

# A generalization of the Turaev cobracket and the minimal self-intersection number of a curve on a surface

Patricia Cahn

ABSTRACT. Goldman and Turaev constructed a Lie bialgebra structure on the free  $\mathbb{Z}$ -module generated by free homotopy classes of loops on a surface. Turaev conjectured that his cobracket  $\Delta(\alpha)$  is zero if and only if  $\alpha$  is a power of a simple class. Chas constructed examples that show Turaev’s conjecture is, unfortunately, false. We define an operation  $\mu$  in the spirit of the Andersen–Mattes–Reshetikhin algebra of chord diagrams. The Turaev cobracket factors through  $\mu$ , so we can view  $\mu$  as a generalization of  $\Delta$ . We show that Turaev’s conjecture holds when  $\Delta$  is replaced with  $\mu$ . We also show that  $\mu(\alpha)$  gives an explicit formula for the minimum number of self-intersection points of a loop in  $\alpha$ . The operation  $\mu$  also satisfies identities similar to the co-Jacobi and coskew symmetry identities, so while  $\mu$  is not a cobracket,  $\mu$  behaves like a Lie cobracket for the Andersen–Mattes–Reshetikhin Poisson algebra.

## CONTENTS

1. Introduction	254
1.1. Related results	256
2. The Goldman–Turaev and Andersen–Mattes–Reshetikhin algebras and the operation $\mu$	257
2.1. The Goldman–Turaev Lie bialgebra	257
2.2. Andersen–Mattes–Reshetikhin algebras of chord diagrams	259
2.3. The operation $\mu$	260
2.4. $\mu(\mathcal{D})$ is independent of the choice of representative of $\mathcal{D}$	262
2.5. Alternative notation for $\mu$	262
2.6. Relationship between $\mu$ , the Goldman–Turaev Lie bialgebra, and the Andersen–Mattes–Reshetikhin algebra of chord diagrams	263
3. Proofs of theorems	264
3.1. Intersection points of powers of loops	264

Received April 14, 2013.

2010 *Mathematics Subject Classification*. 57N05, 57M99 (primary), 17B62 (secondary).

*Key words and phrases*. Self-intersections, curves on surfaces, free homotopy classes, Lie bialgebras.

3.2.	Canceling terms of $\mu$	266
3.3.	Using $\mu$ to compute the minimal self-intersection number of a class which is not primitive	273
4.	An example	275
5.	Algebraic properties of $\mu$	277
5.1.	The definitions of $\mu_i$ and $\mu$	278
5.2.	The maps $\mu$ and $\mu_i$ are well-defined	279
5.3.	Algebraic properties of $\mu$ and $\mu_i$	280
	Acknowledgements	281
	References	281

## 1. Introduction

We work in the smooth category. All manifolds and maps are assumed to be smooth unless stated otherwise, where smooth means  $C^\infty$ .

Goldman [16] and Turaev [22] constructed a Lie bialgebra structure on the free  $\mathbb{Z}$ -module generated by nontrivial free homotopy classes of loops on a surface  $F$ . Turaev [22] conjectured that his cobracket  $\Delta(\alpha)$  is zero if and only if the class  $\alpha$  is a power of a simple class, where we say a free homotopy class is simple if it contains a simple representative. Chas [9] constructed examples showing that, unfortunately, Turaev's conjecture is false. In this paper, we show that Turaev's conjecture is almost true. We define an operation  $\mu$  in the spirit of the Andersen–Mattes–Reshetikhin algebra of chord diagrams, and show that Turaev's conjecture holds on all orientable surfaces when one replaces  $\Delta$  with  $\mu$ .

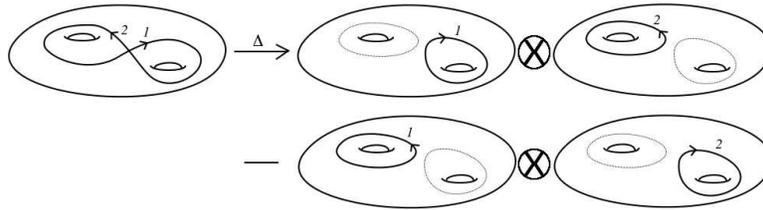


FIGURE 1. Two terms of Turaev's cobracket  $\Delta(\alpha)$  with coefficients  $+1$  and  $-1$ .

Turaev's cobracket  $\Delta(\alpha)$  is a sum over the self-intersection points  $p$  of a loop  $a$  in a free homotopy class  $\alpha$ . Each term of the sum is a tensor product of two free homotopy classes of loops. The two loops are obtained by smoothing  $a$  at the self-intersection point  $p$  according to the orientation of  $a$ . Each tensor product of loops is equipped with a sign (see Figure 1).

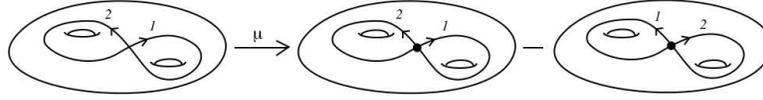


FIGURE 2. Two terms of the operation  $\mu(\alpha)$  with coefficients  $+1$  and  $-1$ .

Turaev’s conjecture is false because it is not uncommon for the same simple tensor of loops to appear twice in the sum  $\Delta(\alpha)$ , but with different signs.

We define the operation  $\mu(\alpha)$  as a sum over the self-intersection points  $p$  of a loop  $a$  in  $\alpha$ , as in the definition of the Turaev cobracket. Rather than smoothing at each self-intersection point to obtain a tensor product of two loops, we glue those loops together to create a wedge of two circles mapped to the surface. This can also be viewed as a chord diagram with one chord.

More precisely, for the reader already familiar with chord diagrams, we define

$$\mu([D]) = \sum_{(t_1, t_2) \in \mathcal{SI}_0} [D_p^+] - [D_p^-],$$

where  $D : S^1 \rightarrow F$  is the given loop,  $[D]$  is its free homotopy class,  $\mathcal{SI}_0$  is the set of self-intersections of  $D$  (excluding those where one of the two loops obtained by smoothing at that self-intersection is trivial),  $[D_p^+]$  is the labeled chord diagram obtained by adding one chord between the preimages  $t_1$  and  $t_2$  of the self-intersection point  $p$  and ordering the two circles in the resulting wedge of circles with a positive labelling, and  $[D_p^-]$  is the same chord diagram but with the labelling reversed. For more details and for the definition of a chord diagram we direct the reader to Section 2.

As a result of replacing the smoothing operation with the gluing operation, terms of  $\mu$  are less likely to cancel than terms of  $\Delta$ , and hence  $\mu(\alpha)$  is less likely to be zero. In fact, Turaev’s conjecture holds when formulated for  $\mu$  rather than  $\Delta$ :

**1.1. Theorem.** *Let  $F$  be an oriented surface with or without boundary, which may or may not be compact. Let  $\alpha$  be a free homotopy class on  $F$ . Then  $\mu(\alpha) = 0$  if and only if  $\alpha$  is a power of a simple class.*

There is a simple relationship between  $\Delta$  and  $\mu$ ; namely, if one smoothes each term of  $\mu$  at the gluing point, and tensors the resulting loops, one obtains a term of  $\Delta$  (see Figure 3). Hence the Turaev cobracket factors

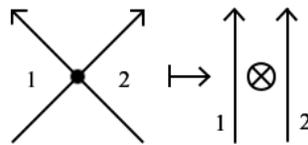


FIGURE 3. The Turaev cobracket factors through  $\mu$ .

through  $\mu$ , and we can view  $\mu$  as a generalization of  $\Delta$ . The relationship between  $\mu$  and  $\Delta$  is analogous to the relationship between the Andersen–Mattes–Reshetikhin Poisson bracket for chord diagrams and the Goldman Lie bracket. While  $\mu$  is not a cobracket, in the final section of the paper, we show that  $\mu$  satisfies identities similar to coskew symmetry and the co-Jacobi identity.

The operation  $\mu$  also gives a formula for the minimum number of self-intersection points of a generic loop in a given free homotopy class  $\alpha$ . By a generic loop, we mean a loop whose self-intersection points are transverse double points. We call this number the *minimal self-intersection number of  $\alpha$*  and denote it by  $m(\alpha)$ . Both Turaev’s cobracket and the operation  $\mu$  give lower bounds on the minimal self-intersection number of a given homotopy class  $\alpha$ . We call a free homotopy class *primitive* if it is not a power of another class in  $\pi_1(F)$ . Any class  $\alpha$  can be written as  $\beta^n$  for some primitive class  $\beta$  and  $n \geq 1$ . It follows easily from the definitions of  $\Delta$  and  $\mu$  that  $m(\alpha)$  is greater than or equal to  $n - 1$  plus half the number of terms in the (reduced) linear combinations  $\Delta(\alpha)$  or  $\mu(\alpha)$ .

**1.2. Definition.** The *number of terms*  $t(L)$  of a reduced linear combination  $L$  of simple tensors of classes of loops, or of classes of chord diagrams, is the sum of the absolute values of the coefficients of the classes.

Chas’ counterexamples to Turaev’s conjecture show that the lower bound given by  $\Delta(\alpha)$  cannot, in general, be used to compute the minimal self-intersection number of  $\alpha$ . In order to compute the minimal self-intersection number of  $\alpha$  using  $\Delta$  on surfaces with boundary, Chas and Krongold showed that one should instead count the number of terms of  $\Delta(\alpha^k)$  for  $k \geq 3$  [12]. However the lower bound given by  $\mu(\alpha)$  is always equal to  $m(\alpha)$ :

**1.3. Theorem.** *Let  $F$  be an oriented surface with or without boundary, which may or may not be compact. Let  $\alpha$  be a nontrivial free homotopy class on  $F$  such that  $\alpha = \beta^n$ , where  $\beta$  is primitive and  $n \geq 1$ . Then the minimal self-intersection number of  $\alpha$  is equal to  $n - 1$  plus half the number of terms of  $\mu(\alpha)$ .*

In order to prove the case of Theorem 1.3 where  $n > 1$ , we make use of the results of Hass and Scott [17] who describe geometric properties of curves with minimal self-intersection (see also [15]).

**1.1. Related results.** We briefly summarize some results related to Turaev’s conjecture and computations of the minimal self-intersection number. Le Donne [18] proved that Turaev’s conjecture is true for genus zero surfaces. For surfaces of positive genus, one might wonder to what extent Turaev’s conjecture is false. Chas and Krongold [12] approach this question by showing that, on surfaces with boundary, if  $\Delta(\alpha) = 0$  and  $\alpha$  is at least a third power of a primitive class  $\beta$ , then  $\beta$  is simple.

A nice history of the problem of determining when a homotopy class is represented by a simple loop is given in Rivin [20]. Birman and Series [6] give an explicit algorithm for detecting simple classes on surfaces with boundary. Cohen and Lustig [14] extend the work of Birman and Series to obtain an algorithm for computing the minimal intersection and self-intersection numbers of curves on surfaces with boundary, and Lustig [19] extends this to closed surfaces. We give an example which shows how one can algorithmically compute  $m(\alpha)$  using  $\mu$  on surfaces with boundary, though generally we do not emphasize algorithmic implications in this paper.

A different algebraic solution to the problem of computing the minimal intersection and self-intersection numbers of curves on a surface is given by Turaev and Viro [24]. However  $\mu$  has a simple relationship to  $\Delta$  and pairs well with the Andersen–Mattes–Reshetikhin Poisson bracket. Chernov [13] used the Andersen–Mattes–Reshetikhin bracket to compute the minimum number of intersection points of loops in given free homotopy classes which are not powers of the same class.

After this work was completed, the techniques developed in this paper allowed the author, together with Chernov, to prove that the Andersen–Mattes–Reshetikhin bracket computes the minimal intersection number of any two distinct free homotopy classes even when the given loops are powers of the same loop [8]. As a result, the author and Chernov obtained a formula for the minimal self-intersection number of a free homotopy class in terms of the Andersen–Mattes–Reshetikhin Poisson bracket. In particular  $m(\alpha)$  is equal to  $\frac{1}{2|pq|}t(\{\alpha^p, \alpha^q\}) + n - 1$ , where  $n$  is the largest positive integer such that  $\alpha = \beta^n$  in  $\pi_1(F)$ , and  $p$  and  $q$  are any two distinct integers. Around the same time, Chas and Gadgil showed that by counting terms of the Goldman bracket  $[\alpha^p, \beta^q]$  for  $p$  and  $q$  large enough, one can compute intersection numbers on surfaces and orbifolds [11].

## 2. The Goldman–Turaev and Andersen–Mattes–Reshetikhin algebras and the operation $\mu$

Before defining  $\mu$ , we review the definitions of the Goldman–Turaev Lie bialgebra and the Andersen–Mattes–Reshetikhin Poisson algebra. Our goal in this section is to emphasize that the relationship between  $\mu$  and the Andersen–Mattes–Reshetikhin Poisson bracket is analogous to the relationship between the Turaev cobracket and the Goldman bracket.

