FUNDAMENTA
MATHEMATICAE
214 (2011)

Homotopy classification of nanophrases with at most four letters

by

Tomonori Fukunaga (Sapporo)

Abstract. We give a homotopy classification of nanophrases with at most four letters. It is an extension of the classification of nanophrases of length 2 with at most four letters, given by the author in a previous paper. As a corollary, we give a stable classification of ordered, pointed, oriented multi-component curves on surfaces with minimal crossing number less than or equal to 2 such that any equivalent curve has no simply closed curves in its components.

1. Introduction. The study of curves via words was initiated by C. F. Gauss [2]. He encoded closed planar curves by words of a certain type, which are now called Gauss words. We can apply this method to encode multi-component curves and links on surfaces. For instance, in [8] and [9] V. Turaev studied stable equivalence classes of curves and links on surfaces by using generalized Gauss words called nanowords (stable equivalence classes of curves and links on surfaces are closely related to the theory of virtual strings and virtual links).

A *knot* is the image of a smooth embedding of S^1 into $\mathbb{R}^3$. Further, a k-component *link* is the image of a smooth embedding of the disjoint union of k circles into $\mathbb{R}^3$. In the theory of knots, we study isotopy classes of knots. When we study knots and links, we often use link diagrams of links. A *knot diagram* is a smooth immersion of S^1 into $\mathbb{R}^2$ with transversal double points such that at each double point one of the two paths is declared to be the overpath and the other the underpath (we call a double point of such an immersion a *crossing*). If a knot diagram D is obtained as the image of a knot by a projection of $\mathbb{R}^3$ to $\mathbb{R}^2$, then we call D a diagram of the knot. A *link diagram* is defined similarly as a smooth immersion of a disjoint union of circles to $\mathbb{R}^2$.

2010 *Mathematics Subject Classification*: Primary 57M27; Secondary 55P99.
Key words and phrases: nanophrases, homotopy, multi-component curves, stable equivalent.
The author is a JSPS research fellow (DC).

DOI: 10.4064/fm214-2-1

In 1999, L. Kauffman introduced the theory of virtual knots and links using combinatorially extended link diagrams called virtual link diagrams. A *virtual knot diagram* is a planar graph of valency four endowed with the following structure: each vertex either has an overcrossing and undercrossing (in other words, *real crossing*) or is marked by a *virtual crossing*. A *virtual link diagram* is defined similarly. Then we define *virtual links* by a set of virtual link diagrams quotiented by an equivalence relation generated by the virtual Reidemeister moves (see [5] for more details).

In [5], Kauffman studied Gauss codes of virtual knots and links, fundamental groups, crystals, racks and quandles of virtual knots and links, and quantum invariants of virtual knots and links.

V. Turaev extended the theory of (pointed ordered) virtual knots and links in the aspect of Gauss codes in [7] and [8]. More precisely a *nanoword* over an alphabet α endowed with an involution $\tau : \alpha \to \alpha$ is a word on an alphabet $\mathcal{A}$ endowed with a projection $\mathcal{A} \ni A \mapsto |A| \in \alpha$ such that every letter appears twice or not at all. In the case where the alphabet α consists of two elements permuted by τ, the notion of a nanoword over α is equivalent to the notion of an open virtual string introduced in [10].

Turaev introduced homotopy equivalence on the set of nanowords over α. For a fixed subset $S \subset \alpha \times \alpha \times \alpha$, the homotopy equivalence relation is generated by three types of moves on nanowords. The first move consists of deleting two consecutive appearances of the same letter. The second move has the form $xAByBAz \mapsto xyz$ where x, y, z are words and A, B are letters such that $|A| = \tau(|B|)$. The third move has the form $xAByACzBCt \mapsto xBAyCAzCBt$ where x, y, z, t are words and A, B, C are letters such that $(|A|, |B|, |C|) \in S$. These moves are suggested by the three local deformations of curves on surfaces (see Fig. 1 and [8] for more details).

In [8] Turaev showed that a stable equivalence class of an oriented pointed curve on a surface can be identified with the homotopy class of a nanoword in a 2-letter alphabet. Moreover, Turaev showed that the stable equivalence class of an oriented pointed knot diagram on a surface can be identified with the homotopy class of a nanoword on a 4-letter alphabet for some S. Moreover Turaev extended this result to multi-component curves. In fact the stable equivalence class of an oriented, ordered, pointed multi-component curve (respectively link diagram) on a surface is identified with the homotopy class of a nanophrase on a 2-letter (respectively 4-letter) alphabet for some S. Roughly speaking, a nanophrase is a sequence of words where concatenation of those words is a nanoword (see also Subsection 3.2 and Section 4 for more details). Thus, using Turaev's theory of words and phrases, we can treat curves on surfaces algebraically.

Any results in Turaev's theory of words yield results on multi-curves and links. For example, if we classify nanophrases up to S-homotopy, then we obtain a classification of ordered, pointed, oriented multi-component curves on surfaces and knot diagrams as a corollary. Furthermore, if we could find other correspondences between the theory of phrases and geometric objects, then we could find common properties of multi-curves, links and other geometric (or knot-theoretic) objects via the theory of words and phrases. Therefore, Turaev's theory of words and phrases can be expected to have a lot of applications.

A homotopy classification of nanowords was given by Turaev in [7]. He classified nanowords of at most six letters. The present author [1] introduced new invariants of nanophrases and gave a homotopy classification of nanophrases of length 2 with at most four letters, using Turaev's classification of nanowords.

The purpose of this paper is to give a classification of nanophrases over an arbitrary alphabet with at most four letters but with no condition on the length. As a corollary, we classify the multi-component curves with minimum crossing number at most 2 which have no "untidy" components up to stable equivalence (Theorem 2.2).

The organization of this paper is as follows. In Sections 2–4 we review the theory of multi-component curves and the homotopy theory of words and phrases. In Section 5 we recall known results on the classification of nanowords and nanophrases up to homotopy and we generalize them to phrases of an arbitrary length. Finally in Section 6 we give the proof of the main theorem of this paper.

2. Stable equivalence of multi-component curves

2.1. Multi-component curves. In this paper a *curve* means the image of a generic immersion of an oriented circle into an oriented surface. The word "generic" means that the curve has only a finite set of self-intersections which are all double and transversal. A *k-component curve* is defined in the same way, with the difference that it may be formed by k curves rather than only one curve. These curves are the *components* of the k-component curve. A k-component curve is *pointed* if each component is endowed with a base point (the origin) distinct from the crossing points. A k-component curve is *ordered* if its components are numbered. Two ordered, pointed curves are *stably homeomorphic* if there is an orientation preserving homeomorphism of their regular neighborhoods in the ambient surfaces mapping the first multi-component curve onto the second and preserving the order, the origins, and the orientations of the components.

Now we define stable equivalence of ordered, pointed multi-component curves [4]: Two ordered, pointed multi-component curves are *stably equivalent* if they can be related by a finite sequence of the following transformations: (i) replacing an ordered, pointed multi-component curve with a stably homeomorphic one; (ii) deformation of a pointed curve in its ambient surface away from the origin (such a deformation may push a branch of the multi-component curve across another branch or a double point but not across the origin of the curves) as in Fig. 1.

