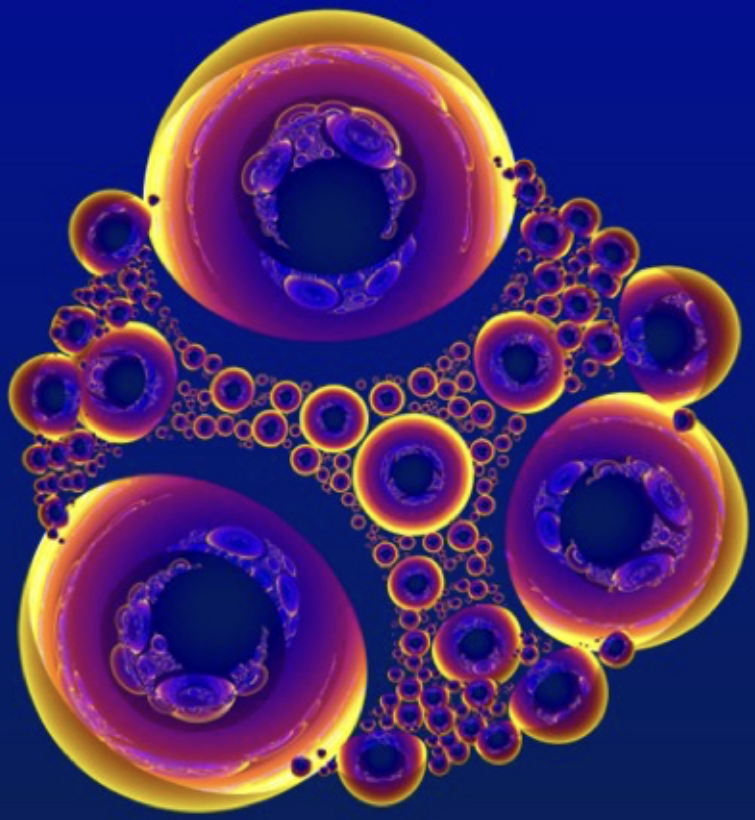
ARNOLD MATHEMATICAL JOURNAL



2026
Volume 12, Issue 1

Arnold Mathematical Journal

Journal of **Institute of Mathematical Sciences**, Stony Brook University, NY
Published by **Association for Mathematical Research**, Davis, CA; Jenkintown PA.

Contents

Mee Song Im, Mikhail Khovanov 1
Foam cobordism and the Sah-Arnoux-Fathi invariant

Alexander Blokh, Lex Oversteegen, Nikita Selinger, Vladlen Timorin, Sandeep Chowdary Vejandla 60
Symmetric cubic polynomials

Thomas Elgin, Nathan Reading, Salvatore Stella 111
Cluster scattering coefficients in rank 2

Allan Allemand, Alexey Kanel-Belov, Rodion Zaytsev 123
Monodromy Groups and the Insolvability of Transcendental Equations in Quadratures

Boris Khesin, Luke Volk 141
Morse-Bott Volume Forms

Arnold Mathematical Journal
Volume 12, Issue 1, 2026

Foam cobordism and the Sah-Arnoux-Fathi invariant

Mee Seong Im Mikhail Khovanov

Received 18 Oct 2024; Accepted 22 Jun 2025

Abstract: This is the first in a series of papers where scissor congruence and K-theoretical invariants are related to cobordism groups of foams in various dimensions. A model example is provided where the cobordism group of weighted 1-foams is identified, via the Sah–Arnoux–Fathi invariant, with the first homology of the group of interval exchange automorphisms and with the Zakharevich first K-group of the corresponding assembler. Several variations on this cobordism group are computed as well.

AMS Classification: 37E05, 37E99, 18M30, 19D99

Key words and phrases: Foams, foam cobordism, Sah-Arnoux-Fathi invariant, interval exchange transformations, train tracks

DOI: 10.56994/ARMJ.012.001.001

1 Introduction

In link homology by a *foam*, one usually means a 2-dimensional finite combinatorial CW-complex F, often embedded in $\mathbb{R}^3$, where each point has one of the three types of neighborhoods shown in Figure 1 below. Foams are used in algebraically-defined link homology to build state spaces of planar graphs, which are then combined into complexes that define the homology of a link [16, 20, 30, 28, 17]. Foams also appear in Kronheimer–Mrowka instanton Floer homology for 3-orbifolds [18].

Locally, the foam structure is that of a two-dimensional spine of a 3-manifold. Often, foams come with extra decorations, such as orientations, weights and other labels on facets.

In this paper, a *closed* 2*-foam* means a foam as above, with additional decorations specified. More generally, one can define a 2*-foam with boundary*, the boundary being a 1-foam. A 1*-foam* is a finite graph, possibly with loops and circle edges without vertices, and additional decorations. Splitting the boundary of a 2-foam F into two disjoint sets of components, $\partial F = (-U_0) \sqcup U_1$, allows one to view F as a cobordism between 1-foams U_0 and U_1. Decorations of U_0, U_1 are induced from those of F.

This paper is the first in a series of papers which aim to use foams, in all dimensions n and with additional decorations, to understand K-theoretical structures. One expects that n-dimensional foams decorated by objects and morphisms of an exact category $\mathcal{C}$, modulo concordances which are $\mathcal{C}$-decorated $(n+1)$-dimensional foams, carry information about the n-th K-theory group $K_n(\mathcal{C})$ of $\mathcal{C}$. Facets, respectively, seams of a foam are decorated by flat connections with objects of $\mathcal{C}$, respectively short exact sequences of $\mathcal{C}$, as fibers of these flat bundles. This relation between decorated foams and algebraic K-theory is started to be studied in [9].

The present paper works out a straightforward example of this correspondence, where the abelian group of suitably decorated one-dimensional foams modulo 2-dimensional cobordisms is identified with the group $\mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R}$, which is the first homology of the group of

interval exchange transformations [36]. The related invariant of interval exchange transformations mapping a group element to its image in the first homology is known as the Sah–Arnoux–Fathi invariant, or SAF invariant, for short [34, 8]. I. Zakharevich interpreted the SAF invariant map via the K_1 group of a suitable *assembler* category [37, 38], and that category plays the role of the exact category $\mathcal{C}$ above. In the present paper, we relate these structures to two-dimensional cobordisms between decorated one-dimensional foams.

In Section 2, we work out this new interpretation of the SAF invariant, as classifying elements of the cobordism group of *weighted* oriented 1-foams. In this construction, edges of an oriented 1-foam are decorated by positive real numbers a, with compatibility relations on these numbers at the vertices. The cobordism group of such foams is identified with the abelianization of the group of interval exchange transformations (IETs) in Theorem 2.6. The isomorphism uses the Sah–Arnoux–Fathi invariant of IETs, extended to arbitrary weighted oriented 1-foams.

Section 3 considers the cobordism group of planar unoriented weighted 1-foams and identifies it with the abelian group generated by brackets $[a, b]$ with $a, b \in \mathbb{R}_{>0}$ modulo the antisymmetry and the 2-cocycle relations (14)-(16). It also looks at a variation on weighted embedded foams, where each facet may carry either a positive or a negative weight. Several other variations on the group of foam cobordisms are studied in Section 4. Constructions of Sections 2 and 3 can perhaps be viewed as first steps exploring the relation between foam cobordisms and dynamical systems.

2 Foams and interval exchange transformations

In this section we interpret the Sah–Arnoux–Fathi invariant of interval exchange transformations [34, 36, 8] via cobordism classes of oriented 1-foams with facets decorated by positive real numbers (called *weighted* or $\mathbb{R}_{>0}$-*decorated* 1-foams).

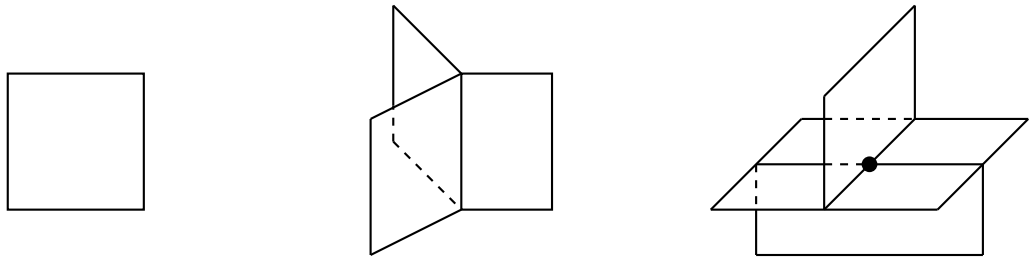


Figure 1: Three types of points on a 2-foam. Left to right: a regular point, seam points on a seam interval, a vertex.

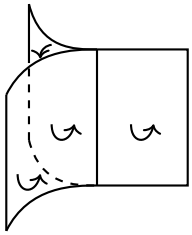


Figure 2: A choice of facet orientations and an order of thin facets near a seam of an oriented 2-foam. Facet orientations are indicated by the three "cap" semicircular arrows. The two thin facets are shown as tangent to each other along the seam, which is a convenient convention for tracking thin facets.

2.1 Oriented 1-foams and 2-foams and cobordisms between 1-foams

In this paper, a *closed 2-foam* denotes a finite combinatorial CW-complex F, where each point is one of the three types and has a neighborhood as depicted in Figure 1; these points are called *regular points*, *seam points* and *vertices* of the 2-foam, respectively. The union of seams and vertices of F is a four-valent graph $s(F)$, possibly with loops and verticeless circles. The connected components of $F \setminus s(F)$ are called the *facets* of F, and the set $s(F)$ is called the *singular points* of F.

A closed 2-foam is *oriented* if

- Each facet is oriented so that along its seams and near its vertices, the orientations match as shown in Figure 2 (for seams) and Figure 3 on the right (for vertices).

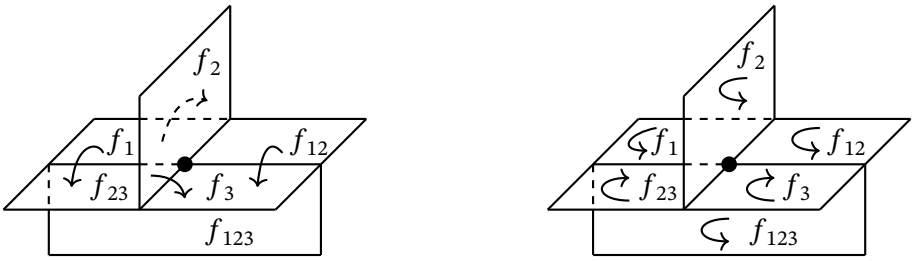


Figure 3: Left: ordering of seams near a vertex, with facets labelled f_1, f_2, f_3, f_{12}, f_{23}, f_{123}. The orderings are from smaller to larger indices: (f_1, f_2), (f_2, f_3), (f_1, f_{23}), (f_{12}, f_3). Right: one out of two possible facet orientations near a vertex is shown. Orientations and orderings must be compatible along each seam, as explained earlier and in Figure 2.

Along each seam, two of the facets are designated as *thin* and the remaining one as *thick*. The orientation of each thin facet matches (flows into) the orientation of the thick facet. Furthermore, the orientations of the two thin facets along a seam are opposite. A facet which is thin at one of its seams may be thick at another seam; this is a crucial difference from the foams mentioned at the beginning of Section 1.

- An order of two thin facets along each seam is fixed (shown by the small curly arrow from one thin facet to the other in Figure 2).

- At each vertex of the foam, the decorations (orientations and orders) of the six adjoint facets along the four seams match as follows (and shown in Figure 3 on the right). The six facets are labeled $f_1, f_2, f_3, f_{12}, f_{13}, f_{123}$. Among the triples of facets $(f_1, f_2, f_{12}), (f_2, f_3, f_{23}), (f_{12}, f_3, f_{123}), (f_1, f_{23}, f_{123})$, one triple for each seam, the first two facets are thin and the last one is thick. The facets are oriented either as shown in Figure 3 on the right or with all orientations opposite (which follows from the orientation requirements along the seams). The orders of the facets along the seams are as shown in Figure 3 on the left, in the direction of increasing indices, or the opposite (decreasing indices).

Figure 4 shows a set of three “parallel cross-sections” of a foam near a vertex, with

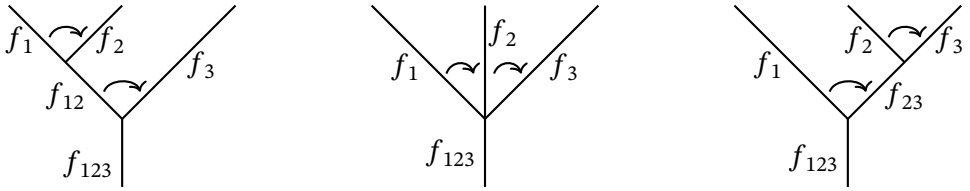


Figure 4: Three parallel cross-sections near a vertex of a 2-foam, with the middle cross-section going through the vertex. Small arrows show the order of facets along the seams.

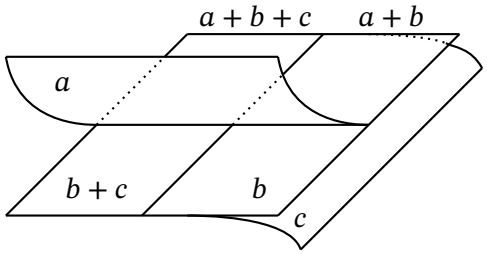


Figure 5: A vertex of a (weighted) 2-foam is analogous to that of a branched surface, c.f. [24, Figure 1.1].

one of the cross-sections going through the vertex. Figure 5 depicts a neighborhood of a vertex taking "tangencies" of the thin facets along the four seams near the vertex into account, analogous to that of a vertex in a branched surface [24, Figure 1.1], see also [27]. (For now, ignore the weights of the facets in Figure 5.)

The 1-foams can be thought of as generic cross-sections of 2-foams. A 1-foam is a finite oriented trivalent graph. At each vertex, there are two *in* edges and one *out* edge or vice versa. We call these *merge* and *split* vertices, correspondingly. They are shown in Figure 6 on the left (ignoring the weights $a, b, a+b$ in that figure). An oriented circle with no vertices on it is allowed as a component of a 1-foam. Loops are *a priori* allowed, although we will not encounter them in the present section due to working with oriented foams (they do appear in Section 3, upon consideration of unoriented foams).

It is convenient to visualize the thin edges at a merge vertex as sharing a tangent line at a vertex and think of a neighborhood of a merge vertex as a generic cross-section

across the seam of the Figure 2 foam. Likewise, a neighborhood of a split vertex of a 1-foam can be visualized as a horizontal cross-section of the rightmost foam in Figure 6. Similar conventions are used in [29, 27].

We define an oriented 2-foam F with boundary as a cobordism between oriented 1-foams U_0, U_1, where we read the morphism from bottom to top. The boundary of F is split into two disjoint 1-foams,

$$\partial F \cong (-\partial_0 F) \sqcup \partial_1 F \cong U_1 \sqcup (-U_0).$$

Away from the boundary F has a local structure that of an oriented 2-foam and collar neighborhoods near U_i, $i = 0, 1$, where it is homeomorphic to the product $U_i \times [0, \epsilon)$, $\epsilon > 0$. The orientations of the facets of F and local orders of the thin facets along the seams of F restrict to orientations of edges of its boundary 1-foams and local orders of the thin edges at the vertices of the boundary 1-foams using the standard convention for an induced orientation of the boundary of a manifold.

For completeness, we mention that an *oriented* 0*-foam* is a finite collection of points with orientations (signs + and −). It is clear how to define oriented 1-foams with boundary.

2.2 Weighted or $\mathbb{R}_{>0}$-decorated foams

Consider oriented 1-foams and 2-foams with edges (for 1-foams) and facets (for 2-foams) decorated by real numbers a for various $a > 0$ and refer to a as the *thickness*, *width*, or *label* of the facet. At a vertex of a 1-foam and a seam of a 2-foam, widths must add as shown in Figure 6. Informally, one can "thicken" the foams and think of intervals $[0, a)$ and $[0, b)$ merging into the interval $[0, a + b) = [0, a) \sqcup [a, a + b)$ at a vertex of a 1-foam and a seam of a 2-foam. This thickening is independent of a facet being thin or thick. The order of thin edges near a vertex (for 1-foams) and order of thin 2-facets near a seam (for 2-foams) matches the order of the intervals in the merge, see Figure 6.

At a vertex of a decorated 2-foam, three thin facets of thickness a_1, a_2, a_3 merge into facets of thickness $a_1 + a_2$ and $a_2 + a_3$, which then merge with the remaining thin facet

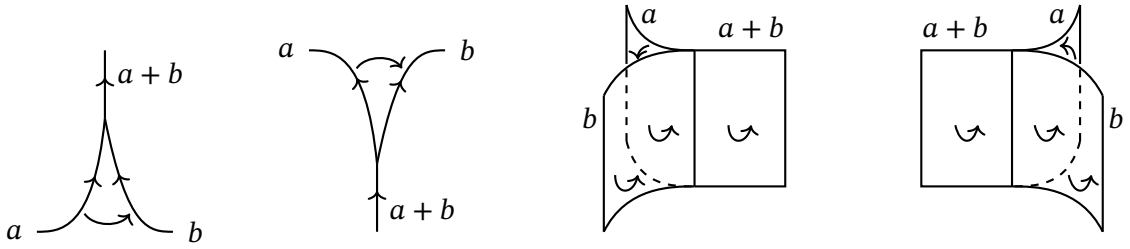


Figure 6: Left: Neighborhoods of a merge and split vertices of a weighted 1-foam, respectively. Right: neighborhoods of a point near a seam of a weighted 2-foam.

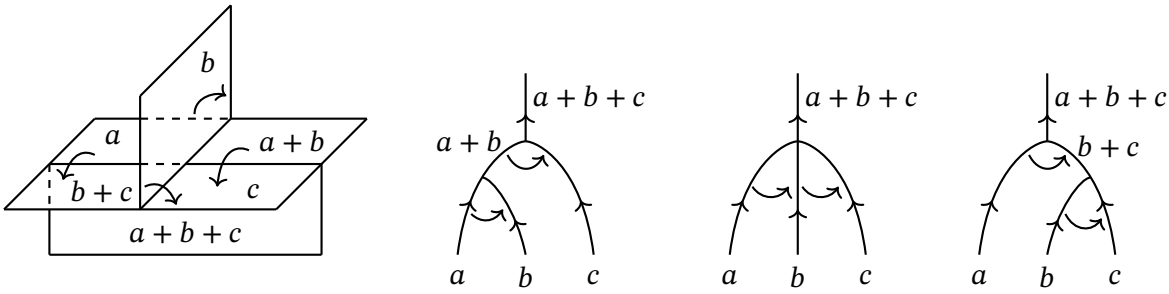


Figure 7: Left: labels near a vertex. Small arrows indicate one of the two possible orders of thin edges at each of the four seams near a vertex (orientations of facets are not shown). Right: three parellel cross-sections of this foam, including one which contains the vertex.

into the facet of thickness $a_1 + a_2 + a_3$, see Figure 7, which also shows three parallel cross-sections of this foam.

Remark 2.1. If desired, one may allow lines and facets to carry the empty interval $[0, 0)$, but this does not seem essential. Such lines and facets can then be deleted from a foam. Namely, remove all 0-weight facets. If a seam had thin facets of weights 0 and $a > 0$ along it, the seam can be hidden and thin and thick facets of thickness a along it merged into a single facet. A seam with thin facets of weights $(0, 0)$ is deleted. This operation is suitably extended to vertices with thin facets of thickness (a, b, c) at it, see Figure 7, depending on which of these three numbers are 0.

We call such foams *weighted foams* or $\mathbb{R}_{>0}$*-decorated foams* or IET*-foams* (see Sec-

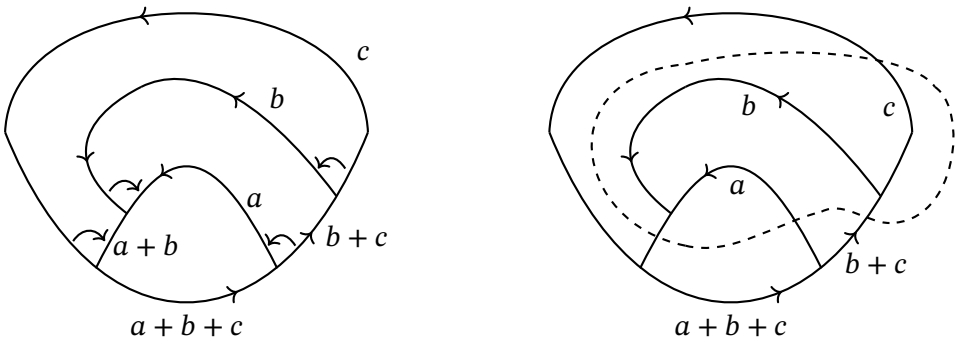


Figure 8: Left: the link of a vertex of a foam, with one possible choice of compatible orientations of facets inducing orientations of edges of the link. Likewise, one out of two possible compatible orders of thin facets at seams is shown. Right: the dashed line cuts the diagram into two pieces that appear as the two boundaries of a vertex cobordism in Figure 16, second row on the right (one of those two diagrams requires orientation reversal vs the diagram above, due to splitting the boundary into top and bottom components).

tion 2.3). The definition is straightforward to extend to all dimensions. A weighted 0-foam is a finite set of points with signs $\{+,-\}$ and weights $a>0$.

Figure 8 shows the link of a vertex of a weighted 2-foam. Weighted 2-foams are analogous to measured branched surfaces and measured laminations [23, 24], but without an embedding into a 3-manifold.

For $n=0,1$ denote by $\mathsf{Cob}^n_{>0}$ the cobordism group of weighted oriented n-foams. An n-foam U defines the trivial element $[U]=0\in\mathsf{Cob}^n_{>0}$ if and only if it bounds a weighted oriented $(n+1)$-foam.

Proposition 2.2. The cobordism group of weighted oriented 0-foams is isomorphic to $\mathbb{R}$:

$$\mathsf{Cob}^0_{>0}\ \cong\mathbb{R}. \tag{1}$$

Proof. A weighted oriented 0-foam is given by a finite collection of points decorated by signs and real weights $a>0$. Merge all $+$-decorated points into one point (adding the

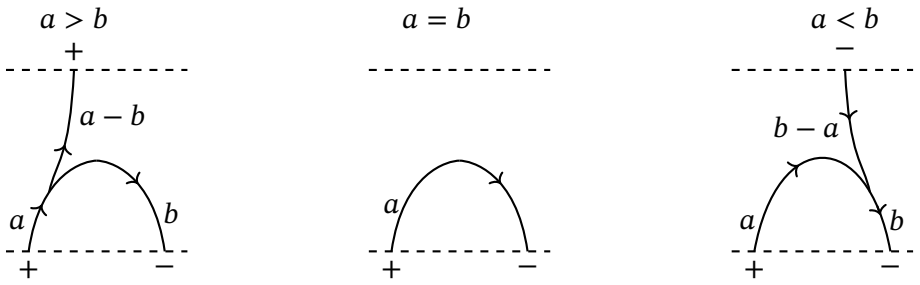


Figure 9: Merging points $(+,a)$ and $(-,b)$ via a cobordism, where $a,b>0$.

weights) and all $-$-decorated points into a point (adding the weights). The result is at most two points $(+,a),(-,b)$, which are cobordant to $(+,a-b)$ if $a>b$, $(-,b-a)$ if $a<b$, and to the empty 0-foam if $a=b$, see Figure 9.

Under the isomorphism in (1) point $(+,a)$, $a\in\mathbb{R}_{>0}$ is sent to $a\in\mathbb{R}$, point $(-,a)$ is sent to $-a\in\mathbb{R}$, and the disjoint union of signed decorated points is converted to the sum of corresponding numbers. □

2.3 Interval exchange transformations and $\mathbb{R}_{>0}$-decorated 1-foams

Pick $r\geq 1$, a decomposition $1=\sum_{i=1}^{r}\lambda_i$, $0<\lambda_i<1, \lambda_i\in\mathbb{R}$ and a permutation $\sigma\in S_r$. Interval exchange transformation $T_{\lambda,\sigma}: [0,1)\longrightarrow[0,1)$ is a bijection of a semiclosed interval to itself given by writing it as the disjoint union of r intervals

$$[0,1)=[0,\lambda_1)\sqcup[\lambda_1,\lambda_1+\lambda_2)\sqcup\ldots\sqcup[1-\lambda_r,1)$$

and permuting the order of intervals according to σ, making the i-th interval $\sigma(i)$-th in the order.

The Sah–Arnoux–Fathi invariant of $T_{\lambda,\sigma}$ is an element of $\mathbb{R}\otimes_{\mathbb{Q}}\mathbb{R}$ given by

$$\mathsf{SAF}(T_{\lambda,\sigma}) := \sum_{i=1}^{r}\lambda_i\otimes t_i = \sum_i\left(\sum_{j:\sigma(j)<\sigma(i)}\lambda_i\otimes\lambda_j-\sum_{j<i}\lambda_i\otimes\lambda_j\right), \tag{2}$$

where $t_i=\sum_{j:\sigma(j)<\sigma(i)}\lambda_j-\sum_{j<i}\lambda_j\in\mathbb{R}$ is the displacement of the i-th interval by σ.

One can write $\mathsf{SAF}(T_{\lambda,\sigma})$ as a linear combination of elements $\lambda_i \otimes \lambda_j - \lambda_j \otimes \lambda_i$, $i, j \leq r$ and view it as an element of $\mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R} = \Lambda^2_{\mathbb{Q}}(\mathbb{R})$, the quotient of $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}$ by the abelian subgroup spanned by $\lambda \otimes \lambda, \lambda \in \mathbb{R}$. Note that we have the decomposition $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R} \cong \Lambda^2_{\mathbb{Q}}(\mathbb{R}) \oplus S^2_{\mathbb{Q}}(\mathbb{R})$, the sum of exterior and symmetric squares, and one is taking the projection onto the first summand. The invariant can also be written as follows:

$$\mathsf{SAF}(T_{\lambda,\sigma}) \;=\; 2 \sum_{i<j\,:\,\sigma(j)<\sigma(i)} \lambda_i \wedge \lambda_j, \tag{3}$$

where $a \wedge b$ denotes the image of $a \otimes b$ under the quotient map $q : \mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R} \longrightarrow \Lambda^2_{\mathbb{Q}}\mathbb{R}$, since $q(a \otimes b - b \otimes a) = 2a \wedge b$.

The SAF invariant of $T_{\lambda,\sigma}$ can be written as the integral

$$\mathsf{SAF}(T_{\lambda,\sigma}) = \int_{[0,1)} 1 \otimes (T_{\lambda,\sigma}(x) - x)dx,$$

see [36, page 2]. This invariant vanishes precisely on the commutator subgroup.

Let $\mathsf{Aut}_{\mathsf{IET}}$ be the group of Interval Exchange Transformations of $[0,1)$, that is, the group of bijections $T_{\lambda,\sigma}$ as above, with the group operation given by the composition of maps. There is a short exact sequence of groups

$$1 \longrightarrow [\mathsf{Aut}_{\mathsf{IET}}, \mathsf{Aut}_{\mathsf{IET}}] \longrightarrow \mathsf{Aut}_{\mathsf{IET}} \xrightarrow{\mathsf{SAF}} \Lambda^2_{\mathbb{Q}}\mathbb{R} \longrightarrow 1. \tag{4}$$

Remark 2.3. I. Zakharevich [38] interpreted the Sah–Arnoux–Fathi invariant as describing K_1 of an appropriate assembler category. Combining this result with constructions of the present paper yields an example of the relation between K_1 group of an appropriate category and the group of 1-foam cobordisms, in a rather special case. In a forthcoming paper, we will discuss the relation between the K_1 group and the cobordism group of decorated 1-foams in greater generality.

Remark 2.4. To each interval exchange transformation $T_{\lambda,\sigma}$ as earlier, we assign a weighted 1-foam with boundary $F_{\lambda,\sigma}$ and a closed weighted 1-foam $\widehat{F}_{\lambda,\sigma}$, as shown in Figure 10. Start with a line of thickness 1 and split it into lines of thickness $\lambda_1, \ldots, \lambda_r$ from left to right. Then permute the points at the top end of the split via the permutation

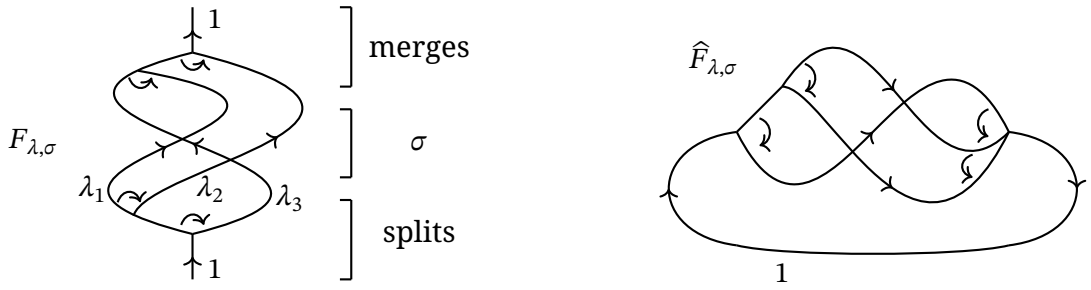


Figure 10: A foam with boundary $F_{\lambda,\sigma}$ associated to IET $T_{\lambda,\sigma}$ and its closure $\widehat{F}_{\lambda,\sigma}$. Intersections are virtual due to having to depict the foam via a projection to the plane.

$\sigma \in S_r$. After that, merge the resulting points into an interval of width 1, and close up top and bottom endpoints, both of thickness 1, into a closed diagram. Denote by $F_{\lambda,\sigma}$ the resulting weighted oriented 1-foam with boundary and by $\widehat{F}_{\lambda,\sigma}$ its closure. Intersections in Figure 10 are virtual, that is, due to having to depict the foam via a projection to the plane.

Notice that, in the cobordism group $\mathrm{Cob}^1_{>0}$, a 1-foam $\widehat{F}_{\lambda,\sigma}$ does not depend on the sequence in which the interval 1 is split into $\lambda_1, \dots, \lambda_r$ as long as in the split, $\lambda_1, \dots, \lambda_r$ go from left to right. For instance, for $r = 3$, the two sequences of splits $1 \to (\lambda_1, \lambda_2 + \lambda_3) \to (\lambda_1, \lambda_2, \lambda_3)$ and $1 \to (\lambda_1 + \lambda_2, \lambda_3) \to (\lambda_1, \lambda_2, \lambda_3)$ give rise to cobordant foams. Likewise, the sequence of merging the intervals back is irrelevant, as long as the order from left to right is $\lambda_{\sigma^{-1}(1)}, \dots, \lambda_{\sigma^{-1}(r)}$. The two 1-foams that differ in that way are then cobordant via a composition of 2-foams that create the vertices, see Figure 7. Likewise, a foam $F_{\lambda,\sigma}$, in the cobordism set of 1-foams with a fixed boundary, does not depend on the order of merges and splits.

Composition of two IETs $T_{\lambda,\sigma}$ and $T_{\lambda',\sigma'}$ is an IET $T_{\lambda'',\sigma''} = T_{\lambda',\sigma'} \circ T_{\lambda,\sigma}$ for suitable (λ'', σ'').

Proposition 2.5. The foams $\widehat{F}_{\lambda'',\sigma''}$ and $\widehat{F}_{\lambda,\sigma} \sqcup \widehat{F}_{\lambda',\sigma'}$ are cobordant. Assigning a 1-foam $\widehat{F}_{\lambda,\sigma}$

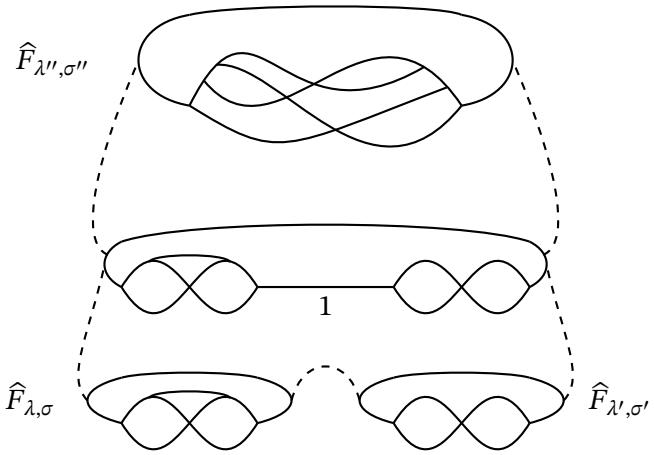


Figure 11: A schematic depiction of the cobordism from $\widehat{F}_{\lambda,\sigma} \sqcup \widehat{F}_{\lambda',\sigma'}$ to $\widehat{F}_{\lambda'',\sigma''}$ in the proof of Proposition 2.5.

to an IET $T_{\lambda,\sigma}$ extends to a homomorphism of groups

$$\phi' \ : \ \mathsf{Aut}_{\mathsf{IET}} \longrightarrow \mathsf{Cob}^1_{>0}. \tag{5}$$

Proof. The cobordism is given by first merging $\widehat{F}_{\lambda,\sigma} \sqcup \widehat{F}_{\lambda',\sigma'}$ into a connected foam, as schematically shown in the bottom half of Figure 11. The interval of thickness 1, labelled in that Figure, is then repeatedly split, by repeatedly applying elementary cobordisms shown in the second row on the right in Figure 16. These elementary cobordisms convert the 1-foam into $\widehat{F}_{\lambda'',\sigma''}$ via a foam concordance between braid-like foams, as schematically depicted in the top half of Figure 11. The rules for computing the composition $T_{\lambda',\sigma'} \circ T_{\lambda,\sigma}$ are easy to translate to a particular sequence of elementary braid-like cobordisms between these two 1-foams. □

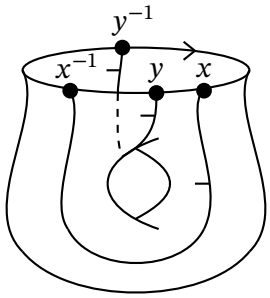


Figure 12: A commutator of IETs x, y is null-cobordant.

2.4 Cobordism group of weighted oriented 1-foams

Since the cobordism group is abelian, homomorphism ϕ' factors modulo the commutator of the automorphism group, giving a homomorphism

$$\phi \ : \ \mathrm{H}_1(\mathsf{Aut}_{\mathsf{IET}}, \mathbb{Z}) \longrightarrow \mathsf{Cob}^1_{>0}. \tag{6}$$

Figure 12 shows a cobordism from a commutator of two elements to the identity (or to the empty 1-foam).

Theorem 2.6. The homomorphism ϕ in (6) is an isomorphism of abelian groups, giving isomorphisms

$$\mathsf{Cob}^1_{>0} \ \cong \ \mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R} \ \cong \ \mathrm{H}_1(\mathsf{Aut}_{\mathsf{IET}}, \mathbb{Z}) \ \cong \ K_1(\mathcal{C}_Z). \tag{7}$$

The second isomorphism is the SAF invariant, and an isomorphism $\mathrm{H}_1(\mathsf{Aut}_{\mathsf{IET}}, \mathbb{Z}) \cong K_1(\mathcal{C}_Z)$ is constructed in [38]. The category $\mathcal{C}_Z$ is the Zacharevich assembler category [38] for the IETs, also see Remark 2.7 below.

Proof. We establish an isomorphism

$$\psi \ : \ \mathsf{Cob}^1_{>0} \xrightarrow{\cong} \mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R} \tag{8}$$

which is compatible with homomorphism ϕ and makes the following diagram commute

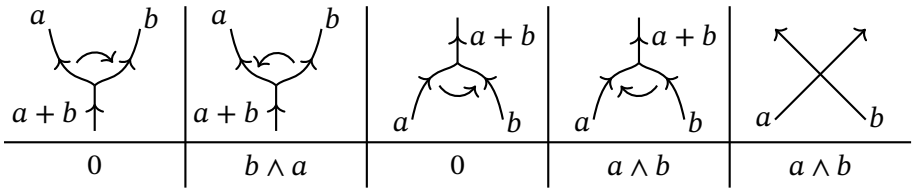
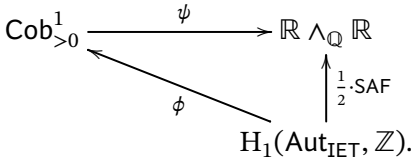


Figure 13: Contributions of splits, merges and intersections to the invariant ν.



Consider a weighted oriented 1-foam U and project it generically to a plane to a diagram D.

The projection has two types of merge points and two types of split points, depending on whether the order of thin edges at a point is clockwise or counterclockwise, see Table 13.

To diagram D, assign an element $\nu(D) \in \mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R}$ as a sum over local contributions:

- A split vertex with a clockwise thin edge order and a merge vertex with a counterclockwise thin edge order contribute 0. See the first split and the first merge (from left to right) in the table in Figure 13.
- For the other orientations, the contributions are shown in the table in Figure 13.
- A crossing of two intervals of lengths a and b contributes $a \wedge b$, with the orientations of the intervals determining the order of a, b in the product, see the table in Figure 13.

Some examples are shown in Figure 14. Note that ν is additive under the disjoint union of diagrams.

We claim that $\nu(D)$ depends only on the 1-foam U, that is, different plane projections result in the same invariant $\nu(D)$. Two such projections differ by local moves in Figure 15

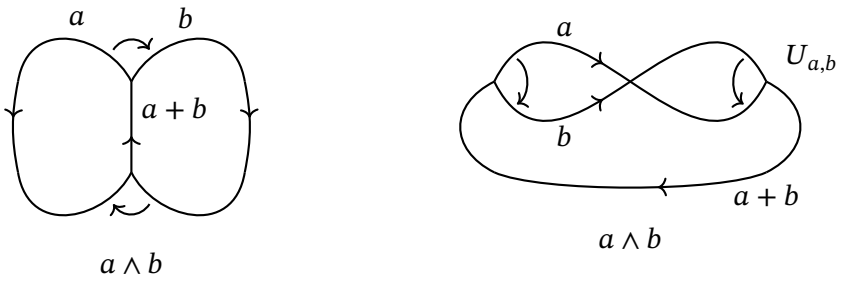


Figure 14: An invariant ν for each of these two foams is $a \wedge b$. For the foam on the left, the merge contributes $a \wedge b$ while the split contributes 0. For the foam on the right, denoted $U_{a,b}$, the intersection contributes $a \wedge b$, while both the merge and the split contribute 0. (These two foams are homeomorphic, through a homeomorphism that preserves all decorations.)

and versions of these moves given by reversing orientation of one or more of the components or reversing the order of thin edges at a vertex. This follows from the list of Reidemeister moves for embedded rigid graphs in Kauffman [14], although our case is easier, since the graphs are projected onto the plane rather than first embedded in $\mathbb{R}^3$ and then projected onto $\mathbb{R}^2$.

It is straightforward to check the invariance of ν under all variations of moves in Figure 15. For example, independent of orientations of a and b lines, invariance of ν under move 1 in Figure 15 is the relation $a \wedge b + b \wedge a = 0$. For move 2 it is $a \wedge a = 0$. Move 4 and its version for the opposite orientation determine two entries of the Figure 13 table given the other three (they determine, for instance, entries 1 and 3 of row 2 given values at entries 2, 4, 5). Move 5 corresponds to the bilinearity property of the tensor product.

Suppose that weighted 1-foams U_0, U_1 are cobordant. A cobordism between them can be realized as a finite sequence of elementary cobordisms shown in Figure 16. We can put a foam in a generic position and then take a sequence of cross-sections with minimal changes of topology between consecutive cross-sections. These changes include

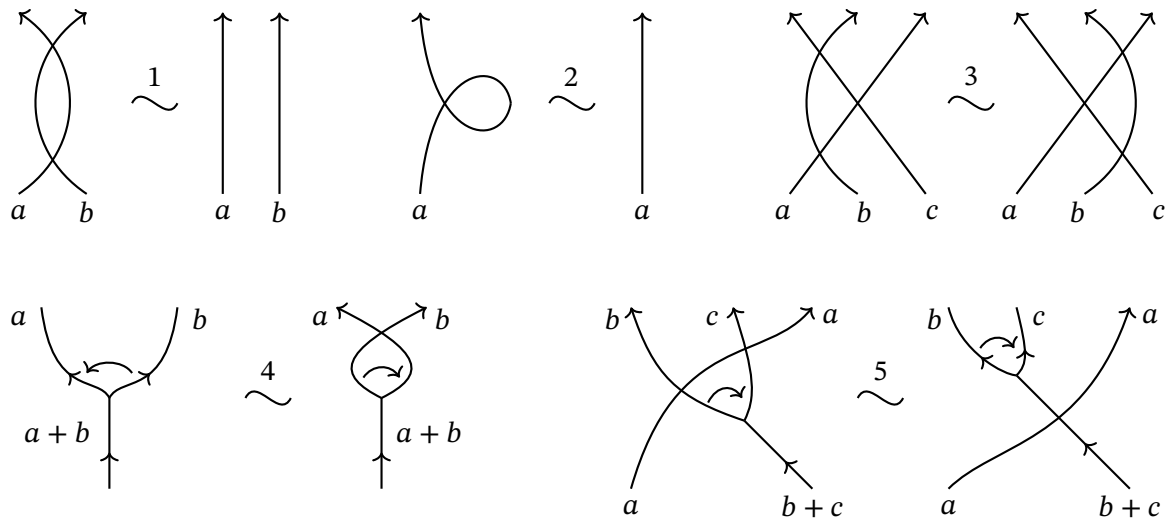


Figure 15: Diagram moves 1-5 that do not change the underlying foam.

(a) going through a vertex, which correspond to the changes in row 2 of Figure 16, (b) going through a local maximum or minimum of a seam, as shown in row 1 of Figure 16, (c) Morse critical point of a facet, as shown in row 3 on the right and the bottom row.

What remains are isotopies. The cobordism on the left of row 3 of the figure has no topology change between its top and bottom boundaries, which are different projections onto the plane of the same decorated 1-foam. This cobordism is included for completeness, and it corresponds to move 4 in Figure 15. Remaining isotopies correspond to the other moves in Figure 15.

For each elementary cobordism, one can pick diagrams for the two 1-foams at its boundaries so that they differ as shown in Figure 16. A direct computation implies that, in each case, the two diagrams have the same invariant ν.

Consequently, the homomorphism ψ in (8) is well-defined. It is clearly surjective, since each generator $a \wedge b$ is the image of some foam. Define a homomorphism

$$\psi_2 \ : \ \mathbb{R} \otimes_{\mathbb{Z}} \mathbb{R} \xrightarrow{\cong} \mathsf{Cob}^1_{>0} \tag{9}$$

by taking $a \otimes b$ to the foam in Figure 14 on the right.

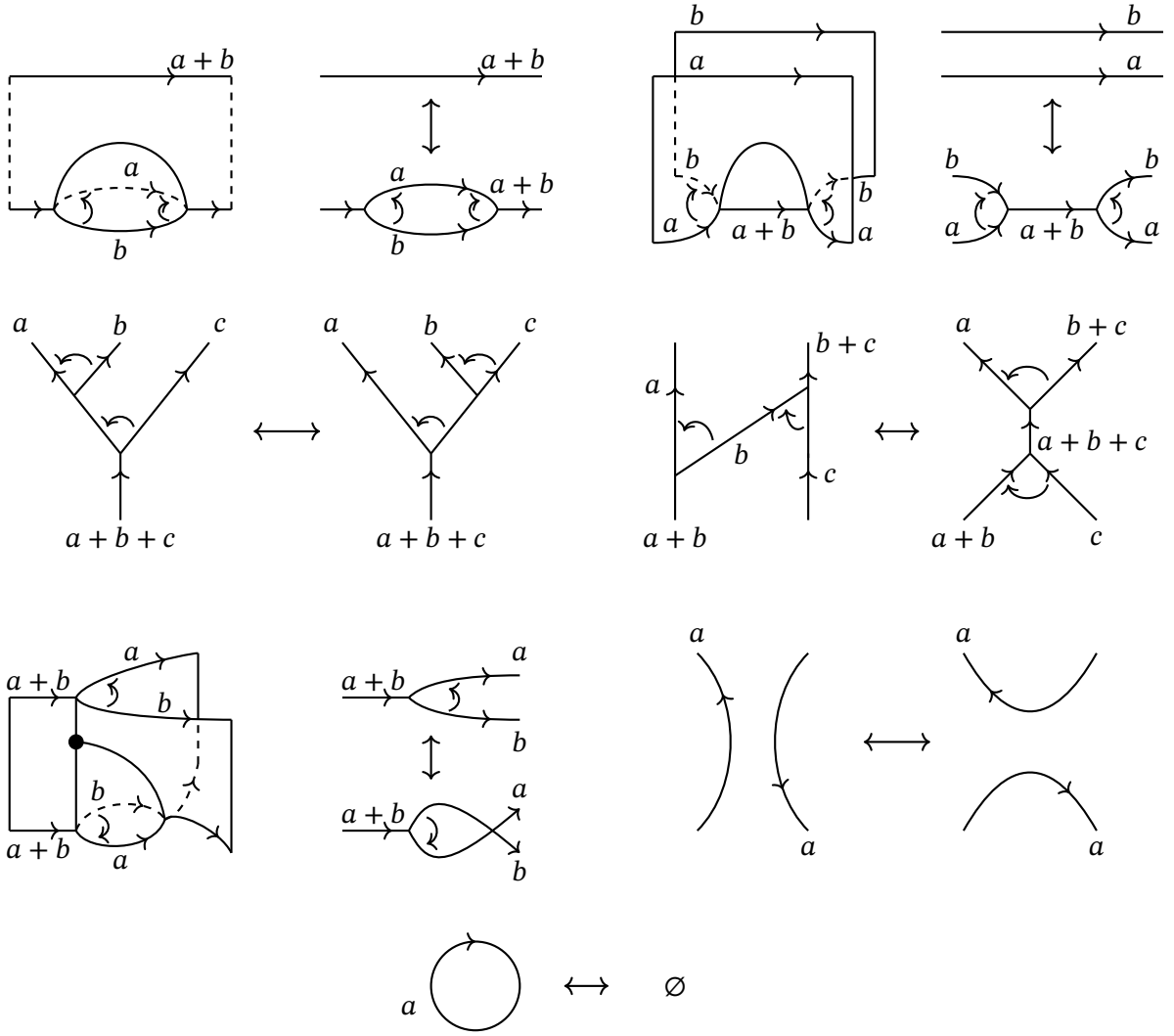


Figure 16: Top left (two pictures): a *singular cap* cobordism. Its reflection in a horizontal plane gives a *singular cup* cobordism. Top right (two pictures): a *singular saddle* cobordism. Second row: the standard cobordisms between these pairs of 1-foams are given by 2-foams with a single vertex. These transformations can also be obtained by splitting the link of a vertex into two halves, see Figure 8. Third row left (two pictures): flipping thin edges at a vertex in a diagram of a 1-foam. Third row right (two pictures): a saddle cobordism relates these two 1-foams. Bottom row: cup and cap cobordisms allow a circle to vanish or appear.