**2.1. The Goldman–Turaev Lie bialgebra.** We now define the Goldman–Turaev Lie Bialgebra on the free  $\mathbb{Z}$ -module generated by the set  $\hat{\pi}$  of free homotopy classes of loops on  $F$ , which we denote by  $\mathbb{Z}[\hat{\pi}]$ . Let  $\alpha, \beta \in \hat{\pi}$ , and let  $a$  and  $b$  be smooth, transverse representatives of  $\alpha$  and  $\beta$ , respectively, and assume that all intersection points of  $a$  and  $b$  are double points. We will use square brackets to denote the free homotopy class of a loop. The set of intersection points  $\mathcal{I}_{a,b}$ , or just  $\mathcal{I}$  when the choice of  $a$  and

$b$  is clear, is defined to be

$$\mathcal{I} = \{(t_1, t_2) \in S^1 \times S^1 : a(t_1) = b(t_2)\}.$$

Let  $a \cdot_p b$  denote the product of  $a$  and  $b$  as based loops in  $\pi_1(F, p)$ , where  $p = a(t_1) = b(t_2)$  and  $(t_1, t_2) \in \mathcal{I}$ .

The Goldman bracket [16] is a linear map  $[\cdot, \cdot] : \mathbb{Z}[\hat{\pi}] \otimes_{\mathbb{Z}} \mathbb{Z}[\hat{\pi}] \rightarrow \mathbb{Z}[\hat{\pi}]$ , defined by

$$[\alpha, \beta] = \sum_{(t_1, t_2) \in \mathcal{I}} \text{sgn}(p; a, b) [a \cdot_p b],$$

where  $\text{sgn}(p; a, b) = 1$  if the orientation given by the pair  $\{a'(t_1), b'(t_2)\}$  of vectors agrees with the orientation of  $F$ , and  $\text{sgn}(p; a, b) = -1$  otherwise.

Next we define the Turaev cobracket [22]. Let  $\alpha$  be a free homotopy class on  $F$ , and let  $a$  be a smooth representative of  $\alpha$  with transverse self-intersection points. Let  $\mathcal{SI}_a$ , or just  $\mathcal{SI}$  when the choice of  $a$  is clear, denote the set of self-intersection points of the loop  $a$ . Let  $\mathbb{D}$  be the diagonal in  $S^1 \times S^1$ . Elements of  $\mathcal{SI}$  will be points in  $S^1 \times S^1 - \mathbb{D}$  modulo the action of  $\mathbb{Z}_2$  which interchanges the two coordinates. Now we define

$$\mathcal{SI} = \{(t_1, t_2) \in (S^1 \times S^1 - \mathbb{D})/\mathbb{Z}_2 : a(t_1) = a(t_2)\}.$$

Let  $p = a(t_1) = a(t_2)$  be a self-intersection point of  $a$ . Let  $[t_1, t_2]$  denote the arc of  $S^1$  going from  $t_1$  to  $t_2$  in the direction of the orientation of  $S^1$ , and let  $[t_2, t_1]$  denote the arc of  $S^1$  going from  $t_2$  to  $t_1$  in the direction of the orientation of  $S^1$ . Since  $p = a(t_1) = a(t_2)$ , then  $a([t_1, t_2])$  and  $a([t_2, t_1])$  are loops. We assign these loops the names  $a_p^1$  and  $a_p^2$  in such a way that the ordered pair of initial velocity vectors  $\{(a_p^1)'(t_i), (a_p^2)'(t_j)\}$  (where  $\{i, j\} = \{1, 2\}$  are chosen so that these vectors are indeed initial rather than final velocity vectors) gives the chosen orientation of  $T_p F$ . Now we let  $\mathcal{SI}_0$  be the subset of  $\mathcal{SI}$  which contains only self-intersection points  $p$  such that the loops  $a_p^i$  are nontrivial:

$$\mathcal{SI}_0 = \{(t_1, t_2) \in \mathcal{SI} : p = a(t_1) = a(t_2), a_p^1, a_p^2 \neq 1 \in \pi_1(F_p)\}.$$

The Turaev cobracket is a linear map  $\Delta : \mathbb{Z}[\hat{\pi}] \rightarrow \mathbb{Z}[\hat{\pi}] \otimes_{\mathbb{Z}} \mathbb{Z}[\hat{\pi}]$  which is given on a single homotopy class by

$$\Delta(\alpha) = \sum_{(t_1, t_2) \in \mathcal{SI}_0} [a_p^1] \otimes [a_p^2] - [a_p^2] \otimes [a_p^1].$$

One can show that the definition of  $\Delta$  is independent of the choice of  $a \in \alpha$  by showing  $\Delta(\alpha)$  does not change under elementary moves for a smooth loop in general position (see Figure 7). Using linearity, this definition of  $\Delta$  can be extended to all of  $\mathbb{Z}[\hat{\pi}]$ .

Together,  $[\cdot, \cdot]$  and  $\Delta$  equip  $\mathbb{Z}[\hat{\pi}]$  with an involutive Lie Bialgebra structure [16, 22]. That is,  $[\cdot, \cdot]$  and  $\Delta$  satisfy (co)skew-symmetry, the (co) Jacobi identity, a compatibility condition, and  $[\cdot, \cdot] \circ \Delta = 0$ . A complete definition of a Lie Bialgebra is given in [9].

**2.2. Andersen–Mattes–Reshetikhin algebras of chord diagrams.**

We now summarize the Andersen–Mattes–Reshetikhin algebra of chord diagrams on  $F$  [1, 2]. A *chord diagram* is a disjoint union of oriented circles  $S_1, \dots, S_k$ , called *core circles*, along with a collection of disjoint arcs  $C_1, \dots, C_l$ , called *chords*, such that:

- (1)  $\partial C_i \cap \partial C_j = \emptyset$  for  $i \neq j$ , and
- (2)  $\bigcup_{i=1}^l \partial C_i = \left(\bigcup_{i=1}^k S_i\right) \cap \left(\bigcup_{i=1}^l C_i\right)$ .

A *geometrical chord diagram on  $F$*  is a smooth map from a chord diagram  $D$  to  $F$  such that each chord  $C_i$  in  $D$  is mapped to a point. A *chord diagram on  $F$*  is a homotopy class of a geometrical chord diagram  $D$ , denoted  $[D]$ .

Let  $M$  denote the free  $\mathbb{Z}$ -module generated by the set of chord diagrams on  $F$  ([2] uses coefficients in  $\mathbb{C}$ , but we use  $\mathbb{Z}$  here for consistency). Let  $N$  be the submodule generated by a set of  $4T$ -relations, one of which is shown in Figure 4. The other relations can be obtained from this one as follows: one can reverse the direction of any arc, and any time a chord intersects an arc whose orientation is reversed, the diagram is multiplied by a factor of  $-1$ .

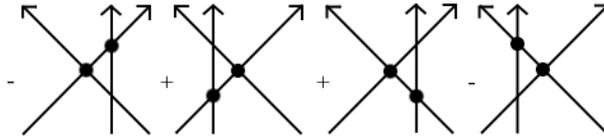


FIGURE 4.  $4T$ -relations

Given two chord diagrams  $\mathcal{D}_1$  and  $\mathcal{D}_2$  on  $F$ , we can form their disjoint union by choosing representatives (i.e., geometrical chord diagrams)  $D_i$  of  $\mathcal{D}_i$ , taking a disjoint union of their underlying chord diagrams, mapping the result to  $F$  as prescribed by the  $D_i$ , and taking its free homotopy class. The disjoint union of chord diagrams  $\mathcal{D}_1 \cup \mathcal{D}_2$  defines a commutative multiplication on  $M$ , giving  $M$  an algebra structure with  $N$  as an ideal. Let  $\mathfrak{ch} = M/N$ , and call this the *algebra of chord diagrams*.

Andersen, Mattes, and Reshetikhin [1, 2] constructed a Poisson bracket on  $\mathfrak{ch}$ , which can be viewed as a generalization of the Goldman bracket for chord diagrams on  $F$  rather than free homotopy classes of loops. Let  $\mathcal{D}_1$  and  $\mathcal{D}_2$  be chord diagrams on  $F$ , and choose representatives  $D_i$  of  $\mathcal{D}_i$ . We define the set of intersection points  $\mathcal{I}_{D_1, D_2}$ , or just  $\mathcal{I}$  when the choice of  $D_1$  and  $D_2$  is clear, to be

$$\mathcal{I} = \{(t_1, t_2) : D_1(t_1) = D_2(t_2)\},$$

where  $t_i$  is a point in the preimage of the geometrical chord diagrams  $D_i$ . For each  $(t_1, t_2) \in \mathcal{I}$  with  $p = D_i(t_i)$ , let  $D_1 \cup_p D_2$  denote the geometrical chord diagram obtained by adding a chord between  $t_1$  and  $t_2$ . Later we will

want to let  $p$  have multiplicity greater than 2. In that case, it is necessary to specify preimages of  $p$  for this notation to be well-defined. Since each copy of  $S^1$  in the chord diagram is oriented, we can define  $\text{sgn}(p; D_1, D_2)$  as before. The Andersen–Mattes–Reshetikhin Poisson bracket  $\{\cdot, \cdot\} : \mathfrak{ch} \times \mathfrak{ch} \rightarrow \mathfrak{ch}$  is defined by

$$\{\mathcal{D}_1, \mathcal{D}_2\} = \sum_{(t_1, t_2) \in \mathcal{I}} \text{sgn}(p; D_1, D_2) [D_1 \cup_p D_2],$$

where square brackets denote the free homotopy class of a geometrical chord diagram. This definition of  $\{\cdot, \cdot\}$  can be extended to all of  $\mathfrak{ch}$  using bilinearity. For a proof that  $\{\cdot, \cdot\}$  does not depend on the choices of  $D_i \in \mathcal{D}_i$ ,  $i = 1, 2$ , see [2]. In particular, it is necessary to check that  $\{\cdot, \cdot\}$  is invariant under elementary moves, including the Reidemeister moves in Figure 7, the moves in Figures 5 and 6, and the  $4T$ -relations.

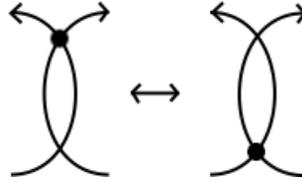


FIGURE 5. An elementary move for chord diagrams (with one of several possible choices of orientations on the arcs).

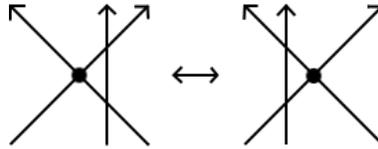


FIGURE 6. An elementary move for chord diagrams (with one of several possible choices of orientations on the arcs).

**2.3. The operation  $\mu$ .** The definition of  $\mu$  given in this section is the simplest for the purposes of computing the minimal self-intersection number of a free homotopy class  $\alpha$ . In this section, we define  $\mu$  only on free homotopy classes. In the final section of this paper, we modify the definition of  $\mu$  in a way that allows us to more easily state an analogue of the co-Jacobi identity, and which allows us to extend the definition of  $\mu$  to certain chord diagrams in the Andersen–Mattes–Reshetikhin algebra.

For this definition of  $\mu$ , we will consider geometrical chord diagrams with one core circle and one chord equipped with a sign  $\epsilon \in \{+, -\}$ . Soon we will see that a generic geometrical chord diagram  $D$  with one core circle and

one signed chord can be viewed as a map of a wedge of two circles to our surface, where the two circles are ordered.

Let  $E$  denote the free  $\mathbb{Z}$ -module generated by chord diagrams on  $F$  with one core circle and one signed chord. Recall that a chord diagram is defined to be the homotopy class of a geometrical chord diagram. Two geometrical chord diagrams with one signed chord represent the same chord diagram on  $F$  if and only if they are related by the usual Reidemeister moves for curves on surfaces (see Figure 7), plus two additional elementary moves for diagrams with signed chords. These moves, with one possible choice of orientation on the branches, are shown in Figures 8 and 9, where  $\epsilon \in \{+, -\}$  denotes the sign on the chord. Note that we do not yet need to take a quotient by 4T-relations on  $E$  because our diagrams have only one chord. We define a linear map  $\mu : \mathbb{Z}[\hat{\pi}] \rightarrow E$ .

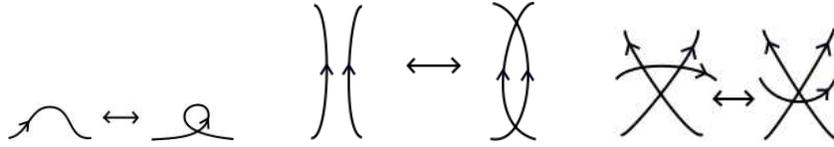


FIGURE 7. Reidemeister moves for curves on surfaces with one possible choice of orientation on each branch.

Let  $D : S^1 \rightarrow F$  be a geometrical chord diagram on  $F$  with one core circle. For each self-intersection point  $p = D(t_1) = D(t_2)$  of  $D$ , we let  $D_p^+$  (respectively  $D_p^-$ ) be the geometrical chord diagram obtained by adding a chord with sign  $\epsilon = +$  (respectively  $\epsilon = -$ ) between  $t_1$  and  $t_2$ .

