Homotopy classification of ribbon tubes and welded string links

BENJAMIN AUDOUX, PAOLO BELLINGERI, JEAN-BAPTISTE MEILHAN
AND EMMANUEL WAGNER

Abstract. Ribbon 2-knotted objects are locally flat embeddings of surfaces in 4-space which bound immersed 3-manifolds with only ribbon singularities. They appear as topological realizations of the theory of welded knotted objects, which is a natural quotient of virtual knot theory. In this paper we consider ribbon tubes and ribbon torus-links, which are natural analogues of string links and links, respectively. We show how ribbon tubes naturally act on the reduced free group, and how this action classifies ribbon tubes up to link-homotopy, that is when allowing each component to cross itself. At the combinatorial level, this provides a classification of welded string links up to self-virtualization. This generalizes a result of Habegger and Lin on usual string links, and the above-mentioned action on the reduced free group can be refined to a general "virtual extension" of Milnor invariants. As an application, we obtain a classification of ribbon torus-links up to link-homotopy.

Mathematics Subject Classification (2010): 57Q45 (primary); 57M25, 57M27, 57Q35 (secondary).

Introduction

The theory of 2-knots, *i.e.* locally flat embeddings of the 2-sphere in 4-space, takes its origins in the mid-twenties from the work of Artin [1]. However, the systematic study of these objects only really began in the early sixties, notably through the work of Kervaire, Fox and Milnor [17,27,28], but also in a series of papers from the Kansai area, Japan, partially referenced below.[1] From this early stage, the class of ribbon 2-knots was given a particular attention. Roughly speaking, a 2-knot is ribbon if it bounds a locally flat immersed 3-ball whose singular set is a finite number of ribbon disks. Introduced by T. Yajima [46] under the name of simply knotted 2-spheres, they were extensively studied by T. Yanagawa in [48–50]. Ribbon 2-knots admit natural generalizations to ribbon 2-knotted objects, such as links and tangles.

[1] See Suzuki's comprehensive survey [43] for a much more complete bibliography on the subject.

One particularly nice feature of these objects is that they admit a diagrammatic representation which allows an explicit presentation of the associated knot groups, via a Wirtinger-type algorithm.

This diagrammatic representation also allows one to use ribbon 2-knotted objects as a (partial) topological realization of welded knot theory. Welded knots are a natural quotient of virtual knots, by the so-called Over Commute relation, which is one of the two forbidden moves in virtual knot theory. What makes this Over Commute relation natural is that the virtual knot group, and hence any virtual knot invariant derived from it, factors through it. These welded knotted objects first appeared in a work of Fenn-Rimanyi-Rourke in the more algebraic context of braids [14]. Although virtual knot theory can be realized topologically as the theory of knots in thickened surfaces modulo handle stabilization [10,33], this is no longer true for the welded quotient. However T. Yajima [46] showed that inflating classical diagrams defines a map, called the Tube map, which sends knots onto ribbon torus-knots, an analogue of ribbon 2-knots involving embedded tori. This map has been generalized by S. Satoh [41] to a surjective map from the welded diagrammatics. But this map fails to be one-to-one, and its kernel is not fully understood yet. Ribbon knotted objects in dimension 4 are therefore closely related to both 3-dimensional topology and welded theory.

In the present paper, we consider two classes of ribbon knotted objects with several components: *ribbon tubes* and *ribbon torus-links*. They are the natural analogues in the ribbon context of string links and links, respectively. As such, the latter ones can be obtained from the former by a natural braid-type closure operation.

We give in Theorem 2.36 the classification of ribbon torus-links up to link-homotopy, that is, up to homotopies in which different components remain disjoint. This theorem should be compared, on one hand, with the link-homotopy classification of links in 3-space by Habegger and Lin [20], and on the other hand, with the result of Bartels and Teichner proving the triviality of 2-links up to link-homotopy [7]. The dichotomy between these two results is striking, and Theorem 2.36 is closer in spirit to the one of Habegger and Lin; in particular, ribbon torus-links are far from being always link-homotopically trivial. Some examples are given in Section 2.5.

The problem of link-homotopy classification in higher dimension, intiated in [15,35], has developed in various directions, for instance with the construction of homotopy invariants (see, *e.g.* [29,30,34,35]), or towards the relationship with link concordance (*e.g.* [12]). Vanishing results of homotopy invariants for embeddings [12,35] have led, on one hand, to consider immersions, and on the other hand, to the result of Bartels and Teichner for spheres [7]. This also naturally connects to the general study of embedded surfaces in 4-space, which is a well-developed subject, see, *e.g.* [11]. Our result suggests that embedded tori in four space form a very particular but interesting case of study, which appears as another natural generalization of 1-dimensional links (as embedded 2-spheres do).

The proof of the above-mentioned Theorem 2.36 closely follows the work of Habegger and Lin [21]. As such, it naturally involves ribbon tubes and their classification up to link-homotopy, that we now state.

Theorem [Theorem 2.34]. *There is a group isomorphism between ribbon tubes up to link-homotopy and the group of basis-conjugating automorphisms of the reduced free group.*

This theorem calls for several comments, which we develop in the next three paragraphs.

Firstly, this theorem says that ribbon tubes up to link-homotopy form a group, in fact the quotient of the welded pure braid group by self-virtualization. Actually, our main results on ribbon tubes are obtained, via the Tube map, as consequences of similar statements for welded string links; see Theorems 3.10 and 3.11. In this context, the welded diagrammatics gives a faithful description of ribbon tubes, and our results can be thus seen as applications of virtual knot theory to the concrete study of topological objects.

Secondly, the classifying invariant underlying the theorem is a 4-dimensional version of Milnor numbers. Through the Tube map, we obtain a general and topologically grounded extension of Milnor invariants to virtual/welded objects, which should be compared with several previous virtual extensions of Milnor invariants proposed in [13,31,32].

Thirdly, since usual string links sit naturally in the welded string link monoid, the above theorem is a generalization of the link-homotopy classification of string links of Habegger and Lin [20]. In particular our 4-dimensional Milnor invariants have the natural feature that they coincide with the classical ones for usual string links through the Tube map. However we emphasize here that our proof of Theorem 2.34 is completely independent from the one of [20].

Our link-homotopy classification results for ribbon torus-links and ribbon tubes lead to two natural questions. The first one is that of the classification in higher dimension (for instance, for 3-dimensional ribbon tori in 5-space); the algebraic counterpart of Theorem 2.34 above suggests that the statement would remain true for codimension two tori in higher dimension. The second question addresses the general case of tori in four space, by removing the ribbon assumption; this is a natural question which seems to us worth studying.

The paper is organized as follows. We begin by setting some notation in Section 1. Section 2 is devoted to the topological aspects of this paper. We introduce ribbon tubes, broken surface diagrams and link-homotopy in our context, and provide various results on these notions. Section 3 focuses on the diagrammatic aspects of the paper; it addresses welded string links and pure braids and their connections with ribbon tubes and configuration spaces. The main tool for proving most of the results of the present paper is the theory of Gauss diagrams, which is reviewed in details in Section 4. The main proofs of the paper are given there. Finally, in Section 5, we define Milnor invariants for ribbon tubes, and show how they provide a natural and general extension of Milnor invariant to virtual knot theory.

1. General settings and notation

Unless otherwise specified, we set n to be a non negative integer, once for all. We begin with some topological notation and setting. Let I be the unit closed interval. Let $[\![1, n]\!]$ be the set of integer between 1 and n. We fix n distinct points $\{p_i\}_{i \in [\![1,n]\!]}$ in I. For every $i \in [\![1, n]\!]$, we choose a disk D_i in the interior of the 2–ball $B^2 = I \times I$ which contains the point p_i in its interior, seen in $\{\frac{1}{2}\} \times I$. We furthermore require that the disks D_i, for $i \in [\![1, n]\!]$, are pairwise disjoint. We denote by $C_i := \partial D_i$ the oriented boundary of D_i. We consider the 3–ball $B^3 := B^2 \times I$ and the 4–ball $B^4 = B^3 \times I$. For m a positive integer and for every submanifold $X \subset B^m \cong B^{m-1} \times I$, we set the notation:

- $\partial_0 X = X \cap (B^{m-1} \times \{0\})$, the "upper boundary" of X;
- $\partial_1 X = X \cap (B^{m-1} \times \{1\})$, the "lower boundary" of X;
- $\partial_* X = \partial X \setminus (\partial_0 X \sqcup \partial_1 X)$, the "lateral boundary" of X;
- $\overset{\star}{X} = X \setminus (\partial_* X \cup \partial(\partial_0 X) \cup \partial(\partial_1 X))$.

By a tubular neighborhood of X, we will mean an open set N such that $N \cap \mathring{B}^m$ is a tubular neighborhood of $\overset{\star}{X}$ in $\mathring{B}^m$ and $\partial_\varepsilon N$ is a tubular neighborhood of $\partial_\varepsilon X$ in $\partial_\varepsilon B^m$ for both $\varepsilon = 0$ and 1.

In the following, an immersion $Y \subset X$ shall be called *locally flat* if and only if it is locally homeomorphic to a linear subspace $\mathbb{R}^k$ in $\mathbb{R}^m$ for some positive integers $k \leq m$, except on ∂X and/or ∂Y, where one of the $\mathbb{R}$-summand should be replaced by $\mathbb{R}_+$. An intersection $Y_1 \cap Y_2 \subset X$ shall be called *flatly transverse* if and only if it is locally homeomorphic to the intersection of two linear subspaces $\mathbb{R}^{k_1}$ and $\mathbb{R}^{k_2}$ in $\mathbb{R}^m$ for some positive integers $k_1, k_2 \leq m$, except on ∂X, ∂Y_1 and/or ∂Y_2, where one of the $\mathbb{R}$–summand should be replaced by $\mathbb{R}_+$.

Throughout this paper, and for various types of objects, diagrammatical or topological, we will consider local moves. A local move is a transformation that changes the object only inside a ball of the appropriate dimension. By convention, we will represent only the ball where the move occurs, and the reader should keep in mind that there is a non represented part, which is identical for each side of the move.

We also define some algebraic notation which will be useful throughout the paper. Let G be a group and $a, b \in G$ two of its elements. We denote by:

- $a^b := b^{-1}ab$ the conjugate of a by b;
- $[a; b] := a^{-1}b^{-1}ab$ the commutator of a and b;

- $\Gamma_k G$, for $k \in \mathbb{N}^*$, the kth term of the lower central series of G inductively defined by $\Gamma_{k+1} G := [G; \Gamma_k G]$ and $\Gamma_1 G = G$;

and, if G is normally generated by elements $g_1, \dots, g_p$, we further denote by:

- $RG := G \big/ \{[g_i; g_i^g] \mid i \in [\![1, p]\!], g \in G\}$ the *reduced version of G*, which is the smallest quotient where each generator commutes with all its conjugates;
- $\text{Aut}_C(G) := \{f \in \text{Aut}(G) \mid \forall i \in [\![1, p]\!], \exists g \in G, f(g_i) = g_i^g\}$, the group of *conjugating automorphisms* of G.

Moreover, the free group on n generators is denoted by F_n. Unless otherwise specified, generators of F_n will be denoted by $x_1, \dots, x_n$. By abuse of notation, same notation will be kept for their images under quotients of F_n. More generally, names of maps and of elements will be kept unchanged when considering quotients.

2. Ribbon tubes

This section is devoted to the topological counterpart of the paper. We first define the considered objects, namely ribbon tubes, and then classify them up to link-homotopy, in terms of automorphisms of the reduced free group.

2.1. Ribbon tubes and their homology

In this section we identify $B^2 \hookrightarrow B^3$ with $B^2 \times \{\frac{1}{2}\}$, so that the disks $\{D_i\}_{i \in [\![1,n]\!]}$ are canonically seen in the interior of B^3.

2.1.1. *Definitions*

Definition 2.1. A *ribbon tube* is a locally flat embedding $T = \underset{i \in [\![1,n]\!]}{\sqcup} A_i$ of n disjoint copies of the oriented annulus $S^1 \times I$ in $\overset{\star}{B}^4$ such that:

- $\partial A_i = C_i \times \{0, 1\}$ for all $i \in [\![1, n]\!]$ and the orientation induced by A_i on ∂A_i coincides with that of C_i;
- there exist n locally flat immersed 3–balls $\underset{i \in [\![1,n]\!]}{\cup} B_i$ such that:

  - $\partial_* B_i = \mathring{A}_i$ for all $i \in [\![1, n]\!]$;
  - $\partial_\varepsilon B_i = D_i \times \{\varepsilon\}$ for all $i \in [\![1, n]\!]$ and $\varepsilon \in \{0, 1\}$;
  - the singular set of $\overset{n}{\underset{i=1}{\cup}} B_i$ is a disjoint union of so-called *ribbon singularities*, *i.e.* flatly transverse disks whose preimages are two disks, one in $\overset{n}{\underset{i=1}{\cup}} \mathring{B}_i$ and the other with interior in $\overset{n}{\underset{i=1}{\cup}} \mathring{B}_i$, and with boundary essentially embedded in $\overset{n}{\underset{i=1}{\sqcup}} \partial_* B_i = \overset{n}{\underset{i=1}{\sqcup}} \mathring{A}_i$.

We denote by rT_n the set of ribbon tubes up to isotopy fixing the boundary circles. It is naturally endowed with a monoidal structure by the stacking product $T \bullet T' := T \underset{\partial_1 T = \partial_0 T'}{\cup} T'$, and reparametrization, and with unit element the *trivial ribbon tube* $\mathbf{1}_n := \underset{i \in [\![1,n]\!]}{\sqcup} C_i \times I$.

Note that this notion of ribbon singularity is a 4–dimensional analogue of the classical notion of ribbon singularity introduced by R. Fox in [18]. Similar ribbon knotted objects were studied, for instance, in [47,48] and [25], and a survey can be found in [43].

Note that the orientation of the I–factor in $S^1 \times I$ induces an order on essential curves embedded in a tube component, since they are simultaneously homotopic to $S^1 \times \{t\}$ for some $t \in I$; we will refer to this order as the "co–orientation order". There is also a second, independent, surface orientation on each tube, which we will not use. The reader is referred to [5, Section 3.4] for a more detailed discussion.

Remark 2.2. There are two natural ways to close a ribbon tube $T \in \mathrm{rT}_n$ into a closed (ribbon) knotted surface in 4–space. First, by gluing the disks $\underset{i \in [\![1,n]\!]}{\sqcup} D_i \times \{0, 1\}$ which bound $\partial_0 T$ and $\partial_1 T$, and gluing a 4–ball along the boundary of B^4, one obtains an n-component ribbon 2–link [46], which we shall call the *disk-closure* of T. Second, by gluing a copy of the trivial ribbon tube $\mathbf{1}_n$ along T, identifying the pair $(B^3 \times \{0\}, \partial_0 T)$ with $(B^3 \times \{1\}, \partial_1 \mathbf{1}_n)$ and $(B^3 \times \{1\}, \partial_1 T)$ with $(B^3 \times \{0\}, \partial_0 \mathbf{1}_n)$, and taking a standard embedding of the resulting $S^3 \times S^1$ in S^4, one obtains an n-component ribbon torus–link [41], which we shall call the *tube-closure* of T. This is a higher dimensional analogue of the usual braid closure operation.

Let us mention here a particular portion of ribbon tube, called *wen* in the literature, that may appear in general, and which we shall encounter in the rest of this paper. Consider an oriented euclidian circle in three-space, and an additional dimension given by time. While time is running, let the circle make a half-turn; this path in 4–space is a wen. One can also think of a wen as an embedding in 4–space of a Klein bottle cut along a meridional circle. There are several topological types of wens, but it was shown in [25] that there are all isotopic in 4–space, so that we can speak of a wen unambiguously. Note that a wen is a surface embedded in 4–space with two boundary components, which do not have the same orientation as is required in the previous definition. Hence a wen is not an element of rT_1, but the square of a wen is. It turns out that the square of a wen is isotopic to the identity (see [25]), a fact which will be used in some proofs of this section. For a more detailed treatment of wens, see for instance [25], [4, Section 2.5.4] and [5, Section 4.5].

Definition 2.3. An element of rT_n is said to be *monotone* if it has a representative which is flatly transverse to the lamination $\underset{t \in I}{\cup} B^3 \times \{t\}$ of B^4.

We denote by rP_n the subset of rT_n whose elements are monotone.

Proposition 2.4. *The set rP_n is a group for the stacking product.*

Proof. The product of two monotone elements is obviously monotone, and an inverse T^{-1} for a monotone ribbon tube T is given by $T^{-1} \cap (B^3 \times \{t\}) := T \cap (B^3 \times \{1 - t\})$ for each $t \in I$. $\qquad \square$

Remark 2.5. It is worth noting here that two monotone ribbon tubes which are equivalent in rT_n are always related by a monotone isotopy, *i.e.* by an isotopy moving only through monotone objects. This is shown in Remark 2.26. As a consequence, rP_n is equal to the group defined as the quotient of monotone ribbon tubes by monotone isotopies.

Remark 2.6. The monotony condition enables I to be considered as a time parameter and the flat transversality forces $T \cap (B^3 \times \{t\})$ to be n disjoint circles for all $t \in I$. A monotone ribbon tube can hence be seen as an element of the fundamental group of the configuration space of n circles in 3–space, that we will denote by PR_n according to [8]. But, since the orientation of ∂T is prescribed by the one of $\bigsqcup_{i \in [\![1,n]\!]} C_i$, these elements are actually in the kernel of the map $PR_n \to \mathbb{Z}_2^n$ constructed in the proof of [8, Proposition 2.2]. The group rP_n is therefore isomorphic to the fundamental group of the configuration space of n circles in 3–space lying in parallel planes, denoted by PUR_n in [8]. This fact can be reinterpreted as seeing monotone ribbon tubes as motions of horizontal rings intermingled with wens. Such wens can be pushed above and since the starting and final orientations for a given circle match, there is an even number of wens and they cancel pairwise.

