SIAM J. DISCRETE MATH.
Vol. 27, No. 3, pp. 1410–1428

LOWER BOUNDING EDIT DISTANCES BETWEEN PERMUTATIONS*

ANTHONY LABARRE†

Abstract. A number of fields, including the study of genome rearrangements and the design of interconnection networks, deal with the connected problems of sorting permutations in as few moves as possible, using a given set of allowed operations, or computing the number of moves the sorting process requires, often referred to as the *distance* of the permutation. These operations often act on just one or two segments of the permutation, e.g., by reversing one segment or exchanging two segments. The *cycle graph* of the permutation to sort is a fundamental tool in the theory of genome rearrangements and has proved useful in settling the complexity of many variants of the above problems. In this paper, we present an algebraic reinterpretation of the cycle graph of a permutation π as an even permutation $\overline{\pi}$ and show how to reformulate our sorting problems in terms of particular factorizations of the latter permutation. Using our framework, we recover known results in a simple and unified way and obtain a new lower bound on the *prefix transposition distance* (where a *prefix transposition* displaces the initial segment of a permutation), which is shown to outperform previous results. Moreover, we use our approach to improve the best known lower bound on the *prefix transposition diameter* from $2n/3$ to $\lfloor 3n/4 \rfloor$ and investigate a few relations between some statistics on π and $\overline{\pi}$.

Key words. theoretical aspects of computational biology, edit distance, symmetric group, transpositions, lower bounds

AMS subject classifications. 68R05, 05A05, 20B30

DOI. 10.1137/13090897X

1. Introduction.

Given a set S of allowed operations and two permutations π and σ of $\{1, 2, \ldots, n\}$, we study the related problems of computing, on the one hand, a sequence of elements of S of minimum length that transforms π into σ and, on the other hand, computing the length of such a sequence, referred to as the *distance* between π and σ. The operations in S usually yield an *edit distance* $d_S(\cdot,\cdot)$ with the property that $d_S(\pi,\sigma) = d_S(\sigma^{-1} \circ \pi, \iota)$ for any two permutations π and σ of the same set, where ι is the *identity permutation* $\langle 1\ 2\ \cdots\ n\rangle$. This property allows us to restrict our attention to *sorting permutations* using a minimum number of operations from S or to computing the distance of a given permutation to the identity permutation rather than to another arbitrary permutation. Two areas in which these questions have applications are the fields of *genome rearrangements* and *interconnection network design*, which we briefly review below.

In genome rearrangements (see Fertin et al. [14] for a survey), the permutation to sort represents an ordering of genes in a given unichromosomal genome and the allowed operations model *mutations* that are known to actually occur in evolution. Rearrangements studied in that context include *reversals* [21], which reverse a segment of the permutation, *block-interchanges* [9], which exchange two not necessarily adjacent segments, and *transpositions* [3], which displace a block of adjacent elements. Those seemingly easy problems turn out to be more challenging than they

*Received by the editors February 7, 2013; accepted for publication (in revised form) May 28, 2013; published electronically August 13, 2013. A significant portion of this work previously appeared in *Proceedings of the* 16*th Annual European Symposium on Algorithms (ESA)*, 2008.
http://www.siam.org/journals/sidma/27-3/90897.html

†Laboratoire d'informatique Gaspard Monge, Université Paris-Est Marne-la-Vallée, 77420 Champs-sur-Marne, France (Anthony.Labarre@univ-mlv.fr).

might appear at first: although a polynomial-time algorithm is known for sorting by block-interchanges or computing the associated distance [9], the same problems were shown to be NP-hard for reversals [6] and more recently for transpositions [5]. *Signed permutations*, where each element can be positive or negative, are a more realistic model of evolution because they model *gene orientation* as well, but that information is not always available (see again Fertin et al. [14] for more details).

In interconnection network design (see Lakshmivarahan, Jwo, and Dhall [25] for a thorough survey), permutations stand, e.g., for processors, or other devices to be connected, and form the vertex set of a graph whose edges correspond to physical connections between two devices. One wants to build a graph with small degree and small *diameter*, among other desirable properties. Akers and Krishnamurthy's landmark paper [2] proposed the idea of choosing a set S that *generates* all permutations of $\{1, 2, \ldots, n\}$ and using the corresponding *Cayley graph*, whose vertex set is the set of all permutations and whose edges connect any two permutations that can be obtained from one another by applying a transformation from S, as an interconnection network. In that setting, sorting algorithms for permutations correspond to *routing algorithms* for the corresponding networks, since a sequence of elements of S transforming π into σ corresponds to a path of the same length in the network. Two kinds of operations that received a lot of attention in that context are *prefix reversals* [17], which reverse the initial segment of the permutation, and *prefix exchanges* [1], which swap the first element of the permutation with another element. Those operations gave birth to the *pancake network* and *star graph* topologies, respectively, which are extensively studied models in that field. We also mention *prefix transpositions*, which displace the initial segment of the permutation and were introduced by Dias and Meidanis [10] in the context of genome rearrangements in the hope that their study would shed light and give insight on the challenging problem of sorting by transpositions. Those more restricted versions of operations studied in the context of genome rearrangements do not lead to problems simpler than their unrestricted counterparts: the sorting and distance computation problems related to prefix exchanges can be solved in polynomial time [1], but the complexity of those problems in the case of prefix transpositions is open, and the problem of sorting by prefix reversals has only recently been showed to be NP-hard [4], more than 30 years after the first works on the subject [17, 18].

The *cycle graph* of a permutation is a ubiquitous structure in the field of genome rearrangements and has proved useful in resolving many questions related to the problems discussed in the above paragraphs. In this paper, we present a new way of encoding the cycle graph of a permutation π as an even permutation $\overline{\pi}$, inspired by a previous work of ours [11], and show how to reformulate *any* sorting problem of the form described above in terms of particular factorisations of the latter permutation. We first illustrate the power of our framework by recovering known lower bounds on the block-interchange and transposition distances in a simple and unified way and then use it to prove a new lower bound on the prefix transposition distance. We prove that our lower bound always outperforms that obtained by Dias and Meidanis [10], and we show experimentally that it is a significant improvement over both that result and the only other known lower bound proved by Chitturi and Sudborough [7]. We then use this new result to improve the previously best known lower bound on the maximal value of the prefix transposition distance from $2n/3$, proved by Chitturi and Sudborough [7], to $\lfloor 3n/4 \rfloor$. Finally, we examine some further properties of the model and establish connections between statistics on π and $\overline{\pi}$.

We note that several similar approaches have been proposed to attack edit distance and sorting problems from an algebraic point of view. Meidanis and Dias

[26] proposed such an algebraic framework in 2000; they considered genomes and rearrangements as cyclic permutations. This differs from our model in that we view genomes and rearrangements as linear orders, i.e., sequences of genes and transformations thereof. Several attempts were later made to adapt Meidanis and Dias' model to *linear* genomes [27, 13], where characterizations of valid operations in that context were also proposed. By contrast, the model we propose is simple, applies directly to every edit distance in the classical sense without the need to adapt the operations under consideration or to view permutations as cycles, and allows the derivation of several results as simple consequences of results obtained by Jerrum [20]. Our model also allows the translation of sorting sequences from one problem to another, meaning that if one can sort a given permutation π using operations from a set S, then one can directly derive a sorting sequence using elements from $f(S)$ for the permutation $f(\pi)$, where $f(\cdot)$ is the function—mentioned above—that encodes the cycle graph of a given permutation.

2. Notation and definitions.

2.1. Permutations and conjugacy classes. Let us start with a quick reminder of basic notions on permutations. (For details, see, e.g., Wielandt [29].)

DEFINITION 2.1. *A* permutation *of a set Ω is a bijective application of Ω onto itself.*

It is convenient to set $\Omega = \{1, 2, \ldots, n\}$, and we will follow this convention here, although we will also sometimes use the set $\{0, 1, 2, \ldots, n\}$. The *symmetric group S_n* is the set of all permutations of a set of n elements, together with the usual function composition $\circ$, applied from right to left. Permutations are denoted by lowercase Greek letters, and we will follow the convention of shortening the traditional two-row notation

$$\pi = \left\langle \begin{matrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{matrix} \right\rangle$$

by keeping only the second row, i.e., $\pi = \langle \pi_1\ \pi_2\ \cdots\ \pi_n \rangle$, where $\pi_i = \pi(i)$.

DEFINITION 2.2. *The* graph $\Gamma(\pi)$ *of the permutation π in S_n is the directed graph with ordered vertex set $(\pi_1, \pi_2, \ldots, \pi_n)$ and arc set $\{(i, j) \mid \pi_i = j, 1 \leq i \leq n\}$.*

The fact that π is a bijection implies that $\Gamma(\pi)$ decomposes in a single way into disjoint cycles (up to the ordering of cycles and of elements within each cycle), leading to another notation for π based on its *disjoint cycle decomposition*. For instance, when $\pi = \langle 4\ 1\ 6\ 2\ 5\ 7\ 3 \rangle$, the disjoint cycle notation is $\pi = (1, 4, 2)(3, 6, 7)(5)$ (notice the parentheses and the commas).