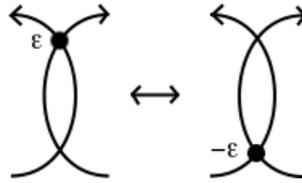


FIGURE 8. Elementary move for chord diagrams with signed chords.

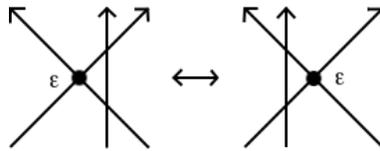


FIGURE 9. Elementary move for chord diagrams with signed chords.

Now we define  $\mu$  on the class of the geometrical chord diagram  $D$  by

$$\mu([D]) = \sum_{(t_1, t_2) \in \mathcal{SI}_0} [D_p^+] - [D_p^-].$$

Using linearity, we can extend this definition to all of  $\mathbb{Z}[\hat{\pi}]$ . It remains to check that  $\mu([D])$  is independent of the choice of representative of  $[D]$ .

**2.4.  $\mu(\mathcal{D})$  is independent of the choice of representative of  $\mathcal{D}$ .** We check that  $\mu$  is invariant under the usual Reidemeister moves:

- (1) *Regular isotopy*: Invariance is clear.
- (2) *First Reidemeister Move*: This follows from the definition of  $\mathcal{SI}_0$ .
- (3) *Second Reidemeister Move*: This follows from the move in Figure 8.
- (4) *Third Reidemeister Move*: This follows from the move in Figure 9.

We note that when checking invariance under the second and third moves, one must consider the case where some of the self-intersection points are in  $\mathcal{SI}$  but not in  $\mathcal{SI}_0$ .

Each signed chord diagram  $D_p^\epsilon$  with one chord of sign  $\epsilon$  and one core circle corresponds to a map from a wedge of circles to  $F$  where the two circles are ordered. Suppose the chord connects two preimages  $t_1$  and  $t_2$  of  $p$ . As before this gives rise to two loops  $a_p^1$  and  $a_p^2$ . If  $\epsilon = +$  we label the loop  $a_p^1$  with a 1 and label  $a_p^2$  with a 2. If  $\epsilon = -$  we label the loop  $a_p^1$  with a 2 and label  $a_p^2$  with a 1. As one can see from Figures 1 and 2, this makes the relationship between  $\Delta$  and  $\mu$  transparent.

**2.5. Alternative notation for  $\mu$ .** Next we will rewrite the definition of  $\mu$  in a way that makes its relationship to  $\Delta$  more transparent. Let  $\phi$  and  $\psi : I = [0, 1] \rightarrow F$  be loops in  $F$  based at  $p$ , such that  $\phi'(0) = \psi'(1)$  and  $\phi'(1) = \psi'(0)$ . We define a geometrical chord diagram  $\phi \bullet_p \psi$  which, intuitively, glues the loops  $\phi$  and  $\psi$  at the point  $p$ . The underlying chord diagram of  $\phi \bullet_p \psi$  contains one core circle  $S^1 = I/\partial I$ , and one chord  $C$  with endpoints at  $0 \in I$  and  $\frac{1}{2} \in I$ . The geometrical chord diagram  $\phi \bullet_p \psi$  maps the chord  $C$  to  $p$ . Then we define  $(\phi \bullet_p \psi)|_{[0, \frac{1}{2}]} = \phi$  and  $(\phi \bullet_p \psi)|_{[\frac{1}{2}, 1]} = \psi$ , and label  $\phi$  with a 1 and  $\psi$  with a 2. By the discussion above, this is equivalent to equipping the chord with a positive (respectively, negative) sign if  $\{\phi'(0), \psi'(0)\}$  form a positive (respectively, negative) frame.

Now we are ready to rewrite the definition of  $\mu$  for  $\alpha \in \hat{\pi}(F)$ . Let  $a$  be a representative of  $\alpha$ , and for each  $(t_1, t_2) \in \mathcal{SI}_0$  with  $p = a(t_1) = a(t_2)$ , let  $a_p^1$  and  $a_p^2$  be the loops we defined for the Turaev cobracket. Now

$$\mu(\alpha) = \sum_{(t_1, t_2) \in \mathcal{SI}_0} [a_p^1 \bullet_p a_p^2] - [a_p^2 \bullet_p a_p^1].$$

We will make frequent use of the following proposition, whose proof is straightforward:

**2.1. Proposition.** *Two generic geometrical chord diagrams  $D_1 = \phi_1 \bullet_p \psi_1$  and  $D_2 = \phi_2 \bullet_p \psi_2$  represent the same chord diagram on  $F$ , i.e., are related by flat Reidemeister moves and the moves in Figures 8 and 9, if and only if there exists  $\gamma \in \pi_1(F, p)$  such that  $\gamma\phi_1\gamma^{-1} = \phi_2$  and  $\gamma\psi_1\gamma^{-1} = \psi_2$ .*

**2.6. Relationship between  $\mu$ , the Goldman–Turaev Lie bialgebra, and the Andersen–Mattes–Reshetikhin algebra of chord diagrams.**

Andersen, Mattes and Reshetikhin [2] show that there is a quotient algebra of  $\mathfrak{ch}$  which corresponds to Goldman’s algebra. Let  $I$  be the ideal generated by the relation in Figure 10. In the quotient  $\mathfrak{ch}/I$ , each chord diagram is identified with the disjoint union of free homotopy classes obtained by smoothing the diagram at the intersections which are images of chords. One can check that  $P : \mathfrak{ch} \rightarrow \mathfrak{ch}/I$  is a Poisson algebra homomorphism and

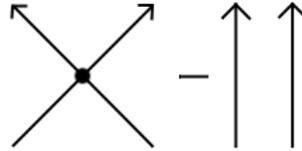


FIGURE 10. Generator of  $I$

$\mathfrak{ch}/I$  is a Poisson algebra with an underlying Lie algebra that corresponds to Goldman’s algebra [2].

There is a similar relationship between the Turaev cobracket and  $\mu$ . Let  $Q$  be the map which smoothes the chord diagram according to its orientation at an intersection which is an image of a chord, and tensors the two resulting homotopy classes together (see Figure 11). Then  $\Delta = Q \circ \mu$ .

**Remark.** Turaev [22, p. 660] notes that the Turaev cobracket can be obtained algebraically from an operation defined in Supplement 2 of [23]. It is possible that  $\mu$  may be obtained from this operation as well. We do not know a way of obtaining Turaev’s operation from  $\mu$ .

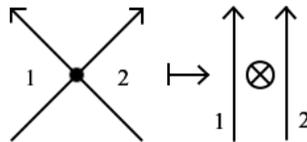


FIGURE 11. The map  $Q$ .

### 3. Proofs of theorems

In this section, we prove Theorems 1.1 and 1.3. Recall that Theorem 1.1 states that  $\mu(\alpha) = 0$  if and only if  $\alpha$  is a power of a simple class. Theorem 1.3 gives an explicit formula for  $m(\alpha)$ . We begin by describing two types of self-intersection points of a loop which is freely homotopic to a power of another loop. Then we prove Theorem 3.2, which describes when certain terms of  $\mu$  cancel. Theorems 1.1 and 1.3 are corollaries of Theorem 3.2.

**3.1. Intersection points of powers of loops.** Our goal is to understand the conditions under which different terms of  $\mu(\alpha)$  cancel, when  $\alpha \in \hat{\pi}$  is a power of another class  $\beta$  in  $\pi_1(F)$ . To do this, we need to distinguish between two different types of self-intersection points of a curve. Suppose we choose a geodesic representative  $g$  of  $\alpha$ . Either all self-intersection points of  $g$  are transverse, or  $g$  has infinitely many self-intersection points, and in particular,  $g$  is a power of another geodesic. Let  $p$  be a point on the image of  $g$  which is not a transverse self-intersection point of  $g$ . Let  $h$  be a geodesic loop such that  $g = h^n$  in  $\pi_1(F, p)$ , and such that there is no geodesic  $f$  such that  $h = f^k$  (it is possible that  $|n| = 1$ ). Now we know that  $h$  has finitely many self-intersection points, all of which are transverse. Let  $m$  be the number of self-intersection points of  $h$ . Since  $F$  is orientable, we can perturb  $g$  slightly to obtain a loop  $g'$  as follows: We begin to traverse  $g$  beginning at  $p$ , but whenever we are about to return to  $p$ , we shift slightly to the left. After doing this  $n$  times, we must return to  $p$  and connect to the starting point. This requires crossing  $n - 1$  strands of the loop, creating  $n - 1$  self-intersection points. We call these *Type 2 self-intersection points*. For each self-intersection point of  $h$ , we get  $n^2$  self-intersection points of  $g$  (see Figure 12). We call these  $mn^2$  self-intersection points *Type 1 self-intersection points*. We note that we are counting self-intersections with multiplicity, as some of the self-intersection points of  $h$  may be images of multiple points in  $\mathcal{S}\mathcal{I}$ . Given a transverse self-intersection point  $p$  of  $h$ , we

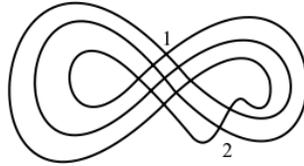


FIGURE 12. Type 1 and Type 2 self-intersection points.

will denote the corresponding set of  $n^2$  Type 1 self-intersection points of  $g'$  by  $\{p_{i,j}\}$ , where  $i \in \{1, \dots, n\}$  is the label on the strand corresponding first branch of  $h$  at  $p$  (i.e., a strand going from top to bottom in Figure 13), and  $j \in \{1, \dots, n\}$  is the label on the strand corresponding to the second branch of  $h$  at  $p$  (i.e., a strand going from left to right in Figure 13). This

relationship between the numbers of self-intersection points of  $g$  and  $h$  can be found in [24] for both orientable and nonorientable surfaces.

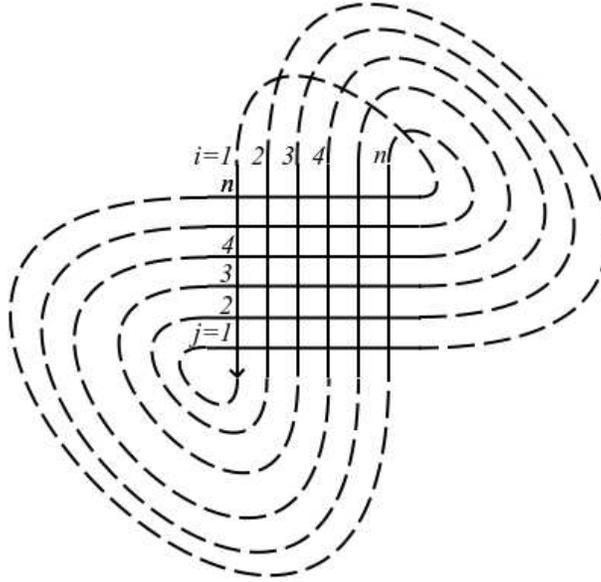


FIGURE 13. Type 1 intersection points.

**3.1. Lemma.** *Let  $g$  be a geodesic representative of  $\alpha \in \hat{\pi}(F)$ , with  $g = h^n$ , and  $h$  and  $n$  are as defined in the paragraph above. Then the contribution to  $\mu(\alpha)$  of a Type 1 self-intersection point  $p_{i,j}$  is*

$$[(XY)^I X \bullet_p (YX)^J Y] - [(YX)^J Y \bullet_p (XY)^I X],$$

where  $X = h_p^1$ ,  $Y = h_p^2$ , and  $I, J \in \mathbb{N}$  such that  $I + J = n - 1$ .

**Proof.** We will compute the contribution to  $\mu$  for a Type 1 self-intersection point  $p_{i,j}$  of  $g'$ , where  $g'$  is the perturbed version of  $g$  described in the above paragraph. These terms are  $[(g')^1_{p_{i,j}} \bullet_{p_{i,j}} (g')^2_{p_{i,j}}]$  and  $-[(g')^2_{p_{i,j}} \bullet_{p_{i,j}} (g')^1_{p_{i,j}}]$ . However, when we record the terms of  $\mu$ , we perturb  $g'$  back to  $g$ , so that the terms we record are geometrical chord diagrams whose images are contained in the image of  $g$  and whose chords are mapped to  $p$ . To compute  $(g')^1_{p_{i,j}}$ , we begin at  $p_{i,j}$  along the branch corresponding to  $X = h_p^1$ , and wish to know how many times we traverse branches corresponding to  $X = h_p^1$  and  $Y = h_p^2$  before returning to  $p_{i,j}$ . The first time we return to  $p_{i,j}$ , we must return along the  $j^{\text{th}}$  branch of  $X$ . Therefore  $[(g')^1_{p_{i,j}}] = [(XY)^I X]$  for some integer  $I \geq 0$ . If we begin at  $p_{i,j}$  along the branch corresponding to  $Y = h_p^2$ , we return to  $p_{i,j}$  for the first time on the  $i^{\text{th}}$  branch of  $Y$ . Therefore  $[(g')^2_{p_{i,j}}] = [(YX)^J Y]$  for some integer  $J \geq 0$ . But if we traverse  $(g')^1_{p_{i,j}}$  followed by  $(g')^2_{p_{i,j}}$ , we must traverse  $g'$  exactly once, so  $I + J = n - 1$ .  $\square$

**3.2. Canceling terms of  $\mu$ .** Throughout this section, we will use the following facts, which hold for a compact surface  $F$  with negative scalar curvature (though compactness is not needed for (3)).

