SIAM J. DISCRETE MATH.
Vol. 11, No. 2, pp. 224–240, May 1998

SORTING BY TRANSPOSITIONS*

VINEET BAFNA† AND PAVEL A. PEVZNER‡

Abstract. Sequence comparison in computational molecular biology is a powerful tool for deriving evolutionary and functional relationships between genes. However, classical alignment algorithms handle only local mutations (i.e., insertions, deletions, and substitutions of nucleotides) and ignore global rearrangements (i.e., inversions and transpositions of long fragments). As a result, the applications of sequence alignment to analyze highly rearranged genomes (i.e., herpes viruses or plant mitochondrial DNA) are rather limited. The paper addresses the problem of *genome* comparison versus classical *gene* comparison and presents algorithms to analyze rearrangements in genomes evolving by *transpositions*. In the simplest form the problem corresponds to *sorting by transpositions*, i.e., sorting of an array using transpositions of arbitrary fragments. We derive lower bounds on *transposition distance* between permutations and present approximation algorithms for sorting by transpositions. The algorithms also imply a nontrivial upper bound on the *transposition diameter* of the symmetric group. Finally, we formulate two *biological* problems in genome rearrangements and describe the first *algorithmic* steps toward their solution.

Key words. computational molecular biology, genome rearrangements, transpositions, the symmetric group, approximation algorithm

AMS subject classifications. 15A15, 15A09, 15A23

PII. S089548019528280X

1. Introduction.

Studies of molecular evolution of herpes viruses raised many more questions than they answered. Genomes of herpes viruses evolve so rapidly that the extremes of present-day phenotypes may appear quite unrelated. As a result, the similarity between many genes in herpes viruses is so low that it is frequently indistinguishable from the background noise (Karlin, Mocarski, and Schachtel [16]). In particular, there is little or no cross-hybridization between DNAs of Epstein–Barr virus EBV and Herpes simplex virus HSV-1 and until recently there was no unambiguous evidence that these herpes viruses actually had a common evolutionary origin (McGeoch [20]). As a result the classical methods of *sequence comparison* are not very useful for such highly diverged genomes and the ventures into the quagmire of molecular phylogeny of herpes viruses may lead to contradictions, since different genes give rise to different evolutionary trees (Griffin and Boursnell [11]). However, recently a new approach to analyze highly diverged genomes was proposed, based on comparison of *gene orders* versus traditional comparison of *DNA sequences* (Sankoff et al. [24]). Since it is often found that the order of genes is much more conserved than the DNA sequence (Franklin [9]) this approach seems to be a method of choice for many "hard-to-analyze" genomes.

*Received by the editors March 10, 1995; accepted for publication (in revised form) June 2, 1997. A preliminary version of this paper appeared in *Proc.* 6*th Annual ACM-SIAM Symposium on Discrete Algorithms*, San Francisco, CA, SIAM, Philadelphia, PA, 1995, pp. 614–623.

http://www.siam.org/journals/sidma/11-2/28280.html

†Bioinformatics, SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406 (bafnav1@mh.us.sbphrd.com). Most of this author's research was carried out while he was at the Pennsylvania State University, University Park, PA and at DIMACS, Piscataway, NJ.

‡Departments of Mathematics and Computer Science, University of Southern California, DRB 155, Los Angeles, CA 90089-1113 (ppevzner@hto-a.usc.edu). The research of this author was supported by NIH grant 1R01 HG00987-01, by an NSF Young Investigator Award, and by NSF grant CCR-9308567.

Analysis of genomes of EBV and HSV-1 reveals that evolution of these herpes viruses involved a number of *inversions* and *transpositions* of large fragments; in particular, an analogue of the gene UL52-BSLF1 (required for DNA replication) in common herpes virus precursor "jumped" from one location in the genome to another (biologists call this event a transposition). The analysis of such rearrangements at the *genome* level might be more conclusive than the analysis at the *gene* level traditionally used in molecular evolution. However, there are almost no computer science results allowing a biologist to analyze genome rearrangements.

Genomes evolve by inversions and transpositions as well as by more simple operations of deletion, insertion, and duplication of fragments. Inversions seem to be a very common rearrangement; in fact, some genomes (for example, many plant mitochondrial DNA) are believed to evolve almost solely by inversions (Palmer and Herbon [23]). A combinatorial problem of *sorting by reversals* (corresponding to genome rearrangements by inversions) has been studied intensively in recent years, and currently there are two software programs which prove to be useful for analyzing rearrangements in animal (Sankoff et al. [24]) and plant (Bafna and Pevzner [3]) organelle DNA. In 1992 Kececiouglu and Sankoff suggested the first performance guarantee algorithm for sorting by reversal (see [17]). Later Bafna and Pevzner [2] devised a 1.75 performance guarantee algorithm for sorting by reversals and proved Gollan's conjecture on the reversal diameter of the symmetric group. See also Kececioglu and Ravi [18] and Hannenhalli and Pevzner [13] for recent progress on genome rearrangements. An interesting problem related to sorting by reversals is the problem of *sorting by prefix reversals*, also known as the *pancake flipping problem* (Gates and Papadimitriou [10]). Improved bounds for sorting by prefix reversals have been obtained recently (see Cohen and Blum [4]; Heydari and Sudborough [14]).

In a study of herpes viruses, Hannenhalli et al. [12] faced the problem of analyzing an entire spectrum of genome rearrangements—in particular, transpositions. As a first approximation, transpositions in genome rearrangements can be modeled in a straightforward but limited manner by *sorting by transpositions*, described below.

We assume that the order of genes in a genome is represented by a permutation $\pi = \pi_1\pi_2, \dots, \pi_n$. Extend the permutation to include $\pi_0 = 0$ and $\pi_{n+1} = n+1$. For a permutation π, a *transposition* $\rho(i,j,k)$ (defined for all $1 \le i < j \le n+1$ and all $1 \le k \le n+1$ such that $k \notin [i,j]$) "inserts" an interval $[i, j-1]$ of π between π_{k-1} and π_k (Fig. 1.1), i.e., $\rho(i,j,k)$ corresponds to a permutation

$$\begin{pmatrix} 1 \ \dots \ i-1 & \boxed{\mathbf{i\ \ i{+}1\ \dots\ \dots\ \dots\ j{-}2\ j{-}1}} & \boxed{\mathrm{j}\ \dots\ \mathrm{k{-}1}} & k\ \dots\ n \\ 1 \ \dots \ i-1 & \boxed{\mathrm{j}\ \dots\ \mathrm{k{-}1}} & \boxed{\mathbf{i\ i{+}1\ \dots\ \dots\ \dots\ j{-}2\ j{-}1}} & k\ \dots\ n \end{pmatrix}.$$

Clearly, $\pi \cdot \rho(i,j,k)$ has the effect of moving genes $\pi_i, \pi_{i+1}, \dots, \pi_{j-1}$ to a new location in a genome. Also, note that for $i < j < k$, $\rho(i,j,k)$ has the effect of exchanging blocks $\pi_i, \dots, \pi_{j-1}$ and $\pi_j, \dots, \pi_{k-1}$, and $\rho(i,j,k) = \rho(j,k,i)$.

Given permutations π and σ, the *transposition distance problem* is to find a series of transpositions $\rho_1, \rho_2, \dots, \rho_t$ such that $\pi \cdot \rho_1 \cdot \rho_2, \dots, \rho_t = \sigma$ and t is minimum. We call t the *transposition distance* between π and σ. Note that transposition distance between π and σ equals the transposition distance between $\sigma^{-1}\pi$ and the *identity* permutation $\imath$. *Sorting* π *by transpositions* is the problem of finding transposition distance $d(\pi)$ between π and $\imath$. Note that the "biological" definition of transpositions used in this paper is different from the usual "algebraic" definition.

