



HAL
open science

Virtual knot theory on a group

Arnaud Mortier

► **To cite this version:**

| Arnaud Mortier. Virtual knot theory on a group. 2014. hal-00958872

HAL Id: hal-00958872

<https://hal.science/hal-00958872>

Preprint submitted on 13 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Virtual knot theory on a group

Arnaud Mortier
arno.mortier@laposte.net

March 13, 2014

Abstract

Given a group endowed with a $\mathbb{Z}/2$ -valued morphism we associate a Gauss diagram theory, and show that for a particular choice of the group these diagrams encode faithfully virtual knots on a given arbitrary surface. This theory contains all of the earlier attempts to decorate Gauss diagrams, in a way that is made precise via *symmetry-preserving maps*. These maps become crucial when one makes use of decorated Gauss diagrams to describe finite-type invariants. In particular they allow us to generalize Grishanov-Vassiliev's formulas and to show that they define invariants of virtual knots.

Contents

1 Preliminary: classical Gauss diagrams and their Reidemeister moves	2
2 Knot and virtual knot diagrams on an arbitrary surface	4
2.1 Thickenings of surfaces	4
2.2 Diagram isotopies and detour moves	5
3 Virtual knot theory on a weighted group	6
3.1 General settings and the main theorem	7
3.1.1 About the orbits of w -moves	11
3.2 Abelian Gauss diagrams	11
3.3 Homological formulas	15
3.3.1 The energy formula	15
3.3.2 The torsion formula	16
4 Finite-type invariants	19
4.1 General algebraic settings	20
4.1.1 The Polyak algebra	20
4.1.2 The symmetry-preserving injections	21
4.1.3 Arrow diagrams and homogeneous invariants	23
4.1.4 Based and degenerate diagrams	25
4.2 Invariance criteria	27
4.2.1 Invariance criterion for w -orbits	28
4.3 Examples and applications	30
4.3.1 Grishanov-Vassiliev's planar chain invariants	30
4.3.2 There is a Whitney index for non nullhomotopic virtual knots	32

Gauss diagrams were introduced in knot theory as a means of representing knots and their finite-type invariants [20, 9], allowing compactification and generalization of formulas due to J.Lannes [14]. Since then, several generalizations have been attempted to adapt them to knot theory in thickened surfaces by decorating them with topological information [7, 11, 17].

Our goal is to construct a unifying “father” framework, and to describe how to get down from there to other versions with less data.

First we define and study (virtual) knot diagrams on an arbitrary surface Σ : these are tetravalent graphs embedded in Σ , some of whose double points (the “real” ones) are pushed and desingularized into a real line bundle over Σ . Defining Gauss diagrams requires a global notion for the

branches at a real crossing to be one “over” the other, and a global notion of writhe of a crossing. It is shown that these notions can be defined simultaneously if and only if Σ is orientable. If it is not, we sacrifice the globality of one property, and take into account its monodromy. It is shown that when the total space of the bundle is orientable, the writhes are globally defined and the monodromy of the “over/under” datum is the first Stiefel-Whitney class of the tangent bundle to Σ , $w_1(\Sigma)$.

In Section 3 is given a definition of Gauss diagrams decorated by elements of a fixed group π , subject to usual Reidemeister moves, and to additional “conjugacy moves”, depending on a fixed group homomorphism $w : \pi \rightarrow \mathbb{F}_2$. It is shown that when there is a surface Σ such that $\pi = \pi(\Sigma)$ and $w_1 = w_1(\Sigma)$, then there is a 1 – 1 correspondence between Gauss diagrams and virtual knot diagrams, that induces a correspondence between the equivalence classes (virtual knot types) on both sides.

A lighter kind of Gauss diagrams, called *abelian*, is defined in Subsection 3.2 following the idea of T.Fiedler’s $H_1(\Sigma)$ -decorated diagrams ([7]) and shown to be equivalent to the above when π is abelian and w is trivial. The little drawback of this version is that it becomes more difficult to compute the homological decoration of an arbitrary loop. Two formulas are presented in 3.3 to sort this out, involving quite unexpected combinatorial tools.

Finally, we describe invariance criteria for the analog of Goussarov-Polyak-Viro’s invariants [10] in this framework. As an application, we obtain a generalization of Grishanov-Vassiliev formulas [11], and a notion of Whitney index for virtual knots whose underlying immersed curve is non nullhomotopic.

Acknowledgements

I wish to thank Christian Blanchet who invited me at the Institut de Mathématiques de Jussieu where this work was done. This work has benefited from discussions with Michael Polyak, Micah Chrisman, and Thomas Fiedler.

1 Preliminary: classical Gauss diagrams and their Reidemeister moves

Definition 1.1. A *classical Gauss diagram* is an equivalence class of an oriented circle in which a finite number of couples of points are linked by an abstract oriented arrow with a sign decoration, up to positive homeomorphism of the circle. A Gauss diagram with n arrows is said to be of *degree* n .

It may happen that one regards Gauss diagrams as topological objects (drawing loops on them, considering their first homology). In that case, one must beware of the fact that *the arrows do not topologically intersect* – that is what is meant by “abstract”. However, the fact that two arrows may *look like* they intersect is something combinatorially well-defined, and interesting for many purposes.

Fact: There is a natural way to associate a Gauss diagram with a knot diagram in the sphere S^2 , from which the knot diagram can be uniquely recovered. Fig.1 illustrates this fact.

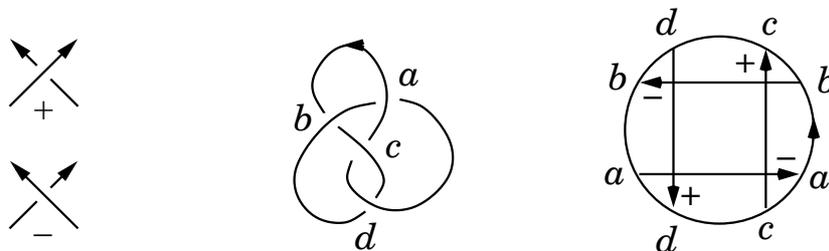


Figure 1: The writhe convention, a diagram of the figure eight knot, and its Gauss diagram – the letters are here only for the sake of clarity.

However, not every Gauss diagram actually comes from a knot diagram in that way. This observation has led to the development of virtual knot theory [13]: basically a virtual knot is a Gauss diagram which does not come from an actual knot. There is a knot-diagrammatic version of these, using virtual crossings subject to virtual Reidemeister moves - that can be thought of as a unique “detour move”. A detour move is naturally any move that leaves the underlying Gauss diagram unchanged.

Of course virtual knot diagrams are also subject to the usual Reidemeister moves, and these do change the face of the Gauss diagram. We call them R-moves for simplicity - and to make it clear whether knot diagrams or Gauss diagrams are considered. Here is a combinatorial description of R-moves.

R₁-moves

An R₁-move is the birth or death of an isolated arrow, as shown in Fig.2 (top-left). There is no restriction on the direction or the sign of the arrow.

R₂-moves

An R₂-move is the birth or death of a pair of arrows with different signs, whose heads are consecutive as well as their tails (Fig.2, top-right).

If one restricts oneself to Gauss diagrams that come from classical knot diagrams, then there is an additional condition as for the creating direction: indeed, two arcs in a knot diagram can be subject to a Reidemeister II move if and only if they *face each other*. In the virtual world, there is no such condition since any two arcs can be brought to face each other by detour moves.

It may be good to know that this condition can be read directly on the Gauss diagram: indeed, two arcs face each other in a knot diagram if one can join them by walking along the diagram and *turning to the left* at each time one meets a crossing. Thanks to the decorations of the arrows, it makes sense for a path in a Gauss diagram to *turn to the left*.

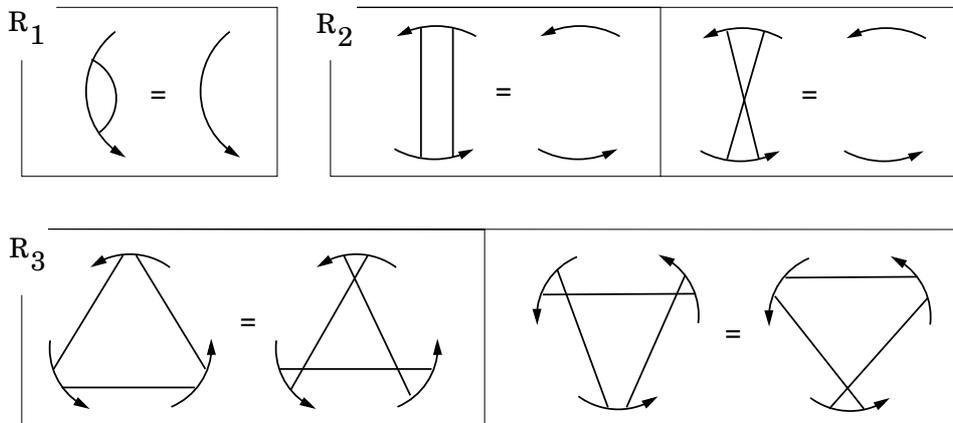


Figure 2: R-moves for Gauss diagrams (see above and below the rules for the decorations)

R₃-moves

Definition 1.2. In a classical Gauss diagram of degree n , the complementary of the arrows is made of $2n$ oriented components. These are called the *edges* of the diagram. In a diagram with no arrow, we still call the whole circle an edge.

Let e be an edge in a Gauss diagram, between two consecutive arrow ends that do not belong to the same arrow. Put

$$\eta(e) = \begin{cases} +1 & \text{if the arrows that bound } e \text{ cross each other} \\ -1 & \text{otherwise} \end{cases},$$

and let $\uparrow(e)$ be the number of arrowheads at the boundary of e . Then define

$$\varepsilon(e) = \eta(e) \cdot (-1)^{\uparrow(e)}.$$

Finally, define $w(e)$ as the product of the writhes of the two arrows at the boundary of e .

An R_3 -move is the simultaneous switch of the endpoints of three arrows as shown on Fig.2 (bottom), with the following conditions:

1. The value of $w(e)\varepsilon(e)$ should be the same for all three visible edges e . This ensures that the piece of diagram containing the three arrows can be represented in a knot-diagrammatic way without making use of virtual crossings.
2. The values of $\uparrow(e)$ should be pairwise different. This ensures that one of the arcs in the knot diagram version actually “goes over” the others.

Remark 1.3. From a simplicial viewpoint, the sign $w(e)\varepsilon(e)$ gives a natural co-orientation of the 1-codimensional strata corresponding to R_3 moves. This is exploited in [15] to construct finite-type 1-cocycles.

2 Knot and virtual knot diagrams on an arbitrary surface

The goal of this section is to examine when and how one can define a couple of equivalent theories “virtual knots – Gauss diagrams” that generalizes knot theory in an arbitrary 3-manifold M . What first appears is that a Gauss diagram depends on a projection; so it seems unavoidable to ask for the existence of a surface Σ (maybe with boundary, non orientable, or non compact), and a “nice” map $p : M \rightarrow \Sigma$. For the *over* and *under* branches at a crossing to be well-defined at least locally, the fibers of p need to be equipped with a total order: this leaves only the possibility of a real line bundle.

2.1 Thickenings of surfaces

Let us now split the discussion according to the two kinds of decorations that one would expect to find on a Gauss diagram: signs (local writhes), and orientation of the arrows.

Local writhes

For a knot in an arbitrary real line bundle, there are situations in which it is possible to switch over and under in a crossing by a mere *diagram isotopy*. For instance, in the non-trivial line bundle over the annulus $\mathbb{S}^1 \times \mathbb{R}$, a full rotation of the closure of the two-stranded elementary braid σ_1 turns it into the closure of σ_1^{-1} (Fig. 3).

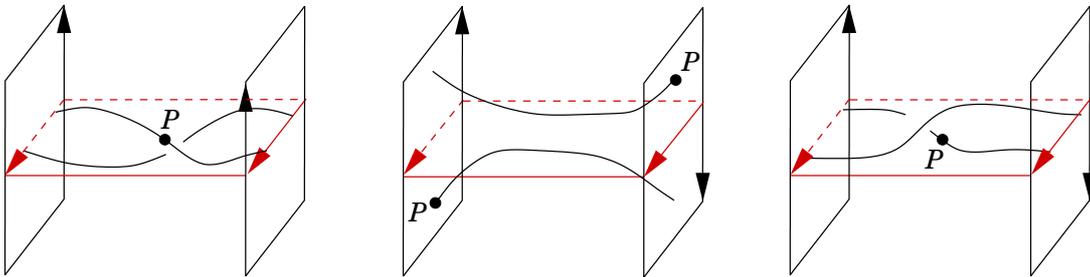


Figure 3: Non trivial line bundle over the annulus – as one reads from left to right, the knot moves towards the right of the picture.

Fig. 3 would be exactly the same (except for the gluing indications) if one considered the trivial line bundle over the Moebius strip. Note that this diagram would then represent a 2-component

link. In fact, it is possible to embed this picture in *any non-orientable total space* of a line bundle over a surface.

This phenomenon reveals the fact that in these cases, there is no way to define the local writhe of a crossing. However, according to [6] (Definition 1.), there is a well-defined writhe as soon as the total space of the bundle is orientable.

Definition 2.1. We call a *thickened surface* a real line bundle over a surface, whose total space is orientable.

Definition-Lemma 2.2. *If $M \rightarrow \Sigma$ is a thickened surface, then its first Stiefel-Whitney class coincides with that of the tangent bundle to Σ . This class induces a homomorphism $w_1(\Sigma) : \pi_1(\Sigma) \rightarrow \mathbb{F}_2$. The couple $(\pi_1(\Sigma), w_1(\Sigma))$ is called the weighted fundamental group of Σ . Note that in particular the thickening of Σ is the trivial bundle $\Sigma \times \mathbb{R}$ if and only if Σ is orientable.*

Arrow orientations

Note that the writhe of a crossing for a knot in $M \rightarrow \Sigma$ depends only on one choice, that of an orientation for M . The important thing is that this choice is *global*, so that it makes sense to compare the writhes of different crossings (they live in “the same” $\mathbb{Z}/2\mathbb{Z}$).

Similarly, for the orientation of the arrows in a Gauss diagram to simultaneously make sense, one needs a global definition of the over/under datum at the crossings; that is, the fibres of $M \rightarrow \Sigma$ should be simultaneously and consistently oriented. In other words, $M \rightarrow \Sigma$ should be the trivial line bundle.

According to our definition of a thickened surface, this happens only if the surface is orientable.

So it seems that one has a choice to make, either restricting one’s attention to orientable surfaces, or taking into account the monodromy of whatever is not globally defined. Additional conjugacy moves will be needed when one defines Gauss diagrams. The convention to consider only fibre bundles with an orientable total space is arbitrary, its only use is to reduce the number of monodromy morphisms to 1 instead of 2.

Virtual knot diagrams on an arbitrary surface

Fix an arbitrary surface Σ and denote its thickening by $M \rightarrow \Sigma$.

Definition 2.3. A *virtual knot diagram* on Σ is a generic immersion $\mathbb{S}^1 \rightarrow \Sigma$ whose every double point has been decorated

- either with the designation “virtual” (which is nothing but a name),
- or with a way to desingularize it locally into M , up to local isotopy.

These diagrams are subject to the usual Reidemeister moves, dictated by local isotopy in M , and to the virtual “detour” moves which are studied in the next section. As explained before, if one chooses an orientation for M , then the real crossings of such a diagram have a well-defined writhe.

2.2 Diagram isotopies and detour moves

Here by *knot diagram* we mean a virtual knot diagram on a fixed arbitrary surface Σ , as defined above. In this case a *diagram isotopy*, usually briefly denoted by $H : \text{Id} \rightarrow h$, is the datum of a diffeomorphism h of Σ together with an isotopy from Id_Σ to h . A *detour move* is a boundary-fixing homotopy of an arc that, before and after the homotopy, goes through only virtual crossings (such an arc is called *totally virtual*). Though both of these processes seem rather simple, it will be useful to understand how they interact.

Lemma 2.4. *A knot diagram obtained from another by a sequence of diagram isotopies alternating with detour moves may always be obtained by a single diagram isotopy followed by detour moves.*

Proof. It is enough to show that a detour move d followed by a diagram isotopy $\text{Id} \rightarrow h$ may be replaced with a diagram isotopy followed by a detour move (without changing the initial and final diagrams). The initial diagram is denoted by D .