- (1) *Nontrivial abelian subgroups of  $\pi_1(F)$  are infinite cyclic.*
- (2) *There is a unique, maximal infinite cyclic group containing each nontrivial  $\alpha \in \pi_1(F)$ .*
- (3) *Two distinct geodesic arcs with common endpoints cannot be homotopic.*
- (4) *Each nontrivial  $\alpha \in \hat{\pi}(F)$  contains a geodesic representative which is unique up to choice of parametrization.*

The first fact holds by Preissman's Theorem. The second fact is true if  $\partial F \neq \emptyset$  because  $\pi_1(F)$  is free. If  $F$  is closed, the second fact follows from the proof of Preissman's Theorem [13]. The third and fourth facts can be found in [7], as Theorems 1.5.3 and 1.6.6 respectively.

We now show that for any free homotopy class  $\alpha$  on a compact surface, it is possible to choose a representative of  $\alpha$  such that no two terms coming from Type 1 intersection points cancel. This proof is based on ideas in [24] and [13]. Later we will see that if  $F = S^2, T^2$ , or the annulus  $A$ , geodesic loop on  $F$  has no Type 1 self-intersection points, so in Theorem 3.2, we only consider surfaces of negative curvature.

**3.2. Theorem.** *Let  $F$  be a compact surface equipped with a metric of negative curvature. Let  $\alpha \in \hat{\pi}(F)$ . If  $g$  is a geodesic representative of  $\alpha$ , then no two terms of  $\mu(\alpha)$  corresponding to Type 1 intersection points of  $g$  cancel.*

**Proof.** Throughout this proof,  $[\cdot]$  denotes a free homotopy class (either of a geometrical chord diagram or a loop),  $[\cdot]_p$  denotes a homotopy class in  $\pi_1(F, p)$ , and  $[\cdot]_q^p$  denotes the homotopy class of a path from  $p$  to  $q$  with fixed endpoints. When we concatenate two paths  $p_1$  and  $p_2$ , we write  $p_1p_2$ , where the path written on the left is the path we traverse first.

We write  $g = h^n$  for some geodesic loop  $h$  and some  $n \geq 1$ , where  $h$  is not a power of another loop. Suppose  $h$  has  $m$  self-intersection points, and let  $g'$  be a perturbation of  $g$  with  $mn^2$  Type 1 self-intersection points and  $n - 1$  Type 2 self-intersection points. Let  $\{p_{i,j} : 1 \leq i, j \leq n\}$  and  $\{q_{k,l} : 1 \leq k, l \leq n\}$  be the sets of  $n^2$  self-intersection points corresponding to the (transverse) self-intersection points  $p$  and  $q$  of  $h$  respectively, with the indexing as defined in the previous section. We assume  $[h]$  is nontrivial, since the theorem clearly holds when  $[h]$  is trivial ( $\mathcal{SI}_0$  is in fact empty).

We wish to show that the terms of  $\mu$  corresponding to points  $p_{i,j}$  and  $q_{k,l}$  cannot cancel. We suppose these terms cancel, and derive a contradiction.

First, we consider the case where  $p = q = h(t_1) = h(t_2)$  for  $(t_1, t_2) \in \mathcal{SI}_0$ , but  $i$  and  $k$  may or may not be equal, and  $j$  and  $l$  may or may not be equal. In other words,  $p_{i,j}$  and  $q_{k,l} = p_{k,l}$  come from the same set of  $n^2$  Type 1 self-intersection points. Let  $X = h_p^1$  and let  $Y = h_p^2$ . If either  $i \neq k$  or  $j \neq l$ ,

then by Lemma 3.1, the terms corresponding to  $p_{i,j}$  and  $p_{k,l}$  are

$$[(XY)^I X \bullet_p (YX)^J Y] - [(YX)^J Y \bullet_p (XY)^I X], \text{ and}$$

$$[(XY)^K X \bullet_p (YX)^L Y] - [(YX)^L Y \bullet_p (XY)^K X],$$

for integers  $I, J, K, L$  such that  $I + J = K + L = n - 1$ . If  $i = k$  and  $j = l$ , then  $p_{i,j} = p_{k,l}$  corresponds to a single element of  $\mathcal{ST}_0$ , so we have just the first two of the above terms. In either case, it suffices to assume that the terms  $[(XY)^I X \bullet_p (YX)^J Y]$  and  $[(YX)^L Y \bullet_p (XY)^K X]$  cancel, where  $I$  and  $K$  may or may not be equal, and  $J$  and  $L$  may or may not be equal.

Suppose that

$$[(XY)^I X \bullet_p (YX)^J Y] = [(YX)^L Y \bullet_p (XY)^K X].$$

Then there exists  $\gamma \in \pi_1(F, p)$  such that

$$(3.1) \quad \gamma[(XY)^I X]_p \gamma^{-1} = [(YX)^L Y]_p$$

and

$$(3.2) \quad \gamma[(YX)^J Y]_p \gamma^{-1} = [(XY)^K X]_p.$$

We multiply Equations (3.1) and (3.2) in both possible orders to obtain the equations

$$(3.3) \quad \gamma[(XY)^I X]_p [(YX)^J Y]_p \gamma^{-1} = [(YX)^L Y]_p [(XY)^K X]_p$$

and

$$(3.4) \quad \gamma[(YX)^J Y]_p [(XY)^I X]_p \gamma^{-1} = [(XY)^K X]_p [(YX)^L Y]_p.$$

Conjugating Equation (3.3) by  $[X]_p \in \pi_1(F, p)$  tells us that  $[X]_p \gamma$  and  $[(XY)^n]_p$  commute, since  $I + J + 1 = K + L + 1 = n$ . Similarly, conjugating Equation (3.4) by  $[Y]_p$  tells us  $[Y]_p \gamma$  and  $[(YX)^n]_p$  commute. Therefore the subgroups  $\langle [X]_p \gamma, [(XY)^n]_p \rangle$  and  $\langle [Y]_p \gamma, [(YX)^n]_p \rangle$  are infinite cyclic, and are generated by elements  $s$  and  $t$  of  $\pi_1(F, p)$ , respectively. Note that these subgroups are nontrivial since  $h$  is nontrivial. Fact (2) states that each nontrivial element of  $\pi_1$  is contained in a unique, maximal infinite cyclic group. Let  $m_1$  and  $m_2$  be the generators of the unique maximal infinite cyclic groups containing  $[(XY)^n]_p$  and  $[(YX)^n]_p$  respectively. Since  $[h] = [XY] = [YX]$  is not freely homotopic to a power of another class, we have that  $\langle m_1 \rangle = \langle [XY]_p \rangle$  and  $\langle m_2 \rangle = \langle [YX]_p \rangle$ . But  $\langle s \rangle$  and  $\langle t \rangle$  are also infinite cyclic groups containing  $[(XY)^n]_p$  and  $[(YX)^n]_p$ , respectively. By the maximality of the  $\langle m_i \rangle$ , we have that  $\langle s \rangle \leq \langle m_1 \rangle$  and  $\langle t \rangle \leq \langle m_2 \rangle$ . This tells us  $[X]_p \gamma$  and  $[Y]_p \gamma$  are powers of  $[XY]_p$  and  $[YX]_p$  respectively, so

$$(3.5) \quad \gamma = [X]_p^{-1} ([XY]_p)^u = [Y]_p^{-1} ([YX]_p)^v.$$

The powers  $u$  and  $v$  can be either zero, positive, or negative. Once we make all possible cancellations in Equation (3.5), we will have two geodesic loops, not necessarily closed geodesics, (one on each side of the equation) formed by products of  $X$ ,  $Y$ , or their inverses, representing the same homotopy class in  $\pi_1(F, p)$ . Therefore these geodesic lassos must coincide. The geodesic

on the left hand side of Equation (3.5) can begin by going along either  $X^{-1}$  or  $Y$  (depending on the sign of  $u$ ), while the geodesic on the right hand side can begin along either  $Y^{-1}$  or  $X$  (depending on the sign of  $v$ ). Therefore  $[X]_p$  and  $[Y]_p$  must either be powers of the same loop, which is impossible, because we assumed  $[h]$  is not a power of another class, or  $[X]_p$  and  $[Y]_p$  must be trivial, which is impossible because of the definition of  $\mathcal{SL}_0$ . Therefore the terms of  $\mu$  corresponding to  $p_{i,j}$  and  $q_{k,l}$  cannot cancel when  $p = q = h(t_1) = h(t_2)$ .

Now we will show that the terms of  $\mu$  which correspond to  $p_{i,j}$  and  $q_{k,l}$  cannot cancel when  $p$  and  $q$  correspond to different ordered pairs in  $\mathcal{SL}_0$ . Let  $X = h_p^1, Y = h_p^2, Z = h_q^1$ , and  $W = h_q^2$ . By Lemma 3.1 the terms which  $p_{i,j}$  and  $q_{k,l}$  contribute to  $\mu$  are:

$$[(XY)^I X \bullet_p (YX)^J Y] - [(YX)^J Y \bullet_p (XY)^I X]$$

and

$$[(ZW)^K Z \bullet_q (WZ)^L W] - [(WZ)^L W \bullet_q (ZW)^K Z],$$

where  $I + J = K + L = n - 1$ . We will suppose that

$$[(XY)^I X \bullet_p (YX)^J Y] = [(WZ)^L W \bullet_q (ZW)^K Z],$$

and derive a contradiction. Switching the orders of the two loops on both sides of the equation gives us the equality

$$[(YX)^J Y \bullet_p (XY)^I X] = [(ZW)^K Z \bullet_q (WZ)^L W],$$

so if we assume that one of these equalities holds, all four terms above will cancel.

As in the case where  $p = q$ , we will use the equality

$$[(XY)^I X \bullet_p (YX)^J Y] = [(WZ)^L W \bullet_q (ZW)^K Z]$$

to find abelian subgroups of  $\pi_1(F, q)$ . To do this, we examine the Gauss diagram of  $h$  with two oriented chords corresponding to the self-intersection points  $p$  and  $q$ . The four possible Gauss diagrams with two oriented chords are pictured in Figure 14.

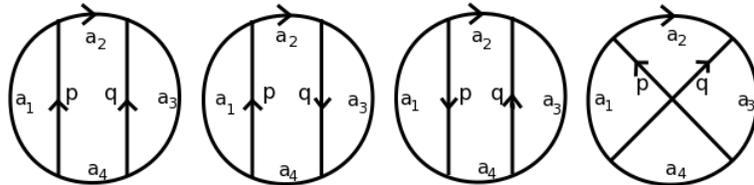


FIGURE 14. (a.) – (d.) The four Gauss diagrams of an oriented loop  $h$  with two self-intersection points.

We use the convention that each oriented chord points from the second branch of  $h$  to the first branch of  $h$ , where the branches of  $h$  at a self-intersection point are ordered according to the orientation of  $F$ . As shown in Figure 14, we let  $a_i, i = 1, \dots, 4$ , denote the arcs between the preimages of  $p$  and  $q$ . We let  $b_i$  denote the image of the arc  $a_i$  under  $h$ .

We first change the basepoint of the first term from  $p$  to  $q$ , replacing  $[(XY)^I X \bullet_p (YX)^J Y]$  by  $[b_2^{-1}(XY)^I X b_2 \bullet_q b_2^{-1}(YX)^J Y b_2]$ . Assuming the terms cancel, we can find  $\gamma \in \pi_1(F, q)$  such that

$$(3.6) \quad \gamma[(WZ)^L W]_q \gamma^{-1} = [b_2^{-1}(XY)^I X b_2]_q$$

and

$$(3.7) \quad \gamma[(ZW)^K Z]_q \gamma^{-1} = [b_2^{-1}(YX)^J Y b_2]_q.$$

Multiplying Equations (3.6) and (3.7) in both possible orders, and using the fact that  $n - 1 = I + J = K + L$ , we have:

$$(3.8) \quad \gamma[(WZ)^n]_q \gamma^{-1} = [b_2^{-1}(XY)^n b_2]_q$$

and

$$(3.9) \quad \gamma[(ZW)^n]_q \gamma^{-1} = [b_2^{-1}(YX)^n b_2]_q.$$

Table 1 lists the values of  $X, Y, Z$  and  $W$  in terms of the  $b_i$  for each Gauss diagram in Figure 14. This allows us to rewrite Equations (3.8) and (3.9)

TABLE 1. The values of  $X, Y, Z$  and  $W$  for each Gauss diagram in Figure 14.