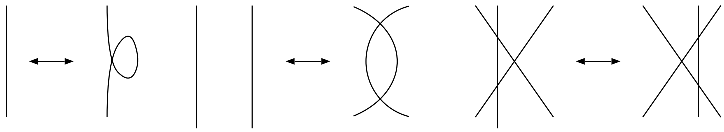

Fig. 1. Three local deformations of curves

We denote by $\mathcal{C}_k$ the set of stable equivalence classes of ordered, pointed k-component curves.

Remark 2.1. The theory of stable equivalence of multi-component curves on surfaces is closely related to the theory of virtual strings. See [3] and [10] for more details.

We will show the following theorem by using Turaev's theory of words.

An ordered, pointed multi-component surface-curve is called *irreducible* if it is not stably equivalent to a surface-curve with a simply closed component.

Theorem 2.2. *Any irreducible ordered, pointed multi-component surface-curve with minimal crossing number less than or equal to* 2 *is stably equivalent to one of the ordered, pointed multi-component curves arising as in also Remark* 2.3 *below. There are exactly* 52 *stable equivalence classes of irreducible ordered, pointed, multi-component surface-curves.*

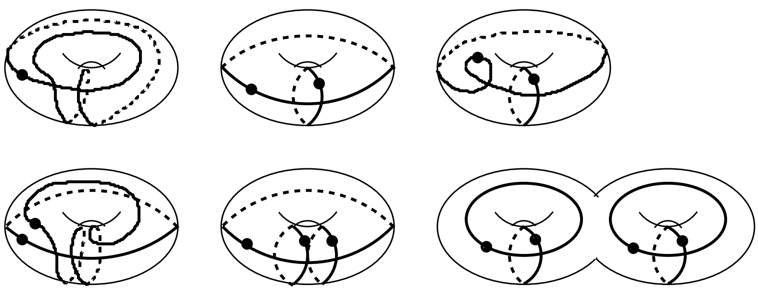

Fig. 2. The list of curves

Remark 2.3. To list the stable equivalence classes of irreducible ordered, pointed multi-component surface-curves with minimal crossing number ≤ 2, in Fig. 2 we just list the multi-component curves without order and orientation of the components. Two different pictures from Fig. 2 never produce equivalent ordered, pointed multi-component surface-curves. On the other hand, two different additional structures (orientation and order) on the same picture may yield equivalent ordered, pointed multi-component surface-curves. More precisely, 2, 2, 8 (respectively 4, 24, 12) different ordered, pointed multi-component surface-curves arise from the upper (respectively lower) row. By Theorem 5.15, two ordered, pointed multi-component surface-curves arising from pictures in Fig. 2 are stably equivalent if and only if the associated nanophrases are homotopic, and we can obtain all the stable equivalence classes of irreducible ordered, pointed multi-component surface-curves with minimal crossing number ≤ 2 by specifying order and orientation for multi-component curves in Fig. 2.

To prove Theorem 2.2, we use Turaev's [7, 8] theory of words and phrases.

3. Turaev's theory of words and phrases. In this section we review the topology of words and phrases.

3.1. Nanowords and their homotopy. An *alphabet* is a set and *letters* are its elements. A *word of length* $n \geq 1$ *on an alphabet* $\mathcal{A}$ is a mapping $w : \hat{n} \to \mathcal{A}$ where $\hat{n} = \{1, \ldots, n\}$. We denote a word of length n by the sequence of letters $w(1) \cdots w(n)$. A word $w : \hat{n} \to \mathcal{A}$ is a *Gauss word* if each element of $\mathcal{A}$ is the image of precisely two elements of $\hat{n}$.

For a set α, an α-*alphabet* is a set $\mathcal{A}$ endowed with a mapping $\mathcal{A} \to \alpha$ called *projection*. The image of $A \in \mathcal{A}$ under this mapping is denoted $|A|$. An *étale word* over α is a pair (an α-alphabet $\mathcal{A}$, a word on $\mathcal{A}$). A *nanoword over* α is a pair (an α-alphabet $\mathcal{A}$, a Gauss word on $\mathcal{A}$). We call an empty étale word in an empty α-alphabet the *empty nanoword*. It is written $\emptyset$ and has length 0.

A *morphism* of α-alphabets $\mathcal{A}_1$, $\mathcal{A}_2$ is a set-theoretic mapping $f : \mathcal{A}_1 \to \mathcal{A}_2$ such that $|A| = |f(A)|$ for all $A \in \mathcal{A}_1$. If f is bijective, then this morphism is an *isomorphism*. Two étale words $(\mathcal{A}_1, w_1)$ and $(\mathcal{A}_2, w_2)$ over α are *isomorphic* if there is an isomorphism $f : \mathcal{A}_1 \to \mathcal{A}_2$ such that $w_2 = f \circ w_1$.

To define homotopy of nanowords we fix a finite set α with an involution $\tau : \alpha \to \alpha$ and a subset $S \subset \alpha \times \alpha \times \alpha$. We call the pair (α, S) *homotopy data*.

Definition 3.1. Let (α, S) be homotopy data. We define *homotopy moves* (1)–(3) as follows:

(1) $(\mathcal{A}, xAAy) \to (\mathcal{A} \setminus \{A\}, xy)$ for all $A \in \mathcal{A}$ and x, y words in $\mathcal{A} \setminus \{A\}$ such that xy is a Gauss word.
(2) $(\mathcal{A}, xAByBAz) \to (\mathcal{A} \setminus \{A, B\}, xyz)$ if $A, B \in \mathcal{A}$ satisfy $|B| = \tau(|A|)$ and x, y, z are words in $\mathcal{A} \setminus \{A, B\}$ such that xyz is a Gauss word.

(3) $(\mathcal{A}, xAByACzBCt) \to (\mathcal{A}, xBAyCAzCBt)$ if $A, B, C \in \mathcal{A}$ satisfy $(|A|, |B|, |C|) \in S$ and x, y, z, t are words in $\mathcal{A}$ such that $xyzt$ is a Gauss word.

Definition 3.2. Let (α, S) be homotopy data. Then nanowords $(\mathcal{A}_1, w_1)$ and $(\mathcal{A}_2, w_2)$ over α are *S-homotopic* (denoted $(\mathcal{A}_1, w_1) \simeq_S (\mathcal{A}_2, w_2)$) if $(\mathcal{A}_2, w_2)$ can be obtained from $(\mathcal{A}_1, w_1)$ by a finite sequence of isomorphisms, S-homotopy moves (1)–(3) and their inverses.

The set of S-homotopy classes of nanowords over α is denoted $\mathcal{N}(\alpha, S)$.

To define S-homotopy of étale words we define *desingularization* of étale words $(\mathcal{A}, w)$ over α as follows: Set $\mathcal{A}^d := \{A_{i,j} := (A, i, j) \mid A \in \mathcal{A}, 1 \leq i < j \leq m_w(A)\}$ with projection $|A_{i,j}| := |A| \in \alpha$ for all $A_{i,j}$ (where $m_w(A) := \mathrm{Card}(w^{-1}(A))$). The word w^d is obtained from w by first deleting all $A \in \mathcal{A}$ with $m_w(A) = 1$; then for each $A \in \mathcal{A}$ with $m_w(A) \geq 2$ and each $i = 1, \ldots, m_w(A)$, we replace the ith entry of A in w by

$$A_{1,i}A_{2,i}\ldots A_{i-1,i}A_{i,i+1}A_{i,i+2}\ldots A_{i,m_w(A)}.$$

The resulting $(\mathcal{A}^d, w^d)$ is a nanoword of length $\sum_{A\in\mathcal{A}} m_w(A)(m_w(A)-1)$ and called the *desingularization* of $(\mathcal{A}, w)$. Then we define S-homotopy of étale words as follows:

Definition 3.3. Let w_1 and w_2 be étale words over α. Then w_1 and w_2 are *S-homotopic* if w_1^d and w_2^d are S-homotopic.