Transpositions generate the *symmetric group* S_n, and we seek a shortest product of *generators* $\rho_1 \cdot \rho_2, \dots, \rho_t$ that equals $\pi \epsilon S_n$. Even and Goldreich [8] show that,

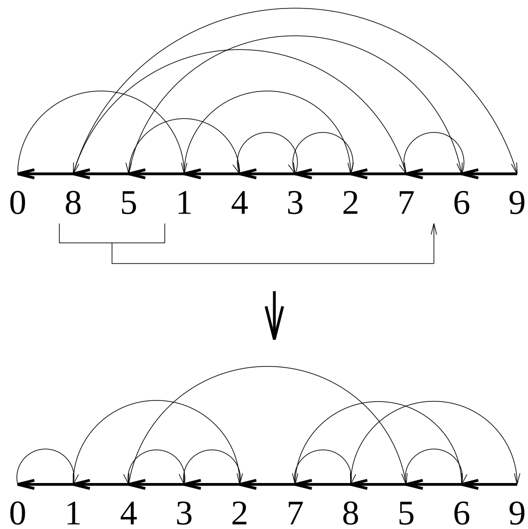

FIG. 1.1. *Transposition* $\rho(1,3,8)$ *on* π *transforms cycle graph* $G(\pi)$ *into* $G(\pi\rho)$.

given a set of generators of a permutation group, determining the shortest product of generators that equals π is NP-hard. In our problem, the generator set is fixed and the complexity status of sorting by transpositions is unknown. The only known polynomially solvable variant of sorting by transpositions is sorting by transpositions $\rho(i, i+1, i+2)$, where the operation is an exchange of adjacent elements. For this problem, polynomial algorithms exist for both linear and circular permutations (Jerrum [15]). Aigner and West [1] found a simple algorithm for sorting by transpositions $\rho(1, 2, i)$ when the operation is reinsertion of the first element.

Sorting by transpositions is a somewhat harder combinatorial problem than the previously studied sorting by reversals; in particular, the transposition diameter of the symmetric group is still unknown. To devise a performance guarantee algorithm for sorting by transpositions, we establish lower bounds for transposition distance based on the notion of the *cycle* graph of a permutation. In section 2 we show that the number of alternating cycles in this edge-colored graph is a bottleneck for sorting by transposition. In section 3 we derive upper bounds for sorting by transposition based on the analysis of *crossing* cycles in the cycle graph. More involved analysis in section 4 provides even better upper bounds in the case where the cycle graph contains *long* cycles. However, this construction breaks for *short* cycles. Somewhat surprisingly, the analysis of *parity* of cycles in the cycle graph provides a compromise and leads to a 1.75 performance guarantee algorithm (section 5). Finally, in section 6 we devise a 1.5 performance guarantee algorithm for sorting by transpositions by exploiting both the structure and parity of crossing cycles in the cycle graph. As an application, we derive a nontrivial upper bound on the transposition diameter of the symmetric group. Algorithms for sorting by reversals and transpositions present the first steps toward the solutions of two open biological problems described in the last section.

2. Lower bounds for sorting by transpositions. For all $0 \le i \le n$, the pair (π_i, π_{i+1}) is a *breakpoint* if $\pi_{i+1} \neq \pi_i + 1$. Observe that the identity permutation is the only permutation with 0 breakpoints, and therefore, sorting a permutation corresponds to decreasing the number of breakpoints. However, this correspondence

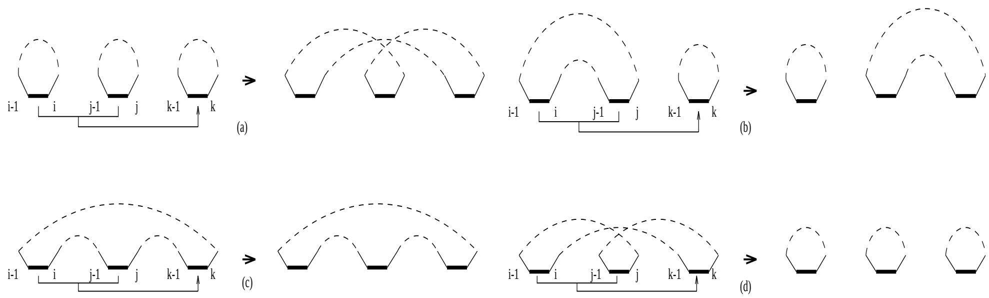

FIG. 2.1. *Transpositions change number of cycles in cycle graphs.*

is not tight in that a permutation with few breakpoints may be more distant from the identity permutation than one with more breakpoints. Also, it is easy to see that a transposition can decrease the number of breakpoints by at most 3, implying a trivial lower bound of $d(\pi) \geq \frac{\#\text{breakpoints}(\pi)}{3}$. However, not all permutations allow transpositions that reduce the number of breakpoints by 3, so the bound is not tight. We introduce the notion of a cycle graph of a permutation and use it to obtain improved lower bounds.

A directed edge-colored *cycle graph* of π, denoted by $G(\pi)$, is the graph with vertex set $\{0, 1, \ldots, n+1\}$ and edge set defined as follows. For all $1 \leq i \leq n+1$, gray edges are directed from $i-1$ to i and black edges from π_i to π_{i-1} (In Fig. 1.1, black edges are shown by thick lines and gray edges are shown by thin lines).

An *alternating cycle* of $G(\pi)$ is a directed cycle in which the edges alternate colors. Observe that for each vertex in $G(\pi)$ every incoming edge is uniquely paired with an outgoing edge of different color. This implies that there is a unique decomposition of the edge set of $G(\pi)$ into alternating cycles. In what follows, we will use *cycle* to refer to an alternating cycle and use *k-cycle* to refer to an alternating cycle of length $2k$. Also, we call a k-cycle *long* if $k > 2$, and *short* otherwise.

There are a total of $2(n+1)$ edges and at most $(n+1)$ cycles in $G(\pi)$ (the identity permutation is the only permutation with $n+1$ cycles). For a permutation π, denote the number of cycles in $G(\pi)$ as $c(\pi)$. Then the sequence of transpositions that sort π must increase the number of cycles from $c(\pi)$ to $n+1$. For a permutation π and a transposition ρ, denote $\Delta c(\rho) = c(\pi\rho) - c(\pi)$ as the change in number of cycles due to transposition ρ.

LEMMA 2.1. $\Delta c(\rho) \in \{2, 0, -2\}$.

Proof. A transposition $\rho(i,j,k)$ involves six vertices of graph $G(\pi)$ $(\pi_{i-1}, \pi_i, \pi_{j-1}, \pi_j, \pi_{k-1}, \pi_k)$ and leads to removing three black edges $((\pi_i, \pi_{i-1}), (\pi_j, \pi_{j-1})$, and $(\pi_k, \pi_{k-1}))$ and adding three new black edges $((\pi_j, \pi_{i-1}), (\pi_i, \pi_{k-1})$, and $(\pi_k, \pi_{j-1}))$.

Three removed edges belong to either three, two, or one cycles in the cycle decomposition of $G(\pi)$. In the case where the removed edges belong to three cycles, $c(\pi\rho) = c(\pi) - 3 + 1$, since these three cycles correspond to one cycle in $G(\pi\rho)$ (Fig. 2.1a). In the case where the removed edges belong to two cycles, $c(\pi\rho) = c(\pi) - 2 + 2$, since these two cycles correspond to two cycles in $G(\pi\rho)$ (Fig. 2.1b). In the case where the removed edges belong to a single cycle C, there are two subcases (Figs. 2.1c and 2.1d). In the subcase shown in Fig. 2.1c, $c(\pi\rho) = c(\pi) - 1 + 1$, since C corresponds to one cycle in $G(\pi\rho)$. In the subcase shown in Fig. 2.1d, $c(\pi\rho) = c(\pi) - 1 + 3$, since C corresponds to three cycles in $G(\pi\rho)$. ∎

Lemma 2.1 immediately gives a lower bound on $d(\pi)$.

THEOREM 2.2. $d(\pi) \geq \frac{n+1-c(\pi)}{2}$.

A cycle in $G(\pi)$ is *odd* if it has an odd number of black edges and *even* otherwise. To establish a better lower bound we analyze odd and even cycles separately. Define $c_{odd}(\pi)$ ($c_{even}(\pi)$) as the number of odd (even) cycles in π. For a permutation π, and a transposition ρ, denote $\Delta c_{odd}(\rho) = c_{odd}(\pi\rho) - c_{odd}(\pi)$ as the change in number of odd cycles due to transposition ρ.

LEMMA 2.3. $\Delta c_{odd}(\rho) \in \{2, 0, -2\}$.

Proof. The proof of Lemma 2.1 implies that the only case when a transposition ρ leads to creating more than two new cycles in $G(\pi\rho)$ is the case presented in Fig. 2.1d. In this case, three cycles are added to $G(\pi)$ and one cycle is removed from $G(\pi)$. If all three added cycles are odd, then the removed cycle is also odd, and $c_{odd}(\pi\rho) = c_{odd}(\pi) - 1 + 3$. Therefore $\Delta c_{odd}(\rho) \leq 2$. This condition, Lemma 2.1, and parity considerations imply $\Delta c_{odd}(\rho) \in \{2, 0, -2\}$. □

As the identity permutation has $n + 1$ odd cycles, Lemma 2.3 implies a better bound.

