development of mirror symmetry. It provides a huge source of computable examples for mathematicians and physicists to understand mirror symmetry [1; 2; 4; 5; 6; 8; 12; 13; 14]. The famous Gross–Siebert program [18; 19; 20] applies toric degenerations to solve the reconstruction problem in mirror symmetry, which is often referred to as the algebro-geometric SYZ program [27].

In this paper, we study the combinatorics of toric vector bundles. The study of toric vector bundles can be dated back to Kaneyama's classification [21] using linear algebra data and also Klyachko's classification [23] using filtrations indexed by rays in the fan. Payne [25; 26] studied toric vector bundles and their moduli in terms of piecewise linear functions defined on cone complexes. Motivated by the work of Payne, the notion of tropical Lagrangian multisections was first introduced by the author of this paper in [28] and generalized to arbitrary 2-dimensional integral affine manifolds with singularities in a joint work with Chan and Ma [9].

We begin by recalling some elementary facts about toric varieties and toric vector bundles in Section 2. In Section 3, we introduce the notion of tropical Lagrangian multisections over a complete fan Σ on $N_{\mathbb{R}} \cong \mathbb{R}^n$. A tropical Lagrangian multisection $\mathbb{L}$ over Σ is a branched covering map $\pi : (L, \Sigma_L, \mu) \to (N_{\mathbb{R}}, \Sigma)$ of connected cone

MSC2020: 14M25, 14Txx.
Keywords: mirror symmetry, toric geometry, toric vector bundles, tropical geometry, symplectic geometry, Lagrangian multisections.

complexes[1] ($\mu : \Sigma_L \to \mathbb{Z}_{>0}$ is the weight or multiplicity map) together with a piecewise linear function $\varphi : L \to \mathbb{R}$. We will introduce three more concepts, namely, *combinatorial union*, *combinatorial indecomposability* and *combinatorial equivalence*. These concepts allow us to break down a tropical Lagrangian multisection into "indecomposable" components. Moreover, these components enjoy some nice properties, for instance, the ramification locus of a combinatorially indecomposable tropical Lagrangian multisection lies in the codimension 2 strata of (L, Σ_L) (Proposition 3.23). Such indecomposability is also related to indecomposability of toric vector bundles as we will see in Section 4 (Theorem 4.7).

In Section 3A, we follow [26] to associate a tropical Lagrangian multisection $\mathbb{L}_{\mathcal{E}}$ to a toric vector bundle $\mathcal{E}$ on X_Σ. Section 4 will be devoted to the converse. Namely, given a tropical Lagrangian multisection $\mathbb{L}$ over a complete fan Σ, we would like to construct a toric vector bundle on X_Σ. We call this the *reconstruction problem*. One should not expect $\mathbb{L}$ to completely determine a toric vector bundle due to its discrete nature, and Payne has already proved in [26] that $\mathbb{L}_{\mathcal{E}}$ only determines the total equivariant Chern class of $\mathcal{E}$. Therefore, we need to introduce some continuous data (Definition 4.1), which are the linear algebra data given by Kaneyama [21]. The set of all such data on $\mathbb{L}$ modulo gauge equivalence will be denoted by $\mathcal{K}(\mathbb{L})$.

A fundamental question that this paper would like to answer is: When is $\mathcal{K}(\mathbb{L}) \neq \varnothing$? In Section 4B, we give a "SYZ-mirror-symmetric" approach to solve this problem. First of all, SYZ mirror symmetry [27] suggests that if a symplectic manifold admits a Lagrangian torus fibration, its complex mirror is obtained by taking the dual torus fibration. Furthermore, the SYZ program also suggests that holomorphic vector bundles are mirror to Lagrangian multisections. Given a Lagrangian multisection whose underlying covering map is unbranched, its SYZ transform was defined in [7; 24]. However, the covering map can be branched over the base of the SYZ fibration. The SYZ program then suggests we first construct the *semiflat bundle*, which is obtained by the usual SYZ transform with the branch locus removed. However, the semiflat bundle would receive nontrivial monodromies around those fibers above the branch locus and thus cannot be extended to the whole mirror space. To perform extension, we need to cancel these monodromies by remembering the ramification locus. The SYZ program suggests that the ramification locus should be remembered by the holomorphic disks bounded by the multisection and certain SYZ fibers. The exponentiation of the generating function of these holomorphic disks is the so-called *wall-crossing automorphism*. A good local example was given by Fukaya [15, Example 4.4]. Moreover, he also pointed out in [15, Section 6.4] that, when the rank is 2, the semiflat bundle needs to be twisted by a nontrivial local system in order to carry out the monodromy cancellation process.

[1] In [28], we assume the domain L is a topological manifold. We extend the definition here by allowing L to be a cone complex, which is not necessarily a manifold.

Going back to our tropical world, we restrict our attention to combinatorially indecomposable tropical Lagrangian multisections. This assumption implies the ramification locus is contained in the codimension 2 stratum $L^{(n-2)}$ of (L, Σ_L) (Proposition 3.23). Following the idea of the SYZ program and Fukaya's proposal, the reconstruction program should consist of two steps. The first step is to equip $L \backslash L^{(n-2)}$ with a suitable $\mathbb{C}^\times$-local system $\mathcal{L}$. Then we construct in Section 4B1 the *semiflat mirror bundle* $\mathcal{E}^{\text{sf}}(\mathbb{L}, \mathcal{L})$ of $(\mathbb{L}, \mathcal{L})$, which is a rank r toric vector bundle defined on the 1-skeleton

$$X_\Sigma^{(1)} := \bigcup_{\tau \in \Sigma(n-1)} X_\tau$$

of X_Σ. In general, the semiflat mirror bundle cannot be extended to X_Σ due to the presence of monodromies of $\pi : L \to N_\mathbb{R}$ around the branch locus $S \subset N_\mathbb{R}$. In order to cancel these monodromies, we will introduce a set of local automorphisms $\Theta := \{\Theta_\tau(\omega')\}_{\tau \in \Sigma(n-1), \omega' \subset S}$ in Section 4B2 to correct the transition maps of $\mathcal{E}^{\text{sf}}(\mathbb{L}, \mathcal{L})$ so that it can be extended to X_Σ. If there exists a $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \backslash L^{(n-2)}$ and a collection of factors Θ that satisfy the *consistency condition* (Definition 4.15), the tropical Lagrangian multisection is called *unobstructed* (Definition 4.17 and see Remark 4.18 for the terminology). Being unobstructed allows us to define a 1-cocycle $\{G_{\sigma_1\sigma_2}\}_{\sigma_1, \sigma_2 \in \Sigma(n)}$ and gives a toric vector bundle $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta)$ over X_Σ. It turns out that all Kaneyama data arise from this construction.

Theorem 4.21. *Suppose $\mathbb{L}$ is combinatorially indecomposable and admits a Kaneyama data $\boldsymbol{g}$. Then there exists a $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \backslash L^{(n-2)}$ and consistent Θ such that $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta) = \mathcal{E}(\mathbb{L}, \boldsymbol{g})$.*

The factors $\{\Theta_\tau(\omega')\}$ should be thought of as wall-crossing automorphisms as described above, which are responsible for Maslov index 0 holomorphic disks bounded by a Lagrangian multisection and certain fibers of the torus fibration $T^*N_\mathbb{R}/M \to N_\mathbb{R}$. Hence our reconstruction program can be regarded as a "tropical SYZ transform".

In the last section, Section 5, we apply our "SYZ construction" to study the unobstructedness of combinatorially indecomposable tropical Lagrangian multisections of rank 2 over a complete fan on $N_\mathbb{R} \cong \mathbb{R}^2$. First of all, not all such objects are unobstructed (Example 5.1). Therefore, we need extra conditions to guarantee unobstructedness. We will define a *slope condition* (Definition 5.8), which is completely determined by the combinatorics of the piecewise linear function $\varphi : L \to \mathbb{R}$ of $\mathbb{L}$. It turns out this combinatorial condition completely determines the obstruction of $\mathbb{L}$.

Theorem 5.9. *A combinatorially indecomposable rank 2 tropical Lagrangian multisection $\mathbb{L}$ over a 2-dimensional complete fan Σ is unobstructed if and only if it satisfies the slope condition.*

From the proof of Theorem 5.9, we can deduce an interesting inequality, bounding the dimension of moduli spaces of toric vector bundles with fixed equivariant Chern classes by the number of rays in Σ.

Corollary 5.10. *If $\mathbb{L}$ is a combinatorially indecomposable rank 2 tropical Lagrangian multisection, then we have the inequality* $\dim_{\mathbb{C}}(\mathcal{K}(\mathbb{L})) \leq \#\Sigma(1) - 1$.

2. Toric varieties and toric vector bundles

We first recall some basics in toric geometry. Standard references are [10; 11; 17]. Throughout, we denote by N a rank n lattice and $M := \mathrm{Hom}_{\mathbb{Z}}(N, \mathbb{Z})$ the dual lattice. We also set $N_{\mathbb{R}} := N \otimes_{\mathbb{Z}} \mathbb{R}$ and $M_{\mathbb{R}} := M \otimes_{\mathbb{Z}} \mathbb{R}$. A *fan* Σ in $N_{\mathbb{R}}$ is a collection of rational strictly convex cones in $N_{\mathbb{R}}$ such that

(1) if $\sigma \in \Sigma$ and $\tau \subset \sigma$ is a face, then $\tau \in \Sigma$ and

(2) if $\sigma_1, \sigma_2 \in \Sigma$, then $\sigma_1 \cap \sigma_2 \in \Sigma$.

Denote by $\Sigma(k)$ the collection of all k-dimensional cones in Σ. For each cone $\sigma \in \Sigma$, one can associate the corresponding dual cone $\sigma^\vee$ in $M_{\mathbb{R}}$, which is defined by

$$\sigma^\vee := \{x \in M_{\mathbb{R}} : \langle x, \xi \rangle \geq 0 \ \ \forall \xi \in \sigma\}.$$

It is also a strictly convex rational cone. For $\tau \subset \sigma$, we have $\sigma^\vee \subset \tau^\vee$. Define

$$U(\sigma) := \mathrm{Spec}(\mathbb{C}[\sigma^\vee \cap M]).$$

There is a $(\mathbb{C}^\times)^n$-action on $U(\sigma)$, given by

$$\lambda \cdot z^m := \lambda^m z^m,$$

for $m \in \sigma^\vee \cap M$. For $\tau \subset \sigma$, we have an open embedding $U(\tau) \to U(\sigma)$. The toric variety X_Σ associated to Σ is defined to be the direct limit

$$X_\Sigma := \varinjlim U(\sigma).$$

The $(\mathbb{C}^\times)^n$-actions on affine charts agree and so induce a $(\mathbb{C}^\times)^n$-action on X_Σ.

Definition 2.1. Let X_Σ be an n-dimensional toric variety. A vector bundle $\mathcal{E}$ on X_Σ is called *toric* if the $(\mathbb{C}^\times)^n$-action on X_Σ lifts to an action on $\mathcal{E}$ which is linear on fibers. Equivalently (see [21]), for each $\lambda \in (\mathbb{C}^\times)^n$, there is a vector bundle isomorphism $\lambda^*\mathcal{E} \cong \mathcal{E}$ covering the identity of X_Σ.

Given a toric vector bundle $\mathcal{E}$ on X_Σ, the $(\mathbb{C}^\times)^n$-action constrains the transition maps of $\mathcal{E}$. Let $G_\sigma : \mathcal{E}|_{U(\sigma)} \to U(\sigma) \times \mathbb{C}^r$ be an equivariant trivialization and

$$G_{\sigma_1\sigma_2} := G_{\sigma_2} \circ G_{\sigma_1}^{-1} : U(\sigma_1 \cap \sigma_2) \times \mathbb{C}^r \to U(\sigma_1 \cap \sigma_2) \times \mathbb{C}^r$$

be the transition map from the affine chart $U(\sigma_1)$ to the chart $U(\sigma_2)$. We can always choose the trivialization $G_\sigma : \mathcal{E}|_{U(\sigma)} \to U(\sigma) \times \mathbb{C}^r$ so that $(\mathbb{C}^\times)^n$ acts diagonally on fibers, that is, the action on $\mathbb{C}[\sigma^\vee \cap M] \otimes_{\mathbb{C}} \mathbb{C}[t_1, \dots, t_r]$ is of form

$$\lambda \cdot (z^m, t_1, \dots, t_r) = (\lambda^m z^m, \lambda^{m^{(1)}(\sigma)} t_1, \dots, \lambda^{m^{(r)}(\sigma)} t_r)$$

for some $m^{(1)}(\sigma), \dots, m^{(r)}(\sigma) \in M$. Since this action extends to X_Σ, we must have

$$G^{(\alpha\beta)}_{\sigma_1\sigma_2}(z) = g^{(\alpha\beta)}_{\sigma_1\sigma_2} z^{m^{(\alpha)}(\sigma_1) - m^{(\beta)}(\sigma_2)}$$

for some $g^{(\alpha\beta)}_{\sigma_1\sigma_2} \in \mathbb{C}$ so that $g^{(\alpha\beta)}_{\sigma_1\sigma_2} \neq 0$ only if $m^{(\alpha)}(\sigma_1) - m^{(\beta)}(\sigma_2) \in (\sigma_1 \cap \sigma_2)^\vee \cap M$.

3. Tropical Lagrangian multisections

In this section, we introduce the notion of tropical Lagrangian multisections. We begin by reviewing some basics about cone complexes. We follow [26] with some small notational changes.

Definition 3.1 [26, Definition 2.1]. A *cone complex* consists of a topological space X together with a finite collection Σ of closed subsets of X and for each $\sigma \in \Sigma$, a finitely generated subgroup $M(\sigma)$ of the group of continuous functions on σ, satisfying the following conditions:

(1) The natural map $\phi_\sigma : \sigma \to (M(\sigma) \otimes_{\mathbb{Z}} \mathbb{R})^\vee$ given by

$$x \mapsto (u \mapsto u(x))$$

maps σ homeomorphically onto a convex rational polyhedral cone.

(2) The preimage of any face of $\phi_\sigma(\sigma)$ is an element of Σ and

$$M(\tau) = \{m|_\tau \mid m \in M(\sigma)\}.$$

(3) The topological space X admits the decomposition

$$X = \bigsqcup_{\sigma \in \Sigma} \mathrm{Int}(\sigma),$$

where $\mathrm{Int}(\sigma)$ denotes the relative interior of σ.

A cone complex (X, Σ) is said to be *connected* if the topological space X is connected. The space of *piecewise linear functions* on (X, Σ) is defined to be

$$PL(X, \Sigma) := \{\varphi : X \to \mathbb{R} \mid \varphi|_\sigma \in M(\sigma) \ \ \forall \sigma \in \Sigma\}.$$

Remark 3.2. The connected components of X are parametrized by minimal cones in Σ. See [26, Remark 2.6].

Definition 3.3 [26, Definition 2.9]. A *morphism* of cone complexes $f:(X', \Sigma_{X'}) \to (X, \Sigma_X)$ is a continuous map $f: X' \to X$ such that for any $\sigma' \in \Sigma_{X'}$, there exists $\sigma \in \Sigma$ such that $f(\sigma') \subset \sigma$ and $f^*M(\sigma) \subset M(\sigma')$.

Definition 3.4 [26, Definition 2.16]. A *weighted cone complex* consists of a cone complex (X, Σ) together with a function $\mu: X \to \mathbb{Z}_{>0}$ such that for any $\sigma \in \Sigma$, $\mu|_{\mathrm{Int}(\sigma)}$ is constant. We simply write $\mu(\sigma)$ for $\mu|_{\mathrm{Int}(\sigma)}$.

If $(X', \Sigma_{X'})$ is weighted by μ, for a surjective morphism $f:(X', \Sigma_{X'}) \to (X, \Sigma_X)$, we can define $\mathrm{Tr}_f(\mu): X \to \mathbb{Z}_{>0}$ by

$$\mathrm{Tr}_f(\mu)(x) := \sum_{x' \in f^{-1}(x)} \mu(x'),$$

called the *trace of* μ *by* f.

Definition 3.5 [26, Definition 2.17]. Let (B, Σ) be a connected cone complex and (L, Σ_L, μ) be a connected weighted cone complex. A *branched covering map* $\pi:(L, \Sigma_L, \mu) \to (B, \Sigma)$ is a surjective morphism of cone complexes such that

(1) for each $\sigma' \in \Sigma_L$, π maps σ homeomorphically to $\pi(\sigma) \in \Sigma$,

(2) for any connected open set $U \subset B$ and connected $V \subset \pi^{-1}(U)$, the function $\mathrm{Tr}_{\pi|_V}(\mu): U \to \mathbb{Z}_{>0}$ is constant.

The morphism $\pi:(L, \Sigma_L, \mu) \to (B, \Sigma)$ is said to be *ramified along* $\tau' \in \Sigma_L$ if $\mu(\tau') > 1$. The number $\mathrm{Tr}_\pi(\mu)$ is called the degree of $\pi:(L, \Sigma_L, \mu) \to (B, \Sigma)$. The subset

$$S' := S'(\mathbb{L}) := \bigcup_{\tau' \in \Sigma_L : \mu(\tau') > 1} \tau' \subset L$$

is called the *ramification locus* of π and $S := S(\mathbb{L}) := \pi(S')$ is called the *branch locus* of π.

Definition 3.6. Let $\pi_1:(L_1, \Sigma_{L_1}, \mu_1) \to (B, \Sigma)$, $\pi_2:(L_2, \Sigma_{L_2}, \mu_2) \to (B, \Sigma)$ be branched covering maps of the same degree. We write $\pi_1 \leq \pi_2$ if there exists a surjective morphism of cone complexes $f:(L_2, \Sigma_{L_2}) \to (L_1, \Sigma_{L_1})$ such that $\pi_1 \circ f = \pi_2$ and $\mathrm{Tr}_f(\mu_2) = \mu_1$.

Definition 3.7 [26, Definition 2.26]. A branched covering map $\pi:(L, \Sigma_L, \mu) \to (B, \Sigma)$ is called *maximal* if it is maximal with respective to the partial ordering given in Definition 3.6.

Given a cone complex (L, Σ_L), we define

$$L^{(n-k)} := \bigcup_{\tau' \in \Sigma_L : \mathrm{codim}(\tau') = k} \tau' \subset L,$$

the codimension k stratum of (L, Σ_L). Payne showed in [26, Proposition 2.30] that if Σ is a complete fan in $N_{\mathbb{R}}$, the ramification locus of any maximal branched covering map $\pi : (L, \Sigma_L, \mu) \to (N_{\mathbb{R}}, \Sigma)$ lies in the codimension 2 stratum $L^{(n-2)}$ of (L, Σ_L). Now we focus on $B = N_{\mathbb{R}} \cong \mathbb{R}^n$ and Σ is a complete fan on $N_{\mathbb{R}}$. In this case, B carries a natural affine structure and Σ turns (B, Σ) into a cone complex. If $\pi : (L, \Sigma_L, \mu) \to (B, \Sigma)$ is a branched covering map, then for any $\sigma' \in \Sigma_L(n)$, we have $\pi^* M = \pi^* M(\sigma) = M(\sigma')$ as $\pi|_{\sigma'} : \sigma' \to \sigma$ is an isomorphism. Hence we can identity $M(\sigma')$ with M via π^* naturally. We can then define

$$\mathrm{Lin}(L) := \{f \in C^0(L, \mathbb{R}) : \exists m \in M \text{ such that } f|_{\sigma'} = m \;\; \forall \sigma' \in \Sigma_L\},$$

to be the space of linear function on L. It is clear that $\mathrm{Lin}(L) \subset PL(L, \Sigma_L)$. Moreover, as L is assumed to be connected, it is clear that $\mathrm{Lin}(L) = \mathrm{Lin}(B) = M$.

Definition 3.8. Let Σ be a complete fan on $N_{\mathbb{R}}$. A *tropical Lagrangian multisection of rank r over* Σ is a quintuple $\mathbb{L} := (L, \Sigma_L, \mu, \pi, \varphi)$, where

(1) (L, Σ_L) is a connected cone complex weighted by μ,

(2) $\pi : (L, \Sigma_L, \mu) \to (N_{\mathbb{R}}, \Sigma)$ is a branched covering map such that $\mathrm{Tr}_\pi(\mu) = r$,

(3) φ is a piecewise linear function on (L, Σ_L).

The number r is called the *rank* of $\mathbb{L}$ and is denoted by $\mathrm{rk}(\mathbb{L})$. The underlying branched covering map of $\mathbb{L}$ is denoted by $\underline{\mathbb{L}}$. A tropical Lagrangian multisection $\mathbb{L}$ is said to be *maximal* if $\underline{\mathbb{L}}$ is maximal.

Remark 3.9. In [28], the author provided a definition of tropical Lagrangian multisections over integral affine manifolds with singularities whose domain of the branched covering map is a topological manifold. While in [9], the authors gave a definition of tropical Lagrangian multisections over 2-dimensional integral affine manifolds with singularities *equipped with polyhedral decomposition*, where they also assumed the domain is also a topological manifold equipped with a polyhedral decomposition that is compatible with the covering map. Of course, if we restrict our attention to the case where the affine manifold is $\mathbb{R}^2$ with polyhedral decomposition being a fan Σ, Definition 3.8 extends Definition 3.6 in [9] because we don't assume L is a topological manifold here.

Remark 3.10. In [2], Abouzaid used the terminology "tropical Lagrangian section" to stand for an honest Lagrangian section of the torus fibration $\mathrm{Log} : (\mathbb{C}^\times)^n \to \mathbb{R}^n$. The term "tropical" in this paper stands for a combinatorial/discrete replacement for Lagrangian multisections, which are supposed to be mirror to vector bundles on X_Σ. However, it is not hard to show that a tropical Lagrangian section ($r = 1$) in our combinatorial sense always produces a tropical Lagrangian section in the sense of Abouzaid by smoothing the piecewise linear function $\varphi : |\Sigma| \to \mathbb{R}$ suitably. Thus our definition is somehow a generalization of Abouzaid's one. Nevertheless, we apologize for any possible confusion with the use of the terminology here.

Definition 3.11. Let $\mathbb{L}_1, \mathbb{L}_2$ be tropical Lagrangian multisections of the same rank. We write $\mathbb{L}_1 \leq \mathbb{L}_2$ if $\underline{\mathbb{L}}_1 \leq \underline{\mathbb{L}}_2$ via some f such that $f^*\varphi_1 = \varphi_2$.

Definition 3.12. Let $\mathbb{L}_1, \mathbb{L}_2$ be tropical Lagrangian multisections over a fan Σ. We write $\mathbb{L}_2 \sim_c \mathbb{L}_1$ if $\mathrm{rk}(\mathbb{L}_1) = \mathrm{rk}(\mathbb{L}_2)$ and there exists a tropical Lagrangian multisection $\mathbb{L}$ over Σ such that $\mathbb{L} \leq \mathbb{L}_i$ for all $i = 1, 2$. We say $\mathbb{L}_1$ is *combinatorially equivalent* to $\mathbb{L}_2$ if there exists a sequence of tropical Lagrangian multisections $\mathbb{L}'_1, \mathbb{L}'_2, \ldots, \mathbb{L}'_k$ such that $\mathbb{L}'_1 = \mathbb{L}_1$, $\mathbb{L}'_k = \mathbb{L}_2$ and $\mathbb{L}'_{i+1} \sim_c \mathbb{L}'_i$ for all $i = 1, \ldots, k-1$.

Remark 3.13. The relation $\sim_c$ is only reflexive and symmetric. The notion of combinatorial equivalence is the transitive closure of $\sim_c$ and hence, an equivalence relation.

Now we define an important class of tropical Lagrangian multisections.

Definition 3.14. A tropical Lagrangian multisection $\mathbb{L} = (L, \Sigma_L, \mu, \pi, \varphi)$ is said to be *k-separated* if it satisfies the following condition: For any $\tau \in \Sigma(k)$ and distinct lifts $\tau^{(\alpha)}, \tau^{(\beta)} \in \Sigma_L(k)$ of τ, we have $\varphi|_{\tau^{(\alpha)}} \neq \varphi|_{\tau^{(\beta)}}$. Note that k-separability implies K-separability for all $K \geq k$. A tropical Lagrangian multisection is said to be *separated* if it is 1-separated.

Remark 3.15. Definition 3.14 holds vacuously for all rank 1 tropical Lagrangian multisections.

We can always "separate" a tropical Lagrangian multisection in the following sense.

Proposition 3.16. *For any tropical Lagrangian multisection $\mathbb{L}$ over Σ, there exists a separated tropical Lagrangian multisection $\mathbb{L}_{\mathrm{sep}}$ over Σ such that $\mathbb{L}_{\mathrm{sep}} \leq \mathbb{L}$. In particular, every tropical Lagrangian multisection is combinatorially equivalent to a separated one.*

Proof. We define a cone complex $(L_{\mathrm{sep}}, \Sigma'_{\mathrm{sep}})$ as follows. Let $\sigma \in \Sigma$. Two lifts $\sigma^{(\alpha)}, \sigma^{(\beta)} \in \Sigma$ of σ are identified if and only if $\varphi|_{\sigma^{(\alpha)}} = \varphi|_{\sigma^{(\beta)}}$. We denote the quotient map $L \to L_{\mathrm{sep}}$ by q. The set of cones is given by

$$\Sigma'_{\mathrm{sep}} := \{q(\sigma') \mid \sigma' \in \Sigma_L\}.$$

The projection map $\pi : L \to N_{\mathbb{R}}$ factors through q and hence descends to a projection $\pi_{\mathrm{sep}} : L_{\mathrm{sep}} \to N_{\mathbb{R}}$. Define $\mu_{\mathrm{sep}} := \mathrm{Tr}_q(\mu)$. It is clear that $\pi_{\mathrm{sep}} : (L_{\mathrm{sep}}, \Sigma'_{\mathrm{sep}}, \mu_{\mathrm{sep}}) \to (N_{\mathbb{R}}, \Sigma)$ is a branched covering map. We define $\varphi_{\mathrm{sep}} : L_{\mathrm{sep}} \to \mathbb{R}$ by

$$\varphi_{\mathrm{sep}}|_{q(\sigma')} = \varphi|_{\sigma'}.$$

It is clear that $\varphi_{\mathrm{sep}}|_{q(\sigma')}$ is independent of the choice of $\sigma' \in \Sigma_L$ and φ_{sep} is continuous. It also follows from construction that $q^*\varphi_{\mathrm{sep}} = \varphi$. Hence $\mathbb{L}_{\mathrm{sep}} \leq \mathbb{L}$. □

Example 3.17. Given a tropical Lagrangian multisection $\mathbb{L}$ as shown in Figure 1, its canonical separation $\mathbb{L}_{\mathrm{sep}}$ is given by gluing $\sigma_0^{(1)}, \sigma_0^{(2)}$ over σ_0.

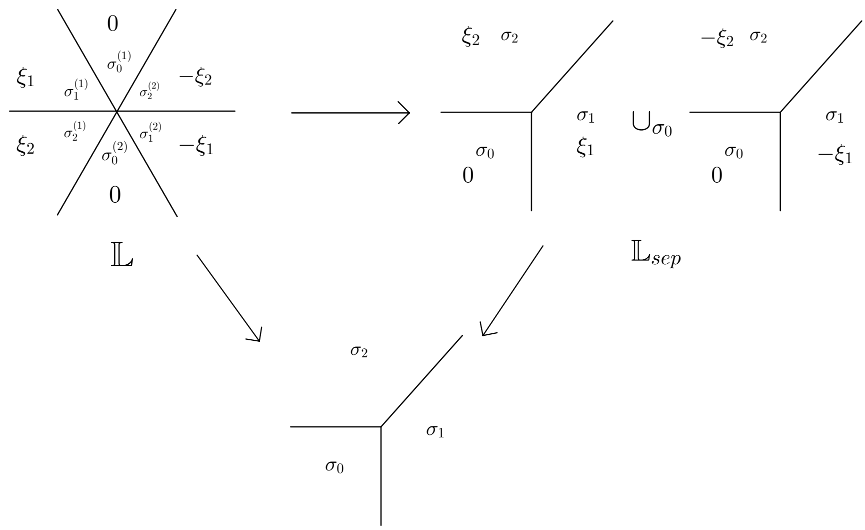


Figure 1. The canonical separation $\mathbb{L}_{\text{sep}}$ of $\mathbb{L}$.

Definition 3.18. The tropical Lagrangian multisection $\mathbb{L}_{\text{sep}}$ constructed in the proof of Proposition 3.16 is called the *canonical separation of* $\mathbb{L}$.

Construction 3.19. There are three natural operations on tropical Lagrangian multisections. As we will see in Proposition 3.26, they correspond to algebraic operations of toric vector bundles.

(1) Given $\mathbb{L} = (L, \Sigma_L, \mu, \pi, \varphi)$, we put $-\mathbb{L} := (L, \Sigma_L, \mu, \pi, -\varphi)$, called the *dual of* $\mathbb{L}$.