3.2. Nanophrases and their homotopy. In [8], Turaev used similar arguments for phrases (sequences of words).

Definition 3.4. A *nanophrase* $(\mathcal{A}, (w_1|\cdots|w_k))$ of length $k \geq 0$ over a set α is a pair consisting of an α-alphabet $\mathcal{A}$ and a sequence of k words $w_1, \ldots, w_k$ on $\mathcal{A}$ such that $w_1 \ldots w_k$ is a Gauss word on $\mathcal{A}$. We denote it simply by $(w_1|\cdots|w_k)$.

By definition, there is a unique *empty nanophrase* of length 0 (the corresponding α-alphabet $\mathcal{A}$ is an empty set).

Remark 3.5. We can consider a nanoword w to be a nanophrase (w) of length 1.

A mapping $f : \mathcal{A}_1 \to \mathcal{A}_2$ is an *isomorphism* of two nanophrases if f is an isomorphism of α-alphabets transforming one nanophrase into the other.

Given homotopy data (α, S), we define homotopy moves on nanophrases as in Section 3.1 with the only difference that the 2-letter subwords AA, AB, BA, AC and BC modified by these moves may occur in different words of the phrase. Isomorphism and homotopy moves generate the equivalence relation $\simeq_S$ of S-homotopy on the classes of nanophrases over α. We denote the set of S-homotopy classes of nanophrases of length k by $\mathcal{P}_k(\alpha, S)$.

4. Nanophrases versus multi-component curves. In [8], Turaev showed that a special case of homotopy theory of nanophrases is equivalent to the study of $\mathcal{C}_k$. More precisely, he showed the following theorem.

THEOREM 4.1 (Turaev [8]). *Let α_0 be the set $\{a, b\}$ with involution $\tau : \alpha_0 \to \alpha_0$ permuting a and b, and S_0 the diagonal of $\alpha_0 \times \alpha_0 \times \alpha_0$. Then there is a canonical bijection of $\mathcal{C}_k$ and $\mathcal{P}_k(\alpha_0, S_0)$.*

The method of making a nanophrase $P(C)$ from an ordered, pointed k-component curve C is as follows. Let us label the double points of C by distinct letters $A_1, \ldots, A_n$. Starting at the origin of the first component of C and following C in the positive direction, we write down the labels of double points which we pass until the return to the origin. Thus we obtain a word w_1. Similarly we obtain words $w_2, \ldots, w_k$ on the alphabet $\mathcal{A} = \{A_1, \ldots, A_n\}$ from the second component, ..., k-th component. Let t_i^1 (respectively, t_i^2) be the tangent vector to C at the double point labeled A_i appearing at the first (respectively, second) passage through this point. Set $|A_i| = a$ if the pair (t_i^1, t_i^2) is positively oriented, and $|A_i| = b$ otherwise. Thus we obtain the required nanophrase $P(C) := (\mathcal{A}, (w_1| \cdots |w_k))$.

By the above theorem, if we classify the homotopy classes of nanophrases, then we obtain a classification of ordered, pointed multi-component curves under stable equivalence as a corollary.

REMARK 4.2. In [6], D. S. Silver and S. G. Williams studied open virtual multi-strings. The theory of open virtual multi-strings is equivalent to the theory of pointed multi-component surface-curves. Silver and Williams constructed invariants of open virtual multi-strings.

5. Classification of nanophrases. In this section, we give a homotopy classification of nanophrases with at most four letters under the assumption that the homotopy data S is the diagonal. We keep this assumption in the remaining part of the paper. Note that it does not restrict generality.

5.1. The case of nanophrases of length 1. In the case of nanophrases of length 1 (in other words, of nanowords), Turaev gave the following classification theorem.

THEOREM 5.1 (Turaev [7]). *Let w be a nanoword of length 4 over α. Then w is either homotopic to the empty nanoword or isomorphic to the nanoword $w_{a,b} := (\mathcal{A} = \{A, B\}, ABAB)$ where $|A| = a$, $|B| = b \in \alpha$ with $a \neq \tau(b)$. Moreover for $a \neq \tau(b)$, the nanoword $w_{a,b}$ is non-contractible and two nanowords $w_{a,b}$ and $w_{a',b'}$ are homotopic if and only if $a = a'$ and $b = b'$.*

REMARK 5.2. In [7], Turaev also gave a classification of nanowords of length 6. But in this paper we do not use that result. The classification problem for nanowords of length 8 or more is still open (see [9]).

5.2. The case of nanophrases of length 2. First we introduce the following notation: $P_a := (A|A)$, $P^{4,0}_{a,b} := (ABAB|\emptyset)$, $P^{3,1}_{a,b} := (ABA|B)$, $P^{2,2I}_{a,b} := (AB|AB)$, $P^{2,2II}_{a,b} := (AB|BA)$, $P^{1,3}_{a,b} := (A|BAB)$ and $P^{0,4}_{a,b} := (\emptyset|ABAB)$ with $|A| = a,\ |B| = b \in \alpha$. If $a = \tau(b)$, then $P^{4,0}_{a,b}$, $P^{2,2I}_{a,b}$, $P^{2,2II}_{a,b}$ and $P^{0,4}_{a,b}$ are homotopic to $(\emptyset|\emptyset)$. So in this paper, if we write $P^{4,0}_{a,b}$, $P^{2,2I}_{a,b}$, $P^{2,2II}_{a,b}$, $P^{0,4}_{a,b}$ then we always assume that $a \neq \tau(b)$.

In [1], the author gave a classification of nanophrases of length 2 with at most four letters.

THEOREM 5.3 ([1]). *Let P be a nanophrase of length 2 with two letters. Then P is not homotopic to $(\emptyset|\emptyset)$ if and only if P is isomorphic to P_a. Moreover P_a and $P_{a'}$ are homotopic if and only if $a = a'$.*

THEOREM 5.4 ([1]). *Let P be a nanophrase of length 2 with four letters. Then P is homotopic to $(\emptyset|\emptyset)$, or homotopic to a nanophrase of length 2 with two letters, or isomorphic to one of the following nanophrases: $P^{4,0}_{a,b}$, $P^{3,1}_{a,b}$, $P^{2,2I}_{a,b}$, $P^{2,2II}_{a,b}$, $P^{1,3}_{a,b}$, $P^{0,4}_{a,b}$. For $(i,j) \in \{(4,0),(3,1),(2,2I),(2,2II),(1,3),(0,4)\}$ and any $a,b \in \alpha$, the nanophrase $P^{i,j}_{a,b}$ is neither homotopic to $(\emptyset|\emptyset)$ nor to any nanophrase of length 2 with two letters. The nanophrases $P^{i,j}_{a,b}$ and $P^{i,j}_{a',b'}$ are homotopic if and only if $a = a'$ and $b = b'$. For $(i,j) \neq (i',j')$, the nanophrases $P^{i,j}_{a,b}$ and $P^{i',j'}_{a',b'}$ are not homotopic for any $a,b,a',b' \in \alpha$.*

COROLLARY 5.5 ([1]). *There are exactly 19 stable equivalence classes of two-component pointed, ordered, oriented, curves on surfaces with minimum crossing number less than or equal to 2.*

In this paper, we give a classification of nanophrases of length at least 3 with four letters.