2.1.2. *Homology groups*

Let T be a ribbon tube with tube components $\bigsqcup_{i \in [\![1,n]\!]} A_i$.

Since T is locally flat in B^4, there is a unique way, up to isotopy, to consider, for all $i \in [\![1, n]\!]$, disjoint tubular neighborhoods $N(A_i) \cong D^2 \times S^1 \times I$ for A_i, with $A_i = \{0\} \times S^1 \times I \subset N(A_i)$. We denote by $N(T) := \bigsqcup_{i \in [\![1,n]\!]} N(A_i)$ a reunion of such tubular neighborhoods and by $W = B^4 \setminus \overset{\star}{N}(T)$ the complement of its interior in B^4.

As a direct application of the Mayer–Vietoris exact sequence, we obtain:

Proposition 2.7. *The homology groups of W are $H_0(W) = \mathbb{Z}$,*

$$H_1(W) = \mathbb{Z}^n = \mathbb{Z}\langle c_i \mid i \in [\![1, n]\!]\rangle, \ H_2(W) = \mathbb{Z}^n = \mathbb{Z}\langle \tau_i \mid i \in [\![1, n]\!]\rangle, \ H_3(W) = \mathbb{Z}$$

and $H_k(W) = 0$ for $k \geq 4$, where c_i is the homology class in $H_1(W)$ of $\partial D^2 \times \{s\} \times \{t\} \subset \partial N(A_i)$ for any $(s, t) \in S^1 \times I$, and τ_i of T is the homology class in $H_2(W)$ of $\partial D^2 \times S^1 \times \{t\} \subset \partial N(A_i)$ for any $t \in I$.

2.2. Broken surface diagrams

Links in 3–space can be described using diagrams, which are their generic projection onto a 2–dimensional plane with extra decoration encoding the 3–dimensional

information. Similarly, it turns out that ribbon knotted objects, which are surfaces in 4–space, can be described using their generic projection onto a 3–space; this leads to the following notion of broken surface diagram.

Definition 2.8. A *broken surface diagram* is a locally flat immersion S of n oriented annuli $\bigsqcup\limits_{i \in [\![1,n]\!]} A_i$ in $\overset{\star}{B}^3$ such that:

- $\partial A_i = C_i \times \{0, 1\}$ for all $i \in [\![1, n]\!]$ and the orientation induced by A_i on ∂A_i coincides with that of C_i;
- the set $\Sigma(S)$ of connected components of singular points in S consists of flatly transverse circles in $\bigcup\limits_{i=1}^{n} \mathring{A}_i$, called *singular circles*.

Remark 2.9. Since we are only interested in broken surface diagrams which represent ribbon tubes, the definition given in this paper is watered down compared to what is commonly used in the literature. Let us recall that in general, a generic projection of an embedded surface in 4–space onto a 3–space has three types of singularities: double points, triple points and branching points (see, for instance, [11,48]). In our context, only the first type does occur.

Moreover, for each element of $\Sigma(S)$, a local ordering is given on the two circle preimages. By convention, this ordering is specified on pictures by erasing a small neighborhood in $\bigcup\limits_{i=1}^{n} \mathring{A}_i$ of the lowest preimage (see Figure 2.1). Note that this is the same convention which is used for usual knot diagrams.
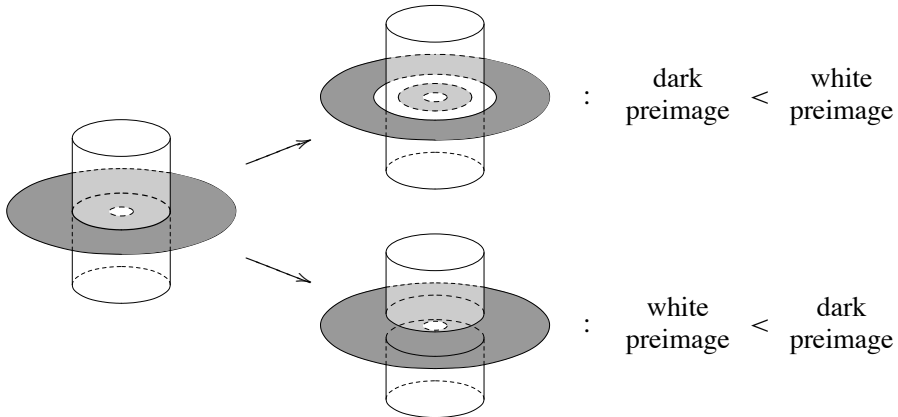

Figure 2.1. Local pictures for a singular circle in a broken surface diagram.

We shall actually consider only a special case of broken surface diagrams, namely symmetric ones, that we now introduce, following essentially Yajima [47].

Definition 2.10. A *ribbon band* is a locally flat immersion $F := \underset{i \in [\![1,n]\!]}{\sqcup} F_i$ of n disks $I \times I$ in $\overset{\star}{B}^3$ such that

- $\partial_\varepsilon F_i = I \times \{\varepsilon\}$ is a diameter for $D_i \subset B^2 \times \{\varepsilon\}$ for each $\varepsilon \in \{0, 1\}$ and each $i \in [\![1, n]\!]$;
- the singular set of $\overset{n}{\underset{i=1}{\cup}} F_i$ is a disjoint union of 1–dimensional ribbon singularities, *i.e.* flatly transverse intervals whose preimages are two intervals, one in $\overset{n}{\underset{i=1}{\cup}} \overset{\circ}{F}_i$ and the other with interior in $\overset{n}{\underset{i=1}{\cup}} \overset{\circ}{F}_i$ and with boundary intersecting distinct connected components of $\partial_* F$.
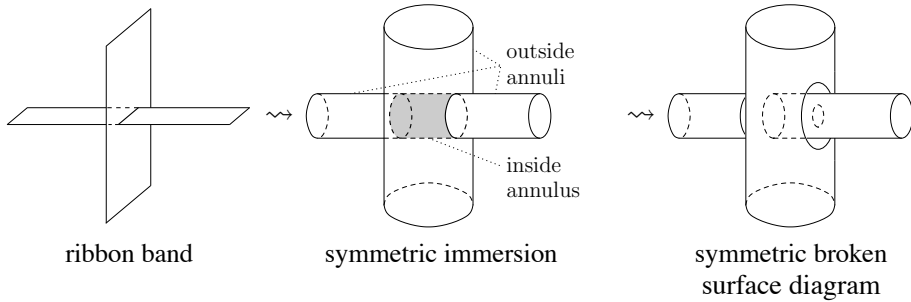

Figure 2.2. From ribbon bands to symmetric broken surface diagrams.

A *symmetric immersion* is a locally flat immersion of n oriented annuli obtained by taking the boundary of a thickening of a ribbon band.

See the left-hand side and middle of Figure 2.2, respectively, for an example of ribbon band and symmetric immersion.

Remark 2.11. A symmetric immersion is naturally endowed with a notion of "inside", which is simply the thickening of the corresponding ribbon band.

As pictured in Figure 2.2, singularities in a symmetric immersion H are singular circles with exactly one essential preimage in H. These essential preimages cut H into smaller annuli. Moreover, the singular circles come in pairs, such that the associated essential preimages delimit an annulus which is entirely contained inside H. Such annuli are called *inside annuli*, and those which are not are called *outside annuli*. See Figure 2.2. Note that every non essential preimage of a singularity belongs to an outside annulus.

Definition 2.12. A broken surface diagram S is said to be *symmetric* if it is a symmetric immersion such that, for each inside annulus, the local orderings between the essential and non essential preimages of its two boundary circles are different. See the right-hand side of Figure 2.2 for an illustration.

Let T be a ribbon tube, and suppose that there is a projection $B^4 \to B^3$ such that the image of T in B^3 has singular locus a union of flatly transverse circles. For each double point, the preimages are naturally ordered by their positions on the projection rays, as in Figure 2.1, so that the projection naturally yields a broken surface diagram. This suggests that broken surface diagrams can be thought of as 3–dimensional representations of ribbon tubes. This is indeed the case, as stated in the next result, which is essentially due to Yanagawa.

Lemma 2.13 ([48]). *Any ribbon tube admits, up to isotopy, a projection to B^3 which is a broken surface diagram. Conversely any broken surface diagram is the projection of a unique ribbon tube.*

Proof. Yanagawa proves that for locally flat embeddings of 2 spheres in $\mathbb{R}^4$, there is an equivalence between the property of being ribbon (property $R(4)$ in [48]) and the property of admitting a projection onto an immersion whose only singular points are transverse double points (property $R(3)$ in [48]). The proof of the equivalence passes through the equivalence with a third notion, which is the fusion of a trivial 2–link (property F in [48]). [48, Lemma 4.3] proves that $R(4)$ and F are equivalent. [48, Corollary 3.3] states that F implies $R(3)$ and Lemma 3.4 states the converse implication. All the arguments are local and apply to the case of ribbon tubes as well. $\square$

More specifically, we have the following:

Lemma 2.14 ([25,47]). *Any ribbon tube can be represented by a symmetric broken surface diagram.*

Proof. By the previous lemma, any ribbon tube T can be represented by a broken surface diagram S. Now we prove that S can be transformed into a symmetric broken surface diagram which still represents the same ribbon tube. For this, we consider the disk–closure S of T as defined in Remark 2.2. We obtain a ribbon diagram for an n–component ribbon 2-link. Now by [25, Theorem 5.2], this diagram is equivalent to a diagram obtained by closing a symmetric broken surface diagram. This equivalence is generated by local moves, which we may assume to avoid a neighborhood of the closing disks *except* possibly for a finite number of moves which consist in discarding some wens across closure disks; we do not perform these latter moves, and leave such wens near the boundary circles instead. We obtain in this way a new broken surface diagram which describes a ribbon tube isotopic to T, and it only remains to get rid of all residual wens. Following again [25], all wens can be pushed down near the bottom circles $\bigcup_{i=1}^{n} C_i \times \{0\}$. The orientations of the two boundary circles of a wen are opposite, and the orientation on ∂T induced by T must agree with that of $\bigsqcup_{i \in [\![1,n]\!]} C_i$, so there are an even number of wens near each bottom circle, and they cancel pairwise. There is thus no more wen and the resulting broken surface diagram for T is symmetric. $\square$

Remark 2.15. If two symmetric broken surface diagrams differ by one of the "Reidemeister" (RI, RII, RIII) or "virtual" (V) moves shown in Figure 2.3, then the associated ribbon tubes are isotopic.

Indeed, both sides of virtual and Reidemeister RII & RIII moves can be locally modelized as occurring in the projection of monotone ribbon tubes. Then, one can use the approach mentioned in Remark 2.6 to prove the statement. Reidemeister move RI is more tricky but can, for example, be checked using Roseman moves defined in [40]. Another approach can also be found in [2]. Nevertheless, it is still unknown whether the correspondence between ribbon tubes and symmetric broken surface diagrams up to Reidemeister and virtual moves is one-to-one; see, *e.g.* [4, Section 3.1.1] or [5, Section 3.1]. This discussion is the key to the relation between ribbon tubes and welded string links defined in Section 3. However, it follows from Corollary 4.34 that this correspondence is one-to-one up to the link-homotopy relation which is defined in the next section.
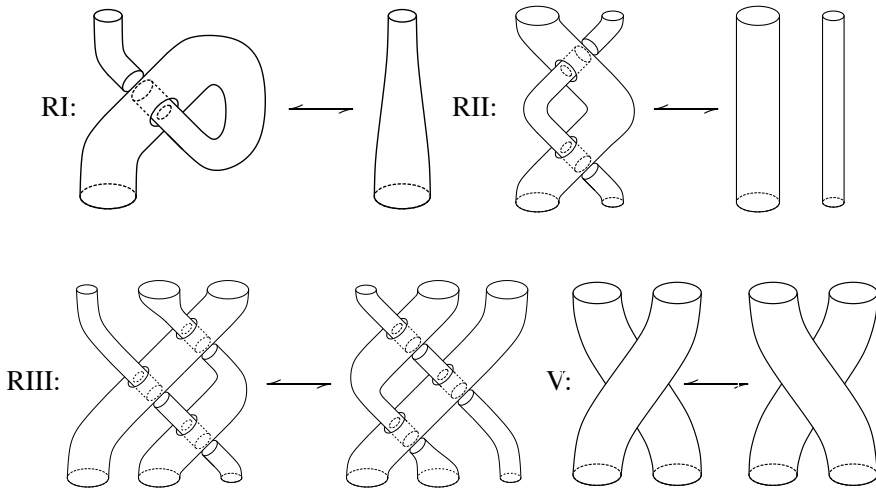

Figure 2.3. Reidemeister and virtual moves for broken diagrams.

2.2.1. *Fundamental group*

Let T be a ribbon tube with tube components $\underset{i \in [\![1,n]\!]}{\bigsqcup} A_i$ and consider $N(T) :=$

$\underset{i \in [\![1,n]\!]}{\bigsqcup} N(A_i)$ and $W = B^4 \setminus \overset{\star}{N}(T)$ as in Section 2.1.2. We also consider a global parametrization (x, y, z, t) of B^4, which is compatible with $B^4 \cong B^3 \times I \cong B^2 \times I \times I$ near $\partial_0 B^4$ and $\partial_1 B^4$, and such that the projection along z maps T onto a symmetric broken surface diagram S. We also fix a base point $e := (x_0, y_0, z_0, t_0)$ with z_0 greater than the highest z–value taken on $N(T)$.

Notation 2.16. We set $\pi_1(T) := \pi_1(W)$ with base point e.

2.2.2. *Meridians and longitudes*

For every point $a := (x_a, y_a, z_a, t_a) \in T$, we define $m_a \in \pi_1(T)$, the *meridian around* a, as $\tau_a^{-1} \gamma_a \tau_a$ where

- τ_a is the straight path from e to $\widetilde{a} := (x_a, y_a, z_0, t_a)$;
- γ_a is the loop in W, unique up to isotopy, based at $\widetilde{a}$ and which enlaces positively T around a.

Definition 2.17. For each $i \in [\![1, n]\!]$ and $\varepsilon \in \{0, 1\}$, we denote by m_i^ε the meridian in $\pi_1(T)$ defined as $m_{a_i^\varepsilon}$ for any $a_i^\varepsilon \in C_i \times \{\varepsilon\}$. If $\varepsilon = 0$, we call it a *bottom meridian*, and if $\varepsilon = 1$, we call it a *top meridian*.

See the left-hand side of Figure 2.4 for an example of a bottom meridian. Note that, for any $\varepsilon \in \{0, 1\}$ and any choice of a_i^ε, the fundamental group of $\partial_\varepsilon W$ based at $(x_0, y_0, z_0, \varepsilon)$ can be identified with the free group $F_n = \langle m_i^\varepsilon \mid i \in [\![1, n]\!] \rangle$.
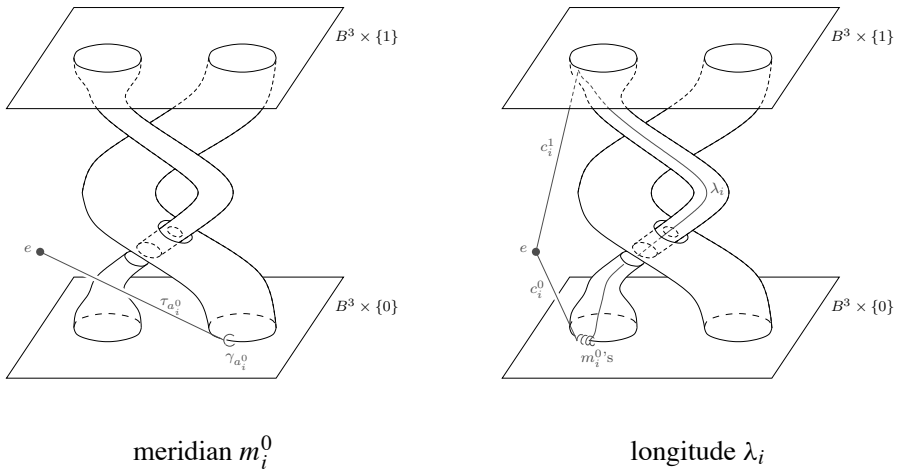

meridian m_i^0 longitude λ_i

Figure 2.4. Examples of meridians and longitude

Now, we define the notion of longitude for T as follows. First, we fix two points $e_i^0 \in \partial_0 N(A_i)$ and $e_i^1 \in \partial_1 N(A_i)$ on each extremity of the boundary of the tubular neighborhood of A_i. A longitude for A_i is defined as the isotopy class of an arc on $\partial N(A_i)$ running from e_i^0 to e_i^1. Since $N(A_i)$ is homeomorphic to $D^2 \times S^1 \times I$, we note that

$$\partial N(A_i) = \left(S^1 \times S^1 \times I\right) \cup \left(D^2 \times S^1 \times \{0\}\right) \cup \left(D^2 \times S^1 \times \{1\}\right),$$

so that the choice of a longitude for A_i is *a priori* specified by two coordinates, one for each of the two S^1–factors in $S^1 \times S^1 \times I$. On one hand, the first S^1–factor

is generated by the meridian m_i^0, so that the first coordinate is given by the linking number with the tube component A_i. It can be easily checked, on the other hand, that two choices of longitude for A_i which only differ by their coordinate in the second S^1–factor are actually isotopic in W.

Definition 2.18. For each $i \in [\![1, n]\!]$, an ith *longitude of* T is the isotopy class of an arc on $\partial N(A_i)$, running from e_i^0 to e_i^1, and closed into a loop with an arc $c_i^0 \cup c_i^1$ defined as follows. For $\varepsilon \in \{0, 1\}$, we denote by $\widetilde{e}_i^\varepsilon$ the point above e_i^ε with z–coordinate z_0; then c_i^ε is the broken line between $e, \widetilde{e}_i^\varepsilon$ and e_i^ε.

This definition is illustrated on the right-hand side of Figure 2.4.