THEOREM 2.4. $d(\pi) \geq \frac{n+1-c_{odd}(\pi)}{2}$.

Define $d(n) = \max_{\pi \in S_n} d(\pi)$ to be the *transposition diameter* of the symmetric group of order n. Observing that for $\pi = n\ n-1, \ldots, 2\ 1$, $c_{odd}(\pi) = 1$ if n is even and $c_{odd}(\pi) \leq 2$ if n is odd, the transposition diameter of the symmetric group S_n is at least $\lfloor \frac{n}{2} \rfloor$. One can verify that $d(n\ n-1, \ldots, 1) \leq \lfloor \frac{n}{2} \rfloor + 1$ for all n and $d(n) = d(n\ n-1, \ldots, 1) = \lfloor \frac{n}{2} \rfloor + 1$ for $3 \leq n \leq 10$.

3. Upper bounds for sorting by transpositions. For $x \in \{2, 0, -2\}$, define an *x-move* on π as a transposition ρ such that $\Delta c(\rho) = x$. In order to sort faster, we would like to use as many 2-moves as possible. In this section, we study the structure of cycles which allow 2-moves and use that to devise a performance guarantee algorithm for sorting by transpositions.

We number the black edges of the cycle graph $G(\pi)$ from 1 to $n+1$ by assigning label i to a black edge from π_i to π_{i-1}. We say that transposition $\rho(i, j, k)$ *acts* on edges i, j, and k. We also say that a transposition $\rho(i, j, k)$ *acts* on a cycle C if all three black edges i, j, and k belong to C. The proof of Lemma 2.1 implies the following simple observations.

LEMMA 3.1. *If a transposition ρ acts on a cycle and creates more than one new cycle in $G(\pi\rho)$, then ρ is a 2-move.*

LEMMA 3.2. *If a transposition ρ acts on edges belonging to exactly two different cycles, then ρ is a 0-move.*

Figure 2.1 presents two different kinds of cycles—*nonoriented* for which no 2-moves are possible (Fig. 2.1c) and *oriented* for which a 2-move is possible (Fig. 2.1d). Below we give a formal definition of oriented and nonoriented cycles.

Consider a k-cycle C visiting (in order) the black edges $i_1, \ldots, i_k$. A cycle C can be written in k possible ways depending on the choice of the first black edge. Below we assume that the initial black edge i_1 of cycle C starts at its "rightmost" vertex in π, i.e., $i_1 = \max_{1 \leq t \leq k} i_t$.

For all $k > 1$, a cycle $C = (i_1, \ldots, i_k)$ is *nonoriented* if $i_1, \ldots, i_k$ is a decreasing sequence; otherwise C is an *oriented* cycle. We will also use a characterization of nonoriented cycles in the terms of *edge directions*. A gray edge joining $\pi_t = i - 1$ with $\pi_s = i$ in $G(\pi)$ is directed *left* if $t > s$ and is directed *right* otherwise. Clearly, a cycle $C = (i_1, \ldots, i_k)$ is nonoriented iff $k > 1$ and C has exactly one right edge (a gray edge between black edges i_k and i_1).

LEMMA 3.3. *If C is an oriented cycle, then there exists a 2-move acting on C. If C is a nonoriented cycle, then there exist no 2-moves acting on C.*

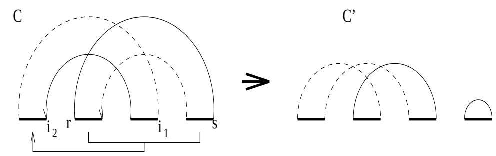

FIG. 3.1. *A* 0*-move creating an oriented cycle.*

Proof. Let $C = (i_1, \dots, i_k)$ be an oriented cycle and let $3 \le t \le k$ be an index such that $i_t > i_{t-1}$. Consider a transposition $\rho(i_{t-1}, i_t, i_1)$ acting on C. This transposition creates a 1-cycle (on vertices $\pi_{i_{t-1}-1}$ and π_{i_t}) and some other cycles. Therefore, by Lemma 3.1, ρ is a 2-move. ∎

Lemmas 3.2 and 3.3 imply the following theorem.

THEOREM 3.4. *For an arbitrary (unsorted) permutation π, there exists either a* 2*-move or a* 0*-move followed by a* 2*-move.*

Proof. If $G(\pi)$ has an oriented cycle then, by Lemma 3.3, a 2-move is possible. Otherwise, consider a nonoriented cycle $C = (i_1, \dots, i_k)$ and let r be a position of the maximal element of π in the interval $[i_2, i_1 - 1]$. Let s be a position of $\pi_r + 1$ in π (Fig. 3.1). Clearly $s \notin [i_2, i_1]$. Without loss of generality, assume that $s > i_1$, and consider a transposition $\rho(r+1, s, i_2)$ (Fig. 3.1). The transposition ρ acts on edges of two different cycles; therefore by Lemma 3.2 ρ is a 0-move. Since ρ changes the direction of the left edge (π_{i_1-1}, π_{i_2}), and does not change direction of the right edge (π_{i_k-1}, π_{i_1}), the cycle C' containing these edges in $G(\pi\rho)$ has at least two right edges. Therefore C' is an oriented cycle allowing a 2-move (Lemma 3.3). ∎

Theorem 3.4 provides an increase of $c(\pi)$ by at least 2 in two consecutive moves and implies the following upper bound for sorting by transpositions.

THEOREM 3.5. *Any permutation π can be sorted in $n + 1 - c(\pi)$ transpositions.*

Theorems 2.2 and 3.5 imply an approximation algorithm for sorting by transpositions with performance guarantee 2. In the following sections, we give a better upper bound by disallowing −2-moves and forcing at least two consecutive 2-moves between any two 0-moves. In our approximation algorithm, we will use only 0- and 2-moves, although we do not have proof that an optimal sequence of transpositions exists which does not use −2-moves.

4. Crossing cycles. Theorem 3.4 shows that the number of 2-moves can be made greater than or equal to the number of 0-moves. In order to improve the performance ratio for sorting by transposition, we need to further increase the number of 2-moves. Theorem 4.7 provides the first step toward such an improvement, but first we need to prove a series of technical lemmas.

Consider a triple of black edges x, y, z belonging to the same cycle C in $G(\pi)$. C induces a cyclic order on x, y, z, and among three possible representations of this order we choose the one starting from the rightmost black edge $\max\{x, y, z\}$ as the canonical representation for a triple (x, y, z). A triple (in a canonical order) is called *nonoriented* if $x > y > z$ and *oriented* otherwise. For example, a triple (k, j, i) in Fig. 2.1c is nonoriented while triple (k, i, j) in Fig. 2.1d is oriented. All triples of a nonoriented cycle are nonoriented. On the other hand, every oriented cycle has at least one oriented triple.

Ordered sequences of integers $\{v_1 < \cdots < v_k\}$ and $\{w_1 < \cdots < w_k\}$ are *interleaving* if either $v_1 < w_1 < v_2 < w_2 < \cdots < v_k < w_k$ or $w_1 < v_1 < w_2 < v_2 \cdots < w_k < v_k$. Sets of integers V and W are interleaving if orderings of V and W are interleaving.

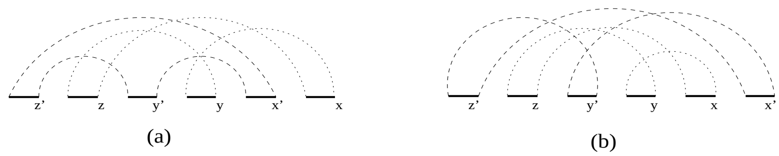

FIG. 4.1. *Crossing and noninterfering cycles.*

Let (x, y, z) be a nonoriented triple, i.e., $x > y > z$. A transposition $\rho(i, j, k)$ is a *shuffling* transposition with respect to a triple (x, y, z) if the sets $\{i, j, k\}$ and $\{x, y, z\}$ interleave.

LEMMA 4.1. *Let (x, y, z) be a triple in a cycle C, and let $i, j, k \notin C$ be black edges in $G(\pi)$. Then $\rho(i, j, k)$ changes the orientation of triple (x, y, z) (i.e., it transforms oriented triple into non-oriented and vice versa) iff ρ is a shuffling transposition for (x, y, z).*

LEMMA 4.2. *If C is nonoriented, then for all triples $(x, y, z) \in C$, transposition $\rho(z, y, x) = \rho(y, x, z)$ transforms C into a nonoriented cycle in $G(\pi\rho)$.*

We will also need the following lemma specifying some 2-moves acting on oriented cycles.