DEFINITION 2.3. *The* length *of a cycle in a graph is the number of vertices it contains, and a k-*cycle *is a cycle of length k.*

The number of cycles in a graph G will be denoted by $c(G)$, and the number of cycles of length k will be denoted by $c_k(G)$. We will also distinguish between cycles of odd (resp., even) length, denoting the number of such cycles in G using $c_{odd}(G)$ (resp., $c_{even}(G)$). It is common practice to omit 1-cycles in the cycle decomposition of (the graph of) a permutation and to call that permutation a k-cycle if the resulting decomposition consists of a single cycle of length $k > 1$. Cycles of length 1 in the disjoint cycle decomposition of a permutation are referred to as *fixed points*.

DEFINITION 2.4. *A permutation π is* even *if the number of even cycles in $\Gamma(\pi)$ is even or, equivalently, if it can be expressed as a product of an even number of 2-cycles.*

The *alternating group* A_n is the subgroup of S_n formed by the set of all even permutations, together with $\circ$. The following notion will be central to this work.

DEFINITION 2.5. *The* conjugate *of a permutation* π *by a permutation* σ*, both in* S_n*, is the permutation* $\pi^\sigma = \sigma\circ\pi\circ\sigma^{-1}$ *and can be obtained by replacing every element* i *in the disjoint cycle decomposition of* π *with* σ_i*. All permutations in* S_n *that can be obtained from one another using this operation form a* conjugacy class *(of* S_n*).*

2.2. Generating sets and edit distances. We are interested in distances between permutations based on operations that can themselves be modeled as permutations. More formally, given a subset S of S_n and two permutations π and σ in S_n, we have two goals:

1. to find a sequence of elements $\alpha, \beta, \ldots, \omega$ from S whose length is minimum and whose product transforms π into σ (or conversely, σ into π),
$$\pi\circ\alpha\circ\beta\circ\cdots\circ\omega = \sigma;$$
2. to find the length of such a sequence, called the S *distance* between π and σ. Distances whose definition is based on a set of allowed operations as described above are often referred to as *edit distances*.

Note that S must be *symmetric*, i.e., $\gamma\in S$ if and only if $\gamma^{-1}\in S$, for the corresponding distance to satisfy the symmetry axiom. An immediate corollary of this property is that for any π in S_n, we have $d(\pi,\iota) = d(\pi^{-1},\iota)$. For any two permutations of the same set to be a finite distance apart, S must also satisfy the following property.

DEFINITION 2.6. *A set* $S\subset S_n$ *is said to* generate S_n*, or to be a* generating set *of* S_n*, if every element of* S_n *can be expressed as the product of a finite number of elements of* S*. We call the elements of* S generators *of* S_n*.*

Moreover, all generating sets we will consider in this paper yield distances that satisfy the following property.

DEFINITION 2.7. *A distance* d *on* S_n *is* left-invariant *if for all* π*,* σ*,* τ *in* S_n*, we have* $d(\pi,\sigma) = d(\tau\circ\pi,\tau\circ\sigma)$*.*

Intuitively, left invariance models the fact that given any two permutations π and σ to be transformed into one another, we can rename the elements of either permutation as we wish without changing the value of the distance between both permutations, as long as we renumber the elements of the other permutation accordingly. Since most of the time we will be considering the distance between a permutation π and the identity permutation ι, we will often abbreviate $d(\pi,\iota)$ to $d(\pi)$.

It can be easily seen that both problems mentioned at the beginning of this section can be reformulated in terms of finding a minimum-length factorization of π that consists only of elements of S, since
$$\pi\circ\alpha\circ\beta\circ\cdots\circ\omega = \iota \Leftrightarrow \pi = \omega^{-1}\circ\cdots\circ\beta^{-1}\circ\alpha^{-1}$$
and S is symmetric. Finally, another parameter of interest in the study of those distances is the largest value they can reach.

DEFINITION 2.8. *The* diameter *of a set* U *under a distance* d *is* $\max_{s,t\in U} d(s,t)$*.*

2.3. Genome rearrangements and the cycle graph. We recall here a few operations that are commonly used in the fields of genome rearrangements and interconnection network design to build generating sets of S_n.

DEFINITION 2.9 (see [9]). *The* block-interchange $\beta(i,j,k,l)$ *with* $1\le i<j\le k<l\le n+1$ *is the permutation that exchanges the closed intervals determined respectively by* i *and* $j-1$ *and by* k *and* $l-1$*:*

$$\left\langle \begin{array}{l} 1\ \cdots\ i-1\ \boxed{i\ \cdots\ j-1}\ j\ j+1\ \cdots\ k-1\ \boxed{k\ \cdots\ l-1}\ l\ l+1\ \cdots\ n \\ 1\ \cdots\ i-1\ \boxed{k\ \cdots\ l-1}\ j\ j+1\ \cdots\ k-1\ \boxed{i\ \cdots\ j-1}\ l\ l+1\ \cdots\ n \end{array} \right\rangle.$$

Two particular cases of block-interchanges are of interest:

1. When $j = k$, the resulting operation exchanges two adjacent intervals and is called a *transposition* [3], denoted by $\tau(i, j, l)$.
2. When $j = i+1$ and $l = k+1$, the resulting operation swaps two not necessarily adjacent elements in respective positions i and k and is called an *exchange*, denoted by $\varepsilon(i, k)$.

We use the notation $bid(\pi)$, $td(\pi)$, and $exc(\pi)$ for the block-interchange distance, the transposition distance, and the exchange distance of π, respectively. The operations we described above can be further restricted by setting $i = 1$ in their definition, thereby transforming them into so-called prefix rearrangements. The corresponding "prefix distances" are defined in an analogous manner with the additional restriction that all operations must act on the initial segment of the permutation. We denote $ptd(\pi)$ and $pexc(\pi)$ the *prefix transposition distance* and *prefix exchange distance* of π, respectively. While sorting by transpositions is NP-hard [5] and the computational complexity of sorting by prefix transpositions is unknown, polynomial-time algorithms exist for sorting by block-interchanges [9], exchanges [20], or prefix exchanges [1], as well as formulas for computing the associated distances.

We will have more to say about sorting by transpositions and sorting by block-interchanges in section 4, where we will give simple proofs of lower bounds on the two corresponding distances, as well as about sorting by prefix transpositions in section 5, where we will prove new and improved lower bounds on the associated distance and diameter. Meanwhile, we conclude this section with the following traditional tool introduced by Bafna and Pevzner [3], which has proved most useful in the study of genome rearrangements.

DEFINITION 2.10. *The* cycle graph *of a permutation π in S_n is the bicolored directed graph $G(\pi)$, whose vertex set $(\pi_0 = 0, \pi_1, \ldots, \pi_n)$ is ordered by positions and whose arc set consists of*

- black *arcs* $\{(\pi_i, \pi_{i-1}) \mid 1 \leq i \leq n\} \cup \{(\pi_0, \pi_n)\}$,
- gray *arcs* $\{(\pi_i, \pi_i + 1) \mid 0 \leq i \leq n\} \cup \{(n, 0)\}$.

The arc set of $G(\pi)$ decomposes in a single way into arc-disjoint *alternating cycles*, i.e., cycles that alternate black and gray arcs. The *length* of an alternating cycle in $G(\pi)$ is the number of black arcs it contains, and a *k-cycle* in $G(\pi)$ is an alternating cycle of length k (note that this differs from Definition 2.3). Figure 2.1 shows an example of a cycle graph, together with its decomposition into a 5-cycle and a 3-cycle.

3. A general lower bounding technique. We now present a framework for obtaining lower bounds on edit distances between permutations in a simple and unified way. To that end, we adapt a bijection previously introduced by Doignon and Labarre [11],

$$f : S_n \to A_{n+1} : \pi \mapsto \overline{\pi} = (0, 1, 2, \ldots, n) \circ (0, \pi_n, \pi_{n-1}, \ldots, \pi_1), \tag{3.1}$$

which in particular maps ι onto $\overline{\iota} = \langle 0\ 1\ 2\ \cdots\ n\rangle$. (Recall that the notation $(0, 1, 2, \ldots, n)$, introduced right after Definition 2.2, refers to an $(n+1)$-cycle, not to the identity permutation.) That mapping allows us to encode the structure of a cycle graph $G(\pi)$ using an even permutation $\overline{\pi}$ in an intuitive way, which corresponds to decomposing the cycle graph into the product of two "monochromatic cycles," namely,