Remark 2.19. As explained above, any two choices of an ith longitude differ by a power of m_i^0, which is detected by the linking number with the tube component A_i. In particular, there is thus a *preferred i^{th} longitude*, which is defined as having linking number zero with A_i. We shall not make use of this fact here, but in Section 5 at the end of this paper.

2.2.3. *Wirtinger presentation*

In the following, we give a presentation for $\pi_1(T)$ in terms of broken surface diagrams. Let S be a symmetric broken surface diagram representing T. According to the notation set in Remark 2.11, we denote by $\mathrm{Out}(S)$ the set of outside annuli of S and by $\mathrm{In}(S)$ the set of inside annuli.

For each inside annulus $\beta \in \mathrm{In}(S)$, we define

- $\alpha_\beta^0 \in \mathrm{Out}(S)$, the outside annulus which contains $\partial\beta$;
- C_β^+, C_β^-, the connected components of $\beta \cap \alpha_\beta^0$ which, respectively, are closer to $\partial_1 S$ and to $\partial_0 S$, according to the co–orientation order (defined after Definition 2.1);
- $\alpha_\beta^+, \alpha_\beta^- \in \mathrm{Out}(S)$, the outside annuli which have C_β^+ and C_β^- as boundary components, respectively;
- $\varepsilon_\beta = 1$ if, according to the local ordering, the preimage of C_β^- in β is higher than the preimage in α_β^0, and $\varepsilon_\beta = -1$ otherwise.

See Figure 2.5 for an illustration.

Proposition 2.20 ([46,48]). *Let T be a ribbon tube and S any broken surface representing T, then*

$$\pi_1(T) \cong \left\langle \mathrm{Out}(S) \mid \alpha_\beta^+ = (\alpha_\beta^-)^{(\alpha_\beta^0)^{\varepsilon_\beta}} \text{ for all } \beta \in \mathrm{In}(S)\right\rangle.$$

In this isomorphism, $\alpha \in \mathrm{Out}(S)$ is sent to m_a, where a is any point on α close to $\partial\alpha$.

Proof. By considering, in B^4, the union of the projection rays from $T \subset B^4$ to $S \subset B^3 = \partial_0 B^4$, the result follows from standard techniques (see, *e.g.*, the proof of [9, Theorem 3.4]). $\qquad\square$
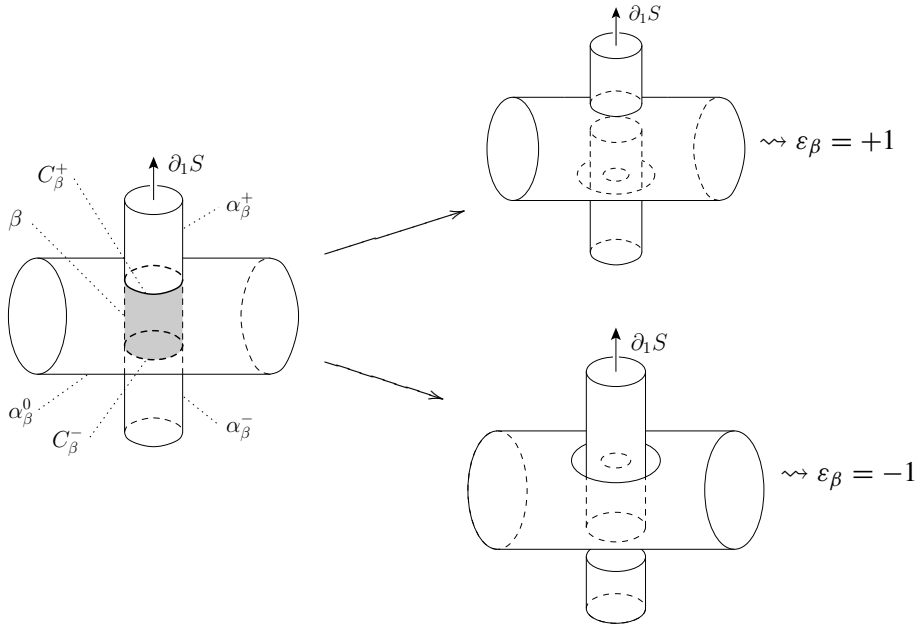
Figure 2.5. Signs associated to inside annuli.

Example 2.21. Let H be the ribbon tube represented in Figure 2.4. We have

$$\pi_1(H) \cong \langle m_1^0, m_1^1, m_2^0 \mid m_1^1 = (m_1^0)^{(m_2^0)^{-1}} \rangle.$$

Corollary 2.22. *The group $\pi_1(T)$ is generated by elements $\{m_a\}_{a \in T}$, and moreover, if $a \in A_i$ for some $i \in [\![1, n]\!]$, then m_a is a conjugate of m_i^ε for both $\varepsilon = 0$ or 1.*

2.2.4. *Reduced fundamental group*

In this section, we define and describe a reduced notion of fundamental group for T. Indeed, Corollary 2.22 states that $\pi_1(T)$ is normally generated by meridians $m_1^\varepsilon, \cdots, m_n^\varepsilon$ for either $\varepsilon = 0$ or 1. Moreover, since top meridians are also conjugates of the bottom meridians and *vice versa*, we can define the following without ambiguity:

Definition 2.23. The *reduced fundamental group of T* is defined as $R\pi_1(T)$, the reduced version of $\pi_1(T)$ seen as normally generated by either bottom meridians or top meridians. For convenience, we also denote $R\pi_1(\partial_\varepsilon W)$ by $R\pi_1(\partial_\varepsilon T)$, for $\varepsilon \in \{0, 1\}$.

It is a consequence of the description of $H_*(W)$ given in Proposition 2.7 that, for $\varepsilon \in \{0, 1\}$, the inclusion $\iota_\varepsilon : \partial_\varepsilon W \hookrightarrow W$ induces isomorphisms at the H_1 and

H_2 levels. Stallings theorem, *i.e.* Theorem 5.1 in [42], then implies that

$$(\iota_\varepsilon)_k : \pi_1(\partial_\varepsilon W)\big/\Gamma_k\pi_1(\partial_\varepsilon W) \xrightarrow{\ \simeq\ } \pi_1(T)\big/\Gamma_k\pi_1(T)$$

are isomorphisms for every $k \in \mathbb{N}^*$. But $\pi_1(\partial_\varepsilon W)$ is the free group F_n generated by meridians $m_1^\varepsilon, \cdots, m_n^\varepsilon$. It follows from Habegger-Lin's Lemma 1.3 in [20] that for $k \geq n$, $R\left(F_n\big/\Gamma_k F_n\right) \cong RF_n$. As a consequence:

Proposition 2.24. *The inclusions ι_0 and ι_1 induce isomorphisms*

$$RF_n \cong R\pi_1(\partial_0 T) \xrightarrow[\iota_0^*]{\ \simeq\ } R\pi_1(T) \xleftarrow[\iota_1^*]{\ \simeq\ } R\pi_1(\partial_1 T) \cong RF_n.$$

Using the isomorphisms of Proposition 2.24, we define, for every ribbon tube T, a map $\varphi_T : RF_n \to RF_n$ by $\varphi_T := \iota_0^{*-1} \circ \iota_1^*$. This can be seen as reading the top meridians as products of the bottom ones. It is straightforwardly checked that $\varphi_{T \bullet T'} = \varphi_T \circ \varphi_{T'}$ and this action on RF_n is obviously invariant under isotopies of ribbon tubes.

It follows from Corollary 2.22 that:

Proposition 2.25. *For every ribbon tube T, φ_T is an element of $\mathrm{Aut}_C(RF_n)$, the group of conjugating automorphisms. More precisely, the action of $T \in rT_n$ on RF_n is given by conjugation of each x_i, for $i \in [\![1, n]\!]$, by the image through ι_0^* of an ith longitude of T.*

Note that, in the reduced free group, this conjugation does not depend on the choice of an ith longitude.

Remark 2.26. If the ribbon tube T is monotone, then the inclusions ι_ε, for $\varepsilon \in \{0, 1\}$, actually induce isomorphisms $\pi_1(\partial_\varepsilon W) = F_n \cong \pi_1(T)$, so that T defines an action in $\mathrm{Aut}_C(F_n)$ which is left invariant by any (possibly non-monotone) isotopy. Theorem 2.6 of [4] shows that this induces an isomorphism between the group of monotone ribbon tubes up to monotone isotopies and $\mathrm{Aut}_C(F_n)$, since this group is isomorphic to the group PUR_n defined in [8] (as noticed in Remark 2.6). It follows that if two monotone ribbon tubes are equivalent in rT_n, then they induce the same action on F_n and hence are isotopic through a monotone isotopy. Put differently, we have that the group PUR_n injects in rT_n.

2.3. Classification of ribbon tubes up to link-homotopy

We now give a classification result for ribbon tubes, up to link-homotopy, in terms of their induced action on the reduced free group.

2.3.1. *Link-homotopy*

Definition 2.27. A *singular ribbon tube* is a locally flat immersion T of n annuli $\underset{i \in [\![1,n]\!]}{\bigsqcup} A_i$ in $\overset{\star}{B}{}^4$ such that

- $\partial A_i = C_i \times \{0, 1\}$ for all $i \in [\![1, n]\!]$ and the orientation induced by A_i on ∂A_i coincides with that of C_i;
- the singular set of T is a single flatly transverse circle, called *singular loop*, whose preimages are two circles embedded in $\overset{n}{\underset{i=1}{\cup}} \mathring{A}_i$, an essential and a non essential one.
- there exist n locally flat immersed 3–balls $\underset{i \in [\![1,n]\!]}{\cup} B_i$ such that

  - $\partial_* B_i = \mathring{A}_i$ and $\partial_\varepsilon B_i = D_i \times \{\varepsilon\}$ for all $i \in [\![1, n]\!]$ and $\varepsilon \in \{0, 1\}$;
  - the singular set of $\overset{n}{\underset{i=1}{\cup}} B_i$ is a disjoint union of flatly transverse disks, all of them being ribbon singularities but one, whose preimages are two disks bounded by the preimages of the singular loop, one in $\overset{n}{\underset{i=1}{\cup}} \partial_* B_i$ and the other with interior in $\overset{n}{\underset{i=1}{\cup}} \mathring{B}_i$.

We say that a singular ribbon tube is *self-singular* if and only if both preimages of the singular loop belong to the same tube component.

Definition 2.28. Two ribbon tubes T_1 and T_2 are said to be *link-homotopic* if and only if there is a 1–parameter family of regular and self-singular ribbon tubes from T_1 to T_2 passing through a finite number of self-singular ribbon tubes.

We denote by $\mathrm{rT}_n^{\mathrm{h}}$ the quotient of rT_n by the link-homotopy equivalence, which is compatible with the monoidal structure of rT_n. Furthermore, we denote by $\mathrm{rP}_n^{\mathrm{h}}$ the image of rP_n in $\mathrm{rT}_n^{\mathrm{h}}$.

Proposition 2.29 ([25]). *The link-homotopy equivalence is generated by* self-circle crossing changes, *which are the operations in B^4 induced by the local move shown in Figure 2.6, which switches the local ordering on the preimages of a given singular circle, where it is required that both preimages are on the same tube component.*
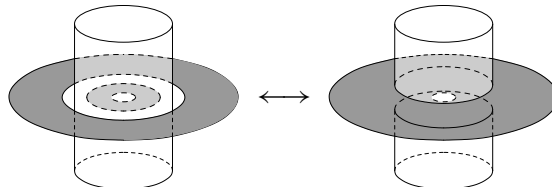

Figure 2.6. A circle crossing change at the level of broken surface diagrams.

Note that a circle crossing change can be seen as a local move among symmetric broken surface diagrams. Indeed, although applying a circle crossing change yields

a surface diagram which is no longer symmetric (see the middle of Figure 2.7), the resulting inside annulus correspond to a piece of tube passing entirely above or below another piece of tube. There is thus no obstruction in B^4 for pushing these two pieces of tube apart, so that their projections don't meet anymore (see the right-hand side of Figure 2.7).
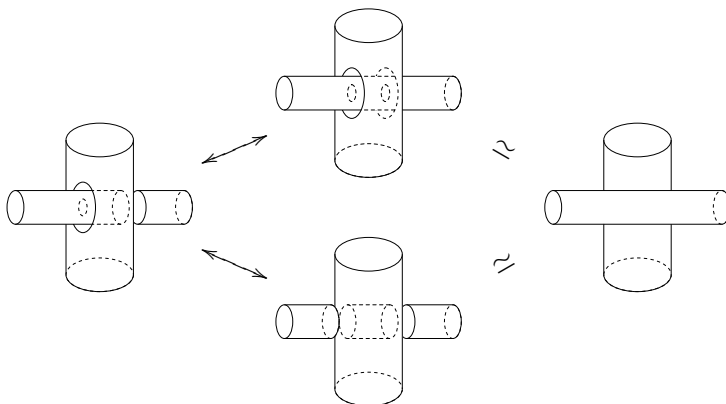

Figure 2.7. A circle crossing change at the symmetric broken surface diagram level.

We now state one of the main results of this paper.

Theorem 2.30. *Every ribbon tube is link-homotopic to a monotone ribbon tube.*

Proof. Using the surjectivity of the Tube map defined in Section 4.5, this is a direct consequence of Theorem 4.12. □

Remark 2.31. Theorem 2.30 can be regarded as a higher-dimensional analogue of Habegger-Lin's result, stating that any string link is link-homotopic to a pure braid [20].

Corollary 2.32. *The set* $\mathrm{rT}_n^{\mathrm{h}}$ *is a group for the stacking product.*

2.3.2. *Actions on the reduced free group and link-homotopy*

In Section 2.2.4, a conjugating automorphism φ_T was associated to any ribbon tube T. It turns out that this automorphism φ_T is invariant under link-homotopy.

Proposition 2.33. *If T_0 and T_1 are two link-homotopic ribbon tubes, then $\varphi_{T_0} = \varphi_{T_1}$.*

Proof. This is a consequence of Lemma 4.20 and the surjectivity of the Tube map defined in Section 4.5. However, this result can be given a more topological proof that we will sketch here.

It is sufficient to prove the proposition in the case of a link-homotopy H passing through a unique singular ribbon tube. We denote then by A the annulus which

contains the unique singular loop and by δ_0 the disk in A which is bounded by this singular loop. By abuse of notation, we will also denote by A the annulus in T_0, in T_1 and in any ribbon tube in-between, which are the deformations of A by H. Moreover, we can assume that the resolution of δ_0 in T_0 creates a ribbon singularity, as on the left-hand side of Figure 7, whereas it doesn't in T_1, as on the right-hand side of the figure.

Following the proof of [20, Lemma 1.6],[2] we consider the complement W_H of a tubular neighborhood of H in $B^5 \cong B^4 \times I$, where I represents the 1–parameter of the link-homotopy. The inclusions $\iota_0 : \partial_0 W_H \hookrightarrow W_H$ and $\iota_1 : \partial_1 W_H \hookrightarrow W_H$ induce the following commutative diagram:

$$
\begin{array}{ccccc}
R\pi_1(\partial_0 T_0) & \xrightarrow{\;\simeq\;} & R\pi_1(\partial_0 W) = R\pi_1(T_0) & \xleftarrow{\;\simeq\;} & R\pi_1(\partial_1 T_0) \\
\| & & \big\downarrow & & \| \\
\mathrm{RF}_n & \xrightarrow{\qquad} & R\pi(H) & \xleftarrow{\qquad} & \mathrm{RF}_n \\
\| & & \big\uparrow & & \| \\
R\pi_1(\partial_0 T_1) & \xrightarrow{\;\simeq\;} & R\pi_1(\partial_1 W) = R\pi_1(T_1) & \xleftarrow{\;\simeq\;} & R\pi_1(\partial_1 T_1),
\end{array}
\tag{2.1}
$$

where $R\pi_1(H) := \pi_1(W_H)/\Omega$ with Ω the normal subgroup generated by all commutators $[m; m^g]$, m being any bottom meridian of T_0 and g any element of $\pi_1(W_H)$. Note that, since top and bottom meridians of T_0 and T_1 are equal in $\pi_1(W_H)$, and since top meridians are conjugates of the bottom ones, both vertical maps to $R\pi(H)$ in (2.1) are well defined.

Let $B_* \subset \mathring{B}^4$ be a 4–ball so that:

- $\overline{B}_* := B_* \times I \subset B^5$ contains δ_0;
- H is trivial outside $\overline{B}_*$;
- $\partial \overline{B}_* \cap H$ is the disjoint union of 4 thickened circles $C_1 \times I$, $C_2 \times I$, $C_3 \times I$ and $C_4 \times I$ where C_1, C_2, C_3 and C_4 are four essential curves in A, numbered according to the co–orientation order.

Up to symmetry, we can assume that δ_0 is totally embedded in α_1, the annulus in A which is cobounded by C_1 and C_2, rather than in α_2, the annulus in A which is cobounded by C_3 and C_4.

Now, we denote by $Z := W_H \setminus \overline{B}_*$ the complement of H outside B_* and by m_i, for $i \in \{1, 2, 3, 4\}$, the meridian in $\pi_1(W_H)$ which enlaces positively $C_i \times I$.

We claim that:

(1) $\pi_1(\partial_0 W_H) = \pi_1(Z)/\{m_2 = m_1; m_4 = m_3^{m_1^{\pm 1}}\}$;

(2) $\pi_1(\partial_1 W_H) = \pi_1(Z)/\{m_2 = m_1; m_4 = m_3\}$;

(3) $\pi_1(W_H) = \pi_1(Z)/\{m_2 = m_1; m_4 = m_3; m_1 m_3 = m_3 m_4\}$.

---

[2] As it was already the case in [20], the theorem of Stallings cannot be used here since $H_2(W_H)$ has an extra summand. The retraction argument of Habegger and Lin could not either be used in our context.