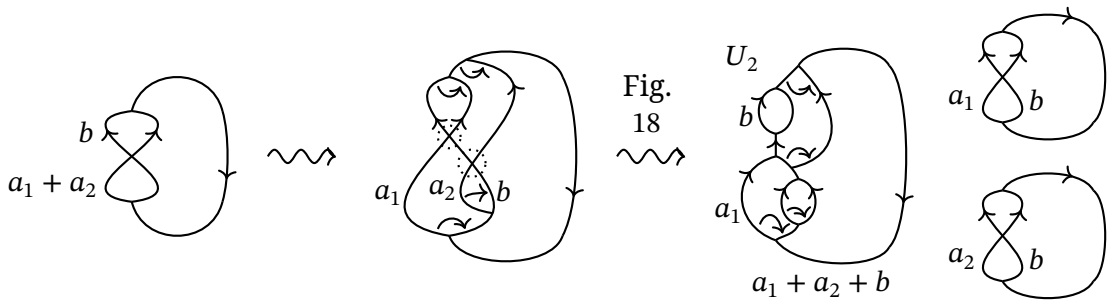


Figure 17: The cobordism for the first arrow splits the line $a_1 + a_2$ into two lines a_1 and a_2. In the second cobordism, two intersection points (with their neighborhoods indicated by dotted circles) are converted into foams $U_{a_1,b}$ and $U_{a_2,b}$, with the remaining braid-like foam U_2 cobordant to a circle of weight $a_1 + a_2 + b$; see Figure 18.

To show that ψ_2 is well-defined, we need to check the relations

$$\psi_2((a_1 + a_2) \otimes b) = \psi_2(a_1 \otimes b) + \psi_2(a_2 \otimes b), \tag{10}$$

$$\psi_2(a \otimes (b_1 + b_2)) = \psi_2(a \otimes b_1) + \psi_2(a \otimes b_2), \tag{11}$$

which also imply $n\,\psi_2(a \otimes b) = \psi_2(na \otimes b) = \psi_2(a \otimes nb)$, and, since $\mathbb{Q}$ is a divisible group, imply $\psi_2(\frac{a}{n} \otimes b) = \psi_2(a \otimes \frac{b}{n})$.

Note a related observation, that the natural quotient map $\mathbb{R} \otimes_{\mathbb{Z}} \mathbb{R} \longrightarrow \mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}$ is an isomorphism, so that (9) holds with $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}$ on the LHS as well.

The 1-foam $U_{a_1+a_2,b}$ associated to $(a_1 + a_2) \otimes b$ is shown in Figure 17 on the left. It is cobordant to the foam U with two crossings shown in the middle of the same figure. A crossing can be split off from any foam, as shown in Figure 18.

Splitting off both crossings from U results in the foam U_1 shown on the right of Figure 17. Foam U_1 is the union of $U_{a_1,b}, U_{a_2,b}$ and a braid-like foam U_2 with no crossings and compatible thin edge orientations at vertices. Foam U_2 is cobordant to the circle of weight $a_1 + a_2 + b$ and, then, to the empty foam. Hence, foams $U_{a_1+a_2,b}$ and $U_{a_1,b} \sqcup U_{a_2,b}$ are cobordant and the relation (10) holds. The relation (11) follows in the same way.

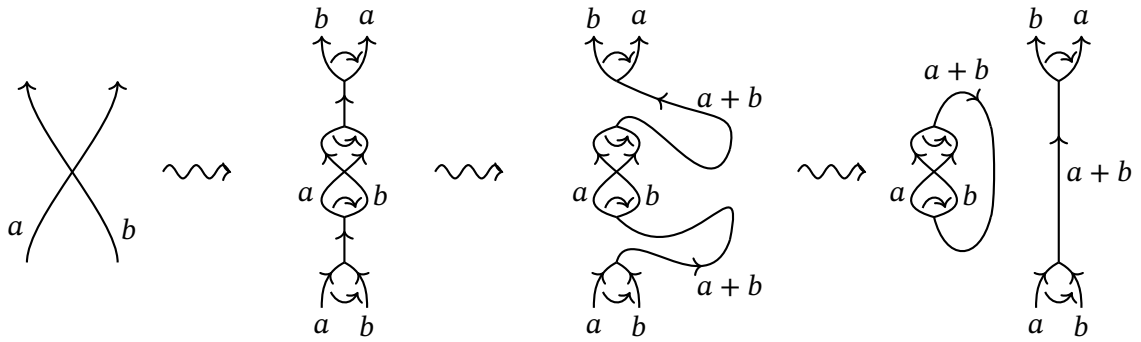


Figure 18: Converting a foam with a crossing to the union of a foam with one less crossing and foam $U_{a,b}$ shown in Figure 14 on the right.

To check that ψ_2 factors through a homomorphism

$$\psi_1 \ : \ \mathbb{R} \wedge_{\mathbb{Z}} \mathbb{R} \xrightarrow{\cong} \mathsf{Cob}^1_{>0}, \tag{12}$$

we observe that

$$\psi_2(a \otimes b + b \otimes a) = 0$$

since the disjoint union $U_{a,b} \sqcup U_{b,a}$ of 1-foams associated to $a \otimes b$ and $b \otimes a$ is null-cobordant.

To see that ψ_1 is surjective, pick a 1-foam U. This foam can be represented as the closure of a braid-like 1-foam B. Choose a diagram D of B where all splits and merges have local ν-invariant 0, see Table 13, with the closure $\widehat{D}$ describing the foam U.

All crossings can be removed from D via cobordisms shown in Figure 18. There, as a first step, parallel lines of thickness a and b above and below the crossing are merged to create two intervals, each of thickness $a + b$. They are then brought near each other and merged via a saddle point cobordism. This results in a disconnected 1-foam which is the disjoint union of foam $U_{a,b}$ and a foam with one fewer crossing versus the original.

The 1-foam $\widehat{D}$ is cobordant to the union $\widehat{D_1} \sqcup D_2$. Here $\widehat{D_1}$ is the closure of a braid-like crossingless diagram D_1 where all merges and splits have local ν-invariant 0, and D_2 is the union of foams U_{a_i,b_i}, $a_i, b_i \in \mathbb{R}_{>0}$ over all crossings (a_i, b_i) of D. The diagram D_1 is cobordant to a circle of some thickness and, hence, null-cobordant. This shows the surjectivity of ψ_1.

Composition $\psi \circ \psi_1$ is clearly identity. That and surjectivity of ψ_1 implies that $\psi_1 \circ \psi$ is the identity map. □

Remark 2.7. On the category side, we can follow Zakharevich [37, 38] and consider the category $\mathcal{C}_Z$ with objects – half-open interval $[a, b) \subset \mathbb{R}$. Morphisms are metric-preserving and order-preserving inclusions of intervals, and the *assembler* structure is given by pairs of morphisms $[a_1, b_1), [a_2, b_2) \xrightarrow{\psi_1, \psi_2} [a, b)$ that cover the interval without overlaps. More generally, given a Zacharevich assembler category $\mathcal{C}$, one can consider n-dimensional foams where facets are decorated by objects of $\mathcal{C}$, $(n-1)$-dimensional seams by coverings of $\mathcal{C}$, and so on. The cobordism group of $\mathcal{C}$-decorated n-foams should then be related to K-theory groups $K_n(\mathcal{C})$ as defined in [37]. Cobordism groups of foams decorated by objects of a category are introduced in [9].

Remark 2.8. It is possible to loosely compare the group $\mathsf{Aut}_{\mathsf{IET}}$ of IET transformations of the interval to the braid group and weighted 1-foams to links (note, though, that 1-foams are not embedded anywhere, while links are embedded in $\mathbb{R}^3$). The closure of a braid is an oriented link and the closure of an IET can be described by an oriented weighted 1-foam. The analogue of the Alexander theorem is simple: any oriented weighted 1-foam is the closure of some element of $\mathsf{Aut}_{\mathsf{IET}}$, and the analogue of the Markov theorem is straightforward to write down as well since 1-foams are not embedded in $\mathbb{R}^3$ (Markov's theorem is known in the harder case of graphs embedded in $\mathbb{R}^3$, see [12, 13, 7]). The analogue of the SAF invariant for oriented links is, perhaps, the sum of the linking numbers $\mathsf{lk}(L_i, L_j)$, $i < j$, over all pairs of components of a link L. This analogy is inspired by Figure 13, where (a, b) crossing adds $a \wedge b$ to the SAF invariant, similar to the formula for the linking number. The SAF invariant is preserved by the cobordisms of the oriented weighted 1-foams, as Theorem 2.6 shows. The linking number is invariant under some cobordisms in $\mathbb{R}^3 \times [0, 1]$ between the links in $\mathbb{R}^3$. More precisely, pick an ordered countable set S and equip a link L with a map $\psi : \mathsf{comp}(L) \longrightarrow S$ from its set $\mathsf{comp}(L)$ of connected components to S. Consider the cobordisms M between such links L, L' equipped with a map $\mathsf{comp}(M) \longrightarrow S$ which is compatible with the maps ψ, ψ' for its boundary links L, L'.

The S-linking number

$$\mathsf{lk}_S(L) \ := \sum_{i,j\,:\,\psi(i)<\psi(j)} \mathsf{lk}(L_i, L_j)$$

is invariant under such cobordisms.

Remark 2.9. In the definition of weighted foams, the abelian semigroup $(\mathbb{R}_{>0}, +)$ can be replaced by an arbitrary commutative semigroup $(H, +)$. One can then form the abelian group Cob^1_H of H-weighted oriented 1-foams modulo cobordisms. The latter are H-weighted oriented 2-foams with boundary. The above arguments extend to an isomorphism

$$H \wedge H \ \cong \ \mathsf{Cob}^1_H, \tag{13}$$

taking $a \wedge b$ to $[U_{a,b}]$. Here $H \wedge H$ is the abelian group generated by symbols $a \wedge b$, $a, b \in H$, with defining relations

$$\begin{aligned} a \wedge b + b \wedge a &= 0, \\ (a_1 + a_2) \wedge b &= a_1 \wedge b + a_2 \wedge b. \end{aligned}$$

In particular, the cobordism group of $\mathbb{R}$-decorated oriented 1-foams is isomorphic to that of $\mathbb{R}_{>0}$-decorated foams, since the natural map $\mathbb{R}_{>0} \wedge \mathbb{R}_{>0} \longrightarrow \mathbb{R} \wedge \mathbb{R}$ induced by the inclusion $\mathbb{R}_{>0} \hookrightarrow \mathbb{R}$ is an isomorphism.

There are several related ways to thicken an (oriented) $\mathbb{R}_{>0}$-weighted 1-foam to a 2-dimensional structure and a cobordism between such foams to a three-dimensional structure.

2.4.1 Lower limit topology

One can thicken an $\mathbb{R}_{>0}$-decorated 1-foam to a 2-dimensional structure by multiplying a 1-facet I carrying label a by $[0, a)$ and then gluing these products at vertices, see Figures 19, 20.

We equip intervals $[0, a)$ with the *lower limit topology* ℓ, with a basis of open sets given by $[a_1, b_1)$, with $0 \leq a_1 < b_1 \leq a$, see Munkres [22, Section 13]. With this topology, there

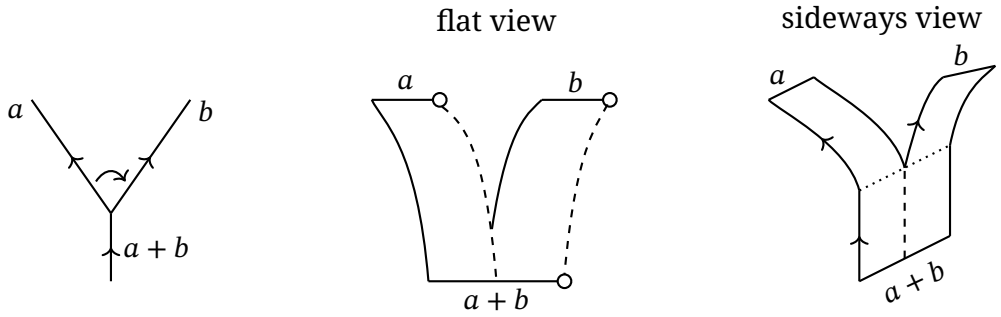


Figure 19: A split and its thickening. An IET 1-foam U can be replaced by a "surface" $T(U)$, which is locally the product $(0,1) \times [0,1)_\ell$, where ℓ denotes the lower limit topology, also see Figure 20.

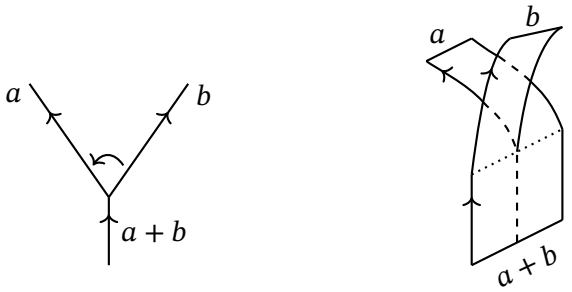


Figure 20: A split with the opposite thin edge orientation vs. the one in Figure 19 and its thickening, shown sideways.

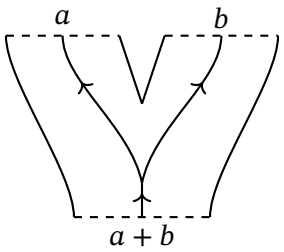


Figure 21: A weighted oriented 1-foam can be thickened to a weighted oriented train track on an oriented surface with boundary.

are homeomorphisms $[0,a) \sqcup [0,b) \cong [0,a+b)$ given by placing $[0,b)$ immediately to the right of $[0,a)$.

In this way, a 1-foam U as above is thickened to a topological space $T(U)$ which is locally homeomorphic to the product $[0,1)_\ell \times (0,1)$. A cobordism between two such 1-foams is thickened to a topological space locally homeomorphic to $[0,1)_\ell \times (0,1)^2$.

This thickening of 1-foams and 2-foams is related to the Zakharevich assembler category, see Remark 2.7 above and the discussion in the introduction.

The topological space $T(U)$ associated to a 1-foam U carries a foliation where connected components of the leaves are locally $x \times (0,1)$ for $0 \leq x < a$. If all leaves are compact (and then necessarily homeomorphic to $\mathbb{S}^1$), the foam U is null-cobordant. The opposite implication fails, since $U \sqcup U^!$ is null-cobordant for any U, where $U^!$ is the mirror image of U.

2.4.2 Train tracks on surfaces

A weighted oriented 1-foam can be thickened to an oriented train track [25] on a surface with boundary, see Figure 21. Transformations of unoriented train tracks that do not change the associated measured foliation or measured lamination [25, Sections 2.1, 2.3] can be interpreted as the cobordisms of train tracks in $S \times [0,1]$, where S is the surface that contains the train track.

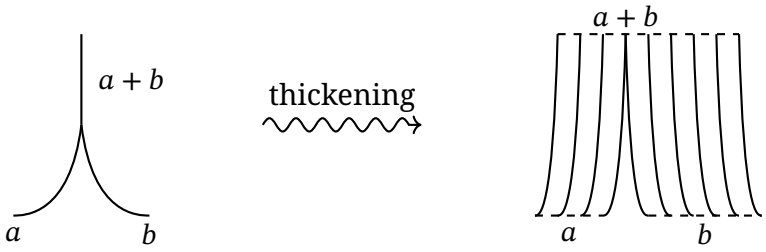


Figure 22: Thickening an embedded (a, b) vertex to a flow.

Remark 2.10. Interval exchange transformations can be thickened to very flat surfaces, or translation surfaces [39, 40], i.e., surfaces with a flat metric and singular points where total angles at these points are multiples of 2π. Oriented weighted 1-foams, equipped with additional data, can likewise be thickened to very flat surfaces (we omit the details).

3 Planar unoriented weighted foams and antisymmetric 2-brackets

Consider unoriented weighted 1-foams U embedded in the plane $\mathbb{R}^2$, and denote an embedded foam also by U. Such a foam U is analogous to a weighted unoriented train track on a surface [25], except that no conditions are imposed on the Euler characteristic of components of the complement of U in $\mathbb{R}^2$ (compare with [25, Section 1.1]). An embedded 1-foam can be thickened to an open subset of $\mathbb{R}^2$ with an unoriented bidirectional flow on it, see Figure 22.

By a cobordism between two unoriented embedded 1-foams U_0, U_1, we mean an unoriented embedded 2-foam $V \subset \mathbb{R}^2 \times [0,1]$ so that $V \cap (\mathbb{R}^2 \times \{i\}) = U_i$, $i = 0, 1$. Note that for any 1-foam U, the disjoint union $U \sqcup U^!$ of U with its mirror image is null-cobordant. See Figure 23 for an example of the mirror image of an unoriented embedded foam.

Denote by $\mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}_{>0}}$ the set of cobordism classes of unoriented embedded 1-foams ("up" in the superscript stands for *unoriented planar*). The disjoint union and mirror image

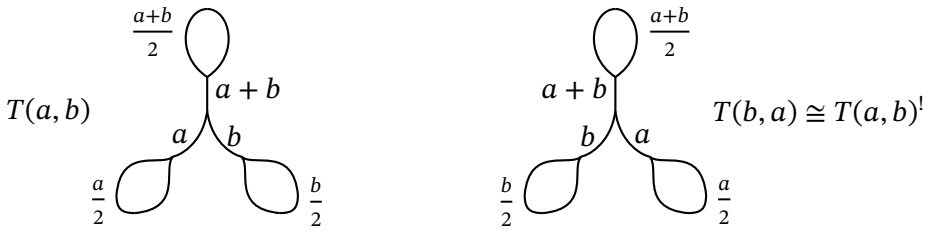


Figure 23: The tripod 1-foam $T(a,b)$ and its mirror image $T(a,b)^! \cong T(b,a)$.

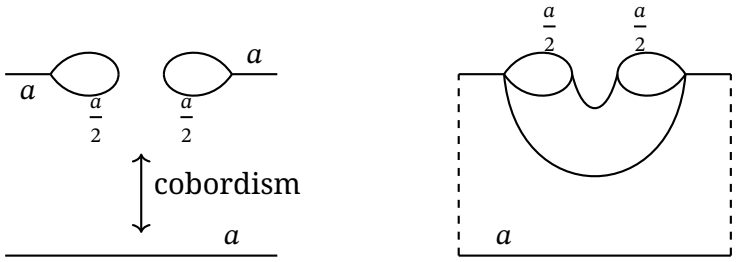


Figure 24: A cobordism between an interval and two looped half-intervals (lollipops).

operations turn this set into an abelian group. Denote by $[U]$ the image of a 1-foam U in that group.

In general, there is no obvious cobordism between U and $U^!$ (and we will see that $[U] \neq [U^!]$, in general).

For $a, b > 0$ denote by $T(a,b)$ the 1-foam shown in Figure 23, which we also call a *tripod foam*. Note that $T(a,b)^! := T(b,a)$.

Proposition 3.1. The group $\mathrm{Cob}^{1,\mathrm{up}}_{\mathbb{R}_{>0}}$ is generated by symbols $[T(a,b)]$ of tripod 1-foams over all $a, b > 0$.

Proof. The cobordism shown in Figures 24, 25 allows to convert an interval into two looped half-intervals. The loop at the end of an a-interval has thickness $a/2$. This cobordism can be applied at each edge of U, as shown in Figure 26 on the left, to cut U into a union of tripod foams and circles. Each circle can further be cut into a barbell foam (the latter is shown in Figure 27, together with a cobordism from it to the empty foam, in

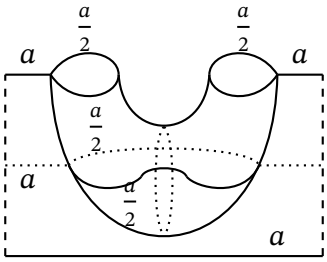


Figure 25: The cobordism from Figure 24 in more detail.

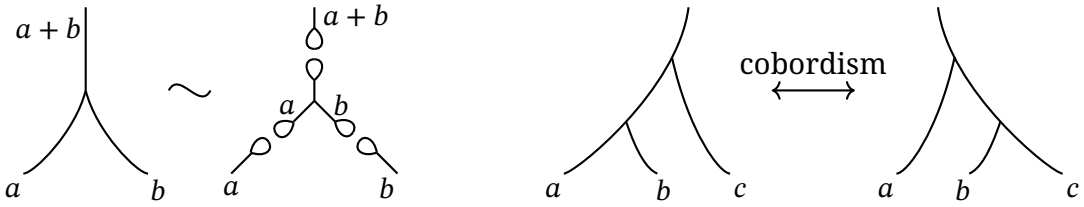


Figure 26: Left: splitting off a vertex of a planar 1-foam into the tripod $T(a,b)$, see also Figure 23 on the left. Right: existence of a cobordism between these 1-foams combined with the splittings on the left corresponds to the relation $[a,b]+[a+b,c]=[b,c]+[a,b+c]$ in $Z^2(\mathbb{R}_{>0})$, see proof of Proposition 3.3 below.

the top right corner of the figure). If a foam U has vertices $v_1,\dots,v_n$ with thin edges at the vertex v_i of thickness (a_i,b_i), going counterclockwise, then $[U]=\sum_{i=1}^{n}[T(a_i,b_i)]$. □

Remark 3.2. Figure 28 shows that $T(a,b)$ is cobordant to $T(a,b-a)$ if $a<b$. Passing to mirror images shows that $T(a,b)$ is cobordant to $T(a-b,b)$ if $a>b$. These cobordisms can be iterated to a foam cobordism version of the Euclidean division algorithm. In particular, iterating these operations we see that $T(a,b)$ is null-cobordant if $b\in\mathbb{Q}a$ (that is, if a and b are proportional over $\mathbb{Q}$). Cobordism between $T(a,a)$ and the empty foam is shown in the second row in Figure 27.

Consider 1-foams in the first two rows of Figure 16, ignoring the orientations of edges and orders of thin edges at vertices and instead viewing the 1-foams as planar (embedded

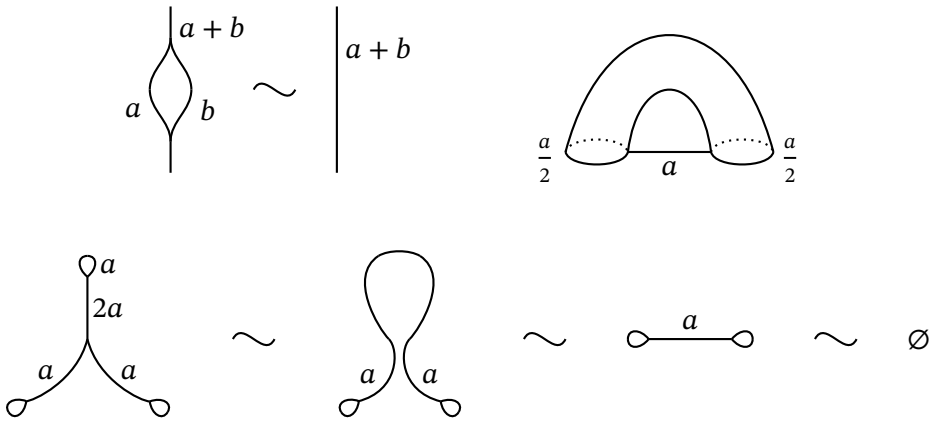


Figure 27: Top left: the two foams there are cobordant, which corresponds to the relation $[a,b]+[b,a]=0$ (see proof of Proposition 3.3 below). Top right: barbell 1-foam is null-cobordant (encoding the last cobordism in the bottom row). Bottom row: foam $T(a,a)$ is cobordant to a barbell foam and null-cobordant (relation $[a,a]=0$).

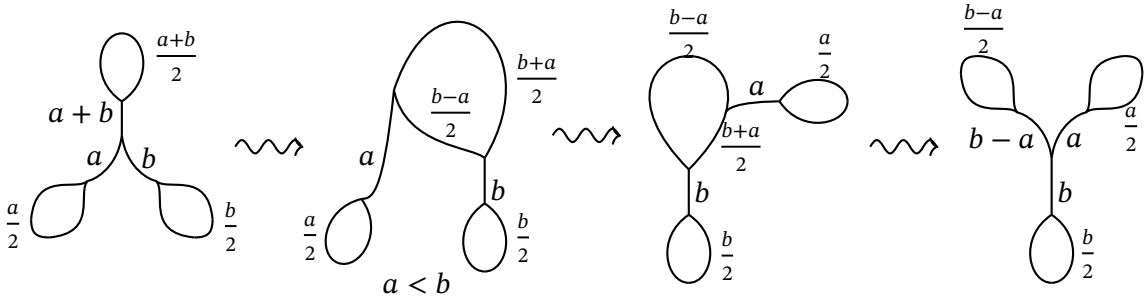


Figure 28: A cobordism between $T(a,b)$ and $T(a,b-a)$, for $a<b$.

in $\mathbb{R}^2$). These 1-foams are cobordant in pairs, via 2-foam cobordisms embedded in $\mathbb{R}^2\times[0,1]$. At the same time, breaking up these 1-foams along edges results in disjoint unions of the foam $T(x,y)$ for various $x,y\in\mathbb{R}_{>0}$. Passing to the cobordism group and replacing $\mathbb{R}_{>0}$ by a commutative semigroup H motivates the following definition.

Given a commutative semigroup $(H,+)$, denote by $Z^2(H)$ the abelian group with generators $[a,b]$, $a,b\in H$, and defining relations

$$[a,a] = 0, \quad a\in H, \tag{14}$$

$$[a,b]+[b,a] = 0, \quad a,b\in H, \tag{15}$$

$$[a,b]+[a+b,c] = [a,b+c]+[b,c], \quad a,b,c\in H. \tag{16}$$

Note that relation (14) does not imply the skew-commutativity relation (15) since the bracket $[a,b]$ is not bilinear. Equations (14) and (15) together are the strong version of the skew-commutativity property in the absence of bilinearity. Equation (16) is reminiscent of the 2-cocycle relation – the difference between the two sides can be interpreted as the signed boundary of a 3-simplex with oriented edges labeled $a,b,c,a+b,b+c,a+b+c$. This is explained in Figure 29. The analogue of symbol $[x,y]$ is an oriented triangle with oriented sides labeled $x,y,x+y$. The oriented boundary of a 3-simplex with sides labeled by a,b,c and their sums is the difference between the right-hand side and left-hand side of equation (16). Also see [11, Section 2.4].

We call $Z^2(H)$ *the antisymmetric 2-bracket* or *antisymmetric 2-cocycle of* H.

A homomorphism $f : H_1 \longrightarrow H_2$ of commutative semigroups induces a homomorphism $Z^2(H_1)\longrightarrow Z^2(H_2)$.

Proposition 3.3. The cobordism group $\mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}_{>0}}$ of planar unoriented weighted 1-foams is isomorphic to $Z^2(\mathbb{R}_{>0})$:

$$\mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}_{>0}} \cong Z^2(\mathbb{R}_{>0}), \tag{17}$$

taking $[T(a,b)]$ to $[a,b]$ for all $a,b>0$.

Proof. Consider the free abelian group Z on generators $[a,b]'$, over all $a,b\in\mathbb{R}_{>0}$. Proposition 3.1 says that there is a surjective homomorphism $\tau : Z\longrightarrow \mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}_{>0}}$ taking $[a,b]'$ to

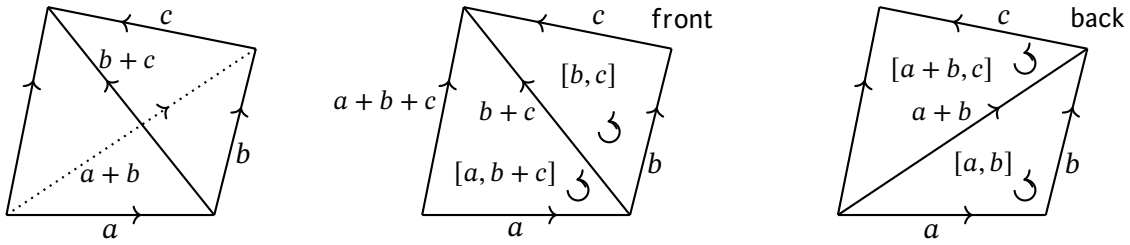


Figure 29: We have $\partial(\Delta^3) = \mathsf{front} - \mathsf{back} = [a, b+c] + [b, c] - ([a, b] + [a+b, c])$. Equation $\partial(\Delta^3) = 0$ is then the relation (16).

$[T(a, b)]$. Furthermore, relations (14)-(16) hold for the images of $[a, b]'$ under τ. Indeed, $T(a, a)$ is null-cobordant, giving the relation $\tau([a, a]') = 0$. The disjoint union $T(a, b) \sqcup T(b, a)$ is null-cobordant, implying

$$\tau([a, b]' + [b, a]') = 0. \tag{18}$$

It is convenient to pair up a- and b-lollipop ends of $T(a, b) \sqcup T(b, a)$ and pass to the 1-foam which is a split of $(a+b)$-strand into a- and b-strands, followed by the merge, see Figure 27 top left. There is a natural cobordism from the split-merge to the $(a+b)$-strand, which is another way to see the relation (18). Ignoring orientations and edge orders, this cobordism is depicted in the top left corner of Figure 16. Likewise, that

$$\tau([a, b]') + \tau([a+b, c]') = \tau([a, b+c]') + \tau([b, c]') \tag{19}$$

follows from the existence of a cobordism between the two ways to merge the parallel a, b, c-strands into $(a+b+c)$-strand, see Figure 26 on the right. For example, there is the one-vertex cobordism between these two 1-foams.

Consequently, the homomorphism τ descends to a surjective homomorphism, also denoted

$$\tau \ : \ Z^2(\mathbb{R}_{>0}) \longrightarrow \mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}>0}. \tag{20}$$

Vice versa, breaking a planar weighted 1-foam into tripods gives a map τ' from planar foams into $\mathbb{Z}$-linear combinations of symbols $[a, b]'$, and we would like to turn τ' into the inverse of τ. A cobordism between two 1-foams can be represented as a composition of

elementary cobordisms, including vertex cobordisms, singular saddles, cups and caps, and the usual saddle, cup and cap cobordisms between 1-manifolds. These cobordisms do not change the linear combination of symbols $[a,b]$ associated to a 1-foam, when viewed as an element of $Z^2(\mathbb{R}_{>0})$.

Note that the relation $[a,a]=0$ in (14) does not come from any elementary cobordism. The tripod $T(a,a)$ is null-cobordant, however, as shown in Figure 27. This discrepancy has the following explanation. When breaking a tripod $T(a,b)$ along every edge to construct the map τ', one adds three more terms to $[a,b]'$ due to the three lollipop vertices of the tripod, so that the composition of τ and τ' is

$$[a,b]' \xrightarrow{\tau} [T(a,b)] \xrightarrow{\tau'} [a,b]' + [a/2,a/2]' + [b/2,b/2]' + [(a+b)/2,(a+b)/2]'.$$

In particular, the composition $\tau'\tau$ differs from the identity due to the presence of three terms $[x,x]'$ for $x \in \{a/2, b/2, (a+b)/2\}$. Setting these terms to 0 in $Z^2(\mathbb{R}_{>0})$ makes the composition $\tau'\tau = \mathsf{id}$, where now τ' is a well-defined map

$$\tau' \ : \ \mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}>0} \longrightarrow Z^2(\mathbb{R}_{>0}), \quad [T(a,b)] \xrightarrow{\tau'} [a,b]. \tag{21}$$

In the other direction, it is clear that $\tau\tau' = \mathsf{id}$. Consequently, the homomorphism τ in (20) is an isomorphism. □

Extending from $\mathbb{R}_{>0}$ to $\mathbb{R}$ and adding bilinearity relations on the symbols $[a,b]$, so that, in addition $[a_1+a_2,b] = [a_1,b]+[a_2,b]$, gives a surjective homomorphism

$$\theta' \ : \ Z^2(\mathbb{R}_{>0}) \longrightarrow \mathbb{R} \wedge_{\mathbb{Z}} \mathbb{R} \cong \mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R}, \tag{22}$$

and, consequently, a surjective homomorphism

$$\theta \ : \ \mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}_{>0}} \longrightarrow \mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R} \tag{23}$$

taking $[T(a,b)]$ to $a \wedge b$ (compare with the SAF invariant, see Section 2). This allows to show that some unoriented planar 1-foams are not null-cobordant.

Corollary 3.4. Planar unoriented foam $T(a,b)$ for $a,b \in \mathbb{R}_{>0}$ is not null-cobordant if $b \notin \mathbb{Q}a$.

That is, if the unoriented planar foam $T(a,b)$ for $a,b \in \mathbb{R}_{>0}$ is null-cobordant, then $b \in \mathbb{Q}a$.

It turns out that the bracket $[a,b]$ is almost bilinear, as explained by the following result.

Proposition 3.5. The kernels of θ' and θ consist of elements of order at most two. For any $a, b_1, b_2 \in \mathbb{R}_{>0}$ the following relation holds in $Z^2(\mathbb{R}_{>0})$:

$$2([a, b_1 + b_2] - [a, b_1] - [a, b_2]) \;=\; 0. \tag{24}$$

Proof. Consider the following three equations:

$$[a, b_1 + b_2] + [b_1, b_2] \;=\; [a + b_1, b_2] + [a, b_1], \tag{25}$$

$$[b_1, b_2 + a] + [b_2, a] \;=\; [b_1 + b_2, a] + [b_1, b_2], \tag{26}$$

$$[b_1, a + b_2] + [a, b_2] \;=\; [a + b_1, b_2] + [b_1, a]. \tag{27}$$

Equation (25) is the 2-cocycle relation, for a, b_1, b_2. Equation (26) is given by cyclicly permuting the terms of the previous equation, $a \longmapsto b_1 \longmapsto b_2 \longmapsto a$. Equation (27) is given by transposing a and b_1 in (25). Writing down the linear combination (25) + (26) − (27) and using that the bracket is antisymmetric gives relation (24).

This argument is borrowed from [5], which shows bilinearity of the difference $[a,b] - [b,a]$ assuming only the 2-cocycle equation (16) for all $a, b \in H$, where H is an abelian group. When the 2-cocycle is, additionally, antisymmetric, via equation (15), the difference $[a,b] - [b,a] = 2[a,b]$. □

The proposition tells us that the bracket $[a,b]$ is "almost" bilinear, with the difference $[a, b_1 + b_2] - [a, b_1] - [a, b_2]$ either 0 or an element of order 2.

Corollary 3.6. The foam $U \sqcup U$, where

$$U \;=\; T(a, b_1 + b_2) \sqcup T(b_1, a) \sqcup T(b_2, a),$$

is null-cobordant for any $a, b_1, b_2 > 0$.

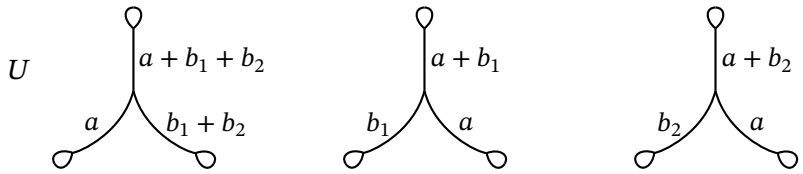


Figure 30: The 1-foam $U = T(a, b_1 + b_2) \sqcup T(b_1, a) \sqcup T(b_2, a)$ in Corollary 3.6.

The foam U is shown in Figure 30.

We do not know whether the scalar 2 can be dropped from equation (24), so that $[a, b]$ is bilinear in a, b. That would be equivalent to foams U in Corollary 3.6 being null-cobordant for all $a, b_1, b_2 > 0$.

To further study abelian groups in Proposition 3.3, it is natural to extend the possible weights of the foam facets from positive to all real numbers. First, we discuss the group $Z^2(H)$ for general commutative semigroups H, having $(\mathbb{R}, +)$ in mind. Note that Proposition 3.5 holds for any commutative semigroup H in place of $\mathbb{R}_{>0}$, so that there is an exact sequence

$$0 \longrightarrow \ker \theta' \longrightarrow Z^2(H) \xrightarrow{\theta'} H \wedge' H \longrightarrow 0,$$

with $2x = 0$ for $x \in \ker \theta'$. Here $H \wedge' H$ is the abelian group which is the quotient of the abelian group closure of $H \otimes_{\mathbb{Z}} H$ by the relations $a \wedge' b + b \wedge' a = 0$ and $a \wedge' a = 0$, by analogy with (14), (15). Symbol $\wedge'$ is used instead of $\wedge$ since the relation $a \wedge a = 0$ is usually not imposed in the definition of the exterior square (but follows for 2-divisible semigroups).

If $0 \in H$, then (16) with the particular triple $(a, b, c) = (a, 0, b)$ implies that $[a, 0] = [0, b]$ for all $a, b \in H$. Specializing to $b = 0$ gives

$$[a, 0] \;=\; [0, a] \;=\; 0, \; \forall a \in H. \tag{28}$$

Proposition 3.7. Assume that $0 \in H$. Then

1. $[a, 0] = 0$ for all $a \in H$,

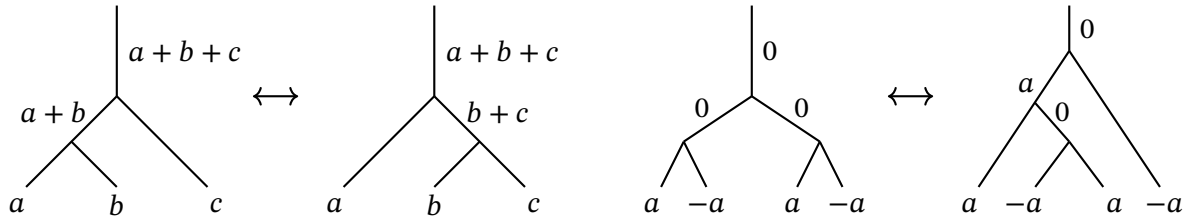


Figure 31: Left: cobordance of these 1-foams matches the 2-cocycle equation (16). Right: Two trees merging $(a, -a, a, -a)$ to 0.

2. If $-a \in H$ (i.e., a is invertible in H), then

$$2[a, -a] = 0, \tag{29}$$

$$[b, -a] = [a, b - a] + [a, -a] \text{ for all } b \in H, \tag{30}$$

$$[2a, -2a] = 0, \tag{31}$$

3. If $-a, -b \in H$, then

$$[-a, -b] = [a, b] + [a, -a] + [b, -b] - [a + b, -a - b]. \tag{32}$$

Proof. See (28) for 1. Notice that relation (16) can be visualized as the "associativity" property for merging a, b, c into $a + b + c$ in two possible ways, where a vertex merging x, y contributes $[x, y]$ to the sum, see Figure 31.

Iterating this associativity relation gives us a relation between any two tree diagrams for merging $(a_1, \dots, a_n)$ into $a_1 + \dots + a_n$. Now apply the relation to the two trees shown in Figure 31 on the right merging $(a, -a, a, -a)$ to 0 and use that $[a, -a] + [-a, a] = 0$ and $[b, 0] = 0$ for any b to conclude that $2[a, -a] = 0$, resulting in relation (29).

For the relation (30), apply (16) to $(b-a, a, -a)$ to get $[b-a, a]+[b, -a] = [b-a, 0]+[a, -a]$. For the relation (31), two of the ways to merge $(a, a, -a, -a)$ to 0 give

$$[a, a] + [-a, -a] + [2a, -2a] = [a, -a] + [a, 0] + [a, -a], \tag{33}$$

resulting in $[2a, -2a] = 2[a, -a] = 0$.

For the relation (32), apply (16) to $(a, -a, -b)$ and $(-a - b, a, b)$. □

Notice that, modulo terms $[x,-x]$, relations (30) and (32) are $[b,-a]\sim[a,b-a]$ and $[-a,-b]\sim[a,b]$.

Remark 3.8. We give an example for which $[a,-a]\neq 0$. Let $H=(\mathbb{Z}/4,+)=\{0,1,2,3\}$. It is tedious but straightforward to check that the map

$$\psi([a,b])=\begin{cases}0 & \text{if } a=0 \text{ or } b=0 \text{ or } a=b,\\ 1 & \text{otherwise}\end{cases} \tag{34}$$

extends to a homomorphism $\psi : Z^2(\mathbb{Z}/4)\longrightarrow \mathbb{Z}/2$. Under this homomorphism, the image of $[1,-1]=[1,3]$ is nontrivial. Via the surjective homomorphism $\mathbb{Z}\longrightarrow\mathbb{Z}/4$ we see that $[1,-1]$ in nontrivial in $Z^2(\mathbb{Z})$ as well. Consequently, $[a,-a]$ is not always 0 in $Z^2(H)$ for $a,-a\in H$.

Elements $[a,-a]$, over all $a,-a\in H$, generate a 2-torsion subgroup in $Z^2(H)$, which we can denote $Z^2_{-}(H)$. This subgroup is trivial if H is 2-divisible, in view of the relation (31). In particular, it is trivial for $H=(\mathbb{R},+)$.

We denote by $\mathbb{R}_{>0}$ the semigroup $(\mathbb{R}_{>0},+)$ and by $\mathbb{R}$ the group $(\mathbb{R},+)$. Semigroup $(\mathbb{R}_{>0},+)$ is not a monoid, that is, $0\notin\mathbb{R}_{>0}$. The inclusion $\mathbb{R}_{>0}\subset\mathbb{R}$ induces a homomorphism

$$\rho \ : \ Z^2(\mathbb{R}_{>0})\longrightarrow Z^2(\mathbb{R}). \tag{35}$$

To differentiate between the elements of the two groups, denote by $[a,b]_{\mathbb{R}}$ the symbol of the pair $a,b\in\mathbb{R}$ viewed as an element of $Z^2(\mathbb{R})$. The map ρ is given by $\rho([a,b])=[a,b]_{\mathbb{R}}$ for $a,b>0$.

Corollary 3.9. In $Z^2(\mathbb{R})$ and for $a,b>0$, the following relations hold:

$$[a,-b]_{\mathbb{R}}=\begin{cases}[b,a-b]_{\mathbb{R}} & \text{if } a>b,\\ [b-a,a]_{\mathbb{R}} & \text{if } a<b,\\ 0 & \text{if } a=b,\end{cases} \tag{36}$$

$$[-a,b]_{\mathbb{R}}=-[b,-a]_{\mathbb{R}},\quad [-a,-b]_{\mathbb{R}}=[a,b]_{\mathbb{R}}. \tag{37}$$

Proof. These relations are obtained by dropping off the terms $[x,-x]$ from the relations in Proposition 3.7. Terms $[x,-x]_{\mathbb{R}}=0$ since $\mathbb{R}$ is 2-divisible. □

Corollary 3.9 implies that ρ is surjective, since the symbol $[a,b]_{\mathbb{R}}$ with at least one of a,b negative can be written as $\pm\rho([a',b'])$ for suitable $a',b'\in\mathbb{R}_{>0}$, or $[a,b]_{\mathbb{R}}=0$.

Proposition 3.10. The homomorphism

$$\rho \ : \ Z^2(\mathbb{R}_{>0}) \longrightarrow Z^2(\mathbb{R})$$

induced by the inclusion $\mathbb{R}_{>0}\subset\mathbb{R}$ is an isomorphism.

Proof. Corollary 3.9 relations can be used to define a map from symbols $[a,b]_{\mathbb{R}}$ with $a,b\in\mathbb{R}$ to signed symbols $[a,b]$ with positive a,b. Consider the map δ defined on symbols as follows and assuming $a,b>0$:

$$\delta([a,b]_{\mathbb{R}}) \ = \ \delta([-a,-b]_{\mathbb{R}}) \ = \ [a,b], \tag{38}$$

$$\delta([a,-b]_{\mathbb{R}}) \ = \ [b,a-b], \text{ if } a>b, \tag{39}$$

$$\delta([a,-b]_{\mathbb{R}}) \ = \ [b-a,a], \text{ if } a<b, \tag{40}$$

$$\delta([-a,b]_{\mathbb{R}}) \ = \ -\delta([b,-a]_{\mathbb{R}}), \tag{41}$$

$$\delta([a,-a]_{\mathbb{R}}) \ = \ 0. \tag{42}$$

We claim that δ extends to a well-defined homomorphism $\delta \ : \ Z^2(\mathbb{R}) \longrightarrow Z^2(\mathbb{R}_{>0})$. This map respects the relations (14) and (15). A tedious case-by-case verification shows that it also respects the relation (16). For example, consider relation (16) for the triple $(a,-b,c)$ where $c>b>a>0$. To check that

$$\delta([a,-b]_{\mathbb{R}})+\delta([a-b,c]_{\mathbb{R}})=\delta([a,c-b]_{\mathbb{R}})+\delta([-b,c]_{\mathbb{R}}),$$

we compute the two sides:

$$\begin{aligned} \mathrm{LHS} &= [b-a,a]+[a+c-b,b-a],\\ \mathrm{RHS} &= [a,c-b]+[c-b,b], \end{aligned}$$

and write

$$\begin{aligned}[a,c-b]+[c-b,b] &= ([c-b,b]+[a,b-a])+[a,b-c]-[a,b-a]\\ &= ([c-b,a]+[a+c-b,b-a])+[a,b-c]-[a,b-a]\\ &= [a+c-b,b-a]+[b-a,a] \;=\; \mathsf{LHS}.\end{aligned}$$

The case $a>b>c$ follows by symmetry, and other cases to consider are $a>b, c>b$; $b>a+c$; $a+c>b, b>a, b>c$. All of them together take care of the relation (16) when only the middle number is negative. The case $(-a,-b,-c)$, i.e., all three numbers are negative, is trivial, but there are many other cases. They follow via straightforward computations which are omitted. □

The group $Z^2(H)$ depends only on the isomorphism class of abelian semigroup H. Thinking of $\mathbb{R}$ an abelian group and using the axiom of choice, one can write $\mathbb{R}\cong\oplus_J\mathbb{Q}$, where the index set J is uncountable. Consequently, $Z^2(\mathbb{R}_{>0})\cong Z^2(\mathbb{R})\cong Z^2(\oplus_J\mathbb{Q})$, giving a more symmetric presentation of $Z^2(\mathbb{R}_{>0})$ since we can now work with positive and *non-positive* generators satisfying the relations (14) - (16). This does not give an explicit description of $Z^2(\mathbb{R}_{>0})$, just a description with more internal symmetries, but in our study of this group we stop here. A natural question would be to understand the kernel of the surjective homomorphism $\theta' : Z^2(\mathbb{R}_{>0})\longrightarrow \Lambda^2_{\mathbb{Q}}(\mathbb{R})$ sending $[a,b]$ to $a\wedge b$. From Proposition 3.5 we know that $2x=0$ for any element $x\in\ker(\theta')$.

Remark 3.11. Note that $Z^2(\mathbb{Q}_{>0})\cong Z^2(\mathbb{Q})=0$. This can be derived from all tripods $T(a,b)$ for $a,b\in\mathbb{Q}_{>0}$ being null-homotopic. A related observation is that thickening $T(a,b)$ with rational a,b results in a foliated planar surface with all leaves closed and diffeomorphic to $\mathbb{S}^1$.