(2) Given two tropical Lagrangian multisections $\mathbb{L}_1, \mathbb{L}_2$ with rank r_1, r_2, respectively, we can construct another tropical Lagrangian multisection $\mathbb{L}_1 \cup_c \mathbb{L}_2$ by gluing $\tau' \in \Sigma_1'$, $\tau'' \in \Sigma_2'$ whenever $\pi_1(\tau') = \pi(\tau'') = \tau$ and $\varphi_1|_{\tau'} = \varphi_1|_{\tau''}$. The domain of $\mathbb{L}_1 \cup_c \mathbb{L}_2$ is denoted by $L_1 \cup_c L_2$ and the quotient map $L_1 \sqcup L_2 \to L_1 \cup_c L_2$ is denoted by q. The set of cones is given by

$$\Sigma_1' \cup_c \Sigma_2' := \{q(\sigma') \mid \sigma' \in \Sigma_1' \sqcup \Sigma_2'\}$$

and the multiplicity map is given by

$$(\mu_1 \cup_c \mu_2)(\bar{\sigma}') := \sum_{\sigma_1' \in \Sigma_1' : q(\sigma_1') = \bar{\sigma}'} \mu_1(\sigma_1') + \sum_{\sigma_2' \in \Sigma_2' : q(\sigma_2') = \bar{\sigma}'} \mu_2(\sigma_2').$$

In particular, the rank of $\mathbb{L}_1 \cup_c \mathbb{L}_2$ is $r_1 + r_2$. Finally, the piecewise linear function is given by

$$(\varphi_1 \cup_c \varphi_2)|_{\bar{\sigma}'} = \begin{cases} \varphi_1|_{\sigma_1'} & \text{if } q(\sigma_1') = \bar{\sigma}' \in q(\Sigma_1'), \\ \varphi_2|_{\sigma'} & \text{if } q(\sigma_2') = \bar{\sigma}' \in q(\Sigma_2'). \end{cases}$$

It follows from the definition of q that $\varphi_1 \cup_c \varphi_2$ is well-defined and continuous. We call the tropical Lagrangian multisection $\mathbb{L}_1 \cup_c \mathbb{L}_2$ the *combinatorial union of* $\mathbb{L}_1, \mathbb{L}_2$.

(3) We define the tropical Lagrangian multisection $\mathbb{L}_1 \times_\Sigma \mathbb{L}_2$ of rank $r_1 r_2$ with domain $L_1 \times_{|\Sigma|} L_2$, the set of cones $\Sigma_1' \times_\Sigma \Sigma_2'$, the multiplicity map

$$\sigma_1' \times \sigma_2' \mapsto \mu_1(\sigma_1)\mu_2(\sigma_2)$$

and the projection $\sigma_1 \times_\sigma \sigma_2 \mapsto \sigma$. The piecewise linear function is given by

$$(x_1, x_2) \mapsto \varphi_1(x_1) + \varphi_2(x_2).$$

Finally, denote the canonical separation of $\mathbb{L}_1 \times_{|\Sigma|} \mathbb{L}_2$ by $\mathbb{L}_1 \times_c \mathbb{L}_2$, called the *combinatorial fiber product of* $\mathbb{L}_1, \mathbb{L}_2$.

Note that $\mathbb{L}_1 \cup_c \mathbb{L}_2, \mathbb{L}_1 \times_c \mathbb{L}_2$ are always separated by construction.

Definition 3.20. Let $\mathbb{L}, \mathbb{L}_1, \mathbb{L}_2$ be tropical Lagrangian multisections over Σ. We say $\mathbb{L}$ is *combinatorially decomposable* by $\mathbb{L}_1, \mathbb{L}_2$ if $\mathbb{L}$ is combinatorially equivalent to $\mathbb{L}_1 \cup_c \mathbb{L}_2$. A tropical Lagrangian multisection is said to be *combinatorially indecomposable* if it is not combinatorially decomposable for all pairs of $\mathbb{L}_1, \mathbb{L}_2$.

Every tropical Lagrangian multisection can be combinatorially decomposed into a union of indecomposable ones. However, such decomposition is not unique most of the time.

Example 3.21. Figure 2 shows a combinatorial indecomposable tropical multisection over the fan of $\mathbb{P}^2$. It is also separated as the piecewise linear function has different slopes along distinct lifts of every ray. This tropical Lagrangian multisection is in fact the associated branched covering map of cone complexes of $T_{\mathbb{P}^2}$. See [26].

Example 3.22. Figure 3 shows a combinatorial indecomposable tropical Lagrangian multisection over the fan $\Sigma_{\mathbb{F}_1}$ of the Hirzebruch surface $\mathbb{F}_1$. The notation $\cup_0$ stands for gluing the two cone complexes (both are $(\mathbb{R}^2, \Sigma_{\mathbb{F}_1})$, but decorated by two different piecewise linear functions) on the left at the origin $0 \in N_{\mathbb{R}}$. Again, it is easy to see that this tropical Lagrangian multisection is also separated.

As Example 3.21 suggests, there is a relation between combinatorial indecomposability and separability.

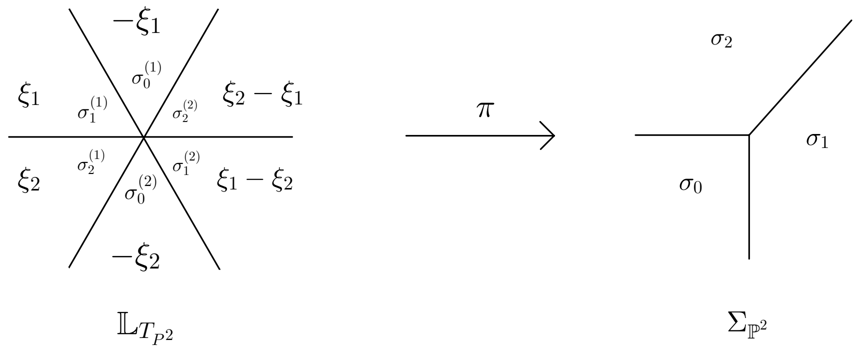


Figure 2. A combinatorially indecomposable tropical Lagrangian multisection over the fan of $\mathbb{P}^2$

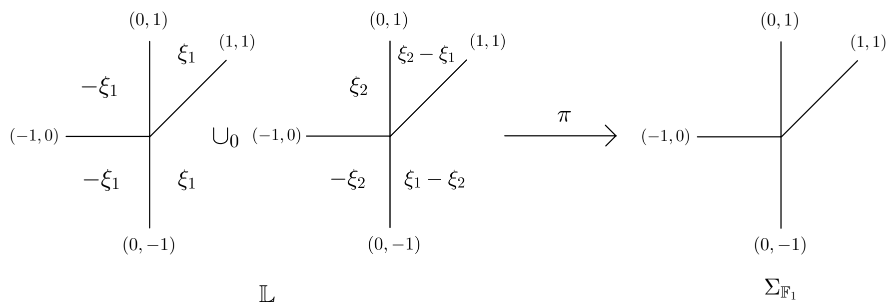


Figure 3. A combinatorially decomposable tropical Lagrangian multisection over the fan of $\mathbb{F}_1$

Proposition 3.23. *Suppose $\mathbb{L}$ is combinatorially indecomposable. Then $\mathbb{L}$ is $(n-1)$-separated and the ramification locus of $\pi : L \to N_{\mathbb{R}}$ lies in the codimension 2 strata $L^{(n-2)}$ of (L, Σ_L). When* $\dim(N_{\mathbb{R}}) = 2$, *the converse is true with the stronger assumption that $\mathbb{L}$ is maximal.*

Proof. We first prove combinatorial indecomposability implies $(n-1)$-separability under the assumption $S'(\mathbb{L}) \subset L^{(n-2)}$. Suppose $\mathbb{L}$ is not $(n-1)$-separated, that is, there exists $\tau \in \Sigma(n-1)$ and distinct lifts $\tau^{(\alpha)}, \tau^{(\beta)} \in \Sigma_L$ such that $\varphi|_{\tau^{(\alpha)}} = \varphi|_{\tau^{(\beta)}}$. Choose a loop $\gamma : [0, 1] \to N_{\mathbb{R}} \backslash S(\mathbb{L})$ so that $\gamma(0) = \gamma(1) \in \mathrm{Int}(\tau)$ and it goes into the interior of each maximal cone once and transverse to the codimension 1 strata. By concatenating γ with itself and using the path lifting lemma, we obtain a lift $\gamma' : [0, 1] \to L \backslash S'(\mathbb{L})$ of γ so that $\gamma'(0) \in \mathrm{Int}(\tau^{(\alpha)})$ and $\gamma'(1) \in \mathrm{Int}(\tau^{(\beta)})$. Let

$$\Sigma_{\gamma'}^{(1)} := \{\sigma' \in \Sigma_L : \mathrm{Int}(\sigma') \cap \gamma' \neq \varnothing\}, \quad \widetilde{L}_{\gamma'}^{(1)} := \bigcup_{\sigma' \in \Sigma_{\gamma'}(n)} \sigma' \subset L.$$

Then there is a cone complex $(L^{(1)}_{\gamma'}, \Sigma_{L^{(1)}_{\gamma'}})$ obtained by gluing $\tau^{(\alpha)}, \tau^{(\beta)}$. Denote the quotient map by $q_1 : \widetilde{L}^{(1)}_{\gamma'} \to L^{(1)}_{\gamma'}$. By considering

$$\Sigma^{(2)}_{\gamma'} = (\Sigma_L \backslash \Sigma^{(1)}_{\gamma'}) \cup \{\tau^{(\alpha)}, \tau^{(\beta)}\}$$

and gluing $\tau^{(\alpha)}, \tau^{(\beta)}$, we obtain another cone complex $(L^{(2)}_{\gamma'}, \Sigma_{L^{(2)}_{\gamma'}})$ and a quotient map q_2. There are two obvious projections

$$\pi^{(i)}_{\gamma'} : L^{(i)}_{\gamma'} \to N_{\mathbb{R}}.$$

We take $\mu_{L^{(i)}_{\gamma'}} := \mathrm{Tr}_{q_i}(\mu|_{\widetilde{L}^{(i)}_{\gamma'}})$ to make $\pi^{(i)}_{\gamma'}$'s into branch covering maps. The function $\varphi|_{\widetilde{L}^{(i)}_{\gamma'}}$ descends to $L^{(i)}_{\gamma'}$ and turn them into two tropical Lagrangian multisections $\mathbb{L}^{(1)}_{\gamma'}$ and $\mathbb{L}^{(2)}_{\gamma'}$. It is then clear that $\mathbb{L} = \mathbb{L}^{(1)}_{\gamma'} \cup_c \mathbb{L}^{(2)}_{\gamma'}$.

Now we handle the general case. Suppose $S'(\mathbb{L}) \not\subset L^{(n-2)}$. Then there is a codimension 1 cone $\tau \in \Sigma(n-1)$ such that $\tau \subset S(\mathbb{L})$. Pass to a cover $f : \mathbb{L}' \to \mathbb{L}$ such that $S'(\mathbb{L}')$ lies in the codimension 2 strata of $(L', \Sigma_{L'})$. Then τ has two distinct lifts $\tau^{(\alpha)}, \tau^{(\beta)} \in \Sigma_{L'}(n-1)$ such that $f^*\varphi|_{\tau^{(\alpha)}} = f^*\varphi|_{\tau^{(\beta)}}$. Hence $\mathbb{L}'$ is not $(n-1)$-separated and hence combinatorially decomposable. But $\mathbb{L}'$ is combinatorially equivalent to $\mathbb{L}$ and so $\mathbb{L}$ is also combinatorially decomposable.

For the converse, note that 1-separability of $\mathbb{L}$ implies any covering morphism of the form $\mathbb{L} \to \mathbb{L}'$ is an isomorphism. Indeed, if $f : \mathbb{L} \to \mathbb{L}'$ is not injective, there exists distinct $\tau^{(\alpha)}, \tau^{(\beta)} \in \Sigma_L(1)$ so that $f(\sigma^{(\alpha)}) = f(\sigma^{(\beta)})$. This implies $\varphi|_{\tau^{(\alpha)}} = \varphi|_{\tau^{(\beta)}}$. As $\tau^{(\alpha)} \neq \tau^{(\beta)}$, this contradicts separability. However, maximality of $\mathbb{L}$ also implies all covering morphism of the form $\mathbb{L}' \to \mathbb{L}$ is an isomorphism. Therefore, if $\mathbb{L}$ is combinatorially decomposable, say by $\mathbb{L}_1, \mathbb{L}_2$, then $\mathbb{L} \cong \mathbb{L}_1 \cup_c \mathbb{L}_2$, which violate maximality. □

Remark 3.24. The converse of Proposition 3.23 is not true without the maximality assumption. For example, let Σ be the fan of $\mathbb{P}^2$ and φ_0, φ_1 be the piecewise linear functions correspond to $\mathcal{O}_{\mathbb{P}^2}, \mathcal{O}_{\mathbb{P}^2}(D_1 + D_2 - 2D_0)$, where D_0, D_1, D_2 are invariant divisors. Then $\mathbb{L}_i := (N_{\mathbb{R}}, \Sigma, 1, \mathrm{id}_{N_{\mathbb{R}}}, \varphi_i)$, $i = 0, 1$ are tropical Lagrangian multisections. Then it is easy to see that $\mathbb{L}_0 \cup_c \mathbb{L}_1$ is separated with the zero cone being the only ramification point. It is obvious that $\mathbb{L}_0 \cup_c \mathbb{L}_1$ is combinatorially decomposable by $\mathbb{L}_0, \mathbb{L}_1$.

3A. ***From toric vector bundles to tropical Lagrangian multisections.*** Let X_Σ be the associated toric variety of Σ. Given a rank r toric vector bundle $\mathcal{E}$ on X_Σ, we can associate a rank r tropical Lagrangian multisection $\mathbb{L}_{\mathcal{E}}$ over Σ by following the construction in [26].

Let $\sigma \in \Sigma$ and $U(\sigma)$ be the affine toric variety corresponding to σ. The toric vector bundle splits equivariantly on $U(\sigma)$ as

$$\mathcal{E}|_{U(\sigma)} \cong \bigoplus_{m(\sigma) \in \boldsymbol{m}(\sigma)} \mathcal{L}_{m(\sigma)},$$

where $\boldsymbol{m}(\sigma) \subset M(\sigma) := M/(\sigma^{\perp} \cap M)$ is a multiset and $\mathcal{L}_{m(\sigma)}$ is the line bundle corresponds to the linear function $m(\sigma) \in M(\sigma)$. We define $\mathbb{L}_{\mathcal{E}}$ as follows. Let $|\Sigma| \to \Sigma$ be the map given by mapping $x \in |\Sigma|$ to the unique cone $\sigma \in \Sigma$ such that $x \in \mathrm{Int}(\sigma)$. Equip Σ with the quotient topology. Define

$$\Sigma_{\mathcal{E}} := \{(\sigma, m(\sigma)) \mid \sigma \in \Sigma, m(\sigma) \in \boldsymbol{m}(\sigma)\}$$

and let $\Sigma_{\mathcal{E}} \to \Sigma$ be the projection

$$(\sigma, m(\sigma)) \mapsto \sigma.$$

We emphasize that although $\boldsymbol{m}(\sigma)$ is a multiset, $\Sigma_{\mathcal{E}}$ is not. Equip $\Sigma_{\mathcal{E}}$ a poset structure

$$(\sigma_1, m(\sigma_1)) \leq (\sigma_2, m(\sigma_2)) \iff \sigma_1 \subset \sigma_2 \text{ and } m(\sigma_2)|_{\sigma_1} = m(\sigma_1)$$

and equip it with the poset topology, namely, a subset $K \subset \Sigma_L$ is closed if and only if

$$\{(\sigma_1, m(\sigma_1)) \mid (\sigma_1, m(\sigma_1)) \leq (\sigma_2, m(\sigma_2))\} \subset K$$

for all $(\sigma_2, m(\sigma_2)) \in K$. Define

$$L_{\mathcal{E}} := |\Sigma| \times_{\Sigma} \Sigma_{\mathcal{E}}.$$

Let the set of cones on $L_{\mathcal{E}}$ be $\Sigma \times_{\Sigma} \Sigma_{\mathcal{E}} \cong \Sigma_{\mathcal{E}}$. The multiplicity $\mu_{\mathcal{E}} : L_{\mathcal{E}} \to \mathbb{Z}_{>0}$ is defined by

$$\mu_{\mathcal{E}}(\sigma, m(\sigma)) := \text{number of times that } m(\sigma) \text{ appears in } \boldsymbol{m}(\sigma).$$

The projection map $\pi_{\mathcal{E}} : L_{\mathcal{E}} \to |\Sigma|$ then induces a rank r branched covering map of cone complexes $\pi_{\mathcal{E}} : (L_{\mathcal{E}}, \Sigma_{\mathcal{E}}, \mu_{\mathcal{E}}) \to (N_{\mathbb{R}}, \Sigma)$. The piecewise linear function $\varphi_{\mathcal{E}} : L_{\mathcal{E}} \to \mathbb{R}$ is tautologically given by

$$\varphi_{\mathcal{E}}|_{(\sigma, m(\sigma))} := \pi_{\mathcal{E}}^* m(\sigma).$$

This gives a tropical Lagrangian multisection $\mathbb{L}_{\mathcal{E}} := (L_{\mathcal{E}}, \Sigma_{\mathcal{E}}, \mu_{\mathcal{E}}, \pi_{\mathcal{E}}, \varphi_{\mathcal{E}})$.

Proposition 3.25. *The tropical Lagrangian multisection $\mathbb{L}_{\mathcal{E}}$ is separated.*

Proof. By construction, if $\omega^{(\alpha)}, \omega^{(\beta)} \in \Sigma_{\mathcal{E}}$ are distinct lifts of some $\omega \in \Sigma$, then $\varphi_{\mathcal{E}}|_{\omega^{(\alpha)}} \neq \varphi_{\mathcal{E}}|_{\omega^{(\alpha)}}$. In particular, slopes on different codimension 1 cones are different. □

Proposition 3.26. *Let $\mathcal{E}, \mathcal{E}_1, \mathcal{E}_2$ be toric vector bundles on X_{Σ}. Then*

(1) $\mathbb{L}_{\mathcal{E}^*} = -\mathbb{L}_{\mathcal{E}}$,

(2) $\mathbb{L}_{\mathcal{E}_1 \oplus \mathcal{E}_2} = \mathbb{L}_{\mathcal{E}_1} \cup_c \mathbb{L}_{\mathcal{E}_2}$,

(3) $\mathbb{L}_{\mathcal{E}_1 \otimes \mathcal{E}_2} = \mathbb{L}_{\mathcal{E}_1} \times_c \mathbb{L}_{\mathcal{E}_2}$.

Proof. They follow from the induced equivariant structure

$$\lambda \cdot f := f(\lambda^{-1} \cdot v),$$
$$\lambda \cdot (v_1 \oplus v_2) := (\lambda \cdot v_1) \oplus (\lambda \cdot v_2),$$
$$\lambda \cdot (v_1 \otimes v_2) := (\lambda \cdot v_1) \otimes (\lambda \cdot v_2),$$

where $f \in \mathcal{E}^*$, $v \in \mathcal{E}$, $v_1 \in \mathcal{E}_1$, $v_2 \in \mathcal{E}_2$. $\square$

The assignment $\mathcal{E} \mapsto \mathbb{L}_{\mathcal{E}}$ is not injective as the following example shows.

Example 3.27. Consider the toric vector bundles

$$\mathcal{E}_1 := \bigoplus_{i=1}^{2} \mathcal{O}_{\mathbb{P}^2}(D_i) \quad \text{and} \quad \mathcal{E}_2 := T_{\mathbb{P}^2} \oplus \mathcal{O}_{\mathbb{P}^2}.$$

Via the Euler sequence

$$0 \to \mathcal{O}_{\mathbb{P}^2} \to \bigoplus_{i=1}^{2} \mathcal{O}_{\mathbb{P}^2}(D_i) \to T_{\mathbb{P}^2} \to 0,$$

$\mathcal{E}_1$, $\mathcal{E}_2$ share the same equivariant Chern class and hence $\mathbb{L}_{\mathcal{E}_1} = \mathbb{L}_{\mathcal{E}_2}$ by Proposition 3.4 of [26]. This example also shows that combinatorially indecomposable components are not unique. Indeed, $\mathbb{L}_{\mathcal{E}_1} = \mathbb{L}_{\mathcal{O}_{\mathbb{P}^2}(D_0)} \cup_c \mathbb{L}_{\mathcal{O}_{\mathbb{P}^2}(D_1)} \cup_c \mathbb{L}_{\mathcal{O}_{\mathbb{P}^2}(D_2)}$, $\mathbb{L}_{\mathcal{E}_2} = \mathbb{L}_{\mathcal{O}_{\mathbb{P}^2}} \cup_c \mathbb{L}_{T_{\mathbb{P}^2}}$, and it is easy to see that $\mathbb{L}_{T_{\mathbb{P}^2}}$ is maximal and separated, hence combinatorially indecomposable.

4. Kaneyama's classification via SYZ-type construction

4A. ***Kaneyama's classification.*** We first rewrite Kaneyama's classification result in terms of the language of tropical Lagrangian multisections. By doing so, some properties of toric vector bundles can be read off from the tropical Lagrangian multisections.

In [21], Kaneyama classified toric vector bundles by both combinatorial and linear algebra data. We can rewrite and refine these data in terms of the language of tropical Lagrangian multisections. Let $\mathbb{L} = (L, \Sigma_L, \mu, \pi, \varphi)$ be a tropical Lagrangian multisection over Σ. For a maximal cone $\sigma' \in \Sigma_L$, we use the notation $m(\sigma')$ to denote the slope of φ on σ', which is an element in M. We also count lifts of a maximal cone with multiplicities (recall that each cone $\sigma' \in \Sigma_L$ has a multiplicity $\mu(\sigma')$).

Definition 4.1. Let $\mathbb{L}$ be a tropical Lagrangian multisection of rank r over Σ. A *Kaneyama data of* $\mathbb{L}$ is a collection $\boldsymbol{g} := \{g_{\sigma_1\sigma_2}\}_{\sigma_1,\sigma_2 \in \Sigma_L(n)} \subset \mathrm{GL}(r, \mathbb{C})$ such that

(G1) for any $\sigma \in \Sigma(n)$, we have $g_{\sigma\sigma} = \mathrm{Id}$, (G2) for any $\sigma_1, \sigma_2 \in \Sigma(n)$, the (α, β)-entry $g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}}$ of $g_{\sigma_1\sigma_2}$ is nonzero only if $\sigma_1^{(\alpha)} \cap \sigma_2^{(\beta)} \neq \varnothing$ and

$$m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)}) \in (\sigma_1 \cap \sigma_2)^\vee \cap M,$$

(G3) for any $\sigma_1, \sigma_2, \sigma_3 \in \Sigma(n)$, we have

$$g_{\sigma_1\sigma_2} g_{\sigma_2\sigma_3} = g_{\sigma_1\sigma_3}.$$

We denote by $\widetilde{\mathcal{K}}(\mathbb{L})$ the set of Kaneyama data on $\mathbb{L}$. Two Kaneyama data $\boldsymbol{g}, \boldsymbol{g}' \in \widetilde{\mathcal{K}}(\mathbb{L})$ are said to be *equivalent* if for any $\sigma \in \Sigma(n)$, there exists $h_\sigma := (h_{\sigma^{(\alpha)}\sigma^{(\beta)}}) \in \mathrm{GL}(r, \mathbb{C})$ such that

(H1) $h_{\sigma^{(\alpha)}\sigma^{(\beta)}} \neq 0$ only if

$$m(\sigma^{(\alpha)}) - m(\sigma^{(\beta)}) \in \sigma^\vee \cap M,$$

(H2) for any $\sigma_1, \sigma_2 \in \Sigma(n)$,

$$h_{\sigma_2} g_{\sigma_1\sigma_2} = g'_{\sigma_1\sigma_2} h_{\sigma_1}.$$

We denote by $\mathcal{K}(\mathbb{L})$ the set of equivalence classes of Kaneyama data on $\mathbb{L}$.

Remark 4.2. In Kaneyama's work [21, pages 74–75], conditions (i) and (i') there are equivalent to continuity of φ, condition (ii) is equivalent to (G1), (G2), (G3) and condition (iii) is equivalent to (H1), (H2).

Theorem 4.3 (a reformulation of [21, Theorem 4.2]). *Let $\mathbb{L}$ be a tropical Lagrangian multisection over Σ. If $\mathbb{L}$ admits a Kaneyama data $\boldsymbol{g}$, then there is a toric vector bundle $\mathcal{E}(\mathbb{L}, \boldsymbol{g})$ over X_Σ such that $\mathbb{L}_{\mathcal{E}(\mathbb{L},\boldsymbol{g})} \leq \mathbb{L}$. Two Kaneyama data $\boldsymbol{g}, \boldsymbol{g}' \in \widetilde{\mathcal{K}}(\mathbb{L})$ are equivalent if and only if $\mathcal{E}(\mathbb{L}, \boldsymbol{g}) \cong \mathcal{E}(\mathbb{L}, \boldsymbol{g}')$ as toric vector bundles.*

Proof. The $(\mathbb{C}^\times)^n$-action on the toric vector bundle $\mathcal{E}_\sigma = \bigoplus_{\alpha=1}^r \mathcal{L}_{m(\sigma^{(\alpha)})}$ on $U(\sigma)$ is given by

$$\lambda \cdot (p, 1(\sigma^{(\alpha)})) := (\lambda \cdot p, \lambda^{m(\sigma^{(\alpha)})} 1(\sigma^{(\alpha)})), \tag{1}$$

where $p \in U(\sigma)$ and $1(\sigma^{(\alpha)})$ is an equivariant holomorphic frame of $\mathcal{L}_{m(\sigma^{(\alpha)})}$. It is straightforward to check that this action is compatible with the transition maps

$$G_{\sigma_1\sigma_2} : 1(\sigma_1^{(\alpha)}) \mapsto \sum_{\beta=1}^r g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}} z^{m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)})} 1(\sigma_2^{(\beta)}).$$

To prove that $\mathbb{L}_{\mathcal{E}(\mathbb{L},\boldsymbol{g})} \leq \mathbb{L}$, we define

$$f_{\sigma'} : \sigma' \to \pi(\sigma') \times \{m(\sigma')\}.$$

By continuity of φ, $\{f_{\sigma'}\}_{\sigma' \in \Sigma_L}$ can be glued to a continuous map $f : L \to L_{\mathcal{E}(\mathbb{L},\boldsymbol{g})}$ which maps cones in Σ_L to cones in $\Sigma_{\mathcal{E}(\mathbb{L},\boldsymbol{g})}$ homeomorphically. By definition, $f^*\varphi_{\mathcal{E}(\mathbb{L},\boldsymbol{g})} = \varphi$ and for any $\sigma \times \{m(\sigma)\}$, we have

$$\begin{aligned}\mathrm{Tr}_f(\mu)(\sigma \times \{m(\sigma)\}) = \sum_{\sigma' : \varphi|_{\sigma'} = m(\sigma)} \mu(\sigma') &= \#\{m \in \boldsymbol{m}(\sigma) : m = m(\sigma)\} \\ &= \mu_{\mathcal{E}(\mathbb{L},\boldsymbol{g})}(\sigma \times \{m(\sigma)\}).\end{aligned}$$

Hence $\mathbb{L}_{\mathcal{E}(\mathbb{L},\boldsymbol{g})} \leq \mathbb{L}$ via f. The last assertion follows from condition (iii) in [21]. □

Suppose $\mathbb{L}$ admits a Kaneyama data $\boldsymbol{g}$. The composition

$$\mathbb{L} \mapsto \mathcal{E}(\mathbb{L}, \boldsymbol{g}) \mapsto \mathbb{L}_{\mathcal{E}(\mathbb{L},\boldsymbol{g})}$$

may not be the identity map. For instance, suppose $\pi : L \to N_{\mathbb{R}}$ is a 2-fold cover conjugate to the square map $z \mapsto z^2$ on $\mathbb{C}$. Let Σ be the fan of $\mathbb{P}^2$. Then there is a natural collection of cones Σ' on L. Equip L with the 0 function. Then, the Kaneyama data $\boldsymbol{g}$ there gives a rank 2 toric vector bundle, which is just $\mathcal{O}_{\mathbb{P}^2}^{\oplus 2}$ with the trivial equivariant structure. But it is clear that the associated tropical Lagrangian multisection of $\mathcal{O}_{\mathbb{P}^2}^{\oplus 2}$ is given by $(N_{\mathbb{R}}, \Sigma, \mu, \mathrm{id}_{N_{\mathbb{R}}}, 0)$, with $\mu(\sigma) = 2$. Nevertheless, the map $\pi : L \to N_{\mathbb{R}}$ gives a branched covering of cone complexes that preserve the function. More generally, we have the following:

Theorem 4.4. *Let $\mathbb{L}_1, \mathbb{L}_2$ be tropical Lagrangian multisections of the same rank r. If $\mathbb{L}_1, \mathbb{L}_2$ are combinatorially equivalent, then there exists a bijection $f_* : \mathcal{K}(\mathbb{L}_1) \to \mathcal{K}(\mathbb{L}_2)$ such that $\mathcal{E}(\mathbb{L}_1, \boldsymbol{g}_1) \cong \mathcal{E}(\mathbb{L}_2, f_*(\boldsymbol{g}_1))$ as toric vector bundles. Conversely, if $\mathcal{E}(\mathbb{L}_1, \boldsymbol{g}_1) \cong \mathcal{E}(\mathbb{L}_2, \boldsymbol{g}_2)$ for some Kaneyama data, then $\mathbb{L}_1$ is combinatorially equivalent to $\mathbb{L}_2$.*

Proof. It suffices to prove that if $\mathbb{L}_2 \leq \mathbb{L}_1$ via some f, then any Kaneyama data of $\mathbb{L}_1$ gives a Kaneyama data of $\mathbb{L}_2$ such that their associated toric vector bundles are the same and vice versa. Let $\sigma_1', \sigma_2' \in \Sigma_1'(n)$ be maximal cones. By the assumption $f^*\varphi_2 = \varphi_1$, we have

$$m(f(\sigma_1')) - m(f(\sigma_2')) = m(\sigma_1') - m(\sigma_2').$$

Moreover, counting with multiplicity, f induces a permutation of the index set $\{1, \dots, r\}$, which parametrizes lifts of a maximal cell. Thus if $\boldsymbol{g}$ is a Kaneyama data of $\mathbb{L}_1$, then we can simply define

$$(f_*g)_{f(\sigma_1^{(\alpha)})f(\sigma_2^{(\beta)})} := g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}},$$

where $\sigma_1^{(\alpha)}, \sigma_2^{(\beta)}$ are preimages of $f(\sigma_1^{(\alpha)}), f(\sigma_2^{(\beta)})$ such that

$$m(f(\sigma_1^{(\alpha)})) = m(\sigma_1^{(\alpha)}), \quad m(f(\sigma_2^{(\beta)})) = m(\sigma_2^{(\beta)}).$$

Although the lifts $\sigma_1^{(\alpha)}, \sigma_2^{(\beta)}$ are not unique, the slopes are and hence f_*g is well-defined. It is straightforward to check that $f_*(\boldsymbol{g}) := \{(f_*g)_{\sigma_1\sigma_2}\}$ is a Kaneyama data for $\mathbb{L}_2$ and any two choices of preimages above differ by a permutation of the equivariant frame $\{1(\sigma^{(\alpha)})\}_{\alpha=1}^r$ and the torus action is preserved. It is then easy to see that there is an isomorphism $\mathcal{E}(\mathbb{L}_1, \boldsymbol{g}) \cong \mathcal{E}(\mathbb{L}_2, f_*(\boldsymbol{g}))$ of toric vector bundles. By pulling back, Kaneyama data on $\mathbb{L}_2$ induces a Kaneyama data on $\mathbb{L}_1$. Modulo equivalence, we obtain the desired bijection. The converse follows from Theorem 4.3. □

Remark 4.5. Theorem 4.4 has the following analog in mirror symmetry. Non-Hamiltonian equivalent Lagrangian branes in a symplectic manifold may give rise to the same mirror object as they can still be isomorphic in the derived Fukaya category. For example, in [7, Example 5.5] gives a Lagrangian immersion and a Lagrangian embedding in a symplectic 2-torus that shares the same mirror sheaf.