Call α the totally virtual arc that is moved by the detour move. By definition, $d(\alpha)$ is boundary-fixing homotopic to α , and is totally virtual too. Thus, $h(d(\alpha))$ and $h(\alpha)$ are totally virtual and boundary-fixing homotopic to each other. Since $h(d(D))$ and $h(D)$ differ only by these two arcs, it follows that there is a detour move taking $h(D)$ to $h(d(D))$. \square

Now an interesting question about diagram isotopies is when two of them lead to diagrams that are equivalent *under detour moves*. Here is a quite useful sufficient condition.

Definition 2.5. Let X and Y be two finite subsets of Σ with the same (positive) cardinality n . A *generalized braid* in $\Sigma \times [0, 1]$ based on the sets X and Y is an embedding β of a disjoint union of segments, such that $\text{Im } \beta \cap (\Sigma \times \{t\})$ has cardinality n for each t , coincides with X at $t = 0$ and with Y at $t = 1$.

Let D be a knot diagram and H a diagram isotopy. Let $p_1 \in P_1, \dots, p_n \in P_n$ denote little neighborhoods of the real crossings of D , and set $\mathcal{P} = \cup P_i$. Then, $\coprod H(p_i, \cdot)$ defines a generalized braid ${}^H\beta$ in $\Sigma \times [0, 1]$ with n strands based on the sets $\{p_1, \dots, p_n\}$ and $\{h(p_1), \dots, h(p_n)\}$. The strand of a braid β that intersects $\Sigma \times \{0\}$ at p_i is denoted by β_i .

Proposition 2.6. *Let D and H be as above. Then, up to detour moves, $h(D)$ only depends on D and the boundary fixing homotopy class of ${}^H\beta$.*

Proof. Let γ be a maximal smooth arc of D outside \mathcal{P} (thus totally virtual). It begins at some P_i and ends at some P_j (of course it may happen that $j = i$). Using little arcs inside of P_i and P_j to join the endpoints of γ with p_i and p_j , one obtains an oriented path ${}^{H\beta_i^{-1}}\gamma {}^{H\beta_j}$.

The obvious retraction of $\Sigma \times [0, 1]$ onto $\Sigma \times \{1\}$ induces a map

$$\pi_1(\Sigma \times [0, 1], h(\mathcal{P}) \times \{1\}) \longrightarrow \pi_1(\Sigma, h(\mathcal{P}))$$

that sends the class $[{}^{H\beta_i^{-1}}\gamma {}^{H\beta_j}]$ to $[h(\gamma)]$. Since the former class is unchanged under boundary-fixing homotopy of γ and ${}^H\beta$, so is the latter, which proves the result. \square

This proposition states that the only relevant datum in a diagram isotopy of a virtual knot is the path followed by the real crossings along the isotopy, *up to homotopy*: the entanglement of these paths with each other or themselves does not matter. It follows that the crossings may be moved one at a time:

Corollary 2.7. *Let D be a knot diagram with its real crossings numbered from 1 to n , and let $H : \text{Id} \rightarrow h$ be a diagram isotopy. Then there is a sequence of diagram isotopies H_1, \dots, H_n , such that $h_n \dots h_1(D)$ coincides with $h(D)$ up to detour moves, and such that H_i is the identity on a neighborhood of each real crossing but the i -th one.*

Remark 2.8. It is to be understood that the i -th crossing of $h_k \dots h_1(D)$ is $h_k \dots h_1(p_i)$.

Proof. Any generalized braid is (boundary-fixing) homotopic to a braid $\beta \subset \Sigma \times [0, 1]$ such that the i -th strand is vertical before the time $\frac{i-1}{n}$ and vertical again after the time $\frac{i}{n}$. Take such a braid β that is homotopic to ${}^H\beta$. Any diagram isotopy H' such that $\beta = {}^{H'}\beta$ factorizes into a product $H_n \dots H_1$ satisfying the last required condition. The fact that $h_n \dots h_1(D)$ and $h(D)$ coincide up to detour moves is a consequence of Proposition 2.6. \square

3 Virtual knot theory on a weighted group

In this section, we define a new Gauss diagram theory, that depends on an arbitrary group π and a homomorphism $w : \pi \rightarrow \mathbb{F}_2 \simeq \mathbb{Z}/2\mathbb{Z}$. These two data together are called a *weighted group*. When (π, w) is the weighted fundamental group of a surface (Definition 2.2), this theory encodes, fully and faithfully, virtual knot diagrams on that surface (Definition 2.3).

3.1 General settings and the main theorem

Definition 3.1. Let π be an arbitrary group and w a homomorphism from π to \mathbb{F}_2 . A *Gauss diagram on π* is a classical Gauss diagram decorated with

- an element of π on each edge if the diagram has at least one arrow.
- a single element of π *up to conjugacy* if the diagram is empty.

Such diagrams are subject to the usual types of R-moves, plus an additional *conjugacy move*, or *w-move* – the dependence on w arises only there. An equivalence class modulo all these moves is called a *virtual knot type on (π, w)* .

A *subdiagram* of a Gauss diagram on π is the result of removing some of its arrows. Removing an arrow involves a merging of its (2, 3, or 4) adjacent edges, and each edge resulting from this merging should be marked with the product in π of the former markings. If all the arrows have been removed, this product is not well-defined, but its conjugacy class is.

The notion of subdiagrams is useful to construct finite-type invariants (see Section 4), but it already allows explicit understanding of

1. The distinction between empty and non empty diagrams in the definition above.
2. The “merge \rightsquigarrow multiply” principle, which is omnipresent, in particular in R-moves.

An **R₁-move** is the local addition or removal of an isolated arrow, surrounding an edge marked with the unit $1 \in \pi$. The markings of the affected edges must satisfy the rule indicated on Fig.4 (top-left). There are no conditions on the decorations of the arrows.

Exceptional case: If the isolated arrow is the only one in the diagram on the left, then the markings a and b on the picture actually correspond to the same edge, and the diagram on the right, with no arrow, must be decorated by $[a]$, the conjugacy class of a .

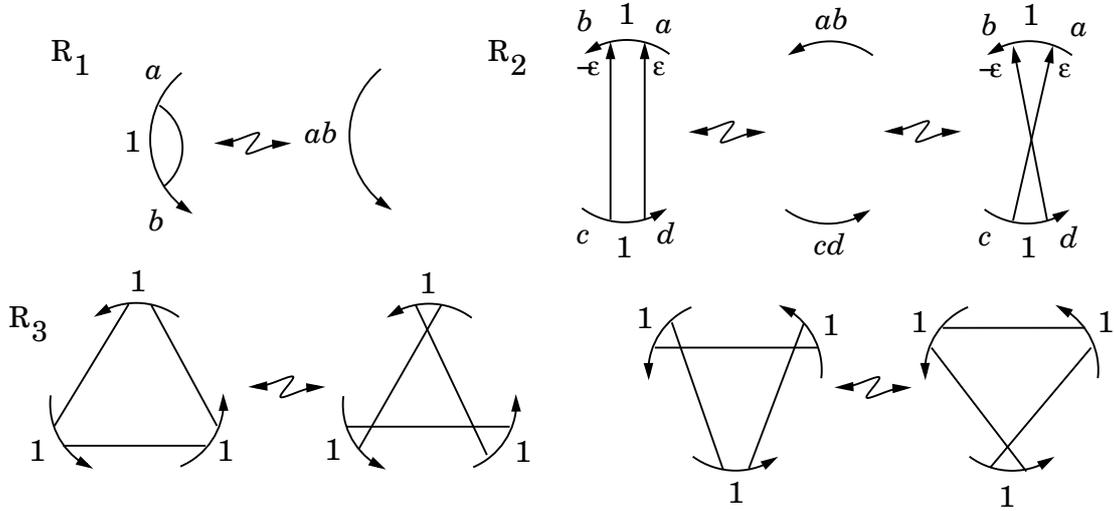


Figure 4: The R-moves for Gauss diagrams on a group – the exceptional cases and the rules for the missing decorations are made precise in Definition 3.1.

An **R₂-move** is the addition or removal of two arrows with opposite writhe and matching orientations as shown on Fig.4 (top-right). The surrounded edges must be decorated with 1 , and the “merge \rightsquigarrow multiply” rule should be satisfied.

Exceptional case of type 1: If the markings a and d (resp. b and c) correspond to the same edge, then the resulting marking shall be cab (resp. abd).

Exceptional case of type 2: If the middle diagram contains no arrow at all, i.e. a and d match and so do b and c , then the (only) marking of the middle diagram shall be $[ab]$.

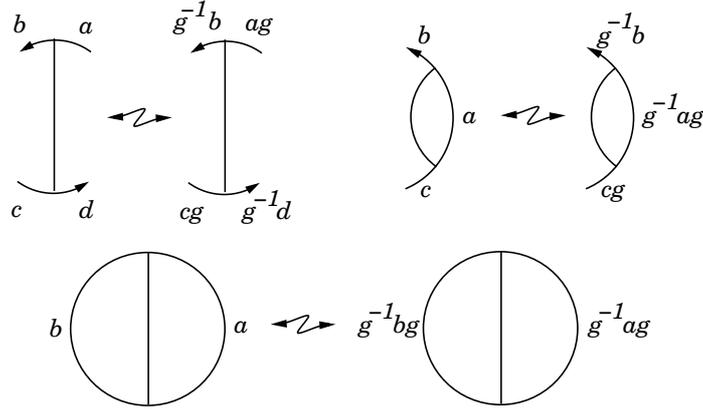


Figure 5: The general conjugacy move (top-left) and its two exceptional cases – in every case the orientation of the arrow switches if and only if $w(g) = -1$.

An R_3 -move may be of the two types shown on Fig.4 (bottom left and right). The surrounded edges must be decorated by 1, the value of $w(\cdot)\varepsilon(\cdot)$ must be the same for all three of them, and the values of $\uparrow(\cdot)$ must be pairwise distinct (see Definition 1.2).

A **conjugacy move** depends on an element $g \in \pi$. It changes the markings of the adjacent edges to an arbitrary arrow as indicated on Fig.5. Besides, if $w(g) = -1$ then the orientation of the arrow is reversed – though its sign remains the same.

Remark 3.2. By composing R-moves and w -moves, it is possible to perform *generalized moves*, which look like R-moves but depend on w . Fig.6 shows some of them.

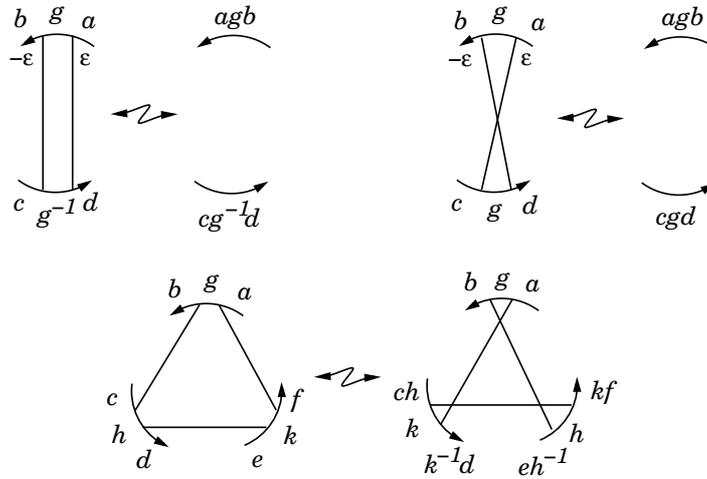


Figure 6: Some generalized moves – for the R_3 picture, it is assumed that $ghk = 1$. **Warning:** the rules for the arrow orientations here depend on the value of $w(g)$.

Theorem 3.3. *Let (Σ, x) be an arbitrary surface with a base point, and denote by (π, w) the weighted fundamental group of (Σ, x) (see Definition 2.2). There is a 1 – 1 correspondence Φ between Gauss diagrams on π up to R-moves and w -moves (i.e. virtual knot types on (π, w)), and virtual knot diagrams on Σ up to diagram isotopy, Reidemeister moves and detour moves (i.e. virtual knot types on Σ).*

Proof. Fix a subset X of Σ homeomorphic to a closed 2-dimensional disc and containing the base point x – so that $\pi = \pi_1(\Sigma, X)$. Also, X being contractible allows one to fix a trivialization of the thickening of Σ over X : this gives meaning to the locally *over* and *under* branches when a knot

diagram has a real crossing in X .

Construction of the bijection. Pick a knot diagram $D \in \Sigma$ and assume that every real crossing of D lies over X . Then D defines a Gauss diagram on π , denoted by $\varphi(D)$: the signs of the arrows are given by the writhes, their orientation is defined by the trivialization of $M \rightarrow \Sigma$ over X , and each edge is decorated by the class in π of the corresponding arc in D . This defines $\varphi(D)$ without ambiguity if D has at least one real crossing. If it does not, then define $\varphi(D)$ as a Gauss diagram without arrows, decorated with the conjugacy class corresponding to the free homotopy class of D . Finally, put

$$\Phi(D) := [\varphi(D)] \text{ mod R-moves and } w\text{-moves.}$$

Invariance of Φ under diagram isotopy and detour moves. It is clear from the definitions that $\varphi(D)$ is strictly unchanged under detour moves on D . Now assume that D_1 and D_2 are equivalent under *usual* diagram isotopy – that is, diagram isotopy that may take real crossings out of X for some time. By Corollary 2.7, it is enough to understand what happens for a diagram isotopy along which only one crossing goes out of X . In that case, $\varphi(D)$ is changed by a w -move performed on the arrow corresponding to that crossing, where the conjugating element g is the loop followed by the crossing along the isotopy. Indeed, since the first Stiefel-Whitney class of the thickening of Σ coincides with that of its tangent bundle, it follows that:

1. The orientation of the fibre (and thus the notions of “over” and “under”) is reversed along g if and only if $w(g) = -1$, which actually corresponds to the rule for arrow orientations in a w -move.
2. The orientation of the fibre over the crossing is reversed along g if and only if a given local orientation of Σ is reversed along g , so that the writhe of the crossing never changes.

Invariance of Φ under Reidemeister moves. Up to conjugacy by a diagram isotopy, it can always be assumed that a Reidemeister move happens inside X . In that case, at the level of $\varphi(D)$, it clearly corresponds to an R-move as described in Definition 3.1.

So far, Φ is a well-defined map from the set of virtual knot types on Σ to the set of virtual knot types on (π, w) .

Construction of an inverse map Ψ . If G is a Gauss diagram without arrows, then define $\psi(G)$ as the totally virtual knot with free homotopy class equal to the marking of G – it is well-defined up to detour moves. If G has arrows, then for each of them draw a crossing inside X with the required writhe, and then join these by totally virtual arcs with the required homotopy classes. The resulting diagram $\psi(G)$ is well-defined up to diagram isotopy and detour moves by this construction. In both cases, put

$$\Psi(D) := \text{virtual knot type of } \psi(D).$$

Let us prove that φ and ψ are inverse maps, so that Ψ will be the inverse of Φ as soon as it is invariant under R-moves and w -moves.

It is clear from the definitions that $\varphi \circ \psi$ coincides with the identity. It is also clear that $\psi \circ \varphi$ is the identity, up to detour moves, for *totally virtual* knot diagrams.

Now fix a knot diagram D with at least one real crossing (and all real crossings inside X). Recall that $\psi \circ \varphi(D)$ is defined up to diagram isotopy and detour moves, so fix a diagram D' in that class. There is a natural correspondence between the set of real crossings of D and those of D' , due to the fact that both identify by construction with the set of arrows of $\varphi(D)$. Pick a diagram isotopy h that takes each real crossing of D to meet its match in D' , *without leaving* X . Then clearly $\varphi(h(D)) = \varphi(D)$, and because $\varphi \circ \psi$ is the identity, one gets

$$\varphi(h(D)) = \varphi(D'). \tag{1}$$

The choice of h ensures that $h(D)$ and D' differ only by totally virtual arcs, and (1) implies that each of these, in $h(D)$, has the same class in $\pi_1(\Sigma, X)$ as its match in D' , which means by definition

that $h(D)$ and D' are equivalent up to detour moves. Thus $\psi \circ \varphi$ is the identity up to diagram isotopy and detour moves.