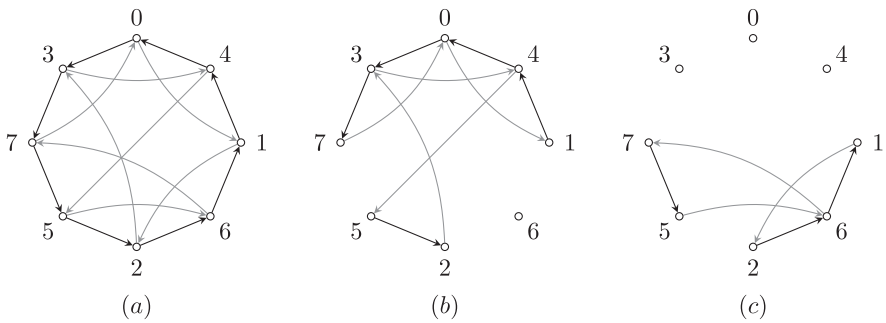


FIG. 2.1. (a) *The cycle graph of* $\langle 4\ 1\ 6\ 2\ 5\ 7\ 3 \rangle$*;* (b), (c) *its decomposition into two alternating cycles.*

the cycle made of all black arcs (i.e., $(0, \pi_n, \pi_{n-1}, \ldots, \pi_1)$) and the cycle made of all gray arcs (i.e., $(0, 1, 2, \ldots, n)$). The construction is perhaps best understood using an example: let $\pi = \langle 4\ 1\ 6\ 2\ 5\ 7\ 3 \rangle$, whose cycle graph is depicted in Figure 2.1(a). Then

$$\overline{\pi} = (0, 1, 2, 3, 4, 5, 6, 7) \circ (0, 3, 7, 5, 2, 6, 1, 4) = (0, 4, 1, 5, 3)(2, 7, 6),$$

and the two disjoint cycles of $\overline{\pi}$ correspond to the two alternating cycles of $G(\pi)$, whose elements they list in the order they are encountered (up to rotation):

1. The first cycle of $G(\pi)$ (Figure 2.1(b)) starts with 0, then visits 4 after following a black-gray path (i.e., a black arc followed by a gray arc), then visits 1 after following a black-gray path, and in the same way visits 5 and 3 before coming back to 0, which corresponds to the first cycle of $\overline{\pi}$.
2. The second cycle of $G(\pi)$ (Figure 2.1(c)) starts with 2, then visits 7 after following a black-gray path, and in the same way visits 6 before coming back to 2, which corresponds to the second cycle of $\overline{\pi}$.

Note that the order in which we decide to follow arcs (first a black arc and then a gray arc) is given by the order in which the two cycles are multiplied. An alternative definition[1] of $\overline{\pi}$ could therefore have been $(0, \pi_n, \pi_{n-1}, \ldots, \pi_1) \circ (0, 1, 2, \ldots, n)$, which can be seen to be equivalent to our definition when conjugated by $(0, n, n-1, \ldots, 1)$ and whose cycles are interpreted exactly as above, with the modification that gray arcs are followed first. Consequently, speaking about cycles of $\overline{\pi}$, of $\Gamma(\overline{\pi})$, or of $G(\pi)$ is equivalent. We will now demonstrate how $f(\cdot)$ can be used to obtain results on the sorting and distance computation problems we discussed in section 2.2. The following lemma expresses how the action of *any* rearrangement operation σ on π is translated on $\overline{\pi}$. We will find it convenient to identify permutations in S_n with their extended versions in S_{n+1} (i.e., we identify π with $\langle 0\ \pi_1\ \pi_2\ \cdots\ \pi_n \rangle$). This allows us to express any permutation π in S_n as follows:

$$\pi = (0, \pi_n, \pi_{n-1}, \ldots, \pi_1) \circ \pi \circ (0, 1, 2, \ldots, n). \tag{3.2}$$

LEMMA 3.1. *For all* π, σ *in* S_n, *we have* $\overline{\pi \circ \sigma} = \overline{\pi} \circ \overline{\sigma}^{\pi}$.

[1]This is actually the definition we used in the conference version of this paper [24].

Proof. By definition, we have

$$\begin{aligned}\overline{\pi\circ\sigma} &= (0,1,2,\dots,n)\circ(0,(\pi\circ\sigma)_n,(\pi\circ\sigma)_{n-1},\dots,(\pi\circ\sigma)_1)\\ &= (0,1,2,\dots,n)\circ\pi\circ(0,\sigma_n,\sigma_{n-1},\dots,\sigma_1)\circ\pi^{-1}\\ &= (0,1,2,\dots,n)\circ(0,\pi_n,\pi_{n-1},\dots,\pi_1)\circ\pi\circ(0,1,2,\dots,n)\\ &\quad\circ(0,\sigma_n,\sigma_{n-1},\dots,\sigma_1)\circ\pi^{-1} \qquad \text{(using (3.2))}\\ &= \overline{\pi}\circ\overline{\sigma}^{\pi}. \qquad \square\end{aligned}$$

We are now ready to prove our main result.

THEOREM 3.2. *Let S be a subset of S_n whose elements are mapped by $f(\cdot)$ onto $S' \subseteq A_{n+1}$. Moreover, let $\mathscr{C}$ be the union of the conjugacy classes (of S_{n+1}) that intersect with S'; then for any π in S_n, any factorization of π into t elements of S yields a factorization of $\overline{\pi}$ into t elements of $\mathscr{C}$.*

Proof. Induction on t. The base case is $\pi \in S$, and clearly $\overline{\pi} \in S' \subseteq \mathscr{C}$. For the induction, let $\pi = \omega\circ\psi\circ\cdots\circ\beta\circ\alpha$, where $\alpha,\beta,\dots,\psi,\omega \in S$, and let $\sigma = \psi\circ\cdots\circ\beta\circ\alpha$; by Lemma 3.1, we have

$$\overline{\pi} = \overline{\omega\circ\psi\circ\cdots\circ\beta\circ\alpha} = \overline{\omega\circ\sigma} = \overline{\omega}\circ\overline{\sigma}^{\omega}.$$

By induction, $\overline{\sigma} = \psi'\circ\circ\cdots\circ\beta'\circ\alpha'$, where $\alpha',\beta',\dots,\psi' \in \mathscr{C}$; therefore

$$\begin{aligned}\omega\circ\overline{\sigma}\circ\omega^{-1} &= \omega\circ\psi'\circ\cdots\circ\beta'\circ\alpha'\circ\omega^{-1}\\ &= \underbrace{\omega\circ\psi'\circ\omega^{-1}}_{\psi''}\circ\,\omega\circ\cdots\circ\omega^{-1}\circ\underbrace{\omega\circ\beta'\circ\omega^{-1}}_{\beta''}\underbrace{\omega\circ\alpha'\circ\omega^{-1}}_{\alpha''},\end{aligned}$$

and $\alpha'',\beta'',\dots,\psi'' \in \mathscr{C}$, which completes the proof. $\square$

We will use Theorem 3.2 in the next two sections to prove lower bounds on several edit distances between permutations.

4. Recovering previous results. We illustrate how to use Theorem 3.2 to recover two previously known results on *bid* and *td*. The general idea is as follows: as we explained in section 2.2, if S is symmetric, then any sorting sequence of length t for π made of elements of S yields a factorization of π into the product of t elements of S, which can in turn be converted, as in the proof of Theorem 3.2, into a factorization of $\overline{\pi}$ into the product of t elements of $S' \subseteq \mathscr{C}$. Therefore, the length of a shortest factorization of $\overline{\pi}$ into the product of elements of $\mathscr{C}$ is a lower bound on the length of a factorization of π into the product of elements of S, and we can obtain a lower bound on the distance of interest by

1. characterizing the set of images of the elements in S by $f(\cdot)$ and
2. computing the distance of $\overline{\pi}$ with respect to $\mathscr{C}$.

Let us now show how we can obtain a lower bound on the block-interchange distance. We start by characterizing the image of a block-interchange by our mapping.

LEMMA 4.1. *For any block-interchange $\beta(i,j,k,l)$ in S_n, we have*

$$\overline{\beta(i,j,k,l)} = (j,l)\circ(i,k).$$

Proof. Equation (3.1) and Definition 2.9 yield

$$\begin{aligned}&(0,1,2,\dots,n)\circ(0,n,n-1,\dots,l,j-1,j-2,\dots,i,k-1,k-2,\dots,\\ &j,l-1,l-2,\dots,k,i-1,i-2,\dots,1)\\ &\quad= (0)(n)(n-1)\cdots(l+1)(l,j)(l-1)\cdots(i+1)(i,k)(i-1)\cdots(1)\\ &\quad= (j,l)\circ(i,k). \qquad \square\end{aligned}$$

Note that (j, l) and (i, k) might not be disjoint, since Definition 2.9 allows for $j = k$ (hence the use of $\circ$ in the expression of $\overline{\beta(i, j, k, l)}$). We can now recover a known lower bound on the block-interchange distance, which is actually the exact distance as shown by Christie [9].

THEOREM 4.2 (see [9]). *For all π in S_n, we have $bid(\pi) \geq \frac{n+1-c(\Gamma(\overline{\pi}))}{2}$.*

Proof. By Theorem 3.2 and Lemma 4.1, a lower bound on $bid(\pi)$ is given by the length of a minimum factorization of $\overline{\pi}$ into the product of pairs of exchanges. Since this length equals $(n+1-c(\Gamma(\overline{\pi})))/2$ (see, e.g., Jerrum [20]), the proof follows. ☐

Let us now characterize the image of a transposition by our mapping.