5.3. Homotopy invariants of nanophrases. In this subsection we introduce some invariants of nanophrases over α (some of them are defined in [1]).

Let Π be the group defined as follows:

$$\Pi := (\{z_a\}_{a\in\alpha} \mid z_a z_{\tau(a)} = 1 \text{ for all } a \in \alpha).$$

DEFINITION 5.6 (cf. [1]). Let $P = (\mathcal{A}, (w_1|\cdots|w_k))$ be a nanophrase of length k over α, and let n_i be the length of nanoword w_i. Set $n = \sum_{1\leq i\leq k} n_i$. Then we define n elements $\gamma^i_1, \gamma^i_2, \ldots, \gamma^i_{n_i}$ $(i \in \{1,\ldots,k\})$ of Π by $\gamma^j_i := z_{|w_j(i)|}$ if $w_j(i) \neq w_l(m)$ for all $l < j$ and all $m < i$ when $l = j$. Otherwise $\gamma^j_i := z_{\tau(|w_j(i)|)}$. Then we define $\gamma(P) \in \Pi^k$ by

$$\gamma(P) := (\gamma^1_1\gamma^1_2\cdots\gamma^1_{n_1}, \gamma^2_1\gamma^2_2\cdots\gamma^2_{n_2}, \ldots, \gamma^k_1\gamma^k_2\cdots\gamma^k_{n_k}).$$

Then we obtain the following proposition.

Proposition 5.6.1. *γ is a homotopy invariant of nanophrases.*

We now define an invariant T of nanophrases.

First we prepare some notation. Since α is a finite set, we obtain the following orbit decomposition of τ:

$$\alpha/\tau = \{\widetilde{a_{i_1}}, \dots, \widetilde{a_{i_l}}, \widetilde{a_{i_{l+1}}}, \dots, \widetilde{a_{i_{l+m}}}\},$$

where $\widetilde{a_{i_j}} := \{a_{i_j}, \tau(a_{i_j})\}$, such that $\mathrm{Card}(\widetilde{a_{i_j}}) = 2$ for all $j \in \{1, \dots, l\}$ and $\mathrm{Card}(\widetilde{a_{i_j}}) = 1$ for all $j \in \{l+1, \dots, l+m\}$ (we fix a complete representative system $\{a_{i_1}, \dots, a_{i_l}, a_{i_{l+1}}, \dots, a_{i_{l+m}}\}$ which satisfies the above condition). Let $\mathcal{A}$ be an α-alphabet. For $A \in \mathcal{A}$ we define $\varepsilon(A) \in \{\pm 1\}$ by

$$\varepsilon(A) := \begin{cases} 1 & \text{if } |A| = a_{i_j} \text{ for some } j \in \{1, \dots, l+m\}, \\ -1 & \text{if } |A| = \tau(a_{i_j}) \text{ for some } j \in \{1, \dots l\}. \end{cases}$$

Let $P = (\mathcal{A}, (w_1|\cdots|w_k))$ be a nanophrase over α, and $A, B \in \mathcal{A}$. Let $K_{(i,j)}$ be $\mathbb{Z}$ if $i \leq l$ and $j \leq l$, otherwise $K_{(i,j)} = \mathbb{Z}/2\mathbb{Z}$. We denote $K_{(1,1)} \times K_{(1,2)} \times \cdots \times K_{(1,l+m)} \times K_{(2,1)} \times \cdots \times K_{(l+m,l+m)}$ by $\prod K_{(i,j)}$. Then we define $\sigma_P(A,B) \in \prod K_{(i,j)}$ as follows: If A and B form $\cdots A \cdots B \cdots A \cdots B \cdots$ in P, $|A| \in \widetilde{a_{i_p}}$ and $|B| = a_{i_q}$ for some $m, n \in \{1, \dots, l+m\}$, or $\cdots B \cdots A \cdots B \cdots A \cdots$ in P, $|A| \in \widetilde{a_{i_p}}$ and $|B| = \tau(a_{i_q})$ for some $p, q \in \{1, \dots, l+m\}$, then

$$\sigma_P(A,B) := (0, \dots, 0, \overset{(p,q)}{\check{1}}, 0, \dots, 0).$$

If $\cdots A \cdots B \cdots A \cdots B \cdots$ in P, $|A| \in \widetilde{a_{i_p}}$ and $|B| = \tau(a_{i_q})$, or $\cdots B \cdots A \cdots B \cdots A \cdots$ in P, $|A| \in \widetilde{a_{i_p}}$ and $|B| = a_{i_q}$, then

$$\sigma_P(A,B) := (0, \dots, 0, \overset{(p,q)}{\check{-1}}, 0, \dots, 0).$$

Otherwise $\sigma_P(A,B) := (0, \dots, 0)$. With the above preparation, we define the invariant T as follows.

Definition 5.7. Let $P = (\mathcal{A}, (w_1|\cdots|w_k))$ be a nanophrase of length k over α. For $A \in \mathcal{A}$ such that there exists $i \in \{1, \dots, k\}$ with $\mathrm{Card}(w_i^{-1}(A)) = 2$, we define $T_P(A) \in \prod K_{(i,j)}$ by

$$T_P(A) := \varepsilon(A) \sum_{B \in \mathcal{A}} \sigma_P(A,B),$$

and $T_P(w_i) \in \prod K_{(i,j)}$ by

$$T_P(w_i) := \sum_{A \in \mathcal{A},\, \mathrm{Card}(w_i^{-1}(A)) = 2} T_P(A).$$

Then we define $T(P) \in (\prod K_{(i,j)})^k$ by

$$T(P) := (T_P(w_1), \dots, T_P(w_k)).$$

PROPOSITION 5.7.1. *T is an invariant of nanophrases over α.*

Proof. It is clear that isomorphism does not change the value of T. Consider the first homotopy move

$$P_1 := (\mathcal{A}, (xAAy)) \to P_2 := (\mathcal{A} \setminus \{A\}, (xy))$$

where x and y are words on $\mathcal{A}$, possibly including the "|" character. Since A and X are unlacement in the phrase P_1 for all $X \in \mathcal{A}$, A does not contribute to $T(P_1)$. So the first homotopy move does not change the value of T.