Using the Seifert–Van Kampen theorem and the Wirtinger presentation for the ribbon tube $\partial_0 H \cap B_*$ and $\partial_1 H \cap B_*$, the first two assertions are rather clear since filling Z with $(\partial_0 H \cap B_*) \times I$ or $(\partial_1 H \cap B_*) \times I$ gives a thickening of, respectively, $\partial_0 W_H$ and $\partial_1 W_H$.

Now we focus our attention on $W_H \cap \overline{B}_*$. Using standard techniques, borrowed, *e.g.*, from the proof of [9, Theorem 3.4], it is easily seen that $\pi_1(W_H \cap \overline{B}_*)$ is generated by m_1 and m_2 and that $m_2 = m_1$ and $m_4 = m_3$. We consider a point of $\partial \delta_0$ and a 4–ball b around it which is transverse to δ_0. The ball b intersects α_1 and α_2 along two transverse disks and b can be seen as the product of these two disks. This product decomposition of b provides a Heegaard splitting of $\partial b = S^3$ into two solid tori whose cores are $\partial b \cap \alpha_1$ and $\partial b \cap \alpha_2$, and whose meridians are m_1 and m_3. But these cores enlace as an Hopf link and since $\pi_1(\text{Hopf link}) \cong \mathbb{Z}^2$, m_1 and m_3 commute. Finally, using the Mayer–Vietoris exact sequence for $W_H \cap \overline{B}_*$ seen as $(\overline{B}_* \setminus N_1) \cap (\overline{B}_* \setminus N_2)$ where N_1 and N_2 are tubular neighborhoods of, respectively, $\alpha_1 \times I$ and $\alpha_2 \times I$, one can compute that the abelianization $H_1(W_H \cap \overline{B}_*)$ of $\pi_1(W_H \cap \overline{B}_*)$ is $\mathbb{Z}^2$, so no other relation can hold in $\pi_1(W_H \cap \overline{B}_*)$. By use of the Seifert–Van Kampen theorem, this proves statement (3).

Now, since m_1, m_2, m_3 and m_4 are all conjugates of the top meridian of A, they commute in the reduced group, and it follows that $R\pi_1(\partial_0 W_H) \cong R\pi_1(\partial_1 W_H) \cong R\pi_1(W_H) \cong R\left(\pi_1(Z)/\{m_2 = m_1; m_4 = m_3\}\right)$. All the maps in the diagram (2.1) are hence isomorphisms and since φ_{T_0} and φ_{T_1} are, respectively, the upper and lower lines of (2.1), they are equal. $\qquad\square$

We can now give the main result of the paper:

Theorem 2.34. *The map $\varphi : \mathrm{rT}_n^{\mathrm{h}} \longrightarrow \mathrm{Aut}_C(\mathrm{RF}_n)$, sending T to φ_T is an isomorphism.*

As pointed out in Section 4.5, this is a consequence of Theorem 4.17. Some examples are given in Section 2.5 below.

2.4. Classification of ribbon torus-links up to link-homotopy

In [21] a structure theorem was given for certain "concordance-type" equivalence relations on links in 3-space. Here we give an analogous structure theorem in the higher dimensional case. Actually, we follow the reformulation given in [23], which was in fact implicit in the proof of [21].

We consider n-component *ribbon torus-links*, that is, locally flat embeddings of n disjoint tori in S^4 which bound locally flat immersed solid tori whose singular set is a finite number of ribbon disks. Denote by rL_n the set of n-component ribbon torus-links up to isotopy. The tube-closure operation defined in Remark 2.2 induces a natural closure map $\hat{} : \mathrm{rT}_n \to \mathrm{rL}_n$, which is easily seen to be surjective. Indeed, given an n-component ribbon torus-link, it is always possible up to isotopy to find a 3-ball intersecting the n components transversally exactly once, along an essential

circle, so that cutting the ribbon torus-link along this ball provides a preimage for the closure operation. We shall refer to such a ball as a "base-ball".

Consider an equivalence relation E on the union, for all $n \in \mathbb{N}^*$, of the sets rT_n and rL_n. We will denote by $E(x)$ the E-equivalence class of a ribbon tube or torus-link x, and we also denote by E the map which sends a ribbon knotted object to its equivalence class. We denote respectively by $E\mathrm{rT}_n$ and $E\mathrm{rL}_n$ the set of E-equivalence classes of ribbon tubes and ribbon torus-links.

The Habegger-Lin Classification Scheme relies on the following set of axioms:

(1) The equivalence relation E is *local*, *i.e.* for all $L_1, L_2 \in \mathrm{rT}_n$ such that $E(L_1) = E(L_2)$, and for all $T_1, T_2 \in \mathrm{rT}_{2n}$ such that $E(T_1) = E(T_2)$, we have:

  (i) $E(\hat{L}_1) = E(\hat{L}_2)$;

  (ii) $E(\mathbf{1}_n \otimes L_1) = E(\mathbf{1}_n \otimes L_2)$, where $\otimes$ denotes the horizontal juxtaposition;

  (iii) $E(L_1 \lhd T_1) = E(L_2 \lhd T_2)$ and $E(T_1 \rhd L_1) = E(T_2 \rhd L_2)$, where the left action $\lhd$ and right action $\rhd$ of rT_{2n} on rT_n are defined in Figure 2.8;
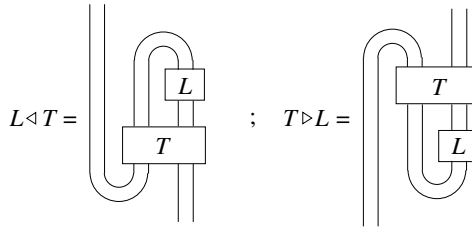
Figure 2.8. Schematical representations of the left and right actions of $T \in \mathrm{rT}_{2n}$ on $L \in \mathrm{rT}_n$.

(2) For all $L \in \mathrm{rT}_n$, there is a ribbon tube L', such that $E(L \cdot L') = E(\mathbf{1}_n)$;

(2′) For all $L \in \mathrm{rT}_n$, $E(L \cdot \overline{L}) = E(\mathbf{1}_n)$, where $\overline{L}$ denotes the image of L under the hyperplane reflexion about $B^3 \times \{\frac{1}{2}\}$;

(3) The equivalence relation E on ribbon torus-links is generated by isotopy of ribbon torus-links and the equivalence relation E on ribbon tubes: if L and L' are two ribbon torus-links such that $E(L) = E(L')$, then there is a finite sequence $L_1, \ldots, L_m$ of ribbon tubes such that L is isotopic to $\hat{L}_1$, L' is isotopic to $\hat{L}_m$, and for all i ($1 \leq i < m$), either $E(L_i) = E(L_{i+1})$ or $\hat{L}_i$ is isotopic to $\hat{L}_{i+1}$.

Let E be a local equivalence relation. Denote respectively by ES_n^R and ES_n^L the right and left stabilizers of the trivial ribbon tube in $E\mathrm{rT}_n$. One can easily check that ES_n^R and ES_n^L are both submonoid of rT_{2n}. Furthermore, the closure operation induces a map $\hat{\ }: E\mathrm{rT}_n \to E\mathrm{rL}_n$ which passes to the quotient by ES_n^R (respectively ES_n^L).

Now, assume in addition that the equivalence relation E satisfies Axiom (2). Then clearly the monoid $E\mathrm{rT}_n$ is a group, and both ES_n^R and ES_n^L are subgroups of $E\mathrm{rT}_{2n}$. If the stronger Axiom (2′) holds, then we actually have $ES_n^R = ES_n^L$.

Theorem 2.35 (Structure Theorem for ribbon torus-links).

- *Let E be a local equivalence relation satisfying axiom* (2). *Then, for $* = R$ or L, the quotient map*

$$\mathrm{rT}_n \longrightarrow E\mathrm{rT}_n \big/ ES_n^*$$

  factors through the closure map, i.e., we have a ribbon torus-link invariant

$$\widetilde{E} : \mathrm{rL}_n \longrightarrow E\mathrm{rT}_n \big/ ES_n^*.$$

  such that the composite map to $E\mathrm{rL}_n$ is E;
- *Furthermore, if Axiom* (3) *also holds, then we have a bijection*

$$E\mathrm{rT}_n \big/ ES_n^* = E\mathrm{rL}_n.$$

This structure theorem is shown by applying verbatim the arguments of [21], as reformulated in [23, Theorem 3.2]. Indeed, although these papers only deal with classical knotted objects, the proof is purely combinatorial and algebraic, and involves no topological argument *except* for [21, Proposition 2.1], whose ribbon tube analogue can actually be shown by a straightforward adaptation of Habegger and Lin's arguments.

We have the following classification result for ribbon torus-links up to link-homotopy.

Theorem 2.36. *The link-homotopy relation on ribbon tubes satisfies Axioms* (1), (2′) *and* (3) *above. Consequently, we have a bijection*

$$\mathrm{rT}_n^{\mathrm{h}} \big/ S_n^+ = \mathrm{rL}_n^{\mathrm{h}},$$

where $\mathrm{rL}_n^{\mathrm{h}}$ is the set of link-homotopy classes of ribbon torus-links and S_n^+ denotes the stabilizer of the trivial ribbon tube in $\mathrm{rT}_n^{\mathrm{h}}$ with respect to the right (or left) action of rT_{2n} on rT_n defined in Figure 2.8.

Proof. The fact that the link-homotopy relation is local (Axiom (1)) is evident. Axiom (2′) holds as a consequence of Theorem 2.30 and the group structure on monotone ribbon tubes as described in the proof of Proposition 2.4. To see that Axiom (3) holds, suppose that the ribbon torus-link L' is obtained from L by applying circle crossing changes at a given set S of self-crossing circles. Let L be the tube-closure of a ribbon tube L_1. As sketched at the beginning of this section, L_1 is specified by the choice of a base-ball intersecting each component of L exactly once, and we may assume that this ball is disjoint from S. Thus L' is the tube-closure of a ribbon-tube L_2, obtained from L_1 by successive circle crossing changes at each of the circles in the set S. This shows that link-homotopy for ribbon torus-links is implied by link-homotopy for ribbon tubes and isotopy, hence concludes the proof. □

2.5. Some examples

We conclude with a few simple examples of applications of the classification Theorems 2.34 and 2.36.

Let H be the ribbon tube represented in Figure 2.4, and denote by H^n the stacking product of n copies of H. Using Proposition 2.20, we have that $\varphi(H^n)$ acts on RF_2 by mapping x_1 to $x_2^n x_1 x_2^{-n}$ and fixing x_2. By Theorem 2.34, this provides an infinite 1-parameter family of 2-component ribbon tubes that are pairwise non link-homotopic.

This result is easily transferred to the case of ribbon torus-links. Namely, consider the tube-closure of H, which is the ribbon torus-link shown on the right-hand side of Figure 2.9, and more generally the tube-closure of H^n for all $n \geq 1$. These torus-links are pairwise non link-homotopic, which is in striking contrast with the case of ribbon 2-links, where all objects are link-homotopically trivial [7]. This can be proved topologically, by considering the homology class of a longitude of a torus component in the complement of the other component, or algebraically, as an application of Theorem 2.36.
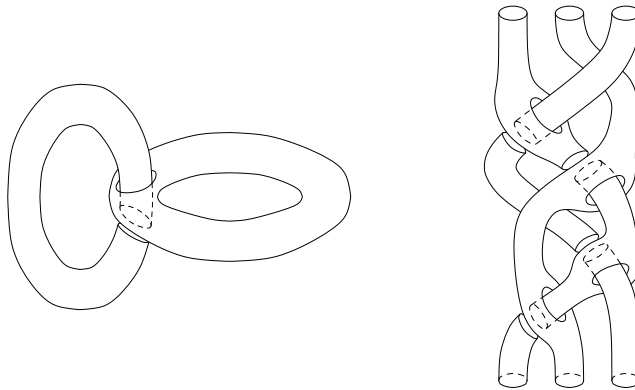
Figure 2.9. Examples of a ribbon torus-link and a ribbon tube that are not link-homotopically trivial.

Further examples involving more components can also be provided. For example, consider the 3-component ribbon tube B shown on the right-hand side of Figure 2.9. Note that B is *Brunnian*, meaning that removing any component yields the trivial 2-component ribbon tube. Unlike for the ribbon tube H, the homology class of a longitude in the complement of the other two components doesn't detect B. But using Proposition 2.20, the automorphism $\varphi(B)$ of RF_3 fixes x_1 and x_2 and maps x_3 to its conjugate by $[x_2^{-1}; x_1^{-1}]$. Theorem 2.34 then implies that B is not link-homotopic to the trivial ribbon tube. One can check that the tube-closure of B is likewise not link-homotopic to the trivial torus-link. One-parameter families of pairwise non link-homotopic 3-component ribbon tubes and torus-links can also be obtained by stacking multiples copies of B.

3. Welded string links

In this section we introduce two classes of diagrammatic objects, welded string links and welded pure braids, and we explain how they are related to the topological objects of the previous section. In particular, we give a classification result for welded string links up to self-virtualization.

3.1. Definitions

Definition 3.1. An n-component *virtual string link diagram* is a locally flat immersion L of n intervals $\bigsqcup_{i \in [\![1,n]\!]} I_i$ in $\overset{\star}{B}^2$, called *strands*, such that:

- Each strand I_i has boundary $\partial I_i = \{p_i\} \times \{0, 1\}$ and is oriented from $\{p_i\} \times \{0\}$ to $\{p_i\} \times \{1\}$ ($i \in [\![1, n]\!]$);
- The singular set $\Sigma(L)$ of L is a finite set of flatly transverse points.

Moreover, for each element of $\Sigma(L)$, a partial ordering is given on the two preimages. If the preimages are comparable, the double point is called *a classical crossing*, if not, it is called *a virtual crossing*. By convention, this ordering is specified on pictures by erasing a small neighborhood of the lowest preimage of classical crossings, and by circling the virtual one (see Figure 3.1).

A classical crossing is said to be *positive* if and only if the basis made of the tangent vectors of the highest and lowest preimages is positive. Otherwise, it is *negative* (see Figure 3.1).
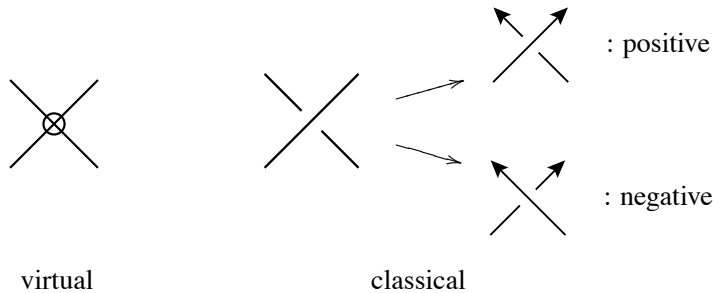

                                                                          : positive

                                                                          : negative

                virtual   classical

Figure 3.1. Virtual and classical crossings.

Up to isotopy, the set of virtual string link diagrams is naturally endowed with a monoidal structure by the stacking product $L \bullet L' := L \underset{\partial_1 L = \partial_0 L'}{\cup} L'$, and with unit element the trivial diagram $\underset{i \in [\![1,n]\!]}{\cup} p_i \times I$.

Definition 3.2. A virtual string link diagram is said to be *monotone* if it is flatly transverse to the lamination $\underset{t \in I}{\cup} I \times \{t\}$ of B^2.

Definition 3.3. Two virtual string link diagrams are equivalent if they are related by a finite sequence of the moves, called *generalized Reidemeister moves*, represented in Figure 3.2. There, all lines are pieces of strands which may belong to the same strand or not, and can have any orientation.

We denote by vSL_n the quotient of n-component virtual string link diagrams up to isotopy and generalized Reidemeister moves, which is compatible with the stacking product. We call its elements n-component *virtual string links*.

Virtual string links are interesting objects on their own, but since we are motivated in the first place by applications to ribbon tubes, we will focus on the following quotient.

Definition 3.4. We define the Over Commute (OC) move as

We denote by $wSL_n := vSL_n/OC$ the quotient of vSL_n up to OC moves, which is compatible with the stacking product. We call its elements n-component *welded string links*.

We denote by wP_n the subset of wSL_n whose elements admit a monotone representative.

Warning 3.5.

classical Reidemeister moves

virtual Reidemeister moves

mixed Reidemeister moves
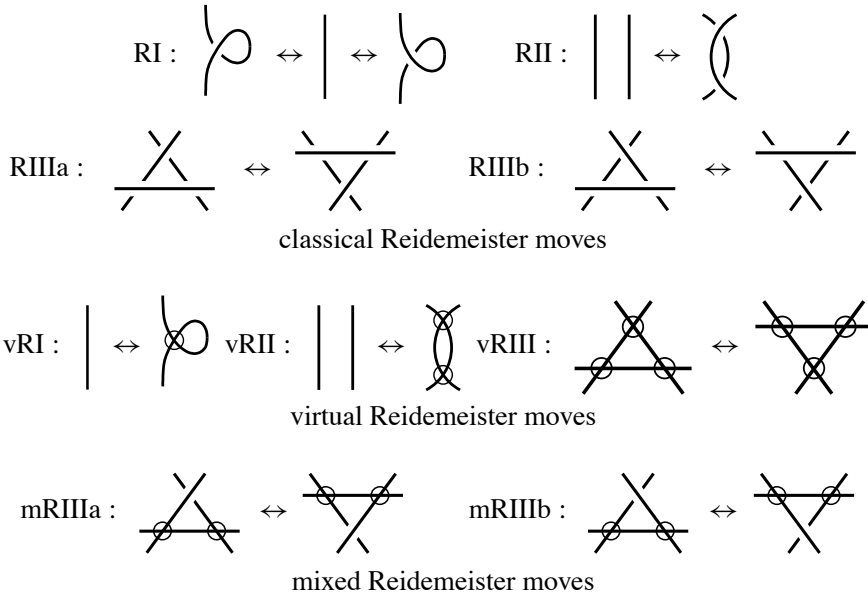
Figure 3.2. Generalized Reidemeister moves on diagrams.