Proposition 4.6. *Suppose that $\mathbb{L} = \mathbb{L}_1 \cup_c \mathbb{L}_2$. Then there exists an embedding $\mathcal{K}(\mathbb{L}_1) \times \mathcal{K}(\mathbb{L}_2) \to \mathcal{K}(\mathbb{L})$.*

Proof. The embedding is given by taking the direct sum of matrices. □

Every tropical Lagrangian multisection can be combinatorially decomposed into combinatorially indecomposable ones. By Proposition 4.6, to obtain Kaneyama data on a general tropical Lagrangian multisection, it suffices to consider its combinatorially indecomposable components.

Theorem 4.7. *If $\mathbb{L}$ is combinatorially indecomposable, then $\mathcal{E}(\mathbb{L}, \boldsymbol{g})$ is indecomposable for any Kaneyama data $\boldsymbol{g}$ of $\mathbb{L}$. The converse is also true if $\mathbb{L}$ can be decomposed into a combinatorial union of two tropical Lagrangian multisections $\mathbb{L}_1, \mathbb{L}_2$ that admits Kaneyama data.*

Proof. If $\mathcal{E}(\mathbb{L}, \boldsymbol{g})$ is decomposable for some $\boldsymbol{g}$, say by $\mathcal{E}_1, \mathcal{E}_2$, then $\mathbb{L}_{\mathcal{E}(\mathbb{L},\boldsymbol{g})} = \mathbb{L}_{\mathcal{E}_1} \cup_c \mathbb{L}_{\mathcal{E}_2}$. Since $\mathbb{L}_{\mathcal{E}(\mathbb{L},\boldsymbol{g})} \leq \mathbb{L}$ by Theorem 4.3, $\mathbb{L}$ is also combinatorially decomposable. Conversely, suppose $\mathbb{L} = \mathbb{L}_1 \cup_c \mathbb{L}_2$ for some unobstructed $\mathbb{L}_1, \mathbb{L}_2$. Let $\boldsymbol{g}_1, \boldsymbol{g}_2$ be some Kaneyama data of $\mathbb{L}_1, \mathbb{L}_2$, respectively. Denote the image of $(\boldsymbol{g}_1, \boldsymbol{g}_2)$ under the embedding $\mathcal{K}(\mathbb{L}_1) \times \mathcal{K}(\mathbb{L}_2) \to \mathcal{K}(\mathbb{L})$ by $\boldsymbol{g}$. Then we have $\mathcal{E}(\mathbb{L}, \boldsymbol{g}) = \mathcal{E}(\mathbb{L}_1, \boldsymbol{g}_1) \oplus \mathcal{E}(\mathbb{L}_2, \boldsymbol{g}_2)$. □

Since sections ($r = 1$) always admit Kaneyama data, we have the following:

Corollary 4.8. *A rank 2 tropical Lagrangian multisection $\mathbb{L}$ is combinatorially indecomposable if and only if $\mathcal{E}(\mathbb{L}, \boldsymbol{g})$ is indecomposable for any Kaneyama data $\boldsymbol{g}$ of $\mathbb{L}$.*

Remark 4.9. The converse of Theorem 4.7 or Corollary 4.8 is not true if we just ask for $\mathcal{E}(\mathbb{L}, \boldsymbol{g})$ to be indecomposable *for some* $\boldsymbol{g}$. For instance, take any indecomposable toric vector bundle $\mathcal{E}$ that contains a toric subbundle. Then $\mathbb{L}_{\mathcal{E}}$ is combinatorially decomposable since $\mathcal{E}$ fits into an exact sequence of toric vector bundles. A concrete example is given by the tangent bundle of the Hirzebruch surface $\mathbb{F}_1$, which is indecomposable. But it contains a line bundle as a toric subbundle. See Corollary 4.1.2 of [22].

4B. ***A mirror symmetric approach.*** Now we go into one of the main themes of this paper. We would like to interpret Kaneyama's result in terms of mirror symmetry. We assume from now on all tropical Lagrangian multisections are combinatorially indecomposable and hence by Proposition 3.23, they are separated and the ramification locus S' always lies in the codimension 2 strata of (L, Σ_L).

4B1. *The semiflat bundle.* For a tropical multisection $\mathbb{L} = (L, \Sigma_L, \mu, \pi, \varphi)$, we have denoted the ramification locus by S' and the branch locus by S. Both of them are assumed to be contained in the codimension 2 strata. We define the 1-skeleton of X_Σ:

$$X_\Sigma^{(1)} := \bigcup_{\tau \in \Sigma(n-1)} X_\tau = X_\Sigma \setminus \bigcup_{\dim(\omega) \leq n-1} U(\omega).$$

The semiflat bundle is a locally free sheaf on $X_\Sigma^{(1)}$. To construct it, we first provide a good open cover for $L \setminus L^{(n-2)}$. For each $\sigma' \in \Sigma_L(n)$, choose a small neighborhood $V_{\sigma'} \subset L \setminus L^{(n-2)}$ contains $\sigma' \setminus L^{(n-2)}$ such that $V_{\sigma_1'} \cap V_{\sigma_2'} \neq \varnothing$ if and only if $\sigma_1' \cap \sigma_2' \in \Sigma_L(n-1)$. See Figure 4. Choose any $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \setminus L^{(n-2)}$. Denote the transition map on $V_{\sigma_1'} \cap V_{\sigma_2'}$ by

$$1_{\sigma_1'} \mapsto g_{\sigma_1'\sigma_2'}^{\mathrm{sf}} 1_{\sigma_2'},$$

where $\sigma_2' \in \Sigma_L(n)$ is the unique lift of σ_2 such that $\sigma_1' \cap \sigma_2' \in \Sigma_L(n-1)$. For a cone $\sigma \in \Sigma$, let $V(\sigma) := U(\sigma) \cap X_\Sigma^{(1)}$. If $\omega \subset S$, then $V(\omega) = \varnothing$. Thus, $\{V(\sigma)\}_{\sigma \in \Sigma(n)}$ forms an open cover of $X_\Sigma^{(1)}$ such that if $\sigma_1 \cap \sigma_2 \in \Sigma(n-1)$, we have $\varnothing \neq V(\sigma_1 \cap \sigma_2) \subset X_\Sigma^{(1)}$. For a maximal cone $\sigma \in \Sigma(n)$, we put

$$\mathcal{E}_\sigma := \bigoplus_{\alpha=1}^{r} \mathcal{L}_{m(\sigma^{(\alpha)})},$$

which is a toric vector bundle defined on $U(\sigma)$. For $\sigma_1, \sigma_2 \in \Sigma(n)$ such that $\sigma_1 \cap \sigma_2 \in \Sigma(n-1)$, we define $G_{\sigma_1\sigma_2}^{\mathrm{sf}} : \mathcal{E}_{\sigma_1}|_{V(\sigma_1 \cap \sigma_2)} \to \mathcal{E}_{\sigma_2}|_{V(\sigma_1 \cap \sigma_2)}$ by

$$G_{\sigma_1\sigma_2}^{\mathrm{sf}} : 1(\sigma_1^{(\alpha)}) \mapsto g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}}^{\mathrm{sf}} z^{m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)})} 1(\sigma_2^{(\beta)}),$$

where $\sigma_2^{(\beta)}$ is uniquely determined by the conditions $\varnothing \neq \sigma_1^{(\alpha)} \cap \sigma_2^{(\beta)} \in \Sigma_L(n-1)$ and $\pi(\sigma_2^{(\beta)}) = \sigma_2$. Since we have no triple intersections, $\{g_{\sigma_1'\sigma_2'}^{\mathrm{sf}}\}$ immediately satisfies the cocycle condition.

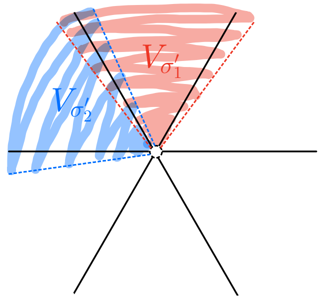


Figure 4. The space $L\backslash L^{(n-2)}$ and the neighborhoods $V_{\sigma_1'}$, $V_{\sigma_2'}$.

Definition 4.10. Let $\mathbb{L} = (L, \Sigma', \mu, \pi, \varphi)$ be a tropical Lagrangian multisection over Σ. Equip $L\backslash L^{(n-2)}$ with a $\mathbb{C}^\times$-local system $\mathcal{L}$. The vector bundle $\mathcal{E}^{\mathrm{sf}}(\mathbb{L}, \mathcal{L})$ is called the *semiflat bundle of* $(\mathbb{L}, \mathcal{L})$.

4B2. *Wall-crossing factors.* After constructing the semiflat bundle $\mathcal{E}^{\mathrm{sf}}(\mathbb{L}, \mathcal{L})$ of $(\mathbb{L}, \mathcal{L})$, we would like to extend $\mathcal{E}^{\mathrm{sf}}(\mathbb{L}, \mathcal{L})$ to the whole space X_Σ. To do this, we may need to correct $G^{\mathrm{sf}}_{\sigma_1\sigma_2}$ by certain factors. Let $\tau \in \Sigma(n-1)$ and $\sigma_1, \sigma_2 \in \Sigma(n)$ be the unique maximal cones so that $\sigma_1 \cap \sigma_2 = \tau$. For each $\omega' \in \Sigma_L$, we define a bundle map $N_\tau(\omega') : \mathcal{E}_{\sigma_1}|_{U(\tau)} \to \mathcal{E}_{\sigma_1}|_{U(\tau)}$ so that with respective to the frame $\{1(\sigma_1^{(\alpha)})\}$, the (α, β)-entry is given by

$$N^{(\alpha\beta)}_{\tau,\sigma_1}(\omega') := \begin{cases} n^{(\alpha\beta)}_{\tau,\sigma_1}(\omega') z^{m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)})} & \text{if } \omega' \subset \sigma_1^{(\alpha)} \cap \sigma_1^{(\beta)}, \alpha \neq \beta \\ & \quad \text{and } m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)}) \in \tau^\vee \cap M, \\ 0 & \text{otherwise} \end{cases}$$

for some $n^{(\alpha\beta)}_{\tau,\sigma_1}(\omega') \in \mathbb{C}$. Note that $N_{\tau,\sigma_1}(\omega') = 0$ if $\omega' \not\subset S'$. Put

$$S'_\tau(\sigma_1) := \{\sigma_1^{(\alpha)} \cap \sigma_1^{(\beta)} \mid m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)}) \in (\sigma_1 \cap \sigma_2)^\vee \cap M\}.$$

By assumption, cones in S'_τ are of codimension ≥ 2. Furthermore, there is a natural bijection $S'_\tau(\sigma_1) \cong S'_\tau(\sigma_2)$. Indeed, for $\sigma_1^{(\alpha)} \cap \sigma_1^{(\beta)} \in S'_\tau$, there exists unique $\sigma_2^{(\alpha')}, \sigma_2^{(\beta')} \in \Sigma_L(n)$ such that $\sigma_1^{(\alpha)} \cap \sigma_2^{(\alpha')}, \sigma_1^{(\beta)} \cap \sigma_2^{(\beta')} \in \Sigma_L(n-1)$. Then

$$\begin{aligned} &m(\sigma_2^{(\alpha')}) - m(\sigma_2^{(\beta')}) \\ &\qquad = (m(\sigma_2^{(\alpha')}) - m(\sigma_1^{(\alpha)})) - (m(\sigma_2^{(\beta')}) - m(\sigma_1^{(\beta)})) + (m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)})). \end{aligned}$$

The first two terms of the right-hand side are in $\tau^\perp \cap M$ by continuity and the last term is in $\tau^\vee \cap M$ by definition. Hence $\sigma_2^{(\alpha')} \cap \sigma_2^{(\beta')} \in S'_\tau$. As $\sigma_2^{(\alpha')}, \sigma_2^{(\beta')}$ are uniquely determined by $\sigma_1^{(\alpha)}, \sigma_1^{(\beta)}$ and vice versa, the assignment $\sigma_1^{(\alpha)} \cap \sigma_1^{(\beta)} \mapsto \sigma_2^{(\alpha')} \cap \sigma_2^{(\beta')}$

gives the desired bijection. Now we define

$$N_{\tau,\sigma_1} := \sum_{\omega' \in S'_\tau(\sigma_1)} N_{\tau,\sigma_1}(\omega').$$

Remark 4.11. Note that the change of frame from $\{1(\sigma_1^{(\alpha)})\}$ to $\{1(\sigma_2^{(\alpha)})\}$ is compatible with the bijection $S'_\tau(\sigma_1) \cong S'_\tau(\sigma_2)$, namely, by choosing suitable $n^{(\alpha\beta)}_{\tau,\sigma_2}(\omega')$, we have $G^{\mathrm{sf}}_{\sigma_1\sigma_2} \circ N_{\tau,\sigma_1}(\omega') \circ G^{\mathrm{sf}}_{\sigma_2\sigma_1} = N_{\tau,\sigma_2}$. As N_{τ,σ_1}, N_{τ,σ_2} are related by a change of frame, to simplify our notation, we simply write $N_\tau(\omega')$ for $N_{\tau,\sigma_1}(\omega')$ and N_τ for N_{τ,σ_1}. We will also write $N_{\sigma_1\sigma_2}$ for N_τ when we want to emphasize the maximal cones σ_1, σ_2 so that $\sigma_1 \cap \sigma_2 = \tau$.

Lemma 4.12. *For any $\tau \in \Sigma(n-1)$ and $\omega' \in \Sigma_L$, $N_\tau(\omega')$ is nilpotent and for any distinct lifts $\omega', \omega'' \in \Sigma_L$ of $\omega \in \Sigma$, $N_\tau(\omega')N_\tau(\omega'') = 0$. In particular, $[N_\tau(\omega'), N_\tau(\omega'')] = 0$ for any lifts $\omega', \omega'' \in \Sigma_L$ of ω.*

Proof. We show that $N_\tau(\omega')^k = 0$ has zero diagonal entries, for all $k \geq 1$. The case $k = 1$ is by definition. Assume $N_\tau(\omega')^k$ has zero diagonal entries for some $k \geq 1$. If there exists α such that

$$\sum_{\beta=1}^{r} (N_\tau(\omega')^k)^{(\alpha\beta)} N_\tau^{(\beta\alpha)}(\omega') \neq 0,$$

there must exist $\beta \neq \alpha$ such that

$$(N_\tau(\omega')^k)^{(\alpha\beta)} N_\tau^{(\beta\alpha)}(\omega') \neq 0,$$

as both $N_\tau(\omega')^k$, $N_\tau(\omega')$ have zero diagonal entries. This implies both $z^{m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\beta)})}$ and $z^{m(\sigma_1^{(\beta)})-m(\sigma_1^{(\alpha)})}$ are regular functions on the affine chart $U(\tau)$. Therefore $z^{m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\beta)})}$ must be invertible on $U(\tau)$ and so $m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)}) \in \tau^\perp$, which means $m(\sigma_1^{(\alpha)})|_\tau = m(\sigma_1^{(\beta)})|_\tau$. This violates $(n-1)$-separability. Hence $N_\tau(\omega')^{k+1}$ has zero diagonal entries too. By induction, we are done.

For the last part, we have

$$\sum_{\beta=1}^{r} n_\tau^{(\alpha\beta)}(\omega') n_\tau^{(\beta\gamma)}(\omega'') z^{m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\gamma)})}.$$

This sum is nonzero only if $\omega' \subset \sigma_1^{(\alpha)} \cap \sigma_1^{(\beta)}$ and $\omega'' \subset \sigma_1^{(\beta)} \cap \sigma_1^{(\gamma)}$ for some β. Then we must have $\omega'' = \omega'$ as $\sigma_1^{(\beta)}$ can only contain one lift of ω. □

Once a choice of $\{n^{(\alpha\beta)}_{\tau,\sigma}(\omega')\}$ is fixed, Lemma 4.12 allows us to define the product of matrices

$$\Theta_\tau := \prod_{\omega' \in S'_\tau} \Theta_\tau(\omega') := \prod_{\omega' \in S'_\tau} \exp(N_\tau(\omega')) \tag{2}$$

unambiguously. Moreover, we have $\det(\Theta_\tau) = 1$ and so Θ_τ is invertible over $\mathbb{C}[\tau^\vee \cap M]$.

Definition 4.13. For $\omega' \in \Sigma_L$, the factors $\{\Theta_\tau(\omega')\}_{\tau\in\Sigma(n-1)}$ are called *wall-crossing automorphisms associated to* ω'.

Remark 4.14. Similar to the notation $N_{\sigma_1\sigma_2}$, we write $\Theta_{\sigma_1\sigma_2}$ for Θ_τ when we want to emphasize the unique maximal cones σ_1, σ_2 so that $\sigma_1 \cap \sigma_2 = \tau$.

Now for $\tau \in \Sigma(n-1)$, put

$$G_{\sigma_1\sigma_2} := G^{\mathrm{sf}}_{\sigma_1\sigma_2} \circ \Theta_{\sigma_1\sigma_2},$$

where $\sigma_1, \sigma_2 \in \Sigma(n)$ are uniquely determined by $\tau = \sigma_1 \cap \sigma_2$. If we express $G_{\sigma_1\sigma_2}$ in terms of the frames $\{1(\sigma_1^{(\alpha)})\}$, $\{1(\sigma_2^{(\gamma)})\}$, we have

$$G_{\sigma_1\sigma_2} : 1(\sigma_1^{(\alpha)}) \mapsto \sum_{\beta=1}^{r} \theta_\tau^{(\alpha\beta)} g^{\mathrm{sf}}_{\sigma_1^{(\beta)}\sigma_2^{(\gamma)}} z^{m(\sigma_1^{(\alpha)})-m(\sigma_2^{(\gamma)})} 1(\sigma_2^{(\gamma)}).$$

In particular, it is easy to choose $n^{(\alpha\beta)}_{\sigma_1\sigma_2}$'s such that

$$G_{\sigma_2\sigma_1} = G^{-1}_{\sigma_1\sigma_2}.$$

We haven't defined $G_{\sigma_1\sigma_2}$ for general $\sigma_1, \sigma_2 \in \Sigma(n)$. To do this, given any $\sigma_1, \sigma_2 \in \Sigma(n)$ such that $\tau := \sigma_1 \cap \sigma_2$, we consider a sequence of maximal cones $\sigma_1 = \sigma_{1'}, \sigma_{2'}, \dots, \sigma_{l'} = \sigma_2 \in \Sigma(n)$ such that $\tau \subset \sigma_{i'}$ and $\sigma_{i'} \cap \sigma_{(i+1)'} \in \Sigma(n-1)$ for all i. Such a sequence always exists since the branch locus S is of codimension at least 2. Then we put

$$G_{\sigma_1\sigma_2} := G_{\sigma_{(l-1)'}\sigma_{l'}}|_{U(\tau)} \circ \cdots \circ G_{\sigma_{1'}\sigma_{2'}}|_{U(\tau)},$$

which is defined on $U(\tau)$. We need to ensure $G_{\sigma_1\sigma_2}$ is independent of the choice of such a sequence of maximal cones.

Definition 4.15. Given a combinatorially indecomposable tropical Lagrangian multisection $\mathbb{L}$ and a $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \backslash L^{(n-2)}$, a collection of wall-crossing automorphisms $\Theta := \{\Theta_\tau(\omega')\}_{\tau\in\Sigma(n-1),\omega'\subset S'}$ defined by (2) is said to be *ω-consistent* if for any cycle of maximal cones

$$\sigma_1, \sigma_2, \dots, \sigma_l, \sigma_{l+1} = \sigma_1$$

such that $\omega \subset \sigma_i$ and $\sigma_i \cap \sigma_{i+1} \in \Sigma(n-1)$ for all i, the composition

$$G_{\sigma_l\sigma_{l+1}}|_{U(\omega)} \circ \cdots \circ G_{\sigma_1\sigma_2}|_{U(\omega)} : \mathcal{E}_{\sigma_1}|_{U(\omega)} \to \mathcal{E}_{\sigma_1}|_{U(\omega)} \tag{3}$$

equals to the identity map on $\mathcal{E}_{\sigma_1}|_{U(\omega)}$. A collection of automorphisms Θ is said to be *consistent* if it is ω-consistent for all $\omega \in \Sigma$.

Proposition 4.16. *A collection of wall-crossing automorphisms Θ is consistent if and only if it is ω-consistent for all $\omega \in \Sigma(n-2)$.*

Proof. Fix $\omega \in \Sigma$. For each cycle of maximal cones

$$\sigma_1, \sigma_2, \dots, \sigma_l, \sigma_{l+1} = \sigma_1$$

that satisfy the condition in Definition 4.15, there is a loop $\gamma : [0, 1] \to N_{\mathbb{R}} \backslash |\Sigma(n-2)|$ such that $\gamma(0) = \gamma(1) \in \mathrm{Int}(\sigma_1)$, intersecting the codimension 1 cones $\mathrm{Int}(\sigma_i \cap \sigma_{i+1})$ transversely for all i. Note that the corresponding composition defined by (3) only depends on the homotopy class of γ. As $\pi_1(N_{\mathbb{R}} \backslash |\Sigma(n-2)|)$ is generated by loops around codimension 2 strata of $(N_{\mathbb{R}}, \Sigma)$, we may write γ in terms of these generators

$$\gamma = \gamma_1 * \cdots * \gamma_k.$$

By choosing sufficiently generic γ_i's, each of them determines a cycle of maximal cones that satisfies the condition stated in Definition 4.15. As the compositions correspond to γ_i's equal to the identity, the composition corresponds to γ also equal to the identity. Hence codimension 2 consistency implies consistency. The converse is trivial. □

It is clear that if Θ is consistent, then $G_{\sigma_1\sigma_2}$ is well-defined for all $\sigma_1, \sigma_2 \in \Sigma(n)$ and the cocycle condition holds on arbitrary triple intersections. Let's make the following definition.

Definition 4.17. A combinatorially indecomposable tropical Lagrangian multisection $\mathbb{L}$ is called *unobstructed* if there exists a $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \backslash L^{(n-2)}$ and a collection of consistent wall-crossing automorphisms Θ. If $\mathbb{L}$ is unobstructed, we denote by $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta)$ the vector bundle associated to the data $(\mathbb{L}, \mathcal{L}, \Theta)$.

Remark 4.18. The notion of (weakly) unobstructed Lagrangian submanifolds was introduced in [16] and [3] for the immersed case. The main feature of an unobstructed Lagrangian submanifolds is that its Floer cohomology is well-defined and hence defines an object in the Fukaya category. In particular, unobstructed Lagrangian submanifolds should have the corresponding mirror objects. As the existence of Kaneyama's data or the data $(\mathcal{L}, \Theta)$ are equivalent to the existence of toric vector bundles, we should think of the tropical Lagrangian multisection can be "realized" by an unobstructed Lagrangian. Thus, we borrow the terminology here.

In defining $G^{\mathrm{sf}}_{\sigma_1\sigma_2}$, we have chosen a 1-cocycle to represent the local system $\mathcal{L}$. When $\mathbb{L}$ is unobstructed, $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta)$ is independent of such choice as the following proposition shows.

Proposition 4.19. *For any isomorphism $\mathcal{L}' \cong \mathcal{L}$ of local system on $L \backslash L^{(n-2)}$, there is an isomorphism $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta) \cong \mathcal{E}(\mathbb{L}, \mathcal{L}', \Theta')$ of toric vector bundles, for some consistent Θ'.*

Proof. Let $f : \mathcal{L} \to \mathcal{L}'$ be an isomorphism of local systems. It induces an isomorphism $F : \pi_*\mathcal{L} \to \pi_*\mathcal{L}'$ of rank r local systems. Locally, F is given by a constant matrix and thus can be regarded as a toric automorphism on a chart $U(\sigma) \subset X_\Sigma$. We also have

$$F_{\sigma_2} \circ G^{\mathrm{sf}}_{\sigma_1\sigma_2} = G'^{\mathrm{sf}}_{\sigma_1\sigma_2} \circ F_{\sigma_1}.$$

If Θ is a consistent data, we simply define Θ' by conjugation by F, that is,

$$\Theta'_{\sigma_1\sigma_2} := F_{\sigma_1} \circ \Theta_{\sigma_1\sigma_2} \circ F^{-1}_{\sigma_1}.$$

Then it is by definition that

$$F_{\sigma_2} \circ G_{\sigma_1\sigma_2} = G'_{\sigma_1\sigma_2} \circ F_{\sigma_1},$$

which means $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta) \cong \mathcal{E}(\mathbb{L}, \mathcal{L}', \Theta')$ as toric vector bundles. □

Combinatorial indecomposability implies the following relation between $\mathcal{E}^{\mathrm{sf}}(\mathbb{L}, \mathcal{L})$ and $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta)$.

Theorem 4.20. *If $\mathbb{L}$ is combinatorially indecomposable, then $\Theta_{\sigma_1\sigma_2}|_{X^{(1)}_\Sigma} = \mathrm{Id}$, for any $\sigma_1, \sigma_2 \in \Sigma(n)$ so that $\sigma_1 \cap \sigma_2 \not\subset S$. In particular, if $\mathbb{L}$ is unobstructed, then $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta)|_{X^{(1)}_\Sigma} = \mathcal{E}^{\mathrm{sf}}(\mathbb{L}, \mathcal{L})$.*

Proof. Let $\tau := \sigma_1 \cap \sigma_2 \in \Sigma(n-1)$. For $m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)}) \in \tau^\vee \cap M$, $(n-1)$-separability implies that there exists a ray $\rho \subset \tau$ so that

$$(m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)}))(v_\rho) > 0,$$

where v_ρ is a generator of ρ. Hence $z^{m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)})}$ vanishes along the divisor $U(\tau) \cap X_\rho$ and in particular, vanishes on $U(\tau) \cap X_\tau$. Hence $\Theta_\tau|_{U(\tau)\cap X_\tau} = \mathrm{Id}$ and this proves $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta)|_{X^{(1)}_\Sigma} = \mathcal{E}^{\mathrm{sf}}(\mathbb{L}, \mathcal{L})$. □

By definition, unobstructedness implies the existence of Kaneyama data. It turns out all Kaneyama data arise from our construction.

