FINITE TYPE INVARIANTS ARE PRESENTED BY GAUSS DIAGRAM FORMULAS

M. GOUSSAROV

1. **Introduction.** This paper was inspired by a question posed by M.Polyak and O.Viro [1]: "Can any Vasiliev invariant be calculated as function of arrow polynomials evaluated on the knot diagram?" The Main Theorem 5 below answers to this question in positive.

The author is grateful to O.Viro for useful discussions and translation of this text.

2. **Long knots.** By a *long knot* we mean a smooth embedding $K : \mathbb{R} \to \mathbb{R}^3$. It is presented by equations

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}$$

such that $x(t) = t$, $y(t) = z(t) = 0$ for sufficiently large $|t|$. In verbal discriptions of mutual position of objects we assume that the z-axis is directed upwards, and in $\mathbb{R}$ the direction to the right corresponds to increasing of t.

Isotopy of long knots is a smooth isotopy in the class of embeddings above. It can be extended to an ambient isotopy of $\mathbb{R}^3$. Denote by $\mathcal{K}$ the set of isotopy classes of long knots.

By a *knot diagram* we call an orthogonal projection of the long knot to the xy-plane enhanced with the information on the order of the branches in the vertical direction at double points. The knot is supposed to be in a general position. Then the projection is a generic immersion: it has no points of multiplicity ≥ 2, and at each double point the branches are transversal to each other. A knot diagram defines the knot up to isotopy. In the classical knot theory long knots are not as standard main objects as the (closed) knots, which are just smooth simple closed curves in $\mathbb{R}^3$. However the isotopy classes of closed knots are in a natural one-to-one correspondence with the isotopy classes of long knots. For technical reasons, long knots are more convenient in the framework of the problem under consideration.

3. **Gauss Diagrams.** A Gauss diagram of a long knot is in a sence a pre-image of its diagram under the map parametrizing the knot. Namely, on the t-axis one marks the pairs of points which are mapped to double points of the knot diagram, and puts the information about the corresponding crossing points. For this, the points of each of these pairs are connected with an arc. The arcs are equipped with the direction from the point mapped to the upper branch towards the point mapped to the lower one. Also, the arcs are equipped with signs, which are equal to the local writhe number of the corresponding crossing point. This sign is $+$, if

1991 *Mathematics Subject Classification.* 57M25.

Key words and phrases. classical knot, Gauss diagram, finite type invaraint, Vassiliev invariant.

in the knot diagram the low branch crosses the upper one from right to left, and $-$, otherwise. Each oriented signed arc is called an *arrow*. From the formal point of view, an arrow stands for a pair of points with all the information above (the ordering from the upper to the lower branch and the writhe number).

Gauss diagram is usually considered up to isotopy of the underlying line. The isotopy classes of Gauss diagrams are in one-to-one correspondence with knot diagrams considered up to isotopy of the plane.

One can position arrows on the line in such a way that the resulting diagram would not correspond to a knot. A diagram of this sort is also called a Gauss diagram. A Gauss diagram, which can be obtained from a long knot, is said to be *real*.

A Gauss diagram D_1 is said to be a *subdiagram* of a Gauss diagram D_2 if all the arrows of D_1 are arrows of D_2. It is easy to see that any Gauss diagram is a subdiagram of a real Gauss diagram.

4. **Finite Type Invariants.** Let G be an abelian group and $\nu : \mathcal{K} \to G$. When we say about the value of ν on a Gauss diagram D, we mean the value of ν on the knot defined by D. We say that ν is a finite type invariant of degree $\leq n$, if for any diagram D of a long knot and any $n+1$ crossing points of D

$$\sum_{D_i} (-1)^{n_i} \nu(D_i) = 0,$$

where D_i runs over 2^{n+1} diagrams which differ from D only in overpasses and underpasses at some of the chosen crossings, and n_i is the number of the crossing points at which D_i differs from D.

In terms of Gauss diagrams, switching of an overpass and underpass corresponds to simultaneous reversing of the orientation and sign of the corresponding arrow. Note that this operation cannot make a real Gauss diagram unreal.

5. **Main Theorem.** *Let $\nu : \mathcal{K} \to G$ be an invariant of degree n of long knots. Then there exists a G-valued function k_ν defined on the set of Gauss diagrams such that*

$$\nu(K) = \sum_C k_\nu(C) \tag{1}$$

where C runs over all the subdiagrams with at most n arrows of a Gauss diagram of K.