LEMMA 4.3. *If (x, y, z) is an oriented triple, then $\rho(y, z, x) = \rho(z, x, y)$ is a 2-move.*

Cycles C and C' are *crossing* if there exists an oriented triple in C and a non-oriented triple in C' that are interleaving (Fig. 4.1a). Cycles C and C' are *non-interfering* if there exist oriented triples in C and C' that are not interleaving (Fig. 4.1b).

LEMMA 4.4. *If permutation π has crossing or noninterfering cycles, then there exist two consecutive 2-moves in π.*

Proof. If cycles C and C' in $G(\pi)$ are crossing, there exist an oriented triple $(x, z, y) \in C$ and a nonoriented triple $(x', y', z') \in C'$ which are interleaving (Fig. 4.1a). By Lemma 4.3, a transposition $\rho(z, y, x)$ defines a 2-move on C. On the other hand, since (x, y, z) and (x', y', z') are interleaving, $\rho(z, y, x)$ is a shuffling transposition with respect to (x', y', z'). Thus, by Lemma 4.1 ρ transforms C' into an oriented cycle in $G(\pi\rho)$ and by Lemma 3.3 provides a second 2-move.

Alternatively, if C and C' are noninterfering, then there exist oriented triples $(x, z, y) \in C$ and $(x', z', y') \in C'$ which are noninterleaving (Fig. 4.1b). By Lemma 4.3, a transposition $\rho(z, y, x)$ defines a 2-move on C. Furthermore, (x', z', y') remains an oriented triple (Lemma 4.1) of C' in $G(\pi\rho)$, which provides a second 2-move. ☐

We say that a transposition *acts* on two cycles C and C' in $G(\pi)$ if it acts on black edges of both C and C'. To prove Theorem 4.7 below, we will need the following observation about transpositions acting on two cycles.

LEMMA 4.5. *Let C be a cycle containing black edges x and y and let D be a cycle containing black edges x' and y'. Let ρ be a transposition acting on three of four black edges x, y, x', y'.*

- *If $\{x, y\}$ does not interleave with $\{x', y'\}$, then ρ creates a cycle with a non-oriented triple.*
- *If $\{x, y\}$ interleaves with $\{x', y'\}$, then ρ creates a cycle with an oriented triple.*

Proof. See Fig. 4.2. All other cases are symmetric. ☐

We say that cycle $C = (i_1, \ldots, i_k)$ *spans* cycle $D = (j_1, \ldots, j_l)$, if $i_k < j_l < j_1 < i_1$. The following lemma illustrates an important property of nonoriented cycles.

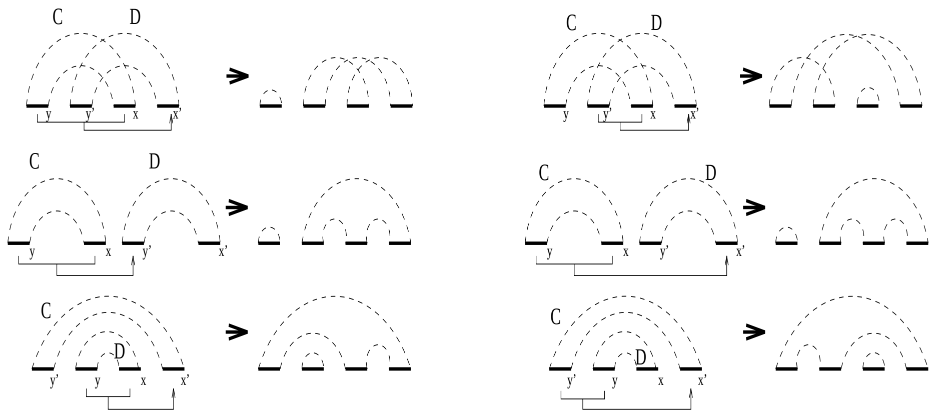

FIG. 4.2. *Transpositions acting on two cycles.*

LEMMA 4.6. *For every nonoriented cycle $C = (\ldots a, \ldots, b \ldots)$, with arbitrary edges a, b, there exists a cycle $D(\ldots, c \ldots, d \ldots)$ such that (a, b) and (c, d) interleave.*

Proof. Let $\pi_c = \max_{i \in [b, a-1]} \pi_i$ and $\pi_d = \pi_c + 1$. Choice of c implies that $d \notin [b, a-1]$, as C is nonoriented $d \neq a$, implying that $d \notin [b, a]$. Therefore, (c, d) and (a, b) interleave. □

THEOREM 4.7. *If there exists a long cycle in $G(\pi)$, then either a 2-move or a 0-move followed by two consecutive 2-moves is possible in π.*

Proof. If $G(\pi)$ has an oriented cycle, then by Lemma 3.3 a 2-move is possible. Also, if there exist nonoriented long cycles C and D with interleaving triples $(r, s, t) \in C$ and $(x, y, z) \in D$, then a 0-move ρ acting on edges z, y, x is a shuffling transposition for C. By Lemma 4.1, ρ transforms C into an oriented cycle C'. By Lemma 4.2 ρ transforms D into a nonoriented cycle D'. It is easy to see that C' and D' are crossing; therefore, by Lemma 4.4 there exist two consecutive 2-moves in $G(\pi\rho)$.

Therefore, assume that no two cycles have interleaving triples. Pick a nonoriented long cycle $C = (i_1, \ldots, i_k)$, such that C is not spanned by any long cycle. Find a cycle $D = (x, \ldots, c, \ldots, d, \ldots, y)$ such that the pairs (c, d) and (i_1, i_k) interleave (Lemma 4.6). Note that if $y < i_k$, then $x < i_1$; otherwise D would span C. On the other hand, if $y > i_k$, then $x > i_1$; otherwise (c, d) and (i_1, i_k) would not interleave. Therefore, either $y < i_k < x < i_1$ or $i_k < y < i_1 < x$. Without loss of generality, we assume the latter. Let s be the rightmost edge in C to the left of y, i.e., $s = \max_{i \in C, i < y} i$. Two cases arise.

$s > i_k$: Find cycle $E = (v, \ldots, c, \ldots, d, \ldots, u)$ such that the pairs (c, d) and (i_k, s) interleave (Lemma 4.6). If $u < i_k$, then $v < s$ because, otherwise, E either spans C ($v > i_1$) or has an interleaving triple with $(i_k, s, i_1) \in C$ ($s < v < i_1$). If $u > i_k$ (Fig. 4.3a), then four cases arise depending on v lying in one of the intervals $[s, y], [y, i_1], [i_1, x]$ or $[x, n+1]$ (Fig. 4.3b-e). The transpositions $\rho(x, y, v)$ in Fig. 4.3a and $\rho(x, y, u)$ in Figs. 4.3b-4.3e are shuffling w.r.t. the triple (i_1, s, i_k) of C, and by Lemma 4.1 transform C into an oriented cycle C' in $G(\pi\rho)$. ρ also transforms D and E into D' and a 1-cycle in $G(\pi\rho)$. From Lemma 4.5, D' is oriented in Fig. 4.3a. In the remaining cases, D' is oriented when $v \in [y, x-1]$ and nonoriented otherwise (Lemma 4.5). Observe that in the first case C' and D' are crossing (Figs. 4.3b, 4.3e); otherwise they are noninterfering (Figs. 4.3c, 4.3d). In either case, two 2-moves are possible in $G(\pi\rho)$(Lemma 4.4).

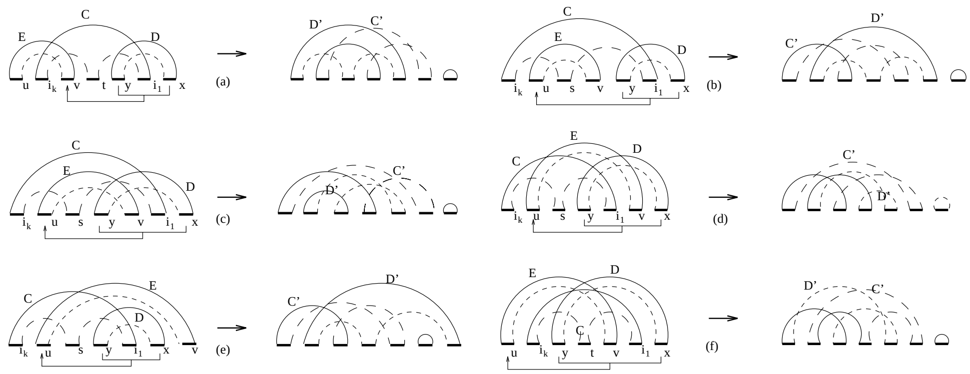

FIG. 4.3. 0-*move leading to two* 2-*moves.*

> **Algorithm** *Tsort(π)*
>
> 1. While $G(\pi)$ has a long cycle, perform either a 2-move or a $0, 2, 2$-move (Theorem 4.7).
> 2. If $G(\pi)$ has only short cycles, perform a 0-move followed by a 2-move (Theorem 3.4).