LEMMA 4.3. *For any transposition $\tau(i, j, l)$, we have*

$$\overline{\tau(i, j, l)} = (i, l, j).$$

Proof. As noted in section 2.2, we have $\tau(i, j, l) = \beta(i, j, j, l)$; Lemma 4.1 yields

$$\overline{\tau(i, j, l)} = \overline{\beta(i, j, j, l)} = (j, l) \circ (i, j) = (i, l, j). \qquad ☐$$

We can now recover the following known lower bound on the transposition distance. Recall that $c_{odd}(\Gamma(\overline{\pi}))$ denotes the number of odd cycles in $\Gamma(\overline{\pi})$.

THEOREM 4.4 (see [3]). *For all π in S_n, we have $td(\pi) \geq \frac{n+1-c_{odd}(\Gamma(\overline{\pi}))}{2}$.*

Proof. By Theorem 3.2 and Lemma 4.3, a lower bound on $td(\pi)$ is given by the length of a minimum factorization of $\overline{\pi}$ into the product of 3-cycles. Since this length equals $(n+1-c_{odd}(\Gamma(\overline{\pi})))/2$ (see, e.g., Jerrum [20]), the proof follows. ☐

5. New results on the prefix transposition distance. Dias and Meidanis [10] initiated the study of sorting by prefix transpositions and derived a lower bound on the corresponding distance using the following concepts.

DEFINITION 5.1. *Given a permutation π in S_n, build the permutation $\widetilde{\pi} = \langle 0\ \pi_1\ \cdots\ \pi_n\ n+1 \rangle$; a pair $(\widetilde{\pi}_i, \widetilde{\pi}_{i+1})$ with $0 \leq i \leq n$ is a* prefix transposition breakpoint *if $\widetilde{\pi}_{i+1} \neq \widetilde{\pi}_i + 1$ or if $i = 0$ and an* adjacency *otherwise.*

The number of prefix transposition breakpoints of π is denoted by $ptb(\pi)$. Noting that a prefix transposition can create at most two adjacencies and that ι is the only permutation with one prefix transposition breakpoint, Dias and Meidanis obtained the following lower bound.

LEMMA 5.2 (see [10]). *For any π in S_n*

$$ptd(\pi) \geq \left\lceil \frac{ptb(\pi) - 1}{2} \right\rceil. \tag{5.1}$$

Chitturi and Sudborough [7] later obtained another lower bound on the prefix transposition distance. They used the following concepts, based on permutations of $\{0, 1, 2, \ldots, n-1\}$ rather than $\{1, 2, \ldots, n\}$.

DEFINITION 5.3. *For a permutation π of $\{0, 1, 2, \ldots, n-1\}$, an ordered pair (π_i, π_{i+1}) is an* antiadjacency *if $\pi_{i+1} = \pi_i - 1 \pmod n$. A* strip *in a permutation π is a maximal interval of π that contains only adjacencies, and a* clan *is a maximal interval of π that contains only antiadjacencies.*

Chitturi and Sudborough proved the following lower bound.

LEMMA 5.4 (see [7]). *For any permutation π of $\{0, 1, 2, \ldots, n-1\}$, let $\Upsilon(\pi)$ denote the set of all clans of π of length at least 3 and $s(\pi)$ denote the number of strips of π. Then*

$$ptd(\pi) \geq \frac{s(\pi) + \frac{\sum_{C \in \Upsilon(\pi)}(|C|-2)}{3}}{2}. \tag{5.2}$$

We will prove a new lower bound on the prefix transposition distance (Theorem 5.8) using our model and the results of Akers, Krishnamurthy, and Hare [1] on computing the prefix exchange distance.

THEOREM 5.5 (see [1]). *For any π in S_n, we have*

$$pexc(\pi) = n + c(\Gamma(\pi)) - 2c_1(\Gamma(\pi)) - \begin{cases} 0 & \text{if } \pi_1 = 1, \\ 2 & \text{otherwise}, \end{cases}$$

where $c_1(\Gamma(\pi))$ denotes the number of 1-cycles in $\Gamma(\pi)$ or equivalently the number of fixed points of π.

5.1. An improved lower bound. Using our theory, we prove a *new* lower bound on $ptd(\pi)$ and show that it always outperforms (5.1). We will find it convenient to express $ptb(\pi)$ as follows.

LEMMA 5.6. *For any π in S_n, we have*

$$ptb(\pi) = n + 1 - c_1(\Gamma(\overline{\pi})) + \begin{cases} 1 & \text{if } \pi_1 = 1, \\ 0 & \text{otherwise}. \end{cases}$$

Proof. The formula results from the observation that among the $n+1$ pairs of adjacent elements in $\widetilde{\pi}$, each adjacency in $\widetilde{\pi}$ gives rise to a 1-cycle in $\Gamma(\overline{\pi})$, and from the fact that if $\pi_1 = 1$, then we counted the 1-cycle that corresponds to $(0, 1)$ as an adjacency, which is contrary to Definition 5.1 and which we correct by adding 1. ☐

As explained in section 4, we can obtain a lower bound on the prefix transposition distance by characterizing the image of a prefix transposition by $f(\cdot)$ and computing the associated distance. We already know that transpositions are mapped onto 3-cycles (see Lemma 4.3); in the case of prefix transpositions, it is easily seen that these 3-cycles will always contain element 0. Therefore, we need to be able to compute the length of a minimum factorization of π in S_n into a product of 3-cycles, where each 3-cycle in the factorization is further required to contain the first element. Let us denote the corresponding distance $d_3^1(\pi)$; the following result shows how to compute it.

LEMMA 5.7. *For any π in A_n, we have*

$$d_3^1(\pi) = \frac{n + c(\Gamma(\pi))}{2} - c_1(\Gamma(\pi)) - \begin{cases} 0 & \text{if } \pi_1 = 1, \\ 1 & \text{otherwise}. \end{cases}$$

Proof. Given a minimum factorization of length ℓ of an even permutation π into prefix exchanges, we can construct a sequence of $\ell/2$ 3-cycles by noting that $(1, j) \circ (1, i) = (1, i, j)$. Therefore $d_3^1(\pi) \leq \ell/2$. On the other hand, assume there exists a shorter sequence of 3-cycles acting on the first element whose product is π; then one can split each of these 3-cycles into two prefix exchanges using the relation above and find a shorter expression for π as a product of prefix exchanges, a contradiction. The result follows from Theorem 5.5. ☐

As a corollary, we obtain the following new lower bound on the prefix transposition distance.

THEOREM 5.8. *For any π in S_n, we have*

$$ptd(\pi) \geq \frac{n + 1 + c(\Gamma(\overline{\pi}))}{2} - c_1(\Gamma(\overline{\pi})) - \begin{cases} 0 & \text{if } \pi_1 = 1, \\ 1 & \text{otherwise}. \end{cases} \tag{5.3}$$

Proof. The proof follows immediately from Theorem 3.2 and Lemma 5.7. ☐

An immediate question is how tight this new lower bound actually is. We will answer this question experimentally in section 6, where we will see that many more permutations are tight with respect to our new result than with respect to the previously known lower bounds. We will in the meantime conclude this section by proving that our lower bound always outperforms that of Dias and Meidanis (given by Lemma 5.2).

THEOREM 5.9. *For all π in S_n, the value of lower bound* (5.3) *is never smaller than that of lower bound* (5.1).

Proof. Assume $\pi \neq \iota$ (otherwise the result trivially holds); this implies that $\Gamma(\overline{\pi})$ has at least one cycle of length at least 2, which means that $c(\Gamma(\overline{\pi})) - c_1(\Gamma(\overline{\pi})) \geq 1$. There are two cases to prove. If $\pi_1 = 1$, then lower bound (5.1) becomes

$$\left\lceil \frac{(n+1-c_1(\Gamma(\overline{\pi}))+1)-1}{2} \right\rceil = \left\lceil \frac{n+1-c_1(\Gamma(\overline{\pi}))}{2} \right\rceil,$$

and lower bound (5.3) satisfies

$$\frac{n+1+c(\Gamma(\overline{\pi}))-2c_1(\Gamma(\overline{\pi}))}{2} \geq \frac{n+2-c_1(\Gamma(\overline{\pi}))}{2} \geq \left\lceil \frac{n+1-c_1(\Gamma(\overline{\pi}))}{2} \right\rceil.$$

On the other hand, if $\pi_1 \neq 1$, then lower bound (5.1) becomes

$$\left\lceil \frac{(n+1-c_1(\Gamma(\overline{\pi})))-1}{2} \right\rceil = \left\lceil \frac{n-c_1(\Gamma(\overline{\pi}))}{2} \right\rceil,$$

and Definition 2.4 implies that for any π in A_n, we have $n \equiv c(\Gamma(\pi)) \pmod 2$. Lower bound (5.3) becomes

$$\begin{aligned}\frac{n+1+c(\Gamma(\overline{\pi}))}{2} - c_1(\Gamma(\overline{\pi})) - 1 &= \left\lceil \frac{n+1+c(\Gamma(\overline{\pi}))-2c_1(\Gamma(\overline{\pi}))-2}{2} \right\rceil \\ &\geq \left\lceil \frac{n-c_1(\Gamma(\overline{\pi}))}{2} \right\rceil. \qquad \square\end{aligned}$$

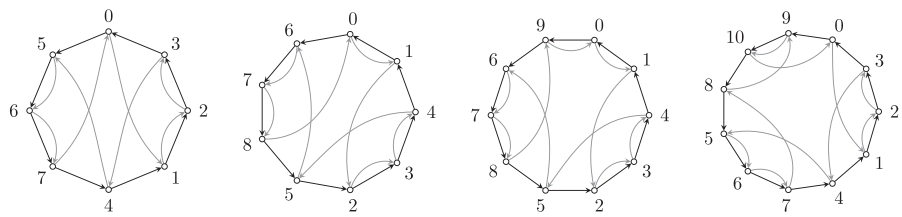
5.2. A tighter lower bound on the prefix transposition diameter. Dias and Meidanis [10] observed that the prefix transposition diameter lies between $n/2$ and $n-1$, and conjectured that it is equal to $n - \lfloor \frac{n}{4} \rfloor$. Chitturi and Sudborough [7, 8] then improved those bounds to $2n/3$ and $n - \log_{9/2} n$, respectively. Using our new lower bound, we further improve the lower bound on the prefix transposition diameter. We prove our result in a constructive way, by building families of permutations whose prefix transposition distance is at least $\lfloor 3n/4 \rfloor$. Figure 5.1, which follows our result, shows examples of such permutations. The proof uses permutations from the following class, which has proved useful in the analysis of several other rearrangement problems [14].