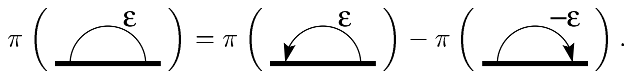
Corollary. *Any integer valued finite type invariant of degree n of long knots can be presented as $\langle A, \cdot \rangle$, where A is a linear combination of arrow diagrams on a line with at most n arrows.*

6. **Mixed Diagrams.** To prove Main Theorem, we need one more class of diagrams. A *mixed diagram* is what can be obtained from a Gauss diagram by forgetting directions of several arrows. The arcs with signs, which are obtained from the arrows, are called *chords*. There are to admissible ways of converting a chord into an arrow:

- To direct it to the left and preserve its sign;
- To direct it to the right, but reverse the sign.

A mixed diagram is said to be *real* if an admissible orientation of all the chords gives rise to a real Gauss diagram. This does not depend on the choice of orientations, provided it is made in an admissible way.

Let $\mathcal{B}$ be a free abelian group generated by mixed diagrams. Let $\mathcal{B}_{re}$ be its subgroup generated by real diagrams and $\mathcal{B}_G$ the subgroup generated by Gauss diagrams. Define a canonical projection $\pi : \mathcal{B} \to \mathcal{B}_G$ by formula



Here ε is a sign and $-\varepsilon$ the opposite sign. We assume that each of the diagrams contains an unshown set of arrows and chords, which is arbitrary but the same for each of the three diagrams.

Any finite type knot invariant ν gives rise to a homomorphism defined on $\mathcal{B}_{re} \cap \mathcal{B}_G$. One can easily check that $\mathcal{B}_{re} \cap \mathcal{B}_G = \pi(B_{re})$. Extend ν over $\mathcal{B}_{re}$ by $\nu(D) = \nu(\pi(D))$. It is clear that if the degree of ν is n then ν vanishes on real mixed diagrams with $> n$ chords.

A diagram C is called a *subdiagram* of a mixed diagram D if C consists of any collection of arrows and all the chords of D. Denote by S the homomorphism $\mathcal{B} \to \mathcal{B}$ mapping each diagram to the sum of all its subdiagrams.

7. **Construction of k_ν.** Let ν be a knot invariant of degree n. In Section 9 below we construct homomorphisms $P, Q : \mathcal{B} \to \mathcal{B}$, which depend only on n and have the following properties:

(A) $P(\mathcal{B}_{re}) \subset \mathcal{B}_{re}$ and $\nu \circ P|_{\mathcal{B}_{re}} = 0$.
(B) Diagram D is said to be *final* if $P(D) = D$. For any mixed diagram D there exists p such that $P^p(D)$ and $Q^p(D)$ are sums of final diagrams.
(C) If the minimal subdiagrams of real final diagrams D_1 and D_2 coincide then $\nu(D_1) = \nu(D_2)$. Note that the minimal subdiagram of a diagram D is the diagram consisting of all chords of D.
(D) For a diagram D, the total number of chords and arrows in each of the diagrams involved in $Q(D)$ is at least the total number of chords and arrows in D.
(E) The set of subdiagrams of a diagram D can be splitted to classes U and V such that $S(P(D)) = \sum_{C \in U} Q(C) + \sum_{C \in V} C$ and $D \in U$.

Now define ν on a subgroup of B generated by all the final diagrams. This can be done so that

- the value on a real final diagram is the same as above,
- the values of ν on final diagrams with equal minimal subdiagrams coincide.

This is possible due to (C).

Now define ν on the whole $\mathcal{B}$. Put

$$\nu(D) = \nu(P^p(D))$$

for any diagram D (recall, that $\mathcal{B}$ is generated by all the mixed diagrams). Here p is the number from (B). Property (A) implies that this agrees with the construction for $\nu|_{\mathcal{B}_{re}}$ given in Section 6. We see that $\nu \circ P = \nu$, since $P^p(D) = P^{p+1}(D)$.

Define $k_\nu(D)$ for a mixed diagram D by induction on the number of arrows in D keeping the relation

$$\nu(D) = \sum_C k_{\nu(C)} \tag{2}$$

where C runs over all the subdiagrams of D. This can be done also by the following explicit formula:

$$k_\nu = \sum_C (-1)^{a(D)-a(C)} k_\nu(C)$$

where C runs over all the subdiagrams of D as above and $a(\cdot)$ is the number of arrows.

8. **Proof of Main Theorem.** The only difference between formulas (1) and (2) is that the former does not contain summands corresponding to the subdiagrams with more than n arrows. Let us prove that for diagrams of this kind $k_\nu = 0$.