Invariance of Ψ under R-moves. Let us treat only the case of R_2 -moves, which contains all the ideas. Let G_1 and G_2 differ by an R_2 -move, and assume that G_1 is the one with more arrows. By appropriate diagram isotopy and detour moves *inside* X , performed on $\psi(G_1)$, it is possible to make the two concerned crossings “face” each other, as in Fig.7 (left). The paths α_1 and α_2 from this picture are totally virtual and trivial in $\pi_1(\Sigma, X)$, thus $\psi(G_1)$ is equivalent to the second diagram of Fig.7 up to detour moves. The fact that at this point, an R-II move is actually possible is a consequence of (in fact equivalent to) the combinatorial conditions defining the R-moves. Denote by D the third diagram of the picture. The “merge \rightsquigarrow multiply” principle that rules R_2 -moves implies that $\varphi(D) = G_2$, so that

$$\psi(G_1) \sim D \sim \psi \circ \varphi(D) = \psi(G_2), \quad (2)$$

where \sim is the equivalence under diagram isotopy, detour moves and Reidemeister moves. It follows that $\psi(G_1)$ and $\psi(G_2)$ have the same knot type.

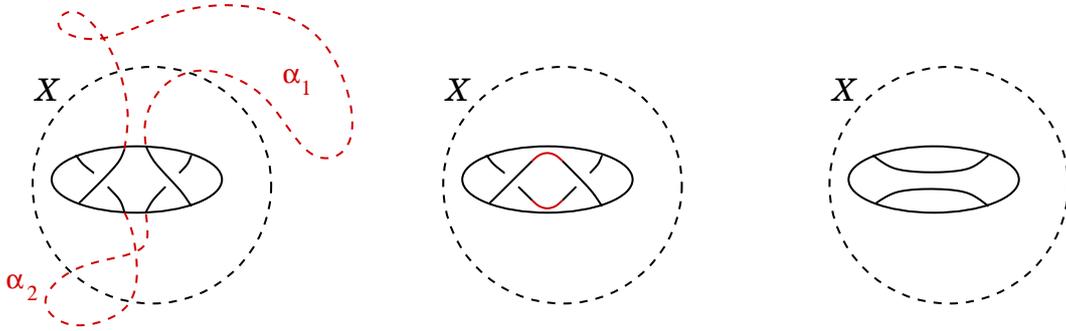


Figure 7: R_2 -moves actually correspond to Reidemeister moves

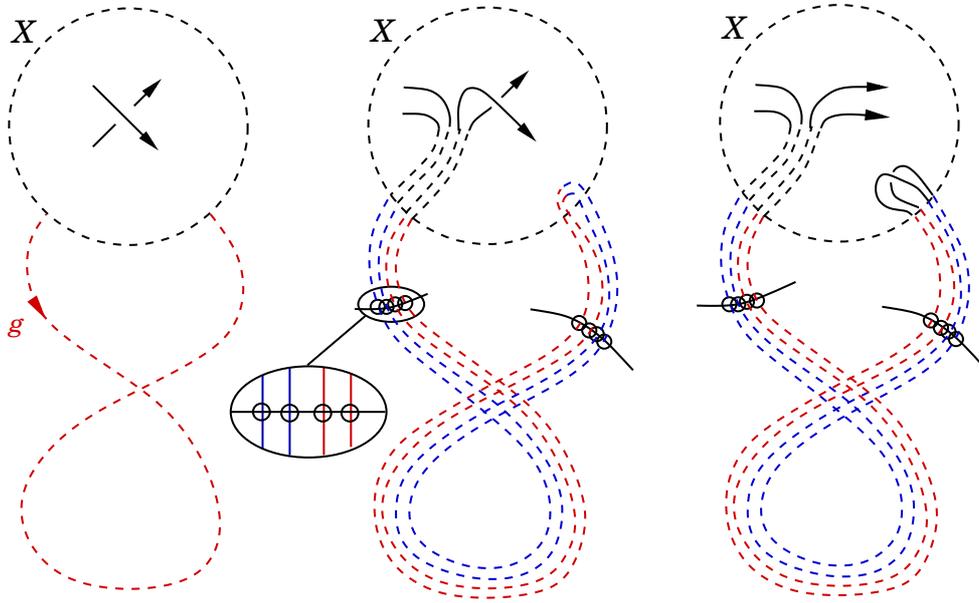


Figure 8: Performing a w -move – the railway trick

Invariance of Ψ under w -moves. Let G_1 and G_2 differ by a w -move on $g \in \pi$. Call c the corresponding crossing on the diagram $\psi(G_1)$. Then, pick two little arcs right before c , one on each branch, and make them follow g by a detour move. At the end, one shall see a totally virtual 4-lane

railway as pictured on Fig.8 (middle): the strands are made parallel, *i.e.* any (virtual) crossing met by either of them is part of a larger picture as indicated by the zoom. This ensures that, using the mixed version of Reidemeister III moves, one can slide the real crossing all along the red part of the railway, ending with the diagram on the right of the picture – let us call it D . The conclusion is identical to that for R-moves: again $\varphi(D) = G_2$ and (2) holds, whence $\psi(G_1)$ and $\psi(G_2)$ have the same knot type. □

3.1.1 About the orbits of w -moves

It could feel natural to try to get rid of w -moves by understanding their orbits in a synthetic combinatorial way. This is what is done in Section 3.2 in the particular case of an abelian group π endowed with the trivial homomorphism $\pi \rightarrow \mathbb{F}_2$.

In general, for a Gauss diagram on π , G , denote by $h_1(G)$ the *set* of free homotopy classes of loops in the underlying topological space of G (it is the set of conjugacy classes in a free group on $\deg(G) + 1$ generators). Also, denote by $h_1(\pi)$ the set of conjugacy classes in π . Then the π -markings of G define a map

$$F_G : h_1(G) \rightarrow h_1(\pi).$$

Observe that the map $G \mapsto F_G$ is invariant under w -moves. This raises a number of questions that amount to technical group theoretic problems, and which will not be answered here (G^w denotes the orbit of G under w -moves):

1. Is the map $G^w \mapsto F_G$ injective?
2. If the answer to 1. is yes, then is G^w determined by a finite number of values of F_G , for instance its values on the free homotopy classes of *simple loops*?
3. Is it possible to detect in a simple manner what maps $h_1(G) \rightarrow h_1(\pi)$ lie in the image of $G^w \mapsto F_G$?

Remark 3.4. Gauss diagrams with decorations in $h_1(\Sigma)$ can be met for example in [11], where they are used to construct knot invariants in a thickened oriented surface Σ – see also Section 4.3. If the answer to Question 1. above is no, then such invariants, which factor through F_G , stand no chance to be complete.

Remark 3.5. Even for diagrams with only one arrow, it still does not seem easy to answer the “simple loop” version of Question 2. Given x, y, h, k in a finite type free group, is it true that

$$h x h^{-1} k y k^{-1} = x y \implies \exists l, \begin{cases} h x h^{-1} = l x l^{-1} \\ k y k^{-1} = l y l^{-1} \end{cases} ?$$

Let us end with an example that shows that the values of F_G on the (finite) set of simple loops running along at most one arrow is not enough (cf. Question 2.). Fig.9 shows a Gauss diagram with such decorations – $\{a, b\}$ is a set of generators for the free group $\pi_1(\Sigma) \simeq \mathbb{F}(a, b)$, where Σ is a 2-punctured disc. These particular values of F_G do not determine the free homotopy class of the red loop γ , as it is shown in Fig.10.

In fact, these two virtual knots are even distinguished by Vassiliev-Grishanov’s planar chain invariants, which means they represent different virtual knot types.

3.2 Abelian Gauss diagrams

In this subsection, π is assumed to be abelian, and w_0 denotes the trivial homomorphism $\pi \rightarrow \mathbb{F}_2$. We describe a version of Gauss diagrams that carries as much information as the previously introduced virtual knot types on (π, w_0) , with two improvements:

- The diagrams are made of less data than in the general version.
- This version is free from conjugacy moves.

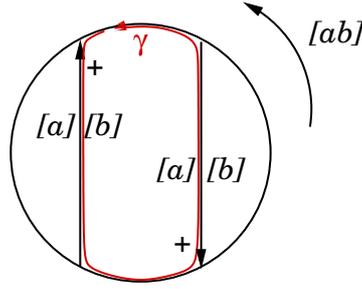


Figure 9: A Gauss diagram with h_1 -decorations that does not define a unique virtual knot

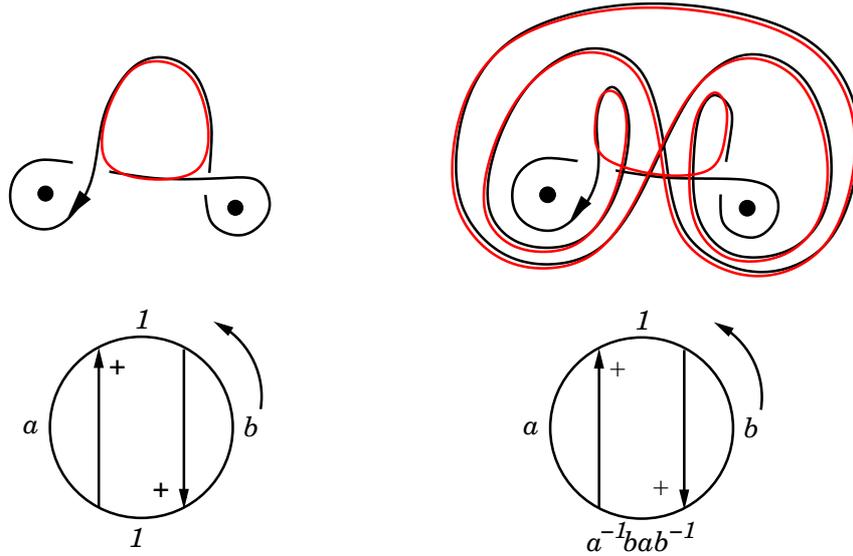


Figure 10: One red loop is trivial, while the other is a commutator

It is inspired from the decorated diagrams introduced by T. Fiedler to study combinatorial invariants for knots in thickened surfaces (see [7, 8] and also [17]).

We use the same notation G for a Gauss diagram and its underlying topological space, which has a 1-dimensional complex structure with edges and arrows as oriented 1-cells. $H_1(G)$ denotes its first integral homology group.

Definition-Lemma 3.6 (fundamental loops). *Let G be a classical Gauss diagram of degree n . There are exactly $n + 1$ simple loops in G respecting the local orientations of edges and arrows, and going along at most one arrow. They are called the fundamental loops of G and their homology classes form a basis of $H_1(G)$.*

Definition 3.7 (abelian Gauss diagram). Let π be an abelian group. An *abelian Gauss diagram on π* is a classical Gauss diagram G decorated with a group homomorphism $\mu : H_1(G) \rightarrow \pi$. It is usually represented by its values on the basis of fundamental loops, that is, one decoration in π for each arrow, and one for the base circle – that last one is called the *global marking* of G .

A Gauss diagram on π determines an abelian Gauss diagram as follows:

- The underlying classical Gauss diagram is the same.
- Each fundamental loop is decorated by the sum of the markings of the edges that it meets (see Fig 11).

This defines an *abelianization map* ab .

Proposition 3.8. *The map ab induces a natural 1 – 1 correspondence between abelian Gauss diagrams on π and equivalence classes of Gauss diagrams on π up to w_0 -moves. Moreover, if*

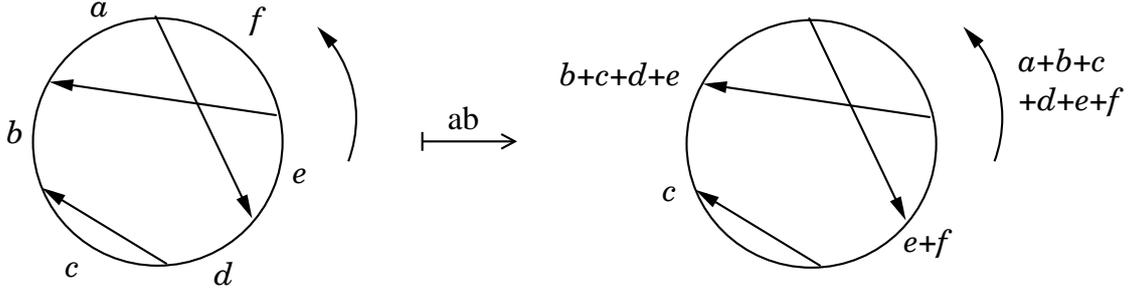


Figure 11: Abelianizing a Gauss diagram on an abelian group

$\pi = \pi_1(\Sigma)$ is the fundamental group of a surface, then these sets are in 1 – 1 correspondence with the set of virtual knot diagrams on Σ up to diagram isotopy and detour moves.

Proof. The proof of the last statement is contained in that of Theorem 3.3 – through the facts that ϕ and ψ are inverse maps up to detour moves and diagram isotopy, and that w -moves at the level of knot diagrams can be performed using only detour moves and diagram isotopies, by the railway trick (Fig.8).

As for the first statement, one easily sees that ab is invariant under w_0 -moves. We have to show that conversely, if $ab(G_1) = ab(G_2)$, then G_1 and G_2 are equivalent under w_0 -moves.

This is clear if G_1 has no arrows, since then $ab(G_1) = G_1$. Now proceed by induction. Since G_1 and G_2 have the same abelianization, they have in particular the same underlying classical Gauss diagram, and there is a natural correspondence between their arrows.

Case 1: No two arrows in G_1 cross each other. Then at least one arrow surrounds a single isolated edge on one side (as in an R_1 -move). Choose such an arrow α and remove it, as well as its match in G_2 . By induction, there is a sequence of w_0 -moves on the resulting diagram G'_1 that turns it into G'_2 . Since the arrows of G'_1 have a natural match in G_1 , those w_0 -moves make sense there, and take every marking of G_1 to be equal to its match in G_2 , except for those in the neighborhood of α . So we may assume that G_1 and G_2 only differ near α as in Fig.12. Since all the unseen markings coincide in G_1 and G_2 , and since $ab(G_1)$ and $ab(G_2)$ have the same global marking, it follows that

$$a + b + c = a' + b' + c'.$$

Thus a w_0 -move on α with conjugating element $g = a' - a$ turns G_1 into G_2 .



Figure 12: Notations for case 1

Case 2: There is at least one arrow α in G_1 that intersects another arrow. By the same process as in case 1, one may assume that G_1 and G_2 only differ near α – see Fig.13, where a , b , c and d actually correspond to pairwise distinct edges since α intersects an arrow. Again, since all the unseen markings coincide in G_1 and G_2 , one obtains

$$a + d = a' + d',$$

and

$$b + c = b' + c',$$

by considering the global marking, and the marking of α , in $\text{ab}(G_1)$ and $\text{ab}(G_2)$. Moreover, there is at least one arrow intersecting α : considering the marking of that arrow gives

$$a + b = a' + b'.$$

The last three equations may be written as

$$a' - a = b - b' = c' - c = d - d',$$

so that, again, a w_0 -move on α with conjugating element $g = a' - a$ turns G_1 into G_2 .



Figure 13: Notations for case 2

□

Remark 3.9. A different proof of this proposition was given in a draft paper, in the special case $\pi = \mathbb{Z}$ ([16], Proposition 2.2). As an exercise, one can show that this proof extends to the case of an arbitrary abelian group.

To make the picture complete, it only remains to understand R-moves in this context.

Definition 3.10 (obstruction loops). Within any local Reidemeister picture like those shown on Fig.2 featuring at least one arrow, there is exactly one (unoriented) simple loop. We call it the *obstruction loop*. Fig.14 shows typical examples.

Definition 3.11 (R-moves). A move from Fig.2 is likely to define an R-move only if the obstruction loop lies in the kernel of the decorating map $H_1(G) \rightarrow \pi$ (which makes sense even though the loop is unoriented). Under that assumption, the *R-moves for abelian Gauss diagrams* are defined by the usual conditions:

- $i = 1$. No additional condition.
- $i = 2$. The arrows head to the same edge, and have opposite signs.
- $i = 3$. The value of $w(e)\varepsilon(e)$ is the same for all three visible edges e , and the values of $\uparrow(e)$ are pairwise different (see Definition 1.2).