Gauss Diagram	$X = h_p^1$	$Y = h_p^2$	$Z = h_q^1$	$W = h_q^2$
(a.)	$b_2 b_3 b_4$	$b_1$	$b_3$	$b_4 b_1 b_2$
(b.)	$b_2 b_3 b_4$	$b_1$	$b_4 b_1 b_2$	$b_3$
(c.)	$b_1$	$b_2 b_3 b_4$	$b_3$	$b_4 b_1 b_2$
(d.)	$b_2 b_3$	$b_4 b_1$	$b_3 b_4$	$b_1 b_2$

just in terms of  $\gamma$  and the  $b_i$ . Note that for diagrams (b.) and (c.), we get the same two equations from 3.8 and 3.9, because the values of  $X$  and  $Y$ , and the values of  $Z$  and  $W$ , are interchanged. Therefore it suffices to consider diagrams (a.), (b.), and (d.). The arguments for diagrams (a.) and (b.) are similar, so we will only examine (a.) and (d.).

**Diagram (a.).** In this case,  $X = b_2 b_3 b_4, Y = b_1, Z = b_3$ , and  $W = b_4 b_1 b_2$  (see Table 1). We express Equations (3.8) and (3.9) in terms of the  $b_i$  to obtain

$$(3.10) \quad \gamma[(b_4 b_1 b_2 b_3)^n]_q \gamma^{-1} = [b_2^{-1}(b_2 b_3 b_4 b_1)^n b_2]_q$$

and

$$(3.11) \quad \gamma[(b_3 b_4 b_1 b_2)^n]_q \gamma^{-1} = [b_2^{-1}(b_1 b_2 b_3 b_4)^n b_2]_q.$$

We conjugate Equation (3.10) by  $[b_3]_q^{-1} \in \pi_1(F, q)$  and Equation (3.11) by  $[b_3b_4b_2]_q \in \pi_1(F, q)$  to obtain the equations

$$[b_3]_q^{-1} \gamma [(b_4b_1b_2b_3)^n]_q \gamma^{-1} [b_3]_q = [(b_4b_1b_2b_3)^n]_q$$

and

$$[b_3b_4b_2]_q \gamma [(b_3b_4b_1b_2)^n]_q \gamma^{-1} [b_3b_4b_2]_q^{-1} = [(b_3b_4b_1b_2)^n]_q.$$

Thus  $[b_3]_q^{-1} \gamma$  and  $[(b_4b_1b_2b_3)^n]_q$  commute, as do  $[b_3b_4b_2]_q \gamma$  and  $[(b_3b_4b_1b_2)^n]_q$ . Since abelian subgroups of  $\pi_1(F, q)$  are infinite cyclic, the subgroups

$$\langle [b_3]_q^{-1} \gamma, [(b_4b_1b_2b_3)^n]_q \rangle \quad \text{and} \quad \langle [b_3b_4b_2]_q \gamma, [(b_3b_4b_1b_2)^n]_q \rangle$$

are generated by elements  $s$  and  $t$  in  $\pi_1(F, q)$  respectively. Each nontrivial element of  $\pi_1(F, q)$  is contained in a unique, maximal infinite cyclic group by Fact (2). Let  $m_1$  and  $m_2$  be the generators of the unique maximal infinite cyclic groups containing  $[(b_4b_1b_2b_3)^n]_q$  and  $[(b_3b_4b_1b_2)^n]_q$  respectively. By assumption,  $[h] = [b_4b_1b_2b_3] = [b_3b_4b_1b_2]$  is not freely homotopic to a power of another class. Therefore  $\langle m_1 \rangle = \langle [b_4b_1b_2b_3]_q \rangle$  and  $\langle m_2 \rangle = \langle [b_3b_4b_1b_2]_q \rangle$ . But  $\langle s \rangle$  and  $\langle t \rangle$  are also infinite cyclic groups containing  $[(b_4b_1b_2b_3)^n]_q$  and  $[(b_3b_4b_1b_2)^n]_q$ , respectively, so by the maximality of  $\langle m_1 \rangle$  and  $\langle m_2 \rangle$ , we have  $\langle s \rangle \leq \langle m_1 \rangle$  and  $\langle t \rangle \leq \langle m_2 \rangle$ . Thus

$$[b_3]_q^{-1} \gamma = ([b_4b_1b_2b_3]_q)^u \quad \text{and} \quad [b_3b_4b_2]_q \gamma = ([b_3b_4b_1b_2]_q)^v$$

for some  $u$  and  $v \in \mathbb{Z}$ . Now

$$\gamma = [b_3(b_4b_1b_2b_3)^u]_q = [b_2^{-1}b_4^{-1}b_3^{-1}(b_3b_4b_1b_2)^v]_q,$$

so the path homotopy classes  $[b_2b_3(b_4b_1b_2b_3)^u]_q^p$  and  $[b_4^{-1}b_3^{-1}(b_3b_4b_1b_2)^v]_q^p$  are equal. Once we cancel  $b_i$  with  $b_i^{-1}$  wherever possible,  $p_1 = b_2b_3(b_4b_1b_2b_3)^u$  and  $p_2 = b_4^{-1}b_3^{-1}(b_3b_4b_1b_2)^v$  will be two geodesic arcs from  $p$  to  $q$  representing the same path homotopy class. Therefore  $p_1$  and  $p_2$  must coincide. Note that some of the  $b_i$  may be trivial. We know  $b_1$  and  $b_3$  cannot be trivial because of the definition of  $\mathcal{SI}_0$ . Given that  $b_2$  or  $b_4$  may be trivial, and that  $u$  and  $v$  may be positive, negative, or zero, we see that  $p_1$  can begin along  $b_2$ ,  $b_3$ , or  $b_1^{-1}$  and  $p_2$  can begin along  $b_4^{-1}$ ,  $b_3^{-1}$ , or  $b_1$ . Thus  $p_1$  and  $p_2$  can only coincide if the beginnings of the arcs  $b_i$  and  $b_j^{\pm 1}$  coincide, and the initial velocity vectors of these arcs coincide for some  $i \neq j$ . This is impossible since  $h$  is a geodesic which is not homotopic to a power of another loop.

**Diagram (d).** In this case,  $X = b_2b_3$ ,  $Y = b_4b_1$ ,  $Z = b_3b_4$ , and  $W = b_1b_2$ ; see Table 1. We rewrite Equations (3.8) and (3.9) in terms of the  $b_i$  to obtain

$$(3.12) \quad \gamma [(b_1b_2b_3b_4)^n]_q \gamma^{-1} = [b_2^{-1}(b_2b_3b_4b_1)^n b_2]_q$$

and

$$(3.13) \quad \gamma [(b_3b_4b_1b_2)^n]_q \gamma^{-1} = [b_2^{-1}(b_4b_1b_2b_3)^n b_2]_q.$$

We conjugate Equation (3.12) by  $[b_1b_2]_q \in \pi_1(F, q)$  and Equation (3.13) by  $[b_3b_2]_q \in \pi_1(F, q)$  to obtain the equations

$$[b_1b_2]_q\gamma[(b_1b_2b_3b_4)^n]_q\gamma^{-1}[b_1b_2]_q^{-1} = [(b_1b_2b_3b_4)^n]_q$$

and

$$[b_3b_2]_q\gamma[(b_3b_4b_1b_2)^n]_q\gamma^{-1}[b_3b_2]_q^{-1} = [(b_3b_4b_1b_2)^n]_q.$$

Thus  $[b_1b_2]_q\gamma$  and  $[(b_1b_2b_3b_4)^n]_q$  commute, as do  $[b_3b_2]_q\gamma$  and  $[(b_3b_4b_1b_2)^n]_q$ . Since abelian subgroups of  $\pi_1(F, q)$  are infinite cyclic, the subgroups

$$\langle [b_1b_2]_q\gamma, [(b_1b_2b_3b_4)^n]_q \rangle \quad \text{and} \quad \langle [b_3b_2]_q\gamma, [(b_3b_4b_1b_2)^n]_q \rangle$$

are generated by elements  $s$  and  $t$  in  $\pi_1(F, q)$  respectively. Each nontrivial element of  $\pi_1(F, q)$  is contained in a unique, maximal infinite cyclic group by Fact (2). Let  $m_1$  and  $m_2$  be the generators of the unique, maximal infinite cyclic groups containing  $[(b_1b_2b_3b_4)^n]_q$  and  $[(b_3b_4b_1b_2)^n]_q$ , respectively. Since  $[h] = [b_1b_2b_3b_4] = [b_3b_4b_1b_2]$  is not freely homotopic to a power of another class, we have  $\langle m_1 \rangle = \langle [b_1b_2b_3b_4]_q \rangle$  and  $\langle m_2 \rangle = \langle [b_3b_4b_1b_2]_q \rangle$ . But  $\langle s \rangle$  and  $\langle t \rangle$  are also infinite cyclic groups containing  $[(b_1b_2b_3b_4)^n]_q$  and  $[(b_3b_4b_1b_2)^n]_q$ , respectively. Thus by the maximality of the  $\langle m_i \rangle$ , we have  $\langle s \rangle \leq \langle m_1 \rangle$  and  $\langle t \rangle \leq \langle m_2 \rangle$ . Hence  $[b_1b_2]_q\gamma = ([b_1b_2b_3b_4]_q)^u$  and  $[b_3b_2]_q\gamma = ([b_3b_4b_1b_2]_q)^v$  for some  $u$  and  $v \in \mathbb{Z}$ . Now

$$\gamma = [b_2^{-1}b_1^{-1}(b_1b_2b_3b_4)^u]_q = [b_2^{-1}b_3^{-1}(b_3b_4b_1b_2)^v]_q,$$

so the path homotopy classes  $[b_1^{-1}(b_1b_2b_3b_4)^u]_q^p$  and  $[b_3^{-1}(b_3b_4b_1b_2)^v]_q^p$  are equal. Once we cancel  $b_i$  with  $b_i^{-1}$  wherever possible,  $p_1 = b_1^{-1}(b_1b_2b_3b_4)^u$  and  $p_2 = b_3^{-1}(b_3b_4b_1b_2)^v$  will be two geodesic arcs from  $p$  to  $q$  representing the same path homotopy class. Therefore  $p_1$  and  $p_2$  must coincide. Again, some of the  $b_i$  may be trivial. Because of the definition of  $\mathcal{SZ}_0$ , adjacent arcs (e.g.  $b_2$  and  $b_3$  or  $b_4$  and  $b_1$ ) cannot both be trivial. If arcs opposite each other (e.g.  $b_2$  and  $b_4$ ) are both trivial, then  $p = q$ ; that case was already examined. So we may assume at most one of the  $b_i$  is trivial. Depending on whether  $u$  is positive, negative, or zero, and on which  $b_i$  is trivial,  $p_1$  can begin along either  $b_1^{-1}$ ,  $b_2$ ,  $b_4^{-1}$ ,  $b_3$ , or  $p_1$  can be trivial (if  $u = 0$  and  $b_1$  is trivial). Similarly,  $p_2$  can begin along  $b_4$ ,  $b_1$ ,  $b_3^{-1}$ , or  $b_2^{-1}$ , or  $p_2$  can be trivial (if  $v = 0$  and  $b_3$  is trivial). Therefore, in order for the  $p_i$  to coincide, either the beginnings and initial velocity vectors of the arcs  $b_i$  and  $b_j^{\pm 1}$  must coincide for some  $i \neq j$ , which is impossible since  $h$  is a geodesic and is not a power of another loop, or both  $p_i$  must be trivial. But if both  $p_i$  are trivial, then  $b_1$  and  $b_3$  are both trivial, and we assumed at most one of the  $b_i$  are trivial, so this is impossible as well.  $\square$

The following lemma allows us to reduce the proofs of Theorems 1.1 and 1.3 to the case where  $F$  is compact.

**3.3. Lemma.** *Suppose  $F$  is noncompact, and let  $g : S^1 \rightarrow F$ . Suppose  $t(\mu([g])) = T$ , where  $[g] \in \hat{\pi}(F)$ . Then there exists a compact subsurface  $F_C$*

of  $F$  containing  $\text{Im}(g)$  such that  $t(\mu([g]_C)) = T$ , where  $[g]_C$  is the class of  $g$  in  $\hat{\pi}(F_C)$ .