FIG. 5.1. *Algorithm* Tsort *for sorting by transpositions.*

$s = i_k$: Let t be the leftmost black edge in C to the right of y, i.e., $t = \min_{i \in C, i > y} i$. As C is a long cycle, $t < i_1$. Find $E = (v, \dots, c, \dots, d, \dots, u)$ such that the pairs (c, d) and (t, i_1) interleave (Lemma 4.6). Cycle E is different from cycle D because, otherwise, E and C would have interleaving triples. If $v > i_1$, then $u > t$ because, otherwise, E either spans C $(u < i_1)$ or has an interleaving triple with $(i_1, t, i_k) \in C$ $(i_1 < u < t)$. This case is similar to the cases shown in Figs. 4.3d, 4.3e. If $v < i_1$, then three cases arise depending on which of the intervals $[0, i_k], [i_k, y]$, or $[y, t]$ contains u. The first of these cases is shown in Fig. 4.3f, while the other two are symmetric to cases in Fig. 4.3c and 4.3e, respectively. In Fig. 4.3f, the transposition $\rho(x, y, u)$ transforms C into a nonoriented cycle C' (Lemma 4.1), and transforms cycles D, E into an oriented cycle D' and a 1-cycle in $G(\pi\rho)$ (Lemma 4.5). Further, C' and D' are crossing, and therefore two 2-moves are possible in $G(\pi\rho)$. ☐

5. Mixing odd and even cycles. Theorem 4.7 guarantees creating at least four cycles in three moves, thus providing $\Delta c(\rho) = \frac{4}{3}$ on average, which is better than $\Delta c(\rho) = 1$, given by Theorem 3.4. However, it can be applied only when $G(\pi)$ has long cycles. In case $G(\pi)$ only has short cycles, the best we can guarantee is a 0-move followed by a 2-move creating four 1-cycles from two 2-cycles (Theorem 3.4). Theorems 3.4 and 4.7 motivate the algorithm *Tsort* (Fig. 5.1).

Does *Tsort* achieve a performance ratio of better than 2? Unfortunately, in the case that $G(\pi)$ has only short cycles, the 0-move followed by a 2-move provides only $\Delta c(\rho) = \frac{4-2}{2} = 1$ on average. However, for these two moves, $\Delta c_{odd}(\rho) = \frac{4-0}{2} = 2$, thus achieving a maximal rate of creating odd cycles from the perspective of Theorem 2.4. On the other hand, Theorem 4.7 does not guarantee yet that $\Delta c_{odd}(\rho) = 2$ for every 2-move. Therefore, if we use either the number of cycles or the number of odd cycles as our objective function, we cannot guarantee a performance ratio better than 2. Somewhat surprisingly, we show that a *mixed* objective function which gives different weights to odd and even cycles leads to an improved performance guarantee.

THEOREM 5.1. Tsort *provides a performance guarantee of* 1.75 *for sorting by transpositions.*

Proof. For arbitrary $x \geq 1$, define the objective function $f(\pi) = xc_{odd}(\pi) + c_{even}(\pi)$, where $c_{odd}(\pi)$ and $c_{even}(\pi)$ are the number of odd and even cycles in $G(\pi)$, respectively. Clearly, for this range of x, $f(\pi)$ is uniquely maximized by the identity permutation, and sorting a permutation corresponds to maximizing f. Observe that the maximum gain any transposition ρ can achieve is $\Delta f(\rho) = f(\pi\rho) - f(\pi) = 2x$. We now evaluate the maximum Δf guaranteed by Theorems 3.4 and 4.7.

In the case that $G(\pi)$ only has short cycles, Theorem 3.4 guarantees that in two moves, four 1-cycles are created from two 2-cycles, implying a gain of $4x - 2$ over two moves, or an average gain of $2x - 1$ in one transposition. In any 2-move, two new cycles are created and, in the worst case (if both are even) we can still guarantee a gain of 2. By construction, a 0-move in Theorem 4.7 either creates a 1-cycle or does not change the number of black edges in any cycle. Therefore $\Delta f \geq 0$ for any 0-move. Moreover, Theorem 4.7 guarantees that any such 0-move is followed by two 2-moves, implying an average gain of $\frac{4}{3}$. It follows that $\Delta f \geq \min\{\frac{4}{3}, 2x - 1\}$ on the average. Comparing the best possible gain of $2x$ against the gain provided by *Tsort*, we get a performance guarantee of

$$\frac{2x}{\min\{\frac{4}{3}, 2x - 1\}}.$$

The best performance is achieved for $x = \frac{7}{6}$, resulting in the approximation ratio 1.75. ◻

6. A 1.5 approximation algorithm for sorting by transposition. In order to improve performance still further, we need to strengthen Theorem 4.7. Note that Theorem 4.7 only guarantees an increase in the number of cycles. However, the identity permutation has $n + 1$ cycles, all of length one, indicating that we need to increase the number of odd cycles. By choosing appropriate 2-moves, we shall ensure that the number of odd cycles increases by at least two in every 2-move.

We call a transposition ρ *valid* if $\Delta c(\rho) = \Delta c_{odd}(\rho)$. For a cycle C containing edges i and j, define $d(i, j)$ as the number of black edges between vertices π_i and π_j in C (in particular, $d(i, j) = 1$ for consecutive black edges i and j).

LEMMA 6.1. *If there exists an oriented cycle in* $G(\pi)$*, then either a valid* 2*-move or a valid* 0*-move followed by two consecutive valid* 2*-moves is possible in* π*.*

Proof. Suppose there is no valid 2-move in π. For an arbitrary oriented cycle C in $G(\pi)$, consider the following set S of oriented triples of C such that the distance between the first and second elements of the triple is odd:

$$S = \{(x, y, z) : \ x, y, z \in C \text{ and } d(x, y) \text{ is odd}\}.$$

The observation that every oriented cycle C has an oriented triple (x, y, z) such that x and y are the consecutive black edges in C implies that S is nonempty. Let (x, y, z) be a triple in S with maximal x.

A transposition ρ acting on edges y, z, and x transforms C into three cycles C_1, C_2, and C_3 consisting of $d(x, y), d(y, z)$, and $d(z, x)$ black edges. As $(x, y, z) \in S$, cycle C_1 is odd. If either C_2 or C_3 is odd, then $\Delta c_{odd}(\rho) = 2$ and ρ is a valid 2-move, contradicting the assumption that there are no valid 2-moves in π. Therefore both

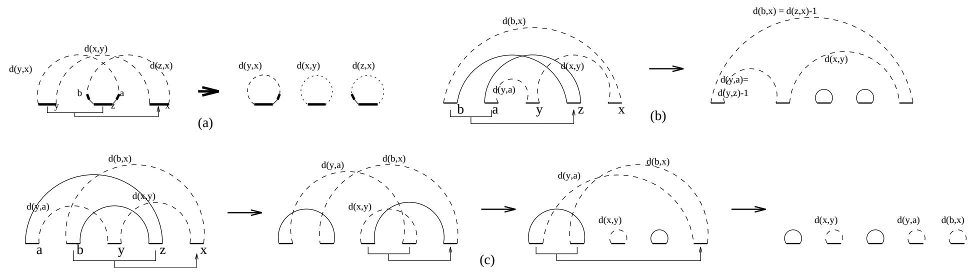

FIG. 6.1. *Valid* 2*-moves and* $0, 2, 2$*-moves on an oriented cycle.*

$d(y, z)$ and $d(z, x)$ are even. As both $d(y, z)$ and $d(z, x)$ are even, the fragments of C from y to z and from z to x contain at least two edges. Let a be a black edge preceding z in C and b be a black edge following z in C (Fig. 6.1a).

If $y < a < x$, then transposition ρ acting on edges y, a, and x creates cycles of length $d(y, z) - 1$ and $d(x, y)$. Both these numbers are odd and, therefore, ρ is a valid 2-move, thus contradicting the assumption. Therefore $a \notin [y, x]$. Symmetric arguments demonstrate that $b \notin [y, x]$.