Proposition 3.10 shows that passing from $\mathbb{R}_{>0}$ to $\mathbb{R}$ does not change the group Z^2. Let us consider unoriented planar 1-foams where edges are labeled by real numbers rather than just positive numbers (planar $\mathbb{R}$-weighted 1-foams). A cobordism between two such foams is given by an unoriented $\mathbb{R}$-decorated 2-foam in $\mathbb{R}^2\times[0,1]$. An $\mathbb{R}$-weighted 2-foam

Figure 32: Going from an $\mathbb{R}$-weighted 1-foam to an $\mathbb{R}_{\geq 0}$-weighted 1-foam (continues in Figures 33, 34).

also has vertices with local structure as in Figures 7 and 5, but now a, b, c are arbitrary real numbers, possibly 0. Denote by $\mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}}$ the cobordism group of $\mathbb{R}$-weighted planar unoriented 1-foams. There is a natural homomorphism

$$\iota \ : \ \mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}_{>0}} \longrightarrow \mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}} \tag{43}$$

given by viewing $\mathbb{R}_{>0}$-weighted 1- and 2-foams as $\mathbb{R}$-weighted foams. Likewise, there is a homomorphism

$$\tau_{\mathbb{R}} \ : \ Z^2(\mathbb{R}) \longrightarrow \mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}} \tag{44}$$

defined analogously to the homomorphism (20). The map $\tau_{\mathbb{R}}$ takes the symbol $[a,b]_{\mathbb{R}}$ to the concordance class of the tripod $T(a,b)$, where now the weights may be non-positive.

Theorem 3.12. Maps ι and $\tau_{\mathbb{R}}$ are isomorphism of abelian groups.

Proof. That $\tau_{\mathbb{R}}$ is an isomorphism can be shown in the same way as for τ, see the proof of Proposition 3.3. Next, observe that formulas in Corollary 3.9 convert symbols $[x,y]_{\mathbb{R}}$ when one of both x, y are negative into symbols with positive entries. We now define the foam counterpart of these formulas. Start with an $\mathbb{R}$-weighted 1-foam U and convert it to an $\mathbb{R}_{\geq 0}$-weighted 1-foam U° as follows. First, convert each line a into a line of weight $|a|$, for $a \in \mathbb{R}^*$, see Figure 32.

At the vertices of U, only the edges of *negative* weight are bent to the opposite side to retain the balance of weights at a vertex. If an edge has positive weight, it is not bent. Figure 33 shows how a single negative edge is bent at a vertex. Figure 34 shows modifications at a vertex if two out of three edges have negative weights. In Figure 35

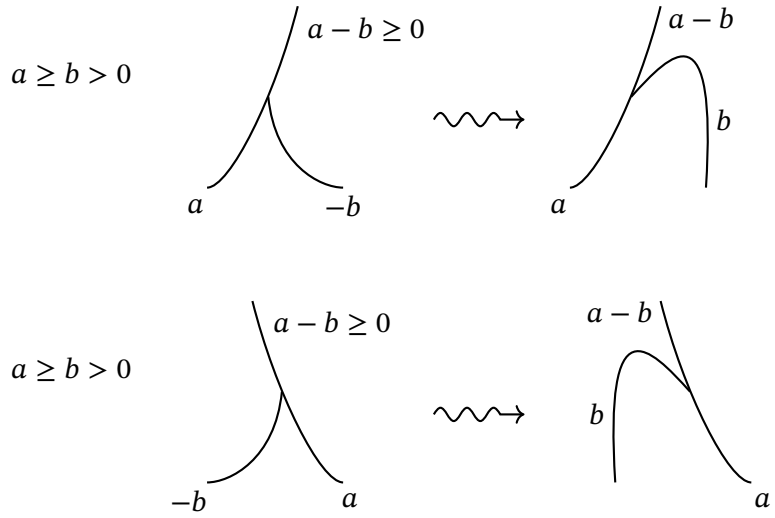


Figure 33: Converting a vertex with one negative weight into a positive vertex; compare with (36), top equality. If $a = b$, the line $a - b = 0$ is removed, see also Figure 35.

top row, we see that an $(a, -a)$ vertex gets smoothed out into part of a segment, and in Figure 35 bottom row, that no bending is necessary at a $(-a, -b)$ vertex, but just weight reversal at all three edges of the vertex.

The foam U° may have edges (and circles) of weight 0. A circle of any weight is null-cobordant even if there is a 1-foam inside the disk that it bounds, by converting the circle to a barbell. Given a 0-edge e, applying Figure 26 (left) transformation at the two endpoints of e produces tripod foams $T(a_1, 0)$ and $T(a_2, 0)$ (or their reflections) for some

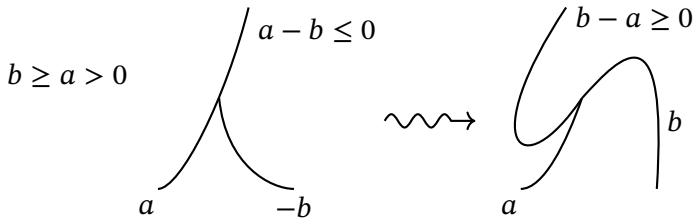


Figure 34: Vertex replacement when two out of three edges are negative. If $a = b$, the line $b - a = 0$ is erased.

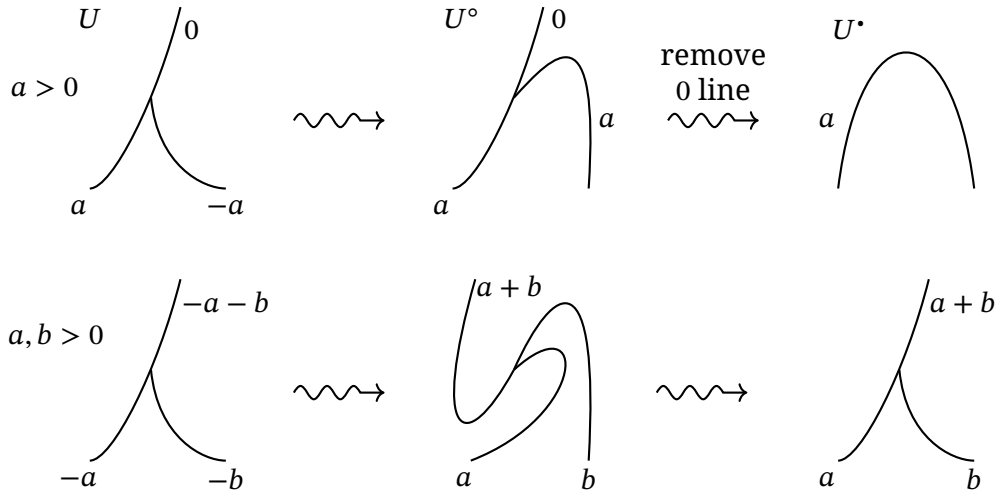


Figure 35: Top row: converting $(a, -a)$-vertex to an undecorated a-segment (analogous to the relation $[a, -a]_{\mathbb{R}} = 0$) by flipping (bending) the $-a$ edge and removing the 0 edge. Bottom row: at a $(-a, -b)$-vertex, all edges are negative so one bends all three edges or one simply reverses all weights (corresponding to the relation $[-a, -b]_{\mathbb{R}} = [a, b]_{\mathbb{R}}$, see Corollary 3.9).

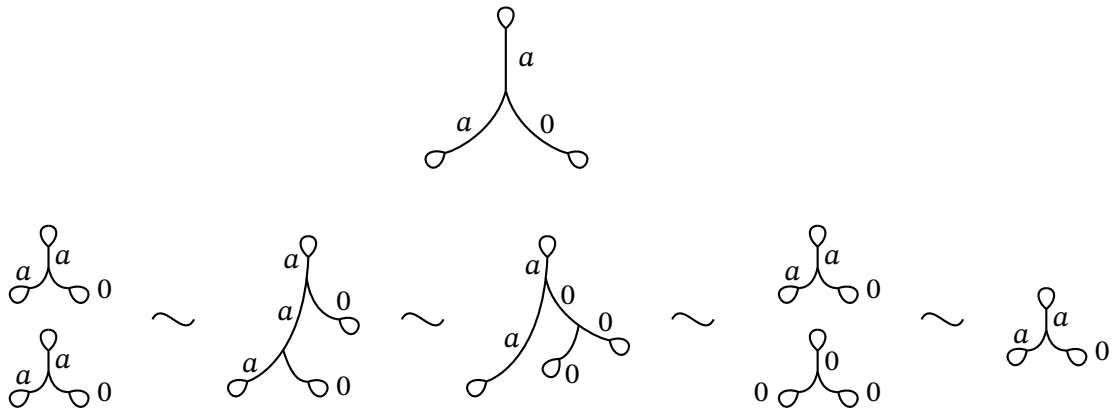


Figure 36: Top: $T(a,0)$ tripod. Bottom row: a cobordism from $T(a,0) \sqcup T(a,0)$ to $T(a,0)$. The first move is the cobordism in Figure 24, the second is $(a,0,0)$-vertex, the third move cuts a 0 edge creating lollipops. The last move vanishes $T(0,0)$, since it is null-cobordant. Gluing $T(a,0) \times [0,1]$ to this cobordism by capping off one $T(a,0)$ on each side shows that $T(a,0)$ is null-cobordant.

a_1, a_2. These foams are null-cobordant (see Figure 36), and cobordant to barbell foams with weights a_1, a_2 (the latter are null-cobordant as well, see Figure 27). Inserting these barbell foams back into the original 1-foam and composing these cobordisms shows that an $\mathbb{R}_{\geq 0}$-weighted 1-foam V with a 0-weight edge e is cobordant to the same foam with edge e deleted. Thus, all edges and circles of weight 0 (components of weight 0) can be deleted from an $\mathbb{R}_{\geq 0}$-weighted foam V, resulting in a cobordant $\mathbb{R}_{>0}$-foam. In particular, this is shown as the second step in the top row of Figure 35.

Denote by $U^\bullet$ the foam U° with weight 0 components removed. The map $U \longmapsto U^\bullet$ from planar $\mathbb{R}$-weighted 1-foams to planar $\mathbb{R}_{>0}$-weighted 1-foams needs to be extended to cobordisms between 1-foams, that is, to 2-foams with boundary.

Suppose that F is an $\mathbb{R}$-weighted 2-foam with boundary U, unoriented and embedded in $\mathbb{R}^2 \times [0,1)$, with $U \cong \partial F \cong F \cap (\mathbb{R}^2 \times \{0\})$. We convert all facets of F with negative labels $-a$ to positive labels $a > 0$.

At each seam of F, two facets merge into one. If one or two of these facets had negative

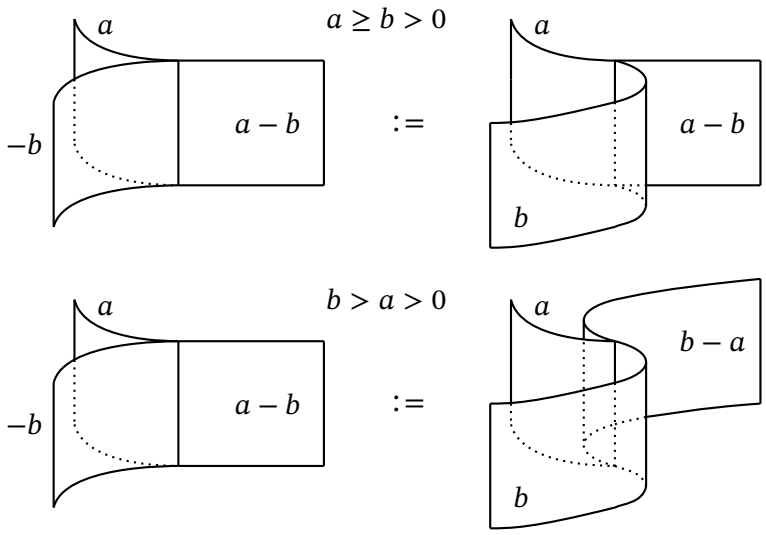


Figure 37: Top row: Converting a seam with a, $-b$ thin facets, with $a \geq b > 0$, into a seam with b, $a-b$ thin facets. Bottom row: Converting a seam when $b > a$.

weights, we make these facets approach the seams from the opposite side, by taking the rules in Figures 33, 34, 35 and multiplying them by the interval to get the corresponding rules for 2-foams. These modifications are depicted in Figure 37.

Next, one produces modification rules at vertices of F, where facets have weights $a, b, c, a+b, b+c, a+b+c$, for some $a, b, c \in \mathbb{R}$. Taking the link of a vertex results in a 2-foam $L(a,b,c)$ on the 2-sphere (see Figure 8 on the left, with orientations and thin edge orders at nodes dropped). Converting it to $L(a,b,c)^{\bullet}$, one needs to check that it is null-cobordant through a $\mathbb{R}_{>0}$-weighted foam and pick a particular cobordism to replace each (a,b,c)-vertex of an $\mathbb{R}$-weighted 2-foam. This is done on a case-by-case basis, and the rest of the proof closely resembles that of Proposition 3.10 towards the end. Here we provide the cobordisms in two out of the many cases here. Instead of the cobordism from $L(a,b,c)^{\bullet}$ to the empty foam we depict cobordisms between two possible ways to merge a, b, c edges into the $a+b+c$ edge, see Figure 26 on the right.

We consider the case when the middle number is negative and write it as $-b$. Since intermediate edges are $a-b$ and $c-b$, there are four cases to consider:

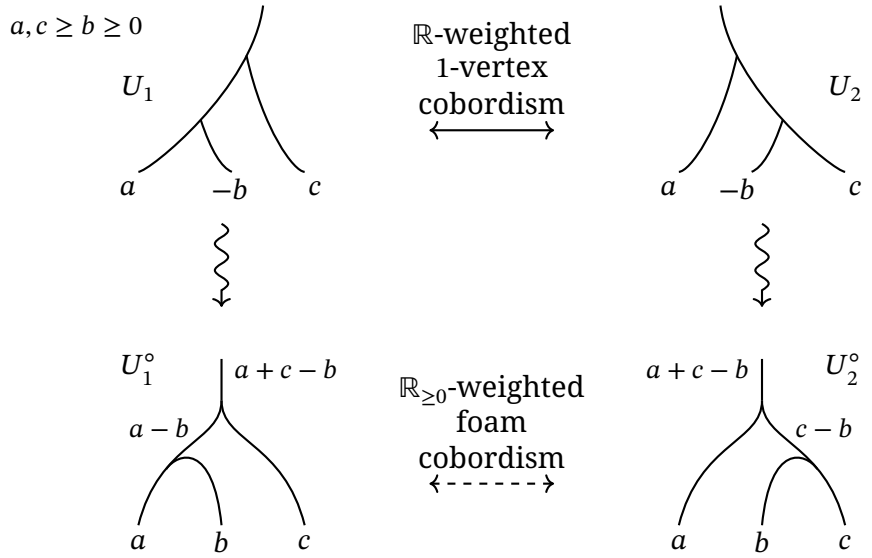


Figure 38: Vertical arrows go from $\mathbb{R}$-weighted foams U_1, U_2 to $\mathbb{R}_{\geq 0}$-weighted foams U_1°, U_2°. Bottom horizontal arrow indicates that we need to produce an $\mathbb{R}_{\geq 0}$-weighted cobordism between foams U_1°, U_2°. Such a cobordism is shown in Figure 39.

1. $a \geq b, c \geq b \geq 0$,
2. $a \geq b \geq c \geq 0$ (case $c \geq b \geq a \geq 0$ is given by reflection),
3. $b \geq a, b \geq c, a + c \geq b$, $a, c \geq 0$,
4. $b \geq a + c$, $a, c \geq 0$.

In each of the cases, one constructs an $\mathbb{R}_{>0}$-weighted cobordism between the corresponding $\mathbb{R}_{>0}$-weighted one foams. Schematically, Figure 38 shows what needs to be done in case (1) above, and similar for the other cases.

Cobordisms between the diagrams that replace the corresponding vertices are shown for cases (1) and (3) in Figures 39 and 40 via sequences of their cross-sections.

Further cases include $L(-a, b, c)$, with the first number negative (that of $L(a, b, -c)$ follows by reflection symmetry). Another case is when two numbers out of three are

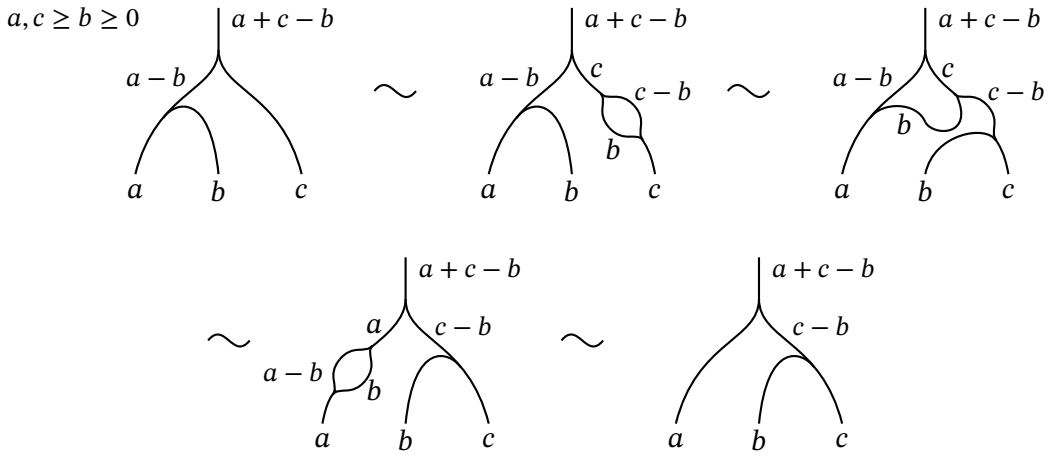


Figure 39: Replacing a vertex of an $\mathbb{R}$-weighted 2-foam for weights $(a, -b, c)$ with $a, c \geq b \geq 0$ by an $\mathbb{R}_{\geq 0}$-weighted 2-foam.

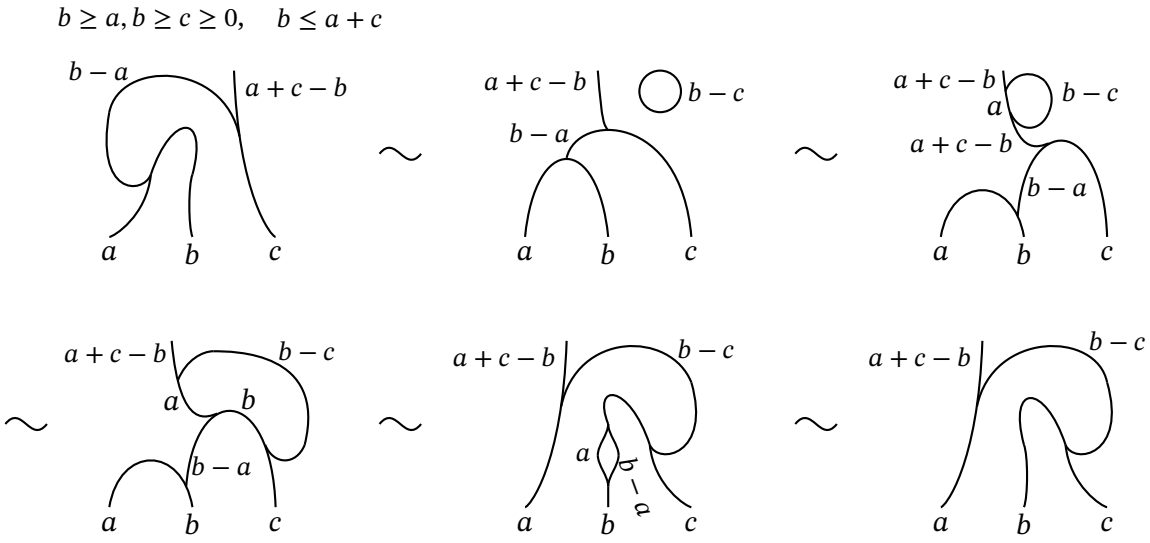


Figure 40: Vertex replacement for weights $(a, -b, c)$ with $b \geq a, b \geq c \geq 0$ and $b \leq a + c$.

negative. The case $L(-a,-b,-c)$ is easy, since no modifications are done at any of the four vertices of the boundary foam. (It is likely that additional symmetries of $L(a,b,c)^\bullet$ can be used to reduce the number of cases but we have not checked that.)

This procedure converts $\mathbb{R}$-weighted 2-foam F with boundary $U \cong \partial F$ to an $\mathbb{R}_{\geq 0}$-weighted embedded foam, denoted F°, with boundary U°.

Next, 0-facets of F° can be removed as well, by analogy and extending our deletion of 0-facets of the foam U°. The resulting 2-foam $F^\bullet \subset \mathbb{R}^2 \times [0,1)$ is $\mathbb{R}_{>0}$-weighted, with the boundary $U^\bullet \subset \mathbb{R}^2 \times \{0\}$. Consequently, our procedure for converting $\mathbb{R}$-weighted 1- and 2-foams into $\mathbb{R}_{>0}$-weighted 1- and 2-foams gives a homomorphism

$$\iota^\bullet \ : \ \mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}} \longrightarrow \mathsf{Cob}^{1,\mathsf{up}}_{\mathbb{R}_{>0}}.$$

It is clear that $\iota^\bullet \circ \iota = \mathsf{id}$, since $\iota^\bullet$ on foams with all facets positive is the identity map.

To show that $\iota \circ \iota^\bullet = \mathsf{id}$ it suffices to check that ι is surjective. For that, it is enough to show that $[T(a,b)]$ is in the image of ι for all $a,b \in \mathbb{R}$. Consider the tripod $T(a,-b)$ for $a \geq b \geq 0$. There are two ways to merge strands $a,-b,b$ into an $a-b+b=a$ strand, with the one-vertex 2-foam cobordism connecting the two ways to merge. This translates into a cobordism between $\mathbb{R}$-weighted 1-foams:

$$T(a,-b) \sqcup T(a-b,b) \sim T(a,0) \sqcup T(-b,b).$$

Foam $T(a-b,b)$ has positive weights. Foam $T(a,0)$ is null-cobordant via $\mathbb{R}_{\geq 0}$-weighted foams, see Figure 36. Foam $T(-b,b)$ is null-cobordant, since $[-b,b]_{\mathbb{R}} = 0$ and $[T(-b,b)]$ is the image of $[-b,b]_{\mathbb{R}}$ under the homomorphism $\tau_{\mathbb{R}}$ in (44). Alternatively, computation in (33) with $-b/2$ in place of a can be converted into a description of a cobordism from $T(-b,b)$ to the empty 1-foam. Consequently, $T(a,-b)$ is cobordant via an $\mathbb{R}$-weighted 2-foam to an $\mathbb{R}_{>0}$-weighted 1-foam $T(b,a-b)$. Reflecting in the plane shows that $T(-b,a)$ is cobordant to $T(a-b,b)$. We leave the remaining cases: $T(a,-b)$ with $b>a>0$ and $T(-a,-b)$, $a,b>0$ to the reader.

Consequently, ι and $\iota^\bullet$ are mutually-inverse isomorphisms. This completes the proof of Theorem 3.12. □

Our constructions and results can be summarized into the following statement.

Theorem 3.13. There is a commutative diagram of isomorphisms of abelian groups

$$\begin{array}{ccc} Z^2(\mathbb{R}_{>0}) & \xrightarrow{\tau} & \mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}>0} \\ \downarrow{\scriptstyle \rho} & & \downarrow{\scriptstyle \iota} \\ Z^2(\mathbb{R}) & \xrightarrow{\tau_{\mathbb{R}}} & \mathsf{Cob}^{1,\mathrm{up}}_{\mathbb{R}} \end{array} \tag{45}$$

The top arrow is given by (20), the bottom arrow $\tau_{\mathbb{R}}$ is the map (44). The left arrow is the map ρ in Proposition 3.10, the right arrow is given by (43).

In particular, the cobordism groups of $\mathbb{R}$-weighted and $\mathbb{R}_{>0}$-weighted planar unoriented 1-foams are isomorphic, and they are isomorphic to the corresponding abelian groups generated by the symbols $[a, b]$ over either all positive real $a, b > 0$ or, alternatively, all real a, b, subject to relations (14) - (16) in each of the two cases.

Remark 3.14. In the isomorphisms τ or $\tau_{\mathbb{R}}$ in Theorem 3.13, commutative semigroup $\mathbb{R}_{>0}$ or commutative group $\mathbb{R}$ can be replaced by any uniquely 2-divisible commutative semigroup H or by a semimodule over $\mathbb{Z}_{>0}[1/2]$. Unique 2-divisibility is needed to consistently split a planar H-weighted 1-foam into a union of tripods, since lollipop loops carry weights $a/2, b/2, (a+b)/2$. These divisions by two must exist and be consistent. One then gets an isomorphism of abelian groups

$$\mathsf{Cob}^{1,\mathrm{up}}_{H} \cong Z^2(H). \tag{46}$$

The group $Z^2(H)$ can be thought of as a universal *antisymmetric* 2-cocycle on H. Antisymmetry condition forces the bracket to be almost bilinear, see Proposition 3.5.

Symmetric 2-cocycles are not almost bilinear, in this sense, and allow for a richer structure. Interestingly, they naturally appear in the theory of formal groups [32, Section 6], with relations (16) and $[a, b] = [b, a]$ interpreted as the infinitesimal version of the formal group law axioms. Formal groups are closely related to cobordism groups of manifolds (to the complex cobordism generalized cohomology theory), see [1, 21, 26, 31, 10] and the references therein.

It seems possible to interpret symmetric 2-cocycles in the framework of foam cobordisms. A step towards such interpretation is to consider unoriented cobordisms (2-foams with boundary) between 1-foams, not embedded anywhere, where 2-foams have oriented seams. One imposes the compatibility condition on seam orientations at vertices of the 2-foam to match the 2-cocycle relation. Absence of an embedding and not keeping track of the order of thin facets at a seam leads to the symmetric relation $[a,b]=[b,a]$. The antisymmetry property vanishes, since in the cobordisms in the top row of Figure 16 the seams are now oriented and the two vertices of the boundary 1-foam for each relation have opposite types, leading us to denote the two brackets by $[a,b]_+$ and $[a,b]_-$ and giving the relation $[a,b]_+ + [b,a]_- = 0$, which simply allows to express one bracket via the other. The bracket $[a,b]_+$, then, satisfies the symmetry property and the 2-cocycle condition.

Remark 3.15. For an interesting cobordism group, we considered planar *unoriented* weighted 1-foams in this section. The planar *oriented* 1-foams do not allow loops and creation of tripod foams $T(a,b)$. The cobordism group of suitably defined planar oriented 1-foams is trivial.

Constructions and results of this section demonstrate the possibility of having interesting cobordism groups of planar objects other than embedded 1-manifolds, with additional decorations, such as (positive) real weights. Notice that the resulting cobordism group has the flavour of scissor congruence groups (for instance, surjecting onto $\Lambda^2_{\mathbb{Q}}\mathbb{R}$, so that $\mathbb{R}$ is essentially viewed as a discrete group, which is typical of scissor congruence).

4 Variations on weighted foams

Here we go back to considering oriented $\mathbb{R}_{>0}$-weighted foams, not embedded anywhere, as in Section 2.

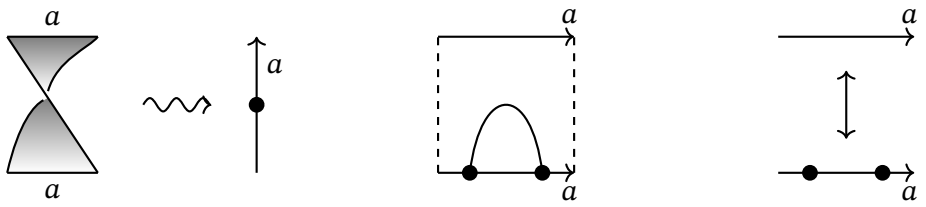


Figure 41: Left: encoding flip by a dot. Middle: a cobordism that cancels a pair of dots on a line. Right: Removing or adding two dots on a line results in a cobordant foam. Our surfaces are not embedded in $\mathbb{R}^3$, and the square of the flip is the identity.

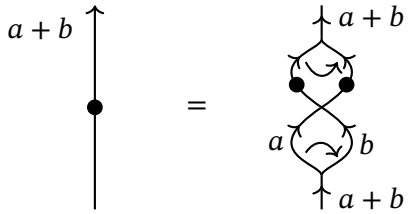


Figure 42: Splitting an $(a+b)$-line flip into a-line and b-line flips.

4.1 Foams with flips

The group of IET automorphisms of the interval can be enhanced with flips $[0,a] \longrightarrow [0,a], x \longmapsto a - x$, see [19], viewed as a subgroup of all measurable automorphisms of $([0,1],|dx|)$. Denote this automorphism group by G_{flip}; it contains $\mathsf{Aut}_{\mathsf{IET}}$ as a subgroup. Arnoux in [3, 4], also see [19], has shown that this group is simple, in particular, $[G_{\mathsf{flip}}, G_{\mathsf{flip}}] = G_{\mathsf{flip}}$.

A flip of an interval $[0,a]$ can be encoded by a dot on a line labeled a, see Figure 41. IET 1-foams and 2-foams can be enhanced by flip dots and flip networks, subject to the following relations:

- Two dots on an interval can cancel via a cobordism,
- A flip line on an $(a+b)$-facet can cross a seam and become two flip lines on a, b facets, reversing the order of the two thin facets at the seam,

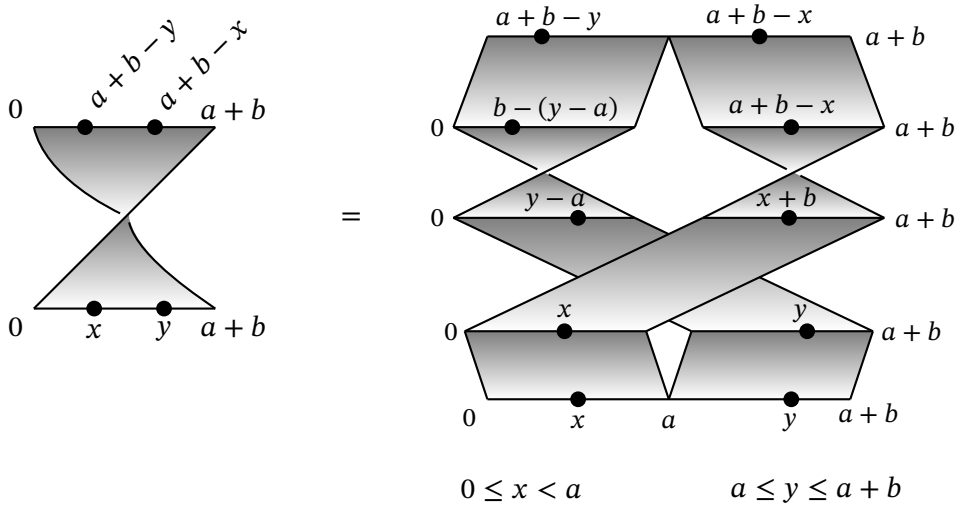


Figure 43: The flip map $x \longmapsto 1 - x$ can be split as shown on the right.

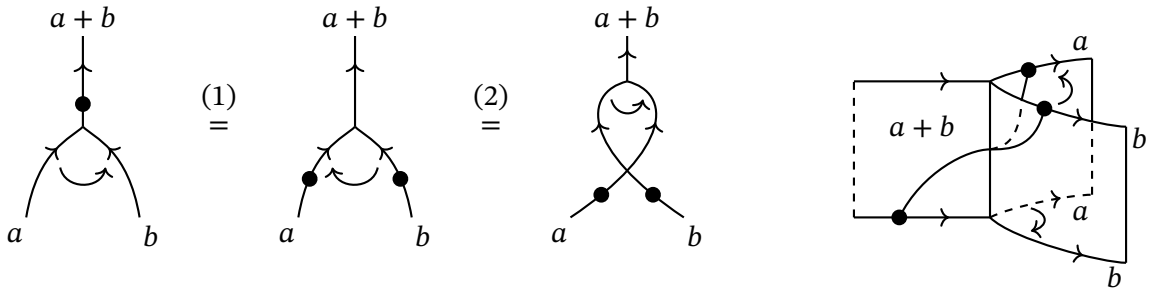


Figure 44: Equation (1) encodes moving a dot through a vertex, which includes flipping the order of thin edges. Relation (2) follows from Move 4 in Figure 15, which does not involve dots. Right: foam cobordism between the two foams in equation (1), with reversed orientation. Here, the direction is reversed during the interaction.

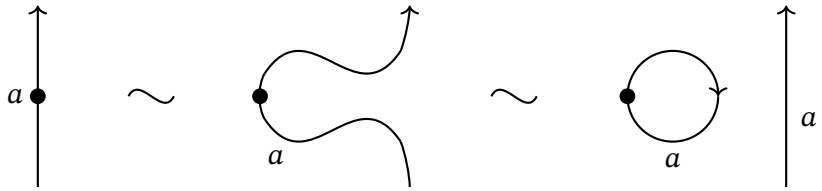


Figure 45: Splitting a dot off the rest of the foam.

as shown in Figures 41 and 44. Figure 42 shows how a flip on an $(a+b)$-line is modified to flips on a- and b-lines, via a concordance of braid-like 1-foams with flips. Figure 43 shows the thickened version of that equivalence transformation. Denote by $\mathsf{Cob}^{1,\mathsf{flip}}_{>0}$ the cobordism group of $\mathbb{R}_{>0}$-decorated 1-foams with flips.

Theorem 4.1. The cobordism group of weighted oriented 1-foams with flips is trivial,

$$\mathsf{Cob}^{1,\mathsf{flip}}_{>0} \;\cong\; 0. \tag{47}$$

Proof. Any $\mathbb{R}_{>0}$-decorated 1-foam with flips U can be represented as the closure $\widehat{U}_0$ of a braid-like foam U_0. To U_0 there is associated the corresponding element $u_0 \in G_{\mathsf{flip}}$. Since G_{flip} is perfect, u_0 can be represented as a product of commutators, $u_0 = \prod_{i=1}^{k}[v_i, w_i]$. Write U_0 as the composition of corresponding 1-foams, $U_0 = \prod_{i=1}^{k}[V_i, W_i]$. The foam for each commutator is cobordant to the interval foam, using the argument as in Figure 12. Consequently $[U] = 0$ in the cobordism group. □

Remark 4.2. Theorem 4.1 can also be proved directly, without invoking the perfectness of G_{flip}. Start with a 1-foam U, possibly with flips. A dot can be split off from the rest of U into an a-circle with a dot, see Figure 45. An a-circle with a dot is null-cobordant, see Figure 46.

Consequently, a 1-foam with flips is cobordant to the same foam without flips, and flips can be removed at any time when constructing a sequence of cobordisms. Present foam U as the closure of a braid-like foam U_0. All crossings in U_0 can be split off from a diagram as in Figure 18, along with the flips.

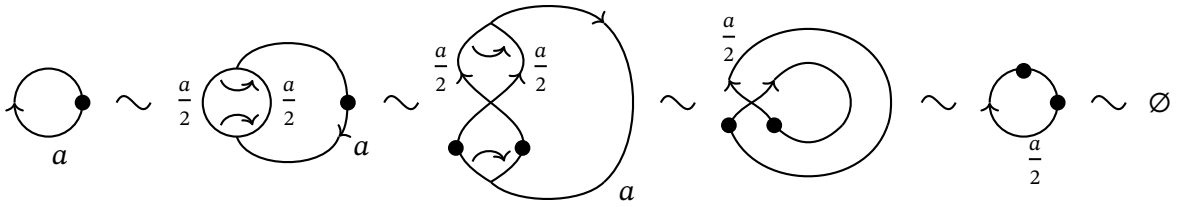


Figure 46: An a-circle with a dot is null-cobordant.

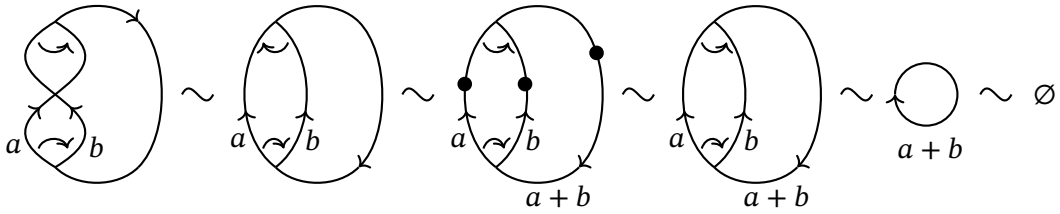


Figure 47: A cobordism from $U_{a,b}$ to the empty foam.

An order of thin edges at a vertex can be reversed by adding three dots, one on each adjacent edge, as shown in relation (1) in Figure 15, with an additional dot added on both sides of the relation on the $(a+b)$-line. Two dots on the $(a+b)$-line on the left hand side can then be cancelled via the Figure 41 cobordism. Dots can be split off as well and removed, being null-cobordant (Figure 46).

A combination of these moves transforms U_0 into a cobordant foam which is a disjoint union of foams U_{a_i,b_i} and a braid-like foam U_2 without crossings, dots, and compatible orders of thin edges at all vertices, see Figure 17 on the right. Foam U_2 is null-cobordant, as explained earlier. Foam $U_{a,b}$ is null-cobordant as well, as shown in Figure 47. Consequently, U is null-cobordant.

4.2 Foams with a map into a topological space

Consider 1-foams and 2-foams equipped with a continuous map into a topological space X. Without loss of generality we can assume that X is a connected CW-complex. One can form the abelian group $\mathrm{Cob}^{1,X}_{\mathbb{R}_{>0}}$ of $\mathbb{R}_{>0}$-decorated oriented 1-foams U with a map $\psi : U \longrightarrow X$

modulo cobordisms. Two 1-foams as above with maps $\psi_i : U_i \longrightarrow X$ are cobordant if there is a 2-foam F with a continuous map $\psi : F \longrightarrow X$ such that $\partial(F,\psi) = (U_1,\psi_1) \sqcup (-U_0,\psi_0)$. For the sake of simplicity, we can further assume that X is a path-connected CW-complex.

The homotopy classes of maps $\psi : U \longrightarrow X$ of a 1-foam into X depend only on the fundamental group $\pi_1(X)$ of X. Denote this group by G and consider G-decorated 1-foams, as follows. A decoration consists of finitely many dots on edges of U, each dot labeled by an element of G. Form the standard model of the classifying space BG, take its 2-skeleton and pass to the Poincaré dual $P(G)$ of the 2-skeleton. A map of a 1-foam to BG can be deformed to a piecewise linear (PL) map into the 1-skeleton BG^1 of BG, which we also denote $\psi : U \longrightarrow \mathsf{BG}^1$. Here we view the 1-skeleton of BG as a subspace of $P(G)$. The inverse image of the 1-skeleton of $P(G)$ is then a collection of points on the edges of U labeled by elements of G. A point labeled $g \in G$ corresponds to the intersections of $\psi(U)$ with the one-cell of $P(G)$ labeled g. Also see [9, Section 2].

A cobordism F between 1-foams U_1, U_2 which is a 2-foam with a map $\psi : F \longrightarrow X$ can be converted to a PL map into the 2-skeleton $P(G)$, also denoted ψ. The inverse image of the 1-skeleton $P(G)^1$ of $P(G)$ is then a one-dimensional PL CW-complex in F with labels on edges, with possible singularities as shown in Figure 48. The edges labeled $1 \in G$ can be erased.

Proposition 4.3. The cobordism group of oriented $\mathbb{R}_{>0}$-decorated 1-foams equipped with a continuous map to a path-connected CW-complex X is given by

$$(\mathbb{R} \otimes_{\mathbb{Z}} \mathrm{H}_1(X,\mathbb{Z})) \oplus (\mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R}). \tag{48}$$

Note that, if $\mathrm{H}_1(X,\mathbb{Z})$ is torsion, the first term vanishes.

Proof. Denote by $[g]$ the image of $g \in G$ in $\mathrm{H}_1(G) = \mathrm{H}_1(X,\mathbb{Z})$ and define a map

$$\gamma_1 \ : \ \mathsf{Cob}^{1,X}_{\mathbb{R}_{>0}} \longrightarrow \mathbb{R} \otimes_{\mathbb{Z}} \mathrm{H}_1(X,\mathbb{Z}) \tag{49}$$

by sending a G-labeled oriented 1-foam U to the sum

$$\gamma_1(U) \ := \ \sum_i a_i \otimes [g_i], \tag{50}$$

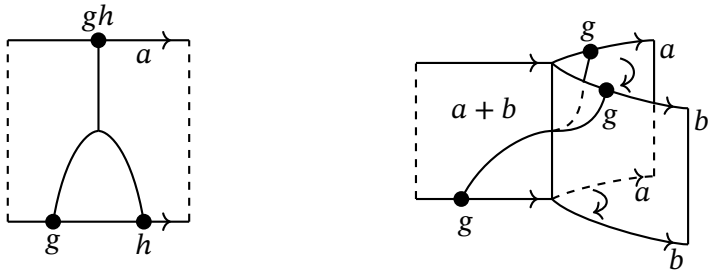


Figure 48: Possible singularities of a network of G-defects on a 2-foam. Trivalent vertices of the network on a facet (left figure) are points in $\psi^{-1}(P(G)^0)$, where $P(G)^0$ is the 0-skeleton of the Poincaré dual $P(G)$. Unlike that in Figure 44 on the right, the order of thin facets does not reverse when a defect line crosses a seam (right figure).

where the sum is over all labels $g_i \in G$ on U and a_i is the thickness of the edge which contains g_i. It is clear that γ_1 is invariant under cobordisms between 1-foams shown in Figure 48. Topology changes of 1-foams under cobordisms, shown in Figure 16, happen away from the one-dimensional network of G-defects on a cobordism, since the latter network can be deformed away from vertices and from local maxima and minima of seams and facets. Consequently, $\gamma_1(U)$ depends only on the cobordism class of U in $\mathsf{Cob}^{1,X}_{\mathbb{R}_{>0}}$, so it is a well-defined map on cobordism classes $[U]$. Define the homomorphism

$$\gamma \;:\; \mathsf{Cob}^{1,X}_{\mathbb{R}_{>0}} \longrightarrow (\mathbb{R} \otimes_{\mathbb{Z}} \mathrm{H}_1(X,\mathbb{Z})) \oplus (\mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R}), \;\; \gamma([U]) = (\gamma_1([U]), \mathsf{SAF}(U)). \tag{51}$$

It is then straightforward to check that γ is an isomorphism of groups. Indeed, the defect points on a 1-foam U can be split away, each on its own circle, from the rest of U, converting U into the disjoint union of a foam U' without G-defects and a union U'' of circles $C_1, \ldots, C_n$ carrying labels $g_1, \ldots, g_n$ and having thicknesses $a_1, \ldots, a_n$. The cobordism class of U' is an element of $\mathbb{R} \wedge_{\mathbb{Q}} \mathbb{R}$, and that of U'' reduces to an element of $\mathbb{R} \otimes_{\mathbb{Z}} \mathrm{H}_1(X,\mathbb{Z})$ as outlined above. The homomorphism γ in (51) has a section s taking a generator $a \otimes g$ in the first term on the RHS to a circle of thickness a carrying label g, and a generator $a \wedge b$ of the second term on the RHS to the web $U_{a,b}$ shown in Figure 14, see the discussion in that section. We have $\gamma \circ s = 1$. Our classification of the cobordism group without

G-decorations in Theorem 2.6 and the above arguments, including splitting off G-points away from the rest of the 1-foam, imply that $s \circ \gamma = 1$, showing that γ is an isomorphism and concluding the proof. □

4.3 Other variations

Remark 4.4. D. Sullivan in [33, 35] has proved that any oriented one-dimensional solenoidal manifold M is the boundary of an oriented solenoidal surface. The idea (tributing his much earlier conversation with B. Edwards) is to represent M as the closure of a braid-like structure, that is, as the mapping torus of a homeomorphism of the Cantor set. Sullivan then uses R. D. Anderson's theorem [2] that the homeomorphism group of the Cantor set is simple and, in particular, perfect. Representing the homeomorphism as a product of commutators allows to realize M as the boundary, schematically shown in Figure 12. This use of braid-like closures is analogous to that in the proofs of Theorems 2.6, 4.1, where an oriented weighted 1-foam is represented as the closure of a braid 1-foam. It is likely that Sullivan's result can be interpreted as the vanishing of K_1 of a suitable assembler category [37, 38], where the assembler structure is that of coverings of the Cantor set.

Remark 4.5. The SAF invariant can be generalized to the Kenyon–Smillie invariant [15, 6], and it is an interesting question to interpret the latter via suitably decorated foam cobordisms.

Remark 4.6. O. Lacourte [19] defines a version of interval exchange transformations for each subgroup $\Gamma \subset \mathbb{R}/\mathbb{Z}$, via the corresponding subgroup $\mathsf{IET}(\Gamma)$ of the group $\mathsf{Aut}_{\mathsf{IET}}$. He establishes an isomorphism between $H_1(\mathsf{IET}(\Gamma))$ and the second skew-symmetric power of $\widetilde{\Gamma}$ over $\mathbb{Z}$, where $\widetilde{\Gamma}$ is the preimage of Γ in $\mathbb{R}$. It is straightforward to extend the results of Section 2 to interpret the above first homology group as the group of foam cobordisms, with the facets of foams carrying weights in $\widetilde{\Gamma} \cap \mathbb{R}_{>0}$ (and see Remark 2.9).

Lacorte also considers the group of IETs with flips. This group is known to be perfect, and Theorem 4.1 is a foam interpretation of this result. Lacourte shows that subgroups

$\mathrm{IET}(\Gamma)$ with flips modulo the commutator may have 2-torsion, and there should be an analogue of this result for foam cobordism.

Acknowledgments

The authors are grateful to David Gepner, Nitu Kitchloo, and Inna Zakharevich for interesting discussions. The authors would like to thank the referees for detailed feedback on our paper. M.K. would like to acknowledge partial support from NSF grant DMS-2204033 and Simons Collaboration Award 994328.