**Proof.** First note that we may assume  $g$  has finitely many transverse self-intersection points (if it did not, we could perturb  $g$  slightly to obtain a loop that does). Suppose that the pair of terms  $[g_p^1 \bullet_p g_p^2]$  and  $[g_q^2 \bullet_q g_q^1]$  of  $\mu([g])$  cancel. Let  $D$  be a chord diagram with one chord attached to one copy of  $S^1$  at its endpoints. Let  $D_p : D \rightarrow F$  and  $D_q : D \rightarrow F$  be the geometrical chord diagrams associated with  $[g_p^1 \bullet_p g_p^2]$  and  $[g_q^2 \bullet_q g_q^1]$  respectively. Since  $[g_p^1 \bullet_p g_p^2]$  and  $[g_q^2 \bullet_q g_q^1]$  cancel, we have a homotopy  $H_{p,q} : D \times [0, 1] \rightarrow F$  between  $D_p$  and  $D_q$ . Note this homotopy takes  $g_p^1$  to  $g_q^2$  and takes  $g_p^2$  to  $g_q^1$ . Since  $\text{Im}(H_{p,q})$  is compact, and since  $g$  has finitely many self-intersection points, we may choose a compact subsurface  $F_C$  of  $F$  containing  $\text{Im}(H_{p,q})$  for all pairs  $(p, q)$  corresponding to terms that cancel. Note that once  $F_C$  contains  $\text{Im}(H_{p,q})$ ,  $F_C$  must also contain the image of  $g$ . If we compute  $\mu([g]_C)$  on  $F_C$ , the terms  $[g_p^1 \bullet_p g_p^2]$  and  $[g_q^2 \bullet_q g_q^1]$  will cancel, since  $H_{p,q}$  can be viewed as a homotopy in  $F_C$ . Thus  $t(\mu([g])) \geq t(\mu([g]_C))$ . Furthermore, any terms which cancel on  $F_C$  must cancel on  $F$ , so the inequality becomes an equality.  $\square$

Now we state our main results. Theorems 1.1 and 1.3 are stated as Corollaries 3.4 and 3.5 of Theorem 3.2, respectively, though for now we restrict Theorem 1.3 to the case where  $\alpha$  is not a power of another class. Recall that Theorem 1.1 states that  $\mu(\alpha)$  is zero if and only if  $\alpha$  is a power of a simple class, and Theorem 1.3 gives a formula for the minimal self-intersection number of  $\alpha$ .

**3.4. Corollary.** *Let  $\alpha \in \hat{\pi}(F)$ . Then  $\mu(\alpha) = 0$  if and only if  $\alpha = \beta^n$  for some  $\beta \in \pi_1(F)$ , where  $m(\beta) = 0$ .*

**Proof.** Suppose  $\alpha = \beta^n$  for some  $\beta \in \pi_1(F)$ , where  $m(\beta) = 0$ . We may assume  $n > 0$ . We will compute  $\mu(\alpha)$ , and see that  $\mu(\alpha) = 0$ . We begin by choosing a simple representative  $h$  of  $\beta$  and a point  $p$  on  $h$ . Then  $g = (h_p)^n$  is a representative of  $\alpha$ . We perturb  $g$  slightly so that it has  $n - 1$  self-intersection points (of Type 2), all with image  $p$ . Now

$$\mu(\alpha) = \sum_{i=1}^{n-1} [(h_p)^i \bullet_p (h_p)^{n-i}] - [(h_p)^{n-i} \bullet_p (h_p)^i],$$

which equals zero, since the positive term corresponding to  $i = k$  cancels with the negative term corresponding to  $i = n - k$ . Note that this argument actually shows that the terms of  $\mu$  corresponding to Type 2 self-intersection points always cancel with each other, even if  $m(\beta) \neq 0$ .

To prove the converse, we assume  $\alpha$  cannot be written as a power of a simple class. By Lemma 3.3, we may assume  $F$  is compact. Therefore, either  $F = S^2, A$ , or  $T^2$ , or  $F$  can be equipped with a metric of negative curvature. In this situation,  $F$  clearly cannot be  $S^2$  or  $A$ .

$F$  also cannot be  $T^2$ . We will show that every free homotopy class on the torus is a power of a simple class (this fact is well-known). Let  $g$  be a representative of  $\alpha$ , and consider  $[g]_p \in \pi_1(T^2, p)$  for some  $p$  on the image of  $g$ . Let  $a$  and  $b$  be the generators of  $\pi_1(T^2, p)$  which lift to the paths from  $(0, 0)$  to  $(1, 0)$  and  $(0, 1)$  respectively under the usual covering map  $\mathbb{R}^2 \rightarrow \mathbb{R}^2/\mathbb{Z}^2$ . Then we can write  $[g]_p = ([a^{e_1}b^{e_2}]_p)^k$  where the  $e_i$  are relatively prime. Taking  $\beta = [a^{e_1}b^{e_2}]_p$ , we have  $m(\beta) = 0$ , since the lift of  $a^{e_1}b^{e_2}$  is homotopic to a path from  $(0, 0)$  to  $(e_1, e_2)$ .

Now we may assume  $F \neq S^2, A$  or  $T^2$ , so we can apply Theorem 3.2. We choose a geodesic representative  $g$  of  $\alpha$ , and write  $g = h^n$  where  $h$  is not a power of another loop in  $\pi_1(F)$ . Suppose  $h$  has  $m$  self-intersection points. Then we can perturb  $g$  to obtain a loop  $g'$  with  $mn^2 + n - 1$  self-intersection points, where  $mn^2$  of the self-intersections are of Type 1, and  $n - 1$  are of Type 2. By Theorem 3.2, no two terms of  $\mu$  corresponding to Type 1 self-intersection points can cancel. As we saw above, all of the Type 2 terms cancel with each other. So, after all cancellations are made,  $2mn^2$  terms of  $\mu$  remain. Since  $m \geq 1$ , we know  $\mu(\alpha) \neq 0$ .  $\square$

**3.5. Corollary.** *Let  $\alpha \in \hat{\pi}(F)$  be primitive. Then  $m(\alpha) = t(\mu(\alpha))/2$ . Thus  $\mu$  computes the minimal self intersection number of  $\alpha$ .*

**Proof.** By Lemma 3.3, we may assume  $F$  is compact. If  $\alpha \neq \beta^n$  for any  $\beta \in \pi_1(F)$  and  $|n| > 1$ , we can choose a geodesic representative  $g$  of  $\alpha$  such that the self-intersection points of its perturbation  $g'$  are all of Type 1. If  $F \neq S^2, A$  or  $T^2$ , then by Theorem 3.2, no two terms of  $\mu(\alpha)$  can cancel. Thus  $m(\alpha) = t(\mu(\alpha))/2$ .

If  $F = S^2, T^2$  or  $A$ , we can choose a representative of  $\alpha$  with no self-intersection points, implying that the equality  $m(\alpha) = t(\mu(\alpha))/2$  holds trivially. This is clear for  $S^2$  and  $A$ . If  $F = T^2$ , we choose a representative  $g$  of  $\alpha$ , and consider  $[g]_p \in \pi_1(T^2, p)$  for some point  $p$  on the image of  $g$ . If  $a$  and  $b$  are generators the of  $\pi_1(T^2, p)$  as in the proof of Corollary 3.4, we can write  $[g]_p = ([a^{e_1}b^{e_2}]_p)^k$  where the  $e_i$  are relatively prime. Since  $\alpha$  is not a power of another class,  $k = \pm 1$ , so  $[g]_p = ([a^{e_1}b^{e_2}]_p)^{\pm 1}$ . This class has a simple representative, as shown in the proof of Corollary 3.4.  $\square$

**3.3. Using  $\mu$  to compute the minimal self-intersection number of a class which is not primitive.** If we combine our results above with those of Hass and Scott in [17], we can use  $\mu$  to compute the minimal self-intersection number of any class  $\alpha$ , even if  $\alpha$  is not primitive. We will use the following result, which we state for orientable surfaces (their result is more general):

**3.6. Lemma** (Hass, Scott). *Let  $f$  be a loop on  $F \neq S^2$  in general position. Suppose that  $f$  is a representative of the class  $\alpha = \beta^n$ , where  $\beta$  is a nontrivial and primitive element of  $\pi_1(F)$ . Let  $\tilde{F}$  be the universal cover of  $F$  and let  $F_\beta$  denote the quotient of  $\tilde{F}$  by the cyclic subgroup of  $\pi_1(F)$  generated by  $\beta$ .*

Let  $f_\beta : S^1 \rightarrow F_\beta$  be the lift of  $f$ , and let  $l$  denote one of the lines in  $\tilde{F}$  above  $f_\beta(S^1)$ . Then  $f$  has least possible self-intersection if and only if  $f_\beta$  has least possible self-intersection and, for all  $\gamma$  in  $\pi_1(F)$  such that  $\gamma$  is not a power of  $\beta$ , the intersection  $\gamma l \cap l$  consists of at most one point.

We now use Lemma 3.6 to get the following corollary, which gives a formula for  $m(\alpha)$  in terms of  $m(\beta)$ , where  $\alpha$  is a power of  $\beta$  and  $\beta$  is primitive. For a similar statement for compact surfaces with boundary, see the remark following Theorem 2 in [21]. (Note that as it seems, the statement in [21] contains a typo, namely  $ps^2 + (p - 1)$  should be  $ps^2 + (s - 1)$ ).

**3.7. Corollary.** *Let  $\alpha \in \hat{\pi}(F)$  be a nontrivial class such that  $\alpha = \beta^n$  and  $\beta$  is a primitive class, where  $F \neq S^2$  and  $n \geq 1$ . Then  $m(\alpha) = n^2m(\beta) + n - 1$ .*

**Proof.** First we remark that one can deduce this fact from the proof of Lemma 3.6 in [17], rather than from the statement of the lemma itself. We choose to do the latter to avoid repeating most of the proof of Lemma 3.6.

Equip  $F$  with a flat or hyperbolic metric. Then  $\tilde{F}$  is diffeomorphic to  $\mathbb{R}^2$ , so  $F_\beta$  is a cylinder. We begin by choosing a geodesic representative  $g$  of  $\alpha$  and perturb  $g$  to obtain a new loop  $g'$ . If  $n = 1$ , we take  $g' = g$ . Otherwise, we lift  $g$  to obtain a geodesic loop  $g_\beta$  in  $F_\beta$ . We then take a  $C^\infty$ -small perturbation of  $g_\beta$  to obtain a loop  $g'_\beta$  with  $n - 1$  self-intersection points, and we let  $g'$  be the projection of  $g'_\beta$  back to  $F$ . We note that  $g'_\beta$  has minimal self-intersection (see, for example, [24]). As mentioned in the discussion leading up to Theorem 3.2, the number of self-intersection points of  $g'$  is  $n^2m(\beta) + n - 1$ . Let  $l'$  denote one of the lines above  $g'$  in  $\tilde{F}$ . To show that  $g'$  has minimal self-intersection, it remains to show that for all  $\gamma \in \pi_1(F)$  such that  $\gamma$  is not a power of  $\beta$ , the intersection  $\gamma l' \cap l'$  consists of at most one point. Suppose that for some  $\gamma$ , the lines  $\gamma l'$  and  $l'$  intersect in two points  $p$  and  $q$ . These lines bound a 2-gon in  $\tilde{F}$ . Let  $A_1$  and  $A_2$  denote the two arcs of the 2-gon in  $\tilde{F}$ . Let  $\rho$  denote the covering map from  $\tilde{F}$  to  $F$ . Since  $A_1$  and  $A_2$  are homotopic with common ends, their projections  $\rho(A_1)$  and  $\rho(A_2)$  must be as well. Furthermore the arcs  $\rho(A_1)$  and  $\rho(A_2)$  are contained in the image of  $g'$ . Since  $g'$  is a perturbation of a geodesic,  $\rho(p)$  and  $\rho(q)$  must be Type 2 self-intersection points of  $g'$ . Thus each  $\rho(A_i)$  is an arc in the image of  $g'$  connecting two Type 2 self-intersection points, and hence must be homotopic to  $g^k$  for some  $k \neq 0$ . This implies  $\gamma$  is a power of  $\beta$ . Hence if  $\gamma$  is not a power of  $\beta$ , the intersection of  $\gamma l'$  and  $l'$  contains at most one point.  $\square$

Recall that, if we make our usual choice of a perturbation of a geodesic  $g' \in \alpha$ , then two terms of  $\mu(\alpha)$  cancel if and only if they correspond to Type 2 self-intersection points. (The observation that terms corresponding to Type 2 self-intersection points must cancel with other such terms is contained in the proof of Corollary 3.4.) Together with Corollary 3.7, this allows us to

use  $\mu$  to compute the minimal self-intersection number of a class which may not be primitive.

**3.8. Corollary.** *Let  $\alpha \in \hat{\pi}(F)$  be a nontrivial class such that  $\alpha = \beta^n$  where  $\beta$  is primitive and  $n \geq 1$ . Then  $m(\alpha) = t(\mu(\alpha))/2 + n - 1$ .*

**Proof.** By Theorem 3.2, we know  $t(\mu(\alpha))/2$  is equal to  $m(\beta)n^2$ , i.e., the number of Type 1 self-intersection points. Since  $m(\alpha) = m(\beta)n^2 + n - 1$ , we have  $m(\alpha) = t(\mu(\alpha))/2 + n - 1$ .  $\square$

### 4. An example

The following example illustrates how one can compute the minimal self-intersection number of a free homotopy class algorithmically using  $\mu(\alpha)$ . We will use  $\mu$  to compute the minimal intersection number of a class

$$\alpha = a_3a_1\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2\bar{a}_2\bar{a}_2$$

on the punctured surface of genus two, with surface word  $a_1a_2\bar{a}_1\bar{a}_2a_3a_4\bar{a}_3\bar{a}_4$ . This homotopy class has Turaev cobracket zero, as shown in Example 5.8 of [9]. A representative of  $\alpha$  with two self-intersection points is pictured in Figure 15. Using  $\mu$ , along with the methods in [13], we show that the minimal self-intersection number of  $\alpha$  is 2.