- The following Under Commute (UC) move

$$\text{UC}: \qquad \longleftrightarrow \qquad ,$$
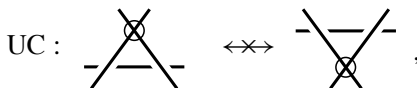
  was forbidden in the virtual context and is still forbidden in the welded context.
- Virtual and welded notions do not coincide, even for $n = 1$, where we get respectively the notion of virtual and welded long knots (see [4]).

Similarly to the ribbon tube case, it is straightforwardly checked that

Proposition 3.6. *The set* wP_n *is a group for the stacking product.*

Remark 3.7. Note that if two monotone virtual string link diagrams are equivalent in wSL_n, then they are related by a monotone transformation, that is by a sequence of isotopies, generalized Reidemeister and OC moves which remain within the set of monotone virtual string links. Indeed, it is a consequence of Remark 4.24 and Proposition 4.30 that any element of wP_n induces an action in $\text{Aut}_C(\text{F}_n)$.[3] But $\text{Aut}_C(\text{F}_n)$ is isomorphic to the group PUR_n, according to [4, Theorem 2.6], and a presentation of the latter is given by monotone virtual string links up to monotone transformations [8]. Since two equivalent monotone welded string links induce the same action on F_n, they are related by a monotone transformation (this is the same argument as in Remark 2.26). It follows that wP_n is isomorphic to the welded pure braid group studied, for instance, in [4]. In other words, we have that the welded pure braid group injects into wSL_n. On that account, we will freely call welded pure braids the elements of wP_n.

Remark 3.8. The subset of vSL_n whose elements admit a monotone representative is, of course, also defined. It is a group for the stacking product, which maps surjectively onto the virtual pure braid group, introduced in [6]. However, the injectivity of this map remains an open question. Indeed, unlike in the welded case addressed in Remark 3.7 above, it is still unknown whether a sequence of isotopies and generalized Reidemeister moves can always be modified into a monotone transformation.

Definition 3.9. Two virtual string link diagrams are related by a *self-virtualization* if one can be obtained from the other by turning a classical self-crossing (*i.e.* a classical crossing where the two preimages belong to the same component) into a virtual one. We call v–equivalence the equivalence relation on wSL_n generated by self-virtualization.

We denote by wSL_n^v the quotient of wSL_n under v–equivalence, which is compatible with the stacking product. We also denote by $\text{wP}_n^v \subset \text{wSL}_n^v$ the subset of elements having a monotone representative.

As we will see in Section 4.5, we have the following result as a straightforward consequences of Theorem 4.12.

---

[3] This fact can also be checked directly, by a straightforward adaptation of Remark 4.24 and Proposition 4.30 to welded string links.

Theorem 3.10. *Every welded string link is monotone up to self-virtualization.*

Furthermore, Theorem 4.17 implies immediately the following classification result for welded string links up to self-virtualization.

Theorem 3.11. $\mathrm{wSL}_n^v \cong \mathrm{wP}_n^v \cong \mathrm{Aut}_C(\mathrm{RF}_n)$ *as monoids.*

The notions of self-virtualization and self-crossing change for welded string links and pure braids, as well as for their usual and virtual counterparts, is further studied in [3].

3.2. Fundamental group for welded string link

Let L be a virtual string link diagram. Any strand of L is cut into smaller pieces by the classical crossings. More precisely, we call *overstrand*, any piece of strand of L whose boundary elements consist of either a strand endpoint or the lowest preimage of a classical crossing, and such that it contains no other lowest preimage of any classical crossing in its interior. Note that any highest preimage of a classical crossing is contained in an overstrand. We denote by $\mathrm{Over}(L)$ and $\mathrm{Cross}(L)$ the sets of, respectively, overstrands of L and classical crossings of L.

We orient all strands from $\partial_0 L$ to $\partial_1 L$. To any $c \in \mathrm{Cross}(L)$, we associate, as show in Figure 3.3:

- ε_c the sign of c;
- s_c^0 the overstrand containing the highest preimage of c;
- s_c^- the overstrand whose exiting boundary component is the lowest preimage of c;
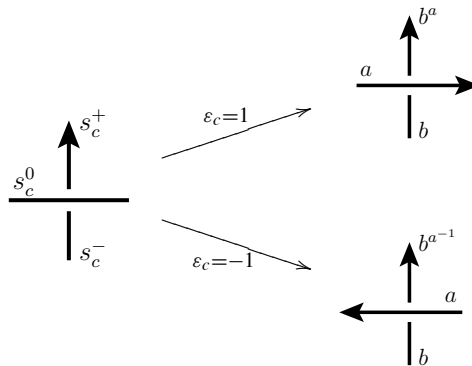- s_c^+ the overstrand whose entering boundary component is the lowest preimage of c.

Figure 3.3. Local relations for $\pi_1(L)$

Definition 3.12. We define the *fundamental group of L* as

$$\pi_1(L) := \left\langle \mathrm{Over}(L) \mid s_c^+ = (s_c^-)^{(s_c^0)^{\varepsilon_c}} \text{ for all } c \in \mathrm{Cross}(L) \right\rangle.$$

See Figure 3.3 for an illustration.

It is well-known that, up to isomorphism, the group associated to a virtual string link diagram is invariant under classical Reidemeister moves. Kauffman proved that virtual and mixed Reidemeister moves do not change the group presentation [26]. It turns out that the "virtual knot group" is also invariant under Over Commute, and is thus a welded invariant, *i.e.* is well defined on wSL_n [26,41].

3.3. Relations with ribbon tubes

It was shown in Section 2.2 that 4–dimensional ribbon tubes can be described by 3–dimensional objects, namely symmetric broken surface diagrams. Following [41] and [46], it is also possible to describe ribbon tubes using 2–dimensional welded string links.

Indeed, let L be a welded string link diagram. One can associate a symmetric broken surface diagram by embedding B^2 into B^3 as $B^2 \times \left\{\frac{1}{2}\right\}$ and considering a tubular neighborhood $N(L)$ of L in B^3 so that $\partial_\varepsilon N(L) = \bigsqcup_{i \in [\![1,n]\!]} D_i \times \{\varepsilon\}$ for $\varepsilon \in \{0,1\}$. The boundary of $N(L)$ then decomposes as a union of 4–punctured spheres, one for each crossing. Then, according to the partial order on the associated crossing, we modify each sphere as shown in Figure 3.4. The result is a symmetric broken surface diagram, to which we can associate a ribbon tube $\mathrm{Tube}(L)$.
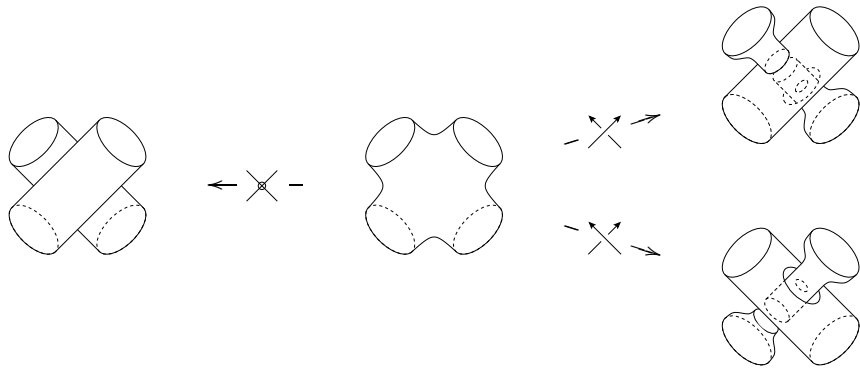

Figure 3.4. Inflating classical and virtual crossings.

There is a one-to-one correspondence between overstrands of L and the outside annuli of the associated symmetric broken surface diagram S, and another one-to-one correspondence between classical crossings of L and inside annuli of S.

Moreover, signs associated to crossing in the definition of $\pi_1(L)$ also correspond to signs associated to inside annuli in the Wirtinger presentation of $\pi_1\big(\text{Tube}(L)\big)$ (compare Figures 2.5 and 3.3). Therefore we have:

Proposition 3.13 ([41, 46]). *For every welded string link diagram L, $\pi_1\big(\text{Tube}(L)\big) \cong \pi_1(L)$.*

But the Tube map is more than just a tool to compute fundamental groups. Indeed, it provides a way to encode ribbon tubes. This has been pointed out by Satoh, but some key ideas already appeared in early works [46] of Yajima.

Proposition 3.14 ([41]). *The map* $\text{Tube} : \text{wSL}_n \to \text{rT}_n$ *is well defined and surjective.*

Proof. It is easily seen that classical Reidemeister moves on welded diagrams correspond to Reidemeister moves on symmetric broken surfaces defined in Remark 2.15, and that virtual and mixed Reidemeister moves, as well as OC, preserve the associated broken surface diagram. It then follows from Remark 2.15 that the Tube is well defined on wSL_n.

Now, a symmetric broken surface diagram S can be seen as given by a ribbon band, in the sense of Figure 2.2. Contracting each band $I \times I$ onto its core $\left\{\frac{1}{2}\right\} \times I$ so that, at each ribbon singularity, the cores intersect transversally, yields a singular string link. Let D be a diagram for this singular string link. By turning the classical crossings of D into virtual ones and its singular crossings into classical with signs corresponding to the initial local ordering on S, we obtain a welded string link diagram which is sent to S by the above process. This operation is illustrated in Figure 3.5.
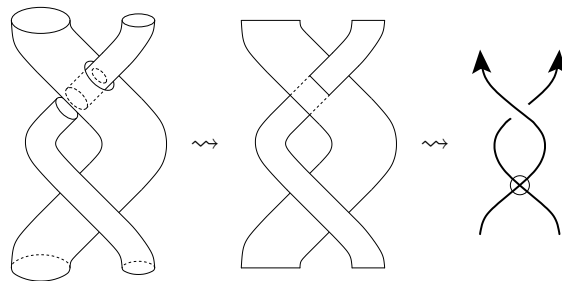

Figure 3.5. Contracting symmetric broken surface diagrams.

Note however that some inside annuli of S may have to be turned around so that the sign rule given in Figure 3.4 can be reversed, as illustrated in Figure 3.6. Surjectivity of the Tube map then follows from Lemma 2.14.   □
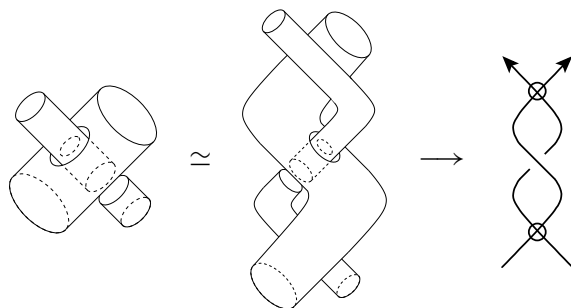
Figure 3.6. Turning inside annuli around.

Remark 3.15. The injectivity of the Tube map is still an open question; see [4,5].[4] However, Brendle and Hatcher proved in [8] that it is an isomorphism when restricting to monotone objects on both sides. This can be compared with the following proposition, although it is a not a consequence of Brendle–Hatcher's result since link-homotopy is inherently a non-monotone transformation.

Proposition 3.16. *The map* Tube : $\mathrm{wSL}_n^v \to \mathrm{rT}_n^{\mathrm{h}}$ *is a well defined group isomorphism*.

Proof. It is easily seen that a self-virtualization corresponds to a self-circle crossing change, see Figure 2.7. It is a surjective group homomorphism by Proposition 3.14. Injectivity is immediate after Corollary 4.34 and Proposition 4.30. □

4. Gauss diagrams

Our main tool for the study of ribbon tubes and welded knotted objects is the theory of Gauss diagrams [16,19,39].

4.1. Definitions

Definition 4.1. A *Gauss diagram* is a set of signed and oriented (thin) arrows between points of n ordered and oriented vertical (thick) strands, up to isotopy of the underlying strands. Endpoints of arrows are called *ends* and are divided in two parts, *heads* and *tails*, defined by the orientation of the arrow (which goes by convention from the tail to the head).

This definition is illustrated on the left-hand side of Figure 4.7.

For all $i \in [\![1, n]\!]$, we will denote the ith strand by I_i. An arrow is said to be *connected* to a strand if it has an end on this strand. An arrow having both ends on the same strand is called a *self-arrow*.

---

[4] In the knot case, the Tube map is known to have nontrivial kernel; see for example [24].

There is a natural stacking product operation, denoted by •, for Gauss diagrams defined by gluing the top endpoints of the strands of the first summand to the bottom endpoints of the strands of the second.

Definition 4.2. A Gauss diagram is said to be *horizontal* if all of its arrows are horizontal.

Remark 4.3. An alternative definition for a Gauss diagram G being horizontal would be to ask for a global order on the set of its arrows such that if two arrows a_1 and a_2 with $a_1 \leq a_2$ have ends e_1 and e_2 on a same strand, then e_1 is below e_2 on this strand. Note in particular that it forbids self-arrows. In the proof of Theorem 4.12, we shall refer to this global order.

Definition 4.4. Two Gauss diagrams are equivalent if they are related by a finite sequence of the following moves, called *Reidemeister moves*:
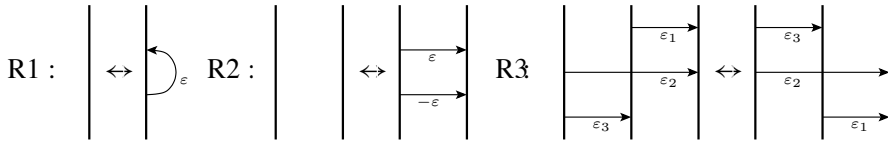
Here, all vertical lines are pieces of strands which may belong to the same strand or not, and can be oriented upward or downward; each ε_*, for $* \in \{\emptyset, 1, 2, 3\}$, is either 1 or −1. Moreover, there is an additional condition for applying move R3: it is required that $\tau_1 \varepsilon_1 = \tau_2 \varepsilon_2 = \tau_3 \varepsilon_3$, where $\tau_i = 1$ if the ith strand (from left to right) is oriented upwards, and −1 otherwise.

We denote by GD_n the quotient of Gauss diagrams up to isotopy and Reidemeister moves, which is compatible with the stacking product.

For the study of ribbon tubes and welded string links, we consider the following quotient of Gauss diagrams.

Definition 4.5. We define the Tail Commute (TC) move as
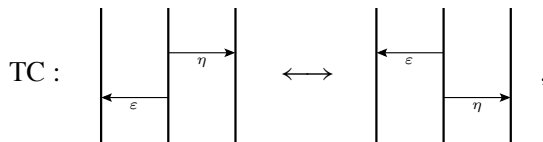
where $\varepsilon, \eta \in \{\pm 1\}$. We denote by $\mathrm{wGD}_n := \mathrm{GD}_n / \mathrm{TC}$ the quotient of GD_n by relation TC, which is compatible with the stacking product. We call its elements *welded Gauss diagrams*. We also denote by $\mathrm{wGP}_n \subset \mathrm{wGD}_n$ the subset of elements which have a horizontal representative.

Remark 4.6. In the welded case, that is when quotiented by TC, the condition $\tau_1\varepsilon_1 = \tau_2\varepsilon_2 = \tau_3\varepsilon_3$ for move R3 can be simplified to $\varepsilon_2\varepsilon_3 = \tau_2\tau_3$. In other words, the signs of the two arrows which have their tails on the same strand should agree if and only if the two others strands have parallel orientations.

The following can be easily checked:

Proposition 4.7. *The set* wGP_n *is a group for the stacking product.*

Definition 4.8. Two Gauss diagrams are related by a *self-arrow* move, denoted by SA, if one can be obtained from the other by removing a self-arrow.

We denote by $\mathrm{wGD}_n^{\mathrm{a}}$ the quotient of wGD_n by SA, which is compatible with the stacking product. We also denote by $\mathrm{wGP}_n^{\mathrm{a}} \subset \mathrm{wGD}_n^{\mathrm{a}}$ the subset of elements which have a horizontal representative.

Two elements of wGD_n are said to be a–*equivalent* if they are sent to the same element in $\mathrm{wGD}_n^{\mathrm{a}}$.

4.2. Commutation of arrows

In this section, we address the notion of commutation of arrows, which means swapping the relative position of two adjacent arrow ends on a strand. The Tail Commute move is a special case of such a commutation, when both ends are tails. More generally, and even in the non welded case, commutations of arrows can always be performed at the cost of some additional surrounding arrows. In the welded case, these additional arrows can be conveniently positioned, as shown below:

Definition 4.9. The C^3 moves are local moves on Gauss diagrams defined in three versions, shown in Figure 4.1. There, ε and η are either 1 or -1, all strands are assumed to be simultaneously upward or downward oriented, and non oriented arrows can have either orientation.
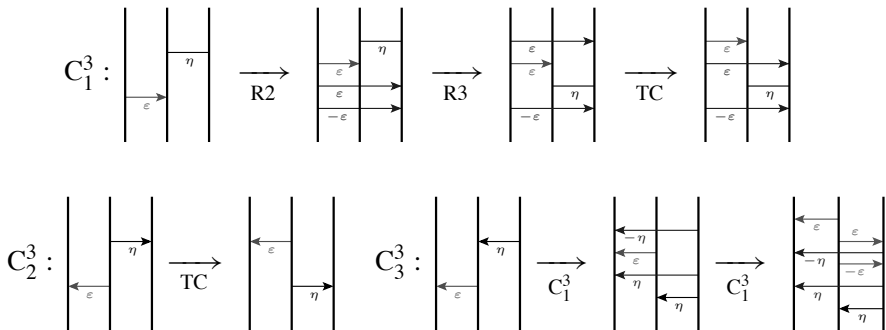

Figure 4.1. The C^3 moves on welded Gauss diagrams.

The C^3 moves can be seen as pushing one arrow across another, at the cost of several additional arrows located below the pushed one. Note that C^3 moves are supported by three pieces of strands, which may or may not belong to the same components.