Consider the second homotopy move

$$P_1 := (\mathcal{A}, (xAByBAz)) \to P_2 := (\mathcal{A} \setminus \{A, B\}, (xyz))$$

where $|A| = \tau(|B|)$, and x, y and z are words on $\mathcal{A}$, possibly including "|". Suppose y does not include "|" and $\mathrm{Card}(\widetilde{|A|}) = 2$ (so $\mathrm{Card}(\widetilde{|B|})$ is also two). Then $T_{P_1}(A) + T_{P_2}(B) = 0$ since

$$\begin{aligned}
T_{P_1}(A) &= \varepsilon(A)\Big(\sigma_{P_1}(A, B) + \sum_{X \in \mathcal{A}\setminus\{B\}} \sigma_{P_1}(A, X)\Big) \\
&= \varepsilon(A) \sum_{X \in \mathcal{A}\setminus\{B\}} \sigma_{P_1}(A, X) = -\varepsilon(B) \sum_{X \in \mathcal{A}\setminus\{A\}} \sigma_{P_1}(B, X) \\
&= -\varepsilon(B)\Big(\sigma_{P_1}(B, A) + \sum_{X \in \mathcal{A}\setminus\{A\}} \sigma_{P_1}(B, X)\Big) \\
&= -T_{P_1}(B).
\end{aligned}$$

Moreover for $X \in \mathcal{A} \setminus \{A, B\}$, $\cdots A \cdots X \cdots A \cdots X \cdots$ (respectively $\cdots X \cdots A \cdots X \cdots A \cdots$) in P_1 if and only if $\cdots B \cdots X \cdots B \cdots X \cdots$ (respectively $\cdots X \cdots B \cdots X \cdots B \cdots$) in P_1, and $|A| = \tau(|B|)$. So

$$\sigma_{P_1}(X, A) + \sigma_{P_1}(X, B) = 0$$

for all $X \in \mathcal{A}$. Hence

$$\begin{aligned}
T_{P_1}(X) &= \varepsilon(X)\Big(\sigma_{P_1}(X, A) + \sigma_{P_1}(X, B) + \sum_{D \in \mathcal{A}\setminus\{A,B\}} \sigma_{P_1}(X, D)\Big) \\
&= \varepsilon(X) \sum_{D \in \mathcal{A}\setminus\{A,B\}} \sigma_{P_1}(X, D) = \varepsilon(X) \sum_{D \in \mathcal{A}\setminus\{A,B\}} \sigma_{P_2}(X, D) \\
&= T_{P_2}(X).
\end{aligned}$$

This implies $T(P_1) = T(P_2)$.

Suppose y does not include "|" and $\mathrm{Card}(\widetilde{|A|}) = 1$ (so $\mathrm{Card}(\widetilde{|B|})$ is also one). In this case also $T_{P_1}(A) + T_{P_2}(B) = 0$ since

$$\begin{aligned}T_{P_1}(A) &= \varepsilon(A)\Big(\sigma_{P_1}(A,B) + \sum_{X\in\mathcal{A}\backslash\{B\}} \sigma_{P_1}(A,X)\Big)\\ &= \varepsilon(A) \sum_{X\in\mathcal{A}\backslash\{B\}} \sigma_{P_1}(A,X) = \varepsilon(B) \sum_{X\in\mathcal{A}\backslash\{A\}} \sigma_{P_1}(B,X)\\ &= \varepsilon(B)\Big(\sigma_{P_1}(B,A) + \sum_{X\in\mathcal{A}\backslash\{A\}} \sigma_{P_1}(B,X)\Big) = T_{P_1}(B),\end{aligned}$$

and all entries of $T_{P_1}(A)$ and $T_{P_2}(B)$ are elements of $\mathbb{Z}/2\mathbb{Z}$. Moreover for $X \in \mathcal{A}\backslash\{A,B\}$, $\cdots A \cdots X \cdots A \cdots X \cdots$ (respectively $\cdots X \cdots A \cdots X \cdots A \cdots$) in P_1 if and only if $\cdots B \cdots X \cdots B \cdots X \cdots$ (respectively $\cdots X \cdots B \cdots X \cdots B \cdots$) in P_1. Since $\widetilde{|A|} = \widetilde{|B|}$ and $\mathrm{Card}(\widetilde{|A|}) = 1$, we have $\sigma_{P_1}(X,A) = \sigma_{P_1}(X,B)$ in $\mathbb{Z}/2\mathbb{Z}$. So $\sigma_{P_1}(X,A) + \sigma_{P_1}(X,B) = 0$ for all $X \in \mathcal{A}$. By the above

$$\begin{aligned}T_{P_1}(X) &= \varepsilon(X)\Big(\sigma_{P_1}(X,A) + \sigma_{P_1}(X,B) + \sum_{D\in\mathcal{A}\backslash\{A,B\}} \sigma_{P_1}(X,D)\Big)\\ &= \varepsilon(X) \sum_{D\in\mathcal{A}\backslash\{A,B\}} \sigma_{P_1}(X,D) = \varepsilon(X) \sum_{D\in\mathcal{A}\backslash\{A,B\}} \sigma_{P_2}(X,D) = T_{P_2}(X).\end{aligned}$$

This implies $T(P_1) = T(P_2)$.

The case when y includes "|" is proved similarly.

Consider the third homotopy move

$$P_1 := (\mathcal{A}, (xAByACzBCt)) \to P_2 := (\mathcal{A}, (xBAyCAzCBt))$$

where $|A| = |B| = |C|$, and x, y, z and t are words on $\mathcal{A}$, possibly with "|". Suppose y and z do not include "|". Note that $\sigma_{P_1}(A,B) = \sigma_{P_2}(A,C)$. So

$$\begin{aligned}T_{P_1}(A) &= \varepsilon(A)\Big(\sigma_{P_1}(A,B) + \sum_{X\in\mathcal{A}\backslash\{B\}} \sigma_{P_1}(A,X)\Big)\\ &= \varepsilon(A)\Big(\sum_{X\in\mathcal{A}\backslash\{C\}} \sigma_{P_2}(A,X) + \sigma_{P_2}(A,C)\Big) = T_{P_2}(A),\end{aligned}$$

and since $\sigma_{P_1}(C,B) = \sigma_{P_2}(C,A)$, we obtain

$$\begin{aligned}T_{P_1}(C) &= \varepsilon(C)\Big(\sigma_{P_1}(C,B) + \sum_{X\in\mathcal{A}\backslash\{B\}} \sigma_{P_1}(C,X)\Big)\\ &= \varepsilon(C)\Big(\sum_{X\in\mathcal{A}\backslash\{C\}} \sigma_{P_2}(C,X) + \sigma_{P_2}(C,A)\Big) = T_{P_2}(C).\end{aligned}$$

Moreover $\sigma_{P_1}(B,A) + \sigma_{P_1}(B,C) = 0$ and $\sigma_{P_2}(B,A) = \sigma_{P_2}(B,C) = 0$. We obtain $T_{P_1}(B) = T_{P_2}(B)$. It is easily checked that $T_{P_1}(E) = T_{P_2}(E)$ for all $E \neq A,B,C$. So we obtain $T(P_1) = T(P_2)$.

The case of y or z including "|" is proved similarly. ∎

Example 5.8. We calculate T for a nanophrase over α_0 with an involution $\tau_0 : a \mapsto b$, $P = (C|ABCA|B)$ with $|A| = a$, $|B| = a$ and $|C| = b$.