References

[1] J. F. Adams. *Stable homotopy and generalised homology*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, Ill.-London, 1974. 46

[2] Richard D. Anderson. The algebraic simplicity of certain groups of homeomorphisms. *Amer. J. Math.*, 80:955–963, 1958. 54

[3] Pierre Arnoux. Échanges d'intervalles et flots sur les surfaces. In *Ergodic theory (Sem., Les Plans-sur-Bex, 1980) (French)*, volume 29 of *Monogr. Enseign. Math.*, pages 5–38. Univ. Genève, Geneva, 1981. 48

[4] Pierre Arnoux. *Un invariant pour les echanges d'intervalles et les flots sur les surfaces*. HAL theses, 1981. Thesis (Ph.D.)–Université de Reims Champagne-Ardenne. 48

[5] Skew of 2-cocycle for trivial group action of abelian group is alternating bihomomorphism. https://groupprops.subwiki.org/wiki/Skew_of_2-cocycle_for_trivial_group_action_of_abelian_group_is_alternating_bihomomorphism, 2012. 32

[6] Kariane Calta. Veech surfaces and complete periodicity in genus two. *J. Amer. Math. Soc.*, 17(4):871–908, 2004. 54

[7] Carmen Caprau, Gabriel Coloma, and Marguerite Davis. The L-move and Markov theorems for trivalent braids. *Involve*, 13(1):21–50, 2020. 21

[8] Hieu Trung Do and Thomas A. Schmidt. New infinite families of pseudo-Anosov maps with vanishing Sah-Arnoux-Fathi invariant. *J. Mod. Dyn.*, 10:541–561, 2016. 3

[9] David Gepner, Mee Seong Im, Mikhail Khovanov, and Nitu Kitchloo. Foams with flat connections and algebraic K-theory. *arXiv preprint arXiv:2405.14465*, pages 1–57, 2024. 2, 21, 52

[10] Yonatan Harpaz. Complex Cobordism and Formal Group Laws. *Online lecture notes, https://www.math.univ-paris13.fr/ harpaz/complex_cobordism.pdf*, pages 1–19, 2012. 46

[11] Mee Seong Im and Mikhail Khovanov. Entropy, cocycles, and their diagrammatics. *arXiv preprint arXiv:2409.08462*, pages 1–82, 2024. 29

[12] Atsushi Ishii. The Markov theorems for spatial graphs and handlebody-knots with Y-orientations. *Internat. J. Math.*, 26(14):1550116, 23, 2015. 21

[13] Ken Kanno and Kouki Taniyama. Braid presentation of spatial graphs. *Tokyo J. Math.*, 33(2):509–522, 2010. 21

[14] Louis H. Kauffman. Invariants of graphs in three-space. *Trans. Amer. Math. Soc.*, 311(2):697–710, 1989. 16

[15] Richard Kenyon and John Smillie. Billiards on rational-angled triangles. *Comment. Math. Helv.*, 75(1):65–108, 2000. 54

[16] Mikhail Khovanov. $\mathfrak{sl}(3)$ link homology. *Algebr. Geom. Topol.*, 4:1045–1081, 2004. 2

[17] Mikhail Khovanov and Nitu Kitchloo. A deformation of Robert-Wagner foam evaluation and link homology. In *Algebraic and topological aspects of representation theory*, volume 791 of *Contemp. Math.*, pages 147–204. Amer. Math. Soc., Providence, RI, 2024. 2

[18] Peter B. Kronheimer and Tomasz S. Mrowka. Tait colorings, and an instanton homology for webs and foams. *J. Eur. Math. Soc. (JEMS)*, 21(1):55–119, 2019. 2

[19] Octave Lacourte. Abelianization of some groups of interval exchanges. *Ann. Inst. Fourier (Grenoble)*, 72(1):59–108, 2022. 48, 54

[20] Marco Mackaay and Pedro Vaz. The universal $\mathfrak{sl}_3$-link homology. *Algebr. Geom. Topol.*, 7:1135–1169, 2007. 2

[21] J. Peter May. *A concise course in algebraic topology*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1999. 46

[22] James R. Munkres. *Topology*. Prentice Hall, Inc., Upper Saddle River, NJ, second edition, 2000. 22

[23] Ulrich Oertel. Incompressible branched surfaces. *Invent. Math.*, 76(3):385–410, 1984. 9

[24] Ulrich Oertel. Measured laminations in 3-manifolds. *Trans. Amer. Math. Soc.*, 305(2):531–573, 1988. 6, 9

[25] Robert C. Penner and John L. Harer. *Combinatorics of train tracks*, volume 125 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1992. 24, 25

[26] Eric Peterson. *Formal geometry and bordism operations*, volume 177 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2019. With a foreword by Matthew Ando. 46

[27] Hoel Queffelec and Kevin Walker. Movie moves for framed foams from multijet transversality. *arXiv preprint arXiv:2205.14947*, pages 1–88, 2022. 6, 7

[28] Louis-Hadrien Robert and Emmanuel Wagner. A closed formula for the evaluation of foams. *Quantum Topol.*, 11(3):411–487, 2020. 2

[29] Louis-Hadrien Robert and Emmanuel Wagner. Symmetric Khovanov-Rozansky link homologies. *J. Éc. polytech. Math.*, 7:573–651, 2020. 7

[30] David E. V. Rose and Paul Wedrich. Deformations of colored $\mathfrak{sl}_N$ link homologies via foams. *Geom. Topol.*, 20(6):3431–3517, 2016. 2

[31] Markus Spitzweck. Algebraic cobordism in mixed characteristic. *Homology Homotopy Appl.*, 22(2):91–103, 2020. 46

[32] Neil P. Strickland. Formal groups. https://strickland1.org/courses/formalgroups/fg.pdf, pages 1–63, 2019. 46

[33] Dennis Sullivan. Solenoidal manifolds. *J. Singul.*, 9:203–205, 2014. 54

[34] William A. Veech. The metric theory of interval exchange transformations. III. The Sah-Arnoux-Fathi invariant. *Amer. J. Math.*, 106(6):1389–1422, 1984. 3

[35] Alberto Verjovsky. Commentaries on the paper “Solenoidal Manifolds" by Dennis Sullivan. *J. Singul.*, 9:245–251, 2014. 54

[36] Yaroslav Vorobets. Notes on the commutator group of the group of interval exchange transformations. *arXiv preprint arXiv:1109.1352*, pages 1–16, 2011. 3, 11

[37] Inna Zakharevich. The K-theory of assemblers. *Adv. Math.*, 304:1176–1218, 2017. 3, 21, 54

[38] Inna Zakharevich. On K_1 of an assembler. *J. Pure Appl. Algebra*, 221(7):1867–1898, 2017. 3, 11, 14, 21, 54

[39] Anton Zorich. Flat surfaces. In *Frontiers in number theory, physics, and geometry. I*, pages 437–583. Springer, Berlin, 2006. 25

[40] Anton Zorich. Geodesics on flat surfaces. In *International Congress of Mathematicians. Vol. III*, pages 121–146. Eur. Math. Soc., Zürich, 2006. 25

Foam cobordism and the Sah-Arnoux-Fathi invariant

AUTHORS

Mee Seong Im
Department of Mathematics
Johns Hopkins University
Baltimore, MD 21218, USA
and
Department of Mathematics
United States Naval Academy
Annapolis, MD 21402, USA
email: meeseong@jhu.edu

Mikhail Khovanov
Department of Mathematics
Johns Hopkins University
Baltimore, MD 21218, USA
and
Department of Mathematics
Columbia University
New York, NY 10027, USA
email: khovanov@jhu.edu

Arnold Mathematical Journal
Volume 12, Issue 1, 2026

Symmetric cubic polynomials

Alexander Blokh Lex Oversteegen Nikita Selinger
Vladlen Timorin Sandeep Chowdary Vejandla

Received 20 Oct 2024; Accepted 22 Jun 2025

Abstract:

We describe a model $\mathcal{M}_3^{comb}$ for the boundary of the connectedness locus $\mathcal{M}_3^{sy}$ of the parameter space of cubic symmetric polynomials $p_c(z) = z^3 - 3c^2z$. We show that there exists a monotone continuous function $\pi : \partial\mathcal{M}_3^{sy} \to \mathcal{M}_3^{comb}$ which is a homeomorphism if $\mathcal{M}_3^{sy}$ is locally connected.

AMS Classification: 37F20;37F10, 37F50

Key words and phrases: Complex dynamics; laminations; Mandelbrot set; Julia set

1 Introduction

A central problem of Complex Dynamics is to describe parameter spaces of holomorphic maps. Investigating the deceptively simple *quadratic family* $f_c(z) = z^2 + c$ led to an explosion of activity in the field. Aided by computer graphics capabilities, mathematicians

DOI: 10.56994/ARMJ.012.001.002

made many interesting discoveries concerning the connectedness locus of the quadratic family, the celebrated *Mandelbrot set* $\mathcal{M}_2$.

One of the first such discoveries, made by Douady and Hubbard [15], was that $\mathcal{M}_2$ is connected. Then the combinatorial description of the structure of the Mandelbrot set was largely carried out in the language of *laminations* introduced by Thurston [42] (see Section 3 and [42] or [6] for precise definitions and other details). Douady constructed a topological *pinched disk* model of $\mathcal{M}_2$; Thurston made this model more explicit and described it in terms of laminations. If $\mathcal{M}_2$ is locally connected (which is still an open question), then it is homeomorphic to its model. The local connectivity of the Mandelbrot set is one of the most important long standing conjectures in the field; if true, it will imply the density of hyperbolicity property of the quadratic family.

The above describes the quadratic version of what one can call the *Douady–Hubbard–Thurston program*, i.e. a two step approach to studying some complex one-dimensional parameter space of polynomials that we now fix. Similar to the quadratic family, its most interesting part is the *connectedness locus*, i.e., the locus of all polynomials in the space with connected Julia sets (in the quadratic case, this is the Mandelbrot set). One needs to prove that the connectedness locus is connected itself. Then two steps are made. On the first step one describes the laminations of the polynomials from the parameter space in question producing in the end the corresponding space of laminations, most likely described itself by a certain *parameter lamination* (like, e.g., Thurston's QML lamination [42]). On the second step one constructs a monotone map from the connectedness locus of interest to the quotient space of the closed unit disk under the parameter lamination. This quotient space can be viewed as a model for the connectedness locus.

The Douady–Hubbard–Thurston program has been implemented for quadratic polynomials, and then for *unicritical polynomials* $z^d + c$ of any degree d, cf. [1, 22, 39, 40], where c is the parameter.

We initiated this program for the space $\mathcal{SCP}$ of *symmetric cubic polynomials* $p_c(z) = z^3 - 3c^2z$ with *marked critical point* c in papers [8, 9] that serve as a prequel to the present

article. Namely, in [8] we investigated the space of *symmetric cubic laminations* and constructed the associated parameter lamination called the *cubic symmetric comajor lamination* C_sCL (see Section 3). In [9] we proved an analog of Lavaurs algorithm for this lamination. Now, let $\mathcal{M}_3^{sy}$ be the *connectedness locus* of $\mathcal{SCP}$. The aim of the present article is to complete the program for the space $\mathcal{SCP}$ and prove the following theorem.

Main Theorem. *The set $\mathcal{M}_3^{sy}$ is a full continuum. There exists a monotone continuous surjective map $\pi : \mathcal{M}_3^{sy} \to \overline{\mathbb{D}}/C_sCL$. If $\mathcal{M}_3^{sy}$ is locally connected, then π is a homeomorphism.*

We are not aware of any other articles in which the Douady–Hubbard–Thurston program is fully implemented for non-unicritical polynomials. We did recently learn of a manuscript by Xavier Buff [12] in which he studies the parameter space of symmetric polynomials of the form $P_{\lambda,d}(z) = \lambda z + z^d$ and shows that the connectedness locus M_d contains the unit disk $\mathbb{D}$ and that every component of $M_d \setminus \overline{\mathbb{D}}$ is homeomorphic to a limb of the Mandelbrot set.

Acknowledgments.

Some figures in the article have been produced with a modified version of *Mandel*, a software written by Wolf Jung. The authors are grateful to the reviewer for careful reading and useful suggestions.

2 Notation and preliminaries

We assume knowledge of basic facts and concepts of complex dynamics. We also use standard notation (such as $J(P)$ for the Julia set of a polynomial P, etc).

Consider the space $\mathcal{SCP}$ of symmetric cubic polynomials $p_c(z) = z^3 - 3c^2z$ with marked critical point c, or, more formally, the space of pairs (p_c, c), which, in turn, is uniquely parameterized by the values of c. Since $p_c = p_{-c}$, every polynomial (except for $p_0(z) = z^3$) shows in $\mathcal{SCP}$ twice. Thus, $\mathcal{SCP}$ is a (branched) two-to-one cover of the moduli space of all

odd cubic polynomials, where the moduli space means the quotient space with respect to complex linear conjugacy. The critical points of p_c are c and $-c$, the corresponding critical values are $-2c^3$ and $2c^3$. The *marked cocritical point* of p_c, i.e., the other preimage of $-2c^3 = p_c(c)$, is $-2c$. Subsets A and B of the plane are said to be *mutually symmetric* if $B = -A$. If $A = -A$ we call a set A *symmetric*. Since p_c is odd, the Julia set $J(p_c)$, the filled Julia set $K(p_c)$, and their complements are symmetric. Observe that $p_c(0) = 0$ and $p_c'(z) = p_c'(-z) = 3z^2 - 3c^2$.

Let $\mathcal{M}_3^{sy}$ be the *connectedness locus* of $\mathcal{SCP}$, i.e., the set of all c for which the Julia set of p_c is connected. It is known that the Julia set of p_c is connected if and only if all forward orbits of critical points of p_c are bounded. Since p_c has mutually symmetric critical orbits, we conclude that $c \in \mathcal{M}_3^{sy}$ if and only if the orbit of c or, equivalently of $-2c$ or $2c$, is bounded.

For any $r > 0$ set $\mathbb{D}_r = \{z \in \mathbb{C} \mid |z| < r\}$, and write $\mathbb{D}$ for $\mathbb{D}_1$. Let $\mathbb{S}^1$ be the unit circle. For a set $A \subset \mathbb{C}$, let $\overline{A}$ be its closure and ∂A be its boundary. We use the terms *periodic orbit* and *cycle* interchangeably. External rays to the Julia set of a polynomial P are denoted $R_P(\theta)$ where θ is the argument of the ray (if there is no ambiguity we may omit the polynomial from our notation; also, we write $R_c(\theta)$ instead of $R_{p_c}(\theta)$).

Let X, Y be topological spaces and $f : X \to Y$ be continuous. Then f is said to be *monotone* if $f^{-1}(y)$ is connected for each $y \in Y$. It is known that if f is monotone and X is a continuum then $f^{-1}(Z)$ is connected for every connected $Z \subset f(X)$.

3 Symmetric cubic laminations

Invariant laminations were introduced in [42]; they play a major role in polynomial dynamics. The preceding papers [8, 9] of this series contain an overview, based on [42] and [6]. Here we follow [8] and [9] (see Section 2 of [8] for a detailed discussion).

If a monic polynomial P has a locally connected Julia set $J(P)$, then $P|_{J(P)}$ is topologically conjugate to a suitable quotient of the d-tupling map $\sigma_d : \mathbb{S}^1 \to \mathbb{S}^1$ (where $\sigma_d(z) = z^d$ if

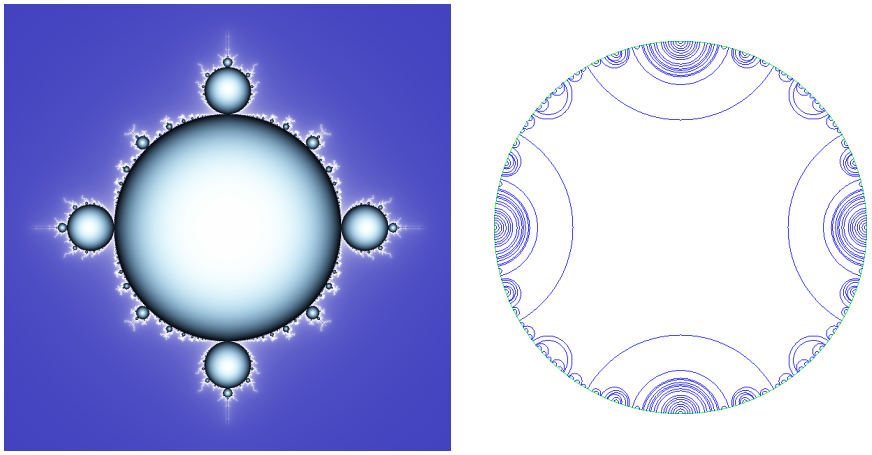


Figure 1: The parameter space of symmetric cubic polynomials $\mathcal{SCP}$ on the left and the Symmetric Cubic Comajor Lamination C_sCL on the right.

$\mathbb{S}^1$ is viewed as the unit circle in $\mathbb{C}$ and as $\sigma_d(\theta) = d\theta$ if $\mathbb{S}^1$ is identified with $\mathbb{R}/\mathbb{Z}$). The quotient is with respect to an equivalence relation $\sim_P$; the *leaves* of the corresponding lamination $\mathcal{L}_P$ are by definition the edges of the convex hulls of all $\sim_P$-classes.

A chord of $\overline{\mathbb{D}}$ with endpoints a, b is denoted by $\overline{ab}$; it is *critical* if $\sigma_d(a) = \sigma_d(b)$ while $a \neq b$. A lamination is normally denoted by $\mathcal{L}$ while the union of all its leaves and $\mathbb{S}^1$ by $\mathcal{L}^*$. For $G \subset \overline{\mathbb{D}}$, denote $G \cap \mathbb{S}^1$ by $\mathcal{V}(G)$ and call the elements of $\mathcal{V}(G)$ *vertices* of G. Call G a *gap* of a lamination $\mathcal{L}$ if it is the closure of a component of $\mathbb{D} \setminus \mathcal{L}^*$. A gap G is said to be *critical* if its image is not a gap or the degree of $\sigma_d|_{\partial G}$ is greater than 1. A *critical set* is a critical leaf or a critical gap. If G is a leaf or a gap of $\mathcal{L}$, then G coincides with the convex hull of $\mathcal{V}(G)$. A gap G is called *infinite (finite)* if and only if $\mathcal{V}(G)$ is infinite (finite).

Let $\mathcal{L}$ be a lamination. The equivalence relation $\sim_{\mathcal{L}}$ induced by $\mathcal{L}$ is defined by declaring that $x \sim_{\mathcal{L}} y$ if and only if there exists a finite concatenation of leaves of $\mathcal{L}$ joining x to y. A lamination $\mathcal{L}$ is called a *q-lamination* if the convex hulls of $\sim_{\mathcal{L}}$-classes are precisely finite gaps or leaves of $\mathcal{L}$. Two distinct chords are called (σ_d-)*siblings* if they have the same σ_d-image.

3.1 Symmetric laminations

Let τ be the rotation of $\overline{\mathbb{D}}$ (or of $\mathbb{S}^1$) by 180° around its center $\mathcal{O}$. Also, given a map $f : X \to X$ we call $x \in X$ *preperiodic of preperiod* $k > 0$ or *k-preperiodic* if $f^k(x)$ is f-periodic while $f^{k-1}(x)$ is not periodic.

Definition 3.1 (Symmetric laminations)**.** A σ_3-invariant lamination $\mathcal{L}$ is called a *symmetric (cubic) lamination* if $\ell \in \mathcal{L}$ implies $\tau(\ell) \in \mathcal{L}$.

Definition 3.2 (length and majors)**.** Given a non-diameter chord ℓ in $\overline{\mathbb{D}}$, define the arc $h(\ell)$ as the shortest arc of $\mathbb{S}^1$ joining the endpoints of ℓ. If ℓ' and ℓ'' are chords such that $h(\ell') \subset h(\ell'')$, then we say that ℓ' is *under* ℓ''. Define the *length* $|\ell|$ of ℓ as the length of $h(\ell)$ divided by 2π in the case when ℓ is not a diameter; if ℓ is a diameter, set $|\ell| = \frac{1}{2}$. Given a symmetric lamination $\mathcal{L}$, call a leaf M a *major of* $\mathcal{L}$ if there are no leaves of $\mathcal{L}$ closer in length to $\frac{1}{3}$ than M.

Let $\Gamma : [0, \frac{1}{2}] \to [0, \frac{1}{2}]$ be a piecewise linear function with slope ± 3 defined as $\Gamma(x) = 3x$ if $0 \le x \le \frac{1}{6}$ and as $\Gamma(x) = |3x - 1|$ if $\frac{1}{6} \le x \le \frac{1}{2}$. Then $|\sigma_3(\ell)| = \Gamma(|\ell|)$. Simple analysis of the dynamics of Γ shows that for any leaf ℓ an eventual image of ℓ has length 0, or length $\frac{1}{2}$, or length which is between $\frac{1}{4}$ and $\frac{5}{12}$.

Suppose that $\frac{1}{4} \le |\ell| \le \frac{5}{12}$. Then there exists a sibling chord ℓ' of ℓ such that the strip S of $\overline{\mathbb{D}}$ between ℓ and ℓ' has two circle arcs on its boundary, each at most $\frac{1}{6}$ long. We also consider chords $\tau(\ell)$ and $\tau(\ell')$ as well as the $\overline{\mathbb{D}}$-strip $\tau(S)$ between them. The union $S \cup \tau(S)$, denoted $\mathrm{SH}(\ell)$, is called the *short strips set* of ℓ.

If a major M of a symmetric lamination $\mathcal{L}$ is critical, there is a unique point $x \in \mathbb{S}^1$ that is not an endpoint of M with the same σ_3-image as M. This point x is called a *comajor (of $\mathcal{L}$)*. If M is not critical, then a leaf M' (similar to ℓ' from the above) and leaves $\tau(M)$ and $\tau(M')$ are also majors of $\mathcal{L}$. We set $\mathrm{SH}(\mathcal{L}) = \mathrm{SH}(M)$ in this case and call this set the *short strips set* of $\mathcal{L}$. Let us stress again that $\mathrm{SH}(\mathcal{L})$ is formed by *two* strips and that $\mathrm{SH}(M) = \mathrm{SH}(\tau(M))$. The third sibling $\overline{c}$ of M that is disjoint from $M \cup M'$, is of length at most $\frac{1}{6}$. It is called a *comajor (of $\mathcal{L}$)*. Similarly we define a *cocritical set* $\overline{\mathrm{co}}(U)$ of a critical set U as the gap, or

the leaf, or the point disjoint from U but with the same image as U.

Because of the symmetry, comajors, majors, etc., come in pairs. A pair of comajors $\overline{c}, \tau(\overline{c})$ of a symmetric lamination $\mathcal{L}$ is called a symmetric *comajor pair*. It is *degenerate* if its elements are points and *non-degenerate* otherwise. For a symmetric lamination $\mathcal{L}$ we often assume that one of its majors is marked; we denote this major by $M_{\mathcal{L}}$ and the corresponding comajor by $\overline{\text{co}}_{\mathcal{L}}$. Observe that if $\mathcal{L}_1$ and $\mathcal{L}_2$ are symmetric laminations such that $\text{SH}(\mathcal{L}_1) \subset \text{SH}(\mathcal{L}_2)$ (e.g., if $\mathcal{L}_2 \subset \mathcal{L}_1$) then the comajors of $\mathcal{L}_1$ are located under the comajors of $\mathcal{L}_2$.

Lemma 3.3 ([8]). *Let ℓ be a leaf of a symmetric lamination $\mathcal{L}$ with $|\ell| \geq \frac{1}{4}$. If $n > 0$ is the least such that $\sigma_3^n(\ell) \subset \text{SH}(\ell)$, then the leaf $\sigma_3^n(\ell)$ non-strictly separates (in $\overline{\mathbb{D}}$) either ℓ from ℓ', or $\tau(\ell)$ from $\tau(\ell')$. Thus, either $\sigma_3^n(\ell)$ equals one of the leaves $\ell, \ell', \tau(\ell), \tau(\ell')$, or it is closer to $\frac{1}{3}$ in length than ℓ. In particular, forward images of majors/comajors of $\mathcal{L}$ never enter the open circle arcs on the boundary of the set* $\text{SH}(\mathcal{L})$.

Lemma 3.3 motivates the next definition.

Definition 3.4 (Legal pairs). If a symmetric pair $\{\overline{c}, \tau(\overline{c})\}$ is either degenerate or satisfies the following conditions:

(a) no two iterated forward images of $\overline{c}, \tau(\overline{c})$ cross, and

(b) no forward image of $\overline{c}$ crosses the interior of $\text{SH}(M_{\overline{c}})$,

then $\{\overline{c}, \tau(\overline{c})\}$ is said to be a *legal pair*.

Lemma 3.5 ([8]). *A legal pair $\{c, \tau(c)\}$ is the comajor pair of the symmetric lamination $\mathcal{L}(c)$. A symmetric pair $\{c, \tau(c)\}$ is a comajor pair if and only if it is legal.*

A symmetric lamination with an infinite gap such that the map σ_3 on it is of degree greater than 1 is called a *Fatou lamination*.

Lemma 3.6 ([8]). *A symmetric lamination is Fatou if and only if it has a preperiodic comajor of preperiod 1.*

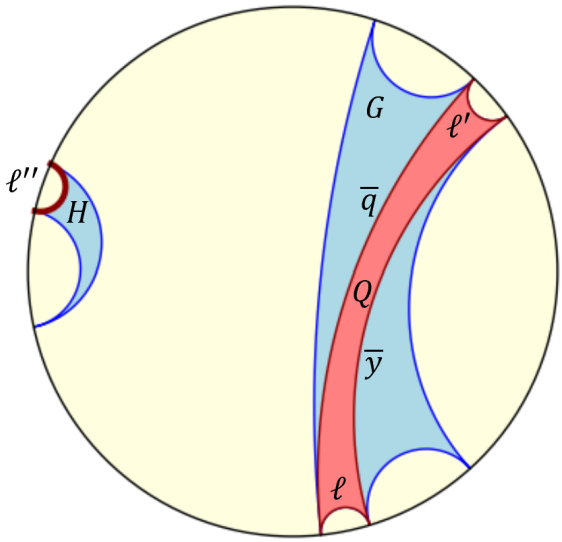


Figure 2: Chords $\overline{q}$ and $\overline{y}$ inserted into a gap G are majors of $\mathcal{L}'$ corresponding to the comajor ℓ''. Note: this figure is intentionally unrealistic; it is not a part of any symmetric cubic lamination. More realistic figures will have Q too narrow.

Suppose that $\mathcal{L}$ is a symmetric q-lamination with two finite critical gaps each of which is preperiodic of preperiod at least two. Then $\mathcal{L}$ is called a *symmetric Misiurewicz lamination*. A symmetric Misiurewicz lamination has a well defined pair of comajors. Suppose that the critical $\sim_{\mathcal{L}}$-classes are gaps G and $\tau(G)$ with at least 6 vertices each. Then there are two cocritical gaps $H \neq G$ and $\tau(H) \neq \tau(G)$ of $\mathcal{L}$ such that $\sigma_3(H) = \sigma_3(G)$ and $\sigma_3(\tau(H)) = \sigma_3(\tau(G))$. One edge of H and one edge of $\tau(H)$ are the comajors of $\mathcal{L}$. Two majors of $\mathcal{L}$ are edges of G that are siblings of the comajor edge of H; two other majors of $\mathcal{L}$ are edges of $\tau(G)$ that are siblings of the comajor edge of $\tau(H)$. While other edges of G and $\tau(G)$ are not siblings of the comajors, they can generate majors of other laminations that are finite tunings of $\mathcal{L}$.

Indeed, suppose that ℓ and ℓ' are two sibling edges of G that are not majors. The convex hull of $\ell \cup \ell'$ is a 4-gon Q with two extra edges $\overline{y}$ and $\overline{q}$ not equal to ℓ or ℓ', see Fig. 2. Construct a new lamination $\mathcal{L}'$ (not a q-lamination) by inserting $\overline{y}$ and $\overline{q}$ in G, pulling them back along the backward orbit of G and then doing the same with $\tau(G)$ and its

backward orbit. The majors of $\mathcal{L}'$ are $\overline{y}$ and $\overline{q}$ and their τ-images. If ℓ'' is a leaf of $\mathcal{L}$ which is not an edge of G and is such that $\sigma_3(\ell'') = \sigma_3(\ell)$ then ℓ'' and $\tau(\ell'')$ are the two comajors of $\mathcal{L}'$. Repeating this construction for all pairs of sibling edges of G but the majors, we see that every edge of the cocritical gap H or $\tau(H)$ is a comajor of a certain symmetric lamination which is a tuning of the original symmetric Misiurewicz lamination $\mathcal{L}$. Call cocritical sets of Misiurewicz laminations *Misiurewicz* cocritical sets. By [8, Theorem 3.9], symmetric laminations have no wandering gaps. Therefore, the above is a full description of finite gaps formed by comajors. The cocritical gaps H and $\tau(H)$ described above will be called *Misiurewicz cocritical gaps*; similarly, if a symmetric Misiurewicz q-lamination has critical 4-gons (not 6-gons or higher as was assumed above) we call its comajors *Misiurewicz cocritical leaves*.

Theorem 3.7 ([8, 9])**.** *The set of non-degenerate comajors of symmetric laminations is a q-lamination $\mathcal{L}$ invariant under τ that induces an equivalence relation $\sim_{\mathcal{L}}$ on $\mathbb{S}^1$. For any non-degenerate comajor $\overline{c}$ (i.e., a leaf of $\mathcal{L}$) one of the following holds.*

1. *It is a two-sided limit leaf in $\mathcal{L}$ which is not eventually periodic.*
2. *It is a preperiodic leaf of $\mathcal{L}$ with preperiod at least 2 which is either a two-sided limit leaf of $\mathcal{L}$ (in which case $\overline{c}$ is a Misiurewicz cocritical leaf), or an edge of a finite gap H of $\mathcal{L}$ whose edges are limits of leaves in $\mathcal{L}$ disjoint from H (in which case H is a Misiurewicz cocritical gap).*
3. *It is a 1-preperiodic comajor of a Fatou lamination and is disjoint from all other leaves of $\mathcal{L}$; all such comajors $\overline{c}$ are dense in $\mathcal{L}$ and all 1-preperiodic angles are endpoint of such comajors.*

Since comajors are leaves of q-laminations, their endpoints are either both not preperiodic, or both preperiodic with the same preperiod and the same period, or both periodic with the same period. All classes of $\sim_{\mathcal{L}}$ from Theorem 3.7 are finite. By Theorem 3.7, periodic points of $\mathbb{S}^1$ are degenerate comajors.

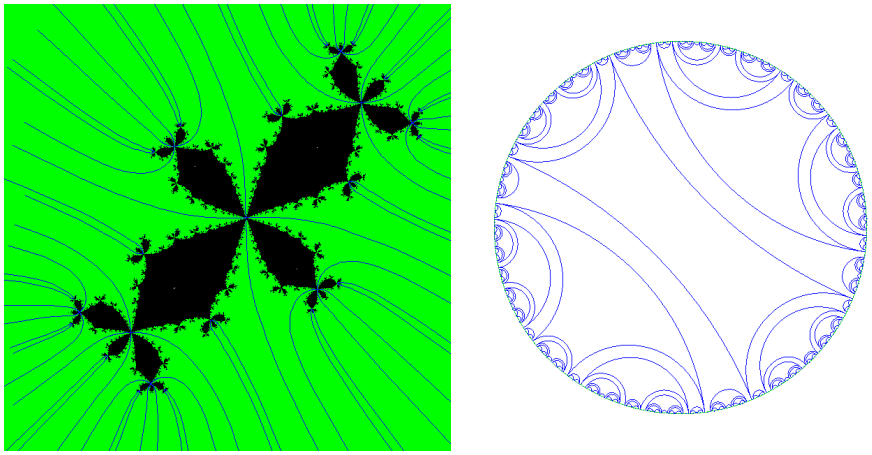


Figure 3: The symmetric cubic lamination with type B comajor $(\frac{5}{48}, \frac{7}{48})$ on the right and the Julia set of a corresponding polynomial with external rays on the left.

Definition 3.8. The q-lamination from Theorem 3.7 is called the *Cubic Symmetric Comajor Lamination* and is denoted by C_sCL. It induces an equivalence relation denoted $\sim_{sy}$. The $\sim_{sy}$-classes corresponding to symmetric Misiurewicz laminations are called *Misiurewicz $\sim_{sy}$-classes*. Denote by $\mathcal{M}^{sy}_{3,comb}$ the quotient space $\overline{\mathbb{D}}/\sim_{sy}$. Let $\mathfrak{pr}_{comb} : \overline{\mathbb{D}} \to \mathcal{M}^{sy}_{3,comb}$ be the corresponding quotient projection.

Theorem 3.7 verifies the *density of hyperbolicity* conjecture for C_sCL.

Lemma 3.9. *Let $\mathcal{L}$ be a symmetric non-empty Fatou lamination. Then one of the following holds.*

(B) There is only one cycle $\mathcal{A}$ of Fatou gaps of $\mathcal{L}$. It has even period $2m$, and is τ-symmetric. The periodic majors M and $\sigma_3^m(M) = \tau(M)$ of $\mathcal{L}$ are edges of critical gaps $U \in \mathcal{A}$ and $V = \sigma_3^m(U) = \tau(U)$. Non-periodic majors of $\mathcal{L}$ are siblings of M and $\tau(M)$ and edges of U and V, respectively. The remaining (i.e., not belonging to a major) $2m$-periodic vertices x, y of U are such that $\sigma_3^m(x) = \tau(y)$ while $\sigma_3^m(y) = \tau(x)$.

(D) There are exactly two cycles of Fatou gaps of the same period, interchanged by τ. Critical gaps U, $V = \tau(U)$ belong to different cycles. The periodic majors M and $\tau(M)$

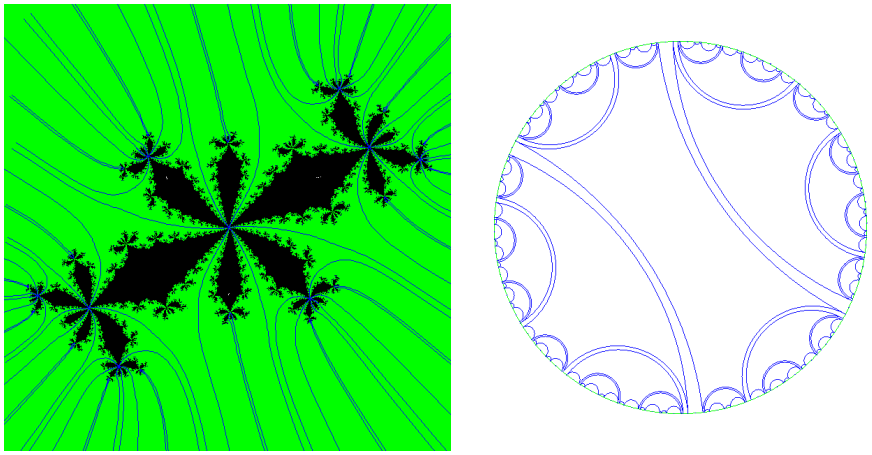


Figure 4: The symmetric cubic lamination with type D comajor $(\frac{7}{78}, \frac{4}{39})$ on the right and the Julia set of a corresponding polynomial with external rays on the left.

of $\mathcal{L}$ are edges of U and V, respectively. Non-periodic majors are siblings of M and $\tau(M)$ and edges of U and V, respectively.

In either case all edges of infinite gaps are eventually mapped to periodic majors. The only periodic orbit of edges of a Fatou gap of $\mathcal{L}$ is the orbit of a major of $\mathcal{L}$ and it has the same period as the Fatou gap.

Lemma 3.9 summarizes the results of [8] (compare [8], Lemma 3.8) dealing with symmetric non-empty (i.e., having some non-degenerate leaves) Fatou laminations. Figures 3 and 4 show examples of polynomials of type B and D and the corresponding laminations.

Proof. All claims of the lemma except for the last one are immediate; observe that the claims concerning the period of the majors follow from Lemma 3.3. To prove the last claim consider an edge ℓ of a critical gap U from a cycle $\mathcal{T}$ of Fatou gaps. It is well-known that any edge of U eventually maps to a periodic or a critical edge. Since, evidently, U has no critical edges, it suffices to prove that the only periodic edge of U is M. Indeed, let $N \neq M$ be a periodic edge of U. Then no image of N can be a point (since N is periodic) or a diameter of $\overline{\mathbb{D}}$ (since otherwise N itself is a diameter invariant under σ_3, hence $N = M$).

Take the closest approach N' in length to $\frac{1}{3}$ among the images of N. By Lemma 3.3 an eventual image of N' that is an edge of U must coincide with M, a contradiction. □

The terminology below is adopted from [31, 34], see also [7].

Definition 3.10. Symmetric Fatou laminations with properties from Lemma 3.9(B) (respectively, Lemma 3.9(D)) are said to be *of type B* (respectively, *of type D*).

We will also need an immediate corollary of Lemma 6.1 of [8].

Corollary 3.11 ([8], Lemma 6.1)**.** *Distinct symmetric Fatou laminations have disjoint comajors.*

Next we consider infinite gaps of C_sCL. One of them plays a special role. Recall that $\mathcal{O}$ is the center of $\mathbb{D}$. Each comajor is of length at most $\frac{1}{6}$. Hence $\mathcal{O}$ does not belong to any comajor; it must then lie inside a gap. The *main gap* G_{main} is by definition the gap of C_sCL that contains $\mathcal{O}$ in its interior.

Theorem 3.12. *The gap G_{main} is infinite, and $\tau(G_{main}) = G_{main}$. Each edge ℓ of G_{main} is a comajor with the same image as the longest edge of a σ_3-invariant symmetric finite rotational gap H and is associated with the symmetric Fatou lamination $\mathcal{L}_{\mathcal{H}}$ formed by H, Fatou gaps of degree greater than 1 attached to H and "rotating" around H, and their pullbacks.*

Proof. If G_{main} is finite, then, by Theorem 3.7, it is a Misiurewicz cocritical gap of preperiod at least 2 of a symmetric lamination $\mathcal{L}$. This is a contradiction, since then the other cocritical set of $\mathcal{L}$ contains $\mathcal{O}$ and intersects the interior of G_{main}. Thus, G_{main} is infinite.

Let ℓ be an edge of G_{main} and the marked comajor of a symmetric lamination $\mathcal{L}$. Since $\mathcal{O} \in G_{main}$, then ℓ cannot be located under another comajor. By Theorem 3.7, the leaf ℓ can only be a 1-preperiodic comajor of a Fatou lamination $\mathcal{L}$. Let U be the marked critical Fatou gap of $\mathcal{L}$ with periodic major M. If M is the limit of leaves of $\mathcal{L}$ (necessarily from outside of U), then ℓ is the limit of leaves ℓ_i so that ℓ is located under ℓ_i for any i. By Lemma 6.6 of [8], this implies that ℓ is the limit of the comajors under which ℓ

is located, a contradiction. Hence M is an edge of a periodic gap and is isolated in $\mathcal{L}$. Clearly, so is $\tau(M)$. Let us remove the grand orbits of M and $\tau(M)$ from $\mathcal{L}$ and consider the resulting family of leaves $\mathcal{L}'$. It easily follows (essentially, by definition) that $\mathcal{L}'$ is again a symmetric lamination. If $\mathcal{L}' \subsetneqq \mathcal{L}$ is non-empty, then, evidently, it has a comajor ℓ' such that ℓ is under ℓ', a contradiction. Thus, $\mathcal{L}'$ is the empty lamination and, so, the grand orbits of M and $\tau(M)$ form the entire $\mathcal{L}$.

Since $\mathcal{L}$ must have a finite invariant gap H, it follows that $\mathcal{L}$ consists of H, Fatou gaps attached to H and "rotating" around H, and their iterated pullbacks. Observe that H itself must be symmetric under τ. By Lemma 3.9, the lamination $\mathcal{L}$ can be of type B or D. By definition, the two periodic majors of $\mathcal{L}$ are the closest to criticality edges of H (in this case it is equivalent to being the longest). There are two comajors of H; just like in the case of symmetric polynomials, either of them can be marked, and so in C_sCL the lamination $\mathcal{L}$ is reflected twice. □

The following notion will allow us to deal with type B and D laminations in a unified fashion.

Definition 3.13 (First (half-)return)**.** Let $\mathcal{L}$ be a symmetric Fatou lamination and U be a critical gap of $\mathcal{L}$. If $\mathcal{L}$ is of type D and the critical gap U is of period n then set $\eta = \sigma_3^n|_U$. If $\mathcal{L}$ is of type B and the critical gap U is of period $2m$ set $\eta = \tau \circ \sigma_3^m|_U$. Thus, η is a self-map of $U \cap \mathbb{S}^1$; it can also be extended linearly over the edges of U and, using a barycentric construction, inside U. The map η is called the *first (half-)return map* of U.

Strictly speaking, η depends on the choice of a symmetric lamination and its gap, however, we will not reflect it in writing to lighten the notation. Lemma 3.14 is left to the reader.

Lemma 3.14. *Let U be a critical gap of a symmetric Fatou lamination $\mathcal{L}$. Then η maps ∂U onto ∂U in a 2-to-1 fashion and is semiconjugate to σ_2 by a monotone map ϕ collapsing edges of U to points. The fixed point set of $\eta|_{\partial U}$ is the periodic major of $\mathcal{L}$. If $\ell \subset \overline{\mathbb{D}}$ is a chord*

whose endpoints are never mapped to the σ_2-fixed point, then the ϕ-preimage of $\ell \cap \mathbb{S}^1$ spans a chord in U that has a unique sibling $\ell' \subset \overline{\mathrm{co}}(U)$.

The leaf/point ℓ' from the last claim of Lemma 3.14 is said to be *induced* by ℓ.

A parabolic quadratic polynomial from the Main Cardioid has a lamination $\mathcal{L}_2$ called *central*; the major of $\mathcal{L}_2 \neq \varnothing$ is an edge shared by a finite invariant gap and a critical periodic Fatou gap.

Theorem 3.15. *Let G be an infinite gap of C_sCL not containing $\mathcal{O}$. Then, for some Fatou lamination $\mathcal{L}$, a cocritical Fatou gap V of $\mathcal{L}$ contains G, and ∂G consists of single points and chords in V corresponding to majors of σ_2-invariant central laminations. In particular, edges of G are 1-preperiodic while other vertices of G have infinite orbits.*

Proof. Let us consider the *ceiling* of the gap G, that is the unique edge ℓ that separates G from $\mathcal{O}$. In other words, the gap G is located under ℓ. By Theorem 3.7, the edge ℓ is a comajor of a Fatou lamination $\mathcal{L}$. Denote by U the critical Fatou gap of $\mathcal{L}$ with periodic major M such that $\sigma_3(M) = \sigma_3(\ell)$. Then all edges of G are associated with symmetric laminations $\mathcal{L}'$ that tune $\mathcal{L}$. Evidently, $\mathcal{L}'|_U$ is invariant under η, the first (half-)return map introduced in Definition 3.13. Recall that η is of degree 2 and is modeled by σ_2; the map $\phi : \partial U \to \mathbb{S}^1$ collapses the edges of U and semiconjugates η with σ_2, as explained above.

Set $V = \overline{\mathrm{co}}(U)$. Periodic majors M' of σ_3 in U correspond under ϕ to periodic majors M'' of σ_2 in $\mathbb{D}$. It follows that the minor $\overline{c}'' = \sigma_2(M'')$ of σ_2 defines a chord $\overline{c}'$ in V that is mapped under σ_3 to the σ_3-image of M'. As is immediate from the definitions, $\overline{c}'$ is a comajor corresponding to M'. Thus, there is a natural correspondence between the quadratic minors and the cubic comajors in V. Under this correspondence, G pairs with the central gap of the quadratic minor lamination. Hence, the edges of G are associated with central quadratic laminations. □

4 Connectedness of $\mathcal{M}_3^{sy}$

Recall that h^k denotes the k-th iteration of a map h.

Lemma 4.1. *The set $\mathcal{M}_3^{sy}$ is invariant under the multiplication by i.*

Proof. We claim that if $f(z) = z^3 - az$ and $g(z) = z^3 + az$ then f^2 and g^2 are conjugate. Indeed, f and $-g$ are conjugate by the map $I : z \mapsto iz$; hence f^2 and $(-g)^2$ are conjugate by I. Since g is an odd function, we have $(-g)^2 = g^2$. Thus, I conjugates f^2 and g^2. Since $p_{ic} = z^3 + 3c^2z$ while $p_c(z) = z^3 - 3c^2z$, then p_c^2 and p_{ic}^2 are conjugate. □

We now need a construction similar to that for quadratic polynomials; in our description below we follow the exposition from [29]. Take a topological disk Δ_c around infinity in $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ that contains no critical points of p_c, does not contain 0, and is such that $p_c(\Delta_c) \subset \Delta_c$ (in particular, $0 \notin p_c^n(\Delta_c)$ for $n \geq 0$). Define Böttcher function $B_c(z) = \lim_{n\to\infty}(p_c^n(z))^{1/3^n}$ on Δ_c, where the root is taken so that the corresponding functions are tangent to the identity at infinity. The existence of a single valued branch follows from the fact that Δ_c is simply connected, and that 0 does not belong to $p_c^n(\Delta_c)$. Recall that the Green function $g_c : \mathbb{C} \to \mathbb{R}$ is defined as $g_c(z) = \lim_{n\to\infty} 3^{-n} \log_+ |p_c^n(z)|$, where $\log_+(t)$ is the maximum of 0 and $\log t$. Then $g_c(z) = \log|B_c(z)|$ for all $z \in \Delta_c$. The *equipotential* $E_c(t) \subset \mathbb{C}$ is defined as the level set $\{g_c = t\}$ of the Green function; this is a real analytic curve for $t > 0$, possibly singular. Note that B_c conjugates p_c and z^3 near infinity.

For $c \in \mathbb{C} \setminus \mathcal{M}_3^{sy}$, set Δ_c to be the exterior of the equipotential of p_c passing through $\pm c$ (the equipotential $E_c(g_c(c))$ has singularities but the exterior of it is a topological disk). Then the required properties of Δ_c are fulfilled. It is easy to see (from the continuous dependence of the Böttcher coordinate on parameters) that the union $\mathcal{U}$ of $\{c\} \times \Delta_c$, where c runs through the complement of $\mathcal{M}_3^{sy}$, is open. Standard arguments show that $B_c(z)$ is analytic in both c and z on $\mathcal{U}$.

If $c \in \mathbb{C} \setminus \mathcal{M}_3^{sy}$, then $(c, -2c)$ is always on the boundary of $\mathcal{U}$, since the value of the Green function of p_c at the cocritical point $-2c$ coincides with those at $\pm c$. However, the map $(c, z) \mapsto B_c(z)$ extends analytically to a neighborhood of $-2c$. Moreover, $-2c$ is a regular

point of this analytic extension in the sense that $z \mapsto B_c(z)$ is a conformal injection in a neighborhood of this point. From now on we will assume that $B_c(z)$ is defined in this neighborhood of $-2c$.

For a fixed $c \notin \mathcal{M}_3^{sy}$ the map B_c is a conformal isomorphism between Δ_c and the set $\{z \in \mathbb{C} : |z| > g_c(c)\} = \mathbb{C} \setminus \overline{\mathbb{D}}_{g_c(c)}$. This defines *initial segments of (dynamical) external rays* $R_c(\theta)$ of p_c, i.e. B_c-preimages of the radial rays of argument θ in $\mathbb{C} \setminus \overline{\mathbb{D}}_{g_c(c)}$. Evidently, these initial segments of external rays of p_c are orthogonal to all equipotentials $E_c(t), t > g_c(c)$. Moreover, as we mentioned above equipotentials can be defined for any $t > 0$. This allows one to give the following definition: a *smooth external ray* R of p_c is a smooth unbounded curve that crosses every equipotential orthogonally and terminates in the Julia set of p_c. All but countably many initial segments of external rays defined above extend as smooth external rays. However countably many initial segments will hit critical points or their eventual preimages (in what follows such points are called *(pre)critical* or eventually critical) and, therefore, will not extend as smooth external rays.

Let $\Psi(c) = B_c(-2c)$ be the Böttcher coordinate of the marked cocritical point in the sense of the analytic continuation mentioned above. Then Ψ is well-defined and holomorphic on $\mathbb{C} \setminus \mathcal{M}_3^{sy}$. Theorem 4.2 is analogous to the corresponding statement for the Mandelbrot set; it proves the first claim of the Main Theorem stated in the Introduction. In what follows we denote the Riemann sphere by $\overline{\mathbb{C}}$.

Theorem 4.2. *The symmetric connectedness locus $\mathcal{M}_3^{sy}$ is a full continuum.*

Proof. Let us show that Ψ maps $\mathbb{C} \setminus \mathcal{M}_3^{sy}$ onto $\mathbb{C} \setminus \overline{\mathbb{D}}$. First we prove that $\Psi(c) \sim \sqrt[3]{2}c$ as $c \to \infty$. Set $z_n = p_c^n(-2c)$ and let

$$r_n = z_{n+1}/z_n^3 = 1 - 3c^2/z_n^2.$$

If $|c| \geq 2$ and $|z| \geq 2|c|$, then $|p_c(z)/z| \geq 4|c|^2 - 3|c|^2 = |c|^2 \geq 4$ and hence, $|p_c(z)| \geq 4|z|$. Thus, $|z_n| \geq 2 \cdot 4^{n-1}|c|^3 \geq 2|c|^3$ for $n \geq 1$ (which implies that $\mathcal{M}_3^{sy} \subset \mathbb{D}_2$) and for any $c \in \mathbb{D}_2$ we have $J(p_c) \subset \mathbb{D}_4$.