Extend k_ν to a homomorphism $\mathcal{B} \to G$. Then (2) can be rewritten as $\nu(D) = k_\nu(S(D))$. Let us prove that $k_\nu(Q(D)) = k_\nu(D)$ by induction on the number of arrows in D. If the number equals -1, there is nothing to check. Suppose that the statement is correct for any diagram whose number of arrows is less then the number of arrows in D and prove the statement for D. Apply k_ν to the relation from property (E):

$$\sum_{C\in U} k_\nu(Q(C)) + \sum_{C\in V} k_\nu(C) = k_\nu(S(P(D))) = \nu(P(D)) = \nu(D) = \\ \sum_{C\in U} k_\nu(C) + \sum_{C\in V} k_\nu(C).$$

From this equality, by cancelling out the second $\sum$'s and those summands of the first ones which are equal by the induction assumption, we obtain the desired equality $k_\nu(Q(D)) = k_\nu(D)$. This completes the induction step.

Let D be a final diagram. Then $\nu(C) = \nu(D)$ for any subdiagram C of D, due to construction of ν. One can see that k_ν takes value $\nu(D)$ on the minimal subdiagram of D and vanishes on all other subdiagrams. Hence $k_\nu(D)$ can be non-trivial only if D has no arrow. On the other hand, if D has more than n chords, then $k_\nu(D) = \nu(D) = 0$. Therefore, if the total number of chords and arrows in a final diagram D is greater than n then $k_\nu(D) = 0$.

Let a Gauss diagram D have more than n arrows and p be the number which exists by (B). Since $k_\nu \circ Q = k_\nu$, we have also $k_\nu(D) = k_\nu(Q^p(D))$. By (D), the expansion of $Q^p(D)$ contains only diagrams with the total number of chords arrows greater than n. By (B) each of these diagrams is final. Then k_ν takes zero value on them. Hence $k_\nu(D) = 0$. □

9. **Construction of P and Q.** We define P and Q on a diagram D. If D has more than n chords, put $P(D) = Q(D) = 0$. Otherwise, we inspect D from left to right untill we meet the first "defect." If this does not happen, we put again $P(D) = Q(D) = D$. By "defects" we mean:

1. the arrowhead of an arrow directed to the left,
2. the left end point of a chord, whose immediate neighbor from the left is an end point of an arrow.

In the first case put $P(D) = Q(D) = D_1 + D_2$, where D_1 is obtained from D by reversing the direction and sign of the "defective" arrow and D_2 is obtained from D by replacing this arrow by a chord with the same sign.

In the second case $P(D)$ is defined according to one of the patterns of Figure 1. Put $Q(D)$ to be the sum of all the subdiagrams of $P(D)$ which contain all the unshown arrows and at least one of the shown ones.

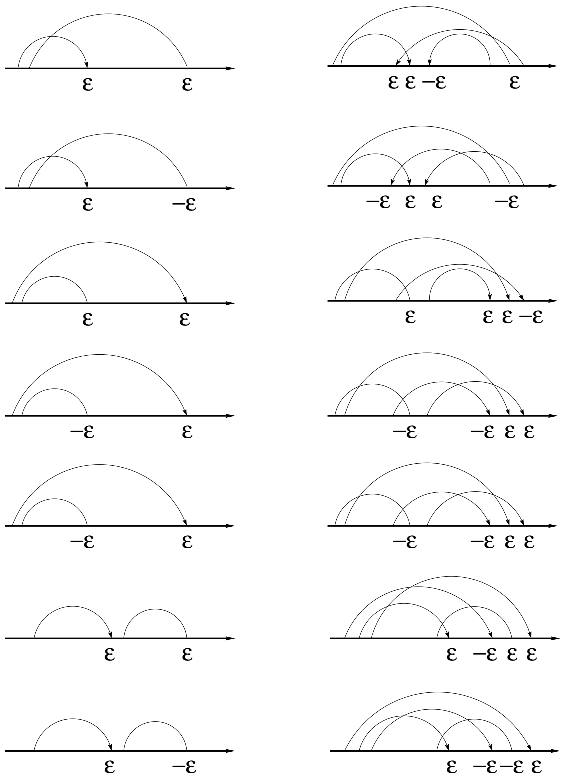


FIGURE 1

10. **Proof of (A).** Recall that we have to prove that $P(\mathcal{B}_{re}) \subset \mathcal{B}_{re}$ and $\nu \circ P|_{\mathcal{B}_{re}} = 0$.

Consider the first case of the construction (see the preceding section). The fact that D_1 is real, was observed in the last sentence of Section 4. Diagram D_2 is real by definition, see Section 6. Equality $\nu \circ P|_{\mathcal{B}_{re}} = 0$ follows from the definition of $\pi(D)$, see also Section 6.

In the second case of the construction, convert chords of D and $P(D)$ to arrows in the same way. Then the knot presented by $P(D)$ can be obtained from the knot presented by D by a combination of the second and third Reidemeister moves. This combination corresponds to dragging of the crossing point represented by the chord under or over one of the branches related to an arrow of D shown in the Figure 1. Therefore the knots are isotopic, diagram $P(D)$ is real and $\nu(P(D)) = \nu(D)$.