Theorem 4.21. *Suppose $\mathbb{L}$ is combinatorially indecomposable and admits a Kaneyama data $\boldsymbol{g}$. Then there exists a $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \backslash L^{(n-2)}$ and consistent Θ such that $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta) = \mathcal{E}(\mathbb{L}, \boldsymbol{g})$.*

Proof. The transition maps of $\mathcal{E}(\mathbb{L}, \boldsymbol{g})$ are of form

$$1(\sigma_1^{(\alpha)}) \mapsto \sum_{\beta=1}^{r} g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}} z^{m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)})} 1(\sigma_2^{(\beta)}).$$

Consider two distinct maximal cones $\sigma_1, \sigma_2 \in \Sigma(n)$ such that $\tau := \sigma_1 \cap \sigma_2 \in \Sigma(n-1)$. For each lift $\sigma_2^{(\beta)}$ of σ_2, let $\sigma_1^{(\alpha')}$ be the unique lift of σ_1 such that $\sigma_1^{(\alpha')} \cap \sigma_2^{(\beta)} \in \Sigma_L(n-1)$. If $\sigma_1^{(\alpha)} \cap \sigma_2^{(\beta)} \subset S'$ then $\alpha \neq \alpha'$ and

$$z^{m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\alpha')})} = z^{m(\sigma_1^{(\alpha)})-m(\sigma_2^{(\beta)})} z^{m(\sigma_2^{(\beta)})-m(\sigma_1^{(\alpha')})}$$

is a regular function since $z^{m(\sigma_2^{(\beta)})-m(\sigma_1^{(\alpha')})}$ is nowhere vanishing on $U(\tau)$. As before $(n-1)$-separability implies the monomial $z^{m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\alpha')})}$ vanishes completely on $V(\tau)$. Hence the transition map of $\mathcal{E}(\mathbb{L}, \boldsymbol{g})|_{X_\Sigma^{(1)}}$ on $V(\tau)$ is given by

$$G^{\mathrm{sf}}_{\sigma_1\sigma_2} := G_{\sigma_1\sigma_2}|_{V(\tau)} : 1(\sigma_1^{(\alpha)}) \mapsto g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}} z^{m(\sigma_1^{(\alpha)})-m(\sigma_2^{(\beta)})} 1(\sigma_2^{(\beta)}),$$

where β is determined by α as before. Then with respect to the cover $\{V_{\sigma'}\}_{\sigma' \in \Sigma_L(n)}$ of $L \backslash L^{(n-2)}$, $\{g_{\sigma_1^{(\alpha)}\sigma_2^{(\beta)}}\}$ gives a $\mathbb{C}^\times$-local system $\mathcal{L}$ on $L \backslash L^{(n-2)}$. For $\sigma_1 \cap \sigma_2 \not\subset S$, we define

$$\Theta_{\sigma_1\sigma_2} := (G^{\mathrm{sf}}_{\sigma_1\sigma_2})^{-1} \circ G_{\sigma_1\sigma_2}.$$

The diagonal entries of $\Theta_{\sigma_1\sigma_2}$ are all equal to 1 and $(n-1)$-separability implies $\Theta_{\sigma_1\sigma_2} - \mathrm{Id}$ is nilpotent (see Lemma 4.12). This allows us to define

$$N_{\sigma_1\sigma_2} := \log(\Theta_{\sigma_1\sigma_2}) = \log(\mathrm{Id} + (\Theta_{\sigma_1\sigma_2} - \mathrm{Id})) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(\Theta_{\sigma_1\sigma_2} - \mathrm{Id})^k}{k}.$$

With respect to the frame $\{1(\sigma_1^{(\alpha)})\}$, the (α, β)-entry of $N_{\sigma_1\sigma_2}$ is given by

$$N^{(\alpha\beta)}_{\sigma_1\sigma_2} = \begin{cases} n^{(\alpha\beta)}_{\sigma_1\sigma_2} z^{m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\beta)})} & \text{if } \alpha \neq \beta \text{ and } m(\sigma_1^{(\alpha)})-m(\sigma_1^{(\beta)}) \in (\sigma_1 \cap \sigma_2)^\vee \cap M, \\ 0 & \text{otherwise,} \end{cases}$$

which can be decomposed as

$$N_{\sigma_1\sigma_2} = \sum_{\omega' \in S'_{\sigma_1\sigma_2}} N_{\sigma_1\sigma_2}(\omega').$$

The collection $\{\Theta_{\sigma_1\sigma_2}\}$ is obviously consistent so that $\mathcal{E}(\mathbb{L}, \mathcal{L}, \Theta) = \mathcal{E}(\mathbb{L}, \boldsymbol{g})$. □

Example 4.22. We look at the 2-fold tropical Lagrangian multisection $\mathbb{L}_{a,b,c}$ over the fan of $\mathbb{P}^2$. Here $a, b, c > 0$. See Figure 5. Choose $\mathcal{L}$ to be the local system on $L \backslash \pi^{-1}(0) \cong \mathbb{R}^2 \backslash \{0\}$ that has monodromy -1 around the minimal cone. Let $z^i_j := Z_i / Z_j$ be the inhomogeneous coordinates on $U(\sigma_i) \cap U(\sigma_j) \subset \mathbb{P}^2$. The semiflat mirror bundle $\mathcal{E}_0(\mathbb{L}_{a,b,c}, \mathcal{L})$ on the $\mathbb{P}^1$-skeleton of $\mathbb{P}^2$ is given by the transition maps

$$\tau^{\mathrm{sf}}_{01} := \begin{pmatrix} -\frac{1}{(z^1_0)^{a+b}} & 0 \\ 0 & \frac{1}{(z^1_0)^c} \end{pmatrix}, \quad \tau^{\mathrm{sf}}_{12} := \begin{pmatrix} \frac{1}{(z^2_1)^a} & 0 \\ 0 & -\frac{1}{(z^2_1)^{b+c}} \end{pmatrix}, \quad \tau^{\mathrm{sf}}_{20} := \begin{pmatrix} 0 & \frac{1}{(z^0_2)^b} \\ -\frac{1}{(z^0_2)^{a+c}} & 0 \end{pmatrix}.$$

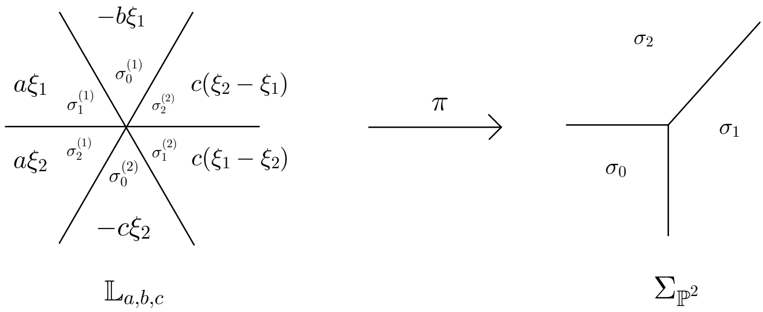


Figure 5. The tropical Lagrangian multisection $\mathbb{L}_{a,b,c}$ over $\Sigma_{\mathbb{P}^2}$.

We choose the wall-crossing factors to be

$$\Theta_{01} := \begin{pmatrix} 1 & 0 \\ -\frac{(z_0^2)^c}{(z_0^1)^b} & 1 \end{pmatrix}, \quad \Theta_{12} := \begin{pmatrix} 1 & -\frac{(z_1^0)^a}{(z_1^2)^c} \\ 0 & 1 \end{pmatrix}, \quad \Theta_{20} := \begin{pmatrix} 1 & 0 \\ -\frac{(z_2^1)^b}{(z_2^0)^a} & 1 \end{pmatrix}.$$

One can see that the resulting toric vector bundle $\mathcal{E}(\mathbb{L}_{a,b,c}, \mathcal{L}, \Theta)$ is actually isomorphic to $E_{a,b,c}$, the toric vector bundle introduced by Kaneyama in [21] using the exact sequence

$$0 \to \mathcal{O}_{\mathbb{P}^2} \to \mathcal{O}(aD_0) \oplus \mathcal{O}(bD_1) \oplus \mathcal{O}(cD_2) \to E_{a,b,c} \to 0.$$

Remark 4.23. From the symplectic point of view, we may think of $\Theta_\tau(\omega')$ as the exponentiation of the generating function of holomorphic disks emitted from the ramification locus ω', bounded by the Lagrangian multisection and certain SYZ fibers of $p : T^*N_{\mathbb{R}}/M \to N_{\mathbb{R}}$. The exponent $m(\sigma_1^{(\alpha)}) - m(\sigma_1^{(\beta)})$ in $\Theta_\tau(\omega')$ should be regarded as the direction of a wall if we use the polytope picture in $M_{\mathbb{R}}$. See [28] for a more detailed discussion in dimension 2.

5. Unobstructedness in dimension 2

In this final section, we would like to determine when $\mathbb{L}$ is unobstructed when $\mathbb{L}$ is a combinatorially indecomposable tropical Lagrangian multisection over a 2-dimensional complete fan. In this case, the ramification locus $S' = L^{(0)} = \pi^{-1}(0)$ is a singleton and $L \backslash \pi^{-1}(0) \cong \mathbb{R}^2 \backslash \{0\}$ topologically. First of all, not all such tropical Lagrangian multisections are unobstructed.

Example 5.1. Consider the tropical Lagrangian multisection $\mathbb{L}$ depicted as in Figure 6. It is easy to see that $\mathbb{L}$ is maximal and separated, which implies combinatorial indecomposability by Proposition 3.23. However, one checks easily that the matrices $G_{\sigma_0\sigma_1}$, $G_{\sigma_1\sigma_2}$ are all upper-triangular while $G_{\sigma_2\sigma_0}$ must have two nonzero off-diagonal entries. Thus $\mathbb{L}$ must be obstructed.

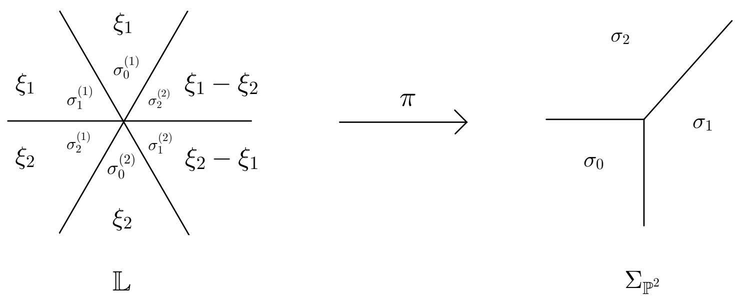


Figure 6. The tropical Lagrangian multisection $\mathbb{L}$ over $\Sigma_{\mathbb{P}^2}$.

Therefore, we need an extra assumption on the piecewise linear function φ to ensure unobstructedness. We begin with two lemmas.

Lemma 5.2. *Suppose $\mathbb{L}$ is a combinatorially indecomposable rank r tropical Lagrangian multisection over a complete 2-dimensional fan Σ. Let $\sigma \in \Sigma(2)$ and $\rho \subset \sigma$ be a ray. Then for $\alpha \neq \beta$, either $m(\sigma^{(\alpha)}) - m(\sigma^{(\beta)}) \in \rho^\vee \cap M$ or $m(\sigma^{(\beta)}) - m(\sigma^{(\alpha)}) \in \rho^\vee \cap M$.*

Proof. Since $\rho \not\subset S$, by separability, $m(\sigma^{(\alpha)})|_\rho \neq m(\sigma^{(\beta)})|_\rho$ if $\alpha \neq \beta$. In particular, $m(\sigma^{(\alpha)}) - m(\sigma^{(\beta)}) \neq 0$. Note that $\rho^\vee$ is a half plane in $M_\mathbb{R}$, we have $m(\sigma^{(\alpha)}) - m(\sigma^{(\beta)})$ or $m(\sigma^{(\beta)}) - m(\sigma^{(\alpha)})$ lies in $\rho^\vee$. Separability implies neither can lie in $\rho^\perp$. Hence only one of them can lie in $\rho^\vee$. □

Being unobstructed also restricts the choice of the local system $\mathcal{L}$.

Lemma 5.3. *If $\mathbb{L}$ is an unobstructed combinatorially indecomposable rank r tropical Lagrangian multisection over a complete 2-dimensional fan Σ, then $\mathcal{L}$ is the unique local system on $L \backslash S'$ that has monodromy $(-1)^{r+1}$ around the unique ramification point of $\pi : L \to N_\mathbb{R}$.*

Proof. Since $G^{\mathrm{sf}}_{\sigma_1\sigma_2}, \dots, G^{\mathrm{sf}}_{\sigma_{k-1}\sigma_k}$ are all diagonal, by taking the determinant of (3), we have

$$(-1)^{r+1} \prod_{\alpha=1}^{r} g^{\mathrm{sf}}_{\sigma_k^{(\alpha)}\sigma_1^{(\alpha+1)}} \prod_{i=1}^{k-1} \prod_{\alpha=1}^{r} g^{\mathrm{sf}}_{\sigma_i^{(\alpha)}\sigma_{i+1}^{(\alpha)}} = 1.$$

Hence the monodromy of $\mathcal{L}$, which is given by the cyclic product of all $g_{\sigma_i^{(\alpha)}\sigma_{i+1}^{(\beta)}}$'s, is equal to $(-1)^{r+1}$. As we are in dimension 2, the monodromy around the ramification point uniquely determines the local system. □

Remark 5.4. When $r = 2$, the choice of the local system $\mathcal{L}$ has appeared in the construction of the semiflat bundle in [15, Section 6.1]. Fukaya pointed out in [15, Remark 6.4] that there should be a Floer theoretic explanation of this local system based on the orientation problem of holomorphic disks. Believing the monomial term $n_{\tau\sigma}^{(\alpha\beta)} z^{m(\sigma^{(\alpha)})-m(\sigma^{(\beta)})}$ corresponds to holomorphic disks, our calculation in Lemma 5.3 suggests that the present of $\mathcal{L}$ is due to the fact that a holomorphic disk only propagates in only *one* direction; $m(\sigma^{(\alpha)})-m(\sigma^{(\beta)})$ or $m(\sigma^{(\beta)})-m(\sigma^{(\alpha)})$ but not both. The number $n_{\tau\sigma}^{(\alpha\beta)}$ will then be the weighted count of holomorphic disks (with extra boundary deformations if necessary, see Remark 5.6).

Therefore, to obtain unobstructedness, it is necessary for us to choose $\mathcal{L}$ to be the unique local system on $L\backslash S'$ that has monodromy $(-1)^{r+1}$. In particular, by Proposition 4.19, we may choose the transition maps of $\mathcal{L}$ to be

$$g^{\mathrm{sf}}_{\sigma_i^{(\alpha)}\sigma_{i+1}^{(\alpha)}} = 1 \quad \text{for all } i<k,\ \alpha=1,\dots,r$$

and

$$g^{\mathrm{sf}}_{\sigma_k^{(r)}\sigma_1^{(1)}} = (-1)^{r+1}, \quad g^{\mathrm{sf}}_{\sigma_k^{(\alpha)}\sigma_1^{(\alpha+1)}} = 1 \qquad \text{for } \alpha<r.$$

We put

$$g^{\mathrm{sf}}_{\sigma_k\sigma_1} := \begin{pmatrix} 0 & \cdots & 0 & (-1)^{r+1} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix},$$

which is the monodromy of the rank r local system $\pi_*\mathcal{L}$ on $N_{\mathbb{R}}\backslash\{0\}$. The consistency condition then becomes

$$\theta_{\sigma_k\sigma_1}\circ\theta_{\sigma_{k-1}\sigma_k}\circ\cdots\circ\theta_{\sigma_1\sigma_2} = g^{\mathrm{sf}}_{\sigma_1\sigma_k},$$

where $\theta_{\sigma_i\sigma_{i+1}}$ is obtained by deleting the monomial part of $\Theta_{\sigma_i\sigma_{i+1}}$. Recalling that $\Theta_{\sigma_i\sigma_{i+1}}$ is of the form $\mathrm{Id}+N_{\sigma_i\sigma_{i+1}}$, we may write the above equation as

$$\prod_{i=1}^{k}(\mathrm{Id}+n_{\sigma_i\sigma_{i+1}}) = g^{\mathrm{sf}}_{\sigma_1\sigma_k}. \tag{4}$$

Thus unobstructedness of $\mathbb{L}$ is equivalent to solving $n_{\sigma_i\sigma_{i+1}}$'s subordinated to the conditions

(N1) $n^{(\alpha\alpha)}_{\sigma_i\sigma_{i+1}} = 0$,

(N2) $n^{(\alpha\beta)}_{\sigma_i\sigma_{i+1}} \neq 0$ only if $m(\sigma_i^{(\alpha)})-m(\sigma_i^{(\beta)}) \in (\sigma_i\cap\sigma_{i+1})^\vee\cap M$.

Note that (N2) gives a combinatorial constraint on φ for solving (4) as expected by Example 5.1. Although (4) is not easy to solve for general r, it has the following interesting consequence.

Theorem 5.5. *Let $\mathbb{L}$ be combinatorially indecomposable rank r tropical Lagrangian multisection over a complete 2-dimensional fan Σ. Then*

$$\dim_{\mathbb{C}}(\mathcal{K}(\mathbb{L})) \leq \tfrac{1}{2}r(r-1)\cdot \#\Sigma(1),$$

where $\mathcal{K}(\mathbb{L})$ is the moduli space of toric vector bundles with equivariant Chern classes determined by $\mathbb{L}$.

Proof. The number of $n_{\sigma_i\sigma_{i+1}}$'s is exactly the number of rays in Σ and each $n_{\sigma_i\sigma_{i+1}}$ has at most $\frac{1}{2}r(r-1)$ free variables. By Theorem 4.21, our construction extracts all the possible Kaneyama data up to equivalence. The inequality follows. □

Remark 5.6. The moduli space $\mathcal{K}(\mathbb{L})$ is parametrized, up to the equivalence defined in Definition 4.1, by the variables $n^{(\alpha\beta)}_{\sigma_i\sigma_{i+1}}$, which only depend on $N_{\sigma_i\sigma_{i+1}}$ or $\Theta_{\sigma_i\sigma_{i+1}}$. As was discussed in Remark 4.23, these parameters are related to holomorphic disks bounded by a Lagrangian multisection and some SYZ fibers. One should expect that these variables are actually mirror to the moduli parameters of A_∞-deformations of the Lagrangian multisection.

Finally, we give an explicit description of the combinatorial obstruction for solving (4) in the case $r=2$. This condition is particularly easy to check. Let's recall Lemma 5.2. In the rank 2 case, it means for any $\sigma_1, \sigma_2 \in \Sigma(2)$ that intersect along an edge, we are always allowed to put 3 nonzero entries in the 2×2 matrices $G_{\sigma_1\sigma_2}$. Without loss of generality, we may arrange $\sigma_1^{(1)}, \sigma_2^{(1)}, \dots, \sigma_k^{(1)}, \sigma_1^{(2)}, \sigma_2^{(2)}, \dots, \sigma_k^{(2)}$ in an anticlockwise manner such that the matrix $G_{\sigma_k\sigma_1}$ is of form

$$\begin{pmatrix} z^{m(\sigma_k^{(1)})-m(\sigma_1^{(1)})} & -z^{m(\sigma_k^{(2)})-m(\sigma_1^{(1)})} \\ z^{m(\sigma_k^{(1)})-m(\sigma_1^{(2)})} & 0 \end{pmatrix}$$

and all the remaining $G_{\sigma_i\sigma_{i+1}}$ are either upper-triangular or lower-triangular.

Definition 5.7. Let $\mathbb{L}$ be a tropical Lagrangian multisection over a complete fan Σ. The *slope matrix* $M_{\sigma_1\sigma_2}$ associated to $\sigma_1, \sigma_2 \in \Sigma(n)$ is the matrix given by

$$M^{(\alpha\beta)}_{\sigma_1\sigma_2} := \begin{cases} m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)}) & \text{if } m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)}) \in (\sigma_1\cap\sigma_2)^\vee \cap M, \\ \infty & \text{otherwise.} \end{cases}$$

One associates to the slope matrix $M_{\sigma_1\sigma_2}$ the *monomial matrix*

$$Z^{(\alpha\beta)}_{\sigma_1\sigma_2} := \begin{cases} z^{m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)})} & \text{if } m(\sigma_1^{(\alpha)}) - m(\sigma_2^{(\beta)}) \in (\sigma_1\cap\sigma_2)^\vee \cap M, \\ 0 & \text{otherwise.} \end{cases}$$

We call a slope matrix *upper-triangular* (resp. *lower-triangular*) if the associated monomial matrix is upper-triangular (resp. lower-triangular).

By Lemma 5.3, the local system $\mathcal{L}$ needs to be chosen to have monodromy -1 around the ramification point. With the above choice of arrangement convention, it is necessary that the coefficient matrix of the composition $G_{\sigma_{k-1}\sigma_k} \circ \cdots \circ G_{\sigma_1\sigma_2}$ takes the form

$$\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}. \tag{5}$$

Definition 5.8. A combinatorially indecomposable rank 2 tropical Lagrangian multisection $\mathbb{L}$ over a complete 2-dimensional fan Σ is said to be satisfying the *slope condition* if under the above arrangement convention, one of the following conditions is satisfied:

(1) If $M_{\sigma_{k-1}\sigma_k}$ is upper-triangular, there is at least one $i < k-1$ such that $M_{\sigma_i\sigma_{i+1}}$ is lower-triangular.

(2) If $M_{\sigma_{k-1}\sigma_k}$ is lower-triangular, there exists some i, j with $1 \le i < j < k-1$, such that $M_{\sigma_j\sigma_{j+1}}$ is upper-triangular and $M_{\sigma_i\sigma_{i+1}}$ is lower-triangular.

Theorem 5.9. *A combinatorially indecomposable rank* 2 *tropical Lagrangian multisection* $\mathbb{L}$ *over a* 2*-dimensional complete fan* Σ *is unobstructed if and only if it satisfies the slope condition.*

Proof. If $\mathbb{L}$ is unobstructed and $G_{\sigma_{k-1}\sigma_k}$ is of upper-triangular type, then it is clear that we need a lower-triangular type matrix to bring it into the required form (5). Suppose $G_{\sigma_{k-1}\sigma_k}$ is of lower-triangular type. There must be some $j < k-1$ so that $G_{\sigma_j\sigma_{j+1}}$ is of upper-triangular type. If there are no $i < j$ for which $G_{\sigma_i\sigma_{i+1}}$ is of lower triangular type, the composition $G_{\sigma_{k-1}\sigma_k} \circ \cdots \circ G_{\sigma_1\sigma_2}$ will then take the form

$$\begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & * \\ * & * \end{pmatrix},$$

which can never have the required form (5). It remains to prove the converse. In the upper-triangular case, let $i < k-1$ be the first index for which $M_{\sigma_i\sigma_{i+1}}$ is lower-triangular. Then

$$(G_{\sigma_{k-1}\sigma_k} \circ G_{\sigma_{k-1}\sigma_{k-2}} \circ \cdots \circ G_{\sigma_{i+1}\sigma_{i+2}}) \circ G_{\sigma_i\sigma_{i+1}} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} = \begin{pmatrix} 1+ab & a \\ b & 1 \end{pmatrix},$$

and by choosing $a = -1,\ b = 1$, we obtain (5). Then we simply choose the remaining matrices to be the identity to obtain $G_{\sigma_{k-1}\sigma_k} \circ \cdots \circ G_{\sigma_1\sigma_2} = G_{\sigma_k\sigma_1}^{-1}$. For the lower-triangular case, let $i < j < k-1$ be the first index for which $M_{\sigma_j\sigma_{j+1}}$ is upper-triangular and $M_{\sigma_i\sigma_{i+1}}$ is lower-triangular. Then we have

$$G_{\sigma_{k-1}\sigma_k} \circ \cdots \circ G_{\sigma_j\sigma_{j+1}} \circ \cdots \circ G_{\sigma_i\sigma_{i+1}} = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} = \begin{pmatrix} 1+bc & b \\ a+c+abc & 1 \end{pmatrix}.$$

Choose $b = -1, c = 1$ and let a be arbitrary. Then the triple product is equal to (5). Again, by choosing the remaining matrices to be the identity, we obtain $G_{\sigma_{k-1}\sigma_k} \circ \cdots \circ G_{\sigma_1\sigma_2} = G_{\sigma_k\sigma_1}^{-1}$. □

The proof of Theorem 5.9 also sharpens the inequality in Theorem 5.5.

Corollary 5.10. *Suppose that* $\mathbb{L}$ *is a combinatorially indecomposable rank* 2 *tropical Lagrangian multisection over a complete* 2*-dimensional fan* Σ*. Then we have* $\dim_{\mathbb{C}}(\mathcal{K}(\mathbb{L})) \leq \#\Sigma(1) - 1$.

Proof. In the proof of Theorem 5.9, the equation $1 + ab = 0$ in the upper-triangular case or $1 + bc = 0$ in the lower-triangular case cut down the dimension by 1. By Theorem 4.21, our construction extracts all the possible toric structures with fixed equivariant Chern class, which is determined by $\mathbb{L}$. Hence $\dim_{\mathbb{C}}(\mathcal{K}(\mathbb{L})) \leq \#\Sigma(1) - 1$. □

Acknowledgment

The author is grateful to Kwokwai Chan and Ziming Nikolas Ma for useful discussions. The author would also like to thank Yong-Geun Oh for his interest in this work. The author was supported by IBS-R003-D1.

References

[1] M. Abouzaid, “Homogeneous coordinate rings and mirror symmetry for toric varieties”, *Geom. Topol.* **10** (2006), 1097–1156. MR Zbl

[2] M. Abouzaid, “Morse homology, tropical geometry, and homological mirror symmetry for toric varieties”, *Selecta Math.* (*N.S.*) **15**:2 (2009), 189–270. MR Zbl

[3] M. Akaho and D. Joyce, “Immersed Lagrangian Floer theory”, *J. Differential Geom.* **86**:3 (2010), 381–500. MR Zbl

[4] K. Chan, “Holomorphic line bundles on projective toric manifolds from Lagrangian sections of their mirrors by SYZ transformations”, *Int. Math. Res. Not.* **2009**:24 (2009), 4686–4708. MR Zbl

[5] K. Chan and N. C. Leung, “Mirror symmetry for toric Fano manifolds via SYZ transformations”, *Adv. Math.* **223**:3 (2010), 797–839. MR Zbl

[6] K. Chan and N. C. Leung, “Matrix factorizations from SYZ transformations”, pp. 203–224 in *Advances in geometric analysis*, edited by S. Janeczko et al., Adv. Lect. Math. (ALM) **21**, International Press, Somerville, MA, 2012. MR

[7] K. Chan and Y.-H. Suen, “SYZ transforms for immersed Lagrangian multisections”, *Trans. Amer. Math. Soc.* **372**:8 (2019), 5747–5780. MR Zbl

[8] K. Chan, D. Pomerleano, and K. Ueda, “Lagrangian sections on mirrors of toric Calabi–Yau 3-folds”, 2016. arXiv 1602.07075

[9] K. Chan, Z. N. Ma, and Y.-H. Suen, “Tropical Lagrangian multi-sections and smoothing of locally free sheaves over degenerate Calabi–Yau surfaces”, *Adv. Math.* **401** (2022), art. id. 108280. MR Zbl

[10] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Graduate Studies in Mathematics **124**, American Mathematical Society, Providence, RI, 2011. MR Zbl

[11] V. I. Danilov, "The geometry of toric varieties", *Uspekhi Mat. Nauk* **33**:2(200) (1978), 85–134. MR Zbl

[12] B. Fang, "Homological mirror symmetry is T-duality for $\mathbb{P}^n$", *Commun. Number Theory Phys.* **2**:4 (2008), 719–742. MR Zbl

[13] B. Fang, C.-C. M. Liu, D. Treumann, and E. Zaslow, "The coherent-constructible correspondence and homological mirror symmetry for toric varieties", pp. 3–37 in *Geometry and analysis*, vol. 2, edited by L. Ji, Adv. Lect. Math. **18**, International Press, Somerville, MA, 2011. MR Zbl

[14] B. Fang, C.-C. M. Liu, D. Treumann, and E. Zaslow, "T-duality and homological mirror symmetry for toric varieties", *Adv. Math.* **229**:3 (2012), 1875–1911. MR Zbl

[15] K. Fukaya, "Multivalued Morse theory, asymptotic analysis and mirror symmetry", pp. 205–278 in *Graphs and patterns in mathematics and theoretical physics*, edited by M. Lyubich and L. Takhtajan, Proc. Sympos. Pure Math. **73**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[16] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, *Lagrangian intersection Floer theory: anomaly and obstruction, I*, AMS/IP Studies in Advanced Mathematics **46.1**, American Mathematical Society, Providence, RI, 2009. MR Zbl

[17] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton University Press, 1993. MR Zbl

[18] M. Gross and B. Siebert, "Mirror symmetry via logarithmic degeneration data, I", *J. Differential Geom.* **72**:2 (2006), 169–338. MR Zbl

[19] M. Gross and B. Siebert, "Mirror symmetry via logarithmic degeneration data, II", *J. Algebraic Geom.* **19**:4 (2010), 679–780. MR Zbl

[20] M. Gross and B. Siebert, "From real affine geometry to complex geometry", *Ann. of Math.* (2) **174**:3 (2011), 1301–1428. MR Zbl

[21] T. Kaneyama, "On equivariant vector bundles on an almost homogeneous variety", *Nagoya Math. J.* **57** (1975), 65–86. MR Zbl

[22] B. Khan and J. Dasgupta, "Toric vector bundles on Bott tower", *Bull. Sci. Math.* **155** (2019), 74–91. MR Zbl

[23] A. A. Klyachko, "Equivariant bundles over toric varieties", *Izv. Akad. Nauk SSSR Ser. Mat.* **53**:5 (1989), 1001–1039. MR

[24] N. C. Leung, S.-T. Yau, and E. Zaslow, "From special Lagrangian to Hermitian–Yang–Mills via Fourier–Mukai transform", *Adv. Theor. Math. Phys.* **4**:6 (2000), 1319–1341. MR Zbl

[25] S. Payne, "Moduli of toric vector bundles", *Compos. Math.* **144**:5 (2008), 1199–1213. MR Zbl

[26] S. Payne, "Toric vector bundles, branched covers of fans, and the resolution property", *J. Algebraic Geom.* **18**:1 (2009), 1–36. MR Zbl

[27] A. Strominger, S.-T. Yau, and E. Zaslow, "Mirror symmetry is T-duality", *Nuclear Phys. B* **479**:1-2 (1996), 243–259. MR Zbl

[28] Y.-H. Suen, "Reconstruction of holomoprhic [sic] tangent bundle of complex projective plane via tropical Lagrangian multi-section", *New York J. Math.* **27** (2021), 1096–1114. MR Zbl

Received March 25, 2022. Revised October 3, 2023.