First, we choose the set $\{a\}$ as a complete representative system of α_0/τ_0. Since this is a one-element set and the orbit of this element is a free orbit, each component of $T(P)$ is an element of $\mathbb{Z}$ for each $i \in \{1,2,3\}$ by the definition of T. Thus $\varepsilon(A) = 1$. Moreover, $\sigma_P(A,B) = 1$ and $\sigma_P(A,C) = 1$. Therefore $T_P(w_1)$ and $T_P(w_3)$ are 0, and

$$T_P(w_2) = \varepsilon(A)(\sigma_P(A,B) + \sigma_P(A,C)) = 2.$$

By the above

$$T(P) = (0,2,0) \in \mathbb{Z}.$$

Remark 5.9. The invariant T is a generalization of the invariants T of nanophrases over α_0 and the one-element set defined in [1]. Using the invariant T defined in this paper, we can classify nanophrases of length 2 with four letters without using Lemma 4.2 of [1].

Next we define another invariant. Let π be the group defined as follows:

$$\pi := (a \in \alpha \mid a\tau(a) = 1,\, ab = ba \text{ for all } a,b \in \alpha) \simeq \Pi/[\Pi,\Pi].$$

Let $P = (\mathcal{A}, (w_1|\cdots|w_k))$ be a nanophrase of length k over α. We define $(w_i,w_j)_P \in \pi$ for $i < j$ by

$$(w_i,w_j)_P := \prod_{A \in \mathrm{Im}(w_i) \cap \mathrm{Im}(w_j)} |A|.$$

Proposition 5.9.1. *If nanophrases P_1 and P_2 over α are homotopic, then $(w_i,w_j)_{P_1} = (w_i,w_j)_{P_2}$.*

Proof. It is clear that isomorphism does not change the value of $(w_i,w_j)_P$. Consider the first homotopy move

$$P_1 := (\mathcal{A}, (xAAy)) \to P_2 := (\mathcal{A} \setminus \{A\}, (xy)).$$

In this move, the letter A appears twice in the same component. So A does not contribute to $(w_i,w_j)_{P_1}$. This implies $(w_i,w_j)_{P_1} = (w_i,w_j)_{P_2}$.

Consider the second homotopy move

$$P_1 := (\mathcal{A}, (xAByBAz)) \to P_2 := (\mathcal{A} \setminus \{A,B\}, (xyz))$$

where $|A| = \tau(|B|)$, and x, y and z are words on $\mathcal{A}$, possibly with "|". Suppose y does not include "|". In this case, A and B are in the same component of the nanophrase P_1. So A and B do not contribute to $(w_i,w_j)_{P_1}$. This implies $(w_i,w_j)_{P_1} = (w_i,w_j)_{P_2}$ for all i, j. Suppose y includes "|". Suppose A and B are in the mth component and the nth component of P_1 respectively. Then

$$(w_m,w_n)_{P_1} = (w_m,w_n)_{P_2} \cdot |A| \cdot |B| = (w_m,w_n)_{P_2} \cdot |A| \cdot \tau(|A|) = (w_m,w_n)_{P_2},$$

and it is clear that $(w_i, w_j)_{P_1} = (w_i, w_j)_{P_2}$ for $(i, j) \neq (m, n)$. So $(w_i, w_j)_{P_1} = (w_i, w_j)_{P_2}$ for all i and j.

Consider the third homotopy move

$$P_1 := (\mathcal{A}, (xAByACzBCt)) \to P_2 := (\mathcal{A}, (xBAyCAzCBt))$$

where $|A| = |B| = |C|$, and x, y, z and t are words on $\mathcal{A}$, possibly with "$|$". Note that the third homotopy move sends a letter in the lth component of P_1 to the lth component of P_2. So $(w_i, w_j)_{P_1}$ is not changed by the third homotopy move.

By the above, $(w_i, w_j)_{P_1}$ is a homotopy invariant of nanophrases. ∎

By the above proposition, we obtain a homotopy invariant of nanophrases

$$((w_1, w_2)_P, (w_1, w_3)_P, \ldots, (w_1, w_k)_P, (w_2, w_3)_P, \ldots, (w_{k-1}, w_k)_P) \in \pi^{\frac{1}{2}k(k-1)}.$$

Example 5.10. We calculate $(w_i, w_j)_P$ for a nanophrase over α_0 with an involution $\tau_0 : a \mapsto b$, $P = (AB|AC|BC)$ with $|A| = a$, $|B| = a$ and $|C| = b$. By definition, $(w_1, w_2)_P = |A|$, $(w_1, w_3)_P = |B|$ and $(w_2, w_3)_P = |C|$. Therefore

$$((w_i, w_j)_P)_{i<j} = (|A|, |B|, |C|) = (a, a, b) \in \pi^3.$$

5.4. The case of nanophrases of length 3 or more. Now using the invariants prepared in the last section and some lemmas, we classify the nanophrases of length 3 or with at most four letters. First recall the following lemmas from [1].

Lemma 5.11. *Let $P_1 = (w_1|\cdots|w_k)$ and $P_2 = (v_1|\cdots|v_k)$ be nanophrases of length k over α. If P_1 and P_2 are homotopic as nanophrases, then w_i and v_i are homotopic as étale words for all $i \in \{1, \ldots, k\}$.*

Lemma 5.12. *Let $P_1 = (w_1|\cdots|w_k)$ and $P_2 = (v_1|\cdots|v_k)$ be nanophrases of length k over α. If P_1 and P_2 are homotopic, then the length of w_i is equal to the length of v_i modulo 2 for all $i \in \{1, \ldots, k\}$.*

The following lemma is checked easily by the definition of homotopy of nanophrases.

Lemma 5.13. *Let $P_1 = (w_1|\cdots|w_k)$ and $P_2 = (v_1|\cdots|v_k)$ be nanophrases over α. If P_1 and P_2 are homotopic, then $(w_1|\cdots|w_l w_{l+1}|\cdots|w_k)$ and $(v_1|\cdots|v_l v_{l+1}|\cdots|v_k)$ are homotopic as nanophrases of length $k-1$ over α for all $l \in \{1, \ldots, k-1\}$.*

Now we give a classification of nanophrases with two letters. Set

$$P_a^{1,1;p,q} := (\emptyset|\cdots|\emptyset|\overset{p}{\breve{A}}|\emptyset|\cdots|\emptyset|\overset{q}{\breve{A}}|\emptyset|\cdots|\emptyset)$$

with $|A| = a$ for $1 \leq p < q \leq k$.