By iterated C^3 moves, one has the following general commutation rule:

Corollary 4.10. *In* wGD_n, *commuting an arrow with a bunch of adjacent arrows all connected to a far away strand is achieved at the cost of additional arrows connected to the far away strand, as shown below:*
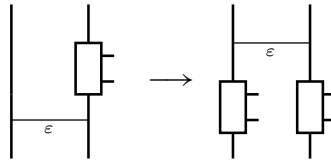
where *denotes a bunch of arrows (not necessarily with the same number of arrows),* ε *is either* 1 *or* -1, *and all orientations are arbitrary.*

Proof. If the two strands are simultaneously upward or downward, it is a direct applications of the C^3 moves as described in Figure 4.1, where the far away strand is the rightmost one. If not, the final step of moves C_1^3 and C_3^3 have to be changed consequently. □

When working up to SA, and dealing with arrows whose ends are on the same two strands, a genuine commutation result holds (without additional arrows), even in the non welded case. We shall refer to such a commutation as a C^2 move.

Proposition 4.11. *Up to self-arrow moves, two arrows with adjacent ends can commute whenever the other two ends are on the same strand.*

Proof. The strategy is to add a self arrow on the strand which does not support the adjacent ends, so that an R3 move involving the three arrows can be performed, and finally to remove the self-arrow. See Figure 4.2 for an example. Nonetheless, the self-arrow has to be choosen so that the R3 move is valid. This means that the global position of the three arrows should be as in the definition of an R3 move (Definition 4.4), and the condition $\tau_1\varepsilon_1 = \tau_2\varepsilon_2 = \tau_3\varepsilon_3$ should hold. However, we have a complete freedom in choosing the self–arrow: orientation, sign, and relative position to the adjacent arrow ends. The choice of orientation ensures that the arrows are in position of a R3 move, while the sign and positions give independant control on $\tau_1\varepsilon_1$, $\tau_2\varepsilon_2$ and $\tau_3\varepsilon_3$. □
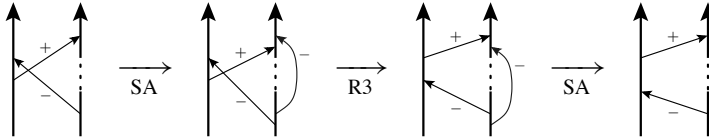
Figure 4.2. Example of a C^2 move.

4.3. Reduction of welded Gauss diagrams to horizontal Gauss diagrams

The main result of this section is the following:

Theorem 4.12. *Every welded Gauss diagram is a–equivalent to a horizontal Gauss diagram. Equivalently, the natural inclusion $\mathrm{wGP}_n^a \hookrightarrow \mathrm{wGD}_n^a$ is onto, that is $\mathrm{wGP}_n^a \cong \mathrm{wGD}_n^a$.*

Proof. We prove the statement by induction on n. For $n = 1$ the statement is trivial since wGD_1^a is reduced to one element.

Now, we assume that the result is true for $n \in \mathbb{N}$ and we consider $G \in \mathrm{wGD}_{n+1}$. We choose a strand I of G and we call I–*arrow* the arrows connected to I. We denote by $G_{\hat{I}}$ the Gauss diagram obtained from G by removing the strand I, that is, by removing all I–arrows and forgetting I. The welded Gauss diagram $G_{\hat{I}}$ is then a diagram on n strands, and by the induction hypothesis, there is a finite sequence S of moves R1, R2, R3, TC and SA which transforms $G_{\hat{I}}$ into a horizontal Gauss diagram B. Now, this sequence cannot be directly performed on T, since most moves require that some ends of arrows are adjacent on strands, and I–arrows may interfere in that. Nevertheless, Figure 4.3 illustrates how to use Corollary 4.10 to push away such I–arrows at the cost of additional I–arrows, and perform the desired moves. There, boxes stand for bunches of I–arrows.

It follows that the whole sequence S can be performed on T, giving a representative of T obtained by adding I–arrows to B.

Now, since B is horizontal, there is a natural total order, from bottom to top, on its arrows. The arrows of B can then be pushed up above the I–arrows, successively in decreasing order, using moves C^3 (for I being the strand on the right). This leads to the following decomposition:
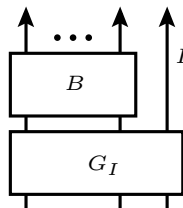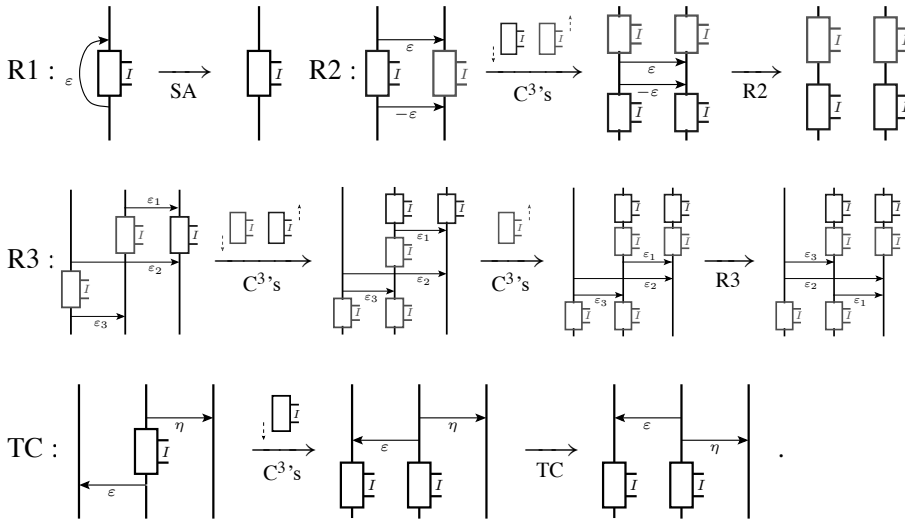
Figure 4.3. Performing Reidemeister and TC moves obstructed by I-arrows.

where B is horizontal and G_I contains only I–arrows. Now, for any strand $J \neq I$ restricted to G_I, all arrows connected to J are also connected to I, so that one can use moves C^2 (for I being the strand on the right), to rearrange the ends on strand J so that their order corresponds to the order induced by I. Applying this operation for all strands $J \neq I$, we eventually obtain a horizontal Gauss diagram $\widetilde{G}_I$, such that G is a–equivalent to the product of $\widetilde{G}_I$ with B, which is still horizontal. This concludes the proof. □

Remark 4.13. As noted above, C^2 moves still hold in GD_n^a, *i.e.* without using TC moves. This implies that any element of GD_2^a is a–equivalent to a horizontal Gauss diagram. This is not likely to hold for GD_n^a with $n \geq 3$.

Corollary 4.14. *The set* wGD_n^a *is a group for the stacking product.*

We now establish a second normalized form for welded Gauss diagrams up to a–equivalence.

Definition 4.15. A Gauss diagram is said to be *ascending* if all tails belong to the lowest halves of the strands, whereas all heads are on the highest halves, that is, if any tail is below any head.

Lemma 4.16. *Every welded Gauss diagram is* a-*equivalent to an ascending Gauss diagram.*

Proof. Consider a Gauss diagram and choose an arbitrary order for its strands. Then, using move C_1^3 repeatedly, each strand can be successively sorted, in the sense that all tails can be pushed below all heads. When performing such C_1^3 moves, the tails (respectively heads) of added arrows are in the neighborhood of a tail (respectively head) of a pre-existent arrow, so none of the performed C_1^3 moves will unsort the already sorted strands. $\qquad\square$

4.4. Classification of welded Gauss diagram up to a-equivalence

In this section we prove the following:

Theorem 4.17. *There is a group isomorphism* $\mathrm{wGD}_n^a \cong \mathrm{Aut}_C(\mathrm{RF}_n)$.

Recall that $\mathrm{Aut}_C(\mathrm{RF}_n)$ is the group of conjugating automorphisms defined in Section 1. The isomorphism between wGD_n^a and $\mathrm{Aut}_C(\mathrm{RF}_n)$ is given explicitly. In Section 4.4.1 we establish the existence of a group homomorphism $\varphi_{G\to A}\colon \mathrm{wGD}_n^a \longrightarrow \mathrm{Aut}_C(\mathrm{RF}_n)$, and in Section 4.4.2 we construct an inverse $\varphi_{A\to G}$ for $\varphi_{G\to A}$.

4.4.1. *Definition of* $\varphi_{G\to A}$

To construct the map $\varphi_{G\to A}$, we need some notation.

Definition 4.18. Let G be a Gauss diagram. A *tail interval* is a pair (h_1, h_2) such that one of the following holds:

- h_1 and h_2 are heads on a same strand, with h_1 lower than h_2, and there is no other head between them;
- h_1 is a head and h_2 is the top endpoint of the strand containing h_1, and there is no other head on this strand above h_1;
- h_2 is a head and h_1 is the bottom endpoint of the strand containing h_2, and there is no other head on this strand below h_2;
- h_1 and h_2 are respectively the bottom and top endpoints of a strand that doesn't contain any head.

Graphically, tail intervals are portions of strand comprised between two successive heads and/or strand endpoints. A tail interval may contain some tails, but no head. We denote by T_G the set of all tail intervals of G.

Notation 4.19. Let G be a Gauss diagram. For every head h, we denote by T_h^+ (respectively T_h^-) the unique tail interval of the form $(h, \ . \ )$ (respectively $(\ . \ , h)$) and by T_h^0 the unique tail interval that contains the tail which is connected to h by an arrow, as illustrated below:
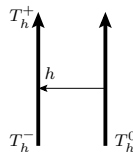
Finally, for every $i \in [\![1, n]\!]$, we denote by T_i^+ and T_i^- the unique tail intervals containing respectively the top and bottom endpoint of the ith strand (note that $T_i^+ = T_i^-$ if the ith strand contains no arrow head).

Lemma 4.20. *For any Gauss diagram G, there is a unique coloring map $\xi_G \colon T_G \to \mathrm{RF}_n$ such that*

(1) *for every $i \in [\![1, n]\!]$, $\xi_G(T_i^-) = x_i$;*

(2) *for every head h, $\xi_G(T_h^+) = \xi_G(T_h^-)^{\xi_G(T_h^0)^{\varepsilon_h}}$, where ε_h is the sign of the arrow that contains h.*

Moreover, for every $i \in [\![1, n]\!]$, $\xi_G(T_i^+)$ only depends on the a*–equivalence class of G.*

See Remark 4.23 for an example. Note that condition (2), together with condition (1), imply that, for each $i \in [\![1, n]\!]$, every tail interval which belongs to I_i is necessarily mapped to a conjugate of x_i.

Proof. First, we deal with the case of horizontal Gauss diagrams. If G contains a single ε–labelled non self-arrow, then the lemma is clear. Indeed, the only possible coloring sends T_1^- to x_1, T_1^+ to $x_1^{x_2^\varepsilon}$ and all other $T_k^+ = T_k^-$ to x_k, as pictured below:
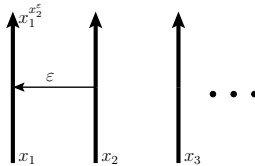
By induction on the number of arrows the result then follows whenever G is horizontal.

Now, each Reidemeister, TC or SA move between two Gauss diagrams G_1 and G_2 induces a one-to-one correspondence between the coloring maps on G_1 and on G_2. These correspondences are shown in Figure 4.4. As usual, for every picture, there is a non represented part which is identical on both side of each move. Note that this is consistent since the ends of the represented tail intervals are pairwise coherently labelled. In particular, in the case of moves SA and R1, the equality $b^a = b$ holds in RF_n since a and b are both conjugate of a same generator x_i for some $i \in [\![1, n]\!]$. By Theorem 4.12, this implies that any Gauss diagram has a unique coloring ξ with the desired property. Moreover, Figure 4.4 also shows that only tail intervals which are not of the form T_i^+ or T_i^- are possibly modified by Reidemeister, TC and SA moves. As a matter of fact, $\xi(T_i^+)$ is kept unchanged for every $i \in [\![1, n]\!]$, and depends on the a–equivalence class of G only. $\qquad\square$
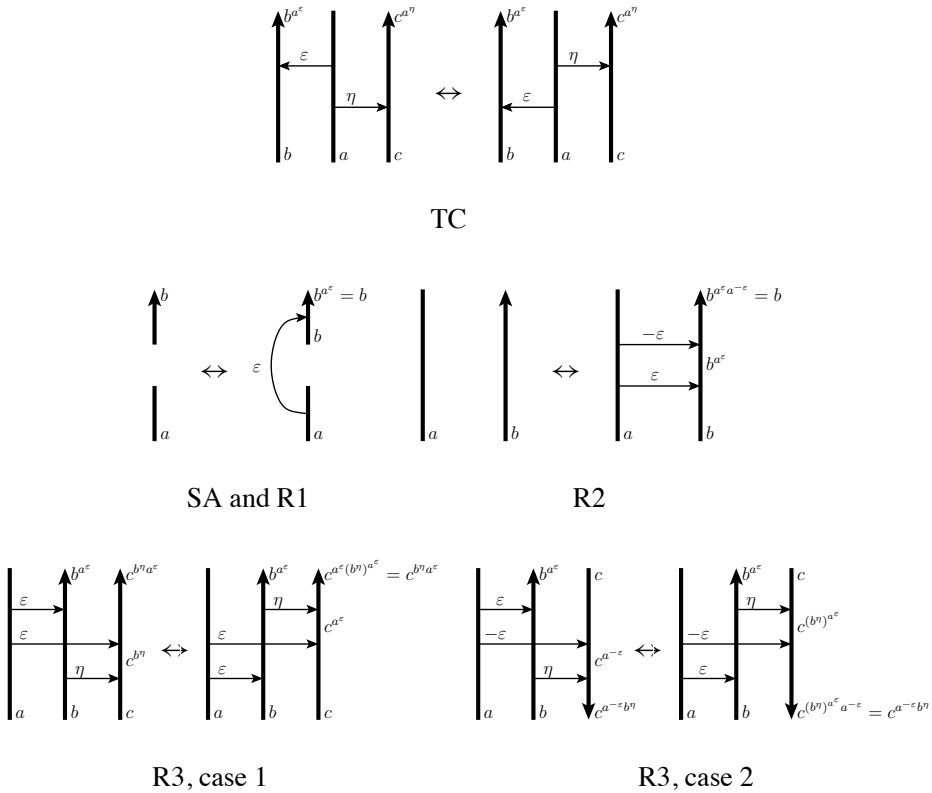
**Figure 4.4.** One-to-one correspondence between the coloring maps ξ.

Remark 4.21. The above proof of Lemma 4.20 can be easily adapted to show that the number of F_n–coloring map $\xi_G : T_G \to F_n$ for a given Gauss diagram G is a well defined invariant on GD_n. Moreover, when such an F_n–coloring map is unique, for instance when $G \in GP_n$, then the value $\big(\xi_G(T_i)\big)_{i \in [\![1,n]\!]} \in \mathbb{F}_n^n$ is also a well-defined invariant.

Definition 4.22. For every $G \in \mathrm{wGD}_n^{\mathrm{a}}$, we define $\varphi_{\mathrm{G} \to \mathrm{A}}(G) \in \mathrm{Aut}_\mathrm{C}(\mathrm{RF}_n)$ as the automorphism of RF_n which sends, for every $i \in [\![1,n]\!]$, the generator x_i to $\xi_G(T_i^+)$.

It is easily checked that $\varphi_{\mathrm{G} \to \mathrm{A}} : \mathrm{wGD}_n^{\mathrm{a}} \to \mathrm{Aut}_\mathrm{C}(\mathrm{RF}_n)$ is a group homomorphism, so that it defines an action of $\mathrm{wGD}_n^{\mathrm{a}}$ on RF_n.

Remark 4.23. If G is ascending, as defined in Definition 4.15, then, for any $i \in [\![1,n]\!]$, the element $g_i \in \mathrm{RF}_n$ by which $\varphi_{\mathrm{G} \to \mathrm{A}}(G)$ conjugates x_i can be read directly on G as $x_{i_1}^{\varepsilon_1} \cdots x_{i_s}^{\varepsilon_s}$ where T_i^- contains exactly the tails $t_1, \cdots, t_s$ in this order and for each $k \in [\![1, s]\!]$, t_k is connected by a ε_k–signed arrow to a head on I_{i_k}. For

example,
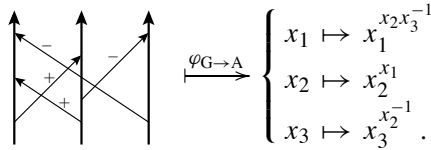
Remark 4.24. According to Remark 4.21, $\varphi_{G\to A}$ can be refined into an action of GP_n on F_n.

4.4.2. *Definition of $\varphi_{A\to G}$*

We will now define an inverse for $\varphi_{G\to A}$. Let us first define, for every $i \in [\![1, n]\!]$, the map

$$\rho_i \colon \mathrm{Aut}_C(\mathrm{RF}_n) \longrightarrow \mathrm{RF}_{n-1}^{(i)} = R\langle x_1, \cdots, \hat{x}_i, \cdots, x_n\rangle$$

as the unique function such that, for every $\varphi \in \mathrm{Aut}_C(\mathrm{RF}_n)$, we have $\varphi(x_i) = x_i^{\rho_k(\varphi)}$.