YAT-HIN SUEN
CENTER FOR GEOMETRY AND PHYSICS
INSTITUTE FOR BASIC SCIENCE
POHANG
SOUTH KOREA

and

SCHOOL OF MATHEMATICS
KOREA INSTITUTE FOR ADVANCED STUDY
SEOUL
SOUTH KOREA

yhsuen@kias.re.kr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.331

SOME EFFECTIVITY RESULTS FOR PRIMITIVE DIVISORS OF ELLIPTIC DIVISIBILITY SEQUENCES

MATTEO VERZOBIO

Let P be a nontorsion point on an elliptic curve defined over a number field K and consider the sequence $\{B_n\}_{n\in\mathbb{N}}$ of the denominators of $x(nP)$. We prove that every term of the sequence of the B_n has a primitive divisor for n greater than an effectively computable constant that we will explicitly compute. This constant will depend only on the model defining the curve.

1. Introduction

Let E be an elliptic curve defined by the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

with coefficients in a number field K. Let $P \in E(K)$ be a nontorsion point and let $\mathcal{O}_K$ be the ring of integers of K. Let us define the fractional ideal

$$(x(nP))\mathcal{O}_K = \frac{A_n}{B_n} \tag{1}$$

with A_n and B_n two relatively prime integral $\mathcal{O}_K$-ideals. We want to study the sequence of integral $\mathcal{O}_K$-ideals $\{B_n\}_{n\in\mathbb{N}}$. These are the so-called *elliptic divisibility sequences*. In particular, we want to study when a term B_n has a primitive divisor, i.e., when there exists a prime ideal $\mathcal{P}$ such that

$$\mathcal{P} \nmid B_1B_2\cdots B_{n-1} \quad \text{but} \quad \mathcal{P} \mid B_n.$$

Silverman [1988, Proposition 10] proved that, if E is defined over $\mathbb{Q}$, then B_n has a primitive divisor for n large enough. This result was generalized for every number field K in [Cheon and Hahn 1999], where the following theorem is proved.

Theorem 1.1 [Cheon and Hahn 1999, main theorem]. *Let E be an elliptic curve defined over a number field K and let P be a nontorsion point in $E(K)$. Consider the sequence $\{B_n\}_{n\in\mathbb{N}}$ of integral $\mathcal{O}_K$-ideals as defined in* (1). *Then B_n has a primitive divisor for all but finitely many $n \in \mathbb{N}$.*

MSC2020: 11B39, 11G05, 11G50.
Keywords: elliptic curves, primitive divisors, elliptic divisibility sequences.

The previous theorem is not effective. Indeed, the proof relies on Siegel's ineffective theorem about integral points on elliptic curves. The aim of this paper is to make the work of [Cheon and Hahn 1999] effective. Indeed, we will explicitly compute a constant C so that B_n has always a primitive divisor for $n > C$.

Theorem 1.2. *Let E be an elliptic curve defined over a number field K and let P be a nontorsion point in $E(K)$. Consider the sequence $\{B_n\}_{n\in\mathbb{N}}$ of integral $\mathcal{O}_K$-ideals as defined in* (1). *There exists a constant $C(E/K, \mathcal{M}) > 0$, effectively computable and depending only on the curve E over the field K equipped with a model $\mathcal{M}$ also defined over K, such that B_n has a primitive divisor for*

$$n > C(E/K, \mathcal{M}).$$

In Section 8, we explicitly compute such a constant $C(E/K, \mathcal{M})$ (see (13)).

Remark 1.3. The dependence on the model $\mathcal{M}$ is necessary. Indeed, given a nontorsion point P on an elliptic curve E and a positive constant C, it is easy to show that we can find a model of E such that B_n does not have a primitive divisor for all $n \leq C$.

Remark 1.4. It is conjectured that, in the case when $\mathcal{M}$ is minimal, the constant $C(E/K, \mathcal{M})$ should depend only on the field K. In [Ingram and Silverman 2012, Theorem 1] it is proved that the number of terms without a primitive divisor of an elliptic divisibility sequence can be bounded by a constant that does not depend on E and P, in the case when E is given by a minimal model, $K = \mathbb{Q}$, and assuming the *abc*-conjecture.

Remark 1.5. We believe that the techniques used in this paper can be applied also to a generalization of elliptic divisibility sequences. Let $\mathcal{O}$ be the endomorphism ring of E and, given $\alpha \in \mathcal{O}$, define B_α as the denominator of $(x(\alpha P))\mathcal{O}_K$. The sequence $\{B_\alpha\}_{\alpha\in\mathcal{O}}$ is a sequence of ideals and one can give a definition of primitive divisors also for these sequences (see [Streng 2008, Section 1]). It has been shown in [Streng 2008, main theorem] that also in this case there are only finitely many terms that do not have a primitive divisor (see also [Verzobio 2021b]). In the case when $\mathrm{End}(E) = \mathbb{Z}$ this is a trivial corollary of Theorem 1.1, but in the case $\mathrm{End}(E) \neq \mathbb{Z}$ (i.e., when E has complex multiplication) this is far from being easy. We believe that using the techniques of this paper one can find an explicit upper bound for the degree of α such that B_α does not have a primitive divisor, in the case when $\mathrm{End}(E)$ is a maximal order and it is a principal ideal domain.

2. Notation

The curve E is defined by the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with coefficients in the number field K.

The following notation will be used during the paper:

Δ is the discriminant of the equation defining the curve.

$\Delta_{E/K}$ is the minimal discriminant of the elliptic curve.

$j(E)$ is the j-invariant of the curve.

$D = [K : \mathbb{Q}]$ is the degree of the number field K.

$\mathcal{O}_K$ is the ring of integers of K.

Δ_K is the discriminant of the field K.

$f_{E/K}$ is the conductor of the curve.

$\sigma_{E/K} = \log|\mathbb{N}_{K/\mathbb{Q}}(\Delta_{E/K})|/\log|\mathbb{N}_{K/\mathbb{Q}}(f_{E/K})|$, where $\mathbb{N}_{K/\mathbb{Q}}$ is the norm of the field extension, is the Szpiro quotient; if E/K has everywhere good reduction (and then $f_{E/K} = 1$), we put $\sigma_{E/K} = 1$.

If $x \in \mathcal{O}_K$ is nonzero, then $\operatorname{gpf}(x)$ is the greatest rational prime p so that $\operatorname{ord}_p(\mathbb{N}_{K/\mathbb{Q}}(x)) > 0$.

If $n \in \mathbb{N}$ is nonzero, then $\omega(n)$ is the number of rational prime divisors of n.

If $x \in K^*$, define $\mathfrak{m}(x) = \max_{\mathcal{P}}\{\operatorname{ord}_{\mathcal{P}}(x)\}$, where the maximum runs over all primes in $\mathcal{O}_K$.

3. Preliminaries

Let M_K be the set of all places of K, take $v \in M_K$, and let $|\cdot|_v$ be the absolute value associated with v. Let n_v be the degree of the local extension $K_v/\mathbb{Q}_v$. We normalize the absolute values as in [Silverman 2009, Section VIII.5, after Example VIII.5.1]. If v is finite, then $|p|_v = p^{-1}$, where p is the rational prime associated to v. If v is infinite, then $|x|_v = \max\{x, -x\}$ for every $x \in \mathbb{Q}$. Thanks to this choice, we have the usual product formula, i.e.,

$$\prod_{v \in M_K} |x|_v^{n_v} = 1$$

for every $x \in K^*$. Define M_K^∞ as the set of infinite places of K and M_K^0 as the set of finite places.

Now, we define the height of a point on the curve; more details can be found in [Silverman 2009, Chapter VIII]. Given $x \in K^*$, define

$$h_v(x) := \max\{0, \log|x|_v\}$$

and

$$h(x) := \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} n_v h_v(x).$$

For every point $R \neq O$ of $E(K)$, define

$$h_v(R) := h_v(x(R))$$

and the height of the point as

$$h(R) := h(x(R)).$$

So, for every $R \in E(K) \setminus \{O\}$,

$$h(R) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v h_v(R).$$

Finally, put

$$h(O) = 0,$$

where O is the identity of the curve.

Given a point R in $E(K)$, define the canonical height as in [Silverman 2009, Proposition VIII.9.1], i.e.,

$$\hat{h}(R) = \frac{1}{2} \lim_{N \to \infty} \frac{h(2^N R)}{4^N}.$$

We recall the properties of the height and of the canonical height that will be necessary for this paper.

- It is known that the difference between the height and the canonical height can be bounded by an explicit constant. In particular, we will use the following result. Let
$$C_E = \frac{h(j(E))}{4} + \frac{h(\Delta)}{6} + 2.14.$$
If E is defined by a Weierstrass equation in short form and with integer coefficients, then, for every $R \in E(K)$,
$$|h(R) - 2\hat{h}(R)| \leq C_E.$$
This is proved in [Silverman 1990, equation 3].
- The canonical height is quadratic, i.e.,
$$\hat{h}(nR) = n^2 \hat{h}(R)$$
for every R in $E(K)$ and $n \in \mathbb{N}$.
- For every nontorsion point $R \in E(K)$,
$$\hat{h}(R) > 0.$$

There exists a positive constant J_E, effectively computable and depending only on E and K, such that

$$J_E \leq \hat{h}(P)$$

for every nontorsion point $P \in E(K)$. Thanks to [Petsche 2006, Theorem 2], we can take

$$J_E = \frac{\log|\mathbb{N}_{K/\mathbb{Q}}(\Delta_{E/K})|}{10^{15} D^3 \sigma_{E/K}^6 \log^2(104613 D \sigma_{E/K}^2)}$$

where $\mathbb{N}_{K/\mathbb{Q}}$ is the norm of the field extension, $D = [K : \mathbb{Q}]$, and

$$\sigma_{E/K} = \frac{\log|\mathbb{N}_{K/\mathbb{Q}}(\Delta_{E/K})|}{\log|\mathbb{N}_{K/\mathbb{Q}}(f_{E/K})|}.$$

If $f_{E/K} = 1$, we put $\sigma_{E/K} = 1$. The conductor $f_{E/K}$ is defined in [Silverman 2009, beginning of Section VIII.11].

In order to prove that B_n has a primitive divisor for all but finitely many terms, Silverman [1988] and Cheon and Hahn [1999] used a theorem of Siegel that says

$$\lim_{n\to\infty} \frac{h_\nu(nP)}{h(nP)} = 0$$

for every $\nu \in M_K$, as is proved in [Silverman 2009, Theorem IX.3.1]. This result is not effective and hence their results are not effective. We will use some results that tell us effectively how this limit goes to 0. As we will show later, for the finite places we will use some results on the formal group of the elliptic curve, and for the infinite places we will use the work in [David 1995]. The idea of using the result of David to study primitive divisors of elliptic divisibility sequences was introduced, as far as we know, by Streng [2008, Section 3].

We conclude this section by showing that we can focus only on the case when E is defined by a Weierstrass equation in short form and with integer coefficients. We will do that in Lemma 3.2. In order to prove that lemma, we need the following.

Lemma 3.1. *Let E be an elliptic curve defined over K by a Weierstrass equation with integer coefficients and let $P \in E(K)$. Let $\nu \in M_K^0$, $\mathcal{P}$ be the associated prime, and p be the associated rational prime. There exists*

$$k \leq p^{\frac{\nu(\Delta(E))}{12}} (2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1) \max\{4, \operatorname{ord}_{\mathcal{P}}(j(E)^{-1})\}$$

such that $\nu(x(kP)) < 0$.

Proof. Let $E_{\mathcal{P}}$ be a minimal model for the elliptic curve over $K_{\mathcal{P}}$ and let $P_{\mathcal{P}}$ be the image of P under the change of variables from E to $E_{\mathcal{P}}$. So, $x(P) = u_{\mathcal{P}}^2 x(P_{\mathcal{P}}) + r_{\mathcal{P}}$ for some $u_{\mathcal{P}}, r_{\mathcal{P}} \in K_{\mathcal{P}}$. By [Silverman 2009, Proposition VII.1.3.d], $\nu(u_{\mathcal{P}}) \geq 0$ and $\nu(r_{\mathcal{P}}) \geq 0$. Note that $12\nu(u_{\mathcal{P}}) = \nu(\Delta(E)) - \nu(\Delta(E_{\mathcal{P}})) \leq \nu(\Delta(E))$ and so

$\nu(u_{\mathcal{P}}) \leq \nu(\Delta(E))/12$. There exists a multiple $kP_{\mathcal{P}}$ with $k \leq \max\{4, \operatorname{ord}_{\mathcal{P}}(j(E_{\mathcal{P}})^{-1})\}$ such that $kP_{\mathcal{P}}$ is not a singular point in $E_{\mathcal{P}}(\mathbb{F}_{\mathcal{P}})$ (see [Silverman 2009, Corollary C.15.2.1]). Observe that $\mathbb{F}_{\mathcal{P}}$ has $\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P})$ elements and then the group of nonsingular points in $E_{\mathcal{P}}(\mathbb{F}_{\mathcal{P}})$ has at most $2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1$ elements. So, the order of $kP_{\mathcal{P}}$ in the group of nonsingular points modulo $\mathcal{P}$ is at most $2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1$. Hence, there exists

$$n_{\mathcal{P}}(P_{\mathcal{P}}) \leq (2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1) \max\{4, \operatorname{ord}_{\mathcal{P}}(j(E_{\mathcal{P}})^{-1})\}$$

such that $n_{\mathcal{P}}(P_{\mathcal{P}})P_{\mathcal{P}}$ reduces to the identity modulo $\mathcal{P}$. Given a point Q in $E(K)$, it is easy to show that Q reduces to the identity modulo $\mathcal{P}$ if and only if $\nu(x(Q)) < 0$. Therefore, $\nu(x(n_{\mathcal{P}}(P_{\mathcal{P}})P_{\mathcal{P}})) < 0$.

From a classic result on formal groups,

$$\nu\big(x(p^{\nu(u_{\mathcal{P}})} n_{\mathcal{P}}(P_{\mathcal{P}})P_{\mathcal{P}})\big) < -2\nu(u_{\mathcal{P}}).$$

For more details on formal groups, see Lemma 5.2 or [Silverman 2009, Corollary IV.4.4]. Using that $\nu(u_{\mathcal{P}}) \geq 0$ and $\nu(r_{\mathcal{P}}) \geq 0$, we have

$$\begin{aligned}\nu\big(x(p^{\nu(u_{\mathcal{P}})} n_{\mathcal{P}}(P_{\mathcal{P}})P)\big) &= \nu\big(u_{\mathcal{P}}^2 x(p^{\nu(u_{\mathcal{P}})} n_{\mathcal{P}}(P_{\mathcal{P}})P_{\mathcal{P}}) + r_{\mathcal{P}}\big) \\ &= \nu\big(x(p^{\nu(u_{\mathcal{P}})} n_{\mathcal{P}}(P_{\mathcal{P}})P_{\mathcal{P}})\big) + 2\nu(u_{\mathcal{P}}) \\ &< 0.\end{aligned}$$

We conclude recalling that $\nu(u_{\mathcal{P}}) \leq \nu(\Delta(E))/12$. □

Lemma 3.2. *Let E/K be an elliptic curve defined over K by a Weierstrass model $\mathcal{M}$. Then, there exists an elliptic curve E' defined over K by a short Weierstrass model $\mathcal{M}'$ with integer coefficients that is isomorphic over K to E, and a positive rational integer $s(E/K, \mathcal{M})$ such that: if Theorem 1.2 holds with $C(E'/K, \mathcal{M}')$ for E', $\mathcal{M}'$, then it holds with*

$$C(E/K, \mathcal{M}) = \max\{C(E'/K, \mathcal{M}'), s(E/K, \mathcal{M})\}$$

for E, $\mathcal{M}$. The constant $s(E/K, \mathcal{M})$ is effectively computable and will be defined during the proof (*see* (2)). *It depends only on E and $\mathcal{M}$.*

Proof. Recall that E is defined by the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Let u be the smallest positive rational integer such that, after the change of variables,

$$(x, y) \to (x', y') = \Big(u^2\Big(x + \frac{a_1^2}{12} + \frac{a_2}{3}\Big), u^3\Big(y + \frac{a_1}{2}x + \frac{a_3}{2}\Big)\Big)$$

we have that E is isomorphic to a curve E' of the form $y'^2 = x'^3 + ax' + b$ with a and b in $\mathcal{O}_K$. Let P' be the image of P under this isomorphism. So, $x' = u^2x + r$

for $u \in \mathbb{Z}_{\neq 0}$ and $r \in K$. Let q be the integral $\mathcal{O}_K$-ideal such that, for every $\nu \in M_K^0$,

$$\nu(q) = \max\{|\nu(u^2)|, -\nu(r)\}.$$

If $r = 0$, we take q such that $\nu(q) = |\nu(u^2)|$ for every $\nu \in M_K^0$. Note that q depends only on E and $\mathcal{M}$. Let B'_n be the elliptic divisibility sequence associated with E' and P'.

Let ν be the absolute value associated with a prime $\mathcal{P}$ coprime with q. We have $\nu(u^2) = 0$ and $\nu(r) \geq 0$. If $\mathcal{P}$ divides B_n, then $\nu(x(nP)) < 0$ and

$$\nu(u^2 x(nP) + r) = \nu(u^2 x(nP)) = \nu(x(nP)) = -\nu(B_n) < 0.$$

Therefore,

$$\nu(B'_n) = \nu(B_n) > 0.$$

In the same way, if $\mathcal{P}$ divides B'_n, then

$$\nu(B'_n) = \nu(B_n) > 0.$$

So, if $\mathcal{P}$ is coprime with q, then $\mathcal{P}$ divides B_n if and only if it divides B'_n.

Let

$$s = \max_{\mathcal{P}|q} \bigl\{ p^{(12\nu_{\mathcal{P}}(u) + \nu_{\mathcal{P}}(\Delta(E)))/12} (2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1)(\max\{4, \operatorname{ord}_{\mathcal{P}}(j(E')^{-1})\}) \bigr\}, \tag{2}$$

where p is the rational prime associated with $\mathcal{P}$.

Assume $n > s$. We will show that, if $\mathcal{P}$ is a primitive divisor of B'_n, then $\mathcal{P}$ is a primitive divisor also for B_n.

Let $\mathcal{P}$ be a primitive divisor of B'_n. Suppose that $\mathcal{P}$ divides q. By Lemma 3.1, there exists

$$\begin{aligned} k &\leq p^{(\nu(\Delta(E')))/12} (2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1)(\max\{4, \operatorname{ord}_{\mathcal{P}}(j(E')^{-1})\}) \\ &= p^{(12\nu(u) + \nu(\Delta(E)))/12} (2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1)(\max\{4, \operatorname{ord}_{\mathcal{P}}(j(E')^{-1})\}) \\ &\leq s \end{aligned}$$

such that $\nu(B'_k) > 0$. But, since $\mathcal{P}$ is a primitive divisor of B'_n we know that $k \geq n$. Hence, $n \leq s$ and this is absurd since we assumed $n > s$. So, $\mathcal{P}$ does not divide q. Since $\mathcal{P}$ is a primitive divisor of B'_n and $\mathcal{P}$ is coprime with q, then $\mathcal{P}$ divides B_n and does not divide B_k for $k < n$. Therefore, it is a primitive divisor for B_n.

In conclusion, if $n > \max\{C(E'/K, \mathcal{M}'), s\}$, then B'_n has a primitive divisor $\mathcal{P}$. As we showed, $\mathcal{P}$ is also a primitive divisor for B_n. Therefore, B_n has a primitive divisor for all $n > \max\{C(E'/K, \mathcal{M}'), s\}$.

Observe that s depends on $j(E')$, $\Delta(E)$, u, r, and q. It is easy to show these five values depend only on E and the model defining the curve. So, we are done. $\square$

From now on, we will assume that E is defined by a short Weierstrass equation with coefficients in $\mathcal{O}_K$ of the form

$$y^2 = x^3 - (g_2/4)x - (g_3/4).$$

Once we prove Theorem 1.2 under this assumption, then we can prove it in general using Lemma 3.2. It is useful to have E in this form in order to apply the work in [David 1995], as we will do in Section 6.

4. Structure of the proof

We start by recalling the structure of the proof of Cheon and Hahn of Theorem 1.1.

(1) If $\mathcal{P}$ is a nonprimitive divisor of B_n, then $\mathcal{P}$ divides $B_{n/q}$ for q a prime divisor of n. Moreover, if ν is the place associated to $\mathcal{P}$, then $h_\nu(nP)$ and $h_\nu\left(\frac{n}{q}P\right)$ are roughly the same.

(2) If B_n does not have a primitive divisor, then, for every $\nu \in M_K^0$, we have

$$h_\nu(nP) \leq \sum_{q|n} h_\nu\left(\frac{n}{q}P\right) + O(\log n),$$

using Step (1). Therefore,

$$\sum_{\nu\in M_K^0} h_\nu(nP) \leq \sum_{\nu\in M_K^0}\left(\sum_{q|n} h_\nu\left(\frac{n}{q}P\right) + O(\log n)\right).$$

(3) For every ν infinite, $h_\nu(nP)$ is negligible compared to $h(nP)$. In particular,

$$\sum_{\nu\in M_K^\infty} h_\nu(nP) = o(n^2).$$

(4) Putting together the inequalities of (2) and (3), we obtain

$$\begin{aligned}
2n^2\hat{h}(P) &= 2\hat{h}(nP)\\
&= h(nP) + O(1)\\
&= \frac{1}{D}\sum_{\nu\in M_K^0} n_\nu h_\nu(nP) + \frac{1}{D}\sum_{\nu\in M_K^\infty} n_\nu h_\nu(nP) + O(1)\\
&\leq \sum_{q|n} h\left(\frac{n}{q}P\right) + o(n^2)\\
&= 2\hat{h}(P)\sum_{q|n}\frac{n^2}{q^2} + o(n^2)\\
&= 2n^2\hat{h}(P)\left(\left(\sum_{q|n}\frac{1}{q^2}\right) + o(1)\right).
\end{aligned}$$

Note that one can use even sharper arguments using a complete inclusion-exclusion to find better inequalities (see for example [Streng 2008, Proof of the main theorem]).

(5) For every n we have $\sum_{q|n} q^{-2} < 1$ and then the inequality of (4) does not hold for n large enough. So, B_n does not have a primitive divisor only for finitely many $n \in \mathbb{N}$.

In order to make this proof effective, we need to make Steps (1) and (3) effective. In Section 5, we bound $h_\nu(nP) - h_\nu((n/q)P)$ as in Step (1). In Section 6, we make effective Step (3).

5. Finite places

Take $\mathcal{P}$ a prime over a valuation $\nu \in M_K^0$. Let p be the rational prime under $\mathcal{P}$. Recall that E is defined by a Weierstrass equation with integer coefficients. The group of points of $E(K_\mathcal{P})$ that reduce to the identity modulo $\mathcal{P}$ is a group that is isomorphic to a formal group, as proved in [Silverman 2009, Proposition VII.2.2]. Observe that, in the hypotheses of this proposition, there is the requirement that E is in minimal form. Anyway, the proof works in the exact same way only requiring that the coefficients of E are integers in $K_\mathcal{P}$, that is our case. Let $Q \in E(K_\mathcal{P})$ and, using the equation defining the elliptic curve, it is easy to show that $3\nu(x(Q)) = 2\nu(y(Q))$ and therefore

$$2\nu\left(\frac{x(Q)}{y(Q)}\right) = -\nu(x(Q)) > 0. \tag{3}$$

Define

$$z(Q) = \frac{x(Q)}{y(Q)} \in K_\mathcal{P}.$$

Lemma 5.1. *Take $\nu \in M_K^0$ and let $\mathcal{P}$ be the associated prime. Define $n_\mathcal{P}$ as the smallest integer such that $n_\mathcal{P} P$ reduces to the identity modulo $\mathcal{P}$. Then, kP reduces to the identity modulo $\mathcal{P}$ if and only if k is a multiple of $n_\mathcal{P}$. Moreover, $\nu(x(kP)) < 0$ if and only if k is a multiple of $n_\mathcal{P}$.*

Proof. Let $E_{\mathrm{ns}}(\mathbb{F}_\mathcal{P})$ be the group of nonsingular points of the curve E reduced (with respect to the given model) modulo $\mathcal{P}$. Suppose by contradiction that kP reduces to the identity but k is not a multiple of $n_\mathcal{P}$. Take q and r the quotient and the remainder of the division of k by $n_\mathcal{P}$. Since $n_\mathcal{P}$ does not divide k, we have that $0 < r < n_\mathcal{P}$. So,

$$rP \equiv nP - kqP \equiv O - O \equiv O \mod \mathcal{P}$$

and this is absurd since $n_{\mathcal{P}}$ is the smallest positive integer such that $n_{\mathcal{P}} P \equiv O \mod \mathcal{P}$. Vice versa, if $k = qn_{\mathcal{P}}$, then

$$kP \equiv q(n_{\mathcal{P}} P) \equiv qO \equiv O \mod \mathcal{P}.$$

Now, we conclude by observing that a point Q reduces to the identity modulo $\mathcal{P}$ if and only if $\nu(x(Q)) < 0$. □

Lemma 5.2. *Let $Q \in E(K)$ be such that $\nu(z(Q)) > 0$. Recall that p is the rational prime such that $\nu(p) > 0$. Then $\nu(z(p^e Q)) \geq e + \nu(z(Q))$. In particular, if $p^e \mid n$, then $\nu(z(nQ)) > e$.*

Proof. By [Silverman 2009, Corollary IV.4.4], $\nu(z(pQ)) \geq 1 + \nu(z(Q))$. Now, we proceed by induction. The case $e = 0$ is trivial. Assume that we know that $\nu(z(p^{e-1}Q)) \geq e - 1 + \nu(z(Q))$. Put $Q' = p^{e-1}Q$ and for the observation at the beginning of the proof we know $\nu(z(pQ')) \geq 1 + \nu(z(Q'))$. Therefore,

$$\nu(z(p^e Q)) = \nu(z(pQ')) \geq 1 + \nu(z(Q')) = 1 + \nu(z(p^{e-1}Q)) \geq e + \nu(z(Q)).$$

Now, we deal with the second part of the lemma. Let $n = p^e n'$ and, by Lemma 5.1, $\nu(z(n'Q)) > 0$. For the first part of the lemma, $\nu(z(nQ)) \geq e + \nu(z(n'Q)) > e$. □

Lemma 5.3. *Let $Q \in E(K)$ be such that $\nu(z(Q)) > \nu(p)/(p-1)$. Then,*

$$\nu(z(nQ)) = \nu(z(Q)) + \nu(n)$$

for all $n \geq 1$.

Proof. This follows by [Silverman 1988, Theorem IV.6.4, Proposition VII.2.2]. □

Definition 5.4. Let S be the set of finite places of K such that $\nu \mid 2$ or ν ramifies over $\mathbb{Q}$. Observe that this set is finite.

Corollary 5.5. *Let $Q \in E(K)$ be such that $\nu(z(Q)) > 0$. If $\nu \notin S$, then*

$$\nu(z(nQ)) = \nu(z(Q)) + \nu(n)$$

for all $n \geq 1$.

Proof. Since $\nu \notin S$, we have $\nu(p) = 1$ and $p - 1 \geq 2$. So, $\nu(z(Q)) \geq 1 > \nu(p)/(p-1)$ and we apply Lemma 5.3. □

Proposition 5.6. *Let E be an elliptic curve defined over a number field K and let $P \in E(K)$ be a nontorsion point. Take $\nu \in M_K^0$, let $\mathcal{P}$ be the associated prime, and p be the rational prime under $\mathcal{P}$. Recall that $n_{\mathcal{P}}$ is the smallest positive integer such that $n_{\mathcal{P}} P$ reduces to the identity modulo $\mathcal{P}$. Assume that $n_{\mathcal{P}} \mid n$ and $n_{\mathcal{P}} \neq n$. Then, one of the following hold:*

- *There exists a prime $q \mid n$ such that $\nu(z((n/q)P)) > 0$ and*

$$\nu(z(nP)) = \nu\Big(z\Big(\frac{n}{q} P\Big)\Big) + \nu(q).$$

- $\nu \in S$ *and*

$$n < n_{\mathcal{P}} p^{\frac{\nu(p)}{p-1}+1}.$$

Proof. Assume $\nu \notin S$ and let $Q = n_{\mathcal{P}} P$. Since $n/n_{\mathcal{P}}$ is an integer greater than 1, there is a prime q that divides it. By Corollary 5.5,

$$\begin{aligned}\nu(z(nP)) - \nu\Big(z\Big(\frac{n}{q}P\Big)\Big) &= \nu\Big(z\Big(\frac{n}{n_{\mathcal{P}}}Q\Big)\Big) - \nu\Big(z\Big(\frac{n}{qn_{\mathcal{P}}}Q\Big)\Big)\\ &= \nu\Big(\frac{n}{n_{\mathcal{P}}}\Big) - \nu\Big(\frac{n}{qn_{\mathcal{P}}}\Big) = \nu(q).\end{aligned}$$

So, we focus on the case $\nu \in S$. Assume that there exists $q \neq p$ such that $q \mid n/n_{\mathcal{P}}$. Then,

$$\nu(z(nP)) = \nu\Big(z\Big(\frac{n}{q}P\Big)\Big)$$

by [Silverman 1988, Corollary IV.4.4] and we are done since $\nu(q) = 0$. Assume now that there is no $q \neq p$ such that $q \mid n/n_{\mathcal{P}}$. So, $n = p^e n_{\mathcal{P}}$ with $e \geq 1$ (since $n \neq n_{\mathcal{P}}$) and recall that we defined $Q = n_{\mathcal{P}} P$.