DEFINITION 5.10. *A permutation π in S_n is a* 2-permutation *if all cycles in $\overline{\pi}$ have length* 2.

Note that the above definition requires $n \equiv 3 \pmod 4$: indeed, $n+1$ must be even in order to obtain a partition of the elements of $\overline{\pi}$ into pairs, and $(n+1)/2$ is also even by the definition of $\overline{\pi}$.

THEOREM 5.11. *For all n, the prefix transposition diameter of S_n is at least $\lfloor 3n/4 \rfloor$.*

Proof. If $n = 1$ or 2, the result is easily verified. For $n \geq 3$, we construct for each value of $n \pmod 4$ a suitable permutation. Figure 5.1 shows an example for each case of the proof.



$n \equiv 3 \pmod 4$ $n \equiv 0 \pmod 4$ $n \equiv 1 \pmod 4$ $n \equiv 2 \pmod 4$

FIG. 5.1. *Cycle graphs of permutations* π *in* S_n *with* $ptd(\pi) \geq \lfloor 3n/4 \rfloor$*, for all values of* n mod 4.

1. If $n \equiv 3 \pmod 4$, then any 2-permutation π in S_n is a valid candidate: indeed, $\overline{\pi}$ contains in this case exactly $(n+1)/2$ cycles of length 2, and Theorem 5.8 yields
$$ptd(\pi) \geq \frac{n+1+\frac{n+1}{2}}{2} - 1 = \frac{3n+3-4}{4} = \frac{3n-1}{4}.$$
2. If $n \equiv 0 \pmod 4$, we build a permutation σ in S_n by inserting a new first element as a fixed point in $\overline{\pi}$, where π is the permutation in S_{n-1} constructed in the previous case. $\overline{\sigma}$ contains $n/2$ cycles of length 2 and one cycle of length 1 that corresponds to the fact that $\sigma_1 = 1$. Theorem 5.8 then yields
$$ptd(\sigma) \geq \frac{n+1+\frac{n}{2}+1}{2} - 1 = \frac{2n+2+n+2-4}{4} = \frac{3n}{4}.$$
3. If $n \equiv 1 \pmod 4$, we build a permutation ξ in S_n by inserting a fixed point anywhere in $\overline{\sigma}$, where σ is the permutation in S_{n-1} built in the previous case. $\overline{\xi}$ contains $(n+1-2)/2$ cycles of length 2 and two cycles of length 1, and $\xi_1 = 1$. Theorem 5.8 then yields
$$ptd(\xi) \geq \frac{n+1+\frac{n+1-2}{2}+2}{2} - 2 = \frac{2n+2+n+1-2+4-8}{4} = \frac{3n-3}{4}.$$
4. If $n \equiv 2 \pmod 4$, we build a permutation τ in S_n by appending a 3-cycle to any permutation $\overline{\pi}$ such that π is a 2-permutation in S_{n-3}. $\overline{\tau}$ contains $(n+1-3)/2$ cycles of length 2 and one cycle of length 3, and Theorem 5.8 yields
$$ptd(\tau) \geq \frac{n+1+\frac{n+1-3}{2}+1}{2} - 1 = \frac{2n+2+n+1-3+2-4}{4} = \frac{3n-2}{4}. \qquad \square$$

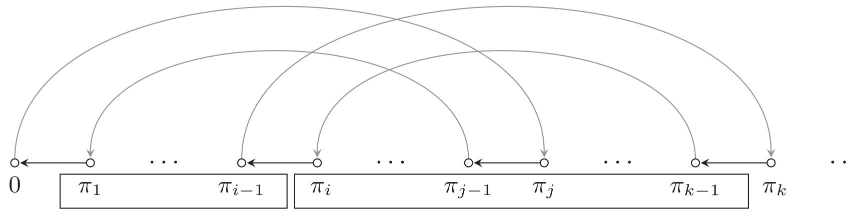
We can actually show that the lower bound on the prefix transposition distance of 2-permutations is tight. In order to do that, we will need the following result. We use the following relation to order black arcs: $(\pi_i, \pi_{i-1}) \prec (\pi_j, \pi_{j-1})$ if $j \geq i$.

LEMMA 5.12 (see [3]). *For any* π *in* S_n*, let* C_1 *be a cycle of length* 2 *in* $G(\pi)$ *with black arcs* a*,* b*; then there exists another cycle* C_2 *in* $G(\pi)$ *containing two black arcs* c *and* d *such that* $a \prec c \prec b \prec d$ *or* $c \prec a \prec d \prec b$*.*

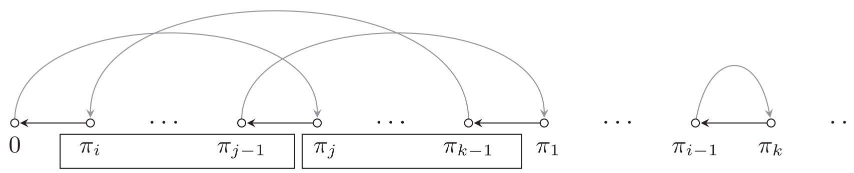
This result can be interpreted in a more visual way in the case of a 2-permutation π by saying that in $G(\pi)$, every 2-cycle intersects with another 2-cycle. We are now ready to prove the following result.

PROPOSITION 5.13. *For any* 2*-permutation* π *in* S_n*, we have* $ptd(\pi) = (3n-1)/4$.

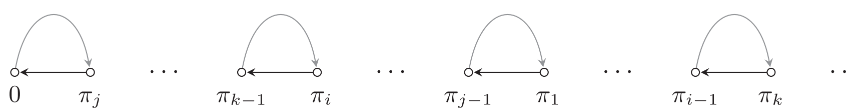
Proof. The lower bound has already been observed in Theorem 5.11. To show that it is also an upper bound, we give an algorithm that sorts π in exactly that number of steps. By Lemma 5.12, every 2-cycle intersects with another 2-cycle, and as observed by Bafna and Pevzner [3], a sequence of two transpositions on any two crossing 2-cycles will transform them into four adjacencies:



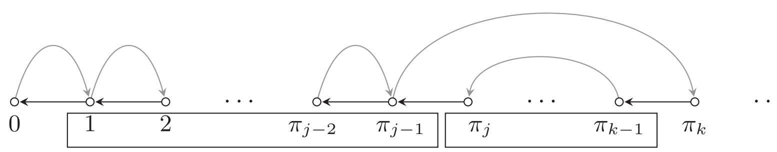
becomes



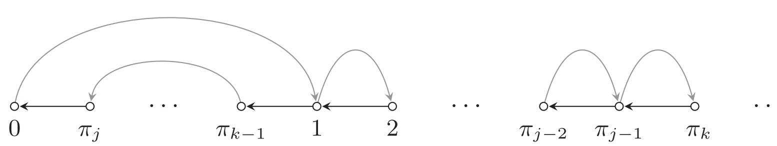
which becomes



We transform the leftmost 2-cycle and any 2-cycle it crosses into four adjacencies using two prefix transpositions, which transforms π into a permutation σ that contains $\frac{n+1}{2} - 2$ cycles of length 2 and fixes the first element. Then, we carry out this process again until σ is sorted, but we need three prefix transpositions at each step, since one move must be wasted to move the fixed points in σ's prefix out of the way, for instance, as follows:



becomes



The algorithm is guaranteed to terminate, since after applying each sequence of three transpositions of the form described above, we obtain either ι or a permutation on which we can repeat the same process by Lemma 5.12. The proof follows from the fact that the number of prefix transpositions used by this algorithm is

$$2 + \frac{3}{2}\left(\frac{n+1}{2} - 2\right) = \frac{8 + 3n - 9}{4} = \frac{3n-1}{4}. \qquad \square$$

Table 6.1

Comparison of all known lower bounds on the prefix transposition distance. Column 3 lists the number of cases where Dias and Meidanis's lower bound is tight [15, p. 48]*, column 4 lists the number of cases where Chitturi and Sudborough's lower bound is tight, and column 5 lists the number of cases where our lower bound is tight.*

n	$n!$	Tight w.r.t. (5.1)	Tight w.r.t. (5.2)	Tight w.r.t. (5.3)
1	1	1	1	1
2	2	2	2	2
3	6	4	4	6
4	24	13	15	22
5	120	41	48	106
6	720	196	255	574
7	5 040	862	1 144	3 782
8	40 320	5 489	7 737	27 471
9	362 880	31 033	44 187	229 167
10	3 628 800	247 006	369 979	2 103 510
11	39 916 800	1 706 816	2 575 693	21 280 564
12	479 001 600	16 302 397	25 791 862	236 651 919

6. Experimental results. We generated all permutations in S_n for $1 \leq n \leq 12$, along with their prefix transposition distance, and compared lower bounds (5.1), (5.2), and (5.3) to the actual distance. Table 6.1 shows the results. It can be observed that many more permutations are tight with respect to our lower bound (column 5) than with respect to that of Dias and Meidanis (column 3) or Chitturi and Sudborough (column 4).