We conclude that $|c| \geq 2$ yields

$$1 - \frac{3}{4|c|^4} \leq |1 - 3c^2/z_n^2| = |r_n| \leq 1 + \frac{3}{4|c|^4}.$$

On the other hand,

$$z_n = r_{n-1} z_{n-1}^3 = r_{n-1} r_{n-2}^3 z_{n-2}^9 = \cdots = r_{n-1} r_{n-2}^3 \cdots r_2^{3^{n-3}} r_1^{3^{n-2}} (2c^3)^{3^{n-1}}$$

which yields that

$$\sqrt[3^n]{z_n} = (\sqrt[3^n]{r_{n-1}}\, \sqrt[3^{n-1}]{r_{n-2}} \cdots \sqrt[2^7]{r_2} \sqrt[9]{r_1})(\sqrt[3]{2})c.$$

Using the above bounds on $|r_n|$, the formula for the sum of the geometric series, and the fact that $\Psi(c) = \lim_{n\to\infty} \sqrt[3^n]{z_n}$ we see that

$$\left(\sqrt[6]{1 - \frac{3}{4|c|^4}}\right)\sqrt[3]{2}|c| \leq \sqrt[3^n]{|z_n|} \leq \left(\sqrt[6]{1 + \frac{3}{4|c|^4}}\right)\sqrt[3]{2}|c|,$$

that yields the following: as $c \to \infty$, $\Psi(c) = \sqrt[3]{2}c(1 + O(1/c^{2/3}))$. Thus, the map Ψ can be continuously extended to ∞ with $\Psi(\infty) = \infty$ implying that it can be done holomorphically and that the local degree of Ψ at ∞ is 1. In particular, ∞ is in the interior of the range of Ψ.

By the above, $\mathcal{M}_3^{sy} \subset \mathbb{D}_2$, in particular, $\mathcal{M}_3^{sy}$ is compact. Let M be the maximum of the continuous function $(z, c) \longmapsto g_c(z)$ on the compact set $\overline{\mathbb{D}}_2 \times \overline{\mathbb{D}}_4$. It follows that $|B_c(z)| \leq e^M$ for all $c \in \overline{\mathbb{D}}_2$ and $z \in \overline{\mathbb{D}}_4$ such that $(c, z) \in \overline{\mathcal{U}}$.

We claim that if $c_n \to \mathcal{M}_3^{sy}$ then $|\Psi(c_n)| \to 1$. Indeed, otherwise there exists a sequence $c_n \to c_0 \in \mathcal{M}_3^{sy}$ with $|\Psi(c_n)| > e^m > 1$. Take k such that $3^k m > M$. Since $|B_c(p_c^k(-2c))| = |(B_c(p_c(-2c))^{3^k}| = |\Psi(c)|^{3^k}$, then $|B_{c_n}(p_{c_n}^k(-2c_n))| > e^M$ which, by the choice of M, implies $|p_{c_n}^k(-2c_n)| > 4$ for all sufficiently large n (indeed, $|c_n| \leq 2$ for large n). By continuity, $p_{c_0}^k(-2c_0) \geq 4$; this shows that the cocritical point of p_{c_0} escapes to infinity contradicting the choice of c_0.

Since ∞ is not on the boundary of the range of Ψ it follows from the above that Ψ is a proper holomorphic map from $\overline{\mathbb{C}} \setminus \mathcal{M}_3^{sy}$ onto $\overline{\mathbb{C}} \setminus \overline{\mathbb{D}}$. Hence it is a branched covering with a well-defined degree. However, the point ∞ has exactly one preimage of degree 1; hence Ψ has degree 1 and is actually a conformal isomorphism. □

The function Ψ gives us an analogue of Böttcher coordinates for the complement of $\mathcal{M}_3^{sy}$. In particular, we can define *external parameter rays* (or simply *parameter rays*) as preimages of radial straight lines under Ψ, namely, $\mathcal{R}_\theta(r) = \Psi^{-1}(re^{2\pi i\theta})$ with $r \in (1,\infty)$. The parameter ray $\mathcal{R}_\theta$ *lands* at a parameter w if $\lim_{r\to 1} \mathcal{R}_\theta(r) = w$. Note that, by definition, the parameter ray $\mathcal{R}_\theta$ consists of the parameters c such that the marked cocritical point $-2c$ of p_c belongs to the dynamical external ray $R_c(\theta)$ of p_c. Also note, that the map Ψ is a conformal isomorphism between $\mathbb{C}\setminus\mathcal{M}_3^{sy}$ and $\mathbb{C}\setminus\overline{\mathbb{D}}$ tangential to the map $z \longmapsto \sqrt[3]{2}z$ at infinity.

5 Hyperbolic components of $\mathcal{M}_3^{sy}$ and their roots

Hyperbolic components of polynomial parameter spaces play an important role in complex dynamics. Here we study them for the parameter space of symmetric cubic polynomials.

5.1 Preliminaries

We start by recalling basic definitions. Let $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a rational function. The *multiplier* $\rho(z)$ of a periodic point z of minimal period n under f is defined as the derivative of the first return map to z, that is, $\rho(z) = (f^n)'(z)$. A periodic point z is said to be *super-attracting* if $\rho(z) = 0$ (which implies that the orbit of z contains a critical point), *attracting* if $|\rho(z)| < 1$, and *parabolic* if $\rho(z)$ is a root of unity (which implies that $(f^{kn})'(z) = 1$ for some k).

A polynomial $p_c \in \mathcal{M}_3^{sy}$ (and the parameter c) is *hyperbolic*/*parabolic* if both of its *finite* critical points are attracted to *finite* attracting/parabolic cycles. To characterize such polynomials we need a lemma.

Lemma 5.1. *If U is an open topological disk and $-U = U$, then $0 \in U$. If A is a full continuum and $-A = A$, then $0 \in A$.*

Proof. Set $s(z) = z^2$; then $U = s^{-1}(s(U))$, and $s : U \to s(U)$ is a branched covering. The claim now follows from the Riemann–Hurwitz formula applied to this covering. Take a

tight symmetric Jordan neighborhood V of A and set $U = s^{-1}(V)$; then $0 \in U$ by the above. Since A is the intersection of all such U, it follows that $0 \in A$. □

We can now describe hyperbolic polynomials p_c more explicitly.

Lemma 5.2. *A polynomial p_c is hyperbolic if and only if it possesses one of the following:*

(a) an invariant symmetric attracting Fatou domain on which p_c is 3-to-1; this happens if and only if $|c| < \sqrt{1/3}$, or

(b) a unique symmetric cycle of attracting Fatou domains of period $2n \geq 2$; there are exactly two mutually symmetric domains in the cycle containing critical points $\pm c$, or

(d) two mutually symmetric attracting cycles of Fatou domains.

Thus, if p_c has an attracting cycle, then p_c is hyperbolic. Also, case (a) is the only case when a hyperbolic polynomial p_c has a unique bounded periodic Fatou domain.

Proof. If there exists a Fatou domain U on which the map is 3-to-1, then U must be symmetric (otherwise, $-U$ is another Fatou domain on which the map is 3-to-1, which is impossible). By Lemma 5.1, this implies that $0 \in U$ and U is invariant. Since there must exist a unique fixed point in U and this point must be attracting, then 0, being a fixed point, must be attracting. Since $p_c'(0) = -3c^2$, the corresponding hyperbolic component of $\mathcal{M}_3^{sy}$ is the round disk of radius $\sqrt{1/3}$ centered at the origin. This corresponds to case (a) and covers $c = 0$ so from now on we assume that $c \neq 0$ and hence p_c has distinct critical points c and $-c$ with mutually symmetric orbits.

If c (resp., $-c$) is attracted to an attracting cycle, then so is $-c$ (resp., c) which implies that if p_c has an attracting cycle, then p_c is hyperbolic. We can also assume that there are no 3-to-1 Fatou domains. Now, if p_c has a cycle A of attracting Fatou domains then by symmetry $-A$ is also a cycle of attracting Fatou domains. Suppose that $A = -A$. By the assumption critical domains in A contain exactly one critical point; since p_c is symmetric, the critical domains in A are mutually symmetric. Moreover, the fact that p_c is symmetric

implies that the first iterate p_c^n that maps either critical domain from A to the other one is the same for both critical domains which implies that the period of A is $2n$. This corresponds to case (b). Otherwise A and $-A$ are distinct cycles of Fatou domains which corresponds to case (d). □

Lemma 5.3 is similar to Lemma 5.2 and its proof is left to the reader.

Lemma 5.3. *Suppose that a polynomial p_c has a parabolic cycle. Then one of the following holds:*

(b) a unique symmetric cycle of parabolic Fatou domains of p_c of period $2n \geq 2$ has exactly two mutually symmetric critical domains;

(d) two parabolic cycles of Fatou domains of p_c are mutually symmetric.

For brevity, a Cremer/Siegel point (cycle) of a polynomial will be referred to as a *CS-point (cycle)*. Now we show that for CS-cycles the situation with symmetric polynomials is similar to that in Lemma 5.2. First we state a part of Theorem 4.3 from [4] combined with results from [18] and [23]. Define a *rational cut* as the union of two external rays with rational arguments that land on the same point called the *vertex* of the cut. If the vertex is a repelling (parabolic) periodic point, then we call the cut *repelling (parabolic)*.

Theorem 5.4 ([4, 18, 23])**.** *Let P be a polynomial and T be CS-cycle. There exists a recurrent critical point c of P and a point $q \in T$ that are not separated by any rational cut of P. Two different objects, each of which is a CS-point, a parabolic domain, or an attracting domain, are always separated by a rational cut.*

Theorem 5.4 is used in the proof of the following lemma.

Lemma 5.5. *Suppose that a polynomial p_c has a CS-cycle T. Then one of the following holds:*

(a) the only non-repelling cycle of p_c is $T = \{0\}$, and neither of the critical points $\pm c$ is separated from T by a rational cut;

(b) the only non-repelling cycle of p_c is T, it is a symmetric cycle of period $2n$;

(d) there are exactly 2 non-repelling cycles, namely, T and $-T$.

Proof. Assume that 0 is a CS-point. Then, by Theorem 5.4, there is a recurrent critical point c not separated from 0 by any rational cut, and the same holds for $-c$. This corresponds to case (a) of the lemma.

If 0 is parabolic or attracting then it attracts at least one critical point of p_c and hence, by symmetry, both of them. In this case p_c has no other non-repelling cycles. So, from now on we assume that 0 is repelling.

If p_c has a symmetric CS-cycle T of period $2n$ then, by Theorem 5.4, it has a point $w \in T$ not separated from a recurrent critical point, say, c, of p_c by a rational cut; hence, $-w \in T$ is not separated from a recurrent critical point $-c$ of p_c by a rational cut. This, again by Theorem 5.4, implies that there are no other non-repelling cycles of p_c. This corresponds to case (b) of the lemma.

Finally, let T and $-T$ be distinct mutually symmetric CS-cycles of p_c. By Theorem 5.4, we may assume that T has a point w not separated from a recurrent critical point, say, c of p_c by a rational cut; hence, $-w \in -T$ is not separated from a recurrent critical point $-c$ of p_c by a rational cut. This implies that there are no other non-repelling cycles of p_c. This corresponds to case (d) of the lemma. □

5.2 Hyperbolic components and multipliers

Let p_{c_0} have a periodic point w of period n such that $(p_{c_0}^n)'(w) \neq 1$. By the implicit function theorem applied to the equation $p_c^n(z) = z$, there is a holomorphic function $\alpha(c)$ defined on an open Jordan disk around c_0 such that $\alpha(c_0) = w$, and $\alpha(c)$ is a periodic point of p_c of period n. Also, the multiplier $(p_c^n)'(\alpha(c))$ is a holomorphic function of c. Hence the set of hyperbolic parameters is an open subset of $\mathcal{M}_3^{sy}$; a connected component $\mathcal{H}$ of this set is called a *hyperbolic component* of $\mathcal{M}_3^{sy}$. For any $c \in \mathcal{H}$ the ω-limit set of the marked critical point is the unique *marked* attracting cycle Q_c; the *period* of $\mathcal{H}$ is the period of Q_c.

Conjecturally, every connected component of the interior of $\mathcal{M}_3^{sy}$ is hyperbolic.

Theorem 5.6 ([41]). *Every component of the Fatou set of a rational function is eventually periodic. In particular, any bounded Fatou domain of a hyperbolic symmetric cubic polynomial eventually maps into a cycle of Fatou domains that contains an attracting cycle.*

Recall that, by Lemma 5.2, the set $\mathbb{D}_{\sqrt{1/3}}$ is a hyperbolic component of $\mathcal{M}_3^{sy}$.

Definition 5.7. The set $\mathbb{D}_{\sqrt{1/3}}$ is called the *main hyperbolic component* of $\mathcal{M}_3^{sy}$ and is denoted $\mathcal{H}_{main}$.

Corollary 5.8 follows from Lemmas 5.2, 5.3 and 5.5.

Corollary 5.8. *A polynomial p_c has one symmetric non-repelling cycle, or two mutually symmetric non-repelling cycles with equal multipliers, or no non-repelling cycles at all.*

Let $\mathcal{H} \neq \mathcal{H}_{main}$ be a hyperbolic component and $c \in \mathcal{H}$. Denote by $F_c^{\pm}$ the critical Fatou domains of p_c that correspond to the critical Fatou gaps $U_{\mathcal{H}}^{\pm}$, respectively, of $\mathcal{L}_{\mathcal{H}}$ (we always assume that the marked critical point c belongs to F_c^+). Let $M_{\mathcal{H}}$ and $M'_{\mathcal{H}}$ be the majors of $\mathcal{L}_{\mathcal{H}}$ that are edges of $U_{\mathcal{H}}^+$; then $\sigma_3(M_{\mathcal{H}}) = \sigma_3(M'_{\mathcal{H}})$ by Lemma 3.9, and we always assume that $M_{\mathcal{H}}$ is periodic. Let $\overline{\text{co}}_{\mathcal{H}}^+$ be the marked comajor (i.e., $\sigma_3(\overline{\text{co}}_{\mathcal{H}}^+) = \sigma_3(M_{\mathcal{H}})$) and let $\overline{\text{co}}_{\mathcal{H}}^-$ be the other comajor of $\mathcal{L}_{\mathcal{H}}$. Similar objects can be defined for any hyperbolic or parabolic parameter c yielding such notation as $U_c^{\pm}$, M_c, M'_c.

Let a hyperbolic component $\mathcal{H}$ be given. All $c \in \mathcal{H}$ satisfy the same option (a), (b) or (d) from Lemma 5.2. According to these three cases, $\mathcal{H}$ is said to be of *type* A, B or D, respectively. Also, given a polynomial p_c with parabolic (or attracting) periodic point z, let $m_r(c)$ be the minimal number such that $p_c^{m_r(c)}$ fixes dynamical external rays landing on z (or the point z itself if it is attracting); note that it does not depend on the choice of a particular ray and is called the *ray period* of p_c. The number $(p_c^{m_r(z)})'(z)$ is called the *ray multiplier* (of p_c) and is denoted $\rho(c)$. Notice that the period of a parabolic point z may be strictly smaller than $m_r(c)$ so that $m_r(c)$ is a multiple of the period.

A similar concept can be defined for a hyperbolic component $\mathcal{H}$. Namely, let $\rho_{\mathcal{H}} : \mathcal{H} \to \mathbb{D}$ be defined as $\rho_{\mathcal{H}}(c) = \rho(c)$. As we show in Theorems 5.9 and 5.10, the function

$\rho_{\mathcal{H}}$ can be extended over $\overline{\mathcal{H}}$. For a parabolic parameter $c \in \partial\mathcal{H}$, this extended function $\rho_{\mathcal{H}}$ may not be equal to $\rho(c)$. To emphasize that, we use the subscript $_{\mathcal{H}}$ in the notation, and call $\rho_{\mathcal{H}}$ the *ray multiplier based on* $\mathcal{H}$. This difference is not present for parameters $c \in \mathcal{H}$, but does show for parameters $c \in \partial\mathcal{H}$.

Theorem 5.9. *For a type D hyperbolic component $\mathcal{H}$ of $\mathcal{M}_3^{sy}$ the map $\rho_{\mathcal{H}}$ can be extended onto $\overline{\mathcal{H}}$ so that $\rho_{\mathcal{H}} : \overline{\mathcal{H}} \to \overline{\mathbb{D}}$ is a homeomorphism conformal on $\mathcal{H}$.*

Proof. The result essentially follows from Theorem C of [21], however, we need to explain how the terminology of Inou–Kiwi relates to ours. Let $\mathcal{L}$ be a cubic invariant q-lamination with at least one cycle of Fatou gaps. With $\mathcal{L}$, one associates a *reduced mapping schema* $T(\mathcal{L})$. Instead of giving a general definition of mapping schemata, we give an explicit description of $T(\mathcal{L})$ in the case when $\mathcal{L}$ is symmetric of type D, that is, when $\mathcal{L}$ has two distinct cycles of Fatou gaps. In this case, $T(\mathcal{L})$ can be represented as the graph with two vertices and two (directed) edges that are loops based at both vertices. Every edge of $T(\mathcal{L})$ is in general equipped with a positive integer called the *degree*; in our specific case, the degrees of both loops are equal to 2. Intuitively, the arrows of $T(\mathcal{L})$ represent the first return maps to the critical Fatou gaps of $\mathcal{L}$. The space $\mathcal{C}(T(\mathcal{L}))$ in our specific case consists of all pairs (q_0, q_1) of monic centered quadratic polynomials q_0, q_1 with connected Julia sets — informally, the two loops of $T(\mathcal{L})$ are replaced with q_0 and q_1.

By definition, the space $\mathfrak{R}(\mathcal{L})$ consists of all monic cubic polynomials f such that

- the filled Julia set $K(f)$ is connected;
- for any (pre)periodic leaf of $\mathcal{L}$ with endpoints α, β, the corresponding external rays $R_f(\alpha)$ and $R_f(\beta)$ land on the same (pre)periodic point of $K(f)$;
- let U_0, U_1 be the critical Fatou gaps of f; the corresponding subcontinua K_0 and K_1 of $K(f)$ are polynomial-like filled Julia sets of certain polynomial-like restrictions of f^n, where n is the period of U_0 and U_1.

Here, one needs to explain in which sense K_0 corresponds to U_0, and similarly with K_1 and U_1. For each $\ell \in \mathcal{L}$, let α and β be the endpoints of ℓ, and write Γ_ℓ for the cut

formed by the external rays $R_f(\alpha)$, $R_f(\beta)$, and their common landing point. Then K_0 corresponding to U_0 means that K_0 lies on the same side of Γ_ℓ as U_0 relative to ℓ, for every $\ell \in \mathcal{L}$. By the Douady–Hubbard straightening theorem, f^n is hybrid equivalent to a unique monic quadratic polynomial q_i near K_i, where $i = 0, 1$. The *Inou–Kiwi straightening map* (abbreviated as *IK-straightening map*) $\chi_{\mathcal{L}} : \mathfrak{R}(\mathcal{L}) \to \mathcal{C}(T(\mathcal{L}))$ takes f to (q_0, q_1). (The fact that q_i are quadratic yields some simplification: in higher degree cases one needs additional normalization called *internal angles assignment* in order to make q_i unique).

Theorem C of [21] can now be formulated as follows. *Denote by* $\mathrm{Hyp}(\mathcal{C}(T(\mathcal{L})))$ *the set of hyperbolic maps contained in* $\mathcal{C}(T(\mathcal{L}))$ *(in our specific case,* $(q_0, q_1) \in \mathcal{C}(T(\mathcal{L}))$ *being hyperbolic means both* q_0 *and* q_1 *are hyperbolic). Then* $\chi_{\mathcal{L}}(\mathfrak{R}(\mathcal{L})) \supset \mathrm{Hyp}(\mathcal{C}(T(\mathcal{L})))$*, the inverse image of* $\mathrm{Hyp}(\mathcal{C}(T(\mathcal{L})))$ *under* $\chi_{\mathcal{L}}$ *is an open set, and the restriction of* $\chi_{\mathcal{L}}$ *onto this open set is biholomorphic.* Now let $\mathcal{H}$ be a given type D hyperbolic component of $\mathcal{M}_3^{sy}$; it lies in some hyperbolic component $\hat{\mathcal{H}}$ of the connectedness locus of all monic centered cubic polynomials. All polynomials from $\hat{\mathcal{H}}$ have the same lamination, say, $\mathcal{L}$. By Theorem C of [21], the restriction of $\chi_{\mathcal{L}}$ to $\hat{\mathcal{H}}$ is a biholomorphic isomorphism between $\hat{\mathcal{H}}$ and the product of the interior Ca of the main cardioid with itself. The image of $\mathcal{H}$ under $\chi_{\mathcal{L}}$ is then the diagonal in $\mathrm{Ca} \times \mathrm{Ca}$, and the restriction of $\chi_{\mathcal{L}}$ is a biholomorphic map between $\mathcal{H}$ and this diagonal (the latter is isomorphic to $\mathbb{D}$ under the multiplier map). It follows that $rp_{\mathcal{H}} : \mathcal{H} \to \mathbb{D}$ is a conformal isomorphism. Since the boundary of any hyperbolic component is contained in a real algebraic curve and, as such, is locally connected, then, clearly, it extends to a homeomorphism $rp_{\mathcal{H}} : \overline{\mathcal{H}} \to \overline{\mathbb{D}}$. Observe that the situation here is similar to the quadratic case [15]. □

Let a polynomial p_c have a parabolic or attracting cycle X_c of type B. Then the period of X_c is an even number $2n$, and $p_c^n(x) = -x$ for every $x \in X_c$. The number $-(p_c^n)'(x)$ does not depend on the point $x \in X_c$, is denoted $r\tilde{p}(c)$, and is called the *ray half-multiplier* (of p_c). Note that $r\tilde{p}(c)$ can be interpreted as the multiplier of the fixed point x of the map $-p_c^n$. As before, the ray half-multiplier depends only on the parameter and is defined as long as p_c is of type B. If $c \in \mathcal{H}$, where $\mathcal{H}$ is of type B, then we write $r\tilde{p}_{\mathcal{H}}$ for the restriction

of the function $r\tilde{\rho}_{\mathcal{H}}(c)$ to $\mathcal{H}$. The *ray (half-)multiplier* of c is defined as either $r\rho(c)$ or $r\tilde{\rho}(c)$ depending on whether c is type D or type B.

Theorem 5.10. *Let $\mathcal{H}$ be a hyperbolic component of $\mathcal{M}_3^{sy}$ of type B. The map $r\tilde{\rho}_{\mathcal{H}}$ can be extended over $\overline{\mathcal{H}}$ so that $r\tilde{\rho}_{\mathcal{H}} : \overline{\mathcal{H}} \to \overline{\mathbb{D}}$ is a homeomorphism which is conformal on $\mathcal{H}$ while the ray multiplier $r\rho_{\mathcal{H}} = r\tilde{\rho}^2_{\mathcal{H}} : \overline{\mathcal{H}} \to \overline{\mathbb{D}}$ is a double-covering.*

Proof. Similarly to Theorem 5.9, Theorem 5.10 also follows from Theorem C of [21]. Let $\hat{\mathcal{H}}$ be the hyperbolic component in the full space of monic centered cubic polynomials containing $\mathcal{H}$. All polynomials from $\hat{\mathcal{H}}$ have the same lamination $\mathcal{L}$ so that the corresponding reduced mapping schema $T(\mathcal{L})$ has two vertices and two directed edges connecting the two vertices in opposite directions; each edge has degree 2. The space $\mathcal{C}(T(\mathcal{L}))$ consists of pairs (q_0, q_1) of monic centered quadratic polynomials such that the Julia set of $q_1 \circ q_0$ is connected. The straightening map of Inou–Kiwi provides a biholomorphic isomorphism between $\hat{\mathcal{H}}$ and the *principal hyperbolic component* in $\mathcal{C}(T(\mathcal{L}))$ consisting of all (q_0, q_1) such that $q_1 \circ q_0$ has an attracting fixed point, and $J(q_1 \circ q_0)$ is a Jordan curve. Clearly, symmetric cubic polynomials are mapped to pairs (q_0, q_1), for which $q_0 = q_1$. The corresponding slice of the principal hyperbolic component is biholomorphic to $\mathbb{D}$, and the corresponding conformal isomorphism is given by the multiplier of the attracting fixed point of $q_0 = q_1$. □

Let us now define the center and the root of a hyperbolic component.

Definition 5.11 (Center and root)**.** Let $\mathcal{H}$ be a hyperbolic component of type B or D. The *center* of $\mathcal{H}$ is the point $c \in \mathcal{H}$ such that p_c has superattracting cycle(s). The *root* of $\mathcal{H}$ is the point $r_{\mathcal{H}} \in \partial\mathcal{H}$ such that $r\rho_{\mathcal{H}}(r_{\mathcal{H}}) = 1$ (if $\mathcal{H}$ is of type D) or $r\tilde{\rho}_{\mathcal{H}}(r_{\mathcal{H}}) = 1$ (if $\mathcal{H}$ is of type B).

Lemma 5.12 justifies the above definition of the roots.

Lemma 5.12. *Suppose that z is a parabolic point of p_c of ray period n. Then there is a major $\overline{\theta\theta'}$ of $\mathcal{L}_c$ such that both $R_c(\theta)$ and $R_c(\theta')$ land on $p_c^k(z)$, for some k with $0 \le k < n$. Moreover, the ray (half-)multiplier of c equals 1.*

Proof. Let U be a period n parabolic domain attached to z, and F be the Fatou gap of $\mathcal{L}_c$ corresponding to U. Replacing F with a suitable Fatou gap from the same cycle, we may assume that F has a unique periodic major $M = \overline{\theta\theta'}$. If U is of type D then the only p_c^n-fixed point in ∂U is z and the only σ_3^n-fixed points in ∂F form the major M. Hence $R_c(\theta)$ and $R_c(\theta')$ land on z. If U is of type B then $n = 2m$ and there are three p_c^n-fixed points in ∂U associated with M and two vertices x, y of F. Let $x', y' \in \partial U$ be points associated with vertices x, y of F. We claim that x', y' are not parabolic. Suppose that x' is parabolic. Then $p_c^m(y') = -x'$ by part (B) of Lemma 3.9, which implies that y' is also parabolic, a contradiction. Hence only M can be associated with z as desired. Let us now prove that the ray (half-)multiplier of c equals 1.

(B) Let p_c be of type B, $z \in \partial F_c^+$ be parabolic, and F_c^+ be of period $2m$; then $-p_c^m(z) = z$, and, by Lemma 3.14, the map $-p_c^m$ fixes the two external rays of p_c landing on z. Hence $(-p_c^m)'(z) = 1$, which implies that $r\tilde{p}(c) = 1$.

(D) Similar to (B) (details are left to the reader). □

The main component $\mathcal{H}_{main}$ of $\mathcal{M}_3^{sy}$ (for which 0 is attracting) is a round disk of radius $\sqrt{3}/3$. For $|c| < \sqrt{3}/3$ the Julia set $J(p_c)$ is a Jordan curve and $p_c|_{J_c}$ is conjugate to σ_3. In other words, the map is neither of type B nor of type D, and neither Theorem 5.9 nor Theorem 5.10 applies. Each polynomial p_c with $|c| = \sqrt{3}/3$ has a neutral fixed point at 0 of multiplier $-3c^2$. If $c_{1,2} = \pm\frac{i}{\sqrt{3}}$ then the corresponding polynomials $p_{c_{1,2}} = z^3 + z$ have multiplier 1 at fixed point 0. It is easy to see that the lamination associated with $z^3 + z$ has the leaf $\overline{0\,\frac{1}{2}}$, which is the horizontal diameter of the unit circle, and two invariant Fatou gaps located, respectively, above and below $\overline{0\,\frac{1}{2}}$; the presence of these invariant gaps completely determines the lamination. In particular, the laminations at these two parameters are different from the empty lamination which corresponds to any polynomial in $\mathcal{H}_{main}$. We will not consider these points (or any other points) as roots of $\mathcal{H}_{main}$ so that $\mathcal{H}_{main}$ has the center (at the origin) but does not have a root.

6 Parabolic polynomials

The following stability property will be used repeatedly. We state it only for symmetric cubic polynomials.

Theorem 6.1 (Lemma B.1 of [18]). *Let w be a repelling periodic point of p_{c_0} such that an external dynamical ray $R_{c_0}(\theta_0)$ with rational argument $\theta_0 \in \mathbb{R}/\mathbb{Z}$ lands on w. Let θ be an angle such that an eventual σ_3-image of θ belongs to the σ_3-orbit of θ_0 and the rays from the (finite) p_{c_0}-orbit of $R_{c_0}(\theta)$ are all smooth and do not land on critical points. Then, for c sufficiently close to c_0, the dynamical external rays from the p_c-orbit of $R_c(\theta)$ are smooth, move continuously with c, and land on preperiodic or repelling periodic points of p_c close to the landing points of the dynamical external rays of p_{c_0} with the same argument.*

Recall that the lamination $\mathcal{L}_c$ is defined for all c such that $K(p_c)$ is locally connected. In particular, if $\mathcal{H}$ is a hyperbolic component, then for any $c \in \mathcal{H}$ the lamination $\mathcal{L}_c$ exists and is independent of the choice of c. Denote it by $\mathcal{L}_{\mathcal{H}}$; clearly, $\mathcal{L}_{\mathcal{H}}$ is a symmetric Fatou lamination.

Definition 6.2 (Geometrically finite and sub-hyperbolic [20]). A polynomial f is *geometrically finite* if all its critical points are preperiodic points or attracted by parabolic or attracting cycles. If, moreover, f has no parabolic points, then it is said to be *sub-hyperbolic*.

Theorem 6.3 based on [20] and additional arguments.

Theorem 6.3 ([20]). *The following holds.*

1. *A parabolic symmetric polynomial p_c is accessible from a hyperbolic component $\mathcal{H}$ of $\mathcal{M}_3^{sy}$ such that $\mathcal{L}_{\mathcal{H}} = \mathcal{L}_c$.*

2. *If $\mathcal{H}'$ is a hyperbolic component and $c' \in \partial\mathcal{H}'$ is such that $\mathcal{L}_{c'} = \mathcal{L}_{\mathcal{H}'}$, then $c' = r_{\mathcal{H}'}$ and $\overline{\mathrm{co}}^+_{c'} = \overline{\mathrm{co}}^+_{\mathcal{H}}$.*

Proof. (1) Let p_c be the given parabolic polynomial. It is of type B or D. By the main theorem of [20], there exists a path f_t of monic centered cubic polynomials such that $f_0 = p_c$, f_t is sub-hyperbolic for $t > 0$, and $f_t|_{J(f_t)}$ is topologically conjugate to p_c on its Julia set. Since p_c is of type B or D, the polynomial f_t is hyperbolic for any $t > 0$ and, hence, there exists a hyperbolic component $\widehat{\mathcal{H}}$ in the space of all monic centered cubic polynomials that contains the path $\{f_t\}_{t>0}$.

We now use the description of $\widehat{\mathcal{H}}$ given in Theorems 5.9 and 5.10. In both cases (type D or type B), the straightening map (ρ or $\tilde{\rho}$, resp.) yields a biholomorphic parametrization of $\widehat{\mathcal{H}}$ by pairs (q_0, q_1) of monic centered quadratic polynomials. A path in $\widehat{\mathcal{H}}$ converging to p_c can be represented as $(q_0(t), q_1(t))$, where $t \in (0, 1]$, and the latter path converges as $t \to 0$ to some (q, q) since both multipliers or half-multipliers have the same limit. It follows that the path in $\widehat{\mathcal{H}}$ represented by $(q_0(t), q_0(t))$ also converges to p_c. On the other hand, this new path consists of symmetric cubic polynomials. Thus p_c is accessible from $\mathcal{H} = \widehat{\mathcal{H}} \cap \mathcal{SCP}$, as desired.

(2) The claim follows from the assumption that $\mathcal{L}_{c'} = \mathcal{L}_{\mathcal{H}'}$ and Lemma 5.12. □

The following is a combinatorial description of hyperbolic components of $\mathcal{M}_3^{sy}$.

Theorem 6.4. *The map taking a hyperbolic component $\mathcal{H}$ (other than $\mathcal{H}_{main}$) of $\mathcal{M}_3^{sy}$ to the corresponding marked comajor $\overline{\mathrm{co}}^+_{\mathcal{H}}$ is a bijection between the hyperbolic components of $\mathcal{M}_3^{sy}$ (with the exception of $\mathcal{H}_{main}$) and 1-preperiodic comajors of C_sCL. Two distinct marked parabolic polynomials must have distinct marked comajors.*

Proof. Bijection follows from results of A. Poirier [35], which in turn extend earlier work of Bielefeld–Fisher–Hubbard [2]. Namely, injectivity is Theorem 1.1 and surjectivity is Theorem 1.3. More precisely, the map defined in [35] sends $\mathcal{H}$ to the corresponding *critical portrait*. On the other hand, there is a bijection between the marked comajors of C_sCL and symmetric cubic Fatou critical portraits of period > 1.

It remains to prove the last claim. Suppose that two parabolic marked polynomials p_c and $p_{c'}$ are associated with the same marked comajor ℓ. By Theorem, 6.3, c is the root

of a unique hyperbolic domain $\mathcal{H}$ such that $\overline{\mathrm{co}}^+_{\mathcal{H}} = \ell$, and, similarly, $p_{c'}$ is the root of a unique hyperbolic domain $\mathcal{H}'$ such that $\overline{\mathrm{co}}^+_{\mathcal{H}'} = \ell$. By the above, $\mathcal{H} = \mathcal{H}'$. Hence $c = c'$ is the root of $\mathcal{H}$, and $p_c = p_{c'}$ is the same marked polynomial. □

We will now study the place of a parabolic parameter c in the parameter space. More precisely, given a hyperbolic component $\mathcal{H}$ we show that $\mathcal{L}_{\mathcal{H}} = \mathcal{L}_{r_{\mathcal{H}}}$. We also consider parameter rays $\mathcal{R}_\theta$ with 1-preperiodic argument θ, show that they land on a parabolic parameter c, and relate θ and the marked comajor associated with p_c.

We will interchangeably use notation 3θ and $\sigma_3(\theta)$ for any $\theta \in \mathbb{S}^1$. Recall also that τ denotes the rotation of $\overline{\mathbb{D}}$ or $\mathbb{S}^1$ by 180°.

Lemma 6.5. *If $\mathcal{L}' \subset \mathcal{L}$ are two symmetric laminations such that $\mathcal{L}' \neq \varnothing$ is Fatou then $\mathcal{L}$ cannot contain a finite periodic gap G located inside a Fatou gap U' of $\mathcal{L}'$ with an edge which is a major of $\mathcal{L}'$.*

Proof. Suppose that the gap G as described in the lemma exists and $M = \overline{xy}$ is a major of $\mathcal{L}'$ and an edge of G. Consider the first (half-)return map η (see Definition 3.13) of U' with respect to $\mathcal{L}'$. Then $\eta|_{\partial U'}$ is two-to-one. If G exists, it is η-invariant, which is impossible. Indeed, $\eta|_{\partial U'}$ is two-to-one and semiconjugate to σ_2 by the map collapsing the edges of U' to points; by Lemma 3.9 all edges of U' are preimages of its major. Hence the existence of G implies the existence of an invariant leaf or gap of σ_2 with a σ_2-fixed endpoint, which is absurd. □

Lemma 6.6. *Every parameter ray $\mathcal{R}_\theta$ at a 1-preperiodic angle θ lands on a parabolic parameter c_0 on the boundary of $\mathcal{M}_3^{sy}$. Moreover, one of the rays $R_{c_0}(\theta \pm \frac{1}{3})$ lands on a parabolic periodic point of p_{c_0}.*

Proof. We claim that every parameter ray $\mathcal{R}_\theta$ at a 1-preperiodic angle θ lands on a parabolic parameter $c_0 \in \partial\mathcal{M}_3^{sy}$, and the dynamical ray $R_{c_0}(3\theta)$ lands on a p_{c_0}-parabolic point. Indeed, let $c_0 \in \partial\mathcal{M}_3^{sy}$ be in the accumulation set of $\mathcal{R}_\theta$. Recall that for $c \in \mathcal{R}_\theta$, the dynamical external ray $R_c(\theta)$ passes through the cocritical point $-2c$ escaping to infinity

under the action of p_c. Thus, the dynamical external ray $R_c(3\theta)$ has a periodic argument 3θ and, therefore, is not smooth because an eventual σ_3-image of 3θ is the argument of an external ray that terminates at the critical point c. In particular, $R_c(3\theta)$ is not smooth.

On the other hand, the dynamical external ray $R_{c_0}(3\theta)$ of p_{c_0} is periodic and lands on a repelling or parabolic periodic point z_0 of p_{c_0}. Hence $R_{c_0}(\theta)$ lands on a non-periodic point z_1 such that $p_{c_0}(z_1) = z_0$. Since $c_0 \in \mathcal{M}_3^{sy}$, all dynamical external rays of p_{c_0} are smooth. If z_0 is a repelling point of p_{c_0}, then, by Theorem 6.1, the dynamical external ray $R_c(3\theta)$ is smooth and lands on a repelling periodic point for all c close to c_0. However, by the previous paragraph, if $c \in \mathcal{R}_\theta$ then $R_c(3\theta)$ is not smooth. This shows that z_0 is a parabolic point of p_{c_0}.

Since the dynamical ray $R_{c_0}(3\theta)$ lands on z_0, the period of z_0 divides the σ_3-period N of 3θ. Since the multiplier of z_0 with respect to $p_{c_0}^N$ is 1, there are finitely many candidates for an accumulation parameter c_0. Indeed, parabolic parameters c_0, for which there exists an N-periodic point of multiplier 1, form an algebraic subset of $\mathcal{SCP}$; this algebraic set is either finite or the entire plane, the latter option being clearly nonsensical. As the accumulation set of a ray is connected, it consists of exactly one such parabolic parameter. It follows also that one of the dynamical external rays $R_{c_0}(\theta \pm \frac{1}{3})$ lands on a parabolic point of p_{c_0}. □

Recall that if $J(f)$ is locally connected, then there is a well-defined σ_d-invariant lamination $\mathcal{L}_f$.

Definition 6.7 (Repelling leaves)**.** Let f be a degree $d > 1$ monic polynomial. For such f, define *repelling* (or *f-repelling*) leaves of $\mathcal{L}_f$ as leaves corresponding to pairs of rays landing on the same repelling periodic point of f or an iterated preimage thereof. Similarly, we can talk of *parabolic* (or *f-parabolic*) leaves of $\mathcal{L}_f$.

Repelling leaves and related laminations were used in [10] in the proof of continuity of the constructed there monotone (except for one point) model of the entire cubic connectedness locus.

Lemma 6.8. *Suppose that a sequence of monic degree $d > 1$ polynomials f_n converges to a polynomial f (necessarily monic of degree d) as $n \to \infty$. Then the following holds.*

1. *Let the dynamical external rays of f_n with periodic arguments $\theta_1, \dots, \theta_m$ land on the same point (neither m nor the angles θ_i depend on n). If the landing points of $R_f(\theta_1)$, $\dots$, $R_f(\theta_m)$ are repelling, then they coincide.*

2. *If $J(f)$ and all $J(f_n)$ are locally connected, $\mathcal{L}_{f_n} = \mathcal{L}$ for some lamination $\mathcal{L}$, and no critical point of f is mapped to a repelling periodic point, then all repelling leaves of $\mathcal{L}_f$ belong to $\mathcal{L}$.*

Proof. The lemma follows from Theorem 6.1. □

We are going to apply Lemma 6.8 to the situation where f is on the boundary of a hyperbolic domain $\mathcal{H}$ in some parameter space of polynomials and $f_n \in \mathcal{H}$ for each n.

Lemma 6.9. *Suppose that a parameter $c \in \mathcal{SCP} \setminus \mathcal{M}_3^{sy}$ is such that the rays $R_c(\theta \pm \frac{1}{3})$ hit a critical point. Let T be a σ_3-periodic polygon whose iterated forward σ_3-images are disjoint from the critical leaves $\overline{\theta + \frac{1}{3}\, \theta - \frac{1}{3}}$ and $\overline{\theta - \frac{1}{6}\, \theta + \frac{1}{6}}$. Then all dynamical external rays of arguments that are vertices of T land on the same point.*

This statement is not new: see [17]; it can also be deduced, e.g., from a more general Theorem 7.1 of [14], which gives a model for the landing pattern of all external rays of p_c (cf. also [13, Theorem 5.4] for a restatement of this result in the language of invariant laminations). For completeness, we give a sketch in our specific situation.

Sketch of a proof. Consider all external rays in the dynamical plane of p_c whose arguments correspond to vertices of T. All these rays land, perhaps at different points. Let $\hat{T}$ be the union of these rays with some continuum so that $\hat{T}$ is connected and disjoint from the forward orbits of the critical points of p_c. Denote by q the period of T. For every $n = 0, 1, \dots$, let $\hat{T}_n$ be the p_c^{qn}-pullback of $\hat{T}$ that contains the same collection of external rays as $\hat{T}$ and the σ_3^{qn}-pullback of the connecting continuum along the orbit of T.

Since p_c is hyperbolic, the sequence $\hat{T}_n$ converges to a connected set, comprising the original periodic rays and a continuum containing their landing points. However, all points of the limiting continuum must be inside of the Julia set, so this continuum has to be a singleton since the Julia set is totally disconnected. □

We can now prove a key theorem describing the limit transition of laminations in our situation.

Theorem 6.10. *If c is parabolic then the following holds.*

1. *If $c = r_{\mathcal{H}}$ for a hyperbolic component $\mathcal{H} \neq \mathcal{H}_{main}$ then $\mathcal{L}_c = \mathcal{L}_{\mathcal{H}}$ and $\overline{\mathrm{co}}^+_{r_{\mathcal{H}}} = \overline{\mathrm{co}}^+_{\mathcal{H}}$.*
2. *If c is the landing point of a parameter ray $\mathcal{R}_\theta$ and θ is 1-preperiodic, then θ is an endpoint of the marked comajor $\overline{\alpha\beta}$ of $\mathcal{L}_c$.*

Proof. (1) By the main result of [19], there exists a parabolic parameter $c_\circ$ with $\mathcal{L}_{c_\circ} = \mathcal{L}_{\mathcal{H}}$ (the surgery of [19] is local, hence it can be performed in a symmetric fashion to yield a symmetric polynomial $p_\circ$). This parabolic parameter $c_\circ$ is the root point of a certain hyperbolic component $\mathcal{H}_\circ$ such that $\mathcal{L}_{c_\circ} = \mathcal{L}_{\mathcal{H}_\circ}$, by Theorem 6.3. Moreover, $c_\circ$ and $\mathcal{H}_\circ$ have the same marked (co)majors. It remains only to establish that $\mathcal{H} = \mathcal{H}_\circ$, and the latter follows from Theorem 6.4.

Alternatively, remove all parabolic leaves from $\mathcal{L}_c$ to get a new symmetric hyperbolic lamination $\mathcal{L}' \subset \mathcal{L}_c$. By Lemma 6.8, $\mathcal{L}' \subset \mathcal{L}_{\mathcal{H}}$. Hence $\mathcal{L}_c$ is a tuning of $\mathcal{L}'$ done in two steps: (I) consistently add to $\mathcal{L}'$ a cycle (in the B case) or two cycles (in the D case) of finite gaps (of the same period as the corresponding cycles of hyperbolic gaps of $\mathcal{L}'$) with attached hyperbolic gaps; (II) pull this finite collection of gaps back. Moreover, $\mathcal{L}_{\mathcal{H}}$ is obtained similarly. Since by Pommerenke-Levin-Yoccoz inequality the combinatorial rotation number of the inserted cycle(s) of finite gaps in both $\mathcal{L}_c$ and $\mathcal{L}_{\mathcal{H}}$ cases is the same, $\mathcal{L}_c = \mathcal{L}_{\mathcal{H}}$.

(2) Let V^+_c be the closure of the maximal open subset of $\mathbb{D}$ containing U^+_c and disjoint from all repelling leaves of $\mathcal{L}_c$. Since there are no fixed return triangles of σ_3 by Lemma

4.4 of [8], either $V^+ = \mathbb{D}$, or the boundary of V_c^+ consists of a Cantor subset of $\mathbb{S}^1$ and countably many pairwise disjoint leaves of $\mathcal{L}_c$. Clearly, $U_c^+ \subset V_c^+$. We claim that if $V_c^+ \neq U_c^+$ then there exists a gap $G \subset V_c^+$ such that all images of U_c^+ inside V_c^+ share an edge with G. Indeed, suppose otherwise. Then it follows from [23] that there are repelling cutpoints of $J(p_c)$ such that the convex hulls of the arguments of dynamical external rays of p_c landing on them separate V_c^+, a contradiction with the definition of V_c^+. This proves the existence of G with the listed properties. Moreover, it follows that the edges of G are the periodic leaves of $\mathcal{L}_c$ inside V_c^+.

By Lemma 6.8, the angles $\theta \pm \frac{1}{3}$ are vertices of V_c^+. By Lemma 6.6, one of the rays $R_c(\theta \pm \frac{1}{3})$ (say, $R_c(\theta + \frac{1}{3})$) lands on a parabolic periodic point z_c of p_c. By way of contradiction, suppose that $\theta \notin \{\alpha, \beta\}$. Set $\overline{AB}$ to be the marked major of $\mathcal{L}_c$; by assumption $\theta + \frac{1}{3} \notin \{A, B\}$. Then $U_c^+ \neq V_c^+$ and we can consider the gap G defined in the previous paragraph. Since at least two rays land on each parabolic point of p_c, the dynamical external rays whose arguments are vertices of G land on z_c. Thus, the three angles A, B and $\theta + \frac{1}{3}$ are vertices of G and the dynamical external rays of arguments A, B and $\theta + \frac{1}{3}$ land on z_c. Note that the major edge of G coincides with the marked major $\overline{AB}$ of $\mathcal{L}_c$.

Similar to the construction of the degree two first (half-)return map η (see Definition 3.13), define a degree two map $\eta' : V_c^+ \to V_c^+$ semiconjugate to σ_2 by collapsing edges of V_c^+ to points. Under this semiconjugacy G maps to a finite σ_2-invariant gap (or leaf) G', and the critical leaf $\overline{\theta - \frac{1}{3}, \theta + \frac{1}{3}}$ projects to a critical leaf ℓ in $\overline{\mathbb{D}}$ with a σ_2-periodic endpoint which is a vertex of G' but not an endpoint of the Thurston major of G' (see [42]). Therefore there exists a finite σ_2-invariant gap T' disjoint from ℓ. Then the lifting of T' gives a finite η-invariant gap $T \subset V_c^+$. Note that G and T have disjoint sets of vertices.