Theorem 3.12. *The map ab induces a natural 1 – 1 correspondence between equivalence classes of abelian Gauss diagrams on π up to R-moves and virtual knot types on (π, w_0) .*

Proof. ab clearly maps an R-move in the non commutative sense to an R-move in the abelian sense. Conversely, if $\text{ab}(G_1)$ and $\text{ab}(G_2)$ differ from an (abelian) R-move, then the vanishing homological obstruction implies that G_1 and G_2 are in a position to perform a “generalized R-move” like the examples pictured on Fig.6. □

Theorems 3.3 and 3.12 together imply the following

Corollary 3.13. *If Σ is an orientable surface with abelian fundamental group, then there is a 1 – 1 correspondence between abelian Gauss diagrams on $\pi_1(\Sigma)$ up to R-moves, and virtual knot types on Σ .*

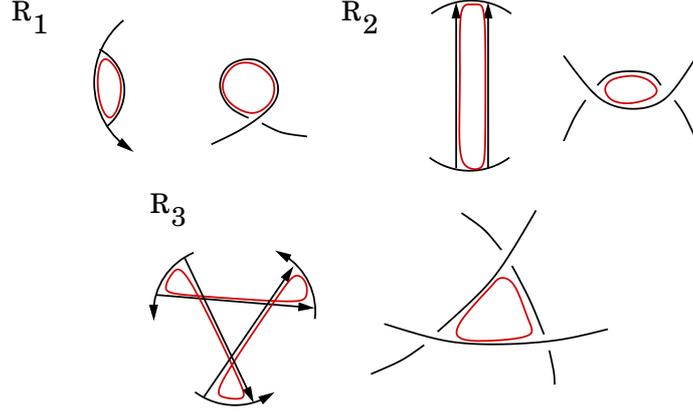


Figure 14: Homological obstruction to R-moves

3.3 Homological formulas

It may seem not easy to compute an arbitrary value of the linear map decorating an abelian Gauss diagram, given only its values on the fundamental loops. To end this section, we give two formulas to fill this gap, by understanding the coordinates of an arbitrary loop in the basis of fundamental loops.

3.3.1 The energy formula

Fix an abelian Gauss diagram G . Observe that as a cellular complex, G has no 2-cells, thus every 1-homology class has a unique set of “coordinates” along the family of edges and arrows. For each 1-cell c (which may be an arrow or an edge), we denote by $\langle \cdot, c \rangle : H_1(G) \rightarrow \mathbb{Z}$ the coordinate function along c . It is a group homomorphism.

Let us denote by $[A] \in H_1(G)$ the class of the fundamental loop associated with an arrow A (Fig.15 left).

Definition-Lemma 3.14 (Energy of a loop). *Fix an edge e in G , and a class $\gamma \in H_1(G)$. The value of*

$$E_e(\gamma) = \langle \gamma, e \rangle - \sum_{\langle [A], e \rangle = 1} \langle \gamma, A \rangle \quad (3)$$

is independent of e . This defines a group homomorphism $E : H_1(G) \rightarrow \mathbb{Z}$.

Proof. Let us compare the values of $E_e(\gamma)$ for an edge e and the edge e' right after it. e and e' are separated by a vertex P , which is the endpoint of an arrow A . There are two possible situations (Fig.15):

1. P is the tail of A . Then $\langle [A], e \rangle = 1$ and $\langle [A], e' \rangle = 0$, so that

$$E_e(\gamma) - E_{e'}(\gamma) = \langle \gamma, e \rangle - \langle \gamma, A \rangle - \langle \gamma, e' \rangle.$$

2. P is the head of A . Then $\langle [A], e \rangle = 0$ and $\langle [A], e' \rangle = 1$, so that

$$E_e(\gamma) - E_{e'}(\gamma) = \langle \gamma, e \rangle + \langle \gamma, A \rangle - \langle \gamma, e' \rangle.$$

In both cases, $E_e(\gamma) - E_{e'}(\gamma)$ is equal to $\langle \partial\gamma, P \rangle$, which is 0 since γ is a cycle. □

Theorem 3.15. *For any $\gamma \in H_1(G)$, one has the decomposition*

$$\gamma = \sum_A \langle \gamma, A \rangle [A] + E(\gamma) [K]. \quad (4)$$

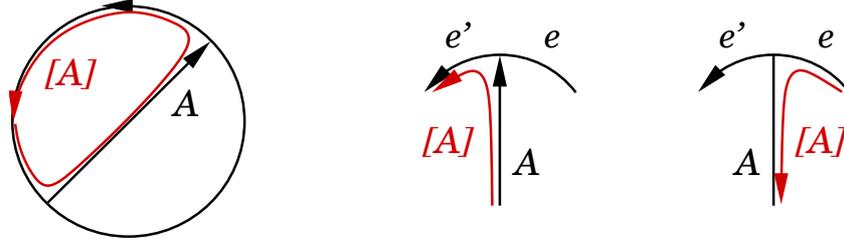


Figure 15: The fundamental loop of an arrow and the two cases in the proof of Lemma 3.14

Proof. This formula is an identity between two group homomorphisms, so it suffices to check it on the basis of fundamental loops, which is immediate. \square

Remark 3.16. The existence of a map E such that Theorem 3.15 holds was clear, since for each arrow A considered as a 1-cell, $[A]$ is the only fundamental loop that involves A . With that in mind, one may read into (3) as follows: $E(\gamma)$ counts the (algebraic) number of times that γ goes through an edge, minus the number of those times that are already taken care of by the fundamental loops of the arrows. This number has to be the same for all edges, so that one recovers a multiple of $[K]$.

3.3.2 The torsion formula

Looking at (4) and Fig.15, one may feel that it would be more natural to have $[K] - [A]$ involved in the formula, instead of $[A]$, for all arrows A such that $\langle \gamma, A \rangle$ is negative – that is, when γ runs along A with the wrong orientation more often than not. The formula then becomes

$$\gamma = \sum_{\langle \gamma, A \rangle > 0} \langle \gamma, A \rangle [A] + \sum_{\langle \gamma, A \rangle < 0} \langle \gamma, A \rangle ([A] - [K]) - \mathcal{T}(\gamma) [K], \quad (5)$$

where

$$-\mathcal{T}(\gamma) = E(\gamma) + \sum_{\langle \gamma, A \rangle < 0} \langle \gamma, A \rangle. \quad (6)$$

Definition 3.17. $\mathcal{T}(\gamma)$ is called the *torsion* of γ .

How is (5) different from (4)?

\ominus On the negative side, unlike the energy, \mathcal{T} is not a group homomorphism. But it actually behaves almost like one:

Lemma 3.18. *Let γ_1 and γ_2 be two homology classes such that*

$$\forall A, \langle \gamma_1, A \rangle \langle \gamma_2, A \rangle \geq 0.$$

Then

$$\mathcal{T}(\gamma_1 + \gamma_2) = \mathcal{T}(\gamma_1) + \mathcal{T}(\gamma_2).$$

Proof. It follows from the definition and the fact that $E(\gamma)$ is a homomorphism. \square

\oplus On the positive side:

Lemma 3.19. *The torsion of a loop in a Gauss diagram G does not depend on the orientations of the arrows of G .*

Proof. By expanding the defining formula,

$$\mathcal{T}(\gamma) = -\langle \gamma, e \rangle + \sum_{\substack{\langle \gamma, A \rangle < 0 \\ \langle [A], e \rangle = 0}} \langle \gamma, A \rangle - \sum_{\substack{\langle \gamma, A \rangle > 0 \\ \langle [A], e \rangle = 1}} \langle \gamma, A \rangle,$$

one sees that reversing an arrow makes its contribution (if non zero) switch from one sum to the other, while $\langle \gamma, A \rangle$ also changes signs. \square

This lemma allows one to expect that $\mathcal{T}(\gamma)$ should admit a very simple combinatorial interpretation. It actually does, but only for a certain family of loops – the ERS loops defined below. Fortunately enough, this family happens to positively generate $H_1(G)$, which allows one to compute the torsion of any loop by using Lemma 3.18.

Definition 3.20. The notation γ is used for loops as well as 1-homology classes. A homology class $\gamma \in H_1(G)$ is said to be

- *ER* (for “edge-respecting”), if for every edge e , $\langle \gamma, e \rangle \geq 0$.
- *simple* if it is the class of a simple (injective) loop, that is, $|\langle \gamma, c \rangle| \leq 1$ for every 1-cell c (edge or arrow).
- *ERS* if it is ER and simple.
- *proper* if it runs along at least one arrow.

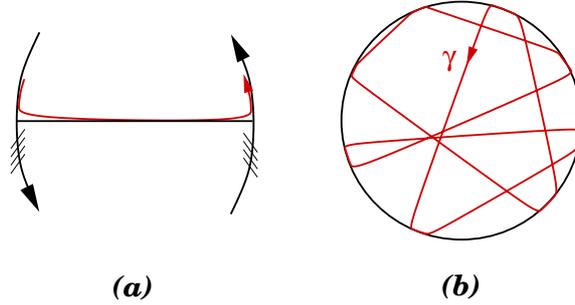


Figure 16: The local and global look of a proper ERS loop

Consider a permutation $\sigma \in \mathfrak{S}(\llbracket 1, n \rrbracket)$, and set

$$\nearrow(\sigma) := \#\{i \in \llbracket 1, n \rrbracket \mid \sigma(i) > i\}.$$

It is easy to check that if σ_0 is the circular permutation $(1\ 2 \dots n)$, then

$$\forall \sigma \in \mathfrak{S}, \nearrow(\sigma) = \nearrow(\sigma_0 \sigma \sigma_0^{-1}).$$

Definition 3.21. The invariance property from above means that \mathcal{T} is well-defined for permutations of a set of n points lying in an abstract oriented circle. We still denote this function by \mathcal{T} , and call it the *torsion* of a permutation.

Let γ be a proper simple loop, then the set of edges e such that $\langle \gamma, e \rangle \neq 0$ can be naturally assimilated to a finite subset of an oriented circle, and γ induces a permutation of this set. Let us denote it by σ_γ .

Theorem 3.22. For all proper ERS loops γ ,

$$\mathcal{T}(\gamma) = \nearrow(\sigma_\gamma).$$

This theorem can be useful in practice, since the torsion of a permutation can be computed at a glance on the braid-like presentation. Observe that

1. Every non proper loop is homologous to a multiple of $[K]$, easy to determine.
2. For every proper loop γ , there is an integer n such that $\tilde{\gamma} = \gamma + n[K]$ is proper, ER, and has zero coordinate along at least one edge. Namely, $n = -\min_e \langle \gamma, e \rangle$.
3. Every class $\tilde{\gamma}$ as above may be decomposed as a sum $\tilde{\gamma} = \sum_i \gamma_i$ such that
 - all the γ_i 's are proper and ERS

$$\rightarrow \forall i, j, A, \langle \gamma_i, A \rangle \langle \gamma_j, A \rangle \geq 0$$

4. By Lemma 3.18, $\mathcal{T}(\tilde{\gamma}) = \sum_i \mathcal{T}(\gamma_i)$, and the $\mathcal{T}(\gamma_i)$'s are given by Theorem 3.22.

This shows that it is possible to compute any homology class by using the torsion formula. Whether it is more interesting than the energy formula depends on the context.

Proof of Theorem 3.22. One may assume that for every arrow A , $\langle \gamma, A \rangle = 1$. Indeed, deleting an arrow avoided by γ , or reversing the orientation of an arrow that γ runs in the wrong direction, have no effect on either side of the formula (notably because of Lemma 3.19). Under this assumption, half of the edges of G are run by γ : call them the *red edges of G* , while the other half are called the *blue edges*. Red and blue edges alternate along the orientation of the circle.

If e is any (red or blue) edge, we define:

$$\lambda(e) := \sum_A \langle [A], e \rangle.$$

Lemma 3.23. *Under the assumption that $\langle \gamma, A \rangle = 1$ for all A , the value of $\lambda(e)$ only depends on the color of the edge e . Moreover,*

$$\lambda(\text{blue}) = \lambda(\text{red}) - 1 = \nearrow(\sigma_\gamma).$$

Let us temporarily admit this result. By the definition of λ ,

$$\begin{aligned} \sum_A [A] &= \sum \text{arrows} + \lambda(\text{red}) \sum(\text{red edges}) + \lambda(\text{blue}) \sum(\text{blue edges}) \\ &\stackrel{\text{Lemma 3.23}}{=} \sum \text{arrows} + \sum(\text{red edges}) + \lambda(\text{blue}) \sum(\text{red and blue edges}) \\ &= \gamma + \lambda(\text{blue})[K] \\ &\stackrel{\text{Lemma 3.23}}{=} \gamma + \nearrow(\sigma_\gamma)[K]. \end{aligned}$$

Since it was assumed that $\langle \gamma, A \rangle = 1$ for every arrow, the definition of \mathcal{T} (5) reads

$$\gamma = \sum_A [A] - \mathcal{T}(\gamma)[K],$$

which terminates the proof of the theorem, up to Lemma 3.23. □

Proof of Lemma 3.23. In the case of $\sigma_0 = (12 \dots n)$ depicted on Fig.17, it is easy to see that $\lambda(\text{red}) = n$ and $\lambda(\text{blue}) = n - 1$, while $\nearrow(\sigma_0) = n - 1$. The lemma being true for one diagram, let us show that it survives elementary changes that cover all the diagrams.

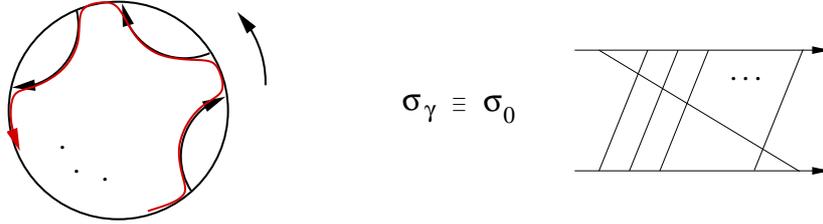


Figure 17: Braid-like representations of permutations are to be read from bottom to top

Notice that for every proper ERS loop γ , σ_γ is a *cycle*, and conversely a permutation that is a cycle uniquely defines an undecorated Gauss diagram *and* a proper ERS loop γ such that for every arrow A , $\langle \gamma, A \rangle = 1$. Thus, covering all possible permutations implies covering all possible diagrams and proper ERS loops. So all we have to check is that the formula survives an operation on σ_γ , of the form:

$$(\dots i j \dots) \rightarrow (\dots j i \dots)$$

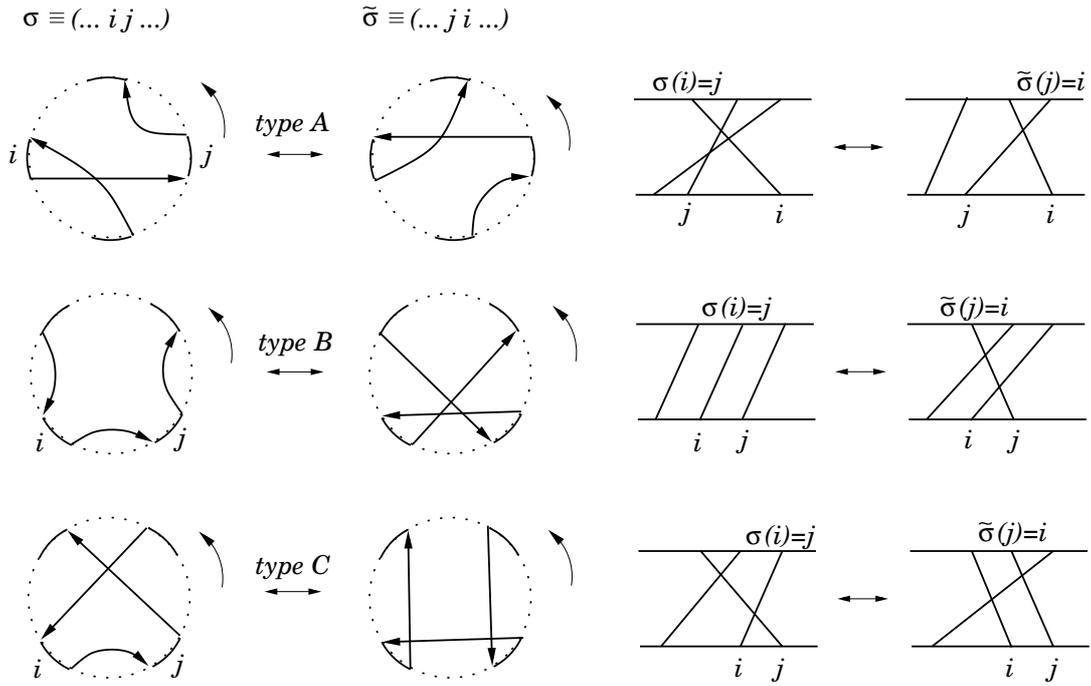


Figure 18: Twist moves on Gauss diagrams

The corresponding move at the level of Gauss diagrams may be of six different types, grouped in three pairs of reverse operations (Fig.18).