If $a > x$, then (a, z, x) is an oriented triple with odd $d(a, z) = 1$, thus contradicting the choice of (x, y, z). Therefore $a < y$. If $b > x$, then (b, a, z) is an oriented triple with odd $d(b, a) = d(b, x) + d(x, y) + d(y, a) = (d(z, x) - 1) + d(x, y) + (d(y, z) - 1)$, thus contradicting the choice of (x, y, z). Therefore $a, b < y$.

The situations described by conditions $b < a$ and $a < b$ are presented in Figs. 6.1b and 6.1c. If $b < a$, then $\rho(b, a, z)$ is a valid 2-move (Fig. 6.1b). If $a < b$, then there exist 2-moves but no valid 2-moves in π. However, there exists a valid 0-move followed by two consecutive valid 2-moves (Fig. 6.1c). □

Fig. 6.1c presents an example of an oriented cycle which does not allow valid 2-moves. This cycle has a complicated "self-interleaving" structure and, in the following, we try to avoid creating such cycles. In order to achieve this goal, we define *strongly oriented* cycles, which have the simplest "self-interleaving" structure among all oriented cycles.

Let $C = (i_1, \ldots, i_k)$ be a cycle in $G(\pi)$ and let $C^* = (i_1 = j_1 > \cdots > j_k)$ be a sequence of black edges of C in decreasing order. Sequences C and C^* coincide for a nonoriented cycle and are different otherwise. Define *strongly oriented* cycles as oriented cycles for which C^* can be transformed into C by a single transposition, i.e., C can be partitioned into strips $C_1 = (i_1, \ldots, i_a)$, $C_2 = (i_{a+1}, \ldots, i_b)$, $C_3 = (i_{b+1}, \ldots, i_c)$, and $C_4 = (i_{c+1}, \ldots, i_k)$ such that $C = C_1C_2C_3C_4$ and $C^* = C_1C_3C_2C_4$ (C_4 might be empty). For example, Fig. 6.1b gives an example of a strongly oriented cycle, as $C = xyabz$ is transformed into $C^* = xzyab$ by a single transposition. Clearly, every strongly oriented cycle has exactly two right edges. On the other hand, not every oriented cycle with two right edges is strongly oriented (Fig. 6.1c).

LEMMA 6.2. *A strongly oriented cycle allows a valid* 2*-move.*

Proof. Depending on whether or not C_4 is empty, there are two kinds of cycles, as shown in Fig. 6.2, with $left + mid + right$ black edges (in Fig. 6.2c, $mid = mid' + mid''$). Dashed lines in the figure represent alternating paths of zero or more edges. In the following, we shall abuse notation by referring to both the sets of edges and their numbers as $left, mid$, and $right$.

In Fig. 6.2a, consider transpositions of the form $\rho(i, j, k)$, where i is the leftmost mid edge, j is the rightmost $right$ edge, and k is a $left$ edge. As all such triples (i, j, k) are oriented; $\rho(i, j, k)$ is a 2-move.

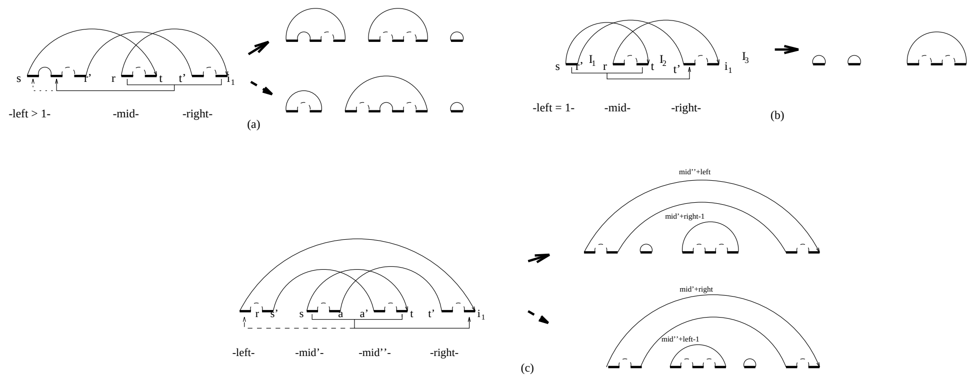

FIG. 6.2. *Strongly oriented cycles:* (a), (b) *First kind.* (c) *Second kind.*

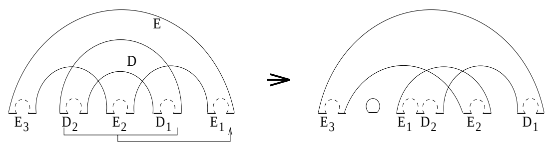

FIG. 6.3. *Transforming two nonoriented cycles into a strongly oriented cycle.*

Figure 6.2a corresponds to the case $left > 1$ and presents two such transpositions, say, $\rho_1(i, j, k_1)$ and $\rho_2(i, j, k_2)$, in which k_1 and k_2 are the two leftmost *left* edges. Both ρ_1 and ρ_2 are 2-moves and create three cycles. One of these cycles is a 1-cycle. If *left* > 1, then an appropriate choice of either ρ_1 or ρ_2 provides at least one more odd cycle, thus indicating that the chosen transposition is a valid 2-move. If *left* $= 1$, then the transposition ρ shown in Fig. 6.2b creates at least two 1-cycles, thus indicating that ρ is a valid 2-move.

In Fig. 6.2c, the transposition ρ inserting a "middle interval" into the leftmost edge creates cycles of length $1, mid'' + left - 1$ and $mid' + right$. On the other hand, a transposition ϱ inserting a middle interval into the rightmost edge creates cycles of length $1, mid'' + left$ and $mid' + right - 1$. Therefore, either ρ or ϱ creates at least two odd cycles, thus ensuring a valid 2-move in π. □

Next, we present two lemmas which show how strongly oriented cycles arise from nonoriented cycles.

LEMMA 6.3. *If ρ is a shuffling transposition on a nonoriented cycle C, then ρ transforms C into a strongly oriented cycle in $G(\pi\rho)$.*

Proof. The proof follows from the definition. □

LEMMA 6.4. *Let $D(x, \ldots, y)$ and $E(x', \ldots, y')$ be two nonoriented cycles in $G(\pi)$ with no interleaving triples, and let ρ be a transposition acting on three of four black edges x, y, x', y'. Then ρ creates a strongly oriented cycle iff D and E have interleaving pairs of edges.*

Proof. Figure 6.3 presents cycles D and E with interleaving pairs of edges, but no interleaving triple. Assume w.l.o.g that the edges of D partition E into three strips $E = E_1E_2E_3$ (E_3 is possibly empty), while the edges of E partition the edges of D into two $D = D_1D_2$. The transposition ρ transforms D and E into a 1-cycle and a cycle F visiting (in order) edges $D_1D_2E_1E_2E_3$. On the other hand, $F^* = D_1E_2D_2E_1E_3$ which can clearly be transformed into F by a transposition.

If D and E have no interleaving pairs of edges, then it is easy to verify that ρ transforms D and E into a 1-cycle and a nonoriented cycle F. □

Every strongly oriented cycle has exactly two right edges, one of which is of the form (r, i_1). Label the other as (s, t). For strongly oriented cycles of the first kind (Fig. 6.2a), define

$$r' = \max_{i \in left} i \quad \text{and} \quad t' = \min_{i \in right} i,$$

and consider three intervals $I_1(C) = [r', r], I_2(C) = [t, t']$, and $I_3 = [0, s] \cup [i_1, n+1]$. For strongly oriented cycles of the second kind (Fig. 6.2c), define

$$s' = \max_{i \in left} i \ , \ t' = \min_{i \in right} i \ , \ a = \max_{i \in mid'} i \text{ and } a' = \min_{i \in mid''} i,$$

and consider intervals $I_1(C) = [s', s], I_2(C) = [t, t']$, and $I_3(C) = [a, a']$.

A strongly oriented cycle C and a nonoriented cycle $C' = (i_1, \ldots, i_k)$ are *strongly crossing* if there exists a black edge x in C' such that each of the sets $I_1(C), I_2(C)$, and $I_3(C)$ contains exactly one element of the triple (i_1, x, i_k). Note that 2-moves for C described in the proof of Lemma 6.2 form shuffling transpositions w.r.t. (i_1, x, i_k). This observation and Lemma 6.2 imply the following.

LEMMA 6.5. *If $G(\pi)$ has strongly crossing cycles, then there exist two consecutive valid 2-moves in $G(\pi)$.*

Next, we modify the concept of "noninterfering" cycles after which we shall have all the tools needed to strengthen Theorem 4.7. A transposition ρ is *safe*, with respect to a strongly oriented cycle $C \in G(\pi)$, if it transforms C into a strongly oriented cycle in $G(\pi\rho)$. The following lemma gives a sufficient condition for a transposition to be safe.