Now consider a point c_t on the parameter ray $\mathcal{R}_\theta$, where the parameter t corresponds to the value of the Green function for $\mathcal{M}_3^{sy}$ at c_t so that c_t converges to c as $t \to 0$. Then the rays $R_{c_t}(\theta \pm \frac{1}{3})$ both hit the marked critical point c_t. By Lemma 6.9, the dynamical external rays of p_{c_t} whose arguments correspond to the vertices of T land on the same repelling periodic point w_t. The point w_t has a well-defined limit w_0 as $t \to 0$. On the other hand, for

every angle γ corresponding to a vertex of T, the ray $R_c(\gamma)$ lands on a repelling periodic point of p_c. By Theorem 6.1 and by Lemma 6.8 this point is close to w_t for small t, hence it must coincide with w_0. Thus, w_0 is repelling for p_c, and the gap or leaf T corresponding to w_0 must belong to $\mathcal{L}_c$, a contradiction. □

Observe that, by Lemma 6.6, there is a dense set of 1-preperiodic angles such that the corresponding parameter rays land on parabolic parameters in $\mathcal{M}_3^{sy}$.

Theorem 6.11. *Let $\overline{\alpha\beta}$ be the marked comajor of a symmetric Fatou lamination $\mathcal{L}$. Then there exists a unique hyperbolic component $\mathcal{H} \neq \mathcal{H}_{main}$ such that the parameter rays $\mathcal{R}_\alpha$, $\mathcal{R}_\beta$ land on the root point $r_\mathcal{H}$ and $\mathcal{L}_{r_\mathcal{H}} = \mathcal{L}_\mathcal{H} = \mathcal{L}$. Moreover, let $c \in \partial\mathcal{H}, c \neq r_\mathcal{H}$ be a parabolic parameter. Then $\mathcal{L}_c \supsetneqq \mathcal{L}_\mathcal{H}$, and the marked comajor of $\mathcal{L}_c$ is located under the marked comajor of $\mathcal{L}_\mathcal{H}$.*

Proof. Consider the marked comajor $\overline{\alpha\beta}$ of a symmetric Fatou lamination $\mathcal{L}$. Then α and β are 1-preperiodic. By Lemma 6.6, the parameter ray $\mathcal{R}_\alpha$ lands on a parabolic parameter c. The angle α is an endpoint of the marked comajor $\overline{\mathrm{co}}_c^+$ of the lamination $\mathcal{L}_c$, by Theorem 6.10. Since distinct 1-preperiodic comajors are disjoint, $\overline{\mathrm{co}}_c^+ = \overline{\alpha\beta}$. By Theorem 6.3, there exists a hyperbolic component $\mathcal{H}$ such that $c = r_\mathcal{H}$ is the root of $\mathcal{H}$. By Theorem 6.10, the lamination $\mathcal{L}_\mathcal{H}$ coincides with $\mathcal{L}_c$. By Theorem 6.4, the set $\mathcal{H}$ is a unique hyperbolic component such that $\overline{\mathrm{co}}_\mathcal{H}^+ = \overline{\alpha\beta}$. It follows that $\mathcal{R}_\beta$ lands on c, too. The fact that $\mathcal{L}_{r_\mathcal{H}} = \mathcal{L}_\mathcal{H} = \mathcal{L}$ follows from Theorem 6.10.

To prove the last claim of the theorem, let $c \in \partial\mathcal{H}, c \neq r_\mathcal{H}$, be a parabolic parameter. Since $c \neq r_\mathcal{H}$, the repelling periodic points of polynomials $p_{c^*} \in \mathcal{H}$ associated with the marked major $M_\mathcal{H}$ of $\mathcal{L}_\mathcal{H}$ converge to a repelling periodic point of p_c of the same period as $c^* \to c$. Hence all the edges of the gaps $U_\mathcal{H}^\pm$ of $\mathcal{L}_\mathcal{H} = \mathcal{L}_{c^*}$ remain edges of $\mathcal{L}_c$. Thus, the critical gaps of $\mathcal{L}_c$ are contained in $U_\mathcal{H}^\pm$, and, hence, $\mathcal{L}_c \supset \mathcal{L}_\mathcal{H}$, and the comajors of $\mathcal{L}_c$ are located under those of $\mathcal{L}_\mathcal{H}$ (this yields the claim of the theorem about the marked comajors). The hyperbolic component with the same marked comajor as $\mathcal{L}_c$ cannot coincide with $\mathcal{H}$, since c is not a root point of $\mathcal{H}$. Hence, $\mathcal{L}_c \neq \mathcal{L}_\mathcal{H}$. □

Call the parameter rays from Theorem 6.11 *characteristic rays of a hyperbolic component* $\mathcal{H} \neq \mathcal{H}_{main}$. Recall that by an arc $(a,b) \subset \mathbb{S}^1$ we always mean the *positively oriented circle arc* with endpoints $a, b \in \mathbb{S}^1$.

Lemma 6.12. *The characteristic rays are the only two strictly preperiodic rays that accumulate on a parabolic parameter c.*

Proof. By Theorems 6.3 and 6.11, c is the landing point of the parameter rays $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ where $\overline{\alpha\beta}$ is a 1-preperiodic comajor. Thus, all 1-preperiodic comajors give rise to cuts in the parameter plane. Hence, if a comajor separates $\overline{\alpha\beta}$ from an angle γ in $\overline{\mathbb{D}}$, then $\mathcal{R}_\gamma$ cannot accumulate on c. Since, by Theorem 3.7, the 1-preperiodic comajors are dense in C_sCL and disjoint from all other comajors, it follows that the only way a parameter ray $\mathcal{R}_\gamma$ can accumulate on c is when γ is a vertex of an infinite gap G with an edge $\overline{\alpha\beta}$. However, by Theorem 3.15, the fact that γ is preperiodic implies that γ is actually 1-preperiodic. Again by Theorem 6.3, this implies that $\mathcal{R}_\gamma$ cannot land on c, as desired. □

Theorem 6.13. *Each parabolic parameter $c \in \partial\mathcal{H}_{main}$ is associated to a comajor $\overline{\alpha\beta}$ which is an edge of G_{main} and, accordingly, to a point of $\overline{\mathbb{D}}/C_sCL$. The parameter rays $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ land on c. For every hyperbolic domain $\mathcal{H} \neq \mathcal{H}_{main}$ the corresponding marked comajor $\overline{\mathrm{co}}^+_{\mathcal{H}} = \overline{\theta_1\theta_2}$ is associated to the parameter rays $\mathcal{R}_{\theta_1}$ and $\mathcal{R}_{\theta_2}$ that land on the root $r_{\mathcal{H}}$ of $\mathcal{H}$ and separate $\mathcal{H}$ from $\mathcal{H}_{main}$.*

Proof. Let $c \in \partial\mathcal{H}_{main}$ be a parabolic parameter. Then 0 is a parabolic point associated to a finite invariant gap T whose vertices are the arguments of dynamical external rays of p_c landing on 0. It follows that there is a unique symmetric lamination $\mathcal{L}$ associated with T which has the gap T, Fatou gaps of degree greater than 1 attached to T and "rotating" around T, and pullbacks of all these gaps (this fully describes $\mathcal{L}$). Let $\overline{\alpha\beta}$ be the marked comajor of $\mathcal{L}$ associated with the marked cocritical point $-2c$ of p_c; by Theorem 6.11, the parameter rays $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ land on c. By Theorem 3.12, all parabolic parameters $c \in \partial\mathcal{H}_{main}$ correspond to edges of the main gap G_{main} of C_sCL and in the end map to the

corresponding points of $\overline{\mathbb{D}}/C_sCL$ that belong to the main domain D_{main} of $\overline{\mathbb{D}}/C_sCL$. The rest of the theorem easily follows. □

7 Misiurewicz parameters

A number $c \in \mathbb{C}$ is a *Misiurewicz parameter* if the p_c-orbits of critical values are strictly preperiodic. If c is a Misiurewicz parameter, then K_c is connected ($c \in \mathcal{M}_3^{sy}$), all p_c-periodic points are repelling (see Theorem 5.4), and $J(p_c)$ is a dendrite (recall that a dendrite is a locally connected continuum that contains no Jordan curves). Recall that a cubic symmetric lamination $\mathcal{L}$ is called a *Misiurewicz* lamination if its critical sets are strictly preperiodic. Such laminations and their comajors are discussed in detail right after Lemma 3.6. By Theorem 3.7(2), Misiurewicz cocritical leaves (gaps) are leaves (gaps) of C_sCL approached from all sides by 1-preperiodic comajors. Since for each polynomial from $\mathcal{SCP}$ a critical point is marked, then each Misiurewicz lamination is considered twice, with either cocritical set marked. Finally, recall that the lamination C_sCL defines a laminational equivalence relation $\sim_{sy}$. For brevity by "$\sim_{sy}$-class" we will mean an equivalence class of $\sim_{sy}$.

Lemma 7.1. *Let p_c be a cubic symmetric polynomial with a dendritic Julia set, and T be the marked critical set of $\mathcal{L}_c$. If $c_n \in \mathcal{SCP} \setminus \mathcal{M}_3^{sy}$ converge to c as $n \to \infty$, then the arguments of the external rays of p_{c_n} hitting the marked critical point c_n converge to vertices of T.*

Proof. Every leaf of $\mathcal{L}_c$ that is not an edge of a gap is approximated by (pre)periodic leaves from both sides, and every edge of a gap G in $\mathcal{L}_c$ is approximated by preperiodic leaves from outside of G. Choose a neighborhood W of T in $\mathbb{D}$ whose boundary is formed by (pre)periodic leaves close to the edges of T, and appropriate circle arcs. By Theorem 6.1, there is a neighborhood $\mathcal{W}$ of c in $\mathcal{SCP}$ such that for $c^* \in \mathcal{W}$, the leaves forming the boundary of W in $\mathbb{D}$ are associated with cuts formed by the dynamical external rays of p_{c^*} (it does not matter whether $J(p_{c^*})$ is connected or not). The part of the dynamical plane of p_{c^*} bounded by all these cuts contains the point c^*. This implies the desired. □

The following theorem is a special case of [26, Theorem 1] but we give a proof for completeness, and also since our special case is much simpler than the general one.

Theorem 7.2. *Let $\mathcal{L}$ be a Misiurewicz lamination. Parameter rays whose arguments are vertices of the marked cocritical set of $\mathcal{L}$ land on a Misiurewicz parameter $\hat{c}$ such that $\mathcal{L}_{\hat{c}} = \mathcal{L}$.*

Proof. Let θ be a k-preperiodic angle with $k > 1$. Choose $c_0 \in \partial \mathcal{M}_3^{sy}$ in the accumulation set of $\mathcal{R}_\theta$. Then the periodic dynamical external ray $R_{c_0}(3^k\theta)$ lands on a periodic point z_0. By Lemma 6.12, the point z_0 is repelling. For a parameter c, consider the union R_c of the rays from the forward orbit of the closure of $R_c(\theta)$; it consists of finitely many rays and their landing points. By Lemma 7.1, there are no critical points among their landing points provided that c is close to c_0.

By Theorem 6.1, for some neighborhood $\mathcal{W}$ of c_0 in $\mathcal{SCP}$, the set R_c depends continuously on $c \in \mathcal{W}$, consists only of smooth rays and their landing points, and contains no critical points of p_c. In particular, $R_{c_0}(\theta)$ lands on the cocritical point $-2c_0$. Thus, p_{c_0} is a symmetric Misiurewicz polynomial, and $\mathcal{L}_{c_0}$ is a symmetric Misiurewicz lamination. Since there are countably many symmetric Misiurewicz polynomials, and the accumulation set of $\mathcal{R}_\theta$ is either a non-degenerate continuum (hence uncountable) or a point, it follows that $\mathcal{R}_\theta$ lands on c_0 and that θ is a vertex of a cocritical set of $\mathcal{L}_{c_0}$.

By Theorem 3.7, all preperiodic angles of preperiod > 1 are partitioned into vertex sets of various Misiurewicz cocritical sets. Thus, for a preperiod > 1 angle θ there exists a unique symmetric lamination $\mathcal{L}$ such that θ is a vertex of a cocritical set of $\mathcal{L}$. Taking into account the fact that each polynomial is counted in $\mathcal{SCP}$ twice (depending on the choice of the marked critical point), and choosing marked cocritical sets accordingly, we see that if $\mathcal{L}$ is a Misiurewicz lamination whose marked cocritical set has vertices $\theta_1, \dots, \theta_m$, then the parameter rays $\mathcal{R}_{\theta_1}, \dots, \mathcal{R}_{\theta_m}$ land on a Misiurewicz parameter $\hat{c}$ such that $\mathcal{L}_{\hat{c}} = \mathcal{L}$. □

8 The Structure of $\mathcal{M}_3^{sy}$

Recall that $\mathcal{M}_{3,comb}^{sy}$ is the quotient space of $\overline{\mathbb{D}}$ under $\sim_{sy}$ and $\mathfrak{pr}_{comb} : \overline{\mathbb{D}} \to \mathcal{M}_{3,comb}^{sy}$ is the corresponding quotient projection (see Definition 3.8). Given a continuum $K \subset \mathbb{C}$, define the *topological hull of* K as the complement to the unbounded complementary component of K. Recall also that a *monotone* map is a map whose point-preimages (fibers) are connected. Theorem 8.4 completes the proof of the Main Theorem.

However first we introduce tools from continuum theory developed in [3] (for basic continuum theory facts see, e.g., [36]) and applied in [3] to the problem of modeling connected polynomial Julia sets. The tools apply to all planar continua.

Let A be a continuum. Then an onto map $\varphi : A \to Y_{\varphi,A}$ is said to be a *finest (monotone) map (onto a locally connected continuum)* if for any other monotone map $\psi : A \to L$ onto a locally connected continuum L there exists a monotone map $h : Y_{\varphi,A} \to L$ such that $\psi = h \circ \varphi$. Observe, that in this situation the map h is automatically monotone because for $x \in L$ we have $h^{-1}(x) = \varphi(\psi^{-1}(x))$. It is easy to see that all sets $Y_{\varphi,A}$ are homeomorphic and all finest maps φ are the same up to a homeomorphism. Thus from now on we may talk of ***the*** *finest model* $Y_A = Y$ *of* A and ***the*** *finest map* $\varphi_A = \varphi$ *of* A *onto* Y.

A planar continuum $Q \subset \mathbb{C}$ is said to be *unshielded* if it coincides with the boundary of its topological hull. Thus, unshielded continua are the boundaries of *full* planar continua.

Theorem 8.1 (Theorem 1 of [3])**.** *Let Q be an unshielded continuum. Then there exist the finest map φ and the finest model Y of Q. Moreover, φ can be extended to a map $\overline{\mathbb{C}} \to \overline{\mathbb{C}}$ which maps ∞ to ∞, in $\overline{\mathbb{C}} \setminus Q$ collapses only those complementary domains to Q whose boundaries are collapsed by φ, and is a homeomorphism elsewhere in $\overline{\mathbb{C}} \setminus Q$.*

It may happen that the finest model is a point (e.g., this is so if the continuum is *indecomposable*, i.e. cannot be represented as the union of two non-trivial subcontinua). In [3] we establish useful sufficient conditions for this not to be the case, then apply Theorem 8.1 to a polynomial P with connected Julia set, and explicitly construct the finest models (see Theorem 8.3 below). Moreover, in [3, Theorem 2] we show that for a

connected polynomial Julia set $J(P)$ the model is *dynamical*, i.e. admitting a self-mapping to which $P|_{J(P)}$ is semiconjugate. The self-mapping in question is actually a topological polynomial. This gives an alternative proof of results of [25] and extends them onto all connected polynomial Julia sets. Denote by $I_\theta(Q)$ the *impression* of the Riemann ray $R_Q(\theta)$ to an unshielded continuum Q which by definition is the same as the impression of the corresponding prime end (abusing the terminology we will call them *impressions of angles* θ).

Definition 8.2. Let Q be an unshielded planar continuum. Declare two points $x, y \in Q$ equivalent if they belong to a finite connected union of impressions of angles, and denote this equivalence relation on Q by $\asymp_Q$. Then consider the intersection $\approx_Q$ of all closed equivalence relations on Q that contain $\asymp_Q$. Declare two angles $\alpha, \beta \in \mathbb{S}^1$ equivalent if their impressions are contained in one $\approx_Q$-class, and denote this equivalence relation on $\mathbb{S}^1$ by $\sim_Q$.

Definition 8.2 offers a constructive version of Theorem 8.1. Moreover, it also suggests a laminational interpretation of the finest model of Q (the second part of Theorem 8.3 is not explicitly proven in [3] but immediately follows and can be established repeating verbatim a part of the proof of [3, Theorem 2]).

Theorem 8.3 (Theorem 18 of [3])**.** *The quotient map $Q \to Q/\approx_Q$ is the finest map of the continuum Q. The equivalence relation $\sim_Q$ is laminational, and the finest model $Q/\approx_Q$ of Q is homeomorphic to $\mathbb{S}^1/\sim_Q$.*

We can now prove Theorem 8.4.

Theorem 8.4. *There is a monotone continuous surjective map $\pi : \mathcal{M}_3^{sy} \to \mathcal{M}_{3,comb}^{sy}$; if $\mathcal{M}_3^{sy}$ is locally connected, π is a homeomorphism.*

Proof. Set $\approx_{\mathcal{M}_3^{sy}} = \approx$ and $\sim_{\mathcal{M}_3^{sy}} = \sim$. By Theorem 8.3, the spaces $\mathcal{M}_3^{sy}/\approx$ and $\mathbb{S}^1/\sim$ are homeomorphic. Thus, it suffices to show that equivalence relations $\sim$ and $\sim_{sy}$ on $\mathbb{S}^1$ coincide. By Theorem 3.7 and the results in Sections 6 and 7, if $\mathbf{g}$ is a $\sim_{sy}$-class whose convex hull does

not intersect an infinite gap of C_sCL, then $\mathbf{g}$ is in fact a $\sim$-class. It remains to consider $\sim_{sy}$-classes $\mathbf{h}$ whose convex hulls intersect infinite gaps of C_sCL. In what follows we use notation and terminology from right before Theorem 3.15. In particular, recall that $\mathcal{O}$ is the center of $\overline{\mathbb{D}}$.

Let $\mathbf{h}$ be a $\sim_{sy}$-class such that $\mathrm{CH}(\mathbf{h})$ intersects ∂G, where G is an infinite gap of C_sCL. If $\mathcal{O} \notin G$, then, by Theorems 3.7 and 3.15, there is an associated with G Fatou lamination $\mathcal{L}$ with critical Fatou gap U and the (half-)return map $\eta : \partial U \to \partial U$ semiconjugate to $\sigma_2 : \mathbb{S}^1 \to \mathbb{S}^1$ by a map ϕ collapsing all edges of U. We have $G \subset U$, and the ϕ-images of the edges of G are the majors of laminations from the Main Cardioid of σ_2.

Let ℓ be the edge of G separating the rest of G from $\mathcal{O}$. By the properties of the Main Cardioid, an infinite gap of C_sCL is attached to G at any edge $\ell' \neq \ell$ of G while the points of ∂G that are not endpoints of edges of G are C_sCL-classes. By Theorem 3.7, ℓ may be an edge of another infinite gap of C_sCL or it may be non-isolated in C_sCL.

By Theorem 6.4, there is a unique hyperbolic component $\mathcal{H}$ with $\overline{\mathrm{co}}^+_{\mathcal{H}} = \ell$. By Theorem 6.11, a dense subset of $\partial\mathcal{H}$ consists of parabolic parameters c with two parameter rays corresponding to the endpoints of an edge of G landing on $c \in \partial\mathcal{H}$. Properties of impressions and the fact that $\partial\mathcal{H}$ is a Jordan curve imply that for all edges of G other than ℓ the corresponding $\sim$-class and $\sim_{sy}$-class coincide.

The situation with ℓ is different and has three cases. Firstly, if ℓ is an edge of another infinite gap of C_sCL not containing $\mathcal{O}$, then the above arguments show that the $\sim$-class associated with ℓ consists only of the endpoints of ℓ. Secondly, ℓ may be the limit of other edges of C_sCL converging to ℓ from outside of G. Let Γ_ℓ be the parameter cut corresponding to ℓ. By Theorem 3.7, Γ_ℓ is approximated by 1-preperiodic parabolic cuts separating Γ_ℓ from impressions of other parameter rays in the complementary component of Γ_ℓ not containing $\mathcal{H}$; thus, in this case, too, the $\sim$-class and the corresponding $\sim_{sy}$-class coincide.

The remaining case is when ℓ is the shared edge of G and the gap G_{main} of C_sCL corresponding to $\mathcal{H}_{main}$. Recall that $\partial\mathcal{H}_{main}$ is the circle of radius $\frac{\sqrt{3}}{3}$ centered at the origin.

For $c \in \partial\mathcal{H}_{main}$, the multiplier at the neutral point 0 is $-3c^2$. Hence there is a dense in $\partial\mathcal{H}_{main}$ set of parabolic parameters c associated to the comajors like ℓ above. The above arguments show that for all edges of G_{main} and for all points of ∂G_{main} that are not endpoints of an edge of G_{main} the same conclusion holds: the $\sim$-classes and the $\sim_{sy}$-classes are the same. □

From now on, we will always denote the modeling projection from Theorem 8.4 by π.

Definition 8.5. Let $\mathcal{H}$ be a hyperbolic component associated with a comajor $\overline{\alpha\beta}$. Then the parameter rays $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ land on the root $r_\mathcal{H}$ of $\mathcal{H}$. Denote by $\mathcal{W}(\overline{\alpha\beta})$ the component of $\mathbb{C} \setminus [\mathcal{R}_\alpha \cup \mathcal{R}_\beta]$ that does not contain $\mathcal{H}_{main}$ and will call this set the *wake generated by* $\overline{\alpha\beta}$.

The next lemma describes wakes in a more dynamical fashion.

Lemma 8.6. *Let $\overline{\alpha\beta}$ be a comajor corresponding to a major $\overline{\alpha'\beta'}$.*

1. *If $c \in \mathcal{W}(\overline{\alpha\beta})$, then the external rays $R_c(\alpha')$ and $R_c(\beta')$ are smooth and land at the same repelling periodic point.*

2. *If $R_c(\alpha')$ and $R_c(\beta')$ are smooth and land at the same repelling periodic point, then $c \in \mathcal{W}(\overline{\alpha\beta}) \cup \mathcal{W}(\tau(\overline{\alpha\beta}))$.*

Proof. (1) The argument given below follows [31]. Set $\mathcal{W} := \mathcal{W}(\overline{\alpha\beta})$, and let q be the period of α and β. The claim holds for the interior of the corresponding hyperbolic domain $\mathcal{H}$. By the properties of C_sCL, parabolic parameters on the boundary of $\mathcal{H}$, and hence the entire $\mathcal{H}$, are separated from $\mathcal{H}_{main}$ by the parameter rays $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ that land on the root $r_\mathcal{H}$ of $\mathcal{H}$. Thus, $\mathcal{H} \subset \mathcal{W}$, and, by Theorem 6.11, for all parameters $c \in \mathcal{H}$ the dynamical external rays $R_c(\alpha')$ and $R_c(\beta')$ land on the same repelling periodic point.

We claim that the rays $R_c(\alpha')$ and $R_c(\beta')$ land for all $c \in \mathcal{W}$. Indeed, by the properties of comajors the circle arc $(\sigma_3^2(\alpha'), \sigma_3^2(\beta'))$ contains no images of $\sigma_3(\alpha')$ or $\sigma_3(\beta')$. Hence points from the orbits of $\sigma_3(\alpha')$ and $\sigma_3(\beta')$ cannot be endpoints of critical chords with the other endpoint in (α', β'). However, these are precisely the critical chords defining the critical

cuts for polynomials $p_c, c \in \mathcal{W}$. It follows that for all $c \in \mathcal{W}$ the rays $R_c(\alpha')$ and $R_c(\beta')$ remain smooth (never pass through a cut) and land as claimed.

The landing point of $R_c(\alpha')$ is repelling except for finitely many values of $c \in \mathcal{W}$, for which it may become parabolic; the multiplier of this parabolic point must be a q-th root of unity. By the maximum modulus principle, such values of c are impossible. It follows that $R_c(\alpha')$ always lands on a repelling point, for $c \in \mathcal{W}$, and the same holds for $R_c(\beta')$. Since the two landing points coincide in $\mathcal{H}$, they also coincide everywhere in $\mathcal{W}$, by Theorem 6.1.

(2) Suppose that $R_c(\alpha')$ and $R_c(\beta')$ are smooth and land at the same repelling periodic point. By Theorem 6.1, this property is stable under small perturbations of c. In particular, if the Julia set of p_c is connected, then we may replace c with a postcritically finite value c' with the same property. Also, if $J(p_c)$ is disconnected, then, upon a slight perturbation of c, one may assume that c lies on a rational ray. Let c' be the landing point of this ray, and replace c' with a nearby postcritically finite parameter c''. Thus, in any case, it is safe to assume that c is postcritically finite, in particular, $J(p_c)$ is locally connected, and the corresponding lamination $\mathcal{L}_c$ determines the topological dynamics of p_c on $J(p_c)$. All periodic leaves of $\mathcal{L}_c$ are repelling.

By the assumption, $\overline{\alpha'\beta'} \in \mathcal{L}_c$. Let S' be the strip formed by $\overline{\alpha'\beta'}$ and its sibling chord so that $S' \cup \tau(S')$ is a short strips set. Some major M_c of $\mathcal{L}_c$ must be contained in S'; it follows that a comajor of $\mathcal{L}_c$ is under $\overline{\alpha\beta}$. Either this is the marked comajor of $\mathcal{L}_c$, in which case $c \in \mathcal{W}(\overline{\alpha\beta})$, or the symmetric one, in which case $c \in \mathcal{W}(\tau(\overline{\alpha\beta}))$. □

Recall that, by Theorem 3.7(3) (i.e., by Theorem 4.15 of [9]), the 1-preperiodic comajors are dense in C_sCL. Using this, we will now relate the dynamics of certain symmetric cubic polynomials with their location in the parameter space. To do this, we will need the original results of Kiwi [25] that we have already mentioned in the beginning of this section. As far as we know, these were the first results where connected but not necessarily locally connected Julia sets of certain polynomials were modeled. The polynomials in question are polynomials with connected Julia set and all cycles repelling. However,

when stating Kiwi's theorem, we use the approach of [3] described in the beginning of this section.

Theorem 8.7 (Theorem 5.12 of [25]). *Let P be a complex polynomial with connected Julia set $J(P)$ and all cycles repelling. Then $\sim_{J_P}$ is a σ_d-invariant laminational equivalence relation defining a q-lamination $\mathcal{L}_{\sim_P}$. Moreover, $P : J(P) \to J(P)$ is monotonically semiconjugate to a topological polynomial $f_{\sim_P} : J_{\sim_P} \to J_{\sim_P}$, the topological Julia set $J_{\sim_P}$ is a dendrite, and the semiconjugacy p_P between $P|_{J(P)}$ and $f_{\sim_P}|_{J_{\sim_P}}$ is one-to-one on all (pre)periodic points of P. If $J(P)$ is locally connected, p_P is a homeomorphism.*

Let P be a symmetric polynomial with connected Julia set and all cycles repelling. Theorem 8.7 allows one to define, for such P, its q-lamination $\mathcal{L}_{\sim_P}$ and, therefore, the associated marked cocritical set C_P. Let us show that then C_P determines the fiber of the modeling projection π from Theorem 8.4 that contains P. To this end, we need a well-known fact concerning the dynamics of dendritic topological Julia sets $J_\sim$ defined by a laminational equivalence relation $\sim$. Still, for the sake of completeness we sketch its proof.

Lemma 8.8. *Let $\sim$ be a σ_d-invariant laminational equivalence relation such that $J_\sim$ is a dendrite. Suppose that $\overline{ab}$ is a periodic chord whose forward σ_d-orbit consists of pairwise unlinked chords that do not cross edges of critical sets of $\mathcal{L}_\sim$. Then $a \sim b$.*

Sketch of the proof. Consider preimages of critical gaps or leaves of $\mathcal{L}_\sim$; choose only those preimages that are themselves gaps of leaves of $\mathcal{L}_\sim$. Take the closure of their union. By the assumptions made, there is a unique component U of the complement to this union that contains $\overline{ab}$ in its closure. The set of all classes of points from $\overline{U} \cap \mathbb{S}^1$ corresponds to a non-degenerate continuum $T \subset J_\sim$ whose orbit, by the assumption, has no iterated images that contain critical points of $f_\sim$ as their cutpoints. On the other hand, by Theorem C of [5], the continuum T is non-wandering and has two distinct iterated images that intersect. The union of the appropriate iterated images of T will then yield connected n-periodic subset of $J_\sim$ whose closure A contains no critical cutpoints. This implies that there are periodic attracting points of $f_\sim^n$ in A, which contradicts the expanding properties of σ_d. □

We can now describe some *fibers*, i.e., point preimages, of π.

Theorem 8.9. *If all cycles of $P \in \mathcal{M}_3^{sy}$ are repelling, then the π-fiber of C_P contains P. Moreover, for any P' in the same fiber, $C_{P'} = C_P$ and therefore $\mathcal{L}_{\sim_{P'}} = \mathcal{L}_{\sim_P}$. If, in addition, P is finitely renormalizable, then the fiber is $\{P\}$.*

Proof. By Theorem 6.13, we may assume that C_P is not a 1-preperiodic comajor. Suppose now that C_P is a finite gap. Then, by Theorem 3.7(3), all edges of C_P are approximated from outside of C_P arbitrarily well by 1-preperiodic comajors that give rise to the associated wakes in the parameter space. Let us consider the edge ℓ of C_P that separates the interior of C_P from the center of the circle. Choose a 1-preperiodic comajor $\overline{y}$ close to ℓ that also separates C_P from the center of the circle; it corresponds to a periodic major ℓ_y, and, by the properties of majors, the iterated σ_3-images of ℓ_y never enter a critical strip defined by ℓ_y or its symmetric counterpart. Thus, ℓ_y satisfies the assumptions of Lemma 8.8. By Lemma 8.8, this implies that the endpoints of ℓ_y are $\sim_P$-equivalent and, hence, are associated with a repelling cut Y of $J(P)$. Therefore, by Lemma 8.6, the parameter of P is located in the wake $\mathcal{W}_{\overline{y}}$ for any $\overline{y}$ with the listed properties.

On the other hand, let $\ell' \neq \ell$ be another edge of C_P. We can approximate it (again by Theorem 3.7(3)) by 1-preperiodic comajors, necessarily from the outside of C_P. These define wakes as before, yet this time the corresponding periodic leaves will not coexist with the critical sets of $\mathcal{L}_P$. Hence, P cannot belong to those wakes, which implies that it belongs to their complements in the plane.

By definition, the intersection of the wakes (described in the first paragraph of the proof), the complements of wakes (described in the second paragraph of the proof), and $\mathcal{M}_3^{sy}$ is the fiber of the modeling map π from Theorem 8.4; this completes the proof of the first claim in the case when C_P is a gap. If C_P is a leaf, a similar (almost verbatim) argument implies the same conclusion. Observe that in this case Theorem 3.7 implies that C_P is approximated by 1-preperiodic comajors from both sides.

The case when $C_P = \{x\}$ is a singleton is a bit different. Recall, that $J_{\sim_P}$ is a dendrite. Denote by $\zeta(X)$ the point in the topological Julia set $J_{\sim_P}$ that corresponds to the convex

hull X of a $\sim_P$-class. By [8, Lemma 3.3], there is a well defined invariant central (i.e. containing the center of the circle) gap (or leaf) $\mathrm{CG}_{\mathcal{L}_{\sim_P}}$ of $\mathcal{L}_\sim$. Let $a = \zeta(\mathrm{CG}_{\mathcal{L}_{\sim_P}})$. Connect a and $\zeta(\{x\})$ with an arc I and consider dynamics of points of I that are located close to $\zeta(\{x\})$. Let $y \in I, y = \zeta(Y)$ be a point close to $\zeta(\{x\})$. Denote by Z_y the component of $J_{\sim_P} \setminus \{y\}$ that contains $\zeta(\{x\})$. Then by Lemma 3.8 (the so-called Short Strips Lemma) of [8] translated into the language of dynamics on $J_{\sim_P}$ we see that the forward orbit of y either never enters Z_y, or, if it does, the first time it enters Z_y is in I again. In the former case it follows from the definitions that the edge ℓ'' of Y that separates the rest of Y from x (i.e., the edge of Y that "faces" x) is a comajor. By Theorem 3.7(3) this implies that there are 1-preperiodic comajors very close to ℓ'' and the arguments from the beginning of the proof apply in this case, too.

Hence we may assume that orbits of all points $y \in I$ close to $\zeta(\{x\})$ always eventually enter their sets Z_y. However there are (pre)periodic points in I arbitrarily close to $\zeta(\{x\})$, and these cannot keep getting mapped to points of I closer and closer to $\zeta(\{x\})$. Thus, there are comajors associated to points of I arbitrarily close to $\zeta(\{x\})$ which implies the desired.

By the above, for every polynomial P' in the fiber of C_P, the associated lamination $\mathcal{L}_{\sim P'}$ has C_P as a marked comajor. It follows that $\mathcal{L}_{\sim P} = \mathcal{L}_{\sim P'}$ as critically marked laminations, cf. Corollary 3.11. To prove the last claim of the lemma, observe that in the case of finitely renormalizable symmetric polynomials the claim follows from [27, Theorem 1.2] stating, in a special case, that, if $\mathcal{L}_{\sim_P} = \mathcal{L}_{\sim_{P'}}$ is at most finitely many times renormalizable, and both P and P' have all cycles repelling, then $P = P'$, up to affine conjugacy. □

To conclude we would like to say that we expect the following claims to hold for the fibers of π:

1. On the union of the boundaries of all hyperbolic components, π is injective.

2. Every nontrivial fiber contains a polynomial whose Julia set has positive area and supports an invariant measurable line field; there is a J-stable component inside

the fiber consisting of such polynomials.

However, we do not prove these claims here.

Acknoledgements

The first named author was partially supported by NSF grant DMS-2349942. The second named author was partially supported by NSF grant DMS–1807558. The work of the fourth named author was supported by the HSE University Basic Research Program.

References

[1] S. Bhattacharya, A. Blokh, D. Schleicher, *Unicritical Laminations*, Fundamenta Mathematicae, **258** (2022), 25–63. 61

[2] B. Bielefeld, Y.Fisher, J. Hubbard, *The Classification of Critically Preperiodic Polynomials as Dynamical Systems*, Journal of the AMS **5** (1992), 721–762. 87

[3] A. Blokh, C. Curry, L. Oversteegen, *Locally connected models for Julia sets*, Advances in Mathematics, **226** (2011), 1621–1661. 97, 98, 102

[4] A. Blokh, D. Childers, G. Levin, L. Oversteegen, D. Schleicher, *An Extended Fatou-Shishikura inequality and wandering branch continua for polynomials*, Advances in Mathematics, **288** (2016), 1121–1174. 79

[5] A. Blokh, G. Levin, *Growing trees, laminations and the dynamics on the Julia set*, Ergodic Theory and Dynamical Systems **22** (2002), 63–97. 102

[6] A. Blokh, D. Mimbs, L. Oversteegen, K. Valkenburg, *Laminations in the language of leaves*, Trans. Amer. Math. Soc., **365** (2013), 5367–5391. 61, 63

[7] A. Blokh, L. Oversteegen, R. Ptacek, V. Timorin, *Laminations from the Main Cubioid,* Disc. and Cont. Dyn. Syst., Ser. A, **36** (2016), 4665–4702. 71

[8] A. Blokh, L. Oversteegen, N. Selinger, V. Timorin, and S. Vejandla, *Symmetric cubic laminations*, Conformal Geometry and Dynamics, **27** (2023), 264–293. 61, 62, 63, 66, 68, 70, 71, 92, 104

[9] A. Blokh, L. Oversteegen, N. Selinger, V. Timorin, and S. Vejandla, *Lavaurs algorithm for cubic symmetric polynomials*, Ergodic Theory and Dynamical Systems. **45**:8 (2025), 2314–2340. 61, 62, 63, 68, 101

[10] A. Blokh, L. Oversteegen, V. Timorin, Y. Wang, *A model of the cubic connectedness locus*, Nonlinearity **38**:7 (2025), Paper No. 075014, 37 pp.; arXiv:2304.11516 (2023) 89

[11] A. Blokh, L. Oversteegen, and V. Timorin, *Slices of parameter space of cubic polynomials*, Trans. Amer. Math. Soc. **375**:8 (2022), 5313–5359.

[12] X. Buff, *Membres de certains lieux de connexité*, Peprint 2003.

[13] D. Calegari, *Saucages and Butcher Paper*, Recent progress in mathematics, 155—200, KIAS Springer Ser. Math., 1, Springer, Singapore (2022) 62

[14] L. DeMarco, C. McMullen, *Trees and the dynamics of polynomials*, Ann. Sci. École Norm. Sup. (4) **41** (2008), no. 3, 337–383. 90

[15] A. Douady, J. H. Hubbard, *Étude dynamique des polynômes complexes. Partie I.* Publications Mathématiques d'Orsay, 84-2. 1984. 75 pp. 90

[16] A. Douady, J. Hubbard, *A proof of Thurston's topological characterization of rational functions,* Acta Math. **171** (1993), no. 2, 263–297. 61, 83

[17] Goldberg, L. R., *On the multiplier of a repelling fixed point*, Inventiones mathematicae **118**:1 (1994), 85–108.

[18] L. Goldberg, J. Milnor, *Fixed points of polynomial maps. II. Fixed point portraits*, Ann. Sci. École Norm. Sup. (4) **26** (1993), no. 1, 51–98. 90

[19] P. Haïssinsky, *Chirurgie parabolique*, C.R. Acad. Sci. Paris, t. **327**, Série I (1998), 195–198. 79, 86

[20] P. Haïssinsky, *Déformation J-équivalente de polynômes géometriquement finis,* Fund. Math. **163** (2000), no. 2, 131–141. 91

[21] H. Inou, J. Kiwi, *Combinatorics and topology of straightening maps I: Compactness and bijectivity*, Adv. Math. **231** (2012), 2666–2733. 86, 87

[22] J. Kahn, M. Lyubich, *Local connectivity of Julia sets for unicritical polynomials*, Ann. of Math.(2) **170**(2009), 413–426. 82, 83, 84

[23] J. Kiwi, *Non-accessible critical points of Cremer polynomials*, Erg. Th. and Dyn. Syst. **20** (2000), no. 5, 1391–1403. 61

[24] J. Kiwi, *Wandering orbit portraits,* Trans. Amer. Math. Soc., **354** (2002), 1473–1485. 79, 92

[25] J. Kiwi, $\mathbb{R}$*eal laminations and the topological dynamics of complex polynomials*, Adv. in Math. **184**, 207–267

[26] J. Kiwi, *Combinatorial continuity in complex polynomial dynamics*, Proceedings of the London Mathematical Society **91**:1 (2005), 215–248. 98, 101, 102

[27] O. Kozlovski, S. van Strien, *Local connectivity and quasi-conformal rigidity of non-renormalizable polynomials,* Proc. of the LMS, **99** (2009), 275–296. 96

[28] Tan Lei, *Local properties of the Mandelbrot set at parabolic points.* The Mandelbrot set, theme and variations, 133–160, London Math. Soc. Lecture Note Ser., 274, Cambridge Univ. Press, Cambridge, 2000. 104

[29] M. Lyubich, *Conformal Geometry and Dynamics of Quadratic Polynomials, vol I-II*, in preparation. http://www.math.stonybrook.edu/ mlyubich/book.pdf

[30] R. Mañé, P. Sad, D. Sullivan, *On the dynamics of rational maps*, Ann. Sci. École Norm. Sup. (4) **16**:2 (1983), 193–217. 74

[31] J. Milnor, *Geometry and dynamics of quadratic rational maps,* Experimental Math., **2** (1993), 37–83.

[32] J. Milnor, *Periodic orbits, externals rays and the Mandelbrot set: an expository account.* Gèomètrie complexe et systèmes dynamiques (Orsay, 1995). Astèrisque No. 261, (2000), xiii, 277–333. 71, 100

[33] J. Milnor, *Dynamics in One Complex Variable*, Annals of Mathematical Studies **160**, Princeton (2006).

[34] J. Milnor, *Cubic polynomial maps with periodic critical orbit I,* in: Complex Dynamics, Families and Friends, ed. D. Schleicher, A.K. Peters (2009), 333-411.

[35] A. Poirier, *On postcritically finite polynomials*, Fund. Math. **203**:2 (2009), 107–163. 71

[36] S. Nadler,*Continuum theory. An introduction,* Monographs and Textbooks in Pure and Applied Mathematics, **158**, Marcel Dekker, Inc., New York, 1992. 87

[37] P. Roesch, Y. Yin, *Bounded critical Fatou components are Jordan domains for polynomials*, Sci. China Math. **65** (2022), 331–358. 97

[38] D. Schleicher, *Rational parameter rays of the Mandelbrot set.* Gèomètrie complexe et systèmes dynamiques (Orsay, 1995). Astèrisque No. 261 (2000), 405–443.

[39] D. Schleicher, *On fibers and local connectivity of Mandelbrot and Multibrot sets.* Fractal geometry and applications: a jubilee of Benoît Mandelbrot. 477–517, Proc. Symp. Pure Math., **72**, P. 1, Amer. Math. Soc., Providence, RI, 2004.

[40] D. Schleicher, *Appendix: Laminations, Julia sets and the Mandelbrot set,* in: Complex Dynamics: Families and Friends, edited by D. Schleicher, AK Peters, Wellesley, MA (2009), 111–137. 61

[41] D. Sullivan, *Quasiconformal homeomorphisms and dynamics. I. Solution of the Fatou-Julia problem on wandering domains.*, Ann. of Math. (2) **122** (1985), no. 3, 401–418. 61

[42] W. P. Thurston, *On the geometry and dynamics of iterated rational maps.* Edited by Dierk Schleicher and Nikita Selinger and with an appendix by Schleicher. Complex dynamics, 3–137, A K Peters, Wellesley, MA, 2009. 81

[43] S. Vejandla, *Cubic symmetric laminations*, PhD Thesis (2021), UAB. 61, 63, 92

AUTHORS

Alexander Blokh
Department of Mathematics
University of Alabama at Birmingham
Birmingham, AL 35294
email: ablokh@math.uab.edu

Lex Oversteegen
Department of Mathematics
University of Alabama at Birmingham
Birmingham, AL 35294
email: overstee@uab.edu

Nikita Selinger
Department of Mathematics
University of Alabama at Birmingham
Birmingham, AL 35294
email: selinger@uab.edu

Vladlen Timorin
Faculty of Mathematics
National Research University Higher School of Economics
6 Usacheva str., Moscow, Russia, 119048
email: vtimorin@hse.ru

Sandeep-Vejandla
Department of Mathematics
University of Alabama at Birmingham
Birmingham, AL 35294
email: vsc4u@uab.edu

Arnold Mathematical Journal
Volume 12, Issue 1, 2026

Cluster scattering coefficients in rank 2

Thomas Elgin Nathan Reading Salvatore Stella

Received 10 Jan 2025; Accepted 4 Jul 2025

Abstract:

We present conjectures on the scattering terms of cluster scattering diagrams of rank 2, supported by significant computational evidence.

AMS Classification: 13F60, 14N35

Key words and phrases: cluster scattering diagram, scattering term

1 Introduction

Scattering diagrams were introduced by Kontsevich, Soibelman, Gross, and Siebert in [14, 16] for the study of mirror symmetry. Cluster algebras were introduced by Fomin and Zelevinsky [7] for the study of total positivity, but quickly found algebraic, geometric, and combinatorial connections to a wide range of mathematical areas. Scattering diagrams became an essential tool in the structural study of cluster algebras when Gross, Hacking, Keel, and Kontsevich defined (and proved the existence of) cluster scattering diagrams [11]. Applying the scattering-diagram machinery of broken lines and theta

DOI: 10.56994/ARMJ.012.001.003

functions that had been introduced and developed in various papers by Carl, Gross, Hacking, Keel, Pumperla, and Siebert [5, 9, 10, 12, 15], they corrected and proved a conjecture of Fock and Goncharov [6] on the cluster variety and several conjectures of Fomin and Zelevinsky on the structure of cluster algebras [8].

The defining data of a cluster scattering diagram is an exchange matrix, meaning a skew-symmetrizable integer matrix. Following the usual terminology in the cluster algebras literature, we will say that the cluster scattering diagram of an $r \times r$ exchange matrix is a cluster scattering diagram of "rank r" (regardless of the rank of the matrix in the usual linear-algebraic sense).

A rank-r scattering diagram lives in a real vector space of dimension r and consists of walls (codimension-1 cones), each decorated by a scattering term (a multivariate formal power series). The cluster scattering diagram is defined by specifying that certain walls must be present and certain walls must not be present and then requiring a consistency condition. Some details of the definition, in rank 2, are given in Section 2. A rank-1 cluster scattering diagram is trivially easy to construct, but a typical rank-2 cluster scattering diagram is already extremely complicated, and completely explicit formulas are not known for coefficients of scattering terms. (This statement depends, of course, on what one calls "explicit". See Section 4.)

Although cluster scattering diagrams of higher rank can be vastly more complicated, there is a sense in which cluster scattering diagrams of rank 2 tell much of the story about higher rank. The importance of rank 2 is seen in the proof of the existence of the cluster scattering diagram in general rank in [11, Appendix C]. That proof constructs the cluster scattering diagram degree by degree (in the sense of degrees of terms in the power series). To move the construction to the next degree, one adds higher order terms and new nontrivial walls to yield consistency up to that degree. The consistency is checked locally about every $(n-2)$-dimensional intersection of walls ("joint"), and the local consistency condition is exactly the condition on a rank-2 cluster scattering diagram. Thus, in some sense, the construction of a general-rank cluster scattering diagram consists

of constructing rank-2 cluster scattering diagrams all throughout the ambient space.

The purpose of this note is to record and share some conjectures on cluster scattering diagrams of rank 2 that are supported by significant computational evidence.

Remark. Since this note was posted on the arXiv, Ryota Akagi has made us aware that some of our conjectures can be proved using results from his paper [1]. Also, an anonymous referee pointed out some additional special cases of our conjectures that are known. Details on these connections are given in Section 4.

2 Cluster scattering diagrams of rank 2

We now give the definition of the cluster scattering diagram, specialized to rank 2 and following the notation of [17]. (See also [11, Example 1.15].) Up to symmetry, we may as well assume that the ***exchange matrix*** is $B = \left[\begin{smallmatrix} 0 & c \\ -b & 0 \end{smallmatrix}\right]$ for $b, c > 0$.

Let N be the lattice $\mathbb{Z}^2$ with the usual basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ and let N° be the sublattice of N generated by $b\mathbf{e}_1$ and $c\mathbf{e}_2$. Let M be the dual lattice to N and let M° be the superlattice of M that is dual to N°, with basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ such that $\langle \mathbf{f}_1, b\mathbf{e}_1 \rangle = \langle \mathbf{f}_2, c\mathbf{e}_2 \rangle = 1$ and $\langle \mathbf{f}_1, \mathbf{e}_2 \rangle = \langle \mathbf{f}_2, \mathbf{e}_1 \rangle = 0$. It is the interplay between N and N°, and dually M° and M, that integrates the skew-symmetrizability of B seamlessly into the construction.