On each diagram in Fig.18, the three moving arrows split the base circle into six regions. One computes the variation of λ separately for each of these regions, and sees that it is the same for each of them. The results are gathered in the following table, proving the lemma.

type of move	variation of λ	variation of $\mathcal{T}(\gamma)$
A	unchanged	unchanged
B (from left to right)	decreases by 1	decreases by 1
C (from left to right)	decreases by 1	decreases by 1

□

4 Finite-type invariants

One of the main points of using Gauss diagrams is their ability to describe finite-type invariants by simple formulas [20, 7, 2, 3]. In the case of classical long knots in 3-space, such formulas actually cover all Vassiliev invariants as was shown by M.Goussarov [9]. In the virtual case, the two notions actually differ (see [13] and also [5, 4]). Finite type invariants for virtual knots that do admit Gauss diagram formulas shall be called GPV invariants [10].

In [17], a simple set of criteria was given to detect a particular family of those formulas, called *virtual arrow diagram formulas*. Most of the examples that are known belong to this family. That includes Chmutov-Khoury-Rossi's formulas for the coefficients of the Conway polynomial [2] (and their generalization by M. Brandenbursky [1]), as well as the formulas from [7, 8, 11] where different kinds of decorated diagrams are used. Note however that the formulas for the invariants extracted from the HOMFLYPT polynomial [3] are arrow diagram formulas only if the variable a is specialized to 1 (which yields back the result of [2]).

In this section, we extend the results from [17] to an arbitrary surface. Then we show how to apply them to any other kind of decorated diagrams found in the literature, by defining *symmetry-preserving maps* which enable one to jump from one theory to another.

4.1 General algebraic settings

We denote by \mathfrak{G}_n (*resp.* $\mathfrak{G}_{\leq n}$) the \mathbb{Q} -vector space freely generated by Gauss diagrams on π of degree n (*resp.* $\leq n$), and set $\mathfrak{G} = \varinjlim \mathfrak{G}_{\leq n}$. Unless π is a finite group, these spaces are not finitely generated, and we define their hat versions $\widehat{\mathfrak{G}}_n$ (*resp.* $\widehat{\mathfrak{G}}_{\leq n}$) as the \mathbb{Q} -spaces of formal series of Gauss diagrams of degree n (*resp.* $\leq n$). Finally, set $\widehat{\mathfrak{G}} = \varinjlim \widehat{\mathfrak{G}}_{\leq n}$. An arbitrary element of $\widehat{\mathfrak{G}}$ is usually denoted by \mathcal{G} and called a *Gauss series*, of degree n if it is represented in $\widehat{\mathfrak{G}}_{\leq n}$ but not in $\widehat{\mathfrak{G}}_{\leq n-1}$. The notation G is saved for single Gauss diagrams.

A Gauss diagram G of degree n has a *group of symmetries* $\text{Aut}(G)$, which is a subgroup of $\mathbb{Z}/2n$, made of the rotations of the circle that leave unchanged a given representative of G (see Subsection 4.1.2). \mathfrak{G} is endowed with the orthonormal scalar product with respect to its canonical basis, denoted by (\cdot, \cdot) , and its normalized version $\langle \cdot, \cdot \rangle$, defined by

$$\langle G, G' \rangle := |\text{Aut}(G)|(G, G'). \quad (7)$$

There is a linear isomorphism $I : \mathfrak{G}_{\leq n} \rightarrow \widehat{\mathfrak{G}}_{\leq n}$, the keystone to the theory, which maps a Gauss diagram of degree n to the formal sum of its 2^n subdiagrams:

$$I(G) = \sum_{\sigma \in \{\pm 1\}^n} G_{(\sigma)}, \quad (8)$$

where $G_{(\sigma)}$ is G deprived from the arrows that σ maps to -1 (see Definition 3.1 for subdiagrams). The inverse map of I is given by

$$I^{-1}(G) = \sum_{\sigma \in \{\pm 1\}^n} \text{sign}(\sigma) G_{(\sigma)}. \quad (9)$$

Definition 4.1. A finite-type invariant for virtual knots in the sense of Goussarov-Polyak-Viro is a virtual knot invariant given by a *Gauss diagram formula*

$$\nu_G : G \mapsto \langle \mathcal{G}, I(G) \rangle, \quad (10)$$

where $\mathcal{G} \in \widehat{\mathfrak{G}}$. Such a formula “counts” the subdiagrams of G , with weights given by the coefficients of \mathcal{G} . Notice that only one of the two arguments of $\langle \cdot, \cdot \rangle$ needs to be a finite sum for the expression to make sense. We do not make a distinction between a virtual knot invariant and the linear form induced on \mathfrak{G} .

4.1.1 The Polyak algebra

A Gauss series $\mathcal{G} \in \widehat{\mathfrak{G}}$ defines a virtual knot invariant if and only if the function $\langle \mathcal{G}, I(\cdot) \rangle$ is zero on the subspace spanned by R-moves and w -moves relators. Hence one has to understand the image of that subspace under I with a simple family of generators. This is the idea of the construction of the Polyak algebra ([19, 10]).

In the present case, \mathcal{P} is defined as the quotient of \mathfrak{G} by

- the relations shown in Fig.19, which we call P_1, P_2, P_3 (or *8T relation*),
- the W relation, which is simply the linear match of w -moves (*i.e.* just replace the “ \rightsquigarrow ” with a “=” in all the relations from Fig.5).

Be careful that unlike R_1 -moves, where an isolated arrow surrounding an edge marked with 1 simply disappears, in a P_1 -move the presence of such an arrow completely kills the diagram. Fig.19 does not feature the π -markings for P_3 to lighten the picture, but they have to follow the usual “merge \rightsquigarrow multiply” rule (see Definition 3.1).

The following proposition extends Theorem 2.D from [10].

Proposition 4.2. *The map I induces an isomorphism $\mathfrak{G}/_{\mathbf{R}, \mathbf{W}} \rightarrow \mathfrak{G}/_{\mathbf{P}, \mathbf{W}} =: \mathcal{P}$. More precisely, I induces an isomorphism between $\text{Span}(\mathbf{R}_i)$ and $\text{Span}(\mathbf{P}_i)$, for $i = 1, 2, 3$, and between $\text{Span}(\mathbf{W})$ and itself. It follows that the map $G \rightarrow I(G) \in \mathcal{P}$ defines a complete invariant for virtual knots.*

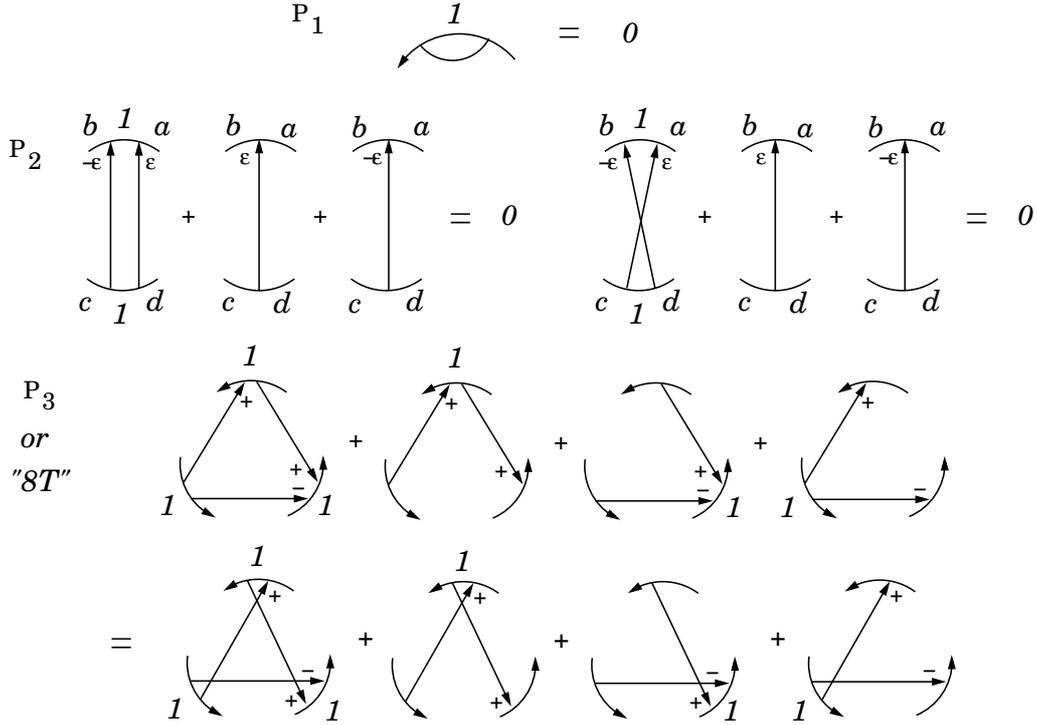


Figure 19: The three kinds of Polyak relations – only one P_3 relation is shown, there is a second one obtained by reversing all the arrow orientations.

4.1.2 The symmetry-preserving injections

Depending on the context, one may have to consider simultaneously different types of Gauss diagrams, with more or less decorations. This subsection presents a natural way to do it, convenient from the viewpoint of Gauss diagram invariants. The construction requires one to choose a kind of combinatorial objects that is the “father” of all other kinds, in the sense of quotienting. We present the construction by taking as the father type that of Gauss diagrams on a group.

In first place, we do not regard Gauss diagrams up to homeomorphisms of the circle: the base circle is assumed to be the unit circle in \mathbb{C} , the endpoints of the arrows are assumed to be located at the $2n$ -th roots of unity, and the arrows are straight line segments. Such a diagram is called *rigid*.

By a “type of rigid Gauss diagrams” we mean an equivalence relation on the set of rigid Gauss diagrams on π , which is required to satisfy two properties:

1. (Degree property) All diagrams in a given equivalence class shall have the same degree.
2. (Stability property) The action of $\mathbb{Z}/2n$ on the set of Gauss diagrams of degree n shall induce an action on the set of degree n equivalence classes.

Since every construction in this subsection is therefore destined to be homogeneous, the degree of all Gauss diagrams is once and for all set equal to n . A *rigid Gauss diagram of type \sim* is an equivalence class under the relation \sim . A *Gauss diagram (of type \sim)* is the orbit of a rigid diagram of type \sim under the action of $\mathbb{Z}/2n$. The corresponding \mathbb{Q} -spaces are respectively denoted by $\mathfrak{G}_{\sim}^{\text{rigid}}$ and \mathfrak{G}_{\sim} .

Since $\mathbb{Z}/2n$ is abelian, two elements from the same orbit have the same stabilizer, hence a Gauss diagram G has a well-defined *group of symmetries* $\text{Aut}(G)$, which is the stabilizer of any of its rigid representatives under the action of $\mathbb{Z}/2n$. Consequently, the space \mathfrak{G}_{\sim} is endowed with a pairing $\langle \cdot, \cdot \rangle$ defined by (7).

Now consider two types of rigid Gauss diagrams, say 1 and 2, such that relation 1 is finer than relation 2 (“1 \prec 2”).

Definition 4.3 (Forgetful projections). A 1-rigid diagram G_1 determines a unique 2-rigid diagram whose $\mathbb{Z}/2n$ -orbit only depends on that of G_1 . This induces a natural surjective map at the level of Gauss diagram spaces, denoted by

$$T_2^1 : \mathfrak{G}_{(1)} \twoheadrightarrow \mathfrak{G}_{(2)}.$$

Note that this map may be not well-defined on the spaces of formal series of Gauss diagrams, if some 2-equivalence class contains infinitely many 1-classes.

Example: the abelianization map ab (Definition 3.7) induces by linearity a forgetful projection from Gauss diagrams on π to abelian diagrams on π , when π is abelian.

Definition 4.4 (Symmetry-preserving injections). In the opposite way, there is a map $\mathfrak{G}_{(2)}^{\text{rigid}} \rightarrow \mathfrak{G}_{(1)}^{\text{rigid}}$ that sends a 2-rigid diagram G_2 to the formal sum of all 1-classes that it contains. When this sum is pushed in $\mathfrak{G}_{(1)}$, the result

- is well-defined: a 2-rigid diagram cannot contain infinitely many rigid representatives of a given Gauss diagram of type 1, since the orbits are finite ($\mathbb{Z}/2n$ is finite).
- only depends on the $\mathbb{Z}/2n$ -orbit of G_2 .

This induces an injective *symmetry-preserving map* at the level of formal series,

$$S_2^1 : \widehat{\mathfrak{G}}_{(2)} \hookrightarrow \widehat{\mathfrak{G}}_{(1)}.$$

S_2^1 is well-defined, componentwise, since 2-rigid diagrams from different $\mathbb{Z}/2n$ -orbits contain 1-rigid diagrams from disjoint sets of $\mathbb{Z}/2n$ -orbits (the images of two different Gauss diagrams do not overlap). It is injective for the same reason.

The terminology is explained by the following fundamental formula:

Lemma 4.5. *With notations as above, for any Gauss diagram G_2 of type 2,*

$$S_2^1(G_2) = \sum_{T_2^1(G_1)=G_2} \frac{|\text{Aut}(G_2)|}{|\text{Aut}(G_1)|} G_1. \quad (11)$$

Informally, the weight given to a preimage of G_2 under T_2^1 is the amount of symmetry that it has lost by the gain of more information. Note that the weights are integers, since $\text{Aut}(G_1)$ identifies with a subgroup of $\text{Aut}(G_2)$.

Proof. Fix a representative G_2^{rigid} of G_2 . By the stability property, $\text{Aut}(G_2)$ acts on the set of 1-classes contained in G_2^{rigid} . Moreover, by definition of $\text{Aut}(G_2)$, two different orbits under that action still lie in different orbits under the action of $\mathbb{Z}/2n$ itself. Therefore there is a 1 – 1 correspondence between the $\text{Aut}(G_2)$ -orbits and the Gauss diagrams that happen in the sum $S_2^1(G_2)$. The stabilizer of a given 1-class G_1^{rigid} is by definition $\text{Aut}(G_1)$, whence the cardinality of the corresponding orbit, which is also the coefficient of G_1 in $S_2^1(G_2)$, is $\frac{|\text{Aut}(G_2)|}{|\text{Aut}(G_1)|}$. \square

Proposition 4.6.

1. *For any three relations such that $1 \prec 2 \prec 3$, the following diagrams commute:*

$$\begin{array}{ccc}
 \mathfrak{G}_{(1)} & & \widehat{\mathfrak{G}}_{(3)} \\
 \downarrow T_2^1 & \searrow T_3^1 & \xrightarrow{S_3^2} \widehat{\mathfrak{G}}_{(2)} \\
 \mathfrak{G}_{(2)} & \xrightarrow{T_3^2} \mathfrak{G}_{(3)} & \downarrow S_3^1 \\
 & & \widehat{\mathfrak{G}}_{(1)}
 \end{array}$$

2. Injections and projections are pairwise \langle, \rangle -adjoint, in the sense that

$$\forall \mathcal{G}_1 \in \mathfrak{G}_{(1)}, \mathcal{G}_2 \in \widehat{\mathfrak{G}}_{(2)}, \quad \langle S_2^1(\mathcal{G}_2), \mathcal{G}_1 \rangle = \langle \mathcal{G}_2, T_2^1(\mathcal{G}_1) \rangle.$$

3. $\text{Im } S_2^1 = \text{Ker}^\perp T_2^1$.