THEOREM 5.14. *Let P be a nanophrase of length k with two letters. Then P is either homotopic to $(\emptyset|\cdots|\emptyset)$ or isomorphic to $P_a^{1,1;p,q}$ for some $p,q\in\{1,\ldots,k\}$, $a\in\alpha$. Moreover $P_a^{1,1;p,q}$ and $P_{a'}^{1,1;p',q'}$ are homotopic if and only if $p=p'$, $q=q'$ and $a=a'$.*

Proof. The first assertion is clear. We show the second assertion. By the definition of $(w_i,w_j)_P$, $(w_i,w_j)_{P_a^{1,1;p,q}}=a$ if $i=p$ and $j=q$. Otherwise $(w_i,w_j)_{P_a^{1,1;p,q}}=1$. For $a\in\alpha$, $a\neq 1$ in π. So if $P_a^{1,1;p,q}$ and $P_{a'}^{1,1;p',q'}$ are homotopic, then $p=p'$, $q=q'$ and $a=a'$. ∎

To state the classification theorem for nanophrases with four letters, we introduce the following notation:

$$\begin{aligned}
P_{a,b}^{4;p} &:= (\emptyset|\cdots|\emptyset|\ A\overset{p}{\check{B}}AB\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{3,1;p,q} &:= (\emptyset|\cdots|\emptyset|\ A\overset{p}{\check{B}}A\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{2,2I;p,q} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}B\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{A}}B\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{2,2II;p,q} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}B\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}A\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,3;p,q} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ B\overset{q}{\check{A}}B\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{2,1,1I;p,q,r} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}B\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{2,1,1II;p,q,r} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{B}}A\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,2,1I;p,q,r} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{A}}B\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,2,1II;p,q,r} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}A\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,1,2I;p,q,r} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{A}}B\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,1,2II;p,q,r} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}A\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,1,1,1I;p,q,r,s} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}\ |\emptyset|\cdots|\emptyset|\ \overset{s}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,1,1,1II;p,q,r,s} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{s}{\check{B}}\ |\emptyset|\cdots|\emptyset),\\
P_{a,b}^{1,1,1,1III;p,q,r,s} &:= (\emptyset|\cdots|\emptyset|\ \overset{p}{\check{A}}\ |\emptyset|\cdots|\emptyset|\ \overset{q}{\check{B}}\ |\emptyset|\cdots|\emptyset|\ \overset{r}{\check{B}}\ |\emptyset|\cdots|\emptyset|\ \overset{s}{\check{A}}\ |\emptyset|\cdots|\emptyset),
\end{aligned}$$

with $|A|=a$, $|B|=b$. If $a=\tau(b)$, then the nanophrases $P_{a,b}^{4;p}$, $P_{a,b}^{2,2I;p,q}$ and $P_{a,b}^{2,2II;p,q}$ are homotopic to $(\emptyset|\cdots|\emptyset)$. So when we write $P_{a,b}^{4;p}$, $P_{a,b}^{2,2I;p,q}$, $P_{a,b}^{2,2II;p,q}$ we always assume that $a\neq\tau(b)$.

Under the above notation the classification of nanophrases with four letters is as follows.

THEOREM 5.15. *Let P be a nanophrase of length k with four letters. Then P is either homotopic to a nanophrase with at most two letters or isomorphic to $P^{X;Y}_{a,b}$ for some $X \in \{4, (3,1), \ldots, (1,1,1,1III)\}$, $Y \in \{1, \ldots, k, (1,2), \ldots, (k-3, k-2, k-1, k)\}$. Moreover $P^{X;Y}_{a,b}$ and $P^{X';Y'}_{a',b'}$ are homotopic if and only if $X = X'$, $Y = Y'$, $a = a'$ and $b = b'$.*

Proof. The first assertion is clear. To prove the rest, it must be shown that (i) if $X \neq X'$, then $P^{X;Y}_{a,b}$ and $P^{X';Y'}_{a,b}$ are not homotopic; and (ii) a 4-letter nanophrase $P^{X;Y}_{a,b}$ is homotopic to $P^{X;Y'}_{a,b}$ if and only if $Y = Y'$, $a = a'$ and $b = b'$. First we split the basic shapes of nanophrases into eight sets:

$$\mathcal{P}_0 = \{(\emptyset|\cdots|\emptyset), P^{1,1;p}_{a}\},$$
$$\mathcal{P}_1 = \{P^{4;p}_{a,b} \mid 1 \leq p \leq k,\ a,b \in \alpha\},$$
$$\mathcal{P}_2 = \{P^{3,1;p,q}_{a,b}, P^{1,3;p,q}_{a,b} \mid 1 \leq p < q \leq k,\ a,b \in \alpha\},$$
$$\mathcal{P}_3 = \{P^{2,2I;p,q}_{a,b}, P^{2,2II;p,q}_{a,b} \mid 1 \leq p < q \leq k,\ a,b \in \alpha\},$$
$$\mathcal{P}_4 = \{P^{2,1,1I;p,q,r}_{a,b}, P^{2,1,1II;p,q,r}_{a,b} \mid 1 \leq p < q < r \leq k,\ a,b \in \alpha\},$$
$$\mathcal{P}_5 = \{P^{1,2,1I;p,q,r}_{a,b}, P^{1,2,1II;p,q,r}_{a,b} \mid 1 \leq p < q < r \leq k,\ a,b \in \alpha\},$$
$$\mathcal{P}_6 = \{P^{1,1,2I;p,q,r}_{a,b}, P^{1,1,2II;p,q,r}_{a,b} \mid 1 \leq p < q < r \leq k,\ a,b \in \alpha\},$$
$$\mathcal{P}_7 = \{P^{1,1,1,1I;p,q,r,s}_{a,b}, P^{1,1,1,1II;p,q,r,s}_{a,b}, P^{1,1,1,1III;p,q,r,s}_{a,b} \mid \\ 1 \leq p < q < r < s \leq k,\ a,b \in \alpha\}.$$

By using the invariants γ, T and $((w_i, w_j)_P)_{i<j}$, we can easily check that two nanophrases $P \in \mathcal{P}_i$ and $P' \in \mathcal{P}_j$ can be homotopic only if $i = j$. This cuts down the number of pairs of nanophrases that need to be considered in (i).

Consider the nanophrases in $\mathcal{P}_1$.

The claim that $P^{4;p}_{a,b}$ is homotopic to $P^{4;p'}_{a',b'}$ if and only if $p = p'$, $a = a'$ and $b = b'$ follows from Theorem 5.1 and Lemma 5.13.

Consider the nanophrases in $\mathcal{P}_2$.

First, $P^{3,1;p,q}_{a,b}$ is not homotopic to $P^{1,3;p',q'}_{a',b'}$: Indeed, suppose otherwise. Then $p = p'$ and $q = q'$, since

$$((w_i, w_j)_{P^{3,1;p,q}_{a,b}})_{i<j} = ((w_i, w_j)_{P^{1,3;p',q'}_{a',b'}})_{i<j}.$$

By Lemma 5.13, $(ABA|B)$ with $|A| = a, |B| = b$ must be homotopic to $(A'|B'A'B')$ with $|A'| = a'$, $|B'| = b'$. However, this contradicts Theorem 5.4.

The claim that $P_{a,b}^{3,1;p,q}$ is homotopic to $P_{a',b'}^{3,1;p',q'}$ if and only if $p = p'$, $q = q'$, $a = a'$ and $b = b'$ follows by comparing $((w_i, w_j)_{P_{a,b}^{3,1;p,q}})_{i<j}$ and $((w_i, w_j)_{P_{a',b'}^{3,1;p',q'}})_{i<j}$.

The claim that $P_{a,b}^{1,3;p,q}$ is homotopic to $P_{a',b'}^{1,3;p',q'}$ if and only if $p = p'$, $q = q'$, $a = a'$ and $b = b'$ is proved similarly.

Consider the nanophrases in $\mathcal{P}_3$.

