

PERMUTATIONS AS GENOME REARRANGEMENTS

SIMON GRÜNING

(Simon Grüning) UNIVERSITY OF ZURICH, RÄMISTRASSE 71, 8006 ZURICH
E-mail address: simon.gruening@uzh.ch.

REMARK 0.1. *RECALL: inversion*

1. BLOCK INTERCHANGES

1.1. The Cycle Graph and Its Relation to bid .

DEFINITION 1.1. A *Block Interchange* is an operation that interchanges two blocks of consecutive entries without rearranging said entries.

EXAMPLE 1.1. $|34|17|562| \rightarrow |562|17|34|$

DEFINITION 1.2. Let the *Block Interchange Distance* between two n -permutation p, q , denoted as $\text{bid}(p, q)$, be the smallest number of block interchanges required to transform p into q . Let $\text{bid}(p) := \text{bid}(p, \text{id})$ be the number of interchanges required to sort p .

REMARK 1.1. $\text{bid}(p, q)$ is a metric and is left-invariant, just as we have seen for $\text{btd}(p, q)$.

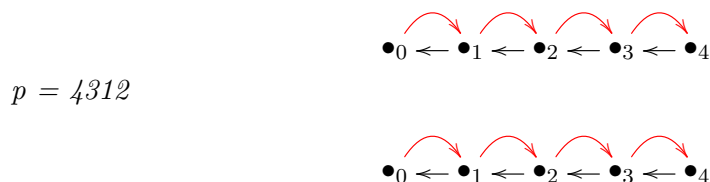
DEFINITION 1.3. Let $p = p_1 p_2 \cdots p_n$ be an n -permutation. The *Cycle Graph* $G(p)$ is a graph of coloured directed edges on the vertex set $\{0, 1, \dots, n, n+1\}$ constructed as follows:

- (1) For every i with $1 \leq i \leq n+1$: Add a black edge from p_i to p_{i-1}
- (2) For every i with $0 \leq i \leq n$: Add a red edge from i to $i+1$

with $p_0 = 0$ and $p_{n+1} = n+1$.

REMARK 1.2. The Cycle Graph has $2n+2$ edges.

EXAMPLE 1.2. $p = 123$



draw cycle graph count cycles find minimal transpositions to id and show how it works

LEMMA 1.1. The cycle graph $G(p)$ has a unique decomposition into edge-disjoint directed cycles in which the colors of the edges alternate.

Proof. It is easy to see that a path of alternating colours cannot branch since by construction there is always only one edge per color directed away from every node. We begin at the 0 node and complete a cycle (this must occur since $G(p)$ is finite). Once we have done so, we remove the edges of this cycle from $G(p)$ and iterate. Notice that in removing full cycles the structure is maintained as we only remove disjointly coloured pairs of edges from every node. The cycles are then edge-disjoint by construction. \square

DEFINITION 1.4. Let $c(G(p))$ be the number of alternating directed cycles in the decomposition of $G(p)$. To avoid confusion, let $c(\Gamma(p))$ denote the number of cycles in the traditional sense of a permutation. From now on the cycles of $G(p)$ will refer to the alternating cycles.

REMARK 1.3. The cycles are only edge-disjoint not vertex-disjoint.

REMARK 1.4. *If there is an alternating path from one node to another, both nodes are in the same cycle.*

REMARK 1.5. *The identity permutation has $n+1$ cycles. In fact, it is the only permutation with this many cycles.*

Before we can find a formula for $\text{bid}(p)$ we shall examine how single block interchanges can influence the block interchange distance of a permutation in the following two lemmas.

DEFINITION 1.5. *Canonical Block Interchange of p .*

LEMMA 1.2. 9.10 *Let $p \in S_n$ with $p \neq \text{id}$. Then there exists a block interchange which increases $c(G(p))$ by two.*

Proof. 3 cases □

LEMMA 1.3. 9.11 *A block interchange cannot increase $c(G(p))$ by more than two.*

Proof. Proof Sketch Only □

THEOREM 1.1. 9.9 *Let $p \in S_n$. Then $\text{bid}(p) = \frac{n+1-c(G(p))}{2}$*

Proof. By the previous two lemmata we have for any $p \in S_n$ that

$$\frac{n+1-c(G(p))}{2} \leq \text{bid}(p) \leq \frac{n+1-c(G(p))}{2}$$

since we require this amount of block interchanges to attain $n+1$ cycles in p , thus culminating our desired equality. □

REMARK 1.6. *We have discovered that we need at most $\lfloor \frac{n}{2} \rfloor$ block interchanges to sort an n -permutation. Furthermore $\text{bid}(p) \in \mathbb{N}$ implies that $(n+1)$ and $c(G(p))$ must either both be odd or even.*

1.2. Average Number of Block Interchanges Required to Sort p .

DEFINITION 1.6. *Define the Hultman Number as*

$$S_H(n, k) := |\{\pi \in S_n : c(G(\pi)) = k\}|$$

EXAMPLE 1.3. $S_H(3, 4) = 1$ and $S_H(3, 2) = 5$. *Draw sample graphs quickly? Maybe at least two. Continue example after 9.14 et al.*

THEOREM 1.2. 9.14 *and Corollary? Proof Sketch?*

REMARK 1.7. *The previous theorem illustrates the connection between $c(G(p))$ and $c(\Gamma(p))$. It allows us to translate ?????????? Skip all the other stuff, useless?*

THEOREM 1.3. 9.20? 9.21, 9.22? *just write theorems and quickly talk about at most.*

THEOREM 1.4. 9.23. *The average number of block interchanges needed to sort an n -permutation is*

$$\frac{1}{2} \left(n - \left\lfloor \frac{n+2}{2} \right\rfloor \right) - \sum_{i=2}^n \frac{1}{i}$$