Proof. 1. The first diagram commutes directly from the definition of the maps T_i^j . As for the maps S_i^j , since they are defined componentwise it is enough to check it for a single diagram G_3 . In that case, it is a consequence of Lemma 4.5 and the relation $T_3^1 = T_3^2 \circ T_2^1$.

2. In both sides, it is clear that only a finite number of terms in \mathcal{G}_2 are relevant, namely those that are projections of some terms of \mathcal{G}_1 under T_2^1 . Thus, by bilinearity, it is enough to consider single diagrams G_1 and G_2 . If $G_2 \neq T_2^1(G_1)$, then both sides are 0. If $G_2 = T_2^1(G_1)$, then

$$\begin{aligned} \langle S_2^1(G_2), G_1 \rangle &= \left\langle \frac{|\text{Aut}(G_2)|}{|\text{Aut}(G_1)|} G_1, G_1 \right\rangle \\ &= |\text{Aut}(G_2)|, \end{aligned}$$

while

$$\begin{aligned} \langle G_2, T_2^1(G_1) \rangle &= \langle G_2, G_2 \rangle \\ &= |\text{Aut}(G_2)|. \end{aligned}$$

3. The inclusion $\text{Im } S_2^1 \subset \text{Ker}^\perp T_2^1$ follows immediately from 2. For the converse, pick a Gauss diagram series \mathcal{G}_1 in $\text{Ker}^\perp T_2^1$. For any two 2-related Gauss diagrams of type 1, G_1 and G_1' , one has

$$\langle \mathcal{G}_1, G_1 - G_1' \rangle = 0.$$

Thus, if G_2 is a Gauss diagram of type 2, one can define $\phi(G_2)$ to be the value of $\langle \mathcal{G}_1, G_1 \rangle$ for any preimage G_1 of G_2 under T_2^1 , and set

$$\mathcal{G}_2 = \sum \frac{\phi(G_2)}{|\text{Aut}(G_2)|} G_2,$$

where the sum runs over all Gauss diagrams of type 2. Finally,

$$\begin{aligned} S_2^1(\mathcal{G}_2) &= \sum \frac{|\text{Aut}(T_2^1(G_1))|}{|\text{Aut}(G_1)|} \frac{\phi(T_2^1(G_1))}{|\text{Aut}(T_2^1(G_1))|} G_1 \\ &= \sum \frac{\langle \mathcal{G}_1, G_1 \rangle}{|\text{Aut}(G_1)|} G_1 \\ &= \mathcal{G}_1 \end{aligned}$$

□

In practice, point 3 is useful in both directions: whether one needs a characterization of the series that lie in the image of some map S (Lemma 4.14), or of the series that define invariants under some kind of moves (Proposition 4.20). Point 2 states that symmetry-preserving maps are the good dictionary to understand invariants that were defined via forgetful projections.

Remark 4.7. Every construction and result in this subsection can be repeated by replacing the set of rigid Gauss diagrams of degree n with any set endowed with the action of an abelian finite group.

4.1.3 Arrow diagrams and homogeneous invariants

Definition 4.8 (see [19, 20]). An *arrow diagram* (on π) is a Gauss diagram G (on π) of which the signs decorating the arrows have been forgotten. As usual, it is considered up to homeomorphisms of the circle.

Arrow diagram spaces $\mathfrak{A}_n, \mathfrak{A}_{\leq n}, \mathfrak{A}$, the hat versions, and the pairings $(,)$ and \langle, \rangle are defined similarly to their signed versions (Subsection 4.1). We use notations A for an arrow diagram and \mathcal{A} for an *arrow diagram series* – i.e. an element of $\widehat{\mathfrak{A}}$.

Arrow diagram formulas

In the language of Subsection 4.1.2, arrow diagrams are a kind of Gauss diagrams satisfying the degree and stability properties – the equivalence relation on rigid Gauss diagrams is given by $G \sim G' \Leftrightarrow$ one may pass from G to G' by writhe changes. Therefore, an arrow diagram A has a well-defined symmetry group $\text{Aut}(A)$, and there are a symmetry-preserving map $S_a : \widehat{\mathfrak{A}} \hookrightarrow \widehat{\mathfrak{G}}$ and a projection $T_a : \mathfrak{G} \rightarrow \mathfrak{A}$. However, for the purpose of defining arrow diagram invariants, we are going to twist these maps a little, by pushing an additional sign into the weights:

Definition 4.9. Define the linear maps $S : \widehat{\mathfrak{A}} \rightarrow \widehat{\mathfrak{G}}$ and $T : \mathfrak{G} \rightarrow \mathfrak{A}$ componentwise by

$$S(A) := \sum_{T_a(G)=A} \text{sign}(G) \frac{\text{Aut}(A)}{\text{Aut}(G)} G \quad (12)$$

$$T(G) := \text{sign}(G) T_a(G) \quad (13)$$

Proposition 4.10.

1. S and T are \langle, \rangle -adjoint, in the sense that

$$\forall \mathcal{G} \in \mathfrak{G}, \mathcal{A} \in \widehat{\mathfrak{A}}, \quad \langle S(\mathcal{A}), \mathcal{G} \rangle = \langle \mathcal{A}, T(\mathcal{G}) \rangle.$$

2. $\text{Im } S = \text{Ker}^\perp T$.

The proof is completely similar to that of Proposition 4.6.

Definition 4.11. A Gauss diagram formula that lies in the image of the map S is called an *arrow diagram formula*.

Homogeneous invariants

Definition 4.12. For each $n \in \mathbb{N}$, there is an orthogonal projection $p_n : \widehat{\mathfrak{G}} \rightarrow \widehat{\mathfrak{G}}_n$ with respect to the scalar product \langle, \rangle . For $\mathcal{G} \in \widehat{\mathfrak{G}}$, the *principal part* of \mathcal{G} is defined by $p_n(\mathcal{G})$, with $n = \text{deg}(\mathcal{G})$. \mathcal{G} is called *homogeneous* if it is equal to its principal part.

Let \mathcal{G} be a homogeneous Gauss series. Then \mathcal{G} satisfies the P_2 and P_3 relations (i.e. $\langle \mathcal{G}, P_i \rangle = 0$ for $i = 2, 3$) if and only if it satisfies the *homogeneous* relations $\langle \mathcal{G}, p_k(P_i) \rangle = 0$ for all k . These are denoted by $P_2^{(n-1),1}$, $P_2^{(n-2),2}$, $P_3^{(n-2),2}$ (or $G6T$) and $P_3^{(n-3),3}$ (or $G2T$). The parenthesized numbers in exponent indicate in each case how many arrows are unseen. P_1 and W relations are already homogeneous and do not get a new name. Some examples are shown on Fig.20; for a full list, just consider the projections of the relations of Fig.19.

Homogeneous relations are also defined for arrow diagram spaces, denoted by AP_1 , $AP_2^{(n-2),2}$, $AP_3^{(n-2),2}$ (or $A6T$), $AP_3^{(n-3),3}$ (or $A2T$) and AW (Lemma 4.14 below explains why $AP_2^{(n-1),1}$ is useless: it reads $0 = 0$). They are the images under T (13) of the homogeneous relations for Gauss diagrams – in particular one should be especially careful to the signs in the $A6T$ relations. A full list is presented in Fig.21.

Lemma 4.13. Let $\mathcal{A} \in \widehat{\mathfrak{A}}$ and let X be a name among P_1 , $P_2^{(n-2),2}$, $P_3^{(n-2),2}$, $P_3^{(n-3),3}$, W . Then

$$\mathcal{A} \in \text{Span}^\perp(\text{AX}) \iff S(\mathcal{A}) \in \text{Span}^\perp(X).$$

where orthogonality is as usual in the sense of \langle, \rangle .

Proof. It is a direct consequence of part 1. of Proposition 4.10. \square

Lemma 4.14. Let $\mathcal{G} \in \widehat{\mathfrak{G}}$. Then \mathcal{G} lies in the image of the map $S : \widehat{\mathfrak{A}} \rightarrow \widehat{\mathfrak{G}}$ if and only if \mathcal{G} satisfies all the homogeneous relations $\langle \mathcal{G}, P_2^{(n-1),1} \rangle = 0$.

Proof. Notice that the $P_2^{(n-1),1}$ relators span the kernel of the map T . Hence the result follows from point 2. of Proposition 4.10. \square

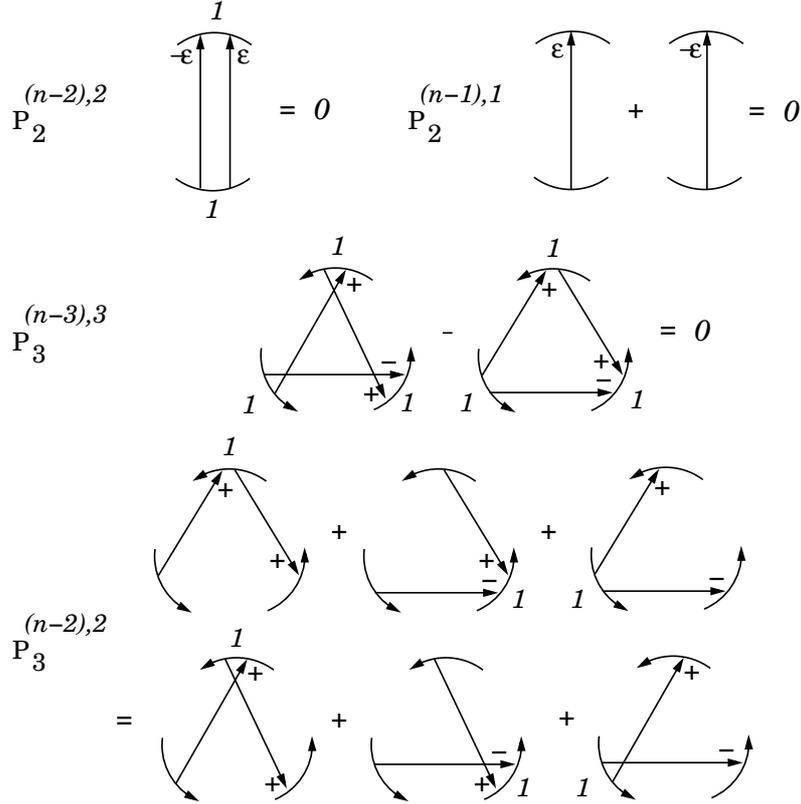


Figure 20: Some homogeneous Polyak relations

The following are proved in a particular case in [17] (Lemma 3.2 and Theorem 2.5); the proof can be readily adapted to the present situation.

Lemma 4.15. *For all $n \geq 3$:*

$$\text{Span}(A2T) \subseteq \text{Span}(A6T) \oplus \text{Span}(AP_2^{(n-2),2}).$$

Theorem 4.16. *Arrow diagram formulas are exactly the linear combinations of homogeneous Gauss diagram formulas.*

4.1.4 Based and degenerate diagrams

Definition 4.17. A *based* Gauss diagram is a Gauss diagram together with a distinguished (*base*) edge. Based arrow diagrams are defined similarly. The corresponding spaces are denoted by \mathfrak{G}_\bullet and \mathfrak{A}_\bullet , in reference to the dot that we use in practice to pinpoint the distinguished edge.

A *degenerate Gauss diagram (with one degeneracy)* is a Gauss diagram in which one edge, whose endpoints belonged to two different arrows, has been shrunk to a point. The spaces of degenerate diagrams are denoted by \mathfrak{DG} and \mathfrak{DA} respectively.

Even though they would encode long knots in the classical theory, based diagrams have only a combinatorial interest here. On the other hand, degenerate diagrams have a natural topological interpretation that is explained in [15].

The space of degenerate arrow diagrams is meant to be quotiented by the so-called *triangle relations*, shown in Fig.22. The quotient space is denoted by \mathfrak{DA}/∇ . These relations originated in the early work of M.Polyak on arrow diagrams ([18, 19], see also [21]).

Definition 4.18. Call a degenerate diagram *monotonic* if an arrowhead and an arrowtail meet at the degenerate point.

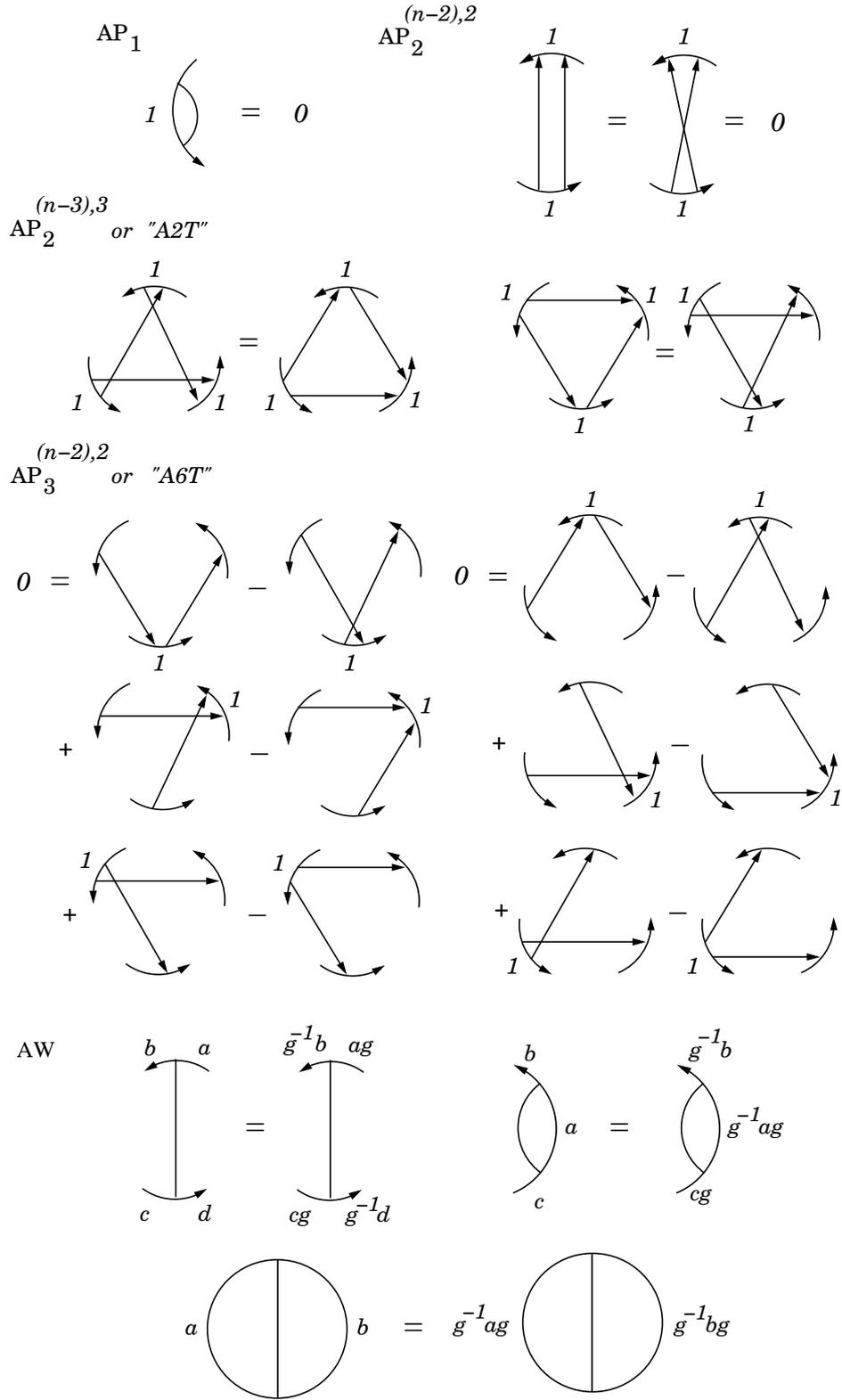


Figure 21: The homogeneous arrow relations – as usual, in AW the orientation of an arrow is reversed if and only if $w(g) = -1$.

Lemma 4.19. $\widehat{\mathfrak{A}}/\nabla$ is naturally isomorphic with the \mathbb{Q} -space of formal series of monotonic arrow diagrams.