Take indeterminates z_1 and z_2 and define $\zeta_1 = z_2^c$ and $\zeta_2 = z_1^{-b}$. (In [11, Example 1.15], z_1 and z_2 are called A_1 and A_2.) Given a vector $n = n_1\mathbf{e}_1 + n_2\mathbf{e}_2 \in N$, we write ζ^n to mean $\zeta_1^{n_1}\zeta_2^{n_2}$, and given $m = m_1\mathbf{f}_1 + m_2\mathbf{f}_2 \in M^\circ$, we write z^m to mean $z_1^{m_1}z_2^{m_2}$.

A ***wall*** is a pair $(\mathfrak{d}, f_\mathfrak{d})$ where $\mathfrak{d}$ is a codimension-1 cone (thus a line through the origin or a ray with vertex at the origin) and $f_\mathfrak{d}$ is a formal power series in ζ_1 and ζ_2 satisfying the following conditions. The condition on $\mathfrak{d}$ is that there must exist a nonzero vector in N with nonnegative entries that is normal to $\mathfrak{d}$. Take n to be such a vector that is primitive in N. The condition on $f_\mathfrak{d}$ is that it has the form $1 + \sum_{\ell \geq 1} c_\ell \zeta^{\ell n}$, or in other words $f_\mathfrak{d}$ is a univariate formal power series in ζ^n with constant term 1. The formal power series $f_\mathfrak{d}$ is called the ***scattering term*** on $\mathfrak{d}$. A ***scattering diagram*** is a collection

$\mathfrak{D}$ of walls, satisfying a finiteness condition that amounts to requiring that all of the relevant computations can be made by taking limits in the sense of formal power series. Specifically, the requirement is that for any $k \geq 0$, all but finitely many walls of $\mathfrak{D}$ are 1 plus terms with total degree $> k$ in ζ_1 and ζ_2.

This definition of a scattering diagram is so broad as to be almost meaningless; the condition that makes a scattering diagram interesting is consistency, which we now define. We begin by describing an action on formal Laurent series that is associated to crossing a wall. Given a wall $\mathfrak{d}$ having normal vector n with nonnegative entries and given a path that crosses the wall transversely, the ***wall-crossing automorphism*** takes a Laurent monomial z^m to $z^m f_{\mathfrak{d}}^{\langle m, \pm n' \rangle}$, where n' is the primitive vector in N° that is a positive scaling of n. The sign in the formula is "$-$" if the curve crosses in the direction that agrees with n or "$-$" if the curve crosses in the direction that disagrees with n. We extend linearly and take limits to act on formal Laurent series. Using standard terminology, we call this a wall-crossing *automorphism*, but we need not be careful about the algebraic structure on which it acts as an automorphism. For us, it is just an action on formal Laurent series.

A ***path-ordered product*** is, loosely speaking, the composition of these wall-crossing automorphisms along a path. More correctly, since a path might intersect infinitely many walls, the path-ordered product is defined as a limit of formal Laurent series. For each $k \geq 0$, consider only those walls whose scattering terms have nontrivial terms of total degree $\leq k$. The finiteness condition on $\mathfrak{D}$ ensures that there are finitely many such walls, so we can compose the wall-crossing automorphism for those walls. The path-ordered product is the limit, in the sense of formal Laurent series, as $k \to \infty$. A scattering diagram is ***consistent*** if the path-ordered product for a small oriented circle about the origin is the identity map.

A wall with nonnegative normal $n = n_1\mathbf{e}_1 + n_2\mathbf{e}_2 \in N$ is ***outgoing*** if the vector $-bn_2\mathbf{f}_1 + cn_1\mathbf{f}_2 \notin \mathfrak{d}$. One can check that, in any case, $-bn_2\mathbf{f}_1 + cn_1\mathbf{f}_2$ is in the line containing $\mathfrak{d}$. Thus $\mathfrak{d}$ is outgoing if and only if it is a ray (rather than a line) and is contained in the fourth quadrant. The ***cluster scattering diagram*** is the unique consistent scattering

diagram containing the walls $(\mathbf{e}_1^\perp, 1 + z_2^c)$ and $(\mathbf{e}_2^\perp, 1 + z_1^{-b})$ such that all other walls are outgoing.

For $i, j \geq 0$, define $\tau(i, j)$ to be the coefficient of $\zeta_1^i \zeta_2^j = z_1^{-jb} z_2^{ic}$ on the wall of the cluster scattering diagram that is orthogonal to $i\mathbf{e}_1 + j\mathbf{e}_2$. To specify or emphasize the exchange matrix, we may write $\tau^{b,c}(i, j)$, but generally, we think of $\tau(i, j)$ as a function of indeterminates b and c. We define $g = \frac{\gcd(ib, jc)}{\gcd(i, j)}$.

Example. Several of the $\tau(i, j)$ are shown below, with i changing in the horizontal direction and j changing in the vertical direction and $\tau(0, 0)$ at the bottom-left.

$$\begin{array}{cccccc}
\vdots & & & & & \\
0 & \frac{g(b-1)(b-2)}{6} & \frac{g(b-1)(3bc-2b-3c+1)}{6} & \frac{g(3bc-3b-3c+2+g)(3bc-3b-3c+1+g)}{6} & & \\
0 & \frac{g(b-1)}{2} & \frac{g(2bc-2b-2c+1+g)}{2} & \frac{g(c-1)(3bc-3b-2c+1)}{6} & & \\
1 & g & \frac{g(c-1)}{2} & \frac{g(c-1)(c-2)}{6} & & \\
1 & 1 & 0 & 0 & \cdots &
\end{array}$$

Example. Take $b = 3$ and $c = 2$ so that the exchange matrix is $\left[\begin{smallmatrix} 0 & 2 \\ -3 & 0 \end{smallmatrix}\right]$. Some of the integers $\tau^{3,2}(i, j)$ are shown below, again with i changing in the horizontal direction and j changing in the vertical direction and $\tau^{3,2}(0, 0)$ at the bottom-left.

$$\begin{array}{cccccccccc}
\vdots & & & & & & & & & \\
0 & 0 & 0 & 1 & 33 & 87 & 286 & 429 & & \\
0 & 0 & 0 & 5 & 327 & 143 & 132 & 143 & & \\
0 & 0 & 1 & 6 & 33 & 42 & 33 & 6 & & \\
0 & 0 & 2 & 6 & 14 & 6 & 2 & 0 & & \\
0 & 1 & 14 & 5 & 14 & 1 & 0 & 0 & & \\
0 & 1 & 2 & 1 & 0 & 0 & 0 & 0 & & \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & & \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots &
\end{array}$$

Remark on symmetry. We point out the obvious symmetry $\tau^{b,c}(i,j) = \tau^{c,b}(j,i)$. Another symmetry is less obvious but not hard to prove using mutation of scattering diagrams: Interpreting (i,j) as a vector $i\mathbf{e}_1 + j\mathbf{e}_2$ in the root lattice of a root system with Cartan matrix $\left[\begin{smallmatrix} 2 & -c \\ -b & 2 \end{smallmatrix}\right]$ and simple roots $\mathbf{e}_1$ and $\mathbf{e}_2$, the integers $\tau^{b,c}(i,j)$ are invariant under the action of the Weyl group on (i,j). This amounts to the symmetries $\tau^{b,c}(i,j) = \tau^{b,c}(i,-j+ci)$ and $\tau^{b,c}(i,j) = \tau^{b,c}(-i+bj,j)$.

3 Conjectures on coefficients of scattering terms

Conjectures.

1. For $i,j > 0$, the coefficient $\tau(i,j)$ is a polynomial in b, c, and g.
2. The polynomial has g as a factor and its degree in g is $\gcd(i,j)$.
3. Its degree in b is $j-1$ and its degree in c is $i-1$.
4. The polynomial $(\max(i,j))!\ \tau(i,j)$ has integer coefficients.
5. $\tau(1,j) = \dfrac{g}{b}\dbinom{b}{j}$
6. $\tau(i,1) = \dfrac{g}{c}\dbinom{c}{i}$
7. $\tau(i,i) = \dfrac{g}{(b-1)(c-1)i+g}\dbinom{(b-1)(c-1)i+g}{i}$
8. $\tau(i,i) = \displaystyle\sum_{\ell=0}^{\infty} \frac{g}{\ell+1}\dbinom{i-1}{\ell}\dbinom{i(bc-b-c)+g-1}{\ell}$
9. $\tau(i,i-1) = \dfrac{g}{i(ib-i+1)}\dbinom{(ib-i+1)(c-1)}{i-1}$
10. $\tau(j-1,j) = \dfrac{g}{j(jc-j+1)}\dbinom{(jc-j+1)(b-1)}{j-1}$
11. $\tau^{b,b}(i,j)$ is a polynomial in b of degree $i+j-1$ that expands positively in the basis $\{\binom{b}{0}, \binom{b}{1}, \binom{b}{2}, \ldots\}$.

12. $\tau^{b,b}(i,j)$ has unimodal log-concave coefficients.

13. $\tau^{1,5}(2j,j) = \frac{1}{j}\sum_{\ell=0}^{\infty} \binom{\ell}{j-\ell+1}\binom{j+\ell-1}{\ell}$.

Comments.

- Conjectures 5 and 6 are symmetric and Conjectures 9 and 10 are symmetric.
- Conjecture 8 is equivalent to Conjecture 7.
- Since $c = b$ in Conjectures 11 and 12, also $g = b$.

Assuming Conjecture 1, write $\tau(i,j\,;k)$ for the coefficient of g^k in $\tau(i,j)$ and similarly $\tau^{b,c}(i,j\,;k)$. Because g is also invariant under the action of the Weyl group on (i,j), we have $\tau^{b,c}(i,j;k) = \tau^{b,c}(i,-j+bi;k)$ and $\tau^{b,c}(i,j;k) = \tau^{b,c}(-i+aj,j;k)$.

Conjectures.

14. $\tau(i,j\,;k)$ is a polynomial of degree $j-k$ in b and degree $i-k$ in c and has a term that is a nonzero constant times $b^{j-k}c^{i-k}$.

15. If $\gcd(i,j) = 1$, then $\tau(ik,jk\,;k) = \dfrac{\tau(i,j\,;1)^k}{k!}$.

16. $\tau(k,jk\,;k-1) = \dfrac{\tau(1,j\,;1)^{k-1}\cdot p_j}{(k-2)!}$, where p_j is a polynomial in b and c that depends only on j, not k.

17. $\tau(ik,k\,;k-1) = \dfrac{\tau(i,1\,;1)^{k-1}\cdot p_i}{(k-2)!}$, where p_i is a polynomial in b and c that depends only on i, not k.

18. If $\gcd(i,j) = 1$, then $\tau(ki,kj\,;k-1)$ has a factor p_{ij} that depends only on i and j, not on k, and the other factors of $\tau(ki,kj\,;k-1)$ also appear as factors of $\tau(i,j\,;1)$.

Comments.

- Assuming Conjecture 2, for fixed i and j, Conjecture 15 is a formula for $\tau(i,j;k)$ for the largest k such that $\tau(i,j;k) \neq 0$ in terms of some $\tau(\,\cdot\,,\,\cdot\,;1)$.

- In Conjecture 18, the factors of $\tau(i, j\,; 1)$ appear to different powers in $\tau(ki, kj\,; k-1)$ for various k. For example,

$$\tau(2,3\,;\,1) = \frac{(b-1)\,(3cb-2b-3c+1)}{6}$$
$$\tau(4,6\,;\,1) = \frac{(b-1)\,p_{23}}{180}$$
$$\tau(6,9\,;\,2) = \frac{(b-1)^2\,(3cb-2b-3c+1)\,p_{23}}{1080}$$
$$\tau(8,12\,;\,3) = \frac{(b-1)^3\,(3cb-2b-3c+1)^2\,p_{23}}{12960}$$

$$\begin{aligned}\text{for}\quad p_{23} = {} & 330b^4c^3 - 720b^4c^2 - 1530b^3c^3 + 525b^4c + 2880b^3c^2 \\ & + 2610b^2c^3 - 128b^4 - 1770b^3c - 4140b^2c^2 \\ & - 1950b\,c^3 + 352b^3 + 2085b^2c + 2520b\,c^2 \\ & + 540c^3 - 328b^2 - 1005cb - 540c^2 + 122b + 165c - 15\end{aligned}$$

Remark on computation. These conjectures are backed up by significant computational evidence. The functions $\tau(i, j)$ of b and c can be computed symbolically up to large values of i and j. Thus, for example, the polynomials shown in the first example in Section 2 are known to be correct for all b and c, rather than only for some specific values of b and c. Similarly, the conjectures on $\tau(i, j)$ in this section have been checked for many values of i and j, and each case that has been checked is true for all b and c.

Computing the functions $\tau(i, j)$ proceeds by induction on $i + j$, by solving, at each step, the equations that describe consistency of the cluster scattering diagram in degree $i + j$ using known values of $\tau(i', j')$ for $i' + j' < i + j$. Thus, it would be significant progress even to find a recursive description of $\tau(i, j)$ whose recursive step is simpler than solving a system of equations.

The code we used in our experiments together with precomputed data up to degree 20 is available at `https://github.com/Etn40ff/scatcoef/`.

4 Related work

Work of Ryota Akagi [1] also aims at explicitly understanding scattering terms in the cluster scattering diagram, but with different approach and conventions. His work is independent of ours, was posted before we publicized any of our conjectures, and also contains results in different directions that we did not conjecture. Akagi has informed us of a forthcoming paper [2] in which he proves Conjectures 1, 2, 3, 5, 6, 11, and 15, as well as part of Conjecture 14 and a result that is similar to Conjectures 16, 17, and 18, using his results from [1].

Tom Bridgeland [3, Theorem 1.5] identifies the cluster scattering diagram with the stability scattering diagram in the case $b = c$, thus realizing $\tau^{b,b}(i, j)$ as the Euler characteristic of a certain moduli scheme of representations of the associated Jacobi algebra.

An anonymous referee pointed out that Conjectures 7 and 9 (and symmetrically 10) generalize results of Reineke and Weist [19]. Specifically, the case $g = 1$ of Conjecture 7 is [19, Corollary 11.2] (which proves a conjecture of Gross and Pandharipande [13, Conjecture 1.4]) and the case $g = 1$ of Conjecture 9 is [19, Theorem 9.4]. In comparing Conjecture 7 with [19, Corollary 11.2], it is useful to notice that after setting $g = 1$, the right side of Conjecture 7 can be rewritten as $\frac{1}{(bc-b-c)i+1}\binom{(b-1)(c-1)i}{i}$.

Gross, Hacking, Keel, and Kontsevich [11, Example 1.15] state the case $b = c$ of Conjecture 8 and attribute its proof to Reineke [18].

Burcroff, Lee, and Mou have recently given a formula for the coefficients of scattering terms as a weighted sum over certain combinatorial objects called tight gradings [4, Theorem 1.1]. They intend to explore some of the conjectures of this paper using tight gradings [4, Section 9.2]. It would also be interesting to understand how the Weyl group symmetry appears in the combinatorics of tight gradings. (See the remark on symmetry at the end of Section 2.)

Acknowledgements.

Thomas Elgin and Nathan Reading were partially supported by the National Science Foundation under award number DMS-2054489. Salvatore Stella was partially supported by INdAM and GNSAGA. We thank the anonymous referees for their helpful comments.

References

[1] Ryota Akagi, *Explicit forms in lower degrees of rank 2 cluster scattering diagrams.* Preprint, 2023 (arXiv:2309.15470). 113, 119

[2] Ryota Akagi, *Some coefficients in rank 2 cluster scattering diagrams.* In preparation, 2025. 119

[3] Tom Bridgeland, *Scattering diagrams, Hall algebras and stability conditions.* Algebr. Geom. **4** (2017), no. 5, 523–561. 119

[4] Amanda Burcroff, Kyungyong Lee, and Lang Mou, *Scattering diagrams, tight gradings, and generalized positivity.* Proc. Natl. Acad. Sci. USA **122** (2025), no. 18, Paper No. e2422893122. 119

[5] Michael Carl, Max Pumperla, and Bernd Siebert, *A tropical view of Landau-Ginzburg models.* Acta Math. Sin. (Engl. Ser.) **40** (2024), no. 1, 329–382. 112

[6] Vladimir V. Fock and Alexander Goncharov, *Cluster $\mathcal{X}$-varieties at infinity.* Preprint, 2011. arXiv:1104.0407. 112

[7] S. Fomin and A. Zelevinsky, *Cluster algebras. I. Foundations.* J. Amer. Math. Soc. **15** (2002), no. 2, 497–529. 111

[8] S. Fomin and A. Zelevinsky, *Cluster Algebras. IV. Coefficients.* Compositio Mathematica **143** (2007), 112–164. 112

[9] Mark Gross, *Mirror symmetry for $\mathbb{P}^2$ and tropical geometry.* Adv. Math. **224** (2010), no. 1, 169–245. 112

[10] Mark Gross, Paul Hacking, and Sean Keel, *Mirror symmetry for log Calabi-Yau surfaces I.* Publ. Math. Inst. Hautes Études Sci. **122** (2015), 65–168. 112

[11] Mark Gross, Paul Hacking, Sean Keel, and Maxim Kontsevich, *Canonical bases for cluster algebras.* J. Amer. Math. Soc. **31** (2018), no. 2, 497–608. 111, 112, 113, 119

[12] Mark Gross, Paul Hacking, and Bernd Siebert, *Theta functions on varieties with effective anti-canonical class.* Mem. Amer. Math. Soc. **278** (2022), no. 1367. 112

[13] Mark Gross and Rahul Pandharipande, *Quivers, curves, and the tropical vertex.* Port. Math. **67** (2010), no. 2, 211–259. 119

[14] Mark Gross and Bernd Siebert, *From real affine geometry to complex geometry.* Ann. of Math. (2) **174** (2011), no. 3, 1301–1428. 111

[15] Mark Gross and Bernd Siebert, *Theta functions and mirror symmetry.* Advances in geometry and mathematical physics, 95–138. Surv. Differ. Geom. **21**. 112

[16] Maxim Kontsevich and Yan Soibelman, *Affine structures and non-Archimedean analytic spaces.* The unity of mathematics, 321–385. Progr. Math. **244**. 111

[17] Nathan Reading, *Scattering fans* Int. Math. Res. Not. IMRN **2020**, no. 23, 9640–9673. 113

[18] Markus Reineke, *Cohomology of quiver moduli, functional equations, and integrality of Donaldson-Thomas type invariants.* Compos. Math. **147** no. 3 (2011), 943–964. 119

[19] Markus Reineke and Thorsten Weist, *Refined GW/Kronecker correspondence.* Math. Ann. **355** (2013), no. 1, 17–56. 119

AUTHORS

Thomas Elgin
Department of Mathematics,
North Carolina State University,
Raleigh, NC, USA
email: thomaselgin97@gmail.com

Nathan Reading
Department of Mathematics,
North Carolina State University,
Raleigh, NC, USA
email: reading@math.ncsu.edu

Salvatore Stella
Dipartimento di Ingegneria e
Scienze dell'Informazione e Matematica,
Università degli Studi dell'Aquila, IT
email: salvatore.stella@univaq.it

Arnold Mathematical Journal
Volume 12, Issue 1, 2026

Monodromy Groups and the Insolvability of Transcendental Equations in Quadratures

Allan Allemand Alexey Kanel-Belov Rodion Zaytsev

Received 31 Aug 2025; Accepted 4 Nov 2025

Abstract:

This paper presents the authors' results in applying Arnold's method to compute the monodromy groups of certain trigonometric complex equations, and also provides a survey of other results in this area.

AMS Classification: 30D05; 20B07, 34B24, 34M35, 41A60

Key words and phrases: Topological Galois theory, monodromy group, unsolvability in quadratures, transcendental equations, braid group.

1 Introduction

Our primary goal in this paper is to compute the monodromy groups for several equations of interest. The transcendental equations investigated here are not arbitrary; they are

DOI: 10.56994/ARMJ.012.001.004

motivated by their appearance in mathematical physics and spectral theory. Equations of the type $\tan(z) - z = 0$ are canonical in this regard: they arose in Fourier's 19th-century analysis of the heat equation, where their roots correspond to the eigenvalues of a Sturm-Liouville operator, as noted by F. Klein [5]. A similar result was obtained by Fourier, but here we present a new perspective and proof using modern methods, which interestingly leads to the same conclusion. A. Elishev demonstrated this fact in modern terms [2], and we include his argument.

Modern analogues, such as the equation studied by Heifetz for acoustic wave propagation, further highlight the physical importance of both the real roots (propagating modes) and the complex roots (resonances) [3]. We include a full summary of his results, as the original was published in a difficult-to-access Soviet journal. This motivates a deeper study of the complete root structure of such equations under parameter variation, which is what their monodromy groups describe.

Our primary tool is the topological Galois theory developed by V. I. Arnold and later by A. G. Khovanskii [4]. This theory connects the solvability of an equation to the algebraic properties of its monodromy group. The notion of solvability can be made precise through a hierarchy of function classes. Starting with elementary functions (built from rational functions, $\exp(z)$, $\log(z)$, and algebraic operations), one can define functions solvable in quadratures (allowing integration), and recursively, in k-**quadratures** (meaning functions obtained by applying up to k nested integrations). The union of all such classes forms the class of functions solvable in generalized quadratures. The main theorem of topological Galois theory, in its strong form due to Khovanskii, addresses this entire hierarchy:

Theorem. *If the solution $z(a)$ of an equation $f(z) = a$, considered as a function of the parameter a, is representable in generalized quadratures, then the monodromy group of the equation is solvable.*

In this paper, we use the contrapositive of this theorem. By showing these monodromy groups to be unsolvable (e.g., containing A_5 or A_∞), we prove that these equations are

not solvable even in the broad class of generalized quadratures.

Our approach follows the constructive spirit of Arnold's method. The core idea is to analyze the behavior of the roots z_i of the equation $f(z) = a$ as the parameter a traverses a closed loop in the complex plane. The roots themselves trace continuous paths, which may not be closed, resulting in a permutation of their initial positions. This permutation defines an element of the monodromy group. Non-trivial permutations are generated when these loops encircle the critical values of $f(z)$ — the images of points where the derivative $f'(z)$ vanishes. Examples of application of this method with illustrations can be found in our previous paper [1].

Therefore, our work in the subsequent sections will be to identify these critical values and deliberately construct paths around them to generate a set of permutations sufficient to prove the unsolvability of the corresponding monodromy group.

2 The Case of a Self-Adjoint Function

Let the function $f(z)$ have the following properties:

1. The function $f(z)$ is self-adjoint, i.e., $f(\overline{z}) = \overline{f(z)}$.

2. The critical points of $f(z)$ (where $f'(z) = 0$) have multiplicity 3.

3. The equation $f(x) = a$ for a real a has a unique real root.

From complex analysis, we know that traversing a critical point must permute three roots in a cycle. Since the function is self-adjoint, its set of roots consists of one real root and pairs of complex conjugate roots. Consequently, when traversing a critical point, the unique real root and a pair of complex conjugate roots must be permuted.

By traversing one critical point, we obtain the permutation of roots $(1\,2\,3)$, where 1 denotes the unique real root. By traversing another critical point, we get the permutation $(1\,4\,5)$. Together, these two permutations generate the alternating group A_5. Thus, the

monodromy group of the equation is unsolvable, which implies that the equation is not solvable in elementary functions.

It remains for us to verify that the following equations satisfy these initial conditions, from which their insolvability will immediately follow. Note that although we do not provide the proof here, the same method is applicable to simpler equations like $\sin(z)-z = a$ and $\cos(z) - z = a$.

2.1 Insolvability of $\sin(\sin(z)) - z = a$

Let us prove the insolvability of the equation

$$\sin(\sin(\dots \sin(z) \dots)) - z = a$$

where the function $f(z)$ is the difference between a composition of sines and the identity function. For simplicity, we will show that conditions 1–3 are met for the simple composition

$$\sin(\sin(z)) - z = a$$

as the proof for more complex compositions is completely analogous. First, let us find the critical points of the function:

$$f(z) = \sin(\sin(z)) - z$$
$$f'(z) = \cos(\sin(z))\cos(z) - 1 = 0 \Rightarrow \cos(\sin(z))\cos(z) = 1$$

Clearly, both $\cos(\sin(z))$ and $\cos(z)$ must be simultaneously equal to either 1 or -1. The latter is impossible, because at the points $z = \pi + 2\pi k$, the sine is zero, and the cosine of zero is one. Consequently, the critical points of the function will be

$$z_k = 2\pi k, k \in \mathbb{Z}$$

At these points,

$$f''(z_k) = -\sin(\sin(z_k))\cos(z_k) - \cos(\sin(z_k))\sin(z_k) = -0 \cdot 1 - 1 \cdot 0 = 0$$

$$f'''(z_k) = -\cos(\sin(z_k))\cos^2(z_k) - \sin(\sin(z_k))\sin(z_k) +$$

$$+\sin(\sin(z_k))\sin(z_k)\cos(z_k) - \cos(\sin(z_k))\cos(z_k) = -1 - 0 + 0 - 1 = -2 \neq 0$$

We find that our points have multiplicity 3. Furthermore, a real root exists for any real a and is unique due to the strict monotonicity of the function, since

$$f'(x) = \cos(\sin(x))\cos(x) - 1 < 0$$

at all points except $2\pi k$. The other roots will come in complex conjugate pairs because the composition of two self-adjoint functions is self-adjoint.

2.2 Insolvability of $\sin(\cos(z)) - z = a$

Now we will show that conditions 1–3 are satisfied for

$$\sin(\cos(z)) - z = a$$

Let us find the critical points and check their multiplicity:

$$f'(z) = -\cos(\cos(z))\sin(z) - 1 = 0 \Rightarrow \cos(\cos(z))\sin(z) = -1 \Rightarrow$$

$$\Rightarrow z_k = -\pi/2 + 2\pi k, k \in \mathbb{Z}$$

Furthermore,

$$f''(z_k) = \sin(\cos(z_k))\sin^2(z_k) - \cos(\cos(z_k))\cos(z_k) = 0 - 0 = 0$$

$$f'''(z_k) = -\cos(\cos(z_k))\sin^3(z_k) + 2\sin(\cos(z_k))\sin(z_k)\cos(z_k) -$$

$$-\sin(\cos(z_k))\sin(z_k)\cos(z_k) + \cos(\cos(z_k))\sin(z_k) = -1 - 0 - 0 - 1 = -2 \neq 0$$

In this case, there is a unique real root because

$$-\cos(\cos(x))\sin(x) - 1 < 0$$

at all points except $-\pi/2 + 2\pi k$. As before, the other roots will come in complex conjugate pairs because the composition of two self-adjoint functions is self-adjoint.

3 Insolvability of $\tan(z) - z = a$

Let us prove the insolvability of the equation

$$\tan(z) - z = a$$

First, we find the critical points:

$$f'(z) = \frac{1}{\cos^2(z)} - 1 = 0 \Rightarrow \cos^2(z) = 1 \Rightarrow z_k = \pi k, k \in \mathbb{Z}$$

Let us check the multiplicity of these points:

$$f''(z_k) = \frac{2\sin(z_k)}{\cos^3(z_k)} = 0$$
$$f'''(z_k) = \frac{2}{\cos^2(z_k)} + \frac{6\sin^2(z_k)}{\cos^4(z_k)} = 2 \neq 0$$

A consequence of Rouché's theorem is that the equation $\tan(z) - z = 0$ has only real roots (a detailed proof is given in [6]). Consequently, if we let the parameter a in the equation $\tan(z) - z = a$ approach zero, three roots will merge into one point.

The function $f(z) = \tan(z) - z$ is self-adjoint, which implies that its roots must come in complex conjugate pairs. In this case, the function $f(x) = \tan(x) - x$ is strictly monotonic on each interval $(-\pi/2 + \pi k, \pi/2 + \pi k)$, which means that the equation has a unique root on each such interval. The situation is similar to the previous cases, except that the real root is not unique. Nevertheless, due to the multiplicity of 3, in the vicinity of a critical point, one real root and a pair of complex conjugate roots will also be permuted. We will construct a path for the parameter in such a way that a real root and a complex conjugate root from the upper half-plane are exchanged.

We modify the method used in the previous section. Let us fix the value $a = -1$ and denote two complex roots as c_1 and $c_2 = \overline{c_1}$, where c_1 is in the upper half-plane. Let us denote the real roots as $z_0, z_1, z_2, ...$, where $z_0 \to 0$ as $a \to 0$.

We start with the configuration of real roots:

$$(z_1, z_2, z_3, ...).$$

Consider a path in the parameter plane of a that starts at $a = -1$ and moves left along the real axis, bypassing each critical value along a small semicircle (we choose the one that causes an exchange between a real root and c_1).

When a reaches the n-th critical point, it loops around it in a small circle and returns along the same path. Let us examine what happens to the roots.

First, c_1 swaps places with z_1, then z_1 with z_2, and so on.

By the time we approach the n-th critical point, the arrangement of roots becomes:

$$(c_1, z_1, z_2, \dots, z_{n-2}, z_n, \dots).$$

The root z_{n-1} moves into the upper half-plane. Traversing the critical point causes a cyclic permutation (z_n, z_{n-1}, c_2). Repeating this traversal twice, we get:

$$(c_1, z_1, \dots, z_{n-3}, z_{n-2}, c_2, \dots).$$

Now z_n is in the upper half-plane, and z_{n-1} is in the lower. As we move back, all previous permutations are reversed: z_n and z_{n-2} swap places, then z_{n-2} and z_{n-3}, and so on.

As a result, we get the configuration:

$$(z_1, z_2, \dots, z_{n-2}, z_n, c_2, \dots),$$

where c_1 returns to its original place, while z_{n-1} remains below.

If we now repeat the same operation for the $(n+1)$-th critical point, we get the permutation:

$$(z_1, z_2, \dots, z_{n-2}, z_n, z_{n+1}, z_{n-1}, \dots),$$

and both complex roots will return to their places. This corresponds to the cycle $(n-1\ n+1\ n)$.

Such 3-cycles for all $n \geq 5$ generate the **infinite alternating group**, denoted A_∞. By analogy with S_∞, this group is a group of finitely supported permutations. More precisely, A_∞ is the subgroup of S_∞ consisting of all *even* permutations on the countable set of roots.

Since any 3-cycle is an even permutation, and any finite alternating group A_k is generated by 3-cycles, our method allows us to construct any permutation from A_k for

any k. As the group A_k is unsolvable for $k \geq 5$, the group A_∞ which contains all of them is also unsolvable.

4 Insolvability of an Equation related to $z^z = a$

We now turn to another classical transcendental equation, $z^z = a$. While a full analysis of its monodromy group is possible, for the purpose of demonstrating the method with simpler derivatives, we will analyze the related, but distinct, equation $e^z + z = a$. It is important to note that these two equations are not equivalent. The analysis below pertains strictly to $e^z + z = a$.

Let us find the critical points of $f(z) = e^z + z$:

$$f'(z) = e^z + 1 = 0 \Rightarrow z_k = (2k+1)\pi i, k \in \mathbb{Z}$$

In this case,

$$f''(z_k) = e^{z_k} = -1 \neq 0$$

Thus, the critical points have multiplicity 2. The function $f(x)$ of a real argument is strictly monotonic because

$$f'(x) = e^x + 1 > 0$$

which means the equation has a unique real root.

First, for simplicity, let us consider the starting point $a = 0$.

Let us consider a path $z(t)$ that leads from this real root to the n-th critical point, then makes a semi-loop around it, swapping the roots, after which the new root returns along the same path.

Now let us derive the trajectory described by the parameter a while the real root moves along the specified path. On the first segment of the path:

$$z(t) = x + it,$$

where x is the real root (i.e., $e^x = -x$), so

$$a(t) = x + it + e^{x+it} = x(1 - \cos t) + i(t - x \sin t).$$

It is important to note that the loops bypass the "dangerous" points a_k, since at $t = (2k + 1)\pi$:

$$a = 2x + i(2k + 1)\pi.$$

Since $x < -\frac{1}{2}$ (which is easy to prove), and the real part of a_k is -1, the point a is to the left of a_k. This also shows that at the end of the first path segment, a is located to the left of a_n.

Let us denote $y_n = (2n + 1)\pi$. Then the second segment of the path is given by the equation:

$$z(s) = s + iy_n + e^{s+iy_n} \Rightarrow a(s) = s - e^s + iy_n.$$

Thus, $a(s)$ also moves along a straight line parallel to the real axis.

On the third segment of the path, after the permutation of the roots, the parameter a describes a loop around a_n, which follows from the general theory. This trajectory is homotopically equivalent to a simple loop around the corresponding point.

Since a plane with a discrete set of points removed is homotopically equivalent to a wedge of circles, any loop around these points can be decomposed into a product of simple loops like the one shown above.

This means that any permutation can be realized by successive transpositions of the real root with the others. On the other hand, the permutation group is transitive, which is true in the general case.

Let us consider the decomposition of a permutation σ that maps r_n to the real root (let us number the roots with natural numbers so that $r_1 = x$ is the real root). As shown above, it will have the form:

$$\sigma = (1\, n_1)(1\, n_2) \dots (1\, n_N).$$

Obviously, the transposition $(1\, n)$ must be part of this product, otherwise the root r_n would remain in its place. Consequently, the group contains all transpositions $(1\, n)$, which

generate the group S_∞ of all permutations of the countable set of roots that affect only a finite number of elements, known as the group of finitely supported permutations. S_∞ is countable and contains all finite symmetric groups S_n. The unsolvability of S_n for $n \geq 5$ implies the unsolvability of S_∞.

A Related Work in Functional Analysis

The equations we have studied are not merely mathematical curiosities; they arise in various physical and analytical contexts. In this appendix, we review two such connections.

A.1 Asymptotic Solutions

For the reader's convenience and due to the difficulty of accessing the original Soviet-era source [3], this section provides a detailed summary of Heifetz's work on the roots of a dispersion equation. The formulas obtained are used for numerical calculations and asymptotic investigation of the roots.

Heifetz considered a point acoustic source located in a homogeneous liquid layer on a homogeneous liquid half-space. This model, despite its simplicity, proves to be very useful in some cases (e.g., shallow sea) and has been the subject of many studies. The calculation of the eigenvalues of the corresponding boundary value problem reduces to solving the transcendental equation

$$mz \cot z = \sqrt{a^2 - z^2},$$

where $m = \rho_1/\rho$, ρ and ρ_1 are the densities of the liquid layer and the lower half-space, respectively, $a = kh\nu$, h is the layer thickness, $k = \omega/c$ is the wave number of the layer, ω is the angular frequency of the source, c is the speed of sound in the layer, c_1 is the speed of sound in the bottom, $\nu^2 = 1 - c^2/c_1^2$. Throughout, it is assumed that $\rho < \rho_1$, $c < c_1$, and the wave number for the bottom $k_1 = \omega/c_1$ is real. Similar equations arise in other problems as well.

Depending on the method of calculating the integral representing the sound pressure field, it is necessary to consider either only the real roots lying on the so-called physical sheet of the Riemann surface of the radical (they generate normal, or propagating modes, and it is these roots that correspond to the eigenvalues of the problem), or also the complex roots lying on the non-physical sheet of the Riemann surface (virtual modes or resonances). The roots of the transcendental equation were previously investigated by Heifetz using asymptotic and numerical methods, and its parametric solution is known. However, the development of algorithms for the numerical solution of this equation, in particular for calculating complex roots, continued to attract the attention of researchers. This work is devoted to refining the location of the roots of the equation, deriving explicit formulas for all roots, and their investigation. Heifetz considered the case $a > 0$ and $m > 1$. The case $0 < m < 1$ can be studied similarly.

Let us describe the location of the roots of the equation. It is obvious that there are no roots on the rays $x = \operatorname{Re} z > a$ and $x < -a$, so to isolate the single-valued branches of the radical, it is convenient to make cuts along these rays. The complex plane thus cut is denoted by D_0. We denote by $(\sqrt{a^2 - z^2})_\pm$ the single-valued and continuous branches of the radical in D_0, defined by the conditions $(\sqrt{a^2 - z^2})|_{z=0} = \pm a$, and consider in D_0 two holomorphic functions $f_\pm(z) = (\sqrt{a^2 - z^2})_\pm \sin z - mz \cos z$. The roots of $f_-(z)$ lie on the physical sheet, and the roots of $f_+(z)$ on the non-physical sheets of the Riemann surface.

To solve the two equations $f_\pm(z) = 0$ in D_0, we transform the original transcendental equation to the form

$$z = \frac{1}{2i} \ln \frac{\sqrt{a^2 - z^2} + imz}{\sqrt{a^2 - z^2} - imz} + \pi j, \quad j = 0, \pm 1, \pm 2, \dots .$$

Thus, we replace our original equation with a countable set of equations

$$l_j^\pm(z) = 0, \quad j = 0, \pm 1, \dots ,$$

where

$$l_j^\pm(z) = z + \frac{i}{2} \ln \frac{(\sqrt{a^2 - z^2})_\pm + imz}{(\sqrt{a^2 - z^2})_\pm - imz} - \pi j.$$

It is easy to see that on the rays $z = iy$, $|y| > \frac{a}{\sqrt{m^2-1}}$, the original equation also has no roots. We denote the region D_0, cut along these rays, by D, and select the branch of the logarithm in D by the condition $\ln 1 = 0$ (the choice of any other branch of the logarithm only leads to a renumbering of the roots). In the simply connected region D, $l_j^{\pm}(z)$ are single-valued analytic functions, the number of whose roots can be found by the argument principle. We find that on the physical sheet, the functions $l_j^-(z)$ for $0 < |j| < \frac{a}{\pi} + \frac{1}{2}$ have one real root x_j^{v-}, with $x_0^{v-} = 0$ and $x_{-j}^{v-} = -x_j^{v-}$, $j > 0$, while for $|j| \geq \frac{a}{\pi} + \frac{1}{2}$, $l_j^-(z)$ have no roots. In other words, the original physical problem has $[a/\pi + 1/2]$ propagating modes. (Here $[x]$ denotes the integer part of the number x.)

On the non-physical sheet, the functions $l_j^+(z)$ for $1 \leq |j| < \frac{a}{\pi} + \frac{1}{2}$ have one real root x_j^+, with $x_{-j}^+ = -x_j^+$, $j > 0$, and for $|j| > \frac{a}{\pi} + \frac{3}{2}$, $l_j^+(z)$ have two complex-conjugate roots $z_j^{\pm}$, $\bar{z}_j^{\pm}$, with $\operatorname{Im} z_j^+ > 0$, $z_{-j}^{\pm} = -\bar{z}_j^{\pm}$, $j > 0$. The function $l_0^+(z)$ has three roots for $a \geq \pi/2$: $z_0 = 0$ and symmetric purely imaginary roots $\pm iy_0$, $y_0 > 0$, and for $0 < a < \pi/2$ – five roots: $z_0 = 0$ and $\pm z'$, $\pm \bar{z}'$. Finally, the function $l_{j_a}^+(z)$, where $j_a = [a/\pi + 1/2]$ (the case of $l_{j_a}^-(z)$ is analogous), also has two roots, which for $a > a_{j_a} = \pi(j_a - 1/2)$ but close to a_{j_a}, are complex-conjugate, approach the real axis as a increases, and at some value of the parameter $a = a_{kr}$ merge to form a double "non-physical" real root x_{kr}, $x_{kr} > a$. With a further increase in a, it splits into two simple real roots, the larger of which moves towards the edge of the cut – the point $x = a$, reaches the edge at $a = \pi(j_a + 1/2)$ and then passes to the physical sheet – thus a new propagating mode is born. For the values a_{kr} and x_{kr}, Heifetz indicated that one can write down transcendental equations, which were not detailed there.

To obtain explicit formulas for the roots, another method is used: we consider only the roots $z_j^{\pm}$ of the functions $l_j^+(z)$, $|j| \geq a/\pi + 1/2$; formulas for all other roots are derived analogously. Regardless of whether $z_j^{\pm}$ is complex or real (for $j = [a/\pi + 1/2]$ and $a \geq a_{kr}$ – see above), it is a root of a quadratic trinomial $z^2 + p_j z + q_j$ with real coefficients. Consider the integrals

$$I_k = \oint_\gamma \frac{z^2 + p_j z + q_j}{z^k l_j^{+2}(z)} dz, \quad k = 3, 4,$$

where γ is a closed contour in D containing the point $z = 0$ inside. By calculating these

integrals, first, by the residue theorem and, second, by deforming the contour γ so that it coincides with the boundary of the region D, we obtain a system of two linear equations with respect to p_j and q_j, from which $z_j = 1/2(-p_j + \sqrt{p_j^2 - 4q_j})$, where

$$p_j = \frac{\delta_1\delta_2 - \delta_0\delta_3}{\delta_1\delta_3 - \delta_2^2}, \quad q_j = \frac{\delta_0\delta_2 - \delta_1^2}{\delta_1\delta_3 - \delta_2^2},$$
$$\delta_0 = d_0 + \pi j B_j - \pi j A_j, \quad \delta_1 = d_1 - \frac{1}{\pi j}C_j + \frac{1}{\pi j}D_j,$$
$$\delta_2 = d_2 - jE_j - \pi j F_j, \quad \delta_3 = d_3 - \frac{1}{\pi j}G_j - \frac{1}{\pi j}H_j.$$

The coefficients $d_k, A_j, \dots, H_j$ are defined through integrals:

$$d_0 = -\frac{1}{\pi j}, \quad d_1 = \frac{m-a}{a\pi^2 j^2}, \quad d_2 = -\frac{(a-m)^2}{a\pi^3 j^3},$$
$$d_3 = \frac{m(3-2m^2)}{6a\pi^2 j^2} - \frac{(a-m)^3}{a^3\pi^4 j^4}.$$

$$A_j = \int_a^\infty \frac{(2x-\pi)f(x)dx}{xg(x)}, \quad C_j = \int_a^\infty \frac{f(x)h(x)dx}{x^2 g(x)},$$
$$E_j = \int_a^\infty \frac{(2x-\pi)f(x)dx}{x^3 g(x)}, \quad G_j = \int_a^\infty \frac{f(x)h(x)dx}{x^4 g(x)},$$

where

$$f(x) = \ln\frac{mx + \sqrt{x^2-a^2}}{mx - \sqrt{x^2-a^2}}, \quad h(x) = \left(x - \frac{\pi}{2}\right)^2 + \pi^2 j^2 + \frac{1}{4}f^2(x),$$
$$g(x) = h^2(x) - 4\pi^2 j^2 (x - \pi/2)^2.$$

$$B_j = 2\int_{a/\sqrt{m^2-1}}^\infty \frac{b(y)dy}{yd(y)}, \quad D_j = \int_{a/\sqrt{m^2-1}}^\infty \frac{c(y)dy}{yd(y)},$$
$$F_j = 2\int_{a/\sqrt{m^2-1}}^\infty \frac{b(y)dy}{y^3 d(y)}, \quad H_j = \int_{a/\sqrt{m^2-1}}^\infty \frac{c(y)dy}{y^4 d(y)},$$

where

$$b(y) = y - \frac{1}{2}\ln\frac{my + \sqrt{a^2+y^2}}{my - \sqrt{a^2+y^2}},$$
$$c(y) = b^2(y) - \pi^2(j^2 - 1/4), \quad d(y) = c^2(y) + 4\pi^2 j^2 b^2(y).$$

These integrals converge slowly, but changes of variables $t = 1/\sqrt{2}f(x)$ or $t = \frac{1}{2}\ln\frac{my+\sqrt{a^2+y^2}}{my-\sqrt{a^2+y^2}}$ transform them into integrals with an exponentially decaying integrand, the calculation of which on a computer poses no difficulties. This method does not require the choice of an initial approximation and provides the same accuracy for roots with any numbers.

Table 1: Values of resonances and absorption coefficients

j	ξ_j	$\bar{\xi}_j$	$\bar{B}_j$	B_j	B_{jT}
6	$0.365801 + 0.730159 \cdot 10^{-3}i$	$0.3659 + 0.68 \cdot 10^{-3}i$	6.342	5.90	6.40
7	$0.346401 + 1.911244 \cdot 10^{-3}i$	$0.3464 + 1.92 \cdot 10^{-3}i$	16.6009	16.7	16.8
8	$0.322795 + 2.903837 \cdot 10^{-3}i$	$0.3228 + 2.90 \cdot 10^{-3}i$	25.2224	25.2	25.7

The table presents the values of the first three resonances $\xi_j = \frac{1}{h\sqrt{(kh)^2-z_j^2}}$ of the considered model, calculated using the formulas above and rounded to 10^{-8}, for the parameter values $h = 100$ m, $m = 2$, $c = 1500$ m/s, $c_1 = 1700$ m/s, $\omega = 2\pi f$, $f = 100$ Hz. They correspond to the complex roots z_j of equation (1) with numbers $j = 6, 7, 8$. We denote by $B_j = (20\lg e)\operatorname{Im}\xi_j$ the absorption coefficient of the j-th mode. For comparison, the table also shows $\bar{\xi}_j$ and $\bar{B}_j$ – the corresponding values of the same quantities, calculated by another method in other works. In addition to numerical calculations, the formulas obtained allow for the study of the asymptotic behavior of the roots. Along with the known asymptotics of large-modulus complex roots $z_j \approx \pi j + \pi/2 + (i/2)\ln\frac{m+1}{m-1} + O(1)$, $j \to \pm\infty$, we obtain, for example, for the roots corresponding to propagating modes, the relation $x_j \approx \pi j + (m/a)\pi j + O(a^{-2})$ as $a \to +\infty$ and for a fixed j, $1 \le j < a/\pi + 1/2$, as well as an asymptotic for complex roots useful in the case of a shallow sea, when there are no normal modes, as $a \to 0$ and for a fixed j: $z_j \approx \pi j + \pi/2 + (i/2)\ln\frac{m+1}{m-1} + O(a)$, with the estimate $O(a)$ being uniform in $j = 1, 2, \ldots$.

A.2 Sturm-Liouville Operators

In a related work, A. Elishev [2] described another way to prove the reality of the roots of the equation $\tan(z) - z = a$. We summarize his argument here, with the author's

permission.

First, it is necessary to prove that all zeros of $\phi(z) = \tan(z) - z$ are in fact given by the square roots of the eigenvalues of some regular Sturm-Liouville operator. Indeed, let us define the space $L^2([0,1], dx)$ as the space of complex-valued square-integrable functions defined on the unit interval $[0,1]$, with the inner product given by

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}dx.$$

Suppose that H_0 is the subspace of $L^2([0,1], dx)$ consisting of twice-differentiable functions f for which f' and f'' are square-integrable, and such that

$$f(0) = 0$$

and

$$f(1) - f'(1) = 0$$

Then the operator $H = -d^2/dx^2$, defined on H_0 is a positive semi-definite operator corresponding to a regular Sturm-Liouville problem with separated boundary conditions. Consequently, by a fundamental result of Sturm-Liouville theory, together with positive semi-definiteness, the eigenvalues of H form an increasing sequence.

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots .$$

Note that each eigenvalue has multiplicity one, and also that H is self-adjoint — this can be checked directly by integration by parts. We establish its positive semi-definiteness.