LEMMA 6.6. *Let C be a strongly oriented cycle, and let $(x, y, z) \notin C$ be a triple such that no edge of C lies in the region between x and y. Then, a transposition acting on (x, y, z) is safe w.r.t. C.*

Let cycles C and C' be strongly oriented. C is *strongly noninterfering* w.r.t. C' if it has a right edge (a, b) such that no black edge of C' lies in the interval $[a, b]$.

LEMMA 6.7. *If $G(\pi)$ has strongly noninterfering cycles, then there exist two consecutive valid 2-moves in $G(\pi)$.*

Proof. Let C be strongly noninterfering w.r.t. C'. Consider a valid 2-move $\rho(x, y, z)$ on C described in the proof of Lemma 6.2. Observe that one of the right edges in C is of the form (x, y) and therefore includes the region $[x, y]$, and the other right edge includes the interval $[y, z]$. Therefore, if C is strongly noninterfering w.r.t. C', then either no black edge of C' lies in $[x, y]$ or no black edge of C' lies in $[y, z]$. In either case, $\rho(x, y, z)$ is safe w.r.t. C' (Lemma 6.6). This implies that a valid 2-move on C' follows a valid 2-move on C'. □

Finally, we can prove a stronger version of Theorem 4.7.

THEOREM 6.8. *If there exists a long cycle in $G(\pi)$, then either a valid 2-move or a valid 0-move followed by two consecutive valid 2-moves is possible in π.*

Proof. We mimic the proof of Theorem 4.7, ensuring that all moves are valid ones.

If $G(\pi)$ has an oriented cycle, then from Lemma 6.1, a valid 2-move or a valid 0-move followed by two valid 2-moves is always possible.

Next, consider the case when $G(\pi)$ has nonoriented cycles C and D with interleaving triples $(r, s, t) \in C$ and $(x, y, z) \in D$. Then, $\rho(x, y, z)$ transforms C into a strongly oriented cycle C' in $G(\pi\rho)$ (Lemma 6.2) and transforms D into a nonoriented cycle D' in $G(\pi\rho)$ (Lemma 4.1). Further observe that each of the intervals $I_1(C'), I_2(C')$,

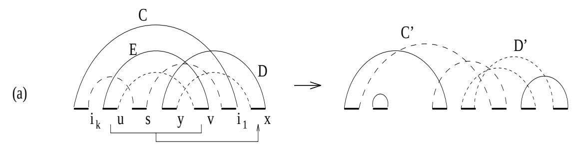

FIG. 6.4. *Transforming $C, D,$ and E into strongly noninterfering cycles.*

and $I_3(C')$ contains exactly one element of a (nonoriented) triple in D'. Therefore, C' and D' are strongly crossing, and from Lemma 6.5, two valid 2-moves are possible in $G(\pi\rho)$.

Therefore, we can assume that $G(\pi)$ has no oriented cycles or cycles with interleaving triples. The proof of theorem holds and we consider them in the following case by case fashion:

Fig. 4.3a. The valid 0-move $\rho(y, x, v)$ transforms D and E into a nonoriented cycle D' (Lemma 6.4) and transforms C into a strongly oriented cycle C' (Lemma 6.2) in $G(\pi\rho)$. Further observe that vertices π_x, π_v, π_y all belong to D' and $\pi_y \in I_1(C')$, $\pi_v \in I_2(C')$, $\pi_x \in I_3(C')$, thereby implying that C' and D' are strongly crossing. From Lemma 6.5, two valid 2-moves are possible in $G(\pi\rho)$.

Fig. 4.3b. The valid 0-move $\rho(y, x, u)$ transforms D and E into a nonoriented cycle D' (Lemma 6.4), and transforms C into a strongly oriented cycle C' (Lemma 6.2) in $G(\pi\rho)$. Observe that $\pi_y \in I_1(C')$ and $\pi_u \in I_2(C')$. Moreover, the choice of s as the rightmost edge to the left of y ensures that there is no edge of C between v and y, and therefore $\pi_v \in I_3(C')$. As vertices π_x, π_v, π_y all belong to D', cycles C' and D' are strongly crossing. From Lemma 6.5, two valid 2-moves are possible in $G(\pi\rho)$.

Fig. 4.3c. In this case, we consider the valid 0-move $\rho(u, v, x)$ (Fig. 6.4) instead of $\rho(y, x, u)$. ρ transforms C into a strongly oriented cycle C' (Lemma 6.2), and transforms D and E into strongly oriented cycle D' (as D and E have no interleaving triples, Lemma 6.4 applies). Define a as the rightmost edge in D to the left of i_1, i.e., $a = \max_{i \in D, i < i_1} i$, and define b as the leftmost edge in C to the right of y, i.e., $b = \min_{i \in C, i > y} i$. Note that $a < b$ because, otherwise, $(i_k, b, i_1) \in C$ and $(y, a, x) \in D$ are interleaving triples. If $b > v$, then there is no edge of C in the interval $[y, v]$, and it follows that C' has no black edge in the region covered by the right edge $(\pi_{y-1}, \pi_x) \in D'$. Therefore D' is strongly noninterfering w.r.t. C'. If $a < b < v$, then there is no black edge of D in the interval $[v, i_1]$, and correspondingly, D' has no black edge in the region covered by the right edge $(\pi_{i_k - 1}, \pi_{i_1}) \in C'$. Therefore, C' is strongly noninterfering w.r.t. D'. In either case, Lemma 6.7 implies that two valid 2-moves are possible in $G(\pi\rho)$.

Fig. 4.3d. The valid 0-move $\rho(x, y, u)$ transforms D and E into strongly oriented cycle D' (as D and E have no interleaving triples, Lemma 6.4 applies) and also transforms C into strongly oriented cycle C' (Lemma 6.2). Let a be the rightmost edge of E to the left of s, and let b the leftmost edge of E to the right of i_1. Note that C has no edge in the region between edges $b \in E$ and $x \in D$ in $G(\pi)$ as $i_1 < b < x$. Also, C has no edge e in the region $[u, a]$ because, otherwise, $(u, a, v) \in E$ would interleave $(e, s, i_1) \in C$. Correspondingly in $G(\pi\rho)$, C' does not have any black edge in the region covered by the right edge $(\pi_{b-1}, \pi_a) \in D'$, implying that D' is strongly noninterfering w.r.t. C'. From Lemma 6.7, two consecutive valid 2-moves are possible in $G(\pi\rho)$.

Algorithm *TransSort(π)*

1. While $G(\pi)$ has a long cycle, perform a valid 2-move or a valid $0, 2, 2$-move (Theorem 6.8).
2. If $G(\pi)$ has only short cycles, perform a good 0-move followed by a valid 2-move (Theorem 6.9).

FIG. 6.5. *Algorithm* TransSort *for sorting by transpositions.*

Fig. 4.3e. The valid 0-move $\rho(x, y, u)$ transforms C into a strongly oriented cycle C' (Lemma 6.2), and cycles D and E into D'. From Lemma 6.4, D' is strongly oriented if D and E have interleaving pairs; otherwise it is nonoriented. In the first case, we use an argument similar to the case in Fig. 4.3d. If D' is nonoriented, then observe that π_y, π_u, π_v all belong to D' and $\pi_y \in I_1(C')$, $\pi_u \in I_2(C')$, and $\pi_v \in I_3(C')$, implying that D' and C' are strongly crossing. From Lemma 6.5, two valid 2-moves are possible in $G(\pi\rho)$.

Fig. 4.3f. The valid 0-move $\rho(x, y, u)$ transforms D and E into strongly oriented cycle D' (as D and E have no interleaving triples, Lemma 6.4 applies) and transforms C into nonoriented C' (Lemma 4.1). Furthermore, $\pi_{i_1}, \pi_t, \pi_{i_k} \in C'$ lie in the regions $I_2(D'), I_1(D')$, and $I_3(D')$, respectively. Therefore, C' and D' are strongly crossing. From Lemma 6.5, two valid 2-moves are possible in $G(\pi\rho)$. □

Theorem 6.8 describes how we can handle the case when $G(\pi)$ has long cycles. For short cycles, we need to formalize the intuitive idea described earlier. Define a 0-move as *good* if it increases the number of odd cycles by two.