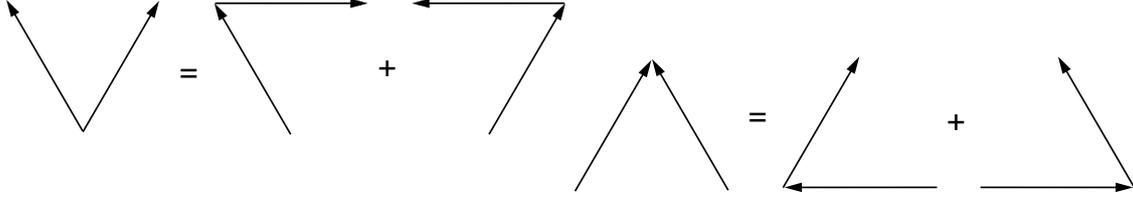


Figure 22: The triangle relations

Proof. It suffices to show that the set of monotonic diagrams forms a basis of \mathfrak{DA}/∇ . It is clearly a generating set thanks to the ∇ relations, and it is free because every non monotonic diagram happens in exactly one relation, and every relation contains exactly one of them. \square

4.2 Invariance criteria

w -invariance

Proposition 4.20. *There is an injective “symmetry-preserving” map $S_{aw}^a : \widehat{\mathfrak{A}}/\widehat{\text{AW}} \hookrightarrow \widehat{\mathfrak{A}}$ defined componentwise by the formula*

$$\alpha \mapsto \sum_{A \in \alpha} \frac{|\text{Aut}(\alpha)|}{|\text{Aut}(A)|} A.$$

If $\mathcal{A} \in \widehat{\mathfrak{A}}$, then the map $G \mapsto \langle \langle \mathcal{A}, G \rangle \rangle = \langle S(\mathcal{A}), I(G) \rangle$ is invariant under w -moves if and only if \mathcal{A} lies in the image of S_{aw}^a .

Proof. The equivalence relation defined by AW-moves and writhe changes on the set of rigid Gauss diagrams on π satisfies the degree and stability properties from Subsection 4.1.2. Hence one may apply the results of that section to get the existence and elementary properties of S_{aw}^a . The last assertion follows from successive application of the W part of Lemma 4.13 and point 3 of Proposition 4.6. \square

This means that an arrow diagram formula must be represented by a series of w -orbits of arrow diagrams. In practice, this condition is most of time satisfied by construction. An important example is the formal sum of all elements contained in a set that is stable under w -moves.

R_1 and R_2 invariance

Proposition 4.21. *Let $\mathcal{A} \in \widehat{\mathfrak{A}}$. Then the function $G \mapsto \langle S(\mathcal{A}), I(G) \rangle$ is invariant under R_1 (resp. R_2) moves if and only if \mathcal{A} satisfies all AP_1 (resp. $\text{AP}_2^{(n-2),2}$) relations.*

Proof. By Proposition 4.2, the R_1 and R_2 invariance are equivalent to the relations $\langle S(\mathcal{A}), P_1 \rangle = 0$ and $\langle S(\mathcal{A}), P_2 \rangle = 0$ being satisfied. Lemma 4.14 implies that the second of these relations is actually equivalent to $\langle S(\mathcal{A}), \text{AP}_2^{(n-2),2} \rangle = 0$. Lemma 4.13 concludes the proof. \square

In practice, these conditions are easy to check naked-eye.

R_3 invariance

Definition 4.22. Say that a based diagram A_\bullet is *nice* if its base edge

- \rightarrow is decorated by 1,
- \rightarrow is bounded by the endpoints of two different arrows.

Lemma 4.23. *Nice based diagrams have a sign ε induced by Definition 1.2.*

Proof. The sign ε from Definition 1.2 takes as an argument the edge of a Gauss diagram. Since it does not depend on the writhes of the arrows, it is well-defined for arrow diagrams with a preferred edge. \square

Definition 4.24. Let A_\bullet be a based arrow diagram. If A_\bullet is not nice, then put $\delta(A_\bullet) = 0$. If A_\bullet is nice, then

1. Shrink the base edge to a point.
2. Multiply the resulting degenerate diagram by $\varepsilon(A_\bullet)$, and call the result $\delta(A_\bullet)$.

This process defines a map $\delta : \widehat{\mathfrak{A}}_\bullet \rightarrow \widehat{\mathfrak{DA}}_\bullet$ – well defined since any monotonic diagram has finitely many preimages. Now let A be an arrow diagram, and denote by $\bullet(A) \in \mathfrak{A}_\bullet$ the sum of all based diagrams that one can form by choosing a base edge in A . Finally, define

$$d : \begin{array}{ccc} \widehat{\mathfrak{A}} & \rightarrow & \widehat{\mathfrak{DA}} \\ \mathcal{A} & \mapsto & \delta(\bullet(\mathcal{A})) \end{array} .$$

Theorem 4.25. Let $\mathcal{A} \in \widehat{\mathfrak{A}}$ satisfy the R_2 invariance condition from Proposition 4.21. Then the following are equivalent:

- The map $G \mapsto \langle S(\mathcal{A}), I(G) \rangle$ is invariant under R_3 -moves.
- $d(\mathcal{A}) = 0$ modulo the triangle relations.
- $\mathcal{A} \in \text{Span}^\perp(A_6T)$.

The proof is similar to that of Theorem 3.6 from [17].

4.2.1 Invariance criterion for w -orbits

As we have seen, an arrow diagram series defines an invariant only if it has a preimage in $\widehat{\mathfrak{A}/\text{AW}}$. The above criteria for invariance under R -moves nicely extend in terms of that preimage. This is especially interesting when w -orbits are well understood, for instance if π is abelian and w is trivial (Proposition 3.8).

w -moves for based and degenerate diagrams

These moves are defined similarly to the regular version (Fig.5, top-left and top-right), with additional moves in the degenerate case, that modify the neighborhood of the arrows meeting at the degenerate point (Fig.23 shows the extreme cases – there are obvious intermediate ones, when only one or two of the unseen arcs is empty). The arrows both change orientations if $w(g) = -1$, and keep the same otherwise.

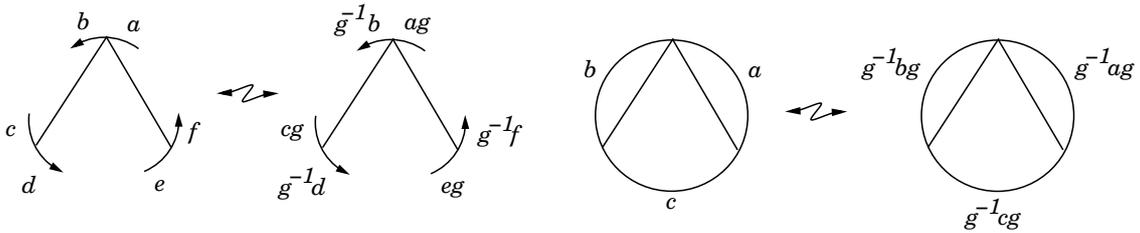


Figure 23: The “degenerate” w -move – the most general and the most exceptional cases. The two arrows change orientations if and only if $w(g) = -1$.

Definition 4.26. Pick a based arrow diagram A_\bullet . If its base edge is bounded by twice the same arrow, then set $\delta_w([A_\bullet]) = 0$. If it is bounded by two different arrows, then

1. pick a nice diagram $A_\bullet^{(1)}$ AW-equivalent to A_\bullet ,
2. set $\delta_w([A_\bullet]) = [\delta(A_\bullet^{(1)})]$.

Finally, set

$$d_w([A]) = \delta_w([\bullet(A)]).$$

As usual, δ_w and d_w are defined componentwise on formal series of w -orbits.

Consistency of the definition. Observe that when the endpoints of an edge belong to two different arrows, then any value can be given to its marking by using the appropriate w -move. This proves that step 1 is always possible – though not in a unique way.

For step 2, first notice that two w -moves performed on different arrows from an arrow diagram commute, so that any finite sequence of w -moves amounts to a sequence made of one move for each arrow. If such a sequence leaves the marking of the base edge unchanged (and equal to 1), then the moves on the two adjacent arrows must involve the same conjugating element g . It follows that whenever A_\bullet and A'_\bullet are nice and lie in the same w -orbit,

$$\rightarrow \varepsilon(A_\bullet) = \varepsilon(A'_\bullet),$$

\rightarrow the degenerate diagrams obtained by shrinking their bases also lie in the same orbit.

Hence δ_w is well-defined. \square

How to handle the quotients by ∇ relations

Note that the quotient of \mathfrak{DA} by the ∇ relations does not fit in the general framework in which symmetry preserving maps were introduced: indeed, it does not come from an equivalence relation at the level of the *set* of diagrams.

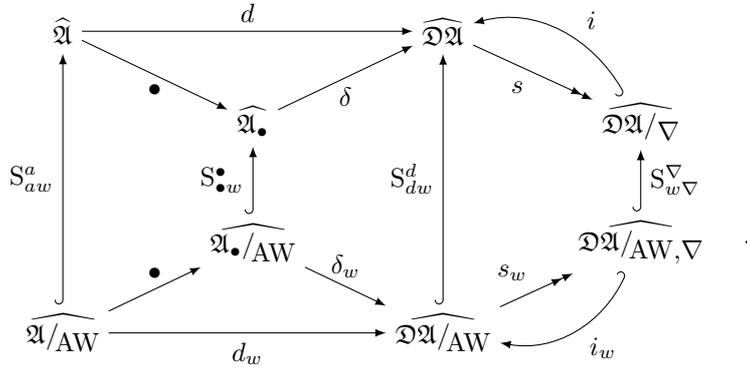
However, the set of classes of monotonic diagrams forms a basis of \mathfrak{DA}/∇ . This induces an injective section i of the projection $s : \mathfrak{DA} \twoheadrightarrow \mathfrak{DA}/\nabla$. Both s and i extend componentwise to formal series of w -orbits.

The same phenomenon happens between $\widehat{\mathfrak{DA}/\text{AW}}$ and $\widehat{\mathfrak{DA}/\text{AW},\nabla}$, because w -moves never change the status of a degenerate diagram – monotonic or not – and the set of monotonic w -orbits still forms a basis of $\widehat{\mathfrak{DA}/\text{AW},\nabla}$. Again there are maps s_w and i_w such that $s_w \circ i_w = \text{Id}$, at the level of formal series.

Finally, this allows us to construct a symmetry-preserving map $S_{w\nabla}^\nabla : \widehat{\mathfrak{DA}/\text{AW},\nabla} \rightarrow \widehat{\mathfrak{DA}/\nabla}$, by regarding the restriction of $S_{dw}^d : \widehat{\mathfrak{DA}/\text{AW}} \rightarrow \widehat{\mathfrak{DA}}$ to the subspaces of monotonic diagrams.

Summary

In the following diagram, the two squares and the two triangles on the left are commutative, as well as the internal and external squares on the right. Except for $S_{w\nabla}^\nabla$, all vertical arrows are symmetry-preserving injections in the usual sense.



Theorem 4.27. *An arrow diagram series \mathcal{A} is an arrow diagram formula if and only if each of the following holds:*

1. \mathcal{A} has a preimage \mathcal{A}_w by S_{aw}^a .

2. \mathcal{A}_w is mapped to 0 in $\widehat{\mathfrak{D}\mathfrak{A}/\mathfrak{AW}, \nabla}$.

3. \mathcal{A}_w satisfies the equations

$$\langle \mathcal{A}_w, T_{aw}^a(\text{AP}_1) \rangle = 0 \quad \text{and} \quad \langle \mathcal{A}_w, T_{aw}^a(\text{AP}_2^{(n-2),2}) \rangle = 0 .$$

Proof. 1 is necessary because of Proposition 4.20, and 3 because of Proposition 4.21 and point 2 of Proposition 4.6. If 1 and 3 are satisfied, then Theorem 4.25 implies that \mathcal{A} defines an arrow diagram formula if and only if $s(d(\mathcal{A})) = 0$. This is equivalent to $S_{w\nabla}^\nabla(s_w(d_w(\mathcal{A}_w))) = 0$ since the diagram commutes, and to $s_w(d_w(\mathcal{A}_w)) = 0$ since $S_{w\nabla}^\nabla$ is injective. \square

Remark 4.28. To apply the above theorem to an element of $\widehat{\mathfrak{A}/\mathfrak{AW}}$, one never needs to push it through symmetry-preserving maps (in the upper half of the diagram). Hence the checkings are done using the most compact expressions. Also, Point 3 can be checked separately on each w -orbit that happens in \mathcal{A}_w , and for each of them it can be checked on any representative diagram by a simple criterion.

For AP_1 relations to hold, the situation at the left of Fig.24 simply must never happen (the 1-marking is invariant under w -moves).

For $\text{AP}_2^{(n-2),2}$, the situations at the middle and at the right of the picture are forbidden:

1. if $w(g) = 1$ and the arrows have “the same” orientation (in the sense of the picture),
2. if $w(g) \neq 1$ and the arrows have “different” orientations.

Again these conditions are stable by w -moves.

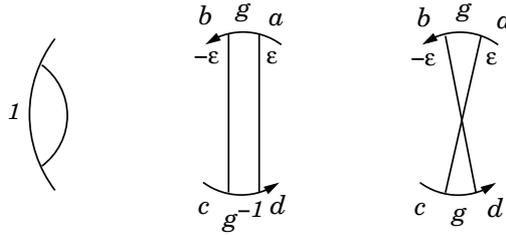


Figure 24: Forbidden situations – the rules for the arrow orientations are explained above.

4.3 Examples and applications

It was noticed by M. Polyak [18] that several families of formulas describing the finite-type invariants extracted from the Conway polynomial [2, 1] actually define invariants of virtual knots. That also led him to conjecture that there should be an invariance criterion such as the one we give here for R_3 -moves.

We describe here two other families of examples: first, Grishanov-Vassiliev formulas [11], which are extended to degenerate cases that were excluded in the original paper, and second, given a surface Σ we describe a regular invariant that can reasonably be used as a definition of the Whitney index for virtual knots whose projection on Σ is a non nullhomotopic curve.

4.3.1 Grishanov-Vassiliev’s planar chain invariants

Definition 4.29. A *naked arrow diagram* is an arrow diagram with every decoration forgotten except for the local orientations. It is called *planar* if no two of its arrows intersect – thus one may regard it as a part of the plane, up to isotopy.

A *chain presentation* of such a diagram with n arrows is a way to number its $n + 1$ bounded complementary components in the plane from 1 to $n + 1$, in such a way that the numbering increases when one goes from the left to the right of an arrow.

Let U_n be the sum of all planar isotopy equivalence classes of chain presentations of naked arrow diagrams of degree n . U_n is called the *universal degree n planar chain* ([11], Definition 1).

Definition 4.30. An h_1 -decorated planar diagram is the result of assigning an element of $h_1(\Sigma) \setminus \{1\}$ to each region in a planar naked arrow diagram.

We consider two ways to construct such diagrams:

1. From the datum of a chain presentation together with a system $\Gamma = \{\gamma_1, \dots, \gamma_{n+1}\}$ – which yields the notion of Γ -decorated diagrams.
2. From a planar arrow diagram on $\pi = \pi_1(\Sigma)$: each region of the diagram receives the conjugacy class of the product of the π -markings at its boundary, in the order induced by the orientation of the circle (the product is not well-defined, but its conjugacy class is).

Call Φ_Γ the sum obtained by decorating every diagram in U_n with a fixed system Γ ([11], Definition 2). Note that some of the summands in U_n – namely those with non trivial symmetries – may lead to the same decorated diagram if some of the γ_i 's are equal; unlike Grishanov-Vassiliev, we do not forbid that. Of course these summands happen with coefficients greater than 1 in Φ_Γ .

To understand Φ_Γ as an arrow diagram series, we apply the machinery of symmetry-preserving maps from Subsection 4.1.2. Even though it is absent from the notations, all considered diagrams are planar:

- \mathfrak{A}^Γ is the \mathbb{Q} -space generated by Γ -decorated planar diagrams (hence of degree n).
- $\mathfrak{A}^{\rightsquigarrow\Gamma}$ is the subspace of \mathfrak{A}_n generated by planar arrow diagrams on $\pi_1(\Sigma)$ that induce Γ -decorated diagrams.
- $\mathfrak{A}_n^{1,2,\dots}$ is the \mathbb{Q} -space generated by all chain presentations of planar naked arrow diagrams of degree n .