We compute  $\mu$ , and find that

$$\begin{aligned} \mu(\alpha) &= a_3a_1\bar{a}_2a_3a_1\bar{a}_2\bar{a}_2 \bullet_p \bar{a}_2a_3a_1\bar{a}_2 - \bar{a}_2a_3a_1\bar{a}_2 \bullet_p a_3a_1\bar{a}_2a_3a_1\bar{a}_2\bar{a}_2 \\ &+ a_3a_1\bar{a}_2\bar{a}_2 \bullet_q \bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2 - \bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2 \bullet_q a_3a_1\bar{a}_2\bar{a}_2. \end{aligned}$$

It is easy to see that terms one and two cannot cancel, since there is no  $t \in \pi_1(F)$  such that  $ta_3a_1\bar{a}_2a_3a_1\bar{a}_2\bar{a}_2t^{-1} = \bar{a}_2a_3a_1\bar{a}_2$ . In general, one can check whether two reduced words in the generators of  $\pi_1$  are in the same conjugacy class by comparing two cyclic lists, since the conjugacy class of a reduced word in a free group consists of all the cyclic permutations of that word. Similarly terms three and four cannot cancel.

Next, we show terms one and four cannot cancel. First we conjugate term one by  $\bar{a}_2$  so that its first loop matches the first loop of term four. The first term becomes

$$\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2 \bullet_p \bar{a}_2\bar{a}_2a_3a_1.$$

If terms one and four cancel, we can find  $s \in \pi_1(F)$  such that  $s$  and  $\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2$  commute, and such that

$$s\bar{a}_2\bar{a}_2a_3a_1s^{-1} = a_3a_1\bar{a}_2\bar{a}_2.$$

Abelian subgroups of  $\pi_1(F)$  are infinite cyclic, so the subgroup

$$\langle s, \bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2 \rangle$$

is generated by some  $u \in \pi_1(F)$ . We will see that  $\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2 = u^{\pm 1}$ . Suppose  $\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2 = u^i$ , and consider this relation in the abelianization of  $\pi_1(F)$ , where the generator  $a_j$  of  $\pi_1(F)$  is sent to the generator of

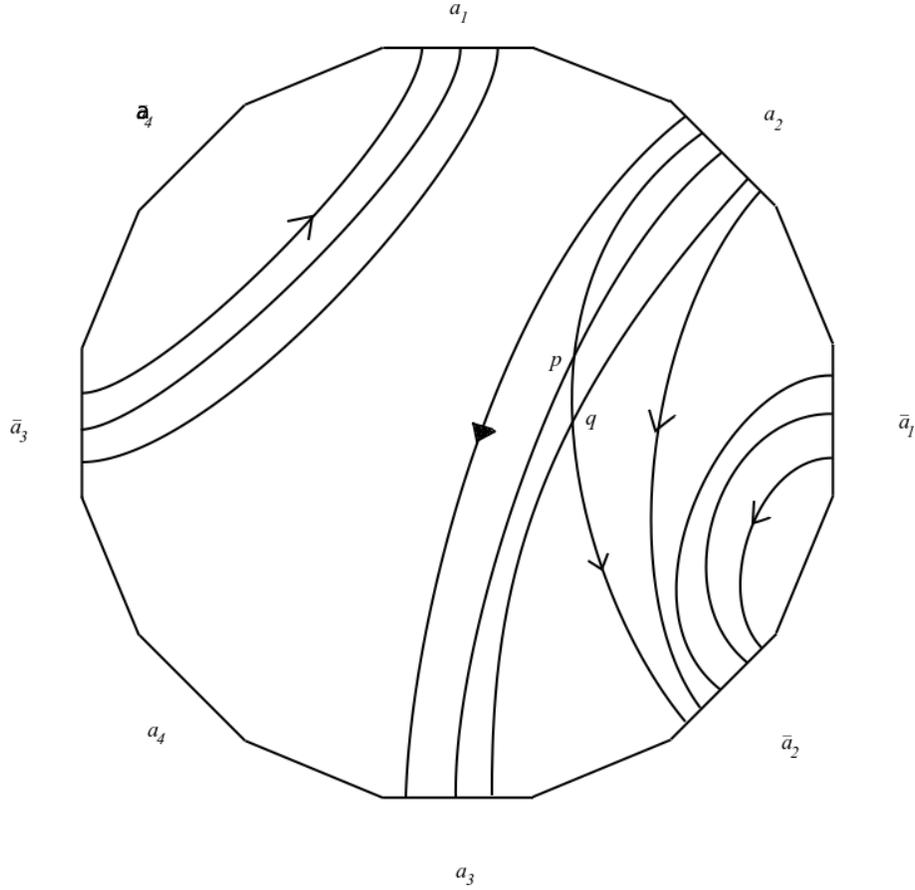


FIGURE 15. The class  $a_3a_1\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2\bar{a}_2\bar{a}_2$  on the punctured surface of genus two. Two match this word with the curve in the picture, begin at the solid arrow.

$\mathbb{Z}^4$  with a 1 in position  $j$  and zeroes elsewhere. Hence this relation becomes  $(2, -3, 2, 0) = i(u_1, u_2, u_3, u_4)$ , which only has solutions when  $i = \pm 1$ . Therefore  $s = (\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2)^k$  for some  $k \in \mathbb{Z}$ . However, the relation

$$(\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2)^k \bar{a}_2\bar{a}_2a_3a_1(\bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2)^{-k} = a_3a_1\bar{a}_2\bar{a}_2$$

cannot hold for any value of  $k$ . Hence none of the terms of  $\mu$  cancel, and  $m(\alpha) = 2$ .

In general, when deciding whether two terms cancel, we first verify that the first loops in each term are indeed in the same conjugacy class. If this is the case, we will have two elements  $s$  and  $t$  that commute ( $t = \bar{a}_2a_3a_1\bar{a}_2a_3a_1\bar{a}_2$  in the example above), and we know  $\langle s, t \rangle = \langle u \rangle$ . Since every nontrivial element of  $\pi_1$  is contained in a unique maximal infinite cyclic subgroup, we can write  $\langle s, t \rangle = \langle u \rangle \leq \langle m \rangle$  where  $m$  is the generator of the maximal infinite cyclic subgroup containing  $u$ . We need to find  $m$

given  $t$ , as our goal is to write  $s$  as a power of  $m$ . To find  $m$  given  $t$ , we first cyclically reduce  $t$  and write the result as a word  $t_r$  in the generators of  $\pi_1$ . Since  $t_r$  is cyclically reduced, then in order to be a power of another element, it must look like a concatenation of  $i$  copies of some cyclically reduced word  $w$  in the generators of  $\pi_1$ , so we can determine  $i$  and  $w$  where  $i$  is as large (in absolute value) as possible. Now  $t_r = w^i$ , and there exists  $c \in \pi_1$  such that  $t = ct_r c^{-1} = (cwc^{-1})^i$ , so  $m = cwc^{-1}$ . We can now write  $s$  as a power of  $m$  and finish the algorithm as in the example above.

### 5. Algebraic properties of $\mu$

We conclude by stating properties of  $\mu$  which allow one to view  $\mu$  as a generalization of a Lie cobracket. In particular, we exhibit analogues of the following properties of the Turaev cobracket  $\Delta$ :

- $\Delta$  satisfies co-skew symmetry:  $\tau \circ \Delta = -\Delta$ , where  $\tau(a \otimes b) = b \otimes a$ .
- $\Delta$  satisfies the co-Jacobi identity:  $(1 + \omega + \omega^2) \circ Id \otimes \Delta \circ \Delta = 0$ , and  $\omega(a \otimes b \otimes c) = b \otimes c \otimes a$ .

We begin by modifying the definition of  $\mu$  given for free loops, and then extend this operation to certain chord diagrams in the Andersen–Mattes–Reshetikhin algebra. For the purposes of computing the minimal self-intersection number, this definition is equivalent to the previous one. However, it is easier to state the analogues of co-skew symmetry and the co-Jacobi identity for this modified definition.

Recall that  $M$  denotes the free  $\mathbb{Z}$ -module generated by the set of chord diagrams on  $F$ , and  $N$  denotes the submodule generated by the  $4T$ -relations in Figure 4 (and relations obtained from them by reversing orientations on arrows). The Andersen–Mattes–Reshetikhin bracket is defined on the quotient  $\mathfrak{ch} = M/N$ . Given a chord  $p$  of a geometrical chord diagram, we say  $p$  is an *external chord* if the endpoints of  $p$  lie on distinct core circles. Otherwise, we call  $p$  an *internal chord*. Now we let  $M_e$  denote the free  $\mathbb{Z}$ -module generated by diagrams with only external chords, and let  $N_e$  denote the submodule of  $M_e$  generated by the relations in Figures 16 and 17, their mirror images, and relations obtained from these by reversing the orientation on any branch in the picture. From these relations, one can obtain the  $4T$  relations of the Andersen–Mattes–Reshetikhin algebra in which the orientations of the arcs in the four pictures are identical. We will define  $\mu$  on  $\mathfrak{ch}_e = M_e/N_e$ . From now on, we also assume that the core circles of our chord diagrams are labeled with the digits  $\{1, \dots, n\}$ , where  $n$  is the number of core circles in the diagram.

Throughout this section we invite the reader to compare our relations with those of Andersen and Turaev [3, 4].

The operation  $\mu$  will add a chord between the preimages of each self-intersection point of each core circle, in the same way as before. However,

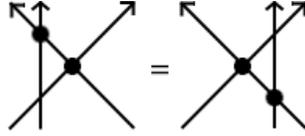


FIGURE 16. We can slide one exterior chord past another.

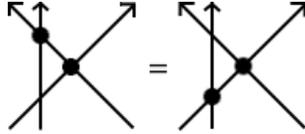


FIGURE 17. We can slide one exterior chord from one core circle to another.

rather than adding a signed chord, whose sign induces a labelling on the resulting two loops in the image of the geometrical chord diagram, we split the abstract diagram into two labeled loops connected by an unsigned chord.

We define two maps  $S_p^+$  and  $S_p^-$  which split a core circle labelled  $i$  into two new core circles labeled  $i$  and  $i + 1$  (see Figures 18 and 19). Suppose  $a : C_i \rightarrow F$  is the restriction of our original chord diagram to the  $i^{th}$  core circle. The map  $S_p^+$  (respectively  $S_p^-$ ) maps the new circle labeled  $i$  (respectively  $i + 1$ ) to  $a_p^1$  and the new circle labeled  $i + 1$  (respectively,  $i$ ) to  $a_p^2$ . The splitting map also increases by one the label on all core circles formerly labeled with numbers greater than or equal to  $i + 1$ . Note that the image of smooth chord diagram under  $S_p^\pm$  may not be smooth at  $p$ , but we can always find a smooth diagram in its homotopy class.

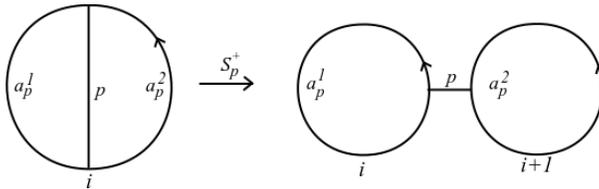


FIGURE 18. The splitting map  $S_p^+$ .

**5.1. The definitions of  $\mu_i$  and  $\mu$ .** Now let  $\mathcal{ST}_0^i$  denote the set of self-intersection points of the  $i^{th}$  core circle in our chord diagram such that neither of the maps formed by the splitting map at that point are homotopy trivial. Let  $D_p$  denote the diagram obtained by adding a chord between  $t_1$  and  $t_2$  in a chord diagram  $D$  where  $(t_1, t_2) \in \mathcal{ST}_0^i$  and  $p = D(t_1) = D(t_2)$ .

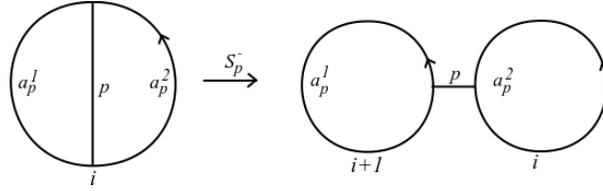


FIGURE 19. The splitting map  $S_p^-$ .

We define

$$\mu_i(\mathcal{D}) = \sum_{p \in ST_0^i} [S_p^+(D_p)] - [S_p^-(D_p)],$$

and we let

$$\mu(\mathcal{D}) = \sum_{i=1}^n \mu_i(\mathcal{D}).$$

Using linearity, we extend this definition so that  $\mu$  becomes a map from  $\mathfrak{ch}_e$  to  $\mathfrak{ch}_e$ .

**Remark.** The identification in Figure 17 arises because  $\mu$  splits the original core circle into two new circles. If we define  $\mu$  without the splitting map, this identification is not needed.