First, $P_{a,b}^{2,2I;p,q}$ and $P_{a',b'}^{2,2II;p',q'}$ are not homotopic: Indeed, suppose they are. Then $p{=}p'$ and $q{=}q'$, since

$$((w_i, w_j)_{P_{a,b}^{2,2I;p,q}})_{i<j} = ((w_i, w_j)_{P_{a',b'}^{2,2II;p',q'}})_{i<j}.$$

By Lemma 5.13, $(AB|AB)$ with $|A| = a$, $|B| = b$ must be homotopic to $(A'B'|B'A')$ with $|A'| = a', |B'| = b'$. However, this contradicts Theorem 5.4.

The claim that $P_{a,b}^{2,2I;p,q}$ and $P_{a',b'}^{2,2I;p',q'}$ are homotopic if and only if $p = p'$, $q = q'$, $a = a'$ and $b = b'$ follows by comparing the values of the invariant $((w_i, w_j)_P)_{i<j}$.

The claim that $P_{a,b}^{2,2II;p,q}$ and $P_{a',b'}^{2,2II;p',q'}$ are homotopic if and only if $p = p'$, $q = q'$, $a = a'$ and $b = b'$ is proved similarly.

Consider the nanophrases in $\mathcal{P}_4$.

First, $P_{a,b}^{2,1,1I;p,q,r}$ and $P_{a',b'}^{2,1,1II;p',q',r'}$ are not homotopic: Indeed, suppose they are. Then $p = p'$, $q = q'$ and $r = r'$, since,

$$((w_i, w_j)_{P_{a,b}^{2,1,1I;p,q}})_{i<j} = ((w_i, w_j)_{P_{a',b'}^{2,1,1II;p',q',r'}})_{i<j}.$$

By Lemma 5.13 the nanophrases $(ABA|B)$ and $(B'A'A'|B')$ are homotopic. However, this contradicts Theorem 5.4.

The claim that $P_{a,b}^{2,1,1I;p,q,r}$ and $P_{a',b'}^{2,1,1I;p',q',r'}$ are homotopic if and only if $p = p'$, $q = q'$ and $r = r'$, $a = a'$ and $b = b'$ follows by comparing the values of the invariant $((w_i, w_j)_P)_{i<j}$.

For the nanophrases in $\mathcal{P}_5$ and $\mathcal{P}_6$, we can prove (i) and (ii) similarly.

Consider the nanophrases in $\mathcal{P}_7$.

First, $P_{a,b}^{1,1,1,1I;p,q,r,s}$ and $P_{a',b'}^{1,1,1,1II;p',q',r',s'}$ are not homotopic: Indeed, if we assume they are homotopic, then $p = p'$, $q = q'$, $r = r'$ and $z = z'$ since

$$((w_i, w_j)_{P_{a,b}^{1,1,1,1I;p,q,r,s}})_{i<j} = ((w_i, w_j)_{P_{a',b'}^{1,1,1,1II;p',q',r',s'}})_{i<j}.$$

So $(A|BAB)$ must be homotopic to $(A'|A'B'B')$ by Lemma 5.13. But this contradicts Theorem 5.4.

Next, $P_{a,b}^{1,1,1,1I;p,q,r,s}$ and $P_{a',b'}^{1,1,1,1III;p',q',r',s'}$ are not homotopic: If we assume they are, then $p = p'$, $q = q'$, $r = r'$ and $z = z'$, so $(A|AB|B)$ must be homotopic to $(A'|\emptyset|A')$ by Lemma 5.13. However, this contradicts the homotopy invariance of $((w_i, w_j)_P)_{i<j}$.

That $P_{a,b}^{1,1,1,1II;p,q,r,s}$ and $P_{a',b'}^{1,1,1,1III;p',q',r',s'}$ are not homotopic follows similarly to the above.

The claim that $P_{a,b}^{1,1,1,1I;p,q,r,s}$ and $P_{a',b'}^{1,1,1,1I;p',q',r',s'}$ are homotopic if and only if $p = p'$, $q = q'$, $r = r'$ and $z = z'$, $a = a'$ and $b = b'$ follows by homotopy invariance of $((w_i, w_j)_P)_{i<j}$. The claims that $P_{a,b}^{1,1,1,1II;p,q,r,s}$ and $P_{a',b'}^{1,1,1,1II;p',q',r',s'}$ are homotopic if and only if $p = p'$, $q = q'$, $r = r'$ and $z = z'$, $a = a'$ and $b = b'$, and that $P_{a,b}^{1,1,1,1III;p,q,r,s}$ and $P_{a',b'}^{1,1,1,1III;p',q',r',s'}$ are homotopic if and only if $p = p'$, $q = q'$, $r = r'$ and $z = z'$, $a = a'$ and $b = b'$, follow similarly.

Now we have completed the homotopy classification of nanophrases with no more than four letters without any condition on length. ■

6. Proof of Theorem 2.2. To complete the proof of Theorem 2.2, we need the following lemma.

LEMMA 6.1. *Let α be an alphabet endowed with an involution τ. The following nanophrases over α:* $(A|A)$, $(AB|AB)$ *with* $|A| \neq \tau(|B|)$, $(AB|BA)$ *with* $|A| \neq \tau(|B|)$, $(ABA|B)$, $(A|BAB)$, $(AB|A|B)$, $(BA|A|B)$, $(A|AB|B)$, $(A|BA|B)$, $(A|B|AB)$, $(A|B|BA)$, $(A|A|B|B)$, $(A|B|A|B)$ *and* $(A|B|B|A)$, *are not homotopic to nanophrases over α which have empty words in their components.*

Proof. This follows easily from Proposition 5.9.1, Lemma 5.12 and Theorem 5.4. ■

Now Theorem 2.2 follows immediately from Theorem 5.15 and Lemma 6.1. It is sufficient to apply the above theorems to $\alpha = \alpha_0$ with the involution $\tau : \alpha_0 \to \alpha_0$ permuting a and b.

Acknowledgements. The author was supported by JSPS within International Training Program (ITP) to make the first draft, and he was supported as a JSPS research fellow (DC) to complete the final version of this paper.

References

[1] T. Fukunaga, *Homotopy classification of nanophrases in Turaev's theory of words*, J. Knot Theory Ramif. 18 (2009), 901–915.

[2] C. F. Gauss, *Werke*, Vol. 8, Teubner, Leipzig, 1900.

[3] T. Kadokami, *Detecting non-triviality of virtual links*, J. Knot Theory Ramif. 12 (2003), 781–803.

[4] N. Kamada and S. Kamada, *Abstract link diagrams and virtual knots*, ibid. 9 (2000), 93–106.

[5] L. H. Kauffman, *Virtual knot theory*, Eur. J. Combin. 20 (1999), 663–691.

[6] D. S. Silver and S. G. Williams, *An invariant for open virtual strings*, J. Knot Theory Ramif. 15 (2006), 143–152.
[7] V. Turaev, *Topology of words*, Proc. London Math. Soc. 95 (2007), 360–417.
[8] —, *Knots and words*, Int. Math. Res. Notices 2006, art. ID 84098, 23 pp.
[9] —, *Lectures on topology of words*, Japan. J. Math. 2 (2007), 1–39.
[10] —, *Virtual strings*, Ann. Inst. Fourier (Grenoble) 54 (2004), 2455–2525.

Tomonori Fukunaga
Department of Mathematics
Hokkaido University
Sapporo 060-0810, Japan
E-mail: fukunaga@math.sci.hokudai.ac.jp

Received 26 April 2009;
in revised form 6 June 2011