We also examined how large the gap between our lower bound and the actual prefix transposition distance can get. Table 6.2 counts permutations whose prefix transposition distance equals our lower bound plus Δ. We note that for $n \leq 9$, all permutations have a prefix transposition distance that is at most our lower bound plus 2 (plus 3 for $n \leq 12$).

Table 6.2

Number of cases where our lower bound underestimates $ptd(\pi)$ *by* Δ.

n	$n!$	$\Delta = 0$	$\Delta = 1$	$\Delta = 2$	$\Delta = 3$
1	1	1	0	0	0
2	2	2	0	0	0
3	6	6	0	0	0
4	24	22	2	0	0
5	120	106	14	0	0
6	720	574	143	3	0
7	5 040	3 782	1 234	24	0
8	40 320	27 471	12 310	539	0
9	362 880	229 167	128 576	5 137	0
10	3 628 800	2 103 510	1 427 966	97 321	3
11	39 916 800	21 280 564	17 532 948	1 103 254	34
12	479 001 600	236 651 919	221 680 237	20 667 140	2 304

7. Further observations on $\overline{\pi}$. Now that we have an alternate representation of the cycle graph of a permutation as another permutation, we would like to examine whether other results can be obtained that could be helpful in getting insight on problems related to length-constrained factorizations of permutations. We investigate in this section a few relations between π and $\overline{\pi}$, starting with relations between the cycle structures of both permutations when subjected to particular operations.

7.1. Cycle structures. A natural question is whether conjugacy classes are preserved by $f(\cdot)$, i.e., whether $\overline{\pi}$ and $\overline{\pi^\sigma}$ are in the same conjugacy class for any choice of π and σ in S_n. The answer is negative in general, as the following counterexample shows: $\pi = (1,3)(2)$ and $\tau = (1,2)(3)$ are conjugate, but $\overline{\pi} = (0,1,2,3) \circ (0,1,2,3) = (0,2)(1,3)$ and $\overline{\tau} = (0,1,2,3) \circ (0,3,1,2) = (0)(3,2,1)$ are not. However, the relation we are interested in holds for two particular cases, whose significance we explain below.

The following result is similar in spirit to Tannier and Sagot's characterization of "inverse breakpoint graphs" of signed permutations [28] and shows that the cycle graphs of a permutation and of its inverse have exactly the same cycle structure.

LEMMA 7.1. *For any π in S_n, we have $\overline{\pi^{-1}} = (\overline{\pi}^{-1})^{(\pi^{-1})}$.*

Proof. The proof is straightforward:

$$\begin{aligned}\overline{\pi^{-1}} &= (0,1,2,\ldots,n) \circ (0, \pi^{-1}_n, \pi^{-1}_{n-1}, \ldots, \pi^{-1}_1)\\ &= \pi^{-1} \circ \pi \circ (0,1,2,\ldots,n) \circ \pi^{-1} \circ (0,n,n-1,\ldots,1) \circ \pi\\ &= \pi^{-1} \circ (0,\pi_1,\pi_2,\ldots,\pi_n) \circ (0,n,n-1,\ldots,1) \circ \pi\\ &= \pi^{-1} \circ \overline{\pi}^{-1} \circ \pi\\ &= (\overline{\pi}^{-1})^{(\pi^{-1})}. \qquad \square\end{aligned}$$

Tannier and Sagot's idea of examining how the cycle graph of π^{-1} evolves when applying a signed reversal to a permutation π—which reverses and flips the signs of the elements of an interval of π—was a key point in their successful attempt at designing an algorithm with an improved running time for sorting permutations by signed reversals. The above relation allows us to derive a simple description of the more general situation (albeit restricted to "traditional", unsigned permutations), i.e., how $\overline{\pi^{-1}}$ changes when an arbitrary rearrangement σ is applied to π.

COROLLARY 7.2. *For all π, σ in S_n, we have $\overline{(\pi \circ \sigma)^{-1}} = (\overline{\sigma}^{-1} \circ \overline{\pi^{-1}})^{\sigma^{-1}}$.*

Proof. Lemma 3.1 yields

$$\begin{aligned}\overline{(\pi \circ \sigma)^{-1}} = \overline{\sigma^{-1} \circ \pi^{-1}} &= \overline{\sigma^{-1}} \circ \overline{\pi^{-1}}^{(\sigma^{-1})}\\ &= \sigma^{-1} \circ \overline{\sigma}^{-1} \circ \sigma \circ \sigma^{-1} \circ \overline{\pi^{-1}} \circ \sigma \quad \text{(using Lemma 7.1)}\\ &= (\overline{\sigma}^{-1} \circ \overline{\pi^{-1}})^{\sigma^{-1}}. \qquad \square\end{aligned}$$

A second particular case of conjugate permutations whose transformation by $f(\cdot)$ yields two conjugate permutations is presented below. We use the notation χ for the *reverse permutation*, i.e., $\chi = \langle n\ n-1\ \cdots\ 1\rangle$.

LEMMA 7.3. *For any π in S_n, we have $\overline{\pi^\chi} = ((\overline{\pi}^{-1})^\chi)^{(0,1,2,\ldots,n)}$.*

Proof. We have by definition

$$\begin{aligned}\overline{\pi^\chi} &= (0,1,2,\ldots,n) \circ (0, \pi^\chi_n, \pi^\chi_{n-1}, \ldots, \pi^\chi_1)\\ &= (0,1,2,\ldots,n) \circ \pi^\chi \circ (0,n,n-1,\ldots,1) \circ (\pi^\chi)^{-1}\\ &= (\pi^\chi \circ (0,n,n-1,\ldots,1) \circ (\pi^\chi)^{-1} \circ (0,1,2,\ldots,n))^{(0,1,2,\ldots,n)}\\ &= (\chi \circ \pi \circ (0,1,2,\ldots,n) \circ \pi^{-1} \circ \chi \circ (0,1,2,\ldots,n))^{(0,1,2,\ldots,n)}\\ &= (\chi \circ \pi \circ (0,1,2,\ldots,n) \circ \pi^{-1} \circ (0,n,n-1,\ldots,1) \circ \chi)^{(0,1,2,\ldots,n)}\\ &= (\chi \circ \overline{\pi}^{-1} \circ \chi)^{(0,1,2,\ldots,n)}. \qquad \square\end{aligned}$$

Conjugating π by χ corresponds to computing its reverse complement: indeed, $\pi \circ \chi = \langle \pi_n\ \pi_{n-1}\ \cdots\ \pi_1\rangle$, and $\chi \circ (\pi \circ \chi) = \langle n+1-\pi_n\ n+1-\pi_{n-1}\ \cdots\ n+1-\pi_1\rangle$. By definition, π and π^χ have the same cycle structure, and by the above result so do their images by $f(\cdot)$. The reverse complement operation is interesting because most (but not all, prefix distances being notable exceptions [23]) genome rearrangement distances are, in addition to being left-invariant, "reverse complement-invariant": for all π and σ in S_n, we have $d(\pi,\sigma) = d(\pi^\chi, \sigma^\chi)$. As a consequence, bounds obtained on the distance between π and ι with respect to a certain set of operations can sometimes be improved by examining π^{-1} or π^χ.

Eriksson et al. [12] introduced another important equivalence relation on permutations that does not preserve their cycle structure in the classical sense but that does preserve the cycle structure of their cycle graphs. This equivalence relation, whose equivalence classes are called *toric permutations*, proved useful in improving bounds on the transposition distance [12, 22]. We will see below that $\overline{\pi}$ provides a simple way of navigating through all cycle graphs of the permutations in the same equivalence class. The equivalence relation uses the following notion.