Lemma 4.25. *For all $i \in [\![1, n]\!]$, the map ρ_i is well defined.*

Proof. By definition of $\mathrm{Aut}_C(\mathrm{RF}_n)$, there exists some $g \in \mathrm{RF}_n$ such that $\varphi(x_i) = x_i^g$. Since for every $g_1, g_2 \in \mathrm{RF}_n$ and every $\varepsilon \in \{\pm 1\}$, we have the equality

$$x_i^{g_1 x_i^\varepsilon g_2} = g_2^{-1} x_i^{-\varepsilon} g_1^{-1} x_i g_1 x_i^\varepsilon g_2 = g_2^{-1} g_1^{-1} x_i g_1 x_i^{-\varepsilon} x_i^\varepsilon g_2 = g_2^{-1} g_1^{-1} x_i g_1 g_2 = x_i^{g_1 g_2}$$

in RF_n, we may assume that g is represented by a word which does not contain x_i, *i.e.* that $g \in \mathrm{RF}_{n-1}^{(i)}$. Now, it remains to prove that this g is uniquely defined in $RF_{n-1}^{(i)}$. Suppose that there exists $g_1, g_2 \in \mathrm{RF}_{n-1}^{(i)}$ such that $x_i^{g_1} = x_i^{g_2}$. Then x_i commutes with $g_1 g_2^{-1}$, that is, $[x_i, g_1 g_2^{-1}] = 1$. Applying the Magnus expansion E, defined in Section 5.1, to the latter equality gives that $X_i.G - G.X_i = 0$, where $E(x_i) = 1 + X_i$ and $E(g_1 g_2^{-1}) = 1 + G$. But since $g_1 g_2^{-1} \in \mathrm{RF}_{n-1}^{(i)}$, the power series G does not contain the variable X_i, which implies that $X_i.G = G.X_i = 0$. So $G = 0$, which by injectivity of the Magnus expansion (see [45]) implies $g_1 = g_2$. $\square$

Now, we define a map

$$\eta \colon \prod_{i=1}^n \mathrm{RF}_{n-1}^{(i)} \longrightarrow \mathrm{wGD}_n^{\mathrm{a}}$$

as follows. Let $(g_1, \cdots, g_n) \in \prod_{i=1}^n \mathrm{RF}_{n-1}^{(i)}$. For each $i \in [\![1, n]\!]$, choose a word $g_i = x_{j_1^i}^{\varepsilon_1^i} \cdots x_{j_{s_i}^i}^{\varepsilon_{s_i}^i}$ representing g_i. Then we define $\eta(g_1, \cdots, g_n)$ as the image in

$\mathrm{wGD}_n^{\mathrm{a}}$ of the Gauss diagram obtained from the trivial one by adding successively, according to the lexicographical order on $\{(i,k)\,|\,i \in [\![1,n]\!], k \in [\![1,s_i]\!]\}$, an ε_k^i–signed arrow with head at the very top of I_i and tail at the very bottom of $I_{j_k^i}$. Note that this actually defines an ascending Gauss diagram. See Figure 4.6 for an example.

Lemma 4.26. *The map η is well defined.*

Proof. The fact that η does not depend on the order of the strands is guaranteed by relation TC, so we only have to show that $\eta(g_1, \cdots, g_n)$ does not depend on the words representing the n variables. Fix $i \in [\![1,n]\!]$. Any two words representing g_i differ by a finite number of the following moves:

(1) $(1 \leftrightarrow x_j^{-\varepsilon} x_j^{\varepsilon})$, for some $j \in [\![1,n]\!]$ and some $\varepsilon \in \{\pm 1\}$;

(2) $(x_j x_j^g \leftrightarrow x_j^g x_j)$, for some $j \in [\![1,n]\!]$ and some word $g := x_{i_1}^{\varepsilon_1} \cdots x_{i_s}^{\varepsilon_s} \in \mathrm{RF}_{n-1}^{(i)}$.

So we only need to check invariance of η under these two types of moves. Invariance under move (1) follows easily from move R2 on Gauss diagrams. Now, let us consider move (2). Using the identity $x_j^{g_1 x_j^{\pm 1} g_2} = x_j^{g_1 g_2}$, we can assume that the word g involved in the move is in $\mathrm{RF}_{n-2}^{(i,j)} := R\langle x_1, \cdots, \hat{x}_i, \hat{x}_j, \cdots, x_n \rangle$. We need to prove that the following equality holds in $\mathrm{wGD}_n^{\mathrm{a}}$


$$(4.1)$$

where B_k^{ε} stands for the Gauss diagram with only one ε–labeled arrow whose tail is on I_k and head on I_i. Note that, in particular, all arrows in the boxes have their head on strand I_i. The proof is done by proving each side to be equal to a third Gauss diagram. This is shown in two main steps, and the reader is encouraged to consult the example given in Figure 4.5 while reading the following proof.

First, we pull the lowest (+–signed) arrow in the left-hand side Gauss diagram in (4.1) across the lowest block (*i.e.* across $B_{i_s}^{-\varepsilon_s} \bullet \cdots \bullet B_{i_1}^{-\varepsilon_1}$) by a sequence of moves TC and C_1^3 (with, in Figure 4.1, I_j being the rightmost strand, I_i the middle one,

and with all strands oriented downwards), and then use move C_2^2 to commute the heads of the two pictured $+$–signed arrows. This produces the following equality in wGD_n^a where $\tilde{B}_k^\varepsilon$ is obtained from B_k^ε by moving the head of each arrow lying on I_i to the strand I_j (preserving the horizontality and height of the arrow):
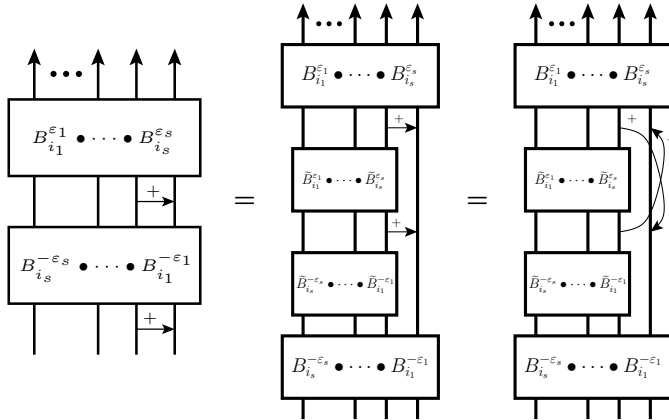
Next, we consider the right-hand side Gauss diagram in (4.1). Using TC moves, we pull the tail of the highest ($+$–signed) arrow just below the tail of the middle ($+$–signed) arrow. Then, we pull its head across the highest block, using moves C_1^3 (with, in Figure 4.1, I_j the rightmost strand, I_i the middle one, but with all strands oriented upwards). This leads to
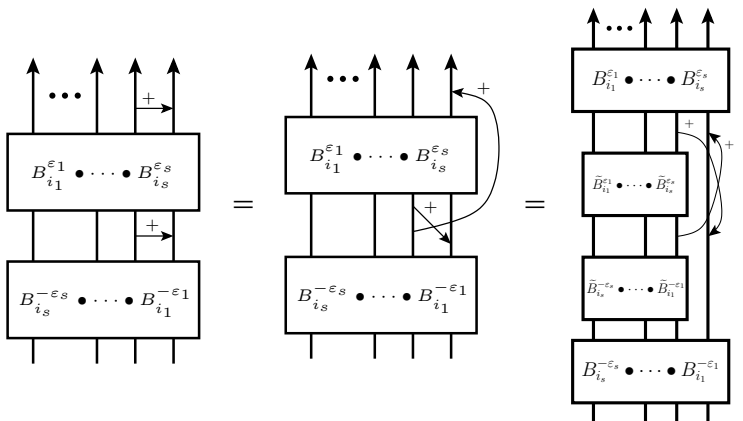
and proves that (4.1) holds in wGD_n^a. $\square$

The example in Figure 4.5 illustrates both steps of the above proof.

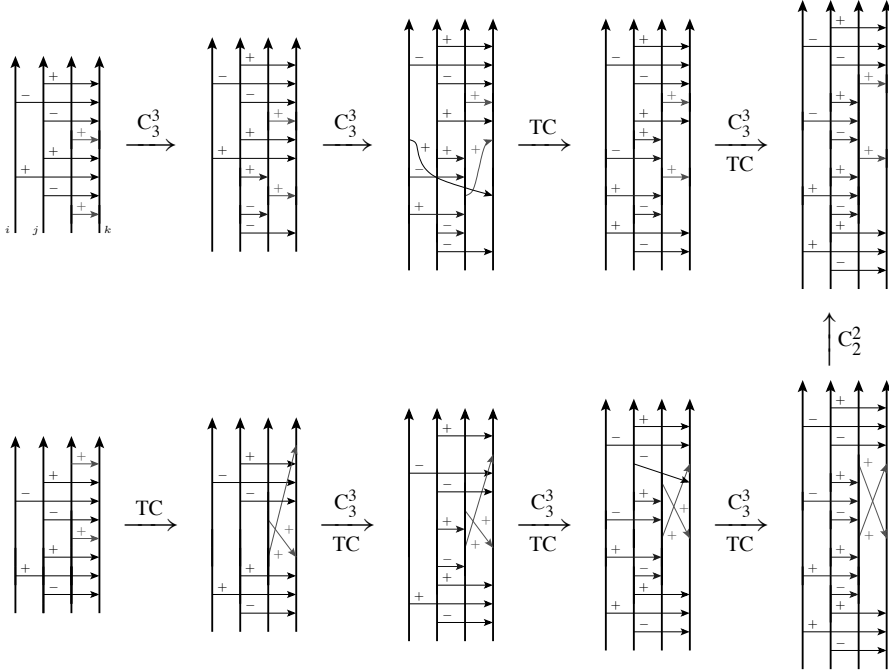**Figure 4.5.** Illustration of the two-step procedure in the proof of equality (4.1).

Definition 4.27. The map $\varphi_{A \to G} \colon \operatorname{Aut}_C(RF_n) \longrightarrow wGD_n^a$ is defined as the composition $\eta \circ \left(\prod_{i=1}^{n} \rho_i \right)$.

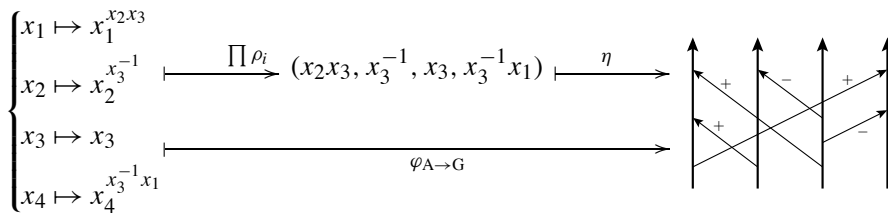The example for $\varphi_{A \to G}$ given in Figure 4.6 is to be compared with Remark 4.23.

Figure 4.6. An example for $\varphi_{A \to G}$.

Lemma 4.28. *The map $\varphi_{A \to G}$ is an inverse for $\varphi_{G \to A}$.*

Proof. The relation $\varphi_{G \to A} \circ \varphi_{A \to G} = \mathrm{Id}_{\mathrm{Aut}_C(\mathrm{RF}_n)}$ holds by construction. The fact that $\varphi_{A \to G} \circ \varphi_{G \to A} = \mathrm{Id}_{\mathrm{wGD}_n^a}$ is clear for ascending Gauss diagrams, which is enough, by Lemma 4.16. $\square$

We conclude with the following corollary which, curiously enough, is not easily checked directly.

Corollary 4.29. *The map $\varphi_{A \to G}$ is a group homomorphism.*

4.5. Relation with welded string link and ribbon tubes

As the notation suggests, welded Gauss diagrams are merely combinatorial tools for describing and studying welded string links. Indeed, to a string link D, one can associate a Gauss diagram G by considering the strands of G as parametrization intervals for the strands of D and add arrows between the preimages of each classical crossing of D, oriented from the overpassing strand to the underpassing one, and signed by the sign of the corresponding crossing. Then Reidemeister moves on D correspond to the eponymous moves on G, OC corresponds to TC and a self-virtualization corresponds to the addition or removal of a self-arrow.

Conversely, to a Gauss diagram G, one can associate a string link by drawing one crossing c_a for each arrow a labelled by the sign of a, and then, for each piece of strand of G comprised between two ends e_1 and e_2 —with e_1 lower than e_2 and a_1 and a_2 the arrows that respectively contain e_1 and e_2— connecting the outgoing endpoint of the overpassing (respectively underpassing) strand of c_{a_1} if e_1 is a tail (respectively a head), to the ingoing endpoint of the overpassing (respectively underpassing) strand of c_{a_2} if e_2 is a tail (respectively a head) —this can always be done at the cost of adding some virtual crossings— and finally closing similarly the string link to its bottom and top endpoints according to the lowest and highest portions of strands of G. See Figure 4.7 for an example. This is well defined up to virtual moves and the correspondence between moves on diagrams and moves on Gauss diagrams mentioned above also holds.
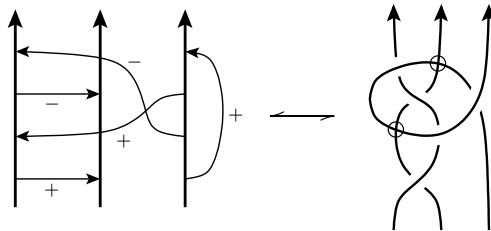

Figure 4.7. Correspondence between string link diagrams and Gauss diagrams.

Moreover, it is easily seen that horizontality for Gauss diagrams corresponds to monotony for string links. As a consequence we have the following statement,

which is considered as folklore in the literature:

Proposition 4.30. $wSL_n \cong wGD_n$, $wP_n \cong wGP_n$, $wSL_n^v \cong wGD_n^a$ *and* $wP_n^v \cong$ wGP_n^a *as monoids. Moreover, under these isomorphisms, there is a one-to-one correspondence between*

- *classical crossings and arrows;*
- *overstrands and tail intervals.*

Note that these correspondences already hold in the virtual setting.

Using these correspondences, constructions and results on welded string link can be transferred to Gauss diagrams and *vice versa*. In particular, we obtain Theorems 3.10 and 3.11 as direct corollaries of Theorems 4.12 and 4.17, respectively.

Furthermore, Satoh's Tube map can now be defined on wGD_n. By abuse of notation, we will still denote it by Tube. Using Notation 4.19, it is a consequence of Proposition 3.13 that

Proposition 4.31. *For every Gauss diagram G one has*

$$\pi_1\big(\text{Tube}(G)\big) \cong \big\langle T_G \mid T_h^+ = (T_h^-)^{(T_h^0)^{\varepsilon_h}} \text{ for each arrow head } h \text{ of } G\big\rangle,$$

where ε_h is the sign of the arrow that contains the head h. Moreover, through this isomorphism and for every $i \in [\![1, n]\!]$, the generator T_i^- (respectively T_i^+) corresponds to the loop that positively enlaces the ith bottom (respectively top) meridian of Tube(G).

Notation 4.32. For every Gauss diagram G, we denote by $W_G : T_G \to R\pi_1\big(\text{Tube}(G)\big)$ the map induced by the isomorphism mentioned in the previous proposition.

Recall that φ denotes the isomorphism of the classification Theorem 2.34. We have

Lemma 4.33. $\varphi_{\text{G} \to \text{A}} = \varphi \circ \text{Tube}$.

Proof. We freely use the notation of Section 2.3.2. Let G be a Gauss diagram. We consider ι_0 and ι_1 the inclusion maps associated to Tube(G). It follows from the Wirtinger presentation that the map $\iota_0^{*-1} \circ W_G$ satisfies conditions (1) and (2) of Lemma 4.20. By uniqueness, it follows that $\iota_0^{*-1} \circ W_G$ is the coloring map ξ_G of this lemma. Moreover, by Proposition 4.31 and the definition of ι_1, we also know that $\iota_1^*(x_i) = W_G(T_i^+)$. So, finally, we get

$$\varphi\big(\text{Tube}(G)\big)(x_i) = \big(\iota_0^{*-1} \circ \iota_1^*\big)(x_i) = \iota_0^{*-1}\big(W_G(T_i^+)\big) = \xi_G(T_i^+). \qquad \square$$

Since $\varphi_{\text{G} \to \text{A}}$ is an isomorphism, we obtain:

Corollary 4.34. *The map* Tube $: wGD_n^a \to rT_n^h$ *is injective.*

Since Tube was already known to be surjective, we obtain that the map $\varphi :$ $rT_n^h \to \text{Aut}_C(RF_n)$ is a group isomorphism, as stated in Theorem 2.34.

5. Virtual Milnor invariants

John Milnor defined in the fifties a family of link invariants, known today as Milnor's $\overline{\mu}$-invariants [36,37]. Given an n–component link L in the 3–sphere, Milnor invariants $\overline{\mu}_I(L)$ of L are defined for any finite sequence I of (possibly repeating) indices in $[\![1, n]\!]$. Roughly speaking, Milnor's $\overline{\mu}$–invariants measure the longitudes in the lower central series of the fundamental group of the link complement. These invariants, however, are in general not well-defined integers, and their definition contains a rather intricate self-recurrent indeterminacy Δ. In [21], Habegger and Lin showed that the indeterminacy in Milnor invariants of a link is equivalent to the indeterminacy in representing this link as the closure of a string link, and that Milnor invariants are actually well defined invariants of string links; the latter integer–valued string link invariants are usually denoted μ_I.

Several authors have given tentative extensions of Milnor invariants to virtual knot theory [13,31,32]. They are all based on various combinatorial approaches of Milnor invariants, and are only partial extensions; in particular, they are limited to the case of link-homotopy invariants, that is, to the case of Milnor invariants indexed by sequences with no repetition. In this section, we give a full extension of Milnor invariants to the virtual setting, in what seems to be the most natural way. Habegger and Lin's observation mentioned above suggests that it is most natural to consider the *string link* case. Moreover, we have seen in the introduction that we are actually seeking for an invariant of *welded* string links, rather than virtual, since Milnor invariant are extracted from the fundamental group, which is a welded invariant. Finally, our construction is purely topological, as we shall see below, and well-behaved with respect to the Tube map, so that it naturally coincides with Milnor's original invariants when restricted to classical string links (see Theorem 5.4).

5.1. Milnor invariants for ribbon tubes

In this subsection we develop a higher-dimensional analogue of Milnor invariants for ribbon tubes. This follows closely, and generalizes the construction given in Section 2.2.4 in the case of the reduced free group. As a consequence, we freely borrow all notation from there.