Assume that $e-1 \geq \nu(p)/(p-1)$. Then, by Lemma 5.2, $\nu(z(p^{e-1}Q)) > e-1 \geq \nu(p)/(p-1)$. Therefore, by Lemma 5.3,

$$\nu(z(nP)) = \nu(z(p^e Q)) = \nu(z(p^{e-1}Q)) + \nu(p) = \nu\Big(z\Big(\frac{n}{p}P\Big)\Big) + \nu(p).$$

It remains the case $e - 1 < \nu(p)/(p-1)$. In this case,

$$n = n_{\mathcal{P}} p^e < n_{\mathcal{P}} p^{\frac{\nu(p)}{p-1}+1}. \qquad \square$$

Remark 5.7. To explicitly compute $\nu(z(nP))$ in the second case of the previous proposition one can use [Stange 2016, Lemma 5.1].

Lemma 5.8. *Let $\nu \in S$, $\mathcal{P}$ be the associated prime, and p be the associated rational prime. It holds that*

$$n_{\mathcal{P}} p^{\frac{\nu(p)}{p-1}+1} \leq \operatorname{gpf}(2\Delta_K)^{\frac{\mathfrak{m}(\Delta_E)}{12}} \max\{4, \mathfrak{m}(j(E)^{-1})\}(2\operatorname{gpf}(2\Delta_K)^D+1)\operatorname{gpf}(2\Delta_K)^{D+1}.$$

See Section 2 for the definition of the constants involved.

Proof. Recall that we are working with an elliptic curve E defined by a Weierstrass equation with integer coefficients. By Lemma 3.1,

$$n_{\mathcal{P}} \leq p^{\frac{\nu(\Delta(E))}{12}} (2\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) + 1) \max\{4, \operatorname{ord}_{\mathcal{P}}(j(E)^{-1})\}.$$

Since $\mathcal{P}$ is a prime over a place in S and the primes that ramify divide the discriminant of the field Δ_K, we have $\mathbb{N}_{K/\mathbb{Q}}(\mathcal{P}) \leq \operatorname{gpf}(2\Delta_K)^D$. Therefore,

$$n_{\mathcal{P}} \leq \operatorname{gpf}(2\Delta_K)^{\frac{\nu(\Delta(E))}{12}} \max\{4, \mathfrak{m}(j(E)^{-1})\}(2\operatorname{gpf}(2\Delta_K)^D + 1).$$

Moreover, $p \leq \operatorname{gpf}(2\Delta_K)$ and $\nu(p)/(p-1) \leq \nu(p) \leq D$. $\square$

Definition 5.9. Define

$$C_1 = \mathrm{gpf}(2\Delta_K)^{\frac{\mathfrak{m}(\Delta_E)}{12}} \max\{4, \mathfrak{m}(j(E)^{-1})\}(2\,\mathrm{gpf}(2\Delta_K)^D + 1)\,\mathrm{gpf}(2\Delta_K)^{D+1}.$$

Proposition 5.10. *Let E be an elliptic curve defined over a number field K and let $P \in E(K)$ be a nontorsion point. Take $\nu \in M_K^0$ and let $\mathcal{P}$ be the associated prime. Assume that $n_{\mathcal{P}} \mid n$, that $n_{\mathcal{P}} \neq n$, and that $n \geq C_1$. Then, there exists a prime $q \mid n$ such that*

$$h_\nu(nP) = h_\nu\Big(\frac{n}{q}P\Big) + 2h_\nu(q^{-1}).$$

Proof. Observe that we are in the hypotheses of Proposition 5.6. By Lemma 5.8 we know that, since $n \geq C_1$, we cannot be in the second case of Proposition 5.6. Therefore, there exists a prime $q \mid n$ such that

$$\nu(z(nP)) = \nu\Big(z\Big(\frac{n}{q}P\Big)\Big) + \nu(q).$$

Observe that, given $Q \in E(K)$ with $\nu(x(Q)) < 0$, then by (3),

$$h_\nu(x(Q)) = \log|x(Q)|_\nu = -2\log\left|\frac{x(Q)}{y(Q)}\right|_\nu = -2\log|z(Q)|_\nu.$$

Therefore,

$$h_\nu(nP) = -2\log|z(nP)|_\nu = -2\log\left|qz\Big(\frac{n}{q}P\Big)\right|_\nu = h_\nu\Big(\frac{n}{q}P\Big) + 2h_\nu(q^{-1}). \qquad \square$$

6. Infinite places

We know that $2n^2\hat{h}(P)$ is close to $h(nP)$ and that

$$h(nP) = \frac{1}{D}\sum_{\nu \in M_K^0} h_\nu(nP) + \frac{1}{D}\sum_{\nu \in M_K^\infty} h_\nu(nP).$$

Thanks to the previous section, we know how to bound $h_\nu(nP)$ for ν finite in the case when B_n does not have a primitive divisor. Now, we need to bound $h_\nu(nP)$ for ν infinite. We show that, for n large enough, $h_\nu(nP)$ is negligible compared to $n^2\hat{h}(P)$.

Recall that we are working with an elliptic curve E defined by the equation $y^2 = x^3 - (g_2/4)x - (g_3/4)$ with $g_2, g_3 \in \mathcal{O}_K$. Fix an embedding $K \hookrightarrow \mathbb{C}$ and consider the group of complex points $E(\mathbb{C})$. We briefly recall the properties of $E(\mathbb{C})$. For the details see [Silverman 2009, Chapter VI]. There is a unique lattice $\Lambda \subseteq \mathbb{C}$ such that $\mathbb{C}/\Lambda$ is isomorphic to $E(\mathbb{C})$ via the map $\phi : z \to (\wp(z), \wp'(z)/2, 1)$ (see [Silverman 2009, Theorem VI.5.1]). Thanks to [Silverman 1994, Proposition 1.1.5], we can take ω_1 and ω_2 two generators of Λ such that $\tau = \omega_2/\omega_1 \in \mathbb{C}$ is in the fundamental domain. In particular, $\Im\tau \geq \sqrt{3}/2$, where $\Im\tau$ is the imaginary part of τ. We need to make this choice in order to use [David 1995, Theorem 2.1].

Before proceeding, we need to define some constants. Let

$$h = \max\{1, h(1 : g_2 : g_3), h(j(E))\},$$

where $h(1 : g_2 : g_3)$ is the usual height on $\mathbb{P}^2$ (for a definition see [Silverman 2009, Section VIII.5]). Let

$$\log V_1 = \max\{h, (3\pi)/(D \cdot \Im\tau)\},$$
$$\log V_2 = \max\big\{h, (3\pi|\omega_2|^2)/(|\omega_1|^2 \cdot D \cdot \Im\tau)\big\}.$$

Let $c_1 := 3.6 \cdot 10^{41}$, that is the constant c_1 of [David 1995, Theorem 2.1] evaluated in $k = 2$. Define

$$C_3 = \max\{30, eh, \log V_1/D, \log V_2/D, D\}, \tag{4}$$

$$C_2 = 54 \cdot c_1 \cdot D^6 \log V_1 \log V_2. \tag{5}$$

Proposition 6.1. *Let E be an elliptic curve defined by the equation*

$$y^2 = x^3 - (g_2/4)x - (g_3/4)$$

for $g_2, g_3 \in K$ and take $P \in E(K)$. Let $z \in \mathbb{C}$ be so that $\phi(z) = P$ and suppose $\log n > C_3$. If $0 \le m_1, n_1, m_2, n_2 \le n$ with $n_1, n_2 \neq 0$, then

$$\log\left|z - \frac{m_1}{n_1}\omega_1 - \frac{m_2}{n_2}\omega_2\right| > -C_2 n^{1/2}.$$

Proof. David [1995, Theorem 2.1] proved that, for all integers $0 \le m_1, n_1, m_2, n_2 \le n$ with $n_1, n_2 \neq 0$, we have

$$\log\left|z - \frac{m_1}{n_1}\omega_1 - \frac{m_2}{n_2}\omega_2\right| > -c_1 D^6 (\log BD)(\log\log B + 1 + \log D + h)^3 \log V_1 \log V_2,$$

where

$$\log B := \max\{eh, \log n, \log V_1/D, \log V_2/D\}.$$

Since $\log n > C_3$, we have $\log n > D$, $\log n > eh > h+1$, and $\log n = \log B$. Hence,

$$c_1 D^6 (\log BD)(\log\log B + 1 + \log D + h)^3 \log V_1 \log V_2 < C_2 \log^4 n.$$

Moreover, since $\log n > 30$, we have

$$\log^4 n < n^{1/2}$$

and then

$$\log\left|z - \frac{m_1}{n_1}\omega_1 - \frac{m_2}{n_2}\omega_2\right| > -C_2 n^{1/2}. \qquad \square$$

7. Proof of Theorem 1.2

Define
$$\rho(n) = \sum_{p|n} \frac{1}{p^2}$$
and $\omega(n)$ as the number of prime divisors of n. It is easy to prove, by direct computation, that
$$\rho(n) < \sum_{p \text{ prime}} \frac{1}{p^2} < \frac{1}{2}.$$

Recall that C_1 is defined in Definition 5.9.

Lemma 7.1. *Let $n \geq C_1$. If B_n does not have a primitive divisor, then there exists an embedding $K \hookrightarrow \mathbb{C}$ such that*
$$\max\{\log|x(nP)|, 0\} \geq 2\hat{h}(P)n^2(1-\rho(n)) - 2\log n - C_E(\omega(n)+1),$$
where with $|x(nP)|$ we mean the absolute value in the embedding and C_E is defined in Section 2.

Proof. Suppose that B_n does not have a primitive divisor and take ν finite. Let $\mathcal{P}$ be the associated prime and assume $\nu(B_n) > 0$. Hence, $n_{\mathcal{P}} \mid n$ but $n \neq n_{\mathcal{P}}$ since B_n does not have a primitive divisor. So, using Proposition 5.10, there is a prime $q_\nu \mid n$ such that
$$h_\nu(nP) = h_\nu\Big(\frac{n}{q_\nu}P\Big) + 2h_\nu(q_\nu^{-1}).$$

Let $M_K^{0,n}$ be the set of finite places ν such that $h_\nu(nP) > 0$. Therefore,
$$\begin{aligned}\sum_{\nu\in M_K^0} n_\nu h_\nu(nP) &= \sum_{\nu\in M_K^{0,n}} n_\nu h_\nu(nP)\\ &\leq \sum_{\nu\in M_K^{0,n}} n_\nu h_\nu\Big(\frac{n}{q_\nu}P\Big) + 2n_\nu h_\nu(q_\nu^{-1})\\ &\leq \bigg(\sum_{q|n} Dh\Big(\frac{n}{q}P\Big) + 2Dh(q^{-1})\bigg).\end{aligned}$$

Here we are using that $h_\nu(kP) \geq 0$ for all $\nu \in M_K$ and all $k \geq 1$. Thus,
$$\begin{aligned}\frac{1}{D}\sum_{\nu\in M_K^\infty} n_\nu h_\nu(nP) &= h(nP) - \frac{1}{D}\sum_{\nu\in M_K^0} n_\nu h_\nu(nP)\\ &\geq 2\hat{h}(nP) - C_E - \sum_{q|n}\Big(h\Big(\frac{n}{q}P\Big) + 2\log q\Big)\end{aligned}$$

$$\begin{aligned}
&\geq 2\hat{h}(nP) - C_E - 2\log n - \sum_{q|n}\Big(2\hat{h}\Big(\frac{n}{q}P\Big) + C_E\Big) \\
&= 2\hat{h}(P)n^2\bigg(1 - \sum_{q|n}\frac{1}{q^2}\bigg) - 2\log n - C_E(\omega(n)+1) \\
&= 2\hat{h}(P)n^2(1-\rho(n)) - 2\log n - C_E(\omega(n)+1).
\end{aligned}$$

Since $h_\nu(nP) \geq 0$ for all $\nu \in M_K$ and $\sum_{\nu\in M_K^\infty} n_\nu = D$, at least one of the $h_\nu(nP)$, for $\nu \in M_K^\infty$, is larger than the right-hand side. Recalling that

$$h_\nu(x(P)) = \max\{\log|x(nP)|_\nu, 0\}$$

we conclude that

$$\max\{\log|x(nP)|, 0\} \geq 2\hat{h}(P)n^2(1-\rho(n)) - 2\log n - C_E(\omega(n)+1). \qquad \square$$

We briefly recall the hypotheses that we are assuming. As we said in the previous section, we are assuming that $E(\mathbb{C}) \cong \mathbb{C}/\Lambda$ with the lattice Λ generated by the complex numbers ω_1 and ω_2. Moreover, we are working with an elliptic curve defined by a Weierstrass equation with integer coefficients and in short form. Recall that C_2 is defined in (5) and define

$$C_4 = 2\max_{\nu\in M_K^\infty}\{\max\{|x(T)|_\nu \mid T \in E(\overline{K})[2]\setminus\{O\}\}\}.$$

Proposition 7.2. *Assume that*

$$2\hat{h}(P)n^2(1-\rho(n)) - 2\log n - C_E(\omega(n)+1) > 0, \tag{6}$$

that $n \geq C_1$, *and that* $\log n \geq C_3$, *as defined in* (4). *If* B_n *does not have a primitive divisor*, *then*

$$\hat{h}(P)n^2 \leq n^{1/2}(2C_2 + 4 + 2C_E + \log C_4). \tag{7}$$

Proof. Fix the embedding $K \hookrightarrow \mathbb{C}$ of Lemma 7.1. Since B_n does not have a primitive divisor, we have

$$\log|x(nP)| \geq 2\hat{h}(P)n^2(1-\rho(n)) - 2\log n - C_E(\omega(n)+1) \tag{8}$$

thanks to Lemma 7.1 and the assumption in (6). Consider the isomorphism $\mathbb{C}/\Lambda \cong E(\mathbb{C})$ as in Section 6 and take $z \in \mathbb{C}$ in the fundamental parallelogram of the period lattice of E such that $\phi(z) = P$. Assume

$$|x(nP)| \geq C_4$$

and let δ be the n-torsion point of $\mathbb{C}/\Lambda$ closest to z (if it is not unique, we choose one of them). Then,

$$\log|x(nP)| \leq -2\log|nz - n\delta| + \log 8 \tag{9}$$

thanks to [Ingram 2009, Lemma 8] (here we are using the assumption $|x(nP)| \geq C_4$). This lemma is stated for $K = \mathbb{Q}$, but the proof works in the exact same way for K number field. Since δ is an n-torsion point, we have

$$\delta = \frac{m_1}{n}\omega_1 + \frac{m_2}{n}\omega_2$$

for $0 \leq m_1, m_2 \leq n$. Using Proposition 6.1 and the assumption that $\log n > C_3$, we have

$$\log|z - \delta| = \log\left|z - \frac{m_1}{n}\omega_1 - \frac{m_2}{n}\omega_2\right| \geq -C_2 n^{1/2}.$$

Applying inequalities (8) and (9) we have

$$\begin{aligned}
\log 8 + 2C_2 n^{1/2} &\geq -2\log|z - \delta| + \log 8 \\
&= 2\log|n| - 2\log|nz - n\delta| + \log 8 \\
&\geq -2\log|nz - n\delta| + \log 8 \\
&\geq \log|x(nP)| \\
&\geq 2\hat{h}(P)n^2(1 - \rho(n)) - 2\log n - C_E(\omega(n) + 1).
\end{aligned} \tag{10}$$

Observe that $\omega(n) \leq \log_2 n$ and $(1 - \rho(n)) > 0.5$. Thus, rearranging (10), we have

$$\begin{aligned}
\hat{h}(P)n^2 &\leq 2\hat{h}(P)n^2(1 - \rho(n)) \\
&\leq 2\log n + C_E(\omega(n) + 1) + \log 8 + 2C_2 n^{1/2} \\
&\leq n^{1/2}(2C_2 + 4 + 2C_E).
\end{aligned}$$

Here we are using that $n^{1/2} > \log n$ thanks to the hypothesis $\log n > C_3$. Recall that we obtained this inequality assuming $|x(nP)| \geq C_4$. If $|x(nP)| < C_4$, applying again (8), we have

$$\begin{aligned}
\log C_4 &\geq \log|x(nP)| \\
&\geq 2\hat{h}(P)n^2(1 - \rho(n)) - 2\log n - C_E(\omega(n) + 1).
\end{aligned}$$

Therefore, one can easily show that, both in the case $|x(nP)| < C_4$ and in the case $|x(nP)| \geq C_4$, we have

$$\hat{h}(P)n^2 \leq n^{1/2}(2C_2 + 4 + 2C_E + \log C_4). \qquad \square$$

We are now ready to prove our main theorem. We will show that (7) does not hold if n is large enough.

Proof of Theorem 1.2. Define

$$C_5 = J_E^{-1}(2C_2 + 4 + 2C_E + \log C_4)$$

and take

$$n > \max\{C_1, C_5^{2/3}, V_1, V_2, \exp(D), (\exp(eh)), e^{30}\}. \tag{11}$$

We want to show that B_n has a primitive divisor.

Observe that, thanks to the assumption in (11) and the definition of C_3 in (4), we have $\log n > C_3$. Moreover,

$$n^{3/2} > C_5 = J_E^{-1}(2C_2 + 4 + 2C_E + \log C_4) > \hat{h}(P)^{-1}(4 + 2C_E)$$

and then

$$n^2 > \log n \cdot \hat{h}(P)^{-1}(4 + 2C_E).$$

Therefore, equation (6) holds. Finally, $n \geq C_1$. Hence, we are in the hypotheses of Proposition 7.2.

We assume that B_n does not have a primitive divisor and we find a contradiction. Since B_n does not have a primitive divisor, we know that we can apply Proposition 7.2 and (7) must hold. But

$$\begin{aligned} n^{3/2} &\geq J_E^{-1}(2C_2 + 4 + 2C_E + \log C_4) \\ &\geq \frac{2C_2 + 4 + 2C_E + \log C_4}{\hat{h}(P)}. \end{aligned}$$

and then (7) does not hold. Therefore, we find a contradiction and then B_n must have a primitive divisor.

In conclusion, define

$$C_6(E/K, \mathcal{M}) = \max\{C_1, V_1, V_2, \exp(D), \exp(eh), e^{30}, C_5^{2/3}\} \tag{12}$$

and B_n has a primitive divisor for $n > C_6(E/K, \mathcal{M})$. Observe that every constant involved in the definition of $C_6(E/K, \mathcal{M})$ does not depend on P and it is effectively computable (we will give more details in the next section). So, we are done.

Recall that we are working under the assumption that E is defined by a short Weierstrass equation with integer coefficients. In order to conclude for the general case, one has to use Lemma 3.2. $\square$

8. Explicit computation

Now, we explicitly write a constant $C(E/K, \mathcal{M})$ such that Theorem 1.2 holds. We assume that E is defined by a short Weierstrass equation with integer coefficients, the general case can be done using Lemma 3.2. Recall that we defined many constants in Section 2.

First of all, we show how to bound $|\tau|$, as defined at the beginning of Section 6. Recall that we are working under the assumption that τ is in the fundamental domain. Hence, we know $|\Re\tau| \le 1/2$ and then we study $\Im\tau$, the imaginary part of τ. Put $q = e^{2\pi i\tau}$ and then

$$|q| = e^{-2\pi\Im\tau}.$$

So,

$$\log|q| = -2\pi\Im\tau.$$

Thanks to [Silverman 1990, Lemma 5.2.b], we have

$$\bigl|\log|q|\bigr| \le 5.7 + \max\{\log|j(E)|, 0\}.$$

Therefore,

$$|\Im\tau| = \frac{\bigl|\log|q|\bigr|}{2\pi} \le \frac{5.7 + \max\{\log|j(E)|, 0\}}{2\pi}.$$

We obtain

$$|\tau|^2 = |\Re\tau|^2 + |\Im\tau|^2 \le \frac{1}{4} + \left(\frac{5.7 + \max\{\log|j(E)|, 0\}}{2\pi}\right)^2.$$

Let

$$\log V_1' = \max\{h, (2\sqrt{3}\pi)/D\},$$

$$\log V_2' = \max\left\{h, \left(2\sqrt{3}\pi\left(\frac{1}{4} + \left(\frac{5.7 + \max\{\log|j(E)|, 0\}}{2\pi}\right)^2\right)\right)/D\right\},$$

$$C_2' = 54 \cdot c_1 \cdot D^6 \log V_1' \log V_2'.$$

By the definitions of V_1, V_2, and C_2 given at the beginning of Section 6, we have $V_1' \ge V_1$, $V_2' \ge V_2$, and $C_2' \ge C_2$. Hence, by (12), Theorem 1.2 holds for

$$(13)\quad C(E/K, \mathcal{M}) = \max\left\{C_1, V_1', V_2', \exp(D), \exp(eh), e^{30}, \left(\frac{2C_2' + 4 + 2C_E + \log C_4}{J_E}\right)^{2/3}\right\}$$

where

$$h = \max\{1, h(1 : g_2 : g_3), h(j(E))\}, \quad c_1 = 3.6 \cdot 10^{41},$$

$$C_1 = \operatorname{gpf}(2\Delta_K)^{\frac{\mathfrak{m}(\Delta_E)}{12}} \max\{4, \mathfrak{m}(j(E)^{-1})\}(2\operatorname{gpf}(2\Delta_K)^D + 1)\operatorname{gpf}(2\Delta_K)^{D+1},$$

$$\log V_1' = \max\{h, (2\sqrt{3}\pi)/D\},$$

$$\log V_2' = \max\left\{h, \left(2\sqrt{3}\pi\left(\frac{1}{4} + \left(\frac{5.7 + \max\{\log|j(E)|, 0\}}{2\pi}\right)^2\right)\right)/D\right\},$$

$$C_2' = 54 \cdot c_1 \cdot D^6 \log V_1' \log V_2',$$
$$C_E = \frac{h(j(E))}{4} + \frac{h(\Delta)}{6} + 2.14,$$
$$C_4 = 2 \max\big\{|x(T)| \mid T \in E(\bar{\mathbb{Q}})[2] \setminus \{O\}\big\},$$
$$J_E = \frac{\log|\mathbb{N}_{K/\mathbb{Q}}(\Delta_{E/K})|}{10^{15} D^3 \sigma_{E/K}^6 \log^2(104613 D \sigma_{E/K}^2)}.$$

9. Examples

We apply our main theorem to a couple of examples.

Example 9.1. Let E be the rational elliptic curve defined by the equation $y^2 = x^3 - 4x + 4$. In this case, $D = \Delta_K = 1$, $h \approx 10.23$, $j(E) = -27648/11$, $\Delta_{E/K} = -2816$, $\sigma_{E/K} \approx 1.78$, and $C_4 \approx 4.76$. Using (13), we have

$$C(E/K, \mathcal{M}) \approx 5.88 \cdot 10^{42} < 6 \cdot 10^{42}.$$

With our methods, even if we optimize all the estimates in the proof, we cannot hope to find a constant for Theorem 1.2 much smaller than the one of Example 9.1. Indeed, in the definition of c_1 and of J_E appear constants that are very large (namely 10^{41} and 10^{15}) and so, even if the other constants involved are small, we cannot find a constant much smaller than 10^{38}. In order to find better constants, one would need to have better constants in the bound of canonical height and in logarithmic approximation.

Now, we present another example where we show the techniques that one can use to find the terms without a primitive divisor.

Example 9.2. We focus on the elliptic curve $y^2 = x^3 - 2x$ and $P = (2, 2) \in E(\mathbb{Q})$. The first terms of the sequence are $B_1 = 1$, $B_2 = 2^2$, $B_3 = 1$, $B_4 = (2^4)(3^2)(7^2)$, and $B_5 = (17)^2(19)^2$. Hence, B_1 and B_3 do not have a primitive divisor. For the terms that have very large indexes, we can use Theorem 1.2. So, we apply Theorem 1.2 with $C(E/K, \mathcal{M})$ as defined in (13). In the definition of $C(E/K, \mathcal{M})$ we substitute J_E with 0.3. Indeed, for every rational nontorsion point of E, we have $\hat{h}(P) > 0.3$ and J_E is a constant such that $J_E < \hat{h}(P)$. The minimum of the canonical height of the rational nontorsion points of E is computed in [LMFDB], where the canonical height is defined as the double of our canonical height. By Theorem 1.2 we have that, for $n \geq 2 \cdot 10^{31}$, B_n has a primitive divisor.

To deal with the terms with indexes smaller than $2 \cdot 10^{31}$, we can use the following techniques. By [Voutier and Yabuta 2012, Theorem 1.3] and [Verzobio 2021a], B_n has a primitive divisor for n even. So, we focus on the terms with odd indexes. As an easy corollary of [Voutier and Yabuta 2012, Lemma 3.4], we have that if B_n does not have a primitive divisor, then $\log B_n \leq 0.18n^2$. So, we can compute the values of B_n and check if the inequality holds (this is much faster than computing the factorization of the terms). As far as we know, the faster way to compute B_n is to use [Verzobio 2022, Theorem 1.9], where it is proved that, for $k \geq 9$,

$$b_k = \frac{b_{k-2}b_{k-6}b_4^2 - b_{k-4}^2 b_6 b_2}{b_{k-8}b_2^2} \tag{14}$$

where $b_k = \pm\sqrt{B_k}$ for an appropriate choice of the sign (for more details, see [Verzobio 2022, Definition B]). One can check that $\log B_n > 0.18n^2$ for $4 \leq n \leq 10^5$ using PARI/GP [2018] and then B_n has a primitive divisor for $4 \leq n \leq 10^5$. So, our bound is too large to be computationally useful and then new methods are needed to bridge the gap.

Acknowledgements

This paper is part of the author's PhD thesis at Università of Pisa. Moreover, this project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101034413.

I thank the referee for many helpful comments.

References

[Cheon and Hahn 1999] J. Cheon and S. Hahn, "The orders of the reductions of a point in the Mordell–Weil group of an elliptic curve", *Acta Arith.* **88**:3 (1999), 219–222. MR Zbl

[David 1995] S. David, "Minorations de formes linéaires de logarithmes elliptiques", *Mém. Soc. Math. France* (*N.S.*) 62 (1995), iv+143. MR Zbl

[Ingram 2009] P. Ingram, "Multiples of integral points on elliptic curves", *J. Number Theory* **129**:1 (2009), 182–208. MR Zbl

[Ingram and Silverman 2012] P. Ingram and J. H. Silverman, "Uniform estimates for primitive divisors in elliptic divisibility sequences", pp. 243–271 in *Number theory, analysis and geometry*, edited by D. Goldfeld et al., Springer, 2012. MR Zbl

[LMFDB] The LMFDB Collaboration, "The L-functions and modular forms database", electronic reference, available at http://www.lmfdb.org.

[PARI/GP 2018] "PARI/GP version 2.11.1", 2018, available at http://pari.math.u-bordeaux.fr.

[Petsche 2006] C. Petsche, "Small rational points on elliptic curves over number fields", *New York J. Math.* **12** (2006), 257–268. MR Zbl

[Silverman 1988] J. H. Silverman, "Wieferich's criterion and the *abc*-conjecture", *J. Number Theory* **30**:2 (1988), 226–237. MR Zbl

[Silverman 1990] J. H. Silverman, "The difference between the Weil height and the canonical height on elliptic curves", *Math. Comp.* **55**:192 (1990), 723–743. MR Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, 1994. MR Zbl

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Mathematics **106**, Springer, 2009. MR Zbl

[Stange 2016] K. E. Stange, "Integral points on elliptic curves and explicit valuations of division polynomials", *Canad. J. Math.* **68**:5 (2016), 1120–1158. MR Zbl

[Streng 2008] M. Streng, "Divisibility sequences for elliptic curves with complex multiplication", *Algebra Number Theory* **2**:2 (2008), 183–208. MR Zbl

[Verzobio 2021a] M. Verzobio, "Primitive divisors of elliptic divisibility sequences for elliptic curves with $j = 1728$", *Acta Arith.* **198**:2 (2021), 129–168. MR Zbl

[Verzobio 2021b] M. Verzobio, "Primitive divisors of sequences associated to elliptic curves with complex multiplication", *Res. Number Theory* **7**:2 (2021), art. id. 37. MR Zbl

[Verzobio 2022] M. Verzobio, "A recurrence relation for elliptic divisibility sequences", *Riv. Math. Univ. Parma* (*N.S.*) **13**:1 (2022), 223–242. MR Zbl

[Voutier and Yabuta 2012] P. Voutier and M. Yabuta, "Primitive divisors of certain elliptic divisibility sequences", *Acta Arith.* **151**:2 (2012), 165–190. MR Zbl

Received March 10, 2022. Revised September 4, 2023.

MATTEO VERZOBIO
INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA
KLOSTERNEUBURG
AUSTRIA

matteo.verzobio@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 325, No. 2, 2023
https://doi.org/10.2140/pjm.2023.325.353

POISSON MANIFOLDS OF STRONG COMPACT TYPE OVER 2-TORI

LUKA ZWAAN

We construct a new class of examples of Poisson manifolds of strong compact type. In particular, we show that all strongly integral affine circles and two-dimensional tori appear as the leaf space of a Poisson manifold of strong compact type.