THEOREM 6.9. *If $G(\pi)$ has only short cycles, a good 0-move followed by a valid 2-move is possible.*

Proof. We mimic the proof of Theorem 3.4. The 0-move takes two cycles of length 2 and creates an oriented cycle of length 3 and a cycle of length 1. A valid 2-move is now possible. □

Our proofs are constructive and immediately imply an $O(n^2)$ algorithm *TransSort* for sorting by transpositions. Finally, Theorems 2.4, 6.8, and 6.9 imply the following.

COROLLARY 6.10. *Algorithm* TransSort *sorts permutation π in no more than $\frac{3}{4} \cdot (n + 1 - c_{odd}(\pi))$ transpositions, thereby ensuring a performance guarantee of* 1.5.

COROLLARY 6.11. *The transposition diameter of the symmetric group S_n is at most $\frac{3}{4}n$.*

7. Open problems. Recent advances in large-scale comparative genetic mapping offer exciting prospects for understanding mammalian genome evolution. The large number of conserved segments in the maps of man and mouse suggest that multiple chromosomal rearrangements have occurred since the divergence of lineages leading to humans and mice. In their pioneering paper, Nadeau and Taylor [21] estimated that just 178 ± 39 rearrangements have occurred since this divergence. This estimate survived a ten-fold increase in the amount of the comparative man/mouse mapping information; the new estimate, based on the latest data (Copeland et al. [5]), almost did not change compared to Nadeau and Taylor [21]. However, the arguments used by Nadeau and Taylor [21] are nonconstructive and do not provide any solution to an open biological problem of reconstructing an evolutionary scenario explaining man and mouse genome rearrangements.

Chromosomal rearrangements include not only inversions and transpositions but *translocations, fussions, fissions, insertions, and deletions* as well. A combinatorial analysis of all such rearrangements to derive a scenario of mammalian evolution is far beyond the possibilities of current algorithms. However, some limited applications of algorithms for inversions and transpositions to study chromosome evolutions are already possible. In particular, extreme conservation of genes on X chromosome across mammalian species provides an opportunity to study evolutionary history of X chromosome independently of the rest of the genomes, thus reducing the computational complexity of the problem. According to Ohno's law (Ohno [22]), gene content of X chromosome is assumed to have remained the same throughout mammalian development for the last 125 million years. However, the order of genes on X chromosome has been disrupted several times. The conservative gene content of X chromosome implies that the only translocations which affected the gene order in X chromosome were translocations between two copies of X chromosome and thus might be ignored for our purposes. A recently discovered violation of the Ohno law by the *Csfgmra* gene (Disteche et al. [7]) does not affect this conclusion, since this gene is located at the very end of the human X chromosome. Davisson [6] and Lyon [19] suggested two conflicting scenarios of rearrangements in X chromosome under the assumption that X chromosome was not involved in translocations. Based on the analysis of the latest data on comparative man/mouse mapping, Bafna and Pevzner [3] found the most parsimonious scenario for evolutionary history of X chromosome and corrected the previously suggested scenarios.

Another open problem on genome rearrangements is related to viral evolution. As was mentioned in the introduction, herpes viruses present a particularly hard case for classical sequence comparison. On the other hand, they present a particularly suitable test case for the study of genome rearrangements, since complete sequences of seven diverse herpes viruses are known. Herpes virus genomes contain from 70 to about 200 genes. Detailed comparison of amino acid sequences of viral proteins resulted in an "alphabet" of about 30 conserved genes which were rearranged in different herpes viruses (Hannenhalli et al. [12]). Three types of arrangements of conserved genes exist, corresponding to the α, β, and γ divisions of herpes viruses. Derived lower bounds for the pairwise genome rearrangements of viral genomes allowed us to construct the most parsimonious scenarios for herpes virus evolution. Moreover, there are only three alternative, equally parsimonious, scenarios of genome rearrangements in herpes viruses with three different Steiner points (Hannenhalli et al. [12]). It is impossible to delineate the true scenario among these three based on the currently available data. However, ongoing efforts to map and sequence different herpes virus genomes provide a warrant that a true evolutionary scenario will be found in the future.

REFERENCES

[1] M. Aigner and D. B. West, *Sorting by insertion of leading element*, J. Combin. Theory, 45 (1987), pp. 306–309.

[2] V. Bafna and P. Pevzner, *Genome rearrangements and sorting by reversals*, SIAM J. Comput., 25 (1996), pp. 272–289.

[3] V. Bafna and P. Pevzner, *Sorting by reversals: Genome rearrangements in plant organelles and evolutionary history of X chromosome*, Molecular Biology and Evolution, 12 (1995), pp. 239–246.

[4] D. Cohen and M. Blum, *Improved bounds for sorting pancakes under a conjecture*, manuscript, 1993.

[5] N. G. Copeland, N. A. Jenkins, D. J. Gilbert, J. T. Eppig, L. J. Maltals, J. C. Miller, W. F. Dietrich, A. Weaver, S. E. Lincoln, R. G. Steen, L. D. Steen, J. H. Nadeau, and E. S. Lander, *A genetic linkage map of the mouse: Current applications and future prospects*, Science, 262 (1993), pp. 57–65.

[6] M. Davisson, *X-linked genetic homologies between mouse and man*, Genomics, 1 (1987), pp. 213–227.

[7] C. M. Disteche, C. I. Brannan, A. Larsen, D. A. Adler, D. F. Schorderet, D. Gearing, N. G. Copeland, N. A. Jenkins, and L. S. Park, *The human pesuodoautosomal GM-CSF receptor α subunit gene is autosomal in mouse*, Nature Genetics, 1 (1992), pp. 333–336.

[8] S. Even and O. Goldreich, *The minimum-length generator sequence problem is NP-hard*, J. Algorithms, 2 (1981), pp. 311–313.

[9] N. Franklin, *Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages $\lambda, \phi 21$ and $P22$*, J. Molecular Evolution, 181 (1985), pp. 75–84.

[10] W. H. Gates and C. H. Papadimitriou, *Bounds for sorting by prefix reversals*, Discrete Math., 27 (1979), pp. 47–57.

[11] A. M. Griffin and M. E. G. Boursnell, *Analysis of the nucleotide sequence of DNA from the region of the thymidine kinase gene of infectious laryngotracheitis virus: potential evolutionary relationships between the herpesvirus subfamilies*, J. General Virology, 71 (1990), pp. 841–850.

[12] S. Hannenhalli, C. Chappey, E. Koonin, and P. Pevzner, *Genome sequence comparison and scenarios for genome rearrangements: A test case*, Genomics, 30 (1995), pp. 299–311.

[13] S. Hannenhalli and P. Pevzner, *Transforming cabbage into turnip*, in Proc. 27th Annual ACM Symposium on Theory of Computing, ACM, New York, 1995, pp. 178–179.

[14] M. Heydari and I. H. Sudborough, *On sorting by prefix reversals and the diameter of pancake networks*, manuscript, 1993.

[15] M. Jerrum, *The complexity of finding minimum-length generator sequences*, Theoret. Comput. Sci., 36 (1985), pp. 265–289.

[16] S. Karlin, E. S. Mocarski, and G. A. Schachtel, *Molecular evolution of herpesviruses: Genomic and protein sequence comparisons*, J. Virology, 68 (1994), pp. 1886–1902.

[17] J. Kececioglu and D. Sankoff, *Exact and approximation algorithms for the inversion distance between two permutations*, Algorithmica, 13 (1995), pp. 180–210.

[18] J D. Kececioglu and R. Ravi, *Of mice and men: Evolutionary distances between genomes under translocations*, in Proc. Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, California, SIAM, Philadelphia, PA, 1995, pp. 604–613.

[19] M. F. Lyon, *X-Chromosome inactivation and the location and expression of X-linked genes*, Amer. J. Hum. Genet., 42 (1988), pp. 8–16.

[20] D. J. McGeoch, *Molecular evolution of large DNA viruses of eukaryotes*, Seminars in Virology, 3 (1992), pp. 399–408.

[21] J. H. Nadeau and B. A. Taylor, *Lengths of chromosomal segments conserved since divergence of man and mouse*, Proc. Nat. Acad. Sci. USA, 81 (1984), pp. 814–818.

[22] S. Ohno, *Sex Chromosomes and Sex-Linked Genes*, Springer-Verlag, Heidelberg, 1967.

[23] J. D. Palmer and L. A. Herbon, *Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence*, J. Molecular Evolution, 27 (1988), pp. 87–97.

[24] D. Sankoff, G. Leduc, N. Antoine, B. Paquin, B. F. Lang, and R. Cedergren, *Gene order comparisons for phylogenetic inference: Evolution of the mitochondrial genome*, Proc. Nat. Acad. Sci. USA, 89 (1992), pp. 6575–6579.