Let T be a ribbon tube with tube components $\underset{i \in [\![1,n]\!]}{\sqcup} A_i$ and let $W = B^4 \backslash \overset{\star}{N}(T)$ be the complement of an open tubular neighborhood of T in B^4. By Proposition 2.7, the inclusion $\iota_\varepsilon : \partial_\varepsilon W \hookrightarrow W$ ($\varepsilon \in \{0, 1\}$) induces isomorphisms both at the level of the first and second homology groups, so by Stallings theorem the maps

$$(\iota_\varepsilon)_k : \pi_1(\partial_\varepsilon W)\big/_{\Gamma_k \pi_1(\partial_\varepsilon W)} \overset{\simeq}{\longrightarrow} \pi_1(T)\big/_{\Gamma_k \pi_1(T)}$$

are isomorphisms for every $k \in \mathbb{N}^*$.

Recall from Remark 2.19 that the preferred i^{th} longitude λ_i of T is the longitude of A_i having linking number zero with A_i. For any $k \in \mathbb{N}$ and $i \in [\![1, n]\!]$, we define $\lambda_i^k := (\iota_0)_k^{-1}(\lambda_i) \in \mathrm{F}_n\big/_{\Gamma_k \mathrm{F}_n}$.

Denote by $\mathbb{Z}\langle\langle X_1, \cdots, X_n\rangle\rangle$ the ring of formal power series in non-commutative variables $X_1, \cdots, X_n$. The *Magnus expansion* $E : F_n \to \mathbb{Z}\langle\langle X_1, \cdots, X_n\rangle\rangle$ is the injective group homomorphism which maps x_i to $1 + X_i$ and x_i^{-1} to $\sum_{k\in\mathbb{N}} (-1)^k X_i^k$, for each $i \in [\![1, n]\!]$.

Proposition 5.1. *For any $m \in \mathbb{N}^*$, any $(i_1, \cdots, i_m) \in [\![1, n]\!]^m$, any $j \in [\![1, n]\!]$ and any integer $k > m$, the coefficient $\mu^{(4)}_{i_1, \cdots, i_m j}(L)$ of $X_{i_1} \cdots X_{i_m}$ in the Magnus expansion $E(\lambda^k_j)$ is a well-defined invariant of T.*

Definition 5.2. The coefficient $\mu^{(4)}_{i_1\ldots i_k}$ of $X_{i_1} \cdots X_{i_m}$ in $E(\lambda^k_j)$ is called a *Milnor $\mu^{(4)}$–invariant* of length $m + 1$.

Here, the (4)-exponent refers to the 4–dimensional nature of ribbon tubes.

For each non negative integer k, we define $rT_n(k)$ to be the submonoid of rT_n consisting of elements with vanishing Milnor $\mu^{(4)}$–invariants of length lower than k. We thus have a descending filtration of monoids $rT_n = rT_n(1) \supset rT_n(2) \supset \cdots \supset rT_n(k) \supset \cdots$, called the *Milnor filtration*.

Now, we generalize the construction given in Section 2.3.2. Indeed, for each non negative integer, a ribbon tube T induces an automorphism $(\iota_0)_k^{-1} \circ (\iota_1)_k$ of $F_n/\Gamma_k F_n$. Actually, this automorphism maps x_i to its conjugate by λ^k_i, for each $i \in [\![1, n]\!]$, so that we have defined a monoid homomorphism

$$A_k : rT_n \to \mathrm{Aut}_C\big(F_n/\Gamma_k F_n\big).$$

One can check that T is in $rT_n(k)$ if and only if λ^k_i is trivial modulo $\Gamma_k F_n$ for all $i \in [\![1, n]\!]$. So for all non negative integer k, we have $rT_n(k) = \mathrm{Ker}(A_k)$, and we can consider the map

$$\mu^{(4)}_{k+1} : rT_n(k) \to F_n/\Gamma_2 F_n \otimes \Gamma_k F_n/\Gamma_{k+1} F_n.$$

which maps T to the sum

$$\mu^{(4)}_{k+1}(T) := \sum_{i\in[\![1,n]\!]} x_i \otimes \lambda^{k+1}_i.$$

We call it *the universal length $k + 1$ Milnor invariant*. This map $\mu^{(4)}_{k+1}$ is strictly equivalent to the collection of all Milnor $\mu^{(4)}$–invariants of length $k + 1$, via the formula $\mu^{(4)}_{k+1}(L) := \sum_{i\in[\![1,n]\!]} x_i \otimes E_k(\lambda^{k+1}_i)$, where E_k denotes the composition of E with the projection onto the degree k part, that is

$$E_k(\lambda^{k+1}_j) := \sum_{i_1, \cdots, i_k \in [\![1,n]\!]^k} \mu^{(4)}_{i_1, \cdots, i_k j} X_{i_1} \cdots X_{i_k}.$$

Note that $E_k(\lambda^{k+1}_j)$ lives in the isomorphic image of $\Gamma_k F_n/\Gamma_{k+1} F_n$ in $\mathbb{Z}\langle\langle X_1, \cdots, X_n\rangle\rangle$.

5.2. Milnor invariants for welded string links

The above 4-dimensional version $\mu^{(4)}$ of Milnor invariants provides, through the Tube map, a natural and general extension μ^w of Milnor invariants to virtual/welded string links.

Definition 5.3. For any sequence I of (possibly repeating) indices in $[\![1, n]\!]$, *Milnor invariant μ_I^w of n-component welded string links is defined by*

$$\mu_I^w := \mu_I^{(4)} \circ \text{Tube}.$$

The construction of Milnor invariants for ribbon tubes given in Section 5.1 is completely parallel to the definition of Milnor invariants of classical string links (as given, *e.g.*, in [20, 22]). Moreover, the Tube map is well-behaved with respect to the fundamental group of the complement; not only does it induce an isomorphism of π_1's, but it also maps meridians to meridians and (preferred) longitudes to (preferred) longitudes, so that the Wirtinger presentations are in one–to–one correspondence (see Proposition 3.13 and the discussion preceding it). As a consequence, we have the following.

Theorem 5.4. *For any sequence I of (possibly repeating) indices in $[\![1, n]\!]$, the restriction of μ_I^w to classical string links coincides with Milnor invariant:*

$$\mu_I^w(L) = \mu_I(L) \text{ , for any classical string link } L.$$

5.3. Comparison with previous works

Let us now discuss previously existing (partial) virtual extensions of Milnor invariants.

The first virtual extension of Milnor invariants is due to Dye and Kauffman [13]. There, it is already noted that these are actually welded invariants. The authors, however, restricted themselves to link–homotopy invariants, and focussed on the link case. This means in particular that the Dye-Kauffman extensions are defined modulo a certain indeterminacy Δ. Their construction follows very closely Milnor's original work [37], using the virtual knot group, and actually only slightly deviates from [37] in the definition of the indeterminacy Δ: since the authors only seek a link–homotopy invariant, a simpler definition can indeed be used. However, the Dye-Kauffman extension does not always coincide with Milnor invariants for classical links, precisely because of this different choice of indeterminacy. Consider for example a 3-component link L obtained from the Borromean rings by inserting a positive clasp between the first and second components; according to Milnor's definition we have $\Delta(123) = 1$, but in Dye–Kauffman's definition we have $\Delta(123) = 0$, so that the invariants $\overline{\mu}_{123}$ of the link L do not agree in both definitions.

Recently, Kotorii gave in [31] another virtual extension of Milnor invariants, using Turaev's theory of nanowords [44] to describe and study virtual links. This

extension is thus purely combinatorial and, again, addresses the link case and is only valid for link–homotopy invariants. Although it indeed coincides with Milnor's original invariants when restricted to classical links, it seems quite challenging to us to generalize directly the Kotorii extension to all Milnor invariants. Indeed, the fact that Milnor invariants are indexed by sequences with no repetition is used to ensure invariance under Reidemeister 3 move (see [31, Propositions 6.4]).

The third virtual extension of Milnor invariants is due to Kravchenko and Polyak [32]. The authors actually provides Gauss diagram formula for μ-invariants, and their extension is therefore closest to the extension given above. The result of [32] holds for string links (*i.e.* provides integer–valued virtual extensions), but as the preceding ones it is only valid for Milnor link–homotopy invariants. Indeed, identifying the Gauss diagram invariants defined in [32] with Milnor invariants relies on Polyak's skein formula [38], which is only valid for sequences without repetitions.

It would be interesting to generalize the Kravchenko–Polyak extension to all sequences (or at least, to identify the resulting Gauss diagram formulas with Milnor invariants). This seem to be a non-obvious problem, in particular because the string link analogue of [37, Theorem 7] (which, roughly speaking, shows how Milnor's link–homotopy invariants suffice to generate all $\overline{\mu}$–invariants of links via cabling) does not hold in full generality, as outlined in [51, Section 3]. Actually, we expect that one way to address this question would be by identifying the extension of Kravchenko–Polyak's Gauss diagram formulas with the topologically–defined use extension μ^w of the present paper; this identification is at least clear in the case of non-repeated indices.

References

[1] E. ARTIN, *Zur Isotopie zweidimensionaler Flächen im* $\mathbb{R}^4$, Abh. Math. Semin. Univ. Hamb. **4** (1925), 174–177.

[2] B. AUDOUX, *On the welded Tube map*, In: "Knot theory and its applications. ICTS program knot theory and its applications (KTH-2013)", IISER Mohali, India, December 10–20, 2013, Providence, RI: American Mathematical Society (AMS), 2016, 261–284.

[3] B. AUDOUX, P. BELLINGERI, J.-B. MEILHAN and E. WAGNER, *On usual, virtual and welded knotted objects up to homotopy*, J. Math. Soc. Japan **69** (2015), 1079–1097.

[4] D. BAR-NATAN and Z. DANCSO, *Finite type invariants of w-knotted objects I: w-knots and the Alexander polynomial*, Algebr. Geom. Topol. **16** (2016), 1063–1133.

[5] D. BAR-NATAN and Z. DANCSO, *Finite type invariants of w-knotted objects II: Tangles, foams and the Kashiwara-Vergne problem*, Math. Ann. (2015), to appear.

[6] V. G. BARDAKOV, *The virtual and universal braids*, Fund. Math. **184** (2004), 1–18.

[7] A. BARTELS and P. TEICHNER, *All two dimensions links are null homotopic*, Geom. Topol. **3** (1999), 235–252.

[8] T. E. BRENDLE and A. HATCHER, *Configuration spaces of rings and wickets*, Comment. Math. Helv. **88** (2013), 131–162.

[9] G. BURDE and H. ZIESCHANG, " Knots", de Gruyter Studies in Mathematics, Vol. 5, Walter de Gruyter & Co., Berlin, second edition, 2003.

[10] J. CARTER, S. KAMADA and M. SAITO, *Stable equivalence of knots on surfaces and virtual knot cobordisms*, J. Knot Theory Ramifications **11** (2002), 311–322.

[11] J. S. Carter and M. Saito, "Knotted Surfaces and their Diagrams", Mathematical Surveys and Monographs, Vol. 55, American Mathematical Society, Providence, RI, 1998.

[12] D. T. Cochran, *Link concordance invariants and homotopy theory*, Invent. Math. **90** (1987), 635–645.

[13] H. A. Dye and L. H. Kauffman, *Virtual homotopy*, J. Knot Theory Ramifications **19** (2010), 935–960.

[14] R. Fenn, R. Rimányi and C. Rourke, *The braid-permutation group*, Topology **36** (1997), 123–135.

[15] R. Fenn and D. Rolfsen, *Spheres may link homotopically in 4-space*, J. Lond. Math. Soc. **34** (1986), 177–184.

[16] T. Fiedler, "Gauss Diagram Invariants for Knots and Links", Mathematics and its Applications, Vol. 532, Kluwer Academic Publishers, Dordrecht, 2001.

[17] R. Fox and J. W. Milnor, *Singularities of 2-spheres in 4-space and cobordism of knots*, Osaka J. Math. **3** (1966), 257–267.

[18] R. H. Fox, *Some problems in knot theory*, In: "Topology of 3-manifolds and Related Topics (Proc. The Univ. of Georgia Institute, 1961)", Prentice-Hall, Englewood Cliffs, N.J., 1962, 168–176

[19] M. Goussarov, M. Polyak and O. Viro, *Finite type invariants of virtual and classical knots*, Topology **39** (2000), 1045–1168.

[20] N. Habegger and X.-S. Lin, *The classification of links up to link-homotopy*, J. Amer. Math. Soc. **3** (1990), 389–419.

[21] N. Habegger and X.-S. Lin, *On link concordance and Milnor's $\overline{\mu}$ invariants*, Bull. Lond. Math. Soc. **30** (1998), 419–428.

[22] N. Habegger and G. Masbaum, *The Kontsevich integral and Milnor's invariants*, Topology **39** (2000), 1253–1289.

[23] N. Habegger and J.-B. Meilhan, *On the classification of links up to finite type*, In: "Topology and Physics", Nankai Tracts Math., Vol. 12, World Sci. Publ., Hackensack, NJ, 2008, 138–150.

[24] A. Ichimori and T. Kanenobu, *Ribbon torus knots presented by virtual knots with up to four crossings*, J. Knot Theory Ramifications **21** (2012), 1240005.

[25] T. Kanenobu and A. Shima, *Two filtrations of ribbon 2-knots*, In: "Proceedings of the First Joint Japan-Mexico Meeting in Topology (Morelia, 1999)", Vol. 121, 2002, 143–168.

[26] L. H. Kauffman, *Virtual knot theory*, European J. Combin. **20** (1999), 663–690.

[27] M. Kervaire and J. W. Milnor, *On 2-spheres in 4-manifolds*, Proc. Natl. Acad. Sci. USA **47** (1961), 1651–1657.

[28] M. A. Kervaire, *Les nœuds de dimensions supérieures*, Bull. Soc. Math. France **93** (1965), 225–271.

[29] P. Kirk, *Links homotopy with one codimension two component*, Trans. Amer. Math . Soc **319** (1990), 663–688.

[30] U. Koschorke, *A generalization of Milnor's μ-invariants to higher dimensional link maps*, Topology **36** (1997), 301–324.

[31] Y. Kotorii, *The Milnor $\overline{\mu}$ invariants and nanophrases*, J. Knot Theory Ramifications **22** (2013), 1250142.

[32] O. Kravchenko and M. Polyak, *Diassociative algebras and Milnor's invariants for tangles*, Lett. Math. Phys. **95** (2011), 297–316.

[33] G. Kuperberg, *What is a virtual link?*, Algebr. Geom. Topol. **3** (2003), 587–591.

[34] G.-S. Li, *An invariant of link homotopy in dimension four*, Topology **36** (1997), 881–897.

[35] W. S. Massey and D. Rolfsen, *Homotopy classification of higher dimensional links*, Indiana Univ. Math. J. **34** (1986), 375–391.

[36] J. Milnor, *Link groups*, Ann. of Math. **59** (1954), 177–195.

[37] J. Milnor, *Isotopy of links. Algebraic geometry and topology*, In: "A symposium in honor of S. Lefschetz", Princeton University Press, Princeton, N. J., 1957, 280–306.

[38] M. POLYAK, *Skein relations for Milnor's μ-invariants*, Algebr. Geom. Topol. **5** (2005), 1471–1479 (electronic).

[39] M. POLYAK and O. VIRO, *Gauss diagram formulas for Vassiliev invariants*, Internat. Math. Res. Notices **11** (1994), 445–453 (electronic).

[40] D. ROSEMAN, *Reidemeister-type moves for surfaces in four-dimensional space*, In: "Knot Theory (Warsaw, 1995)", Banach Center Publ., Vol. 42, Polish Acad. Sci., Warsaw, 1998, 347–380.

[41] S. SATOH, *Virtual knot presentation of ribbon torus-knots*, J. Knot Theory Ramifications **9** (2000), 531–542.

[42] J. STALLINGS, *Homology and central series of groups*, J. Algebra **2** (1965), 170–181.

[43] S. SUZUKI, *Knotting problems of 2-spheres in 4-sphere*, Math. Sem. Notes Kobe Univ. **4** (1976), 241–371.

[44] V. TURAEV, *Lectures on topology of words*, Jpn. J. Math. **2** (2007), 1–39.

[45] D. S. W. MAGNUS and A. KARRASS, *Combinatorial group theory*, Pure and Appl. Math., Vol. XIII, Interscience, NY, 1966.

[46] T. YAJIMA, *On the fundamental groups of knotted 2-manifolds in the 4-space*, J. Math. Osaka City Univ. **13** (1962), 63–71.

[47] T. YAJIMA, *On simply knotted spheres in R^4*, Osaka J. Math. **1** (1964), 133–152.

[48] T. YANAGAWA, *On ribbon 2-knots: The 3-manifold bounded by the 2-knots*, Osaka J. Math. **6** (1969), 447–464.

[49] T. YANAGAWA, *On ribbon 2-knots II: The second homotopy group of the complementary domain*, Osaka J. Math. **6** (1969), 465–473.

[50] T. YANAGAWA, *On ribbon 2-knots III: On the unknotting Ribbon 2-knots in S^4*, Osaka J. Math. **7** (1970), 165–172.

[51] A. YASUHARA, *Classification of string links up to self delta-moves and concordance*, Algebr. Geom. Topol. **9** (2009), 265–275.

Aix Marseille Université
I2M, UMR 7373
13453 Marseille, France
benjamin.audoux@univ-amu.fr

Université de Caen
LMNO UMR 6139
14032 Caen, France
paolo.bellingeri@unicaen.fr

Université Grenoble Alpes
IF
38000 Grenoble, France
jean-baptiste.meilhan@ujf-grenoble.fr

IMB UMR5584, CNRS
Univ. Bourgogne Franche-Comté
F-21000 Dijon, France
emmanuel.wagner@u-bourgogne.fr