1. Introduction

Like symplectic geometry, Poisson geometry started from the mathematical formalisation of classical mechanics. Roughly speaking, a Poisson manifold is a smooth manifold equipped with a *Poisson bracket* on its space of smooth functions, which allows one to formulate Hamiltonian dynamics. Examples of Poisson manifolds include symplectic manifolds and duals of Lie algebras, an early glimpse into the deep connection with symplectic geometry and Lie theory. Unlike symplectic manifolds, Poisson manifolds are very flexible in nature. For instance, every manifold admits a Poisson structure and there is no local classification of Poisson structures. For this reason it is common to restrict one's attention to specific classes of Poisson manifolds, where one can formulate deep results about their geometry. In this paper we are concerned with *Poisson manifolds of compact type* (PMCTs). PMCTs are the "compact objects" in Poisson geometry. They were first introduced in [Crainic and Fernandes 2005] and their role in the theory is analogous to the one played by compact Lie algebras in Lie theory. Just as there is the special class of compact *semisimple* Lie algebras among compact Lie algebras, there is an important distinguished class among PMCTs, namely that of Poisson manifolds of *strong* compact type (PMSCTs). A simple class of examples of PMSCTs is given by compact symplectic manifolds with finite fundamental group, but it is difficult to construct examples that are not symplectic. The first such example was given in [Martínez Torres 2014], building on work of

This work was partially supported by NSF grant DMS-2003223.
MSC2020: 53D17.
Keywords: Poisson geometry, symplectic groupoid, integral affine structure, K3 surface.

[Kotschick 2006]. There a regular PMSCT is constructed whose symplectic leaves are all diffeomorphic to a K3 surface and whose leaf space is diffeomorphic to a circle. One can form new PMSCTs by taking products of the aforementioned examples, but apart from these no other examples are known. In this paper we use the construction of [Martínez Torres 2014] to obtain new examples of PMSCTs. It is known that the leaf space of a PMSCT must be a compact integral affine orbifold and in the example of [Martínez Torres 2014] this is the "standard" integral affine structure on the circle. In this work we show that *all* strongly integral affine circles and two-dimensional tori can appear as the leaf space of a PMSCT.

In order to explain our main result, recall that a Poisson structure on a manifold M is a Lie bracket on $C^\infty(M)$ which is a derivation in each entry. Equivalently, a Poisson structure is a bivector $\pi \in \mathfrak{X}^2(M)$ satisfying $[\pi, \pi] = 0$. This is the definition we work with in this paper. Every Poisson manifold has a partition into symplectic manifolds. This *symplectic foliation* can be viewed as a singular foliation integrating the (singular) distribution $\pi^\#(T^*M) \subset TM$. If π has constant rank, this is actually a regular foliation. In this case the Poisson manifold is called *regular*.

The "global" objects in Poisson geometry are the so-called *symplectic groupoids*. A symplectic groupoid is a Lie groupoid $\mathcal{G} \rightrightarrows M$ carrying a multiplicative symplectic form $\Omega \in \Omega^2(\mathcal{G})$. A Poisson manifold (M, π) is called *integrable* if there exists some symplectic groupoid $(\mathcal{G} \rightrightarrows M, \Omega)$ for which the target map $\boldsymbol{t} : (\mathcal{G}, \Omega) \to (M, \pi)$ is a Poisson map (see [Crainic et al. 2021]). PMCTs are defined as those Poisson manifolds that are integrated by a source connected, Hausdorff symplectic groupoid having a certain compactness property. Contrary to the case of Lie groups and Lie algebras, there are multiple notions of compactness for Lie groupoids, namely a Lie groupoid $\mathcal{G} \rightrightarrows M$ is called

- *proper* if the anchor map $(\boldsymbol{s}, \boldsymbol{t}) : \mathcal{G} \to M \times M$ is proper;
- *source proper*, or $\boldsymbol{s}$-proper, if the source map is proper;
- *compact* if the space of arrows $\mathcal{G}$ is compact.

Accordingly, we say that (M, π) is of *proper/source proper/compact type* if it admits a source connected, Hausdorff symplectic groupoid of proper/source proper/compact type, respectively.

The types just defined depend on the choice of integration of (M, π). However, just like for Lie groups, there is a unique "largest" integration, namely the one with 1-connected source fibres. This is often called the *Weinstein groupoid*. We say that an integrable Poisson manifold has *strong* proper/source proper/compact type if its Weinstein groupoid is Hausdorff and has the corresponding type. As mentioned above, we will focus here on Poisson manifolds of strong compact type.

Unlike general Poisson manifolds, PMCTs have a rich geometry transverse to their associated symplectic foliation. For example, the leaf space of a regular PMCT inherits the structure of an integral affine orbifold. Roughly speaking this means that the leaf space has an orbifold atlas where the transitions are integral affine maps. The precise statement can be found in [Crainic et al. 2019a; 2019b], where many other properties of PMCTs are discussed.

As mentioned above, the first example of a PMSCT that is not symplectic was given in [Martínez Torres 2014]. The construction there is inspired by [Kotschick 2006], where nontrivial results on the geometry of K3 surfaces are used to construct a free symplectic circle action with contractible orbits. The orbit space of such an action is a PMSCT with smooth leaf space a circle endowed with its standard integral affine structure (that is, the one it inherits as a quotient of $\mathbb{R}$ by $\mathbb{Z}$ acting by translations). In general, it is not known whether any compact integral affine orbifold can appear as the leaf space of a PMSCT. On the one hand constructing strong PMCTs is a difficult problem on its own, and on the other not much is known about the classification of compact integral affine manifolds in dimension greater than two. The integral affine structures on a circle are easily classified, and the classification of integral affine structures on compact 2-dimensional manifolds was obtained in [Mishachev 1996; Sepe 2010]. The main result of this paper is the following.

Main theorem. *Any strongly integral affine circle or two-dimensional torus can be realised as the leaf space of a PMSCT.*

Here by a *strongly integral affine structure* we mean an integral affine structure with integral translational part (see [Sepe 2013, Remark 5.10] and Remark 4.2).

Our strategy to prove this result is as follows. Using the geometry of K3 surfaces one constructs a universal family of marked Kähler K3 surfaces (see Section 3) to which one can apply a general method from [Crainic et al. 2019b] to obtain PMSCTs. Using this construction together with the classification of integral affine 2-tori from [Mishachev 1996], one obtains examples of PMSCTs for all isomorphism classes of strongly integral affine 2-tori.

This paper is organised as follows. In Section 2, we provide some background on PMCTs and we recall the general method of constructing regular PMSCTs from [Crainic et al. 2019b]. In Section 3 we recall the relevant results on K3 surfaces that are needed for our construction. The resulting examples of PMSCTs have symplectic foliation a fibration over S^1 or $\mathbb{T}^2$ with typical fibre the smooth manifold underlying a K3 surface. The symplectic structures on the fibres vary in a controlled fashion which ensures that the Weinstein groupoid is a compact symplectic groupoid. Finally, Section 4 is dedicated to the actual constructions, which includes some lengthy computations. We treat the circle case first and this includes the original example from [Martínez Torres 2014]. Lastly, we construct the PMSCTs with leaf space the strongly integral affine 2-tori.

2. Background & general construction of PMSCTs

The construction we give below is based on two results on PMCTs:

(a) The leaf space carries an integral affine orbifold structure (see [Crainic et al. 2019b, Section 3]).

(b) The *linear variation theorem* (see [Crainic et al. 2019b, Sections 4–5]).

We briefly recall these results before giving the general construction. Here we only need to consider the case of 1-connected leaves. In this case the leaf space is smooth, since this assumption implies that the monodromy groupoid of the symplectic foliation is proper and has trivial isotropy groups. Then both (a) and (b) above simplify significantly.

2A. ***The integral affine structure on the leaf space.*** Recall that an integral affine structure on a manifold B is given by an atlas whose transition functions are integral affine maps. Equivalently, it is specified by a lattice $\Lambda \subset T^*B$ locally spanned by closed 1-forms.

Consider a regular, $\boldsymbol{s}$-connected, proper symplectic groupoid $(\mathcal{G}, \Omega) \rightrightarrows (M, \pi)$. As mentioned above, we assume that the associated symplectic foliation $\mathcal{F}_\pi$ has 1-connected leaves so that the leaf space B is a smooth manifold. We obtain a lattice $\widetilde{\Lambda} \subset \nu^*(\mathcal{F}_\pi)$ as follows:

(1) For each $x \in M$, the kernel of the exponential map $\mathfrak{g}_x \to \mathcal{G}_x$ gives a lattice in $\mathfrak{g}_x$.

(2) The isomorphism $\mathfrak{g}_x \cong \nu_x^*(\mathcal{F}_\pi)$ induced by Ω allows us to transport it to the conormal space.

This lattice descends to an integral affine structure $\Lambda \subset T^*B$ on B.

2B. ***The linear variation theorem.*** We assume now in addition that $(\mathcal{G}, \Omega) \rightrightarrows (M, \pi)$ is source proper. Denoting the symplectic leaf corresponding to $b \in B$ by (S_b, ω_b), we form the vector bundle

$$\mathcal{H}^2 := \bigsqcup_{b \in B} H^2(S_b, \mathbb{R}) \to B$$

and the lattice

$$\mathcal{H}^2_{\mathbb{Z}} := \bigsqcup_{b \in B} \operatorname{im}\bigl(H^2(S_b, \mathbb{Z}) \to H^2(S_b, \mathbb{R})\bigr)$$

inside it. Associated to this we have the *Gauss–Manin connection* ∇ on $\mathcal{H}^2$, uniquely determined by requiring the sections of $\mathcal{H}^2_{\mathbb{Z}}$ to be parallel. Note that π gives us a section $\varpi \in \Gamma(\mathcal{H}^2)$, $b \mapsto [\omega_b]$.

The Gauss–Manin connection allows us to study the variation of ϖ: parallel transport makes $\mathcal{H}^2$ into a $\Pi_1(B)$-representation and we define the *variation map* $\mathrm{var}_\varpi : \Pi_1(B) \to \mathcal{H}^2$ to be

$$[\gamma] \mapsto \gamma_*(\varpi_{\gamma(0)}) \in \mathcal{H}^2_{\gamma(1)}.$$

On the other hand, we also have the *linear variation map* $\mathrm{var}^{\mathrm{lin}}_\varpi : TB \to \mathcal{H}^2$ given by

$$v \mapsto \nabla_v \varpi$$

and the *affine variation map* $\mathrm{var}^{\mathrm{aff}}_\varpi := \varpi + \mathrm{var}^{\mathrm{lin}}_\varpi$.

The linear variation theorem relates the variation and affine variation maps by means of the *developing map* associated to the integral affine manifold (B, Λ). Associated to the lattice $\Lambda^* \subset TB$ we have a canonical flat connection on TB (not to be confused with ∇ above). This makes TB into a TB-representation, and since the connection is torsion-free the identity map $TB \to TB$ is an algebroid cocycle. The developing map is defined to be the groupoid cocycle $\mathrm{dev} : \Pi_1(B) \to TB$ integrating it.

Remark 2.1. One can show that after fixing $b \in B$ and a basis of Λ_b this boils down to the classical notion of developing map defined on the universal covering space (see [Crainic et al. 2019b, Section 4.2]):

$$\mathrm{dev}_b : \widetilde{B} \to T_b B \simeq \mathbb{R}^q.$$

We can now state the linear variation theorem as follows.

Theorem 2.2 [Crainic et al. 2019b, Theorem 4.4.2]. *One has a commutative diagram*

$$\begin{tikzcd} \Pi_1(B) \arrow[rr, "\mathrm{var}_\varpi"] \arrow[dr, "\mathrm{dev}"'] & & \mathcal{H}^2 \\ & TB \arrow[ur, "\mathrm{var}^{\mathrm{aff}}_\varpi"'] & \end{tikzcd}$$

This rather abstract formulation can locally be made explicit. Let $b_0 \in B$ and choose an integral affine chart (U, φ) centered at b_0 such that $\varphi(U)$ is convex and such that $M \to B$ trivialises over U. This induces a trivialisation $\Phi : \mathcal{H}^2|_U \cong U \times H^2(S_{b_0}, \mathbb{R})$. The chart induces an identification $T_{b_0}B \cong \mathbb{R}^q$ and allows us to consider "straight line" paths from $b \in U$ to b_0. Restricting to such paths the above diagram becomes

$$\begin{tikzcd} U \arrow[rr, "b \mapsto \Phi([\omega_b])"] \arrow[dr, "\varphi"'] & & H^2(S_{b_0}, \mathbb{R}) \\ & \mathbb{R}^q \arrow[ur, "v \mapsto {[\omega_{b_0}]} + \sum_i v_i c_i"'] & \end{tikzcd}$$

where $c_i \in H^2(S_{b_0}, \mathbb{Z})$ are the Chern classes of the torus bundle $\boldsymbol{s}^{-1}(x) \to S_{b_0}$, where $x \in S_{b_0}$ (see [Crainic et al. 2019b, Corollary 4.4.4]). This local formulation is reminiscent of the linear variation theorem from [Duistermaat and Heckman 1982]. In other words, Theorem 2.2 can be viewed as a global formulation and generalisation of the classical Duistermaat–Heckman theorem.

2C. ***The construction.*** The construction we describe in this section yields a PMSCT with 1-connected symplectic leaves, whose leaf space is a complete integral affine manifold. This means that the leaf space is a quotient of $\mathbb{R}^q$ by a free and proper action of a discrete group of integral affine transformations. Note that if the Markus conjecture holds true, then in fact every compact integral affine manifold is of this type (see [Goldman 2022, Section 8.6]). This allows us to give an explicit formulation of the linear variation, similar to the discussion following Theorem 2.2. The setup is as follows.

Let $E \to \mathbb{R}^q$ be a fibre bundle with typical fibre S, a compact 1-connected manifold, and assume that E admits a Poisson structure π_E whose symplectic leaves are precisely the fibres of this bundle. As in Section 2B we have

(i) the vector bundle $\mathcal{H}^2 \to \mathbb{R}^q$ whose fibres are the degree two cohomology groups of the symplectic leaves,

(ii) the lattice $\mathcal{H}^2_{\mathbb{Z}} \subset \mathcal{H}^2$ of integral cohomology,

(iii) the associated Gauss–Manin connection ∇ and

(iv) the section $\varpi \in \Gamma(\mathcal{H}^2)$ induced by π_E.

Next, let $\Gamma \subset \mathrm{Aff}_{\mathbb{Z}}(\mathbb{R}^q) = \{x \mapsto Ax + v \mid A \in \mathrm{GL}(q, \mathbb{Z}),\ v \in \mathbb{R}^q\}$ be a discrete group of integral affine transformations acting freely and properly on $\mathbb{R}^q$, and assume that there is a Poisson action of Γ on (E, π_E) making the projection $E \to \mathbb{R}^q$ equivariant. Then setting $M := E/\Gamma$ and $B := \mathbb{R}^q/\Gamma$, we get a (smooth) fibre bundle $p : M \to B$, again with typical fibre S, and a Poisson structure π on M whose leaves are the fibres of p. In other words, (M, π) is a regular Poisson manifold with leaf space B. Note also that B, being a quotient $\mathbb{R}^q/\Gamma$, naturally inherits an integral affine structure.

We can now state the general method of constructing PMSCTs. It is a reformulation of [Crainic et al. 2019b, Proposition 4.4.6].

Proposition 2.3. *Let $(M = E/\Gamma, \pi)$ be constructed as above. Assume that there exists a ∇-flat section $s \in \Gamma(\mathcal{H}^2)$ and linearly independent sections $c_1, \dots, c_q \in \Gamma(\mathcal{H}^2_{\mathbb{Z}})$ such that*

$$\varpi = s + \sum_{i=1}^{q} \mathrm{pr}^i \cdot c_i, \tag{2-1}$$

where $\mathrm{pr}^i : \mathbb{R}^q \to \mathbb{R}$ denotes projection onto the i-th coordinate. Then (M, π) is of strong $\boldsymbol{s}$-proper type and the induced integral affine structure on B agrees with the *one coming from the quotient* $\mathbb{R}^q/\Gamma$. *In particular*, *if* B *is compact then* (M,π) *is a PMSCT.*

Proof. Pulling back the integral affine structure on B along $p: M \to B$ yields a transverse integral affine structure on the symplectic foliation $\mathcal{F}_\pi$, i.e., a lattice in its conormal bundle. We denote this lattice by $\widetilde{\Lambda} \subset \nu^*(\mathcal{F}_\pi)$. The main point is that for all $x \in M$, the monodromy group $N_x(M,\pi)$ is equal to the lattice $\widetilde{\Lambda}_x$. In fact, using the description of the monodromy groups for regular Poisson manifolds as the variation of symplectic areas (see [Crainic and Fernandes 2004, Section 6]) this follows directly from (2-1). The integrability criteria for Poisson manifolds then imply that (M,π) is integrable. Furthermore, since S has trivial fundamental group, the isotropy groups of the Weinstein groupoid $\Sigma(M,\pi)$ fit into the exact sequence

$$\cdots \to \pi_2(S,x) \xrightarrow{\partial_x} \nu_x^*(\mathcal{F}_\pi) \to \Sigma_x(M,\pi) \to 0,$$

where ∂_x is the monodromy map at x. Therefore, from our previous discussion, it follows that $\Sigma_x(M,\pi) \simeq \nu_x^*(\mathcal{F}_\pi)/\widetilde{\Lambda}_x$, i.e., that the isotropy group at x is compact. Since this holds for all $x \in M$ and since S is also compact, this shows that the Weinstein groupoid is $\boldsymbol{s}$-proper.

Finally, since $\widetilde{\Lambda} \subset \nu^*(\mathcal{F}_\pi)$ is closed, Hausdorffness of the Weinstein groupoid follows from [Alcalde-Cuesta and Hector 1995, Theorem 1.1]. □

3. Background on K3 surfaces and the Poisson structure on the universal family

We start by listing some definitions and results concerning K3 surfaces, after which we describe the moduli spaces and universal families for K3 surfaces. These results can be found in [Barth et al. 1984]. Finally, following [Martínez Torres 2014], we use the Calabi–Yau theorem to turn the universal family into a Poisson manifold and the strong Torelli theorem to establish a Poisson action on it, setting us up to apply our construction.

Definition 3.1. A K3 surface is a compact, 1-connected complex surface with trivial canonical bundle.

Every K3 surface is Kähler (see [Siu 1983]). All K3 surfaces have the same underlying smooth manifold S (see [Barth et al. 1984, Corollary VIII.8.6]); this will be the model fibre used in Proposition 2.3. The intersection form on $H^2(S,\mathbb{Z})$ turns it into a lattice and this lattice is isomorphic to the aptly named $K3$ *lattice*, which we denote by $(L,(\cdot,\cdot))$. It is the unique even, unimodular lattice of signature (3, 19) (see [Barth et al. 1984, Proposition VIII.3.2(ii)]). Explicitly, we have $L = U^{\oplus 3} \oplus (-E_8)^{\oplus 2}$, where $U = \mathbb{Z}^{\oplus 2}$ with form given by $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ and $E_8 = \mathbb{Z}^{\oplus 8}$ with form given by the Cartan matrix of E_8; it is important for us that this form is positive definite. We also set $L_{\mathbb{R}} := L \otimes \mathbb{R}$ and $L_{\mathbb{C}} := L \otimes \mathbb{C}$; note that these are models for the real and complex cohomology, respectively.

3A. ***The Torelli theorem.***

Definition 3.2. Let X, X' be K3 surfaces. A $\mathbb{Z}$-module isomorphism $H^2(X', \mathbb{Z}) \to H^2(X, \mathbb{Z})$ is a *Hodge isometry* if

(i) it preserves the intersection form;

(ii) its $\mathbb{C}$-linear extension preserves the Hodge decomposition.

A Hodge isometry is called *effective* if its $\mathbb{R}$-linear extension maps some Kähler class of X' to one of X.

Effectiveness of a Hodge isometry is equivalent to requiring it to map the Kähler cone of X' to that of X (see [Barth et al. 1984, Proposition VIII.3.10]).

Theorem 3.3 (Torelli [Barth et al. 1984, Corollary VIII.11.4]). *Let X, X' be $K3$ surfaces. Then for any effective Hodge isometry $\varphi : H^2(X', \mathbb{Z}) \to H^2(X, \mathbb{Z})$ there exists a unique biholomorphism $f : X \to X'$ such that $f^* = \varphi$.*

This result is ultimately used to obtain the action in Proposition 2.3.

3B. ***Moduli spaces and universal families.*** There are two moduli spaces and corresponding families for K3 surfaces: one takes into account the Kähler structure and the other only considers the complex structure. We start now with the latter.

Definition 3.4. A *marked $K3$ surface* is a pair (X, φ) consisting of a K3 surface X and a *marking* φ, i.e., an isometry $\varphi : H^2(X, \mathbb{Z}) \to L$. Two marked K3 surfaces are *equivalent* if there is a bihomolorphism between them intertwining the markings. The *moduli space of marked $K3$ surfaces* is the set of equivalence classes:

$$M_1 := \{(X, \varphi)\}/\sim .$$

It follows immediately from the definition that any K3 surface admits, up to scalar multiplication, a unique nowhere vanishing holomorphic 2-form. In fact, one can show that, again up to scalar multiplication, there is a bijection between complex structures on S and closed, complex 2-forms $\sigma \in \Omega^2(S, \mathbb{C})$ satisfying $\sigma \wedge \sigma = 0$ and $\sigma \wedge \bar{\sigma} > 0$. This motivates the following definitions. We will use the same letter to denote a marking $\varphi : H^2(X, \mathbb{Z}) \to L$ and the induced maps $\varphi : H^2(X, \mathbb{R}) \to L_{\mathbb{R}}$ and $\varphi : H^2(X, \mathbb{C}) \to L_{\mathbb{C}}$.

Definition 3.5. The *period domain* is given by

$$\Omega := \{[\sigma] \in \mathbb{P}(L_{\mathbb{C}}) \mid (\sigma, \sigma) = 0, (\sigma, \bar{\sigma}) > 0\}.$$

We define the *period map* $\tau_1 : M_1 \to \Omega$ by

$$[(X, \varphi)] \mapsto [\varphi(\sigma_X)],$$

where σ_X is a nowhere vanishing holomorphic 2-form on X.

Theorem 3.6 [Barth et al. 1984, Theorem VIII.12.1]. *The moduli space M_1 admits the structure of a* 20*-dimensional complex manifold such that the period map $\tau_1 : M_1 \to \Omega$ becomes a surjective local biholomorphism. Furthermore, there exists a universal family $\mathcal{U} \to M_1$ of marked $K3$ surfaces.*

Remark 3.7. Recall that a family is universal if any other family is locally the pullback of it by a unique map (see [Barth et al. 1984, Section I.10]). The fibre of the universal family $\mathcal{U} \to M_1$ over any $t \in M_1$ is a marked K3 surface (X_t, φ_t) such that $[(X_t, \varphi_t)] = t$. Furthermore, these markings vary smoothly in the sense that they induce local trivialisations of the bundle $\bigcup_{t\in M_1} H^2(X_t, \mathbb{R})$.

There are still some inconveniences present here. It can be shown that M_1 is not Hausdorff, and that the period map τ_1 is not injective (see [Barth et al. 1984, Remark VIII.12.2]). These problems disappear when taking into account the Kähler structure.

Definition 3.8. We define M_2 to be the subset of the bundle

$$\bigsqcup_{t\in M_1} H^2(X_t, \mathbb{C})$$

consisting of all Kähler classes.

It can be shown that M_2 is a real-analytic manifold of dimension 60 (see [Barth et al. 1984, Lemma VIII.9.3] and its proof). One should think of a point in M_2 as an equivalence class of marked K3 surfaces together with a specified Kähler class. Note that there is a projection map $\mathrm{pr} : M_2 \to M_1$.

Inspired by some analysis of the Kähler cone of K3 surfaces (see [Barth et al. 1984, Sections VIII.3 and VIII.9]) one makes the following definitions.

Definition 3.9. Set

$$K\Omega := \{(k, [\sigma]) \in L_{\mathbb{R}} \times \Omega \mid (k, k) > 0, (k, \sigma) = 0\}.$$

The *refined period domain* is then given by

$$K\Omega^0 := \{(k, [\sigma]) \in K\Omega \mid (k, d) \neq 0 \text{ for all } d \in L \text{ such that } (d, d) = -2,\ (d, \sigma) = 0\}.$$

The *refined period map* $\tau_2 : M_2 \to K\Omega^0$ is defined as

$$(t, k) \mapsto (\varphi_t(k), \tau_1(t)).$$

Theorem 3.10 [Barth et al. 1984, Theorems VIII.12.3 and VIII.14.1]. *The refined period map is a diffeomorphism.*

We set $K\mathcal{U} := (\mathrm{pr} \circ \tau_2^{-1})^*\mathcal{U}$. This is a real-analytic family (i.e., fibre bundle) over $K\Omega^0$ with extra data attached: the fibre over $(k, [\sigma])$ is a triple (X, φ, ω) consisting of a K3 surface X, a marking $\varphi : H^2(X, \mathbb{Z}) \to L$ and a Kähler class $\omega \in H^2(X, \mathbb{R})$

such that $\varphi(\omega) = k$. These markings vary smoothly in the same sense as before, and hence so do the Kähler classes.

The family $K\mathcal{U} \to K\Omega^0$ is universal for real-analytic "marked Kähler K3 families", i.e., real-analytic families of K3 surfaces equipped with smoothly varying markings and Kähler classes.

3C. ***The Poisson structure.*** Recall the following special version of the Calabi–Yau theorem (see, e.g., [Barth et al. 1984, Theorem I.15.1]).

Theorem 3.11. *Let X be a compact complex manifold with vanishing first Chern class. Then for any Kähler class $\omega \in H^2(X, \mathbb{R})$ there exists a unique Ricci flat Kähler metric whose Kähler form belongs to ω.*

This theorem applies in particular to K3 surfaces, and thus we can use it to endow the fibres of $K\mathcal{U} \to K\Omega^0$ with smoothly varying Kähler forms, turning it into a Poisson manifold (see also [Martínez Torres 2014, Section 2.1.3]).

Corollary 3.12. *The family $K\mathcal{U}$ admits a regular Poisson structure $\pi_{K\mathcal{U}}$ whose symplectic leaves are the fibres of $K\mathcal{U} \to K\Omega^0$. Moreover the symplectic form on the fibre X over $(k, [\sigma])$ with marking φ is the Kähler form associated to the unique Ricci flat Kähler metric on X with Kähler class $\varphi^{-1}(k)$.*

3D. ***The action.*** We will construct an action on $K\mathcal{U}$ by the group $O(L)$ of isometries of the K3 lattice. Note that there is an obvious induced action of $O(L)$ on $K\Omega^0$.

Proposition 3.13. *There is a Poisson action of $O(L)$ on $(K\mathcal{U}, \pi_{K\mathcal{U}})$ with respect to which the projection $K\mathcal{U} \to K\Omega^0$ is equivariant.*

Proof. Fix $\gamma \in O(L)$ and $p \in K\Omega^0$. Using the notation from above, denote the triple over p by $(X_p, \varphi_p, \omega_p)$ and similarly for $\gamma(p)$. It is easy to see that

$$\varphi_p^{-1} \circ \gamma^{-1} \circ \varphi_{\gamma(p)} : H^2(X_{\gamma(p)}, \mathbb{Z}) \to H^2(X_p, \mathbb{Z})$$

is an effective Hodge isometry, so that by Theorem 3.3 we obtain a biholomorphism $f_\gamma^p : X_p \to X_{\gamma(p)}$. The universality of the family then gives neighbourhoods U and V of p and of $\gamma(p)$ respectively and an isomorphism $(\Psi, \psi) : K\mathcal{U}|_U \to K\mathcal{U}|_V$ extending f_γ^p: through the biholomorphism f_γ^p, $K\mathcal{U}$ becomes a deformation of X_p at two basepoints, p and $\gamma(p)$. Since $K\mathcal{U}$ is universal, these two deformations are locally isomorphic. Writing $\Psi_q : X_q \to X_{\psi(q)}$ for the fiberwise maps, it then follows that for all $q \in U$ we have that

$$\Psi_q^* = \varphi_q^{-1} \circ \gamma^{-1} \circ \varphi_{\psi(q)} : H^2(X_{\psi(q)}, \mathbb{Z}) \to H^2(X_q, \mathbb{Z}).$$

This implies first of all that $\psi = \gamma|_U$, from which it follows that $\Psi_q = f_\gamma^q$, since biholomorphisms of K3 surfaces are uniquely determined by their induced maps on degree 2 integral cohomology (see [Barth et al. 1984, Proposition VIII.11.3]). Thus these fibrewise biholomorphisms f_γ^p, $p \in K\Omega^0$, together form an automorphism $F_\gamma : K\mathcal{U} \to K\mathcal{U}$. It is immediate from the above construction that $F_{\mathrm{id}} = \mathrm{id}$, and from the uniqueness part of Theorem 3.3 it follows that $F_{\gamma\circ\gamma'} = F_\gamma \circ F_{\gamma'}$ for all $\gamma, \gamma' \in O(L)$, meaning that we have an action of $O(L)$ on $K\mathcal{U}$. This action makes $K\mathcal{U} \to K\Omega^0$ equivariant by construction. Finally, from the uniqueness part of the Calabi–Yau theorem it follows that each f_γ^p preserves the symplectic forms on the fibres, meaning that the action is by Poisson maps. □

4. The examples

From our work in Section 3 we have a Poisson manifold $(K\mathcal{U}, \pi_{K\mathcal{U}})$ with leaf space $K\Omega^0$ such that

(i) the cohomology classes of the symplectic forms on the leaves are described in terms of the leaf space $K\Omega^0$ (Corollary 3.12);

(ii) the natural action of $O(L)$ on $K\Omega^0$ lifts to a Poisson action on $(K\mathcal{U}, \pi_{K\mathcal{U}})$ (Proposition 3.13).