Definition 7.4. *The* circular permutation *obtained from a permutation π in S_n is $\pi^\circ = 0\ \pi_1\ \pi_2\ \cdots\ \pi_n$ with indices taken modulo $n+1$ so that $0 = \pi^\circ_0 = \pi^\circ_{n+1}$.*

This circular permutation can be read starting from any position, and the original "linear" permutation is reconstructed by taking the element following 0 as π_1 and removing 0. For x in $\{0, 1, 2, \ldots, n\}$, let $\overline{x}^m = (x+m) \pmod{n+1}$, and define the following operation on circular permutations:

$$m + \pi^\circ = \overline{0}^m\ \overline{\pi_1}^m\ \overline{\pi_2}^m\ \cdots\ \overline{\pi_n}^m.$$

Definition 7.5. *For any π in S_n, the* toric permutation $\pi^\circ_\circ$ *is the set of permutations in S_n reconstructed from all circular permutations $m + \pi^\circ$ with $0 \le m \le n$.*

Definition 7.6. *Two permutations π, σ in S_n are* torically equivalent *if $\sigma \in \pi^\circ_\circ$ (or $\pi \in \sigma^\circ_\circ$), which we also write as $\pi \equiv^\circ_\circ \sigma$.*

Let us illustrate those notions using our running example $\pi = \langle 4\ 1\ 6\ 2\ 5\ 7\ 3\rangle$; we have $\pi^\circ = 0\ 4\ 1\ 6\ 2\ 5\ 7\ 3$, and

$$\begin{array}{rcl}
0+\pi^\circ & = & 0\,4\,1\,6\,2\,5\,7\,3 \\
1+\pi^\circ & = & 1\,5\,2\,7\,3\,6\,0\,4 \\
2+\pi^\circ & = & 2\,6\,3\,0\,4\,7\,1\,5 \\
3+\pi^\circ & = & 3\,7\,4\,1\,5\,0\,2\,6 \\
4+\pi^\circ & = & 4\,0\,5\,2\,6\,1\,3\,7 \\
5+\pi^\circ & = & 5\,1\,6\,3\,7\,2\,4\,0 \\
6+\pi^\circ & = & 6\,2\,7\,4\,0\,3\,5\,1 \\
7+\pi^\circ & = & 7\,3\,0\,5\,1\,4\,6\,2,
\end{array}$$

which yields $\pi^\circ_\circ = \{\langle 4\ 1\ 6\ 2\ 5\ 7\ 3\rangle, \langle 4\ 1\ 5\ 2\ 7\ 3\ 6\rangle, \langle 4\ 7\ 1\ 5\ 2\ 6\ 3\rangle, \langle 2\ 6\ 3\ 7\ 4\ 1\ 5\rangle, \langle 5\ 2\ 6\ 1\ 3\ 7\ 4\rangle, \langle 5\ 1\ 6\ 3\ 7\ 2\ 4\rangle, \langle 3\ 5\ 1\ 6\ 2\ 7\ 4\rangle, \langle 5\ 1\ 4\ 6\ 2\ 7\ 3\rangle\}$. Hultman [19] proved the following interesting result.

Lemma 7.7 (see [19]). *For all π in S_n and $0 \le m \le n$, every cycle in $G(\pi)$ is mapped onto a cycle in $G(\sigma)$, where σ is the permutation obtained from $m + \pi^\circ$.*

In other words, if $\pi \equiv^\circ_\circ \sigma$, then $\overline{\pi}$ and $\overline{\sigma}$ are conjugate. We show below how one can iterate over the cycle graphs of all elements in $\pi^\circ_\circ$.

Lemma 7.8. *For all π, σ in S_n, if $\sigma^\circ = m + \pi^\circ$, then $\overline{\sigma} = \overline{\pi}^{(0,1,2,\ldots,n)^m}$.*

Proof. By Equation 3.1, we have

$$\begin{aligned}
\overline{\sigma} &= (0,1,2,\ldots,n) \circ (\sigma_0, \sigma_n, \sigma_{n-1}, \ldots, \sigma_1) \\
&= (0,1,2,\ldots,n) \circ (m+\pi_0, m+\pi_n, m+\pi_{n-1}, \ldots, m+\pi_1),
\end{aligned}$$

since by hypothesis $\sigma^\circ = m + \pi^\circ$.

On the other hand, the mapping $(\pi_0, \pi_n, \pi_{n-1}, \ldots, \pi_1) \mapsto (m+\pi_0, m+\pi_n, m+\pi_{n-1}, \ldots, m+\pi_1)$ consists in replacing each element of the cycle with its value plus $m \pmod{n+1}$, which is by definition equivalent to conjugating $(\pi_0, \pi_n, \pi_{n-1}, \ldots, \pi_1)$ by $(0, 1, 2, \ldots, n)^m$. ∎

COROLLARY 7.9. *For all π in S_n, we have $\{\overline{\sigma} \mid \sigma \in \pi^\circ_\circ\} = \{\overline{\pi}^{(0,1,2,\ldots,n)^m} \mid 0 \leq m \leq n\}$.*

Other relations between the cycle structure of π and that of $\overline{\pi}$ can easily be derived from previous work. The following relation allows us to bound the number of odd cycles of $\overline{\pi}$.

THEOREM 7.10 (see [22]). *For all π in S_n, we have $td(\pi) \leq n - c_{odd}(\Gamma(\pi))$.*

The following result is an immediate corollary of Theorems 4.4 and 7.10.

COROLLARY 7.11. *For all π in S_n, we have $2c_{odd}(\Gamma(\pi)) \leq n - 1 + c_{odd}(\Gamma(\overline{\pi}))$.*

Similarly, the following result is an immediate corollary of Theorem 4.2 and of the characterization of exchanges as restricted block-interchanges.

COROLLARY 7.12. *For all π in S_n, we have $2c(\Gamma(\pi)) \leq n - 1 + c(\Gamma(\overline{\pi}))$.*

7.2. Descents of π and cycles of $\overline{\pi}$. Aside from relations between cycle structures, we can also establish relations between pairs of elements of π and cycles of $\overline{\pi}$. An example of such a relation is the fact that the number of adjacencies in $\langle 0\ \pi_1\ \pi_2\ \cdots\ \pi_n\ n+1\rangle$ equals $c_1(\Gamma(\overline{\pi}))$. We will prove that a less obvious relation connects the *descents* of π (defined below) and the cycles of $\overline{\pi}$.

DEFINITION 7.13. *A* descent *in a permutation π is a pair (π_{i-1}, π_i) such that $\pi_i < \pi_{i-1}$.*

For instance, the permutation $\langle 4 \downarrow 1\ 6 \downarrow 2\ 5\ 7 \downarrow 3\rangle$ has three descents, indicated by vertical arrows.

DEFINITION 7.14. *A cycle C in $G(\pi)$* contains *a descent (π_{i-1}, π_i) if (π_i, π_{i-1}) is a black arc of C.*

We now derive bounds on the number of descents contained by cycles in $G(\pi)$.

LEMMA 7.15. *For all π in S_n, every cycle of length $\ell \geq 2$ in $G(\pi)$ contains at most $\ell - 1$ descents and at least one descent of π.*

Proof. For clarity, let us write the vertices of C in the order in which C visits them, starting with the element whose position in π is maximal: we get $C = (\pi_{i_1}, \pi_{j_1}, \pi_{i_2}, \pi_{j_2}, \ldots, \pi_{i_k}, \pi_{j_k})$, where i_1 (resp.,j_k) is the largest (resp.,smallest) position of an element of $\widetilde{\pi}$ appearing in C. We identify here $\widetilde{\pi}_0$ and $\widetilde{\pi}_{n+1} \equiv \widetilde{\pi}_0 \pmod{n+1}$. Recall that (π_{i_x}, π_{j_x}) for $1 \leq x \leq k$ is a black arc of C and that by Definition 2.10 the following relation holds:

$$\pi_{i_x} = \pi_{j_{x-1}} + 1 \text{ for all } 1 \leq x \leq k, \text{ and } \pi_{i_1} = \pi_{j_k} + 1. \tag{7.1}$$

1. For the upper bound, assume on the contrary that C contains ℓ descents; then every black edge of C corresponds to a descent, and we have

$$\pi_{j_x} > \pi_{i_x} \text{ for } 1 \leq x \leq k. \tag{7.2}$$

By alternating between the conditions specified by (7.2) and (7.1), we obtain

$$\begin{aligned}\pi_{j_k} > \pi_{i_k} = \pi_{j_{k-1}} + 1 > \pi_{i_{k-1}} + 1 = \pi_{j_{k-2}} + 2 > \cdots = \pi_{j_1} + k - 1 \\ > \pi_{i_1} + k - 1 = \pi_{j_k} + k,\end{aligned}$$

which is clearly a contradiction.

2. For the lower bound, assume on the contrary that C contains no descent; we have

$$\pi_{j_x} < \pi_{i_x} \text{ for } 1 \leq x \leq k. \tag{7.3}$$

By alternating between the conditions specified by (7.3) and (7.1), we obtain

$$\begin{aligned} \pi_{i_1} - 1 = \pi_{j_k} < \pi_{i_k} = \pi_{j_{k-1}} + 1 < \pi_{i_{k-1}} + 1 = \pi_{j_{k-2}} + 2 < \cdots = \pi_{j_1} + k - 1 \\ < \pi_{i_1} + k - 1. \end{aligned}$$

For the above relations to hold, elements from the set $A = \{\pi_{i_k}, \pi_{i_{k-1}} + 1, \ldots, \pi_{i_2} + k - 2\}$ can only be assigned values from the set $B = \{\pi_{i_1} + 1, \pi_{i_1} + 2, \ldots, \pi_{i_1} + k - 2\}$. However, we have $k - 1 = |A| > |B| = k - 2$, which clearly makes it impossible to obtain a permutation.