The explicit form of the eigenfunctions (unique up to a scalar factor for each eigenvalue) is a trivial exercise. First, we obtain for $\lambda_0 = 0$

$$\psi_0(x) = x.$$

Then, for $\lambda \neq 0$, $\lambda = k^2$, the boundary value problem is

$$\psi'' + k^2\psi = 0, \quad 0 < x < 1,$$
$$\psi(0) = 0,$$
$$\psi(1) - \psi'(1) = 0.$$

The general solution is given by $\psi(x) = A\cos(kx) + B\sin(kx)$. The first boundary condition implies $A = 0$, while the second yields the equation

$$\sin(k) - k\cos(k) = 0$$

or (if $\cos(k) \neq 0$)

$$\tan(k) - k = 0$$

Consequently, the spectrum of H is given by the set

$$\{0\} \cup \{\lambda \in \mathbb{C} : \tan(\sqrt{\lambda}) - \sqrt{\lambda} = 0\}.$$

Now, since H is self-adjoint, its spectrum is real, which means that any solution to

$$\tan(z) - z = 0$$

is either real or purely imaginary, $z = iy$. The purely imaginary case leads to the equation for $y \in \mathbb{R}$:

$$\tanh(y) - y = 0$$

which has no (real) solutions other than $y = 0$. Thus, the proof of the lemma is complete.

Elishev notes that each positive eigenvalue λ corresponds to two solutions of the equation, given by $+\sqrt{\lambda}$ and $-\sqrt{\lambda}$. This is expected since $\phi(z)$ is an odd function.

The situation described in this comment is typical for the realization of the zeros of meromorphic functions as spectra of self-adjoint operators. Often they correspond to the Hamiltonians of quantum-mechanical systems; several other transcendental equations can be handled in this way.

The example provided here by Elishev (along with other easily obtainable examples of meromorphic functions whose zeros are spectra of self-adjoint operators) is quite straightforward; at the same time, on the opposite end of the complexity scale lies, as is well known, the conjectured spectral realization of the non-trivial zeros of the Riemann zeta function. In fact, as Elishev notes, there is a significant amount of evidence in favor of a spectral approach to the Riemann hypothesis, developed, in particular, in the deep

and far-reaching work of Alain Connes and his school. Simple examples, such as the one considered in this comment, can, in Elishev's opinion, serve as illustrations of the powerful methods of operator theory and its potential applications to long-standing open problems.

References

[1] A. Belov-Kanel, A. Malistov, and R. Zaytsev. Solvability of equations in elementary functions. *Journal of Knot Theory and Its Ramifications*, 29(02):2040005, 2020. 125

[2] Alexei Elishev. A brief comment on the paper "solvability of equations in elementary functions" by kanel-belov, malistov and zaytsev. *Journal of Knot Theory and Its Ramifications*, 29(02):2040007, 2020. 124, 136

[3] A. I. Heifetz. On the roots of a transcendental equation in the problem of shallow sea acoustics. *Akusticheskii Zhurnal*, 31(2):258–263, 1985. (in Russian). English translation in: Soviet Physics, Acoustics, 31(2), 1985. 124, 132

[4] A. G. Khovanskii. *Topological Galois Theory. Solvability and Unsolvability of Equations in Finite Terms*. MCCME, Moscow, 2008. (in Russian). 124

[5] Felix Klein. *Lectures on the Development of Mathematics in the 19th Century*, volume 1. Nauka, Moscow, 1989. (in Russian). 124

[6] R. V. Zaytsev. Topological theory of solvability of equations in elementary functions. Master's thesis, Higher School of Economics, 2021. (in Russian). 128

AUTHORS

Allan Allemand
Lomonosov Moscow State University,
Higher School of Economics
Suleyman Rustam, 33A, 64
1022 Baku, Azerbaijan
email: allansuleykin@gmail.com

Alexey Kanel-Belov
Bar-Ilan University
Ramat Gan, 5290002
52900 Ramat Gan, Israel
email: kanelster@gmail.com

Rodion Zaytsev
California Institute of Technology
1200 E California Blvd, Pasadena, CA
91125, United States
email: r.zayt@mail.ru

Arnold Mathematical Journal
Volume 12, Issue 1, 2026

Morse-Bott Volume Forms

Boris Khesin Luke Volk

Received 11 Jun 2025; Accepted 24 Aug 2025

Abstract: A Morse-Bott volume form on a manifold is a top-degree form which vanishes along a non-degenerate critical submanifold. We prove that two such forms are diffeomorphic (by a diffeomorphism fixed on the submanifold) provided that their relative cohomology classes with respect to the submanifold coincide. For a zero submanifold of codimension at least 2, this means that two Morse-Bott volume forms with the same zero set are diffeomorphic if and only if they have equal total volumes. We show how "Moser's trick" for establishing equivalence of non-degenerate volume forms can be adapted to this setting.

AMS Classification: 58K50, 57R70, 58A10, 28A75

Key words and phrases: Morse-Bott function, volume form, Moser's method, Euler vector field, Madelung transform

1 Introduction

Background

In his 1965 paper [14], Moser showed that any two volume forms η_0, η_1 with the same total volume on a compact connected oriented manifold M are related by a diffeomorphism Φ of M via pullback, $\Phi^*\eta_1 = \eta_0$. To construct such a diffeomorphism,

DOI: 10.56994/ARMJ.012.001.005

Moser's method was to connect the forms η_0 and η_1 by a path $(\eta_t)_{t\in[0,1]}$ in the same cohomology class and to look for a family of diffeomorphisms $(\Phi_t)_{t\in[0,1]}$ such that $\Phi_0 = \mathrm{id}_M$ and $\Phi_t^*\eta_t = \eta_0$. The latter is achieved by solving the corresponding infinitesimal version of the equation on the vector field generating the flow Φ_t and invoking the existence theorem for ODEs guaranteeing the existence of a flow for a given vector field as well as verifying the conditions for its existence for all $t \in [0,1]$. The strategy has dubbed variably as "Moser's trick", "Moser's path method", or the "homotopy method".

This method has seen a wide variety of applications. Moser also applied the method in [14] to symplectic structures and twisted volume forms on non-orientable manifolds. Banyaga described Moser's approach for volume forms on manifolds with boundary in [2], while Bruveris *et al.* [4] extended it to volume forms on manifolds with corners. Cardona and Miranda [6] considered an analogue of Moser's result for equivalence of top-degree forms transverse to the zero section with a shared zero hypersurface. Other authors have considered solutions to the so-called "pullback equation" $\Phi^*\eta_1 = \eta_0$ in more analytic contexts, see e.g. a summary of equivalence results for k-forms for any k for Hölder spaces in [7].

Main result

Let M be a compact connected oriented n-dimensional manifold, which we equip with a reference (non-vanishing) volume form μ. In this paper, we consider volume forms on M which have a quadratic degeneration along an oriented submanifold $\Gamma \subset M$.

Definition 1.1. A *Morse-Bott volume form* for Γ on M is a non-negative n-form $\eta \in \Omega^n(M)$ with zero set Γ such that the ratio of n-forms $f \stackrel{\text{def}}{=} \eta/\mu$ is a Morse-Bott function $f : M \to \mathbb{R}$ for which each component of Γ is a non-degenerate critical submanifold.

Note that the critical zero set Γ must have Morse-Bott index 0 since the function f is non-negative on M. Furthermore, the Morse-Bott property of η does not depend on choice of the reference form μ. We prove the necessary and sufficient conditions for diffeomorphism equivalence of such Morse-Bott volume forms:

Theorem 1.2. *Let η_0 and η_1 be Morse-Bott volume forms for $\Gamma \subset M$ such that their*

relative cohomology classes with respect to Γ *coincide:*

$$[\eta_0] = [\eta_1] \in H^n(M, \Gamma)\,.$$

Then there exists a diffeomorphism $\Phi : M \to M$ *such that* $\Phi^*\eta_1 = \eta_0$ *and which restricts to the identity on* Γ*.*

We treat this as two different cases: when the submanifold Γ is a hypersurface, and when its codimension is at least 2.

Corollary 1.3. *If the shared zero submanifold* $\Gamma \subset M$ *of two Morse-Bott volume forms* η_0 *and* η_1 *on* M *is of codimension at least 2, the forms are diffeomorphic,*

$$\Phi^*\eta_1 = \eta_0 \text{ with } \Phi|_\Gamma = \mathrm{id}_\Gamma,$$

if and only if they have equal total volumes of M*,*

$$\int_M \eta_0 = \int_M \eta_1.$$

If Γ is a hypersurface in M, i.e. it has codimension 1, it can be separating or not. Either case is covered by the following corollary:

Corollary 1.4. *If the shared zero submanifold* $\Gamma \subset M$ *has codimension* $= 1$*, two Morse-Bott volume forms* η_0 *and* η_1 *are diffeomorphic,*

$$\Phi^*\eta_1 = \eta_0 \text{ with } \Phi|_\Gamma = \mathrm{id}_\Gamma,$$

if and only if they have coinciding volumes for each connected component M_i *of* $M \setminus \Gamma$*:*

$$\int_{M_i} \eta_0 = \int_{M_i} \eta_1 \quad \text{for all } i.$$

The same result also holds for volume forms which have hypersurface $\Gamma \subset M$ as a non-critical zero set. Let η_0 and η_1 be two n-forms on M with the same *non-critical zero set* Γ, i.e. it is a non-critical zero set for each of the corresponding functions η_i/μ, $i = 0, 1$. Note that Γ must be a compact oriented hypersurface in this case.

Theorem 1.5. *Two* n*-forms* η_0 *and* η_1 *with the same non-critical zero set* $\Gamma \subset M$ *are*

diffeomorphic,

$$\Phi^*\eta_1 = \eta_0 \text{ with } \Phi|_\Gamma = \mathrm{id}_\Gamma,$$

if and only if they represent the same relative cohomology classes $[\eta_0] = [\eta_1] \in H^n(M, \Gamma)$*, or equivalently, they have coinciding volumes of each connected component* M_i *of* $M \setminus \Gamma$.

This theorem strengthens one of the results of Cardona and Miranda [6], who proved that two folded volumes forms with the same non-critical zero hypersurface $\Gamma \subset M$ can be mapped to each other by a diffeomorphism taking Γ to itself, although not necessarily the identity on Γ. (Note that while n-forms change sign across the non-critical zero set $\Gamma \subset M$, the term *folded*, or *transversally vanishing*, "volume forms" became standard and we adapt it in this paper.)

Remark 1.6. The assumption of the orientability of M and Γ can be weakened to require only orientability of $M \setminus \Gamma$. For $\operatorname{codim}\Gamma \geq 2$ this reduces to orientability of M. In the case of $\operatorname{codim}\Gamma = 1$ in a nonorientable M, instead of volume forms one needs to consider densities, or pseudo-forms changing sign along orientation-reversing paths. Theorem 1.2 naturally extends to this setting: its proof combines tubular neighbourhood embeddings, which hold in any setting, with adapting the classical Moser theorem to each (orientable) connected component of $M \setminus \Gamma$. (As an example, it is easy to construct a Morse-Bott density on the even-dimensional real projective space with a hyperplane as a critical set: for instance take the product of function x_1^2 with the standard volume element on the sphere S^{2k} and project it to $\mathbb{RP}^{2k}$ via the antipodal map.) An extension of Theorem 1.5 to the nonorientable setting might be more subtle, requiring certain averaging on the orientation cover, cf. normal forms for Morse functions and densities in the $n = 2$ area-preserving case in [10].

Motivation

A motivation for this problem comes from the Madelung transform, which establishes an equivalence of quantum mechanics and equations of compressible fluids [12]. Namely, let a wave function $\psi : M \to \mathbb{C}$ on a manifold M satisfy the non-linear Schrödinger (NLS) equation,

$$i\partial_t\psi + \Delta\psi + V\psi + f(|\psi|^2)\psi = 0,$$

where $V : M \to \mathbb{R}$ and $f : \mathbb{R}_+ \to \mathbb{R}$. Then the *Madelung transform* $\psi = \sqrt{\rho} e^{i\theta}$ allows one to rewrite the quantum mechanics of the NLS equation in a "hydrodynamical form" as equations of a barotropic-type fluid on the velocity field $v \overset{\text{def}}{=} \nabla\theta$ and the density ρ as follows:

$$\begin{cases} \partial_t v + \nabla_v v + \nabla\Big(V + f(\rho) - \frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Big) = 0 \\ \partial_t \rho + \operatorname{div}(\rho v) = 0\,. \end{cases}$$

The Madelung transform $\psi \longmapsto (\rho, \theta)$ is well-defined provided that ψ does not vanish on M and it is understood modulo a phase factor ($\psi \sim \psi e^{i\alpha}$), while θ is understood to be modulo an additive constant on M. Moreover, by confining to the unit sphere of normalized wave functions ψ and the space $\operatorname{Dens}(M)$ of normalized densities ρ, the Madelung transform can be understood as the map $\mathbb{CP}^\infty(M, \mathbb{C} \setminus 0) \to T^* \operatorname{Dens}(M)$. It turns out to be a symplectomorphism for the corresponding natural symplectic structures on those spaces, and a Kähler map between the Fubini-Study and Fisher-Rao metrics respectively, see [12].

However, the presence of zeros of the (complex-valued) wave function ψ brings substantial complications. A non-critical zero set Γ of ψ has codimension 2 in M, and the corresponding density function ρ can be understood as a Morse-Bott volume form for $\Gamma \subset M$. The fact that ψ is univalued on M imposes the "quantization constraint" on the phase function θ: its change along any path in M going around Γ must be a multiple of 4π, see numerous discussions in [9, 15]. The above equivalence theorems for the Morse-Bott volume forms allow one to deal with zero submanifolds of wave functions by using more convenient "normal forms" of the corresponding densities around zeros.

2 The Morse-Bott Lemma

Morse-Bott functions have local normal forms which will allow us to more easily handle behaviour near the zero set of Morse-Bott volume forms. A ("local") normal form in a neighbourhood of a point can be found, e.g., in [3]. Below we outline a

proof of a ("semi-global") normal form in a neighbourhood of the critical set Γ using Euler-like vector fields, following [13].

Euler-Like vector fields

Given a submanifold $\Gamma \subseteq M$ of codimension k, we denote the *normal bundle* of Γ in M by:

$$\nu(M,\Gamma) \stackrel{\text{def}}{=} TM|_\Gamma / T\Gamma.$$

Morphisms between pairs (M,Γ) and (M',Γ') are smooth maps $f:\ M \to M'$ taking Γ to Γ'. Given a morphism $f:\ (M,\Gamma) \to (M',\Gamma')$, we associate to it the linear map $\nu(f)$ defined as follows:

$$\begin{aligned} \nu(f):\ \nu(M,\Gamma) &\to \nu(M',\Gamma') \\ v + T\Gamma &\longmapsto f_* v + T\Gamma', \end{aligned}$$

which we call the *linearisation* of f.

The *Euler vector field to* Γ is the vector field $\mathcal{E}$ on the normal bundle $\nu(M,\Gamma)$ which is the Euler vector field in the usual sense on each fibre. That is, if $x \in \Gamma$ and the fibre $\nu(M,\Gamma)_x$ is given the coordinates y^i, then:

$$\mathcal{E}_x = \sum_{i=1}^{k} y^i \frac{\partial}{\partial y^i}.$$

If a vector field X on M is tangent to Γ (i.e., for each $p \in \Gamma$, $X_p \in T_p\Gamma$) then X can be seen as a morphism $X:\ (M,\Gamma) \to (TM, T\Gamma)$ of pairs. We say that X is *Euler-like* for Γ if its linearisation,

$$\nu(X):\ \nu(M,\Gamma) \to \nu(TM,T\Gamma) \cong T\nu(M,\Gamma),$$

is the Euler vector field to Γ.

Example 2.1. For $M = \mathbb{R}^n$ and $\Gamma = \{0\}$, we have that $\nu(M,\Gamma) = T_0\mathbb{R}^n$ and a vector field

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i}$$

is Euler-like if for all $1 \le i \le n$ we have $X^i(0) = 0$ and $DX^i|_0 = x^i$.

Tubular neighbourhood embeddings

A *tubular neighbourhood embedding* of Γ is a neighbourhood $U \subseteq \nu(M,\Gamma)$ of the zero section in the normal bundle and an embedding $\varphi : U \to M$ such that:

(i) For each $x \in \Gamma$, $\varphi(0_x) = x$. That is, $\varphi|_\Gamma = \mathrm{id}_\Gamma$ after identifying Γ with the zero section in $\nu(M,\Gamma)$.

(ii) The linearisation $\nu(\varphi) : \nu(U,\Gamma) \cong \nu(M,\Gamma) \to \nu(M,\Gamma)$ is the identity, $\mathrm{id}_{\nu(M,\Gamma)}$.

The benefit of Euler-like vector fields is their correspondence with tubular neighbourhood embeddings as the following theorem summarizes:

Theorem 2.2. *An Euler-like vector field X for (M,Γ) determines a unique maximal (with respect to inclusion) tubular neighbourhood embedding $\varphi : U \to M$ of Γ with $U \subseteq \nu(M,\Gamma)$ such that $\varphi^* X = \mathcal{E}$.*

We refer for the proof to [5].

Fibre-wise polynomial functions

We say that $f : M \to \mathbb{R}$ is *Morse-Bott for* $\Gamma \subset M$ if Γ is a non-degenerate critical submanifold of f. Without loss of generality, we assume $f|_\Gamma = 0$. The Euler vector field to Γ gives a handy method for identifying fibre-wise homogeneous polynomials on $\nu(M,\Gamma)$.

Proposition 2.3. *Let $\mathcal{E}$ be the Euler vector field to Γ and $f \in C^\infty(\nu(M,\Gamma))$ a function on $\nu(M,\Gamma)$. If $\mathcal{L}_\mathcal{E} f = kf$ for some $k \in \mathbb{N}$, then f is fibre-wise a homogeneous polynomial of order k.*

Proof. This is a fibre-wise application of Euler's homogeneous function theorem. ∎

The following proof of the Morse-Bott lemma (as sketched in [13]) makes use of Euler-like vector fields, and can be regarded as a semi-global, fibre-wise version of the Morse lemma with parameters.

Theorem 2.4 (Morse-Bott lemma [13])**.** *If $f : M \to \mathbb{R}$ is Morse-Bott for (M,Γ) and $f|_\Gamma = 0$, then there exists a tubular neighbourhood embedding $\varphi : U \to M$ (with $U \subseteq \nu(M,\Gamma)$) such that $\varphi^* f$ is fibre-wise a homogeneous polynomial of degree 2 (i.e. fibre-wise quadratic).*

Proof. Without loss of generality, take M to be a tubular neighbourhood of Γ, so M sits inside $\nu(M,\Gamma)$. Because f is Morse-Bott, for each $x \in \Gamma$, the functions

$$g_x \overset{\text{def}}{=} f|_{\nu(M,\Gamma)_x \cap M} : \nu(M,\Gamma)_x \cap M \to \mathbb{R}$$

are *Morse* functions, each with a non-degenerate critical point at 0_x, where the Hessian $Hg_x|_{0_x}$ is non-degenerate. In coordinates y^i on the fibre $\nu(M,\Gamma)_x$ near 0_x (we suppress the subscript x),

$$g(y) = \frac{1}{2}\sum_{i,j} A_{ij}(y) y^i y^j,$$

where $A(y) = (A_{ij}(y))$ is a symmetric matrix-valued function such that $A(0) = Hg|_{0_x}$. We compute:

$$\frac{\partial g}{\partial y^k} = \left(\frac{1}{2}\sum_{i,j} y^i y^j \frac{\partial A_{ij}}{\partial y^k}\right) + \sum_i A_{ik} y^i = \sum_i B_{ki} y^i,$$

here $B_{ki} = A_{ik} + \frac{1}{2}\sum_j \frac{\partial A_{ij}}{\partial y^k} y^j$. Note that $Hg|_{0_x} = A(0) = B(0)$, and so the matrix-valued function B is invertible in a neighbourhood of $y = 0$. We will now construct an Euler-like vector field on $\nu(M,\Gamma)_x$, analogous to the construction in the proof of the Morse lemma in [13]. Let X be a vector field (implicitly depending on x, more accurately notated X_x) in this neighbourhood near zero by:

$$X = \sum_{i,j} (AB^{-1})_{ij}(y)\, y^i \frac{\partial}{\partial y^j}.$$

Since $A(0)B^{-1}(0) = I$, near zero $\nu(X) = \mathcal{E}$, and so X is Euler-like. Then we have:

$$X(g) = \sum_{i,j} (AB^{-1})_{ij}(y) y^i \frac{\partial g}{\partial y^j} = \sum_{i,j,k} (AB^{-1})_{ij}(y) B_{jk}(y) y^i y^k = \sum_{i,k} (AB^{-1}B)_{ik}(y) y^i y^k = 2g.$$

Now define a vector field Y on M by $Y(x,y) = X_x(y)$. By construction, Y is Euler-like and, by Theorem 2.2, Y determines a tubular neighbourhood embedding $\varphi : U \to M$ such that $\varphi^* Y = \mathcal{E}$. Note that for each $x \in \Gamma$, the vector field X_x was defined to be tangent to the fibre at x, and so Y only flows along the fibres, its flow being $\phi_t(x,y) = (x, \phi_t^{X_x}(y))$, where $\phi_t^{X_x}$ is the flow of X_x. We then can compute that:

$$(\mathcal{L}_Y f)(x,y) = \left.\frac{d}{dt}\right|_{t=0} f(x, \phi_t^{X_x}(y)) = \left.\frac{d}{dt}\right|_{t=0} g_x(\phi_t^{X_x}(y)) = (\mathcal{L}_{X_x} g_x)(y) = 2g_x(y) = 2f(x,y),$$

hence applying φ^* to both sides yields $\mathcal{L}_\mathcal{E}(\varphi^* f) = 2\varphi^* f$. By Proposition 2.3, $\varphi^* f$ must be fibre-wise a homogeneous polynomial of degree 2. ■

In our case, where the Morse-Bott functions associated with Morse-Bott volume forms necessarily have index 0 (i.e. correspond to positive-definite quadratic forms), we have a convenient normal form:

Theorem 2.5. *If $f : M \to \mathbb{R}$ is Morse-Bott for $\Gamma \subset M$ of index 0 and $f|_\Gamma = 0$, then there exist coordinates (x, y) in the tubular neighbourhood U of Theorem 2.4 (x parametrising Γ, and y the fibres) such that:*

$$(\varphi^* f)(x, y) = |y|^2.$$

Proof. This theorem is equivalent to the existence of *bundle metrics* (also known as *Euclidean metrics*, see [11]) on the normal bundle. It is based on the partition of unity and the fact that the space of positive definite quadratic forms in n-variables is a convex cone. This allows one to combine the diffeomorphism φ constructed in Theorem 2.4 taking f to its quadratic part with a fibre-wise linear map L, so that the composition $L \circ \varphi$ takes f to the fibre-wise quadratic function given by the length-squared in the fibre. ■

Corollary 2.6. *Suppose that f_0 and f_1 are Morse-Bott functions for $\Gamma \subset M$, both with a Morse-Bott index of 0. Then there exists a neighbourhood U of Γ and a diffeomorphism φ defined on U such that $\varphi^* f_1 = f_0$.*

The analogue of the corollary for a maximal Morse-Bott index $k = \operatorname{codim}(\Gamma)$ follows similarly. A generalization of this result for any index is proven for fibre-wise quadratic functions on general vector bundles in [8] subject to the constraint that the positive- and negative-definite parts of f_0 and f_1 give the same splittings of the vector bundle.

Example 2.7. Note that the existence of the universal semi-global normal form in Corollary 2.6 for Morse-Bott functions of index 0 is based on the contractibility of the cone of symmetric positive definite matrices, allowing one to connect any two such functions in a tubular neighborhood of their critical set and apply Moser's trick (discussed in more detail in the next section in the context of forms). Morse-Bott functions of non-max/minimal indices might be non-isotopic, as can be seen in the

following example.

The space of non-degenerate quadratic forms $ax^2 + 2bxy + cy^2$ in two variables is split into three components by the double cone $b^2 - ac = 0$ in the 3D space $(a, b, c) \in \mathbb{R}^3$, see Figure 1. This allows the following simple construction of a pair of non-isotopic Morse-Bott functions, as they realize contractible and non-contractible loops in the set of forms of index 1.

Let $M = S^2 \times S^1 \subset \mathbb{R}^3 \times S^1$ be a 3-manifold regarded as a bundle over $S^1 = \{\theta \mod 2\pi\}$, where $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. As one of the functions one can take the restriction $g|_{S^2 \times S^1}$ of the function $g(x, y, z, \theta) = x^2 - y^2$ from $\mathbb{R}^3 \times S^1$ to M, independent of the variables z and θ. Its critical set $\Gamma \stackrel{\text{def}}{=} \{(x, y, z) \in \mathbb{R}^3 \mid x = y = 0, z = \pm 1\} \subset M$ consists of two (north and south) circles, both having the Morse-Bott index 1.

The other function is also defined by the restriction $f|_{S^2 \times S^1}$ of a fibre-wise quadratic f in $\mathbb{R}^3 \times S^1$, where:

$$f(x, y, z, \theta) = (x^2 - y^2) \cos\theta + 2xy \sin\theta.$$

For each θ, the restriction of $f(\cdot, \theta)$ to the sphere has non-degenerate critical points of index 1 at the north and south poles of the fibre S^2 and it defines a non-contractible loop in the space of quadratic forms, as illustrated in Figure 1. As a result, there is no isotopy between f and g.

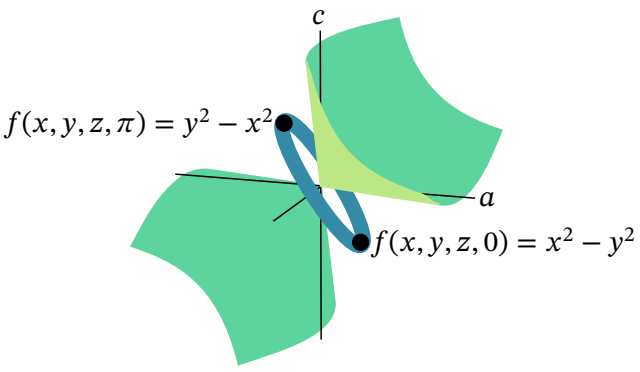


Figure 1: Example of a non-trivial loop in the space of quadratic forms $ax^2+2bxy+cy^2$, as parametrized by (a, b, c).

3 Proofs of main results

In this section, we will prove Theorems 1.2 and 1.5 with their corollaries.

Proof of Theorem 1.2 on Morse-Bott volume forms

Assume that η_0 and η_1 are Morse-Bott volume forms for $\Gamma \subset M$ such that their relative cohomology classes with respect to Γ coincide, $[\eta_0] = [\eta_1] \in H^n(M, \Gamma)$. We will show that there exists a diffeomorphism $\Phi : M \to M$ such that $\Phi^* \eta_1 = \eta_0$ which restricts to the identity on Γ.

Proof. First consider the local problem. Let N be a tubular neighbourhood of Γ, which we identify with a neighbourhood in the normal bundle, $\nu(M, \Gamma) \subseteq M$. Two Morse-Bott volume forms in N can be expressed as $\rho_0 \stackrel{\text{def}}{=} f\mu$ to $\rho_1 \stackrel{\text{def}}{=} h\mu$, where f and h are Morse-Bott functions having Γ as a non-degenerate minimum (i.e. a critical submanifold of index 0) and μ is a reference (non-vanishing) volume form on $N \subset M$. The forms ρ_i can be thought of as the restrictions $\rho_i \stackrel{\text{def}}{=} \eta_i|_N$ of globally defined Morse-Bott forms η_i to the neighbourhood N.

By Corollary 2.6, there exists a diffeomorphism F (of a possibly smaller neighbourhood of Γ) taking h to f, but changing the reference form μ. So without loss of generality we assume that, after application of $F : N \to N$ the Morse-Bott volume forms are $\rho_0 = f\mu$ and $\rho_1 = f\phi\,\mu$ for some non-vanishing function $\phi \in C^\infty(N)$. Next, the function f can be assumed fibre-wise quadratic in $N \subset \nu(M, \Gamma)$ by Theorem 2.4. We are looking for a diffeomorphism of N pulling back ρ_1 to ρ_0 while remaining the identity on Γ.

To apply Moser's trick, we consider the interpolation $\rho_t \stackrel{\text{def}}{=} (1-t)\rho_0 + t\rho_1$ of these forms and seek a family of diffeomorphisms $(\psi_t)_{t\in[0,1]}$ such that $\psi_t^* \rho_t = \rho_0$. Applying $\psi_{t*}\frac{d}{dt}$ to this relation, we get that:

$$\mathcal{L}_{X_t}\rho_t + \dot{\rho}_t = 0,$$

where $(X_t)_{t\in[0,1]}$ is the time-dependent vector field whose flow is $(\psi_t)_{t\in[0,1]}$, and $\dot{\rho}_t = \rho_1 - \rho_0$. We will show that there exists such a smooth vector field X_t vanishing on Γ

for $t \in [0,1]$.

Let $\mathcal{E}$ be the Euler vector field to Γ, defined on $N \subset \nu(M,\Gamma)$, and let g_s be the flow of $-\mathcal{E}$ towards Γ. The following expression for a primitive for $\dot\rho_t$ is similar to the one in the proof of the Poincaré lemma:

$$\dot\rho_t = -\int_0^\infty \frac{d}{ds} g_s^*(\dot\rho_t)\, ds = -\int_0^\infty g_s^*\left(-\mathcal{L}_{\mathcal{E}}\dot\rho_t\right)\, ds$$
$$= \int_0^\infty g_s^*\left(d\,\iota_{\mathcal{E}}\dot\rho_t + \cancel{\iota_{\mathcal{E}} d\dot\rho_t}\right)\, ds = \int_0^\infty d\,\iota_{g_s^*\mathcal{E}}(g_s^*\dot\rho_t)\, ds.$$

Now note that $g_s^* f = f e^{-2s}$ since f is fibre-wise quadratic, while $\dot\rho_t = \rho_1 - \rho_0 = f(\phi\mu - \mu)$, hence:

$$\dot\rho_t = \int_0^\infty d\left(\iota_{g_s^*\mathcal{E}}(f e^{-2s} g_s^*(\phi\mu - \mu))\right)\, ds = d\,[f \underbrace{\int_0^\infty \iota_{g_s^*\mathcal{E}}(e^{-2s} g_s^*(\phi\mu - \mu))\, ds}_{-\beta}],$$

and therefore, locally near Γ, one has $-\dot\rho_t = d(f\beta)$ for some $(n-1)$-form β on N. Note that for all $p \in \Gamma$, we have that $g_s^*\mathcal{E}|_p = 0$, and so $\beta|_\Gamma = 0$.

On the other hand, $\mathcal{L}_{X_t}\rho_t = d\,\iota_{X_t}\rho_t$, while $\rho_t = f((1-t)\mu + t\phi\mu)$. Hence to solve the equation $\mathcal{L}_{X_t}\rho_t + \dot\rho_t = 0$ for the field X_t or, equivalently, $d\,\iota_{X_t}\rho_t = d(f\beta)$, it suffices to solve:

$$f\,\iota_{X_t}((1-t)\mu + t\phi\mu) = f\beta.$$

This amounts to solving the equation $\iota_{X_t}((1-t)\mu + t\phi\mu) = \beta$ for a family of vector fields X_t on N. Note that the volume form $(1-t)\mu + t\phi\mu$ interpolates between μ and $\phi\mu$ and hence it is non-vanishing for all $t \in [0,1]$. Hence the field X_t solving $\iota_{X_t}((1-t)\mu + t\phi\mu) = \beta$ exists on $N \setminus \Gamma$, and it is smooth and must vanish on Γ. Note also that due to this vanishing condition, solutions starting sufficiently close to Γ exist for the whole interval $t \in [0,1]$. Hence the time-1 map of the flow ψ_t corresponding to the vector field X_t provides the required diffeomorphism of some neighbourhood of Γ. Without loss of generality, we can assume that this is the neighbourhood $N \supset \Gamma$, and this completes the proof of the local statement.

To prove the existence of a smooth globally-defined field on M whose flow takes

η_1 to η_0, we first (smoothly and arbitrarily) extend the field X_t from N to the whole of M. Now consider a smaller tubular neighbourhood U of Γ, sitting compactly within N, $\Gamma \subset U \subset \overline{U} \subset N \subset M$, and pick a bump function $b : M \to [0,1]$ which is 1 on U and 0 on $M \setminus N$. This allows one to define the time-dependent vector field $Y_t \stackrel{\text{def}}{=} bX_t$ on M whose time-1 flow map $G : M \to M$ satisfies $G^*\eta_1|_U = \eta_0|_U$ and $G|_\Gamma = \mathrm{id}_\Gamma$.

Consider the pull-back action of the map G on the Morse-Bott form η_1: it is a new form $\zeta_1 \stackrel{\text{def}}{=} G^*\eta_1$ which coincides with η_0 in the neighbourhood $U \supset \Gamma$, but outside of U, the form ζ_1 is only known to be non-vanishing and, by assumption, representing the same relative cohomology class in $H^n(M,\Gamma)$ as the form η_0.

We will now apply Moser's method again to find a diffeomorphism mapping ζ_1 to η_0 everywhere on M. For this we consider the interpolation $\zeta_t \stackrel{\text{def}}{=} (1-t)\eta_0 + t\zeta_1$ between them, joining $\zeta_0 \stackrel{\text{def}}{=} \eta_0$ and ζ_1. Note that all these forms coincide in the tubular neighbourhood U, $\zeta_t|_U = \zeta_0|_U$ for all $t \in [0,1]$. We will be seeking a family of diffeomorphisms $(\varphi_t)_{t\in[0,1]}$ such that $\varphi_t^*\zeta_t = \zeta_0$. Applying $\varphi_{t*}\frac{d}{dt}$ to this relation, we get that:

$$\mathcal{L}_{Z_t}\zeta_t = \zeta_0 - \zeta_1,$$

where $(Z_t)_{t\in[0,1]}$ is the time-dependent vector field whose flow is $(\varphi_t)_{t\in[0,1]}$. Note that ζ_0 and ζ_1 represent the same class in $H^n(M,\Gamma)$ and $(\zeta_0 - \zeta_1)|_U = 0$. We wish to find a primitive $(n-1)$-form for $\zeta_0 - \zeta_1$ which is zero on U. This can be done in a number of ways, cf. [1, 4, 6, 7], for instance via the following consideration.

Since Γ is a deformation retract of its tubular neighbourhood U, the forms ζ_0 and ζ_1 represent the same relative cohomology class in $H^n(M,U) = H^n(M,\Gamma)$. By the definition of relative cohomology, there exists a $\omega \in \Omega^{n-1}(M)$ and $\theta \in \Omega^{n-2}(U)$ such that:

$$\zeta_0 - \zeta_1 = d\omega, \quad i^*\omega = d\theta,$$

for the inclusion $i : U \hookrightarrow M$. Pick a bump function $\tilde{b} : M \to [0,1]$ equal to 1 on a smaller tubular neighbourhood $\tilde{U}$ compactly contained in U and equal to 0 on $M \setminus U$. Then define:

$$\tilde{\omega} \stackrel{\text{def}}{=} \omega - d(\tilde{b}\theta) \in \Omega^{n-1}(M),$$

where $d\tilde{\omega} = d\omega = \zeta_0 - \zeta_1$ and $\tilde{\omega}|_{\tilde{U}} = 0$. To complete Moser's trick, we now want to solve

the equation $\iota_{Z_t}\zeta_t = \tilde{\omega}$ for an unknown vector field Z_t. The Morse-Bott form $\zeta_t|_{M\setminus\Gamma}$ is a volume form for all $t \in [0,1]$ (with $\zeta_t|_U = \zeta_0|_U$) and so Z_t has a unique solution on $M \setminus \Gamma$. Furthermore, the solution Z_t vanishes on the tubular neighbourhood $\tilde{U} \supset \Gamma$, since $\tilde{\omega}|_{\tilde{U}} = 0$. This allows one to define Z_t on the whole manifold M (extending it by zero to Γ itself). The corresponding flow of Z_t on the compact manifold M exists for $t \in [0,1]$, and it is the identity on $\tilde{U} \supset \Gamma$.

Thus Moser's trick yields that the time-1 flow map $H : M \to M$ is a diffeomorphism satisfying $H^*\zeta_1 = \zeta_0$ and $H|_\Gamma = \mathrm{id}_\Gamma$. Finally, we define the diffeomorphism Φ of M as the composition $\Phi \stackrel{\text{def}}{=} H\circ G\circ F$, where F (extended to M) is from the Morse-Bott lemma (Corollary 2.6), G identifies the Morse-Bott volume forms in a neighbourhood of Γ, and H relates the forms outside of a neighbourhood of Γ while keeping fixed the neighbourhood itself. The composition satisfies $\Phi^*\eta_1 = \eta_0$ and $\Phi|_\Gamma = \mathrm{id}_\Gamma$, as desired. ■

Note that the last part of the proof boils down to construction of a vector field with a prescribed divergence, while controlling its support outside of a neighbourhood of the critical set. This topic has a long history. In the C^∞-case for manifolds it was considered in [1]. The arguments above can be regarded as an extension of the boundary case in [2, 4].

We now turn to the corollaries of this theorem, are different manifestations of the theorem depending on the codimension of the critical submanifold Γ.

Corollary 1.3. *If the shared zero submanifold $\Gamma \subset M$ of two Morse-Bott volume forms η_0 and η_1 on M is of codimension at least 2, the forms are diffeomorphic, $\Phi^*\eta_1 = \eta_0$ with $\Phi|_\Gamma = \mathrm{id}_\Gamma$, if and only if they have equal total volumes of M,*

$$\int_M \eta_0 = \int_M \eta_1.$$

Proof. If $\Gamma \subset M$ has codimension $k \geq 2$, the condition of $[\eta_0] = [\eta_1] \in H^n(M,\Gamma)$ is equivalent to the condition that they have equal total volumes. This can be seen from the long exact sequence for relative cohomology:

$$\cdots \to H^{n-1}(\Gamma) \to H^n(M,\Gamma) \to H^n(M) \to H^n(\Gamma) \to 0,$$

where the constraint on the codimension implies $H^{n-1}(\Gamma) = H^n(\Gamma) = 0$. This implies an isomorphism $H^n(M, \Gamma) \cong H^n(M)$ by exactness of the sequence:

$$0 \to H^n(M, \Gamma) \to H^n(M) \to 0.$$

■

Corollary 1.4. *If the shared zero submanifold $\Gamma \subset M$ has codimension 1, two Morse-Bott volume forms η_0 and η_1 are diffeomorphic, $\Phi^*\eta_1 = \eta_0$ with $\Phi|_\Gamma = \mathrm{id}_\Gamma$, if and only if they have coinciding volumes for each connected component M_i of $M \setminus \Gamma$:*

$$\int_{M_i} \eta_0 = \int_{M_i} \eta_1 \quad \textit{for all } i.$$

Proof. If $\Gamma \subset M$ has codimension $k = 1$, one has only a surjection $H^n(M, \Gamma) \to H^n(M)$ from $H^n(\Gamma) = 0$. In general, it is not necessarily an isomorphism, due to the possible disconnectedness of $M \setminus \Gamma$. If $M \setminus \Gamma$ consists of several connected components, $M \setminus \Gamma = \bigsqcup_{i \in I} M_i$, the forms η_0 and η_1 represent the same cohomology class relative to Γ if and only they have equal volumes on each component M_i. Indeed, by definition of relative de Rham cohomology, the forms η_0 and η_1 represent the same cohomology class relative to Γ whenever their difference is exact, $d\omega = \eta_1 - \eta_0$ for some $\omega \in \Omega^{n-1}(M)$. Hence we have:

$$\int_{M_i} (\eta_1 - \eta_0) = \int_{M_i} d\omega = \int_{\partial M_i} \omega = 0\,,$$

where the last equality follows from the fact that ω is exact on $\partial M_i \subset \Gamma$. ■

Non-critical zero sets

A *folded volume form* on an oriented n-dimensional manifold M is a top-degree form $\eta \in \Omega^n(M)$ which is transverse to the zero section of the determinant bundle $\Lambda^n T^*M$. Here we outline how the following strengthening of the result of Cardona and Miranda [6] on the equivalence of two folded volume forms with the same zero sets can be proven using a similar strategy through Euler-like vector fields.

Note that the zero set $\Gamma \subset M$ of a folded volume form η is an oriented hypersurface Γ (possibly consisting of several components). By fixing a reference volume form, μ,

the hypersurface Γ has a *defining function* $\eta/\mu \stackrel{\text{def}}{=} f : M \to \mathbb{R}$ satisfying $\Gamma = f^{-1}(0)$ and $df|_x \neq 0$ for each $x \in \Gamma$. We have the following analogue of the Morse-Bott lemma (Theorem 2.4):

Lemma 3.1. *If $f : M \to \mathbb{R}$ is a defining function for a hypersurface Γ, then there exists a tubular neighbourhood embedding $\varphi : U \to M$ (with $U \subseteq \nu(M, \Gamma)$) such that $\varphi^* f$ is fibre-wise linear.*

Proof. This is achieved via the inverse function theorem by taking f as the coordinate in the fibres of the one-dimensional normal bundle. In a sense, this is Hadamard's lemma depending on the parameter $x \in \Gamma$. The Euler-like vector field of Theorem 2.4 becomes $X = f\frac{\partial}{\partial f}$ in this coordinate. ∎

Now we prove Theorem 1.5, that two folded volume forms η_0 and η_1 on a compact oriented manifold M with the same non-critical zero set $\Gamma \subset M$ are diffeomorphic, $\Phi^*\eta_1 = \eta_0$ with $\Phi|_\Gamma = \mathrm{id}_\Gamma$, if and only if they represent the same relative cohomology classes $[\eta_0] = [\eta_1] \in H^n(M, \Gamma)$, or equivalently, they have coinciding volumes of each connected component M_i of $M \setminus \Gamma$. The lemma above allows one to set up Moser's trick, much like how Corollary 2.6 was used in the proof of Theorem 1.2.

Proof. Consider the defining functions f_0 and f_1 of Γ corresponding to the folded volume forms η_0 and η_1 for a fixed volume form μ, i.e. $f_i = \eta_i/\mu$. Hadamard's lemma guarantees the existence of a non-vanishing $\phi \in C^\infty(M)$ such that $\eta_1 = \phi\,\eta_0$.

Taking a tubular neighbourhood $N \subset M$ of Γ (and identifying it with a neighborhood in $\nu(M, \Gamma)$), one comes to a setting analogous to the proof of Theorem 1.2: here we are constructing a diffeomorphism equivalence between two folded volume forms $\eta_0 = f\mu$ and $\eta_1 = f\phi\,\mu$ in the neighbourhood N, where μ is a volume form, ϕ is a non-vanishing function, and f is a defining function for Γ. Without loss of generality, we can assume f to be fibre-wise linear (otherwise applying Lemma 3.1 to make it so).

Now we are looking for a diffeomorphism of a neighbourhood $N \subseteq M$ of Γ pulling back η_1 to η_0 which is the identity on Γ. The rest of the proof using Moser's trick follows *mutatis mutandis* the proof of Theorem 1.2, except that now the function f is

fibre-wise *linear*, and hence the pullback $g_s^* f$ of f by the inverse flow g_s of the Euler vector field to Γ will be fe^{-s}, with the factor of 2 replaced by 1.

As before, the diffeomorphism of the neighbourhood N extends to a global diffeomorphism $G : M \to M$. Then the existence of a diffeomorphism H relating the form on $M \setminus N$ is based on the equality of the corresponding relative cohomology classes $[\eta_0] = [\eta_1] \in H^n(M, \Gamma)$, or equivalently, on the equality of the volumes of connected components M_i of $M \setminus N$. The desired diffeomorphism Φ is obtained by composing the corresponding diffeomorphisms G and H as in the proof of Theorem 1.2 (with the application of Hadamard's lemma instead if the diffeomorphism F). ■

Note that due to the local-to-global nature of the construction of the diffeomorphism, it immediately follows from the above that if η_0 and η_1 are top-degree forms with a shared zero set Γ such that each component of Γ is either Morse-Bott or non-critical (i.e. Γ is of mixed-type) and such that their relative cohomology classes in $H^n(M \setminus \Gamma)$ coincide, then there is a diffeomorphism $\Phi : M \to M$ such that $\Phi^* \eta_1 = \eta_0$ restricting to the identity on Γ.

Acknowledgments

We are indebted to Alexander Givental for key suggestions on the proof, and to Yael Karshon for fruitful discussions. We are also grateful to the anonymous referee for useful suggestions. B.K. was partially supported by an NSERC Discovery Grant.

References

[1] Artur Avila. On the regularization of conservative maps. *Acta Mathematica*, 205(1):5–18, 2010. 153, 154

[2] Augustin Banyaga. Formes-volume sur les variétés à bord. *Enseignement Math*, 20(2):127–131, 1974. 142, 154

[3] Augustin Banyaga and David E. Hurtubise. A proof of the Morse-Bott lemma. *Expositiones Mathematicae*, 22(4):365–373, 2004. 145

[4] Martins Bruveris, Peter W. Michor, Adam Parusiński, and Armin Rainer. Moser's theorem on manifolds with corners. *Proc. Amer. Math. Soc.*, 146(11):4889–4897, 2018. 142, 153, 154

[5] Henrique Bursztyn, Hudson Lima, and Eckhard Meinrenken. Splitting theorems for Poisson and related structures. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2019(754):281–312, 2019. 147

[6] Robert Cardona and Eva Miranda. On the volume elements of a manifold with transverse zeroes. *Regular and Chaotic Dynamics*, 24:187–197, 2019. 142, 144, 153, 155

[7] Gyula Csató, Bernard Dacorogna, and Olivier Kneuss. *The pullback equation for differential forms*, volume 83. Springer Science & Business Media, 2011. 142, 153

[8] Giacomo Dossena. Sylvester's law of inertia for quadratic forms on vector bundles. *preprint arXiv:1307.2171*, 2013. 149

[9] L. Fritsche and M. Haugk. Stochastic foundation of quantum mechanics and the origin of particle spin. *preprint arXiv:0912.3442*, 2009. 145

[10] Anton Izosimov, Boris Khesin, and Ilia Kirillov. Coadjoint orbits of area-preserving diffeomorphisms of non-orientable surfaces. *J. of Symplectic Geometry*, 23(1):1–35, 2025. 144

[11] Jürgen Jost. *Riemannian geometry and geometric analysis*, volume 4. Springer, 2005. 149

[12] Boris Khesin, Gerard Misiołek, and Klas Modin. Geometry of the Madelung transform. *Archive for Rational Mechanics and Analysis*, 234:549–573, 2019. 144, 145

[13] Eckhard Meinrenken. Euler-like vector fields, normal forms, and isotropic embeddings. *Indagationes Mathematicae*, 32(1):224–245, 2021. 146, 147, 148

[14] Jürgen Moser. On the volume elements on a manifold. *Transactions of the American Mathematical Society*, 120(2):286–294, 1965. 141, 142

[15] Timothy C. Wallstrom. Inequivalence between the Schrödinger equation and the Madelung hydrodynamic equations. *Phys. Rev. A*, 49:1613–1617, Mar 1994. 145

AUTHORS

Boris Khesin
Department of Mathematics,
University of Toronto;
Toronto, Canada
email: khesin@math.toronto.edu

Luke Volk
Department of Mathematics,
University of Toronto;
Toronto, Canada
email: luke.volk@mail.utoronto.ca