In order to apply the construction described in Section 2, we need to find a suitable embedding $\mathbb{R}^q \hookrightarrow K\Omega^0$ and a suitable subgroup $\Gamma \subset O(L)$. We rephrase Proposition 2.3 in the current setting in order to make this more precise. For a different version of this result see also [Martínez Torres 2014, Theorem 1].

Corollary 4.1. *Assume that we have an embedding $f : \mathbb{R}^q \to K\Omega^0$ and a subgroup $\Gamma \subset O(L)$ such that*

(i) *there exist $a \in L_\mathbb{R}$ and linearly independent $a_1, \dots, a_q \in L$ such that the $L_\mathbb{R}$-component of f has the form*
$$(x_1, \dots, x_q) \mapsto a + \sum_{i=1}^q x_i a_i;$$

(ii) *the action of Γ on $K\Omega^0$ preserves the image of f;*

(iii) *the induced action on $\mathbb{R}^q$ is free, proper and by integral affine maps.*

Then $M := f^*K\mathcal{U}/\Gamma$ with the Poisson structure induced from $\pi_{K\mathcal{U}}$ is a Poisson manifold of strong $\mathbf{s}$-proper type with leaf space $B := \mathbb{R}^q/\Gamma$. If B is compact, M is a PMSCT.

Remark 4.2. We can now explain why our construction leads to PMSCTs with *strongly* integral affine leaf spaces. On the one hand, because of Theorem 2.2, we are forced to consider embeddings with integral variation, i.e., the a_i must lie in the integral lattice L. On the other hand, to apply Theorem 3.3 we need to consider isometries of integral cohomology, i.e., we need to act by elements of $O(L)$. These two technical limitations together only allow for strongly integral affine leaf spaces in the examples.

Remark 4.3. At the level of the symplectic groupoid, one can see that the leaf space being strongly integral affine implies that the restriction of the symplectic form to the identity component of the isotropy (a torus bundle) lies in the integral cohomology. See [Sepe 2013, Remark 5.10].

We now recall the classification of strongly integral affine structures for S^1 and $\mathbb{T}^2$.

Theorem 4.4. *The strongly integral affine circles are, up to isomorphism, the quotients* $\mathbb{R}/\mathbb{Z}$ *where the* $\mathbb{Z}$*-action is generated by* $x \mapsto x + p$, *for a fixed* $p \in \mathbb{Z}_{\geq 1}$.

Proof. It is easy to see that all integral affine circles are complete. Hence, it suffices to classify, up to conjugation, embeddings $\mathbb{Z} \to \mathrm{Aff}_{\mathbb{Z}}(\mathbb{R})$ inducing free and proper actions. These are precisely the actions generated by $x \mapsto x + a$ with $a > 0$. Restricting to *strongly* integral affine circles yields the result. ☐

Theorem 4.5. *The strongly integral affine* 2*-tori, up to isomorphism, are quotients* $\mathbb{R}^2/\mathbb{Z}^2$, *where the* $\mathbb{Z}^2$*-actions fall into one of the following types*:

(I) *An action generated by* $(x, y) \mapsto (x + p, y)$ *and* $(x, y) \mapsto (x, y + q)$, *where* $p, q \in \mathbb{Z}_{\geq 1}$ *and* $p|q$.

(II) *An action generated by* $(x, y) \mapsto (x + p, y)$ *and* $(x, y) \mapsto (x + ny, y + q)$, *where* $n, p, q \in \mathbb{Z}_{\geq 1}$.

Proof. The classification of all integral affine structures on 2-tori is given in [Mishachev 1996, Theorem A]. Restricting to strongly integral affine structures and using the Smith normal form for matrices with integer entries to simplify the possibilities from type (I) yields the above classification. ☐

Remark 4.6. The integral affine 2-tori of type (I) are (isomorphic to) products of integral affine circles. Thus to find examples of PMSCTs with leaf space of this type one can simply take products of PMSCTs with leaf space S^1, constructed in Section 4A. This yields Poisson manifolds of dimension 10 whose leaves are products of K3 surfaces. However, the examples we construct in Section 4B are six-dimensional Poisson manifolds with K3 surfaces as symplectic leaves and thus result in “smaller” examples.

Remark 4.7. Continuing the previous remark, note that by taking products we can also realise *some* higher-dimensional integral affine tori as the leaf space of a PMSCT, namely those that are isomorphic to a product of some of the integral affine circles and 2-tori classified above.

Before we move on to the examples, we establish some notation. Recall that $L = U^{\oplus 3} \oplus (-E_8)^{\oplus 2}$. We denote the standard bases of the three copies of U by $\{u, v\}$, $\{x, y\}$ and $\{z, t\}$, so that $(u, v) = (x, y) = (z, t) = 1$ with all other combinations yielding zero. Recall also that $-E_8$ is even and negative definite. Finally, let $\{e_1, \dots, e_8\}$ be a set of real numbers such that the set

$$\{1, e_1, \dots, e_8, e_1^2, e_1e_2, \dots, e_7^2, e_7e_8, e_8^2\}$$

consisting of $1, e_1, \dots, e_8$ and their pairwise products is linearly independent over the integers, or equivalently the rationals. The existence of such a set is guaranteed by [Mordell 1953]. We then set $e := (e_1, \dots, e_8) \in (-E_8)_{\mathbb{R}}$, scaling if necessary such that $|(e, e)| \leq \frac{1}{2}$, and we set $a := (0, e)$, $b := (e, 0) \in (-E_8)_{\mathbb{R}}^{\oplus 2} \subset L_{\mathbb{R}}$.

Let us outline the strategy for the examples below. In each case, we start by defining f and Γ. It is fairly straightforward to check items (ii) and (iii) from Corollary 4.1 and that the image of f is contained in $K\Omega$. It then remains to show that it is actually contained in $K\Omega^0$. This is the more involved part of the computations.

4A. ***The PMSCTs with leaf space the circle.*** We will construct a PMSCT whose leaf space is a strongly integral affine circle, i.e., we want the action of $\mathbb{Z}$ on $\mathbb{R}$ generated by $x \mapsto x + p$ with $p \in \mathbb{Z}_{\geq 1}$. The case $p = 1$ is the one treated in [Martínez Torres 2014] and the computations carried out below for general p are an obvious generalisation of the computations there.

Consider the map $f : \mathbb{R} \to L_{\mathbb{R}} \times \mathbb{P}(L_{\mathbb{C}})$ defined by

$$s \mapsto \big(2u + v + sy, [x - su + 2y + a + i(z + 2t + b)]\big)$$

and the map $\varphi : L \to L$ defined by $u \mapsto u$, $v \mapsto v + py$, $x \mapsto x - pu$, $y \mapsto y$ on the first two copies of U and as the identity on the other summands of L. It is easily checked that φ is an isometry and that

$$\varphi \cdot f(s) = f(s + p).$$

This implies that the image of f is invariant under the action of $\Gamma := \langle \varphi \rangle$, and also that the induced action on $\mathbb{R}$ is the one we need.

To show that the image of f is contained in $K\Omega$, let $s \in \mathbb{R}$. Setting $f_1(s) = 2u + v + sy$, $f_2(s) = x - su + 2y + a$ and $f_3(s) = z + 2t + b$, we see that

$$\begin{aligned}
(f_2(s), f_2(s)) &= (x - su + 2y + a, x - su + 2y + a) \\
&= 4(x, y) + (a, a) \\
&= 4 + (e, e) \geq \tfrac{7}{2} > 0, \\
(f_3(s), f_3(s)) &= (z + 2t + b, z + 2t + b) \\
&= 4(z, t) + (b, b) \\
&= 4 + (e, e) \geq \tfrac{7}{2} > 0, \\
(f_2(s), f_3(s)) &= (x - su + 2y + a, z + 2t + b) \\
&= 0.
\end{aligned}$$

These computations imply that $[f_2(s)+if_3(s)] \in \Omega$. Since

$$\begin{aligned}
(f_1(s), f_1(s)) &= (2u+v+sy, 2u+v+sy) = (2u, v)+(v, 2u) = 4 > 0,\\
(f_1(s), f_2(s)) &= (2u+v+sy, x-su+2y+a)\\
&= -s(v,u)+s(y,x) = -s+s = 0,\\
(f_1(s), f_3(s)) &= (2u+v+sy, z+2t+b) = 0,
\end{aligned}$$

we see that $f(s) \in K\Omega$.

It remains to check that $f(s) \in K\Omega^0$ for all $s \in \mathbb{R}$.

Proof. Assume that we have $d \in L$ such that $(d,d) = -2$ and $(d, f_1(s)) = (d, f_2(s)) = (d, f_3(s)) = 0$. We need to find a contradiction. Let us write

$$d = Au + Bv + Cx + Dy + Ez + Ft + d_1 + d_2,$$

with $A, \dots, F \in \mathbb{Z}$ and d_i in the i-th copy of $-E_8$. Since E_8 is even and positive definite, we can write $(d_i, d_i) = -2n_i$, for $n_i \in \mathbb{Z}_{\geq 0}$. The above conditions then translate into three equations:

$$AB + CD + EF = n_1 + n_2 - 1, \tag{4-1}$$

$$2B + A + Cs = 0, \tag{4-2}$$

$$D - Bs + 2C + (d_2, e) = 0, \tag{4-3}$$

$$F + 2E + (d_1, e) = 0. \tag{4-4}$$

This is where the seemingly strange choice of e comes in. There exist $k_1, \dots, k_8 \in \mathbb{Z}$ such that $(d_1, e) = \sum_i k_i e_i$ and since $\{1, e_1, \dots, e_8\}$ is linearly independent over the integers by choice of e, it follows from (4-4) that we must have $F + 2E = k_1 = \cdots = k_8 = 0$. Since the bilinear form on $-E_8$ is nondegenerate, it follows that $d_1 = 0$ and thus that $n_1 = 0$.

Case $C = 0$: Equation (4-2) yields $2B + A = 0$, and (4-1) becomes

$$2B^2 + 2E^2 = 1 - n_2.$$

This implies that $B = E = 0$ and $n_2 = 1$. But then $d_2 \neq 0$ and (4-3) becomes

$$D + (d_2, e) = 0,$$

which together with $d_2 \neq 0$ contradicts the "linear independence" assumption on e.

Case $C \neq 0$: From (4-2) we get

$$s = -\frac{2B + A}{C},$$

and substituting this into (4-3) yields

$$AB + CD = -2C^2 - 2B^2 - (d_2, e).$$

Combining this with (4-1) gives

$$2B^2 + 2C^2 + 2E^2 + C(d_2, e) = 1 - n_2.$$

From the properties of e we get $Cd_2 = 0$, implying that $d_2 = 0$ and thus also that $n_2 = 0$, so that we are left with

$$2B^2 + 2C^2 + 2E^2 = 1,$$

which is absurd since $B, C, E \in \mathbb{Z}$. □

4B. ***The PMSCTs with leaf space a torus of type (I).*** Here we construct a PMSCT with leaf space the torus $\mathbb{T}^2$ with an integral affine structure of type (I). This means that we want the action of $\mathbb{Z}^2$ on $\mathbb{R}^2$ generated by $(x, y) \mapsto (x + p, y)$ and $(x, y) \mapsto (x, y + q)$, with $p, q \in \mathbb{Z}_{\geq 1}$.

Consider the map $f : \mathbb{R}^2 \to L_{\mathbb{R}} \times \mathbb{P}(L_{\mathbb{C}})$ defined by

$$(s, r) \mapsto \big(2u + v + sy + rt, [x - su + 2y + a + i(z - ru + 2t + b)]\big),$$

the map $\varphi : L \to L$ as in the previous example and the map $\psi : L \to L$ defined by $u \mapsto u$, $v \mapsto v + qt$, $x \mapsto x$, $y \mapsto y$, $z \mapsto z - qu$, $t \mapsto t$ on two copies of U and as the identity on the other summands of L. It is easily checked that these are isometries and that

$$\varphi \cdot f(s, r) = f(s + p, r),$$
$$\psi \cdot f(s, r) = f(s, r + q).$$

This implies that the image of f is invariant under the action of $\Gamma := \langle \varphi, \psi \rangle$, and also that the induced action on $\mathbb{R}^2$ is as desired.

To show that the image of f is contained in $K\Omega$, let f_1, f_2, f_3 be the three “components” of f, as before, and let $(s, r) \in \mathbb{R}^2$. We compute

$$\begin{aligned}
(f_2(s, r), f_2(s, r)) &= (x - su + 2y + a, x - su + 2y + a) \\
&= 4(x, y) + (a, a) \\
&= 4 + (e, e) \geq \tfrac{7}{2} > 0, \\
(f_3(s, r), f_3(s, r)) &= (z - ru + 2t + b, z - ru + 2t + b) \\
&= 4(z, t) + (b, b) \\
&= 4 + (e, e) \geq \tfrac{7}{2} > 0, \\
(f_2(s, r), f_3(s, r)) &= (x - su + 2y + a, z - ru + 2t + b) \\
&= 0
\end{aligned}$$

and conclude that $[f_2(s,r)+if_3(s,r)]\in\Omega$. Also,

$$\begin{aligned}(f_1(s,r),\,f_1(s,r)) &= (2u+v+sy+rt,\,2u+v+sy+rt)\\ &= (2u,\,v)+(v,\,2u)=4>0,\\ (f_1(s,r),\,f_2(s,r)) &= (2u+v+sy+rt,\,x-su+2y+a)\\ &= -s(u,\,v)+s(x,\,y)=-s+s=0,\\ (f_1(s,r),\,f_3(s,r)) &= (2u+v+sy+rt,\,z-ru+2t+b)\\ &= -r(u,\,v)+r(z,\,t)=-r+r=0\end{aligned}$$

implies that $f(s,r)\in K\Omega$.

It remains to check that $f(s,r)\in K\Omega^0$ for all $(s,r)\in\mathbb{R}^2$.

Proof. Let $d\in L$ such that $(d,d)=-2$ and $(d,\,f_1(s,r))=(d,\,f_2(s,r))=(d,\,f_3(s,r))=0$ and as before write

$$d=Au+Bv+Cx+Dy+Ez+Ft+d_1+d_2,$$

and $(d_i,d_i)=-2n_i$ for $n_i\in\mathbb{Z}_{\geq 0}$. We need to find a contradiction. The relevant equations now become

$$AB+CD+EF=n_1+n_2-1, \tag{4-5}$$

$$2B+A+Cs+Er=0, \tag{4-6}$$

$$D-Bs+2C+(d_2,e)=0, \tag{4-7}$$

$$F-Br+2E+(d_1,e)=0. \tag{4-8}$$

Case $B=0$: The assumptions on e, together with (4-7) and (4-8), imply that $D+2C=F+2E=0$ and $d_1=d_2=0$, so that $n_1=n_2=0$. But then (4-5) becomes

$$2C^2+2E^2=1,$$

which is impossible.

Case $B\neq 0$: From (4-7) and (4-8) we get

$$s=\frac{D+2C+(d_2,e)}{B},\qquad r=\frac{F+2E+(d_1,e)}{B}.$$

Substituting this into (4-6) gives

$$AB+CD+EF=-2B^2-2C^2-2E^2-C(d_2,e)-E(d_1,e),$$

and combining this with (4-5) we obtain

$$2B^2+2C^2+2E^2+C(d_2,e)+E(d_1,e)=1-n_1-n_2.$$

The assumptions on e imply that $Cd_2 + Ed_1 = 0$, so that this becomes

$$2B^2 + 2C^2 + 2E^2 = 1 - n_1 - n_2.$$

This is impossible under the assumption $B \neq 0$, since $n_i \in \mathbb{Z}_{\geq 0}$. $\square$

4C. ***The PMSCTs with leaf space a torus of type (II).*** In this example we will construct a PMSCT whose leaf space is a torus with an induced integral affine structure of type (II), namely one induced by the action of $\mathbb{Z}^2$ on $\mathbb{R}^2$ generated by $(x, y) \mapsto (x + p, y)$ and $(x, y) \mapsto (x + ny, y + q)$, where $n, p, q \in \mathbb{Z}_{\geq 1}$.

Consider the map $f : \mathbb{R}^2 \to L_{\mathbb{R}} \times \mathbb{P}(L_{\mathbb{C}})$ defined by

$$\begin{aligned}&(s,r)\\ &\mapsto \big(2u+v+sy+rt, [qx+(nr^2-qs)u-nrz+2qy+a+i(z-ru+2q^2t+2nqry+b)]\big),\end{aligned}$$

the map $\varphi : L \to L$ defined as before and the map $\psi : L \to L$ defined by $u \mapsto u$, $v \mapsto v + qt$, $x \mapsto x - nz + qnu$, $y \mapsto y$, $z \mapsto z - qu$, $t \mapsto t + ny$ on the copies of U and the identity on the other summands of L. It is easily checked that these are isometries and that

$$\begin{aligned}\varphi \cdot f(s,r) &= f(s+p,r),\\ \psi \cdot f(s,r) &= f(s+nr, r+q).\end{aligned}$$

This implies that the image of f is invariant under the action of $\Gamma := \langle \varphi, \psi \rangle$, and also that the induced action on $\mathbb{R}^2$ is the desired one. To show that the image of f is contained in $K\Omega$, denote once more by f_1, f_2, f_3 the "components" of f, and let $(s, r) \in \mathbb{R}^2$. Since

$$\begin{aligned}(f_2(s,r), f_2(s,r)) &= \big(qx+(nr^2-qs)u-nrz+2qy+a, qx+(nr^2-qs)u-nrz\\ &\qquad\qquad +2qy+a\big)\\ &= 4q^2(x,y)+(a,a)\\ &= 4q^2+(e,e) \geq \tfrac{7}{2} > 0,\\ (f_3(s,r), f_3(s,r)) &= (z-ru+2q^2t+2nqry+b, z-ru+2q^2t+2nqry+b)\\ &= 4q^2(z,t)+(b,b)\\ &= 4q^2+(e,e) \geq \tfrac{7}{2} > 0,\\ (f_2(s,r), f_3(s,r)) &= (qx+(nr^2-qs)u-nrz+2qy+a, z-ru+2q^2t+2nqry+b)\\ &= 2nq^2r(x,y)-2nq^2r(z,t) = 2nq^2r-2nq^2r = 0,\end{aligned}$$

we get that $[f_2(s,r)+if_3(s,r)]\in\Omega$. The computations

$$\begin{aligned}(f_1(s,r),\, f_1(s,r)) &= (2u+v+sy+rt, 2u+v+sy+rt)\\ &= (2u,v)+(v,2u)=4>0,\\ (f_1(s,r),\, f_2(s,r)) &= (2u+v+sy+rt, qx+(nr^2-qs)u-nrz+2qy+a)\\ &= (nr^2-qs)(u,v)+qs(x,y)-nr^2(z,t)\\ &= nr^2-qs+qs-nr^2=0,\\ (f_1(s,r),\, f_3(s,r)) &= (2u+v+sy+rt, z-ru+2q^2t+2nqry+b)\\ &= -r(u,v)+r(z,t)=-r+r=0\end{aligned}$$

show that $f(s,r)\in K\Omega$.

It remains to show that $f(s,r)\in K\Omega^0$ for all $(s,r)\in\mathbb{R}^2$.

Proof. Let $d\in L$ such that $(d,d)=-2$ and $(d,\, f_1(s))=(d,\, f_2(s))=(d,\, f_3(s))=0$. Like before we write

$$d=Au+Bv+Cx+Dy+Ez+Ft+d_1+d_2,$$

and we set $(d_i,d_i)=-2n_i$ with $n_i\in\mathbb{Z}_{\geq 0}$. The goal is to find a contradiction. The main equations are now

$$AB+CD+EF=n_1+n_2-1, \tag{4-9}$$

$$2B+A+Cs+Er=0, \tag{4-10}$$

$$Dq+B(nr^2-qs)-Fnr+2Cq+(d_2,e)=0, \tag{4-11}$$

$$F-Br+2Eq^2+2Cnqr+(d_1,e)=0. \tag{4-12}$$

Case $B-2Cnq=0$: Equation (4-12) tells us that $d_1=0$ and $F+2Eq^2=0$.

Subcase $C=0$: This implies that $B=0$, so that (4-9) becomes

$$2E^2q^2=1-n_2.$$

This is only possible if $E=0$ and $n_2=1$, but then also $F=0$ and (4-11) becomes

$$Dq+(d_2,e)=0,$$

which would imply that $d_2=0$, contradicting $n_2=1$.

Subcase $C\neq 0$: Equation (4-10) tells us that

$$s=-\frac{2B+A+Er}{C},$$

and with (4-11) we obtain

$$2Cn^2qr^2-2Fnr+2Anq^2+C(8n^2q^3+2q)+Dq+(d_2,e)=0.$$

Since $C, n, q \neq 0$ and $r \in \mathbb{R}$, we must have that

$$F^2 \geq 2Cq\big[2Anq^2 + C(8n^2q^3 + 2q) + Dq + (d_2, e)\big].$$

But $F = -2Eq^2$ and $B = 2Cnq$, so combining this with (4-9) yields

$$q(1 - n_2) \geq C^2(8n^2q^3 + 2q) + C(d_2, e).$$

Since $C \neq 0$, this is certainly impossible when C and (d_2, e) have the same parity. So let us assume that they have opposite parity, so that the equation becomes

$$q(1 - n_2) \geq C^2(8n^2q^3 + 2q) - |C| \cdot |(d_2, e)|. \tag{4-13}$$

Now both d_2 and e lie in the same copy of $-E_8$, and since $(\cdot, \cdot)$ is negative definite on $-E_8$ we can use the Cauchy–Schwarz inequality to obtain

$$|(d_2, e)| \leq \sqrt{|(d_2, d_2)| \cdot |(e, e)|} = \sqrt{2 \cdot |(e, e)| n_2} \leq \sqrt{n_2},$$

using that we chose e such that $|(e, e)| \leq \frac{1}{2}$. Now, in order for (4-13) to hold we certainly must have

$$C^2(8n^2q^3 + 2q) - \sqrt{n_2} \cdot |C| + qn_2 - q \leq 0$$

and it is easily seen that this is not possible for $0 \neq C \in \mathbb{Z}$.

Case $B - 2Cnq \neq 0$: We immediately distinguish two cases: $B = 0$ and $B \neq 0$.

Subcase $B = 0$: We claim that $F \neq 0$. Indeed, if we had $F = 0$, equation (4-11) would become

$$Dq + 2Cq + (d_2, e) = 0,$$

meaning that $d_2 = 0$, so $n_2 = 0$, and $D + 2C = 0$. But then (4-9) becomes

$$2C^2 = 1 - n_1,$$

which can only hold if $C = 0$ and $n_1 = 1$. But then (4-12) becomes

$$2Eq^2 + (d_1, e) = 0,$$

which implies $d_1 = 0$, contradicting $n_1 = 1$. So we see indeed that $F \neq 0$. But then (4-11) and (4-12) yield

$$r = -\frac{F + 2Eq^2 + (d_1, e)}{2Cnq} = \frac{Dq + 2Cq + (d_2, e)}{Fn}.$$

This becomes

$$2CDnq^2 + 4C^2nq^2 + 2Cnq(d_2, e) + F^2n + 2EFnq^2 + Fn(d_1, e) = 0,$$

and the assumptions on e imply that $2Cqd_2 + Fd_1 = 0$ and

$$2CDnq^2 + 4C^2nq^2 + F^2n + 2EFnq^2 = 0.$$

Since $B = 0$, combining this with (4-9) we obtain

$$4C^2nq^2 + F^2n = 2nq^2(1 - n_1 - n_2).$$

Both C and F are nonzero, meaning that this is impossible.

Subcase $B \neq 0$: We can write

$$r = \frac{F + 2Eq^2 + (d_1, e)}{B - 2Cnq}, \qquad s = \frac{Dq + Bnr^2 - Fnr + 2Cq + (d_2, e)}{Bq}.$$

This yields

$$s = \frac{(B - 2Cnq)^2(2Cq + Dq + (d_2, e)) + Bn(F + 2Eq^2 + (d_1, e))^2}{Bq(B - 2Cnq)^2} - \frac{Fn(B - 2Cnq)(F + 2Eq^2 + (d_1, e))}{Bq(B - 2Cnq)^2}$$

and substituting this into (4-10) and using the assumptions on e (actually, finally using them to their full potential), this reduces to

$$\begin{aligned} 0 = {} & 2B^2q(B-2Cnq)^2 + ABq(B-2Cnq)^2 \\ & + C\big((B-2Cnq)^2(2Cq+Dq) + Bn(F+2Eq^2)^2 - Fn(B-2Cnq)(F+2Eq^2)\big) \\ & + BEq(B-2Cnq)(F+2Eq^2). \end{aligned}$$

Some rewriting turns this into

$$\begin{aligned} 0 = {} & q(B - 2Cnq)^2\big(2B^2 + 2C^2 + AB + CD\big) + BCn(F + 2Eq^2)^2 \\ & - CFn(B - 2Cnq)(F + 2Eq^2) + BEq(B - 2Cnq)(F + 2Eq^2), \end{aligned}$$

and some easy computations show that the second line is equal to

$$EFq(B - 2Cnq)^2 + 2q(BEq + CFn)^2,$$

so that altogether we obtain

$$q(B - 2Cnq)^2\big(2B^2 + 2C^2 + AB + CD + EF\big) + 2q(BEq + CFn)^2 = 0.$$

Combining this with (4-9) we get

$$2\big((B - 2Cnq)^2(B^2 + C^2) + (BEq + CFn)^2\big) = (B - 2Cnq)^2(1 - n_1 - n_2).$$

But this is impossible, since $B - 2Cnq \neq 0$, $B \neq 0$ and $n_1, n_2 \geq 0$, giving us the desired contradiction. □

Acknowledgements

I would like to thank Marius Crainic for his help as supervisor of my master thesis, when I first worked on this topic. I would also like to thank David Martínez Torres for a fruitful discussion regarding his paper that this work is based on. Finally, I would like to thank my PhD advisor Rui Loja Fernandes for his feedback during the writing of this paper.

References

[Alcalde-Cuesta and Hector 1995] F. Alcalde-Cuesta and G. Hector, “Intégration symplectique des variétés de Poisson régulières”, *Israel J. Math.* **90**:1-3 (1995), 125–165. MR Zbl

[Barth et al. 1984] W. Barth, C. Peters, and A. Ven., “Compact complex surfaces”, Ergebnisse der Mathematik und ihrer Grenzgebiete **3**, Springer, 1984.

[Crainic and Fernandes 2004] M. Crainic and R. L. Fernandes, “Integrability of Poisson brackets”, *J. Differential Geom.* **66**:1 (2004), 71–137. MR

[Crainic and Fernandes 2005] M. Crainic and R. L. Fernandes, “Rigidity and flexibility in Poisson geometry”, pp. 53–68 in *Travaux mathématiques*, edited by C. Molitor-Braun et al., Sém. Math. Luxembourg **16**, Univ. Luxembourg, 2005. MR

[Crainic et al. 2019a] M. Crainic, R. L. Fernandes, and D. Martínez Torres, “Poisson manifolds of compact types (PMCT 1)”, *J. Reine Angew. Math.* **756** (2019), 101–149. MR

[Crainic et al. 2019b] M. Crainic, R. L. Fernandes, and D. Martínez Torres, *Regular Poisson manifolds of compact types* (*PMCT 2*), Astérisque **413**, 2019. MR

[Crainic et al. 2021] M. Crainic, R. L. Fernandes, and I. Mărcuţ, *Lectures on Poisson geometry*, Graduate Studies in Mathematics **217**, Amer. Math. Soc., Providence, RI, 2021. MR

[Duistermaat and Heckman 1982] J. J. Duistermaat and G. J. Heckman, “On the variation in the cohomology of the symplectic form of the reduced phase space”, *Invent. Math.* **69**:2 (1982), 259–268. MR Zbl

[Goldman 2022] W. M. Goldman, *Geometric structures on manifolds*, Graduate Studies in Mathematics **227**, Amer. Math. Soc., Providence, RI, 2022. MR Zbl

[Kotschick 2006] D. Kotschick, “Free circle actions with contractible orbits on symplectic manifolds”, *Math. Z.* **252**:1 (2006), 19–25. MR Zbl

[Martínez Torres 2014] D. Martínez Torres, “A Poisson manifold of strong compact type”, *Indag. Math.* (*N.S.*) **25**:5 (2014), 1154–1159. MR Zbl

[Mishachev 1996] K. N. Mishachev, “The classification of Lagrangian bundles over surfaces”, *Differential Geom. Appl.* **6**:4 (1996), 301–320. MR Zbl

[Mordell 1953] L. J. Mordell, “On the linear independence of algebraic numbers”, *Pacific J. Math.* **3** (1953), 625–630. MR Zbl

[Sepe 2010] D. Sepe, “Classification of Lagrangian fibrations over a Klein bottle”, *Geom. Dedicata* **149** (2010), 347–362. MR Zbl

[Sepe 2013] D. Sepe, “Universal Lagrangian bundles”, *Geom. Dedicata* **165** (2013), 53–93. MR Zbl

[Siu 1983] Y. T. Siu, “Every $K3$ surface is Kähler”, *Invent. Math.* **73**:1 (1983), 139–150. MR Zbl

Received June 8, 2023. Revised September 1, 2023.

LUKA ZWAAN
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN
URBANA, IL
UNITED STATES
lukaz2@illinois.edu

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 325 No. 2 August 2023