Finally, note that π and $\widetilde{\pi}$ can be regarded as equivalent as far as descents are concerned, since $(\widetilde{\pi}_0, \widetilde{\pi}_1)$ and $(\widetilde{\pi}_n, \widetilde{\pi}_{n+1})$ cannot be descents. ◻

The following result is a direct corollary of the above.

PROPOSITION 7.16. *For any* 2*-permutation* π *in* S_n*, the number of descents of* π *is* $(n+1)/2$.

Proof. By definition, $\overline{\pi}$ contains exactly $(n+1)/2$ cycles of length 2, and by Lemma 7.15 each of these cycles contains exactly one descent of π. ◻

8. Conclusions.

We presented a new framework for reformulating any edit distance problem on permutations as a minimum-length factorization problem on a related even permutation, under the implicit assumption that the edit operations are revertible. This approach is based on a new representation of a structure known as the *cycle graph*, which pervades the field of genome rearrangements in several different forms; it previously allowed us to enumerate permutations whose cycle graph decomposes into a given number of alternating cycles [11] and allowed us in this work to recover two previously known results in a simple and unified way. Moreover, we used our approach to derive a new lower bound on the prefix transposition distance that, as we showed both theoretically and experimentally, is a significant improvement over previous results. From that result, we deduced an improved lower bound on the prefix transposition diameter of the symmetric group, whose exact value is still unknown. Finally, we investigated other relations between permutations and their cycle graphs that we hope will prove useful in obtaining new results.

Several interesting questions and leads for future work arise. First, our method provides an automated way of obtaining lower bounds on distances between permutations; is there an analogous way of obtaining *upper* bounds instead? Second, we initiated the study of relations between statistics on a permutation and statistics on the permutation that corresponds to its cycle graph. Can other relations be deduced and used to prove other results, including tighter bounds on the distances of interest? Third, permutations are but one structure for which the cycle graph has been defined. Other structures, such as *signed* permutations, give rise to a more general structure known as the *breakpoint graph*. As mentioned in the introduction, analogues of $f(\cdot)$ have been proposed in the signed setting too; are there other generalizations that could apply to other ways of modeling genomes (e.g., posets, set systems)? Finally, another question is whether Cayley graphs obtained from genome rearrangement operations can yield good interconnection networks. For instance, (signed) reversals generalize the operations that generate the (burnt) pancake network, exchanges generalize the operations that generate the star network, and prefix transpositions generalize the operations that generate the birotator graphs (see [25] for definitions). It seems likely that collaborations between researchers in both fields could be fruitful in investigating this topic.

Acknowledgments. The author wishes to thank Jean-Paul Doignon for suggesting the new definition of $\overline{\pi}$, which greatly simplifies formulas, as well as Galvão and Dias, whose freely available source code [16] allowed the extension of Tables 6.1 and 6.2 by two lines.

REFERENCES

[1] S. B. Akers, B. Krishnamurthy, and D. Harel, *The star graph: An attractive alternative to the n-cube*, in Proceedings of the Fourth International Conference on Parallel Processing (ICPP), Pennsylvania State University Press, 1987, pp. 393–400.
[2] S. B. Akers and B. Krishnamurthy, *A group-theoretic model for symmetric interconnection networks*, IEEE Trans. Comput., 38 (1989), pp. 555–566.
[3] V. Bafna and P. A. Pevzner, *Sorting by transpositions*, SIAM J. Discrete Math., 11 (1998), pp. 224–240.
[4] L. Bulteau, G. Fertin, and I. Rusu, *Pancake flipping is hard*, in Proceedings of the 37th International Symposium on Mathematical Foundations of Computer Science (MFCS), B. Rovan, V. Sassone, and P. Widmayer, eds., Lecture Notes in Comput. Sci. 7464, 2012, Springer, New York, pp. 247–258.
[5] L. Bulteau, G. Fertin, and I. Rusu, *Sorting by transpositions is difficult*, SIAM J. Discrete Math., 26 (2012), pp. 1148–1180.
[6] A. Caprara, *Sorting permutations by reversals and Eulerian cycle decompositions*, SIAM J. Discrete Math., 12 (1999), pp. 91–110.
[7] B. Chitturi and I. H. Sudborough, *Bounding prefix transposition distance for strings and permutations*, in Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS), Los Alamitos, CA, 2008, IEEE Computer Society Press.
[8] B. Chitturi and I. H. Sudborough, *Bounding prefix transposition distance for strings and permutations*, Theoret. Comput. Sci., 421 (2012), pp. 15–24.
[9] D. A. Christie, *Sorting permutations by block-interchanges*, Inform. Process. Lett., 60 (1996), pp. 165–169.
[10] Z. Dias and J. Meidanis, *Sorting by prefix transpositions*, in Proceedings of the Ninth International Symposium on String Processing and Information Retrieval (SPIRE), A. H. F. Laender and A. L. Oliveira, eds., Lecture Notes in Comput. Sci. 2476, 2002, Springer-Verlag, New York, pp. 65–76.
[11] J.-P. Doignon and A. Labarre, *On Hultman numbers*, J. Integer Sequences, 10 (2007).
[12] H. Eriksson, K. Eriksson, J. Karlander, L. Svensson, and J. Wästlund, *Sorting a bridge hand*, Discrete Math., 241 (2001), pp. 289–300.
[13] P. Feijão and J. Meidanis, *Extending the algebraic formalism for genome rearrangements to include linear chromosomes*, in Proceedings of the Seventh Brazilian Symposium on Bioinformatics (BSB), M. C. Pereira de Souto and M. G. Kann, eds., Lecture Notes in Comput. Sci. 7409, 2012, Springer, New York, pp. 13–24.
[14] G. Fertin, A. Labarre, I. Rusu, É. Tannier, and S. Vialette, *Combinatorics of genome rearrangements*, in Computational Molecular Biology, MIT Press, Cambridge, MA, 2009.
[15] V. J. Fortuna, *Distâncias de transposição entre genomas*, M.S. thesis, Universidade Estadual de Campinas, São Paulo, Brazil, 2005.
[16] G. R. Galvão and Z. Dias, *Rearrangement distance database*, http://mirza.ic.unicamp.br:8080/bioinfo/index.jsf (2011).
[17] W. H. Gates and C. H. Papadimitriou, *Bounds for sorting by prefix reversal*, Discrete Math., 27 (1979), pp. 47–57.
[18] E. Györi and G. Turán, *Stack of pancakes*, Stud. Sci. Math. Hungar., 13 (1978), pp. 133–137.
[19] A. Hultman, *Toric Permutations*, M.S. thesis, Department of Mathematics, KTH, Stockholm, Sweden, 1999.
[20] M. R. Jerrum, *The complexity of finding minimum-length generator sequences*, Theoret. Comput. Sci., 36 (1985), pp. 265–289.
[21] J. Kececioglu and D. Sankoff, *Exact and approximation algorithms for sorting by reversals, with application to genome rearrangement*, Algorithmica, 13 (1995), pp. 180–210.

[22] A. Labarre, *New bounds and tractable instances for the transposition distance*, IEEE/ACM Trans. Computat. Biology Bioinformatics, 3 (2006), pp. 380–394.

[23] A. Labarre, *Combinatorial Aspects of Genome Rearrangements and Haplotype Networks*, Ph.D. thesis, Université Libre de Bruxelles, Brussels, Belgium, 2008.

[24] A. Labarre, *Edit distances and factorisations of even permutations*, in Proceedings of the 16th Annual European Symposium on Algorithms (ESA), D. Halperin and K. Mehlhorn, eds., Lecture Notes in Comput. Sci. 5193, 2008, Springer-Verlag, New York, pp. 635–646.

[25] S. Lakshmivarahan, J.-S. Jwo, and S. K. Dhall, *Symmetry in interconnection networks based on Cayley graphs of permutation groups: A survey*, Parallel Comput., 19 (1993), pp. 361–407.

[26] J. Meidanis and Z. Dias, *An alternative algebraic formalism for genome rearrangements*, in Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment and Evolution of Gene Families, Springer, New York, 2000, pp. 213–223.

[27] C. Mira and J. Meidanis, *Sorting by block-interchanges and signed reversals*, in Proceedings of the Fourth International Conference on Information Technology (ITNG), IEEE, 2007, pp. 670–676.

[28] É. Tannier and M.-F. Sagot, *Sorting by reversals in subquadratic time*, in Proceedings of the 15th Annual Symposium on Combinatorial Pattern Matching (CPM), Lecture Notes in Comput. Sci. 3109, Springer-Verlag, New York, 2004, pp. 1–13.

[29] H. Wielandt, *Finite Permutation Groups*, Academic Press, New York, 1964.