**5.2. The maps  $\mu$  and  $\mu_i$  are well-defined.** We need to verify that  $\mu_i(\mathcal{D})$  does not depend on the choice of diagram  $D$  in  $\mathcal{D}$ . Clearly  $\mu_i$  does not change when  $D$  undergoes regular isotopy. Applying the first Reidemeister move to  $D$  does not affect  $\mu_i([D])$  because we only sum over self-intersection points such that the two new maps produced by  $S_p^\pm$  are not homotopy trivial. The fact that  $\mu_i$  does not change under other elementary moves follows from either other elementary moves or identifications we make:

- *Second Reidemeister move.* This follows from the move in Figure 5.
- *Third Reidemeister move.* This follows from the move in Figure 6.
- *The move in Figure 5.* Because the existing chord is an exterior chord, we do not sum over the intersection point in this diagram because it must be an intersection point of different core circles.
- *The move in Figure 6.* This follows from the following identifications in Figures 16 and 17. Note that the  $4T$  relations in the Andersen–Mattes–Reshetikhin algebra imply that their bracket is invariant under this move. In our case, the relations only contain two terms because the fact that the chord is an exterior chord implies we only sum over at most one of the intersection points before and after the move in Figure 6.
- *The identification in Figure 16.* This follows from the following identifications in Figures 16 and 17.

- *The identification in Figure 17.* Because the chords in the picture are exterior chords, the one intersection point cannot be a self-intersection point, so we do not sum over it.

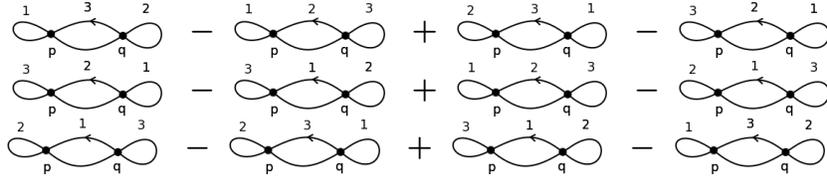


FIGURE 20. The generalized co-Jacobi identity.

**5.3. Algebraic properties of  $\mu$  and  $\mu_i$ .** Let  $\tau_i$  be a map which swaps labels the  $i$  and  $i + 1$  in a chord diagram with enumerated core circles. We have the following analogue of co-skew symmetry:

**5.1. Proposition.**  $\tau_i \circ \mu_i = -\mu_i$ .

**Proof.** Clear. □

Let  $\omega_i$  be a map which cyclically permutes the labels  $i, i + 1, i + 2$ . Specifically,  $\omega_i$  decreases the labels  $i + 1$  and  $i + 2$  by one, and sends  $i$  to  $i + 2$ . We have the following analogue of the co-Jacobi identity:

**5.2. Proposition.** For all  $1 \leq i \leq n - 1$ ,  $(1 + \omega_i + \omega_i^2) \circ \mu_{i+1} \circ \mu_i = 0$ .

**Proof.** It suffices to show that, for a diagram with one core circle  $C_1$ , we have  $(1 + \omega_1 + \omega_1^2) \circ \mu_2 \circ \mu_1 = 0$ . Each diagram in the sum  $\mu_2 \circ \mu_1(C_1)$  has three core circles and two chords corresponding to two self-intersection points of  $C_1$ , which we call  $p$  and  $q$ . The idea of the proof is as follows:  $\mu_2 \circ \mu_1(C_1)$  is a sum of four terms, two of which are positive and two of which are negative. Relative to some fixed initial order on the three core circles, the labellings on each diagram form an even or odd permutation in  $S_3$ , the symmetric group on three elements. The labellings on two of the four terms of  $\mu_2 \circ \mu_1(C_1)$  form even permutations, and the other two form odd permutations. Of the “even” terms, one has coefficient equal to  $-1$  and one has coefficient equal to  $+1$ . The same holds for the “odd” terms. Therefore, when we apply  $(1 + \omega_1 + \omega_1^2)$  to the terms with coefficient equal to  $+1$ , we get six terms with coefficient equal to  $+1$ , one for each element of  $S_3$ . When we apply  $(1 + \omega_1 + \omega_1^2)$  to the terms with coefficient equal to  $-1$ , we get the same six terms with negative coefficients, so all terms cancel. To verify the above claims, one can examine all possible Gauss diagrams of  $C_1$  with two noncrossing arrows, corresponding to the self-intersection points  $p$  and  $q$  (there are three such diagrams). Figure 16 lists the terms in the sum  $(1 + \omega_i + \omega_i^2) \circ \mu_{i+1} \circ \mu_i$  before cancellations are made for a sample free loop. □

We conclude by stating the relationship between  $\mu_i$  and the Turaev cobracket. Let  $E$  be a map which erases all chords from a diagram, and tensors the resulting loops, putting  $C_i$  in the  $i^{\text{th}}$  position of the tensor product. Let  $\Delta_i = Id \otimes \dots \otimes Id \otimes \Delta \otimes Id \otimes \dots \otimes Id$ , where  $\Delta$  is in the  $i^{\text{th}}$  position.

**5.3. Proposition.**  $E \circ \mu_i = \Delta_i \circ E$ .

**Proof.** This follows from the fact that  $E \circ \mu_1 = \Delta([\alpha])$  for any free loop  $\alpha$  on  $F$ .  $\square$

**Remark.** One might hope to find analogues of the compatibility and involutivity conditions in the Goldman–Turaev Lie bialgebra:

- $\Delta$  and  $[\cdot, \cdot]$  satisfy the compatibility condition

$$\Delta \circ [\alpha, \beta] = [\alpha, \Delta(\beta)] + [\Delta(\alpha), \beta],$$

where  $[\alpha, \beta \otimes \gamma]$  is given by  $[\alpha, \beta] \otimes \gamma + \beta \otimes [\alpha, \gamma]$ .

- The Lie bialgebra formed by  $[\cdot, \cdot]$  and  $\Delta$  is involutive, i.e.,  $[\cdot, \cdot] \circ \Delta = 0$ .

However, we do not see a way of doing this unless we allow internal chords in our chord diagrams, and once we do this, the maps involved in analogues of these identities (and  $\mu$  in particular) are not well-defined.

## Acknowledgements

I would like to thank my advisor Vladimir Chernov for his guidance, many helpful conversations, and for reading drafts of this paper. I would also like to thank Yong Su for computing examples using a preliminary definition of the operation  $\mu$ .

## References

- [1] ANDERSEN, JØRGEN ELLEGAARD; MATTES, JOSEF; RESHETIKHIN, NICOLAI. Quantization of the algebra of chord diagrams. *Math. Proc. Cambridge Philos. Soc.* **124** (1998), no. 3, 451–467. MR1636568 (99m:58040), Zbl 0943.57007, arXiv:q-alg/9701018, doi:10.1017/S0305004198002813.
- [2] ANDERSEN, JØRGEN ELLEGAARD; MATTES, JOSEF; RESHETIKHIN, NICOLAI. The Poisson structure on the moduli space of flat connections and chord diagrams. *Topology* **35** (1996), no. 4, 1069–1083. MR1404925 (98e:57005), Zbl 0857.58009, doi:10.1016/0040-9383(95)00059-3.
- [3] ANDERSEN, JØRGEN ELLEGAARD; TURAEV, VLADIMIR. Higher skein modules. *J. Knot Theory Ramifications* **8** (1999), no. 8, 963–984. MR1723433 (2001c:57013), Zbl 0991.57015, arXiv:math/9812071, doi:10.1142/S0218216599000626.
- [4] ANDERSEN, JØRGEN ELLEGAARD; TURAEV, VLADIMIR. Higher skein modules. II. Topology, ergodic theory, real algebraic geometry, 21–30. Amer. Math. Soc. Transl. Ser 2, 202, *Amer. Math. Soc., Providence, R.I.*, 2001. MR1819178 (2002e:57006), Zbl 1008.57011.
- [5] BENEDETTI, RICCARDO; PETRONIO, CARLO. Lectures on hyperbolic geometry. Universitext. *Springer-Verlag, Berlin*, 1992. xiv+330 pp. ISBN: 3-540-55534-X. MR1219310 (94e:57015), Zbl 0768.51018.
- [6] BIRMAN, JOAN S.; SERIES, CAROLINE. An algorithm for simple curves on surfaces. *J. London Math. Soc. (2)* **29** (1984), no. 2, 331–342. MR0744104 (85m:57002), Zbl 0507.57006.

- [7] BUSER, PETER. Geometry and spectra of compact Riemann surfaces. Progress in Mathematics, 106. *Birkhäuser Boston, Inc., Boston, MA*, 1992. xiv+454 pp. ISBN: 0-8176-3406-1. MR1183224 (93g:58149), Zbl 0770.53001.
- [8] CAHN, PATRICIA; CHERNOV, VLADIMIR. Intersections of loops and the Andersen–Mattes–Reshetikhin algebra. To appear in *J. London Math. Soc.*, 2013. arXiv:1105.4638.
- [9] CHAS, MOIRA. Combinatorial Lie bialgebras of curves on surfaces. *Topology* **43** (2004), no. 3, 543–568. MR2041630 (2005b:57038), Zbl 1050.57014, arXiv:math/0105178, doi:10.1016/j.top.2003.03.003.
- [10] CHAS, MOIRA. Minimal intersection of curves on surfaces. *Geom. Dedicata* **144** (2010), 25–60. MR2580416 (2011c:57001), Zbl 1186.57015, arXiv:0706.2439, doi:10.1007/s10711-009-9385-y.
- [11] CHAS, MOIRA; GADGIL, SIDDHARTHA. The Goldman bracket determines intersection numbers for surfaces and orbifolds. 2012. arXiv:1209.0634.
- [12] CHAS, MOIRA; KRONGOLD, FABIANA. Algebraic characterization of simple closed curves via Turaev’s cobracket. 2010. arXiv:1009.2620.
- [13] CHERNOV, VLADIMIR. Graded Poisson algebras on bordism groups of garlands and their applications. 2007. arXiv:math/0608153.
- [14] COHEN, MARSHALL; LUSTIG, MARTIN. Paths of geodesics and geometric intersection numbers. I. Combinatorial group theory and topology, (Altañ Utah, 1984), 479–500. *Ann. of Math. Stud.*, 111, *Princeton University Press, Princeton, NJ*, 1987. MR0895629 (88m:57016), Zbl 0624.57010.
- [15] FREEDMAN, MICHAEL; HASS, JOEL; SCOTT, PETER. Closed geodesics on surfaces. *Bull. London Math. Soc.* **14** (1982), no. 5, 385–391. MR0671777 (84b:53042), Zbl 0476.53026.
- [16] GOLDMAN, WILLIAM M. Invariant functions on Lie groups and Hamiltonian flows of surface group representations. *Invent. Math.* **85** (1986), no. 2, 263–302. MR0846929 (87j:32069), Zbl 0619.58021, doi:10.1007/BF01389091.
- [17] HASS, JOEL; SCOTT, PETER. Intersections of curves on surfaces. *Israel J. Math.* **51** (1985), no. 1–2, 90–120. MR0804478 (86k:57007), Zbl 0576.57009, doi:10.1007/BF02772960.
- [18] LE DONNE, ATTILIO. On Lie bialgebras of loops on orientable surfaces. *J. Knot Theory Ramifications* **17** (2008), no. 3, 351–359. MR2402511 (2009e:57028), Zbl 1152.17009, arXiv:math/0510659, doi:10.1142/S0218216508006178.
- [19] LUSTIG, MARTIN. Paths of geodesics and geometric intersection numbers. II. Combinatorial group theory and topology (Altañ, Utah, 1984), 501–543, *Ann. of Math. Stud.*, 111, *Princeton University Press, Princeton, NJ*, 1987. MR0895630 (88m:57017), Zbl 0662.57002.
- [20] RIVIN, IGOR. Simple curves on surfaces. *Geom. Dedicata* **87** (2001), no. 1–3, 345–360. MR1866856 (2003c:57018), Zbl 1002.53027, arXiv:math/9907041, doi:10.1023/A:1012010721583.
- [21] TAN, SER PEOW. Self-intersections of curves on surfaces. *Geom. Dedicata* **62** (1996), no. 2, 209–225. MR1407382 (97g:57001), Zbl 0977.53037, doi:10.1007/BF00147812.
- [22] TURAEV, VLADIMIR G. Skein quantization of Poisson algebras of loops on surfaces. *Ann. Sci. École Norm. Sup. (4)* **24** (1991), no. 6, 635–704. MR1142906 (94a:57023), Zbl 0758.57011.
- [23] TURAEV, V. G. Intersections of loops in two-dimensional manifolds. *Mat. Sb.* **106** (1978), no. 4, 566–588. MR0507817 (80d:57019). English translation in *Math. U.S.S.R. Sb.* **35** (1979), 229–250. Zbl 0422.57005, doi:10.1070/SM1979v035n02ABEH001471.

- [24] TURAEV, V. G.; VIRO, O. YA. Intersection of loops in two-dimensional manifolds. II. Free loops. *Mat. Sb.* **121** (1983), no. 3, 359–369 (Russian). MR0708001 (85e:57007). English translation in *Math. USSR, Sb.* **49** (1984), 357–366. Zbl 0543.57007, doi: 10.1070/SM1984v049n02ABEH002714.

PATRICIA CAHN, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA, DAVID RITTENHOUSE LAB. 209 SOUTH 33RD STREET, PHILADELPHIA, PA 19104-6395, USA  
pcahn@math.upenn.edu

This paper is available via <http://nyjm.albany.edu/j/2013/19-14.html>.