Note that $\mathfrak{A}_n^{1,2,\dots}$ is not defined by an equivalence relation on rigid Gauss diagrams on $\pi_1(\Sigma)$. The father of all types of Gauss diagrams here is the type of planar Gauss diagrams on $\pi_1(\Sigma)$ endowed with a chain presentation, such that the chain presentation and the π -markings induce the same h_1 -decorated diagram – it is the pullback of Diagram (14) below.

Let us denote by S_Γ^a the symmetry-preserving map $\widehat{\mathfrak{A}}^\Gamma \hookrightarrow \widehat{\mathfrak{A}}_n$, and set

$$\tilde{\Phi}_\Gamma = S_\Gamma^a(\Phi_\Gamma).$$

Theorem 4.31. For any system of non-trivial conjugacy classes $\Gamma = (\gamma_1, \dots, \gamma_{n+1})$, $\tilde{\Phi}_\Gamma$ defines an arrow diagram formula for virtual knots.

The reason why $\tilde{\Phi}_\Gamma$ coincides as an invariant with Grishanov-Vassiliev's Φ_Γ is Point 2 of Proposition 4.6. This theorem improves Theorem 1 of [11], since we remove the assumption that Γ is unambiguous (*i.e.* any of the γ_i 's may coincide) and show that $\tilde{\Phi}_\Gamma$ is an invariant for virtual knots.

Proof. 1. First, notice that an equivalence class of arrow diagrams under AW-moves determines an h_1 -decorated diagram. Thus the map S_Γ^a factorizes through S_{aw}^a (by point 1. of Proposition 4.6), wherefrom Proposition 4.20 implies that $\tilde{\Phi}_\Gamma$ defines an invariant under w -moves.

2. The fact that no γ_i may be trivial gives immediately the condition of R_1 and R_2 invariance from Proposition 4.21.

3. For R_3 , the more convenient here is to check condition 3 of Theorem 4.25. In any $A6T$ relation, only three diagrams possibly have pairwise non intersecting arrows, and either all three of them do, either no one does. This yields two kinds of reduced relations. Let us consider only that on Fig.25 (ignore the markings i, j, k for now), the other case is completely similar. Write the relator of the picture as $A_1 - A_2 - A_3$ in the order of reading. One has to prove that

$$\langle \tilde{\Phi}_\Gamma, A_1 \rangle = \langle \tilde{\Phi}_\Gamma, A_2 \rangle + \langle \tilde{\Phi}_\Gamma, A_3 \rangle.$$

The three spaces defined previously fit into the diagram

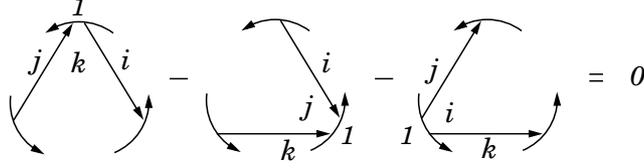


Figure 25: One of the two reduced 6-term relations for planar diagrams

$$\begin{array}{ccc}
 \mathfrak{A}^{\rightsquigarrow \Gamma} & & \mathfrak{A}_n^{1,2,\dots} \\
 \searrow T_\Gamma^a & & \swarrow T_\Gamma^{1,2,\dots} \\
 & \mathfrak{A}^\Gamma &
 \end{array} \quad . \quad (14)$$

Pick a planar arrow diagram A of degree n , and consider the set $\text{Dec}(A)$ of all rigid representatives of all preimages of $T_\Gamma^a(A)$ under the map $T_\Gamma^{1,2,\dots}$. By Definition 4.4, the cardinality of that set is the sum of all coefficients of $S_\Gamma^{1,2,\dots}(T_\Gamma^a(A))$, which means, since the sum of all generators of $\mathfrak{A}_n^{1,2,\dots}$ is U_n , that:

$$\#\text{Dec}(A) = \left(U_n, S_\Gamma^{1,2,\dots}(T_\Gamma^a(A)) \right).$$

But no diagram decorated with a chain presentation admits non-trivial symmetries, whence, by successive applications of Proposition 4.6 2.,

$$\begin{aligned}
 \#\text{Dec}(A) &= \left\langle U_n, S_\Gamma^{1,2,\dots}(T_\Gamma^a(A)) \right\rangle \\
 &= \left\langle \Phi_\Gamma, T_\Gamma^a(A) \right\rangle \\
 &= \left\langle \tilde{\Phi}_\Gamma, A \right\rangle.
 \end{aligned}$$

Now it remains to see that

$$\#\text{Dec}(A_1) = \#\text{Dec}(A_2) + \#\text{Dec}(A_3).$$

Look at Fig.25 again, now considering the markings i, j, k . Each element of $\text{Dec}(A_1)$ determines either an element of $\text{Dec}(A_2)$, or an element of $\text{Dec}(A_3)$, as indicated by the picture, depending on whether $i < j$ or $i > j$. This separates $\text{Dec}(A_1)$ into two parts, respectively in bijection with $\text{Dec}(A_2)$ and $\text{Dec}(A_3)$, and terminates the proof. \square

4.3.2 There is a Whitney index for non nullhomotopic virtual knots

In the classical framework, the Whitney index is an invariant of regular plane curve homotopies which, together with the total writhe number, classifies the representatives of any given knot type up to regular isotopy. In other words, these invariants count the (algebraic) number of Reidemeister I moves that *have to* happen in a sequence of moves connecting two given diagrams. Here we describe such an invariant for *virtual* knot diagrams whose underlying curve on Σ is not homotopically trivial.

The classical Whitney index

Let $\delta : \mathbb{S}^1 \rightarrow \mathbb{R}^2$ be a smooth immersion (with non-vanishing differential). There is an associated Gauss map

$$\begin{array}{ccc}
 \Gamma : \mathbb{S}^1 & \rightarrow & \mathbb{S}^1 \\
 p & \mapsto & u_p(\delta) \quad ,
 \end{array}$$

where $u_p(\delta)$ is the unitary tangent vector to δ at the point p . It depends on a trivialization of the tangent space to \mathbb{R}^2 .

Definition 4.32 (usual Whitney index). The index of the above Gauss map only depends on the homotopy class of δ within the space of smooth immersions. It is called the *rotation number*, or *Whitney index*, of δ .

Given a projection $\mathbb{R}^3 \rightarrow \mathbb{R}^2$, a generic isotopy of a knot $S^1 \rightarrow \mathbb{R}^3$ is called a *regular isotopy* if the corresponding sequence of Reidemeister moves does not involve R-I. Clearly, the Whitney index of planar loops induces an invariant of regular isotopy classes of knots. In practice, it can be easily computed by looking at the Seifert circles of the projection (each of them contributes by $+1$ or -1). The total writhe number of a classical knot projection is also an invariant of regular isotopy. These two invariants together satisfy the following

Lemma 4.33 (see [12]). *Two equivalent knot diagrams are regularly equivalent if and only if their projections have the same total writhe and the same Whitney index.*

The proof (which we omit) relies essentially on the *Whitney trick* (see [12]) and the following variation table:

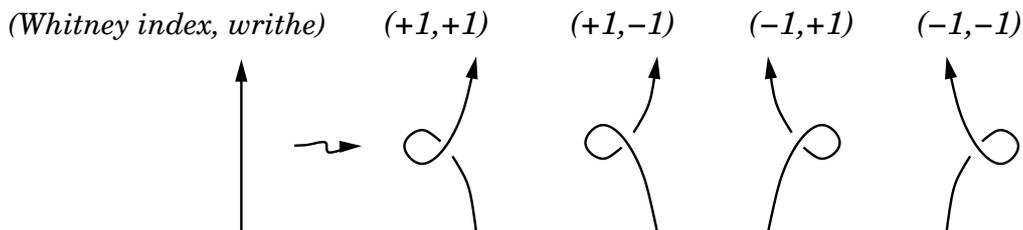


Figure 26: R-I moves sorted by their effect on the regular invariants

The virtual framework

We want to define a Whitney index for virtual knots, that satisfies a version of Lemma 4.33.

Given the virtual Reidemeister I moves, it does not seem reasonable to hope for counting the degree of a Gauss map. Relatedly, the Seifert circles are not embedded any more, and they do not have a well-defined contribution (at least not in the previous sense). Even when one looks only at real knot diagrams, the Whitney index is no more invariant when virtual moves are allowed: see Fig.27. In other words, though virtual moves do not connect knot diagrams that are not isotopic in the usual sense, they do add bridges between different regular isotopy classes.

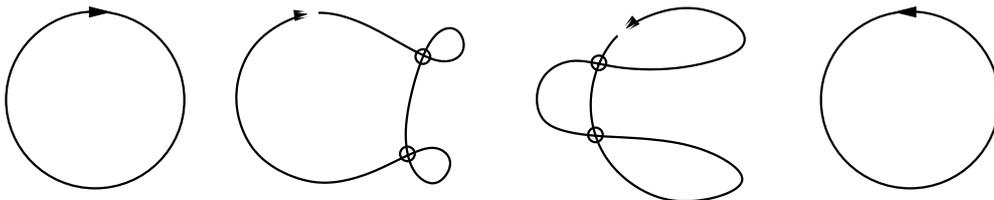


Figure 27: A “regular” sequence that changes the usual Whitney index

From now on, let Σ be an orientable surface with non trivial fundamental group.

Definition 4.34. Two virtual knot diagrams in Σ are called *regularly equivalent* if they are connected by a sequence of moves that does not involve the *real* Reidemeister I move.

In [13], regular equivalence is defined as equivalence under all moves but the real R-I *and* the virtual R-I too. Here, our goal is to define maps on the set of Gauss diagrams, so implicitly they must be invariant under all detour moves anyway.

Lemma 4.35. *The h_1 -decorated diagram series v_l and v_r from Fig.28 define invariants of regular equivalence. Moreover, two virtual knot diagrams from the same knot type with non trivial homotopy class are regularly equivalent as soon as both v_l and v_r coincide on them.*

$$v_l = \sum_{g \in h_1(\Sigma) \setminus \{1\}} \begin{array}{c} \text{---} 1 \text{---} \\ \text{---} g \text{---} \end{array} \quad v_r = \sum_{g \in h_1(\Sigma) \setminus \{1\}} \begin{array}{c} \text{---} g \text{---} \\ \text{---} 1 \text{---} \end{array}$$

Figure 28: The regular invariants for non nullhomotopic virtual knots

Proof. First assertion. First, notice that since Σ is orientable, $w = w_1(T\Sigma)$ is trivial and the w -moves never change the orientation of an arrow. Hence a w -orbit of planar Gauss diagrams on $\pi_1(\Sigma)$ determines an $h_1(\Sigma)$ -decorated diagram (see Definition 4.30). In restriction to the diagrams that happen in v_l and v_r , this forgetful map is actually a 1 – 1 correspondence. Hence v_l and v_r can be regarded as series of w -orbits of Gauss diagrams, so that they satisfy the w invariance criterion (Proposition 4.20). The invariance under R_2 and R_3 moves follows from Theorem 4.27.

Second assertion. The proof is essentially the same as that of Lemma 4.33. A little loop can run along a knot diagram without using R-I moves even where there are virtual crossings. The table from Fig.26 becomes, respectively (for the couple (v_l, v_r)):

$$(0, +1) \quad (-1, 0) \quad (+1, 0) \quad (0, -1).$$

This is the essential reason for which one needs the assumption that the knot diagrams are not nullhomotopic: otherwise the increase would be $(0, 0)$ for all R-I moves. \square

Finally, looking at the above table, which details how v_l and v_r control the R-I moves, it appears that:

1. $v_r + v_l$ behaves like the total writhe number under Reidemeister moves: it has the same “derivative”. It follows that these differ by a constant that depends only on the virtual knot type, *i.e.* a virtual knot invariant in the usual sense (see Fig.29).

$$\sum_{g, h \in \pi_1(\Sigma) \setminus \{1\}} \begin{array}{c} g \\ \text{---} \text{---} \text{---} \\ \text{---} h \end{array}$$

Figure 29: The difference between $v_r + v_l$ and the (usual) total writhe number

2. *In restriction to real knot diagrams*, $v_r - v_l$ has the same derivative as the Whitney index, with respect to Reidemeister moves. Hence there is an invariant of real knots c such that for every real knot diagram D , $v_r(D) - v_l(D) + c(D)$ is the usual Whitney index of D .

Hence we have proved:

Definition-Lemma 4.36. *Call $v_r - v_l$ the Whitney index of non nullhomotopic virtual knot diagrams, and call $v_r + v_l$ their writhe number. These make Lemma 4.33 hold for non nullhomotopic virtual knot diagrams.*

References

- [1] M. Brandenbursky. Link invariants via counting surfaces. *ArXiv e-prints*, September 2012.
- [2] Sergei Chmutov, Michael Cap Houry, and Alfred Rossi. Polyak-Viro formulas for coefficients of the Conway polynomial. *J. Knot Theory Ramifications*, 18(6):773–783, 2009.
- [3] Sergei Chmutov and Michael Polyak. Elementary combinatorics of the HOMFLYPT polynomial. *Int. Math. Res. Not. IMRN*, (3):480–495, 2010.

- [4] Micah W. Chrisman. Twist lattices and the Jones-Kauffman polynomial for long virtual knots. *J. Knot Theory Ramifications*, 19(5):655–675, 2010.
- [5] Micah W. Chrisman. On the Goussarov-Polyak-Viro finite-type invariants and the virtualization move. *J. Knot Theory Ramifications*, 20(3):389–401, 2011.
- [6] Thomas Fiedler. A small state sum for knots. *Topology*, 32(2):281–294, 1993.
- [7] Thomas Fiedler. *Gauss diagram invariants for knots and links*, volume 532 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.
- [8] Thomas Fiedler. Gauss diagram invariants for knots which are not closed braids. *Math. Proc. Cambridge Philos. Soc.*, 135(2):335–348, 2003.
- [9] Mikhail Goussarov. Finite-type invariants are presented by Gauss diagram formulas, 1998. Translated from Russian by O. Viro.
- [10] Mikhail Goussarov, Michael Polyak, and Oleg Viro. Finite-type invariants of classical and virtual knots. *Topology*, 39(5):1045–1068, 2000.
- [11] Sergei A. Grishanov and Victor A. Vassiliev. Fiedler type combinatorial formulas for generalized Fiedler type invariants of knots in $M^2 \times \mathbf{R}^1$. *Topology Appl.*, 156(14):2307–2316, 2009.
- [12] Louis H. Kauffman. *On knots*, volume 115 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1987.
- [13] Louis H. Kauffman. Virtual knot theory. *European J. Combin.*, 20(7):663–690, 1999.
- [14] Jean Lannes. Sur les invariants de Vassiliev de degré inférieur ou égal à 3. *Enseign. Math. (2)*, 39(3-4):295–316, 1993.
- [15] Arnaud Mortier. Finite-type 1-cocycles of knots and virtual knots given by Polyak-Viro formulas. In preparation.
- [16] Arnaud Mortier. Gauss diagrams of real and virtual knots in the solid torus. *ArXiv e-prints*, January 2012.
- [17] Arnaud Mortier. Polyak type equations for virtual arrow diagram invariants in the annulus. *Journal of Knot Theory and Its Ramifications*, 22(07):1350034, 2013.
- [18] Michael Polyak. Talk at Swiss Knots 2011 – “3 stories about $\nabla = \nabla + \overleftarrow{\nabla}$ ”. http://drorbn.net/dbnvp/SK11_Polyak.php. Videography by Pierre Dehornoy and Dror Bar-Natan.
- [19] Michael Polyak. On the algebra of arrow diagrams. *Lett. Math. Phys.*, 51(4):275–291, 2000.
- [20] Michael Polyak and Oleg Viro. Gauss diagram formulas for Vassiliev invariants. *Internat. Math. Res. Notices*, (11):445ff., approx. 8 pp. (electronic), 1994.
- [21] Michael Polyak and Oleg Viro. On the Casson knot invariant. *J. Knot Theory Ramifications*, 10(5):711–738, 2001. Knots in Hellas '98, Vol. 3 (Delphi).