11. **Proof of (B).** Let $l(D)$ be the number of left end points of chords of D which are to the left with respect to the first "defect". As easy to see, $l(D') \geq l(D)$ for each diagram in the expansion of both $P(D)$ and $Q(D)$. However the number of the left end points of chords (as well as the number of chords) in a nonfinal dagram is at most n. Therefore it suffices to prove that the diagram cannot change infinitely many times in subsequent iterations of P or Q without changing l.

Consider the number of arrowheads on the ray to the left from $(l(D)) + 1$-th left end point of chords. For any diagram involved into the expansion of $P(D)$ or $Q(D)$

it is not greater than for D. If it is the same for one of these diagrams, then it has less arrowtails on the same ray. This can happen only finitely many times.

12. **Proof of (C).** Let ν be an invariant of degree $\leq n$ and D_1, D_2 real final diagrams with the same collection of chords. If each of them has more than n chords, then $\nu(D_1) = \nu(D_2) = 0$. If there is no chords, the diagrams represent descending and hence trivial knots.

In other cases convert all the chords to arrows in D_1 and D_2 in the same way. It suffices to prove that for any conversion like this the resulting diagrams represent the same knot. To prove this, consider the knot diagrams corresponding to these Gauss diagrams. Mark the crossing points originated from the chords of D_1 and D_2.

Since D_1 and D_2 are final, they do not have an end point of an arrow to the left from the leftmost end point of a chord. The left end point of a chord, which is not the leftmost, follows immediately after other end point of a chord. Select on the knot diagrams the following parts of the knot:

1. the part of the string to the left of the first crossing point (this crossing is marked),
2. a small neighborhood of all the crossing points,
3. the arcs between crossing points along which the knot arrives for the first time to a marked point (these arcs start also at marked points).

On the plane the selected arcs comprise two embedded trees rooted at the beginning of the string. The arcs contain only marked crossings. The marked crossings corresponding to each other are organized in the same way. Therefore the selected trees can be transformed to each other by a plane isotopy with compact support. Let us make them coinciding by this kind of isotopy. Then make the corresponding pieces of the knots coinciding in $\mathbb{R}^3$ by an isotopy preserving the xy coordinates.

Keeping change only z, we make $z(t) = x(t) - t$ everywhere besides the selected part of the knots. There is no obstruction for achieving this, because in the corresponding part of the Gauss diagrams all the arrows are directed to right, i.e. at a crossing point the smaller value of z corresponds to the larger t.

Now connect by a homotopy the projection of a nonselected arc of one of the knots to the projection of its counter-part in another knot. Doing this we make sure that

- the end points of the arcs were fixed and there connection with the selected part was kept smooth,
- the points corresponding to suficiently large values of t were fixed,
- the image of the moving arcs did not go over the projection of the selected part.

This is possible, since the complement of the projection of the selected part is contracible (it is a complement of a tree rooted at infinity). The homotopy can be lifted to an isotopy between the knots such that it is fixed on the selected parts and on the rest the inequality $z(t) = x(t) - t$ holds true. It is this equality that ensures absence of self-intersections during the homotopy.

13. **Proof of (D) and** (E)**.** The definition of Q obviously implies (D).

If a subdiagram has a "defect" then the diagram also has it. The diagram and subdiagram have the same number of chords. Therefore if $P(D) = D$ then the same is true for any subdiagram of D and hence (E) holds true.

Consider the first case of the construction for P and Q. Include into U the subdiagram which contain the arrow that is transformed. The second summands in $Q(C)$ for $C \in U$ comprise $S(D_2)$ (here we use the notations of Section 9). The first summands in $Q(C)$ are the subdiagrams of $P(D)$ which contain a new arrow. The subdiagrams which do not contain this arrow, comprise V.

Consider the second case. Compose U of the subdiagrams which contains the arrow shown on the left hand side in Figure 1. These subdiagrams contain the same "defect" as the whole diagram D. Therefore $\sum_{C \in U} Q(C)$ is the sum of all the subdiagrams of $P(D)$ containing at least one of the arrows shown in the right hand side of Figure 1. Subdiagrams which do not contain the arrows shown are the same in D and $P(D)$. They comprise V.

References

[1] M.Polyak and O.Viro, *Gauss diagram formulas for Vassiliev invariants,* International Math. Research Notices, (1994), No. 11, 445–453

Translated by O.VIRO

POMI, Fontanka 27, St. Petersburg, 191011, Russia
E-mail address: `goussar@@pdmi.ras.ru`